Analysis

Marc Mertens

Email: marc.lisp@gmail.com

November 7, 2024

Table of contents

Chapter 1
Elements of set theory

1.1 Basic concepts about classes and sets

Every book about mathematical subjects must be based on one form of set theory. Because the focus of this book is on mathematical analysis instead of the foundations of mathematics, I have decided to use Von Neumann's set theory instead of the set theory of Fraenkel, Skolem and Zermelo. The benefit of Von Neumann's theory is that it is nearer to the naive set theory of Cantor. This book assumes that the basics of mathematical logic are understood, more specifically that the reader knows the meaning of the following terms:

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

and how to use them. Axiomatic set theory is based on two undefined concepts: **class** and the **membership** relation between classes (noted as $\in$). Intuitive you can think of a class as a collection and $x \in A$ to mean that x is part of the collection where A stands for. We introduce then axioms that state which are true statements about these undefined concepts. Further we introduce different definitions that helps us to simplify our notation. To start with, we define the concept of $\notin$ [not member of]

Definition 1.1. *Let A be a class then $x \notin A$ is equivalent with saying $\neg(x \in A)$.*

Next we introduce **sets** and **elements**, they are two notation for the same thing, we use **set** if we want to stress that it is a kind of collection and **element** that is also a member of a class in contrast to a class that does not have to be a member. A **set** or **element** is something that is a member of a class.

Definition 1.2. *We say that a **class** x is a **element** if $x \in A$ where A is a class. Another name for a **element** is a **set***

From here on we use the following convention: elements are noted in **lower-case** and classes are noted in **upper-case**. Next we define equality of classes.

Definition 1.3. *Let A, B classes then we say that $A = B$ if and only if*

$$\forall X \text{ we have } A \in X \Rightarrow B \in X \wedge B \in X \Rightarrow A \in X$$

Less formally, two classes A and B are equal if every class that contains A or B must contains B or A.

Once we have defined equality we can define inequality

Definition 1.4. *Let A and B classes then $A \neq B$ is equivalent with $\neg(A = B)$*

If two classes are equal, we expect them to contain the same elements, this is stated in the first set axiom.

Axiom 1.5. (Axiom of extent)

$$A = B \Leftrightarrow [x \in A \Rightarrow x \in B \wedge x \in B \Rightarrow x \in A]$$

Less formally A is equal to B if and only if ever element of A is a element of B and every element of B is a element of A, in other words A and B have the same elements.

Definition 1.6. *Let A and B classes then A is a sub-class of B noted by $A\subseteq B$ iff*

$$x\in A\Rightarrow x\in B$$

So A is a sub-class of B iff every element of A is also a element of B.

Definition 1.7. *Let A and B classes then A is a proper sub-class of B noted by $A\subseteq B$ iff*

$$x\in A\Rightarrow x\in B\wedge A\neq B$$

So A is a proper sub-class of B iff A is different from B and every element of A is also a element of B.

Theorem 1.8. *Let A,B,C be classes then the following holds:*

1. $A=A$
2. $A=B\Rightarrow B=A$
3. $A=B\wedge B=C\Rightarrow A=C$
4. $A\subseteq B\wedge B\subseteq A\Rightarrow A=B$
5. $A\subseteq B\wedge B\subseteq C\Rightarrow A\subseteq C$
6. $A=B\Rightarrow A\subseteq B$

Proof.

1. $x\in A\Rightarrow x\in A$ and $x\in A\Rightarrow x\in A$ are obviously true, hence using the Axiom of Extent [axiom: 1.5] it follows that $A=A$
2. As $A=B$ we have using the Axiom of Extent [axiom: 1.5] that $x\in A\Rightarrow x\in B\wedge x\in B\Rightarrow x\in A$ which is equivalent with $x\in B\Rightarrow x\in A\wedge x\in A\Rightarrow x\in B$. Using the Axiom of Extent [axiom: 1.5] it follows that $B=A$
3. As $A=B\wedge B=A$ we have by he Axiom of Extent [axiom: 1.5] that

$$x\in A \Rightarrow x\in B \tag{1.1}$$
$$x\in B \Rightarrow x\in A \tag{1.2}$$
$$x\in B \Rightarrow x\in C \tag{1.3}$$
$$x\in C \Rightarrow x\in B \tag{1.4}$$

   From [eq: 1.1] and [eq: 1.3] it follows that $x\in A\Rightarrow x\in C$ and from [eq: 1.4] and [eq: 1.2] it follows that $x\in C\Rightarrow x\in A$. Using the Axiom of Extent [axiom: 1.5] it follows then that $A=C$.

4. From $A\subseteq B\wedge B\subseteq A$ it follows that $x\in A\Rightarrow x\in B\wedge x\in B\Rightarrow x\in A$, so by the Axiom of Extent [axiom: 1.5] we have $A=b$
5. As $A\subseteq B\wedge B\subseteq C$ that $x\in A\Rightarrow x\in B$ and $x\in B\Rightarrow x\in C$ proving that $x\in A\Rightarrow x\in C$ or $A\subseteq C$
6. If $x\in A$ then as $A=B$ we have by the axiom of extension [axiom: 1.5] that $x\in B$, hence $A\subseteq B$.

□

One way to create a new class is to specify a predicate that a object must satisfies and then take the class of all objects that satisfies this predicate. The problem with this construction is that it can lead to paradoxes like the famous Russell paradox. Consider the predicate $R(x)=x\notin x$, this predicate is true for x if x is not a member of itself and consider the class that contains all classes that has not them self as member. Does this class contain itself yes or no? If the class contain itself then by definition $R(x)$ should be true so the class should not contain itself leading to a contradiction. If the class does not contain itself then it satisfies $R(x)$, hence it is a member of itself again leading to a contradiction. So we can not test the predicate $R(x)$ for all classes and thus can not define the class of all classes for which $R(x)$ is true. The axiom of class construction allows us to create a new class in a safe way.

Axiom 1.9. (Axiom of Construction) *Let $P(x)$ be a statement about x [using mathematical logic] then there exists a class C such that $x\in C$ iff x is a element and $P(x)$ is true.*

***Notation* 1.10.** *This class C is noted as $C=\{x|P(x)\}$, note the use of lower cases for x, which is a visual indicator that x is a element.*

Note that that C consists of **elements** for which $P(x)$ is true, it is not enough that $P(x)$ is true to belong to C. A object must belong to a class [be a element or equivalently be a set] and $P(x)$ must be true to be a member of C. Let's see how that solves Russell's paradox. Define the class $R = \{x | x \notin x\}$ [Russel's class] and check if $R \in R$ or $R \notin R$ is true:

$R \in R$. Then R is a element and $R \notin R$ giving the contradiction $R \in R \wedge R \notin R$

$R \notin R$. Then R is not a element or $R \in R$ which as $R \notin R$ gives that R is not a element

So we have that R is not a element and indeed because of this that $R \notin R$. You can ask yourself if there actually exists elements, none of the axioms up to now can be used to get elements [or equivalent sets], for this we need extra axioms.
The axiom of construction can be used as a way of creating a sub-class of a given class.

Definition 1.11. *Let A be a class and $P(x)$ a predicate then $\{x \in A | P(x)\} = \{x | x \in A \wedge P(x)\}$*

Using the axiom of construction [axiom: 1.9] we can then define the universal class $\mathcal{U}$.

Definition 1.12. (Universal class) *The universal class $\mathcal{U}$ is defined by $\mathcal{U} = \{x | x = x\}$*

The universal class contains all the elements, as is expressed in the following theorem.

Theorem 1.13. *If x is a element then $x \in \mathcal{U}$*

Proof. Let x be a element then, as $x = x$ [see theorem: 1.8] we have that $x \in \mathcal{U}$ □

We use now the axiom of construction to define the union and intersection of two classes.

Definition 1.14. *Let A, B be two classes then the union of A and B, noted as $A \bigcup B$ is defined by*

$$A \bigcup B = \{x | x \in A \vee x \in B\}$$

Definition 1.15. *Let A, B be two classes then the union of A and B, noted as $A \bigcap B$ is defined by*

$$A \bigcap B = \{x | x \in A \wedge x \in B\}$$

Next we define the empty class, the class that does not contains a element.

Definition 1.16. *The empty class $\varnothing$ is defined by*

$$\varnothing = \{x | x \neq x\}$$

Theorem 1.17. *$\varnothing$ does not contains elements, meaning if x is a element then $x \notin \varnothing$*

Proof. We proof this by contradiction, so assume that there exists a element $x \in \varnothing$ then $x \neq x$, contradicting $x = x$ [see theorem: 1.8]. □

Theorem 1.18. *If A is a class then*

1. *$\varnothing \subseteq A$*
2. *$A \subseteq \mathcal{U}$*
3. *If $A \subseteq \varnothing$ then $A = \varnothing$*

Proof.

1. We proof this by contra-position, as $\varnothing \subseteq A$ is equivalent with $x \in \varnothing \Rightarrow x \in A$. We must proof that $x \notin A \Rightarrow x \notin \varnothing$. Well if $x \notin A$ then certainly $x \notin \varnothing$ [Theorem: 1.17] so that $x \notin A \Rightarrow x \notin \varnothing$.
2. If $x \in A$ then x is a element, hence $x \in \mathcal{U}$ by [Theorem: 1.13]
3. By (1) we have $\varnothing \subseteq A$ which together with $A \subseteq \varnothing$ proves by [theorem: 1.8] that $A = \varnothing$. □

We also have that every class with no elements is equal to the empty set [there is only one empty set]

Theorem 1.19. *If A is a a class such that $x \in A$ yields a contradiction then $A = \varnothing$*

Proof. Let $x \in A$ then we have a contradiction, so $x \in A$ must be false and thus $x \in A \Rightarrow x \in \varnothing$ is vacuously true which proves that $A \subseteq \varnothing$, combining this with [theorem: 1.18,1.8] proves that $A = \varnothing$ □

Corollary 1.20. *Let A be a class such that $A \neq \varnothing$ then $\exists x$ such that $x \in A$*

Proof. We proof this by contradiction. Assume that $\forall x$ we have $x \notin A$ then $x \in A$ yields the contradiction $x \in A \wedge x \notin A$, hence by [theorem: 1.19] $A = \varnothing$ which contradicts $A \neq \varnothing$. □

Definition 1.21. *Two classes A, B are disjoint iff $A \bigcap B = \varnothing$*

We define now the complement of a class

Definition 1.22. *Let A be a class then the complement of A noted by A^c is defined by*

$$A^c = \{x | x \notin A\}$$

Something similar to the complement of a class is the difference between two classes

Definition 1.23. *Let A, B be classes then the difference between A and B noted by $A \setminus B$ is defined by*

$$A \setminus B = \{x | x \in A \wedge x \notin B\} \underset{\text{shorter notation}}{=} \{x \in A | x \in B\}$$

We can express the difference of two classes using the intersection and the complement.

Theorem 1.24. *Let A, B be classes then*

$$A \setminus B = A \bigcap B^c$$

Proof. Let $x \in A \setminus B$ then $x \in A \wedge x \notin B$ so that $x \in A \wedge x \in B^c$, further if $x \in A \bigcap B^c$ then $x \in A \wedge x \notin B$. Using then the axiom of extent [axiom: 1.5]. □

1.2 Class operations

Theorem 1.25. *Let A, B, C are classes then we have*

1. $A \subseteq A \bigcup B$
2. $B \subseteq A \bigcup B$
3. $A \bigcap B \subseteq A$
4. $A \bigcap B \subseteq B$
5. $A \setminus B \subseteq A$
6. $(A \bigcup B) \setminus C = (A \setminus C) \bigcup (B \setminus C)$
7. $(A \bigcap B) \setminus C = (A \setminus C) \bigcap B = A \bigcap (B \setminus C)$
8. *If C is a class such that $A \subseteq C$ and $B \subseteq C$ then $A \bigcup B \subseteq C$*
9. *If C is a class such that $A \subseteq C$ and D a class such that $B \subseteq D$ then $A \bigcup B \subseteq C \bigcup D$*
10. *If C is a class such that $C \subseteq A$ and $C \subseteq B$ then $C \subseteq A \bigcap B$*
11. *If C is a class such that $A \subseteq C$ and D a class such that $B \subseteq D$ then $A \bigcap B \subseteq C \bigcap D$*

Proof.

1. If $x \in A$ then $x \in A \vee x \in B$ proving that $x \in A \bigcup B$, hence $A \subseteq A \bigcup B$
2. If $x \in B$ then $x \in A \vee x \in B$ proving that $x \in A \bigcup B$, hence $B \subseteq A \bigcup B$
3. If $x \in A \bigcap B$ then $x \in A \wedge x \in B$, hence $x \in A$ so that $x \in A$, hence $A \bigcap B \subseteq A$
4. If $x \in A \bigcap B$ then $x \in A \wedge x \in B$, hence $x \in B$ so that $x \in A$, hence $A \bigcap B \subseteq B$

5. If $x\in A\backslash B$ then $x\in A\wedge x\notin B$ so that $A\backslash B\subseteq A$
6. We have

$$\begin{aligned}(A\bigcup B)\backslash C &\underset{[\text{theorem: } 1.24]}{=} (A\bigcup B)\bigcap C^c\\ &\underset{[\text{theorem: } 1.30]}{=} (A\bigcap C^c)\bigcup (B\bigcap C^c)\\ &\underset{[\text{theorem: } 1.24]}{=} (A\backslash C)\bigcup (B\backslash C)\end{aligned}$$

7.
   a. $(A\bigcap B)\backslash C=(A\bigcap B)\bigcap C^c=A\bigcap (B\bigcap C^c)=(A\bigcap C^c)\bigcap B=(A\backslash C)\bigcap B$
   b. $(A\bigcap B)\backslash C=(A\bigcap B)\bigcap C^c=A\bigcap (B\bigcap C^c)=A\bigcap (B\backslash C)$
8. If $x\in A\bigcup B$ then $x\in A \underset{A\subseteq C}{\Rightarrow} x\in C$ or $x\in B \underset{B\subseteq C}{\Rightarrow} x\in C$ proving that $x\in C$
9. Using (1) $A\subseteq C\bigcup D$ and $B\subseteq C\bigcup D$, so using (6) we have $A\bigcup B\subseteq C\bigcup D$
10. If $x\in C$ then $x\in A$ and $x\in B$ so that $x\in A\bigcap B$
11. If $x\in A\bigcap B$ then $x\in A \underset{A\subseteq C}{\Rightarrow} x\in C$ and $x\in B \underset{B\subseteq D}{\Rightarrow} x\in D$ hence $x\in C\bigcap D$.

□

Theorem 1.26. *If A,B,C 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* $A\bigcap B=A$
3. *If* $A\subseteq B$ *then* $C\backslash B\subseteq C\backslash A$
4. *If* $B\subseteq A$ *then* $A\backslash(A\backslash B)=B$

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 [theorem: 1.25] 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 [theorem: 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 $A\subseteq B$.
3. If $x\in C\backslash B$ then $x\in C$ and $x\notin B$. If $x\in A$ then as $A\subseteq B$ we would have $x\in B$ contradicting $x\notin B$ hence $x\notin A$ proving that $x\in C\backslash A$.
4. We have

$$\begin{aligned}A\backslash(A\backslash B) &\underset{[\text{theorem: } 1.24]}{=} A\bigcap (A\backslash B)^c\\ &\underset{[\text{theorem: } 1.24]}{=} A\bigcap (A\bigcap B^c)^c\\ &\underset{[\text{theorem: } 1.29]}{=} A\bigcap (A^c\bigcup (B^c)^c)\\ &\underset{[\text{theorem: } 1.28]}{=} A\bigcap (A^c\bigcup B)\\ &\underset{[\text{theorem: } 1.30]}{=} (A\bigcap A^c)\bigcup (A\bigcap B)\\ &\underset{[\text{theorem: } 1.24]}{=} (A\backslash A)\bigcup (A\bigcap B)\\ &\underset{[\text{theoremc: } 1.32]}{=} \varnothing\bigcup (A\bigcap B)\\ &\underset{[\text{theoremc: } 1.32]}{=} A\bigcap B\\ &\underset{[\text{theorem: } 1.26]}{=} B\end{aligned}$$

□

Theorem 1.27. **(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 [theorem: 1.25 we have $A\bigcap B\subseteq A$, hence using [theorem: 1.26] we have that $A\bigcup(A\bigcap B)=A$
2. By [theorem: 1.25] we have $A\subseteq A\bigcup B$, hence using [theorem: 1.26] we have that $A\bigcap(A\bigcup B)=A$ □

Theorem 1.28. *Let A be a class then $(A^c)^c=A$*

Proof. If $x\in(A^c)^c$ then x is a element and $x\notin A$ then $x\in A$ [for if $x\notin A$ we have $x\in A^c$]. If $x\in A$ then $x\notin A^c$ so that $x\in(A^c)^c$. □

Theorem 1.29. **(DeMorgan's Law)** *For all classes A, B, C we have*

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

Proof.

1. If $x\in(A\bigcup B)^c$ then $x\notin A\bigcup B$, so that $\neg(x\in A\vee x\in B)=x\notin A\wedge x\notin B$ proving that $x\in A^c\bigcap B^c$. If $x\in A^c\bigcap B^c$ then $x\notin A\wedge x\notin B=\neg(x\in A\vee x\in B)$, so that $x\notin A\bigcup B$ or $x\in(A\bigcup B)^c$. The proof follows then from the axiom of extent [axiom: 1.5]
2. If $x\in(A\bigcap B)^c$ then $x\notin A\bigcap B$, so that $\neg(x\in A\wedge x\in B)=x\notin A\vee x\notin B$ proving that $x\in A^c\bigcup B^c$. If $x\in A^c\bigcup B^c$ then $x\notin A\vee x\notin B=\neg(x\in A\wedge x\in B)$, so that $x\in(A\bigcap B)^c$. The proof follows then from axiom of extent [axiom: 1.5] □

Theorem 1.30. *Let A, B, C be classes then we have:*

commutativity.

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

idem potency.

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

associativity.

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

Distributivity.

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

Proof.

commutatitivity.

1. This follows from [axiom: 1.5] and

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

2. This follows from [axiom: 1.5] and

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

idem potency.

1. This follows from [axiom: 1.5] and

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

2. This follows from [axiom: 1.5] and

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

associativity.

1. This follows from [axiom: 1.5] and

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

2. This follows from [axiom: 1.5] and

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

Distributivity.

1. This follows from [axiom: 1.5] and

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

2. This follows from [axiom: 1.5] and

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

$\square$

Theorem 1.31. *Let A, B, C be classes then we have*

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

Proof.

1.

$$\begin{aligned}
A \setminus (B \bigcup C) &\underset{\text{theorem:}1.24}{=} A \bigcap (B \bigcup C)^c \\
&\underset{\text{theorem: } 1.29}{=} A \bigcap (B^c \bigcap C^c) \\
&\underset{\text{associativity}}{=} (A \bigcap B^c) \bigcap C^c \\
&\underset{\text{idem potency}}{=} ((A \bigcap A) \bigcap B^c) \bigcap C^c \\
&\underset{\text{associativity}}{=} (A \bigcap (A \bigcap B^c)) \bigcap C^c \\
&\underset{\text{commutativity}}{=} ((A \bigcap B^c) \bigcap A) \bigcap C^c \\
&\underset{\text{associativity}}{=} (A \bigcap B^c) \bigcap (A \bigcap C^c) \\
&\underset{\text{theorem:}1.24}{=} (A \setminus B) \bigcap (A \setminus C) \\
A \setminus (B \bigcup C) &\underset{\text{theorem:}1.24}{=} A \bigcap (B \bigcup C)^c \\
&\underset{\text{theorem: } 1.29}{=} A \bigcap (B^c \bigcap C^c) \\
&\underset{\text{associativity}}{=} (A \bigcap B^c) \bigcap C^c \\
&\underset{\text{theorem:}1.24}{=} (A \setminus B) \setminus C
\end{aligned}$$

2.

$$\begin{aligned}
A \setminus (B \bigcap C) &\underset{\text{theorem:}1.24}{=} A \bigcap (B \bigcap C)^c \\
&\underset{\text{theorem: } 1.29}{=} A \bigcap (B^c \bigcup C^c) \\
&\underset{\text{Distributivity}}{=} (A \bigcap B^c) \bigcup (A \bigcap C^c) \\
&\underset{\text{theorem:}1.24}{=} (A \setminus B) \bigcup (A \setminus C)
\end{aligned}$$

□

Theorem 1.32. *Let A be a class then we have:*

1. $\varnothing \bigcup A = A$
2. $\varnothing \bigcap \varnothing = \varnothing$
3. $A \bigcup \mathcal{U} = \mathcal{U}$
4. $A \bigcap \mathcal{U} = A$
5. $A \setminus A = \varnothing$

Proof.

1. As $\varnothing \subseteq A$ [theorem: 1.18] we have by [theorem: 1.26] that $\varnothing \bigcup A = A$
2. As $\varnothing \subseteq A$ [theorem: 1.18] we have by [theorem: 1.26] that $\varnothing \bigcap A = A$
3. As $A \subseteq \mathcal{U}$ [theorem 1.18] we have by [theorem: 1.26] that $A \bigcap \mathcal{U} = A$
4. As $A \subseteq \mathcal{U}$ [theorem 1.18] we have by [theorem: 1.26] that $A \bigcap \mathcal{U} = A$
5. Let $x \in A \setminus A$ then $x \in A \wedge x \notin A$ a contradiction, so by [theorem: 1.19] we have that $A \setminus A = \varnothing$ □

1.3 Cartesian products

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

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**.*

Lemma 1.34. *If a,b are elements such that $a=b$ then $\{a\}=\{b\}$*

Proof. If $z\in\{a\}$ then $z=a$ which by $a=b$ and [theorem: 1.8] proves that $z=b$ hence $z\in\{b\}$. Likewise if $z\in\{b\}$ then $z=b$ which by $a=b$ and [theorem: 1.8] proves that $z=a$ hence $z\in\{a\}$. Using the axiom of extent [axiom: 1.5] it follows then that $\{a\}=\{b\}$ □

If a,b are elements then we can use the axiom of construction [axiom: 1.9] to define the class $\{x|x=a\vee x=b\}$ consisting of two elements. This leads to the following definition.

Definition 1.35. *If a,b are elements then $\{a,b\}=\{x|x=a\vee x=b\}$ is called a **unordered pair**.*

The next axiom ensures we can construct new elements from given elements.. It allows us to create classes that has as members pairs of elements.

Axiom 1.36. (Axiom of Pairing) *If a,b are elements then $\{a,b\}$ is a element*

Lemma 1.37. *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 1.38. *If a is a element then $\{a\}$ is a element*

Proof. As a is a element we have by the axiom of pairing [axiom: 1.36] that $\{a,a\}$ is a element, which as $\{a\}\underset{\text{lemma: 1.37}}{=}\{a.,a\}$ proves that $\{a\}$ is a element. □

The following lemma characterize equality of unordered pairs and will be used later to characterize equality of ordered pairs.

Lemma 1.39. *If x,y,x',y' are elements then*

$$\{x,y\}=\{x',y'\} \textit{ implies } (x=x'\wedge y=y')\vee(x=y'\wedge y=x')$$

Proof. Lets's consider the following possible cases x,y:

- $\boldsymbol{x=y}$**.** Then $\{x,y\}\underset{\text{lemma: 1.37}}{=}\{x\}=\{x',y'\}$. From $x'\in\{x',y'\}=\{x\}$ it follows that $x=x'$ and from $y'\in\{x',y'\}=\{x\}$ it follows that $y=x$. As $x=x'$ it follows from [theorem: 1.8] that $y=x'$. So we have that $(x=x'\wedge y=y')$ from which it follows that

$$(x=x'\wedge y=y')\vee(x=y'\wedge y=x')$$

- $\boldsymbol{x\neq y}$**.** Then as $x\in\{x,y\}=\{x',y'\}$ we have by [axiom: 1.5] that $x\in\{x',y'\}$, so by definition we have for x either
  - $\boldsymbol{x=x'}$**.** Then as $y\in\{x,y\}=\{x',y'\}$ we have by [axiom: 1.5] that $y\in\{x',y'\}$, so by definition we have for y either:
    - $\boldsymbol{y=x'}$**.** As $x=x'\underset{\text{theorem: 1.8}}{\Rightarrow}x=y$ we contradict $x\neq y$ so this case does not apply
    - $\boldsymbol{y=y'}$**.** Then $(x=x'\wedge y=y')$ hence $(x=x'\wedge y=y')\vee(x=y'\wedge y=x')$
  - $\boldsymbol{x=y'}$**.** Then as $y\in\{x,y\}=\{x',y'\}$ we have by [axiom: 1.5] that $y\in\{x',y'\}$, so by definition we have for y either:
    - $\boldsymbol{y=x'}$**.** Then $(x=y'\wedge y=x')$ hence $(x=x'\wedge y=y')\vee(x=y'\wedge y=x')$
    - $\boldsymbol{y=y'}$**.** As $x=y'\underset{\text{theorem: 1.8}}{\Rightarrow}x=y$ we contradict $x\neq y$ so this case does not apply

So in all cases that apply we have

$$(x = x' \wedge y = y') \vee (x = y' \wedge y = x') \qquad \square$$

Lemma 1.40. *If x, y, x', y' are elements such that $(x = x' \wedge y = y') \vee (x = y' \wedge y = x')$ then $\{x, y\} = \{x, y'\}$*

Proof. Let $z \in \{x, y\}$ then either:

- $\boldsymbol{z = x}$**.** then if $x = x' \wedge y = y'$ we have using [theorem: 1.8] that $z = x'$, hence by definition $z \in \{x', y'\}$ and if $x = y' \wedge y = x'$ we have using [theorem: 1.8] that $z = y'$, hence by definition $x \in \{x', y'\}$
- $\boldsymbol{z = y}$**.** then if $x = x' \wedge y = y'$ we have using [theorem: 1.8] that $z = y'$, hence by definition $z \in \{x', y'\}$ and if $x = y' \wedge y = x'$ we have using [theorem: 1.8] that $z = x'$, hence by definition $x \in \{x', y'\}$

which proves that

$$\{x, y\} \subseteq \{x', y'\} \tag{1.5}$$

Let $z \in \{x', y'\}$ then either:

- $\boldsymbol{z = x'}$**.** then if $x = x' \wedge y = y'$ we have using [theorem: 1.8] that $z = x$, hence by definition $z \in \{x, y\}$ and if $x = y' \wedge y = x'$ we have using [theorem: 1.8] that $z = y$, hence by definition $x \in \{x, y\}$
- $\boldsymbol{z = y}$**.** then if $x = x' \wedge y = y'$ we have using [theorem: 1.8] that $z = y$, hence by definition $z \in \{x, y\}$ and if $x = y' \wedge y = x'$ we have using [theorem: 1.8] that $z = x$, hence by definition $x \in \{x, y\}$

which proves that

$$\{x', y'\} \subseteq \{x, y\} \tag{1.6}$$

Using [theorem: 1.8] on [eq: 1.5,1.6] proves that

$$\{x = y\} = \{x' = y'\} \qquad \square$$

The above lemma actually shows that the order of the elements in unordered pairs do not matter, to remedy this we construct a ordered pair.

Definition 1.41. *If a, b are elements then*

$$(a, b) = \{\{a\}, \{a, b\}\}$$

***Note* 1.42.** *As $\{a\}, \{a, b\}$ are elements we have again that $\{\{a\}, \{a, b\}\}$ is a element, hence (a, b) is also a element.*

Next we show that the order of elements is important for a tuple

Theorem 1.43. *Let x, y, x', y' are elements then*

$$(x, y) = (x', y') \Leftrightarrow x = x' \wedge y = y'$$

Proof.

- $\boldsymbol{\Rightarrow}$**.** If $(x, y) = (x', y')$ then by definition

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

  By [lemma: 1.39] we have either:

  - $\boldsymbol{\{x\} = \{x'\} \wedge \{x, y\} = \{x', y'\}}$**.** then, as $x \in \{x\}$, we have by definition $x = x'$, using [lemma: 1.39] again we have either:
    - $\boldsymbol{x = x' \wedge y = y'}$**.** Then $x = x' \wedge y = y'$
    - $\boldsymbol{x = y' \wedge y = x'}$**.** Then by [theorem: 1.8] and $x = x'$ we have $y' = x'$ so that by [theorem: 1.8] again $y = y'$. Hence we have $x = x' \wedge y = y'$
  - $\boldsymbol{\{x\} = \{x', y'\} \wedge \{x, y\} = \{x'\}}$**.** Then as $x', y' \in \{x', y'\} = \{x\}$ we have $x' = x \wedge y' = x$, as $x, y \in \{x, y\} = \{x'\}$ we have $x = x' \wedge y = x'$. Using [theorem: 1.8] on $y' = x \wedge x = x' \wedge y = x'$ we have $y = y'$. Hence $x = x' \wedge y = y'$.

  So in all cases we have

  $$x = x' \wedge y = y'$$

⇐. As $x = x'$ it follows from [lemma: 1.34] that $\{x\} = \{x'\}$, from $x = x' \wedge y = y'$ we have by [lemma: 1.40] that $\{x, y\} = \{x', y'\}$. Using [lemma: 1.40] gives then that $\{\{x\}, \{x, y\}\} = \{\{x'\}, \{x', y'\}\}$ which by definition gives

$$(x, y) = (x', y')$$

□

We are now ready to define the Cartesian product of two classes, using the axiom of construction [axiom: 1.9].

Definition 1.44. *If A, B are classes then the* ***Cartesian product*** *of A and B noted by $A \times B$ is defined as*

$$A \times B = \{z | z = (a, b) \wedge a \in A \wedge b \in B\}$$

Notation **1.45.** *Instead of writing $\{z | z = (a, b) \wedge a \in A \wedge b \in A\}$ we use in the future the shorter notation $\{(a, b) | a \in A \wedge b \in B\}$*

A special case of the Cartesian product is the Cartesian product of empty sets.

Example 1.46. $\varnothing = \varnothing \times \varnothing$

Proof. If $z \in \varnothing \times \varnothing$ then there exists a $x, y \in \varnothing$ such that $z = (x, y)$ which contradict $x, y \notin \varnothing$ [theorem: 1.17] hence by 1.19 we have $\varnothing \times \varnothing = \varnothing$. □

Theorem 1.47. *Let A be a class then $A \times \varnothing = \varnothing$ and $\varnothing \times A = \varnothing$*

Proof. If $z \in A \times \varnothing$ then $z = (x, y)$ where $y \in \varnothing$, which contradicts $y \notin \varnothing$ [theorem: 1.17], so using [theorem: 1.19] we have that

$$A \times \varnothing = \varnothing$$

Likewise if $x \in \varnothing \times A$ then $z = (x, y)$ where $x \in \varnothing$, which contradicts $x \notin \varnothing$ [theorem: 1.17], so using [theorem: 1.19] we have that

$$\varnothing \times A = \varnothing$$

□

Theorem 1.48. *If A, B, C, D are classes then we have:*

1. *If $A \subseteq B \wedge C \subseteq D$ then $A \times C \subseteq B \times D$*
2. *Let $A \neq \varnothing \wedge C \neq \varnothing$ then if $A \times C \subseteq B \times D$ it follows that $A \subseteq B \wedge C \subseteq D$*
3. *Let $A \neq \varnothing \wedge B \neq \varnothing \wedge C \neq \varnothing$ then $A \times C = B \times D \Leftrightarrow A = B \wedge C = D$*

Proof.

1. Let $z \in A \times C$ then there exists a $x \in A$ and $y \in C$ such that $z = (x, y)$. As $A \subseteq B \wedge C \subseteq D$ it follows that $x \in B \wedge y \in D$ so that $z = (x, y) \in B \times D/$ Hence

$$A \times C \subseteq B \times D$$

2. Let $x \in A$ then, as $C \neq \varnothing$, we have by [corollary: 1.20] the existence of a $y \in C$, then $(x, y) \in A \times C$ which as $A \times C \subseteq B \times D$ proves that $(x, y) \in B \times D$. By definition we have then that $x \in B$ proving

$$A \subseteq B$$

Likewise, let $y \in C$ then, as $A \neq \varnothing$ we have by [corollary: 1.20] the existence of a $x \in A$, hence $(x, y) \in A \times C$, which as $A \times C \subseteq B \times D$, proves $(x, y) \in B \times D$ and by definition $y \in D$. Hence

$$C \subseteq D$$

3.

⇒. First as $A \times C = B \times D$ we have by [theorem: 1.8] that $A \times C \subseteq B \times D$, using (2) proves then that

$$A \subseteq B \wedge C \subseteq B \quad (1.7)$$

Next as $A \times C = B \times D$ we have by [theorem: 1.8] that $B \times D \subseteq A \times C$, using (2) proves then that

$$B \subseteq A \wedge C \subseteq D \quad (1.8)$$

Combining then [eq 1.7,1.8] with [theorem: 1.8] proves

$$A = B \wedge C = D$$

⇐. As $A = B \wedge C = D$ we have by [theorem: 1.8] that $A \subseteq B$, $C \subseteq D$, $B \subseteq A$,$D \subseteq C$ which using (1) gives that $A \times C \subseteq B \times D \wedge B \times D \subseteq A \times C$. Using [theorem: 1.8 it follows then that

$$A \times C = B \times D$$ □

Theorem 1.49. *Let A,B,C and D be classes then 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

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

5. We have

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

6. We have

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

7. We have

$$\begin{aligned}
(A \times C) \setminus (B \times C) &\underset{(6)}{=} ((A \setminus C) \times B) \bigcup (A \times (C \setminus C)) \\
&\underset{[\text{theorem: } 1.32]}{=} ((A \setminus C) \times B) \bigcup (A \times \varnothing) \\
&\underset{[\text{theorem: } 1.47}{=} ((A \setminus C) \times B) \bigcup \varnothing \\
&\underset{[\text{theorem: } 1.32]}{=} (A \setminus C) \times B
\end{aligned}$$

8. We have

$$\begin{aligned}
(A \times B) \setminus (A \times C) &\underset{(6)}{=} ((A \setminus A) \times B) \bigcup (A \times (B \setminus C)) \\
&\underset{[\text{theorem: } 1.32]}{=} (\varnothing \times B) \bigcup (A \times (B \setminus C)) \\
&\underset{[\text{theorem: } 1.47}{=} \varnothing \bigcup (A \times (B \setminus C)) \\
&\underset{[\text{theorem: } 1.32]}{=} A \times (B \setminus C)
\end{aligned}$$

□

1.4 Sets

Remember that that another name for **element** is **set** [definition: 1.2]. Up to now we have used the name **element**, because we want to think of a element as a member of a class. However a element is also a class and can contain other elements. If we want to stress the collection aspect then we use the word **set** instead of **element**. The convention is to use uppercase to represent a set and lower cases for a element. Of course set and element are the same thing, we just want to stress different aspects of the same thing. Note that we have two kinds of classes classes that are a member of another class and classes that are not a member of a class. This leads to the following definition.

Definition 1.50. *A class A is a **set** [or **element**] if there exists a class B such that $A \in B$. A class that is never a member of another class is called a **proper class**.*

Up to know we had axioms that given a element/set create a new element/set, but we have not ensured the existence of a element/set. To this we must first define the concept of a successor set.

Definition 1.51. *A set S is a **successor set** iff*

1. *$\varnothing \in S$*
2. *If $X \in S$ then $X \bigcup \{X\} \in S$*

Of course nothing proves that successor set's exists, to ensure the existence of a successor set we have the axiom of infinity.

Axiom 1.52. (Axiom of Infinity) *There exists a* ***successor set***

This axiom ensures that we have at least one set. We can then use the other axioms about elements/sets to create new elements. Later we will use the Axiom of Infinity to create the Natural Numbers, form which we build all the other numbers (integers, rationals, reals, complex numbers). The Axiom of Infinity ensures also that the empty class is actually a set.

Theorem 1.53. *$\varnothing$ is a set*

Proof. The Axiom of Infinity [axiom: 1.52] ensures the existence of a successor set S. By definition we have then that $\varnothing \in S$ which proves that $\varnothing$ is a set. □

So now we have two sets to start with, the successor set and the empty set. We can use the Axiom of Pairing [axiom: 1.36] to create new sets like singletons, unordered pairs and pairs. We introduce now extra axioms to create new sets given existing sets.

Axiom 1.54. (Axiom of Subsets) *Every sub-class of a set is a set*

As a application we proof that the intersection of two sets is a set

Theorem 1.55. *Let A, B be sets then $A \bigcap B$ is a set*

Proof. By [theorem: 1.25] we have that $A \bigcap B \subseteq A$, so by the axiom of infinity [axiom: 1.52] it follows that $A \bigcap B$ is a set. □

We define now a more general concept of union and intersection

Definition 1.56. *Let $\mathcal{A}$ be a class then using the Axiom of Construction [axiom: 1.9] we define $\bigcup \mathcal{A} = \{x | \exists y \in \mathcal{A}$ such that $x \in y\}$*

Definition 1.57. *Let $\mathcal{A}$ be a class then using the Axiom of Construction [axiom: 1.9] we define $\bigcap \mathcal{A} = \{x | \forall y \in \mathcal{A}$ we have $x \in y\}$*

Example 1.58. Let A be a class then

1. $\bigcup \{A\} = A$
2. $\bigcap \{A\} = A$
3. $\bigcup \varnothing = \varnothing$

Proof.

1.
$$\begin{aligned} x \in \bigcup \{A\} &\Leftrightarrow \exists y \in \{A\} \text{ with } x \in y \\ &\underset{y \in \{A\} \Leftrightarrow y = A}{\Leftrightarrow} x \in A \end{aligned}$$
proving that
$$\bigcup \{A\} = A$$

2.
$$\begin{aligned} x \in \bigcap \{A\} &\Leftrightarrow \forall y \in \{A\} \text{ we have } x \in y \\ &\underset{y \in \{A\} \Leftrightarrow y = A}{\Leftrightarrow} x \in A \end{aligned}$$
proving that
$$\bigcap \{A\} = A$$

3. Assume that $x \in \varnothing$ then $\exists y \in \varnothing$ such that $x \in y$ which lead by the definition of $\varnothing$ [definition: 1.16] to the contradiction that $y \neq y$. □

Example 1.59. Let A and B classes then

1. $\bigcup \{A, B\} = A \bigcup B$
2. $\bigcap \{A, B\} = A \bigcap B$

Proof.

1.
$$
\begin{aligned}
x\in\bigcup\{A,B\} &\Leftrightarrow \exists y\in\{A,B\} \text{ with } x\in y\\
&\underset{y\in\{A,B\}\Leftrightarrow y=A\vee y=B}{\Leftrightarrow} x\in A\vee x\in B\\
&\Leftrightarrow x\in A\bigcup B
\end{aligned}
$$

   proving that
$$\bigcup\{A,B\}=A\bigcup B$$

2.
$$
\begin{aligned}
x\in\bigcap\{A,B\} &\Leftrightarrow \forall y\in\{A,B\} \text{ with } x\in y\\
&\underset{y\in\{A,B\}\Leftrightarrow y=A\vee y=B}{\Leftrightarrow} x\in A\wedge x\in B\\
&\Leftrightarrow x\in A\bigcap B
\end{aligned}
$$

   proving that
$$\bigcap\{A,B\}=A\bigcap B$$

□

Theorem 1.60. *If $\mathcal{A}$ is a class*

1. *If $A\in\mathcal{A}$ then $\bigcap\mathcal{A}\subseteq A$*
2. *If $A\in\mathcal{A}$ then $A\subseteq\bigcup\mathcal{A}$*
3. *If $\forall A\in\mathcal{A}$ we have $C\subseteq A$ then $C\subseteq\bigcap\mathcal{A}$*
4. *If $\forall A\in\mathcal{A}$ we have $A\subseteq C$ then $\bigcup\mathcal{A}\subseteq C$*
5. *If $\mathcal{A}\neq\varnothing$ then $\bigcap\mathcal{A}$ is a set*

Proof.

1. Let $A\in\mathcal{A}$ then if $x\in\bigcap\mathcal{A}$ we have by definition of $\bigcap\mathcal{A}$ that $x\in A$. Hence $\bigcap\mathcal{A}\subseteq A$
2. If $x\in A$ then $\exists y\in\mathcal{A}$ such that $x\in y$ [take $y=A$] so that $x\in\bigcup\mathcal{A}$
3. If $x\in C$ then $\forall A\in\mathcal{A}$ we have as $C\in A$ that $x\in A$ so that $x\in\bigcap\mathcal{A}$
4. If $x\in\bigcup\mathcal{A}$ then $\exists A\in\mathcal{A}$ such that $x\in A$ which as $A\subseteq C$ proves that $x\in A$
5. As $\mathcal{A}\neq\varnothing$ there exists a $A\in\mathcal{A}$, which by definition means that A is a set. Using (1) we have $\bigcap\mathcal{A}\subseteq A$, applying then the Axiom of Subsets [axiom: 1.54] it follows that $\bigcap\mathcal{A}$ is a set. □

The above is not applicable for unions, however we state the Axiom of Unions that will ensure that $\bigcup\mathcal{A}$ is a set if $\mathcal{A}$ is a set

Axiom 1.61. (Axiom of Unions) *If $\mathcal{A}$ is a set then $\bigcup\mathcal{A}$ is a set*

A consequence of the above axiom is that the union of two sets is a set

Theorem 1.62. *Let A,B be tow sets then $A\bigcup B$ is a set*

Proof. Using the Axiom of Pairing [axiom: 1.36] we have that $\{A,B\}$ is a set. Further

$$
\begin{aligned}
x\in A\bigcup B &\Leftrightarrow x\in A\vee x\in B\\
&\Leftrightarrow \exists C\in\{A,B\} \text{ with } x\in C\\
&\Leftrightarrow \bigcup\{A,B\}
\end{aligned}
$$

proving by the Axiom of Union [axiom: 1.61] we have that $A\bigcup B$ is a set. □

Definition 1.63. *Let A be a set then we use the Axiom of Construction to define $\mathcal{P}(A)$ by*

$$\mathcal{P}(A)=\{B|B\subseteq A\}$$

We introduce now the Axiom of Power Sets to ensure that $\mathcal{P}(A)$ is a set, called the **power set** of A.

Axiom 1.64. (Axiom of Power Sets) *If A is a set then $\mathcal{P}(A)$ is a set*

Theorem 1.65. *If A is a set and $P(X)$ a predicate then $\{X|X\subseteq A\wedge P(X)\}$ is a set.*

Proof. If $B\in\{X|X\subseteq A\wedge P(X)\}$ then $B\subseteq A$ so that $B\in\mathcal{P}(A)$, proving that

$$\{X|X\subseteq A\wedge P(X)\}\subseteq\mathcal{P}(A)$$

Using the Axiom of Power Sets [axiom: 1.64] $\mathcal{P}(A)$ is a set, so we can use the Axiom of Subsets to prove that $\{X|X\subseteq A\wedge P(X)\}$ is a set. □

Lemma 1.66. *If A,B are classes then $A\times B\subseteq\mathcal{P}(\mathcal{P}(A\bigcup B))$*

Proof. Let $z\in A\times B$ then there exists a $x\in A$ and a $y\in B$ so that $z=(x,y)$. Now if $e\in\{x\}$ then $e=x$ proving that $e\in A$, hence we have, by definition of the union, that $\{x\}\subseteq A\bigcup B$. By definition of the $\mathcal{P}(A\bigcup B)$ set it follows then that

$$\{x\}\in\mathcal{P}\left(A\bigcup B\right)$$

Likewise if $e\in\{x,y\}$ then either $e=x\Rightarrow e\in A$ or $e=y\Rightarrow e\in B$, hence ,by definition of the union, we have $\{x,y\}\subseteq A\bigcup B$. Using the definition $\mathcal{P}(A\bigcup B)$ we have then

$$\{x,y\}\in\mathcal{P}\left(A\bigcup B\right)$$

Now if $e\in\{\{x\},\{x,y\}\}$ then either $e=\{x\}\in\mathcal{P}(A\bigcup B)$ or $e=\{z,y\}\in\mathcal{P}(A\bigcup B)$ which proves that $\{\{x\},\{x,y\}\}\subseteq\mathcal{P}(A\bigcup B)$ or

$$z\in\{\{x\},\{x,y\}\}\in\mathcal{P}\left(\mathcal{P}\left(A\bigcup B\right)\right)$$

giving finally

$$A\times B\subseteq\mathcal{P}\left(\mathcal{P}\left(A\bigcup B\right)\right)$$ □

Theorem 1.67. *If A and B are sets then $A\times B$ is a set*

Proof. As A,B are sets we have by [theorem: 1.62] that $A\bigcup B$ is a set, using the Axiom of Power sets [axiom: 1.64] it follows that $\mathcal{P}(A\bigcup B)$ is a set, using the Axiom of Power sets [axiom: 1.64] again proves that $\mathcal{P}(\mathcal{P}(A\bigcup B))$ is a set. Finally by [lemma: 1.66] we have that $A\times B\subseteq\mathcal{P}(\mathcal{P}(A\bigcup B))$, which using the Axiom of Subsets [axiom: 1.54] proves that

$$A\times B \text{ is a set}$$ □

We show now that given a set we can find a element that is not in the set.

Theorem 1.68. *Let S be a set then there exists a element x such that $x\notin S$*

Proof. Use the axiom of construction [axiom: 1.9] to define $T=\{x\in S|x\notin x\}$ which as $T\subseteq S$ is a set by [axiom: 1.54] is a set. Assume now that $T\in S$ then we have either:

***T* ∈ *T*.** Then we have $T\notin T$ contradicting $T\in T$.

***T* ∉ *T*.** Then we have by definition of T that $T\in T$ contradicting $T\notin T$

As in all cases we have a contradiction the assumption is wrong and we must have $T\notin S$. □

Chapter 2

Partial Functions and Functions

2.1 Pairs and Triples

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 2.1. *If A, B are classes then $\langle A, B\rangle$ is defined by $\langle A, B\rangle = (A \times \{\varnothing\}) \bigcup (B \times \{\{\varnothing\}\})$*

We show now that from $\langle A, B\rangle = \langle A',' B\rangle$ it follows that $A = A' \wedge B = B'$, first we need some lemma's

Lemma 2.2. *We have $\varnothing \neq \{\varnothing\}$*

Proof. Assume that $\{\varnothing\} = \varnothing$ then, as $\varnothing \in \{\varnothing\}$ it follows that $\varnothing$ which is a contradiction, hence

$$\varnothing \neq \{\varnothing\}$$

□

Lemma 2.3. *If A, B, C, D are classes then $\langle A, B\rangle = \langle C, D\rangle \Leftrightarrow A = C \wedge B = D$*

Proof.

⇒. Assume that $\langle A, B\rangle = \langle C, D\rangle$ then by definition

$$(A \times \{\varnothing\}) \bigcup (B \times \{\{\varnothing\}\}) = (C \times \{\varnothing\}) \bigcup (D \times \{\{\varnothing\}\}) \tag{2.1}$$

Let now $x \in A$ then $(x, \varnothing) \in (A \times \{\varnothing\})$ so that by the axiom of extent [axiom: 1.5] and [eq: 2.1]

$$(x, \varnothing) \in (C \times \{\varnothing\}) \bigcup (D \times \{\{\varnothing\}\})$$

which by the definition of the union gives

$$(x, \varnothing) \in C \times \{\varnothing\} \vee (x, \varnothing) \in D \times \{\{\varnothing\}\} \tag{2.2}$$

Now if $(x, \varnothing) \in D \times \{\{\varnothing\}\}$ then $\varnothing \in \{\{\varnothing\}\}$ or $\varnothing = \{\varnothing\}$ which is impossible by [lemma: 2.2] so that by [eq: 2.2] we have $(x, \varnothing) \in C \times \{\varnothing\}$, hence $x \in C$. This proves that

$$A \subseteq C \tag{2.3}$$

Likewise, let $x \in C$ then $(x, \varnothing) \in (C \times \{\varnothing\})$ so that by the axiom of extent [axiom: 1.5] and [eq: 2.1

$$(x, \varnothing) \in (A \times \{\varnothing\}) \bigcup (B \times \{\{\varnothing\}\})$$

which by the definition of the union gives

$$(x, \varnothing) \in A \times \{\varnothing\} \vee (x, \varnothing) \in B \times \{\{\varnothing\}\} \tag{2.4}$$

Now if $(x, \varnothing) \in B \times \{\{\varnothing\}\}$ then $\varnothing \in \{\{\varnothing\}\}$ or $\varnothing = \{\varnothing\}$ which is impossible by [lemma: 2.2] so that by [eq: 2.4] we have $(x, \varnothing) \in C \times \{\varnothing\}$, hence $x \in A$. This proves that

$$C \subseteq A \tag{2.5}$$

Combining [eq: 2.3,2.5] with [theorem: 1.8] proves

$$A = C$$

Further if $x \in B$ then $(x, \{\varnothing\}) \in B \times \{\{\varnothing\}\}$ so that by the axiom of extent [axiom: 1.5] and [eq: 2.1]

$$(x, \{\varnothing\}) \in (C \times \{\varnothing\}) \bigcup (D \times \{\{\varnothing\}\})$$

or using the definition of the union that

$$(x, \{\varnothing\}) \in C \times \{\varnothing\} \vee (x, \{\varnothing\}) \in D \times \{\{\varnothing\}\} \tag{2.6}$$

If $(x, \{\varnothing\}) \in C \times \{\varnothing\}$ then $\{\varnothing\} \in \{\varnothing\}$ or $\{\varnothing\} = \varnothing$ which is impossible by [lemma: 2.2], so by [eq: 2.6] we have that $(x, \{\varnothing\}) \in D \times \{\{\varnothing\}\}$, hence $x \in D$. This proves that

$$B \subseteq D \tag{2.7}$$

Likewise, if $x \in D$ then $(x, \{\varnothing\}) \in D \times \{\{\varnothing\}\}$ so that by the axiom of extent [axiom: 1.5] and [eq: 2.1]

$$(x, \{\varnothing\}) \in (A \times \{\varnothing\}) \bigcup (B \times \{\{\varnothing\}\})$$

or using the definition of the union that

$$(x, \{\varnothing\}) \in A \times \{\varnothing\} \vee (x, \{\varnothing\}) \in B \times \{\{\varnothing\}\} \tag{2.8}$$

If $(x, \{\varnothing\}) \in A \times \{\varnothing\}$ then $\{\varnothing\} \in \{\varnothing\}$ or $\{\varnothing\} = \varnothing$ which is impossible by [lemma: 2.2], so by [eq: 2.8] we have that $(x, \{\varnothing\}) \in B \times \{\{\varnothing\}\}$, hence $x \in B$. This proves that

$$D \subseteq B \tag{2.9}$$

Combining [eq: 2.7,2.9] with [theorem: 1.8] proves

$$B = D$$

$\Leftarrow$. Assume that $A = C \wedge B = D$ then

$$\begin{aligned}
x \in \langle A, B \rangle &\Leftrightarrow x \in (A \times \{\varnothing\}) \bigcup (B \times \{\{\varnothing\}\}) \\
&\Leftrightarrow x \in A \times \{\varnothing\} \vee x \in B \times \{\{\varnothing\}\} \\
&\Leftrightarrow (x = (a, \varnothing) \wedge a \in A) \vee (x = (b, \{\varnothing\}) \wedge b \in B) \\
&\underset{\text{[axiom: 1.5]}}{\Leftrightarrow} (x = (a, \varnothing) \wedge a \in C) \vee (x = (b, \{\varnothing\}) \wedge b \in D) \\
&\Leftrightarrow x \in (C \times \{\varnothing\}) \bigcup (D \times \{\{\varnothing\}\}) \\
&\Leftrightarrow e \in \langle C, D \rangle
\end{aligned}$$

so that by the Axiom of Extent [axiom: 1.5]

$$\langle A, B \rangle = \langle C, D \rangle$$

□

We can now easily extend $\langle A, B \rangle$ to a triple $\langle A, B, C \rangle$.

Definition 2.4. *Let A, B, C be classes then $\langle A, B, C \rangle$ is defined by*

$$\langle A, B, C \rangle = \langle \langle A, B \rangle, C \rangle$$

Lemma 2.5. *Let A, B, C, D, E, F be classes then*

$$\langle A, B, C \rangle = \langle D, E, F \rangle \Leftrightarrow A = D \wedge B = E \wedge C = F$$

Proof.

$\Rightarrow$. Assume that $\langle A, B, C \rangle = \langle D, E, F \rangle$ then by definition $\langle \langle A, B \rangle, C \rangle = \langle \langle D, E \rangle, F \rangle$, by [lemma: 2.3] then $C = F \wedge \langle A, B \rangle = \langle D, E \rangle$, using [lemma: 2.3] again proves then $A = D \wedge B = E$.

$\Leftarrow$. Assume that $A = D \wedge B = E \wedge C = F$ then by [lemma: 2.3] $\langle A, B \rangle = \langle D, E \rangle$, using [lemma: 2.3] again we have $\langle \langle A, B \rangle, C \rangle = \langle \langle D, E \rangle, F \rangle$ which by definition proves that

$$\langle A, B, C \rangle = \langle D, E, F \rangle$$

□

2.2 Partial functions and Functions

The concept of a function as a mapping of one value to a unique value is used throughout mathematics, especially in analysis, which is essential a theory of functions. Note that a function maps a value x to a **unique** value y which in the context of a set theory defines a pair (x, y). This leads to the following definition of a graph.

2.2.1 Partial function

Definition 2.6. (Graph) *A graph is a sub class of $\mathcal{U}\times\mathcal{U}$, or in other words a graph is a collection of pairs.*

Definition 2.7. *A triple $\langle A,B,f\rangle$ where A,B are classes and f a graph is a **partial function between A and B** if*

1. $f\subseteq A\times B$
2. *If* $(x,y)\in f\wedge(x,y')\in f$ *then* $y=y'$

*We call A the **source** of the partial function, B the **destination** of the partial function and f the **graph** of the partial function.*

Remark 2.8. Instead of writing $\langle A,B,f\rangle$ for a partial function between A and B we use the notation $f:A\rightarrow B$ or $A\xrightarrow{f}$B. Further the condition (2) ensures that only one value can be associated with x. So it is useful to use a special notation for this unique value, especially if we have a expression to calculate this unique value.

Definition 2.9. *Let $f:A\rightarrow B$ be a partial function then $(x,y)\in f$ is equivalent with $y=f(x)$*

From now on we will use the Axiom of Construction [axiom: 1.9] to define different classes related to partial functions without explicitly mentioning this. It is assumed that the reader understand when to use this axiom.

Definition 2.10. *Let $f:A\rightarrow B$ be a partial function then its domain noted as $\mathrm{dom}(f)$ and range noted as $\mathrm{range}(f)$ is defined by*

$$\mathrm{dom}(f)=\{x|\exists y \textit{ such that } (x,y)\in f\}$$

$$\mathrm{range}(f)=\{y|\exists x \textit{ such that } (x,y)\in f\}$$

Theorem 2.11. *If $f:A\rightarrow B$ is a partial function then $\mathrm{dom}(f)\subseteq A$ and $\mathrm{range}(f)\subseteq B$*

Proof. If $x\in\mathrm{dom}(f)$ then $\exists y$ such that $(x,y)\in f \underset{f\subseteq A\times B}{\Rightarrow}(x,y)\in A\times B$ proving that $x\in A$, hence

$$\mathrm{dom}(f)\subseteq A$$

Further if $y\in\mathrm{range}(f)$ then $\exists x$ such that $(x,y)\in f \underset{f\subseteq A\times B}{\Rightarrow}(x,y)\in A\times B$ proving that $y\in B$, hence

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

□

Corollary 2.12. *If A,B are sets and $f:A\rightarrow B$ a partial function then $\mathrm{dom}(f)$ and $\mathrm{range}(f)$ are sets*

Proof. Using [theorem: 2.11] we have that $\mathrm{dom}(f)\subseteq A$ and $\mathrm{range}(f)\subseteq B$, so applying the Axiom of Subsets [axiom: 1.54] proves that $\mathrm{dom}(f)$ and $\mathrm{range}(f)$ are sets. □

Definition 2.13. *Let $f:A\rightarrow B$ be a partial function and C a class such that $C\subseteq A$ then **the image of C by f** noted as $f(C)$ is defined by*

$$f(C)=\{y|\exists x\in C \textit{ such that } (x,y)\in f\}$$

Remark 2.14. Note that we use a conflicting notation here. On one hand $y=f(x)$ can be interpreted as $(x,y)\in f$, on the other hand it can also means that y is the image of x by f. We adopt the following convention. If lower cases are used as in $y=f(x)$ we interpret this as $(x,y)\in f$ and if we use uppercase like in $f(C)$ we are talking about images. In case of doubt $(f)(C)$ always refers to the image.

Definition 2.15. *Let $f:A\rightarrow B$ be a partial function and C a class then **the preimage of C by f** noted as $f^{-1}(C)$ is defined by*

$$f^{-1}(C)=\{x|\exists y\in C \textit{ such that } (x,y)\in f\}$$

Note 2.16. In contrast with most text books we do not require that $C\subseteq B$, this will give us more flexibility if we compose partial functions.

Theorem 2.17. *Let $f:A\rightarrow B$ be a partial function, $C\subseteq A$ and D a class then we have:*

1. $f(C)\subseteq\mathrm{range}(f)$
2. $f^{-1}(D)\subseteq\mathrm{dom}(f)$

3. $f(A) = \text{range}(f)$
4. $f^{-1}(B) = \text{dom}(f)$
5. *If* $E \subseteq C$ *then* $f(E) \subseteq f(C)$
6. *If* $E \subseteq D$ *then* $f^{-1}(E) \subseteq f^{-1}(D)$

and if in addition A, B *are sets then* $f(C)$ *and* $f^{-1}(D)$ *are sets*

Proof.

1. If $y \in f(C)$ then there exists a $x \in C$ such that $(x, y) \in f$, so $y \in \text{range}(f)$. Hence
$$f(C) \subseteq \text{range}(f)$$
2. If $x \in f^{-1}(D)$ then there exists a $y \in D$ such that $(x, y) \in f$, which proves that $x \in \text{dom}(f)$, hence
$$f^{-1}(D) \subseteq \text{dom}(f)$$
3. If $y \in \text{range}(f)$ then $\exists x$ such that $(x, y) \in f$, which as $f \subseteq A \times B$ proves that $x \in A$, hence $y \in f(A)$, or $\text{range}(f) \subseteq f(A)$. From (1) we have $f(A) \subseteq \text{range}(f)$, so using [theorem: 1.8]
$$f(A) = \text{range}(f)$$
4. If $x \in \text{dom}(f)$ then $\exists y$ such that $(x, y) \in f$, which as $f \subseteq A \times B$ proves that $y \in B$, giving $x \in f^{-1}(B)$, hence $\text{dom}(f) \subseteq f^{-1}(B)$. From (2) we have $f^{-1}(B) \subseteq \text{dom}(f)$, so using [theorem: 1.8]
$$f^{-1}(B) = \text{dom}(f)$$
5. If $y \in f(E)$ then $\exists x \in E$ such that $(x, y) \in f$, as $E \subseteq C$ we have $x \in C$ and still $(x, y) \in f$ so that $y \in f(C)$ proving
$$f(E) \subseteq f(C)$$
6. If $x \in f^{-1}(E)$ there $\exists y \in E$ such that $(x, y) \in f$, as $E \subseteq D$ we have $y \in D$ and still $(x, y) \in f$ so that $x \in f^{-1}(D)$ proving
$$f^{-1}(E) \subseteq f^{-1}(D)$$

Finally if A, B are sets then using [theorem: 2.12] $\text{range}(f)$ and $\text{dom}(f)$ are sets, applying then the Axiom of Subsets [axiom: 1.54] proves that $f(C)$ and $f^{-1}(D)$ are sets. □

Next we define the composition of two partial functions.

Definition 2.18. (Composition of graphs) *Let* f, g *be two graphs then* $f \circ g$ *is defined by*

$$g \circ f = \{z | z = (x, y) \text{ such that } \exists u \text{ with } (x, u) \in f \wedge (u, y) \in g\}$$

Theorem 2.19. *Let* $f: A \to B$ *and* $g: C \to D$ *be partial functions then*

$$g \circ f: A \to D$$

is a partial function. We call $g \circ f: A \to D$ *the* ***composiiton*** *of* $f: A \to B$ *and* $g: C - D$

Proof. If $(x, y) \in g \circ f$ then there exist a u such that $(x, u) \in f$ and $(u, y) \in g$, as f, g are partial functions we have that $f \subseteq A \times B$ and $g \subseteq C \times D$. So $(x, u) \in A \times B$ and $(u, y) \in C \times D$. So $x \in A$ and $y \in D$ proving that $(x, y) \in A \times D$. Hence

$$g \circ f \subseteq A \times D$$

Further if $(x, y) \in g \circ f \wedge (x, y') \in g \circ f$ then there exists u, v such that $(x, u) \in f \wedge (u, y) \in g \wedge (x, v) \in f \wedge (v, y') \in g$. From $(x, u) \in f \wedge (x, v) \in f$ it follows [as f is a partial function] that $u = v$. So $(u, y) \underset{\text{u=v and [theorem: 1.43]}}{=} (u, y') \in g$. Hence as g is a partial function it follows that $y = y'$. To summarize

$$\text{If } (x, y) \in g \circ f \wedge (x, y') \in g \circ f \text{ then } y = y'$$

So all the requirements for $g \circ f: A \to D$ to be a partial function are satisfied. □

Note 2.20. In contrast with most textbooks we do not require that $B=C$ in this theorem, there is no need for this because for partial functions $\mathrm{dom}(f\circ g)$ can be different from A. Later we will compose functions and then we will need a extra condition for C.

Theorem 2.21. (Associativity of Composition) *Let $f:A\rightarrow B$, $g:C\rightarrow D$ and $h:E\rightarrow F$ be partial functions then $h\circ(g\circ f)=(h\circ g)\circ f$*

Proof. If $(x.z)\in h\circ(g\circ f)$ then $\exists u$ such that $(x,u)\in g\circ f$ and $(u,z)\in h$. As $(x,u)\in g\circ f$ there exists a v such that $(x,v)\in f$ and $(v,u)\in g$. As $(v,u)\in g\wedge(u,z)\in h$ we have that $(v,z)\in h\circ g$, as $(x,v)\in f$ it follows $(x,z)\in(h\circ g)\circ f$.
If $(x,z)\in(h\circ g)\circ f$ there $\exists u$ such that $(x,u)\in f$ and $(u,z)\in h\circ g$. As $(u,z)\in h\circ g$ there $\exists v$ such that $(u,v)\in g$ and $(v,z)\in h$. From $(x,u)\in f$ and $(u,v)\in g$ we have that $(x,v)\in g\circ f$. As $(v,z)\in h$ we have that $(x,z)\in h\circ(h\circ f)$.
Using the Axiom of Extent [axiom: 1.5] it follows that

$$h\circ(g\circ f)=(h\circ g)\circ f$$

□

Let's look now at the domain and range of of the composition of two partial functions.

Theorem 2.22. *Let $f:A\rightarrow B$ and $g:C\rightarrow D$ be partial functions then for $g\circ f:A\rightarrow D$ we have*

1. $\mathrm{dom}\,(g\circ f)=\mathrm{dom}(f)\bigcap f^{-1}(\mathrm{dom}(g))$
2. $\mathrm{range}(g\circ f)=g(\mathrm{range}(f)\bigcap \mathrm{dom}(g))$
3. $\mathrm{range}(g\circ f)\subseteq\mathrm{range}(g)$

Proof.

1. If $x\in\mathrm{dom}(g\circ f)$ then there exist a z such that $(x,z)\in g\circ f$. So there exist a y such that $(x,y)\in f$ and $(y,z)\in g$, hence $x\in\mathrm{dom}(f)$ and $y\in\mathrm{dom}(g)\underset{(x,y)\in f}{\Rightarrow}x\in f^{-1}(\mathrm{dom}(g))$. So $x\in\mathrm{dom}(f)\bigcap f^{-1}(\mathrm{dom}(g))$. Hence

$$\mathrm{dom}(g\circ f)\subseteq\mathrm{dom}(f)\bigcap f^{-1}(\mathrm{dom}(g)) \tag{2.10}$$

   If $x\in\mathrm{dom}(f)\bigcap f^{-1}(\mathrm{dom}(g))$ then $x\in\mathrm{dom}(f)$ so that $\exists y$ such that $(x,y)\in f$ and $x\in f^{-1}(\mathrm{dom}(g))$ so that $\exists y'\in\mathrm{dom}(g)$ such that $(x,y')\in f$. As f is a partial function it follows that $y=y'$. So $y\in\mathrm{dom}(g)$, from which it follows that $\exists z$ such that $(y,z)\in g$. As we have $(x,y)\in f$ and $(y,z)\in g$ it follows that $(x.z)\in g\circ f$ or $x\in\mathrm{dom}(g\circ f)$. This proves that $\mathrm{dom}(f)\bigcap f^{-1}(\mathrm{dom}(g))\subseteq\mathrm{dom}(g\circ f)$, combining this with [eq: 2.10] allows us to use [theorem: 1.8] to get

$$\mathrm{dom}(g\circ f)=\mathrm{dom}(f)\bigcap f^{-1}(\mathrm{dom}(g))$$

2. If $z\in\mathrm{range}(g\circ f)$ then there exists a $x\in A$ such that $(x,z)\in g\circ f$, so there exist a y such that $(x,y)\in f\wedge(y,z)\in g$. Then $y\in\mathrm{range}(f)$ and $y\in\mathrm{dom}(g)$ or $y\in\mathrm{range}(f)\bigcap\mathrm{dom}(g)$, which as $(y,z)\in g$ proves that $z\in g(\mathrm{range}(f)\bigcap\mathrm{dom}(g))$. Hence

$$\mathrm{range}(g\circ f)\subseteq g\big(\mathrm{range}(f)\bigcap\mathrm{dom}(g)\big) \tag{2.11}$$

   If $z\in g(\mathrm{range}(f)\bigcap\mathrm{dom}(g))$ then $\exists y\in\mathrm{range}(f)\bigcap\mathrm{dom}(g)$ such that $(y,z)\in g$. From $y\in\mathrm{range}(f)$ it follows that there exist a x such that $(x,y)\in f$. So $(x,z)\in g\circ f$ proving that $x\in\mathrm{range}(g\circ f)$, hence $g(\mathrm{range}(f)\bigcap\mathrm{dom}(g))\subseteq\mathrm{range}(g\circ f)$. Combining this with [eq: 2.11] allows us to use [theorem: 1.8] to get

$$\mathrm{range}(g\circ f)=g\big(\mathrm{range}(f)\bigcap\mathrm{dom}(g)\big)$$

3. 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 $(x,y)\in f\wedge(y,z)\in g$. Hence $z\in\mathrm{range}(g)$.

□

Theorem 2.23. *If $f:A\rightarrow B$ and $g:C\rightarrow D$ are partial functions then we have*

1. *If $E\subseteq A$ then $(g\circ f)(E)=g(f(E))$*
2. *If $E\subseteq D$ then $(g\circ f)^{-1}(E)=f^{-1}(g^{-1}(E))$*

Proof.

1. If $z\in(g\circ f)(E)$ then there exists a $x\in E$ such that $(x,z)\in g\circ f$. So by definition there exist a y such that $(x,y)\in f\wedge(y,z)\in g$. From $(x,y)\in f$ it follows that $y\in f(E)$ and as $(y,z)\in g$ it follows that $z\in g(f(E))$. Hence

$$(g\circ f)(E)\subseteq g(f(E)) \tag{2.12}$$

On the other hand if $z \in g(f(E))$ there exist a $y \in f(E)$ such that $(y, z) \in g$. As $y \in f(E)$ there exists a $x \in E$ such that $(x, y) \in f$. From $(x, y) \in f \wedge (y, z) \in g$ it follows that $(x, z) \in g \circ f$ so that [as $x \in E$] $z \in (g \circ f)(E)$. Proving $g(f(E)) \subseteq (g \circ f)(E)$, combining this with [eq 2.12] and [theorem: 1.8] gives

$$(g \circ f)(E) = g(f(E))$$

2. If $x \in (g \circ f)^{-1}(E)$ then there exist a $z \in E$ such that $(x, z) \in g \circ f$, hence $\exists y$ such that $(x, y) \in f \wedge (y, z) \in g$. So by definition $y \in g^{-1}(E)$ and as $(x, y) \in f$ it follows that $x \in f^{-1}(g^{-1}(E))$. Hence

$$(g \circ f)^{-1}(E) \subseteq f^{-1}(g^{-1}(E)) \tag{2.13}$$

If $x \in f^{-1}(g^{-1}(E))$ then there exist a $y \in g^{-1}(E)$ such that $(x, y) \in f$, as $y \in g^{-1}(E)$ then there exist a $z \in E$ such that $(y, z) \in g$. From $z \in E \wedge (x, y) \in f \wedge (y, z) \in g$ it follows that $x \in (g \circ f)^{-1}(E)$ proving that $f^{-1}(g^{-1}(E)) \subseteq (g \circ f)^{-1}(E)$. Combining this with [eq: 2.13] and [theorem: 1.8] gives

$$(g \circ f)^{-1}(E) = f^{-1}(g^{-1}(E))$$

□

2.2.2 Functions

Definition 2.24. *A partial function $f: A \rightarrow B$ is a* ***function*** *iff* $\operatorname{dom}(f) = A$

So every function is also a partial function, hence statements about partial functions applies also for functions. One special benefiit of functions is the following.

Theorem 2.25. *If $f: A \rightarrow B$ is a function then for $C \subseteq A$ we have*

1. $C \subseteq f^{-1}(f(C))$
2. $f(f^{-1}(C)) \subseteq C$

Proof.

1. If $x \in C \subseteq A$ then as $A = \operatorname{dom}(f)$ there exist a y such that $(x, y) \in f$ so that $y \in f(C)$, which as $(x, y) \in f$ proves that $x \in f^{-1}(f(C))$. Hence we have $C \subseteq f^{-1}(f(C))$
2. If $y \in f(f^{-1}(C))$ then there exist a $x \in f^{-1}(C)$ such that $(x, y) \in f$. As $x \in f^{-1}(C)$ there exist a $c \in C$ such that $(x, c) \in f$. Hence as $(x, y), (x, c) \in f$ we must have that $y = c \in C$ proving that $f(f^{-1}(C)) \subseteq C$. □

Proposition 2.26. *A partial function $f: A \rightarrow B$ is a function iff $A \subseteq \operatorname{dom}(f)$*

Proof. As $A \subseteq \operatorname{dom}(f)$ and $\operatorname{dom}(f) \subseteq A$ [theorem: 2.11] we have by [theorem: 1.8] that

$$\operatorname{dom}(f) = A$$

□

Example 2.27. Let A, B be elements and define $f = \{(0, A), (1, B)\}$ then $f: \{0, 1\} \rightarrow \{A, B\}$ is a function

Proof. If $(x, y) \in f$ then

$$(x, y) = (0, A) \Rightarrow x = 1 \in \{0, 1\} \wedge y = A \in \{A, B\} \text{ so that } (x, y) \in \{0, 1\} \times \{A, B\}$$

or

$$(x, y) = (1, B) \Rightarrow x = 1 \in \{0, 1\} \wedge y = B \in \{A, B\} \text{ so that } (x, y) \in \{0, 1\} \times \{A, B\}$$

proving that

$$f \subseteq \{0, 1\} \times \{A, B\}$$

If $(x, y), (x, y') \in f$ then for (x, y) we have either:

$\boldsymbol{(x, y) = (0, A)}$**.** Then $x = 0$ and $y = A$ so that $(x', y') = (0, y') \in f \Rightarrow y' = A$ hence $y = y'$

$\boldsymbol{(x, y) = (1, B)}$**.** Then $x = 1$ and $y = B$ so that $(x', y') = (1, y') \in f \Rightarrow y' = B$ hence $y = y'$

which proves that

$$f: \{0, 1\} \rightarrow \{A, B\} \text{ is a partial function}$$

If $x \in \{0,1\}$ then either $x=0$ so that $(0,A) \in f$ or $x=1$ so that $(1,B) \in f$, so $\{0,1\} \subseteq \operatorname{dom}(f)$. Using [proposition: 2.26] it follows that

$$f: \{0,1\} \rightarrow \{A,B\} \text{ is a function}$$ □

Although the composition of functions $f: A \rightarrow B$ and $g: C \rightarrow D$ is a partial function [see theorem: 2.19], it does not have to be a function as we need the extra requirement that $\operatorname{dom}(g \circ f) = A$. So we must have a extra condition on C. This is expressed in the following theorem,

Theorem 2.28. *Let $f: A \rightarrow B$ and $g: C \rightarrow D$ functions with $f(A) \subseteq C$ then $g \circ f: C \rightarrow D$ is also a function with* $\operatorname{range}(g \circ f) = g(\operatorname{range}(f))$

Proof. Using [theorem: 2.19] we have that

$$g \circ f: A \rightarrow D \text{ is a partial function}$$

Using [theorem: 2.25] we have that $A \subseteq f^{-1}(f(A))$ and by [theorem: 2.17] together with $f(A) \subseteq C$ we have $f^{-1}(f(A)) \subseteq f^{-1}(C)$ proving that

$$A \subseteq f^{-1}(C) \tag{2.14}$$

Further using [theorem: 2.22] we have

$$\begin{aligned}
\operatorname{dom}(g \circ f) &= \operatorname{dom}(f) \bigcap f^{-1}(\operatorname{dom}(g)) \\
&\underset{f,g \text{ are functions}}{=} A \bigcap f^{-1}(C) \\
&\underset{[\text{theorem: } 2.14]}{=} A
\end{aligned}$$

which proves that

$$g \circ f \text{ is a function}$$

Finally

$$\begin{aligned}
\operatorname{range}(g \circ f) &\underset{[\text{theorem: } 2.22}{=} g\big(\operatorname{range}(f) \bigcap \operatorname{dom}(g)\big) \\
&\underset{f \text{ is a function}}{=} g\big(\operatorname{range}(f) \bigcap C\big) \\
&\underset{[\text{theorem: } 2.17]}{=} g\big(f(A) \bigcap C\big) \\
&\underset{f(A) \subseteq C}{=} g(f(A)) \\
&\underset{[\text{theorem: } 2.17]}{=} g(\operatorname{range}(f))
\end{aligned}$$

□

Next we define the class of all the graphs of functions between two classes

Note 2.29. Be aware that some books calls partial functions functions and functions mappings.

Definition 2.30. *Let A, B be two classes then we define the class B^A [using the Axiom of Construction] as*

$$B^A = \{f \mid f: A \rightarrow B \textit{ is a function}\}$$

Note 2.31. B^A is not the class of functions between A and B, but the class of graphs of functions between A and B. This distinction is important because it makes the following theorem possible.

Example 2.32. Let A be a class then $A^{\varnothing} = \{\varnothing\}$

Proof. Let $f \in A^{\varnothing}$ then $f: \varnothing \Rightarrow A$ is a function, so that $f \subseteq \varnothing \times A = \varnothing$ or $f = \varnothing$ □

Lemma 2.33. *If $f: A \rightarrow B$ is a function and $B \subseteq C$ then $f: A \rightarrow C$ is a function*

Proof. As $f: A \rightarrow B$ is a function we have $f \subseteq A \times B$ which as by [theorem: 1.48] $A \times B \subseteq A \times C$ means that $f \subseteq A \times C$'. Further as $f: A \rightarrow B$ is a function we we have also $\operatorname{dom}(f) = A$ and if $(x,y),(x,y') \in f$ then $y = y'$. So by definition$f: A \rightarrow C$ is a function. □

Theorem 2.34. *Let A, B, C be classes such that $B \subseteq C$ then $B^A \subseteq C^A$*

Proof. Let $f \in B^A$ then $f: A \rightarrow B$ is a function, using the above lemma [lemma: 2.33] we have that $f: A \rightarrow C$ is a function, hence $f \in C^A$ proving that

$$B^A \subseteq C^A$$

□

We have also the following relation between $A \times B$ and B^C

Theorem 2.35. *Let A, B be two classes then we have:*

1. $B^A \subseteq A \times B$
2. *If A, B are sets then B^A is a set*

Proof.

1. If $f \in B^A$ then $f: A \rightarrow B$ is a function so that $f \subseteq A \times B$ proving that $B^A \subseteq A \times B$
2. If A, B are sets then by [theorem: 1.67] we have that $A \times B$ is a set. So using the Axiom of Subsets [axiom: 1.54] we have that f is a set, □

Theorem 2.36. *Let A, B, C be classes then $A^C \bigcap B^C = (A \bigcap B)^C$*

Proof. First by [theorem: 1.25] we have $A \bigcap B \subseteq A$ and $A \bigcap B \subseteq B$ it follows from the above theorem [theorem: 2.34] that $(A \bigcap B)^C \subseteq A^C$ and $(A \bigcap B)^C \subseteq B^C$. Applying then [theorem: 1.26] gives

$$\left(A \bigcap B\right)^C \subseteq A^C \bigcap B^C \tag{2.15}$$

For the opposite inclusion, let $f \in A^C \bigcap B^C$ then $f \in A^C \wedge f \in B^C$ so that $f: C \rightarrow A$ and $f: C \rightarrow B$ are functions. Then we have that $f \subseteq C \times A$ and $f \subseteq C \times B$ so that

$$f \subseteq (C \times A) \bigcap (C \times B) \underset{1.49}{=} \left(C \bigcap C\right) \times \left(A \bigcap B\right) \underset{[\text{theorem: } 1.30]}{=} C \times \left(A \bigcap B\right)$$

Further as $f: A \rightarrow C$ is a function we have $(x, y), (x, y') \in f$ and $\operatorname{dom}(f) = C$ so that

$$f: C \rightarrow \left(A \bigcap B\right) \text{ is a function}$$

proving that $f \in (A \bigcap B)^I$. So $A^C \bigcap B^C \subseteq (A \bigcap B)^C$ which combined with [eq: 2.15] gives

$$A^C \bigcap B^C = \left(A \bigcap B\right)^C$$

□

We have the follow trivial fact about a function

Proposition 2.37. *Let $f: A \rightarrow B$ be a function then if $f(A) \underset{[\text{theorem: } 2.17]}{=} \operatorname{range}(f) \subseteq C$ we have that $f: A \rightarrow C$ is a function.*

Proof. If $(x, y) \in f$ then $y \in \operatorname{range}(f)$ hence as $\operatorname{range}(f) \subseteq C$ $y \in C$. As $f \subseteq A \times B$ we have also $x \in A$ so that $(x, y) \in C \times B$. Hence $f \subseteq A \times C$, further if $(x, y), (x, y') \in f$ we have as $f: A \rightarrow B$ is a function that $y = y'$. So

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

As $\operatorname{range}(f) = A$ (because $f: A \rightarrow B$ is a function] we have that $f: A \rightarrow C$ a function □

We have the following trivial proposition about the equality of two functions

Proposition 2.38. *Two functions $f: A \rightarrow B$ and $g: A \rightarrow B$ are equal if*

$$[(x, y) \in f \Rightarrow (x, y) \in g \wedge (x, y) \in g \Rightarrow (x, y) \in f]$$

Proof. Note that the statement $f: A \rightarrow B$ and $g: A \rightarrow B$ are equal is equivalent with $\langle A, B, f \rangle = \langle A, B, g \rangle$, which by 2.5 is equivalent with $A = A \wedge B = B \wedge f = g$, As $A = A$ and $B = B$ are true this is equivalent with $f = g$. Now by the Axiom of Extent [axiom: 1.5] we have that

$$f = g \Leftrightarrow [(x, y) \in f \Rightarrow (x, y) \in g \wedge (x, y) \in g \Rightarrow (x, y) \in f]$$

□

If $f: A \rightarrow B$ is a function then for every $x \in A$ we have a unique $y \in B$ such that $(x, y) \in f$. Furthermore in many cases we have actually a expression valid for every $x \in A$ to calculate this unique value. To express this we use the following notation.

Definition 2.39. *If $f: A \rightarrow B$ is a function then*

$$\boldsymbol{y = f(x)} \textit{ or } \boldsymbol{f(x) = y} \textit{ is equivalent with } \boldsymbol{(x, y) \in f}$$

and

$$\boldsymbol{f(x) = E(x)} \textit{ where } \boldsymbol{E(x)} \textit{ is a expression depending on } \boldsymbol{x} \textit{ is equivalent with } \boldsymbol{(x, E(x)) \in f}$$

Further if $D \subseteq B$ then $\boldsymbol{f(x) \in D}$ *is the same as* $\boldsymbol{\exists y \in D}$ ***such that*** $\boldsymbol{y = f(x)}$ ***or*** $\boldsymbol{(x, y) \in f}$

Example 2.40. Let $3 \cdot x + 1$ be the value associated with x, so $f = \{z | z = (x, 3 \cdot x + 1) \in f \wedge x \in A\}$, then we can use the following equivalent notations to define our function

$$f: A \rightarrow B \text{ is defined by } x \rightarrow 3 \cdot x + 1$$

If we have defined a function $f: A \rightarrow B$ using a expression and we want to refer to the expression of the function we use the notation $f(x)$. Hence we define a function also as

$$f: A \rightarrow B \text{ is defined by } x \rightarrow f(x) = 3 \cdot x + 1$$

or

$$f: A \rightarrow B \text{ is defined by } x \rightarrow f(x) \text{ where } f(x) = 3 \cdot x + 1$$

or

$$f: A \rightarrow B \text{ is defined by } f(x) = 3 * x + 1$$

In all of the above cases we actually means that $\langle f, A, B \rangle$ is a function with $f = \{z | z = (x, 3 \cdot x + 1) \wedge x \in A\}$.

Using the above notation we can reformulate [proposition: 2.38] in a form that is easier to work with if we use expressions to define a function.

Proposition 2.41. *Two functions $f: A \rightarrow B$ and $g: A \rightarrow B$ are equal if and only if*

$$\forall x \in A \;\; f(x) = g(x)$$

Proof. Assume that $f: A \rightarrow B$ and $g: A \rightarrow B$ are equal then if $x \in A$ we have $\exists y \in B$ such that $(x, y) \in f$ or $y = f(x)$, using [proposition: 2.38] we have also $(x, y) \in g$ hence $y = g(x)$ which proves that $f(x) = g(x)$.
On the other hand assume that $\forall x \in A \;\; f(x) = g(x)$ then if $(x, y) \in f$ we have $y = f(x) = g(x)$ so that $(x, y) \in g$. If $(x, y) \in g$ then $y = g(x) = f(x)$ or $(x, y) \in g$. Using [proposition: 2.38] we have then that $f: A \rightarrow B$ and $g: A \rightarrow B$ are equal. □

Using the new notation, composition of function is written as

Theorem 2.42. *If $f: A \rightarrow B$ and $g: C \rightarrow D$ are two functions with $f(A) \subseteq C$ then*

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

Proof. Take $z = (g \circ f)(x)$ then $(x, z) \in g \circ f$ so that $\exists y$ such that $(x, y) \in f$ and $(y, z) \in g$. Hence $y = f(x)$ and $z = g(y)$ so that $z = g(f(x))$, proving $(g \circ f)(x) = g(f(x))$. □

Image and pre-image can also be expressed in the new notation.

Proposition 2.43. *Let $f: A \rightarrow B$ a function, $C \subseteq A$ and $D \subseteq B$ then*

1. *$y \in f(C) \Leftrightarrow \exists x \in A$ such that $y = f(x)$*
2. *$x \in f^{-1}(D) \Leftrightarrow f(x) \in D$*

Proof.

1.

$$\begin{aligned} y \in f(C) &\Leftrightarrow \exists x \in C \text{ such that } (x, y) \in f \\ &\Leftrightarrow \exists x \in C \text{ such that } y = f(x) \end{aligned}$$

2.

$$\begin{aligned} x \in f^{-1}(C) &\Leftrightarrow \exists y \in D \text{ such that } (x, y) \in f \\ &\Leftrightarrow \exists y \in D \text{ such that } y = f(x) \\ &\Leftrightarrow f(x) \in D \end{aligned}$$

□

Let's now look at some example of functions:

Example 2.44. (Empty Function) $\varnothing: \varnothing \to B$

Proof. First $\varnothing \subseteq \varnothing \times B$ by [theorem: 1.18], if $x \in \mathrm{dom}(\varnothing)$ then $\exists y \in \varnothing$ such that $(x, y) \in \varnothing$ which is a contradiction, so by [theorem: 1.19] we have that $\mathrm{dom}(\varnothing) = \varnothing$. And finally $(x, y) \in \varnothing \wedge (x, y') \in \varnothing \Rightarrow y = y'$ is satisfied vacuously as $(x, y) \in \varnothing \wedge (x, y') \in \varnothing$ is never true. □

Example 2.45. (Constant Function) Let A, B classes and $c \in B$ then $C_c: A \to B$ is defined by $C_c(x) = c$ or formally $C_c = \{z | z = (x, c) | x \in A\} = A \times \{c\}$

Proof. If $(x, y) \in C_c$ then $x \in A$ and $y = c \in B$ so that $C_c \subseteq A \times B$. If $(x, y) \in C_c \wedge (x, y') \in C_c$ then $y = c \wedge y' = c$ so that $y = y'$. So

$$C_c: A \to B \text{ is a partial function}$$

Finally if $x \in A$ then $(x, c) \in C_c$ so that $A \subseteq \mathrm{dom}(C_c)$ which by [proposition: 2.26] proves that

$$C_c: A \to B \text{ is a function}$$

□

Example 2.46. (Characteristics Function) Let A be a class and $B \subseteq A$ then $\mathcal{X}_{A,B}: A \to \{0, 1\}$ is defined by $\mathcal{X}_{A,B} = (B \times \{1\}) \bigcup ((A \setminus B) \times \{0\})$ [so that $\mathcal{X}_{A,B}(x) = \begin{cases} 1 \text{ if } x \in B \\ 0 \text{ if } x \in A \setminus B \end{cases}$

Proof. If $(x, y) \in \mathcal{X}_{A,B}$ then either $(x, y) \in (B \times \{1\}) \Rightarrow x \in B \underset{B \subseteq A}{\Rightarrow} x \in A$ and $y = 1 \in \{0, 1\}$ or $(x, y) \in ((A \setminus B), \{0\}) \Rightarrow x \in A \setminus B \Rightarrow x \in A$ and $y = 1 \in \{0, 1\}$ so that

$$\mathcal{X}_{A,B} \subseteq A \times \{0, 1\}$$

Also if $(x, y), (x, y') \in \mathcal{X}_{A,B}$ then for (x, y) we have either:

$\boldsymbol{(x, y) \in B \times \{1\}}$**.** then $x \in B$ so that $(x, y') \in B \times \{1\}$ hence $y = 1 = y'$

$\boldsymbol{(x, y) \in (A \setminus B) \times \{0\}}$**.** then $x \in A \setminus B$ so that $(x, y') \in (A \setminus B) \times \{0\}$ hence $y = 0 = y'$

or in all cases $y = y'$ and $x \in B \bigcup (A \setminus B) = A$. Hence $\mathcal{X}_{A,B}: A \to \{0, 1\}$ is a function. □

Example 2.47. (Identity Function) Let A be a class then $\mathrm{Id}_A: A \to B$ is defined by

$$\mathrm{Id}_A = \{z | z = (x, x) \wedge x \in A\}$$

Proof. Trivially we have $\mathrm{Id}_A \subseteq A \times A$. If $(x, y), (x, y') \in \mathrm{Id}_A$ then $(x, y) = (x, x) = (x, y')$ proving that $y = x = y'$. Hence $I_d: A \to A$ is a partial function. Further if $x \in A$ then $(x, x) \in \mathrm{Id}_A$ so that $x \in \mathrm{dom}(\mathrm{Id}_A)$ or $\mathrm{dom}(\mathrm{Id}_A) \subseteq A$ which by [proposition: 2.26] proves that

$$\mathrm{Id}_A: A \to A \text{ is a function}$$

□

Proposition 2.48. *Let $f: A \to B$ be a partial function then $f = f \circ \mathrm{Id}_A$ and $f = \mathrm{Id}_B \circ f$*

Proof.

1. If $(x, y) \in f$ then as $f \subseteq A \times B$ we have $x \in A \wedge x \in B$, by the definition of Id_A we have $(x, x) \in \mathrm{Id}_A$, as $(x, y) \in f$ we have $(x, y) \in \mathrm{Id}_A \circ f$. If $(x, y) \in f \circ \mathrm{Id}_A$ then $\exists x'$ such that $(x, x') \in \mathrm{Id}_A \wedge (x', y) \in f$. By definition of Id_A we have that $\exists z \in A$ such that $(x, x') = (z, z)$ hence $x = x'$ so that $(x, y) \in f$. Using the Axiom of Extent [axiom: 1.5] we have then that

$$f = f \circ \mathrm{Id}_A$$

2. If $(x, y) \in f$ then as $f \subseteq A \times B$ we have $x \in A \wedge x \in B$, by the definition of Id_B we have $(y, y) \in \mathrm{Id}_B$, so $(x, y) \in \mathrm{Id}_B \circ f$. If $(x, y) \in \mathrm{Id}_B \circ f$ then $\exists y'$ such that $(x, y') \in f \wedge (y, y')$, from the definition of Id_B we have that $y = y'$ so that $(x, y) \in f$. Using the Axiom of Extent [axiom: 1.5] we have then that

$$f = \mathrm{Id}_B \circ f$$

□

As a function $f: A \rightarrow B$ is a partial function with $\mathrm{dom}(f) = A$ we can refine [theorem: 2.17].

Theorem 2.49. *If $f: A \rightarrow B$ is a function $C \subseteq B$ and $D \subseteq B$ then we have*

1. $f(C) \subseteq B$
2. $f^{-1}(D) \subseteq A$
3. $f(A) = \mathrm{range}(f)$
4. $f^{-1}(B) = A$

Proof. This follows from 2.17 taking in account that $A = \mathrm{dom}(f)$ □

2.2.3 Injectivity, Surjectivity and bijectivity

First we define injectivity and surjectivity of partial functions.

Definition 2.50. *Let $f: A \rightarrow B$ be a partial function then we say that:*

1. *f is **injective** iff if $(x, y) \in f \wedge (x', y) \in f$ implies $x = x'$*
2. *f is **surjective** iff $\mathrm{range}(f) = B$*

Proposition 2.51. *A partial function $f: A \rightarrow B$ is surjective if $B \subseteq \mathrm{range}(f)$*

Proof. By [theorem: 2.11] $\mathrm{range}(f) \subseteq B$, so if $B \subseteq \mathrm{range}(f)$ it follows from [theorem: 1.8] that $B = \mathrm{range}(f)$, proving surjectivity. □

Using the notation $y = f(x)$ is the same as $(x, y) \in f$ we have

Theorem 2.52. *Let $f: A \rightarrow B$ be a function then*

1. *f is injective if and only if $\forall x, x \in A$ with $f(x) = f(x')$ we have $x = x'$*
2. *If $B \subseteq C$ and $f: A \rightarrow B$ is injective then $f: A \rightarrow C$ is injective*
3. *f is surjective if and only if $\forall y \in B$ there exists a $x \in A$ such that $y = f(x)$*

Proof.

1.
   ⇒. Let $x, x' \in A$ then if $y = f(x) = f(x')$ we have $(x, y) \in f$ and (x', y) so that $x = x'$

   ⇐. If $(x, y) \in f$ and $(x', y) \in f$ then $y = f(x) \wedge y = f(x')$ so that $f(x) = f(x')$ so that $x = x'$
2. This is trivial because injectivity is a property of the graph of a function.
3.
   ⇒. As $B = \mathrm{range}(f)$ we have $y \in B$ then $\exists x$ such that $(x, y) \in f \Rightarrow y = f(x)$ which as $f \subseteq A \times B$ proves that $x \in A$. So $\forall y \in B$ $\exists x \in A$ such that $y = f(x)$

   ⇐. Let $y \in B$ then $\exists x \in A$ such that $y = f(x)$ or $(x, y) \in f$ proving that $B \subseteq \mathrm{range}(f)$, using [proposition: 2.51] we have that f is surjective □

Example 2.53. Let A, B be classes, $B \subseteq A$ then $i_B: B \rightarrow A$ defined by $i_B = \{(x, x) | x \in B\}$ is a injective function. This function is called the **inclusion** function.

Proof. First if $(x, y) \in i_B$ then $\exists b \in B$ such that $(x, y) = (b, b)$ so that $x = b \in B \wedge y = b \in B \subseteq A$ proving that

$$i_B \subseteq B \times A$$

Further if $(x, y), (x, y') \in i_B$ then $\exists b, b' \in B$ such that $(x, y) = (b, b) \wedge (x, y') = (b', b')$, so that $x = b \wedge y = b \wedge x = b' \wedge y' = b'$, hence $y = y'$. So

$$i_B: B \rightarrow A \text{ is a partial function}$$

If $x \in B$ then $(x, x) \in i_B$ proving that $A \subseteq \mathrm{dom}(i_b)$ so using [proposition: 2.26] it follows that

$$i_B: B \rightarrow A \text{ is a function}$$

Finally if $(x, y), (x', y) \in i_B$ then there exists $b, b' \in B$ such that $(x, y) = (b, b) \wedge (x', y) = (b', b')$, so that $x = b \wedge y = b \wedge x' = b' \wedge y = b'$, hence $x = x'$, proving injectivity.

□

The following axiom ensures that the image of a set by a surjection is a set.

Axiom 2.54. (Axiom of Replacement) *If A is a set and $f: A \rightarrow B$ a surjection then B is a set.*

Proposition 2.55. *If $f: A \rightarrow B$ is a a function and $C \subseteq A, D \subseteq B$ then*

1. *$C \subseteq f^{-1}(f(C))$*
2. *If f is injective then $C = f^{-1}(f(C))$*
3. *If f is surjective then $D = f(f^{-1}(D))$*

Proof.

1. This is stated in [theorem: 2.25]
2. If $x \in f^{-1}(f(C))$ then $\exists y \in f(C)$ such that $(y, x) \in f^{-1}$, hence $(x, y) \in f$. As $y \in f(C)$ there exists a $x' \in C$ such that $(x', y) \in f$. Given that f is injective it follows from $(x, y), (x', y) \in f$ that $x = x'$, so as $x' \in C$ it follow that $x \in C$. Hence $f^{-1}(f(C)) \subseteq C$ which combined with (1) proves
$$C = f^{-1}(f(C))$$
3. If $y \in f(f^{-1}(D))$ then $\exists x \in f^{-1}(D)$ such that $(x, y) \in f$, hence $\exists z \in D$ such that $(z, x) \in f^{-1} \Rightarrow (x, z) \in f$, As f is a function we have $y = z$ so that $y \in D$. Hence
$$f(f^{-1}(D)) \subseteq D \tag{2.16}$$
If $y \in D$ then as f is a surjection there exist a $x \in A$ such that $(x, y) \in f$, hence $x \in f^{-1}(D)$ proving that $y \in f(f^{-1}(D))$. So $D \subseteq f(f^{-1}(D))$ which together with [eq: 2.16] proves
$$D = f(f^{-1}(D))$$
□

The importance of injectivity is that it allows us to define the inverse of a partial function. First we define the inverse graph of the graph of a partial function.

Definition 2.56. *Let $f: A \rightarrow B$ be a partial function then the **inverse of the graph f** noted as f^{-1} is defined by*

$$f^{-1} = \{z : z = (z, y) \text{ where } (y, x) \in f\}$$

Theorem 2.57. *Let $f: A \rightarrow B$ be a **injective** partial function then $f^{-1}: B \rightarrow A$ is a partial function*

Proof. If $(x, y) \in f^{-1}$ then $(y, x) \in f$ which, as $f \subseteq A \times B$, gives $(y, x) \in A \times B$, so $x \in B \wedge y \in A$, proving $(x, y) \in B \times Y$. Hence

$$f^{-1} \subseteq B \times A$$

Further if $(x, y) \in f^{-1}$ and $(x, y') \in f^{-1}$ then $(y, x) \in f \wedge (y, x') \in f$ which, as f is injectivity proves that $y = y'$. So all the conditions are satisfied to make $f^{-1}: B \rightarrow A$ a partial function. □

Note 2.58. The requirement that f is injective is needed to make f^{-1} is a partial function. For example assume that $A = \{1, 2, 3\}$, $B = \{10, 20\}$ and $f = \{(1, 10), (2, 10), (3, 20)\}$ then $f^{-1} = \{(10, 1), (10, 2), (20, 3)\}$ which is not the graph of a partial function.

If f is a injective function then the above theorem ensures that f^{-1} is a partial function however f^{-1} can be a graph of a function if we restrict the source of the inverse function.

Theorem 2.59. *If $f: A \rightarrow B$ is a injective function then $f^{-1}: f(A) \rightarrow A$ is a function*

Proof. First if $(x, y) \in f^{-1}$ then $(y, x) \in f \subseteq A \times B$ so that $y \in A \wedge x \in B$, as $(y, x) \in f$ we have that $x \in f(A)$, hence $(x, y) \in f(A) \times A$. So $f^{-1} \subseteq f(A) \times B$. Further if $(x, y), (x, y') \in f^{-1}$ then $(y, x), (y', x) \in f$ which as f is injective proves $y = y'$. Hence

$$f^{-1}: f(A) \rightarrow A \text{ is a partial function}$$

Further if $x \in f(A)$ then there exists a $y \in A$ such that $(y, x) \in f$, hence $(x, y) \in f^{-1}$ so that $x \in \mathrm{dom}(f^{-1})$, proving that $f(A) \subseteq \mathrm{dom}(f^{-1})$. Hence

$$f^{-1}: f(A) \rightarrow A \text{ is a function}$$ □

Corollary 2.60. *If $f: A \rightarrow B$ is a function, $A \neq \varnothing$ then $f: A \rightarrow B$ is injective if and only if there exist a function $g: B \rightarrow A$ such that $g \circ f = \mathrm{Id}_A$*

Proof.

⇒. Using the above [theorem: 2.59] we have that $f^{-1}: f(A) \rightarrow A$ is a function. As $A \neq 0$ there exist a $a \in A$ so we can consider the constant function $C_a: B \setminus f(A) \rightarrow A$ [see example: 2.45]. As $f(A) \bigcap (B \setminus f(A)) = \varnothing$ and $B = f(A) \bigcup (B \setminus f(A))$ we have by [theorem: 2.79] that

$$g = C_a \bigcup f^{-1}: B \rightarrow A$$

is a function. If $(x, y) \in g \circ f$ then $\exists$z such that $(x, z) \in f \wedge (z, y) \in g$. As $(x, z) \in f$ we have that $(z, x) \in f^{-1} \subseteq C_a \bigcup f^{-1} = g$, as also $(z, y) \in g$ and g is function, we have that $y = x$ so that $(x, y) = (x, x) \in \mathrm{Id}_A$ hence

$$g \circ f \subseteq \mathrm{Id}_A$$

Further if $(x, y) \in \mathrm{Id}_A$ then $x = y$, as $x \in A = \mathrm{dom}(f)$ there exist a $z \in B$ such that $(x, z) \in f \Rightarrow (z, x) \in f^{-1} \subseteq C_a \bigcup f^{-1} = g$ proving that $(x, y) = (x, x) \in g \circ f$. Hence

$$\mathrm{Id}_A \subseteq g \circ f$$

proving that

$$g \circ f = \mathrm{Id}_A$$

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

$$\begin{array}{rcl} (x, y), (x', y) \in f \subseteq A \times B & \underset{y \in B, \mathrm{dom}(g) = B}{\Rightarrow} & \exists z \vdash (y, z) \in g \\ & \Rightarrow & (x, z), (x', z) \in g \circ f = \mathrm{Id}_A \\ & \Rightarrow & x = z = x' \\ & \Rightarrow & x = x' \end{array}$$

□

Definition 2.61. *A function $f: A \rightarrow B$ is a* ***bijection*** *iff the function is* ***injective*** *and* ***surjective****.*

Definition 2.62. *Two classes A and B are bijective iff there exists a bijection between A and B*

Example 2.63. The function $\varnothing: \varnothing \rightarrow \varnothing$ is a bijection.

Proof. By [example: 2.44] $\varnothing: \varnothing \rightarrow \varnothing$ is a function. To prove that is a bijection we have:

injectivity. $\forall (x, y), (x', y) \in \varnothing$ we have $x = x'$ is satisfied vacuously.

surjectivity. $\forall y \in \varnothing$ there exist a $x \in \varnothing$ such that $(x, y) \in \varnothing$ is satisfied vacuously. □

Example 2.64. Let A be a class then $\mathrm{Id}_A: A \rightarrow A$ [example: 2.47] is a bijection

Proof. Let $(x, y) \in \mathrm{Id}_A \wedge (x', y) \in \mathrm{IdA}$ then $\exists z, z' \in A$ such that $(x, y) = (z, z) \wedge (x', y) = (z', z')$. So using [theorem: 1.43] $x = z \wedge y = z \wedge x = z' \wedge y = z'$. Using [theorem: 1.8] repeatedly gives then $x = x'$ proving that

$$\mathrm{Id}_A \text{ is injective}$$

If $y \in A$ then by definition $(y, y) \in \mathrm{Id}_A$ so that $\mathrm{range}(\mathrm{Id}_A) \subseteq A$. Using [theorem: 2.51] it follows that

$$\mathrm{Id}_A \text{ is surjective}$$ □

Example 2.65. Let $I = \{0\}$ B a class and take $f: I \rightarrow \{B\}$ defined by $f = \{(0, B)\}$ is a bijection

Proof. As $0 \in \{0\}$ and $B \in \{B\}$ it follows that $(0, B) \in \{0\} \times \{B\}$, hence $f = \{(0, B)\} \subseteq \{0\} \times \{B\}$. If $(x, y), (x, y') \in f = \{0\} \times \{B\}$ then $y = B = y'$, further $\operatorname{dom}(f) = \{0\} = I$. So we conclude that $f: \{0\} \rightarrow \{B\}$ is indeed a function. Further if $y \in \{B\}$ then $y = B$ and as $(0, B) \in f$ it follows that $y \in \operatorname{range}(f)$ or $\{B\} \subseteq \operatorname{range}(f)$, which by [theorem: 2.51] proves that f is surjective. Finally if $(x, y), (x', y) \in f = \{(0, B)\}$ then $x = 0 = x'$ proving that $f: \{0\} \rightarrow \{B\}$ is a bijection. □

Proposition 2.66. *If $f: A \rightarrow B$ is a injective function then $f: A \rightarrow f(A)$ is a bijection*

Proof. As injectivity is a property of the graph of a function, the function $f: A \rightarrow B$ is still injective. Further $\operatorname{range}(f) \underset{[\text{theorem: } 2.17]}{=} f(A)$ which proves surjectivity. □

Theorem 2.67. If *$f: A \rightarrow B$ is a bijection then $f^{-1}: B \rightarrow A$ is a function*

Proof. As $f: A \rightarrow B$ is injective and surjective we have that $f(A) = B$ and by [theorem: 2.59] that $f^{-1}: f(A) \rightarrow B$ is a function. Hence $f^{-1}: B \rightarrow A$ is a function. □

Theorem 2.68. *If $f: A \rightarrow B$ is a bijection then for $f^{-1}: B \rightarrow A$ we have*

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

Proof.

1. Let $y \in (f^{-1})(C)$ then there exist a $x \in C$ such that $(x, y) \in f^{-1}$, hence $(y, x) \in f$ so that $y \in f^{-1}(C)$ proving
$$(f^{-1})(C) \subseteq f^{-1}(C)$$
If $x \in f^{-1}(C)$ then there exist a $y \in C$ such that $(x, y) \in f$ hence $(y, x) \in f^{-1}$ proving that $x \in (f^{-1})(C)$, hence $f^{-1}(C) \subseteq (f^{-1})(C)$. Combining this with the above gives
$$(f^{-1})(C) = f^{-1}(C)$$
2. Let $x \in (f^{-1})^{-1}(C)$ then there exist a $y \in C$ such that $(x, y) \in f^{-1}$ hence $(y, x) \in f$ so that $x \in f(C)$ proving
$$(f^{-1})^{-1}(C) \subseteq f(C)$$
If $x \in f(C)$ then there exist a $y \in C$ such that $(y, x) \in f$ hence $(x, y) \in f^{-1}$ proving that $x \in (f^{-1})^{-1}(C)$, hence $f(C) \subseteq (f^{-1})^{-1}(C)$. Combining this with the above gives
$$(f^{-1})^{-1}(C) = f(C)$$
□

Theorem 2.69. *If $f: A \rightarrow B$ is bijective then*

1. *$f \circ f^{-1} = \mathrm{Id}_B$*
2. *$f^{-1} \circ f = \mathrm{Id}_A$*

Proof. First $f^{-1}: B \rightarrow A$ is a function by [theorem: 2.67].

1. Let $(x, y) \in f \circ f^{-1}$ then $\exists z$ such that $(x, z) \in f^{-1} \Rightarrow (z, x)$ and $(z, y) \in f$. As f^{-1} is the graph of a function we have that $x = y$. Further from $(x, z) \in f^{-1} \subseteq B \times A$ it follow that $x \in B$. Hence $(x, y) = (x, x) \in \mathrm{Id}_B$, proving that
$$f \circ f^{-1} \subseteq \mathrm{Id}_B \tag{2.17}$$
If $(x, y) \in \mathrm{Id}_B$ then $\exists z \in B$ such that $(x, y) = (z, z)$ so that $x = y \in B$, As $B = \operatorname{dom}(f^{-1})$ there exists a u such that $(y, u) \in f^{-1} \Rightarrow (u, y) \in f$ so that $(y, y) \in f \circ f^{-1} \underset{x=y}{\Rightarrow} (x, y) \in f \circ f^{-1}$. So $\mathrm{Id}_B \subseteq f \circ f^{-1}$. Combining this with [eq: 2.17] proves that
$$f \circ f^{-1} = \mathrm{Id}_B$$
2. Let $(x, y) \in f^{-1} \circ f$ then $\exists z$ such that $(x, z) \in f \Rightarrow (z, x) \in f^{-1}$ and $(z, y) \in f^{-1}$. As f^{-1} is the graph of a function we have that $x = y$. Further from $(x, z) \in f \subseteq A \times B$ it follows that $x \in A$. Hence $(x, y) = (x, x) \in \mathrm{Id}_A$, proving that
$$f^{-1} \circ f \subseteq I_A \tag{2.18}$$

If $(x, y) \in \mathrm{Id}_A$ then $\exists z \in A$ such that $(x, y) = (z, z)$ so that $x = y \in A$, As $A = \mathrm{dom}(f)$ there exists a u such that $(x, u) \in f \Rightarrow (u, x) \in f^{-1}$ so that $(x, x) \in f^{-1} \circ f \underset{x=y}{\Rightarrow} (x, y) \in f^{-1} \circ f$. So $\mathrm{Id}_B \subseteq f^{-1} \circ f$. Combining this with [eq: 2.18] proves that

$$f^{-1} \circ f = \mathrm{Id}_A$$

□

Corollary 2.70. *If $f: A \rightarrow B$ is bijection then*

1. *$\forall x \in A$ we have $(f^{-1})(f(x)) = x$*
2. *$\forall y \in B$ we have $f((f^{-1})(y)) = y$*

Proof.

1. If $x \in A$ then $(f^{-1})(f(x)) = ((f^{-1}) \circ f)(x) \underset{\text{[theorem:}}{=} \mathrm{Id}_A(x) = x$
2. If $y \in B$ then $f((f^{-1})(y)) \underset{\text{[theorem:}}{=} \mathrm{Id}_B(y) = y$

□

Corollary 2.71. *Let $f: A \rightarrow B$ a function then the following are equivalent:*

1. *$f: A \rightarrow B$ is a bijection*
2. *There exists a function $g: B \rightarrow A$ such that $f \circ g = \mathrm{id}_B$ and $g \circ f = \mathrm{Id}_A$*

Proof.

1 ⇒ 2. This follows from [theorem: 2.69] by taking $g = f^{-1}$

2 ⇒ 1. Let $(x, y), (x', y) \in f \subseteq A \times B$, as $y = \mathrm{dom}(g)$ there exists a z such that $(y, z) \in g$, hence $(x, z), (x', z) \in g \circ f = \mathrm{Id}_A$ so that $x = z = x'$ proving that

$$f: A \rightarrow B \text{ is injective}$$

Further if $y \in B$ then $(y, y) \in \mathrm{Id}_B = f \circ g$ so there exists a $z \in A$ such that $(y, z) \in g$ and $(z, y) \in f$. Proving that $B \subseteq \mathrm{range}(f)$ so by [proposition: 2.51]

$$f: A \rightarrow B \text{ is a surjection}$$

□

The inverse of a bijection is again a bijection

Corollary 2.72. *If $f: A \rightarrow B$ is a bijection then $f^{-1}: B \rightarrow A$ is a bijection*

Proof. If $f: A \rightarrow B$ is a bijection then by [theorem: 2.69] $f \circ f^{-1} = \mathrm{Id}_B$ and $f^{-1} \circ f = \mathrm{Id}_A$ which by [theorem: 2.71] proves that $f^{-1}: B \rightarrow A$ is a bijection. □

Proposition 2.73. *If $f: A \rightarrow B$ is a bijection then we have:*

1. *If $g: B \rightarrow A$ is such that $f \circ g = \mathrm{Id}_B$ and $g \circ f = \mathrm{Id}_A$ then $g = f^{-1}$*
2. *$(f^{-1})^{-1} = f$*

Proof.

1. We have

$$\begin{aligned}
f \circ g = \mathrm{Id}_B \quad &\Rightarrow \quad f^{-1} \circ (f \circ g) = f^{-1} \circ \mathrm{Id}_B \\
&\underset{\text{[proposition: 2.48}}{\Rightarrow} f^{-1} \circ (f \circ g) = f^{-1} \\
&\underset{\text{[theorem: 2.21]}}{\Rightarrow} (f^{-1} \circ f) \circ g = f^{-1} \\
&\underset{\text{[function: 2.69]}}{\Rightarrow} \mathrm{Id}_B \circ g = f^{-1} \\
&\underset{\text{[proposition: 2.48}}{\Rightarrow} g = f^{-1}
\end{aligned}$$

2. We have

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

which by the Axiom of Extent [axiom: 1.5] proves

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

□

Composition preserves injectivity, surjectivity and bijectivity.

Theorem 2.74. *We have*

1. If $f: A \to B$ *and* $g: C \to D$ *are injective functions with* $f(A) \subseteq C$ *then* $g \circ f: A \to D$ *is a injective function.*
2. *If* $f: A \to B$ *and* $g: C \to D$ *are injective functions with* $f(A) \subseteq C$ *then* $g \circ f: A \to g(f(A))$ *is a bijective function.*
3. *If* $f: A \to B$ *is a function and* $g: C \to D$ *a surjective function so that* $f(A) = C$ *then* $g \circ f: A \to D$ *is a surjective function.*
4. *If* $f: A \to B$ *is a injective function and* $g: C \to D$ *a bijective function so that* $f(A) = C$ *then* $g \circ f: A \to D$ *is a bijective function.*
5. *If* $f: A \to B$ *is a injective function and* $g: C \to D$ *a bijective function so that* $f(A) = C$ *then* $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

Proof.

1. Let $(x, z), (x', z) \in g \circ f$ then $\exists u, v$ such that

$$(x, u) \in f \wedge (x', v) \in f \wedge (u, y) \in g \wedge (v, y) \in g$$

As g is injective we have $u = v$, but that means from the above that $(x, u) \in f \wedge (x', u) \in f$, which as f is injective proves

$$x = x'$$

2. Using (1) we have that $g \circ f: A \to D$ is injective so that by [theorem: 2.66] $g \circ f: A \to (g \circ f)(A)$ is a bijection. Further by [theorem: 2.23] $(g \circ f)(A) = g(f(A))$ so that $g \circ f: A \to g(f(A))$ is a bijection.
3. Let $z \in D$ then as g is surjective there $\exists y \in C$ such that $(y, z) \in g$. As $f(A) = C$ there exists a $x \in A$ such that $(x, y) \in f$. But then $(x, z) \in g \circ f$ proving that $g \circ f$ is surjective.
4. Using (1) and (2) proves that $g \circ f: A \to D$ is injective and surjective and thus by definition bijective.
5. By (3) $g \circ f$ is a bijection, so by [theorem: 2.69] we have that

$$\begin{aligned}
(g \circ f)^{-1} \circ (g \circ f) = \mathrm{Id}_A &\underset{\text{[associativity: 2.21]}}{\Rightarrow} ((g \circ f)^{-1} \circ g) \circ f = \mathrm{Id}_A \\
&\Rightarrow (((g \circ f)^{-1} \circ g) \circ f) \circ f^{-1} = \mathrm{Id}_A \circ f^{-1} \\
&\underset{\text{[proposition: 2.48]}}{\Rightarrow} (((g \circ f)^{-1} \circ g) \circ f) \circ f^{-1} = f^{-1} \\
&\underset{\text{[associativity: 2.21]}}{\Rightarrow} ((g \circ f)^{-1} \circ g) \circ (f \circ f^{-1}) = f^{-1} \\
&\underset{\text{[theorem: 2.69]}}{\Rightarrow} ((g \circ f)^{-1} \circ g) \circ \mathrm{Id}_B = f^{-1} \\
&\underset{\text{[proposition: 2.48]}}{\Rightarrow} (g \circ f)^{-1} \circ g = f^{-1} \\
&\Rightarrow ((g \circ f)^{-1} \circ g) \circ g^{-1} = f^{-1} \circ g^{-1} \\
&\underset{\text{[associativity: 2.21]}}{\Rightarrow} (g \circ f^{-1}) \circ (g \circ g^{-1}) = f^{-1} \circ g^{-1} \\
&\underset{\text{[theorem: 2.69]}}{\Rightarrow} (g \circ f)^{-1} \circ \mathrm{Id}_A = f^{-1} \circ g^{-1} \\
&\underset{\text{[proposition: 2.48]}}{\Rightarrow} (g \circ f)^{-1} = f^{-1} \circ g^{-1}
\end{aligned}$$

□

In the special case that $B = C$ we have

Corollary 2.75. *We have*

1. If $f: A \to B$ *and* $g: B \to C$ *are injective functions then* $g \circ f: A \to C$ *is a injective function.*
2. *If* $f: A \to B$ *and* $g: B \to C$ *are surjective functions then* $g \circ f: A \to C$ *is a surjective function.*
3. *If* $f: A \to B$ *and* $g: B \to C$ *are bijective function then* $g \circ f: A \to C$ *is a bijective function.*
4. *If* $f: A \to B$ *and* $g: B \to C$ *are bijective function then* $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

Proof.

1. This follows from [theorem: 2.74 (1)] because $f(A) \subseteq B$.
2. This follows from [theorem: 2.74 (2)] because if f is surjective we have $f(A) = B$.
3. This follows from (1) and (2)
4. This follows from [theorem: 2.74 (4)] because if f is bijective, hence surjective, we have $f(A) = B$ □

The following is a example of a bijection between a class and the class of functions in this set.

Theorem 2.76. *Let A be a class then there exists a bijection between A and $A^{\{0\}}$*

Proof. Given $x \in A$ define the function $f_x: \{0\} \to \{x\}$ where $f_x = \{(0, x)\}$ [see [example: 2.65] to prove that this is a function (even a bijection)]. So $f_x \in \{x\}^{\{0\}}$, which as $\{x\} \subseteq A$ proves by [theorem: 2.34] that $f_x \in A^{\{0\}}$. Define now $f = \{z | z = (x, f_x) \text{ where } x \in A\}$. If $(x, y) \in f$ we have $x \in A$ and thus $y = f_x \in A^{\{0\}}$ hence $(x, y) \in A \times A^{\{0\}}$. Also if $(x, y), (x, y') \in A$ then $y = f_x$ and $y' = f_x$ so that $y = y'$. Further for every $x \in A$ we have by the definition of f that $(x, f_x) \in f$. So we conclude that

$$f: A \to A^{\{0\}} \text{ is a function}$$

Assume now that $(x, y), (x', y) \in f$ then $f_x = y = f_{x'}$, so that $\{(0, x)\} = \{(0, x')\}$, hence $(0, x) = (0, x')$, from which it follows that $x = x'$. this proves that

$$f: A \to A^{\{0\}} \text{ is a injective function}$$

If $y \in A^{\{0\}}$ then $y: \{0\} \to A$ is a function, hence $0 \in \{0\} = \text{dom}(y)$, so there exists a z such that $(0, z) \in y \subseteq \{0\} \times A$ proving that $z \in A$. Hence

$$\{(0, z)\} \subseteq y \wedge z \in A \tag{2.19}$$

If $(u, v) \in y \subseteq \{0\} \times A$ then $u = 0$ so that $(0, u) \in y$, which, as $(0, z) \in y$ and y is a function, proves that $u = z$ or $(u, v) = (0, z) \in \{(0, z)\}$. So $y \subseteq \{(0, z)\}$ which combined with [eq: 2.19] proves that $\{(0, z)\} = y$. As $f_z = \{(0, z)\} = y$ we have that $(z, y) \in f$ which proves that

$$f \text{ is a surjection}$$ □

Theorem 2.77. *If A is a class then there is a bijection between $\mathcal{P}(A)$ and $\{0, 1\}^A$ where $0 = \varnothing$ and $1 = \{\varnothing\}$ are different elements.*

Proof. Define $\gamma: \mathcal{P}(A) \to \{0, 1\}^A$ by $\gamma = \{z | z = (B, \mathcal{X}_{A,B}) \text{ where } B \in \mathcal{P}(A)\}$ where $\mathcal{X}_{A,B} = (B \times \{1\}) \bigcup ((A \setminus B) \times \{0\})$ is the graph of the Characteristic function [example: 2.46]. If $(B, f) \in \gamma$ then $B \in \mathcal{P}(A)$ and $f = \mathcal{X}_{A,B}$, as $B \in \mathcal{P}(A) \Rightarrow B \subseteq A$ it follow using [example: 2.46] that $\mathcal{X}_{A,B}: A \to \{0, 1\}$ is a function. So $(B, f) \in \{0, 1\}^A$ giving

$$\gamma \subseteq \mathcal{P}(A) \times (\{0, 1\}^A)$$

If $(B, f), (B, g) \in \gamma$ then $f = \mathcal{X}_{A,B}$ and $g = \mathcal{X}_{A,B}$ so that $f = g$, also by the definition of γ we have that $\text{dom}(\gamma) = \mathcal{P}(A)$, hence

$$\gamma: \mathcal{P}(A) \to \{0, 1\}^A \text{ is a function}$$

If $(B, f), (B', f) \in \gamma$ then $\mathcal{X}_{A,B} = \mathcal{X}_{A,B'}$ so that

$$\begin{aligned} x \in B \quad & \Leftrightarrow \quad \mathcal{X}_{A,B}(x) = 1 \\ & \underset{\mathcal{X}_{A,B} = \mathcal{X}_{A,B'}}{\Leftrightarrow} \quad \mathcal{X}_{A,B'}(x) = 1 \\ & \Leftrightarrow \quad x \in B' \end{aligned}$$

proving that $B = B'$. Hence

$$\gamma: \mathcal{P}(A) \to \{0, 1\}^A \text{ is injective}$$

Let $f \in \{0, 1\}^A$, define $B = \{x \in A | (x, 1) \in f\} \subseteq A$, then $B \in \mathcal{P}(A)$.
If $(x, y) \in f$ then we have for x either:

***x* ∈ *B*.** Then $(x, 1) \in f$ and as $(x, y) \in f$ we have that $y = 1$ so that $(x, y) = (x, 1) \in \mathcal{X}_{A,B}$

***x* ∉ *B*.** Then $(x, 0) \in f$ and as $(x, y) \in f$ we have that $y = 0$ so that $(x, y) = (x, 0) \in \mathcal{X}_{A,B}$ [as $x \in A \setminus B$]

proving that

$$f \subseteq \mathcal{X}_{A,B} \tag{2.20}$$

If $(x, y) \in \mathcal{X}_{A,B}$ then we have for x either:

- $\boldsymbol{x \in B}$**.** Then as $(x, 1) \in \mathcal{X}_{A,B}$ we must have that $y = 1$, using the definition of B we have also $(x, 1) \in f \Rightarrow (x, y) \in f$
- $\boldsymbol{x \notin B}$**.** Then $x \in A \setminus B$ so that $(x, 0) \in \mathcal{X}_{A,B}$ hence we must have that $y = 0$. As $(x, 0) \in f$ [if $(x, 1) \in f$ then $x \in B$ a contradiction] it follows that $(x, y) = (x, 0) \in f$

proving that $\mathcal{X}_{A,B} \subseteq f$, which combined with 2.20 gives

$$\mathcal{X}_{A,B} = f \tag{2.21}$$

So given $f \in \{0,1\}^A$ we have found a $B \in \mathcal{P}(A)$ such that $\mathcal{X}_{A,B} \underset{[\text{eq: } 2.21]}{=} f$, hence $(B, f) \in \gamma$ proving that

$$\gamma: \mathcal{P}(A) \rightarrow \{0,1\}^A \text{ is a surjective}$$

□

2.2.4 Restriction of a Function/Partial Function

Sometimes we only want to work with functions whose graphs satisfies certain conditions. It could be that the graph of a function does not satisfies these, but that the restriction of this graph to a sub-class satisfies the conditions. For example, the function $f: \mathbb{R} \rightarrow \mathbb{R}$ defined by $f(x) = \left\{ \begin{smallmatrix} 1 \text{ if } x < 1 \\ 0 \text{ if } 1 \leqslant x \end{smallmatrix} \right.$ is not continuous, as it is discontinuous at 1. However restricting this function to $\mathbb{R} \setminus \{1\}$ produces a continuous function. This is the idea of the next definition

Definition 2.78. *Let $f: A \rightarrow B$ be a partial function and $C \subseteq A$ a sub-class of A then the restriction of f to C noted by $f_{|C}$ is defined by*

$$f_{|C} = \{z | z = (x, y) \in f \wedge x \in C\} = f \bigcap (C \times B)$$

which defines the partial function

$$f_{|C}: C \rightarrow B$$

Proof. We must of course proof that $\{z|z = (x, y) \in f \wedge x \in C\} = f \bigcap (C \times B)$ and that $f_{|C}: C \rightarrow B$ is indeed a partial function. If $(x, y) \in \{z|z = (x, y) \in f \wedge x \in C\}$ then $(x, y) \in f \subseteq A \times B \Rightarrow y \in B$ and $x \in C$, so that $(x, y) \in f \wedge (x, y) \in C \times B$, hence $(x, y) \in f \bigcap (C \times B)$. If $(x, y) \in f \bigcap (C \times B)$ then $(x, y) \in f \wedge (x, y) \in C \times B \Rightarrow x \in C$, proving that $(x, y) \in \{z|z = (x, y) \in f \wedge x \in C\}$. So we have that

$$f_{|C} = \{z | z = (x, y) \in f \wedge x \in C\} = f \bigcap (C \times B)$$

From the above it follows, using [theorem: 1.25], that

$$f_{|C} \subseteq C \times B$$

Finally, if $(x, y), (x, y') \in f_{|C}$ then $(x, y), (x, y') \in f$ so that $y = y'$. Hence we have that $f_{|C}: C \rightarrow B$ is a partial function. □

Theorem 2.79. *Let $f: A \rightarrow C$ and $g: B \rightarrow C$ be two partial functions such that $A \bigcap B = \varnothing$ then*

1. *$f \bigcup g: A \bigcup B \rightarrow C$ is a partial function*
2. *$f = (f \bigcup g)_{|A}$ and $g = (f \bigcup g)_{|B}$*
3. *$\operatorname{dom}(f \bigcup g) = \operatorname{dom}(f) \bigcup \operatorname{dom}(g)$*
4. *If $f: A \rightarrow C$ and $g: B \rightarrow C$ are functions then $f \bigcup g: A \bigcup B \rightarrow C$ are functions*

Proof.

1. As $f: A \rightarrow C$ and $g: B \rightarrow C$ are functions we have that $f \subseteq A \times C$ and $g \subseteq B \times C$ so that by [theorem: 1.25]

   $$f \bigcup g \subseteq (A \times C) \bigcup (B \times C) \underset{[\text{theorem: } 1.49]}{=} (A \bigcup B) \times C$$

   Let $(x, y), (x, y') \in f \bigcup g$. Assume that $y \neq y'$ then we can not have that $(x, y), (x, y') \in f$ for then, as f is a function, we would have $y = y'$, likewise we can not have that $(x, y), (x, y') \in g$, for then, as g is a function, we would have that $y = y'$. So we must that either $(x, y) \in f \wedge (x, y') \in g$ or $(x, y) \in g \wedge (x, y') \in f$, but then we would have $x \in A \bigcap B$ which contradicts $A \bigcap B = \varnothing$. So we must have that $y = y'$. Summarized

   $$\text{If } (x, y), (x, y) \in f \bigcup g \text{ then we have } y = y'$$

2. As $f\subseteq A\times C$ we have by [theorem :1.25] that

$$f\bigcap(B\times C)\subseteq(A\times C)\bigcap(B\times C)\underset{\text{[theorem: 1.49]}}{=}\left(A\bigcap B\right)\times C=\varnothing\times C\underset{\text{[theorem: 1.47}}{=}\varnothing$$

proving using [theorem: 1.18] that

$$f\bigcap(B\times C)=\varnothing \tag{2.22}$$

As $g\subseteq B\times C$ we have by [theorem :1.25] that

$$g\bigcap(A\times C)\subseteq(B\times C)\bigcap(A\times C)\underset{\text{[theorem: 1.49]}}{=}\left(A\bigcap B\right)\times C=\varnothing\times C\underset{\text{[theorem: 1.47}}{=}\varnothing$$

proving using [theorem: 1.18 that

$$g\bigcap(A\times C)=\varnothing \tag{2.23}$$

Further we have

$$\begin{array}{rcl}\left(f\bigcup g\right)_{|A} & = & \left(f\bigcup g\right)\bigcap(A\times C)\\ & \underset{\text{[theorem: 1.30}}{=} & \left(f\bigcap(A\times C)\right)\bigcup\left(g\bigcap(A\times C)\right)\\ & \underset{\text{[eq: 2.23]}}{=} & \left(f\bigcap(A\times C)\right)\bigcup\varnothing\\ & \underset{\text{[theorem: 1.32]}}{=} & f\bigcap(A\times C)\\ & \underset{f\subseteq A\times C\,\text{snd}\,\text{[theorem:1.26]}}{=} & f\\ \left(f\bigcup g\right)_{|B} & = & \left(f\bigcup g\right)\bigcap(B\times C)\\ & \underset{\text{[theorem: 1.30}}{=} & \left(f\bigcap(B\times C)\right)\bigcup\left(g\bigcap(B\times C)\right)\\ & \underset{\text{[eq: 2.22]}}{=} & \varnothing\bigcup\left(\bigcap(B\times C)\right)\\ & \underset{\text{[theorem: 1.32]}}{=} & g\bigcap(B\times C)\\ & \underset{g\subseteq B\times C\,\text{snd}\,\text{[theorem:1.26]}}{=} & g\end{array}$$

3.

$$\begin{array}{rcl}x\in\operatorname{dom}\left(f\bigcup g\right) & \Leftrightarrow & \exists y\text{ such that }(x,y)\in f\bigcup g\\ & \Leftrightarrow & \exists y\text{ such that }(x,y)\in f\vee(x,y)\in g\\ & \Rightarrow & x\in\operatorname{dom}(f)\vee x\in\operatorname{dom}(g)\\ & \Rightarrow & x\in\operatorname{dom}(f)\bigcup\operatorname{dom}(g)\\ x\in\operatorname{dom}(f)\bigcup\operatorname{dom}(g) & \Rightarrow & x\in\operatorname{dom}(f)\vee x\in\operatorname{dom}(g)\\ & \Rightarrow & (\exists y\text{ such that }(x,y)\in f)\vee(\exists y'\text{ such that }(x,y)\in g)\\ & \Rightarrow & \left(\exists y\text{ such that }(x,y)\in f\bigcup g\right)\vee\left(\exists y'\text{ such that }(x,y)\in f\bigcup g\right)\\ & \Rightarrow & x\in\operatorname{dom}\left(f\bigcup g\right)\end{array}$$

so

$$\operatorname{dom}\left(f\bigcup g\right)=\operatorname{dom}(f)\bigcup\operatorname{dom}(g)$$

4. As $f:A\rightarrow C$ and $g:B\rightarrow C$ are functions we have that $A=\operatorname{dom}(f)$, $B=\operatorname{dom}(g)$. So that

$$\operatorname{dom}\left(f\bigcup g\right)\underset{(3)}{=}\operatorname{dom}(f)\bigcup\operatorname{dom}(g)=A\bigcup B$$

proving that

$$f\bigcup g:A\bigcup B\rightarrow C\text{ is a function}$$ □

Corollary 2.80. *Let $f:A\rightarrow B$ and $g:C\rightarrow D$ be functions such that $A\bigcap C=\varnothing$ then*

$$f\bigcup g:A\bigcup C\rightarrow B\bigcup D$$

is a function.

Proof. Using [theorem: 2.33] we have that $f: A \rightarrow B \bigcup D$ and $g: C \rightarrow B \bigcup D$ are functions. Applying then the previous theorem [theorem: 2.79] proves that $f \bigcup g: A \bigcup C \rightarrow B \bigcup D$ is a function. □

Corollary 2.81. *Let $f: A \rightarrow B$ and $g: C \rightarrow D$ be bijections with $A \bigcap C = \varnothing$ and $B \bigcap D = \varnothing$ then*

$$f \bigcup g: A \bigcup C \rightarrow B \bigcup D$$

is a bijection.

Proof. Using the previous theorem [theorem: 2.80] we have that $f \bigcup g: A \bigcup C \rightarrow B \bigcup D$ is a function. Now we have:

injectivity. If $(x, y), (x', y) \in f \bigcup g \subseteq (A \bigcup C) \times (B \bigcup D)$ we have the following possibilities for y:

- ***y* ∈ *B*.** As $f \subseteq A \times B$ and $g \subseteq C \times D$ we can not have $(x, y), (x', y) \in g$ [for then $y \in D \Rightarrow y \in B \bigcap D = \varnothing$], as g is injective we have $x = x'$.
- ***y* ∈ *D*.** As $f \subseteq A \times B$ and $g \subseteq C \times D$ we can not have $(x, y), (x', y) \in f$ [for then $y \in B \Rightarrow y \in B \bigcap D = \varnothing$], as f is injective we have $x = x'$.

so in all cases we have $x = x'$ proving injectivity of $f \bigcup g: A \bigcup C \rightarrow B \bigcup D$.

surjectivity. If $y \in B \bigcup D$ then we have either:

- ***y* ∈ *B*.** Then as f is surjective there exist a $x \in A \subseteq A \bigcup C$ such that $(x, y) \in f \subseteq f \bigcup g$.
- ***y* ∈ *D*.** Then as g is surjective there exist a $x \in C \subseteq A \bigcup C$ such that $(x, y) \in g \subseteq f \bigcup g$.

proving that in all cases there exist a $x \in A \bigcup C$ such that $(x, y) \in f \bigcup g$. □

Corollary 2.82. *Let $f: A \rightarrow B$ be a function a, b elements such that $a \notin A$ then*

$$g: A \bigcup \{a\} \rightarrow B \bigcup \{b\} \text{ defined by } g = \{(a, b)\} \bigcup f$$

is a function.

***Note* 2.83.** *A alternative definition of g is $g(x) = \begin{cases} b & \text{if } x = a \\ f(x) & \text{if } x \in A \end{cases}$*

Proof. Using [example: 2.45] we have that $C_b: \{a\} \rightarrow \{b\}$ where $C_b = \{(x, b) | x \in \{a\}\} = \{(a, b)\}$ is a function. As $A \bigcap \{a\}$ we can use the previous corollary [corollary: 2.80] so that

$$h: A \bigcup \{a\} \rightarrow B \bigcup \{b\} \text{ where } h = \{(a, b)\} \bigcup f \text{ is a function}$$ □

Theorem 2.84. *Let $f: A \rightarrow B$ be a partial function and $C \subseteq A$ a sub-class of A then we have:*

1. $\operatorname{dom}(f_{|C}) = C \bigcap \operatorname{dom}(f)$
2. $\operatorname{range}(f_{|C}) = f(C)$
3. *If $D \subseteq C$ then $f_{|C}(D) = f(D)$ and $(f_{|C})_{|D} = f_{|D}$*
4. *If $E \subseteq B$ then $(f_{|C})^{-1}(E) = C \bigcap f^{-1}(E)$*
5. *If $f: A \rightarrow B$ is injective then $f_{|C}: C \rightarrow B$ is injective*

Proof.

1. If $x \in \operatorname{dom}(f_{|C})$ then there exists a y such that $(x, y) \in f_{|C}$, hence $x \in C$ and $(x, y) \in f$ or $x \in C$ and $x \in \operatorname{dom}(f)$, so that $x \in C \bigcap \operatorname{dom}(f)$. Hence

$$\operatorname{dom}(f_{|C}) \subseteq C \bigcap \operatorname{dom}(f) \tag{2.24}$$

Further if $x \in C \bigcap \operatorname{dom}(f)$ then $x \in C$ and $x \in \operatorname{dom}(f)$, so there exists a y such that $(x, y) \in f$, hence $(x, y) \in f_{|C}$ or $x \in \operatorname{dom}(f_{|C})$. So $C \bigcap \operatorname{dom}(f) \subseteq \operatorname{dom}(f_{|C})$ which together with [eq: 2.24] gives

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

2. If $y \in \operatorname{range}(f_{|C})$ then $\exists x$ such that $(x, y) \in f_{|C}$, hence $(x, y) \in f$ and $x \in C$, so that $y \in f(C)$. On the other hand if $y \in f(C)$ there exists a $x \in C$ such that $(x, y) \in f$, hence $(x, y) \in f_{|C}$ so that $y \in \operatorname{range}(f_{|C})$. Hence using the Axiom of Extent [axiom: 1.5] we have

$$\operatorname{range}(f_{|C}) = f(C)$$

3. If $y \in f_{|C}(D)$ then $\exists x \in D$ such that $(x, y) \in f_{|C}$, hence $(x, y) \in f$ so that $y \in f(D)$. On the other hand if $y \in f(D)$ then $\exists x \in D$ such that $(x, y) \in f$, which as $x \in D \subseteq C \Rightarrow x \in C$ proves that $(x, y) \in f_{|C}$, so $y \in f_{|C}(D)$. Hence using the Axiom of Extent [axiom: 1.5] we have

$$f_{|C}(D) = f(D)$$

Further

$$(f_{|C})_{|D} = \left(f \bigcap (C \times B)\right) \bigcap (D \times B) \underset{D \times B \subseteq C \times B}{=} f \bigcap (D \times B) = f_{|D}$$

4. If $x \in (f_{|C})^{-1}(E)$ then there exist a $y \in E$ such that $(x, y) \in f_{|C}$, hence $x \in C$ and $(x, y) \in f \Rightarrow x \in f^{-1}(E)$, so that $x \in C \bigcap f^{-1}(E)$. Further if $x \in C \bigcap f^{-1}(E)$ then $x \in C$ and $x \in f^{-1}(E)$, so there exist a $y \in E$ such that $(x, y) \in f \underset{x \in C}{\Rightarrow} (x, y) \in f_{|C}$, hence $x \in (f_{|C})^{-1}(E)$. Hence using the Axiom of Extent [axiom: 1.5] we have

$$(f_{|C})^{-1}(E) = C \bigcap f^{-1}(E)$$

5. If $(x, y), (x', y) \in f_{|C}$ then as $f_{|C} \subseteq f$ we have $(x, y), (x', y) \in f$ which as f is injective proves $y = y'$ □

Theorem 2.85. *Let $f: A \rightarrow B$ be a **partial** function then $f_{|\operatorname{dom}(f)} = f$*

Proof. If $(x, y) \in f$ then by definition $x \in \operatorname{dom}(f)$ hence $(x, y) \in f_{|\operatorname{dom}(f)}$, further if $(x, y) \in f_{|\operatorname{dom}(f)}$ then $(x, y) \in f$ and $x \in \operatorname{dom}(f)$, so evidently $(x, y) \in f$. Hence using the Axiom of Extent [axiom: 1.5] we have

$$f_{|\operatorname{dom}(f)} = f$$ □

Theorem 2.86. *Let $f: A \rightarrow B$ be a injective **partial** function and $C \subseteq A$ then $(f^{-1})_{|f(C)} = (f_{|C})^{-1}$*

Proof. Let $(x, y) \in (f^{-1})_{|f(C)}$ then $x \in f(C)$ and $(x, y) \in f^{-1} \Rightarrow (y, x) \in f$, as $x \in f(C)$ there exists a $z \in C$ such that $(z, x) \in f$. As f is injective we have that $z = y$, proving that $y \in C$, which as $(y, x) \in f$ gives $(y, x) \in f_{|C}$ so that $(x, y) \in (f_{|C})^{-1}$. Hence

$$(f^{-1})_{|f(C)} \subseteq (f_{|C})^{-1} \tag{2.25}$$

If $(x, y) \in (f_{|C})^{-1}$ then $(y, x) \in f_{|C}$ so that $y \in C$ and $(y, x) \in f$. Hence $x \in f(C)$ and as $(y, x) \in f$ gives $(x, y) \in f^{-1}$ we have $(x, y) \in (f^{-1})_{|f(C)}$. This proves that $(f_{|C})^{-1} \subseteq (f^{-1})_{|f(C)}$, combing this with [eq: 2.25] gives:

$$(f^{-1})_{f(C)} = (f_{|C})^{-1}$$ □

Theorem 2.87. *Let $f: A \rightarrow B$ and $g: C \rightarrow D$ be **partial** functions and $E \subseteq A$ then*

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

Proof. Let $(x, z) \in (f \circ g)_{|E}$ then $(x, z) \in f \circ g$ and $x \in E$. Hence $\exists y$ such that $(x, y) \in f \wedge (y, z) \in g$, as $x \in E$ $(x, y) \in f_{|E}$. From $x \in E$ and $(x, y) \in f$ it follows also that $y \in f(E)$, hence as $(y, z) \in g$ we have that $(y, z) \in g_{|f(E)}$. From $(x, y) \in f_{|E}$ and $(y, z) \in g_{|f(E)}$ it follows that $(x, z) \in g_{|f(E)} \circ f_{|E}$ so that

$$(g \circ f)_{|E} \subseteq g_{|f(E)} \circ f_{|E} \tag{2.26}$$

If $(x, z) \in g_{|f(E)} \circ f_{|E}$ then there exists a y such that $(x, y) \in f_{|E}$ and $(y, z) \in g_{|f(E)}$, so $x \in E$, $(x, y) \in f$, $y \in f(E)$ and $(y, z) \in g$. Hence $x \in E$ and $(x, z) \in g \circ f$ proving that $(x, z) \in (g \circ f)_{|E}$. So $g_{|f(E)} \circ f_{|E} \subseteq (g \circ f)_{|E}$ which combined with [eq: 2.26] gives

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

Theorem 2.88. *Let $f: A \rightarrow B$ be a function and $C \subseteq A$ a sub-class of A then*

1. *$f_{|C}: C \rightarrow B$ is a function*
2. *$f_{|C}: C \rightarrow f(C)$ is a surjective function*

3. *If $f: A \rightarrow B$ is injective then $f_{|C}: C \rightarrow f(C)$ is a bijection*
4. *If $f: A \rightarrow B$ is a bijection then $f_{|C}: C \rightarrow f(C)$ is a bijection*

Proof.

1. Using [definition: 2.78] we have that $f_{|C}: C \rightarrow B$ is a partial function, as by [theorem: 2.84] $\mathrm{dom}(f_{|C}) = C \bigcap \mathrm{dom}(f) \underset{f \text{ is a function}}{=} C \bigcap A \underset{C \subseteq A}{=} C$, it follows that $f_{|C}: C \rightarrow B$ is a function.
2. Using [theorem: 2.37] we have that $f_{|C}: C \rightarrow f(C)$ is a function. Further if $y \in f(C)$ then there exist a $x \in C$ such that $(x, y) \in f$. As $x \in C$ we have that $(x, y) \in f \bigcap (C \times B) = f_{|C}$. Hence $f: C \rightarrow f(C)$ is a surjective function.
3. Let $(x, y), (x', y) \in f_{|C}$ then as $f_{|C} \subseteq f$ we have $(x, y), (x', y) \in f$ so that $x = x'$ proving that $f_{|C}: C \rightarrow f(C)$ is injective. As by (2) $f_{|C}: C \rightarrow f(C)$ is surjective it follows that
$$f_{|C}: C \rightarrow f(C) \text{ is a bijection}$$
4. As a bijection is injective thie follows from (3). □

Theorem 2.89. *Let $f: A \rightarrow B$ be a bijection then f $a \in A$ then $f_{|A \setminus \{a\}}: A \setminus \{a\} \rightarrow B \setminus \{b\}$ is a bijection [where $(a, b) \in f$] or in other words $f_{|A \setminus \{a\}}: A \setminus \{a\} \rightarrow B \setminus f(a)$ is a bijection.*

Proof. Let $(x, y) \in f_{|A \setminus \{a\}} = f \bigcap ((A \setminus \{a\}) \times B)$ then $(x, y) \in f$ and $x \neq a$. Assume that $y = b$ then $(x, b) \in f$ and as $(a, b) \in f$ we must have by injectivity of f that $x = a$ contradicting $x \neq a$, hence we have that $y \in B \setminus \{b\}$. So we have that $f_{|A \setminus \{a\}} \subseteq (A \setminus \{a\}) \times ((B \setminus \{b\}))$ and

$$f_{|A \setminus \{a\}}: A \setminus \{a\} \rightarrow B \setminus \{b\} \text{ is a function}$$

If $(x, y), (x', y) \in f_{|A \setminus \{a\}}$ then $(x, y), (x', y) \in f \underset{f \text{ is injective}}{\Rightarrow} x = x'$ proving that $f_{|A \setminus \{a\}}$ is injective. So

$$f_{|A \setminus \{a\}} \text{ is injective}$$

Finally if $y \in B \setminus \{b\}$ then as $f: A \rightarrow B$ is a bijection hence surjective, there exist a $x \in A$ such that $(x, y) \in f$. Assume that $x = a$ then $(a, y) \in f$ which as $(a, b) \in f$ gives $y = b$ contradicting $y \in B \setminus \{b\}$. Hence we must have $x \in A \setminus \{a\}$ so that $(x, y) \in f \bigcap ((A \setminus \{a\}) \times B) = f_{|A \setminus \{a\}}$ proving that

$$f: A \setminus \{a\} \rightarrow B \setminus \{b\} \text{ is surjective}$$ □

Theorem 2.90. *If $f: A \rightarrow B$ is a bijection and $C \subseteq A$ then*

1. $(f^{-1})_{|f(C)} = (f_{|C})^{-1}$
2. $(f^{-1})_{|f(C)}: f(C) \rightarrow C$ *is a bijection*

Proof.

1. If $(y, x) \in (f^{-1})_{|f(C)}$ then $y \in f(C)$ and $(y, x) \in f^{-1} \Rightarrow (x, y) \in f$. As $y \in f(C)$ there exist a $x' \in C$ such that $(x', y) \in f$. As f is injective it follow that $x = x'$ so that $x \in C$, hence $(x, y) \in f_{|C}$ or $(y, x) \in (f_{|C})^{-1}$. From this it follows that
$$(f^{-1})_{|f(C)} \subseteq (f_{|C})^{-1} \tag{2.27}$$
If $(y, x) \in (f_{|C})^{-1}$ then $(x, y) \in f_{|C}$ so that $x \in C$ and $(x, y) \in f$ proving that $y \in f(C)$. Further as $(x, y) \in f$ it follows that $(y, x) \in f^{-1}$. Hence $(y, x) \in (f^{-1})_{|f(C)}$ so that $(f_{|C})^{-1} \subseteq (f^{-1})_{|f(C)}$. Combining this result with [eq: 2.27] proves
$$(f^{-1})_{|f(C)} = (f_{|C})^{-1}$$
2. We have

   injectivity. Let $(y, x), (y', x) \in (f^{-1})_{|f(C)}$ then $(y, x), (y', x) \in f^{-1} \Rightarrow (x, y), (x, y') \in f$. Hence as f is a function we have that $y = y'$ which proves that $(f^{-1})_{|f(C)}$ is injective.

   surjectivity. Let $x \in C$ then as $C \subseteq A$ $x \in A$. As $f^{-1}: B \rightarrow A$ is a bijection by [theorem: 2.72] there exist a $y \in B$ such that $(y, x) \in f^{-1}$ or $(x, y) \in f$ so that $y \in f(C)$. Hence $(y, x) \in (f^{-1})_{|f(C)}$ proving surjectivity. □

The following theorem will be usefull for manifolds later

Theorem 2.91. *Let $f: A \rightarrow B$ and $g: C \rightarrow D$ be injections then we have*

1. *$f: A \rightarrow f(A)$ and $g: C \rightarrow f(C)$ are bijections*
2. $\mathrm{dom}(f \circ g^{-1}) = g(A \bigcap C)$
3. *$f \circ g^{-1}: g(A \bigcap C) \rightarrow f(A \bigcap C)$ is a bijection*
4. $f \circ g^{-1} = (f \circ g^{-1})_{|g(A \bigcap C)} = f_{|A \bigcap C} \circ (g^{-1})_{|g(A \bigcap C)} = f_{|(A \bigcap C)} \circ (g_{|A \bigcap C})^{-1}$

Proof.

1. This follows from [proposition: 2.66]
2. If $z \in \mathrm{dom}(f \circ g^{-1})$ then $\exists x$ such that $(z, x) \in f \circ g^{-1}$, hence $\exists y$ such that $(z, y) \in g^{-1}$ and $(y, z) \in f$, from which it follows that $(y, z) \in g$ and $(y, z) \in f$. As $g \subseteq C \times B$ and $f \subseteq A \times B$ it follows that $y \in A$ and $y \in C$ so that $y \in A \bigcap C$, as $(y, z) \in g$ we have $z \in g(A \bigcap C)$. This proves
$$\mathrm{dom}(g \circ f^{-1}) \subseteq g\left(A \bigcap C\right) \tag{2.28}$$
If $z \in g(A \bigcap C)$ then $\exists y \in A \bigcap C$ such that $(y, z) \in g$, hence $(z, y) \in g^{-1}$. As f is a function we have that $A = \mathrm{dom}(f)$, hence as $y \in A \bigcap C \Rightarrow y \in A$, there exists a x such that $(y, x) \in f$. As $(z, y) \in g^{-1}$ we have $(z, x) \in f \circ g^{-1}$ proving that $z \in \mathrm{dom}(f \circ g^{-1})$. Hence $g(A \bigcap C) \subseteq \mathrm{dom}(g \circ f^{-1})$ which combined with [eq: 2.28].
$$\mathrm{dom}(g \circ f^{-1}) = g\left(A \bigcap C\right)$$
3.
   injectivity. If $(x, y), (x', y) \in f \circ g^{-1}$ then $\exists z, z'$ such that $(x, z), (x', z') \in f$ and $(z, y), (z', y) \in g^{-1}$. Hence $(y, z), (y, z') \in g$ so that $z = z'$ [as g^{-1} is a function] hence $(x, z), (x', z) \in f$ giving $x = x'$.

   surjectivity. If $y \in f(A \bigcap C)$ then $\exists x \in A \bigcap C$ such that $(x, y) \in f$. As $A \bigcap C \subseteq C$ we have that $x \in C$, so as $g: C \rightarrow B$ is a function there exist a z such that $(x, z) \in g$, hence $(z, x) \in g^{-1}$. As $(x, y) \in f$ it follows that $(z, y) \in f \circ g^{-1}$.
4. We have
$$\begin{array}{rcl} (f \circ g^{-1}) & \underset{[\text{theorem: } 2.85]}{=} & (f \circ g^{-1})_{\mathrm{dom}(f \circ g^{-1})} \\ & \underset{(1)}{=} & (f \circ g^{-1})_{g(A \bigcap C)} \\ & \underset{[\text{theorem: } 2.87]}{=} & f_{|g^{-1}(g(A \bigcap C))} \circ (g^{-1})_{g(A \bigcap C)} \\ & \underset{g \text{ is injective and } [\text{theorem: } 2.55]}{=} & f_{|A \bigcap C} \circ (g^{-1})_{|g(A \bigcap C)} \\ & \underset{[\text{theorem: } 2.86]}{=} & f_{|A \bigcap C} \circ (g_{|A \bigcap C})^{-1} \end{array}$$

□

2.2.5 Set operations and (Partial) Functions

Theorem 2.92. *Let $f: A \rightarrow B$ be 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 $C, D \subseteq B$ then $f^{-1}(C \setminus D) = f^{-1}(C) \setminus f^{-1}(D)$*
4. *If $D \subseteq B$ then $f^{-1}(B \setminus D) = A \setminus f^{-1}(D)$*
5. *If $C, D \subseteq A$ then $f(C) \setminus f(D) \subseteq f(C \setminus D)$*
6. *If $C, D \subseteq A$ and f is **injective** then $f(C) \setminus f(D) = f(C \setminus D)$*

Proof.

1. Let $y \in f(C)$ then there exist a $x \in C$ such that $(x, y) \in f$, as $C \subseteq D$ we have $x \in D$ so that $y \in f(D)$
2. If $x \in f^{-1}(C)$ there exists a $y \in C$ such that $(x, y) \in f$, as $C \subseteq D$ then $y \in D$ so that $x \in f^{-1}(D)$

3. If $x \in f^{-1}(C \setminus D)$ then $\exists y \in C \setminus D$ such that $(x, y) \in f$. As $y \in C \setminus D$ we have that $y \in C$ and $y \notin D$, from $y \in C$ it follows that $x \in f^{-1}(C)$. Assume that also $x \in f^{-1}(D)$ then $\exists y' \in D$ such that $(x, y') \in f$ which, as f is a function and $(x, y) \in f$, proves that $y = y'$, hence $y \in D$ contradicting $y \notin D$, so we must have $x \notin f^{-1}(D)$, hence $x \in f(C) \setminus f(D)$ proving
$$f^{-1}(C \setminus D) \subseteq f^{-1}(C) \setminus f^{-1}(D) \tag{2.29}$$
If $x \in f^{-1}(C) \setminus f^{-1}(D)$ then $x \in f^{-1}(C)$ and $x \notin f^{-1}(D)$. As $x \in f^{-1}(C)$ there exists a $y \in C$ such that $(x, y) \in f$. Assume that $y \in D$, then as $(x, y) \in f$ we have $x \in f^{-1}(D)$ contradicting $x \notin f^{-1}(D)$, so we must have $y \notin D$. Hence $y \in C \setminus D$ which proves that $x \in f^{-1}(C \setminus D)$ or $f^{-1}(C) \setminus f^{-1}(D) \subseteq f^{-1}(C \setminus D)$. Combining this with [eq: 2.29] proves
$$f^{-1}(C \setminus D) = f^{-1}(C) \setminus f^{-1}(D)$$
4. As $D \subseteq B \subseteq B$ we have by (3) that
$$\begin{aligned} f^{-1}(B \setminus D) &= f^{-1}(B) \setminus f^{-1}(D) \\ &\underset{[\text{theorem: } 2.49]}{=} A \setminus f^{-1}(D) \end{aligned}$$
5. If $y \in f(C) \setminus f(D)$ then $y \in f(C)$ and $y \notin f(D)$. From $y \in f(C)$ it follows that $\exists x \in C$ such that $(x, y) \in f$. Assume that $x \in D$ then as $(x, y) \in f$ we have $y \in f(D)$ contradicting $y \notin f(D)$, so we must have $x \notin D$, proving that $x \in C \setminus D$. Hence $y \in f(C \setminus D)$ or
$$f(C) \setminus f(D) \subseteq f(C \setminus D)$$
6. If $y \in f(C \setminus D)$ then $\exists x \in C \setminus D$ such that $x \in C$, $x \notin D$ and $(x, y) \in f$. From $x \in C$ it follows that $y \in f(C)$. Assume that $y \in f(D)$ then there exist a $x' \in D$ such that $(x', y) \in f$, as f is **injective** we have $x = x'$ so that $x \in D$ contradicting $x \notin D$, hence $y \notin f(D)$. This proves that $y \in f(C) \setminus f(D)$ or $f(C \setminus D) \subseteq f(C) \setminus f(D)$ which combined with (3) gives
$$f(C) \setminus f(D) = f(C \setminus D)$$
□

Theorem 2.93. *If $f: A \rightarrow B$ is a function, $E, F \subseteq A$ and $C, D \subseteq B$ then we have*

1. $f(E \bigcup F) = f(E) \bigcup f(F)$
2. $f^{-1}(C \bigcup D) = f^{-1}(C) \bigcup f^{-1}()$
3. $f(E \bigcap F) \subseteq f(E) \bigcap f(F)$
4. *If f is injective then* $f(E \bigcap F) = f(E) \bigcap f(F)$
5. $f^{-1}(C \bigcap D) = f^{-1}(C) \bigcap f^{-1}(D)$

Proof.

1. Let $y \in f(E \bigcup F)$ then there exist a $x \in E \bigcup F$ with $(x, y) \in f$. So $x \in E$ proving that $y \in f(E)$ or $x \in F$ proving $y \in f(F)$. So it follows that $y \in f(E) \bigcup f(F)$ or
$$f\left(E \bigcup F\right) \subseteq f(E) \bigcup f(F) \tag{2.30}$$
If $y \in f(E) \bigcup f(F)$ then we have the following possibilities

   ***y* ∈ *f*(*E*).** Then $\exists x \in E$ such that $(x, y) \in f$. As by the definition of a union $x \in E \bigcup F$, it follows that $y \in f(E \bigcup F)$

   ***y* ∈ *f*(*F*).** Then $\exists x \in F$ such that $(x, y) \in f$. As by the definition of a union $x \in E \bigcup F$, it follows that $y \in f(E \bigcup F)$

   So in all cases we have $y \in f(E \bigcup F)$. Hence $f(E) \bigcup f(F) \subseteq f(E \bigcup F)$ which combined with [eq: 2.30] proves
$$f\left(E \bigcup F\right) = f(E) \bigcup f(F)$$
2. If $x \in f^{-1}(C \bigcup D)$ there exists a $y \in C \bigcup D$ such that $(x, y) \in f$. From $y \in C \bigcup D$ we have $y \in C$ hence $x \in f^{-1}(C)$ or $y \in D$ hence $x \in f^{-1}(D)$. So $x \in f^{-1}(C) \bigcup f^{-1}(D)$ proving
$$f^{-1}\left(C \bigcup D\right) \subseteq f^{-1}(C) \bigcup f^{-1}(D) \tag{2.31}$$
If $x \in f^{-1}(C) \bigcup f^{-1}(D)$ then we have the following possibilities to consider:

   $x \in f^{-1}(C)$. Then $\exists y \in C$ such that $(x, y) \in f$. As by the definition of a union $y \in C \bigcup D$ it follows that $x \in f^{-1}(C \bigcup D)$

   $x \in f^{-1}(D)$. Then $\exists y \in D$ such that $(x, y) \in f$. As by the definition of a union $y \in C \bigcup D$ it follows that $x \in f^{-1}(C \bigcup D)$

So in all cases we have $x \in f^{-1}(C \bigcup D)$, proving $f^{-1}(C) \bigcup f^{-1}(D) \subseteq f^{-1}(C \bigcup D)$ which combined with [eq 2.31] proves

$$f^{-1}\left(C \bigcup D\right) = f^{-1}(C) \bigcup f^{-1}(D)$$

3. If $y \in f(E \bigcap F)$ then $\exists x \in E \bigcap F$ such that $(x, y) \in f$. From $x \in E \bigcap F$ we have that $x \in E$ hence $y \in f(E)$ and $x \in F$, so that $y \in f(F)$. Hence $y \in f(E) \bigcap f(F)$ or

$$f\left(E \bigcap F\right) \subseteq f(E) \bigcap f(F)$$

4. Using (3) we have that

$$f\left(E \bigcap F\right) \subseteq f(E) \bigcap f(F) \tag{2.32}$$

Let $y \in f(E) \bigcap f(F)$ then we have $y \in f(E)$ so that $\exists x \in E$ such that $(x, y) \in f$ and $y \in f(F)$ so that $\exists x' \in F$ such that $(x', y) \in f$. As f is injective and $(x, y), (x', y) \in f$ we have $x = x'$ so that $x \in E \bigcap F$, proving that $f(E) \bigcap f(F) \subseteq f(E \bigcap F)$. Combining this result with [eq: 2.32] gives

$$f\left(E \bigcap F\right) = f(E) \bigcap f(F)$$

5. If $x \in f^{-1}(C \bigcap D)$ then $\exists y \in C \bigcap D$ such that $y \in C$, so that $x \in f^{-1}(C)$ and $y \in D$, so that $x \in f^{-1}(D)$. Hence $x \in f^{-1}(C) \bigcap f^{-1}(D)$ proving

$$f^{-1}\left(C \bigcap D\right) \subseteq f^{-1}(C) \bigcap f^{-1}(D) \tag{2.33}$$

If $x \in f^{-1}(C) \bigcap f^{-1}(D)$ then $x \in f^{-1}(C)$ so there exists a $y \in C$ such that $(x, y) \in f$ and $x \in f^{-1}(D)$ so $\exists y' \in D$ such that $(x, y') \in f$. As f is a function $y = y'$ proving $y \in C \bigcap D$, hence $x \in f^{-1}(C \bigcap D)$. So $f^{-1}(C) \bigcap f^{-1}(D) \subseteq f^{-1}(C \bigcap D)$, combining this with [eq: 2.33] gives

$$f^{-1}\left(C \bigcap D\right) = f^{-1}(C) \bigcap f^{-1}(D)$$ □

Up to now we define a function $f: A \rightarrow B$ by specifying what the classes f, A, B are. However in many cases we have a parameterized expression [based on function calls and operators) to define f. Then we have the following

Proposition 2.94. *Let A, B be classes and suppose that there exists a parameterized expression $F(x)$ that calculates a* ***unique*** *value for* ***every*** *$x \in A$ then we can define the function $f: A \rightarrow B$ by $f = \{z | z = (x, F(x)) \wedge x \in A\}$*

Proof. If $(x, y), (x, y') \in f$ then there exists $a, a' \in A$ such that $(x, y) = (a, F(a)) \wedge (x, y') \in (a', F(a'))$, hence $x = a \wedge x = a' \wedge y = f(a) \wedge y' = F(a') \Rightarrow a = a' \wedge y' = F(a) \wedge y = F(a)$ proving that $y = y'$. So

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

If $x \in A$ then as $F(x)$ is defined on every $x \in A$ we have that $(x, F(x)) \in f$ so that $x \in \mathrm{dom}(f)$. So $A \subseteq \mathrm{dom}(f)$ we have by 2.26 that

$$f: A \rightarrow B \text{ is a function}$$ □

This leads to a notation that we will gradually start to use

Notation 2.95. *The function definition $f: A \rightarrow B$ defined by $f(x) = F(x)$ [where $E(x)$ is a parameterized expression that calculates a unique value for every $x \in A$] is equivalent with*

$$f = \{z | z = (x, E x) \wedge x \in E\} = \{(x, E(x)) | x \in X\}$$

Example 2.96. $f: \mathbb{R} \rightarrow \mathbb{R}$ is defined by $f(x) = \cos(4 \cdot x)$

2.2.6 Indexed sets

In many cases we have to deal with sets indexed by a index, which is actually a function in another form. We will use this in toplogy and vector spaces.

Definition 2.97. (indexed set) *Let $f: I \rightarrow A$ be a surjection then A is called a* ***indexed set*** *and noted as*

$$A = \{f(i) | i \in I\} \text{ or } A = \{f_i | i \in I\}$$

So

$$x \in A \Leftrightarrow \exists i \in I \text{ such that } x = f(i) \text{ or } x = f_i$$

I is called the index of the indexed set $\{f_i | i \in I\}$.

Definition 2.98. (unique indexed set) *$A = \{f_i | i \in I\}$ is a **unique indexed set** if $f: I \to A$ is a bijection. So*

$$x \in A \Leftrightarrow \exists i \in I \text{ such that } x = f(i) \text{ or } x = f_i$$

and

$$\text{If } x_i = x_j \text{ then } i = j$$

Example 2.99. Every set can be writen as a unique indexed set indexed by itself. So if A is a set then $A = \{\mathrm{Id}_X(i) | i \in I\}$.

2.3 Families

2.3.1 Family

We introduce now the idea of a indexed family which is essential a function of a class to another class. It is essential another notation for a function where the emphasis is on the objects in a collection and a way of indexing these objects and less on the function itself

Definition 2.100. *A Let I, B be classes then a family*

$$\boldsymbol{\{x_i\}_{i \in I} \subseteq B}$$

is actually a function

$$\boldsymbol{f: I \to B}$$

Further $\boldsymbol{x_i}$ is another notation for $\boldsymbol{f(i)}$ so that $\boldsymbol{y = f_i}$ is equivalent with $\boldsymbol{y = f(i)}$ or $\boldsymbol{(i, y) \in f}$

Note 2.101. In the above definition $\{x_i\}_{i \in I}$ only make sense if you specify what the defining function is. To avoid excessive notation, we assume that if we write $\boldsymbol{\{x\}_{i \in I} \subseteq B}$ that the defining function is $\boldsymbol{x: I \to B}$. However this is sometimes not feasible and in that case we state what the defining function of $\{x_i\}_{i \in I}$ is.

Note 2.102. In most cases we use the notation $\{x_i\}_{i \in I} \subseteq B$ to indicate three parts of a function namely $x: I \to B$ where x is the graph of a function, I is the domain and B is the target of the function. In some cases we are not interested in the target of the function [we can always assume then that the target is the universal class $\mathcal{U}$] in this case we just use the notation $\{x_i\}_{i \in I}$ for the family.

Example 2.103. The empty function $\varnothing: \varnothing \to V$ [see example: 2.44] defines a family that is noted as $\{\varnothing_i\}_{i \in \varnothing} \subseteq V$. Further if $\{x_i\}_{i \in \varnothing} \subseteq V$ is a family where the index set is empty then $\{x_i\}_{i \in \varnothing} \subseteq V = \{\varnothing_i\}_{i \in \varnothing} \subseteq V$

Proof. $\{x_i\}_{i \in \varnothing} \subseteq V$ is defined by the function $x: \varnothing \to V$, as $x \subseteq \varnothing \times V = \varnothing$ we have that $x = \varnothing$ so that $\{x_i\}_{i \in \varnothing} \subseteq V = \{\varnothing_i\}_{i \in \varnothing} \subseteq V$. □

Proposition 2.104. *For the family $\{x_i\}_{i \in I} \subseteq \varnothing$ we have $I = \varnothing$ so that $\{x_i\}_{i \in I} \subseteq \varnothing = \{\varnothing_i\}_{i \in \varnothing} \subseteq \varnothing$*

Proof. Let $f: I \to \varnothing$ be the function that defines the family then as f is a function we have that $f(I) = \varnothing$. So if $x \in I$ then $\exists y \in \varnothing$ such that $(x, y) \in f \subseteq I \times \varnothing = \varnothing$ a contradiction, hence we must have $I = \varnothing$. □

Example 2.105. Let A, B be classes then family $\{(\mathrm{Id}_A)_a\}_{a \in A} \subseteq B$ defined by the function $\mathrm{Id}_A: A \to B$ is noted as $\{x\}_{x \in A}$.

We can now define the concept of a sub family

Definition 2.106. *Let $\{A_i\}_{i \in I} \subseteq B$ be a family of objects in B defined by the function $f: I \to B$ and $J \subseteq I$ then $\{A_i\}_{i \in J} \subseteq B$ is the family defined by the function $f_{|J}: J \to B$ [see: theorem: 2.88 for the proof that $f_J: I \to B$ is a function]*

Definition 2.107. *Let I, J, A, B be classes such that $I \bigcap J = \varnothing$ and*

$$\{x\}_{i \in I} \subseteq A \text{ defined by the function } f: I \to A$$

$$\{y_i\}_{i \in J} \subseteq B \text{ defined by the function } g: J \to B$$

then $\{z_i\}_{i\in I\bigcup J}\subseteq A\bigcup B$ *defined by* $z_i=\left\{\begin{smallmatrix} A_i\text{if } i\in I \\ B_i \textit{ if } i\in J\end{smallmatrix}\right.$ *is the family defined by the function*

$$f\bigcup g: I\bigcup J\rightarrow A\bigcup B$$

[see theorem: 2.80 for the proof that $f\bigcup g: I\bigcup J\rightarrow A\bigcup B$ *is indeed a function]*

Definition 2.108. *If* I, J *are classes then* $\{x_{i,j}\}_{(i,j)\in I\times J}\subseteq A$ *is defined by a function* $x: I\times J\rightarrow A$*, based on this we can define the following families:*

1. $\forall i\in I$ $\{x_{i,j}\}_{j\in J}$ *is defined by the function* $x_{i,*}: J\rightarrow A$ *where* $x_{i,*}(j)=x(i,j)=x_{i,j}$
2. $\forall j\in J$ $\{x_{i,j}\}_{i\in I}$ *is defined by the function* $x_{*,j}: I\rightarrow A$ *where* $x_{*,j}(i)=x(i,j)=x_{i,j}$

Composition of functions can also also be represented via the above family notation,

Definition 2.109. *If you have a function* $\boldsymbol{f: I\rightarrow J}$ *and a family* $\boldsymbol{\{x_j\}_{j\in J}\subseteq A}$ *[defined by the function* $\boldsymbol{g: J\rightarrow A}$*] then*

$$\boldsymbol{\{x_{f(i)}\}_{i\in I}}$$

is the family represented by the function

$$\boldsymbol{g\circ f: I\rightarrow A}$$

So a family is just another notation for a function. We introduce also a new notation for the range of this function.

Definition 2.110. *If* $\{x_i\}_{i\in I}$ *is a family of objects in* B *[standing for the function* $f: I\rightarrow B$*] then we define* $\{x_i|i\in I\}$ *by*

$$\{x_i|i\in I\}=\text{range}(f)=f(I)$$

The motivation for this definition is the following theorem

Theorem 2.111. *If* $\{x_i\}_{i\in I}\subseteq B$ *is a family of objects in* B *with associated function* f *then*

$$x\in\{x_i|i\in I\}\Leftrightarrow\exists i\in I \text{ such that } x=x_i$$

Proof. As $\{x_i\}_{i\in I}\subseteq B$ is equivalent with $f: I\rightarrow B$ we have

$$\begin{array}{rcl} z\in\{x_i|i\in I\} & \underset{\text{define}}{\Leftrightarrow} & z\in\text{range}(x) \\ & \Leftrightarrow & \exists i \text{ with } (i,z)\in f \\ & \underset{f\subseteq I\times B}{\Leftrightarrow} & \exists i \text{ with } i\in I\wedge(i,z)\in f \\ & \Leftrightarrow & \exists i\in I \text{ with } (i,z)\in f \\ & \Leftrightarrow & \exists i\in I \text{ with } z=f(i) \\ & \Leftrightarrow & \exists i\in I \text{ with z=x\_i} \\ & & \square \end{array}$$

Theorem 2.112. *If* $\{x_i\}_{i\in I}\subseteq B$ *is a family such that* I *and* B *are sets then* $\{x_i|i\in I\}$ *is a set*

Proof. $\{x_i\}_{i\in I}\subseteq B$ is actually the function $x: I\rightarrow B$ where $\text{range}(x)=\{x_i|i\in I\}$. As $I$and B are sets, it follows from [theorem: 2.12] that range(x) is a set, hence $\{x_i|i\in I\}$ is a set. □

Up to now we consider a family as a indexed collection of objects. What is actually a object, in set theory it is a class which can be either a set or a proper class. A class is a collection so we can talk about the union of these collection. The convention is then to use upper case instead of lower case. If we want to deal with the union and intersection of the objects [considered as collections] in the family we use also a different notation.

Notation 2.113. *If* $\{A_i\}_{i\in I}\subseteq B$ *is a family of objects in* B *[standing for the function* $A: I\rightarrow B$*] then* $\bigcup_{i\in I}A_i$ *is defined by*

$$\bigcup_{i\in I}A_i=\bigcup\{\text{range}(A)\} \textit{ [definition: 1.56]}$$

Definition 2.114. *A family $\{A_i\}_{i\in I}\subseteq B$ is **pairwise disjoint** iff $\forall i,j\in I$ with $i\neq j$ we have $A_i\bigcap A_j=\varnothing$.*

Notation 2.115. *If $\{A_i\}_{i\in I}\subseteq B$ is pairwise disjoint and we want to indicate this fact when we write the union of the family then we use the notation $\bigsqcup_{i\in I}A_i$. So $\bigsqcup_{i\in I}A_i$ is actually the same as $\bigcup_{i\in I}A_i$, but also relating the information that $\{A_i\}_{i\in I}$ is pairwise disjoint.*

Using this new notation we have the following characterization of the union

Theorem 2.116. *If $\{A_i\}_{i\in I}\subseteq B$ is a family of objects in B then*

$$x\in\bigcup_{i\in I}A_i\Leftrightarrow\exists i\in I \text{ such that } x\in A_i$$

Proof. As $\{A_i\}_{i\in I}\subseteq B$ is actually the function $A: I\rightarrow B$ where $\bigcup_{i\in I}A_i=\bigcup\operatorname{range}(A)$. Then we have

$$\begin{aligned}
x\in\bigcup_{i\in I}A_i &\underset{\text{definition}}{\Leftrightarrow} x\in\bigcup\operatorname{range}(A)\\
&\underset{[\text{definition: } 1.56]}{\Leftrightarrow} \exists y\in\operatorname{range}(A) \text{ such that } x\in y\\
&\Leftrightarrow \exists i \text{ such that } (i,y)\in A \text{ and } x\in y\\
&\underset{A\subseteq I\times B}{\Leftrightarrow} \exists i\in I \text{ such that } (x,y)\in A \text{ and } x\in y\\
&\Leftrightarrow \exists i\in I \text{ such that } y=A_i \text{ and } x\in y\\
&\Leftrightarrow \exists i\in I \text{ such that } x\in A_i
\end{aligned}$$

□

Corollary 2.117. *If $\{A_j\}_{j\in J}\subseteq B$ is a family and $f: I\rightarrow J$ is a surjection then*

$$\bigcup_{j\in J}A_j=\bigcup_{i\in I}A_{f(i)}$$

Proof. If $x\in\bigcup_{j\in J}A_j$ then by [theorem: 2.116] there exist a $j\in J$ such that $x\in A_j=A(j)$. As f is surjective we have by [theorem: 2.52] that there exist a $i\in I$ such that $j=f(i)$. Hence $x\in A(f(i))=(A\circ f)(i)$. So by [theorem: 2.116] and the definition of $\bigcup_{i\in I}A_{f(i)}$ we have $x\in\bigcup_{i\in I}A_{f(i)}$. Hence

$$\bigcup_{j\in J}A_j\subseteq\bigcup_{i\in I}A_{f(i)} \tag{2.34}$$

If $x\in\bigcup_{i\in I}A_{f(i)}$ then there exist a $i\in I$ such that $x\in(A\circ f)(i)$, which, as using [theorem: 2.22] $(A\circ f)(i)\in\operatorname{range}(A)$, means that there exists a $j\in J$ such that $A_j=(A\circ f)(i)$. Hence $x\in A_j$ proving by [theorem: 2.116] that $x\in\bigcup_{j\in J}A_j$. So $\bigcup_{i\in I}A_{f(i)}\subseteq\bigcup_{j\in J}A_j$ which combined with [eq: 2.34] gives

$$\bigcup_{j\in J}A_j=\bigcup_{i\in I}A_{f(i)}$$

□

Theorem 2.118. *If $\{A_i\}_{i\in I}\subseteq B$ is a family of objects in B where I and B are sets then $\bigcup_{i\in I}A_i$ is a set.*

Proof. As $\{A_i\}_{i\in I}\subseteq B$ is another way of saying $A: I\rightarrow B$ and I and B are sets, it follows from [theorem: 2.12] that $\operatorname{range}(A)$ is a set. Using the Axiom of Union [axiom: 1.61] $\bigcup\operatorname{range}(A)$ is a set, so by definition $\bigcup_{i\in I}A_i$. is a set. □

Example 2.119. Let $\{A_i\}_{i\in\varnothing}\subseteq B$ be the family defined by $A=\varnothing$ [the empty function $\varnothing:\varnothing\rightarrow B$ [see example: 2.44[] then $\bigcup_{i\in\varnothing}A_i=\varnothing$

Proof. Let $y\in\operatorname{range}(A)=\operatorname{range}(\varnothing)$ then x such that $(x,y)\in\varnothing$, a contradiction. Hence $\operatorname{range}(A)=\varnothing$. So

$$\bigcup_{i\in\varnothing}A=\bigcup\operatorname{range}(A)=\bigcup\varnothing\underset{1.58}{=}\varnothing$$

□

Definition 2.120. *If $\{A_i\}_{i\in I}\subseteq B$ is a family of objects in B [standing for the function $A: I\rightarrow B$] then $\bigcap_{i\in I}A_i$ is defined by*

$$\bigcap_{i\in I}A_i=\bigcap\operatorname{range}(A) \text{ [definition: 1.57]}$$

Theorem 2.121. *If* $\{A_i\}_{i\in I}\subseteq B$ *then* $x\in\bigcap_{i\in I}A_i\Leftrightarrow\forall i\in I$ *we have* $x\in A_i$

Proof. $\{A_i\}_{i\in I}\subseteq B$ is actually the function $A: I\rightarrow B$ where $\bigcap_{i\in I}A_i=\bigcap\text{range}(A)$.

$$\begin{array}{rcl} x\in\bigcap_{i\in I}A_i & \underset{\text{definition}}{\Leftrightarrow} & x\in\bigcap\text{range}(A)\\ & \underset{[\text{definition: }1.57]}{\Leftrightarrow} & \forall y\in\text{range}(A)\text{ we have }x\in y\\ & \underset{y\in\text{range}(A)\Leftrightarrow\exists i\text{ with }(i,y)\in A}{\Leftrightarrow} & \forall i\in I\text{ with }(i,y)\in A\text{ we have }x\in y\\ & \Leftrightarrow & \forall i\in I\text{ with }y=A_i\text{ we have }x\in y\\ & \Leftrightarrow & \forall i\in I\text{ we have }x\in A_i \end{array}$$

□

Theorem 2.122. *If* $\{A_i\}_{i\in I}\subseteq B$ *is a family of objects in* B *such that* $I\neq\varnothing$ *then* $\bigcap_{i\in I}A_i$ *is a set.*

Proof. $\{A_i\}_{i\in I}\subseteq B$ is actually the function $A: I\rightarrow B$ where $\bigcap_{i\in I}A_i=\bigcap\text{range}(A)$. As $I\neq\varnothing$ there exists a $i\in I$. Given that A is a function it follows that $\text{dom}(A)=I$, so there exists a y such that $(i,y)\in A$ or $y\in\text{range}(A)$. So $\varnothing\neq\text{range}(A)$ which by [theorem: 1.57] proves that $\bigcap\text{range}(A)$ is a set, hence $\bigcap_{i\in I}A_i$ is a set. □

Example 2.123. Let $I=\{0\}$, B a class and take $A: I\rightarrow\{B\}$ defined by $A=\{(0,B)\}$, defining the family $\{A_i\}_{i\in\{0\}}\subseteq\{B\}$ where $A_0=B$. For this family we have $\bigcap_{i\in\{0\}}A_i=B$ and $\bigcup_{i\in\{0\}}A_i=B$

Proof. Using [example: 2.65] it follows that $A: I\rightarrow\{B\}$ is bijection, hence a function, so that $\{A_i\}_{i\in\{0\}}\subseteq\{B\}$ is a well defined family. Further as A is a bijection we have that

$$\text{range}(A)=\{B\}$$

Finally

$$\bigcup_{i\in\{0\}}A_i=\bigcup\text{range}(A)=\bigcup\{B\}\underset{[\text{example: }1.58]}{=}A$$

and

$$\bigcap_{i\in\{0\}}A_i=\bigcap\text{range}(A)=\bigcap\{B\}\underset{[\text{example: }1.58]}{=}A$$

□

Example 2.124. Let C,D classes, $I=\{0,1\}$ and take $A: I\rightarrow\{C,D\}$ defined by $A=\{(0,C),(1,D)\}$ [see example: 2.27], defining the family $\{A_i\}_{i\in\{0,1\}}\subseteq\{C,D\}$ where $A_0=C$ and $A_1=D$. For this family we have $\bigcup_{i\in\{0,1\}}A_i=C\bigcup D$ and $\bigcap_{i\in\{0,1\}}A_i=C\bigcap D$.

Proof. If $y\in\text{range}(A)$ then $\exists x$ such that $(x,y)\in A=\{(0,C),(1,D)\}$, so that $(x,y)=(0,C)\Rightarrow y=C$ or $(x,y)=(1,D)\Rightarrow y=D$, proving that $x\in\{C,D\}$. Further if $y\in\{C,D\}$ then $y=C\Rightarrow(0,C)\in A\Rightarrow y\in\text{range}(A)$ or $y=D\Rightarrow(1,D)\in A\Rightarrow y\in\text{range}(A)$. So we have

$$\text{range}(A)=\{C,D\}$$

Finally

$$\bigcup_{i\in\{0,1\}}A_i=\bigcup\text{range}(A)=\bigcup\{C,D\}\underset{[\text{example: }1.59]]}{=}C\bigcup D$$

and

$$\bigcap_{i\in\{0,1\}}A_i=\bigcap\text{range}(A)=\bigcap\{C,D\}\underset{[\text{example: }1.59]]}{=}C\bigcap D$$

□

2.3.2 Properties of the union and intersection of families

To save space, from now on we use [theorem: 2.116] and [theorem: 2.121] about union and intersection of families without explicit referring to these theorems.

Theorem 2.125. *If* $\{A_i\}_{i\in I}\subseteq B$ *is a family then we have:*

1. $\forall i\in I$ *we have* $A_i\subseteq\bigcup_{i\in I}A_i$

2. $\forall i \in I$ *we have* $\bigcap_{i \in I} A_i \subseteq A_i$
3. *If* $\forall i \in I$ *we have that* $A_i \subseteq C$ *then* $\bigcup_{i \in I} A_i \subseteq C$
4. *If* $\forall i \in I$ *we have* $C \subseteq A_i$ *then* $C \subseteq \bigcap_{i \in I} A_i$

Proof.

1. Let $i \in I$ and assume that $x \in A_i$ then $\exists i \in I$ such that $x \in A_i$, so $x \in \bigcup_{i \in I} A_i$, proving that $A_i \subseteq \bigcup_{i \in I} A_i$.
2. Let $i \in I$ then if $x \in \bigcap_{i \in I} A_i$ we have $\forall j \in I$ that $x \in A_j \underset{i \in I}{\Rightarrow} x \in A_i$, proving that $\bigcap_{i \in I} A_i \subseteq A_i$
3.

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

4.

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

□

Theorem 2.126. *If* $\{A_i\}_{i \in I} \subseteq B$ *is a family then*

1. *If* $J \subseteq I$ *then*
   a. $\bigcup_{i \in J} A_i \subseteq \bigcup_{i \in I} A_i$
   b. $\bigcap_{i \in I} A_i \subseteq \bigcap_{i \in J} A_i$
2. *If* $I = J \bigcup K$ *then*
   a. $\bigcup_{i \in I} A_i = (\bigcup_{i \in J} A_i) \bigcup (\bigcup_{i \in K} A_i)$
   b. $\bigcap_{i \in I} A_i = (\bigcap_{i \in J} A_i) \bigcap (\bigcap_{i \in K} A_i)$

Proof.

1.
   a. If $x \in \bigcup_{i \in J} A_i$ then $\exists i \in J$ such that $x \in A_i$, as $J \subseteq I$ we have $i \in I$ with $x \in A_i$, so that $x \in \bigcup_{i \in I} A_i$.
   b. If $x \in \bigcap_{i \in I} A_i$ then $\forall i \in I$ we have $x \in A_i$, as $J \subseteq I$ we have also $\forall i \in J$ that $x \in A_i$, hence $x \in \bigcap_{i \in J} A_i$.
2.
   a. As by [theorem: 1.25] $J, K \subseteq I$ we have using (1) that $\bigcup_{i \in J} A_i \subseteq \bigcup_{i \in I} A_i$ and $\bigcup_{i \in K} A_i \subseteq \bigcup_{i \in I} A_i$. Using [theorem: 1.25 it follows that

$$\left(\bigcup_{i \in J} A_i\right) \bigcup \left(\bigcup_{i \in K} A_i\right) \subseteq \bigcup_{i \in I} A_i \tag{2.35}$$

If $x \in \bigcup_{i \in I} A$ then $\exists i \in I$ such that $x \in A_i$, as $I = J \bigcup K$ we have $i \in J \Rightarrow x \in \bigcup_{i \in J} A_i$ or $i \in K \Rightarrow x \in \bigcup_{i \in K} A_i$, which proves that $x \in (\bigcup_{i \in J} A_i) \bigcup (\bigcup_{i \in K} A_i)$. Hence

$$\bigcup_{i \in I} A_i \subseteq \left(\bigcup_{i \in J} A_i\right) \bigcup \left(\bigcup_{i \in K} A_i\right)$$

which combined with [eq: 2.35] proves

$$\bigcup_{i \in I} A_i = \left(\bigcup_{i \in J} A_i\right) \bigcup \left(\bigcup_{i \in K} A_i\right)$$

b. As by [theorem: 1.25] $J, K \subseteq I$ we have using (1) that $\bigcap_{i\in I} A_i \subseteq \bigcap_{i\in J} A_i$ and $\bigcap_{i\in I} A_i \subseteq \bigcap_{i\in K} A_i$. Using [theorem: 1.25] it follows that

$$\bigcap_{i\in I} A_i \subseteq \left(\bigcap_{i\in J} A_i\right) \bigcap \left(\bigcap_{i\in K} A_i\right) \tag{2.36}$$

If $x \in (\bigcap_{i\in J} A) \bigcap (\bigcap_{i\in K} A)$ then $x \in \bigcap_{i\in J} A_i$ and $x \in \bigcap_{i\in K} A_i$. So $\forall i \in J$ we have $x \in A_i$ and $\forall i \in K$ we have $x \in A_i$. Hence as $\forall i \in I$ we have $i \in J \Rightarrow x \in A_i$ or $i \in K \Rightarrow x \in A_i$ it follows that $x \in \bigcap_{i\in I} A_i$. So $(\bigcap_{i\in J} A) \bigcap (\bigcap_{i\in K} A) \subseteq \bigcap_{i\in I} A_i$ which combined with [eq: 2.36] proves

$$\bigcap_{i\in I} A_i = \left(\bigcap_{i\in J} A_i\right) \bigcap \left(\bigcap_{i\in K} A_i\right)$$

□

Theorem 2.127. *Let $\{A_i\}_{i\in I} \subseteq C$ and $\{B_i\}_{i\in I} \subseteq D$ be two families such that $\forall i \in I$ we have $A_i \subseteq B_i$ then*

1. $\bigcup_{i\in I} A_i \subseteq \bigcup_{i\in I} B_i$
2. $\bigcap_{i\in I} A_i \subseteq \bigcap_{i\in I} B_i$

Proof.

1. If $x \in \bigcup_{i\in I} A_i$ there exist a $i \in I$ such that $x \in A_i \underset{A_i \subseteq B_i}{\Rightarrow} x \in B_i$, hence $x \in \bigcup_{i\in I} B_i$
2. If $x \in \bigcap_{i\in I} A_i$ then $\forall i \in I$ we have $x \in A_i \underset{A_i \subseteq B_i}{\Rightarrow} x \in B_i$ proving $x \in \bigcap_{i\in I} B_i$ □

We have also the distributive laws for union and intersection [theorem: 1.30]

Theorem 2.128. (Distributivity) *Let $\{A_i\}_{i\in I} \subseteq B$ be a family and C a class then*

1. $C \bigcap (\bigcup_{i\in I} A_i) = \bigcup_{i\in I} (C \bigcap A_i)$
2. $C \bigcup (\bigcap_{i\in I} A_i) = \bigcap_{i\in I} (C \bigcup A_i)$
3. $C \bigcap (\bigcap_{i\in I} A_i) = \bigcap_{i\in I} (C \bigcap A_i)$
4. $C \bigcup (\bigcup_{i\in I} A_i) = \bigcup_{i\in I} (C \bigcup A_i)$

Proof.

1. If $x \in C \bigcap (\bigcup_{i\in I} A_i)$ then $x \in C$ and $x \in \bigcup_{i\in I} A_i \Rightarrow \exists i \in I$ such that $x \in A_i$. Hence $x \in C \bigcap A_i$, proving by [theorem: 2.125] that $x \in \bigcup_{i\in I} A_i$. So

$$C \bigcap \left(\bigcup_{i\in I} A_i\right) \subseteq \bigcup_{i\in I} (C \bigcap A_i) \tag{2.37}$$

If $x \in \bigcup_{i\in} (C \bigcap A_i)$ then there exist a $i \in I$ such that $x \in C$ and $x \in A_i \Rightarrow x \in \bigcup_{i\in I} A_i$, so $x \in C \bigcap (\bigcup_{i\in I} A_i)$, proving that $\bigcup_{i\in I} (C \bigcap A_i) \subseteq C \bigcap (\bigcup_{i\in I} A_i)$. Combining this with [eq: 2.37] proves

$$C \bigcap \left(\bigcup_{i\in I} A_i\right) = \bigcup_{i\in I} (C \bigcap A_i)$$

2. If $x \in C \bigcup (\bigcap_{i\in I} A_i)$ then we have the following cases to consider:

   ***$x \in C$*.** then $\forall i \in I$ we have $x \in C \bigcup A_i$ hence $x \in \bigcap_{i\in I} (C \bigcup A_i)$

   ***$x \in \bigcap_{i\in I} A_i$*.** then $\forall i \in I$ we have $x \in A_i$ hence $x \in \bigcap_{i\in I} (C \bigcup A_i)$

   which proves that

$$C \bigcup \left(\bigcap_{i\in I} A_i\right) \subseteq \bigcap_{i\in I} (C \bigcup A_i) \tag{2.38}$$

   If $x \in \bigcap_{i\in I} (C \bigcup A_i)$ then we have two cases to consider:

   ***$x \in C$*.** then $x \in C \bigcup (\bigcap_{i\in I} A_i)$

$x \notin C$. then, as $\forall i \in I$ we have $x \in C \bigcup A_i \underset{x \notin C}{\Rightarrow} x \in A_i$, it follows that $x \in \bigcap_{i \in I} A_i$ hence $x \in C \bigcup (\bigcap_{i \in I} A_i)$

In all cases we have $x \in C \bigcup (\bigcap_{i \in I} A_i)$ proving that $\bigcap_{i \in I} (C \bigcup A_i) \subseteq C \bigcup (\bigcap_{i \in I} A_i)$, combining this with [eq: 2.38] gives

$$C \bigcup \left(\bigcap_{i \in I} A_i \right) \subseteq \bigcap_{i \in I} (C \bigcup A_i)$$

3. We have

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

Proving

$$C \bigcap \left(\bigcap_{i \in I} A_i \right) = \bigcap_{i \in I} (C \bigcap A_i)$$

4. We have

$$\begin{aligned} x \in C \bigcup \left(\bigcup_{i \in I} A_i \right) &\Leftrightarrow x \in C \vee x \in \bigcup_{i \in I} A_i \\ &\Leftrightarrow x \in C \vee \exists i \in I \text{ with } x \in A_i \\ &\Leftrightarrow \exists i \in I \text{ with } (x \in C \vee x \in A_i) \\ &\Leftrightarrow \exists i \in I \text{ we have } x \in C \bigcup A_i \end{aligned}$$

proving that

$$C \bigcup \left(\bigcup_{i \in I} A_i \right) = \bigcup_{i \in I} (C \bigcup A_i)$$

□

Theorem 2.129. *Let $\{A_i\}_{i \in I} \subseteq C$ and $\{B_i\}_{i \in I} \subseteq D$ be two families then*

1. $(\bigcup_{i \in I} A_i) \bigcup (\bigcup_{i \in I} B_i) = \bigcup_{i \in I} (A_i \bigcup B_i)$
2. $\bigcup_{i \in I} (A_i \bigcap B_i) \subseteq (\bigcup_{i \in I} A_i) \bigcap (\bigcup_{i \in I} B_i)$

Proof.

1. First as $\forall i \in I$ we have by [theorem: 1.25] that $A_i \subseteq A_i \bigcup B_i$ and $B_i \subseteq A_i \bigcup B_i$ so it follows using [theorem: 2.127] that $\bigcup_{i \in I} A_i \subseteq \bigcup_{i \in I} (A_i \bigcup B_i)$ and $\bigcup_{i \in I} B_i \subseteq \bigcup_{i \in I} (A_i \bigcup B_i)$. Applying then [theorem: 1.25] gives

$$\left(\bigcup_{i \in I} A_i \right) \bigcup \left(\bigcup_{i \in I} B_i \right) \subseteq \bigcup_{i \in I} (A_i \bigcup B_i) \tag{2.39}$$

If now $x \in \bigcup_{i \in I} A_i \bigcup B_i$ then $\exists i \in I$ such that $x \in A_i \bigcup B_i$, then we have $x \in A_i \Rightarrow x \in \bigcup_{i \in I} A_i$ or $x \in B_i \Rightarrow x \in \bigcup_{i \in I} B_i$. So $x \in (\bigcup_{i \in I} A_i) \bigcup (\bigcup_{i \in I} B_i)$ proving that $\bigcup_{i \in I} (A_i \bigcup B_i) \subseteq (\bigcup_{i \in I} A_i) \bigcup (\bigcup_{i \in I} B_i)$ which combined with 2.39 gives

$$\left(\bigcup_{i \in I} A_i \right) \bigcup \left(\bigcup_{i \in I} B_i \right) = \bigcup_{i \in I} (A_i \bigcup B_i)$$

2. As $\forall i \in I$ we have by [theorem: 1.25] that $A_i \bigcap B_i \subseteq A_i$ and $A_i \bigcap B_i \subseteq A_i$, $B_i \subseteq A_i \bigcup B_i$ it follows using [theorem: 2.127] that $\bigcup_{i \in I} (A_i \bigcap B_i) \subseteq \bigcup_{i \in I} A_i$ and $\bigcup_{i \in I} (A_i \bigcap B_i) \subseteq \bigcup_{i \in I} B_i$. Using then [theorem: 1.25] we have

$$\bigcup_{i \in I} (A_i \bigcap B_i) \subseteq \left(\bigcup_{i \in I} A_i \right) \bigcup \left(\bigcup_{i \in I} B_i \right)$$ □

We have also a variant of the deMorgan's laws [theorem: 1.29]

Theorem 2.130. (deMorgan's Law) *Let $\{A_i\}_{i\in I}\subseteq B$ be a family then we have*

1. $(\bigcup_{i\in I} A_i)^c = \bigcap_{i\in I} (A_i)^c$
2. $(\bigcap_{i\in I} A_i)^c = \bigcup_{i\in I} (A_i)^c$
3. *If C is a class then* $C\setminus(\bigcup_{i\in I} A_i) = \bigcap_{i\in I} (C\setminus A_i)$
4. *If C is a class then* $C\setminus(\bigcap_{i\in I} A_i) = \bigcup_{i\in I} (C\setminus A_i)$

Proof.

1.

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

proving that

$$\left(\bigcup_{i\in I} A_i\right)^c = \bigcap_{i\in I} (A_i)^c$$

2.

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

proving that

$$\left(\bigcap_{i\in I} A_i\right)^c = \bigcup_{i\in I} (A_i)^c$$

3. We have

$$\begin{aligned}
C\setminus\left(\bigcup_{i\in I} A_i\right) &\underset{[\text{theorem: } 1.24]}{=} C\bigcap\left(\bigcup_{i\in I} A_i\right)^c\\
&\underset{(1)}{=} C\bigcap\left(\bigcap_{i\in I} (A_i)^c\right)\\
&\underset{[\text{theorem: } 2.128]}{=} \bigcap_{i\in I}\left(C\bigcap (A_i)^c\right)\\
&\underset{[\text{theorem: } 1.24]}{=} \bigcap_{i\in I}(C\setminus A_i)
\end{aligned}$$

4. We have

$$\begin{aligned} C\setminus\left(\bigcap_{i\in I}A_i\right) &\underset{[\text{theorem: }1.24]}{=} C\bigcap\left(\bigcap_{i\in I}A_i\right)^c \\ &\underset{(2)}{=} C\bigcap\left(\bigcup_{i\in I}(A_i)^c\right) \\ &\underset{[\text{theorem: }2.128]}{=} \bigcup_{i\in I}\left(C\bigcap(A_i)^c\right) \\ &= \bigcup_{i\in I}(C\setminus A_i) \end{aligned}$$

□

Theorem 2.131. *If* $\{A_i\}_{i\in I}\subseteq B$ *is a family and* A *a class then we have*

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

Proof.

1.

$$\begin{aligned} \left(\bigcup_{i\in I}A_i\right)\setminus A &\underset{[\text{theorem: }1.24]}{=} \left(\bigcup_{i\in I}A_i\right)\bigcap A^c \\ &\underset{[\text{theorem: }1.30]}{=} A^c\bigcap\left(\bigcup_{i\in I}A_i\right) \\ &\underset{[\text{theorem: }2.128]}{=} \bigcup_{i\in I}\left(A^c\bigcap A_i\right) \\ &\underset{[\text{theorem: }1.30]}{=} \bigcup_{i\in I}\left(A_i\bigcap A^c\right) \\ &\underset{[\text{theorem: }1.24]}{=} \bigcup_{i\in I}(A_i\setminus A) \end{aligned}$$

2.

$$\begin{aligned} \left(\bigcap_{i\in I}A_i\right)\setminus A &\underset{[\text{theorem: }1.24]}{=} \left(\bigcap_{i\in I}A_i\right)\bigcap A^c \\ &\underset{[\text{theorem: }1.30]}{=} A^c\bigcap\left(\bigcap_{i\in I}A_i\right) \\ &\underset{[\text{theorem: }2.128]}{=} \bigcap_{i\in I}\left(A^c\bigcap A_i\right) \\ &\underset{[\text{theorem: }1.30]}{=} \bigcap_{i\in I}\left(A_i\bigcap A^c\right) \\ &\underset{[\text{theorem: }1.24]}{=} \bigcap_{i\in I}(A_i\setminus A) \end{aligned}$$

3.

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

4.

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

5.

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

6.

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

$\square$

Theorem 2.132. *Let $\{A_i\}_{i\in I}\subseteq B$ a family then*

1. *If $j\in I$ then $\left(\bigcup_{i\in I\setminus\{j\}}A_i\right)\bigcup A_j=\bigcup_{i\in I}A_i$*
2. *$\bigcup_{i\in I}A_i=\bigcup_{i\in\{j\in I|A_j\neq\varnothing\}}A_i$*
3. *If $\exists i\in I$ such that $A_i=\varnothing$ then $\bigcap_{i\in I}A_i=\varnothing$*

Proof.

1. If $x\in\left(\bigcup_{i\in I\setminus\{j\}}A_i\right)\bigcup A_j$ then either $x\in A_j\subseteq\bigcup_{i\in I}A_i$ [see: 2.125], so that $x\in\bigcup_{i\in I}A_i$ or $x\in\bigcup_{i\in I\setminus\{j\}}A_i\Rightarrow\exists k\in I\setminus\{j\}$ with $x\in A_k$ which as $k\in I$ proves $x\in\bigcup_{i\in I}A_i$. If $x\in\bigcup_{i\in I}A_i$ then $\exists i\in I$ such that $x\in A_i$, we have then for i either $i\in I\setminus\{j\}$ so that $x\in\bigcup_{i\in I\setminus\{j\}}A_i$ or $i=j$ giving $x\in A_j$, proving that $x\in\left(\bigcup_{i\in I\setminus\{j\}}A_i\right)\bigcup A_j$.
2. As $\{j\in I|A_j\neq\varnothing\}\subseteq I$ we have by [theorem: 2.126] that

$$\bigcup_{i\in\{j\in I|A_j\neq\varnothing\}}A_i\subseteq\bigcup_{i\in I}A_i \tag{2.40}$$

Further if $x\in\bigcup_{i\in I}A_i$ then there exist a $i\in I$ such that $x\in A_i$. As $x\in A_i$ we must have that $A_i\neq\varnothing$ or $i\in\{j\in I|A_j\neq\varnothing\}$, proving that $x\in\bigcup_{i\in\{j\in I|A_j\neq\varnothing\}}A_i$. So

$$\bigcup_{i\in I}A_i\subseteq\bigcup_{i\in\{j\in I|A_j\neq\varnothing\}}A_i$$

combining this with [eq: 2.40] proves

$$\bigcup_{i\in I}A_i\subseteq\bigcup_{i\in\{j\in I|A_j\neq\varnothing\}}A_i$$

3. Assume that $i\in I$ such that $A_i=\varnothing$. If $x\in\bigcap_{j\in I}A_j$ we have $\forall j\in I$ that $x\in A_j$, so for sure $x\in A_i$ which contradicts $A_i=\varnothing$. Hence we have that $\bigcap_{j\in I}A_j=\varnothing$.

□

Theorem 2.133. *If $\{A_i\}_{i\in I}\subseteq C$ a family and $\forall i\in I$ $\{B_{i,j}\}_{j\in J}\subseteq C$ a family such that $A_i=\bigcup_{j\in J}B_{i,j}$ then*

$$\bigcup_{i\in I}A_i=\bigcup_{(i,j)\in I\times J}B_{i,j}$$

in othere words

Proof. If $x\in\bigcup_{i\in I}A_i$ then $\exists i\in I$ such that $x\in A_i=\bigcup_{j\in J}B_i$, hence $\exists j\in J$ such that $x\in B_{i,j}$. So as $(i,j)\in I\times J$ we have that $x\in\bigcup_{(i,j)\in I\times J}B_{i,j}$. Further if $x\in\bigcup_{(i,j)\in I\times J}B_{i,j}$ then $\exists(i,j)\in I\times J$ such that $x\in B_{i,j}$, which, as $A_i=\bigcup_{j\in J}B_{i,j}$, proves that $x\in A_i$, hence $x\in\bigcup_{i\in I}A_i$. So we conclude that

$$\bigcup_{i\in I}A_i=\bigcup_{(i,j)\in I\times J}B_{i,j}$$

□

Theorem 2.134. *If $f:A\rightarrow B$ is a function, $\{A_i\}_{i\in}\subseteq\mathcal{P}(A)$ and $\{B_i\}_{i\in I}\subseteq\mathcal{P}(B)$ are families of sub-classes of A and B then*

1. *$f(\bigcup_{i\in I}A_i)=\bigcup_{i\in I}f(A_i)$*
2. *$f^{-1}(\bigcup_{i\in I}B_i)=\bigcup_{i\in I}f^{-1}(B_i)$*
3. *$f(\bigcap A_{i\in I})\subseteq\bigcap_{i\in I}f(A_i)$*
4. *If f is injective and $I\neq\varnothing$ then $f(\bigcap_{i\in I}A_i)=\bigcap_{i\in I}f(A_i)$*
5. *$f^{-1}(\bigcap_{i\in I}B_i)=\bigcup_{i\in I}f^{-1}(B_i)$*

Proof.

1. If $y\in f(\bigcup_{i\in I}A_i)$ then $\exists x\in\bigcup_{i\in I}A_i$ such that $(x,y)\in f$, hence $\exists i\in I$ such that $x\in A_i$, which as $(x,y)\in f$ proves that $y\in f(A_i)$. So $y\in\bigcup_{i\in I}f(A_i)$ giving

$$f\left(\bigcup_{i\in I}A_i\right)\subseteq\bigcup_{i\in I}f(A_i) \tag{2.41}$$

If $y\in\bigcup_{i\in I}f(A_i)$ then there exists a $i\in I$ such that $y\in f(A_i)$, hence $\exists x\in A_i$ such that $(x,y)\in f$, as $x\in A_i$ this implies $x\in\bigcup_{i\in I}A_i$, so we have that $y\in f(\bigcup_{i\in I}A_i)$. Hence $\bigcup_{i\in I}f(A_i)\subseteq f(\bigcup_{i\in I}A_i)$, which combined with [eq: 2.41] gives

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

2. If $x\in f^{-1}(\bigcup_{i\in I}B_i)$ then there exists a $y\in\bigcup_{i\in I}B_i$ such that $(x,y)\in f$, hence $\exists i\in I$ such that $y\in B_i$. So $x\in f^{-1}(B_i)$ which as $i\in I$ implies that $x\in\bigcup_{i\in I}f^{-1}(B_i)$ or

$$f^{-1}\left(\bigcup_{i\in I}B_i\right)\subseteq\bigcup_{i\in I}f^{-1}(A_i) \tag{2.42}$$

If $x\in\bigcup_{i\in I}f^{-1}(A_i)$ then there exists a $i\in I$ such that $x\in f^{-1}(A_i)$, so $\exists y\in A_i$ with $(x,y)\in f$. As from $y\in A_i$ we have $y\in\bigcup_{i\in I}$ it follows that $x\in f^{-1}(\bigcup_{i\in I}A_i)$. This proves that $\bigcup_{i\in I}f^{-1}(A_i)\subseteq f^{-1}(\bigcup_{i\in I}A_i)$ which combined with [eq: 2.42] gives

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

3. If $y\in f(\bigcap_{i\in I}A_i)$ then there exists a $x\in\bigcap_{i\in I}A_i$ such that $(x,y)\in f$. From $x\in\bigcap_{i\in I}A_i$ it follows that $\forall i\in I\ x\in A_i$, which as $(x,y)\in f$ proves that $\forall i\in I\ x\in f(A_i)$ or $x\in\bigcap_{i\in I}f(A_i)$. So

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

4. Let $y\in\bigcap_{i\in I}f(A_i)$ then $\forall i\in I$ we have $y\in f(A_i)$. As $I\neq\varnothing$ there exists a $i\in I$ and we must thus have that $y\in f(A_i)$. So there exists a $x\in A_i$ such that $(x,y)\in f$. Assume that $x\notin\bigcap_{i\in I}A_i$ then $\exists j\in I$ such that $x\notin A_j$. However as $j\in I$ we must have that $y\in f(A_j)$, so there exists a $x'\in A_j$ such that $(x',y)\in f$. As f is injective and $(x,y),(x',y)\in f$ we must have $x=x'$, but this means that $x\in A_j$ contradicting $x\notin A_j$. So the assumption that $x\notin\bigcap_{i\in I}A_i$ is wrong, hence $x\in\bigcap A_i$. As $(x,y)\in f$ we have $y\in f(\bigcap_{i\in I}A_i)$, proving that $\bigcap_{i\in I}f(A_i)\subseteq f(\bigcap_{i\in I}A_i)$, which combined with (3) proves

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

5. If $x\in f^{-1}(\bigcap_{i\in I}B_i)$ then there exists a $y\in\bigcap_{i\in I}B_i$ such that $(x,y)\in f$. Hence $\forall i\in I$ we have that $y\in B_i\underset{(x,y)\in f}{\Rightarrow}x\in f^{-1}(B_i)$ proving that $x\in\bigcap_{i\in I}f^{-1}B_i$. So

$$f^{-1}\left(\bigcap_{i\in I}B_i\right)\subseteq\bigcap_{i\in I}f^{-1}(B_i)\tag{2.43}$$

If $x\in\bigcap_{i\in I}f^{-1}(B_i)$ then $\forall i\in I$ we have $x\in f^{-1}(B_i)$ or $\exists y\in B_i$ with $(x,y)\in f$. So $y\in\bigcap_{i\in I}B_i$ which as $(x,y)\in f$ proves that $x\in f^{-1}(\bigcap_{i\in I}B_i)$. So $\bigcap_{i\in I}f^{-1}(B_i)\subseteq f^{-1}(\bigcap_{i\in I}B_i)$ which combined with 2.43 gives

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

2.4 Product of a family of sets

The Cartesian product $A\times B$ consists of all the possible pairs that you can form, where the first element is a element of A and the second element is a element of B. We want now to construct a generalized product of a family of classes consisting of tuples whose elements are indexed by the index of the family.

Definition 2.135. (Product of a family of sets) *Let $\{A_i\}_{i\in I}\subseteq B$ a family then the* ***product of*** $\boldsymbol{\{A_i\}_{i\in I}}$ *noted as* $\boldsymbol{\prod_{i\in I}A_i}$ *is defined by*

$$\prod_{i\in I}A_i=\left\{f: f\in\left(\bigcup_{i\in I}A_i\right)^I\text{ where }\forall i\in I\text{ we have }f(i)\in A_i\right\}$$

If $x\in\prod_{i\in I}A_i$ then x_i is defined as

$$x_i=x(i)$$

Here $(\bigcup_{i\in I}A_i)^I$ is the class of function graphs of functions between I and $\bigcup_{i\in I}A_i$ [definition: 2.30] and $f(i)$ is the unique y such that $(i,y)\in f$. So $\prod_{i\in I}A_i$ is the class of graphs of functions from I to $\bigcup_{i\in I}A_i$ such that $\forall i\in I\ f_i=f(i)\in A_i$.

Theorem 2.136. *Let $\{A_i\}_{i\in I}\subseteq B$ be such that $\exists i_0\in I$ with $A_{i_0}=\varnothing$ then $\prod_{i\in I}A_i=\varnothing$*

Proof. Let $x\in\prod_{i\in I}A_i$ then we have that $x: I\to\bigcup_i A_i$ is a function such that $x(i_0)\in A_{i_0}$, contradicting $A_{i_0}=\varnothing$. Hence we must have that $\prod_{i\in I}A_i=\varnothing$. □

The following shows that the product of a family of only one class is 'almost' that class itself.

Example 2.137. Let $\{A_i\}_{i\in\{0\}}\subseteq\{B\}$ be the family in [example: 2.123] defined by $A:\{0\}\to\{B\}$ where $A=\{(0,B)\}$ then there exists a bijection between B and $\prod_{i\in\{0\}}A_i$ or as $A_0=B$ there exists a bijection between A_0 and $\prod_{i\in\{0\}}A_i$.

Proof. First using [example: 2.123] we have

$$B=\bigcup_{i\in\{0\}}A_i\tag{2.44}$$

hence

$$\left(\bigcup_{i\in\{0\}} A_i\right)^{\{0\}} = B^{\{0\}} \tag{2.45}$$

Let $f\in B^{\{0\}} \underset{\text{[eq: 2.45]}}{=} \left(\bigcup_{i\in\{0\}} A_i\right)$ then if $i\in\{0\}$ we must have $i=1$ hence $f(i)=f(0)\in B=A(0)=A_0$ proving that $\forall i\in\{0\}$ we have $f(i)\in A_i$. Hence $f\in\prod_{i\in\{0\}} A_i$ from which it follows that $B^{\{0\}}\subseteq\prod_{i\in\{0\}} A_i$. As clearly $\prod_{i\in\{0\}} A_i\subseteq\left(\bigcup_{i\in\{0\}} A_i\right)^{\{0\}} \underset{\text{[eq: 2.45]}}{=} B^{\{0\}}$ we have that

$$\prod_{i\in\{0\}} A_i = B^{\{0\}}$$

Now by [theorem: 2.76] there exists a bijection between B and $B^{\{0\}}$ which by the above proves the example. □

The next theorem shows that the product of a family of two classes is 'almost' the Cartesian product of these classes.

Theorem 2.138. *Let $\{A_i\}_{i\in\{0,1\}}\subseteq\{C,D\}$ be the family in [example: 2.124] defined by $A:\{0,1\}\rightarrow\{C,D\}$ where $A=\{(0,C),(1,D)\}$ then there exists a bijection between $A\times B$ and $\prod_{i\in\{0,1\}} A_i$*

Proof. First using [example: 2.124]: we have that

$$\bigcup_{i\in\{0,1\}} A_i = C\bigcup D \tag{2.46}$$

so that

$$\left(\bigcup_{i\in\{0,1\}} A_i\right)^{\{0,1\}} = \left(C\bigcup D\right)^{\{0,1\}} \tag{2.47}$$

So

$$\prod_{i\in\{0,1\}} A_i = \left\{f \mid f\in\left(C\bigcup D\right)^{\{0,1\}} \text{ where } f(0)\in C\wedge f(1)\in D\right\} \tag{2.48}$$

Given $(c,d)\in C\times D\Rightarrow c\in C\wedge d\in D$, define $f_{c,d}=\{(0,c),(1,d)\}$. If $(x,y)\in f_{c,d}$ we have either

$$(x,y)=(0,c)\Rightarrow x=0\in\{0,1\}\wedge y=c\in C\subseteq C\bigcup D\Rightarrow(x,y)\in\{0,1\}\times\left(C\bigcup D\right)$$

or

$$(x,y)=(1,d)\Rightarrow x=1\in\{0,1\}\wedge y=d\in D\subseteq C\bigcup D\Rightarrow(x,y)\in\{0,1\}\times\left(C\bigcup D\right)$$

proving that

$$f_{a,b}\subseteq\{0,1\}\times\left(C\bigcup D\right)\wedge f_{a,b}(0)\in C\wedge f_{a,b}(1)\in D \tag{2.49}$$

If $(x,y),(x,y')\in f_{c,d}$ then either

$$(x,y)=(0,c)\Rightarrow x=0\Rightarrow(0,y')\in f_{c,d}\Rightarrow(0,y')=(0,c)\Rightarrow y'=c\Rightarrow y=y'$$

or

$$(x,y)=(1,d)\Rightarrow x=1\Rightarrow(1,y')\in f_{c,d}\Rightarrow(1,y')=(1,d)\Rightarrow y'=d\Rightarrow y=y'.$$

Together with [eq: 2.49] this proves that

$$f_{a,b}:\{0,1\}\rightarrow C\bigcup D \text{ is a partial function} \tag{2.50}$$

If $x\in\{0,1\}$ then either $x=0\Rightarrow(0,c)\in f_{c,d}$ or $x=1\Rightarrow(1,d)\in f_{c,d}$ proving that $\{0,1\}\subseteq\operatorname{dom}(f_{c,d})$ which by [theorem: 2.26] proves that

$$f_{c,d}:\{0,1\}\rightarrow C\bigcup D \text{ is a function} \tag{2.51}$$

As by [eq: 2.49] $f_{c,d}(0)\in C\wedge f_{c,d}(1)\in D$ proving that

$$f_{c,d}\in\prod_{i\in\{0,1\}} A_i \tag{2.52}$$

Define now γ by $\gamma=\{((c,d),f_{c,d})|(c,d)\in C\times D\}$. If $(x,y)\in\gamma$ then $x=(c,d)\in C\times D$ and $y=f_{c,d} \underset{\text{[eq: 2.52]}}{\Rightarrow}$, hence $y\in(C\bigcup D)^{\{0,1\}}$. This proves that $(x,y)\in(C\times D)\times\left(\prod_{i\in\{0,1\}} A_i\right)$ or

$$\gamma\subseteq(C\times D)\times\left(\prod_{i\in\{0,1\}} A_i\right) \tag{2.53}$$

If $(x,y),(x,y')\in\gamma$ then $\exists(c,d)\in C\times D$ such that $(x,y)=((c,d),f_{c,d})$ and $(x,y')=((c,d),f_{c,d})$ so that $y=f_{c,d}=y'$ hence $y=y'$. Combining this with [eq:2.53] proves that

$$\gamma: C\times D\rightarrow\left(\prod_{i\in\{0,1\}}A_i\right)\text{ is a partial function}\tag{2.54}$$

If $(c,d)\in C\times D$ then by definition of γ we have $((c,d),f_{c,d})\in\gamma$ so that $(c,d)\in\mathrm{dom}(\gamma)$ proving that $C\times D\subseteq\mathrm{dom}(\gamma)$. By [theorem: 2.26] and [eq: 2.54] we have

$$\gamma: C\times D\rightarrow\left(\prod_{i\in\{0,1\}}A_i\right)\text{ is a function}\tag{2.55}$$

If $(x,y),(x',y)\in\gamma$ then there exists $(c,d),(c',d')\in C\times D$ such that $x=(c,d)\wedge x'=(c',d')$ and $f_{c,d}=y=f_{c',d'}$. As $(0,c)\in f_{c,d}=f_{c',d'}$ we have $(0,c)=(0,c')$ giving $c=c'$ and from $(1,d)\in f_{c,d}=f_{c',d'}$ we have $(1,d)=(1,d')$ giving $d=d'$. So $(c,d)=(c',d')$ proving that

$$\gamma: C\times D\rightarrow\left(\prod_{i\in\{0,1\}}A_i\right)\text{ is a injection}$$

If $g\in\prod_{i\in\{0,1\}}A_i$ then $g:\{0,1\}\rightarrow C\bigcup D$ is a function and $g(0)\in C\wedge g(1)\in D$ So there exists a $c\in C$ such that $(0,c)\in g$ and there exists a $d\in D$ such that $(1,d)\in g$. So $g=\{(0,c),(1,d)\}=f_{c,d}$ which proves that

$$\gamma: C\times D\rightarrow\left(\prod_{i\in\{0,1\}}A_i\right)\text{ is a surjection}\qquad\square$$

Theorem 2.139. *Let $\{A_i\}_{i\in I}\subseteq A$ and $\{B_i\}_{i\in I}\subseteq B$ classes such that $\forall i\in I$ $A_i\subseteq B_i$ 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 $x\in(\bigcup_{i\in I}A_i)^I$ and $\forall i\in I$ $x(i)\in A_i$. Using [theorem: 2.127] it follows that $\bigcup_{i\in I}A_i\subseteq\bigcup_{i\in I}B_i$, applying [theorem: 2.34] proves then $(\bigcup_{i\in I}A_i)^I\subseteq(\bigcup_{i\in I}B_i)^I$, so that

$$x\in\left(\bigcup_{i\in I}B_i\right)^I$$

If $i\in I$ then $x(i)\in A_i$, which as $A_i\subseteq B_i$ gives $x(i)\in B_i$, combining this with the above proves that $x\in\prod_{i\in I}B_i$. Hence we have

$$\prod_{i\in I}A_i\subseteq\prod_{i\in I}B_i\qquad\square$$

Theorem 2.140. *Let $\{A_i\}_{i\in I}\subseteq C$ and $\{B_i\}_{i\in I}\subseteq D$ are two families then*

$$\left(\prod_{i\in I}A_i\right)\bigcap\left(\prod_{i\in I}B_i\right)=\prod_{i\in I}\left(A_i\bigcap B_i\right)$$

Proof. First, as $\forall i\in I$ we have by [theorem: 1.25] $A_i\bigcap B_i\subseteq A_i$ and $A_i\bigcap B_i\subseteq B_i$ it follows that by [theorem: 2.139]

$$\prod_{i\in I}\left(A_i\bigcap B_i\right)\subseteq\prod_{i\in I}A_i\text{ and }\prod_{i\in I}\left(A_i\bigcap B_i\right)\subseteq\prod_{i\in I}B_i$$

so that by [theorem: 1.25]

$$\prod_{i\in I}\left(A_i\bigcap B_i\right)\subseteq\left(\prod_{i\in I}A_i\right)\bigcap\left(\bigcup_{i\in I}B_i\right)\tag{2.56}$$

Now for the opposite inclusion Let $x\in(\prod_{i\in I}A_i)\bigcap(\prod_{i\in I}B_i)$ then $x\in\prod_{i\in I}A_i$ and $x\in\prod_{i\in I}B_i$. So $x\in(\bigcup_{i\in I}A_i)^I\wedge\forall i\in I\vDash x(i)\in A_i$ and $x\in(\bigcup_{i\in I}B_i)^I\wedge\forall i\in I\vDash x(i)\in B_i$. Hence

$$\begin{array}{ll}x: I\rightarrow\bigcup_{i\in I}A_i & \text{is a function}\\ x: I\rightarrow\bigcup_{i\in I}B_i & \text{is a function}\\ \forall i\in I & \text{we have } x(i)\in A_i\bigcap B_i\end{array}$$

Now if $(i,y)\in x$ we have $i\in I$ [as $x\subseteq I\times(\bigcup_i A_i)$] and $y=x(i)\in A_i\bigcap B_i\subseteq\bigcup_{i\in I}(A_i\bigcap B_i)$ so that $(i,y)\in I\times(\bigcup_{i\in I}(A_i\bigcap B_i))$ giving

$$x\subseteq I\times\left(\bigcup_{i\in I}(A_i\bigcap B_i)\right)\text{ and }\forall i\in I\text{ we have }x(i)\in A_i\bigcap B_i \tag{2.57}$$

Further as $x: I\rightarrow\bigcup_{i\in I}A_i$ is a function we have $\forall (i,y),(i,y')$ that $y=y'$ and that $\operatorname{dom}(x)=I$. Combining this with [eq: 2.57] proves that $f: I\rightarrow\bigcup_{i\in I}(A_i\times B_i)$ is a function and $\forall i\in I$ we have $x(i)\in A_i\bigcap B_i$. This proves that $x\in\prod_{i\in I}(A_i\bigcap B_i)$ giving $(\prod_{i\in I}A_i)\bigcap(\prod_{i\in I}B_i)\subseteq\prod_{i\in I}(A_i\bigcap B_i)$ which combined with 2.56 gives finally

$$\prod_{i\in I}(A_i\bigcap B_i)\subseteq\left(\prod_{i\in I}A_i\right)\bigcap\left(\bigcup_{i\in I}B_i\right)$$

□

We can easily generalize the above theorem.

Theorem 2.141. *Let I, J be sets and $\{\{A_{i,j}\}_{j\in J}\}_{i\in I}$ be a family of sets then*

$$\bigcap_{i\in I}\left(\prod_{j\in J}A_{i,j}\right)=\prod_{j\in J}\left(\bigcap_{i\in I}A_{i,j}\right)$$

Proof. As $\forall i\in I, j\in I$ we have $\bigcap_{i\in I}A_{i,j}\subseteq A_{i,j}$ it follows from [theorem: 2.139] that $\prod_{j\in J}(\bigcap_{i\in I}A_{i,j})\subseteq\prod_{j\in J}A_{i,j}$. Hence by [theorem: 2.125]

$$\prod_{j\in J}\left(\bigcap_{i\in I}A_{i,j}\right)\subseteq\bigcap_{i\in I}\left(\prod_{j\in J}A_{i,j}\right) \tag{2.58}$$

If $x\in\bigcap_{i\in I}(\prod_{j\in J}A_{i,j})$ then $\forall i\in I\; x\in\prod_{j\in J}A_{i,j}$ hence $x: J\rightarrow A_{i,j}$ is a function such that $\forall j\in J$ we have $x(j)\in A_{i,j}$. Hence $x(j)\in\bigcap_{i\in I}A_{i,j}$ $\forall j\in J$ so that $x\in\prod_{j\in J}(\bigcap_{i\in I}A_{i,j})$, so

$$\bigcap_{i\in I}\left(\prod_{j\in J}A_{i,j}\right)\subseteq\prod_{j\in J}\left(\bigcap_{i\in I}A_{i,j}\right)$$

which combined with [eq: 2.58] gives

$$\bigcap_{i\in I}\left(\prod_{j\in J}A_{i,j}\right)=\prod_{j\in J}\left(\bigcap_{i\in I}A_{i,j}\right)$$

□

The following theorem is a motivation for the notation A^B for the graphs of functions from B to A.

Theorem 2.142. *Let I,B be classes and consider the family $\{A_i\}_{i\in I}\subseteq\{B\}$ based on the constant function $A: I\rightarrow\{B\}$ where $A=C_B=I\times\{B\}$ so that $\forall i\in I\; A(i)=B$ [see example: 2.45] then $\prod_{i\in I}A_i=A^I$*

Proof. For I we have the following cases to consider:

$I=\varnothing$. Using [example: 2.32] we have that

$$\left(\bigcup_{i\in\varnothing}A_i\right)^{\varnothing}=\{\varnothing\}$$

Further $\forall i\in\varnothing$ we have $\varnothing(i)\in A_i$ is satisfied vacuously proving that $\varnothing\in\prod_{i\in\varnothing}A_i$ so that $\{\varnothing\}\subseteq\prod_{i\in\varnothing}A_i\subseteq(\bigcup_{i\in\varnothing}A_i)^{\varnothing}=\{\varnothing\}$ or taking $I=\varnothing$

$$\prod_{i\in I}A_i=A^I$$

$I\neq\varnothing$. If $y\in\operatorname{range}(A)$ then $\exists x$ such that $(x,y)\in C_B=I\times\{B\}$, so that $y\in\{B\}$. Hence

$$\operatorname{range}(A)\subseteq\{B\} \tag{2.59}$$

As $I\neq\varnothing$ there exists a $i\in I$, which by the definition of C_B means that $(i,B)\in C_B$, hence $B\in\operatorname{range}(A)$. So if $y\in\{B\}$ then $y=B\in\operatorname{range}(A)$ proving that $\{B\}\subseteq\operatorname{range}(A)$ which combined with [eq: 2.59] gives

$$\operatorname{range}(A)=\{B\}$$

hence

$$\bigcup_{i\in I} A_i = \bigcup (\text{range}(A)) = \bigcup \{B\} \underset{[\text{example: } 1.58]}{=} B$$

so that

$$\left(\bigcup_{i\in I} A_i\right)^I = B^I \tag{2.60}$$

Now if $f \in B^I$ then $\forall i \in I$ we have $f(i) \in B = A(i) = A_i$ proving that

$$f \in \{f | f \in B^i \wedge \forall i \in I\, f(i) \in A_i\} \underset{[\text{eq: } 2.60]}{=} \left\{f | f \in \left(\prod_{i\in I} A_i\right)^I \wedge \forall i \in I\, f(i) \in A_i\right\} = \prod_{i\in I} A_i$$

proving that

$$B^I \subseteq \prod_{i\in I} A_i \tag{2.61}$$

Further

$$\prod_{i\in I} A_i = \left\{f | f \in \left(\prod_{i\in I} A_i\right)^I \wedge \forall i \in I\, f(i) \in A_i\right\} \subseteq \left\{f | f \in \left(\prod_{i\in I} A_i\right)^I\right\} = \{f | f \in B^I\} = B^I$$

which combined with [eq: 2.61] proves that

$$B^I = \prod_{i\in I} A_i$$

□

Theorem 2.143. *Let I, J, B be classes, $f: I \to J$ a bijection and $\{A_j\}_{j\in J}$ then*

$$\beta: \prod_{j\in J} A_j \to \prod_{i\in I} A_{f(i)} \text{ where } \beta(x) = x \circ f$$

is a bijection.

Proof. First as $f: I \to J$ is a bijection, hence surjective, we have by [theorem: 2.117] that

$$\bigcup_{j\in J} A_j = \bigcup_{i\in I} A_{f(i)} \tag{2.62}$$

Let $x \in \prod_{j\in J} A_j$ then $x \in (\bigcup_{j\in J} A_j)^J$, which is equivalent with $x: J \to \bigcup_{j\in J} A_j$ is a function, and $\forall j \in J$ we have $x(j) \in A_j$. So $x \circ f: I \to \bigcup_{j\in J} A_j \underset{[\text{eq: } 2.62]}{=} \bigcup_{i\in I} A_{f(i)}$ is a function, proving that $x \circ f \in (\bigcup_{i\in I} A_{f(i)})^I$, further if $i \in I$ then $(x \circ f)(i) = x(f(i)) \in A_{f(i)}$, hence

$$x \circ f \in \prod_{i\in I} A_{f(i)} \tag{2.63}$$

So

$$\beta: \prod_{j\in J} A_j \to \prod_{i\in I} A_{f(i)}$$

is indeed a function. To prove that it is a bijection note:

injectivity. Assume that $\beta(x) = \beta(y)$ then

$$\begin{aligned} x \circ f = y \circ f &\underset{f \text{ is bijective}}{\Rightarrow} (x \circ f) \circ f^{-1} = (y \circ f) \circ f^{-1} \\ &\Rightarrow x \circ (f \circ f^{-1}) = y \circ (f \circ f^{-1}) \\ &\Rightarrow x \circ \mathrm{Id}_J = y \circ \mathrm{Id}_J \\ &\Rightarrow x = y \end{aligned}$$

surjectivity. If $y \in \prod_{i\in I} A_{f(i)}$ then $y: I \to \bigcup_{i\in I} A_{f(i)} \underset{[\text{eq: } 2.62]}{=} \bigcup_{j\in J} A_j$ is a function and $\forall i \in I$ we have $y(i) \in A_{f(i)}$. As $f^{-1}: J \to I$ is a bijection we have that $y \circ f^{-1}: J \to \bigcup_{j\in J} A_j$ is a function, so that $y \circ f^{-1} \in (\bigcup_{j\in J} A_j)^J$, and $(y \circ f^{-1})(j) = y(f^{-1}(j)) \in A_{f(f^{-1}(j))} = A_j$. So that

$$y \circ f^{-1} \in \prod_{j\in J} A_j$$

Finally $\beta(y \circ f^{-1}) = (y \circ f^{-1}) \circ f = y \circ (f^{-1} \circ f) = y \circ \mathrm{Id}_I = y$ proving surjectivity. □

Definition 2.144. *Let $\{A_i\}_{i\in I}\subseteq B$ be a family and $J\subseteq I$ then $\prod_{i\in J}A_i$ is the product based on the sub-family $\{A_i\}_{i\in J}\subseteq B$ [see definition: 2.106] or equivalently*

$$\prod_{i\in J} A_i=\left\{f: f\in\left(\bigcup_{i\in J} A_i\right)^J \text{ where } \forall i\in J \text{ we have } f(i)\in A_i\right\}$$

The following theorem will be used later in induction arguments.

Theorem 2.145. *Let $\{A_i\}_{i\in I}\subseteq B$, $i\in I$ and $b\in A_i$ then*

$$\text{if } x\in\prod_{j\in I\setminus\{i\}} A_j \text{ we have } y\in\prod_{i\in I} A_j$$

where y is defined by

$$y_j=y(j)=\begin{cases} b \text{ if } j=i\\ x_j \text{ if } j\in I\setminus\{i\}\end{cases}\underset{\text{def}}{=}\begin{cases} b \text{ if } j=i\\ x(j) \text{ if } j\in I\setminus\{i\}\end{cases}$$

Proof. If $x\in\prod_{j\in I\setminus\{i\}}A_j$ then $x\in\left(\bigcup_{j\in I\setminus\{i\}}A_i\right)^{I\setminus\{i\}}$ so that $x: I\setminus\{i\}\rightarrow\bigcup_{j\in I\setminus\{i\}}A_j$ is a function. As $i\notin(I\setminus\{i\})$, $I=(I\setminus\{i\})\bigcup\{i\}$ and $\bigcup_{j\in I}A_j \underset{\text{[theorem: 2.132]}}{=} A_i\bigcup\left(\bigcup_{j\in I\setminus\{i\}}A_j\right)$ we have by [theorem: 2.82] that

$$y: I\rightarrow\bigcup_{i\in I} A_i \text{ where } y(j)=\begin{cases} b \text{ if } j=i\\ x(j) \text{ if } j\in I\setminus\{i\}\end{cases}$$

is a function, so

$$y\in\left(\bigcup_{i\in I} A_i\right)^I \tag{2.64}$$

Further if $j\in I$ then either $j=i$ so that $y_j=y(i)=b\in A_i=A_j$ or $j\in I\setminus\{i\}$ then $y_j=y(j)=x(j)=x_j\in A_j$. Hence

$$\forall j\in I \text{ we have } y_j\in A_j \tag{2.65}$$

From [eq: 2.64] and [eq: 2.65] it follows by

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

□

We introduce now the projection operator

Definition 2.146. *Let $\{A_i\}_{i\in I}\subseteq B$ be family then for $i\in I$ we define the projection function*

$$\pi_i: \prod_{j\in I} A_j\rightarrow A_i$$

where

$$\pi_i=\left\{z|z=(x,x(i))|x\in\prod_{j\in I} A_j\right\}$$

In other words $(x,y)\in\pi_i\Leftrightarrow x\in\prod_{j\in I}A_j$ and $y=x(i)\Leftrightarrow(i,y)\in x$

Proof. This definition only make sense if $\forall i\in I$ that $\pi_i: \prod_{j\in I}A_j\rightarrow A_i$ is a function. First if $(x,y)\in\pi_i$ we have that $x\in\prod_{j\in I}A_j$ and $y=x(i)$ giving $y\in A_i$, so $(x,y)\in(\prod_{i\in I}A_i)\times A_i$. Hence

$$\pi_i\subseteq\left(\prod_{i\in I} A_i\right)\times A_i \tag{2.66}$$

If $(x,y),(x,y')\in\pi_i$ then $y=x(i)\wedge y'=x(i)$ proving that $y=y'$ or

$$\pi_i: \prod_{j\in I} A_j\rightarrow A_i \text{ is a partial function}$$

If $x\in\prod_{j\in I}A_j$ then by definition $(x,x(i))\in\pi_i$ proving that $x\in\mathrm{dom}(\pi_i)$ proving that $\prod_{j\in I}A_i\subseteq\mathrm{dom}(\pi_i)$, which by [theorem: 2.26] gives

$$\pi_i: \prod_{j\in I} A_j\rightarrow A_i \text{ is a function}$$

□

We are not yet finished with the product of a family of classes, however for some of the theorems we need the Axiom of Choice. For example to prove that the projection function is a surjection we need the Axiom of Choice.

Chapter 3
Relations

3.1 Relation

The idea of a relation is that we can specify which elements of a class are related to each other. You do this by specifying a class of pairs.

Definition 3.1. *Let A be a class then a relation in A is a sub-class of $A \times A$*

Notation 3.2. *So a relation is a set of pairs from elements of the same class, to avoid confusion with the graph of a function we use the following notation:*
If $R \subseteq A \times A$ is relation then instead of writing $(x, y) \in R$ we write $x R y$

Example 3.3. Let A be a class then $A \times A$ is a relation [as $A \times A \subseteq A \times A$]

We define now the following properties that a relation can have

Definition 3.4. *If A is a class and $R \subseteq A \times A$ a relation then we say that R is*

reflexive. *iff $\forall x \in A$ we have*

$$x R x$$

in other words every element is related to itself.

symmetric. *iff*

$$x R y \Rightarrow y R x$$

in other words if one element is related to a second element then the second element is related to the first element.

anti symmetric. *iff*

$$x R y \wedge y R x \Rightarrow x = y$$

in other words if on element is related to a second element and the second element is related to the first element then the two elements are the same.

transitive. *iff*

$$x R y \wedge y R z \Rightarrow x R z$$

in other words if one element is related to a second element and the second element is related to the third element then the first element is also related to the third element.

3.2 Equivalence relations

3.2.1 Equivalence relations and equivalence classes

Note that for classes and equality we have by [theorem: 1.8] that

- $A = A$
- $A = B \Rightarrow B = A$
- $A = B \wedge B = C \Rightarrow A = C$

If we want to create a relation that defines a kind of equality then it must behave in the same way as the equality for classes. This it he idea behind the following definition.

Definition 3.5. (Equivalence Relation) *If A is a class then a relation R is a* ***equivalence relation*** *iff it is reflexive, symmetric and transitive or in other words if*

reflectivity. $\forall x \in A\ x R x$

symetricity. $x R y \Rightarrow y R x$

transitivity. $x R y \wedge y R z \Rightarrow x R z$

Given a set A and a equivalence relation in A then it is useful to partition the set in subsets containing all the elements that are equivalent with each other. To do this we must first define what a partition of a set is.

Definition 3.6. *Let A be a set then a* ***partition*** *of A is a family $\{A_i\}_{i\in I} \subseteq \mathcal{P}(A)$ of non empty subsets of A [$\forall i \in I$ we have $A_i \neq \varnothing$] such that:*

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

Note 3.7. Condition (2) in the above definition is a weaker condition that pairwise disjointedness. For example if we define the family $(A_i)_{i\in\{1,2,3\}}$ by $A_1 = \{1\}, A_2 = \{1\}$ and $A_3 = \{2\}$ then this family is not pairwise disjoint as $1 \neq 2$ and $A_1 \bigcap A_2 \neq \varnothing$, however (2) is clearly satisfied.

We can also reformulate the definition of a partition of A in the following way

Theorem 3.8. *Let A be a set and $\{A_i\}_{i\in I} \subseteq \mathcal{P}(A)$ a family of non empty subsets of A then we have the following equivalences*

1. *$\{A_i\}_{i\in I} \subseteq \mathcal{P}(A)$ is a partition of A*
2. *$\{A_i\}_{i\in I} \subseteq \mathcal{P}(A)$ satisfies*
   a. *$\forall x \in A$ there exists a $i \in I$ such that $x \in A_i$*
   b. *$\forall i, j \in I$ with $A_i \bigcap A_j \neq \varnothing$ we have $A_i = A_j$*
3. *$\{A_i\}_{i\in I} \subseteq \mathcal{P}(A)$ satisfies*
   a. *$\forall x \in A$ there exists a $i \in I$ such that $x \in A_i$*
   b. *$\forall i, j \in I$ with $A_i \neq A_j$ we have $A_i \bigcap A_j = \varnothing$*

Proof.

1 ⇒ 2.

a) If $x \in A$ then as $A = \bigcup_{i\in I} A_i$ there exists a $i \in I$ such that $x \in A_i$

b) Let $i, j \in I$ with $A_i \bigcap A_j \neq \varnothing$. As by definition of a partition $A_i \bigcap A_j = \varnothing \vee A_i = A_j$ we must have that $A_i = A_j$.

2 ⇒ 3.

a) This is trivial

b) Let $i, j \in I$ with $A_i \neq A_j$. Assume that $A_i \bigcap A_j \neq \varnothing$ then by (2.b) we have $A_i = A_j$ contradicting $A_i = A_j$, so we must have that $A_i \bigcap A_j = \varnothing$

3 ⇒ 1.

a) Using (3.a) it follows that $A \subseteq \bigcup_{i\in I} A_i$. If $z \in \bigcup_{i\in I} A_i$ then there exists a $i \in I$ such that $x \in A_i$ [theorem: 2.116], hence as $A_i \in \mathcal{P}(A) \Rightarrow A_i \subseteq A$ it follows that $x \in A$, proving that $\bigcup_{i\in I} A_i \subseteq A$. So we have that

$$\bigcup_{i\in I} A_i = A$$

b) Let $i, j \in I$ then if $A_i \neq A_j$ we have by (3b) that $A_i \bigcap A_j = \varnothing$, so we have that $A_i = A_j \vee A_i \bigcap A_j = \varnothing$. □

We show now how a equivalence relation can be used to partition a set.

Definition 3.9. *Let A be a set and R a equivalence relation in A then given x we define the* ***equivalence class*** *of x noted by $R[x]$ by*

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

Note 3.10. Because $R[x]\subseteq A$ and A is a set we have by the axiom of subset 1.54 that $R[x]$ is a set.

We have the following important property for equivalence classes

Theorem 3.11. *Let A be a set with a equivalence relation R in A then*

1. $\forall x\in A$ *we have* $x\in R[x]$
2. $\forall x,y\in A$ *we have*

$$x\,R\,y\Leftrightarrow R[x]=R[y]$$

3. $\forall x\in A$ *we have*

$$y\in R[x]\Leftrightarrow R[x]=R[y]$$

Proof.

1. If $x\in A$ then using reflexivity we have $x\,R\,x$ so that $x\in R[x]$
2.
   - **$\Rightarrow$.** Let $z\in R[x]$ then $x\,R\,z$, further from $x\,R\,y$ we have $y\,R\,x$, so using transitivity it follows that $y\,R\,z$ or $z\in R[y]$. If $z\in R[y]$ then $y\,R\,z$ so as $x\,R\,y$ we have by transitivity that $x\,R\,z$ or that $z\in R$.
   - **$\Leftarrow$.** Using (1) $x\in R[x]\underset{R[x]=R[y]}{\Rightarrow}x\in R[y]$ proving that $r\,R\,y$
3.
   - **$\Rightarrow$.** If $y\in R[x]$ then $y\,R\,x$ hence by (2) $R[x]=R[y]$
   - **$\Leftarrow$.** If $R[x]=R[y]$ then $y\,R\,x$ proving that $y\in R[x]$ □

We define now a function that maps a element of as set on its equivalence class and use it to define a family of equivalence classes indexed by the elements of the set.

Definition 3.12. *Let A be a set and R a equivalence relation in A then $\{R[x]\}_{x\in A}\subseteq\mathcal{P}(X)$ is the family defined by the function $R[]$: $A\rightarrow\mathcal{P}(A)$ where $R[](x)=R[x]$*

Note 3.13. As $x\in R[x]$ we have that $\{R[x]\}_{x\in A}$ is a non empty family of subsets of A

Proof. We must of course prove that this a function. First $R[x]$ is defined for every $x\in A$ and calculates a unique set, further $R[x]\subseteq A\Rightarrow R[x]\in\mathcal{P}(A)$. So by [proposition: 2.94] $R[]: A\rightarrow\mathcal{P}[A]$ is a function. □

Theorem 3.14. *Let A be a set and R a equivalence relation in A then $\{R[x]\}_{x\in A}$ is a partition of A*

Proof. We use [theorem: 3.8] to prove this

1. If $x\in A$ then by [theorem: 3.11] we have that $x\in R[x]$ so that $x\in\bigcup_{x\in A}R[x]$
2. Let $x,y\in A$ such that $R[x]\bigcap R[y]\neq\varnothing$ then there exists a

$$z\in R[x]\bigcap R[y]\Rightarrow z\,R\,x\wedge z\,R\,y\underset{\text{symmetry}}{\Rightarrow}x\,R\,z\wedge z\,R\,y\underset{\text{transitivity}}{\Rightarrow}x\,R\,y$$

   Using the above together with [theorem: 3.11] we have then that $R[x]=R[y]$

So by [theorem: 3.8] it follows that $\{R[x]\}_{x\in A}\subseteq\mathcal{P}(A)$ is a partition of A □

We have also the opposite of the above theorem in that a partition defines a equivalence relation that generates the same partition.

Theorem 3.15. *Let A be a set and $\{A_i\}_{i\in I}\subseteq\mathcal{P}(A)$ a partition of A. Define $R\subseteq A\times A$ by*

$$R=\{(x,y)|\exists i\in I\text{ such that }x\in A_i\wedge y\in A_i\}$$

then we have:

1. *R is a equivalence relation*
2. *$\forall i \in I$ there exists a $x \in A$ such that $R[x] = A_i$*
3. *$\forall x \in A$ there exists a $i \in I$ such that $R[x] = A_i$*

we call R is the called the ***equivalence relation associated with the partition*** $\{A_i\}_{i \in I} \subseteq \mathcal{P}(A)$

Proof.

1. We have:
   a. If $x \in A = \bigcup_{i \in I} A_i$ then $\exists i \in I$ such that $x \in A_i$ so that $(x, x) \in R$ or $x R x$
   b. If $x R y$ or $(x, y) \in R$ then $\exists i \in I$ such that $x \in A_i \wedge y \in A_i \Rightarrow y \in A_i \wedge x \in A_i$. Hence $(y, x) \in R$ or $y R x$.
   c. If $x R y \wedge y R z$ then $\exists i \in I$ such that $x, y \in A_i$ and $\exists j \in I$ such that $y, z \in A_j$. So $y \in A_i \bigcap A_j$ or $A_i \bigcap A_j \neq \varnothing$, by [theorem: 3.8] we have that $A_i = A_j$, hence $x, z \in A_i$ proving that $(x, z) \in R$ or $x R z$.
2. If $i \in I$ then as $A_i \neq \varnothing$ [a partition is a family of non empty subsets] there exists a $x \in A_i$. Take $y \in A_i$ then $x, y \in A_i$ or $y R x$ proving that $y \in R[x]$. So
$$A_i \subseteq R[x]$$
Take $y \in R[x]$ then $y R x$ so there exist a $j \in I$ such that $x, y \in A_j$, hence $A_i \bigcap A_j \neq \varnothing$ which by [theorem: 3.8] proves that $A_i = A_j$, so that $y \in A_i$. So $R[x] \subseteq A_i$ giving
$$A_i = R[x]$$
3. If $x \in A$ then $\exists i \in I$ such that $x \in A_i$. Take $y \in A_i$ then $x, y \in A_i$ or $y R x$ proving that $y \in R[x]$, hence
$$A_i \subseteq R[x]$$
Take $y \in R[x]$ then $y R x$ so there exist a $j \in I$ such that $x, y \in A_j$, hence $A_i \bigcap A_j \neq \varnothing$ which by [theorem: 3.8] proves that $A_i = A_j$, so that $y \in A_i$. So $R[x] \subseteq A_i$ giving
$$A_i = R[x]$$

□

Definition 3.16. *Let A be a set and R a equivalence relation then A/R is defined by*

$$A/R = \{R[x] | x \in A\}$$

Note 3.17. As $\forall x \in X$ $R[x] \in \mathcal{P}(A)$ it follows that

$$R/X \in \mathcal{P}(A).$$

As A is a set it follows from the Axiom Power [axiom: 1.64] that $P(A)$ is a set, applying the Axiom of Subsets [axiom: 1.54] we have

$$R/X \text{ is a set}$$

Theorem 3.18. *Let A be a set and R a equivalence relation then*

$$A = \bigcup_{I \in A/R} I \text{ and } \forall I, J \in A/R \text{ with } I \neq J \text{ we have } I \bigcap J = \varnothing$$

Proof. If $x \in A$ then by [theorem: 3.11] $x \in R[x] \in A/R$ so that $x \in \bigcup_{I \in A/R} I$ hence

$$A \subseteq \bigcup_{I \in A/R} I$$

Further, as $\forall I \in A/R$ we have that $\exists x \in A$ such that $I = R[x] \subseteq A$, it follows that $\bigcup_{I \in A/R} I \subseteq A$ which combined with the above gives

$$A = \bigcup_{I \in A/R} I$$

Further if $I, J \in A/R$ with $I \neq J$ then $\exists x, y \in A$ such that $I = R[x]$ and $J = R[y]$. If $z \in R[x] \bigcap R[y]$ then by [theorem: 3.11] we have $R[x] = R[z] = R[y]$ so that $I = J$ contradicting $I \neq J$. Hence we must have that

$$I \bigcap J = \varnothing$$

□

3.2.2 Functions and equivalence relations

In this section we show how a function can be decomposed as the composition of a surjection, a bijection and injection. First we examine the relation between functions and equivalence relations.
We can use functions to generate a equivalence relation on the domain of the function based on a equivalence relation on the target of the function.

Theorem 3.19. *$f: A \rightarrow B$ a function and R a equivalence relation in A then*

$$f\langle R \rangle = \{(x, y) | f(x) \, R \, f(y)\} \subseteq A \times A$$

is a equivalence relation in A

Proof.

reflectivity. If $x \in A$ then $f(x) \in B$ so that $f(x) \, R \, f(x)$ hence by definition $x R x$

symmetric. If $x R y$ then $f(x) \, R \, f(y)$ so that $f(y) \, R \, f(x)$ proving $y R x$

transitivity. If $x R y \wedge y R z$ then $f(x) \, R \, f(y) \wedge f(y) \, R \, f(z)$ so that $f(x) \, R \, f(x)$ proving that $x R z$. □

A equivalence relation on a set induce a equivalence relation on a subset

Theorem 3.20. *Let A be a class, $B \subseteq A$ a sub-class and R a equivalence relation in R then $R_{|B}$ defined by*

$$R_{|B} = \{(x, y) | x \in B \wedge y \in B \wedge x R y\} = R \bigcap (B \times B)$$

is a equivalence relation.

Proof.

reflexivity. If $x \in B$ then $x R x$ so that $x R_{|B} x$

symmetric. If $x R_{|B} y \Rightarrow x \in B \wedge y \in B \wedge \text{xRy} \Rightarrow \text{yR}_{|B} x$

transitivity. If $x R_{|B} y \wedge y R_{|B} z$ then $x, y, z \in \mathbb{R}$ and $x R y \wedge y R z$ so that $x, z \in B$ and $x R z$ proving $x R_{|B} z$ □

Theorem 3.21. *If $f: A \rightarrow B$ is a function then R_f defined by*

$$R_f = \{(x, y) \in A \times A | f(x) = f(y)\}$$

is a relation. R_f is called the ***equivalence relation determined by f***

Proof.

reflexivity. If $x \in A$ then $f(x) = f(x)$ proving that $x R_f x$

symmetric. If $x R_f y$ then $f(x) = f(y) \Rightarrow f(y) = f(x)$ proving that $y R_f x$

transitivity. If $x R_f y \wedge y R_f x$ then $f(x) = f(y)$ and $f(y) = f(z)$ so that $f(x) = f(x)$ hence $x R_f z$ □

We can also do the opposite and associate a function with a equivalence relation

Theorem 3.22. (Canonical Function) *Let A be a set and R a equivalence relation in A then:*

1. *$f_R: A \rightarrow A/R$ defined by $f_R(x) = R[x]$ is a surjective function.*
2. *$R_{R_f} = R$*

$f_R: A \rightarrow A/R$ is called the **Canonical function associated with *R***

Proof.

1. As for every $x \in A$ we have the unique $R[x] \in R/X$ it follows from [proposition: 2.94] that

$$f_R: A \to A/R \text{ is a function}$$

Let $y \in R/X$ then $\exists x \in A$ such that $y = R[x]$ so that $(x, y) = (x, R[x]) \in f_R$ proving that $y \in \text{range}(f_R)$. So $R/X \subseteq \text{range}(f_R)$ which by [theorem: 2.51] proves that

$$f_R: A \to A/R \text{ is surjective}$$

2. We have

$$\begin{aligned} (x, y) \in R &\Leftrightarrow x R y \\ &\underset{\text{[theorem: 3.11]}}{\Leftrightarrow} R[x] = R[y] \\ &\Leftrightarrow f_R(x) = f_R(y) \\ &\Leftrightarrow (x, y) \in R_{f_R} \end{aligned}$$

$\square$

We use the above to decompose every function as the composition of a surjection, bijection and injection.

Theorem 3.23. *Let A, B be sets and $f: A \to B$ a function and define the following functions:*

a) $s_f: A/R_f \to f(A)$ where $s_f = \{(R_f[x], f(x)) | x \in A\}$

b) $i_{f(A)}: f(A) \to B$ where $i_{f(A)} = \{(x, x) | x \in f(A)\}$ [the inclusion function see [example: 2.53]

c) $f_{R_f}: A \to A/R_f$ where $f_{R_f}(x) = R_f[x]$ [theorem: 3.22]

then

1. *$s_f: A/R_f \to f(A)$ is a bijection*
2. *$i_{f(A)}: f(A) \to B$ is a injective function*
3. *$f_{R_f}: A \to A/R_f$ is a surjective function*
4. *$f = i_{f(A)} \circ (s_f \circ f_{R_f}) \underset{\text{[theorem: 2.21]}}{=} (i_{f(A)} \circ s_f) \circ f_{R_f}$*

Proof. Using [example: 2.53] and [theorem: 3.22] we have that

$$i_{f(A)}: f(A) \to B \text{ is a injective function}$$

and

$$f_{R_f}: A \to A/R_f \text{ is surjective function}$$

We proceed now to prove that s_f is a bijection. If $(x, y) \in s_f$ then there exists a $a \in A$ such that $(x, y) = (R_f[a], f(a))$ hence $x = R_f[a] \in A/R_f$ and $y = f(a) \Rightarrow (a, y) \in f \Rightarrow y \in f(A)$. So that $(x, y) \in (A/R_f) \times f(A)$ or

$$s_f \subseteq (A/R_f) \times f(A)$$

If $(x, y), (x, y') \in s_f$ then there exists $a, a' \in A$ such that

$$(x, y) = (R_f[a], f(a)) \wedge (x, y') = (R_f[a'], f(a'))$$

or

$$x = R_f[a] \wedge y = f(a) \wedge x = R_f[a'] \wedge y' = f(a') \tag{3.1}$$

From the above $R_f[a] = x = R_f[a']$, which using [theorem: 3.11] means that $a R_f a'$, so by the definition of R_f [theorem: 3.21] we have $f(a) = f(a')$. As by [eq: 3.1] $y = f(a) \wedge y' = f(a')$ it follows that $y = y'$. So

$$s_f: A/R_f \to f(A) \text{ is a partial function}$$

If $x \in A/R_f$ then $\exists a \in A$ such that $x = [a]$, hence if we take $y = f(A)$ we have that $(x, y) = ([a], f(a)) \in s_f$ proving that $x \in \text{dom}(s_f)$. So $A/R_f \subseteq \text{dom}(f)$ which by [proposition: 2.26] proves that

$$s_f: A/R_f \to f(A) \text{ is a function}$$

Let $(x, y), (x', y) \in s_f$ then $\exists a, a' \in A$ such that $(x, y) = (R_f[a], f(a))$ and $(x', y) = (R_f[a'], f(a'))$, hence

$$x = R_f[a] \wedge x' = R_f[a'] \wedge y = f(a) \wedge y = f(a') \tag{3.2}$$

From $f(a) = y = f(a')$ it follows that $f(a) = f(a')$, which by the definition of R_f [theorem: 3.21] proves that $a R_f a'$. Using [theorem: 3.11 it follows that $R_f[a] = R_f[a']$ or using [eq: 3.2] that $x = x'$. So we have proved that

$$s_f: A / R_f \rightarrow f(A) \text{ is injective} \tag{3.3}$$

Let $y \in f(A)$ then there exist a $a \in A$ such that $(a, y) \in f \Rightarrow y = f(a)$. But then $(R_f[a], y) = (R_f[a], f(a)) \in s_f$ proving that $y \in \text{range}(s_f)$. So $A / R_f \subseteq \text{range}(s_f)$ which by [proposition: 2.51] proves that

$$s_f: A / R_f \rightarrow f(A) \text{ is surjective} \tag{3.4}$$

Combining [eq: 3.3] and [eq: 3.4] it follows that

$$s_f: A / R_f \rightarrow f(A) \text{ is a bijection}$$

Now we proceed to prove that $f = (i_{f(A)} \circ s_f) \circ f_{R_f}$. Let $(x, u) \in (i_{f(A)} \circ s_f) \circ f_{R_f}$ then $\exists y$ such that $(x, y) \in f_{R_f} \wedge (y, u) \in i_{f(A)} \circ s_f$, from $(y, u) \in i_{f(A)} \circ s_f$ $\exists$z such that $(y, z) \in s_f \wedge (z, u) \in i_{f(A)}$, summarized

$$(x, y) \in f_{R_f} \wedge (y, z) \in s_f \wedge (z, u) \in i_{f(A)} \tag{3.5}$$

From $(x, y) \in f_{R_f}$ it follows that $\exists a \in A$ such that $(x, y) = (a, R_f[a])$ or

$$x = a \wedge y = R_f[a] \tag{3.6}$$

From $(y, z) \in s_f$ it follows that $\exists a' \in A$ such that $(y, z) = (R_f[a'], f(a'))$ or $y = R_f[a'] \wedge z = f(a')$. As $y \underset{\text{[eq: 3.6]}}{=} R_f[a]$ we have that $R_f[a] = R_f[a']$, which by [theorem: 3.11] proves that $a R_f a'$, so by the definition of R_f we have $f(a) = f(a')$ hence $z = f(a)$. From $(z, u) \in i_{f(A)}$ it follows that $z = u$ hence $u = f(a)$. As $x \underset{\text{[eq: 3.6]}}{=} a$ it follows that $(x, u) = (a, f(a)) \in f$. Hence

$$(i_{f(A)} \circ s_f) \circ f_{R_f} \subseteq f \tag{3.7}$$

Finally if $(x, y) \in f$ then as $f \subseteq A \times B$ proves that $x \in A$ and $f(x) = y \in f(A)$. Hence $(R_f[x], f(x)) \in s_f$, $(x, R_f[x]) \in f_{R_f}$ and $(f(x), y) = (f(x), f(x)) \in i_{f(A)}$. So that $(R_f[x], y) \in i_{f(A)} \circ s_f$ and $(x, R_f[x]) \in f_{R_f}$ proving that $(x, y) \in (i_{f(A)} \circ s_f) \circ f_{R_f}$. So $f \subseteq (i_{f(A)} \circ s_f) \circ f_{R_f}$ which combined with [eq: 3.7] gives

$$f = (i_{f(A)} \circ s_f) \circ f_{R_f}$$ □

Notation 3.24. *For the rest of this book we use the standard convention of noting a equivalence relation as $\sim$, The definition of $\sim$ should then be clear from the context. If many equivalence relations are used in the same context we use superscripts like $\sim_{\mathbb{R}}$ and $\sim_{\mathbb{Z}}$ to avoid conflicts.*

3.3 Partial ordered classes

3.3.1 Order relation

First we define a partial order relation that allows us to compare two elements and specify which element 'lies before' another element.

Definition 3.25. (Pre-order) *Let A be a class then a relation $R \subseteq A \times A$ in A is a pre-order if it is* ***reflexive*** *and* ***transitive*** *or in other words:*

- ***reflectivity.*** *$\forall x \in A$ we have $x R x$*
- ***transitivity.*** *If $x R y \wedge y R z$ then $x R z$*

Definition 3.26. *$\langle A, R \rangle$ is a pre-ordered class iff A is a class and R is a pre-order in A*

A order relation is a pre-order with one extra condition

Definition 3.27. (Order relation) *If A is a class then a relation $R \subseteq A \times A$ in A is a* ***order*** *if it is a pre-order that is anti-symmetric or in other words:*

- ***reflectivity.*** *$\forall x \in A$ we have $x R x$*

anti-symmetry. *If* $x R y \wedge y R x$ *then* $x = y$

transitive. *If* $x R y \wedge y R z$ *then* $x R z$

Definition 3.28. (Partial ordered class) $\langle A, R\rangle$ *is a* ***partial ordered class*** *if* A *is a class and* R *is a order.*

Notation 3.29. *We use the standard convention of noting a pre-order relation as* $\leqslant$*, The definition of* $\leqslant$ *should then be clear from the context. If many pre-order relations are used in the same context we use superscripts like* $\leqslant_{\mathbb{R}}$ *and* $\leqslant_{\mathbb{Z}}$ *or* $\preccurlyeq$ *to avoid conflicts.*

Definition 3.30. *If* $\langle A, \leqslant\rangle$ *is a pre-ordered or partial class and* $x, y, z \in A$ *then we define:*

$$\begin{array}{lll} x \leqslant y \leqslant z & \text{is the same as} & x \leqslant y \wedge y \leqslant z \\ x \leqslant y < z & \text{is the same as} & x \leqslant y \wedge y < z \\ x < y \leqslant z & \text{is the same as} & x < y \wedge y \leqslant z \\ x < y < z & \text{is the same as} & x < y \wedge y < z \end{array}$$

Definition 3.31. *If* $\langle A, \leqslant\rangle$ *is a pre-ordered class [or partial ordered class] then* $x < y$ *is equivalent with* $x \leqslant y \wedge x \neq y$

Theorem 3.32. *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. $(x < y \vee x = y) \Leftrightarrow (x \leqslant y)$

or in other words

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

Proof.

1. If $x \leqslant y \wedge y < z$ then $x \leqslant y \wedge y \leqslant z \wedge y \neq z$, so that $x \leqslant z$ and $y \neq z$. Assume that $x = z$ then $z \leqslant y \underset{y \leqslant z}{=} z = y$ contradicting $y \neq z$, so we must have $x \neq z$, which together with $x \leqslant z$ gives
$$x < z$$
2. If $x < y \wedge y \leqslant z$ then $x \leqslant y \wedge y \leqslant z \wedge x \neq y$, so that $x \leqslant z$ and $x \neq y$. Assume that $x = z$ then $y \leqslant x \underset{x \leqslant y}{\Rightarrow} y = x$ contradicting $x \neq y$, so we must have $x \neq z$, which together with $x \leqslant z$ gives
$$x < z$$
3. If $x < y \wedge y < z$ then $x \neq y \wedge x \leqslant y \wedge y < z$ so that by (1) we have $x < z$
4. We have
$$\begin{array}{rcl} (x < y \vee x = y) & \Leftrightarrow & ((x \leqslant y \wedge x \neq y) \vee x = y) \\ & \Leftrightarrow & ((x \leqslant y \vee x = y) \wedge (x \neq y \vee x = y)) \\ & \Leftrightarrow & x \leqslant y \vee x = y \\ & \Leftrightarrow & x \leqslant y \end{array}$$

□

Example 3.33. Let A be a class of classes and $\leqslant$ defined by $\leqslant = \{(x, y) \in \mathcal{A} \times \mathcal{A} | x \subseteq y\}$ then $\langle \mathcal{A}, \leqslant\rangle$ is a partial ordered class

Proof.

reflectivity. If $A \in \mathcal{C}$ then by [theorem: 1.8] $A \subseteq A$ so that $A \leqslant A$

anti-symmetric. If $A \leqslant B$ and $B \leqslant A$ then $A \subseteq B \wedge B \subseteq A$ so that by [theorem: 1.8] $A = B$

transitivity. If $A \leqslant B \wedge B \leqslant C$ then $A \subseteq B \wedge B \subseteq C$ so that by [theorem: 1.8] $A \subseteq C$ or $A \leqslant C$ □

Every pre-order can be used as the base to create a order relation as is expressed in the following theorem. The basic idea is that $x \leqslant y \wedge y \leqslant x \Rightarrow x = y$ is missing from a pre-order. By defining a equivalence relation $\sim$ such that $x \sim y$ if $x \leqslant y \wedge y \leqslant x$ we turn this in equality of equivalence classes. This is a typical example about the use of equivalence relations, they allow you to define a new type of equality, so that objects that are not equal have associated equivalence classes that are equal.

Theorem 3.34. *Let $\langle A, \leqslant \rangle$ be a pre-ordered set then we have*

1. $\sim \subseteq A \times A$ *defined by* $\sim = \{(x, y) \in A | x \leqslant y \wedge y \leqslant x\}$ *is a equivalence relation*
2. *Define* $\preccurlyeq \subseteq (A/\sim) \times (A/\sim)$ *by*

$$\preccurlyeq = \{(x, y) \in (A/\sim) \times (A/\sim) | \exists x' \in \sim[x] \text{ and } \exists y' \in \sim[y] \text{ such that } x' \leqslant y'\}$$

   then $\preccurlyeq$ *is a order relation in* $A/\sim$. *So* $\langle A/\sim, \preccurlyeq \rangle$ *is a partial ordered set*
3. $\forall x, y \in A$ *we have* $x \leqslant y \Leftrightarrow \sim[x] \preccurlyeq \sim[y]$

Proof.

1. To prove that $\sim$ is a equivalence relation note:

   reflectivity. If $x \in A$ then $x \leqslant x$ proving that $x \sim x$

   symmetric. If $x \sim y$ then $x \leqslant y \wedge y \leqslant x \Rightarrow y \leqslant x \wedge x \leqslant y$ so that $y \sim x$

   transitive. If $x \sim y$ and $y \sim z$ then $x \leqslant y \wedge y \leqslant x \wedge y \leqslant z \wedge z \leqslant y$ so that $x \leqslant z$ and $z \leqslant x$ or $x \sim z$

2. To prove that $\preccurlyeq$ is a order relation we must prove reflectivity, symmetry and transitivity:

   reflexivity. Take $\sim[x]$ then as $x \leqslant x$ there exists a $u \in \sim[x]$ and $v \in \sim[x]$ such that $u \leqslant v$ [just take $u = x = v$] so that

$$\sim[x] \preccurlyeq \sim[x]$$

   symmetry. Let $\sim[x] \leqslant \sim[y]$ and $\sim[y] \leqslant \sim[x]$ then $\exists x', x'' \in \sim[x]$, $\exists y'y'' \in \sim[y]$ such that

$$x' \leqslant y' \wedge y'' \leqslant x''$$

   From $\exists x', x'' \in \sim[x]$, $\exists y'y'' \in \sim[y]$ we have

$$x' \leqslant x \wedge x \leqslant x' \wedge x'' \leqslant x \wedge x \leqslant x'' \wedge y' \leqslant y \wedge y \leqslant y' \wedge y'' \leqslant y \wedge y \leqslant y''$$

   From $x \leqslant x'$ and $x' \leqslant y'$ we have $x \leqslant y'$, as $y' \leqslant y$ we have

$$x \leqslant y$$

   From $y \leqslant y''$ and $y'' \leqslant x''$ we have $y \leqslant x''$, as $x'' \leqslant x$ it follows that

$$y \leqslant x$$

   Finally from $x \leqslant y$ and $y \leqslant x$ we have that $x \sim y$ which by [theorem: 3.11] gives

$$\sim[x] = \sim[y]$$

   transitivity. Assume that $\sim[x] \preccurlyeq \sim[y]$ and $\sim[y] \preccurlyeq \sim[z]$ then we have the existence of $x' \in \sim[x]$, $y', y'' \in \sim[y]$ and $z' \in \sim[z]$ such that

$$x' \leqslant y' \wedge y'' \leqslant z'$$

   From $x' \in \sim[x]$, $y', y'' \in \sim[y]$ and $z' \in \sim[z]$ it follows that

$$x' \leqslant x \wedge x \leqslant x' \wedge y' \leqslant y \wedge y \leqslant y' \wedge y'' \leqslant y \wedge y \leqslant y'' \wedge z' \leqslant z \wedge z \leqslant z'$$

   From $x \leqslant x'$ and $x' \leqslant y'$ we have $x \leqslant y'$, as $y' \leqslant y$ we have $x \leqslant y$, as $y \leqslant y''$ it follows that $x \leqslant y''$, from $y'' \leqslant z'$ we have that $x \leqslant z'$ and finally from $z' \leqslant z$ it follows that $x \leqslant z$. Hence

$$\sim[x] \preccurlyeq \sim[z]$$

3.

⇒. If $x \leqslant y$ then as $x \in \sim[x]$ and $y \in \sim[y]$ we have $\sim[x] \preccurlyeq \sim[y]$

⇐. If $\sim[x] \preccurlyeq \sim[y]$ then $\exists x' \in \sim[x]$ and $\exists y' \in \sim[y]$ such that

$$x' \leqslant y'$$

From $x' \in \sim[x]$ and $y' \in \sim[y]$ we have that

$$x' \leqslant x \wedge x \leqslant x' \wedge y' \leqslant y \wedge y \leqslant y'$$

From $x \leqslant x'$ and $x' \leqslant y'$ it follows that $x \leqslant y'$ and as $y' \leqslant y$ it follows that

$$x \leqslant y$$

□

Given a partial ordered class then we can induce the order on a sub-class making the sub-class also a partial ordered class.

Theorem 3.35. *If $\langle A, \leqslant \rangle$ is a partial ordered sets and $B \subseteq A$ then $\leqslant_{|B}$ defined by*

$$\leqslant_{|B} = \leqslant \bigcap B \times B = B$$

is a order relation in B making $\langle B, \leqslant_{|B} \rangle$ a partial ordered set.

Proof.

reflectivity. If $x \in B$ then $x \leqslant x$ or $(x, x) \in \leqslant \underset{x \in B}{\Rightarrow} (x, x) \in \leqslant \bigcap (B \times B)$ hence $x \leqslant_{|B} y$

symmetry. If $x \leqslant_{|B} y \wedge y \leqslant_{|B} x \Rightarrow x \leqslant y \wedge y \leqslant x \Rightarrow x = y$

transitivity. If $x \leqslant_{|B} y \wedge y \leqslant_{|B} z \Rightarrow x \leqslant y \wedge y \leqslant z \Rightarrow x \leqslant z \underset{x, z \in B}{\Rightarrow} x \leqslant_{|B} z$ □

Convention 3.36. To avoid excessive usage notation we write $\langle B, \leqslant \rangle$ instead of $\langle B, \leqslant_{|B} \rangle$

The following shows a technique of defining a partial order on the Cartesian product of partial ordered set.

Theorem 3.37. (Lexical ordering) *Let $\langle A, \leqslant_A \rangle$ and $\langle B, \leqslant_B \rangle$ be partial ordered classes then $\leqslant_{A \times B}$ defined by*

$$\leqslant_{A \times B} = \{((x, y), (u, v)) \in (A \times B) \times (A \times B) | (x \neq u \wedge x \leqslant_A u) \vee (x = y \wedge y \leqslant_B v\}$$

is a order in $A \times B$ making $\langle (A \times B) \times (A \times B), \leqslant_{A \times B} \rangle$ a partial ordered set

Proof.

reflexivity. If $(x, y) \in A \times B$ then $x \leqslant_A x \wedge y \leqslant_B y$ proving that $(x, y) \leqslant_{A \times B} (x, y)$

symmetry. Let $(x, y) \leqslant_{A \times B} (u, v) \wedge (u, v) \leqslant_{A \times B} (x, y)$. If $x \neq u$ we would have $x \leqslant_A u \wedge u \leqslant_A x \Rightarrow x = u$ a contradiction. So we must have that $x = u$ but then $y \leqslant_B v \wedge v \leqslant_{|B} y \Rightarrow y = v$ proving that

$$(x, y) = (u, v)$$

transitivity. Let $(x, y) \leqslant_{A \times B} (u, v) \wedge (u, v) \leqslant_{A \times B} (r, s)$ then we have to consider the following cases:

***x* = *u*.** Then $y \leqslant_B v$ and we have the following possibilities

***u* = *r*.** Then $v \leqslant_B s$ so that $y \leqslant_B s$ which as $x = r$ proves that

$$(x, y) \leqslant_{A \times B} (r, s)$$

***u* ≠ *r*.** Then $u \leqslant_A r \underset{x = u}{\Rightarrow} x \leqslant_A r$ which as $x \neq r$ proves that

$$(x, y) \leqslant_{A \times B} (r, s)$$

***x* ≠ *u*.** Then $x \leqslant_A u$ and we have the following possibilities

***u* = *r*.** Then $x \leqslant_A u \underset{u = r}{\Rightarrow} x \leqslant_A r$ and $x \neq r$ so that

$$(x, y) \leqslant_{A \times B} (r, s)$$

$\boldsymbol{u \neq r}$**.** Then $u \leqslant_A r$ so that $x \leqslant_A r$. If $x = r$ then we would have $x \leqslant_A u \wedge u \leqslant_A x$ giving $x = u$ contradicting $x \neq u$. So we must have $x \neq r$ which as $x \leqslant_A r$ gives

$$(x, y) \leqslant_{A \times B} (r, s)$$

□

Definition 3.38. *Let $\langle A, \leqslant \rangle$ be a partial ordered class then $x, y \in A$ are **comparable** if $x \leqslant y$ or $y \leqslant x$*

Theorem 3.39. *Let $\langle A, \leqslant \rangle$ be a partial ordered class and $x, y \in A$ comparable elements then we have either $x \leqslant y$ or $y < x$*

Proof. As x, y are comparable then we have $x \leqslant y \vee y \leqslant x$, consider the following cases:

$\boldsymbol{x \leqslant y}$**.** hen $x \leqslant y$

$\boldsymbol{\neg(x \leqslant y)}$**.** then we must have $y \leqslant x$. If $x = y$ then as $x \leqslant x$ we have $x \leqslant y$ contradicting $\neg(x \leqslant y)$ so that $x \neq y$ proving $y < x$.

Hence we have

$$x \leqslant y \vee y < x$$

□

Definition 3.40. *A pre-ordered class $\langle A, \leqslant \rangle$ is a **totally ordered class** iff*

$$\forall x, y \in A \text{ we have } x \leqslant y \vee y \leqslant x$$

*In other words $\langle A, \leqslant \rangle$ is a **totally ordered class** if every pair of elements are comparable. Other names used in the literature are **fully ordered class** or **linear ordered class**.*

Definition 3.41. (chain) *Let $\langle A, \leqslant \rangle$ be a partial ordered class and $C \subseteq A$ then C is called a **chain** if $\forall x, y \in C$ we have that $x \leqslant y$ or $y \leqslant x$.*

Example 3.42. Let $\langle A, \leqslant \rangle$ be a partial ordered class then $\varnothing$ is a chain

Proof. The condition $\forall x, y \in \varnothing$ we have that x, y are comparable is satisfied vacuously. □

Theorem 3.43. *Let $\langle A, \leqslant \rangle$ be a partial ordered class and $B \subseteq A$ a chain then $\langle B, \leqslant_{|B} \rangle$ is a totally ordered class*

Proof. Using [theorem: 3.35] we have that $\langle B, \leqslant_{|B} \rangle$ is a partial ordered class. Let $x, y \in B$ then as B is a chain we have that $\forall x, y \in B$ $x \leqslant y \vee y \leqslant x$ or using the definition of $\leqslant_{|B}$ that $x \leqslant_{|B} y \vee y \leqslant_{|B} x$. □

Theorem 3.44. *Let $\langle A, \leqslant \rangle$ be a totally ordered class and $B \subseteq A$ then B is a chain [hence by [theorem: 3.43] $\langle B, \leqslant_{|B} \rangle$ is a totally ordered class]*

Proof. If $x, y \in B$ then $x, y \in A$ and as A is totally ordered we have $x \leqslant y \vee y \leqslant x$ so B is a chain □

Theorem 3.45. *Let $\langle A, \leqslant_A \rangle$ and $\langle B, \leqslant_B \rangle$ be totally ordered classes then $\langle A \times B, \leqslant_{A \times B} \rangle$ is a totally ordered class.*

Proof. First $\langle A \times B, \leqslant_{A \times B} \rangle$ is a partially ordered class by [theorem: 3.37]. If $(x, y), (x', y') \in A \times B$ 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 $(x, y) \leqslant (x', y')$
- $\boldsymbol{y' \leqslant y}$**.** 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 3.46. (Initial Segment) *If $\langle A, \leqslant \rangle$ is a partial ordered class, $a \in A$ then a **initial segment of A determined by a** noted as $S_{A,a}$ is defined by*

$$S_{A,a} = \{x \in A | x < a\}$$

We have the following trivial result for initial segments.

Proposition 3.47. *If $\langle A, \leqslant \rangle$ is a partial ordered class and $a, b \in A$ such that $a \leqslant b$ then $S_{A,a} \subseteq S_{A,b}$*

Proof. If $x \in S_{A,a}$ then $x < a \underset{a \leqslant b}{\Rightarrow} x < b$ proving that $x \in S_{A,b}$ □

Theorem 3.48. *If $\langle A, \leqslant \rangle$ is a partial ordered class and P is a initial segment of A and Q is a initial segment of P [using the induced order $\leqslant_{|P}$] then A is a initial segment of A*

Proof. Using the hypothesis there exists $a \in A$ such that $P = \{x \in A | x < a\}$ and a $b \in P$ such that $Q = \{x \in P | x < b\}$. Consider then the initial segment $S_{A,b} = \{x \in A | x < b\}$ of A determined by a then we have

$$\begin{array}{rcl}
x \in S_{A,b} & \Rightarrow & x \in A \wedge x < b \\
 & \underset{b < a \Rightarrow x < b \Rightarrow x < a}{\Rightarrow} & x \in A \wedge x < a \wedge x < b \\
 & \Rightarrow & x \in P \wedge x < b \\
 & \Rightarrow & x \in P \wedge x <_{|P} b \\
 & \Rightarrow & x \in Q \\
x \in Q & \Rightarrow & x \in P \wedge x <_{|P} b \\
 & \Rightarrow & x \in P \wedge x < b \\
 & \underset{P \subseteq A}{\Rightarrow} & x \in A \wedge x < b \\
 & \Rightarrow & x \in S_{A,b}
\end{array}$$

Hence $Q = S_{A,b}$ a initial segment of A □

3.3.2 Order relations and functions

Functions between two partial ordered classes can be classified based on the fact that they preserve or not preserve the order relation. This is expressed in the next definition.

Definition 3.49. *Let $\langle A, \leqslant_A \rangle$, $\langle B, \leqslant_B \rangle$ be partial ordered classes and $f: A \rightarrow B$ a function then:*

1. *$f: \langle A, \leqslant_A \rangle \rightarrow B$ is **increasing** if $\forall x, y \in A$ with $x \leqslant y$ we have $f(x) \leqslant f(y)$. Another name that is used is **a order homeomorphism** [a homeomorphism is a function that preserver a certain operation, in this case the order relation]*
2. *$f: \langle A, \leqslant_A \rangle \rightarrow \langle B, \leqslant_B \rangle$ is **strictly increasing** if $\forall x, y \in A$ with $x < y$ we have $f(x) < f(y)$*
3. *$f: \langle A, \leqslant_A \rangle \rightarrow \langle B, \leqslant_B \rangle$ is **decreasing** if $\forall x, y \in A$ with $x \leqslant y$ we have $f(y) \leqslant f(x)$*
4. *$f: \langle A, \leqslant_A \rangle \rightarrow \langle B, \leqslant_B \rangle$ is **strictly decreasing** if $\forall x, y \in A$ with $x < y$ we have $f(y) < f(x)$*
5. *$f: \langle A, \leqslant_A \rangle \rightarrow \langle B, \leqslant_B \rangle$ is a **order isomorphism** if $\forall x, y \in A$ with $x \leqslant y \Leftrightarrow f(x) \leqslant f(y)$*

Definition 3.50. *Two partial ordered classes $\langle A, \leqslant_A \rangle$ and $\langle B, \leqslant_B \rangle$ are **order isomorphic** noted as $A \cong B$ if there exists order isomorphism between A and B.*

Theorem 3.51. *Let $\langle A, \leqslant_A \rangle$, $\langle B, \leqslant_B \rangle$ be two partial ordered classes, $D \subseteq B$ and*

$$f: \langle A, \leqslant_A \rangle \rightarrow \langle D, (\leqslant_B)_{|D} \rangle \text{ be a order homeomorphism [see theorem: 3.35 for } \langle D, (\leqslant_B)_{|D} \rangle$$

then

$$f: \langle A, \leqslant_A \rangle \rightarrow \langle B, \leqslant_B \rangle \text{ is a order homeomorphism}$$

Proof. The proof is trivial for if $x, y \in A$ with $x \leqslant_A y$ then $f(x) (\leqslant_B)_{|D} f(y) \underset{3.35}{\Rightarrow} f(x) \leqslant_B f(y)$ □

Theorem 3.52. *Let $\langle A, \leqslant_A \rangle$, $\langle B, \leqslant_B \rangle$, $\langle C, \leqslant_C \rangle$ be partial ordered classes, $D \subseteq B$*

1. *If $D \subseteq B$ is equiped with the induced order from $\langle B, \leqslant_B \rangle$ [see theorem: 3.35] and*

$$f: \langle A, \leqslant_A \rangle \rightarrow \langle D, \leqslant_B \rangle \text{ and } g: \langle B, \leqslant_B \rangle \rightarrow \langle C, \leqslant_C \rangle \text{ are order homeomorphisms}$$

then

$$g\circ f:\langle A,\leqslant_A\rangle\rightarrow\langle C,\leqslant_C\rangle \text{ is a order homeomorphism}$$

2. *If $D\subseteq B$ is equiped with the induced order from $\langle B,\leqslant_B\rangle$ [see theorem: 3.35] and*

$$f:\langle A,\leqslant_A\rangle\rightarrow\langle D,\leqslant_B\rangle \text{ and } g:\langle B,\leqslant_B\rangle\rightarrow\langle C,\leqslant_C\rangle \text{ are strictly increasing}$$

then

$$g\circ f:\langle A,\leqslant_A\rangle\rightarrow\langle C,\leqslant_C\rangle \text{ is stritly increasing}$$

3. *If $D\subseteq B$ is equiped with the induced order from $\langle B,\leqslant_B\rangle$ [see theorem: 3.35] and*

$$f:\langle A,\leqslant_A\rangle\rightarrow\langle D,\leqslant_B\rangle \text{ and } g:\langle B,\leqslant_B\rangle\rightarrow\langle C,\leqslant_C\rangle \text{ are order isomorphism}$$

then

$$g\circ f:\langle A,\leqslant_A\rangle\rightarrow\langle g(f(A)),\leqslant_C\rangle \text{ is a order isomorphism}$$

or as $D \underset{f:a\rightarrow D \text{ is bijective}}{=} f(A)$

$$g\circ f:\langle A,\leqslant_A\rangle\rightarrow\langle g(D),\leqslant_C\rangle \text{ is a order isomorphism}$$

Proof.

1. Let $x,y\in A$ with $x\leqslant_A y$ then $f(x)\leqslant f_B(y)$ hence $(g\circ f)(x)=g(f(x))\leqslant_C g(f(y))=(g\circ f)()$.
2. Let $x,y\in A$ with $x<_A y$ then $f(x)<_B f(y)$ hence $(g\circ f)(x)=g(f(x))<_C g(f(y))=(g\circ f)()$.
3. Using [theorem: 2.74] we have that $g\circ f:A\rightarrow g(D)=g(f(A))$ is a bijection. Let $x,y\in A$. If $x\leqslant_A y$ then $f(x)\leqslant_B f(y)$ hence $(g\circ f)(x)=g(f(x))\leqslant_C g(f(y))=(g\circ f)(y)$. Also if $(g\circ f)(x)\leqslant_C(g\circ f)(y)$ then $g(f(x))\leqslant_C g(f(y))$ so that $f(x)\leqslant_B f(y)$, giving $x\leqslant_A y$. □

Theorem 3.53. *If $\langle A,\leqslant_A\rangle$ and $\langle B,\leqslant_B\rangle$ are partially ordered classes and*

$$f:\langle A,\leqslant_A\rangle\rightarrow\langle B,\leqslant_B\rangle \text{ a order isomorphism}$$

then $\forall x,y\in A$ we have

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

Proof.

⇒. If $x<_A y$ then $x\neq y$ and $x\leqslant_A y\Rightarrow f(x)\leqslant_B f(y)$. Assume that $f(x)=f(y)$ then as f is a bijection we would have $x=y$ contradicting $x\neq y$. So we must have that $f(x)\neq f(y)$ hence

$$f(x)<_B f(y)$$

⇐. As $f(x)<_B f(y)$ we have that $f(x)\neq f(y)$ so that we must have $x\neq y$. Further as f is a isomorphism we have $x\leqslant_A y$. So

$$x<_A y$$

□

Theorem 3.54. *If $\langle A,\leqslant_A\rangle$ and $\langle B,\leqslant_B\rangle$ are partially ordered classes and $f:A\rightarrow B$ a bijection then*

$$f:\langle A,\leqslant_A\rangle\rightarrow\langle B,\leqslant_B\rangle \text{ is a order isomorphism} \Leftrightarrow f:\langle A,\leqslant_A\rangle\rightarrow\langle B,\leqslant_B\rangle \text{ and } f^{-1}:\langle B,\leqslant_B\rangle\rightarrow\langle A,\leqslant_A\rangle \text{ are increasing functions}$$

Proof. As $f:A\rightarrow B$ is a bijection we have by [theorems: 2.67, 2.72] that $f^{=1}:B\rightarrow A$ is a bijection.

⇒. As $f:\langle A,\leqslant_A\rangle\rightarrow\langle B,\leqslant_B\rangle$ is a isomorphism we have that $\forall x,y\in A$ with $x\leqslant_A y\Rightarrow f(x)\leqslant f(b)$ hence $f:A\rightarrow B$ is increasing. If x, $y\in B$ with $x\leqslant_B y$ then

$$f(f^{-1}(x))=(f\circ f^{-1})(x)\underset{[\text{theorem: } 2.69}{=}x\leqslant_B y=(f\circ f^{-1})(y)=f(f^{-1}(y))$$

which as f is a isomorphism proves that $f^{-1}(x)\leqslant_A f^{-1}(y)$, hence f^{-1} is increasing.

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

Theorem 3.55. *If $\langle A,\leqslant_A\rangle,\langle C,\leqslant_C\rangle$ and $\langle B,\leqslant_B\rangle$ are partially ordered classes then*

1. $1_A:\langle A,\leqslant_A\rangle\rightarrow\langle A,\leqslant_A\rangle$ *is a order isomorphism*
2. *If $f:\langle A,\leqslant_A\rangle\rightarrow\langle B,\leqslant_B\rangle$ is a order isomorphism then $f^{-1}:\langle B,\leqslant_B\rangle\rightarrow\langle A,\leqslant_A\rangle$ is a order isomorphism*
3. *If $f:\langle A,\leqslant_A\rangle\rightarrow B$ and $g:\langle B,\leqslant_B\rangle\rightarrow\langle C,\leqslant_C\rangle$ are order isomorphism's then*

$$g\circ f:\langle A,\leqslant_A\rangle\rightarrow\langle C,\leqslant_C\rangle \text{ is a order isomorphism}$$

Proof.

1. By 2.47 we have that $\mathrm{Id}_A: A\rightarrow A$ is a bijection then, as $x=I_A(x)$ and $y=\mathrm{Id}_A(y)$, we have $x\leqslant y\Leftrightarrow\mathrm{Id}_A(x)\leqslant\mathrm{Id}_A(y)$.
2. If $f:A\rightarrow B$ is a isomorphism then by [theorem: 2.72] we have that $f^{-1}:B\rightarrow A$ is a bijection. By the previous theorem [theorem: 3.54] we have that f^{-1} is increasing. Further as by 2.73 $f=(f^{-1})^{-1}$ and by [theorem: 3.54] f is increasing it follows that $(f^{-1})^{-1}$ is increasing. Using then [theorem: 3.54] it follows that f^{-1} is a isomorphism.
3. This follows from [theorem: 3.52] □

Theorem 3.56. *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: 3.55] □

Theorem 3.57. *Let $\langle A,\leqslant_A\rangle$. be a totally ordered class and $\langle B,\leqslant_B\rangle$ is a partially ordered class then a bijective and increasing function $f:\langle A,\leqslant_A\rangle\rightarrow\langle B,\leqslant_B\rangle$ 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 [theorem: 3.38] we have the following exclusive cases

1. $x\leqslant_A y$ in this case our 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. □

3.3.3 Min, max, supremum and infinum

Definition 3.58. *Let $\langle X,\leqslant\rangle$ be a pre-ordered class and $A\subseteq X$ then*

1. *m is a **maximal element** of A iff $m\in A$ and if $\forall x\in A$ with $m\leqslant x$ we have $x=m$*
2. *m is a **minimal element** of A iff $m\in A$ and if $\forall x\in A$ with $x\leqslant m$ we have $x=m$*

Definition 3.59. *If $\langle X,\leqslant\rangle$ is a partial ordered class and $A\subseteq X$ then*

1. *m is the **greatest element** of A iff $m\in A$ and $\forall x\in A$ we have $x\leqslant m$*
2. *m is the **least element** of A iff $m\in A$ and $\forall x\in A$ we have $m\leqslant x$*

Note 3.60. There is a subtle difference between the definition of a maximal (minimal) element and the greatest (least) element. If m is the greatest (least) element of A then every element in A is comparable with m, which is not the case if m is a maximal (minimal) element of A.

Note 3.61. The empty set $\varnothing$ can not have a maximal, minimal element, greatest element or least element.

Theorem 3.62. *If $\langle X,\leqslant\rangle$ is a partial ordered class and $A\subseteq X$ then*

1. *If m,m' are greatest elements of A then $m=m'$*
2. *If m,m' are least elements of A then $m=m'$*

The unique greatest element of A (if it exist) is called the maximum of A and noted as $\max(A)$, the unique least element of A (if it exist) is called the minimum of A and noted as $\min(A)$

Proof.

1. If m, m' are greatest elements of A then as $m, m' \in A$ we have $m \leqslant m' \wedge m' \leqslant m$ so that $m = m'$.
2. If m, m' are least elements of A then as $m, m' \in A$ we have $m \leqslant m' \wedge m' \leqslant m$ so that $m = m'$. □

Example 3.63. Let $\langle X, \leqslant \rangle$ be a totally ordered set, $x, y \in X$ then $\max(\{x, y\})$ and $\min(\{x, y\})$ exist.

Proof. As $\langle X, \leqslant \rangle$ is a totally ordered set we have either:

$\boldsymbol{x \leqslant y}$**.** Then as $x, y \in X$, $x \leqslant x$, $y \leqslant y$ and $x \leqslant y$ we have $\min(\{x, y\}) = x$ and $\max(\{x, y\}) = y$.

$\boldsymbol{y \leqslant x}$**.** Then as $x, y \in X$, $x \leqslant x$, $y \leqslant y$ and $y \leqslant x$ we have $\min(\{x, y\}) = y$ and $\max(\{x, y\}) = x$. □

Theorem 3.64. *Let $\langle X, \leqslant \rangle$ be a totally ordered set and $A \subseteq X$, $B \subseteq X$ then we have:*

1. *If $\max(A)$ and $\max(B)$ exist then $\max(A \bigcup B)$ exist and*

$$\max\left(A \bigcup B\right) = \max(\{\max(A), \max(B)\})$$

2. *If $\min(A)$ and $\min(B)$ exist then $\min(A \bigcup B)$ exist and*

$$\min\left(A \bigcup B\right) = \min(\{\min(A), \min(B)\})$$

Proof.

1. We have the following possibilities for $\max(A)$ and $\max(B)$:

   $\boldsymbol{\max(A) \leqslant \max(B)}$**.** Then $\max(B) \in B \subseteq A \bigcup B$ and $\forall x \in A \bigcup B$ we have either:

   $\boldsymbol{x \in A}$**.** Then $x \leqslant \max(A) \leqslant \max(B)$ so that $x \leqslant \max(B)$

   $\boldsymbol{x \in B}$**.** Then $x \leqslant \max(B)$

   Hence $\forall x \in A \bigcup B$ we have that $x \leqslant \max(B)$ proving that

$$\max\left(A \bigcup B\right) = \max(B) \underset{[\text{example: } 3.63]}{=} \max(\{\max(A), \max(B)\})$$

   $\boldsymbol{\max(B) \leqslant \max(A)}$**.** Then $\max(A) \in A \subseteq A \bigcup B$ and $\forall x \in A \bigcup B$ we have either:

   $\boldsymbol{x \in A}$**.** Then $x \leqslant \max(A)$

   $\boldsymbol{x \in B}$**.** Then $x \leqslant \max(B) \leqslant \max(A)$

   Hence $\forall x \in A \bigcup B$ we have that $x \leqslant \max(A)$ proving that

$$\max\left(A \bigcup B\right) = \max(A) \underset{[\text{example: } 3.63]}{=} \max(\{\max(A), \max(B)\})$$

2. We have the following possibilities for $\min(A)$ and $\min(B)$:

   $\boldsymbol{\min(A) \leqslant \min(B)}$**.** Then $\min(A) \in A \subseteq A \bigcup B$ and $\forall x \in A \bigcup B$ we have either:

   $\boldsymbol{x \in A}$**.** Then $\min(A) \leqslant x$

   $\boldsymbol{x \in B}$**.** Then $\min(A) \leqslant \min(B) \leqslant x$ so that $\min(A) \leqslant x$

   Hence $\forall x \in A \bigcup B$ we have that $\min(A) \leqslant x$ proving that

$$\min\left(A \bigcup B\right) = \min(A) \underset{[\text{example: } 3.63]}{=} \min(\{\min(A), \min(B)\})$$

   $\boldsymbol{\min(B) \leqslant \min(A)}$**.** Then $\min(B) \in B \subseteq A \bigcup B$ and $\forall x \in A \bigcup B$ we have either:

   $\boldsymbol{x \in A}$**.** Then $\min(B) \leqslant \min(A) \leqslant x$ so that $\min(B) \leqslant x$

   $\boldsymbol{x \in B}$**.** Then $\min(B) \leqslant x$

   Hence $\forall x \in A \bigcup B$ we have that $\min(B) \leqslant x$ proving that

$$\min\left(A \bigcup B\right) = \min(B) \underset{[\text{example: } 3.63]}{=} \min(\{\min(A), \min(B)\})$$

□

Theorem 3.65. *If $\langle X, \leqslant\rangle$ is a partial ordered class and $A \subseteq X$ such that $\min(A)$ and $\max(A)$ exist then $\min(A) \leqslant \max(A)$*

Proof. As $\min(A) \in A$ we have by definition that $\min(A) \leqslant \max(A)$. □

Theorem 3.66. *Let $\langle X, \leqslant\rangle$ be a partial ordered class, $A \subseteq X$, $B \subseteq X$ then*

1. *If $\max(A)$ and $\max(B)$ exist and $\forall x \in A\ \exists y \in B$ such that $x \leqslant y$ then $\max(A) \leqslant \max(B)$*
2. *If $\min(A)$ and $\min(B)$ exist $\forall x \in B\ \exists y \in A$ such that $y \leqslant x$ then then $\min(A) \leqslant \min(B)$*

Proof.

1. As $\max(A) \in A$ there exist a $y \in B$ such that $\max(A) \leqslant y$, as $y \leqslant \max(B)$ we have
$$\max(A) \leqslant \max(B)$$
2. As $\min(B) \in A$ there exist a $y \in A$ such that $y \leqslant \min(B)$, as $\min(A) \leqslant y$ we have
$$\min(A) \leqslant \max(A)$$
□

Definition 3.67. *If $\langle X, \leqslant\rangle$ is a partial ordered class and $A \subseteq X$ then*

1. *$u \in X$ is a **upper bound** of A if $\forall a \in A\ a \leqslant u$.*
2. *A is **bounded above** if it has a upper bound.*
3. *$l \in X$ is a **lower bound** of A if $\forall x \in A\ l \leqslant a$*
4. *A is **bounded below** if it has a lower bound.*
5. *$\upsilon(A) = \{x \in X | x$ is a upper bound of $A\}$ [the class of upper bound of A]*
6. *$\lambda(A) = \{x \in X | x$ is a lower bound of $A\}$ [the class of lower bounds of A]*

Example 3.68. If $\langle X, \leqslant\rangle$ then $\upsilon(\varnothing) = X$ and $\lambda(\varnothing) = X$

Proof. Let $x \in X$ then as $\forall a \in \varnothing\ a \leqslant x$ [or $x \leqslant a$] is vacuously satisfied $X \subseteq \upsilon(A)$ and $X \subseteq \lambda(A)$, which as $\upsilon(X) \subseteq X$ and $\lambda(X) \subseteq X$ proves $\upsilon(A) = X = \lambda(A)$. □

Definition 3.69. *If $\langle X, \leqslant\rangle$ is a partial ordered class and $A \subseteq X$ then*

1. *If $\min(\upsilon(A))$ exists then $\min(\upsilon(A))$ is called the supremum of A and noted as $\sup(A)$.*
2. *If $\max(\lambda(A))$ exists then $\max(\lambda(A))$ is called the infinum of A and noted as $\inf(A)$*

In other words if $\upsilon(A)$ has a least element then the supremum of A is this unique, by [theorem: 3.62], element. So $\sup(A)$ is the least upper bound of A [if it exist] and it is itself a upper bound. If $\lambda(A)$ has a least element then the infinum of A is this unique, by [theorem: 3.62], element. So $\inf(A)$ is the greatest lower bound [if it exist] and it is itself a lower bound.

Example 3.70. Let $\mathcal{A}$ be a class of classes and $\langle \mathcal{A}, \leqslant\rangle$ the partial class where
$$\leqslant = \{(x, y) \in \mathcal{A} \times \mathcal{A} | x \subseteq y\}$$
[see example: 3.33] and $\mathcal{B} \subseteq \mathcal{A}$ we have that

1. If $\bigcap \mathcal{B} \in \mathcal{A}$ then $\inf(\mathcal{B})$ exist and $\inf(\mathcal{B}) = \bigcap \mathcal{B}$
2. If $\bigcup \mathcal{B} \in \mathcal{A}$ then $\sup(\mathcal{B})$ exist and $\sup(\mathcal{B}) = \bigcup \mathcal{B}$

Proof.

1. If $B \in \mathcal{B}$ then by [theorem: 1.60] $\bigcap \mathcal{B} \subseteq B \Rightarrow \bigcap \mathcal{B} \leqslant B$ so that $\bigcap \mathcal{B} \in \lambda(\mathcal{B})$. Now if $C \in \lambda(\mathcal{B})$ then $\forall B \in \mathcal{B}$ we have that $C \leqslant B \Rightarrow C \subseteq B$, so that by [theorem: 1.60] we have $C \subseteq \bigcap \mathcal{B} \Rightarrow C \leqslant \bigcap \mathcal{B}$ so that $\bigcap \mathcal{B}$ is the greatest element of $\lambda(\mathcal{B})$ proving that $\inf(\mathcal{B})$ exists and $\inf(\mathcal{B}) = \bigcap \mathcal{B}$.
2. If $B \in \mathcal{B}$ then by [theorem: 1.60] $B \subseteq \bigcup \mathcal{B} \Rightarrow B \leqslant \bigcup \mathcal{B}$ so that $\bigcup \mathcal{B} \in \upsilon(\mathcal{B})$. Now if $C \in \upsilon(\mathcal{B})$ then $\forall B \in \mathcal{B}$ we have that $B \leqslant C \Rightarrow B \subseteq C$, so that by [theorem: 1.60] we have $\bigcup \mathcal{B} \subseteq C \Rightarrow \bigcup \mathcal{B} \leqslant C$ so that $\bigcup \mathcal{B}$ is the lowest element of $\upsilon(\mathcal{B})$ proving that $\sup(\mathcal{B})$ exists and $\sup(\mathcal{B}) = \bigcup \mathcal{B}$. □

The following theorem will be used a lot of time when dealing with supremums and infinums.

Theorem 3.71. *Let $\langle X, \leqslant \rangle$ be a totally ordered set and $A \subseteq X$ then*

1. *If* $\sup(A)$ *exists then* $\forall x \in X$ *with* $x < \sup(A)$ *there* $\exists a \in A$ *such that* $x < a \wedge a \leqslant \sup(A)$
2. *If* $\inf(A)$ *exist then* $\forall x \in X$ *with* $\inf(A) < x$ *there* $\exists a \in A$ *such that* $\inf(A) \leqslant a \wedge a < x$

Proof. First as $\langle X, \leqslant \rangle$ is totally ordered we have $\forall x, y \in X$ that x, y are comparable, hence by [theorem: 3.39], we have $x \leqslant y \wedge y < x/$

1. Let $x \in X$ such that $x < \sup(A)$. Assume that $\forall a \in A$ we have $\neg(x < a)$ so that $a \leqslant x$, so x is a upper bound of A, hence $x \in \upsilon(A)$, so that $\sup(A) = \min(\upsilon(A)) \leqslant x$, which, as $x < \sup(A)$, leads to the contradiction $x < x$. So we must have that $\exists a \in A$ such that $x < a$, further as $\sup(A)$ is a upper bound we have that $a \leqslant \sup(A)$. So
$$\exists a \in A \;\; x < a \wedge a \leqslant \sup(A)$$
2. Let $x \in X$ such that $\inf(A) < x$. Assume that $\forall a \in A$ we have $\neg(a < x)$ so that $x \leqslant a$, so x is a lower bound of A, hence $x \in \lambda(A)$, so that $x \leqslant \max(\lambda(A)) = \inf(A)$, which, as $\inf(A) < x$, leads to the contradiction $x < x$. So we must have that $\exists a \in A$ such that $a \leqslant x$, further as $\inf(A)$ is a lower bound we have we have that $\inf(A) \leqslant a$. So
$$\exists a \in A \; \inf(A) \leqslant a \wedge a < x$$ □

Lemma 3.72. *If $\langle X, \leqslant \rangle$ is a partially ordered class and $A \subseteq X, B \subseteq X$ with $A \subseteq B$ then*

1. *If* $\max(A)$ *and* $\max(B)$ *exist then* $\max(A) \leqslant \max(B)$
2. *If* $\min(A)$ *and* $\min(B)$ *exists then* $\min(B) \leqslant \min(A)$

Proof.

1. As $\max(A) \in A$ and $A \subseteq B$ we have that $\max(A) \in B$ so that $\max(A) \leqslant \max(B)$
2. As $\min(A) \in A$ and $A \subseteq B$ we have that $\min(A) \in B$ so that $\min(B) \leqslant \min(A)$ □

Lemma 3.73. *If $\langle X, \leqslant \rangle$ is a partially ordered class and $A \subseteq X, B \subseteq X$ with $A \subseteq B$ then*

1. $\upsilon(B) \subseteq \upsilon(A)$
2. $\lambda(B) \subseteq \lambda(A)$

Proof.

1. Let $x \in \upsilon(B)$ then $\forall a \in A$ we have, as $A \subseteq B$ that $a \in B$ hence $a \leqslant x$ proving that x is a upper bound of A or $x \in \upsilon(A)$.
2. Let $x \in \lambda(B)$ then $\forall a \in A$ we have as $A \subseteq B$ hat $a \in B$ hence $x \leqslant a$ proving that x is a lower bound of A or $x \in \lambda(A)$. □

Theorem 3.74. *Let $\langle X, \leqslant \rangle$ be a partial ordered class and $A \subseteq X$, $B \subseteq Y$ such that $A \subseteq B$ then*

1. *If* $\sup(A)$ *and* $\sup(B)$ *exist then* $\sup(A) \leqslant \sup(B)$
2. *If* $\inf(A)$ *and* $\inf(B)$ *exist then* $\inf(B) \leqslant \inf(A)$

Proof.

1. Using [lemma: 3.73] we have that $\upsilon(B) \subseteq \upsilon(A)$ so that by [lemma: 3.72]
$$\sup(A) = \min(\upsilon(A)) \leqslant \min(\upsilon(B)) = \sup(B)$$
2. Using [lemma: 3.73] we have that $\lambda(B) \subseteq \lambda(A)$ so that by [lemma: 3.72]
$$\inf(B) = \max(\lambda(B)) \leqslant \max(\lambda(A)) = \inf(A)$$ □

Theorem 3.75. *Let $\langle X, \leqslant \rangle$ be a partial ordered class and $A \subseteq X, B \subseteq X$ then*

1. *If* $\sup(A)$, $\sup(B)$ *exists and* $\forall a \in A$ $\exists b \in B$ *such that* $a \leqslant b$ *then* $\sup(A) \leqslant \sup(B)$
2. *If* $\inf(A)$ *and* $\inf(B)$ *exist and* $\forall a \in A$ $\exists b \in B$ *such that* $b \leqslant a$ *then* $\inf(B) \leqslant \inf(A)$

Proof.

1. Let $a \in A$ then $\exists b \in B$ such that $a \leqslant b$, as $b \leqslant \sup(B)$ it follows that $a \leqslant \sup(B)$. Hence $\sup(B) \in \upsilon(A)$. So $\sup(A) = \min(\upsilon(A)) \leqslant \sup(A)$, hence
$$\sup(A) \leqslant \sup(B)$$

2. Let $a \in A$ then $\exists b \in B$ such that $b \leqslant a$, as $\inf(B) \leqslant b$ it follows that $\inf(B) \leqslant a$. Hence $\inf(B) \in \lambda(A)$, So $\inf(B) \leqslant \max(\lambda(A)) = \inf(A)$, hence
$$\inf(B) \leqslant \inf(A)$$
□

We have by definition that $\sup(A)$ exists if $\min(\upsilon(A))$ exists and $\inf(A)$ exist if $\max(\lambda(A))$ exist. The following theorem shows that there is a weaker condition for the existence of $\sup(A)$ and $\inf(A)$.

Theorem 3.76. *Let $\langle X, \leqslant\rangle$ be a partial ordered class and $A \subseteq X$ then*

1. *If $\lambda(A)$ has a supremum then A has a infinum and $\sup(\lambda(A)) = \inf(A)$*
2. *If $\upsilon(A)$ has a infinum then A has a supremum and $\inf(\upsilon(A)) = \sup(A)$*

Proof.

1. If $a \in A$ then $\forall y \in \lambda(A)$ we have $y \leqslant a$ so that $a \in \upsilon(\lambda(A))$. As $\sup(\lambda(A)) = \min(\upsilon(\lambda(A)))$ we have that $\sup(\lambda(A)) \leqslant a$. As $a \in A$ was arbitrary chosen we have that
$$\sup(\lambda(A)) \in \lambda(A) \tag{3.8}$$
If $x \in \lambda(A)$, then, as $\sup(\lambda(A))$ is a upper bound of $\lambda(A)$, we have $x \leqslant \sup(\lambda(A))$. So
$$\forall x \in \lambda(A) \text{ we have } x \leqslant \sup(\lambda(A)) \tag{3.9}$$
Using [eq: 3.8] and [eq: 3.9] it follows that $\sup(\lambda(A)) = \max(\lambda(A)) = \inf(A)$ or
$$\sup(\lambda(A)) = \inf(A)$$

2. If $a \in A$ then $\forall y \in \upsilon(A)$ we have $a \leqslant y$ so that $a \in \lambda(\upsilon(A))$. As $\inf(\upsilon(A)) = \max(\lambda(\upsilon(A)))$ we have that $a \leqslant \inf(\upsilon(A))$. As $a \in A$ was arbitrary chosen we have that
$$\inf(\upsilon(A)) \in \upsilon(A) \tag{3.10}$$
If $x \in \upsilon(A)$, then, as $\inf(\upsilon(A))$ is a lower bound of $\upsilon(A)$, we have $\inf(\upsilon(A)) \leqslant x$. So we have that
$$\forall x \in \upsilon(A) \text{ we have that } \inf(\upsilon(A)) \leqslant x \tag{3.11}$$
Using [eq: 3.10] and [eq: 3.11] it follows that $\inf(\upsilon(A)) = \min(\upsilon(A)) = \sup(A)$ or
$$\inf(\upsilon(A)) = \sup(A)$$
□

In general it is not guaranteed that $\sup(A)$ or $\inf(A)$ exists. However there exists partial order classes that guarantees the existence of a supremum for non empty sub-classes that are bounded above.

Definition 3.77. (Conditional Completeness) *A partial ordered class $\langle X, \leqslant\rangle$ is **conditional complete** if every non empty sub-class of A that is bounded above has a supremum.*

The next theorem shows that conditional completeness can also be defined based on bounded below and infinum.

Theorem 3.78. *If $\langle A, \leqslant\rangle$ is a partial ordered class then the following are equivalent*

1. *Every non empty sub-class of X that is bounded above has a supremum [$\langle X, \leqslant\rangle$ is conditional complete]*
2. *Every non empty sub-class of X that is bounded below has a infinum*

Proof.

1 ⇒ 2. Let $A \subseteq X$ a non empty sub-class that is bounded below. As $A \neq \varnothing$ there exists a $a \in A$, further by definition of $\lambda(A)$ we have $\forall y \in \lambda(A)$ that $y \leqslant a$ so $\lambda(A)$ is bounded above. As A is bounded below we have that $\lambda(A) \neq \varnothing$. So by the hypothesis $\sup(\lambda(A))$ exist. Applying then [theorem: 3.76] proves
$$\inf(A) \text{ exist}$$

2 ⇒ 1. Let $A \subseteq X$ a non empty sub-class that is bounded above. As $A \neq \varnothing$ there exists a $a \in A$, further by definition of $\upsilon(A)$ we have $\forall y \in \upsilon(A)$ that $a \leqslant y$ so $\upsilon(A)$ is bounded below. As A is bounded above we have that $\upsilon(A) \neq \varnothing$. So by the hypothesis $\inf(\upsilon(A))$ exist. Applying then [theorem: 3.76] proves
$$\sup(A) \text{ exist}$$
□

Next we show that a order isomorphism preserves the concepts of greatest element, least element, upper bound, lower bound, supremum and infinum.

Lemma 3.79. *Let $\langle X, \leqslant_X \rangle$, $\langle Y, \leqslant_Y \rangle$ be partial ordered classes, $f: \langle X, \leqslant_X \rangle \rightarrow \langle Y, \leqslant_Y \rangle$ is a order isomorphism, $A \subseteq X$ and $B \subseteq Y$ then*

1. *If u is a upper bound of B then $(f^{-1})(u)$ is a upper bound of $f^{-1}(B)$*
2. *If l is a lower bound of B then $(f^{-1})(l)$ is a lower bound of $f^{-1}(B)$*
3. *If u is a upper bound of A then $f(u)$ is a upper bound of $f(A)$*
4. *If l is a lower bound of A then $f(u)$ is a lower bound of $f(A)$*
5. *$f(\upsilon(A)) = \upsilon(f(A))$*
6. *$f(\lambda(A)) = \lambda(f(A))$*
7. *If $\max(A)$ exist then $\max(f(A))$ exist and $\max(f(A)) = f(\max(A))$*
8. *If $\min(A)$ exist then $\min(f(A))$ exist and $\min(f(A)) = f(\min(A))$*
9. *If $\sup(A)$ exist then $\sup(f(A))$ exist and $\sup(f(A)) = f(\sup(A))$*
10. *If $\inf(A)$ exist then $\inf(f(A))$ exist and $\inf(f(A)) = f(\inf(A))$*

Proof. First using [theorem: 3.54] we have that $f: \langle X, \leqslant_X \rangle \rightarrow \langle Y, \leqslant_Y \rangle$ and $f^{-1}: \langle Y, \leqslant_Y \rangle \rightarrow \langle X, \leqslant_X \rangle$ are increasing.

1. Let $x \in f^{-1}(B)$ then $\exists y \in B$ such that $y = f(x)$, as u is a upper bound of B, we have that $y \leqslant_B u$. So $x \underset{[\text{theorem: } 2.70]}{=} (f^{-1})(f(x)) = (f^{-1})(y) \leqslant_A (f^{-1})(u)$, proving that $(f^{-1})(u)$ is a upper bound of $f^{-1}(B)$.
2. Let $x \in f^{-1}(B)$ then $\exists y \in B$ such that $y = f(x)$, as l is a lower bound of B we have that $l \leqslant_B y$. So $(f^{-1})(l) \leqslant_A (f^{-1})(y) = (f^{-1})(f(x)) \underset{[\text{theorem: } 2.70]}{=} x$, proving that $(f^{-1})(l)$ is a lower bound of $f^{-1}(B)$.
3. If $y \in f(A)$ then $\exists x \in A$ such that $y = f(x)$. As u is a upper bound of A we have that $x \leqslant_A u$, so $y = f(x) \leqslant_B f(u)$ proving that $f(u)$ is a upper bound of $f(A)$.
4. If $y \in f(A)$ then $\exists x \in A$ such that $y = f(x)$, As l is a lower bound of A we have that $l \leqslant_A x$, so $f(l) \leqslant_B f(x) = y$ proving that $f(l)$ is a lower bound of $f(A)$.
5. If $y \in f(\upsilon(A))$ then there $\exists x \in \upsilon(A)$ such that $y = f(x)$. As $x \in \upsilon(A)$, x is a upper bound of B, so that by (3) $y = f(x)$ is a upper bound of $f(A)$. Hence
$$f(\upsilon(A)) \subseteq \upsilon(f(A)) \tag{3.12}$$
If $y \in \upsilon(f(A))$ then by (1) $(f^{-1})(y)$ is a upper bound of $f^{-1}(f(A)) \underset{[\text{theorem: } 2.55]}{=} A$ so that $(f^{-1})(y) \in \upsilon(A)$. So $y \underset{[\text{theorem: } 2.70]}{=} f((f^{-1})(y)) = y \in f(\upsilon(A))$. Hence $\upsilon(f(A)) \subseteq f(\upsilon(A))$ which combined with [eq: 3.12] proves
$$f(\upsilon(A)) = \upsilon(f(A))$$
6. If $y \in f(\lambda(A))$ then there $\exists x \in \lambda(A)$ such that $y = f(x)$. As $x \in \lambda(A)$, x is a lower bound of A, so that by (4) $y = f(x)$ is a lower bound of $f(A)$. Hence
$$f(\lambda(A)) \subseteq \lambda(f(A)) \tag{3.13}$$
If $y \in \lambda(f(A))$ then by (2) $(f^{-1})(y)$ is a lower bound of $f^{-1}(f(A)) \underset{[\text{theorem: } 2.55]}{=} A$ so that $(f^{-1})(y) \in \lambda(A)$. So $y \underset{[\text{theorem: } 2.70]}{=} f((f^{-1})(y)) = y \in f(\lambda(A))$. Hence $\lambda(f(A)) \subseteq f(\lambda(A))$ which combined with [eq: 3.12] proves
$$f(\lambda(A)) = \lambda(f(A))$$
7. If $\max(A)$ exist then $\max(A) \in A$ giving $f(\max(A)) \in f(A)$. Let $y \in f(A)$ then $\exists x \in A$ such that $y = f(x)$, as $\max(A)$ exist we have $x \leqslant_A \max(A)$ so that $y = f(x) \leqslant_B f(\max(A))$. So
$$\max(f(A)) \text{ exist and } \max(f(A)) = f(\max(A))$$
8. If $\min(A)$ exist then $\min(A) \in A$ giving $f(\min(A)) \in f(A)$. Let $y \in f(A)$ then $\exists x \in A$ such that $y = f(x)$, as $\min(A)$ exist we have $\min(A) \leqslant_A x$ so that $f(\min(A)) \leqslant_B f(x) = y$. So
$$\min(f(A)) \text{ exist and } \min(f(A)) = f(\min(A))$$

9. If $\sup(A)$ exists then $\min(\upsilon(A))$ exists and $\sup(A)=\min(\upsilon(A))$. Using (8) $\min(f(\upsilon(A)))$ exist, As $f(\upsilon(A)) \underset{(5)}{=} \upsilon(f(A))$ we have that $\min(\upsilon(f(A)))$ exist and
$$\sup(f(A))=\min(\upsilon(f(A))) \underset{(5)}{=} \min(f(\upsilon(A))) \underset{(8)}{=} f(\min(\upsilon(A)))=f(\sup(A))$$
10. If $\inf(A)$ exists then $\max(\lambda(A))$ exists and $\inf(A)=\max(\lambda(A))$. Using (7) $\max(f(\lambda(A)))$ exist, As $f(\lambda(A)) \underset{(6)}{=} \lambda(f(A))$ we have that $\max(\lambda(f(A)))$ exist and
$$\inf(f(A))=\max(\lambda(f(A))) \underset{(6)}{=} \max(f(\lambda(A))) \underset{(7)}{=} f(\max(\lambda(A)))=f(\inf(A))$$
□

Theorem 3.80. *Let $\langle X,\leqslant_X\rangle$ be a conditional complete partial ordered set, $\langle Y,\leqslant_Y\rangle$ a partial ordered class and $f:\langle X,\leqslant_X\rangle\rightarrow\langle Y,\leqslant_Y\rangle$ a order isomorphism then $\langle Y,\leqslant_Y\rangle$ is conditionally complete.*

Proof. Let $A\subseteq Y$ be such that A is bounded above and non empty. Let u be a upper bound of A then by [lemma: 3.79] we have that $(f^{-1})(u)$ is a upper bound of $f^{-1}(A)$. As $A\neq\varnothing$ there exists a $a\in A$ which as f is surjective means that $\exists x$ such that $a=f(x)$ hence $x\in f^{-1}(A)$ proving that $f^{-1}(A)\neq\varnothing$. As $\langle X,\leqslant_X\rangle$ is conditional complete $\sup(f^{-1}(A))$ exist. Using [lemma: 3.79] $\sup(f(f^{-1}(A)))$ exist which as $A \underset{\text{[theorem: 2.55]}}{=} f(f^{-1}(A))$ proves that $\sup(A)$ exist. So $\langle X,\leqslant_Y\rangle$ is conditional complete. □

3.3.4 Well ordering

Definition 3.81. *A partial ordered class $\langle X,\leqslant\rangle$ is **well ordered** is every non empty sub-class of X has a least element. In other words if $\forall A\in\mathcal{P}(X)$ $\min(A)$ exist.*

Theorem 3.82. *If $\langle X,\leqslant_X\rangle$, $\langle Y,\leqslant_Y\rangle$ are partial ordered sets, $f:\langle X,\leqslant_X\rangle\rightarrow\langle Y,\leqslant_Y\rangle$ a order isomorphism then if $\langle X,\leqslant_X\rangle$ is well ordered $\langle Y,\leqslant_Y\rangle$ is well ordered.*

Proof. Let $A\subseteq Y$ be a non empty subclass of Y. Then $\exists a\in A$ and as f is a bijection there exist a $x\in X$ such that $y=f(x)$, from which it follows that $x\in f^{-1}(A)$. So
$$f^{-1}(A)\neq\varnothing$$
As $\langle X,\leqslant_X\rangle$ is well ordered we have that $f^{-1}(A)$ has a least element, hence
$$\exists m'\in f^{-1}(A)\,\text{such that}\,\forall a\in f^{-1}(A) \text{ we have } m'\leqslant_X a$$
Take now $m=f(m')$ then as $m'\in f^{-1}(A)$ we have that
$$m\in A \tag{3.14}$$
Further if $a\in A$ then as f is surjective there exists a $b\in X$ such that $a=f(b)$ or $b\in f^{-1}(A)$, so that $m'\leqslant_X b$. As f is a order isomorphism we have $m=f(m')\leqslant_Y f(b)=a$. Hence we have proved that
$$\forall a\in A \text{ we have } m\leqslant a \tag{3.15}$$
From [eq: 3.14] and [eq: 3.15] we conclude finally that $\langle Y,\leqslant_Y\rangle$ is well ordered. □

Theorem 3.83. *If $\langle X,\leqslant\rangle$ is a partial ordered class, $B\subseteq X$ then for $\langle B,\leqslant_{|B}\rangle$ [see theorem: 3.35] we have*

1. *If $\langle X,\leqslant\rangle$ is totally ordered then $\langle B,\leqslant_{|B}\rangle$ is totally ordered*
2. *If $\langle X,\leqslant\rangle$ is well ordered then $\langle B,\leqslant_{|B}\rangle$ is totally ordered*

Proof.

1. If $x,y\in B\Rightarrow x,y\in X$ hence $x\leqslant y\vee y\leqslant x$ so that $x\leqslant_{|B} y\vee y\leqslant_{|B} x$.
2. If $C\subseteq B$ is a non empty class then as $B\subseteq X$ we have $\varnothing\neq C\subseteq X$. So there exists a least element c of C. So $c\in C$ and $\forall x\in C$ we have $c\leqslant x \underset{x\in B}{\Rightarrow} c\leqslant_{|B} x$ proving that c is a least element of C using the order relation $\leqslant_{|B}$. □

Well ordering is a stronger condition then conditional completeness and totally ordering

Theorem 3.84. *Let $\langle X,\leqslant\rangle$ is a well ordered class then*

1. *$\langle X,\leqslant\rangle$ is totally ordered*

2. *$\langle X, \leqslant \rangle$ is conditional complete*
3. *$\forall x, y \in X$ we have $x \leqslant y$ or $y < x$*

Proof.

1. If $x, y \in X$ then $\{x, y\}$ is a non empty sub-class of X and must have a least element. If x is the least element then $x \leqslant y$ and if y is the least element then $y \leqslant x$, so $\langle X, \leqslant \rangle$ is totally ordered.
2. If A is a non empty sub-class of X that is bounded above then $\upsilon(A) \neq \varnothing$. Using well ordering we have that $\sup(A) = \min(\upsilon(A))$ exist.
3. As by (1) $\langle X, \leqslant \rangle$ is totally ordered we have that x and y are comparable, hence by [theorem: 3.39] we have $x \leqslant y \vee y < x$. □

One difference between the order relation on the set of whole numbers $\mathbb{Z}$ and the set of real numbers $\mathbb{R}$ is that there does not exist a whole number between 1 and 2 while for the real numbers there is the real number 1.5 between 1 and 2. This leads to the following definition.

Definition 3.85. (Immediate successor) *Let $\langle X, \leqslant \rangle$ be a partial ordered set and $x, y \in X$ then y is the **immediate** successor of x iff*

1. *$x < y$*
2. *$\neg(\exists z \in X$ such that $x < z \wedge z < y)$ [in words there does not exists a $x \in X$ such that $x < z < y$]*

Theorem 3.86. *Let $\langle X, \leqslant \rangle$ be a well ordered class then every element that is not a greatest element of X has a immediate successor.*

Proof. Using [theorem: 3.84] we have that $\langle X, \leqslant \rangle$ is totally ordered. Let $x \in X$ such that x is not a greatest element in X. Take $B = \{y \in X | x < y\}$ then if $B = \varnothing$ we have that $X \setminus B = X$ so $\forall r \in X$ we have $r \notin B$ or $\neg(x < r)$, by [theorem: 3.84] we have that $r \leqslant x$, proving that x is a greatest element of X which contradicts or hypothesis.. So we must have that $B \neq \varnothing$, by well ordering there exist a least element b of B, which as $b \in B$ gives $x < b$. Assume that there exist a $a \in X$ such that $x < a \wedge a < b$, then we must have that $a \in B$ and $a < b$. As b is the least element of B and $a \in B$ we have $b < a$ leading to the contradiction $a < a$. So b is a immediate successor of x/ □

Definition 3.87. *Let $\langle X, \leqslant \rangle$ be a partial ordered class then $B \subseteq A$ is a **section** of X if*

$$\forall x \in X \text{ we have } \forall y \in B \textit{ with } x \leqslant y \textit{ that } x \in B$$

Lemma 3.88. *Let $\langle X, \leqslant \rangle$ be a well ordered class and $B \subseteq X$ then*

$$B \textit{ is a section} \Leftrightarrow B = X \textit{ or } B \textit{ is a initial segment of } X \textit{ [definition: 3.46]}$$

Proof.

⇒. Let B be a section of X then if $B = X$ we are done. So we must prove the theorem for $B \neq X$ or equivalently $X \setminus B \neq \varnothing$. Because X is well ordered, there a exists a least element $l \in X \setminus B$. Consider the initial segment $S_{X,l} = \{x \in X | x < l\}$ [see definition: 3.46]. Let $x \in S_{X,l}$ so that $x < l$. Assume that $x \notin B$ then $x \in X \setminus B$ so, as l is a least element of $X \setminus B$, we have $l \leqslant x$ which combined with $x < l$ leads to the contradiction $l < l$. So we must have that $x \in B$ which proves that

$$S_{X,l} \subseteq B \tag{3.16}$$

Let $x \in B$, as X is well ordered we have by [theorem: 3.84] that $l \leqslant x \vee x < l$. Assume that $l \leqslant x$ then, as B is a section, we have $l \in B$ contradicting $l \in X \setminus B$ [as l is least element of $X \setminus B$]. So we must have $x < l$ or $x \in S_{X,l}$ so $B \subseteq S_{X,l}$. Combining this result with [eq: 3.16] proves

$$S_{X,l} = B$$

⇐. If $X = B$ then $\forall x \in X$ we have $\forall y \in B = X$ with $x \leqslant y$ that trivially $x \in X = B$, so B is a section. If B is initial segment then there exist a $l \in X$ such that $B = \{y \in X | y < l\}$. Take $x \in X$ then if $y \in B$ with $x \leqslant y$ we have $y < l$ so that $x < l$ hence $x \in B$, proving that B is a section. □

A application of the above lemma is Transfinite Induction.

Theorem 3.89. (Transfinite Induction) *Let $\langle X, \leqslant \rangle$ be a well ordered class and let $P(x)$ a proposition about x [a statement about x that can be true or false] such that*

$$\forall x \in X \textit{ such that, if } P(y) \textit{ is true for every } y < x \textit{ then } P(x) \textit{ is true} \tag{3.17}$$

then

$$\forall x \in X \ P(x) \textit{ is true}$$

Proof. We prove this by contradiction. Assume that $\exists x\in X$ such that $P(x)$ is false, then $B=\{x\in X|\mathcal{P}(x) \text{ is false}\}$ is non empty. As X is well ordered there exist a least element $l\in B$. Take $x\in X$ with $x<l$ then $x\notin B$ [for if $x\in B$ then $l\leqslant x$, which combined with $x<l$ gives the contradiction $l<l$] so that $P(x)$ is true. By the hypothesis [eq: 3.17] we have that $P(l)$ is true, which means that $l\notin B$ contradicting $l\in B$. So we must have that $\forall x\in X\ P(x)$ is true. $\square$

Lemma 3.90. *Let $\langle X,\leqslant\rangle$ be a well ordered class, $B\subseteq X$ and $f:\langle X,\leqslant\rangle\rightarrow\langle B,\leqslant\rangle$ a order isomorphism then $\forall x\in X$ we have $x\leqslant f(x)$*

Proof. We prove this by contradiction. Assume that that$\exists x\in X$ such that $\neg(x\leqslant f(x))$. As $\langle X,\leqslant\rangle$ if well ordered we have by [theorem: 3.84] that $f(x)<x$, hence $C=\{x\in X|f(x)<x\}\neq\varnothing$. By well ordering there exists a least element c of C. As $c\in C$ we have that $f(c)<c$, hence by [theorem: 3.53] $f(f(c))<f(c)$ so that $f(c)\in C$. As c is the least element of C we have $c\leqslant f(c)$, which combined with $f(c)<c$ gives the contradiction $c<c$. So we must have $\forall x\in X$ that $x\leqslant f(x)$. $\square$

Theorem 3.91. *Let $\langle X,\leqslant\rangle$ be a well ordered class then there does not exist a order isomorphism from X to a sub-class of an initial segment of X.*

Proof. We prove this by contradiction. So assume that there exists a initial segment $S_{X,a}=\{y\in X|y<a\}$ of X, a $B\subseteq S_{X,\alpha}$ and a order isomorphism $f:\langle X,\leqslant\rangle\rightarrow\langle B,\leqslant\rangle$. Using the previous lemma [lemma: 3.90] we have that $a\leqslant f(a)$, so $f(a)\notin S_{X,a}$ [for if $f(a)\in S_{X,a}$ then $f(a)<a$ leading to the contradiction $a<a$]. However as $\operatorname{range}(f)=B\subseteq S_{X,a}$ we must have that $f(a)\in S_{X,a}$ and we reach a contradiction. $\square$

Corollary 3.92. *Let $\langle X,\leqslant\rangle$ be a well ordered class then there does not exist a order isomorphism between X and initial segment of X*

Proof. As a initial segment is a sub-class of itself this follows from the previous theorem [theorem: 3.91] $\square$

Theorem 3.93. *If $\langle X,\leqslant_X\rangle$, $\langle Y,\leqslant_Y\rangle$ are well ordered classes then if X is order isomorphic with an initial segment of Y we have that Y is not order isomorphic with any sub-class of X.*

Proof. Let $S_{Y,y}$ be a initial segment of Y and $f:\langle X,\leqslant_X\rangle\rightarrow\langle S_{Y,y},\leqslant_Y\rangle$ a order isomorphism. Assume that there exist a $A\subseteq X$ and a order isomorphism $g:\langle Y,\leqslant_Y\rangle\rightarrow\langle A,\leqslant_A\rangle$, As by [lemma: 2.33],[theorem: 2.52] and the fact that 'increasing' is a property of the graph of a function,we have that $g:\langle Y,\leqslant_Y\rangle\rightarrow\langle X,\leqslant_X\rangle$ is a injective increasing function. Using [theorem: 2.74],[theorem: 3.52] we have that $f\circ g$: $\langle Y,\leqslant_Y\rangle\rightarrow\langle S_{Y,y},\leqslant_Y\rangle$ is a injective increasing function, hence $f\circ g: Y\rightarrow(f\circ g)(Y)$ is a bijective function [see theorem: 2.66] which is increasing, hence by [theorem: 3.57] we have that $f\circ g:\langle Y,\leqslant_Y\rangle\rightarrow\langle(f\circ g)(Y),\leqslant_Y\rangle$ is a order isomorphism. As $(f\circ g)(Y)\subseteq\operatorname{range}(f)$ [see theorem: 2.22] and $\operatorname{range}(f)\subseteq S_{Y,y}$ we have a order isomorphism between Y and a sub-class of a initial segment of Y. By [theorem: 3.91] this is impossible so the assumption is false, hence Y is not order isomorphic to a an initial segment of Y. $\square$

Corollary 3.94. *If $\langle X,\leqslant_X\rangle$, $\langle Y,\leqslant_Y\rangle$ are well ordered classes such that X is order isomorphic with Y then*

1. *X can not be order isomorphic with a initial segment of Y*
2. *Y can not be order isomorphic with a initial segment of X*

Proof. We prove this by contradiction. First by the hypothesis we have $X\cong Y$ and by [theorem: 3.56] $Y\cong X$.

1. If X is order isomorphic with a initial segment of Y then as $Y\cong X$ we have that Y is order isomorphic with a sub-class of X, which by [theorem: 3.93] is not allowed.
2. If Y is order isomorphic with a initial segment of X then as $X\cong Y$ we have that X is order isomorphic with a sub-class of Y, which by [theorem: 3.93] is not allowed. $\square$

Lemma 3.95. *Let $\langle X,\leqslant\rangle$ be a well ordered class and $a,b\in X$ with $a<b$ then $S_{X,a}$ is a initial segment of $S_{X,b}$ [using the order $\leqslant_{|S_{X,y}}$]*

Proof. First if $x\in S_{X,a}$ then $x<a\underset{a<b}{\Rightarrow}x<b$ so that $x\in S_{X,b}$, hence

$$S_{X,a}\subseteq S_{X,b}$$

Now if $x\in S_{X,b}$ and $y\in S_{X,a}$ is such that $x\leqslant_{|S_{X_B}}y$ then $x\leqslant y\underset{y\in S_{X,a}\Rightarrow y<a}{\Rightarrow}x<a$ hence $x\in S_{X,a}$. So $S_{X,a}$ is a section of $S_{X,b}$, as $a\notin S_{X,a}\wedge a\in S_{X,b}$ [for $a<b$] we have $S_{X,a}\neq S_{X,b}$ so that, using [theorem: 3.88], $S_{X,a}$ is a initial segment of $S_{X,b}$. $\square$

Theorem 3.96. *Let Let $\langle X,\leqslant_X\rangle$ and $\langle Y,\leqslant_Y\rangle$ be well ordered classes then exactly one of the following cases hold*

1. *X is order isomorphic with Y*
2. *X is order isomorphic with an initial segment of Y*
3. *Y is order isomorphic with an initial segment of X*

Proof. Define

$$C=\{x\in X|\exists y\in Y \text{ such that } S_{X,x}\cong S_{Y,y}\} \tag{3.18}$$

and

$$F=\{(x,y)\in C\times Y|S_{X,x}\cong S_{Y,y})\tag{3.19}$$

We prove now that F is the graph of a order isomorphism between C and $F(C)$. We have trivially from the definition of F that

$$F\subseteq C\times Y\tag{3.20}$$

Let $(x,y),(x,y')\in F$, then $\mathcal{S}_{X,x}\cong S_{Y,y}$ and $S_{x,x}\cong S_{Y,y'}$ so by [theorem: 3.56]

$$S_{Y,y}\cong S_{Y,y'}\tag{3.21}$$

Assume that $y\neq y'$ then, as $\langle Y,\leqslant_Y\rangle$ is well ordered we have by [theorem: 3.84] either:

- $\boldsymbol{y\leqslant y'}$**.** then $y<y'$ so that by the previous lemma [lemma: 3.95] we have that $S_{Y,y}$ is a initial segment of $S_{Y,y'}$. Using [corollary: 3.92] we have then that $S_{Y,y'}$ is not order isomorphic with $S_{Y,y}$ contradicting [eq: 3.21].
- $\boldsymbol{y'<y}$**.** then by the previous lemma [lemma: 3.95] we have that $S_{Y,y'}$ is a initial segment of $S_{Y,y}$. Using [corollary: 3.92] we have then that $S_{Y,y}$ is not order isomorphic with $S_{Y,y'}$ contradicting [eq: 3.21].

as in all cases we have a contradiction, the assumption must be wrong. Hence

$$\text{If } (x,y),(x,y')\in F \text{ then } y=y'\tag{3.22}$$

Further if $x\in C$ then by definition of C there exists a $y\in Y$ such that $S_{X,x}=S_{Y,y}$ hence $(x,y)\in F$ proving that

$$C\subseteq \operatorname{dom}(F)\tag{3.23}$$

If $(x,y),(x',y)\in F$ then $S_{X,x}\cong S_{Y,y}$ and $\mathcal{S}_{X,x'}\cong S_{Y,y}$ so by [theorem: 3.56] we have that

$$S_{X,x}\cong S_{X,x'}\tag{3.24}$$

Assume that $x\neq x'$ then, as $\langle X,\leqslant_X\rangle$ is well ordered we have by [theorem: 3.84] either:

- $\boldsymbol{x\leqslant x'}$**.** then $x<x'$ so that by the previous lemma [lemma: 3.95] we have that $S_{X,x}$ is a initial segment of $S_{X,x'}$. Using [corollary: 3.92] we have then that $S_{X,x'}$ is not order isomorphic with $S_{X,x}$ contradicting [eq: 3.24].
- $\boldsymbol{x'\leqslant x}$**.** then by the previous lemma [lemma: 3.95] we have that $S_{X,x'}$ is a initial segment of $S_{X,x}$. Using [corollary: 3.92] we have then that $S_{X,x}$ is not order isomorphic with $S_{X,x'}$ contradicting [eq: 3.24].

as in all cases we have a contradiction, the assumption must be wrong. Hence

$$\text{If } (x,y),(x',y)\in F \text{ we have } x=x'\tag{3.25}$$

Combining [eq: 3.20], [eq: 3.22], [eq: 3.23] and [eq: 3.25] it follows that $F:C\rightarrow Y$ is a injective function. Applying then [proposition: 2.66] gives if we define $D=F(C)$

$$F:C\rightarrow D \text{ is a bijection}\tag{3.26}$$

Take $x,y\in C$ such that $x\leqslant_X y$ then by definition of F we have

$$S_{X,x}\cong S_{Y,F(x)} \text{ and } S_{X,y}\cong S_{Y,F(y)}\tag{3.27}$$

Assume now that $\neg(F(x)\leqslant_Y F(y))$ then as $\langle Y,\leqslant_Y\rangle$ is well ordered we have by [theorem: 3.84] that $F(y)<_Y F(x)$. So using [theorem: 3.95] we have that $S_{Y,F(y)}$ is a initial segment of $S_{Y,F(x)}$. As $x\leqslant_X y$ it follows that $S_{X,x}\subseteq S_{X,y}$ [see proposition: 3.47]. So we have using [eq: 3.27]

a) $S_{X,y}$ is order isomorphic with $S_{Y,F(y)}$ a initial segment of $S_{Y,F(x)}$

b) $S_{F(x)}$ is order isomorphic with $S_{X,x}$ a sub-class of $S_{X,y}$

Using [theorem: 3.93] we see that (a) and (b) can not be all true, hence our assumption is false so that $F(x)\leqslant F(y)$. Hence we have that $F:C\rightarrow D$ is a increasing bijection which by [theorem: 3.57] proves that

$$F:\langle C,\leqslant_X\rangle\rightarrow\langle D,\leqslant_Y\rangle \text{ is a order isomorphism or } C\cong D\tag{3.28}$$

Next we prove that

$$C \text{ is a section of } X\tag{3.29}$$

Proof. Let $x \in X$ and take $c \in C$ such that $x \leqslant_X c$. As $S_{X,c} \cong S_{Y,F(c)}$ there exist a order isomorphism

$$g: S_{X,c} \rightarrow S_{Y,F(c)} \tag{3.30}$$

Now as $x \leqslant_X c$ we have by [proposition: 3.47] that $S_{X,x} \subseteq S_{X,c}$. Hence by 2.88 we have that

$$g_{|S_{X,x}}: S_{X,x} \rightarrow S_{X,c} \text{ is a function} \tag{3.31}$$

Further if $y \in S_{X,x}$ we have that $y <_X x$, so as g is a order isomorphism we have $g(y) <_Y g(x)$ proving that $g_{|S_{X,x}}(y) = g(y) \in S_{Y,g(x)}$ or $\text{range}(g_{|S_{X,x}}) \subseteq S_{Y,g(x)}$. So by [theorem: 2.37] it follows that

$$g_{|S_{X,x}}: S_{X,x} \rightarrow S_{Y,g(x)} \text{ is a function} \tag{3.32}$$

As g is a isomorphism and thus injective it follows from [theorem: 2.84] that

$$g_{|S_{X,x}}: S_{X_x} \rightarrow S_{Y,g(x)} \text{ is injective} \tag{3.33}$$

Further if $y \in S_{Y,g(x)}$ then $y <_Y g(x)$, as $g(x) \in S_{Y,F(c)}$ [see eq: 3.30] we have $g(x) <_Y F(c)$ so that $y <_Y F(c)$ proving $y \in S_{Y,F(c)}$. As g is surjective there exist a $u \in S_{X,c}$ such that $y = g(u)$. Assume that $x \leqslant_X u$ then $g(x) \leqslant_Y g(u) = y$, as $y <_Y g(x)$ this gives the contradiction $g(x) < g(x)$. So we have $\neg(x \leqslant u)$ which, as $\langle X, \leqslant_X \rangle$ is well ordered, gives by [theorem: 3.84] that $u <_X x$ so that $u \in S_{X,x}$. So for $y \in S_{Y,g(x)}$ we found a $u \in S_{X,x}$ such that $g_{|S_{X,x}}(u) = g(u) = y$ proving that

$$g_{|S_{X,x}}: S_{X,x} \rightarrow S_{Y,g(x)} \text{ is surjective} \tag{3.34}$$

Further if $u, v \in S_{X,x}$ are such that $u \leqslant_X v$ so that $g_{|S_{X,x}}(u) = g(u) \leqslant_X g(v) = g_{|S_{X,x}}(v)$ proving that

$$g_{|S_{X,x}}: S_{X,x} \rightarrow S_{Y,g(x)} \text{ is increasing} \tag{3.35}$$

Combining [eq: 3.31], [eq: 3.32], [eq: 3.34], [eq: 3.35] we have that $g_{|S_{X,x}}: \langle S_{X,x}, \leqslant_X \rangle \rightarrow \langle S_{Y,g(x)}, \leqslant_Y \rangle$ is a order isomorphism so that $S_{X,x} \cong S_{Y,g(x)}$ hence $x \in C$. Proving that C is as section of X. □

Next we prove that

$$D \text{ is a section of } Y \tag{3.36}$$

Proof. Let $y \in Y$ and take $d \in D$ such that $y \leqslant_Y d$. As $d \in D = \text{range}(F)$ there exist a $c \in C$ such that $F(c) = d$, so $S_{X,c} \cong S_{Y,d} \underset{\text{[theorem: 3.56]}}{\Rightarrow} S_{Y,d} \cong S_{X,c}$. So there exist a order isomorphism

$$f: S_{Y,d} \rightarrow S_{X,c} \tag{3.37}$$

Now from $y \leqslant_D d$ we have by [theorem: 3.47] $S_{Y,y} \subseteq S_{Y,d}$. Hence by 2.88 we have that

$$f_{|S_{Y,y}}: S_{Y,y} \rightarrow S_{X,c} \text{ is a function} \tag{3.38}$$

If $x \in S_{Y,y}$ then $x <_Y y$ so, as f is a order isomorphism, $f_{|S_{Y,y}}(x) = f(x) <_X f(y)$, we have that $f_{|S_{Y,y}}(x) \in S_{Y,f(y)}$, so $\text{range}(f_{|S_{Y,y}}) \subseteq S_{X,f(y)}$. By [theorem: 2.37] it follows that

$$f_{|S_{Y,y}}: S_{Y,y} \rightarrow S_{X,f(y)} \text{ is a function} \tag{3.39}$$

As f is a isomorphism and injective it follows from [theorem: 2.84] that

$$f_{|S_{Y,y}}: S_{Y,y} \rightarrow S_{X,f(y)} \text{ is injective} \tag{3.40}$$

If $x \in S_{X,f(y)}$ then $x <_X f(y)$, as by [eq: 3.37] $f(y) \in S_{X,c}$, we have $f(y) < c$, so that $x <_X c$ or $x \in S_{X,c}$. As f is surjective there exists a $u \in S_{Y,d}$ such that $f(u) = x$. As $u \in S_{Y,d}$ we have that $u <_Y d$. Assume now that $y \leqslant_Y u$ then, as f is a order isomorphism, $f(y) \leqslant_X f(u) = x$, which as $x <_X f(y)$ gives the contradiction $x <_X x$. So we must have that $\neg(y \leqslant_Y u)$, which, as $\langle Y, \leqslant_Y \rangle$ is well ordered, gives by [theorem: 3.84] that $u <_Y y$ or $u \in S_{Y,y}$. So for $x \in S_{X,f(y)}$ there exist a $u \in S_{Y,y}$ such that $f(u) = x$, proving that

$$f_{|S_{Y,y}}: S_{Y,y} \rightarrow S_{X,f(y)} \text{ is surjective} \tag{3.41}$$

Further if $u, v \in S_{Y,y}$ is such that $u \leqslant v$ then $f_{|S_{Y,y}}(u) = f(u) \leqslant f(v) = f_{|S_{U,y}}(v)$ proving that

$$f_{|S_{Y,y}}: S_{Y,y} \rightarrow S_{X,f(y)} \text{ is increasing} \tag{3.42}$$

Combining [eq: 3.39], [eq: 3.40], [eq: 3.41] and [eq: 3.42] we have that

$$f_{|S_{Y,y}}: \langle S_{Y,y}, \leqslant \rangle \rightarrow \langle S_{X,f(y)}, \leqslant_X \rangle \text{ is a order isomorphism,}$$

hence $S_{Y,y} \cong S_{X,f(y)}$. As $f(y) \in S_{X,c} \subseteq X$ and $y \in Y$ it follows from the definition of C that $f(y) \in C$, hence by definition of F $(f(y), y) \in F$ or $y = F(f(y)) \in F(C) = D$, giving $y \in D$. Proving that D is a section of Y. □

To summarize [eq: 3.28], [eq: 3.29] and [eq: 3.36] we have

$$C \cong D \wedge C \text{ is a segment of } X \wedge D \text{ is a segment of } Y \tag{3.43}$$

Assume now that C is a initial segment of X and D is a initial segment of Y then there exist a $r \in X$ and a $s \in Y$ such that $C = S_{X,r}$ and $D = S_{Y,s}$. By 3.43 we have that $S_{X,r} \cong S_{Y,s}$ which by definition of C means that $r \in C$ or as $C = S_{X,r}$ that $r < r$ a contradiction. So we have that

$$\neg(C \text{ is a initial segment of } X \wedge D \text{ is a initial segment of } Y) \tag{3.44}$$

As C is a section of X we have by [theorem: 3.88] that

$$X = C \text{ or } C \text{ is a initial segment of } X \tag{3.45}$$

Like wise, as D is a section of Y we have by [theorem: 3.88] that

$$Y = D \text{ or } D \text{ is a initial segment of } Y \tag{3.46}$$

We have taking [eq: 3.45] and [eq: 3.46] in account that either:

$X = C \wedge Y = D$. then by [eq: 3.43]

$$X \cong Y$$

Using theorem [theorem: 3.94] and the above we have that

$$X \text{ is not order isomorphic with a sub-class of Y}$$

$$Y \text{ is not order isomorphic with a sub-class of X}$$

$X = C \wedge Y \neq D$. then by [eq: 3.46] we have that D is a initial segment of Y, which as by [eq: 3.43] $X = C \cong D$ prove that

$$X \text{ is order isomorphic with a initial segment of } Y$$

If Y is order isomorphic with a initial segment of X then by [theorem: 3.93] we have that X is not order isomorphic to a subset of Y contradicting $X \cong D$ and $X \cong Y$. So

$$Y \text{ is not order isomorphic to a initial segment of } X$$

$$X \text{ is not order isomorphic to } Y$$

$X \neq C \wedge Y = D$. then by [eq: 3.45] we have that C is a initial segment of X, which as by [eq: 3.43] $C \cong D \underset{\text{[theorem: 3.56]}}{\Rightarrow} Y = D \cong C$ proves that

$$Y \text{ is order isomorphic with a initial segment of } X$$

If X is order isomorphic with a initial segment of Y then by [theorem: 3.93] we have that Y is not order isomorphic to a subset of X contradicting $Y \cong C$ and $Y \cong X$. So

$$X \text{ is not order isomorphic to a initial segment of } Y$$

$$X \text{ is not order isomorphic to } Y$$

$X \neq C \wedge Y \neq D$. Using [eq: 3.45] and [eq: 3.46] we have that C is a initial segment of X and D is a initial segment of Y which contradicts [eq: 3.44]. Hence this case does not apply. □

Corollary 3.97. *Let $\langle X, \leqslant \rangle$ be a well ordered class and $Y \subseteq X$ then we have either (but not both):*

1. *Y is order isomorphic with X*
2. *X is order isomorphic with a initial segment of X*

Proof. If $Y \subseteq X$ then $\langle Y, \leqslant_{|Y} \rangle$ is a well ordered class [see theorem: 3.83], so using the previous [theorem: 3.96] we have either:

1. Y is order isomorphic with X
2. Y is order isomorphic with a initial segment of X

3. X is order isomorphic with a initial segment of Y. By [theorem: 3.93] we may not have that Y is order isomorphic with a sub-class of X. As by [theorem: 3.56] $Y \cong Y$ and Y is a sub-class of X we reach a contradiction, so this case never applies. □

3.4 Axiom of choice

The axiom of choice in it's many equivalent forms like

$$\text{Hausdorff}'s\,\text{Maximal Principle}$$
$$\text{Zorn}'s\,\text{Lemma}$$
$$\text{Well} - \text{Ordering Theorem}$$

plays a major role in some fundamental theorems about the product of sets, the existence of a basis for a vector space, etc.

Definition 3.98. *Let A be a class then $\mathcal{P}'(A)$ is defined as*

$$\mathcal{P}'(A) = \mathcal{P}(A) \setminus \{\varnothing\}$$

In other words it is the collection of all non empty sub sets of a set

It turns out that if A is a set then $\mathcal{P}'(A)$ is also a set.

Theorem 3.99. *If A is a set then $\mathcal{P}'(A)$ is a set*

Proof. Using the Axiom of Power [axiom 1.64] we have that $\mathcal{P}(A)$ is a set. As $\mathcal{P}'(A) \subseteq \mathcal{P}(A)$ [see [theorem: 1.25] it follow from the Axiom of Subsets [axiom: 1.54] that $\mathcal{P}'(A)$ is a set. □

Definition 3.100. (Choice Function) *Let A be a set then a **choice function for A** is a function $f: \mathcal{P}'(A) \to A$ such that $\forall B \in \mathcal{P}'(A)$ we have $f(B) \in B$*

So a choice function picks out one element out of each subset of A and the axiom of choice ensures the existence of a choice function for a set.

Axiom 3.101. (Axiom of Choice) *If A is a set then there exist a choice function for A*

As a application of the axiom of choice we have the following theorem

Theorem 3.102. *If $f: A \to B$ is a surjective function then there exists a injective function $g: B \to A$ such that $f \circ g = \mathrm{Id}_B$*

Proof. By the axiom of choice there exists a choice function

$$c: \mathcal{P}'(A) \to A \text{ such that } \forall A \in \mathcal{P}'(A) \text{ we have } c(A) \in A$$

If $f: A \to B$ is surjective. Then $\forall y \in B$ we have that $f^{-1}(\{y\})$ is a non empty subset of $A \Rightarrow f^{-1}(\{y\}) \in \mathcal{P}'(A)$. Define then the function

$$g: B \to Y \text{ by } g(y) = c(f^{-1}(\{y\}))$$

Now if $y \in Y$ then, as c is a choice function, $c(f^{-1}(\{y\})) \in f^{-1}(\{y\})$ so that $f(c(f^{-1}(\{y\}))) = y$. Hence we have that $(f \circ g)(y) = f(g(y)) = f(c(f^{-1}(\{y\}))) = y$ or

$$f \circ g = \mathrm{Id}_B$$

If $g(y) = g(y')$ then we have $f(g(y)) = f(g(y')) \underset{f \circ g = \mathrm{Id}_B}{\Rightarrow} \mathrm{Id}_B(y) = \mathrm{Id}_B(y') \Rightarrow y = y'$ proving that

$$g: B \to Y \text{ is injective}$$ □

The important thing to remember in the above is that the axiom of choice ensures the existence of $g: B \to A$ but does not give a way to construct the function g itself.
We have the following equivalent statements of the axiom of choice

Theorem 3.103. *The following are equivalent*

1. *The Axiom of Choice*

2. *Let $\mathcal{A}$ be a set of sets such that:*

   a. $\forall A \in \mathcal{A}$ *we have* $A \neq \varnothing$

   b. $\forall A, B \in \mathcal{A}$ *with* $A \neq B$ *we have* $A \bigcap B = \varnothing$

   *then there exist a set C called the **choice set for $\boldsymbol{\mathcal{A}}$** such that*

   a. $C \subseteq \bigcup \mathcal{A}$

   b. $\forall A \in \mathcal{A}$ *we have* $A \bigcap C \neq \varnothing$ *and if* $y, y' \in A \bigcap C$ *then* $y = y'$

   In other words C consists of exactly one element from each $A \in \mathcal{A}$.

3. *If $\{A_i\}_{i \in I} \subseteq \mathcal{A}$ is a family of non empty sets [$\forall i \in I$ we have $A_i \neq \varnothing$] where $I, \mathcal{A}$ are sets then there exists a function $f: I \to \bigcup_{i \in I} A_i$ such that $\forall i \in I$ we have $f(i) \in A_i$*

Proof.

1 ⇒ 2. Take $U = \bigcup \mathcal{A}$ [see definition: 1.56]. As $\mathcal{A}$ is a set we have by the Axiom of Union [axiom: 1.61] that U is a set. So we can apply the Axiom of Choice [axiom: 3.101] to get a function

$$c: \mathcal{P}'(U) \to U \text{ such that } \forall A \in \mathcal{P}'(U) \text{ we have } c(A) \in A$$

If $A \in \mathcal{A}$ then $A \neq \varnothing$ and using [theorem: 1.60] we have $A \subseteq U$ proving that $A \in \mathcal{P}'(U)$ hence

$$\mathcal{A} \subseteq \mathcal{P}'(U)$$

so we can take the **image** of $\mathcal{A}$ by c

$$C = c(\mathcal{A})$$

We have now:

a) If $x \in C$ then $\exists A \in \mathcal{A}$ such that $x = (c)(A)$, which as c is a choice function means that $x \in A$ hence, by [theorem: 1.60], we have that $x \in \bigcup \mathcal{A}$ proving that

$$C \subseteq \bigcup \mathcal{A}$$

b) Let $A \in \mathcal{A}$ then $(c)(A) \in c(\mathcal{A}) = C$ and, as c is a choice function, $(c)(A) \in A$ [note: $(c)(A)$ is function application and $c(\mathcal{A})$ is the image of $\mathcal{A}$ by c]. Hence

$$A \bigcap C \neq \varnothing$$

If $y, y' \in A \bigcap C$ then as $y, y' \in C = c(\mathcal{A})$ there exist $Y, Y' \in \mathcal{A}$ such that $y = (c)(Y)$ and $y' = (c)(Y')$, as c is a choice function we have $y = (c)(Y) \in Y$ and $y' = (c)(Y') \in Y'$. Assume that $Y \neq Y'$ then we have the contradiction $y, y' \in Y \bigcap Y' = \varnothing$, so we have that Y=Y' but then $y = c(Y) = c(Y') = y'$ proving that $y = y'$. So

$$y, y' \in A \bigcap C \Rightarrow y = y'$$

so (2.a) and (2.b) is proved.

2 ⇒ 1. Let A be a set and let $B \in \mathcal{P}'(A)$ then $\varnothing \neq B \subseteq A$. Define now

$$P_B = \{(B, x) | x \in B\} \tag{3.47}$$

If $(B, x) \in P_B$ then as $B \in \mathcal{P}'(A)$ and $x \in B \subseteq A$ we have $(B, x) \in \mathcal{P}'(A) \times A$ or

$$P_B \subseteq \mathcal{P}'(A) \times A \text{ or } P_B \in \mathcal{P}(\mathcal{P}'(A) \times A) \tag{3.48}$$

As $B \neq \varnothing$ we have that $\exists b \in B$ so that $(B, p) \in P_B$ proving that

$$\forall B \in \mathcal{P}'(A) \text{ we have } P_B \neq \varnothing \tag{3.49}$$

If $x \in P_B \bigcap P_{B'}$ then $\exists b \in B$ and $b' \in B$ such that $(B, b) = x = (B', b')$ proving that $B = B'$, hence $P_B = P_{B'}$. From this it follows that

$$\forall B, B' \in \mathcal{P}'(A) \text{ we have If } P_B \neq P_{B'} \text{ then } P_B \bigcap P_{B'} = \varnothing \tag{3.50}$$

Define

$$\mathcal{A} = \{P_B | B \in \mathcal{P}'(A)\} \subseteq \mathcal{P}(\mathcal{P}'(A) \times A) \tag{3.51}$$

As A is a set we have by [theorem: 3.99] that $\mathcal{P}'(A)$ is a set, using [theorem: 1.67] it follow that $\mathcal{P}'(A)\times A$ is a set, applying the Axiom of Power sets [axiom: 1.64] proves that $\mathcal{P}(\mathcal{P}'(A)\times A)$ is a set. As by [eq: 3.51] we have that $\mathcal{A}\subseteq\mathcal{P}(\mathcal{P}'(A)\times A)$ we can use the Axiom of Sub Sets [axiom: 1.54] giving

$$\mathcal{A}\,\text{is}\,a\,\text{set} \tag{3.52}$$

So the conditions for the hypothesis (2) are satisfied by [eq: 3.52],[eq: 3.49] and [eq: 3.50] hence there exist a choice set C for $\mathcal{A}$ such that:

$$C\subseteq\bigcup\mathcal{A}\text{ and }\forall B\in\mathcal{A}\text{ we have }B\bigcap C\neq\varnothing\text{ and if }y,y'\in B\bigcap C\text{ then }y=y' \tag{3.53}$$

If $x\in C$ then $\exists y\in\mathcal{A}$ such that $x\in y$. As $y\in\mathcal{A}$ there exists a $B\in\mathcal{P}'(A)$ such that $y=P_B=\{(B,x)|x\in B\}$, hence there exist a $b\in B$ such that $x=(B,b)\in P_B\subseteq\mathcal{P}'(A)\times A$ [see eq: 3.48] proving that

$$C\subseteq\mathcal{P}'(A)\times A \tag{3.54}$$

If $(B,y),(B,y')\in C$ then $(B,y),(B,y')\in P_B\bigcap C\underset{\text{[eq: 3.53}}{\Rightarrow}(B,y)=(B,y')$ proving that $y=y'$, so

$$\text{If }(B,y),(B,y')\in C\text{ then }y=y' \tag{3.55}$$

Let $B\in\mathcal{P}'(A)$ then $P_B\in\mathcal{A}$ so that by [eq: 3.53] $P_B\bigcap C\neq\varnothing$ hence there exist a $y\in B$ such that $(B,y)\in C$ proving that

$$\mathcal{P}'(A)\subseteq\operatorname{dom}(C) \tag{3.56}$$

From [eq: 3.54], [eq: 3.55] and [eq: 3.56] it follows that

$$C:\mathcal{P}'(A)\to A\text{ is a function} \tag{3.57}$$

Let $B\in\mathcal{P}'(A)$ then $(B,C(B))\in C\subseteq\bigcup\mathcal{A}$ so that $\exists B'\in\mathcal{P}'(A)$ such that $(B,C(B))\in P_{B'}$ hence $B=B'$ and $C(B)\in B'=B$ proving that $\forall B\in\mathcal{P}'(A)$ we have $C(B)\in B$, so that

$$C:\mathcal{P}'(A)\to A\text{ is a choice function}$$

proving (1)

1 ⇒ 3. Let $\{A_i\}_{i\in I}\subseteq\mathcal{A}$ be a family of non empty sets where $I,\mathcal{A}$ are sets. Then using [theorem: 2.118] it follows that $\bigcup_{i\in I}A_i$ is a set. Using the Axiom of Choice [axiom: 3.101] there exist a choice function

$$c:\mathcal{P}'\left(\bigcup_{i\in I}A_i\right)\to\bigcup_{i\in I}A_i\text{ where }\forall A\in\mathcal{P}'\left(\bigcup_{i\in I}A_i\right)c(A)\in A$$

Let $A:I\to\mathcal{A}$ be the function that defines $\{A_i\}_{i\in I}\subseteq\mathcal{A}$ then $\forall i\in I$ we have that $A(i)=A_i\subseteq\bigcup_{i\in I}A_i$ [see: theorem: 2.125] or $A(i)\in\mathcal{P}(\bigcup_{i\in I}A_i)$, further as $A_i\neq\varnothing$ we have that $A_i\in\mathcal{P}'(\bigcup_{i\in I}A_i)$, hence $\operatorname{range}(A)\subseteq\mathcal{P}'(\bigcup_{i\in I}A_i)$. Using [theorem: 2.37] it follows that $A:I\to\mathcal{P}'(\bigcup_{i\in I}A_i)$ is also a function. If we take $f=c\circ A$ then

$$f:I\to\bigcup_{i\in I}A_i\text{ is a function and }\forall i\in I\text{ we have }f(i)=c(A(i))=c(A_i)\in A_i$$

proving (3).

3 ⇒ 1. Let A be a set and define the family $\{B_C\}_{C\in\mathcal{P}'A}\subseteq\mathcal{P}'(A)$ by $B=\mathrm{Id}_{\mathcal{P}'(A)}:\mathcal{P}'(A)\to\mathcal{P}'(A)$ [see example: 2.47]. For every $C\in\mathcal{P}'(A)$ we have $B_C=\mathrm{Id}(C)=C\neq\varnothing$, further as A is a set we have by [theorem: 3.99] that $\mathcal{P}'(A)$ is a set. So the conditions for (3) are satisfied and by (3) there exist a function

$$f:\mathcal{P}'(A)\to\bigcup_{C\in\mathcal{P}'(A)}B_C\text{ such that }\forall C\in\mathcal{P}'(A)\text{ we have }f(C)\in B_C=\mathrm{Id}(C)=C \tag{3.58}$$

Let $x\in\bigcup_{C\in\mathcal{P}'(A)}B_C$ then $\exists C\in\mathcal{P}'(A)$ such that $x\in B_C=\mathrm{Id}_{\mathcal{P}'(A)}(C)=C\subseteq A\Rightarrow x\in A$. So $\bigcup_{C\in\mathcal{P}'(A)}B_C\subseteq A$. Using then [theorem: 2.33] we have

$$f:\mathcal{P}'(A)\to A\text{ is a function with }\forall C\in\mathcal{P}'(A)\text{ we have }f(C)\in C$$

which proves that $f:\mathcal{P}'(A)\to A$ is a choice function for A, proving (1). □

Theorem 3.104. *Let A,B be sets such that $\forall a\in A$ there exist a $b\in B$ satisfying $P(a,b)$ [where $P(a,b)$ is a predicate] then there exist a function $f:A\to B$ such that $\forall a\in A\ P(a,f(a))$. Furthermore as a function defines a family you can also say that there is a family $\{f_a\}_{a\in A}\subseteq B$ such that $\forall a\in A\ P(a,f_a)$.*

Proof. Let $a \in A$ then by the hypothesis the set $\{b \in B | P(a, b)\}$ is non empty. This allows us to define the function

$$\mathcal{A}: A \rightarrow \mathcal{P}'(B) \text{ by } \mathcal{A}(a) = \{b \in B | P(a, b)\} \subseteq B$$

defining the family

$$\{\mathcal{A}_a\}_{a \in A} \subseteq \mathcal{P}'(B)$$

Applying then [theorem: 3.103 (3)] there exist a function

$$f: A \rightarrow \bigcup_{a \in A} \mathcal{A}_a$$

such that $\forall a \in A$ we have $f(a) \in \mathcal{A}_a$ so that $f(a)$ satisfies $P(a, f(a))$. Further as $\forall a \in A$ $\mathcal{A}_a \subseteq B$ we have that $\bigcup_{a \in A} \mathcal{A}_a \subseteq B$, hence we have that

$$f: A \rightarrow B$$

is a function such that $\forall a \in A$ $P(a, f(a))$. □

As a application of the Axiom of Choice we have the following theorems about the product of a family of sets. First we prove that the projection function is surjective.

Theorem 3.105. *Let $\{A_i\}_{i \in I} \subseteq \mathcal{A}$ be a family of* ***non empty sets*** *where $I, \mathcal{A}$ are sets then $\forall i \in I$ we have that the projection function*

$$\pi_i: \prod_{j \in I} A_j \rightarrow A_i \text{ defined by } \pi_j(x) = x(j) \text{ [see definition: 2.146]}$$

is a surjection.

Proof. Let $i \in I$ and take $x \in A_i$. Consider the family $\{A_j\}_{j \in I \setminus \{i\}}$ [see definition: 2.106] then $\forall j \in I \setminus \{i\}$ we have $A_j \neq \varnothing$. So we can use [theorem: 3.103 (3)] to find a function

$$f: I \setminus \{i\} \rightarrow \bigcup_{j \in I \setminus \{i\}} A_j \text{ such that } \forall j \in I \setminus \{i\} \text{ we have } f(j) \in A_j$$

By the definition of the product of a family of sets we have that

$$f \in \prod_{j \in I \setminus \{i\}} A_j$$

Define now $g: I \rightarrow \bigcup_{j \in I} A_j$ by $g(j) = \left\{ \begin{smallmatrix} x \text{ if } j = i \\ f(j) \text{ if } j \in I \setminus \{i\} \end{smallmatrix} \right.$ then by [theorem: 2.145] we have that $g \in \prod_{i \in I} A_i$. Finally by $\pi_i(g) = g(i) = x$ proving surjectivity. □

Second we prove that the product of a family of sets is not empty if and only if every set in the family is non empty.

Theorem 3.106. *Let $\{A\}_{i \in I} \subseteq \mathcal{A}$ be a family of sets where $I, \mathcal{A}$ are sets then we have*

$$\prod_{i \in I} A_i \neq \varnothing \Leftrightarrow \forall i \in I \text{ we have } A_i \neq \varnothing$$

Proof.

⇒. We prove this by contradiction, so assume that $\exists i \in I$ such that $A_i = \varnothing$. As $\prod_{i \in I} A_i \neq \varnothing$ there exists a $x \in \prod_{i \in I} A_i$ such that $\forall j \in I$ $x_j \in A_j$, in particular we would have $x_i \in A_i$ contradicting $A_i = \varnothing$. So we must have that $\forall i \in I$ we have $A_i \neq \varnothing$.

⇐. If $\forall i \in I$ we have $A_i \neq \varnothing$ we have by [theorem: 3.103 (3)] that there exist a function

$$f: I \rightarrow \bigcup_{i \in I} A_i \text{ such that } \forall i \in I \text{ we have } f(i) \in \mathrm{A}_i$$

which by definition of the product means that $f \in \prod_{i \in I} A_i$ proving that

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

□

We can rephrase the above theorem in another way.

Theorem 3.107. *Let I be a set, B a set and $\forall i \in I$ there exist a $A_i \subseteq A_i$*

The Axiom of Choice has also important consequences for partial ordered sets.

Theorem 3.108. *Let $\langle X, \leqslant \rangle$ be a partial ordered **set** such that:*

1. *X has a least element p*
2. *Every chain [see definition:3.41] of X has a supremum*

then there is a element $x \in X$ which has no immediate successor [see definition: 3.85]

Proof. We prove this by contradiction, so assume that $\forall x \in X$ there exist a immediate successor. Given $x \in X$ define $T_x = \{y | y$ is a immediate successor of $x\}$ then $T_x \neq \varnothing$ so that $T_x \in \mathcal{P}'(X)$. Using the Axiom of Choice [axiom: 3.101] there exist a choice function

$$c: \mathcal{P}'(A) \rightarrow A \text{ such that } \forall A \in \mathcal{P}'(X) \text{ we have } c(A) \in A \tag{3.59}$$

As $\forall x \in X$ we have $T_x \in \mathcal{P}'(X)$ so that $c(T_x)$ is well defined we can use [proposition: 2.94] to define the function

$$\operatorname{succ}: X \rightarrow X \text{ by } \operatorname{succ}(x) = c(T_x).$$

If $x \in X$ then $\operatorname{succ}(x) = c(T_x) \in T_x$ so that $\operatorname{succ}(x)$ is a immediate successor of x, to summarize

$$\operatorname{succ}: X \rightarrow X \text{ is a function such that } \forall x \in X \operatorname{succ}(x) \text{ is a immediate successor of } x \tag{3.60}$$

Before we can reach the contradiction we need to have some definitions and sub lemmas.

Definition 3.109. *$A \subseteq X$ is a **p-sequence** iff*

1. *$p \in A$*
2. *If $x \in A$ then $\operatorname{succ}(x) \in A$*
3. *If $C \subseteq A$ is a chain then $\sup(C) \in A$ [note that by hypothesis (2) $\sup(C)$ exist]*

Note 3.110. X is a p-sequence so there exist p-sequences.

Proof. First $p \in X$ by the hypothesis (1), second if $x \in X$ then by [eq: 3.60] $\operatorname{succ}(X) \in X$ and finally if C is chain then by definition of the supremum $\sup(C) \in X$ □

Lemma 3.111. *Every intersection of a set of p-sequences is a p-sequence*

Proof. Let $\mathcal{A}$ be a set of p-sequences then

1. $\forall A \in \mathcal{A}$ A is a p-sequence hence $p \in A$ so that $p \in \bigcap \mathcal{A}$
2. If $x \in \bigcap \mathcal{A}$ then $\forall A \in \mathcal{A}$ we have $p \in A$ which as A is a p-sequence gives that $\operatorname{succ}(x) \in A$ hence $\operatorname{succ}(x) \in \bigcap \mathcal{A}$
3. If $C \subseteq \bigcap \mathcal{A}$ is a chain then $\forall A \in \mathcal{A}$ we have $C \subseteq A$ and as A is a p-sequence we have that $\sup(C) \in A$ so that $\sup(A) \in \bigcap \mathcal{A}$

so by definition of a p-sequence we have that

$$\bigcap \mathcal{A} \text{ is a p-sequence}$$ □

From the above lemma [lemma: 3.111] we have that $\bigcap \{A \in \mathcal{P}(X) | A \text{ is a p-sequence}\}$ is a p-sequence and by definition $p \in \bigcap \{A \in \mathcal{P}(X) | A$ is a p-sequence$\}$. Further if A is a p-sequence then $\bigcap \{A \in \mathcal{P}(X) | A \text{ is a p-sequence}\} \subseteq A$. Summarized

$$P = \bigcap \{B \in \mathcal{P}(X) | B \text{ is a p-sequence }\} \text{ is a p-sequence} \wedge p \in P \wedge \text{If } A \text{ is a p-sequence} \Rightarrow P \subseteq A \tag{3.61}$$

Definition 3.112. *A element $x \in P$ is **select** if x is comparable with every element in P.*

Lemma 3.113. *If $x \in P$ is select then $\forall y \in P$ with $y < x$ have $\operatorname{succ}(y) \leqslant x$*

Proof. If $y \in P$ with $y < x$ then as P is a p-sequence we have by [definition: 3.109 (2)] that $\operatorname{succ}(y) \in P$. Now as x is select we have that $x, \operatorname{succ}(y)$ are comparable, hence by [theorem: 3.39] we have either $\operatorname{succ}(y) \leqslant x$ or $x < \operatorname{succ}(y)$. If $x < \operatorname{succ}(y)$ then from $y < x$ it follows that $y < x \wedge x < \operatorname{succ}(y)$ contradicting the fact that by [eq: 3.60] $\operatorname{succ}(y)$ is the immediate successor of y. Hence we must have that

$$\operatorname{succ}(y) \leqslant x$$ □

Lemma 3.114. *If x is select then $A_x = \{y \in P | y \leqslant x \vee \operatorname{succ}(x) \leqslant y\}$ is a p-sequence*

Proof.

1. As p is a least element of X we have that $p \leqslant x$ so that $p \in A_x$
2. Let $y \in A_x$ Then we have either:

   $y = x$. Then $\operatorname{succ}(x) = \operatorname{succ}(y) \Rightarrow \operatorname{succ}(x) \leqslant \operatorname{succ}(y)$ so that $\operatorname{succ}(y) \in A_x$.

   $y < x$. Then as $y \in A_x \subseteq P$ we have by the previous lemma [lemma: 3.113] that $\operatorname{succ}(y) < x \Rightarrow \operatorname{succ}(y) \leqslant x$ so that $\operatorname{succ}(y) \in A_x$.

   $\operatorname{succ}(x) \leqslant y$. As $\operatorname{succ}(y)$ is the immediate successor of y we have $y < \operatorname{succ}(y)$ so that $\operatorname{succ}(x) < \operatorname{succ}(y) \Rightarrow \operatorname{succ}(x) \leqslant \operatorname{succ}(y)$ proving that $\operatorname{succ}(y) \in A_x$.

   so in all cases we have

   $$\operatorname{succ}(y) \in A_x$$

3. If $C \subseteq A_x$ is a chain then we have the following excluding cases:

   $\exists y \in C$ with $\operatorname{succ}(x) \leqslant y$. Then as $y \leqslant \sup(C)$ we have that $\operatorname{succ}(x) \leqslant \sup(C)$ so that $\sup(C) \in A_x$.

   $\forall y \in C$ we have $\neg(\operatorname{succ}(x) \leqslant y)$. Now $\forall y \in C$ as $y \in C \subseteq A_x$ we have either $y \leqslant x$ or $\operatorname{succ}(y) \leqslant y$. As $\neg(\operatorname{succ}(x) \leqslant y)$ is true we must have $y \leqslant x$ and thus x is a upper bound of C. So by definition of the supremum as the least upper bound of C we must have that $\sup(C) \leqslant x$, hence $\sup(C) \in A_x$

   So in all cases we have

   $$\sup(C) \in A_x$$

From (1),(2) and (3) it follows then that

$$A_x \text{ is a p-sequence}$$

□

Corollary 3.115. *If x is select then $\forall y \in P$ we have $y \leqslant x$ or $\operatorname{succ}(x) \leqslant y$*

Proof. As A_x is a p-sequence by the previous lemma [lemma: 3.114] we have by [eq: 3.61] that $P \subseteq A_x$ and as by definition of A_x $A_x \subseteq P$ it follows that

$$P = A_x$$

□

Lemma 3.116. *The set $\{x \in X | x \text{ is select}\}$ is a p-sequence.*

Proof.

1. As p is a least element of X we have $\forall x \in P$ that $p \leqslant x$ so it is comparable with every element of p, hence p is select, so $p \in \{x \in X | s \text{ is select}\}$.
2. If $x \in \{x \in X | x \text{ is select}\}$ then x is select and by [corollary: 3.115] we have $\forall y \in P$ either:

   $y \leqslant x$. Then as $\operatorname{succ}(x)$ is the immediate successor of x we have $x < \operatorname{succ}(x)$ so that $y < \operatorname{succ}(x) \Rightarrow y \leqslant \operatorname{succ}(x)$ proving that $\operatorname{succ}(x)$ is comparable with y

   $\operatorname{succ}(x) \leqslant y$. Then $\operatorname{succ}(x)$ is comparable with y

   from the above it follows that $\operatorname{succ}(x)$ is comparable with every $y \in P$ hence

   $$\operatorname{succ}(x) \in \{x \in X | x \text{ is selected}\}$$

3. Let $C \subseteq \{x \in X | x \text{ is select}\}$ be a chain. Then as $C \subseteq X$ we have the hypothesis (3) that $\sup(C)$ exist. Then $\forall y \in P$ we have the following possibilities for C:

   $\exists x \in C$ with $y \leqslant x$. Then $x \leqslant \sup(C)$ so that $y \leqslant \sup(C)$ so that $\sup(C)$ is comparable with y

   $\forall x \in C$ we have $\neg(y \leqslant x)$. Then given $x \in C$ we have as $C \subseteq \{x \in X | x \text{ is select}\}$ that x is select. By [corollary: 3.115] we have either $y \leqslant x$ which is not allowed or $\operatorname{succ}(x) \leqslant y$. As $\operatorname{succ}(x)$ is a immediate successor of x we have $x < \operatorname{succ}(x)$ so that $x < y$ proving that y is a upper bound of C. Hence $\sup(C) \leqslant y$ proving that $\sup(C)$ is comparable with y

   So in all cases we have that $\sup(C)$ is comparable with y proving that $\sup(C)$ is select and thus that $\sup(C) \in \{x \in X | x \text{ is select}\}$

From (1),(2),(3) it follows then that $\{x\in X|x \text{ is select}\}$ is a p-sequence. $\square$

Now for the last corollary in the proof.

Corollary 3.117. *P is a chain*

Proof. As by the previous lemma [lemma: 3.116] $\{x\in X|x \text{ is select}\}$ is a p-sequence it follows from [eq: 3.61] that $P\subseteq\{x\in X|x \text{ is select}\}$. So if $x,y\in P$ then x is select and as $y\in P$ comparable with y, proving that P is a chain. $\square$

We are now finally able to reach a contradiction and prove the theorem. As P is a chain we have by hypothesis (2) that $\sup(P)$ exist. Now as P is a p-sequence [see eq: 3.61] we have by [definition: 3.109 (3)] that $\sup(P)\in P$ and by [definition: 3.109 (2)] that $\text{succ}(\sup(P))\in P$ so that $\text{succ}(\sup(P))\leqslant\sup(P)$. As $\text{succ}(\sup(P))$ is the immediate successor of $\sup(P)$ we have that $\sup(P)<\text{succ}(\sup(P))$. Hence $\sup(P)<\sup(P)$ which is a contradiction. $\square$

This was a long proof but it will be used in the following important theorem.

Definition 3.118. *A partial ordered set $\langle X,\leqslant\rangle$ is **Hausdorff maximal** if there exist a chain C such that if D is a chain with $C\subseteq D$ then $C=D$. In other words C is maximal when using the order relation defined by $\subseteq$.*

We show now that as a consequence of the Axiom of choice every partial ordered set is Hausdorff maximal.

Theorem 3.119. (Hausdorff's Maximal Theorem) *Let $\langle X,\leqslant\rangle$ be a partial ordered set then it is Hausdorff maximal. In other words there exists a chain C such that if D is a chain such that $C\subseteq D$ then $C=D$.*

Proof. Define the set of all chain of X

$$\mathcal{C}=\{A\in\mathcal{P}(X)|A \text{ is a chain in } \langle X,\leqslant\rangle\}$$

Using the fact $\mathcal{P}(X)$ is a set by the Axiom of Power Sets [axiom: 1.64] we have by the Axiom of Subsets [axiom: 1.54] and the fact that $\mathcal{C}\subseteq\mathcal{P}(X)$ it follows that

$$\mathcal{C} \text{ is a set} \tag{3.62}$$

Using [example: 3.33] we have that

$$\langle\mathcal{C},\preccurlyeq\rangle \text{ where } \preccurlyeq=\{(x,y)\in\mathcal{C}\times\mathcal{C}|x\subseteq y\} \text{ is a partial ordered set}$$

As $\forall A\in\mathcal{C}$ we have $\varnothing\subseteq A\Rightarrow\varnothing\preccurlyeq A$ and $\varnothing$ is a chain [see example: 3.42] in $\langle X,\leqslant\rangle$ it follows that

$$\mathcal{C} \text{ has a least element [using } \preccurlyeq] \tag{3.63}$$

Let $\mathcal{D}$ a chain in $\langle\mathcal{C},\preccurlyeq\rangle$ then if $x,y\in\bigcup\mathcal{D}$ there exists $A,B\in\mathcal{D}\subseteq\mathcal{C}$ such that $x\in A\wedge y\in B$ where A,B are chains in $\langle X,\leqslant\rangle$. As $\mathcal{D}$ is a chain we have either:

- ***$A\subseteq B$*.** Then $x,y\in B$ which as B is a chain [using $\leqslant$] means that x,y are comparable [using the order $\leqslant$]
- ***$B\subseteq A$*.** Then $x,y\in A$ which as A is a chain [using $\leqslant$] means that x,y are comparable [using the order $\leqslant$]

From the above it follows that $\bigcup\mathcal{D}$ is a chain in $\langle X,\leqslant\rangle$ hence $\bigcup\mathcal{D}\in\mathcal{C}$. Hence by [example: 3.70] it follows that $\bigcup\mathcal{D}=\sup(\mathcal{D})$ [using $\preccurlyeq$]. So we have proved that

$$\text{Every chain of } \langle\mathcal{C},\preccurlyeq\rangle \text{ has a supremum} \tag{3.64}$$

Now the conditions for [theorem: 3.108] are satisfied by [eq: 3.62], [eq: 3.63] and [eq: 3.64] so we have

$$\exists C\in\mathcal{C} \text{ [so } C \text{ is a chain in } \langle X,\leqslant\rangle] \text{ which has no immediate successor [using } \preccurlyeq] \tag{3.65}$$

Let now D be a chain in $\langle X,\leqslant\rangle$ [so that $D\in\mathcal{C}$] such that $C\subseteq D$. Take $d\in D$ and assume that $d\notin C$ then $C\subset C\bigcup\{d\}$ [as $C\bigcup\{d\}\nsubseteq C\Rightarrow C\neq C\bigcup\{d\}$] so that $C\prec C\bigcup\{d\}$. As C has no immediate successor [using $\prec$] there must be a $H\in\mathcal{C}$ such that $C\prec H\wedge H\prec C\bigcup\{d\}$ or $C\subset H\wedge H\subset C\bigcup\{d\}$. As $C\subset H$ there exists a $h\in H$ such that $h\notin C$, but then as $H\subset C\bigcup\{d\}$ we must have $h\in\{d\}$ or $h=d$, so $d\in H$. Now as $H\subset C\bigcup\{d\}$ there exists a $y\in C\bigcup\{d\}$ such that $y\notin H$, we can not have $y=d$ [as $d\in H$] so we must have $y\in C$ but then as $C\subset H$ we have $y\in H$ contradicting $y\in H$. So we must have $d\in C$. As $d\in D$ was chosen arbitrary we have that $D\subseteq C$ or $C=D$ which proves maximality. $\square$

We state now Zorn's lemma but not prove it yet, it will be show to be directly dependent on the Hausdorff maximal principle, which in turn depends on the Axiom of Choice. So if we accept the Axiom of Choice [which we do as it is a expressed as a Axiom] then Zorn's lemma applies.

Lemma 3.120. **(Zorn's Lemma)** *Let $\langle X, \leqslant \rangle$ be a partial ordered set such that every chain has a upper bound then X has a maximal element.*

We prove now that the Hausdorff Maximal principle implies Zorn's lemma.

Theorem 3.121. *Let $\langle X, \leqslant \rangle$ be Hausdorff Maximal then Zorn's lemma follows.*

Proof. Let $\langle X, \leqslant \rangle$ be a partial ordered set such that every chain in X has a upper bound. As $\langle X, \leqslant \rangle$ is Hausdorff maximal [definition: 3.118] there exist a chain C such that for every chain D with $C \subseteq D$ we have $C = D$. As C is a chain it has by the hypothesis a upper bound u for C. Assume now that u is not a maximal element of X, then by the definition of a maximal element [definition: 3.58] there exist a $x \in X$ with $u \leqslant x$ and $u \neq x$ so that $u < x$. If $x \in C$ then as u is a upper bound of C we have $x \leqslant u$ so that $u < u$ a contradiction. So we must have that $x \notin C$. Consider now $r, s \in C \bigcup \{x\}$ then we have to consider the following possibilities:

$r = x \wedge s = x$. Then by reflectivity we have $r \leqslant s$, so r, s are comparable.

$r = x \wedge s \neq x$. Then $s \in C$ so that $s \leqslant u$, which as $u \leqslant x$ proves that $s \leqslant x \underset{r=x}{\Rightarrow} s \leqslant r$, so r, s are comparable.

$r \neq x \wedge s = x$. Then $r \in C$ so that $r \leqslant u$, which as $u \leqslant x$ proves that $r \leqslant x \underset{s=x}{\Rightarrow} r \leqslant s$, so r, s are comparable.

$r \neq x \wedge s \neq x$. Then $r, s \in C$, which as C is a chain proves that r, s are comparable

From the above it follows that $C \bigcup \{x\}$ is a chain such that $C \subseteq C \bigcup \{x\}$ giving by maximality of C that $C = C \bigcup \{x\}$ contradicting $x \notin C$. Hence the assumption that u is not a maximal element of X is false. So u is a maximal element of X. □

We show now that Zorn's lemma implies well ordering.

Theorem 3.122. *Zorn's lemma implies that given a set X there exist a order relation $\leqslant$ on X such that $\langle X, \leqslant \rangle$ is well ordered [see 3.81]*

Proof. Just like the proof of [theorem: 3.108] this proof will consist of many sub lemma's.
Let X be a set and define the class

$$\mathcal{A} = \{(B, R) | B \in \mathcal{P}(A) \wedge R \text{ a order relation on } B \text{ so that } \langle B, R \rangle \text{ is well ordered}\}$$

Define now $\preccurlyeq \in \mathcal{A} \times \mathcal{A}$ by

$$\preccurlyeq = \{((B, R), (B', R')) | B \subseteq B' \wedge R \subseteq R' \wedge \text{If } x \in B \wedge y \in B' \backslash B \text{ then } (x, y) \in R'\}$$

then we have that

$$\langle \mathcal{A}, \preccurlyeq \rangle \text{ is a order relation} \tag{3.66}$$

Proof. We have to prove reflexivity, anti-symmetry and transitivity:

reflectivity. If $(B, R) \in \mathcal{A}$ then we have

1. $B \subseteq B$
2. $R \subseteq R$
3. If $x \in B \wedge y \in B \setminus B \underset{\text{[theorem: 1.32]}}{=} \varnothing$ which can not occur so that $(x, y) \in R$ is satisfied vacuously

proving that $(B, R) \preccurlyeq (B, R)$

anti-symmetry. If $(B, R) \preccurlyeq (B', R') \wedge (B', R') \preccurlyeq (B, R)$ then $B \subseteq B' \wedge R \subseteq R' \wedge B' \subseteq B \wedge R' \subseteq R$ proving that $B = B'$ and $R = R'$ so that $(B, R) = (B', R')$

transitivity. Let $(B, R) \preccurlyeq (B', R')$ and $(B', R') \preccurlyeq (B'', R'')$ then we have

1. $B \subseteq B' \wedge B' \subseteq B'' \Rightarrow B \subseteq B''$
2. $R \subseteq R' \wedge R' \subseteq R'' \Rightarrow R \subseteq R''$
3. If $x \in B \wedge y \in B'' \backslash B$ we have for y to consider the following possibilities

   $y \in B'$. Then $y \in B' \backslash B$ so that $(x, y) \in R' \underset{R' \subseteq R''}{\Rightarrow} (x, y) \in R''$

   $y \notin B'$. Then $y \in B'' \backslash B'$ so that $(x, y) \in R''$

   so in all cases we have $(x, y) \in R''$.

proving $(B,R)\preccurlyeq(B'',R'')$. □

We now have the following sub lemma:

Lemma 3.123. *If $\mathcal{C}\subseteq\mathcal{A}$ is a chain in $\langle\mathcal{A},\preccurlyeq\rangle$ then if*

$$B_{\mathcal{C}}=\bigcup\{B|\exists R \textit{ such that } (B,R)\in\mathcal{C}\}$$

$$R_{\mathcal{C}}=\bigcup\{R|\exists B \textit{ such that } (B,R)\in\mathcal{C}\}$$

then

$$(B_{\mathcal{C}},R_{\mathcal{C}})\in\mathcal{A}$$

Proof. First note that if $(B,R)\in\mathcal{C}$ then

$$B\in\{B|\exists R \text{ such that } (B,R)\in\mathcal{C}\}$$

and

$$R\in\{R|\exists B \text{ such that } (B,R)\in\mathcal{C}\}$$

or

$$\forall(B,R)\in\mathcal{C} \text{ we have } B\subseteq B_{\mathcal{C}}\wedge R\subseteq R_{\mathcal{C}} \tag{3.67}$$

1. If $x\in B_{\mathcal{C}}$ then $\exists(B,R)\in\mathcal{C}$ such that $x\in B$, as $\mathcal{C}\subseteq\mathcal{A}$ we have $(B,R)\in\mathcal{C}$, so that $B\in\mathcal{P}(A)$, hence $B\subseteq A$, proving that $x\in A$. In other words $B_{\mathcal{C}}\subseteq A$ or $B\in\mathcal{P}(A)$.
2. We must prove that $R_{\mathcal{C}}$ is a a order relation on $B_{\mathcal{C}}$:

   reflectivity. If $x\in B_{\mathcal{C}}$ then $\exists(B,R)\in\mathcal{C}$ such that $x\in B$, as R is a order relation we have that $(x,x)\in R$ so that by [eq: 3.67] $(x,x)\in R_{\mathcal{C}}$

   anti-symmetry. If $(x,y)\in R_{\mathcal{C}}\wedge(y,x)\in R_{\mathcal{C}}$ then $\exists(B,R),(B',R')\in\mathcal{C}$ such that $(x,y)\in R$ and $(y,x)\in R'$. As $\mathcal{C}$ is a chain we have either:

   $(B,R)\preccurlyeq(B',R')$. Then $R\subseteq R'$ so that $(x,y)\in R'\wedge(y,x)\in R'$, which as R' is a order relation proves that $x=y$.

   $(B',R')\preccurlyeq(B,R)$. Then $R'\subseteq R$ so that $(x,y)\in R\wedge(y,x)\in R$, which as R is a order relation proves that $x=y$.

   transitivity. If $(x,y)\in R_{\mathcal{C}}\wedge(y,z)\in R_{\mathcal{C}}$ then $\exists(B,R),(B',R')\in\mathcal{C}$ such that $(x,y)\in R$ and $(y,x)\in R'$. As $\mathcal{C}$ is a chain we have either:

   $(B,R)\preccurlyeq(B',R')$. Then $R\subseteq R'$ so that $(x,y)\in R'\wedge(y,z)\in R'$, which as R' is a order relation proves that $(x,z)\in R'$, hence $(x,z)\in R_{\mathcal{C}}$ [see eq: 3.67].

   $(B',R')\preccurlyeq(B,R)$. Then $R'\subseteq R$ so that $(x,y)\in R\wedge(y,z)\in R$, which as R is a order relation proves that $(x,z)\in R$, hence $(x,z)\in R_{\mathcal{C}}$ [see eq: 3.67].
3. Next we have to prove well ordering of $\langle B_{\mathcal{C}},R_{\mathcal{C}}\rangle$. Let $D\subseteq B_{\mathcal{C}}$ and $D\neq\varnothing$. Then there exist a $x\in D$ so that $x\in B_{\mathcal{C}}$, hence there exist a $(B,R)\in\mathcal{C}$ such that $x\in B$ or $x\in D\bigcap B$ proving that $D\bigcap B\neq\varnothing$. As $\mathcal{C}\subseteq\mathcal{A}$ we have by the definition of $\mathcal{A}$ that $\langle B,R\rangle$ is well ordered, hence there exist a least element $b\in B$. So

   $$\forall y\in B \text{ we have } (b,y)\in R \tag{3.68}$$

   We prove now that

   $$b \text{ is a least element of } D$$

   Proof. If $x\in D$ then $\exists(B',R')$ such that $x\in B'$. For x and B we the following possible cases:

   $x\in B$. Then by [eq: 3.68] we have that $(b,x)\in R$ so that by [eq: 3.67] $(b,x)\in R_{\mathcal{C}}$.

   $x\notin B$. Then $x\in B'\setminus B\wedge b\in B$. As $\mathcal{C}$ is a chain we have the following cases:

   $(B,R)\preccurlyeq(B',R')$. Then by definition of $\preccurlyeq$ we have $(b,x)\in R'$ so that by [eq: 3.67] $(b,x)\in R_{\mathcal{C}}$

   $(B',R')\preccurlyeq(B,R)$. Then $B'\subseteq B$ and as $x\in B'$ we have $x\in B$ contradicting $x\notin B$. So this case never occurs.

   So in all cases that apply we have $(b,x)\in R_{\mathcal{C}}$ proving that b is a least element of D. □

   As we have proved that every non empty $D\subseteq B_C$ has a least element [using the order $R_{\mathcal{C}}$ it follows that $\langle B_{\mathcal{C}},R_{\mathcal{C}}\rangle$ is well ordered.

From (1),(2) and (3) it follows that

$$(B_{\mathcal{C}}, R_{\mathcal{C}}) \in \mathcal{A}$$

□

Lemma 3.124. *If $\mathcal{C}$ is a chain in $\langle \mathcal{A}, \preccurlyeq \rangle$ then $(B_{\mathcal{C}}, R_{\mathcal{C}})$ is a upper bound of $\mathcal{C}$*

Proof. Let $(B, R) \in \mathcal{C}$ then

1. $B \subseteq B_{\mathcal{C}}$ [see eq: 3.67]
2. $R \subseteq R_{\mathcal{C}}$ [see eq: 3.67]
3. Let $x \in B$ and $y \in B_{\mathcal{C}} \setminus B$ then $\exists (B', R') \in \mathcal{C}$ such that $y \in B'$ or as $y \in B_{\mathcal{C}} \setminus B$ that

   $$y \in B' \setminus B$$

   As $\mathcal{C}$ is a chain we have either $(B, R) \preccurlyeq (B', R')$ or $(B', R') \preccurlyeq (B, R)$. If $(B', R') \preccurlyeq (B, R)$ then $B' \subseteq B$, as $y \in B'$ we would have $y \in B$ contradiction $y \in B_{\mathcal{C}} \setminus B$. So we have

   $$(B, R) \preccurlyeq (B', R')$$

   As $x \in B$ and $y \in B' \setminus B$ we have by definition of $\preccurlyeq$ and the above that $(x, y) \in R'$ which as $R' \subseteq R_C$ [see eq: 3.67] proves that $(x, y) \in R_{\mathcal{C}}$

So by the definition of $\preccurlyeq$ we have by (1),(2) and (3) that

$$(B, R) \preccurlyeq (B_{\mathcal{C}}, R_{\mathcal{C}})$$

□

Using Zorn's [lemma: 3.120] together with the above lemma [lemma: 3.124] we have

$$\exists (B_m, R_m) \in \mathcal{A} \text{ such that } (B_m, R_m) \text{ is a maximum element of } \mathcal{A} \tag{3.69}$$

We prove now by contradiction that

$$B_m = X$$

Proof. Assume that $X \neq B_m$. Then as $B_m \in \mathcal{P}(X) \Rightarrow B_m \subseteq X$ there exist a

$$x \in X \setminus B_m \Rightarrow x \notin B_m.$$

Define

$$R^* = R_m \bigcup \{(b, x) | b \in B_m\} \bigcup \{(x, x)\} \tag{3.70}$$

Then if $(r, s) \in R_m \bigcap \{(b, x) | b \in B_m\}$ we have as $R_m \subseteq B_m \times B_m$ that $s \in B_m \wedge s = x \notin B_m$ a contradiction, if $(r, s) \in R_m \bigcap \{(x, x)\}$ then $r \in B_m \wedge r = x \notin B_m$ a contradiction and finally if $(r, s) \in \{(b, x) | b \in B_m\} \bigcap \{(x, x)\}$ then $r \in B_m \wedge r = x \notin B_m$ a contradiction. So we have

$$R_m \bigcap \{(b, x) | b \in B_m\} = \varnothing \wedge R_m \bigcap \{(x, x)\} = \varnothing \wedge \{(b, x) | b \in B_m\} \bigcap \{(x, x)\} = \varnothing \tag{3.71}$$

Further if $(x, r) \in R^{\star}$ then we have either $(x, r) \in R_m \Rightarrow x \in B_m$ contradicting $x \notin B_m$, $(x, r) \in \{(b, x) | b \in B_m\} \Rightarrow x \in B_m$ contradicting $x \notin B_m$ or $(x, r) \in \{(x, x)\} \Rightarrow r = x$. To summarize we have

$$\text{If } (x, r) \in R^* \text{ then } r = x \tag{3.72}$$

We prove now that $\langle B_m \bigcup \{x\}, R^* \rangle$ is well ordered.

Proof. First we have:

reflexivity. If $r \in B_m \bigcup \{x\}$ then we have either:

> **$r \in B_m$.** Then as $\langle B_m, R_m \rangle$ is a partial order we have $(r, r) \in R_m \subseteq R^*$.
>
> **$r \notin B_m$.** Then $r \in \{x\}$ so that $r = x$ hence $(r, r) = (x, x) \in \{(x, x)\} \subseteq R^*$

proving that $(r, r) \in R^*$.

anti-symmetry. If $(r, s) \in R^*$ and $(s, r) \in R^*$ then we have by [eq: 3.70] for (r, s) either:

> **$(r, s) \in R_m$.** Then as $R_m \subseteq B_m \times B_m$ we have $r, s \in B_m$ so that $r \neq x \neq s$ so that $(s, r) \in R_m$ [if $(s, r) \in \{(b, x) | b \in B\} \bigcup \{(x, x)\}$ then $r = x$ contradicting $r \neq x$], which as $\langle B_m, R_m \rangle$ is a partial order gives that $r = s$.
>
> **$(r, s) \in \{(b, x) | b \in B_m\}$.** Then $s = x$ so that $(x, r) = (s, r) \in R^* \underset{\text{[eq: 3.72]}}{\Rightarrow} r = x = s$ hence $s = r$.

$(r,s)\in\{(x,x)\}$. Then $r=x=s\Rightarrow r=s$.

proving $r=s$

transitivity. If $(r,s)\in R^*\wedge(s,t)\in R^*$ then we have by [eq: 3.70] that:

$(r,s)\in R_m$. We have the following case for (s,t):

$(s,t)\in R_m$. Then as $\langle B_m,R_m\rangle$ is a partial ordered we have $(r,t)\in R_m\subseteq R^*$.

$(s,t)\in\{(b,x)|b\in B_m\}$. Then $t=x$ and $r\in B_m$ so that $(r,t)\in\{(b,x)|b\in B_m\}\subseteq R^*$.

$(s,t)\in\{(x,x)\}$. Then $t=x$ and $r\in B_m$ so that $(r,t)\in\{(b,x)|x\in B_m\}\subseteq R^*$.

$(r,s)\in\{(b,x)|b\in B_m\}$. Then $s=x$ so that $(s,t)=(x,t)\in R^*\underset{\text{[eq: 3.72]}}{\Rightarrow}t=x$. As $r\in B_m$ we have $(r,t)\in\{(b,x)|b\in B_m\}\subseteq R^*$.

$(r,s)\in\{(x,x)\}$. Then $r=x\wedge t=x$ so that $(x,t)=(s,t)\in R^*\underset{\text{[eq: 3.72]}}{\Rightarrow}t=x$ hence $(r,t)=(x,x)\in\{(x,x)\}\subseteq R^\star$.

proving $(r,t)\in R^*$.

Hence

$$\left\langle B_m\bigcup\{x\},R^*\right\rangle \text{ is partial ordered}$$

If $\varnothing\neq C\subseteq B_m\bigcup\{x\}$ is non empty then we have for $C\bigcap B_m$ the following possibilities:

$C\bigcap B_m\neq\varnothing$. Then as $\varnothing\neq C\bigcap B_m\subseteq B_m$ and $\langle B_m,R_m\rangle$ is well ordered [see definition of $\mathcal{A}$] there exist a least element $l\in C\bigcap B_m$ so

$$\forall r\in C\bigcap B_m \text{ we have } (l,r)\in R_m \tag{3.73}$$

Now if $r\in C$ we have either:

$r\in B_m$. then $r\in C\bigcap B_m$ so that by the above [eq: 3.73] $(l,r)\in R_m\subseteq R^*$

$r\notin B_m$. then as $C\subseteq B_m\bigcup\{x\}$ we have $r=x$ so $(l,r)\in\{(b,x)|b\in B_m\}\bigcup\{(x,x)\}\subseteq R^*$

proving that $(l,r)\in R^*$. Hence

$$C \text{ has a least element[using } \left\langle B\bigcup\{x\},R^*\right]$$

$C\bigcap B_m=\varnothing$. Then $C=\{x\}$ so that $\forall r\in C$ we have $r=x$ so that $(r,x)=(x,x)\in\{(x,x)\}\subseteq R^*$ proving that x is a least element of C.

So in all cases we have that C has a least element, hence

$$\left\langle B_m\bigcup\{x\},R^*\right\rangle \text{ is well ordered} \qquad \square$$

Now as $B_m\bigcup\{x\}\subseteq X$, we have by the definition of $\mathcal{A}$ and the above that

$$\left(B_m\bigcup\{x\},R^*\right)\in\mathcal{A}$$

Next we have:

1. $B_m\subseteq B_m\bigcup\{x\}$
2. $R_m\subseteq R^*$
3. If $r\in B_m$ and $s\in(B_m\bigcup\{x\})\setminus B_m$ then $s=x$ so that $(r,s)=(r,x)\in\{(b,x)|b\in B_m\}\subseteq R^*$

proving that $(B_m,R_m)\preccurlyeq(B_m\bigcup\{x\},R^*)$. As (B_m,R_m) is a maximal element of $\langle\mathcal{A},\preccurlyeq\rangle$ we must have $(B_m,R_m)=(B_m\bigcup\{x\},R^*)$ so that $B=B\bigcup\{x\}$ which as $x\notin B_m$ leads to a contradiction. Hence the assumption that $X\neq B_m$ is wrong and we must have that

$$X=B_m \qquad \square$$

As $\langle B_m,R_m\rangle$ is a well ordered the above proves that there exists a partial order R_m such that

$$\langle X,R_m\rangle=\langle B_m,R_m\rangle \text{ is well-ordered [by definition of } \mathcal{A}\ B_m \text{ is well ordered]} \qquad \square$$

We show now that Well Ordering implies the Axiom of Choice.

Theorem 3.125. *Assume that for every X there exist a order relation such that $\langle X,\leqslant\rangle$ is well ordered then there exists a function c: $\mathcal{P}'(X)\rightarrow X$ such that $\forall A\in\mathcal{P}'(X)$ we have $c(A)\in A$ (Axiom of Choice).*

Proof. Let X be a set then by the hypothesis there exist a order $\leqslant$ on X such that $\langle X, \leqslant\rangle$ is well ordered. Define now $c = \{(A, x) | A \in \mathcal{P}'(X) \wedge x$ is a least element of $A\}$. If $(A, x) \in c$ then $A \in \mathcal{P}'(X)$ and x is a least element of A, so that $x \in A \subseteq X$ proving that $(A, x) \in \mathcal{P}'(X) \times X$. So $c \subseteq \mathcal{P}'(X) \times X$. If $(A, x), (A, x') \in c$ then x and x' are least elements of A, which are unique by [theorem: 3.62] so that $x = x'$. Hence we have that

$$c: \mathcal{P}'(X) \rightarrow X \text{ is a partial function}$$

If $A \in \mathcal{P}'(X)$ then $A \neq \varnothing$ so by well ordering A has a least element l so that $(A, l) \in c$, so $\mathcal{P}'(A) \subseteq \mathrm{dom}(c)$. Hence by [proposition: 2.26] we have that

$$c: \mathcal{P}'(X) \rightarrow X \text{ is a function}$$

If $(A, x) \in c$ then x is the least element of A so that $c(A) = x \in A$ proving that

$$c: \mathcal{P}'(X) \rightarrow X \text{ is a choice function for } X$$

□

We are now ready to specify the different equivalent statements of the Axiom of Choice

Theorem 3.126. *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*

Proof.

- **1 ⇒ 2.** This follows from [theorem: 3.119]
- **2 ⇒ 3.** This follows from [theorem: 3.121]
- **3 ⇒ 4.** This follows from [theorem: 3.122]
- **4 ⇒ 1.** This follows from [theorem: 3.125] □

As in most of works about mathematics we assume the Axiom of Choice. To summarize the consequences of the Axiom of Choice we have [taking in account [theorem: 3.103] that the following statements are true.

Theorem 3.127.

Axiom of Choice. *Let X be a set then there exist a function $c: \mathcal{P}'(X) \rightarrow X$ such that $\forall A \in \mathcal{P}'(X)$ we have $c(A) \in A$.*

Existence of Choice set. *Let $\mathcal{A}$ be a set of sets such that*

a) $\forall A \in \mathcal{A}$ we have $A \neq \varnothing$

b) $\forall A, B \in \mathcal{A}$ with $A \neq B$ we have $A \bigcap B = \varnothing$

then there exist a set C [called the ***choice set of $\mathcal{A}$****] such that*

a) $C \subseteq \bigcup \mathcal{A}$

b) $\forall A \in \mathcal{A}$ we have $A \bigcap C \neq \varnothing$ and if $y, y' \in A \bigcap C$ then $y = y'$

Axiom of Choice alternative. *If $\{A_i\}_{i \in I} \subseteq \mathcal{A}$ is a family of non empty sets [$\forall i \in I$ we have $A_i \neq \varnothing$] where $I, \mathcal{A}$ are sets then there exists a function $f: I \rightarrow \bigcup_{i \in I} A_i$ such that $\forall i \in I$ we have $f(i) \in A_i$*

Hausdorff's Maximal Theorem. *If $\langle X, \leqslant\rangle$ is a partial ordered set then there exists a chain $C \subseteq X$ such that for every chain $D \subseteq X$ with $C \subseteq D$ we have $C = D$*

Zorn's Lemma. *If $\langle X, \leqslant\rangle$ is a partial ordered set such that every chain has a upper bound then X has a maximal element.*

Well-Ordering Theorem. *For every set there exists a order relation making $\langle X, \leqslant\rangle$ well-ordered.*

There is a kind of extension of Zorn's lemma to pre-ordered sets if we change the definition of maximal element slightly.

Theorem 3.128. *Let $\langle X, \leqslant\rangle$ be a pre-ordered set [see definitions: 3.26, 3.25] such that every chain has a upper bound then there exists a $m \in X$ such that $\forall x \in X$ with $m \leqslant x$ we have $x \leqslant m$*

Proof. Using [theorem: 3.34] we have the following

1. $\sim \subseteq X \times X$ defined by $\sim = \{(x, y) \in X | x \leqslant y \wedge y \leqslant x\}$ is a equivalence relation
2. Define $\preccurlyeq \subseteq (X/\sim) \times (X/\sim)$ by
$$\preccurlyeq = \{(x, y) \in (X/\sim) \times (X/\sim) | \exists x' \in \sim[x] \text{ and } \exists y' \in \sim[y] \text{ such that } x' \leqslant y'\}$$
then $\preccurlyeq$ is a order relation in $X/\sim$. So $\langle X/\sim, \preccurlyeq\rangle$ is a partial ordered set
3. $\forall x, y \in A$ we have $x \leqslant y \Leftrightarrow \sim[x] \preccurlyeq \sim[y]$

Let $C \subseteq X/\sim$ be a chain [using the order $\preccurlyeq$] and construct $C' = \bigcup C$. If $x, y \in C'$ then $\exists \sim[x'], \sim[y']$ such that $x \in \sim[x']$ and $y \in \sim[y']$, so $x \sim x'$ and $y \sim y'$ or $x \leqslant x' \wedge x' \leqslant x$ and $y \leqslant y' \wedge y' \leqslant y$. As C is a chain [using $\preccurlyeq$] we have the following possibilities:

$\sim[x'] \preccurlyeq \sim[y']$. then $x' \leqslant y'$ and as $x \leqslant x'$ and $y' \leqslant y$ we have $x \leqslant y$

$\sim[y'] \preccurlyeq \sim[x']$. then $y' \leqslant x'$ and as $y \leqslant y'$ and $x' \leqslant x$ we have $y \leqslant x$

proving that x, y are comparable. Hence

$$C' \text{ is a chain [using } \leqslant]$$

By the hypothesis we have that there exist a upper bound u of C' [using $\leqslant$], in other words

$$\exists u \in X \text{ such that } \forall x \in C' \text{ we have } x \leqslant u$$

Take now $\sim[z] \in C$ then $z \in \sim[z] \subseteq C'$ so that $z \leqslant u$ and thus by (3) $\sim[z] \preccurlyeq \sim[u]$. So $\sim[u]$ is a upper bound of C. As we just have proved that every chain in $X/\sim$ has a upper bound and $\langle X/\sim, \preccurlyeq\rangle$ is a partial order, it follows from Zorn's lemma that there exist a maximal element $\sim[m]$ in $X/\sim$. So by [definition: 3.58] we have

$$\forall \sim[x] \in X/\sim \text{ with } \sim[m] \preccurlyeq \sim[x] \text{ we have } \sim[x] = \sim[m]$$

If now $x \in X$ such that $x \leqslant m$ then by (3) we have $\sim[x] \preccurlyeq \sim[m]$ hence by the above we have $\sim[x] = \sim[m]$ so that $x \sim m$ hence $x \leqslant m$. □

From this point on we will gradually start to use the simpler notations for functions and families that are mentioned in the references [definition: 2.39], [theorem: 2.41], [theorem: 2.42], [theorem: 2.52], [theorem: 2.94], [notation: 2.95], [theorem: 2.116] and [theorem: 2.121] without explicit referring to them. This to avoid excessive notation and difference of notation between this text and standard mathematical practice.

As a interesting application of the Axiom of Choice we prove that every function can be restricted to a injection or bijection.

Theorem 3.129. *Let X, Y be sets, $f: X \to Y$ a function then there exist a $Z \subseteq X$ such that:*

1. *$f_{|Z}: Z \to Y$ is a injection*
2. *$f_{|Z}(Z) = f(X)$*
3. *$f_{|Z}: Z \to f(X)$ is a bijection*

Proof.

1. Define
$$\mathcal{A} = \{f^{-1}(\{y\}) | y \in f(X)\}.$$
If $A \in \mathcal{A}$ then $\exists y \in f(X)$ such that $A = f^{-1}(\{y\}) \subseteq X$ and as $y \in f(X)$ there exists a $x \in X$ such that $f(x) = y \in \{y\} \Rightarrow x \in f^{-1}(\{y\}) = A$, proving that $A \neq \varnothing$. So we have proved that
$$\mathcal{A} \subseteq \mathcal{P}'(X)$$
By the Axiom of Choice [axiom: 3.101] there exist a function
$$c: \mathcal{P}'(X) \to X \text{ such that } \forall A \in \mathcal{P}'(X)\ (c)(A) \in A$$
Take
$$Z = c(\mathcal{A}) \subseteq X$$
and consider the restriction of f to Z
$$f_{|Z}: Z \to Y$$

Let $x, y \in Z$ such that $f_{|Z}(x) = f_{|Z}(y) \underset{x,y \in Z}{\Rightarrow} f(x) = f(y)$. As $x, y \in Z = c(\mathcal{A})$ there exists $A_x \in \mathcal{A} \wedge A_y \in \mathcal{A}$ such that $x = (c)(A_x) \in A_x$ and $y = (c)(A_y) \in A_x$. As $A_x, A_y \in \mathcal{A}$ there exist $x', y' \in f(X)$ such that $A_x = f^{-1}(\{x'\})$ and $A_y = f^{-1}(\{y'\})$. Then $f(x) \underset{x \in A_x}{=} x'$ and $f(y) \underset{y \in A_y}{=} y'$. As $f(x) = f(y)$ we have $x' = y'$ so that $A_x = f^{-1}(\{x'\}) = f^{-1}(\{y'\}) = A_y$. So $x = (c)(A_x) = (c)(A_y) = y$, proving that $x = y$.

2. If $y \in f(X)$ then $f^{-1}(\{y\}) \in \mathcal{A}$ so to that x=$(c)(f^{-1}(\{y\})) \in c(\mathcal{A}) = Z$. Further as $(c)(f^{-1}(\{y\})) \in f^{-1}(\{y\})$ we have that $f(x) = f((c)(f^{-1}(\{y\}))) \in \{y\}$ so that $y = f(x) \in f(Z)$, proving that $f(X) \subseteq f(Z)$. As $Z \subseteq X$ we have by [theorem: 2.17] that $f(Z) \subseteq f(X)$ so that

$$f(X) = f(Z)$$

3. From (2) we have that $f_{|Z}: Z \rightarrow f(X)$ is surjective which together with (1) proves bijectivity. □

Another application of the Axiom of Choice is the following.

Theorem 3.130. *Let I, X, Y be sets, $f: X \rightarrow Y$ a function and $\{y_i\}_{i \in I} \subseteq f(X)$ then $\exists \{x_i\}_{i \in I} \subseteq X$ such that $\forall i \in I$ we have $f(x_i) = y_i$*

Proof. Define

$$\{A_i\}_{i \in I} \text{ by } A_i = f^{-1}(\{y_i\}) \subseteq X$$

Let $i \in I$ then, as $y_i \in f(X)$, there exist a $x \in X$ such that $f(x) = y_i$, hence $x \in f^{-1}(\{y_i\}) = A_i$. This proves that

$$\forall i \in I \; A_i \neq \varnothing$$

Using the Axiom of Choice [axiom: 3.103(3)] there exist a function

$$x: I \rightarrow \bigcup_{i \in I} A_i \text{ such that } \forall i \in I \text{ we have } x(i) \in A_i$$

If $i \in I$ we have that that $x(i) \in A_i$ so that $f(x(i)) \in \{y_i\} \Rightarrow f(x_i) = y_i$, this together with the fact that we can extend the function $x: I \rightarrow \bigcup_{i \in I} A_i$ to $f: I \rightarrow X$ proves that we have found a function

$$x: I \rightarrow X \text{ such that } \forall i \in I \text{ we have } f(x_i) = y_i$$

This function defines then the family $\{x_i\}_{i \in I} \subseteq X$ satisfying $\forall i \in I \; f(x_i) = y_i$. □

3.5 Generalized Intervals

Definition 3.131. *Let $\langle A, \leqslant \rangle$ be a partial ordered class then if $a, b \in A$ 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. $]-\infty, a] = \{x \in A | x \leqslant a\}$
6. $]-\infty, a[= \{x \in A | x < a\}$
7. $[a, \infty[= \{x \in A | a \leqslant x\}$
8. $]a, \infty[= \{x \in A | a < x\}$
9. $]-\infty, \infty[= A$

these are the intervals in A

Theorem 3.132. *Let $\langle A, \leqslant \rangle$ be a partial ordered class and $a, b \in A$ then we have:*

1. *If $b < a$ then $[a, b] = [a, b[=]a, b] =]a, b[= \emptyset$*
2. *If $b \leqslant a$ then $]a, b] = [a, b[=]a, b[= \emptyset$*

Proof.

1. Let $b<a$ then we have

   a. If $z\in[a,b]$ then $b<a\wedge a\leqslant z\wedge z\leqslant b$ so that $b<z\wedge z\leqslant b$ a contradiction.

   b. If $z\in]a,b]$ then $b<a\wedge a<z\wedge z\leqslant b$ so that $b<z\wedge z\leqslant b$ a contradiction.

   c. If $z\in[a,b[$ then $b<a\wedge a\leqslant z\wedge z<b$ so that $b<z\wedge z<b$ a contradiction.

   d. If $z\in[a,b]$ then $b<a\wedge a<z\wedge z\leqslant b$ so that $b<z\wedge z<b$ a contradiction.

2. Let $b\leqslant a$ then we have

   a. If $z\in[a,b[$ then $b\leqslant a\wedge a\leqslant z\wedge z<b$ so that $b\leqslant z\wedge z<b$ a contradiction.

   b. If $z\in]a,b]$ then $b\leqslant a\wedge a<z\wedge z\leqslant b$ so that $b<z\wedge z\leqslant b$ a contradiction.

   c. If $z\in]a,b[$ then $b\leqslant a\wedge a<z\wedge z<b$ so that $b<z\wedge z<b$ a contradiction. □

Definition 3.133. (Generalized Interval) *Let $\langle A,\leqslant\rangle$ be a partial ordered class then $I\subseteq A$ is a **generalized interval** if $\forall x,y\in I$ we have $[x,y]\subseteq I$.*

We have the following equivalent conditions for a generalized interval.

Theorem 3.134. *Let $\langle A,\leqslant\rangle$ be a partial ordered class that is conditional complete and $I\subseteq A$ then then*

$$I \text{ is a generalized interval} \Leftrightarrow \forall x,y\in I \text{ we have }]x,y[\subseteq I$$

Proof.

$\Rightarrow$. Let $x,y\in I$ then if $z\in]x,y[$ we have $x<z\wedge z<y\Rightarrow x\leqslant z\wedge z\leqslant y\Rightarrow z\in[x,y]\subseteq I$ proving that $]x,y[\subseteq I$

$\Leftarrow$. Let $x,y\in I$ then if $z\in[x,y]$ we have either

- **$x=y$.** Then $x\leqslant z\wedge z\leqslant x\Rightarrow z=x\in I$
- **$x<y$.** Then we have either
  - **$z=x$.** Then $z\in I$
  - **$z=y$.** Then $z\in I$
  - **$z\neq x\wedge z\neq y$.** Then $x<z\wedge z<y$ proving that $z\in]x,y[\subseteq I$

So in all cases we have that $z\in I$ proving that $[x,y]\subseteq I$ □

Theorem 3.135. *Let $\langle A,\leqslant\rangle$ be a fully ordered class that is conditional complete and $\emptyset\neq I\subseteq A$ a generalized interval then we have either*

1. *If I is bounded below and above [so that by conditional completeness $\inf(I)$ and $\sup(I)$ exists] we have:*

   a. *If $\inf(I)\in I\wedge\sup(I)\in I$ then $I=[\inf(I),\sup(I)]$*

   b. *If $\inf(I)\notin I\wedge\sup(I)\in I$ then $I=]\inf(I),\sup(I)]$*

   c. *If $\inf(I)\in I\wedge\sup(I)\notin I$ then $I=[\inf(I),\sup(I)[$*

   d. *If $\inf(I)\notin I\wedge\sup(I)\notin I$ then $I=]\inf(I),\sup(I)[$*

2. *If I is bounded below and not bounded above [so that by conditional completeness $\inf(I)$ exists] we have*

   a. *If $\inf(I)\in I$ then $I=[\inf(I),\infty[$*

   b. *If $\inf(I)\notin I$ then $I=]\inf(I),\infty[$*

3. *If I is bounded above and not bounded below [so that by conditional completeness $\sup(I)$ exists] we have*

   a. *If $\sup(I)\in I$ then $I=]-\infty,\sup(I)]$*

   b. *If $\sup(I)\notin I$ then $I=]-\infty,\sup(I)[$*

4. *If I is not bounded above and not bounded below we have $I=A$*

In other words a generalized interval I has either one of the following forms

$$\begin{array}{ll}[a,b] & a,b\in A \text{ and } a\leqslant b\\ ]a,b] & a,b\in A \text{ and } a<b\\ [a,b[& a,b\in A \text{ and } a<b\\ ]a,b[& a,b\in A \text{ and } a<b\\ [a,\infty[& \\ ]a,\infty[& \\ ]-\infty,a] & \\ ]-\infty,a[& \\ ]-\infty,\infty[& \end{array}$$

Proof.

1.

a. Let $z\in I$ then using the definition of a supremum and infinum as a upper and lowerbound that $\inf(I)\leqslant z\wedge z\leqslant\sup(I)$ so that

$$I\subseteq[\inf(I),\sup(I)] \tag{3.74}$$

As I is a generalized interval and $\sup(I)\in I\wedge\inf(I)\in I$ we have further that $[\inf(I),\sup(I)]\subseteq I$ which using 3.74 proves

$$I=[\inf(I),\sup(I)]$$

b. Let $z\in I$ then using the definition of a supremum and infinum as upper and lower bound and the fact that $\inf(I)\notin I$ we have that $\inf(I)<z\wedge z\leqslant\sup(I)$ so that

$$I\subseteq]\inf(I),\sup(I)] \tag{3.75}$$

If $z\in]\inf(I),\sup(I)]$ we have $\inf(I)<z\wedge z\leqslant\sup(I)$. Using [theorem: 3.71] there exists a $x\in I$ such that $\inf(I)\leqslant x<z$ or $z\in[x,\sup I]\subseteq I$ [as I is a generalized interval and $x,\sup(I)\in I$]. This proves that $]\inf(I),\sup(I)]\subseteq I$. Combining this with [eq: 3.75] proves

$$I=]\inf(I),\sup(I)]$$

c. Let $z\in I$ then using the definition of a supremum and infinum as upper and lower bound and the fact that $\sup(I)\notin I$ we have that $\inf(I)\leqslant z\wedge z<\sup(I)$ so that

$$I\subseteq[\inf(I),\sup(I)[\tag{3.76}$$

If $z\in[\inf(I),\sup(I)[$ we have $\inf(I)\leqslant z\wedge z<\sup(I)$. Using [theorem: 3.71] there exists a $y\in I$ such that $z<y\leqslant\sup(I)$ or $z\in[\inf(I),y]\subseteq I$ [as I is a generalized interval and $\inf(I),y\in I$]. This proves that $[\inf(I),\sup(I)[\subseteq I$. Combining this with [eq: 3.76] proves

$$I=[\inf(I),\sup(I)[$$

d. Let $z\in I$ then using the definition of a supremum and infinum as upper and lower bound and the fact that $\inf(I)\notin I\wedge\sup(I)\notin I$ we have that $\inf(I)<z\wedge z<\sup(I)$ so that

$$I\subseteq]\inf(I),\sup(I)[\tag{3.77}$$

If $z\in]\inf(I),\sup(I)[$ we have $\inf(I)<z\wedge z<\sup(I)$. Using [theorem: 3.71] there exists $x,y\in I$ such that $\inf(I)\leqslant x<z\wedge z<y\leqslant\sup(I)$ or $z\in]x,y[\subseteq I$ [as I is a generalized interval and $x,y\in I$]. This proves that $]\inf(I),\sup(I)[\subseteq I$. Combining this with [eq: 3.77] gives

$$I=]\inf(I),\sup(I)[$$

2.

a. Let $z\in I$ then using the definition of the infinum as a lower bound we have that $\inf(I)\leqslant z$ proving that

$$I\subseteq[\inf(I),\infty[\tag{3.78}$$

If $z\in[\inf(I),\infty[$ then $\inf(I)\leqslant z$. As I is not bounded above there exists a $x\in I$ such that $\neg(x\leqslant z)\underset{A\text{ is fully ordered}}{\Rightarrow}z<x$. Hence $z\in[\inf(I),x]\subseteq I$ [as I is a generalized interval and $\inf(I),x\in I$]. So $[\inf(I),\infty[\subseteq I$ and combining this with [eq: 3.78] proves

$$I=[\inf(I),\infty[$$

b. Let $z \in I$ then using the definition of the infinum as a lower bound and the fact that $\inf(I) \notin I$ we have that $\inf(I) < z$ proving that

$$I \subseteq]\inf(I), \infty[\tag{3.79}$$

If $z \in]\inf(I), \infty[$ then $\inf(I) < z$. Using [theorem: 3.71] there exists a $x \in I$ such that $\inf(I) \leqslant x < z$. As I is not bounded above there exists a $y \in I$ such that $\neg(y \leqslant z) \underset{A \text{ is fully ordered}}{\Rightarrow} z < y$. Hence $z \in]x, y[\subseteq I$ [as I is a generalized interval and x, $y \in I$]. So $]\inf(I), \infty[\subseteq I$ and combining this with [eq: 3.79] gives

$$I =]\inf(I), \infty[$$

3.

a. Let $z \in I$ then using the definition of the supremum as a upper bound we have $z \leqslant \sup(I)$ proving that

$$I \subseteq]-\infty, \sup(I)] \tag{3.80}$$

If $z \in]-\infty, \sup(I)]$ then $z \leqslant \sup(I)$. As I is not bounded below there exists a $x \in I$ such that $\neg(z \leqslant x) \underset{A \text{ is fully ordered}}{\Rightarrow} x < z$. Hence $z \in [x, \sup(I)] \subseteq I$ [as I is a generalized interval and $x, \sup(I) \in I$]. So $]-\infty, \sup(I)] \subseteq I$ which combined with [eq: 3.80] gives

$$I =]-\infty, \sup(I)]$$

b. Let $z \in I$ then using the definition of the supremum as a upper bound and the fact that $\sup(I) \notin I$ we have that $z < \sup(I)$ proving that

$$I \subseteq]-\infty, \sup(I)[\tag{3.81}$$

If $z \in]-\infty, \sup(I)[$ then $z < \sup(I)$. Using [theorem: 3.71] there exists a $y \in I$ such that $z < y \leqslant \sup(I)$. As I us not bounded below there exists a $x \in I$ such that $\neg(z \leqslant x) \underset{A \text{ is fully ordered}}{\Rightarrow} x < z$. Hence $z \in]x, y[\subseteq I$ [as I is a generalized interval and x, $y \in I$]. So $]-\infty, \sup(I)[\subseteq I$ which combined with [eq: 3.81] proves

$$I =]-\infty, \sup(I)[$$

4. If $z \in A$ then as I is not bounded below and not bounded above there exists $x, y \in I$ such that $x < z \wedge z < y$ giving $z \in]x, y[\subseteq I$ [as I is a generalized interval and $x, y \in I$]. This proves that $A \subseteq I$ and as trivially we have $I \subseteq A$ it follows that

$$I = A$$

□

We have also the opposite of the above theorem

Theorem 3.136. *Let $\langle A, \leqslant \rangle$ be a fully ordered class that is conditional complete then we have given $a, b \in A$ that*

1. *$[a, b]$ is a generalized interval*
2. *$[a, b[$ is a generalized interval*
3. *$]a, b]$ is a generalized interval*
4. *$]a, b[$ is a generalized interval*
5. *$[a, \infty[$ is a generalized interval*
6. *$]a, \infty[$ is a generalized interval*
7. *$]-\infty, a]$ is a generalized interval*
8. *$]-\infty, a[$ is a generalized interval*
9. *A is a generalized interval*

Proof.

1. If $x, y \in [a, b]$ and $z \in [x, y]$ then we have $x \leqslant z \underset{a \leqslant x}{\Rightarrow} a \leqslant z$ and $z \leqslant y \underset{y \leqslant b}{\Rightarrow} z \leqslant b$ proving that $z \in [a, b]$. Hence $[x, y] \subseteq [a, b]$ from which it follows that

$$[a, b] \text{ is a generalized interval}$$

2. If $x, y \in [a, b[$ and $z \in [x, y]$ then we have $x \leqslant z \underset{a \leqslant x}{\Rightarrow} a \leqslant z$ and $z \leqslant y \underset{y < b}{\Rightarrow} z < b$ proving that $z \in [a, b[$. Hence $[x, y] \subseteq [a, b[$ from which it follows that

$$[a, b[\text{ is a generalized interval}$$

3. If $x, y \in]a, b]$ and $z \in [x, y]$ then we have $x \leqslant z \underset{a < x}{\Rightarrow} a < z$ and $z \leqslant y \underset{y \leqslant b}{\Rightarrow} z \leqslant b$ proving that $z \in]a, b]$. Hence $[x, y] \subseteq]a, b]$ from which it follows that

$$]a, b] \text{ is a generalized interval}$$

4. If $x, y \in]a, b[$ and $z \in [x, y]$ then we have $x \leqslant z \underset{a < x}{\Rightarrow} a < z$ and $z \leqslant y \underset{y < b}{\Rightarrow} z < b$ proving that $z \in]a, b[$. Hence $[x, y] \subseteq]a, b[$ from which it follows that

$$]a, b[\text{ is a generalized interval}$$

5. If $x, y \in [a, \infty[$ and $z \in [x, y]$ we have $x \leqslant z \underset{a \leqslant x}{\Rightarrow} a \leqslant z$ proving that $z \in [a, \infty[$. Hence $[x, y] \subseteq [a, \infty[$ from which it follows that

$$[a, \infty[\text{ is a generalized interval}$$

6. If $x, y \in]a, \infty[$ and $z \in [x, y]$ we have $x \leqslant z \underset{a < x}{\Rightarrow} a < z$ proving that $z \in]a, \infty[$. Hence $[x, y] \subseteq]a, \infty[$ from which it follows that

$$]a, \infty[\text{ is a generalized interval}$$

7. If $x, y \in]-\infty, a]$ and $z \in [x, y]$ we have $z \leqslant y \underset{y \leqslant a}{\Rightarrow} z \leqslant a$ proving that $z \in]-\infty, a]$. Hence $[x, y] \subseteq]-\infty, a]$ from which it follows that

$$]-\infty, a] \text{ is a generalized interval}$$

8. If $x, y \in]-\infty, a[$ and $z \in [x, y]$ we have $z \leqslant y \underset{y < a}{\Rightarrow} z < a$ proving that $z \in]-\infty, a[$. Hence $[x, y] \subseteq]-\infty, a[$ from which it follows that

$$]-\infty, a[\text{ is a generalized interval}$$

9. If $x, y \in A$ then trivially $[x, y] \subseteq A$ hence A is a generalized interval. □

Theorem 3.137. *Let $\langle A, \leqslant \rangle$ be a totally ordered class then*

1. *$\forall a, b \in A$ we have if $]-\infty, a] =]-\infty, b]$ then $a = b$*
2. *$\forall a, b \in A$ we have if $]-\infty, a[=]-\infty, b[$ then $a = b$*
3. *$\forall a, b \in A$ we have if $[a, \infty[= [b, \infty[$ then $a = b$*
4. *$\forall a, b \in A$ we have if $]a, \infty[=]b, \infty[$ then $a = b$*
5. *$\forall a, b, c, d$ with $[a, b] \neq \emptyset$ then we have if $[a, b] = [c, d]$ then $a = c \wedge b = d$*
6. *$\forall a, b, c, d$ with $]a, b] \neq \emptyset$ then we have if $]a, b] =]c, d]$ then $a = c \wedge b = d$*
7. *$\forall a, b, c, d$ with $[a, b[\neq \emptyset$ then we have if $[a, b[= [c, d[$ then $a = c \wedge b = d$*
8. *$\forall a, b, c, d$ with $]a, b[\neq \emptyset$ then we have if $]a, b[=]c, d[$ then $a = c \wedge b = d$*

Proof.

1. As $a \in]-\infty, a] =]-\infty, b]$ we have $a \leqslant b$, further from $b \in]-\infty, b] =]-\infty, a]$ we have $b \leqslant a$, hence $a = b$.
2. As A is totally ordered we have for a, b either

   ***a* = *b*.** then (2) is proved

   ***a* < *b*.** then $a \in]-\infty, b[=]-\infty, a[$ giving the contradiction $a < a$

   ***b* < *a*.** then $b \in]-\infty, a[=]-\infty, b[$ giving the contradiction $b < b$

   so the only valid conclusion is that $a = b$
3. As $a \in [a, \infty[= [b, \infty[$ we have $b \leqslant a$, further from $b \in [b, \infty[= [a, \infty[$ we have $a \leqslant b$, hence $a = b$
4. As A is totally ordered we have for a, b either

   ***a* = *b*.** then (4) is proved

   ***a* < *b*.** then $b \in]a, \infty[=]b, \infty[$ giving the contradiction $b < b$

   ***b* < *a*.** then $a \in]b, \infty[=]a, \infty[$ giving the contradiction $a < a$

so the only valid conclusion is that $a=b$

5. As $[a,b]=[c,d]\neq\varnothing$ it follows from [theorem: 3.132] that $a\leqslant b$ and $c\leqslant d$ so that $a,b\in[a,b]=[c,d]\Rightarrow c\leqslant a\wedge b\leqslant d$ and $c,d\in[c,d]=[a,b]\Rightarrow a\leqslant c\wedge d\leqslant b$ proving that $a=c\wedge b=d$.

6. As $]a,b]=]c,d]\neq\varnothing$ it follows from [theorem: 3.132] that $a<b\wedge c<d$ hence $b\in]a,b]\wedge d\in]c,d]$. From $b\in]a,b]=]c,d]$ we have $b\leqslant d$ and from $d\in]c,d]=]a,b]$ we have $d\leqslant b$ hence we have

$$b=d$$

Assume that $a<c$ then as $c<d=b$ we have $c\in]a,b]=]c,d]$ so that $c<c$ a contradiction, hence we must have

$$c\leqslant a \tag{3.82}$$

Assume that $c<a$ then as $a<b=d$ we have $a\in]c,d]=]a,b]$ so that $a<a$ a contradiction hence we must have $a\leq c$ which combined with [eq: 3.82] proves that

$$a=c$$

7. As $[a,b[=[c,d[\neq\varnothing$ we have by [theorem: 3.132] that $a<b\wedge c<d$ so that $a\in[a,b[\wedge c\in[c,d[$. From $a\in[a,b[=[c,d[$ it follows that $c\leqslant a$ and from $c\in[c,d[=[a,b[$ we have $a\leqslant c$ hence we have

$$a=c$$

Assume that $b<d$ then as $c=a<b$ we have $b\in[c,d[=[a,b[$ so that $b<b$ a contradiction, hence

$$b\leqslant d \tag{3.83}$$

Assume that $d<b$ then as $a=c<d<b$ we have $d\in[a,b[=[c,d[$ so that $d<d$ a contradiction, hence $d\leqslant b$ which combined with [eq: 3.83] proves

$$b=d$$

8. As $\emptyset\neq]a.b[=]c,d[$ there exists a $x\in A$ such that $x\in]a,b[=]c,d[$ giving

$$a<x<b\wedge c<x<d \tag{3.84}$$

Assume that $a<c$ then by [eq: 3.84] $a<c<x<b$ so that $c\in]a,b[=]c,d[$ leading to the contradiction $c<c$ hence

$$c\leqslant a \tag{3.85}$$

Assume that $c<a$ then by [eq: 3.84] $c<a<x<d$ so that $a\in]c,d[=]a,b[$ leading to the contradiction $a<a$ hence $a\leqslant c$ which combined with [eq: 3.85] gives

$$a=c$$

Assume that $b<d$ then by [eq: 3.84] $c<x<b<d$ so that $b\in]c,d[=]a,b[$ leading to the contradicton $b<b$, hence

$$d\leqslant b \tag{3.86}$$

Assume that $d<b$ then by [eq: 3.84] $a<x<d<b$ so that $d\in]a,b[=]c,d[$ leading to the contradiction $d<d$ proving that $b\leqslant d$. Combining this with [eq: 3.86] results in

$$b=d$$

□

Chapter 4

Algebraic constructs

Before we define the different number systems, like the natural numbers, whole numbers, rational numbers, real numbers and complex numbers, we define the algebraic operations and structures that we can define on them. In this way we abstract away the algebraic operations and algebraic structures. First we define the concept of a operator which is short notation for the application of a function with two arguments between a set and itself.

Definition 4.1. (Operator) *Let X be a set then a **operator** is function*

$$f: X \times X \to X$$

To avoid using excessive notation we use infix notation instead of the classic function call notation, so

$$f(x, y) \text{ is noted as } x f y$$

4.1 Groups

Definition 4.2. *A semi-group is a pair $\langle G, \odot \rangle$ where G is a set and $\odot$ a operator $\odot: G \times G \to G$ such that:*

neutral element. *$\exists e \in G$ such that $\forall x \in G$ we have $x \odot e = x = e \odot x$*

associativity. *$\forall x, y, z \in G$ we have $(x \odot y) \odot z = x \odot (y \odot z)$*

Theorem 4.3. *If $\langle G, \odot \rangle$ is a semi-group then*

1. *$G \neq \varnothing$*
2. *G has only one neutral element*

Proof.

1. As G is a group there exist a neutral element $e \in G$ so that $G \neq \varnothing$
2. Assume that there exists two neutral elements e, e' then we have

$$e \underset{e' \text{ is neutral element}}{=} e \odot e' \underset{e \text{ is neutral element}}{=} e'$$ □

Example 4.4. Let X be a set then $\langle X^X, \circ \rangle$ is a semi group [see definition: 2.30]. Here X^X is the set of function graphs between X and X and $\circ$ is the composition between functions.

Proof. As X is a set we have by [theorem: 2.35] that X^X is a set. Further if $f, g \in X^X$ then $f: X \to$ and $g: X \to X$ are functions, so that by [theorem: 2.28] $f \circ g: X \to X$ is a function, hence $f \circ g \in X^X$. So

$$\circ: X^X \times X^X \to X^X \text{ defined by } \circ(f, g) = f \circ g$$

is a function. The neutral element is Id_X because $\forall f \in X^X$ we have

$$f \circ \mathrm{Id}_X \underset{[\text{theorem: } 2.48}{=} f \underset{[\text{theorem: } 2.48}{=} \mathrm{Id}_X \circ f$$

□

A group is a semi-group with the extra condition that is has a inverse element.

Definition 4.5. *A **group** $\langle X, \odot\rangle$ is a semi-group with the extra condition*

Inverse Element. $\forall x \in G$ *there* $\exists y \in G$ *such that*

$$x \odot y = e = y \odot x$$

where e is the neutral element of the group.

One benefit that a group has is the canceling property

Theorem 4.6. *If $x, y, z \in \langle G, \odot\rangle$ then $x \odot z = y \odot z$ then $x = y$*

Proof. We have

$$\begin{array}{lcl} x \odot z = y \odot z & \Rightarrow & (x \odot z) \odot z^{-1} = (y \odot z) \odot z^{-1} \\ & \underset{\text{associativity}}{\Rightarrow} & x \odot (z \odot z^{-1}) = y \odot (z \odot z^{-1}) \\ & \underset{\text{inverse element}}{\Rightarrow} & x \odot e = y \odot e \\ & \Rightarrow & x = y \\ & & \square \end{array}$$

Theorem 4.7. *If $\langle G, \odot\rangle$ is group then every element has a unique inverse element. So*

$$\forall x \in G\ \exists! y \in G \text{ such that } x \odot y = x = y \odot x$$

this unique element is noted as x^{-1} [or sometimes as $-x$]

Proof. Let $x \in G$ and assume that y, y' are inverse elements for x then we have

$$x \odot y = e = y \odot x \text{ and } x \odot y' = e = y' \odot x$$

So that

$$y = y \odot e = y \circ (x \odot y') = (y \odot x) \odot y' = e \odot y' = y'$$ $\square$

Theorem 4.8. *If $\langle G, \odot\rangle$ is a group then $\forall x, y \in G$ we have $(x \odot y)^{-1} = y^{-1} \odot x^{-1}$*

Proof. We have

$$\begin{array}{rcl} (x \odot y) \odot (y^{-1} \odot x^{-1}) & = & x \odot (y \odot (y^{-1} \odot x^{-1})) \\ & = & x \odot ((y \odot y^{-1}) \odot x^{-1}) \\ & = & x \odot (e \odot x^{-1}) \\ & = & x \odot x^{-1} \\ & = & e \\ (y^{-1} \odot x^{-1}) \odot (x \odot y) & = & y^{-1} \odot (x^{-1} \odot (x \odot y)) \\ & = & y^{-1} \odot ((x^{-1} \odot x) \odot y) \\ & = & y^{-1} \odot (e \odot y) \\ & = & y^{-1} \odot y \\ & = & e \\ & & \square \end{array}$$

Theorem 4.9. *If $\langle G, \odot\rangle$ is a group then $\forall x \in G$ we have $(x^{-1})^{-1} = x$ and $e^{-1} = e$ where e is the neutral element.*

Proof. If $x \in G$ then $x \odot x^{-1} = e = x^{-1} \odot x$ and $(x^{-1})^{-1} \odot x^{-1} = e = x^{-1} \odot (x^{-1})^{-1}$. So

$$\begin{array}{rcl} x & = & x \odot e \\ & = & x \odot (x^{-1} \odot (x^{-1})^{-1}) \\ & = & (x \odot x^{-1}) \odot (x^{-1})^{-1} \\ & = & e \circ (x^{-1})^{-1} \\ & = & (x^{-1})^{-1} \end{array}$$

Further

$$e^{-1} = e \cdot e^{-1} = e$$

□

Theorem 4.10. *If $\langle G, \odot \rangle$ then $\forall x, y \in X$ we have $x = y \Leftrightarrow x^{-1} = y^{-1}$ [and by contra position $x \neq y \Leftrightarrow x^{-1} \neq y^{-1}$]*

Proof.

⇒. $e = x^{-1} \cdot x = x^{-1} \cdot y$ and $e = x \cdot x^{-1} = y \cdot x^{-1}$ proving by uniqueness of the inverse [see theorem: 4.7] that $y^{-1} = x^{-1}$

⇐. If $x^{-1} = y^{-1}$ then by the above we have $(x^{-1})^{-1} = (y^{-1})^{-1}$ it follows from [theorem: 4.9] that $x = y$.

□

Definition 4.11. *A semi-group or group $\langle G, \odot \rangle$ is Abelian or **commutative** iff*

$$\forall x, y \in G \text{ we have } x \odot y = y \odot x$$

Definition 4.12. *Let $\langle G, \odot \rangle$ be a semi-group then $F \subseteq G$ is a sub-semi-group iff*

1. *$\forall x, y \in F$ we have $x \odot y \in F$*
2. *$e \in F$ [e is the neutral element of G]*

Definition 4.13. *Let $\langle G, \odot \rangle$ be groups then $F \subseteq G$ is a sub-group iff*

1. *$\forall x, y \in F$ we have $x \odot y \in F$*
2. *$e \in F$ [e is the neutral element of G]*
3. *$\forall x \in F$ we have $x^{-1} \in F$*

The following show how sub-semi-groups and sub-groups can be used to reduce the work for proving the group axioms.

Theorem 4.14. *Let $\langle G, \odot \rangle$ be a semi-group and $F \subseteq G$ a sub-semi-group then*

1. *$\langle F, \odot_{|F \times F} \rangle$ is a semi group with the same neutral element as $\langle G, \odot \rangle$*
2. *If $\langle G, \odot \rangle$ is Abelian then $\langle F, \odot_{|F \times F} \rangle$ is Abelian*

To avoid excessive notation we use $\odot$ instead of $\odot_{|F \times F}$ if it is clear from the context which operation should be used.

Proof. First as G is a set we have by the Axiom of Subsets [axiom: 1.54] that G is a set.

1. For $\langle F, \odot_{|F \times F} \rangle$

   neutral element. By definition of a subgroup $e \in F$. Let $x \in F$ then

   $$e \odot_{|F \times F} x \underset{e, x \in F}{=} e \odot x = x = x \odot e = x \odot_{|F \times F} e$$

   associativity. Let $x, y, z \in F$ then

   $$(x \odot_{|F \times F} y) \odot_{|F \times F} z = (x \circ y) \circ z = x \circ (y \circ z) = x \odot_{|F \times F} (y \odot_{|F \times F} z)$$

2. Let $x, y \in F$ then

   $$x \odot_{|F \times F} y = x \odot y = y \odot x = y \odot_{|F \times F} x$$

□

Theorem 4.15. *Let $\langle G, \odot \rangle$ be a semi-group $F \subseteq G$ a sub semi-group of $\langle G, \odot \rangle$ and $H \subseteq F$ a sub semi-group of $\langle G, \odot_{|F \times F} \rangle$ then H is a sub semi-group of $\langle G, \odot \rangle$*

Proof.

1. $\forall x, y \in H \subseteq F$ we have $x \odot_{|F \times F} y \in H$ which as $(x, y) \in F \times F$ proves that $x \odot y = x \odot_{|F \times F} y \in H$
2. if e is the neutral element of $\langle G, \odot \rangle$ then by [theorem: 4.14] e is also the neutral element of F, hence $e \in H$.

□

Theorem 4.16. *Let $\langle G, \odot \rangle$ be a group and $F \subseteq G$ a sub-group then*

1. *$\langle F, \odot_{|F \times F} \rangle$ is a group with same neutral element as $\langle G, \odot \rangle$ and for every $x \in F$ it's inverse element in $\langle G, \odot \rangle$ is also the inverse element in $\langle F, \odot_{|F \times F|} \rangle$.*

2. If $\langle G, \odot \rangle$ is Abelian then $\langle F, \odot_{|F \times F} \rangle$ is Abelian

To avoid excessive notation we use $\odot$ instead of $\odot_{|F \times F}$ if it is clear from the context which operation should be used.

Proof.

1. For $\langle F, \odot_{|F \times F} \rangle$ we have

   neutral element. Let $x \in F$ then $e \odot_{|F \times F} x \underset{e, x \in F}{=} e \odot x = x = x \odot e = x \circ,_{|F \times F} e$

   associativity. Let $x, y, z \in F$ then

   $$(x \odot_{|F \times F} y) \odot_{|F \times F} z = (x \circ y) \circ z = x \circ (y \circ z) = x \odot_{|F \times F} (y \odot_{|F \times F} z)$$

   inverse element. Let $x \in F$ then also $x^{-1} \in F$ then

   $$(x \odot_{|F \times F} x^{-1}) = x \odot x^{-1} = e = x^{-1} \odot c = x^{-1} \odot_{|F \times F} x$$

2. Let $x, y \in F$ then

   $$x \odot_{|F \times F} y = x \odot y = y \odot x = y \odot_{|F \times F} x$$

□

Theorem 4.17. *Let $\langle G, \odot \rangle$ be a [semi-]group $F \subseteq G$ a sub [semi-]group of $\langle G, \odot \rangle$ and $H \subseteq F$ a sub [semi-]group of $\langle G, \odot_{|F \times F} \rangle$ then H is a sub [semi-]group of $\langle G, \odot \rangle$*

Proof.

1. $\forall x, y \in H \subseteq F$ we have $x \odot_{|F \times F} y \in H$ which as $(x, y) \in F \times F$ proves that $x \odot y = x \odot_{|F \times F} y \in H$.
2. If e is the neutral element of $\langle G, \odot \rangle$ then by [theorem: 4.16] e is also the neutral element of F, hence $e \in H$.
3. If $x \in H$ then $x \in G$ and by [theorem: 4.16 it's inverse element x^{-1} in $\langle G, \odot \rangle$ is also its inverse element in $\langle F, \odot_{|F \times F} \rangle$ hence $x^{-1} \in H$.

□

Example 4.18. Let X be a set, $\langle X^X, \circ \rangle$ the semi-group used in [example: 4.4] then $\langle \mathcal{B}[X], \circ \rangle$ is a group where $\mathcal{B}[X] = \{f \in X^X | f: X \rightarrow X \text{ is a bijection}\}$.

Proof. First we prove that $\mathcal{B}[X]$ is a sub-semi-group

1. $\forall f, g \in \mathcal{B}[X]$ we have that $f: X \rightarrow X$ and $g: X \rightarrow X$ are bijections so that by [theorem: 2.74] $f \circ g$ is a bijection so that $f \circ g \in \mathcal{B}[X]$
2. $\mathrm{Id}_X: X \rightarrow X$ is by [theorem: 2.64] a bijection so that $\mathrm{Id}_X \in \mathcal{B}[X]$

Applying then [theorem: 4.14] proves that

$$\langle \mathcal{B}[X], \circ \rangle \text{ is a semi-group}$$

Let $f \in \mathcal{B}[X]$ then $f: X \rightarrow X$ is a bijection and by [theorems: 2.69,2.72] we have that $f^{-1}: X \rightarrow X$ is a bijection, so that $f^{-1} \in \mathcal{B}[X]$ and $f \circ \mathrm{Id}_X = f = \mathrm{Id}_X \circ f$.

□

Definition 4.19. (Group Homeomorphism) *If $\langle F, \odot \rangle$ and $\langle G, \oplus \rangle$ be semi-groups then a function $f: F \rightarrow G$ is a **group homeomorphism between** $\langle F, \odot \rangle$ **and** $\langle G, \oplus \rangle$ iff $\forall x, y \in F$ we have $f(x \odot y) = f(x) \oplus g(y)$.*

***Notation* 4.20.** *We use the following notation for a group homeomorphism between $\langle F. \odot \rangle$ and $\langle G, \oplus \rangle$*

$$f: \langle F, \odot \rangle \rightarrow \langle G, \oplus \rangle \textit{ is a group homeomorphism}$$

Theorem 4.21. *If $\langle F, \odot \rangle$ and $\langle G, \oplus \rangle$ are semi-groups, $H \subseteq G$ a sub-semi-group of $\langle G, \oplus \rangle$ and*

$$f: \langle F, \odot \rangle \rightarrow \langle H, \oplus_{|H \times H} \rangle \textit{ is a group homeomorphism}$$

then

$$f: \langle F, \odot \rangle \rightarrow \langle G, \oplus \rangle \textit{ is a group homeomorphism}$$

Proof. Let $x, y \in F$ then we have

$$f(x \odot y) = f(x) \oplus_{|H \times H} f(y) = f(x) \oplus f(y)$$

□

Theorem 4.22. *If $\langle F,\odot\rangle$ and $\langle G,\oplus\rangle$ be semi groups with neutral elements e_F, e_G and $f: F\rightarrow G$ a* ***group homeomorphism*** *then:*

1. $f(e_F)=e_G$
2. *If $\langle F,\odot\rangle$ and $\langle G,\oplus\rangle$ are groups then $\forall x\in F$ we have $f(x^{-1})=f(x)^{-1}$*
3. *$f(F)$ is a sub-[semi-]group of $\langle G,\oplus\rangle$ if $\langle F,\odot\rangle$ is a [semi-]group*

Proof.

1.

$$\begin{aligned} e_G &= f(e_F)^{-1}\oplus f(e_F) \\ &= f(e_F)^{-1}\oplus f(e_F\odot e_F) \\ &= f(e_F)^{-1}\oplus (f(e_F)\oplus f(e_F)) \\ &= (f(e_F)^{-1}\oplus f(e_F))\oplus f(e_F) \\ &= e_G\oplus f(e_F) \\ &= f(e_F) \end{aligned}$$

2. If $x\in F$ then

$$f(x^{-1})\oplus f(x)=f(x^{-1}\odot x)=f(e_F)\underset{(1)}{=}e_G$$

and

$$f(x)\oplus f(x^{-1})=f(x\odot x^{-1})=f(e_F)\underset{(1)}{=}e_G$$

so that $f(x)^{-1}=f(x^{-1})$

3. If $x,y\in f(F)$ then their exists $u,v\in F$ such that $x=f(u)$ and $y=f(v)$, then we have

$$x+y=f(u)\oplus f(v)=f(u\odot v)\in f(F)$$

Also

$$e_G\underset{(1)}{=}f(e_F)\in f(F)$$

Finally if $\langle F,\odot\rangle$ and $\langle G,\oplus\rangle$ are groups and $x\in f(F)$ then there exists a $u\in F$ such that $x=f(u)$, then we have

$$x^{-1}\underset{(2)}{=}f(u)^{-1}=f(u^{-1})\in f(F)$$ □

Definition 4.23. (Group Isomorphism) *If $\langle F,\odot\rangle$ and $\langle G,\oplus\rangle$ are semi-groups then a* ***group isomorphism*** *is a* ***bijection*** *$f: F\rightarrow G$ such that*

$$f:\langle F,\odot\rangle\rightarrow\langle G,\oplus\rangle \text{ is a group homeomorphism}$$

Theorem 4.24. *Let $\langle F,\odot\rangle$, $\langle G,\oplus\rangle$ be semi groups and*

$$f:\langle F,\odot\rangle\rightarrow\langle G,\oplus\rangle \text{ is a group isomorphism}$$

then

$$f^{-1}:\langle G,\oplus\rangle\rightarrow\langle F,\odot\rangle \text{ is a group isomorphism}$$

Proof. As $f: F\rightarrow G$ is a bijection we have by [theorem: 2.72] that $f^{-1}: G\rightarrow F$ is a bijection. Take $x,y\in G$ then we have

$$\begin{aligned} f^{-1}(x\oplus y) &= f^{-1}(\mathrm{Id}_G(x)\oplus \mathrm{Id}_G(y)) \\ &\underset{\text{[theorem: 2.69}}{=} f^{-1}((f\circ f^{-1})(x)\oplus (f\circ f^{-1})(y)) \\ &\underset{\text{[theorem: 2.42]}}{=} f^{-1}(f(f^{-1}(x))\oplus f(f^{-1}(y))) \\ &\underset{f\text{ is homeomorphism}}{=} f^{-1}(f(f^{-1}(x)\odot f^{-1}(y))) \\ &\underset{\text{[theorem: 2.42]}}{=} (f^{-1}\circ f)(f^{-1}(x)\odot f^{-1}(y)) \\ &\underset{\text{[theorem: 2.69}}{=} \mathrm{Id}_F(f^{-1}(x)\odot f^{-1}(y)) \\ &= f^{-1}(x)\odot f^{-1}(y) \end{aligned}$$

Further if e_F, e_G are the neutral elements of $\langle F, \oplus_F\rangle$, $\langle G, \oplus_G\rangle$ then

$$\begin{aligned} e_F &= \mathrm{Id}_F(e_f) \\ &= (f^{-1}\circ f)(e_F) \\ &= f^{-1}(f(e_F)) \\ &= f^{-1}(e_G) \end{aligned}$$

proving that

$$f^{-1}: F \rightarrow G \text{ is a group isomorphism}$$

□

Theorem 4.25. *If $\langle A, \oplus_A\rangle$, $\langle B, \oplus_B\rangle$ and $\langle C, \oplus_C\rangle$ are [semi-]groups then*

1. *If D is a sub [semi-]group of $\langle B, \oplus_B\rangle$ and*

   $$f: \langle A, \oplus_A\rangle \rightarrow \langle D, \oplus_B\rangle \text{ and } g: \langle B, \oplus_B\rangle \rightarrow \langle C, \oplus_C\rangle \text{ are group homeomorphism}$$

   then

   $$g\circ f: \langle A, \oplus_A\rangle \rightarrow \langle C, \oplus_C\rangle \text{ is a group homeomorphism}$$

   and

   $$g(f(A)) \text{ is a sub [semi-]group of } \langle C, \oplus_C\rangle$$

2. *If D is a sub group of $\langle B, \oplus_B\rangle$ and*

   $$f: \langle A, \oplus_A\rangle \rightarrow \langle D, \oplus_B\rangle \text{ and } g: \langle B, \oplus_B\rangle \rightarrow \langle C, \oplus_C\rangle \text{ are group isomorphisms}$$

   then

   $$g\circ f: \langle A, \oplus_A\rangle \rightarrow \langle g(f(A)), \oplus_C\rangle \text{ is a group isomorphism}$$

   or as $g(D) \underset{f:A\rightarrow D\text{ is injective}}{=} g(f(A))$ that

   $$g\circ f: \langle A, \oplus_A\rangle \rightarrow \langle g(D), \oplus_C\rangle \text{ is a group isomorphism}$$

Proof.

1. Let $x, y \in A$ then we have

   $$\begin{aligned} (g\circ f)(x\oplus_A y) &= g(f(x\oplus_A y)) \\ &\underset{f\text{ is a homeomorphism}}{=} g(f(x)\oplus_B f(y)) \\ &\underset{g\text{ is a homeomorphism}}{=} g(f(x))\oplus_C g(f(y)) \\ &= (g\circ f)(x)\oplus_C (g\circ f)(y) \end{aligned}$$

   proving that $g\circ f$ is a group homeomorphism. Finally using [theorem: 4.22] we have then that

   $$g(f(A)) \text{ is a sub group}$$

2. Using [theorem: 2.74] we have that $g\circ f: A\rightarrow g(f(A))$ is a bijection which combined with (1) proves that $g\circ f: \langle A, \oplus_A\rangle \rightarrow \langle C, \oplus_C\rangle$ is a group isomorphism. □

The following theorem show how we can define a group on the product of a family of groups.

Theorem 4.26. *Let $\{\langle A_i, \odot_i\rangle\}_{i\in I}$ be a family of semi-groups then we have*

1. *If $x, y \in \prod_{i\in I} A_i$ then $(x\odot y) \in \prod_{i\in I} A_i$ where $x\odot y$ is defined by $(x\odot y)_i = x_i \odot_i y_i$*
2. *If we define $\odot: (\prod_{i\in I} A_i) \times (\prod_{i\in I} A_i) \rightarrow \prod_{i\in I} A_i$ by $\odot(x, y) = x\odot y$ then*

   $$\left\langle \prod_{i\in I} A_i, \odot \right\rangle$$

   is a semi-group with neutral element e defined by $(e)_i = e_i$ where e_i is the neutral element of $\langle A_i, \odot_i\rangle$
3. *If $\forall i \in I$ we have that $\langle A_i, \odot_i\rangle$ is Abelian then $\langle \prod_{i\in I} A_i, \odot\rangle$ is Abelian.*
4. *If $\forall i \in I$ we have that $\langle A_i, \odot_i\rangle$ is a group then $\langle \prod_{i\in I} A_i, \odot\rangle$ is a group where the inverse x^{-1} for each $x \in \prod_{i\in I} A_i$ is defined by $(x^{-1})_i = (x_i)^{-1}$ [here $(x_i)^{-1}$ is the inverse of x_i in the group $\langle A_i, \odot_i\rangle$*

Proof.

1. If $x,y\in\prod_{i\in I} A_i$ then x is a function $x: I\rightarrow\bigcup_{i\in I} A_i$ such that $\forall i\in I\ x_i=x(i)\in A_i$ and y is a function $y: I\rightarrow\bigcup_{i\in I} A_i$ such that $\forall i\in I\ y_i=y(i)\in A_i$. So if we define $x\odot y$ by $(x\odot y)(i)=(x\odot y)_i=x_i\odot_i y_i=x(i)\odot_i y(i)$ then $x\odot y: I\rightarrow\bigcup_{i\in I} A_i$ is a function and $\forall i\in I$ we have $(x\odot y)(i)=x(i)\odot_i y(i)\in A_i$ [as $\langle A_i,\odot_i\rangle$ is a semi-group]. Hence $x\odot y\in\prod_{i\in I} A_i$

2. We have

   associativity. Let $x,y,z\in\prod_{i\in I} A_i$ then we have for $i\in I$

$$\begin{aligned}(x\odot(y\odot z))(i) &= x(i)\odot_i(y\odot z)(i)\\ &= x(i)\odot_i(y(i)\odot_i z(i))\\ &\underset{\langle A_i,\odot_i\rangle\text{ is a semi group}}{=} (x(i)\odot y(i))\odot z(i)\\ &= (x\odot y)(i)\odot_i z(i)\\ &= ((x\odot y)\odot z)(i)\end{aligned}$$

   so that

$$x\odot(y\odot z)=(x\odot y)\odot z$$

   neutral element. Let $x\in\prod_{i\in I} A_i$ then $\forall i\in I$

$$\begin{aligned}(x\odot e)(i) &= x(i)\odot_i e(i)\\ &= x(i)\odot_i e_i\\ &\underset{\langle A_i,\odot_i\rangle\text{ is a semi group}}{=} x(i)\\ (e\odot x)(i) &= e(i)\odot_i x(i)\\ &= e_i\odot_i x(i)\\ &\underset{\langle A_i,\odot_i\rangle\text{ is a semi group}}{=} x(i)\end{aligned}$$

   so that

$$x\odot e=x=e\odot x$$

3. Let $x,y\in\prod_{i\in I} A_i$ then $\forall i\in I$ we have

$$(x\circ y)(i)=x(i)\odot_i y(i)\underset{\langle A_i,\odot_i\rangle\text{ is Abelian}}{=} y(i)\odot_i x(i)=(y\odot x)(i)$$

   so that $x\odot y=y\odot x$

4. Let $x\in\prod_{i\in I} A_i$ then we have $\forall i\in I$ that

$$\begin{aligned}(x\odot x^{-1})(i) &= x(i)\odot_i(x^{-1})(i)\\ &= x(i)\odot_i(x_i)^{-1}\\ &= x_i\odot_i(x_i)^{-1}\\ &\underset{\langle A_i,\odot_i\rangle\text{ is a group}}{=} e_i\\ &= e(i)\\ (x^{-1}\odot x)(i) &= (x^{-1})(i)\odot_i x(i)\\ &= (x_i)^{-1}\odot_i x(i)\\ &= (x_i)^{-1}\odot_i x_i\\ &= e_i\\ &= e(i)\end{aligned}$$

   So that $x\odot x^{-1}=e=x^{-1}\odot x$. Which as by (2) $\langle\prod_{i\in I} A_i,\odot\rangle$ is a semi group proves that $\langle\prod_{i\in I} A_i,\odot\rangle$ is a group. □

The following five definitions will be later used in Linear Algebra.

Definition 4.27. *Let $\langle G,\odot\rangle$ be a group with neutral element e and let X be a set then we have the following definitions:*

1. *A **left group action** is a function $\triangleright: G\times X\rightarrow X$ where $\triangleright(g,x)\underset{\text{notation}}{=} g\triangleright x$ such that*
   a. *$\forall x\in X$ we have $e\triangleright x=x$*
   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* $\lhd: X \times G \rightarrow X$ *where* $\lhd(x, g) \underset{\text{notation}}{=} x \lhd g$ *such that*
   a. $\forall x \in X$ *we have* $x \lhd e = x$
   b. $\forall g, g' \in G$ *and* $\forall x \in X$ *we have* $x \lhd (g \odot g') = (x \lhd g) \lhd g'$

Definition 4.28. *Let* $\langle G, \odot \rangle$ *be a group, X a set,* $\rhd$ *a left group action and* $g \in G$ *then we define*

$$g_{\rhd}: X \rightarrow X \text{ by } g_{\rhd}(x) = g \rhd x$$

Definition 4.29. *Let* $\langle G, \odot \rangle$ *be a group, X,* $\lhd$ *a right group action and* $g \in G$ *then we define*

$$g_{\lhd}: X \rightarrow X \text{ by } g_{\lhd}(x) = x \lhd g$$

Definition 4.30. *Let* $\langle G, \odot \rangle$ *be a group with neutral element e and let X be a set then we have the following definitions for a left group action* $\rhd$

1. $\rhd$ *is **faithful** if*

$$g_{\rhd} = \mathrm{Id}_X \text{ if and only if } g = e$$

   or equivalently

$$\{g \in G | \forall x \in X \text{ we have } g \rhd x = x\} = \{e\}$$

2. $\rhd$ *is **transitive** iff* $\forall x_1, x_2$ *there exist a* $g \in G$ *such that* $g \rhd x_1 = x_2$
3. $\rhd$ *is **free** iff* $\forall x \in X$ *we have* $\{g \in G | g \rhd x = x\} = \{e\}$

Definition 4.31. *Let* $\langle G, \odot \rangle$ *be a group with neutral element e and let X be a set then we have the following definitions for a right group action* $\lhd$

1. $\rhd$ *is **faithful** if*

$$g_{\lhd} = \mathrm{Id}_X \text{ if and only if } g = e$$

   or equivalently

$$\{g \in G | \forall x \in X \text{ we have } g \lhd x = x\} = \{e\}$$

2. $\rhd$ *is **transitive** iff* $\forall x_1, x_2$ *there exists a* $g \in G$ *such that* $g \lhd x_1 = x_2$
3. $\rhd$ *is **free** iff* $\forall x \in X$ *we have* $\{g \in G | g \lhd x = x\} = \{e\}$

4.2 Rings

Definition 4.32. (Ring) *A triple* $\langle R, \oplus, \odot \rangle$ *is a ring iff*

1. *R is a set*
2. $\langle R, \oplus \rangle$ *is a Abelian group or* $\oplus: R \times R \rightarrow R$ *is a operator such that*

   associativity. $\forall x, y, z \in R$ *we have* $x \oplus (y \oplus z) = (x \oplus y) \oplus z$

   neutral element. $\exists 0 \in R$ *such that* $\forall x \in R$ *we have* $0 \oplus x = x = x \oplus 0$

   inverse element. $\forall x \in R$ *there exist a* $-x$ *such that* $x \oplus (-x) = 0 = (-x) \oplus x$

   commutativity. $\forall x, y \in R$ *we have* $x \oplus y = y \oplus x$

   $\oplus$ *is called the sum operator of the ring.*
3. $\odot: R \times R \rightarrow R$ *is a operator so that*

   distributivity. $\forall x, y, z \in R$ *we have* $x \odot (y \oplus z) = (x \odot y) \oplus (x \odot z)$

   neutral element. $\exists 1 \in R$ *such that* $\forall x \in R$ *we have* $1 \odot x = x = x \odot 1$

   commutativity. $\forall x, y \in R$ *we have* $x \odot y = y \odot x$

   associativity. $\forall x, y, z \in R$ *we have* $x \odot (y \odot z) = (x \odot y) \odot z$

   $\odot$ *is called the multiplication operator of the ring.*

Definition 4.33. *If $\langle R, \oplus, \odot\rangle$ is a ring then a **zero divisor of R** is a $x \in R\backslash\{0\}$ so that $\exists y \in R\backslash\{0\}$ such that $x \odot y = 0$*

Definition 4.34. *A ring $\langle R, \oplus, \odot\rangle$ is a **integral domain** if it does not contains a **zero divisor***

Definition 4.35. (Sub-ring) *If $\langle R, \oplus, \odot\rangle$ is a ring then a subset $S \subseteq R$ is a sub ring iff*

1. *$\forall x, y \in S$ we have $x \oplus y \in S$ and $x \odot y \in S$*
2. *$\forall x \in S$ we have $-x \in S$ [the inverse element for $\oplus$]*
3. *$1 \in S$ [the neutral element for $\odot$]*
4. *$0 \in S$ [the neutral element for $\oplus$]*

Theorem 4.36. *If $\langle R. \oplus, \odot\rangle$ is a ring $F \subseteq R$ a sub ring of $\langle R, \oplus, \odot\rangle$ then*

$$F \text{ is a sub group of } \langle R, \oplus\rangle \text{ and } F \text{ is a sub semi-group of } \langle R, \odot\rangle$$

Proof. This follows directly from [definitions: 4.12, 4.13 and 4.35] □

Theorem 4.37. *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 a ring with the same neutral elements for addition and multiplications and for each $x \in S$ its inverse element is also the inverse element in $\langle S, \oplus_{|S\times S}, \odot_{|S\times S|}\rangle$. For simplicity we note this ring as $\langle S, \oplus, \odot\rangle$*

Proof.

1. S is a set as R is a set by the Axiom of Subsets [axiom: 1.54].
2. $\langle S, \oplus_{|S\times S}\rangle$ is a Abelian group by [theorem: 4.14]
3. $\odot: R \times R \rightarrow R$ is a operator so that

   Distributivity. $\forall x, y, z \in S$ we have

   $$x \odot_{|S\times S} (y \oplus_{|S\times S} z) = x \odot (y \oplus z) = (x \odot y) \oplus (x \odot z) = (x \odot_{|S\times S} y) \oplus_{|S\times S} (x \odot_{|S\times S} z)$$

   neutral element. For $1 \in R$ we have $\forall x \in S$ that

   $$1 \odot_{|S\times S} x = 1 \odot x = x = x \odot 1 = x \odot_{|S\times S} 1$$

   commutativity. $\forall x, y \in S$ we have

   $$x \odot_{|S\times S} y = x \odot y = y \odot x = y \odot_{|S\times S} x$$

   associativity. $\forall x, y, z \in S$ we have

   $$x \odot_{|S\times S} (y \odot_{|S\times S} z) = x \odot (y \odot z) = (x \odot y) \odot z = (x \odot_{|S\times S} y) \odot_{|S\times S} z$$ □

Theorem 4.38. *Let $\langle R, \oplus, \odot\rangle$ be a ring $F \subseteq R$ a sub-ring of $\langle R, \oplus, \odot\rangle$ and $H \subseteq F$ a sub-ring of $\langle R, \oplus_{|F\times F}, \odot_{|F\times F}\rangle$ then H is a sub-ring of $\langle R, \oplus, \odot\rangle$*

Proof.

1. $\forall x, y \in H \subseteq F$ we have $x \oplus_{|F\times F} y \in H$ which as $(x, y) \in F \times F$ proves that

   $$x \oplus y = x \oplus_{|F\times F} y \in H$$

   and $x \odot_{|F\times F} y \in H$ which as $(x, y) \in F \times F$ proves that

   $$x \odot y = x \odot_{|F\times F} y \in H$$

2. If 0 is the additive neutral element of $\langle R, \oplus, \odot\rangle$ then by [theorem: 4.37] 0 is also the additive neutral element of $\langle F, \oplus_{|F\times F}, \odot_{|F\times F}\rangle$hence $0 \in H$.
3. If 1 is the multiplicative neutral element of $\langle R, \oplus, \odot\rangle$ then by [theorem: 4.37] 1 is also the multiplicative neutral element of $\langle F, \oplus_{|F\times F}, \odot_{|F\times F}\rangle$ hence $1 \in H$.
4. If $x \in H$ then by [theorem: 4.37] its inverse element $-x$ is also the inverse element in $\langle F, \oplus_{|F\times F}, \odot_{|F\times F}\rangle$ hence $x^{-1} \in H$. □

The following theorem shows that the neutral element for the sum in a ring is actual a absorbing element.

Theorem 4.39. *Let $\langle X, \oplus, \odot\rangle$ be a ring with 0 the neutral element for $\oplus$ then $\forall x \in R$ we have*

$$x \odot 0 = 0 = 0 \odot x$$

Proof. If $x \in R$ then

$$\begin{aligned} 0 &\underset{\text{inverse element}}{=} (0 \odot x) \oplus -(0 \odot x) \\ &\underset{0 \oplus 0 = 0}{=} ((0 \oplus 0) \odot x) \oplus (-(0 \odot x)) \\ &\underset{\text{distributivity}}{=} [(0 \odot x) \oplus (0 \odot x)] \oplus (-(0 \odot x)) \\ &\underset{\text{associativity}}{=} (0 \odot x) \oplus [(0 \odot x) + (-(0 \odot x))] \\ &\underset{\text{inverse element}}{=} (0 \odot x) \oplus 0 \\ &\underset{\text{inverse element}}{=} 0 \odot x \\ &\underset{\text{commutativity}}{=} x \odot 0 \end{aligned}$$

□

Theorem 4.40. *Let $\langle X, \oplus, \odot\rangle$ be a ring with 0 the neutral element for $\oplus$ and 1 the neutral element for $\odot$ then we have:*

1. $\forall x \in X$ *we have* $-x = (-1) \odot x$
2. $\forall x, y \in X$ *we have* $-(x \odot y) = (-x) \odot y = x \odot (-y)$
3. $\forall x, y \in X$ *we have* $x \odot y = (-x) \odot (-y)$
4. $(-1) \odot (-1) = 1$

Proof.

1. Let $x \in X$ then

$$x + (-1) \odot x \underset{\text{commutative}}{=} (-1) \odot x + x = (-1) \odot x + 1 \odot x = ((-1) + 1) \odot x = 0 \odot x \underset{[\text{theorem: } 4.39]}{=} 0$$

which as the inverse element is unique by [theorem: 4.7] proves that $-x = (-1) \odot x$.

2. If $x, y \in X$ then

$$-(x \odot y) \underset{(1)}{=} (-1) \odot (x \odot y) = ((-1) \odot x) \odot y \underset{(1)}{=} (-x) \odot y$$

and

$$-(x \odot y) = -(y \odot x) \underset{(1)}{=} (-1) \odot (y \odot x) = ((-1) \odot y) \odot x \underset{(1)}{=} (-y) \odot = x \odot (-y)$$

3. If $x, y \in X$ then we have

$$(-x) \cdot (-y) \underset{(2)}{=} -(x \odot (-y)) \underset{(2)}{=} -(-(x \odot y)) \underset{[\text{theorem: } 4.9]}{=} x \odot y$$

4. We have $(-1) \odot (-1) \underset{(2)}{=} 1 \odot 1 = 1$

□

Definition 4.41. *Let $\langle A, \oplus_A, \odot_A\rangle$ and $\langle B, \oplus_B, \odot_B\rangle$ be rings then a function $f: A \rightarrow B$ is is a* ***ring homeomorphism*** *iff*

1. $\forall x, y \in A$ *we have* $f(x \oplus_A y) = f(x) \oplus_B f(y)$
2. $\forall x, y \in A$ *we have* $f(x \odot_A y) = f(x) \odot_B f(y)$
3. $f(1_A) = 1_B$ *where* 1_A *is the multiplicative neutral element in* A *and* 1_B *is the multiplicative neutral element in* B.

***Notation* 4.42.** *As ring homeomorphism between $\langle A, \oplus_A, \odot_A\rangle$ and $\langle B, \oplus_B, \odot_B\rangle$ is noted as*

$$f: \langle A, \oplus_A, \odot_A\rangle \rightarrow \langle B, \oplus_B, \odot_B\rangle \text{ is a ring homeomorphism}$$

Note 4.43. Note that a ring homeomorphism $f: \langle A, \oplus_A, \odot_A\rangle \rightarrow \langle B, \oplus_B, \odot_B\rangle$ is automatically a

$$f: \langle A, \oplus_A\rangle \rightarrow \langle B, \oplus_B\rangle \text{ and } f: \langle A, \odot\rangle \rightarrow \langle B, \odot_B\rangle \text{ group homeomorphism}$$

Theorem 4.44. *If $\langle F, \odot_F, \oplus_F\rangle$ and $\langle G, \oplus_G, \odot_G\rangle$ are rings, $H \subseteq G$ a sub-ring of $\langle G, \oplus\rangle$ and*

$$f: \langle F, \oplus_F, \odot_F\rangle \rightarrow \langle H, (\oplus_G)_{|H \times H}, (\odot_G)_{|H \times H}\rangle \text{ is a ring homeomorphism}$$

then

$$f:\langle F,\oplus_F,\odot_F\rangle\rightarrow\langle G,\oplus_G,\odot_G\rangle \textit{ is a ring homeomorphism}$$

Proof. Let $x,y\in F$ then we have

$$f(x\oplus_F y)=f(x)(\oplus_G)_{|H\times H}f(y)=f(x)\oplus_G f(y)$$

and

$$f(x\odot_F y)=f(x)(\odot_G)_{|H\times H}f(y)=f(x)\odot_G f(y)$$

and finally as the neutral element in $\langle H,(\oplus_G)_{|H\times H},(\odot_G)_{|H\times H}\rangle$ is $1_B\in B$

$$f(1_A)=1_B$$

□

Using 4.22 we have then the following theorem.

Theorem 4.45. *If $\langle A,\oplus_A,\odot_A\rangle$ and $\langle B,\oplus_B,\odot_B\rangle$ are rings with additive units $0_A,0_B$, multiplicative units $1_A,1_B$ and*

$$f:\langle A,\oplus_A,\odot_A\rangle\rightarrow\langle B,\oplus_B,\odot_B\rangle \textit{ a ring homeomorphism}$$

1. $f(0_A)=0_B$
2. $\forall a\in A$ *we have* $f(-a)=-f(a)$
3. $f(A)$ *is a sub-ring of* $\langle B,\oplus_B,\odot_B\rangle$

Proof. Be careful the same symbol will be used in the context of $\langle A,\oplus_A,\odot_A\rangle$ and $\langle B,\oplus_B,\odot_B\rangle$.

1. We have

$$\begin{array}{rcl} 0_B &=& (-f(0_A))\oplus_B f(0_A)\\ &=& (-f(0_A))\oplus_B f(0_A\oplus_A 0_A)\\ &=& (-f(0_A))\oplus_B (f(0_A)\oplus_B f(0_A))\\ &=& ((-f(0_A))\oplus_B f(0_A))\oplus_B f(0_A)\\ &=& 0_B\oplus_B f(0_A)\\ &=& f(0_A)\end{array}$$

2. We have

$$\begin{array}{rcl} f(-x)\oplus_B f(x) &=& f(x\oplus_A(-x))\\ &=& f(0_A)\\ &\underset{(1)}{=}& 0_B\\ f(x)\oplus_B f(-x) &=& f(x\oplus_A(-x))\\ &=& f(0_A)\\ &\underset{(1)}{=}& 0_A\end{array}$$

   so that

$$f(-x)=-f(x)$$

3. Let $x,y\in f(A)$ then $\exists u,v\in A$ such that $x=f(u)$ and $y=f(v)$ then we have

$$x\oplus_B y=f(u)\oplus_B f(v)=f(u\oplus_A v)\in f(A)$$

   and

$$x\odot_B y=f(u)\odot_B f(v)=f(u\odot_A v)\in f(A)$$

   and

$$-x=-f(x)\underset{(2)}{=}f(-x)\in f(A)$$

   and

$$0_B\underset{(1)}{=}f(0_A)\in f(A)$$

and

$$1_B = f(1_A)$$

□

Definition 4.46. *If $\langle A, \oplus_A, \odot_A\rangle$ and $\langle B, \oplus_B, \odot_B\rangle$ are rings then a function $f: A \to B$ is a ring isomorphism if it is a bijection and*

$$f: \langle A, \oplus_A, \odot_A\rangle \to \langle B, \oplus_B, \odot_B\rangle \text{ is a ring homeomorphism}$$

Theorem 4.47. *If $\langle A, \oplus_A, \odot_A\rangle$ and $\langle B, \oplus_B, \odot_B\rangle$ are rings and*

$$f: \langle A, \oplus_A, \odot_A\rangle \to \langle B, \oplus_B, \odot_B\rangle \text{ a ring homeomorphism [or ring isomorphism]}$$

then

$$f: \langle A, \oplus_A\rangle \to \langle B, \oplus_B\rangle \text{ is a group homeomorphism [or group isomorphism]}$$

and

$$f: \langle A, \odot_A\rangle \to \langle B, \odot_B\rangle \text{ is a group homeomorphism [or group isomorphism]}$$

Proof. This follows directly from [definitions: 4.19, 4.23, 4.41 and 4.46]. □

Theorem 4.48. *If $\langle A, \oplus_A, \odot_A\rangle$, $\langle B, \oplus_B, \odot_B\rangle$ and $\langle C, \oplus_C, \odot_C\rangle$ are rings then*

1. *If D is a sub-ring of $\langle B, \oplus_B, \odot_B\rangle$ and*

$$f: \langle A, \oplus_A, \odot_A\rangle \to \langle D, \oplus_B, \odot_B\rangle \text{ and } g: \langle B, \oplus_B, \odot_B\rangle \to \langle C, \oplus_C, \odot_C\rangle \text{ are ring homeomorphisms}$$

   then

$$g \circ f: \langle A, \oplus_A, \odot_A\rangle \to \langle C, \oplus_C, \odot_C\rangle \text{ is a ring homeomorphism}$$

   and

$$g(f(A)) \text{ is a sub-ring of } \langle C, \oplus_C, \odot_C\rangle$$

2. *If D is a sub-ring of $\langle B, \oplus_B, \odot_B\rangle$ and*

$$f: \langle A, \oplus_A, \odot_A\rangle \to \langle D, \oplus_B, \odot_B\rangle \text{ and } g: \langle B, \oplus_B, \odot_B\rangle \to \langle C, \oplus_C, \odot_C\rangle \text{ are ring isomorphisms}$$

   then

$$g \circ f: \langle A, \oplus_A, \odot_A\rangle \to \langle g(f(A)), \oplus_C, \odot_C\rangle \text{ is a ring isomorphism}$$

   or as $g(f(A)) \underset{f: A \to D \text{ is a bijection}}{=} g(D)$

$$g \circ f: \langle A, \oplus_A, \odot_A\rangle \to \langle g(f(A)), \oplus_C, \odot_C\rangle \text{ is a ring isomorphism}$$

Proof.

1. Let $x, y \in A$ then

$$\begin{array}{rcl} (g \circ f)(x \oplus_A y) & = & g(f(x \oplus_A y)) \\ & \underset{f \text{ is a homeomorphism}}{=} & g(f(x) \oplus_B f(y)) \\ & \underset{g \text{ is a homeomorphism}}{=} & g(f(x)) \oplus_C g(f(y)) \\ & = & (g \circ f)(x) \oplus_C (g \circ f)(y) \\ (g \circ f)(x \odot_A y) & = & g(f(x \odot_A y)) \\ & \underset{f \text{ is a homeomorphism}}{=} & g(f(x) \odot_B f(y)) \\ & \underset{g \text{ is a homeomorphism}}{=} & g(f(x)) \odot_C g(f(y)) \\ & = & (g \circ f)(x) \odot_C (g \circ f)(y) \\ (g \circ f)(1_A) & = & g(f(1_A)) \\ & \underset{f \text{ is a homeomorphism}}{=} & g(1_B) \\ & \underset{g \text{ is a homeomorphism}}{=} & 1_C \end{array}$$

   proving that $g \circ f$ is a ring homeomorphism. Finally using [theorem: 4.45] we have then that

$$g(f(A)) \text{ is a sub-ring of } \langle C, \oplus_C, \odot_C\rangle$$

2. Using [theorem: 2.75] we have that $g \circ f: A \rightarrow g(f(A))$ is a bijection which combined with (1) proves that $g \circ f: \langle A, \oplus_A, \odot_A \rangle \rightarrow \langle g(f(A)), \oplus_C, \odot_C \rangle$ is a ring isomorphism. □

Definition 4.49. *$\langle R, \oplus, \odot, \leqslant \rangle$ is a ordered ring if*

1. *$\langle R, \oplus, \odot \rangle$ is a ring*
2. *$\langle R, \leqslant \rangle$ is totally ordered*
3. *$\forall x, y, z \in R$ with $x < y$ we have $x \oplus z < y \oplus z$*
4. *$\forall x, y \in R$ with $0 < x$ and $0 < y$ we have $0 < x \odot y$*

If in addition $\langle R, \oplus, \odot \rangle$ is a integral domain then we say that $\langle R, \oplus, \odot, \leqslant \rangle$ is a ordered integral domain.

Theorem 4.50. *If $\langle R, \oplus, \odot, \leqslant \rangle$ is a ordered ring then we have :*

1. *$\forall x, y, z \in R$ we have $x < y \Leftrightarrow x \oplus z < y \oplus z$*
2. *$\forall x, y, z \in R$ we have $x \leqslant y \Leftrightarrow x \oplus z \leqslant y \oplus z$*
3. *$\forall x, y \in R$ we have $x < y \Leftrightarrow 0 < y \oplus (-x)$*
4. *$\forall x, y \in R$ we have $x \leqslant y \Leftrightarrow 0 \leqslant y \oplus (-x)$*
5. *$\forall x, y \in R$ we have $x < y \Leftrightarrow -y < -x$*
6. *$\forall x, y \in R$ we have $x \leqslant y \Leftrightarrow -y \leqslant -x$*
7. *$\forall x, y, z \in R$ with $0 < z$ we have $x < y \Leftrightarrow x \odot z < y \odot z$*
8. *$\forall x, y, z \in R$ with $0 < z$ we have $x \leqslant y \Leftrightarrow x \odot z \leqslant y \odot z$*
9. *$\forall x, y, z \in \mathbb{R}$ with $0 \leqslant z$ and $x \leqslant y$ we have $x \odot \leqslant y \odot z$*
10. *$\forall x, y, z \in R$ with $z < 0$ we have $x < y \Leftrightarrow y \odot z < x \odot z$*
11. *$\forall x, y, z \in R$ with $z < 0$ we have $x \leqslant y \Leftrightarrow y \odot z \leqslant x \odot z$*
12. *$\forall x, y, z \in R$ with $z \leqslant 0$ and $x \leqslant y$ we have $y \odot z \leqslant x \odot z$*
13. *$\forall x \in R$ we have $0 \leqslant x \cdot x \underset{\text{def}}{=} x^2$, further if $x \neq 0$ then $0 < x \odot x \underset{\text{def}}{=} x^2$*
14. *$0 \leqslant 1$*
15. *$\forall x, y \in R$ with $0 < x < y$ we have that $x^2 < y^2$ where $x^2 = x \odot x$ and $y^2 = y \odot y$*
16. *$\forall x, y \in R$ with $0 \leqslant x \leqslant y$ we have that $x^2 \leqslant y^2$ where $x^2 = x \odot x$ and $y^2 = y \odot y$*

Proof.

1.

 ⇒. This follows directly from the definition of a ordered ring.

 ⇐. If $x \oplus z \leqslant y \oplus z$ then from the definition of a ordered ring we have

$$x = x + (z \oplus (-z)) = (x \oplus z) \oplus (-z) < (y \oplus z) + (-z) = y \oplus (z \oplus (-z)) = y$$

 so that $x < y$

2.

$$\begin{array}{rcl} x \leqslant y & \Leftrightarrow & x = y \vee x < y \\ & \underset{(1)}{\Leftrightarrow} & x = y \vee x \oplus z < y \oplus z \\ & \underset{x = y \Leftrightarrow x \oplus z = y \oplus z}{\Leftrightarrow} & x \oplus z = y \oplus z \vee x \oplus z < y \oplus z \\ & \Leftrightarrow & x \oplus z \leqslant y \oplus z \end{array}$$

3.

$$\begin{array}{rcl} x < y & \underset{(1)}{\Leftrightarrow} & x \oplus (-x) < y \oplus (-x) \\ & \Leftrightarrow & 0 < y \oplus (-x) \end{array}$$

4.

$$\begin{aligned} x \leqslant y & \underset{(2)}{\Leftrightarrow} x \oplus (-x) \leqslant y \oplus (-x) \\ & \Leftrightarrow 0 \leqslant y \oplus (-x) \end{aligned}$$

5.

$$\begin{aligned} x < y & \underset{(3)}{\Leftrightarrow} 0 < y \oplus (-x) \\ & \underset{\text{commutativity}}{\Leftrightarrow} 0 < (-x) \oplus y \\ & \underset{(1)}{\Leftrightarrow} 0 + (-y) < ((-x) \oplus y) \oplus (-y) \\ & \Leftrightarrow -y < -x \end{aligned}$$

6.

$$\begin{aligned} x \leqslant y & \underset{(4)}{\Leftrightarrow} 0 < y \oplus (-x) \\ & \underset{\text{commutativity}}{\Leftrightarrow} 0 < (-x) \oplus y \\ & \underset{(1)}{\Leftrightarrow} 0 + (-y) < ((-x) \oplus y) \oplus (-y) \\ & \Leftrightarrow -y < -x \end{aligned}$$

7. Let $0 < z$ then we have

⇒. As $x < y$ we have by (3) that $0 < y \oplus (-x)$ so by the definition of a ordered ring that

$$\begin{aligned} 0 & < (y \oplus (-x)) \odot z \\ & \underset{\text{distributivity}}{=} y \odot z \oplus (-x) \odot z \\ & \underset{[\text{theorem: } 4.40]}{=} y \odot z \oplus (-(x \odot z)) \end{aligned}$$

which using (3) prove that $x \odot z < y \odot z$.

⇐. Then $x \odot z < y \odot z$ and using the totally ordering we have for x, y either:

$\boldsymbol{x = y}$**.** Then $x \odot z = y \odot z$ contradicting $x \odot z < y \odot z$ so this case never occurs.

$\boldsymbol{y < x}$**.** Then by (3) we have $0 < x \oplus (-y)$ so that by the definition of a ordered field we have

$$\begin{aligned} 0 & < (x \oplus (-y)) \odot z \\ & \underset{\text{distributivity}}{=} x \odot z \oplus (-y) \odot z \\ & \underset{[\text{theorem: } 4.40]}{=} x \odot z \oplus (-(y \odot z)) \end{aligned}$$

which using (3) proves that $y \odot z < x \odot z$ contradicting $x \odot z < y \odot z$. So this case does not occur.

$\boldsymbol{x < y}$**.** the remaining case.

So the only remaining case is $x < y$

8.

⇒. As $0 < z$ and $x \leqslant y$ we have for x, y either:

$\boldsymbol{x = y}$**.** Then $x \odot z = y \odot z$ hence $x \odot z \leqslant y \odot z$

$\boldsymbol{x < y}$**.** Then by (7) we have $x \odot z < y \odot z$ hence $x \odot z \leqslant y \odot z$

So in all cases we have $x \odot z \leqslant y \odot z$.

⇐. Assume that $x \odot z \leqslant y \odot z$. By totally ordering we have for x, y either $x \leqslant y$ or $y < x$. If $y < x$ then by (7) we have $y \odot z < x \odot z$ contradicting $x \odot z \leqslant y \odot z$ so we must have that $x \leqslant y/$

9. Let $0 \leqslant z$ and $x \leqslant y$ then we have for z either:

$\boldsymbol{z = 0}$**.** Then $x \odot z = x \odot 0 \underset{[\text{theorem: } 4.39]}{=} 0 \underset{[\text{theorem: } 4.39]}{=} y \odot 0 = y \odot z$ so that $x \odot z \leqslant y \odot z$

$\boldsymbol{0 < z}$**.** Then by (8) we have $x \odot z < y \odot z$

So in all cases we have $x \odot z \leqslant y \odot z$

10. As $z < 0$ we have by (5) that $0 < (-z)$ so that

$$\begin{aligned} x < y & \underset{(7)}{\Leftrightarrow} x \odot (-z) < y \odot (-z) \\ & \underset{[\text{theorem: } 4.40]}{\Leftrightarrow} -(x \odot z) < -(y \odot z) \\ & \underset{(5)}{\Leftrightarrow} y \odot z < z \odot z \end{aligned}$$

11. As $z < 0$ we have by (5) that $0 < (-z)$ so that

$$\begin{aligned} x \leqslant y & \underset{(8)}{\Leftrightarrow} x \odot (-z) \leqslant y \odot (-z) \\ & \underset{[\text{theorem: } 4.40]}{\Leftrightarrow} -(x \odot z) \leqslant -(y \odot z) \\ & \underset{(6)}{\Leftrightarrow} y \odot z \leqslant z \odot z \end{aligned}$$

12. As $z \leqslant 0$ we have by (5) that $0 \leqslant -z$ so that by (9) we have

$$\begin{aligned} x \leqslant y & \underset{(9)}{\Rightarrow} x \odot (-z) \leqslant y \odot (-z) \\ & \underset{[\text{theorem: } 4.40]}{\Leftrightarrow} -(x \odot z) \leqslant -(y \odot z) \\ & \underset{(6)}{\Leftrightarrow} y \odot z \leqslant z \odot z \end{aligned}$$

13. If $x \in \mathbb{R}$ then we have either:

    $0 < x$. Then we have by (7) that $0 \underset{[\text{theorem: } 4.39]}{=} 0 \odot x < x \odot x$ hence $0 < x \odot x$

    $0 = x$. Then we have $0 = 0 \odot 0 = x \odot x$ so that $0 \leqslant x \odot x$

    $x < 0$. Then we have by (7) $0 \underset{[\text{theorem: } 4.39]}{=} 0 \odot (-x) \leqslant (-x) \odot (-x) \underset{[\text{theorem: } 4.40]}{=} x \odot x$ hence $0 < x \odot x$

14. Using (13) we have $0 \leqslant 1 \odot 1 = 1$

15. Let $0 < x < y$ then by (7) we have $x \odot x < y \odot x$ and $x \odot y < y \odot y$ so that $x \odot x < y \odot y$.

16. Let $0 \leqslant x < \leqslant y$ then by (8) we have $x \odot x \leqslant y \odot x$ and $x \odot y \leqslant y \odot y$ so that $x \odot x < \leqslant y \odot y$.

□

4.3 Fields

A ring has no inverse for a multiplicative element, one of the reasons for this is that is is difficult to say what the inverse of 0 is, as expressed in the following computation

$$1 = 0 \odot 0^{-1} \underset{[\text{theorem: } 4.39]}{=} 0$$

so that we have

$$\forall x \in R \text{ that } x = 1 \odot x = 0 \odot x \underset{[\text{theorem: } 4.39]}{=} 0$$

and we end up with $R = \{0\}$, which is not a useful ring. However we can avoid this problem if we exclude the 0 of the list of elements that has a inverse element. This is the idea behind a field.

Definition 4.51. *A triple $\langle F. \oplus, \odot \rangle$ is a field if $\langle F, \oplus, \odot \rangle$ is a ring and additional*

$$\forall x \in F \setminus \{0\} \; \exists b \in F \text{ such that } x \odot b = 1 = b \odot x$$

$$0 \neq 1$$

where 1 is the neutral element for $\odot$. In other words $\langle F, \oplus, \odot \rangle$ is a field iff

1. *F is a set*
2. *$\langle F, \oplus \rangle$ is a Abelian group or $\oplus: F \times F \rightarrow F$ is a operator such that*

    associativity. *$\forall x, y, z \in F$ we have $x \oplus (y \oplus z) = (x \oplus y) \oplus z$*

neutral element. *$\exists 0 \in F$ such that $\forall x \in F$ we have $0 \oplus x = x = x \oplus 0$*

inverse element. *$\forall x \in F$ there exist a $-x$ such that $x \oplus (-x) = 0 = (-x) \oplus x$*

commutativity. *$\forall x, y \in F$ we have $x \oplus y = y \oplus x$*

$\oplus$ is called the sum operator of the field.

3. *$\odot: F \times F \to F$ is a operator so that*

   Distributivity. *$\forall x, y, z \in F$ we have $x \odot (y \oplus z) = (x \odot y) \oplus (x \odot z)$*

   neutral element. *$\exists 1 \in F$ such that $\forall x \in F$ we have $1 \odot x = x = x \odot 1$*

   commutativity. *$\forall x, y \in F$ we have $x \odot y = y \odot x$*

   associativity. *$\forall x, y, z \in F$ we have $x \odot (y \odot z) = (x \odot y) \odot z$*

   inverse element. *$\forall x \in F \setminus \{0\}$ $\exists b \in F$ such that $x \odot b = 1 = b \odot x$*

4. $1 \neq 0$
   $\odot$ is called the multiplication operator of the field.

The inverse if it exist is unique

Theorem 4.52. *If $\langle F, \oplus, \odot \rangle$ is field then $\forall x \in F \setminus \{0\}$ there exist a **unique** inverse element for $\odot$. We note this element as x^{-1}.*

Proof. Let $x \in F \setminus \{0\}$ and assume that $y, y' \in F$ such that $y \odot x = 1 = x \odot y$ and $y' \odot x = 1 = x \odot y'$ then we have

$$y = y \odot 1 = y \circ (x \odot y') = (y \odot x) \odot y' = 1 \odot y' = y'$$

□

Theorem 4.53. *If $\langle F, \oplus, \odot \rangle$ is field then $\forall x \in F \setminus \{0\}$ we have $(x^{-1})^{-1} = x$*

Proof. First if $x^{-1} = 0$ then $x = x \odot 1 = x \odot (x \odot x^{-1}) = (x \odot x) \odot x^{-1} = (x \odot x) \odot 0 \underset{[\text{theorem: } 4.39]}{=} 0$ contradicting $x \in F \setminus \{0\}$. So we must have that $x^{-1} \neq 0$ hence $(x^{-1})^{-1}$ is defined. Further

$$\begin{array}{rcl} 1 = x^{-1} \odot (x^{-1})^{-1} & \Rightarrow & x \odot 1 = x \odot (x^{-1} \odot (x^{-1})^{-1}) \\ & \underset{\text{neutral element}}{\Rightarrow} & x = x \odot (x^{-1} \odot (x^{-1})^{-1}) \\ & \underset{\text{associativity}}{\Rightarrow} & x = (x \odot x^{-1}) \odot (x^{-1})^{-1} \\ & \underset{\text{inverse element}}{\Rightarrow} & x = 1 \odot (x^{-1})^{-1} \\ & \underset{\text{neutral element}}{\Rightarrow} & x = (x^{-1})^{-1} \end{array}$$

□

Theorem 4.54. *If $\langle F, \oplus, \odot \rangle$ is a field then $\forall x, y \in F \setminus \{0\}$ we have $x^{-1} = y^{-1} \Leftrightarrow x = y$*

Proof. We have

$\Rightarrow$. If $x = y$ then trivially $x^{-1} = y^{-1}$

$\Leftarrow$. Then we have

$$\begin{array}{rcl} x^{-1} = y^{-1} & \Rightarrow & x^{-1} \odot y = y^{-1} \odot y \\ & \underset{\text{inverse element}}{\Rightarrow} & x^{-1} \odot y = 1 \\ & \Rightarrow & x \odot (x^{-1} \odot y) = x \odot 1 \\ & \underset{\text{neutral element}}{\Rightarrow} & x \odot (x^{-1} \odot y) = x \\ & \underset{\text{associativity}}{\Rightarrow} & (x \odot x^{-1}) \odot y = x \\ & \underset{\text{inverse element}}{\Rightarrow} & 1 \odot y = x \\ & \underset{\text{neutral element}}{\Rightarrow} & y = x \end{array}$$

□

Theorem 4.55. *If $\langle F, \oplus, \odot \rangle$ is a field then $\forall x, y \in F$ we have $(x \odot y)^{-1} = x^{-1} \odot y^{-1}$*

Proof.

$$\begin{aligned}(x^{-1}\circ y^{-1})\odot(x\odot y) &\underset{\text{commutativity}}{=} (x\odot y)\odot(x^{-1}\circ y^{-1})\\ &\underset{\text{commutativity}}{=} (x\odot y)\odot(y^{-}\odot x^{-1})\\ &\underset{\text{associativity}}{=} ((x\odot y)\odot y^{-1})\odot x^{-1}\\ &\underset{\text{associativity}}{=} (x\odot(y\odot y^{-1}))\odot x^{-1}\\ &\underset{\text{inverse element}}{=} (x\odot 1)\odot x^{-1}\\ &\underset{\text{neutral element}}{=} x\odot x^{-1}\\ &\underset{\text{inverse element}}{=} 1\end{aligned}$$

proving by the uniqueness of the inverse [theorem: 4.52] that

$$(x\odot y)^{-1}=x^{-1}\odot y^{-1}$$

□

Theorem 4.56. *Let $\langle F,\oplus,\odot\rangle$ be a field and $x,y\in F$ and $z\in F\setminus\{0\}$ then $x=y\Leftrightarrow x\odot z=y\odot z$*

Proof. As $z\neq 0$ we have that z^{-1} exist.

⇒. If $x=y$ then clearly $x\odot z=y\odot z$

⇐. If $x\odot z\oplus y\odot z$ then

$$x=x\odot 1=x\odot(z\odot z^{-1})=(x\odot z)\odot z^{-1}\underset{x\odot z=y\odot z}{=}(y\odot z)\odot z^{-1}=y\odot(z\odot z^{-1})=y\odot 1=y$$ □

Theorem 4.57. *If $\langle F,\oplus,\odot\rangle$ is a field and $x,y\in F$ then $x\odot y=0\Leftrightarrow x=0\vee y=0$*

Proof.

⇒. If $x\odot y=0$ then for x we have either

$x=0$. Then clearly $0\vee y=0$ is true.

$x\neq 0$. Then $0=x^{-1}\odot 0=x^{-1}\odot(x\odot y)=(x^{-1}\odot x)\odot y=1\odot y=y$ so that $y=0$ or $x=0\vee y=0$ is true.

⇐. If $x=0$ or $y=0$ we have by [theorem: 4.39] and the fact that a field is a ring that $x\odot y=0$ □

Corollary 4.58. *If $\langle F,\oplus,\odot\rangle$ is a field then $\langle F,\oplus,\odot\rangle$ is a integral domain*

Proof. From the definition of a field [definition: 4.51] it follows that $\langle F,\oplus,\odot\rangle$ is a ring. Assume that $x\in F$ is a zero divisor [see 4.33 then $x\neq 0$ and $\exists y\in F\setminus\{0\}$ such that $x\cdot y=0$. However by [theorem: 4.57] we have $x=0\vee y=0$ contradicting $x\neq 0\neq y$. So F does not have zero divisors proving that $\langle F,\oplus,\odot\rangle$ is a integral domain. □

Definition 4.59. *If $\langle F,\oplus,\odot\rangle$ is a field then a subset $S\subseteq F$ is a sub-field iff the following is satisfied*

1. *$\forall x,y\in F$ we have $x\oplus y\in F$ and $x\odot y\in F$*
2. *$\forall x\in F$ we have $-x\in F$ [the inverse element for $\oplus$]*
3. *$1\in F$ [the neutral element for $\odot$]*
4. *$0\in F$ [the neutral element for $\oplus$]*
5. *$\forall x\in F\setminus\{0\}$ we have $x^{-1}\in F$*

Theorem 4.60. *If $\langle F.\oplus,\odot\rangle$ is a field $G\subseteq G$ a sub ring of $\langle F,\oplus,\odot\rangle$ then*

$$G\,\text{is a subring of}\,\langle F,\oplus,\odot\rangle,\ G\textit{ is a sub group of }\langle F,\oplus\rangle\textit{ and }G\textit{ is a sub semi-group of }\langle F,\odot\rangle$$

Proof. This follows directly from [definitions: 4.35, 4.59] and [theorem: 4.36]. □

Theorem 4.61. *If $\langle F,\oplus,\odot\rangle$ is a field and $S\subseteq F$ is a sub-field then $\langle S,\oplus_{|S\times S},\odot_{|S\times S}\rangle$ is a field with the same additive and multiplicative neutral element as $\langle F,\oplus,\odot\rangle$. Further if $x\in S$ then the additive inverse element in $\langle F,\oplus,\odot\rangle$ is also the inverse in $\langle S,\oplus_{/S\times S},\odot_{/S\times S}\rangle$ and if $x\in S\setminus\{0\}$ then the multiplicative inverse element in $\langle F,\oplus,\odot\rangle$ is also the multiplicative element in $\langle S,\oplus_{/S\times S},\odot_{/S\times S}\rangle$.*

Proof. Using [theorem: 4.37] it follows that $\langle S,\oplus_{|S\times S},\odot_{|S\times S}\rangle$ is a ring. Further if $x\in F\setminus\{0\}$ then $1\in S$ and $x^{-1}\in S$, further $x\odot_{|S}x^{-1}=x\odot x^{-1}=1=x^{-1}\odot x=x^{-1}\odot_{|S}x$ proving that $\langle S,\oplus_{|S\times S},\odot_{|S\times S}\rangle$ is a field. □

Theorem 4.62. *Let $\langle F,\oplus,\odot\rangle$ be a field $G\subseteq F$ a sub-field of $\langle F,\oplus,\odot\rangle$ and $H\subseteq G$ a sub-field of $\langle G,\oplus_{|G\times G},\odot_{|G\times G}\rangle$ then H is a sub-field of $\langle G,\odot\rangle$*

Proof.

1. $\forall x,y\in H\subseteq G$ we have $x\oplus_{|G\times G}y\in H$ which as $(x,y)\in G\times G$ proves that
$$x\oplus y=x\oplus_{|G\times G}G\in H$$
and $x\odot_{|G\times G}y\in H$ which as $(x,y)\in G\times G$ proves that
$$x\odot y=x\odot_{|G\times G}y\in H$$
2. If 0 is the additive neutral element of $\langle F,\oplus,\odot\rangle$ then by [theorem: 4.61] 0 is also the additive neutral element of $\langle G,\oplus_{|G\times G},\odot_{|G\times G}\rangle$hence $0\in H$.
3. If 1 is the multiplicative neutral element of $\langle F,\oplus,\odot\rangle$ then by [theorem: 4.61] 1 is also the multiplicative neutral element of $\langle G,\oplus_{|G\times G},\odot_{|G\times G}\rangle$ hence $1\in H$.
4. If $x\in H$ then by [theorem: 4.61] its additive inverse element $-x$ is also the additive inverse element in $\langle G,\oplus_{|G\times G},\odot_{|G\times G}\rangle$ hence $x^{-1}\in H$.
5. If $x\in H\setminus\{0\}$ then by [theorem: 4.61] its multiplicative inverse element x^{-1} is also the multiplicative inverse element in $\langle G,\oplus_{|G\times G},\odot_{|G\times G}\rangle$ hence $x^{-1}\in H$. □

Definition 4.63. *If $\langle A,\odot_A,\oplus_A\rangle$ and $\langle B,\odot_B,\oplus_B\rangle$ are fields with multiplicative units $1_A,1_B$ then a function $f\colon A\rightarrow B$ is a field homeomorphism iff*

1. *$\forall x,y\in A$ we have $f(x\odot_A y)=f(x)\odot_B f(y)$*
2. *$\forall x,y\in A$ we have $f(x\oplus_A y)=f(x)\oplus_B f(y)$*
3. *$f(1_A)=1_B$*

***Notation* 4.64.** *A field homeomorphism between $\langle A,\odot_A,\oplus_A\rangle$ and $\langle B,\odot_B,\oplus_B\rangle$ is noted as*
$$f\colon\langle A,\odot_A,\oplus_A\rangle\rightarrow\langle B,\odot_B,\oplus_B\rangle \textit{ is a field homeomorphism}$$

Note 4.65. If $\langle A,\odot_A,\oplus_A\rangle$ and $\langle B,\odot_B,\oplus_B\rangle$ are fields and
$$f\colon\langle A,\odot_A,\oplus_A\rangle\rightarrow\langle B,\odot_B,\oplus_B\rangle \text{ a field homeomorphism}$$
then
$$f\colon\langle A,\odot_A,\oplus_A\rangle\rightarrow\langle B,\odot_B,\oplus_B\rangle \text{ is a ring homeomorphism}$$
and
$$f\colon\langle A,\oplus_A\rangle\rightarrow\langle B,\oplus_B\rangle \text{ and } f\colon\langle A,\odot_A\rangle\rightarrow\langle B,\odot_B\rangle \text{ are group homeomorphisms}$$

Proof. As a field is also a ring [see definitions: 4.32 and 4.51] we have by [definitions: 4.41 and 4.63] that
$$f\colon\langle A,\odot_A,\oplus_A\rangle\rightarrow\langle B,\odot_B,\oplus_B\rangle \text{ is a ring homeomorphism}$$
Finally using [note: 4.43] we have that
$$f\colon\langle A,\oplus_A\rangle\rightarrow\langle B,\oplus_B\rangle \text{ and } f\colon\langle A,\odot_A\rangle\rightarrow\langle B,\odot_B\rangle \text{ are group homeomorphisms}$$ □

Definition 4.66. *Let $\langle A,\odot_A,\oplus_A\rangle$ and $\langle B,\odot_B,\oplus_B\rangle$ be fields then a field homeomorphism*
$$f\colon\langle A,\oplus_A\rangle\rightarrow\langle B,\oplus_B\rangle \textit{ and } f\colon\langle A,\odot_A\rangle\rightarrow\langle B,\odot_B\rangle$$
is a field isomorphism if it is also a bijection.

Theorem 4.67. *If $\langle A,\oplus_A,\odot_A\rangle$ and $\langle B,\oplus_B,\odot_B\rangle$ are rings and*
$$f\colon\langle A,\oplus_A,\odot_A\rangle\rightarrow\langle B,\oplus_B,\odot_B\rangle \textit{ a field homeomorphism [or field isomorphism]}$$

then

$$f: \langle A, \oplus_A, \odot_A \rangle \rightarrow \langle B, \oplus_B, \odot_B \rangle \text{ a ring homeomorphism [or ring isomorphism]}$$

and

$$f: \langle A, \oplus_A \rangle \rightarrow \langle B, \oplus_B \rangle \text{ is a group homeomorphism [or group isomorphism]}$$

and

$$f: \langle A, \odot_A \rangle \rightarrow \langle B, \odot_B \rangle \text{ is a group homeomorphism [or group isomorphism]}$$

Proof. This follows directly from [definitions: 4.41, 4.46, 4.63 and 4.66] and [theorem: 4.47] □

Theorem 4.68. *If* $\langle F, \odot_F, \oplus_F \rangle$ *and* $\langle G, \oplus_G, \odot_G \rangle$ *are fields,* $H \subseteq G$ *a sub-field of* $\langle G, \oplus \rangle$ *and*

$$f: \langle F, \oplus_F, \odot_F \rangle \rightarrow \langle H, (\oplus_G)_{|H \times H}, (\odot_G)_{|H \times H} \rangle \text{ is a field homeomorphism}$$

then

$$f: \langle F, \oplus_F, \odot_F \rangle \rightarrow \langle G, \oplus_G, \odot_G \rangle \text{ is a field homeomorphism}$$

Proof. Let $x, y \in F$ then we have

$$f(x \oplus_F y) = f(x)(\oplus_G)_{|H \times H} f(y) = f(x) \oplus_G f(y)$$

and

$$f(x \odot_F y) = f(x)(\odot_G)_{|H \times H} f(y) = f(x) \odot_G f(y)$$

and finally as the neutral element in $\langle H, (\oplus_G)_{|H \times H}, (\odot_G)_{|H \times H} \rangle$ is $1_B \in B$

$$f(1_A) = 1_B$$

□

Theorem 4.69. *If* $\langle A, \odot_A, \oplus_A \rangle$ *and* $\langle B, \odot_B, \oplus_B \rangle$ *are fields with multiplicative units* $1_A, 1_B$ *and*

$$f: \langle A, \oplus_A \rangle \rightarrow \langle B, \oplus_B \rangle \text{ and } f: \langle A, \odot_A \rangle \rightarrow \langle B, \odot_B \rangle \text{ is a field isomorphism}$$

then

$$f^{-1}: \langle B, \odot_B \rangle \rightarrow \langle A, \odot_A \rangle \text{ is a field isomorphism}$$

Proof. First using [theorem: 2.72] we have that $f^{-1}: B \rightarrow A$ is a bijection. Further we have:

1. Take $x, y \in B$ then we have

$$\begin{array}{rcl} f^{-1}(x \oplus_B y) & = & f^{-1}(\mathrm{Id}_B(x) \oplus_B \mathrm{Id}_B(y)) \\ & \underset{\text{[theorem: 2.69]}}{=} & f^{-1}((f \circ f^{-1})(x) \oplus_B (f \circ f^{-1})(y)) \\ & \underset{\text{[theorem: 2.42]}}{=} & f^{-1}(f(f^{-1}(x)) \oplus_B f(f^{-1}(y))) \\ & \underset{f \text{ is homeomorphism}}{=} & f^{-1}(f(f^{-1}(x) \oplus_A f^{-1}(y))) \\ & \underset{\text{[theorem: 2.42]}}{=} & (f^{-1} \circ f)(f^{-1}(x) \oplus_A f^{-1}(y)) \\ & \underset{\text{[theorem: 2.69]}}{=} & \mathrm{Id}_A(f^{-1}(x) \oplus_A f^{-1}(y)) \\ & = & f^{-1}(x) \oplus_A f^{-1}(y) \end{array}$$

2. Take $x, y \in B$ then we have

$$\begin{array}{rcl} f^{-1}(x \odot_B y) & = & f^{-1}(\mathrm{Id}_B(x) \odot_B \mathrm{Id}_B(y)) \\ & \underset{\text{[theorem: 2.69]}}{=} & f^{-1}((f \circ f^{-1})(x) \odot_B (f \circ f^{-1})(y)) \\ & \underset{\text{[theorem: 2.42]}}{=} & f^{-1}(f(f^{-1}(x)) \odot_B f(f^{-1}(y))) \\ & \underset{f \text{ is homeomorphism}}{=} & f^{-1}(f(f^{-1}(x) \odot_A f^{-1}(y))) \\ & \underset{\text{[theorem: 2.42]}}{=} & (f^{-1} \circ f)(f^{-1}(x) \odot_A f^{-1}(y)) \\ & \underset{\text{[theorem: 2.69]}}{=} & \mathrm{Id}_A(f^{-1}(x) \odot_A f^{-1}(y)) \\ & = & f^{-1}(x) \odot f^{-1}(y) \end{array}$$

3. From $f(1_A) = 1_B$ it follows that

$$f^{-1}(1_B) = f^{-1}(f(1_B)) = (f^{-1} \circ f)(1_B) \underset{[\text{theorem: } 2.69}{=} \mathrm{Id}_A(1_B) = 1_B$$ □

Theorem 4.70. *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 $f: \langle A, \oplus_A, \odot_A \rangle \rightarrow \langle B, \oplus_B, \odot_B \rangle$ 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* $f(a^{-1}) = (f(a))^{-1}$
4. $f(A)$ *is a sub-field of* $\langle B, \oplus_B, \odot_B \rangle$

Proof.

1. We have

$$\begin{aligned} 0_B &= (-f(0_A)) \oplus_B f(0_A) \\ &= (-f(0_A)) \oplus_B f(0_A \oplus_A 0_A) \\ &= (-f(0_A)) \oplus_B (f(0_A) \oplus_B f(0_A)) \\ &= ((-f(0_A)) \oplus_B f(0_A)) \oplus_B f(0_A) \\ &= 0_B \oplus f(0_A) \\ &= f(0_A) \end{aligned}$$

2. We have

$$\begin{aligned} f(-x) \oplus_B f(x) &= f((-x) \oplus_A x) \\ &= f(0_A) \\ &\underset{(1)}{=} 0_B \\ f(x) \oplus_B f(-x) &= f(x \oplus_A (-x)) \\ &= f(0_A) \\ &\underset{(1)}{=} 0_A \end{aligned}$$

so that

$$f(-x) = -f(x)$$

3. If $a \in A$ with $a \neq 0_A$ then

$$\begin{aligned} f(a^{-1}) \odot_B f(a) &= f(a^{-1} \odot_A a) \\ &= f(1_A) \\ &= 1_B \\ f(a) \odot_B f(a^{-1}) &= f(a \odot_A a^{-1}) \\ &= f(1_A) \\ &= 1_B \end{aligned}$$

so that

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

4. Let $x, y \in f(A)$ then $\exists u, v \in A$ such that $x = f(u)$ and $y = f(v)$ so we have

$$x \oplus_B y = f(u) \oplus_B f(v) = f(u \oplus_A v) \in f(A)$$

and

$$x \odot_B y = f(u) \odot_B f(v) = f(u \odot_A v) \in f(A)$$

and

$$-x = -f(x) \underset{(2)}{=} f(-x) \in f(A)$$

and if $x \neq 0_B$ then

$$x^{-1} = f(u)^{-1} \underset{(3)}{=} f(u^{-1}) \in f(A)$$

and

$$0_B \underset{(1)}{=} f(0_A) \in f(A)$$

and by definition of a field homeomorphism

$$1_B = f(1_A)$$

□

Theorem 4.71. *If $\langle A, \oplus_A, \odot_A \rangle$, $\langle B, \oplus_B, \odot_B \rangle$ and $\langle C, \oplus_C, \odot_C \rangle$ are fields then*

1. *If $D \subseteq B$ is a sub-field of $\langle B, \oplus_B, \odot_B \rangle$ and*

$$f: \langle A, \oplus_A, \odot_A \rangle \to \langle D, \oplus_B, \odot_B \rangle \text{ and } g: \langle B, \oplus_B, \odot_B \rangle \to \langle C, \oplus_C, \odot_C \rangle \text{ are field homeomorphism}$$

then

$$g \circ f: \langle A, \oplus_A, \odot_A \rangle \to \langle C, \oplus_C, \odot_C \rangle \text{ is a field homeomorphism}$$

and

$$g(f(A)) \text{ is a sub-field of } \langle C, \oplus_C, \odot_C \rangle$$

2. *If $D \subseteq B$ is a sub-field of $\langle B, \oplus_B, \odot_B \rangle$ and*

$$f: \langle A, \oplus_A, \odot_A \rangle \to \langle B, \oplus_B, \odot_B \rangle \text{ and } g: \langle B, \oplus_B, \odot_B \rangle \to \langle C, \oplus_C, \odot_C \rangle \text{ are field isomorphisms}$$

then

$$g \circ f: \langle A, \oplus_A, \odot_A \rangle \to \langle g(f(A)), \oplus_C, \odot_C \rangle \text{ is a field isomorphism}$$

or as $g(f(A)) \underset{f: A \to D \text{ is a bijection}}{=} g(D)$

$$g \circ f: \langle A, \oplus_A, \odot_A \rangle \to \langle g(D), \oplus_C, \odot_C \rangle \text{ is a field isomorphism}$$

Proof.

1. Let $x, y \in A$ then

$$\begin{array}{rcl}
(g \circ f)(x \oplus_A y) & = & g(f(x \oplus_A y)) \\
& \underset{f \text{ is a homeomorphism}}{=} & g(f(x) \oplus_B f(y)) \\
& \underset{g \text{ is a homeomorphism}}{=} & g(f(x)) \oplus_C g(f(y)) \\
& = & (g \circ f)(x) \oplus_C (g \circ f)(y) \\
(g \circ f)(x \odot_A y) & = & g(f(x \odot_A y)) \\
& \underset{f \text{ is a homeomorphism}}{=} & g(f(x) \odot_B f(y)) \\
& \underset{g \text{ is a homeomorphism}}{=} & g(f(x)) \odot_C g(f(y)) \\
& = & (g \circ f)(x) \odot_C (g \circ f)(y) \\
(g \circ f)(1_A) & = & g(f(1_A)) \\
& \underset{f \text{ is a homeomorphism}}{=} & g(1_B) \\
& \underset{g \text{ is a homeomorphism}}{=} & 1_C
\end{array}$$

Finally using [theorem: 4.70] we have then that

$$g(f(A)) \text{ is a sub-field of } \langle C, \oplus_C, \odot_C \rangle$$

2. Using [theorem: 2.75] we have that $g \circ f: A \to g(f(A))$ is a bijection which combined with (1) proves that $g \circ f$ is a field isomorphism. □

Definition 4.72. *$\langle F, \oplus, \odot, \leqslant \rangle$ is a ordered field if*

1. *$\langle F, \oplus, \odot \rangle$ is a field*
2. *$\langle F, \leqslant \rangle$ is totally ordered*

3. *$\forall x, y, z \in F$ with $x \leqslant y$ we have $x + z \leqslant y + z$*
4. *$\forall x, y \in F$ with $0 < x$ and $0 < y$ we have $0 < x \odot y$*

Theorem 4.73. *If $\langle F, \oplus, \odot, \leqslant \rangle$ is a ordered ring then we have :*

1. *$\forall x, y, z \in F$ we have $x < y \Leftrightarrow x \oplus z < y \oplus z$*
2. *$\forall x, y, z \in F$ we have $x \leqslant y \Leftrightarrow x \oplus z \leqslant y \oplus z$*
3. *$\forall x, y \in F$ we have $x < y \Leftrightarrow 0 < y \oplus (-x)$*
4. *$\forall x, y \in F$ we have $x \leqslant y \Leftrightarrow 0 \leqslant y \oplus (-x)$*
5. *$\forall x, y \in F$ we have $x < y \Leftrightarrow -y < -x$*
6. *$\forall x, y \in F$ we have $x \leqslant y \Leftrightarrow -y \leqslant -x$*
7. *$\forall x, y, z \in F$ with $0 < z$ we have $x < y \Leftrightarrow x \odot z < y \odot z$*
8. *$\forall x, y, z \in F$ with $0 < z$ we have $x \leqslant y \Leftrightarrow x \odot z \leqslant y \odot z$*
9. *$\forall x, y, z \in F$ with $0 \leqslant z$ and $x \leqslant y$ we have $x \odot \leqslant y \odot z$*
10. *$\forall x, y, z \in F$ with $z < 0$ we have $x < y \Leftrightarrow y \odot z < x \odot z$*
11. *$\forall x, y, z \in F$ with $z < 0$ we have $x \leqslant y \Leftrightarrow y \odot z \leqslant x \odot z$*
12. *$\forall x, y, z \in F$ with $z \leqslant 0$ and $x \leqslant y$ we have $y \odot z \leqslant x \odot z$*
13. *$\forall x \in F$ we have $0 \leqslant x \cdot x \underset{\text{def}}{=} x^2$, further if $0 \neq x$ then $0 < x \odot x \underset{\text{def}}{=} x^2$*
14. *$0 \leqslant 1$*
15. *$\forall x, y \in F$ with $0 < x < y$ we have that $x^2 < y^2$ where $x^2 = x \odot x$ and $y^2 = y \odot y$*
16. *$\forall x, y \in F$ with $0 \leqslant x \leqslant y$ we have that $x^2 \leqslant y^2$ where $x^2 = x \odot x$ and $y^2 = y \odot y$*
17. *$\forall x \in F$ with $0 < x$ we have $0 < x^{-1}$*
18. *$\forall x, y \in F$ we have $0 < x < y \Leftrightarrow 0 < y^{-1} < x^{-1}$*
19. *$\forall x, y \in F$ we have $0 < x \leqslant y \Leftrightarrow 0 < y^{-1} \leqslant x^{-1}$*
20. *$\forall x \in F$ with $0 < x < 1$ we have $1 < x^{-1}$*
21. *$\forall x \in F$ with $0 < x \leqslant 1$ we have $1 \leqslant x^{-1}$*

Proof. Using [definitions: 4.49, 4.72] and the fact that a field is automatically a ring we have that $(1 - 14)$ follows from [theorem: 4.50] so that we only have to proof (15-17). So

1. This follows from [theorem: 4.50 (1)].
2. This follows from [theorem: 4.50 (2)].
3. This follows from [theorem: 4.50 (3)].
4. This follows from [theorem: 4.50 (4)].
5. This follows from [theorem: 4.50 (5)].
6. This follows from [theorem: 4.50 (6)].
7. This follows from [theorem: 4.50 (7)].
8. This follows from [theorem: 4.50 (8)].
9. This follows from [theorem: 4.50 (9)].
10. This follows from [theorem: 4.50 (10)].
11. This follows from [theorem: 4.50 (11)].
12. This follows from [theorem: 4.50 (12)].
13. This follows from [theorem: 4.50 (13)].
14. This follows from [theorem: 4.50 (14)].

15. This follows from [theorem: 4.50 (15)].
16. This follows from [theorem: 4.50 (16)].
17. Let $x \in R$ with $0 < x$. By the totally ordering we have for x^{-1}:

    $x^{-1} = 0$. Then $1 = x^{-1} \odot x = 0 \odot x \underset{\text{[theorem: 4.39]}}{=} 0$ so that $x = x \odot 1 = x \odot 0 \underset{\text{[theorem: 4.39]}}{=} 0$ contradicting $0 < x$. So this case does not occur.

    $x^{-1} < 0$. Then by (7) we have that $1 = x^{-1} \odot x < 0 \odot x \underset{\text{[theorem: 4.39]}}{=} 0$ so that $1 < 0$ contradicting $0 \leqslant 1$, so this case will never occur.

    $0 < x^{-1}$. This is the remaining case proving (15).

18.

    $\Rightarrow$. If $0 < x < y$ then by (15) we have that $0 < x^{-1}$ and $0 < y^{-1}$. Hence

$$\begin{aligned} x < y &\underset{(7)}{\Rightarrow} x \odot x^{-1} < y \odot x^{-1} \\ &\Rightarrow 1 < y \odot x^{-1} \\ &\Rightarrow 1 < x^{-1} \odot y \\ &\underset{(7)}{\Rightarrow} 1 \odot y^{-1} < (x^{-1} \odot y) \odot y^{-1} \\ &\Rightarrow y^{-1} < x^{-1} \end{aligned}$$

    $\Leftarrow$. If $0 < y^{-1} < x^{-1}$ then by (15) we have $0 < (x^{-1})^{-1}$ and $0 < (y^{-1})^{-1}$ which using [theorem: 4.53] proves that $0 < x$ and $0 < y$ so that

$$\begin{aligned} y^{-1} < x^{-1} &\underset{(7)}{\Rightarrow} y^{-1} \odot x < x^{-1} \odot x \\ &\Rightarrow y^{-1} \odot x < 1 \\ &\Rightarrow x \odot y^{-1} < 1 \\ &\underset{(7)}{\Rightarrow} (x \odot y^{-1}) \odot y < 1 \odot y \\ &\Rightarrow x < y \end{aligned}$$

19.

    $\Rightarrow$. If $0 < x \leqslant y$ then by (15) we have that $0 < x^{-1}$ and $0 < y^{-1}$. Hence

$$\begin{aligned} x \leqslant y &\underset{(7)}{\Rightarrow} x \odot x^{-1} \leqslant y \odot x^{-1} \\ &\Rightarrow 1 \leqslant y \odot x^{-1} \\ &\Rightarrow 1 \leqslant x^{-1} \odot y \\ &\underset{(7)}{\Rightarrow} 1 \odot y^{-1} \leqslant (x^{-1} \odot y) \odot y^{-1} \\ &\Rightarrow y^{-1} \leqslant x^{-1} \end{aligned}$$

    $\Leftarrow$. If $0 < y^{-1} \leqslant x^{-1}$ then by (15) we have $0 < (x^{-1})^{-1}$ and $0 < (y^{-1})^{-1}$ which using [theorem: 4.53] proves that $0 < x$ and $0 < y$ so that

$$\begin{aligned} y^{-1} \leqslant x^{-1} &\underset{(7)}{\Rightarrow} y^{-1} \odot x \leqslant x^{-1} \odot x \\ &\Rightarrow y^{-1} \odot x \leqslant 1 \\ &\Rightarrow x \odot y^{-1} \leqslant 1 \\ &\underset{(7)}{\Rightarrow} (x \odot y^{-1}) \odot y \leqslant 1 \odot y \\ &\Rightarrow x \leqslant y \end{aligned}$$

20. As $0 < x < 1$ we have by (16) that $0 < 1^{-1} < x^{-1}$ hence $1 < x^{-1}$.
21. As $0 < x \leqslant 1$ we have by (16) that $0 < 1^{-1} \leqslant x^{-1}$ hence $1 \leqslant x^{-1}$.

$\square$

Chapter 5
Natural Numbers

Now we build a tower of different sets of numbers each one based on a previous one.

$$\begin{array}{c}\text{Natural Numbers } (\mathbb{N}_0)\\ \Downarrow\\ \text{Integer Numbers } (\mathbb{Z})\\ \Downarrow\\ \text{Rational Numbers } (\mathbb{Q})\\ \Downarrow\\ \text{Real Numbers } (\mathbb{R})\\ \Downarrow\\ \text{Complex Numbers } (\mathbb{C})\end{array}$$

Table 5.1.

We start with the the Natural Numbers. Using the set of Natural numbers we can introduce the concept of finite sets, infinite sets, denumerable sets, countable sets, mathematical induction and recursion. Further we will introduce a total ordering on $\mathbb{N}_0$ and prove that $\langle\mathbb{N}_0, \leqslant\rangle$ is well ordered. Finally two arithmetic operators sum $(+)$ and product $(\cdot)$ are introduced, unfortunately it will turn out that $\langle\mathbb{N}_0, +\rangle$ is not a group but only a semi-group so that for example $x+1=0$ has no solution. To solve this we introduce a new set of numbers, the integer numbers $(\mathbb{Z})$ and embed $\mathbb{N}_0$ in $\mathbb{Z}$ by creating a group and order isomorphism $i_{\mathbb{N}_0}: \mathbb{N}_0 \rightarrow \mathbb{Z}_0^+$ where $\mathbb{Z}_0^+ \subseteq \mathbb{Z}$ is the embedding of $\mathbb{N}_0$ in $\mathbb{Z}$. Although we will have succeeded in making $\langle\mathbb{Z}, +, \cdot\rangle$ a ring a equation like $2 \cdot x = 1$ has no solution, for this we must have a field. Hence the next step is to construct a new set of numbers, the rational numbers $(\mathbb{Q})$, so that $\langle\mathbb{Q}, +, \cdot\rangle$ forms a field. We will embed then $\mathbb{Z}, \mathbb{N}_0$ in $\mathbb{Q}$ by creating ring, group and order isomorphisms $i_{\mathbb{Z}\rightarrow\mathbb{Q}}: \mathbb{Z} \rightarrow \mathbb{Z}_{\mathbb{Q}}$, $i_{\mathbb{N}_0\rightarrow\mathbb{Q}}: \mathbb{N}_0 \rightarrow \mathbb{N}_{0,\mathbb{Q}}$ where $\mathbb{Z}_{\mathbb{Q}} \subseteq \mathbb{Q}$ is the embedding of $\mathbb{Z}$ in $\mathbb{Q}$ and $\mathbb{N}_{0,\mathbb{Q}}$ is the embedding of $\mathbb{N}_0$ in $\mathbb{Q}$. However the totally ordered set $\langle\mathbb{Q}, \leqslant\rangle$ is not conditional complete which is important for analysis. So we create the set of real numbers $\mathbb{R}$ so that $\langle\mathbb{R}, \leqslant\rangle$ is conditionally complete. Next we create embeddings $\mathbb{N}_{0,\mathbb{R}}$, $\mathbb{Z}_{\mathbb{R}}$, $\mathbb{Q}_{\mathbb{R}}$ in $\mathbb{R}$ for $\mathbb{N}_0$, $\mathbb{Z}$ and $\mathbb{Q}$. It turns out that a equation like $x^2 = -1$ has no solution, so we create a new set of numbers, the complex numbers $(\mathbb{C})$ to solve this defect. Then we embed $\mathbb{N}_0$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ in $\mathbb{C}$ in the form $\mathbb{N}_{0,\mathbb{C}}$, $\mathbb{Z}_{\mathbb{C}}$, $\mathbb{Q}_{\mathbb{C}}$ and $\mathbb{R}_{\mathbb{C}}$ such that $\mathbb{N}_{0,\mathbb{C}} \subseteq \mathbb{Z}_{\mathbb{C}} \subseteq \mathbb{Q}_{\mathbb{C}} \subseteq \mathbb{R}_{\mathbb{C}} \subseteq \mathbb{C}$. From then on, for the rest of this book we work with $\mathbb{C}$ and these embeddings, to avoid excessive notation we use the symbols $\mathbb{N}_0$, $\mathbb{Z}$. $\mathbb{Q}$ and $\mathbb{R}$ for these embeddings.

5.1 Definition of the Natural Numbers

We are now ready to define the first set of numbers namely the natural numbers which forms the basic of the other number systems but also of the important concepts of finite, infinite sets, countable sets, recursion and mathematical induction. To define the set of natural numbers recall the following definitions and axiom.

Definition 5.1. (Successor Set) *A **set** A is a **successor set** iff*

1. $\varnothing \in A$
2. *If* $X \in A \Rightarrow X \bigcup \{X\} \in A$

[see definition: 1.51]

Axiom 5.2. (Axiom of Infinity) *There exists a successor set [see axiom: 1.52].*

Definition 5.3. (Natural numbers) *The set of **natural numbers** $\mathbf{N_0}$ **is defined** by*

$$\mathbb{N}_0 = \bigcap \{S \mid S \text{ is a successor set}\}$$

Theorem 5.4. $\mathbb{N}_0$ *is a set*

Proof. By the axiom of infinity it follows that $\{S|S \text{ is a successor set}\} \neq \varnothing$ so that by [theorem: 1.60 (5)] $\bigcap \{S|S \text{ is a successor set}\}$ is a set. □

Theorem 5.5. *If* $n \in \mathbb{N}_0$ *then* $n \bigcup \{n\} \in \mathbb{N}_0$

Proof. If $n \in \mathbb{N}_0$ then for $\forall A \in \{S|S \text{ is a successor set}\}$ we have $n \in A$ so that by definition of a successor set we have $n \bigcup \{n\} \in A$ so that $n \bigcup \{n\} \in \bigcap \{S|S \text{ is a successor set}\} = \mathbb{N}_0$. □

The above theorem allows us to define the successor function

Definition 5.6. (Successor Function) *The function defined by*

$$s: \mathbb{N}_0 \rightarrow \mathbb{N}_0 \text{ where } s(n) = n \bigcup \{n\}$$

is called the ***successor function****.*

The set $\mathbb{N}_0$ is not empty as is shown in the next theorem.

Theorem 5.7. $\varnothing \in \mathbb{N}_0$

Proof. If A is a successor set then by definition $\varnothing \in A$ so that $\varnothing \in \bigcap \{A|A \text{ is a successor set}\}$ □

Further using the successor function we have that $s(\varnothing)$, $s(s(\varnothing))$ etc. are all elements of $\mathbb{N}_0$,. we introduce a special notation for this elements that corespondents with the notation used for counting.

Notation 5.8. *We define the numbers 0,1,2,3,... as follows*

1. $0 = \varnothing$
2. $1 = s(0) = s(\varnothing) = \varnothing \bigcup \{\varnothing\} = \{\varnothing\} = \{0\}$
3. $2 = s(1) = s(\varnothing) \bigcup \{s(\varnothing)\} = \{\varnothing\} \bigcup \{\{\varnothing\}\} = \{\varnothing, \{\varnothing\}\} = \{0, 1\}$
4. $3 = s(2) = \{\varnothing, \{\varnothing\}\} \bigcup \{\{\varnothing, \{\varnothing\}\}\} = \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\} = \{0, 1, 2\}$
5. ...

The notation $\mathbb{N}_0$ may seem a little bit strange, the fact is that many mathematicians don't consider 0 a natural number. To express that $0 \in \mathbb{N}_0$ we add the 0 subscript. If we want to indicate that $0 \notin \mathbb{N}_0$ we use the following definition.

Definition 5.9. $\mathbb{N} = \mathbb{N}_0 \setminus \{0\}$

Theorem 5.10. *If* $n \in \mathbb{N}_0$ *then* $s(n) \neq 0$ *in particula* $0 \neq 1$

Proof. By definition we have $s(n) = n \bigcup \{n\}$ so that $n \in s(n)$ proving that $s(n) \neq \varnothing = 0$ □

We introduce now the very important principle of **Mathematical Induction**.

Theorem 5.11. (Mathematical Induction) *If* $X \subseteq \mathbb{N}_0$ *such that*

1. $0 \in X$
2. *If* $n \in X$ *then* $s(n) \in X$

then $X = \mathbb{N}_0$

Proof. By $(1), (2)$ it follows that X is a successor set so that $X \in \{A|A \text{ is a successor set}\}$ hence by [theorem: 1.60] $\mathbb{N}_0 = \bigcap \{A|A \text{ is a successor set}\} \subseteq X$, which together with $X \subseteq \mathbb{N}$ proves that $X = \mathbb{N}$. □

Theorem 5.12. *Let* $n, m \in \mathbb{N}_0$ *then if* $m \in s(n)$ *we have* $m \in n \vee m = n$

Proof. If $m \in s(n) = n \bigcup \{n\}$ then we have either $m \in n$ or $m \in \{n\} \Rightarrow m = n$ □

Definition 5.13. *A set A is* ***transitive*** *if $\forall x \in A$ we have $x \subseteq A$.*

As a application of mathematical induction we prove that every natural number is transitive, this fact will be used later, when we define a order relation on $\mathbb{N}_0$ to prove transitivity, hence the name for this property.

Theorem 5.14. *$\forall n \in \mathbb{N}_0$ we have that n is transitive [in other words $\forall x \in n$ we have $x \subseteq n$]*

Proof. We prove this by mathematical induction, let $\mathcal{S} = \{n \in \mathbb{N}_0 | n \text{ is transitive}\}$ then clearly $S \subseteq \mathbb{N}_0$. Further we have

- $\boldsymbol{0 \in S}$**.** Because $\forall x \in \emptyset \vdash x \subseteq \emptyset$ is satisfied vacuously.
- $\boldsymbol{n \in S \Rightarrow s(n) \in S}$**.** If $n \in S$ then we have for $m \in s(n)$ by the previous theorem [theorem: 5.12] the following cases:
  - $\boldsymbol{m \in n}$**.** Then as $n \in S$, n is transitive so that $m \subseteq n \subseteq n \bigcup \{n\} = s(n)$
  - $\boldsymbol{m = n}$**.** Then $m = n \subseteq n \bigcup \{n\} = s(n)$

  So $\forall m \in s(n)$ we have $m \subseteq s(n)$ which proves that $s(n)$ is transitive, hence $s(n) \in S$

Using mathematical induction [see theorem: 5.11] it follows then that $S = \mathbb{N}_0$. So if $n \in \mathbb{N}_0$ then $n \in S$ or n is transitive. □

Another application of transitivity and mathematical induction is the following theorem.

Theorem 5.15. *If $n \in \mathbb{N}_0$ then $n \neq s(n)$*

Proof. Let $S = \{n \in \mathbb{N}_0 | n \neq s(n)\}$ then we have

- $\boldsymbol{0 \in S}$**.** By [theorem: 5.10] $0 \neq s(0)$.
- $\boldsymbol{n \in S \Rightarrow s(n) \in S}$**.** Assume that $s(s(n)) = s(n)$. As $s(s(n)) = s(n) \bigcup \{s(n)\}$ we have that $s(n) \in s(s(n)) = s(n)$, so $s(n) \in n \bigcup \{n\}$. As $n \in S$ we have that $n \neq s(n)$ so we must have that $s(n) \in n$. As by [theorem: 5.14] $s(n)$ is transitive it follows that $s(n) \subseteq n$, further we have that $n \subseteq n \bigcup \{n\} = s(n)$. So we conclude that $n = s(n)$ proving $n \notin S$ which contradicts $n \in S$. So we must have that $s(s(n)) \neq s(n)$ proving that $s(n) \in S$.

Using mathematical induction it follows then that $\mathbb{N}_0 = S$ so if $n \in \mathbb{N}_0$ then $n \in S$ and thus $n \neq s(n)$. □

The next theorem shows that the successor function is a injection.

Theorem 5.16. *If $n, m \in \mathbb{N}_0$ is such that $s(n) = s(m)$* then n=m. *In other words*

$$s: \mathbb{N}_0 \rightarrow \mathbb{N}_0 \text{ is injective}$$

Proof. As $n \in n \bigcup \{n\} = s(n) = s(m)$ and $m \in m \bigcup \{m\} = s(m) = s(n)$ we have that $n \in s(m) \wedge m \in s(n)$. Using [theorem: 5.12] this becomes

$$(n \in m \vee n = m) \wedge (m \in n \vee n = m) \Rightarrow (n \in m \wedge m \in n) \vee n = m$$

If $n = m$ we are done. So we must look at the case that $m \in n \wedge n \in m$. By transitivity [theorem: 5.14] we have then $n \subseteq m$ and $m \subseteq n$ proving that $n = m$. □

The above theorems are part of what is in number theory the Peano Axioms.

Theorem 5.17. (Peano Axioms) $\mathbb{N}_0$ *satisfies the following so called Peano Axioms*

1. $0 \in \mathbb{N}_0$
2. *If $n \in \mathbb{N}_0$ then $s(n) \in \mathbb{N}_0$*
3. *$\forall n \in \mathbb{N}_0$ we have that $s(n) \neq 0$*
4. *If $X \subseteq \mathbb{N}_0$ is such that*
   - *a.* $0 \in X$
   - *b.* $n \in X \Rightarrow s(n) \in X$

   then $X = \mathbb{N}_0$

5. *If $n, m \in \mathbb{N}_0$ is such that $s(n) = s(m)$ then $n = m$*

Proof.

1. See [theorem: 5.7]
2. See [definition: 5.6]
3. See [theorem: 5.10]
4. See [theorem: 5.11]
5. See [theorem: 5.16] □

Theorem 5.18. *If $n \in \mathbb{N}_0 \wedge n \neq 0$ then $\exists! m \in \mathbb{N}_9$ such that $n = s(m)$*

Proof. We use mathematical induction to prove this. So let

$$S = \{n \in \mathbb{N}_0 | (n = 0) \vee (\exists! m \in \mathbb{N}_0 \text{ such that } n = s(m))\} \subseteq \mathbb{N}_0$$

then we have:

$0 \in S$. As $0 = 0$ we have that $0 \in S$.

$n \in S \Rightarrow s(n) \in S$. Consider $s(n)$ then by [theorem: 5.10] $s(n) \neq 0$, further we have that $m = n$ satisfies $s(n) = s(m)$ proving the existence part. Assume that there is another $m' \in \mathbb{N}_0$ such that $s(n) = s(m')$, then by [theorem: 5.16] we have $n = m'$, proving uniqueness. So $s(n) \in S$.

Mathematical induction [see: 5.11] proves then that $\mathbb{N}_0 = S$. So if $n \in \mathbb{N}_0$ with $n \neq 0$ we have as $n \in S$ that $\exists! m \in \mathbb{N}_0$ such that $n = s(m)$. □

5.2 Recursion

Recursion will be used to essential define things in terms of itself. It is the mathematical equivalent of iteration in many programming languages. Actually, functional languages that are mathematical oriented, like Haskell, have no iteration and loop constructs at all and relay fully on recursion. Recursion is based on the definition of a recursive function that takes the role of iterating. The following theorem ensures the existence of such a function.

Theorem 5.19. (Recursion) *Let A be a set, $a \in A$ and $f: A \rightarrow A$ a function then there exists a **unique** function*

$$\lambda: \mathbb{N}_0 \rightarrow A$$

such that

1. $\lambda(0) = a$
2. *$\forall n \in \mathbb{N}_0$ we have $\lambda(s(n)) = f(\lambda(n))$*

Proof. Define

$$\mathcal{G} = \{G | G \subseteq \mathbb{N}_0 \times A \text{ such that } (0, a) \in G \text{ and } \forall n \in \mathbb{N}_0 \text{ that } (n, x) \in G \Rightarrow (s(n), f(x)) \in G\}$$

Define $G = \mathbb{N}_0 \times A$ then as $0 \in \mathbb{N}_0$ and $a \in A$ we have $(0, a) \in \mathbb{N}_0 \times A$. Further if $(n, x) \in \mathbb{N}_0 \times A$ then $n \in \mathbb{N}_0$ and $x \in A$ so that $s(n) \in \mathbb{N}_0$ and $f(x) \in A$, hence $(s(n), f(x)) \in \mathbb{N}_0 \times A$. So

$$\mathbb{N}_0 \times A \in \mathcal{G} \tag{5.1}$$

We prove now that

$$\text{If } \lambda = \bigcap \mathcal{G} \text{ then } \lambda \in \mathcal{G},\ \lambda \subseteq N_0 \times A \text{ and } (0, a) \in \lambda \tag{5.2}$$

Proof.

1. By [eq: 5.1] we have $\mathbb{N}_0 \times A \in \mathcal{G}$ so that by [theorem: 1.60] $\bigcap \mathcal{G} \subseteq \mathbb{N}_0 \times A$ hence $\lambda \subseteq \mathbb{N}_0 \times A$
2. $\forall G \in \mathcal{G}$ we have by definition that $(0, a) \in G$ hence $(0, a) \in \bigcap \mathcal{G}$ or $(0, a) \in \lambda$
3. If $(n, x) \in \bigcap \mathcal{G}$ then $\forall G \in \mathcal{G}$ we have $(n, x) \in G \Rightarrow (s(n), f(x)) \in G$, so that $(s(n), f(x)) \in \bigcap \mathcal{G}$.

Using (1),(2) and (3) it follows that $\bigcap \mathcal{G} \in \mathcal{G}$. $\square$

If $x \in \operatorname{dom}(\lambda)$ then $\exists y$ such that $(x, y) \in \lambda \subseteq \mathbb{N}_0 \times A$ [see eq: 5.2] so that $x \in \mathbb{N}_0$, hence

$$\operatorname{dom}(\lambda) \subseteq \mathbb{N}_0 \tag{5.3}$$

As by [eq: 5.2] $(0, a) \in \lambda$ we have that

$$0 \in \operatorname{dom}(\lambda) \tag{5.4}$$

If $n \in \operatorname{dom}(\lambda)$ then then $\exists x$ such that $(n, x) \in \lambda$, as by [eq: 5.2] $\lambda \in \mathcal{G}$, we have $(s(n), f(x)) \in \lambda$ so that $s(n) \in \operatorname{dom}(\lambda)$. In other words we have

$$\text{if } n \in \operatorname{dom}(\lambda) \text{ then } s(n) \in \operatorname{dom}(\lambda) \tag{5.5}$$

Now [eq: 5.3], [eq: 5.4] and [eq: 5.5] are the conditions for mathematical induction [theorem: 5.11], so we have proved that

$$\operatorname{dom}(\lambda) = \mathbb{N}_0 \tag{5.6}$$

We use now mathematical induction to prove that λ is the graph of a function. Let

$$S = \{n \in \mathbb{N}_0 | \exists! x \text{ such that } (n, x) \in \lambda\} \subseteq \mathbb{N}_0$$

then we have:

$0 \in S$. By [eq: 5.2] we have $(0, a) \in \lambda$. Assume that $\exists x \in A$ with $x \neq a$ such that $(0, x) \in \lambda$, then $(0, a) \neq (0, x)$. Define now $\beta = \lambda \setminus \{(0, x)\}$ then we have

1. $\beta \subseteq \lambda \subseteq \mathbb{N}_0 \times A$
2. As $(0, a) \neq (0, x)$ and $(0, a) \in \lambda$ we have $(0, a) \in \beta$
3. If $(n, y) \in \beta \underset{\beta \subseteq \lambda}{\Rightarrow} (n, y) \in \lambda$ so that $(s(n), f(x)) \in \lambda$, as by [theorem: 5.10] $s(n) \neq 0$ we have that $(s(n), f(x)) \neq (0, x)$, hence $(s(n), f(y)) \in \beta$

From (1),(2) and (3) it follows that $\beta \in \mathcal{G}$ so that by [theorem: 1.60] $\lambda = \bigcap \mathcal{G} \subseteq \mathcal{B}$ which as $(0, x) \in \lambda$ would give $(0, x) \in \beta = \lambda \setminus \{(0, x)\}$ a contradiction. So the assumption is wrong and we must have that $x = a$, proving uniqueness, hence that $0 \in S$.

$n \in S \Rightarrow s(n) \in S$. As $n \in S$ there exist a **unique** $x \in \mathcal{S}$ such that $(n, x) \in \lambda$. As $(n, x) \in \lambda$ we have as $\lambda \in \mathcal{G}$ that $(s(n), f(x)) \in \lambda$. Assume now that $\exists y$ such that $(s(n), y) \in \lambda$ and $f(x) \neq y$. Define then $\beta = \lambda \setminus \{(s(n), y)\}$ then we have:

1. $\beta \subseteq \lambda \subseteq \mathbb{N}_0 \times A$
2. As by [theorem: 5.15] $s(n) \neq 0$ we have that $(0, a) \neq (s(n), y)$, as further $(0, a) \in \lambda$ it follows that $(0, a) \in \beta$
3. If $(m, z) \in \beta$ then $(m, z) \in \lambda$ so that $(s(m), f(z)) \in \lambda$ we must now consider two cases for $s(n), s(m)$:

   $s(m) = s(n)$. Then by [theorem: 5.16] we have $n = m$ so that $(n, z) = (m, z) \in \lambda$. As $n \in S$ and we have $(n, x) \in \lambda$ it follows that $z = x$. So that $(s(m), f(z)) = (s(n), f(x)) \neq (s(n), y)$ [as we assumed that $y \neq f(x)$] hence we have that $(s(m), f(z)) \in \beta$.

   $s(m) \neq s(n)$. then $(s(m), f(z)) \neq (s(n), y)$ so that $(s(m), f(z)) \in \beta$

   So we have prove that if $(m, z) \in \beta$ then $(s(m), f(z)) \in \beta$

From (1),(2) and (3) it follows that $\beta \in \mathcal{G}$ but then using [theorem: 1.60] we have that $\lambda = \bigcap \mathcal{G} \subseteq \beta$ which as $(s(n), y) \in \lambda$ leads to $(s(n), y) \in \beta = \lambda \setminus \{(s(n), y)\}$ a contradiction. So the assumption is wrong and we must have that $y = f(x)$ proving **uniqueness**, hence we have that $s(n) \in S$.

Using mathematical induction [theorem: 5.11] it follows that $\mathbb{N}_0 = S$. So if $(n, x), (n, x') \in \lambda$ then $n \in \mathbb{N}_0 = S$ so that $y = y'$ giving

$$\text{If } (n, x), (n, x') \in \lambda \text{ then } x = x' \tag{5.7}$$

From [eq: 5.2], [eq: 5.6] and [eq: 5.7] it follows that

$$\lambda: \mathbb{N}_0 \to A \text{ is a function} \tag{5.8}$$

As $\lambda \in \mathcal{G}$ we have that $(0, a) \in \lambda \Rightarrow a = \lambda(0)$, further if $n \in \mathbb{N}_0 = \operatorname{dom}(\lambda)$ then $\exists x$ such that $(n,x) \in \lambda$ and $(s(n), f(x)) \in \lambda$, Now $(n, x) \in \lambda$ is equivalent with $\lambda(n) = x$ and $(s(n), f(x)) \in \lambda$ is equivalent with $\lambda(s(n)) = f(x) = f(\lambda(n))$. So we have for λ that

$$\lambda(0) = a \text{ and } \forall n \in \mathbb{N}_0 \text{ we have } \lambda(s(n)) = f(\lambda(n)) \tag{5.9}$$

So we have proved the existence of our function, next we must prove that this function is unique. Assume that there exist another function

$$\beta: \mathbb{N}_0 \to A \text{ such that } \beta(0) = a \text{ and } \forall n \in \mathbb{N}_0 \text{ we have } \lambda(s(n)) = f(\lambda(n))$$

We proceed by mathematical induction, so define $T = \{n \in \mathbb{N}_0 | \lambda(n) = \beta(\lambda)\}$ then we have

$0 \in T$. As $\lambda(0) = a = \beta(0)$ we have that $0 \in T$.

$n \in T \Rightarrow s(n) \in T$. As $n \in T$ we have $\lambda(n) = \beta(n)$ but then $\lambda(s(n)) = f(\lambda(n)) = \beta(s(n))$ so that $s(n) \in T$

Using mathematical induction [theorem: 5.11] we have then $T = \mathbb{N}_0$. So $\forall n \in \mathbb{N}_0$ we have $n \in T$ hence $\lambda(n) = \beta(n)$ which by [theorem: 2.41] proves that

$$\lambda = \beta$$

□

Corollary 5.20. *If A is a set, $a \in A$ and $f: A \to A$ a function then there exists a unique function*

$$\lambda: \mathbb{N}_0 \to A$$

such that

1. $\lambda(0) = a$
2. $\forall n \in \mathbb{N}_0$ *we have* $\lambda(s(n)) = f(\lambda(n))$
3. *If* $a \notin f(A)$ *and* $f: A \to A$ *is injective then* λ *is injective*

Proof. The first part is easy. Using recursion [theorem: 5.19] there exists a function

$$\lambda: \mathbb{N}_0 \to A$$

such that

$$\lambda(0) = a \text{ and } \forall n \in \mathbb{N}_0 \text{ we have } \lambda(s(n)) = f(\lambda(n))$$

We use now mathematical induction to prove (3). Assume that $a \notin f(A)$ and take

$$S = \{n \in \mathbb{N}_0 | \forall m \in \mathbb{N}_0 \text{ with } \lambda(n) = \lambda(m) \text{ we have n=m}\}$$

then we have:

$0 \in S$. If $\lambda(m) = \lambda(0)$ then as $\lambda(0) = a$ we have that $\lambda(m) = a$. Assume that $m \neq 0$ then by [theorem: 5.18] there exists a $k \in \mathbb{N}_0$ such that $m = s(k)$ so that $a = \lambda(m) = \lambda(s(k)) = f(\lambda(k))$, which proves that $a \in f(A)$ contradicting $a \notin f(A)$. Hence we must have $m = 0$ so that $0 \in S$.

$n \in S \Rightarrow s(n) \in S$. Let $m \in \mathbb{N}_0$ such that $\lambda(s(n)) = \lambda(m)$. Assume that $m = 0$ then $\lambda(s(n)) = \lambda(m) = \lambda(0) = a$ so that $f(\lambda(n)) = \lambda(s(n)) = a$, resulting in $a \in f(A)$ contradicting $a \notin f(A)$. Hence we must have that $m \neq 0$. Using [theorem: 5.18] there exists a $k \in \mathbb{N}_0$ such that $m = s(k)$, from $\lambda(s(n)) = \lambda(m)$ it follows then that $\lambda(s(n)) = \lambda(s(k))$ so that $f(\lambda(n)) = \lambda(s(n)) = \lambda(s(k)) = f(\lambda(k))$. As f is injective we have $\lambda(n) = \lambda(k)$. Now as $n \in S$ we must have $n = k$ or $s(n) = s(k) = m$. This proves that $\forall m \in \mathbb{N}_0$ with $\lambda(s(n)) = \lambda(m)$ we have $s(n) = m$, hence $s(n) \in S$

Using mathematical induction [theorem: 5.11] it follows that $\mathbb{N}_0 = S$. So if $n, m \in \mathbb{N}_0$ is such that $\lambda(n) = \lambda(m)$ then $n \in S$ and as $m \in \mathbb{N}_0$ we have $n = m$, proving that

$$\lambda \text{ is injective}$$

□

Remark 5.21. To understand how recursion works in the above theorem consider the following, Let $f: A \to A$ a function, $a \in A$ and $\lambda: \mathbb{N}_0 \to A$ such that $\lambda(0) = a$ and $\lambda(s(n)) = f(\lambda(n))$

$$\begin{array}{rcl}
\lambda(0) & = & a \\
\lambda(1) = \lambda(s(n)) & = & f(\lambda(0)) = f(a) \\
\lambda(2) = \lambda(s(1)) & = & f(\lambda(1)) = f(f(a)) \\
\lambda(3) = \lambda(s(2)) & = & f(\lambda(2)) = f(f(f(a))) \\
& \ldots & \\
\lambda(n) & = & \overbrace{f(f(\ldots(f(a))))}^{\text{n times}}
\end{array}$$

so $\lambda(n)$ is the result of applying f n-times on a value a. If $a \notin f(A)$ and f is injective then λ is injective and we would have that $f(a)$, $f(f(z)), f(f(f(a))), \ldots, \overbrace{f(f(\ldots(f(a))))}^{\text{n times}}$ are all different numbers.

To see the conditions for injectivity of λ consider the following two examples:

Example 5.22. Define $f\colon \{1,2,3\} \rightarrow f(\{1,2,3\})$ by $f(i) = \begin{cases} 2 \text{ if } i=1 \\ 3 \text{ if } i=2 \\ 2 \text{ if } i=1 \end{cases}$ (so f is not injective) and $a=3$

then we have

$$\begin{array}{rcl} \lambda(0) & = & 3 \\ \lambda(1) & = & f(3)=2 \\ \lambda(2) & = & f(f(3))=f(2)=1 \\ \lambda(3) & = & f(f(f(3)))=f(1)=2 \\ \lambda(4) & = & f(f(f(3)))=f(2)=1 \\ & \ldots & \end{array}$$

So that $\lambda\colon \mathbb{N}_0 - A$ is clearly not injective.

Example 5.23. Take $f\colon \{1,2,3\} \rightarrow \{1,2,3\}$ by $f(i) = \begin{cases} 2 \text{ if } i=1 \\ 3 \text{ if } i=2 \\ 1 \text{ if } i=3 \end{cases}$ so that f is injective and $a=2$ so that $a \in f(\{1,2,3\})$ then we have

$$\begin{array}{rcl} \lambda(0) & = & 2 \\ \lambda(1) & = & f(2)=1 \\ \lambda(2) & = & f(f(2))=f(1)=2 \\ \lambda(3) & = & f(f(f(2)))=f(2)=1 \\ & \ldots & \end{array}$$

So that $\lambda\colon \mathbb{N}_0 \rightarrow \{1,2,3\}$ is not injective.

We can rephrase the above remark in the iteration principle that is useful in proofs using mathematical induction.

Theorem 5.24. (Iteration) *Let A be a non empty set and $f\colon A \rightarrow A$ a function. Then $\forall n \in \mathbb{N}_0$ there exist a function*

$$(f)^n\colon A \rightarrow A$$

such that

1. $(f)^0 = \mathrm{Id}_A$
2. $(f)^{s(n)} = f \circ (f)^n$

Proof. Let $a \in A$ and use the recursion [theorem: 5.19] to find a **unique** function

$$\lambda_a\colon \mathbb{N}_0 \rightarrow A \text{ such that } \lambda_a(0) = a \text{ and } \forall n \in \mathbb{N}_0\ \lambda_a(s(n)) = f(\lambda_a(n))$$

Define now

$$(f)^n\colon A \rightarrow A \text{ where } (f)^n(a) = \lambda_a(n)$$

Then we have

1. $\forall a \in A$ we have that $(f)^0(a) = \lambda_a(0) = a$ so that
$$(f)^0 = \mathrm{Id}_A$$
2. $\forall a \in A$ we have that $(f)^{s(n)}(a) = \lambda_a(s(n)) = f(\lambda_a(n)) = f((f)^n(a)) = (f \circ (f)^n)(a)$ so that
$$(f)^{s(n)} = f \circ (f)^n$$

□

As illustration of iteration let $f: A \rightarrow A$ then we have

$$\begin{array}{rcl}(f)^0 &=& \mathrm{Id}_A \\ (f)^1 = (f)^{s(0)} &=& f \circ (f)^0 = f \circ \mathrm{Id}_A = f \\ (f)^2 = (f)^{s(1)} &=& f \circ (f)^1 = f \circ f \\ (f)^3 = (f)^{s(2)} &=& f \circ (f)^2 = f \circ f \circ f \\ & \ldots & \\ (f)^n &=& \overbrace{f \circ \ldots \circ f}^{n \, \mathrm{times}}\end{array}$$

We can apply the above on a group to define new operations on the group.

Example 5.25. 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 $n \in \mathbb{N}$ $a \langle \oplus \rangle n = (\oplus_a)^n(e)$ where e is the neutral element in the group . So

$$\begin{array}{rcl}a\langle\oplus\rangle 0 &=& (\oplus_a)^0(e) = \mathrm{Id}_A(e) = e \\ a\langle\oplus\rangle 1 &=& (\oplus_a)^1(e) = \oplus_a(e) = a \oplus e = e \\ a\langle\oplus\rangle 2 &=& (\oplus_a)^2(e) = \oplus_a(\oplus_a(e)) = a \oplus (a \oplus e) = a \oplus (a \oplus e) = a \oplus a \\ a\langle\oplus\rangle 3 &=& (\oplus_a)^3(e) = (\oplus_a(\oplus_a(\oplus_a(e)))) = a \oplus (a \oplus (a \oplus e)) = a \oplus a \oplus a \\ & \ldots & \\ a\langle\oplus\rangle n &=& \overbrace{a \oplus \cdots \oplus a}^{\text{n times}}\end{array}$$

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 hence we have

1. Additive group $\langle A, + \rangle$ with neutral element 0 gives

$$\begin{array}{rcl}a \cdot 0 &=& 0 \\ a \cdot 1 &=& a \\ a \cdot 2 &=& a+a \\ a \cdot 3 &=& a+a+a \\ & \ldots & \\ a \cdot n &=& \overbrace{a + \cdots + a}^{\text{n times}}\end{array}$$

2. Multiplicative group $\langle A, \cdot \rangle$ with neutral element 1 gives

$$\begin{array}{rcl}a^0 &=& 1 \\ a^1 &=& a \\ a^2 &=& a \cdot a \\ a^3 &=& a \cdot a \cdot a \\ & \ldots & \\ a^n &=& \overbrace{a \cdot \cdots \cdot a}^{\text{n times}}\end{array}$$

Recursion is mostly used in it's step form to define recursive functions.

Theorem 5.26. (Recursion on $\mathbb{N}_0$ Step Form) *Let A be a set, $a \in A$ and $g: \mathbb{N} \times A \rightarrow A$ a function then there exist a **unique** function $\lambda: \mathbb{N}_0 \rightarrow A$ such that*

1. $\lambda(0) = a$
2. $\forall n \in \mathbb{N}_0$ *we have* $\lambda(s(n)) = g(n, \lambda(n))$

Proof. First we define the projection functions

$$\pi_1 \colon \mathbb{N}_0 \times A \to \mathbb{N}_0 \text{ where } \pi_1(n,x) = n$$

$$\pi_2 \colon \mathbb{N}_0 \times A \to A \text{ where } \pi_2(n,x) = x$$

Define now

$$\gamma \colon \mathbb{N}_0 \times A \to \mathbb{N}_0 \times A \text{ where } \gamma(x) = (s(\pi_1(x)), g(\pi_1(x), \pi_2(x))) \tag{5.10}$$

Using the iteration [theorem: 5.24] on the above functions gives $\forall n \in \mathbb{N}_0$ the existence of the function

$$(\gamma)^n \colon \mathbb{N}_0 \times A \to \mathbb{N}_0 \times A \text{ such that } (\gamma)^0 = \mathrm{Id}_{\mathbb{N}_0 \times A} \text{ and } \forall n \in \mathbb{N}_0 \text{ we have } (\gamma)^{s(n)} = \gamma \circ (\gamma)^n \tag{5.11}$$

We prove now by mathematical induction that $\forall n \in \mathbb{N}_0$ $\pi_1((\gamma)^n(0,a)) = n$. So let

$$S = \{n \in \mathbb{N}_0 | \pi_1((\gamma)^n(0,a)) = n\}$$

then we have:

$0 \in S$. As $\pi_1((\gamma)^0(0,a)) \underset{[\text{eq: } 5.11]}{=} \pi_1(\mathrm{Id}_{\mathbb{N}_0 \times A}(0,a)) = \pi_1(0,a) = 0$ we have that $0 \in S$

$n \in S \Rightarrow s(n) \in S$. We have

$$\begin{array}{lcl} \pi_1((\gamma)^{s(n)}(0,a)) & \underset{[\text{eq: } 5.11]}{=} & \pi_1((\gamma \circ (\gamma)^n)(0,a)) \\ & = & \pi_1(\gamma((\gamma)^n(0,n))) \\ & \underset{[\text{eq: } 5.10]}{=} & \pi_1(\pi_1((\gamma)^n(0,n)), g(\pi_1((\gamma)^n(0,a)), \pi_2((\gamma)^n(0,a)))) \\ & \underset{n \in S \Rightarrow \pi_1((\gamma)^n(0,a)) = n}{=} & \pi_1(n, g(n, \pi_2((\gamma)^n(0,a)))) \\ & = & n \end{array}$$

proving that $s(n) \in S$

Using mathematical induction [theorem: 5.11] we have $\mathbb{N}_0 = S$, hence

$$\forall n \in \mathbb{N}_0 \text{ we have } \pi_1((\gamma)^n(0,a)) = n \tag{5.12}$$

Define now

$$\lambda \colon \mathbb{N}_0 \to A \text{ by } \gamma(n) = \pi_2((\gamma)^n(0,a)) \tag{5.13}$$

then we have:

1. $\lambda(0) = \pi_2((\gamma)^0(0,a)) = \pi_2(\mathrm{Id}_{\mathbb{N}_0 \times A}(0,a)) = \pi_2(0,a) = a$
2. If $n \in \mathbb{N}_0$ then

$$\begin{array}{lcl} \lambda(s(n)) & = & \pi_2((\gamma)^{s(n)}(0,a)) \\ & \underset{[\text{eq: } 5.11]}{=} & \pi_2((\gamma \circ (\gamma)^n)(0,a)) \\ & = & \pi_2(\gamma((\gamma)^n(0,a))) \\ & \underset{[\text{eq: } 5.10}{=} & \pi_2(\pi_1((\gamma)^n(0,a)), g(\pi_1((\gamma)^n(0,a)), \pi_2((\gamma)^n(0,a)))) \\ & = & g(\pi_1((\gamma)^n(0,a)), \pi_2((\gamma)^n(0,a))) \\ & \underset{[\text{eq: } 5.12}{=} & g(n, \pi_2((\gamma)^n(0,a))) \\ & \underset{[\text{eq: } 5.13}{=} & g(n, \lambda(n)) \end{array}$$

This proves the existence of the function we are searching for. Now for uniqueness assume that there is a

$$\beta \colon \mathbb{N}_0 \to A \text{ such that } \beta(0) = a \text{ and } \forall n \in \mathbb{N}_0 \text{ that } \beta(s(n)) = g(n, \beta(n))$$

Define now $B = \{n \in \mathbb{N}_0 | \lambda(n) = \beta(n)\}$ then we have:

$0 \in B$. As $\beta(0) = a = \lambda(0)$ it follows that $0 \in B$.

$n \in B \Rightarrow s(n) \in B$. As

$$\beta(s(n)) = g(n, \beta(n)) \underset{n \in B}{=} g(n, \lambda(n)) = \lambda(s(n))$$

we have that $s(n)\in B$

Using mathematical induction we have $B=\mathbb{N}_0$, so $\forall n\in\mathbb{N}_0$ we have $n\in B$ hence $\beta(n)=\lambda(n)$ proving that

$$\beta=\lambda$$

□

Up to now we have used the successor function $s:\mathbb{N}_0\rightarrow\mathbb{N}_0$ in the recursion and induction theorems. Once we have introduced the arithmetic of the natural numbers, we will rewrite these theorems by a version where $s(n)$ is replaced by $n+1$.

5.3 Arithmetic of the Natural numbers

We use recursion to define the sum of two natural numbers.

Definition 5.27. *Let $m,n\in\mathbb{N}_0$ then the addition operator $+$ is defined by*

$$+:\mathbb{N}_0\times\mathbb{N}_0\rightarrow\mathbb{N}_0 \text{ where } n+m\underset{notation}{=}+(n,m)=(s)^m(n)$$

Here $s:\mathbb{N}_0\rightarrow\mathbb{N}_0$ is the successor function [definition: 5.6] and we use the iteration principle from [theorem: 5.24] to define $(s)^n$.

Example 5.28. Using this definition we can easily calculate that $1+1=2$

Proof. $1+1=(s)^1(1)=(s\circ(s)^0)(s)=s((s)^0(1))=s(\mathrm{Id}_{\mathbb{N}_0}(1))=s(1)=2$ □

We will show now that $\langle\mathbb{N}_0,+\rangle$ forms a Abelian semi-group.

Theorem 5.29. (Neutral Element) *Let $n\in\mathbb{N}_0$ then $n+0=n=0+n$*

Proof.

1. $n+0=(s)^0(n)=\mathrm{Id}_{\mathbb{N}_0}(n)=n$
2. For the $0+n=n$ we use mathematical induction. So let $S=\{n\in\mathbb{N}_0|0+n=n\}$ then we have:

   $0\in S$. As $0+0\underset{(1)}{=}0$ proving $0\in S$

   $n\in S\Rightarrow s(n)\in S$. We have $0+s(n)=(s)^{s(n)}(0)=(s\circ(s)^n)(0)=s((s)^n(0))\underset{n\in S}{=}s(n)$ proving that $s(n)\in S$

   Using mathematical induction 5.11 we have $S=\mathbb{N}_0$. So if $n\in\mathbb{N}_0\Rightarrow n\in S$ then $0+n=n$. □

Theorem 5.30. $\forall n\in\mathbb{N}_0$ *we have $n+1=s(n)=1+n$*

Proof.

1. $n+1=(s)^1(n)=(s\circ(s)^0)(n)=s((s)^0(n))=s(\mathrm{Id}_{\mathbb{N}_0}(n))=s(n)$
2. For $1+n=s(n)$ we use induction, so define $S=\{n\in\mathbb{N}_0|1+n=s(n)\}$ then we have:

   $0\in S$. $1+0\underset{[\text{theorem: } 5.29]}{=}1=s(0)$

   $n\in S\Rightarrow n+1\in S$.

   $$1+s(n)=(s)^{s(n)}(1)=(s\circ(s)^n)(1)=s((s)^n(1))=s(1+n)\underset{n\in S}{=}s(s(n))$$

   proving that $s(n)\in S$.

   By mathematical induction [theorem: 5.11] we have $S=\mathbb{N}_0$ completing the proof. □

Lemma 5.31. *If $n,m\in\mathbb{N}$ 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 5.32. (Associativity) *If $n,m,k\in\mathbb{N}$ then $(n+m)+k=n+(m+k)$*

Proof. The proof is by mathematical induction, so given $n,m\in\mathbb{N}_0$ define

$$S_{n,m}=\{k\in\mathbb{N}|(n+m)+k=n+(m+k)\}$$

then we have:

$0\in S_{n,m}$. $(n+m)+0 \underset{\text{[theorem: 5.29]}}{=} n+m \underset{\text{[theorem: 5.29]}}{=} n+(m+0)\Rightarrow 0\in S_{n,m}$

$k\in S_{n,m}\Rightarrow s(k)\in S_{n,m}$. We have

$$\begin{aligned}(n+m)+s(k) &\underset{\text{[lemma: 5.31]}}{=} s((n+m)+k)\\ &\underset{k\in S}{=} s(n+(m+k))\\ &\underset{\text{[lemma: 5.31]}}{=} (n+s(m+k))\\ &\underset{\text{[lemma: 5.31]}}{=} (n+(m+s(k)))\end{aligned}$$

proving that $s(k)\in S_{n,m}$.

By mathematical induction [theorem: 5.11] we have $\mathbb{N}_0=S_{n,m}$. So if $n,m,k\in\mathbb{N}_0$ then $k\in S_{n,m}\Rightarrow(n+m)+k=n+(m+k)$ □

Theorem 5.33. (Commutativity) *If $n,m\in\mathbb{N}$ then $n+m=m+n$*

Proof. This is done again by induction. Let $n\in\mathbb{N}_0$ and define

$$S_n=\{k\in\mathbb{N}_0|n+k=k+n\}$$

then we have:

$0\in S_n$. Using [theorem: 5.29] it follows that $n+0=0+n$ proving that $0\in S_n$

$k\in S_n\Rightarrow s(k)\in S_n$. We have

$$\begin{aligned}n+s(k) &\underset{\text{[lemma: 5.31]}}{=} s(n+k)\\ &\underset{k\in S_{n,m}}{=} s(k+n)\\ &\underset{\text{[theorem: 5.30]}}{=} 1+(k+n)\\ &\underset{\text{[theorem: 5.32]}}{=} (1+k)+n\\ &\underset{\text{[theorem: 5.30]}}{=} s(k)+n\end{aligned}$$

Using mathematical induction [theorem: 5.11] we have that $S_n=\mathbb{N}_0$, So if $n,m\in\mathbb{N}\Rightarrow m\in S_n\Rightarrow n+m=m+n$. □

We can summarize the above theorems as follows.

Theorem 5.34. *$\langle\mathbb{N}_0,+\rangle$ forms a Abelian semi-group with neutral element 0*

Proof.

neutral element. This follows from [theorem: 5.29].

associativity. This follows from [theorem: 5.32].

commutativity. This follows from [theorem: 5.33] □

Next we use recursion to define multiplication in $\mathbb{N}_0$ and prove that $\langle\mathbb{N}_0,\cdot\rangle$ is a Abelian group.

Definition 5.35. (Multiplication) *Given $n\in\mathbb{N}_0$ define*

$$\alpha_n:\mathbb{N}_0\rightarrow\mathbb{N}_0 \text{ by } \alpha_n(m)=n+m$$

Then we define the multiplication operator as follows

$$\cdot:\mathbb{N}_0\times\mathbb{N}_0\rightarrow\mathbb{N}_0 \text{ by } n\cdot m=.(n,m)=(\alpha_n)^m(0)$$

Using the above definition we have

We have the following examples to see how multiplication works by repeating summation

$$\begin{array}{rcl}2\cdot 0 & = & (\alpha_2)^0(0)=\mathrm{Id}_{\mathbb{N}}(0)=0\\ 2\cdot 1 & = & (\alpha_2)^1(0)=(\alpha_2)^{s(0)}(0)=(\alpha_2\circ(\alpha_2)^0)(0)=\alpha_2(0)=2+0=2\\ 2\cdot 2 & = & (\alpha_2)^2(0)=(\alpha_2)^{s(1)}(0)=(\alpha_2((\alpha_2)^1(0)))=\alpha_2(2)=2+2=4\\ & \ldots. & \end{array}$$

Theorem 5.36. (Absorbing Element) *If* $n\in\mathbb{N}_0$ *then* $n\cdot 0=0=0\cdot n$

Proof.

1. $n\cdot 0=(\alpha_n)^0(0)=\mathrm{Id}_{\mathbb{N}_0}(0)=0$
2. We prove by induction that $0\cdot n=0$, so let $S=\{n\in\mathbb{N}_0|0\cdot n=0\}$ then we have:

   $\boldsymbol{0\in S}$. This follows from $0\cdot 0\underset{(1)}{=}0$

   $\boldsymbol{n\in S\Rightarrow s(n)\in S}$. We have

$$\begin{array}{rcl}0\cdot s(n) & = & (\alpha_0)^{s(n)}(0)\\ & = & (\alpha\circ(\alpha_0)^n)(0)\\ & = & \alpha_0((\alpha_0)^n(0))\\ & = & \alpha_0(0\cdot n)\\ & \underset{n\in S}{=} & \alpha_0(0)\\ & = & 0+0\\ & \underset{[\text{theorem: }5.29}{=} & 0\end{array}$$

   proving that $s(n)\in\mathcal{S}$.

   By induction [theorem: 5.11] we have that $S=\mathbb{N}_0$ hence the theorem follows. □

Theorem 5.37. (Neutral Element) *If* $n\in\mathbb{N}_0$ *then* $n\cdot 1=n=1\cdot n$

Proof.

1.

$$\begin{array}{rcl}n\cdot 1 & = & (\alpha_n)^1(0)\\ & = & (\alpha_n)^{s(0)}(0)\\ & = & (\alpha_n\circ(\alpha_n)^0)(0)\\ & = & \alpha_n((\alpha_n)^0(0))\\ & = & \alpha_n(\mathrm{Id}(0))\\ & = & \alpha_n(0)\\ & = & n+0\\ & \underset{[\text{theorem: }5.29]}{=} & n\end{array}$$

2. We prove $1\cdot n$ by induction, so let $S=\{n\in\mathbb{N}_0|1\cdot n=n\}$ then we have:

   $\boldsymbol{0\in S}$. This follows from $1\cdot 0\underset{[\text{theorem: }5.36]}{=}0$

   $\boldsymbol{n\in S\Rightarrow s(n)\in S}$. We have

$$\begin{array}{rcl}1\cdot s(n) & = & (\alpha_1)^{s(n)}(n)\\ & = & (\alpha_1\circ(\alpha_1)^n)(0)\\ & = & a_1((\alpha_1)^n(0))\\ & = & \alpha_1(1\cdot n)\\ & \underset{n\in S}{=} & \alpha_1(n)\\ & = & 1+n\\ & \underset{[\text{theorem: }5.30}{=} & s(n)\end{array}$$

proving that $s(n) \in S$.

By induction [theorem: 5.11] it follows that $S = \mathbb{N}_0$ completing the proof. $\square$

Lemma 5.38. *If $n, m \in \mathbb{N}_0$ then $n \cdot s(m) = n + n \cdot m \underset{\text{[theorem: 5.33]}}{=} n \cdot m + n$.*

Proof. $n \cdot s(m) = (\alpha_n)^{s(m)}(0) = (\alpha_n \circ (\alpha_n)^m)(0) = a_n((\alpha_n)^m(0)) = \alpha_n(n \cdot m) = n + n \cdot m$. $\square$

Theorem 5.39. (Distributivity) *$\forall n, m, k \in \mathbb{N}_0$ we have $(n+m) \cdot k = n \cdot k + m \cdot k$.*

Proof. We use induction to prove this. So given $n, m \in \mathbb{N}_0$ let

$$S_{n,m} = \{k \in \mathbb{N}_0 | (n+m) \cdot k = n \cdot k + m \cdot k\}$$

then we have:

$0 \in S_{n,m}$. $(n+m) \cdot 0 \underset{\text{[theorem: 5.36]}}{=} 0 \underset{\text{[theorem: 5.29]}}{=} 0 + 0 \underset{\text{[theorem: 5.36]}}{=} n \cdot 0 + m \cdot 0$

$n \in S_{n,m} \Rightarrow s(n) \in S_{n,m}$. We have

$$\begin{aligned} (n+m) \cdot s(k) & \underset{\text{[lemma: 5.38}}{=} (n+m) \cdot k + (n+m) \\ & \underset{k \in S_{n,m}}{=} (n \cdot k + m \cdot k) + (n+m) \\ & \underset{\text{[theorem: 5.32}}{=} n \cdot k + (m \cdot k + (n+m)) \\ & \underset{\text{[theorem: 5.33]}}{=} n \cdot k + (m \cdot k + (m+n)) \\ & \underset{\text{[theorem: 5.32}}{=} n \cdot k + ((m \cdot k + m) + n) \\ & \underset{\text{[theorem: 5.33]}}{=} n \cdot k + (n + (m \cdot k + m)) \\ & \underset{\text{[theorem: 5.32}}{=} (n \cdot k + n) + (m \cdot k + m) \\ & \underset{\text{[lemma: 5.38}}{=} n \cdot s(k) + m \cdot s(k) \end{aligned}$$

proving that $s(k) \in S_{n,m}$.

By induction [theorem: 5.11] it follows that $\mathbb{N}_0 = S_{n,m}$. So if $n, m, k \in \mathbb{N}_0$ then $k \in S_{n,m}$ giving $(n+m) \cdot k = n \cdot k + m \cdot k$. $\square$

Theorem 5.40. (Commutativity) *If $n, m \in \mathbb{N}_0$ then $n \cdot m = m \cdot n$.*

Proof. We prove this by induction so given $n \in \mathbb{N}_0$ let $S_n = \{m \in \mathbb{N}_0 | n \cdot m = m \cdot n\}$ then we have:

$0 \in S_n$. Using [theorem: 5.36] we have $n \cdot 0 = 0 = 0 \cdot n$ proving that $0 \in S_n$.

$m \in S_n \Rightarrow s(m) \in S_n$. We have

$$\begin{aligned} n \cdot s(m) & \underset{\text{[lemma: 5.38}}{=} n + n \cdot m \\ & \underset{m \in S_n}{=} n + m \cdot n \\ & \underset{\text{[theorem: 5.37]}}{=} 1 \cdot n + m \cdot n \\ & \underset{\text{[theorem 5.39]}}{=} (1+n) \cdot n \\ & \underset{\text{[theorem: 5.30}}{=} s(m) \cdot n \end{aligned}$$

proving that $s(m) \in S_n$.

Using induction [theorem: 5.11] it follows that $\mathbb{N}_0 = S_n$. So if $n, m \in \mathbb{N}_0$ then $m \in S_n$ hence $n \cdot m = m \cdot n$. $\square$

Theorem 5.41. (Associativity) *If $n, m, k \in \mathbb{N}_0$ then $(n \cdot m) \cdot k = n \cdot (m \cdot k)$*

Proof. We prove this by induction. So 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:

$0 \in S_{n,m}$. This follows from $(n \cdot m) \cdot 0 \underset{[\text{theorem: } 5.36]}{=} 0 \underset{[\text{theorem: } 5.36]}{=} n \cdot 0 = \underset{[\text{theorem: } 5.36]}{=} n \cdot (m \cdot 0)$

$k \in S_{n,m} \Rightarrow s(k) \in S_{n,m}$. We have

$$\begin{aligned} (n \cdot m) \cdot s(k) & \underset{[\text{theorem: } 5.38]}{=} (n \cdot m) \cdot k + n \cdot m \\ & \underset{k \in S_{n,m}}{=} n \cdot (m \cdot k) + n \cdot m \\ & \underset{[\text{theorem: } 5.40]}{=} (m \cdot k) \cdot n + m \cdot n \\ & \underset{[\text{theorem: } 5.39]}{=} ((m \cdot k) + m) \cdot n \\ & \underset{[\text{theorem: } 5.40]}{=} n \cdot ((m \cdot k) + m) \\ & \underset{[\text{theorem: } 5.38]}{=} n \cdot (m \cdot s(k)) \end{aligned}$$

proving that $s(k) \in S_{n,m}$.

Using induction we have then that $\mathbb{N}_0 = S_{n,m}$. So if $n, m, k \in \mathbb{N}_0$ we have $k \in S_{n,m}$ giving $(n \cdot m) \cdot k = n \cdot (m \cdot k)$. □

To summarize the above we have the following;

Theorem 5.42. *$\langle \mathbb{N}_0, \cdot \rangle$ is a Abelian semi-group with neutral element 1.*

Proof.

neutral element. This follows from [theorem: 5.37]

associativity. This follows from [theorem: 5.41]

commutativity. This follows from [theorem: 5.40] □

Although there is no inverse element for addition in $\mathbb{N}_0$ [this will be solved by the set of whole numbers], we can still solve equations as is expressed in the next theorem.

Theorem 5.43. *If $n, m, k \in \mathbb{N}_0$ then if $n + k = m + k$ it follows that $n = m$*

Proof. We prove this by induction. So given $n, m \in \mathbb{N}_0$ define $S = \{k \in \mathbb{N}_0 | \forall n, m \in \mathbb{N}_0 \text{ with } n + k = m + k \text{ we have n=m}\}$ then we have:

$0 \in S$. If $n, m \in \mathbb{N}_0$ are such that that $n + 0 = m + 0$ then we have $n \underset{5.29}{=} n + 0 = m + 0 \underset{5.29}{=} m$ or $n = m$ which proves that $0 \in S$

$k \in S \Rightarrow s(k) \in S$. If $n, m \in \mathbb{N}_0$ are such that $n + s(k) = m + s(k)$ then we have by [theorem: 5.30] that $n + (1 + k) = m + (1 + k)$ or using [theorem: 5.32] that $(n + 1) + k = (m + 1) + k$. As $k \in S$ it follows that $n + 1 = m + 1$ or using [theorem: 5.30] that $s(n) = s(m)$. Finally using [theorem: 5.16] we have $n = m$. So $s(k) \in S$.

Using induction we have then that $\mathbb{N}_0 = S$. So if $n, m, k \in \mathbb{N}_0$ then as $k \in S$ we have if $n + k = m + k$ that $n = m$. □

Note 5.44. We do not have a equivalent theorem for the product of two natural numbers, for example $0 \cdot 0 = 1 \cdot 0$ but we don't have that $1 = 0$.

5.4 Order relation on the natural numbers

Theorem 5.45. *If we define the relation $\leqslant$ by*

$$\leqslant = \{(n, m) \in \mathbb{N}_0 \times \mathbb{N}_0 | n \in m \vee n = m\}$$

then

$$\langle \mathbb{N}_0, \leqslant \rangle \text{ is a partial ordered set}$$

Proof.

reflectivity. If $n \in \mathbb{N}_0$ then $n = n \Rightarrow n \in n \vee n = n$ so that $n \leqslant n$.

anti-symmetry. If $n \leqslant m \wedge m \leqslant n$ then we have

$$\begin{aligned}(n \in m \vee n = m) \wedge (m \in n \vee m = n) &\Rightarrow (n \in m \vee n = m) \wedge (m \in n \vee n = m)\\ &\Rightarrow (n \in m \wedge m \in n) \vee n = m\\ &\underset{[\text{theorem: } 5.14]}{\Rightarrow} (n \subseteq m \wedge m \subseteq n) \vee n = m\\ &\Rightarrow n = m \vee n = m\\ &\Rightarrow n = m\end{aligned}$$

transitivity. If $n \leqslant m \wedge m \leqslant k$ then we have the following possibilities to consider

1. $n \in m \wedge m \in k$ then by [theorem: 5.14] $n \in m \wedge m \subseteq k \Rightarrow n \in k \Rightarrow n \leqslant k$
2. $n \in m \wedge m = k$ then $n \in k$ so that $n \leqslant k$
3. $n = m \wedge m \in k$ then $n \in k$ so that $n \leqslant k$
4. $n = m \wedge m = k$ then $n = k \Rightarrow n \leqslant k$

So in all cases we have $n \leqslant k$ proving transitivity. □

Theorem 5.46. *$\forall n \in \mathbb{N}_0$ we have $0 \leqslant n$*

Proof. We prove this by induction, so let $S = \{n \in \mathbb{N}_0 | 0 \leqslant n\}$ then we have:

$0 \in S$. $0 = 0$ so that $0 \leqslant 0$ proving that $0 \in S$.

$n \in S \Rightarrow s(n) \in S$. As $s(n) = n \bigcup \{n\}$ we have that $n \in s(n)$ so that $n \leqslant s(n)$, as $n \in S$ $0 \leqslant n$, so by transitivity we have that $0 \leqslant s(n)$. Hence we have $s(n) \in S$.

Using induction [theorem: 5.11] it follows that $\mathbb{N}_0 = S$ proving the theorem. □

Theorem 5.47. *$\forall n \in \mathbb{N}_0$ we have $n < s(n)$ [in other words using [theorem: 5.30] we have $n < n + 1$]*

Proof. From $n \in n \bigcup \{n\} = s(n)$ we have that $n \leqslant s(n)$ and by [theorem: 5.15] $n \neq s(n)$ so that $n < s(n)$. □

Theorem 5.48. *If $n \in \mathbb{N}_0$ then $k \in n \Leftrightarrow k < n$.*

Proof.

$\Rightarrow$. We proceed by induction, so let $S = \{n \in \mathbb{N}_0 | \text{If } k \in n \Rightarrow k < n\}$ then we have:

$0 \in S$. As $0 = \varnothing$ so that $k \in 0$ is never true hence $k \in n \Rightarrow k < n$ is true, proving that $0 \in S$.

$n \in S \Rightarrow s(n) \in S$. If $k \in s(n) = n \bigcup \{n\}$ then we have the following cases to consider:

$k \in n$. As $n \in S$ we have $k < n$, further from [theorem: 5.47] we have $n < s(n)$ so that $k < s(n)$.

$k = n$. By [theorem: 5.47] we have $n < s(n)$ so that $k < s(n)$.

So in all cases we have $k < s(n)$ proving that $s(n) \in S$.

By the induction [theorem: 5.11] it follows that $\mathbb{N}_0 = S$, proving the theorem.

$\Leftarrow$. If $k < n$ then $k \neq n$ and $k \leqslant n \Rightarrow k \in n \vee k = n$ so that $k \in n$. □

Theorem 5.49. *If $n, m \in \mathbb{N}_0$ then we have that*

1. *$n < 0$ is false.*
2. *If $n \leqslant 0$ then $n = 0$.*
3. *$n < m \wedge m < n$ is false.*
4. *$n \leqslant m \wedge m < n$ is false.*
5. *$n < m \wedge m \leqslant n$ is false.*

Proof.

1. If $n<0$ then by [theorem: 5.48] we have $n\in 0=\varnothing$ which is false.
2. If $n\leqslant 0$ then we have either $n<0$ [which by (1) is false] or $n=0$.
3. If $n<m\wedge m<n$ then $n\leqslant m\wedge m\leqslant n\Rightarrow n=m$ and $n\neq m$ which is a contradiction.
4. If $n\leqslant m\wedge m<n$ then $n\leqslant m\wedge m\leqslant n\Rightarrow n=m$ and $n\neq m$ which is a contradiction.
5. If $n<m\wedge m\leqslant n$ then $n\leqslant m\wedge m\leqslant n\Rightarrow n=m$ and $n\neq m$ which is a contradiction. □

Theorem 5.50. *$\forall n,m\in\mathbb{N}_0$ with $n<m$ we have $s(n)\leqslant m$ [in other words using [theorem: 5.30] $n<m$ implies $n+1\leqslant m$].*

Proof. We proof this by induction, so given $n\in\mathbb{N}_0$, define $S_n=\{m\in\mathbb{N}_0|n<m\Rightarrow s(n)\leqslant m\}$ then we have:

$0\in S_n$. By [theorem: 5.49] $n<0$ is false, so $n<0\Rightarrow s(n)\leqslant m$ is true, proving that $0\in S_n$.

$m\in S_n\rightarrow s(m)\in S_n$. Let $n<s(m)$ then we have $n\neq s(m)$ and $n\leqslant s(m)$ so that $n\in s(m)=m\bigcup\{m\}$, hence we have to look at:

$n\in m$. By [theorem: 5.48] we have $n<m$, as $m\in S_n$ we have $s(n)\leqslant m$, as by [theorem: 5.47] $m<s(m)$ it follows by transitivity that $s(n)\leqslant s(m)$ [actually even $s(n)<s(m)$].

$n=m$. Then $s(n)=s(m)$ so that $s(n)\leqslant s(m)$.

So we have $s(m)\in S_n$

Using induction [theorem: 5.11] it follows that $\forall n,m\in\mathbb{N}_0$ with $n<m$ we have as $m\in S_n$ such that $s(n)\leqslant m$. □

Theorem 5.51. *$\langle\mathbb{N}_0,\leqslant\rangle$ is a well ordered set.*

Proof. We prove this by contradiction. Assume that there exist a A such that $\varnothing\neq A\subseteq\mathbb{N}_0$ with no 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 that $S_A\bigcap A=\varnothing$ [for if $l\in S_A\bigcap A$ then $l\in A$ and $\forall m\in A$ we have $l\leqslant m$ so that l is a least element of A]. For S_A we have

$0\in S_A$. If $m\in A$ we have by [theorem: 5.46] that $0\leqslant m$ so that $0\in S_A$.

$n\in S_A\Rightarrow s(n)\in S_A$. As $n\in S_A$ we have $\forall m\in A$ that $n\leqslant m$, $S_A\bigcap A=\varnothing$ so we have $n\neq m$ so that $n<m$, using then [theorem: 5.50] proves $s(n)\leqslant m$. Hence $s(n)\in S_A$

Using mathematical induction we have $S_A=\mathbb{N}_0$, so that $S_A\bigcap A=\mathbb{N}_0\bigcap A=A\neq\varnothing$ contradicting $S_A\bigcap A=\varnothing$. As the assumption gives a contradiction every non empty subset of $\mathbb{N}_0$ has a least element and $\langle\mathbb{N}_0,\leqslant\rangle$ must be well ordered. □

As a consequence of the above we have:

Corollary 5.52. *$\langle\mathbb{N}_0,\leqslant\rangle$ is totally ordered and conditional complete.*

Proof. As $\langle\mathbb{N}_0,\leqslant\rangle$ is well ordered by [theorem: 5.51] we have by [theorem: 3.84] that $\langle\mathbb{N}_0,\leqslant\rangle$ is totally ordered and conditional complete. □

Corollary 5.53. *If $x,y\in\mathbb{N}_0$ then we have either $x\leqslant y$ or $y<x$*

Proof. As $\langle\mathbb{N}_0,\leqslant\rangle$ is well ordered the corollary follows from [theorem: 3.84]. □

Theorem 5.54. *$\forall n,m\in\mathbb{N}$ then $n<m\Leftrightarrow s(n)<s(m)$*

Proof.

$\Rightarrow$. From [theorem: 5.50] we have $s(n)\leqslant m$, as by [theorem: 5.47] $m<s(m)$ it follows that $s(n)<s(m)$.

$\Leftarrow$. Assume that $m\leqslant n$ then by [theorem: 5.47] we have $n<s(n)$ so that $n<s(m)$, using [theorem: 5.50] we have $s(n)\leqslant s(m)$, combining this with $s(n)<s(m)\Rightarrow s(n)\neq s(m)\wedge s(n)\leqslant s(m)$ gives the contradiction $s(n)=s(m)\wedge s(n)\neq s(m)$, so we have

$$\neg(m\leqslant n)$$

Using [corollary: 5.53] we have $m\leqslant n$ or $n<m$ so that we must have

$$n<m$$

□

Theorem 5.55. *If $n,m,k\in\mathbb{N}_0$ then we have*

$$n<m\Leftrightarrow n+k<m+k$$

which, using [theorem: 5.43], implies that

$$n\leqslant m\Leftrightarrow n+k\leqslant m+k$$

Proof. We use induction , so let $S=\{k\in\mathbb{N}_0|\text{If } m,n\in\mathbb{N}_0 \text{ then } n<m\Leftrightarrow n+k<m+k\}$ then we have:

$0\in S$. If $k=0$ then for $n,m\in\mathbb{N}_0$ we have, as by [theorem: 5.29] $n=n+0\wedge m=m+0$ that $n<m\Leftrightarrow n+0<m+0$. So $0\in S$.

$k\in S\Rightarrow s(k)\in S$. then we have

$$\begin{array}{rcl} n<m & \underset{k\in S}{\Leftrightarrow} & n+k<m+k \\ & \underset{\text{[theorem: 5.50]}}{\Leftrightarrow} & s(n+k)<s(m+k) \\ & \underset{\text{[theorem: 5.31]}}{\Leftrightarrow} & n+s(k)<m+s(k) \end{array}$$

proving that $s(k)\in S$

Induction [theorem: 5.11] proves then $\mathbb{N}_0=S$ completing the proof. □

Corollary 5.56. *If $n\in\mathbb{N}_0$ then we have:*

1. *If $k\in\mathbb{N}_0\setminus\{0\}$ then $n<n+k$*
2. *If $k\in\mathbb{N}_0$ then $n\leqslant n+k$*

Proof.

1. If $k\neq 0$ then $0<k$ so that by the above theorem [theorem: 5.55] we have

$$n\underset{\text{[theorem: 5.29]}}{=}0+n<n+k$$

2. As $0\leqslant 0$ it follows from the above theorem [theorem: 5.55] we have that

$$n\underset{\text{[theorem: 5.29]}}{=}0+n\leqslant n+k$$

□

Theorem 5.57. *If $n,k\in\mathbb{N}_0$ then $n+k=0$ implies $n=k=0$.*

Proof. Suppose that $k\neq 0$ then as $0\leqslant n\underset{\text{[theorem: 5.56]}}{\Rightarrow}0\leqslant n<n+k=0$ so that $0\neq 0$ a contradiction, so $k=0$. But then $n=n+0=n+k=0$. □

Theorem 5.58. *If $n,m\in\mathbb{N}_0$ with $n<s(m)$ then $n\leqslant m$.*

Note 5.59. As by [theorem: 5.30] $s(m)=m+1$ this is equivalent with $n<m+1\Rightarrow n\leqslant m$

Proof. Using [corollary: 5.53] we have that either $n\leqslant m$ or $m<n$. If $m<n$ then by [theorem: 5.50] $s(m)\leqslant n$, which combined with the hypothesis $n<s(m)$ gives the contradiction $n<m$. Hence we must have $n\leqslant m$. □

Theorem 5.60. *If $n,m\in\mathbb{N}_0$ with $n<m$ then $\exists!k\in\mathbb{N}_0\setminus\{0\}$ such that $m=n+k$.*

Proof. First we prove existence by induction, so let

$$S_n=\{m\in\mathbb{N}_0|\text{If } n<m \text{ then there exist a } k\in\mathbb{N}_0 \text{ such that } k\neq 0 \text{ and } m=n+k\}$$

then we have:

$0\in S_n$. As $n<0$ is false by [theorem: 5.49], the condition is satisfied vacuously, proving that $0\in S_n$.

$m\in S_n\Rightarrow s(m)\in S_n$. If $n<s(m)$ then we have by [theorem: 5.58] that $n\leqslant m$ so that we have the following possibilities to consider:

$n=m$. Then $n+1\underset{\text{[theorem: 5.30]}}{=}s(n)=s(m)$, as $1=s(0)\neq 0$ we have if we take $k=1$ that $k\neq 0$ and $n+k=s(m)$, proving that $s(m)\in S_n$

$n<m$. Then as $m\in S_n$ there exist a $l\in\mathbb{N}_0$ such that $l\neq 0$ and $n+l=m$. Now

$$s(m)=s(n+l)\underset{\text{[theorem: 5.31]}}{=}n+s(l)$$

Take $k = s(l)$ then $n + k = s(m)$, further by [theorems: 5.46, 5.47] we have $0 \leqslant l \wedge l < s(l) = k$ so that $0 < k$ hence $k \neq 0$. This proves that in this case we also have $s(m) \in S_n$.

Induction [see theorem: 5.11] proves then that $\mathbb{N}_0 = S_n$. Hence if $n, m \in \mathbb{N}_0$ we have $m \in S_n$ so that if $n < m$ there exist a $k \in \mathbb{N}_0$ such that $k \neq 0$ and $m = n + k$.
Now for uniqueness assume that $n < m$ and there exists $k, l \in \mathbb{N}_0$ such that

$$k + n \underset{[\text{theorem: } 5.33]}{=} n + k = m = n + l \underset{[\text{theorem: } 5.33]}{=} l + n$$

then by [theorem: 5.43] $k = l$. □

Corollary 5.61. *If $n, m \in \mathbb{N}_0$ then $n < m \Leftrightarrow \exists! k \in \mathbb{N}_0 \setminus \{0\}$ such that $n + k = m$*

Proof.

⇒. This follows from the previous theorem [theorem: 5.60].

⇐. Let $k \in \mathbb{N}_0 \setminus \{0\}$ such that $n + k = m$. As $k \in \mathbb{N}_0 \setminus \{0\}$ we have $0 < k$ so that by [theorem: 5.55] $0 + n < k + n \underset{[\text{theorems: } 5.29, 5.33]}{\Rightarrow} n < n + k = m$. □

Corollary 5.62. *If $n, m \in \mathbb{N}_0$ then $n \leqslant m \Leftrightarrow \exists! k \in \mathbb{N}_0$ such that $m = n + k$*

Proof.

⇒. If $n \leqslant m$ then we have either:

$\boldsymbol{n = m}$**.** Then $m \underset{[\text{theorem: } 5.29]}{=} n + 0$ where $0 \in \mathbb{N}_0$.

$\boldsymbol{n < m}$**.** Then by the previous corollary [corollary: 5.61] there exists a $k \in \mathbb{N}_0 \setminus \{0\} \subseteq \mathbb{N}_0$ such that $m = n + k$.

proving existence. For uniqueness assume that $n + k = m = n + l$ then

$$k + n \underset{[\text{theorem: } 5.33]}{=} n + k = m = n + l \underset{[\text{theorem: } 5.33]}{=} l + n$$

proving by [theorem: 5.43] that $k = l$.

⇐. As $k \in \mathbb{N}_0$ we have either:

$\boldsymbol{k = 0}$**.** Then $m = n + 0 \underset{[\text{theorem: } 5.29]}{=} n$ so that $n \leqslant m$.

$\boldsymbol{0 < k}$**.** Then by the previous corollary [corollary: 5.61] we have $n < m$ so that $n \leqslant m$. □

The above corollary ensures that the following definition is well defined.

Definition 5.63. *If $n, m \in \mathbb{N}_0$ with $n \leqslant m$ then the **unique** $k \in \mathbb{N}_0$ such that $m = n + k$ is noted as $m - n$. So we have that $n + (m - n) \underset{[\text{theorem: } 5.33]}{=} (m - n) + n = m$ and using [theorem: 5.29] that $n - n = 0$.*

Note 5.64. The condition $n \leqslant m$ is essential for the existence of $n - m$ as this is needed for [corollary: 5.62]. Later when we define the set $\mathbb{Z}$ of integers we will relax this condition.

Theorem 5.65. *If $n, m, k \in \mathbb{N}_0$ is such that $n \leqslant k$ then*

$$(k + m) - n = (k - n) + m = (m + k) - n$$

Proof. As $n \leqslant k$ we have by [theorem: 5.56] $n \leqslant k + m$ so that $(k + m) - n$ and $k - n$ are well defined. Now

$$\begin{aligned} ((k - n) + m) + n & \underset{[\text{theorem: } 5.32]}{=} (k - n) + (m + n) \\ & \underset{[\text{theorem: } 5.33}{=} (k - n) + (n + m) \\ & \underset{[\text{theorem: } 5.33]}{=} ((k - n) + n) + m \\ & \underset{\text{defiition}}{=} k + m \end{aligned}$$

So we have that

$$(k + m) - n = (k - n) + m$$

Further using commutativity [theorem: 5.33] we have that $(m+k)-n=(k+m)-n$ so that

$$(m+k)-n=(k-n)+m$$

□

Theorem 5.66. *If $n,k\in\mathbb{N}_0$ then $(n+k)-n=k=(k+n)-n$*

Proof. As $n\leqslant n$ we can us the previous theorem [see theorem: 5.65] so that

$$(k+n)-n=(n+k)-n=(n-n)+k=0+k=k$$

□

Theorem 5.67. *Let $n\in\mathbb{N}_0$ then we have Let $n,m\in\mathbb{N}_0$ such that $n<m$ then $n\leqslant m-1$*

Proof. As $n<m$ we have by [theorem: 5.60] a $k\in\mathbb{N}_0\backslash\{0\}$ such that $m=n+k$. As $0\neq k$ we have by [theorem: 5.18] that there exist a $l\in\mathbb{N}_0$ such that $k=s(l)=l+1$, so $m=(n+l)+1$ which by [definition 5.63] means that $n+l=m-1$. Further by [theorem: 5.56] we have $n\leqslant n+l$ so that $n\leqslant m-1$. □

Corollary 5.68. *Let $n\in\mathbb{N}_0$ and $m\in\mathbb{N}=\mathbb{N}_0\backslash\{0\}$ then $n<m\Leftrightarrow n\leqslant m-1$*

Proof.

$\Rightarrow$. This follows from the previous theorem [theorem: 5.67]

$\Leftarrow$. By [theorem: 5.47] we have $(m-1)<(m-1)+1=m$ we have from $n\leqslant m-1$ that $n<m$. □

Theorem 5.69. *Let $n\in\mathbb{N}_0$ and $m\in\mathbb{N}_0\backslash\{0\}$ then $(m-1)\cdot n=n\cdot(m-1)=n\cdot m-n$*

Proof. As $0<m$ we have by [theorem: 5.50] that $1=s(0)\leqslant m$ so that $m-1$ is well defined. Now

$$n+(m-1)\cdot n\underset{\text{commutativity}}{=}(m-1)\cdot n+n=(m-1)\cdot n+1\cdot n=((m-1)+1)\cdot n=m\cdot n=n.m$$

so that $(m-1)\cdot n=n\cdot m-n$ and by commutativity [see theorem: 5.33] $n\cdot(m-1)=n\cdot m-n$ □

Theorem 5.70. *If $n,m,i\in\mathbb{N}_0$ then*

1. *If $n\leqslant i<m$ then $0\leqslant i-n<m-n$*
2. *If $n\leqslant i\leqslant m$ then $0\leqslant i-n\leqslant m-n$*

Proof.

1. As $n\leqslant i<m$ we have $n<m$. From [corollary: 5.53] it follows that $0\leqslant i-n\vee i-n<0$ and $i-n<m-n\vee m-n\leqslant i-n$. Now by [theorem: 5.49] we have that $i-n<0$ is false so we must have that $0\leqslant i-n$. If $m-n\leqslant i-n$ then by [theorem: 5.55] $m=(m-n)+n\leqslant(i-n)+n=n$ proving that $m\leqslant n$ which by [theorem: 5.49] contradicts with $n<m$, so we must have $i-n<m-n$.
2. As $n\leqslant i\leqslant m$ we have $n\leqslant m$. From [corollary: 5.53] it follows that $0\leqslant i-n\vee i-n<0$ and $i-n\leqslant m-n\vee m-n<i-n$. Now by [theorem: 5.49] we have that $i-n<0$ is false so we must have that $0\leqslant i-n$. If $m-n<i-n$ then by [theorem: 5.55] $m=(m-n)+n<(i-n)+n=n$ proving that $m<n$ which by [theorem: 5.49] contradicts with $n\leqslant m$, so we must have $i-n\leqslant m-n$. □

Theorem 5.71. *If $k,n,m\in\mathbb{N}_0$ such that $k\leqslant n\wedge k\leqslant m$ then we have*

$$n\leqslant m\Leftrightarrow n-k\leqslant m-k$$

Proof.

$\Rightarrow$. Using [theorem: 5.53] we have either $m-k<n-k$ or $n-k\leqslant m-k$, if $m-k<n-k$ we have by [theorem: 5.55] that $(m-k)+k<(n-k)+k$ so that $m<n$ which as $n\leqslant m$ gives the contradiction $m<m$, so we have $n-k\leqslant m-k$.

$\Leftarrow$. Using [theorem: 5.55] we have that $(n-k)+k\leqslant(m-k)+k$ so that $n\leqslant m$. □

Theorem 5.72. *If $n\in\mathbb{N}_0$ then there does not exist a $k\in\mathbb{N}_0$ such that $n<k<s(n)$*

Proof. Assume that $\exists k\in\mathbb{N}_0$ such that $n<k<s(n)$. As $n<k$ we have by [theorem: 5.50] that $s(n)\leqslant k$ which combined with $k<s(n)$ gives $s(n)<s(n)$ a contradiction. □

Theorem 5.73. *If* $\varnothing \neq A \subseteq \mathbb{N}_0$ *is a set such that* $\sup(A)$ *exist then* $\sup(A) \in A$

Proof. We have the following cases for $\sup(A)$ to consider:

$\boldsymbol{\sup(A) = 0.}$ As $A \neq \varnothing$ there exist a $x \in A$, further as the $\sup(A)$ is a upper bound of A we have that $x \leqslant 0$, which by [theorem: 5.49] proves that $x = 0 = \sup(A)$, giving that $\sup(A) = x \in A$.

$\boldsymbol{\sup(A) \neq 0.}$ Using [theorem: 5.18] there exist a $k \in \mathbb{N}_0$ such that $s(k) = \sup(A)$. As $\langle \mathbb{N}_0, \leqslant \rangle$ is totally ordered [see theorem: 5.52] and $k < s(k) = \sup(A)$, it follows from the properties of the supremum [theorem: 3.71] that there exist a $a \in A$ such that $k < a \leqslant \sup(A) = s(k)$. As we can not have $k < a < s(k)$ [see theorem: 5.72], it follows that $a = \sup(A)$ so that $\sup(A) \in A$. □

Theorem 5.74. *If* $n, m, r, s \in \mathbb{N}_0$ *then*

1. *If* $n < m \wedge r < s$ *then* $n + r < m + s$
2. *If* $n \leqslant m \wedge r \leqslant s$ *then* $n + r \leqslant m + r$
3. *If* $n < m \wedge r \leqslant s$ *then* $n + r < m + r$
4. *If* $n \leqslant m \wedge r < s$ *then* $n + m < m + r$

Proof.

1. Using [theorem: 5.55] to follows that $n + r < m + r$ and $r + m < s + m \underset{[\text{theorem: } 5.33]}{\Rightarrow} n + r < m + s$ proving, using transitivity, that $n + r < m + 1$.
2. Using [theorem: 5.55] to follows that $n + r \leqslant m + r$ and $r + m \leqslant s + m \underset{[\text{theorem: } 5.33]}{\Rightarrow} n + r \leqslant m + s$ proving, using transitivity, that $n + r < m + 1$.
3. Using [theorem: 5.55] to follows that $n + r \leqslant m + r$ and $r + m < s + m \underset{[\text{theorem: } 5.33]}{\Rightarrow} n + r < m + s$ proving, using transitivity, that $n + r < m + 1$.
4. Using [theorem: 5.55] to follows that $n + r < m + r$ and $r + m \leqslant s + m \underset{[\text{theorem: } 5.33]}{\Rightarrow} n + r < m + s$ proving, using transitivity, that $n + r < m + 1$. □

Theorem 5.75. *Let* $n, m \in \mathbb{N}_0 \setminus \{0\}$ *then* $n \cdot m \in \mathbb{N}_0 \setminus \{0\}$.

Proof. As $m \neq 0$ it follows from [theorem: 5.18] that $\exists k \in \mathbb{N}_0$ such that $m = s(k)$. So $n \cdot m = n \cdot s(k) \underset{[\text{theorem: } 5.38]}{=} n + n \cdot k$. Further as $n \neq 0$ we have that $0 < n$, so that by [theorem: 5.55] $n \underset{[\text{theorem: } 5.29]}{=} n + 0 \leqslant n + n \cdot k = n \cdot m$, using transitivity gives then finally $0 < n \cdot m$. □

Theorem 5.76. *If* $n, m \in \mathbb{N}_0$ *such that* $n < m$ *then*

1. *If* $k \in \mathbb{N}_0 \setminus \{0\}$ *then* $n \cdot k < m \cdot k$
2. *If* $k \in \mathbb{N}_0$ *then* $n \cdot k \leqslant m \cdot k$

Proof.

1. As $n < m$ we have by [theorem: 5.60] that there exist a $l \in \mathbb{N}_0 \setminus \{0\}$ such that $m = n + l$. So

$$m \cdot k = (n + l) \cdot k \underset{[\text{theorem: } 5.39]}{=} n \cdot k + l \cdot k.$$

As $l, k \in \mathbb{N}_0 \setminus \{0\}$ we have by [theorem: 5.75] that $l \cdot k \neq 0$ so that $0 < l \cdot k$, hence using [theorem: 5.55] we have that

$$n \cdot k \underset{[\text{theorem: } 5.29]}{=} 0 + n \cdot k < l \cdot k + n \cdot k \underset{[\text{theorem: } 5.33]}{=} n \cdot k + l \cdot k = m \cdot k$$

so that

$$n \cdot k < m \cdot k$$

2. If $k \in \mathbb{N}_0$ then we have either:

$\boldsymbol{k = 0.}$ Then by [theorem: 5.36] we have $n \cdot k = 0 = m \cdot k$ so that $n \cdot k \leqslant m \cdot l$.

$\boldsymbol{k \neq 0.}$ Then by (1) $n \cdot k < m \cdot k \Rightarrow n \cdot k \leqslant m \cdot k$. □

Theorem 5.77. *If* $n, m \in \mathbb{N}_0$ *such that* $\exists k \in \mathbb{N}_0 \setminus \{n\}$ *such that* $n \cdot k = m \cdot k$ *then* $n = m$.

Proof. Using [corollary: 5.53] we have that $n<m$, $m<n$ or $n=m$. If $n<m$ then by [theorem: 5.76] $n\cdot k<m\cdot k$ contradicting $n\cdot k=m\cdot k$, likewise if $m<n$ then by [theorem: 5.76] $m\cdot k<n\cdot k$ contradicting $n\cdot k=m\cdot k$. So we must have $n=m$. □

Theorem 5.78. (Archimedean Property) *If $x,y\in\mathbb{N}_0$ and $x\neq 0$ then there exists a $z\in\mathbb{N}_0\backslash\{0\}$ such that $y<z\cdot x$*

Proof. For y we have two possibilities:

- **$y=0$.** As $x\neq 0$ we have $y=0<x\underset{[\text{theorem: } 5.37]}{=}1\cdot x$, so using $z=1$ proves the theorem.
- **$y\neq 0$.** Using [corollary: 5.53] we have for $x,y\in\mathbb{N}_0$ either:
  - **$y\leqslant x$.** Then as $1<s(1)=2$ [see theorem: 5.47] we have $x\underset{[\text{theorem: } 5.37]}{=}1\cdot x<2\cdot x$ [see: theorem: 5.76], hence $y<2\cdot x$, so using $z=2$ proves the theorem.
  - **$x<y$.** Using [theorem: 5.60] there exist $k\in\mathbb{N}_0\backslash\{0\}$ such that

    $$y=x+k \tag{5.14}$$

    As $0<x$ we have by [theorem: 5.50] $1=s(0)\leqslant x$ so that by multiplication with k we have [see theorem: 5.76] that

    $$k=1\cdot k\leqslant x\cdot k \tag{5.15}$$

    As $0\neq k<s(k)$ and $x\neq 0$ we have by [see theorem: 5.76] that $k\cdot x<s(k)\cdot x\Rightarrow x\cdot k<x\cdot s(k)$ combining this with [eq: 5.15] gives that

    $$k<x\cdot s(k) \tag{5.16}$$

    Using [theorem: 5.55] we have

    $$x+k=k+x<\cdot s(k)+x=x+x\cdot s(k)=x\cdot 1+x\cdot s(k)\underset{\text{distributivity}}{=}x\cdot(1+s(k))$$

    or using [eq: 5.14] that $y<x\cdot(s+s(k))$. So if we take $z=1+s(k)$ we have that $y<x\cdot z$ which as also $0<1<1+s(k)$ proves the theorem. □

Theorem 5.79. (Division) *If $m\in\mathbb{N}_0$ and $n\in\mathbb{N}_0\backslash\{0\}$ then there exists a **unique** $r\in\mathbb{N}_0$ and a unique $q\in\mathbb{N}_0$ such that*

$$m=n\cdot q+r \text{ and } 0\leqslant r<n$$

Proof. First we prove existence of q and r. As $n\in\mathbb{N}_0\backslash\{0\}$ $n\neq 0$ so that $0<n$. For m we have the following cases to consider:

- **$m=0$.** In this case taking $q=0$ and $r=0$ gives $n\cdot 0+0\underset{[\text{theorem: } 5.36]}{=}0+0\underset{[\text{theorem: } 5.29]}{=}0\cdot m+0$ and $0\leqslant 0<n$, so $q=0=r$ satisfies $m=n\cdot q+r$ and $0\leqslant r<n$.
- **$0<m$.** Then we have the following cases for n to consider:
  - **$n=1$.** Take $q=m$ and $r=0$ then $n\cdot q+r=1\cdot m+0\underset{[\text{theorem: } 5.29, 5.37]}{=}m$ and $0\leqslant 0<n$, so q,r satisfies $m=n\cdot q+r$ and $0\leqslant r<n$.
  - **$n\neq 1$.** Then as $0<n\underset{[\text{theorem: } 5.50]}{\Rightarrow}1=s(0)\leqslant n$ we have $1<n$. By [theorem: 5.76] it follows that $m=1\cdot m<n\cdot m$, so if we define

    $$A_{n,m}=\{x\in\mathbb{N}_0|m<n\cdot x\wedge x\leqslant m\}$$

    then $m\in A_{n,m}$ proving that

    $$A_{n,m}\neq\varnothing$$

    As $\langle\mathbb{N}_0,\leqslant\rangle$ is well ordered [see theorem: 5.51] there exist a least element

    $$q'=\min(A_{n,m})$$

    If $q'=0$ then as $q'\in A_{n,m}$ we would have $m<n\cdot 0\underset{[\text{theorem: } 5.36]}{=}0$ a contradiction, hence we must have that $0<q'$. So by [theorem: 5.18] there exist a $q\in\mathbb{N}_0$ such that $s(q)=q'$. As $q<s(q)=q'$ [see theorem: 5.47] we must have that $q\notin A_{n,m}$, which, as $q<q'\leqslant m$, means that $n\cdot q\leqslant m$. From this we have by [theorem: 5.76] the existence of a $r\in\mathbb{N}_0$ such that

    $$m=n\cdot q+r$$

    Using [corollary: 5.53] w have either $n\leqslant r$ or $r<n$. If $n\leqslant r$ then by [theorem: 5.55] we have $n+n\cdot q\leqslant r+n\cdot q=n\cdot q+r=m$, hence

    $$n\cdot q'=n\cdot s(q)\underset{[\text{theorem: } 5.38}{=}n+n\cdot q\leqslant m$$

As $q' \in A_{n,m}$ we have by definition that $m < n \cdot q'$ which combined with the above yields the contradiction $m < m$. So we must have

$$0 \leqslant r < n$$

To summarize we have found q, r such that $m = n \cdot q + r$ and $0 \leqslant r < n$ proving existence.

Now to prove uniqueness. Assume that $n \cdot q + r = m = n \cdot q'' + r''$ and $0 \leqslant q < n$, $0 \leqslant q'' < n$ with $q \neq q''$ then by [corollary: 5.53] we have either $q < q''$, $q'' < q$ or $q = q'$. For the cases $q < q''$ or $q'' < q$ we have

$q < q''$. Then by [theorem: 5.50]

$$\begin{aligned}
s(q) \leqslant q'' &\underset{\text{[theorem: 5.76}}{\Rightarrow} s(q) \cdot n \leqslant q'' \cdot n \\
&\underset{\text{[theorem 5.40]}}{\Rightarrow} n \cdot s(q) \leqslant q'' \cdot n \\
&\underset{\text{[theorem: 5.38]}}{\Rightarrow} n \cdot q + n \leqslant q'' \cdot n \\
&\underset{\text{[theorem: 5.55]}}{\Rightarrow} n \cdot q + n + r + r'' \leqslant q'' \cdot n + r + r'' \\
&\Rightarrow m + n + r'' \leqslant m + r \\
&\underset{\text{[theorem: 5.55]}}{\Rightarrow} n + r'' \leqslant r \\
&\underset{\text{[theorem: 5.56]}}{\Rightarrow} n \leqslant n + r'' \leqslant r
\end{aligned}$$

contradicting $r < n$.

$q'' < q$. Then by [theorem: 5.50]

$$\begin{aligned}
s(q'') \leqslant q &\underset{\text{[theorem: 5.76}}{\Rightarrow} s(q'') \cdot n \leqslant q \cdot n \\
&\underset{\text{[theorem 5.40]}}{\Rightarrow} n \cdot s(q'') \leqslant q \cdot n \\
&\underset{\text{[theorem: 5.38]}}{\Rightarrow} n \cdot q'' + n \leqslant q \cdot n \\
&\underset{\text{[theorem: 5.55]}}{\Rightarrow} n \cdot q'' + n + r + r'' \leqslant q \cdot n + r + r'' \\
&\Rightarrow m + n + r \leqslant m + r'' \\
&\underset{\text{[theorem: 5.55]}}{\Rightarrow} n + r \leqslant r'' \\
&\underset{\text{[theorem: 5.56]}}{\Rightarrow} n \leqslant n + r \leqslant r''
\end{aligned}$$

contradicting $r < n$.

So we must have that $q = q''$ but then $r + n \cdot q = n \cdot q + r = m = n \cdot q + r'' = r'' + n \cdot q$ proving by [theorem: 5.55] that $r = r''$. □

5.5 Other forms of Mathematical Induction and Recursion

In this section we rewrite the theorem of induction and recursion using $n + 1$ instead of $s(n)$ [see theorem: 5.30]. First we introduce some definitions.

Definition 5.80. *Let $n \in \mathbb{N}_0$ then $\{n, \ldots \infty\}$ is defined as*

$$\{n, \ldots \infty\} = \{i \in \mathbb{N}_0 | n \leqslant i\}$$

Note 5.81. $\{0, \ldots, \infty\} = \{x \in \mathbb{N}_0 | 0 \leqslant x\} \underset{\text{[theorem: 5.46]}}{=} \mathbb{N}_0$

Definition 5.82. *Let $n, m \in \mathbb{N}_0$ then $\{n, \ldots, m\}$ is defined as*

$$\{n, \ldots, m\} = \{i \in \mathbb{N}_0 | n \leqslant i \wedge i \leqslant m\}$$

We have now the following variation on mathematical induction.

Theorem 5.83. (Mathematical Induction) *If $n \in \mathbb{N}_0$ and $X \subseteq \{n, \ldots, \infty\}$ is such that*

1. $n \in X$

2. *If* $i \in X$ *then* $i+1 \in X$

then $X = \{n, \ldots, \infty\}$.

Proof. Take $S = \{i \in \mathbb{N}_0 | i+n \in X\}$ then we have:

$0 \in S$. As $0+n \underset{[\text{theorem: } 5.29]}{=} n \in X$ we have $0 \in S$.

$i \in S \Rightarrow s(i) \in S$. As $i \in S$ we have $i+n \in X$ so that by the hypothesis $(i+n)+1 \in X$. Now

$$\begin{aligned}(i+n)+1 &\underset{[\text{theorem: } 5.32]}{=} i+(n+1)\\ &\underset{[\text{theorem: } 5.33]}{=} i+(1+n)\\ &\underset{[\text{theorem: } 5.32]}{=} (i+1)+n\\ &\underset{[\text{theorem: } 5.30}{=} s(i)+n\end{aligned}$$

so that $s(i)+n \in X$, proving $s(i) \in S$.

By mathematical induction we have that $S = \mathbb{N}_0$. If $i \in \{n, \ldots \infty\}$ then $n \leqslant i$ so by [theorem: 5.62] $\exists k \in \mathbb{N}_0$ such that $i = n+k \underset{[\text{theorem: } 5.33]}{=} k+n \underset{k \in \mathbb{N}_0 = S}{\Rightarrow} i \in X$. Hence $\{n, \ldots, \infty\} \subseteq X$ which together with $X \subseteq \{1, \ldots, n\}$ proves that

$$X = \{1, \ldots, \infty\}$$

□

For recursion we have the following theorems that follows from [theorem: 5.20], [theorem: 5.24] and [theorem: 5.26] by replacing $s(n)$ by its equivalent form $n+1$.

Theorem 5.84. *Let A be a set, $a \in A$ and $f: A \rightarrow A$ a function then there exist a **unique** function*

$$\lambda: \mathbb{N}_0 \rightarrow A$$

such that:

1. $\lambda(0) = a$
2. $\forall n \in \mathbb{N}_0$ *we have* $\lambda(n+1) = f(\lambda(n))$

Further if $f: A \rightarrow A$ is injective and $a \notin f(A)$ then $\lambda: \mathbb{N}_0 \rightarrow A$ is injective.

Theorem 5.85. *Let A be a set, $f: A \rightarrow A$ a function then $\forall n \in \mathbb{N}_0$ there exist a function*

$$(f)^n: A \rightarrow A$$

such that:

1. $(f)^0 = \mathrm{Id}_A$
2. $(f)^{n+1} = f \circ (f)^n$

Theorem 5.86. *Let A be a set, $a \in A$ and $g: \mathbb{N}_0 \times A \rightarrow A$ then there exist a **unique** function*

$$\lambda: \mathbb{N}_0 \rightarrow A$$

such that:

1. $\lambda(0) = a$
2. $\forall n \in \mathbb{N}_0$ $\lambda(n+1) = g(n, \lambda(n))$

Corollary 5.87. *Let A be a set, $a \in A$ and $g: \mathbb{N}_0 \times A \rightarrow A$ then there exist a **unique** function*

$$\lambda: \mathbb{N}_0 \rightarrow A$$

such that:

1. $\lambda(0) = a$
2. $\forall n \in \{1, \ldots, \infty\}$ $\lambda(n) = g(n-1, \lambda(n-1))$

Proof. Using [theorem: 5.86] there exists a **unique** $\lambda: \mathbb{N}_0 \rightarrow A$ such that

$$\lambda(0) = a \text{ and } \forall n \in \mathbb{N}_0 \ \lambda(n+1) = g(n, \lambda(n)) \tag{5.17}$$

Let $n \in \{1, \ldots, \infty\}$ then $1 \leqslant n$ so by [definition: 5.63] we have that $n - 1 \in \mathbb{N}_0$ such that $n = (n-1) + 1$, hence $\lambda(n) = \lambda((n-1)+1) = g(n-1, \lambda(n-1))$. □

Theorem 5.88. *Let A be a set, $a \in A$, $n \in \mathbb{N}_0$ and $g: \{0, \ldots, n-1\} \times A \rightarrow A$ a function then there exists a **unique** function $\lambda: \{0, \ldots, n\} \rightarrow A$ satisfying*

$$\begin{array}{rcl} \lambda(0) & = & a \\ \forall i \in \{0, \ldots, n-1\} \textit{ we have } \lambda(i+1) & = & g(i, \lambda(i)) \end{array}$$

Proof. Define

$$g': \mathbb{N}_0 \times A \rightarrow A \text{ by } g'(i, x) = \begin{cases} g(i, x) \text{ if } i \in \{0, \ldots, n-1\} \\ x \text{ if } i \in \{n, \ldots, \infty\} \end{cases}$$

then by [corollary: 5.87] there exists a $\beta: \mathbb{N}_0 \rightarrow A$ such that

$$\beta(0) = a \tag{5.18}$$

$$\forall i \in \mathbb{N}_0 \text{ we have } \beta(i+1) = g'(i, \beta(i)) \tag{5.19}$$

Define now $\lambda: \{0, \ldots, n\} \rightarrow A$ by $\lambda = \beta_{|\{0, \ldots, n\}}$ then we have

$$\lambda(0) = \beta_{|\{0, \ldots, n\}}(0) \underset{0 \in \{0, \ldots, n\}}{=} \beta(0) \underset{[\text{theorem: } 5.18}{=} a$$

and $\forall i \in \{0, \ldots, n-1\}$ we have

$$\begin{array}{rcl} \lambda(i+1) & = & \beta_{|\{0, \ldots, n\}}(i+1) \\ & \underset{i+1 \in \{0, \ldots, n\}}{=} & \beta(i+1) \\ & \underset{[\text{theorem: } 5.19]}{=} & g'(i, \beta(i)) \\ & \underset{i \in \{0, \ldots, n-1\}}{=} & g'(i, \lambda(i)) \\ & \underset{i \in \{0, \ldots, n-1\}}{=} & g(i, \lambda(i)) \end{array}$$

so that we found a function $\lambda: \{0, \ldots, n\} \rightarrow A$ such that

$$\begin{array}{rcl} \lambda(0) & = & a \\ \forall i \in \{0, \ldots, n-1\} \text{ we have } \lambda(i+1) & = & g(i, \lambda(i)) \end{array}$$

Next we must prove uniqueness so let $\gamma: \{0, \ldots, n\} \rightarrow A$ be such that

$$\begin{array}{rcl} \gamma(0) & = & a \\ \forall i \in \{0, \ldots, n-1\} \text{ we have } \gamma(i+1) & = & g(i, \gamma(i)) \end{array}$$

and define $S = \{i \in \mathbb{N}_0 | i \notin \{0, \ldots, n\} \vee \lambda(i) = \gamma(i)\}$ then we have:

$0 \in S$. As $\lambda(0) = a = \gamma(0)$ we have $0 \in S$

$i \in S \Rightarrow i + 1 \in S$. then for $i + 1$ we have either:

> ***$i + 1 \in \{0, \ldots, n\}$.*** Then $i + 1 \leqslant n$ so that $i < n$ and as $i \in S$ we have $0 \leqslant i$, so it follows that $i \in \{0, \ldots, m\}$. Further $\lambda(i+1) = g(i, \lambda(i)) \underset{i \in \{0, \ldots, m\} \text{ and } i \in S}{=} g(i, \gamma(i)) = \gamma(i+1)$ proving that $i + 1 \in S$
>
> ***$i + 1 \notin \{0, \ldots, n\}$.*** Then $i + 1 \in S$

so in all cases we have $i + 1 \in S$.

By mathematical induction [theorem: 5.83] we have that $S = \mathbb{N}_0$. If $i \in \{0, \ldots, n\} \subseteq \mathbb{N}_0$ we have $i \in S$ which as $i \in \{0, \ldots, n\}$ gives $\lambda(i) = \gamma(i)$ so that $\lambda = \gamma$. □

In the above the function $\lambda: \mathbb{N}_0 \rightarrow A$ is specified by saying what $a \in A$ is and what the function $g: N_0 \times A \rightarrow A$ is. There exist a more intuitive way of specifying these requirement as is expressed in the following definitions.

Definition 5.89. *Let A be a set, $a \in A$ then we can define a function as follows:*

$$f: \mathbb{N}_0 \rightarrow A$$

is defined by:

1. $f(0) = a$
2. $f(n+1) = G(n, \lambda(n))$

where $G(n, \lambda(n))$ is a expression of two parameters. The above is equivalent with the function defined by [theorem: 5.86] where $a \in A$ and $g: \mathbb{N}_0 \times A \rightarrow A$ is defined by $g(n, x) = G(n, x)$.

Another way to define a recursive function is based on [corollary: 5.87]

Definition 5.90. *Let A be a set, $a \in A$ then we define $f: \mathbb{N}_0 \rightarrow A$ as follows*

$$f(n) = \begin{cases} a \textit{ if } n = 0 \\ G(n-1, f(n-1)) \textit{ if } n \in \{1, \ldots \infty\} \end{cases}$$

Which is equivalent with the function defined by [theorem: 5.87] where $a \in A$ and $g: \mathbb{N}_0 \times A \rightarrow A$ is defined by $g(n, x) = G(n, x)$.

We can use the above to define functions by recursion.

Definition 5.91. *Let A be a set, $a \in A$, $n \in \mathbb{N}_0$ then we define the function*

$$\lambda: \{0, \ldots, n\} \rightarrow A$$

by

$$\begin{array}{rcl} \lambda(0) & = & a \\ \forall i \in \{0, \ldots, n-1\} \textit{ we have } \lambda(i+1) & = & G(i, \lambda(i)) \end{array}$$

where $G(n, \lambda(n))$ is a expression of two parameters. The above is equivalent with the function defined by [theorem: 5.88] where $a \in A$ and $g: \{0, \ldots, n-1\} \times A \rightarrow A$ is defined by $g(n, x) = G(n, x)$.

Example 5.92. (Faculty) $\text{fac}: \mathbb{N}_0 \rightarrow \mathbb{N}_0$ is defined by

$$\text{fac}(n) = \begin{cases} 1 \text{ if n=0} \\ n \cdot \text{fac}(n-1) = ((n-1)+1) \cdot \text{fac}(n-1) \end{cases}$$

this is the function defined by [corollary: 5.87] where $a = 1$ and $g: \mathbb{N}_0 \times A \rightarrow A$ is define by $g(n, x) = (n+1) \cdot x$ then we have

$$\begin{array}{rcl} \text{fac}(0) & = & 1 \\ \text{fac}(1) & = & g(0, \text{fac}(0)) = (0+1) \cdot \text{fac}(0) = 1 \cdot \text{fac}(0) = 1 \cdot 1 = 1 \\ \text{fac}(2) & = & g(1, \text{fac}(1)) = (1+1) \cdot \text{fac}(1) = 2 \cdot \text{fac}(1) = 2 \cdot 1 = 2 \\ \text{fac}(3) & = & g(2, \text{fac}(2)) = (2+1) \cdot \text{fac}(2) = 3 \cdot \text{fac}(2) = 3 \cdot 2 = 6 \\ & \ldots & \\ \text{fac}(n) & = & g(n-1, \text{fac}(n-1)) = ((n-1)+1) \cdot \text{fac}(n-1) = n \cdot \text{fac}(n-1) \\ & \ldots & \end{array}$$

or in other words without using g

$$\begin{array}{rcl} \text{fac}(0) & = & 1 \\ \text{fac}(1) & = & 1 \cdot \text{fac}(0) = 1 \cdot 1 = 1 \\ \text{fac}(2) & = & 2 \cdot \text{fac}(1) = 2 \cdot 1 = 2 \\ \text{fac}(3) & = & 3 \cdot \text{fac}(2) = 3 \cdot 2 = 6 \\ & \ldots & \\ \text{fac}(n) & = & n \cdot \text{fac}(n-1) \\ & \ldots & \end{array}$$

which is exactly what we mean by the definition

$$\mathrm{fac}(n) = \begin{cases} 1 \text{ if n=0} \\ n \cdot \mathrm{fac}(n-1) \text{ if } n \in \{1, \ldots, \infty\} \end{cases}$$

Chapter 6
Finite and Infinite Sets

6.1 Equipotence

First we define the concept of equipotency which allows us to state that two sets have the same size without actually counting the number of elements. The latter will turn out to be impossible for every set.

Definition 6.1. *Two sets A and B are* ***equipotent*** *if there exist a bijection $f: A \rightarrow B$. We note this as $A \approx B$.*

Theorem 6.2. *Let A, B, C be sets then*

1. *$A \approx A$*
2. *If $A \approx B$ then $B \approx A$*
3. *If $A \approx B \wedge B \approx C$ then $A \approx C$*

Proof.

1. $\mathrm{Id}: A \rightarrow A$ is a bijection [see example: 2.64] proving that $A \approx A$
2. As $A \approx B$ there exist a bijection $f: A \rightarrow B$ but then by [theorem: 2.72] $f^{-1}|B \rightarrow A|$ is also a bijection, so that $B \approx A$.
3. If $A \approx B$ and $B \approx C$ then there exists bijections $f: A \rightarrow B$ and $g: B \rightarrow C$, using [theorem: 2.74] we have that $g \circ f: A \rightarrow C$ is a bijection, so $A \approx C$. □

Next we define a relation that says one set is smaller or equal to another set.

Definition 6.3. *Let A, B be sets then $A \preccurlyeq B$ if there exist a $C \subseteq B$ such that $A \approx C$.*

The following relation expresses that one set is smaller then another set.

Definition 6.4. *Let A, B be sets then $A \prec B$ if $A \preccurlyeq B$ and $\neg(A \approx B)$*

Clearly we have the following:

Theorem 6.5. *If A is a set then $\mathcal{P}(A) \approx 2^A$*

Proof. As $2 = s(1) = s(s(0)) = s(\{\varnothing\}) = \{\varnothing\} \bigcup \{\{\varnothing\}\} = \{\varnothing, \{\varnothing\}\} = \{0, 1\}$ we have that $2^A = \{0, 1\}^A$, finally using [theorem: 2.77] there exist a bijection $\mathcal{P}(A)$ and $\{0, 1\}^A$. □

Theorem 6.6. *Let A, B be sets then $A \preccurlyeq B$ if and only if there exist a injection $f: A \rightarrow B$*

Proof.

⇒. If $A \preccurlyeq B$ then there exist a set $C \subseteq B$ and a bijection $f: A \rightarrow C$, as a bijection is injective we have that $f: A \rightarrow C$ is injective and finally by [theorem: 2.52] $f: A \rightarrow B$ is a injection.

⇐. If $f: A \rightarrow C$ is a injection then by [theorem: 2.66] $f: A \rightarrow f(A)$ is a bijection where $f(A) \subseteq B$ proving that $A \preccurlyeq B$. □

Theorem 6.7. *If A is a set then there exist no surjection between A and $\mathcal{P}(A)$*

Proof. We prove this by contradiction. So assume that there exists a surjective function

$$f: A \to \mathcal{P}(A)$$

Define

$$B = \{x \in A | x \notin f(x)\}$$

As $B \subseteq A$ we have that $B \in \mathcal{P}(A)$ and by surjectivity there exists a $a \in A$ such that $f(a) = B$. If $a \in B$ then $a \notin f(a) = B$ leading to the contradiction $a \in B \wedge a \notin B$, so we must have $a \notin B = f(a)$ giving the contradiction $a \in B \wedge a \notin B$. So the the assumption must be wrong hence there is no surjection between A and $\mathcal{P}(A)$. □

Corollary 6.8. *If A is a set then no subset of A can be equipotent with $\mathcal{P}(A)$ or 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 to consider:

- **$B = A$.** Then by [theorem: 6.7] we can not have a surjection between B and $\mathcal{P}(A)$, which as a bijection is surjection, proves that there is no bijection between B and $\mathcal{P}(A)$. So B is not equipotent with $\mathcal{P}(A)$.
- **$B \subset A$.** Then $A \setminus B \neq \varnothing, B \bigcap (A \setminus B) = \varnothing$ and $A = (A \setminus B) \bigcup B$. Assume now that B is equipotent with $\mathcal{P}(A)$ then a bijection $g: B \to \mathcal{P}(A)$ exist, take the constant function $C_\varnothing: A \setminus B \to \mathcal{P}(A)$ where $C_\varnothing(x) = \varnothing$ and form then using [theorem: 2.79] the function

  $$f = g \bigcup C_\varnothing: A \to \mathcal{P}(A)$$

  If $C \in \mathcal{P}(A)$ then, as g is bijective $\exists x \in B$ such that $(x, C) \in g \subseteq f$ or $f(x) = C$, hence f is a surjection which is not allowed by [theorem: 6.7]. So B is not equipotent with $\mathcal{P}(A)$.

If $B \approx 2^A$ then, as by [theorem: 6.5] $2^A \approx \mathcal{P}(A)$, we have by [theorem: 6.2]] that $B \approx \mathcal{P}(A)$ which we have just shown to be impossible. So B can not be equipotency with 2^A. □

Theorem 6.9. *If A, B are sets, $A \neq \varnothing$ then there exists a injection $f: A \to B$ if and only there exist a surjection $g: B \to A$*

Proof.

- **⇒.** Let $f: A \to B$ be a injection then by [theorem: 2.60] there exist a $g: B \to A$ such that $g \circ f = \mathrm{Id}_A$. If $x \in A$ then $y = \mathrm{Id}_A(y) = (g \circ f)(y) = g(f(y))$ so that g is surjective.
- **⇐.** Let $g: B \to A$ be a surjection then by [theorem: 3.102] there exist a injective function $f: A \to B$. □

Corollary 6.10. *If A, B are sets then $A \preccurlyeq B$ if and only if there exist a surjection $f: B \to A$*

Proof. This follows from [theorem: 6.6] and the above [theorem: 6.9]. □

Theorem 6.11. *Let A, B, C, D classes with $A \bigcap C = \varnothing = B \bigcap D$, $A \approx B$ and $C \approx D$ then*

$$\left(A \bigcup C\right) \approx \left(B \bigcup D\right)$$

Proof. As $A \approx B$ and $C \approx D$ then there exists bijections $f: A \to B$ and $g: C \to D$ then by [theorem: 2.81] there exists a bijection $f \bigcup g: A \bigcup C \to B \bigcup D$. Hence $A \bigcup C \approx B \bigcup D$. □

Theorem 6.12. *If A, B, C, D are sets such that $A \approx B$ and $C \approx D$ then $A \times C \approx B \times D$*

Proof. As $A \approx B$ and $C \approx D$ there exist bijections $f: A \to B$ and $g: C \to D$. Define

$$h: A \times C \to B \times D \text{ by } h(x, y) = (f(x), g(x))$$

then we have:

- **injectivity.** If $h(x, y) = h(x', y')$ then $(f(x), g(x)) = (f(x'), g(x'))$ so that $f(x) = f'(x)$ and $g(x) = g(x')$, as f and g are injective we have $x = x'$ and $y = y'$ so that $(x, y) = (x', y')$.
- **surjectivity.** If $(r, s) \in B \times D$ then as f, g are surjective there exists $x \in A$, $y \in C$ such that $r = f(x)$ and $s = g(y)$ so that $h(x, y) = (f(x), g(y)) = (r, s)$. □

Theorem 6.13. *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 $C \approx D$ then there exists bijections $f: A \rightarrow B$ and $g: D \rightarrow C$. If $x \in A^C$ then $x: C \rightarrow A$ is a function, so $x \circ g: D \rightarrow A$ is a function, hence $f \circ (x \circ g): D \rightarrow B$ is a function, proving that $f \circ (x \circ g) \in B^D$. Define now $h: A^C \rightarrow B^D$ by $h(x) = f \circ (x \circ g)$ then we have:

injectivity. If $x, y \in A^C$ satisfies $h(x) = h(y)$ then

$$\begin{array}{rcl} f \circ (x \circ g) = f \circ (y \circ g) & \Rightarrow & f^{-1} \circ (f \circ (x \circ g)) = f^{-1} \circ (f \circ (y \circ g)) \\ & \Rightarrow & (f^{-1} \circ f) \circ (x \circ g) = (f^{-1} \circ f) \circ (y \circ g) \\ & \Rightarrow & \mathrm{Id}_A \circ (x \circ g) = \mathrm{Id}_A \circ (y \circ g) \\ & \Rightarrow & x \circ g = y \circ g \\ & \Rightarrow & (x \circ g) \circ g^{-1} = (y \circ g) \circ g^{-1} \\ & \Rightarrow & x \circ (g \circ g^{-1}) = y \circ (g \circ g^{-1}) \\ & \Rightarrow & x \circ \mathrm{Id}_C = y \circ \mathrm{Id}_C \\ & \Rightarrow x = y & \end{array}$$

surjectivity. If $y \in B^D$ then $y: D \rightarrow B$ is a function so that $y \circ g^{-1}: C \rightarrow B$ is a function, hence $f^{-1} \circ (y \circ g^{-1}): C \rightarrow A$ is a function or $f^{-1} \circ (y \circ g^{-1}) \in A^C$. Further

$$\begin{array}{rcl} h(f^{-1} \circ (y \circ g^{-1})) & = & f \circ ((f^{-1} \circ (y \circ g^{-1})) \circ g) \\ & = & f \circ ((f^{-1} \circ y) \circ (g^{-1} \circ g)) \\ & = & f \circ ((f^{-1} \circ y) \circ \mathrm{Id}_D) \\ & = & f \circ (f^{-1} \circ y) \\ & = & (f \circ f^{-1}) \circ y \\ & = & \mathrm{Id}_B \circ y \\ & = & y \end{array}$$

□

Theorem 6.14. *If A, B are sets such that $A \approx B$ then $\mathcal{P}(A) \approx \mathcal{P}(B)$ and $2^A \approx 2^B$*

Proof. As $A \approx B$ and $2 \approx 2$ [see theorem: 6.2], we have by [theorem: 6.13] that $2^A \approx 2^B$. Further by [theorem: 6.8] $\mathcal{P}[A] \approx 2^A$ and $\mathcal{P}(B) \approx 2^B$, so by [theorem: 6.2] it follows that $\mathcal{P}(A) = \mathcal{P}(B)$. □

6.2 Finite, Infinite and Denumerable sets

6.2.1 Finite and Infinite sets

Applying the concept of initial segments [see definition: 3.46] on $\langle \mathbb{N}_0 . \leqslant \rangle$ we have the following definition.

Definition 6.15. *Let $n \in \mathbb{N}_0$ then S_n is defined by*

$$S_n = \{m \in \mathbb{N}_0 | m < n\}$$

Actual we have already encountered the initial segments for $\langle \mathbb{N}_0, \leqslant \rangle$ because they are actual the natural numbers as is proved in the following theorem.

Theorem 6.16. *$\forall n \in \mathbb{N}_0$ we have $n = S_n$.*

Proof. We prove this by induction. So let $S = \{n \in \mathbb{N}_0 | n = S_n\}$ then we have:

$\mathbf{0 \in S}$**.** If $x \in S_0$ then $x < 0$ which by [theorem: 5.49] is false, so $S_0 = \varnothing = 0$ proving that $0 \in S$.

$\boldsymbol{n \in S \Rightarrow s(n) \in S}$**.** As $n \in S$ we have that $n = S_n$ so that

$$s(n) = n \bigcup \{n\} = S_n \bigcup \{n\}$$

If $m \in s(n)$ then we have the following possibilities to consider:

$\boldsymbol{m = n}$**.** Then by [theorem: 5.47] we have that $m < s(m) = s(n)$ so that $m \in S_{s(n)}$

$\boldsymbol{m \in S_n}$**.** Then $m < n$ which as by [theorem: 5.47] $n < s(n)$ proves that $m < s(n)$ hence $m \in S_{s(n)}$

this proves that

$$s(n) \subseteq S_{s(n)} \tag{6.1}$$

If $m \in S_{s(n)}$ then $m < s(n)$, now by [theorem: 5.53] we have either $n < m$ or $m \leqslant n$. If $n < m$ then by [theorem: 5.50] we have $s(n) \leqslant m$ so that by transitivity we have $m < m$ a contradiction. So we must have that $m \leqslant n$, if $m = n$ then $m \in n \bigcup \{n\} = s(n)$ and if $m < n$ then $m \in S_n \subseteq S_n \bigcup \{n\} = s(n)$. So in all cases we have $m \in s(n)$ proving that $S_{s(n)} \subseteq s(n)$, combining this with [eq: 6.1] gives

$$s(n) = S_{s(n)}$$

proving that $s(n) \in S$.

Using induction [theorem: 5.11] it follows that $S = \mathbb{N}_0$ proving the theorem. □

Theorem 6.17. *Let $n, m \in \mathbb{N}_0$ then*

$$n \leqslant m \Leftrightarrow S_n \subseteq S_m.$$

In other words as $n = S_n$ and $m = S_m$ we have

$$n \leqslant m \Leftrightarrow n \subseteq m$$

Proof.

⇒. If $x \in S_n$ then $x < n$ which as $n \leqslant m$ proves that $x < m$ so that $x \in S_m$, hence $S_n \subseteq S_m$.

⇐. By definition if $n \leqslant m$ then either $n = m \underset{n = S_n, m = S_m}{\Rightarrow} S_n = S_m \Rightarrow S_n \subseteq S_m$ or $n \in m$ which by [theorem: 5.14] we have that $n \subseteq m \underset{n = S_n, m = S_m}{\Rightarrow} S_n \subseteq S_m$. □

Theorem 6.18. *Let $n, m \in \mathbb{N}_0$ with $n \leqslant m$ then*

$$\beta: \{n, \ldots, m\} \rightarrow S_{(m-n)+1} \text{ where } \beta(i) = i - n$$

is a bijection with inverse

$$\beta^{-1}: S_{(m-n)+1} \rightarrow \{n, \ldots, m\} \text{ where } \beta^{-1}(i) = i + n$$

Proof. We have for the function $\beta: \{n, \ldots, m\} \rightarrow S_{(m-n)+1}$ where $\beta(i) = i - n$ the following:

injectivity. If $k, l \in \{n, \ldots, m\}$ such that $\beta(k) = \beta(l)$ then $k - n = l - n$, so by [theorem: 5.43] $k = (k - n) + n = (l - n) + n = l$ proving that $k = l$.

surjectivity. If $k \in S_{(m-n)+1}$ then $0 \leqslant k < (m - n) + 1$ so that by [theorem: 5.58] $0 \leqslant k \leqslant m - n$, then by [theorem: 5.55] we have that $n = 0 + n \leqslant k + n \leqslant (m - n) + n = m$. If we take $i = k + n$ then we have $0 \leqslant i \leqslant m$ and further $i - n = (k + n) - n \underset{\text{[theorem: 5.66]}}{=} k$ proving that $\beta(i) = k$.

So $\beta: \{n, \ldots, m\} \rightarrow S_{(m-n)+1}$ is a bijection. Further we have if $k \in S_{(m-n)+1}$ that $k = \beta(\beta^{-1}(k)) = \beta^{-1}(k) - n$ so that by [theorem: 5.43] $k + n = (\beta^{-1}(k) - n) + n = \beta^{-1}(k)$ proving that

$$\beta^{-1}: S_{(m-n)+1} \rightarrow \{n, \ldots, m\} \text{ is defined by } \beta^{-1}(k) = k + n$$ □

We define now the concept of a finite set.

Definition 6.19. (Finite Set) *A set A is **finite** if $\exists n \in \mathbb{N}_0$ such that $n \approx A$*

Example 6.20. $\varnothing$ is finite.

Proof. $\varnothing: \varnothing \rightarrow \varnothing$ is a bijection by [example: 2.63], so as $0 = \varnothing$ we have that $0 \approx \varnothing$. □

Lemma 6.21. *If $n \in \mathbb{N}_0$ then $n \approx \{1, \ldots, n\}$*

Proof. If $n \in \mathbb{N}_0$ then we have either:

$n = 0$. Then $n = 0 = \varnothing$ and $\{1, \ldots, 0\} = \varnothing$ so that $n \approx \{1, \ldots, n\}$

$n \neq 0$. Then $0 < n$ and we have for $\beta: \{1, \ldots, n\} \rightarrow S_n$ defined by $\beta(i) = i - 1$ that it satisfies:

injectivity. If $\beta(i) = \beta(j)$ then $i - 1 = j - 1$ so that $i = j$

surjectivity. If $j \in S_n$ then $0 \leqslant j < n$ so that $0 < j + 1 \leqslant n \Rightarrow 1 \leqslant j + 1 \leqslant n$, so if we take $i = j + 1$ we have that $i \in \{1, \ldots, n\}$ and $\beta(i) = (j + 1) - 1 = j$

proving β is a bijection. This proves that

$$\{1, \ldots, n\} \approx S_n \underset{[\text{theorem: } 6.16]}{=} n$$ □

Theorem 6.22. *As set A is finite if and only if there exist a $n \in \mathbb{N}_0$ such that $\{1, \ldots, n\} \approx A$.*

Proof. We have

$$\begin{aligned} A \text{ is finite} \quad & \Leftrightarrow \quad \exists n \in \mathbb{N}_0 \text{ such that } A \approx n \\ & \Leftrightarrow \sim_{[\text{theorem: } 6.21]} \exists n \in \mathbb{N}_0 \text{ such that } A \approx \{1, \ldots, n\} \end{aligned}$$

□

Proof. First, if $n \in \mathbb{N}_0$ with $0 < n$ we have for $\beta: \{1, \ldots, n\} \rightarrow S_n$ defined by $\beta(i) = i - 1$ that it satisfies:

injectivity. If $\beta(i) = \beta(j)$ then $i - 1 = j - 1$ so that $i = j$

surjectivity. If $j \in S_n$ then $0 \leqslant j < n$ so that $0 < j + 1 \leqslant n \Rightarrow 1 \leqslant j + 1 \leqslant n$, so if we take $i = j + 1$ we have that $i \in \{1, \ldots, n\}$ and $\beta(i) = (j + 1) - 1 = j$

proving β is a bijection. This proves that

$$\forall n \in \mathbb{N}_0 \setminus \{0\} \; \{1, \ldots, n\} \approx S_n \underset{[\text{theorem: } 6.16]}{=} n \tag{6.2}$$

Now for the final proof:

⇒. Then for A we have either:

$A = \varnothing$. Then we have as $\{1, \ldots, 0\} = \varnothing$ that $\{1, \ldots, 0\} \approx A$

$A \neq \varnothing$. Then there exists a $n \in \mathbb{N}_0$ such that $n \approx A$, as $0 \underset{\text{def}}{=} \varnothing$ we must have that $0 < n$. Combining this with [eq: 6.2] gives

$$\{1, \ldots, n\} \approx A$$

⇐. As $\exists n \in \mathbb{N}_0$ such that $A \approx \{1, \ldots, n\}$ we have for n either:

$n = 0$. Then $\{1, \ldots, n\} = \varnothing$ so that $A \approx \varnothing \Rightarrow A = \varnothing$ hence A is finite.

$n \neq 0$. Then by [eq: 6.2] $\{1, \ldots, n\} \approx n$ and as $A \approx \{1, \ldots, n\}$ we have that $A \approx n$ proving that A is finite. □

Definition 6.23. (Infinite Set) *A set A is **infinite** if A is not **finite**.*

Definition 6.24. (Denumerable Set) *A set $A$is **denumerable** or **infinite countable** if*

$$\mathbb{N}_0 \approx A.$$

Definition 6.25. (Countable Set) *A set A is countable if it is **finite** or **denumerable.***

Theorem 6.26. *If A, B are sets such that $A \approx B$ then we have*

1. *If A is finite then B is finite*
2. *If A is denumerable then B is denumerable*
3. *If A is countable then B is countable.*

Proof.

1. As A is finite there exists a $n \in \mathbb{N}_0$ such that $n \approx A$ which as $A \approx B$ proves by [theorem: 6.2] that $n \approx B$ hence B is finite.
2. As A is denumerable $\mathbb{N}_0 \approx A$ which as $A \approx B$ proves by [theorem: 6.2] that $\mathbb{N}_0 \approx B$ hence B is finite.
3. As A is countable it is either finite or denumerable, (1) and (2) ensures then that B is either finite or denumerable. □

Lemma 6.27. *If A is a **denumerable** set and $a \in A$ then $A \setminus \{a\}$ is a denumerable set.*

Proof. As A is denumerable there exist a bijection $f: \mathbb{N}_0 \rightarrow A$. As $a \in A$ we have by surjectivity that $\exists n \in \mathbb{N}_0$ such that $f(n) = a$. Define now

$$g: \mathbb{N}_0 \rightarrow A \text{ where } g(i) = \begin{cases} f(i) \text{ if } i < n \\ f(i + 1) \text{ if } n \leqslant i \end{cases}$$

which, as $\{x\in\mathbb{N}_0|x<n\}\bigcap\{x\in\mathbb{N}_0|n\leqslant x\}=\varnothing$ and $\mathbb{N}_0=\{x\in\mathbb{N}_0|x<n\}\bigcup\{x\in\mathbb{N}_0|n\leqslant x\}$, is a function As for bijectivity we have:

injectivity. If $g(i)=g(i')$ then for i,i' we have either:

$i<n\wedge i'<n$. Then $f(i)=g(i)=g(i')=f(i')$ which as f is injective proves that $i=i'$.

$i<n\wedge n\leqslant i'$. Then $f(i)=g(i)=g(i')=f(i'+1)$ which as f is injective proves that $i=i'+1$, Now as $n\leqslant i'<i'+1=i$ and $i<n$ we reach the contradiction $n<n$, so this case is not possible.

$n\leqslant i\wedge i'<n$. Then $f(i+1)=g(i)=g(i')=f(i')$ which as f is injective proves that $i+1=i'$. Now as $n\leqslant i<i+1=i'$ and $i'<n$ we reach the contradiction $n<n$, so this case is not possible.

$n\leqslant i\wedge n\leqslant i'$. Then $f(i+1)=g(i)=g(i')=f(i'+1)$, hence, as f is injective, we have $i+1=i+1'$ or by [theorem: 5.43] $i=i'$.

So in all valid cases we have $i=i'$ proving injectivity.

surjectivity. If $y\in A\setminus\{x\}$ then there exists by surjectivity of f a $i\in\mathbb{N}_0$ such that $f(i)=y$. We can not have $i=n$, because we would then have $f(i)=f(n)=y\notin A\setminus\{y\}$. So we have either

$i<n$. Then $g(i)=f(i)=y$

$n<i$. Then by [theorem: 5.67] $n\leqslant i-1$, so $g(i-1)=f((i-1)+1)=f(i)=y$

proving surjectivity. □

Lemma 6.28. *Let $n\in\mathbb{N}_0$ then n has no denumerable subset. In particular, as $n\subseteq n$, n is not denumerable.*

Proof. We prove this by induction, so define

$$S=\{n\in\mathbb{N}_0|n \text{ does not contain a denumerable subset}\}$$

then we have:

$0\in S$. As $0=\varnothing$ we have if $A\subseteq 0$ that $A=\varnothing$. If now $\mathbb{N}_0\approx A$ then there exists a bijection $f\colon\mathbb{N}_0\to A$ so that $f(0)\in A=\varnothing$ which is a contradiction. So 0 does not contains a denumerable subset.

$n\in S\Rightarrow n+1\in S$. We proceed by contradiction, so assume that there exist a $A\subseteq n+1=s(n)=n\bigcup\{n\}$ which is denumerable. If $n\notin A$ then $A\subseteq n$ which is impossible because $n\in S$, so we must have that $n\in A$. Let $a\in A\setminus\{n\}\subseteq n\bigcup\{n\}$ then, as $a\neq n$, $a\in n$ proving that $A\setminus\{n\}\subseteq n$. Now by the previous lemma [lemma: 6.27] we have, as A is denumerable, that $A\setminus\{n\}$ is denumerable which is forbidden as $n\in S$. So the assumption is wrong, hence every subset of $s(n)$ is not denumerable, proving that $n+1\in S$.

Using induction [see theorem: 5.83] it follows that $S=\{0,\ldots,\infty\}=\mathbb{N}_0$ proving the lemma. □

Theorem 6.29. *Let A be a set then A is infinite if and only if A contains a denumerable subset.*

Proof.

$\Rightarrow$. Let A be a infinite set. Using the well ordering theorem [see theorem: 3.127] there exists a order relation $\leqslant_A$ such that $\langle A,\leqslant_A\rangle$ is a well ordered set. Using [theorem: 3.96] and the fact that $\langle\mathbb{N},\leqslant\rangle$ is well ordered [see theorem: 5.51] we have exactly one of the following cases:

$\langle\mathbb{N}_0,\leqslant\rangle$ is order isomorphic with $\langle A,\leqslant_A\rangle$. This implies that $A\approx\mathbb{N}_0$ so that A is a denumerable subset of itself.

$\langle\mathbb{N}_0,\leqslant\rangle$ is order isomorphic with an initial segment of $\langle A,\leqslant_A\rangle$. This implies that A has a denumerable subset [the initial segment].

$\langle A,\leqslant_A\rangle$ is order isomorphic with an initial segment of $\langle\mathbb{N}_0,\leqslant\rangle$. So there exists a $n\in\mathbb{N}_0$ such that $A\approx S_n\underset{\text{[theorem: 6.16]}}{=}n$ so that A is finite, contradicting the fact that A is infinite. Hence this case does not apply.

So in all applicable cases we have that A contains a denumerable subset.

$\Leftarrow$. Let $B\subseteq A$ be a denumerable subset of A. Assume that A is finite then there exists a $n\in\mathbb{N}_0$ such that $n\approx A$, hence there exist a bijection $f\colon A\to n$. As $B\subseteq A$ we have that $f_{|B}\colon B\to f(B)$ is a bijection [see theorems: 2.84, 2.66] so that $B\approx f(B)$, as B is denumerable $\mathbb{N}_0\approx B$, so by [theorem: 6.2] it follows that $\mathbb{N}_0\approx f(B)\subseteq n$. So there exists a denumerable subset of n which by [theorem: 6.28] is impossible. Hence A is not finite which by definition means that A is infinite. □

Corollary 6.30. $\mathbb{N}_0$ *is infinite.*

Proof. As $\mathbb{N}_0\approx\mathbb{N}_0$ $\mathbb{N}_0$ is denumerable, clearly $\mathbb{N}_0\subseteq\mathbb{N}_0$ so by the previous theorem [theorem: 6.29] we have that $\mathbb{N}_0$ is infinite. □

Corollary 6.31. *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 [theorem: 6.29] the existence of a denumerable set $C \subseteq B$, but then $C \subseteq A$ and thus A has a denumerable subset. Using [theorem: 6.29] it follows that A is infinite. □

Corollary 6.32. *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 6.33. *If A and B are finite sets then $A \bigcup B$ is a finite set.*

Proof. As A is finite we have by [theorem: 6.32] that $A \setminus B$ is finite. So there exists $n, m \in \mathbb{N}_0$ such that $n \approx A \setminus B$ and $m \approx B$, hence we have two bijections

$$f: A \setminus B \rightarrow n \underset{[\text{theorem: } 6.16]}{=} S_n \text{ and } g': B \rightarrow m \underset{[\text{theorem: } 6.16]}{=} S_m \tag{6.3}$$

Define

$$C = \{i \in \mathbb{N}_0 | n \leqslant i \wedge i < n + m\}$$

If $b \in B$ then $g'(b) \in S_n$, hence $0 \leqslant g'(b) < m$ so that by [theorem: 5.55] $n = 0 + n \leqslant g'(b) + n < m + n$ or $g'(b) + n \in C$. So

$$g: B \rightarrow C \text{ where } g(i) = g'(i) + n \tag{6.4}$$

defines a function. Further we have:

- **injectivity.** If $g(b) = g(b')$ then $g'(b) + n = g'(b') + n$, so using [theorem: 5.43] $g'(b) = g'(b')$, hence, as g' is injective, we have $b = b'$.
- **surjectivity.** If $i \in C$ then $n \leqslant i < n + m$, using [theorem: 5.60] there exist a $k \in \mathbb{N}_0$ such that $n + k = i$. If $m \leqslant k$ then by [theorem: 5.55] $n + m \leqslant n + k = i < n + m$ a contradiction. So $k < m$ and thus $k \in S_m$. As g' is surjective there exists a $b \in B$ such that $g'(b) = k$ and thus $g(b) = g'(b) = k + n = i$.

proving that

$$g: B \rightarrow C \text{ is a bijection} \tag{6.5}$$

Further if $i \in n \bigcap C = S_n \bigcap C$ then $i < n \wedge n \leqslant i$ yielding the contradiction $i < i$ so we have that

$$n \bigcap C \neq \varnothing \tag{6.6}$$

If $i \in n \bigcup C$ then either

- **$i \in n$.** Then, as $n = S_n$, we have $i < n$ which as $n \leqslant n + m$ proves that $i < n + m$ hence $i \in S_{n+m}$.
- **$i \in C$.** Then $i < n + m$ so that $i \in S_{n+m}$

proving

$$n \bigcup C \subseteq S_{n+m} \tag{6.7}$$

If $i \in S_{n+m}$ then $i < n + m$, further we have either $i < n$ so that $i \in S_n = n$ or $n \leqslant i$ giving $i \in C$, hence $i \in n \bigcup C$ or $S_{n+m} \subseteq n \bigcup C$ which by [eq: 6.7] proves that

$$n \bigcup C = S_{n+m} \tag{6.8}$$

Using [eq: 6.3], [eq: 6.5], [eq 6.6],[eq: 6.8], $A \bigcup B = (A \setminus B) \bigcup B$ and $(A \setminus B) \bigcap B = \varnothing$ allows use to use [theorem: 2.81] to get the bijection

$$f \bigcup g: A \bigcup B \rightarrow S_{n+m}$$

proving that

$$A \bigcup B \approx S_{n+m}$$

□

Lemma 6.34. *If $\{A_i\}_{i \in S_n}$ is such that $\forall i \in S_n$ A_i is finite then $\bigcup_{i \in S_n} A_i$ is finite.*

Proof. We use induction to prove this, so define

$$S = \left\{ n \in \mathbb{N}_0 | \text{If } \{A_i\}_{i \in S_n} \text{ satisfies } \forall i \in S_n \ A_i \text{ is finite then } \bigcup_{i \in S_n} A_i \text{ is finite} \right\}$$

then we have:

- **$0 \in S$.** If $n = 0$ then $S_0 = 0 = \varnothing$ so that $\bigcup_{i \in S_0} A_i = \bigcup_{i \in \varnothing} A_i \underset{[\text{example: } 2.119}{=} \varnothing$ which is finite, hence $0 \in S$.

$\boldsymbol{n \in S \Rightarrow n+1 \in S}$. Let $\{A_i\}_{i \in n+1}$ a family of finite sets. As $S_{n+1} = n+1 = s(n) = n \bigcup \{n\} = S_n \bigcup \{n\}$ and $n \notin S_n$ we have that $S_{n+1} \setminus \{n\} = S_n$. So

$$\bigcup_{i \in S_{n+1}} A_i \underset{[\text{theorem: } 2.132]}{=} \left(\bigcup_{i \in S_{n+1} \setminus \{n\}} A_i \right) \bigcup A_n = \left(\bigcup_{i \in S_n} A_i \right) \bigcup A_n$$

As $n \in S$ we have that $\bigcup_{i \in S_n} A_i$ is finite which, as A_n is also finite, proves, using [theorem: 6.33] that $(\bigcup_{i \in S_n} A_i) \bigcup A_n$ is finite. So $\bigcup_{i \in S_{n+1}} A_i$ is finite proving that $n+1 \in S$.

Mathematical induction [see theorem: 5.83] proves then the lemma. □

Theorem 6.35. *If $\{A_i\}_{i \in I}$ is a such that I is finite and $\forall i \in I$ A_i is finite then $\bigcup_{i \in I} A_i$ is finite.*

Proof. As I is finite there exists a $n \in \mathbb{N}_0$ and a bijection $f: S_n \rightarrow I$ so that by [theorem: 2.117] we have that

$$\bigcup_{i \in I} A_i = \bigcup_{i \in S_n} A_{f(i)} \tag{6.9}$$

Using the previous lemma [lemma: 6.34] it follows that $\bigcup_{i \in S_n} A_{f(i)}$ is finite, hence using [eq: 6.9] we have

$$\bigcup_{i \in I} A_i \text{ is finite}$$ □

Theorem 6.36. *A set A is infinite if and only if $\exists B \subset A$ such that $B \approx A$. In other words A is infinite if and only if A is equipotent with a proper subset of itself.*

Proof.

$\boldsymbol{\Rightarrow}$. If A is infinite then by [theorem: 6.29] there exist a denumerable $B \subseteq A$. So there exists a bijection $f: \mathbb{N}_0 \rightarrow B$. Define now the function [taking in account that $(A \setminus B) \bigcap B = \varnothing$ and $A = (A \setminus B) \bigcup B)$]

$$g: A \rightarrow A \text{ where } g(x) = \begin{cases} x \text{ if } x \in A \setminus B \\ f(f^{-1}(x)+1) \text{ if } x \in B \end{cases}$$

where $f^{-1}: B \rightarrow \mathbb{N}_0$ is the inverse of f.
Then we have:

$$g(A) = A \setminus \{f(0)\} \tag{6.10}$$

Proof. If $y \in g(A)$ then there exists a $x \in A$ such that $y = g(x)$, we have for x either:

$\boldsymbol{x \in A \setminus B}$. Then $y = g(x) = x$ so that $y \in A \setminus B$ or as $f(0) \in B$ that $y \in A \setminus \{f(0)\}$.

$\boldsymbol{x \in B}$. If $f(0) = f(f^{-1}(x)+1)$ we have, as f is a bijection hence injective, that $0 = f^{-1}(x)+1$ which contradicts $0 < f^{-1}(x)+1$. So we must have that

$$f(0) \neq f(f^{-1}(x)+1) = y.$$

proving $y \in A \setminus \{f(0)\}$.

So we conclude that

$$g(A) \subseteq A \setminus \{f(0)\} \tag{6.11}$$

If $y \in A \setminus \{f(0)\}$ then we have either:

$\boldsymbol{y \in B}$. If $f^{-1}(y) = 0$ we would have that $y = f(f^{-1}(y)) = f(0)$ contradicting $y \in A \setminus \{f(0)\}$. So we have that $f^{-1}(y) \neq 0$ or $0 < f^{-1}(y)$, using [theorem: 5.67] we have then that $0 \leqslant f^{-1}(y) - 1$. Take then $x = f(f^{-1}(y) - 1) \in B \subseteq A$ then we have:

$$\begin{aligned} g(x) &= f(f^{-1}(x)+1) \\ &= f(f^{-1}(f(f^{-1}(y)-1))+1) \\ &= f((f^{-1}(y)-1)+1) \\ &= f(f^{-1}(y)) \\ &= y \end{aligned}$$

so that $y \in g(A)$.

$\boldsymbol{y \notin B}$**.** Then $y \in A \setminus B$ so that $g(y) = y$ proving that $y \in g(A)$

So we conclude that $A \setminus \{f(0)\} \subseteq g(A)$ which combined with [eq: 6.11] proves $g(A) = A \setminus \{f(0)\}$. □

Next we proof that $g: A \rightarrow A$ is injective

Proof. Let $x, x' \in A$ such that $g(x) = g(x')$ then for x, x' we have to consider the following possible cases:

$\boldsymbol{x \in B \wedge x' \in B}$**.** then $f(f^{-1}(x) + 1) = g(x) = g(x') = f(f^{-1}(x') + 1)$ so that

$$\begin{aligned} f(f^{-1}(x) + 1) = f(f^{-1}(x') + 1) & \underset{f \text{ is injective}}{\Rightarrow} & f^{-1}(x) + 1 = f^{-1}(x') + 1 \\ & \Rightarrow & f^{-1}(x) = f^{-1}(x') \\ & \underset{f^{-1} \text{ is injective}}{\Rightarrow} & x = x' \end{aligned}$$

$\boldsymbol{x \in B \wedge x' \notin B}$**.** Then $f(f^{-1}(x) + 1) = g(x) = g(x') = x'$ so that $f(f^{-1}(x) + 1) \notin B$ contradicting $f: \mathbb{N}_0 \rightarrow B$. So this case does not apply.

$\boldsymbol{x \notin B \wedge x' \in B}$**.** Then $x = g(x) = g(x') = f(f^{-1}(x) + 1)$ so that $f(f^{-1}(x) + 1) \notin B$ contradicting $f: \mathbb{N}_0 \rightarrow B$, So this case never occurs.

$\boldsymbol{x \notin B \wedge x' \notin B}$**.** Then $x = g(x) = g(x') = x'$. □

So we have proved that

$$g: A \rightarrow A \text{ is injective} \tag{6.12}$$

Using [eq: 6.10] and [eq: 6.12] proves that $g: A \rightarrow A \setminus \{f(0)\}$ is a bijection or

$$A \approx A \setminus \{f(0)\}$$

Further as $f(0) \in B \subseteq A$ we have that $A \neq A \setminus \{f(0)\}$ giving $A \setminus \{f(0)\} \subset A$. Hence we have proved that A is equipotent with a proper subset of itself.

⇐. Assume that there exists a proper subset $B \subset A$ such that $A \approx B$ then there exists a bijection $f: A \rightarrow B$, resulting in the injection [see theorem: 2.52]

$$f: A \rightarrow A \text{ with } f(A) = B \subset A$$

As $f(A) \subset A$ there exists a $a \in A$ such that $a \notin f(A)$. Using recursion [theorem: 5.84] there exist a injection $\lambda: \mathbb{N}_0 \rightarrow A$ such that $\lambda(0) = a$ and $\forall n \in \mathbb{N}_0$ $\lambda(n+1) = f(\lambda(n))$. Hence we have a bijection $\lambda: \mathbb{N}_0 \rightarrow \lambda(A)$ proving that $\lambda(A)$ is denumerable, as $\lambda(A) \subseteq A$ it follows from [theorem: 6.29] that A is infinite. □

The following theorem allows you to quantify the number of elements in a finite set.

Theorem 6.37. *If $n, m \in \mathbb{N}_0$ such that $n \approx m$ then $n = m$.*

Proof. Assume that $n \approx m$ then by [theorem: 5.53] we have either $n < m$, $m < n$ or $n = m$. If

$\boldsymbol{n < m}$**.** Then $\forall i \in n = S_n$ we have $i < n < m \Rightarrow i < m$ so that $i \in S_m = m$ which as $n \neq m$ proves that $n \subset m$. So m is equipotent to a proper subset of itself which by [theorem: 6.36] would mean that m is infinite contradicting the fact that m is finite [as $m \approx m$].

$\boldsymbol{m < n}$**.** Then $\forall i \in m = S_m$ we have $i < m < n \Rightarrow i < n$ so that $i \in S_n = n$ which as $n \neq m$ proves that $m \subset n$. So n is equipotent to a proper subset of itself which by [theorem: 6.36] would mean that n is infinite contradicting the fact that n is finite [as $n \approx n$].

So the only option left is

$$n = m$$ □

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 $n' \approx A$ then $n \approx n'$, hence $n = n'$. This leads to the following definition.

Definition 6.38. *If A is a **finite** set then $\exists! n \in \mathbb{N}_0$ such that $n \approx A$. This unique number is noted as $\# A$, so $\# A \approx A$. $\# A$ can be interpreted as the number of elements in A.*

Theorem 6.39. *If A is a set then $A = \varnothing \Leftrightarrow \# A = 0$*

Proof.

⇒. If $A=\varnothing$ then by [example: 2.63] $\varnothing: \varnothing \rightarrow \oslash$ is a bijection, so as $0=\varnothing$ we have $\#\varnothing=0$.

⇐. If $\#A=0$ then as $0=\varnothing$ there exists a bijection $f: \varnothing \rightarrow A$, Assume that $A \neq \varnothing$ then there exist a $y \in A$ and as f is a bijection we would have a $x \in \varnothing$ such that $f(x)=y$ contradicting the fact that $\forall x\ x \notin \varnothing$. □

Theorem 6.40. *If A, B are finite sets then $A \times B$ is finite and $\#(A \times B)=\#A \cdot \#B$*

Proof. We have for A, B to consider the following possibilities:

$\boldsymbol{A=\varnothing \vee B=\varnothing}$**.** Then $0=\varnothing \approx A$ and $0=\varnothing \approx B$ so that $\#A=0=\#B$, further by [theorem: 1.47] $0=\varnothing=A \times B$ hence $\#(A \times B)=0=\#A \cdot \#B$.

$\boldsymbol{A \neq \varnothing \wedge B \neq \varnothing}$**.** Take $n=\#A \neq 0$ and $m=\#B \neq 0$ then there exist bijections $f: B \rightarrow n=S_n$ and $g: A \rightarrow m=S_m$. Now $\forall x \in A, \forall y \in B$ we have $f(x)<n$ and $g(y)<m$, using [theorem: 5.67] we have $g(y) \leqslant m-1$. So by [theorem: 5.76]

$$n \cdot g(x)=g(x) \cdot n \leqslant (m-1) \cdot n \underset{\text{[theorem: 5.69]}}{=} m \cdot n-n,$$

further by [theorem: 5.74] we have

$$(m \cdot n-n)+f(x)<(m \cdot n-n)+n=m \cdot n=n \cdot m$$

This allows us to define the function

$$h: A \times B \rightarrow S_{n \cdot m} \text{ where } h(x, y)=n \cdot g(x)+f(x)$$

then we have:

injectivity. If $h(x, y)=h(x', y')$ then $n \cdot g(x)+f(x)=n \cdot g(x')+f(x')$. As $0 \leqslant f(x)<n$ and $0 \leqslant f(x')<n$ it follows from [theorem: 5.79] that $g(x)=g(x')$ and $f(x)=f(x')$ which as f, g are bijections gives $x=x'$ and $y=y'$ so that $(x, y)=(x', y')$.

surjectivity. If $z \in S_{n \cdot m}$ then $0 \leqslant z<n \cdot m$, using [theorem: 5.79] there exist a q, r such that $z=q \cdot n+r$ and $0 \leqslant r<n$. If $m \leqslant q \underset{\text{[theorem: 5.76]}}{\Rightarrow} m \cdot n \leqslant q \cdot n \underset{\text{[theorem: 5.55]}}{\Rightarrow} m \cdot n+r \leqslant q \cdot n+r=z<n \cdot m$ so that $n \cdot m+r<n \cdot m$ or $r+n \cdot m<0+n \cdot \underset{\text{[theorem: 5.55]}}{\Rightarrow} r<0$ a contradiction, hence $q<m$. So we have proved that $r \in S_n$ and $q \in S_m$, as f, g are bijections there exists $x \in A$, $y \in B$ such that $f(x)=r$ and $g(y)=q$. So $h(x, y)=n \cdot g(x)+f(x)=n \cdot q+r=z$.

Hence we have $A \times B \approx S_{n \cdot m}$ proving that $A \times B$ is finite and $\#(A \times B)=n \cdot m=\#A \cdot \#B$. □

Theorem 6.41. *If A, B are finite sets such that $A \bigcap B=\varnothing$ then $\#(A \bigcup B)=\#A+\#B$*

Proof. Let $n=\#A$, $m=\#B$ then there exist bijections $f: A \rightarrow S_n$ and $g: B \rightarrow S_m$. If $x \in A$ then $f(x)<n<n+m$ and if $x \in B$ then $g(x)<m \Rightarrow n+g(x)<n+m$, as further $A \bigcap B=\varnothing$ we can define the function

$$h: A \bigcup B \rightarrow S_{n+m} \text{ where } h(x)=\begin{cases} f(x) \text{ if } x \in A \\ n+g(x) \text{ if } x \in B \end{cases}$$

We prove now that this is a bijection.

injectivity. If $h(x)=h(x')$ then we have the following cases to consider for $x, x' \in A \bigcup B$:

$\boldsymbol{x \in A \wedge x' \in A}$**.** Then $f(x)=h(x)=h(x')=f(x')$ which as f is a bijection gives $x=x'$.

$\boldsymbol{x \in A \wedge x' \in B}$**.** Then $f(x)=h(x)=h(x')=n+g(x')$, now as $f(x)<n$ we have

$$n+g(x')=f(x)<n+0$$

so that by [theorem: 5.55] $g(x')<0$, a contradiction. So this case will never occur.

$\boldsymbol{x \in B \wedge x' \in A}$**.** Then $n+g(x)=h(x)=h(x')=f(x')$, now as $f(x')<n$ we have

$$n+g(x)=f'(x)<n+0$$

so that by [theorem: 5.55] $g(x)<0$, a contradiction. So this case will never occur.

$\boldsymbol{x \in B \wedge x \in B}$**.** Then $g(x)+n=n+g(x)=h(x)=h(x')=n+g(x')=g(x')+n$ so that by [theorem: 5.55] $g(x)=g(x')$, which as g is a bijection proves that $x=x'$.

surjectivity. If $y \in S_{n+m}$ then $y<n+m$ and we have the following cases for y to consider:

$\boldsymbol{y<n}$**.** Then $y \in S_n$ so that by surjectivity of f we have a $x \in A$ such that $f(x)=y$, hence $h(x)=f(x)=y$

$n \leqslant y$. Then $n \leqslant y < n+m$, by [theorem: 5.70] we have then that $0 \leqslant y-m < (n+m)-n \underset{\text{[theorem: 5.66]}}{=} m$, proving that $y-n \in S_m$. As g is a surjection there exists a $x \in B$ such that $g(x) = y-n$, hence $h(n) = n+g(x) = n+(y-n) = y$. □

Theorem 6.42. *If A is a finite set and $B \subseteq A$ then:*

1. *B is finite*
2. *$A \setminus B$ is finite*
3. *$\#B \leqslant \#A$*
4. *If $B \subset A$ then $\#B < \#A$*
5. *$\#A = \#B + \#(A \setminus B)$*

Proof. As A is finite there exist $n \in \mathbb{N}_0$ and a bijection $f: n = S_n \rightarrow A$. We have then to consider the following possibilities:

$B = A$. Then obviously B is finite, $A \setminus B = \varnothing$ is also finite, $\#B = \#A \Rightarrow \#B \leqslant \#A$ and $\#B + \#(A \setminus B) = \#A + \#\varnothing \doteq \#A + 0 = \#A$, So (1), (2),(3), (4) and (5) are satisfied.

$B = \varnothing$. Then clearly B is finite, $A \setminus B = A$ is finite, $\#B = 0 \leqslant \#A$ and $\#B + \#(A \setminus B) = 0 + \#A = \#A$, further if $B \subset A$ then $A \neq \varnothing$ so that $\#B = 0 < \#A$.

$\varnothing \neq B \subset A$. As every subset of a finite set is finite [see theorem: 6.32] we have that B and $A \setminus B$ are finite, further as $B \subset A$ we have that $A \setminus B \neq \varnothing$ so that

$$0 < \#(A \setminus B).$$

As $B \bigcap (A \setminus B) = \varnothing$ and $A \bigcup B = (A \setminus B) \bigcup B$ it follows from [theorem: 6.41] that

$$\#A = \#B + \#(A \setminus B)$$

Now if $\#A \leqslant \#B$ then as $0 < \#(A \setminus B)$ it follows from [theorem: 5.74] that

$$\#A = \#A + 0 < \#B + \#(A \setminus B) = \#A$$

a contradiction, so we must have that

$$\#B < \#A$$

So (1),(2),(3),(4) and (5) are satisfied. □

Corollary 6.43. *If A, B are sets, A is finite and $f: A \rightarrow B$ is a surjection then B is finite and $\#B \leqslant \#A$.*

Proof. If $B = \varnothing$ then B is finite and $\#B = 0 \leqslant \#A$ proving the theorem in this case. If $B \neq \varnothing$ then by [theorem: 6.9] there exist as injection $g: B \rightarrow A$, leading by [theorem: 2.66] to a bijection $g: B \rightarrow g(B)$, hence $B \approx g(B)$. As $g(B) \subseteq A$ we have by [theorem: 6.42] that $g(B)$ is finite and $\#g(B) \leqslant \#A$. Finally as $\#g(B) \approx B$ and $B \approx g(B)$ it follows that $\#B = \#g(B) \leqslant \#(A)$. □

corollary

Corollary 6.44. *If A, B are sets, A finite and $f: A \rightarrow B$ a function then $f(A)$ is finite*

Proof. As $f: A \rightarrow f(A)$ is a surjection it follows from the previous theorem [theroem: 6.43]. □

Theorem 6.45. *Let I be a finite set and $\{x_i\}_{i \in I} \subseteq X$ a finite family of elements in X then $\{x_i | i \in I\}$ is finite and $\#(\{x_i | i \in I\}) \leqslant \#I$*

Proof. Define the function $f: I \rightarrow \{x_i | i \in I\}$ by $f(i) = x_i$ then if $y \in \{x_i | i \in I\}$ there exist a $i \in I$ such that $y = x_i$, hence $y = f(i)$. This proves that $f: I \rightarrow \{x_i | i \in I\}$ is a surjection, so by the previous corollary [corollary: 6.43] we have as I is finite that $\{x_i | i \in I\}$ is finite and $\#(\{x_i | i \in I\}) \leqslant \#I$. □

Theorem 6.46. *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 $g: n \rightarrow f(A)$, as $f: A \rightarrow f(A)$ is a bijection we have that $f^{-1}: f(A) \rightarrow A$ is a bijection so that $f^{-1} \circ g: n \rightarrow A$ is a bijection, hence A is finite, contradicting the fact that A is infinite. So the assumption is wrong hence B is infinite. □

Corollary 6.47. *Let I be a infinite set and $\{A_i\}_{i \in I}$ a family such that $\forall i \in I$ we have $A_i \neq \varnothing$ and $\forall i, j \in I$ with $i \neq j$ we have $A_i \bigcap A_j \neq \varnothing$ then $\bigcup_{i \in I} A_i$ is infinite.*

Proof. Using the Axiom of Choice [see 3.103] there exist a function $f: I \rightarrow \bigcup_{i \in I} A_i$ such that $\forall i \in I$ $f(i) \in A_i$. Let $i, j \in I$ such that $f(i) = f(j)$ then $f(i) \in A_i$ and $f(j) = f(i) \in A_j$ so that $A_i \bigcap A_j \neq \varnothing$, hence $i = j$ [for if $i \neq j$ then $A_i \bigcap A_j = \varnothing$]. Hence $f: I \rightarrow \bigcup_{i \in I} A_i$ is a injection which by [theorem: 6.46] proves that $\bigcup_{i \in I} A_i$ is infinite. □

Theorem 6.48. *Let $\langle X, \leqslant \rangle$ be a totally ordered set, $\varnothing \neq A \subseteq X$ a finite set then* $\max(A)$ *and* $\min(A)$ *exists.*

Proof. We prove this by induction on $\#A$, so let

$$S = \{n \in \{1, \ldots, \infty\} | \text{If } A \subseteq X \text{ with } \#A = n \text{ then} \max(A) \text{ and } \min(A) \text{ exists}\}$$

then we have:

$1 \in S$. As $\#A = 1 = \{0\}$ there exists a bijection $f: \{0\} \rightarrow A$ so that $A = \{f(0)\}$ and $\max(A) = f(0) = \min(A)$.

$n \in S \Rightarrow n + 1 \in S$. Let $A \subseteq X$ with $\#A = n + 1$ then $n + 1 = s(n) = n \bigcup \{n\}$, so that there exists a bijection $f: n \bigcup \{n\} \rightarrow A$. If $n \in n$ then $n < n$ a contradiction so we have $n \notin n$. Take now

$$f_{|n}: n \rightarrow A \setminus \{f(n)\}$$

then by [theorem: 2.84] $f_{|n}$ is injective. Further if $y \in A \setminus \{f(n)\}$ then, as f is a bijection, there exists a $i \in n + 1$ such that $f(i) = y$, we can not have $i = n$ [because then $f(i) = f(n)$], so $i \neq n \Rightarrow i \in n$, proving that $f_{|n|}(i) = f(i) = y$. Hence $f_{|n}: n \rightarrow A$ is a surjection, which together with injectivity proving that

$$f_{|n}: n \rightarrow A \setminus \{f(n)\} \text{ is a bijection hence } \#(A \setminus \{f(n)\}) = n$$

As $n \in S$ we have that $M = \max(A \setminus \{f(n)\})$ and $m = \min(A \setminus \{f(n)\})$ exists. We have now for $M, f(n)$ to consider the following possibilities::

$M \leqslant f(n)$. Then $\forall x \in A \setminus \{f(n)\}$ we have $x \leqslant M \leqslant f(n) \Rightarrow x \leqslant f(n)$ and for $x = f(n)$ $x \leqslant f(n)$. So $\forall x \in A$ we have $x \leqslant f(n)$, proving that $\max(A)$ exist and $\max(A) = f(n)$.

$f(n) < M$. Then $\forall x \in A$ we have $x \leqslant M$ so that $\max(A)$ exist and $\max(A) = M$

For $m, f(n)$ we need to consider:

$m \leqslant f(n)$. Then $\forall x \in A$ we have $m \leqslant x$ so that $\min(A)$ exist and $\min(A) = m$.

$f(n) < m$. Then $\forall x \in A \setminus \{f(n)\}$ we have $m \leqslant x$ so that $f(n) < m$ and for $x = f(n)$ $x \leqslant f(n)$. So $\forall x \in A$ we have $f(n) \leqslant x$ proving that $\min(A)$ exist and that $f(n) = \min(A)$.

As $\min(A)$ and $\max(A)$ exist it follows that $n + 1 \in S$

Using induction [see theorem:5.83] it follows that $\{1, \ldots, \infty\} = S$. Assume now that $\varnothing \neq A \subseteq X$ such that A is finite we must have that $\#A \in \{1, \ldots, \infty\}$ [for if $\#A = 0$ then $A = \varnothing$], so that $\min(A)$ and $\max(A)$ exist. □

Theorem 6.49. *If A is a finite set and $f: \mathbb{N}_0 \rightarrow A$ a function then $\exists a \in A$ such that $f^{-1}(\{a\})$ is infinite.*

Proof. Assume that $\forall a \in A$ $f^{-1}(\{a\})$ is finite. As A is finite we have for the family $\{f^{-1}(\{a\})\}_{a \in A}$ by [theorem: 6.35] that $\bigcup_{a \in A} f^{-1}(\{a\})$ is finite. Now

$$\begin{aligned} x \in \bigcup_{a \in A} f^{-1}(\{a\}) &\Leftrightarrow \exists a \in A \text{ such that } x \in f^{-1}(\{a\}) \\ &\Leftrightarrow \exists a \in A \text{ such that } f(x) \in \{a\} \\ &\Leftrightarrow \exists a \in A \text{ such that } f(x) = a \\ &\Leftrightarrow x \in f^{-1}(A) \end{aligned}$$

So that $\mathbb{N}_0 = f^{-1}(A) = \bigcup_{a \in A} f^{-1}(\{a\})$ from which it follows that $\mathbb{N}_0$ is finite contradicting the fact that $\mathbb{N}_0$ is infinite [by theorem: 6.30]. So the assumption is wrong, hence $\exists a \in A$ such that $f^{-1}(\{a\})$ is infinite. □

Corollary 6.50. *If A is finite and $f: \mathbb{N}_0 \rightarrow A$ a function then $\exists a \in A$ such that $\forall n \in \mathbb{N}_0$ there exist a $m \in \{n, \ldots, \infty\}$ so that $f(m) = a$.*

Proof. By the preceding theorem [theorem: 6.49] there exist a $a \in A$ such that $f^{-1}(\{a\})$ is infinite. Assume now that $\exists n \in \mathbb{N}_0$ such that $\forall m \in \{n, \ldots, \infty\}$ we have $f(m) \neq a$. If $m \in f^{-1}(\{a\})$ then $f(m) \in \{a\} \Rightarrow f(m) = a$, so we must have that $m \notin \{n, \ldots, \infty\}$, hence $m < n$ or $m \in S_n$. So we have proved that $f^{-1}(\{a\}) \subseteq S_n$ a finite set, giving by [theorem: 6.42] that $f^{-1}(\{a\})$ is finite contradicting the fact that $f^{-1}(\{a\})$ is infinite. So the assumption must be wrong, hence $\forall n \in \mathbb{N}_0$ there exists a $m \in \{n, \ldots, \infty\}$ such that $f(m) = a$. □

6.2.2 Finite families

We show now that every finite family of elements of a totally ordered set can be sorted.

Theorem 6.51. *Let $\langle X, \leqslant \rangle$ be a totally ordered set, $n \in \mathbb{N}_0$ and $\{x_i\}_{i \in S_{n+1}} \subseteq X$ then there exists a bijection $\beta: S_{n+1} \to S_{n+1}$ such that $\forall i \in S_n$ we have $x_{\beta(i)} \leqslant x_{\beta(n)}$.*

Proof. We prove this by induction, so let

$$S = \{n \in \mathbb{N}_0 | \forall \{x_i\}_{i \in S_{n+1}} \subseteq X \text{ there exist a bijection } \beta: S_{n+1} \to S_{n+1} \text{ such that } \forall i \in S_n \; x_{\beta(i)} \leqslant x_{\beta(n)}\}$$

then we have:

$0 \in S$. If $\{x_i\}_{i \in S_1 = \{0\}} \subseteq X$ then for the bijection $\beta = \mathrm{Id}_{S_1}: S_1 \to S_1$ we have $\forall i \in S_0 = \varnothing$ that $x_{\beta(i)} \leqslant x_{\beta(0)}$ is satisfied vacuously, proving that $0 \in S$.

$n \in S \Rightarrow n+1 \in S$. Let $\{x_i\}_{i \in S_{(n+1)+1}} \subseteq X$ then for $\{x_i\}_{i \in S_{n+1}}$ we have, as $n \in S$, the existence of a bijection $\alpha: S_{n+1} \to S_{n+1}$ such that $\forall i \in S_n \; x_{\alpha(i)} \leqslant x_{\alpha(n)}$. For x_{n+1} we have now two cases to consider:

$x_{\alpha(n)} \leqslant x_{n+1}$. Define

$$\beta: S_{(n+1)+1} \to S_{(n+1)+1} \text{ by } \beta(i) = \begin{cases} \alpha(i) \text{ if } i \in S_{n+1} \\ n+1 \text{ if } i = n+1 \end{cases}$$

then we have:

injectivity. Let $i, j \in S_{(n+1)+1}$ be such that $\beta(i) = \beta(j)$ then we have the following possibilities:

$i \in S_{n+1} \wedge j \in S_{n+1}$. Then $\alpha(i) = \beta(i) = \beta(j) = \alpha(j)$ which as α is a bijection proves that $i = j$.

$i \in S_{n+1} \wedge j = n+1$. Then $\alpha(i) = \beta(i) = \beta(j) = n+1$ from which it follows that $n+1 = \alpha(i) \in S_{n+1}$ giving the contradiction $n+1$<n+1. So this case never occurs.

$i = n+1 \wedge j \in S_{n+1}$. Then $n+1 = \beta(i) = \beta(j) = \alpha(j)$ from which it follows that $n+1 = \alpha(j) \in S_{n+1}$ giving the contradiction $n+1$<n+1. So this case never occurs.

$i = n+1 \wedge j = n+1$. Then $i = j$

surjectivity. If $j \in S_{(n+1)+1}$ then we have the following possibilities:

$j = n+1$. Then $n+1 = \beta(n+1)$.

$j \in S_n$. Then as α is a bijection there exist a $i \in S_n$ such that $j = \alpha(i) \underset{i \in S_n}{\Rightarrow} j = \beta(i)$.

So $\beta: S_{(n+1)+1} \to S_{(n+1)+1}$ is a bijection. Let now $i \in S_{n+1}$ then we have the following possibilities:

$i = n$. Then $x_{\beta(i)} = x_{\alpha(i)} = x_{\alpha(n)} \leqslant x_{n+1} = x_{\beta(n+1)}$.

$i \in S_n$. Then $x_{\beta(i)} = x_{\alpha(i)} \leqslant x_{\alpha(n)} \leqslant x_{n+1} = x_{\beta(n+1)}$.

which proves that in this case we have $n+1 \in S$.

$x_{n+1} < x_{\alpha(n)}$. Define

$$\beta: S_{(n+1)+1} \to S_{(n+1)+1} \text{ by } \beta(i) = \begin{cases} \alpha(i) \text{ if } i \in S_n \\ n+1 \text{ if } i = n \\ \alpha(n) \text{ if } i = n+1 \end{cases}$$

then we have:

injectivity. Let $i, j \in S_{(n+1)+1}$ such that $\beta(i) = \beta(j)$ then we have the following possibilities:

$i \in S_n \wedge j \in S_n$. Then $\alpha(i) = \beta(i) = \beta(j) = \alpha(j)$ which as β is a bijection gives $i = j$.

$i \in S_n \wedge j = n$. Then $\alpha(i) = \beta(i) = \beta(j) = n+1$ so that $n+1 = \alpha(i) \in S_{n+1}$ giving the contradiction $n+1 < n+1$, so this case never occurs.

$i \in S_n \wedge j = n+1$. Then $\alpha(i) = \beta(i) = \beta(j) = \alpha(n)$, which as α is a bijection, gives $i = n$ contradicting $i \in S_n \Rightarrow i < n$, so this case never occurs.

$i = n \wedge j \in S_n$. Then $n+1 = \beta(i) = \beta(j) = \alpha(j)$ so that $n+1 = \alpha(j) \in S_{n+1}$ giving the contradiction $n+1 < n+1$, so this case never occurs.

$i = n \wedge j = n$. Then $i = j$.

$i = n \wedge j = n+1$. Then $n+1 = \beta(i) = \beta(j) = \alpha(n)$ so that $n+1 = \alpha(n) \in S_{n+1}$ giving the contradiction $n+1 < n+1$, so this case never occurs.

$i = n+1 \wedge j \in S_n$. Then $\alpha(n) = \beta(i) = \beta(j) = \alpha(j)$, which as α is a bijection gives $n = j \in S_n$ resulting in the contradiction $n < n$, so this case never occurs.

$i = n+1 \wedge j = n$. Then $\alpha(n) = \beta(i) = \beta(j) = n+1$ so that $n+1 = \alpha(n) \in S_{n+1}$ leading to the contradiction $n+1 < n+1$, so this case never occur.

$i = n+1 \wedge j = n+1$. Then $i = j$.

surjectivity. Let $j \in S_{(n+1)+1}$ then we have the following possibilities to check:

$j = n+1$. then $\beta(n) = j$

$j \in S_{n+1}$. then as α is a bijection there exist a $i \in S_{n+1}$ so that $\alpha(i) = j$. If $i = n$ then $\beta(n+1) = \alpha(n) = j$ and if $i \in S_n$ then $\beta(i) = \alpha(i) = j$.

So $\beta: S_{(n+1)+1} \rightarrow S_{(n+1)+1}$ is a bijection. Let now $i \in S_{n+1}$ then we have to consider the following possibilities:

$i = n$. Then $x_{\beta(i)} = x_{n+1} \leqslant x_{\alpha(n)} = x_{\beta(n+1)}$.

$i \in S_n$. Then $x_{\beta(i)} = x_{\alpha(i)} \leqslant x_{\alpha(n)} = x_{\beta(n+1)}$,

which proves that in this case $n+1 \in S$.

Mathematical induction [see theorem: 5.83] proves then that $S = \mathbb{N}_0$. □

Corollary 6.52. *Let $\langle X, \leqslant \rangle$, $n, m \in \mathbb{N}_0$ such that $n \leqslant m$ and $\{x_i\}_{i \in \{n,\ldots,m\}} \subseteq X$ then there exist a bijection $\alpha: \{n,\ldots,m\} \rightarrow \{n,\ldots,m\}$ such that $\forall i \in \{n,\ldots,m-1\}$ we have $x_{\alpha(i)} \leqslant x_{\alpha(m)}$*

Proof. Using [theorem: 6.18] there exists bijections

$$\beta: \{n,\ldots,m\} \rightarrow S_{(m-n)+1} \text{ where } \beta(i) = i - n \tag{6.13}$$

and

$$\beta^{-1}: S_{(m-n)+1} \rightarrow \{n,\ldots,m\} \text{ where } \beta(i) = i + n \tag{6.14}$$

Let $\{x_i\}_{i \in \{n,\ldots,m\}} \subseteq X$ then for $\{x_{\beta^{-1}(i)}\}_{i \in S_{(m-n)+1}}$ we have by [theorem: 6.51] a bijection

$$\gamma: S_{(m-n)+1} \rightarrow S_{(m-n)+1} \text{ such that } \forall i \in S_{m-n} \text{ we have } x_{\beta^{-1}(\gamma(i))} \leqslant x_{\beta^{-1}(\gamma(m-n))} \tag{6.15}$$

Define now the bijection

$$\alpha = \beta^{-1} \circ \gamma \circ \beta: \{n,\ldots,m\} \rightarrow \{n,\ldots,m\}$$

If $k \in \{n,\ldots,m-1\}$ then $n \leqslant k \leqslant m-1 < m$ so that by [theorem: 5.70] we have $0 \leqslant k-n < m-n$ or $0 \leqslant \beta(k) < m-n$. So $\beta(k) \in S_{m-n}$ and thus by [eq: 6.15] we have that

$$x_{\beta^{-1}(\gamma(\beta(k)))} \leqslant x_{\beta^{-1}(\gamma(m-n))} \underset{\beta(m)=m-n}{=} x_{\beta^{-1}(\gamma(\beta(m)))} \tag{6.16}$$

Hence

$$\begin{array}{rcl} x_{\alpha(k)} & = & x_{(\beta^{-1}\circ\gamma\circ\beta)(k)} \\ & = & x_{\beta^{-1}(\gamma(\beta(k)))} \\ & \leqslant_{[\text{eq: } 6.16]} & x_{\beta^{-1}(\gamma(\beta(m)))} \\ & = & x_{(\beta^{-1}\circ\gamma\circ\beta)(m)} \\ & = & x_{\alpha(m)} \end{array}$$

So we have found a bijection $\alpha: \{n,\ldots,m\} \rightarrow \{n,\ldots,m\}$ such that $\forall k \in \{n,\ldots,m-1\}$ $x_{\alpha(k)} \leqslant x_{\alpha(m)}$ □

Theorem 6.53. *Let $\langle X, \leqslant \rangle$ be a totally ordered set, $n \in \mathbb{N}_0$ and $\{x_i\}_{i \in S_{n+1}} \subseteq X$ then there exists a bijection $\beta: S_{n+1} \rightarrow S_{n+1}$ such that*

$$\forall i \in S_n \text{ we have } x_{\beta(i)} \leqslant x_{\beta(i+1)}$$

Proof. We proof this by induction, so let

$$S = \{n \in \mathbb{N}_0 | \forall \{x_i\}_{i \in n+1} \subseteq X \text{ there exist a bijection } \beta: S_{n+1} \rightarrow S_{n+1} \text{ such that } \forall i \in S_n\ x_{\beta(i)} \leqslant x_{\beta(i+1)}\}$$

then we have:

$0 \in S$. Then $S_0 = \varnothing$ and $S_1 = \{0\}$. Let $\{x_i\}_{i \in S_1 = \{0\}} \subseteq X$ then, for the bijection $\beta: S_1 \rightarrow S_1$ where $\beta = \mathrm{Id}_{S_1}$, we have that, $\forall i \in S_0 = \varnothing$ $x_{\beta(i)} \leqslant x_{\beta(i+1)}$, is satisfied vacuously.

$n \in S \Rightarrow n+1 \in S$. Let $\{x_i\}_{i \in S_{(n+1)+1}} \subseteq X$ then by the previous theorem [theorem: 6.51] there exists a bijection

$$\alpha: S_{(n+1)+1} \rightarrow S_{(n+1)+1} \text{ such that } \forall i \in S_{n+1}\ x_{\alpha(i)} \leqslant x_{\alpha(n+1)} \tag{6.17}$$

Consider now $\{x_{\alpha(i)}\}_{i \in S_{n+1}}$ then as $n \in S$ we have the existence of a bijection

$$\gamma: S_{n+1} \rightarrow S_{n+1} \text{ such that } \forall i \in S_n \text{ we have } x_{\alpha(\gamma(i))} \leqslant x_{\alpha(\gamma(i+1))} \tag{6.18}$$

Define now

$$\beta: S_{(n+1)+1} \rightarrow S_{(n+1)+1} \text{ by } \beta(i) = \begin{cases} \alpha(\gamma(i)) \text{ if } i \in S_{n+1} \\ \alpha(n+1) \text{ if } i = n+1 \end{cases}$$

then we have:

injectivity. Let $k, l \in S$ be such that $\beta(k) = \beta(l)$ then we must consider the following possibilities:

$k \in S_{n+1} \wedge l \in S_{n+1}$. Then

$$(\alpha \circ \gamma)(k) = \alpha(\gamma(k)) = \beta(k) = \beta(l) = (\alpha(\gamma(l))) = (\alpha \circ \gamma)(l)$$

which as $\alpha \circ \lambda$ is a bijection proves that $k = l$.

$k \in S_{n+1} \wedge l = n+1$. Then $\alpha(n+1) = \beta(l) = \beta(k) = \alpha(\gamma(k))$ which, as α is a bijection, gives $n+1 = \gamma(k)$, as $\gamma(k) \in S_{n+1} \Rightarrow \gamma(k) < n+1$ we reach the contradiction $n+1 < n+1$, so this case never occurs.

$k = n+1 \wedge l \in S_{n+1}$. Then $\alpha(n+1) = \beta(k) = \beta(l) = \alpha(\gamma(l))$ which, as α is a bijection, gives $n+1 = \gamma(l)$, as $\gamma(l) \in S_{n+1} \Rightarrow \gamma(l) < n+1$, we reach the contradiction $n+1 < n+1$, so this case never occurs.

$k = n+1 \wedge l = n+1$. then $k = l$

surjectivity. If $k \in S_{(n+1)+1}$ we have, as α is a bijection, that there exist a $l \in S_{(n+1)+1}$ such that $\alpha(l) = k$, for l we have then the following possibilities:

$l = n+1$. Then $\beta(n+1) = \alpha(n+1) = k$

$l \in S_{n+1}$. Then as γ is a bijection there exist a $i \in S_{n+1}$ such that $l = \gamma(i)$, hence $\beta(i) = \alpha(\gamma(i)) = \alpha(l) = k$.

Further if $i \in S_{n+1}$ we have the following possibilities to consider:

$i = n$. Then $\gamma(n) \in S_{n+1}$ so that by [eq: 6.17] $x_{\alpha(\gamma(i))} \leqslant x_{\alpha(n+1)} = x_{\beta(n+1)}$ hence

$$x_{\beta(i)} = x_{\alpha(\gamma(i))} \leqslant x_{\beta(n+1)} = x_{\beta(i+1)}$$

$i \in S_n$. Then by [eq: 6.18] we have $x_{\alpha(\gamma(i))} \leqslant x_{\alpha(\gamma(i+1))}$ so that

$$x_{\beta(i)} = x_{\alpha(\gamma(i))} \leqslant x_{\alpha(\gamma(i+1))} = x_{\beta(i+1)}$$

Hence $\forall i \in S_{n+1}$ we have $x_{\beta(i)} \leqslant x_{\beta(i+1)}$ proving that $n+1 \in S$.

Mathematical induction [see theorem: 5.83] proves that $S = \mathbb{N}_0$ and thus the theorem. □

Corollary 6.54. *Let $\langle X, \leqslant \rangle$ be a totally ordered set, $n, m \in \mathbb{N}_0$ with $n \leqslant m$ and $\{x_i\}_{i \in \{n, \ldots, m\}} \subseteq X$ then there exist a bijection α: $\{n, \ldots, m\} \rightarrow \{n, \ldots, m\}$ such that $\forall i \in \{n, \ldots, m-1\}$ $x_{\alpha(i)} \leqslant x_{\alpha(i+1)}$*

Proof. Using [theorem: 6.18] there exists bijections

$$\beta: \{n, \ldots, m\} \rightarrow S_{(m-n)+1} \text{ where } \beta(i) = i - n \tag{6.19}$$

and

$$\beta^{-1}: S_{(m-n)+1} \rightarrow \{n, \ldots, m\} \text{ where } \beta(i) = i + n \tag{6.20}$$

Let $\{x_i\}_{i \in \{n, \ldots, m\}} \subseteq X$ then for $\{x_{\beta^{-1}(i)}\}_{i \in S_{(m-n)+1}}$ we have by [theorem: 6.53] a bijection

$$\gamma: S_{(m-n)+1} \rightarrow S_{(m-n)+1} \text{ such that } \forall i \in S_{m-n} \text{ we have } x_{\beta^{-1}(\gamma(i))} \leqslant x_{\beta^{-1}(\gamma(i+1))} \tag{6.21}$$

Define now the bijection

$$\alpha = \beta^{-1} \circ \gamma \circ \beta: \{n, \ldots, m\} \rightarrow \{n, \ldots, m\}$$

If $k \in \{n, \ldots, m-1\}$ then $n \leqslant k \leqslant m-1 < m$ so that by [theorem: 5.70] we have $0 \leqslant k-n < m-n$ or $0 \leqslant \beta(k) < m-n$. So $\beta(k) \in S_{m-n}$ and thus by [eq: 6.21] we have that

$$x_{\beta^{-1}(\gamma(\beta(k)))} \leqslant x_{\beta^{-1}(\gamma(\beta(k)+1))}$$

Now $\beta(k+1) = (k+1) - n \underset{\text{[theorem: 5.65]}}{=} (k-n) + 1 = \beta(k) + 1$ so that by the above we have

$$x_{\beta^{-1}(\gamma(\beta(k)))} \leqslant x_{\beta^{-1}\gamma(\beta(k+1))} \tag{6.22}$$

Hence

$$\begin{array}{rcl} x_{\alpha(k)} & = & x_{(\beta^{-1}\circ\gamma\circ\beta)(k)} \\ & = & x_{\beta^{-1}(\gamma(\beta(k)))} \\ & \leqslant_{\text{[eq: 6.22]}} & x_{\beta^{-1}(\gamma(\beta(k+1)))} \\ & = & x_{(\beta^{-1}\circ\gamma\circ\beta)(k+1)} \\ & = & x_{\alpha(k+1)} \end{array}$$

So we have found a bijection $\alpha: \{n, \ldots, m\} \to \{n, \ldots, m\}$ such that $\forall k \in \{n, \ldots, m-1\}$ $x_{\alpha(k)} \leqslant x_{\alpha(k+1)}$ □

The next theorem allows use later to apply induction on the product of a finite family of sets.

Theorem 6.55. *Let $n \in \mathbb{N}_0$ and let $\{A_i\}_{i \in S_{n+1}}$ a family of sets then*

$$\prod_{i \in S_{n+1}} A_i \approx \left(\prod_{i \in S_n} A_i \right) \times A_n$$

Proof. If $x \in \prod_{i \in S_{n+1}} A_i$ then $x \in \left(\bigcup_{i \in S_{n+1}} A_i\right)^{S_{n+1}}$ such that $\forall i \in S_{n+1}$ we have $x(i) \in A_i$ or equivalently $x: S_{n+1} \to \bigcup_{i \in S_{n+1}} A_i$ is a function so that $\forall i \in S_{n+1}$ we have $x(i) \in A_i$. As $\forall i \in S_n$ we have $x(i) \in A_i \subseteq \bigcup_{i \in S_n} A_i$, it follows that $x_{|S_n}: S_n \to \bigcup_{i \in S_n} A_i$ is a function, so $x_{|S_n} \in \prod_{i \in S_n} A_i$. Hence we can define the following function

$$\beta: \left(\prod_{i \in S_{n+1}} A_i \right) \to \left(\prod_{i \in S_n} A_i \right) \times A_n \text{ by } \beta(x) \to (x_{|S_n}, x(n))$$

Then we have:

injectivity. If $\beta(x) = \beta(y)$ then $(x_{|S_n}, x(n)) = (y_{|S_n}, y(n))$ or $x_{|S_n} = y_{|S_n}$ and $x(n) = y(n)$. So if $i \in S_{n+1}$ we have either $i \in S_n$ then $x(i) = x_{|S_n}(i) = y_{|S_n}(i) = y(i)$ or $i = n$ and then $x(i) = x(n) = y(n) = y(i)$, proving that $x = y$.

surjectivity. Let $(y, a) \in (\prod_{i \in S_n} A_i) \times A_n$ then $y \in \prod_{i \in S_n} A_i$ and $a \in A_n$. Define then the function:

$$x: S_{n+1} \to \bigcup_{i \in S_{n+1}} A_i \text{ by } x(i) = \left\{ \begin{array}{l} y(i) \text{ if } i \in S_n \\ a \text{ if } i = n \end{array} \right.$$

Then $\forall i \in S_{n+1}$ we have either $i \in S_n$ giving $x(i) = y(i) \in A_i$ or $i = n$ giving $x(i) = x(n) = a \in A_n$, proving that $x \in \prod_{i \in S_{n+1}} A_i$. Further as clearly $x_{|S_n} = y$ and $x(n) = a$ we have that $\beta(x) = y$. □

We use the above theorem to prove that the product of a finite family of finite sets is finite.

Theorem 6.56. *Let $n \in \mathbb{N}_0 \setminus \{0\}$ and $\{A_i\}_{i \in S_n}$ be such that $\forall i \in S_n$ A_i is finite then $\prod_{i \in S_n} A_i$ is finite.*

Proof. we proof this by induction so define

$$S = \left\{ n \in \{1, \ldots, \infty\} | \text{If } \{A_i\}_{i \in S_n} \text{satisfies } \forall i \in S_n \quad A_i \text{ is finite then } \prod_{i \in S_n} A_i \text{ is finite} \right\}$$

then we have:

$1 \in S$. Using [example: 2.137] there exist a bijection $\beta: A_0 \to \prod_{i \in \{0\}} A_i$, hence as $S_1 = \{0\}$ $A_0 \approx \prod_{i \in S_1} A_i$. As A_0 is finite there exist a $k \in \mathbb{N}_0$ such that $k \approx A_0$ proving that $k \approx \prod_{i \in S_0} A_i$ or that $\prod_{i \in S_1} A_i$ is finite. So $1 \in S$.

$n \in S$ then $n + 1 \in S$. Let $\{A_i\}_{i \in S_{n+1}} A_i$ be such that that $\forall i \in S_{n+1}$ we have that A_i is finite. As $n \in S$ we have that $\prod_{i \in S_n} A_i$ is finite so using [theorem: 6.40] it follows that $(\prod_{i \in S_n} A_i) \times A_n$ is finite. Hence $\exists k \in \mathbb{N}_0$ such that $k \approx (\prod_{i \in S_n} A_i) \times A_n$. Using [theorem: 6.55] we have $(\prod_{i \in S_n} A_i) \times A_n \approx \prod_{i \in S_{n+1}} A_i$ proving that $k \approx \prod_{i \in S_{n+1}} A_i$. Hence $\prod_{i \in S_{n+1}} A_i$ is finite proving that $n + 1 \in S$.

Using mathematical induction it follows that $S=\{1,\ldots,\infty\}$ proving the theorem. □

Corollary 6.57. *Let I be a non empty finite set and $\{A_i\}_{i\in I}$ is such that $\forall i\in I$ we have A_i is finite then $\prod_{i\in I} A_i$ is finite.*

Proof. As I is finite and $I\neq\varnothing$ there exists a $n\in\mathbb{N}_0\setminus\{0\}$ such that $k\approx I$, so there exist a bijection $f: S_k\rightarrow I$. Using [theorem: 2.143] we have that there exists a bijection $\beta: \prod_{i\in I} A_i\rightarrow\prod_{i\in S_k} A_{f(i)}$ hence $\prod_{i\in I} A_i\approx\prod_{i\in S_k} A_{f(i)}$. By [theorem: 6.56] we have that $\prod_{i\in S_k} A_{f(i)}$ is finite so there exists a $m\in\mathbb{N}_0$ such that $m\approx\prod_{i\in S_k} A_{f(i)}$, hence $m\approx\prod_{i\in I} A_i$, proving that $\prod_{i\in I} A_i$ is finite. □

6.2.3 Denumerable sets

Lemma 6.58. *Every subset of $\mathbb{N}_0$ is either finite or denumerable*

Proof. By [theorem: 5.51[$\langle\mathbb{N}_0,\leqslant\rangle$ is a well ordered set, hence by [theorem: 3.97] we have for $N\subseteq\mathbb{N}_0$ either:

1. N is order isomorphic with $\mathbb{N}_0$ hence $N\approx\mathbb{N}_0$ proving that N is denumerable.
2. N is order isomorphic with a initial segment of $\mathbb{N}_0$ so there exists a $n\in\mathbb{N}_0$ such that $N\approx S_n$ proving that N is finite. □

Theorem 6.59. *Every subset of a denumerable set is finite or denumerable.*

Proof. Let A be a denumerable set and $B\subseteq A$. As A is denumerable there exists a bijection

$$\beta: A\rightarrow\mathbb{N}_0$$

Using [theorem: 2.84] and [theorem: 2.66] we have that $\beta_{|B}: B\rightarrow\beta(B)$ is a bijection so that

$$\beta(B)\approx B$$

as $\beta(B)\subseteq\mathbb{N}$ we have by the previous lemma [lemma: 6.58] that either:

$\boldsymbol{\beta(B)\approx\mathbb{N}_0}$. Then by [theorem: 6.2] $B\approx\mathbb{N}_0$ proving that B is denumerable.

$\boldsymbol{\beta(B)}$ is finite. Then there exists a $n\in\mathbb{N}_0$ such that $\beta(B)\approx n$, by [theorem: 6.2] $B\approx n$ proving that B is finite. □

Theorem 6.60. $\mathbb{N}_0\times\mathbb{N}_0\approx\mathbb{N}_0$*, in other words* $\mathbb{N}_0\times\mathbb{N}_0$ *is denumerable/*

Proof. First define the function

$$f: \mathbb{N}_0\times\mathbb{N}_0\rightarrow\mathbb{N}_0\times\mathbb{N}_0 \text{ where } f(k,m)=\begin{cases}(0,k+1) \text{ if m=0}\\(k+1,m-1) \text{ if } m\in\mathbb{N}_0\setminus\{0\}\end{cases}$$

If $f(k,m)=f(k',m')$ we have the following cases for m,m'

$\boldsymbol{m=0\wedge m'=0}$. Then $m=m'$ and $(0,k+1)=f(k,m)=f(k',m')=(0,k'+1)$ so that $k+1=k'+1 \underset{\text{[theorem: 5.43]}}{\Rightarrow} k=k'$ hence $(k,m)=(k',m')$.

$\boldsymbol{m=0\wedge m'\in\mathbb{N}_0\setminus\{0\}}$. Then $(0,k+1)=f(k,m)=f(k',m')=(k'+1,m'-1)$ so that $0=k'+1$ which as $0<s(k')=k'+1$ is a contradiction, so this case does not occur.

$\boldsymbol{m\in\mathbb{N}_0\setminus\{0\}\wedge m'=0}$. Then $(k+1,m-1)=f(k,m)=f(k',m')=(0,k'+1)$ so that $0=k+1$ which as $<s(k)=k+1$ is a contradiction, so this case does not occur.

$\boldsymbol{m\in\mathbb{N}_0\setminus\{0\}\wedge m'\in\mathbb{N}_0\setminus\{0\}}$. Then $(k+1,m-1)=f(k,m)=f(k',m')=(k'+1,m'-1)$ so that $k+1=k'+1 \underset{\text{[theorem: 5.43]}}{\Rightarrow} k=k'$ and $m-1=m'-1 \underset{\text{[theorem: 5.43]}}{\Rightarrow} m=(m-1)+1=(m'-1)+1=m'$ so that $(k,m)=(k',m')$

The above proves that

$$f: \mathbb{N}_0\times\mathbb{N}_0\rightarrow\mathbb{N}_0\times\mathbb{N}_0 \text{ is injective} \tag{6.23}$$

Assume that $f(k,m)=(0,0)$ then if $m=0$ we have $(0,0)=(0,k+1)$ giving the contradiction $0=k+1$ and if $m\neq 0$ we have $(k+1,m-1)$ giving the contradiction $0=k+1$. So the assumption is incorrect hence

$$(0,0)\notin f(\mathbb{N}_0\times\mathbb{N}_0) \tag{6.24}$$

Using [eq: 6.23] and [eq: 6.24] we can use recursion [see theorem: 5.84] to get a **injective** function

$$\lambda: \mathbb{N}_0\rightarrow\mathbb{N}_0\times\mathbb{N}_0 \text{ such that } \lambda(0)=(0,0) \text{ and } \forall n\in\mathbb{N}_0 \text{ we have } \lambda(n+1)=f(\lambda(n))$$

We prove now the following proposition about λ:

Proposition 6.61. *If there exist a $n,m\in\mathbb{N}_0$ such that $\lambda(n)=(0,m)$ then if $k,l\in\mathbb{N}_0$ is such that $k+l=m$ we have $\lambda(n+k)=(k,l)$.*

Proof. We proof this by induction so let

$$S_{n,m}=\{k\in\mathbb{N}_0|\text{For } l\in\mathbb{N}_0 \text{ with } k+l=m \text{ we have } \lambda(n+k)=(k,l)\}$$

then we have:

$0\in S_{n,m}$. If $l\in\mathbb{N}_0$ such that $k+l=m$ then $l=m$ and $\lambda(n+k)=\lambda(n)=(0,m)\underset{k=0\wedge l=m}{=}(k,l)$ proving that $0\in S_{n,m}$.

$k\in S_{n,m}\Rightarrow k+1\in S_{n,m}$. If $l\in\mathbb{N}_0$ such that $(k+1)+l=m$ then we have $k+(l+1)=m$ and as $k\in S_{n,m}$ it follows that

$$\lambda(n+k)=(k,l+1) \tag{6.25}$$

Further

$$\begin{array}{rcl}\lambda(n+(k+1)) & = & \lambda((n+k)+1)\\ & = & f(\lambda(n+k))\\ & \underset{[\text{eq: } 6.25]}{=} & f(k,l+1)\\ & \underset{l+1\neq 0}{=} & (k+1,(l+1)-1)\\ & \underset{[\text{theorem: } 5.66]}{=} & (k+1,l)\end{array}$$

proving that $k+1\in S_{n,m}$.

Using induction [theorem: 5.83] it follows that $S_{n,m}=\mathbb{N}_0$ proving the proposition. □

We prove now using induction that $\lambda:\mathbb{N}_0\rightarrow\mathbb{N}_0\times\mathbb{N}_0$ is surjective. So let

$$S=\{n\in\mathbb{N}_0|\text{For } (k,m)\in\mathbb{N}_0\times\mathbb{N}_0 \text{ with } k+m=n \text{ there exists a } l\in\mathbb{N}_0 \text{ such that } \lambda(l)=(k,m)\}$$

$0\in S$. If $(k,m)\in\mathbb{N}_0\times\mathbb{N}_0$ is such that $k+m=0$ then we must have $k=m=0$, as $\lambda(0)=(0,0)=(k,l)$ we have $0\in S$.

$n\in S$ then $n+1\in S$. Let $(k,m)\in\mathbb{N}_0$ be such that $k+m=n+1$, then for k we have to consider the following cases:

$k=0$. Then $m=k+m=n+1$ so that $(k,m)=(0,m)=(0,n+1)=f(n,0)$. As $n\in S$ and $n=n+0$ there exist a $l\in\mathbb{N}_0$ such that $\lambda(l)=(n,0)$. So

$$\lambda(l+1)=f(\lambda(l))=f(n,0)=(0,n+1)\underset{k=0}{=}(k,m)$$

$k\neq 0$. Then $0<k$ so that $0\leqslant k-1$, further as $0\neq m+1$ we have that

$$f(k-1,m+1)=((k-1)+1,(m+1)-1)=(k,m)$$

Let $k'=(k+m)-1\underset{\text{k+m=n+1}}{=}(n+1)-1=n$ and $l'=0$ then $k'+l'=n$ so that, as $n\in S$, there exist a $l\in\mathbb{N}_0$ such that

$$\lambda(l)=(k',l')=((k+m)-1,0) \tag{6.26}$$

Hence

$$\begin{array}{rcl}\lambda(l+1) & = & f(\lambda(l))\\ & \underset{[\text{eq: } 6.26]}{=} & f((k+m)-1,0)\\ & = & (0,k+m)\end{array}$$

Combining the above with [proposition: 6.61] we have that $\lambda((l+1)+k)=(k,m)$, so that $n+1\in S$.

By mathematical induction [theorem: 5.83] it follows that $S=\mathbb{N}_0$. So if $(k,m)\in\mathbb{N}_0\times\mathbb{N}_0$ we have that $k+m\in\mathbb{N}_0=S$ so that $\exists n\in\mathbb{N}_0$ $\lambda(n)=(k,m)$ which proves that λ is a surjection. Hence as λ is also injective it follows that $\lambda:\mathbb{N}_0\rightarrow\mathbb{N}_0\times\mathbb{N}_0$ is a bijection, proving that $\mathbb{N}_0\times\mathbb{N}_0$ is denumerable. □

Corollary 6.62. *If A,B are denumerable then $A\times B$ is denumerable*

Proof. As A,B are denumerable we have $\mathbb{N}_0\approx A$ and $\mathbb{N}_0\approx B$, proving by [theorem: 6.12] that $\mathbb{N}_0\times\mathbb{N}_0\approx A\times B$. Finally as $\mathbb{N}_0\approx\mathbb{N}_0\times\mathbb{N}_0$ it follows that that $\mathbb{N}_0\approx A\times B$. □

Corollary 6.63. *If $n\in\mathbb{N}_0\backslash\{0\}$ then $n\times\mathbb{N}_0$ is denumerable*

Proof. As $n = S_n \subseteq \mathbb{N}_0$ we have by [theorem: 1.48] that $n \times \mathbb{N}_0 \subseteq \mathbb{N}_0 \times \mathbb{N}_0$ so that by [theorem: 6.59]

$$n \times \mathbb{N}_0 \text{ is either finite or denumerable}$$

As $n \neq 0$ we have that $n \neq \varnothing$ so there exist a $m \in n$, define then

$$\beta: \mathbb{N}_0 \rightarrow \{m\} \times \mathbb{N}_0 \text{ by } \beta(i) = (m, i)$$

then we have:

injectivity. If $\beta(i) = \beta(i')$ then $(m, i) = (m, i')$ giving $i = i'$

surjectivity. If $(x, y) \in \{m\} \times \mathbb{N}_0$ then $x = m$ and $y \in \mathbb{N}_0$ so that $\beta(y) = (m, y) = (x, y)$

So $\beta: \mathbb{N}_0 \rightarrow \{m\} \times \mathbb{N}_0$ is a bijection proving that $\{m\} \times \mathbb{N}_0$ is denumerable. As $\{m\} \times \mathbb{N}_0 \subseteq n \times \mathbb{N}_0$ it follows by [theorem: 6.29] that $n \times \mathbb{N}_0$ is not finite so $n \times \mathbb{N}_0$ must be denumerable. □

Corollary 6.64. *If A is a non empty finite set and B a denumerable set then $A \times B$ and $B \times A$ are denumerable sets.*

Proof. As $A \neq \varnothing$ and finite there exist a $n \notin \mathbb{N}_0 \backslash \{0\}$ such that $n \approx A$, as B is denumerable $\mathbb{N}_0 \times B$ we have by [theorem: 6.12] that

$$n \times \mathbb{N}_0 \approx A \times B$$

which as $\mathbb{N}_0 \approx \mathbb{N}_0 \times \mathbb{N}_0$ [see corollary: 6.63] proves that $\mathbb{N}_0 \approx A \times B$, hence

$$A \times B \text{ is denumerable}$$

Define the function

$$\beta: A \times B \rightarrow B \times A \text{ by } \beta(x, y) = (y, x)$$

then we have

injectivity. If $\beta(x, y) = \beta(x', y')$ then $(y, x) = \beta(x, y) = \beta(x', y') = (y', x')$ so that $x = x' \wedge y = y'$ proving that $(x, y) = (x', y')$.

surjectivity. If $(x, y) \in B \times A$ we have that $(y, x) \in A \times B$ so that $\beta(y, x) = (x, y)$.

proving that

$$\beta: A \times B \rightarrow B \times A \text{ is a bijection}$$

hence $A \times B \approx B \times A$, which as $A \times B \approx \mathbb{N}_0$ proves that

$$B \times A \text{ is denumerable.}$$

□

Theorem 6.65. *If $\{A_i\}_{i \in I}$ is such that $I \neq \varnothing \wedge I$ is finite and $\forall i \in I$ A_i is denumerable then $\bigcup_{i \in I} A_i$ is denumerable. In other words the union of a finite family of denumerable sets is denumerable.*

Proof. As I is finite and non empty there exist $n_0 \in \mathbb{N}_0 \backslash \{0\}$ and a bijection $\beta: n_0 \rightarrow I$. Further as $\forall i \in I$ A_i is denumerable there exist a bijection $\alpha_i: \mathbb{N}_0 \rightarrow A_i$. Define now the function

$$g: n_0 \times \mathbb{N}_0 \rightarrow \bigcup_{i \in I} A_i \text{ by } g(n, m) = \alpha_{\beta(n)}(m)$$

Now if $y \in \bigcup_{i \in I} A_i$ there exist a $l \in I$ such that $y \in A_l$, as β is bijective there exists a $n \in n_0$ such that $\beta(n) = l$. As $\alpha_l: \mathbb{N}_0 \rightarrow A_l$ is a bijection there exist a $m \in \mathbb{N}_0$ such that $\alpha_l(m) = y$. So

$$g(n, m) = \alpha_{\beta(n)}(m) = \alpha_l(m) = y$$

proving that

$$g: n_0 \times \mathbb{N}_0 \rightarrow \bigcup_{i \in I} A_i \text{ is surjective}$$

Now by [theorem: 6.63] there exist a bijection $\gamma: \mathbb{N}_0 \rightarrow n_0 \times \mathbb{N}_0$ so that

$$g \circ \gamma: \mathbb{N}_0 \rightarrow \bigcup_{i \in I} A_i \text{ is surjective}$$

Using [theorem: 6.10] we have that $\bigcup_{i \in I} A_i \preccurlyeq \mathbb{N}_0$ which by [definition: 6.3] gives that $\exists E \subseteq \mathbb{N}_0$ such that $\bigcup_{i \in I} A_i \approx E$. Using [theorem: 6.59] we have that E is either finite or E is denumerable so that $\bigcup_{i \in I} A_i$ is either finite or denumerable. As $n_0 \neq 0 \Rightarrow 0 < n_0$ we have $0 \in S_{n_0} = n_0$, so that $\beta(0) \in I$, hence $A_{\beta(0)} \subseteq \bigcup_{i \in I} A_i$, which, as $A_{\beta(0)}$ is denumerable, proves by [theorem: 6.29] that $\bigcup_{i \in I} A_i$ is not finite. So we must have that $\bigcup_{i \in I} A_i$ is enumerable. □

Theorem 6.66. *If $\{A_i\}_{i \in I}$ is such that I is denumerable and $\forall i \in I$ A_i is denumerable then $\bigcup_{i \in I} A_i$ is denumerable. In other words every union of a denumerable family of denumerable sets is denumerable.*

Proof. As I is denumerable there exist a bijection $\beta: \mathbb{N}_0 \to I$. Further as $\forall i \in I$ A_i is denumerable there exist a bijection $\alpha_i: \mathbb{N}_0 \to A_i$. Define now the function

$$g: \mathbb{N}_0 \times \mathbb{N}_0 \to \bigcup_{i \in I} A_i \text{ by } g(n,m) = \alpha_{\beta(n)}(m)$$

Now if $y \in \bigcup_{i \in I} A_i$ there exist a $l \in I$ such that $y \in A_l$, as β is bijective there exists a $n \in \mathbb{N}_0$ such that $\beta(n) = l$. As $\alpha_l: \mathbb{N}_0 \to A_l$ is a bijection there exist a $m \in \mathbb{N}_0$ such that $\alpha_l(m) = y$. So

$$g(n,m) = \alpha_{\beta(n)}(m) = \alpha_l(m) = y$$

proving that

$$g: \mathbb{N}_0 \times \mathbb{N}_0 \to \bigcup_{i \in I} A_i \text{ is surjective}$$

Now by [theorem: 6.60] there exist a bijection $\gamma: \mathbb{N}_0 \to \mathbb{N}_0 \times \mathbb{N}_0$ so that

$$g \circ \gamma: \mathbb{N}_0 \to \bigcup_{i \in I} A_i \text{ is surjective}$$

Using [theorem: 6.10] we have that $\bigcup_{i \in I} A_i \preccurlyeq \mathbb{N}_0$ which by [definition: 6.3] gives that $\exists E \subseteq \mathbb{N}_0$ such that $\bigcup_{i \in I} A_i \approx \mathbb{N}_0$. Using [theorem: 6.59] we have that E is either finite or E is denumerable so that $\bigcup_{i \in I} A_i$ is either finite or denumerable. As $A_{\beta(0)} \subseteq \bigcup_{i \in I} A_i$ and $A_{\beta(0)}$ is denumerable it follows from [theorem: 6.29] that $\bigcup_{i \in I} A_i$ is not finite. So we must have that $\bigcup_{i \in I} A_i$ is enumerable. □

6.2.4 Countable Sets

Remember that a countable set is a set that is either finite or denumerable.

Theorem 6.67. *Every subset of a denumerable set is countable*

Proof. This follows from [theorem: 6.59] and the definition of countable sets. □

Theorem 6.68. *Every subset of a countable set is countable*

Proof. If A is countable then A is either denumerable or finite. If A is finite then by [theorem: 6.42] every subset of A is finite hence countable. If A is denumerable then by [theorem: 6.67] every subset of A is countable. □

Theorem 6.69. *Let A be a non empty set then the following are equivalent:*

1. *A is countable*
2. *There exists a surjection $\beta: \mathbb{N}_0 \to A$*
3. *There exists a injection $\alpha: A \to \mathbb{N}_0$*
4. *There exist a denumerable set B and a injection $\alpha: A \to B$*

Proof.

1 ⇒ 2. If A is countable then we have either:

A is finite. Then $\exists n \in \mathbb{N}_0$ and a bijection $\alpha: n = S_n \to A$. As $A \neq \varnothing$ there exist a $a \in A$, this allows us to define the function

$$\beta: \mathbb{N}_0 \to A \text{ where } \beta(i) = \begin{cases} \alpha(i) \text{ if } i < n \\ a \text{ if } n \leqslant i \end{cases}$$

If $y \in A$ then as α is surjective we have that $\exists i \in S_n = n$ such that $\alpha(i) = y$ so that $\beta(i) = \alpha(i) = y$ proving hat $\beta: \mathbb{N}_0 \to A$ is surjective.

A is denumerable. Then $\mathbb{N}_0 \approx A$ so there exist a bijection, hence surjection, $\beta: \mathbb{N}_0 \to A$.

2 ⇒ 3. Given that there exists a surjection $\beta: \mathbb{N}_0 \to A$ and $A \neq \varnothing$ we have by [theorem: 6.9] the existence of a injection $\alpha: A \to \mathbb{N}_0$.

3 ⇒ 4. As B is denumerable we have $\mathbb{N}_0 \approx B$ so there exist a bijection $\beta: \mathbb{N}_0 \rightarrow B$. by (3) there exist a injection $\alpha: A \rightarrow \mathbb{N}_0$, hence we have the injection $\beta \circ \alpha: A \rightarrow B$.

4 ⇒ 1. As B is denumerable there exist a bijection $\beta: B \rightarrow \mathbb{N}_0$ so that we have a injection $\beta \circ \alpha: A \rightarrow \mathbb{N}_0$. Using [theorem: 2.66] it follows that $\beta \circ \alpha: A \rightarrow (\beta \circ \alpha)(A) \subseteq \mathbb{N}_0$ is a bijection hence

$$A \approx (\beta \circ \alpha)(A) \subseteq \mathbb{N}_0$$

Using [theorem: 6.58] we have that $(\beta \circ \alpha)(A)$ is either finite or denumerable. If $(\beta \circ \alpha)(A)$ is finite then there exist a $n \in \mathbb{N}_0$ such that $n \approx (\beta \circ \alpha)(A)$, hence $n \approx A$ proving that A is finite, hence countable. If $(\beta \circ \alpha)(A)$ is denumerable then $\mathbb{N}_0 \approx (\beta \circ \alpha)(A)$ so that $\mathbb{N}_0 \approx A$ proving that A is denumerable hence countable. So we reach the conclusion that A is countable. □

Theorem 6.70. *If $\{A_i\}_{i \in I}$ is such that I is denumerable and $\forall i \in I$ A_i is countable then $\bigcup_{i \in I} A_i$ is countable. In other words every union of a denumerable family of countable sets is countable.*

Proof. As I is denumerable there exist a bijection $\beta: \mathbb{N}_0 \rightarrow I$. Further as $\forall i \in I$ A_i is denumerable there exist a surjection $\alpha_i: \mathbb{N}_0 \rightarrow A_i$ [see theorem: 6.69]. Define now the function

$$g: \mathbb{N}_0 \times \mathbb{N}_0 \rightarrow \bigcup_{i \in I} A_i \text{ by } g(n,m) = \alpha_{\beta(n)}(m)$$

Now if $y \in \bigcup_{i \in I} A_i$ there exist a $l \in I$ such that $y \in A_l$, as β is bijective there exists a $n \in \mathbb{N}_0$ such that $\beta(n) = l$. As $\alpha_l: \mathbb{N}_0 \rightarrow A_l$ is a surjection there exist a $m \in \mathbb{N}_0$ such that $\alpha_l(m) = y$. So

$$g(n,m) = \alpha_{\beta(n)}(m) = \alpha_l(m) = y$$

which proves that

$$g: \mathbb{N}_0 \times \mathbb{N}_0 \rightarrow \bigcup_{i \in I} A_i \text{ is surjective}$$

Now by [theorem: 6.60] there exist a bijection $\gamma: \mathbb{N}_0 \rightarrow \mathbb{N}_0 \times \mathbb{N}_0$ so that

$$g \circ \gamma: \mathbb{N}_0 \rightarrow \bigcup_{i \in I} A_i \text{ is surjective}$$

Using [theorem: 6.69] it follows that $\bigcup_{i \in I} A_i$ is countable. □

Theorem 6.71. *If A, B are countable sets then we have $A \times B$ is countable.*

Proof. For A, B we have the following possibilities:

***A* is finite and *B* is finite.** Then by [theorem: 6.40] $A \times B$ is finite hence countable.

***A* is finite and *B* is denumerable.** Then by [theorem: 6.64] $A \times B$ is denumerable hence countable.

***A* is denumerable and *B* is finite.** Then by [theorem: 6.64] $A \times B$ is denumerable hence countable.

***A* is denumerable and *B* is denumerable.** Then by [theorem: 6.62] $A \times B$ is denumerable hence countable. □

Lemma 6.72. *Let $n \in \mathbb{N}_0 \setminus \{0\}$ and $\{A_i\}_{i \in S_n}$ such that $\forall i \in S_n$ A_i is countable then $\prod_{i \in S_n} A_i$ is countable.*

Proof. We proof this by induction, so define

$$S = \left\{ n \in \{1, \ldots, \infty\} | \text{If } \{A_i\}_{i \in S_n} \text{ satisfies } \forall i \in S_n\ A_i \text{ is countable then } \prod_{i \in S_n} A_i \text{ is countable} \right\}$$

then we have:

$1 \in S$. As $S_1 = \{0\}$ we can use [example: 2.137] to find a bijection $\beta: A_0 \rightarrow \prod_{i \in \{0\}} A_i = \prod_{i \in S_1} A_i$ proving that $A_0 \approx \prod_{i \in S_1} A_i$, hence $\prod_{i \in S_1} A_i$ is countable [see theorem" 6.26].

$n \in S \Rightarrow n + 1 \in S$. Let $\{A_i\}_{i \in S_{n+1}}$ be such that $\forall i \in S_{n+1}$ A_i is countable. As $n \in S$ we have that $\prod_{i \in S_n} A_i$ is countable, so by [theorem: 6.71] we have that $(\prod_{i \in S_n} A_i) \times A_n$ is countable. Finally by [theorem: 6.55] we have $\prod_{i \in S_{n+1}} A_i \approx (\prod_{i \in S_n} A_i) \times A_n$ so that $\prod_{i \in S_{n+1}} A_i$ is countable [see theorem: 6.26]. Hence $n + 1 \in S$

Mathematical induction proves then that $S = \{1, \ldots, \infty\}$ proving the theorem. □

Theorem 6.73. *If I is non empty and finite and $\{A_i\}_{i \in I}$ such that $\forall i \in I$ A_i is countable then $\prod_{i \in I} A_i$ is countable.*

Proof. As I is finite and non empty there exists a $n \in \mathbb{N}_0 \backslash \{0\}$ such that $n \approx I$ hence there exist a bijection $f: n = S_n \rightarrow I$, Using [theorem: 2.143] there exists a bijection $\beta: \prod_{i \in I} A_i \rightarrow \prod_{i \in S_n} A_{f(i)}$ so that $\prod_{i \in S_n} A_{f(i)} \approx \prod_{i \in I} A_i$. Using the previous lemma [lemma: 6.72] $\prod_{i \in S_n} A_{f(i)}$ is countable, hence by [theorem: 6.26] $\prod_{i \in I} A_i$ is countable. □

6.3 Finite product of sets

We turn now our attention to the finite product of sets. Using the general definition of a product of a family of sets as is discussed in [definition: 2.135] we can define the finite product of sets.

Definition 6.74. *If $n \in \mathbb{N}$ and $\{A_i\}_{i \in \{1, \ldots, n\}} \subseteq B$ a finite family of sets then $\prod_{i=1}^n A_i$ is defined as*

$$\prod_{i=1}^{n} A_i = \prod_{i \in \{1, \ldots, n\}} A_i \text{ [see definition: 2.135]}$$

In other words

$$\prod_{i=1}^{n} A_i \underset{\text{[defintion: 2.135]}}{=} \left\{ f \mid f \in \left(\bigcup_{i \in \{1, \ldots, n\}} A_i \right)^{\{1, \ldots, n\}} \text{ where } \forall i \in \{1, \ldots, n\} \text{ we have } f(i) \in A_i \right\}$$

So if $x \in \prod_{i=1}^n A_i$ then $x: \{1, \ldots, n\} \rightarrow \bigcup_{i \in \{1, \ldots, n\}} A_i$ is a function with $\forall i \in \{1, \ldots, n\}$ $x(i) \in A_i$. As x_i is another notation for $x(i)$ we can introduce a new notation for $x \in \prod_{i=1}^n A_i$.

***Notation* 6.75.** *$x \in \prod_{i=1}^n A_i$ is noted as $(x_1, \ldots, x_n)$ which is equivalent with saying that $x: \{1, \ldots, n\} \rightarrow \bigcup_{i \in \{1, \ldots, n\}} A_i$ is a function with $\forall i \in \{1, \ldots, n\}$ $x_i = x(i) \in A_i$. Using this new notation we have the much shorter specification of $\prod_{i=1}^n A_i$.*

$$x \in \prod_{i=1}^{n} A_i \Leftrightarrow x = (x_1, \ldots, x_n) \text{ and } \forall i \in \{1, \ldots, n\} \text{ we have } x_i \in A_i$$

Using the above notation and definition we can also rephrase the projection operators [see definition: 2.146]

Definition 6.76. *If $n \in \mathbb{N}$ and $\{A_i\}_{i \in \{1, \ldots, n\}} \subseteq B$ a finite family of sets then $\forall i \in \{1, \ldots, n\}$*

$$\pi_i: \prod_{i=1}^{n} A_i \rightarrow A_i \text{ is defined by } \pi_i(x) = \pi_i(x_1, \ldots, x_i) = x_i$$

Theorem 6.77. *If $n \in \mathbb{N}$ and $\{A_i\}_{i \in \{1, \ldots, n\}} \subseteq B$ a finite family of sets then $\forall i \in \{1, \ldots, n\}$ we have that $\pi_i: \prod_{i=1}^n A_i \rightarrow A_i$ is a surjection*

Proof. This was proved in [theorem: 3.105] □

If we have only two sets involed we use the following notation to distinguish the product of sets from the cartesian product.

Definition 6.78. *Let A, B two sets then for the family $\{X_i\}_{i \in \{1,2\}} \subseteq A \bigcup B$ where $X_1 = A \wedge X_2 = B$ we have*

$$A \cdot B = \prod_{i \in \{1, \ldots, 2\}} X_i = \prod_{i \in \{1,2\}} X_i$$

so

$$A \cdot B = \left\{ f \mid f \in \left(A \bigcup B \right)^{\{1,2\}} \text{ where } f(1) \in A \wedge f(2) \in B \right\}$$

We use the notation $A \cdot B$ to make a distinction with the cartesian product $A \times B$, of course by [theorem: 2.138] we have that $A \cdot B$ and $A \times B$ are bijective.

Theorem 6.79. *Let A, B, C, D be sets then we have*

1. $(A \cdot B) \bigcap (C \cdot D) = (A \bigcap C) \cdot (B \bigcap D)$

2. *If $A \subseteq B$ then $A \cdot C \subseteq B \cdot C$*
3. *If $B \subseteq C$ then $A \cdot B \subseteq A \cdot C$*
4. $A \cdot (B \bigcup C) = (A \cdot B) \bigcup (A \cdot C)$
5. $(A \bigcup B) \cdot C = (A \cdot C) \bigcup (B \cdot C)$

Proof.

1. This follows from the definition of $\cdot$ and [theorem: 2.140]
2. This follows from the definition of $\cdot$ $C \subseteq C$ and [theorem: 2.139]
3. This follows from the definition of $\cdot$ $A \subseteq A$ and [theorem: 2.139]
4. As $B, C \subseteq B \bigcup C$ we have by (3) that $A \cdot B \subseteq A \cdot (B \bigcup C)$ and $A \cdot C \subseteq A \cdot (B \bigcup C)$ so that
$$(A \cdot B) \bigcup (A \cdot C) \subseteq A \cdot \left(B \bigcup C\right) \tag{6.27}$$
For the opposite inclusion let $x \in A \cdot (B \bigcup C)$ then $x: \{1, 2\} \rightarrow A \bigcup (B \bigcup C)$ is a function so that $x(1) \in A \wedge x(2) \in B \bigcup C$, so that either $x(1) \in A \wedge x(2) \in B$ or $x(1) \in A \wedge x(2) \in C$ proving that $x \in A \cdot B$ or $x \in A \cdot C$, hence $x \in (A \cdot B) \bigcup (A \cdot C)$, This proves that $A \cdot (B \bigcup C) \subseteq (A \cdot B) \bigcup (A \cdot C)$ which combined with [eq: 6.27] results in
$$A \cdot \left(B \bigcup C\right) = (A \cdot B) \bigcup (A \cdot C)$$
5. As $A, B \subseteq A \bigcup B$ we have by (2) that $A \cdot C \subseteq (A \bigcup B) \cdot C$ and $B \cdot C \subseteq (A \bigcup B) \cdot C$ so that
$$(A \cdot C) \bigcup (B \cdot C) \subseteq \left(A \bigcup B\right) \cdot C \tag{6.28}$$
For the opposite inclusion let $x \in (A \bigcup B) \cdot C$ then $x: \{1, 2\} \rightarrow (A \bigcup B) \bigcup C$ is a function so that $x(1) \in (A \bigcup B) \wedge x(2) \in C$, so that either $x(1) \in A \wedge x(2) \in C$ or $x(1) \in B \wedge x(2) \in C$ proving that $x \in A \cdot C$ or $x \in B \cdot C$, hence $x \in (A \cdot C) \bigcup (B \cdot C)$, This proves that $(A \bigcup B) \cdot C \subseteq (A \cdot C) \bigcup (B \cdot C)$ which combined with [eq: 6.28] results in
$$\left(A \bigcup B\right) \cdot C = (A \cdot C) \bigcup (B \cdot C)$$ □

The following theorem will be useful in induction arguments.

Theorem 6.80. *Let $n \in \mathbb{N}$ and $\{X_i\}_{i \in \{1, \ldots, n\}} \subseteq B$ a finite family of sets then $\prod_{i \in \{1, \ldots, n+1\}} X_i$ is bijective with $\left(\prod_{i \in \{1, \ldots, n\}} X_i\right) \times X_{n+1}$*

Proof. Define
$$\beta: \prod_{i \in \{1, \ldots, n+1\}} X_i \rightarrow \left(\prod_{i \in \{1, \ldots, n\}} X_i\right) \times X_{n+1} \text{ by } \beta(x) = (x_{|\{1, \ldots, n\}}, x(n+1))$$
then we have:

injectiviy. If $\beta(x) = \beta(y)$ then $(x_{|\{1, \ldots, n\}}, x(n+1)) = (y_{|\{1, \ldots, n\}}, y(n+1))$, hence $x_{|\{1, \ldots, n\}} = y_{|\{1, \ldots, n\}}$ and $x(n+1) = y(n+1)$ or $\forall i \in \{1, \ldots, n\}$
$$x(i) = x_{|\{1, \ldots, n\}}(x) = y_{|\{1, \ldots, n\}}(i) = y(i) \wedge x(n+1) = y(n+1)$$
proving that $x = y$.

surjectivity. If $z \in \left(\prod_{i \in \{1, \ldots, n\}} X_i\right) \times X_{n+1}$ then $z = (z_1, z_2)$ where $z_1 \in \prod_{I \in \{1, \ldots, n\}} X_i$ and $z_2 \in X_{n+1}$. Define the function
$$x: \{1, \ldots, n+1\} \rightarrow \bigcup_{i \in \{1, \ldots, n\}} X_i \text{ by } x(i) = \left\{ \begin{array}{l} z_2 \in X_{n+1} \text{ if } i = n+1 \\ z_1(i) \in X_i \text{ if } i \in \{1, \ldots, n\} \end{array} \right. \in X_i$$
then $x \in \prod_{i \in \{1, \ldots, n\}} X_i$. Further $\forall i \in \{1, \ldots, n\}$ $x_{|\{1, \ldots, n\}}(i) = x(i) = z(i)$ proving that $x_{|\{1, \ldots, n\}} = z_1$ so that
$$\beta(x) = (x_{|\{1, \ldots, n\}}, x(n+1)) = (z_1, z_2) = z$$
proving surjectivity. □

Next we consider the special case where for $\forall i \in \{1, \ldots, n\}$ $A_i = A$ [see also theorem: 2.142]

Definition 6.81. *Let $n\in\mathbb{N}$, A a set then A^n is defined by*

$$A^n=\prod_{i=1}^n A_i \text{ where } \{A_i\}_{i\in\{1,\ldots,n\}}\subseteq\{A\} \text{ is defined by } C_A\colon\{1,\ldots,n\}\rightarrow\{A\}$$

so that $\forall i\in\{1,\ldots,n\}$ $A_i=C_A(i)=A$. So as $A=\bigcup_{i\in\{1,\ldots,n\}}A_i$ we have that

$$A^n=\{f|f\in A^{\{1,\ldots,n\}} \text{ where } \forall i\in\{1,\ldots,n\} \text{ we have } f(i)\in A\}=A^{\{1,\ldots,n\}}$$

Using [notation: 6.75] we can write also

$$x\in A^n\Leftrightarrow x=(x_1,\ldots,x_n) \text{ and } \forall i\in\{1,\ldots,n\} \text{ we have } x_i\in A$$

Theorem 6.82. *Let $n\in\mathbb{N}$, A a set then $A^n=A^{\{1,\ldots,n\}}\underset{\text{def}}{=}\{f|f\colon\{1,\ldots,n\}\rightarrow A \text{ is a function}\}$*

Proof. This follows from [theorem: 2.142] □

Theorem 6.83. *If A is a set then*

$$\beta\colon A\rightarrow A^1 \text{ defined by } \beta(a)=\{(1,a)\}$$

is a bijection.

Proof. If $a\in A$ then $(1,a)\subseteq\{1\}\times A=\{1,...,1\}\times A$ so that $\beta(a)\colon\{1,...,1\}\rightarrow A$ is indeed a function, proving that $\beta(A)\in A^1$. Further we have

injectivity. If $\beta(a)=\beta(b)$ then $\{(1,a)\}=\{(1,b)\}\Rightarrow(1,a)=(1,b)\Rightarrow a=b$

surjectivity. If $f\in A^1=A^{\{1\}}$ then $f\colon\{1\}\rightarrow A$ is a function, so that $f\subseteq\{1\}\times A$, hence $\exists a\in A$ such that $f=\{(1,a)\}=\beta(a)$ □

Once we have defined the finite power of a set we can consider the finite power of semi-groups

Theorem 6.84. *Let $\langle G,+\rangle$ be a semi-group withe neutral element 0, $n\in\mathbb{N}$ then if we define*

$$+^n\colon G^n\times G^n\rightarrow G^n \text{ by } x+^n y\colon\{1,\ldots,n\}\rightarrow G \text{ where } (x+^n y)(i)=x(i)+y(i)$$

$$0^n\colon\{1,\ldots,n\}\rightarrow G \quad \text{by } 0^n(i)=0$$

or using [notation: 6.75]

$$\forall x,y\in G^n \text{ we have } x+^n y=(x_1+y_1,\ldots,x_n+y_n) \text{ where } (x+^n y)_i=x_i+y_i$$

$$0^n\in G^n \text{ is } \left(\underbrace{0,\ldots,0}_{n}\right) \text{ so that } (0^n)_i=0$$

then we have:

1. *$\langle G^n,+^n\rangle$ is a semi-group with neutral element 0^n.*
2. *If $\langle G,+\rangle$ is Abelian semi-group then $\langle G^n,+^n\rangle$ is Abelian.*
3. *If $\langle G,+\rangle$ is a group where $-x$ is the inverse element of $x\in G$ then $\langle G^n,+^n\rangle$ is a group where for $x\in G^n$ the inverse element*

$$\sim x\colon\{1,\ldots,n\}\rightarrow G \text{ is defined by } (\sim x)(i)=-x(i)$$

or using [notation: 6.75]

$$\sim x=\sim(x_1,\ldots,x_n)=(-x_1,\ldots,-x_n) \text{ so that } (\sim x)_i=-x_i$$

Proof. This was actuall proved for a more general case in [theorem: 4.26]. □

Note 6.85. As usual, for the rest of this text we reduce the number of symbols by using $+,-,0$ instead of $+^n,\sim,0^n$.

Chapter 7
The integer numbers

In this chapter we will introduce the set of integers and embed the natural numbers in it. Just as with $\mathbb{N}_0$ we will introduce a order relation, a sum operator, a product operator, neutral elements for addition and multiplication as well as inverse elements for the integers. If we would use different symbols to note these, we introduce a lot of excessive notation clutter. So we use the same symbols for the natural numbers and the integers, and use context to determine the meaning of the symbols involved. A practice also used in programming languages [where it is called 'overloading', the following table should help you in determining the meaning of the different symbols based on the context of there usage.

Context	Expression	Operator
$n,m\in\mathbb{N}_0$	n+m	sum in $\langle\mathbb{N}_0,+\rangle$
$n,m\in\mathbb{N}_0$	$n\cdot m$	product in $\langle\mathbb{N}_0,\cdot\rangle$
$n,m\in\mathbb{N}_0$	$n\leqslant m$	order in $\langle\mathbb{N}_0,\leqslant\rangle$
$n,m\in\mathbb{N}_0$	$n<m$	strict order in $\langle\mathbb{N}_0,\leqslant\rangle$
$n,m\in\mathbb{N}_0$	$n-m$	subtraction in $\langle\mathbb{N}_0,+\rangle$
$n\in\mathbb{N}_0$	$n+0$ or $0+n$	neutral element in $\langle\mathbb{N}_0,+\rangle$
$n\in\mathbb{N}_0$	$n\cdot 1$ or $1\cdot n$	neutral element in $\langle\mathbb{N}_0,\cdot\rangle$
$n\in\mathbb{N}_0$	$-n$	inverse element in $\langle\mathbb{N}_0,+\rangle$
$n,m\in\mathbb{Z}$	n+m	sum in $\langle\mathbb{Z},+\rangle$
$n,m\in\mathbb{Z}$	$n\cdot m$	product in $\langle\mathbb{Z},\cdot\rangle$
$n,m\in\mathbb{Z}$	$n\leqslant m$	order in $\langle\mathbb{Z}\leqslant\rangle$
$n,m\in\mathbb{Z}$	$n<m$	strict order in $\langle\mathbb{Z},\leqslant\rangle$
$n,m\in\mathbb{Z}$	$n-m$	subtraction in $\langle\mathbb{Z},-\rangle$
$n\in\mathbb{Z}$	$n+0$ or $0+n$	neutral element in $\langle\mathbb{Z},+\rangle$
$n\in\mathbb{Z}$	$n\cdot 1$ or $1\cdot n$	neutral element in $\langle\mathbb{Z},\cdot\rangle$
$n\in\mathbb{Z}$	$-n$	inverse element in $\langle\mathbb{Z},+\rangle$

7.1 Definition and arithmetic

One major defect of $\mathbb{N}_0$ is that $n-m$, defined to be the unique natural number such that $(n-m)+m=n$, is only defined for $m\leqslant n$. If this limitation did not exist then we can easily find a inverse for a number n, just take $-n=0-n$, then $(-n)+n=(0-n)+n=0$. The purpose of this chapter is to define a new set of numbers, the set of integers, that does not have this defect. One strategy is adding to the set of natural numbers the set of numbers of the form $n-m$ where $n<m$. The numbers of the form $n-m$ where $m\leqslant n$ is then the set of non negative integers and represent the set of natural numbers and the numbers $n-m$ where $n<m$ forms the set of negative numbers. Of course the expression $n-m$ is only a formal expression and is not defined in $\mathbb{N}_0$ if $n<m$ hence we use pairs to of natural numbers to define integers. So a integer is a pair (n,m) where $n,m\in\mathbb{N}_0$ that must be interpreted as the **formal** expression $n-m$ if $n<m$ and the **real** expression $n-m$ if $m\leqslant n$. However we encounter then another problem, the representations are not **unique**. For example we know that for the natural number 3 we have that $3=3-0=4-1=5-2=6-3,\ldots$, so that $(3,0),(4,1),(5,2),(6,3),\ldots$, must all represent the same number 3. How can we see if two representations of a natural number are the same? If (n,m) and (n',m') are representations of the same **natural** number then $m\leqslant n$ and $m'\leqslant n'$ and we must have

$$\begin{aligned} n-m=n'-m' &\Rightarrow (n-m)+m=(n'-m')+m \\ &\Rightarrow n=(n'-m')+m \\ &\Rightarrow n=m+(n'-m') \\ &\Rightarrow n+m'=(m+(n'-m'))+m' \\ &\Rightarrow n+m'=m+((n-\!\!-m')+m') \\ &\Rightarrow n+m'=m+n' \end{aligned}$$

So (n,m) and (n',m') with $m\leqslant n$ and $m'\leqslant n'$ represent the same **natural** number if $n+m'=m+n'$. As we don't use subtraction anymore we can extends this test also to the cases where $n<m$ or $n'<m$. So we say that two representations (n,m) and (n',m') represent the same **integer** if $n+m'=m+n'$. Hence if we define the relation $(n,m)\sim(n',m')$ iff $n+m'=m+n'$ and prove that is a equivalence relation then the equivalence classes will be our integers.

Theorem 7.1. *The relation* $\sim\subseteq(\mathbb{N}_0\times\mathbb{N}_0)\times(\mathbb{N}_0\times\mathbb{N}_0)$ *defined by*

$$\sim=\{((n,m),(n',m'))|n+m'=m+n'\}$$

is a equivalence relation.

Proof.

reflexivity. If $(n,m)\in\mathbb{N}_0\times\mathbb{N}_0$ then $n+m\underset{\text{[theorem: 5.33]}}{=}m+n$ so that $(n,m)\sim(n,m)$.

symmetry. If $(n,m)\sim(n',m')$ then $n+m'=m+n'\underset{\text{[theorem: 5.33]}}{\Rightarrow}n'+m=m'+n$ so that $(n',m')\sim(n,m)$.

transitivity. We have

$$\begin{aligned}(n,m)\sim(n',m')\wedge(n',m')\sim(n'',m'') &\Rightarrow n+m'=m+n'\wedge n'+m''=m'+n''\\ &\Rightarrow (n+m')+(n'+m'')=(m+n')+(m'+n'')\\ &\Rightarrow (n+m'')+(m'+n')=(m+n'')+(n'+m')\\ &\Rightarrow (n+m'')+(n'+m')=(m+n'')+(n'+m')\\ &\Rightarrow (n+m'')=(m+n'')\end{aligned}$$

so that $(n,m)\sim(n'',m'')$. □

Next we define the set of integers.

Definition 7.2. *The set of integers* $\mathbb{Z}$ *is defined by* $(\mathbb{N}_0\times\mathbb{N}_0)/\sim$ *or in other words*

$$\mathbb{Z}=\{\sim[(n,m)]|(n,m)\in\mathbb{N}_0\times\mathbb{N}_0\}$$

Theorem 7.3. *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. $n+(m+k)=(n+m)+k=(m+n)+k=m+(n+k)$ so that $(n,m)\sim(n+k,m+k)$. Hence by [theorem: 3.11] $\sim[(n,m)]=\sim[(n+k,m+k)]$. □

Theorem 7.4. *If* $\sim[(n,m)],\sim[(r,s)],\sim[(n',m')]$ *and* $\sim[(r',s')]$ *are elements of* $\mathbb{Z}$ *such that* $\sim[(n,m)]=\sim[(n',m')]$ *and* $\sim[(r,s)]=\sim[(r',s')]$ *then* $\sim[(n+r,m+s)]=\sim[(n'+r',m'+s')]$

Proof. As $\sim[(n,m)]=\sim[(n',m')]\wedge\sim[(r,s)]=\sim[(r',s')]$ we have

$$n+m'=m+n'\wedge r+s'=s+r' \tag{7.1}$$

then

$$\begin{aligned}(n+r)+(m'+s') &= (n+m')+(r+s')\\ &\underset{\text{[eq: 7.1]}}{=} (m+n')+(s+r')\\ &= (m+s)+(n'+r')\end{aligned}$$

so that $(n+r,m+s)\sim(n'+r',m'+s')$ proving that

$$\sim[(n+r,m+s)]=\sim[(n'+r',m'+s')]$$ □

The above theorem ensure that the following definition is well defined:

Definition 7.5. *The sum operator* $+:\mathbb{Z}\times\mathbb{Z}\to\mathbb{Z}$ *is defined by*

$$\sim[(n,m)]+\sim[(r,s)]=\sim[(n+r,m+s)]$$

Lemma 7.6. *If* $n\in\mathbb{N}_0$ *then* $\sim[(n,n)]=\sim[(0,0)]$

Proof. As $n+0=n+0$ we have $(n,n)\sim(0,0)$ so that $\sim[(n,n)]=\sim[(0,0)]$. $\square$

Theorem 7.7. ($\langle\mathbb{Z},+\rangle$ is a Abelian group) *so*

- ***Associativity.*** *$\forall n,m,k\in\mathbb{Z}$ we have $(n+m)+k=n+(m+k)$.*
- ***Neutral element.*** *$\forall n\in\mathbb{Z}$ we have that $n+0=0+n$ where $0=\sim[(0,0)]$.*
- ***Inverse element.*** *$\forall n\in\mathbb{Z}$ there exist a inverse element $-n$ such that $(-n)+n=0=n+(-n)$. More specifically if $x=\sim[(n,m)]$ then $-x=[(m,n)]$.*
- ***Commutativity.*** *$\forall n,m\in\mathbb{N}_0$ we have $n+m=m+n$.*

Proof.

Associativity. If $n=\sim[(n_1,m_1)]$,$m=\sim[(n_2,m_2)]$ and $k=\sim[(n_3,m_3)]$ then we have

$$\begin{aligned}(n+m)+k &= (\sim[(n_1,m_1)]+\sim[(n_2,m_2)])+\sim[(n_3,m_3)]\\ &= \sim[(n_1+n_2,m_1+m_2)]+\sim[(n_3,m_3)]\\ &= [\sim((n_1+n_2)+n_3,(m_1+m_2)+m_3)]\\ &= \sim[(n_1+(n_2+n_3),m_1+(m_2+m_3))]\\ &= \sim[(n_1,m_1)]+\sim[(n_2+n_3,m_2+m_3)]\\ &= \sim[(n_1,m_1)]+(\sim[(n_2,m_2)]+\sim[(n_3,m_3)])\\ &= n+(m+k)\end{aligned}$$

Commutativity. If $n=\sim[(n_1,m_1)]$ and $m=\sim[(n_2,m_2)]$ then

$$\begin{aligned}\sim[(n_1,m_1)]+\sim[(n_2,m_2)] &= \sim[(n_1+n_2,m_1+m_2)]\\ &= \sim[(n_2+n_1,m_2+m_1)]\\ &= \sim[(n_2,m_2)]+\sim[(n_1,m_1)]\end{aligned}$$

Neutral element. If $k=\sim[(n,m)]\in\mathbb{Z}$ then

$$\begin{aligned}0+k &\underset{\text{commutativity}}{=} k+0\\ &= \sim[(n,m)]+\sim[(0,0)]\\ &= \sim[(n+0,m+0)]\\ &= \sim[(n,m)]\\ &= k\end{aligned}$$

Inverse element. If $k=\sim[(n,m)]$

$$\begin{aligned}k+(-k) &\underset{\text{commutativity}}{=} (-k)+k\\ &= \sim[(m,n)]+\sim[(n,m)]\\ &= \sim[(m+n,n+m)]\\ &= \sim[(n+m,n+m)]\\ &\underset{[\text{theorem: }7.6]}{=} \sim[(0,0)]\end{aligned}$$

$\square$

The following introduce the difference operator that is now defined for all integers.

Definition 7.8. *Let $n,m\in\mathbb{Z}_0^+$ then we have $n-m=n+(-m)$*

Now to define multiplication in $\mathbb{Z}$, note that (n,m) is to be interpreted as $n-m$. So if $x=(n,m)$ and $y=(r,s)$ are two integers then $x\cdot y=(n,m)\cdot(r,s)$ is to be interpreted as the formal expression $(n-m)\cdot(r-s)$. Which if we formally evaluate it gives

$$\begin{aligned}(n-m)\cdot(r-s) &= n\cdot r-n\cdot s-m\cdot r+m\cdot s\\ &= n\cdot r+m\cdot s-(m\cdot r+n\cdot s)\end{aligned}$$

which suggest us that $(n,m)\cdot(r,s)$ should be equal to $(n\cdot r+m\cdot s, m\cdot r+n\cdot s)$, of course this is based on the representation of x and y. The next theorem proves that this product is independent of the representation, allowing us to define the product.

Theorem 7.9. *If $\sim[(n,m)], \sim[(r,s)], \sim[(n',m')]$ and $\sim[(r',s')]$ are elements of $\mathbb{Z}$ such that $\sim[(n,m)]=\sim[(n',m')]$ and $\sim[(r,s)]=\sim[(r',s')]$ then*

$$\sim[(n\cdot r+m\cdot s, m\cdot r+n\cdot s)]=\sim[(n'\cdot r'+m'\cdot s', m'\cdot r'+n'\cdot s')]$$

Proof. As $\sim[(n,m)]=\sim[(n',m')]\wedge\sim[(r,s)]=\sim[(r',s')]$ we have

$$n+m'=m+n'\wedge r+s'=s+r' \tag{7.2}$$

So we have

$$\begin{array}{rcl} n\cdot r+m'\cdot r & = & (n+m')\cdot r\\ & \underset{[\text{eq: } 7.2]}{=} & (m+n')\cdot r\\ & = & m\cdot r+n'\cdot r\\ m\cdot s+n'\cdot s & = & (m+n')\cdot s\\ & \underset{[\text{eq: } 7.2]}{=} & (n+m')\cdot s\\ & = & n\cdot s+m'\cdot s\\ m'\cdot s+m'\cdot r' & = & m'\cdot(s+r')\\ & \underset{[\text{eq: } 7.2]}{=} & m'\cdot(r+s')\\ & = & m'\cdot r+m'\cdot s'\\ n'\cdot r+n'\cdot s' & = & n'\cdot(r+s')\\ & \underset{[\text{eq: } 7.2]}{=} & n'\cdot(s+r')\\ & = & n'\cdot s+n'\cdot r' \end{array}$$

so after summing (underlining common terms).

$$\begin{array}{l} n\cdot r+\underbrace{m'\cdot r}_{1}+m\cdot s+\underbrace{n'\cdot s}_{2}+\underbrace{m'\cdot s}_{3}+m'\cdot r'+\underbrace{n'\cdot r}_{4}+n'\cdot s' \;=\\ m\cdot r+\underbrace{n'\cdot r}_{4}+n\cdot s+\underbrace{m'\cdot s}_{3}+\underbrace{m'\cdot r}_{1}+m'\cdot s'+\underbrace{n'\cdot s}_{2}+n'\cdot r' \end{array}$$

Using [theorem: 5.43] to eliminate common terms in the above gives:

$$n\cdot r+m\cdot s+m'\cdot r'+n'\cdot s' \;=\; m\cdot r+n\cdot s+m'\cdot s'+n'\cdot r'$$

So that

$$(n\cdot r+m\cdot s, m\cdot r+n\cdot s) \;\sim\; (n'\cdot r'+m'\cdot s', m'\cdot r'+n'\cdot s')$$

Hence

$$\sim[(n\cdot r+m\cdot s, m\cdot r+n\cdot s)]=\sim[(n'\cdot r'+m'\cdot s', m'\cdot r'+n'\cdot s')]$$ □

The above theorem ensures that the following definition is sensible.

Definition 7.10. *The multiplication operator $\cdot:\mathbb{Z}\times\mathbb{Z}\rightarrow\mathbb{Z}$ is defined by*

$$\sim[(n,m)]\cdot\sim[(r,s)]=\sim[(n\cdot r+m\cdot s, m\cdot r+n\cdot s)]$$

Theorem 7.11. *$\langle\mathbb{Z},+,\cdot\rangle$ is a* ***integral domain*** *[definition: 4.34], more specific:*

1. *$\langle\mathbb{Z},+\rangle$ is a* ***Abelian*** *group [see: 7.7]*
2. *$\langle\mathbb{Z},\cdot\rangle$ is a* ***Abelian semi-group****.*

   Associativity. *$\forall n,m,k\in\mathbb{Z}$ we have $n\cdot(m\cdot k)=(n\cdot m)\cdot k$*

   Neutral Element. *There exist a $1=\sim[(1,0)]$ such that $\forall n\in\mathbb{N}_0$ we have $n\cdot 1=n=1\cdot n$.*

   Commutativity. *$\forall n,m\in\mathbb{Z}$ we have $n\cdot m=m\cdot n$.*

3. Further we have:

Distributivity. *$\forall n,m,k\in\mathbb{Z}$ we have $n\cdot(m+k)=n\cdot m+n\cdot k$*

There does not exist a zero divisor. *If $n,m\in\mathbb{Z}$ is such that $n\cdot m=0\Rightarrow n=0\vee m=0$*

Proof.

1. This is already proved in [theorem: 7.7].
2.

Commutativity. If $\sim[(n,m)],\sim[(r,s)]\in\mathbb{Z}$ we have

$$\begin{array}{rcl}\sim[(n,m)]\cdot\sim[(r,s)] & = & \sim[(n\cdot r+m\cdot s,m\cdot r+n\cdot s)]\\ & = & \sim[(r\cdot n+s\cdot m,s\cdot n+r\cdot m)]\\ & = & \sim[(r,s)]\cdot\sim[(n,m)]\end{array}$$

Associativity. Let $\sim[(a,b)],\sim[(c,d)],\sim[(e,f)]\in\mathbb{Z}$ then

$$\begin{array}{rl}\sim[(a,b)]\cdot(\sim[(c,d)]\cdot\sim[(e,f)]) & =\\ \sim[(a,b)]\cdot(\sim[(c\cdot e+d\cdot f,d\cdot e+c\cdot f)]) & =\\ \sim[(a\cdot(c\cdot e+d\cdot f)+b\cdot(d\cdot e+c\cdot f),b\cdot(c\cdot e+d\cdot f)+a\cdot(d\cdot e+c\cdot f))] & =\\ \sim\left[\left(\underbrace{a\cdot(c\cdot e)}_{1}+\underbrace{a\cdot(d\cdot f)}_{2}+\underbrace{b\cdot(d\cdot e)}_{3}+\underbrace{b\cdot(c\cdot f)}_{4},\underbrace{b\cdot(c\cdot e)}_{5}+\underbrace{b\cdot(d\cdot f)}_{6}+\underbrace{a\cdot(d\cdot e)}_{7}+\underbrace{a\cdot(c\cdot f)}_{8}\right)\right] & =\\ \sim\left[\left(\underbrace{(a\cdot c)\cdot e}_{1}+\underbrace{(b\cdot d)\cdot e}_{3}+\underbrace{(b\cdot c)\cdot f}_{4}+\underbrace{(a\cdot d)\cdot f}_{2},\underbrace{(b\cdot c)\cdot e}_{5}+\underbrace{(a\cdot d)\cdot e}_{7}+\underbrace{(a\cdot c)\cdot f}_{8}+\underbrace{(b\cdot d)\cdot f}_{6}\right)\right] & \\ \sim[(a\cdot c+b\cdot d)\cdot e+(b\cdot c+a\cdot d)\cdot f,(b\cdot c+a\cdot d)\cdot e+(a\cdot c+b\cdot d)\cdot f] & =\\ \sim[(a\cdot c+b\cdot d,b\cdot c+a\cdot d)]\cdot\backsim[(e,f)] & =\\ (\sim[(a,b)]\cdot\sim[(c,d)])\cdot\sim[(e,f)] & \end{array}$$

Neutral element. If $n=\sim[(n,m)]\in\mathbb{Z}$ then we have

$$\begin{array}{rcl}n\cdot 1 & \underset{\text{commutativity}}{=} & 1\cdot n\\ & = & \sim[(1,0)]\cdot\sim[(n,m)]\\ & = & \sim[(1\cdot n+0\cdot m,0\cdot n+1\cdot m)]\\ & = & \sim(n,m)\end{array}$$

3. Further we have:

Distributivity. If $\sim[(a,b)],\sim[(c,d)],\sim[(e,f)]\in\mathbb{Z}$ then

$$\begin{array}{rl}\sim[(a,b)]\cdot(\sim[(c,d)]+\sim[(e,f)]) & =\\ \sim[(a,b)]\cdot\sim[(c+e,d+f)] & =\\ \sim[(a\cdot(c+e)+b\cdot(d+f),b\cdot(c+e)+a\cdot(d+f))] & =\\ \sim\left[\left(\underbrace{a\cdot c}_{1}+\underbrace{a\cdot e}_{2}+\underbrace{b\cdot d}_{3}+\underbrace{b\cdot f}_{4},\underbrace{b\cdot c}_{5}+\underbrace{b\cdot e}_{6}+\underbrace{a\cdot d}_{7}+\underbrace{a\cdot f}_{8}\right)\right] & =\\ \sim\left[\left(\underbrace{a\cdot c}_{1}+\underbrace{b\cdot d}_{3}+\underbrace{a\cdot e}_{2}+\underbrace{b\cdot f}_{4},\underbrace{b\cdot c}_{5}+\underbrace{a\cdot d}_{7}+\underbrace{b\cdot e}_{6}+\underbrace{a\cdot f}_{8}\right)\right] & =\\ \sim[(a\cdot c+b\cdot d,b\cdot c+a\cdot d)]+\sim[(a\cdot e+b\cdot f,b\cdot e+a\cdot f)] & =\\ \sim[(a,b)]\cdot\sim[(c,d)]+\sim[(a,b)]\cdot\sim[(e,f)] & =\end{array}$$

There does not exist a zero divisor. Let $n=\sim\{(a,b)\}$, $m=\sim[(c,d)]$ such that $n\cdot m=0$ then

$$\sim[(a,b)]\cdot\sim[(c,d)]=\sim[(a\cdot c+b\cdot d,b\cdot c+a\cdot d)]=\sim[(0,0)]$$

so we have that $(a\cdot c+b\cdot d)+0=(b\cdot c+a\cdot d)+0$ giving

$$a\cdot c+b\cdot d=b\cdot c+a\cdot d \tag{7.3}$$

Assume that $n \neq 0$ then $\sim[(a,b)] \neq \sim[(0,0)]$ so that $a+0 \neq b+0$ so that $a \neq b$, hence we have the following cases to consider:

a < b. Then using [theorem: 5.60] there exists a $k \in \mathbb{N}_0 \setminus \{0\}$ such that $a+k=b$, so substituting this in [eq: 7.3] gives

$$\begin{array}{rl} a \cdot c + (a+k) \cdot d = (a+k) \cdot c + a \cdot d & \Rightarrow \\ \underbrace{a \cdot c}_{1} + \underbrace{a \cdot d}_{2} + k \cdot d = \underbrace{a \cdot c}_{1} + k \cdot c + \underbrace{a \cdot d}_{2} & \Rightarrow \\ k \cdot d = k \cdot c & \underset{k \neq 0 \wedge [\text{theorem: } 5.77]}{\Rightarrow} \\ d = c & \end{array}$$

So $m = \sim[(c,d)] = \sim[(d,d)] \underset{[\text{theorem: } 7.6]}{=} \sim[(0,0)] = 0$.

b < a. Then using [theorem: 5.60] there exists a $k \in \mathbb{N}_0 \setminus \{0\}$ such that $b+k=a$, so substituting this in [eq: 7.3] gives

$$\begin{array}{rl} (b+k) \cdot c + b \cdot d = b \cdot c + (b+k) \cdot d & \Rightarrow \\ \underbrace{b \cdot c}_{1} + k \cdot c + \underbrace{b \cdot d}_{2} = \underbrace{b \cdot c}_{1} + \underbrace{b \cdot d}_{2} + k \cdot d & \Rightarrow \\ k \cdot c = k \cdot d & \underset{k \neq 0 \wedge [\text{theorem: } 5.77]}{\Rightarrow} \\ c = d & \end{array}$$

So $m = \sim[(c,d)] = \sim[(d,d)] \underset{[\text{theorem: } 7.6]}{=} [(0,0)] = 0$.

So if $n \cdot m = 0$ then we have either $n \neq 0$ but then $m = 0$ or $n = 0$ proving that $n \cdot m = 0 \Rightarrow n = 0 \vee m = 0$. □

Example 7.12. $1+1=2$ where $2 = \sim[(2,0)]$

Proof. $1+1 = \sim[(1,0)] + \sim[(1,0)] = \sim[(1+1, 0+0)] = \sim[(1+1, 0)] \underset{[\text{example: } 5.28]}{=} \sim[(2,0)] = 2$ □

Lemma 7.13. *$\forall n \in \mathbb{N}_0 \setminus \{0\}$ we have that $\sim[(n,0)] \neq 0$*

Proof. We prove this by contradiction so assume that $\sim[(n,0)] = 0 = \sim[(0,0)]$ then $n+0=0 \Rightarrow n=0$ contradicting $n \in \mathbb{N}_0 \setminus \{0\}$. So $\sim[(n,0)] \neq 0$. □

Corollary 7.14. *$\forall z \in \mathbb{Z}$ such that $z = -z$ we have $z = 0$*

Proof. If $z = -z$ we have that $z + z = (-z) + z = 0$. So $(1+1) \cdot z = z \cdot 1 + z \cdot 1 = z + z = 0$, hence

$$(1+1) \cdot z = 0$$

As $1+1 = \sim[(1,0)] + \sim[(1,0)] = \sim[(2,0)]$ and $2 \neq 0$ we have by [lemma: 7.13] that $1+1 \neq 0$, using [theorem: 7.11] on the above proves then that $z = 0$. □

Theorem 7.15. *Let $n,k,r \in \mathbb{Z}$ with $r \neq 0$ then $n \cdot r = k \cdot r$ implies $n = k$.*

Proof.

$$\begin{array}{rcl} n \cdot r = k \cdot r & \Rightarrow & n \cdot r + (-(k \cdot r)) = (k \cdot r) + (-(k \cdot r)) \\ & \Rightarrow & n \cdot r + (-(k \cdot r)) = 0 \\ & \underset{[\text{theorem: } 4.40]}{\Rightarrow} & n \cdot r + (-k) \cdot r = 0 \\ & \Rightarrow & (n + (-k)) \cdot r = 0 \end{array}$$

As by [theorem: 7.11] $\langle \mathbb{Z}, +, \cdot \rangle$ is a integral domain and $r \neq 0$ we have $n + (-k) = 0$ so that $(n + (-k)) + k = 0 + k$ or $n + ((-k) + k) = k$ proving $n = k$. □

7.2 Order relation on the set of integers

First we define the set of non negative integers.

Definition 7.16. $\mathbb{Z}_0^+ = \{\sim[(n,0)] | n \in \mathbb{N}_0\} \subseteq \mathbb{Z}$

We have the following properties for the set on non negative integers.

Theorem 7.17. *The set of non negative integers satisfies*

1. $\mathbb{Z}_0^+$ *is a sub semi-group of* $\langle \mathbb{Z}, + \rangle$ *[hence by [theorem: 4.14]* $\langle \mathbb{Z}_0^+, + \rangle$ *is a Abelian semi-group].*
2. $\mathbb{Z}_0^+$ *is a sub semi-group of* $\langle \mathbb{Z}, \cdot \rangle$ *[hence by [theorem: 4.14]* $\langle \mathbb{Z}_0^+, \cdot \rangle$ *is a Abelian semi-group].*
3. *The function* $i_{\mathbb{N}_0 \rightarrow \mathbb{Z}}: \mathbb{N}_0 \rightarrow \mathbb{Z}_0^+$ *defined by* $i_{\mathbb{N}_0 \rightarrow \mathbb{Z}}(n) = \sim[(n,0)]$ *is a bijection and*
   a. $i_{\mathbb{N}_0 \rightarrow \mathbb{Z}}: \langle \mathbb{N}_0, + \rangle \rightarrow \langle \mathbb{Z}_0^+, + \rangle$ *is a group isomorphism*
   b. $i_{\mathbb{N}_0 \rightarrow \mathbb{Z}}: \langle \mathbb{N}_0, \cdot \rangle \rightarrow \langle \mathbb{Z}_0^+, \cdot \rangle$ *is a group isomorphism*
4. *For every* $z \in \mathbb{Z}$ $\exists x, y \in \mathbb{Z}_0^+$ *such that* $z = x - y$

Proof.

1. Let $z, z' \in \mathbb{Z}$ then $z = \sim[(n,0)]$ and $z' = \sim[(n',0)]$ so that
$$z + z' = \sim[(n,0)] + \sim[(n',0)] = \sim[(n+n', 0+0)] = \sim[(n+n',0)] \in \mathbb{Z}_0^+$$
further
$$0 = \sim[(0,0)] \in \mathbb{Z}_0^+.$$
Using [definition: 4.12] it follows that $\mathbb{Z}_0^+$ is a sub semi-group of $\langle \mathbb{Z}, + \rangle$.
2. Let $z, z' \in \mathbb{Z}$ then $z = \sim[(n,0)]$ and $z' = \sim[(n',0)]$ so that
$$z \cdot z' = \sim[(n,0)] \cdot \sim[(n',0)] = \sim[(n \cdot n' + 0 \cdot 0, 0 \cdot n' + n \cdot 0)] = \sim[(n \cdot n', 0)] \in \mathbb{Z}_0^+$$
further
$$1 = \sim[(1,0)] \in \mathbb{Z}_0^+$$
Using [definition: 4.12] it follows that $\mathbb{Z}_0^+$ is a sub semi-group of $\langle \mathbb{Z}, \cdot \rangle$.
3. First we show that $i_{\mathbb{N}_0 \rightarrow \mathbb{Z}}$ is a bijection:

   injectivity. If $i_{\mathbb{N}_0 \rightarrow \mathbb{Z}}(n) = i_{\mathbb{N}_0 \rightarrow \mathbb{Z}}(m)$ then $\sim[(n,0)] = \sim[(m,0)]$ so that $n + 0 = 0 + m \Rightarrow n = m$.

   surjectivity. If $z \in \mathbb{Z}_0^+$ there exist a $n \in \mathbb{N}_0$ such that $z = \sim[(n,0)] = i_{\mathbb{N}_0 \rightarrow \mathbb{Z}}(n)$.

   Next we have:
   a. First $i_{\mathbb{N}_0 \rightarrow \mathbb{Z}}(n+m) = \sim[(n+m,0)] = \sim[(n,0)] + \sim[(m,0)] = i_{\mathbb{N}_0 \rightarrow \mathbb{Z}}(n) + i_{\mathbb{N}_0 \rightarrow \mathbb{Z}}(m)$. Secondly $i_{\mathbb{N}_0 \rightarrow \mathbb{Z}}(0) = \sim[(0,0)] = 0 \in \mathbb{Z}_0^+$.
   b. First
$$\begin{array}{rcl} i_{\mathbb{N}_0 \rightarrow \mathbb{Z}}(n) \cdot i_{\mathbb{N}_0 \rightarrow \mathbb{Z}}(m) & = & \sim[(n,0)] \cdot \sim[(m,0)] \\ & = & \sim[(n \cdot m + 0 \cdot m, 0 \cdot n + n \cdot 0)] \\ & = & \sim[(n \cdot m, 0)] \\ & = & i_{\mathbb{N}_0 \rightarrow \mathbb{Z}}(n \cdot m) \end{array}$$
   Second $i_{\mathbb{N}_0 \rightarrow \mathbb{Z}}(1) = \sim[(1,0)] = 1 \in \mathbb{Z}_0^+$.
4. Let $z \in \mathbb{Z}$ then $z = \sim[(n,m)]$, take $x = \sim[(n,0)] \in \mathbb{Z}_0^+$ and $y = \sim[(m,0)] \in \mathbb{Z}_0^+$ then we have
$$x - y = x + (-y) = \sim[(n,0)] + \sim[(0,m)] = \sim[(n,m)] = z$$
□

Next we define the set of non positive numbers.

Definition 7.18. $\mathbb{Z}_0^- = \{-n | n \in \mathbb{Z}_0^+\} = \{(0,n) | n \in \mathbb{N}_0\} \subseteq \mathbb{Z}$

Definition 7.19. $\mathbb{Z}^+ = \mathbb{Z}_0^+ \setminus \{0\}$ *and* $\mathbb{Z}^- = \mathbb{Z}_0^- \setminus \{0\}$

The following theorem shows the relation between $\mathbb{Z}_0^+$ and $\mathbb{Z}_0^-$.

Theorem 7.20. $\mathbb{Z} = \mathbb{Z}_0^+ \bigcup \mathbb{Z}_0^-$ *and* $\{0\} = \mathbb{Z}_0^+ \bigcap \mathbb{Z}_0^-$

Proof. As $\mathbb{Z}_0^+ \subseteq \mathbb{Z}$ and $\mathbb{Z}_0^- \subseteq \mathbb{Z}$ it follows that

$$\mathbb{Z}_0^+ \bigcup \mathbb{Z}_9^- \subseteq \mathbb{Z} \tag{7.4}$$

Let $z \in \mathbb{Z}$ then $\exists n, m \in \mathbb{N}_0$ such that $z = \sim[(n, m)]$, for n, m we have either:

n $\leqslant$ ***m*.** then using [theorem: 5.62] there exist a $k \in \mathbb{N}_0$ such that $m = n + k$ so that

$$z = \sim[(n, n + k)] \tag{7.5}$$

Now for $(0, k)$ and $(n, n + k)$ we have $0 + (n + k) = n + k$ so that $(0, k) \sim (n, n + k)$ proving that $\sim[(0, k)] = \sim[(n, n + k)] \underset{\text{[eq: 7.5]}}{=} z$, proving that $z \in \mathbb{Z}_0^- \subseteq \mathbb{Z}_0^+ \bigcup \mathbb{Z}_0^-$.

***m* < *n*.** Then using [theorem: 5.62] there exist a $k \in \mathbb{N}_0$ such that $n = m + k$ so that

$$z = \sim[(m + k, m)] \tag{7.6}$$

Now for $(k, 0)$ and $(m + k, m)$ we have $k + m = 0 + m + k$ so that $(k, 0) \sim (m + k, m)$ proving that $\sim[(k, 0)] = \sim[(m + k, m)] \underset{\text{[eq: 7.6]}}{=} z$, proving that $z \in \mathbb{Z}_0^+ \subseteq \mathbb{Z}_0^+ \bigcup \mathbb{Z}_0^-$.

From the above we have $\mathbb{Z} \subseteq \mathbb{Z}_0^+ \bigcup \mathbb{Z}_0^-$ which by [eq: 7.4] proves that

$$\mathbb{Z} = \mathbb{Z}_0^+ \bigcup \mathbb{Z}_0^-$$

As $0 = \sim[(0, 0)] \in \mathbb{Z}_0^+$ and $0 = \sim[(0, 0)] \in \mathbb{Z}_0^-$ we have that $\{0\} \in \mathbb{Z}_0^+ \bigcap \mathbb{Z}_0^-$. Let $z \in \mathbb{Z}_0^+ \bigcap \mathbb{Z}_0^-$ then there exists $n, m \in \mathbb{N}_0$ such that $z = \sim[(n, 0)] = \sim[(0, m)]$ hence $n + 0 = 0 + m \Rightarrow n = m,$. So $z = \sim[(n, 0)] = \sim[(0, n)] = -z$. Applying then [theorem: 7.14] it follows that $z = 0$ or $\mathbb{Z}_0^+ \bigcap \mathbb{Z}_0^- \subseteq \{0\}$. Hence

$$\mathbb{Z}_0^+ \bigcap \mathbb{Z}_0^- = \{0\}$$

□

We can now define a order relation on $\mathbb{Z}$.

Theorem 7.21. $\langle \mathbb{Z}, \leqslant \rangle$ *where*

$$\leqslant = \{(x, y) \in \mathbb{Z} \times \mathbb{Z} | y + (-x) \in \mathbb{Z}_0^+ |\}$$

is a totally ordered set.

Proof.

reflexivity. If $x \in \mathbb{Z}$ then $x + (-x) = 0 \in \mathbb{Z}_0^+$ so that $x \leqslant x$.

anti symmetry. Let $x, y \in \mathbb{Z}$ with $x \leqslant y$ and $y \leqslant x$ then

$$y + (-x) \in \mathbb{Z}_0^+ \ \wedge \ x + (-y) \in \mathbb{Z}_0^+$$

then $\exists n, m \in \mathbb{N}_0$ such that

$$y + (-x) = \sim[(n.0)] \wedge x + (-y) = \sim[(m, 0)]$$

so taking the sum we have

$$\begin{aligned} \sim[0, 0] &= 0 \\ &= y + (-x) + x + (-y) \\ &= \sim[(n, 0)] + \sim[(m, 0)] \\ &= \sim[(n + m, 0)] \end{aligned}$$

Hence $0 + 0 = 0 + n + m$ so that $n + m = 0$ which by [theorem: 5.57] proves that $n = m = 0$ so that $y + (-x) = \sim[(n, 0)] = \sim[(0, 0)] = 0$. Hence $x = 0 + x = (y + (-x)) + x = y + ((-x) + x) = y$ from which it follows that $x = y$.

transitivity. If $x \leqslant y$ and $y \leqslant z$ then $y + (-x) \in \mathbb{Z}_0^+$ and $z + (-y) \in \mathbb{Z}_0^+$. Then we have

$$\begin{aligned} z + (-x) &= (z + (-x)) + 0 \\ &= (z + (-x)) + (y + (-y)) \\ &= (y + (-x)) + (z + (-y)) \end{aligned}$$

which as $y+(-x), z+(-y)\in\mathbb{Z}_0^+$ proves by [theorem: 7.17] that $z+(-x)\in\mathbb{Z}_0^+$ proving that $x\leqslant z$.

total ordering. If $x,y\in\mathbb{N}_0$ then we have for $x+(-y)\in\mathbb{Z}\underset{\text{[theorem: 7.20]}}{=}\mathbb{Z}_0^+\bigcup\mathbb{Z}_0^-$ either:

$\boldsymbol{x+(-y)\in\mathbb{Z}_0^+}$**.** Then $y\leqslant x$

$\boldsymbol{x+(-y)\in\mathbb{Z}_0^-}$**.** Then $-(x+(-y))\in\mathbb{Z}_0^+$, further

$$\begin{array}{rcl}-(x+(-y)) & \underset{\text{[theorem: 4.8]}}{=} & -x+(-(-y))\\ & \underset{\text{[theorem: 4.9]}}{=} & -x+y\\ & = & y+(-x)\end{array}$$

proving that $x\leqslant y$. □

Theorem 7.22. $\mathbb{Z}_0^+=\{x\in\mathbb{Z}|0\leqslant x\}$ *and* $\mathbb{Z}_0^-=\{x\in\mathbb{Z}|x\leqslant 0\}$

Proof. First we have

$$x+(-0)\underset{\text{[theorem: 4.9]}}{=}x+0=x \tag{7.7}$$

Now

$$\begin{array}{rcl}x\in\mathbb{Z}_0^+ & \underset{\text{[eq: 7.7]}}{\Leftrightarrow} & x+(-0)\in\mathbb{Z}_0^+\\ & \Leftrightarrow & 0\leqslant x\\ & \Leftrightarrow & x\in\{x\in\mathbb{Z}|0\leqslant x\}\end{array}$$

proving

$$\mathbb{Z}_0^+=\{x\in\mathbb{Z}|0\leqslant x\}$$

Further

$$\begin{array}{rcl}x\in\mathbb{Z}_0^- & \Leftrightarrow & -x\in\mathbb{Z}_0^+\\ & \Leftrightarrow & 0+(-x)\in\mathbb{Z}_0^+\\ & \Leftrightarrow & x\leqslant 0\\ & \Leftrightarrow & x\in\{x\in\mathbb{Z}|x\leqslant 0\}\end{array}$$

proving

$$\mathbb{Z}_0^-=\{x\in\mathbb{Z}|x\leqslant 0\}$$ □

Theorem 7.23. *If* $x=\sim[n,m]\in\mathbb{Z}$ *then we have*

1. $0\leqslant x\Leftrightarrow m\leqslant n$
2. $0<x\Leftrightarrow m<n$
3. *If* $0<x$ *then* $1\leqslant x$

Proof.

1.

$$\begin{array}{rcl}0\leqslant x & \underset{\text{[theorem: 7.22]}}{\Leftrightarrow} & x\in\mathbb{Z}_0^+\\ & \Leftrightarrow & \exists k\in\mathbb{N}_0\,\text{such that}\,x=\sim[(k,0)]\\ & \Leftrightarrow & \exists k\in\mathbb{N}_0\text{ such that }n+0=m+k\Leftrightarrow n=m+k\\ & \underset{\text{[theorem: 5.62]}}{\Leftrightarrow} & m\leqslant n\end{array}$$

2. First

$$\begin{array}{rcl}x\neq 0 & \Leftrightarrow & \sim[(n,m)]\neq\sim[(0,0)]\\ & \Leftrightarrow & n+0\neq m+0\\ & \Leftrightarrow & n\neq m\end{array}$$

then

$$\begin{aligned} 0<x &\Leftrightarrow x\neq 0\wedge 0\leqslant x \\ &\Leftrightarrow n\neq m\wedge 0\leqslant x \\ &\underset{(1)}{\Leftrightarrow} n\neq m\wedge m\leqslant n \\ &\Leftrightarrow m<n \end{aligned}$$

3. If $0<x$ then by (2) $m<n$ so that by [theorem: 5.50]

$$m+1\leqslant n.$$

Now

$$x+(-1)=\sim[(n,m)]+\sim[(0,1)]=\sim[(n,m+1)]$$

so that $0\leqslant x+(-1)$, hence $x+(-1)\in\mathbb{Z}_0^+$ from which we conclude that

$$1\leqslant x$$

□

Corollary 7.24. *$\forall n\in\mathbb{N}_0$ we have $0\leqslant\sim[(n,0)]$ further if $n\neq 0$ then $0<\sim[(n,0)]$*

Proof. By [theorem: 5.46] we have $\forall n\in\mathbb{N}_0$ that $0\leqslant n$ so that by [theorem: 7.23] [$0\leqslant\sim[(n,0)]$, further if $n\neq 0$ then $0<n$, hence by [theorem: 7.23] we have that $0<\sim[(n,0)]$ □

Example 7.25. $0<1$ and $0<2$ where $1,2\in\mathbb{Z}$

Proof. This follows directly from [corollary: 7.24] and the fact that for $1,2\in\mathbb{N}_0$ we have $0<1$ and $0<2$. □

Theorem 7.26. *If $x,y\in\mathbb{Z}$ and $0<x\wedge 0<y$ then $0<x\cdot y$.*

Proof. $x=\sim[(n.m)]$ and $y=\sim[(r,s)]$ then by [theorem: 7.23] we have $m<n$ and $s<r$, so by [theorem: 5.60] there exists $k,l\in\mathbb{N}_0\setminus\{0\}$ such that $n=m+k$ and $r=s+l$. Hence

$$\begin{aligned} n\cdot r+m\cdot s &= (m+k)\cdot(s+l)+m\cdot s \\ &= \underbrace{m\cdot s}_{1}+\underbrace{m\cdot l}_{2}+\underbrace{k\cdot s}_{3}+k\cdot l+\underbrace{m\cdot s}_{4} \\ m\cdot r+n\cdot s &= m\cdot(s+l)+(m+k)\cdot s \\ &= \underbrace{m\cdot s}_{1}+\underbrace{m\cdot l}_{2}+\underbrace{m\cdot s}_{4}+\underbrace{k\cdot s}_{3} \end{aligned}$$

so that

$$n\cdot r+m\cdot s=m\cdot r+n\cdot s+k\cdot l$$

As $0\neq k\Rightarrow 0<k$ and $0\neq l\Rightarrow 0<l$ it follows from [theorem: 5.76] that $0<k\cdot l$ so that $k,l\in\mathbb{N}_0\setminus\{0\}$, using the above together with [theorem: 5.61] proves that

$$m\cdot r+n\cdot s<n\cdot r+m\cdot s \tag{7.8}$$

now

$$x\cdot y = \sim[(n\cdot r+m\cdot s,m\cdot r+n\cdot s)]$$

Combining the above with [eq: 7.8] and [theorem: 7.23 proves finally:

$$0<x\cdot y$$

□

Theorem 7.27. *$\langle\mathbb{Z},+,\cdot,\leqslant\rangle$ is a ordered integral domain [definition: 4.49]*

Proof. Using [theorem: 7.11] $\langle\mathbb{Z},+,\cdot\rangle$ is a integral domain and using [theorem: 7.21] we have that $\langle\mathbb{Z},\leqslant\rangle$ is totally ordered. Next

1. For $n,m,k\in\mathbb{Z}$ with $n<m$ we have

$$m+(-n)\in\mathbb{Z}_0^+ \text{ and } n\neq m\Rightarrow n+k\neq m+k \tag{7.9}$$

Further

$$
\begin{array}{rcl}
(m+k)+(-(n+k)) & \underset{[\text{theorem: } 4.8]}{=} & (m+k)+((-n)+(-k)) \\
& \underset{\text{commutativity}}{=} & (m+k)+((-k)+(-n)) \\
& \underset{\text{associativity}}{=} & m+(k+((-k)+(-n))) \\
& \underset{\text{associativity}}{=} & m+((k+(-k))+(-n)) \\
& = & m+(0+(-n)) \\
& = & m+(-n)
\end{array}
$$

which by [eq: 7.9] proves that $(m+k)+(-(n+k)) \in \mathbb{Z}_0^+$. Hence $n+k \leqslant m+k$ and $n+k \neq m+k$ proving that

$$n+k < m+k$$

2. Let $n, m \in \mathbb{Z}$ with $0 < m$ and $0 < m$ then by [theorem: 7.26] we have $0 < n \cdot m$. □

Theorem 7.28. *Let $x, y \in \mathbb{Z}$ with $x < y$ then*

1. $x+1 \leqslant y$
2. $x \leqslant y+(-1)$

Proof.

1. If $x < y$ then by [theorems 7.27, 4.50] $0 < y+(-x)$, using [theorem: 7.23] we have $1 \leqslant y+(-x)$ so that using [theorem: 7.23] $x+1 \leqslant (y+(-x))+x = y$.
2. By (1) $x+1 \leqslant y$ so that by [theorem: 7.23] $x = (x+1)+(-1) \leqslant y+(-1)$ □

Theorem 7.29. *Define $i_{\mathbb{N}_0 \rightarrow \mathbb{Z}} \colon \mathbb{N}_0 \rightarrow \mathbb{Z}_0^+$ by $i_{\mathbb{N}_0 \rightarrow \mathbb{Z}}(n) = \sim[(n,0)]$ then*

$$i_{\mathbb{N}_0 \rightarrow \mathbb{Z}} \colon \langle \mathbb{N}_0, \leqslant \rangle \rightarrow \langle \mathbb{Z}_0^+, \leqslant \rangle \textit{ is a order isomorphism}$$

Proof. Using [theorem: 7.17 (3)] it follows that

$$i_{\mathbb{N}_0 \rightarrow \mathbb{Z}} \colon \mathbb{N}_0 \rightarrow \mathbb{Z}_0^+ \text{ is a bijection}$$

Further we have:

$$
\begin{array}{rcl}
i_{\mathbb{N}_0 \rightarrow \mathbb{Z}}(x) \leqslant i_{\mathbb{N}_0 \rightarrow \mathbb{Z}}(y) & \Leftrightarrow & i_{\mathbb{N}_0 \rightarrow \mathbb{Z}}(y) + (-i_{\mathbb{N}_0 \rightarrow \mathbb{Z}}(x)) \in \mathbb{Z}_0^+ \\
& \Leftrightarrow & \sim[(y,0)] + (-(\sim[(x,0)])) \in \mathbb{Z}_0^+ \\
& \Leftrightarrow & \sim[(y,0)] + \sim[(0,x)] \in \mathbb{Z}_0^+ \\
& \Leftrightarrow & \sim[(y,x)] \in \mathbb{Z}_0^+ \\
& \underset{[\text{theorem: } 7.23]}{\Leftrightarrow} & x \leqslant y
\end{array}
$$

□

The above theorem allows us to transfer properties of $\mathbb{N}_0$ to $\mathbb{Z}_0^+$ as is expressed in the following theorems.

Theorem 7.30. (Archimedean property) *If $x, y \in \mathbb{Z}$ with $0 < x$ then there exist a $k \in \mathbb{Z}_0^+$ such that $y < k \cdot x$.*

Proof. We have the following cases for y:

$\boldsymbol{y \leqslant 0}$**.** Take $k = 1 \in \mathbb{Z}_0^+$ then as $y \leqslant 0 < x = 1 \cdot x = k \cdot x$ proving that $y < k \cdot x$

$\boldsymbol{0 < y}$**.** Then $y \in \mathbb{Z}_0^+$. Using [theorem: 7.17] $i_{\mathbb{N}_0 \rightarrow \mathbb{Z}} \colon \langle \mathbb{N}_0, \cdot \rangle \rightarrow \langle \mathbb{Z}_0^+, \cdot \rangle$ defined by $i_{\mathbb{N}_0 \rightarrow \mathbb{Z}}(n) = \sim[(n,0)]$ is a group isomorphism. Take $n = (i_{\mathbb{N}_0 \rightarrow \mathbb{Z}})^{-1}(x)$ and $m = (i_{\mathbb{N}_0 \rightarrow \mathbb{Z}})^{-1}(y)$ then $x = i_{\mathbb{N}_0 \rightarrow \mathbb{Z}}(n)$ and $n \neq 0$ [otherwise $x = i_{\mathbb{N}_0 \rightarrow \mathbb{Z}}(n) = i_{\mathbb{N}_0 \rightarrow \mathbb{Z}}(0) = 0$]. Using the Archimedean property of the natural numbers [see theorem: 5.78] there exists a $l \in \mathbb{N}_0$ such that $m < l \cdot n$. So by [theorem: 7.29] we have that

$$i_{\mathbb{N}_0 \rightarrow \mathbb{Z}}(m) < i_{\mathbb{N}_0 \rightarrow \mathbb{Z}}(l \cdot n) \underset{[\text{theorem: } 7.17]}{=} i_{\mathbb{N}_0 \rightarrow \mathbb{Z}}(l) \cdot i_{\mathbb{N}_0 \rightarrow \mathbb{Z}}(n) \tag{7.10}$$

Take $k = i_{\mathbb{N}_0 \rightarrow \mathbb{Z}}(l) \in \mathbb{Z}_9^+$ then as $i_{\mathbb{N}_0 \rightarrow \mathbb{Z}}(n) = i_{\mathbb{N}_0 \rightarrow \mathbb{Z}}((i_{\mathbb{N}_0 \rightarrow \mathbb{Z}})^{-1}(x)) = x$ and $i_{\mathbb{N}_0 \rightarrow \mathbb{Z}}(m) = i_{\mathbb{N}_0 \rightarrow \mathbb{Z}}((i_{\mathbb{N}_0 \rightarrow \mathbb{Z}})^{-1}(y)) = y$ we have by [eq: 7.10] that

$$y < k \cdot x$$

□

Theorem 7.31. *$\langle \mathbb{Z}_0^+, \leqslant \rangle$ is a well-ordered set*

Proof. Using [theorem: 7.29] we have that $i_{\mathbb{N}_0 \rightarrow \mathbb{Z}}: \mathbb{N}_0 \rightarrow \mathbb{Z}_0^+$ is a order isomorphism, further by [theorem: 5.51] $\langle \mathbb{N}_0, \leqslant \rangle$ is well ordered. so using [theorem: 3.90] we conclude that

$$\langle \mathbb{Z}_0^+, \leqslant \rangle \text{ is well ordered}$$

□

Theorem 7.32. *$\langle \mathbb{Z}_0^+, \leqslant \rangle$ is conditional complete [see definition: 3.77].*

Proof. As by [theorem: 7.31] $\langle \mathbb{Z}_0^+, \leqslant \rangle$ is well-ordered it follows from [theorem: 3.84] it follows that $\langle \mathbb{Z}_0^+, \leqslant \rangle$ is conditional complete. □

Theorem 7.33. *If $A \subseteq \mathbb{Z}_0^+$ is such that $A \neq \varnothing$ and $\sup(A)$ exists then $\sup(A) \in A$.*

Proof. By [theorem: 7.29]

$$i_{\mathbb{N}_0 \rightarrow \mathbb{Z}}: \langle \mathbb{N}_0, \leqslant \rangle \rightarrow \langle \mathbb{Z}_0^+, \leqslant \rangle \text{ is a order isomorphism}$$

which by [theorem: 3.55] means that

$$(i_{\mathbb{N}_0 \rightarrow \mathbb{Z}})^{-1}: \langle \mathbb{Z}_0^+, \leqslant \rangle \rightarrow \langle \mathbb{N}_0, \leqslant \rangle \text{ is a order isomorphism}$$

Assume that $M = \sup(A)$ exists then by [theorem: 3.79] $\sup((i_{\mathbb{N}_0 \rightarrow \mathbb{Z}})^{-1}(A))$ exist and $\sup((i_{\mathbb{N}_0 \rightarrow \mathbb{Z}})^{-1}(A)) = (i_{\mathbb{N}_0 \rightarrow \mathbb{Z}})^{-1}(M)$. By [theorem: 5.73] we have that $\sup((i_{\mathbb{N}_0 \rightarrow \mathbb{Z}})^{-1}(A)) \in (i_{\mathbb{N}_0 \rightarrow \mathbb{Z}})^{-1}(A)$ so that $(i_{\mathbb{N}_0 \rightarrow \mathbb{Z}})^{-1}(M) \in (i_{\mathbb{N}_0 \rightarrow \mathbb{Z}})^{-1}(A)$, hence $M = i_{\mathbb{N}_0 \rightarrow \mathbb{Z}}((i_{\mathbb{N}_0 \rightarrow \mathbb{Z}})^{-1}(M)) \in (i_{\mathbb{N}_0 \rightarrow \mathbb{Z}})((i_{\mathbb{N}_0 \rightarrow \mathbb{Z}})^{-1}(A)) = A$ or

$$\sup(A) \in A$$

□

Definition 7.34. (Absolute Value) *If $x \in \mathbb{Z}$ then $|x|$ is defined by*

$$|x| = \begin{cases} x & \text{if } 0 \leqslant x \\ -x & \text{if } x < 0 \end{cases}$$

Theorem 7.35. *If $x, y \in \mathbb{Z}$ then $|x \cdot y| = |x| \cdot |y|$*

Proof. We have the following possibilities for x, y:

$0 \leqslant x \wedge 0 \leqslant y$. Then $|x| = x$ and $|y| = y$. Further by [theorems: 7.27, 4.50] $0 = 0 \cdot y \leqslant x \cdot y$, hence $x \cdot y = |x \cdot y|$. So we have that $|x \cdot y| = |x| \cdot |y|$.

$0 \leqslant x \wedge y < 0$. Then $x = |x|$ and $-y = |y|$, further by [theorems: 7.27, 4.50] $x \cdot y \leqslant 0 \cdot y = 0$, hence $|x \cdot y| = -(x \cdot y)$. So

$$|x| \cdot |y| = x \cdot (-y) \underset{\text{[theorem: 4.40]}}{=} -(x \cdot y) = |x \cdot y|.$$

$x < 0 \wedge 0 \leqslant y$. Then $-x = |x|$ and $y = |y|$, further by [theorems: 7.27, 4.50] $x \cdot y \leqslant 0 \cdot y = 0$, hence $|x \cdot y| = -(x \cdot y)$. So

$$|x| \cdot |y| = (-x) \cdot y \underset{\text{[theorem: 4.40]}}{=} -(x, y) = |x \cdot y|$$

$x < 0 \wedge y < 0$. Then $-x = |x|$, $-y = |y|$, further by [theorems: 7.27, 4.50] $0 = 0 \cdot y < x \cdot y$, hence $|x \cdot y| = x \cdot y$. So

$$|x| \cdot |y| = (-x) \cdot (-y) \underset{\text{[theorem: 4.40]}}{=} -(-(x \cdot y)) \underset{\text{[theorem: 4.9]}}{=} x \cdot y| = x \cdot y|$$

□

Theorem 7.36. *If $x \in \mathbb{Z}$ then $x \leqslant |x|$*

Proof. If $0 \leqslant x$ then $x = |x|$ so that trivially $x \leqslant |x|$, if $x < 0$ then by [theorems: 7.27, 4.50] $0 < -x = |x|$ so that by transitivity $x < |x|$ or $x \leqslant |x|$. □

Theorem 7.37. *$\forall x \in \mathbb{Z}$ we have $|x| = 0 \Leftrightarrow x = 0$*

Proof.

$\Rightarrow$. If $x = 0$ then $0 \leqslant x$ so that $|x| = x = 0$ hence $|x| = 0$

$\Leftarrow$. If $|x| = 0$ then if $x < 0$ we would have $-x = |x| = 0$ so that $-x = 0 \Rightarrow x = 0$ contradicting $x < 0$. So we must have $0 \leqslant x$, hence $x = |x| = 0$ proving that $x = 0$. □

We introduce now division, just as it was done for the natural numbers.

Theorem 7.38. (Division Algorithm) *If $n,m\in\mathbb{Z}$ and $0<n$ then there exists **unique** $r\in\mathbb{Z}_0^+$, $q\in\mathbb{Z}$ such that $0\leqslant r<n$ and $m=n\cdot q+r$*

Proof. First we prove existence, let $n,m\in\mathbb{Z}$ with $0<n$. Define

$$A_{n,m}=\{m+n\cdot q|q\in\mathbb{Z}\wedge 0\leqslant m+n\cdot q\}\subseteq\mathbb{Z}_0^+.$$

Using $0<n$ and the Archimedean property of $\mathbb{Z}$ [see theorem: 7.30] there exist a $k\in\mathbb{Z}_0^+$ such that $-m<n\cdot k$, using [theorems: 7.27, 4.50] it follows that $0<n\cdot k+(-(-m))=n\cdot k+m=m+n\cdot k$ proving that $m+n\cdot k\in A_{n,m}$, hence $A_{n,m}\neq\varnothing$. As $\langle\mathbb{Z}_0^+,\leqslant\rangle$ is well-ordered [see theorem: 7.31] $A_{n,m}$ has a least element, hence

$$\exists r'\in A_{n,m}\text{ such that }\forall a\in A_{n,m}\text{ we have }r'\leqslant a \tag{7.11}$$

As $r'\in A_{n,m}$ there exist a $q'\in\mathbb{Z}$ such that

$$r'=m+n\cdot q'\text{ and }0\leqslant r' \tag{7.12}$$

Assume that $n<r'$ then by [theorems: 7.27, 4.50] $k=r'+(-n)\in\mathbb{Z}_0^+\setminus\{0\}$ so that $r'=n+k$. Hence $m+n\cdot q'=n+k$, so that $0<k=m+n\cdot q'+(-n)=m+(q'-1)\cdot n$ proving that $k\in A_{n,m}$. Now $0<n\underset{\text{[theorems: 7.27, 4.50]}}{\Rightarrow}k<n+k=r'\Rightarrow k<r'$, as $k\in A_{n,m}$ we have by [eq: 7.11] $r'\leqslant k$, giving the contradiction $k<k$. So we must have that $r'\leqslant n$ or

$\boldsymbol{r'=n}$**.** In this case we we have that $m+n\cdot q'=r'=n$, hence

$$m=n+(-(n\cdot q'))=n\cdot 1+n\cdot(-q')=n\cdot(1+(-q'))$$

So by taking $q=(1+(-q'))$ and $r=0<n$ we have

$$m=n\cdot q+r\text{ and }0\leqslant r<n$$

$\boldsymbol{r'<n}$**.** Then as $r'=m+n\cdot q'$ we have $m=r'+(-(n\cdot q'))=r'+n\cdot(-q')$, so taking $q=-q'$ and $r=r'$ then

$$m=n\cdot q+r\ \text{ and }\ 0\leqslant r'<n$$

Now for uniqueness assume that there exists $q_1,q_2\in\mathbb{Z}$ and $r_1,r_2\in\mathbb{Z}_0^+$ such that

$$m=n\cdot q_1+r_1\wedge m=n\cdot q_2+r_2\wedge 0\leqslant r_1<n\wedge 0\leqslant r_2<n$$

Then

$$\begin{aligned}
n\cdot q_1+r_1=n\cdot q_2+r_2 &\Rightarrow n\cdot q_1+(-(n\cdot q_2))=r_2+(-r_1)\\
&\Rightarrow n\cdot(q_1+(-q_2))=r_2+(-r_1) \qquad (7.13)\\
n\cdot q_1+r_1=n\cdot q_2+r_2 &\Rightarrow n\cdot q_2+(-(n\cdot q_1))=r_1+(-r_2)\\
&\Rightarrow n\cdot(q_2+(-q_1))=r_1+(-r_2) \qquad (7.14)
\end{aligned}$$

Assume now that $r_1\neq r_2$ then we have either:

$\boldsymbol{r_1<r_2}$**.** Then by [theorems: 7.27, 4.50] $0<r_2+(-r_1)\underset{\text{[eq: 7.13]}}{=}n\cdot(q_1+(-q_2))$, hence $0\cdot n<(q_1+(-q_2))\cdot n$, as $0<n$ we must have by [theorems: 7.27, 4.50] that $0<q_1+(-q_2)$. Using [theorem: 7.23] we have

$$1\leqslant q_1+(-q_2) \tag{7.15}$$

As $r_2<n$ we have by [theorems: 7.27, 4.50] that $r_2+(-r_1)<n+(-r_1)$, further as $(-r_1)\leqslant 0$ we have by [theorems: 7.27, 4.50] that $n+(-r_1)\leqslant n$ so that $r_2+(-r_1)<n$. Using this with [eq: 7.13] gives $n\cdot(q_1+(-q_2))<n=1\cdot n$, hence using [theorems: 7.27, 4.50] we have that $q_2+(-q_1)<1$, contradicting [eq: 7.15]. So this case never occurs.

$\boldsymbol{r_2<r_1}$**.** Then by [theorems: 7.27, 4.50] $0<r_1+(-r_2)\underset{\text{[eq: 7.14]}}{=}n\cdot(q_2+(-q_1))$, hence $0\cdot n<(q_2+(-q_1))\cdot n$ as $0<n$ we must have by [theorems: 7.27, 4.50] that $0<q_2+(-q_1)$. Using [theorem: 7.23] we have

$$1\leqslant q_2+(-q_1) \tag{7.16}$$

As $r_1<n$ we have by [theorems: 7.27, 4.50] that $r_1+(-r_2)<n+(-r_2)$, further as $(-r_2)\leqslant 0$ we have by [theorems: 7.27, 4.50] that $n+(-r_2)\leqslant n$ so that $r_1+(-r_2)<n$. Using this with [eq: 7.14] gives $n\cdot(q_2+(-q_1))<n=1\cdot n$, hence using [theorems: 7.27, 4.50] we have that $q_1+(-q_2)<1$, contradicting [eq: 7.16]. So this case never occurs.

As all the cases lead to a contradiction the assumption $r_1\neq r_2$ is wrong. Hence

$$r_1=r_2$$

So $n\cdot q_1+r_1\underset{r_1=r_2}{=}n\cdot q_2+r_1$ giving, by adding $-r_1$ to both sides, that $n\cdot q_1=n\cdot q_2$. Applying [theorem: 7.15] proves then

$$q_1=q_2$$

□

Definition 7.39. *If $n,m\in\mathbb{Z}$ then we say that n divides m noted as $n|m$ if there exist a $q\in\mathbb{Z}$ such that $q\cdot n=m$, we call n a **divisor** of m.*

Example 7.40. Every integer is a divisor of 0.

Proof. If $n\in\mathbb{Z}$ then $n\cdot 0=0$ □

Example 7.41. If $n\in\mathbb{Z}$ then $1|n$

Proof. As $1\cdot n=n$ we have by definition $1|n$. □

Theorem 7.42. *Let $m\in\mathbb{Z}$ then if $n|m$ we have that $(-n)|m$. In other words if n is a divisor of m then $-n$ is a divisor of m. So as $|n|=\left\{\begin{smallmatrix}-n \text{ if } n<0\\ n \text{ if } 0\leqslant n\end{smallmatrix}\right.$ we have also that $n|m\Rightarrow|n||m$.*

Proof. If $n|m$ then there exist a q such that $n\cdot q=m$, then $(-n)\cdot(-q)=n\cdot q=m$ so that $(-n)|m$. □

Theorem 7.43. *If $m\in\mathbb{Z}$ and $n\in\mathbb{Z}\backslash\{0\}$ a divisor of m then there exists a **unique** q such that $n\cdot q=m$*

Proof. Existence follows from the definition of divisor. Now for uniqueness assume that $q_1,q_2\in\mathbb{Z}$ such that $n\cdot q_1=m=n\cdot q_2$ then by [theorem: 7.15] $q_1=q_1$. □

Definition 7.44. *If $m\in\mathbb{Z}$ and $n\in\mathbb{Z}\backslash\{0\}$ then the unique number q such that $m=n\cdot q$ is called the **quotient** of n and m and is noted as $m:n$. So $n\cdot(m:n)=m$.*

Definition 7.45. (Common Divisor) *If $n,m\in\mathbb{Z}$ then d is a **common divisor** of n and m if $d|n$ and $d|m$.*

Lemma 7.46. *If $n,m\in\mathbb{Z}$ such that $m\neq 0$ and $n|m$ then $n\leqslant|m|$*

Proof. As $n|m$ there exist a $q\in\mathbb{Z}$ such that $n\cdot q=m$, as $m\neq 0$ we must have $q\neq 0$ [otherwise $m=n\cdot q=0$]. For n,m we have now the following possibilities to consider:

- $\boldsymbol{0<m\wedge n\leqslant 0}$. In this case we have $n\leqslant 0<m\leqslant|m|$ so that $n\leqslant|m|$
- $\boldsymbol{0<m\wedge 0<n}$. If $q\leqslant 0\underset{q\neq 0}{\Rightarrow}q<0\underset{0<n\wedge[\text{theorems: 7.27, 4.50}]}{\Rightarrow}q\cdot n<0\cdot n=0$ so that $m=q\cdot n<0$ contradicting $0<m$, hence we must have that $0<q$. Using [theorem: 7.23] we have $1\leqslant q$ so that by [theorems: 7.27, 4.50] $n=1\cdot n\leqslant q\cdot n=m=|m|$, hence $n\leqslant|m|$.
- $\boldsymbol{m<0\wedge n\leqslant 0}$. Then $0<-m=|m|$ so that $n\leqslant 0<|m|$ giving $n\leqslant|m|$.
- $\boldsymbol{m<0\wedge 0<n}$. If $0\leqslant q\underset{q\neq 0}{\Rightarrow}0<q\underset{0<n\wedge[\text{theorems: 7.27, 4.50}]}{\Rightarrow}0=0\cdot n<q\cdot n=m$ contradicting $m<0$, hence $q<0$, so that $0<-q$. Using [theorem: 7.23] we have then

$$1\leqslant -q\underset{[\text{theorems:}7.27,4.50]}{\Rightarrow}n=1\cdot n\leqslant(-q)\cdot n=-(q\cdot n)=|m|$$

  proving that $n\leqslant|m|$.

So in all cases we have

$$n\leqslant|m|$$

□

Theorem 7.47. *Let $n,m\in\mathbb{Z}$ with $n\neq 0$ then $\max(\{d\in\mathbb{Z}_0^+|d$ is a common divisor of n and $m\})$ exist and $0<1\leqslant\max(\{d\in\mathbb{Z}_0^+|d$ is a common divisor of n and $m\})$*

Proof. Let $n,m\in\mathbb{Z}$ and define $D_{n,m}=\{d\in\mathbb{Z}_0^+|d$ is a common divisor of n and $m\}$. By [example: 7.41] 1 is a common divisor of n and m, which as $0<1$ means that $1\in D_{n,m}$ so that $D_{n,m}\neq\varnothing$. Let $d\in D_{n,m}$ then as $d|n$ and $n\neq 0$ we have by [lemma: 7.46] that $d\leqslant|n|$ so that $D_{n,m}$ has a upper bound. As $\langle\mathbb{Z}_0^+,\leqslant\rangle$ is conditional complete [see theorem: 7.32] it follows that $\sup(D_{n,m})$ exist, using [theorem: 7.33] $\sup(D_{n,m})\in D_{n,m}$. Hence $\max(D_{n,m})$ $\max(D_{n,m})$ exists. □

The above theorem ensures that the following definition is well defined,

Definition 7.48. *Let* $n, m \in \mathbb{Z}_0^+$ *with* $n \neq 0$ *then*

$$\gcd(n, m) = \max(\{d \in \mathbb{Z}_0^+ | d \textit{ is a common divisor if } n \textit{ and } m\} \geqslant 1 > 0$$

$\gcd(n, m)$ *is called the* ***greatest common divisor*** *of* n *and* m.

Theorem 7.49. *If* $n, m \in \mathbb{Z}$ *with* $m \neq 0$ *then we have*

1. $\{d \in \mathbb{Z} | d | (n: \gcd(n, m)) \wedge d | (m: \gcd(n, m))\} = \{1, -1\}$
2. $\gcd(n: \gcd(n, m), m: \gcd(n, m)) = 1$

Proof. As $\gcd(n, m) | n$ and $\gcd(n, m) | m$ the quotients $n: \gcd(n, m)$ and $m: \gcd(n, m)$ are well defined.

1. Take $n' = n: \gcd(n, m)$ and $m' = m: \gcd(n, m)$ then $n = n' \cdot \gcd(n, m)$ and $m = m' \cdot \gcd(n, m)$. If $d | n'$ and $d | m'$ there exists n'', $m'' \in \mathbb{Z}$ such that $n'' \cdot d = n'$ and $m'' \cdot d = m'$. Multiplying both sides by $\gcd(n, m)$ gives

$$(d \cdot \gcd(n, m)) \cdot n'' = (n'' \cdot d) \cdot \gcd(n, m) = n' \cdot \gcd(n, m) = n \tag{7.17}$$

and

$$(d \cdot \gcd(n, m)) \cdot m'' = (m'' \cdot d) \cdot \gcd(n, m) = m' \cdot \gcd(n, m) = m \tag{7.18}$$

proving that $d \cdot \gcd(n, m) | n$ and $d \cdot \gcd(n, m) | m$ Using [theorem: 7.42] and $0 < \gcd(n, m)$ we have that

$$|d| \cdot \gcd(n, m) | n \text{ and } |d| \cdot \gcd(n, m) | m$$

So by the definition of $\gcd(n, m)$ we have then that

$$|d| \cdot \gcd(n, m) \leqslant \gcd(n, m) = 1 \cdot \gcd(n, m)$$

As $0 < \gcd(n, m)$ we have by [theorems: 7.27, 4.50] and the above that

$$|d| \leqslant 1$$

If $d = 0$ then by [eq: 7.18] $m = 0$ contradicting $m \neq 0$ so we have $d \neq 0$, proving by [theorem: 7.37] that $|d| \neq 0$ which as $0 \leqslant |d|$ implies that $0 < |d|$ or using [theorem: 7.23] $1 \leqslant |d|$, which by the above proves that $|d| = 1$ hence $d = 1$ or $d = -1$. So

$$\{d \in \mathbb{Z} | d | (n: \gcd(n, m)) \wedge d | (m: \gcd(n, m))\} = \{1, -1\}$$

2. We have

$$\begin{aligned} \gcd(n: \gcd(n, m), m: \gcd(n, m)) &= \max(\{d \in \mathbb{Z}_0^+ | d | (n: \gcd(n, m)) \wedge d | (m: \gcd(n, m))\}) \\ &\underset{(1)}{=} \max(\{1, -1\}) \\ &= 1 \end{aligned}$$

□

Definition 7.50. *A* $z \in \mathbb{Z}$ *is* ***even*** *if* $2 | z$ *and* ***odd*** *is* z *is not even.*

Theorem 7.51. *Let* $z \in \mathbb{Z}$ *then we have*

1. z *is even* $\Leftrightarrow \exists m \in \mathbb{Z}$ *such that* $z = 2 \cdot m$
2. z *is odd* $\Leftrightarrow \exists m \in \mathbb{Z}$ *such that* $z = 2 \cdot m + 1$

Proof.

1.

$$\begin{aligned} z \text{ is even} &\Leftrightarrow 2 | z \\ &\Leftrightarrow \exists m \in \mathbb{Z} \text{ such that } z = 2 \cdot m \end{aligned}$$

2. Using the Division Algorithm [see: theorem: 7.38] there exists unique $q, r \in \mathbb{Z}$ such that $z = 2 \cdot q + r$ and $0 \leqslant r < 2$ proving that $r \in \{0, 1\}$. So

$$\begin{aligned} z \text{ is odd} &\Leftrightarrow z \text{ is not even} \\ &\underset{r = 0 \Rightarrow z \text{ is evn}}{\Leftrightarrow} z = 2 \cdot q + 1 \end{aligned}$$

□

Theorem 7.52. *If $z \in \mathbb{Z}$ then we have*

1. *z is even $\Leftrightarrow z^2 = z \cdot z$ is even*
2. *z is odd $\Leftrightarrow z^2 = z \cdot z$ is odd*

Proof.

1. If z is even then $z = 2 \cdot m$ so that $z \cdot z = (2 \cdot m) \cdot (2 \cdot m) = 2 \cdot (2 \cdot (m \cdot m))$ proving that $z \cdot z$ is even. If $z.z$ is even then if z is odd we have $z = 2 \cdot m + 1$ so that

$$\begin{aligned} z \cdot z &= (2 \cdot m + 1) \cdot (2 \cdot m + 1) \\ &= 2 \cdot (m \cdot (2 \cdot m + 1)) + 2 \cdot m + 1 \\ &= 2 \cdot (m \cdot (2 \cdot m + 1) + m) + 1 \end{aligned}$$

   proving that $z \cdot z$ is odd contradiction the fact that $z \cdot z$ is even, hence z should be even.

2. This follows from (1) by contra position. □

7.3 Denumerability of the Integers

Theorem 7.53. *$\mathbb{Z}_0^+$, $\mathbb{Z}_0^+$ and $\mathbb{Z}$ are all denumerable*

Proof. Using [theorem: 7.17 (3)] there exists a bijection $i_{\mathbb{N}_0 \rightarrow \mathbb{Z}}: \mathbb{N}_0 \rightarrow \mathbb{Z}_0^+$ so that $\mathbb{N}_0 \approx \mathbb{Z}_0^+$, hence

$$\mathbb{Z}_0^+ \text{ is denumerable}$$

Define now $\beta: \mathbb{Z}_0^+ \rightarrow \mathbb{Z}_0^-$ by $\beta(n) = -n$ then we have

injectivity. If $\beta(n) = \beta(n')$ then $-n = -n' \Rightarrow n = (-(-n)) = (-(-n')) = n'$

surjectivity. If $n \in \mathbb{Z}_0^- = \{-n | n \in \mathbb{Z}_0^-\}$ there exists $m \in \mathbb{Z}^+$ such that $n = -m = \beta(m)$

Hence $\beta: \mathbb{Z}_0^+ - < \mathbb{Z}_0^-$ is a bijection proving that $\mathbb{Z}_0^+ \approx \mathbb{Z}_0^-$. So using [theorem: 6.26] it follows that

$$\mathbb{Z}_0^- \text{ is denumerable}$$

Finally as $\mathbb{Z} \underset{\text{[theorem: 7.20]}}{=} \mathbb{Z}_0^+ \bigcup \mathbb{Z}_0^-$ it follows by [theorem: 6.65] that

$$\mathbb{Z} \text{ is denumerable}$$

□

Chapter 8
The Rational Numbers

In this chapter we will introduce the set of rational numbers and embed the integer numbers in it. Just as with $\mathbb{Z}$ and $\mathbb{N}_0$ we will introduce a order relation, a sum operator, a product operator, neutral elements for addition and multiplication as well as inverse elements. To avoid excesive notation we use the same symbols for the natural numbers, integers and rational numbers and use context to determine the meaning of the symbols involved. The following table should help you in determining the meaning of the different symbols based on the context of their usage.

Context	Expression	Operator
$n,m\in\mathbb{N}_0$	n+m	sum in $\langle\mathbb{N}_0,+\rangle$
$n,m\in\mathbb{N}_0$	$n\cdot m$	product in $\langle\mathbb{N}_0,\cdot\rangle$
$n,m\in\mathbb{N}_0$	$n\leqslant m$	order in $\langle\mathbb{N}_0,\leqslant\rangle$
$n,m\in\mathbb{N}_0$	$n<m$	strict order in $\langle\mathbb{N}_0,\leqslant\rangle$
$n,m\in\mathbb{N}_0$	$n-m$	subtraction in $\langle\mathbb{N}_0,+\rangle$
$n\in\mathbb{N}_0$	$n+0$ or $0+n$	neutral element in $\langle\mathbb{N}_0,+\rangle$
$n\in\mathbb{N}_0$	$n\cdot 1$ or $1\cdot n$	neutral element in $\langle\mathbb{N}_0,\cdot\rangle$
$n\in\mathbb{N}_0$	$-n$	inverse element in $\langle\mathbb{N}_0,+\rangle$
$n,m\in\mathbb{Z}$	n+m	sum in $\langle\mathbb{Z},+\rangle$
$n,m\in\mathbb{Z}$	$n\cdot m$	product in $\langle\mathbb{Z},\cdot\rangle$
$n,m\in\mathbb{Z}$	$n\leqslant m$	order in $\langle\mathbb{Z}\leqslant\rangle$
$n,m\in\mathbb{Z}$	$n<m$	strict order in $\langle\mathbb{Z},\leqslant\rangle$
$n,m\in\mathbb{Z}$	$n-m$	subtraction in $\langle\mathbb{Z},-\rangle$
$n\in\mathbb{Z}$	$n+0$ or $0+n$	neutral element in $\langle\mathbb{Z},+\rangle$
$n\in\mathbb{Z}$	$n\cdot 1$ or $1\cdot n$	neutral element in $\langle\mathbb{Z},\cdot\rangle$
$n\in\mathbb{Z}$	$-n$	inverse element in $\langle\mathbb{Z},+\rangle$
$q,r\in\mathbb{Q}$	q+r	sum in $\langle\mathbb{Q},+\rangle$
$q,r\in\mathbb{Q}$	$q\cdot r$	product in $\langle\mathbb{Q},\cdot\rangle$
$q,r\in\mathbb{Q}$	$q\leqslant r$	order in $\langle\mathbb{Q}\leqslant\rangle$
$q,r\in\mathbb{Q}$	$q<r$	strict order in $\langle\mathbb{Q},\leqslant\rangle$
$q,e\in\mathbb{Q}$	$q-r$	subtraction in $\langle\mathbb{Q},-\rangle$
$q,r\in\mathbb{Q}$	q/r	division in $\langle\mathbb{Q},\cdot\rangle$
$q\in\mathbb{Q}$	$q+0$ or $0+q$	neutral element in $\langle\mathbb{Q},+\rangle$
$q\in\mathbb{Q}$	$q\cdot 1$ or $1\cdot q$	neutral element in $\langle\mathbb{Q},\cdot\rangle$
$q\in\mathbb{Q}$	$-q$	inverse element in $\langle\mathbb{Q},+\rangle$

8.1 Definition and arithmetic

One of the problems that the integer numbers have is that the quotient of two numbers n and m is only defined if n divides m. The following example shows this issue.

Example 8.1. If x is a even number and y is a odd number then x can not divide y.

Proof. As x is even there exists a $n\in\mathbb{Z}$ such that $x=2\cdot n$ and as y is odd y is not even. Assume that $x|y$ then there exists a $q\in\mathbb{Z}$ such that $y=x\cdot q$ but then $y=(2\cdot n)\cdot q=2\cdot(n\cdot q)$ proving that y is even, contradicting the fact that y is odd. □

The rational numbers will resolve this defect. Just as we have done with set of integers we work with pairs of integers (n,m) that will be interpreted as the quotient $\frac{n}{m}$. We have to be careful however for if $m=0$ then the quotient only exist if $n=0$ and then every integer is a quotient. So we should only consider pairs (n,m) where $n\in\mathbb{Z}$ and $m\in\mathbb{Z}\setminus\{0\}$. Further we have that $\frac{8}{4}=\frac{6}{3}=\frac{4}{2}=\frac{2}{1}=2$ so we have to define a equivalence relation and work with equivalence classes.

Definition 8.2. $\mathbb{Z}^*=\mathbb{Z}\setminus\{0\}$

Theorem 8.3. *The relation $\simeq\subseteq(\mathbb{Z}\times\mathbb{Z}^*)\times(\mathbb{Z}\times\mathbb{Z}^*)$ defined by*

$$\simeq=\{((n,m),(r,s))\in(\mathbb{Z}\times\mathbb{Z}^*)\times(\mathbb{Z}\times\mathbb{Z}^*)|n\cdot s=m\cdot r\}$$

is a equivalence relation in $\mathbb{Z}\times\mathbb{Z}^*$.

Proof.

reflexivity. If $\{n,m\}\in\mathbb{Z}\times\mathbb{Z}^*$ then $n\cdot m \underset{[\text{theorem: } 7.7]}{=} m\cdot n$ so that $(n,m)\simeq(n,m)$

symmetry. If $(n,m)\simeq(r,s)$ then $n\cdot s=m\cdot r \underset{[\text{theorem: } 7.7]}{\Rightarrow} r\cdot m=s\cdot n$ proving that $(r,s)\simeq(n,m)$

transitivity. If $(n,m)\simeq(k.l)$ and $(k,l)\simeq(r,s)$ then $n\cdot l=m\cdot k$ and $k\cdot s=l\cdot r$, further

$$\begin{aligned}
(n\cdot l)\cdot s=(m\cdot k)\cdot s &\underset{[\text{theorem: } 7.11]}{\Rightarrow} (n\cdot s)\cdot l=m\cdot(k\cdot s)\\
&\Rightarrow (n\cdot s)\cdot l=m\cdot(l\cdot r)\\
&\underset{[\text{theorem: } 7.11]}{\Rightarrow} (n\cdot s)\cdot l=(m\cdot r)\cdot l\\
&\underset{l\neq 0\wedge[\text{theorem: } 7.15]}{\Rightarrow} n\cdot s=m\cdot r\\
&\Rightarrow (n,m)\simeq(r,s)
\end{aligned}$$

□

Definition 8.4. *The set of rational numbers noted as $\mathbb{Q}$ is defined as*

$$\mathbb{Q}=(\mathbb{Z}\times\mathbb{Z}^*)/\simeq$$

or using the definition of $(\mathbb{Z}\times\mathbb{Z}^*)/\simeq$

$$\mathbb{Q}=\{\simeq[(n,m)]|(n,m)\in\mathbb{Z}\times\mathbb{Z}^*\}$$

We note $\simeq[(n,m)]\in\mathbb{Q}$ as $\frac{n}{m}$, n is called the ***numerator*** *and m is called the* ***denominator****. Using this notation we have that $\frac{n}{m}=\frac{n'}{m'}\Leftrightarrow n\cdot m'=m\cdot n'$. In this new notation we have*

$$\mathbb{Q}=\left\{\frac{n}{m}|(n,m)\in\mathbb{Z}^*\right\}$$

Theorem 8.5. *If $k\in\mathbb{Z}^*$ and $(n,m)\in\mathbb{Z}\times\mathbb{Z}^*$ then*

1. $\frac{n}{m}=\frac{n\cdot k}{m\cdot k}$
2. $\frac{0}{n}=\frac{0}{1}$
3. $\frac{n}{m}=\frac{0}{1}\Leftrightarrow n=0$
4. $\frac{n}{m}\neq\frac{0}{1}\Leftrightarrow n\neq 0$

Proof.

1. First as $k\neq 0$ and $m\neq 0$ we have that $m\cdot k\neq 0$ so that $\frac{n\cdot k}{m\cdot k}\in\mathbb{Q}$. Further

   $$n\cdot(m\cdot k) \underset{[\text{theorem: } 7.11]}{=} m\cdot(n\cdot k)$$

   proving that

   $$\frac{n}{m}=\frac{n\cdot k}{m\cdot k}$$

2. As $0\cdot 1=0=n\cdot 0$ we have $\frac{0}{n}=\frac{0}{1}$
3. $\frac{n}{m}=\frac{0}{1}\Leftrightarrow n\cdot 1=m\cdot 0\Leftrightarrow n=0$
4. This follows from (3) by contra-position. □

Theorem 8.6. *Let* $\frac{n}{m}, \frac{n'}{m'}, \frac{r}{s}, \frac{r'}{s'} \in \mathbb{Q}$ *are such that* $\frac{n}{m} = \frac{n'}{m'}$ *and* $\frac{r}{s} = \frac{r'}{s'}$ *then*

$$\frac{n \cdot s + r \cdot m}{m \cdot s}, \frac{n' \cdot s' + r' \cdot m'}{m' \cdot s'} \in \mathbb{Q} \text{ and } \frac{n \cdot s + r \cdot m}{m \cdot s} = \frac{n' \cdot s' + r' \cdot m'}{m' \cdot s'}$$

Proof. First as $m \neq 0, m' \neq 0, s \neq 0$ and $s' \neq 0$ we have $m \cdot s \neq 0, m' \cdot s' \neq 0$ so that

$$\frac{n \cdot s + r \cdot m}{m \cdot s}, \frac{n' \cdot s' + r' \cdot m'}{m' \cdot s'} \in \mathbb{Q}$$

As $\frac{n}{m} = \frac{n'}{m'}$ and $\frac{r}{s} = \frac{r'}{s'}$ we have

$$n \cdot m' = m \cdot n' \wedge r \cdot s' = s \cdot r' \tag{8.1}$$

$$\begin{aligned}
(n \cdot s + r \cdot m) \cdot (m' \cdot s') &= (n \cdot s) \cdot (m' \cdot s') + (r \cdot m) \cdot (m' \cdot s') \\
&\underset{[\text{theorem: } 7.11]}{=} (n \cdot m') \cdot (s \cdot s') + (r \cdot s') \cdot (m \cdot m') \\
&\underset{[\text{eq: } 8.1]}{=} (m \cdot n') \cdot (s \cdot s') + (s \cdot r') \cdot (m \cdot m') \\
&\underset{[\text{theorem: } 7.11]}{=} (n' \cdot s) \cdot (m \cdot s) + (r' \cdot m') \cdot (m \cdot s) \\
&\underset{[\text{theorem: } 7.11]}{=} (n' \cdot s + r' \cdot m') \cdot (m \cdot s) \\
&\underset{[\text{theorem: } 7.11]}{=} (m \cdot s) \cdot (n' \cdot s + r' \cdot m')
\end{aligned}$$

proving that

$$\frac{n \cdot s + r \cdot m}{m \cdot s} = \frac{n' \cdot s' + r' \cdot m'}{m' \cdot s'}$$ □

The above theorem ensures that the following is well-defined, independent of the representation.

Definition 8.7. *The sum operator* $+: \mathbb{Q} \times \mathbb{Q} \to \mathbb{Q}$ *is defined by*

$$\frac{n}{m} + \frac{r}{s} = \frac{n \cdot s + m \cdot r}{m \cdot s}$$

Theorem 8.8. $\langle \mathbb{Q}, + \rangle$ *is a* ***Abelian group*** *with neutral element* $0 = \frac{0}{1}$ *and for every* $\frac{n}{m} \in \mathbb{Q}$ *the inverse element* $\frac{-n}{m}$.

Proof.

associativity. Let $\frac{a}{b}, \frac{c}{d}, \frac{e}{f} \in \mathbb{Q}$ then

$$\begin{aligned}
\frac{a}{b} + \left(\frac{c}{d} + \frac{e}{f}\right) &= \frac{a}{b} + \frac{c \cdot f + d \cdot e}{d \cdot f} \\
&= \frac{a \cdot (d \cdot f) + b \cdot (c \cdot f + d \cdot e)}{b \cdot (d \cdot f)} \\
&\underset{[\text{theorem: } 7.11]}{=} \frac{(a \cdot d) \cdot f + (c \cdot b) \cdot f + (b \cdot d) \cdot e}{(b \cdot d) \cdot f} \\
&\underset{[\text{theorem: } 7.11]}{=} \frac{(a \cdot d + c \cdot b) \cdot f + (b \cdot d) \cdot e}{(b \cdot d) \cdot f} \\
&= \frac{a \cdot d + c \cdot b}{b \cdot d} + \frac{e}{f} \\
&= \left(\frac{a}{b} + \frac{c}{d}\right) + \frac{e}{f}
\end{aligned}$$

commutativity. Let $\frac{a}{b}, \frac{c}{d} \in \mathbb{Q}$ then

$$\begin{aligned}
\frac{a}{b} + \frac{c}{d} &= \frac{a \cdot d + b \cdot c}{b \cdot d} \\
&\underset{[\text{theorem: } 7.11]}{=} \frac{c \cdot b + d \cdot a}{d \cdot b} \\
&= \frac{c}{d} + \frac{a}{b}
\end{aligned}$$

neutral element. Let $\frac{a}{b} \in \mathbb{Q}$ then

$$\begin{aligned}\frac{a}{b}+\frac{0}{1} &\underset{\text{commutativity}}{=} \frac{a \cdot 1 + b \cdot 0}{b \cdot 1} \\ &\underset{\text{[theorem: 7.11] and [theorem: 4.39]}}{=} \frac{a}{b}\end{aligned}$$

inverse element. Let $\frac{a}{b} \in \mathbb{Q}$ then we have

$$\begin{aligned}\frac{a}{b}+\frac{-a}{b} &= \frac{-a}{b}+\frac{a}{b} \\ &= \frac{(-a) \cdot b + b \cdot a}{b \cdot b} \\ &\underset{\text{[theorem: 7.11]}}{=} \frac{b \cdot ((-a)+a)}{b \cdot b} \\ &= \frac{b \cdot 0}{b \cdot b} \\ &\underset{\text{[theorem: 4.39]}}{=} \frac{0}{b \cdot b} \\ &\underset{\text{[theorem: 8.5]}}{=} \frac{0}{1} \\ &= 0\end{aligned}$$

□

Definition 8.9. *If $x, y \in \mathbb{Q}$ then $x - y = x + (-y)$*

Next we define multiplication.

Theorem 8.10. *If $\frac{n}{m}, \frac{n'}{m'}, \frac{r}{s}, \frac{r'}{s'} \in \mathbb{Q}$ such that $\frac{n}{m} = \frac{n'}{m'}$ and $\frac{r}{s} = \frac{r'}{s'}$ then*

$$\frac{n \cdot r}{m \cdot s}, \frac{n' \cdot r'}{m' \cdot s'} \in \mathbb{Q} \text{ and } \frac{n \cdot r}{m \cdot s} = \frac{n' \cdot r'}{m' \cdot s'}$$

Proof. First as $m \neq 0, m' \neq 0, s \neq 0$ and $s' \neq 0$ we have that $m \cdot s \neq 0$ and $m' \cdot s' \neq 0$ so that $\frac{n \cdot r}{m \cdot s}, \frac{n' \cdot r'}{m' \cdot s'} \in \mathbb{Q}$. As $\frac{n}{m} = \frac{n'}{m'}$ and $\frac{r}{s} = \frac{r'}{s'}$ we have also that

$$n \cdot m' = m \cdot n' \wedge r \cdot s' = s \cdot r' \tag{8.2}$$

$$\begin{aligned}(n \cdot r) \cdot (m' \cdot s') &\underset{\text{[theorem: 7.11]}}{=} (n \cdot m') \cdot (r \cdot s') \\ &\underset{\text{[eq: 8.2]}}{=} (m \cdot n') \cdot (s \cdot r') \\ &= (m \cdot s) \cdot (n' \cdot r')\end{aligned}$$

so that

$$\frac{n \cdot r}{m \cdot s} = \frac{n' \cdot r'}{m' \cdot s'}$$

□

The above theorem ensures that the next definition is well defined.

Definition 8.11. *The product operator $\cdot: \mathbb{Q} \times \mathbb{Q} \to \mathbb{Q}$ is defined by*

$$\frac{n}{m} \cdot \frac{r}{s} = \frac{n \cdot r}{m \cdot s}$$

Theorem 8.12. *$\langle \mathbb{Q}, +, \cdot \rangle$ is a field [see definition: 4.51] more specifically:*

1. *$\langle \mathbb{Q}, + \rangle$ is a Abelian group [see theorem: 8.8]*
2. *$\cdot: \mathbb{Q} \times \mathbb{Q} \to \mathbb{Q}$ satisfies*

   distributivity. *$\forall x, y, z \in \mathbb{Q}$ we have $x \cdot (y + z) = x \cdot y + x \cdot z$*

   commutativity. *$\forall x, y \in \mathbb{Q}$ we have $x \cdot y = y \cdot x$*

neutral element. *$\forall x \in \mathbb{Q}$ $\frac{1}{1} \cdot x = 1 = x \cdot \frac{1}{1}$, so $1 \underset{\text{definition}}{=} \frac{1}{1}$ is the neutral element.*

associativity. *$\forall x, y, z \in \mathbb{Q}$ $(x \cdot y) \cdot z = x \cdot (y \cdot z)$*

inverse element. *$\forall x \in \mathbb{Q} \setminus \{0\} = \mathbb{Q}^*$ there exists a $x^{-1} \cdot x = x \cdot x^{-1}$. More specific if $x = \frac{a}{b} \neq 0$ then $x^{-1} = \frac{b}{a}$.*

3. $0 \neq 1$

Proof.

1. This follows from [theorem: 8.8].
2. We have:

distributivity. Let $\frac{a}{b}, \frac{c}{d}, \frac{e}{f} \in \mathbb{Q}$ then

$$\begin{aligned}
\frac{a}{b} \cdot \frac{c}{d} + \frac{a}{b} \cdot \frac{e}{f} &= \frac{a \cdot c}{b \cdot d} + \frac{a \cdot e}{b \cdot f} \\
& \frac{(a \cdot c) \cdot (b \cdot f) + (b \cdot d) \cdot (a \cdot e)}{(b \cdot d) \cdot (b \cdot f)} \\
&\underset{[\text{theorem: } 7.11]}{=} \frac{b \cdot (a \cdot (c \cdot f)) + b \cdot (a \cdot (d \cdot e))}{b \cdot (b \cdot (d \cdot f))} \\
&\underset{[\text{theorem: } 7.11]}{=} \frac{b \cdot (a \cdot (c \cdot f) + a \cdot (d \cdot e))}{b \cdot (b \cdot (d \cdot f))} \\
&\underset{b \neq 0 \wedge [\text{theorem: } 8.5]}{=} \cdot \frac{a \cdot (c \cdot f) + a \cdot (d \cdot e)}{b \cdot (d \cdot f)} \\
&= \frac{a}{b} \cdot \left(\frac{c \cdot f + d \cdot e}{d \cdot f} \right) \\
&= \frac{a}{b} \cdot \left(\frac{c}{d} + \frac{e}{f} \right)
\end{aligned}$$

commutativity. Let $\frac{a}{b}, \frac{c}{d} \in \mathbb{Q}$ then

$$\begin{aligned}
\frac{a}{b} \cdot \frac{c}{d} &= \frac{a \cdot c}{b \cdot d} \\
&\underset{[\text{theorem: } 7.11]}{=} \frac{c \cdot a}{d \cdot b} \\
&= \frac{c}{d} \cdot \frac{a}{b}
\end{aligned}$$

neutral element. Let $\frac{a}{b} \in \mathbb{Q}$ then

$$\begin{aligned}
\frac{1}{1} \cdot \frac{a}{b} &\underset{\text{commutativity}}{=} \frac{a}{b} \cdot \frac{1}{1} \\
&= \frac{a \cdot 1}{b \cdot 1} \\
&\underset{[\text{theorem: } 7.11]}{=} \frac{a}{b}
\end{aligned}$$

associativity. Let $\frac{a}{b}, \frac{c}{d}, \frac{e}{f} \in \mathbb{Q}$ then

$$\begin{aligned}
\frac{a}{b} \cdot \left(\frac{c}{d} \cdot \frac{e}{f} \right) &= \frac{a}{b} \cdot \frac{c \cdot e}{d \cdot f} \\
&= \frac{a \cdot (c \cdot e)}{b \cdot (d \cdot f)} \\
&\underset{[\text{theorem: } 7.11]}{=} \frac{(a \cdot c) \cdot e}{(b \cdot d) \cdot f} \\
&= \frac{a \cdot c}{b \cdot d} \cdot \frac{e}{f} \\
&= \left(\frac{a}{b} \cdot \frac{c}{d} \right) \cdot \frac{e}{f}
\end{aligned}$$

inverse element. Let $\frac{a}{b}\in\mathbb{Q}\setminus\{0\}$ then $\frac{a}{b}\neq\frac{0}{1}\underset{[\text{theorem: }8.5]}{\Rightarrow} a\neq 0$ so that $\frac{b}{a}\in\mathbb{Q}$, then

$$\begin{aligned}\frac{a}{b}\cdot\frac{b}{a} &\underset{\text{commutativity}}{=} \frac{b}{a}\cdot\frac{a}{b}\\ &= \frac{b\cdot a}{a\cdot b}\\ &\underset{[\text{theorem: }7.11]}{=} \frac{1\cdot(a\cdot b)}{1\cdot(a\cdot b)}\\ &\underset{a\cdot b\neq 0\wedge[\text{theorem: }8.5]}{=} \frac{1}{1}\end{aligned}$$

3. If $\frac{0}{1}=\frac{1}{1}$ then $0\cdot 1=1\cdot 1$ o that $0=1$ which is impossible by [theorem: 5.10]. □

Example 8.13. $1+1=2$ and $2^{-1}=\frac{1}{2}$ where $2=\frac{2}{1}$.

Proof. $\frac{1}{1}+\frac{1}{1}=\frac{1\cdot 1+1\cdot 1}{1\cdot 1}=\frac{1+1}{1}\underset{[\text{theorem: }7.12]}{=}\frac{2}{1}=2$, so $2^{-1}=\left(\frac{2}{1}\right)^{-1}=\frac{1}{2}$ □

Theorem 8.14. *Let $q,r\in\mathbb{Q}$ and $s\neq 0$ then*

1. $q=r\Leftrightarrow q\cdot s=r\cdot s$
2. $q\neq r\Leftrightarrow q\cdot s\neq r\cdot s$

Proof.

1.

⇒. If $q=r$ then $q\cdot s=r\cdot s$

⇐. We have

$$\begin{aligned} q\cdot s=r\cdot s &\underset{s\neq 0}{\Rightarrow} (q\cdot s)\cdot s^{-1}=(r\cdot s)\cdot s^{-1}\\ &\underset{[\text{theorem: }8.12]}{\Rightarrow} q\cdot(s\cdot s^{-1})=r\cdot(s\cdot s^{-1})\\ &\underset{[\text{theorem: }8.12]}{\Rightarrow} q\cdot 1=r\cdot 1\\ &\underset{[\text{theorem: }8.12]}{\Rightarrow} q\cdot s=r\cdot s\end{aligned}$$

2. This follows by contra-position. □

8.2 Order Relation

Definition 8.15. *The set of non negative rational numbers $\mathbb{Q}_0^+$ and the set of non positive numbers $\mathbb{Q}_0^-$ is defined by:*

$$\begin{aligned}\mathbb{Q}_0^+ &= \left\{\frac{a}{b}|(a,b)\in\mathbb{Z}\times\mathbb{Z}^*\wedge a\cdot b\in\mathbb{Z}_0^+\right\}\underset{[\text{theorem: }7.22]}{=}\left\{\frac{a}{b}|(a,b)\in\mathbb{Z}\times\mathbb{Z}^*\wedge 0\leqslant a\cdot b\right\}\\ \mathbb{Q}_0^- &= \left\{\frac{a}{b}|(a,b)\in\mathbb{Z}\times\mathbb{Z}^*\wedge a\cdot b\in\mathbb{Z}_0^-\right\}\underset{[\text{theorem: }7.22]}{=}\left\{\frac{a}{b}|(a,b)\in\mathbb{Z}\times\mathbb{Z}^*\wedge a\cdot b\leqslant 0\right\}\end{aligned}$$

Theorem 8.16. $\mathbb{Q}=\mathbb{Q}_0^+\bigcup\mathbb{Q}_0^-$ *and* $\{0\}=\mathbb{Q}_0^+\bigcap\mathbb{Q}_0^-$

Proof. If $q\in\mathbb{Q}$ then $\exists(a,b)\in\mathbb{Z}\times\mathbb{Z}^*$ such that $q=\frac{a}{b}$, as $a\cdot b\in\mathbb{Z}\underset{[\text{theorem: }7.20]}{=}\mathbb{Z}_0^+\bigcup\mathbb{Z}_0^-$ we have either:

$\boldsymbol{a\cdot b\in\mathbb{Z}_0^+}$**.** Then $q=\frac{a}{b}\in\mathbb{Q}_0^+$

$\boldsymbol{a\cdot b\in\mathbb{Z}_0^-}$**.** Then $q=\frac{a}{b}\in\mathbb{Q}_0^-$

proving that

$$\mathbb{Q}\subseteq\mathbb{Q}^+\bigcup\mathbb{Q}_0^-$$

As trivially $\mathbb{Q}_0^+ \subseteq \mathbb{Q}$ and $\mathbb{Q}_0^- \subseteq \mathbb{Q}$ we have that $\mathbb{Q}_0^+ \bigcup \mathbb{Q}_0^- \subseteq \mathbb{Q}$, which by the above proves that

$$\mathbb{Q} = \mathbb{Q}_0^+ \bigcup \mathbb{Q}_0^-$$

If $q \in \{0\}$ then $q = \frac{0}{1}$ so that $0 \cdot 1 = 0 \in \mathbb{Z}_0^+$ and $0 \cdot 1 = 0 \in \mathbb{Z}_0^-$ so that $q \in \mathbb{Q}_0^+ \bigcap \mathbb{Q}_0^-$ proving that

$$\{0\} \subseteq \mathbb{Q}_0^+ \bigcap \mathbb{Q}_0^- \tag{8.3}$$

If $q \in \mathbb{Q}_0^+ \bigcap \mathbb{Q}_0^-$ then there exist $(a, b), (c, d) \in \mathbb{Z} \times \mathbb{Z}^* \Rightarrow b \neq 0 \neq d$ such that $a \cdot b \in \mathbb{Z}_0^+ \Rightarrow 0 \leqslant a \cdot b$, $c \cdot d \in \mathbb{Z}_0^- \Rightarrow c \cdot d \leqslant 0$ and $\frac{a}{b} = \frac{c}{d} \Rightarrow a \cdot d = b \cdot c$. Assume that $a \neq 0$ then we have either:

$0 < a$. Assume that $b < 0$ then by [theorems: 7.27, 4.50] $a \cdot b < 0$ contradicting $0 \leqslant a \cdot b$, so we must have that $0 \leqslant b$ which as $b \neq 0$ gives

$$0 < b \tag{8.4}$$

As $d \neq 0$ we have by [theorems: 7.27, 4.50] that $0 < d \cdot d$ so that by [theorems: 7.27, 4.50]

$$0 < a \cdot (d \cdot d) = (a \cdot d) \cdot d \underset{a \cdot d = b \cdot c}{=} (b \cdot c) \cdot d = (c \cdot d) \cdot b \tag{8.5}$$

Using [theorems: 7.27, 4.50] on [eq: 8.4] and [eq: 8.5] we have that $0 < c \cdot d$ contradicting $c \cdot d \leqslant 0$.

$a < 0$. Assume that $0 < b$ then by [theorems: 7.27, 4.50] $a \cdot b < 0$ contradicting $0 \leqslant a \cdot b$, so we must have that $0 \leqslant b$ which as $b \neq 0$ gives

$$b < 0 \tag{8.6}$$

As $d \neq 0$ we have by [theorems: 7.27, 4.50] that $0 < d \cdot d$ so that by [theorems: 7.27, 4.50] $a \cdot (d \cdot d) < 0$, hence as $a \cdot d = b \cdot c$

$$(c \cdot d) \cdot b = (b \cdot c) \cdot d = (a \cdot d) \cdot d = a \cdot (d \cdot d) < 0 \tag{8.7}$$

Using [theorems: 7.27, 4.50] on [eq: 8.6] and [eq: 8.7] we have that $0 < c \cdot d$ contradicting $c \cdot d \leqslant 0$.

As in all cases we reach a contradiction the assumption $a \neq 0$ is wrong, so $a = 0$ or $q = \frac{0}{b} \underset{\text{[theorem: 8.5]}}{=} \frac{0}{1} = 0$. Hence $\mathbb{Q}_0^+ \bigcap \mathbb{Q}_0^- \subseteq \{0\}$ which combined with [eq: 8.3] proves that

$$\mathbb{Q}_0^+ \bigcap \mathbb{Q}_0^- = \{0\}$$

□

Theorem 8.17. $\mathbb{Q}_0^- = \{-q | q \in \mathbb{Q}_0^+\}$

Proof. If $q \in \mathbb{Q}_0^-$ then $\exists (n, m) \in \mathbb{Z} \times \mathbb{Z}^*$ with $n \cdot m \in \mathbb{Z}_0^-$ such that $q = \frac{n}{m}$. By the definition of $\mathbb{Z}_0^-$ it follows that $\exists k \in \mathbb{Z}_0^+$ such that $n \cdot m = -k$. Hence

$$(-n) \cdot m \underset{\text{[theorem: 4.40]}}{=} -(n \cdot m) = -(-(k)) \underset{\text{[theorem: 4.9]}}{=} k \in \mathbb{Z}_0^+,$$

proving that $-q = \frac{-n}{m} \in \mathbb{Q}_0^+$. Using [theorem: 4.9] we have $q = -(-q)$ so that $q \in \{-q | q \in \mathbb{Q}_0^+\}$ or that

$$\mathbb{Q}_0^- \subseteq \{-q | q \in \mathbb{Q}_0^+\} \tag{8.8}$$

If $q \in \{-q | q \in \mathbb{Q}_0^+\}$ then $\exists (n, m) \in \mathbb{Z} \times \mathbb{Z}^*$ with $n \cdot m \in \mathbb{Z}_0^+$ such that $q = -\frac{n}{m} = \frac{-n}{m}$, as $(-n) \cdot m = -(n, m) \in \mathbb{Z}_0^-$, it follows that $q \in \mathbb{Q}_0^-$. Hence $\{-q | q \in \mathbb{Q}_0^+\} \subseteq \mathbb{Q}_0^-$ which together with [eq: 8.8] gives

$$\mathbb{Q}_0^- = \{-q | q \in \mathbb{Q}_-^+\}$$

□

Theorem 8.18. *$\langle \mathbb{Q}_0^+, + \rangle$ is a sub semi-group of $\langle \mathbb{Q}, + \rangle$ [hence $\langle \mathbb{Q}_0^+, + \rangle$ is a semi-group]*

Proof. By [theorem: 8.16]

$$0 \in \mathbb{Q}_0^+ \tag{8.9}$$

If $q, r \in \mathbb{Q}_0^+$ then there exists $(a, b), (c, d) \in \mathbb{Z} \times \mathbb{Z}^*$ with $a \cdot b, c \cdot d \in \mathbb{Z}_0^+$ such that $q = \frac{a}{b}$ and $\frac{c}{d}$. Then we have

$$q + r = \frac{a \cdot d + b \cdot c}{b \cdot d}$$

so we have to check that $(a \cdot d + b \cdot c) \cdot (b \cdot d) \in \mathbb{Z}_0^+$. Hence

$$\begin{aligned} (a \cdot d + b \cdot c) \cdot (b \cdot d) &\underset{\text{[theorem: 7.11]}}{=} (a \cdot d) \cdot (b \cdot d) + (b \cdot c) \cdot (b \cdot d) \\ &\underset{\text{[theorem: 7.11]}}{=} (a \cdot b) \cdot (d \cdot d) + (c \cdot d) \cdot (b \cdot b) \end{aligned} \tag{8.10}$$

Now using [theorems: 7.27, 4.50] we have that $0\leqslant d\cdot d\wedge 0\leqslant b\cdot b$, as $0\leqslant a\cdot b\wedge 0\leqslant c\cdot d$, we have by [theorems: 7.27, 4.50] that $0\leqslant (a\cdot b)\cdot(d\cdot d)\wedge 0\leqslant (c\cdot d)\cdot(b\cdot b)$ or $(a\cdot b)\cdot(d\cdot d),(c\cdot d)\cdot(b\cdot b)\in\mathbb{Z}_0^+$. Using [theorem: 7.17] it follows that $(a\cdot b)\cdot(d\cdot d)+(c\cdot d)\cdot(b\cdot b)\in\mathbb{Z}_0^+$, hence by [eq: 8.10] $(a\cdot d+b\cdot c)\cdot(b\cdot d)\in\mathbb{Z}_0^+$ so that $q+r\in\mathbb{Q}_0^+$. So

$$\forall q,r\in\mathbb{Q}_0^+ \text{ we have } q+r\in\mathbb{Q}_0^+ \tag{8.11}$$

Finally [eq: 8.9] and [eq: 8.11] proves that $\langle\mathbb{Q}_0^+,+\rangle$ is a semi-group. □

Next we define the relation that will later become a order relation on $\mathbb{Q}$.

Definition 8.19. (Order Relation) $\leqslant\subseteq\mathbb{Q}\times\mathbb{Q}$ *is defined as*

$$\leqslant=\{(q,r)\in\mathbb{Q}\times\mathbb{Q}|r+(-q)\in\mathbb{Q}_0^+\}$$

So $q\leqslant r$ *if and only if* $r+(-q)\in\mathbb{Q}_0^+$

Theorem 8.20. $\mathbb{Q}_0^+=\{q\in\mathbb{Q}|0\leqslant q\}$ *and* $\mathbb{Q}_0^-=(q\in\mathbb{Q}|q\leqslant 0)$.

Proof.

$$\begin{aligned} q\in\mathbb{Q}_0^+ &\underset{q=q+(-0)}{\Leftrightarrow} q+(-0)\in\mathbb{Q}_0^+\\ &\underset{\mathbb{Q}_0^+\subseteq\mathbb{Q}}{\Leftrightarrow} q\in\mathbb{Q}\wedge 0\leqslant q\\ &\Leftrightarrow q\in\{q\in Q|0\leqslant q\}\end{aligned}$$

proving that

$$\mathbb{Q}_0^+=\{q\in\mathbb{Q}|0\leqslant q\}$$

Further

$$\begin{aligned} q\in\mathbb{Q}_0^- &\underset{[\text{theorem: } 8.17]}{\Leftrightarrow} -q\in\mathbb{Q}_0^+\\ &\underset{0+(-q)=-q}{\Leftrightarrow} 0+(-q)\in\mathbb{Q}_0^+\\ &\underset{\mathbb{Q}_0^+\subseteq\mathbb{Q}}{\Leftrightarrow} q\in\mathbb{Q}\wedge q\leqslant 0\\ &\Leftrightarrow q\in\{q\in\mathbb{Q}|q\leqslant 0\}\end{aligned}$$

proving that

$$\mathbb{Q}_0^-=\{q\in\mathbb{Q}|q\leqslant 0\}$$

□

Theorem 8.21. *If* $q,r\in\mathbb{Q}$ *then*

1. $q\leqslant r\Leftrightarrow 0\leqslant r+(-q)$
2. $q<r\Leftrightarrow 0<r+(-q)$

Proof.

1. We have

$$\begin{aligned} q\leqslant r &\Leftrightarrow r+(-q)\in\mathbb{Q}_0^+\\ &\underset{[\text{theorem: } 8.20]}{\Leftrightarrow} 0\leqslant r+(-q)\end{aligned}$$

2. We have

$$\begin{aligned} q<r &\Leftrightarrow q\neq r\wedge q\leqslant r\\ &\Leftrightarrow q+(-q)\neq r+(-q)\wedge q\leqslant r\\ &\Leftrightarrow 0\neq r+(-q)\wedge q\leqslant r\\ &\underset{(1)}{\Leftrightarrow} 0\neq r+(-q)\wedge 0\leqslant r+(-q)\\ &\Leftrightarrow 0<r+(-q)\end{aligned}$$

□

Theorem 8.22. *If $q \in \mathbb{Q}$ satisfies $0 \leqslant q \wedge q \leqslant 0$ then $q = 0$*

Proof. As $0 \leqslant q$ and $q \leqslant 0$ we have by [theorem: 8.20] that $q \in \mathbb{Q}_0^+ \wedge \mathbb{Q}_0^-$, so $q \in \mathbb{Q}_0^+ \bigcap \mathbb{Q}_0^- \underset{[\text{theorem: } 8.16]}{=} \{0\}$. Proving $q = 0$. □

Theorem 8.23. *$\langle \mathbb{Q}, \leqslant \rangle$ is a totally ordered set.*

Proof.

reflectivity. If $q \in \mathbb{Q}_0^+$ then $q + (-q) = 0 \in \{0\} \underset{[\text{theorem: } 8.16]}{=} \mathbb{Q}_0^+ \bigcap \mathbb{Q}_0^- \subseteq \mathbb{Q}_0^+$ so that $q \leqslant q$.

anti symmetry. If $q \leqslant r$ and $r \leqslant q$ then $r + (-q) \in \mathbb{Q}_0^+$, $q + (-r) \in \mathbb{Q}_0^+ \Rightarrow -(q + (-r)) \in \mathbb{Q}_0^-$ so that by [theorem: 8.20] $0 \leqslant r + (-q)$ and $r + (-q) = -(q + (-r)) \leqslant 0$. Using [theorem: 8.22] it follows that $r + (-q) = 0$ so that $r = q$.

transitivity. If $q \leqslant r$ and $r \leqslant s$ then $r + (-q), s + (-r) \in \mathbb{Q}_0^+$ so that by [theorem: 8.18] we have that

$$(r + (-q)) + (s + (-r)) \in \mathbb{Q}_0^+ \tag{8.12}$$

As $(r + (-q)) + (s + (-r)) \underset{[\text{theorem: } 8.8]}{=} s + (-q)$ we have by [eq: 8.12] that $s + (-q) \in \mathbb{Q}_0^+$ proving that

$$q \leqslant s$$

totally order. If $q, r \in \mathbb{Q}$ then as $r + (-q) \in \mathbb{Q} \underset{[\text{theorem: } 8.16]}{=} \mathbb{Q}_0^+ \bigcup \mathbb{Q}_0^-$ we have the following possibilities:

$r + (-q) \in \mathbb{Q}_0^+$. Then $q \leqslant r$

$r + (-q) \in \mathbb{Q}_0^-$. Then by [theorem: 8.17] we have that $-(r + (-q)) \in \mathbb{Q}_0^+$. Further

$$-(r + (-q)) \underset{[\text{theorem: } 4.40]}{=} (-r) + (-(-q)) = q + (-r)$$

so that $q + (-r) \in \mathbb{Q}_0^+$ or $r \leqslant q$. □

Lemma 8.24. *Let $q \in \mathbb{Q}$ then $0 < q \Leftrightarrow \exists n, m \in \mathbb{Z}$ with $0 < n \wedge 0 < m$ such that $q = \frac{n}{m}$*

Proof.

⇒. As $0 < q$ we have $0 \neq q$ and $0 \leqslant q \underset{[\text{theorem: } 8.20]}{\Rightarrow} q \in \mathbb{Q}_0^+$, so there exists $(n', m') \in \mathbb{Z} \times \mathbb{Z}^*$ with $0 \leqslant n' \cdot m'$ such that $q = \frac{n'}{m'}$, as $m' \in \mathbb{Z}^*$ we have $m' \neq 0$, further by [theorem: 8.5] $n' \neq 0$. So we have the following resting cases to consider for n', m':

$0 < n' \wedge 0 < m'$. Then $q = \frac{n'}{m'}$ so if we take $n = n'$ and $m = m'$ we have $0 < n \wedge 0 < m$ such that $q = \frac{n}{m}$.

$0 < n' \wedge m' < 0$. Then by [theorems: 7.27, 4.50] we have $n' \cdot m' < 0$ contradicting $0 \leqslant n' \cdot m'$ so this is not a valid case.

$n' < 0 \wedge 0 < m'$. Then by [theorems: 7.27, 4.50] we have $n' \cdot m' < 0$ contradicting $0 \leqslant n' \cdot m'$ so this is not a valid case.

$n' < 0 \wedge m' < 0$. Then by [theorem: 7.27, 4.50] we have $0 < -n' \wedge 0 < -m'$ so that $\frac{-n'}{-m'} = \frac{n' \cdot (-1)}{m' \cdot (-1)} \underset{[\text{theorem: } 8.5}{=} \frac{n'}{m'} = q$. So if we take $n = -n'$ and $m = -m'$ then $0 < n \wedge 0 < m$ and $q = \frac{n}{m}$.

So in all valid cases we found a $n, m \in \mathbb{Z}$ with $0 < n \wedge 0 < m$ and $q = \frac{n}{m}$.

⇐. If $\exists n, m \in \mathbb{Z}$ with $0 < n \wedge 0 < m$ such that $q = \frac{n}{m}$ then by [theorem: 7.26] we have that $0 < n \cdot m$ so that $0 \leqslant q$, further by [theorem: 8.5] and $n \neq 0$ we have $q \neq 0$, hence $0 < q$. □

Example 8.25. $0 < 1$ where $0, 1 \in \mathbb{Q}$

Proof. As $1 = \frac{1}{1} \in \mathbb{Q}$ and $1 = 1 \cdot 1 \in \mathbb{Z}$ and $0 < 1$ [see example: 7.25] it follows from [lemma: 8.24] that $0 < 1$. □

Corollary 8.26. *If $q \in \mathbb{Q}$ then*

1. $q < q + 1$
2. $q - 1 = q + (-1) < q$

Proof.

1. $(q + 1) + (-q) = (q + (-q)) + 1 = 1$ and by [example: 8.25] $0 < 1$ so that $q \leqslant q + 1$. If $q = q + 1$ we have $0 = q + (-q) = (q + 1) + (-q) = 1$ contradicting $0 < 1$ so we must have that

$$q < q + 1$$

2. $q+(-(q+(-1))) \underset{[\text{theorem: } 4.8]}{=} q+((-q)+(-(-1)))=(q+(-q))+(-(-1))=0+1=1$ and by [example: 8.25] $0<1$ so that $q+(-1)\leqslant q$. If $q+(-1)=q$ then $q=q+1$ contradicting (1) so $q-1\neq q$ and we have

$$q-1=q-1<q$$

□

Theorem 8.27. *Let $n\in\mathbb{Z}$ and $m\in\mathbb{Z}^*$ then we have*

1. $n=m\Leftrightarrow\frac{n}{m}=1$
2. *If* $0<m$ *then*
   a. $n<m\Leftrightarrow\frac{n}{m}<1$
   b. $m<n\Leftrightarrow 1<\frac{n}{m}$
   c. $n\leqslant m\Leftrightarrow\frac{n}{m}\leqslant 1$
   d. $m\leqslant n\Leftrightarrow 1\leqslant\frac{n}{m}$
3. *If* $m<0$ *then*
   a. $n<m\Leftrightarrow 1<\frac{n}{m}$
   b. $m<n\Leftrightarrow\frac{n}{m}<1$
   c. $n\leqslant m\Leftrightarrow 1\leqslant\frac{n}{m}$
   d. $m\leqslant n\Leftrightarrow\frac{n}{m}\leqslant 1$

Proof.

1.

   ⇒. If $n=m$ then $\frac{n}{m}=\frac{n}{n}=\frac{1\cdot n}{1\cdot n}\underset{[\text{theorem: } 8.5]}{=}\frac{1}{1}=1$.

   ⇐. If $\frac{n}{m}=1$ then $\frac{n}{m}=\frac{1}{1}$ so that $n\cdot 1=m\cdot 1$ proving that $n=m$.

2.

   a. Note that

$$1+\left(-\frac{n}{m}\right)=\frac{1}{1}+\frac{-n}{m}=\frac{m+(-n)}{m}$$

   So

$$\begin{aligned} n<m &\Leftrightarrow 0<m+(-n)\\ &\underset{[\text{theorem: } 8.24]}{\Leftrightarrow} 0<\frac{m+(-n)}{m}\\ &\Leftrightarrow 1+\left(-\frac{n}{m}\right)\\ &\Leftrightarrow \frac{n}{m}<1 \end{aligned}$$

   b. Note that

$$\frac{n}{m}+(-1)=\frac{n}{m}+\frac{-1}{1}=\frac{n+(-m)}{m}$$

   So

$$\begin{aligned} m<n &\Leftrightarrow 0<m+(-n)\\ &\underset{[\text{theorem: } 8.24]}{\Leftrightarrow} 0<\frac{m+(-n)}{m}\\ &\Leftrightarrow 0<\frac{n}{m}+(-1)\\ &\Leftrightarrow 1<\frac{n}{m} \end{aligned}$$

   c. This follows from (1) and (2.a)

   d. This follows from (1) and (2.b)

3. As $m<0$ we have by [theorem: 7.27, 4.50] that $0<-m$ so we have
   a. Then
$$\begin{array}{rcl} n<m & \underset{[\text{theorem: } 7.27,\ 4.50]}{\Leftrightarrow} & -m<-n \\ & \underset{(2.b)}{\Leftrightarrow} & 1<\dfrac{-n}{-m} \\ & \underset{[\text{theorem: } 7.27,\ 4.40]}{\Leftrightarrow} & 1<\dfrac{(-1)\cdot n}{(-1)\cdot m} \\ & \underset{[\text{theorem: } 8.5]}{=} & 1<\dfrac{n}{m} \end{array}$$
   b. Then
$$\begin{array}{rcl} m<n & \underset{[\text{theorem: } 7.27,\ 4.50]}{\Leftrightarrow} & -n<-m \\ & \underset{(2,a)}{\Leftrightarrow} & \dfrac{-n}{-m}<1 \\ & \underset{[\text{theorem: } 7.27,\ 4.40]}{\Leftrightarrow} & \dfrac{(-1)\cdot n}{(-1)\cdot m}<1 \\ & \underset{[\text{theorem: } 8.5]}{=} & \dfrac{n}{m}<1 \end{array}$$
   c. This follows from (1) and (3.a)
   d. This follows from (1) and (3.b) □

Theorem 8.28. *If $q,r\in\mathbb{Q}$ such that $0<q$ and $0<r$ then $0<q\cdot r$.*

Proof. As $0<q\wedge 0<r$ we have by [lemma: 8.24] the existence of $a,b,c,d\in\mathbb{Z}$ with $0<a$, $0<b$, $0<c$, $0<d$ such that $q=\frac{a}{b}$ and $r=\frac{c}{d}$. So by applying [theorems: 7.27, 4.50] we have $0<a\cdot c\wedge 0<b\cdot d$, hence $0<(a\cdot c)\cdot(b\cdot d)$, so that $q\cdot r=\frac{a\cdot c}{b\cdot d}\in\mathbb{Q}_0^+$ or $0\leqslant q\cdot r$. As $0<a\cdot c$ we have by [theorem: 8.5] that $q\cdot r\neq 0$, so
$$0<q\cdot r$$
□

Theorem 8.29. *$\langle\mathbb{Q},+,\cdot,\leqslant\rangle$ is a ordered field*

Proof. First using [theorem: 8.12] $\langle\mathbb{Q},+,\cdot\rangle$ is a field. Next

1. For $r,q,s\in\mathbb{Q}$ with $r<q$ we have
$$q+(-r)\in\mathbb{Q}_0^+ \text{ and } r\neq q\Rightarrow r+s\neq q+s \tag{8.13}$$
   Further
$$\begin{array}{rcl} (q+s)+(-(r+s)) & \underset{[\text{theorem: } 4.8]}{=} & (q+s)+((-r)+(-s)) \\ & \underset{\text{commutativity}}{=} & (q+s+((-s)+(-r)) \\ & \underset{\text{associativity}}{=} & q+(s+((-s)+(-r))) \\ & \underset{\text{associativity}}{=} & q+((s+(-s))+(-r)) \\ & = & q+(0+(-r)) \\ & = & q+(-r) \end{array}$$
   which by [eq: 8.13] proves that $(q+s)+(-(r+s))\in\mathbb{Q}_0^+$. Hence $r+s\leqslant q+s$ and $r+s\neq q+s$ proving that
$$r+s<q+s$$
2. Let $q,r\in\mathbb{Q}$ with $0<q$ and $0<r$ then by [theorem: 8.28] $0<q\cdot r$. □

Theorem 8.30. *$\forall q\in\mathbb{Q}$ with $q\neq 0$ we have $-(q^{-1})=(-q)^{-1}$*

Proof. If $q\neq 0$ then $-q\neq 0$ so q^{-1} and $(-q)^{-1}$ exists, further $q=\frac{a}{b}$ where $a,b\neq 0$. Now
$$-(q^{-1})=-\left(\frac{b}{a}\right)=\left(\frac{-b}{a}\right)=\left(\frac{a}{-b}\right)^{-1}=\left(\frac{-1}{-1}\cdot\frac{a}{-b}\right)^{-1}=\left(\frac{-a}{b}\right)^{-1}=(-q)^{-1}$$
□

Next we embed the set of integer numbers in the set of rational numbers.

Definition 8.31. $\mathbb{Z}_{\mathbb{Q}} = \left\{\frac{z}{1} | z \in \mathbb{Z}\right\} \subseteq \mathbb{Q}$

Theorem 8.32. *$\mathbb{Z}_{\mathbb{Q}}$ is a sub-ring [see definition: 4.35] of $\langle \mathbb{Q}, +, \cdot \rangle$ and for*

$$i_{\mathbb{Z}\rightarrow\mathbb{Q}}: \mathbb{Z} \rightarrow \mathbb{Z}_{\mathbb{Q}} \text{ defined by } i_{\mathbb{Z}\rightarrow\mathbb{Q}}(z) = \frac{z}{1}$$

we have

$$i_{\mathbb{Z}->\mathbb{Q}}: \langle \mathbb{Z}, \leqslant \rangle \rightarrow \langle \mathbb{Z}_{\mathbb{Q}}, \leqslant \rangle \text{ is a order isomorphism}$$

and

$$i_{\mathbb{Z}\rightarrow\mathbb{Q}}: \langle \mathbb{Z}, +, \cdot \rangle \rightarrow \langle \mathbb{Z}_{\mathbb{Q}}, +, \cdot \rangle \text{ is a ring isomorphism}$$

Proof. If $q, r \in \mathbb{Z}_{\mathbb{Q}}$ then $\exists n, m \in \mathbb{Z}$ such that $q = \frac{n}{1}$ and $r = \frac{m}{1}$ so we have

$$q + r = \frac{n}{1} + \frac{m}{1} = \frac{n \cdot 1 + 1 \cdot m}{1 \cdot 1} = \frac{n + m}{1}$$

proving that $q + r \in \mathbb{Z}_{\mathbb{Q}}$. Further we have

$$q \cdot r = \frac{n}{1} \cdot \frac{m}{1} = \frac{n \cdot m}{1 \cdot 1} = \frac{n \cdot m}{1}$$

proving that $q \cdot r \in \mathbb{Z}_{\mathbb{Q}}$. Also $-q = \frac{-n}{1} \in \mathbb{Z}_{\mathbb{Q}}$. So we have

$$\forall q, r \in \mathbb{Z}_{\mathbb{Q}} \text{ we have } q + r \in \mathbb{Z}_{\mathbb{Q}},\ q \cdot r \in \mathbb{Z}_{\mathbb{Q}} \text{ and } -q \in \mathbb{Z}_{\mathbb{Q}}$$

Further we have $0 = \frac{0}{1}$ and $1 = \frac{1}{1}$ so that

$$0, 1 \in \mathbb{Z}_{\mathbb{Q}}$$

Hence we have that

$$\mathbb{Z}_{\mathbb{Q}} \text{ is a sub-ring of } \langle \mathbb{Q}, +, \cdot \rangle$$

Now for $i_{\mathbb{Z}\rightarrow\mathbb{Q}}: \mathbb{Z} \rightarrow \mathbb{Z}_{\mathbb{Q}}$ we have:

injectivity. If $i_{\mathbb{Z}\rightarrow\mathbb{Q}}(x) = i_{\mathbb{Z}\rightarrow\mathbb{Q}}(y)$ then $\frac{x}{1} = \frac{y}{1}$ so that $x \cdot 1 = 1 \cdot y \Rightarrow x = y$.

surjectivity. This follows from the definition of $\mathbb{Z}_{\mathbb{Q}}$.

proving that

$$i_{\mathbb{Z}\rightarrow\mathbb{Q}}: \mathbb{Z} \rightarrow \mathbb{Z}_{\mathbb{Q}} \text{ is a bijection} \tag{8.14}$$

Further if $x, y \in \mathbb{Z}$ then $i_{\mathbb{Z}\rightarrow\mathbb{Q}}(y) + (-i_{\mathbb{Z}\rightarrow\mathbb{Q}}(x)) = \frac{y}{1} + \left(-\frac{x}{1}\right) = \frac{y}{1} + \frac{-x}{1} = \frac{y \cdot 1 + 1 \cdot (-x)}{1 \cdot 1} = \frac{y + (-x)}{1}$ so

$$i_{\mathbb{Z}\rightarrow\mathbb{Q}}(y) + (-i_{\mathbb{Z}\rightarrow\mathbb{Q}}(x)) = \frac{y + (-x)}{1} \tag{8.15}$$

Hence we have

$$\begin{aligned}
i_{\mathbb{Z}\rightarrow\mathbb{Q}}(x) \leqslant i_{\mathbb{Z}\rightarrow\mathbb{Q}}(y) \quad &\Leftrightarrow \quad i_{\mathbb{Z}\rightarrow\mathbb{Q}}(y) + (-i_{\mathbb{Z}\rightarrow\mathbb{Q}}(x)) \in \mathbb{Q}_0^+ \\
&\underset{\text{[eq: 8.15]}}{\Leftrightarrow} \quad \frac{y + (-x)}{1} \in \mathbb{Q}_0^+ \\
&\underset{\text{def } \mathbb{Q}_0^+}{\Leftrightarrow} \quad (y + (-x)) \cdot 1 \in \mathbb{Z}_0^+ \\
&\Leftrightarrow \quad y + (-x) \in \mathbb{Z}_0^+ \\
&\Leftrightarrow \quad x \leqslant y
\end{aligned}$$

which combined with [eq: 8.14] proves that

$$i_{\mathbb{Z}\rightarrow\mathbb{Q}}: \langle \mathbb{Z}, \leqslant \rangle \rightarrow \langle \mathbb{Z}_{\mathbb{Q}}, \leqslant \rangle \text{ is a order isomorphism}$$

Now for the proof that $i_{\mathbb{Z}\rightarrow\mathbb{Q}}$ is a ring isomorphism.

$$i_{\mathbb{Z}\rightarrow\mathbb{Q}}(x + y) = \frac{x + y}{1} = \frac{x \cdot 1 + 1 \cdot y}{1 \cdot 1} = \frac{x}{1} + \frac{y}{1} = i_{\mathbb{Z}\rightarrow\mathbb{Q}}(x) + i_{\mathbb{Z}\rightarrow\mathbb{Q}}(y)$$

and

$$i_{\mathbb{Z}\rightarrow\mathbb{Q}}(x \cdot y) = \frac{x \cdot y}{1} = \frac{x}{1} \cdot \frac{y}{1} = i_{\mathbb{Z}\rightarrow\mathbb{Q}}(x) \cdot i_{\mathbb{Z}\rightarrow\mathbb{Q}}(y)$$

and

$$i_{\mathbb{Z}\rightarrow\mathbb{Q}}(1)=\frac{1}{1}=1$$

proving with [eq: 8.14] that

$$i_{\mathbb{Z}\rightarrow\mathbb{Q}}\colon\langle\mathbb{Z},+,\cdot\rangle\rightarrow\langle\mathbb{Z}_{\mathbb{Q}},+,\cdot\rangle \text{ is a ring isomorphism}$$

□

Theorem 8.33. $\mathbb{Q}=\{x\cdot y^{-1}|x\in\mathbb{Z}_{\mathbb{Q}}\wedge y\in\mathbb{Z}_{\mathbb{Q}}\setminus\{0\}\}$

Proof. If $q\in\mathbb{Q}$ there exists $x'\in\mathbb{Z}$ and $y'\in\mathbb{Z}^*=\mathbb{Z}\setminus\{0\}$ such that $q=\frac{x'}{y'}$. Define now $x=\frac{x'}{1}=i_{\mathbb{Z}\rightarrow\mathbb{Q}}(x')$ and $y=\frac{y'}{1}=i_{\mathbb{Z}\rightarrow\mathbb{Q}}(y')$ then by the previous theorems [theorem: 8.32] we have that $x\in\mathbb{Z}_{\mathbb{Q}}$ and $y\in\mathbb{Z}_{\mathbb{Q}}\setminus\{0\}$. Further

$$q=\frac{x'}{y'}=\frac{x'}{1}\cdot\frac{1}{y'}=\frac{x'}{1}\cdot\left(\frac{y'}{1}\right)^{-1}=x\cdot y^{-1}\in\{x\cdot y^{-1}|x\in\mathbb{Z}_{\mathbb{Q}}\wedge y\in\mathbb{Z}_{\mathbb{Q}}\setminus\{0\}\}$$

proving

$$\mathbb{Q}\subseteq\{x\cdot y^{-1}|x\in\mathbb{Z}_{\mathbb{Q}}\wedge y\in\mathbb{Z}_{\mathbb{Q}}\setminus\{0\}\} \tag{8.16}$$

If $q\in\{x\cdot y^{-1}|x\in\mathbb{Z}_{\mathbb{Q}}\wedge y\in\mathbb{Z}_{\mathbb{Q}}\setminus\{0\}\}$ there exists a $x\in\mathbb{Z}_{\mathbb{Q}}$ and $y\in\mathbb{Z}_{\mathbb{Q}}\setminus\{0\}$ such that $q=x\cdot y^{-1}$. Using [theorem: 8.32] again there exists $x'\in\mathbb{Z}$ and $y'\in\mathbb{Z}\setminus\{0\}$ such that $x=i_{\mathbb{Z}\rightarrow\mathbb{Q}}(x')=\frac{x'}{1}$ and $y=i_{\mathbb{Z}\rightarrow\mathbb{Q}}(y')=\frac{y'}{1}$ so that

$$q=x\cdot y^{-1}=\frac{x'}{1}\cdot\left(\frac{1}{y'}\right)^{-1}=\frac{x'}{1}\cdot\frac{1}{y'}=\frac{x'}{y'}\in\mathbb{Q}$$

proving that $\{x\cdot y^{-1}|x\in\mathbb{Z}_{\mathbb{Q}}\wedge y\in\mathbb{Z}_{\mathbb{Q}}\setminus\{0\}\}\subseteq\mathbb{Q}$ which combined with [eq: 8.16] results in

$$\mathbb{Q}=\{x\cdot y^{-1}|x\in\mathbb{Z}_{\mathbb{Q}}\wedge y\in\mathbb{Z}_{\mathbb{Q}}\setminus\{0\}\}$$

□

Definition 8.34. $\mathbb{N}_{0,\mathbb{Q}}=(i_{\mathbb{Z}\rightarrow\mathbb{Q}}\circ i_{\mathbb{N}_0\rightarrow\mathbb{Q}})(\mathbb{N}_0)\subseteq\mathbb{Q}$ *where*

$$i_{\mathbb{N}_0\rightarrow\mathbb{Z}}\colon\mathbb{N}_0\rightarrow\mathbb{Z} \textit{is defined by } i_{\mathbb{N}_0\rightarrow\mathbb{Z}}(n)=\sim[n,0]$$
$$i_{\mathbb{Z}\rightarrow\mathbb{Q}}\colon\mathbb{Z}\rightarrow\mathbb{Q} \textit{ is defined by } i_{\mathbb{Z}\rightarrow\mathbb{Q}}(z)=\frac{z}{1}$$

Theorem 8.35. *We have that*

1. $\mathbb{N}_{0,\mathbb{Q}}=\{\frac{n}{1}|n\in\mathbb{Z}_0^+\}$
2. *If we define* $i_{\mathbb{N}_0\rightarrow\mathbb{Q}}\colon\mathbb{N}_0\rightarrow\mathbb{N}_{0,\mathbb{Q}}$ *by* $i_{\mathbb{N}_0\rightarrow\mathbb{Q}}=i_{\mathbb{Z}\rightarrow\mathbb{Q}}\circ i_{\mathbb{N}_0\rightarrow\mathbb{Z}}$ *then we have:*
   a. $\langle\mathbb{N}_{0,\mathbb{Q}},+\rangle$ *is a sub semi-group of* $\langle\mathbb{Q},+\rangle$ *and*
   $$i_{\mathbb{N}_0\rightarrow\mathbb{Q}}\colon\langle\mathbb{N}_0,+\rangle\rightarrow\langle\mathbb{N}_{0,\mathbb{Q}},+\rangle \textit{ is a group isomorphism}$$
   b. $\langle\mathbb{N}_{0,\mathbb{Q}},\cdot\rangle$ *is a sub semi-group of* $\langle\mathbb{Q},\cdot\rangle$ *and*
   $$i_{\mathbb{N}_0\rightarrow\mathbb{Q}}\colon\langle\mathbb{N}_0,\cdot\rangle\rightarrow\langle\mathbb{N}_{0,\mathbb{Q}},\cdot\rangle \textit{ is a group isomorphism}$$
   c.
   $$i_{\mathbb{N}_0\rightarrow\mathbb{Q}}\colon\langle\mathbb{N}_0,\leqslant\rangle\rightarrow\langle\mathbb{N}_{0,\mathbb{Q}},\leqslant\rangle \textit{ is a group isomorphism}$$
3. $\forall n\in\mathbb{N}_{0,\mathbb{Q}}$ *we have that* $0\leqslant n$*, if* $n\neq 0$ *then* $0<1\leqslant n$

Proof.

1. We have

$$\begin{array}{rcl} x\in\mathbb{N}_{0,\mathbb{Q}} & \Leftrightarrow & \exists n\subset\mathbb{N}_0 \text{ such that } x=(i_{\mathbb{Z}\rightarrow\mathbb{Q}}\circ i_{\mathbb{N}_0\rightarrow\mathbb{Z}})(n) \\ & \Rightarrow & x=i_{\mathbb{Z}\rightarrow\mathbb{Q}}(i_{\mathbb{N}_0\rightarrow\mathbb{Z}}(n)) \\ & \Rightarrow & x=\dfrac{\sim[(n,0)]}{1} \\ & \underset{\sim[(n,0)]\in\mathbb{Z}_0^+}{\Rightarrow} & x\in\left\{\dfrac{n}{1}|n\in\mathbb{Z}_0^+\right\} \end{array}$$

   proving that

$$\mathbb{N}_{0,\mathbb{Q}}\subseteq\left\{\frac{n}{1}|n\in\mathbb{Z}_0^+\right\} \tag{8.17}$$

Further

$$\begin{array}{rcl} x\in\left\{\frac{n}{1}|n\in\mathbb{Z}_0^+\right\} & \Rightarrow & \exists n\in\mathbb{Z}_0^+ \text{ such that } x=\frac{n}{1} \\ & \underset{\text{definition of } \mathbb{Z}_0^+}{\Rightarrow} & \exists n'\in\mathbb{N}_0 \text{ such that } n=\sim[(n',0)] \\ & \Rightarrow & x=\frac{\sim[(n,0)]}{1} \\ & \Rightarrow & x=i_{\mathbb{Z}\rightarrow\mathbb{Q}}(i_{\mathbb{N}_0\rightarrow\mathbb{Z}}(n')) \\ & \Rightarrow & (i_{\mathbb{Z}\rightarrow\mathbb{Q}}\circ i_{\mathbb{N}_0\rightarrow\mathbb{Z}})(n') \\ & \Rightarrow & x\in\mathbb{N}_{0,\mathbb{Q}} \end{array}$$

proving that $\left\{\frac{n}{1}|n\in\mathbb{Z}_0^+\right\}\subseteq\mathbb{N}_{0,\mathbb{Q}}$ which combined with [eq: 8.17] gives

$$\mathbb{N}_{0,\mathbb{Q}}=\left\{\frac{n}{1}|n\in\mathbb{Z}_0^+\right\}$$

2.

a. Using [theorem: 7.17] and [theorem: 8.32] we have that $\mathbb{Z}_0^+$ is a sub semi group of $\langle\mathbb{Z},+\rangle$ and that

$$i_{\mathbb{N}_0\rightarrow\mathbb{Z}}\colon\langle\mathbb{N}_0,+\rangle\rightarrow\langle\mathbb{Z}_0^+,+\rangle \text{ and } i_{\mathbb{Z}\rightarrow\mathbb{Q}}\colon\langle\mathbb{Z},+\rangle\rightarrow\langle\mathbb{Z}_{\mathbb{Q}},+\rangle \text{ are group isomorphisms}$$

Applying then [theorems: 4.25, 4.17] we find that $(i_{\mathbb{Z}\rightarrow\mathbb{Q}}\circ i_{\mathbb{N}_0\rightarrow\mathbb{Z}})(\mathbb{N}_0)$ is a sub semi-group of $\langle\mathbb{Q},+\rangle$ and that $i_{\mathbb{Z}\rightarrow\mathbb{Q}}\circ i_{\mathbb{N}_0\rightarrow\mathbb{Z}}\colon\langle\mathbb{N}_0,+\rangle\rightarrow\langle(i_{\mathbb{Z}\rightarrow\mathbb{Q}}\circ i_{\mathbb{N}_0\rightarrow\mathbb{Z}})(\mathbb{N}_0),+\rangle$ is a group isomorphism. Using then the definition of $\mathbb{N}_{0,\mathbb{Q}}$ and $i_{\mathbb{N}_0\rightarrow\mathbb{Q}}$ we get then finally

$$\mathbb{N}_{0,\mathbb{Q}} \text{ is a subgroup of } \langle\mathbb{Q},+\rangle \text{ and } i_{\mathbb{N}_0\rightarrow\mathbb{Q}}\colon\langle\mathbb{N}_0,+\rangle\rightarrow\langle\mathbb{N}_{0,\mathbb{Q}},+\rangle \text{ is a group isomorphism.}$$

b. Using [theorem: 7.17] and [theorem: 8.32] we have that $\mathbb{Z}_0^+$ is a sub semi group of $\langle\mathbb{Z},\cdot\rangle$ and that

$$i_{\mathbb{N}_0\rightarrow\mathbb{Z}}\colon\langle\mathbb{N}_0,\cdot\rangle\rightarrow\langle\mathbb{Z}_0^+,\cdot\rangle \text{ and } i_{\mathbb{Z}\rightarrow\mathbb{Q}}\colon\langle\mathbb{Z},\cdot\rangle\rightarrow\langle\mathbb{Z}_{\mathbb{Q}},\cdot\rangle \text{ are group isomorphisms}$$

Applying then [theorems: 4.25, 4.17] we find that $(i_{\mathbb{Z}\rightarrow\mathbb{Q}}\circ i_{\mathbb{N}_0\rightarrow\mathbb{Z}})(\mathbb{N}_0)$ is a sub semi-group of $\langle\mathbb{Q},\cdot\rangle$ and that $i_{\mathbb{Z}\rightarrow\mathbb{Q}}\circ i_{\mathbb{N}_0\rightarrow\mathbb{Z}}\colon\langle\mathbb{N}_0,\cdot\rangle\rightarrow\langle(i_{\mathbb{Z}\rightarrow\mathbb{Q}}\circ i_{\mathbb{N}_0\rightarrow\mathbb{Z}})(\mathbb{N}_0),\cdot\rangle$ is a group isomorphism. Using then the definition of $\mathbb{N}_{0,\mathbb{Q}}$ and $i_{\mathbb{N}_0\rightarrow\mathbb{Q}}$ we get then finally

$$\mathbb{N}_{0,\mathbb{Q}} \text{ is a subgroup of } \langle\mathbb{Q},\cdot\rangle \text{ and } i_{\mathbb{N}_0\rightarrow\mathbb{Q}}\colon\langle\mathbb{N}_0,\cdot\rangle\rightarrow\langle\mathbb{N}_{0,\mathbb{Q}},\cdot\rangle \text{ is a group isomorphism}$$

3. Using [theorem: 7.29] and [theorem: 8.32] we have that

$$i_{\mathbb{N}_0\rightarrow\mathbb{Z}}\colon\langle\mathbb{N}_0,\leqslant\rangle\rightarrow\langle\mathbb{Z}_0^+,\leqslant\rangle \text{ and } i_{\mathbb{Z}\rightarrow\mathbb{Q}}\colon\langle\mathbb{Z},\leqslant\rangle\rightarrow\langle\mathbb{Z}_{\mathbb{Q}},\leqslant\rangle \text{ are order isomorphisms}$$

So using [theorem: 3.52] we have that $i_{\mathbb{Z}\rightarrow\mathbb{Q}}\circ i_{\mathbb{N}_0\rightarrow\mathbb{Z}}\colon\langle\mathbb{N}_0,\leqslant\rangle\rightarrow\langle(i_{\mathbb{N}_0\rightarrow\mathbb{Z}}\circ i_{\mathbb{Z}\rightarrow\mathbb{Q}})(\mathbb{N}_0),\leqslant\rangle$ is a order isomorphism. Using then the definition of $\mathbb{N}_{0,\mathbb{Q}}$ and $i_{\mathbb{N}_0\rightarrow\mathbb{Q}}$ we get then finally that

$$i_{\mathbb{N}_0\rightarrow\mathbb{Q}}\colon\langle\mathbb{N}_0,\leqslant\rangle\rightarrow\langle\mathbb{N}_{0,\mathbb{Q}},\leqslant\rangle \text{ is a order isomorphism}$$

4. If $n\in\mathbb{N}_{0,\mathbb{Q}}$ then there exist a $n'\in\mathbb{N}_0$ such that

$$n=i_{\mathbb{N}_0\rightarrow\mathbb{Q}}(n')=i_{\mathbb{Z}\rightarrow\mathbb{Q}}(i_{\mathbb{N}_0\rightarrow\mathbb{Z}}(n'))=i_{\mathbb{Z}\rightarrow\mathbb{Q}}(\sim[(n',0)])=\frac{\sim[(n',0)]}{1} \tag{8.18}$$

now $\sim[(n',0)]\cdot 1=\sim[(n',0)]\in\mathbb{Z}_0^+$ so that $n\in\mathbb{Q}_0^+$ or

$$0\leqslant n$$

If $n\neq 0$ then by [eq: 8.18] $n'\neq 0\Rightarrow 0<n'$ so by [theorem: 5.50] $1=s(0)\leqslant n'$. Using (3) we have that

$$1=i_{\mathbb{N}_0\rightarrow\mathbb{Q}}(1)\leqslant i_{\mathbb{N}_0\rightarrow\mathbb{Q}}(n')=n$$

as by [example: 8.25] $0<1$ we have

$$0<1\leqslant n$$

□

Theorem 8.36. (Archimedean Property) *If $x,y\in\mathbb{Q}$ with $0<x$ then there exist a $n\in\mathbb{N}_{0,\mathbb{Q}}$ such that $y<n\cdot x$*

Proof. For $y\in\mathbb{Q}$ we have the following possibilities to consider:

***y* ⩽ 0.** Take $1\in\mathbb{Q}$ then by [theorem: 8.35] $1\in\mathbb{N}_{0,\mathbb{Q}}$ so if we take $n=1$ then $y\leqslant 0<x=1\cdot x=n\cdot x$, hence $y<n\cdot x$.

$0<y$. As also $0<x$ we have by [theorem: 8.24] the existence of $p,q,r,s\in\mathbb{Z}$ with $0<p$, $0<q$, $0<r$, $0<s$ such that $x=\frac{p}{q}$ and $y=\frac{r}{s}$. As $0<p\wedge 0<s$ we have by [theorem: 7.26] that $0<p\cdot s$. Using the Archimedean property of $\mathbb{Z}$ [see theorem: 7.30] there exist a $n'\in\mathbb{Z}_0^+$ such that $q\cdot r<n'\cdot(p\cdot s)$ or

$$0<n'\cdot(p\cdot s)+(-(q\cdot r)) \tag{8.19}$$

As $n'\in\mathbb{Z}_0^+$ there exists a $n''\in\mathbb{N}_0$ such that $n'=\sim[(n'',0)]$ so that if we take $n=\frac{n'}{1}=\frac{\sim[(n'',0)]}{1}$ we have $n=i_{\mathbb{Z}\rightarrow\mathbb{Q}}(i_{\mathbb{N}_0\rightarrow\mathbb{Z}}(n''))$ so that

$$n\in\mathbb{N}_{0,\mathbb{Q}} \tag{8.20}$$

Now

$$\begin{aligned} n\cdot x-y &= \frac{n'}{1}\cdot x+(-y)\\ &= \frac{n'}{1}\cdot\frac{p}{q}+\frac{-r}{s}\\ &= \frac{n'\cdot p}{1\cdot q}+\frac{-r}{s}\\ &= \frac{n'\cdot p}{q}+\frac{-r}{s}\\ &= \frac{(n'\cdot p)\cdot s+q\cdot(-r)}{q\cdot s}\\ &= \frac{n'\cdot(p\cdot s)+(-(q\cdot r))}{q\cdot s} \end{aligned} \tag{8.21}$$

As $0<q\wedge 0<s \underset{\text{[theorem: 7.26]]}}{\Rightarrow} 0<q\cdot s$ and $0<n'\cdot(p\cdot s)+(-(q\cdot r))$ [see eq: 8.19] it follows using [theorem: 8.24] that $0<n\cdot x-y$ hence

$$y<n\cdot x \text{ where } n\in\mathbb{N}_{0,\mathbb{Q}}$$

□

Theorem 8.37. ($\mathbb{Q}$ is dense) *If $x,y\in\mathbb{Q}$ with $x<y$ then there exist a $q\in\mathbb{Q}$ such that $x<q<y$.*

Proof. As $x<y$ we have by [theorems 8.29, 4.73] that $x+x<y+x=x+y$ and $x+y<y+y$. Further $x+x=1\cdot x+1\cdot x=(1+1)\cdot x \underset{\text{[example: 8.13]}}{=} \frac{2}{1}\cdot x$ and $y+y=1\cdot y+1\cdot y=(1+1)\cdot y \underset{\text{[example: 8.13]}}{=} \frac{2}{1}\cdot y$. So

$$\frac{2}{1}\cdot x<x+y \text{ and } x+y<\frac{2}{1}\cdot y \tag{8.22}$$

As $0<1<1+1=2=\frac{2}{1}$ we have by [theorems: 8.29, 4.73] $0<\left(\frac{2}{1}\right)^{-1}=\frac{1}{2}$, so using [theorems: 8.29, 4.73] on [eq: 8.22] gives $x<\frac{1}{2}\cdot(x+y)$ and $\frac{1}{2}\cdot(x+y)<y$. So if $q=\frac{1}{2}\cdot(x+y)$ we have that

$$x<q<y$$

□

Theorem 8.38. *$\langle\mathbb{N}_{0,\mathbb{Q}},\leqslant\rangle$ is well ordered*

Proof. Using [theorem: 8.35] we have that $i_{\mathbb{N}_0\rightarrow\mathbb{Q}}:\mathbb{N}_0\rightarrow\mathbb{N}_{0,\mathbb{Q}}$ is a order isomorphism, further by [theorem: 5.51] $\langle\mathbb{N}_0,\leqslant\rangle$ is well ordered. so using [theorem: 3.82] we conclude that

$$\langle\mathbb{N}_{0,\mathbb{Q}},\leqslant\rangle \text{ is well ordered}$$

□

Theorem 8.39. *$\langle\mathbb{N}_{0,\mathbb{Q}},\leqslant\rangle$ is conditional complete*

Proof. As by [theorem: 8.38] $\langle\mathbb{N}_{0,\mathbb{Q}},\leqslant\rangle$ is well-ordered it follows from [theorem: 3.84] that $\langle\mathbb{Z}_0^+,\leqslant\rangle$ is conditional complete. □

Although $\langle\mathbb{N}_{0,\mathbb{Q}},\leqslant\rangle$ is conditional complete we have that $\langle\mathbb{Q},\leqslant\rangle$ is not conditional complete as we will prove now. First we need a lemma that essentially says that $\sqrt{2}$ is not a rational number.

Lemma 8.40. *$\forall q\in\mathbb{Q}$ we have $q^2 \underset{\text{def}}{=} q\cdot q\neq 2=\frac{2}{1}$.*

Proof. We prove this by contradiction. Assume that $\exists q'\in\mathbb{Q}$ such that $q'\cdot q'=2$, then $q'\neq 0$ [if $q'=0\Rightarrow q'\cdot q'=0=2$]. Take $q=\left\{\begin{array}{l} -q' \text{ if } q'<0\\ q' \text{ if q'>0} \end{array}\right.$ then

$$0<q \text{ and } q\cdot q=2$$

Using [theorem: 8.24] there exist $n, m \in \mathbb{Z}$ with $0 < n \wedge 0 < m$ such that $q = \frac{n}{m}$. Take now $n' = n \colon \gcd(n, m)$ and $m' = m \colon \gcd(n, m)$ then as $m \neq 0 \Rightarrow m' \neq 0$ we have by [theorem: 7.49] that

$$\{d \in \mathbb{Z} | d | n' \wedge d | m'\} = \{1, -1\} \tag{8.23}$$

Now

$$\frac{n'}{m'} \underset{0 < \gcd(n,m) \wedge [\text{theorem: } 8.5]}{=} \frac{n' \cdot \gcd(n, m)}{m' \cdot \gcd n, m} = \frac{n}{m} = q$$

so that

$$\frac{n' \cdot n'}{m' \cdot m'} = \frac{n'}{m'} \cdot \frac{n'}{m'} = q \cdot q = \frac{2}{1}$$

Hence $(n' \cdot n') \cdot 1 = (m' \cdot m') \cdot 2$ or $n' \cdot n' = 2 \cdot (m' \cdot m')$ proving that $n' \cdot n'$ is even, using [theorem: 7.52] it follows then that n' is even. So there exist a $k \in \mathbb{Z}$ such that $n' = 2 \cdot k$. Then $2 \cdot (m' \cdot m') = (2 \cdot k) \cdot (2 \cdot k) = 2 \cdot (2 \cdot (k \cdot k))$, hence by [theorem: 7.15] $m' \cdot m' = 2 \cdot (k \cdot k)$ proving that $m' \cdot m'$ is even, by [theorem: 7.52] m' is even hence $\exists l \in \mathbb{Z}$ such that $m' = 2 \cdot l$. So $2 | n'$ and $2 | m'$ which by [eq: 8.23] means that $2 = 1$ or $2 = -1$ both of which are false, so we reach a contradiction. □

Theorem 8.41. *$\langle \mathbb{Q}, \leqslant \rangle$ is not conditional complete, so there exist a non empty subset of $\mathbb{Q}$ that is bounded above but does not have a least upper bound.*

Proof. In this prove we make use of the fact that there does not exist a $q \in \mathbb{Q}$ such that $q \cdot q = 2$. So define

$$A = \left\{ q \in \mathbb{Q} | 0 < q \wedge q \cdot q < \frac{2}{1} \right\} \subseteq \mathbb{Q}$$

As $0 < \frac{4}{3}$ and $\frac{2}{1} + \left(-\left(\frac{4}{3}\right) \cdot \left(\frac{4}{3}\right)\right) = \frac{18 - 16}{9} = \frac{2}{8} > 0$ so that $\frac{4}{3} \cdot \frac{4}{3} < 2$ we have that

$$\frac{4}{3} \in A \Rightarrow \varnothing \neq A \tag{8.24}$$

Let $x \in A$ then $0 < x$ and $x \cdot x < \frac{2}{1}$. Assume that $\frac{2}{1} < x$ then by multiplying both sides by x we have by [theorems: 8.29, 4.73] that $\frac{2}{1} \cdot x < x \cdot x < \frac{2}{1} = 1 \cdot \frac{2}{1}$, we have by [theorems: 8.29, 4.73] that $x < 1 < \frac{2}{1}$ contradicting $\frac{2}{1} < x$. So we must have that $x \leqslant \frac{2}{1}$ hence

$$\frac{2}{1} \text{ is a upper bound of } A \tag{8.25}$$

Assume now that $u = \sup(A)$ exist. As $\frac{4}{3} + (-1) = \frac{4}{3} + \frac{-1}{1} = \frac{4 + (-3)}{3} = \frac{1}{3} > 0$ it follows that $1 < \frac{4}{3} \in A$ so that $0 < 1 < u$ and as $\frac{2}{1}$ is a upper bound of A we have

$$0 < 1 < u \leqslant \frac{2}{1} \tag{8.26}$$

Now for $u \cdot u$ we have by [theorem: 8.40] that $u \cdot u \neq \frac{2}{1}$ so we have only to consider the following possibilities:

$\boldsymbol{u \cdot u < \frac{2}{1}}$**.** So $0 < \frac{2}{1} + (-u \cdot u)$ and by the Archimedean property [see theorem: 8.36] there exist a $n' \in \mathbb{N}_{0, \mathbb{Q}}$ such that

$$\frac{5}{1} < n' \cdot \left(\frac{2}{1} - u \cdot u\right)$$

Using [theorem 8.35] we have that $\exists n \in \mathbb{Z}_0^+$ such that $n' = \frac{n}{1}$, further $n \neq 0$ [otherwise $\frac{5}{1} < \frac{0}{1} \cdot \left(\frac{2}{1} - u \cdot u\right) = 0$] so there exist a $n \in \mathbb{Z}_0^+ \setminus \{0\}$ such that

$$\frac{5}{1} < \frac{n}{1} \cdot \left(\frac{2}{1} - u \cdot u\right)$$

multiplying both sides by $\frac{1}{n} = \left(\frac{n}{1}\right)^{-1}$ gives

$$\frac{5}{n} < \frac{2}{1} - u \cdot u \tag{8.27}$$

Now

$$\begin{aligned}\left(u + \frac{1}{n}\right) \cdot \left(u + \frac{1}{n}\right) &= u \cdot u + u \cdot \frac{1}{n} + u \cdot \frac{1}{n} + \frac{1}{n} \cdot \frac{1}{n} \\ &= u \cdot u + \frac{2}{1} \cdot u \cdot \frac{1}{n} + \frac{1}{n} \cdot \frac{1}{n}\end{aligned}$$

and thus

$$\left(u+\frac{1}{n}\right)\cdot\left(u+\frac{1}{n}\right)<\frac{2}{1} \Leftrightarrow u\cdot u+\frac{2}{1}\cdot u\cdot\frac{1}{n}+\frac{1}{n}\cdot\frac{1}{n}<\frac{2}{1}$$
$$\Leftrightarrow \frac{2}{1}\cdot u\cdot\frac{1}{n}+\frac{1}{n}\cdot\frac{1}{n}<\frac{2}{1}-u\cdot u$$
$$\Leftrightarrow \frac{2}{n}\cdot u+\frac{1}{n}\cdot\frac{1}{n}<\frac{2}{1}-u\cdot u \tag{8.28}$$

As $0<n$, so that by [theorem: 7.23] $1\leqslant n$, hence $0\leqslant n-1$. Now

$$\frac{1}{n}-\frac{1}{n}\cdot\frac{1}{n}=\frac{1}{n}-\frac{1}{n\cdot n}=\frac{n\cdot n-n\cdot 1}{n\cdot n}=\frac{n\cdot(n-1)}{n\cdot n}=\frac{n-1}{n}\geqslant 0$$

giving

$$\frac{1}{n}\cdot\frac{1}{n}\leqslant\frac{1}{n} \tag{8.29}$$

Further as $u\leqslant\frac{2}{1}$ [see eq: 8.26] we have by [theorems: 8.29, 4.73] and $0<\frac{2}{n}$ [as $0<n\wedge 0<2$]

$$u\cdot\frac{2}{n}\leqslant\frac{2}{n}\cdot\frac{2}{1}=\frac{4}{n} \tag{8.30}$$

So

$$\begin{aligned}
\frac{2}{n}\cdot u+\frac{1}{n}\cdot\frac{1}{n} &\leqslant_{[\text{eq: } 8.29]} \frac{2}{n}\cdot u+\frac{1}{n}\\
1 &\leqslant_{[\text{eq: } 8.30]} \frac{4}{n}+\frac{1}{n}\\
&= \left(\frac{4}{1}+\frac{1}{1}\right)\cdot\frac{1}{n}\\
&= \frac{5}{1}\cdot\frac{1}{n}\\
&= \frac{5}{n}\\
&<_{[\text{eq: } 8.27]} \frac{2}{1}-u\cdot u
\end{aligned}$$

So by [eq: 8.28] we have

$$\left(u+\frac{1}{n}\right)\cdot\left(u+\frac{1}{n}\right)<\frac{2}{1}$$

By [eq: 8.26] $0<u\Rightarrow 0<u+\frac{1}{n}$ which together with the above proves that $u+\frac{1}{n}\in A$, so $u+\frac{1}{n}\leqslant\sup(A)=u$, which as $u<u+\frac{1}{n}$ leads to the contradiction $u<u$. So this case is impossible.

$\boldsymbol{\frac{2}{1}<u\cdot u}$**.** So $0<u\cdot u+\frac{-2}{1}$ and using the Archimedean property there exist a $n'\in\mathbb{N}_{0,\mathbb{Q}}$ such that

$$\frac{2}{1}\cdot u<n'\cdot\left(u\cdot u+\frac{-2}{1}\right) \tag{8.31}$$

Using [theorem: 8.35] there exist a $n\in\mathbb{Z}_0^+$ such that $n'=\frac{n}{1}$. If $n=0\Rightarrow n'=0$ so that $\frac{2}{1}\cdot u<0\cdot\left(u\cdot u+\frac{-2}{1}\right)=0\Rightarrow u<0$ contradicting $0<u$ [see eq: 8.26], hence we must have that $n\neq 0$ or $0<n$, so $(n')^{-1}=\frac{1}{n}$ exist and $0<\frac{1}{n}$. Next

$$\begin{aligned}
\frac{2}{1}\cdot u<n'\cdot\left(u\cdot u+\frac{-2}{1}\right) &\Rightarrow \frac{2}{1}\cdot u<\frac{n}{1}\cdot\left(u\cdot u+\frac{-2}{1}\right)\\
&\underset{0<\frac{1}{n}\wedge[\text{theorems: } 8.29,\ 4.73]}{\Rightarrow} \left(\frac{2}{1}\cdot u\right)\cdot\frac{1}{n}<\left(\frac{n}{1}\cdot\left(u\cdot u+\frac{-2}{1}\right)\right)\cdot\frac{1}{n}\\
&\Rightarrow \frac{2}{n}\cdot u<u\cdot u+\frac{-2}{1}\\
&\Rightarrow \frac{2}{1}<u\cdot u+\frac{-2}{n}\cdot u
\end{aligned}$$

which as $0<\frac{1}{n}\Rightarrow 0<\frac{1}{n}\cdot\frac{1}{n}$ proves that

$$\frac{2}{1}<u\cdot u+\frac{-2}{n}\cdot u+\frac{1}{n}\cdot\frac{1}{n} \tag{8.32}$$

As $0<n$ we have that $1\leqslant n$ which by [theorem: 8.27] gives $\frac{1}{n}\leqslant 1$ and as $1<u$ we have $\frac{1}{n}<u$, hence

$$0<u+\left(\frac{-1}{n}\right)\underset{[\text{theorems: } 8.29,\ 4.73]}{\Rightarrow}0<\left(u+\frac{-1}{n}\right)\cdot\left(u+\frac{-1}{n}\right) \tag{8.33}$$

As $0<\frac{1}{n}\underset{[[\text{theorems: } 8.29,\ 4.73]}{\Rightarrow}\frac{-1}{n}<0$ so that $u+\frac{-1}{n}<u$, as $u=\sup(A)$ and $\langle\mathbb{Q},\leqslant\rangle$ is totally ordered we have by [theorem: 3.71] that there exist a $q\in A$ such that

$$u+\frac{-1}{n}<q\leqslant u \tag{8.34}$$

Multiplying both sides of [eq: 8.34] by $u+\frac{-1}{n}$ we have by [theorems: 8.29, 4.73] that

$$\left(u+\frac{-1}{n}\right)\cdot\left(u+\frac{-1}{n}\right)<q\cdot\left(u+\frac{-1}{n}\right)$$

Further as by [eq: 8.33] $0<u+\frac{-1}{n}<q\Rightarrow 0<q$ we have, by multiplying both sides of [eq: 8.34] by q, that

$$\left(u+\frac{-1}{n}\right)\cdot q<q\cdot q.$$

Hence $\left(u+\frac{-1}{n}\right)\cdot\left(u+\frac{-1}{n}\right)<q\cdot q$ and as $q\in A$ we have also $q\cdot q<\frac{2}{1}$ so that

$$\left(u+\frac{-1}{n}\right)\cdot\left(u+\frac{-1}{n}\right)<\frac{2}{1} \tag{8.35}$$

Next

$$\begin{aligned}\left(u+\frac{-1}{n}\right)\cdot\left(u+\frac{-1}{n}\right) &= u\cdot u+\frac{-1}{n}\cdot u+\frac{-1}{n}\cdot n+\frac{-1}{n}\cdot\frac{-1}{n}\\ &= u\cdot u+\frac{-2}{n}+\frac{1}{n}\cdot\frac{1}{n}\end{aligned}$$

which by [eq: 8.35] proves that $u\cdot u+\frac{-2}{n}+\frac{1}{n}\cdot\frac{1}{n}<\frac{2}{1}$, combining this with [eq: 8.32] result in $\frac{2}{1}<\frac{2}{1}$ a contradiction. So this case is impossible.

As all possible cases are impossible, the assumption is wrong hence A has no supremum and $\langle\mathbb{Q},\leqslant\rangle$ is not conditional complete. □

So we have that $\langle\mathbb{N}_{0,\mathbb{Q}},\leqslant\rangle$ is conditional complete but $\langle\mathbb{Q},\leqslant\rangle$ is not. This defect will be resolved by introducing the set of real numbers that will extend the set of rationals.

8.3 Denumerability of the rationals

Theorem 8.42. $\mathbb{N}_{0,\mathbb{Q}}$ is denumerable.

Proof. Using [theorem: 8.35] $i_{\mathbb{N}_0\to Q}:\mathbb{N}_0\to\mathbb{N}_{0,\mathbb{Q}}$ is a bijection, hence $\mathbb{N}_0\approx\mathbb{Q}$ proving that $\mathbb{Q}$ is denumerable. □

Theorem 8.43. $\mathbb{Z}_{\mathbb{Q}}$ *is denumerable*

Proof. Using [theorem: 7.53] we have that $\mathbb{Z}$ is denumerable, further by [theorem: 8.32]

$$i_{\mathbb{Z}\to\mathbb{Q}}:\mathbb{Q}\to\mathbb{Z}_{\mathbb{Q}}$$

is a bijection, hence $\mathbb{N}_0\approx\mathbb{Q}\approx\mathbb{Z}_{\mathbb{Q}}$, proving that $\mathbb{Z}_{\mathbb{Q}}$ is denumerable. □

Theorem 8.44. $\mathbb{Q}$ *is denumerable*

Proof. Define the mapping $f:\mathbb{Z}\times\mathbb{Z}\to\mathbb{Q}$ by

$$f(x,y)\ =\ \begin{cases}\frac{x}{y}\text{ if }(x,y)\in\mathbb{Z}\times\mathbb{Z}^{*}\\ 0\text{ if }(x,y)\in\mathbb{Z}\times\{0\}\end{cases}$$

If $q\in\mathbb{Q}$ then there exist a $(x,y)\in\mathbb{Z}\times\mathbb{Z}^{\star}\subseteq\mathbb{Z}\times\mathbb{Z}$ such that $q=\frac{x}{y}=f(x,y)$ proving that

$$f:\mathbb{Z}\times\mathbb{Z}\to\mathbb{Q}\text{ is a surjection}$$

As $\mathbb{Z}$ is denumerable [see theorem: 7.53] we have by [theorem: 6.62] that $\mathbb{Z}\times\mathbb{Z}$ is denumerable, hence there exist a bijection $g:\mathbb{N}_0\to\mathbb{Z}\times\mathbb{Z}$, so $f\circ g:\mathbb{N}_0\to\mathbb{Q}$ is a surjection. By [theorem: 6.69] $\mathbb{Q}$ is countable, hence either finite or denumerable. As $\mathbb{N}_{0,\mathbb{Q}}\subseteq\mathbb{Q}$ and $\mathbb{N}_{0,\mathbb{Q}}$ is denumerable, it follow from [theorem: 6.29] that $\mathbb{Q}$ is not finite, hence we must have that $\mathbb{Q}$ is denumerable. □

Chapter 9

The real numbers

In this chapter we will introduce the set of real numbers and embed the natural, integer and rational numbers in it. Just as with $\mathbb{Q}$, $\mathbb{Z}$ and $\mathbb{N}_0$ we will introduce a order relation, a sum operator, a product operator, neutral elements for addition and multiplication as well as inverse elements. If we would use different symbols for these we introduce a lot of excessive notation clutter. So we use the same symbols for the natural numbers, integers, rational numbers and real numbers and use context to determine the meaning of the symbols involved. The following table should help you in determining the meaning of the different symbols based on the context of there usage.

Context	Expression	Operator
$n,m\in\mathbb{N}_0$	n+m	sum in $\langle\mathbb{N}_0,+\rangle$
$n,m\in\mathbb{N}_0$	$n\cdot m$	product in $\langle\mathbb{N}_0,\cdot\rangle$
$n,m\in\mathbb{N}_0$	$n\leqslant m$	order in $\langle\mathbb{N}_0,\leqslant\rangle$
$n,m\in\mathbb{N}_0$	$n<m$	strict order in $\langle\mathbb{N}_0,\leqslant\rangle$
$n,m\in\mathbb{N}_0$	$n-m$	subtraction in $\langle\mathbb{N}_0,+\rangle$
$n\in\mathbb{N}_0$	$n+0$ or $0+n$	neutral element in $\langle\mathbb{N}_0,+\rangle$
$n\in\mathbb{N}_0$	$n\cdot 1$ or $1\cdot n$	neutral element in $\langle\mathbb{N}_0,\cdot\rangle$
$n\in\mathbb{N}_0$	$-n$	inverse element in $\langle\mathbb{N}_0,+\rangle$
$n,m\in\mathbb{Z}$	n+m	sum in $\langle\mathbb{Z},+\rangle$
$n,m\in\mathbb{Z}$	$n\cdot m$	product in $\langle\mathbb{Z},\cdot\rangle$
$n,m\in\mathbb{Z}$	$n\leqslant m$	order in $\langle\mathbb{Z}\leqslant\rangle$
$n,m\in\mathbb{Z}$	$n<m$	strict order in $\langle\mathbb{Z},\leqslant\rangle$
$n,m\in\mathbb{Z}$	$n-m$	subtraction in $\langle\mathbb{Z},-\rangle$
$n\in\mathbb{Z}$	$n+0$ or $0+n$	neutral element in $\langle\mathbb{Z},+\rangle$
$n\in\mathbb{Z}$	$n\cdot 1$ or $1\cdot n$	neutral element in $\langle\mathbb{Z},\cdot\rangle$
$n\in\mathbb{Z}$	$-n$	inverse element in $\langle\mathbb{Z},+\rangle$
$q,r\in\mathbb{Q}$	q+r	sum in $\langle\mathbb{Q},+\rangle$
$q,r\in\mathbb{Q}$	$q\cdot r$	product in $\langle\mathbb{Q},\cdot\rangle$
$q,r\in\mathbb{Q}$	$q\leqslant r$	order in $\langle\mathbb{Q}\leqslant\rangle$
$q,r\in\mathbb{Q}$	$q<r$	strict order in $\langle\mathbb{Q},\leqslant\rangle$
$q,e\in\mathbb{Q}$	$q-r$	subtraction in $\langle\mathbb{Q},-\rangle$
$q,r\in\mathbb{Q}$	q/r	division in $\langle\mathbb{Q},\cdot\rangle$
$q\in\mathbb{Q}$	$q+0$ or $0+q$	neutral element in $\langle\mathbb{Q},+\rangle$
$q\in\mathbb{Q}$	$q\cdot 1$ or $1\cdot q$	neutral element in $\langle\mathbb{Q},\cdot\rangle$
$q\in\mathbb{Q}$	$-q$	inverse element in $\langle\mathbb{Q},+\rangle$
$q,r\in\mathbb{R}$	q+r	sum in $\langle\mathbb{R},+\rangle$
$q,r\in\mathbb{R}$	$q\cdot r$	product in $\langle\mathbb{R},\cdot\rangle$
$q,r\in\mathbb{R}$	$q\leqslant r$	order in $\langle\mathbb{R},\leqslant\rangle$
$q,r\in\mathbb{R}$	$q<r$	strict order in $\langle\mathbb{R},\leqslant\rangle$
$q,e\in\mathbb{R}$	$q-r$	subtraction in $\langle\mathbb{R},-\rangle$
$q,r\in\mathbb{R}$	q/r	division in $\langle\mathbb{R},\cdot\rangle$
$q\in\mathbb{R}$	$q+0$ or $0+q$	neutral element in $\langle\mathbb{R},+\rangle$
$q\in\mathbb{R}$	$q\cdot 1$ or $1\cdot q$	neutral element in $\langle\mathbb{R},\cdot\rangle$
$q\in\mathbb{R}$	$-q$	inverse element in $\langle\mathbb{R},+\rangle$

9.1 Definition and Arithmetic on $\mathbb{R}$

9.1.1 Definition of the real numbers

Definition 9.1. (Dedekind Cut) *A subset $\alpha \subseteq \mathbb{Q}$ is a Dedekind's cut of $\mathbb{Q}$ if the following properties hold*

1. $\alpha \neq \varnothing$
2. $\alpha \neq \mathbb{Q}$ *[which as $\alpha \subseteq \mathbb{Q}$ implies that $\mathbb{Q} \setminus \alpha \neq \varnothing$]*
3. $\forall q \in \alpha \wedge \forall r \in \mathbb{Q} \setminus \alpha$ *we have* $q < r$
4. α *does not have a greatest element [or maximum]*

So a Dedekind cut α divides $\mathbb{Q}$ in two disjoint pieces so that $\alpha \neq \varnothing \neq Q \setminus \alpha$ where every element in α is strict lower of elements in $\mathbb{Q} \setminus \alpha$ and α has not a greatest element. The collection of Dedekind cuts will from the set of real numbers.

Definition 9.2. *The set of real numbers, noted as $\mathbb{R}$ is the set of Dedekind cuts of $\mathbb{Q}$ hence*

$$\mathbb{R} = \{\alpha \subseteq \mathbb{Q} | \alpha \text{ is a Dedekind cut}\}$$

Lemma 9.3. $\forall \alpha \in \mathbb{R}$ *we have* $\forall q \in \alpha$ *and* $\forall r \in \mathbb{Q}$ *with* $r \leqslant q$ *that* $r \in \alpha$.

Proof. Let $\alpha \in \mathbb{R}$, $q \in \alpha$ and $r \in \mathbb{Q}$ with $r \leqslant q$. Assume that $r \notin \alpha$ then $r \in \mathbb{Q} \setminus \alpha$ and by [definition: 9.1] we have $q < r$ contradicting $r \leqslant q$. Hence we must have that $r \in \alpha$. □

We prove now that every rational number can be associated with a Dedekind cut of $\mathbb{Q}$.

Theorem 9.4. (Rational cuts) *If $q \in \mathbb{Q}$ then $\alpha_q = \{r \in \mathbb{Q} | r < q\}$ is a Dedekind cut. Dedekind cuts of this forms are called rational cuts. Furthermore we have:*

1. $\alpha_q = \alpha_r \Leftrightarrow q = r$
2. α *is a rational cut* $\Leftrightarrow$ $q = \min(\mathbb{Q} \setminus \alpha)$ *exist and in that case* $\alpha = \alpha_q$

Proof. First we prove that given $q \in \mathbb{Q}$ $\alpha_q = \{r \in \mathbb{Q} | r < q\}$ is a cut.

1. By [theorem: 8.26] $q - 1 < q$ hence $q - 1 \in \alpha_q$ proving that $\alpha_q \neq \varnothing$.
2. As $q < q$ is false we have that $q \in \mathbb{Q} \setminus \alpha_q$ so that $\mathbb{Q} \setminus \alpha_q \neq \varnothing$.
3. If $r \in \alpha_q$ and $s \in \mathbb{Q} \setminus \alpha_q$ then $r < q$ and $\neg(s < q) \Rightarrow q \leqslant s$ so that $r < s$.
4. Assume that m is a greatest element of α_q then $m \in \alpha_q$ and $\forall r \in \alpha_q$ we have $r \leqslant m$. As $m \in \alpha_q$ we have that $m < q$, using the density of $\mathbb{Q}$ [see theorem: 8.37] there exist a $r \in \mathbb{Q}$ such that $m < r < q$. As $r < q$ we have that $r \in \alpha_q$ so that $r \leqslant m$ contradicting $m < r$. So the assumption is false proving that α_q has no greatest element.

Next we prove (1) and (2)

1.

   ⇒. If $\alpha_q = \alpha_r$ then if $q \neq r$ we have either

   q < r. then $q \in \alpha_r$ and so that $q \in \alpha_q$ resulting in the contradiction $q < q$.

   r < q. then $r \in \alpha_q$ and so that $r \in \alpha_r$ resulting in the contradiction $r < r$.

   so we must have $q = r$.

   ⇐. $s \in \alpha_q \Leftrightarrow s \in \mathbb{Q} \wedge s < q \underset{q=r}{\Leftrightarrow} s \in \mathbb{Q} \wedge s < r \Leftrightarrow s \in \alpha_r$ hence $\alpha_q = \alpha_s$

2.

⇒. If α is a rational cut then there exist a $q\in\mathbb{Q}$ such that $\alpha=\{r\in\mathbb{Q}|r<q\}$. So

$$\begin{aligned} s\in\mathbb{Q}\backslash\alpha &\Leftrightarrow s\in\mathbb{Q}\wedge\neg(s<q)\\ &\Leftrightarrow s\in\mathbb{Q}\wedge q\leqslant s\\ &\Leftrightarrow s\in\{s\in\mathbb{Q}|q\leqslant s\}\end{aligned}$$

proving that $\mathbb{Q}\backslash\alpha=\{s\in\mathbb{Q}|q\leqslant s\}$. So $q\in\{s\in\mathbb{Q}|q\leqslant s\}$ and $\forall s\in Q\backslash\alpha$ we have $q\leqslant s$ proving that $q=\min(\mathbb{Q}\backslash a)$ and $\alpha=\{r\in\mathbb{Q}|r<q\}=\alpha_q$

⇐. If $q=\min(\mathbb{Q}\backslash\alpha)$ exists then $q\in\mathbb{Q}\backslash\alpha$ and $\forall r\in\mathbb{Q}\backslash a$ we have $q\leqslant r$. If now $r\in\alpha$ then by the definition of a cut we have $r<q$, hence $r\in\{r\in\mathbb{Q}|r<q\}=\alpha_q$. Further if $r\in\alpha_q$ then $r<q$, assume that $r\notin\alpha$ then we have $q\leqslant r$ contradicting $r<q$, so we must have that $r\in\alpha$. Hence we have that

$$\alpha=\alpha_q \text{ where } q=\min(\mathbb{Q}\backslash\alpha)$$ □

Corollary 9.5. $\mathbb{R}\neq 0$

Proof. As $0,1\in\mathbb{Q}$ we have that $\alpha_0,\alpha_1\in\mathbb{R}$ proving that $\mathbb{R}\neq\varnothing/$ □

We embed now the rational numbers in the set of reals.

Definition 9.6. *The set* $\mathbb{Q}_{\mathbb{R}}$ *is defined by*

$$\mathbb{Q}_{\mathbb{R}}=\{\alpha_q|q\in\mathbb{Q}\}\subseteq\mathbb{R}$$

where $\alpha_q=\{r\in\mathbb{Q}|r<q\}$

To make the above a embedding we need a bijection between $\mathbb{Q}$ and $\mathbb{Q}_{\mathbb{R}}$ and once we have defined sum, product and order prove that it is a field and order isomorphism. We start with providing a bijection.

Theorem 9.7. $i_{\mathbb{Q}\rightarrow\mathbb{R}}:\mathbb{Q}\rightarrow\mathbb{Q}_{\mathbb{R}}$ *defined by* $i_{\mathbb{Q}}(q)=\alpha_q$ *is a bijection.*

Proof. We have

reflexivity. If $i_{\mathbb{Q}\rightarrow\mathbb{R}}(q)=i_{\mathbb{Q}\rightarrow\mathbb{R}}(r)$ then $\alpha_q=\alpha_r$ so that by [theorem: 9.4] $q=r$.

surjective. If $\alpha\in\mathbb{Q}_{\mathbb{Q}}$ we have a $q\in\mathbb{Q}$ such that $\alpha=\alpha_q=i_{\mathbb{Q}\rightarrow\mathbb{R}}(q)$ □

Corollary 9.8. *The set* $\mathbb{Q}_{\mathbb{R}}$ *is denumerable.*

Proof. As $\mathbb{Q}$ is denumerable we have that $\mathbb{N}_0\approx\mathbb{Q}$, further from the previous theorem [theorem: 9.7] we have $\mathbb{Q}\approx\mathbb{Q}_{\mathbb{R}}$ so that $\mathbb{N}_0\approx\mathbb{Q}_{\mathbb{R}}$. Hence $\mathbb{Q}_{\mathbb{R}}$ is denumerable. □

Theorem 9.9. (Gap theorem) *If* $\alpha\in\mathbb{R}$ *then* $\forall\varepsilon\in\mathbb{Q}$ *with* $0<\varepsilon$ *there* $\exists q\in\alpha$ *and* $\exists r\in\mathbb{Q}\backslash\alpha$ *such that*

$$r-q=r+(-q)<\varepsilon$$

Proof. Let $\alpha\in\mathbb{R}$ and $\varepsilon\in\mathbb{Q}\backslash\{0\}$. By the definition of a cut there exist a $q'\in\alpha$ and a $r'\in\mathbb{Q}\backslash\alpha$ such that $q'<r'$, so $0<r'+(-q')=r'-q'$ and we have by the Archimedean property [see theorem: 8.36] the existence of a $k\in\mathbb{N}_{0,\mathbb{Q}}$ such that $r'-q'<k\cdot\varepsilon$. If $k=0$ then we would have that $0<r'-k'<0$ a contradiction, so $k\neq 0$ which by [theorem: 8.35] proves that $0<k$. Applying [theorems: 8.29, 4.73] we have that $0<k^{-1}$, so multiplying both sides of $r'-q'<k\cdot\varepsilon$ gives

$$k^{-1}\cdot(r'-q')<\varepsilon \tag{9.1}$$

Define now

$$A=\{n\in\mathbb{N}_{0,\mathbb{Q}}|q'+(n\cdot k^{-1})\cdot(r'-q')\notin\alpha\}\subseteq\mathbb{N}_{0,\mathbb{Q}}$$

As $q'+(k\cdot k^{-1})\cdot(r'-q')=q'+(r'-q')=r'\in\mathbb{Q}\backslash\alpha$ it follows that $k\in A$ so that $A\neq 0$, as $\mathbb{N}_{0,\mathbb{Q}}$ is well ordered [see theorem: 8.38] it follows that $k'=\min(A)$ exist. If $k'=0$ then as $k'\in A$ we would have $q'=q'+(0\cdot k^{-1})\cdot(r'-q')\notin\alpha$ contradicting $q'\in a$, so we must have that $k'\neq 0$ and using [theorem: 8.35] it follows that $1\leqslant k'$, hence $0\leqslant k'-1$, giving by [theorem: 8.35] that $k'-1\in\mathbb{N}_{0,\mathbb{Q}}$. As by [theorem: 8.26] $k'-1<k'$ we have, as $k'=\min(A)$, that $k'-1\notin A$ so that

$$q'+((k'-1)\cdot k^{-1})\cdot(r'-q')\in\alpha$$

Define now $q = q' + ((k' - 1) \cdot k^{-1}) \cdot (r' - q')$ and $r = q' + (k' \cdot k^{-1}) \cdot (r' - q')$ then we have

$$q \in \alpha \text{ and } r \in \mathbb{Q} \backslash \alpha$$

Next

$$\begin{aligned} r - q &= (q' + (k' \cdot k^{-1}) \cdot (r' - q')) - (q' + ((k' - 1) \cdot k^{-1}) \cdot (r' - q')) \\ &= (q' + k' - (q' + (k' - 1))) \cdot k^{-1} \cdot (r' - q') \\ &= k^{-1} \cdot (r' - q') \\ &< \varepsilon \quad \text{[see eq: 9.1]} \end{aligned}$$

□

Theorem 9.10. (Negative cut) *If $\alpha \in \mathbb{R}$ then $-\alpha$ defined by*

$$-\alpha = \{r | -r \in \mathbb{Q} \backslash \alpha \text{ such that } \exists t \in \mathbb{Q} \backslash \alpha \vDash t < -r\}$$

*is a Dedekind cut called the **negative cut**.*

Proof.

1. As α is a Dedekind cut we have by [definition: 9.1 (2)] that $\mathbb{Q} \backslash \alpha \neq \varnothing$ so there exist a $q \in \mathbb{Q} \backslash \alpha$. Assume that $q + 1 \in \alpha$ then by [definition: 9.1 (3)] we have $q + 1 < q$ a contradiction, so we must have that $q + 1 \notin \alpha$ or $q + 1 \in \mathbb{Q} \backslash \alpha$. Hence we have $-(-(q + 1)) = q + 1 \in \mathbb{Q} \backslash \alpha$ and $q \in \mathbb{Q} \backslash \alpha$ with $q < q + 1 = -(-(q + 1))$ proving that $-(q + 1) \in -\alpha$ or that

$$-\alpha \neq \varnothing$$

2. As α is a Dedekind cut we have by [definition: 9.1 (1)] that $\alpha \neq \varnothing$ so there exist a $q \in \alpha$ hence $q \notin \mathbb{Q} \backslash \alpha$. If $-q \in -\alpha$ then $q = -(-q) \in \mathbb{Q} \backslash \alpha$ contradicting $q \notin \mathbb{Q} \backslash \alpha$ hence we must have that $-q \notin -\alpha$ proving that

$$-\alpha \neq \mathbb{Q}$$

3. Let $q \in -\alpha$ and $s \in \mathbb{Q} \backslash -\alpha$. Assume that $s \leqslant q$ then by [theorems: 8.29, 4.73]

$$-q \leqslant -s \tag{9.2}$$

As $q \in -\alpha$ we have that

$$-q \in \mathbb{Q} \backslash \alpha \text{ and } \exists t \in \mathbb{Q} \backslash \alpha \vDash t < -q \tag{9.3}$$

If $-s \in \alpha$ then as $-q \in \mathbb{Q} \backslash \alpha$ we have by [definition: 9.1 (3)] that $-s < -q$ contradicting [eq: 9.2] hence we must have that $-s \notin \alpha$ so that $-s \in \mathbb{Q} \backslash \alpha$. Using [eq: 9.2] and [eq: 9.3] we have $\exists t \in \mathbb{Q} \backslash \alpha$ such that $t < -q \leqslant -s$ so we have that $s \in -\alpha$ contradicting $s \in \mathbb{Q} \backslash -\alpha$. So the assumption is wrong and we have

$$q < s$$

4. Assume that $-\alpha$ has a greatest element m then

$$m \in -\alpha \text{ and } \forall r \in -\alpha \text{ we have } r \leqslant m \tag{9.4}$$

As $m \in -a$ we have that

$$-m \in \mathbb{Q} \backslash \alpha \text{ and } \exists t \in \mathbb{Q} \backslash \alpha \vDash t < -m \underset{\text{[theorems: 8.29, 4.73]}}{\Rightarrow} m < -t \tag{9.5}$$

For $\mathbb{Q} \backslash \alpha$ we have now two cases to consider:

$\min(\mathbb{Q} \backslash \alpha)$ does not exist. As $t \in \mathbb{Q} \backslash \alpha$ and $\min(\mathbb{Q} \backslash \alpha)$ does not exist there exist a $s \in \mathbb{Q} \backslash \alpha$ such that $s < t$, hence we have $-(-t) \in \mathbb{Q} \backslash \alpha \wedge s < -(-t)$ proving that $-t \in -\alpha$ hence by [eq: 9.4] that $-t \leqslant m$ contradicting [eq 9.5].

$\min(\mathbb{Q} \backslash \alpha)$ exist. As $-m \in \mathbb{Q} \backslash \alpha$ we have $\min(\mathbb{Q} \backslash \alpha) \leqslant -m$, further as $t \in \mathbb{Q} \backslash \alpha \wedge t < -m$ we have $-m \neq \min(\mathbb{Q} \backslash \alpha)$ so that $\min(\mathbb{Q} \backslash \alpha) < -m$. Using the density of $\mathbb{Q}$ [see 8.37] there exist a $s \in \mathbb{Q}$ such that

$$\min(\mathbb{Q}) < s < -m \tag{9.6}$$

If $s \in \alpha$ then as $\min(\mathbb{Q} \backslash \alpha) \in \mathbb{Q} \backslash \alpha$ we have by [definition: 9.1 (3)] that $s < \min(\mathbb{Q} \backslash \alpha)$ contradicting $\min(\mathbb{Q} \backslash \alpha) < s$, so we must have that $s \in \mathbb{Q} \backslash \alpha$. Hence $s = -(-s) \in \mathbb{Q} \backslash \alpha$, $\min(\mathbb{Q} \backslash \alpha) \in \mathbb{Q} \backslash \alpha$ and $\min(\mathbb{Q}) < s = -(-s)$ proving that $-s \in -\alpha$. Using [eq: 9.4] it follows that $-s \leqslant m$ or $-m \leqslant s$ contradicting [eq: 9.6]

So in all cases we reach a contradiction so that the assumption is wrong. Hence

$$-\alpha \text{ has no greatest element}$$ □

For rational cuts there is a simple expression for negative cuts.

Theorem 9.11. *If* $q \in \mathbb{Q}$ *then* $-\alpha_q = \alpha_{-q}$

Proof. Using [theorem: 9.4] we have that

$$\min(\mathbb{Q} \backslash \alpha_q) \text{ exist and } q = \min(\mathbb{Q} \backslash \alpha_q)$$

If $x \in -\alpha_q$ then $-x \in \mathbb{Q} \backslash \alpha_q$ $\exists t \in \mathbb{Q} \backslash \alpha$ such that $t < -x$ so that $-x \neq \min(\mathbb{Q} \backslash \alpha) = q$. As $\alpha_q = \{r \in \mathbb{Q} | r < q\}$ and $-x \in \mathbb{Q} \backslash \alpha_q$ we have $q \leqslant -x$ or $x \leqslant -q$ which as $-x \neq q \Rightarrow x \neq -q$, gives $x < -q$. Hence $x \in \{r \in \mathbb{Q} | r < -q\} = \alpha_{-q}$ proving that

$$-\alpha_q \subseteq \alpha_{-q} \tag{9.7}$$

If $x \in \alpha_{-q}$ then $x < -q$, so that $q < -x$ hence $-x \notin \{x \in \mathbb{Q} | x < q\} = \alpha_q$ and $q < -x$ where $q = \min(\mathbb{Q} \backslash \alpha_q) \in \mathbb{Q} \backslash \alpha_q$ proving that $x \in -\alpha_q$. So $\alpha_{-q} \subseteq \alpha_q$, combining this with [eq: 9.7] gives

$$-\alpha_q = \alpha_{-q}$$ □

9.1.2 Arithmetic in $\mathbb{R}$

9.1.2.1 Addition in $\mathbb{R}$

Definition 9.12. *If* $\alpha, \beta \in \mathbb{R}$ *then we define* $\alpha + \beta$ *by*

$$\alpha + b = \{q + r | q \in \alpha \wedge r \in \beta\}$$

Before we can use the above definition to define the addition operator in $\mathbb{R}$ we must prove that $\alpha + \beta$ is a Dedekind cut, hence a element of $\mathbb{R}$. First we need a little lemma.

Lemma 9.13. $\forall \alpha \in \mathbb{R}$ *and* $\forall \varepsilon \in \mathbb{Q}$ *with* $0 < \varepsilon$ *there exist a* $r \in \alpha$ *such that* $r + \varepsilon \in \mathbb{Q} \backslash \alpha$

Proof. Let $\alpha \in \mathbb{R}$ and $\varepsilon \in \mathbb{Q}$ such that $0 < \varepsilon$. Using [theorem: 9.9] there exist a $q \in \alpha$ and a $r \in \mathbb{Q} \backslash \alpha$ such that $r - q < \varepsilon$. Assume that $q + \varepsilon \in \alpha$ then we have by the definition of a cut that $q + \varepsilon < r$ so that $\varepsilon < r - q$ contradicting $r - q < \varepsilon$. Hence we must have that $q + \varepsilon \notin \alpha$ or $q + \varepsilon \in \mathbb{Q} \backslash \alpha$. □

Theorem 9.14. $\forall \alpha, \beta \in \mathbb{R}$ *we have that* $\alpha + \beta \in \mathbb{R}$*, hence* $+: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ *where* $+(\alpha, \beta) = \alpha + \beta$ *is a operator on* $\mathbb{R}$.

Proof. Given Dedekind cuts α and β we must prove that $\alpha + \beta$ is a Dedekind cut.

1. As $\alpha \neq \varnothing$ and $\beta \neq \varnothing$ it follows that $\exists a \in \alpha$ and $\exists b \in \beta$ so that $a + b \in \{q + r | q \in \alpha \wedge r \in \beta\} = \alpha + \beta$, proving that

$$\alpha + \beta \neq \varnothing$$

2. Given $\varepsilon = \frac{1}{2} \in \mathbb{Q}$ we can as $0 < \varepsilon$ use [lemma: 9.13] to find a $r' \in \alpha$ and a $s' \in \beta$ such that $q' + \varepsilon \in \mathbb{Q} \backslash \alpha$ and $r' + \varepsilon \in \mathbb{Q} \backslash \beta$. Assume that $q' + y' + 1 \in \alpha + \beta$ then there exists a $q \in \alpha$ and $r \in \beta$ such that

$$q' + y' + 1 = q + r \tag{9.8}$$

As $q \in \alpha \wedge q' + \varepsilon \in \mathbb{Q} \backslash \alpha$ and $r \in \beta \wedge r' + \varepsilon \in \mathbb{Q} \backslash \beta$ it follows from the definition of Dedekind cuts that $q < q' + \varepsilon$ and $r < r' + \varepsilon$ so that $q + r < q' + r' + 2 \cdot \varepsilon = q' + r' + 1 \underset{\text{[eq: 9.8]}}{=} q + r$ giving the contradiction that $q + r < q + r$. So we must have that $q' + r' + 1 \notin \alpha + \beta$ proving that

$$\alpha + \beta \neq \mathbb{Q}$$

3. Let $s \in \alpha + \beta$ and $t \in \mathbb{Q} \backslash (\alpha + \beta)$ then there exists a $q \in \alpha$ and a $r \in \beta$ such that $s = q + r$. Assume now that that $t \leqslant s$, then $t \leqslant q + r$, so that $t - r \leqslant q$, by [theorem: 9.3] it follows then that $t - r \in \alpha$. From this and the fact that $r \in \beta$ it follows that $t = (t - r) + r \in \alpha + \beta$ contradicting $t \in \mathbb{Q} \backslash (\alpha + \beta)$, hence we must have that $s < t$.

4. Assume that $\alpha+\beta$ has a greatest element m then we have

$$m\in\alpha+\beta \text{ and } \forall q\in\alpha+\beta \text{ we have } q\leqslant m \tag{9.9}$$

As $m\in\alpha+\beta$ there exists a $q\in\alpha$ and a $r\in\beta$ such that $m=q+r$. As α has no greatest element there exist a $q'\in\alpha$ such that $q<q'$, hence $m=q+r<q'+r$ which as $q'+r\in\alpha+\beta$ contradicts [eq: 9.9]. So the assumption is wrong, hence $\alpha+\beta$ has no greatest element. □

Theorem 9.15. *$\langle\mathbb{R},+\rangle$ is a Abelian group with neutral element $0=\alpha_0=\{q\in\mathbb{Q}|q<0\}$ and if $\alpha\in\mathbb{R}$ then $-\alpha$ [the negative cut of α] is the inverse element of α.*

Proof. We make use of the fact that $\langle\mathbb{Q},+\rangle$ is a Abelian group [see theorem: 8.8]. So we have

associativity. If $\alpha,\beta,\gamma\in\mathbb{R}$ then

$$\begin{aligned} z\in(\alpha+\beta)+\gamma &\Leftrightarrow z=r+s\wedge r\in(a+\beta)\wedge s\in\gamma \\ &\Leftrightarrow z=(q+t)+s\wedge q\in\alpha\wedge t\in\beta\wedge s\in\gamma \\ &\Leftrightarrow z=q+(t+s)\wedge q\in\alpha\wedge t\in\beta\wedge s\in\gamma \\ &\Leftrightarrow z=q+r\wedge q\in\alpha\wedge r\in\beta+\gamma \\ &\Leftrightarrow z\in\alpha+(\beta+\gamma) \end{aligned}$$

proving that $(\alpha+\beta)+\gamma=\alpha+(\beta+\gamma)$.

commutativity. If $\alpha,\beta\in\mathbb{R}$ then

$$\begin{aligned} z\in\alpha+\beta &\Leftrightarrow z=r+s\wedge r\in\alpha\wedge s\in\beta \\ &\Leftrightarrow z=s+r\wedge r\in\alpha\wedge s\in\beta \\ &\Leftrightarrow z\in\beta+\alpha \end{aligned}$$

neutral element. Let $\alpha\in\mathbb{R}$ and take $\alpha_0=\{q\in\mathbb{Q}|q<0\}$. If $q\in\alpha+\alpha_0$ then there exists $r\in\alpha$ and $s\in\alpha_0$ such that $q=r+s$, as $s\in\alpha_0$ we have that $s<0$ so that $q=r+s<r$, using [theorem: 9.3] it follows then that $q\in\alpha$. Hence we have that

$$\alpha+\alpha_0\subseteq\alpha \tag{9.10}$$

If $q\in\alpha$ then as α has no maximum there exist a $r\in\alpha$ such that $q<r$, so $q-r<0$ so that $q-r\in\alpha_0$, hence $q=(q-r)+r\in\alpha+\alpha_0$. So $\alpha\subseteq\alpha+\alpha_0$ which together with [eq: 9.10] proves that

$$\alpha=\alpha+\alpha_0\underset{\text{commutativity}}{=}\alpha_0+\alpha$$

inverse element. Let $\alpha\in\mathbb{R}$ and take

$$-\alpha\underset{\text{[theorem: 9.10]}}{=}\{r|-r\in\mathbb{Q}\backslash\alpha \text{ such that } \exists t\in\mathbb{Q}\backslash\alpha\vDash t<-r\}$$

then we have the following cases to consider:

min $(\mathbb{Q}\backslash\alpha)$ does not exist. If $q\in\alpha_0$ then $q<0$ so that by [theorems: 8.29, 4.73] $0<-q$, by [theorem: 9.13] there exist a $r\in\alpha$ such that $-(q+(-r))=r+(-q)\in\mathbb{Q}\backslash\alpha$, as min $(\mathbb{Q}\backslash\alpha)$ does not exist there exist a $s\in\mathbb{Q}\backslash\alpha$ such that $s<r+(-q)=-(q+(-r))$. So we conclude that $q+(-r)\in-\alpha$, hence $q=(q+(-r))+r\in(-\alpha)+\alpha$ giving

$$\alpha_0\subseteq(-\alpha)+\alpha \tag{9.11}$$

If $q\in(-\alpha)+\alpha$ there exist a $r\in-\alpha$ and $s\in\alpha$ such that $q=r+s$. As $r\in-\alpha$ we have that $-r\in\mathbb{Q}\backslash\alpha$, using [definition: 9.1 (3)] we have then $s<-r$ so that $q=s+r<0$ proving that $q\in\alpha_0$. Hence $(-\alpha)+\alpha\in\alpha_0$ which by [eq: 9.11] proves that

$$\alpha_0=(-\alpha)+\alpha$$

min $(\mathbb{Q}\backslash\alpha)$ exist. Let $m=\min(\mathbb{Q}\backslash\alpha)$ then by [theorem: 9.4]

$$\alpha=\alpha_m=\{q\in\mathbb{Q}|q<m\}$$

Further by [theorem: 9.11] we have then that

$$-\alpha=-\alpha_m=\alpha_{-m}$$

so that

$$\alpha+(-\alpha)=\alpha_m+\alpha_{-m}$$

If $q\in\alpha+(-\alpha)$ there exist a $r\in\alpha_m$ and a $s\in\alpha_{-m}$ such that $q=r+s$. As $r\in\alpha_m$ we have $r<m$ and as $s\in\alpha_{-m}$ $s<-m$ so that $q=r+s<m+(-m)=0$ proving that $q\in\alpha_0$. Hence

$$\alpha+(-\alpha)\subseteq\alpha_0 \tag{9.12}$$

Further if $q\in\alpha_0$ then $q<0$ then as $0<\frac{1}{2}$ we have that $\frac{1}{2}\cdot x<0$ so that $m+\frac{1}{2}\cdot q<m$ and $-m+\frac{1}{2}\cdot q<-m$ so that $m+\frac{1}{2}\cdot q\in a_m$ and $-m+\frac{1}{2}\cdot q\in\alpha_{-m}$ hence

$$\left(m+\frac{1}{2}\cdot q\right)+\left(-m+\frac{1}{2}\cdot q\right)\in\alpha_m+\alpha_{-m}=\alpha+(-\alpha)$$

which as $\left(m+\frac{1}{2}\cdot q\right)+\left(-m+\frac{1}{2}\cdot q\right)=\frac{1}{2}\cdot q+\frac{1}{2}\cdot q=\left(\frac{1}{2}+\frac{1}{2}\right)\cdot q=q$ proves that $q\in\alpha+(-\alpha)$. So $\alpha_0\subseteq\alpha+-\alpha\cdot$ which combined with [eq: 9.12] gives

$$\alpha_0=\alpha+(-\alpha)\underset{\text{commutativity}}{=}(-\alpha)+\alpha$$

□

9.1.2.2 Multiplication

Before we can define multiplication we have to divide the set of real numbers in the positive real numbers, the negative real numbers and the 0 element. We will then define multiplication of positive real numbers and extend it to all the real numbers.

Definition 9.16. *The set of positive real numbers noted by $\mathbb{R}^+$ and negative real numbers noted by $\mathbb{R}^-$ is defined by*

$$\mathbb{R}^+=\{\alpha\in\mathbb{R}|0\in\alpha\}\subseteq\mathbb{R}$$
$$\mathbb{R}^-=\{\alpha|-\alpha\in\mathbb{R}^+\}\subseteq\mathbb{R}$$

Further we define the set $\mathbb{R}_0^+$ of non negative numbers and $\mathbb{R}_0^-$ of non positive numbers by

$$\mathbb{R}_0^+=\mathbb{R}^+\bigcup\{0\}$$
$$\mathbb{R}_0^-=\mathbb{R}^-\bigcup\{0\}$$

The following theorem shows that $\mathbb{R}_0^+\neq\mathbb{R}^+$ and $\mathbb{R}_0^-\neq\mathbb{R}^-$

Theorem 9.17. $\mathbb{R}=\mathbb{R}^+\bigcup\mathbb{R}^-\bigcup\{0\}$ *where* $\mathbb{R}^+\bigcap\mathbb{R}^-=\varnothing$, $\mathbb{R}^+\bigcap\{0\}=\varnothing$ *and* $\mathbb{R}^-\bigcap\{0\}=\varnothing$

Note 9.18. Be careful here, 0 can mean either $0\in\mathbb{Z}$ or $0\in\mathbb{R}$ in which case $0=\alpha_0$

Proof. As $\{0\}\subseteq\mathbb{R}$, $\mathbb{R}^+\subseteq\mathbb{R}$ and $\mathbb{R}^-\subseteq\mathbb{R}$ we have

$$\mathbb{R}^+\bigcup\mathbb{R}^-\bigcup\{0\}\subseteq\mathbb{R} \tag{9.13}$$

If $\alpha\in\mathbb{R}$ then we have either:

- $\mathbf{0\in\alpha}$. then $\alpha\in\mathbb{R}^+$ so that $\alpha\in\mathbb{R}^+\bigcup\mathbb{R}^-\bigcup\{0\}$
- $\mathbf{0\notin\alpha}$. then we have either:
  - $\mathbf{min\,(\mathbb{Q}\setminus\alpha)}$ **does not exist.** As $0\notin\alpha$ we have $-0=0\in\mathbb{Q}\setminus\alpha$ and as $\min(\mathbb{Q}\setminus\alpha)$ does not exist there exist a $s\in\mathbb{Q}\setminus\alpha$ such that $s<0=-0$ so that $0\in-\alpha$. Hence $-\alpha\in\mathbb{R}^+$ proving that $\alpha\in\mathbb{R}^-\subseteq\mathbb{R}^+\bigcup\mathbb{R}^-\bigcup\{0\}$.
  - $\mathbf{min\,(\mathbb{Q}\setminus\alpha)}$ **exist.** Then by [theorem: 9.4] $\alpha=\alpha_m$ where $m=\min(\mathbb{Q}\setminus\alpha)$
    - $\mathbf{0=m}$. Then $\alpha=\alpha_0=0$ so that $\alpha\in\{0\}\subseteq\mathbb{R}^+\bigcup\mathbb{R}^-\bigcup\{0\}$.
    - $\mathbf{0<m}$. Then $0\in\alpha_m=\alpha$ so that $\alpha\in\mathbb{R}^+\subseteq\mathbb{R}^+\bigcup\mathbb{R}^-\bigcup\{0\}$.
    - $\mathbf{m<0}$. Then $0\notin\alpha_m=\alpha$ so that $-0=0\in\mathbb{Q}\setminus\alpha$ and as $m<0=-0$, $m\in\mathbb{Q}\setminus\alpha$ it follows that $0\in-\alpha$, proving that $-\alpha\in\mathbb{R}^+$, hence $\alpha\in\mathbb{R}^-\subseteq\mathbb{R}^+\bigcup\mathbb{R}^-\bigcup\{0\}$

So in all cases we have $\alpha\in\mathbb{R}^+\bigcup\mathbb{R}^-\bigcup\{0\}$ proving $\mathbb{R}\subseteq\mathbb{R}^+\bigcup\mathbb{R}^-\bigcup\{0\}$ which combined with [eq: 9.13] proves

$$\mathbb{R}=\mathbb{R}^+\bigcup\mathbb{R}^-\bigcup\{0\}$$

Now as $0=\alpha_0=\{q\in\mathbb{Q}|q<0\}$ we have that $0\notin\alpha_0$ hence $0=\alpha_0\notin\mathbb{R}^+$ proving that

$$\mathbb{R}^+\bigcap\{0\}=\varnothing$$

Using [theorem: 9.11] it follows that $-\alpha_0=\alpha_{-0}=\alpha_0$ so that $-0=-\alpha_0\notin\mathbb{R}^+$ hence $0\notin\mathbb{R}^-$ proving that

$$\mathbb{R}^-\bigcap\{0\}=\varnothing$$

Finally if $\alpha\in\mathbb{R}^+\bigcap\mathbb{R}^-$ then $0\in\alpha$ and $0\in-\alpha$, as $0\in-\alpha$ then at least $-0\in\mathbb{Q}\backslash\alpha$ so that $0=-0\notin\alpha$ contradicting $0\in\alpha$. So we have

$$\mathbb{R}^+\bigcap\mathbb{R}^-=\varnothing$$ □

Defining multiplication in $\mathbb{R}$ is difficult. we first define multiplication for $\mathbb{R}^+$ and extend it later to $\mathbb{R}$.

Definition 9.19. *Given $\alpha,b\in\mathbb{R}^+$ we define $A=\alpha\odot\beta$ by*

$$\begin{aligned}\alpha\odot\beta &= \mathbb{Q}_0^-\bigcup\{s\cdot t|(s,t)\in\alpha\times\beta\wedge 0<s\wedge 0<t\}\\ &= \{r\in\mathbb{Q}|r\leqslant 0\}\bigcup\{s\cdot t|(s,t)\in\alpha\times\beta\wedge 0<s\wedge 0<t\}\end{aligned}$$

Theorem 9.20. $\forall\alpha,\beta\in\mathbb{R}^+$ *we have that* $\alpha\odot\beta\in\mathbb{R}^+$

Proof. First we prove that $\alpha\odot\beta$ is a Dedekind cut.

1. As $0\in\{r\in\mathbb{Q}|r\leqslant 0\}$ it follows that

$$\alpha\odot\beta\neq\varnothing$$

2. As $\alpha,\beta\in\mathbb{R}^+$ it follows that $0\in\alpha\wedge 0\in\beta$ and as α,β do not have a greatest element we have

$$\exists s_1\in\alpha,\exists t_1\in\beta \text{ such that } 0<s_1\wedge 0<t_1 \tag{9.14}$$

As $0<1$ we have by [theorem: 9.13] that

$$\exists s_2\in\alpha,\exists t_2\in\beta \text{ such that } s_2+1\in\mathbb{Q}\backslash\alpha\wedge t_2+1\in\mathbb{Q}\backslash\beta \tag{9.15}$$

Take now

$$s=\max(\{s_1,s_2\}) \text{ and } t=\max(\{t_1,t_2\}) \tag{9.16}$$

If $s\notin\alpha$ then $s\in\mathbb{Q}\backslash\alpha$ so that by [definition: 9.1 (3)] and [eqs: 9.14, 9.15] we have $s1<s$ and $s_2<s$ contradicting the fact that $s\in\{s_1,s_2\}$, so we must have that $s\in\alpha$. Likewise if $t\notin\beta$ then $t\in\mathbb{Q}\backslash\beta$ so that by [definition: 9.1 (3)] and [eq: 9.14, 9.15] we have $t_1<t$ and $t_2<t$ contradicting the fact that $t\in\{t_1,t_2\}$, so we must have that $s\in\beta$. So we have

$$s\in\alpha\wedge t\in\beta \text{ and by [eqs: 9.14,9.16] that } 0<s\wedge 0<t \tag{9.17}$$

If $s+1\in\alpha$ then by [definition: 9.1 (3)] and [eq: 9.15] we have $s+1<s_2+1\Rightarrow s<s_2$ contradicting $s=\max(s_1,s_2)$. Likewise if $t+1\in\beta$ then by [definition: 9.1 (3)] and [eq: 9.15] we have $t+1<t_2+t\Rightarrow t<t_2$ contradicting $t=\max(t_1,t_2)$. So we must have

$$s+1\in\mathbb{Q}\backslash\alpha \text{ and } t+1\in\mathbb{Q}\backslash\alpha \tag{9.18}$$

Assume now that $s\cdot t+s+t+1\in\alpha\odot\beta$. As $0<s\wedge 0<t$ we have that $0<s\cdot t$ giving $0<s\cdot t+s+t+1$ so that $s\cdot t+s+t+1\notin\{q\in\mathbb{Q}|q\leqslant 0\}$ so we must have that

$$s\cdot t+s+t+1\in\{s\cdot t|(s,t)\in\alpha\times\beta\wedge 0<s\wedge 0<t\}$$

hence there exists $s'\in\alpha$ and $t'\in\beta$ with $0<s'\wedge 0<t'$ such that $s\cdot t+s+t+1=s'\cdot t'$. Using [definition: 9.1 (3)] and [eq: 9.18] we have that $s'<s+1$ and $t'<t+1$ so $s'\cdot t'<(s+1)\cdot t'$ and $t'\cdot(s+1)<(t+1)\cdot(s+1)$, hence $s'\cdot t'<(s+1)\cdot(t+1)=s\cdot t+s+t+1=s'\cdot t'$ giving the contradiction $s'\cdot t'<s'\cdot t'$. Hence the assumption is false so that $s\cdot t+s+t+1\notin\alpha\odot\beta$ proving that

$$\alpha\odot\beta\neq\mathbb{Q}$$

3. Let $q\in\alpha\odot\beta$ and $r\in\mathbb{Q}\backslash\alpha\odot\beta$ then for q we have either:

   $q\in\{r\in\mathbb{Q}|r\leqslant 0\}$. Then $q\leqslant 0$ further as $r\in\mathbb{Q}\backslash\alpha\odot\beta$ we have that $r\notin\{r\in\mathbb{Q}|r\leqslant 0\}$ so that $0<r$ from which it follows that $q<r$.

$q \notin \{r \in \mathbb{Q} | r \leqslant 0\}$. Then $q \in \{s \cdot t | (s, t) \in \alpha \times \beta \wedge 0 < s \wedge 0 < t\}$ so that

$$\exists s' \in \alpha, \exists t' \in \beta \text{ with } 0 < s' \wedge 0 < t' \text{ such that } q = s' \cdot t' \tag{9.19}$$

Assume now that $r \leqslant q$. As $r \in \mathbb{Q} \setminus \alpha \odot \beta$ we have that $r \neq q$ [as $q \in \alpha \odot \beta$] and $r \notin \{r \in \mathbb{Q} | r \leqslant 0\}$ so that $0 < r$. Hence we have $0 < r < q$ or multiplying by r^{-1} [which by [theorems: 8.29, 4.73] exists and $0 < r^{-1}$] we have $1 = r \cdot r^{-1} < q \cdot r^{-1}$ or if we define $t = q \cdot r^{-1}$, it follows that

$$1 < t \text{ and } t \cdot r = q \tag{9.20}$$

Using the above, we have by [theorems: 8.29, 4.73] that $0 < t^{-1} < 1$ so that by multiplying by s' we have, as $0 < s'$, that

$$t^{-1} \cdot s' < s' \tag{9.21}$$

If now $t^{-1} \cdot s' \notin \alpha$ then $t^{-1} \cdot s' \in \mathbb{Q} \setminus \alpha$ which, as $s' \in \alpha$, means by [definition: 9.1 (3)] that $s' < t^{-1} \cdot s'$ contradicting [eq: 9.21]. Hence we must have that

$$t^{-1} \cdot s' \in \alpha \tag{9.22}$$

As by [eq: 9.19] $t' \in \beta$ we have using the above that $(t^{-1} \cdot s') \cdot t' \in \{s \cdot t | (s, t) \in \alpha \times \beta \wedge 0 < s \wedge 0 < t\}$ so that

$$(t^{-1} \cdot s') \cdot t' \in \alpha \odot \beta \tag{9.23}$$

Now

$$\begin{aligned}
(t^{-1} \cdot s') \cdot t' &= t^{-1} \cdot (s' \cdot t') \\
&\underset{[\text{eq: } 9.19]}{=} t^{-1} \cdot q \\
&\underset{[\text{eq: } 9.20]}{=} t^{-1} \cdot (t \cdot r) \\
&= r
\end{aligned}$$

which combined with [eq: 9.23] proves that $r \in \alpha \odot \beta$ contradicting the fact $r \in \mathbb{Q} \setminus \alpha \odot \beta$, hence the assumption is wrong and we must have

$$q < r$$

4. Assume now that $\alpha \odot \beta$ has a greatest element m then we have

$$m \in \alpha \odot \beta \text{ and } \forall r \in \alpha \odot \beta \text{ we have } r \leqslant m \tag{9.24}$$

As $m \in \alpha \odot \beta$ we have the following cases to consider:

$m \in \{r \in \mathbb{Q} | r \leqslant 0\}$. Then $m \leqslant 0$. As $\alpha, \beta \in \mathbb{R}^+$ we have that $0 \in \alpha$ and $0 \in \beta$ which, as α, β have no greatest element, there exists $s \in \alpha$ and $t \in \beta$ such that $0 < s$ and $0 < t$, hence $s \cdot t \in \{s \cdot t | (s, t) \in \alpha \times \beta \wedge 0 < s \wedge 0 < t\}$ proving that

$$s \cdot t \in \alpha \odot \beta \text{ and thus } s \cdot t \leqslant m$$

As $0 < s \wedge 0 < t$ we have that $0 < s \cdot t$ so, as $m \leqslant 0$, we have $m < s \cdot t$ contradicting the above.

$m \notin \{r \in \mathbb{Q} | r \leqslant 0\}$. Then $0 < m$ and $m \in \{s \cdot t | (s, t) \in \alpha \times \beta \wedge 0 < s \wedge 0 < t\}$ hence there exists $s \in \alpha$ and $t \in \beta$ with $0 < s$ and $0 < t$ such that

$$m = s \cdot t \tag{9.25}$$

As α, β has no greatest element there exists $s' \in \alpha$ and $t' \in \beta$ such that $0 < s < s'$ and $0 < t < t'$. As $0 < s \wedge 0 < t$ we have $s \cdot t < s' \cdot t$ and $t \cdot s' < s' \cdot t'$ so that $s \cdot t < s' \cdot t'$ or using [eq: 9.25]

$$m < s' \cdot t' \tag{9.26}$$

Further as $s' \cdot t' \in \{s \cdot t | (s, t) \in \alpha \times \beta \wedge 0 < s \wedge 0 < t\}$ we have that $s' \cdot t' \in \alpha \odot \beta$ so that by [eq: 9.24] $s' \cdot t' \leqslant m$ contradicting [eq: 9.26].

As in all cases we have a contradiction the assumption must be wrong, so $\alpha \odot \beta$ has no greatest element.

By (1),(2),(3) and (4) we have that $\alpha \odot \beta$ is a Dedekind cut, hence

$$\alpha \odot \beta \in \mathbb{R}$$

Finally as $0 \in \{r \in \mathbb{Q} | r \leqslant 0\}$ we have $0 \in \alpha \odot \beta$ proving that

$$\alpha \odot \beta \in \mathbb{R}^+$$

□

After we have defined multiplication in $\mathbb{R}^+$ we want to specify the neutral element for $\odot$

Theorem 9.21. *$\forall \alpha \in \mathbb{R}^+$ we have $\alpha_1 \odot \alpha = \alpha$*

Proof. Let $x \in \alpha_1 \odot \alpha$ then we have either:

- $\boldsymbol{x \leqslant 0}$**.** As $\alpha \in \mathbb{R}^+$ we have that $0 \in \alpha$ and as $x \leqslant 0$ it follows from [theorem: 9.3] that $x \in \alpha$.
- $\boldsymbol{0 < x}$**.** Then $x \notin \mathbb{Q}_0^-$ so there exists a $s \in \alpha_1$ and a $t \in \alpha$ with $0 < s \wedge 0 < t$ such that $x = s \cdot t$. From $s \in \alpha_1$ it follows that $s < 1$ so, as $0 < t$ we have that $x = s \cdot t < t \Rightarrow x < t$. As $t \in \alpha$ it follows from [theorem: 9.3] that $x \in \alpha$.

As in all cases $x \in \alpha$ it follows that

$$\alpha_1 \odot \alpha \subseteq \alpha \tag{9.27}$$

If $x \in \alpha$ then we have either:

- $\boldsymbol{x \leqslant 0}$**.** Then $x \in \{q \in \mathbb{Q} | q \leqslant 0\}$ so that $x \in \alpha_1 \odot \alpha$
- $\boldsymbol{0 < x}$**.** As α has no greatest element [see definition: 9.1 (4)] there exist a $t \in \alpha$ such that $0 < x < t$. Then as by [theorems: 8.29, 4.73] $0 < t^{-1}$ we have that $0 < x \cdot t^{-1} < t \cdot t^{-1} = 1$ so that $x \cdot t^{-1} \in \alpha_1$. Now $0 < t$, $0 < x \cdot t^{-1}$ so that $x = (x \cdot t^{-1}) \cdot t \in a_1 \odot \alpha$. Hence $\alpha \subseteq \alpha_1 \odot \alpha$ which combined with [eq: 9.27] results in

$$\alpha_1 \circ \alpha = \alpha$$

□

Theorem 9.22. *$\forall \alpha, \beta \in \mathbb{R}^+$ we have $\alpha \odot \beta = \beta \odot \alpha$*

Proof. Then we have

$$\begin{array}{rcl} q \in \alpha \odot \beta & \Leftrightarrow & q \in \{r \in \mathbb{Q} | r \leqslant 0\} \bigcup \{s \cdot t | (s,t) \in \alpha \times \beta \wedge 0 < s \wedge 0 < t\} \\ & \Leftrightarrow & q \leqslant 0 \vee \exists (s,t) \in \alpha \times \beta \text{ with } 0 < s \wedge 0 < t \text{ such that } q = s \cdot t \\ & \underset{\text{[theorem: 8.12}}{\Leftrightarrow} & q \leqslant 0 \vee \exists (s,t) \in \alpha \times \beta \text{ with } 0 < s \wedge 0 < t \text{ such that } q = t \cdot s \\ & \Leftrightarrow & q \leqslant 0 \vee \exists (t,s) \in \beta \times \alpha \text{ with } 0 < t \wedge 0 < s \text{ such that } q = s \cdot t \\ & \Leftrightarrow & q \in \beta \odot \alpha \end{array}$$

proving that

$$\alpha \circ \beta = \beta \odot \alpha$$

□

Theorem 9.23. *Let $\alpha, \beta, \gamma \in \mathbb{R}^+$ then we have that $\alpha \odot (\beta \odot \gamma) = (\alpha \odot \beta) \odot \gamma$*

Proof. Using the definition of $\odot$ we have

$$\alpha \odot \beta = \mathbb{Q}_0^- \bigcup \{s \cdot t | (s,t) \in \alpha \times \beta \text{ with } 0 < s \wedge 0 < t\} \tag{9.28}$$

$$\beta \odot \gamma = \mathbb{Q}_0^- \bigcup \{s \cdot t | (s,t) \in \beta \times \gamma \text{ with } 0 < s \wedge 0 < t\} \tag{9.29}$$

$$\alpha \odot (\beta \odot \gamma) = \mathbb{Q}_0^- \bigcup \{s \cdot t | (s,t) \in \alpha \times (\beta \odot \gamma) \text{ with } 0 < s \wedge 0 < t\} \tag{9.30}$$

$$(\alpha \odot \beta) \odot \gamma = \mathbb{Q}_0^- \bigcup \{s \cdot t | (s,t) \in (\alpha \odot \beta) \times \gamma \text{ with } 0 < s \wedge 0 < t\} \tag{9.31}$$

Let $x \in \alpha \odot (\beta \odot \gamma)$ then we have either

- $\boldsymbol{x \leqslant 0}$**.** Then we have $x \in \mathbb{Q}_0^-$ proving by [eq: 9.31] that $x \in (\alpha \odot \beta) \odot \gamma$
- $\boldsymbol{0 < x}$**.** Then by [eq: 9.30] there exist $q \in \alpha$ and $r \in \beta \odot \gamma$ with $0 < q \wedge 0 < r$ such that

$$x = q \cdot r \tag{9.32}$$

  As $0 < r$ and $r \in \beta \odot \gamma$ it follows from [eq: 9.29] that there exists a $s \in \beta$ and $t \in \gamma$ with $0 < s$ and $0 < t$ such that $r = s \cdot t$ hence $x = q \cdot r = q \cdot (s \cdot t) = (q \cdot s) \cdot t$ proving that

$$x = (q \cdot s) \cdot t \tag{9.33}$$

  As $q \in \alpha \wedge s \in \beta \wedge 0 < \alpha \wedge 0 < \beta$ we have by [eq: 9.28] that $q \cdot s \in \alpha \odot \beta$. Further as $0 < q \wedge 0 < s$ we have $0 < q \cdot s$ which together with $t \in \gamma \wedge 0 < \gamma$ proves that $(q \cdot s) \cdot t \in (\alpha \odot \beta) \odot \gamma$ or using [eq: 9.33] that

$$x \in (\alpha \odot \beta) \odot \gamma$$

So we have proved that

$$\alpha \odot (\beta \odot \gamma) \subseteq (\alpha \odot \beta) \odot \gamma \tag{9.34}$$

Let $x \in (\alpha \odot \beta) \odot \gamma$ then we have either:

$x \leqslant 0$. Then we have $x \in \mathbb{Q}_0^-$ proving by [eq: 9.31] that $x \in \alpha \odot (\beta \odot \gamma)$.

$0 < x$. Then by [eq: 9.31] we have that there exists a $q \in \alpha \odot \beta$ and a $r \in \gamma$ with $0 < q \wedge 0 < r$ such that

$$x = q \cdot r \tag{9.35}$$

As $0 < q$ and $q \in \alpha \odot \beta$ it follows from [eq: 9.28] that there exists a $s \in \alpha$ and $t \in \beta$ with $0 < s \wedge 0 < t$ such that $q = s \cdot t$. Hence $x = q \cdot r = (s \cdot t) \cdot r = s \cdot (t \cdot r)$ giving

$$x = s \cdot (t \cdot r) \tag{9.36}$$

As $t \in \beta \wedge r \in \gamma \wedge 0 < \beta \wedge 0 < \gamma$ we have by [eq: 9.29] that $t \cdot r \in \beta \odot \gamma$. Further as $0 < t \wedge 0 < r$ we have $0 < t \cdot r$ which together with $s \in \alpha \wedge 0 < s$ proves that $s \cdot (t \cdot r) \in \alpha \odot (\beta \odot \gamma)$ or using [eq: 9.36] we have that

$$x \in \alpha \odot (\beta \odot \gamma)$$

So we have proved that $(\alpha \odot \beta) \odot \gamma \subseteq \alpha \odot (\beta \odot \gamma)$ which combined with [eq: 9.34] gives

$$(\alpha \odot \beta) \odot \gamma - \alpha \odot (\beta \odot \gamma)$$

□

Theorem 9.24. *$\forall \alpha, \beta, \gamma \in \mathbb{R}^+$ we have that $\alpha \odot (\beta + \gamma) = \alpha \odot \beta + \alpha \odot \gamma$*

Proof. Let $x \in \alpha \cdot (\beta + \gamma)$ then we have either:

$x \leqslant 0$. Then $0, x \in \mathbb{Q}_0^-$ so that $x \in \alpha \odot \beta$ and $0 \in \alpha \odot \gamma$ hence $x = x + 0 \in \alpha \odot \beta + \alpha \odot \gamma$.

$0 < x$. Then $x = s \cdot t$ where $s \in \alpha \wedge 0 < s$ and $t \in \beta + \gamma \wedge 0 < t$. As $t \in \beta + \gamma$ there exists $u \in \beta$ and $v \in \gamma$ such that $t = u + v$. Using [theorem: 8.12] we have that

$$x = s \cdot t = s \cdot (u + v) = s \cdot u + s \cdot v \tag{9.37}$$

We have now the following possibilities for u and v:

$u \leqslant 0 \wedge v \leqslant 0$. Then $t = u + v \leqslant 0$ giving the contradiction $0 < t \leqslant 0$ so this case will not occur.

$u \leqslant 0 \wedge 0 < v$. Then as $0 < s$ we have $s \cdot u \leqslant 0 \Rightarrow s \cdot u \in \mathbb{Q}_0^- \Rightarrow s \cdot u \in \alpha \odot \beta$, further as $0 < s \wedge 0 < v$ we have that $s \cdot u \in \alpha \odot \gamma$. Hence $x \underset{[\text{eq: } 9.37]}{=} s \cdot u + s \cdot v \in \alpha \odot \beta + \alpha \odot \gamma$.

$0 < u \wedge v \leqslant 0$. Then as $0 < s \wedge 0 < u$ we have that $s \cdot u \in \alpha \odot \beta$, further $s \cdot v \leqslant 0 \Rightarrow s \cdot v \in \mathbb{Q}_0^- \Rightarrow s \cdot v \in \alpha \odot \gamma$. Hence $x \underset{[\text{eq: } 9.37]}{=} s \cdot u + s \cdot v \in \alpha \odot \beta + s \odot \gamma$

$0 < u \wedge 0 < v$. Then as $0 < s \wedge 0 < u \wedge 0 < v$ we have that $s \cdot u \in \alpha \circ \beta$ and $s \cdot v \in \alpha \odot \gamma$. Hence $x \underset{[\text{eq: } 9.37]}{=} s \cdot u + s \cdot v \in \alpha \odot \beta + s \odot \gamma$.

So in call cases we have that $x \in \alpha \odot \beta + \alpha \odot \gamma$ proving that

$$\alpha \odot (\beta + \gamma) \subseteq \alpha \odot \beta + \alpha \odot \gamma \tag{9.38}$$

For the opposite inclusion let $x \in \alpha \odot \beta + \alpha \odot \gamma$. Then

$$x = r + t \text{ where } r \in \alpha \odot \beta \text{ and } t \in \alpha \odot \gamma \tag{9.39}$$

We must now consider the following cases for x:

$x \leqslant 0$. Then $x \in \mathbb{Q}_0^-$ so that $x \in \alpha \odot (\beta + \gamma)$

$0 < x$. Then we have to look at the following sub cases:

$r \leqslant 0 \wedge t \leqslant 0$. Then $x \underset{[\text{eq: } 9.39]}{=} r + t \leqslant 0$ a contradicting $0 < x$, so this case does not occur.

$r \leqslant 0 \wedge 0 < t$. Then as $t \notin \mathbb{Q}_0^+$ there exists $u \in \alpha$ and $v \in \gamma$ with $0 < u \wedge 0 < v$ such that $t = u \cdot v$. As $\beta \in \mathbb{R}^+ \Rightarrow 0 \in \beta$ and $v \in \gamma$ we have that $(0 + v) \in \beta + \gamma$, further as $0 < u$ and $0 < 0 + v$ it follows that $t = u \cdot v = u \cdot (0 + v) \in \alpha \odot (\beta + \gamma)$. Since $r \leqslant 0$ we have $x \underset{[\text{eq: } 9.39]}{=} r + t \leqslant 0 + t = t$, which, as $t \in \alpha \odot (\alpha + \beta)$, proves by [theorem: 9.3] that

$$x \in \alpha \odot (\beta + \gamma)$$

$0<r\wedge t\leqslant 0$. Then as $r\notin\mathbb{Q}_0^-$ there exists $u\in\alpha$ and $v\in\beta$ with $0<u\wedge 0<v$ such that $r=u\cdot v$. As $\gamma\in\mathbb{R}^+\Rightarrow 0\in\gamma$ and $v\in\beta$ we have that $v+0\in\beta+\gamma$, further as $0<u$ and $0<0+v$ it follows that $r=u\cdot v=u\cdot(v+0)\in\alpha\odot(\beta+\gamma)$. Since $t\leqslant 0$ we have $x\underset{[\text{eq: }9.39]}{=}r+t\leqslant r+0=r$, which, as $r\in\alpha\odot(\beta+\gamma)$, proves by [theorem: 9.3] that

$$x\in\alpha\odot(\beta+\gamma)$$

$0<r\wedge 0<t$. Then as $r,t\notin\mathbb{Q}_0^-$ there exists $u,u'\in\alpha$, $v\in\beta$ and $v'\in\gamma$ such that

$$r=u\cdot v\wedge t=u'\cdot v'\wedge 0<u\wedge 0<v\wedge 0<u'\wedge 0<v' \tag{9.40}$$

For u,u' we must now examine the following possibilities:

$u=u'$. Then

$$x\underset{[\text{eqs: }9.39,\ 9.40]}{=}u\cdot v+u'\cdot v'=u\cdot v+u\cdot v'=u\cdot(v+v') \tag{9.41}$$

so as $0<u\wedge 0<v+v'$ we have that $u\cdot(v+v')\in\alpha\odot(\beta+\gamma)$ hence

$$x\in\alpha\odot(\beta+\gamma)$$

$u<u'$. Then as $0<u'\wedge 0<v+v'\wedge u'\in\alpha\wedge v+v'\in\beta+\gamma$ we have

$$u'\cdot(v+v')\in\alpha\odot(\beta+\gamma)$$

Further from $u<u'$, $0<v$ we have that $u\cdot v<u'\cdot v$, hence

$$x\underset{[\text{eq: }9.39,9.40]}{=}u\cdot v+u'\cdot v'<u'\cdot v+u'\cdot v'=u'\cdot(v+v')\in\alpha\odot(\beta+\gamma)$$

which by [theorem: 9.3] proves that

$$x\in\alpha\odot(\beta+\gamma)$$

$u'<u$. Then as $0<u\wedge 0<v+v'\wedge u\in a\wedge v+v'\in\beta+\gamma$ we have

$$u\cdot(v+v')\in\alpha\odot(\beta+\gamma)$$

Further from $u'<u$, $0<v'$ it follows that $u'\cdot v'<u\cdot v'$, hence

$$x\underset{[\text{eq: }9.39,9.40]}{=}u\cdot v+u'\cdot v'<u\cdot v+u\cdot v'=u\cdot(v+v')\in\alpha\odot(\beta+\gamma)$$

which by [theorem: 9.3] proves that

$$x\in\alpha\odot(\beta+\gamma)$$

So in all cases we have $x\in\alpha\odot(\beta+\gamma)$ proving that $\alpha\odot\beta+\alpha\odot\gamma\subseteq\alpha\odot(\beta+g)$ which combined with [eq: 9.38] gives

$$\alpha\odot(\beta+\gamma)=\alpha\odot\beta+\alpha\odot\gamma$$

□

Theorem 9.25. *Let $\alpha\in\mathbb{R}^+$ then $\operatorname{inv}(\alpha)$ defined by*

$$\begin{aligned}\operatorname{inv}(\alpha) &= \{r\in\mathbb{Q}|r\leqslant 0\}\bigcup\{s^{-1}|s\in\mathbb{Q}\setminus\alpha\textit{ with }0<s\textit{ and }\exists t\in\mathbb{Q}\setminus\alpha\vdash t<s\}\\ &\underset{\mathbb{Q}_0^-=\{r\in\mathbb{Q}|r\leqslant 0\}}{=} \mathbb{Q}_0^-\bigcup\{s^{-1}|s\in\mathbb{Q}\setminus\alpha\textit{ with }0<s\textit{ and }\exists t\in\mathbb{Q}\setminus\alpha\vdash t<s\}\end{aligned}$$

is a Dedekind cut such that $\operatorname{inv}(\alpha)\in\mathbb{R}^+$.

Proof. We have

1. As $0\in\mathbb{Q}_0^-$ it follow that $0\in\operatorname{inv}(\alpha)$ proving that

$$\operatorname{inv}(\alpha)\neq\varnothing$$

2. As $\alpha\in\mathbb{R}^+$ we have $0\in\alpha$ and as α has no greatest element there exist a $s\in\alpha$ such that $0<s$. Hence s^{-1} exist and by [theorems: 8.29, 4.73] $0<s^{-1}$ so that $s^{-1}\notin\mathbb{Q}_0^-$. Assume that $s^{-1}\in\operatorname{inv}(\alpha)$ then as $s^{-1}\notin\mathbb{Q}_0^-$ we must have that

$$s^{-1}\in\{s^{-1}|s\in\mathbb{Q}\setminus\alpha\text{ with }0<s\text{ and }\exists t\in\mathbb{Q}\setminus\alpha\vDash t<s\}$$

so that $\exists t \in \mathbb{Q} \setminus \alpha$ such that $s^{-1} = t^{-1} \underset{\text{[theorem: 4.54]]}}{\Rightarrow} s = t$. So $s \in \mathbb{Q} \setminus \alpha$ contradicting $s \in \alpha$ hence $s^{-1} \notin \text{inv}(\alpha)$ proving that

$$\text{inv}(\alpha) \neq \mathbb{Q}$$

3. Let $q \in \text{inv}(\alpha)$ and $r \in \mathbb{Q} \setminus \text{inv}(\alpha)$. For q we have the following possibilities:

$q \leqslant 0$. Then as $r \in \mathbb{Q} \setminus \text{inv}(\alpha)$ we have $r \notin \text{inv}(\alpha)$ hence $r \notin \mathbb{Q}_0^-$ so that $0 < r$ giving

$$q < r$$

$0 < q$. Then $q \notin \mathbb{Q}_0^-$ hence, as $q \in \text{inv}(\alpha)$, we have:

$$q \in \{s^{-1} | s \in \mathbb{Q} \setminus \alpha \text{ with } 0 < s \text{ and } \exists t \in \mathbb{Q} \setminus \alpha \vdash t < s\}$$

so there exist a $s \in \mathbb{Q} \setminus \alpha$ with $0 < s$ and $\exists t \in \mathbb{Q} \setminus \alpha \vdash t < s$ such that $q = s^{-1}$, as $q^{-1} = (s^{-1})^{-1} \underset{\text{[theorem: 4.53]}}{=} s$ we have that

$$q^{-1} \in \mathbb{Q} \setminus \alpha,\ 0 < q^{-1} \text{ and } \exists t \in \mathbb{Q} \setminus \alpha \vdash t < q^{-1} \tag{9.42}$$

Further as $r \in \mathbb{Q} \setminus \text{inv}(\alpha)$ we have that $r \notin \mathbb{Q}_0^-$ giving $0 < r$ so that by [theorems: 8.29, 4.73]

$$0 < r \text{ and } 0 < r^{-1} \tag{9.43}$$

For r^{-1} we have the following possibilities:

$r^{-1} \in \alpha$. Then as $q^{-1} \in \mathbb{Q} \setminus \alpha$ [see eq: 9.42] we have by [definition: 9.1 (3)] that $r^{-1} < q^{-1}$, so as $0 < r^{-1}$ we have by [theorems: 8.29, 4.73] that

$$q < r$$

$r^{-1} \notin \alpha$. Then $r^{-1} \in \mathbb{Q} \setminus \alpha$ and we have to look at the following possibilities

$\forall t \in \mathbb{Q} \vDash r^{-1} \leqslant t$. Then as $q^{-1} \in \mathbb{Q} \setminus \alpha$ [see eq: 9.42] we have that $r^{-1} \leqslant q^{-1}$. If $r^{-1} = q^{-1}$ we have by [eq: 9.42] a $t \in \mathbb{Q} \setminus \alpha$ such that $t < r^{-1}$ contradicting $\forall t \in \mathbb{Q} \vDash r^{-1} \leqslant t$, hence $r^{-1} \neq q^{-1}$. So $0 < r^{-1} < q^{-1}$ and by [theorems: 8.29, 4.73]

$$q < r$$

$\exists t \in \mathbb{Q}$ such that $t < r^{-1}$. Then as $r^{-1} \in \mathbb{Q} \setminus \alpha$ and $0 < r^{-1}$ we have that $r = (r^{-1})^{-1} \in \{s^{-1} | s \in \mathbb{Q} \setminus \alpha \text{ with } 0 < s \text{ and } \exists t \in \mathbb{Q} \setminus \alpha \vdash t < s\}$ so that $r \in \text{inv}(\alpha)$ contradicting $r \in \mathbb{Q} \setminus \text{inv}(\alpha)$ so this case does not occur.

Hence is all valid cases we have

$$q < r$$

4. Assume that $\text{inv}(\alpha)$ has a greatest element m then we have

$$m \in \text{inv}(\alpha) \text{ and } \forall s \in \text{inv}(\alpha) \text{ we have } s \leqslant m \tag{9.44}$$

For m we have to look at the following possibilities:

$m \leqslant 0$. Using [definition: 9.1 (2)] $\varnothing \neq \mathbb{Q} \setminus \alpha$ so there exist a $r \in \mathbb{Q} \setminus \alpha$. As $\alpha \in \mathbb{R}^+$ we have that $0 \in \alpha$ so that by [definition: 9.1 (3)] that

$$0 < r \underset{r < r+1}{\Rightarrow} 0 < r + 1 \text{ and by [theorems: 8.29, 4.73] } 0 < (r+1)^{-1} \tag{9.45}$$

If $r + 1 \in \alpha$ then as $r \in \mathbb{Q} \setminus \alpha$ we have by [definition: 9.1 (3)] that $r + 1 < r$ a contradiction, so we must have that $r + 1 \notin \alpha$ or $r + 1 \in \mathbb{Q} \setminus \alpha$. As further $r < r + 1$ and $0 < r + 1$ it follows that $(r+1)^{-1} \in \{s^{-1} | s \in \mathbb{Q} \setminus \alpha \text{ with } 0 < s \text{ and } \exists t \in \mathbb{Q} \setminus \alpha \vdash t < s\}$ so that $(r+1)^{-1} \in \text{inv}(\alpha)$ hence by [eq: 9.44] $(r+1)^{-1} \leqslant m \leqslant 0$ contradicting $0 < (r+1)^{-1}$ [see eq: 9.45]. So we end with a contradiction.

$0 < m$. Then $m \notin \mathbb{Q}_0^-$ so that, as $m \in \text{inv}(\alpha)$, we have $m \in \{s^{-1} | s \in \mathbb{Q} \setminus \alpha \text{ with } 0 < s \text{ and } \exists t \in \mathbb{Q} \setminus \alpha \vdash t < s\}$ so there exist a $s \in \mathbb{Q} \setminus \alpha$ with $0 < s$ and $\exists t \in \mathbb{Q} \setminus \alpha \vdash t < s$ such that $m = s^{-1}$, hence $m^{-1} = s$ so that:

$$m^{-1} \in \mathbb{Q} \setminus \alpha,\ 0 < m^{-1} \text{ and } \exists t \in \mathbb{Q} \setminus \alpha \text{ such that } t < m^{-1} \tag{9.46}$$

As $t \in \mathbb{Q} \setminus \alpha$ we have that $t \notin \mathbb{Q}_0^-$ so that $0 < t$, further as $t < m^{-1}$ we have by the density of $\mathbb{Q}$ [see theorem: 8.37] that there exist a $s \in \mathbb{Q}$ such that $t < s < m^{-1}$, hence

$$0 < t < s < m^{-1} \text{ which by [theorems: 8.29, 4.73] gives also } m < s^{-1} \tag{9.47}$$

If $s \in \alpha$ then as $t \in \mathbb{Q} \backslash \alpha$ [see eq: 9.46] we have by [definition: 9.1 (3)] that $s < t$ contradicting $t < s < m^{-1}$, so we must have $s \notin \alpha$, hence $s \in \mathbb{Q} \backslash \alpha$, so as $s \notin a$ we have $s \notin \mathbb{Q}_0^-$, so that $0 < s$, which together with $t \in \mathbb{Q} \backslash \alpha$ and $t < s$ proves that $s^{-1} \in \operatorname{inv}(\alpha)$. Using [eq: 9.44] it follows that $s^{-1} \leqslant m$ which contradicts [eq: 9.47]. So this case ends also in a contradiction.

As all possible cases ends in a contradiction the assumption must be false resulting in

$$\operatorname{inv}(\alpha) \text{ has no greatest element}$$

(1),(2),(3),(4) proves that

$$\operatorname{inv}(\alpha) \text{ is a Dedekind cut}$$

Further as $0 \in \mathbb{Q}_0^-$ we have that $0 \in \operatorname{inv}(\alpha)$ hence

$$\operatorname{inv}(\alpha) \in \mathbb{R}^+$$ □

We prove now that $\operatorname{inv}(\alpha)$ is the multiplicative inverse for $\mathbb{R}^+$.

Theorem 9.26. *If $\alpha \in \mathbb{R}^+$ then $\alpha \odot \operatorname{inv}(\alpha) = \alpha_1$*

Proof. If $x \in \alpha \odot \operatorname{inv}(\alpha)$ then we have for x either:

$\boldsymbol{x \leqslant 0}$. Then as $0 < 1$ we have $x < 1$ hence $x \in \alpha_1$

$\boldsymbol{0 < x}$. Then $x \notin \mathbb{Q}_0^-$ we have as $x \in \alpha \odot \operatorname{inv}(\alpha)$ that $\exists s \in \alpha \wedge \exists t \in \operatorname{inv}(\alpha)$ with $0 < s$ and $0 < t$ such that $x = s \cdot t$. For t we have the following cases:

$\boldsymbol{t \leqslant 0}$. Then from $0 < s$ we have that $x = s \cdot t \leqslant 0 < 1$ hence $x \in \alpha_1$.

$\boldsymbol{0 < t}$. Then $t \notin \mathbb{Q}_0^-$ which as $t \in \operatorname{inv}(\alpha)$ means that there exist a $s \in \mathbb{Q} \backslash \alpha$ such that $0 < s$ and $\exists r \in \mathbb{Q} \backslash \alpha \vDash r < s$ such that $t = s^{-1}$. As $t^{-1} = (s^{-1})^{-1} = s$ we have that $t^{-1} \in \mathbb{Q} \backslash \alpha$. Using [definition: 9.1 (3)] we have $s < t^{-1}$ so that $x = s \cdot t < 1$ proving that $x \in \alpha_1$.

As in all cases $x \in \alpha_1$ we have that

$$\alpha \odot \operatorname{inv}(\alpha) \subseteq \alpha_1 \tag{9.48}$$

Now for the opposite inclusion, let $x \in \alpha_1$ then $x < 1$ and we have either:

$\boldsymbol{x \leqslant 0}$. Then $x \in \mathbb{Q}_0^-$ so that $x \in \alpha \odot \operatorname{inv}(\alpha)$.

$\boldsymbol{0 < x}$. As $\alpha \in \mathbb{R}^+$ we have that $0 \in \alpha$, further as α is a Dedekind cut, α has no greatest element [see definition: 9.1 (4)] so

$$\exists s_1 \in \alpha \text{ such that } 0 < s_1 \tag{9.49}$$

As $0 < x < 1$ we have $0 < 1 - x$, and by [theorems: 8.29, 4.73] $0 < x^{-1}$ so that by applying [theorems: 8.29, 4.73] repeately we have that $0 < s_1 \cdot (1 - x) \cdot x^{-1}$ or

$$\text{If } \varepsilon = s_1 \cdot (1 - x) \cdot x^{-1} \text{ then } 0 < \varepsilon \tag{9.50}$$

We have by [theorem: 9.13] that there exist a $s_2 \in \alpha$ such that $s_2 + \varepsilon \in \mathbb{Q} \backslash \alpha$. As α has no maximal element and $s_2 \in \alpha$ there exist a $s_3 \in \alpha$ such that $s_2 < s_3$ then $s_2 + \varepsilon < s_3 + \varepsilon$. If $s_3 + \varepsilon \in \alpha$ then by [definition: 9.1 (3)] we have as $s_2 + \varepsilon \in \mathbb{Q} \backslash \alpha$ that $s_3 + \varepsilon < s_2 + s_3$ contradicting $s_2 + \varepsilon < s_3 + \varepsilon$ hence we must have that

$$s_3 + \varepsilon \in \mathbb{Q} \backslash \alpha \text{ and } s_2 + \varepsilon < s_3 + \varepsilon \wedge s_2 + \varepsilon \in \mathbb{Q} \backslash \alpha \tag{9.51}$$

For s_1,s_2 we have either:

$\boldsymbol{s_3 < s_1}$. Then $s_3 + \varepsilon < s_1 + \varepsilon$. If $s_1 + \varepsilon \in \alpha$ then as $s_3 + \varepsilon \in \mathbb{Q} \backslash \alpha$ we have by [definition: 9.1 (3)] that $s_1 + \varepsilon < s_3 + \varepsilon$ contradicting $s_3 + \varepsilon < s_1 + \varepsilon$ so we must have that

$$s_1 + \varepsilon \in \mathbb{Q} \backslash \alpha \tag{9.52}$$

As $0 \in \alpha$ and $x_1 + \varepsilon \in \mathbb{Q} \backslash \alpha$ we have by [definition: 9.1 (3)] that

$$0 < s_1 + \varepsilon \text{ and by [theorems: 8.29, 4.73] } 0 < (s_1 + \varepsilon)^{-1} \tag{9.53}$$

By [eq: 9.52], [eq: 9.53] and the fact that $s_3 + \varepsilon < s_1 + \varepsilon$, $s_3 + \varepsilon \in \mathbb{Q} \backslash \alpha$ we have by the definition of $\operatorname{inv}(\alpha)$ we have that

$$(s_1 + \varepsilon)^{-1} \in \operatorname{inv}(\alpha)$$

As $0 < s_1 \in \alpha$, $0 < (s_1 + \varepsilon)^{-1} \in \operatorname{inv}(\alpha)$ [see eqs: 9.49, 9.52, 9.53] it follows from the definition of $\odot$ that

$$s_1 \cdot (s_1 + \varepsilon)^{-1} \in \alpha \odot \operatorname{inv}(\alpha) \tag{9.54}$$

Now

$$\begin{array}{rcl} s_1\cdot(s_1+\varepsilon)^{-1} & \underset{[\text{eq: }9.50]}{=} & s_1\cdot(s_1+s_1\cdot(1-x)\cdot x^{-1})^{-1}\\ & = & s_1\cdot(s_1+s_1\cdot x^{-1}-s_1\cdot x^{-1}\cdot x)^{-1}\\ & = & s_1\cdot(s_1+s_1\cdot x^{-1}-s_1)^{-1}\\ & = & s_1\cdot(s_1\cdot x^{-1})^{-1}\\ & \underset{[\text{theorem: }4.55]}{=} & s_1\cdot(s_1^{-1}\cdot(x^{-1})^{-1})\\ & = & (x^{-1})^{-1}\\ & = & x \end{array}$$

proving using [eq: 9.54[that

$$x\in\alpha\odot\mathrm{inv}(\alpha)$$

***$s_1\leqslant s_3$*.** Then as $0\in\alpha$ and $s_3+\varepsilon\in\mathbb{Q}\setminus\alpha$ [see eq: 9.51] it follows from [definition: 9.1] that

$$0<s_3+\varepsilon \text{ and by [theorems: 8.29, 4.73] } 0<(s_3+\varepsilon)^{-1} \tag{9.55}$$

So using the definition of $\mathrm{inv}(\alpha)$ together with [eqs: 9.51, 9.55] that

$$(s_3+\varepsilon)^{-1}\in\mathrm{inv}(\alpha) \tag{9.56}$$

Now by [eq: 9.49] $0<s_1\leqslant s_3\in\alpha$ and $0<(s_3+\varepsilon)^{-1}\in\mathrm{inv}(\alpha)$ [see eq: 9.55, 9.51] so that

$$s_3\cdot(s_3+\varepsilon)^{-1}\in\alpha\odot\mathrm{inv}(\alpha) \tag{9.57}$$

Now

$$\begin{array}{rcl} s_1\leqslant s_3 & \underset{0<1-x}{\Rightarrow} & s_1\cdot(1-x)\leqslant s_3\cdot(1-x)\\ & \underset{0<x^{-1}}{\Rightarrow} & s_1\cdot(1-x)\cdot x^{-1}\leqslant s_3\cdot(1-x)\cdot x^{-1}\\ & \Rightarrow & s_3+s_1\cdot(1-x)\cdot x^{-1}\leqslant s_3+s_3\cdot(1-x)\cdot x^{-1}\\ & \underset{[\text{theorems: }8.29,\,4.73]}{\Rightarrow} & (s_3+s_1\cdot(1-x)\cdot x^{-1})^{-1}\geqslant(s_3+s_3\cdot(1-x)\cdot x^{-1})^{-1}\\ & \underset{[\text{eq: }9.50]}{\Rightarrow} & (s_3+\varepsilon)^{-1}\geqslant(s_3+s_3\cdot(1-x)\cdot x^{-1})^{-1}\\ & \underset{0<s_1\leqslant s_3}{\Rightarrow} & s_3\cdot(s_3+\varepsilon)^{-1}\geqslant s_3\cdot(s_3+s_3\cdot(1-x)\cdot x^{-1})^{-1} \end{array} \tag{9.58}$$

Further

$$\begin{array}{rcl} s_3\cdot(s_3+\varepsilon)^{-1} & \geqslant_{[\text{eq: }9.58]} & s_3\cdot(s_3+s_3\cdot(1-x)\cdot x^{-1})^{-1}\\ & = & s_3\cdot(s_3+s_3\cdot x^{-1}-s_3\cdot x\cdot x^{-1})^{-1}\\ & = & s_3\cdot(s_3+s_3\cdot x^{-1}-s_3)^{-1}\\ & = & s_3\cdot(s_3\cdot x^{-1})^{-1}\\ & \underset{[\text{theorem: }4.55]}{=} & s_3\cdot s_3^{-1}\cdot(x^{-1})^{-1}\\ & = & x \end{array}$$

which by [theorem: 9.3] and [eq: 9.57]proves that

$$x\in\alpha\odot\mathrm{inv}(\alpha)$$

As in all cases $x\in\alpha\odot\mathrm{in}(\alpha)$ it follows that $\alpha_1\subseteq\alpha\odot\mathrm{inv}(\alpha)$ which combined with [eq: 9.48] proves finally that

$$\alpha_1=\alpha\odot\mathrm{inv}(\alpha)$$

□

We prove now that $\mathbb{R}\times\mathbb{R}$ is the disjoint union of sets of the form $A\times B$ where $A,B\in\{\mathbb{R}^+,\mathbb{R}^-,\{0\}\}$

Theorem 9.27. $\mathbb{R}\times\mathbb{R}$ *can be expressed as follows*

$$\mathbb{R}\times\mathbb{R}=(\mathbb{R}^+\times\mathbb{R}^+)\bigcup(\mathbb{R}^+\times\mathbb{R}^-)\bigcup(\mathbb{R}^-\times\mathbb{R}^+)\bigcup(\mathbb{R}^-\times\mathbb{R}^-)\bigcup\left((\mathbb{R}\times\{0\})\bigcup(\{0\}\times\mathbb{R})\right)$$

where

$$(\mathbb{R}^+\times\mathbb{R}^+)\bigcap(\mathbb{R}^+\times\mathbb{R}^-)=\varnothing$$
$$(\mathbb{R}^+\times\mathbb{R}^+)\bigcap(\mathbb{R}^-\times\mathbb{R}^+)=\varnothing$$
$$(\mathbb{R}^+\times\mathbb{R}^+)\bigcap(\mathbb{R}^-\times\mathbb{R}^-)=\varnothing$$
$$(\mathbb{R}^+\times\mathbb{R}^+)\bigcap\big((\mathbb{R}\times\{0\})\bigcup(\{0\}\times\mathbb{R})\big)=\varnothing$$
$$(\mathbb{R}^+\times\mathbb{R}^-)\bigcap(\mathbb{R}^-\times\mathbb{R}^+)=\varnothing$$
$$(\mathbb{R}^+\times\mathbb{R}^-)\bigcap(\mathbb{R}^-\times\mathbb{R}^-)=\varnothing$$
$$(\mathbb{R}^+\times\mathbb{R}^-)\bigcap\big((\mathbb{R}\times\{0\})\bigcup(\{0\}\times\mathbb{R})\big)=\varnothing$$
$$(\mathbb{R}^-\times\mathbb{R}^+)\bigcap(\mathbb{R}^-\times\mathbb{R}^-)=\varnothing$$
$$(\mathbb{R}^-\times\mathbb{R}^+)\bigcap\big((\mathbb{R}\times\{0\})\bigcup(\{0\}\times\mathbb{R})\big)=\varnothing$$
$$(\mathbb{R}^-\times\mathbb{R}^-)\bigcap\big((\mathbb{R}\times\{0\})\bigcup(\{0\}\times\mathbb{R})\big)=\varnothing$$

Proof. First note that by [theorem: 9.17]

$$\mathbb{R}=\mathbb{R}^+\bigcup\mathbb{R}^-\bigcup\{0\}=\bigcup_{A\in\{\mathbb{R}^+,\mathbb{R}^-,\{0\}\}}A \tag{9.59}$$

and

$$\mathbb{R}^+\bigcap\mathbb{R}^-=\varnothing\text{ and }\mathbb{R}^+\bigcap\mathbb{R}^+\bigcap\{0\}=\varnothing\text{ and }\mathbb{R}^-\bigcap\{0\}=\varnothing \tag{9.60}$$

First as $\mathbb{R}^+\subseteq\mathbb{R}$, $\mathbb{R}^-\subseteq\mathbb{R}$, $\{0\}\subseteq\mathbb{R}$ and $\mathbb{R}\subseteq\mathbb{R}$ we have by [theorem: 1.48] that $\mathbb{R}^+\times\mathbb{R}^+\subseteq\mathbb{R}\times\mathbb{R}$, $\mathbb{R}^+\times\mathbb{R}^-\subseteq\mathbb{R}\times\mathbb{R}$, $\mathbb{R}^-\times\mathbb{R}^+\subseteq\mathbb{R}\times\mathbb{R}$, $\mathbb{R}^-\times\mathbb{R}^-\subseteq\mathbb{R}\times\mathbb{R}$, $((\mathbb{R}\times\{0\})\bigcup(\{0\}\times\mathbb{R}))\subseteq\mathbb{R}\times\mathbb{R}$ so that

$$(\mathbb{R}^+\times\mathbb{R}^+)\bigcup(\mathbb{R}^+\times\mathbb{R}^-)\bigcup(\mathbb{R}^-\times\mathbb{R}^+)\bigcup(\mathbb{R}^-\times\mathbb{R}^-)\bigcup\big((\mathbb{R}\times\{0\})\bigcup(\{0\}\times\mathbb{R})\big)\subseteq\mathbb{R}\times\mathbb{R} \tag{9.61}$$

Let $(x,y)\in\mathbb{R}\times\mathbb{R}$ then $x\in\mathbb{R}\underset{\text{[eq: 9.59]}}{=}\mathbb{R}^+\bigcup\mathbb{R}^-\bigcup\{0\}$ and $y\in\mathbb{R}\underset{\text{[eq: 9.59]}}{=}\mathbb{R}^+\bigcup\mathbb{R}^-\bigcup\{0\}$ so that for x we have either:

$\boldsymbol{x\in\mathbb{R}^+}$**.** Then for y we have either:

$\boldsymbol{y\in\mathbb{R}^+}$**.** Then $(x,y)\in\mathbb{R}^+\times\mathbb{R}^+$ so that

$$(x,y)\in(\mathbb{R}^+\times\mathbb{R}^+)\bigcup(\mathbb{R}^+\times\mathbb{R}^-)\bigcup(\mathbb{R}^-\times\mathbb{R}^+)\bigcup(\mathbb{R}^-\times\mathbb{R}^-)\bigcup\big((\mathbb{R}\times\{0\})\bigcup(\{0\}\times\mathbb{R})\big)$$

$\boldsymbol{y\in\mathbb{R}^-}$**.** Then $(x,y)\in\mathbb{R}^+\times\mathbb{R}^-$ so that

$$(x,y)\in(\mathbb{R}^+\times\mathbb{R}^+)\bigcup(\mathbb{R}^+\times\mathbb{R}^-)\bigcup(\mathbb{R}^-\times\mathbb{R}^+)\bigcup(\mathbb{R}^-\times\mathbb{R}^-)\bigcup\big((\mathbb{R}\times\{0\})\bigcup(\{0\}\times\mathbb{R})\big)$$

$\boldsymbol{y\in\{0\}}$**.** Then $(x,y)\in\mathbb{R}\times\{0\}$ so that

$$(x,y)\in(\mathbb{R}^+\times\mathbb{R}^+)\bigcup(\mathbb{R}^+\times\mathbb{R}^-)\bigcup(\mathbb{R}^-\times\mathbb{R}^+)\bigcup(\mathbb{R}^-\times\mathbb{R}^-)\bigcup\big((\mathbb{R}\times\{0\})\bigcup(\{0\}\times\mathbb{R})\big)$$

$\boldsymbol{x\in\mathbb{R}^-}$**.** Then for y we have either:

$\boldsymbol{y\in\mathbb{R}^+}$**.** Then $(x,y)\in\mathbb{R}^-\times\mathbb{R}^+$ so that

$$(x,y)\in(\mathbb{R}^+\times\mathbb{R}^+)\bigcup(\mathbb{R}^+\times\mathbb{R}^-)\bigcup(\mathbb{R}^-\times\mathbb{R}^+)\bigcup(\mathbb{R}^-\times\mathbb{R}^-)\bigcup\big((\mathbb{R}\times\{0\})\bigcup(\{0\}\times\mathbb{R})\big)$$

$\boldsymbol{y\in\mathbb{R}^-}$**.** Then $(x,y)\in\mathbb{R}^-\times\mathbb{R}^-$ so that

$$(x,y)\in(\mathbb{R}^+\times\mathbb{R}^+)\bigcup(\mathbb{R}^+\times\mathbb{R}^-)\bigcup(\mathbb{R}^-\times\mathbb{R}^+)\bigcup(\mathbb{R}^-\times\mathbb{R}^-)\bigcup\big((\mathbb{R}\times\{0\})\bigcup(\{0\}\times\mathbb{R})\big)$$

$\boldsymbol{y\in\{0\}}$**.** Then $(x,y)\in\mathbb{R}\times\{0\}$ so that

$$(x,y)\in(\mathbb{R}^+\times\mathbb{R}^+)\bigcup(\mathbb{R}^+\times\mathbb{R}^-)\bigcup(\mathbb{R}^-\times\mathbb{R}^+)\bigcup(\mathbb{R}^-\times\mathbb{R}^-)\bigcup\big((\mathbb{R}\times\{0\})\bigcup(\{0\}\times\mathbb{R})\big)$$

$\boldsymbol{x\in\{0\}}$**.** Them $(x,y)\in\{0\}\times\mathbb{R}$ so that

$$(x,y)\in(\mathbb{R}^+\times\mathbb{R}^+)\bigcup(\mathbb{R}^+\times\mathbb{R}^-)\bigcup(\mathbb{R}^-\times\mathbb{R}^+)\bigcup(\mathbb{R}^-\times\mathbb{R}^-)\bigcup\big((\mathbb{R}\times\{0\})\bigcup(\{0\}\times\mathbb{R})\big)$$

So $(x,y)\in(\mathbb{R}^+\times\mathbb{R}^+)\bigcup(\mathbb{R}^+\times\mathbb{R}^-)\bigcup(\mathbb{R}^-\times\mathbb{R}^+)\bigcup(\mathbb{R}^-\times\mathbb{R}^-)\bigcup((\mathbb{R}\times\{0\})\bigcup(\{0\}\times\mathbb{R}))$ proving that $\mathbb{R}\times\mathbb{R}\subseteq(\mathbb{R}^+\times\mathbb{R}^+)\bigcup(\mathbb{R}^+\times\mathbb{R}^-)\bigcup(\mathbb{R}^-\times\mathbb{R}^+)\bigcup(\mathbb{R}^-\times\mathbb{R}^-)\bigcup((\mathbb{R}\times\{0\})\bigcup(\{0\}\times\mathbb{R}))$ which combined with [eq: 9.61] gives

$$\mathbb{R}\times\mathbb{R}=(\mathbb{R}^+\times\mathbb{R}^+)\bigcup(\mathbb{R}^+\times\mathbb{R}^-)\bigcup(\mathbb{R}^-\times\mathbb{R}^+)\bigcup(\mathbb{R}^-\times\mathbb{R}^-)\bigcup\big((\mathbb{R}\times\{0\})\bigcup(\{0\}\times\mathbb{R})\big) \tag{9.62}$$

Next we have by [theorem: 1.49] and [theorem: 1.47] that

$$
\begin{aligned}
(\mathbb{R}^+ \times \mathbb{R}^+) \bigcap (\mathbb{R}^+ \times \mathbb{R}^-) &= \left(\mathbb{R}^+ \bigcap \mathbb{R}^+\right) \times \left(\mathbb{R}^+ \bigcap \mathbb{R}^-\right) \\
&\underset{\text{[eq: 9.60]}}{=} \left(\mathbb{R}^+ \bigcap \mathbb{R}^+\right) \times \varnothing \\
&= \varnothing \\
(\mathbb{R}^+ \times \mathbb{R}^+) \bigcap (\mathbb{R}^- \times \mathbb{R}^+) &= \left(\mathbb{R}^+ \bigcap \mathbb{R}^-\right) \times \left(\mathbb{R}^+ \bigcap \mathbb{R}^+\right) \\
&\underset{\text{[eq: 9.60]}}{=} \varnothing \times \left(\mathbb{R}^+ \bigcap \mathbb{R}^+\right) \\
&= \varnothing \\
(\mathbb{R}^+ \times \mathbb{R}^+) \bigcap (\mathbb{R}^- \times \mathbb{R}^-) &= \left(\mathbb{R}^+ \bigcap \mathbb{R}^-\right) \times \left(\mathbb{R}^+ \bigcap \mathbb{R}^-\right) \\
&\underset{\text{[eq: 9.60]}}{=} \varnothing \times \varnothing \\
&= \varnothing \\
(\mathbb{R}^+ \times \mathbb{R}^+) \bigcap \left((\mathbb{R} \times \{0\}) \bigcup (\{0\} \times \mathbb{R})\right) &= \left((\mathbb{R}^+ \times \mathbb{R}^-) \bigcap (\mathbb{R} \times \{0\})\right) \bigcup \left((\mathbb{R}^+ \times \mathbb{R}^+) \bigcap (\{0\} \times \mathbb{R})\right) \\
&= \left(\left(\mathbb{R}^+ \bigcap \mathbb{R}\right) \times \left(\mathbb{R}^- \bigcap \{0\}\right)\right) \bigcup \left(\left(\mathbb{R}^+ \bigcap \{0\}\right) \times \left(\mathbb{R}^+ \bigcap \mathbb{R}\right)\right) \\
&\underset{\text{[eq: 9.60]}}{=} \left(\left(\mathbb{R}^+ \bigcap \mathbb{R}\right) \times \varnothing\right) \bigcup \left(\varnothing \times \left(\mathbb{R}^+ \bigcap \mathbb{R}\right)\right) \\
&= \varnothing \bigcup \varnothing \\
&= \varnothing \\
(\mathbb{R}^+ \times \mathbb{R}^-) \bigcap (\mathbb{R}^- \times \mathbb{R}^+) &= \left(\mathbb{R}^+ \bigcap \mathbb{R}^-\right) \times \left(\mathbb{R}^- \bigcap \mathbb{R}^+\right) \\
&\underset{\text{[eq: 9.60]}}{=} \varnothing \times \varnothing \\
&= \varnothing \\
(\mathbb{R}^+ \times \mathbb{R}^-) \bigcap (\mathbb{R}^- \times \mathbb{R}^-) &= \left(\mathbb{R}^+ \bigcap \mathbb{R}^-\right) \times \left(\mathbb{R}^- \bigcap \mathbb{R}^-\right) \\
&\underset{\text{[eq: 9.60]}}{=} \varnothing \times \left(\mathbb{R}^+ \bigcap \mathbb{R}^-\right) \\
&= \varnothing \\
(\mathbb{R}^+ \times \mathbb{R}^-) \bigcap \left((\mathbb{R} \times \{0\}) \bigcup (\{0\} \times \mathbb{R})\right) &= \left((\mathbb{R}^+ \times \mathbb{R}^-) \bigcap (\mathbb{R} \times \{0\})\right) \bigcup \left((\mathbb{R}^+ \times \mathbb{R}^-) \bigcap (\{0\} \times \mathbb{R})\right) \\
&= \left(\left(\mathbb{R}^+ \bigcap \mathbb{R}\right) \times \left(\mathbb{R}^- \bigcap \{0\}\right)\right) \bigcup \left(\left(\mathbb{R}^+ \bigcap \{0\}\right) \times \left(\mathbb{R}^- \bigcap \mathbb{R}\right)\right) \\
&\underset{\text{[eq: 9.60]}}{=} \left(\left(\mathbb{R}^+ \bigcap \mathbb{R}\right) \times \varnothing\right) \bigcup \left(\varnothing \times \left(\mathbb{R}^- \bigcap \mathbb{R}\right)\right) \\
&= \varnothing \bigcup \varnothing \\
&= \varnothing \\
(\mathbb{R}^- \times \mathbb{R}^+) \bigcap (\mathbb{R}^- \times \mathbb{R}^-) &= \left(\mathbb{R}^- \bigcap \mathbb{R}^-\right) \times \left(\mathbb{R}^+ \bigcap \mathbb{R}^-\right) \\
&= \left(\mathbb{R}^- \bigcap \mathbb{R}^-\right) \times \varnothing \\
&= \varnothing \\
(\mathbb{R}^- \times \mathbb{R}^+) \bigcap \left((\mathbb{R} \times \{0\}) \bigcup (\{0\} \times \mathbb{R})\right) &= \left(\left((\mathbb{R}^- \times \mathbb{R}^+) \bigcap (\mathbb{R} \times \{0\})\right) \bigcup \left((\mathbb{R}^- \times \mathbb{R}^+) \bigcap (\{0\} \times \mathbb{R})\right)\right) \\
&= \left(\left(\mathbb{R}^- \bigcap \mathbb{R}\right) \times \left(\mathbb{R}^+ \bigcap \{0\}\right)\right) \bigcup \left(\left(\mathbb{R}^- \bigcap \{0\}\right) \times \left(\mathbb{R}^+ \bigcap \mathbb{R}\right)\right) \\
&\underset{\text{[eq: 9.60]}}{=} \left(\left(\mathbb{R}^- \bigcap \mathbb{R}\right) \times \varnothing\right) \bigcup \left(\varnothing \times \left(\mathbb{R}^+ \bigcap \mathbb{R}\right)\right) \\
&= \varnothing \bigcup \varnothing \\
&= \varnothing \\
(\mathbb{R}^- \times \mathbb{R}^-) \bigcap \left((\mathbb{R} \times \{0\}) \bigcup (\{0\} \times \mathbb{R})\right) &= \left((\mathbb{R}^- \times \mathbb{R}^-) \bigcap (\mathbb{R} \times \{0\})\right) \bigcup \left((\mathbb{R}^- \times \mathbb{R}^-) \bigcap (\{0\} \times \mathbb{R})\right) \\
&= \left(\left(\mathbb{R}^- \bigcap \mathbb{R}\right) \times \left(\mathbb{R}^- \bigcap \{0\}\right)\right) \bigcup \left(\left(\mathbb{R}^- \bigcap \{0\}\right) \times \left(\mathbb{R}^- \bigcap \mathbb{R}\right)\right) \\
&\underset{\text{[eq: 9.60]}}{=} \left(\left(\mathbb{R}^- \bigcap \mathbb{R}\right) \times \varnothing\right) \bigcup \left(\varnothing \times \left(\mathbb{R}^- \bigcap \mathbb{R}\right)\right) \\
&= \varnothing \bigcup \varnothing \\
&= \varnothing
\end{aligned}
$$

$\square$

We use now [theorem: 9.20] and [theorem: 9.27] to define the multiplication operator on $\mathbb{R}$.

Definition 9.28. *The multiplication operator* $\cdot : \mathbb{R} \times \mathbb{R} \Rightarrow \mathbb{R}$ *is defined as*

$$\alpha \cdot \beta = \begin{cases} \alpha \odot \beta \text{ if } (\alpha, b) \in \mathbb{R}^+ \times \mathbb{R}^+ \\ -((-\alpha) \odot \beta) \text{ if } (\alpha, \beta) \in \mathbb{R}^- \times \mathbb{R}^+ \\ -(\alpha \odot (-\beta)) \text{ if } (\alpha, \beta) \in \mathbb{R}^+ \times \mathbb{R}^- \\ (-\alpha) \odot (-\beta) \text{ if } (\alpha, \beta) \in \mathbb{R}^- \times \mathbb{R}^- \\ 0 \text{ if } (\alpha, \beta) \text{ if } (\alpha, \beta) \in (\mathbb{R} \times \{0\}) \bigcup (\{0\} \times \mathbb{R}) \end{cases}$$

If we want to prove something about multiplication then we have 5 cases to consider for the definition of the multiplication operator. The following lemma allows to reduce the amount work.

Lemma 9.29. $\forall \alpha, \beta \in \mathbb{R} \times \mathbb{R}$ *we have* $-(\alpha \cdot \beta) = (-\alpha) \cdot \beta = \alpha \cdot (-\beta)$

Proof. We have to consider the following exclusive 5 cases [see theorem: 9.27]:

$(\boldsymbol{\alpha}, \boldsymbol{\beta}) \in \mathbb{R}^+ \times \mathbb{R}^+$**.** Then

$$\begin{aligned} (-\alpha) \cdot \beta &\underset{-\alpha \in \mathbb{R}^- \wedge \beta \in \mathbb{R}^+}{=} -((-(-\alpha)) \odot \beta) \\ &\underset{-(-\alpha) = \alpha}{=} -(\alpha \odot \beta) \\ &\underset{\alpha \in \mathbb{R}^+ \wedge \beta \in \mathbb{R}^+}{=} -(\alpha \cdot \beta) \\ \alpha \cdot (-\beta) &\underset{\alpha \in \mathbb{R}^+ \wedge -\beta \in \mathbb{R}^-}{=} -(\alpha \odot (-(-\beta))) \\ &= -(\alpha \odot \beta) \\ &\underset{\alpha \in \mathbb{R}^+ \wedge \beta \in \mathbb{R}^+}{=} -(\alpha \cdot \beta) \end{aligned}$$

$(\boldsymbol{\alpha}, \boldsymbol{\beta}) \in \mathbb{R}^+ \times \mathbb{R}^-$**.** Then

$$\begin{aligned} -(\alpha \cdot \beta) &\underset{\alpha \in \mathbb{R}^+ \wedge \beta \in \mathbb{R}^-}{=} -(-(\alpha \odot (-\beta))) \\ &= \alpha \odot (-\beta) \\ &\underset{\alpha \in \mathbb{R}^+ \wedge -\beta \in \mathbb{R}^-}{=} \alpha \cdot (-\beta) \qquad (9.63) \\ (-\alpha) \cdot \beta &\underset{-\alpha \in \mathbb{R}^- \wedge \beta \in \mathbb{R}^-}{=} (-(-\alpha)) \odot (-\beta) \\ &= \alpha \odot (-\beta) \\ &\underset{\alpha \in \mathbb{R}^+ \wedge -\beta \in \mathbb{R}^+}{=} \alpha \cdot (-\beta) \\ &\underset{[\text{eq:}9.63}{=} -(\alpha \cdot \beta) \end{aligned}$$

$(\boldsymbol{\alpha}, \boldsymbol{\beta}) \in \mathbb{R}^- \times \mathbb{R}^+$**.** Then

$$\begin{aligned} -(\alpha \cdot \beta) &\underset{\alpha \in \mathbb{R}^- \wedge \beta \in \mathbb{R}^+}{=} -(-((-\alpha) \odot \beta)) \\ &= (-\alpha) \odot \beta \\ &\underset{-\alpha \in \mathbb{R}^+ \wedge \beta \in \mathbb{R}^+}{=} (-\alpha) \cdot \beta \qquad (9.64) \\ \alpha \cdot (-\beta) &\underset{\alpha \in \mathbb{R}^- \wedge -\beta \in \mathbb{R}^-}{=} (-\alpha) \odot (-(-\beta)) \\ &= (-\alpha) \odot \beta \\ &\underset{-\alpha \in \mathbb{R}^+ \wedge \beta \in \mathbb{R}^+}{=} (-\alpha) \cdot \beta \\ &\underset{[\text{eq: } 9.64]}{=} -(\alpha \cdot \beta) \end{aligned}$$

$(\alpha,\beta)\in\mathbb{R}^-\times\mathbb{R}^-$. Then

$$\begin{array}{rcl} -(\alpha\cdot\beta) & \underset{\alpha\in\mathbb{R}^-\wedge\beta\in\mathbb{R}^-}{=} & -((-\alpha)\odot(-\beta)) \\ (-\alpha)\cdot\beta & \underset{-\alpha\in\mathbb{R}^+\wedge\beta\in\mathbb{R}^-}{=} & -((-\alpha)\odot(-\beta)) \\ & \underset{[\text{eq: } 9.65]}{=} & -(\alpha\cdot\beta) \\ \alpha\cdot(-\beta) & \underset{\alpha\in\mathbb{R}^-\wedge-\beta\in\mathbb{R}^+}{=} & -((-\alpha)\odot(-\beta)) \\ & \underset{[\text{eq: } 9.65]}{=} & -(\alpha\cdot\beta) \end{array} \tag{9.65}$$

$(\alpha,\beta)\in(\{0\}\times\mathbb{R})\bigcup(\mathbb{R}\times\{0\})$. Then

$$\begin{array}{rcl} -(\alpha\cdot\beta) & \underset{(\alpha,\beta)\in(\{0\}\times\mathbb{R})\bigcup(\mathbb{R}\times\{0\})}{=} & 0 \\ & \underset{(-\alpha,\beta)\in(\{0\}\times\mathbb{R})\bigcup(\mathbb{R}\times\{0\})}{=} & (-\alpha)\cdot\beta \\ -(\alpha\cdot\beta) & \underset{(\alpha,\beta)\in(\{0\}\times\mathbb{R})\bigcup(\mathbb{R}\times\{0\})}{=} & 0 \\ & \underset{(\alpha,-\beta)\in(\{0\}\times\mathbb{R})\bigcup(\mathbb{R}\times\{0\})}{=} & \alpha\cdot(-\beta) \end{array}$$

□

Lemma 9.30. *If $\alpha,\beta\in\mathbb{R}^+$ and $\gamma\in\mathbb{R}^-$ then $\alpha\cdot(\beta+\gamma)=\alpha\cdot\beta+\alpha\cdot\gamma$.*

Proof. First we proof that

$$\forall\alpha,\beta\in\mathbb{R}^+\wedge\gamma\in\mathbb{R}^- \text{ such that } \beta+\gamma\in\mathbb{R}^+ \text{ we have } \alpha\cdot(\beta+\gamma) \tag{9.66}$$

Proof. As $\beta,-\gamma\in\mathbb{R}^+$ we have $0\in\beta\wedge 0\in-\gamma$ so that $0=0+0\in\beta+(-\gamma)$ proving that

$$\beta+(-\gamma)\in\mathbb{R}^+\wedge\beta+\gamma\in\mathbb{R}^+ \tag{9.67}$$

Now

$$\begin{array}{rcl} \alpha\cdot\beta+\alpha\cdot\beta & \underset{\alpha,\beta\in\mathbb{R}^+}{=} & \alpha\odot\beta+\alpha\odot\beta \\ & \underset{\alpha,\beta\in\mathbb{R}^+ \text{ and [theorem: } 9.24]}{=} & \alpha\odot(\beta+\beta) \\ & = & \alpha\odot((\beta+(-\gamma))+(\beta+\gamma)) \\ & \underset{[\text{eq: } 9.67+\text{theorem: } 9.24]}{=} & (\alpha\odot(\beta+(-\gamma)))+\alpha\odot(\beta+\gamma) \\ & \underset{[\beta,-\gamma\in\mathbb{R}^+ + \text{theorem: } 9.24]}{=} & \alpha\odot\beta+\alpha\odot(-\gamma)+\alpha\odot(\beta+\gamma) \\ & = & \alpha\odot\beta+(-(-(\alpha\odot(-\gamma))))+\alpha\odot(\beta+\gamma) \\ & \underset{\alpha,\beta\in\mathbb{R}^+,\gamma\in\mathbb{R}^-,\beta+\gamma\in\mathbb{R}^+}{=} & \alpha\cdot\beta+(-(\alpha\cdot\gamma))+\alpha\cdot(\beta+\gamma) \end{array}$$

so after adding $-(\alpha\cdot\beta)+\alpha\cdot\gamma$ to both sides gives

$$\begin{array}{rcl} \alpha\cdot(\beta+\gamma) & = & \alpha\cdot\beta+\alpha\cdot\beta+(-(\alpha\cdot\beta))+\alpha\cdot\gamma \\ & = & \alpha\cdot\beta+\alpha\cdot\gamma \end{array}$$

□

For $\beta+\gamma\in\mathbb{R}$ we have three cases to consider:

$\beta+\gamma\in\mathbb{R}^+$. Then $\alpha\cdot(\beta+\gamma)\underset{[\text{eq: } 9.66]}{=}\alpha\cdot\beta+\alpha\cdot\gamma$

$\beta+\gamma\in\mathbb{R}^-$. Then $(-\beta)+(-\gamma)\underset{[\text{theorem: } 4.8]}{=}-(\beta+\gamma)\in\mathbb{R}^+$. So if we take $\gamma'=-\beta\in\mathbb{R}^-$ and $\beta'=-\gamma\in\mathbb{R}^+$ we have that $\beta'+\gamma'=-(\beta+\gamma)\in\mathbb{R}^+$, so we can apply [eq: 9.66] resulting in

$$\alpha\cdot(\beta'+\gamma')=\alpha\cdot\beta'+\alpha\cdot\gamma'$$

which after substituting the formulas for β',γ' gives

$$\alpha\cdot((-\gamma)+(-\beta))=\alpha\cdot(-\gamma)+\alpha\cdot(-\beta) \tag{9.68}$$

Now we have

$$\begin{array}{rcl}\alpha\cdot(\beta+\gamma) & = & \alpha\cdot(-(-(\beta+\gamma)))\\ & \underset{[\text{theorem: } 9.29]}{=} & -(\alpha\cdot(-(\beta+\gamma)))\\ & = & -(\alpha\cdot((-\gamma)+(-\beta)))\\ & \underset{[\text{eq: } 9.68]}{=} & -(\alpha\cdot(-\gamma)+\alpha\cdot(-\beta))\\ & = & -(\alpha\cdot(-\beta)+\alpha\cdot(-\gamma))\\ & \underset{[\text{theorem: } 4.8]}{=} & (-(\alpha\cdot(-\beta)))+(-(\alpha\cdot(-\gamma)))\\ & \underset{[\text{theorem: } 9.29]}{=} & \alpha\cdot(-(-\beta))+\alpha\cdot(-(-\gamma))\\ & = & \alpha\cdot\beta+\alpha\cdot\gamma\end{array}$$

$\beta+\gamma=0$. Then $\gamma=-\beta$ and we have

$$\begin{array}{rcl}\alpha\cdot(\beta+\gamma) & = & \alpha\cdot 0\\ & = & 0\\ & = & \alpha\cdot\beta+(-(\alpha\cdot\beta))\\ & \underset{[\text{theorem: } 9.29]}{=} & \alpha\cdot\beta+\alpha\cdot(-\beta)\\ & = & \alpha\cdot\beta+\alpha\cdot\gamma\end{array}$$

□

We are finally ready to prove that $\langle\mathbb{R},+,\cdot\rangle$ is a field

Definition 9.31. *If $\alpha\in\mathbb{R}\setminus\{0\}$ then we define α^{-1} by* $\alpha^{-1}=\begin{cases}\text{inv}(\alpha) & \text{if } \alpha\in\mathbb{R}^+\\ -\text{inv}(-\alpha) & \text{if } \alpha\in\mathbb{R}^-\end{cases}$

Proof. As by [theorem: 9.17] $\mathbb{R}\setminus\{0\}=\mathbb{R}^+\bigcup\mathbb{R}^-$ and $\mathbb{R}^+\bigcap\mathbb{R}^-=\varnothing$ α^{-1} is well defined. □

Theorem 9.32. *$\langle\mathbb{R},+,\cdot\rangle$ is a field where*

1. *$+:\mathbb{R}\times\mathbb{R}\to\mathbb{R}$is defined in [theorem: 9.15]*
2. *$\cdot:\mathbb{R}\times\mathbb{R}\to\mathbb{R}$ is defined in [definition: 9.28]*
3. *$0=\alpha_0$ is the additive neutral element [see theorem: 9.15]*
4. *$1=\alpha_1$ is the multiplicative neutral element.*
5. *$\forall\alpha\in\mathbb{R}$ the additive inverse is the negative cut of α [see theorem: 9.15]*
6. *$\forall\alpha\in\mathbb{R}\setminus\{0\}$ we have the multiplicative inverse is defined by [definition: 9.31]*

Proof.

1. Using [theorem: 9.15] $\langle\mathbb{R},+\rangle$ is a Abelian group with neutral element $0=\alpha_0$ and $\forall\alpha\in\mathbb{R}$ the negative cut $-\alpha$ as inverse.
2. For the multiplication operator $\cdot:\mathbb{R}\times\mathbb{R}\to\mathbb{R}$ we have:

   commutativity. Let $(\alpha,\beta)\in\mathbb{R}\times\mathbb{R}$ then using [theorem: 9.27] we have to consider the following cases:

   $(\alpha,\beta)\in\mathbb{R}^+\times\mathbb{R}^+$. Then

$$\begin{array}{rcl}\alpha\cdot\beta & \underset{\alpha,\beta\in\mathbb{R}^+}{=} & \alpha\odot\beta\\ & \underset{[\text{theorem: } 9.22]}{=} & \beta\odot\alpha\\ & = & \beta\cdot\alpha\end{array}$$

   $(\alpha.\beta)\in\mathbb{R}^+\times\mathbb{R}^-$. Then

$$\begin{array}{rcl}\alpha\cdot\beta & \underset{\alpha\in\mathbb{R}^+\wedge\beta\in\mathbb{R}^-}{=} & -(\alpha\odot(-\beta))\\ & \underset{[\text{theorem: } 9.22]}{=} & -((-\beta)\odot\alpha)\\ & \underset{\alpha\in\mathbb{R}^+\wedge-\beta\in\mathbb{R}^+}{=} & -((-\beta)\cdot\alpha)\\ & \underset{[\text{theorem: } 9.29]}{=} & -(-(\beta\cdot\alpha))\\ & = & \beta\cdot\alpha\end{array}$$

$(\alpha, \beta) \in \mathbb{R}^- \times \mathbb{R}^+$. Then

$$\begin{array}{rcl} \alpha \cdot \beta & \underset{\alpha \in \mathbb{R}^- \wedge \beta \in \mathbb{R}^+}{=} & -((-\alpha) \odot \beta) \\ & \underset{[\text{theorem: } 9.22]}{=} & -(\beta \odot (-\alpha)) \\ & \underset{-\alpha \in \mathbb{R}^+ \wedge \beta \in \mathbb{R}^+}{=} & -(\beta \cdot (-\alpha)) \\ & \underset{[\text{theorem: } 9.29]}{=} & -(-(\beta \cdot \alpha)) \\ & = & \beta \cdot \alpha \end{array}$$

$(\alpha, \beta) \in \mathbb{R}^- \times \mathbb{R}^-$. Then

$$\begin{array}{rcl} \alpha \cdot \beta & \underset{a, \beta \in \mathbb{R}^-}{=} & (-\alpha) \odot (-\beta) \\ & \underset{[\text{theorem: } 9.22]}{=} & (-\beta) \odot (-\alpha) \\ & \underset{a, \beta \in \mathbb{R}^-}{=} & \beta \cdot \alpha \end{array}$$

$(\alpha, \beta) \in (\{0\} \times \mathbb{R}) \bigcup (\mathbb{R} \times \{0\})$. Then

$$\begin{array}{rcl} \alpha \cdot \beta & = & 0 \\ & = & \beta \cdot \alpha \end{array}$$

neutral element. First note that as $0 \in \alpha_1$ we have $\alpha_1 \in \mathbb{R}^+$. Let $\alpha \in \mathbb{R} \underset{[\text{theorem: } 9.17]}{=} \mathbb{R}^+ \bigcup \mathbb{R}^- \bigcup \{0\}$ then we have either:

$\alpha \in \mathbb{R}^+$. Then we have

$$\begin{array}{rcl} \alpha \cdot \alpha_1 & \underset{\text{commutativity}}{=} & \alpha_1 \cdot \alpha \\ & \underset{\alpha_1, \alpha \in \mathbb{R}^+}{=} & \alpha_1 \odot \alpha \\ & \underset{[\text{theorem: } 9.21]}{=} & \alpha \end{array}$$

$\alpha \in \mathbb{R}^-$. Then we have

$$\begin{array}{rcl} \alpha \cdot \alpha_1 & \underset{\text{commutativity}}{=} & \alpha_1 \cdot \alpha \\ & \underset{\alpha_1 \in \mathbb{R}^+ \wedge \alpha \in \mathbb{R}^-}{=} & -(\alpha_1 \odot (-\alpha)) \\ & \underset{[\text{theorem: } 9.21]}{=} & -(-\alpha) \\ & = & \alpha \end{array}$$

$\alpha = 0$. Then we have

$$\begin{array}{rcl} \alpha \cdot \alpha_1 & = & \alpha_1 \cdot \alpha \\ & = & 0 \\ & = & \alpha \end{array}$$

inverse element. Let $\alpha \in \mathbb{R} \backslash \{0\}$ then by [theorem: 9.17] we have to consider:

$\alpha \in \mathbb{R}^+$. Then $\operatorname{inv}(\alpha) \in \mathbb{R}^+$ [see theorem: 9.25] and

$$\begin{array}{rcl} \alpha^{-1} \cdot \alpha & \underset{\text{commutativity}}{=} & \alpha \cdot \alpha^{-1} \\ & \underset{\alpha \in \mathbb{R}^+}{=} & \alpha \cdot \operatorname{inv}(\alpha) \\ & \underset{\operatorname{inv}(\alpha) \in \mathbb{R}^+ \wedge \alpha \in \mathbb{R}^+}{=} & \alpha \odot \operatorname{inv}(\alpha) \\ & \underset{[\text{theorem: } 9.26]}{=} & \alpha_1 \end{array}$$

$\alpha \in \mathbb{R}^-$. Then $-\alpha \in \mathbb{R}^+$ and by [theorem: 9.17] $\mathrm{inv}(-\alpha) \in \mathbb{R}^+$, further

$$\begin{aligned}\alpha^{-1} \cdot \alpha &\underset{\text{commutativity}}{=} \alpha \cdot \alpha^{-1} \\ &\underset{\alpha \in \mathbb{R}^-}{=} \alpha \cdot (-(\mathrm{inv}(-\alpha))) \\ &\underset{[\text{theorem: } 9.29]}{=} -(\alpha \cdot \mathrm{inv}(-\alpha)) \\ &\underset{[\text{theorem: } 9.29]}{=} (-\alpha) \cdot \mathrm{inv}(-\alpha) \\ &\underset{[\text{theorem:: } 9.26]}{=} \alpha_1\end{aligned}$$

associativity. As $\mathbb{R} \underset{[\text{theorem: } 9.17]}{=} \mathbb{R}^+ \bigcup \mathbb{R}^- \bigcup \{0\}$ we have for $\alpha, \beta, \gamma \in \mathbb{R}$ the following 27 cases to consider:

$\alpha \in \mathbb{R}^+ \wedge \beta \in \mathbb{R}^+ \wedge \gamma \in \mathbb{R}^+$. Then we have

$$\begin{aligned}\alpha \cdot (\beta \cdot \gamma) &= \alpha \odot (\beta \odot \gamma) \\ &\underset{[\text{theorem: } 9.23]}{=} (\alpha \odot \beta) \odot \gamma \\ &= (\alpha \cdot \beta) \cdot \gamma\end{aligned}$$

$\alpha \in \mathbb{R}^+ \wedge \beta \in \mathbb{R}^+ \wedge \gamma \in \mathbb{R}^-$. Then we have

$$\begin{aligned}\alpha \cdot (\beta \cdot \gamma) &= \alpha \cdot (-(-(\beta \cdot \gamma))) \\ &\underset{[\text{theorem: } 9.29]}{=} \alpha \cdot (-(\beta \cdot (-\gamma))) \\ &\underset{[\text{theorem: } 9.29]}{=} -(\alpha \cdot (\beta \cdot (-\gamma))) \\ &= -(\alpha \odot (\beta \odot (-\gamma))) \\ &\underset{[\text{theorem: } 9.23]}{=} -((\alpha \odot \beta) \odot (-\gamma)) \\ &= -((\alpha \cdot \beta) \cdot (-\gamma)) \\ &\underset{[\text{theorem: } 9.29]}{=} (\alpha \cdot \beta) \cdot (-(-\gamma)) \\ &= (\alpha \cdot \beta) \cdot \gamma\end{aligned}$$

$\alpha \in \mathbb{R}^+ \wedge \beta \in \mathbb{R}^+ \wedge \gamma = 0$. Then we have

$$\begin{aligned}\alpha \cdot (\beta \cdot \gamma) &= \alpha \cdot 0 \\ &= 0 \\ &= (\alpha \cdot \beta) \cdot 0 \\ &= (\alpha \cdot \beta) \cdot \gamma\end{aligned}$$

$\alpha \in \mathbb{R}^+ \wedge \beta \in \mathbb{R}^- \wedge \gamma \in \mathbb{R}^+$. Then we have

$$\begin{aligned}\alpha \cdot (\beta \cdot \gamma) &= \alpha \cdot ((-(-\beta)) \cdot \gamma) \\ &\underset{[\text{theorem: } 9.29]}{=} \alpha \cdot (-((-\beta) \cdot \gamma)) \\ &\underset{[\text{theorem: } 9.29]}{=} -(\alpha \cdot ((-\beta) \cdot \gamma)) \\ &= -(\alpha \odot ((-\beta) \odot \gamma)) \\ &\underset{[\text{theorem: } 9.23]}{=} -((\alpha \odot (-\beta)) \odot \gamma) \\ &= -((\alpha \cdot (-\beta)) \cdot \gamma) \\ &\underset{[\text{theorem: } 9.29]}{=} (-(\alpha \cdot (-\beta))) \cdot \gamma \\ &\underset{[\text{theorem: } 9.29]}{=} (\alpha \cdot (-(-\beta))) \cdot \gamma \\ &= (\alpha \cdot \beta) \cdot \gamma\end{aligned}$$

$\boldsymbol{\alpha \in \mathbb{R}^+ \wedge \beta \in \mathbb{R}^- \wedge \gamma \in \mathbb{R}^-}$**.** Then we have

$$\begin{array}{rcl}
\alpha \cdot (b \cdot \gamma) & = & \alpha \cdot ((-(-\beta)) \cdot (-(-\gamma))) \\
& \underset{[\text{theorem: } 9.29]}{=} & \alpha \cdot (-((-\beta) \cdot (-(-\gamma)))) \\
& \underset{[\text{theorem: } 9.29]}{=} & \alpha \cdot (-(-((-\beta) \cdot (-\gamma)))) \\
& = & \alpha \cdot ((-\beta) \cdot (-\gamma)) \\
& = & \alpha \odot ((-\beta) \odot (-\gamma)) \\
& \underset{[\text{theorem: } 9.23]}{=} & (\alpha \odot (-\beta)) \odot (-\gamma) \\
& = & (\alpha \cdot (-\beta)) \cdot (-\gamma) \\
& \underset{[\text{theorem: } 9.29]}{=} & -((\alpha \cdot (-\beta)) \cdot \gamma) \\
& = & -((-(\alpha \cdot \beta)) \cdot \gamma) \\
& \underset{[\text{theorem: } 9.29]}{=} & -(-((a \cdot \beta) \cdot \gamma)) \\
& = & (\alpha \cdot \beta) \cdot \gamma
\end{array}$$

$\boldsymbol{\alpha \in \mathbb{R}^+ \wedge \beta \in \mathbb{R}^- \wedge \gamma = 0}$**.** Then we have

$$\begin{array}{rcl}
\alpha \cdot (\beta \cdot \gamma) & = & \alpha \cdot (\beta \cdot 0) \\
& = & \alpha \cdot 0 \\
& = & 0 \\
& = & (\alpha \cdot \beta) \cdot 0 \\
& = & (\alpha \cdot \beta) \cdot \gamma
\end{array}$$

$\boldsymbol{\alpha \in \mathbb{R}^+ \wedge \beta = 0 \wedge \gamma \in \mathbb{R}^+}$**.** Then we have

$$\begin{array}{rcl}
\alpha \cdot (\beta \cdot \gamma) & = & \alpha \cdot (0 \cdot \gamma) \\
& = & \alpha \cdot 0 \\
& = & 0 \\
& = & 0 \cdot \gamma \\
& = & (\alpha \cdot 0) \cdot \gamma \\
& = & (\alpha \cdot \beta) \cdot \gamma
\end{array}$$

$\boldsymbol{\alpha \in \mathbb{R}^+ \wedge \beta = 0 \wedge \gamma \in \mathbb{R}^-}$**.** Then we have

$$\begin{array}{rcl}
\alpha \cdot (\beta \cdot \gamma) & = & \alpha \cdot (0 \cdot \gamma) \\
& = & \alpha \cdot 0 \\
& = & 0 \\
& = & 0 \cdot \gamma \\
& = & (\alpha \cdot 0) \cdot \gamma \\
& = & (\alpha \cdot \beta) \cdot \gamma
\end{array}$$

$\boldsymbol{\alpha \in \mathbb{R}^+ \wedge \beta = 0 \wedge \gamma = 0}$**.** Then we have

$$\begin{array}{rcl}
\alpha \cdot (\beta \cdot \gamma) & = & \alpha \cdot (0 \cdot \gamma) \\
& = & \alpha \cdot 0 \\
& = & 0 \\
& = & 0 \cdot \gamma \\
& = & (\alpha \cdot 0) \cdot \gamma \\
& = & (\alpha \cdot \beta) \cdot \gamma
\end{array}$$

$\boldsymbol{\alpha \in \mathbb{R}^- \wedge \beta \in \mathbb{R}^+ \wedge \gamma \in \mathbb{R}^+}$**.** Then we have

$$\begin{array}{rcl}
\alpha \cdot (\beta \cdot \gamma) & = & (-(-\alpha)) \cdot (\beta \cdot \gamma) \\
& \underset{[\text{theorem: } 9.29]}{=} & -((-\alpha) \cdot (\beta \cdot \gamma)) \\
& = & -((-\alpha) \odot (\beta \odot \gamma))
\end{array}$$

$$
\begin{aligned}
&\underset{[\text{theorem: } 9.23]}{=} -(((-\alpha)\odot\beta)\odot\gamma)\\
&= -(((-\alpha)\cdot\beta)\cdot\gamma)\\
&\underset{[\text{theorem: } 9.29]}{=} -((-(\alpha\cdot\beta))\cdot\gamma)\\
&\underset{[\text{theorem: } 9.29]}{=} -(-((\alpha\cdot\beta)\cdot\gamma))\\
&= (\alpha\cdot\beta)\cdot\gamma
\end{aligned}
$$

$\alpha\in\mathbb{R}^-\wedge\beta\in\mathbb{R}^+\wedge\gamma\in\mathbb{R}^-$. Then we have

$$
\begin{aligned}
\alpha\cdot(\beta\cdot\gamma) &= (-(-\alpha))\cdot(\beta\cdot(-(-\gamma)))\\
&\underset{[\text{theorem: } 9.29]}{=} -((-\alpha)\cdot(\beta\cdot(-(-\gamma))))\\
&\underset{[\text{theorem: } 9.29]}{=} -((-\alpha)\cdot(-(\beta\cdot(-\gamma))))\\
&\underset{[\text{theorem: } 9.29]}{=} -(-((-\alpha)\cdot(\beta\cdot(-\gamma))))\\
&= (-\alpha)\cdot(\beta\cdot(-\gamma))\\
&= (-\alpha)\odot(\beta\odot(-\gamma))\\
&\underset{[\text{theorem: } 9.23]}{=} ((-\alpha)\odot\beta)\odot(-\gamma)\\
&= ((-\alpha)\cdot\beta)\cdot(-\gamma)\\
&\underset{[\text{theorem: } 9.29]}{=} -(((-\alpha)\cdot\beta)\cdot\gamma)\\
&\underset{[\text{theorem: } 9.29]}{=} -((-(\alpha\cdot\beta))\cdot\gamma)\\
&\underset{[\text{theorem: } 9.29]}{=} -(-((\alpha\cdot\beta)\cdot\gamma))\\
&= (\alpha\cdot\beta)\cdot\gamma
\end{aligned}
$$

$\alpha\in\mathbb{R}^-\wedge\beta\in\mathbb{R}^+\wedge\gamma=0$. Then we have

$$
\begin{aligned}
\alpha\cdot(\beta\cdot\gamma) &= \alpha\cdot(\beta\cdot 0)\\
&= \alpha\cdot 0\\
&= 0\\
&= (\alpha\cdot\beta)\cdot 0\\
&= (\alpha\cdot\beta)\cdot\gamma
\end{aligned}
$$

$\alpha\in\mathbb{R}^-\wedge\beta\in\mathbb{R}^-\wedge\gamma\in\mathbb{R}^+$. Then we have

$$
\begin{aligned}
\alpha\cdot(\beta\cdot\gamma) &= (-(-\alpha))\cdot((-(-\beta))\cdot\gamma)\\
&\underset{[\text{theorem: } 9.29]}{=} -((-\alpha)\cdot((-(-\beta))\cdot\gamma))\\
&\underset{[\text{theorem: } 9.29]}{=} -((-\alpha)\cdot(-((-\beta)\cdot\gamma)))\\
&\underset{[\text{theorem: } 9.29]}{=} -(-((-\alpha)\cdot((-\beta)\cdot\gamma)))\\
&= (-\alpha)\cdot((-\beta)\cdot\gamma)\\
&= (-\alpha)\odot((-\beta)\odot\gamma)\\
&\underset{[\text{theorem: } 9.23]}{=} ((-\alpha)\odot(-\beta))\odot\gamma\\
&= (\alpha\cdot\beta)\cdot\gamma
\end{aligned}
$$

$\alpha\in\mathbb{R}^-\wedge\beta\in\mathbb{R}^-\wedge\gamma\in\mathbb{R}^-$. Then we have

$$
\begin{aligned}
\alpha\cdot(\beta\cdot\gamma) &= \alpha\cdot((-\beta)\odot(-\gamma))\\
&= -((-\alpha)\odot((-\beta)\odot(-\gamma)))\\
&\underset{[\text{theorem: } 9.23]}{=} -(((-\alpha)\odot(-\beta))\odot(-\gamma))\\
&= -((\alpha\cdot\beta)\cdot(-\gamma))\\
&\underset{[\text{theorem: } 9.29]}{=} -(-((\alpha\cdot\beta)\cdot\gamma))\\
&= (\alpha\cdot\beta)\cdot\gamma
\end{aligned}
$$

$\boldsymbol{\alpha \in \mathbb{R}^- \wedge \beta \in \mathbb{R}^- \wedge \gamma = 0.}$ Then we have

$$\begin{aligned}
\alpha \cdot (\beta \cdot \gamma) &= \alpha \cdot (\beta \cdot 0) \\
&= \alpha \cdot 0 \\
&= 0 \\
&= (\alpha \cdot \beta) \cdot 0 \\
&= (\alpha \cdot \beta) \cdot \gamma
\end{aligned}$$

$\boldsymbol{\alpha \in \mathbb{R}^- \wedge \beta = 0 \wedge \gamma \in \mathbb{R}^+.}$ Then we have

$$\begin{aligned}
\alpha \cdot (\beta \cdot \gamma) &= \alpha \cdot (0 \cdot \gamma) \\
&= \alpha \cdot 0 \\
&= 0 \\
&= 0 \cdot \gamma \\
&= (\alpha \cdot 0) \cdot \gamma \\
&= (\alpha \cdot \beta) \cdot \gamma
\end{aligned}$$

$\boldsymbol{\alpha \in \mathbb{R}^- \wedge \beta = 0 \wedge \gamma \in \mathbb{R}^-.}$ Then we have

$$\begin{aligned}
\alpha \cdot (\beta \cdot \gamma) &= \alpha \cdot (0 \cdot \gamma) \\
&= \alpha \cdot 0 \\
&= 0 \\
&= 0 \cdot \gamma \\
&= (\alpha \cdot 0) \cdot \gamma \\
&= (\alpha \cdot \beta) \cdot \gamma
\end{aligned}$$

$\boldsymbol{\alpha \in \mathbb{R}^- \wedge \beta = 0 \wedge \gamma = 0.}$ Then we have

$$\begin{aligned}
\alpha \cdot (\beta \cdot \gamma) &= \alpha \cdot (0 \cdot \gamma) \\
&= \alpha \cdot 0 \\
&= 0 \\
&= 0 \cdot \gamma \\
&= (\alpha \cdot 0) \cdot \gamma \\
&= (\alpha \cdot \beta) \cdot \gamma
\end{aligned}$$

$\boldsymbol{\alpha = 0 \wedge \beta \in \mathbb{R}^+ \wedge \gamma \in \mathbb{R}^+.}$ Then we have

$$\begin{aligned}
\alpha \cdot (\beta \cdot \gamma) &= 0 \cdot (\beta \cdot \gamma) \\
&= 0 \\
&= 0 \cdot \gamma \\
&= (0 \cdot \beta) \cdot \gamma \\
&= (\alpha \cdot \beta) \cdot \gamma
\end{aligned}$$

$\boldsymbol{\alpha = 0 \wedge \beta \in \mathbb{R}^+ \wedge \gamma \in \mathbb{R}^-.}$ Then we have

$$\begin{aligned}
\alpha \cdot (\beta \cdot \gamma) &= 0 \cdot (\beta \cdot \gamma) \\
&= 0 \\
&= 0 \cdot \gamma \\
&= (0 \cdot \beta) \cdot \gamma \\
&= (\alpha \cdot \beta) \cdot \gamma
\end{aligned}$$

$\boldsymbol{\alpha = 0 \wedge \beta \in \mathbb{R}^+ \wedge \gamma = 0}$**.** Then we have

$$\begin{array}{rcl}\alpha\cdot(\beta\cdot\gamma) & = & 0\cdot(\beta\cdot\gamma)\\ & = & 0\\ & = & 0\cdot\gamma\\ & = & (0\cdot\beta)\cdot\gamma\\ & = & (\alpha\cdot\beta)\cdot\gamma\end{array}$$

$\boldsymbol{\alpha = 0 \wedge \beta \in \mathbb{R}^- \wedge \gamma \in \mathbb{R}^+}$**.** Then we have

$$\begin{array}{rcl}\alpha\cdot(\beta\cdot\gamma) & = & 0\cdot(\beta\cdot\gamma)\\ & = & 0\\ & = & 0\cdot\gamma\\ & = & (0\cdot\beta)\cdot\gamma\\ & = & (\alpha\cdot\beta)\cdot\gamma\end{array}$$

$\boldsymbol{\alpha = 0 \wedge \beta \in \mathbb{R}^- \wedge \gamma \in \mathbb{R}^-}$**.** Then we have

$$\begin{array}{rcl}\alpha\cdot(\beta\cdot\gamma) & = & 0\cdot(\beta\cdot\gamma)\\ & = & 0\\ & = & 0\cdot\gamma\\ & = & (0\cdot\beta)\cdot\gamma\\ & = & (\alpha\cdot\beta)\cdot\gamma\end{array}$$

$\boldsymbol{\alpha = 0 \wedge \beta \in \mathbb{R}^- \wedge \gamma = 0}$**.** Then we have

$$\begin{array}{rcl}\alpha\cdot(\beta\cdot\gamma) & = & 0\cdot(\beta\cdot\gamma)\\ & = & 0\\ & = & 0\cdot\gamma\\ & = & (0\cdot\beta)\cdot\gamma\\ & = & (\alpha\cdot\beta)\cdot\gamma\end{array}$$

$\boldsymbol{\alpha = 0 \wedge \beta = 0 \wedge \gamma \in \mathbb{R}^+}$**.** Then we have

$$\begin{array}{rcl}\alpha\cdot(\beta\cdot\gamma) & = & 0\cdot(\beta\cdot\gamma)\\ & = & 0\\ & = & 0\cdot\gamma\\ & = & (0\cdot\beta)\cdot\gamma\\ & = & (\alpha\cdot\beta)\cdot\gamma\end{array}$$

$\boldsymbol{\alpha = 0 \wedge \beta = 0 \wedge \gamma \in \mathbb{R}^-}$**.** Then we have

$$\begin{array}{rcl}\alpha\cdot(\beta\cdot\gamma) & = & 0\cdot(\beta\cdot\gamma)\\ & = & 0\\ & = & 0\cdot\gamma\\ & = & (0\cdot\beta)\cdot\gamma\\ & = & (\alpha\cdot\beta)\cdot\gamma\end{array}$$

$\boldsymbol{\alpha = 0 \wedge \beta = 0 \wedge \gamma = 0}$**.** Then we have

$$\begin{array}{rcl}\alpha\cdot(\beta\cdot\gamma) & = & 0\cdot(\beta\cdot\gamma)\\ & = & 0\\ & = & 0\cdot\gamma\\ & = & (0\cdot\beta)\cdot\gamma\\ & = & (\alpha\cdot\beta)\cdot\gamma\end{array}$$

distributivity. As $\mathbb{R} \underset{[\text{theorem: } 9.17]}{=} \mathbb{R}^+ \bigcup \mathbb{R}^- \bigcup \{0\}$ we have for $\alpha, \beta, \gamma \in \mathbb{R}$ the following 27 cases to consider:

$\alpha \in \mathbb{R}^+ \wedge \beta \in \mathbb{R}^+ \wedge \gamma \in \mathbb{R}^+$. Then

$$\begin{aligned}\alpha \cdot (\beta + \gamma) &= \alpha \odot (\beta + \gamma) \\ &\underset{[\text{theorem: } 9.24]}{=} \alpha \odot \beta + \alpha \odot \gamma \\ &= \alpha \cdot \beta + \alpha \cdot \gamma\end{aligned}$$

$\alpha \in \mathbb{R}^+ \wedge \beta \in \mathbb{R}^+ \wedge \gamma \in \mathbb{R}^-$. Then

$$\alpha \cdot (b + \gamma) \underset{[\text{lemma: } 9.30]}{=} \alpha \cdot \beta + \alpha \cdot \gamma$$

$\alpha \in \mathbb{R}^+ \wedge \beta \in \mathbb{R}^+ \wedge \gamma = 0$. Then

$$\begin{aligned}\alpha \cdot (\beta + \gamma) &= \alpha \cdot (\beta + 0) \\ &= \alpha \cdot \beta \\ &= \alpha \cdot \beta + 0 \\ &= \alpha \cdot \beta + \alpha \cdot 0 \\ &= \alpha \cdot \beta + \alpha \cdot \gamma\end{aligned}$$

$\alpha \in \mathbb{R}^+ \wedge \beta \in \mathbb{R}^- \wedge \gamma \in \mathbb{R}^+$. Then

$$\begin{aligned}\alpha \cdot (\beta + \gamma) &= \alpha \cdot (\gamma + \beta) \\ &\underset{[\text{lemma: } 9.30]}{=} \alpha \cdot \gamma + \alpha \cdot \beta \\ &= \alpha \cdot \beta + \alpha \cdot \gamma\end{aligned}$$

$\alpha \in \mathbb{R}^+ \wedge \beta \in \mathbb{R}^- \wedge \gamma \in \mathbb{R}^-$. Then

$$\begin{aligned}\alpha \cdot (\beta + \gamma) &= \alpha \cdot (-(-(\beta + \gamma))) \\ &\underset{[\text{theorem: } 9.29]}{=} -(\alpha \cdot (-(\beta + \gamma))) \\ &\underset{[\text{theorem: } 4.8]}{=} -(\alpha \cdot ((-\beta) + (-\gamma))) \\ &= -(\alpha \odot ((-\beta) + (-\gamma))) \\ &\underset{[\text{theorem: } 9.24]}{=} -(\alpha \odot (-\beta) + \alpha \odot (-\gamma)) \\ &= -(\alpha \cdot (-\beta) + \alpha \cdot (-\gamma)) \\ &\underset{[\text{theorem: } 4.8]}{=} (-(\alpha \cdot (-\beta))) + (-(\alpha \cdot (-\gamma))) \\ &\underset{[\text{theorem: } 9.29]}{=} \alpha \cdot (-(-\beta)) + \alpha \cdot (-(-\gamma)) \\ &= \alpha \cdot \beta + \alpha \cdot \gamma\end{aligned}$$

$\alpha \in \mathbb{R}^+ \wedge \beta \in \mathbb{R}^- \wedge \gamma = 0$. Then

$$\begin{aligned}\alpha \cdot (\beta + \gamma) &= \alpha \cdot (\beta + 0) \\ &= \alpha \cdot \beta \\ &= \alpha \cdot \beta + 0 \\ &= \alpha \cdot \beta + \alpha \cdot 0 \\ &= \alpha \cdot \beta + \alpha \cdot \gamma\end{aligned}$$

$\alpha \in \mathbb{R}^+ \wedge \beta = 0 \wedge \gamma \in \mathbb{R}^+$. Then

$$\begin{aligned}\alpha \cdot (\beta + \gamma) &= \alpha \cdot (0 + \gamma) \\ &= \alpha \cdot \gamma \\ &= 0 + \alpha \cdot \gamma \\ &= \alpha \cdot \beta + \alpha \cdot \gamma\end{aligned}$$

$\alpha \in \mathbb{R}^+ \wedge \beta = 0 \wedge \gamma \in \mathbb{R}^-$. Then

$$\begin{aligned}\alpha\cdot(\beta+\gamma) &= \alpha\cdot(0+\gamma)\\ &= \alpha\cdot\gamma\\ &= 0+\alpha\cdot\gamma\\ &= \alpha\cdot\beta+\alpha\cdot\gamma\end{aligned}$$

$\alpha \in \mathbb{R}^+ \wedge \beta = 0 \wedge \gamma = 0$. Then

$$\begin{aligned}\alpha\cdot(\beta+\gamma) &= \alpha\cdot(0+\gamma)\\ &= \alpha\cdot\gamma\\ &= 0+\alpha\cdot\gamma\\ &= \alpha\cdot\beta+\alpha\cdot\gamma\end{aligned}$$

$\alpha \in \mathbb{R}^- \wedge \beta \in \mathbb{R}^+ \wedge \gamma \in \mathbb{R}^+$. Then

$$\begin{aligned}\alpha\cdot(\beta+\gamma) &= -((-\alpha)\odot(\beta+\gamma))\\ &\underset{\text{[theorem: 9.24]}}{=} -((-\alpha)\odot\beta+(-\alpha)\odot\gamma)\\ &\underset{\text{[theorem: 4.8]}}{=} (-((-\alpha)\odot\beta))+(-((-\alpha)\odot\gamma))\\ &= (-((-\alpha)\cdot\beta))+(-((-\alpha)\cdot\gamma))\\ &\underset{\text{[theorem: 9.29]}}{=} (-(-\alpha))\cdot\beta+(-(-\alpha))\cdot\gamma\\ &= \alpha\cdot\beta+\alpha\cdot\gamma\end{aligned}$$

$\alpha \in \mathbb{R}^- \wedge \beta \in \mathbb{R}^+ \wedge \gamma \in \mathbb{R}^-$. Then

$$\begin{aligned}\alpha\cdot(\beta+\gamma) &= (-(-\alpha))\cdot(\beta+\gamma)\\ &\underset{\text{[theorem: 9.29]}}{=} -((-\alpha)\cdot(\beta+\gamma))\\ &\underset{\text{[lemma: 9.30]}}{=} -((-\alpha)\cdot\beta+(-\alpha)\cdot\gamma)\\ &\underset{\text{[theorem: 4.8]}}{=} (-((-\alpha)\cdot\beta))+(-((-\alpha)\cdot\gamma))\\ &\underset{\text{[theorem: 9.29]}}{=} (-(-\alpha))\cdot\beta+(-(-\alpha))\cdot\gamma\\ &= \alpha\cdot\beta+\alpha\cdot\gamma\end{aligned}$$

$\alpha \in \mathbb{R}^- \wedge \beta \in \mathbb{R}^+ \wedge \gamma = 0$. Then

$$\begin{aligned}\alpha\cdot(\beta+\gamma) &= \alpha\cdot(\beta+0)\\ &= \alpha\cdot\beta\\ &= \alpha\cdot\beta+0\\ &= \alpha\cdot\beta+\alpha\cdot 0\\ &= \alpha\cdot\beta+\alpha\cdot\gamma\end{aligned}$$

$\alpha \in \mathbb{R}^- \wedge \beta \in \mathbb{R}^- \wedge \gamma \in \mathbb{R}^+$. Then

$$\begin{aligned}\alpha\cdot(\beta+\gamma) &= \alpha\cdot(\gamma+\beta)\\ &= (-(-\alpha))\cdot(\gamma+\beta)\\ &\underset{\text{[theorem: 9.29]}}{=} -((-\alpha)\cdot(\gamma+\beta))\\ &\underset{\text{[lemma: 9.30]}}{=} -((-\alpha)\cdot\gamma+(-\alpha)\cdot\beta)\\ &\underset{\text{[theorem: 4.8]}}{=} (-((-\alpha)\cdot\gamma))+(-((-\alpha)\cdot\beta))\\ &\underset{\text{[theorem: 9.29]}}{=} (-(-\alpha))\cdot\gamma+(-(-\alpha))\cdot\beta\\ &= \alpha\cdot\gamma+\alpha\cdot\beta\\ &= \alpha\cdot\beta+\alpha\cdot\gamma\end{aligned}$$

$\alpha \in \mathbb{R}^- \wedge \beta \in \mathbb{R}^- \wedge \gamma \in \mathbb{R}^-$. Then

$$\begin{array}{rcl}\alpha \cdot (\beta + \gamma) & = & (-(-\alpha)) \cdot (\beta + \gamma) \\ & \underset{[\text{theorem: } 9.29]}{=} & -((-\alpha) \cdot (\beta + \gamma)) \\ & \underset{[\text{theorem: } 9.29]}{=} & (-\alpha) \cdot (-(\beta + \gamma)) \\ & \underset{[\text{theorem: } 4.8]}{=} & (-\alpha) \cdot ((-\beta) + (-\gamma)) \\ & = & (-\alpha) \odot ((-\beta) + (-\gamma)) \\ & \underset{[\text{theorem: } 9.24]}{=} & (-\alpha) \odot (-\beta) + (-\alpha) \odot (-\gamma) \\ & = & (-\alpha) \cdot (-\beta) + (-\alpha) \cdot (-\gamma) \\ & \underset{[\text{theorem: } 9.29]}{=} & (-(\alpha \cdot (-\beta))) + (-(\alpha \cdot (-\gamma))) \\ & \underset{[\text{theorem: } 9.29]}{=} & (-(-(\alpha \cdot \beta))) + (-(-(\alpha \cdot \gamma))) \\ & = & \alpha \cdot \beta + \alpha \cdot \gamma\end{array}$$

$\alpha \in \mathbb{R}^- \wedge \beta \in \mathbb{R}^- \wedge \gamma = 0$. Then

$$\begin{array}{rcl}\alpha \cdot (\beta + \gamma) & = & \alpha \cdot (\beta + 0) \\ & = & \alpha \cdot \beta \\ & = & \alpha \cdot \beta + 0 \\ & = & \alpha \cdot \beta + \alpha \cdot 0 \\ & = & \alpha \cdot \beta + \alpha \cdot \gamma\end{array}$$

$\alpha \in \mathbb{R}^- \wedge \beta = 0 \wedge \gamma \in \mathbb{R}^+$. Then

$$\begin{array}{rcl}\alpha \cdot (\beta + \gamma) & = & \alpha \cdot (0 + \gamma) \\ & = & \alpha \cdot \gamma \\ & = & 0 + \alpha \cdot \gamma \\ & = & \alpha \cdot 0 + \alpha \cdot \gamma \\ & = & \alpha \cdot \beta + \alpha \cdot \gamma\end{array}$$

$\alpha \in \mathbb{R}^- \wedge \beta = 0 \wedge \gamma \in \mathbb{R}^-$. Then

$$\begin{array}{rcl}\alpha \cdot (\beta + \gamma) & = & \alpha \cdot (0 + \gamma) \\ & = & \alpha \cdot \gamma \\ & = & 0 + \alpha \cdot \gamma \\ & = & \alpha \cdot 0 + \alpha \cdot \gamma \\ & = & \alpha \cdot \beta + \alpha \cdot \gamma\end{array}$$

$\alpha \in \mathbb{R}^- \wedge \beta = 0 \wedge \gamma = 0$. Then

$$\begin{array}{rcl}\alpha \cdot (\beta + \gamma) & = & \alpha \cdot (0 + \gamma) \\ & = & \alpha \cdot \gamma \\ & = & 0 + \alpha \cdot \gamma \\ & = & \alpha \cdot 0 + \alpha \cdot \gamma \\ & = & \alpha \cdot \beta + \alpha \cdot \gamma\end{array}$$

$\alpha = 0 \wedge \beta \in \mathbb{R}^+ \wedge \gamma \in \mathbb{R}^+$. Then

$$\begin{array}{rcl}\alpha \cdot (\beta + \gamma) & = & 0 \cdot (\beta + \gamma) \\ & = & 0 \\ & = & 0 + 0 \\ & = & 0 \cdot \beta + 0 \cdot \gamma \\ & = & \alpha \cdot \beta + \alpha \cdot \gamma\end{array}$$

$\boldsymbol{\alpha = 0 \wedge \beta \in \mathbb{R}^+ \wedge \gamma \in \mathbb{R}^-}$**.** Then

$$\begin{aligned}\alpha\cdot(\beta+\gamma) &= 0\cdot(\beta+\gamma)\\ &= 0\\ &= 0+0\\ &= 0\cdot\beta+0\cdot\gamma\\ &= \alpha\cdot\beta+\alpha\cdot\gamma\end{aligned}$$

$\boldsymbol{\alpha = 0 \wedge \beta \in \mathbb{R}^+ \wedge \gamma = 0}$**.** Then

$$\begin{aligned}\alpha\cdot(\beta+\gamma) &= 0\cdot(\beta+\gamma)\\ &= 0\\ &= 0+0\\ &= 0\cdot\beta+0\cdot\gamma\\ &= \alpha\cdot\beta+\alpha\cdot\gamma\end{aligned}$$

$\boldsymbol{\alpha = 0 \wedge \beta \in \mathbb{R}^- \wedge \gamma \in \mathbb{R}^+}$**.** Then

$$\begin{aligned}\alpha\cdot(\beta+\gamma) &= 0\cdot(\beta+\gamma)\\ &= 0\\ &= 0+0\\ &= 0\cdot\beta+0\cdot\gamma\\ &= \alpha\cdot\beta+\alpha\cdot\gamma\end{aligned}$$

$\boldsymbol{\alpha = 0 \wedge \beta \in \mathbb{R}^- \wedge \gamma \in \mathbb{R}^-}$**.** Then

$$\begin{aligned}\alpha\cdot(\beta+\gamma) &= 0\cdot(\beta+\gamma)\\ &= 0\\ &= 0+0\\ &= 0\cdot\beta+0\cdot\gamma\\ &= \alpha\cdot\beta+\alpha\cdot\gamma\end{aligned}$$

$\boldsymbol{\alpha = 0 \wedge \beta \in \mathbb{R}^- \wedge \gamma = 0}$**.** Then

$$\begin{aligned}\alpha\cdot(\beta+\gamma) &= 0\cdot(\beta+\gamma)\\ &= 0\\ &= 0+0\\ &= 0\cdot\beta+0\cdot\gamma\\ &= \alpha\cdot\beta+\alpha\cdot\gamma\end{aligned}$$

$\boldsymbol{\alpha = 0 \wedge \beta = 0 \wedge \gamma \in \mathbb{R}^+}$**.** Then

$$\begin{aligned}\alpha\cdot(\beta+\gamma) &= 0\cdot(\beta+\gamma)\\ &= 0\\ &= 0+0\\ &= 0\cdot\beta+0\cdot\gamma\\ &= \alpha\cdot\beta+\alpha\cdot\gamma\end{aligned}$$

$\boldsymbol{\alpha = 0 \wedge \beta = 0 \wedge \gamma \in \mathbb{R}^-}$**.** Then

$$\begin{aligned}\alpha\cdot(\beta+\gamma) &= 0\cdot(\beta+\gamma)\\ &= 0\\ &= 0+0\\ &= 0\cdot\beta+0\cdot\gamma\\ &= \alpha\cdot\beta+\alpha\cdot\gamma\end{aligned}$$

$\boldsymbol{\alpha = 0 \wedge \beta = 0 \wedge \gamma = 0.}$ Then

$$\begin{array}{rcl}\alpha\cdot(\beta+\gamma) & = & 0\cdot(\beta+\gamma)\\ & = & 0\\ & = & 0+0\\ & = & 0\cdot\beta+0\cdot\gamma\\ & = & \alpha\cdot\beta+\alpha\cdot\gamma\end{array}$$

3. Assume that $\alpha_0 = \alpha_1$ which by [theorem: 9.7] proves that in $\mathbb{Q}$ we have $0 = 1$. However by [theorem: 8.12] $\langle Q, +, \cdot\rangle$ is a field so that $0 \neq 1$ and we reach a contradiction. Hence we must have that

$$0 = \alpha_0 \neq \alpha_1 = 1$$

□

Remember that $x + (-y)$ has a shorthand notation $x - y$, in the same way we have some shorthand notations for multiplication with a inverse element.

Notation 9.33. *If $x, y \in \mathbb{R}$ $x \neq 0$ then we use the following shorthand notation*

1. *x^{-1} is noted as $1/x$*
2. *$y \cdot x^{-1}$ is noted as y/x*

We show now how the rational numbers as a field can be embedded in the field of the real numbers. The primary candidate for this are the rational cuts, so lets review some of the properties of the rational cuts. First we need two little lemmas.

Lemma 9.34. *If $r \in \mathbb{Q}$ such that $\alpha_r \in \mathbb{R}^+$ then $\operatorname{inv}(\alpha_r) = \alpha_{r^{-1}}$*

Proof. As $\alpha_r \in \mathbb{R}^+$ we have $0 \in \alpha_r$ hence $0 < r \Rightarrow 0 < r^{-1}$. Let $x \in \operatorname{inv}(\alpha_r)$ then we have for x the following to consider:

$\boldsymbol{x \leqslant 0.}$ Then $x \leqslant 0 < r^{-1}$ proving that $x \in \alpha_{r^{-1}}$.

$\boldsymbol{0 < x.}$ Then $x \notin \mathbb{Q}_0^-$ so there exists a $s \in \mathbb{Q}\backslash\alpha_r$ such that $0 < s$ and $\exists t \in \mathbb{Q}\backslash\alpha \vdash t < s$ such that $x = s^{-1}$. Hence $s = x^{-1}$ and $x^{-1} \in \mathbb{Q}\backslash\alpha_r$ and $\exists t \in \mathbb{Q}\backslash\alpha_r$ such that $t < x^{-1}$. As α_r is a rational cut we have by [theorem: 9.4] that $r = \min(\mathbb{Q}\backslash\alpha_r)$ hence $\forall t \in \mathbb{Q}\backslash\alpha_r$ we have $r \leqslant t$, from which we conclude that $x^{-1} \neq r$. As $x^{-1} \in \mathbb{Q}\backslash\alpha_r \Rightarrow r \leqslant x^{-1}$ we conclude that $0 < r < x^{-1}$ or using [theorems: 8.29, 4.73] that $x < r^{-1}$. Hence we have $x \in \alpha_{r^{-1}}$

So it follows that

$$\operatorname{inv}(\alpha_r) \subseteq \alpha_{r^{-1}} \tag{9.69}$$

Let $x \in \alpha_{r^{-1}}$ then for x we have either:

$\boldsymbol{x \leqslant 0.}$ Then $x \in \mathbb{Q}_0^-$ so that $x \in \operatorname{inv}(\alpha_r)$.

$\boldsymbol{0 < x.}$ Then as $x \in \alpha_{r^{-1}}$ we have that $0 < x < r^{-1}$ so that by [theorems: 8.29,4.73] $0 < r < x^{-1}$, hence $x^{-1} \notin \alpha_r$ or $x^{-1} \in \mathbb{Q}\backslash\alpha_r$, further $\mathbb{Q}\backslash\alpha \ni \min(\mathbb{Q}\backslash\alpha_r) \underset{\text{[theorem: 9.4]}}{=} r < x^{-1}$. Summarized we have $x^{-1} \in \mathbb{Q}\backslash\alpha_r \wedge 0 < x^{-1} \wedge \exists t \in \mathbb{Q}\backslash\alpha_r \vdash t < x^{-1}$ proving that $x = (x^{-1})^{-1} \in \operatorname{inv}(\alpha_r)$.

So in all cases $x \in \operatorname{inv}(\alpha_r)$, hence $\alpha_{r^{-1}} \subseteq \operatorname{inv}(\alpha_r)$, combining this with [eq: 9.69] gives

$$\operatorname{inv}(\alpha_r) = \alpha_{r^{-1}}$$

□

Lemma 9.35. *Let $\alpha_r, \alpha_s \in \mathbb{R}^+$ then $\alpha_r \odot \alpha_s = \alpha_{r\cdot s}$*

Proof. As $\alpha_r, \alpha_s \in \mathbb{R}^+$ we have $0 \in \alpha_r \wedge 0 \in \alpha_s$ so that

$$0 < r \wedge 0 < s \tag{9.70}$$

Let $x \in \alpha_r \odot \alpha_s$ then we have the following possibilities:

$\boldsymbol{x \leqslant 0.}$ Then as $0 < r \wedge 0 < s$ we have that $0 < r \cdot s$ so that $x < r \cdot s$ proving that $x \in \alpha_{r\cdot s}$

$\boldsymbol{0 < x.}$ Then we have $x \notin \mathbb{Q}_0^-$ so there exists $u \in \alpha_r$ and $v \in \alpha_s$ so that $x = u \cdot v$ with $0 < u$ and $0 < v$. As $u \in \alpha_r$ and $v \in \alpha_s$ we have $u < r$ and $v < s$. So $u \cdot v < r \cdot v$ and $r \cdot v < r \cdot s$ hence $x = u \cdot v < r \cdot s$.

So we conclude that

$$\alpha_r \odot \alpha_s \subseteq \alpha_{r\cdot s} \tag{9.71}$$

Let $x \in \alpha_{r \cdot s}$ then $x < r \cdot s$. For x we have now the following cases to consider:

$x \leqslant 0$. Then $x \in \mathbb{Q}_0^-$ so that $x \in \alpha_r \cdot \alpha_s$.

$0 < x$. As $0 < r$ we have by the density of $\mathbb{Q}$ [see theorem: 8.37] the existence of $\varepsilon_1 \in \mathbb{Q}$ such that $0 < \varepsilon_1 < r$. From $x < r \cdot s$ it follows that $0 < r \cdot s - x$ hence, as $0 < s \Rightarrow 0 < s^{-1}$ we have that $0 < (r \cdot s - x) \cdot s^{-1} = r - x \cdot s^{-1}$. Using density of $\mathbb{Q}$ again there exist a $\varepsilon_2 \in \mathbb{Q}$ such that $0 < \varepsilon_2 < r - x \cdot s^{-1}$. Take now $\varepsilon = \min(\varepsilon_1, \varepsilon_2)$ then we have

$$0 < \varepsilon \leqslant \varepsilon_1 < r \text{ and } 0 < \varepsilon \leqslant \varepsilon_2 < r - x \cdot s^{-1} \tag{9.72}$$

From the above we have $x \cdot s^{-1} < r - \varepsilon$ or as $0 < x \wedge 0 < s^{-1} \Rightarrow 0 < x \cdot s^{-1}$ that $0 < x \cdot s^{-1} < r - \varepsilon$ allowing us to apply [theorems: 8.29, 4.73] giving

$$0 < (r - \varepsilon)^{-1} < (x \cdot s^{-1})^{-1} \underset{\text{[theorem: 4.55]}}{=} (s^{-1})^{-1} \cdot x^{-1} = s \cdot x^{-1}$$

multiplying both sides x we get by [theorems: 8.29, 4.73] and $0 < x$ that $0 < x \cdot (r - \varepsilon)^{-1} < s$ so that

$$0 < x \cdot (r - \varepsilon)^{-1} \in \alpha_s \tag{9.73}$$

As $0 < \varepsilon < r$ [see eq: 9.72] we have $0 < r - \varepsilon < r$ so that

$$0 < r - s \in \alpha_r \tag{9.74}$$

Now $(x \cdot (r - \varepsilon)^{-1}) \cdot (r - s) = x$ which combined with [eqs: 9.73, 9.74] proves that $x \in \alpha_r \cdot \alpha_s$.

So in all cases we have $x \in \alpha_r \cdot \alpha_s$ hence it follows that $\alpha_{r \cdot s} \subseteq \alpha_r \cdot \alpha_s$. Combining this with [eq: 9.71] proves that

$$\alpha_{r \cdot s} = \alpha_r \cdot \alpha_s$$

□

Theorem 9.36. *Let $r, s \in \mathbb{Q}$ then we have*

1. $\alpha_r + \alpha_s = \alpha_{r+s}$
2. $-\alpha_r = \alpha_{-r}$
3. $\alpha_r \cdot \alpha_s = \alpha_{r \cdot s}$
4. *If $\alpha_r \neq 0$ then* $1/\alpha_r \underset{\text{notation}}{=} (\alpha_r)^{-1} = \alpha_{r^{-1}}$

Proof.

1. Let $x \in \alpha_r + \alpha_s$ then there exists $u \in \alpha_r$ and $v \in \alpha_s$ such that $x = u + v$. As $u \in \alpha_r$ and $v \in \alpha_s$ we have that $u < r$ and $v < s$ so that $u + v < r + v$ and $v + r < s + r$ giving $x = u + v < r + s$ proving that $x \in \alpha_r + \alpha_s$. Hence we have

$$\alpha_r + \alpha_s \subseteq \alpha_{r+s} \tag{9.75}$$

Let $x \in \alpha_{r+s}$ then $x < r + s$ hence $x - r < s$. Using the density of $\mathbb{Q}$ [see theorem: 8.37] there exist a $z \in \mathbb{Q}$ such that $x - r < z < s$. Then $z \in \alpha_s$ and if we define $\varepsilon = z - (x - r)$ we have $0 < \varepsilon \Rightarrow -\varepsilon < 0$. So $r - \varepsilon = r + (-\varepsilon) < r$ proving that $r - \varepsilon \in \alpha_r$. Hence

$$(r - \varepsilon) + z \in \alpha_r + \alpha_s \tag{9.76}$$

Now

$$\begin{aligned} (r - \varepsilon) + z &= r - (z - (x - r)) + z \\ &= r - z + x - r + z \\ &= x \end{aligned}$$

so that by [eq: 9.76] $x \in \alpha_r + \alpha_s$. Hence $\alpha_{r+s} \subseteq \alpha_r + \alpha_s$ which together with [eq: 9.75] proves

$$\alpha_{r+s} = \alpha_r + \alpha_s$$

2. This is stated in [theorem: 9.11]

3. Using [theorem: 9.27] we have to look at the following five cases:

$\alpha_r \in \mathbb{R}^+ \wedge \alpha_s \in \mathbb{R}^+$. Then

$$\begin{aligned} \alpha_r \cdot \alpha_s &= \alpha_r \odot \alpha_s \\ &\underset{\text{[lemma: 9.35]}}{=} \alpha_{r \cdot s} \end{aligned}$$

$\boldsymbol{\alpha_r \in \mathbb{R}^+ \wedge \alpha_s \in \mathbb{R}^-}$**.** Then

$$\begin{array}{rcl}
\alpha_r \cdot \alpha_s & = & -(\alpha_r \odot (-\alpha_s)) \\
& \underset{(2)}{=} & -(\alpha_r \odot \alpha_{-s}) \\
& \underset{[\text{lemma: } 9.35]}{=} & -\alpha_{r \cdot (-s)} \\
& \underset{[\text{theorem: } 4.40}{=} & -\alpha_{-(r \cdot s)} \\
& \underset{(2)}{=} & \alpha_{-(-(r \cdot s))} \\
& = & \alpha_{r \cdot s}
\end{array}$$

$\boldsymbol{\alpha_r \in \mathbb{R}^- \wedge \alpha_s \in \mathbb{R}^+}$**.** Then

$$\begin{array}{rcl}
\alpha_r \cdot \alpha_s & = & -((-\alpha_r) \odot \alpha_s) \\
& = & -(\alpha_{-r} \odot \alpha_s) \\
& \underset{[\text{lemma: } 9.35]}{=} & -(\alpha_{(-r) \cdot s}) \\
& \underset{[\text{theorem: } 4.40}{=} & -(\alpha_{-(r \cdot s)}) \\
& \underset{(2)}{=} & \alpha_{-(-(r \cdot s))} \\
& = & \alpha_{r \cdot s}
\end{array}$$

$\boldsymbol{\alpha_r \in \mathbb{R}^- \wedge \alpha_s \in \mathbb{R}^-}$**.** Then

$$\begin{array}{rcl}
\alpha_r \cdot \alpha_s & = & (-\alpha_r) \odot (-\alpha_s) \\
& \underset{(2)}{=} & \alpha_{-r} \odot \alpha_{-s} \\
& \underset{[\text{lemma: } 9.35]}{=} & \alpha_{(-r) \cdot (-s)} \\
& \underset{[\text{theorem: } 4.40]}{=} & \alpha_{r \cdot s}
\end{array}$$

$\boldsymbol{(\alpha_r, \alpha_s) \in (\{0\} \times \mathbb{R}) \bigcup (\mathbb{R} \times \{0\})}$**.** Then we have two sub cases:

$\boldsymbol{(\alpha_r, \alpha_s) \in \{0\} \times \mathbb{R}}$**.** Then $\alpha_r = 0 = \alpha_0 \Rightarrow_{[\text{theorem: } 9.7]} r = 0$ and

$$\begin{array}{rcl}
\alpha_r \cdot \alpha_s & = & 0 \cdot \alpha_s \\
& = & 0 \\
& = & \alpha_0 \\
& = & \alpha_{0 \cdot s} \\
& = & \alpha_{r \cdot s}
\end{array}$$

$\boldsymbol{(\alpha_r, \alpha_s) \in \mathbb{R} \times \{0\}}$**.** Then $\alpha_s = 0 = \alpha_0 \Rightarrow_{[\text{theorem: } 9.7]} s = 0$ and

$$\begin{array}{rcl}
\alpha_r \cdot \alpha_s & = & \alpha_r \cdot 0 \\
& = & 0 \\
& = & \alpha_0 \\
& = & \alpha_{r \cdot 0} \\
& = & \alpha_{r \cdot s}
\end{array}$$

So in all cases we have

$$\alpha_r \cdot \alpha_s = \alpha_{r \cdot s}$$

4. Let $\alpha_r \in \mathbb{R} \backslash \{0\}$ then we have the following possibilities:

$\boldsymbol{\alpha_r \in \mathbb{R}^+}$**.** Then $(\alpha_r)^{-1} = \text{inv}(\alpha_r) \underset{[\text{lemma: } 9.34]}{=} \alpha_{r^{-1}}$

$\alpha_r \in \mathbb{R}^-$. Then

$$\begin{array}{ccc}(\alpha_r)^{-1} & = & -\mathrm{inv}(-\alpha_r)\\ & \underset{(2)}{=} & -\mathrm{inv}(\alpha_{-r})\\ & \underset{\text{[lemma: 9.34]}}{=} & -\alpha_{(-r)^{-1}}\\ & \underset{\text{[theorems: 8.30]}}{=} & -\alpha_{-(r^{-1})}\\ & \underset{(2)}{=} & \alpha_{-(-(r^{-1}))}\\ & = & \alpha_{r^{-1}}\end{array}$$

□

We show now that $\mathbb{Q}_{\mathbb{R}}$ is a embedding of $\mathbb{Q}$ in $\mathbb{R}$ that conserves the field structure.

Theorem 9.37. *For $\mathbb{Q}_{\mathbb{R}} = \{\alpha_r | r \in \mathbb{Q}\}$ [definition: 9.6] we have:*

1. *$\mathbb{Q}_{\mathbb{R}}$ is a sub-field of $\langle \mathbb{R}, +, \cdot \rangle$*
2. *The function $i_{\mathbb{Q}\rightarrow\mathbb{R}}: \langle \mathbb{Q}, +, \cdot \rangle \rightarrow \langle \mathbb{R}, +, \cdot \rangle$ defined by $i_{\mathbb{Q}\rightarrow\mathbb{R}}(q) = \alpha_q$ is a field isomorphism*

Proof.

1. Let $x, y \in \mathbb{Q}_{\mathbb{R}}$ then we have that $\exists r, s \in \mathbb{Q}$ such that $x = \alpha_r$ and $y = \alpha_s$. Then we have:

   a. $x + y = \alpha_r + \alpha_s \underset{\text{[theorem: 9.36]}}{=} \alpha_{r+s} \in \mathbb{Q}_{\mathbb{R}}$

   b. $x \cdot y = \alpha_r \cdot \alpha_s \underset{\text{[theorem: 9.36]}}{=} \alpha_{r \cdot s} \in \mathbb{Q}_{\mathbb{R}}$

   c. If $x \neq 0$ then $x^{-1} = (\alpha_r)^{-1} \underset{\text{[theorem: 9.36]}}{=} \alpha_{r^{-1}} \in \mathbb{Q}_{\mathbb{R}}$

   d. $0 = \alpha_0 \in \mathbb{Q}_{\mathbb{R}}$

   e. $1 = \alpha_1 \in \mathbb{Q}_{\mathbb{R}}$

   which proves that $\mathbb{Q}_{\mathbb{R}}$ is a sub-field of $\langle \mathbb{R}, +, \cdot \rangle$.

2. Using [theorem: 9.7] it follows that

$$i_{\mathbb{Q}\rightarrow\mathbb{R}}: \mathbb{Q} \rightarrow \mathbb{R} \text{ is a bijection}$$

   Next we have to prove the homeomorphism properties:

   a. If $r, s \in \mathbb{Q}$ then $i_{\mathbb{Q}\rightarrow\mathbb{R}}(r+s) = \alpha_{r+s} \underset{\text{[theorem: 9.36]}}{=} \alpha_r + \alpha_s = i_{\mathbb{Q}\rightarrow\mathbb{R}}(r) + i_{\mathbb{Q}\rightarrow\mathbb{R}}(s)$

   b. If $r, s \in \mathbb{Q}$ then $i_{\mathbb{Q}\rightarrow\mathbb{R}}(r \cdot s) = \alpha_r \cdot \alpha_s \underset{\text{[theorem: 9.36]}}{=} \alpha_{r \cdot s} = i_{\mathbb{Q}\rightarrow\mathbb{R}}(r) \cdot i_{\mathbb{Q}\rightarrow\mathbb{R}}(s)$

   c. $i_{\mathbb{Q}\rightarrow\mathbb{R}}(1) = \alpha_1 = 1$ □

9.2 Order relation on $\mathbb{R}$

Theorem 9.38. *If $\alpha, \beta \in \mathbb{R}^+$ then*

1. *$\alpha + \beta \in \mathbb{R}^+$*
2. *$\alpha \cdot \beta \in \mathbb{R}^+$*
3. *$\alpha^{-1} \in \mathbb{R}^+$*

Proof.

1. As $\alpha, \beta \in \mathbb{R}^+$ we have that $0 \in \alpha$ and $0 \in \beta$ then $0 = 0 + 0 \in \alpha + \beta$ proving that

$$\alpha + \beta \in \mathbb{R}^+$$

2. As $0 \in \mathbb{Q}_0^-$ we have $\alpha \odot \beta \in \mathbb{R}^+$ so that

$$\alpha \cdot \beta \underset{\alpha,\beta \in \mathbb{R}^+}{=} \alpha \odot \beta \in \mathbb{R}^+$$

3. We have $\alpha^{-1} \underset{\alpha \in \mathbb{R}^+}{=} \mathrm{inv}(\alpha) \in \mathbb{R}^+$

□

We define now the relation on $\mathbb{R}$ that later will be proved to be a order relation, this definition mirrors the definition of order in $\mathbb{Z}$ and $\mathbb{Q}$ and is the reason why we have defined $\mathbb{R}^+$. One problem is that $0 \notin \mathbb{R}^+$, so we have first to define $<$ and base $\leqslant$ on $<$.

Definition 9.39. $< \subseteq \mathbb{R} \times \mathbb{R}$ *is defined by*

$$< = \{(\alpha, \beta) \in \mathbb{R} \times \mathbb{R} | \beta + (-\alpha) \in \mathbb{R}^+\}$$

or in other words for $\alpha, \beta \in \mathbb{R}$ *we have*

$$\alpha < \beta \Leftrightarrow \beta - \alpha = \beta + (-\alpha) \in \mathbb{R}^+$$

Definition 9.40. $\leqslant \subseteq \mathbb{R} \times \mathbb{R}$ *is defined by*

$$\leqslant = \{(\alpha, \beta) \in \mathbb{R} \times \mathbb{R} | \alpha = \beta \vee \beta + (-\alpha) \in \mathbb{R}^+\} = \{(\alpha, \beta) \in \mathbb{R} \times \mathbb{R} | \alpha = \beta \vee \alpha < \beta\}$$

or in other words for $\alpha, \beta \in \mathbb{R}$ *we have*

$$\alpha \leqslant \beta \Leftrightarrow \alpha = \beta \vee \alpha < \beta \Leftrightarrow \alpha = \beta \vee \beta - \alpha = \beta + (-\alpha) \in \mathbb{R}^+$$

The following theorem shows a simpler way to decide if $\alpha < \beta$ or $\alpha \leqslant \beta$.

Theorem 9.41. $\forall \alpha, \beta \in \mathbb{R}$ *we have*

1. $\alpha < \beta \Leftrightarrow \alpha \subset \beta$ *[strict inclusion]*
2. $\alpha \leqslant \beta \Leftrightarrow \alpha \subseteq \beta$

Proof.

1.

⇒. As $\alpha < \beta$ we have that $\beta + (-\alpha) \in \mathbb{R}^+$ so that

$$0 \in \beta + (-\alpha) \tag{9.77}$$

Let $r \in \alpha$. If $-r \in -\alpha$ then by the definition of a negative cut [see definition: 9.10] we have $r = -(-r) \in \mathbb{Q} \backslash \alpha$ contradicting $r \in \alpha$. Hence we must have that $-r \notin -\alpha$ or

$$-r \in \mathbb{Q} \backslash -\alpha \tag{9.78}$$

As $0 \in \beta + (-\alpha)$ [see eq: 9.77] there exists $s \in \beta$ and a $t \in -\alpha$ such that $0 = s + t$ or $s = -t$. As $t \in -\alpha$ and $-r \in \mathbb{Q} \backslash -\alpha$ [see eq: 9.78] it follows from [definition: 9.1 (3)] that $t < -r$ or $r < -t = s \in \beta$. So $r < s \in \beta$ which by [theorem: 9.3] proves that $r \in \beta$. Hence we have

$$\alpha \subseteq \beta \tag{9.79}$$

If now $\alpha = \beta$ then $\beta + (-\alpha) = \beta + (-\beta) = 0 = a_0$, as by [eq: 9.77] $0 \in \beta + (-\alpha)$ we find that $0 \in \alpha_0 = \{q \in \mathbb{Q} | q < 0\}$ a contradiction, so $\alpha \neq \beta$, which combined with [eq: 9.79] gives

$$\alpha \subset \beta$$

⇐. As $\alpha \subset \beta$ there exist a $r \in \beta$ such that $r \notin \alpha$ or

$$r \in \mathbb{Q} \backslash \alpha \tag{9.80}$$

As by [definition: 9.1 (4)] $\max(\beta)$ does not exist we have

$$\exists r' \in \beta \text{ such that } r < r'$$

If $r' \in \alpha$ then as $r \in \mathbb{Q} \backslash \alpha$ we have by [definition: 9.1 (3)] that $r' < r$ contradicting $r < r'$, hence $r' \notin a$ or

$$r' \in \mathbb{Q} \backslash \alpha$$

So we have that $-(-r') \in \mathbb{Q} \backslash \alpha$ and $r < r' = -(-r')$ where $r \in \mathbb{Q} \backslash \alpha$ which by the definition of a negative cut [see definition: 9.10] proves that $-r' \in -\alpha$. As $r' \in \beta$ we have that $0 = r' + (-r') \in \beta + (-\alpha)$ proving that $\beta + (-\alpha) \in \mathbb{R}^+$ or that

$$\alpha < \beta$$

2.

⇒. As $\alpha \leqslant \beta$ we have by (1) that $\alpha = \beta \vee \alpha \subset \beta$ so that $\alpha \subseteq \beta$

⇐. If $\alpha \subseteq \beta$ then $\alpha = \beta \wedge \alpha \subset \beta$ so that by (1) $\alpha = \beta \vee \alpha < b$ proving $a \leqslant \beta$ □

Corollary 9.42. *For* $\langle \mathbb{R}, \leqslant \rangle$ *we have that* $0 < 1$

Proof. Note that $1 = \alpha_1 = \{q \in \mathbb{Q} | q < 1\}$ and $0 = \alpha_0 = \{q \in \mathbb{Q} | q < 0\}$. So if $q \in \alpha_0$ we have $q < 0 < 1 \Rightarrow q \in \alpha_1$ proving that

$$\alpha_0 \subseteq \alpha_1$$

As in $\mathbb{Q}$ we have $0 < 1$ [see example: 8.25] we have by the density of $\mathbb{Q}$ [see theorem: 8.37] that there exist a $q \in \mathbb{Q}$ such that $0 < q < 1$ hence $q \in \alpha_1$ but $q \notin \alpha_0$ proving that

$$\alpha_0 \subset \alpha_1$$

which by the previous theorem [theorem: 9.41] proves that

$$0 < 1$$ □

Theorem 9.43. $\langle \mathbb{R}, \leqslant \rangle$ *is a totally ordered set*

Proof.

reflexivity. If $\alpha \in \mathbb{R}$ then $\alpha \subseteq \alpha$ so that by [theorem: 9.41] $\alpha \leqslant \alpha$

anti symmetry. If $\alpha \leqslant \beta$ and $\beta \leqslant \alpha$ then by [theorem: 9.41] we have $\alpha \subseteq \beta$ and $\beta \subseteq \alpha$ hence $\alpha = \beta$.

transitivity. If $\alpha \leqslant \beta$ and $\beta \leqslant \gamma$ then by [theorem: 9.41] we have $\alpha \subseteq \beta \wedge \beta \subseteq \gamma$ so that $\alpha \subseteq \gamma$ which by [theorem: 9.41] proves that $\alpha \leqslant \gamma$

totally ordering. Let $\alpha, \beta \in \mathbb{R}$ then for $\alpha + (-\beta)$ we have, as $\mathbb{R} \underset{[\text{theorem: } 9.17]}{=} \mathbb{R}^+ \bigcup \mathbb{R}^- \bigcup \{0\}$, either:

$\boldsymbol{\alpha + (-\beta) \in \mathbb{R}^+}$. Then $\alpha < \beta \Rightarrow \alpha \leqslant \beta$

$\boldsymbol{\alpha + (-\beta) \in \mathbb{R}^-}$. Then $\beta + (-\alpha) = -(\alpha + (-\beta)) \in \mathbb{R}^+$ so that $\beta < \alpha \Rightarrow \beta \leqslant \alpha$

$\boldsymbol{\alpha + (-\beta) = 0}$. Then $\alpha = \beta$ hence $\alpha \leqslant \beta$ □

Theorem 9.44. *We have the following for* $\mathbb{R}$

1. $\mathbb{R}^+ = \{\alpha \in \mathbb{R} | 0 < \alpha\}$
2. $\mathbb{R}_0^+ = \{\alpha \in \mathbb{R} | 0 \leqslant \alpha\}$
3. $\mathbb{R}^- = \{\alpha \in \mathbb{R} | \alpha < 0\}$
4. $\mathbb{R}_0^- = \{\alpha \in \mathbb{R} | \alpha \leqslant 0\}$

Proof.

1. $\alpha \in \mathbb{R}^+ \Leftrightarrow \alpha + (-0) \in \mathbb{R}^+ \Leftrightarrow 0 < \alpha \Leftrightarrow \alpha \in \{\alpha \in \mathbb{R} | 0 < \alpha\}$
2. $\alpha \in \mathbb{R}_0^+ \Leftrightarrow \alpha = 0 \vee \alpha + (-0) \in \mathbb{R}^+ \Leftrightarrow \alpha = 0 \vee 0 < \alpha \Leftrightarrow 0 \leqslant \alpha \Leftrightarrow \alpha \in \{\alpha \in \mathbb{R} | 0 \leqslant \alpha\}$.
3. $\alpha \in \mathbb{R}^- \Leftrightarrow -\alpha \in \mathbb{R}^+ \Leftrightarrow 0 + (-\alpha) \in \mathbb{R}^+ \Leftrightarrow \alpha < 0 \Leftrightarrow \alpha \in \{\alpha \in \mathbb{R} | \alpha < 0\}$
4. $\alpha \in \mathbb{R}_0^- \Leftrightarrow \alpha = 0 \vee -\alpha \in \mathbb{R}^- \Leftrightarrow \alpha = 0 \vee 0 + (-\alpha) \in \mathbb{R}^+ \Leftrightarrow \alpha = 0 \vee \alpha < 0 \Leftrightarrow \alpha \leqslant 0 \Leftrightarrow \alpha \in \{\alpha \in \mathbb{R} | \alpha \leqslant 0\}$ □

Theorem 9.45. $\langle \mathbb{R}, +, \cdot, \leqslant \rangle$ *is a ordered field*

Proof. First using [theorem: 9.32] we have that $\langle \mathbb{R}, +, \cdot \rangle$ is a field. Second we have

1. If $x, y, z \in \mathbb{R}$ with $x < y$ then

$$y + (-x) \in \mathbb{R}^+, \tag{9.81}$$

further

$$\begin{array}{rcl}(y+z)+(-(x+z)) & \underset{[\text{theorem: } 4.8]}{=} & (y+z)+((-x)+(-z)) \\ & = & (y+z)+((-z)+(-x)) \\ & = & y+(z+((-z)+(-x))) \\ & = & y+((z+(-z))+(-x)) \\ & = & y+(0+(-x)) \\ & = & y+(-x)\end{array}$$

which by [eq: 9.81] proves that $(y+z)+(-(x+z))\in\mathbb{R}^+$ or

$$x+z<y+z$$

2. If $x,y\in\mathbb{R}$ with $0<x$ and $0<y$ then by [theorem: 9.44] $x,y\in\mathbb{R}^+$. So by [theorem: 9.38] we have that $x\cdot y\in\mathbb{R}^+$ which by [theorem: 9.44 proves that

$$0<x\cdot y$$

□

Corollary 9.46. *Let $\alpha\in\mathbb{R}$*

1. *$\forall\alpha\in\mathbb{R}$ we have $\alpha<\alpha+1$*
2. *$\alpha-1<\alpha$*

Proof.

1. *As $0<1$ [see corollary: 9.42] we have by [theorems: 9.45, 4.73] that $\alpha=0+\alpha<1+\alpha=\alpha+1$*
2. *As $0<1$ we have by [theorems: 9.45, 4.73] that $-1<0$ hence using [theorems: 9.45, 4.73] again we have $\alpha-1=(-1)+\alpha<0+\alpha=\alpha$.*

□

Theorem 9.47. *If $\alpha,\beta\in\mathbb{R}_0^+$ [so that $0\leqslant\alpha\wedge 0\leqslant\beta$] such that $\alpha+\beta=0$ then $\alpha=0=\beta$*

Proof. As $0\leqslant\alpha$, $0\leqslant\beta$ then we have either

$0<\alpha$. Then $\beta=0+\beta<\alpha+\beta=0$ hence $\beta<0$ contradicting $0\leqslant\beta$, so this case does not occur.

$0<\beta$. Then $\alpha=\alpha+0<\alpha+\beta=0$ hence $\alpha<0$ contradicting $0\leqslant\alpha$, so this case does not occur.

$\alpha=\beta=0$. This is the only resting case proving that $\alpha=\beta=0$ □

Lemma 9.48. *Let $r,s\in\mathbb{Q}$ then we have*

1. *$r<s\Leftrightarrow\alpha_r<\alpha_s$*
2. *$r\leqslant s\Leftrightarrow\alpha_r<\alpha_s$*

Proof.

1.

$\Rightarrow$. If $x\in\alpha_r$ then $x<r$ which as $r<s$ proves that $x<s$ hence $x\in\alpha_s$, so $\alpha_r\subseteq\alpha_s$. Further as $r<s$ we have by the density of $\mathbb{Q}$ [see theorem: 8.37] that there exists a $q\in\mathbb{Q}$ such that $r<q<s$ hence $q\in\alpha_s$ and $q\notin r$, proving that $\alpha_r\subset\alpha_s$. By [theorem: 9.41] it follows then that

$$\alpha_r<\alpha_s$$

$\Leftarrow$. If $\alpha_r<\alpha_s$ then we have by [theorem: 9.41] that $\alpha_r\subset\alpha_s$. Assume that $s\leqslant r$ then $\forall t\in\alpha_s$ we have $t<s\leqslant r\Rightarrow t<r\Rightarrow t\in\alpha_r$ proving that $\alpha_s\subseteq\alpha_r$ contradicting $\alpha_r\subset\alpha_s$. As the assumption $s\leqslant r$ leads to a contradiction we must have that $r<s$.

2.

$$\begin{array}{rcl}r\leqslant s & \Leftrightarrow & r=s\vee r<s \\ & \Leftrightarrow & \alpha_r=\alpha_s\vee r<s \\ & \underset{(1)}{\Leftrightarrow} & \alpha_r=a_s\vee\alpha_r<\alpha_s \\ & \Leftrightarrow & \alpha_r\leqslant\alpha_s\end{array}$$

□

The above lemma allows us to show that the embedding of $\mathbb{Q}$ in $\mathbb{R}$ by $i_{\mathbb{Q}\rightarrow\mathbb{R}}$ is not only preserving the field structure but also the order.

Theorem 9.49. *The field isomorphism $i_{\mathbb{Q}\rightarrow\mathbb{R}}: \langle\mathbb{Q}, +, \cdot\rangle \rightarrow \langle\mathbb{Q}_{\mathbb{R}}, +, \cdot\rangle$ defined by $i_{\mathbb{Q}\rightarrow\mathbb{R}}(r) = \alpha_r$ [see theorem: 9.37] is a order isomorphism between $\langle\mathbb{Q}, \leqslant\rangle$ and $\langle\mathbb{Q}_{\mathbb{R}}, \leqslant\rangle$*

Proof. Using [theorem: 9.37] we have that $i_{\mathbb{Q}\rightarrow\mathbb{R}}: \mathbb{Q} \rightarrow \mathbb{Q}_{\mathbb{R}}$ is a bijection. Further for $r, s \in \mathbb{Q}$ we have

$$\begin{aligned} r \leqslant s &\underset{[\text{theorem: } 9.48]}{\Leftrightarrow} \alpha_r \leqslant \alpha_s \\ &\Leftrightarrow i_{\mathbb{Q}\rightarrow\mathbb{R}}(r) \leqslant i_{\mathbb{Q}\rightarrow\mathbb{R}}(s) \end{aligned}$$

□

We have seen in [theorem: 8.41] that the rational numbers are not conditional complete [causing $\langle\mathbb{Q}_{\mathbb{R}}, \leqslant\rangle$ to be not conditional complete], the prime reason that we construct the real numbers is that the real numbers are conditional complete.

Theorem 9.50. *$\langle\mathbb{R}, \leqslant\rangle$ is conditional complete [definition: 3.77] in other words*

$$\forall S \subseteq \mathbb{R} \textit{ with } S \neq \varnothing \textit{ such that } \exists \upsilon \in \mathbb{R} \textit{ such that } \forall \alpha \in S \textit{ we have } \alpha \leqslant \upsilon \textit{ we have that } \sup(S) \textit{ exist}$$

Using [theorem: 3.78] this is equivalent with

$$\forall S \subseteq \mathbb{R} \textit{ with } S \neq \varnothing \textit{ such that } \exists \lambda \in \mathbb{R} \textit{ such that } \forall \alpha \in S \textit{ we have } \lambda \leqslant \alpha \textit{ we have that } \inf(S) \textit{ exist}$$

Proof. Let $S \subseteq \mathbb{R}$ with $S \neq 0$ such that there exists a $\upsilon \in S$ such that $\forall \alpha \in S$ we have $\alpha \leqslant \upsilon$. Define γ by

$$\gamma = \{q \in \mathbb{Q} | \exists \alpha \in S | q \in \alpha\}$$

First we prove that $\gamma \in \mathbb{R}$ [or γ is a Dedekind cut]

1. As $S \neq \varnothing$ there exist a $\alpha \in S \subseteq \mathbb{R}$. As α is a Dedekind cut we have by [definition: 9.1 (1)] that $\alpha \neq \varnothing$. Hence $\exists q \in \alpha \subseteq \mathbb{Q}$ so that $q \in \gamma$, proving that

$$\gamma \neq \varnothing$$

2. If $r \in \gamma$ then $\exists \alpha \in S$ such that $r \in \alpha$, as υ is a upper bound of S we have that $\alpha \leqslant \upsilon$, so using 9.41] $\alpha \subseteq \upsilon$ proving that $r \in \upsilon$ so that $\gamma \subseteq \upsilon$. As $\upsilon \in \mathbb{R}$ we have by [definition: 9.1 (2)] that $\upsilon \neq \mathbb{Q}$ so that $\exists q \in \mathbb{Q}$ such that $q \notin \upsilon$, which as $\gamma \subseteq \upsilon$ proves that $q \notin \gamma$. Hence

$$\gamma \neq \mathbb{Q}$$

3. Let $r \in \gamma$ and $s \in \mathbb{Q} \setminus \gamma$. As $r \in \gamma$ there exists a $\alpha \in S$ such that $r \in \alpha$ and as $s \in \mathbb{Q} \setminus \gamma$ we have that $\forall \zeta \in S$ we have $s \notin \zeta$, so in particular $s \notin \alpha$ hence $s \in \mathbb{Q} \setminus \alpha$. Using [definition: 9.1 (3)] we have that $r < s$. So

$$\text{If } r \in \gamma \wedge s \in \mathbb{Q} \setminus \gamma \text{ then } r < s$$

4. Assume that γ has a greatest element m then

$$m \in \gamma \text{ and } \forall r \in \gamma \text{ we have } r \leqslant m \tag{9.82}$$

Now as $m \in \gamma$ there exist a $\alpha \in S$ such that $m \in \alpha$. As by [definition: 9.1 (4)] α has no greatest element there exist a $s \in \alpha$ such that $m < s$. As $s \in \alpha \in S$ it follows that $s \in \gamma$ so by [eq: 9.82] we must have that $s \leqslant m$ contradicting $m < s$. So the assumption is wrong and we have

$$\gamma \text{ has no greatest element}$$

From (1),(2),(3) and (4) we conclude that γ is a Dedekind cut, hence

$$\gamma \in \mathbb{R}$$

Next we proof that γ is a upper bound of S. So let $\alpha \in S$ then if $q \in \alpha$ we have by definition that $q \in \gamma$ proving that $\alpha \subseteq \gamma$ which by [theorem: 9.41] results in $\alpha \leqslant \gamma$. Hence

$$\gamma \text{ is a upper bound of S}$$

Finally let $\lambda \in \upsilon(S) = \{\alpha \in \mathbb{R} | \alpha \text{ is a upper bound of } S\}$. If $q \in \gamma$ there exist a $\alpha \in \mathcal{S}$ such that $q \in \alpha$, as λ is a upper bound of S we have $\alpha \leqslant \lambda \underset{[\text{theorem: } 9.41]}{\Rightarrow} \alpha \subseteq \lambda$, so $q \in \lambda$, proving that $\gamma \subseteq \lambda$ or by [theorem: 9.41] that $\gamma \leqslant \lambda$. Hence γ is the least element of $\upsilon(S)$ which by definition proves that

$$\sup(S) \text{ exist}$$

□

9.3 Embeddings in $\mathbb{R}$

First remember that by [theorems: 9.37,9.49] we have a embedding of $\mathbb{Q}$ in $\mathbb{R}$ by the order and field isomorphism $i_{\mathbb{Q}\rightarrow\mathbb{R}}: \mathbb{Q}\rightarrow\mathbb{Q}_{\mathbb{R}}$ defined by $i_{\mathbb{Q}\rightarrow\mathbb{R}}(r)=\alpha_r$. We show now that there exist also embeddings of $\mathbb{N}_0$ and $\mathbb{Z}$ in $\mathbb{R}$.

Definition 9.51. $\mathbb{Z}_{\mathbb{R}}=(i_{\mathbb{Q}\rightarrow\mathbb{R}}\circ i_{\mathbb{Z}\rightarrow\mathbb{Q}})(\mathbb{Z})$ *where*

$$i_{\mathbb{Z}\rightarrow\mathbb{Q}}: \mathbb{Z}\rightarrow\mathbb{Z}_{\mathbb{Q}}\subseteq\mathbb{Q} \text{ is defined by } i_{\mathbb{Z}\rightarrow\mathbb{Q}}(z)=\frac{z}{1} \text{ [theorem: 8.32]}$$

and

$$i_{\mathbb{Q}\rightarrow\mathbb{R}}: \mathbb{Q}\rightarrow\mathbb{Q}_{\mathbb{R}} \text{ is defined by } i_{\mathbb{Q}\rightarrow\mathbb{R}}(r)=\alpha_r \text{ [theorem: 9.49]}$$

so that

$$\mathbb{Z}_{\mathbb{R}}\subseteq\mathbb{Q}_{\mathbb{R}}$$

Theorem 9.52. *For* $\langle\mathbb{Z}_{\mathbb{R}}.+,\cdot\rangle$ *and* $i_{\mathbb{Z}\rightarrow\mathbb{R}}: \mathbb{Z}\rightarrow\mathbb{Z}_{\mathbb{R}}$ *defined by* $i_{\mathbb{Z}\rightarrow\mathbb{R}}=i_{\mathbb{Z}\rightarrow\mathbb{Q}}\circ i_{\mathbb{Q}\rightarrow\mathbb{R}}$ *we have*

1. $\mathbb{Z}_{\mathbb{R}}$ *is a sub ring of* $\langle\mathbb{R},_{,+,\cdot}\rangle$ *and* $i_{\mathbb{Z}\rightarrow\mathbb{R}}: \langle\mathbb{Z},+,\cdot\rangle\rightarrow\langle\mathbb{Z}_{\mathbb{R}},+,\cdot\rangle$ *is a ring isomorphism.*
2. $\mathbb{Z}_{\mathbb{R}}$ *is a sub group of* $\langle\mathbb{R},+\rangle$ *and* $i_{\mathbb{Z}\rightarrow\mathbb{R}}: \langle\mathbb{Z},+\rangle\rightarrow\langle\mathbb{Z}_{\mathbb{R}},+\rangle$ *is a group isomorphism.*
3. $\mathbb{Z}_{\mathbb{R}}$ *is a sub semi-group of* $\langle\mathbb{R},\cdot\rangle$ *and* $i_{\mathbb{Z}\rightarrow\mathbb{R}}: \langle\mathbb{Z},\cdot\rangle\rightarrow\langle\mathbb{Z}_{\mathbb{R}},\cdot\rangle$ *is a group isomorphism.*
4. $i_{\mathbb{Z}\rightarrow\mathbb{R}}: \langle\mathbb{Z},\leqslant\rangle\rightarrow\langle\mathbb{Z}_{\mathbb{R}},\leqslant\rangle$ *is a order isomorphism*
5. $\mathbb{Z}_{\mathbb{R}}$ *is denumerable*

Proof. First note that by the definition of $\mathbb{Z}_{\mathbb{R}}$ we have that

$$\mathbb{Z}_{\mathbb{R}}=i_{\mathbb{Q}\rightarrow\mathbb{R}}(i_{\mathbb{Z}\rightarrow\mathbb{Q}}(\mathbb{Z})) \tag{9.83}$$

Second we have

1. Using [theorems: 8.32, 9.49] we have that

   $$i_{\mathbb{Z}\rightarrow\mathbb{Q}}: \langle\mathbb{Z},+\rangle\rightarrow\langle\mathbb{Q},+\rangle \text{ and } i_{\mathbb{Q}\rightarrow\mathbb{R}}: \langle\mathbb{Q},+\rangle\rightarrow\langle\mathbb{Q}_{\mathbb{R}},+\rangle \text{ are ring isomorphism}$$

   So using [theorem: 4.48] and [theorem: 8.32] we have that

   $$i_{\mathbb{Q}\rightarrow\mathbb{R}}(i_{\mathbb{Z}\rightarrow\mathbb{Q}}(\mathbb{Z})) \text{ is a sub-ring of } \langle\mathbb{Q}_{\mathbb{R}},+,\cdot\rangle$$

   and

   $$i_{\mathbb{Q}-\mathbb{R}}\circ i_{\mathbb{Z}\rightarrow\mathbb{Q}}: \langle\mathbb{Z},\leqslant\rangle\rightarrow\langle i_{\mathbb{Q}\rightarrow\mathbb{R}}(i_{\mathbb{Z}\rightarrow\mathbb{Q}}(\mathbb{Z})),+,\cdot\rangle \text{ is a ring isomorphism}$$

   Using [eq: 9.83] and [theorem: 4.38, 9.37] we have that

   $$\mathbb{Z}_{\mathbb{R}} \text{ is a sub-ring of } \langle\mathbb{R},+,\cdot\rangle$$

   and

   $$i_{\mathbb{Z}\rightarrow\mathbb{R}}=i_{\mathbb{Q}\rightarrow\mathbb{R}}\circ i_{\mathbb{Z}\rightarrow\mathbb{Q}}: \langle\mathbb{Z}+,\cdot\rangle\rightarrow\langle\mathbb{Z}_{\mathbb{R}},+,\cdot\rangle \text{ is a ring isomorphism}$$

2. This follows from (1) and [theorems: 4.36 and 4.47]
3. This follows from (1) and [theorems: 4.36 and 4.47]
4. Using [theorems: 8.32, 9.49] we have that

   $$i_{\mathbb{Z}\rightarrow\mathbb{Q}}: \langle\mathbb{Z},\leqslant\rangle\rightarrow\langle\mathbb{Z}_{\mathbb{Q}},\leqslant\rangle \text{ and } i_{\mathbb{Q}\rightarrow\mathbb{R}}: \langle\mathbb{Q},\leqslant\rangle\rightarrow\langle\mathbb{Q}_{\mathbb{R}},\leqslant\rangle \text{ are order isomorphisms}$$

   So using [theorem: 3.52] we have that

   $$i_{\mathbb{Q}\rightarrow\mathbb{R}}\circ i_{\mathbb{Z}\rightarrow\mathbb{Q}}: \langle\mathbb{Z},\leqslant\rangle\rightarrow\langle i_{\mathbb{Q}\rightarrow\mathbb{R}}(i_{\mathbb{Z}\rightarrow\mathbb{Q}}(\mathbb{Z})),\leqslant\rangle \text{ is a order isomorphism}$$

   hence using [eq: 9.83] we have that

   $$i_{\mathbb{Z}\rightarrow\mathbb{R}}=i_{\mathbb{Q}\rightarrow\mathbb{R}}\circ i_{\mathbb{Z}\rightarrow\mathbb{Q}}: \langle\mathbb{Z},\leqslant\rangle\rightarrow\langle\mathbb{Z}_{\mathbb{R}},\leqslant\rangle \text{ is a order isomorphism}$$

5. Using (4) we have that $Z\approx Z_{\mathbb{R}}$ which as by [theorem: 7.53] $\mathbb{N}_0\approx\mathbb{Z}$ proves that $\mathbb{N}_0\approx\mathbb{Z}_R$ proving that $\mathbb{Z}_{\mathbb{R}}$ is denumerable. ☐

We can use the same technique to embed the set of natural numbers in $\mathbb{R}$.

Definition 9.53. $\mathbb{N}_{0,\mathbb{R}}=(i_{\mathbb{Z}\rightarrow\mathbb{R}}\circ i_{\mathbb{N}_0\rightarrow\mathbb{Z}})(\mathbb{N}_0)$ *where*

$$i_{\mathbb{N}_0}:\mathbb{N}_0\rightarrow\mathbb{Z}_0^+\subseteq\mathbb{Z} \textit{ is defined by } i_{\mathbb{N}_0}(z)=\sim[(z,0)] \textit{ [theorem: 7.17]} \text{ and}$$

$$i_{\mathbb{Z}\rightarrow\mathbb{R}}:\mathbb{Z}\rightarrow\mathbb{Z}_{\mathbb{R}} \textit{ [theorem: 9.52]}$$

so that

$$\mathbb{N}_{0,\mathbb{R}}\subseteq\mathbb{Z}_{\mathbb{R}}$$

In many cases we want to exclude 0 from the embedded real numbers, hence the following definition.

Definition 9.54. $\mathbb{N}_{\mathbb{R}}=\mathbb{N}_{0,\mathbb{R}}\setminus\{0\}$

Theorem 9.55. *For* $\langle\mathbb{N}_{0,\mathbb{R}}.+,\cdot\rangle$ *and* $i_{\mathbb{N}_0\rightarrow\mathbb{R}}:\mathbb{N}_0\rightarrow\mathbb{N}_{0,\mathbb{R}}$ *defined by* $i_{\mathbb{N}_0\rightarrow\mathbb{R}}=i_{\mathbb{Z}\rightarrow\mathbb{R}}\circ i_{\mathbb{N}_0\rightarrow\mathbb{Z}}$ *we have*

1. $\mathbb{N}_{0,\mathbb{R}}$ *is a sub-semi-group of* $\langle\mathbb{R},+\rangle$ *and* $i_{\mathbb{N}_0\rightarrow\mathbb{R}}:\langle\mathbb{N}_0,+\rangle\rightarrow\langle\mathbb{N}_{0,\mathbb{R}},+\rangle$ *is a group isomorphism.*
2. $\mathbb{N}_{0,\mathbb{R}}$ *is a sub-semi-group of* $\langle\mathbb{R},\cdot\rangle$ *and* $i_{\mathbb{N}_0\rightarrow\mathbb{R}}:\langle\mathbb{N}_0,\cdot\rangle\rightarrow\langle\mathbb{N}_{0,\mathbb{R}},\cdot\rangle$ *is a group isomorphism.*
3. $i_{\mathbb{N}_0\rightarrow\mathbb{R}}:\langle\mathbb{N}_0,\leqslant\rangle\rightarrow\langle\mathbb{N}_{0,\mathbb{R}},\leqslant\rangle$ *is a order isomorphism.*
4. $\mathbb{N}_{0,\mathbb{R}}$ *is denumerable.*
5. $\langle\mathbb{N}_{0,\mathbb{R}},\leqslant\rangle$ *is well ordered.*

Proof. First note that by the definition of $\mathbb{N}_{0,\mathbb{R}}$ we have that

$$\mathbb{N}_{0,\mathbb{R}}=i_{\mathbb{Z}\rightarrow\mathbb{R}}(i_{\mathbb{N}_0}(\mathbb{N}_0)) \tag{9.84}$$

1. Using [theorems: 7.17 and 9.52] we have that

   $$i_{\mathbb{N}_0\rightarrow\mathbb{Z}}:\langle\mathbb{N}_0,+\rangle\rightarrow\langle\mathbb{Z}_0^+,+\rangle \text{ and } i_{\mathbb{Z}\rightarrow\mathbb{R}}:\langle\mathbb{Z},+\rangle\rightarrow\langle\mathbb{Z}_{\mathbb{R}},+\rangle \text{ are a group isomorphisms}$$

   So using [theorem: 4.25] and [theorem: 7.17] we have that

   $$i_{\mathbb{Z}\rightarrow\mathbb{R}}(i_{\mathbb{N}_0\rightarrow\mathbb{Z}}(\mathbb{N}_0)) \text{ is a sub group of } \langle\mathbb{Z}_{\mathbb{R}},+\rangle$$

   and

   $$i_{\mathbb{N}_0\rightarrow\mathbb{R}}=i_{\mathbb{Z}\rightarrow\mathbb{R}}\circ i_{\mathbb{N}_0\rightarrow\mathbb{Z}}:\langle\mathbb{N}_0,+\rangle\rightarrow\langle i_{\mathbb{Z}\rightarrow\mathbb{R}}(i_{\mathbb{N}_0\rightarrow\mathbb{Z}}(\mathbb{N}_0)),+\rangle \text{ is a group isomorphism}$$

   Using [eq: 9.84] and [theorems: 4.17, 9.52] it follows that

   $$\mathbb{N}_{0,\mathbb{R}} \text{ is a sub semi-group of } \langle\mathbb{R},+\rangle$$

   and

   $$i_{\mathbb{N}_0\rightarrow\mathbb{R}}=i_{\mathbb{Z}\rightarrow\mathbb{R}}\circ i_{\mathbb{N}_0\rightarrow\mathbb{Z}}:\langle\mathbb{N}_0,+\rangle\rightarrow\langle\mathbb{N}_{0,\mathbb{R}},+\rangle \text{ is a group isomorphism}$$

2. Using [theorems: 7.17 and 9.52] we have that

   $$i_{\mathbb{N}_0\rightarrow\mathbb{Z}}:\langle\mathbb{N}_0,\cdot\rangle\rightarrow\langle\mathbb{Z}_0^+,\cdot\rangle \text{ and } i_{\mathbb{Z}\rightarrow\mathbb{R}}:\langle\mathbb{Z},\cdot\rangle\rightarrow\langle\mathbb{Z}_{\mathbb{R}},\cdot\rangle \text{ are group isomorphisms}$$

   So using [theorem: 4.25] and [theorem: 7.17] we have that

   $$i_{\mathbb{Z}\rightarrow\mathbb{R}}(i_{\mathbb{N}_0\rightarrow\mathbb{Z}}(\mathbb{N}_0)) \text{ is a sub semi-group of } \langle\mathbb{Z}_{\mathbb{R}},\cdot\rangle$$

   and

   $$i_{\mathbb{N}_0\rightarrow\mathbb{R}}=i_{\mathbb{Z}\rightarrow\mathbb{R}}\circ i_{\mathbb{N}_0\rightarrow\mathbb{Z}}:\langle\mathbb{N}_0,\cdot\rangle\rightarrow\langle i_{\mathbb{Z}\rightarrow\mathbb{R}}(i_{\mathbb{N}_0\rightarrow\mathbb{Z}}(\mathbb{N}_0)),\cdot\rangle \text{ is a group isomorphism}$$

   Using [eq: 9.84] and [theorems: 4.17, 9.52] it follows that

   $$\mathbb{N}_{0,\mathbb{R}} \text{ is a sub semi-group of } \langle\mathbb{R},+\rangle$$

   and

   $$i_{\mathbb{N}_0\rightarrow\mathbb{R}}=i_{\mathbb{Z}\rightarrow\mathbb{R}}\circ i_{\mathbb{N}_0\rightarrow\mathbb{Z}}:\langle\mathbb{N}_0,\cdot\rangle\rightarrow\langle\mathbb{N}_{0,\mathbb{R}},\cdot\rangle \text{ is a group isomorphism}$$

3. Using [theorems: 7.29 and 9.52] we have that

   $$i_{\mathbb{N}_0\rightarrow\mathbb{Z}}:\langle\mathbb{N}_0,\leqslant\rangle\rightarrow\langle\mathbb{Z}_0^+,\leqslant\rangle \text{ and } i_{\mathbb{Z}\rightarrow\mathbb{R}}:\langle\mathbb{Z},\leqslant\rangle\rightarrow\langle\mathbb{Z}_{\mathbb{R}}.\leqslant\rangle$$

so using [theorem: 3.52]

$$i_{\mathbb{N}_0\rightarrow\mathbb{R}} = i_{\mathbb{Z}\rightarrow\mathbb{R}} \circ i_{\mathbb{N}_0\rightarrow\mathbb{Z}}: \langle\mathbb{N}_0, \leqslant\rangle \rightarrow \langle i_{\mathbb{Z}\rightarrow\mathbb{R}}(i_{\mathbb{N}_0\rightarrow\mathbb{Z}}(\mathbb{N}_0)), \leqslant\rangle \text{ is a order isomorphism}$$

Using [eq: 9.84] it follows that

$$i_{\mathbb{N}_0\rightarrow\mathbb{R}} = i_{\mathbb{Z}\rightarrow\mathbb{R}} \circ i_{\mathbb{N}_0\rightarrow\mathbb{Z}}: \langle\mathbb{N}_0, \leqslant\rangle \rightarrow \langle\mathbb{N}_{0,\mathbb{R}}, \leqslant\rangle \text{ is a order isomorphism}$$

4. Using (3) we have that $\mathbb{N}_0 \approx \mathbb{N}_{0,\mathbb{R}}$ proving that $\mathbb{N}_{0,\mathbb{R}}$ is denumerable.
5. By [theorem: 5.51] $\langle\mathbb{N}_0, \leqslant\rangle$ is well ordered, further by (3) $i_{\mathbb{N}_0\rightarrow\mathbb{R}}: \langle\mathbb{N}_0, \leqslant\rangle \rightarrow \langle\mathbb{N}_{0,\mathbb{R}}, \leqslant\rangle$ is a order isomorphism, hence by [theorem: 3.82] it follows that

$$\langle\mathbb{N}_{0,\mathbb{R}}, \leqslant\rangle \text{ is well ordered}$$ □

For the relation between $\mathbb{N}_{0,\mathbb{R}}$, $\mathbb{Z}_\mathbb{R}$, $\mathbb{Q}_\mathbb{R}$ and $\mathbb{R}$ we have

Theorem 9.56. *We have the following relation between the embbeddings of* $\mathbb{N}_0, \mathbb{Z}, \mathbb{Q}$ *in* $\mathbb{R}$

$$\mathbb{N}_{0,\mathbb{R}} \subseteq \mathbb{Z}_\mathbb{R} \subseteq \mathbb{Q}_\mathbb{R} \subseteq \mathbb{R}$$

Proof. Using [definition: 9.6] we have that $\mathbb{Q}_\mathbb{R} \subseteq \mathbb{R}$, using [definition: 9.51] we have that $\mathbb{Z}_\mathbb{R} \subseteq \mathbb{Q}_\mathbb{R}$ and finally by [definition: 9.53] it follows that $\mathbb{N}_{0,\mathbb{R}} \subseteq \mathbb{Z}_\mathbb{R}$ □

Finally we can define $\mathbb{Q}_\mathbb{R}$ in terms of $\mathbb{Z}$.

Theorem 9.57. $\mathbb{Q}_\mathbb{R} = \{x/y | x \in \mathbb{Z}_\mathbb{R} \wedge y \in \mathbb{Z}_\mathbb{R} \setminus \{0\}\}$ *where* q/r *is a shorthand for* $q \cdot r^{-1}$ *[see notation: 9.33]*

Proof. Using [theorems: 8.32, 9.37, 9.52] we have the following

$$i_{\mathbb{Z}\rightarrow\mathbb{Q}}: \langle\mathbb{Z}, +, \cdot\rangle \rightarrow \langle\mathbb{Z}_\mathbb{Q}, +, \cdot\rangle \text{ is a ring isomorphism} \tag{9.85}$$

$$i_{\mathbb{Q}\rightarrow\mathbb{R}}: \langle\mathbb{Q}, +, \cdot\rangle \rightarrow \langle\mathbb{Q}_\mathbb{R}, +, \cdot\rangle \text{ is a field isomorphism} \tag{9.86}$$

$$i_{\mathbb{Q}\rightarrow\mathbb{R}} \circ i_{\mathbb{Z}\rightarrow\mathbb{Q}}: \langle\mathbb{Z}, +, \cdot\rangle \rightarrow \langle\mathbb{Z}_\mathbb{R}, +, \cdot\rangle \text{ is a ring isomorphism} \tag{9.87}$$

Let $q \in \mathbb{Q}_\mathbb{R}$ then by [eq: 9.86] there exist a $q' \in \mathbb{Q}$ such that $q = i_{\mathbb{Q}\rightarrow\mathbb{R}}(q')$. As $q' \in \mathbb{Q}$ we have by [theorem: 8.33] that there exists $x \in \mathbb{Z}_\mathbb{Q}$ and $y \in \mathbb{Z}_\mathbb{Q} \setminus \{0\}$ such that $q' = x \cdot y^{-1}$, hence we have

$$\begin{aligned} q &= i_{\mathbb{Q}\rightarrow\mathbb{R}}(q') \\ &= i_{\mathbb{Q}\rightarrow\mathbb{R}}(x \cdot y^{-1}) \\ &= i_{\mathbb{Q}\rightarrow\mathbb{R}}(x) \cdot i_{\mathbb{Q}\rightarrow\mathbb{R}}(y^{-1}) \\ &= i_{\mathbb{Q}\rightarrow\mathbb{R}}(x) \cdot (i_{\mathbb{Q}\rightarrow\mathbb{R}}(y))^{-1} \\ &= i_{\mathbb{Q}\rightarrow\mathbb{R}}(x) / i_{\mathbb{Q}\rightarrow\mathbb{R}}(y) \end{aligned} \tag{9.88}$$

As $x \in \mathbb{Z}_\mathbb{Q}$ and $y \in \mathbb{Z}_\mathbb{Q} \setminus \{0\}$ there exists by [eq: 9.85] a $x' \in \mathbb{Z}$ and $y' \in \mathbb{Z}$ such that $x = i_{\mathbb{Z}\rightarrow\mathbb{Q}}(x')$ and $y = i_{\mathbb{Z}\rightarrow\mathbb{Q}}(y')$. So by [eq: 9.87] we have that $i_{\mathbb{Q}\rightarrow\mathbb{R}}(x) = (i_{\mathbb{Q}\rightarrow\mathbb{R}} \circ i_{\mathbb{Z}\rightarrow\mathbb{Q}})(x') \in \mathbb{Z}_R$ and $i_{\mathbb{Q}\rightarrow\mathbb{R}}(y) = (i_{\mathbb{Q}\rightarrow\mathbb{R}} \circ i_{\mathbb{Z}\rightarrow\mathbb{Q}})(y') \in \mathbb{Z}_\mathbb{R} \setminus \{0\}$. Combining this with [eq: 9.88] we have that $q \in \{x/y | x \in \mathbb{Z}_\mathbb{R} \wedge y \in \mathbb{Z}_\mathbb{R} \setminus \{0\}\}$ proving that

$$\mathbb{Q}_\mathbb{R} \subseteq \{x/y | x \in \mathbb{Z}_\mathbb{R} \wedge y \in \mathbb{Z}_\mathbb{R} \setminus \{0\}\} \tag{9.89}$$

If $q \in \{x/y | x \in \mathbb{Z}_\mathbb{R} \wedge y \in \mathbb{Z}_\mathbb{R} \setminus \{0\}\}$ then there exists $x \in \mathbb{Z}_\mathbb{R}$ and $y \in \mathbb{Z}_\mathbb{R} \setminus \{0\}$ such that $q = x/y$. Using [eq: 9.87] there exists $x' \in \mathbb{Z}$ and $z' \in \mathbb{Z} \setminus \{0\}$ such that $x = i_{\mathbb{Q}\rightarrow\mathbb{R}}(i_{\mathbb{Z}\rightarrow\mathbb{Q}}(x'))$ and $y = i_{\mathbb{Q}\rightarrow\mathbb{R}}(i_{\mathbb{Z}\rightarrow\mathbb{Q}}(y'))$. From [eq: 9.85] it follows that $i_{\mathbb{Z}\rightarrow\mathbb{Q}}(x') \in \mathbb{Z}_\mathbb{Q}$ and $i_{\mathbb{Z}\rightarrow\mathbb{Q}}(y') \in \mathbb{Z}_\mathbb{Q} \setminus \{0\}$ which by [theorem: 8.33] gives $i_{\mathbb{Z}\rightarrow\mathbb{Q}}(x') \cdot (i_{\mathbb{Z}\rightarrow\mathbb{Q}}(y'))^{-1} \in \mathbb{Q}$. So by [eq: 9.86]

$$i_{\mathbb{Q}\rightarrow\mathbb{R}}(i_{\mathbb{Z}\rightarrow\mathbb{Q}}(x') \cdot (i_{\mathbb{Z}\rightarrow\mathbb{Q}}(y'))^{-1}) \in \mathbb{Q}_\mathbb{R} \tag{9.90}$$

Further

$$\begin{aligned} i_{\mathbb{Q}\rightarrow\mathbb{R}}(i_{\mathbb{Z}\rightarrow\mathbb{Q}}(x') \cdot (i_{\mathbb{Z}\rightarrow\mathbb{Q}}(y'))^{-1}) &= i_{\mathbb{Q}\rightarrow\mathbb{R}}(i_{\mathbb{Z}\rightarrow\mathbb{Q}}(x')) \cdot i_{\mathbb{Q}\rightarrow\mathbb{R}}((i_{\mathbb{Z}\rightarrow\mathbb{Q}}(y'))^{-1}) \\ &= i_{\mathbb{Q}\rightarrow\mathbb{R}}(i_{\mathbb{Z}\rightarrow\mathbb{Q}}(x')) \cdot (i_{\mathbb{Q}\rightarrow\mathbb{R}}(i_{\mathbb{Z}\rightarrow\mathbb{Q}}(y')))^{-1} \\ &= x \cdot y^{-1} \\ &= q \end{aligned}$$

proving by [eq: 9.90] that $q \in \mathbb{Q}_\mathbb{R}$. Hence $\{x/y | x \in \mathbb{Z}_\mathbb{R} \wedge y \in \mathbb{Z}_\mathbb{R} \setminus \{0\}\} \subseteq \mathbb{Q}_\mathbb{R}$ which combined with [eq: 9.89] proves finally that

$$\mathbb{Q}_\mathbb{R} = \{x/y | x \in \mathbb{Z}_\mathbb{R} \wedge y \in \mathbb{Z}_\mathbb{R} \setminus \{0\}\}$$ □

Chapter 10

The complex numbers

One problem that exist in the set of real numbers is that the equation $x^2 = -1$ has no solution because $0 < 1 \Rightarrow -1 < 0$ and by [theorems: 9.45, 4.73] $0 \leqslant x^2$. This problem wil be solved by introducing the set of complex numbers. Note that to avoid having to use different symbols for neutral elements, inverse elements, sum, product etc. we use context to derive the meaning of the different symbols.

Context	Expression	Operator
$n, m \in \mathbb{N}_0$	n+m	sum in $\langle \mathbb{N}_0, + \rangle$
$n, m \in \mathbb{N}_0$	$n \cdot m$	product in $\langle \mathbb{N}_0, \cdot \rangle$
$n, m \in \mathbb{N}_0$	$n \leqslant m$	order in $\langle \mathbb{N}_0, \leqslant \rangle$
$n, m \in \mathbb{N}_0$	$n < m$	strict order in $\langle \mathbb{N}_0, \leqslant \rangle$
$n, m \in \mathbb{N}_0$	$n - m$	subtraction in $\langle \mathbb{N}_0, + \rangle$
$n \in \mathbb{N}_0$	$n + 0$ or $0 + n$	neutral element in $\langle \mathbb{N}_0, + \rangle$
$n \in \mathbb{N}_0$	$n \cdot 1$ or $1 \cdot n$	neutral element in $\langle \mathbb{N}_0, \cdot \rangle$
$n \in \mathbb{N}_0$	$-n$	inverse element in $\langle \mathbb{N}_0, + \rangle$
$n, m \in \mathbb{Z}$	n+m	sum in $\langle \mathbb{Z}, + \rangle$
$n, m \in \mathbb{Z}$	$n \cdot m$	product in $\langle \mathbb{Z}, \cdot \rangle$
$n, m \in \mathbb{Z}$	$n \leqslant m$	order in $\langle \mathbb{Z} \leqslant \rangle$
$n, m \in \mathbb{Z}$	$n < m$	strict order in $\langle \mathbb{Z}, \leqslant \rangle$
$n, m \in \mathbb{Z}$	$n - m$	subtraction in $\langle \mathbb{Z}, - \rangle$
$n \in \mathbb{Z}$	$n + 0$ or $0 + n$	neutral element in $\langle \mathbb{Z}, + \rangle$
$n \in \mathbb{Z}$	$n \cdot 1$ or $1 \cdot n$	neutral element in $\langle \mathbb{Z}, \cdot \rangle$
$n \in \mathbb{Z}$	$-n$	inverse element in $\langle \mathbb{Z}, + \rangle$
$q, r \in \mathbb{Q}$	q+r	sum in $\langle \mathbb{Q}, + \rangle$
$q, r \in \mathbb{Q}$	$q \cdot r$	product in $\langle \mathbb{Q}, \cdot \rangle$
$q, r \in \mathbb{Q}$	$q \leqslant r$	order in $\langle \mathbb{Q} \leqslant \rangle$
$q, r \in \mathbb{Q}$	$q < r$	strict order in $\langle \mathbb{Q}, \leqslant \rangle$
$q, e \in \mathbb{Q}$	$q - r$	subtraction in $\langle \mathbb{Q}, - \rangle$
$q, r \in \mathbb{Q}$	q / r	division in $\langle \mathbb{Q}, \cdot \rangle$
$q \in \mathbb{Q}$	$q + 0$ or $0 + q$	neutral element in $\langle \mathbb{Q}, + \rangle$
$q \in \mathbb{Q}$	$q \cdot 1$ or $1 \cdot q$	neutral element in $\langle \mathbb{Q}, \cdot \rangle$
$q \in \mathbb{Q}$	$-q$	inverse element in $\langle \mathbb{Q}, + \rangle$
$q, r \in \mathbb{R}$	q+r	sum in $\langle \mathbb{R}, + \rangle$
$q, r \in \mathbb{R}$	$q \cdot r$	product in $\langle \mathbb{R}, \cdot \rangle$
$q, r \in \mathbb{R}$	$q \leqslant r$	order in $\langle \mathbb{R}, \leqslant \rangle$
$q, r \in \mathbb{R}$	$q < r$	strict order in $\langle \mathbb{R}, \leqslant \rangle$
$q, e \in \mathbb{R}$	$q - r$	subtraction in $\langle \mathbb{R}, - \rangle$
$q, r \in \mathbb{R}$	q / r	division in $\langle \mathbb{R} \cdot \rangle$
$q \in \mathbb{R}$	$q + 0$ or $0 + q$	neutral element in $\langle \mathbb{R}, + \rangle$
$q \in \mathbb{R}$	$q \cdot 1$ or $1 \cdot q$	neutral element in $\langle \mathbb{R}, \cdot \rangle$
$q \in \mathbb{R}$	$-q$	inverse element in $\langle \mathbb{R}, + \rangle$
$q, r \in \mathbb{C}$	q+r	sum in $\langle \mathbb{C}, + \rangle$
$q, r \in \mathbb{C}$	$q \cdot r$	product in $\langle \mathbb{C}, \cdot \rangle$
$q, r \in \mathbb{C}$	$q \leqslant r$	order in $\langle \mathbb{C}, \leqslant \rangle$
$q, r \in \mathbb{C}$	$q < r$	strict order in $\langle \mathbb{C}, \leqslant \rangle$
$q, e \in \mathbb{C}$	$q - r$	subtraction in $\langle \mathbb{C}, - \rangle$
$q, r \in \mathbb{C}$	q / r	division in $\langle \mathbb{C}, \cdot \rangle$
$q \in \mathbb{C}$	$q + 0$ or $0 + q$	neutral element in $\langle \mathbb{C}, + \rangle$
$q \in \mathbb{C}$	$q \cdot 1$ or $1 \cdot q$	neutral element in $\langle \mathbb{C}, \cdot \rangle$
$q \in \mathbb{C}$	$-q$	inverse element in $\langle \mathbb{C}, + \rangle$

10.1 Definition and arithmetic's

Definition 10.1. *The space* $\mathbb{C}$ *of complex numbers together with two operators* $+: \mathbb{C} \times \mathbb{C} \rightarrow \mathbb{C}$ *and* $\cdot: \mathbb{C} \times \mathbb{C} \rightarrow \mathbb{C}$ *is is defined as follows*

1. $\mathbb{C} = \mathbb{R} \times \mathbb{R} = \{(x, y) | x, y \in \mathbb{R}\}$
2. $+: \mathbb{C} \times \mathbb{C} \rightarrow \mathbb{C}$ *where* $(x, y) + (x', y') = (x + x', y + y')$
3. $\cdot: \mathbb{C} \times \mathbb{C} \rightarrow \mathbb{C}$ *where* $(x, y) \cdot (x', y') = (x \cdot x' - y \cdot y', x \cdot y' + y \cdot x')$

Just as $\langle\mathbb{R},+,\cdot\rangle$ is a field we have that $\langle\mathbb{C},+,\cdot\rangle$ is a field.

Theorem 10.2. *$\langle\mathbb{C},+,\cdot\rangle$ is a field where*

1. *The additive neutral element is* $(0,0)$
2. *Then multiplicative neutral element is* $(1,0)$
3. *The additive inverse element of* $(x,y)=(-x,-y)$
4. *The multiplicative inverse element for* $(x,y)\neq(0,0)$ *is*

$$(x/(x^2+y^2),-y/(x^2+y^2))$$

As usual we use the following notation convention based on context:

1. *The additive neutral element is noted as* 0.
2. *The multiplicative neutral element is noted as* 1.
3. *The additive inverse of* $z\in\mathbb{C}$ *is noted as* $-z$.
4. *The multiplicative inverse of* $x\in\mathbb{C}\setminus\{0\}$ *is noted as* x^{-1}.

Proof. First we prove that $\langle\mathbb{C},+\rangle$ is a Abelian group:

associativity. If $(x,y),(x',y'),(x'',y'')\in\mathbb{C}$ then we have

$$\begin{aligned}(x,y)+((x',y')+(x'',y'')) &= (x,y)+(x'+x'',y'+y'')\\ &= (x+(x'+x''),y+(y'+y''))\\ &= ((x+x')+x'',(y+y')+y'')\\ &= (x+x',y+y')+(x'',y'')\\ &= ((x,y)+(x',y'))+(x'',y'')\end{aligned}$$

commutativiy. If $(x,y),(x',y')\in\mathbb{C}$ then

$$(x,y)+(x',y')=(x+x',y+y')=(x'+x,y'+y)=(x',y')+(x,y)$$

neutral element. If $(x,y)\in\mathbb{C}$ then

$$\begin{aligned}(x,y)+(0,0) &\underset{\text{commutativity}}{=} (0,0)+(x,y)\\ &= (0+x,0+y)\\ &= (x,y)\end{aligned}$$

inverse element. If $(x,y)\in\mathbb{C}$ then

$$\begin{aligned}(x,y)+(-x,-y) &= (-x,-y)+(x,y)\\ &= ((-x)+x,(-y)+y)\\ &= (0,0)\end{aligned}$$

Next we prove the rest of the axioms for a field for the multiplication

associativity. If $(x,y),(x',y'),(x'',y'')\in\mathbb{C}$ then

$$\begin{aligned}(x,y)\cdot((x',y')\cdot(x'',y'')) &=\\ (x,y)\cdot(x'\cdot x''-y'\cdot y'',x'\cdot y''+y'\cdot x'') &=\\ (x\cdot(x'\cdot x''-y'\cdot y'')-y\cdot(x'\cdot y''+y'\cdot x''),x\cdot(x'\cdot y''+y'\cdot x''))+y\cdot(x'\cdot x''-y'\cdot y'') &=\\ (x\cdot(x'\cdot x'')-x\cdot(y'\cdot y'')-y\cdot(x'\cdot y'')-y\cdot(y'\cdot x''),x\cdot(x'\cdot y'')+x\cdot(y'\cdot x'')+y\cdot(x'\cdot x'')-y\cdot(y'\cdot y'')) &=\\ \left(\underbrace{(x\cdot x')\cdot x''}_{1}-\underbrace{(x\cdot y')\cdot y''}_{2}\underbrace{-(y\cdot x')\cdot y''}_{2}\underbrace{-(y\cdot y')\cdot x''}_{1},\underbrace{(x\cdot x')\cdot y''}_{3}+\underbrace{(x\cdot y')\cdot x''+(y\cdot x')\cdot x''}_{4}\underbrace{-(y\cdot y')\cdot y''}_{3}\right) &=\\ \left(\underbrace{(x\cdot x'-y\cdot y')\cdot x''}_{1}\underbrace{-(x\cdot y'+y\cdot y')\cdot y''}_{2},\underbrace{(x\cdot x'-y\cdot y')\cdot y''}_{3}+\underbrace{(x\cdot y'+y\cdot x')\cdot x''}_{4}\right) &=\\ (x\cdot x'-y\cdot y',x\cdot y'+y\cdot y')\cdot(x'',y'') &=\\ ((x,y)\cdot(x',y'))\cdot(x'',y'') &=\end{aligned}$$

commutativity. If $(x,y),(x',y')\in\mathbb{C}$ then

$$(x,y)\cdot(x',y')=(x\cdot x'-y\cdot y',x\cdot y'+y\cdot x')=(x'\cdot x-y'\cdot y,y'\cdot x+x'\cdot y)=(x',y')\cdot(x,y)$$

neutral element. If $(x,y)\in\mathbb{C}$ then

$$\begin{array}{rcl}(x,y)\cdot(1,0) & \underset{\text{commutativity}}{=} & (1,0)\cdot(x,y)\\ & = & (1\cdot x-0\cdot y,1\cdot y+0\cdot x)\\ & = & (x,y)\end{array}$$

inverse element. Let $(x,y)\in\mathbb{C}\setminus\{(0,0)\}$ then $x\neq 0\vee y\neq 0$, by [theorems: 9.45, 4.73] it follows that $0<x^2\vee 0<y^2$ giving $0<x^2+y^2$, so that $(x^2+y^2)\neq 0$ hence $(x/(x^2+y^2),-y/(x^2+y^2))$ is well defined. Now

$$\begin{array}{rc}(x,y)\cdot(x/(x^2+y^2),-y/(x^2+y^2)) & \underset{\text{commutativity}}{=}\\ (x/(x^2+y^2),-y/(x^2+y^2))\cdot(x,y) & =\\ (x^2/(x^2+y^2)+y^2/(x^2+y^2),x\cdot y/(x^2+y^2)+(-y)\cdot x/(x^2+y^2)) & =\\ ((x^2+y^2)/(x^2+y^2),(x\cdot y-y\cdot x)/(x^2+y^2)) & =\\ (1,0) & =\end{array}$$

distributivity. If $(x,y),(x',y'),(x'',y'')\in\mathbb{C}$ then

$$\begin{array}{rcl}(x,y)\cdot((x',y')+(x'',y'')) & = & (x,y)\cdot(x'+x'',y'+y'')\\ & = & (x\cdot(x'+x'')-y\cdot(y'+y''),x\cdot(y'+y'')+y\cdot(x'+x''))\\ & = & (x\cdot x'+x\cdot x''-y\cdot y'-y\cdot y'',x\cdot y'+x\cdot y''+y\cdot x'+y\cdot x'')\\ & = & ((x\cdot x'-y\cdot y')+(x\cdot x''-y\cdot y''),(x\cdot y'+y\cdot x')+(x\cdot y''+y\cdot x''))\\ & = & (x\cdot x'-y\cdot y',x\cdot y'+y\cdot x')+(x\cdot x''-y\cdot y'',x\cdot y''+y\cdot x'')\\ & = & (x,y)\cdot(x',y')+(x,y)\cdot(x'',y'')\end{array}$$

Finally as in $\mathbb{R}$ we have that $1\neq 0$ [see theorem: 9.32] we have that

$$0=(0,0)\neq(1,0)=1$$

□

Just as with the integers, rationals and real numbers we introduce the following shorthand notation.

Notation 10.3. *If $x,y\in\mathbb{C}$ then we have the following notation conventions:*

1. *$x+(-y)$ is noted as $x-y$*
2. *If $y\in\mathbb{C}\setminus\{0\}$ then $x\cdot y^{-1}$ is noted as x/y and y^{-1} is noted as $1/y$*

10.2 Embedding of $\mathbb{N}_0,\mathbb{Z},\mathbb{Q}$ and $\mathbb{R}$ in $\mathbb{C}$

10.2.1 Embeddings

Definition 10.4. *We define $\mathbb{R}_{\mathbb{C}},\mathbb{Q}_{\mathbb{C}},\mathbb{Z}_{\mathbb{C}},\mathbb{N}_{0,\mathbb{C}}$ as follows*

$$\mathbb{R}_{\mathbb{C}}=\{(x,0)|x\in\mathbb{R}\}$$
$$\mathbb{Q}_{\mathbb{C}}=\{(x,0)|x\in\mathbb{Q}_{\mathbb{R}}\}$$
$$\mathbb{Z}_{\mathbb{C}}=\{(x,0)|x\in\mathbb{Z}_{\mathbb{R}}\}$$
$$\mathbb{N}_{0,\mathbb{C}}=\{(x,0)|x\in\mathbb{N}_{0,\mathbb{R}}\}$$

It turns out that $\mathbb{R}_{\mathbb{C}},\mathbb{Q}_{\mathbb{C}},\mathbb{Z}_{\mathbb{C}},\mathbb{N}_{\mathbb{C}}$ are embeddings of $\mathbb{R},\mathbb{Q},\mathbb{Z}$ and $\mathbb{N}_0$ in $\mathbb{C}$.

Theorem 10.5. *We have that*

1. *$\mathbb{R}_{\mathbb{C}}$ is a sub-field of $\langle\mathbb{C},+,\cdot\rangle$ and*

$$i_{\mathbb{R}\rightarrow\mathbb{C}}:\langle\mathbb{R},+,\cdot\rangle\rightarrow\langle\mathbb{R}_{\mathbb{C}},+,\cdot\rangle \text{ defined by } i_{\mathbb{R}\rightarrow\mathbb{C}}(r)=(r,0)$$

is a field isomorphism.

2. $\mathbb{Q}_{\mathbb{C}}$ *is a sub-field of* $\langle \mathbb{C}, +, \cdot \rangle$ *and if we define* $i_{\mathbb{Q}\rightarrow\mathbb{C}} = i_{\mathbb{R}\rightarrow\mathbb{C}} \circ i_{\mathbb{Q}\rightarrow\mathbb{R}}$ *then*
$$i_{\mathbb{Q}\rightarrow\mathbb{C}}: \langle \mathbb{Q}, +, \cdot \rangle \rightarrow \langle \mathbb{Q}_{\mathbb{C}}, +, \cdot \rangle \text{ is a field isomorphism}$$
3. $\mathbb{Z}_{\mathbb{C}}$ *is a sub-ring of* $\langle \mathbb{C}, +, \cdot \rangle$ *and if we define* $i_{\mathbb{Z}\rightarrow\mathbb{C}} = i_{\mathbb{R}\rightarrow\mathbb{C}} \circ i_{\mathbb{Z}\rightarrow\mathbb{R}}$ *then*
$$i_{\mathbb{Z}\rightarrow\mathbb{C}}: \langle \mathbb{Z}, +, \cdot \rangle \rightarrow \langle \mathbb{Z}_{\mathbb{C}}, +, \cdot \rangle \text{ is a field isomorphism}$$
4. $\mathbb{N}_{0,\mathbb{C}}$ *is a sub semi-group of* $\langle \mathbb{C}, + \rangle$ *and if we define* $i_{\mathbb{N}_0\rightarrow\mathbb{C}} = i_{\mathbb{R}\rightarrow\mathbb{C}} \circ i_{\mathbb{N}_0\rightarrow\mathbb{R}}$ *then*
$$i_{\mathbb{N}_0\rightarrow\mathbb{C}}: \langle \mathbb{N}_0, + \rangle \rightarrow \langle \mathbb{N}_{0,\mathbb{C}}, + \rangle \text{ is a group isomorphism}$$
5. $\mathbb{N}_{0,\mathbb{C}}$ *is a sub semi-group of* $\langle \mathbb{C}, \cdot \rangle$ *and if we define* $i_{\mathbb{N}_0\rightarrow\mathbb{C}} = i_{\mathbb{R}\rightarrow\mathbb{C}} \circ i_{\mathbb{N}_0\rightarrow\mathbb{R}}$ *then*
$$i_{\mathbb{N}_0\rightarrow\mathbb{C}}: \langle \mathbb{N}_0, \cdot \rangle \rightarrow \langle \mathbb{N}_{0,\mathbb{C}}, \cdot \rangle \text{ is a group isomorphism}$$
6. $\mathbb{N}_{0,\mathbb{C}} \subseteq \mathbb{Z}_{\mathbb{C}} \subseteq \mathbb{Q}_{\mathbb{C}} \subseteq \mathbb{R}_{\mathbb{C}} \subseteq \mathbb{C}$
7. $\mathbb{N}_{0,\mathbb{C}}, \mathbb{Z}_{\mathbb{C}}$ *and* $\mathbb{Q}_{\mathbb{C}}$ *are denumerable*

Proof.

1. If $x \in \mathbb{R}_{\mathbb{C}}$ then $\exists x' \in \mathbb{R}$ such that $x = (x', 0) = i_{\mathbb{R}\rightarrow\mathbb{C}}(x') \in i_{\mathbb{R}\rightarrow\mathbb{C}}(\mathbb{R})$. Also if $x \in i_{\mathbb{R}\rightarrow\mathbb{C}}(\mathbb{R})$ then $\exists x' \in \mathbb{R}$ such that $x = i_{\mathbb{R}\rightarrow\mathbb{C}}(x') = (x', 0) \in \mathbb{R}_{\mathbb{C}}$. So we have that
$$\mathbb{R}_{\mathbb{C}} = i_{\mathbb{R}\rightarrow\mathbb{C}}(\mathbb{R}) \tag{10.1}$$
Let $x, y \in \mathbb{R}$ then we have
$$i_{\mathbb{R}\rightarrow\mathbb{C}}(x + y) = (x + y, 0) = (x, 0) + (y, 0) = i_{\mathbb{R}\rightarrow\mathbb{C}}(x) + i_{\mathbb{R}\rightarrow\mathbb{C}}(y)$$
and
$$i_{\mathbb{R}\rightarrow\mathbb{C}}(x \cdot y) = (x \cdot y, 0) = (x \cdot y - 0 \cdot 0, x \cdot 0 + 0 \cdot y) = (x, 0) \cdot (y, 0) = i_{\mathbb{R}\rightarrow\mathbb{C}}(x) \cdot i_{\mathbb{R}\rightarrow\mathbb{C}}(y)$$
$$i_{\mathbb{R}\rightarrow\mathbb{C}}(1) = (1, 0) = 1$$
proving that
$$i_{\mathbb{R}\rightarrow\mathbb{C}}: \langle \mathbb{R}, +, \cdot \rangle \rightarrow \langle \mathbb{C},, +, \cdot \rangle \text{ is a field homeomorphism} \tag{10.2}$$
As $\langle \mathbb{R}, +, \cdot \rangle$ is a field we can use [theorem: 4.70] together with the above proving that $i_{\mathbb{R}\rightarrow\mathbb{C}}(\mathbb{R})$ is a sub-field of $\langle \mathbb{C}, +, \cdot \rangle$. Combining this with [eq: 10.1] gives
$$\mathbb{R}_{\mathbb{C}} \text{ is a sub-field of } \langle \mathbb{C} \dotplus, \cdot \rangle$$
Further if $i_{\mathbb{R}\rightarrow\mathbb{C}}(x) = i_{\mathbb{R}\rightarrow\mathbb{C}}(y)$ then $(x, 0) = (y, 0)$ proving that $x = y$ hence
$$i_{\mathbb{R}\rightarrow\mathbb{C}}: \mathbb{R} \rightarrow \mathbb{C} \text{ is injective}$$
so that $i_{\mathbb{R}\rightarrow\mathbb{C}}: \mathbb{R} \rightarrow i_{\mathbb{R}\rightarrow\mathbb{R}}(\mathbb{R}) \underset{[\text{eq: } 10.1]}{=} \mathbb{R}_{\mathbb{C}}$ is a bijection. Hence we have that
$$i_{\mathbb{R}\rightarrow\mathbb{C}}: \langle \mathbb{R}, +, \cdot \rangle \rightarrow \langle \mathbb{R}_{\mathbb{C}}, +, \cdot \rangle \text{ is a field isomorphism} \tag{10.3}$$
2. By [theorem: 9.37] we have that
$$\langle \mathbb{Q}_{\mathbb{R}}, +, \cdot \rangle \text{ is a sub-field of } \langle \mathbb{R}, +, \cdot \rangle \text{ and } i_{\mathbb{Q}\rightarrow\mathbb{R}}: \langle \mathbb{Q}, +, \cdot \rangle \rightarrow \langle \mathbb{Q}_{\mathbb{R}}, +, \cdot \rangle \text{ is a field isomorphism} \tag{10.4}$$
Using then the above together with [eq: 10.3] we have by [theorem: 4.71] that
$$i_{\mathbb{R}\rightarrow\mathbb{C}}(i_{\mathbb{Q}\rightarrow\mathbb{R}}(\mathbb{Q})) \text{ is a sub-field of } \langle \mathbb{R}_{\mathbb{C}}, +, \cdot \rangle \tag{10.5}$$
and
$$i_{\mathbb{R}\rightarrow\mathbb{C}} \circ i_{\mathbb{Q}\rightarrow\mathbb{R}}: \langle \mathbb{Q}, +, \cdot \rangle \rightarrow \langle i_{\mathbb{R}\rightarrow\mathbb{C}}(i_{\mathbb{Q}\rightarrow\mathbb{R}}(\mathbb{Q})), +, \cdot \rangle \text{ is a field isomorphism} \tag{10.6}$$
Now if $x \in i_{\mathbb{R}\rightarrow\mathbb{C}}(i_{\mathbb{Q}\rightarrow\mathbb{R}}(\mathbb{Q}))$ $\exists x' \in \mathbb{Q}$ such that $x = i_{\mathbb{R}\rightarrow\mathbb{C}}(i_{\mathbb{Q}\rightarrow\mathbb{R}}(x')) = (i_{\mathbb{Q}\rightarrow\mathbb{R}}(x'), 0)$, using [eq: 10.4] we have that $i_{\mathbb{Q}\rightarrow\mathbb{R}}(x') \in \mathbb{Q}_{\mathbb{R}}$ so that $x \in \mathbb{Q}_{\mathbb{C}}$. Also if $x \in \mathbb{Q}_{\mathbb{C}}$ then there exist a $x' \in \mathbb{Q}_{\mathbb{R}}$ such that $x = (x', 0) = i_{\mathbb{R}\rightarrow\mathbb{C}}(x')$, using [eq: 10.4] there exists a $x'' \in \mathbb{Q}$ such that $x' = i_{\mathbb{Q}\rightarrow\mathbb{R}}(x'')$ so that $x = i_{\mathbb{R}\rightarrow\mathbb{C}}(i_{\mathbb{Q}\rightarrow R}(x'')) \in i_{\mathbb{R}\rightarrow\mathbb{C}}(i_{\mathbb{Q}\rightarrow\mathbb{R}}(\mathbb{Q}))$. Hence we have that
$$\mathbb{Q}_{\mathbb{C}} = i_{\mathbb{R}\rightarrow\mathbb{C}}(i_{\mathbb{Q}\rightarrow\mathbb{R}}(\mathbb{Q}))$$

which combined with [eqs: 10.5,10.6] and [theorem: 4.62] proves that

$$\mathbb{Q}_{\mathbb{C}} \text{ is a sub-field of } \langle \mathbb{C}, +, \cdot \rangle$$

and

$$i_{\mathbb{Q}\rightarrow\mathbb{C}} = i_{\mathbb{R}\rightarrow\mathbb{C}} \circ i_{\mathbb{Q}\rightarrow\mathbb{R}}: \langle \mathbb{Q}, +, \cdot \rangle \rightarrow \langle \mathbb{Q}_{\mathbb{C}}, +, \cdot \rangle \text{ is a field isomorphism}$$

3. By [theorem: 9.52] we have that

$$\langle \mathbb{Z}_{\mathbb{R}}, +, \cdot \rangle \text{ is a sub-ring of } \langle \mathbb{R}, +, \cdot \rangle \text{ and } i_{\mathbb{Z}\rightarrow\mathbb{R}}: \langle \mathbb{Z}, +, \cdot \rangle \rightarrow \langle \mathbb{Z}_{\mathbb{R}}, +, \cdot \rangle \text{ is a field isomorphism} \tag{10.7}$$

Using then the above together with [eq: 10.3] we have by [theorem: 4.48] that

$$i_{\mathbb{R}\rightarrow\mathbb{C}}(i_{\mathbb{Z}\rightarrow\mathbb{R}}(\mathbb{Z})) \text{ is a sub-ring of } \langle \mathbb{R}_{\mathbb{C}}, +, \cdot \rangle \tag{10.8}$$

and

$$i_{\mathbb{R}\rightarrow\mathbb{C}} \circ i_{\mathbb{Z}\rightarrow\mathbb{R}}: \langle \mathbb{Z}, +, \cdot \rangle \rightarrow \langle i_{\mathbb{R}\rightarrow\mathbb{C}}(i_{\mathbb{Z}\rightarrow\mathbb{R}}(\mathbb{Z})), +, \cdot \rangle \text{ is a ring isomorphism} \tag{10.9}$$

Now if $x \in i_{\mathbb{R}\rightarrow\mathbb{C}}(i_{\mathbb{Z}\rightarrow\mathbb{R}}(\mathbb{Z}))$ $\exists x' \in \mathbb{Z}$ such that $x = i_{\mathbb{R}\rightarrow\mathbb{C}}(i_{\mathbb{Z}\rightarrow\mathbb{R}}(x')) = (i_{\mathbb{Z}\rightarrow\mathbb{R}}(x'), 0)$, using [eq: 10.7] we have that $i_{\mathbb{Z}\rightarrow\mathbb{R}}(x') \in \mathbb{Z}_{\mathbb{R}}$ so that $x \in \mathbb{Z}_{\mathbb{C}}$. Also if $x \in \mathbb{Z}_{\mathbb{C}}$ then there exist a $x' \in \mathbb{Z}_{\mathbb{R}}$ such that $x = (x', 0) = i_{\mathbb{R}\rightarrow\mathbb{C}}(x')$, using [eq: 10.7] there exists a $x'' \in \mathbb{Z}$ such that $x' = i_{\mathbb{Z}\rightarrow\mathbb{R}}(x'')$ so that $x = i_{\mathbb{R}\rightarrow\mathbb{C}}(i_{\mathbb{Z}\rightarrow R}(x'')) \in i_{\mathbb{R}\rightarrow\mathbb{C}}(i_{\mathbb{Z}\rightarrow\mathbb{R}}(\mathbb{Z}))$. Hence we have that

$$\mathbb{Z}_{\mathbb{C}} = i_{\mathbb{R}\rightarrow\mathbb{C}}(i_{\mathbb{Z}\rightarrow\mathbb{R}}(\mathbb{Z}))$$

which combined with [eqs: 10.8,10.9] and [theorem: 4.17] proves that

$$\mathbb{Z}_{\mathbb{C}} \text{ is a sub-ring of } \langle \mathbb{C}, +, \cdot \rangle$$

and

$$i_{\mathbb{Z}\rightarrow\mathbb{C}} = i_{\mathbb{R}\rightarrow\mathbb{C}} \circ i_{\mathbb{Z}\rightarrow\mathbb{R}}: \langle \mathbb{Z}, +, \cdot \rangle \rightarrow \langle \mathbb{Z}_{\mathbb{C}}, +, \cdot \rangle \text{ is a ring isomorphism}$$

4. By [theorem: 9.55] we have that

$$\langle \mathbb{N}_{0,\mathbb{R}}, + \rangle \text{ is a sub semi-group of } \langle \mathbb{R}, + \rangle \text{ and } i_{\mathbb{N}_0\rightarrow\mathbb{R}}: \langle \mathbb{N}_0, + \rangle \rightarrow \langle \mathbb{N}_{0,\mathbb{R}}, + \rangle \text{ is a group isomorphism} \tag{10.10}$$

Using then the above together with [eq: 10.3] we have by [theorem: 4.48] that

$$i_{\mathbb{R}\rightarrow\mathbb{C}}(i_{\mathbb{N}_0\rightarrow\mathbb{R}}(\mathbb{N}_0)) \text{ is a sub semi-group of } \langle \mathbb{R}_{\mathbb{C}}, + \rangle \tag{10.11}$$

and

$$i_{\mathbb{R}\rightarrow\mathbb{C}} \circ i_{\mathbb{N}_0\rightarrow\mathbb{R}}: \langle \mathbb{N}_0, + \rangle \rightarrow \langle i_{\mathbb{R}\rightarrow\mathbb{C}}(i_{\mathbb{C}\rightarrow\mathbb{R}}(\mathbb{N}_0)), + \rangle \text{ is a group isomorphism} \tag{10.12}$$

Now if $x \in i_{\mathbb{R}\rightarrow\mathbb{C}}(i_{\mathbb{N}_0\rightarrow\mathbb{R}}(\mathbb{N}_0))$ $\exists x' \in \mathbb{N}_0$ such that $x = i_{\mathbb{R}\rightarrow\mathbb{C}}(i_{\mathbb{N}_0\rightarrow\mathbb{R}}(x')) = (i_{\mathbb{N}_0\rightarrow\mathbb{R}}(x'), 0)$, using [eq: 10.10] we have that $i_{\mathbb{N}_0\rightarrow\mathbb{R}}(x') \in \mathbb{N}_{0,\mathbb{R}}$ so that $x \in \mathbb{N}_{0,\mathbb{C}}$. Also if $x \in \mathbb{N}_{0,\mathbb{C}}$ then there exist a $x' \in \mathbb{C}_{\mathbb{R}}$ such that $x = (x', 0) = i_{\mathbb{R}\rightarrow\mathbb{C}}(x')$, using [eq: 10.10] there exists a $x'' \in \mathbb{N}_0$ such that $x' = i_{\mathbb{N}_0\rightarrow\mathbb{R}}(x'')$ so that $x = i_{\mathbb{R}\rightarrow\mathbb{C}}(i_{\mathbb{N}_0\rightarrow R}(x'')) \in i_{\mathbb{R}\rightarrow\mathbb{C}}(i_{\mathbb{N}_0\rightarrow\mathbb{R}}(\mathbb{N}_0))$. Hence we have that

$$\mathbb{N}_{0,\mathbb{C}} = i_{\mathbb{R}\rightarrow\mathbb{C}}(i_{\mathbb{N}_0\rightarrow\mathbb{R}}(\mathbb{N}_0)) \tag{10.13}$$

which combined with [eqs: 10.11,10.12] and [theorem: 4.17] proves that

$$\mathbb{N}_{0,\mathbb{C}} \text{ is a sub-semi-group of } \langle \mathbb{C}, + \rangle$$

and

$$i_{\mathbb{N}_0\rightarrow\mathbb{C}} = i_{\mathbb{R}\rightarrow\mathbb{C}} \circ i_{\mathbb{N}_0\rightarrow\mathbb{R}}: \langle \mathbb{N}_0, + \rangle \rightarrow \langle \mathbb{N}_{0,\mathbb{C}}, + \rangle \text{ is a group isomorphism}$$

5. By [theorem: 9.55] we have that

$$\langle \mathbb{N}_{0,\mathbb{R}}, \cdot \rangle \text{ is a sub semi-group of } \langle \mathbb{R}, \cdot \rangle \text{ and } i_{\mathbb{N}_0\rightarrow\mathbb{R}}: \langle \mathbb{N}_0, \cdot \rangle \rightarrow \langle \mathbb{N}_{0,\mathbb{R}}, \cdot \rangle \text{ is a group isomorphism} \tag{10.14}$$

Using then the above together with [eq: 10.3] we have by [theorem: 4.48] that

$$i_{\mathbb{R}\rightarrow\mathbb{C}}(i_{\mathbb{N}_0\rightarrow\mathbb{R}}(\mathbb{N}_0)) \text{ is a sub semi-group of } \langle \mathbb{R}_{\mathbb{C}}, \cdot \rangle \tag{10.15}$$

and

$$i_{\mathbb{R}\rightarrow\mathbb{C}} \circ i_{\mathbb{N}_0\rightarrow\mathbb{R}}: \langle \mathbb{N}_0, \cdot \rangle \rightarrow \langle i_{\mathbb{R}\rightarrow\mathbb{C}}(i_{\mathbb{C}\rightarrow\mathbb{R}}(\mathbb{N}_0)), \cdot \rangle \text{ is a group isomorphism} \tag{10.16}$$

Combining [eq: 10.13] with [eqs: 10.15,10.16] proves that

$$\mathbb{N}_{0,\mathbb{C}} \text{ is a sub-field of } \langle \mathbb{C}, \cdot \rangle$$

and

$$i_{\mathbb{N}_0 \to \mathbb{C}} = i_{\mathbb{R} \to \mathbb{C}} \circ i_{\mathbb{N}_0 \to \mathbb{R}}: \langle \mathbb{N}_0, \cdot \rangle \to \langle \mathbb{N}_{0,\mathbb{C}}, \cdot \rangle \text{ is a group isomorphism}$$

6. By [theorem: 9.56] $\mathbb{N}_{0,\mathbb{R}} \subseteq \mathbb{Z}_\mathbb{R} \subseteq \mathbb{Q}_\mathbb{R} \subseteq \mathbb{R}$ so that

$$\{(x,0)|x \in \mathbb{N}_{0,\mathbb{R}}\} \subseteq \{(x,0)|x \in \mathbb{Z}_\mathbb{R}\} \subseteq \{(x,0)|x \in \mathbb{Q}_\mathbb{R}|\} \subseteq \{(x,0)|x \in \mathbb{R}\} \subseteq \mathbb{C}$$

proving that

$$\mathbb{N}_{0,\mathbb{C}} \subseteq \mathbb{Z}_\mathbb{C} \subseteq \mathbb{Q}_\mathbb{C} \subseteq \mathbb{R}_\mathbb{C} \subseteq \mathbb{C}$$

7. Using (2),(3) and (4) we have that $\mathbb{Q} \approx \mathbb{Q}_\mathbb{C}$, $\mathbb{Z} \approx \mathbb{Z}_\mathbb{C}$ and $\mathbb{N}_0 \approx \mathbb{N}_{0,\mathbb{C}}$, further by [theorems: 7.53 and 8.44] we have that $\mathbb{N}_0 \approx \mathbb{Z}$ and $\mathbb{N}_0 \approx \mathbb{Q}$. So $\mathbb{N}_0 \approx \mathbb{N}_{0,\mathbb{C}}$, $\mathbb{N}_0 \approx \mathbb{Z}_\mathbb{C}$ and $\mathbb{N}_0 \approx \mathbb{Q}$. □

Theorem 10.6. *$\mathbb{Q}_\mathbb{C} = \{n/m | n \in \mathbb{Z}_\mathbb{C} \wedge m \in \mathbb{Z}_\mathbb{C} \setminus \{0\}\}$ where n/m is a shorthand for $n \cdot m^{-1}$ [see notation: 10.3]*

Proof. Let $x \in \mathbb{Q}_\mathbb{C}$ then by definition there exists a $q \in \mathbb{Q}_\mathbb{R}$ such that $x = (q,0) \underset{\text{[theorem: 10.5]}}{=} i_{\mathbb{R} \to \mathbb{C}}(q)$. Using then [theorem: 9.57] we have $q = n \cdot m^{-1}$ where $n \in \mathbb{Z}_\mathbb{R}$ and $m \in \mathbb{Z}_\mathbb{R} \setminus \{0\}$. So we have

$$\begin{aligned} x &= i_{\mathbb{R}\to\mathbb{C}}(q) \\ &= i_{\mathbb{R}\to\mathbb{C}}(n \cdot m^{-1}) \\ &= i_{\mathbb{R}\to\mathbb{C}}(n) \cdot i_{\mathbb{R}\to\mathbb{C}}(m^{-1}) \\ &\underset{\text{[theorem: 4.70]}}{=} i_{\mathbb{R}\to\mathbb{C}}(n) \cdot (i_{\mathbb{R}\to\mathbb{C}}(m))^{-1} \\ &\underset{\text{[theorem: 10.5]}}{=} (n,0) \cdot (m,0)^{-1} \end{aligned}$$

which, as by definition of $\mathbb{Z}_\mathbb{C}$ $(n,0) \in \mathbb{Z}_\mathbb{C}$ and $(m,0) \in \mathbb{Z}_\mathbb{C} \setminus \{0\}$, proves that $x \in \{n/m | n \in \mathbb{Z}_\mathbb{C} \wedge m \in \mathbb{Z}_\mathbb{C} \setminus \{0\}\}$. Hence

$$\mathbb{Q}_\mathbb{C} \subseteq \{n/m | n \in \mathbb{Z}_\mathbb{C} \wedge m \in \mathbb{Z}_\mathbb{C} \setminus \{0\}\} \tag{10.17}$$

If $x \in \{n/m | n \in \mathbb{Z}_\mathbb{C} \wedge m \in \mathbb{Z}_\mathbb{C} \setminus \{0\}\}$ then there exists $n \in \mathbb{Z}_\mathbb{C}$ and $m \in \mathbb{Z}_\mathbb{C} \setminus \{0\}$ such that $x = n \cdot m^{-1}$. So there exists $n' \in \mathbb{Z}_\mathbb{R}$ and $m' \in \mathbb{Z}_\mathbb{R} \setminus \{0\}$ such that $n = (n',0)$ and $m = (m',0)$. Now by [theorem: 9.57] $n' \cdot m'^{-1} \in \mathbb{Q}_\mathbb{R}$ so that $(n' \cdot m'^{-1}, 0) \in \mathbb{Q}_\mathbb{C}$. As

$$\begin{aligned} (n' \cdot m'^{-1}, 0) &\underset{\text{[theorem: 10.5]}}{=} i_{\mathbb{R}\to\mathbb{C}}(n' \cdot m'^{-1}) \\ &= i_{\mathbb{R}\to\mathbb{C}}(n') \cdot i_{\mathbb{R}\to\mathbb{C}}(m'^{-1}) \\ &\underset{\text{[theorem: 4.70]}}{=} i_{\mathbb{R}\to\mathbb{C}}(n') \cdot (i_{\mathbb{R}\to\mathbb{C}}(m'))^{-1} \\ &= n \cdot m^{-1} \\ &= x \end{aligned}$$

proving that $x \in \mathbb{Q}_\mathbb{C}$. Hence $\{n/m | n \in \mathbb{Z}_\mathbb{C} \wedge m \in \mathbb{Z}_\mathbb{C} \setminus \{0\}\} \subseteq \mathbb{Q}_\mathbb{C}$ which combined with [eq: 10.17]

$$\mathbb{Q}_\mathbb{C} = \{n/m | n \in \mathbb{Z}_\mathbb{C} \wedge m \in \mathbb{Z}_\mathbb{C} \setminus \{0\}\}$$ □

10.2.2 Order relation

As $\langle \mathbb{R}, \leqslant \rangle$ is totally ordered by [theorem: 9.43] we could use [theorem: 3.37] to define a lexical order on $\mathbb{C}$. However we can't guarantee $\langle \mathbb{C}, +, \cdot \rangle$ is a ordered field. The proof is by contradiction, so assume that $\langle \mathbb{C}, +. \cdot, \leqslant \rangle$ is a ordered field then by [theorem: 4.73]

1. If $x < y$ then $x + z < y + z$
2. If $x < y$ and $0 < z$ then $x \cdot z < y \cdot z$

Now for $i = (0,1)$ we have that $i \cdot i = (0,1) \cdot (0,1) = (0 \cdot 0 - 1 \cdot 1, 0 \cdot 1 + 1 \cdot 0) = (-1, 0) = -1$ and as $\langle \mathbb{C}, \leqslant \rangle$ must be totally ordered we have for i either:

$0 < i$. Then by (2) $0 < i \cdot i = -1$ so that $0 < -1$, hence by (1) we have $1 = 0 + 1 < 1 + (-1) = 0$ or $1 < 0$. But then by (2) we have $i = 1 \cdot i < 0 \cdot i = i$ or $i < 0$ contradicting $0 > i$.

$i<0$. Then by (1) we have that $0=i+(-i)<0+(-i)=-i$ so that $0<-i$. Hence using (2) we have that $0<(-i)\cdot(-i) \underset{\text{[theorem: 4.40]}}{=} i\cdot i=-1$ or $0<-1$ but then by (1) $1=0+1<(-1)+1<0$ or $1<0$. But then by (2) we have $-i=1\cdot(-i)<0\cdot(-i)=0$ or $-i<0$ contradicting that we found that $0<-i$.

However we can still have a order relation on the sub-field $\mathbb{R}_{\mathbb{C}}$ that satisfies (1) and (2) as will be showed in the following.

Definition 10.7. *The relation $\leqslant \subseteq \mathbb{R}_{\mathbb{C}}\times\mathbb{R}_{\mathbb{C}}$ is defined by*

$$\leqslant=\{((x,0),(y,0))\in\mathbb{R}_{\mathbb{C}}|x\leqslant y\}$$

Note: that in $x\leqslant y$ we use the order of $\langle\mathbb{R},\leqslant\rangle$

Theorem 10.8. *Using the above order relation we have that*

1. *$\langle\mathbb{R}_{\mathbb{C}},\leqslant\rangle$ is a totally ordered set*
2. *$i_{\mathbb{R}\rightarrow\mathbb{C}}\colon\langle\mathbb{R},\leqslant\rangle\rightarrow\langle\mathbb{R}_{\mathbb{C}},\leqslant\rangle$ is a order isomorphism*

Proof.

1. We have

   reflexivity. If $(x,0)\in\mathbb{R}_{\mathbb{C}}$ then as $\langle\mathbb{R},\leqslant\rangle$ is totally ordered we have $x\leqslant x$ so that $(x,0)\leqslant(x,0)$

   anti symmetry. If $(x,0)\leqslant(y,0)\wedge(y,0)\leqslant(x,0)$ then $x\leqslant y\wedge y\leqslant x$ which, as $\langle\mathbb{R},\leqslant\rangle$ is totally ordered, proves that $x=y$ hence $(x,0)=(y,0)$.

   transitivity. If $(x,0)\leqslant(y,0)\wedge(y,0)\leqslant(z,0)$ then $x\leqslant y\wedge y\leqslant z$ which, as $\langle\mathbb{R},\leqslant\rangle$ is totally ordered, proves that $x\leqslant z$, hence $(x,0)\leqslant(z,0)$

   totally order. If $(x,0),(y,0)\in\mathbb{R}_{\mathbb{C}}$ then we have as $\langle\mathbb{R},\leqslant\rangle$ is totally ordered that either $x\leqslant y\Rightarrow(x,0)\leqslant(y,0)$ or $y\leqslant x\Rightarrow(y,0)\leqslant(x,0)$.

2. First by [theorem:10.5] $i_{\mathbb{R}\rightarrow\mathbb{C}}\colon\mathbb{R}\rightarrow\mathbb{R}_{\mathbb{C}}$ is a bijection and second

$$\begin{aligned} i_{\mathbb{R}\rightarrow\mathbb{R}_{\mathbb{C}}}(x)\leqslant i_{\mathbb{R}\rightarrow\mathbb{C}}(y) &\Leftrightarrow (x,0)\leqslant(y,0)\\ &\underset{\text{definition}}{\Leftrightarrow} x\leqslant y \end{aligned}$$

□

Theorem 10.9. *We have that*

1. *$i_{\mathbb{R}\rightarrow\mathbb{C}}\colon\langle\mathbb{R},\leqslant\rangle\rightarrow\langle\mathbb{R}_{\mathbb{C}},\leqslant\rangle$ is a order isomorphism*
2. *$i_{\mathbb{Q}\rightarrow\mathbb{C}}\colon\langle\mathbb{Q},\leqslant\rangle\rightarrow\langle\mathbb{Q}_{\mathbb{C}},\leqslant\rangle$ is a order isomorphism*
3. *$i_{\mathbb{Z}\rightarrow\mathbb{C}}\colon\langle\mathbb{Z},\leqslant\rangle\rightarrow\langle\mathbb{Z}_{\mathbb{C}},\leqslant\rangle$ is a order isomorphism*
4. *$i_{\mathbb{N}_0\rightarrow\mathbb{C}}\colon\langle\mathbb{N}_0,\leqslant\rangle\rightarrow\langle\mathbb{N}_{0,\mathbb{C}},\leqslant\rangle$ is a order isomorphism*

Proof.

1. This is expressed by [theorem: 10.8].
2. Using (1) and [theorem: 9.49] we have that

$$i_{\mathbb{R}\rightarrow\mathbb{C}}\colon\langle\mathbb{R},\leqslant\rangle\rightarrow\langle\mathbb{R}_{\mathbb{C}},\leqslant\rangle \text{ and } i_{\mathbb{Q}\rightarrow\mathbb{R}}\colon\langle\mathbb{Q},\leqslant\rangle\rightarrow\langle\mathbb{Q}_{\mathbb{R}},\leqslant\rangle \text{ are order isomorphisms}$$

   so using [theorem 3.52]

$$i_{\mathbb{Q}\rightarrow\mathbb{C}}=i_{\mathbb{R}\rightarrow\mathbb{C}}\circ i_{\mathbb{Q}\rightarrow\mathbb{R}}\colon\langle\mathbb{Q},\leqslant\rangle\rightarrow\langle i_{\mathbb{R}\rightarrow\mathbb{C}}(i_{\mathbb{Q}\rightarrow\mathbb{R}}(\mathbb{Q})),\leqslant\rangle \text{ is a order isomorphism}$$

   As by [theorem: 10.5] $i_{\mathbb{R}\rightarrow\mathbb{C}}(i_{\mathbb{Q}\rightarrow\mathbb{R}}(\mathbb{Q}))=\mathbb{Q}_{\mathbb{C}}$ we have that

$$i_{\mathbb{Q}\rightarrow\mathbb{C}}=i_{\mathbb{R}\rightarrow\mathbb{C}}\circ i_{\mathbb{Q}\rightarrow\mathbb{R}}\colon\langle\mathbb{Q},\leqslant\rangle\rightarrow\langle\mathbb{Q}_{\mathbb{C}},\leqslant\rangle \text{ is a order isomorphism}$$

3. Using (1) and [theorem: 9.52] we have that

$$i_{\mathbb{R}\rightarrow\mathbb{C}}\colon\langle\mathbb{R},\leqslant\rangle\rightarrow\langle\mathbb{R}_{\mathbb{C}},\leqslant\rangle \text{ and } i_{\mathbb{Z}\rightarrow\mathbb{R}}\colon\langle\mathbb{Z},\leqslant\rangle\rightarrow\langle\mathbb{Z}_{\mathbb{R}},\leqslant\rangle \text{ are order isomorphisms}$$

   so using [theorem 3.52]

$$i_{\mathbb{Z}\rightarrow\mathbb{C}}=i_{\mathbb{R}\rightarrow\mathbb{C}}\circ i_{\mathbb{Z}\rightarrow\mathbb{R}}\colon\langle\mathbb{Z},\leqslant\rangle\rightarrow\langle i_{\mathbb{R}\rightarrow\mathbb{C}}(i_{\mathbb{Z}\rightarrow\mathbb{R}}(\mathbb{Z})),\leqslant\rangle \text{ is a order isomorphism}$$

As by [theorem: 10.5] $i_{\mathbb{R}\rightarrow\mathbb{C}}(i_{\mathbb{Z}\rightarrow\mathbb{R}}(\mathbb{Z}))=\mathbb{Z}_{\mathbb{C}}$ we have that

$$i_{\mathbb{Z}\rightarrow\mathbb{C}}=i_{\mathbb{R}\rightarrow\mathbb{C}}\circ i_{\mathbb{Z}\rightarrow\mathbb{R}}\colon \langle\mathbb{Z},\leqslant\rangle\rightarrow\langle\mathbb{Z}_{\mathbb{C}},\leqslant\rangle \text{ is a order isomorphism}$$

4. Using (1) and [theorem: 9.55] we have that

$$i_{\mathbb{R}\rightarrow\mathbb{C}}\colon \langle\mathbb{R},\leqslant\rangle\rightarrow\langle\mathbb{R}_{\mathbb{C}},\leqslant\rangle \text{ and } i_{\mathbb{N}_0\rightarrow\mathbb{R}}\colon \langle\mathbb{N}_0,\leqslant\rangle\rightarrow\langle\mathbb{N}_{0,\mathbb{R}},\leqslant\rangle \text{ are order isomorphisms}$$

so using [theorem 3.52]

$$i_{\mathbb{N}_0\rightarrow\mathbb{C}}=i_{\mathbb{R}\rightarrow\mathbb{C}}\circ i_{\mathbb{N}_0\rightarrow\mathbb{R}}\colon \langle\mathbb{N}_0,\leqslant\rangle\rightarrow\langle i_{\mathbb{R}\rightarrow\mathbb{C}}(i_{\mathbb{N}_0\rightarrow\mathbb{R}}(\mathbb{N}_0)),\leqslant\rangle \text{ is a order isomorphism}$$

As by [theorem: 10.5] $i_{\mathbb{R}\rightarrow\mathbb{C}}(i_{\mathbb{N}_0\rightarrow\mathbb{R}}(\mathbb{N}_0))=\mathbb{N}_{0,\mathbb{C}}$ we have that

$$i_{\mathbb{N}_0\rightarrow\mathbb{C}}=i_{\mathbb{R}\rightarrow\mathbb{C}}\circ i_{\mathbb{N}_0\rightarrow\mathbb{R}}\colon \langle\mathbb{N}_0,\leqslant\rangle\rightarrow\langle\mathbb{N}_{0,\mathbb{C}},\leqslant\rangle \text{ is a order isomorphism}$$

□

Corollary 10.10. *We have* $0,1\in\mathbb{R}_{\mathbb{C}}$ *with* $0<1$

Proof. Using [theorem: 10.5]we have that $i_{\mathbb{R}\rightarrow\mathbb{C}}(0)=0$ and $i_{\mathbb{R}\rightarrow\mathbb{C}}(1)=1$ which as in $\mathbb{R}$ $0<1$ proves that $0=i_{\mathbb{R}\rightarrow\mathbb{C}}(0)<i_{\mathbb{R}\rightarrow\mathbb{C}}(1)=1$ □

Theorem 10.11. $\forall x\in\mathbb{N}_{0,\mathbb{C}}$ *we have* $0\leqslant x$

Proof. By [theorems: 10.5, 10.9] we have that

$$i_{\mathbb{N}_0\rightarrow\mathbb{C}}\colon \langle\mathbb{N}_0,+\rangle\rightarrow\langle\mathbb{N}_{0,\mathbb{C}},+\rangle \text{ is a group isomorphism hence } i_{\mathbb{N}_0}(0)=0 \tag{10.18}$$

and

$$i_{\mathbb{N}_0\rightarrow\mathbb{C}}\colon \langle\mathbb{N}_0,\leqslant\rangle\rightarrow\langle\mathbb{N}_{0,\mathbb{C}},\leqslant\rangle \text{ is a order isomorphism} \tag{10.19}$$

If $x\in\mathbb{N}_{0,\mathbb{C}}$ we have by [eq: 10.18] that $\exists x'\in\mathbb{N}_0$ such that $x=i_{\mathbb{N}_0\rightarrow\mathbb{C}}(x')$. Using [theorem: 5.46] we have that $0\leqslant x'$. So we have $0\underset{\text{[eq: 10.18]}}{=}i_{\mathbb{N}_0\rightarrow\mathbb{C}}(0)\leqslant_{\text{[theorem: 10.18]}}i_{\mathbb{N}_0\rightarrow\mathbb{C}}(x')=x$ proving that $0\leqslant x$. □

We use now the order isomorphism and field isomorphism to transfer the properties of $\langle\mathbb{R},\leqslant\rangle$ and $\langle\mathbb{R},+,\cdot\rangle$ to $\langle\mathbb{R}_{\mathbb{C}},\leqslant\rangle$ and $\langle\mathbb{R}_{\mathbb{C}},+,\cdot\rangle$.

Definition 10.12. *Using the above order relation we can define* $\mathbb{R}^+_{\mathbb{C}}$, $\mathbb{R}^+_{0,\mathbb{C}}$, $\mathbb{R}^-_{\mathbb{C}}$ *and* $\mathbb{R}^-_{0,\mathbb{C}}$ *by*

1. $\mathbb{R}^+_{\mathbb{C}}=\{x\in\mathbb{R}_{\mathbb{C}}|0<x\}$
2. $\mathbb{R}^+_{0,\mathbb{C}}=\{x\in\mathbb{R}_{\mathbb{C}}|0\leqslant x\}$
3. $\mathbb{R}^-_{\mathbb{C}}=\{x\in\mathbb{R}_{\mathbb{C}}|x<0\}$
4. $\mathbb{R}^-_{0,\mathbb{C}}=\{x\in\mathbb{R}_{\mathbb{C}}|x\leqslant 0\}$

Theorem 10.13. $\langle\mathbb{R}_{\mathbb{C}},+,\cdot,\leqslant\rangle$ *is a ordered field*

Proof. Using [theorems: 10.5, 10.9] we have that

$$\langle\mathbb{R}_{\mathbb{C}},+,\cdot\rangle \text{ is a field}$$

and

$$i_{\mathbb{R}\rightarrow\mathbb{C}}\colon \langle\mathbb{R},\leqslant\rangle\rightarrow\langle\mathbb{R}_{\mathbb{C}},\leqslant\rangle \text{ is a order isomorphism} \tag{10.20}$$

and

$$i_{\mathbb{R}\rightarrow\mathbb{C}}\colon \langle\mathbb{R}\dotplus,\cdot\rangle\rightarrow\langle\mathbb{R},+,\cdot\rangle \text{ is a field isomorphism} \tag{10.21}$$

Then we have

1. If $x,y,z\in\mathbb{R}_{\mathbb{C}}$ with $x<y$ then $\exists\alpha,\beta,\gamma\in\mathbb{R}$ such that $x=i_{\mathbb{R}\rightarrow\mathbb{C}}(\alpha)$, y=$i_{\mathbb{R}\rightarrow\mathbb{C}}(\beta)$ and $z=i_{\mathbb{R}\rightarrow\mathbb{C}}(\gamma)$. As $x<y\Rightarrow i_{\mathbb{R}\rightarrow\mathbb{C}}(\alpha)<i_{\mathbb{R}\rightarrow\mathbb{C}}(\beta)$ we have by [eq: 10.20] that $\alpha<\beta$. By [theorem: 9.45] we have that $\alpha+\gamma<\beta+\gamma$. Using then [eqs: 10.20, 10.21] we have that

$$\begin{aligned} x+z &= i_{\mathbb{R}\rightarrow\mathbb{C}}(\alpha)+i_{\mathbb{R}\rightarrow\mathbb{C}}(\gamma)\\ &= i_{\mathbb{R}\rightarrow\mathbb{C}}(\alpha+\gamma)\\ &< i_{\mathbb{R}\rightarrow\mathbb{C}}(\beta+\gamma)\\ &= i_{\mathbb{R}\rightarrow\mathbb{C}}(\beta)+i_{\mathbb{R}\rightarrow\mathbb{C}}(\gamma)\\ &= y+z \end{aligned}$$

proving that

$$x + x < y + z$$

2. If $x, y \in \mathbb{R}_{\mathbb{C}}$ with $0 < x$ and $0 < y$ then there exist $\alpha, \beta \in \mathbb{R}$ such that $x = i_{\mathbb{R} \rightarrow \mathbb{C}}(\alpha)$ and $y = i_{\mathbb{R} \rightarrow \mathbb{C}}(\beta)$. As $0 < x$ and $0 < y$ we have by [eq: 10.20] that $0 < \alpha$ and $0 < \beta$, using [theorem: 9.45] it follows that $0 < \alpha \cdot \beta$. Using then [eqs: 10.20, 10.21] we have that

$$0 = i_{\mathbb{R} \rightarrow \mathbb{C}}(0) < i_{\mathbb{R} \rightarrow \mathbb{C}}(\alpha \cdot \beta) = i_{\mathbb{R} \rightarrow \mathbb{C}}(\alpha) \cdot i_{\mathbb{R} \twoheadrightarrow \mathbb{C}}(\beta) = x \cdot y$$

proving that

$$0 < x \cdot y$$ □

As $\langle \mathbb{R}, +, \cdot, \leqslant \rangle$ is a ordered field we have automatically the following properties:

Theorem 10.14. *For $\langle \mathbb{R}_{\mathbb{C}}, +, \cdot, \leqslant \rangle$ we have*

1. $\forall x, y, z \in \mathbb{R}_{\mathbb{C}}$ *we have* $x < y \Leftrightarrow x + z < y + z$
2. $\forall x, y, z \in \mathbb{R}_{\mathbb{C}}$ *we have* $x \leqslant y \Leftrightarrow x + z \leqslant y + z$
3. $\forall x, y \in \mathbb{R}_{\mathbb{C}}$ *we have* $x < y \Leftrightarrow 0 < y + (-x)$
4. $\forall x, y \in \mathbb{R}_{\mathbb{C}}$ *we have* $x \leqslant y \Leftrightarrow 0 \leqslant y + (-x)$
5. $\forall x, y \in \mathbb{R}_{\mathbb{C}}$ *we have* $x < y \Leftrightarrow -y < -x$
6. $\forall x, y \in \mathbb{R}_{\mathbb{C}}$ *we have* $x \leqslant y \Leftrightarrow -y \leqslant -x$
7. $\forall x, y, z \in \mathbb{R}_{\mathbb{C}}$ *with* $0 < z$ *we have* $x < y \Leftrightarrow x \cdot z < y \cdot z$
8. $\forall x, y, z \in \mathbb{R}_{\mathbb{C}}$ *with* $0 < z$ *we have* $x \leqslant y \Leftrightarrow x \cdot z \leqslant y \cdot z$
9. $\forall x, y, z \in \mathbb{R}_{\mathbb{C}}$ *with* $0 \leqslant z$ *and* $x \leqslant y$ *we have* $x \cdot z \leqslant y \cdot z$
10. $\forall x, y, z \in \mathbb{R}_{\mathbb{C}}$ *with* $z < 0$ *we have* $x < y \Leftrightarrow y \cdot z < x \cdot z$
11. $\forall x, y, z \in \mathbb{R}_{\mathbb{C}}$ *with* $z < 0$ *we have* $x \leqslant y \Leftrightarrow y \cdot z \leqslant x \cdot z$
12. $\forall x, y, z \in \mathbb{R}_{\mathbb{C}}$ *with* $z \leqslant 0$ *and* $x \leqslant y$ *we have* $y \cdot z \leqslant x \cdot z$
13. $\forall x \in \mathbb{R}_{\mathbb{C}}$ *we have* $0 \leqslant x \cdot x \underset{\text{def}}{=} x^2$, *further if* $0 \neq x$ *then* $0 < x \cdot x = x^2$
14. $0 \leqslant 1$ *[actually by [corollary: 10.10] we have* $0 < 1$*]*
15. $\forall x, y \in \mathbb{R}_{\mathbb{C}}$ *with* $0 < x < y$ *we have that* $x^2 < y^2$ *where* $x^2 = x \cdot x$ *and* $y^2 = y \cdot y$
16. $\forall x, y \in \mathbb{R}_{\mathbb{C}}$ *with* $0 \leqslant x \leqslant y$ *we have that* $x^2 \leqslant y^2$ *where* $x^2 = x \cdot x$ *and* $y^2 = y \cdot y$
17. $\forall x \in \mathbb{R}_{\mathbb{C}}$ *with* $0 < x$ *we have* $0 < x^{-1}$
18. $\forall x, y \in \mathbb{R}_{\mathbb{C}}$ *we have* $0 < x < y \Leftrightarrow 0 < y^{-1} < x^{-1}$
19. $\forall x, y \in \mathbb{R}_{\mathbb{C}}$ *we have* $0 < x \leqslant y \Leftrightarrow 0 < y^{-1} \leqslant x^{-1}$
20. $\forall x \in \mathbb{R}_{\mathbb{C}}$ *with* $0 < x < 1$ *we have* $1 < x^{-1}$
21. $\forall x \in \mathbb{R}_{\mathbb{C}}$ *with* $0 < x \leqslant 1$ *we have* $1 \leqslant x^{-1}$

Proof. This follows from [theorem: 4.73] □

Corollary 10.15. *If $x \in \mathbb{R}_{\mathbb{C}}$ then we have*

1. $x < x + 1$
2. $x - 1 < x$

Proof.

1. As $0 < 1$ [see corollary: 10.10] we have $x = 0 + x < 1 + x = x + 1$ [see theorem: 10.14].
2. From (1) we have $x < x + 1$ so that $x - 1 = x + (-1) < (x + 1) + -1 = x$ [see theorem: 10.14]. □

As for conditonal completeness we have the following theorems

Theorem 10.16. *$\langle \mathbb{Z}_{\mathbb{C}}, \leqslant \rangle$ is conditional complete [definition: 3.77] in other words*

$$\forall S \subseteq \mathbb{Z}_{\mathbb{C}} \text{ with } S \neq \varnothing \text{ such that } \exists \upsilon \in \mathbb{Z}_{\mathbb{C}} \text{ such that } \forall \alpha \in S \text{ we have } \alpha \leqslant \upsilon \text{ we have that } \sup(S) \text{ exist}$$

Using [theorem: 3.78] this is equivalent with

$$\forall S \subseteq \mathbb{Z}_{\mathbb{C}} \text{ with } S \neq \varnothing \text{ such that } \exists \lambda \in \mathbb{Z}_{\mathbb{C}} \text{ such that } \forall \alpha \in S \text{ we have } \lambda \leqslant \alpha \text{ we have that } \inf(S) \text{ exist}$$

Proof. Using [theorem: 10.9] we have that

$$i_{\mathbb{Z}\rightarrow\mathbb{C}}: \langle \mathbb{Z}, \leqslant \rangle \rightarrow \langle \mathbb{Z}_{\mathbb{C}}, \leqslant \rangle \text{ is a order isomorphism}$$

and by [theorem: 7.32] $\langle \mathbb{Z}, \leqslant \rangle$ is conditional complete. Hence by [theorem: 3.80]

$$\langle \mathbb{Z}_{\mathbb{C}}, \leqslant \rangle \text{ is conditional complete}$$

□

Theorem 10.17. *$\langle \mathbb{Q}_{\mathbb{C}}, \leqslant \rangle$ is not conditional complete.*

Proof. Using [theorems: 10.9 and 3.55] we have that

$$(i_{\mathbb{Q}\rightarrow\mathbb{C}})^{-1}: \langle \mathbb{Q}_{\mathbb{C}}, \leqslant \rangle \rightarrow \langle \mathbb{Q}, \leqslant \rangle \text{ is a order isomorphism}$$

Assume that $\langle \mathbb{Q}_{\mathbb{C}}, \leqslant \rangle$ is conditional complete then by the above and [theorem: 3.80] $\langle \mathbb{Q}, \leqslant \rangle$ is conditional complete, contradiction the fact that by [theorem: 8.41] $\langle \mathbb{Q}, \leqslant \rangle$ is not conditional complete. So the assumption is wrong and $\langle \mathbb{Q}_{\mathbb{R}}, \leqslant \rangle$ is not conditional complete. □

Theorem 10.18. *$\langle \mathbb{R}_{\mathbb{C}}, \leqslant \rangle$ is conditional complete [definition: 3.77] in other words*

$$\forall S \subseteq \mathbb{R}_{\mathbb{C}} \text{ with } S \neq \varnothing \text{ such that } \exists \upsilon \in \mathbb{R}_{\mathbb{C}} \text{ such that } \forall \alpha \in S \text{ we have } \alpha \leqslant \upsilon \text{ we have that } \sup(S) \text{ exist}$$

Using [theorem: 3.78] this is equivalent with

$$\forall S \subseteq \mathbb{R}_{\mathbb{C}} \text{ with } S \neq \varnothing \text{ such that } \exists \lambda \in \mathbb{R}_{\mathbb{C}} \text{ such that } \forall \alpha \in S \text{ we have } \lambda \leqslant \alpha \text{ we have that } \inf(S) \text{ exist}$$

Proof. Using [theorem: 10.9] we have that

$$i_{\mathbb{R}\rightarrow\mathbb{C}}: \langle \mathbb{R}, \leqslant \rangle \rightarrow \langle \mathbb{R}_{\mathbb{C}}, \leqslant \rangle \text{ is a order isomorphism}$$

and by [theorem: 9.50] $\langle \mathbb{R}, \leqslant \rangle$ is conditional complete. Hence by [theorem: 3.80]

$$\langle \mathbb{R}_{\mathbb{C}}, \leqslant \rangle \text{ is conditional complete}$$

□

Theorem 10.19. *$\langle \mathbb{N}_{0,\mathbb{C}}, \leqslant \rangle$ is well ordered [definition: 3.81]*

Proof. Using [theorem: 10.9] we have that

$$i_{\mathbb{N}_0\rightarrow}: \langle \mathbb{N}_0, \leqslant \rangle \rightarrow \langle \mathbb{N}_{0,\mathbb{C}}, \leqslant \rangle \text{ is a order isomorphism}$$

As by [theorems: 5.51] $\langle \mathbb{N}_0, \leqslant \rangle$ is well ordered we have by the above and [theorem: 3.82] that $\langle \mathbb{N}_{0,\mathbb{C}}, \leqslant \rangle$ is well ordered. □

Corollary 10.20. (Irrational numbers) *$\mathbb{Q}_{\mathbb{C}} \subset \mathbb{R}_{\mathbb{C}}$ so that $\mathbb{R}_{\mathbb{C}} \setminus \mathbb{Q}_{\mathbb{C}} \neq \varnothing$. The set $\mathbb{R}_{\mathbb{C}} \setminus \mathbb{Q}$ is called the set of* ***irrational numbers****.*

Proof. By [theorem: 10.17] $\langle \mathbb{Q}_{\mathbb{C}}, \leqslant \rangle$ is not conditional complete. Hence there exists a non empty $S \subseteq \mathbb{Q}_{\mathbb{C}}$ with a upper bound such that $\{u \in \mathbb{Q}_{\mathbb{C}} | u \text{ is a upper bound of } S\}$ has no least element. As $\langle \mathbb{R}_{\mathbb{C}}, \leqslant \rangle$ is conditional complete we have that $s = \min(\{u \in \mathbb{R}_{\mathbb{C}} | u$ is a upper bound of $S\})$ exist. Assume now that $s \in \mathbb{Q}_{\mathbb{C}}$ then if $u \in \{u \in \mathbb{Q}_{\mathbb{C}} | u \text{ is a upper bound of } S\}$ we have, as $\mathbb{Q}_{\mathbb{C}} \subseteq \mathbb{R}_{\mathbb{C}}$, that $u \in \{u \in \mathbb{R}_{\mathbb{C}} | u \text{ is a upper bound of } S\}$ so that $u \leqslant s$. Hence s is a least element of $\{u \in \mathbb{Q}_{\mathbb{C}} | u \text{ is a upper bound of } S\}$ contradicting the fact that $\{u \in \mathbb{Q}_{\mathbb{C}} | u \text{ is a upper bound of } S\}$ has no least element. So $s \notin \mathbb{Q}_{\mathbb{C}}$ or $\mathbb{R}_{\mathbb{C}} \setminus \mathbb{Q}_{\mathbb{C}} \neq \varnothing$. □

Theorem 10.21. *Let $\varnothing \neq S \subseteq \mathbb{R}_{\mathbb{C}}$ then for $-S = \{-s | s \in S\}$ we have*

1. *If $\sup(S)$ exist then $\inf(-S)$ exist and $\inf(-S) = -\sup(S)$.*

2. *If* $\inf(S)$ *exist then* $\sup(-S)$ *exist and* $\sup(-S) = -\inf(S)$

Proof. First

$$\begin{aligned} s \in -S &\Leftrightarrow \exists s' \in S \text{ such that } s = -s' \\ &\Leftrightarrow \exists s' \in S \text{ such that } -s = s' \\ &\Leftrightarrow -s \in S \end{aligned}$$

proving that

$$-S = \{s | -s \in S\}$$

1. Let $s \in -S$ then $-s \in S$ so that $-s \leqslant \sup(S) \Rightarrow -\sup(S) \leqslant s$. So $-S$ is bounded below by $-\sup(S)$, as $\mathbb{R}_{\mathbb{C}}$ is conditional complete [see theorem: 10.18] we have that

$$\inf(-S) \text{ exist and } -\sup(S) \leqslant \inf(-S) \tag{10.22}$$

Assume that $-\sup(S) < \inf(-S)$ then $-\inf(-S) < \sup(S)$ so that by the definition of the supremum there exist a $s \in S$ such that $-\inf(-S) < s \Rightarrow -s < \inf(-S)$ which, as $-s \in -S \Rightarrow \inf(-S) \leqslant -s$, results in the contradiction $-s < -s$. Hence we must have that $\inf(-S) \leqslant -\sup(S)$ which combined with [eq: 10.22] gives

$$\inf(-S) = -\sup(S)$$

2. Let $s \in -S$ then $-s \in S$ so that $\inf(S) \leqslant -s \Rightarrow s \leqslant -\inf(S)$. So $-S$ is bounded above by $-\inf(S)$, as $\mathbb{R}_{\mathbb{C}}$ is conditional complete [see theorem: 10.18] we have that

$$\sup(-S) \text{ exist and } \sup(-S) \leqslant -\inf(S) \tag{10.23}$$

Assume that $\sup(-S) < -\inf(S)$ then $\inf(S) < -\sup(-S)$ so that by the definition of the infinum there exist a $s \in S$ such that $s < -\sup(-S) \Rightarrow \sup(-S) < -s$ which, as $-s \in -S \Rightarrow -s \leqslant \sup(-S)$, results in the contradiction $-s < -s$. Hence we must have that $-\inf(S) \leqslant \sup(-S)$ which combined with [eq: 10.23] gives

$$\sup(-S) = -\inf(S)$$

□

Theorem 10.22. *Let* $\varnothing \neq S \subseteq \mathbb{R}_{\mathbb{C}}$ *and* $\alpha \in \mathbb{R}_{\mathbb{C}}$ *with* $0 \leqslant \alpha$ *then we have for* $\alpha \cdot S = \{\alpha \cdot s | s \in S\}$ *that*

1. *If* $\sup(S)$ *exist then* $\sup(\alpha \cdot S)$ *exist and* $\sup(\alpha \cdot S) = \alpha \cdot \sup(S)$
2. *If* $\inf(S)$ *exist then* $\inf(\alpha \cdot S)$ *exist and* $\inf(\alpha \cdot S) = \alpha \cdot \inf(S)$

Proof.

1. For α we have two possible cases:

   $\boldsymbol{\alpha = 0}$**.** Then $\alpha \cdot S = \{0\}$ so that $\sup(\alpha \cdot S) = 0 = 0 \cdot \sup(S) = \alpha \cdot \sup(S)$.

   $\boldsymbol{0 < \alpha}$**.** As $\sup(S)$ is a upper bound of S we have that $\forall s \in S\ s \leqslant \sup(S)$. If $x \in \alpha \cdot S$ then there exist a $s \in S$ such that $x = \alpha \cdot s \leqslant_{0 \leqslant \alpha} \alpha \cdot \sup(S)$ proving that

   $$\alpha \cdot \sup(S) \text{ is a upper bound of } \alpha \cdot S$$

   Hence as $\mathbb{R}_{\mathbb{C}}$ is conditional complete [see theorem: 10.18] we have that

   $$\sup(\alpha \cdot S) \text{ exist}$$

   and as the supremum is the lowest upper bound that

   $$\sup(\alpha \cdot S) \leqslant \alpha \cdot \sup(S) \tag{10.24}$$

   Assume that $\sup(\alpha \cdot S) < \alpha \cdot \sup(S)$ then $\frac{1}{\alpha} \cdot \sup(\alpha \cdot S) < \sup(S)$, hence there exist a $s \in S$ such that $\frac{1}{\alpha} \cdot \sup(\alpha \cdot S) < s \Rightarrow \sup(\alpha \cdot S) < \alpha \cdot s$, as $\alpha \cdot s \in \alpha \cdot S$ we have $\alpha \cdot s \leqslant \sup(\alpha \cdot S)$ giving the contradiction $\alpha \cdot s < \alpha \cdot s$. So we must have that $\alpha \cdot \sup(S) \leqslant \sup(\alpha \cdot S)$ which combined with [eq: 10.24] proves that

   $$\sup(\alpha \cdot S) = \sup(\alpha \cdot S)$$

   So in all cases $\sup(\alpha \cdot S)$ exists and $\sup(\alpha \cdot S) = \alpha \cdot \sup(S)$.

2. For α we have two possible cases:

   $\alpha = 0$. Then $\alpha \cdot S = \{0\}$ so that $\inf(\alpha \cdot S) = 0 = 0 \cdot \inf(S) = \alpha \cdot \inf(S)$.

   $0 < \alpha$. As $\inf(S)$ is a lower bound of S we have that $\forall s \in S$ $\inf(S) \leqslant s$. If $x \in \alpha \cdot S$ then there exist a $s \in S$ such that $\alpha \cdot \inf(S) \leqslant_{0 \leqslant \alpha} \alpha \cdot s = x$ proving that

   $$\alpha \cdot \inf(S) \text{ is a upper bound of } \alpha \cdot S$$

   Hence as $\mathbb{R}_{\mathbb{C}}$ is conditional complete [see theorem: 10.18] we have that

   $$\inf(\alpha \cdot S) \text{ exist}$$

   and as the infinum is the highest lower bound that

   $$\alpha \cdot \inf(S) \leqslant \inf(\alpha \cdot S) \tag{10.25}$$

   Assume that $\alpha \cdot \inf(S) < \inf(\alpha \cdot S)$ then $\inf(S) < \frac{1}{\alpha} \cdot \inf(\alpha \cdot S)$, hence there exist a $s \in S$ such that $s < \frac{1}{\alpha} \cdot \inf(\alpha \cdot S) \Rightarrow \alpha \cdot s < \inf(\alpha \cdot S)$, as $\alpha \cdot s \in \alpha \cdot S$ we have $\inf(\alpha \cdot S) \leqslant \alpha \cdot s$ giving the contradiction $\alpha \cdot s < \alpha \cdot s$. So we must have that $\inf(\alpha \cdot S) \leqslant \alpha \cdot \inf(S)$ which combined with [eq: 10.25] proves that

   $$\inf(\alpha \cdot S) = \inf(\alpha \cdot S)$$

   So in all cases $\inf(\alpha \cdot S)$ exists and $\inf(\alpha \cdot S) = \alpha \cdot \inf(S)$.

□

Theorem 10.23. *Let $S, T \subseteq \mathbb{R}_{\mathbb{C}}$ with $S \neq \varnothing \neq T$ then for*

$$S + T = \{\alpha + \beta | \alpha \in S \wedge \beta \in T\}$$

we have

1. *If* $\sup(S), \sup(T)$ *exists then* $\sup(S + T)$ *exist and* $\sup(S + T) = \sup(S) + \sup(T)$
2. *If* $\inf(S), \inf(T)$ *exists then* $\inf(S + T)$ *exist and* $\inf(S + T) = \inf(S) + \inf(T)$

Proof. First as $S \neq \varnothing \neq T$ there exists $s \in S$ and $t \in T$ so that $s + t \in S + T$ hence

$$S + T \neq \varnothing$$

1. Let $q \in S + T$ then $\exists s \in S$ and $\exists t \in T$ such that $q = s + t$, as $s \leqslant \sup(S)$ we have $q = s + t \leqslant \sup(S) + t$, further as $t \leqslant \sup(T)$ it follows that $\sup(S) + t \leqslant \sup(S) + \sup(T)$ giving $q \leqslant \sup(S) + \sup(T)$. So $\sup(S) + \sup(T)$ is a upper bound of $S + T$ which as $S + T \neq \varnothing$ and $\langle \mathbb{R}_{\mathbb{C}}, \leqslant \rangle$ is conditional complete [see theorem: 10.18] proves that

   $$\sup(S + T) \text{ exist and } \sup(S + T) \leqslant \sup(S) + \sup(T) \tag{10.26}$$

   Assume now that $\sup(S + T) < \sup(S) + \sup(T)$ then for $\varepsilon = \sup(S) + \sup(T) - \sup(S + T)$ we have $0 < \varepsilon$. So $-\varepsilon < 0$ and as $0 < 2 \Rightarrow 0 < 2^{-1}$ we have that $-(\varepsilon/2) < 0$. So $\sup(S) - \varepsilon/2 < \sup(S)$ and $\sup(T) - \varepsilon/2 < \sup(T)$. As $\langle \mathbb{R}_{\mathbb{C}}, \leqslant \rangle$ is totally ordered we have by [theorem: 3.71] that there exists $s \in S$ and $t \in T$ such that $\sup(S) - \varepsilon/2 < s$ and $\sup(T) - \varepsilon/2 < t$. So

   $$\begin{aligned} s + t &> \sup(S) - \varepsilon/2 + \sup(T) - \varepsilon/2 \\ &= \sup(S) + \sup(T) - (\varepsilon + \varepsilon)/2 \\ &= \sup(S) + \sup(T) - \varepsilon \\ &= \sup(S) + \sup(T) - \sup(S) - \sup(T) + \sup(S + T) \\ &= \sup(s + t) \end{aligned} \tag{10.27}$$

   As $s + t \in S + T$ we have that $s + t \leqslant \sup(S + T)$ contradicting [eq: 10.27], so the assumption is wrong and we must have $\sup(S) + \sup(T) \leqslant \sup(S + T)$ which combined with [eq: 10.26] proves that

   $$\sup(S + T) = \sup(S) + \sup(T)$$

2. Let $q \in S + T$ then $\exists s \in S$ and $\exists t \in T$ such that $q = s + t$, as $\inf(S) \leqslant s$ we have $\inf(S) + t \leqslant s + t = q$, further as $\inf(T) \leqslant t$ it follows that $\inf(S) + \inf(T) \leqslant \inf(S) + t$ giving $\inf(S) + \inf(T) \leqslant q$. So $\inf(S) + \inf(T)$ is a lower bound of $S + T$ which as $S + T \neq \varnothing$ and $\langle \mathbb{R}_{\mathbb{C}}, \leqslant \rangle$ is conditional complete [see theorem: 10.18] proves that

   $$\inf(S + T) \text{ exist and } \inf(S) + \inf(T) \leqslant \inf(S + T) \tag{10.28}$$

Assume now that $\inf(S)+\inf(T)<\inf(S+T)$ then for $\varepsilon=\inf(S+T)-\inf(S)-\inf(T)$ we have $0<\varepsilon$. As $0<2\Rightarrow 0<2^{-1}$ we have that $0<\varepsilon/2$. So $\inf(S)<\inf(S)+\varepsilon/2$ and $\inf(T)<\inf(T)+\varepsilon/2$. As $\langle\mathbb{R}_{\mathbb{C}},\leqslant\rangle$ is totally ordered we have by [theorem: 3.71] that there exists $s\in S$ and $t\in T$ such that $s<\inf(S)+\varepsilon/2$ and $t<\inf(T)+\varepsilon/2$. So

$$\begin{aligned} s+t &< \inf(S)+\varepsilon/2+\inf(T)+\varepsilon/2 \\ &= \inf(S)+\inf(T)+\varepsilon \\ &= \inf(S)+\inf(T)+\inf(S+T)-\inf(S)-\inf(T) \\ &= \inf(S+T) \end{aligned} \tag{10.29}$$

As $s+t\in S+T$ we have that $\inf(S+T)\leqslant s+t$ contradicting [eq: 10.29], so the assumption is wrong and we must have $\inf(S+T)\leqslant\inf(S)+\inf(T)$ which combined with [eq: 10.28] proves that

$$\inf(S+T)=\inf(S)+\inf(T)$$

□

Corollary 10.24. *Let $S\subseteq\mathbb{R}$ with $S\neq\varnothing$ and $\alpha\in\mathbb{R}$ then for $S+\alpha=\{s+\alpha|s\in S\}$ we have that*

1. *If $\sup(S)$ exists then $\sup(S+\alpha)$ exists and $\sup(S+\alpha)=\sup(S)+\alpha$*
2. *If $\inf(S)$ exists then $\inf(S+\alpha)$ exists and $\inf(S+\alpha)=\inf(S)+\alpha$*

Proof. *First*

$$\begin{aligned} x\in S+\alpha &\Leftrightarrow \exists s\in S \text{ such that } x=s+\varepsilon \\ &\underset{t\in\{\alpha\}\Leftrightarrow t=\alpha}{\Leftrightarrow} \exists s\in S\wedge\exists t\in\{\alpha\} \text{ such that } x=s+t \\ &\Leftrightarrow x\in S+\{\alpha\} \end{aligned}$$

hence we have

$$S+\alpha=S+\{\alpha\}$$

Now we have

1. *If $\sup(S)$ exists then by [theorem 10.23] that $\sup(S+\{a\})$ exist and $\sup(S+\{\alpha\})=\sup(S)+\sup(\{\alpha\})$ which as $S+\{\alpha\}=S+\alpha$ and $\sup(\{\alpha\})=\alpha$ proves that*

$$\sup(S+\alpha) \text{ exist and } \sup(S+\alpha)=\sup(S)+\alpha$$

2. *If $\inf$ exists then by [theorem 10.23] that $\inf(S+\{a\})$ exist and $\inf(S+\{\alpha\})=\inf(S)+\inf(\{\alpha\})$ which as $S+\{\alpha\}=S+\alpha$ and $\inf(\{\alpha\})=\alpha$ proves that*

$$\inf(S+\alpha) \text{ exist and } \inf(S+\alpha)=\inf(S)+\alpha$$

□

We have simular theorems for the minimum and maximum.

Theorem 10.25. *Let $S,T\subseteq\mathbb{R}_{\mathbb{C}}$ with $S\neq\varnothing\neq T$ then for*

$$S+T=\{\alpha+\beta|\alpha\in S\wedge\beta\in T\}$$

we have

1. *If $\max(S),\max(T)$ exists then $\max(S+T)$ exist and $\max(S+T)=\max(S)+\max(T)$*
2. *If $\min(S),\min(T)$ exists then $\min(S+T)$ exist and $\min(S+T)=\min(S)+\min(T)$*

Proof.

1. As $\max(S)$, $\max(T)$ exist we have

$$\max(S)\in S\wedge\max(T)\in S\wedge\forall s\in S\ s\leqslant\max(S)\wedge\forall t\in T\ t\leqslant\max(T)$$

As $\max(S)\in S$ and $\max(T)\in T$ it follows by definition that

$$\max(S)+\max(T)\in S+T \tag{10.30}$$

Further if $x\in S+T$ then there exists a $s\in S$ and a $t\in T$ such that $x=s+t$, hence, as $s\leqslant\max(S)\wedge t\leqslant\max(T)$, we have $x=s+t\leqslant\max(S)+\max(T)$. This proves that

$$\forall x\in S+T \text{ we have } x\leqslant\max(S)+\max(T) \tag{10.31}$$

Finally using the definition of the maximum together with [eqs: 10.30, 10.31] we have that

$$\max(S+T) \text{ exists and } \max(S+T) = \max(S) + \max(T)$$

2. As $\min(S)$, $\min(T)$ exist we have

$$\min(S) \in S \wedge \min(T) \in S \wedge \forall s \in S \ \min(S) \leqslant s \wedge \forall t \in T \ \min(T) \leqslant t$$

As $\min(S) \in S$ and $\min(T) \in T$ it follows by definition that

$$\min(S) + \min(T) \in S + T \tag{10.32}$$

Further if $x \in S+T$ then there exists a $s \in S$ and a $t \in T$ such that $x = s+t$, hence, as $\min(S) \leqslant s \wedge \min(T) \leqslant t$, we have $\min(S) + \min(T) \leqslant s+t = x$. This proves that

$$\forall x \in S+T \text{ we have } \min(S) + \min(T) \leqslant s+t \tag{10.33}$$

Finally using the definition of the minimum together with [eqs: 10.32, 10.33] we have that

$$\min(S+T) \text{ exists and } \min(S+T) = \min(S) + \min(T)$$

□

Corollary 10.26. *Let $\varnothing \neq S \subseteq \mathbb{R}_{\mathbb{C}}$ and $\alpha \in \mathbb{R}_{\mathbb{C}}$ then we have*

1. *If $\max(S)$ exist then $\max(S+\alpha)$ exist and $\max(S+\alpha) = \max(S) + \alpha$*
2. *If $\min(S)$ exist then $\min(S+\alpha)$ exist and $\min(S+\alpha) = \min(S) + \alpha$*

Proof.

1. Define $T = \{\alpha\}$ then $S+T = S+\alpha$ and $\max(T)$ exist and $\max(T) = \alpha$. Hence by]theorem: 10.25] $\max(S+\alpha) = \max(S+T)$ exist and

$$\max(S+\alpha) = \max(S) + \max(\{\alpha\}) = \max(S) + \alpha$$

2. Define $T = \{\alpha\}$ then $S+T = S+\alpha$ and $\min(T)$ exist and $\min(T) = \alpha$. Hence by]theorem: 10.25] $\min(S+\alpha) = \min(S+T)$ exist and

$$\min(S+\alpha) = \min(S) + \min(\{\alpha\}) = \min(S) + \alpha$$

□

Theorem 10.27. *Let $\varnothing \neq S \subseteq \mathbb{R}$ and $\alpha \in \mathbb{R}$ such that $0 \leqslant \alpha$ then for $\alpha \cdot S = \{\alpha \cdot x | x \in S\}$ we have:*

1. $\max(\alpha \cdot S) = \alpha \cdot \max(S)$
2. $\min(\alpha \cdot S) = \alpha \cdot \min(S)$

Proof.

1. First as $\max(S) \in S$ we have that $\alpha \cdot \max(S) \in \alpha \cdot S$. If $x \in \alpha \cdot S$ then there exist a $s \in S$ such that $x = \alpha \cdot s$, as $s \leqslant \max(S)$ and $0 \leqslant \alpha$ we have by [theorem: 10.14] that $x = \alpha \cdot s \leqslant \alpha \cdot \max(S)$ hence $\max(\alpha \cdot S) = \alpha \cdot \max(S)$.
2. First as $\min(S) \in S$ we have that $\alpha \cdot \min(S) \in \alpha \cdot S$. If $x \in \alpha \cdot S$ then there exist a $s \in S$ such that $x = \alpha \cdot s$, as $\min(S) \leqslant s$ and $0 \leqslant \alpha$ we have by [theorem: 10.14] that $\alpha \cdot \min(S) \leqslant \alpha \cdot s = x$ hence $\min(\alpha \cdot S) = \alpha \cdot \min(S)$. □

Theorem 10.28. *Let $x, y \in \mathbb{Z}_{\mathbb{C}}$ with $x < y$ then*

1. $x + 1 \leqslant y$
2. $x \leqslant y - 1$

Proof. By [theorems: 10.5 and 10.9] we have that

$$i_{\mathbb{Z} \rightarrow \mathbb{C}}: \langle \mathbb{Z}, \leqslant \rangle \rightarrow \langle \mathbb{Z}_{\mathbb{C}}, \leqslant \rangle \text{ is a order isomorphism} \tag{10.34}$$

$$i_{\mathbb{Z} \rightarrow \mathbb{Z}_{\mathbb{C}}}: \langle \mathbb{Z}, +, \cdot \rangle \rightarrow \langle \mathbb{Z}_{\mathbb{C}}, +, \cdot \rangle \text{ is a field isomorphism} \tag{10.35}$$

Let $x, y \in \mathbb{Z}_{\mathbb{C}}$ then by [eq: 10.34] there exists $x', y' \in \mathbb{Z}$ such that

$$x = i_{\mathbb{Z} \rightarrow \mathbb{C}}(x') \wedge y = i_{\mathbb{Z} \rightarrow \mathbb{C}}(y') \wedge x' < y' \tag{10.36}$$

1. Using [theorem: 7.28] we have that $x' + 1 \leqslant y'$ so that

$$x + 1 \underset{\text{[eq: 10.35]}}{=} i_{\mathbb{Z} \rightarrow \mathbb{C}}(x') + i_{\mathbb{Z} \rightarrow \mathbb{C}}(1) \underset{\text{[eq: 10.34]}}{=} i_{\mathbb{Z} \rightarrow \mathbb{C}}(x'+1) \leqslant_{\text{[eq: 10.34]}} i_{\mathbb{Z}}(y') = y$$

proving that

$$x+1 \leqslant y$$

2. As by (1) we have that $x+1 \leqslant y$ hence $x=(x+1)+(-1) \leqslant y+(-1)=y-1$ □

Theorem 10.29. (Archimedean Property) *If $x, y \in \mathbb{R}_{\mathbb{C}}$ with $0<x$ then $\exists n \in \mathbb{N}_{0,\mathbb{C}}$ such that $y<n \cdot x$*

Proof. For y we have either

$\boldsymbol{y \leqslant 0}$**.** Then for $n=1$ we have $y \leqslant 0<x=1 \cdot x=n \cdot x$ proving that $y<n \cdot x$

$\boldsymbol{0<y}$**.** We prove this by contradiction. Assume that $\forall n \in \mathbb{N}_{0,\mathbb{C}}$ we have $n \cdot x \leqslant y$. Define

$$A=\{n \cdot x | n \in \mathbb{N}_{0,\mathbb{C}}\}$$

then $\forall t \in A$ we have $t \leqslant y$ so that y is a upper bound of A and as $x=1 \cdot x \in A$ $A \neq \varnothing$. By [theorem: 10.18] $\langle \mathbb{R}_{\mathbb{C}}, \leqslant \rangle$ is conditional complete so that $\sup(A)$ exists. As $0<x$ we have $-x<0$ so that $\sup(A)-x<\sup(A)$, given that by [theorem: 10.8] $\langle \mathbb{R}, \leqslant \rangle$ is totally ordered, we have by [theorem: 3.71] that $\exists t \in A$ such that $\sup(A)-x<t$. Using the definition of A we have then that $\exists n \in \mathbb{N}_{0,\mathbb{C}}$ such that $t=n \cdot x$ hence $\sup(A)-x<n \cdot x$, so that

$$\sup(A)<n \cdot x+x=(n+1) \cdot x \tag{10.37}$$

As $n+1 \in \mathbb{N}_{0,\mathbb{C}}$ we have that $(n+1) \cdot x \in A$ so that $(n+1) \cdot x \leqslant \sup(A)$ contradicting [eq: 10.37]. So our assumption is wrong hence

$$\exists n \in \mathbb{N}_{0,\mathbb{R}} \text{ such that } y<n \cdot x$$ □

Corollary 10.30. *Let $x \in \mathbb{R}_{\mathbb{C}}$ then we have*

1. *$\exists n \in \mathbb{N}_{0,\mathbb{C}}$ such that $x<n$*
2. *$\exists n \in \mathbb{N}_{0,\mathbb{C}}$ such that $-n<x$*
3. *$\exists n \in \mathbb{Z}_{\mathbb{C}}$ such that $n \leqslant x<n+1$*
4. *$\exists n \in \mathbb{Z}_{\mathbb{C}}$ such that $n<x \leqslant n+1$*
5. *$\exists n \in \mathbb{Z}_{\mathbb{C}}$ such that $n-1 \leqslant x<n$*
6. *$\exists n \in \mathbb{Z}_{\mathbb{C}}$ such that $n-1<x \leqslant n$*
7. *If $0<x$ then $\exists n \in \mathbb{N}_{\mathbb{C}}$ such that $0<1/n<x$*

Proof.

1. As $0<1$ [see corollary: 10.10] we have by the Archimedean property [see theorem: 10.29] that there exist a $n \in \mathbb{N}_{0,\mathbb{C}}$ such that $x<n \cdot 1=n$
2. By (1) there exists a $n \in \mathbb{N}_0$ such that $-x<n$ hence $-n<-(-x)=x$
3. By (1) $A=\{n \in \mathbb{N}_{0,\mathbb{C}} | x<n\} \neq \varnothing$ and by the well ordering of $\langle \mathbb{N}_{0,\mathbb{C}}, \leqslant \rangle$ [see theorem: 10.19] there exists a least element $m \in A$. As $m-1<m$ [see corollary: 10.15] we have that $m-1 \notin A$ hence $m-1 \leqslant x$ and as $m \in A$ we have also $x<m$. Take $n=m-1$ then $n \leqslant x<m=n+1$.
4. Using (2) there exist a $m \in \mathbb{N}_{0,\mathbb{C}}$ such that $m \leqslant x<m+1$. As $m \leqslant x$ we have the following possibilities to consider:

   $\boldsymbol{m=x}$**.** Take then $n=m-1$ so that $n+1=m$ then we have $n<x \leqslant n+1$

   $\boldsymbol{m<x}$**.** Take then $n=m$ so that $n<x \leqslant n+1$
5. Using (2) there exist a $m \in \mathbb{N}_{0,\mathbb{C}}$ such that $m \leqslant x<m+1$, take then $n=m+1$ so that $m=n-1$, hence $n-1 \leqslant x<n$.
6. Using (3) there exist a $m \in \mathbb{N}_{0,\mathbb{C}}$ such that $m<x \leqslant m+1$, take then $n=m+1$ so that $m=n-1$, hence $n-1<x \leqslant n$.
7. Using the Archimedean Property [see theorem: 10.29] there exists a $n \in \mathbb{N}_{0,\mathbb{C}}$ such that $1<n \cdot x$. If $n=0$ we would have $1<0$ a contradiction so $0 \neq n$. Using [theorem: 10.11] we have that $0 \leqslant n$, so that $0<n$. Applying then [theorem: 10.14] we have $0<n^{-1}=1/n$ which using [theorems: 10.14] on $1<n \cdot x$ gives $0<1/n=1 \cdot n^{-1}<(n \cdot x) \cdot n^{-1}=x$. □

Corollary 10.31. *Let $x, y \in \mathbb{R}_{\mathbb{C}}$ then we have*

1. *If $\forall n \in \mathbb{N}_{\mathbb{C}}$ we have that $x \leqslant y+1/n$ then $x \leqslant y$*

2. *If* $\forall \varepsilon \in \mathbb{R}^+_{\mathbb{C}}$ *we have that* $x \leqslant y + \varepsilon$ *then* $x \leqslant y$
3. *Let* $a \in \mathbb{R}^+_{0,\mathbb{C}}$ *then if* $\forall \varepsilon \in \mathbb{R}^+_{\mathbb{C}}$ *we have that* $x \leqslant y + \varepsilon \cdot a$ *then* $x \leqslant y$
4. *Let* $a \in \mathbb{R}^+_{0,\mathbb{C}}$ *then if* $\forall n \in \mathbb{N}_{\mathbb{C}}$ *we have that* $x \leqslant y + a/n$ *then* $x \leqslant y$

Proof.

1. Assume that $y < x$ then we have $0 < x - y$ so by [corollary: 10.30] there exist a $n \in \mathbb{N}_{\mathbb{C}}$ such that $1/n < x - y$. As we also have that $x \leqslant y + 1/n \Rightarrow x - y \leqslant 1/n$ we reach the contradiction $1/n < 1/n$. So the assumption is wrong and we must have that
$$x \leqslant y$$
2. Assume that $y < x$ then we have $0 < x - y$ so by [corollary: 10.30] there exist a $n \in \mathbb{N}_{\mathbb{C}}$ such that $1/n < x - y$. Take $\varepsilon = 1/n$ then we have also $x \leqslant y + \varepsilon \Rightarrow x - y \leqslant \varepsilon = \frac{1}{n}$ so we reach the contradiction $1/n < 1/n$. So the assumption is wrong and we must have that
$$x \leqslant y$$
3. As $a \in \mathbb{R}^+_{0,\mathbb{C}}$ we have two possibilities to consider:

   $\boldsymbol{a = 0}$**.** Then if we take $\varepsilon = 1 \in \mathbb{R}^+_{\mathbb{C}}$ we have from $x \leqslant y + a \cdot \varepsilon = y + 0 \cdot 1 = y$ that $x \leqslant y$

   $\boldsymbol{0 < a}$**.** Then $0 < a^{-1} = 1/a$, take $\varepsilon \in \mathbb{R}^+_{\mathbb{C}}$ then $0 < \varepsilon/a$, hence $\delta = \varepsilon/a \in \mathbb{R}^+_{\mathbb{C}}$, by the assumption we have $x \leqslant y + a \cdot \delta = y + (\varepsilon/a) \cdot a = y + \varepsilon$. So we have $\forall \varepsilon \in \mathbb{R}^+_{\mathbb{C}}$ that $x \leqslant y + \varepsilon$ which by (2) proves that
$$x \leqslant y$$
4. As $a \in \mathbb{R}^+_{0,\mathbb{C}}$ we have two possibilities to consider:

   $\boldsymbol{a = 0}$**.** Then if we take $n = 1 \in \mathbb{N}_{\mathbb{C}}$ we have from $x \leqslant y + a/n = y + (1/1) \cdot 0 = y$ that $x \leqslant y$

   $\boldsymbol{0 < a}$**.** Take $n \in \mathbb{N}_{\mathbb{C}}$ then
$$0 < n \underset{\text{[theorems: 10.14]}}{\Rightarrow} 0 < a \cdot n \underset{\text{[theorems: 10.14]}}{\Rightarrow} 0 < (a \cdot n)^{-1} = a^{-1} \cdot n^{-1}$$
   so that by [corollary: 10.30] there exists a $m \in \mathbb{N}_{\mathbb{R}}$ such that $1/m < a^{-1} \cdot n^{-1}$. By assumption we have now $x \leqslant y + a/m <_{\text{[theorem: 10.14]}} y + a \cdot (a^{-1} \cdot n^{-1}) = y + 1/n$. So we have $\forall n \in \mathbb{N}_{\mathbb{R}}$ that $x \leqslant y + 1/n$ which by (1) implies that
$$x \leqslant y$$
□

The next theorem shows how the embedded rational numbers are dense in the set of real numbers.

Theorem 10.32. (Density Theorem) *If* $x \in y \in \mathbb{R}_{\mathbb{C}}$ *such that* $x < y$ *then we have*

1. $\exists q \in \mathbb{Q}_{\mathbb{C}}$ *such that* $x < q < y$
2. $\exists r \in \mathbb{R}_{\mathbb{C}} \setminus \mathbb{Q}_{\mathbb{C}}$ *such that* $x < r < y$

In other words if we have two different real numbers then we can always find a rational number and a irrational number that lies between the two real numbers.

Proof.

1. We first prove the case for $0 \leqslant x$, then we have either

   $\boldsymbol{0 < x}$**.** From $x < y$ we have that $0 < y - x$ so that by [theorem: 10.14] $0 < (y - x)^{-1}$. Using [corollary: 10.30] there exists a $n \in \mathbb{N}_{0,\mathbb{C}}$ such that $0 < (y - x)^{-1} < n$. As $0 < y - x$ we have that $1 = (y - x)^{-1} \cdot (y - x) < n \cdot (y - x) = n \cdot y - n \cdot x$ so that
$$1 + n \cdot x < n \cdot y \tag{10.38}$$
   and from $0 < (y - x)^{-1} < n$ that
$$0 < n \underset{\text{[theorem: 10.14]}}{\Rightarrow} 0 < n^{-1} = 1/n \tag{10.39}$$
   Using [corollary: 10.30] there exist a $m \in \mathbb{N}_{0,\mathbb{R}}$ such that
$$m - 1 \leqslant n \cdot x < m \tag{10.40}$$

Multiplying by n^{-1} gives then by [theorem: 10.14] that $x=(n\cdot x)\cdot n^{-1}<m\cdot n^{-1}=m/n$. Take $q=m/n$ then, as n, $m\in\mathbb{N}_{0,\mathbb{C}}\subseteq\mathbb{Z}_{\mathbb{C}}$ and $n\neq 0$ we have by [theorem: 10.6] that $q\in\mathbb{Q}_{\mathbb{C}}$. So we have

$$x<q\in\mathbb{Q}_{\mathbb{C}} \tag{10.41}$$

From [eq: 10.40] we have that $m\leqslant n\cdot x+1<n\cdot y$ so that $m<n\cdot y$ and by multiplying both sides by n^{-1} that $q=m/n<y$. Combining this with [eq: 10.41] gives finally

$$x<q<y \text{ where } q\in\mathbb{Q}_{\mathbb{C}}$$

$\boldsymbol{x=0.}$ Then $0<y$ and by [corollary: 10.30] there exists a $n\in\mathbb{N}_{\mathbb{C}}$ such that $0<1/n<y$. As $1,n\in\mathbb{N}_{0,\mathbb{C}}\subseteq\mathbb{Z}_{\mathbb{C}}$ and $n\neq 0$ we have by [theorem: 10.6] that $1/n\in\mathbb{Q}_{\mathbb{C}}$ hence if we take $q=1/n$ we have that

$$x<q<y \text{ where } q\in\mathbb{Q}_{\mathbb{C}}$$

So we have proved that

$$\forall x,y\in\mathbb{R}_{\mathbb{C}} \text{ with } 0\leqslant x\wedge x<y \text{ there exist a } q\in\mathbb{Q}_{\mathbb{C}} \text{ such that } x<q<y \tag{10.42}$$

So the only case left to prove is where $x<0$. Then for y we have either

$\boldsymbol{0<y.}$ Then if we take $q=0\in\mathbb{N}_{0,\mathbb{C}}\subseteq\mathbb{Z}_{\mathbb{C}}\subseteq\mathbb{Q}_{\mathbb{C}}$ so that

$$x<q<y \text{ with } q\in\mathbb{Q}_{\mathbb{C}}$$

$\boldsymbol{y\leqslant 0.}$ Then $x<y\leqslant 0$ or using [theorem: 10.14] $0\leqslant -y<-x$. Take $x'=-y$ and $y'=-x$ then $0\leqslant x'<y'$ and by [eq: 10.42] there exists a $q'\in\mathbb{Q}_{\mathbb{C}}$ such that $x'<q'<y'$. So $-y<q'<-x$ or $x<-q'<y$. As $\langle\mathbb{Q}_{\mathbb{C}},+\rangle$ is a sub group of $\langle\mathbb{C},+\rangle$ it follows that $-q'\in\mathbb{Q}_{\mathbb{C}}$ hence if we take $q=-q'$ then we have

$$x<q<y \text{ where } q\in\mathbb{Q}_{\mathbb{C}}$$

2. Finally we prove the case for the irrational numbers. By [theorem: 10.20] $\mathbb{R}_{\mathbb{C}}\setminus\mathbb{Q}_{\mathbb{C}}\neq\varnothing$ so there exist a $z\in\mathbb{R}\setminus\mathbb{Q}_{\mathbb{R}}$, then from $x<y$ we have that $x+z<y+z$. Using (1) there exist a $r\in\mathbb{Q}_{\mathbb{R}}$ such that $x+z<r<y+z$ or $x<r-z<y$. Take $q=r-z$ then we have $x<q<y$. Assume that $q\in\mathbb{Q}_{\mathbb{R}}$ then as $q,r\in\mathbb{Q}_{\mathbb{R}}$ and $\langle\mathbb{Q}_{\mathbb{R}},+\rangle$ is a sub group of $\langle\mathbb{R},+\rangle$ we have that $z=r-q\in\mathbb{Q}_{\mathbb{R}}$ contradicting $z\in\mathbb{R}\setminus\mathbb{Q}_{\mathbb{R}}$. So we must have that $q\in\mathbb{R}\setminus\mathbb{Q}_{\mathbb{R}}$ and we conclude that

$$x<q<y \text{ where } x\in\mathbb{Q}_{\mathbb{R}}$$ □

10.2.3 Recursion and mathematical induction in $\mathbb{C}$

The embedding $\mathbb{N}_{0,\mathbb{C}}$ of $\mathbb{N}_0$ in $\mathbb{C}$ is important because it allows use to extend recursion and induction using $\mathbb{N}_{0,\mathbb{C}}$ instead of $\mathbb{N}_0$.

Definition 10.33. *Let $n,m\in\mathbb{N}_{0,\mathbb{C}}$ then we define*

$$\{n,\ldots,m\}=\{i\in\mathbb{N}_{0,\mathbb{C}}|n\leqslant i\leqslant m\}\subseteq\mathbb{N}_{0.\mathbb{C}}$$

$$\{n,,\ldots,\infty\}=\{i\in\mathbb{N}_{0,\mathbb{C}}|n\leqslant i\}\subseteq\mathbb{N}_{0,C\mathbb{C}}$$

Lemma 10.34. *Let $n,m\in\mathbb{N}_{0,\mathbb{C}}$ then we have*

1. $i_{\mathbb{N}_0\rightarrow\mathbb{C}}(\{n,\ldots,m\})=\{i_{\mathbb{N}_0\rightarrow\mathbb{C}}(n),\ldots,i_{\mathbb{N}_0\rightarrow\mathbb{C}}(m)\}$
2. $i_{\mathbb{N}_0\rightarrow\mathbb{C}}(\{m,\ldots,\infty\})=\{i_{\mathbb{N}_0\rightarrow\mathbb{C}}(n),\ldots,\infty\}$
3. $\mathbb{N}_{0,\mathbb{C}}=\{0,\ldots,\infty\}$

Proof. Using [theorems: 10.9] we have that

$$i_{\mathbb{N}_0\rightarrow\mathbb{C}}\colon\langle\mathbb{N}_0,\leqslant\rangle\rightarrow\langle\mathbb{N}_{0,\mathbb{C}},\leqslant\rangle \text{ is a order isomorphism} \tag{10.43}$$

then we have:

1. If $y\in i_{\mathbb{N}_0\rightarrow\mathbb{C}}(\{n,\ldots,m\})$ then $\exists x\in\{n,\ldots,m\}$ such that $y=i_{\mathbb{N}_0\rightarrow\mathbb{C}}(x)$. As $x\leqslant n\wedge x\leqslant m$ we have by [eq: 10.43] that $i_{\mathbb{N}_0\rightarrow\mathbb{C}}(n)\leqslant i_{\mathbb{N}_0\rightarrow\mathbb{C}}(x)\wedge i_{\mathbb{N}_0\rightarrow\mathbb{C}}(x)\leqslant i_{\mathbb{N}_0\rightarrow\mathbb{C}}(m)$ proving that $i_{\mathbb{N}_0\rightarrow\mathbb{C}}(n)\leqslant y\wedge y\leqslant i_{\mathbb{N}_0\rightarrow\mathbb{C}}(m)$ hence $y\in\{i_{\mathbb{N}_0\rightarrow\mathbb{C}}(n),\ldots,i_{\mathbb{N}_0\rightarrow\mathbb{C}}(m)\}$, so we have

$$i_{\mathbb{N}_0\rightarrow\mathbb{C}}(\{n,\ldots,m\})\subseteq\{i_{\mathbb{N}_0\rightarrow\mathbb{C}}(n),\ldots,i_{\mathbb{N}_0\rightarrow\mathbb{C}}(m)\} \tag{10.44}$$

Further if $y \in \{i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(n), \ldots, i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(m)\} \subseteq \mathbb{N}_{0,\mathbb{C}}$ then $i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(n) \leqslant y \wedge y \leqslant i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(m)$ and by [eq: 10.43] $\exists x \in \mathbb{N}_0$ such that $y = i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(x)$, hence

$$i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(n) \leqslant i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(x) \wedge i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(x) \leqslant i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(m).$$

Using [eq: 10.43] again we have then that $n \leqslant x \wedge x \leqslant m$ so that $x \in \{n, \ldots, m\}$ proving that $y \in i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(\{n, \ldots, m\})$. So we have $\{i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(n), \ldots, i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(m)\} \subseteq i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(\{n, \ldots, m\})$ which combined with [eq: 10.44] gives

$$i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(\{n, \ldots, m\}) = \{i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(n), \ldots, i_{\mathbb{N}_0 \rightarrow \mathbb{R}}(m)\}$$

2. If $y \in i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(\{n,\ldots,\infty\})$ then $\exists x \in \{n,\ldots,\infty\}$ such that $y = i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(x)$. As $n \leqslant x$ we have by [eq: 10.43] that $i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(n) \leqslant i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(x) = y$ hence $y \in \{i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(n), \ldots, \infty\}$, so we have

$$i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(\{n, \ldots, \infty\}) \subseteq \{i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(n), \ldots, \infty\} \tag{10.45}$$

Further if $y \in \{i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(n), \ldots, \infty\} \subseteq \mathbb{N}_{0,\mathbb{C}}$ then $i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(n) \leqslant y$ and by [eq: 10.43] $\exists x \in \mathbb{N}_0$ such that $y = i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(x)$, hence $i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(n) \leqslant i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(x)$ Using [eq: 10.43] again we have then that $n \leqslant x$ so that $x \in \{n, \ldots, \infty\}$ proving that $y \in i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(\{n, \ldots, \infty\})$. So we have $\{i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(n), \ldots, \infty\} \subseteq i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(\{n, \ldots, \infty\})$ which combined with [eq: 10.45] gives

$$i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(\{n, \ldots, \infty\}) = \{i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(n), \ldots, \infty\}$$

3. Using [note: 5.81] we have that $\mathbb{N}_0 = \{0, \ldots, \infty\}$ so that

$$\mathbb{N}_{0,\mathbb{C}} \underset{\text{[eq 10.43]}}{=} i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(\mathbb{N}_0) = i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(\{0, \ldots, \infty\}) \underset{(2)}{=} \{i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(0), \ldots, \infty\} = \{0, \ldots, \infty\}$$ □

Next we state Mathematical Induction for $\mathbb{N}_{0,\mathbb{C}}$ using Mathematical Induction using $\mathbb{N}_0$.

Theorem 10.35. *Let $k \in \mathbb{N}_{0,\mathbb{C}}$ and $S \subseteq \{k, \ldots, \infty\}$ such that*

1. *$k \in S$*
2. *If $n \in S$ then $n + 1 \in S$*

then $S = \{k, \ldots, \infty\}$

Proof. Using [theorems: 10.5, 3.55] we have that

$$i_{\mathbb{N}_0 \rightarrow \mathbb{C}}: \langle \mathbb{N}_0, + \rangle \rightarrow \langle \mathbb{N}_{0,\mathbb{C}}, + \rangle \text{ is a group isomorphisms and } i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(1) = 1 \tag{10.46}$$

Define $T = (i_{\mathbb{N}_0 \rightarrow \mathbb{C}})^{-1}(S)$ and take $k' = (i_{\mathbb{N}_0 \rightarrow \mathbb{C}})^{-1}(k)$ then we have

$\boldsymbol{k' \in T}$. As $i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(k') = i_{\mathbb{N}_0 \rightarrow \mathbb{C}}((i_{\mathbb{N}_0 \rightarrow \mathbb{C}})^{-1}(k)) = k \in S$ it follows that $k' \in T$

$\boldsymbol{n \in T \Rightarrow n + 1 \in T}$. As $n \in T = (i_{\mathbb{N}_0 \rightarrow \mathbb{C}})^{-1}(S)$ we have that $n' = i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(n) \in S$, by the hypothesis we have then that $n' + 1 \in S$. Hence

$$i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(n + 1) \underset{\text{[eq: 10.46]}}{=} i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(n) + i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(1) \underset{\text{[eq: 10.46]}}{=} n' + 1 \in S$$

so that $n + 1 \in (i_{\mathbb{N}_0 \rightarrow \mathbb{R}})^{-1}(S) = T$.

Using [theorem: 5.83] it follows that $T = \{k', \ldots \infty\}$ so

$$i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(\{k', \ldots, \infty\}) = i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(T) = i_{\mathbb{N}_0 \rightarrow \mathbb{C}}((i_{\mathbb{N}_0 \rightarrow \mathbb{C}})^{-1}(S)) \underset{\text{[eq: 10.46] and [theorem: 2.55]}}{=} S \tag{10.47}$$

Next

$$\begin{aligned} i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(\{k', \ldots, \infty\}) & \underset{\text{[lemma: 10.34]}}{=} \{i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(k'), \ldots, \infty\} \\ & = \{k, \ldots, \infty\} \end{aligned}$$

which combined with [eq: 10.47] gives finally

$$S = \{k, \ldots, \infty\}$$ □

As a example of using induction we have the following theorem.

Theorem 10.36. *Let $\langle A, \leqslant \rangle$ be a partial ordered set then we have*

1. *If $n \in \mathbb{N}_0$ and $\{x_i\}_{i \in \{n, \ldots, \infty\}} \subseteq A$ a sequence such that $\forall i \in \{n,\ldots,\infty\}$ $x_i \leqslant x_{i+1}$ then $\forall k \in \{n,\ldots,\infty\}$ we have $\forall i \in \{0,\ldots,k\}$ that $x_i \leqslant x_k$*

2. *If $n, m \in \mathbb{N}_0$ with $n \leqslant m$ and $\{x_i\}_{i \in \{n, \ldots, m\}} \subseteq A$ a finite family such that $\forall i \in \{n, \ldots, m-1\}$ $x_i \leqslant x_{i+1}$ then $\forall k \in \{n, \ldots, m\}$ we have $\forall i \in \{n, \ldots, k\}$ we have $x_i \leqslant x_k$.*

Proof.

1. We prove this by induction, take

$$S = \{k \in \{n, \ldots, \infty\} | \forall i \in \{n, \ldots, k\} \text{we have } x_i \leqslant x_k\}$$

then we have:

$n \in S$. If $i \in \{n, \ldots, n\}$ then $i = n$ so that $x_i = x_n \leqslant x_n$ proving that $n \in S$.

$k \in S \Rightarrow k+1 \in S$. Let $i \in \{n, \ldots, k+1\}$ then for i we have either:

$i = k+1$. Then $x_i = x_{k+1} \leqslant x_{k+1}$

$i \in \{n, \ldots, k\}$. Then as $k \in S$ we have $x_i \leqslant x_k$ and s by the hypothesis $x_k \leqslant x_{k+1}$ it follows that $x_i \leqslant x_{k+1}$.

so in all cases we have $x_i \leqslant x_{k+1}$ which proves that $k+1 \in S$

Mathematical induction [theorem: 10.35] proves then that $S = \{n, \ldots, \infty\}$ so $\forall k \in \{n, \ldots, \infty\}$ we have $\forall i \in \{n, \ldots, k\}$ that $x_i \leqslant x_k$.

2. This also proved by induction, take

$$S = \{k \in \{n, \ldots, \infty\} | k \in \{m+1, ., \infty\} \vee \forall i \in \{n, \ldots, k\} \text{ we have } x_i \leqslant x_k\}$$

then we have:

$n \in S$. If $i \in \{n, \ldots, n\}$ then $i = n$ and $x_i = x_n \leqslant x_n$ proving that $n \in S$.

$k \in S \Rightarrow k+1 \in S$. For $k+1$ we have either:

$k+1 \in \{m+1, \ldots, \infty\}$. Then $k+1 \in S$

$k+1 \notin \{m+1, \ldots, \infty\}$. Then $k < k+1 < m+1$ so that $k \notin \{m+1, \ldots, \infty\}$. For $i \in \{n, \ldots, k+1\}$ we have either:

$i = k+1$. Then $x_i = x_{k+1} \leqslant x_{k+1}$

$i \in \{n, \ldots, k\}$. Then as $k \in S$ and $k \notin \{m+1, \ldots, \infty\}$ we must have that $x_i \leqslant x_k$ which as by the hypothesis $x_k \leqslant x_{k+1}$ proves that $x_i \leqslant x_{k+1}$.

so in all cases we have $x_i \leqslant x_{k+1}$ proving that $k+1 \in S$.

Using mathematical induction it follows then that $S = \{n, \ldots, \infty\}$. So if $k \in \{n, \ldots, m\}$ then $k \in S$ and $k \notin \{m+1, \ldots, \infty\}$ so that $\forall i \in \{n, \ldots, k\}$ we have $x_i \leqslant x_k$. □

We turn now to recursion.

Theorem 10.37. *Let A be a set, $a \in A$ and $f: A \rightarrow A$ a function then there exist a **unique** function*

$$\lambda: \mathbb{N}_{0,\mathbb{C}} \rightarrow A$$

such that:

1. *$\lambda(0) = a$*
2. *$\forall n \in \mathbb{N}_{0,\mathbb{R}}$ we have $\lambda(n+1) = f(\lambda(n))$*

Proof. Using [theorems: 10.5,4.24] we have that

$$i_{\mathbb{N}_0 \rightarrow \mathbb{C}}: \langle \mathbb{N}_0, + \rangle \rightarrow \langle \mathbb{N}_{0,\mathbb{C}}, + \rangle \text{ and } (i_{\mathbb{N}_0 \rightarrow \mathbb{C}})^{-1}: \langle \mathbb{N}_{0,\mathbb{C}}, + \rangle \rightarrow \langle \mathbb{N}_0, + \rangle \text{ are group isomorphisms} \tag{10.48}$$

and

$$i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(1) = 1 \tag{10.49}$$

Now by [theorem: 5.84] there exist a **unique** function $\beta: \mathbb{N}_0 \rightarrow A$ such that

1. $\beta(0) = a$
2. $\forall n \in \mathbb{N}_0$ we have $\beta(n+1) = f(\beta(n))$

Define $\lambda: \mathbb{N}_{0,\mathbb{C}} \rightarrow A$ by $\lambda = \beta \circ (i_{\mathbb{N}_0 \rightarrow \mathbb{C}})^{-1}$ then we have:

1. $\lambda(0) = (\beta \circ (i_{\mathbb{N}_0 \rightarrow \mathbb{C}})^{-1})(0) = \beta((i_{\mathbb{N}_0 \rightarrow \mathbb{C}})^{-1}(0)) = \beta(0) = a$

2. Let $n \in \mathbb{N}_{0,\mathbb{C}}$ and take $n' = (i_{\mathbb{N}_0 \rightarrow \mathbb{C}})^{-1}(n) \in \mathbb{N}_0$ then

$$(i_{\mathbb{N}_0 \rightarrow \mathbb{C}})^{-1}(n+1) = (i_{\mathbb{N}_0 \rightarrow \mathbb{C}})^{-1}(n) + (i_{\mathbb{N}_0 \rightarrow \mathbb{C}})^{-1}(1) \underset{\text{[eq: 10.49]}}{=} (i_{\mathbb{N}_0 \rightarrow \mathbb{R}})^{-1}(n) + 1 = n' + 1$$

So that

$$\begin{array}{rcl} \lambda(n+1) & = & (\beta \circ (i_{\mathbb{N}_0 \rightarrow \mathbb{C}})^{-1})(n+1) \\ & = & \beta((i_{\mathbb{N}_0 \rightarrow \mathbb{C}})^{-1}(n+1)) \\ & = & \beta(n'+1) \\ & = & f(\beta(n')) \\ & = & f(\beta((i_{\mathbb{N}_0 \rightarrow \mathbb{C}})^{-1}(n))) \\ & = & f(\lambda(n)) \end{array}$$

Hence we have proved the existence of $\lambda: \mathbb{N}_{0,\mathbb{C}} \rightarrow A$.

Next we prove uniqueness, so assume that there is another $\gamma: \mathbb{N}_{0,\mathbb{C}} \rightarrow A$ such that $\gamma(0) = a$ and $\gamma(n+1) = f(\gamma(n))$. Then $(\gamma \circ i_{\mathbb{N}_0 \rightarrow \mathbb{C}})$: $\mathbb{N}_0 \rightarrow A$ is such that

$$(\gamma \circ i_{\mathbb{N}_0 \rightarrow \mathbb{C}})(0) = \gamma(i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(0)) = \gamma(0) = a$$

and

$$\begin{array}{rcl} (\gamma \circ i_{\mathbb{N}_0 \rightarrow \mathbb{R}})(n+1) & = & \gamma(i_{\mathbb{N}_0 \rightarrow \mathbb{R}}(n+1)) \\ & = & \gamma(i_{\mathbb{N}_0 \rightarrow \mathbb{R}}(n) + i_{\mathbb{N}_0 \rightarrow \mathbb{R}}(1)) \\ & = & \gamma(i_{\mathbb{N}_0 \rightarrow \mathbb{R}}(n) + 1) \\ & = & f(\gamma(i_{\mathbb{N}_0 \rightarrow \mathbb{R}}(n))) \\ & = & f((\gamma \circ i_{\mathbb{N}_0 \rightarrow \mathbb{R}})(n)) \end{array}$$

As β is unique we have by the above that $\beta = \gamma \circ i_{\mathbb{N}_0 \rightarrow \mathbb{R}}$, so that

$$\lambda = \beta \circ (i_{\mathbb{N}_0 \rightarrow \mathbb{C}})^{-1} = (\gamma \circ i_{\mathbb{N}_0 \rightarrow \mathbb{C}}) \circ (i_{\mathbb{N}_0 \rightarrow \mathbb{C}})^{-1} = \gamma \circ (i_{\mathbb{N}_0 \rightarrow \mathbb{C}} \circ (i_{\mathbb{N}_0 \rightarrow \mathbb{C}})^{-1}) = \gamma \circ \mathrm{id}_{\mathbb{N}_{0,\mathbb{C}}} = \gamma$$

proving uniqueness. □

Theorem 10.38. (Iteration) *Let A be a set and $f: A \rightarrow A$ a function then $\forall n \in \mathbb{N}_{0,\mathbb{C}}$ there exist a function*

$$(f)^n: A \rightarrow A$$

such that:

1. $(f)^0 = \mathrm{Id}_A$
2. $(f)^{n+1} = f \circ (f)^n$

Proof. Let $a \in A$ and use the recursion [theorem: 10.37] to find a **unique** function

$$\lambda_a: \mathbb{N}_{0,\mathbb{R}} \rightarrow A \text{ such that } \lambda_a(0) = a \text{ and } \forall n \in \mathbb{N}_{0,\mathbb{R}}\ \lambda_a(n+1) = f(\lambda_a(n))$$

Define now

$$(f)^n: A \rightarrow A \text{ where } (f)^n(a) = \lambda_a(n)$$

Then we have

1. $\forall a \in A$ we have that $(f)^0(a) = \lambda_a(0) = a$ so that
$$(f)^0 = \mathrm{Id}_A$$
2. $\forall a \in A$ we have that $(f)^{n+1}(a) = \lambda_a(n+1) = f(\lambda_a(n)) = f((f)^n(a)) = (f \circ (f)^n)(a)$ so that
$$(f)^{n+1} = f \circ (f)^n$$
□

Theorem 10.39. *Let A be a set, $a \in A$ and $g: \mathbb{N}_{0,\mathbb{R}} \times A \rightarrow A$ then there exist a* ***unique*** *function*

$$\gamma: \mathbb{N}_{0,\mathbb{R}} \rightarrow A$$

such that:

1. $\lambda(0) = a$

2. $\forall n \in \mathbb{N}_{0,\mathbb{R}}\ \lambda(n+1) = g(n, \lambda(n))$

Proof. By [theorems: 10.5,4.24] we have that

$$i_{\mathbb{N}_0 \to \mathbb{C}}: \langle \mathbb{N}_0, + \rangle \to \langle \mathbb{N}_{0,\mathbb{C}}, + \rangle \text{ and } (i_{\mathbb{N}_0 \to \mathbb{C}})^{-1}: \langle \mathbb{N}_{0,\mathbb{C}}, + \rangle \to \langle \mathbb{N}_0, + \rangle \text{ are group isomorphisms} \tag{10.50}$$

and

$$i_{\mathbb{N}_0 \to \mathbb{C}}(1) = 1 \tag{10.51}$$

Define now

$$h: \mathbb{N}_0 \times A \to A \text{ by } h(n, a) = g((i_{\mathbb{N}_0 \to \mathbb{C}})(n), a)$$

Using [theorem: 5.86] there exist a $\beta: \mathbb{N}_0 \to A$ such that

1. $\beta(0) = a$
2. $\forall n \in \mathbb{N}_{0,\mathbb{R}}$ we have $\beta(n+1) = h(n, \beta(n))$

Define now

$$\lambda: \mathbb{N}_{0,\mathbb{C}} \to A \text{ by } \lambda = \beta \circ (i_{\mathbb{N}_0 \to \mathbb{C}})^{-1}$$

then we have:

1. $\lambda(0) = \beta((i_{\mathbb{N}_0 \to \mathbb{C}})^{-1}(0)) \underset{\text{[eq: 10.50]}}{=} \beta(0) = a$
2. If $n \in \mathbb{N}_{0,\mathbb{C}}$ then

$$\begin{array}{rcl} \lambda(n+1) & = & \beta((i_{\mathbb{N}_0 \to \mathbb{C}})^{-1}(n+1)) \\ & \underset{\text{[eq: 10.50]}}{=} & \beta((i_{\mathbb{N}_0 \to \mathbb{C}})^{-1}(n) + (i_{\mathbb{N}_0 \to \mathbb{C}})^{-1}(1)) \\ & \underset{\text{[eq: 10.51]}}{=} & \beta((i_{\mathbb{N}_0 \to \mathbb{C}})^{-1}(n) + 1) \\ & = & h((i_{\mathbb{N}_0 \to \mathbb{C}})^{-1}(n), \beta((i_{\mathbb{N}_0 \to \mathbb{C}})^{-1}(n))) \\ & = & h((i_{\mathbb{N}_0 \to \mathbb{C}})^{-1}(n), \lambda(n)) \\ & = & g(i_{\mathbb{N}_0 \to \mathbb{C}}((i_{\mathbb{N}_0 \to \mathbb{C}})^{-1}(n)), \lambda(n)) \\ & = & g(, \lambda(n)) \end{array}$$

which proves existence. Now for uniqueness, assume that there is a

$$\gamma: \mathbb{N}_{0,\mathbb{C}} \to A \text{ such that } \beta(0) = a \text{ and } \forall n \in \mathbb{N}_{0,\mathbb{C}} \text{ that } \beta(n+1) = g(n, \beta(n))$$

Define now $S = \{n \in \mathbb{N}_{0,\mathbb{C}} | \lambda(n) = \gamma(n)\}$ then we have:

$0 \in S$. As $\lambda(0) = a = \gamma(0)$ it follows that $0 \in S$.

$n \in S \Rightarrow n+1 \in S$. As

$$\lambda(n+1) = g(n, \lambda(n)) \underset{n \in S}{=} g(n, \gamma(n)) = \gamma(n+1)$$

we have that $n+1 \in S$

Using mathematical induction [theorem: 10.35] we have $S = \mathbb{N}_{0,\mathbb{C}}$, so $\forall n \in \mathbb{N}_{0,\mathbb{C}}$ we have $n \in S$ hence $\lambda(n) = \gamma(n)$ proving that

$$\lambda = \gamma$$

□

Corollary 10.40. *Let A be a set, $a \in A$ and $g: \mathbb{N}_{0,\mathbb{C}} \times A \to A$ then there exist a **unique** function*

$$\lambda: \mathbb{N}_{0,\mathbb{C}} \to A$$

such that:

1. $\lambda(0) = a$
2. $\forall n \in \{1, \ldots, \infty\}\ \lambda(n) = g(n-1, \lambda(n-1))$

Proof. Using [theorem: 10.39] there exists a **unique** $\lambda: \mathbb{N}_{0,\mathbb{C}} \to A$ such that

$$\lambda(0) = a \text{ and } \forall n \in \mathbb{N}_0\ \lambda(n+1) = g(n, \lambda(n))$$

Let $n \in \{1, \ldots, \infty\}$ then $1 \leqslant n$ so that $n - 1 \in \mathbb{N}_{0,\mathbb{C}}$ such that $n = (n-1) + 1$, hence $\lambda(n) = \lambda((n-1)+1) = g(n-1, \lambda(n-1))$. □

Theorem 10.41. *Let A be a set, $a \in A$, $n \in \mathbb{N}_0$ and $g: \{0,\ldots,n-1\} \times A \rightarrow A$ a function then there exists a **unique** function $\lambda: \{0,\ldots,n\} \rightarrow A$ satisfying*

$$\begin{aligned} \lambda(0) &= a \\ \forall i \in \{0,\ldots,n-1\} \text{ we have } \lambda(i+1) &= g(i, \lambda(i)) \end{aligned}$$

Proof. Define

$$g': \mathbb{N}_0 \times A \rightarrow A \text{ by } g'(i,x) = \begin{cases} g(i,x) \text{ if } i \in \{0,\ldots,n-1\} \\ x \text{ if } i \in \{n,\ldots,\infty\} \end{cases}$$

then by [corollary: 5.87] there exists a $\beta: \mathbb{N}_0 \rightarrow A$ such that

$$\beta(0) = a \tag{10.52}$$

$$\forall i \in \mathbb{N}_0 \text{ we have } \beta(i+1) = g'(i, \beta(i)) \tag{10.53}$$

Define now $\lambda: \{0,\ldots,n\} \rightarrow A$ by $\lambda = \beta_{|\{0,\ldots,n\}}$ then we have

$$\lambda(0) = \beta_{|\{0,\ldots,n\}}(0) \underset{0 \in \{0,\ldots,n\}}{=} \beta(0) \underset{[\text{theorem: } 10.52]}{=} a$$

and $\forall i \in \{0,\ldots,n-1\}$ we have

$$\begin{aligned} \lambda(i+1) &= \beta_{|\{0,\ldots,n\}}(i+1) \\ &\underset{i+1 \in \{0,\ldots,n\}}{=} \beta(i+1) \\ &\underset{[\text{theorem: } 10.53]}{=} g'(i, \beta(i)) \\ &\underset{i \in \{0,\ldots,n-1\}}{=} g'(i, \lambda(i)) \\ &\underset{i \in \{0,\ldots,n-1\}}{=} g(i, \lambda(i)) \end{aligned}$$

so that we found a function $\lambda: \{0,\ldots,n\} \rightarrow A$ such that

$$\begin{aligned} \lambda(0) &= a \\ \forall i \in \{0,\ldots,n-1\} \text{ we have } \lambda(i+1) &= g(i, \lambda(i)) \end{aligned}$$

Next we must prove uniqueness so let $\gamma: \{0,\ldots,n\} \rightarrow A$ be such that

$$\begin{aligned} \gamma(0) &= a \\ \forall i \in \{0,\ldots,n-1\} \text{ we have } \gamma(i+1) &= g(i, \gamma(i)) \end{aligned}$$

and define $S = \{i \in \mathbb{N}_0 | i \notin \{0,\ldots,n\} \vee \lambda(i) = \gamma(i)\}$ then we have:

- $\boldsymbol{0 \in S}$. As $\lambda(0) = a = \gamma(0)$ we have $0 \in S$
- $\boldsymbol{i \in S \Rightarrow i+1 \in S}$. then for $i+1$ we have either:
  - $\boldsymbol{i+1 \in \{0,\ldots,n\}}$. $\boldsymbol{i+1 \in \{0,\ldots,n\}}$. Then $i+1 \leqslant n$ so that $i < n$ and as $i \in S$ we have $0 \leqslant i$, so it follows that $i \in \{0,\ldots,n\}$. Further

    $$\lambda(i+1) = g(i, \lambda(i)) \underset{i \in \{0,\ldots,n\} \text{ and } i \in S}{=} g(i, \gamma(i)) = \gamma(i+1)$$

    proving that $i+1 \in S$
  - $\boldsymbol{i+1 \notin \{0,\ldots,n\}}$. Then $i+1 \in S$

  so in all cases we have $i+1 \in S$.

By mathematical induction [theorem: 5.83] we have that $S = \mathbb{N}_0$. If $i \in \{0,\ldots,n\} \subseteq \mathbb{N}_0$ we have $i \in S$ which as $i \in \{0,\ldots,n\}$ gives $\lambda(i) = \gamma(i)$ so that $\lambda = \gamma$. □

The three previous theorems gives a way of defining functions by recursions as is expressed in the following two definitions.

Definition 10.42. *Let A be a set, $a \in A$ then we can define a function as follows:*

$$f: \mathbb{N}_{0,\mathbb{C}} \rightarrow A$$

is defined by:

1. $f(0) = a$
2. $f(n+1) = G(n, \lambda(n))$

where $G(n, \lambda(n))$ *is a expression of two parameters. The above is equivalent with the function defined by [theorem: 10.39] where* $a \in A$ *and* $g: \mathbb{N}_{0,\mathbb{C}} \times A \rightarrow A$ *is defined by* $g(n, x) = G(n, x)$.

Another way to define a recursive function is based on [corollary: 10.40]

Definition 10.43. *Let* A *be a set,* $a \in A$ *then we define* $f: \mathbb{N}_{0,\mathbb{C}} \rightarrow A$ *as follows*

$$f(n) = \begin{cases} a \text{ if } n = 0 \\ G(n-1, f(n-1)) \text{ if } n \in \{1, \ldots \infty\} \end{cases}$$

Which is equivalent with the function defined by [theorem: 10.40] where $a \in A$ *and* $g: \mathbb{N}_{0,\mathbb{C}} \times A \rightarrow A$ *is defined by* $g(n, x) = G(n, x)$.

Definition 10.44. *Let* A *be a set,* $a \in A$, $n \in \mathbb{N}_0$ *then we define the function*

$$\lambda: \{0, \ldots, n\} \rightarrow A$$

by

$$\begin{array}{rcl} \lambda(0) & = & a \\ \forall i \in \{0, \ldots, n-1\} \text{ we have } \lambda(i+1) & = & G(i, \lambda(i)) \end{array}$$

where $G(n, \lambda(n))$ *is a expression of two parameters. The above is equivalent with the function defined by [theorem: 10.41] where* $a \in A$ *and* $g: \{0, \ldots, n-1\} \times A \rightarrow A$ *is defined by* $g(n, x) = G(n, x)$.

As a application of recursion we show how we can define a product of a natural number and a element of a field by repeating addition.

Definition 10.45. *Let* $\langle F, \oplus, \odot \rangle$ *be a field with additive neutral element* e *then we define*

$$*: \mathbb{N}_{0,\mathbb{C}} \times F \rightarrow F \text{ where } n * f = f_*(n)$$

where $f_*: F \rightarrow F$ *is defined by*

$$f_*(0) = e$$

$$f_*(n+1) = f \oplus f_*(n)$$

Example 10.46. Let $\langle F, \oplus, \odot \rangle$ be a field with additive neutral element e then

$$\begin{array}{rcl} 0 * f & = & f_*(0) = e \\ 1 * f & = & f_*(1) = f \oplus f_*(0) = f \oplus e = f \\ 2 \cdot f & = & f_*(2) = f \oplus f_*(1) = f \oplus f \\ 3 \cdot f & = & f_*(3) = f \oplus f_*(2) = f \oplus f \oplus f \\ & \ldots & \\ n \cdot f & = & \underbrace{f \oplus \cdots \oplus f}_{n} \end{array}$$

The above allows us to define a field of characteristics zero, which we will use if talk about determinant functions.

Definition 10.47. *Let* $\langle F, \oplus, \odot \rangle$ *be a field with additive neutral element* e *and multiplicative unit* u *then* $\langle F, \oplus, \odot \rangle$ *is of* ***characteristics zero*** *if* $\forall n \in \mathbb{N}_{\mathbb{C}}$ *we have* $n * u \neq e$

10.3 Power in $\mathbb{C}$

We are ready now to define power in $\mathbb{C}$

Definition 10.48. *Let $x \in \mathbb{C}$ then $x^{(\cdot)}: \mathbb{N}_{0,\mathbb{C}} \rightarrow \mathbb{C}$ is defined by $n \rightarrow x^{(\cdot)}(n) = x^n$ where*

$$\begin{aligned} x^0 &= 1 \\ x^{n+1} &= x \cdot x^n \end{aligned}$$

Note 10.49. Let $x \in \mathbb{C}$ then $x^0 = 1$, $x^1 = x$, $x^2 = x \cdot x$

Proof. $x^0 \underset{\text{def}}{=} 1$, $x^1 = x \cdot x^0 = x \cdot 1 = x$ and $x^2 = x \cdot x^1 = x \cdot x$ □

Theorem 10.50. *Let $x, y \in \mathbb{R}_{\mathbb{C}}$ then we have*

1. *If $0 \leqslant x < y$ then $x^2 < y^2$*
2. *If $0 \leqslant x \leqslant y$ then $x^2 \leqslant y^2$*

Proof. 10.14

1. As $0 \leqslant x < y$ we have $0 < y$. For x we have as $0 \leqslant x$ either

   $\boldsymbol{x = 0}$**.** Then by [theorem: 10.14] $0 = 0 \cdot y < y \cdot y = y^2$. Hence $x^2 = 0 \cdot 0 = 0 < y^2$.

   $\boldsymbol{0 < x}$**.** Then multiplying both sides of $x < y$ by x we have by [theorem: 10.14] that $x^2 = x \cdot x < y \cdot x$. Further multiplying both sides of $x < y$ by y we have by [theorem: 10.14] that $x \cdot y < y \cdot y = y^2$, Hence $x^2 < y \cdot x < y^2$.

2. For x, y we have as $0 \leqslant x \leqslant y$ either

   $\boldsymbol{x = y}$**.** Then $x^2 = x \cdot x = y \cdot y = y^2$ hence $x^2 \leqslant y^2$.

   $\boldsymbol{x < y}$**.** Then by (1) we have $x^2 < y^2$ hence $x^2 \leqslant y^2$. □

Theorem 10.51. *Let $n \in \mathbb{N}_{0,\mathbb{C}}$ then we have*

1. *If $x \in \mathbb{N}_{0,\mathbb{C}}$ then $x^n \in \mathbb{N}_{0,\mathbb{C}}$*
2. *If $x \in \mathbb{Z}_{\mathbb{C}}$ then $x^n \in \mathbb{Z}_{\mathbb{C}}$*
3. *If $x \in \mathbb{Q}_{\mathbb{C}}$ then $x^n \in \mathbb{Q}_{\mathbb{C}}$*
4. *if $x \in \mathbb{R}_{\mathbb{C}}$ then $x^n \in \mathbb{R}_{\mathbb{C}}$*
5. *If $x \in \mathbb{R}^+$ then $x^n \in \mathbb{R}^+$ [in other words if $0 < x$ then $0 < x^n$]*

Proof. This is easily proved by induction [see: 10.35]

1. Take $S_x = \{n \in \mathbb{N}_{0,\mathbb{C}} | x^n \in \mathbb{N}_{0,\mathbb{C}}\}$ then we have

   $\boldsymbol{0 \in S_x}$**.** As $x^0 = 1 \in \mathbb{N}_{0,\mathbb{C}}$ we have that $0 \in S_x$

   $\boldsymbol{n \in S_x \Rightarrow n + 1 \in S_x}$**.** As $n \in S$ we have $x^n \in \mathbb{N}_{0.\mathbb{C}}$ and by the hypothesis $x \in \mathbb{N}_{0,\mathbb{C}}$ so using [theorem: 10.5] $x^{n+1} = x \cdot x^n \in \mathbb{N}_{0,\mathbb{C}}$ proving that $n + 1 \in S_x$.

2. Take $S_x = \{n \in \mathbb{N}_{0,\mathbb{C}} | x^n \in \mathbb{Z}_{\mathbb{C}}\}$ then we have

   $\boldsymbol{0 \in S_x}$**.** As $x^0 = 1 \in \mathbb{N}_{0,\mathbb{C}} \subseteq \mathbb{Z}_{\mathbb{C}}$ we have that $0 \in S_x$

   $\boldsymbol{n \in S_x \Rightarrow n + 1 \in S_x}$**.** As $n \in S$ we have $x^n \in \mathbb{Z}_{\mathbb{C}}$ and by the hypothesis $x \in \mathbb{Z}_{\mathbb{C}}$ so using [theorem: 10.5] $x^{n+1} = x \cdot x^n \in \mathbb{Z}_{\mathbb{C}}$ proving that $n + 1 \in S_x$.

3. Take $S_x = \{n \in \mathbb{N}_{0,\mathbb{C}} | x^n \in \mathbb{Q}_{\mathbb{C}}\}$ then we have

   $\boldsymbol{0 \in S_x}$**.** As $x^0 = 1 \in \mathbb{N}_{0,\mathbb{C}} \subseteq \mathbb{Q}_{\mathbb{C}}$ we have that $0 \in S_x$

   $\boldsymbol{n \in S_x \Rightarrow n + 1 \in S_x}$**.** As $n \in S$ we have $x^n \in \mathbb{Q}_{\mathbb{C}}$ and by the hypothesis $x \in \mathbb{Q}_{\mathbb{C}}$ so using [theorem: 10.5] $x^{n+1} = x \cdot x^n \in \mathbb{Q}_{\mathbb{C}}$ proving that $n + 1 \in S_x$.

4. Take $S_x = \{n \in \mathbb{N}_{0,\mathbb{C}} | x^n \in \mathbb{R}_{\mathbb{C}}\}$ then we have

   $\boldsymbol{0 \in S_x}$**.** As $x^0 = 1 \in \mathbb{N}_{0,\mathbb{C}} \subseteq \mathbb{R}_{\mathbb{C}}$ we have that $0 \in S_x$

   $\boldsymbol{n \in S_x \Rightarrow n + 1 \in S_x}$**.** As $n \in S$ we have $x^n \in \mathbb{R}_{\mathbb{C}}$ and by the hypothesis $x \in \mathbb{R}_{\mathbb{C}}$ so using [theorem:10.5] $x^{n+1} = x \cdot x^n \in \mathbb{R}_{\mathbb{C}}$ proving that $n + 1 \in S_x$.

5. Take $S_x = \{n \in \mathbb{N}_{0,\mathbb{C}} | 0 < x^n\}$ then we have

   $0 \in S_x$. Using [corollary: 10.10] $0 < 1 = x^0$ so that $0 \in S_x$.

   $n \in S_x \Rightarrow n+1 \in S_x$. As $0 < x$ and $n \in S \Rightarrow 0 < x^n$ we have by [theorem: 10.14] that $0 < x \cdot x^n = x^{n+1}$ proving that $n+1 \in S_x$ □

Theorem 10.52. *If $n, m \in \mathbb{N}_{0,\mathbb{C}}$ then $\forall x \in \mathbb{C}$ we have $x^{n+m} = x^n \cdot x^m$*

Proof. This is proved by induction, so let $x \in \mathbb{C}, n \in \mathbb{N}_{0,\mathbb{C}}$ and define

$$S_{n,x} = \{m \in \mathbb{N}_{0,\mathbb{C}} | x^{n+m} = x^n \cdot x^m\}$$

then we have:

$0 \in S_{n,x}$. We have $x^{n+0} = x^n = x^n \cdot 1 = x^n \cdot x^0$ proving that $0 \in S_{n,x}$.

$m \in S_{n,x} \Rightarrow m+1 \in S_{n,x}$. Then

$$\begin{array}{rcl} x^{n+(m+1)} & = & x^{(n+m)+1} \\ & = & x \cdot x^{(n+m)} \\ & = & x^{n+m} \cdot x \\ & \underset{m \in S_{n,x}}{=} & (x^n \cdot x^m) \cdot x \\ & = & x^n \cdot (x^m \cdot x) \\ & = & x^n \cdot (x \cdot x^m) \\ & = & x^n \cdot x^{m+1} \end{array}$$

proving that $m+1 \in S_{n,x}$

So $\forall n \in \mathbb{N}_{0,\mathbb{C}}$ we have by mathematical induction [theorem: 10.35] that $S_{n,x} = \mathbb{N}_{0,\mathbb{C}}$. So if $n, m \in \mathbb{N}_{0,\mathbb{C}}$ then $m \in S_{n,x}$ so that $x^{n+m} = x^n \cdot x^m$. □

Theorem 10.53. *If $n, m \in \mathbb{N}_{\mathbb{C}}$ then $\forall x \in \mathbb{C}$ we have $(x^n)^m = x^{n \cdot m}$*

Proof. This is proved by induction, let $S_{n,x} = \{m \in \mathbb{N} | (x^n)^m = x^{n \cdot m}\}$ then we have:

$1 \in S_{n,x}$. As $(x^n)^1 = x^n = x^{n \cdot 1}$ we have $1 \in S$.

$m \in S_{n,x} \Rightarrow m+1 \in S_{n,x}$. We have

$$(x^n)^{m+1} = (x^n)^m \cdot x^n \underset{m \in S}{=} x^{n \cdot m} \cdot x^n \underset{\text{[theorem: 10.52]}}{=} x^{n \cdot m + n} = x^{n \cdot (m+1)}$$

proving that $m+1 \in S_{n,x}$. □

Theorem 10.54. *Let $x \in \mathbb{C} \setminus \{0\}$ then $(x^{-1})^n = (x^n)^{-1}$ or in other words $(1/x)^n = 1/x^n$*

Proof. Theorem we prove this by induction, take $S = \{n \in \mathbb{N}_{\mathbb{C}} | (x^{-1})^n = (x^n)^{-1}\}$ then we have:

$0 \in S$. As $(1/x)^0 = 1 = (1)^{-1} = (x^0)$ proving that $0 \in S$.

$n \in S \rightarrow n+1 \in S$. We have

$$\begin{array}{rcl} (x^{-1})^{n+1} & = & (x^{-1}) \cdot (x^{-1})^n \\ & \underset{n \in S}{=} & (x^{-1}) \cdot (x^n)^{-1} \\ & \underset{\text{[theorem: 4.55]}}{=} & (x \cdot x^n)^{-1} \\ & = & (x^{n+1})^{-1} \end{array}$$

proving that $n+1 \in S$. □

Theorem 10.55. *Let $n \in \mathbb{N}_{0,\mathbb{C}}$ then we have*

1. *If $n \neq 0$ then $x^n = 0 \Leftrightarrow 0$ [note that by definition $0^0 = 1$ so we must have $n \neq 0$]*
2. *$1^n = 1$*
3. *$(-1)^n = 1 \vee (-1)^n = -1$*

4. $(-1)^{2 \cdot n} = 1$

5. $(-1)^{2 \cdot n+1} = -1$

Proof.

1. We have to prove that $\forall n \in \mathbb{N}_{\mathbb{C}}$ we have $x^n = 0 \Leftrightarrow x = 0$

   ⇒. We prove this by induction so let $S = \{n \in \mathbb{N}_{\mathbb{C}} | \text{If } x^n = 0 \text{ then } x = 0\}$ then we have:

   $\boldsymbol{1 \in S.}$ This is trivial because $x = x^1 = 0 \Rightarrow x = 0$

   $\boldsymbol{n \in S \rightarrow n+1 \in S.}$ If $x^{n+1} = 0$ then as $x^{n+1} = x \cdot x^n$ it follows that $0 = x \cdot x^n$. Assume that $x \neq 0$ then as $\mathbb{C}$ is a field x^{-1} exist so that

$$0 = x^{-1} \cdot 0 = x^{-1} \cdot (x \cdot x^n) = (x^{-1} \cdot x) \cdot x^n = 1 \cdot x_n = x_n \Rightarrow x^m = 0$$

   As $m \in S$ it follows that $x = 0$ contradicting the assumption $x \neq 0$, hence the assumption is wrong and we must have $x = 0$. This proves that $n+1 \in S$.

   ⇐. As $n \in \mathbb{N}_{\mathbb{C}}$ we have that $0^n = 0 \cdot 0^{n-1} = 0$.

   so that $0^n = 0^{m+1} = 0 \cdot 0^m = 0$

2. We proceed by induction, so let

$$S = \{n \in \mathbb{N}_{0,\mathbb{C}} | 1^n = 1\}$$

   then we have:

   $\boldsymbol{0 \in S.}$ $1^0 = 1$ by definition, proving that $0 \in S$

   $\boldsymbol{n \in S \Rightarrow n+1 \in S.}$ $1^{n+1} = 1 \cdot 1^n \underset{n \in S}{=} 1 \cdot 1 = 1$ proving that $n+1 \in S$

3. Again we use induction, so let

$$S = \{n \in \mathbb{N}_{0,\mathbb{C}} | (-1)^n = 1 \vee (-1)^n = -1\}$$

   then we have:

   $\boldsymbol{0 \in S.}$ $(-1)^0 = 1$ proving that $0 \in S$.

   $\boldsymbol{n \in S \Rightarrow n+1 \in S.}$ As $n \in S$ we have either:

   $\boldsymbol{(-1)^n = 1.}$ Then $(-1)^{n+1} = (-1) \cdot (-1)^n = (-1) \cdot 1 = -1$ so the $n+1 \in S$

   $\boldsymbol{(-1)^n = -1.}$ Then $(-1)^{n+1} = (-1) \cdot (-1)^n = (-1) \cdot (-1) \underset{\text{[theorem: 4.40]}}{=} 1 \cdot 1 = 1$ so that $n+1 \in S$

4. $(-1)^{2 \cdot n} = (-1)^{(1+1) \cdot n} = (-1)^{n+n} \underset{\text{[theorem: 10.52]}}{=} (-1)^n \cdot (-1)^n \underset{\text{[theorem: 4.40] and (3)}}{=} 1$

5. $(-1)^{2 \cdot n+1} = (-1) \cdot (-1)^{2 \cdot n} \underset{(4)}{=} (-1) \cdot 1 = -1$ □

Theorem 10.56. *For $\mathbb{R}_{\mathbb{C}}$ we have*

1. $\forall \alpha \in \mathbb{R}_{\mathbb{C}}$ *with* $0 < \alpha < 1$ *and* $n \in \{1, \ldots, \infty\}$ *that* $0 < \alpha^n < 1$

2. $\forall \alpha \in \mathbb{R}_{\mathbb{C}}$ *with* $0 < \alpha < 1$ *and* $n \in \{2, \ldots, \infty\}$ *that* $0 < \alpha^n < \alpha$

3. $\forall \alpha, \beta \in \mathbb{R}_{\mathbb{C}}$ *such that* $1 \leqslant \alpha$ *and* $n \in \mathbb{N}_{\mathbb{C}}$ *we have:*

   a. If $\alpha < \beta$ *then* $\alpha < \beta^n$

   b. If $\alpha \leqslant \beta$ *then* $\alpha \leqslant \beta^n$

Proof.

1. We proof this by induction on n. So let $S = \{n \in \{1, \ldots, \infty\} | 0 < \alpha^n < 1\}$ then we have:

   $\boldsymbol{1 \in S.}$ As $0 < \alpha < 1$ we have $0 < \alpha^1 < 1$ so that $1 \in S$

   $\boldsymbol{n \in S \Rightarrow n+1 \in S.}$ As $n \in S$ we have $0 < \alpha^n < 1$, so using [theorem: 10.14] we have that $0 = 0 \cdot \alpha < \alpha^n \cdot \alpha < 1 \cdot \alpha = \alpha < 1$ or $0 < \alpha^{n+1} < 1$. Hence $n+1 \in \boldsymbol{S}$

   proving that $S = \{1, \ldots, \infty\}$ or $\forall n \in \{1, \ldots, \infty\}$ we have $0 < \alpha^n < 1$

2. As $n\in\{2,\ldots,\infty\}$ we have $2\leqslant n\Rightarrow 1=2+(-1)\leqslant n+(-1)=n-1$ so that $(n-1)\in\{1,\ldots,\infty\}$. Using (1) we have $0<\alpha^{n-1}<1$ which as $0<\alpha$ gives by [theorem: 10.14]

$$0<0\cdot\alpha<\alpha^{n-1}\cdot\alpha<1\cdot\alpha=\alpha.$$

or as $\alpha^{n-1}\cdot\alpha=\alpha^n$ that $0<\alpha^n<\alpha$.

3. Let $\alpha,\beta\in\mathbb{R}_{\mathbb{C}}$ such that $1\leqslant\alpha$ and $n\in\mathbb{N}_{\mathbb{C}}$

a. If $\alpha<\beta$ we have to prove that $\alpha<\beta^n$. Let $S=\{n\in\{1,\ldots,n\}|\alpha<\beta^n\}$ then we have:

$\boldsymbol{1\in S}$. As $\alpha<\beta=\beta^1$ we have that $1\in S$

$\boldsymbol{n\in S\Rightarrow n+1\in S}$. As $1\leqslant\alpha<\beta\Rightarrow 1<\beta$ we have by [theorem: 10.14] that

$$\alpha=1\cdot\alpha<\beta\cdot\alpha=\alpha\cdot\beta \tag{10.54}$$

As $n\in S$ we have that $\alpha<\beta^n$ which by [theorem: 10.14] gives $\alpha\cdot\beta<\beta^n\cdot\beta=\beta^{n+1}$, combining this with [eq: 10.54] proves $\alpha<\beta^{n+1}$. So $n+1\in S$.

Hence $S=\{1,\ldots,n\}=\mathbb{N}_{\mathbb{C}}$ hence $\forall n\in\mathbb{N}_{\mathbb{C}}$ we have $\alpha<\beta^n$.

b. If $\alpha\leqslant\beta$ we have to prove that $\alpha\leqslant\beta^n$. Let $S=\{n\in\{1,\ldots,n\}|\alpha\leqslant\beta^n\}$ then we have:

$\boldsymbol{1\in S}$. As $\alpha\leqslant\beta=\beta^1$ we have that $1\in S$

$\boldsymbol{n\in S\Rightarrow n+1\in S}$. As $1\leqslant\alpha\leqslant\beta\Rightarrow 1\leqslant\beta$ we have by [theorem: 10.14] that

$$\alpha=1\cdot\alpha\leqslant\beta\cdot\alpha=\alpha\cdot\beta \tag{10.55}$$

As $n\in S$ we have that $\alpha\leqslant\beta^n$ which by [theorem: 10.14] gives $\alpha\cdot\beta\leqslant\beta^n\cdot\beta=\beta^{n+1}$, combining this with [eq: 10.55] proves $\alpha\leqslant\beta^{n+1}$. So $n+1\in S$.

So $S=\{1,\ldots,n\}=\mathbb{N}$ hence $\forall n\in\mathbb{N}$ we have $\alpha\leqslant\beta^n$. □

Theorem 10.57. *$\forall n\in\mathbb{N}_{0,\mathbb{C}}$ we have that $n<2^n$*

Proof. This is proved by induction so let $S=\{n\in\mathbb{N}_{0,\mathbb{C}}|n<2^n\}$ then we have

$\boldsymbol{0\in S}$. As $n=0<1=2^0$ we have that $0\in S$

$\boldsymbol{n\in S\Rightarrow n+1\in S}$. For $n+1$ we have the following cases to consider:

$\boldsymbol{n+1=1}$. Then $n+1=1<2=2^1=2^{n+1}$ proving that in this case $n+1\in S$

$\boldsymbol{1<n+1}$. Then by [theorem: 10.28] we have $1+1\leqslant n+1$ so that $1\leqslant n$, adding n to both sides gives then $n+1\leqslant n+n=(1+1)\cdot n=2\cdot n$. Further as $n\in S$ we have $n<2^n$ so that $2\cdot n<2\cdot 2^n=2^{n+1}$, hence we have $n+1<2^{n+1}$, proving that $n+1\in S$. □

Corollary 10.58. *$\forall x\in\mathbb{R}_{\mathbb{C}}$ there exists a $n\in\mathbb{N}_{0,\mathbb{C}}$ such that $x<2^n$*

Proof. Let $x\in\mathbb{R}_{\mathbb{C}}$ then by [corollary: 10.30] there exist a $n\in\mathbb{N}_{0,\mathbb{C}}$ such that $x<n$, using [theorem: 10.57] $n<2^n$ so we have $x<2^n$. □

Lemma 10.59. *Let $x\in\mathbb{R}_{\mathbb{C}}$ with $1<x$ then $\forall n\in\mathbb{N}_{0,\mathbb{C}}$ we have*

$$n\cdot(x-1)\leqslant x^n-1$$

Proof. We prove this by induction, so let $S=\{n\in\mathbb{N}_{0,\mathbb{C}}|n\cdot(x-1)\leqslant x^n-1\}$ then we have:

$\boldsymbol{0\in S}$. If $n=0$ then $n\cdot(x-1)=0\cdot(x-1)=0\leqslant 0=1-1=x^0-1=x^n-1$ proving that $0\in S$.

$\boldsymbol{n\in S\Rightarrow n+1\in S}$. As $n\in S$ we have $n\cdot(x-1)\leqslant x^n-1$ which as $0<x$ prove that

$$x\cdot n\cdot(x-1)\leqslant x\cdot(x^n-1) \tag{10.56}$$

As $1<x$ we have that $0<x-1\underset{0\leqslant n}{\Rightarrow}0<n\cdot(x-1)$ so that

$$n\cdot(x-1)=1\cdot n\cdot(x-1)<x\cdot n\cdot(x-1)$$

which combined with [eq: 10.56] proves that

$$n\cdot(x-1)\leqslant x\cdot(x^n-1) \tag{10.57}$$

Further

$$\begin{array}{rcl} x^{n+1}-1 & = & x\cdot x^n-1\\ & = & x\cdot x^n-x+x-1\\ & = & x\cdot(x^n-1)+(x-1)\\ & \geqslant_{[\text{eq: } 10.57]} & n\cdot(x-1)+(x-1)\\ & = & (n+1)\cdot(x-1) \end{array}$$

proving that $n+1\in S$. □

Theorem 10.60. *If $m\in\mathbb{N}_{0,\mathbb{C}}$ and $x\in\mathbb{R}_{\mathbb{C}}$ with $1<x$ then $\exists n\in\mathbb{N}_{0,\mathbb{C}}$ such that $m<x^n$.*

Proof. For $m\in\mathbb{N}_{0,\mathbb{C}}$ we have either:

- $\boldsymbol{m=0.}$ Then $m=0<1=x^0$ proving the theorem in this case.
- $\boldsymbol{m=1.}$ Then $m=1<x=x^1$ proving the theorem in this case.
- $\boldsymbol{1<m.}$ Then $0<m-1$ and $0<x-1\Rightarrow 0<(x-1)^{-1}$ so that $(m-1)/(x-1)$ is defined. By [corollary: 10.30] there exist a $n\in\mathbb{N}_{0,\mathbb{C}}$ such that $(m-1)/(x-1)<n$. As $0<x-1$ we have $m-1<n\cdot(x-1)$, by [lemma: 10.59] $n\cdot(x-1)\leqslant(x^n-1)$, hence $m-1\leqslant x^n-1$ or $m<x^n$, proving the theorem in this case. □

Theorem 10.61. *If $x\in\mathbb{R}_{\mathbb{C}}$ with $0<x<1$ and $n,m\in\mathbb{N}_{\mathbb{C}}$ such that $n<m$ then $x^m<x^n$.*

Proof. We prove this by induction, let $m\in\mathbb{N}_{0,\mathbb{C}}$ and take $S_m=\{n\in\{1,\ldots,\infty\}|x^{m+n}<x^m\}$ then we have:

- $\boldsymbol{1\in S_m.}$ As $0<x$ we have by [theorem: 10.51] that $0<x^m$, hence from $x<1$ we have that $x^{m+1}=x\cdot x^m<1\cdot x^m=x^m$, proving that $1\in S_m$.
- $\boldsymbol{n\in S_m\Rightarrow n+1\in S_m.}$ As $x<1$ and by [theorem: 10.51] $0<x^{m+n}$ we have that

$$x^{m+(n+1)}=x\cdot x^{m+n}<x^{m+n}\underset{n\in S_m\Rightarrow x^{m+n}<x^m}{\Rightarrow}x^{m+(n+1)}<x^m$$

proving that $n+1\in S_m$.

Using mathematical induction [see theorem: 10.35] we have that $S_m=\{1,\ldots,\infty\}$. So take $n,m\in\mathbb{N}_{0,\mathbb{R}}$ with $n<m$ then $k=m-n>0\Rightarrow k\geqslant 1$ so that $k\in\{1,\ldots,\infty\}=S_n$ hence $x^m=x^{n+k}<x^n$ completing the proof. □

Theorem 10.62. *If $x\in\mathbb{R}_{\mathbb{C}}$ with $1\leqslant x$ then $\forall n,m\in\mathbb{N}_{\mathbb{C},0}$ with $n\leqslant m$ we have $x^n\leqslant x^m$*

Proof. We prove this by induction so take $S_n=\{m\in\{0,\ldots,\infty\}|x^n\leqslant x^{n+m}\}$ then we have:

- $\boldsymbol{0\in S_n.}$ As $x^n=x^{n+0}\Rightarrow x^n\leqslant x^{n+0}$ proving that $0\in S$.
- $\boldsymbol{m\in S\Rightarrow m+1\in S.}$ As $1\leqslant x$ $0<x$ it follows from [theorem: 10.56] that $0<x^{n+m}$, hence as $1\leqslant x$ it follows from [theorem: 10.14] that $x^{n+m}=1\cdot x^{n+m}\leqslant x\cdot x^{n+m}=x^{n+(m+1)}$. Given that $x^n\leqslant x^{n+m}$ [because $m\in S$] it follows that $x^n\leqslant x^{n+(m+1)}$, hence $m+1\in S$.

Using mathematical induction [see theorem: 10.35] we have that $S_n=\{0,\ldots,\infty\}$. So if $n,m\in\mathbb{N}_{\mathbb{C},0}$ with $n\leqslant m$ then $k=m-n\in\{0,\ldots,\infty\}=S_n$ and $x^n\leqslant x^{n+k}=x^{n+(m-n)}=x^m$ completing the proof. □

Theorem 10.63. *If $\varepsilon\in\mathbb{R}_{\mathbb{C}}^+$, $x\in\mathbb{R}_{\mathbb{C}}$ with $0<x<1$ then $\exists N\in\mathbb{N}_{\mathbb{C}}$ such that $\forall n\in\mathbb{N}_{\mathbb{C}}$ $N\leqslant n$*

$$0<x^n<\varepsilon$$

Proof. As $0<\varepsilon\Rightarrow\varepsilon\neq 0$ so that $1/\varepsilon$ is defined. By a consequence of the Archimedean property [theorem: 10.30] there exist a $M\in\mathbb{N}_{\mathbb{C}}$ such that

$$0<1/\varepsilon<M$$

As $0<x<1$ we have $1<\frac{1}{x}$ so that by [theorem: 10.60] there exist a $N'\in\mathbb{N}_{\mathbb{C},0}$ such that $M<\left(\frac{1}{x}\right)^N$, hence $0<1/\varepsilon<\left(\frac{1}{x}\right)^N\underset{[\text{theorem: } 10.54]}{=}1/x^N$ so that

$$0<x^{N'}<\varepsilon$$

Take $N=N'+1$ then if $n\in\mathbb{N}_{0,\mathbb{C}}$ with $N\leqslant n$ then $N'<n$. So we have by [theorem: 10.61] and $0<x<1$ that $x^n<x^{N'}<\varepsilon$. As $0<x$ we have by [theorem: 10.51] that

$$0<x^n<\varepsilon$$

□

10.4 The square root in $\mathbb{R}_{\mathbb{C}}$

Theorem 10.64. *The function*

$$(.)^2 \colon \mathbb{R}_{0,\mathbb{C}}^{+} \rightarrow \mathbb{R}_{0,\mathbb{C}}^{+} \text{ defined by } (.)^2(x) = x^2 = x \cdot x$$

is a bijection.

Proof. We have

injectivity. Let $x, y \in \mathbb{R}_{0,\mathbb{C}}^{+} = \{x \in \mathbb{R}_{\mathbb{C}} \wedge 0 \leqslant x\}$ be such that $x^2 = y^2$ then we have the following possibilities:

$x = 0$. Assume that $y \neq 0$ then $0 < y$ so that by [theorems: 10.14] $0 < y \cdot y = y^2 = x^2 = x \cdot x = 0$ leading to the contradiction $0 < 0$. Hence we have $y = 0$ so that $x = y$.

$y = 0$. Assume that $x \neq 0$ then $0 < x$ so that by [theorems: 10.14] $0 < x \cdot x = x^2 = y^2 = y \cdot y = 0$ leading to the contradiction $0 < 0$. Hence we have $x = 0$ so that $x = y$.

$0 < x \wedge 0 < y$. Assume that $x \neq y$ then we have either:

$x < y$. Then by [theorems: 10.14] we have $x \cdot y < y \cdot y = y^2$ and $x^2 = x \cdot x < y \cdot x = x \cdot y$ so that $x^2 < y^2$ contradicting $x^2 = y^2$.

$y < x$. Then by [theorems: 10.14] we have $y \cdot x < x \cdot x = x^2$ and $y^2 = y \cdot y < x \cdot y = y \cdot x$ so that $y^2 < x^2$ contradicting $x^2 = y^2$.

As the assumption $x \neq y$ leads to a contradiction in all cases we must have that $x = y$.

So in all cases we have $x = y$ proving injectivity.

surjectivity. If $y \in \mathbb{R}_{0,\mathbb{C}}^{+} = \{x \in \mathbb{R}_{\mathbb{C}} | 0 \leqslant x\}$ then $0 \leqslant y$ and we have the fallowing possibilities to consider:

$y = 0$. Then $0^0 = 0 \cdot 0 = 0 = y$, hence if $x = 0$ we have $y = x^2$

$y = 1$. Then $1^2 = 1 \cdot 1 = 1 = y$, hence if $x = 1$ we have $y = x^2$

$0 < y \wedge y \neq 1$. Take then $S_y = \{t \in \mathbb{R}_{\mathbb{C}} | 0 \leqslant t \wedge t^2 \leqslant y\}$. As $0^2 = 0 < y$ we have $0 \in S_y$ hence

$$S_y \neq \varnothing \tag{10.58}$$

As $y \neq 1$ we have either:

$y < 1$. Assume that $\exists t \in S_y$ such that $1 < t$ then by [theorems: 10.14] we have $t < t \cdot t = t^2$, as $t \in S_y$ we have $t^2 \leqslant y < 1$ so that $t < 1$ contradicting $1 < t$. Hence we have $\forall t \in S_y$ that $t \leqslant 1$ proving that S_y is bounded above.

$1 < y$. Assume that $\exists t \in S_y$ such that $y < t$ then as $1 < y$ we have $1 < t$ so that $t < t \cdot t = t^2 \leqslant y$ contradicting the assumption $y < t$. So the assumption is wrong and we have $\forall t \in S_Y$ that $t \leqslant y$ proving that S_y is bounded above.

So in all cases we have that

$$S_y \text{ is bounded above} \tag{10.59}$$

As $\langle \mathbb{R}_{\mathbb{C}}, \leqslant \rangle$ is conditionally complete [see theorem: 10.18] we have thanks to [eqs: 10.58, 10.59] that

$$s_y = \sup(S_y) \text{ exist} \tag{10.60}$$

For $y \neq 1$ we consider again the following possibilities:

$y < 1$. As $0 < y$ we have by [theorems: 10.14] that $y^2 = y \cdot y < 1 \cdot y = y$ so that $y \in S_y$ so $y \leqslant s_y \underset{0<y}{\Rightarrow} 0 < s_y$.

$1 < y$. Then $1^2 = 1 < y$ so as $0 < 1$ we have $1 \in S_y$, hence $1 \leqslant s_y$ proving that $0 < s_y$.

So in all cases we have

$$0 < s_y \tag{10.61}$$

Let $\varepsilon \in \mathbb{R}_{\mathbb{C}}$ such that $0 < \varepsilon < s_y$. Then $0 < s_y - \varepsilon < s_y < s_y + \varepsilon$, so that by [theorems: 10.14]

$$\begin{gathered}
(s_y - \varepsilon)^2 = (s_y - \varepsilon) \cdot (s_y - \varepsilon) < s_y \cdot (s_y - \varepsilon) \\
(s_y - \varepsilon) \cdot s_y < s_y \cdot s_y = s_y^2 \\
s_y \cdot (s_y + \varepsilon) < (s_y + \varepsilon) \cdot (s_y + \varepsilon) = (s_y + \varepsilon)^2 \\
s_y^2 = s_y \cdot s_y < (s_y + \varepsilon) \cdot s_y
\end{gathered}$$

So that

$$(s_y - \varepsilon)^2 < s_y^2 < (s_y + \varepsilon)^2 \tag{10.62}$$

As s_y is a upper bound of S_y and $s_y < s_y + \varepsilon$ we must have that $s_y + \varepsilon \notin S_y$, which, as $0 < s_y < s_y + \varepsilon$ proves that

$$y < (s_y + \varepsilon)^2 \tag{10.63}$$

As $\langle \mathbb{R}_{\mathbb{C}}, \leqslant \rangle$ is totally ordered and $s_y - \varepsilon < s_y$ we have by [theorem: 3.71] that $\exists f \in S_y \Rightarrow 0 \leqslant f$ such that $s_y - \varepsilon < f$. As $0 < s_y - \varepsilon$ we have by [theorems: 10.14] that $(s_y - \varepsilon) \cdot f < f \cdot f = f^2$ and $(s_y - \varepsilon)^2 = (s_y - \varepsilon) \cdot (s_y - \varepsilon) < f \cdot (s_y - \varepsilon)$ so that $(s_y - \varepsilon)^2 < f^2$. As $f \in S_y$ it follows that $f^2 \leqslant y$ so that

$$(s_y - \varepsilon)^2 < y \tag{10.64}$$

Using [theorems: 10.14] on [eqs: 10.63, 10.64] we have that

$$-(s_y + \varepsilon)^2 < -y < -(s_y - \varepsilon)^2 \tag{10.65}$$

Adding [eq: 10.62] to [eq: 10.65] gives

$$(s_y - \varepsilon)^2 - (s_y + \varepsilon)^2 < s_y^2 - y < (s_y + \varepsilon)^2 - (s_y - \varepsilon)^2 \tag{10.66}$$

Now

$$\begin{aligned} (s_y - \varepsilon)^2 - (s_y + \varepsilon)^2 &= s_y^2 - 2 \cdot \varepsilon \cdot s_y + \varepsilon^2 - (s_y^2 + 2 \cdot \varepsilon * s_y + \varepsilon^2) \\ &= -4 \cdot \varepsilon \cdot s_y \end{aligned}$$

which combined with [eq: 10.66] gives

$$\forall \varepsilon \in \mathbb{R}_{\mathbb{C}} \text{ with } 0 < \varepsilon < s_y \text{ we have } -4 \cdot \varepsilon \cdot s_y^2 < s_y^2 - y < 4 \cdot \varepsilon \cdot s_y \tag{10.67}$$

or using [theorems: 10.14]

$$\forall \varepsilon \in \mathbb{R}_{\mathbb{C}} \text{ with } 0 < \varepsilon < s_y \text{ we have } -4 \cdot \varepsilon \cdot s_y^2 < y - s_y^2 < 4 \cdot \varepsilon \cdot s_y \tag{10.68}$$

Now as $0 < s_y$ we have by [theorem: 10.32] a $\varepsilon_0 \in \mathbb{R}_{\mathbb{C}}$ such that

$$0 < \varepsilon_0 < s_y \tag{10.69}$$

For $s_y^2 - y$ we can have now the following possibilities:

$s_y^2 - y < 0$. Take then $\delta = y - s_y^2$ then $0 < \delta$. Take $\varepsilon = \min(\delta / (4 \cdot s_y), \varepsilon_0)$ then we have as $0 < 4, s_y, \varepsilon_0$ by [theorems: 10.14] that $0 < \varepsilon$ and $\varepsilon \leqslant \varepsilon_0 < s_y$ so we have by [eq: 10.68] that $\delta = y - s_y^2 < 4 \cdot \varepsilon \cdot s_y$. As $4 \cdot \varepsilon \cdot s_y \leqslant (\delta / (4 \cdot s_y)) \cdot 4 \cdot s_y = \delta$ we have the contradiction $\delta < \delta$. So this case does not occur.

$0 < s_y^2 - y$. Take then $\delta = s_y^2 - y$ then $0 < \delta$. Take $\varepsilon = \min(\delta / (4 \cdot s_y), \varepsilon_0)$ then we have as $0 < 4, s_y, \varepsilon_0$ by [theorems: 10.14] that $0 < \varepsilon$, further $\varepsilon \leqslant \varepsilon_0 < s_y$, so we have by [eq: 10.67] that $\delta = s_y^2 - y < 4 \cdot \varepsilon \cdot s_y$. As $4 \cdot \varepsilon \cdot s_y \leqslant (\delta / (4 \cdot s_y)) \cdot 4 \cdot s_y = \delta$ we have the contradiction $\delta < \delta$. So this case does not occur.

$s_y^2 - y = 0$. Then $y = s_y^2$

So the only valid case is where $y = s_y^2$, hence if we take $x = s_y$ then $x^2 = y$.

In all cases we have found a $x \in \mathbb{R}_{\mathbb{C}}$ such that $y = x^2 = (.)^2(x)$ which proves surjectivity. □

Definition 10.65. (Square Root) *Using the previous theorem [theorem: 10.64] we have that*

$$(.)^2 \colon \mathbb{R}_{0,\mathbb{C}}^+ \rightarrow \mathbb{R}_{0,\mathbb{C}}^+ \text{ defined by } (.)^2(x) = x \cdot x$$

is a bijection so that we have a inverse bijection

$$((.)^2)^{-1} \colon \mathbb{R}_{0.\mathbb{C}}^+ \rightarrow \mathbb{R}_{0,\mathbb{C}}^+$$

this inverse bijection is called the square root mapping and noted by

$$\sqrt{.} \colon \mathbb{R}_{0,\mathbb{C}}^+ \rightarrow \mathbb{R}_{0,\mathbb{C}}^+ \text{ where } \sqrt{.} = ((.)^2)^{-1}$$

Hence if $x \in \mathbb{R}_0^+$ *then* $\sqrt{(x^2)} = (\sqrt{.} \circ (.)^2)(x) = i_{\mathbb{R}_{0,\mathbb{C}}^+}(x) = x$ *and* $(\sqrt{x})^2 = ((.)^2 \circ \sqrt{.})(x) = i_{\mathbb{R}_{0,\mathbb{C}}^+}(x)$

Note 10.66. The requirement that $x \in \mathbb{R}_{0,\mathbb{C}}^+$ in the above is required because $(.)^2 \colon \mathbb{R}_{\mathbb{C}} \rightarrow \mathbb{R}_{\mathbb{C}}$ is not injective [for example $(1)^2 = 1 = (-1)^2$]

Example 10.67. $\sqrt{0} = 0$ and $\sqrt{1} = 1$

Proof. First, as by [corollary: 10.10] $0<1$ so that $\sqrt{0}$ and $\sqrt{1}$ are well defined. As $0^2=0\cdot 0=0$ we have $\sqrt{0}=((.)^2)^{-1}(0)=0$. Further from $1^2=1\cdot 1=1$ we have that $\sqrt{1}=((.)^2)^{-1}(1)=1$. □

Note 10.68. $2\in\mathbb{R}_0^+$so $\sqrt{2}$ exist but $\sqrt{2}\notin\mathbb{Q}_{\mathbb{C}}$ so that $\sqrt{2}\in\mathbb{R}_{\mathbb{C}}\setminus\mathbb{Q}_{\mathbb{C}}$ or $\sqrt{2}$ is a irrational number.

Proof. Assume that $\sqrt{2}\in\mathbb{Q}_{\mathbb{C}}$ then as by [theorem: 10.5] $i_{\mathbb{Q}\rightarrow\mathbb{C}}$: $\langle\mathbb{Q},+,\cdot\rangle\rightarrow\langle\mathbb{Q}_{\mathbb{C}},+,\cdot\rangle$ is a ring isomorphism there exist a $q\in\mathbb{Q}$ such that $i_{\mathbb{Q}\rightarrow\mathbb{C}}(q)=\sqrt{2}$, hence:

$$\begin{aligned} i_{\mathbb{Q}\rightarrow\mathbb{C}}(q\cdot q) &= i_{\mathbb{Q}\rightarrow\mathbb{C}}(q)\cdot i_{\mathbb{Q}\rightarrow\mathbb{C}}(q) \\ &= \sqrt{2}\cdot\sqrt{2} \\ &= 2 \\ &= 1+1 \\ &= i_{\mathbb{Q}\rightarrow\mathbb{C}}(1)+i_{\mathbb{Q}_{\mathbb{C}}}(1) \\ &= i_{\mathbb{Q}\rightarrow\mathbb{C}}(1+1) \\ &= i_{\mathbb{Q}\rightarrow\mathbb{C}}(2) \end{aligned}$$

so that by injectivity we have that $q\cdot q=2$ which by [theorem: 8.40] is impossible. □

Theorem 10.69. $\sqrt{.}:\mathbb{R}_0^+\rightarrow\mathbb{R}_0^+$ *is strictly increasing.*

Proof. We prove this by contradiction. Let $x,y\in\mathbb{R}_{0,\mathbb{C}}^+$ be such that $x<y$ and assume that $\sqrt{y}\leqslant\sqrt{x}$, then as $\sqrt{x},\sqrt{y}\in\mathbb{R}_{0,\mathbb{C}}^+$ we can use [theorems: 10.14] getting $y=\sqrt{y}\cdot\sqrt{y}\leqslant\sqrt{x}\cdot\sqrt{y}$ and $\sqrt{y}\cdot\sqrt{x}\leqslant\sqrt{x}\cdot\sqrt{x}=x$, so $y\leqslant x$ contradicting $x<y$. Hence we must have that $\sqrt{x}<\sqrt{y}$. □

Theorem 10.70. *If* $x\in\mathbb{R}_{\mathbb{C}}$ *and* $a\in\mathbb{R}_{0,\mathbb{C}}^+$ *then we have*

1. $x^2=a\Leftrightarrow x=\sqrt{a}\vee x=-\sqrt{a}$
2. $x^2\leqslant a\Leftrightarrow -\sqrt{a}\leqslant x\leqslant\sqrt{a}$
3. $x^2<a\Leftrightarrow -\sqrt{a}<x<\sqrt{a}$

Proof.

1.

⇒. For x we have either:

$0\leqslant x$. Then by [definition: 10.65] we have $x=\sqrt{x^2}=\sqrt{a}$

$x<0$. Then $0<-x$ and $(-x)^2=x^2=a$ so that by [definition: 10.65] that $-x=\sqrt{(-x)^2}=\sqrt{a}$ giving $x=-\sqrt{a}$

⇐. If $x=\sqrt{a}$ then $x^2=(\sqrt{a})^2=a$ and if $x=-\sqrt{a}$ we have $x^2=(-\sqrt{a})^2=(\sqrt{a})^2=a$

2.

⇒. For x we have either

$0\leqslant x$. As $0\leqslant\sqrt{a}$ it follows that $-\sqrt{a}\leqslant 0\leqslant x$. Further assume that $\sqrt{a}<x$ then by [theorems: 10.14] and $0\leqslant\sqrt{a},x$ we have $a=(\sqrt{a})^2<x^2$ contradicting $x^2\leqslant a$, so we must have $x\leqslant\sqrt{a}$. So $-\sqrt{a}\leqslant x\leqslant\sqrt{a}$.

$x<0$. Then as $0\leqslant\sqrt{a}\Rightarrow -\sqrt{a}\leqslant 0$ we have $x<\sqrt{a}$. Further assume that $\sqrt{a}<-x$ we have by[theorems: 10.14] and $0\leqslant\sqrt{a},-x$ that $a=(\sqrt{a})^2<(-x)^2=x^2$ contradicting $x^2\leqslant a$, hence $-x\leqslant\sqrt{a}$ or $-\sqrt{a}\leqslant x$. So $-\sqrt{a}\leqslant x<\sqrt{a}\Rightarrow -\sqrt{a}\leqslant x\leqslant\sqrt{a}$.

⇐. For x we have either:

$0\leqslant x$. Then from $x\leqslant\sqrt{a}$ and $0\leqslant\sqrt{a}$ we have by [theorems: 10.14] that $x^2\leqslant(\sqrt{a})^2=a$.

$x<0$. Then from $-\sqrt{a}\leqslant x\Rightarrow -x\leqslant\sqrt{a}$ and $0\leqslant\sqrt{a},-x$ we have by [theorems: 10.14] that $x^2=(-x)^2\leqslant(\sqrt{a})^2=a$.

3.

⇒. If $x=\sqrt{a}$ or $x=-\sqrt{a}$ then by (1) $x^2=a$ contradicting $x^2<a$ hence we must have that $x\neq\sqrt{a}$ and $x\neq-\sqrt{a}$. Using (2) we have that $-\sqrt{a}\leqslant x\leqslant\sqrt{a}$ so that $-\sqrt{a}<x<\sqrt{a}$.

⇐. As $-\sqrt{a}<x<\sqrt{a}$ we have $x\neq\sqrt{a}$ and $x\neq-\sqrt{a}$ so that by (1) $x^2\neq a$, further by (2) we have $x^2\leqslant a$ so that $x^2<a$. □

Corollary 10.71. *If $x\in\mathbb{R}_{\mathbb{C}}$ then $x\leqslant|x|=\sqrt{x^2}$*

Proof. For $x\in\mathbb{R}_{\mathbb{C}}$ we have either:

$0\leqslant x$. Then $|x|=x$ so that $|x|^2=x^2$

$x<0$. Then $|x|=-x$ so that $|x|^2=(-x)\cdot(-x)=(-1)\cdot(-1)\cdot x\cdot x=x^2$

so that

$$|x^2|=x^2$$

As $|x|\in\mathbb{R}_{0,C}^+$ we have that

$$|x|=\sqrt{|x|^2}=\sqrt{x^2}$$

□

Theorem 10.72. *If $x,y\in\mathbb{R}_{0,\mathbb{C}}^+$ 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 $(.)^2\colon\mathbb{R}_{0,\mathbb{C}}^+\rightarrow\mathbb{R}_{0,\mathbb{C}}^+$ is a bijection and thus injective that $\sqrt{x\cdot y}=\sqrt{x}\cdot\sqrt{y}$. □

Theorem 10.73. *Let $x,y\in\mathbb{R}_{0,\mathbb{C}}^+$ then $\sqrt{x+y}\leqslant\sqrt{x}+\sqrt{y}$*

Proof. We prove this by contradiction, so assume that $\sqrt{x}+\sqrt{y}<\sqrt{x+y}$. Then by [theorem: 10.14] we have that $(\sqrt{x}+\sqrt{y})^2<(\sqrt{x+y})^2=x+y$. Now

$$(\sqrt{x}+\sqrt{y})^2=(\sqrt{x})^2+(\sqrt{y})^2+2\cdot\sqrt{x}\cdot\sqrt{y}=x+y+2\cdot\sqrt{x}\cdot\sqrt{y}$$

so that we have that $x+y+2\cdot\sqrt{x}\cdot\sqrt{y}<x+y$ or $2\cdot\sqrt{x}\cdot\sqrt{y}<0$ giving

$$\sqrt{x}\cdot\sqrt{y}<0$$

As $0\leqslant\sqrt{x}$ and $0\leqslant\sqrt{y}$ we have by [theorem: 10.14] that $0\leqslant\sqrt{x}\cdot\sqrt{y}$ contradicting the above. So we must have that $\sqrt{x+y}\leqslant\sqrt{x}+\sqrt{y}$. □

10.5 Operations on Complex numbers

10.5.1 Notation of complex numbers

Note that the additive neutral element $(0,0)$ is noted as 0 and the multiplicative unit element $(1,0)$ is noted as 1. In the following definition we have also a special notation for $(0,1)$.

Definition 10.74. *$i\in\mathbb{C}$ is defined as $i=(0,1)$ so that as $0\neq 1$ we have that $i\in\mathbb{C}\setminus\mathbb{R}_{\mathbb{C}}$.*

Theorem 10.75. *For i we have*

1. *$i\cdot i=-1$*
2. *If $z\in\mathbb{C}$ then there exists **unique** $x,y\in\mathbb{R}_{\mathbb{C}}$ such that $z=x+i\cdot y$*

Proof.

1. $i\cdot i=(0,1)\cdot(0,1)=(0\cdot 0-1\cdot 1,0\cdot 1+1\cdot 0)=(-1,0)=-(1,0)=-1$
2. If $z\in\mathbb{C}$ then $\exists x',y'\in\mathbb{R}$ such that $z=(x',y')$, define then $x=(x',0)\in\mathbb{R}_{\mathbb{C}}$ and $y=(y',0)\in\mathbb{R}_{\mathbb{C}}$ then we have

$$\begin{aligned}x+i\cdot y &= (x',0)+(0,1)\cdot(y',0)\\ &= (x',0)+(0\cdot y'-1\cdot 0,0\cdot 0+1\cdot y')\\ &= (x',0)+(0,y')\\ &= (x',y')\\ &= z\end{aligned}$$

Further if $u, v \in \mathbb{R}_{\mathbb{C}}$ such that $z = u + i \cdot v$ then $\exists u', v' \in \mathbb{R}$ such that $u = (u', 0) \wedge v = (v', 0)$ then we have

$$\begin{aligned}(z', y') &= z \\ &= u + i \cdot v \\ &= (u', 0) + (0, 1) \cdot (v', 0) \\ &= (u', 0) + (0 \cdot v' - 1 \cdot 0, 0 \cdot 0 + 1 \cdot v') \\ &= (u', 0) + (0, v') \\ &= (u', v')\end{aligned}$$

so that $x' = u'$ and $y' = v'$ hence $x = (x', 0) = (u', 0) = u$ and $y = (y', 0) = (v', 0) = v$ proving uniqueness. □

Note 10.76. We have now two ways to **uniquely** represent $z \in \mathbb{C}$ either $z = x + i \cdot y$ where $x, y \in \mathbb{R}_{\mathbb{C}}$ or $z = (x, y)$ where $x, y \in \mathbb{R}$. In this book we follow the common practice of using the first representation.

Definition 10.77. (Real & Imaginary part) *The above theorem allows us to define the following functions*

$$\mathrm{Re}\colon \mathbb{C} \rightarrow \mathbb{R}_{\mathbb{C}} \text{ is defined by } \mathrm{Re}(x + i \cdot y) = x$$

and

$$\mathrm{Img}\colon \mathbb{C} \rightarrow \mathbb{R}_{\mathbb{C}} \text{ by } \mathrm{Img}(x + i \cdot y) = y$$

So $\forall z \in \mathbb{C}$ *we have that*

$$z = \mathrm{Re}(z) + i \cdot \mathrm{Img}(z)$$

Theorem 10.78. *We have the following properties for* Re *and* Img

1. $z \in \mathbb{R}_{\mathbb{C}} \Leftrightarrow z = \mathrm{Re}(z)$
2. $z \in \mathbb{R}_{\mathbb{C}} \Leftrightarrow \mathrm{Img}(z) = 0$
3. $\forall z_1, z_2 \in \mathbb{C}$ *we have* $\mathrm{Re}(z_1 + z_2) = \mathrm{Re}(z_1) + \mathrm{Re}(z_2)$
4. $\forall z_1, z_2 \in \mathbb{C}$ *we have* $\mathrm{Img}(z_1 + z_2) = \mathrm{Img}(z_1) + \mathrm{Img}(z_2)$
5. $\forall z \in \mathbb{C}$ *and* $\alpha \in \mathbb{R}$ *then* $\mathrm{Re}(\alpha \cdot z) = \alpha \cdot \mathrm{Re}(z)$.
6. $\forall z \in \mathbb{C}$ *and* $\alpha \in \mathbb{R}$ *then* $\mathrm{Img}(\alpha \cdot z) = \alpha \cdot \mathrm{Img}(z)$

Proof.

1.
   - **⇒.** If $z \in \mathbb{R}_{\mathbb{C}}$ then $z = z + i \cdot 0$ where $z, 0 \in \mathbb{R}_{\mathbb{C}}$ so that $\mathrm{Re}(z) = z$
   - **⇐.** If $z = \mathrm{Re}(z)$ then as $\mathrm{Re}(z) \in \mathbb{R}_{\mathbb{C}}$ we have $z \in \mathbb{R}_{\mathbb{C}}$
2.
   - **⇒.** If $z \in \mathbb{R}_{\mathbb{C}}$ then $z = z + i \cdot 0$ where $z, 0 \in \mathbb{R}_{\mathbb{C}}$ so that $\mathrm{Img}(z) = 0$
   - **⇐.** If $\mathrm{Img}(z) = 0$ then $z = \mathrm{Re}(z) + \mathrm{Img}(z) = \mathrm{Re}(z) \in \mathbb{R}_{\mathbb{C}}$
3.
$$\begin{aligned}\mathrm{Re}(z_1 + z_2) &= \mathrm{Re}((\mathrm{Re}(z_1) + i \cdot \mathrm{Img}(z_1)) + (\mathrm{Re}(z_2) + i \cdot \mathrm{Img}(z_2))) \\ &= \mathrm{Re}((\mathrm{Re}(z_1) + \mathrm{Re}(z_2)) + i \cdot (\mathrm{Img}(z_1) + \mathrm{Img}(z_2))) \\ &= \mathrm{Re}(z_1) + \mathrm{Re}(z_2)\end{aligned}$$
4.
$$\begin{aligned}\mathrm{Img}(z_1 + z_2) &= \mathrm{Img}((\mathrm{Re}(z_1) + i \cdot \mathrm{Img}(z_1)) + (\mathrm{Re}(z_2) + i \cdot \mathrm{Img}(z_2))) \\ &= \mathrm{Img}((\mathrm{Re}(z_1) + \mathrm{Re}(z_2)) + i \cdot (\mathrm{Img}(z_1) + \mathrm{Img}(z_2))) \\ &= \mathrm{Img}(z_1) + \mathrm{Img}(z_2)\end{aligned}$$

5.

$$\begin{aligned}\mathrm{Re}(\alpha\cdot z) &\underset{(1)}{=} \mathrm{Re}(\mathrm{Re}(\alpha)\cdot(\mathrm{Re}(z)+i\cdot\mathrm{Img}(z)))\\ &= \mathrm{Re}(\mathrm{Re}(\alpha)\cdot\mathrm{Re}(z)+i\cdot(\mathrm{Re}(\alpha)\cdot\mathrm{Img}(z)))\\ &= \mathrm{Re}(\alpha)\cdot\mathrm{Re}(z)\\ &\underset{(1)}{=} \alpha\cdot\mathrm{Re}(z)\end{aligned}$$

6.

$$\begin{aligned}\mathrm{Img}(\alpha\cdot z) &\underset{(1)}{=} \mathrm{Img}(\mathrm{Re}(\alpha)\cdot(\mathrm{Re}(z)+i\cdot\mathrm{Img}(z)))\\ &= \mathrm{Img}(\mathrm{Re}(\alpha)\cdot\mathrm{Re}(z)+i\cdot(\mathrm{Re}(\alpha)\cdot\mathrm{Img}(z)))\\ &= \mathrm{Re}(\alpha)\cdot\mathrm{Re}(z)\\ &= \alpha\cdot\mathrm{Re}(z)\end{aligned}$$

□

Theorem 10.79. *For i we have*

1. $i^2=-1$
2. $\forall n\in\mathbb{N}_0$ *we have*
   a. $i^{4\cdot n}=1$
   b. $i^{4\cdot n+1}=i$
   c. $i^{4\cdot n+2}=-1$
   d. $i^{4\cdot n+3}=-i$

Proof.

1. $i^2=i\cdot i \underset{[\text{theorem: } 10.75]}{=} -1$
2. For $n\in\mathbb{N}$ we have:
   a. We use induction to prove this, so let $S=\{n\in\mathbb{N}|i^{4\cdot n}=1\}$ then we have

      $0\in S$. As $i^{4\cdot 0}=i^0=1$ we have that $1\in S$.

      $n\in S\Rightarrow n+1\in S$. We have

$$\begin{aligned}i^{4\cdot(n+1)} &= i^{4\cdot n+4}\\ &\underset{[\text{theorem: } 10.52]}{=} i^{4\cdot n}\cdot i^4\\ &\underset{n\in S}{=} 1\cdot i^4\\ &= i^{2+2}\\ &\underset{[\text{theorem: } 10.52]}{=} i^2\cdot i^2\\ &\underset{(1)}{=} (-1)\cdot(-1)\\ &= 1\end{aligned}$$

      proving that $n\in S$.

   b. $i^{4\cdot n+1} \underset{[\text{theorem: } 10.52]}{=} i^{4\cdot n}\cdot i \underset{(2.a)}{=} i$
   c. $i^{4\cdot n+2} \underset{[\text{theorem: } 10.52]}{=} i^{4\cdot n}\cdot i^2 \underset{(2\cdot a)}{=} i^2 \underset{(1)}{=} -1$
   d. $i^{4\cdot n+3} \underset{[\text{theorem: } 10.52]}{=} i^{4\cdot n}\cdot i^3 \underset{(2\cdot a)}{=} i^3 \underset{[\text{theorem: } 10.52]}{=} i^2\cdot i \underset{(2)}{=} (-1)\cdot i=-i$

□

10.5.2 Norm on $\mathbb{C}$

Definition 10.80. (Conjugate) *Let $z\in\mathbb{C}$ then the **conjugate** of z noted as $\bar{z}$ is defined as*

$$\bar{z}=\mathrm{Re}(z)-i\cdot\mathrm{Img}(z)$$

Theorem 10.81. *The conjugate has the following properties*

1. $\bar{i}=-i$
2. $\forall z\in\mathbb{C}$ *we have* $\overline{(\bar{z})}=z$
3. $\forall z_1,z_2\in\mathbb{C}$ *we have* $\overline{z_1+z_2}=\overline{z_1}+\overline{z_2}$
4. $\forall z_1,z_2\in\mathbb{C}$ *we have* $\overline{z_1\cdot z_2}=\overline{z_1}\cdot\overline{z_2}$
5. $\forall z\in\mathbb{C}\setminus\{0\}$ *we have* $\overline{z^{-1}}=(\bar{z})^{-1}$
6. $\forall z_1\in\mathbb{C}$, $\forall z_2\in\mathbb{C}\setminus\{0\}$ *we have* $\overline{z_1/z_2}=\overline{z_1}/\overline{z_2}$
7. $z\in\mathbb{R}_{\mathbb{C}}\Leftrightarrow z=\bar{z}$
8. $\forall z\in\mathbb{C}$ *we have that* $z\cdot z=\mathrm{Re}(z)^2+\mathrm{Img}(z)^2\in\mathbb{R}_0^+$
9. $\forall z\in\mathbb{C}$ *we have* $z+\bar{z}=2\cdot\mathrm{Re}(z)$
10. $\forall z\in\mathbb{C}$ *we have* $\overline{-z}=-\bar{z}$
11. *Let* $n\in\mathbb{N}_{\mathbb{C}}$ *and* $\{z_i\}_{i\in\{1,\ldots,n\}}\subseteq\mathbb{C}$ *then* $\overline{\sum_{i=1}^n z_i}=\sum_{i=1}^n\overline{z_i}$
12. *Let* $n\in\mathbb{N}_{\mathbb{C}}$ *and* $z\in\mathbb{C}$ *then* $\overline{z^n}=(\bar{z})^n$

Proof.

1. As $i=0+i\cdot 1$ we have that $\mathrm{Re}(i)=0$ and $\mathrm{Img}(i)=1$ so that $\bar{i}=\mathrm{Re}(i)-i\cdot\mathrm{Img}(i)=0-i=-i$
2.
$$\begin{aligned}\overline{(\bar{z})} &= \overline{\mathrm{Re}(z)-i\cdot\mathrm{Img}(z)}\\ &= \overline{(\mathrm{Re}(z)+i\cdot(-\mathrm{Img}(z)))}\\ &= \mathrm{Re}(z)-i\cdot(-\mathrm{Img}(z))\\ &= \mathrm{Re}(z)+i\cdot\mathrm{Img}(z)\\ &= z\end{aligned}$$
3.
$$\begin{aligned}\overline{z_1+z_2} &= \mathrm{Re}(z_1+z_2)-i\cdot\mathrm{Img}(z_1+z_2)\\ &\underset{[\text{theorem: }10.78]}{=} \mathrm{Re}(z_1)+\mathrm{Re}(z_2)-i\cdot(\mathrm{Img}(z_1)+\mathrm{Img}(z_2))\\ &= (\mathrm{Re}(z_1)-i\cdot\mathrm{Img}(z_1))+(\mathrm{Re}(z_2)-i\cdot\mathrm{Img}(z_2))\\ &= \overline{z_1}+\overline{z_2}\end{aligned}$$
4.
$$\begin{aligned}\overline{z_1\cdot z_2} &= \overline{(\mathrm{Re}(z_1)+i\cdot\mathrm{Img}(z_1))\cdot(\mathrm{Re}(z_2)+i\cdot\mathrm{Img}(z_2))}\\ &= \overline{\mathrm{Re}(z_1)\cdot\mathrm{Re}(z_2)-\mathrm{Img}(z_1)\cdot\mathrm{Img}(z_2)+i\cdot(\mathrm{Re}(z_1)\cdot\mathrm{Img}(z_2)+\mathrm{Img}(z_1)\cdot\mathrm{Re}(z_2))}\\ &= \mathrm{Re}(z_1)\cdot\mathrm{Re}(z_2)-\mathrm{Img}(z_1)\cdot\mathrm{Img}(z_2)-i\cdot(\mathrm{Re}(z_1)\cdot\mathrm{Img}(z_2)+\mathrm{Img}(z_1)\cdot\mathrm{Re}(z_2))\\ &= (\mathrm{Re}(z_1)-i\cdot\mathrm{Img}(z_1))\cdot(\mathrm{Re}(z_2)-i\cdot\mathrm{Img}(z_2))\\ &= \overline{z_1}\cdot\overline{z_2}\end{aligned}$$
5. We have $\overline{z^{-1}}\cdot\bar{z}=\bar{z}\cdot\overline{z^{-1}}\underset{(4)}{=}\overline{z\cdot z^{-1}}=\bar{1}=\overline{1+i\cdot 0}=1-i\cdot 0=1$ proving that $(\bar{z})^{-1}=\overline{z^{-1}}$
6. $\overline{z_1/z_2}=\overline{z_1\cdot z_2^{-1}}\underset{(4)}{=}\overline{z_1}\cdot\overline{z_2^{-1}}\underset{(5)}{=}\overline{z_1}\cdot(\overline{z_2})^{-1}=\overline{z_1}/\overline{z_2}$
7.

   ⇒. If $z\in\mathbb{R}_{\mathbb{C}}$ then $\bar{z}=\mathrm{Re}(z)-i\cdot\mathrm{Im}(z)\underset{[\text{theorem: }10.78]}{=}\mathrm{Re}(z)-i\cdot 0=\mathrm{Re}(z)\underset{[\text{theorem: }10.78]}{=}z$

⇐. If $z=\bar{z}$ then $\mathrm{Re}(z)+i\cdot\mathrm{Img}(z)=\mathrm{Re}(z)-i\cdot\mathrm{Img}(z)\Rightarrow 2\cdot\mathrm{Img}(z)=0\Rightarrow\mathrm{Img}(z)=0$ so that $z=\mathrm{Re}(z)\in\mathbb{R}_{\mathbb{C}}$ [see theorem: 10.78]

8.
$$\begin{aligned} z\cdot\bar{z} &= (\mathrm{Re}(z)+i\cdot\mathrm{Img}(z))\cdot(\mathrm{Re}(z)-i\cdot\mathrm{Img}(z)) \\ &= \mathrm{Re}(z)\cdot\mathrm{Re}(z)-i\cdot\mathrm{Re}(z)\cdot\mathrm{Img}(z)+i\cdot\mathrm{Img}(z)\cdot\mathrm{Re}(z)+i\cdot(-i)\cdot\mathrm{Img}(z)\cdot\mathrm{Img}(z) \\ &= \mathrm{Re}(z)^2+\mathrm{Img}(z)^2 \end{aligned}$$

As $\mathrm{Re}(z),\mathrm{Img}(z)\in\mathbb{R}_{\mathbb{C}}$ we have by [theorem: 10.14] that $0\leqslant\mathrm{Re}(z)^2+\mathrm{Img}(z)^2$ so that $z\cdot\bar{z}\in\mathbb{R}^+_{0,\mathbb{C}}$.

9. $z+\bar{z}=(\mathrm{Re}(z)+i\cdot\mathrm{Img}(z))+(\mathrm{Re}(z)-i\cdot\mathrm{Img}(z))=2\cdot\mathrm{Re}(z)$

10.
$$\begin{aligned} \overline{-z} &= \overline{-(\mathrm{Re}(z)+i\cdot\mathrm{Img}(z))} \\ &= \overline{-\mathrm{Re}(z)+i\cdot(-\mathrm{Img}(z))} \\ &= -\mathrm{Re}(z)-i\cdot(-\mathrm{Img}(z)) \\ &= -(\mathrm{Re}(z)-i\cdot\mathrm{Img}(z)) \\ &= -\bar{z} \end{aligned}$$

11. We prove this by induction, so let
$$S=\left\{n\in\mathbb{N}|\forall\{z_i\}_{i\in\{1,\ldots,n\}}\subseteq\mathbb{C}\text{ we have }\overline{\sum_{i=1}^n z_i}=\sum_{i=1}^n\overline{z_i}\right\}$$
then we have:

$\boldsymbol{1\in\mathcal{S}}$. this follows from $\overline{\sum_{i=1}^1 z_i}=\overline{z_1}=\sum_{i=1}^1\overline{z_i}$ for $\{z_i\}_{i\in\{1,\ldots,1\}}\subseteq\mathbb{C}$

$\boldsymbol{n\in\mathcal{S}\Rightarrow n+1\in\mathcal{S}}$. we have for $\{z_i\}_{i\in\{1,\ldots,n+1\}}$ that
$$\begin{aligned} \overline{\sum_{i=1}^{n+1} z_i} &= \overline{z_{n+1}+\sum_{i=1}^n z_i} \\ &\underset{(3)}{=} \overline{z_{n+1}}+\overline{\sum_{i=1}^n z_i} \\ &\underset{n\in\mathcal{S}}{=} \overline{z_{n+1}}+\sum_{i=1}^n\overline{z_i} \\ &= \sum_{i=1}^{n+1}\overline{z_i} \end{aligned}$$

proving that $n+1\in\mathcal{S}$.

12. We prove this by induction, so let $S=\{n\in\mathbb{N}_{\mathbb{C},0}|\overline{z^n}=(\bar{z})^n\}$ then we have:

$\boldsymbol{0\in S}$. As $\overline{z^0}=\bar{1}\underset{(7)}{=}1=(\bar{z})^0$ proving that $0\in S$.

$\boldsymbol{n\in S\Rightarrow n+1\in S}$. Then we have
$$\overline{z^{n+1}} = \overline{z^n\cdot z}\underset{(4)}{=}\overline{z^n}\cdot\bar{z}\underset{n\in S}{=}(\bar{z})^n\cdot\bar{z}=(\bar{z})^{n+1}$$
□

Using the above theorem we have that $\forall z\in\mathbb{C}\ z\cdot\bar{z}\in\mathbb{R}^+_{0,\mathbb{C}}$ so that $\sqrt{z\cdot\bar{z}}\in\mathbb{R}^+_{0,\mathbb{C}}$ is well defined, hence the following definition makes sense.

Definition 10.82. (complex norm)
$$||:\mathbb{C}\rightarrow\mathbb{R}^+_{0,\mathbb{C}}\textit{ is defined by }|z|=\sqrt{z\cdot\bar{z}}\underset{\text{[theorem: 10.81]}}{=}\sqrt{\mathrm{Re}(z)^2+\mathrm{Img}(z)^2}$$

Theorem 10.83. *$\forall z,z'\in\mathbb{C}$ we have*

1. $|z\cdot z'|=|z|\cdot|z'|$
2. $|\bar{z}|=|z|$

3. $\mathrm{Re}(z) \leqslant |\mathrm{Re}(z)| \leqslant_{\mathbb{R}} |z|$ *and* $\mathrm{Img}(z) \leqslant |\mathrm{Img}(z)| \leqslant_{\mathbb{R}} |z|$
4. $z \cdot \overline{z} = |z|^2$
5. $|z + z'| \leqslant |z| + |z'|$
6. $|-z| = |z|$
7. $|1| = |-1| = |i| = 1$
8. $|z| \leqslant |\mathrm{Re}(z)| + |\mathrm{Img}(z)|$
9. $|z| = 0 \Leftrightarrow z = 0$
10. *If* $z \neq 0$ *then* $|z^{-1}| = |z|^{-1}$ *or in other words* $|1/z| = 1/|z|$
11. *If* $z \in \mathbb{R}_{\mathbb{C}}$ *then*
$$|z| = \begin{cases} z \text{ if } 0 \leqslant z \\ -z \text{ if } z < 0 \end{cases}$$
12. *If* $n \in \mathbb{N}_{\mathbb{C},0}$ *then* $|z^n| = |z|^n$

Proof.

1.
$$\begin{aligned} |z \cdot z'| &= \sqrt{(z \cdot z') \cdot \overline{(z \cdot z')}} \\ &\underset{[\text{theorem: } 10.81]}{=} \sqrt{(z \cdot z') \cdot (\overline{z} \cdot \overline{z'})} \\ &= \sqrt{(z \cdot \overline{z}) \cdot (z' \cdot \overline{z'})} \\ &\underset{[\text{theorem: } 10.72]}{=} \sqrt{z \cdot \overline{z}} \cdot \sqrt{z' \cdot \overline{z'}} \\ &= |z| \cdot |z'| \end{aligned}$$

2. $|\overline{z}| = \sqrt{\overline{z} \cdot \overline{(\overline{z})}} \underset{[\text{theorem: } 10.81]}{=} \sqrt{\overline{z} \cdot z} = \sqrt{z \cdot \overline{z}} = |z|$

3. Using [corollary: 10.71] we have that
$$\mathrm{Re}(z) \leqslant \sqrt{\mathrm{Re}(x)^2} = \sqrt{\mathrm{Re}(x)^2 + 0^2} = |\mathrm{Re}(z)|$$
and
$$\mathrm{Img}(z) \leqslant \sqrt{(\mathrm{Img}(z))^2} = \sqrt{0 + \mathrm{Img}(z)^2} = |\mathrm{Img}(z)|$$
Further as $\mathrm{Re}(z)^2 \leqslant \mathrm{Re}(z)^2 + \mathrm{Img}(z)^2$ and $\mathrm{Img}(z)^2 \leqslant \mathrm{Re}(z)^2 + \mathrm{Img}(z)^2$ and [theorem: 10.69] $\sqrt{\cdot}$ is increasing it follows that
$$\sqrt{\mathrm{Re}(z)^2} \leqslant \sqrt{\mathrm{Re}(z)^2 + \mathrm{Img}(z)^2} = |z| \text{ and } \sqrt{\mathrm{Img}(z)^2} \leqslant \sqrt{\mathrm{Re}(z)^2 + \mathrm{Img}(z)^2} = |z|$$
so that $\mathrm{Re}(z) \leqslant |\mathrm{Re}(z)| \leqslant_{\mathbb{R}} |z|$ and $\mathrm{Img}(z) \leqslant |\mathrm{Img}(z)| \leqslant_{\mathbb{R}} |z|$.

4. $|z|^2 = (\sqrt{z \cdot \overline{z}})^2 \underset{0 \leqslant z \cdot \overline{z}}{=} z \cdot \overline{z}$

5. We have
$$\begin{aligned} |z + z'|^2 &\underset{(5)}{=} (z + z') \cdot \overline{(z + z')} \\ &\underset{[\text{theorem: } 10.81]}{=} (z + z') \cdot (\overline{z} + \overline{z'}) \\ &= z \cdot \overline{z} + z \cdot \overline{z'} + z' \cdot \overline{z} + z' \cdot \overline{z'} \\ &\underset{(4)}{=} |z|^2 + |z'|^2 + z \cdot \overline{z'} + z' \cdot \overline{z} \\ &\underset{[\text{theorem: } 10.81]}{=} |z|^2 + |z'|^2 + z \cdot \overline{z'} + \overline{z \cdot \overline{(z')}} \\ &\underset{[\text{theorem: } 10.81]}{=} |z|^2 + |z'|^2 + 2 \cdot \mathrm{Re}(z \cdot \overline{z'}) \\ &\leqslant_{(3)} |z|^2 + |z'|^2 + 2 \cdot |z \cdot \overline{z'}| \\ &\underset{(1)}{=} |z|^2 + |z'|^2 + 2 \cdot |z| \cdot |\overline{z'}| \\ &= (|z| + |z'|)^2 \end{aligned}$$

which as $\sqrt{.}$ is increasing [see theorem: 10.69] and $0 \leqslant |z+z'|, |z|+|z'|$ we have that

$$|z+z'| = \sqrt{|z+z'|^2} \leqslant \sqrt{(|z|+|z'|)^2} = |z|+|z'|$$

6. $|-z| = \sqrt{(-z)\cdot\overline{(-z)}} \underset{\text{[theorem: 10.81]}}{=} \sqrt{(-z)\cdot(-\bar{z})} = \sqrt{z\cdot\bar{z}} = |z|$

7.

$$|-1| \underset{(6)}{=} |1| = \sqrt{1\cdot\bar{1}} \underset{1\in\mathbb{R}_{\mathbb{C}} \text{ and [theorem: 10.81]}}{=} \sqrt{1\cdot 1} = \sqrt{1} \underset{\text{[example: 10.67]}}{=} 1$$

and

$$|i| = \sqrt{i\cdot\bar{i}} \underset{\text{[theorem: 10.81]}}{=} \sqrt{-(i\cdot i)} = \sqrt{1} = 1$$

8. $|z| = |\mathrm{Re}(x) + i\cdot\mathrm{Img}(z)| \leqslant_{(5)} |\mathrm{Re}(z)| + |i\cdot\mathrm{Img}(z)| \underset{(4)}{=} |\mathrm{Re}(z)| + |i|\cdot|\mathrm{img}(z)| \underset{(7)}{=} |\mathrm{Re}(z)| + |\mathrm{Img}(z)|$

9.

⇒. If $z=0$ then $z = 0 + i\cdot 0$ so that $\mathrm{Re}(x) = \mathrm{Img}(z) = 0$ hence

$$|z| = \sqrt{\mathrm{Re}(z)^2 + \mathrm{Img}(z)^2} = \sqrt{0^2+0^2} \underset{\text{[example: 10.67]}}{=} 0$$

⇐. If $|z|=0$ then $\sqrt{\mathrm{Re}(z)^2+\mathrm{Img}(z)^2} = 0 \underset{\text{[example: 10.67]}}{=} \sqrt{0}$ so as $\sqrt{.}$ is injective it follows that $\mathrm{Re}(z)^2+\mathrm{Img}(z)^2 = 0$. Using [theorem: 10.14] we have that $0 \leqslant \mathrm{Re}(z)^2, \mathrm{Img}(z)^2$ so that by [theorem: 10.14] again we have that

$$\mathrm{Re}(z)^2, \mathrm{Img}(z)^2 \leqslant \mathrm{Re}(z)^2 + \mathrm{Img}(z)^2$$

If now either $\mathrm{Re}(z) \neq 0$ or $\mathrm{Img}(z) \neq 0$ then by [theorem: 10.14] $0 < \mathrm{Re}(z)$ or $0 < \mathrm{Img}(z)$ which would give $0 < \mathrm{Re}(z)^2 + \mathrm{Img}(z)^2$ contradicting $\mathrm{Re}(z)^2 + \mathrm{Img}(z)^2 = 0$, Hence we have that $\mathrm{Re}(z) = 0 = \mathrm{Img}(z)$ so that $z = \mathrm{Re}(z) + i\cdot\mathrm{Img}(z) = 0$.

10. If $z \in \mathbb{C}\setminus\{0\}$ then as $1 = z\cdot z^{-1}$ we have $1 \underset{(7)}{=} |1| = |z\cdot z^{-1}| \underset{(1)}{=} |z|\cdot|z^{-1}|$ so that $|z^{-1}| = |z|^{-1}$

11. If $z \in \mathbb{R}^+_{0,\mathbb{C}}$ then we have $\mathrm{Re}(z) \underset{\text{[theorem: 10.78]}}{=} z \in \mathbb{R}^+_{0,\mathbb{C}}$ and $\mathrm{Img}(z) = 0$ so that

$$z = \mathrm{Re}(z) = \sqrt{\mathrm{Re}(z)^2} = \sqrt{\mathrm{Re}(z)^2 + \mathrm{Img}(z)^2} = |z|$$

proving that

$$\forall z \in \mathbb{R}^+_{0,\mathbb{C}} \text{ we have } z = |z| \tag{10.70}$$

Now if $z \in \mathbb{R}_{\mathbb{C}}$ we have either

$0 \leqslant z$. Then $z \in \mathbb{R}^+_{0,\mathbb{C}}$ and by [eq: 10.70] we have $z = |z|$

$z < 0$. Then $0 < -z$ so that by [eq: 10.70] we have $-z = |-z| \underset{(6)}{=} |z|$ proving $-z = |z|$

12. We prove this by induction so let $S = \{i \in \{0,\ldots,\infty\} \mid |z^n| = |z|^n\}$ then we have:

$0 \in S$. As $|z^0| = |1| \underset{(11)}{=} 1 = |1|^0$ proving that $0 \in S$.

$n \in S \Rightarrow n+1 \in S$. We have

$$\begin{aligned} |z^{n+1}| &= |z\cdot z^n| \\ &\underset{(1)}{=} |z|\cdot|z^n| \\ &\underset{n\in S}{=} |z|^{n+1} \end{aligned}$$

proving that $n+1 \in S$.

□

10.5.3 Finite sets

We define now a characterization of finite sets in terms of $\mathbb{N}_{0,\mathbb{C}}$.

Lemma 10.84. *If $n \in \mathbb{N}_0$ then $\{1, \ldots, n\} \approx \{1, \ldots, i_{\mathbb{N}_0 \to \mathbb{C}}(n)\}$*

Proof. First we have by [theorem: 10.5] that $i_{\mathbb{N}_0 \to \mathbb{Z}}: \mathbb{N}_0 \to \mathbb{N}_{0,\mathbb{C}}$ is a bijection, hence $i_{\mathbb{N}_0 \to \mathbb{C}}$ is injective. Further we have by [theorem: 10.34] that

$$i_{\mathbb{N}_0 \to \mathbb{C}}(\{1, \ldots, n\}) = \{i_{\mathbb{N}_0 \to \mathbb{C}}(1), \ldots, i_{\mathbb{N}_0, \to \mathbb{C}}(n)\} \underset{\text{[theorem: 10.5}}{=} \{1, \ldots, i_{\mathbb{N}_0 \to \mathbb{C}}(n)\}$$

so that

$$(i_{\mathbb{N}_0 \to \mathbb{C}})_{|\{1, \ldots, n\}}: \{1, \ldots, n\} \to \{1, \ldots, i_{\mathbb{N}_0}(n)\}$$

is a bijection. □

Theorem 10.85. *We have the following characterization of finite sets then*

$$I \text{ is finite} \Leftrightarrow \text{there exists a unique } k \in \mathbb{N}_{\mathbb{C},0} \text{ such that } \{1, \ldots, k\} \approx I$$

Proof.

⇒. As I is finite we have by [theorem: 6.22] that $\exists n \in \mathbb{N}_0$ such that $\{1, \ldots, n\} \approx I$. Hence if we take $k = i_{\mathbb{N}_0 \to \mathbb{C}}(n) \in \mathbb{N}_{\mathbb{C},0}$ we have by [lemma: 10.84] that

$$\{1, \ldots, n\} \approx \{1, \ldots, i_{\mathbb{N}_0 \to \mathbb{C}}(n)\} = \{1, \ldots, k\}$$

proving that

$$\{1, \ldots, k\} \approx I$$

Now for uniqueness, assume that there is another $l \in \mathbb{N}_{\mathbb{C},0}$ such that $\{1, \ldots, l\} \approx I$. Take then $m = (i_{\mathbb{N}_0 \to \mathbb{C}})^{-1}(l)$ we have by [lemma: 10.84] that

$$\{1, \ldots, m\} \approx \{1, \ldots, i_{\mathbb{N}_0 \to \mathbb{C}}(m)\} = \{1, \ldots, i_{\mathbb{N}_0 \to \mathbb{C}}((i_{\mathbb{N}_0 \to \mathbb{C}})^{-1}(l))\} = \{1, \ldots, l\} \approx I$$

so that $\{1, \ldots, n\} \approx \{1, \ldots, m\}$, which, as by [theorem: 6.21] $\{1, \ldots, n\} \approx n \wedge \{1, \ldots, m\} \approx m$, gives that $n \approx m$. So that $(i_{\mathbb{N}_0 \to \mathbb{C}})^{-1}(k) = n = m = (i_{\mathbb{N}_0 \to \mathbb{C}})^{-1}(l)$ giving as $(i_{\mathbb{N}_0 \to \mathbb{C}})^{-1}$ is a bijection that $k = l$.

⇐. Let $k \in \mathbb{N}_{0,\mathbb{C}}$ be such that $I \approx \{1, \ldots, k\}$ then if we define $n \in \mathbb{N}_0$ by $n = (i_{\mathbb{N}_0 \to \mathbb{C}})^{-1}(k)$ we have by [lemma: 10.84] that

$$\{1, \ldots, n\} \approx \{1, \ldots, i_{\mathbb{N}_0 \to \mathbb{C}}(n)\} = \{1, \ldots, i_{\mathbb{N}_0 \to \mathbb{C}}((i_{\mathbb{N}_0 \to \mathbb{C}})^{-1}(k))\} = \{1, \ldots, k\} \approx I$$

proving that $\{1, \ldots, n\} \approx I$, hence by [theorem: 6.22] I is finite. □

Corollary 10.86. *Let I be a non empty set then*

$$I \text{ is finite} \Leftrightarrow \text{there exists a } \textbf{unique } k \in \mathbb{N}_{\mathbb{C}} \text{ such that } \{0, \ldots, k-1\} \approx I$$

Proof. Let $k \in \mathbb{N}_{\mathbb{C}} \Rightarrow k - 1 \in \mathbb{N}_{\mathbb{C},0}$. Define

$$\tau: \{0, \ldots, k-1\} \to \{1, \ldots, k\} \text{ by } \tau(i) = i + 1$$

then we have:

injectivity. If $\gamma(i) = \gamma(j)$ then $i + 1 = j + 1$ hence $i = j$

surjectity. If $j \in \{1, \ldots, k\}$ then $1 \leqslant j \leqslant k \Rightarrow 0 \leqslant j - 1 \leqslant k - 1$ proving that $j - 1 \in \{0, \ldots, k-1\}$. As $\gamma(j-1) = (j-1) + 1 = j$ surjectivity is proved.

So we have that

$$\text{If } k \in \mathbb{N}_{\mathbb{C}} \text{ then } \{0, \ldots, k-1\} \approx \{1, \ldots, k\}$$

⇒. If I is finite there exists by [theorem: 10.85] a $k \in \mathbb{N}_{\mathbb{C},0}$ such that $\{1, \ldots, k\} \approx I$ hence there exists a bijection

$$\alpha: \{1, \ldots, k\} \to I.$$

If $k = 0$ then $\{1, \ldots, k\} = \varnothing$ so that $I = \alpha(\varnothing) = \varnothing$ contradicting $I \neq \varnothing$., hence $k \in \mathbb{N}_{\mathbb{C}}$. As $\{0, \ldots, k-1\} \approx \{1, \ldots, k\}$ and $\{1, \ldots, k\} \approx I$ it follows that

$$\{0, \ldots, k-1\} \approx I$$

As for uniqueness. If $l \in \mathbb{N}_{\mathbb{C}}$ such that $\{0, \ldots, l-1\} \approx I$ then as $\{0, \ldots, l-1\} \approx \{1, \ldots, l\}$ we have also $\{1, \ldots, l\} \approx I$. Using the previous theorem [theorem: 10.85] it follows then that $k = l$.

$\Leftarrow$. If $k \in \mathbb{N}_{\mathbb{C}}$ such that $\{0, \ldots, k-1\} \approx I$ then as $\{0, \ldots, k-1\} \approx \{1, \ldots, k\}$ it follows that $\{1, \ldots, k\} \approx I$. Using the previous theorem [theorem: 10.85] it follows then that I is finite. $\square$

Definition 10.87. *Let I be a finite set then* $\operatorname{card}(I)$ *is defined as follows*

$$\operatorname{card}(I) = i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(\#(I))$$

where $\#(I)$ is defined in [definition: 6.38]

Theorem 10.88. *If I is a finite set then we have that*

$$\operatorname{card}(I) \text{ is the unique } k \in \mathbb{N}_{0,\mathbb{C}} \text{ such that } \{1, \ldots, k\} \approx I$$

In particular

$$I = \varnothing \Leftrightarrow \operatorname{card}(I) = 0$$

Proof. By definition we have that $\#(I) = n$ where $n \approx I$ then

$$\operatorname{card}(I) = i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(\#(I)) = i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(n)$$

As by [theorem: 6.21] we have that $\{1, \ldots, n\} \approx n$, further by [lemma: 10.84]

$$\{1, \ldots, n\} \approx \{1, \ldots, i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(n)\} = \{1, \ldots, \operatorname{card}(I)\}$$

so that $I \approx \{1, \ldots, \operatorname{card}(I)\}$. Finally we have

$$I = \varnothing \underset{[\text{theorem: } 6.39]}{\Leftrightarrow} \#(I) = 0 \Leftrightarrow i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(\#(I)) = 0 \Leftrightarrow \operatorname{card}(I) = 0$$ $\square$

Theorem 10.89. *If A, B are sets such that A is finite then*

1. *If $B \subseteq A$ then B is finite and* $\operatorname{card}(B) \leqslant \operatorname{card}(A)$
2. *If $B \subset A$ then B is finite and* $\operatorname{card}(B) < \operatorname{card}(A)$

Proof.

1. Using [theorem: 6.42] it follows that B is finite and $\#(B) \leqslant \#(A)$, so, as by [theorem: 10.9] is a order isomorphism, we have $\operatorname{card}(B) = i_{\mathbb{N}_9 \rightarrow \mathbb{C}}(\#(B)) \leqslant i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(\#(A)) = \operatorname{card}(B)$.
2. Using [theorem: 6.42] it follows that B is finite and $\#(B) < \#(A)$, so, as by [theorem: 10.9] is a order isomorphism, we have $\operatorname{card}(B) = i_{\mathbb{N}_9 \rightarrow \mathbb{C}}(\#(B)) < i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(\#(A)) = \operatorname{card}(B)$ $\square$

Theorem 10.90. *If A, B are finite sets then*

1. *$A \times B$ is finite and* $\operatorname{card}(A \times B) = \operatorname{card}(A) \cdot \operatorname{card}(B)$
2. *$A \bigcup B$ is finite and $A \bigcap B = \varnothing$ then* $\operatorname{card}(A \bigcup B) = \operatorname{card}(A) + \operatorname{card}(B)$

Proof.

1. By [theorem: 6.40] we have that $A \times B$ is finite and $\#(A \times B) = \#(A) \cdot \#(B)$. Further we have that

$$\begin{array}{rcl}\operatorname{card}(A \times B) & = & i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(\#(A \times B)) \\ & = & i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(\#(A) \cdot \#(B)) \\ & = & i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(\#(A)) \cdot i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(\#(B)) \\ & = & \operatorname{card}(A) \cdot \operatorname{card}(B)\end{array}$$

2. By [theorems: 6.33,6.41] we have that $A \bigcup B$ is finite and $\#(A \bigcup B) = \#(A) + \#(B)$. Further we have that

$$\begin{array}{rcl}\operatorname{card}\left(A \bigcup B\right) & = & i_{\mathbb{N}_0 \rightarrow \mathbb{C}}\left(\#\left(A \bigcup B\right)\right) \\ & = & i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(\#(A) + \#(B)) \\ & = & i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(\#(A)) + i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(\#(B)) \\ & = & \operatorname{card}(A) + \operatorname{card}(B)\end{array}$$

$\square$

Corollary 10.91. *If A is a finite set and $a \notin A$ then* $\operatorname{card}(A \bigcup \{a\}) = \operatorname{card}(A) + 1$

Proof. As $A \notin A$ we have that $A \bigcap \{a\} = \varnothing$ so by [theorem: 10.90] we have

$$\operatorname{card}\left(A \bigcup \{a\}\right) = \operatorname{card}(A) + \operatorname{card}(\{a\}) = \operatorname{card}(A) + 1$$

□

Corollary 10.92. *If A is a finite set and $a \in A$ then* $\operatorname{card}(A \setminus \{a\}) = \operatorname{card}(A) - 1$

Proof. As $a \notin A \setminus \{a\}$ and $A = (A \setminus \{a\}) \bigcup \{a\} = A$ we have that

$$\operatorname{card}(A) \underset{[\text{corollary: } 10.91]}{=} \operatorname{card}(A \setminus \{a\}) + 1$$

so that

$$\operatorname{card}(A \setminus \{a\}) = \operatorname{card}(A) - 1$$

□

Corollary 10.93. *If A is a finite set and $B \subseteq A$ such that* $\operatorname{card}(B) = \operatorname{card}(A)$ *then* $B = A$

Proof. Assume that $B \neq A$ then as $B \subseteq A$ there exist a $a \in A$ such that $a \notin B$. Hence $B \bigcup \{a\} \subseteq A$ so that $\operatorname{card}(B \bigcup \{a\}) \leqslant \operatorname{card}(A)$, further we have by [corollary: 10.91] that $\operatorname{card}(B \bigcup \{a\}) = \operatorname{card}(B) + 1$ so that $\operatorname{card}(B) + 1 \leqslant \operatorname{card}(A)$ or $\operatorname{card}(B) < \operatorname{card}(A)$ contradicting $\operatorname{card}(A) = \operatorname{card}(B)$. Hence we must have that $A = B$. □

Theorem 10.94. *Let I be a finite set, $\{x_i\}_{i \in I} \subseteq X$ then $\{x_i | i \in I\}$ is finite and*

$$\operatorname{card}(\{x_i | i \in I\}) \leqslant \operatorname{card}(I)$$

Proof. This follows from [theorem: 6.45]. □

10.5.4 Extended real numbers

Finally we define the set of extended real numbers which usefull if we have to work with numbers that are bigger or lower then every real number. This will be usefull later for limits, dimensions.

Lemma 10.95. *There exists at least two different elements that are not element of $\mathbb{R}_{\mathbb{C}}$*

Proof. Using [definitions: 9.1, 9.2] it follows that $\varnothing \notin \mathbb{R}$ and $\mathbb{Q} \notin \mathbb{R}$ and as $0 \in \mathbb{Q}$ we have $\varnothing \neq \mathbb{Q}$ So that $(\varnothing, 0) \notin \mathbb{R}_{\mathbb{C}}$ and $(\mathbb{Q}, 0) \notin \mathbb{R}_{\mathbb{C}}$ and $(\varnothing, 0) \neq (\mathbb{Q}, 0)$ □

Definition 10.96. *The set of extended real numbers $\overline{\mathbb{R}}$ is defined as*

$$\overline{\mathbb{R}} = \mathbb{R}_{\mathbb{C}} \bigcup \{\infty, -\infty\}$$

where $\infty, -\infty \notin \mathbb{R}_{\mathbb{C}}$ and $\infty \neq -\infty$

Definition 10.97. *A $x \in \overline{\mathbb{R}}$ is a called a finite real number if $x \in \mathbb{R}_{\mathbb{C}}$ so $\mathbb{R}_{\mathbb{C}}$ is the set of finite real numbers.*

Definition 10.98. $\overline{\leqslant} \in \overline{\mathbb{R}} \times \overline{\mathbb{R}}$ *is defined as follows*

$$\begin{array}{rl} \{(-\infty, -\infty), (-\infty, \infty), (\infty, \infty)\} & \bigcup \ \{(x, \infty) | x \in \mathbb{R}_{\mathbb{C}}\} \\ & \bigcup \ \{(-\infty, x) | x \in \mathbb{R}_C\} \\ & \bigcup \ \{(x, y) \in \mathbb{R}_{\mathbb{C}} \times \mathbb{R}_{\mathbb{C}} | x \leqslant y\} \end{array}$$

Note 10.99. As $\{-\infty, \infty\} \bigcap \mathbb{R} = \emptyset$ and $-\infty \neq \infty$ we have $\forall x \in \mathbb{R}_{\mathbb{C}}$ we have $-\infty < x$ and $x < \infty$

Theorem 10.100. $\langle \overline{\mathbb{R}}, \overline{\leqslant} \rangle$ *is fully ordered.*

Proof.

reflexitivity. The following cases occurs for $x \in \overline{\mathbb{R}}$

$x = \infty$. then by definition $x \overline{\leqslant} x$

$x = -\infty$. then by definition $x \overline{\leqslant} x$

$x \in \mathbb{R}_{\mathbb{C}}$. then as $x \leqslant_{\mathbb{R}} x \Rightarrow x \overline{\leqslant} x$

proving reflexitivity.

anti-symmetry. Let $x, y \in \overline{\mathbb{R}}$ with $x \overline{\leqslant} y \wedge y \overline{\leqslant} x$ then the following cases must be considered for $x, y \in \overline{\mathbb{R}}$:

$x = \infty \wedge y = \infty$. then $x = y$

$x = -\infty \wedge y = \infty$. then as by the definition $\neg(y \overline{\leqslant} x)$ this case will not apply.

$x \in \mathbb{R} \wedge y = \infty$. then as by definition $y \overline{\nleqslant} x$ this case does not apply

$x = \infty \wedge y = -\infty$. then as by definition $\neg(x \overline{\leqslant} y)$ this case does not apply

$x = -\infty \wedge y = -\infty$. then $x = y$

$x \in \mathbb{R} \wedge y = -\infty$. then asby definition $\neg(x \overline{\leqslant} y)$ the case does not apply

$x = \infty \wedge y \in \mathbb{R}$. then as by definition $\neg(x \overline{\leqslant} y)$ this case does not apply

$x = -\infty \wedge y \in \mathbb{R}$. then as by definiton $\neg(y \overline{\leqslant} x)$ this case does not apply

$x \in \mathbb{R} \wedge y \in \mathbb{R}$. then by defintion $x \leqslant y$ and $y \leqslant x$ from which it follows that $x = y$

so in all the cases where $x \overline{\leqslant} y \wedge y \overline{\leqslant} x$ we have $x = y$

transitivity. Let $x, y, z \in \overline{\mathbb{R}}$ with $x \overline{\leqslant} y \wedge y \overline{\leqslant} z$ then the following cases must be considered for $x, y, z \in \overline{\mathbb{R}}$:

$x = \infty \wedge y = \infty \wedge z = \infty$. then $x \overline{\leqslant} z$

$x = -\infty \wedge y = \infty \wedge z = \infty$. then $x \overline{\leqslant} z$

$x \in \mathbb{R}_{\mathbb{C}} \wedge y = \infty \wedge z = \infty$. then $x \overline{\leqslant} z$

$x = \infty \wedge y = -\infty \wedge z = \infty$. then as $\neg(x \overline{\leqslant} y)$ this cases does not count

$x = -\infty \wedge y = -\infty \wedge z = \infty$. then $x \overline{\leqslant} z$

$x \in \mathbb{R}_{\mathbb{C}} \wedge y = -\infty \wedge z = \infty$. then as $\neg(x \overline{\leqslant} y)$ this case does not count

$x = \infty \wedge y \in \mathbb{R}_{\mathbb{C}} \wedge z = \infty$. then as $\neg(x \overline{\leqslant} y)$ this case does not count

$x = -\infty \wedge y \in \mathbb{R}_{\mathbb{C}} \wedge z = \infty$. then $x \overline{\leqslant} z$

$x \in \mathbb{R}_{\mathbb{C}} \wedge y \in \mathbb{R}_{\mathbb{C}} \wedge z = \infty$. then $x \overline{\leqslant} z$

$x = \infty \wedge y = \infty \wedge z = -\infty$. then as $\neg(y \overline{\leqslant} x)$ this case does not count

$x = -\infty \wedge y = \infty \wedge z = -\infty$. then as $\neg(y \overline{\leqslant} x)$ this case does not count

$x \in \mathbb{R}_{\mathbb{C}} \wedge y = \infty \wedge z = -\infty$. then as $\neg(y \overline{\leqslant} x)$ this case does not count

$x = \infty \wedge y = -\infty \wedge z = -\infty$. then as $\neg(x \overline{\leqslant} y)$ this case does not count

$x = -\infty \wedge y = -\infty \wedge z = -\infty$. then $x \overline{\leqslant} z$

$x \in \mathbb{R}_{\mathbb{C}} \wedge y = -\infty \wedge z = -\infty$. then as $\neg(x \overline{\leqslant} y)$ this case does not count

$x = \infty \wedge y \in \mathbb{R}_{\mathbb{C}} \wedge z = -\infty$. then as $\neg(y \overline{\leqslant} z)$ this case does not count

$x = -\infty \wedge y \in \mathbb{R}_{\mathbb{C}} \wedge z = -\infty$. then as $\neg(y \overline{\leqslant} z)$ this case does not count

$x \in \mathbb{R}_{\mathbb{C}} \wedge y \in \mathbb{R}_{\mathbb{C}} \wedge z = -\infty$. then as $\neg(y \overline{\leqslant} z)$ this case does not count

$x = \infty \wedge y = \infty \wedge z \in \mathbb{R}_{\mathbb{C}}$. then as $\neg(y \overline{\leqslant} z)$ this case does not count

$x = -\infty \wedge y = \infty \wedge z \in \mathbb{R}_{\mathbb{C}}$. then as $\neg(y \overline{\leqslant} z)$ this case does not count

$x \in \mathbb{R}_{\mathbb{C}} \wedge y = \infty \wedge z \in \mathbb{R}_{\mathbb{C}}$. then as ¬(y$\overline{\leqslant}$z) this case does not count

$x = \infty \wedge y = -\infty \wedge z \in \mathbb{R}_{\mathbb{C}}$. then as $\neg(x \overline{\leqslant} y)$ this case does not count

$x - \infty \wedge y = -\infty \wedge z \in \mathbb{R}_{\mathbb{C}}$. then $x \overline{\leqslant} z$

$x \in \mathbb{R}_{\mathbb{C}} \wedge y = -\infty \wedge z \in \mathbb{R}_{\mathbb{C}}$. then as ¬(x$\overline{\leqslant}$y) this case does not count

$x = \infty \wedge y \in \mathbb{R}_{\mathbb{C}} \wedge z \in \mathbb{R}_{\mathbb{C}}$. then as ¬(x$\overline{\leqslant}$y) this case does not count

$x=-\infty \wedge y \in \mathbb{R}_{\mathbb{C}} \wedge z \in \mathbb{R}_{\mathbb{C}}$. then $x \overline{\leqslant} z$

$x \in \mathbb{R}_{\mathbb{C}} \wedge y \in \mathbb{R}_{\mathbb{C}} \wedge z \in \mathbb{R}_{\mathbb{C}}$. then $x \overline{\leqslant} z$

so in all cases that count we have $x \overline{\leqslant} z$

fully-ordered. The following cases must be considered for $x, y \in \overline{\mathbb{R}}$:

$x=\infty \wedge y=\infty$. then $x \overline{\leqslant} y$

$x=-\infty \wedge x=\infty$. then $x \overline{\leqslant} y$

$x \in \mathbb{R}_{\mathbb{C}} \wedge y=\infty$. then $x \overline{\leqslant} y$

$x=\infty \wedge y=-\infty$. then $y \overline{\leqslant} x$

$x=-\infty \wedge y=-\infty$. then $x \overline{\leqslant} y$

$x \in \mathbb{R}_{\mathbb{C}} \wedge y=-\infty$. then $y \overline{\leqslant} x$

$x=\infty \wedge y \in \mathbb{R}_{\mathbb{C}}$. then $y \overline{\leqslant} x$

$x=-\infty \wedge y \in \mathbb{R}_{\mathbb{C}}$. then $x \overline{\leqslant} y$

$x \in \mathbb{R}_{\mathbb{C}} \wedge y_{\mathbb{C}} \in \mathbb{R}$. then either $x \leqslant y \Rightarrow x \overline{\leqslant} y$ or $y \leqslant x \Rightarrow y \overline{\leqslant} x$

so in all possible cases we have either $x \overline{\leqslant} y$ or $y \overline{\leqslant} x$ □

Notation 10.101. *From now on, to avoid excessive nottion, we use* $\leqslant$ *to note* $\overline{\leqslant}$.

10.5.5 Conventions

Now we are finished with the tower of different types of numbers. From now on for the rest of this book we work only with $\mathbb{N}_{0,\mathbb{C}}$, $\mathbb{Z}_{\mathbb{C}}$, $\mathbb{Q}_{\mathbb{C}}$, $\mathbb{R}_{\mathbb{C}}$ and $\mathbb{C}$. To avoid excesive use of subscripts we leave out the subscripts. Hence for the rest of this book we use the following conventions:

Name of set	Symbol	Meaning
Natural Number	$\mathbb{N}_0$	$\mathbb{N}_{0,\mathbb{C}}$
Positive Natural Numbers	$\mathbb{N}$	$\mathbb{N}_{0,\mathbb{C}} \setminus \{0\}$
Integers	$\mathbb{Z}$	$\mathbb{Z}_{\mathbb{C}}$
Rational Numbers	$\mathbb{Q}$	$\mathbb{Q}_{\mathbb{C}}$
Real Numbers	$\mathbb{R}$	$\mathbb{R}_{\mathbb{C}}$
Extended Real Numbers	$\overline{\mathbb{R}}$	$\overline{\mathbb{R}}$
Non negative real numbers	$\mathbb{R}_0^+$	$\{x \in \mathbb{R} \mid 0 \leqslant x\}$
Positive real numbers	$\mathbb{R}^+$	$\{x \in \mathbb{R} \mid 0 < x\}$
Non positve numbers	$\mathbb{R}_0^-$	$\{x \in \mathbb{R} \mid x \leqslant 0\}$
Negative real numbers	$\mathbb{R}^-$	$\{x \in \mathbb{R} \mid x < 0\}$
Complex numbers	$\mathbb{C}$	$\mathbb{C}$

Using this notation we have that

$\mathbb{N}_0 \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}$
$\mathbb{N}_0 \subseteq \mathbb{R}_0^+ \subseteq \mathbb{R} \subseteq \mathbb{C}$
$\mathbb{N} \subseteq \mathbb{R}^+ \subseteq \mathbb{R}_0^+ \subseteq \mathbb{R} \subseteq \mathbb{C}$
$\mathbb{R}^- \subseteq \mathbb{R}_0^- \subseteq \mathbb{R} \subseteq \mathbb{C}$
$\forall z \in \mathbb{C}$ there exists unique $x.y \in \mathbb{R}$ such that $z = x + i \cdot y$

Chapter 11
Linear Algebra

11.1 Sums and products

11.1.1 Definition and properties

First we define the concept of a the sum of a finite family in a semi-group.

Definition 11.1. (Finite Sum) *Let $\langle A, +\rangle$ be a semi-group, $n \in \mathbb{N}_0$ and $\{x_i\}_{i \in \{0,\ldots,n\}} \subseteq A$ a finite family of elements of A then*

$$\sum_{i=0}^{(.)} x_i \colon \{0,\ldots,n\} \rightarrow A$$

is recursively defined [see definition: 10.41] by

$$\begin{array}{rcl} \displaystyle\sum_{i=0}^{0} x_i & = & x_0 \\ \forall i \in \{0,\ldots,n-1\} \displaystyle\sum_{i=0}^{i+1} x_i & = & \left(\displaystyle\sum_{i=0}^{i} x_i\right) + x_{i+1} \end{array}$$

Remark 11.2. We typical use the $\sum$ symbol in this text to represent a finite sum in a semi-group. However on the same set we can have many operators that makes the set a semi-group (for example in a ring, field or vector space we have a operator that represent addition and a operator that represent multiplication. To express these differences we use $\sum$ for addition and Π for multiplication. The following table gives a overview of the use of these symbols.

Structure	Operator	Symbol
additive semi-group	$+$	$\sum$
multiplicative semi-group	$\cdot$	$\prod$
ring	$+$	$\sum$
	$\cdot$	$\prod$
field	$+$	$\sum$
	$\cdot$	$\prod$
vector space	$+$	$\sum$
	$\cdot$	$\prod$
X^X	$\circ$	$\prod$

All the theorems, propositions, lemma's. corollaries and definitions in this section using the $\sum$ symbol also apply when we use the $\prod$ symbol. The meaning of $\sum$, $\prod$ symbol is always derived from the context where the operator is specified.

Example 11.3. Let $\{2^i\}_{i \in \{0,\ldots,3\}} \subseteq \mathbb{R}$ then for the field $\langle \mathbb{R}, +, \cdot\rangle$ $\sum_{i=0}^{3} 2^i = 15$

Proof.

$$\begin{aligned}\sum_{i=0}^{3} 2^i &= \left(\sum_{i=0}^{2} 2^i\right)+2^3\\ &= \left(\left(\sum_{i=0}^{1} 2^i\right)+2^2\right)+2^3\\ &= \left(\left(\left(\sum_{i=0}^{0} 2^i\right)+2^1\right)+2^2\right)+2^3\\ &= ((2^0+2^1)+2^2)+2^3\\ &= ((1+2)+4)+8\\ &= 15\end{aligned}$$

□

Theorem 11.4. *If $\langle A,+\rangle$ is a semi-group, $n\in\mathbb{N}_0$ and $\{x_i\}_{i\in\{0,\ldots,n\}}$ then we have $\forall k\in\{0,\ldots,n\}$ that*

$$\sum_{i=0}^{k} x_i = \begin{cases} x_0 \text{ if } k=0\\ (\sum_{i=0}^{k-1} x_i)+x_k \text{ if } k\in\{1,.,n\}\end{cases}$$

Proof. For $k\in\{0,\ldots,n\}$ we have either:

$k=0$. Then $\sum_{i=0}^{k} x_i=\sum_{i=0}^{0} x_i=x_0$

$k\in\{1,\ldots,n\}$. Then $l=k-1\in\{0,\ldots,n-1\}$ so that

$$\sum_{i=0}^{k} x_i=\sum_{i=0}^{l+1} x_i=\left(\sum_{i=0}^{l} x_i\right)+x_{l+1}=\left(\sum_{i=0}^{k-1} x_i\right)+x_k$$

□

Theorem 11.5. *Let $\langle A,+\rangle$ be a semi-group, $n\in\mathbb{N}_0$ and $\{x_i\}_{i\in\{0,\ldots,n\}}\subseteq A$ is such that $\forall i\in\{0,\ldots,n\}$ $x_i=0$ then*

$$\sum_{i=0}^{n} x_i=0$$

Proof. We prove this by mathematical induction so let

$$S=\left\{n\in\mathbb{N}_0|\forall\{x_i\}_{i\in\{0,\ldots,n\}}\subseteq A \text{ with } \forall i\in\{0,\ldots,n\}\ x_i=0 \text{ we have } \sum_{i=0}^{n} x_i=0\right\}$$

then we have:

$0\in S$. If $\{x_i\}_{i\in\{0,\ldots,0\}}\subseteq A$ with $\forall i\in\{0,\ldots,0\}=\{0\}$ $x_i=0$ we have that $\sum_{i=0}^{0} x_i=x_0=0$ proving that $0\in S$.

$n\in S\Rightarrow n+1\in S$. If $\{x_i\}_{i\in\{0,\ldots,n+1\}}\subseteq A$ with $\forall i\in\{0,\ldots,n+1\}$ $x_i=0$ then we have

$$\begin{aligned}\sum_{i=0}^{n+1} x_i &= \left(\sum_{i=0}^{n} x_i\right)+x_{n+1}\\ &= \left(\sum_{i=0}^{n} x_i\right)+0\\ &= \sum_{i=0}^{n} x_i\\ &\underset{n\in S}{=} 0\end{aligned}$$

proving that $n+1\in S$.

□

Theorem 11.6. *Let $\langle A,+\rangle$ be a **Abelian** semi-group, $n\in\mathbb{N}_0$ and $\{x_i\}_{i\in\{0,\ldots,n\}}\subseteq A$, $\{y_i\}_{i\in\{0,\ldots,n\}}\subseteq A$ 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 mathematical induction so let

$$S=\left\{n\in\mathbb{N}_0|\forall\{x_i\}_{i\in\{0,\ldots,n\}},\{y_i\}_{i\in\{0,\ldots,n\}}\subseteq A\text{ we have }\sum_{i=0}^{n}(x_i+y_i)=\sum_{i=0}^{n}x_i+\sum_{i=0}^{n}y_i\right\}$$

then we have:

$\boldsymbol{0\in S.}$ If $\{x_i\}_{i\in\{0\}},\{y_i\}_{i\in\{0\}}\subseteq A$ then $\sum_{i=0}^{0}(x_i+y_i)=x_0+y_0=\sum_{i=0}^{0}x_i+\sum_{i=0}^{0}y_i$ proving that $0\in S$.

$\boldsymbol{n\in S\Rightarrow n+1\in S.}$ If $\{x_i\}_{i\in\{0,\ldots,n+1\}},\{y_i\}_{i\in\{0,\ldots,n+1\}}\subseteq A$ then

$$\begin{aligned}\sum_{i=0}^{n+1}(x_i+y_i) &= \left(\sum_{i=0}^{n}(x_i+y_i)\right)+(x_{n+1}+y_{n+1})\\ &\underset{n\in S}{=} \left(\sum_{i=0}^{n}x_i+\sum_{i=0}^{n}y_i\right)+(x_{n+1}+y_{n+1})\\ &\underset{\text{associativity and commutativity}}{=} \left(\sum_{i=0}^{n}x_i+x_{n+1}\right)+\left(\sum_{i=0}^{n}y_i+y_{n+1}\right)\\ &= \sum_{i=0}^{n+1}x_i+\sum_{i=0}^{n+1}y_i\end{aligned}$$

proving that $n+1\in S$. □

Theorem 11.7. *Let $\langle A,+\rangle$ is a group, such that $\forall x\in A$ $-x$ is the inverse of x, $n\in\mathbb{N}_0$ and $\{x_i\}_{i\in\{0,\ldots,n\}}\subseteq A$ then we have*

$$\sum_{i=0}^{n}(-x_i)=-\sum_{i=0}^{n}x_i$$

Proof. We prove this by induction so let

$$S=\left\{n\in\mathbb{N}_0|\forall\{x_i\}_{i\in\{0,\ldots,n\}}\subseteq A\text{ we have }\sum_{i=0}^{n}(-x_i)=-\sum_{i=0}^{n}x_i\right\}$$

then we have:

$\boldsymbol{0\in S.}$ If $\{x_i\}_{i\in\{0\}}\subseteq A$ then $\sum_{i=0}^{0}(-x_i)=-x_0=-\sum_{i=0}^{0}x_i$ proving that $0\in S$.

$\boldsymbol{n\in S\Rightarrow n+1\in S.}$ If $\{x_i\}_{i\in\{0,\ldots,n+1\}}\subseteq A$ then

$$\begin{aligned}\sum_{i=0}^{n+1}(-x_i) &= \left(\sum_{i=0}^{n}(-x_i)\right)+(-x_{n+1})\\ &\underset{n\in S}{=} -\sum_{i=0}^{n}x_i+(-x_{n+1})\\ &= -\left(\left(\sum_{i=0}^{n}x_i\right)+x_{n+1}\right)\\ &= -\sum_{i=0}^{n+1}x_i\end{aligned}$$

hence we have $n+1\in S$ □

Theorem 11.8. *Let $\langle R,+,\cdot\rangle$ be a ring, $\alpha\in R$, $n\in\mathbb{N}_0$ and $\{x_i\}_{i\in\{0,\ldots,n\}}\subseteq A$ then we have*

$$\sum_{i=0}^{n}\alpha\cdot x_i=\alpha\cdot\sum_{i=0}^{n}x_i$$

Proof. As usually the prove is by induction. So let

$$S=\left\{n\in\mathbb{N}_0|\forall\{x_i\}_{i\in\{0,\ldots,n\}}\subseteq R \text{ we have } \sum_{i=0}^{n}(\alpha_i\cdot x_i)=\alpha\cdot\sum_{i=0}^{n}x_i\right\}$$

then we have:

$\boldsymbol{0\in S.}$ Let $\{x_i\}_{i\in\{0,\ldots,0\}}\subseteq R$ then $\sum_{i=0}^{0}(\alpha\cdot x_i)=\alpha\cdot x_0=\alpha\cdot\sum_{i=0}^{0}x_i$, proving that $0\in S$.

$\boldsymbol{n\in S\Rightarrow n+1\in S.}$ Let $\{x_i\}_{i\in\{0,\ldots,n+1\}}\subseteq A$ then

$$\begin{aligned}\sum_{i=0}^{n+1}(\alpha\cdot x_i) &= \left(\sum_{i=0}^{n}(\alpha\cdot x_i)\right)+\alpha\cdot x_{n+1}\\ &\underset{n\in S}{=} \alpha\cdot\sum_{i=0}^{n}x_i+\alpha\cdot x_{n+1}\\ &= \alpha\cdot\left(\left(\sum_{i=0}^{n}x_i\right)+x_{n+1}\right)\\ &= \alpha\cdot\sum_{i=0}^{n+1}x_i\end{aligned}$$

proving that $n+1\in S$. □

Theorem 11.9. *Let $\langle A,+\rangle$ be a* ***Abelian*** *semi-group, $n_1,n_2\in\mathbb{N}_0$ and*

$$\{x_{i,j}\}_{(i,j)\in\{0,\ldots,n_1\}\times\{0,\ldots,n_2\}}\subseteq A$$

then

$$\sum_{i=0}^{n_1}\left(\sum_{j=0}^{n_2}x_{i,j}\right)=\sum_{j=0}^{n_2}\left(\sum_{i=0}^{n_1}x_{i,j}\right)$$

See [definition: 2.108] for the definition of $\{x_{i,k}\}_{k\in\{0,\ldots,n_2\}}$ and $\{x_{k,j}\}_{k\in\{0,\ldots,n_1\}}$ used in $\sum_{j=0}^{n_2}x_{i,j}$ and $\sum_{i=0}^{n_1}x_{i,j}$.

Proof. We prove this by induction so let $n\in\{0,\ldots,n_1\}$ and take

$$S_n=\left\{m\in\{0,\ldots,n_2\}|\forall\{x_{i,j}\}_{(i,j)\in\{0,\ldots,n\}\times\{0,\ldots,m\}}\subseteq A \text{ then } \sum_{i=0}^{n}\left(\sum_{j=0}^{m}x_i\right)=\sum_{j=0}^{m}\left(\sum_{i=0}^{n}x_{i,j}\right)\right\}$$

then we have:

$\boldsymbol{0\in S_n.}$ Take $\{x_{i,j}\}_{(i,j)\in\{0,\ldots,n\}\times\{0\}}$ then

$$\sum_{i=0}^{n}\left(\sum_{j=0}^{0}x_{i,j}\right)=\sum_{i=0}^{n}x_{i,0}=\sum_{j=0}^{0}\left(\sum_{i=0}^{n}x_{i,j}\right)$$

proving that $0\in S_n$.

$\boldsymbol{m\in S_n\Rightarrow m+1\in S_n.}$ Take $\{x_{i,j}\}_{(i,j)\in\{0,\ldots,n\}\times\{0,\ldots,m+1\}}$ then we have

$$\begin{aligned}\sum_{i=0}^{n}\left(\sum_{j=0}^{m+1}x_{i,j}\right) &= \sum_{i=0}^{n}\left(\left(\sum_{j=0}^{m}x_{i,j}\right)+x_{i,m+1}\right)\\ &\underset{[\text{theorem: }11.6]}{=} \sum_{i=0}^{n}\left(\sum_{j=0}^{m}x_{i,j}\right)+\sum_{i=0}^{n}x_{i,m+1}\\ &\underset{m\in S}{=} \sum_{j=0}^{m}\left(\sum_{i=0}^{n}x_{i,j}\right)+\sum_{i=0}^{n}x_{i,m+1}\\ &= \sum_{j=0}^{m+1}\left(\sum_{i=0}^{n}x_{i,j}\right)\end{aligned}$$

proving that $m+1\in S_n$.

Using mathematical induction we have then that $S_n=\mathbb{N}_0$. So if $n_1,n_2\in\mathbb{N}_0$ and $\{x_{i.j}\}_{(i,j)\in\{0,\ldots,n_1\}\times\{0,\ldots,n_2\}}\subseteq A$ then $n_1\in\mathbb{N}_0$ and $n_2\in\mathbb{N}_0=S_{n_1}$ so that

$$\sum_{i=0}^{n_1}\left(\sum_{j=0}^{n_2} x_{i,j}\right)=\sum_{j=0}^{n_2}\left(\sum_{i=0}^{n_1} x_{i,j}\right)$$

□

Theorem 11.10. *Let $\langle A,+\rangle$ be a Abelian group, $n\in\mathbb{N}$ and $\{x_i\}_{i\in\{0,\ldots,n\}}\subseteq A$ then for $\{x_{i+1}-x_i\}_{i\in\{0,\ldots,n-1\}}\subseteq A$ we have*

$$\sum_{i=0}^{n-1}(x_{i+1}-x_i)=x_n-x_0$$

Proof. We prove this by induction, so let

$$S=\left\{n\in\mathbb{N}|\text{If }\{x_i\}_{i\in\{0,\ldots,n\}}\subseteq A\text{ then }\sum_{i=0}^{n-1}(x_{i+1}-x_i)=x_n-x_0\right\}$$

then we have:

$1\in S$. If $\{x_i\}_{i\in\{0,1\}}\subseteq A$ then $\sum_{i=0}^{1-1}x_i=\sum_{i=0}^{0}(x_{i+1}-x_i)=x_1-x_0$ proving that $1\in S$.

$n\in S\Rightarrow n+1\in S$. If $\{x_i)_{i\in\{0,\ldots,n+1\}}\subseteq A$ then

$$\begin{aligned}\sum_{i=0}^{(n+1)-1}(x_{i+1}-x_i) &= \sum_{i=0}^{n}(x_{i+1}-x_i)\\ &= \left(\sum_{i=0}^{n-1}(x_{i+1}-x_i)\right)+(x_{n+1}-x_n)\\ &\underset{n\in S}{=} (x_n-x_0)+(x_{n+1}-x_n)\\ &\underset{\text{associativity and commutativity}}{=} x_{n+1}-x_0\end{aligned}$$

proving that $n+1\in S$. □

Definition 11.11. *Let A be a set, $n,m\in\mathbb{N}_0$ with $n\leqslant m$ and $\{x_i\}_{i\in\{n,\ldots,m\}}\subseteq A$ a family then then the family $\{x_{i+n}\}_{i\in\{0,\ldots,n\}}$ is defined as $x\circ\beta$ where $\beta:\{0,\ldots,m-n\}\to\{n,\ldots m\}$ is the function defined by*

$$\beta(i)=n+i$$

[see definition: 2.109]

Up to now we have only defined the finite sum of a family of elements indexed by $\{0,\ldots,n\}$, we extend now this definition to a more general index set.

Definition 11.12. *If $\langle A,+\rangle$ is a semi-group, $n,m\in\mathbb{N}_0$ with $n\leqslant m$ and $\{x_i\}_{i\in\{n,\ldots,m\}}\subseteq A$ then we define $\sum_{i=n}^{m}x_i$ by*

$$\sum_{i=n}^{m}x_i=\sum_{i=0}^{m-n}x_{n+i}$$

***Note* 11.13.** *If $n=0$ then $m-n=m$ and $\{x_i\}_{i\in\{n,\ldots,m\}}=\{x_{0+i}\}_{i\in\{0,\ldots,m\}}=\{x_i\}_{i\in\{0,\ldots,m\}}$ and $\sum_{i=n}^{m}x_i=\sum_{i=0}^{m}x_{0+i}=\sum_{i=0}^{m}x_i$ as expected.*

Theorem 11.14. *If $\langle A,+\rangle$ is a semi-group, $n,m\in\mathbb{N}_0$ with $n\leqslant m$ and $\{x_i\}_{i\in\{n,\ldots,m\}}\subseteq A$ then for $k\in\{0,\ldots,n\}$ we have*

$$\sum_{i=n}^{m}x_i = \sum_{i=k}^{m+k-n}x_{n-k+i}$$

Proof. We have as $k\in\{0,\ldots,n\}\subseteq\mathbb{N}_0$ that $k\leqslant n\leqslant m$ that

$$\begin{aligned}\sum_{i=k}^{m+k-n} x_{(n-k)+i} &\underset{\text{def}}{=} \sum_{i=0}^{(m+k-n)-k} x_{(n-k)+i+k}\\ &= \sum_{i=0}^{m-n} x_{n+i}\\ &= \sum_{i=n}^{m} x_i\end{aligned}$$

$\square$

Using the above definition we can rephrase [theorems: 11.4, 11.5, 11.6, 11.9, 11.7, 11.10]

Theorem 11.15. *If $\langle A,+\rangle$ is a semi-group, $n,m\in\mathbb{N}_0$ with $n\leqslant m$ and $\{x_i\}_{i\in\{n,\ldots,m\}}\subseteq A$ then we have for $k\in\{n,\ldots,m\}$ that*

$$\sum_{i=n}^{k} x_i=\begin{cases} x_n \text{ if } k=n\\ \left(\sum_{i=n}^{k-1} x_i\right)+x^k \text{ if } k\in\{n+1,\ldots,m\}\end{cases}$$

Proof. Let $k\in\{n,\ldots,m\}$ then we have either:

***k = n*.** Then $\sum_{i=n}^{k} x_i=\sum_{i=n}^{n} x_i=\sum_{i=0}^{n-n} x_{n+i}=\sum_{i=0}^{0} x_{n+i}=x_{n+0}=x_n$

***k* ∈ {*n* + 1, …, *m*}.** Then as $n<m$ we have $m-n\in\mathbb{N}$ and for $l\in\{1,\ldots,m-n\}$

$$\sum_{i=0}^{l} x_{n+i} \underset{[\text{theorem: } 11.4]}{=} \left(\sum_{i=0}^{l-1} x_{n+i}\right)+x_{n+l} \tag{11.1}$$

Further

$$\begin{aligned}\sum_{i=n}^{k} x_i &= \sum_{i=0}^{k-n} x_{n+i}\\ &\underset{k-n\in\{1,\ldots,m-n\}\text{ and [eq: 11.1]}}{=} \left(\sum_{i=0}^{(k-n)-1} x_{n+i}\right)+x_{n+(k-n)}\\ &= \left(\sum_{i=n}^{k-1} x_i\right)+x_k\end{aligned}$$

$\square$

Theorem 11.16. *If $\langle A,+\rangle$ is a semi-group, $n,m\in\mathbb{N}_0$ with $n\leqslant m$ and $\{x_i\}_{i\in\{n,\ldots,m\}}\subseteq A$ is such that $\forall i\in\{n,\ldots,m\}$ $x_i=0$ then*

$$\sum_{i=n}^{m} x_i=0$$

Proof. Note that $\forall i\in\{0,\ldots,m-n\}$ $n+i\in\{n,\ldots,m\}$, so that $x_{n+i}=0$, hence

$$\sum_{i=n}^{m} x_i \underset{\text{def}}{=} \sum_{i=0}^{m-n} x_{n+i} \underset{[\text{theorem: } 11.5]}{=} 0$$ $\square$

Theorem 11.17. *Let $\langle A,+\rangle$ be a Abelian semi-group, $n,m\in\mathbb{N}_0$ with $n\leqslant m$ and $\{x_i\}_{i\in\{n,\ldots,m\}},\{y_i\}_{i\in\{n,\ldots,m\}}\subseteq A$ then*

$$\sum_{i=n}^{m}(x_i+y_i)=\sum_{i=n}^{m} x_i+\sum_{i=n}^{m} y_i$$

Proof. We have

$$\sum_{i=n}^{m}(x_i+y_i)=\sum_{i=0}^{m-n}(x_{n+i}+y_{n+i}) \underset{[\text{theorem: } 11.6]}{=} \sum_{i=0}^{m-n} x_i+\sum_{i=0}^{m-n} y_i=\sum_{i=n}^{m} x_i+\sum y_{i\,i=n}^{\,m}$$ $\square$

Theorem 11.18. *Let $\langle R,+,\cdot\rangle$ be a ring, $\alpha\in R$, $n,m\in\mathbb{N}_0$ with $n\leqslant m$ and $\{x_i\}_{i\in\{n,\ldots,m\}}\subseteq A$ then we have*

$$\sum_{i=n}^{m}\alpha\cdot x_i=\alpha\cdot\sum_{i=n}^{m}x_i$$

Proof.

$$\sum_{i=n}^{m}\alpha\cdot x_i=\sum_{i=0}^{m-n}\alpha\cdot x_{n+i}\underset{[\text{theorem: }11.8]}{=}\alpha\cdot\sum_{i=0}^{m-n}x_{n+i}=\alpha\cdot\sum_{i=n}^{m}x_i$$ □

Theorem 11.19. *Let $\langle A,+\rangle$ be a Abelian semi-group, $n_1,n_2,m_1,m_2\in\mathbb{N}_0$ with $n_1\leqslant m_1$, $n_2\leqslant m_2$ and $\{x_{i,j}\}_{(i,j)\in\{n_1,\ldots,m_1\}\times\{n_2,\ldots,m_2\}}\subseteq A$ then*

$$\sum_{i=n_1}^{m_1}\left(\sum_{j=n_2}^{m_2}x_{i,j}\right)=\sum_{j=n_2}^{m_2}\left(\sum_{i=n_1}^{m_1}x_{i,j}\right)$$

See [definition: 2.108] for $\{x_{i,k}\}_{k\in\{n_2,\ldots,m_2\}}$ and $\{x_{k,j}\}_{k\in\{n_1,\ldots,m_1\}}$

Proof. We have

$$\begin{aligned}
\sum_{i=n_1}^{m_1}\left(\sum_{j=n_2}^{m_2}x_{i,j}\right) &= \sum_{i=0}^{m_1-n_1}\left(\sum_{j=n_2}^{m_2}x_{n_1+i,j}\right)\\
&= \sum_{i=0}^{m_1-n_1}\left(\sum_{j=0}^{m_2-n_2}x_{n_1+i,m_2+j}\right)\\
&\underset{[\text{theorem: }11.9]}{=} \sum_{j=0}^{m_2-n_2}\left(\sum_{i=0}^{m_1-n_1}x_{n_1+i,m_2+j}\right)\\
&= \sum_{j=0}^{m_2-n_2}\left(\sum_{i=n_1}^{m_1}x_{i,m_2+j}\right)\\
&= \sum_{j=n_2}^{m_2}\left(\sum_{i=n_1}^{m_1}x_{i,j}\right)
\end{aligned}$$

□

Theorem 11.20. *Let $\langle A,+\rangle$ is a group, $n,m\in\mathbb{N}_0$ with $n\leqslant m$ and $\{x_i\}_{i\in\{n,\ldots,m\}}\subseteq A$ then we have*

$$\sum_{i=n}^{m}(-x_i)=-\sum_{i=n}^{m}x_i$$

Proof. We have

$$\sum_{i=n}^{m}(-x_i)=\sum_{i=0}^{m-n}(-x_{n+i})\underset{[\text{theorem: }11.7]}{=}-\sum_{i=0}^{m-n}x_i=-\sum_{i=n}^{m}x_i$$ □

Theorem 11.21. *Let $\langle A,+\rangle$ be a Abelian group, $n,m\in\mathbb{N}_0$ with $n<m$ and $\{x_i\}_{i\in\{n,\ldots,m\}}\subseteq A$ then for $\{x_{i+1}-x_i\}_{i\in\{n,\ldots,m-n-1\}}$ we have*

$$\sum_{i=n}^{m-1}(x_{i+1}-x_i)=x_m-x_n$$

Proof. As $n<m$ we have that $m-n\in\mathbb{N}$, so using [theorem: 11.10] we have

$$\sum_{i=0}^{m-n-1}(x_{(n+i)+1}-x_{n+i})=x_{n+(m-n)}-x_{n+0}=x_m-x_n$$

so that

$$\sum_{i=n}^{m-1} (x_{i+1}-x_i) = \sum_{i=0}^{m-n-1} (x_{(n+i)+1}-x_{n+i}) = x_m - x_n$$ □

Theorem 11.22. *If $\langle A,+\rangle$ is a semi-group, $n\in\mathbb{N}$ and $\{x_i\}_{i\in\{0,\ldots,n\}}$ then we have*

$$\sum_{i=0}^{n} x_i = x_0 + \sum_{i=1}^{n} x_i$$

Proof. We prove this by induction so let

$$S = \left\{ n\in\mathbb{N} | \text{if } \{x_i\}_{i\in\{0,\ldots,n\}}\subseteq A \text{ then } \sum_{i=0}^{n} x_i = x_0 + \sum_{i=1}^{n} x_i \right\}$$

then we have:

$1\in S$. If $\{x_i\}_{i\in\{0,\ldots,1\}}$ then $\sum_{i=0}^{1} x_i = (\sum_{i=0}^{0} x_i) + x_1 = x_0 + x_1 \underset{[\text{theorem: } 11.15]}{=} x_0 + \sum_{i=1}^{1} x_i$ proving that $1\in S$.

$n\in S \Rightarrow n+1\in S$. Let $\{x_i\}_{i\in\{0,\ldots,n+1\}}\subseteq A$ then we have

$$\begin{aligned}
\sum_{i=0}^{n+1} x_i &= \left(\sum_{i=0}^{n} x_i\right) + x_{n+1} \\
&\underset{n\in S}{=} \left(x_0 + \sum_{i=1}^{n} x_i\right) + x_{n+1} \\
&\underset{\text{associativity}}{=} x_0 + \left(\left(\sum_{i=1}^{n} x_i\right) + x_{n+1}\right) \\
&\underset{[\text{theorem: } 11.15]}{=} x_0 + \sum_{i=1}^{n+1} x_i
\end{aligned}$$

proving that $n+1\in S$ □

11.1.2 Associativity

Theorem 11.23. *Let $\langle A,+\rangle$ be a semi-group $n,m\in\mathbb{N}_0$, with $n<m$, $\{x_i\}_{i\in\{n,\ldots m\}}\subseteq A$ and $k\in\{n,\ldots,m-1\}$ then*

$$\sum_{i=n}^{m} x_i = \sum_{i=n}^{k} x_i + \sum_{i=k+1}^{m} x_i$$

or taking $k=n$ that

$$\sum_{i=n}^{m} x_i = \sum_{i=n}^{n} x_i + \sum_{i=n+1}^{m} x_i = x_n + \sum_{i=n+1}^{m} x_i$$

Proof. Let $n\in\mathbb{N}$ and define

$$S = \left\{ n\in\mathbb{N} | \forall\{x_i\}_{i\in\{0,\ldots,n\}}\subseteq A \text{ we have } \forall k\in\{0,\ldots,n-1\} \sum_{i=0}^{n} x_i = \sum_{i=0}^{k} x_i + \sum_{i=k+1}^{n} x_i \right\}$$

then we have:

$1\in S$. Let $\{x_i\}_{i\in\{0,\ldots,1\}=\{0,1\}}\subseteq A$ and let $k\in\{0,\ldots,1-1\}=\{0\}$ then

$$\sum_{i=0}^{1} x_i = \left(\sum_{i=0}^{0} x_i\right) + x_1 = \left(\sum_{0}^{k} x_i\right) + x_1 = \sum_{i=0}^{k} x_i + \sum_{i=0}^{0} x_{1+i} = \sum_{i=0}^{k} x_i + \sum_{i=1}^{1} x_i$$

proving that $1 \in S$.

$n \in S \Rightarrow n+1 \in S$. Let $\{x_i\}_{i \in \{0,\ldots,n+1\}} \subseteq A$ and take $k \in \{0,\ldots,(n+1)-1\} = \{0,\ldots,n\}$ then we have for k either:

$k = n$. Then

$$\begin{aligned}\sum_{i=0}^{n+1} x_i &= \left(\sum_{i=0}^{n} x_i\right) + x_{n+1} \\ &= \sum_{i=0}^{n} x_i + \sum_{i=0}^{0} x_{(n+1)+i} \\ &= \sum_{i=0}^{n} x_i + \sum_{i=n+1}^{n+1} x_i \\ &\underset{k=n}{=} \sum_{i=0}^{k} x_i + \sum_{i=k+1}^{n+1} x_i\end{aligned}$$

proving that $n+1 \in S$

$k \in \{0,\ldots,n-1\}$. Then we have

$$\begin{aligned}\sum_{i=0}^{n+1} x_i &= \left(\sum_{i=0}^{n} x_i\right) + x_{n+1} \\ &\underset{n \in S \wedge k \in \{0,\ldots,n-1\}}{=} \left(\sum_{i=0}^{k} x_i + \sum_{i=k+1}^{n} x_i\right) + x_{n+1} \\ &= \sum_{i=0}^{k} x_i + \left(\left(\sum_{i=k+1}^{n} x_i\right) + x_{n+1}\right) \\ &= \sum_{i=0}^{k} x_i + \left(\left(\sum_{i=0}^{n-(k+1)} x_{(k+1)+i}\right) + x_{(k+1)+(n-(k+1))+1}\right) \\ &= \sum_{i=0}^{k} x_i + \sum_{i=0}^{(n-(k+1))+1} x_{(k+1)+i} \\ &= \sum_{i=0}^{k} x_i + \sum_{i=k+1}^{n+1} x_i\end{aligned}$$

proving that $n+1 \in S$.

So by mathematical induction we have

$$\forall n \in \mathbb{N},\ \{x_i\}_{i \in \{0,\ldots,n\}} \subseteq A \text{ we have } \forall k \in \{0,\ldots,n-1\} \text{ that } \sum_{i=0}^{n} x_i = \sum_{i=0}^{k} x_i + \sum_{i=k+1}^{n} x_i \tag{11.2}$$

Take now $n, m \in \mathbb{N}_0$ with $n < m \Rightarrow m - n \in \mathbb{N}$ then for $\{x_i\}_{i \in \{n,\ldots,m\}}$ and $k \in \{n,\ldots,m-1\} \Rightarrow k - n \in \{0,\ldots,m-n\}$ we have

$$\begin{aligned}\sum_{i=n}^{m} x_i &= \sum_{i=0}^{m-n} x_{n+i} \\ &\underset{k-n \in \{0,\ldots,m-n\} \text{and [eq: 11.2]}}{=} \sum_{i=0}^{k-n} x_{n+i} + \sum_{i=(k-n)+1}^{m-n} x_{n+i} \\ &= \sum_{i=n}^{k} x_i + \sum_{i=k+1}^{m} x_i\end{aligned}$$

proving the theorem. □

Theorem 11.24. (Associativity) *Let $\langle A,+\rangle$ be a semi-group, $n\in\mathbb{N}$, let $\{(b_i,e_i)\}_{i\in\{0,\ldots,n\}}\subseteq\mathbb{N}_0\times\mathbb{N}_0$ so that*

1. $\forall i\in\{0,\ldots,n\}\ b_i\leqslant e_i$
2. $\forall i\in\{0,\ldots,n-1\}\ e_i+1=b_{i+1}$

then for $\{x_i\}_{i\in\{b_0,\ldots,e_n\}}\subseteq A$ we have

$$\sum_{i=b_0}^{e_n} x_i=\sum_{i=0}^{n}\left(\sum_{j=b_i}^{e_i} x_j\right)$$

Proof. We prove this by induction so let

$$S=\left\{n\in\mathbb{N}|\forall\{(b_i,e_i)\}_{i\in\{0,\ldots,n\}}\text{ satisfying (1),(2)},\forall\{x_i\}_{i\in\{b_0,\ldots,e_n\}}\subset A\vDash\sum_{i=b_0}^{e_n} x_i=\sum_{i=0}^{n}\left(\sum_{j=b_i}^{e_i} x_j\right)\right\}$$

then we have:

$1\in S$. Then for $\{(b_i,e_i)\}_{i\in\{0,1\}}$ we have $b_0\leqslant e_0\wedge b_1\leqslant e_1\wedge e_0+1=b_1\Rightarrow e_0\in\{b_0,\ldots,e_1-1\}$, for $\{x_i\}_{i\in\{b_0,\ldots,e_1\}}\subseteq A$ we have then

$$\begin{aligned}\sum_{i=b_0}^{e_1} x_i &\underset{\text{[theorem: 11.23]}}{=} \sum_{j=b_0}^{e_0} x_j+\sum_{j=e_0+1}^{e_1} x_j\\ &= \sum_{j=b_0}^{e_0} x_j+\sum_{j=b_1}^{e_1} x_j\\ &= \sum_{i=0}^{0}\left(\sum_{j=b_i}^{e_i} x_j\right)+\sum_{j=b_1}^{e_1} x_j\\ &= \sum_{i=0}^{1}\left(\sum_{j=b_i}^{e_i} x_j\right)\end{aligned}$$

proving that $1\in S$

$n\in S\Rightarrow n+1\in S$. Let $\{(b_i,e_i)\}_{i\in\{0,\ldots,n+1\}}\subseteq\mathbb{N}_0\times\mathbb{N}_0$ with $\forall i\in\{1,\ldots,n+1\}\ b_i\leqslant e_i$, $\forall i\in\{0,\ldots,(n+1)-1\}=\{0,\ldots,n\}\ e_i+1=b_{i+1}$ and $\{x_i\}_{i\in\{b_0,\ldots,e_{n+1}\}}\subseteq A$. In particular we have $b_n\leqslant e_n=b_{n+1}-1\leqslant e_{n+1}-1$. Further $\forall i\in\{0,\ldots,n\}$ we have $b_i\leqslant e_i<e_i+1=b_{i+1}$ proving by [theorem: 10.36] that $b_0\leqslant b_n\leqslant e_n$ so that $e_n\in\{b_0,\ldots,e_{n+1}-1\}$. Next

$$\begin{aligned}\sum_{j=b_0}^{e_{n+1}} x_j &\underset{e_n\in\{b_0,\ldots,e_{n+1}-1\}\ \text{[theorem: 11.23]}}{=} \sum_{j=b_0}^{e_n} x_j+\sum_{j=e_n+1}^{e_{n+1}} x_j\\ &\underset{e_n+1=b_{n+1}}{=} \sum_{j=b_0}^{e_n} x_j+\sum_{j=b_{n+1}}^{e_{n+1}} x_j\\ &\underset{n\in S}{=} \sum_{i=0}^{n}\left(\sum_{j=b_i}^{e_i} x_j\right)+\sum_{j=b_{n+1}}^{e_{n+1}} x_j\\ &= \sum_{i=0}^{n+1}\left(\sum_{j=b_i}^{e_i} x_j\right)\end{aligned}$$

proving that $n+1\in S$. □

11.1.3 Commutativity

We will now generalize commutativity to finite sum. First we must introduce the concept of permutations.

Definition 11.25. (Permutation) *If I is a set then a bijection $\sigma: I \rightarrow I$ is called a* ***permutation of I****. The set of all the permutation graphs of I is noted as S_I hence*

$$S_I = \{\sigma \in I^I | \sigma: I \rightarrow I \text{ is a bijection}\}$$

Theorem 11.26. Let *I be a set then $\langle S_I, \circ \rangle$ is a group with neutral element Id_I and $\forall \sigma \in S_I$ σ^{-1} as inverse element.*

Proof. As the composition of two bijections is a bijection [see theorem: 2.75] we have that

$$\circ: S_I \times S_I \rightarrow S_I \text{ is a operator}$$

Further we have:

associativity. If $\sigma, \beta, \gamma \in \mathcal{S}_I$ then $(\sigma \circ \beta) \circ \gamma \underset{\text{[function: 2.21]}}{=} \sigma \circ (\beta \circ \gamma)$

neutral element. As $\mathrm{Id}_I: I \rightarrow I$ is a bijection [see: 2.64] we have that $\mathrm{Id}_I \in S_I$. Further by [theorem: 2.48] we have

$$\sigma = \sigma \circ \mathrm{Id}_i = \mathrm{Id}_I \circ \sigma$$

inverse element. If $\sigma \in S_I$ then $\sigma: I \rightarrow I$ is a bijection, hence by [theorem: 2.72] $\sigma^{-1}: I \rightarrow I$ is a bijection so that $\sigma^{-1} \in S_I$. Hence $\sigma \circ \sigma^{-1} \underset{\text{[theorem: 2.69]}}{=} \mathrm{Id}_I \underset{\text{[theorem: 2.69]}}{=} \sigma^{-1} \circ \sigma$ □

Theorem 11.27. *Let I be a set and $\sigma \in S_I$ then if $i \in I$ with $\sigma(i) = i$ then $\sigma_{|I \setminus \{i\}} \in S_{I \setminus \{i\}}$*

Proof. As $\sigma: I \rightarrow I$ is a bijection we have by [theorem: 2.89] that

$$\sigma_{|I \setminus \{i\}}: I \setminus \{i\} \rightarrow I \setminus \{\sigma(i)\} \underset{\sigma(i) = i}{=} I \setminus \{i\}$$

is a bijection. □

We define now a special type of permutation a transposition

Theorem 11.28. (transposition) *Let I be a set, $i, j \in I$ then for $\left(i \underset{I}{\leftrightarrow} j\right)$ defined by*

$$\left(i \underset{I}{\leftrightarrow} j\right)(k) = \begin{cases} k \text{ if } k \in I \setminus \{i, j\} \\ i \text{ if } k = j \\ j \text{ if } k = i \end{cases}$$

we have that

$$\left(i \underset{I}{\leftrightarrow} j\right) \in S_I$$

The permutation $\left(i \underset{I}{\longleftrightarrow} j\right)$ is called a ***transposition*** *of i and j.*

Proof. We have

injectivity. If $\left(i \underset{I}{\leftrightarrow} j\right)(k) = \left(i \underset{I}{\leftrightarrow} j\right)(l)$ then for k, l we have either:

$k = i \wedge l = i$. Then trivially $k = i$

$k = i \wedge l = j$. Then $l = j = \left(i \underset{I}{\leftrightarrow} j\right)(k) = \left(i \underset{I}{\leftrightarrow} j\right)(l) = i = k$ so that $k = l$

$k = i \wedge l \neq i, j$. Then $j = \left(i \underset{I}{\leftrightarrow} j\right)(k) = \left(i \underset{I}{\leftrightarrow} j\right)(l) = l \neq i, j$ a contradiction so this case never occurs.

$k = j \wedge l = i$. Then $l = i = \left(i \underset{I}{\leftrightarrow} j\right)(k) = \left(i \underset{I}{\leftrightarrow} j\right)(l) = j = k$ so that $k = l$.

$k = j \wedge l = j$. Then trivially $k = l$.

$k = j \wedge l \neq i, j$. Then $i = \left(i \underset{I}{\leftrightarrow} j\right)(k) = \left(i \underset{I}{\leftrightarrow} j\right)(l) = l \neq i, j$ a contradiction so this case never occurs.

$k \neq i, j \wedge l = i$. Then $i, j \neq k = \left(i \underset{I}{\leftrightarrow} j\right)(k) = \left(i \underset{I}{\leftrightarrow} j\right)(l) = j$ a contradiction so this case never occurs.

$k \neq i, j \wedge l = j$. Then $i, j \neq k = \left(i \underset{I}{\leftrightarrow} j\right)(k) = \left(i \underset{I}{\leftrightarrow} j\right)(l) = i$ a contradiction so this case never occurs.

$k \neq i, j \wedge l \neq i, j$. Then $k = \left(i \underset{I}{\leftrightarrow} j\right)(k) = \left(i \underset{I}{\leftrightarrow} j\right)(l) = l$ so that $k = l$

Hence in all valid cases we have $k = l$ proving injectivity.

surjectivity. If $l \in I$ then we have either:

$l = i$. Then for $k = j$ we have $\left(i \underset{I}{\leftrightarrow} j\right)(k) = \left(i \underset{I}{\leftrightarrow} j\right)(j) = i = l$

$l = j$. Then for $k = i$ we have $\left(i \underset{I}{\leftrightarrow} j\right)(k) = \left(i \underset{I}{\leftrightarrow} j\right)(i) = j = l$

$l \neq i, j$. Then for $k = l$ we have $\left(i \underset{I}{\leftrightarrow} j\right)(k) = \left(i \underset{I}{\leftrightarrow} j\right)(l) = l$

so in all cases we found a $k \in I$ such that $\left(i \underset{I}{\leftrightarrow} j\right)(k) = l$. □

Theorem 11.29. *Let I be a set then we have*

1. *If* $\left(i \underset{I}{\leftrightarrow} i\right) = \mathrm{Id}_I(i)$
2. *If $i, j \in I$ then* $\left(i \underset{I}{\leftrightarrow} j\right) \circ \left(i \underset{I}{\leftrightarrow} j\right) = \mathrm{Id}_I$
3. $\left(i \underset{I}{\leftrightarrow} j\right) = \left(j \underset{n}{\leftrightarrow} i\right)$
4. *If $J \subseteq I$ and $i, j \in J$ then* $\left(i \underset{J}{\longleftrightarrow} j\right) = \left(i \underset{I}{\leftrightarrow} j\right)_{|J}$

Proof.

1. If $k \in I$ then we have either:

   $k = i$. Then $\left(i \underset{I}{\leftrightarrow} j\right)(k) = j \underset{i=j}{=} i = k = \mathrm{Id}_I(k)$

   $k = j$. Then $\left(i \underset{I}{\leftrightarrow} j\right)(k) = j \underset{i=j}{=} i = k = \mathrm{Id}_I(k)$

   $k \neq i, j$. Then $\left(i \underset{I}{\leftrightarrow} j\right)(k) = j \underset{i=j}{=} i = k = \mathrm{Id}_I(k)$

   proving that $\left(i \underset{I}{\leftrightarrow} j\right)$=Id_I

2. If $k \in I$ then we have either:

   $k = i$. Then $\left(i \underset{I}{\leftrightarrow} j\right)\left(\left(i \underset{I}{\longleftrightarrow} j\right)(k)\right) = \left(i \underset{I}{\leftrightarrow} j\right)(j) = i = k$

   $k = j$. Then $\left(i \underset{I}{\leftrightarrow} j\right)\left(\left(i \underset{I}{\longleftrightarrow} j\right)(k)\right) = \left(i \underset{I}{\leftrightarrow} j\right)(i) = j = k$

   $k \neq i, j$. Then $\left(i \underset{I}{\leftrightarrow} j\right)\left(\left(i \underset{I}{\longleftrightarrow} j\right)(k)\right) = k$

   proving that $\left(i \underset{I}{\leftrightarrow} j\right) \circ \left(i \underset{I}{\leftrightarrow} j\right) = \mathrm{Id}_I$.

3. If $k \in I$ then we have either:

   $k = i$. Then $\left(i \underset{I}{\leftrightarrow} j\right)(k) = j = (j \leftrightarrow i)(k)$

   $k = j$. Then $\left(i \underset{I}{\leftrightarrow} j\right)(k) = i = (j \leftrightarrow i)(k)$

   $k \neq i, j$. $\left(i \underset{I}{\leftrightarrow} j\right)(k) = k = (j \leftrightarrow i)(k)$

   proving that $\left(i \underset{I}{\leftrightarrow} j\right) = (j \leftrightarrow i)$.

4. If $k \in J$ then we have either:

   $k = i$. Then $\left(i \underset{J}{\leftrightarrow} j\right)(k) = j = \left(i \underset{I}{\leftrightarrow} j\right)(k)$

   $k = j$. Then $\left(i \underset{J}{\leftrightarrow} j\right)(k) = i = \left(i \underset{I}{\leftrightarrow} j\right)(k)$

$\boldsymbol{k \neq i, j.}$ Then $(i \underset{J}{\leftrightarrow} j)(k) = k = (i \underset{I}{\leftrightarrow} j)(k)$

proving that $\left(i \underset{J}{\longleftrightarrow} j\right) = (i \underset{I}{\leftrightarrow} j)_{|J}$. □

Lemma 11.30. *Let $n \in \mathbb{N}_0$ and $\sigma \in S_{\{0,\ldots,n+1\}}$ is such that $\sigma(n+1) \neq n+1$ then for $k = \sigma^{-1}(n+1)$ we have for γ_σ defined by*

$$\gamma_\sigma \colon \{0,\ldots,n\} \rightarrow \{0,\ldots,n\} \text{ defined by } \gamma_\sigma(i) = \begin{cases} \sigma\left(\left(n \underset{\{0,\ldots n\}}{\leftrightarrow} k\right)(i)\right) & \text{if } i \in \{0,\ldots,n-1\} \\ \sigma(n+1) & \text{if } i = n \end{cases}$$

that

$$\gamma_\sigma \in S_{\{0,\ldots,n\}}$$

Proof. First for $i \in [0,\ldots,n]$ we have either:

$\boldsymbol{i \in \{0,\ldots,n-1\}.}$ Assume that $\sigma\left(\left(n \underset{\{0,\ldots n\}}{\leftrightarrow} k\right)(i)\right) = n+1$ then as $\sigma(k) = n+1$ we have as σ is injective that $\left(n \underset{\{0,\ldots n\}}{\leftrightarrow} k\right)(i) = k = \left(n \underset{\{0,\ldots n\}}{\leftrightarrow} k\right)(n)$ so that $i = n$ contradicting $i \in \{0,\ldots,n-1\}$. So we must have that $\sigma\left(\left(n \underset{\{0,\ldots n\}}{\leftrightarrow} k\right)(i)\right) \neq n+1$ or $\gamma_\sigma(i) \neq n+1 \Rightarrow \gamma_\sigma(i) \in \{0,\ldots,n\}$.

$\boldsymbol{i = n.}$ Then as $\sigma(n+1) \neq n+1$ we have that $\sigma(n+1) \in \{0,\ldots,n\}$ or $\gamma_\sigma(i) \in \{0,\ldots,n\}$

So we have that $\forall i \in \{1,\ldots,n\}$ that $\gamma_\sigma(i) \in \{0,\ldots,n\}$ hence we have indeed a function

$$\gamma_\sigma \colon \{0,\ldots,n\} \rightarrow \{0,\ldots,n\}$$

Next we have to prove that it is a bijection.

injectivity. Let $r, s \in \{0,\ldots,n\}$ such that $\gamma_\sigma(r) = \gamma_\sigma(s)$ then we have for r, s either:

$\boldsymbol{r = n \wedge s = n.}$ Then trivially $r = s$.

$\boldsymbol{r = n \wedge s \neq n.}$ Then we have for s either:

$\boldsymbol{s = k.}$ Then $\sigma(n+1) = \gamma_\sigma(r) = \gamma_\sigma(s) = \sigma\left(\left(n \underset{\{0,\ldots n\}}{\leftrightarrow} k\right)(s)\right) = \sigma(n)$ leading to the contradiction $n+1 = n$, hence this case never occurs.

$\boldsymbol{s \neq k.}$ Then $\sigma(n+1) = \gamma_\sigma(r) = \gamma_\sigma(s) = \sigma\left(\left(n \underset{\{0,\ldots n\}}{\leftrightarrow} k\right)(s)\right) = \sigma(s)$ so that $s = n+1$ contradicting $s \in \{1,\ldots,n\}$, hence this case does not occur.

$\boldsymbol{r \neq n \wedge s = n.}$ Then we have for r either:

$\boldsymbol{r = k.}$ Then $\sigma(n+1) = \gamma_\sigma(s) = \gamma_\sigma(r) = \sigma\left(\left(n \underset{\{0,\ldots n\}}{\leftrightarrow} k\right)(r)\right) = \sigma(n)$ leading to the contradiction $n+1 = n$, so this case never occurs.

$\boldsymbol{r \neq k.}$ Then $\sigma(n+1) = \gamma_\sigma(s) = \gamma_\sigma(r) = \sigma\left(\left(n \underset{\{0,\ldots n\}}{\leftrightarrow} k\right)(r)\right) = \sigma(r)$ so that $n+1 = r$ contradicting $r \in \{1,\ldots,n\}$, hence this case does not occur.

$\boldsymbol{r \neq n \wedge s \neq n.}$ Then $\sigma\left(\left(n \underset{\{0,\ldots n\}}{\leftrightarrow} k\right)(r)\right) = \gamma_\sigma(r) = \gamma_\sigma(s) = \sigma\left(\left(n \underset{\{0,\ldots n\}}{\leftrightarrow} k\right)(s)\right)$ so that $\left(n \underset{\{0,\ldots n\}}{\leftrightarrow} k\right)(r) = \left(n \underset{\{0,\ldots n\}}{\leftrightarrow} k\right)(s)$ or $r = s$ [as σ and $\left(n \underset{\{0,\ldots n\}}{\leftrightarrow} k\right)$ are injections).

So in all valid cases we have $r = s$ proving injectivity.

surjectivity. Let $r \in \{0,\ldots,n\}$ then by the surjectivity of σ there exist a $s \in \{0,\ldots,n+1\}$ such that $\sigma(s) = r$. For s we have then either:

$\boldsymbol{s = n+1.}$ Then for $n \in \{0,\ldots,n\}$ we have $\gamma_\sigma(n) = \sigma(n+1) = \sigma(s) = r$ so that $\gamma_\sigma(n) = r$.

$\boldsymbol{s = n.}$ If $s = k$ then $r = \sigma(s) = \sigma(k) = n+1$ contradicting $r \in \{0,\ldots,n\}$. So we have that $n = s \neq k$, hence $k \in \{0,\ldots,n-1\}$, so that

$$\gamma_\sigma(k) = \sigma\left(\left(n \underset{\{0,\ldots n\}}{\leftrightarrow} k\right)(k)\right) = \sigma(n) = \sigma(s) = r$$

proving that $\gamma_\sigma(k) = r$.

$s\in\{0,\ldots,n-1\}$. If $s=k$ then $r=\sigma(s)=\sigma(k)=n+1$ so that $r=n+1$ contradicting $r\in\{0,\ldots,n\}$ so we must have that $s\neq k,n$. Hence

$$\gamma_\sigma(s)=\sigma\Big(\Big(n \underset{\{0,\ldots n\}}{\leftrightarrow} k\Big)(s)\Big)=\sigma(s)=r$$

So in all cases we found a $l\in\{0,\ldots,n\}$ such that $\gamma_\sigma(l)=r$ proving surjectivity. □

Theorem 11.31. (Commutativity) *Let $\langle A,+\rangle$ be a **Abelian** semi-group, $n\in\mathbb{N}_0$, $\{x_i\}_{i\in\{0,\ldots,n\}}\subseteq A$ and $\sigma\in S_{\{0,\ldots,n\}}$ a permutation then*

$$\sum_{i=0}^{n} x_i=\sum_{i=0}^{n} x_{\sigma(i)}$$

Proof. We prove this by induction so take

$$S=\left\{n\in\mathbb{N}_0|\text{If } \{x_i\}_{i\in\{0,\ldots,n\}}\subseteq A \text{ and } \sigma\in S_{\{0,\ldots,n\}} \text{ then } \sum_{i=0}^{n} x_i=\sum_{i=0}^{n} x_{\sigma(i)}\right\}$$

then we have:

$0\in S$. Let $\{x_i\}_{i\in\{0\}}\subseteq A$ and $\sigma\in S_{\{0\}}$ then $\sigma=\mathrm{Id}_{\{0\}}$ so that

$$\sum_{i=0}^{0} x_i=x_0=x_{\mathrm{Id}_{\{0\}}(0)}=x_{\sigma(0)}=\sum_{i=0}^{n} x_{\sigma(i)}$$

proving that $0\in S$.

$n\in S\Rightarrow n+1\in S$. Let $\{x_i\}_{i\in\{0,\ldots,n+1\}}$ and $\sigma\in S_{\{0,\ldots,n+1\}}$ then for $n+1$ we have either

$n+1=1$. Then $S_{\{0,\ldots,n+1\}}=S_{\{0,1\}}$ so that for $\sigma(1)$ we have either:

$\sigma(1)=0$. Then as $\sigma:\{0,1\}\to\{0,1\}$ is a bijection we must have $\sigma(0)=1$, so

$$\sum_{i=0}^{1} x_i=x_0+x_1=x_1+x_0=x_{\sigma(0)}+x_{\sigma(1)}=\sum_{i=0}^{1} x_{\sigma(i)}$$

$\sigma(1)=1$. Then as $\sigma:\{0,1\}\to\{0,1\}$ is a bijection we must have $\sigma(0)=0$, so

$$\sum_{i=0}^{1} x_i=x_0+x_1=x_{\sigma(0)}+x_{\sigma(1)}=\sum_{i=0}^{1} x_{\sigma(i)}$$

so in all cases $\sum_{i=0}^{1} x_i=\sum_{i=0}^{1} x_{\sigma(i)}$ proving that $n+1\in S$.

$1<n+1$. Now for $\sigma(n+1)$ we have either:

$\sigma(n+1)=n+1$. Then by [theorem: 11.27] we have that $\sigma_{|\{0,\ldots,n\}}\in S_{\{0,\ldots,n\}}$ which as $n\in S$ proves that

$$\sum_{i=0}^{n} x_i\underset{n\in S}{=}\sum_{i=0}^{n} x_{\sigma_{|\{0,\ldots,n\}}(i)}=\sum_{i=0}^{n} x_{\sigma(i)} \tag{11.3}$$

So

$$\begin{aligned}\sum_{i=0}^{n+1} x_i &= \left(\sum_{i=0}^{n} x_i\right)+x_{n+1}\\ &\underset{[\text{eq: } 11.3]}{=} \left(\sum_{i=0}^{n} x_{\sigma(i)}\right)+x_{n+1}\\ &= \left(\sum_{i=0}^{n} x_{\sigma(i)}\right)+x_{\sigma(n+1)}\\ &= \sum_{i=0}^{n+1} x_{\sigma(i)}\end{aligned}$$

proving that $n+1 \in S$.

$\boldsymbol{\sigma(n+1) \in \{0,\ldots,n\}}$. Take then $k = \sigma^{-1}(n+1) \in \{0,\ldots,n\}$ then we have by [lemma: 11.30] that $\gamma_\sigma \in S_{\{0,\ldots,n\}}$ where

$$\gamma_\sigma(i) = \begin{cases} \sigma\left(\left(n \underset{\{0,\ldots n\}}{\leftrightarrow} k\right)(i)\right) \text{ if } i \in \{0,\ldots,n-1\} \\ \sigma(n+1) \text{ if } i = n \end{cases} \tag{11.4}$$

so as $n \in S$ we have that:

$$\begin{aligned} \sum_{i=0}^{n} x_i &\underset{n \in S}{=} \sum_{i=0}^{n} x_{\gamma_\sigma(i)} \\ &= \left(\sum_{i=0}^{n-1} x_{\gamma_\sigma(i)}\right) + x_{\gamma_\sigma(n+1)} \\ &\underset{11.4}{=} \left(\sum_{i=0}^{n-1} x_{\sigma\left(\left(n \underset{\{0,\ldots,n\}}{\leftrightarrow} k\right)(i)\right)}\right) + x_{\sigma(n+1)} \end{aligned} \tag{11.5}$$

Further we have

$$\left(n \underset{\{0,\ldots,n\}}{\leftrightarrow} k\right) \in S_{\{0,\ldots,n\}} \tag{11.6}$$

we have

$$\begin{aligned} \sum_{i=}^{n+1} x_{\sigma(i)} &= \\ \left(\sum_{i=0}^{n} x_{\sigma(i)}\right) + x_{\sigma(n+1)} &\underset{n \in S \wedge [\text{eq: } 11.6]}{=} \\ \left(\sum_{i=0}^{n} x_{\sigma\left(\left(n \underset{\{0,\ldots,n\}}{\leftrightarrow}\right)(i)\right)}\right) + x_{\sigma(n+1)} &= \\ \left(\left(\sum_{i=0}^{n-1} x_{\sigma\left(\left(n \underset{\{0,\ldots,n\}}{\leftrightarrow}\right)(i)\right)}\right) + x_{\sigma\left(\left(n \underset{\{0,\ldots,n\}}{\leftrightarrow} k\right)(n)\right)}\right) + x_{\sigma(n+1)} &= \\ \left(\sum_{i=0}^{n-1} x_{\sigma\left(\left(n \underset{\{0,\ldots,n\}}{\leftrightarrow}\right)(i)\right)}\right) + \left(x_{\sigma\left(\left(n \underset{\{0,\ldots,n\}}{\leftrightarrow} k\right)(n)\right)} + x_{\sigma(n+1)}\right) &= \\ \left(\sum_{i=0}^{n-1} x_{\sigma\left(\left(n \underset{\{0,\ldots,n\}}{\leftrightarrow}\right)(i)\right)}\right) + \left(x_{\sigma(n+1)} + x_{\sigma\left(\left(n \underset{\{0,\ldots,n\}}{\leftrightarrow} k\right)(n)\right)}\right) &= \\ \left(\left(\sum_{i=0}^{n-1} x_{\sigma\left(\left(n \underset{\{0,\ldots,n\}}{\leftrightarrow}\right)(i)\right)}\right) + x_{\sigma(n+1)}\right) + x_{\sigma\left(\left(n \underset{\{0,\ldots,n\}}{\leftrightarrow} k\right)(n)\right)} &\underset{[\text{eq: } 11.5]}{=} \\ \left(\sum_{i=0}^{n} x_i\right) + x_{\sigma\left(\left(n \underset{\{0,\ldots,n\}}{\leftrightarrow} k\right)(n)\right)} &= \\ \left(\sum_{i=0}^{n} x_i\right) + x_{\sigma(k)} &\underset{k = \sigma^{-1}(n+1)}{=} \\ \left(\sum_{i=0}^{n} x_i\right) + x_{n+1} &= \\ \sum_{i=0}^{n+1} x_i &= \end{aligned}$$

proving that $n+1 \in S$.

proving that $n+1\in S$.

So in all cases we have $n+1\in S$.

Mathematical induction proves then that $S=\mathbb{N}_0$ and the theorem. □

11.1.4 Generalized sum

Next we define the sum of elements over a finite index set.

Definition 11.32. *Let $\langle A,+\rangle$ be a Abelian semi-group with neutral element 0, I a finite set and $\{x_i\}_{i\in I}$ then we define $\sum_{i\in I}x_i$ as follows:*

1. *If $I=\varnothing$ then $\sum_{i\in I}x_i=0$*
2. *If $I\neq\varnothing$ then as I is finite there exist by [corollary: 10.86] a **unique** $n\in\mathbb{N}$ and a bijection $\beta:\{0,\ldots,n-1\}\rightarrow I$, $\sum_{i\in I}x_i$ is then defined by*

$$\sum_{i\in I}x_i=\sum_{i=0}^{n-1}x_{\beta(i)}$$

***Note* 11.33.** *If $\langle A,\cdot\rangle$ is a multiplicative Abelian group [for example for a ring or field we want to make distinction between summation and multiplication] with neutral element 1, I a finite set and $\{x_i\}_{i\in I}$ then we use $\prod$ instead of $\sum$ and have then that $\prod_{i\in I}x_i$ ss defined by*

1. *If $I=\varnothing$ then $\prod_{i\in I}x_i=1$*
2. *If $I\neq\varnothing$ then I is finite there exist by [corollary: 10.86] a **unique** $n\in\mathbb{N}$ and a bijection $\beta:\{0,...,n-1\}\rightarrow I$, $\sum_{i\in I}x_i$ is then defined by*

$$\prod_{i\in I}x_i=\prod_{i=0}^{n-1}x_{\beta(i)}$$

Of course all staments about summation applies also to mulitplication if you replace $\sum$ by $\prod$ and 0 by 1.

Proof. We must for (2) prove that $\sum_{i\in I}x_i$ is unique. So assume that $\gamma:\{0,\ldots,n-1\}\rightarrow I$ is another bijection then $\beta^{-1}\circ\gamma:\{0,\ldots,n-1\}\rightarrow\{0,\ldots,n-1\}$ is a bijection, hence $\beta^{-1}\circ\gamma\in S_{\{0,\ldots,n-1\}}$ so that

$$\sum_{i=0}^{n-1}x_{\beta(i)}\underset{\text{[theorem: 11.31]}}{=}\sum_{i=0}^{n-1}x_{\beta(\beta^{-1}(\gamma(i)))}=\sum_{i=0}^{n-1}x_{\gamma(i)}$$ □

Example 11.34. Let $\langle A,+\rangle$ be a Abelian semi-group and $\{x_i\}_{i\in\{k\}}\subseteq A$ then $\sum_{i\in\{k\}}x_i=x_k$

Proof. As $\beta:\{0\}=\{0,\ldots,(1-1)\}\rightarrow\{k\}$ by $\beta(0)=k$ we have that $\sum_{i\in\{k\}}x_i=\sum_{i=0}^{0}x_{\beta(i)}=x_{\beta(0)}=x_k$ □

This new definition is equivalent with the previous definition as the following theorem shows.

Theorem 11.35. *Let $\langle A,+\rangle$ be a Abelian semi-group with neutral element 0, $n,m\in\mathbb{N}_0$ with $n\leqslant m$ and $\{x_i\}_{i\in\{n,\ldots,m\}}$ then $\sum_{i=n}^{m}x_i=\sum_{i\in\{n,\ldots,m\}}x_i$*

Proof. Define $\sigma:\{0,\ldots m-n\}\rightarrow\{n,\ldots,m\}$ by $\sigma(i)=n+i$ then we have

injectivity. If $\sigma(i)=\sigma(j)$ then $n+i=n+j$ hence $i=j$.

surjectivity. If $j\in\{n,\ldots,m\}\Rightarrow n\leqslant j\leqslant m\Rightarrow 0\leqslant j-n\leqslant m-n$ then for $i=j-n$ we have $j=i+n=\beta(i)$ and $0\leqslant i\leqslant m-n$

so that $\sigma:\{0,\ldots,m-n\}\rightarrow\{n,\ldots,m\}$ is a bijection and by definition

$$\sum_{i\in\{n,\ldots,m\}}x_i=\sum_{i=0}^{m-n}x_{\sigma(i)}=\sum_{i=0}^{m-n}x_{n+i}\underset{\text{def}}{=}\sum_{i=n}^{m}x_i$$

□

Theorem 11.36. *Let $\langle A, +\rangle$ be a Abelian semi-group with neutral element 0, I a finite set, $\{x_i\}_{i\in I}$ and $\sigma: J \rightarrow I$ a bijection then*

$$\sum_{i\in I} x_i = \sum_{j\in J} x_{\sigma(j)}$$

Proof. For I we have either:

$I = \varnothing$. Then as $\sigma: J \rightarrow I$ is a bijection we have $J = \varnothing$ so that $\sum_{i\in I} x_i = 0 = \sum_{j\in J} x_{\sigma(j)}$

$I \neq \varnothing$. Then there exists a $n \in \mathbb{N}$ such that $\beta: \{0, \ldots, n-1\} \rightarrow I$ such that

$$\sum_{i\in I} x_i = \sum_{i=0}^{n-1} x_{\beta(i)}$$

As $\sigma^{-1} \circ \beta: \{0, \ldots, n-1\} \rightarrow J$ is a bijection we have

$$\sum_{j\in J} x_{\sigma(j)} = \sum_{i=0}^{n-1} x_{\sigma(\sigma^{-1}(\beta(i)))} = \sum_{i=0}^{n-1} x_{\beta(i)} = \sum_{i\in I} x_i$$

□

Using this more general definition of a finite sum we can rewrite [theorems: 11.16, 11.6, 11.8, 11.9 and 11.7].

Theorem 11.37. *Let $\langle A, +\rangle$ be a Abelian semi-group with neutral element 0, I a finite set and $\{x_i\}_{i\in I} \subseteq A$ with $\forall i \in I\ x_i = 0$ then $\sum_{i\in I} x_i = 0$.*

Proof. For I we have either

$I = \varnothing$. Then $\sum_{i\in I} x_i = \varnothing$

$I \neq \varnothing$. Then there exists a $n \in \mathbb{N}$ and a bijection $\beta: \{0, \ldots, n-1\} \rightarrow I$ such that

$$\sum_{i\in I} x_i = \sum_{i=0}^{n-1} x_{\beta(i)}$$

As $\forall i \in \{0, \ldots, n-1\}\quad \beta(i) \in I$ we have $x_{\beta(i)} = 0$ so that by [theorem: 11.5]

$$\sum_{i=0}^{n-1} x_{\beta(i)} = 0,$$

hence we conclude that

$$\sum_{i\in I} x_i = 0$$

□

Theorem 11.38. *Let $\langle A, +\rangle$ be a Abelian semi-group, I a finite set and $\{x_i\}_{i\in I} \subseteq A$, $\{y_i\}_{i\in I} \subseteq A$ then*

$$\sum_{i\in I} (x_i + y_i) = \sum_{i\in I} x_i + \sum_{i\in I} y_i$$

Proof. For I we have either:

$I = \varnothing$. Then we have

$$\sum_{i\in I} (x_i + y_i) = 0 = 0 + 0 = \sum_{i\in I} x_i + \sum_{i\in I} y_i$$

$I \neq \varnothing$. Then there exists a $n \in \mathbb{N}$ and a bijection $\beta: \{0, \ldots, n-1\} \rightarrow I$ such that

$$\sum_{i\in I} (x_i + y_i) = \sum_{i=0}^{n-1} (x_{\beta(i)} + y_{\beta(i)}),\ \sum_{i\in I} x_i = \sum_{i=0}^{n-1} x_{\beta(i)} \text{ and } \sum_{i\in I} y_i = \sum_{i=0}^{n-1} y_{\beta(i)}$$

By [theorem: 11.6] we have that $\sum_{i=0}^{n-1} (x_{\beta(i)} + y_{\beta(i)}) = \sum_{i=0}^{n-1} x_{\beta(i)} + \sum_{i=0}^{n-1} y_{\beta(i)}$ proving that

$$\sum_{i \in I} (x_i + y_i) = \sum_{i \in I} x_i + \sum_{i \in I} y_i$$ □

Theorem 11.39. *Let $\langle R, +, \cdot \rangle$ be a ring, $\alpha \in R$, I a finite set and $\{x_i\}_{i \in I} \subseteq R$ then*

$$\sum_{i \in I} \alpha \cdot x_i = \alpha \cdot \sum_{i \in I} x_i$$

Proof. For I we have either:

***I* = ∅.** Then we have

$$\sum_{i \in I} \alpha \cdot x_i = 0 = \alpha \cdot 0 = \alpha \cdot \sum_{i \in I} x_i$$

***I* ≠ ∅.** Then there exists a $n \in \mathbb{N}$ and a bijection $\beta: \{0, \ldots, n-1\} \to I$ such that

$$\sum_{i \in I} \alpha \cdot x_i = \sum_{i=0}^{n-1} \alpha \cdot x_{\beta(i)} \underset{[\text{theorem: } 11.8]}{=} \alpha \cdot \sum_{i=0}^{n-1} x_{\beta(i)} = \alpha \cdot \sum_{i \in I} x_i$$ □

Theorem 11.40. *Let $\langle A, +, \cdot \rangle$ be a Abelian semi-group, I a finite set and $\{x_i\}_{i \in I} \subseteq A$ then*

$$\sum_{i \in I} (-x_i) = -\sum_{i \in I} x_i$$

Proof. For I we have either:

***I* = ∅.** Then we have

$$\sum_{i \in I} (-x_i) = 0 = -0 = -\sum_{i \in I} x_i$$

***I* ≠ ∅.** Then there exists a $n \in \mathbb{N}$ and a bijection $\beta: \{0, \ldots, n-1\} \to I$ such that

$$\sum_{i \in I} (-x_i) = \sum_{i=0}^{n-1} (-x_{\beta(i)}) \underset{[\text{theorem: } 11.7]}{=} -\sum_{i=0}^{n-1} x_{\beta(i)} = -\sum_{i \in I} x_i$$ □

Theorem 11.41. *Let $\langle A, + \rangle$ be a Abelian semi-group, I, J finite sets and $\{x_{i,j}\}_{(i,j) \in I \times J} \subseteq A$ then*

$$\sum_{i \in I} \left(\sum_{j = J} x_{i,j} \right) = \sum_{j \in I} \left(\sum_{i \in I} x_{i,j} \right)$$

Proof. For I, J we have either:

***I* = ∅.** Then we have

$$\sum_{i \in I} \left(\sum_{j \in J} x_{i,j} \right) = 0 \underset{[\text{theorem: } 11.37]}{=} \sum_{j \in J} 0 = \sum_{j \in J} \left(\sum_{i \in I} x_{i,j} \right)$$

***J* = ∅.** Then we have

$$\sum_{i \in I} \left(\sum_{j \in J} x_{i,j} \right) = \sum_{i \in I} 0 \underset{[\text{theorem: } 11.37]}{=} 0 = \sum_{j \in J} \left(\sum_{i \in I} x_{i,j} \right)$$

***I* ≠ ∅ ∧ *J* ≠ ∅.** Then there exists $n, m \in \mathbb{N}$ and bijections

$$\alpha: \{0, \ldots, n-1\} \to I,\ \beta: \{0, \ldots, m-1\} \to J$$

such that

$$\forall j\in J \text{ we have } \sum_{i\in I} x_{i,j} = \sum_{i=0}^{n-1} x_{\alpha(i),j} \text{ and } \forall i\in I \text{ we have } \sum_{j\in J} x_{i,j} = \sum_{j=0}^{m-1} x_{i,\beta(j)} \tag{11.7}$$

so that

$$\begin{aligned}
\sum_{i\in I}\left(\sum_{j\in J} x_{i,j}\right) &\underset{[\text{eq: } 11.7]}{=} \sum_{i\in I}\left(\sum_{j=0}^{m-1} x_{i,\beta(j)}\right)\\
&= \sum_{i=0}^{n-1}\left(\sum_{j=0}^{m-1} x_{\alpha(i),\beta(j)}\right)\\
&\underset{[\text{theorem: } 11.9]}{=} \sum_{j=0}^{m-1}\left(\sum_{i=0}^{n-1} x_{\alpha(i),\beta(j)}\right)\\
&= \sum_{j\in J}\left(\sum_{i=0}^{n-1} x_{\alpha(i),j}\right)\\
&\underset{[\text{eq: } 11.7]}{=} \sum_{j\in J}\left(\sum_{i\in I} x_{i,j}\right)
\end{aligned}$$

□

We have the equivalent of [theorem: 11.23]

Lemma 11.42. *Let $\langle A,+\rangle$ be a Abelian semi-group with neutral element 0, I a finite set such that $I=M\bigcup N$ and $M\bigcap N=\varnothing$ and $\{x_i\}_{i\in I}\subseteq A$ then we have*

$$\sum_{i\in I} x_i = \sum_{i\in M} x_i + \sum_{i\in N} x_i$$

Proof. As I is finite we have by [theorem: 6.32] that M,N are finite and we have either:

$M=\varnothing$. Then $N=I$ and

$$\sum_{i\in I} x_i = \sum_{i\in N} x_i = \sum_{i\in N} x_i + 0 = \sum_{i\in N} x_i + \sum_{i\in M} x_i$$

$N=\varnothing$. Then $M=I$ and

$$\sum_{i\in I} x_i = \sum_{i\in M} x_i = 0 + \sum_{i\in M} x_i = \sum_{i\in N} x_i + \sum_{i\in M} x_i$$

$M,N\neq\varnothing$. Then there exists $n,m\in\mathbb{N}$ such $\alpha:\{0,\ldots,n-1\}\to N$ and $\beta:\{0,\ldots,m-1\}\to M$ are bijections and

$$\sum_{i\in N} x_i = \sum_{i=0}^{n-1} x_{\alpha(i)} \text{ and } \sum_{i\in M} x_i = \sum_{i=0}^{m-1} x_{\beta(i)} \tag{11.8}$$

Define now

$$\gamma:\{0,\ldots,n+m-1\}\to N\bigcup M \text{ by } \gamma(i) = \begin{cases} \alpha(i) \text{ where } i\in\{0,\ldots,n-1\}\\ \beta(i-n) \text{ where } i\in\{n,\ldots,n+m-1\}\end{cases}$$

then we have:

injectivity. If $\gamma(i)=\gamma(j)$ then for i,j we have either:

$i\in\{0,\ldots,n-1\}\wedge j\in\{0,\ldots,n-1\}$. Then $\alpha(i)=\gamma(i)=\gamma(j)=\alpha(j)$ which as α is injective proves $i=j$.

$i\in\{0,\ldots,n-1\}\wedge j\in\{n,\ldots,n+m-1\}$. Then $\alpha(i)=\gamma(i)=\gamma(j)=\beta(j-n)$ so that $\alpha(i)\in N\bigcap M$ contradicting $N\bigcap M=\varnothing$, so this case never occurs.

$i\in\{n,\ldots,n+m-1\}\wedge j\in\{0,\ldots,n-1\}$. Then $\beta(i-n)=\gamma(i)=\gamma(j)=\alpha(j)$ so that $\alpha(j)\in N\bigcap M$ contradicting $N\bigcap M=\varnothing$, so this case never occurs.

$i\in\{n,\ldots,n+m-1\}\wedge j\in\{n,\ldots,n+m-1\}$. Then $\beta(i-n)=\gamma(i)=\gamma(j)=\beta(j-n)$ which as β is injective gives $i-n=j-n$ proving $i=j$.

So in all valid cases we have $i=j$ proving injectivity.

surjectivity. If $y\in I=N\bigcup M$ then we have either:

$y\in N$. Then as α is surjective there exist a $i\in\{0,\ldots,n-1\}$ such that $y=\alpha(i)$ which as $\gamma(i)=\alpha(i)$ proves that $\gamma(i)=y$.

$y\in M$. Then as β is surjective there exist a $i\in\{0,\ldots,m-1\}$ such that $y=\beta(i)$. For $j=i+n$ we have $n\leqslant i\leqslant n+m-1$ and $\gamma(j)=\beta(j-n)=\beta(i)=y$

So

$$\gamma\colon\{0,\ldots,n+m-1\}\rightarrow N\bigcup M=I \text{ is a bijection}$$

Hence

$$\begin{aligned}
\sum_{i\in I}x_i &= \sum_{i=0}^{n+m-1}x_{\gamma(i)}\\
&\underset{[\text{theorem: } 11.23]}{=} \sum_{i=0}^{n-1}x_{\gamma(i)}+\sum_{i=n}^{n+m-1}x_{\gamma(i)}\\
&= \sum_{i=0}^{n-1}x_{\alpha(i)}+\sum_{i=n}^{n+m-1}x_{\beta(i-n)}\\
&= \sum_{i=0}^{n-1}x_{\alpha(i)}+\sum_{i=0}^{m-1}x_{\beta((i+n)-n)}\\
&= \sum_{i=0}^{n-1}x_{\alpha(i)}+\sum_{i=0}^{m-1}x_{\beta(i)}\\
&\underset{[\text{theorem: } 11.8]}{=} \sum_{i\in N}x_i+\sum_{i\in M}x_i
\end{aligned}$$

□

We have now the equivalence of [theorem: 11.24]

Theorem 11.43. *Let $\langle A,+\rangle$ be a Abelian semi-group, $n\in\mathbb{N}_0$, $\{I_i\}_{i\in\{0,\ldots,n\}}$ a family of finite sets such that $\forall i,j\in I \text{ with } i\neq j\ I_i\bigcap I_j=\varnothing$ then for $\{x_i\}_{i\in\bigcup_{j\in\{0,\ldots,n\}}I_j}\subseteq A$ we have*

$$\sum_{i\in\bigcup_{j\in\{0,\ldots,n\}}I_j}x_i=\sum_{i=0}^{n}\left(\sum_{j\in I_i}x_j\right)$$

Proof. We prove this by induction, so take

$$S=\left\{n\in\mathbb{N}_0|\forall\{I_i\}_{i\in\{0,\ldots,n\}}\text{ pairwise disjoint family of sets, }\forall\{x_i\}_{i\in\bigcup_{j\in\{0,\ldots,n\}}I_j}\subseteq A\text{ we have }\sum_{i\in\bigcup_{j\in\{0,\ldots,n\}}I_j}x_i=\sum_{i=0}^{n}\left(\sum_{j\in I_i}x_j\right)\right\}$$

then we have:

$0\in S$. Let $\{I_i\}_{i\in\{0\}}$ be a pairwise disjoint family of sets then $\bigcup_{i\in\{0\}}I_i=I_0$, let $\{x_i\}_{i\in\bigcup_{j\in\{0\}}I_j}\subseteq A$ then $\sum_{i\in\bigcup_{j\in\{0\}}I_j}x_i=\sum_{i\in I_0}x_i=\sum_{i=0}^{0}\left(\sum_{j\in I_i}x_j\right)$ proving that $0\in S$.

$n\in S\Rightarrow n+1\in S$. Let $\{I_i\}_{i\in\{0,\ldots,n+1\}}$ be a pairwise disjoint family of finite sets and $\{x_i\}_{i\in\bigcup_{j\in\{0,\ldots,n+1\}}I_j}$. Take $J=\bigcup_{i\in\{0,\ldots,n\}}I_i$ then by [theorem: 6.35] J is finite and

$$J\bigcap I_{n+1}=\left(\bigcup_{i\in\{0,\ldots,n\}}I_i\right)\bigcap I_{n+1}=\bigcup_{i\in\{0,\ldots,n\}}(I_i\bigcap I_{n+1})=\bigcup_{i\in\{0,\ldots,n\}}\varnothing=\varnothing$$

proving

$$\bigcup_{j\in\{0,\ldots,n+1\}}I_j=J\bigcup I_{n+1},\ J,I_{n+1}\text{ are finite and }J\bigcap I_{n+1}=\varnothing \tag{11.9}$$

So we have

$$\begin{aligned}
\sum_{i\in\bigcup_{i\in\{0,\ldots n\}}I_i} x_i &\underset{\text{[eq: 11.9] and [lemma: 11.42]}}{=} \sum_{i\in J} x_i+\sum_{i\in I_{n+1}} x_i\\
&= \sum_{i\in\bigcup_{j\in\{0,\ldots,n\}}I_j} x_i+\sum_{i\in I_{n+1}} x_i\\
&\underset{n\in S}{=} \sum_{i=0}^{n}\left(\sum_{j\in I_i} x_j\right)+\sum_{i\in I_{n+1}} x_i\\
&= \sum_{i=0}^{n+1}\left(\sum_{j\in I_i} x_i\right)
\end{aligned}$$

proving that $n+1\in S$. □

Corollary 11.44. *Let $\langle A,+\rangle$ be a Abelian group, I a finite set, $\{I_i\}_{i\in I}$ a family of finite sets such that $\forall i,j\in I$ with $i\neq j$ we have $I_i\bigcap I_j=\varnothing$ and $\{x_i\}_{i\in\bigcup_{j\in I}I_j}\subseteq A$ then*

$$\sum_{i\in\bigcup_{j\in I}I_j} x_i=\sum_{i\in I}\left(\sum_{j\in I_i} x_j\right)$$

Proof. For I we have either:

***$I=\varnothing$*.** Then $\bigcup_{j\in I}I_j=\varnothing$ so that

$$\sum_{i\in\bigcup_{j\in I}I_j} x_i=\sum_{i\in\varnothing} x_i=0=\sum_{i\in\varnothing}\left(\sum_{j\in I_j} x_j\right)$$

***$I\neq\varnothing$*.** Then there exist a $n\in\mathbb{N}$ and a bijection $\beta:\{0,\ldots,n-1\}\to I$. Using [theorem: 2.117] we have that

$$\bigcup_{j\in I} I_j=\bigcup_{j\in\{0,\ldots,n-1\}} I_{\beta(j)}$$

so that

$$\begin{aligned}
\sum_{i\in\bigcup_{j\in I}I_j} x_i &= \sum_{i\in\bigcup_{j\in\{0,\ldots,n-1\}}I_{\beta(j)}} x_i\\
&\underset{\text{[theorem: 11.43]}}{=} \sum_{i=0}^{n-1}\left(\sum_{j\in I_{\beta(i)}} x_j\right)\\
&= \sum_{i\in I}\left(\sum_{j\in I_i} x_j\right)
\end{aligned}$$

□

Corollary 11.45. *Let $\langle A,+\rangle$ be a Abelian semi-group, I,J finite sets and $\{x_{i,j}\}_{(i,j)\in I\times J}$ then*

$$\sum_{(i,j)\in I\times J} x_{i,j}=\sum_{i\in I}\left(\sum_{j\in J} x_{i,j}\right)\underset{\text{[theorem: 11.41]}}{=}\sum_{j\in J}\left(\sum_{i\in I} x_{i,j}\right)$$

Proof. Given $i\in I$ let $J_i=\{i\}\times J$ then as $I=\bigcup_{i\in I}\{i\}$ we have by [theorem: 2.131] that

$$I\times J=\bigcup_{i\in I} J_i$$

Further if $i,j\in I$ with $i\neq j$ then if $x\in J_i\bigcap J_j=(\{i\}\times J)\bigcap(\{j\}\times J)$ then $\exists k,l\in J$ such that $(i,k)=x=(j,l)$ giving $i=j$ contradicting $i\neq j$. So

$$\forall i,j\in I \text{ with } i\neq j \text{ we have } \{i\}\times J\bigcap\{j\}\times J=\varnothing \text{ or } J_i\neq J_j$$

So we can apply [theorem: 11.44] giving

$$\begin{aligned}\sum_{(i,j)\in I\times J} x_{i,j} &= \sum_{(i,j)\in\bigcup_{k\in I}J_k} x_{i,j}\\ &\underset{[\text{theorem: }11.44]}{=} \sum_{k\in I}\left(\sum_{(i,j)\in J_k} x_{i,j}\right)\\ &\underset{(i,j)\in\{k\}\times J\Leftrightarrow j=k}{=} \sum_{k\in I}\left(\sum_{(k,j)\in J_k} x_{k,j}\right)\end{aligned}\tag{11.10}$$

Given $k\in I$ define $\beta_k: J\rightarrow J_k$ by $\beta_k(j)=(k,j)\in\{k\}\times J=J_k$ then we have

injectivity. If $\beta_k(i)=\beta_k(j)$ then $(k,i)=(k,j)$ giving $i=j$

surjectivity. If $(k,j)\in\{k\}\times J$ then $\beta_k(j)=(k,j)$

so that $\beta_k: K\rightarrow J_k$ is a bijection, hence by [theorem: 11.36]

$$\sum_{j\in J} x_{k,j}=\sum_{j\in J} x_{\beta_k(j)} \underset{[\text{theorem: }11.36]}{=} \sum_{(k,j)\in J_k} x_{k,j}$$

which combined with [eq: 11.10] gives

$$\sum_{(i,j)\in I\times J} x_{i,j}=\sum_{k\in I}\left(\sum_{j\in J} x_{k,j}\right)$$

□

Corollary 11.46. *Let $\langle R,+,\cdot\rangle$ be a ring, I, J finite sets and $\{x_i\}_{i\in I}\subseteq R$, $\{y_i\}_{i\in J}\subseteq R$ then*

$$\sum_{(i,j)\in I\times J}(x_i\cdot x_j)=\left(\sum_{i\in I} x_i\right)\cdot\left(\sum_{i\in J} y_i\right)$$

Proof. As $\langle R,+\rangle$ is a Abelian group we have

$$\begin{aligned}\sum_{(i,j)\in I\times J}(x_i\cdot y_j) &\underset{[\text{corollary: }11.45]}{=} \sum_{i\in I}\left(\sum_{j\in J}(x_i\cdot y_j)\right)\\ &\underset{[\text{theorem: }11.39]}{=} \sum_{i\in I}\left(x_i\cdot\sum_{j\in J} y_j\right)\\ &\underset{[\text{theorem: }11.39]}{=} \left(\sum_{i\in I} x_i\right)\cdot\left(\sum_{j\in J} y_j\right)\\ &= \left(\sum_{i\in I} x_i\right)\cdot\left(\sum_{i\in J} y_i\right)\end{aligned}$$

□

Theorem 11.47. *Let $\langle A,+\rangle$ be a Abelian semi-group, $n\in\mathbb{N}$, $\langle A^n,+\rangle$ the semi-group based on $\langle A,+\rangle$ [see theorem: 6.84] then we have:*

1. *Let $k,l\in\mathbb{N}_0$ with $k\leqslant l$ and $\{x_i\}_{i\in\{k,\ldots,l\}}\subseteq A^n$ then $\forall i\in\{1,\ldots,n\}$ we have*

$$\left(\sum_{j=k}^{l} x_j\right)_i=\sum_{j=k}^{l}(x_j)_i$$

2. *Let J be a finite set and $\{x_j\}_{j\in J}\subseteq A^n$ then for $i\in\{1,\ldots,n\}$ we have $(\sum_{j\in J} x_j)_i=\sum_{j\in J}(x_j)_i$.*

Proof.

1. We use induction to prove this for the case $\sum_{j=0}^{l} x_j$, so define

$$S=\left\{l\in\mathbb{N}_0|\text{If }\{x_i\}_{i\in\{0,\ldots,l\}}\subseteq A^n\text{ then }\left(\sum_{j=0}^{l} x_j\right)_i=\sum_{j=0}^{l}(x_j)_i\right\}$$

then we have:

$0\in S$. If $\{x_j\}_{j\in\{0\}}\subseteq A^n$ then for $i\in\{1,\ldots,n\}$ we have $(\sum_{j=0}^{0} x_j)_i=(x_0)_i=\sum_{j=0}^{0}(x_j)_i$ proving that $0\in S$.

$l \in S \Rightarrow l+1 \in S$. Let $\{x_j\}_{j\in\{0,\ldots,l+1\}} \subseteq A^n$ then we have

$$\begin{aligned}\left(\sum_{j=0}^{l+1} x_j\right)_i &= \left(\left(\sum_{j=0}^{l} x_j\right) + x_{l+1}\right)_i \\ &\underset{[\text{theorem: } 6.84]}{=} \left(\sum_{j=0}^{l} x_j\right)_i + (x_{l+1})_i \\ &\underset{l\in S}{=} \left(\sum_{j=0}^{l} (x_j)_i\right) + (x_{l+1})_j \\ &= \sum_{j=0}^{l+1} (x_j)_i\end{aligned}$$

proving that $l+1 \in S$.

By mathematical induction it follows then that

$$\forall l \in \mathbb{N}_0 \text{ we have for } \{x_j\}_{j\in\{1,\ldots,l\}} \subseteq A^n \text{ and } i \in \{1,\ldots,n\} \text{ that } \left(\sum_{j=0}^{l} x_j\right)_i = \sum_{j=0}^{l} (x_j)_i \tag{11.11}$$

Let now $k, l \in \mathbb{N}_0$ with $k \leqslant l$ and $\{x_j\}_{j\in\{k,\ldots,l\}} \subseteq A^n$ then we have for $i \in \{1,\ldots,n\}$

$$\left(\sum_{j=k}^{l} x_j\right)_i = \left(\sum_{j=0}^{l-k} x_{j+k}\right)_i \underset{[\text{eq: } 11.11]}{=} \sum_{j=0}^{l-k} (x_{j+k})_i = \sum_{j=k}^{l} (x_j)_i$$

2. If I is a finite set, $i \in \{1,\ldots,n\}$ and $\{x_j\}_{j\in I} \subseteq I$ then we have for I either:

$I = \varnothing$. Then

$$\left(\sum_{j\in I} x_j\right)_i = \left(\sum_{j\in\varnothing} x_j\right)_i = 0_i = \sum_{j\in\varnothing} (x_j)_i$$

$I \neq \varnothing$. Then there exists a $l \in \mathbb{N}_0$ and bijection $\beta: \{0,\ldots,l\} \rightarrow I$ then

$$\left(\sum_{j\in I} x_j\right)_i \underset{\text{def}}{=} \left(\sum_{j=0}^{l} x_{\beta(j)}\right)_i \underset{[\text{theorem: } 11.11]}{=} \sum_{j=0}^{l} (x_{\beta(j)})_i = \sum_{j\in I} (x_j)_i$$ □

Theorem 11.48. *For the field $\langle \mathbb{K}, \cdot, + \rangle$ where $\mathbb{K} = \mathbb{C}$ or $\mathbb{K} = \mathbb{R}$ we have*

1. *If $n \in \mathbb{N}_0$ and $\{\alpha_i\}_{i\in\{0,\ldots,n\}} \subseteq \mathbb{K}$ is defined by $\alpha_i = \alpha$ then*
   a. $\sum_{i=0}^{n} \alpha_i = (n+1) \cdot \alpha$
   b. $\prod_{i=0}^{n} \alpha_i = \alpha^{(n+1)}$
2. *If $n, m \in \mathbb{N}_0$ with $n \leqslant m$ and $\{\alpha_i\}_{i\in\{n,\ldots,m\}} \subseteq \mathbb{K}$ is defined by $\alpha_i = \alpha$ then*
   a. $\sum_{i=n}^{m} \alpha_i = (m-n+1) \cdot \alpha$
   b. $\prod_{i=n}^{m} \alpha_i = \alpha^{m-n+1}$
3. *If I is a finite set then for $\{\alpha_i\}_{i\in I} \subseteq \mathbb{K}$ defined by $\alpha_i = \alpha$ we have*
   a. $\sum_{i\in I} \alpha_i = \mathrm{card}(I) \cdot \alpha$
   b. $\prod_{i\in I} \alpha_i = \alpha^{\mathrm{card}(I)}$

Proof.

1.

a. We prove this by induction, so let

$$S = \left\{ n \in \mathbb{N}_0 | \text{If } \{\alpha_i\}_{i\in\{0,\ldots,n\}} \subseteq \mathbb{K} \text{ is defined by } \alpha_i = \alpha \text{ then } \sum_{i=0}^{n} \alpha_i = (n+1) \cdot \alpha \right\}$$

then we have:

$0\in S$. Then $\sum_{i=0}^{0} \alpha_i = \alpha_0 = \alpha = (0+1)\cdot\alpha$ proving that $0\in S$

$n\in S \Rightarrow n+1\in S$. Let $\{\alpha_i\}_{i\in\{1,\ldots,n+1\}} \subseteq \mathbb{K}$ be defined by $\alpha_i = \alpha$. Then

$$\begin{array}{rcl}\sum_{i=0}^{n+1} \alpha_i & = & \sum_{i=0}^{n} \alpha_i + \alpha_{n+1} \\ & = & \sum_{i=0}^{n} \alpha_i + \alpha \\ & \underset{n\in S}{=} & (n+1)\cdot\alpha + \alpha \\ & = & ((n+1)+1)\cdot\alpha\end{array}$$

proving that $n+1\in S$

b. We prove this by induction, so let

$$S = \left\{ n\in\mathbb{N}_0 | \text{If } \{\alpha_i\}_{i\in\{0,\ldots,n\}} \subseteq \mathbb{K} \text{ is defined by } \alpha_i = \alpha \text{ then } \prod_{i=0}^{n} \alpha_i = (n+1)\cdot\alpha \right\}$$

then we have:

$0\in S$. Then $\prod_{i=0}^{0} \alpha_i \underset{[\text{theorem: } 11.34]}{=} \alpha_0 = \alpha = \alpha^1 = \alpha^{(0+1)}$ proving that $0\in S$

$n\in S \Rightarrow n+1\in S$. Let $\{\alpha_i\}_{i\in\{1,\ldots,n+1\}} \subseteq \mathbb{R}$ be defined by $\alpha_i = \alpha$. Then

$$\begin{array}{rcl}\prod_{i=0}^{n+1} \alpha_i & = & \left(\prod_{i=0}^{n} \alpha_i\right)\cdot\alpha_{n+1} \\ & = & \left(\prod_{i=0}^{n} \alpha_i\right)\cdot\alpha \\ & \underset{n\in S}{=} & \alpha^{n+1}\cdot\alpha \\ & = & \alpha^{(n+1)+1}\end{array}$$

proving that $n+1\in S$

2.

a. $\sum_{i=n}^{m} \alpha_i = \sum_{i=0}^{m-n} \alpha_{i+n} \underset{(1.a)}{=} (m-n+1)\cdot\alpha$

b. $\prod_{i=n}^{m} \alpha_i = \prod_{i=0}^{m-n} \alpha_{i+n} \underset{(1.a)}{=} \alpha^{m-n+1}$

3.

a. For I we have either:

$I=\varnothing$. Then $\sum_{i\in I} \alpha_i = \sum_{i\in\varnothing} \alpha_i = 0$

$I\neq\varnothing$. Then $\{0,\ldots,\text{card}(I)-1\} \approx \{1,\ldots,\text{card}(i)\}$ so that there exists a bijection $\beta\colon \{0,\ldots,\text{card}(I)-1\} \to I$ so that

$$\sum_{i\in I} \alpha_i = \sum_{0}^{\text{card}(I)-1} \alpha_{\beta(i)} \underset{(1.a)}{=} (\text{card}(I)-1+1)\cdot\alpha = \text{card}(I)\cdot\alpha$$

b. For I we have either:

$I=\varnothing$. Then $\prod_{i\in I} \alpha_i = \prod_{i\in\varnothing} \alpha_i = 1$

$I\neq\varnothing$. Then $\{0,\ldots,\text{card}(I)-1\} \approx \{1,\ldots,\text{card}(i)\}$ so that there exists a bijection $\beta\colon \{0,\ldots,\text{card}(I)-1\} \to I$ so that

$$\prod_{i\in I} \alpha_i = \prod_{0}^{\text{card}(I)-1} \alpha_{\beta(i)} \underset{(2.a)}{=} \alpha^{(\text{card}(I)-1+1)} = \alpha^{\text{card}(I)}$$

□

Theorem 11.49. *If $\langle F, +, \cdot \rangle$ is a field then we have*

1. *Let I be a finite set, $i \in I$, $x, y \in F$ and $\{\alpha_j\}_{j \in I} \subseteq F$ such that $\alpha_i = x + y$ then*
$$\prod_{j \in I} \alpha_j = \prod_{j \in I} \beta_j + \prod_{j \in I} \gamma_j$$
where $\{\beta_j\}_{j \in I}$ and $\{\gamma_j\}_{j \in I}$ are defined by
$$\beta_j = \begin{cases} x & \text{if } j = i \\ \alpha_j & \text{if } j \in I \setminus \{i\} \end{cases} \text{ and } \gamma_j = \begin{cases} y & \text{if } j = i \\ \alpha_j & \text{if } j \in I \setminus \{i\} \end{cases}$$
2. *Let I be a finite set, $i \in I$, $x, \beta \in F$, and $\{\alpha_j\}_{j \in I} \subseteq F$ such that $\alpha_i = \beta \cdot x$ then*
$$\prod_{j \in I} \alpha_j = \beta \cdot \prod_{j \in I} \beta_j$$
where $\{\beta_j\}_{j \in I}$ is defined by $\beta_j = \left\{ \begin{smallmatrix} x & \text{if } j = i \\ \alpha_j & \text{if } j \in I \setminus \{i\} \end{smallmatrix} \right.$.
3. *Let $n, m \in \mathbb{N}_0$ with $n \leqslant m$, $i \in \{n, \ldots, m\}$, $x, y \in F$ and $\{\alpha_j\}_{j \in \{n, \ldots, m\}} \subseteq F$ such that $\alpha_i = x + y$ then*
$$\prod_{j=n}^{m} \alpha_j = \prod_{j=n}^{m} \beta_j + \prod_{j=n}^{m} \gamma_j$$
where $\{\beta_j\}_{j \in \{n, \ldots, m\}}$ and $\{\gamma_j\}_{j \in \{n, \ldots, m\}}$ are defined by
$$\beta_j = \begin{cases} x & \text{if } j = i \\ \alpha_j & \text{if } j \in \{n, \ldots, m\} \setminus \{i\} \end{cases} \text{ and } \gamma_j = \begin{cases} y & \text{if } j = i \\ \alpha_j & \text{if } j \in \{n, \ldots, m\} \setminus \{i\} \end{cases}$$
4. *Let $n, m \in \mathbb{N}_0$ with $n \leqslant m$, $i \in \{n, \ldots, m\}$, $x, \beta \in F$, and $\{\alpha_j\}_{j \in \{n, \ldots, m\}} \subseteq F$ such that $\alpha_i = \beta \cdot x$ then*
$$\prod_{j=n}^{n} \alpha_j = \beta \cdot \prod_{j=m}^{n} \beta_j$$
where $\{\beta_j\}_{j \in \{n, \ldots, m\}}$ is defined by $\beta_j = \left\{ \begin{smallmatrix} x & \text{if } j = i \\ \alpha_j & \text{if } j \in \{n, \ldots, m\} \setminus \{i\} \end{smallmatrix} \right.$.

Proof.

1. We have
$$\begin{aligned}
\prod_{j \in I} \alpha_j &\underset{[\text{theorem: } 11.43]}{=} \left(\prod_{j \in I \setminus \{i\}} \alpha_j \right) \cdot \left(\prod_{j \in \{i\}} \alpha_j \right) \\
&\underset{[\text{example: } 11.34]}{=} \left(\prod_{j \in I \setminus \{i\}} \alpha_j \right) \cdot \alpha_i \\
&= \left(\prod_{j \in I \setminus \{i\}} \alpha_j \right) \cdot (x + y) \\
&= \left(\prod_{j \in I \setminus \{i\}} \alpha_j \right) \cdot x + \left(\prod_{j \in I \setminus \{i\}} \alpha_j \right) \cdot y \\
&= \left(\prod_{j \in I \setminus \{i\}} \beta_j \right) \cdot \beta_i + \left(\prod_{j \in I \setminus \{i\}} \gamma_j \right) \cdot \gamma_i \\
&\underset{[\text{theorem: } 11.34]}{=} \left(\prod_{j \in I \setminus \{i\}} \beta_j \right) \cdot \left(\prod_{j \in \{i\}} \beta_j \right) + \left(\prod_{j \in I \setminus \{i\}} \gamma_j \right) \cdot \left(\prod_{j \in \{i\}} \gamma_j \right) \\
&\underset{[\text{theorem: } 11.43]}{=} \prod_{j \in I} \beta_j + \prod_{j \in I} \gamma_j
\end{aligned}$$

2. We have

$$\begin{aligned}\prod_{j\in I}\alpha_j &\underset{[\text{theorem: } 11.43]}{=} \left(\prod_{j\in I\setminus\{i\}}\alpha_j\right)\cdot\left(\prod_{j\in\{i\}}\alpha_j\right)\\ &\underset{[\text{theorem: } 11.34]}{=} \left(\prod_{j\in I\setminus\{i\}}\alpha_j\right)\cdot\alpha_i\\ &= \left(\prod_{j\in I\setminus\{i\}}\alpha_j\right)\cdot(\beta\cdot x)\\ &= \beta\cdot\left(\left(\prod_{j\in I\setminus\{i\}}\alpha_j\right)\cdot x\right)\\ &= \beta\cdot\left(\left(\prod_{j\in I\setminus\{i\}}\beta_j\right)\cdot\beta_i\right)\\ &\underset{[\text{theorem: } 11.34]}{=} \beta\cdot\left(\left(\prod_{j\in I\setminus\{i\}}\beta_j\right)\cdot\left(\prod_{j\in\{i\}}\beta_j\right)\right)\\ &\underset{[\text{theorem: } 11.43]}{=} \beta\cdot\prod_{j\in I}\beta_j\end{aligned}$$

3. We have

$$\begin{aligned}\prod_{i=n}^{m}\alpha_i &\underset{[\text{theorem: } 11.35]}{=} \prod_{i\in\{n,\ldots,m\}}\alpha_i\\ &\underset{(1)}{=} \prod_{i\in\{n,\ldots,m\}}\beta_i+\prod_{i\in\{n,\ldots,m\}}\gamma_i\\ &\underset{[\text{theorem: } 11.35]}{=} \prod_{i=n}^{m}\beta_i+\prod_{i=n}^{m}\gamma_i\end{aligned}$$

4. We have

$$\begin{aligned}\prod_{i=n}^{m}\alpha_i &\underset{[\text{theorem: } 11.35]}{=} \prod_{i\in\{n,\ldots,m\}}\alpha_i\\ &\underset{(2)}{=} \beta\cdot\prod_{i\in\{n,\ldots,m\}}\beta_i\\ &\underset{[\text{theorem: } 11.35]}{=} \beta\cdot\prod_{i=n}^{m}\beta_i\end{aligned}$$

$\square$

Theorem 11.50. *If $\langle F,\cdot,+\rangle$ is a field then*

1. *Let $n\in\mathbb{N}_0$ and $\{\alpha_i\}_{i\in\{0,\ldots,n\}}\subseteq F\setminus\{0\}$ then we have $\prod_{i=0}^{n}\alpha_i\neq 0$*
2. *Let $n,m\in\mathbb{N}_0$ and $\{\alpha_i\}_{i\in\{n,\ldots,m\}}\subseteq F\setminus\{0\}$ then we have $\prod_{i=n}^{m}\alpha_i\neq 0$*
3. *If I is a finite set and $\{\alpha_i\}_{i\in I}\subseteq F\setminus\{0\}$ then we have $\prod_{i\in I}\alpha_i\neq 0$*

Proof.

1. We prove this by induction, so let

$$S=\left\{n\in\mathbb{N}_0|\text{If } \{\alpha_i\}_{i\in\{0,\ldots,n\}}\subseteq F\setminus\{0\} \text{ then } \prod_{i=0}^{n}\alpha_i\neq 0\right\}$$

then we have:

$0\in S$. Then $\prod_{i=0}^{0}\alpha_i\underset{[\text{theorem: } 11.34]}{=}\alpha_0\neq 0$ proving that $0\in S$.

$n\in S\Rightarrow n+1\in S$. Let $\{\alpha_i\}_{i\in\{0,\ldots,n+1\}}\subseteq F\setminus\{0\}$ then we have

$$\prod_{i=0}^{n+1}\alpha_i = \left(\prod_{i=0}^{n}\alpha_i\right)\cdot\alpha_{n+1}$$

As $n\in S$ we have that $\prod_{i=0}^{n}\alpha_i\neq 0$ so as, $\alpha_{n+1}\neq 0$, so as a field is a integral domain [see: 4.58], we have that $\prod_{i=0}^{n+1}\alpha_i\neq 0$. So $n+1\in S$.

2. $\prod_{i=n}^{m}\alpha_i=\prod_{i=0}^{m-n}\alpha_i\neq 0$ [using (1)]
3. If I is finite then we have either:

   $\boldsymbol{I=\varnothing}$**.** Then $\prod_{i\in S}\alpha_i=\prod_{i\in\varnothing}\alpha_i=1\neq 0$.

   $\boldsymbol{I\neq\varnothing}$**.** Then there exists a bijection $\beta:\{0,\ldots,n-1\}\rightarrow I$ such that

$$\sum_{i\in I}\alpha_i=\sum_{i=0}^{n-1}\alpha_{\beta(i)}\neq 0 \text{ (by (1))}$$

□

Theorem 11.51. *If $\langle R,+,\cdot\rangle$ is a commutative ring with neutral element 0, I a finite set, $\{x_i\}_{i\in I}\subseteq R$ such that $\exists i\in I$ with $x_i=0$ then*

$$\prod_{j\in I}x_j=0$$

Proof. This follows from

$$\prod_{j\in I}x_j\underset{[\text{theorem: }11.41]}{=}\left(\prod_{j\in I\setminus\{i\}}x_j\right)\cdot\prod_{j\in\{i\}}x_j\underset{[\text{theorem: }11.34]}{=}\left(\prod_{j\in I\setminus\{i\}}x_j\right)\cdot x_i=\left(\prod_{j\in I\setminus\{i\}}x_j\right)\cdot 0=0$$

□

Theorem 11.52. *Let I be a finite set and $\{x_i\}_{i\in I}\subseteq\mathbb{R}$ then we have*

1. *If $\forall i\in I$ we have $0\leqslant x_i$ then*

$$0\leqslant\sum_{i\in I}x_i$$

   Further if additional there exist a $i\in I$ such that $0<x_i$ then

$$0<\sum_{i\in I}x_i$$

2. *If $\{y_i\}_{i\in I}\subseteq\mathbb{R}$ is a finite family such that $\forall i\in I$ $x_i\leqslant y_i$ then*

$$\sum_{i\in I}x_i\leqslant\sum_{i\in I}y_i$$

   Further if additional there exist a $i\in I$ such that $x_i<y_i$ then

$$\sum_{i\in I}x_i<\sum_{i\in I}y_i$$

Proof.

1. We prove this by induction on the size of I so let

$$S=\left\{n\in\mathbb{N}_0|\text{If card}(I)=n\text{ then }\forall\{x_i\}_{i\in I}\subseteq R\text{ with }\forall i\in I\ 0\leqslant x_i\text{ we have }0\leqslant\sum_{i\in I}x_i\right\}$$

   then we have:

   $\boldsymbol{0\in S}$**.** If $\text{card}(I)=0$ then $I=\varnothing$ such that $\sum_{i\in I}x_i=\sum_{i\in\varnothing}x_i\underset{[\text{definition: }11.32]}{=}0$ so that $0\leqslant\sum_{i\in I}x_i$.

   $\boldsymbol{n\in S\Rightarrow n+1\in S}$**.** Let I be such that $\text{card}(I)=n+1$ then $I\neq\varnothing$ so that $\exists i_0\in I$ and by [theorem: 10.92] $\text{card}(I\setminus\{i_0\})=n$. Let $\{x_i\}_{i\in I}\subseteq\mathbb{R}$ then

$$\begin{aligned}\sum_{i\in I}x_i&\underset{[\text{theorem: }11.43]}{=}\sum_{i\in I\setminus\{i_0\}}x_i+\sum_{i\in\{i_o\}}x_i\\&\underset{[\text{theorem: }11.34]}{=}\sum_{i\in I\setminus\{i_0\}}x_i+x_{i_0}\end{aligned}$$

   As $n\in S$ we have that $0\leqslant\sum_{i\in I\setminus\{i_0\}}x_i$, further by the hypothesis $0\leqslant x_{i_0}$ hence it follows that $n+1\in S$.

   Hence using mathematical induction we have proved that if I is finite and $\{x_i\}_{i\in I}\subseteq\mathbb{R}$ satisfies $\forall i\in I$ $0\leqslant x_i$ then

$$0\leqslant\sum_{i\in I}x_i$$

Further if there exist a $k\in I$ such that $0<x_k$ then we have

$$0<\sum_{i\in I\setminus\{k\}} x_i+x_k \underset{[\text{theorem: } 11.34]}{=} \sum_{i\in I\setminus\{k\}} x_i+\sum_{i\in\{k\}} x_i \underset{[\text{theorem: } 11.43]}{=} \sum_{i\in I} x_i$$

2. As $\forall i\in I$ we have $x_i\leqslant y_i$ it follows that $0\leqslant y_i-x_i$ so that by (1)

$$\begin{array}{rcl} 0 & \leqslant & \sum_{i\in I}(y_i-x_i) \\ & \underset{[\text{theorems: } 11.40,\ 11.38]}{=} & \sum_{i\in I} y_i-\sum_{i\in I} x_i \end{array}$$

proving that

$$\sum_{i\in I} x_i\leqslant\sum_{i\in I} y_i$$

Further if in addition there exist a $k\in I$ such that $x_i<y_i$ then

$$\begin{array}{rcl} 0 & < & \sum_{i\in I}(y_i-x_i) \\ & \underset{[\text{theorems: } 11.40,\ 11.38]}{=} & \sum_{i\in I} y_i-\sum_{i\in I} x_i \end{array}$$

proving that

$$\sum_{i\in I} x_i<\sum_{i\in I} y_i$$

□

11.2 Vector spaces

11.2.1 Definition

Definition 11.53. *A vector space $\langle V,\oplus,\odot\rangle$ over a field $\langle F,+,\cdot\rangle$ is a Abelian group $\langle V,\oplus\rangle$ together with a map $\odot: F\times V\rightarrow V$ satisfying*

1. *$\forall\alpha\in F$ and $\forall x,y\in V$ we have $\alpha\odot(x\oplus y)=\alpha\odot x+\alpha\odot y$*
2. *$\forall\alpha,\beta\in F$ and $x\in V$ we have $(\alpha+\beta)\odot x=\alpha\odot x\oplus\beta\odot y$*
3. *$\forall\alpha,\beta\in F$ and $x\in V$ we have $(\alpha\cdot\beta)\odot x=\alpha\odot(\beta\odot x)$*
4. *If 1 is the multiplicative neutral element of $\langle F,+,\cdot\rangle$ then $1\odot x=x$*

The map $\odot$ is called the scalar product and elements of V are called vectors.

Note 11.54. Some books call a vector space a linear space, which is maybe clearer as we later will introduce linear (in)dependent sets and linear combinations. In this book we follow the convention of most books about mathematics and physics and use the term vector space.

Theorem 11.55. *Let $\langle V,\oplus,\odot\rangle$ be a vector space over the field $\langle F,+,\cdot\rangle$ where*

a) 0_v is the neutral element of the Abelian group $\langle V,\oplus\rangle$

b) 0_f is the additive neutral element of $\langle F,+,\cdot\rangle$

c) 1 is the multiplicative neutral element of $\langle F,+,\cdot\rangle$

d) For $x\in V$ $-x$ is the inverse of x in $\langle V,\oplus\rangle$

then we have

1. *$\forall x\in V$ we have $0_f\odot x=0_v$*
2. *$\forall x\in V$ we have $(-1)\odot x=-x$*
3. *$\forall\alpha\in F$ we have $\alpha\odot 0_v=0_v$*
4. *If $\alpha\in F\setminus\{0_f\}$ then for $x\in V$ with $\alpha\odot x=0_v$ we have $x=0_v$*
5. *If $x\in V\setminus\{0_v\}$ then for $\alpha\in F$ with $\alpha\odot x=0_v$ we have $\alpha=0_f$.*

Proof.

1. If $x \in V$ then

$$\begin{array}{rcl} 0_v & = & (0_f \odot x) \oplus (-(0_f \odot x)) \\ & \underset{0_f = 0_f + 0_f}{=} & ((0_f + 0_f) \odot x) \oplus (-(0_f \odot x)) \\ & = & ((0_f \odot x) \oplus (0_f \odot x)) \oplus (-(0_f \odot x)) \\ & = & (0_f \odot x) \oplus ((0_f \odot x) \oplus (-(0_f \odot x))) \\ & = & (0_f \odot x) \oplus 0_v \\ & = & 0_f \odot x \end{array}$$

2. If $x \in V$ then

$$\begin{array}{rcl} x \oplus ((-1) \odot x) & = & (1 \odot x) \oplus ((-1) \odot x) \\ & = & (1 + (-1)) \odot x \\ & = & 0_f \odot x \\ & \underset{(1)}{=} & 0_v \end{array}$$

so that by the definition of a inverse element we have

$$(-1) \odot x = -x$$

3. If $\alpha \in F$ then

$$\begin{array}{rcl} \alpha \odot 0_v & \underset{(1)}{=} & \alpha \odot (0_f \odot 0_v) \\ & = & (\alpha \cdot 0_f) \odot 0_v \\ & \underset{[\text{theorem: } 4.39]}{=} & 0_f \odot 0_v \\ & \underset{(1)}{=} & 0_v \end{array}$$

4. Let $\alpha \in F \setminus \{0_f\}$ and take $x \in V$ such that $\alpha \odot x = 0_v$. As $\alpha \neq 0_f$ we have that α^{-1} exist. So that

$$0_v \underset{(3)}{=} \alpha^{-1} \odot 0_v = \alpha^{-1} \odot (\alpha \odot x) = (\alpha^{-1} \cdot \alpha) \odot x = 1 \odot x = x$$

5. Let $x \in V \setminus \{0_v\}$ and $\alpha \in F$ such that $\alpha \odot x = 0_v$. Assume that $\alpha \neq 0_f$ then by (4) we have that $x = 0_v$ contradicting $x \in V \setminus \{0_v\}$, so we must have that $\alpha = 0_f$. □

Just as we have sub-groups, sub-rings and sub-fields we have sub-spaces of a vector space.

Definition 11.56. *Let $\langle V, \oplus, \odot \rangle$ be a vector space over $\langle F, +, \cdot \rangle$ then $W \subseteq V$ is a sub-space of $\langle V, \oplus, \odot \rangle$ if*

1. $W \neq \varnothing$
2. $\forall \alpha \in F$ *and* $x, y \in V$ *we have that* $(\alpha \odot x) \oplus y \in W$

Theorem 11.57. *Let $\langle V, \oplus, \odot \rangle$ be a vector space over $\langle F, +, \cdot \rangle$ and $W \subseteq V$ a sub-space then $\langle W, \oplus_{|W \times W}, \odot_{|F \odot W} \rangle$ is a vector space over $\langle F, +, \cdot \rangle$ where the neutral element of W is the neutral element of V.*

Proof. First we prove that $\langle W, \oplus_{W \times W} \rangle$ is a Abelian group:.

$\oplus_{|W \times W}$ is a operator on W. If $x, y \in W$ then $x \oplus y = (1 \odot x) \oplus y \in W$ so that $\oplus_{|W \times W}$ is indeed a function between $W \times W$ and W.

associativity. If $x, y, z \in W$ then

$$x \oplus_{|W \times W} (y \oplus_{|W \times W} z) \underset{x, y, z \in W}{=} x \oplus (y \oplus z) = (x \oplus y) \oplus z = (x \oplus_{|W \times W} y) \oplus_{|W \times W|} x$$

commutativity. If $x, y \in W$ then $x \oplus_{|W \times W} y = x \oplus y = y \oplus x = y \oplus_{|W \times W} x$

neutral element. Then as $W \neq \varnothing$ there exist a $w \in W$ so that

$$0_v = (-w) \oplus w \underset{[\text{theorem: } 11.55\ (2)]}{=} (-1) \odot w \oplus w \in W$$

proving that $0_v \in W$. Further $\forall x \in W$ we have

$$x \oplus_{|W \times W} 0_v \underset{\text{commutativity}}{=} 0_v \oplus_{|W \times W} x = 0_v \oplus x = x$$

inverse element. If $x \in W$ then $-x \underset{\text{[theorem: 11.55 (2)]}}{=} ((-1) \odot x) = ((-1) \odot x) \oplus 0_v \in W$ so that $-x \in W$. Further $x \oplus_{|W \times W} (-x) \underset{\text{commutativity}}{=} (-x) \oplus_{|W \times W} x = (-x) \oplus x = 0_v$

Further we have:

1. If $\alpha \in F$ and $x \in W$ then $\alpha \odot x = \alpha \odot x \oplus 0_v \in W$ so that $\odot_{|F \times W}$ is a function between $F \times W$ and W.
2. If $\alpha \in F$ and $x, y \in W$ then $\alpha \odot_{|F \times W} (x \oplus_{|W \times W} y) = \alpha \odot (x \oplus y) = \alpha \odot x \oplus \alpha \odot y = \alpha \odot_{|F \times W} x \oplus_{|W \times W} \beta \odot_{|F \times W} y$
3. If $\alpha, \beta \in F$ and $x \in W$ then $(\alpha + \beta) \odot_{|F \times W} x = (\alpha + \beta) \odot x = \alpha \odot x \oplus \beta \odot x = \alpha \odot_{|F \times W} x \oplus_{|W \times W} \beta \odot_{|W \times W} x$
4. If $\alpha, \beta \in F$ and $x \in W$ then $(\alpha \cdot \beta) \odot_{|F \times W} x = (\alpha \cdot \beta) \odot x = \alpha \odot (\beta \odot x) = \alpha \odot_{|F \times W} (\beta \odot_{|F \times W} x)$
5. If $x \in W$ then $1 \odot_{|F \times W} x = 1 \odot x = x$ □

Note 11.58. To avoid excessive use of subscripts we follow for the rest of this book the convention that if $\langle V, \oplus, \odot \rangle$ is a vector space over $\langle F, +, \cdot \rangle$ and $W \in V$ is a sub-space of $\langle V, \oplus, \odot \rangle$ we use $\oplus$ instead of $\oplus_{|W \times W}$ and $\odot$ instead of $\odot_{|F \times W}$. Using this convention we have then that $\langle W, \oplus, \odot \rangle$ is a vector space over $\langle F, +, \cdot \rangle$.

11.2.2 Examples of vector spaces

Example 11.59. Let $\langle F, +, \cdot \rangle$ be a field, e a element and

1. $\oplus: \{e\} \times \{e\} \rightarrow \{e\}$ defined by $e \oplus e = e$
2. $\odot: F \times \{e\} \rightarrow \{e\}$ defined by $\alpha \odot e = e$

then $\langle \{e\}, \oplus, \odot \rangle$ is a vector space over $\langle F, +, \cdot \rangle$ with neutral element e and the inverse of e is e. This vector space is called the **trivial vector space**.

Proof. First we prove that $\langle \{e\}, \oplus \rangle$ is a Abelian group:

associativity.

$$\forall x, y, z \in \{e\} \text{ we have } x \oplus (y \oplus z) = e \oplus (e \oplus e) = e \oplus e = e = e \oplus e = (e \oplus e) \oplus e = (x \oplus y) \oplus z$$

neutral element. $\forall x \in \{e\}$ we have $x = e$ so that $x \oplus e = e \oplus e = e = e \oplus e = e \oplus x$ proving that e is the neutral element.

inverse element. $\forall x \in \{e\}$ we have x=e so that $x \oplus e = e \oplus e = e$ proving that e is the inverse element of x.

commutativity. $\forall x, y \in \{e\}$ we have $x = e = y$ so that $x \oplus y = e \oplus e = e = e \oplus e = y \oplus x$

For the remaining axioms we have

1. $\forall \alpha \in F$ and $x, y \in \{e\}$ we have as $x = e = y$ that
$$\alpha \odot (x + y) = \alpha \odot (e \oplus e) = \alpha \odot e = e = e \oplus e = \alpha \odot e \oplus \alpha \odot e = \alpha \odot x \oplus \alpha \odot y$$
2. $\forall \alpha, \beta \in F$ and $x \in \{e\}$ we have as $x = e$ that
$$(\alpha + \beta) \odot x = (\alpha + b) \odot e = e = e \oplus e = \alpha \odot e + \beta \odot e = \alpha \odot x + \beta \odot x$$
3. $\forall \alpha, \beta \in F$ and $x \in \{e\}$ we have as $x = e$ that
$$(\alpha \cdot \beta) \odot x = (\alpha \cdot \beta) \odot e = e = \alpha \odot e = \alpha \odot (\beta \odot e)$$
4. $\forall x \in \{e\}$ we have $1 \odot x = 1 \odot e = e = x$ □

Every field is a vector spaces over itself.

Theorem 11.60. *If $\langle F, +, \cdot \rangle$ is a field then $\langle F, +, \cdot \rangle$ is a vector space over itself*

Proof. As $\langle F, +, \cdot \rangle$ is a field we have by definition of a field [see definition: 4.51] that $\langle F, + \rangle$ is a Abelian group. Further for the rest of the vector axioms we have:

1. $\forall \alpha, x, y \in F$ we have $\alpha \cdot (x + y) \underset{\text{distributivity}}{=} \alpha \cdot x + \alpha \cdot y$
2. $\forall \alpha, \beta, x \in F$ we have $(\alpha + \beta) \cdot x \underset{\text{distributivity}}{=} \alpha \cdot x + \beta \cdot y$

3. $\forall \alpha, \beta, x \in F$ we have $(\alpha \cdot \beta) \cdot x \underset{\text{associativity}}{=} \alpha \cdot (\beta \cdot x)$
4. $\forall x \in F$ we have for the multiplicative neutral element of $\langle F, +, \cdot \rangle$ that $1 \cdot x = x$ □

Using the above and the fact that $\langle \mathbb{Q}, +, \cdot \rangle$, $\langle \mathbb{R}, +, \cdot \rangle$ and $\langle \mathbb{C}, +, \cdot \rangle$ are fields [see theorem: 10.5] we have:

Example 11.61. $\langle \mathbb{Q}, +, \cdot \rangle$ is a vector space over $\langle \mathbb{Q}, +, \cdot \rangle$, $\langle \mathbb{R}, +, \cdot \rangle$ is a vector space over $\langle \mathbb{R}, +, \cdot \rangle$ and $\langle \mathbb{C}, +, \cdot \rangle$ is a vector space over $\langle \mathbb{C}, +, \cdot \rangle$.

Be aware that a vector space depends also on the field used, so for example $\mathbb{C}$ can be used to define another vector space, as is shown in the next example.

Example 11.62. $\langle \mathbb{C}, +, \cdot \rangle$ is a vector space over $\langle \mathbb{R}, +, \cdot \rangle$

Proof. As $\langle \mathbb{C}, +, \cdot \rangle$ is a field we have by definition of a field [see definition: 4.51] that $\langle \mathbb{C}, + \rangle$ is a Abelian group. Further for the rest of the vector axioms we have as $\mathbb{R} \subseteq \mathbb{C}$ that

1. $\forall \alpha \in \mathbb{R}$, $\forall x, y \in \mathbb{C}$ we have $\alpha \cdot (x + y) \underset{\text{distributivity}}{=} \alpha \cdot x + \alpha \cdot y$
2. $\forall \alpha, \beta \in \mathbb{R}$, $\forall x \in \mathbb{C}$ we have $(\alpha + \beta) \cdot x = \alpha \cdot x + \beta \cdot x$
3. $\forall \alpha, \beta \in \mathbb{R}$, $\forall x \in \mathbb{C}$ we have $(\alpha \cdot \beta) \cdot x = \alpha \cdot (\beta \cdot x)$
4. As $1 \in \mathbb{R} \subseteq \mathbb{C}$ we have $\forall x \in \mathbb{C}$ $1 \cdot x$ □

Definition 11.63. *A vector space $\langle V, \oplus, \odot \rangle$ over $\langle \mathbb{R}, +, \cdot \rangle$ is called a **real** vector space and a vector space $\langle V, \oplus, \odot \rangle$ over $\langle \mathbb{C}, +, \cdot \rangle$ is called a **complex** vector space. If we want to refer to a vector space $\langle V, \oplus, \odot \rangle$ over either $\langle \mathbb{R}, +, \cdot \rangle$ or $\langle \mathbb{R}, +, \cdot \rangle$ then we say that $\langle V, \oplus, \odot \rangle$ is a vector space over $\langle \mathbb{K}, +, \cdot \rangle$.*

Next we use an existing vector space and pairwise addition and scalar multiplication to define a function space. Later will use this to define more special function spaces that are sub-spaces of this function space.

Theorem 11.64. (function space) *Let $\langle V, \oplus, \odot \rangle$ be a vector space over $\langle F, +, \cdot \rangle$, X a set then and define for $V^X = \{f | f: X \to V$ is a function$\}$ the following operations:*

$$\boxplus: V^X \times V^X \to V^X \textit{ defined by } f \boxplus g \text{ where } (f \boxplus g)(x) = f(x) \oplus g(y) \textit{ [pairwise addition]}$$

$$\boxdot: F \times V^X \to V^X \textit{ defined by } \alpha \boxdot f \textit{ where } (\alpha \boxdot f)(x) = \alpha \odot f(x)$$

then

$$\langle V^X, \boxplus, \boxdot \rangle \textit{ is a vector space over } \langle F, +, \cdot \rangle$$

where:

a) $C_o: X \to V$ is defined by $C_0(x)0e$ [the constant function [see example: 2.45].

b) If $f \in V^X$ then $-f$ defined by $(-f)(x) = -f(x)$ is the inverse element of f.

Proof. First we prove that $\langle V^X, \boxplus, \boxdot \rangle$ is a Abelian group:

associativity. Let $f, g, h \in V^X$ then $\forall x \in X$ we have

$$\begin{aligned}
(f \boxplus (g \boxplus h))(x) &= f(x) \oplus (g \boxplus h)(x) \\
&= f(x) \oplus (g(x) \oplus h(x)) \\
&= (f(x) \oplus g(x)) \oplus h(x) \\
&= (f \boxplus g)(x) \oplus h(x) \\
&= ((f \boxplus g) \boxplus h)(x)
\end{aligned}$$

proving that $f \boxplus (g \boxplus h) = (f \boxplus g) \boxplus h$

commutativity. Let $f, g \in V^X$ then $\forall x \in X$ we have

$$\begin{aligned}
(f \boxplus g)(x) &= f(x) \oplus g(x) \\
&= g(x) \oplus f(x) \\
&= (g \boxplus f)(x)
\end{aligned}$$

so that

$$f \boxplus g = g \boxplus f$$

neutral element. Let $f \in V^X$ then $\forall x \in X$ we have

$$(f \boxplus C_0)(x) = f(x) \oplus C_0(x) = f(x) \oplus 0 = f(x)$$

so that

$$C_0 \boxplus f \underset{\text{commutativity}}{=} f \boxplus C_0 = f$$

inverse element. Let $f \in V^X$ then $\forall x \in X$ we have

$$(f \boxplus (-f))(x) = f(x) \oplus (-f)(x) = f(x) \oplus (-(f(x))) = 0 = C_9(x)$$

so that $(-f) + f \underset{\text{commutativity}}{=} f + (-f) = C_0$

For the remaining axioms of a vector space we have:

1. If $\alpha \in F$ and $f, g \in^X$ then $\forall x \in X$ we have

$$\begin{aligned}(\alpha \boxdot (f \boxplus g))(x) &= \alpha \odot (f \boxplus g)(x) \\ &= \alpha \odot (f(x) \oplus g(x)) \\ &= \alpha \odot f(x) \oplus \alpha \odot g(x) \\ &= (\alpha \boxdot f)(x) \oplus (\alpha \boxdot g)(x) \\ &= (\alpha \boxdot f \boxplus \alpha \boxdot g)(x)\end{aligned}$$

   proving that $\alpha \boxdot (f \boxplus g) = \alpha \boxdot f \boxplus \alpha \boxdot g$.

2. If $\alpha, \beta \in F$ and $f \in V^X$ then $\forall x \in X$ we have

$$\begin{aligned}((\alpha + \beta) \boxdot f)(x) &= (\alpha + \beta) \odot f(x) \\ &= \alpha \odot f(x) + \beta \odot f(x) \\ &= (\alpha \boxdot f)(x) \oplus (\beta \boxdot f)(x) \\ &= (\alpha \boxdot f \boxplus \beta \boxdot f)(x)\end{aligned}$$

   so that $(\alpha + \beta) \boxdot f = \alpha \boxdot f \boxplus \beta \boxdot f$.

3. If $\alpha, \beta \in F$ and $f \in V^X$ then $\forall x \in X$ we have

$$\begin{aligned}((\alpha \cdot \beta) \boxdot f)(x) &= (\alpha \cdot \beta) \odot f(x) \\ &= \alpha \cdot (\beta \cdot f(x)) \\ &= \alpha \cdot (\beta \boxdot f)(x) \\ &= (\alpha \boxdot (\beta \boxdot f))(x)\end{aligned}$$

   proving that $(\alpha \cdot \beta) \boxdot f = \alpha \boxdot (\beta \boxdot f)$

4. Let $f \in V^X$ then we have $\forall x \in X$ that $(1 \boxdot f)(x) = 1 \odot f(x) = f(x)$ so that $1 \boxdot f = f$. □

Up to now we have used different operator symbols for addition and multiplication in the different structures. To simplify notation we start using the same symbol $+$ for addition in the vector space and field and the same symbol $\cdot$ for the scalar product and multiplication in the field. We also use 0 for the additive neutral element in the vector space and the field and note the inverse of x in the vector space and the field as $-x$.

Referring to the power of a set [see definition: 6.81] we can construct a vector space on the power of a vector space.

Theorem 11.65. *Let $n \in \mathbb{N}$, $\langle V, +, \cdot \rangle$ be a vector space over $\langle F, +, \cdot \rangle$ then $\langle V^n, +, \cdot \rangle$ is a vector space over $\langle F, +, \cdot \rangle$ where*

$$+ \colon V^n \times V^n \to V^n \text{ is defined by } (x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n)$$

$$\cdot \colon F \times V^n \to V^n \text{ is defined by } \alpha \cdot (x_1, \ldots, x_n) = (\alpha \cdot x_1, \ldots, \alpha \cdot x_n)$$

$$0 = \underbrace{(0, \ldots, 0)}_{n} \in V^n \text{ is the additive neutral element in } V^n$$

$$\forall (x_1, \ldots, x_n) \in V^n \text{ the additive negative is } (-x_1, \ldots, -x_n)$$

Proof. Note that using [theorem: 6.82] we have that

$$V^n = V^{\{1,\ldots,n\}} = \{f \mid f\colon \{1,\ldots,n\} \rightarrow V \text{ is a function}\}$$

and $x = (x_1, \ldots, x_n)$ is equivalent with $x\colon \{1, \ldots, n\} \rightarrow V$ is a function where $\forall i \in \{1, \ldots, n\}$ we have $x(i) = x_i$. So

$$+\colon V^n \times V^n \rightarrow V^n \text{ is defined by } (x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n)$$

is equivalent with:

$$+\colon V^n \times V^n \rightarrow V^n \text{ is defined by } (x + y) \text{ where } (x + y)(i) = x(i) + y(i)$$

and

$$\cdot\colon F \times V^n \rightarrow V^n \text{ is defined by } \alpha \cdot (x_1, \ldots, x_n) = (\alpha \cdot x_1, \ldots, \alpha \cdot x_n)$$

is equivalent with:

$$\cdot\colon F \times V^n \rightarrow V^n \text{ is defined by } \alpha \cdot x \text{ where } (\alpha \cdot x)(i) = \alpha \cdot x(i)$$

and

$$(0, \ldots, 0) \in V^n \text{ is } C_0 \text{ where } C_0(i) = 0$$

and finally

$$\forall (x_1, \ldots, n) \in V^n \text{ the additive negative is } (-x_1, \ldots, -x_n)$$

is equivalent with

$$\forall x \in V^n \text{ the additive negative is } -x \text{ where } (-x)(i) = -x(i)$$

Combining these equivalent definitions with [theorem: 11.64] proves then that $\langle V^n, +, \cdot\rangle$ is a vector space. □

As a application of the above theorem we have that

Corollary 11.66. *Let $n \in \mathbb{N}$ and $\langle F, +, \cdot\rangle$ a field then $\langle F^n \dot{+}, \cdot\rangle$ is a vector space over $\langle F, +, \cdot\rangle$*

Proof. This follows from [theorems: 11.60, 11.65] □

Using [examples: 11.61, 11.62] and the above theorem [theorem: 11.65] we have then the following examples of vector spaces:

Example 11.67. Let $n \in \mathbb{N}$ then

1. $\langle \mathbb{Q}^n, +, \cdot\rangle$ is a vector space over $\langle Q, +, \cdot\rangle$
2. $\langle \mathbb{R}^n, +, \cdot\rangle$ is a vector space over $\langle \mathbb{R}, +, \cdot\rangle$
3. $\langle \mathbb{C}, +, \cdot\rangle$ is a vector space over $\langle \mathbb{C}, +, \cdot\rangle$
4. $\langle \mathbb{C}, +, \cdot\rangle$ is a vector space over $\langle \mathbb{R}, +, \cdot\rangle$

Remark 11.68. Note that in the proof of the previous theorem we use the fact that $V^n = \{f \mid f\colon \{1,\ldots,n\} \rightarrow V \text{ is a function}\}$, however this is not standard practice. Most books prefers to work with the notation $x \in V^n \Leftrightarrow x = (x_1, \ldots, x_n)$ such that $x_i \in V$ $\forall i \in \{1, \ldots, n\}$. Likewise for $\prod_{i \in \{1,\ldots,n\}} V_i$ most books use the notation $x \in \prod_{i \in \{1,\ldots,n\}} V_i \Leftrightarrow x = (x_1, \ldots, x_n)$ such that $x_i \in V_i$ $\forall i \in \{1, \ldots, n\}$. This a standard that we will follow for the rest of this book. If needed you can use [definition: 6.74] and [theorem: 6.82] to fallback to the original definitions.

Definition 11.69. *Let $\langle V, +, \cdot\rangle$ be a vector space over a field $\langle F, +, \cdot\rangle$, $A, B \subseteq V$, $G \subseteq F, x \in V$ and $\alpha \in F$ then we have the following definitions:*

1. $x + A \underset{\text{def}}{=} \{x + y \mid y \in A\} \underset{\text{commutativity}}{=} \{y + x \mid y \in A\} \underset{\text{def}}{=} A + x$
2. $A + B = \{x + y \mid x \in A \wedge y \in B\}$
3. $\alpha \cdot A = \{\alpha \cdot x \mid x \in A\}$
4. $G \cdot A = \{\gamma \cdot x \mid \gamma \in F \wedge x \in A\}$
5. $A - B = \{x - y \mid x \in A \wedge y \in B\}$

Theorem 11.70. *If $\langle V, +, \cdot\rangle$ is a vector space over $\langle F, +, \cdot\rangle$ 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. $$\begin{aligned} y \in x + \left(A \bigcup B\right) &\underset{(1)}{\Leftrightarrow} y - x \in A \bigcup B \\ &\Leftrightarrow (y - x \in A) \vee (y - x \in B) \\ &\Leftrightarrow (y \in x + A) \vee (y \in x + B) \\ &\Leftrightarrow y \in (x + A) \bigcup (x + B) \end{aligned}$$
3. $$\begin{aligned} y \in x + \left(A \bigcap B\right) &\underset{(1)}{\Leftrightarrow} y - x \in A \bigcap B \\ &\Leftrightarrow (y - x \in A) \wedge (y - x \in B) \\ &\Leftrightarrow (y \in x + A) \wedge (y \in x + B) \\ &\Leftrightarrow y \in (x + A) \bigcap (x + B) \end{aligned}$$
4. $$\begin{aligned} z \in (x + (y + A)) &\underset{(1)}{\Leftrightarrow} z - x \in y + A \\ &\underset{(1)}{\Leftrightarrow} (z - x) - y \in A \\ &\Leftrightarrow (z - (x + y)) \in A \\ &\Leftrightarrow z \in (x + y) + A \end{aligned}$$

$\square$

Theorem 11.71. *If* $\langle V_i, +_i, \cdot_i \rangle_{i \in I}$ *is a family of vector spaces over a field* $\langle F, +, \cdot \rangle$ *then if we define:*

1. $+ : \prod_{i \in I} V_i \times \prod_{i \in I} V_i \rightarrow \prod_{i \in I} V_i$ *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 theorem 4.26]*
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 [theorem 4.26] 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 vector space that $\alpha \cdot_i x(i) \in V_i$ so 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
$$\begin{aligned} (\alpha \cdot (x + y))(i) &= \alpha \cdot_i (x + y)(i) \\ &= \alpha \cdot_i (x(i) +_i y(i)) \\ &\underset{\langle V_i, +_i, \cdot_i \rangle \text{ is a vector space}}{=} \alpha \cdot_i x(i) +_i \alpha \cdot_i y(i) \\ &= (\alpha \cdot x)(i) +_i (\alpha \cdot y)(i) \\ &= (\alpha \cdot x + \alpha \cdot y)(i) \end{aligned}$$
so that $\alpha \cdot (x + y) = \alpha \cdot x + \alpha \cdot y$
2. If $\alpha, \beta \in F$ and $x \in \prod_{i \in I} V_i$ then $\forall i \in I$ we have
$$\begin{aligned} ((\alpha + \beta) \cdot x)(i) &= (\alpha + \beta) \cdot_i x(i) \\ &\underset{\langle V_i, +_i, \cdot_i \rangle \text{ is a vector space}}{=} \alpha \cdot_i x(i) +_i \beta \cdot_i x(i) \\ &= (\alpha \cdot x)(i) +_i (\beta \cdot x)(i) \\ &= (\alpha \cdot x + \beta \cdot y)(i) \end{aligned}$$

so that $(\alpha+\beta)\cdot x=\alpha\cdot x+\beta\cdot x$.

3. If $\alpha,\beta\in F$ and $x\in\prod_{i\in I}V_i$ then $\forall i\in I$ we have

$$\begin{array}{lcl}((\alpha\cdot\beta)\cdot x)(i) & = & (\alpha\cdot\beta)\cdot_i x(i)\\ & \underset{\langle V_i,+_i,\cdot_i\rangle \text{ is a vector space}}{=} & \alpha\cdot_i(\beta\cdot_i x(i))\\ & = & \alpha\cdot_i(\beta\cdot x)(i)\\ & = & (\alpha\cdot(\beta\cdot x))(i)\\ & = & (\alpha\cdot(\beta\cdot x))(i)\end{array}$$

   So that $(\alpha\cdot\beta)\cdot x=\alpha\cdot(\beta\cdot x)$.

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 $1\cdot x=x$ □

11.2.3 Factor spaces of a vector space

A factor space allows us to threat a sub-space of a vector space as a neutral elment in higher level of vector space. To create a factor space we first define a equivalence relation, the factor space is then the set of equivalence classes. We add then operators to the factor space based on the operators for addition and scalar product of the original vector space.

Theorem 11.72. *Let $\langle X,+,\cdot\rangle$ be a vector space over a field $\langle F,+,\cdot\rangle$ and Y a sub-space of X then $\sim_Y$ defined by $\sim_Y=\{(x,y)\in X\times X|x-y\in Y\}$ is a equivalence relation.*

Proof.

reflexitivity. As Y is a sub-space of x we have that $0\in Y$ [see theorem: 11.57], hence if $x\in X$ we have that $x-x=0\in Y$ proving that $\forall x\in X$ $x\sim_Y x$.

symmetry. If $x\sim_Y y$ then $x-y\in Y$, so as Y is a vector space [as a sub-space of X] we have that $y-x=-(x-y)\in Y$, hence $y\sim_Y x$.

transitivity. If $x\sim_Y y$ and $y\sim_Y z$ then $x-y\in Y$ and $y-z$ so that $x-z=(x-y)+(y-z)\in Y$ [as Y is a vector space], hence $x\sim_Y z$. □

We form now the set of equivalence classes and endow it with a addtion and scallar product operator based on the operators of the original vector space.

Theorem 11.73. *Let X be a vector space over a field F and Y a subspace of X then we use [definition: 3.16] to define X/Y as follows:*

$$X/Y=X/\sim_Y\underset{\text{[definition: 3.16]}}{=}\{\sim_Y[x]|x\in X\}$$

where

$$\sim_Y[x]=\{y\in X|y\sim_Y x\}=\{y\in X|y-x\in Y\}\underset{y-x\in Y\Leftrightarrow y\in x+Y}{=}x+Y$$

Using this definition we have that the following are well defined functions

$$+: X/Y\times X/Y\rightarrow X/Y \text{ defined by } \sim_Y[x]+\sim_Y[y]=\sim[x+y]$$

$$\cdot: F\times X/Y\rightarrow X/Y \text{ defined by } \alpha\cdot\sim[x]=\sim[\alpha\cdot x]$$

Using the above operators we have that

$$\langle X/Y,+,\cdot\rangle \text{ is a vector space over } \langle F,+,\cdot\rangle$$

where the neutral element is $\sim[0]=Y$ and for every $\sim[x]$ $\sim[-x]$ is the inverse element

Proof. First we must prove that $+$ and $\cdot$ are well defined.

1. Assume that $\sim_Y[x]=\sim_Y[x']$ and $\sim_Y[y]=\sim_Y[y']$ then we have by [theorem: 3.11 that $x\sim_Y x'$ and $y\sim_Y y'$. Hence $x-x'\in Y$ and $y-y'\in Y$ so that $(x+y)-(x'+y')=(x-x')+(y-y')\in Y$ or $x+y\sim_Y x'+y'$. Using then [theorem: 3.11] it follows that $\sim_Y[x+y]=\sim_Y[x'+y']$. So

$$+: X/Y\times X/Y\rightarrow X/Y \text{ defined by } \sim_Y[x]+\sim_Y[y]=\sim[x+y] \text{ is a function}$$

2. Let $\alpha \in F$. Assume that $\sim_Y[x] = \sim_Y[x']$ then we have by [theorem: 3.11 that $x \sim_Y x'$ or $x - x' \in Y$. Hence $\alpha \cdot x - \alpha \cdot x' = \alpha \cdot (x - x') \in Y$ so that $\sim[\alpha \cdot x] = \sim_Y[\alpha \cdot x']$. Using then [theorem: 3.11] it follows that $\sim_Y[\alpha \cdot x] = \sim_Y[\alpha \cdot x']$. So

$$\cdot: F \times X/Y \rightarrow X/Y \text{ defined by } \alpha \cdot \sim[x] = \sim[\alpha \cdot x] \text{ is a function}$$

Next note that $x \in \sim_Y[0] \Leftrightarrow x - 0 \in Y \Leftrightarrow x \in Y$ which proves that

$$\sim_Y[0] = Y$$

Next we have to prove the vector space axioms:

1. $\langle X/Y, +\rangle$ is a Abelian group:

   associativity. Let $\sim_Y[x]$, $\sim_Y[y]$, $\sim_Y[z] \in X/Y$ then

$$\begin{aligned}\sim_Y[x] + (\sim_Y[y] + \sim_Y[z]) &= \sim_Y[x] + \sim_Y[y+z] \\ &= \sim_Y[x + (y+z)] \\ &= \sim_Y[(x+y)+z] \\ &= \sim_Y[x+y] + \sim_Y[z] \\ &= (\sim_Y[x] + \sim_Y[y]) + \sim_Y[z]\end{aligned}$$

   commutatitivity. Let $\sim_Y[x]$, $\sim_Y[y] \in X/Y$ then

$$\sim_Y[x] + \sim_Y[y] = \sim_Y[x+y] = \sim_Y[y+x] = \sim_Y[y] + \sim_Y[z]$$

   neutral element. Let $\sim_Y[x] \in X/Y$ then

$$\sim_Y[x] + \sim_Y[0] \underset{\text{commutatitivity}}{=} \sim_Y[0] + \sim_Y[x] = \sim_Y[0+x] = \sim_Y[x]$$

   inverse element. Let $\sim_Y[x] \in X/Y$ then

$$\sim_Y[x] + \sim_Y[-x] \underset{\text{commutatitivity}}{=} \sim_Y[-x] + \sim_Y[x] = \sim_Y[(-x)+x] = \sim_Y[0]$$

2. Scalar product axioms

   a. Let $\alpha \in F$ and $\sim_Y[x]$, $\sim_Y[y] \in X/Y$ then

$$\begin{aligned}\alpha \cdot (\sim_Y[x] + \sim_Y[y]) &= \alpha \cdot \sim_Y[x+y] \\ &= \sim_Y[\alpha \cdot (x+y)] \\ &= \sim_Y[\alpha \cdot x + \alpha \cdot y] \\ &= \sim_Y[\alpha \cdot x] + \sim_Y[\alpha \cdot y] \\ &= \alpha \cdot \sim_Y[x] + \alpha \cdot \sim_Y[y]\end{aligned}$$

   b. Let $\alpha, \beta \in F$ and $\sim[x] \in X/Y$ then

$$\begin{aligned}(\alpha + \beta) \cdot \sim_Y[x] &= \sim_Y[(\alpha+\beta) \cdot x] \\ &= \sim_Y[\alpha \cdot x + \beta \cdot x] \\ &= \sim_Y[\alpha \cdot x] + \sim_Y[\beta \cdot x] \\ &= \alpha \cdot \sim_Y[x] + \beta \cdot \sim_Y[x]\end{aligned}$$

   c. Let $\alpha, \beta \in F$ and $\sim[x] \in X/Y$ then

$$\begin{aligned}(\alpha \cdot \beta) \cdot \sim_Y[x] &= \sim_Y[(\alpha \cdot \beta) \cdot x] \\ &= \sim_Y[\alpha \cdot (\beta \cdot x)] \\ &= \alpha \cdot \sim_Y[\beta \cdot x] \\ &= \alpha \cdot (\beta \cdot \sim_Y[x])\end{aligned}$$

   d. Let $\sim_Y[x] \in X/Y$ then

$$1 \cdot \sim_Y[x] = \sim_Y[1 \cdot x] = \sim_Y[x]$$

□

Theorem 11.74. (Projection Function) *Let X be a vector space over a field F and Y a subspace of X then the* ***projection function*** *is defined by*

$$\pi_Y: X \rightarrow X/Y \text{ where } \pi_Y(x) = \sim_Y[x]$$

This function has the following properties:

1. $\pi_Y \in \text{Hom}(X, X:Y)$ *[or π_Y is linear]*
2. $\pi_Y(X) = X \setminus Y$ *[surjectivity]*
3. $\forall x \in Y$ we have $\pi_Y(x) = \sim_Y[0] = 0$ *the neutral element of X/Y*

Proof.

1. Let $x, y \in X$ and $\alpha \in F$ then we have

$$\pi_Y(x+y) = \sim_Y[x+y] = \sim_Y[x] + \sim_Y[y] = \pi_Y(x) + \pi_Y(y)$$

and

$$\pi_Y[\alpha \cdot x] = \sim_Y[\alpha \cdot x] = \alpha \cdot \sim_Y[x] = \alpha \cdot \pi_Y(x)$$

2. $\pi_Y(X) = \{\pi_Y(x) | x \in X\} = \{\sim_Y[x] | x \in X\} = X \setminus Y$
3. If $x \in Y$ then we have $x - 0 = x \in Y$ so that $x \sim_Y 0$, hence by [theorem: 3.11] we have that $\sim_Y[x] = \sim_Y[0]$. Hence

$$\forall x \in Y \text{ we } \quad \pi_Y(x) = \sim[0] = 0$$

□

11.3 Basis of a vector space

11.3.1 Finite sums on a vector space

As a vector space is also a Abelian group we can talk about finite sums of vectors in the vector space.

Theorem 11.75. *If $\langle V, +, \cdot \rangle$ is a vector space over a field $\langle F, +, \cdot \rangle$ then we have:*

1. *If $\alpha \in F$, $n, m \in \mathbb{N}_0$ with $n \leqslant m$ and $\{x_i\}_{i \in \{n, \ldots, m\}} \subseteq V$ then $\sum_{i=n}^{m} \alpha \cdot x_i = \alpha \cdot \sum_{=n}^{m} x_i$*
2. *If $\alpha \in F$, I a finite set, $\{x_i\}_{i \in I} \subseteq V$ then $\sum_{i \in I} \alpha \cdot x_i = \alpha \cdot \sum_{i \in I} x_i$*

Proof.

1. We use induction, so take:

$$S = \left\{ n \in \mathbb{N}_0 | \forall \{x_i\}_{i \in \{0, \ldots, n\}} \text{ we have } \sum_{i=0}^{n} \alpha \cdot x_i = \alpha \cdot \sum_{i=0}^{n} x_i \right\}$$

then we have:

$0 \in S$. If $\{x_i\}_{i \in \{0\}}$ then $\sum_{i=0}^{0} \alpha \cdot x_i = \alpha \cdot x_0 = \alpha \cdot \sum_{i=0}^{0} x_i$ proving that $0 \in S$

$n \in S \Rightarrow n+1 \in S$. If $\{x_i\}_{i \in \{0, \ldots, n+1\}} \subseteq V$ then

$$\begin{aligned} \sum_{i=0}^{n+1} \alpha \cdot x_i &= \left(\sum_{i=0}^{n} \alpha \cdot x_i \right) + \alpha \cdot x_{n+1} \\ &\underset{n \in S}{=} \alpha \cdot \left(\sum_{i=0}^{n} x_i \right) + \alpha \cdot x_{n+1} \\ &= \alpha \cdot \left(\left(\sum_{i=0}^{n} x_i \right) + x_{n+1} \right) \\ &= \alpha \cdot \sum_{i=0}^{n+1} x_i \end{aligned}$$

proving that $n+1 \in S$.

So by mathematical induction

$$\forall \{x_i\}_{i\in\{0,\ldots,n\}} V \text{ we have } \sum_{i=0}^{n} \alpha\cdot x = \alpha\cdot\sum_{i=0}^{n} x_i \tag{11.12}$$

Let now $n,m\in\mathbb{N}_0$ with $n\leqslant m$ then we have

$$\sum_{i=n}^{m} \alpha\cdot x_i = \sum_{i=0}^{m-n} \alpha\cdot x_{n+i} \underset{\text{[eq: 11.12]}}{=} \alpha\cdot\sum_{i=0}^{m-n} x_{n+i} = \alpha\cdot\sum_{i=n}^{m} x_i$$

2. If I is finite and $\{x_i\}_{i\in I}$ with $\forall i\in I\ x_i = x$ then we have either:

   ***I* = ∅.** Then

$$\sum_{i\in I} \alpha\cdot x_i = 0 = \alpha\cdot 0 = \alpha\cdot\sum_{i\in I} x_i$$

   ***I* ≠ ∅.** Then there exist a $n\in\mathbb{N}$ and a bijection $\beta:\{0,\ldots,n-1\}\rightarrow I$ such that

$$\sum_{i\in I} \alpha\cdot x_i = \sum_{i=0}^{n-1} \alpha\cdot x_{\beta(i)} \underset{\text{[eq: 11.12]}}{=} \alpha\cdot\sum_{i=0}^{n-1} x_{\beta(i)} = \alpha\cdot\sum_{i\in I} x_i$$

□

Theorem 11.76. *If $\langle V,+,\cdot\rangle$ is a vector space over a field $\langle F,+,\cdot\rangle$ and $x\in V$ then we have:*

1. *If $n,m\in\mathbb{N}_0$ with $n\leqslant m$ $\{\alpha_i\}_{i\in\{n,\ldots m\}}\subseteq F$ then*

$$\sum_{i=n}^{m} \alpha_i\cdot x = \left(\sum_{i=n}^{m} \alpha_i\right)\cdot x$$

2. *If I is a finite set, $\{\alpha\}_{i\in I}\subseteq F$ then*

$$\sum_{i\in I} \alpha_i\cdot x = \left(\sum_{i\in I} \alpha_i\right)\cdot x$$

Proof.

1. We will use induction in this proof, so let

$$S = \left\{ n\in\mathbb{N}_0 | \forall\{\alpha_i\}_{i\in\{0,\ldots,n\}}\subseteq F \text{ we have } \sum_{i=0}^{n} (\alpha_i\cdot x) = \left(\sum_{i=0}^{n} \alpha_i\right)\cdot x \right\}$$

   then we have:

   0 ∈ *S*. If $\{\alpha_i\}_{i\in\{0\}}$ then we have $\sum_{i=0}^{0} (\alpha_i\cdot x) = \alpha_0\cdot x = (\sum_{i=0}^{0} \alpha_i)\cdot x$ proving that $0\in S$.

   ***n* ∈ *S* ⇒ *n* + 1 ∈ *S*.** Let $\{\alpha_i\}_{i\in\{0,\ldots,n+1\}}\subseteq F$ then we have

$$\begin{aligned}
\sum_{i=0}^{n+1} (\alpha_i\cdot x) &= \left(\sum_{i=0}^{n} (\alpha_i\cdot x)\right) + \alpha_{n+1}\cdot x \\
&\underset{n\in S}{=} \left(\sum_{i=0}^{n} \alpha_i\right)\cdot x + \alpha_{n+1}\cdot x \\
&= \left(\left(\sum_{i=0}^{n} \alpha_i\right) + \alpha_{n+1}\right)\cdot x \\
&= \left(\sum_{i=0}^{n+1} \alpha_i\right)\cdot x
\end{aligned}$$

   proving that $n+1\in S$.

   By mathematical induction it follows then that

$$\forall n\in\mathbb{N}_0 \text{ and } \{\alpha_i\}_{i\in\{0,\ldots,n\}} \text{ we have } \sum_{i=0}^{n} \alpha_i\cdot x = \left(\sum_{i=0}^{n} \alpha_i\right)\cdot x \tag{11.13}$$

Let now $n,m\in\mathbb{N}_0$ with $n\leqslant m$ then for $\{\alpha_i\}_{i\in\{n,\ldots,m\}}\subseteq F$ then we have

$$\sum_{i=n}^{m}\alpha_i\cdot x=\sum_{i=0}^{m-n}\alpha_{n+i}\cdot x\underset{[\text{eq: }11.13]}{=}\left(\sum_{i=0}^{m-n}\alpha_i\right)\cdot x=\left(\sum_{i=n}^{m}\alpha_i\right)\cdot x$$

2. I is a finite set and $\{\alpha_i\}_{i\in I}\subseteq F$ then we have either:

   $I=\varnothing$. Then

   $$\sum_{i\in I}\alpha_i\cdot x=0=0\cdot x=\left(\sum_{i\in I}\alpha_i\right)\cdot x$$

   $I\neq\varnothing$. Then there exists a $n\in\mathbb{N}$ and a bijection $\beta:\{0,\ldots,n-1\}\rightarrow I$ such that

   $$\sum_{i\in I}\alpha_i\cdot x=\sum_{i=0}^{n-1}\alpha_{\beta(i)}\cdot x\underset{[\text{eq: }11.13]}{=}\left(\sum_{i=0}^{n-1}\alpha_i\right)\cdot x=\left(\sum_{i\in I}\alpha_i\right)\cdot x$$

□

Theorem 11.77. *If $\langle V,+,\cdot\rangle$ is a vector space over $\langle\mathbb{K},+,\cdot\rangle$ where $\mathbb{K}=\mathbb{C}$ or $\mathbb{R}$ then we have:*

1. *If $n,m\in\mathbb{N}_0$ with $n\leqslant m$ and $\{x_i\}_{i\in\{n,\ldots,m\}}\subseteq V$ is such that $\forall i\in\{n,\ldots,.m\}$ $x_i=x$ then*

   $$\sum_{i=n}^{m}x_i=(m-n+1)\cdot x$$

2. *If I is a finite set and $\{x_i\}_{i\in I}\subseteq V$ is such that $\forall i\in I$ $x_i=x$ then*

   $$\sum_{i\in I}x_i=\operatorname{card}(I)\cdot x$$

Proof.

1. We use induction for the proof, so let

   $$S=\left\{n\in\mathbb{N}_0|\forall\{x_i\}_{i\in\{0,\ldots,n\}}\subseteq\mathbb{K}\text{ with }\forall i\in\{0,\ldots,n\}\,x_i=x\text{ we have }\sum_{i=0}^{n}x_i=(n+1)\cdot x\right\}$$

   then we have:

   $0\in S$. If $\{x_i\}_{i\in\{0\}}\subseteq V$ is such that $x_0=x$ then $\sum_{i=0}^{0}x_i=x_0=x=1\cdot x=(0+1)\cdot x$ proving that $0\in S$.

   $n\in S\Rightarrow n+1\in S$. If $\{x_i\}_{i\in\{0,\ldots,n+1\}}\subseteq V$ is such that $\forall i\in\{0,\ldots,n+1\}$ $x_i=x$ then we have

   $$\sum_{i=0}^{n+1}x_i=\left(\sum_{i=0}^{n}x_i\right)+x_{n+1}=\left(\sum_{i=0}^{n}x_i\right)+x\underset{n\in S}{=}n\cdot x+x=(n+1)\cdot x$$

   proving that $n+1\in S$.

   By mathematical induction we have then that

   $$\forall\{x_i\}_{i\in\{0,\ldots,n\}}\subseteq V\text{ with }\forall i\in\{0,\ldots,n\}\,x_i=x\text{ we have }\sum_{i=0}^{n}x_i=(n+1)\cdot x\tag{11.14}$$

   If now $n,m\in\mathbb{N}_0$ with $n\leqslant m$ then

   $$\sum_{i=n}^{m}x_i=\sum_{i=0}^{m-n}x_{n+i}\underset{[\text{eq: }11.14]}{=}(m-n)+1\cdot x$$

2. For I is finite and $\{x_i\}_{i\in I}$ with $\forall i\in I$ $x_i=x$ then we have either:

   $I=\varnothing$. Then

   $$\sum_{i=I}x_i=0=0\cdot x=\operatorname{card}(I)\cdot x$$

   $I\neq\varnothing$. Then there exist a $n\in\mathbb{N}$ and a bijection $\beta:\{0,\ldots,n-1\}\rightarrow I\Rightarrow\{0,\ldots,n-1\}\approx I$ such that

   $$\sum_{i\in I}x_i=\sum_{i=0}^{n-1}x_{\beta(i)}=(n-1+1)\cdot x=n\cdot x\underset{[\text{theorem: }10.88]}{=}\operatorname{card}(I)\cdot x$$

□

Theorem 11.78. *Let $\langle V,+,\cdot\rangle$ be a vector space over $\langle F,+,\cdot\rangle$ and $W\subseteq V$ a sub-space of $\langle V,+,\cdot\rangle$ then:*

1. *$\forall n,m\in\mathbb{N}_0$ with $n\leqslant m$ and $\{x_i\}_{i\in\{n,\ldots,m\}}\subseteq W$ we have $\sum_{i=n}^m x_i\in W$*
2. *If I is a finite set and $\{x_i\}_{i\in I}\subseteq W$ then $\sum_{i\in I} x_i\in W$*

Proof.

1. We prove this by induction, so let

$$S=\left\{n\in\mathbb{N}_0|\forall\{x_i\}_{i\in\{0,\ldots,n\}}\subseteq W \text{ we have } \sum_{i=0}^n x_i\in W\right\}$$

then we have:

$0\in S$. If $\{x_i\}_{i\in\{0\}}\subseteq W$ then $x_0\in W$ so that $\sum_{i=0}^0 x_i=x_0\in W$ proving that $0\in W$.

$n\in S\Rightarrow n+1\in S$. Let $\{x_i\}_{i\in\{0,\ldots,n+1\}}$ then we have

$$\begin{aligned}\sum_{i=0}^{n+1} x_i &= \left(\sum_{i=0}^n x_i\right)+x_{n+1}\\ &= 1\cdot\left(\sum_{i=0}^n x_i\right)+x_{n+1}\\ &\in W\ [\text{as } n\in S\Rightarrow\sum_{i=0}^n x_i\in S\wedge x_{n+1}\in S]\end{aligned}$$

proving that $n+1\in\mathcal{S}$.

using induction we have then that

$$\forall n\in\mathbb{N}_0 \text{ and } \{x_i\}_{i\in\{0,\ldots,n\}}\subseteq W \text{ then } \sum_{i=0}^n x_i\in W \tag{11.15}$$

Let now $n,m\in\mathbb{N}_0$ with $n\leqslant m$ and $\{x_i\}_{i\in\{n,\ldots,m\}}\subseteq W$ then we have

$$\sum_{i=n}^m x_i=\sum_{i=0}^{m-n} x_{n+i}\in_{[\text{eq: } 11.15]} W$$

2. Let I be finite then we have either:

$I=\varnothing$. By [theorem: 11.57] that $\langle W,+,\cdot\rangle$ is a vector space with $0\in W$. Hence $\sum_{i\in I} x_i=0\in W$.

$I\neq\varnothing$. As $I\neq$ there exists a $\in$N and a bijection β: $\{0,\ldots,n-1\}\to I$ such that

$$\sum_{i\in I} x_i=\sum_{i=0}^{n-1} x_{\beta(i)}\in W \text{ [see eq: 11.15]}$$

□

11.3.2 Linear (in)dependency

11.3.2.1 Finite distinct set

In the previous section we have defined sums and products over finite sets by using the fact that for a finite set there exists a $n\in\mathbb{N}_0$ and a bijection between $\{1,\ldots,n\}$ and the finite set. To further develop the theory of dependent, independent sets in a vector space we have to work with these finite sums/products and it will be usefully to introduce a notation that expresses this relation between the finite set, $\{1,\ldots,n\}$ and the bijection.

Definition 11.79. (Finite distinct set) *Let $n\in\mathbb{N}_0$ and x: $\{1,\ldots,n\}\to A$ a bijection then we note this as $A=\{x_1,\ldots,x_n\}$, further if $i\in\{1,\ldots,n\}$ then x_i is a shorthand for $x(i)$. In other words if we write $A=\{x_1,\ldots,x_n\}$ then we actually means that $n\in\mathbb{N}_0$ and there exist a bijection x: $\{1,\ldots,n\}\to A$.*

The notation $A=\{x_1,\ldots,x_n\}$ is a way of specifying how the content of the set A must be calculated. In essence this is a extension of the notation $A=\{1.3.5\}$ which defines the set A by specifying that $x\in A\Leftrightarrow x=1\vee x=2\vee x=3$. Using the above definition we have then the following theorem.

Theorem 11.80. *If $A=\{x_1,\ldots,x_n\}$ then we have:*

1. $A=\{x(i)|i\in\{1,\ldots,n\}\}=\{x_i|i\in\{1,\ldots,n\}\}$
2. $\forall i\in\{1,\ldots,n\}$ $x_i=x(i)\in A$
3. $y\in A\Leftrightarrow\exists i\in\{1,\ldots,n\}$ *such that* $x_i=x(i)=y$
4. $\forall i,j\in I$ *with* $x_i=x_j$ *we have* $i=j$
5. $\forall i,j\in I$ *with* $i\neq j$ *we have* $x_i\neq x_j$

Proof. As $A=\{x_1,\ldots,x_n\}$ $n\in\mathbb{N}_0$ and $x:\{1,\ldots,n\}\rightarrow A$ is a bijection. Hence we have:

1. $A=x(\{1,\ldots,n\})=\{x(i)|i\in\{1,\ldots,n\}\}\underset{x_i\underset{\text{notation}}{=}x(i)}{=}\{x_i|i\in\{1,\ldots,n\}\}$.
2. $\forall i\in I$ we have $x_i=x(i)\in A$.
3. This follows from the fact that a bijection is surjective and (1)
4. This follows from the fact that a bijection is injective.
5. Let $i,j\in\{1,\ldots,n\}$ with $i\neq j$. If $x_i=x_j\Rightarrow x(i)=x(j)$ we have by injectivity that $i=j$ contradicting $i\neq j$, so we must have $x_i\neq x_j$.

Remark 11.81. If in a context we have defined $A=\{x_1,\ldots,x_n\}$ or equivalently the bijection $x\{\{1,\ldots,n\}\}\rightarrow A$ then we will identify the set A with $\{x_1,\ldots,x_n\}$. We will do this also if we are not really interested in the exact details of either A or x. Rephrasing the above theorem using this remark we have for $n\in\mathbb{N}_0$ and $\{x_1,\ldots,x_n\}$ that:

1. $\{x_1,\ldots,x_n\}=\{x(i)|i\in\{1,\ldots,n\}\}=\{x_i|i\in\{1,\ldots,n\}\}$
2. $\forall i\in\{1,\ldots,n\}$ $x_i=x(i)\in A$
3. $y\in\{x_1,\ldots,x_n\}\Leftrightarrow\exists i\in\{1,\ldots,n\}$ such that $x_i=x(i)=y$
4. $\forall i,j\in I$ with $x_i=x_j$ we have $i=j$
5. $\forall i,j\in I$ with $i\neq j$ we have $x_i\neq x_j$

Definition 11.82. *If $A=\{x_1,\ldots,x_n\}$ and $m\in\{1,\ldots,n\}$ then $B=\{x_1,\ldots,x_m\}$ represents the bijection $x_{|\{1,\ldots,m\}}:\{1,\ldots,m\}\rightarrow x(\{1,\ldots,m\})\subseteq A$.*

Remark 11.83. If we use the previous remark on the previous definition we have that $\forall n\in\mathbb{N}_0$ and $m\in\{1,\ldots,n\}$ that

$$\{x_1,\ldots,x_m\}\subseteq\{x_1,\ldots,x_n\}$$

Actually every finite set can be written as a finite distinct set.

Theorem 11.84. *We have*

$$A\text{ is finite}\Leftrightarrow A=\{x_1,\ldots,x_n\}\text{ where }n=\operatorname{card}(A)$$

Note that the representation $A=\{x_1,\ldots,x_n\}$ is not unique as there could be many bijections between $\{1,\ldots,n\}$ and A □

Proof.

⇒. If A is finite then by [theorem: 10.88] there exists a bijection $x:\{1,\ldots,n\}\rightarrow A$ where $n=\operatorname{card}(A)$. Hence $A=\{x_1,\ldots,x_n\}$ where $n=\operatorname{card}(A)$.

⇐. As $A=\{x_1,\ldots,x_n\}$ $x:\{1,\ldots,n\}\rightarrow A$ is a bijection so that A is finite. □

Example 11.85. $A=\varnothing\Leftrightarrow A=\{x_1,\ldots,x_0\}$

Proof.

⇒. Let $A=\varnothing$ then as $\{1,...,0\}=\varnothing$ we have for $x=\varnothing$ by [example: 2.63] the empty bijection $x=\varnothing:\{1,...,0\}\rightarrow\varnothing$ so that $A=\{x_1,...,x_0\}$.

⇐. If $A=\{x_1,\ldots,x_0\}$ then we have the bijection $x:\{1,\ldots,0\}=\varnothing\rightarrow A$ so that $A=x(\varnothing)=\varnothing$. □

The following lemma will be used in the Steinitz Lemma later.

Lemma 11.86. *$\forall n\in\mathbb{N}_0$ we have $\{x_1,\ldots,x_n\}\subseteq\{x_1,\ldots,x_{n+1}\}\setminus\{x_{n+1}\}$*

Proof. Let $A=\{x_1,\ldots,x_{n+1}\}$ and $B=\{x_1,\ldots,x_m\}$ then there exists a bijection $x:\{1,\ldots,n\}\rightarrow A$ and $x_{|\{1,\ldots,m\}}:\{1,\ldots,m\}\rightarrow x(\{1,\ldots,.m\})=B$. If $y\in A\setminus\{x_{n+1}\}$ then $\exists i\in\{1,\ldots,n+1\}$ such that $y=x_i$, we can not have $i=n+1$ because then $y=x_{n+1}$, hence $i\in\{1,\ldots,n\}$ and $y\in x(\{1,\ldots,n\})=B$. Hence we have

$$A\setminus\{x_{n+1}\}\subseteq x(\{1,\ldots,n\})=B$$

On the other hand if $y\in x(\{1,\ldots,n\})=B$ there exists a $i\in\{1,\ldots,n\}$ such that $y=x_i=x(i)\in A$. Assume that $y=x_{n+1}=x(n+1)$ then by injectivity we have from $x(i)=y=x(n+1)$ that $n+1$ contradicting $i\in\{1,\ldots,n\}$ hence $y\neq x_{n+1}$ proving that $y\in A\setminus\{x_{n+1}\}$. So

$$x(\{1,\ldots,x_n\})\subseteq A\setminus\{x_{n+1}\}$$

□

If $x:\{1,\ldots,n\}\rightarrow A$ is a bijection and $\beta:\{1,\ldots,n\}\rightarrow\{1,\ldots,n\}$ another bijection then we can consider the bijection $x\circ\beta:\{1,\ldots,n\}\rightarrow A$ so that we have the finite distinct set

$$A=\{(x\circ\beta)_1,\ldots,(x\circ\beta)_n\}$$

and given $i\in\{1,\ldots,n\}$ we have $(x\circ\beta)_i\underset{\text{def}}{=}(x\circ\beta)(i)=x(\beta(i))\underset{\text{def}}{=}x_{\beta(i)}$. So the following definition seems logical

Definition 11.87. *Let $A=\{x_1,\ldots,x_n\}$ be a finite disjoint set and $\beta:\{1,\ldots,n\}\rightarrow\{1,\ldots,n\}$ a bijection then*

$$A=\{x_{\beta(1)},\ldots,x_{\beta(n)}\}$$

is the finite distinct set defined by the bijection

$$x\circ\beta:\{1,\ldots,n\}\rightarrow A.$$

Definition 11.88. (Ordered Family) *A family of the form $\{x_i\}_{i\in\{1,\ldots,n\}}\subseteq X$ where $n\in\mathbb{N}_0$ is called a **ordered family**.*

The set of all members of a ordered family is always finite as is expressed in the following theorem.

Theorem 11.89. *If $\{x_i\}_{i\in\{1,\ldots,n\}}\subseteq X$, $n\in\mathbb{N}_0$ is a ordered family then $\{x_i|i\in\{1,\ldots,n\}\}$ is finite. with $\operatorname{card}(\{x_i|i\in\{1,\ldots,n\}\})\leqslant n$.*

Proof. As $\{x_i\}_{i\in\{1,\ldots,n\}}\subseteq X$ is actually the function $x:\{1,\ldots,n\}\rightarrow X$ we have that

$$x:\{1,\ldots,n\}\rightarrow x(\{1,\ldots,n\})=\{x(i)|i\in I\}=\{x_i|i\in I\}\text{ is a surjection}$$

So by [theorem: 6.43] $\{x_i|i\in\{1,\ldots,n\}\}$ is finite and $\operatorname{card}(\{x_i|i\in\{1,\ldots,n\}\})\leqslant n$. □

Theorem 11.90. *Let X be a set, $n\in\mathbb{N}_0$, $\{x_i\}_{i\in\{1,\ldots,n\}}\subseteq X$ and $Y\subseteq\{x_i|i\in\{1,\ldots,n\}\}$ then there exist a bijection*

$$\beta:\{1,\ldots,\operatorname{card}(Y)\}\rightarrow J\subseteq\{1,\ldots,n\}$$

such that

$$Y=\{x_{\beta(1)},\ldots,x_{\beta(\operatorname{card}(Y))}\}$$

Proof. Let $x:\{1,\ldots,n\}\rightarrow X$ be the function defining the family $\{x_i\}_{i\in\{1,\ldots,n\}}$. Using a consequence of the Axiom of Choice [see theorem: 3.129] there exists a $I\subseteq\{1,\ldots,n\}$ such that

$$x_{|I}:I\rightarrow x(\{1,\ldots,n\})=\{x|i\in\{1,\ldots,n\}\}\text{ is a bijection}$$

Let $J=(x_{|I})^{-1}(Y)\subseteq I$ then for

$$x_{|J}\underset{[\text{theorem: } 2.84]}{=}(x_{|I})_{|J}:J\rightarrow Y$$

we have:

- **injectivity.** Let $i,j\in J$ such that $x_{|J}(i)=x_{|J}(j)$ then $x(i)=x(j)$ and as $J\subseteq I$ we have $x_{|I}(i)=x_{|I}(j)$. So by injectivity of $x_{|I}$ it follows that $i=j$.
- **surjectivity.** Let $y\in Y$ then, as $x_{|I}:I\rightarrow\{x_i|i\in\{1,\ldots,n\}\}$ is a bijection and $Y\subseteq\{x_i|i\in\{1,\ldots,n\}\}$, there exists a $i\in I$ such that $y=x_{|I}(i)=x(i)$. As $y\in Y$ we have $i\in(x_{|I})^{-1}(Y)=J$ hence $y=x(i)\underset{i\in J}{=}x_{|J}(i)$.

From the above we conclude that

$$x_{|J}:J\rightarrow Y\text{ is a bijection}$$

As Y is finite [using theorems: 10.89, 11.89] we have that J is finite with $\operatorname{card}(J)=\operatorname{card}(Y)$. Hence there exists a bijection

$$\beta: \{1,\ldots,\operatorname{card}(Y)\} \to J$$

As the composition of two bijections is again a bijection we have that

$$x_{|J} \circ \beta: \{1,\ldots,\operatorname{card}(Y)\} \to Y \text{ is a bijection}$$

so that $Y=\{(x_{|J}\circ\beta)_1,\ldots,(x_{|J}\circ\beta)_{\operatorname{card}(Y)}\}$. As $\forall i \in \{1,\ldots,\operatorname{card}(Y)\}$ we have $(x_{|J}\circ\beta)_i = x(\beta(i)) = x_{\beta(i)}$ we have

$$Y=\{x_{\beta(1)},\ldots,x_{\beta(\operatorname{card}(Y))}\}$$

□

11.3.2.2 Span of a set

Definition 11.91. *Let $\langle V,+,\cdot\rangle$ be a vector space over a field $\langle F,+,\cdot\rangle$ and $W\subseteq V$ then $v\in V$ is a linear combination of W if there exists a finite set $I\subseteq W$ and a $\{\alpha_i\}_{i\in I}\subseteq F$ such that*

$$v=\sum_{i\in I}\alpha_i\cdot i$$

For finite sets (that can always be representation as a finite distinct sets) we have a simpler definition of a linear combination in terms of the finite distinct representation of a finite set.

Theorem 11.92. *Let $\langle V,+,\cdot\rangle$ be a vector space over a field $\langle F,+,\cdot\rangle$, $n\in\mathbb{N}_0$, $W=\{w_1,\ldots,w_n\}\subseteq V$ a finite set and $v\in V$ then we have:*

1. *If $\exists\{\alpha_u\}_{u\in W}\subseteq F$ such that $v=\sum_{u\in W}\alpha_u\cdot w$ then there exists a $\{\beta_i\}_{i\in\{1,\ldots,n\}}\subseteq F$ such that $v=\sum_{i\in\{1,\ldots,n\}}\beta_i\cdot w_i$ and $\{\alpha_u|u\in W\}=\{\beta_i|i\in\{1,\ldots,n\}\}$.*
2. *If $\exists\{\alpha_i\}_{i\in\{1,\ldots,n\}}\subseteq F$ such that $v=\sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot w_i$ then there exists a $\{\beta_w\}_{w\in W}\subseteq F$ such that $v=\sum_{u\in W}\beta_u\cdot u$ and $\{\alpha_i|i\in\{1,\ldots,n\}\}=\{\beta_w|w\in W\}$.*

Proof. First as $W=\{w_1,\ldots,w_n\}$ we have the following two bijections

$$w:\{1,\ldots,n\}\to W \text{ and } w^{-1}: W\to\{1,\ldots,n\} \tag{11.16}$$

Now for the rest of the proof:

1. Define $\{\beta_i\}_{i\in\{1,\ldots,n\}}\subseteq F$ by the function $\beta=\alpha\circ w:\{1,...,n\}\to F$ [so that $\forall i\in\{1,...,n\}$ $\beta_i=(\alpha\circ\omega)=\alpha(w(i))=\alpha_{w(i)}$] then we have

$$\begin{aligned}\sum_{i\in\{1,\ldots,n\}}\beta_i\cdot w_i &= \sum_{i\in\{1,\ldots,n\}}\alpha_{w(i)}\cdot w(i)\\ &\underset{[\text{eq:11.16}]\wedge[\text{theorem: 11.36}]}{=} \sum_{u\in W}\alpha_u\cdot u\\ &= v\end{aligned}$$

   Further we have:

$$\begin{aligned}\{\alpha_u|u\in W\} &= \alpha(W)\\ &\underset{w:\{1,\ldots,n\}\to W\text{ is bijective}}{=} \alpha(w(\{1,\ldots,n\}))\\ &= (\alpha\circ w)(1,\ldots,n)\\ &= \beta(\{1,\ldots,n\})\\ &= \{\beta_i|i\in\{1,\ldots,n\}\}\end{aligned}$$

2. Define $\{\beta_u\}_{u\in W}\subseteq F$ by the function $\beta=\alpha\circ w^{-1}: W\to F$ [so that $\forall i\in\{1,...,n\}$ $\beta_u=(\alpha\circ w^{-1})=\alpha(w^{-1}(u))=\alpha_{w^{-1}(u)}\}$ then we have

$$\begin{aligned}\sum_{u\in W}\beta_u\cdot u &= \sum_{u\in W}\alpha_{w^{-1}(u)}\cdot w(w^{-1}(u))\\ &\underset{[\text{eq:11.16}]\wedge[\text{theorem: 11.36}]}{=} \sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot w(i)\\ &= \sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot w_i\end{aligned}$$

Further we have:

$$\begin{array}{lll}\{\alpha_i|i\in\{1,\ldots,n\}\} & = & \alpha(\{1,\ldots,n\}) \\ & \underset{w^{-1}:W\rightarrow\{1,\ldots,n\}\text{ is bijective}}{=} & \alpha((w^{-1})(W)) \\ & = & (\alpha\circ w)(W) \\ & = & \beta(W) \\ & = & \{\beta_u|u\in W\}\end{array}$$

□

We have a more general version of the above theorem.

Theorem 11.93. *Let $\langle V,+,\cdot\rangle$ be a vector space over a field $\langle F,+,\cdot\rangle$,$n\in\mathbb{N}_0$ and $\{w_i\}_{i\in\{1,\ldots,n\}}\subseteq V$ then we have:*

1. *If $\{\alpha_i\}_{i\in\{1,\ldots,n\}}\subseteq F$ then there exist a $\{\beta_v\}_{v\in\{w_i|i\in\{1,\ldots,n\}\}}\subseteq F$ such that*
$$\sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot w_i=\sum_{v\in\{w_i|i\in\{1,\ldots,n\}\}}\beta_v\cdot v$$
2. *If $\{\alpha_w\}_{w\in\{w_i|i\in\{1,\ldots n\}\}}\subseteq F$ then there exist a $\{\beta_i\}_{i\in\{1,\ldots,n\}}\subseteq F$ such that*
$$\sum_{v\in\{w_i|i\in\{1,\ldots,n\}\}}\alpha_v\cdot v=\sum_{i\in\{1,\ldots,n\}}\beta_i\cdot w_i$$

***Note* 11.94.** *Using [theorem: 11.89] $\{w_i|i\in\{1,\ldots,n\}\}$ is finite so that the sums over $\{w_i|i\in\{1,\ldots,n\}\}$ are well defined.*

Proof.

1. By [theorem: 11.89] we have that
$$\{w_i|i\in\{1,\ldots,n\}\}\text{ is finite}\tag{11.17}$$
Let $w\in\{w_i|i\in\{1,\ldots,n\}\}$ and define $I_w=\{i\in\{1,\ldots,n\}|w_i=w\}$. If $j\in\{1,\ldots,n\}$ we have $w_j\in\{w_i|i\in\{1,\ldots,n\}\}$ so that $j\in I_{w_j}$ proving that $\{1,\ldots,n\}\subseteq\bigcup_{w\in\{w_i|i\in\{1,\ldots,n\}\}}I_w$. Further as $I_w\subseteq\{1,\ldots,n\}$ it follows that $\bigcup_{w\in\{w_i|i\in\{1,\ldots,n\}\}}I_w\subseteq\{1,\ldots,n\}$. Hence we have
$$\{1,\ldots,n\}=\bigcup_{w\in\{w_i|i\in\{1,\ldots,n\}\}}I_w\tag{11.18}$$
Let $w,u\in\{w_i|i\in\{1,\ldots,n\}\}$ with $w\neq u$, if $k\in I_w\bigcap I_u$ then $w=w_k=u$ contradicting $w\neq u$, hence we must have
$$\forall w,u\in\{w_i|i\in\{1,\ldots,n\}\}\text{ with }w\neq u\text{ we have }I_w\bigcap I_u=\varnothing\tag{11.19}$$
Define now
$$\{\beta_w\}_{w\in\{w_i|i\in\{1,\ldots,n\}\}}\subseteq F\text{ by }\beta_w=\sum_{i\in I_w}\alpha_i\tag{11.20}$$
Then we have
$$\begin{array}{lll}\displaystyle\sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot w_i & \underset{\text{[theorem: 11.44] and [eqs: 11.17, 11.18, 11.19]}}{=} & \displaystyle\sum_{w\in\{w_i|i\in\{1,\ldots,n\}\}}\left(\sum_{i\in I_w}\alpha_i\cdot w_i\right) \\ & \underset{i\in I_w\Rightarrow w_i=w}{=} & \displaystyle\sum_{w\in\{w_i|i\in\{1,\ldots,n\}\}}\left(\sum_{i\in I_w}\alpha_i\cdot w\right) \\ & \underset{\text{[theorem: 11.76]}}{=} & \displaystyle\sum_{w\in\{w_i|i\in\{1,\ldots,n\}\}}\left(\left(\sum_{i\in I_w}\alpha_i\right)\cdot w\right) \\ & \underset{\text{[theorem: 11.20]}}{=} & \displaystyle\sum_{w\in\{w_i|i\in\{1,\ldots,n\}\}}\beta_w\cdot w\end{array}$$
2. Let $w:\{1,\ldots,n\}\rightarrow V$ be the function defining $\{w_i\}_{i\in\{1,\ldots,n\}}$ then by a consequence of the Axiom of Choice [see theorem: 3.129] there exists a $I\subseteq\{1,\ldots,n\}$ such that
$$w_{|I}:I\rightarrow w(\{1,\ldots,n\})=\{w_i|i\in\{1,\ldots,n\}\}\text{ is a bijection}$$

Define now

$$\{\beta_i\}_{i\in\{1,\ldots,n\}}\subseteq F \text{ by } \beta_i=\begin{cases}\alpha_{w(i)} \text{ if } i\in I\\ 0 \text{ if } i\in\{1,\ldots,n\}\setminus I\end{cases}$$

then we have

$$\begin{aligned}\sum_{i\in\{1,\ldots,n\}}\beta_i\cdot w_i &\underset{[\text{theorem: } 11.43]}{=} \sum_{i\in\{1,\ldots,n\}\setminus I}\beta_i\cdot w_i+\sum_{i\in I}\beta_i\cdot w_i\\ &= \sum_{i\in\{1,\ldots,n\}\setminus I}0\cdot w_i+\sum_{i\in I}\alpha_{w(i)}\cdot w(i)\\ &= \sum_{i\in I}\alpha_{w_{|I}(i)}\cdot w_{|I}(i)\\ &\underset{[\text{theorem: } 11.36]}{=} \sum_{v\in\{w_i|i\in\{1,\ldots,n\}\}}\alpha_w\cdot w\end{aligned}$$

□

The set of all linear combinations of a set is called the span of the set.

Definition 11.95. *Let $\langle V,+,\cdot\rangle$ be a vector space over a field $\langle F,+,\cdot\rangle$ and $W\subseteq V$ then*

$$\operatorname{span}(W)=\left\{v\in V|\exists\{\alpha_w\}_{w\in I}\subseteq F,\ I \text{ finite and } I\subseteq W \text{ such that } v=\sum_{w\in I}\alpha_w\cdot w\right\}$$

We can reformulate this definition using ordered families of vectors.

Theorem 11.96. *Let $\langle V,+,\cdot\rangle$ be a vector space over a field $\langle F,+,\cdot\rangle$ and $W\subseteq V$ then*

$$\operatorname{span}(W)=\left\{v\in V|\exists n\in\mathbb{N}_0,\{u_i\}_{i\in\{1,\ldots,n\}}\subseteq W \text{ and } \{\alpha_i\}_{i\in\{1,\ldots,n\}}\subseteq F \text{ with } v=\sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot u_i\right\}$$

Proof. Let

$$S=\left\{v\in V|\exists n\in\mathbb{N}_0,\{u_i\}_{i\in\{1,\ldots,n\}}\subseteq W \text{ and } \{\alpha_i\}_{i\in\{1,\ldots,n\}}\subseteq F \text{ with } v=\sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot u_i\right\}$$

If $v\in\operatorname{span}(W)$ then there exist a finite $I\subseteq W$ and $\{\beta_u\}_{u\in I}\subseteq F$ so that $v=\sum_{u\in I}\beta_u\cdot u$. As I is finite we have by [theorem: 11.84] that $I=\{u_1,\ldots,u_n\}\subseteq W$ hence $\{u_i\}_{i\in\{1,\ldots,n\}}\subseteq W$. Using [theorem: 11.92] there exists a $\{\alpha_i\}_{i\in\{1,\ldots,n\}}\subseteq F$ such that $v=\sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot u_i$ proving that $v\in S$. Hence

$$\operatorname{span}(W)\subseteq S \tag{11.21}$$

Let $v\in S$ then there exists a $n\in\mathbb{N}_0$, a $\{u_i\}_{i\in\{1,\ldots,n\}}\subseteq W$ and $\{\alpha_i\}_{i\in\{1,\ldots,n\}}\subseteq F$ such that

$$v=\sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot u_i \tag{11.22}$$

Let

$$U=\{u_i|i\in\{1,\ldots,n\}\}\subseteq W \tag{11.23}$$

then by [theorem: 11.89]

$$U \text{ is finite} \tag{11.24}$$

For $u\in U$ define $I_u=\{i\in\{1,\ldots,n\}|u_i=u\}\subseteq\{1,\ldots,n\}$ then we have

$$\bigcup_{u\in U}I_u\subseteq\{1,\ldots,n\} \tag{11.25}$$

Further if $i\in\{1,\ldots,n\}$ then $u_i\in I_{u_i}$ and as $u_i\in\{u_i|i\in\{1,\ldots,n\}\}\underset{[\text{eq: } 11.23]}{=}U$ we have $i\in\bigcup_{u\in U}I_u$, so $\{1,\ldots,n\}\subseteq\bigcup_{u\in U}I_u$. Combining this with [eq: 11.25] gives

$$\bigcup_{u\in U}I_u=\{1,\ldots,n\} \tag{11.26}$$

Further if $u,v\in U$ with $u\neq v$ then if $i\in I_u\bigcap I_v$ we have that $u=u_i=v$ contradicting $u\neq v$ hence

$$\forall u,v\in U \text{ with } u\neq v \text{ we have } I_u\bigcap I_v=\varnothing \tag{11.27}$$

As $\forall u \in U \; I_u \subseteq \{1, \ldots, \}$ we have by [theorem: 10.89] that I_u is finite and we can define

$$\{\beta_u\}_{u \in U} \text{ by } \beta_u = \sum_{i \in I_u} \alpha_i \tag{11.28}$$

Then

$$\begin{aligned}
v &\underset{[\text{eq: } 11.22]}{=} \sum_{i \in \{1, \ldots, n\}} \alpha_i \cdot u_i \\
&\underset{[\text{theorem: } 11.44] \wedge [\text{eqs:} 11.26, 11.27]}{=} \sum_{u \in U} \left(\sum_{i \in I_u} \alpha_i \cdot u_i \right) \\
&\underset{i \in I_u \Rightarrow u_i = u}{=} \sum_{u \in U} \left(\sum_{i \in I_u} \alpha_i \cdot u \right) \\
&\underset{[\text{theorem: } 11.76]}{=} \sum_{u \in U} \left(\left(\sum_{i \in I_u} \alpha_i \right) \cdot u \right) \\
&\underset{[\text{eq: } 11.28]}{=} \sum_{u \in U} \beta_u \cdot u
\end{aligned}$$

which as $U \subseteq W$ is finite [see eqs: 11.23, 11.24] proves that $v \in \text{span}(W)$. Hence $S \subseteq \text{span}(W)$. Combining this with [eq: 11.21] results in $S = \text{span}(W)$ or

$$\text{span}(W) = \left\{ v \in V \,|\, \exists n \in \mathbb{N}_0, \{u_i\}_{i \in \{1, \ldots, n\}} \subseteq W \text{ and } \{\alpha_i\}_{i \in \{1, \ldots, n\}} \subseteq F \text{ with } v = \sum_{i \in \{1, \ldots, n\}} \alpha_i \cdot u_i \right\}$$

□

This definition for the span of a set has to use a finite $I \subseteq W$ so that the the sum is well defined as a finite sum (for infinite sums we have to introduce the concept of limits and topology). For finite sets we have a simpler version of the span.

Theorem 11.97. *Let $\langle V, +, \cdot \rangle$ be a vector space over a field $\langle F, +, \cdot \rangle$ and $\{w_i\}_{i \in \{1, \ldots, n\}} \subseteq V$ then*

$$\text{span}(\{w_i | i \in \{1, \ldots, n\}\}) = \left\{ v \in V \,|\, \exists \{\alpha_i\}_{i \in \{1, \ldots, n\}} \subseteq F \text{ such that } v = \sum_{i \in \{1, \ldots, n\}} \alpha_i \cdot w_i \right\}$$

***Note* 11.98.** *If $W \subseteq V$ is a finite set then by [theorem: 11.84⟨⟩ $W = \{w_1, \ldots, n\}$ where $n = \text{card}(W)$ so that the above becomes*

$$\text{span}(W) = \left\{ v \in V \,|\, \exists \{\alpha_i\}_{i \in \{1, \ldots, n\}} \subseteq F \text{ such that } v = \sum_{i \in \{1, \ldots, n\}} \alpha_i \cdot w_i \right\}$$

Proof. Let $x \in \text{span}(\{w_i | i \in \{1, \ldots, n\}\})$ then by definition there exists a finite $I \subseteq \{w_i | i \in \{1, \ldots, n\}\}$ and a $\{\alpha_v\}_{v \in I}$ such that

$$x = \sum_{v \in I} \alpha_v \cdot v \tag{11.29}$$

Define now

$$\{\beta_v\}_{v \in \{w_i | i \in \{1, \ldots, n\}\}} \subseteq F \text{ by } \beta_v = \begin{cases} \alpha_v \text{ if } v \in I \\ 0 \text{ if } v \in \{w_i | i \in \{1, \ldots, n\}\} \setminus I \end{cases}$$

then using [theorem: 11.93] there exists a $\{\gamma_i\}_{i \in \{1, \ldots, n\}} \subseteq F$ such that

$$\sum_{v \in \{w_i | i \in \{1, \ldots, n\}\}} \beta_v \cdot v = \sum_{i \in \{1, \ldots, n\}} \gamma_i \cdot w_i \tag{11.30}$$

Hence

$$\begin{aligned}
\sum_{i \in \{1, \ldots, n\}} \gamma_i \cdot w_i &\underset{[\text{eq: } 11.30]}{=} \sum_{v \in \{w_i | i \in \{1, \ldots, n\}\}} \beta_v \cdot v \\
&\underset{[\text{theorem: } 11.43]}{=} \sum_{v \in \{w_i | i \in \{1, \ldots, n\}\} \setminus I} \beta_v \cdot v + \sum_{v \in I} \beta_v \cdot v \\
&= \sum_{v \in \{w_i | i \in \{1, \ldots, n\}\} \setminus I} 0 \cdot v + \sum_{v \in I} \alpha_v \cdot v \\
&= \sum_{v \in I} \alpha_v \cdot v \\
&\underset{[\text{eq: } 11.29]}{=} x
\end{aligned}$$

proving that $x\in\left\{v\in V|\exists\{\alpha_i\}_{i\in\{1,\ldots,n\}}\subseteq F\text{ such that }v=\sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot w_i\right\}$. Hence we have

$$\mathrm{span}(\{w_i|i\in\{1,\ldots,n\}\})\subseteq\left\{v\in V|\exists\{\alpha_i\}_{i\in\{1,\ldots,n\}}\subseteq F\text{ such that }v=\sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot w_i\right\}$$

Fort the opposite inclusion, let $x\in\left\{v\in V|\exists\{\alpha_i\}_{i\in\{1,\ldots,n\}}\subseteq F\text{ such that }v=\sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot w_i\right\}$ then there exist a $\{\alpha_i\}_{i\in\{1,\ldots,n\}}\subseteq F$ such that $x=\sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot w_i$. Using [theorem: 11.93] there exists a $\{\beta_v\}_{v\in\{w_i|i\in\{1,\ldots,n\}\}}\subseteq F$ such that $x=\sum_{v\in\{w_i|i\in\{1,\ldots,n\}\}}\beta_v\cdot v$, which, as $\{w_i|i\in\{1,\ldots,n\}\}$ is finite and a subset of itself, proves that $x\in\mathrm{span}(\{w_i|i\in\{1,\ldots,n\}\})$. Hence we have also

$$\left\{v\in V|\exists\{\alpha_i\}_{i\in\{1,\ldots,n\}}\subseteq F\text{ such that }v=\sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot w_i\right\}\subseteq\mathrm{span}(\{w_i|i\in\{1,\ldots,n\}\})$$ □

Example 11.99. Let $\langle V,+,\cdot\rangle$ be a vector space over a field $\langle F,+,\cdot\rangle$ then $\mathrm{span}(\varnothing)=\{0\}$

Proof. Let $v\in\mathrm{span}(\varnothing)$ then there exists a finite $I\subseteq\varnothing\Rightarrow I=\varnothing$ and a $\{\alpha_w\}_{w\in\varnothing}\subseteq F$ such that

$$v=\sum_{w\in\varnothing}\alpha_w\cdot w\underset{[\text{definition: }11.32]}{=}0$$

proving that

$$\mathrm{span}(\varnothing)=0$$ □

Example 11.100. Let $\langle V,+,\cdot\rangle$ be a vector space over a field $\langle F,+,\cdot\rangle$ then $\mathrm{span}(\{0\})=\{0\}$

Proof. Let $v\in\mathrm{span}(\{0\})$ then there exists a finite $I\subseteq\{0\}$ and a $\{\alpha_u\}_{u\in I}\subseteq F$ such that $v=\sum_{u\in I}\alpha_u\cdot u$. For I we have either:

$\boldsymbol{I=\varnothing}$**.** Then $v=\sum_{u\in I}\alpha_u\cdot u=\sum_{u\in\varnothing}\alpha_u\cdot u\underset{[\text{definition: }11.32]}{=}0$

$\boldsymbol{I=\{0\}}$**.** Then $v=\sum_{u\in I}\alpha_u\cdot u=\sum_{u\in\{0\}}\alpha_u\cdot u\underset{[\text{theorem: }11.37]}{=}\alpha_0\cdot 0=0$

proving that

$$\mathrm{span}(\{0\})=\{0\}$$ □

The coefficients in a span of a set can be assumed to be different from the neutral element as is show in the next theorem.

Theorem 11.101. *Let $\langle V,+,\cdot\rangle$ be a vector space over a field $\langle F,+,\cdot\rangle$ and $W\subseteq V$ then*

$$v\in\mathrm{span}(W)$$

$$\Updownarrow$$

$$\textit{There exists a finite } I\subseteq W \textit{ and a } \{\alpha_w\}_{w\in I}\subseteq F\setminus\{0\}\textit{ such that } v=\sum_{w\in I}\alpha_w\cdot w$$

Proof.

⇒. As $v\in\mathrm{span}(W)$ there exist by definition a $J\subseteq W$ and a $\{\beta_w\}_{w\in J}\subseteq F$ such that

$$v=\sum_{w\in J}\alpha_w\cdot w$$

Let $I=\{w\in J|\alpha_w\neq 0\}\subseteq J$ and define $\{\alpha_w\}_{w\in I}\subseteq F\setminus\{0\}$ by $\alpha_w=\beta_w$ then we have

$$\begin{aligned}
v &= \sum_{w\in J}\beta_w\cdot w\\
&\underset{[\text{theorem: }11.43]}{=} \sum_{w\in J\setminus I}\beta_w\cdot w+\sum_{w\in I}\beta_w\cdot w\\
&= \sum_{w\in J\setminus I}0\cdot w+\sum_{w\in I}\beta_w\cdot w\\
&\underset{[\text{theorem: }11.37]}{=} \sum_{w\in I}\beta_w\cdot w\\
&= \sum_{w\in I}\alpha_w\cdot w
\end{aligned}$$

⇐. As $F\setminus\{0\}\subseteq F$ this follows from the definition of $\mathrm{span}(W)$. □

Theorem 11.102. *Let $\langle V, +, \cdot \rangle$ be a vector space over a field $\langle F, +, \cdot \rangle$ and $W \subseteq V$ then*

$$v \in \operatorname{span}(W)$$

$$\Updownarrow$$

$$\exists U = \{u_1, \ldots, u_n\} \textit{with } U \subseteq W \textit{ and } \{\alpha_i\}_{i \in \{1, \ldots, n\}} \subseteq F \setminus \{0\} \textit{ such that } v = \sum_{i \in \{1, \ldots, n\}} \alpha_i \cdot w_i$$

Proof.

⇒. As $v \in \operatorname{span}(W)$ there exists by the previous theorem [theorem: 11.101] a finite $U \subseteq W$ and a $\{\beta_u\}_{u \in U} \subseteq F \setminus \{0\}$ such that $v = \sum_{u \in U} \beta_u \cdot u$. Using [theorem: 11.84] we have that

$$U = \{u_1, \ldots, u_n\}$$

where $n = \operatorname{card}(U)$. Applying then [theorem: 11.92] gives a $\{\alpha_i\}_{i \in \{1, \ldots, n\}} \subseteq F$ such that

$$v = \sum_{i \in \{1, \ldots, n\}} \alpha_i \cdot u_i$$

and

$$\{\beta_u | u \in U\} = \{\alpha_i | i \in \{1, \ldots, n\}\}$$

As $\{\beta_u\}_{u \in U} \subseteq F \setminus \{0\}$ it follows that

$$\{\alpha_i\}_{i \in \{1, \ldots, n\}} \subseteq F \setminus \{0\}$$

⇐. As $U = \{u_1, \ldots, u_n\}$, $U \subseteq W$, $\{\alpha_i\}_{i \in \{1, \ldots, n\}} \subseteq F \setminus \{0\} \subseteq F$ and $v = \sum_{i \in \{1, \ldots, n\}} \alpha_i \cdot w_i$ it follows by the definition of the span of a set that $n \in \operatorname{span}(W)$. □

Theorem 11.103. *Let $\langle V, +, \cdot \rangle$ be a vector space over a field $\langle F, +, \cdot \rangle$ and $W \subseteq V$ then*

$$W \subseteq \operatorname{span}(W)$$

Proof. Let $w \in W$ then $\{w\}$ is finite and $\{w\} \subseteq W$. Define $\{\alpha_u\}_{u \in \{w\}}$ by $\alpha_u = 1$ then

$$\sum_{u \in \{w\}} \alpha_u \cdot u \underset{[\text{theorem: } 11.34]}{=} \alpha_w \cdot w = 1 \cdot w = w$$

proving that $w \in \operatorname{span}(W)$.Hence $W \subseteq \operatorname{span}(W)$. □

Theorem 11.104. *Let $\langle V, +, \cdot \rangle$ be a vector space over a field $\langle F, +, \cdot \rangle$ and $U, W \subseteq V$ with $U \subseteq W$ then*

$$\operatorname{span}(U) \subseteq \operatorname{span}(W)$$

Proof. Let $v \in \operatorname{span}(U)$ then there exists a finite $I \subseteq U$ and a $\{\alpha_i\}_{i \in I} \subseteq F$ such that $v = \sum_{w \in I} \alpha_w \cdot w$. As $U \subseteq W$ we have $I \subseteq W$ so that $v \in \operatorname{span}(W)$. □

Theorem 11.105. *Let $\langle V, +, \cdot \rangle$ be a vector space over a field $\langle F, +, \cdot \rangle$ and $W \subseteq V$ then*

$$\operatorname{span}(W) \textit{ is a sub-space of } \langle V, +, \cdot \rangle$$

Proof. As $\varnothing \subseteq W$ we have $\{0\} \underset{[\text{example: } 11.99]}{=} \operatorname{span}(\varnothing) \subseteq_{[\text{theorem: } 11.104]} \operatorname{span}(W)$ proving that

$$0 \in W \text{ hence } W \neq \varnothing$$

Let $x, y \in \operatorname{span}(W)$ and $\alpha \in F$ then there exists finite $I, J \subseteq W$ and $\{\alpha_w\}_{w \in I} \subseteq F$, $\{\beta_w\}_{w \in J} \subseteq F$ such that

$$x = \sum_{w \in I} \alpha_w \cdot w \text{ and } y = \sum_{w \in J} \beta_w \cdot w$$

As $I, J$are finite we have by [theorem: 6.33] that $K = I \bigcup J = (I \setminus J) \bigcup (I \bigcap J) \bigcup (J \setminus I)$ is finite. Define

$$\{\gamma_w\}_{w \in K} \subseteq F \text{ by } \gamma_w = \begin{cases} \alpha \cdot \alpha_w \text{ if } w \in I \setminus J \\ \alpha \cdot \alpha_w + \beta_w \text{ if } w \in I \bigcap J \\ \beta_w \text{ if } w \in J \setminus I \end{cases}$$

then we have

$$
\begin{aligned}
\sum_{w \in K} \gamma_w \cdot w &\underset{[\text{theorem: } 11.43]}{=} \\
\sum_{w \in I \setminus J} \gamma_w \cdot w + \sum_{w \in I \bigcap J} \gamma_w \cdot w + \sum_{w \in J \setminus I} \gamma_w \cdot w &= \\
\sum_{w \in I \setminus J} (\alpha \cdot \alpha_w) \cdot w + \sum_{w \in I \bigcap J} (\alpha \cdot \alpha_w + \beta_w) \cdot w + \sum_{w \in J \setminus I} \beta_w \cdot w &= \\
\sum_{w \in I \setminus J} \alpha \cdot (\alpha_w \cdot w) + \sum_{w \in I \bigcap J} (\alpha \cdot (\alpha_w \cdot w) + \beta_w \cdot w) + \sum_{w \in J \setminus I} \beta_w \cdot w &\underset{[\text{theorem: } 11.75]}{=} \\
\alpha \cdot \sum_{w \in I \setminus J} \alpha_w \cdot w + \sum_{w \in I \bigcap J} (\alpha \cdot (\alpha_w \cdot w) + \beta_w \cdot w) + \sum_{w \in J \setminus I} \beta_w \cdot w &\underset{[\text{theorem: } 11.38]}{=} \\
\alpha \cdot \sum_{w \in I \setminus J} \alpha_w \cdot w + \sum_{w \in I \bigcap J} \alpha \cdot (\alpha_w \cdot w) + \sum_{i \in I \bigcap J} \beta_w \cdot w + \sum_{w \in J \setminus I} \beta_w \cdot w &\underset{[\text{theorem: } 11.75]}{=} \\
\alpha \cdot \sum_{w \in I \setminus J} \alpha_w \cdot w + \alpha \cdot \sum_{w \in I \bigcap J} \alpha_w \cdot w + \sum_{i \in I \bigcap J} \beta_w \cdot w) + \sum_{w \in J \setminus I} \beta_w \cdot w &= \\
\alpha \cdot \left(\sum_{w \in I \setminus J} \alpha_w \cdot w + \sum_{w \in I \bigcap J} \alpha_w \cdot w \right) + \left(\sum_{w \in J \setminus I} \beta_w \cdot w + \sum_{i \in I \bigcap J} \beta_w \cdot w) \right) &\underset{[\text{theorem: } 11.43]}{=} \\
\alpha \cdot \sum_{w \in I} \alpha_w \cdot w + \sum_{w \in J} \beta_w \cdot w &= \\
\alpha \cdot x + y &
\end{aligned}
$$

proving that

$$\alpha \cdot x + y \in \operatorname{span}(W)$$

Hence by the definition of a sub-space [see definition: 11.56] it follows that $\operatorname{span}(W)$ is a sub-space of $\langle V, +, \cdot \rangle$. □

Theorem 11.106. *Let $\langle V, +, \cdot \rangle$ be a vector space over a field $\langle F, +, \cdot \rangle$ and $W \subseteq V$ a sub-space of V then*

$$W = \operatorname{span}(W)$$

Proof. Let $v \in \operatorname{span}(W)$ then there exists a finite $I \subseteq W$ and $\{\alpha_w\}_{w \in I} \subseteq F$ such that

$$v = \sum_{w \in I} \alpha_w \cdot w$$

As W is a sub-space we have $\forall w \in I$ that $\alpha_w \cdot w \in w$, by [theorem: 11.78] it follows then that $v = \sum_{w \in I} \alpha_w \cdot w \in W$. Hence $\operatorname{span}(W) \subseteq W$ which together with [theorem: 11.103] proves

$$W = \operatorname{span}(W)$$

□

Corollary 11.107. *Let $\langle V, +, \cdot \rangle$ be a vector space over a field $\langle F, +, \cdot \rangle$ and $W \subseteq V$ then*

$$\operatorname{span}(W) = \operatorname{span}(\operatorname{span}(W))$$

Proof. By [theorem: 11.105] $\operatorname{span}(W)$ is a sub-space of $\langle V, +, \cdot \rangle$, hence using [theorem: 11.106] we have

$$\operatorname{span}(W) = \operatorname{span}(\operatorname{span}(w))$$

□

Theorem 11.108. *Let $\langle V, +, \cdot \rangle$ be a vector space over a field $\langle F, +, \cdot \rangle$, $W \subseteq V$ and $w \in \operatorname{span}(W)$ then*

$$\operatorname{span}(W) = \operatorname{span}\left(W \bigcup \{w\}\right)$$

Proof. For $w \in \operatorname{span}(W)$ we have either:

***w* ∈ *W*.** Then $W = W \bigcup \{w\}$ so that $\operatorname{span}(W) = \operatorname{span}(W \bigcup \{w\})$

***w* ∉ *W*.** Then as $w \in \operatorname{span}(W)$ there exists a finite $I \subseteq W$ and a $\{\alpha_u\}_{u \in I} \subseteq F$ such that

$$w = \sum_{u \in I} \alpha_u \cdot u \tag{11.31}$$

Let $v \in \operatorname{span}(W \bigcup \{w\})$ then there exists a finite $J \subseteq W \bigcup \{w\}$ and $\{\beta_u\}_{u \in J} \subseteq F$ such that

$$v = \sum_{u \in J} \beta_u \cdot u \tag{11.32}$$

For J we have either:

$w \notin J$. Then $J \subseteq W$ so that $v \in \operatorname{span}(W)$

$w \in J$. Take the finite set $K = (I \bigcup J) \setminus \{w\} \underset{\text{[theorem: 1.25]}}{=} (I \setminus \{w\}) \bigcup (J \setminus \{w\})$ then, as $I \subseteq W, J \subseteq W \bigcup \{w\} \Rightarrow I \setminus \{w\}, J \setminus \{w\} \subseteq W$, we have $K \subseteq W$. Further

$$\begin{aligned} K &= (I \bigcup J) \setminus \{w\} \\ &= ((I \setminus J) \bigcup (I \bigcap J) \bigcup (J \setminus I)) \setminus \{w\} \\ &= ((I \setminus J) \setminus \{w\}) \bigcup ((I \bigcap J) \setminus \{w\}) \bigcup ((J \setminus I) \setminus \{w\}) \\ &\underset{w \notin W \Rightarrow w \notin I \Rightarrow w \in I \bigcap J}{=} (I \setminus J) \bigcup (I \bigcap J) \bigcup ((J \setminus I) \setminus \{w\}) \end{aligned} \tag{11.33}$$

and we have

$$(I \setminus J) \bigcap (I \bigcap J) = \varnothing \wedge (I \setminus J) \bigcap ((J \setminus I) \setminus \{w\}) = \varnothing \wedge (I \bigcap J) \bigcap ((J \setminus I) \setminus \{w\}) = \varnothing \tag{11.34}$$

Define now

$$\{\gamma_u\}_{u \in K} \subseteq F \text{ by } \gamma_u = \begin{cases} \beta_u \text{ if } u \in ((J \setminus I) \setminus \{w\}) \\ \beta_w \cdot \alpha_u + \beta_u \text{ if } u \in I \bigcap J \\ \beta_w \cdot \alpha_u \text{ if } u \in I \setminus J \end{cases}$$

then we have

$$\begin{aligned}
\sum_{u \in K} \gamma_u \cdot u &\underset{\text{[theorem: 11.43]}}{=} \\
\sum_{u \in I \setminus J} \gamma_u \cdot u + \sum_{u \in I \bigcap J} \gamma_u \cdot u + \sum_{u \in (J \setminus I) \setminus \{u\}} \gamma_u \cdot u &= \\
\sum_{u \in I \setminus J} (\beta_w \cdot \alpha_u) \cdot u + \sum_{u \in I \bigcap J} (\beta_w \cdot \alpha_u + \beta_u) \cdot u + \sum_{u \in (J \setminus I) \setminus \{w\}} \beta_u \cdot u &= \\
\sum_{u \in I \setminus J} \beta_w \cdot (\alpha_u \cdot u) + \sum_{u \in I \bigcap J} (\beta_w \cdot (\alpha_u \cdot u) + \beta_u \cdot u) + \sum_{u \in (J \setminus I) \setminus \{w\}} \beta_u \cdot u &\underset{\text{[theorem: 11.75]}}{=} \\
\beta_w \cdot \sum_{u \in I \setminus J} \alpha_u \cdot u + \sum_{u \in I \bigcap J} (\beta_w \cdot (\alpha_u \cdot u) + \beta_u \cdot u) + \sum_{u \in (J \setminus I) \setminus \{w\}} \beta_u \cdot u &\underset{\text{[theorem: 11.38]}}{=} \\
\beta_w \cdot \sum_{u \in I \setminus J} \alpha_u \cdot u + \sum_{u \in I \bigcap J} \beta_w \cdot (\alpha_u \cdot u) + \sum_{i \in I \bigcap J} \beta_u \cdot u + \sum_{u \in (J \setminus I) \setminus \{w\}} \beta_u \cdot u &\underset{\text{[theorem: 11.75]}}{=} \\
\beta_w \cdot \sum_{u \in I \setminus J} \alpha_u \cdot u + \beta_w \cdot \sum_{u \in I \bigcap J} \alpha_u \cdot u + \sum_{i \in I \bigcap J} \beta_u \cdot u + \sum_{u \in (J \setminus I) \setminus \{w\}} \beta_u \cdot u &= \\
\beta_w \cdot \left(\sum_{u \in I \setminus J} \alpha_u \cdot u + \sum_{u \in I \bigcap J} \alpha_u \cdot u \right) + \sum_{i \in I \bigcap J} \beta_u \cdot u + \sum_{u \in (J \setminus I) \setminus \{w\}} \beta_u \cdot u &\underset{\text{[theorem: 11.43]}}{=} \\
\beta_w \cdot \sum_{u \in I} \alpha_u \cdot u + \sum_{i \in I \bigcap J} \beta_u \cdot u + \sum_{u \in (J \setminus I) \setminus \{w\}} \beta_u \cdot u &\underset{\text{[eq: 11.31]}}{=} \\
\beta_w \cdot w + \sum_{i \in I \bigcap J} \beta_u \cdot u + \sum_{u \in (J \setminus I) \setminus \{w\}} \beta_u \cdot u &\underset{w \notin I \bigcap J}{=} \\
\beta_w \cdot w + \sum_{i \in (I \bigcap J) \setminus \{w\}} \beta_u \cdot u + \sum_{u \in (J \setminus I) \setminus \{w\}} \beta_u \cdot u &\underset{\text{[theorem: 11.43]}}{=} \\
\beta_w \cdot w + \sum_{i \in J \setminus \{w\}} \beta_u \cdot u &\underset{\text{[theorem: 11.34]}}{=} \\
\sum_{i \in \{w\}} \beta_u \cdot u + \sum_{i \in J \setminus \{w\}} \beta_u \cdot u &\underset{\text{[theorem: 11.43]}}{=} \\
\sum_{i \in J} \beta_u \cdot u &\underset{\text{[eq: 11.32]}}{=} \\
v &
\end{aligned}$$

proving, as K is finite and $K \subseteq W$, that $v \in \operatorname{span}(W)$.

So in all cases we have $v \in \operatorname{span}(W)$ proving that $\operatorname{span}(W \bigcup \{w\}) \subseteq \operatorname{span}(W)$. As further $W \subseteq W \bigcup \{w\}$ we have by [theorem: 11.104] that $\operatorname{span}(W) \subseteq \operatorname{span}(W \bigcup \{w\})$ proving that $\operatorname{span}(W) = \operatorname{span}(W \bigcup \{w\})$.

So in all cases we have

$$\operatorname{span}(W) = \operatorname{span}\left(W \bigcup \{w\}\right)$$

completing the proof. □

11.3.2.3 Linear (in)dependent sets

Definition 11.109. (Linear Dependency) *Let $\langle V, +, \cdot \rangle$ be a vector space over the field $\langle F, +, \cdot \rangle$ and $W \subseteq V$ then W* ***is linear dependent*** *if there exists a finite $I \subseteq W$ and a $\{\alpha_w\}_{w \in I} \subseteq F$ satisfying $\exists u \in I$ with $\alpha_u \neq 0$ such that $\sum_{u \in I} \alpha_u \cdot u = 0$.*

Example 11.110. Let $\langle V, +, \cdot \rangle$ be a vector space over the field $\langle F, +, \cdot \rangle$ and $W \subseteq V$ such that $0 \in W$ then W is linear dependent.

Proof. Take $I = \{0\}$ and $\{\alpha_u\}_{u \in \{0\}}$ by $\alpha_0 = 1$ then we have $\sum_{u \in \{0\}} \alpha_u \cdot u \underset{[\text{theorem: } 11.34]}{=} \alpha_0 \cdot 0 = 0$ proving that W is linear dependent. □

Linear dependent sets can also be described as sets where one of the vector can be written as a linear combination of some other vectors, this is proved in the next theorem.

Theorem 11.111. *Let $\langle V, +, \cdot \rangle$ be a vector space over the field $\langle F, +, \cdot \rangle$ and $W \subseteq V$ then*

$$W \text{ is linear dependent}$$
$$\Updownarrow$$
$$\exists w \in W \text{ so that there exists a finite } I \subseteq W \setminus \{w\} \text{ and a } \{\alpha_u\}_{u \in I} \subseteq F \text{ such that } w = \sum_{u \in I} \alpha_u \cdot u$$

Proof.

⇒. As W is a linear dependent there exists a finite $J \subseteq W$ and a $\{\beta_u\}_{u \in J} \subseteq F$ satisfying $\exists w \in J$ with $\beta_w \neq 0$ such that $\sum_{u \in J} \beta_u \cdot u = 0$. So we have

$$\begin{aligned} 0 &= \sum_{u \in J} \beta_u \cdot u \\ &\underset{[\text{theorem: } 11.43]}{=} \sum_{u \in J \setminus \{w\}} \beta_u \cdot u + \sum_{u \in \{w\}} \beta_u \cdot u \\ &\underset{[\text{theorem: } 11.34]}{=} \sum_{u \in J \setminus \{w\}} \beta_u \cdot u + \beta_w \cdot w \end{aligned}$$

proving as $\beta_w \neq 0$ that

$$\begin{aligned} w &= (\beta_w)^{-1} \cdot \left(-\sum_{u \in J \setminus \{w\}} \beta_u \cdot u\right) \\ &\underset{[\text{theorem: } 11.40]}{=} (\beta_w)^{-1} \cdot \sum_{u \in J \setminus \{w\}} -(\beta_u \cdot u) \\ &= (\beta_w)^{-1} \cdot \sum_{u \in J \setminus \{w\}} (-\beta_u) \cdot u \\ &\underset{[\text{theorem: } 11.75]}{=} \sum_{u \in J \setminus \{w\}} ((\beta_w)^{-1}((-\beta_u) \cdot u)) \\ &= \sum_{u \in J \setminus \{w\}} ((\beta_w)^{-1}(-\beta_u)) \cdot u \end{aligned}$$

So if we define $I = J \setminus \{w\}$ and $\{\alpha_u\}_{u \in I}$ by $\beta_u = (\beta_w)^{-1}(-\beta_u)$ then $I \subseteq W \setminus \{w\}$ and

$$\sum_{u \in I} \alpha_u \cdot u = w$$

⇐. By the hypothesis there exists a $w \in W$, a finite $I \subseteq W \setminus \{w\}$ and a $\{\alpha_u\}_{u \in I} \subseteq F$ such that

$$w = \sum_{u \in I} \alpha_u \cdot u$$

Let $J = I \bigcup \{w\}$ then J is finite and $J \subseteq W$ and define

$$\{\beta_u\}_{u \in J} \text{ by } \beta_u = \begin{cases} -1 \text{ if } u = w \\ \alpha_u \text{ if } u \in I \end{cases}$$

then

$$\begin{aligned}
\sum_{u \in J} \beta_u \cdot u & \underset{[\text{theorem: } 11.43]}{=} \sum_{u \in I} \beta_u \cdot u + \sum_{u \in \{w\}} \beta_u \cdot u \\
& = \sum_{u \in I} \alpha_u \cdot u + \sum_{u \in \{w\}} \beta_u \cdot u \\
& \underset{[\text{theorem: } 11.34]}{=} \sum_{u \in I} \alpha_u \cdot u + \beta_w \cdot w \\
& = \sum_{u \in I} \alpha_u \cdot u + (-1) \cdot w \\
& = w + (-w) \\
& = 0
\end{aligned}$$

which as $\beta_w = -1 \neq 0$ proves that W is linear dependent. □

For finite sets we have a easier condition for linear dependency.

Theorem 11.112. *Let $\langle V, +, \cdot \rangle$ be a vector space over the field $\langle F, +, \cdot \rangle$, $n \in \mathbb{N}$ and $W = \{w_1, \ldots, w_n\} \subseteq V$ a finite set then we have the following equivalences:*

1. *W is linear dependent*
2. *$\exists \{\alpha_u\}_{u \in W} \subseteq F$ satisfying $\exists u \in W$ with $\alpha_u \neq 0$ such that $\sum_{w \in W} \alpha_u \cdot u = 0$*
3. *$\exists u \in W$ and a $\{\alpha_w\}_{w \in W \setminus \{u\}} \subseteq F$ such that $u = \sum_{w \in W \setminus \{u\}} \alpha_q \cdot w$*
4. *$\exists \{\alpha_i\}_{i \in \{1, \ldots, n\}} \subseteq F$ satisfying $\exists i \in \{1, \ldots, n\}$ with $\alpha_i \neq 0$ such that $\sum_{i \in \{1, \ldots, n\}} \alpha_i \cdot w_i = 0$.*
5. *$\exists k \in \{1, \ldots, n\}$ and a $\{\alpha_i\}_{i \in \{1, \ldots, n\} \setminus \{k\}} \subseteq F$ such that $w_k = \sum_{i \in \{1, \ldots, n\}\{k\}} \alpha_i \cdot w_i$*

Proof.

1 ⇒ 2. As W is linear dependent there exist by [definition: 11.109] a finite set $I \subseteq W$ and a $\{\beta_u\}_{u \in I} \subseteq F$ satisfying $\exists u_0 \in I$ with $\beta_{u_0} \neq 0$ such that $\sum_{u \in I} \beta_u \cdot u = 0$. Define for the finite set W

$$\{\alpha_u\}_{u \in W} \subseteq F \text{ by } \alpha_w = \begin{cases} 0 \text{ if } w \in W \setminus I \\ \beta_w \text{ if } w \in I \end{cases}$$

then we have for $u_0 \in I \subseteq W$ that $\alpha_{u_0} = \beta_{u_0} \neq 0$ and

$$\begin{aligned}
\sum_{u \in W} \alpha_u \cdot u & \underset{[\text{theorem: } 11.43]}{=} \sum_{u \in W \setminus I} \alpha_u \cdot u + \sum_{u \in I} \alpha_u \cdot u \\
& = \sum_{u \in W \setminus I} 0 \cdot u + \sum_{u \in I} \beta_u \cdot u \\
& \underset{[\text{theorem: } 11.37]}{=} 0 + \sum_{u \in I} \beta_u \cdot u \\
& = 0
\end{aligned}$$

2 ⇒ 1. As W is finite and $W \subseteq W$ this follows from the definition of linear dependency [see definition: 11.109]

2 ⇒ 3. By the hypothesis there exist a $\{\beta_w\}_{w \in W} \subseteq W$ satisfying $\exists u \in W$ with $\beta_u \neq 0$ such that $\sum_{w \in W} \beta_w \cdot w = 0$. Hence

$$\begin{aligned}
0 & = \sum_{w \in W} \beta_w \cdot w \\
& \underset{[\text{theorem: } 11.43]}{=} \sum_{w \in W \setminus \{u\}} \beta_w \cdot w + \sum_{w \in \{u\}} \beta_w \cdot w \\
& \underset{[\text{theorem: } 11.34]}{=} \sum_{w \in W \setminus \{u\}} \beta_w \cdot w + \beta_u \cdot u
\end{aligned}$$

which, as $\beta_u \neq 0$, gives that

$$
\begin{array}{rcl}
u & = & (\beta_u)^{-1} \cdot \left(-\sum_{w \in W \setminus \{u\}} \beta_w \cdot w\right) \\
& \underset{\text{[theorem: 11.40]}}{=} & (\beta_u)^{-1} \cdot \left(\sum_{w \in W \setminus \{u\}} -(\beta_w \cdot w)\right) \\
& \underset{\text{[theorem: 11.75]}}{=} & \sum_{w \in W \setminus \{u\}} (\beta_u)^{-1} \cdot (-\beta_w) \cdot w
\end{array}
$$

So if we define $\{\alpha_w\}_{w \in W \setminus \{u\}} \subseteq F$ by $\alpha_w = (\beta_u)^{-1} \cdot (-\beta_w)$ then

$$u = \sum_{w \in W \setminus \{u\}} \alpha_w \cdot w$$

proving (3).

3 ⇒ 2. By the hypothesis $\exists u \in W$ and a $\{\beta_w\}_{w \in W \setminus \{u\}} \subseteq F$ such that $u = \sum_{w \in W \setminus \{u\}} \beta_w \cdot w$. Define $\{\alpha_w\}_{w \in W} \subseteq F$ by $\alpha_w = \left\{ \begin{array}{l} -1 \text{ if } w = u \\ \beta_w \text{ if w‹W\{u} \end{array} \right.$ then

$$
\begin{array}{rcl}
\sum_{w \in W} \alpha_w \cdot w & \underset{\text{[theorem: 11.43]}}{=} & \sum_{w \in W \setminus \{u\}} \alpha_w \cdot w + \sum_{w \in \{u\}} \alpha_w \cdot w \\
& \underset{\text{[theorem: 11.34]}}{=} & \sum_{w \in W \setminus \{u\}} \alpha_w \cdot w + \alpha_u \cdot u \\
& = & u + (-1) \cdot u \\
& = & 0
\end{array}
$$

which as $\alpha_u = -1 \neq 0$ proves (2).

2 ⇒ 4. By the hypothesis $\exists \{\beta_u\}_{u \in W} \subseteq F$ satisfying $\exists u_0 \in W$ with $\beta_{u_0} \neq 0$ such that $\sum_{u \in W} \beta_u \cdot u = 0$. Using [theorem: 11.92] there exist a $\{\alpha_i\}_{i \in \{1, \ldots, n\}} \subseteq F$ such that

$$\sum_{i \in \{1, \ldots, n\}} \alpha_i \cdot w_i = \sum_{u \in W} \beta_u \cdot u = 0$$

and

$$\{\beta_u | u \in W\} = \{\alpha_i | i \in \{1, \ldots, n\}\}$$

As $\beta_{u_0} \in \{\beta_u | u \in W\} = \{\alpha_i | i \in \{1, \ldots, n\}\}$ there exists a $i_0 \in \{1, \ldots, n\}$ such that $\alpha_{i_0} = \beta_{u_0} \neq 0$. Hence we found a $\{\alpha_i\}_{i \in \{1, \ldots, n\}} \subseteq F$ satisfying $\exists i_0 \in \{1, \ldots, N\}$ with $\alpha_{i_0} \neq 0$ such that $\sum_{i \in \{1, \ldots, n\}} \alpha_i \cdot w$

4 ⇒ 2. By the hypothesis there exist a $\{\beta_i\}_{i \in \{1, \ldots, n\}} \subseteq F$ satisfying $\exists i_0 \in \{1, \ldots, n\}$ with $\beta_{i_0} \neq 0$ such that $\sum_{i \in \{1, \ldots, n\}} \beta_i \cdot w_i = 0$. Using [theorem: 11.92] there exist a $\{\alpha_u\}_{u \in W} \subseteq F$ such that

$$\sum_{u \in W} \alpha_u \cdot u = \sum_{i \in \{1, \ldots, n\}} \beta_i \cdot w_i = 0$$

and

$$\{\beta_i | i \in \{1, \ldots, n\}\} = \{\alpha_u | u \in W\}.$$

As $\beta_{i_0} \in \{\beta_i | i \in \{1, \ldots, n\}\} = \{\alpha_u | u \in W\}$ there exist a $u_0 \in W$ such that $\alpha_{u_0} = \beta_{i_0} \neq 0$. Hence we found a $\{\alpha_u\}_{u \in W} \subseteq F$ satisfying $\exists u_0 \in W$ with $\alpha_{u_0} \neq 0$ such that $\sum_{u \in W} \alpha_u \cdot u = 0$.

4 ⇒ 5. By the hypothesis there exist a $\{\alpha_i\}_{i \in \{1, \ldots, n\}} \subseteq F$ satisfying $\exists k \in \{1, \ldots, n\}$ with $\alpha_k \neq 0$ such that $\sum_{i \in \{1, \ldots, n\}} \alpha_i \cdot w_i = 0$. Then we have

$$
\begin{array}{rcl}
0 & = & \sum_{i \in \{1, \ldots, n\}} \alpha_i \cdot w_i \\
& \underset{\text{[theorem: 11.43]}}{=} & \sum_{i \in \{1, \ldots, n\} \setminus \{k\}} \alpha_i \cdot w_i + \sum_{i \in \{k\}} \alpha_i \cdot w_i \\
& \underset{\text{[theorem: 11.34]}}{=} & \sum_{i \in \{1, \ldots, n\} \setminus \{k\}} \alpha_i \cdot w_i + \alpha_k \cdot w_k
\end{array}
$$

which as $\alpha_k \neq 0$ gives

$$\begin{array}{rcl} w_k & = & (\alpha_k)^{-1} \cdot \left(- \sum_{i \in \{1, \ldots, n\} \setminus \{k\}} \alpha_i \cdot w_i \right) \\ & \underset{[\text{theorem: } 11.40]}{=} & (\alpha_k)^{-1} \cdot \sum_{i \in \{1, \ldots, n\} \setminus \{k\}} -(\alpha_i \cdot w_i) \\ & \underset{[\text{theorem: } 11.75]}{=} & \sum_{i \in \{1, \ldots, n\} \setminus \{k\}} (\alpha_k)^{-1} \cdot (-(\alpha_i \cdot w_i)) \\ & = & \sum_{i \in \{1, \ldots, n\} \setminus \{k\}} ((\alpha_k)^{-1} \cdot (-\alpha_i)) \cdot w_i \end{array}$$

Hence if $\{\beta_i\}_{i \in \{1, \ldots, n\} \setminus \{k\}}$ is defined by $\beta_i = (\alpha_k)^{-1} \cdot (-\alpha_i)$ we have

$$w_k = \sum_{i \in \{1, \ldots, n\} \setminus \{k\}} \beta_i \cdot w_i$$

5 ⇒ 4. By the hypothesis there $\exists k \in \{1, \ldots, n\}$ and a $\{\beta_i\}_{i \in [1, \ldots, n] \setminus \{k\}} \subseteq F$ such that

$$w_k = \sum_{i \in \{1, \ldots, n\} \setminus \{k\}} \beta_i \cdot w_i$$

Define then

$$\{\alpha_i\}_{i \in \{1, \ldots, n\}} \subseteq F \text{ by } \alpha_i = \begin{cases} -1 \text{ if } i = k \\ \beta_i \text{ if } i \in \{1, \ldots, n\} \setminus \{k\} \end{cases}$$

then $\alpha_k = -1 \neq 0$ and

$$\begin{array}{rcl} \sum_{i \in \{1, \ldots, n\}} \alpha_i \cdot w_i & \underset{[\text{theorem: } 11.43]}{=} & \sum_{i \in \{1, \ldots, n\} \setminus \{k\}} \alpha_i \cdot w_i + \sum_{i \in \{k\}} \alpha_i \cdot w_i \\ & \underset{[\text{theorem: } 11.34]}{=} & \sum_{i \in \{1, \ldots, n\} \setminus \{k\}} \alpha_i \cdot w_i + \alpha_k \cdot w_k \\ & = & w_k + (-1) \cdot w_k \\ & = & 0 \end{array}$$

□

The opposite of a linear dependent set is a linear independent set.

Definition 11.113. (Linear Independence) *Let $\langle V, +, \cdot \rangle$ be a vector space over the field $\langle F, +, \cdot \rangle$ and $W \subseteq V$ then W* ***is linear independent*** *if W is* ***not linear dependent***.

We have the following equivalent definition of a linear independent set.

Theorem 11.114. *Let $\langle V, +, \cdot \rangle$ be a vector space over the field $\langle F, +, \cdot \rangle$ and $W \subseteq V$ then*

W is linear independent

$$\Updownarrow$$

If $\{\alpha_u\}_{u \in I} \subseteq F$, I a finite subset of W satisfies $\sum_{u \in I} \alpha_u \cdot u = 0$ then $\forall u \in I$ we have $\alpha_u = 0$

Proof.

⇒. As W is independent we have by definition that W is not dependent. Let $\{\alpha\}_{u \in I} \subseteq W$, I a finite subset of W that satisfies $\sum_{u \in I} \alpha_u \cdot u = 0$. Assume that $\exists u \in W$ with $\alpha_u \neq 0$ then by definition W is linear dependent and we have a contradiction, so we must have that $\forall u \in I$ $\alpha_u = 0$.

⇐. Assume that W is linear dependent, then there exists a finite $I \subseteq W$ and a $\{\alpha_u\}_{u \in I} \subseteq F$ satisfying $\exists w \in I$ with $\alpha_w \neq 0$ such that $\sum_{u \in I} \alpha_u \cdot u = 0$ which contradict the hypothesis which says that $\forall w \in I$ $\alpha_w = 0$. □

For finite sets we can simplify the conditions for linear independency.

Theorem 11.115. *Let $\langle V,+,\cdot\rangle$ be a vector space over the field $\langle F,+,\cdot\rangle$ and $W=\{w_1,\ldots,w_n\}\subseteq V$ a* ***finite*** *set then we have the following equivalences:*

1. *W is linear independent*
2. *$\forall\{\alpha_u\}_{u\in W}\subseteq F$ such that $\sum_{w\in W}\alpha_u\cdot u=0$ we have $\forall u\in W$ that $\alpha_u=0$*
3. *$\forall\{\alpha_i\}_{i\in\{1,\ldots,n\}}\subseteq F$ such that $\sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot w_i=0$ we have $\forall i\in\{1,\ldots,n\}$ $\alpha_i=0$.*

Proof.

1 ⇒ 2. As W is finite and $W\subseteq W$ this follows from [theorem: 11.114].

2 ⇒ 1. Let $I\subseteq W$ be a finite set and $\{\alpha_u\}_{u\in I}\subseteq F$ satisfying $\sum_{u\in I}\alpha_u\cdot u=0$. Define

$$\{\beta_u\}_{u\in W}\subseteq F \text{ by } \beta_u=\begin{cases}0 \text{ if } u\in W\setminus I\\ \alpha_u \text{ if } u\in I\end{cases}$$

then we have as W is finite

$$\begin{aligned}\sum_{u\in W}\beta_u\cdot u &\underset{[\text{theorem: } 11.43]}{=} \sum_{u\in W\setminus I}\beta_u\cdot u+\sum_{u\in I}\beta_u\cdot u\\ &= \sum_{u\in W\setminus I}0\cdot u+\sum_{u\in I}\alpha_u\cdot u\\ &\underset{[\text{theorem: } 11.37]}{=} \sum_{u\in I}\alpha_u\cdot u\\ &= 0\end{aligned}$$

By the hypothesis we have that $\forall u\in W$ $\beta_u=0$, so if $u\in I\subseteq W$ we have $\alpha_u=\beta_u=0$. Hence we have that W is linear independent.

2 ⇒ 3. Let $\{\alpha_i\}_{i\in\{1,\ldots,n\}}\subseteq F$ such that $\sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot w_i=0$. Using [theorem: 11.92] there exist a $\{\beta_u\}_{u\in W}\subseteq F$ such that

$$\sum_{u\in W}\beta_u\cdot u=\sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot w_i=0$$

and $\{\alpha_i|i\in\{1,\ldots,n\}\}=\{\beta_u|u\in W\}$. By the hypothesis we have $\forall u\in W$ that $\beta_u=0$. If $i\in\{1,\ldots,n\}$ then $\alpha_i\in\{\beta_u|u\in W\}$ so that $\exists u\in W$ such that $\alpha_i=\beta_u=0$. Hence $\forall i\in\{1,\ldots,n\}$ we have that $\alpha_i=0$.

3 ⇒ 2. Let $\{\alpha_u\}_{u\in W}\subseteq F$ such that $\sum_{u\in W}\alpha_u\cdot u=0$. Using [theorem: 11.92] there exist a $\{\beta_i\}_{i\in\{1,\ldots,n\}}\subseteq F$ such that

$$\sum_{i\in\{1,\ldots,n\}}\beta_i=\sum_{u\in W}\alpha_u\cdot u=0$$

and $\{\alpha_u|u\in W\}=\{\beta_i|i\in\{1,\ldots,n\}\}$. By the hypothesis we have that $\forall i\in\{1,\ldots,n\}$ $\beta_i=0$. If $u\in W$ then $\alpha_u\in\{\beta_i|i\in\{1,\ldots,n\}\}$ so there exists a $i\in\{1,\ldots,n\}$ such that $\alpha_u=\beta_i=0$. Hence $\forall u\in W$ we have $\alpha_u=0$. □

Example 11.116. Let $\langle V,+,\cdot\rangle$ be a vector space over the field $\langle F,+,\cdot\rangle$ then $\varnothing$ is linear independent.

Proof. Let $I\subseteq\varnothing$, I a finite set then $I=\varnothing$ so that if $\{\alpha_u\}_{u\in I}\subseteq F$ with $\sum_{u\in\varnothing}\alpha_u\cdot u=0$ we have that $\forall u\in\varnothing$ $\alpha_u=0$ is satisfied vacuously. □

Example 11.117. Let $\langle V,+,\cdot\rangle$ be a vector space over the field $\langle F,+,\cdot\rangle$ and $v\in V$ such that $v\neq 0$ then $\{v\}$ is a linear independent set.

Proof. Let $\{\alpha_u\}_{u\in\{v\}}\subseteq F$ be such that $\sum_{u\in\{v\}}\alpha_u\cdot u=0$ then by [theorem: 11.34] $\alpha_v\cdot v=0$. Assume that $\alpha_v\neq 0$ then $v=(\alpha_v)^{-1}\cdot(\alpha_v\cdot v)=0$ contradicting $v\neq 0$. Hence $\alpha_v=0$ or $\forall u\in\{v\}$ $\alpha_u=0$ proving by [theorem: 11.115] that the finite set $\{v\}$ is linear dependent. □

Theorem 11.118. *Let $\langle V,+,\cdot\rangle$ be a vector space over the field $\langle F,+,\cdot\rangle$ and $U,W\subseteq V$ with $U\subseteq W$ then we have*

1. *If W is linear independent then U is is linear independent, in other words every subset of a linear independent set is linear independent.*
2. *If U is linear dependent then W is linear dependent, in other words every super set of a linear dependent set is linear dependent.*

Proof.

1. Let $I\subseteq U$ be a finite set then as $U\subseteq V$ we have $I\subseteq V$. So if $\{\alpha_u\}_{u\in I}\subseteq F$ is such that $\sum_{\in I}\alpha_u\cdot u=0$ we have as V is linear independent that $\forall u\in I$ we have $\alpha_u=0$. Hence U is linear independent.

2. As U is linear dependent there exists a finite $I \subseteq U$ and a $\{\alpha_u\}_{u \in I} \subseteq F$ not all zeroes such that $\sum_{u \in I} \alpha_u \cdot u = 0$. As $U \subseteq V$ we have $I \subseteq V$ proving that V is linear dependent. □

The following lemma shows, how given a finite linear independent set $\{e_1, \ldots, e_n\}$ and a set that spans the vector space we can create a new spanning set by replacing n vectors in the spanning set by the vectors in $\{e_1, \ldots, e_n\}$. This lemma will be crucial in introducing the concept of the dimension of a vector space.

Lemma 11.119. (Steinitz Lemma) *Let $\langle V, +, \cdot \rangle$ be a vector space over a field $\langle F, +, \cdot \rangle$, $n \in \mathbb{N}_0$, $E = \{e_1, \ldots, e_n\} \subseteq V$ a linear independent finite set and $T \subseteq V$ such that $\operatorname{span}(T) = V$ then there exist a $T_n \subseteq T$ with $\operatorname{card}(T_n) = n$ such that*

$$\operatorname{span}\big((T \setminus T_n) \bigcup \{e_1, \ldots, e_n\}\big) = V$$

Proof. We prove this by induction so let

$$S = \big\{k \in \mathbb{N}_0 | \text{If } 0 \leqslant k \leqslant n \text{ then } \exists T_k' \subseteq T \text{ with } \operatorname{card}(T_k') = k \text{ such that } \operatorname{span}\big((T \setminus T_k') \bigcup \{e_1, \ldots, e_k\}\big) = V\big\}$$

then we have:

$0 \in S$. Take $T_0' = \varnothing$ then $\operatorname{card}(T_0') = 0$ and

$$V = \operatorname{span}(T) = \operatorname{span}\big((T \setminus \varnothing) \bigcup \varnothing\big) \underset{\text{[example: 11.85]}}{=} \operatorname{span}\big((T \setminus T_0') \bigcup \{e_1, \ldots, e_0\}\big)$$

proving that $0 \in S$.

$k \in S \Rightarrow k+1 \in S$. If $0 \leqslant k+1 \leqslant n$ then $k < n$ and, as $k \in S \subseteq \mathbb{N}_0$, $0 \leqslant k < n$. So using the fact that $k \in S$ there exist a $T_k' \subseteq T$ with $\operatorname{card}(T_k') = k$ such that $\operatorname{span}((T \setminus T_k') \bigcup \{e_1, \ldots, e_k\}) = V$ Define

$$T_k = (T \setminus T_k') \bigcup \{e_1, \ldots, e_k\} \tag{11.35}$$

then

$$\operatorname{span}(T_k) = V \tag{11.36}$$

As $\{e_1, \ldots, e_{k+1}\} \subseteq \{e_1, \ldots, e_n\}$ a linear independent set it follows from [theorem: 11.118] that

$$\{e_1, \ldots, e_{k+1}\} \text{ is linear independent} \tag{11.37}$$

As $e_{k+1} \in V = \operatorname{span}(T_k)$ we have by [theorem: 11.101] that

$$\text{there exists a finite } I \subseteq T_k \text{ and } \{\lambda_u\}_{u \in I} \text{ such that } e_{k+1} = \sum_{u \in I} \lambda_u \cdot u \tag{11.38}$$

Assume that $I \subseteq \{e_1, \ldots, e_k\} \underset{\text{[theorem: 11.86]}}{=} \{e_1, \ldots, e_{k+1}\} \setminus \{e_{k+1}\}$ then we have by the above and [theorem: 11.111] that $\{e_1, \ldots, e_{k+1}\}$ is linear dependent contradicting [eq: 11.37]. Hence $I \nsubseteq \{e_1, \ldots, e_{k+1}\}$ so that

$$\exists t \in I \text{ with } t \notin \{e_1, \ldots, e_{k+1}\} \tag{11.39}$$

As $t \in I \subseteq T_k \underset{\text{[eq: 11.35]}}{=} (T \setminus T_k') \bigcup \{e_1, \ldots, e_k\} \subseteq (T \setminus T_k') \bigcup \{e_1, \ldots, e_{k+1}\}$ we must have by the above that $t \in T \setminus T_k'$ hence $t \notin T_k'$. Using [theorem: 10.91] it follows then that $\operatorname{card}(T_k' \bigcup \{t\}) = \operatorname{card}(T_k') + 1 = k + 1$. To summarize we have

$$t \in T \text{ and } \operatorname{card}(T_{k+1}') = k + 1 \text{ where } T_{k+1}' = T_k' \bigcup \{t\} \tag{11.40}$$

Further

$$\begin{aligned}
T_k \setminus \{t\} &\underset{\text{[eq: 11.35]}}{=} \big((T \setminus T_k') \bigcup \{e_1, \ldots, e_k\}\big) \setminus \{t\} \\
&= ((T \setminus T_k') \setminus \{t\}) \bigcup (\{e_1, \ldots, e_k\} \setminus \{t\}) \\
&\underset{\text{[eq: 11.39]}}{=} ((T \setminus T_k') \setminus \{t\}) \bigcup \{e_1, \ldots, e_k\} \\
&\underset{\text{[theorem: 1.31]}}{=} (T \setminus (T_k' \bigcup \{t\})) \bigcup \{e_1, \ldots, e_k\}
\end{aligned} \tag{11.41}$$

Let

$$T_{k+1} = (T \setminus T_{k+1}') \bigcup \{e_1, \ldots, e_{k+1}\} \tag{11.42}$$

Then we have

$$\begin{array}{rcl} T_{k+1} & = & (T\setminus T'_{k+1})\bigcup \{e_1,\ldots,e_{k+1}\} \\ & = & (T\setminus T'_{k+1})\bigcup (\{e_1,\ldots,e_k\}\bigcup \{e_{k+1}\}) \\ & = & ((T\setminus T'_{k+1})\bigcup \{e_1,\ldots,e_k\})\bigcup \{e_{k+1}\} \\ & \underset{[\text{eq: } 11.40]}{=} & ((T\setminus (T'_k\bigcup \{t\}))\bigcup \{e_1,\ldots,e_k\})\bigcup \{e_{k+1}\} \\ & \underset{[\text{eq: } 11.41]}{=} & (T_k\setminus\{t\})\bigcup \{e_{k+1}\} \end{array}$$

proving that

$$T_{k+1} = (T_k\setminus\{t\})\bigcup \{e_{k+1}\} \tag{11.43}$$

Further

$$\begin{array}{rcl} e_{k+1} & \underset{[\text{eq: } 11.38]}{=} & \sum_{u\in I} \lambda_u\cdot u \\ & \underset{[\text{theorem: } 11.43]}{=} & \sum_{u\in I\setminus\{t\}} \lambda_u\cdot u + \sum_{u\in\{t\}} \lambda_u\cdot u \\ & \underset{[\text{theorem: } 11.34]}{=} & \sum_{u\in I\setminus\{t\}} \lambda_u\cdot u + \lambda_t\cdot t \end{array} \tag{11.44}$$

As $(I\setminus\{t\})\bigcup \{e_{k+1}\} \subseteq_{[\text{eq: } 11.38]} (T_k\setminus\{t\})\bigcup \{e_{k+1}\} \underset{[\text{eq: } 11.43]}{=} T_{k+1}$ we have, taking in account that I is finite that

$$J\subseteq T_{k+1} \text{ and } J \text{ is finite where } J = (I\setminus\{t\})\bigcup \{e_{k+1}\} \tag{11.45}$$

For e_{k+1} we have either:

$\boldsymbol{e_{k+1}\in I\setminus\{t\}}$**.** Then by [eq: 11.44]

$$\begin{array}{rcl} e_{k+1} & = & \sum_{u\in I\setminus\{t\}} \lambda_u\cdot u+\lambda_t\cdot t \\ & \underset{[\text{theorem: } 11.43]}{=} & \sum_{u\in I\setminus\{t,e_{k+1}\}} \lambda_u\cdot u + \sum_{u\in\{e_{k+1}\}} \lambda_u\cdot u+\lambda_t\cdot t \\ & \underset{[\text{theorem: } 11.34]}{=} & \sum_{u\in I\setminus\{t,e_{k+1}\}} \lambda_u\cdot u+\lambda_{e_{k+1}}\cdot e_{k+1}+\lambda_t\cdot t \end{array}$$

as $\{\lambda_u\}_{u\in I}\subseteq F\setminus\{0\}$ we have $\lambda_t\neq 0$, so

$$t = ((\lambda_t)^{-1}\cdot(1-\lambda_{k+1}))\cdot e_{k+1} - (\lambda_t)^{-1}\cdot \sum_{u\in I\setminus\{t,e_{k+1}\}} \lambda_u\cdot u \tag{11.46}$$

Then for $J \underset{[\text{eq: } 11.45]}{=} (I\setminus\{t\})\bigcup e_{k+1} = (I\setminus\{t,e_{k+1}\})\bigcup \{e_{k+1}\}$ we can define

$$\{\alpha_u\}_{u\in J}\subseteq F \text{ by } \alpha_u = \begin{cases} (\lambda_t)^{-1}\cdot(1-\lambda_{k+1}) \text{ if } u = e_{k+1} \\ (-(\lambda_t)^{-1})\cdot\lambda_u \text{ if } u\in I\setminus\{t,e_{k+1}\} = J\setminus\{e_{k+1}\} \end{cases}$$

giving:

$$\begin{array}{rcl} \sum_{u\in J}\alpha_u\cdot u & \underset{[\text{theorem: } 11.43]}{=} & \sum_{u\in J\setminus\{e_{k+1}\}} \alpha_u\cdot u + \sum_{u\in\{e_{k+1}\}} \alpha_u\cdot u \\ & \underset{[\text{theorem: } 11.34]}{=} & \sum_{u\in J\setminus\{e_{k+1}\}} \alpha_u\cdot u+\alpha_{e_{k+1}}\cdot e_{k+1} \\ & = & \sum_{u\in J\setminus\{e_{k+1}\}} ((-(\lambda_t)^{-1})\cdot\lambda_u)\cdot u+(\lambda_t)^{-1}\cdot(1-\lambda_{k+1})\cdot e_{k+1} \\ & \underset{[\text{theorem: } 11.75]}{=} & (-(\lambda_t)^{-1})\cdot \sum_{u\in J\setminus\{e_{k+1}\}} \lambda_u\cdot u+(\lambda_t)^{-1}\cdot(1-\lambda_{k+1})\cdot e_{k+1} \\ & \underset{[\text{eq: } 11.46]}{=} & t \end{array}$$

which proves, as by [eq: 11.45] $J\subseteq T_{k+1}$, that

$$t\in \operatorname{span}(T_{k+1})$$

$\boldsymbol{e_{k+1}\notin I\setminus\{t\}}$**.** Using [eq: 11.44] we have, as $\{\lambda_u\}_{u\in I}\subseteq F\setminus\{0\}$ implies $\lambda_t\neq 0$, that

$$t=(\lambda_t)^{-1}\cdot e_{k+1}-(\lambda_t)^{-1}\cdot\sum_{u\in I\setminus\{t\}}\lambda_u\cdot u \tag{11.47}$$

Further from $e_{k+1}\notin I\setminus\{t\}$ it follows that

$$J\setminus\{e_{k+1}\}\underset{[\text{eq: }11.45]}{=}\left((I\setminus\{t\})\bigcup\{e_{k+1}\}\right)\setminus\{e_{k+1}\}=(I\setminus\{t\})\setminus\{e_{k+1}\}=I\setminus\{t\}$$

giving

$$J\setminus\{e_{k+1}\}=I\setminus\{t\} \tag{11.48}$$

As $e_{k+1}\in(I\setminus\{t\})\bigcup\{e_{k+1}\}\underset{[\text{eq: }11.45]}{=}J$ we can define

$$\{\alpha_u\}_{u\in J}\subseteq F \text{ by } \alpha_u=\begin{cases}(\lambda_t)^{-1} \text{ if } u=e_{k+1}\\ (-(\lambda_t)^{-1})\cdot\lambda_u \text{ if } u\in J\setminus\{e_{n+1}\}\end{cases}$$

then we have:

$$\begin{aligned}
\sum_{u\in J}\alpha_u\cdot u &\underset{[\text{theorem: }11.43]}{=} \sum_{u\in J\setminus\{e_{k+1}\}}\alpha_u\cdot u+\sum_{u\in\{e_{k+1}\}}\alpha_u\cdot u\\
&\underset{[\text{theorem: }11.34]}{=} \sum_{u\in J\setminus\{e_{k+1}\}}\alpha_u\cdot u+\alpha_{e_{k+1}}\cdot e_{k+1}\\
&= \sum_{u\in J\setminus\{e_{k+1}\}}((-(\lambda_t)^{-1})\cdot\lambda_u)\cdot u+(\lambda_t)^{-1}\cdot e_{k+1}\\
&\underset{[\text{theorem: }11.75]}{=} (-(\lambda_t)^{-1})\cdot\sum_{u\in J\setminus\{e_{k+1}\}}\lambda_u\cdot u+(\lambda_t)^{-1}\cdot e_{k+1}\\
&\underset{[\text{eq: }11.48]}{=} (-(\lambda_t)^{-1})\cdot\sum_{u\in I\setminus\{t\}}\lambda_u\cdot u+(\lambda_t)^{-1}\cdot e_{k+1}\\
&\underset{[\text{eq: }11.47]}{=} t
\end{aligned}$$

proving, as by [eq: 11.45] $J\subseteq T_{k+1}$, that

$$t\in\operatorname{span}(T_{k+1})$$

So in all cases we have $t\in\operatorname{span}(T_{k+1})$ which using [theorem: 11.108] proves that

$$\operatorname{span}\left(T_{k+1}\bigcup\{t\}\right)=\operatorname{span}(T_{k+1}) \tag{11.49}$$

As

$$\begin{aligned}
T_{k+1}\bigcup\{t\} &\underset{[\text{eq: }11.43]}{=} (T_k\setminus\{t\})\bigcup\{e_{k+1}\}\bigcup\{t\}\\
&= \left((T_k\setminus\{t\})\bigcup\{t\}\right)\bigcup\{e_{k+1}\}\bigcup\{t\}\\
&= T_k\bigcup\{e_{k+1}\}\\
&\supseteq T_k
\end{aligned}$$

we have by [theorem: 11.104]

$$V\underset{[\text{eq: }11.36]}{=}\operatorname{span}(T_k)\leqslant\operatorname{span}\left(T_{k+1}\bigcup\{t\}\right)\underset{[\text{eq: }11.49]}{=}\operatorname{span}(T_{k+1})\subseteq V$$

proving that

$$\operatorname{span}\left((T\setminus T'_{k+1})\bigcup\{e_1,\ldots,e_{k+1}\}\right)\underset{[\text{eq: }11.42]}{=}\operatorname{span}(T_{k+1})=V$$

Using [eq: 11.40] we have $\operatorname{card}(T'_{k+1})=k+1$, $t\in T$ and $T'_{k+1}=T'_k\bigcup\{t\}\subseteq T$ [as $t\in T\wedge T'_k\subseteq T$. Hence by the definition of S proves that $k+1\in S$ finishing the induction step.

Mathematical induction proves then that $S=\mathbb{N}_0$. So as $n\in\{1,\ldots,n\}\subseteq\mathbb{N}_0=S$ we have that there exists a $T_n\subseteq T$ with $\operatorname{card}(T_n)=n$ such that $\operatorname{span}((T\setminus T_n)\bigcup\{e_1,\ldots,e_n\})=V$. □

Corollary 11.120. *Let $\langle V,+,\cdot\rangle$ be a vector space over a field $\langle F,+,\cdot\rangle$. Suppose that $E=\{e_1,\ldots,e_n\}\subseteq V$ is a finite linear independent set and there exist a finite $T\subseteq V$ such that $\operatorname{span}(T)=V$ then $n\leqslant\operatorname{card}(T)$.*

Proof. As T is finite we have by [theorem: 11.84] that $T=\{t_1,\ldots,t_m\}$ where $m=\operatorname{card}(T)$. Assume that $m<n$ then $\{e_1,\ldots,e_m\}\subseteq\{e_1,\ldots,e_n\}$ a linear independent set. Hence we have by [theorem: 11.118] that

$$\{e_1,\ldots,e_m\} \text{ is linear independent}$$

Using the Steinitz lemma [lemma: 11.119] there exists a $T_m \subseteq T$ with $\text{card}(T_m) = m$ such that

$$\text{span}\big((T \setminus T_m) \bigcup \{e_1, \ldots, e_m\}\big) = V \tag{11.50}$$

As $T_m \subseteq T$ and $\text{card}(T_m) = m = \text{card}(T)$ we have by [theorem: 10.93] that $T_m = T$ or $T \setminus T_m = \varnothing$ so that $(T \setminus T_m) \bigcup \{e_1, \ldots, e_m\} = \{e_1, \ldots, e_m\}$ hence by [eq: 11.50]

$$\text{span}(\{e_1, \ldots, e_m\}) = V$$

As $e_{m+1} \in V$ there exists by the above a finite

$$I \subseteq \{e_1, \ldots, e_m\} \underset{\text{[theorem: 11.86]}}{=} \{e_1, \ldots, e_{m+1}\} \setminus \{e_{m+1}\}$$

and a $\{\alpha_u\}_{u \in I}$ such that

$$e_{m+1} = \sum_{u \in I} \alpha_u \cdot u$$

Using [theorem:11.111] it follows then that $\{e_1,\ldots,e_{m+1}\}$ is linear dependent. By [theorem: 11.118] and the fact that $\{e_1,\ldots,e_{m+1}\} \subseteq \{e_1,\ldots, e_n\}$ [because $m+1 \leqslant n$] we reach the conclusion that $\{e_1, \ldots, e_n)$ is linear dependent which contradicts the hypothesis that $\{e_1, \ldots, e_n\}$ is linear independent. So the assumption $m < n$ leads to a contradiction, hence $n \leqslant m = \text{card}(T)$. □

Corollary 11.121. *Let $\langle V, +, \cdot \rangle$ be a vector space over a field $\langle F, +, \cdot \rangle$. Suppose that $E \subseteq V$ is a linear independent set and $T \subseteq F$ is a finite set such that* $\text{span}(T) = V$ *then E is finite.*

Proof. Assume that E is infinite then by [theorem: 6.29] there exists a denumerable set $B \subseteq E$. So there exists a bijection

$$\beta: \mathbb{N}_0 \rightarrow B \subseteq E$$

As $\{1, \ldots, \text{card}(T) + 1\} \subseteq N_0$ we have the bijection

$$e = \beta_{|\{1, \ldots \text{card}(T)+1\}}: \{1, \ldots, \text{card}(T) + 1\} \rightarrow \beta(\{1, \ldots, \text{card}(T) + 1\}) \subseteq B \subseteq E$$

defining

$$E' = \{e_1, \ldots, e_{\text{card}(T)+1}\} \subseteq E$$

As E is linear independent we have by [theorem: 11.118] that E' is linear independent. Using then the previous corollary [corollary: 11.120] we have that $\text{card}(T) + 1 \leqslant \text{card}(T)$ leading to the contradiction $\text{card}(T) \neq \text{card}(T)$. Hence E must be finite. □

11.3.2.4 Linear (in)dependent families

Up to now we have talked about linear (in)dependency of sets. However if want to work later with skew-symmetric mappings (needed for the theory of determinant functions) we need the concept of linear (in)dependency of ordered families.

Definition 11.122. *Let $\langle V, +, \cdot \rangle$ be a vector space over a field $\langle F, +, \cdot \rangle$, $n \in \mathbb{N}$ then a family $\{x_i\}_{i \in \{1, \ldots, n\}} \subseteq V$ is **linear dependent** if $\exists \{x_i\}_{i \in \{1, \ldots, n\}} \subseteq F$ satisfying $\exists i \in \{1, \ldots, n\}$ with $\alpha_i \neq 0$ such that*

$$\sum_{i \in \{1, \ldots, n\}} \alpha_i \cdot x_i$$

Just as for linear dependent sets there exist a equivalent definition for linear dependent families.

Theorem 11.123. *Let $\langle V, +, \cdot \rangle$ be a vector space over a field $\langle F, +, \cdot \rangle$, $n \in \mathbb{N}$ and $\{x_i\}_{i \in \{1, \ldots, n\}} \subseteq V$ then*

$$\{x_i\}_{i \in \{1, \ldots, n\}} \textit{ is linear dependent}$$
$$\Updownarrow$$
$$\exists k \in \{1, \ldots, n\} \textit{ and } \exists \{\alpha_i\}_{i \in \{1, \ldots, n\} \setminus \{k\}} \subseteq F \textit{ such that } x_k = \sum_{i \in \{1, \ldots, n\} \setminus \{k\}} \alpha_i \cdot x_i$$

Proof.

- **⇒.** As $\{x_i\}_{i \in \{1, \ldots, n\}}$ is linear dependent there exist a $\{\beta_i\}_{i \in \{1, \ldots, n\}} \subseteq F$ satisfying $\exists k \in \{1, \ldots, n\}$ such that $\beta_k \neq 0$ such that $\sum_{i \in \{1, \ldots, n\}} \beta_i \cdot x_i = 0$. Hence

$$\begin{aligned} 0 &= \sum_{i \in \{1, \ldots, n\}} \beta_i \cdot x_i \\ &\underset{\text{[theorem: 11.43]}}{=} \sum_{i \in \{1, \ldots, n\} \setminus \{k\}} \beta_i \cdot x_i + \sum_{i \in \{k\}} \beta_i \cdot x_i \\ &\underset{\text{[theorem: 11.34]}}{=} \sum_{i \in \{1, \ldots, n\} \setminus \{k\}} \beta_i \cdot x_i + \beta_k \cdot x_k \end{aligned}$$

so that

$$\begin{aligned}x_k &= -(\beta_k)^{-1}\cdot\sum_{i\in\{1,\ldots,n\}\setminus\{k\}}\beta_i\cdot x_i\\ &= \sum_{i\in\{1,\ldots,n\}\setminus\{k\}}((\beta_k)^{-1}\cdot(-\beta_i))\cdot x_i\\ &= \sum_{i\in\{1,\ldots,n\}\setminus\{k\}}\alpha_i\cdot x_i\end{aligned}$$

where

$$\{\alpha_i\}_{i\in\{1,\ldots,n\}}\subseteq F \text{ is defined by } \alpha_i=((\beta_k)^{-1}\cdot(-\beta_i))$$

⇐. Assume that $x_k=\sum_{i\in\{1,\ldots,n\}\setminus\{k\}}\beta_i\cdot x_i$ where $\beta\alpha_i\}_{i\in\{1,\ldots,n\}\setminus\{k\}}\subseteq F$. Define

$$\{\alpha_i\}_{i\in\{1,\ldots,n\}}\subseteq F \text{ by } \alpha_i=\begin{cases}-1 \text{ if } i=k\\ \beta_i \text{ if } i\in\{1,\ldots,n\}\setminus\{k\}\end{cases}$$

then we have

$$\begin{aligned}\sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot x_i &\underset{[\text{theorem: } 11.43]}{=} \sum_{i\in\{1,\ldots,n\}\setminus\{k\}}\alpha_i\cdot x_i+\sum_{i\in\{k\}}\alpha_i\cdot x_i\\ &\underset{[\text{theorem: } 11.34]}{=} \sum_{i\in\{1,\ldots,n\}\setminus\{k\}}\alpha_i\cdot x_i+\alpha_k\cdot x_k\\ &= \sum_{i\in\{1,\ldots,n\}\setminus\{k\}}\beta_i\cdot x_i+(-1)\cdot x_k\\ &= x_k=x_k\\ &= 0\end{aligned}$$

proving, as $\alpha_k=-1\neq 0$, linear dependency. □

Theorem 11.124. *Let $\langle V,+,\cdot\rangle$ be a vector space over a field $\langle F,+,\cdot\rangle$, $n\in\mathbb{N}$ and $\{x_k\}_{k\in\{1,\ldots,n\}}\subseteq X$ a ordered family then:*

1. *If there exists $i,j\in\{1,\ldots,n\}$ with $i\neq j$ and $x_i=x_j$ then $\{x_k\}_{k\in\{1,\ldots,n\}}$ is linear dependent.*
2. *If there exists a $i\in\{1,\ldots,n\}$ such that $x_i=0$ then $\{x_k\}_{k\in\{1,\ldots,n\}}$ is linear dependent.*

Proof.

1. Define

$$\{\alpha_k\}_{k\in\{1,\ldots,n\}}\subseteq F \text{ by } \alpha_k=\begin{cases}1 \text{ if } k=i\\ -1 \text{ if } k=j\\ 0 \text{ if } k\in\{1,\ldots,n\}\setminus\{i,j\}\end{cases}$$

then we have

$$\begin{aligned}\sum_{k\in\{1,\ldots,n\}}\alpha_k\cdot x_k &\underset{[\text{theorem: } 11.43]}{=} \sum_{k\in\{1,\ldots,n\}\setminus\{i,j\}}\alpha_k\cdot x_k+\sum_{k\in\{i\}}\alpha_k\cdot x_k+\sum_{k\in\{j\}}\alpha_k\cdot x_k\\ &\underset{[\text{theorem: } 11.34]}{=} \sum_{k\in\{1,\ldots,n\}\setminus\{i,j\}}\alpha_k\cdot x_k+\alpha_i\cdot x_i+\alpha_j\cdot x_j\\ &= \sum_{k\in\{1,\ldots,n\}\setminus\{i,j\}}0\cdot x_k+1\cdot x_i+(\alpha_j-1)\cdot x_j\\ &= x_i-x_j\\ &\underset{x_i=x_j}{=} 0\end{aligned}$$

which as $\alpha_i=1\neq 0$ proves linear dependency.

2. Define

$$\{a_k\}_{k\in\{1,\ldots,n\}}\subseteq F \text{ by } \alpha_k=\begin{cases}1 \text{ if } k=i\\ 0 \text{ if } k\in\{1,\ldots,n\}\setminus\{i\}\end{cases}$$

then we have

$$\begin{aligned}\sum_{k\in\{1,\ldots,n\}}\alpha_k\cdot x_k &\underset{[\text{theorem: } 11.43]}{=} \sum_{k\in\{1,\ldots,n\}\setminus\{i\}}\alpha_k\cdot x_k+\sum_{k\in\{i\}}\alpha_k\cdot x_k\\ &\underset{[\text{theorem: } 11.34]}{=} \sum_{k\in\{1,\ldots,n\}\setminus\{i\}}\alpha_k\cdot x_k+\alpha_i\cdot x_i\\ &= \sum_{k\in\{1,\ldots,n\}\setminus\{i\}}0\cdot x_k+1\cdot 0\\ &\underset{[\text{theorem: } 11.37]}{=} 0\end{aligned}$$

which as $\alpha_i=1\neq 0$ proves linear dependency. □

The following theorem shows the relation between linear dependency of a family $\{x_i\}_{i\in\{1,\ldots,n\}}\subseteq V$ and the linear dependency of its associated set $\{x_i|i\in\{1,\ldots,n\}\}$.

Theorem 11.125. *Let $\langle V,+,\cdot\rangle$ be a vector space over a field $\langle F,+,\cdot\rangle$, $n\in\mathbb{N}$ and $\{x_i\}_{i\in\{1,\ldots,n\}}\subseteq X$ a ordered family such that $\{x_i|i\in\{1,\ldots,n\}\}$ is linear dependent then $\{x_i\}_{i\in\{1,\ldots,n\}}$ is linear dependent.*

Proof. Let $W=\{x_i|i\in\{1,\ldots,n\}\}$ then by [theorem: 10.94] W is finite and

$$\operatorname{card}(W)\leqslant\operatorname{card}(\{1,\ldots,n\})=n.$$

As $\{x_i|i\in\{1,\ldots,n\}\}$ is linear dependent there exist a $I\subseteq W$ and $\{\alpha_i\}_{i\in I}\subseteq F$, satisfying $\exists l\in I$ with $\alpha_l\neq 0$, such that

$$\sum_{i\in I}\alpha_i\cdot i=0$$

By definition of a family, $\{x_i\}_{i\in\{1,\ldots,n\}}$ is defined by a function $x:\{1,\ldots,n\}\rightarrow x(\{1,\ldots,n\})=W$. Using a consequence of the Axiom of Choice [see theorem: 3.129] there exists a $J\subseteq\{1,\ldots,n\}$ so that

$$x_{|J}:J\rightarrow W \text{ is a bijection}$$

Take $K=(x_{|J})^{-1}(I)\subseteq J\subseteq\{1,\ldots,n\}$ then we have the bijection

$$x_{|K}\underset{[\text{theorem: } 2.84]}{=}(x_{|j})_{|K}:K\rightarrow x_{|K}(K)=x(K)=I$$

Define now

$$\{\beta_k\}_{k\in\{1,\ldots,n\}} \text{ by } \beta_k=\begin{cases}\alpha_{(x_{|K})(i)} \text{ if } k\in K\\ 0 \text{ if } k\in\{1,\ldots,n\}\setminus K\end{cases}$$

then

$$\begin{aligned}\sum_{i\in\{1,\ldots,n\}}\beta_i\cdot x_i &\underset{[\text{theorem: } 11.43]}{=} \sum_{i\in\{1,\ldots,n\}\setminus K}\beta_i\cdot x_i+\sum_{i\in K}\beta_i\cdot x_i\\ &= \sum_{i\in\{1,\ldots,n\}\setminus K}0\cdot x_i+\sum_{i\in K}\alpha_{|x_{|K}(i)}\cdot x_{|K}(i)\\ &= \sum_{i\in K}\alpha_{|x_{|K}(i)}\cdot x_{|K}(i)\\ &\underset{[\text{theorem: } 11.36]}{=} \sum_{i\in I}\alpha_i\cdot i\\ &= 0\end{aligned}$$

which as $\beta_{(x_{|K})^{-1}(l)}=\alpha_{x_{|K}((x_{|K})^{-1}(l))}=\alpha_l\neq 0$ proves that

$$\{x_i\}_{i\in\{1,\ldots,n\}} \text{ is linear dependent}$$ □

The opposite of the above theorem is not true as the following example shows.

Example 11.126. Let $\langle V,+,\cdot\rangle$ be a vector space over a field $\langle F,+,\cdot\rangle$ and $v\in V$ with $v\neq 0$ then $\{x_i\}_{i\in\{1.2\}}\subseteq X$ defined by $x_i=v$ is linear dependent, however $\{x_i|i\in\{1,.2\}\}$ linear independent.

Proof. Define $\{\alpha_i\}_{i\in\{1,2\}}$ by $\alpha_1=1$ and $\alpha_2=-1$ then $\alpha_1\neq 0$ and $\sum_{i\in\{1,.2\}}\alpha_i\cdot x_i=(-1)\cdot x+1\cdot x=0$ proving that linear dependency. Further $\{x_i|i\in\{1,2\}\}=\{v\}$ is by [example: 11.117] linear independent. □

Definition 11.127. *Let $\langle V,+,\cdot\rangle$ be a vector space over a field $\langle F,+,\cdot\rangle$, $n\in\mathbb{N}$ then $\{x_i\}_{i\in\{1,\ldots,n\}}\subseteq V$ is linear independent if $\{x_i\}_{i\in\{1,\ldots,n\}}\subseteq V$ is **not** linear dependent.*

Just as with linear independent set we have a equivalent definition of linear dependency.

Theorem 11.128. *Let $\langle V,+,\cdot\rangle$ be a vector space over the field $\langle F,+,\cdot\rangle$,$n\in\mathbb{N}$ and $\{x_i\}_{i\in\{1,\ldots,n\}}\subseteq V$ then*

$$\{x_i\}_{i\in\{1,\ldots,n\}}\subseteq V \text{ is linear independent}$$
$$\Updownarrow$$
$$\forall\{\alpha_i\}_{i\in\{1,\ldots,n\}}\subseteq F \text{ such that } \sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot x_i=0 \text{ we have } \forall i\in\{1,\ldots,n\}\ \alpha_i=0$$

Proof.

⇒. As $\{\alpha_i\}_{i\in\{1,\ldots,n\}}\subseteq F$ is linear independent we have by definition that $\{\alpha_i\}_{i\in\{1,\ldots,n\}}\subseteq F$ is not dependent. Hence for every $\{\alpha_i\}_{i\in\{1,\ldots,n\}}\subseteq F$ satisfying $\exists i\in I$ with $\alpha_i\neq 0$ we must have $\sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot x_i\neq 0$. So if $\sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot x_i=0$ we must have $\forall i\in\{1,\ldots,n\}$ that $\alpha_i=0$ otherwise we reach the contradiction that $\sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot x_i\neq 0$.

⇐. Assume that $\{\alpha_i\}_{i\in\{1,\ldots,n\}}\subseteq F$ is linear dependent then there exists a $\{\alpha_i\}_{i\in\{1,\ldots,n\}}\subseteq V$ satisfying $\exists k\in\{1,\ldots,n\}$ with $\alpha_k\neq 0$ such that $\sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot x_i=0$. From the hypothesis it follows then that $\forall i\in\{1,\ldots,n\}$ $\alpha_i=0$ contradicting $\alpha_k\neq 0$. So we must have that $\{x_i\}_{i\in\{1,\ldots,n\}}\subseteq F$ is linear independent. □

The following theorem shows the relation between linear independency of a family $\{x_i\}_{i\in\{1,\ldots,n\}}\subseteq V$ and the linear independency of its associated set $\{x_i|i\in\{1,\ldots,n\}\}$.

Theorem 11.129. *Let $\langle V,+,\cdot\rangle$ be a vector space over a field $\langle F,+,\cdot\rangle$, $n\in\mathbb{N}$ and $\{x_i\}_{i\in\{1,\ldots,n\}}\subseteq V$ a linear independent family then*

1. *$\{x_i|i\in\{1,\ldots,n\}\}$ is a linear independent set*
2. *$\{x_i|i\in\{1,\ldots,n\}\}=\{v_1,\ldots,v_n\}$ [or by definition: 11.79] that $x:\{1,\ldots,n\}\rightarrow\{x_i|i\in\{1,\ldots,n\}\}$ is a bijection].*

Proof.

1. We prove this by contradiction, so assume that $\{x_i|i\in\{1,...,n\}\}$ is linear dependent then by [theorem: 11.125] $\{x_i\}_{i\in\{1,\ldots,n\}}$ is linear dependent, contradicting the linear independency of $\{x_i\}_{i\in\{1,\ldots,n\}}$. Hence we must have that $\{x_i|i\in\{1,...,n\}\}$ is linear independent.
2. Let $x:\{1,\ldots,n\}\rightarrow X$ be the function that defines $\{x_i\}_{i\in\{1,\ldots,n\}}\subseteq X$ [where $x_i\underset{\text{notation}}{=}x(i)$]. Let $i,j\in\{1,\ldots,n\}$ such that $x(i)=x(j)$ and assume that $i\neq j$. Define

$$\{\alpha_k\}_{k\in\{1,\ldots,n\}}\subseteq F \text{ by } \alpha_k=\begin{cases}1 \text{ if } k=i\\-1 \text{ if } k=j\\0 \text{ if } k\in\{1,\ldots,n\}\setminus\{i,j\}\end{cases}$$

then we have

$$\begin{aligned}\sum_{k\in\{1,\ldots,n\}}\alpha_k\cdot x_k &\underset{\text{[theorem: 11.43]}}{=} \sum_{k\in\{1,\ldots,n\}\setminus\{i,j\}}\alpha_k\cdot x_k+\sum_{k\in\{i\}}\alpha_k\cdot x_k+\sum_{k\in\{j\}}\alpha_k\cdot x_k\\ &\underset{\text{[theorem: 11.34]}}{=} \sum_{k\in\{1,\ldots,n\}\setminus\{i,j\}}\alpha_k\cdot x_k+\alpha_i\cdot x_i+\alpha_j\cdot x_j\\ &= \sum_{k\in\{1,\ldots,n\}\setminus\{i,j\}}0\cdot x_k+1\cdot x_i+(-1)\cdot x_j\\ &\underset{\text{[theorem: 11.37]}}{=} x_i-x_j\\ &= x(i)-x(j)\\ &= 0\end{aligned}$$

which, as $\{x_i\}_{i\in\{1,\ldots,n\}}\subseteq V$ is linear independent, results in the contradiction $0=\alpha_1=-1$. Hence $\forall i,j\in\{1,\ldots,n\}$ with $x(i)=x(j)$ we must have $i=j$ proving that x is injective. So

$$x:\{1,\ldots,n\}\rightarrow x(\{1,\ldots,n\})=\{x_i|i\in\{1,\ldots,n\}\}\text{ is a bijection}$$

or by [definition: 11.79] that $\{x_i|i\in\{1,\ldots,n\}\}=\{x_1,\ldots,x_n\}$. □

11.3.3 Basis of a vector space

Definition 11.130. *Let $\langle V,+,\cdot\rangle$ be a vector space over a field $\langle F,+,\cdot\rangle$ then $B\subseteq V$ is a basis of $\langle V,+,\cdot\rangle$ if*

1. *B is linear independent*
2. $\operatorname{span}(B)=V$

Note 11.131. The basis that we define here is called a Hamel basis in books about vector spaces. Every vector in a vector space can be written as a linear combination of the Hamel basis. So we don't have to bother with infinite sums and convergence of these sums, although the basis itself does not have to be finite. There are other concepts of basis in a vector space that work with infinite sums (for example Fourier series) but these are not covered in this section.

Example 11.132. Let $\langle\{0\},+,\cdot\rangle$ be the trivial space over a field $\langle F,+,\cdot\rangle$ then $\varnothing$ is the only basis for $\langle\{0\},+,\cdot\rangle$

Proof. By [example: 11.99] we have $\operatorname{span}(\varnothing)=\{0\}$ and by [example: 11.116] $\varnothing$ is linear independent, proving that

$$\varnothing\text{ is a basis for }\langle\{0\},+,\cdot\rangle$$

If B is another basis of $\langle\{0\},+,\cdot\rangle$ then $B\subseteq\{0\}$. If $B=\{0\}$ then $0\in B$ so that by [theorem: 11.110] B is linear dependent contradicting linear Independence, hence we must have that $B=\varnothing$. □

A basis of a vector space allows us to write every vector of the vector space as a finite linear combination of vectors in the basis.

Theorem 11.133. *Let $\langle V,+,\cdot\rangle$ be a vector space over a field $\langle F,+,\cdot\rangle$ and $B\subseteq V$ then*

$$B\textit{ is a basis of }\langle V,+,\cdot\rangle$$

$$\Updownarrow$$

$$\forall v\in V\textit{ their exists a finite }I\subseteq B\textit{ and a }\textbf{unique (1)}\ \{\alpha_u\}_{u\in I}\subseteq F\textit{ such that }v=\sum_{u\in I}\alpha_u\cdot u$$

(1) If $\{\beta_u\}_{u\in I}\subseteq F$ satisfies $v=\sum_{u\in I}\beta$ then $\{\alpha_u\}_{u\in I}=\{\beta_u\}_{u\in I}$

Proof.

⇒. As $\operatorname{span}(B)=V$ and $v\in V$ there exists a finite $I\subseteq B$ and a $\{\alpha_u\}_{u\in I}\subseteq F$ such that

$$v=\sum_{u\in I}\alpha_u\cdot u$$

proving existence. For uniqueness let $\{\beta_u\}_{u\in I}\subseteq F$ such that

$$v=\sum_{u\in I}\beta_u\cdot u$$

Define

$$\{\gamma_u\}_{u\in I}\subseteq F\text{ by }\gamma_u=\alpha_u-\beta_u$$

then we have

$$\begin{aligned}\sum_{u\in I}\gamma_u\cdot u &= \sum_{u\in I}(\alpha_u-\beta_u)\cdot u\\ &\underset{[\text{theorems: 11.6, 11.40}]}{=} \sum_{u\in I}\alpha_u\cdot u-\sum_{u\in I}\beta_u\cdot u\\ &= v-v\\ &= 0\end{aligned}$$

As B is linear independent we have $\forall u \in I$ that $\alpha_u - \beta_u = \gamma_u = 0$ proving that

$$\{\alpha_u\}_{u \in I} \subseteq F = \{\beta_u\}_{u \in I} \subseteq F$$

$\Leftarrow$. Let $v \in V$ then by the hypothesis there exists a finite $I \subseteq B$ and a $\{\alpha_u\}_{u \in I} \subseteq F$ such that $v = \sum_{u \in I} \alpha_u \cdot u$ proving that $v \in \operatorname{span}(B)$. So $V \subseteq \operatorname{span}(B) \subseteq V$ hence

$$\operatorname{span}(B) = V$$

As for linear independence let $I \subseteq B$ a finite set and $\{\alpha_u\}_{u \in I} \subseteq F$ such that $\sum_{u \in I} \alpha_u \cdot u = 0$. Define $\{\beta_u\}_{u \in I} \subseteq F$ by $\beta_u = 0$ then $\sum_{u \in I} \beta_u \cdot u = \sum_{u \in I} 0 \cdot u \underset{[\text{theorem: } 11.37]}{=} 0$, by uniqueness we have $\forall u \in I$ $\alpha_u = \beta_u = 0$. So by definition

$$B \text{ is linear independent}$$

□

The uniqueness in the above is only true if the index sets are equal, however we can ensure uniqueness also for the index set if we require that the coefficients $\{\alpha_i\}_{i \in I}$ are non zero.

Theorem 11.134. *Let $\langle V, +, \cdot \rangle$ be a vector space over a field $\langle F, +, \cdot \rangle$ and B a basis of $\langle V, +, \cdot \rangle$ then $\forall v \in V$ there exist a finite $I \subseteq B$ and $\{\alpha_u\}_{u \in I} \subseteq F$ such that $v = \sum_{u \in I} \alpha_u \cdot u$. Further if there exists another $J \subseteq B$ and a $\{\beta_u\}_{u \in J} \subseteq F$ such that $v = \sum_{u \in J} \beta_u \cdot u$ then*

$$\begin{array}{ccc} \forall u \in I \setminus J & \text{we have} & \alpha_u = 0 \\ \forall u \in J \setminus I & \text{we have} & \beta_u = 0 \\ \multicolumn{3}{c}{\forall u \in I \bigcap J \text{ we ahve } \alpha_u = \beta_u} \end{array}$$

Proof. Let $v \in V$ then using the previous theorem [theorem: 11.133] there exist a finite $I \subseteq B$ and a $\{\alpha_i\}_{i \in I} \subseteq F$ such that $v = \sum_{i \in I} \alpha_u \cdot u$. Assume that there exists another finite $J \subseteq B$ and a $\{\beta_i\}_{i \in J} \subseteq F \setminus \{0\}$ such that $v = \sum_{i \in J} \beta_i \cdot i$. Define

$$\{\gamma_i\}_{i \in I \bigcup J} \subseteq F \text{ by } \gamma_i = \begin{cases} \alpha_i \text{ if } i \in I \\ 0 \text{ if } i \in (I \bigcup J) \setminus I \end{cases} \text{ and } \{\delta_i\}_{i \in I \bigcup J} \text{ by } \delta_i = \begin{cases} \beta_i \text{ if } i \in J \\ 0 \text{ if } i \in (I \bigcup J) \setminus J \end{cases}$$

then we have

$$\begin{aligned} \sum_{i \in I \bigcup J} \alpha_i \cdot i &\underset{[\text{theorem: } 11.43]}{=} \sum_{i \in (I \bigcup J) \setminus I} \alpha_i \cdot i + \sum_{i \in I} \alpha_i \cdot i \\ &= \sum_{i \in (I \bigcup J) \setminus I} 0 \cdot i + v \\ &\underset{[\text{theorem: } 11.37]}{=} 0 + v \\ &= v \\ &= 0 + v \\ &\underset{[\text{theorem: } 11.37]}{=} \sum_{i \in (I \bigcup J) \setminus J} 0 \cdot i + v \\ &= \sum_{i \in (I \bigcup J) \setminus J} \beta_i \cdot i + \sum_{i \in J} \beta_i \cdot i \\ &\underset{[\text{theorem: } 11.43]}{=} \sum_{i \in I \bigcup J} \beta_i \cdot i \end{aligned}$$

Applying then the previous theorem [theorem: 11.133] on the above we have

$$\forall i \in I \bigcup J \text{ that } \gamma_i = \delta_i$$

If $i \in I \setminus J$ then $i \in (I \bigcup J) \setminus J$ hence $0 = \delta_i = \gamma_i = \alpha_i$ so that

$$\forall i \in I \setminus J \text{ we have } \alpha_i = 0$$

If $i \in J \setminus I$ then $i \in (I \bigcup J) \setminus I$ hence $0 = \gamma_i = \delta_i = \beta_i$ so that

$$\forall i \in J \setminus I \text{ we have } \beta_i = 0$$

Finally if $i \in I \bigcap J$ then $\alpha_i = \gamma_i = \delta_i = \beta_i$ hence

$$\forall i \in I \bigcap J \text{ we have } \alpha_i = \beta_i$$

□

Theorem 11.135. *Let $\langle V,+,\cdot\rangle$ be a vector space over a field $\langle F,+,\cdot\rangle$ and B a basis of $\langle V,+,\cdot\rangle$ then $\forall v\in V$ there exists a **unique pair** $(I,\{\alpha_u\}_{u\in I}\subseteq F\setminus\{0\})$ such that I is finite, $I\subseteq B$ and*

$$v=\sum_{u\in I}\alpha_u\cdot u$$

Proof. Let $v\in V$. By [theorem: 11.133] there exist a finite $K\subseteq B$ and a $\{\alpha_i\}_{i\in K}\subseteq F$ such that $v=\sum_{i\in K}\alpha_i\cdot u$. Define now $I\subseteq K$ by $I=\{i\in K|\alpha_i\neq 0\}$ so that $\{\alpha_i\}_{i\in I}\subseteq F\setminus\{0\}$ then we have

$$\begin{aligned}
v &= \sum_{i\in K}\alpha_i\cdot i\\
&\underset{[\text{theorem: } 11.43]}{=} \sum_{i\in K\setminus I}\alpha_i\cdot i+\sum_{i\in I}\alpha_i\cdot i\\
&= \sum_{i\in K\setminus I}0\cdot i+\sum_{i\in K}\alpha_i\cdot i\\
&\underset{[\text{theorem: } 11.37]}{=} 0+\sum_{i\in I}\alpha_i\cdot i\\
&= \sum_{i\in I}\alpha_i\cdot i
\end{aligned}$$

which proves existence. As for uniqueness, assume that there exists a finite $J\subseteq B$ and a $\{\beta_i\}_{i\in J}\subseteq F\setminus\{0\}$ such that $v=\sum_{i\in J}\beta_i\cdot i$. Then by the previous theprem [theorem: 11.134] we have that

$$\forall i\in I\setminus J\ \alpha_i=0\wedge\forall i\in J\setminus I\ \beta_i=0\wedge\forall i\in I\bigcap J\ \alpha_i=\beta_i$$

As by the hypothesis $\forall i\in I\ \alpha_i\neq 0$ and $\forall i\in J\ \beta_i\neq 0$ we must have that $I\setminus J=\varnothing=J\setminus I$ hence

$$I=J$$

and $\forall i\in I=J=I\bigcap J$ that $\alpha_i=\beta_i$ or

$$\{\alpha_i\}_{i\in I}=\{\beta_i\}_{i\in J}$$

□

If B is a finite we have a simpler equivalence.

Theorem 11.136. *Let $\langle V,+,\cdot\rangle$ be a vector space over a field $\langle F,+,\cdot\rangle$ and $B\subseteq V$ where B is **finite** then*

$$B \textit{ is a basis of } \langle V,+,\cdot\rangle$$

$$\Updownarrow$$

$$\forall v\in V \textit{ their exists a } \textbf{\textit{unique}}\ \{\alpha_u\}_{u\in B}\subseteq F \textit{ such that } v=\sum_{u\in B}\alpha_u\cdot e_u$$

Proof.

⇒. If B is a basis we have by [theorem: 11.133] that there exists a finite $I\subseteq B$ and a $\{\beta_u\}_{u\in I}\subseteq F$ such that $v=\sum_{u\in I}\beta_u\cdot u$. Define

$$\{\alpha_u\}_{u\in B}\subseteq F \text{ by } \alpha_u=\begin{cases}0 \text{ if } u\in B\setminus I\\ \beta_u \text{ if } u\in I\end{cases}$$

then we have

$$\begin{aligned}
\sum_{u\in B}\alpha_u\cdot u &\underset{[\text{theorem: } 11.43]}{=} \sum_{u\in B\setminus I}\alpha_u\cdot u+\sum_{u\in I}\alpha_u\cdot u\\
&= \sum_{u\in B\setminus I}0\cdot u+\sum_{u\in I}\beta_u\cdot u\\
&\underset{[\text{theorem: } 11.37]}{=} \sum_{u\in I}\beta_u\cdot u\\
&= v
\end{aligned}$$

proving existence. For uniqueness let $\{\lambda_u\}_{u\in B}\subseteq F$ such that

$$v=\sum_{u\in B}\lambda_u\cdot u$$

Define

$$\{\gamma_u\}_{u\in B} \subseteq F \text{ by } \gamma_u = \alpha_u - \lambda_u$$

then we have

$$\begin{aligned}\sum_{u\in B} \gamma_u \cdot u &= \sum_{u\in B} (\alpha_u - \lambda_u) \cdot u \\ &\underset{[\text{theorems: } 11.6,\ 11.40]}{=} \sum_{u\in B} \alpha_u \cdot u - \sum_{u\in B} \lambda_u \cdot u \\ &= v - v \\ &= 0\end{aligned}$$

As B is linear independent we have $\forall u \in B$ that $\alpha_u - \lambda_u = \gamma_u = 0$ proving that

$$\{\alpha_u\}_{u\in B} \subseteq F = \{\lambda_u\}_{u\in B} \subseteq F$$

$\Leftarrow$. If $v \in V$ then by the hypothesis there exist a $\{\alpha_u\}_{u\in B} \subseteq F$ such that $v = \sum_{u\in B} \alpha_u \cdot u$ proving that $v \in \operatorname{span}(B)$. So $V \subseteq \operatorname{span}(B) \subseteq V$ hence

$$\operatorname{span}(B) = V$$

As for linear independence let $I \subseteq B$ a finite set and $\{\alpha_u\}_{u\in I} \subseteq F$ such that $\sum_{u\in I} \alpha_u \cdot u = 0$. Define $\{\beta_u\}_{u\in I} \subseteq F$ by $\beta_u = 0$ then $\sum_{u\in I} \beta_u \cdot u = \sum_{u\in I} 0 \cdot u \underset{[\text{theorem: } 11.37]}{=} 0$. By uniqueness we have $\forall u \in I$ $\alpha_u = \beta_u = 0$. Using [theorem: 11.115] it follows then that

$$B \text{ is linear independent}$$

□

A other usefully alternative characterization of a finite basis is the following

Theorem 11.137. *Let $\langle V, +, \cdot\rangle$ be a vector space over a field $\langle F, +, \cdot\rangle$ and $E = \{e_1, \ldots, e_n\} \subseteq V$ a finite set then*

$$B \textit{ is a basis of } \langle V, +, \cdot\rangle$$

$$\Updownarrow$$

$$\forall v \in V \textit{ their exists a } \boldsymbol{unique}\ \{\alpha_i\}_{i\in\{1,\ldots,n\}} \subseteq F \textit{ such that } v = \sum_{i\in\{1,\ldots,n\}} \alpha_i \cdot e_i$$

Proof. $E = \{e_1, \ldots, e_n\}$ is defined by the bijection

$$e: \{1, \ldots, n\} \rightarrow E \text{ where } n = \operatorname{card}(E)$$

Further we have

$\Rightarrow$. Let $v \in V$ then by [theorem: 11.136] there exists a $\{\alpha_u\}_{u\in B} \subseteq F$ such that

$$v = \sum_{u\in W} \alpha_u \cdot u$$

Define

$$\{\beta_i\}_{i\in\{1,\ldots,n\}} \subseteq F \text{ by } \beta_i = \alpha_{e(i)}$$

then we have

$$\begin{aligned}\sum_{i\in\{1,\ldots,n\}} \beta_i \cdot e_i &= \sum_{i\in\{1,\ldots,n\}} \alpha_{e(i)} \cdot e(i) \\ &\underset{[\text{theorem: } 11.36]}{=} \sum_{u\in B} \alpha_u \cdot u \\ &= v\end{aligned}$$

proving existence. As for uniqueness let $\{\lambda_i\}_{i\in\{1,\ldots,n\}} \subseteq F$ be another family such that $v = \sum_{i\in\{1,\ldots,n\}} \lambda_i \cdot e_i$ and define

$$\{\gamma_i\}_{i\in\{1,\ldots,n\}} \subseteq F \text{ by } \gamma_i = \beta_i - \lambda_n$$

then we have

$$\begin{aligned}\sum_{i\in\{1,\ldots,n\}}\gamma_i\cdot e_i &= \sum_{i\in\{1,\ldots,n\}}(\beta_i-\lambda_i)\cdot e_i\\ &= \sum_{i\in\{1,\ldots,n\}}(\beta_i\cdot e_i-\lambda_i\cdot e_i)\\ &\underset{[\text{theorems: } 11.6,\ 11.40]}{=} \sum_{i\in\{1,\ldots,n\}}\beta_i\cdot e_i-\sum_{i\in\{1,\ldots,n\}}\lambda_i\cdot e_i\\ &= v-v\\ &= 0\end{aligned}$$

As E is linear independent it follows then from the above and [theorem: 11.115] that $\forall i\in\{1,\ldots,n\}$ $\gamma_i=0\Rightarrow\beta_i=\lambda_i$. Hence $\{\beta_i\}_{i\in\{1,\ldots,n\}}=\{\lambda_i\}_{i\in\{1,\ldots,n\}}$.

⇐. Let $v\in V$ then by the hypothesis there exists a $\{\alpha_i\}_{i\in\{1,\ldots,n\}}\subseteq F$ such that

$$v=\sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot e_i$$

Define then

$$\{\beta_u\}_{u\in B}\subseteq F \text{ by } \beta_u=\alpha_{e^{-1}(u)}$$

then we have

$$\begin{aligned}\sum_{u\in B}\beta_u\cdot u &= \sum_{u\in B}\alpha_{e^{-1}(u)}\cdot e(e^{-1}(u))\\ &\underset{[\text{theorem: } 11.36]}{=} \sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot e(i)\\ &= \sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot e_i\\ &= v\end{aligned}$$

proving that $v\in\mathrm{span}(B)$. Hence $V\subseteq\mathrm{span}(B)\subseteq V$ so that

$$V=\mathrm{span}(V)$$

Let $\{\gamma_i\}_{i\in\{1,\ldots,n\}}\subseteq F$ be such that $0=\sum_{i\in\{1,\ldots,n\}}\gamma_i\cdot e_i$ and define $\{\zeta_i\}_{i\in\{1,\ldots,n\}}$ by $\gamma_i=0$ then we have also $\sum_{i\in\{1,\ldots,n\}}\zeta_i\cdot e_i=0$. Hence by uniqueness we have $\forall i\in\{1,\ldots,n\}$ that $\gamma_i=\zeta_i=0$. Using then [theorem: 11.115] it follows that E is linear independent. □

Scaling a finite basis of a vector space results in a finite basis of the same vector space as is show in the next theorem.

Theorem 11.138. *Let $\langle V,+,\cdot\rangle$ be a vector space over a field $\langle F,+,\cdot\rangle$, $\{\alpha_i\}_{i\in\{1,\ldots,n\}}\subseteq F\setminus\{0\}$ and a finite basis $X=\{x_1,\ldots,x_n\}\subseteq V$ of V then*

$$y:\{1,\ldots,n\}\rightarrow\{\alpha_i\cdot x_i|i\in\{1,\ldots,n\}\} \text{ defined by } y(i)=\alpha_i\cdot x_i \text{ is a bijection}$$

defining the finite set

$$Y=\{y_1,\ldots,y_n\}=\{\alpha_1\cdot x_1,\ldots,\alpha_n\cdot x_n\}$$

that is a basis of V. In other words if $\{x_1,\ldots,x_n\}$ is a basis for V then if $\{\alpha_i\}_{i\in\{1,\ldots,n\}}\subseteq F\setminus\{0\}$ we have that $\{\alpha_1\cdot x_1,\ldots,\alpha_n\cdot x_n\}$ is a basis of V.

Proof. Define

$$y:\{1,\ldots,n\}\rightarrow V \text{ by } y(i)=\alpha_i\cdot x_i$$

Assume that $\exists i,j\in\{1,\ldots,n\}$ with $y(i)=y(j)\Rightarrow\alpha_i\cdot x_i=\alpha_j\cdot x_j$ such that $i\neq j$. Then as $\{\alpha_k\}_{k\in\{1,\ldots,n\}}\subseteq F\setminus\{0\}$ we have $x_j=(\alpha_i/\alpha_j)\cdot x_i$. Define then

$$\{\beta_k\}_{k\in\{1,\ldots,n\}} \text{ by } \beta_k=\begin{cases}\alpha_i/\alpha_j \text{ if } k=i\\ -1 \text{ if } k=j\\ 0 \text{ if } k\in\{1,\ldots,n\}\setminus\{i,j\}\end{cases}$$

then we have

$$\begin{aligned}\sum_{k\in\{1,\ldots,n\}}\alpha_k\cdot x_k &\underset{[\text{theorem: } 11.43]}{=} \sum_{k\in\{1,\ldots,n\}\setminus\{i,j\}}\alpha_k\cdot x_k+\sum_{k\in\{i\}}\alpha_k\cdot x_k+\sum_{k\in\{j\}}\alpha_k\cdot x_k\\ &\underset{[\text{theorem: } 11.34]}{=} \sum_{k\in\{1,\ldots,n\}\setminus\{i,j\}}\alpha_k\cdot x_k+\alpha_i\cdot x_i+\alpha_j\cdot x_j\\ &= \sum_{k\in\{1,\ldots,n\}\setminus\{i,j\}}0\cdot x_k+(\alpha_i/\alpha_j)\cdot x_i+(-1)\cdot x_j\\ &= x_j+(-x_j)\\ &= 0\end{aligned}$$

As $\{x_1,...,x_n\}$ is linear independent [because it is a basis] we have by [theorem: 11.115] that $\alpha_i/\alpha_j=0\Rightarrow\alpha_i=0$ contradicting $\{\alpha_i\}_{i\in\{1,\ldots,n\}}\subseteq F\setminus\{0\}$. So the assumption must be wrong, hence $\forall i,j\in\{1,\ldots,n\}$ with $y(i)=y(j)$ we have $i=j$, proving that y is injective. So

$$y:\{1,\ldots,n\}\rightarrow y(\{1,\ldots,n\})=\{y_i|i\in\{1,\ldots,n\}\}=\{\alpha_i\cdot x_i|i\in\{1,\ldots,n\}\}\text{ is a bijection}\tag{11.51}$$

defining the finite set

$$Y=\{y_1,\ldots,y_n\}=\{\alpha_1\cdot x_1,\ldots,\alpha_n\cdot x_n\}\tag{11.52}$$

For linear independency of $\{y_1,\ldots,y_n\}$, let $\{\gamma_k\}_{k\in\{1,\ldots,n\}}\subseteq F$ be such that $\sum_{k\in\{1,\ldots,n\}}\alpha_k\cdot y_k=0$ and define

$$\{\rho_k\}_{k\in\{1,\ldots,n\}}\subseteq F\text{ by }\rho_k=\alpha_k\cdot\gamma_k$$

then we have

$$\begin{aligned}\sum_{i\in\{1,\ldots,n\}}\rho_k\cdot x_k &= \sum_{i\in\{1,\ldots,n\}}(\alpha_k\cdot\gamma_k)\cdot x_k\\ &= \sum_{i\in\{1,\ldots,n\}}\gamma_k\cdot y_k\\ &= 0\end{aligned}$$

As $\{x_1,\ldots,x_n\}$ is linear independent it follows, using [theorem: 11.115] that $\forall k\in\{1,\ldots,n\}$ $\rho_k=0\Rightarrow\alpha_k\cdot\gamma_k=0\underset{\alpha_k\neq 0}{\Rightarrow}\gamma_k=0$, proving by [theorem: 11.115] that

$$\{y_1,\ldots,y_n\}\text{ is linear independent}\tag{11.53}$$

Let $v\in V$ then, as $\operatorname{span}(\{x_1,\ldots,x_n\})=V$, there exists by [theorem: 11.97] a $\{\omega_k\}_{k\in\{1,\ldots,n\}}\subseteq F$ such that $v=\sum_{k\in\{1,\ldots,n\}}\omega_k\cdot x_k$. As $\{\alpha_k\}_{k\in\{1,\ldots,n\}}\subseteq F\setminus\{0\}$ we can define

$$\{\xi_k\}_{k\in\{1,\ldots,n\}}\subseteq F\text{ by }\varepsilon_k=\omega_k\cdot(\alpha_k)^{-1}$$

then

$$\begin{aligned}\sum_{k\in\{1,\ldots,n\}}\xi_k\cdot y_k &= \sum_{k\in\{1,\ldots,n\}}(\omega_k\cdot(\alpha_k)^{-1})\cdot(\alpha_k\cdot x_k)\\ &= \sum_{k\in\{1,\ldots,n\}}\omega_k\cdot x_k\\ &= v\end{aligned}$$

Which proves that $V\subseteq\operatorname{span}(\{y_1,\ldots,y_n\})\subseteq V$ so that

$$V=\operatorname{span}(\{y_1,\ldots,y_n\})=\operatorname{span}(\{\alpha_1\cdot x_1,\ldots,\alpha_n\cdot x_n\})$$

combining the above with [eq: 11.53] proves finally that

$$\{y_1,\ldots,y_n\}\text{ is a basis of }V$$ □

Up to now we are not certain that every vector space has a basis. We will now use Zorn's lemma, a consequence of the Axiom of Choice, to prove that every vector space has a basis. We start with proving that every linear independent set that is a subset of a spanning set can be extended to a basis.

Theorem 11.139. *Let $\langle V,+,\cdot\rangle$ be a vector space over a field $\langle F,+,\cdot\rangle$, $R\subseteq W\subseteq V$ such that*

1. *R is linear independent*

2. $\operatorname{span}(W) = V$

then there exists a basis B of $\langle V, +, \cdot\rangle$ such that $R \subseteq B \subseteq W$

Proof. Define the set of linear independent sets between R and B

$$\mathcal{A} = \{X \subseteq V | R \subseteq X \subseteq W \text{ and } X \text{ is linear independent}\} \tag{11.54}$$

then we have, as R is linear independent and $R \subseteq R \subseteq W$, that

$$R \in \mathcal{A} \tag{11.55}$$

Using [example: 3.33] we can partial order $\mathcal{A}$ by the inclusion operator, in other words

$$\langle \mathcal{A}, \subseteq\rangle \text{ is a partial ordered set} \tag{11.56}$$

Let $\mathcal{C} \subseteq \mathcal{A}$ be a **non empty** chain [see definition: 3.41]. If $X \in \mathcal{C}$ then $X \in \mathcal{A}$ hence $R \subseteq X \subseteq W$, so that $R \subseteq \bigcup_{X \in \mathcal{C}} X \subseteq W$ giving

$$R \subseteq B_{\mathcal{C}} \subseteq W \text{ and } \forall X \in \mathcal{C} \text{ we have } X \subseteq B_{\mathcal{C}} \text{ where } B_{\mathcal{C}} = \bigcup_{X \in \mathcal{C}} X \tag{11.57}$$

We use now mathematical induction on the size of finite sets to proof that

$$\forall A \subseteq B_{\mathcal{C}} \text{ such that } A \text{ is finite there } \exists X \in \mathcal{C} \text{ such that } A \subseteq X \tag{11.58}$$

Proof. Define

$$\mathcal{S} = \{n \in \mathbb{N}_0 | \text{If } A \subseteq B_{\mathcal{C}} \text{ with } \{1, \ldots, n\} \approx A \text{ then } \exists X \in \mathcal{C} \text{ such that } A \subseteq X\}$$

then we have:

$0 \in \mathcal{S}$. If $A \subseteq B_{\mathcal{C}}$ with $A \approx \{1, \ldots, 0\} = \varnothing$ then by [theorem: 10.85] there exist a bijection

$$\beta \colon \{1, \ldots, 0\} = \varnothing \rightarrow A$$

hence $A = \beta(\varnothing)$ proving that $A = \varnothing$. As $\mathcal{C} \neq \varnothing$ there exists a $X \in \mathcal{C}$ and trivially $A = \varnothing \subseteq X$, proving $0 \in \mathcal{S}$.

$n \in \mathcal{S} \Rightarrow n + 1 \in \mathcal{S}$. Let $A \subseteq B_{\mathcal{C}}$ with $\{1, \ldots, n + 1\} \approx A$ then there exist a bijection

$$\beta \colon \{1, \ldots, n + 1\} \rightarrow A.$$

Take $A' = A \setminus \{\beta(n + 1)\}$ and consider $\beta_{|\{1, \ldots, n\}} \colon \{1, \ldots, n\} \rightarrow A'$ then we have:

injectivity. If $k, l \in \{1, \ldots, n\}$ and $\beta_{|\{1, \ldots, n\}}(k) = \beta_{|\{1, \ldots, n\}}(l)$ then $\beta(k) = \beta(l)$ which, as β is a bijection, proves that $k = l$

surjectivity. If $y \in A'$ then $y \in A$ and $y \neq \beta(n + 1)$. As β is a bijection there exists a $i \in \{1, \ldots, n + 1\}$ such that $\beta(i) = y$. If $i = n + 1$ then $y = \beta(i) = \beta(n + 1)$ contradicting $y \neq \beta(n + 1)$, so $i \in \{1, \ldots, n\}$, hence $y = \beta(i) = \beta_{|\{1, \ldots, n\}}(i)$.

proving that $\beta_{|\{1, \ldots, n\}} \colon \{1, \ldots, n\} \rightarrow A'$ is a bijection hence $\{1, \ldots, n\} \approx A'$. As $n \in S$ there exist a $X' \in \mathcal{C}$ such that $A' \subseteq X'$. Further as $\beta(n + 1) \in A \subseteq B_{\mathcal{C}} = \bigcup_{X \in \mathcal{C}} X$ there exists a $X \in \mathcal{C}$ such that $\beta(n + 1) \in X$. Now as $\mathcal{C}$ is a chain we have either:

$X' \subseteq X$. Then as $A \setminus \{\beta(n + 1)\} = A' \subseteq X' \subseteq X$ and $\beta(n + 1) \in X$ we have that $A \subseteq X$ proving that $n + 1 \in S$

$X \subseteq X'$. Then as $A \setminus \{\beta(n + 1)\} = A' \subseteq X'$ and $\beta(n + 1) \in X \subseteq X'$ we have that $A \subseteq X'$ proving that $n + 1 \in S$

so in both cases we have that $n + 1 \in S$.

By mathematical induction it follows that $S = \mathbb{N}_0$. So for a finite $A \subseteq B_{\mathcal{C}}$ we have by [theorem: 10.85] a $n \in \mathbb{N}_0$ such that $\{1, \ldots, n\} \approx A$. Hence as $n \in \mathbb{N}_0 = S$ there exists a $X \in \mathcal{C}$ such that $A \subseteq X$. □

We prove now that $B_{\mathcal{C}}$ is linear independent. Let $I \subseteq B_C$ be a finite set then by [eq: 11.58] we have

$$\exists Y \in \mathcal{C} \text{ such that } I \subseteq Y \tag{11.59}$$

Further as $Y \in \mathcal{C} \subseteq \mathcal{A}$ we have by [eq: 11.54] that

$$R \subseteq Y \subseteq W \text{ and } Y \text{ is linear independent} \tag{11.60}$$

So if $\{\alpha_u\}_{u \in I} \subseteq F$ is such that $\sum_{u \in I} \alpha_u \cdot u = 0$ then as $I \subseteq Y$ and Y is linear independent we have by the definition of linear independence that $\forall u \in I$ $\alpha_u = 0$, which, as $I \subseteq B_{\mathcal{C}}$, I finite was chosen arbitrary, proves that

$$B_{\mathcal{C}} \text{ is linear independent} \tag{11.61}$$

Combining this with [eqs: 11.54, 11.57] it follows that

$$B_{\mathcal{C}} \in \mathcal{A}$$

So for every **non empty** chain $\mathcal{C} \subseteq \mathcal{A}$ we have found a $B_{\mathcal{C}} \in \mathcal{A}$ such that $\forall X \in \mathcal{C}$ we have $X \subseteq \bigcup_{X \in \mathcal{C}} X = B_{\mathcal{C}}$ hence $\mathcal{C}$ has a upper bound $B_{\mathcal{C}}$. Further as by [eq: 11.55] $R \in \mathcal{A}$ we have for the empty chain $[\mathcal{C} = \varnothing]$ that $\forall X \in \varnothing\ X \subseteq R$ is satisfied variously, hence R is a upper bound of the empty chain.. So we have that

$$\text{every chain } \mathcal{C} \text{ in } \mathcal{A} \text{ has a upper bound in } \mathcal{A}$$

By Zorn's lemma [see theorem: 3.120] there exists a maximal element in $\mathcal{A}$, in other words

$$\exists B \in \mathcal{A} \text{ such that } \forall X \in \mathcal{A} \text{ we have } X \subseteq B \tag{11.62}$$

and as $B \in \mathcal{A}$ we have

$$R \subseteq B \subseteq W \text{ and } B \text{ is linear independent} \tag{11.63}$$

Let now $w \in W$ then we have either:

$\boldsymbol{w \in B}$**.** As by [theorem: 11.103] $B \subseteq \operatorname{span}(B)$ it follows that $w \in \operatorname{span}(B)$.

$\boldsymbol{w \notin B}$**.** Assume that $B \bigcup \{w\}$ is linear independent then as $R \subseteq B \subseteq B \bigcup \{w\} \subseteq W$ it follows that $B \bigcup \{w\} \in \mathcal{A}$, using [eq: 11.62] it follows that $B \bigcup \{w\} \subseteq B$ so that $w \in B$ contradicting $w \notin B$. Hence we must have that

$$B \bigcup \{w\} \text{ is linear dependent}$$

Hence there exists a finite $I \subseteq B \bigcup \{w\}$ and a $\{\alpha_u\}_{u \in I} \subseteq F$ satisfying $\exists u_0 \in I$ with $\alpha_{u_0} \neq 0$ such that $\sum_{u \in I} \alpha_u \cdot u = 0$. If $w \notin I$ then $I \subseteq B$ which as B is linear independent would mean that $\forall u \in I$ we have $\alpha_u = 0$ contradicting $\alpha_{u_0} \neq 0$. So we must have

$$w \in I \tag{11.64}$$

Further

$$\begin{aligned} 0 &= \sum_{u \in I} \alpha_u \cdot u \\ &\underset{[\text{theorem: } 11.43]}{=} \sum_{u \in I \setminus \{w\}} \alpha_u \cdot u + \sum_{u \in \{w\}} \alpha_u \cdot u \\ &\underset{[\text{theorem: } 11.34]}{=} \sum_{u \in I \setminus \{w\}} \alpha_u \cdot u + \alpha_w \cdot w \end{aligned} \tag{11.65}$$

Note as $I \subseteq B \bigcup \{w\}$ we have $I \setminus \{w\} \subseteq (B \bigcup \{w\}) \setminus \{w\} \subseteq B$ so that

$$I \setminus \{w\} \subseteq B \tag{11.66}$$

Assume $\alpha_w = 0$ then from [eq: 11.65] we have $0 = \sum_{u \in I \setminus \{w\}} \alpha_u \cdot u$ which, as $I \setminus \{w\} \subseteq B$ [see eq: 11.66] and B is linear independent, gives $\forall u \in I \setminus \{w\}\ \alpha_u = 0$. So, as we assumed $\alpha_w = 0$, we have $\forall u \in I$ that $\alpha_u = 0$ contradicting $\alpha_{u_0} \neq 0$. Hence we must have that $\alpha_w \neq 0$, and applying this on [eq: 11.65] proves that

$$w = (\alpha_w)^{-1} \cdot \left(- \sum_{u \in I \setminus \{w\}} \alpha_u \cdot u \right) \underset{[\text{theorems: } 11.40, 11.75]}{=} \sum_{u \in I} ((\alpha_u)^{-1} \cdot (-\alpha_u)) \cdot u = \sum_{u \in I} \lambda_u \cdot u$$

where $\{\lambda_u\}_{u \in I \setminus \{W\}} \subseteq F$. So as $I \setminus \{w\} \subseteq B$ it follows that

$$w \in \operatorname{span}(B)$$

So, as in all cases we have $w \in \operatorname{span}(B)$, it follows that $W \subseteq \operatorname{span}(B)$, hence

$$V = \operatorname{span}(W) \subseteq_{[\text{theorem: } 11.104]} \operatorname{span}(\operatorname{span}(B)) \underset{[\text{theorem: } 11.107]}{=} \operatorname{span}(B) \subseteq V$$

proving that

$$\operatorname{span}(B) = V \tag{11.67}$$

By [eqs: 11.63, 11.67] we have that B is a basis of $\langle V, +, \cdot \rangle$ and $R \subseteq B \subseteq W$ proving the theorem. □

Corollary 11.140. *Let $\langle V, +, \cdot \rangle$ be a vector space over the field $\langle F, +, \cdot \rangle$ and R a linear independent set then there exists a basis B of $\langle V, +, \cdot \rangle$ such that $R \subseteq B$. In other words a linear independent set of a vector space can be extended to a basis of the vector space.*

Proof. As R is linear independent, $R \subseteq V$ and $\operatorname{span}(V) \underset{[\text{theorem: } 11.106]}{=} V$ we have by [theorem: 11.139] that there exist a basis B of $\langle V, +, \cdot \rangle$ with $R \subseteq B \subseteq V$. □

Corollary 11.141. *Let $\langle V,+,\cdot\rangle$ be a vector space over the field $\langle F,+,\cdot\rangle$ and $W\subseteq V$ such that* $\operatorname{span}(W)=V$ *then there exist a basis B of $\langle V,+,\cdot\rangle$ such that $B\subseteq W$. In other words every spanning set of a vector space can be reduced to a basis of the vector space.*

Proof. For V we have either:

$\boldsymbol{V=\{0\}}$**.** By [example: 11.132] $\varnothing$ is a basis for V and trivially $\varnothing\subseteq W$.

$\boldsymbol{V\neq\{0\}}$**.** As $V=\operatorname{span}(W)$ and $\operatorname{span}(\varnothing)\underset{\text{[example: 11.99]}}{=}\{0\}$ we must have $W\neq\varnothing$, also if $W=\{0\}$ then by [example: 11.100] $\operatorname{span}(W)=\{0\}$. Hence there exist a $w\in W$ such that $w\neq 0$. Using [example: 11.117] we have that $R=\{w\}$ is linear independent. As further $R=\{w\}\subseteq W$ we have by [theorem: 11.139] that there exist a basis B of $\langle V,+,\cdot\rangle$ such that $R\subseteq B\subseteq W$. □

Corollary 11.142. *Let $\langle V,+,\cdot\rangle$ be a vector space over the field $\langle F,+,\cdot\rangle$ then there exist a basis $B\subseteq V$ of $\langle V,+,\cdot\rangle$.*

Proof. As $\operatorname{span}(V)\underset{\text{[theorem: 11.106]}}{=}V$ we have by [corollary: 11.141] that there exist a basis B of $\langle V,+,\cdot\rangle$. □

11.3.4 Dimension of a vector space

Lemma 11.143. *Let $\langle V,+,\cdot\rangle$ be a vector space over a field $\langle F,+,\cdot\rangle$ that has a* ***finite*** *basis $B\subseteq V$ then every basis of $\langle V,+,\cdot\rangle$ is* ***finite****. Further for every basis A of $\langle V,+,\cdot\rangle$ we have* $\operatorname{card}(B)=\operatorname{card}(A)$

Proof. As B is a basis we have that

$$\operatorname{span}(B)=V,\ B \text{ is linear independent and } B \text{ is finite}$$

Let A be another basis of $\langle V,+,\cdot\rangle$ then

$$\operatorname{span}(A)=V \text{ and } A \text{ is linear independent}$$

By [corollary: 11.121] it follows that A is finite. Further as A,B are finite we have by [theorem: 11.84 that

$$A=\{a_1,\ldots,a_n\},\ B=\{b_1,\ldots,b_m\} \text{ where } n=\operatorname{card}(A) \text{ and } m=\operatorname{card}(B)$$

From $V=\operatorname{span}(B)$ and the linear independence of A it follows from [corollary: 11.120] that

$$n\leqslant\operatorname{card}(B)=m$$

From $V=\operatorname{span}(A)$ and the linear independency of B it follows from [corollary: 11.120] that

$$m\leqslant\operatorname{card}(A)=n$$

So we have that $n=m$ proving that

$$\operatorname{card}(A)=\operatorname{card}(B)$$

□

Corollary 11.144. *Let $\langle V,+,\cdot\rangle$ be a vector space over a field $\langle F,+,\cdot\rangle$ that has a* ***infinite*** *basis $B\subseteq V$ then every basis of $\langle V,+,\cdot\rangle$ is* ***infinite****.*

Proof. Assume that there exists a finite basis of $\langle V,+,\cdot\rangle$ then by [lemma: 11.143] B is finite a contradiction. So every basis of $\langle V,+,\cdot\rangle$ must be infinite. □

Definition 11.145. *Let $\langle V,+,\cdot\rangle$ be a vector space over a field $\langle F,+,\cdot\rangle$ then V is* ***finite dimensional*** *if V has a finite basis and V is* ***infinite dimensional*** *if V has infinite basis. Using [corollary: 11.142], [lemma: 11.143] and [corollary: 11.144] we have that:*

1. *V is either finite dimensional or infinite dimensional but not both.*
2. *If V is finite dimensional then every basis of V is finite and have the same cardinality.*
3. *If V is infinite dimensional then every basis of V is infinite.*

The above definition allows us to define $\dim(V)$.

Definition 11.146. *Let $\langle V,+,\cdot\rangle$ be a vector space over a field $\langle F,+,\cdot\rangle$ then* $\dim(V)$ *is defined by*

$$\dim(V)=\begin{cases}\infty \text{ if } B \text{ is infinite}\\ \operatorname{card}(B) \text{ if } B \text{ is finite}\end{cases}$$

where B is basis of V. By [corollary: 11.142], [lemma: 11.143] and [corollary: 11.144] this is well defined.

Example 11.147. Let $\langle \{0\}, +, \cdot \rangle$ be the trivial vector space over a field $\langle F, +, \cdot \rangle$ [see: 11.59] then

$$\dim(\{0\}) = 0$$

Further if $\langle V, +, \cdot \rangle$ is a vector space over the field $\langle F, +, \cdot \rangle$ with $\dim(X) = 0$ then $V = \{0\}$

Proof. By [example: 11.132] $\varnothing$ is a basis of $\langle \{0\}, +, \cdot \rangle$ so that $\dim(\{0\}) = \text{card}(\varnothing) = 0$. Further if $\langle V, +, \cdot \rangle$ is a vector space over the field $\langle F, +, \cdot \rangle$ with $\dim(X) = 0$ then if B is a basis for $\langle V, +, \cdot \rangle$ we must have that $\text{card}(B) = \dim(X) = 0$. Hence $B = \varnothing$, so by [theorem: 11.99] $V = \text{span}(B) = \text{span}(\varnothing) = \{0\}$. $\square$

Theorem 11.148. *Let $\langle V, +, \cdot \rangle$ be a vector space over a field $\langle F, +, \cdot \rangle$, $n \in \mathbb{N}$ and $\{e_i\}_{i \in \{1, \ldots, n\}} \subseteq V$ then $\text{span}(\{e_i | i \in \{1, \ldots, n\}\})$ is finite dimensional and $\dim(\text{span}(\{e_i | i \in \{1, \ldots, n\}\})) \leqslant n$*

Proof. First using [theorem: 11.89] we have that

$$\{e_i | i \in \{1, \ldots, n\}\} \text{ is finite and } \text{card}(\{e_i | i \in \{1, \ldots, n\}\}) \leqslant n$$

Using [theorem: 11.106, 11.57] we have that $\text{span}(\{e_i | i \in \{1, \ldots, n\}\})$ is a vector space and trivially equal to itself, hence using [theorem: 11.141] there exists a basis $B \subseteq \{e_i | i \in \{1, \ldots, n\}\}$ of $\text{span}(\{e_i | i \in \{1, \ldots, n\}\})$. Further as $B \subseteq \{e_i | i \in \{1, \ldots, n\}\}$ we have by [theorem: 10.89] that B is finite and

$$\dim(\text{span}(\{e_i | i \in \{1, \ldots, n\}\})) = \text{card}(B) \leqslant \text{card}(\{e_i | i \in \{1, \ldots, n\}\}) \leqslant n$$ $\square$

In a finite dimensional vector space we have that every linear independent set with the same number of elements as the dimension of the vector space is automatically a basis of the vector space.

Theorem 11.149. *Let $\langle V, +, \cdot \rangle$ be a finite dimensional vector space over a field $\langle F, +, \cdot \rangle$ with $\dim(V) = n$ and $B \subseteq V$ a finite linear independent set with $\text{card}(B) = n$ then B is a basis for $\langle V, +, \cdot \rangle$.*

Proof. As $\dim(V) = n$ there exist a basis B' with $\text{card}(B') = n$. Using [theorem: 11.140] there exist a basis B'' such that $B \subseteq B''$. From [theorem: 11.143] we have that $\text{card}(B'') = \text{card}(B') = n$. So we have that $B \subseteq B''$ and $\text{card}(B) = \text{card}(B'')$, applying then [theorem: 10.93] proves that $B = B''$, hence B is a basis. $\square$

Corollary 11.150. *Let $\langle V, +, \cdot \rangle$ be a finite dimensional vector space over a field $\langle F, +, \cdot \rangle$ with $\dim(V) = n$ then if $\{v_i\}_{i \in \{1, \ldots, n\}} \subseteq V$ is linear independent we have that*

1. *$\{v_i | i \in \{1, \ldots, n\}\} = \{v_1, \ldots, v_n\}$ [or by [definition: 11.79] that $v: \{1, \ldots, n\} \rightarrow \{v_i | i \in \{1, \ldots, n\}\}$ is a bijection].*
2. *$\{v_i | i \in \{1, \ldots, n\}\}$ is a basis of V*
3. *$\dim(V) = n$*

Proof.

1. This follows from [theorem: 11.129 (2)]
2. As $\{v_i\}_{i \in \{1, \ldots, n\}} \subseteq V$ is linear independent we have by [theorem: 11.129] that

   $$\{v_i | i \in \{1, \ldots, n\}\} \text{ is linear independent}$$

   From (1) it follows that $\text{card}(\{v_i | i \in \{1, \ldots, n\}\}) = n$, so applying the previous theorem [theorem: 11.149], we have that

   $$\{v_i | i \in \{1, \ldots, n\}\} \text{ is a basis for } V$$

3. From (1) it follows that $\text{card}(\{v_i | i \in \{1, \ldots, n\}\}) = n$ and from (2) that $\{v_i | i \in \{1, \ldots, n\}\}$ is a basis of v hence $\dim(V) = n$. $\square$

Theorem 11.151. *Let $\langle V, +, \cdot \rangle$ be a finite dimensional vector space over a field $\langle F, +, \cdot \rangle$ with $\dim(V) = n$ then for every $\{v_i\}_{i \in \{1, \ldots, n\}} \subseteq V$ with $\text{span}(\{v_i | i \in \{1, \ldots, n\}\}) = V$ we have that*

$$\{v_i | i \in \{1, \ldots, n\}\} \textit{ is a basis of } V$$

Proof. As $\text{span}(\{v_i | i \in \{1, \ldots, n\}\}) = V$ there exists by [theorem: 11.141] a basis $B \subseteq \{v_i | i \in \{1, \ldots, n\}\}$. As $\dim(V) = n$ we must have that $\text{card}(B) = n$, further by [theorem: 10.94] $\text{card}(\{v_i | i \in \{1, \ldots, n\}\}) \leqslant n$ and as $B \subseteq \{v_i | i \in \{1, \ldots, n\}\}$ we have by [theorem: 10.89] $n = \text{card}(B) \leqslant \text{card}(\{v_i | i \in \{1, \ldots, n\}\})$. So we have that $\text{card}(\{v_i | i \in \{1, \ldots, n\}\}) = \text{card}(B)$. Finally from [theorem: 10.93] it follows that $B = \{v_i | i \in \{1, \ldots, n\}\}$ hence $\{v_i | i \in \{1, \ldots, n\}\}$ is a basis for V. $\square$

Theorem 11.152. *Let $\langle V,+,\cdot\rangle$ be a finite dimensional vector space over a field $\langle F,+,\cdot\rangle$ and W a sub-space of $\langle V,+,\cdot\rangle$ then $\dim(W)\leqslant\dim(V)$ and if V is finite dimensional W is finite dimensional.*

Proof. For $\dim(V)$ we have for V either:

- ***V* is infinite dimensional.** Then V has a infinite basis so that by definition $\dim(V)=\infty$, using the definition of the order on the extended real numbers it follows that $\dim(W)\leqslant\dim(V)$.
- ***V* is finite dimensional.** Let $B_W\subseteq W$ a basis of W then B_W is linear independent and $B_W\subseteq V=\operatorname{span}(V)$. Using [theorem: 11.139] there exist then a basis B_V of $\langle V,+,\cdot\rangle$ such that $B_W\subseteq B_V\subseteq V$. As V is finite dimensional B_V is finite and using [theorem: 10.89] it follows that B_W is finite and $\operatorname{card}(B_W)\leqslant\operatorname{card}(B_V)$. Hence we have that W is finite dimensional with

$$\dim(W)=\operatorname{card}(B_W)\leqslant\operatorname{card}(B_V)=\dim(V)$$

□

Theorem 11.153. *Let $\langle F,+,\cdot\rangle$ a field and consider the vector space of $\langle F,+,\cdot\rangle$ over $\langle F,+,\cdot\rangle$ [see theorem: 11.60]. Then if $1\in F$ is the multiplicative neutral element we have that $\{1\}$ is a basis of $\langle F,+,\cdot\rangle$, so $\dim(F)=1$*

Proof. First if $\{\alpha_i\}_{i\in\{1\}}\subseteq F$ is such that $\sum_{i\in\{1\}}\alpha_i\cdot i=0$ then as

$$0=\sum_{i\in\{1\}}\alpha_i\cdot i\underset{\text{[theorem: 11.34]}}{=}\alpha_1\cdot 1=\alpha_1$$

hence $\forall i\in\{1\}$ $\alpha_i=0$ proving that

$$\{1\}\text{ is linear independent}$$

Further if $v\in F$ then, if we take $\{\alpha_i\}_{i\in\{1\}}\subseteq F$ defined by $\alpha_1=v$, we have

$$v=v\cdot 1\underset{\text{[theorem: 11.34]}}{=}\sum_{i\in\{1\}}\alpha_i\cdot e_i$$

proving that

$$\operatorname{span}(\{1\})=F$$

Hence we have that

$$\{1\}\text{ is a basis for }\langle F,+,\cdot\rangle\text{ and }\dim(F)=\operatorname{card}(\{1\})=1$$

□

The above theorem proves automatically the following .

Corollary 11.154. *We have [see: 11.61]*

1. *The vector space $\langle\mathbb{Q},+,\cdot\rangle$ over $\langle\mathbb{Q},+,\cdot\rangle$ has as basis $\{1\}$ and $\dim(\mathbb{Q})=1$*
2. *The vector space $\langle\mathbb{R},+,\cdot\rangle$ over $\langle\mathbb{R},+,\cdot\rangle$ has as basis $\{1\}$ and $\dim(\mathbb{R})=1$*
3. *The vector space $\langle\mathbb{C},+,\cdot\rangle$ over $\langle\mathbb{C},+,\cdot\rangle$ has as basis $\{1\}$ and $\dim(\mathbb{C})=1$*

Be aware that that the field for a vector space is important as the following shows.

Theorem 11.155. *The vector space $\langle\mathbb{C},+,\cdot\rangle$ over $\langle\mathbb{R},+,\cdot\rangle$ [see example: 11.62] has as basis $\{1,i\}$ so that $\dim(\mathbb{C})=2$*

Proof. Let $\{\alpha_u\}_{u\in\{1,i\}}\subseteq\mathbb{R}$ be such that $\sum_{u\in\{1,i\}}\alpha_u\cdot u=0$ then we have

$$\begin{aligned}0+0\cdot i &= \sum_{u\in\{1,i\}}\alpha_u\cdot u\\ &= \alpha_1\cdot 1+\alpha_i\cdot i\\ &= \alpha_1+\alpha_i\cdot i\end{aligned}$$

so that $\alpha_1=0=\alpha_i$ proving that

$$\{1,i\}\text{ is linear independent}$$

Further if $v\in\mathbb{C}$ then there exists $x,y\in\mathbb{R}$ so that $v=x+i\cdot y$. Hence if we define $\{\alpha_u\}_{u\in\{1,i\}}$ by $\alpha_1=x$ and $\alpha_i=y$ we have

$$\sum_{u\in\{1,i\}}\alpha_i\cdot e_i=\alpha_1\cdot 1+\alpha_i\cdot i=x\cdot 1+y\cdot i=x+i\cdot y=v$$

proving that

$$\text{span}(\{1,i\}) = \mathbb{C}$$

□

Now for a basis of higher dimensional space we introduce the Kronecker delta.

Definition 11.156. (Kronecker delta) *Let $n \in \mathbb{N}_0$ and $\langle F,+,\cdot\rangle$ a field with additive neutral element 0 and multiplicative neutral element 1 then $\{\delta_{i,j}^{[n]}\}_{(i,j)\in\{1,\ldots,n\}\times\{1,\ldots,n\}}$ is defined by*

$$\delta_{i,j}^{[n]} = \begin{cases} 1 \text{ if } i=j \\ 0 \text{ if } i \neq j \end{cases}$$

If n is know from the context then we write $\delta_{i,j}$ instead of $\delta_{i,j}^{[n]}$

Theorem 11.157. *If $n \in \mathbb{N}$ and $\langle F,+,\cdot\rangle$ a field, $j \in \{1,\ldots,n\}$ and $\{\alpha_i\}_{i\in\{1,\ldots,n\}} \subseteq F$ then*

$$\sum_{i\in\{1,\ldots,n\}} \delta_{i,j}\cdot\alpha_i = \alpha_j$$

Proof. We have

$$\begin{aligned}
\sum_{i\in\{1,\ldots,n\}} \delta_{i,j}\cdot\alpha_i &\underset{[\text{theorem: } 11.43]}{=} \sum_{i\in\{1,\ldots,n\}\setminus\{j\}} \delta_{i,j}\cdot\alpha_i + \sum_{i\in\{j\}} \delta_{i,j}\cdot\alpha_i \\
&\underset{[\text{theorem: } 11.34]}{=} \sum_{i\in\{1,\ldots,n\}\setminus\{j\}} \delta_{i,j}\cdot\alpha_i + \delta_{j,j}\cdot\alpha_j \\
&= \sum_{i\in\{1,\ldots,n\}\setminus\{j\}} (0\cdot\alpha_i) + 1\cdot\alpha_j \\
&\underset{[\text{theorem: } 11.37]}{=} \alpha_j
\end{aligned}$$

□

We have a similar theorem for vector spaces.

Theorem 11.158. *If $n \in \mathbb{N}$, $\langle X,+,\cdot\rangle$ a vector space over a field $\langle F,+,\cdot\rangle$ a field, $j \in \{1,\ldots,n\}$ and $\{x_i\}_{i\in\{1,\ldots,n\}} \subseteq X$ then*

$$\sum_{i\in\{1,\ldots,n\}} \delta_{i,j}\cdot x_i = x_j$$

Proof. We have

$$\begin{aligned}
\sum_{i\in\{1,\ldots,n\}} \delta_{i,j}\cdot x_i &\underset{[\text{theorem: } 11.43]}{=} \sum_{i\in\{1,\ldots,n\}\setminus\{j\}} \delta_{i,j}\cdot x_i + \sum_{i\in\{j\}} \delta_{i,j}\cdot x_i \\
&\underset{[\text{theorem: } 11.34]}{=} \sum_{i\in\{1,\ldots,n\}\setminus\{j\}} \delta_{i,j}\cdot x_i + \delta_{j,j}\cdot x_j \\
&= \sum_{i\in\{1,\ldots,n\}\setminus\{j\}} (0\cdot x_i) + 1\cdot x_j \\
&\underset{[\text{theorem: } 11.37]}{=} x_j
\end{aligned}$$

□

Theorem 11.159. *let $n \in \mathbb{N}$, $\langle F,+,\cdot\rangle$ a field and $\langle F^n,+,\cdot\rangle$ the vector space over $\langle F,+,\cdot\rangle$ [see theorem: 11.66] define*

$$e: \{1,\ldots,n\} \rightarrow F^n \text{ where } e(i) \in F^n \text{ is defined by } (e(i))_j = \delta_{i,j}$$

then e is injective so that

$$e: \{1,\ldots,n\} \rightarrow e(\{1,\ldots,n\}) = E \text{ is a bijection}$$

defining $E = \{e_1,\ldots,e_n\}$. Then $E = \{e_1,\ldots,e_n\}$ is a basis for F^n so that $\dim(F^n) = n$. Further

$$\forall x = (x_1,\ldots,x_n) \in F^n \text{ we have } x = \sum_{i\in\{1,\ldots,n\}} x_i\cdot e_i$$

Proof. Let $i, j \in \{1, \ldots, n\}$ such that $e_i = e_j$. Assume that $i \neq j$ then we have

$$1 = \delta_{i,i} = (e_i)(i) = (e_j)(i) = \delta_{i,j} = 0$$

leading to the contradiction $1 = 0$ [see definition: 4.51], hence we must have that $e: \{1, \ldots, n\} \to F^n$ is a injection and thus that $e: \{1, \ldots, n\} \to F_n$ is a bijection defining the finite distinct set

$$E = \{e_1, \ldots, e_n\}$$

Let now $x = (x_1, \ldots, x_n) \in F^n$ then for $\{x_i\}_{i \in \{1, \ldots, n\}} \subseteq F$ we have that $\forall j \in \{1, \ldots, n\}$

$$\begin{aligned} \left(\sum_{i \in \{1, \ldots, n\}} x_i \cdot e_i \right)_j &\underset{[\text{theorem: } 11.47]}{=} \sum_{i \in \{1, \ldots, n\}} (x_i \cdot e_i)_j \\ &= \sum_{i \in \{1, \ldots, n\}} x_i \cdot (e_i)_j \\ &= \sum_{i \in \{1, \ldots, n\}} x_i \cdot \delta_{i,j} \\ &\underset{[\text{theorem: } 11.157]}{=} x_j \end{aligned}$$

proving that

$$\sum_{i \in \{1, \ldots, n\}} x_i \cdot e_i = x$$

so that $x \in \operatorname{span}(\{e_1, \ldots, e_n\})$, proving $F_n \subseteq \operatorname{span}(\{e_1, \ldots, e_n\}) \subseteq F^n$. Hence we have

$$\operatorname{span}(E) = F^n \tag{11.68}$$

Further if $\{\alpha_i\}_{i \in \{1, \ldots, n\}} \subseteq F$ is such that $\sum_{i \in \{1, \ldots, n\}} \alpha_i \cdot e_i = 0$ then for every $j \in \{1, \ldots, n\}$ we have:

$$\begin{aligned} 0 &= 0_j \\ &= \left(\sum_{i \in \{1, \ldots, n\}} \alpha_i \cdot e_i \right)_j \\ &\underset{[\text{theorem: } 11.47]}{=} \sum_{i \in \{1, \ldots, n\}} \alpha_i \cdot (e_i)_j \\ &= \sum_{i \in \{1, \ldots, n\}} \alpha_i \cdot \delta_{i,j} \\ &\underset{[\text{theorem: } 11.157]}{=} \alpha_j \end{aligned}$$

proving by [theorem: 11.115] that

$$E \text{ is linear independent} \tag{11.69}$$

From [eqs: 11.68, 11.69] it follows that

$$E \text{ is a basis of } F^n$$

□

11.4 Linear mappings

11.4.1 Linear mappings

Similar to the concepts of group, ring, field homeomorphisms we have also mappings that preserves the structure of vector spaces, these are called linear mappings. Linear mappings are very important in Banach spaces, Hilbert spaces, differential analysis and so on.

Definition 11.160. (Linear Mapping) *If $\langle X, +, \cdot \rangle$ and $\langle Y, +, \cdot \rangle$ are vector spaces over a field $\langle F, +, \cdot \rangle$ then a function $L: X \to Y$ is a **linear mapping** if*

1. *$\forall x, y \in X$ we have $L(x + y) = L(x) + L(y)$*
2. *$\forall x \in X$, $\forall \alpha \in F$ we have that $L(\alpha \cdot x) = \alpha \cdot L(x)$*

The set of graphs of linear mappings is $\mathrm{Hom}(X,Y)$ *so that*

$$\mathrm{Hom}\,(X,Y)=\{L\in Y^X|L:X\to Y \text{ is a linear mapping}\}$$

If $L\in\mathrm{Hom}(X,X)$ *then* $L:X\to Y$ *is called a* ***linear transformation****.*

Theorem 11.161. *Let* $\langle X,+,\cdot\rangle$, $\langle Y,+,\cdot\rangle$ *are vector spaces over a field* $\langle F,+,\cdot\rangle$ *and* $L\in\mathrm{Hom}(X,Y)$ *then*

1. *If* $x\in X$ *and* $A\subseteq X$ *then* $L(x+A)=L(x)+L(A)$
2. *If* $A,B\subseteq X$ *then* $L(A+B)=L(A)+L(B)$
3. *If* $A\subseteq X$ *and* $\alpha\in F$ *then* $L(\alpha\cdot A)=\alpha\cdot L(A)$

Proof.

1. If $y\in L(x+A)$ then there exist a $z\in x+A$ such that $y=L(z)$. As $z\in z+A$ there exist a $a\in A$ such that $z=x+a$ hence $y=L(z)=L(x+a)=L(x)+L(a)\in L(x)+L(A)$. Likewise if $y\in L(x)+L(A)$ then there exist a $z\in L(A)$ such that $y=L(x)+z$. As $z\in L(A)$ there exist a $a\in A$ such that $z=L(a)$. Hence $y=L(x)+z=L(x)+L(a)=L(x+a)\in L(x+A)$.
2. If $y\in L(A+B)$ then there exists a $a\in A$, $b\in B$ such that $y=L(a+b)=L(a)+L(b)$ so that $y\in L(a)+L(b)$. Likewise if $y\in L(A)+L(B)$ there exist a $a\in A$, $b\in B$ such that $y=L(a)+L(b)=L(a+b)$ proving, as $a+b\in A+B$, that $y\in L(A+B)$.
3. If $y\in L(\alpha\cdot A)$ then there exist a $a\in A$ such that $y=L(\alpha\cdot a)=\alpha\cdot L(a)$ so that $y\in\alpha\cdot L(A)$. Likewise if $y\in\alpha\cdot L(A)$ then there a $a\in A$ such that $y=\alpha\cdot L(a)=L(\alpha\cdot a)$ proving that $y\in L(\alpha\cdot A)$. □

Theorem 11.162. *Let* $\langle X,+,\cdot\rangle$ *and* $\langle Y,+,\cdot\rangle$ *are vector spaces over a field* $\langle F,+,\cdot\rangle$ *and* $L\in\mathrm{Hom}(X,Y)$ *is a linear mapping then* $L(0)=0$.

Proof. $L(0)=L(0\cdot 0)=0\cdot L(0)=0$ □

A equivalent definition for a linear mapping is the following theorem:

Theorem 11.163. *If* $\langle X,+,\cdot\rangle$ *and* $\langle Y,+,\cdot\rangle$ *are vector spaces over a field* $\langle F,+,\cdot\rangle$ *then we have for a function* $L:X\to Y$ *that*

$$L:X\to Y \text{ is a linear mapping}$$
$$\Updownarrow$$
$$\forall x,y\in X \text{ and } \alpha\in F \text{ we have } L(x+\alpha\cdot y)=L(x)+\alpha\cdot L(y)$$

Proof.

⇒. Let $x,y\in X$ and $\alpha\in F$ we have as L is linear that

$$L(x+\alpha\cdot y)=L(x)+L(\alpha\cdot y)=L(x)+\alpha\cdot L(y)$$

⇐. Let $x,y\in X$ and $\alpha\in F$ then we have

$$L(x+y)=L(x+1\cdot y)=L(x)+1\cdot L(y)=L(x)+L(y)$$

and

$$L(\alpha\cdot x)=L(0+\alpha\cdot x)=L(0)+\alpha\cdot L(x)\underset{\text{[theorem: 11.162]}}{=}0+\alpha\cdot L(x)=\alpha\cdot L(x)$$

proving by definition that $L:X\to Y$ is linear. □

Example 11.164. Consider the vector space $\langle\mathbb{C},+,\cdot\rangle$ and $\langle\mathbb{R},+,\cdot\rangle$ over $\langle\mathbb{R},+,\cdot\rangle$ then we have that

1. $\mathrm{Img}:\mathbb{C}\to\mathbb{R}$ defined by $\mathrm{Img}(x+i\cdot y)=y$ is linear
2. $\mathrm{Re}:\mathbb{C}\to\mathbb{R}$ defined by $\mathrm{Re}(x+i\cdot y)=x$ is linear
3. If $\{x_i\}_{i\in I}\subseteq\mathbb{C}$ is family with finite support then
   a. $\mathrm{Img}(\sum_{i\in I}x_i)=\sum_{i\in I}\mathrm{Img}(x_i)$
   b. $\mathrm{re}(\sum_{i\in I}x_i)=\sum_{i\in I}\mathrm{Re}(x_i)$

Proof.

1. This follows from [theorem: 10.78].
2. This follows from [theorem: 10.78].

3. This follows from (1), (2) and [theorem: 11.180] □

Example 11.165. Let $n \in \mathbb{N}$, $\langle F, +, \cdot \rangle$ be a field and $\langle F^n, +, \cdot \rangle$ be the vector space over $\langle F, +, \cdot \rangle$ defined by [theorem: 11.66] then $\forall i \in \{1, \ldots, n\}$ we have that the projection mapping

$$\pi_i: F^n \rightarrow F$$

is a linear mapping.

Proof. If $x, y \in F^n$ and $\alpha \in F$ then we have

$$\pi_i(x + y) = (x + y)_i = x_i + y_i = \pi_i(x) + \pi_i(y)$$

and

$$\pi_i(\alpha \cdot x) = \alpha \cdot x_i = \alpha \cdot \pi_i(x)$$ □

We can extend the above example to a more general case.

Example 11.166. Let $\{\langle X_i, +_i, \cdot_i \rangle\} i \in I$ be a family of vector spaces over a field $\langle F, +, \cdot \rangle$ and $\langle \prod_{i \in I} X_i, +, \cdot \rangle$ the vector space defined in [theorem: 11.71] then $\forall i \in I$

$$\pi_i: \prod_{i \in I} X_i \rightarrow X_i \text{ defined by } \pi_i(x) = x_i \text{ is a linear mapping}$$

Proof. Let $x, y \in \prod_{i \in I} X_i$ and $\alpha \in F$ then we have

$$\pi_i(x + y) = (x + y)_i \underset{[\text{theorem: } 11.71]}{=} x_i +_i y_i = \pi_i(x) +_i \pi_i(y)$$

and

$$\pi_i(\alpha \cdot x) = (\alpha \cdot x)_i = \alpha \cdot_i x_i = \alpha \cdot_i \pi_i(x)$$ □

Definition 11.167. (Linear Isomorphism) *If $\langle X, +, \cdot \rangle$ and $\langle Y, +, \cdot \rangle$ are vector spaces over a field $\langle F, +, \cdot \rangle$ then a mapping $L: X \rightarrow Y$ is a **linear isomorphism** iff*

1. *$L: X \rightarrow Y$ is a bijection*
2. *$L: X \rightarrow Y$ is a linear mapping*

*If between two vector spaces a linear isomorphism exists then we say that the **vector spaces are isomorphic**.*

Example 11.168. Let $\langle \mathbb{R}^2, +, \cdot \rangle$ and $\langle \mathbb{C}, +, \cdot \rangle$ be vector spaces over the field $\langle R, +, \cdot \rangle$ [see theorems: 11.159, 11.155] then $\mathcal{C}: \mathbb{R}^2 \rightarrow \mathbb{C}$ defined by $\mathcal{C}((x, y)) = x + i \cdot y$ is a linear isomorphism so that $\langle \mathbb{R}^2, +, \cdot \rangle$ is isomorphic with $\langle \mathbb{C}, +, \cdot \rangle$.

Proof. If $\mathcal{C}((x, y)) = \mathcal{C}((x', y'))$ then $x + i \cdot y = x' + i \cdot y'$ then by [theorem: 10.75] $x = x'$ and $y = y'$ so that $(x, y) = (x', y')$, hence

$$\mathcal{C}: \mathbb{R}^2 \rightarrow \mathbb{C} \text{ is injective}$$

Further if $z \in \mathbb{C}$ then $z = x + i \cdot y = \mathcal{C}((x, y))$ proving surjectivity. combining this with the above proves that

$$\mathcal{C}: \mathbb{R}^2 \rightarrow \mathbb{C} \text{ is a bijection}$$

Let $(x, y), (x', y') \in \mathbb{R}^2$ and $\alpha \in \mathbb{R}$

$$\begin{aligned}
\mathcal{C}((x, y) + (x', y')) &= \mathcal{C}((x + x', y + y')) \\
&= (x + x') + i \cdot (y + y') \\
&= (x + i \cdot y) + (x' + i \cdot y') \\
&= \mathcal{C}((x, y)) + \mathcal{C}((x', y')) \\
\mathcal{C}(\alpha \cdot (x, y)) &= \mathcal{C}((\alpha \cdot x, \alpha \cdot y)) \\
&= (\alpha \cdot x) + i \cdot (\alpha \cdot y) \\
&= \alpha \cdot (x + i \cdot y) \\
&= \alpha \cdot \mathcal{C}((x, y))
\end{aligned}$$

proving that

$$\mathcal{C}: \mathbb{R}^2 \rightarrow \mathbb{C} \text{ is a isomorphism}$$ □

Theorem 11.169. *Let $\langle X,+,\cdot\rangle$ and $\langle Y,+,\cdot\rangle$ be vector spaces over the field $\langle F,+,\cdot\rangle$ and*

$$L: X\rightarrow Y \text{ a linear isomorphism}$$

then

$$L^{-1}: Y\rightarrow X \text{ is a linear isomorphism}$$

Proof. Using [theorem: 2.72] we have that

$$L^{-1}: Y\rightarrow X \text{ is a bijection}$$

Let $x,y\in Y$ and $\alpha\in F$ then we have

$$\begin{array}{rcl} L^{-1}(x)+L^{-1}(y) & = & L^{-1}(L(L^{-1}(x)+L^{-1}(y))) \\ & = & L^{-1}(L(L^{-1}(x))+L(L^{-1}(y))) \\ & = & L^{-1}(x+y) \end{array}$$

Likewise

$$\begin{array}{rcl} \alpha\cdot L^{-1}(x) & = & L^{-1}(L(\alpha\cdot L^{-1}(x))) \\ & = & L^{-1}(\alpha\cdot L(L^{-1}(x))) \\ & = & L^{-1}(\alpha\cdot x) \end{array}$$

□

We look now at the composition of linear mappings.

Theorem 11.170. *Let $\langle X,+,\cdot\rangle$, $\langle Y,+,\cdot\rangle$ and $\langle Z,+,\cdot\rangle$ be vector spaces over the field $\langle F,+,\cdot\rangle$ and $L_1\in\mathrm{Hom}(X,Y)$, $L_2\in\mathrm{Hom}(Y,Z)$ then $L_2\circ L_1\in\mathrm{Hom}(X,Z)$*

Proof. Let $x,y\in X$ and $\alpha\in f$ then we have for the function

$$L_2\circ L_1: X\rightarrow Z$$

$$\begin{array}{rcl} (L_2\circ L_1)(x+y) & = & L_2(L_1(x+y)) \\ & = & L_2(L_1(x)+L_1(y)) \\ & = & L_2(L_1(x))+L_2(L_1(y)) \\ & = & (L_2\circ L_1)(x)+(L_2\circ L_1)(y) \\ (L_2\circ L_1)(\alpha\cdot x) & = & L_2(L_1(\alpha\cdot x)) \\ & = & L_2(\alpha\cdot L_1(x)) \\ & = & \alpha\cdot L_2(L_1(x)) \\ & = & \alpha\cdot(L_2\circ L_1)(x) \end{array}$$

□

Theorem 11.171. *Let $\langle X,+,\cdot\rangle$, $\langle Y,+,\cdot\rangle$ and $\langle Z,+,\cdot\rangle$ be vector spaces over the field $\langle F,+,\cdot\rangle$ then we have:*

1. *If $L\in\mathrm{Hom}(Y,Z)$ and $L_1,L_2\in\mathrm{Hom}(X,Y)$ then $L\circ(L_1+L_2)=L\circ L_1+L\circ L_2$*
2. *If $\alpha\in F$, $L_1\in\mathrm{Hom}(X,Y)$ and $L_2\in\mathrm{Hom}(Y,Z)$ then $L_2\circ(\alpha\cdot L_1)=\alpha\cdot(L_2\circ L_1)$*

Proof.

1. Let $x\in X$ then we have

$$\begin{array}{rcl} (L\circ(L_1+L_2))(x) & = & L((L_1+L_2)(x)) \\ & = & L(L_1(x)+L_2(x)) \\ & = & L(L_1(x))+L(L_2(x)) \\ & = & (L\circ L_1)(x)+(L\circ L_2)(x) \\ & = & (L\circ L_1+L\circ L_2)(x) \end{array}$$

proving that

$$L\circ(L_1+L_2)=L\circ L_1+L\circ L$$

2. Let $x \in X$ then we have:

$$\begin{aligned}(L_2 \circ (\alpha \cdot L_1))(x) &= L_2((\alpha \cdot L_1)(x)) \\ &= L_2(\alpha \cdot L_1(x)) \\ &= \alpha \cdot L_2(L_1(x)) \\ &= \alpha \cdot (L_2 \circ L_1)(x) \\ &= (\alpha \cdot (L_2 \circ L_1))(x)\end{aligned}$$

   proving that

$$L_2 \circ (\alpha \cdot L_1) = \alpha \cdot (L_2 \circ L_1)$$

□

It turns out that $\mathrm{Hom}(X, X)$ together with the composition operator $\circ$ forms a semi-group allowing us to define the composition of more then one linear transformations.

Corollary 11.172. *Let $\langle X, +, \cdot\rangle$, $\langle Y, +, \cdot\rangle$ and $\langle Z, +, \cdot\rangle$ be vector spaces over the field $\langle F, +, \cdot\rangle$ and $L_1: X \to Y$ and $L_2: Y \to Z$ linear isomorphism then $L_2 \circ L_1: X \to Z$ is a linear isomorphism.*

Proof. This follows from [theorems: 2.74 and 11.170] □

Theorem 11.173. *Let $\langle X, +, \cdot\rangle$ be a vector space over a field $\langle F, +, \cdot\rangle$ then*

$$\langle \mathrm{Hom}(X, X), \circ\rangle$$

is a semi-group with neutral element Id_X.

Proof. First using the previous theorem [theorem: 11.170] we have that the following mapping is well defined

$$\circ: \mathrm{Hom}(X, X) \times \mathrm{Hom}(X, X) \to \mathrm{Hom}(X, X) \text{ where } \circ(L, K) = L \circ K$$

Further we have for the operator $\circ$ that

associativity. $\forall L_1, L_2, L_3 \in \mathrm{Hom}(X, X)$ we have

$$L_1 \circ (L_2 \circ L_3) \underset{[\text{theorem: } 2.21]}{=} (L_1 \circ L_2) \circ L_3$$

neutral element. $\forall L \in \mathrm{Hom}(X, X)$ we have

$$L \circ \mathrm{Id}_X \underset{[\text{theorem: } 2.47]}{=} L \underset{[\text{theorem: } 2.47]}{=} \mathrm{Id}_X \circ L$$

□

Using $\circ$ as the product operator in the semi-group we can define the finite product (composition) of linear transformations.

Definition 11.174. *Let $n \in \mathbb{N}$, X a finite dimensional vector space with $\dim(X) = n$ and $\{L_i\}_{i \in \{1, \ldots, m\}} \subseteq \mathrm{Hom}(X, X)$ then*

$$L_1 \circ \cdots \circ L_n = \prod_{i=1}^{m} L_i$$

[see remark: 11.2]

In [theorem: 11.64] it is proved that $\langle Y^X, +, \cdot\rangle$ is a vector space where for $f, g \in Y^X$, $\alpha \in F$:

1. $f + g: X \to Y$ is defined by $(f + g)(x) = f(x) + g(x)$
2. $\alpha \cdot f: X \to Y$ is defined by $(\alpha \cdot f)(x) = \alpha \cdot f(x)$
3. $C_0: X \to Y$ is defined by $0(x) = 0$
4. $(-f): X \to Y$ is defined by $(-f)(x) = -f(x)$

As $\mathrm{Hom}(X, Y) \subseteq Y^X$, a natural question to ask is: Is $\mathrm{Hom}(X, Y)$ a sub-space of $\langle Y^X, +, \cdot\rangle$. It turns out that the answer is yes, as is shown in the next theorem.

Theorem 11.175. *Let $\langle X, +, \cdot\rangle$ and $\langle Y, +, \cdot\rangle$ be vector spaces over the field $\langle F, +, \cdot\rangle$ then*

$$\mathrm{Hom}(X, Y) \textit{ is a sub-space of } \langle Y^X, +, \cdot\rangle$$

Applying then [theorem: 11.57] proves that $\langle \mathrm{Hom}(X, Y), +, \cdot\rangle$ is vector space over the field $\langle F, +, \cdot\rangle$.

Proof. Let $L_1, L_2 \in \operatorname{Hom}(X, Y)$, $\alpha \in F$ then for $x, y \in X$ and $\gamma \in F$ we have

$$\begin{aligned}
(\alpha \cdot L_1 + L_2)(x + y) &= (\alpha \cdot L_1)(x + y) + L_2(x + y)\\
&= \alpha \cdot L_1(x + y) + L_2(x + y)\\
&= \alpha \cdot (L_1(x) + L_1(y)) + L_2(x) + L_2(y)\\
&= \alpha \cdot L_1(x) + \alpha \cdot L_1(y) + L_2(x) + L_2(y)\\
&= \alpha \cdot L_1(x) + L_2(x) + \alpha \cdot L_1(y) + L_2(y)\\
&= (\alpha \cdot L_1)(x) + L_2(y) + (\alpha \cdot L_1)(y) + L_2(y)\\
&= (\alpha \cdot L_1 + L_2)(x) + (\alpha \cdot L_1 + L_2)(y)\\
(\alpha \cdot L_1 + L_2)(\gamma \cdot x) &= (\alpha \cdot L_1)(\gamma \cdot x) + L_2(\gamma \cdot x)\\
&= \alpha \cdot L_1(\gamma \cdot x) + L_2(\gamma \cdot x)\\
&= \alpha \cdot (\gamma \cdot L_1(x)) + \gamma \cdot L_2(x)\\
&= \gamma \cdot (\alpha \cdot L_1(x)) + \gamma \cdot L_2(x)\\
&= \gamma \cdot ((\alpha \cdot L_1)(x) + L_2(x))\\
&= \gamma \cdot (\alpha \cdot L_1 + L_2)(x)\\
C_0(x + y) &= 0\\
&= 0 + 0\\
&= 0(x) + 0(y)\\
C_0(\gamma \cdot x) &= 0\\
&= \gamma \cdot 0\\
&= \gamma \cdot 0(x)
\end{aligned}$$

proving that

$$\alpha \cdot L_1 + L_2 \in \operatorname{Hom}(X, Y) \text{ and } C_0 \in \operatorname{Hom}(X, Y) \Rightarrow \operatorname{Home}(X, Y) \neq \varnothing$$

hence by [definition: 11.56] we have that $\operatorname{Hom}(X, Y)$ is a sub-space of $\langle Y^X, +, \cdot \rangle$ □

Definition 11.176. (Dual Space) *Let $\langle X, +, \cdot \rangle$ be a vector space over over a field $\langle F, +, \cdot \rangle$ [which is a vector space over itself] then $\operatorname{Hom}(X, F)$ is called the **dual space** of X and noted as X^*.*

Theorem 11.177. *Let $\langle X, +, \cdot \rangle$ be a vector space over over a field $\langle F, +, \cdot \rangle$ [which is a vector space over itself] and $x \in X$ with $x \neq 0$ then there exists a $L \in \operatorname{Hom}(X, F) = X^*$ such that $L(x) = 1$*

Proof. As $x \neq 0$ we have by [theorem: 11.117] that $\{x\}$ is a linear independent, using [theorem: 11.140] there exists a basis B of $\langle V, +, \cdot \rangle$ such that $\{x\} \subseteq B$. Let $y \in X$ then by [theorem: 11.133] there exists a **unique** $\{\alpha_w^{(y)}\}_{w \in B} \subseteq F$ such that $y = \sum_{w \in B} \alpha_w^{(y)} \cdot w$. This allows us to define

$$L: X \rightarrow F \text{ by } L(y) = \alpha_x^{(y)}$$

As $x \in B$ we have that for $\{\alpha_w\}_{w \in B}$ defined by $\alpha_w = \begin{cases} 1 \text{ if } w = x \\ 0 \text{ if } w \in B \setminus \{x\} \end{cases}$ that

$$\begin{aligned}
\sum_{w \in B} \alpha_w \cdot w &\underset{\text{[theorem: 11.43]}}{=} \sum_{w \in B \setminus \{x\}} \alpha_w \cdot w + \sum_{w \in \{x\}} \alpha_w \cdot w\\
&\underset{\text{[theorem: 11.34]}}{=} \sum_{w \in B \setminus \{x\}} \alpha_w \cdot w + \alpha_x \cdot x\\
&= \sum_{w \in B \setminus \{x\}} 0 \cdot w + 1 \cdot x\\
&\underset{\text{[theorem: 11.37]}}{=} 1 \cdot x\\
&= x\\
&= \sum_{w \in B} \alpha_w^{(x)} \cdot w
\end{aligned}$$

so that by uniqueness of the extension in a basis we have $\{\alpha_w\}_{w \in B} = \{\alpha_w^{(x)}\}_{w \in B}$ then $L(x) = \alpha_x^{(x)} = \alpha_x = 1$ or

$$L(x) = 1 \tag{11.70}$$

Further we have if $y_1, y_2 \in X$ then

$$\begin{aligned} y_1 + y_2 &= \sum_{w \in B} \alpha_w^{(y_1)} \cdot w + \sum_{w \in B} \alpha_w^{(y_2)} \cdot w \\ &\underset{[\text{theorem: } 11.38]}{=} \sum_{w \in B} \left(\alpha_w^{(y_1)} + \alpha_w^{(y_2)}\right) \cdot w \\ &= \sum_{w \in B} \alpha_w^{(y_1 + y_2)} \cdot w \end{aligned}$$

so that by uniqueness we have $L(y_1 + y_2) = \alpha_x^{(y_1 + y_2)} = \alpha_x^{(y_1)} + \alpha_x^{(y_2)} = L(y_1) + L(y_2)$ so that

$$L(y_1 + y_2) = L(y_1) + L(y_2) \tag{11.71}$$

Finally if $y \in X$ and $\alpha \in F$ we have

$$\begin{aligned} \alpha \cdot w &= \alpha \cdot \sum_{w \in B} \alpha_w^{(y)} \cdot w \\ &\underset{[\text{theorem: } 11.75]}{=} \sum_{w \in B} \alpha \cdot \left(\alpha_w^{(y)} \cdot w\right) \\ &= \sum_{w \in B} \left(\alpha \cdot \alpha_w^{(y)}\right) \cdot w \\ &= \sum_{w \in B} \alpha_w^{(\alpha \cdot y)} \cdot w \end{aligned}$$

so that by uniqueness we have $L(\alpha \cdot y) = \alpha_x^{(\alpha \cdot y)} = \alpha \cdot \alpha_x^{(y)} = \alpha \cdot L(y)$ proving

$$L(\alpha \cdot y) = \alpha \cdot L(y) \tag{11.72}$$

The theorem is then proved by [eqs: 11.70, 11.71, 11.72]. □

Theorem 11.178. *Let $\langle X, +, \cdot \rangle$ be a vector space over over a field $\langle F, +, \cdot \rangle$ [which is a vector space over itself]. Let $x, y \in X$ such that $\forall L \in \mathrm{Hom}(X, Y)$ we have $L(x) = L(y)$ then $x = y$.*

Proof. Let $x, y \in X$ such that $\forall L \in \mathrm{Hom}(X, Y)$ $L(x) = L(y)$. Assume that $x \neq y$ then $x - y \neq 0$ so by the previous theorem [theorem: 11.177] there exists a $L \in \mathrm{Hom}(X, Y)$ such that $L(x - y) = 1$, hence we have $0 \underset{L(x) = L(y)}{=} L(x) - L(y) = L(x - y) = 1$ leading to the contradiction $1 = 0$, hence we must have that $x = y$. □

Theorem 11.179. *Let $\langle X, +, \cdot \rangle$ be a vector space over a field $\langle F, +, \cdot \rangle$, $n \in \mathbb{N}$ and $\{e_i\}_{i \in \{1, \ldots, n\}} \subseteq X$ a linear independent family then there exist a $L \in \mathrm{Hom}(X, F)$ such that*

$$L(e_1) = 1 \text{ and } \forall i \in \{2, \ldots, n\}\ L(e_i) = 0$$

Proof. As $\{e_i\}_{i \in \{1, \ldots, n\}} \subseteq X$ is linear independent we have by [theorem: 11.129] that $E = \{e_1, \ldots, e_n\}$ is a linear independent set. Using [theorem: 11.140] then there exist a basis B of X such that $E \subseteq B$. Let $y \in X$ then there exist by [theorem: 11.134] a finite $I \subseteq B$ and a $\{\alpha_i\}_{i \in I} \subseteq F$ such that

$$y = \sum_{i \in I} \alpha_i \cdot i$$

Let

$$L_y^{\{a_i\}_{i \in I}} = \begin{cases} 0 \text{ if } e_1 \notin I \\ \alpha_{e_1} \text{ if } e_1 \in I \end{cases}$$

then if there is another $J \subseteq B$ and $\{\beta_i\}_{i \in J} \subseteq F$ such that

$$y = \sum_{i \in J} \beta_i \cdot i$$

and let

$$L_y^{\{\beta_i\}_{i \in \mathrm{IJ}}} = \begin{cases} 0 \text{ if } e_1 \notin J \\ \beta_{e_1} \text{ if } e_1 \in J \end{cases}$$

then by [theorem: 11.134] we have that

$$\begin{array}{ll} \forall u \in I \setminus J & \alpha_u = 0 \\ \forall u \in J \setminus I & \beta_u = 0 \\ \forall u \in I \bigcap J & \alpha_u = \beta_u \end{array}$$

For e_1 we have then either

$e_1 \in I \wedge e_1 \in J$. Then $e_1 \in I \bigcap J$ so that $L_y^{\{\beta_i\}_{i \in J}} = \beta_{e_1} = \alpha_{e_1} = L_y^{\{\alpha_i\}_{i \in I}}$

$e_1 \notin I \wedge e_1 \in J$. Then $L_y^{\{\beta_i\}_{i \in J}} = \beta_{e_1} = 0 = L_y^{\{\alpha_i\}_{i \in I}}$

$e_1 \in I \wedge e_1 \notin J$. Then $L_y^{\{\beta_i\}_{i \in J}} = 0 = \alpha_{e_1} = L_y^{\{\beta_i\}_{i \in J}}$

$e_1 \notin I \wedge e_1 \notin J$. Then $L_y^{\{\beta_i\}_{i \in J}} = 0 = L_y^{\{\beta_i\}_{i \in J}}$

proving that in all cases $L_y^{\{a_i\}_{i \in I}} = L_y^{\{\beta_i\}_{i \in J}}$, hence the following function is well defined:

$$L: X \rightarrow F \text{ by } L(y) = L_y^{\{a_i\}_{i \in I}}$$

Let now $\alpha \in F$, $x, y \in X$ then there exists finite $I, J \subseteq B$ and $\{\alpha_i\}_{i \in I} \subseteq F$, $\{\beta_i\}_{i \in J} \subseteq F$ such that

$$x = \sum_{i \in I} \alpha_i \cdot i \text{ and } y = \sum_{i \in I} \beta_i \cdot i$$

Define

$$\{\gamma_i\}_{i \in I \bigcup J} \subseteq F \text{ by } \gamma_i = \begin{cases} \alpha_i \text{ if } i \in I \setminus J \\ \alpha \cdot \beta_i \text{ if } i \in J \setminus I \\ \alpha_i + \alpha \cdot \beta_i \text{ if } i \in I \bigcap J \end{cases}$$

then we have:

$$\begin{array}{rcl} \sum_{i \in I \bigcup J} \gamma_i \cdot i & \underset{[\text{theorem: } 11.44]}{=} & \sum_{i \in I \bigcap J} \gamma_i \cdot i + \sum_{i \in I \setminus J} \gamma_i \cdot i + \sum_{i \in J \setminus I} \gamma_i \cdot i \\ & = & \sum_{i \in I \bigcap J} (\alpha_i + \alpha \cdot \beta_i) \cdot i + \sum_{i \in I \setminus J} \alpha_i \cdot i + \sum_{i \in J \setminus I} \alpha \cdot \beta_i \cdot i \\ & \underset{[\text{theorem: } 11.38, 11.75]]}{=} & \underbrace{\sum_{i \in I \bigcap J} \alpha_i \cdot i}_{1} + \underbrace{\alpha \cdot \sum_{i \in I \bigcap J} \beta_i \cdot i}_{2} + \underbrace{\sum_{i \in I \setminus J} \alpha_i \cdot i}_{1} + \underbrace{\alpha \cdot \sum_{i \in J \setminus I} \beta_i \cdot i}_{2} \\ & \underset{[\text{theorem: } 11.44]}{=} & \underbrace{\sum_{i \in I} \alpha_i \cdot i}_{1} + \underbrace{\alpha \cdot \sum_{i \in J} \beta_i \cdot i}_{2} \\ & = & x + \alpha \cdot y \end{array}$$

from the above we have that

$$\begin{array}{rcl} L(x + \alpha \cdot y) & = & L_{x + \alpha \cdot y}^{\{\gamma_i\}_{i \in I \bigcup J}} \\ & = & \begin{cases} 0 \text{ if } e_1 \notin I \bigcup J \\ \gamma_{e_1} \text{ if } e_2 \in I \bigcup J \end{cases} \\ & = & \begin{cases} 0 \text{ if } e_1 \notin I \bigcup J \\ \alpha_{e_1} + \alpha \cdot \beta_{e_1} \text{ if } e_2 \in I \bigcup J \end{cases} \\ & \underset{e_1 \notin I \bigcup J \Leftrightarrow e_1 \notin I \wedge e_1 \notin J}{=} & \begin{cases} 0 \text{ if } e_1 \notin I \\ \alpha_{e_1} \text{ if } e_2 \in I \bigcup J \end{cases} + \alpha \cdot \begin{cases} 0 \text{ if } e_1 \notin J \\ \beta_{e_1} \text{ if } e_2 \in I \bigcup J \end{cases} \\ & = & L_x^{\{\alpha_i\}_{i \in I}} + \alpha \cdot L_y^{\{\beta_i\}_{i \in J}} \\ & = & L(x) + \alpha \cdot L(y) \end{array}$$

proving by [theorem: 11.163] that

$$L \in \operatorname{Hom}(X, F)$$

Define now $\{\zeta_i\}_{i \in \{e_1\}} \subseteq B$ by $\zeta_{e_1} = 1$ then $\sum_{i \in \{e_1\}} \zeta_i \cdot i \underset{[\text{theorem: } 11.34]}{=} e_1$ so that

$$L(e_1) = \begin{cases} 0 \text{ if } e_1 \notin \{e_1\} \\ 1 \text{ if } e_1 \in \{e_1\} \end{cases} = 1$$

If $i\in\{2,\ldots,n\}$ then as $\{e_i|i\in\{1,\ldots,n\}=\{e_1,\ldots,e_n\}$ $e_1\neq e_i\Rightarrow e_1\notin\{e_i\}$, define $\{\eta_k\}_{k\in\{e_i\}}\subseteq F$ by $\eta_k=1$ then $\sum_{k\in\{e_i\}}\eta_k\cdot k\underset{[\text{theorem: } 11.34]}{=}e_i$ so that

$$L(e_i)=\left\{\begin{array}{l}0 \text{ if } e_1\notin\{e_i\}\\ 1 \text{ if } e_1\in\{e_i\}\end{array}\right.=0$$

Hence we have

$$L(e_1)=1 \text{ and } \forall i\in\{2,\ldots,n\}\ L(e_i)=0$$

□

Theorem 11.180. *Let $\langle X,+,\cdot\rangle$ and $\langle Y,+,\cdot\rangle$ be vector spaces over a field $\langle F,+,\cdot\rangle$ and $L\in\mathrm{Hom}(X,Y)$ a linear mapping between X and Y then:*

1. *If $n\in\mathbb{N}_0$ and $\{\alpha_i\}_{i\in\{0,\ldots,n\}}\subseteq F$ and $\{x_i\}_{i\in\{0,\ldots,n\}}\subseteq X$ then*

$$L\left(\sum_{i=0}^{n}\alpha_i\cdot x\right)=\sum_{i=0}^{n}\alpha_i\cdot L(x_i)$$

2. *If $n,m\in\mathbb{N}_0$ with $n\leqslant m$ and $\{\alpha_i\}_{i\in\{n,\ldots,m\}}\subseteq F$ and $\{x_i\}_{i\in\{n,\ldots,m\}}\subseteq X$ then*

$$L\left(\sum_{i=n}^{m}\alpha_i\cdot x\right)=\sum_{i=n}^{m}\alpha_i\cdot L(x_i)$$

3. *If $W\subseteq X$ is a finite set and $\{\alpha_w\}_{w\in W}\subseteq F$ then*

$$L\left(\sum_{w\in W}\alpha_w\cdot w\right)=\sum_{w\in W}\alpha_w\cdot L(w)$$

Proof.

1. This is proved by induction, let

$$S=\left\{n\in\mathbb{N}_0|\text{If } \{\alpha_i\}_{i\in\{0,\ldots,n\}}\subseteq F \text{ and } \{x_i\}_{i\in\{0,\ldots,n\}}\subseteq X \text{ are families then } L\left(\sum_{i=0}^{n}\alpha_i\cdot x\right)=\sum_{i=0}^{n}\alpha_i\cdot L(x_i)\right\}$$

then we have:

$0\in S$. If $\{\alpha_i\}_{i\in\{0,\ldots,0\}}\subseteq F$ and $\{x_i\}_{i\in\{0,\ldots,0\}}\subseteq X$ are families then we have

$$\begin{aligned}L\left(\sum_{i=0}^{0}a_i\cdot x_i\right)&\underset{\text{def}}{=}L(a_0\cdot x_o)\\&=\alpha_0\cdot L(x_0)\\&\underset{\text{def}}{=}\sum_{i=0}^{0}\alpha_i\cdot L(x_i)\end{aligned}$$

proving that $0\in S$.

$n\in S\Rightarrow n+1\in S$. Let $\{\alpha_i\}_{i\in\{0,\ldots,n+1\}}\subseteq F$ and $\{x_i\}_{i\in\{0,\ldots,n+1\}}\subseteq X$ be families then we have

$$\begin{aligned}L\left(\sum_{i=0}^{n+1}\alpha_i\cdot x_i\right)&\underset{\text{def}}{=}L\left(\sum_{i=0}^{n}\alpha_i\cdot x_i+\alpha_{n+1}\cdot x_{n+1}\right)\\&=L\left(\sum_{i=0}^{n}\alpha_i\cdot x_i\right)+\alpha_{n+1}\cdot L(x_{n+1})\\&\underset{n\in S}{=}\sum_{i=0}^{n}\alpha_i\cdot L(x_i)+\alpha_{n+1}\cdot L(x_{n+1})\\&\underset{\text{def}}{=}\sum_{i=0}^{n+1}\alpha_i\cdot L(x_i)\end{aligned}$$

proving that $n+1\in S$

2. Let $n,m\in\mathbb{N}_0$ with $n\leqslant m$ and $\{\alpha_i\}_{i\in\{n,\ldots,m\}}\subseteq F$ and $\{x_i\}_{i\in\{n,\ldots,m\}}\subseteq X$ then we have

$$L\left(\sum_{i=n}^{m}\alpha_i\cdot x_i\right)\underset{\text{def}}{=}L\left(\sum_{i=0}^{m-n}\alpha_{i+n}\cdot x_{i+n}\right)\underset{(1)}{=}\sum_{i=0}^{m-n}\alpha_{i+n}\cdot L(x_{i+n})\underset{\text{def}}{=}\sum_{i=n}^{m}\alpha_i\cdot L(x_i)$$

3. If W is finite we have either:

   ***W* = ∅.** Then $\sum_{w\in W}\alpha_w\cdot w=0$ so that

$$L\left(\sum_{w\in W}\alpha_w\cdot w\right)=L(0)\underset{[\text{theorem: }11.162]}{=}0=\sum_{w\in W}\alpha_w\cdot L(w)$$

   ***W* ≠ ∅.** Then there exist by definition a $n\in\mathbb{N}$ and a $\beta:\{0,\ldots,n-1\}\to W$ such that

$$\sum_{w\in W}\alpha_w\cdot w=\sum_{i=0}^{n-1}\alpha_{w(i)}\cdot w(i)$$

   so that

$$L\left(\sum_{w\in W}\alpha_w\cdot w\right)=L\left(\sum_{i=0}^{n-1}\alpha_{w(i)}\cdot w(i)\right)\underset{(1)}{=}\sum_{i=0}^{n-1}\alpha_{w(i)}\cdot L(w(i))\underset{[\text{theorem: }11.36]}{=}\sum_{w\in W}\alpha_w\cdot L(w)$$

□

11.4.2 Kernel and image of a linear mapping

Definition 11.181. (Kernel of a Linear Mapping) *Let $\langle X,+,\cdot\rangle$ and $\langle Y,+,\cdot\rangle$ be vector spaces over a field $\langle F,+,\cdot\rangle$ and $L\in\mathrm{Hom}(X,Y)$ then the **kernel** of L noted as $\ker(L)$ is defined by*

$$\ker(L)=\{x\in X|L(x)=0\}=L^{-1}(\{0\})$$

Definition 11.182. (Range of a Linear Mapping) *Let $\langle X,+,\cdot\rangle$ and $\langle Y,+,\cdot\rangle$ be vector spaces over a field $\langle F,+,\cdot\rangle$ and $L\in\mathrm{Hom}(X,Y)$ then the range of L noted as $\mathbf{range}(L)$ is defined by*

$$\mathbf{range}(L)=L(X)$$

Theorem 11.183. *Let $\langle X,+,\cdot\rangle$ and $\langle Y,+,\cdot\rangle$ are vector spaces over a field $\langle F,+,\cdot\rangle$ and $L\in\mathrm{Hom}(X,Y)$ then*

1. *$\ker(L)$ is a sub-space of X*
2. *$\mathrm{range}(L)$ us a sub-space of Y*

Proof.

1. First as $L(0)\underset{[\text{theorem: }11.162]}{=}0$ we have that

$$0\in\ker(L)\Rightarrow 0\neq\ker(L)$$

   Second if $x,y\in\ker(L)$ and $\alpha\in F$ then we have

$$L(x+y)=L(x)+L(y)\underset{[\text{theorem: }11.162]}{=}0+0=0$$

   and

$$L(\alpha\cdot x)=\alpha\cdot L(x)\underset{[\text{theorem: }11.162]}{=}\alpha\cdot 0=0$$

   so that

$$x+y\in\ker(L)\text{ and }\alpha\cdot x\in\ker(L)$$

So we have that $\ker(L)$ is a sub-space of X.

2. If $x, y \in \text{range}(L) = L(X)$ and $\alpha \in F$ then we have a $x', y' \in X$ such that $x = L(x')$ and $y = L(y')$. For $x' + y' \in X$ we have then that

$$L(x' + y') = L(x') + L(y') = x + y$$

so that

$$x + y \in L(X) = \text{range}(L)$$

Likewise for $\alpha \cdot x' \in X$ we have

$$L(\alpha \cdot x') = \alpha \cdot L(x') = \alpha \cdot x$$

proving that

$$\alpha \cdot x \in L(X) = \text{range}(L)$$ □

As $\text{range}(L)$ is a vector space the following definition make sense.

Definition 11.184. (Rank of a Linear Mapping) *Let $\langle X, +, \cdot \rangle$ and $\langle Y, +, \cdot \rangle$ be vector spaces over a field $\langle F, +, \cdot \rangle$ and $L \in \text{Hom}(X, Y)$ then the r**ank of L** noted as $\text{rank}(L)$ is defined by*

$$\text{rank}(L) = \dim(\text{range}(L)) = \dim(L(X))$$

Theorem 11.185. *Let $\langle X, +, \cdot \rangle$ and $\langle Y, +, \cdot \rangle$ are vector spaces over a field $\langle F, +, \cdot \rangle$ and $L \in \text{Hom}(X, Y)$ then*

$$L \textit{ is injective} \Leftrightarrow \ker(L) = \{0\}$$

Proof.

⇒. As $L(0) \underset{\text{[theorem: 11.162]}}{=} 0$ we have $0 \in \ker(L)$ so that $\{0\} \subseteq \ker(L)$. Let $x \in \ker(L)$ then $L(x) = 0 \underset{\text{[theorem: 11.162]}}{=} L(0)$, as L is injective we have $x = 0$ proving that $\ker(L) \subseteq \{0\}$. Hence we have

$$\ker(L) = \{0\}$$

⇐. Let $x, y \in X$ such that $L(x) = L(y)$ then $L(x - y) = L(x) - L(y) = 0$ so that $x - y \in \ker(L) = \{0\}$. Hence $x - y = 0$ or $x = y$ proving that L is injective. □

Theorem 11.186. *Let $\langle X, +, \cdot \rangle$ and $\langle Y, +, \cdot \rangle$ be vector spaces over a field $\langle F, +, \cdot \rangle$, $W \subseteq X$ and $L \in \text{Hom}(X, Y)$ then*

$$L(\text{span}(W)) \subseteq \text{span}(L(W))$$

Further if L is injective [or equivalently $\ker(L) = \{0\}$] then

$$L(\text{span}(W)) = \text{span}(L(W))$$

Proof. For the first part. If $y \in L(\text{span}(W))$ then $\exists x \in \text{span}(W)$ such that $y = L(x)$. As $x \in \text{span}(W)$ there exists by [theorem: 11.96] a finite $\{v_i\}_{i \in \{1, \ldots, n\}} \subseteq W$ and $\{\alpha_i\}_{i \in 1, \ldots, n} \subseteq F$ such that

$$x = \sum_{i \in \{1, \ldots, n\}} \alpha_i \cdot v_i$$

Define

$$\{u_i\}_{i \in \{1, \ldots, n\}} \subseteq L(W) \text{ by } u_i = L(v_i)$$

then we have

$$\begin{aligned} L(x) &= L\left(\sum_{i \in \{1, \ldots, n\}} \alpha_i \cdot w_i\right) \\ &\underset{\text{[theorem: 11.180]}}{=} \sum_{i \in \{1, \ldots, n\}} \alpha_i \cdot L(w_i) \end{aligned}$$

By [theorem: 11.96] it follows that $y = L(x) \in \text{span}(W)$ proving that

$$L(\text{span}(W)) \subseteq \text{span}(L(W)) \tag{11.73}$$

For the second part. Let $y \in \text{span}(L(W))$ then there exists a finite $J \subseteq L(W)$ and a $\{\alpha_i\}_{i \in J} \subseteq F$ such that

$$y = \sum_{u \in J} \alpha_u \cdot u \tag{11.74}$$

Let $I = L^{-1}(J)$ then if $i \in I$ we have that $L(i) \in J$, as $J \subseteq L(W)$ there exist a $j \in W$ such that $L(i) = L(j)$ which, as L is injective, proves that $i = j \in W$. Hence

$$I \subseteq W \tag{11.75}$$

Define

$$\beta: I \to J \text{ by } \beta(i) = L(i)$$

then we have

injectivity. If $i, j \in I$ such that $\beta(i) = \beta(j)$ then $L(i) = L(j)$ which, as L is injective, proves that $i = j$.

surjectivity. If $j \in J$ then as $J \subseteq L(W)$ we have that $\exists i \in W$ such that $j = L(i)$. As $L(i) = j \in J$ we have that $i \in L^{-1}(J) = I$, so there exists a $i \in L^{-1}(J)$ such that $j = L(i) = \beta(i)$.

Hence we have

$$\beta: I \to J \text{ defined by } \beta(i) = L(i) \text{ is a bijection} \tag{11.76}$$

As J is finite we have also that

$$I \text{ is finite} \tag{11.77}$$

Define

$$\{\gamma_i\}_{i \in I} \subseteq F \text{ by } \gamma_i = \alpha_{\beta(i)} \tag{11.78}$$

By [eqs: 11.75, 11.77] and the definition of a span we have that

$$\sum_{i \in I} \gamma_i \cdot i \in \text{span}(W)$$

Further

$$\begin{aligned} L\left(\sum_{i \in I} \gamma_i \cdot i\right) &\underset{[\text{theorem: } 11.180]}{=} \sum_{i \in I} \gamma_i \cdot L(i) \\ &= \sum_{i \in I} \alpha_{\beta(i)} \cdot \beta(i) \\ &\underset{[\text{theorem: } 11.36]}{=} \sum_{i \in J} \alpha_i \cdot i \\ &= y \end{aligned}$$

proving that $y \in L(\text{span}(W))$. Hence $\text{span}(L(W)) \subseteq L(\text{span}(W))$. Combining this with [eq: 11.73 proves that

$$L(\text{span}(W)) = \text{span}(L(W))$$ □

Theorem 11.187. *Let $\langle X, +, \cdot \rangle$, $\langle Y, +, \cdot \rangle$ be vector spaces over a field $\langle F, +, \cdot \rangle$, $W \subseteq X$ and $L \in \text{Hom}(X, Y)$ a injective linear mapping. If W is linear independent then $L(W)$ is linear independent.*

Proof. Let $J \subseteq L(W)$ be a finite subset and $\{\alpha_i\}_{i \in J} \subseteq F$ such that

$$\sum_{i \in J} \alpha_i \cdot i = 0 \tag{11.79}$$

Let $I = L^{-1}(J)$ then if $i \in I$ we have that $L(i) \in J$, as $J \subseteq L(W)$ there exist a $j \in W$ such that $L(i) = L(j)$ which, as L is injective, proves that $i = j \in W$. Hence

$$I \subseteq W \tag{11.80}$$

Define

$$\beta: I \to J \text{ by } \beta(i) = L(i)$$

then we have

injectivity. If $i, j \in I$ such that $\beta(i) = \beta(j)$ then $L(i) = L(j)$ which, as L is injective, proves that $i = j$.

surjectivity. If $j \in J$ then as $J \subseteq L(W)$ we have that $\exists i \in W$ such that $j = L(i)$. As $L(i) = j \in J$ we have that $i \in L^{-1}(J) = I$, so there exists a $i \in L^{-1}(J)$ such that $j = L(i) = \beta(i)$.

$$\beta: I \to J \text{ defined by } \beta(i) = L(i) \text{ is a bijection} \tag{11.81}$$

Hence as J is finite we have also that

$$I \text{ is finite} \tag{11.82}$$

Define now

$$\{\gamma_i\}_{i \in I} \subseteq F \text{ by } \gamma_i = \alpha_{\beta(i)}$$

Then we have

$$\begin{aligned} L\left(\sum_{i \in I} \gamma_i \cdot i\right) & \underset{[\text{theorem: } 11.180]}{=} \sum_{i \in I} \gamma_i \cdot L(i) \\ & = \sum_{i \in I} \alpha_{\beta(i)} \cdot \beta(i) \\ & \underset{[\text{theorem: } 11.36] \text{ and } [\text{eq: } 11.81]}{=} \sum_{i \in J} \alpha_i \cdot i \\ & \underset{[\text{theorem: } 11.79]}{=} 0 \end{aligned}$$

So $\sum_{i \in I} \gamma_i \cdot i \in \ker(L)$, as L is injective we have by [theorem: 11.185] that $\ker(L) = \{0\}$, so that

$$\sum_{i \in I} \gamma_i \cdot i = 0$$

As I is finite, $I \subseteq W$ and W is linear independent we have by [theorem: 11.114] that $\forall i \in I \;\; \gamma_i = 0$. Hence, if $i \in J$ then $\beta^{-1}(i) \in I$ so that $\alpha_i = \alpha_{\beta(\beta^{-1}(i))} = \gamma_{\beta^{-1}(i)} = 0$, proving that $\forall i \in J$ we have $\alpha_i = 0$. Applying then [theorem: 11.114] proves that $L(W)$ is linear independent. □

Corollary 11.188. *Let $\langle X, +, \cdot \rangle$ and $\langle Y, +, \cdot \rangle$ be vector spaces over a field $\langle F, +, \cdot \rangle$, $B \subseteq X$ a basis of $\langle X, +, \cdot \rangle$ and $L \in \mathrm{Hom}(X, Y)$ a injective linear mapping then $L(B)$ is a basis of* $\mathrm{range}(L)$.

Proof. As $B \subseteq X$ we have

$$L(B) \subseteq L(X) = \mathrm{range}(X)$$

As B is a basis of X we have that B is linear independent and $\mathrm{span}(B) = X$. As B is linear independent we have by [theorem: 11.187] that

$$L(B) \text{ is linear independent}$$

As B is a basis we have $\mathrm{span}(B) = X$ so that

$$\begin{aligned} \mathrm{range}(L) & = L(X) \\ & = L(\mathrm{span}(B)) \\ & \underset{[\text{theorem: } 11.186]}{=} \mathrm{span}(L(B)) \end{aligned}$$

So $L(B)$ is a basis of $\mathrm{range}(L)$. □

Corollary 11.189. *Let $\langle X, +, \cdot \rangle$ and $\langle Y, +, \cdot \rangle$ be finite dimensional vector spaces over a field $\langle F, +, \cdot \rangle$ such that $\dim(X) = \dim(Y)$ and $L \in \mathrm{Hom}(X, Y)$ a injective linear mapping then $L: X \to Y$ is a linear isomorphism*

Proof. As X is finite dimensional there exists a basis $B \subseteq X$ with $\mathrm{card}(B) = \dim(X)$. Further as L is injective we have that $L_{|B}: B \to L(B)$ is a bijection so that $\mathrm{card}(L(B)) = \mathrm{card}(B) = \dim(X)$. By [theorem: 11.187] we have that $L(B) \subseteq Y$ is linear independent. Using this, the fact that $\mathrm{card}(L(B)) = \dim(X) = \dim(Y)$ it follows from [theorem: 11.149] that $L(B)$ is a basis of Y. Hence

$$L(X) = L(\mathrm{span}(B)) \underset{[\text{theorem: } 11.186]}{=} \mathrm{span}(L(B)) \underset{L(B) \text{ is a basis of Y}}{=} Y$$

proving that L is surjective, hence as L is also injective it follows that $L: X \rightarrow Y$ is a bijection and by definition that

$$L \text{ is a linear isomorphism}$$

□

Corollary 11.190. *Let $\langle X, +, \cdot\rangle$ be a finite dimensional vector space over a filed $\langle F, +, \cdot\rangle$ and $L \in \operatorname{Hom}(X, X)$ a injective linear transformation then $L: X \rightarrow X$ is a linear isomorphism.*

Proof. As clearly $\dim(X) = \dim(X)$ this follows from the previous theorem. □

Corollary 11.191. *Let $\langle X, +, \cdot\rangle$ and $\langle Y, +, \cdot\rangle$ be vector spaces over a field $\langle F, +, \cdot\rangle$ and $L \in \operatorname{Hom}(X, Y)$ a linear isomorphism then we have:*

1. *If X is infinite dimensional then Y is infinite dimensional*
2. *If X is finite dimensional then Y is finite dimensional*
3. $\dim(X) = \dim(Y)$

Proof. If L is a linear isomorphism then

$$L: X \rightarrow Y \text{ is a bijection so that } \operatorname{range}(L) = L(X) = Y \tag{11.83}$$

Let B be a basis for X then

$$L_{|B}: B \rightarrow L(B) \text{ is a bijection or } B \approx L(B) \tag{11.84}$$

Further by the previous theorem [theorem: 11.188]

$$L(B) \text{ is a basis for } \operatorname{range}(L) \underset{\text{[eq: 11.83]}}{=} Y \tag{11.85}$$

So we have

1. If X is infinite dimensional then B is infinite so that by [theorem: 6.46] and [eq: 11.84] it follows that $L(B)$ is infinite. Hence Y is infinite dimensional.
2. If X is finite dimensional then B is finite so that by [theorem: 6.43] and [eq: 11.84] it follows that $L(B)$ is finite. Hence Y is finite dimensional.
3. For B we have either:

   ***B* is infinite.** Then $L(B)$ is infinite and $\dim(X) = \infty = \dim(Y)$

   ***B* is finite.** Then there exists a $n \in \mathbb{N}_0$ such that $\{1, \ldots, n\} \approx B \approx L(B)$ so that

   $$\dim(X) = n = \dim(Y)$$

□

The above theorem is a special case of a more general theorem.

Theorem 11.192. *Let $\langle X, +, \cdot\rangle$ be a finite dimensional vector space over the field $\langle F, +, \cdot\rangle$, $\langle Y, +, \cdot\rangle$ a vector space over $\langle F, +, \cdot\rangle$ and $L \in \operatorname{Hom}(X, Y)$ then*

$$\dim(X) = \dim(\ker(L)) + \operatorname{rank}(L)$$

Proof. We divide the proof in two cases:

$\ker(L) = \{0\}$. Then by [theorem: 11.185] L is injective so that

$$L: X \rightarrow L(X) = \operatorname{range}(L)$$

is a isomorphism, hence by [theorem: 11.191] $\langle \operatorname{range}(L), +, \cdot\rangle$ is finite dimensional and $\dim(X) = \dim(\operatorname{range}(L)) = \operatorname{rank}(L)$. Further $\dim(\ker(L)) = \dim(\{0\}) \underset{\text{]example: 11.147]}}{=} 0$ so that

$$\dim(X) = \operatorname{rank}(L) = 0 + \operatorname{rank}(L) = \dim(\ker(L)) + \operatorname{rank}(L)$$

$\ker(L) \neq \{0\}$. Let B_k be a basis [hence a linear independent set] of $\langle \ker(L), +, \cdot\rangle$ and use [theorem: 11.140] to find a basis B_x of $\langle X, +, \cdot\rangle$ such that $B_k \subseteq B_x$. As B_x is finite [for X is finite dimensional] we have that B_k is finite. Further, as B_x is a disjoint union of $B_x \setminus B_k$ and B_k, we have by [theorem: 10.90] that

$$\dim(X) = \operatorname{card}(B_x) = \operatorname{card}(B_x \setminus B_k) + \operatorname{card}(B_k) = \operatorname{card}(B_x \setminus B_k) + \dim(\ker(L)) \tag{11.86}$$

Consider the function

$$L_{|B_x \setminus B_k}: B_x \setminus B_k \rightarrow L(B_x \setminus B)$$

then we have

injectivity. Let $x, y \in B_x \setminus B_k$ such that $L_{|B_x \setminus B_k}(x) = L_{|B_x \setminus B_k}(y)$ then $L(x) = L(y)$ so that $L(x - y) = L(x) - L(y) = 0$, hence $x - y \in \ker(L)$. As B_k is a basis for $\ker(L)$ and B_k is finite we have by [theorem: 11.136] that there exists a $\{\alpha_u\}_{u \in B_k}$ such that

$$x - y = \sum_{u \in B_k} \alpha_u \cdot u$$

so that

$$0 = \sum_{u \in B_k} \alpha_u \cdot u + 1 \cdot y + (-1) \cdot x \tag{11.87}$$

Now

$$\begin{array}{rcl} B_x & \underset{B_k \subseteq B_x}{=} & (B_x \setminus B_k) \bigcup B_k \\ & = & ((B_x \setminus B_k) \setminus \{x, y\}) \bigcup \{x, y\} \bigcup B_k \\ & = & ((B_x \setminus B_k) \setminus \{x, y\}) \bigcup \{x\} \bigcup \{y\} \bigcup B_k \end{array}$$

Assume that $x \neq y$ then $\{x\} \bigcap \{y\} = \varnothing$, further $((B_x \setminus B_k) \setminus \{x, y\}) \bigcap \{x\} = \varnothing$, $((B_x \setminus B_k) \setminus \{x, y\}) \bigcap \{y\}$, $((B_x \setminus B_k) \setminus \{x, y\}) \bigcap B_k = \varnothing$, $\{x\} \bigcap B_k \underset{x, y \in B_x \setminus B_k}{=} \varnothing$ and $\{x\} \bigcap B_k \underset{x, y \in B_x \setminus B_k}{=} \varnothing$ so that

$$B_x \text{ is the disjoint union of } ((B_x \setminus B_k) \setminus \{x, y\}), \{x\}, \{y\} \text{ and } B_k \tag{11.88}$$

The above allows us to define

$$\{\gamma_u\}_{u \in B_x} \subseteq F \text{ by } \gamma_u = \begin{cases} \alpha_u \text{ if } u \in B_k \\ 1 \text{ if } u \in \{x\} \\ -1 \text{ if } u \in \{y\} \\ 0 \text{ if } u \in ((B_x \setminus B_k) \setminus \{x, y\}) \end{cases}$$

then we have

$$\begin{array}{rc} \sum_{u \in B_x} \gamma_u \cdot u & \underset{[\text{theorem: } 11.43]}{=} \\ \sum_{u \in B_k} \gamma_u \cdot u + \sum_{u \in \{x\}} \gamma_u \cdot u + \sum_{u \in \{y\}} \gamma_u \cdot u + \sum_{u \in ((B_x \setminus B_k) \setminus \{x, y\})} \gamma_u \cdot u & \underset{[\text{theorem: } 11.34]}{=} \\ \sum_{u \in B_k} \gamma_u \cdot u + \gamma_x \cdot x + \gamma_y \cdot y + \sum_{u \in ((B_x \setminus B_k) \setminus \{x, y\})} \gamma_u \cdot u & = \\ \sum_{u \in B_k} \alpha_u \cdot u + 1 \cdot x + (-1) \cdot y + \sum_{u \in ((B_x \setminus B_k) \setminus \{x, y\})} 0 \cdot u & \underset{[\text{theorem: } 11.37]}{=} \\ \sum_{u \in B_k} \alpha_u \cdot u + 1 \cdot x + (-1) \cdot y & \underset{[\text{eq: } 11.87]}{=} \\ 0 & \end{array}$$

As B_x is linear independent we have that $\forall u \in B_k \; \gamma_u = 0$ contradicting the fact that $\gamma_x = 1$. Hence we must have that $x = y$ proving injectivity.

surjectivity. This is trivial.

So we have proved that

$$L_{|B_x \setminus B_k}: B_x \setminus B_k \rightarrow L(B_x \setminus B_k) \text{ is a isomorphism} \tag{11.89}$$

Hence , as $L(B_x \setminus B_k) \approx B_x \setminus B_k \subseteq B_x$ a finite set, we have that

$$L(B_x \setminus B_k) \text{ is finite} \tag{11.90}$$

If $y \in L(X)$ then there exists a $x \in X$ such that $y = L(x)$. As B_x is a basis there exists a $\{\alpha_u\}_{u \in B_x} \subseteq F$ such that $x = \sum_{u \in B_x} \alpha_u \cdot u$. So that

$$\begin{aligned} y &= L(x) \\ &= L\left(\sum_{u \in B_x} \alpha_u \cdot u\right) \\ &\underset{[\text{theorem: } 11.180]}{=} \sum_{u \in B_x} \alpha_u \cdot L(u) \\ &\underset{[\text{theorem: } 11.43]}{=} \sum_{u \in B_x \setminus B_k} \alpha_u \cdot L(u) + \sum_{u \in B_k} \alpha_u \cdot L(u) \\ &= \sum_{u \in B_x \setminus B_k} \alpha_u \cdot L(u) + \sum_{u \in B_k} \alpha_u \cdot 0 \\ &\underset{[\text{theorem: } 11.37]}{=} \sum_{u \in B_x \setminus B_k} \alpha_u \cdot L(u) \end{aligned} \tag{11.91}$$

Define now

$$\{\beta_u\}_{u \in L(B_x \setminus B_k)} \text{ by } \beta_u = \alpha_{(L_{|B_x \setminus B_k})^{-1}(u)}$$

Then we have

$$\begin{aligned} \sum_{u \in L(B_x \setminus B_k)} \beta_u \cdot u &= \sum_{u \in L(B_x \setminus B_k)} \alpha_{(L_{|B_x \setminus B_k})^{-1}(u)} \cdot (L_{|B_x \setminus B_k}((L_{|B_x \setminus B_k})^{-1}(u))) \\ &\underset{[\text{theorem: } 11.36]}{=} \sum_{u \in B_x \setminus B_k} \alpha_u \cdot L_{|B_x \setminus B_k}(u) \\ &= \sum_{u \in B_x \setminus B_k} \alpha_u \cdot L(u) \\ &\underset{[\text{eq: } 11.91]}{=} y \end{aligned}$$

proving, as $L(B_x \setminus B_k)$ is finite [see eq: 11.90], that $y \in \operatorname{span}(L(B_x \setminus B_k))$. Hence

$$L(X) \subseteq \operatorname{span}(L(B_x \setminus B_k)) \tag{11.92}$$

For the opposite inclusion:

$$\begin{aligned} B_x \setminus B_k \subseteq X &\Rightarrow L(B_x \setminus B_k) \subseteq L(X) \\ &\underset{[\text{theorem: } 11.104]}{\Rightarrow} \operatorname{span}(L(B_x \setminus B_k)) \subseteq \operatorname{span}(L(X)) \\ &\underset{[\text{theorems: } 11.183, 11.106]}{\Rightarrow} \operatorname{span}(L(B_x \setminus B_k)) \subseteq L(x) \end{aligned}$$

which combined with [eq: 11.92] gives

$$L(X) = \operatorname{span}(L(B_x \setminus B_k)) \tag{11.93}$$

As $B_x \setminus B_k \subseteq B_x$ a linear independent set, it follows from [theorem: 11.118] that $B_x \setminus B_k$ is linear independent. As $L_{|B_x \setminus B_k}: B_x \setminus B_k \rightarrow L(B \setminus B_y)$ is a isomorphism [see eq: 11.89] , we have by [theorem: 11.187]

$$L(B_x \setminus B_k) \text{ is linear independent}$$

The above together with [eq: 11.93] prove that

$$L(B_x \setminus B_k) \text{ is a basis of } L(X)$$

Using [eq: 11.89] we have that $\operatorname{card}(B_x \setminus B_k) = \operatorname{card}(L(B_x \setminus B_k)) = \dim(L(x)) = \operatorname{rank}(L)$, Substituting this in [eq: 11.86] gives finally

$$\dim(X) = \operatorname{rank}(L) + \dim(\ker(L))$$

So in all cases we have proved that

$$\dim(X) = \operatorname{rank}(L) + \dim(\ker(L))$$

□

11.4.3 Internal Direct Sum

Definition 11.193. *Let $\langle V, +, \cdot \rangle$ be a vector space over a field $\langle F, +, \cdot \rangle$ and $X \subseteq V$, $Y \subseteq V$ then*

$$X + Y = \{v \mid \exists (x, y) \in X \times Y \text{ such that } v = x + y\} \subseteq V$$

Theorem 11.194. *Let $\langle V,+,\cdot\rangle$ be a vector space over a field $\langle F,+,\cdot\rangle$ and X,Y sub-spaces of V the $X+Y$ is a sub-space of V.*

Proof. Let $x,y\in V_1+V_2$ and $\alpha\in F$ then there exists $v_1,w_1\in V_1$ and $v_2,w_2\in V_2$ such that $x=v_1+w_1$ and $y=v_2+w_2$. Then as V_1,V_2 are sub-spaces we have that $\alpha\cdot v_1+w_1\in V_1$ and $\alpha\cdot v_2+w_2\in V_2$ so that

$$\alpha\cdot x+y=\alpha\cdot(v_1+w_1)+(v_2+w_2)=(\alpha\cdot v_1+w_1)+(\alpha\cdot v_2+w_2)$$

hence we have

$$\alpha\cdot x+y\in V_1+V_2.$$

Further as V_1,V_2 are sub-spaces we have that $0\in V_1$ and $0\in V_2$ so that $0=0+0\in V_1+V_2$ proving that

$$V_1+V_2\neq\varnothing$$ □

Definition 11.195. (Internal Direct Sum) *Let $\langle V,+,\cdot\rangle$ be a vector space over a field $\langle F,+,\cdot\rangle$ then*

$$V=X\oplus Y\ [V \text{ is the internal direct sum of } X \text{ and } y]$$

if

1. *X and Y are sub-spaces of V*
2. *$\forall v\in V$ there exist a **unique** $x\in X\wedge y\in Y$ such that $v=x+y$*

We have the following relation between $X+Y$ and $X\oplus Y$.

Theorem 11.196. *Let $\langle V,+,\cdot\rangle$ be a vector space over a field $\langle F,+,\cdot\rangle$, X,Y sub spaces of V then*

$$V=X\oplus Y\Leftrightarrow V=X\bigcup Y \text{ and } X\bigcap Y=\{0\}$$

Proof.

⇒. As X,Y are sub-spaces of V we have by [theorem: 11.57] that $0\in X\wedge 0\in Y$ so that $0\in X\bigcap Y$ proving that

$$\{0\}\in X\bigcap Y \tag{11.94}$$

On the other hand if $v\in X\bigcap Y$ then $v\in X\wedge v\in Y$ and $v=v+0=0+v$. As $V=X\oplus Y$ we must have that $v=0\wedge 0=v$ or $v=0$. Hence $X\bigcap Y\subseteq\{0\}$ which combined with [eq: 11.94] proves that

$$X\bigcap Y=\{0\}$$

⇐. If $v\in V$ then as $V=X+Y$ there exists a $x\in X$ and $y\in Y$ such that $v=x+y$. Assume that there exists also a $x'\in X,y'\in Y$ such that $v=x'+y'$. Then $0=v-v=(x+y)-(x'+y')$ so that $x-x'=y-y'$. As X,Y are sub-spaces, hence vector spaces [see theorem: 11.57] we have $x-x'\in X\wedge y'-y\in Y$ so that $(x-x'),(y'-y)\in X\bigcap Y=\{0\}$ proving that $x=x'\wedge y=y'$. Hence we found a **unique** $x\in X$, $y\in Y$ such that $v=x+y$ which proves that

$$V=X\oplus Y$$ □

Theorem 11.197. *Let $\langle V,+,\cdot\rangle$ be a vector space over a field $\langle F,+,\cdot\rangle$ and V_1,V_2 finite dimensional sub spaces of V so that $V_1+V_2=V_1\oplus V_2$ [or equivalently that $V_1\bigcap V_2=\{0\}$] then V_1+V_2 is finite dimensional and*

$$\dim(V_1+V_2)=\dim(V_1\oplus V_2)=\dim(V_1)+\dim(V_2)$$

Further if $E_1=\{e_1^1,\ldots,e_{n_1}^1\}$ and $E_2=\{e_1^2,\ldots,e_{n_2}^2\}$ are bases for V_1 and V_2 then

$$E_1\bigcap E_2=\varnothing \text{ and } E_1\bigcup E_2 \text{ is a basis for } V_1\oplus V_2$$

Proof. Let $n_1=\dim(V_1)$, $n_2=\dim(V_2)$ and $E_1=\{e_1^1,\ldots,e_{n_1}^1\}$, $E_2=\{e_1^2,\ldots,e_{n_2}^2\}$ bases for V_1, V_2 respectively. Using [theorem: 11.110] it follows that $0\notin E_1\wedge 0\notin E_2$ so that

$$0\notin E_1\bigcap E_2\subseteq V_1\bigcap V_2=\{0\}$$

which proves that

$$E_1\bigcap E_2=\varnothing \tag{11.95}$$

Let $\{\alpha_i\}_{i\in E_1\bigcup E_2}\subseteq F$ be such that

$$\sum_{i\in E_1\bigcup E_2}\alpha_i\cdot i=0 \tag{11.96}$$

As $V_1=\mathrm{span}(E_1)$ and $V_2=\mathrm{span}(E_2)$ we have that

$$\sum_{i\in E_1}\alpha_i\cdot i\in V_1\wedge\sum_{i\in E_2}\alpha_i\cdot i\in V_2$$

Further we have

$$\begin{aligned}\underbrace{0}_{\in V_1}+\underbrace{0}_{e\in V_2} &= \underbrace{0}_{\in V_1\oplus V_2}\\ &= \sum_{i\in E_1\bigcup E_2}\alpha_i\cdot i\\ &\underset{[\text{theorem: } 11.43]}{=} \underbrace{\sum_{i\in E_1}\alpha_i\cdot i}_{\in V_1}+\underbrace{\sum_{i\in E_2}\alpha_i\cdot i}_{\in V_2}\end{aligned}$$

so by [definition: 11.195] we have that

$$\sum_{i\in E_1}\alpha_i\cdot i=0\wedge\sum_{i\in E_2}\alpha_i\cdot i=0$$

By the linear independency of E_1,E_2 it follows then that $\forall i\in E_1\ \alpha_i=0$ and $\forall i\in E_2\ \alpha_i=0$ or $\forall i\in E_1\bigcup E_2$ we have α_i. This proves that

$$E_1\bigcup E_2 \text{ is linear independent} \tag{11.97}$$

Further if $v\in V_1\oplus V_2$ there exist a $v_1\in V_1$ and $v_2\in V_2$ such that $v=v_1+v_2$, as E_1,E_2 are bases for V_1,V_2 there exists $\{\alpha_i\}_{i\in E_1}\subseteq V$, $\{\beta_i\}_{i\in E_2}\subseteq V_2$ such that $v_1=\sum_{i\in E_1}\alpha_i\cdot i$ and $v_2=\sum_{i\in E_2}\beta_i\cdot i$. Define

$$\{\gamma_i\}_{i\in E_1\bigcup E_2}\subseteq F \text{ by } \gamma_i=\begin{cases}\alpha_i \text{ if } i\in E_1\\ \beta_i \text{ if } i\in E_2\end{cases}$$

then

$$\begin{aligned}\sum_{i\in E_1\bigcup E_2}\gamma_i\cdot i &\underset{[\text{theorem: } 11.43]}{=} \sum_{i\in E_1}\gamma_i\cdot i+\sum_{i\in E_2}\gamma_i\cdot i\\ &= \sum_{i\in E_1}\alpha_i\cdot i+\sum_{i\in E_2}\beta_i\cdot i\\ &= v_1+v_2\\ &= v\end{aligned}$$

Hence we have that

$$V_1\oplus V_2\subseteq\mathrm{span}\left(E_1\bigcup E_2\right) \tag{11.98}$$

If $x\in E_1$ [or $x\in E_2$] then $x=x+0\in V_1\oplus V_2$ [or $x=0+x\in V_1\oplus V_2$] so that $E_1\bigcup E_2\subseteq V_1\oplus V_2$, hence by [theorem: 11.104] we have that $\mathrm{span}(E_1\bigcup E_2)\subseteq\mathrm{span}(V_1\oplus V_2)\underset{[\text{theorem: } 11.106]}{=}V_1\oplus V_2$. Combining this with [eq: 11.98] gives $V_1\oplus V_2=\mathrm{span}(E_1\bigcup E_2)$, hence we have, taking in account [eq: 11.97] that

$$E_1\bigcup E_2 \text{ is a basis for } V_1\oplus V_2$$

Finally

$$\dim(V_1\oplus V_2)=\mathrm{card}\left(E_1\bigcup E_2\right)\underset{[\text{eq:}11.95] \text{ and } [\text{theorem: } 10.90]}{=}\mathrm{card}(V_1)+\mathrm{card}(V_2)=\dim(V_1)+\dim(V_2)$$

□

The following is a trivial example of the decomposition of V in two sub-spaces.

Example 11.198. Let $\langle V,+,\cdot\rangle$ be a vector space over a field $\langle F,+,\cdot\rangle$ then $V=V\oplus\varnothing$

Proof. Clearly $V,\varnothing$ are sub-spaces of V and if $v\in V$ we have that $v=0+v$ so that $V=\{x+y|x\in\{0\}\wedge y\in V\}$. Finally $\{0\}\bigcap V=\{0\}$ □

So every vector space can be written as a trivial direct sum, however every vector space containing a sub-space can be written as a direct sum of this subspace and another sub-space even if the sub-space is different from $\{0\}$.

Theorem 11.199. (Fundamental theorem of subspaces) *Let $\langle V,+,\cdot\rangle$ be a vector space over a field $\langle F,+,\cdot\rangle$ and X a sub-space of $\langle V,+,\cdot\rangle$ then there exists a subspace Y of V such that $V=X\oplus Y$.*

Proof. As X is a sub-space, hence a vector space there exists by [theorem: 11.142] a basis $B_X\subseteq X$ for $\langle X,+,\cdot\rangle$. As B_X is linear independent there exists by [theorem: 11.140] a basis B for $\langle V,+,\cdot\rangle$ such that $B_X\subseteq B$. We have now either:

$\boldsymbol{B_X=B}$**.** Then $X=\operatorname{span}(B_X)=\operatorname{span}(B)=V$ so that $X=V$ and thus by [example: 11.198] we have that for $Y=\{0\}$ that

$$V=X\oplus Y$$

$\boldsymbol{B_X\neq B}$**.** Take $Y=\operatorname{span}(B\setminus B_X)$ then by [theorem: 11.105] Y is a sub-space of $\langle V,+,\cdot\rangle$. Let $v\in V$ then as $\operatorname{span}(B)=V$ there exists a finite set $I\subseteq B$ and $\{\alpha_i\}_{i\in I}\subseteq F$ such that

$$v=\sum_{i\in I}\alpha_i\cdot i\underset{[\text{theorem: }11.43]}{=}\sum_{i\in I\bigcap B_X}\alpha_i\cdot i+\sum_{i\in I\setminus B_X}\alpha_i\cdot i=x+y$$

where $x=\sum_{i\in I\bigcap B_X}\alpha_i\cdot i\in\operatorname{span}(I\bigcap B_X)$ and $y=\sum_{i\in I\setminus B_x}\alpha_i\cdot i\in\operatorname{span}(I\setminus B_X)$. As $I\setminus B_X\subseteq B\setminus B_X$ and $I\bigcap B_X\subseteq B_X$ we have by [theorem: 11.104] that $\operatorname{span}(I\setminus B_X)\subseteq\operatorname{span}(B\setminus B_X)=Y$ and $\operatorname{span}(I\bigcap B_X)\subseteq\operatorname{span}(B_X)=X$ so that $x\in X$ and $y\in Y$, proving that $V\subseteq X+Y\subseteq V$, hence

$$V=X+Y \tag{11.99}$$

As X,Y are vector spaces we have that $0\in X\bigcap Y$ so that $\{0\}\subseteq X\bigcap Y$. Let $z\in X\bigcap Y$ then $z\in X=\operatorname{span}(B_X)$ and $z\in Y=\operatorname{span}(B\setminus B_X)$, so there exists finite sets $I\subseteq B_X$, $J\subseteq B\setminus B_X$ and families $\{\alpha_i\}_{i\in I}\subseteq F$, $\{\beta_i\}_{i\in J}\subseteq F$ such that

$$\sum_{i\in I}\alpha_i\cdot i=z=\sum_{i\in J}\beta_i\cdot i \tag{11.100}$$

As $I\bigcap J\subseteq B_X\bigcap(B\setminus B_X)=\varnothing$ we can define

$$\{\gamma_i\}_{i\in I\bigcup J}\text{ by }\gamma_i=\begin{cases}\alpha_i\text{ if }i\in I\\-\beta_i\text{ if }i\in J\end{cases}$$

then

$$\begin{aligned}\sum_{i\in I\bigcup J}\gamma_i\cdot i &\underset{[\text{theorem: }11.43]}{=} \sum_{i\in I}\gamma_i\cdot i+\sum_{i\in J}\gamma_i\cdot i\\ &= \sum_{i\in I}\alpha_i\cdot i+\sum_{i\in J}(-\beta)\cdot i\\ &\underset{[\text{theorem: }11.40]}{=} \sum_{i\in I}\alpha_i\cdot i-\sum_{i\in J}\beta_i\cdot i\\ &\underset{[\text{eq: }11.100]}{=} z-z\\ &= 0\end{aligned}$$

As $I\bigcup J\subseteq B_X\bigcup(B\setminus B_X)=B$ a linear independent set we have that $\forall i\in I\bigcup J$ $\gamma_i=0$, hence $\forall i\in I$ we have $\alpha_i=\gamma_i=0$, so $z=\sum_{i\in I}\alpha_i\cdot i=\sum_{i\in I}0\cdot i\underset{[\text{theorem: }11.37]}{=}0$. So it follows that $X\bigcap Y\subseteq\{0\}$ proving as $\{0\}\subseteq X\bigcap Y$ that $X\bigcap Y=\{0\}$. Combining this with [eq: 11.99] gives finally

$$V=X\oplus Y$$

□

Theorem 11.200. *Let $\langle V,+,\cdot\rangle$ be a vector space over a field $\langle F,+,\cdot\rangle$, X,Y a sub-spaces of $\langle V,+,\cdot\rangle$ then there exists mappings*

$$\pi_X: V\rightarrow X\text{ and }\pi_Y: V\rightarrow Y$$

such that:

1. $\pi_X\in\operatorname{Hom}(V,X)$ *and* $\pi_Y\in\operatorname{Hom}(V,X)$
2. $\forall v\in V\ v=\pi_X(v)+\pi_Y(v)$
3. $(\pi_X)_{|X}=\mathrm{Id}_X$, $(\pi_X)_{|Y}=C_0$, $(\pi_Y)_{|X}=C_0$ *and* $(\pi_Y)_{|Y}=\mathrm{Id}_Y$

Proof. Let $v\in V=X\oplus Y$ then by [theorem: 11.196] there exists **unique** $v_X\in X$ and $v_Y\in Y$ such that $v=v_X+v_Y$. This allows us to define $\pi_X: V\rightarrow X$ and $\pi_Y: V\rightarrow Y$ by $\pi_X(v)=v_X$ and $\pi_Y(v)=v_Y$ so that

$$v=\pi_X(v)+\pi_Y(v)$$

Let $x, y \in V$ and $\alpha \in F$ then $x = \pi_X(x) + \pi_Y(x)$, $y = \pi_X(y) + \pi_Y(y)$. So

$$x + y = \pi_X(x) + \pi_Y(x) + \pi_X(y) + \pi_Y(y) = \underbrace{(\pi_X(x) + \pi_X(y))}_{\in X} + \underbrace{(\pi_Y(x) + \pi_Y(y))}_{\in Y}$$

which by definition proves

$$\pi_X(x + y) = \pi_X(x) + \pi_X(y) \text{ and } \pi_Y(x + y) = \pi_Y(x) + \pi_Y(y)$$

Further

$$\alpha \cdot x = \alpha \cdot (\pi_X(x) + \pi_Y(y)) = \underbrace{\alpha \cdot \pi_X(x)}_{\in X} + \underbrace{\alpha \cdot \pi_X(x)}_{\in Y}$$

which by definition proves

$$\pi_X(\alpha) = \alpha \cdot \pi_X(x) \text{ and } \pi_Y(x \cdot y) = \alpha \cdot \pi_Y(x)$$

Hence we have

$$\pi_X \in \mathrm{Hom}(V, X) \text{ and } \pi_Y \in \mathrm{Hom}(V, Y)$$

Finally if $v \in X$ then $v = v + 0 = \pi_X(v) + \pi_Y(v)$ and if $v \in Y$ then $v = 0 + v = \pi_X(v) + \pi_Y(v)$ so that

$$\forall v \in X\ v_X(v) = v = \mathrm{Id}_X(v),\ v_Y(v) = 0 = C_0(v) \text{ and } \forall v \in Y\ v_X(v) = 0 = C_0(v),\ v_Y(v) = v = \mathrm{Id}_Y(v)$$

proving that

$$(v_X)_{|X} = \mathrm{id}_X \wedge (v_X)_{|Y} = C_0 \wedge (v_Y)_{|Y} = \mathrm{Id}_Y \wedge (v_Y)_{|X} = C_0$$

□

Theorem 11.201. *Let $\langle V, +, \cdot\rangle$ be a vector space over a field $\langle F, +, \cdot\rangle$, X a sub-space of $\langle V, +, \cdot\rangle$, $\langle Z, +, \cdot\rangle$ a vector space over the same field $\langle F, +, \cdot\rangle$ and $L \in \mathrm{Hom}(X, Z)$ then there exists a $K \in \mathrm{Hom}(V, Z)$ such that $K_{|X} = L$. In other words K is a extension of L.*

Proof. Using [theorem: 11.199] there exists a sub-space Y of $\langle V, +, \cdot\rangle$ such that

$$V = X \oplus Y$$

Using [theorem: 11.200] there exists mappings

$$\pi_X: V \rightarrow X \text{ and } \pi_Y: V \rightarrow Y$$

such that:

1. $\pi_X \in \mathrm{Hom}(V, X)$ and $\pi_Y \in \mathrm{Hom}(V, X)$
2. $\forall v \in V\ v = \pi_X(v) + \pi_Y(v)$
3. $(\pi_X)_{|X} = \mathrm{Id}_X$, $(\pi_X)_{|Y} = C_0$, $(\pi_Y)_{|X} = C_0$ and $(\pi_Y)_{|Y} = \mathrm{Id}_Y$

Define then

$$K: V \rightarrow Z \text{ by } K = L \circ \pi_X$$

then if $x, y \in V$ and α we have

$$K(x + y) = L(\pi_X(x + y)) \underset{(1)}{=} L(\pi_X(x) + \pi_X(y)) = L(\pi_X(x)) + L(\pi_X(y)) = K(x) + K(y)$$

and

$$K(\alpha \cdot x) = L(\pi_X(\alpha \cdot x)) \underset{(1)}{=} L(\alpha \cdot \pi_X(x)) = \alpha \cdot L(\pi_X(x)) = \alpha \cdot K(x)$$

proving that

$$K \in \mathrm{Hom}(V, Z)$$

Finally if $x \in X$ then $K(x) = L(\pi_X(x)) \underset{(3)}{=} L(x)$ proving that

$$K_{|X} = L$$

□

11.5 Permutations

In [definition: 11.25] we have introduced the idea of permutations, bijections of a set on itself, that forms a group under function compositions. We consider now the special case of permutations of sets of the form $\{1, \ldots, n\}$ where $n \in \mathbb{N}$.

Definition 11.202. *Let $n \in \mathbb{N}$ then*

$$P_n = \{\sigma \in \{1, \ldots, n\}^{\{1, \ldots, n\}} \text{ such that } \sigma: \{1, \ldots, n\} \rightarrow \{1, \ldots, n\} \text{ is a bijection}\}$$

In other words using [definition: 11.25] $P_n = S_{\{1,\ldots,n\}}$

Example 11.203. Let $n \in \mathbb{N}$ then for

$$\rho_n: \{1, \ldots n\} \rightarrow \{1, \ldots, n\} \text{ defined by } \rho_n(i) = n - i + 1$$

we have $\rho_n \in P_n$.

ρ_n is the reversion operation, for example:

- $\rho_3(1) = 3 - 1 + 1 = 3$
- $\rho_3(2) = 3 - 2 + 1 = 2$
- $\rho_3(3) = 3 - 3 + 1 = 1$

Proof. We have:

injectivity. If $\rho_n(i) = \rho_n(j)$ then $n - i + 1 = n - j + 1$ so that $-i = -j$ hence $i = j$.

surjectivity. If $k \in \{1, \ldots, n\}$ take then $j = n - k + 1 \in \{1, \ldots, n\}$ so that $\rho_n(n - k + 1) = n - (n - k + 1) + 1 = k$

proving bijectivity, hence $\rho_n \in P_n$. □

Next we will prove that P_n is a finite set, first we show how we can extend a permutation in P_n to a permutation in P_{n+1}. Then we will show that $\{1,\ldots,n+1\} \times P_n \approx P_{n+1}$ and finally we will use mathematical induction to prove that $\forall n \in \mathbb{N}$ P_n is finite and $\text{card}(P_n) = \text{fac}(n) \underset{\text{noted}}{=} n!$. First we define $\text{fac}(n)$ using recursion.

Definition 11.204. *Let* $\text{fac}: \mathbb{N}_0 \rightarrow \mathbb{N}$ *is defined by*

$$\text{fac}(n) = \begin{cases} 1 \text{ if } n = 0 \\ n \cdot \text{fac}(n-1) \text{ if } n \in \mathbb{N} \end{cases}$$

We will use the common convention and write $n!$ *for* $\text{fac}(n)$*.*

Next we extend a permutation $\sigma \in P_n$ to a permutation $\sigma^{[i]}$ of P_{n+1}.

Lemma 11.205. *Let* $n \in \mathbb{N}$*,* $i \in \{1, \ldots, n+1\}$ *and* $\sigma \in P_n$ *then define*

$$\sigma^{[i]}: \{1, \ldots, n+1\} \rightarrow \{1, \ldots, n+1\}$$

by

1. *If* $i = n + 1$ *then*

$$\sigma^{[n+1]}(j) = \begin{cases} n+1 \text{ if } j = n+1 \\ \sigma(j) \text{ if } j \in \{1, \ldots, n\} \end{cases}$$

2. *If* $i \in \{1, \ldots, n\}$ *take* $k = \sigma^{-1}(i)$ *[hence* $\sigma(k) = i$*]*

$$\sigma^{[i]}(j) = \begin{cases} i \text{ if } j=n+1 \\ n+1 \text{ if } j = k \\ \sigma(j) \text{ if } j \in \{1, \ldots, n\} \setminus \{k\} \end{cases}$$

then we have that

$$\sigma^{[i]} \in P_{n+1} \text{ and } \sigma^{[i]}(n+1) = i$$

Proof. For $i \in \{1, \ldots, n+1\}$ we have either:

$\boldsymbol{i = n+1}$**.** Then for $\sigma^{[n+1]}: \{1, \ldots, n+1\} \rightarrow \{1, \ldots, n+1\}$ we have by definition

$$\sigma^{[n+1]}(n+1) = n+1$$

Further we have

injectivity. If $\sigma^{[n+1]}(k) = \sigma^{[n+1]}(l)$ then we have for k either:

$\boldsymbol{k \in \{1, \ldots, n\}}$**.** Then $\sigma^{[n+1]}(k) = \sigma(k) \in \{1, \ldots, n\}$ so that $n + 1 \neq \sigma^{[n+1]}(k) = \sigma^{[n+1]}(l)$, hence $\sigma^{[n+1]}(l) \in \{1, \ldots, n\}$ so that $l \in \{1,\ldots,n\}$ [as $\sigma^{[n+1]}(n+1) = n+1 \notin \{1,\ldots,n\}$]. So $\sigma(k) = \sigma^{[n+1]}(k) = \sigma^{[n+1]}(l) = \sigma(l)$ which as σ is a bijection proves that $k = l$.

$k=n+1$. If $l\in\{1,\ldots,n\}$ then $\sigma^{[n+1]}(l)=\sigma(l)\in\{1,\ldots,n\}$ so that $\sigma^{[n+1]}(l)\neq n+1=\sigma^{[n+1]}(k)$ contradicting $\sigma^{[n+1]}(k)=\sigma^{[n+1]}(l)$, hence $k=n+1=l$.

surjectivity. If $j\in\{1,\ldots,n+1\}$ then we have either:

$j=n+1$. Then $n+1\in\{1,\ldots,n+1\}$ and $\sigma^{[n+1]}(n+1)=n+1=j$

$j\in\{1,\ldots,n\}$. Then, as $\sigma:\{1,\ldots,n\}\rightarrow\{1,\ldots,n\}$ is a bijection, we have a $k\in\{1,\ldots,n\}\subseteq\{1,\ldots,n+1\}$ such that $j=\sigma(k)=\sigma^{[i]}(k)$

Hence $\sigma^{[n+1]}:\{1,\ldots,n+1\}\rightarrow\{1,\ldots,n+1\}$ is a bijection or

$$\sigma^{[n+1]}\in P_{n+1}$$

$i\in\{1,\ldots,n\}$. Let

$$k=\sigma^{-1}[i]\Rightarrow k\in\{1,\ldots,n\}\wedge\sigma(k)=i \tag{11.101}$$

then we have for $\sigma^{[i]}:\{1,\ldots,n+1\}\rightarrow\{1,\ldots,n+1\}$ that

injectivity. If $\sigma^{[i]}(r)=\sigma^{[i]}(s)$ then we have to consider the following cases for r,s:

$r=n+1\wedge s=n+1$. Then $r=s$

$r=k\wedge s=n+1$. Then

$$n+1=\sigma^{[i]}(k)=\sigma^{[i]}(r)=\sigma^{[i]}(s)=\sigma^{[i]}(n+1)=i\in\{1,\ldots,n\}$$

leading to the contradiction $n+1\leqslant n$, so this case does not occur.

$r\in\{1,\ldots,n\}\setminus\{k\}\wedge s=n+1$. Then

$$\sigma(k)\underset{[\text{eq: } 11.101]}{=}i=\sigma^{[i]}(n+1)=\sigma^{[i]}(s)=\sigma^{[i]}[r]=\sigma[r]$$

so that, as σ is injective, $r=k$ contradicting $r\in\{1,\ldots,n\}\setminus\{k\}$. So this case does not occur.

$r=n+1\wedge s=k$. Then

$$n+1=\sigma^{[i]}(k)=\sigma^{[i]}(s)=\sigma^{[i]}(r)=\sigma^{[i]}(n+1)=i\in\{1,\ldots,n\}$$

leading to the contradiction $n+1\leqslant n$. So this case does not occur.

$r=k\wedge s=k$. Then $r=s$

$r\in\{1,\ldots,n\}\setminus\{k\}\wedge s=k$. Then

$$n+1=\sigma^{[i]}[k]=\sigma^{[i]}(s)=\sigma^{[i]}(r)=\sigma(r)\in\{1,\ldots,n\}\setminus\{k\}$$

leading to the contradiction $n+1\leqslant n$. Hence this case does not occur.

$r=n+1\wedge s\in\{1,\ldots,n\}\setminus\{k\}$. Then

$$\sigma(k)=i=\sigma^{[i]}(n+1)=\sigma^{[i]}(r)=\sigma^{[i]}(s)=\sigma(s),$$

which, as σ is injective, proves that $k=s$ contradicting $s\in\{1,\ldots,n\}\setminus\{k\}$. Hence this case does not occur.

$r=k\wedge s\in\{1,\ldots,n\}\setminus\{k\}$. Then

$$n+1=\sigma^{[i]}(k)=\sigma^{[i]}(r)=\sigma^{[i]}(s)=\sigma(s)\in\{1,\ldots,n\}$$

leading to the contradiction $n+1\leqslant n$, hence this case does not apply.

$r\in\{1,\ldots,n\}\setminus\{k\}\wedge s\in\{1,\ldots,n\}\setminus\{k\}$. Then $\sigma(r)=\sigma^{[i]}(r)=\sigma^{[i]}(s)=\sigma(s)$, which as σ is injective proves that $r=s$.

So in all valid cases we have $r=s$ proving injectivity.

surjectivity. Let $j\in\{1,\ldots,n+1\}$ then we have either:

$j=n+1$. Then $\sigma^{[i]}(k)=n+1=j$

$j=i$. Then $\sigma^{[i]}(n+1)=i=j$

$j\in\{1,\ldots,n\}\setminus\{i\}$. Then as $\sigma:\{1,...,n\}\rightarrow\{1,...,n\}$ is a bijection there exists $l\in\{1,...,n\}$ such that $\sigma(l)=j$. If $l=k$ then $j=\sigma(l)=\sigma(k)=i$ contradicting $j\in\{1,...,n\}\setminus\{i\}$. So we must have that $l\in\{1,...,n\}\setminus\{k\}$, hence $\sigma^{[i]}(l)=\sigma(l)=j$

proving surjectivity.

From the above it follows that $\sigma^{[i]}:\{1,\ldots,n+1\}\rightarrow\{1,\ldots,n+1\}$ is bijective and thus that

$$\sigma^{[i]}\in P_{n+1}$$

□

Lemma 11.206. *Let $n\in\mathbb{N}$ then $\{1,\ldots,n+1\}\times P_n\approx P_{n+1}$*

Proof. Let $n\in\mathbb{N}$ and define the function

$$\beta:\{1,\ldots,n+1\}\times P_n\approx P_{n+1} \text{ by } \beta(i,\sigma)=\sigma^{[i]}$$

then we have:

injectivity. Assume that $\beta(k_1,\sigma_1)=\beta(k_2,\sigma_2)$ then $\sigma_1^{[k_1]}=\sigma_2^{[k_2]}$. So that

$$k_1\underset{[\text{theorem: } 11.205]}{=}\sigma_1^{[k_1]}(n+1)=\sigma_2^{[k_2]}(n+1)\underset{[\text{theorem: } 11.205]}{=}k_2$$

proving that

$$k_1=k_2 \tag{11.102}$$

We have now to look at two cases for k_1,k_2 corresponding with the definitions of $\sigma_1^{[k_1]}$, $\sigma_2^{[k_2]}$

$k_1=k_2=n+1$. Then if $i\in\{1,\ldots,n\}$ we have

$$\sigma_1(i)\underset{[\text{theorem: } 11.205]}{=}\sigma_1^{[k_1]}(i)=\sigma_2^{[k_2]}(i)\underset{[\text{theorem: } 11.205]}{=}\sigma_2(i)$$

proving that

$$\sigma_1=\sigma_2$$

$k_1=k_2\neq n+1$. Let $l_1=(\sigma_1)^{-1}(k_1)$ and $l_2=(\sigma_2)^{-1}(k_2)$ then we have

$$\sigma_1^{[k_1]}(l_1)\underset{[\text{theorem: } 11.205]}{=}n+1\underset{[\text{theorem: } 11.205]}{=}\sigma^{[k_2]}(l_2)$$

proving as $\sigma_1^{[k_1]},\sigma_2^{[k_2]}$ are bijections that $l_1=l_2$. Now if $i\in\{1,\ldots,n\}$ then we have either:

$i=l_1=l_2$. Then $\sigma_1(i)=\sigma_1(l_1)\underset{[\text{theorem: } 11.205]}{=}k_1=k_2\underset{[\text{theorem: } 11.205]}{=}\sigma_2(l_2)=\sigma_2(i)$

$i\neq l_1=l_2$. Then $\sigma_1(i)\underset{[\text{theorem: } 11.205]}{=}\sigma_1^{[k_1]}(i)=\sigma_2^{[k_2]}(i)\underset{[\text{theorem: } 11.205]}{=}\sigma_2(i)$

proving that

$$\sigma_1=\sigma_2$$

So in all cases we have $\sigma_1=\sigma_2$ which together with [eq: 11.102] gives $(k_1,\sigma_1)=(k_2,\sigma_2)$, proving injectivity.

surjectivity. Let $\rho\in P_{n+1}$ then for ρ we have two cases to consider:

$\rho(n+1)=n+1$. Take $\sigma=\rho_{|\{1,\ldots,n\}}$ then by [theorem: 11.27] $\rho_{|\{1,\ldots,n\}}\in P_n$. Further $\forall i\in\{1,\ldots,n+1\}$ we have

$$\begin{aligned}
\beta(n+1,\rho_{|\{1,\ldots,n\}})(i) &= (\rho_{|\{1,\ldots,n\}})^{[n+1]}(j)\\
&\underset{[\text{theorem: } 11.205]}{=} \begin{cases} n+1 \text{ if i=n+1}\\ \rho_{|\{1,\ldots,n+1\}}(i) \text{ if } i\in\{1,\ldots,n\}\end{cases}\\
&= \begin{cases} n+1 \text{ if i=n+1}\\ \rho(i) \text{ if } i\in\{1,\ldots,n\}\end{cases}\\
&= \begin{cases} \rho(n+1) \text{ if i=n+1}\\ \rho(i) \text{ if } i\in\{1,\ldots,n\}\end{cases}\\
&= \rho(i)
\end{aligned}$$

proving

$$\beta(n+1,\rho_{|\{1,\ldots,n\}})=\rho$$

$\boldsymbol{\rho(n+1)\neq n+1}$. Then $\rho(n+1)\in\{1,\ldots,n\}$. Define

$$k=\rho(n+1) \text{ and } l=\rho^{-1}(n+1) \text{ so that } \rho(l)=n+1 \tag{11.103}$$

Then we have $k\in\{1,\ldots,n\}$ and $l\in\{1,\ldots,n\}$ [for if $l=n+1$ we have $\rho(n+1)=\rho(l)\underset{[\text{eq: } 11.103]}{=}n+1$ contradicting $\rho(n+1)\neq n+1$]. Further if $i\in\{1,\ldots,n\}\setminus\{l\}$ we have that $\rho(i)\in\{1,\ldots,n\}$ [for if $\rho(i)=n+1\underset{[\text{eq: } 11.103]}{=}\rho(l)$ then, as ρ is injective, $i=l$ contradicting $i\in\{1,\ldots,n\}\setminus\{l\}$. To summarize we have

$$k,l\in\{1,\ldots,n\} \text{ and } \forall i\in\{1,\ldots,n\}\setminus\{l\} \quad \rho(i)\in\{1,\ldots,n\} \tag{11.104}$$

Define now

$$\sigma:\{1,\ldots,n\}\to\{1,\ldots,n\} \text{ by } \sigma(i)=\begin{cases} k \text{ if } i=l \\ \rho(i) \text{ if } i\in\{1,\ldots,n\}\setminus\{l\}\end{cases}$$

Then we have:

injectivity. Let $i,j\in\{1,\ldots,n\}$ such that $\sigma(i)=\sigma(j)$ then we have either:

$\boldsymbol{\sigma(i)=\sigma(j)=k}$. If $i\neq l$ then $i\in\{1,\ldots,n\}\setminus\{l\}$ so that $\sigma(i)=\rho(i)$, hence $\rho(i)=k\underset{[\text{eq: } 11.103]}{=}\rho(n+1)$ which as ρ is injective proves that $i=n+1$ contradicting $i\in\{1,\ldots,n\}\setminus\{l\}$, hence we must have that $i=l$. Likewise if $j\neq l$ then $j\in\{1,\ldots,n\}\setminus\{l\}$ so that $\sigma(j)=\rho(j)$, hence $\rho(j)=k\underset{[\text{eq: } 11.103]}{=}\rho(n+1)$ which as ρ is injective proves that $j=n+1$ contradicting $j\in\{1,\ldots,n\}\setminus\{l\}$. Hence we must have that $j=l$. So we have $i=l=j$ proving injectivity in this case.

$\boldsymbol{\sigma(i)=\sigma(j)\neq k}$. If $i=l$ then $\sigma(i)=\sigma(l)=k$ contradicting $\sigma(i)\neq k$, likewise if $j=l$ then $\sigma(j)=\sigma(l)=k$ contradicting $\sigma(j)\neq k$. So we have $i,j\in\{1,\ldots,n\}\setminus\{l\}$ so that $\rho(i)=\sigma(i)=\sigma(j)=\rho(j)$. As ρ is injective we have that $i=j$, proving injectivity in this case.

surjectivity. Let $j\in\{1,\ldots,n\}$ then we have either:

$\boldsymbol{j=k}$. Then $\sigma(l)=k=j$ where $l\in\{1,\ldots,n\}$ proving surjectivity in this case.

$\boldsymbol{j\neq k}$. Then $j\in\{1,\ldots,n\}\setminus\{k\}\subseteq\{1,\ldots,n+1\}$. By surjectivity of $\rho:\{1,\ldots,n+1\}\to\{1,\ldots,n+1\}$ there exists a $i\in\{1,\ldots,n+1\}$ such that $\rho(i)=j$. Assume that $i=n+1$ then

$$j=\rho(i)=\rho(i)=\rho(n+1)\underset{[\text{eq: } 11.103]}{=}k$$

contradicting $j\neq k$. So we must have that $i\in\{1,\ldots,n\}$. Also if $i=l$ then $j=\rho(i)=\rho(l)\underset{[\text{eq: } 11.103]}{=}n+1$ contradicting $j\in\{1,\ldots,n\}$, hence $i\neq l$ or $i\in\{1,\ldots,n\}\setminus\{l\}$, so that $\sigma(i)=\rho(i)=j$, proving surjectivity in this case.

So we have proved that $\sigma:\{1,\ldots,n\}\to\{1,\ldots,n\}$ is bijection or

$$\sigma\in P_n$$

Consider now $\sigma^{[k]}\in P_{n+1}$ and take $i\in\{1,\ldots,n+1\}$ then for i we have either:

$\boldsymbol{i=n+1}$. Then we have

$$\begin{aligned}\sigma^{[k]}(i) &= \sigma^{[k]}(n+1)\\ &\underset{[\text{theorem: } 11.205]}{=} k\\ &\underset{[\text{eq: } 11.103]}{=} \rho(n+1)\\ &= \rho(i)\end{aligned}$$

$\boldsymbol{i=l}$. Then as $\sigma(l)=k$ we have $l=\sigma^{-1}(k)$ and

$$\begin{aligned}\sigma^{[k]}(i) &= \sigma^{[k]}(l)\\ &\underset{[\text{theorem: } 11.205]\wedge\rho(n+1)\neq n+1\wedge l=\sigma^{-1}(k)}{=} n+1\\ &\underset{[\text{eq: } 11.103]}{=} \rho(l)\\ &= \rho(i)\end{aligned}$$

$i\in\{1,\ldots,n\}\setminus\{l\}$. As $\sigma(l)=k$ we have

$$i\neq l=\sigma^{-1}(k) \text{ and } i\neq n+1 \tag{11.105}$$

so that

$$\begin{aligned}\sigma^{[k]}(i) &\underset{\text{[theorem: 11.205] and [eq: 11.105]}}{=} \sigma(i)\\ &= \rho(i)\end{aligned}$$

So in all cases we have $\rho(i)=\sigma^{[k]}(i)=\beta(k,\sigma)(i)$ for every $i\in\{1,\ldots,n+1\}$. Hence $\rho=\beta(k,\sigma)$ proving surjectivity of $\beta:\{1,\ldots,n+1\}\times P_n\rightarrow P_{n+1}$.

So we have proved that $\beta:\{1,\ldots,n+1\}\times P_n\rightarrow P_n$ is a bijection hence

$$\{1,\ldots,n+1\}\times P_n\approx P_{n+1}$$

□

Finally we can use mathematical induction to prove that P_n is finite and calculate its cardinality.

Theorem 11.207. *Let $n\in\mathbb{N}$ then P_n is finite and $\operatorname{card}(P_n)=n!$*

Proof. Define

$$S=\{n\in\mathbb{N}|P_n \text{ is finite and } \operatorname{card}(P_n)=n!\}$$

then we have:

$1\in S$. Let $\sigma\in P_1$ then, as $\sigma:\{1\}\rightarrow\{1\}$ is a bijection, we must have that $\sigma=\mathrm{Id}_{\{1\}}$, so that $P_n=\{\mathrm{Id}_{\{1\}}\}$. Hence P is finite and $\operatorname{card}(P_n)=1=1!$ proving that $1\in S$.

$n\in S\Rightarrow n+1\in S$. First $\{1,\ldots,n+1\}$ is finite with $\operatorname{card}(\{1,\ldots,n+1\})\underset{\text{[theorem: 10.88]}}{=}n+1$, second, as $n\in S$, P_n is finite with $\operatorname{card}(P_n)=n!$. By [theorem:10.90] it follows then that $\{1,\ldots,n+1\}\times P_n$ is finite with

$$\operatorname{card}(\{1,\ldots,n+1\}\times P_n)=\operatorname{card}(\{1,\ldots,n+1\})\cdot\operatorname{card}(P_n)=(n+1)\cdot n!=(n+1)!$$

Using [lemma: 11.206] we have

$$\{1,\ldots,n+1\}\times P_n\approx P_{n+1}$$

so that P_{n+1} is finite and $\operatorname{card}(P_{n+1})=(n+1)!$, proving that $n+1\in S$. □

Theorem 11.208. *Let $n\in\mathbb{N}$ then $\langle P_n,\circ\rangle$ is a group called the permutation group*

Proof. Using [definition: 11.25] we have that $P_n=S_{\{1,\ldots,n\}}$, the rest follows from [theorem: 11.26]. □

Note 11.209. P_n is not commutative

Proof. For example if $f_1=\begin{pmatrix}1&\rightarrow&3\\2&\rightarrow&2\\3&\rightarrow&1\end{pmatrix}$ and $f_2=\begin{pmatrix}1&\rightarrow&2\\2&\rightarrow&1\\3&\rightarrow&3\end{pmatrix}$ then $f_1\circ f_2=\begin{pmatrix}1&\rightarrow&2\\2&\rightarrow&3\\3&\rightarrow&1\end{pmatrix}$ and $f_2\circ f_1=\begin{pmatrix}1&\rightarrow&3\\2&\rightarrow&1\\3&\rightarrow&2\end{pmatrix}$ so that $f_1\circ f_2\neq f_2\circ f_1$ □

Definition 11.210. *Let $n\in\mathbb{N}$. As $\langle P_n,\circ\rangle$ forms a group we can define the finite product of permutations [see remark: 11.2]. Given a family $\{\sigma_i\}_{i\in\{1,\ldots,n\}}\subseteq P_n$ we define $\sigma_1\circ\cdots\circ\sigma_n$ by*

$$\sigma_1\circ\cdots\circ\sigma_n=\prod_{i=1}^{n}\sigma_i$$

(see [definition: 11.12] using $\circ$ for product).

Theorem 11.211. *Let $m,n\in\mathbb{N}$ and $\{\sigma_i\}_{i\in\{1,\ldots,m\}}\subseteq P_n$ then we have*

1. *If $m=1$ then $\sigma_1\circ\cdots\circ\sigma_1=\sigma_1$*
2. *If $1<m$ then*

$$\begin{aligned}\sigma_1\circ\cdots\circ\sigma_m &= (\sigma_1\circ\cdots\circ\sigma_{m-1})\circ\sigma_m\\ &= \sigma_1\circ(\sigma_2\circ\cdots\circ\sigma_m)\end{aligned}$$

Proof.

1. We have

$$\sigma_1 \circ \cdots \circ \sigma_1 = \prod_{i=1}^{1} \sigma_1 \underset{[\text{theorem: } 11.15]}{=} \sigma_1$$

2. If $1 < m$ then we have

$$\begin{array}{rcl}
\sigma_1 \circ \cdots \circ \sigma_m & = & \prod_{i=1}^{m} \sigma_i \\
& \underset{[\text{theorem: } 11.15]}{=} & \left(\prod_{i=1}^{m-1} \sigma_i\right) \circ \sigma_m \\
& = & (\sigma_1 \circ \cdots \circ \sigma_{m-1}) \circ \sigma_m \\
\sigma_1 \circ \cdots \circ \sigma_m & = & \prod_{i=1}^{m} \sigma_i \\
& \underset{\text{def}}{=} & \prod_{i=0}^{m-1} \sigma_{i+1} \\
& \underset{[\text{theorem: } 11.22}{=} & \sigma_{0+1} \circ \prod_{i=1}^{m-1} \sigma_{i+1} \\
& \underset{\text{def}}{=} & \sigma_1 \circ \prod_{i=2}^{m} \sigma_i \\
& = & \sigma_1 \circ (\sigma_2 \circ \cdots \circ \sigma_m)
\end{array}$$

□

11.5.1 Transpositions

Definition 11.212. (Transposition) *Let $n \in \mathbb{N}$ and $i, j \in \{1, \ldots, n\}$ then a **transposition** $\left(i \underset{n}{\leftrightarrow} j\right) \in P_n$ is defined by*

$$\left(i \underset{n}{\leftrightarrow} j\right) \underset{[\textit{definition: } 11.28]}{=} \left(i \underset{\{1,\ldots,n\}}{\leftrightarrow} j\right)$$

or using [definition: 11.28]

$$\left(i \underset{n}{\leftrightarrow} j\right)(k) = \begin{cases} k \text{ if } k \in \{1, \ldots, n\} \setminus \{i, j\} \\ j \text{ if } k = i \\ i \text{ if } k = j \end{cases}$$

*If $i \neq j$ then $\left(i \underset{n}{\leftrightarrow} j\right)$ is a **strict transposition** [note that for a transposition to be strict we must have that $n \in \mathbb{N} \setminus \{1\}$].*

Theorem 11.213. *Let $n \in \mathbb{N}$ then we have for $i, j \in \{1, \ldots, n\}$*

1. *If $i = j$ then $\left(i \underset{n}{\leftrightarrow} j\right) = \mathrm{Id}_{\{1,\ldots,n\}}$*
2. $\left(i \underset{n}{\leftrightarrow} j\right) = \left(j \underset{n}{\leftrightarrow} i\right)$
3. $\left(i \underset{n}{\leftrightarrow} j\right) \circ \left(i \underset{n}{\leftrightarrow} j\right) = \mathrm{Id}_{\{1,\ldots,n\}}$
4. $\left(i \underset{n}{\leftrightarrow} j\right) = \left(i \underset{n+1}{\leftrightarrow} j\right)_{|\{1,\ldots,n\}}$

Proof. This is proved in [theorem: 11.29] □

We can always extend a permutation on $\{1, \ldots, n\}$ to a permutation on $\{1, \ldots, n+1\}$ as the following theorem shows.

Theorem 11.214. *Let $n, m \in \mathbb{N}$ with $n < m$ then for every $\sigma \in P_n$ define*

$$\sigma^{|\{1,\ldots m\}}: \{1, \ldots, m\} \Rightarrow \{1, \ldots, m\}$$

by

$$\sigma^{|\{1,\ldots,m\}}(i) = \begin{cases} i \text{ if } i \in \{n+1, \ldots, m\} = \{1, \ldots, m\} \setminus \{1, \ldots, n\} \\ \sigma(i) \text{ if } i \in \{1, \ldots, n\} \end{cases} \in \{1, \ldots, m\}$$

then we have

1. $\sigma^{|\{1,\ldots,m\}} \in P_m$
2. $\forall i, j \in \{1, \ldots, n\}$ *we have* $\left(i \underset{n}{\leftrightarrow} j\right)^{|\{1,\ldots,m\}} = \left(i \underset{m}{\leftrightarrow} j\right)$
3. *If* $\sigma, \rho \in P_n$ *then* $(\sigma \circ \rho)^{|\{1,\ldots,m\}} = \sigma^{|\{1,\ldots,m\}} \circ \rho^{|\{1,\ldots,m\}}$
4. *If* $k \in \mathbb{N}$ *and* $\{\sigma_i\}_{i \in \{1,\ldots,k\}} \subseteq P_n$ *then*

$$(\sigma_1 \circ \cdots \circ \sigma_n)^{|\{1,\ldots,m\}} = (\sigma_1)^{|\{1,\ldots,m\}} \circ \cdots \circ (\sigma_n)^{|\{1,\ldots,m\}}$$

Proof.

1. We have:

   injectivity. For $i, j \in \{1, \ldots, m\}$ with $\sigma^{|\{1,\ldots,m\}}(i) = \sigma^{|\{1,\ldots,m\}}(j)$ we have either:

   $\boldsymbol{i, j \in \{1, \ldots, n\}}$**.** Then we have $\sigma(i) = \sigma^{|\{1,\ldots,m\}}(i) = \sigma^{|\{1,\ldots,m\}}(j) = \sigma(j)$ proving, as σ is injective, that $i = j$.

   $\boldsymbol{i \in \{1, \ldots, n\} \wedge j \in \{n+1, \ldots, m\}}$**.** Then $\sigma(i) = \sigma^{|\{1,\ldots,m\}}(i) = \sigma^{|\{1,\ldots,m\}}(j) = j$, hence as $\sigma(i) \in \{1, \ldots, n\}$ we have that $j \in \{1, \ldots, n\}$ contradicting $j \in \{n+1, \ldots, m\}$. So this case does not apply.

   $\boldsymbol{i \in \{n+1, \ldots, m\} \wedge j \in \{1, \ldots, n\}}$**.** Then $\sigma(j) = \sigma^{|\{1,\ldots,m\}}(j) = \sigma^{|\{1,\ldots,m\}}(i) = i$, hence as $\sigma(j) \in \{1, \ldots, n\}$ we have that $i \in \{1, \ldots, n\}$ contradicting $i \in \{n+1, \ldots, m\}$. So this case does not apply.

   $\boldsymbol{i, j \in \{n+1, \ldots, m\}}$**.** Then we have $i = \sigma^{|\{1,\ldots,m\}}(i) = \sigma^{|\{1,\ldots,m\}}(j) = j$ or $i = j$.

   So in all cases $i = j$.

   surjectivity. If $j \in \{1, \ldots, m\}$ then either $j \in \{n+1, \ldots, m\}$ so that $j = \sigma^{|n+1}(j)$ or $j \in \{1, \ldots, n\}$ and then there exists a $i \in \{1, \ldots, n\}$ so that $j = \sigma(i) = \sigma^{|\{1,\ldots,m\}}(i)$.

   So $\sigma^{|\{1,\ldots m\}}\colon \{1, \ldots, m\} \Rightarrow \{1, \ldots, m\}$ is a bijection or $\sigma^{|\{1,\ldots,m\}} \in P_m$

2. Let $k \in \{1, \ldots, m\}$ then we have either:

   $\boldsymbol{k \in \{1, \ldots, n\}}$**.** Then as $i, j \in \{1, \ldots, n\}$ we have

$$\left(i \underset{n}{\leftrightarrow} j\right)^{|\{1,\ldots,m\}}(k) = \left(i \underset{n}{\leftrightarrow} j\right)(k) = \begin{cases} i \text{ if } k = j \\ j \text{ if } k = i \\ k \text{ if } i \in \{1, \ldots, n\} \setminus \{i, j\} \end{cases} = \left(i \underset{m}{\leftrightarrow} j\right)(k)$$

   $\boldsymbol{k \in \{n+1, \ldots, m\}}$**.** Then as $i, j \in \{1, \ldots, n\}$ we have that $i, j \neq k$ so that

$$\left(i \underset{n}{\leftrightarrow} j\right)^{|\{1,\ldots,m\}}(k) = k = \left(i \underset{m}{\leftrightarrow} j\right)(k)$$

   So in all cases $\left(i \underset{n}{\leftrightarrow} j\right)^{|\{1,\ldots,m\}}(k) = \left(i \underset{m}{\leftrightarrow} j\right)(k)$ proving that

$$\left(i \underset{n}{\leftrightarrow} j\right)^{|\{1,\ldots,m\}} = \left(i \underset{m}{\leftrightarrow} j\right)$$

3. Let $i \in \{1, \ldots, m\}$ then we have either:

   $\boldsymbol{i \in \{1, \ldots, n\}}$**.** Then

$$\begin{aligned} (\sigma \circ \rho)^{|\{1,\ldots,m\}}(i) &= (\sigma \circ \rho)(i) \\ &= \sigma(\rho(i)) \\ &= \sigma^{|\{1,\ldots,m\}}(\rho^{|\{1,\ldots,m\}}(i)) \\ &= (\sigma^{|\{1,\ldots,m\}} \circ \rho^{|\{1,\ldots,m\}})(i) \end{aligned}$$

   $\boldsymbol{i \in \{n+1, \ldots, m\}}$**.** Then

$$\begin{aligned} (\sigma \circ \rho)^{|\{1,\ldots,m\}}(i) &= i \\ &= \sigma^{|\{1,\ldots,m\}}(i) \\ &= \sigma^{|\{1,\ldots,m\}}(\rho^{|\{1,\ldots,m\}}(i)) \\ &= (\sigma^{|\{1,\ldots,m\}} \circ \rho^{|\{1,\ldots,m\}})(i) \end{aligned}$$

So $\forall i \in \{1,\ldots,m\}$ we have $(\sigma\circ\rho)^{|\{1,\ldots,m\}}(i) = (\sigma^{|\{1,\ldots,m\}}\circ\rho^{|\{1,\ldots,m\}})(i)$ proving that

$$(\sigma\circ\rho)^{|\{1,\ldots,m\}} = \sigma^{|\{1,\ldots,m\}}\circ\rho^{|\{1,\ldots,m\}}$$

4. We prove this by induction, so let

$$S = \{k\in\mathbb{N}|\text{If } \{\sigma_i\}_{i\in\{1,\ldots,k\}}\subseteq P_n \text{ then } (\sigma_1\circ\cdots\circ\sigma_k)^{|\{1,\ldots,m\}} = (\sigma_1)^{|\{1,\ldots,m\}})\circ\cdots\circ(\sigma_k)^{|\{1,\ldots,m\}}\}$$

then we have:

$1\in S$. Let $\{\sigma_i\}_{i\in\{1,\ldots,1\}}\subseteq P_n$ then

$$\begin{aligned}(\sigma_1\circ\cdots\circ\sigma_n)^{|\{1,\ldots,m\}} &\underset{[\text{theorem: } 11.211]}{=} (\sigma_1)^{|\{1,\ldots,m\}}\\ &\underset{[\text{theorem: } 11.211]}{=} ((\sigma_1)^{|\{1,\ldots,m\}})\circ\cdots\circ(\sigma_k)^{|\{1,\ldots,m\}}\end{aligned}$$

proving that $1\in S$.

$n\in S\Rightarrow n+1\in S$. Let $\{\sigma_i\}_{i\in\{1,\ldots,n+1\}}\subseteq P_n$ then we have

$$\begin{aligned}(\sigma_1\circ\cdots\circ\sigma_{n+1})^{|\{1,\ldots,m\}} &\underset{[\text{theorem: } 11.211]}{=}\\ ((\sigma_1\circ\cdots\circ\sigma_n)\circ\sigma_{n+1})^{|\{1,\ldots,m\}} &\underset{(3)}{=}\\ (\sigma_1\circ\cdots\circ\sigma_n)^{|\{1,\ldots,m\}}\circ(\sigma_{n+1})^{|\{1,\ldots,m\}} &\underset{n\in S}{=}\\ ((\sigma_1)^{|\{1,\ldots,m\}}\circ\cdots\circ(\sigma_n)^{|\{1,\ldots,m\}})\circ(\sigma_{n+1})^{|\{1,\ldots,m\}} &\underset{[\text{theorem: } 11.211]}{=}\\ ((\sigma_1)^{|\{1,\ldots,m\}}\circ\cdots\circ(\sigma_{n+1})^{|\{1,\ldots,m\}})\end{aligned}$$

proving that $n+1\in S$

□

Corollary 11.215. *Let $n\in\mathbb{N}$ and $\sigma\in P_{n+1}$ such that $\sigma(n+1)=n+1$ [so that by [theorem: 11.27] $\sigma_{|\{1,\ldots,n\}}\in P_n$] then*

$$(\sigma_{|\{1,\ldots,n\}})^{|\{1,\ldots,n+1\}} = \sigma$$

Proof. *Let $k\in\{1,\ldots,n+1\}$ then we have either:*

$k\in\{1,\ldots,n\}$. *Then $(\sigma_{|\{1,\ldots,n\}})^{|\{1,\ldots,n+1\}}(k) = \sigma_{|\{1,\ldots,n\}}(k) = \sigma(k)$*

$k=n+1$. *Then $(\sigma_{|\{1,\ldots,n\}})^{|\{1,\ldots,n+1\}}(k) = n+1 = \sigma(n+1)$* □

So $\forall k\in\{1,\ldots,n+1\}$ we have $(\sigma_{|\{1,\ldots,n\}})^{|\{1,\ldots,n+1\}}(k)=\sigma(k)$ proving that

$$(\sigma_{|\{1,\ldots,n\}})^{|\{1,\ldots,n+1\}} = \sigma$$

We show now that every permutation of more then one element can be written as the composition of disjoint transpositions.

Theorem 11.216. *Let $n\in\mathbb{N}\setminus\{1\}$ and $\sigma\in P_n$ then there exists a family $\{(i_k \underset{n}{\leftrightarrow} j_k)\}_{k\in\{1,\ldots,l\}}\subseteq P_n$ with $\forall k\in\{1,\ldots,l\}$ $i_k\neq j_k$ such that*

$$\sigma = (i_1\underset{n}{\leftrightarrow} j_1)\circ\cdots\circ(i_l\underset{n}{\leftrightarrow} j_l)$$

Proof. We use mathematical induction to prove this, so let

$$S = \{n\in\{2,\ldots,\infty\}|\text{If } \sigma\in P_n \text{ then } \sigma = (i_1\underset{n}{\leftrightarrow} j_1)\circ\cdots\circ(i_l\underset{n}{\leftrightarrow} j_l) \text{ where } \{(i_k\underset{n}{\leftrightarrow} j_k)\}_{k\in\{1,\ldots,l\}}\subseteq P_n \text{ satisfies } i_k\neq j_k\,\forall k\in\{1,\ldots,l\}\}$$

then we have:

$2\in S$. Then for $\sigma\in P_2$ we have for $\sigma(1)$ either:

$\sigma(1)=1$. Then we must have, as σ is injective, that $\sigma(2)=2$ hence $\sigma=\mathrm{Id}_{\{1,\ldots,2\}}$. Define $\{(i_k\leftrightarrow j_k)\}_{k\in\{1,\ldots,2\}}$ by $(i_1\underset{2}{\leftrightarrow} j_1) = (i_2\underset{2}{\leftrightarrow} j_2) = (1\underset{2}{\leftrightarrow} 2)$, a family of **strict** transpositions. Then

$$(i_1\underset{2}{\leftrightarrow} j_1)\circ\cdots\circ(i_2\underset{2}{\leftrightarrow} j_2) = (1\underset{2}{\leftrightarrow} 2)\circ(1\underset{2}{\leftrightarrow} 2) \underset{[\text{theorem: } 11.213]}{=} \mathrm{Id}_{\{1,\ldots,2\}} = \sigma$$

proving that in this cases $2\in S$.

$\boldsymbol{\sigma(1)=2}$. Then we must have, as σ is injective, that $\sigma(2)=1$ hence $\sigma=\left(1\underset{2}{\leftrightarrow}2\right)$. Define $\{(i_k\leftrightarrow j_k)\}_{k\in\{1,\ldots,1\}}$ by $\left(i_1\underset{2}{\leftrightarrow}j_1\right)=\left(1\underset{2}{\leftrightarrow}2\right)$, a family of strict transpositions. Then

$$\left(i_1\underset{2}{\leftrightarrow}j_1\right)\circ\cdots\circ\left(i_1\underset{2}{\leftrightarrow}j_1\right)=\left(1\underset{2}{\leftrightarrow}2\right)=\sigma$$

$\boldsymbol{n\in S\Rightarrow n+1\in S}$. Let $\sigma\in P_{n+1}$ then for $\sigma(n+1)$ we have either:

$\boldsymbol{\sigma(n+1)=n+1}$. Using [theorem: 11.27] we have that $\sigma_{|\{1,\ldots,n\}}\in P_n$. As $n\in S$ there exists a family of strict transpositions $\left\{\left(i_k\underset{n}{\leftrightarrow}j_k\right)\right\}_{k\in\{1,\ldots,l\}}\subseteq P_n$ such that

$$\sigma_{|\{1,\ldots,n\}}=\left(i_1\underset{n}{\leftrightarrow}j_1\right)\circ\cdots\circ\left(i_l\underset{n}{\leftrightarrow}j_l\right)$$

So that

$$\begin{aligned}
\sigma &\underset{[\text{theorem: }11.215]}{=} \left(\sigma_{|\{1,\ldots,n\}}\right)^{|\{1,\ldots,n+1\}}\\
&= \left(\left(i_1\underset{n}{\leftrightarrow}j_1\right)\circ\cdots\circ\left(i_l\underset{n}{\leftrightarrow}j_l\right)\right)^{|\{1,\ldots,n+1\}}\\
&\underset{[\text{theorem: }11.214]}{=} \left(\left(i_1\underset{n}{\leftrightarrow}j_1\right)^{|\{1,\ldots,n+1\}}\circ\cdots\circ\left(i_l\underset{n}{\leftrightarrow}j_l\right)^{|\{1,\ldots,n+1\}}\right)\\
&\underset{[\text{theorem: }11.214]}{=} \left(i_1\underset{n+1}{\leftrightarrow}j_1\right)\circ\cdots\circ\left(i_l\underset{n+1}{\leftrightarrow}j_l\right)
\end{aligned}$$

proving that in this case $n+1\in S$

$\boldsymbol{\sigma(n+1)\neq n+1}$. Then as σ is a bijection there exists a $i\in\{1,\ldots,n+1\}\setminus\{n+1\}=\{1,\ldots,n\}$ such that $\sigma(i)=n+1$. Define then

$$\rho=\sigma\circ\left(i\underset{n+1}{\leftrightarrow}n+1\right)\tag{11.106}$$

where

$$\left(i\underset{n+1}{\leftrightarrow}n+1\right)\text{ is a strict transposition as }i\neq n+1$$

then we have $\rho(n+1)=\left(\sigma\left(\left(i\underset{n+1}{\leftrightarrow}n+1\right)(i)\right)\right)=\sigma(i)=n+1$, so using [theorem: 11.27] we have that $\rho_{|\{1,\ldots,n\}}\in P_n$. As $n\in S$ there exists a $\{(r_k\underset{n}{\leftrightarrow}s_k)\}_{k\in\{1,\ldots,l\}}\subseteq P_n$ such that

$$\rho_{|\{1,\ldots,n\}}=\left((r_1\underset{n}{\leftrightarrow}s_1)\circ\cdots\circ(r_l\underset{n}{\leftrightarrow}s_l)\right)$$

So that

$$\begin{aligned}
\rho &\underset{[\text{theorem: }11.215]}{=} \left(\rho_{|\{1,\ldots,n\}}\right)^{|\{1,\ldots,n+1\}}\\
&= \left((r_1\underset{n}{\leftrightarrow}s_1)\circ\cdots\circ(r_l\underset{n}{\leftrightarrow}s_l)\right)^{|\{1,\ldots,n+1\}}\\
&\underset{[\text{theorem: }11.214]}{=} \left((r_1\underset{n}{\leftrightarrow}s_1)^{|\{1,\ldots,n+1\}}\circ\cdots\circ(r_l\underset{n}{\leftrightarrow}s_l)^{|\{1,\ldots,n+1\}}\right)\\
&\underset{[\text{theorem: }11.214]}{=} \left(r_1\underset{n+1}{\leftrightarrow}s_1\right)\circ\cdots\circ\left(r_l\underset{n+1}{\leftrightarrow}s_l\right)
\end{aligned}\tag{11.107}$$

Define now

$$\left\{\left(i_k\underset{n+1}{\leftrightarrow}j_k\right)\right\}_{k\in\{1,\ldots,l+1\}}\text{ by }\left(i_k\underset{n}{\leftrightarrow}j_k\right)=\begin{cases}\left(i\underset{n+1}{\leftrightarrow}n+1\right)\text{ if }k=l+1\\ \left(r_k\underset{n+1}{\leftrightarrow}s_k\right)\text{ if }k\in\{1,\ldots,l\}\end{cases}$$

then we have

$$\begin{aligned}
\sigma &\underset{[\text{theorem: }11.213]}{=} \sigma\circ\left(\left(i\underset{n+1}{\leftrightarrow}n+1\right)\circ\left(i\underset{n+1}{\leftrightarrow}n+1\right)\right)\\
&\underset{\text{associativity}}{=} \left(\sigma\circ\left(i\underset{n+1}{\leftrightarrow}n+1\right)\right)\circ\left(i\underset{n+1}{\leftrightarrow}n+1\right)\\
&\underset{[\text{eq: }11.106]}{=} \rho\circ\left(i\underset{n+1}{\leftrightarrow}n+1\right)\\
&\underset{[\text{theorem: }11.107]}{=} \left(\left(r_1\underset{n+1}{\leftrightarrow}s_1\right)\circ\cdots\circ\left(r_l\underset{n+1}{\leftrightarrow}s_l\right)\right)\circ\left(i\underset{n+1}{\leftrightarrow}n+1\right)\\
&= \left(\left(i_1\underset{n+1}{\leftrightarrow}j_1\right)\circ\cdots\circ\left(i_l\underset{n+1}{\leftrightarrow}j_l\right)\right)\circ\left(i_{n+1}\underset{n+1}{\leftrightarrow}j_{n+1}\right)\\
&\underset{[\text{theorem: }11.211]}{=} \left(i_1\underset{n+1}{\leftrightarrow}j_1\right)\circ\cdots\circ\left(i_{l+1}\underset{n+1}{\leftrightarrow}j_{l+1}\right)
\end{aligned}$$

proving that $n+1\in S$

So in all cases we have $n+1\in S$.

Mathematical induction proves then that $S=\mathbb{N}\setminus\{1\}=\{2,\ldots\infty\}$ completing the proof. □

TODO check this lemma, the next theorem and corollary 11.225

Lemma 11.217. *Let $n\in\mathbb{N}$, $i,j\in\{1,\ldots,n\}$, $m\in\mathbb{N}$ and $\{(k_r,l_r)\}_{r\in\{1,\ldots,m\}}\subseteq\{1,\ldots,n\}$ then if*

$$(i\underset{n}{\leftrightarrow}j)=(k_1\underset{n}{\leftrightarrow}l_1)\circ\ldots\circ(k_m\underset{n}{\leftrightarrow}l_m)$$

then

$$\left(i\underset{n+1}{\leftrightarrow}j\right)=\left(k_1\underset{n+1}{\leftrightarrow}l_1\right)\circ\ldots\circ\left(k_m\underset{n+1}{\leftrightarrow}l_n\right)$$

Proof. First we prove by induction that

$$(k_1\underset{n}{\leftrightarrow}l_1)\circ\ldots\circ(k_m\underset{n}{\leftrightarrow}l_m)=\left(\left(k_1\underset{n+1}{\leftrightarrow}l_1\right)\circ\ldots\circ\left(k_m\underset{n+1}{\leftrightarrow}l_m\right)\right)_{|\{1,\ldots,n\}}\tag{11.108}$$

So let

$$S=\left\{m\in\mathbb{N}|\text{If }\{(k_r,l_r)\}_{r\in\{1,\ldots,m\}}\subseteq\{1,\ldots,n\}\text{ then }(k_1\underset{n}{\leftrightarrow}l_1)\circ\ldots\circ(k_m\underset{n}{\leftrightarrow}l_m)=\left(\left(k_1\underset{n+1}{\leftrightarrow}l_1\right)\circ\ldots\circ\left(k_m\underset{n+1}{\leftrightarrow}l_m\right)\right)_{|\{1,\ldots,n\}}\right\}$$

then we have:

$1\in S$. As

$$\begin{aligned}(k_1\underset{n}{\leftrightarrow}l_1)\circ\ldots\circ(k_1\underset{n}{\leftrightarrow}l_1) &= (k_1\underset{n}{\leftrightarrow}l_1)\\ &\underset{[\text{theorem: }11.213]}{=} \left(k_1\underset{n+1}{\leftrightarrow}l_1\right)_{|\{1,\ldots,n\}}\\ &= \left(\left(k_1\underset{n+1}{\leftrightarrow}l_1\right)\circ\ldots\circ\left(k_1\underset{n+1}{\leftrightarrow}l_1\right)\right)_{|\{1,\ldots,n\}}\end{aligned}$$

proving that $1\in S$.

$m\in S\Rightarrow m+1\in S$. Consider $(k_1\underset{n}{\leftrightarrow}l_1)\circ\ldots\circ(k_{m+1}\underset{n}{\leftrightarrow}l_{m+1})$. Let $k\in\{1,\ldots,n\}$ then we have

$$\begin{aligned}((k_1\underset{n}{\leftrightarrow}l_1)\circ\ldots\circ(k_{m+1}\underset{n}{\leftrightarrow}l_{m+1}))(k) &=\\ (((k_1\underset{n}{\leftrightarrow}l_1)\circ\ldots\circ(k_m\underset{n}{\leftrightarrow}l_m))\circ(k_{m+1}\underset{n}{\leftrightarrow}l_{m+1}))(k) &\underset{m\in S}{=}\\ \left(\left(\left(k_1\underset{n+1}{\leftrightarrow}l_1\right)\circ\ldots\circ\left(k_m\underset{n+1}{\leftrightarrow}l_m\right)\right)_{|\{1,\ldots,n\}}\circ(k_{m+1}\underset{n}{\leftrightarrow}l_{m+1})\right)(k) &=\\ \left(\left(k_1\underset{n+1}{\leftrightarrow}l_1\right)\circ\ldots\circ\left(k_m\underset{n+1}{\leftrightarrow}l_m\right)\right)_{|\{1,\ldots,n\}}((k_{m+1}\underset{n}{\leftrightarrow}l_{m+1})(k)) &\underset{(k_{m+1}\underset{n}{\leftrightarrow}l_{m+1})(k)\in\{1,\ldots,n\}}{=}\\ \left(\left(k_1\underset{n+1}{\leftrightarrow}l_1\right)\circ\ldots\circ\left(k_m\underset{n+1}{\leftrightarrow}l_m\right)\right)((k_{m+1}\underset{n}{\leftrightarrow}l_{m+1})(k)) &=\\ \left(\left(\left(k_1\underset{n+1}{\leftrightarrow}l_1\right)\circ\ldots\circ\left(k_m\underset{n+1}{\leftrightarrow}l_m\right)\right)\circ(k_{m+1}\underset{n}{\leftrightarrow}l_{m+1})\right)(k) &=\\ \left(\left(k_1\underset{n+1}{\leftrightarrow}l_1\right)\circ\ldots\circ\left(k_{m+1}\underset{n+1}{\leftrightarrow}l_{m+1}\right)\right)(k) &\underset{k\in\{1,\ldots,n\}}{=}\\ \left(\left(k_1\underset{n+1}{\leftrightarrow}l_1\right)\circ\ldots\circ\left(k_{m+1}\underset{n+1}{\leftrightarrow}l_{m+1}\right)\right)_{|\{1,\ldots,n\}}(k)\end{aligned}$$

proving that

$$(k_1\underset{n}{\leftrightarrow}l_1)\circ\ldots\circ(k_{m+1}\underset{n}{\leftrightarrow}l_{m+1})=\left(\left(k_1\underset{n+1}{\leftrightarrow}l_1\right)\circ\ldots\circ\left(k_{m+1}\underset{n+1}{\leftrightarrow}l_{m+1}\right)\right)_{|\{1,\ldots,n\}}$$

so that

$$n+1\in S$$

Next we use induction to prove that

$$((k_1\underset{n}{\leftrightarrow}l_1)\circ\ldots\circ(k_m\underset{n}{\leftrightarrow}l_m))(n+1)=n+1\tag{11.109}$$

So let

$$S=\left\{m\in\mathbb{N}|\text{If }\{(k_r,l_r)\}_{r\in\{1,\ldots,m\}}\subseteq\{1,\ldots,n\}\text{ then }\left(\left(k_1\underset{n+1}{\leftrightarrow}l_1\right)\circ\ldots\circ\left(k_m\underset{n+1}{\leftrightarrow}l_m\right)\right)(n+1)=n+1\right\}$$

then we have:

$1\in S$. Then $\left(\left(k_1\underset{n+1}{\leftrightarrow}l_1\right)\circ\ldots\circ\left(k_1\underset{n+1}{\leftrightarrow}l_1\right)\right)(n+1)=\left(k_1\underset{n+1}{\leftrightarrow}l_1\right)(n+1)\underset{n+1\neq k_1,l_1}{=}n+1$ proving that $1\in S$.

$\boldsymbol{m\in S\Rightarrow m+1\in S.}$ As we have

$$\begin{array}{rc}\left(\left(k_1 \underset{n+1}{\leftrightarrow} l_1\right)\circ\ldots\circ\left(k_{m+1}\underset{n+1}{\leftrightarrow} l_{m+1}\right)\right)(n+1) & = \\ \left(\left(\left(k_1 \underset{n+1}{\leftrightarrow} l_1\right)\circ\ldots\circ\left(k_m\underset{n+1}{\leftrightarrow} l_m\right)\right)\circ\left(k_{m+1}\underset{n+1}{\leftrightarrow} l_{m+1}\right)\right)(n+1) & = \\ \left(\left(k_1 \underset{n+1}{\leftrightarrow} l_1\right)\circ\ldots\circ\left(k_m\underset{n+1}{\leftrightarrow} l_m\right)\right)\left(\left(k_{m+1}\underset{n+1}{\leftrightarrow} l_{m+1}\right)(n+1)\right) & \underset{n+1\neq k_{m+1},l_{m+1}}{=} \\ \left(\left(k_1 \underset{n+1}{\leftrightarrow} l_1\right)\circ\ldots\circ\left(k_m\underset{n+1}{\leftrightarrow} l_m\right)\right)(n+1) & \underset{m\in S}{=} \\ n+1 & \end{array}$$

it is proved that

$$m+1\in S$$

Now we are ready for the rest of the proof. Let $k\in\{1,\ldots,n+1\}$ then we have either:

$\boldsymbol{k=n+1.}$ Then we have

$$\begin{array}{rcl}\left(i\underset{n+1}{\leftrightarrow} j\right)(k) & = & \left(i\underset{n+1}{\leftrightarrow} j\right)(n+1) \\ & \underset{n+1\neq i,j}{=} & n+1 \\ & \underset{\text{[eq: 11.109]}}{=} & \left(\left(k_1 \underset{n+1}{\leftrightarrow} l_1\right)\circ\ldots\circ\left(k_m\underset{n+1}{\leftrightarrow} l_m\right)\right)(n+1) \\ & = & \left(\left(k_1 \underset{n+1}{\leftrightarrow} l_1\right)\circ\ldots\circ\left(k_m\underset{n+1}{\leftrightarrow} l_m\right)\right)(k)\end{array}$$

$\boldsymbol{k\in\{1,\ldots,n\}.}$ Then we have

$$\begin{array}{rcl}\left(i\underset{n+1}{\leftrightarrow} j\right)(k) & = & \left(i\underset{n+1}{\leftrightarrow} j\right)_{|\{1,\ldots,n\}}(k) \\ & \underset{\text{[theorem: 11.29]}}{=} & \left(i\underset{n}{\leftrightarrow} j\right)(k) \\ & = & \left(\left(k_1\underset{n}{\leftrightarrow} l_1\right)\circ\ldots\circ\left(k_m\underset{n}{\leftrightarrow} l_m\right)\right)(k) \\ & \underset{\text{[theorem: 11.108]}}{=} & \left(\left(k_1 \underset{n+1}{\leftrightarrow} l_1\right)\circ\ldots\circ\left(k_m\underset{n+1}{\leftrightarrow} l_m\right)\right)_{|\{1,\ldots,n\}}(k) \\ & = & \left(\left(k_1 \underset{n+1}{\leftrightarrow} l_1\right)\circ\ldots\circ\left(k_m\underset{n+1}{\leftrightarrow} l_m\right)\right)(k)\end{array}$$

So in all cases we have $\left(i\underset{n+1}{\leftrightarrow} j\right)(k)=\left(\left(k_1 \underset{n+1}{\leftrightarrow} l_1\right)\circ\ldots\circ\left(k_m\underset{n+1}{\leftrightarrow} l_m\right)\right)(k)$ proving that

$$\left(i\underset{n+1}{\leftrightarrow} j\right)=\left(k_1 \underset{n+1}{\leftrightarrow} l_1\right)\circ\ldots\circ\left(k_m\underset{n+1}{\leftrightarrow} l_m\right)$$

□

Theorem 11.218. *Let $n\in\mathbb{N}\setminus\{1\}$, $i,j\in\{1,\ldots,n\}$ with $i\neq j$ then there exxists a $m\in\mathbb{N}$ and $\{k_r\}_{r\in\{1,\ldots,m\}}\subseteq\{1,\ldots,n-1\}$ such that $\left(i\underset{n}{\leftrightarrow} j\right)=\left(k_1\underset{n}{\leftrightarrow} k_1+1\right)\circ\ldots\circ\left(k_m\underset{n}{\leftrightarrow} k_{m+1}\right)$*

Proof. We prove this by recursion on n so define

$S=\left\{n\in\{2,\ldots,\infty\}|\text{If}\,(i,j)\in\{1,\ldots,n\}\text{ with }i\neq j\text{ then there exist a }\{k_r\}_{r\in\{1,\ldots,m\}}\text{ such that }\left(i\underset{n}{\leftrightarrow} j\right)=\left(k_1\underset{n}{\leftrightarrow} k_1+1\right)\circ\ldots\circ\left(k_m\underset{n}{\leftrightarrow} k_{m+1}\right)\right\}$

then we have:

$\boldsymbol{2\in S.}$ As $\left(i\underset{n}{\leftrightarrow} j\right)\underset{\text{[theorem: 11.213]}}{=}\left(j\underset{n}{\leftrightarrow} i\right)$ we may assume that $i<j$ so we have that $i=1$ and $j=2=i+1$ so that $\left(i\underset{2}{\leftrightarrow} j\right)=\left(1\underset{2}{\leftrightarrow} 2\right)$ proving that $2\in S$.

$\boldsymbol{n\in S\Rightarrow n+1\in S.}$ Let $(i,j)\in\{1,\ldots,n+1\}$ with $i\neq j$. As $\left(i\underset{n+1}{\leftrightarrow} j\right)\underset{\text{[theorem: 11.213]}}{=}\left(j\underset{n+1}{\leftrightarrow} i\right)$ we may assume that $i<j$ so we have the following cases to consideder:

$\boldsymbol{i=n\wedge j=n+1.}$ Then $\left(i\underset{n+1}{\leftrightarrow} j\right)=\left(i\underset{n+1}{\leftrightarrow} 1\right)$ proving that $n+1\in S$ in this case.

$\boldsymbol{i<n\wedge j=n+1.}$ As $n\in S$ there exists $\{k_r\}_{r\in\{1,\ldots,m\}}\subseteq\{1,\ldots,n-1\}$ such that

$$\left(i\underset{n}{\leftrightarrow} n\right)=\left(k_1\underset{n}{\leftrightarrow} k_1+1\right)\circ\ldots\circ\left(k_m\underset{n}{\leftrightarrow} k_m+1\right)$$

Given that $i, n \in \{1, \ldots, n\}$ and $\forall r \in \{1, \ldots, m\}$ we have $k_r, k_r + 1 \in \{1, \ldots, n\}$ it follows from [lemma: 11.217] that

$$\left(i \underset{n+1}{\leftrightarrow} n\right) = \left(k_1 \underset{n+1}{\leftrightarrow} k_1 + 1\right) \circ \ldots \circ \left(k_m \underset{n+1}{\leftrightarrow} k_m + 1\right) \tag{11.110}$$

Consider now

$$\rho = \left(n \underset{n+1}{\leftrightarrow} n+1\right) \circ \left(i \underset{n+1}{\leftrightarrow} n\right) \circ \left(n \underset{n+1}{\leftrightarrow} n+1\right)$$

then given $k \in \{1, \ldots, n+1\}$ we have either:

$\boldsymbol{k = n+1.}$ Then we have

$$\begin{array}{rcl}
\rho(k) & = & \left(n \underset{n+1}{\leftrightarrow} n+1\right)\left(\left(i \underset{n+1}{\leftrightarrow} n\right)\left(\left(n \underset{n+1}{\leftrightarrow} n+1\right)(n+1)\right)\right) \\
 & = & \left(n \underset{n+1}{\leftrightarrow} n+1\right)\left(\left(i \underset{n+1}{\leftrightarrow} n\right)(n)\right) \\
 & = & \left(n \underset{n+1}{\leftrightarrow} n+1\right)(i) \\
 & \underset{i \neq n, n+1}{=} & i \\
 & = & \left(i \underset{n+1}{\leftrightarrow} n+1\right)(n+1) \\
 & \underset{j = n+1 \wedge k = n+1}{=} & \left(i \underset{n+1}{\leftrightarrow} j\right)(k)
\end{array}$$

$\boldsymbol{k = n.}$ Then we have

$$\begin{array}{rcl}
\rho(k) & = & \left(n \underset{n+1}{\leftrightarrow} n+1\right)\left(\left(i \underset{n+1}{\leftrightarrow} n\right)\left(\left(n \underset{n+1}{\leftrightarrow} n+1\right)(n)\right)\right) \\
 & = & \left(n \underset{n+1}{\leftrightarrow} n+1\right)\left(\left(i \underset{n+1}{\leftrightarrow} n\right)(n+1)\right) \\
 & \underset{n+1 \neq i, n}{=} & \left(n \underset{n+1}{\leftrightarrow} n+1\right)(n+1) \\
 & = & n \\
 & \underset{n \neq i, n+1}{=} & \left(i \underset{n+1}{\leftrightarrow} n+1\right)(n) \\
 & \underset{n = k \wedge j = n+1}{=} & \left(i \underset{n+1}{\leftrightarrow} j\right)(k)
\end{array}$$

$\boldsymbol{k = i.}$ Then

$$\begin{array}{rcl}
\rho(k) & = & \left(n \underset{n+1}{\leftrightarrow} n+1\right)\left(\left(i \underset{n+1}{\leftrightarrow} n\right)\left(\left(n \underset{n+1}{\leftrightarrow} n+1\right)(i)\right)\right) \\
 & \underset{i \neq n, n+1}{=} & \left(n \underset{n+1}{\leftrightarrow} n+1\right)\left(\left(i \underset{n+1}{\leftrightarrow} n\right)(i)\right) \\
 & = & \left(n \underset{n+1}{\leftrightarrow} n+1\right)(n) \\
 & = & n+1 \\
 & = & \left(i \underset{n+1}{\leftrightarrow} n+1\right)(i) \\
 & \underset{i = k \wedge j = n+1}{=} & \left(i \underset{n+1}{\leftrightarrow} j\right)(k)
\end{array}$$

$\boldsymbol{k \in \{1, \ldots, n-1\} \setminus \{i\}.}$ Then

$$\begin{array}{rcl}
\rho(k) & = & \left(n \underset{n+1}{\leftrightarrow} n+1\right)\left(\left(i \underset{n+1}{\leftrightarrow} n\right)\left(\left(n \underset{n+1}{\leftrightarrow} n+1\right)(k)\right)\right) \\
 & \underset{k \neq n, n+1}{=} & \left(n \underset{n+1}{\leftrightarrow} n+1\right)\left(\left(i \underset{n+1}{\leftrightarrow} n\right)(k)\right) \\
 & \underset{k \neq i, n}{=} & \left(n \underset{n+1}{\leftrightarrow} n+1\right)(k) \\
 & \underset{k \neq n, n+1}{=} & k \\
 & \underset{k \neq i, n+1}{=} & \left(i \underset{n+1}{\rightarrow} n+1\right)(k) \\
 & \underset{j = n+1}{=} & \left(i \underset{n+1}{\leftrightarrow} j\right)(k)
\end{array}$$

So we have that $\forall k \in \{1, \ldots, n+1\}$ that

$$\left(i \underset{n+1}{\leftrightarrow} j\right)(k) = \rho(k) = \left(\left(n \underset{n+1}{\leftrightarrow} n+1\right) \circ \left(i \underset{n+1}{\leftrightarrow} n\right) \circ \left(n \underset{n+1}{\leftrightarrow} n+1\right)\right)(k)$$

proving that

$$\left(i \underset{n+1}{\leftrightarrow} j\right) = \left(n \underset{n+1}{\leftrightarrow} n+1\right) \circ \left(i \underset{n+1}{\leftrightarrow} n\right) \circ \left(n \underset{n+1}{\leftrightarrow} n+1\right)$$

or combining the above with [eq: 11.110] proves that

$$\left(i \underset{n+1}{\leftrightarrow} j\right) = \left(n \underset{n+1}{\leftrightarrow} n+1\right) \circ \left(k_1 \underset{n+1}{\leftrightarrow} k_1+1\right) \circ \ldots \circ \left(k_m \underset{n+1}{\leftrightarrow} k_m+1\right) \circ \left(n \underset{n+1}{\leftrightarrow} n+1\right)$$

so that $n+1 \in S$ in this case.

$\boldsymbol{i < j \wedge j \in \{1, \ldots, n\}}$**.** Then as $i, j \in \{1, \ldots, n\}$ we have by [theorem: 11.213] that $\left(i \underset{n}{\leftrightarrow} j\right) = \left(i \underset{n+1}{\leftrightarrow} j\right)_{|\{1,\ldots,n\}}$. As $n \in S$ it follows that there exist a $\{k_r\}_{r \in \{1,\ldots,m\}} \subseteq \{1, \ldots, n-1\}$ such that

$$\left(i \underset{n}{\leftrightarrow} j\right) = \left(k_1 \underset{n}{\leftrightarrow} k_1+1\right) \circ \ldots \circ \left(k_m \underset{n}{\leftrightarrow} k_m+1\right)$$

Given that $i, j \in \{1, \ldots, n\}$ and $\forall r \in \{1, \ldots, m\}$ we have $k_r, k_r+1 \in \{1, \ldots, n\}$ it follows from [lemma: 11.217] that

$$\left(i \underset{n+1}{\leftrightarrow} j\right) = \left(k_1 \underset{n+1}{\leftrightarrow} k_1+1\right) \circ \ldots \circ \left(k_m \underset{n+1}{\leftrightarrow} k_m+1\right)$$

which proves that $n+1 \in S$ in this case.

So in all possible cases we have $n+1 \in S$ which completes the induction argument. □

Permutations are typically used to permute the arguments of functions with several arguments. Remember that X^n is defined as $X^n \underset{[\text{definition: } 6.81]}{=} A^{\{1,\ldots,n\}}$ so that the following definition makes sense.

Definition 11.219. *Let X, Y be sets, $n \in \mathbb{N}$ and $f: X^n \to Y$ a function and $\sigma \in P_n$ then*

$$\sigma f: X^n \to Y$$

is defined by

$$(\sigma f)(x) = f(x \circ \sigma)$$

or using [notation: 6.75]

$$\sigma f(x) = f(x_{\sigma(1)}, \ldots, x_{\sigma(n)})$$

Theorem 11.220. *Let X be a set, $\langle Y, +, \cdot\rangle$ a vector space over a field $\langle F, +, \cdot\rangle$, $\sigma \in P_n$, then we have for the vector space $\langle Y^X, +, \cdot\rangle$ [see example: 11.64] that:*

1. *$\forall f, g \in Y^X$ we have $\sigma(f+g) = \sigma f + \sigma g$*
2. *$\forall f \in Y^X$, $\alpha \in F$ we have $\sigma(\alpha \cdot f) = \alpha \cdot \sigma f$*

Proof.

1. Let $x \in X$ then

$$(\sigma(f+g))(x) = (f+g)(x \circ \sigma) = f(x \circ \sigma) + g(x \circ \sigma) = (\sigma f)(x) + (\sigma g)(x) = (\sigma f + \sigma g)(x)$$

proving that

$$\sigma(f+g) = \sigma f + \sigma g$$

2. Let $x \in X$ then

$$(\sigma(\alpha \cdot f))(x) = (\alpha \cdot f)(x \circ \sigma) = \alpha \cdot f(x \circ \sigma) = \alpha \cdot ((\sigma f)(x)) = (\alpha \cdot \sigma f)(x)$$

proving that

$$\sigma\,(\alpha \cdot f) = \alpha \cdot \sigma f$$

□

Theorem 11.221. *Let X, Y be sets, $n \in \mathbb{N}$ and $f: X^n \to Y$ a function then*

$$\mathrm{Id}_{\{1,\ldots,n\}} f = f$$

Proof. $\forall x \in X^n$ we have $(\mathrm{Id}_{\{1,\ldots,n\}} f)(x) = f(x \circ \mathrm{Id}_{\{1,\ldots,n\}}) = f(x)$ proving that

$$\mathrm{Id}_{\{1,\ldots,n\}} f = f. \qquad \square$$

Theorem 11.222. *Let X, Y be sets, $n \in \mathbb{N}$ and $f: X^n \rightarrow Y$ a function and $\sigma, \rho \in P_n$ then*

$$\sigma(\rho f) = (\sigma \circ \rho) f$$

Proof. $\forall x \in X^n$ we have

$$\sigma(\rho f)(x) = (\rho f)(x \circ \sigma) = f((x \circ \sigma) \circ \rho) = f(x \circ (\sigma \circ \rho)) = (\sigma \circ \rho) f(x)$$

so that $\sigma(\rho f) = (\sigma \circ \rho) f$. $\square$

Theorem 11.223. *Let $n \in \mathbb{N}$, $\sigma \in P_n$, X a set, $\langle Y, +, \cdot \rangle$ a ring and $f, g: X^n \rightarrow Y$ then we have using point wise product and adding of functions that*

1. *$\forall \alpha \in Y$ $\sigma(\alpha \cdot f) = \alpha \cdot (\sigma f)$*
2. *$\sigma(f + g) = \sigma f + \sigma g$*

Proof.

1. Let $x \in X^n$ then $(\sigma(\alpha \cdot f))(x) = (\alpha \cdot f)(x \circ \sigma) \underset{\text{def}}{=} \alpha \cdot (f(x \circ \sigma)) = \alpha \cdot (\sigma f)(x)$ proving that

$$\sigma(\alpha \cdot f) = \alpha \cdot (\sigma f)$$

2. Let $x \in X^n$ then

$$(\sigma(f + g))(x) = (f + g)(x \circ \sigma) \underset{\text{def}}{=} f(x \circ \sigma) = g(x \circ \sigma) = (\sigma f)(x) + (\sigma g)(x) = (\sigma f + \sigma g)(x)$$

   proving that

$$\sigma(f + g) = \sigma f + \sigma g \qquad \square$$

Theorem 11.224. *Let X, Y be sets, $n \in \mathbb{N}$ and $f: X^n \rightarrow Y$ a function such that $\forall i, j \in \{1, \ldots, n\}$ with $i \neq j$ we have $\left(i \underset{n}{\leftrightarrow} j\right) f = f$ then if $l \in \mathbb{N}$ and $\left\{\left(i_k \underset{n}{\leftrightarrow} j_k\right)\right\}_{k \in \{1, \ldots, l\}} \subseteq P_n$ we have*

$$\left(\left(i_1 \underset{n}{\leftrightarrow} j_1\right) \circ \cdots \circ \left(i_l \underset{n}{\leftrightarrow} j_l\right)\right) f = f$$

Proof. We prove this by induction on l, so take:

$S = \left\{l \in \mathbb{N} | \text{If } f: X^n \rightarrow Y \text{ satisfies } \forall i, j \in \{1, \ldots, n\} \text{ with } i \neq j \text{ we have } \left(i \underset{n}{\leftrightarrow} j\right) f = f \text{ then for every } \left\{\left(i_k \underset{n}{\leftrightarrow} j_k\right)\right\}_{k \in \{1, \ldots, l\}} \subseteq P_n \text{ we have } \left(\left(i_1 \underset{n}{\leftrightarrow} j_1\right) \circ \cdots \circ \left(i_l \underset{n}{\leftrightarrow} j_l\right)\right) f = f\right\}$

then we have:

$1 \in S$. Let $f: X^n \rightarrow Y$ satisfies $\forall i, j \in \{1, \ldots, n\}$ $\left(i \underset{n}{\leftrightarrow} j\right) f = f$ and $\left\{\left(i_k \underset{n}{\leftrightarrow} j_k\right)\right\}_{k \in \{1, \ldots, 1\}} \subseteq P_n$ then we have

$$\left(\left(i_1 \underset{n}{\leftrightarrow} j_1\right) \circ \cdots \circ \left(i_l \underset{n}{\leftrightarrow} j_l\right)\right) f \underset{[\text{theorem: } 11.211]}{=} \left(i_1 \underset{n}{\leftrightarrow} j_1\right) f \tag{11.111}$$

For i_1, j_1 we have either:

$i_1 = j_1$. Then
$\left(i_1 \underset{n}{\leftrightarrow} j_1\right) f \underset{[\text{theorem: } 11.211]}{=} \mathrm{Id}_{\{1,, \ldots, n\}} f \underset{[\text{theorem: } 11.221]}{=} f$

$i_1 \neq j_1$. Then by the hypothesis we have $\left(i_1 \underset{n}{\leftrightarrow} j_1\right) f = f$

combining this with [eq: 11.111] proves that $\left(\left(i_1 \underset{n}{\leftrightarrow} j_1\right) \circ \cdots \circ \left(i_l \underset{n}{\leftrightarrow} j_l\right)\right) f = f$ hence $1 \in S$.

$l \in S \Rightarrow l + 1 \in S$. If $f: X^n \rightarrow Y$ satisfies $\forall i, j \in \{1, \ldots, n\}$ $\left(i \underset{n}{\leftrightarrow} j\right) f = f$ and $\left\{\left(i_k \underset{n}{\leftrightarrow} j_k\right)\right\}_{k \in \{1, \ldots, l+1\}} \subseteq P_n$. For i_{l+1}, j_{l+1} we have either:

$i_{l+1} = j_{l+1}$. Then
$\left(i_{l+1} \underset{n}{\leftrightarrow} j_{l+1}\right) f \underset{[\text{theorem: } 11.211]}{=} \mathrm{Id}_{\{1,, \ldots, n\}} f \underset{[\text{theorem: } 11.221]}{=} f$

$i_{l+1} \neq j_{l+1}$. Then by the hypothesis we have $\left(i_{l+1} \underset{n}{\leftrightarrow} j_{l+1}\right) f = f$

so in all cases we have

$$\left(i_{l+1} \underset{n}{\leftrightarrow} j_{i+1}\right) f = f \tag{11.112}$$

Further

$$\begin{aligned}
\left(\left(i_1 \underset{n}{\leftrightarrow} j_1\right) \circ \cdots \circ \left(i_{l+1} \underset{n}{\leftrightarrow} j_{l+1}\right)\right) f & \underset{\text{[eq: 11.211]}}{=} \\
\left(\left(\left(i_1 \underset{n}{\leftrightarrow} j_1\right) \circ \cdots \circ \left(i_l \underset{n}{\leftrightarrow} j_l\right)\right) \circ \left(i_{l+1} \underset{n}{\leftrightarrow} j_{l+1}\right)\right) f & \underset{\text{[theorem: 11.222]}}{=} \\
\left(\left(i_1 \underset{n}{\leftrightarrow} j_1\right) \circ \cdots \circ \left(i_l \underset{n}{\leftrightarrow} j_l\right)\right)\left(\left(i_{l+1} \underset{n}{\leftrightarrow} j_{l+1}\right) f\right) & \underset{\text{[eq: 11.112]}}{=} \\
\left(\left(i_1 \underset{n}{\leftrightarrow} j_1\right) \circ \cdots \circ \left(i_l \underset{n}{\leftrightarrow} j_l\right)\right) f & \underset{n \in S}{=} \\
f &
\end{aligned}$$

proving that $l+1 \in S$. □

We can extend the above theorem to general permutation's.

Corollary 11.225. *Let X, Y be sets, $n \in \mathbb{N}$ and $f: X^n \rightarrow Y$ a function such that $\forall i, j \in \{1, \ldots, n\}$ with $i \neq j$ we have $\left(i \underset{n}{\leftrightarrow} j\right) f = f$ then $\forall \sigma \in P_n$ we have $\sigma f = f$.*

Proof. If $n = 1$ then $\sigma = \mathrm{Id}_{\{1\}}$ so that $\sigma f = \mathrm{Id}_{\{1\}} f \underset{\text{[theorem: 11.221]}}{=} f$. To complete the proof we have to look at the case where $n \in \mathbb{N} \setminus \{1\}$. Using [theorem: 11.216] there exists a $\left\{\left(i_k \underset{n}{\leftrightarrow} j_k\right)\right\}_{k \in \{1, \ldots, l\}} \subseteq P_n$ such that $\sigma = \left(\left(i_1 \underset{n}{\leftrightarrow} j_1\right) \circ \cdots \circ \left(i_l \underset{n}{\leftrightarrow} j_l\right)\right)$. Further

$$\sigma f = \left(\left(i_1 \underset{n}{\leftrightarrow} j_1\right) \circ \cdots \circ \left(i_l \underset{n}{\leftrightarrow} j_l\right)\right) f \underset{\text{[theorem: 11.224]}}{=} f$$

proving the corollary. □

The above shows that in order to prove that a multi parameter function is stable under a permutation we only have to show that it is stable under arbitrary transpositions. Actually we can weaken this condition. A multi parameter function is stable under a permutation if it is stable under transpositions of the form $\left(i \underset{n}{\leftrightarrow} i+1\right)$.

Corollary 11.226. *Let X, Y be sets, $n \in \mathbb{N}$ and $f: X^n \rightarrow Y$ a function such that $\forall i \in \{1, \ldots, n-1\}$ we have $\left(i \underset{n}{\leftrightarrow} i+1\right) f = f$ then $\forall \sigma \in P_n$ we have $\sigma f = f$.*

Proof. If $n = 1$ then $\sigma = \mathrm{Id}_{\{1\}}$ so that $\sigma f = \mathrm{Id}_{\{1\}} f \underset{\text{[theorem: 11.221]}}{=} f$. To complete the proof we have to look at the case where $n \in \mathbb{N} \setminus \{1\}$. Let $i, j \in \{1, \ldots, n\}$ with $i \neq j$ then by [theorem: 11.218] there exist a $\{k_r\}_{r \in \{1, \ldots, m\}} \subseteq \{1, \ldots, n-1\}$ such that $\left(i \underset{n}{\leftrightarrow} j\right) = \left(k_1 \underset{n}{\leftrightarrow} k_1 + 1\right) \circ \ldots \circ \left(k_m \underset{n}{\leftrightarrow} k_m + 1\right)$. By the hypothese we have $\forall r \in \{1, \ldots, n-1\}$ $\left(k_r \underset{n}{\leftrightarrow} k_r + 1\right) f = f$, hence using [theorem: 11.224] it follows that $\left(\left(k_1 \underset{n}{\leftrightarrow} k_1 + 1\right) \circ \ldots \circ \left(k_m \underset{n}{\leftrightarrow} k_m + 1\right)\right) f = f$. So as $\left(i \underset{n}{\leftrightarrow} j\right) = \left(k_1 \underset{n}{\leftrightarrow} k_1 + 1\right) \circ \ldots \circ \left(k_m \underset{n}{\leftrightarrow} k_m + 1\right)$ it follows that $\left(i \underset{n}{\leftrightarrow} j\right) f = f$. As $i, j \in \{1, \ldots, n\}$ was chosen arbitrary we can use the previous corollary [corollary: 11.225] to prove $\sigma f = f$ for every $\sigma \in P_n$. □

11.5.2 Sign of a Permutation

In the following definition we will use the multiplicative Abelian semi-group $\langle \mathbb{Z}, \cdot \rangle$, so we use the symbol $\prod$ instead of $\sum$ [see remark: 11.2]. Further let $n \in \mathbb{N}_0$ then

$$\{(i, j) \in \{1, \ldots, n\} \times \{1, \ldots, n\} | i < j\} \subseteq \{1, \ldots, n\} \times \{1, \ldots, n\} \text{ a finite set [see theorems: 10.85,10.90]}$$

so that by [theorem: 6.42] $\{(i, j) \in \{1, \ldots, n\} \times \{1, \ldots, n\} | i < j\}$ is a finite set. Hence the following definition makes sense.

Definition 11.227. *Let $n \in \mathbb{N}$ then we define*

$$\Phi_n: \mathbb{Z}^n \rightarrow \mathbb{Z} \text{ by } \Phi_n(x) = \prod_{(i,j) \in \{(i,j) \in \{1, \ldots, n\} \times \{1, \ldots, n\} | i < j\}} (x_i - x_j)$$

Example 11.228.

1. $\Phi_1(x) = 1$, as $\{(i, j) \in \{1\} \times \{1\} | i < j\} = \varnothing$ we have $\Phi_1(x) = \prod_{(i,j) \in \varnothing} (x_i - x_j) = 1$ [the neutral element of $\langle \mathbb{Z}, \cdot \rangle$].

2. If $x = (x_1, x_2) \in \mathbb{Z}^2$ then $\Phi_2(x) = x_2 - x_1$ because $\{(i, j) \in \{1, 2\} \times \{1, 2\} | i < j\} = \{(1, 2)\}$ so that

$$\Phi_2(x) = \prod_{(i,j) \in \{(1,2)\}} (x_i - x_j) = x_1 - x_2$$

3. If $x = (x_1, x_2, x_3)$ then $\Phi_3(x) = (x_1 - x_2) \cdot (x_1 - x_3) \cdot (x_2 - x_3)$ because

$$\{(i, j) \in \{1, 2, 3\} \times \{1, 2, 3\} | i < j\} = \{(1, 2), (1, 3), (2, 3)\}$$

so that

$$\Phi_3(x) = \prod_{(i,j) \in \{(1,2),(1,3),(2,3)\}} (x_i - x_j) = (x_1 - x_2) \cdot (x_1 - x_3) \cdot (x_2 - x_3)$$

4. …

Lemma 11.229. *Let $n \in \mathbb{N}$ then for $(x_1, \ldots, x_n) \in \mathbb{Z}^n$ defined by $x_i = i$ we have $\Phi_n(x_1, \ldots, x_n) \neq 0$, in other words $\Phi_n(1, \ldots, n) \neq 0$.*

Proof. If we define $(x_1, \ldots, x_n)$ by $x_i = i$ then $\forall (i, j) \in \{(i, j) \in \{1, \ldots, n\} \times \{1, \ldots, m\} | i < j\}$ we have $x_i - x_j = i - j < 0$, so that $\{x_i - x_j\}_{(i,j) \in \{(i,j) \in \{1,\ldots,n\} \times \{1,\ldots,m\} | i<j\}} \subseteq \mathbb{R} \setminus \{0\}$. Using [theorem: 11.50] it follows then that $\Phi_n(x_1, \ldots, x_n) = \prod_{(i,j) \in \{(i,j) \in \{1,\ldots,n\} \times \{1,\ldots,m\} | i<j\}} (x_i - x_j) \neq 0$. □

Lemma 11.230. *Let $n \in \mathbb{N}_0 \setminus \{1\} = \{2, \ldots, \infty\}$ then if $k, l \in \{1, \ldots, n\}$ with $k \neq l$ we have for the strict transposition $\left(k \underset{n}{\leftrightarrow} l\right)$ that*

$$\left(k \underset{n}{\leftrightarrow} l\right) \Phi_n = (-1) \cdot \Phi_n$$

Proof. First we may always assume that $k < l$ [otherwise exchange k and l]. Define

$$I = \{(i, j) \in \{1, \ldots, n\} \times \{1, \ldots, n\} | i < j\}$$

Then we have the following excluding possibilities for $(i, j) \in I$

- $\boldsymbol{i = k}$**.** Then as $k < l$ and $i < j$ we have for j either:
  - $\boldsymbol{j < l}$**.** Then $(i, j) \in J_0$ where $J_0 = \{(k, j) \in I | k < j < l\} \subseteq I$
  - $\boldsymbol{j = l}$**.** Then $(i, j) \in J_1$ where $J_1 = \{(k, l)\} \subseteq I$
  - $\boldsymbol{l < j}$**.** Then $(i, j) \in J_2$ where $J_2 = \{(k, j) \in I | l < j\} \subseteq I$
- $\boldsymbol{i \neq k}$**.** Then we have for i either:
  - $\boldsymbol{i = l}$**.** Then as $i < j$ we have $(i, j) \in J_3 = \{(l, j) \in I | l < j\} \subseteq I$
  - $\boldsymbol{i \neq l}$**.** Then we have for j either:
    - $\boldsymbol{j = k}$**.** Then as $i < j \Rightarrow i < k$ we have $(i, j) \in J_4$ where $J_4 = \{(i, k) \in I | i < k\} \subseteq I$
    - $\boldsymbol{j = l}$**.** Then as $i = k$ we have either:
      - $\boldsymbol{i < k}$**.** Then $(i, j) \in J_5$ where $J_5 = \{(i, l) \in I | i < k\} \subseteq I$
      - $\boldsymbol{k < i}$**.** Then $(i, j) \in J_6$ where $J_6 = \{(i, l) \in I | k < i < l\} \subseteq I$
    - $\boldsymbol{j \neq l, k}$**.** Then $(i, j) \in J_7$ where $J_7 = \{(i, j) \in I | i \neq k, l \wedge j \neq k, l\} \subseteq I$

So

$$I = J_0 \bigcup J_1 \bigcup J_2 \bigcup J_3 \bigcup J_4 \bigcup J_5 \bigcup J_6 \bigcup J_7$$

Further for the possible intersections of $J_0, J_1, J_2, J_3, J_4, J_5, J_6$ and J_7 [using commutativity] we have

- $\boldsymbol{(i, j) \in J_0 \bigcap J_1}$**.** Then $i = k < j < l \wedge j = l$ a contradiction.
- $\boldsymbol{(i, j) \in J_0 \bigcap J_2}$**.** Then $i = k < j < l \wedge l < j$ a contradiction.
- $\boldsymbol{(i, j) \in J_0 \bigcap J_3}$**.** Then $i = k < j < l \wedge l < j$ a contradiction.
- $\boldsymbol{(i, j) \in J_0 \bigcap J_4}$**.** Then $i = k < j < l \wedge j = k$ a contradiction.
- $\boldsymbol{(i, j) \in J_0 \bigcap J_5}$**.** Then $i = k < j < l \wedge j = l$ a contradiction.
- $\boldsymbol{(i, j) \in J_0 \bigcap J_6}$**.** Then $i = k < j < l \wedge j = l$ a contradiction.

$(i,j)\in J_0\bigcap J_7$. Then $i=k<j<l\wedge i\neq k$ a contradiction.

$(i,j)\in J_1\bigcap J_2$. Then $i=k\wedge j=l\wedge l<j$ a contradiction

$(i,j)\in J_1\bigcap J_3$. Then $i=k\wedge j=l\wedge l<j$ a contradiction.

$(i,j)\in J_1\bigcap J_4$. Then $i=k\wedge j=l\wedge i<k$ a contradiction.

$(i,j)\in J_1\bigcap J_5$. Then $i=k\wedge j=l\wedge i<k$ a contradiction.

$(i,j)\in J_1\bigcap J_6$. Then $i=k\wedge j=l\wedge k<i$ a contradiction

$(i,j)\in J_1\bigcap J_7$. Then $i=k\wedge j=l\wedge i\neq k$ a contradiction.

$(i,j)\in J_2\bigcap J_3$. Then $i=k\wedge l<j\wedge i=l$ contradicting $k<l$.

$(i,j)\in J_2\bigcap J_4$. Then $i=k\wedge l<j\wedge j=k$ contradicting $i<j$.

$(i,j)\in J_2\bigcap J_5$. Then $i=k\wedge l<j\wedge i<k$ a contradiction.

$(i,j)\in J_2\bigcap J_6$. Then $i=k\wedge l<j\wedge k<i$ a contradiction.

$(i,j)\in J_2\bigcap J_7$. Then $i=k\wedge l<j\wedge i\neq k$ a contradiction.

$(i,j)\in J_3\bigcap J_4$. Then $i=l\wedge l<j\wedge i<k$ giving $l<k$ contradicting $k<l$.

$(i,j)\in J_3\bigcap J_5$. Then $i=l\wedge l<j\wedge i<k$ giving $l<k$ contradicting $k<l$.

$(i,j)\in J_3\bigcap J_6$. Then $i=l\wedge l<j\wedge l=j$ a contradicting.

$(i,j)\in J_3\bigcap J_7$. Then $i=l\wedge l<j\wedge i\neq l$ a contradiction.

$(i,j)\in J_4\bigcap J_5$. Then $i<k=j\wedge j=l$ contradicting $k<l$.

$(i,j)\in J_4\bigcap J_6$. Then $i<k=j\wedge j=l$ contradicting $k<l$.

$(i,j)\in J_4\bigcap J_7$. Then $i<k=j\wedge j\neq k$ a contradiction.

$(i,j)\in J_5\bigcap J_6$. Then $i<k\wedge j=l\wedge k<i$ a contradiction.

$(i,j)\in J_5\bigcap J_7$. Then $i<k<i\wedge j=l\wedge j\neq l$ a contradiction.

$(i,j)\in J_6\bigcap J_7$. Then $k<i\wedge j=l\wedge j\neq l$ a contradiction.

So

$$I=\bigcup_{i\in\{1,\ldots,7\}} J_i \text{ and } \forall i,j\in\{1,\ldots,7\} \text{ with } i\neq j \text{ we have } J_i\bigcap J_j=\varnothing$$

So that

$$\begin{aligned}
\Phi_n(x_1,\ldots,x_n) &= \prod_{(i,j)\in I}(x_i-x_j)\\
&\underset{[\text{theorem: } 11.43]}{=} \prod_{m=0}^{7}\left(\prod_{(i,j)\in J_m}(x_i-x_j)\right)\\
&= \prod_{m=0}^{7} Q_m
\end{aligned}\tag{11.113}$$

where

$$Q_m=\prod_{(i,j)\in J_m}(x_i-x_j)\tag{11.114}$$

and

$$\begin{aligned}
\left(\left(k\underset{n}{\leftrightarrow}l\right)\Phi_n\right)(x_1,\ldots,x_n) &= \prod_{(i,j)\in I}\left(x_{(k\underset{n}{\leftrightarrow}l)(i)}-x_{(k\underset{n}{\leftrightarrow}l)(j)}\right)\\
&\underset{[\text{theorem: } 11.43]}{=} \prod_{m=0}^{7}\left(\prod_{(i,j)\in J_m}\left(x_{(k\underset{n}{\leftrightarrow}l)(i)}-x_{(k\underset{n}{\leftrightarrow}l)(j)}\right)\right)\\
&= \prod_{m=0}^{7} R_m
\end{aligned}\tag{11.115}$$

where

$$R_m = \prod_{(i,j)\in J_m} \left(x_{(k \underset{n}{\leftrightarrow} l)(i)} - x_{(k \underset{n}{\leftrightarrow} l)(j)}\right) \tag{11.116}$$

Now we have:

***R*₀.**

$$\begin{aligned} R_0 &= \prod_{(i,j)\in J_0} \left(x_{(k \underset{n}{\leftrightarrow} l)(i)} - x_{(k \underset{n}{\leftrightarrow} l)(j)}\right) \\ &= \prod_{(i,j)\in\{(k,j)\in I \mid k<j<l\}} \left(x_{(k \underset{n}{\leftrightarrow} l)(i)} - x_{(k \underset{n}{\leftrightarrow} l)(j)}\right) \\ &= \prod_{(i,j)\in\{(k,j)\in I \mid k<j<l\}} (x_l - x_j) \\ &= \prod_{(i,j)\in J_0} (x_l - x_j) \end{aligned} \tag{11.117}$$

Define now

$$\beta: J_6 = \{(i,l)\in I \mid k<i<l\} \rightarrow J_0 = \{(k,j)\in I \mid k<j<l\} \text{ by } \beta(i,j) = (k,i)$$

then we have:

injectivity. Let $(i,j),(i',j')\in J_6$ with $\beta(i,j)=\beta(i',j')$ then $j=l=j'$ and $(k,i)=(k,i')\Rightarrow i=i'$ proving that $(i,j)=(i',j')$.

surjectivity. Let $(r,s)\in J_0$ then $k=r$ and $k<s<l$ so that $(s,s)\in J_7$ which, as $\beta(s,s)=(k,s)=(r,s)$, proves surjectivity.

proving that

$$\beta: J_6 \rightarrow J_0 \text{ is a bijection}$$

Next

$$\begin{aligned} \prod_{(i,j)\in J_0} (x_l - x_j) &= \prod_{(i,j)\in J_0} (x_l - x_{(i,j)_2}) \\ &\underset{\text{[theorem: 11.36]}}{=} \prod_{(i,j)\in J_6} (x_l - x_{(\beta(i,j))_2}) \\ &= \prod_{(i,j)\in J_6} (x_l - x_i) \\ &= \prod_{(i,j)\in J_6} ((-1)\cdot(x_i - x_l)) \\ &\underset{\text{[theorem: 11.38]}}{=} \left(\prod_{(i,j)\in J_6} (-1)\right)\cdot \prod_{(i,j)\in J_6} (x_i - x_l) \\ &\underset{\text{[theorem: 11.48]}}{=} (-1)^{\text{card}(J_6)}\cdot \prod_{(i,j)\in\{(i,l)\in I \mid k<i<l\}} (x_i - x_l) \\ &= (-1)^{\text{card}(J_6)}\cdot \prod_{(i,j)\in\{(i,l)\in I \mid k<i<l\}} (x_i - x_j) \\ &= (-1)^{\text{card}(j_6)}\cdot \prod_{(i,j)\in J_6} (x_i - x_j) \\ &= (-1)^{\text{card}(j_6)}\cdot Q_6 \end{aligned}$$

combining the above with [eq: 11.117] gives

$$R_0 = (-1)^{\text{card}(j_6)}\cdot Q_6 \tag{11.118}$$

***R*₁.** Then

$$\begin{aligned}
R_1 &= \prod_{(i,j)\in J_1} \left(x_{(k \underset{n}{\leftrightarrow} l)(i)} - x_{(k \underset{n}{\leftrightarrow} l)(j)}\right) \\
&= \prod_{(i,j)\in\{(k,l)\}} \left(x_{(k \underset{n}{\leftrightarrow} l)(i)} - x_{(k \underset{n}{\leftrightarrow} l)(j)}\right) \\
&\underset{[\text{theorem: } 11.34}{=} x_{(k \underset{n}{\leftrightarrow} l)(k)} - x_{(k \underset{n}{\leftrightarrow} l)(l)} \\
&= x_l - x_k \\
&= (-1)\cdot(x_k - x_l) \\
&\underset{[\text{theorem: } 11.34}{=} (-1)\cdot \prod_{(i,j)\in\{(k,l)\}} (x_i - x_j) \\
&= (-1)\cdot \prod_{(i,j)\in J_1} (x_i - x_j) \\
&= (-1)\cdot Q_1
\end{aligned}$$

proving that

$$R_1 = (-1)\cdot Q_1 \tag{11.119}$$

***R*₂.** Then

$$\begin{aligned}
R_3 &= \prod_{(i,j)\in J_2} \left(x_{(k \underset{n}{\leftrightarrow} l)(i)} - x_{(k \underset{n}{\leftrightarrow} l)(j)}\right) \\
&= \prod_{(i,j)\in\{(k,j)\in I|l<j\}} \left(x_{(k \underset{n}{\leftrightarrow} l)(i)} - x_{(k \underset{n}{\leftrightarrow} l)(j)}\right) \\
&= \prod_{(i,j)\in\{(k,j)\in I|l<j\}} (x_l - x_j) \\
&= \prod_{(i,j)\in J_2} (x_l - x_j)
\end{aligned} \tag{11.120}$$

Define

$$\beta\colon J_3 = \{(l,j)\in I|l<j\} \rightarrow J_2 = \{(k,j)\in I|l<j\} \text{ by } \beta(i,j) = (k,j)$$

then we have:

injectivity. Let $(i,j),(i',j')\in J_3$ with $\beta(i,j)=\beta(i',j')$ then $i=l=i'$ and $(k,j)=(k,j')\Rightarrow j=j'$ so that $(i,j)=(i',j')$.

surjectivity. Let $(r,s)\in J_2$ then $r=k$ and $l<s$ so that $(l,s)\in J_3$ which, as $\beta(l,s)=(k,s)=(r,s)$, proves surjectivity.

Hence

$$\beta\colon J_3 \rightarrow J_2 \text{ is a bijection}$$

Next

$$\begin{aligned}
\prod_{(i,j)\in J_2} (x_l - x_j) &= \prod_{(i,j)\in J_2} (x_l - x_{(i,j)_2}) \\
&\underset{[\text{theorem: } 11.36]}{=} \prod_{(i,j)\in J_3} (x_l - x_{(\beta(i,j))_2}) \\
&= \prod_{(i,j)\in J_3} (x_l - x_j) \\
&= \prod_{(i,j)\in\{(l,j)\in I|l<j\}} (x_l - x_j) \\
&= \prod_{(i,j)\in\{(l,j)\in I|l<j\}} (x_i - x_j) \\
&= \prod_{(i,j)\in J_3} (x_i - x_j) \\
&= Q_3
\end{aligned}$$

which combined with [eq: 11.120] proves that

$$R_2 = Q_3 \tag{11.121}$$

***R*₃.** We have

$$\begin{aligned}
R_3 &= \prod_{(i,j)\in J_3} \left(x_{(k\underset{n}{\leftrightarrow}l)(i)} - x_{(k\underset{n}{\leftrightarrow}l)(j)}\right)\\
&= \prod_{(i,j)\in\{(l,j)\in I|l<j\}} \left(x_{(k\underset{n}{\leftrightarrow}l)(i)} - x_{(k\underset{n}{\leftrightarrow}l)(j)}\right)\\
&= \prod_{(i,j)\in\{(k,j)\in I|l<j\}} (x_k - x_j)\\
&= \prod_{(i,j)\in J_3} (x_k - x_j)
\end{aligned} \tag{11.122}$$

Define

$$\beta: J_2 = \{(k,j)\in I|l<j\} \rightarrow J_3 = \{(l,j)\in I|l<j\} \text{ by } \beta(i,j) = (l,j)$$

then we have

injectivity. Let $(i,j),(i',j')\in J_2$ with $\beta(i,j)=\beta(i',j')$ then $i=k=i'$ and $(l,j)=(l,j')\Rightarrow j=j'$ proving $(i,j)=(i',j')$.

surjectivity. Let $(r,s)\in J_3$ then $r=l$ and $l<s$, then $(k,s)\in J_3$ and $\beta(k,s)=(l,s)=(r,s)$ proving surjectivity.

hence we have

$$\beta: J_2 \rightarrow J_3 \text{ is a bijection}$$

Next

$$\begin{aligned}
\prod_{(i,j)\in J_3} (x_k - x_j) &= \prod_{(i,j)\in J_3} (x_k - x_{(i,j)_2})\\
&\underset{[\text{theorem: } 11.36]}{=} \prod_{(i,j)\in J_2} (x_k - x_{(\beta(i,j))_2})\\
&= \prod_{(i,j)\in J_2} (x_k - x_{(l,j)_2})\\
&= \prod_{(i,j)\in J_2} (x_k - x_j)\\
&= \prod_{(i,j)\in\{(k,j)\in I|l<j\}} (x_k - x_j)\\
&= \prod_{(i,j)\in\{(k,j)\in I|l<j\}} (x_i - x_j)\\
&= \prod_{(i,j)\in J_2} (x_i - x_j)\\
&= Q_2
\end{aligned}$$

proving together with [eq: 11.122] that

$$R_3 = Q_2 \tag{11.123}$$

***R*₄.** We have

$$\begin{aligned}
R_4 &= \prod_{(i,j)\in J_4} \left(x_{(k\underset{n}{\leftrightarrow}l)(i)} - x_{(k\underset{n}{\leftrightarrow}l)(j)}\right)\\
&= \prod_{(i,j)\in\{(i,k)\in I|i<k\}} \left(x_{(k\underset{n}{\leftrightarrow}l)(i)} - x_{(k\underset{n}{\leftrightarrow}l)(j)}\right)\\
&= \prod_{(i,j)\in\{(i,k)\in I|i<k\}} (x_i - x_l)\\
&= \prod_{(i,j)\in J_4} (x_i - x_l)
\end{aligned} \tag{11.124}$$

Define

$$\beta: J_5 = \{(i,l)\in I|i<k\} \rightarrow J_4 = \{(i,k)\in I|i<k\} \text{ by } \beta(i,j) = (i,k)$$

then we have:

injectivity. Let $(i,j),(i',j')\in J_5$ with $\beta(i,j)=\beta(i',j')$ then $j=l=j'$ and $(i,k)=(i',k)\Rightarrow i=i'$ so that $(i,j)=(i',j')$.

surjectivity. Let $(r,s)\in J_4$ then $s=k$ and $r<k$ then for (r,l) we have $(r,l)\in J_5$ and $\beta(r,l)=(r,k)=(r,s)$ proving surjectivity.

So

$$\begin{aligned}\prod_{(i,j)\in J_4}(x_i-x_l) &= \prod_{(i,j)\in J_4}(x_{(i,j)_1}-x_l)\\ &\underset{[\text{theorem: } 11.36]}{=} \prod_{(i,j)\in J_5}(x_{(\beta(i,j))_1}-x_l)\\ &= \prod_{(i,j)\in J_5}(x_{(i,k)_1}-x_l)\\ &= \prod_{(i,j)\in J_5}(x_i-x_l)\\ &= \prod_{(i,j)\in\{(i,l)\in I|i<k\}}(x_i-x_l)\\ &= \prod_{(i,j)\in\{(i,l)\in I|i<k\}}(x_i-x_j)\\ &= \prod_{(i,j)\in J_5}(x_i-x_j)\\ &= Q_5\end{aligned}$$

combining this with [eq: 11.124] gives

$$R_4=Q_5 \tag{11.125}$$

***R*$_5$.** We have

$$\begin{aligned}R_5 &= \prod_{(i,j)\in J_5}\left(x_{(k\underset{n}{\leftrightarrow}l)(i)}-x_{(k\underset{n}{\leftrightarrow}l)(j)}\right)\\ &= \prod_{(i,j)\in\{(i,l)\in I|i<k\}}\left(x_{(k\underset{n}{\leftrightarrow}l)(i)}-x_{(k\underset{n}{\leftrightarrow}l)(j)}\right)\\ &= \prod_{(i,j)\in\{(i,l)\in I|i<k\}}(x_i-x_k)\\ &= \prod_{(i,j)\in J_5}(x_i-x_k)\end{aligned} \tag{11.126}$$

Define

$$\beta: J_4=\{(i,k)\in I|i<k\}\rightarrow J_5=\{(i,l)\in I|i<k\} \text{ by } \beta(i,j)=(i,l)$$

then we have:

injectivity. Let $(i,j),(i',j')\in J_4$ with $\beta(i,j)=\beta(i',j')$ then $j=k=j'$ and $(i,l)=(i',l)\Rightarrow i=i'$ proving that $(i,j)=(i',j')$.

surjectivity. Let $(r,s)\in J_5$ then $s=l$ and $r<k$ then for (r,k) we have $(r,k)\in J_4$ and $\beta(r,k)=(r,l)=(r,s)$.

proving that

$$\beta: J_4\rightarrow J_5 \text{ is a bijection}$$

Next

$$\begin{aligned}\prod_{(i,j)\in J_5}(x_i-x_k) &= \prod_{(i,j)\in J_5}(x_{(i,j)_1}-x_k)\\ &\underset{[\text{theorem: } 11.36]}{=} \prod_{(i,j)\in J_4}(x_{(\beta(i,j))_1}-x_k)\\ &= \prod_{(i,j)\in J_4}(x_{(i,l)_1}-x_k)\\ &= \prod_{(i,j)\in J_4}(x_i-x_k)\\ &= \prod_{(i,j)\in\{(i,k)\in I|i<k\}}(x_i-x_k)\end{aligned}$$

$$\begin{aligned} &= \prod_{(i,j)\in\{(i,k)\in I|i<k\}} (x_i - x_j) \\ &= \prod_{(i,j)\in J_4} (x_i - x_j) \\ &= Q_4 \end{aligned}$$

combining the above with [eq: 11.126] proves that

$$R_5 = Q_4 \tag{11.127}$$

***R*₆.** We have

$$\begin{aligned} R_6 &= \prod_{(i,j)\in J_6} \left(x_{(k \underset{n}{\leftrightarrow} l)(i)} - x_{(k \underset{n}{\leftrightarrow} l)(j)}\right) \\ &= \prod_{(i,j)\in\{(i,l)\in I|k<i<l\}} \left(x_{(k \underset{n}{\leftrightarrow} l)(i)} - x_{(k \underset{n}{\leftrightarrow} l)(j)}\right) \\ &= \prod_{(i,j)\in\{(i,l)\in I|k<i<l\}} (x_i - x_k) \\ &= \prod_{(i,j)\in J_6} (x_i - x_k) \end{aligned} \tag{11.128}$$

Define

$$\beta: J_0 = \{(k,j) \in I | k < j < l\} \to J_6 = \{(i,l) \in I | k < i < l\} \text{ by } \beta(i,j) = (j,l)$$

then we have:

injectivity. Let $(i,j),(i',j') \in J_0$ with $\beta(i,j) = \beta(i',j')$ then $i = k = i'$ and $(j,l) = (j',l) \Rightarrow j = j'$ hence $(i,j) = (i',j')$.

surjectivity. Let $(r,s) \in J_6$ then $s = l$ and $k < r < l$ so that for (k,r) we have $(k,r) \in J_0$ and $\beta(k,r) = (r,l) = (r,s)$.

hence

$$\beta: J_0 \to J_6 \text{ is a bijection so that } \mathrm{card}(J_0) = \mathrm{card}(J_6) \tag{11.129}$$

Further

$$\begin{aligned} \prod_{(i,j)\in J_6} (x_i - x_k) &= \prod_{(i,j)\in J_6} (x_{(i,j)_1} - x_k) \\ &\underset{[\text{theorem: } 11.36]}{=} \prod_{(i,j)\in J_0} (x_{(\beta(i,j))_1} - x_k) \\ &= \prod_{(i,j)\in J_0} (x_{(j,l)_1} - x_k) \\ &= \prod_{(i,j)\in J_0} (x_j - x_k) \\ &= \prod_{(i,j)\in J_0} ((-1)\cdot(x_k - x_j)) \\ &\underset{[\text{theorem: } 11.38]}{=} \left(\prod_{(i,j)\in J_0} (-1)\right) \cdot \prod_{(i,j)\in J_0} (x_k - x_j) \\ &\underset{[\text{theorem: } 11.48]}{=} (-1)^{\mathrm{card}(J_0)} \cdot \prod_{(i,j)\in J_0} (x_k - x_j) \\ &\underset{[\text{eq: } 11.129]}{=} (-1)^{\mathrm{card}(J_6)} \cdot \prod_{(i,j)\in J_0} (x_k - x_j) \\ &= (-1)^{\mathrm{card}(J_6)} \cdot \prod_{(i,j)\in\{(k,j)\in I|k<j<l\}} (x_k - x_j) \\ &= (-1)^{\mathrm{card}(J_6)} \cdot \prod_{(i,j)\in\{(k,j)\in I|k<j<l\}} (x_i - x_j) \\ &= (-1)^{\mathrm{card}(J_6)} \cdot \prod_{(i,j)\in J_0} (x_i - x_j) \\ &= (-1)^{\mathrm{card}(J_6)} \cdot Q_0 \end{aligned}$$

which together with [eq: 11.128] proves that

$$R_6 = (-1)^{\operatorname{card}(J_6)} \cdot Q_0 \tag{11.130}$$

***R*₇.** We have

$$\begin{aligned} R_7 &= \prod_{(i,j)\in J_7} \left(x_{(k \underset{n}{\leftrightarrow} l)(i)} - x_{(k \underset{n}{\leftrightarrow} l)(j)}\right) \\ &= \prod_{(i,j)\in\{(i,j)\in I \mid i\neq k,l \wedge j\neq k,l\}} \left(x_{(k \underset{n}{\leftrightarrow} l)(i)} - x_{(k \underset{n}{\leftrightarrow} l)(j)}\right) \\ &= \prod_{(i,j)\in\{(i,j)\in I \mid i\neq k,l \wedge j\neq k,l\}} (x_i - x_j) \\ &= \prod_{(i,j)\in J_7} (x_i - x_j) \\ &= Q_7 \end{aligned}$$

then

$$R_7 = Q_7 \tag{11.131}$$

Finally we have

$$\begin{array}{rc} \left(\left(k \underset{n}{\leftrightarrow} l\right)\Phi_n\right)(x_1,\ldots,x_n) & \underset{[\text{eq: } 11.115]}{=} \\ R_1\cdot R_2\cdot R_3\cdot R_4\cdot R_5\cdot R_6\cdot R_7 & \underset{[\text{eqs: } 11.118,11.119,11.121,11.123,11.125,11.127,11.130,11.131]}{=} \\ (-1)^{\operatorname{card}(J_6)}\cdot Q_6\cdot(-1)\cdot Q_1\cdot Q_3\cdot Q_2\cdot Q_5\cdot Q_4\cdot(-1)^{\operatorname{card}(J_6)}\cdot Q_0\cdot Q_7 & = \\ -(Q_1\cdot Q_2\cdot Q_3\cdot Q_4\cdot Q_5\cdot Q_6\cdot Q_7) & = \\ -\Phi_n(x_1,\ldots,x_n) & \end{array}$$

proving that

$$\left(k \underset{n}{\leftrightarrow} l\right)\Phi_n = -\Phi_n$$

□

Corollary 11.231. *Let* $n,m\in\mathbb{N}$ *and* $\left\{\left(i_k \underset{n}{\leftrightarrow} j_k\right)\right\}_{k\in\{1,\ldots,m\}} \subseteq P_n$ *such that* $\forall k\in\{1,\ldots,m\}\ i_k\neq j_k$ *then*

$$\left(\left(i_1 \underset{n}{\leftrightarrow} j_1\right)\circ\cdots\circ\left(i_m \underset{n}{\leftrightarrow} j_m\right)\right)\Phi_n = (-1)^m\cdot\Phi_n$$

Proof. We prove this by induction, so let

$$S = \left\{m\in\mathbb{N} \middle| \left(\left(i_1 \underset{n}{\leftrightarrow} j_1\right)\circ\cdots\circ\left(i_m \underset{n}{\leftrightarrow} j_m\right)\right)\Phi_n = (-1)^m\cdot\Phi_n\right\}$$

then we have:

$1\in S$. $\left(\left(i_1 \underset{n}{\leftrightarrow} j_1\right)\circ\cdots\circ\left(i_1 \underset{n}{\leftrightarrow} j_1\right)\right)\Phi_n = \left(i_1 \underset{n}{\leftrightarrow} j_1\right)\Phi_n \underset{[\text{theorem: } 11.230]}{=} (-1)\cdot\Phi_n = (-1)^1\cdot\Phi_n$ proving that $1\in S$.

$m\in S\Rightarrow m+1\in S$. We have

$$\begin{array}{rc} \left(\left(i_1 \underset{n}{\leftrightarrow} j_1\right)\circ\cdots\circ\left(i_{m+1} \underset{n}{\leftrightarrow} j_{m+1}\right)\right)\Phi_n & \underset{[\text{theorem: } 11.211]}{=} \\ \left(\left(\left(i_1 \underset{n}{\leftrightarrow} j_1\right)\circ\cdots\circ\left(i_m \underset{n}{\leftrightarrow} j_m\right)\right)\circ\left(i_{m+1} \underset{n}{\leftrightarrow} j_{m+1}\right)\right)\Phi_n & \underset{[\text{theorem: } 11.222]}{=} \\ \left(\left(i_1 \underset{n}{\leftrightarrow} j_1\right)\circ\cdots\circ\left(i_m \underset{n}{\leftrightarrow} j_m\right)\right)\left(\left(i_{m+1} \underset{n}{\leftrightarrow} j_{m+1}\right)\Phi_n\right) & \underset{[\text{theorem: } 11.230]}{=} \\ \left(\left(i_1 \underset{n}{\leftrightarrow} j_1\right)\circ\cdots\circ\left(i_m \underset{n}{\leftrightarrow} j_m\right)\right)((-1)\cdot\Phi_n) & \underset{[\text{theorem: } 11.223]}{=} \\ (-1)\cdot\left(\left(i_1 \underset{n}{\leftrightarrow} j_1\right)\circ\cdots\circ\left(i_m \underset{n}{\leftrightarrow} j_m\right)\right)(\Phi_n) & \underset{n\in S}{=} \\ (-1)\cdot((-1)^m\cdot\Phi_n) & = \\ (-1)^{m+1}\cdot\Phi_n & \end{array}$$

proving that $m+1\in S$ □

Theorem 11.232. *Let* $n\in\mathbb{N}$*and* $\sigma\in P_n$ *then there exist a* ***unique*** $\varepsilon_\sigma\in\{-1,1\}$ *such that*

$$\sigma\Phi_n = \varepsilon_\sigma\cdot\Phi_n$$

Proof. If $n \in \mathbb{N}$ then we have either

***n* = 1.** Then if $\sigma \in P_1$ we have that $P_1 = \mathrm{Id}_{\{1\}}$ so that $\sigma\Phi_1 = \sigma\mathrm{Id}_{\{1\}} \underset{\text{[theorem: 11.221]}}{=} \Phi_1$ so that if we take $\varepsilon_\sigma = 1$ then $\sigma\Phi_1 = \varepsilon_\sigma\Phi_1$.

***n* ∈ ℕ \ {1}.** Using [theorem: 11.216] there exist a $\left\{\left(i_k \underset{n}{\leftrightarrow} j_k\right)\right\}_{k \in \{1,\ldots,m\}} \subseteq P_n$ with $\forall k \in \{1,\ldots,m\}$ $i_k \neq j_k$ such that

$$\sigma = \left(i_1 \underset{n}{\leftrightarrow} j_1\right) \circ \cdots \circ \left(i_m \underset{n}{\leftrightarrow} j_m\right)$$

Hence by [theorem: 11.231] we have that $\left(\left(i_1 \underset{n}{\leftrightarrow} j_1\right) \circ \cdots \circ \left(i_m \underset{n}{\leftrightarrow} j_m\right)\right)\Phi_n = (-1)^m \cdot \Phi_n$. So if we take $\varepsilon_\sigma = (-1)^m$ we have $\sigma\Phi_n = \varepsilon_\sigma \cdot \Phi_n$

So in all cases there exists a ε_σ

$$\sigma\Phi_n = \varepsilon_\sigma \cdot \Phi_n$$

proving existence. Now for uniqueness assume that there exists a $\delta_\sigma \in \{1, -1\}$ such that $\sigma\Phi_n = \delta_\sigma \cdot \Phi_n$ then $\varepsilon_\sigma \cdot \Phi_n = \delta_\sigma \cdot \Phi_n$. hence $\varepsilon_\sigma \cdot \Phi_n(1,\ldots,n) = \delta_\sigma \cdot \Phi_n(1,\ldots,n)$. By [theorem: 11.229] $\Phi_n(1,\ldots,n) \neq 0$ so, after multiplying both sides by $\Phi_n(1,\ldots,n)^{-1}$, we have that $\varepsilon_\sigma = \delta_\sigma$. □

The above theorem ensures that the following definition makes sense.

Definition 11.233. *Let $n \in \mathbb{N}$and $\sigma \in P_n$ then $\mathrm{sign}(\sigma) \in \{-1, 1\}$ is the* ***unique*** *number such that*

$$\sigma\Phi_n = \mathrm{sign}(\sigma) \cdot \Phi_n$$

A permutation σ is called ***even*** *if $\mathrm{sign}(\sigma) = 1$ and* ***odd*** *if $\mathrm{sign}(\sigma) = -1$*

The concept of **even** or even permutation follows from the following.

Remark 11.234. Let $n \in \mathbb{N} \setminus \{1\}$ and $\sigma \in P_n$ then by [theorem: 11.216] there exist a family $\left\{\left(i_k \underset{n}{\leftrightarrow} j_k\right)\right\}_{k \in \{1,\ldots,m\}} \subseteq P_n$ of **strict** transpositions such that

$$\sigma = \left(i_1 \underset{n}{\leftrightarrow} j_1\right) \circ \cdots \circ \left(i_m \underset{n}{\leftrightarrow} j_m\right)$$

so that

$$\sigma\Phi_n = \left(\left(i_1 \underset{n}{\leftrightarrow} j_1\right) \circ \cdots \circ \left(i_m \underset{n}{\leftrightarrow} j_m\right)\right)\Phi_n \underset{\text{[theorem: 11.231]}}{=} (-1)^m\Phi_n$$

proving that

$$\mathrm{sign}(\sigma) = (-1)^m$$

In other words if σ can be written as a odd number of strict transpositions σ is odd and if σ can be written as a even number of transpositions σ is even.

Theorem 11.235. *If $n \in \mathbb{N}$ then we have*

1. *$\forall \sigma, \tau \in P_n$ that $\mathrm{sign}(\sigma \circ \tau) = \mathrm{sign}(\sigma) \cdot \mathrm{sign}(\tau)$*
2. *If $i, j \in \{1,\ldots,n\}$ with $i \neq j$ then $\mathrm{sign}\left(\left(i \underset{n}{\leftrightarrow} j\right)\right) = -1$*
3. *$\mathrm{sign}(\mathrm{Id}_{\{1,\ldots,n\}}) = 1$*
4. *$\forall \sigma \in P_n$ $\mathrm{sign}(\sigma) = \mathrm{sign}(\sigma^{-1})$*

Proof.

1.

$$\begin{array}{rcl} (\sigma \circ \tau)\Phi_n & \underset{\text{[theorem: 11.222]}}{=} & \sigma(\tau\Phi_n) \\ & = & \sigma(\mathrm{sign}(\tau) \cdot \Phi_n) \\ & \underset{\text{[theorem:11.223]}}{=} & \mathrm{sign}(\tau) \cdot (\tau\Phi_n) \\ & = & \mathrm{sign}(\tau) \cdot \mathrm{sign}(\sigma) \cdot \Phi_n \\ & = & \mathrm{sign}(\sigma) \cdot \mathrm{sign}(\tau) \cdot \Phi_n \end{array}$$

proving that

$$\mathrm{sign}(\sigma \circ \tau) = \mathrm{sign}(\sigma) \cdot \mathrm{sign}(\tau)$$

2. As $(i \underset{n}{\leftrightarrow} j)\Phi \underset{[\text{theorem: } 11.230]}{=} (-1) \cdot \Phi$ we have
$$\text{sign}((i \underset{n}{\leftrightarrow} j)) = -1$$
3. As $\text{Id}_{\{1,\ldots,n\}}\Phi_n \underset{[\text{theorem: } 11.221]}{=} \Phi_n$ we have
$$\text{sign}(\text{Id}_{\{1,\ldots,n\}}) = 1$$
4. We have
$$1 \underset{(2)}{=} \text{sign}(\text{Id}_{\{1,\ldots,n\}}) = \text{sign}(\sigma \circ \sigma^{-1}) \underset{(1)}{=} \text{sign}(\sigma) \cdot \text{sign}(\sigma^{-1})$$
so that $\text{sign}(\sigma) = \text{sign}(\sigma) \cdot 1 = (\text{sign}(\sigma) \cdot \text{sign}(\sigma)) \cdot \text{sign}(\sigma^{-1}) \underset{\text{sing}(\sigma) \in \{-1,1\}}{=} \text{sign}(\sigma^{-1})$ □

The following permutation will be used in determinant functions later.

Definition 11.236. *Let $n \in \mathbb{N}$ and $i, j \in \{1, \ldots, n\}$ then we define*
$$(i \underset{n}{\rightsquigarrow} j): \{1, \ldots, n\} \rightarrow \{1, \ldots, n\}$$
as follows:

1. *If $i = j$ then $(i \underset{n}{\rightsquigarrow} j) = \text{Id}_{\{1,\ldots,n\}}$*
2. *If $i < j$ then for $k \in \{1, \ldots, n\}$ we have*
$$(i \underset{n}{\rightsquigarrow} j)(k) = \begin{cases} k \text{ if } 1 \leqslant k < i \\ k+1 \text{ if } i \leqslant k < j \\ i \text{ if } k = j \\ k \text{ if } j < k \leqslant n \end{cases}$$
3. *If $j < i$ then for $k \in \{1, \ldots, n\}$ we have*
$$(i \underset{n}{\rightsquigarrow} j)(k) = \begin{cases} k \text{ if } 1 \leqslant k < j \\ i \text{ if } k = j \\ k-1 \text{ if } j < k \leqslant i \\ k \text{ if } i < k \leqslant n \end{cases}$$

Example 11.237. To get a idea of the mapping $(i \underset{n}{\rightsquigarrow} j)$ we can look at some examples:

1. $n = 6$
$$(2 \underset{6}{\rightsquigarrow} 5)\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 4 \\ 5 \\ 2 \\ 6 \end{pmatrix}$$
2. $n = 6$
$$(5 \underset{6}{\rightsquigarrow} 2)\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{pmatrix} = \begin{pmatrix} 1 \\ 5 \\ 2 \\ 3 \\ 4 \\ 6 \end{pmatrix}$$

So we interpret $(i \underset{n}{\rightsquigarrow} j)$ as removing i from its position an inserting it after (before) j if $i < j$ $(j < i)$.

Next we prove that $(i \underset{n}{\rightsquigarrow} j)$ is actual a permutation and decompose it in a composition of transpositions so that we can calculate its sign. First we need a little lemma:

Definition 11.238. *Let $n \in \mathbb{N}$, $i, j \in \{1, \ldots, n\}$ then we define*

1. *If $i < j$ then $\theta^{i<j}: \{1, \ldots, j-i\} \rightarrow \{i+1, \ldots, j\}$ by $\theta^{i<j}(l) = j - (l-1)$*

2. *If* $j<i$ *then* $\theta^{i>j}\colon\{1,\ldots,i-j\}\rightarrow\{j,\ldots,i-1\}$ *by* $\theta^{i>j}(l)=j+(l-1)$

Proof. Of course we must prove that the range is correct.

1. We have

$$\begin{array}{rcl} l\in\{1,\ldots,j-i\} & \Leftrightarrow & 1\leqslant l\leqslant j-i\\ & \Leftrightarrow & 0\leqslant l-1\leqslant j-i-1\\ & \Leftrightarrow & i-j+1\leqslant-(l-1)\leqslant 0\\ & \Leftrightarrow & i+1\leqslant j-(l-1)\leqslant j\\ & \Leftrightarrow & i+1\leqslant\theta^{i<j}(l)\leqslant j\\ & \Leftrightarrow & \theta^{i<j}(l)\in\{i+1,\ldots,j\} \end{array}$$

2. We have

$$\begin{array}{rcl} l\in\{1,\ldots,i-j\} & \Leftrightarrow & 1\leqslant l\leqslant i-j\\ & \Leftrightarrow & 0\leqslant l-1\leqslant i-j-1\\ & \Leftrightarrow & j\leqslant j+(l-1)\leqslant i-1\\ & \Leftrightarrow & j\leqslant\theta^{i>j}(l)\leqslant i-1\\ & \Leftrightarrow & \theta^{i>j}(l)\in\{j,\ldots,i-1\}\\ & & \square \end{array}$$

Theorem 11.239. *If* $n\in\mathbb{N}$ *and* $i,j\in\{1,\ldots,n\}$ *then we have:*

1. *If* $i<j$ *then*
$$\left(i\underset{n}{\rightsquigarrow}j\right)=\left(i\underset{n}{\leftrightarrow}\theta^{i<j}(1)\right)\circ\cdots\circ\left(i\underset{n}{\leftrightarrow}\theta^{i<j}(j-i)\right)$$
2. *If* $j<i$ *then*
$$\left(i\underset{n}{\rightsquigarrow}j\right)=\left(i\underset{n}{\leftrightarrow}\theta^{i>j}(1)\right)\circ\cdots\circ\left(i\underset{n}{\leftrightarrow}\theta^{i>j}(i-j\right)$$
3. $\left(i\underset{n}{\rightsquigarrow}j\right)\in P_n$
4. $\operatorname{sign}\left(\left(i\underset{n}{\rightsquigarrow}j\right)\right)=(-1)^{|i-j|}$

Proof.

1. Let $n\in\mathbb{N}$. We proceed now by induction on $k=j-i$, so let

$$S_n=\left\{k\in\mathbb{N}|\text{If } i,j\in\{1,\ldots,n\}\text{with } i<j\wedge j-i=k\text{ then }\left(i\underset{n}{\rightsquigarrow}j\right)=\left(i\underset{n}{\leftrightarrow}\theta^{i<j}(1)\right)\circ\cdots\circ\left(i\underset{n}{\leftrightarrow}\theta^{i<j}(j-i)\right)\right\}$$

then we have:

$1\in S$. Let $i,j\in\{1,\ldots,n\}$ with $i<j\wedge j-i=1\Rightarrow i<j\wedge j=i+1$. Then

$$\theta^{i<j}(1)=j-(1-1)=j=i+1 \tag{11.132}$$

and we have for $l\in\{1,\ldots,n\}$ either:

$1\leqslant l<i$. Then

$$\begin{array}{rcl} \left(i\underset{n}{\rightsquigarrow}j\right)(l) & = & l\\ & \underset{l<i<i+1\Rightarrow l\neq i,i+1}{=} & \left(i\underset{n}{\leftrightarrow}i+1\right)(l)\\ & \underset{[\text{eq: } 11.132]}{=} & \left(i\underset{n}{\leftrightarrow}\theta^{i<j}(1)\right)(l) \end{array}$$

$i\leqslant l<j$. Then $i\leqslant l<i+1$ so that $l=i$ and

$$\begin{array}{rcl} \left(i\underset{n}{\rightsquigarrow}j\right)(l) & = & l+1\\ & \underset{l=i}{=} & i+1\\ & = & \left(i\underset{n}{\leftrightarrow}i+1\right)(i)\\ & \underset{l=i}{=} & \left(i\underset{n}{\leftrightarrow}i+1\right)(l)\\ & \underset{[\text{eq: } 11.132]}{=} & \left(i\underset{n}{\leftrightarrow}\theta^{i<j}(1)\right)(l) \end{array}$$

$\boldsymbol{l=j}$**.** Then $l=j=i+1$ and

$$\begin{aligned}(i\underset{n}{\rightsquigarrow}j)(l) &= i\\ &= (i\underset{n}{\leftrightarrow}i+1)(i+1)\\ &\underset{l=i+1}{=} (i\underset{n}{\leftrightarrow}i+1)(l)\\ &\underset{\text{[eq: 11.132]}}{=} (i\underset{n}{\leftrightarrow}\theta^{i<j}(1))(l)\end{aligned}$$

$\boldsymbol{j<l\leqslant n}$**.** Then

$$\begin{aligned}(i\underset{n}{\rightsquigarrow}j)(l) &= l\\ &\underset{i<j<l\Rightarrow i+1\leqslant j<l\Rightarrow l\neq i,i+1}{=} (i\underset{n}{\leftrightarrow}i+1)(l)\\ &\underset{\text{[eq: 11.132]}}{=} (i\underset{n}{\leftrightarrow}\theta^{i<j}(1))(l)\end{aligned}$$

So we have in all cases $(i\underset{n}{\rightsquigarrow}j)(l)=(i\underset{n}{\leftrightarrow}\theta^{i<j}(1))(l)$ hence

$$(i\underset{n}{\rightsquigarrow}j)=(i\underset{n}{\leftrightarrow}\theta^{i<j}(1))\underset{j-i=1}{=}(i\underset{n}{\leftrightarrow}\theta^{i<j}(1))\circ\cdots\circ(i\underset{n}{\leftrightarrow}\theta^{i<j}(j-i))$$

proving that $1\in S_n$.

$\boldsymbol{k\in S\Rightarrow k+1\in S}$**.** Let $i,j\in\{1,\ldots,n\}$ with $i<j$ and $j-i=k+1$. First we prove that

$$(i\underset{n}{\rightsquigarrow}j)=(i\underset{n}{\leftrightarrow}j)\circ(i\underset{n}{\rightsquigarrow}j-1) \tag{11.133}$$

Proof. For $l\in\{1,\ldots,n\}$ we have the following cases to consider:

$\boldsymbol{1\leqslant l<i}$**.** Then

$$\begin{aligned}(i\underset{n}{\leftrightarrow}j)((i\underset{n}{\rightsquigarrow}j-1)(l)) &= (i\underset{n}{\leftrightarrow}j)(l)\\ &\underset{l<i<j\Rightarrow l\neq i,j}{=} l\\ &\underset{1\leqslant l<i}{=} (i\underset{n}{\rightsquigarrow}j)(l)\end{aligned}$$

$\boldsymbol{i\leqslant l<j-1}$**.** Then

$$\begin{aligned}(i\underset{n}{\leftrightarrow}j)((i\underset{n}{\rightsquigarrow}j-1)(l)) &= (i\underset{n}{\leftrightarrow}j)(l+1)\\ &\underset{i<i+1\leqslant l+1<j\Rightarrow l+1\neq i,j}{=} l+1\\ &\underset{i\leqslant l+1<j}{=} (i\underset{n}{\rightsquigarrow}j)(l)\end{aligned}$$

$\boldsymbol{l=j-1}$**.** Then

$$\begin{aligned}(i\underset{n}{\leftrightarrow}j)((i\underset{n}{\rightsquigarrow}j-1)(l)) &= (i\underset{n}{\leftrightarrow}j)(i)\\ &= j\\ &\underset{i<j\Rightarrow i\leqslant j-1=l<j}{=} (i\underset{n}{\rightsquigarrow}j)(j-1)\\ &\underset{l=j-1}{=} (i\underset{n}{\rightsquigarrow}j)(l)\end{aligned}$$

$\boldsymbol{j-1<l\leqslant n}$**.** Then $j\leqslant l$ so we have either:

$\boldsymbol{j=l}$**.** Then

$$\begin{aligned}(i\underset{n}{\leftrightarrow}j)((i\underset{n}{\rightsquigarrow}j-1)(l)) &= (i\underset{n}{\leftrightarrow}j)(l)\\ &\underset{l=j}{=} i\\ &\underset{l=j}{=} (i\underset{n}{\rightsquigarrow}j)(l)\end{aligned}$$

$\boldsymbol{j<l}$**.** Then

$$\begin{aligned}(i\underset{n}{\leftrightarrow}j)((i\underset{n}{\rightsquigarrow}j-1)(l)) &\underset{j-1<j<l}{=} (i\underset{n}{\leftrightarrow}j)(l)\\ &\underset{i<j<l\Rightarrow i,j\neq l}{=} l\\ &\underset{j<l\leqslant n}{=} (i\underset{n}{\rightsquigarrow}j)(l)\end{aligned}$$

So in all cases $(i \underset{n}{\leftrightarrow} j)((i \underset{n}{\rightsquigarrow} j-1)(l)) = (i \underset{n}{\rightsquigarrow} j)(l)$ or $(i \underset{n}{\rightsquigarrow} j) = (i \underset{n}{\leftrightarrow} j) \circ (i \underset{n}{\rightsquigarrow} j-1)$ proving [eq: 11.133]. □

As $\theta^{i<j}(1) = j + (1-1) = j$ we have by [eq: 11.133] that

$$(i \underset{n}{\rightsquigarrow} j) = (i \underset{n}{\leftrightarrow} \theta^{i<j}(1)) \circ (i \underset{n}{\rightsquigarrow} j-1) \tag{11.134}$$

Further as $(j-1) - i \underset{j-i=k+1}{=} k$ we have as $k \in S_n$ that

$$(i \underset{n}{\rightsquigarrow} j-1) = (i \underset{n}{\leftrightarrow} \theta^{i<j-1}(1)) \circ \cdots \circ (i \underset{n}{\leftrightarrow} \theta^{i<j-1}(j-1-i))$$

Let $l \in \{1, \ldots, j-i\}$ then $\theta^{i<j-1}(l) = j-1-(l-1) = j-((l+1)-1) = \theta^{i<j}(l+1)$ so we have that

$$(i \underset{n}{\rightsquigarrow} j-1) = (i \underset{n}{\leftrightarrow} \theta^{i<j}(2)) \circ \cdots \circ (i \underset{n}{\leftrightarrow} \theta^{i<j-1}(j-i)) \tag{11.135}$$

Further

$$\begin{aligned}
(i \underset{n}{\rightsquigarrow} j) &\underset{\text{[eq: 11.134]}}{=} (i \underset{n}{\leftrightarrow} \theta^{i<j}(1)) \circ (i \underset{n}{\rightsquigarrow} j-1) \\
&\underset{\text{[eq: 11.135}}{=} (i \underset{n}{\leftrightarrow} \theta^{i<j}(1)) \circ ((i \underset{n}{\leftrightarrow} \theta^{i<j}(2)) \circ \cdots \circ (i \underset{n}{\leftrightarrow} \theta^{i<j-1}(j-i))) \\
&\underset{\text{[theorem: 11.211]}}{=} (i \underset{n}{\leftrightarrow} \theta^{i<j}(1)) \circ \cdots \circ (i \underset{n}{\leftrightarrow} \theta^{i<j-1}(j-i))
\end{aligned}$$

proving that $k+1 \in \mathcal{S}_n$.

2. Let $n \in \mathbb{N}$. We proceed now by induction on $k = i - j$, so let

$$S_n = \{k \in \mathbb{N} | \text{If } i, j \in \{1, \ldots, n\} \text{with } j < i \wedge i - j = k \text{ then } (i \underset{n}{\rightsquigarrow} j) = (i \underset{n}{\leftrightarrow} \theta^{i>j}(1)) \circ \cdots \circ (i \underset{n}{\leftrightarrow} \theta^{i>j}(i-j))\}$$

then we have:

$1 \in S$. Let $i, j \in \{1, \ldots, n\}$ with $j < i$ and $i - j = 1$ so that $j = i - 1$. Then

$$\theta^{i>j}(1) = j + (1-1) = j = i - 1 \tag{11.136}$$

For $l \in \{1, \ldots, n\}$ we have either:

$1 \leqslant l < j$. Then we have

$$\begin{aligned}
(i \underset{n}{\rightsquigarrow} j)(l) &= l \\
&\underset{l<j<i \Rightarrow l \neq j, i-1}{=} (i \underset{n}{\leftrightarrow} i-1)(l) \\
&\underset{\text{[eq: 11.136]}}{=} (i \underset{n}{\leftrightarrow} \theta^{i>j}(1))(l)
\end{aligned}$$

$l = j$. Then we have $l = j = i - 1$

$$\begin{aligned}
(i \underset{n}{\rightsquigarrow} j)(l) &= i \\
&= (i \underset{n}{\leftrightarrow} i-1)(i-1) \\
&\underset{l=i-1}{=} (i \underset{n}{\leftrightarrow} i-1)(l) \\
&\underset{\text{[eq: 11.136]}}{=} (i \underset{n}{\leftrightarrow} \theta^{i>j}(1))(l)
\end{aligned}$$

$j < l \leqslant i$. Then we have $i - 1 < l \leqslant i$ so that $l = i$ and

$$\begin{aligned}
(i \underset{n}{\rightsquigarrow} j)(l) &= l - 1 \\
&\underset{l=i}{=} i - 1 \\
&= (i \underset{n}{\leftrightarrow} i-1)(i) \\
&\underset{l=i}{=} (i \underset{n}{\leftrightarrow} i-1)(l) \\
&\underset{\text{[eq: 11.136]}}{=} (i \underset{n}{\leftrightarrow} \theta^{i>j}(1))(l)
\end{aligned}$$

$i < l \leqslant n$. Then we have

$$\begin{aligned}
(i \underset{n}{\rightsquigarrow} j)(l) &= l \\
&\underset{j<i<l}{=} (i \underset{n}{\leftrightarrow} i-1)(l) \\
&\underset{\text{[eq: 11.136]}}{=} (i \underset{n}{\leftrightarrow} \theta^{i>j}(1))(l)
\end{aligned}$$

So in all cases we have $\left(i \underset{n}{\rightsquigarrow} j\right)(l) = \left(i \underset{n}{\leftrightarrow} \theta^{i>j}(1)\right)(l)$, hence

$$\left(i \underset{n}{\rightsquigarrow} j\right) = \left(i \underset{n}{\leftrightarrow} \theta^{i>j}(1)\right) \underset{i-j=1}{=} \left(i \underset{n}{\leftrightarrow} \theta^{i>j}(1)\right) \circ \cdots \circ \left(i \underset{n}{\leftrightarrow} \theta^{i>j}(i-j)\right)$$

proving that $1 \in S_n$.

$\boldsymbol{k \in S \Rightarrow k+1 \in S.}$ Let $i, j \in \{1, \ldots, n\}$ with $j < i$ and $i - j = k + 1$. First we prove that

$$\left(i \underset{n}{\rightsquigarrow} j\right) = \left(i \underset{n}{\leftrightarrow} j\right) \circ \left(i \underset{n}{\rightsquigarrow} j+1\right) \tag{11.137}$$

Proof. For $l \in \{1, \ldots, n\}$ we have either:

$\boldsymbol{1 \leqslant l < j.}$ Then

$$\begin{aligned}\left(i \underset{n}{\leftrightarrow} j\right)\left(\left(i \underset{n}{\rightsquigarrow} j+1\right)(l)\right) &\underset{1 \leqslant l < j < j+1}{=} \left(i \underset{n}{\leftrightarrow} j\right)(l) \\ &\underset{l<j<i}{=} l \\ &\underset{l<j}{=} \left(i \underset{n}{\rightsquigarrow} j\right)(l)\end{aligned}$$

$\boldsymbol{l = j.}$ Then

$$\begin{aligned}\left(i \underset{n}{\leftrightarrow} j\right)\left(\left(i \underset{n}{\rightsquigarrow} j+1\right)(l)\right) &\underset{l=j<j+1}{=} \left(i \underset{n}{\leftrightarrow} j\right)(l) \\ &\underset{l=j}{=} \left(i \underset{n}{\leftrightarrow} j\right)(j) \\ &= i \\ &\underset{l=j}{=} \left(i \underset{n}{\rightsquigarrow} j\right)(l)\end{aligned}$$

$\boldsymbol{j < l \leqslant i.}$ Then we have $j + 1 \leqslant l$ so that either:

$\boldsymbol{l = j+1.}$ Then

$$\begin{aligned}\left(i \underset{n}{\leftrightarrow} j\right)\left(\left(i \underset{n}{\rightsquigarrow} j+1\right)(l)\right) &= \left(i \underset{n}{\leftrightarrow} j\right)(i) \\ &= j \\ &\underset{l=j+1}{=} l-1 \\ &\underset{j<l\leqslant i}{=} \left(i \underset{n}{\rightsquigarrow} j\right)(l)\end{aligned}$$

$\boldsymbol{j+1 < l.}$ Then

$$\begin{aligned}\left(i \underset{n}{\leftrightarrow} j\right)\left(\left(i \underset{n}{\rightsquigarrow} j+1\right)(l)\right) &\underset{j+1<l\leqslant i}{=} \left(i \underset{n}{\leftrightarrow} j\right)(l-1) \\ &\underset{j<j+1\leqslant l-1<l\leqslant i}{=} l-1 \\ &\underset{j<l\leqslant i}{=} \left(i \underset{n}{\rightsquigarrow} j\right)(l)\end{aligned}$$

$\boldsymbol{i < l \leqslant n.}$ Then

$$\begin{aligned}\left(i \underset{n}{\leftrightarrow} j\right)\left(\left(i \underset{n}{\rightsquigarrow} j+1\right)(l)\right) &\underset{i<l\leqslant n}{=} \left(i \underset{n}{\leftrightarrow} j\right)(l) \\ &\underset{j<i<l}{=} l \\ &\underset{j<i<l}{=} \left(i \underset{n}{\rightsquigarrow} j\right)(l)\end{aligned}$$

So in all cases we have $\left(i \underset{n}{\leftrightarrow} j\right)\left(\left(i \underset{n}{\rightsquigarrow} j+1\right)(l)\right) = \left(i \underset{n}{\rightsquigarrow} j\right)(l)$, hence $\left(i \underset{n}{\rightsquigarrow} j\right) = \left(i \underset{n}{\leftrightarrow} j\right) \circ \left(i \underset{n}{\rightsquigarrow} j+1\right)$, proving [eq: 11.137]. □

Now $\theta^{i>j}(1) = j + (1-1) = j$ which combined with [eq: 11.137] gives

$$\left(i \underset{n}{\rightsquigarrow} j\right) = \left(i \underset{n}{\leftrightarrow} \theta^{i>j}(1)\right) \circ \left(i \underset{n}{\rightsquigarrow} j+1\right) \tag{11.138}$$

Further as $i - (j+1) = i - j - 1 = k + 1 - 1 = k \in \mathcal{S}_n$ we have that

$$\left(i \underset{n}{\rightsquigarrow} j+1\right) = \left(i \underset{n}{\leftrightarrow} \theta^{i>j+1}(1)\right) \circ \cdots \circ \left(i \underset{n}{\leftrightarrow} \theta^{i>j+1}(i-(j+1))\right)$$

Now if $l \in \{1, \ldots, i-(j+1)\}$ then

$$\theta^{i>j+1}(l) = j + 1 + (l-1) = j + ((l+1)-1) = \theta^{i>j}(l+1)$$

so that

$$\left(i \underset{n}{\rightsquigarrow} j+1\right) = \left(i \underset{n}{\leftrightarrow} \theta^{i>j}(2)\right) \circ \cdots \circ \left(i \underset{n}{\leftrightarrow} \theta^{i>j}(i-j)\right) \tag{11.139}$$

Next

$$\left(i \underset{n}{\rightsquigarrow} j\right) \underset{\text{[eq: 11.138]}}{=} \left(i \underset{n}{\leftrightarrow} \theta^{i>j}(1)\right) \circ \left(i \underset{n}{\rightsquigarrow} j+1\right) \underset{\text{[eq: 11.139]}}{=} \left(i \underset{n}{\leftrightarrow} \theta^{i>j}(1)\right) \circ \left(\left(i \underset{n}{\leftrightarrow} \theta^{i>j}(2)\right) \circ \cdots \circ \left(i \underset{n}{\leftrightarrow} \theta^{i>j}(i-j)\right)\right) \underset{\text{[theorem: 11.211]}}{=} \left(i \underset{n}{\leftrightarrow} \theta^{i>j}(1)\right) \circ \cdots \circ \left(i \underset{n}{\leftrightarrow} \theta^{i>j}(i-j)\right)$$

proving that $k+1 \in S_n$.

3. Let $i, j \in \{1, \ldots, n\}$ then we have either:

   $\boldsymbol{i=j}$**.** Then $\left(i \underset{n}{\rightsquigarrow} j\right) \underset{\text{def}}{=} \mathrm{Id}_{1,\ldots,n} \in P_n$

   $\boldsymbol{i<j}$**.** Then by (1) $\left(i \underset{n}{\rightsquigarrow} j\right) = \left(i \underset{n}{\leftrightarrow} \theta^{i<j}(1)\right) \circ \cdots \circ \left(i \underset{n}{\leftrightarrow} \theta^{i<j}(j-i)\right) \in P_n$

   $\boldsymbol{j<i}$**.** Then by (2) $\left(i \underset{n}{\rightsquigarrow} j\right) = \left(i \underset{n}{\leftrightarrow} \theta^{i>j}(1)\right) \circ \cdots \circ \left(i \underset{n}{\leftrightarrow} \theta^{i>j}(i-j)\right) \in P_n$

4. Let $i, j \in \{1, \ldots, n\}$ then we have either:

   $\boldsymbol{i=j}$**.** Then $|i-j| = 0$ and

$$\mathrm{sign}\left(\left(i \underset{n}{\rightsquigarrow} j\right)\right) = \mathrm{sign}(\mathrm{Id}_{\{1,\ldots,\}}) \underset{\text{[theorem: 11.235]}}{=} 1 = (-1)^0 = (-1)^{|i-j|}$$

   $\boldsymbol{i<j}$**.** Then as $\theta^{i<j}(\{1,\ldots,j-i\}) \underset{\text{[definition: 11.238]}}{=} \{i+1,\ldots,j\}$ so that $\forall l \in \{1,\ldots,j-i\}$ we have that $i \neq \theta^{i<j}(l)$, so $\left(i \underset{n}{\leftrightarrow} \theta^{i<j}(l)\right)$ is a strict transposition, hence as $i-j = |i-j|$

$$\mathrm{sign}\left(\left(i \underset{n}{\rightsquigarrow} j\right)\right) \underset{(1)}{=} \mathrm{sign}\left(\left(i \underset{n}{\leftrightarrow} \theta^{i<j}(1)\right) \circ \cdots \circ \left(i \underset{n}{\leftrightarrow} \theta^{i<j}(j-i)\right)\right) \underset{\text{[theorem: 11.234]}}{=} (-1)^{|i-j|}$$

   $\boldsymbol{j<i}$**.** Then as $\theta^{i>j}(\{1,\ldots,i-j\}) \underset{\text{[definition: 11.238]}}{=} \{j,\ldots,i-1\}$ so that $\forall l \in \{1,\ldots,i-j\}$ we have that $i \neq \theta^{i<j}(l)$, so $\left(i \underset{n}{\leftrightarrow} \theta^{i>j}(l)\right)$ is a strict transposition, hence as $j-i = |i-j|$

$$\mathrm{sign}\left(\left(i \underset{n}{\rightsquigarrow} j\right)\right) \underset{(1)}{=} \mathrm{sign}\left(\left(i \underset{n}{\leftrightarrow} \theta^{i>j}(1)\right) \circ \cdots \circ \left(i \underset{n}{\leftrightarrow} \theta^{i>j}(j-i)\right)\right) \underset{\text{[theorem: 11.234]}}{=} (-1)^{|i-j|}$$

□

11.6 Multilinear mappings

From now on, unless specified otherwise, instead of saying that $\langle X,+,\cdot\rangle$, $\langle Y,+,\cdot\rangle$ are vector spaces over a field $\langle F,+,\cdot\rangle$ we just says that X and Y are vector spaces over a field F, the addition operators and multiplication operators follows then from the context.
We use the following notation to specify what the i-the element of a tuple is.

Definition 11.240. *Let $n \in \mathbb{N}$ and $\{X_i\}_{i\in\{1,\ldots,n\}}$ a family of sets, $i \in \{1,\ldots,n\}$ and $x \in \prod_{i\in\{1,\ldots,n\}} X_i$ then*

$$\left(\cdots \underbrace{a}_{i} \cdots\right) \text{is equivalent with } x_i = a$$

using this notation we have

$$y = \left(x_1, \ldots \underbrace{a}_{i} \ldots, x_n\right) \text{ is equivalent with saying that } y_j = \begin{cases} x_i \text{ if } j = i \\ x_j \text{ if } j \in \{1,\ldots,n\} \setminus \{i\} \end{cases}$$

Another notation that we use is

$$y = (x_1, x_2, \ldots, x_n) \text{ is equivalent with } y_j = \begin{cases} x_1 \text{ if } j = 1 \\ x_j \text{ if } j \in \{1,\ldots,n\} \setminus \{i\} \end{cases}$$

Likewise

$$y = (x_1, \ldots, x_{n-1}, a) \text{ is equivalent with } y_j = \begin{cases} x_n \text{ if } j = n \\ x_j \text{ if } j \in \{1,\ldots,n-1\} \end{cases}$$

The notation

$$z = (x_1, \ldots, x_n, y_1, \ldots, y_m) \text{ is equivalent with } z_j = \begin{cases} x_j \text{ if } j \in \{1,\ldots,n\} \\ y_{j-n} \text{ if } j \in \{n+1,\ldots,n+m\} \end{cases}$$

Finally the notation

$$z = (x_1, \ldots, x_n, y_{n+1}, \ldots, x_m) \text{ is equivalent with } z_j = \begin{cases} x_j \text{ if } j \in \{1,\ldots,n\} \\ y_j \text{ if } j \in \{n+1,\ldots,m\} \end{cases}$$

Proposition 11.241. *Let $n \in \mathbb{N}$, $\{X_i\}_{i \in \{1,\ldots,n\}}$ a family of sets, $\sigma \in P_n$, $i \in \{1,\ldots,n\}$ and $x \in \prod_{i \in \{1,\ldots,n\}} X_i$ such that $x = \left(x_1, \ldots \underbrace{a}_{i} \ldots, x_n\right)$ then*

$$x \circ \sigma = \left((x \circ \sigma)_1, \ldots \underbrace{a}_{\sigma^{-1}(i)} \ldots, (x \circ \sigma)_n\right)$$

Proof. Let $i \in \{1, \ldots n\}$ and $x = \left(x_1, \ldots \underbrace{a}_{i} \ldots, x_n\right)$ then $x(i) = x_i = a$. As

$$(x \circ \sigma)_{\sigma^{-1}(i)} = x(\sigma(\sigma^{-1}(i))) = x(i) = a$$

we have by definition that

$$x \circ \sigma = \left((x \circ \sigma)_1, \ldots \underbrace{a}_{\sigma^{-1}(i)} \ldots, (x \circ \sigma)_n\right)$$

□

Definition 11.242. *Let $n \in \mathbb{N}$, $i \in \{1, \ldots, n\}, \{X_j\}_{j \in \{1,\ldots,n\}}$ a family of vector spaces over a field F and Y a vector space over the same field then a function $L: \prod_{j \in \{1,\ldots,n\}} X_j \rightarrow Y$ is a multilinear mapping if $\forall i \in \{1, \ldots, n\}$ we have*

1. *$\forall u, v \in X_i$ then $\forall x \in \prod_{i \in \{1,\ldots,n\}} X_i$ such that $x_i = u + v$ we have*

$$L(x) = L(y) + L(z)$$

   where y, z are defined by

$$y_k = \begin{cases} u \text{ if } k = i \\ x_k \text{ if } k \in \{1, \ldots, n\} \setminus \{i\} \end{cases} \text{ and } z_k = \begin{cases} v \text{ if } k = i \\ x_k \text{ if } k \in \{1, \ldots, n\} \setminus \{i\} \end{cases}$$

2. *$\forall u \in X_i$, $\forall \alpha \in F$ then $\forall x \in \prod_{i \in \{1,\ldots,n\}} X_i$ such that $x_i = \alpha \cdot u$ we have*

$$L(x) = \alpha \cdot L(y)$$

   where y is defined by

$$y_k = \begin{cases} u \text{ if } k = i \\ x_k \text{ if } k \in \{1, \ldots, n\} \setminus \{i\} \end{cases}$$

Using [definition: 11.240] this can be written more clearly as

1. *$\forall u, v \in X_i$ and $\forall \left(x_1, \ldots \underbrace{u + v}_{i} \ldots, x_n\right) \in \prod_{j \in \{1,\ldots,n\}} X_j$*

$$L\left(x_1, \ldots \underbrace{u + v}_{i} \ldots, x_n\right) = L\left(x_1, \ldots \underbrace{u}_{i} \ldots, x_n\right) + L\left(x_1, \ldots \underbrace{v}_{i} \ldots, x_n\right)$$

2. *$\forall \alpha \in F$, $\forall u \in X_i$ and $\forall \left(x_1, \ldots \underbrace{\alpha \cdot u}_{i} \ldots, x_n\right) \in \prod_{j \in \{1,\ldots,n\}} X$*

$$L\left(x_1, \ldots \underbrace{\alpha \cdot u}_{i} \ldots, x_n\right) = \alpha \cdot L\left(x_1, \ldots \underbrace{u}_{i} \ldots, x_n\right)$$

The set of graphs of multilinear mappings is noted as $\mathrm{Hom}(X_1 \ldots X_n; Y)$, more specific:

$$\mathrm{Hom}(X \ldots X_n; Y) = \left\{ L \in Y^{\prod_{i \in \{1,\ldots,n\}} X_i} | L: \prod_{i \in \{1,\ldots,n\}} X_i \rightarrow Y \text{ is multilinear} \right\}$$

Theorem 11.243. *Let $n \in \mathbb{N}$, $\{X_i\}_{i \in \{1,\ldots,n\}}$ be family of vector spaces over a field F, Y a vector space over the same field F and $L \in \mathrm{Hom}(X \ldots X_n; Y)$ then if $x \in \prod_{i \in \{1,\ldots,n\}} X_i$ such that $\exists i \in \{1, \ldots, n\}$ with $x_i = 0$ we have that $L(x) = 0$.*

Proof. If $x \in \prod_{i \in \{1,\ldots,n\}} X_i$ such that $\exists i \in \{1, \ldots, n\}$ with $x_i = 0$ then

$$x = \left(x_1, \ldots \underbrace{0}_{i} \ldots, x_n\right)$$

so that

$$\begin{aligned}
L(x) &= L\left(x_1, \ldots \underbrace{0}_{i} \ldots, x_n\right) \\
&= L\left(x_1, \ldots \underbrace{0 \cdot 0}_{i} \ldots, x_n\right) \\
&= 0 \cdot \left(x_1, \ldots \underbrace{0}_{i} \ldots, x_n\right) \\
&= 0
\end{aligned}$$

□

Just as with linear functions we have a alternative and simpler condition for multilinearity.

Theorem 11.244. *Let $n \in \mathbb{N}$, $\{X_j\}_{j \in \{1, \ldots, n\}}$ a family of vector spaces over a field F and Y a vector space over the same field then for $L: \prod_{i \in \{1, \ldots, n\}} X_i \to Y$ we have*

$$L: \prod_{i \in \{1, \ldots, n\}} X_i \to Y \text{ is a multilinear mapping}$$

$$\Updownarrow$$

$$\forall i \in \{1, \ldots, n\} \text{ then } \forall \left(x_1, \ldots \underbrace{u + \alpha \cdot v}_{i} \ldots x_n\right) \in \prod_{j \in \{1, \ldots, n\}} X_i \text{ where } \alpha \in F,\ u, v \in X_i$$

we have

$$L\left(x_1, \ldots \underbrace{u + \alpha \cdot v}_{i} \ldots x_n\right) = L\left(x_1, \ldots \underbrace{u}_{i} \ldots x_n\right) + \alpha \cdot L\left(x_1, \ldots \underbrace{v}_{i} \ldots x_n\right)$$

Proof. We have:

⇒. Let $i \in \{1, \ldots, n\}$, $u, v \in X_i$ and $\alpha \in F$ then we have

$$\begin{aligned}
L\left(x_1, \ldots \underbrace{u + \alpha \cdot v}_{i} \ldots x_n\right) &\underset{\text{multilinearity}}{=} \\
L\left(x_1, \ldots \underbrace{u}_{i} \ldots x_n\right) + L\left(x_1, \ldots \underbrace{\alpha \cdot v}_{i} \ldots x_n\right) &\underset{\text{multilinearity}}{=} \\
L\left(x_1, \ldots \underbrace{u}_{i} \ldots x_n\right) + \alpha \cdot L\left(x_1, \ldots \underbrace{v}_{i} \ldots x_n\right) &
\end{aligned}$$

⇐. Let $i \in \{1, \ldots, n\}$, $u, v \in X_i$ and $\alpha \in F$ then we have:

$$\begin{aligned}
L\left(x_1, \ldots \underbrace{u + v}_{i} \ldots x_n\right) &= \\
L\left(x_1, \ldots \underbrace{u + 1 \cdot v}_{i} \ldots x_n\right) &= \\
L\left(x_1, \ldots \underbrace{u}_{i} \ldots x_n\right) + 1 \cdot L\left(x_1, \ldots \underbrace{v}_{i} \ldots x_n\right) &= \\
L\left(x_1, \ldots \underbrace{u}_{i} \ldots x_n\right) + L\left(x_1, \ldots \underbrace{v}_{i} \ldots x_n\right) &
\end{aligned}$$

and

$$\begin{aligned}
L\left(x_1, \ldots \underbrace{\alpha \cdot v}_{i} \ldots, x_n\right) &= \\
L\left(x_1, \ldots \underbrace{0 + \alpha \cdot v}_{i} \ldots, x_n\right) &= \\
L(x_1, \ldots \underbrace{0} \ldots, x_n) + L\left(x_1, \ldots \underbrace{\alpha \cdot v}_{i} \ldots, x_n\right) &\underset{\text{[theorem: 11.243]}}{=} \\
\alpha \cdot L\left(x_1, \ldots \underbrace{u}_{i} \ldots, x_n\right) &
\end{aligned}$$

□

Given a family of linear mappings we can construct a multilinear mapping.

Theorem 11.245. *Let $n\in\mathbb{N}$, $\{X_i\}_{i\in\{1,\ldots,n\}}$ a family of vector spaces over a field F and $\{L_i\}_{i\in\{1,\ldots,n\}}$ a family satisfying $\forall i\in\{1,\ldots,n\}$ $L_i\in\operatorname{Hom}(X_i,F)$ then if we define*

$$\prod_{i\in\{1,\ldots,n\}} L_i\colon \prod_{i\in\{1,\ldots,n\}} X_i\to F \text{ by } \left(\prod_{i\in\{1,\ldots,n\}} L_i\right)(x_1,\ldots,x_n)=\prod_{i\in\{1,\ldots,n\}} L_i(x_i)$$

we have

$$\prod_{i\in\{1,\ldots,n\}} L_i\in\operatorname{Hom}(X_1\ldots X_n;F)$$

Proof. Let $i\in\{1,\ldots,n\}$, $x\in\prod_{j\in\{1,\ldots,n\}\setminus\{i\}} X_j$, $u,v\in X_i$ and $\alpha\in F$ then we have for

$$r=\left(x_1,\ldots\underbrace{u+\alpha\cdot v}_{i}\ldots,x_n\right)$$

$$s=\left(x_1,\ldots\underbrace{u}_{i}\ldots,x_n\right)$$

$$t=\left(x_1,\ldots\underbrace{v}_{i}\ldots,x_n\right)$$

$$\begin{aligned}
\left(\prod_{j\in\{1,\ldots,n\}} L_j\right)\left(x_1,\ldots\underbrace{u+\alpha\cdot v}_{i}\ldots,x_n\right) &=\\
\left(\prod_{j\in\{1,\ldots,n\}} L_j\right)(r) &=\\
\prod_{j\in\{1,\ldots,n\}} L_j(r_j) &\underset{[\text{theorem: } 11.43]}{=}\\
\left(\prod_{j\in\{1,\ldots,n\}\setminus\{i\}} L_j(r_j)\right)\cdot L_i(r_i) &=\\
\left(\prod_{j\in\{1,\ldots,n\}\setminus\{i\}} L_j(r_j)\right)\cdot L_i(u+\alpha\cdot v) &=\\
\left(\prod_{j\in\{1,\ldots,n\}\setminus\{i\}} L_j(r_j)\right)\cdot (L_i(u)+\alpha\cdot L_i(v)) &=\\
\left(\prod_{j\in\{1,\ldots,n\}\setminus\{i\}} L_j(r_j)\right)\cdot L_i(u)+\alpha\cdot\left(\left(\prod_{j\in\{1,\ldots,n\}\setminus\{i\}} L_j(r_j)\right)\cdot L_i(v)\right) &=\\
\left(\prod_{j\in\{1,\ldots,n\}\setminus\{i\}} L_j(s_j)\right)\cdot L_i(s_i)+\alpha\cdot\left(\left(\prod_{j\in\{1,\ldots,n\}\setminus\{i\}} L_j(t_j)\right)\cdot L_i(t_i)\right) &\underset{[\text{theprem: } 11.43]}{=}\\
\prod_{j\in\{1,\ldots,n\}} L_j(s_j)+\alpha\cdot\prod_{j\in\{1,\ldots,n\}} L_j(t_j) &=\\
\left(\prod_{j\in\{1,\ldots,n\}} L_j\right)(s)+\alpha\cdot\left(\prod_{j\in\{1,\ldots,n\}} L_j\right)(t) &=\\
\left(\prod_{j\in\{1,\ldots,n\}} L_j\right)\left(x_1,\ldots\underbrace{u}_{i}\ldots,x_n\right)+\alpha\cdot\left(\prod_{j\in\{1,\ldots,n\}} L_j\right)\left(x_1,\ldots\underbrace{v}_{i}\ldots,x_n\right)
\end{aligned}$$

proving by [theorem: 11.244] that $\prod_{j\in\{1,\ldots,n\}} L_j$ is a multilinear mapping. □

A special case of multilinear mappings is where $\{X_i\}_{i\in\{1,\ldots,n\}}$ satisfies $\forall i\in\{1,\ldots,n\}$ $X_i=X$, so that $\prod_{i\in\{1,\ldots,n\}} X_i=X^n$. This leads to the definition of n-linear functions.

Definition 11.246. *Let $n \in \mathbb{N}$ and X, Y vector spaces over a field F then a multilinear mapping*

$$L: X^n \to Y$$

*is called a **n-linear mapping**. The set of graphs of n-linear mappings is noted as $\mathrm{Hom}^n(X;Y)$ so*

$$\mathrm{Hom}^n(X;Y) = \mathrm{Hom}\left(\underbrace{X\ldots X}_{n}; Y\right)$$

Example 11.247. Let $n \in \mathbb{N}$ and F a field then $L_\otimes: F^n \to F$ defined by

$$L_\otimes(x) = \prod_{i \in \{1,\ldots,n\}} x_i$$

is multilinear. So

$$L_\otimes \in \mathrm{Hom}(F^n; F)$$

Proof. This follows from [theorem: 11.49] □

To be able to use induction in proofs about multilinear mappings we have the following theorem.

Theorem 11.248. *Let $n \in \mathbb{N}$, $\{X_i\}_{i \in \{1,\ldots,n+1\}}$ be a family of vector spaces over a field F, Y a vector space over the same field F, $a \in X_{n+1}$ and $L \in \mathrm{Hom}(X_1 \ldots X_{n+1}; Y)$ then if we define*

$$L_{\{\ldots a\}}: \prod_{j \in \{1,\ldots,n\}} X_j \to Y \text{ by } L_{\{\ldots a\}}(x_1, \ldots, x_n) = L(x_1, \ldots, x_n, a)$$

then

$$L_{\{\ldots a\}} \in \mathrm{Hom}(X_1 \ldots X_n; Y)$$

Proof. Let $i \in \{1, \ldots, n\}$, $\alpha \in F$, $u, v \in X_i$. If $\left(x_1, \ldots \underbrace{u + \alpha \cdot v}_{i} \ldots, x_n\right) \in \prod_{j \in \{1,\ldots,n\}} X_j$ then

$$\begin{aligned}
L_{\{\ldots a\}}\left(x_1, \ldots \underbrace{u + \alpha \cdot v}_{i} \ldots, x_n\right) &= \\
L\left(x_1, \ldots \underbrace{u + \alpha \cdot v}_{i} \ldots, x_n, a\right) &= \\
L\left(x_1, \ldots \underbrace{u}_{i} \ldots, x_n, a\right) + \alpha \cdot L\left(x_1, \ldots \underbrace{v}_{i} \ldots, x_n, a\right) &= \\
L_{\{\ldots a\}}\left(x_1, \ldots \underbrace{u}_{i} \ldots, x_n, \right) + \alpha \cdot L_{\{\ldots a\}}\left(x_1, \ldots \underbrace{v}_{i} \ldots, x_n, \right)
\end{aligned}$$

□

Definition 11.249. *Let $n \in \mathbb{N}$, $i \in \{1, \ldots, n\}$, $\{X_i\}_{j \in \{1,\ldots,n\}}$ a family of sets, Y a set, $x \in \prod_{j \in \{1,\ldots,n\} \setminus \{i\}} X_i$ and*

$$f: \prod_{j \in \{1,\ldots,n\}} X_i \to Y$$

then $f\left(x_1, \ldots \underbrace{*}_{i} \ldots, x_n\right): \prod_{i \in \{1,\ldots,n\} \setminus \{i\}} X_i \to Y$ is defined by

$$f\left(x_1, \ldots \underbrace{*}_{i} \ldots, x_n\right)(x) = f\left(x_1, \ldots \underbrace{x}_{i} \ldots, x_n\right)$$

Theorem 11.250. *Let $n \in \mathbb{N}$, $i \in \{1, \ldots, n\}, \{X_j\}_{j \in \{1,\ldots,n\}}$ a family of vector spaces over a field F and Y a vector space over the same field $L \in \mathrm{Hom}(X_1 \ldots X_n; Y)$ and $x \in \prod_{j \in \{1,\ldots,n\} \setminus \{i\}} X_i$ then*

$$L\left(x_1, \ldots \underbrace{*}_{i} \ldots, x_n\right) \in \mathrm{Hom}(X_i, Y)$$

Proof. Let $\alpha \in F$, $u, v \in X_i$ then

$$\begin{array}{rl} L\left(x_1, \ldots \underbrace{*}_{i} \ldots, x_n\right)(u+\alpha \cdot v) & = \\ L\left(x_1, \ldots \underbrace{u+\alpha \cdot v}_{i} \ldots, x_n\right) & = \\ L\left(x_1, \ldots \underbrace{u}_{i} \ldots, x_n\right) + \alpha \cdot L\left(x_1, \ldots \underbrace{v}_{i} \ldots, x_n\right) & = \\ L\left(x_1, \ldots \underbrace{*}_{i} \ldots, x_n\right)(u) + \alpha \cdot L\left(x_1, \ldots \underbrace{*}_{i} \ldots, x_n\right)(v) & \end{array}$$

□

Example 11.251. Let $n \in \mathbb{N}$, $\{X_i\}_{i \in \{1, \ldots, n\}}$ a family of vector spaces over a field F and Y a vector space over the same field F then $C_0 \in \mathrm{Hom}(X_1 \ldots X_n; Y)$

Proof. Let $x \in \prod_{i=1}^n X_i$ $i \in \{1, \ldots, n\}$, $\alpha \in F$ and $u, v \in X_i$ then

$$C_0\left(x_1, \ldots \underbrace{u+\alpha \cdot v}_{i} \ldots, x_n\right) = 0 = 0 + 0 = C_0\left(x_1, \ldots \underbrace{u}_{i} \ldots, x_n\right) + \alpha \cdot C_0\left(x_1, \ldots \underbrace{v}_{i} \ldots, x_n\right)$$ □

Theorem 11.252. *Let $n \in \mathbb{N}$, $\{X_i\}_{i \in \{1, \ldots, n\}}$ be a family of vector spaces over a field F and Y a vector space over the same field then $\mathrm{Hom}(X_1 \ldots X_n; Y)$ is a sub-space of $\langle Y^{\prod_{i=1}^n X_i}, +, \cdot \rangle$. So by [theorem: 11.57] $\langle \mathrm{Hom}(X_1 \ldots X_n; Y), +, \cdot \rangle$ is a vector space.*

Proof. First using [exercise: 11.251] we have that

$$C_0 \in \mathrm{Hom}(X_1 \ldots X_n; Y) \Rightarrow \mathrm{Hom}(X_1 \ldots X_n; Y) \neq \varnothing$$

Second if $\alpha \in F$ and $L_1, L_2 \in \mathrm{Hom}(X_1 \ldots X_n; Y)$ and $i \in \{1, \ldots, n\}$ we have

$$\begin{array}{rl} (\alpha \cdot L_1 + L_2)(x_1, \ldots, \underbrace{u+v}_{} \ldots, x_n) & = \\ \alpha \cdot L_1\left(x_1, \ldots \underbrace{u+v}_{i} \ldots, x_n\right) + L_2\left(x_1, \ldots \underbrace{u+v}_{i} \ldots, x_n\right) & = \\ \underbrace{\alpha \cdot L_1\left(x_1, \ldots \underbrace{u}_{i} \ldots, x_n\right)}_{1} + \underbrace{\alpha \cdot L_1\left(x_1, \ldots \underbrace{v}_{i} \ldots, x_n\right)}_{2} + \underbrace{L_2\left(x_1, \ldots \underbrace{u}_{i} \ldots, x_n\right)}_{1} + \underbrace{L_2\left(x_1, \ldots \underbrace{v}_{i} \ldots, x_n\right)}_{2} & = \\ \underbrace{(\alpha \cdot L_1 + L_2)\left(x_1, \ldots \underbrace{u}_{i} \ldots, x_n\right)}_{1} + \underbrace{(\alpha \cdot L_1 + L_2)\left(x_1, \ldots \underbrace{v}_{i} \ldots, x_n\right)}_{2} & = \end{array}$$

and

$$\begin{array}{rl} (\alpha \cdot L_1 + L_2)\left(x_1, \ldots x_1, \ldots \underbrace{\gamma \cdot u}_{i} \ldots, x_n \ldots, x_n\right) & = \\ \alpha \cdot L_1\left(x_1, \ldots \underbrace{\gamma \cdot u}_{i} \ldots, x_n\right) + L_2\left(x_1, \ldots \underbrace{\gamma \cdot u}_{i} \ldots, x_n\right) & = \\ \alpha \cdot \gamma \cdot L_1\left(x_1, \ldots \underbrace{u}_{i} \ldots, x_n\right) + \gamma \cdot L_2\left(x_1, \ldots \underbrace{u}_{i} \ldots, x_n\right) & = \\ \gamma \cdot \left(\alpha \cdot L_1\left(x_1, \ldots \underbrace{u}_{i} \ldots, x_n\right) + L_2\left(x_1, \ldots \underbrace{u}_{i} \ldots, x_n\right)\right) & = \\ \gamma \cdot (\alpha \cdot L_1 + L_2)\left(\underbrace{u}_{i}\right) & \end{array}$$

proving that

$$\alpha \cdot L_1 + L_2 \in \mathrm{Hom}(X_1 \ldots X_n; Y)$$ □

Theorem 11.253. *Let $n \in \mathbb{N}$ and $\{X_i\}_{i \in \{1, \ldots, n\}}$ be a family of vector spaces over a field F, Y and Z vector spaces over the same field F, $L \in \operatorname{Hom}(X_1 \ldots X_n; Y)$ and $S \in \operatorname{Hom}(Y, Z)$ then $S \circ H \in \operatorname{Hom}(X_1, \ldots X_n; Y)$*

Proof. Let $i \in \{1, \ldots, n\}$, $x = \left(x_1, \ldots \underbrace{u + \alpha \cdot v}_{i} \ldots, x_n\right) \in \prod_{j \in \{1, \ldots, n\} \setminus \{i\}} X_i$, $\alpha \in F$ and $u, v \in X_i$ then

$$\begin{aligned}
(S \circ L)\left(x_1, \ldots \underbrace{u + \alpha \cdot v}_{i} \ldots, x_n\right) &= \\
S\left(L\left(x_1, \ldots \underbrace{u + \alpha \cdot v}_{i} \ldots, x_n\right)\right) &= \\
S\left(L\left(x_1, \ldots \underbrace{u}_{i} \ldots, x_n\right) + \alpha \cdot L\left(x_1, \ldots \underbrace{v}_{i} \ldots, x_n\right)\right) &= \\
S\left(L\left(x_1, \ldots \underbrace{u}_{i} \ldots, x_n\right)\right) + \alpha \cdot S\left(L\left(x_1, \ldots \underbrace{v}_{i} \ldots, x_n\right)\right) &= \\
(S \circ L)\left(x_1, \ldots \underbrace{u}_{i} \ldots, x_n\right) + \alpha \cdot (S \circ L)\left(x_1, \ldots \underbrace{v}_{i} \ldots, x_n\right) &=
\end{aligned}$$

□

Theorem 11.254. *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 the same field F, $\{\alpha_i\}_{i \in \{1, \ldots, n\}} \subseteq F$ and $L \in \operatorname{Hom}(X_1 \ldots X_n; Y)$ then if $(x_1, \ldots, x_n) \in \prod_{i \in \{1, \ldots, n\}} X_i$ we have*

$$L(\alpha_1 \cdot x, \ldots, \alpha_n \cdot x_n) = \left(\prod_{i=1}^{n} \alpha_i\right) \cdot L(x_1, \ldots, x_n)$$

Proof. We prove this by induction, so let

$$S = \left\{k \in \mathbb{N} | \text{If } k \leqslant n \text{ and } L \in \operatorname{Hom}(X_1, \ldots, X_k; Y) \text{ then } L(\alpha_1 \cdot x_1, \ldots, \alpha_k \cdot x_k) = \left(\prod_{i=1}^{k} \alpha_i\right) \cdot L(x_1, \ldots, x_k)\right\} \text{then we have:}$$

$1 \in S$. Then $L(\alpha_1 \cdot x_1, \ldots, \alpha_1 \cdot x_1) = L(\alpha_1 \cdot x_1) = \alpha_1 \cdot L(x_1) = (\prod_i^k \alpha_i) \cdot L(x_1) = (\prod_i^k \alpha_i) \cdot L(x_1, \ldots, x_1)$ proving that $1 \in S$.

$k \in S \Rightarrow k + 1 \in S$. If $k + 1 \leqslant n$ and $L \in \operatorname{Hom}(X_1 \ldots X_{k+1}; Y)$ then by [theorem: 11.248] $L_{\{\ldots x_{k+1}\}} \in \operatorname{Hom}(X_1 \ldots X_k; Y)$ so that as $k \in S$ we have

$$L_{\{\ldots x_k\}}(\alpha_1 \cdot x_1, \ldots, \alpha_k \cdot x_k) = \left(\prod_{i=1}^{k} \alpha_i\right) \cdot L_{\{\ldots x_k\}}(x_1, \ldots, x_k) \tag{11.140}$$

Then we have

$$\begin{aligned}
L(\alpha_1 \cdot x_1, \ldots, \alpha_{k+1} \cdot x_{k+1}) &= \alpha_{k+1} \cdot L(\alpha_1 \cdot x_1, \ldots, \alpha_k \cdot x_k, x_{k+1}) \\
&= \alpha_{k+1} \cdot L_{\{\ldots, x_{k+1}\}}(\alpha_1 \cdot x_1, \ldots, a_k \cdot x_k) \\
&\underset{[\text{eq:: } 11.140]}{=} \alpha_{k+1} \cdot \left(\prod_{i=1}^{k} \alpha_i\right) \cdot L_{\{\ldots x_k\}}(x_1, \ldots, x_k) \\
&= \left(\prod_{i=1}^{k+1} \alpha_i\right) \cdot L_{\{\ldots x_k\}}(x_1, \ldots, x_k) \\
&= \left(\prod_{i=1}^{k+1} \alpha_i\right) \cdot L(x_1, \ldots, x_{k+1})
\end{aligned}$$

proving that $k + 1 \in S$.

Mathematical induction proves then that $S = \mathbb{N}$, so as $n \leqslant n$ we have that

$$L(\alpha_1 \cdot x, \ldots, \alpha_n \cdot x_n) = \left(\prod_{i=1}^{n} \alpha_i\right) \cdot L(x_1, \ldots, x_n)$$

□

Theorem 11.255. *Let $n\in\mathbb{N}$, $\{X_i\}_{i\in\{1,\ldots,n\}}$ be family of vector spaces over a field F, Z a vector space over the same field F, $\{Y_i\}_{i\in\{1,\ldots,n\}}$ a family such that $\forall i\in\{1,\ldots,n\}$ Y_i is a sub-space of X_i and $L\in\mathrm{Hom}(Y_1\ldots Y_n,Z)$ then their exist a $K\in\mathrm{Hom}(X_1\ldots X_n;Z)$ such that $K_{|\prod_{i\in\{1,\ldots,n\}}X_i}=L$. In other words K is a multilinear mapping extending L.*

Proof. Take $i\in\{1,\ldots,n\}$. As Y_i is a sub-space of X_i there exists by [theorem: 11.199] a sub-space Z_i of X_i such that $X_i=Y_i\oplus Z_i$. Using [theorem: 11.200] there exists $\pi_{Y_i}: X_i\rightarrow Y_i$ such that $\pi_{Y_i}\in\mathrm{Hom}(X_i,Y_i)$ and $(\pi_{Y_i})_{|Y_i}=\mathrm{Id}_{Y_i}$. Define

$$K: \prod_{i\in\{1,\ldots,n\}} X_i\rightarrow Z \text{ by } K(x_1,\ldots,x_n)=L(\pi_{Y_1}(x_1),\ldots,\pi_{Y_n}(x_n))$$

Let $i\in\{1,\ldots,n\}$ $u,v\in X_i$ and $\alpha\in F$ and $\left(x_1,\ldots\underbrace{\alpha\cdot u+v}_{i}\ldots,x_n\right)\in\prod_{j\in\{1,\ldots n\}}X_i$ then we have:

$$\begin{aligned}
K\left(x_1,\ldots\underbrace{\alpha\cdot u+v}_{i}\ldots,x_n\right) &=\\
L\left(\left(\pi_{Y_1}(x_1),\ldots\underbrace{\pi_{Y_i}(u+\alpha\cdot v)}_{i}\ldots,\pi_{Y_n}(x_n)\right)\right) &=\\
L\left(\left(\pi_{Y_1}(x_1),\ldots\underbrace{\pi_{Y_i}(u)+\alpha\cdot\pi_{Y_i}(v)}_{i}\ldots,\pi_{Y_n}(x_n)\right)\right) &=\\
L\left(\left(\pi_{Y_1}(x_1),\ldots\underbrace{\pi_{Y_i}(u)}_{i}\ldots,\pi_{Y_n}(x_n)\right)\right)+\alpha\cdot L\left(\left(\pi_{Y_1}(x_1),\ldots\underbrace{\pi_{Y_i}(v)}_{i}\ldots,\pi_{Y_n}(x_n)\right)\right) &=\\
K\left(x_1,\ldots,\underbrace{u}_{i}\ldots,x_n\right)+\alpha\cdot K\left(x_1,\ldots,\underbrace{v}_{i}\ldots,x_n\right)
\end{aligned}$$

proving that

$$K\in\mathrm{Hom}(X_1,\ldots,X_n;Y)$$

Finally we if $(x_1,\ldots,x_n)\in\prod_{i\in\{1,\ldots,n\}}Y_i$ so that $\forall i\in\{1,\ldots,n\}$ $x_i\in Y_i$ so that

$$\pi_i(x_i)=(\pi_i)_{|Y_i}=\mathrm{Id}_{Y_i}(x_i)=x_i$$

hence

$$K(x_1,\ldots,x_n)=L(\pi_{Y_1}(x_1),\ldots,\pi_{\mathrm{Yn}}(x_n))=L(x_1,\ldots,x_n)$$

proving that

$$K_{|\prod_{i\in\{1,\ldots,n\}}Y_i}=L$$

□

TODO

Theorem 11.256. *Let $n\in\mathbb{N}$, X a finite dimensional vector space over a field F with basis $E=\{e_1,\ldots,e_n\}$, Y a vector space over F and $L\in\mathrm{Hom}^n(X;Y)$ then $\forall x\in X$ we have*

$$L(x)=L(x_1,\ldots,x_n)=\sum_{k\in\{1,\ldots,n\}^n}\left(\prod_{j\in\{1,\ldots,n\}}\alpha_j^{k_j}\right)\cdot L(e_{k_1},\ldots,e_{k_n})$$

where $\forall i\in\{1,\ldots,n\}$ $\{\alpha_j^i\}_{j\in\{1,\ldots,n\}}$ is the ***unique family*** *satisfying $x_i=\sum_{j\in\{1,\ldots,n\}}\alpha_i^j\cdot e_j$ [which exists by [theorem: 11.137]*

Proof. We prove this by induction

$$S_n=\left\{m\in\mathbb{N}|\text{If } m\leqslant n \text{ then } L(x_1,\ldots,x_n)=\sum_{k\in\{1,\ldots,n\}^m}\left(\prod_{j\in\{1,\ldots,m\}}\alpha_{k_j}^j\right)\cdot L(e_{k_1},\ldots,e_{k_m},x_{m+1},\ldots,x_n)\right\}$$

where

$$(e_{\rho_1},\ldots,e_{\rho_m},x_{m+1},\ldots,x_n) \text{ is defined by } (e_1,\ldots,e_m,x_{m+1},\ldots,x_n)_k=\begin{cases} e_{\rho_k} \text{ if } k\in\{1,\ldots,m\}\\ x_k \text{ if } k\in\{m+1,\ldots,x_n\}\end{cases}$$

then we have

$1 \in S_n$. Then we have

$$\begin{aligned} L(x_1, \ldots, x_n) &= L\left(\sum_{i \in \{1, \ldots, n\}} \alpha_1^i \cdot e_i, x_2, \ldots, x_n\right) \\ &\underset{\text{[theorems: 11.25011.75]}}{=} \sum_{i \in \{1, \ldots, n\}} L(\alpha_1^i \cdot e_i, x_2, \ldots, x_n) \\ &\underset{\text{[theorem: 11.250]}}{=} \sum_{i \in \{1, \ldots, n\}} \alpha_1^i \cdot L(e_i, x_2, \ldots, x_n) \\ &= \sum_{i \in \{1, \ldots, n\}} \left(\prod_{j \in \{1\}} \alpha_j^i\right) \cdot L(e_i, x_2, \ldots, x_n) \end{aligned} \tag{11.141}$$

Define $\beta: \{1, \ldots, n\}^1 \to \{1, \ldots, n\}$ by $\beta(\rho) = \rho_1$ then we have

injectivity. If $\beta(\rho) = \beta(\sigma)$ then $\forall k \in \{1\}$ we have $\rho_k = \rho_1 = \beta(\rho) = \beta(\sigma) = \sigma_1 = \sigma_k$ so that $\rho = \sigma$.

surjectivity. If $i \in \{1, \ldots, n\}$ then define $\rho \in \{1, \ldots, n\}^1$ by $\rho_1 = i$ then $\beta(\rho) = \rho_1 = i$

so that

$$\beta: \{1, \ldots, n\}^1 \to \{1, \ldots, n\} \text{ is a bijection}$$

then we have

$$\begin{aligned} \sum_{\rho \in \{1, \ldots, n\}^1} \left(\prod_{j \in \{1\}} \alpha_j^{\rho_j}\right) \cdot L(e_{\rho_j}, x_2, \ldots, x_n) &= \\ \sum_{\rho \in \{1, \ldots, n\}^1} \left(\prod_{j \in \{1\}} \alpha_j^{\rho_1}\right) \cdot L(e_{\rho_1}, x_2, \ldots, x_n) &= \\ \sum_{i \in \{1, \ldots, n\}^1} \left(\prod_{j \in \{1\}} \alpha_j^{\beta(\rho)}\right) \cdot L(e_{\beta(\rho)}, x_2, \ldots, x_n) &\underset{\text{[theorem: 11.36]}}{=} \\ \sum_{i \in \{1, \ldots, n\}} \left(\prod_{j \in \{1\}} \alpha_j^i\right) \cdot L(e_i, x_2, \ldots, x_n) &\underset{\text{[eq: 11.141}}{=} \\ L(x_1, \ldots, x_n) & \end{aligned}$$

proving that $1 \in S_n$.

$m \in S_n \Rightarrow m + 1 \in S_n$. If $m + 1 \leqslant n$ then we have

$$\begin{aligned} L(x_1, \ldots, x_n) &\underset{m \in S_n}{=} \\ \sum_{\rho \in \{1, \ldots, n\}^m} \left(\prod_{j \in \{1, \ldots, m\}} \alpha_j^{\rho_j}\right) \cdot L(e_{\rho_1}, \ldots, e_{\rho_m}, x_{m+1}, \ldots, x_n) &= \\ \sum_{\rho \in \{1, \ldots, n\}^m} \left(\prod_{j \in \{1, \ldots, m\}} \alpha_j^{\rho_j}\right) \cdot L\left(e_{\rho_1}, \ldots, e_{\rho_m}, \sum_{i \in \{1, \ldots, n\}} \alpha_{m+1}^i \cdot x_i, x_{m+2}, \ldots, x_n\right) &= \\ \sum_{\rho \in \{1, \ldots, n\}^m} \left(\prod_{j \in \{1, \ldots, m\}} \alpha_j^{\rho_j}\right) \cdot \sum_{i \in \{1, \ldots, n\}} \alpha_{m+1}^i \cdot L(e_{\rho_1}, \ldots, e_{\rho_m}, e_i, x_{m+2}, \ldots x_n) &= \\ \sum_{\rho \in \{1, \ldots, n\}^m} \left(\sum_{i \in \{1, \ldots, n\}} \left(\left(\prod_{j \in \{1, \ldots, m\}} \alpha_j^{\rho_j}\right) \cdot \alpha_{m+1}^i\right) \cdot L(e_{\rho_1}, \ldots, e_{\rho_m}, e_i, x_{m+2}, \ldots x_n)\right) &= \end{aligned}$$

so that

$$L(x_1, \ldots, x_n) = \sum_{\rho \in \{1, \ldots, n\}^m} \left(\sum_{i \in \{1, \ldots, n\}} A_{\rho,i}\right) \tag{11.142}$$

where

$$A_{\rho,i} = \left(\left(\prod_{j \in \{1, \ldots, m\}} \alpha_j^{\rho_j}\right) \cdot \alpha_{m+1}^i\right) \cdot L(e_{\rho_1}, \ldots, e_{\rho_m}, e_i, x_{m+2}, \ldots x_n) \tag{11.143}$$

Given $\rho \in \{1, \ldots, n\}^m$ define

$$I_\rho = \{\rho\} \times \{1, \ldots, n\} \text{ and } \pi_2: I_\rho \rightarrow \{1, \ldots, n\} \text{ by } \pi_2(\rho, k) = k \tag{11.144}$$

then we have for π_2

injectivity. If $\pi_2(\omega, k) = \pi_2(\sigma, l)$ then as $(\omega, k), (\sigma, l) \in I_\rho = \{\rho\} \times \{1, \ldots, n\}$ we have $\omega = \rho = \sigma$, further, as $k = \pi_2(\omega, k) = \pi_2(\sigma, l) = l$, we have $(\omega, k) = (\sigma, l)$.

surjectivity. If $k \in \{1, \ldots, n\}$ then for $(\rho, k) \in I_\rho$ we have $\pi_2(\rho, k) = k$ proving surjectivity.

hence:

$$\forall \rho \in \{1, \ldots, n\} \text{ we have } \pi_2: I_\rho \rightarrow \{1, \ldots, n\} \text{ is a bijection}$$

so that

$$\begin{aligned}\sum_{k \in I_\rho} A_{k_1, k_2} &\underset{k \in I_\rho \Rightarrow k_1 = \rho}{=} \sum_{k \in I_\rho} A_{\rho, \pi_2(k)} \\ &\underset{[\text{theorem: } 11.36]}{=} \sum_{i \in \{1, \ldots, n\}} A_{\rho, i}\end{aligned}$$

which by substituting this in [eq: 11.142] gives

$$L(x_1, \ldots, x_n) = \sum_{\rho \in \{1, \ldots, n\}^m} \left(\sum_{k \in I_\rho} A_{k_1, k_2} \right) \tag{11.145}$$

If $\rho, \sigma \in \{1, \ldots, n\}^m$ with $\rho \neq \sigma$ then if $k \in I_\rho \bigcap I_\sigma$ we have $(\rho, i) = k = (\sigma, j)$ so that $\rho = \sigma$ contradicting $\rho \neq \sigma$. Hence $I_\rho \bigcap I_\sigma = \varnothing$, so if we define

$$I = \bigcup_{\rho \in \{1, \ldots, n\}^m} I_\rho$$

we have

$$\sum_{\rho \in I} A_{k_1, k_2} \underset{[\text{theorem: } 11.44]}{=} \sum_{\rho \in \{1, \ldots, n\}^m} \left(\sum_{k \in I_\rho} A_{k_1, k_2} \right)$$

which combined with [eq: 11.145] gives

$$L(x_1, \ldots, x_n) = \sum_{\rho \in I} A_{k_1, k_2} \tag{11.146}$$

Define now

$$\eta: \{1, \ldots, n\}^{m+1} \rightarrow \bigcup_{\rho \in \{1, \ldots, n\}^m} I_\rho = I \text{ by } \eta(k) = (k_{|\{1, \ldots, m\}}, k(m+1)) = (k_{|\{1, \ldots, m\}}, k_{m+1})$$

then we have:

injectivity. If $\eta(k) = \eta(l)$ then we have $(k_{|\{1, \ldots, m\}}, k(m+1)) = (l_{|\{1, \ldots, m\}}, l(m+1))$ then $k_{|\{1, \ldots, m\}} = l_{|\{1, \ldots, m\}} \Rightarrow \forall i \in \{1, \ldots, m\}$ $k(i) = l(i)$ and $k(m+1) = l(m+1)$ so that $k = l$.

surjectivity. If $z \in \bigcup_{\rho \in \{1, \ldots, n\}^m} I_\rho$ then there exists a $\rho \in \{1, \ldots, n\}^m$ so that $z \in I_\rho$. Hence there exists a $k \in \{1, \ldots, n\}$ such that $z = (\rho, k)$. Define $l \in \{1, \ldots, n\}^{m+1}$ by $l_i = \left\{ \begin{smallmatrix} \rho_i & \text{if } i \in \{1, \ldots, m\} \\ k & \text{if } i = m+1 \end{smallmatrix} \right.$ then $\eta(l) = (l_{|\{1, \ldots, m\}}, l(m+1)) = (\rho, k) = z$.

proving that

$$\eta: \{1, \ldots, n\}^{m+1} \rightarrow I \text{ is a bijection}$$

Using [theorem: 11.36] we have

$$\sum_{k \in I} A_{k_1, k_2} = \sum_{k \in \{1, \ldots, n\}^{m+1}} A_{\eta(k)_1, \eta(k)_2}$$

which combined with [eq: 11.146] proves

$$L(x_1, \ldots, x_n) = \sum_{k \in \{1, \ldots, n\}^{m+1}} A_{\eta(k)_1, \eta(k)_2} \tag{11.147}$$

Now for $k \in \{1, \ldots, n\}^{m+1}$ we have by [eq: 11.143]

$$\begin{aligned} A_{\eta(k)_1, \eta(k)_2} &= \\ A_{k_{|\{1,\ldots,m\}}, k_{m+1}} &= \\ \left(\left(\prod_{j \in \{1,\ldots,m\}} \alpha_j^{(k_{|\{1,\ldots,m\}})_j}\right) \cdot \alpha_{m+1}^{k_{m+1}}\right) \cdot L(e_{(k_{|\{1,\ldots,m\}})_1}, \ldots, e_{(k_{|\{1,\ldots,m\}})_m}, e_{k_{m+1}}, x_{m+2}, \ldots x_n) &= \\ \left(\left(\prod_{j \in \{1,\ldots,m\}} \alpha_j^{k_j}\right) \cdot \alpha_{m+1}^{k_{m+1}}\right) \cdot L(e_{k_1}, \ldots, e_{k_m}, e_{k_{m+1}}, x_{m+2}, \ldots x_n) &= \\ \left(\prod_{j \in \{1,\ldots,m+1\}} \alpha_j^{k_j}\right) \cdot L(e_{k_1}, \ldots, e_{k_{m+1}}, x_{m+2} \ldots, x_n) & \end{aligned}$$

which combined with [eq: 11.147] gives

$$L(x_1, \ldots, x_n) = \sum_{k \in \{1,\ldots,n\}^{m+1}} \left(\prod_{j \in \{1,\ldots,m+1\}} \alpha_j^{k_j}\right) \cdot L(e_{k_1}, \ldots, e_{k_{m+1}}, x_{m+2} \ldots, x_n)$$

proving that $m+1 \in S$.

By mathematical induction we have that $S_n = \mathbb{N}$. So as $n \leqslant n$ we have

$$L(x_1, \ldots, x_n) = \sum_{k \in \{1,\ldots,n\}^n} \left(\prod_{j \in \{1,\ldots,n\}} \alpha_j^{k_j}\right) \cdot L(e_{k_1}, \ldots, x_n)$$ □

11.7 Determinant Functions

We will make extensive use of the signature of a permutation in defining skew-symmetric functions. So it will be useful to have the concept of -1 in a field and examine its relation with $-1 \in \mathbb{Z}$.

Definition 11.257. *If F is a field and 1 is the multiplicative unit then -1 is the additive inverse of 1. Hence we have* $1 + (-1) = 0 = (-1) + 1$

Proposition 11.258. *If F is a field then* $(-1) \cdot (-1) = 1$

Proof. $0 = (-1) \cdot 0 = (-1) \cdot ((-1) + 1) = (-1) \cdot (-1) + (-1) \cdot 1 = (-1) \cdot (-1) + (-1)$ so that

$$1 = 0 + 1 = ((-1) \cdot (-1) + (-1)) + 1 = (-1) \cdot (-1)$$ □

Proposition 11.259. If *F is a field then $-1 \neq 0$ and $(-1)^{-1} = -1$.*

Proof. Assume that $-1 = 0$ then $1 = 1 + 0 = 1 + (-1) = 0$ contradicting that for a field $1 \neq 0$ [see definition: 4.51], so we must have that $(-1) \neq 0$ and thus that $(-1)^{-1}$ exist. Further, as $1 = (-1) \cdot (-1)$ we have that

$$(-1)^{-1} = (-1)^{-1} \cdot 1 = (-1)^{-1} \cdot ((-1) \cdot (-1)) = ((-1)^{-1} \cdot (-1)) \cdot (-1) = 1 \cdot (-1) = -1$$ □

Proposition 11.260. *If F is a field then $\forall f \in F$ we have $(-1) \cdot f = -f$*

Proof. As we have

$$f + (-1) \cdot f = (-1) \cdot f + f = (-1) \cdot f + 1 \cdot f = ((-1) + 1) \cdot f = 0 \cdot f = 0$$

□

Proposition 11.261. *If V is a vector space over a field F then $\forall x \in X$ we have $-x = (-1) \cdot x$*

Proof. Let $x \in X$ then

$$(-1) \cdot x + x = x + (-1) \cdot x = 1 \cdot x + (-1) \cdot x = (1 + (-1)) \cdot x = 0 \cdot x = 0$$

so that $-x = (-1) \cdot x$ $\square$

Next we define the power in a field by recursion.

Definition 11.262. *Let F be a field, $a \in F$ and $n \in \mathbb{N}_0$ then*

$$a^n = \begin{cases} 1 \text{ if } n = 0 \\ a \cdot a^{n-1} \text{ if } n \in \mathbb{N} \end{cases}$$

Proposition 11.263. *Let F be a field, $a \in F$ and $n, m \in \mathbb{N}_0$ then $a^n \cdot a^m = a^{n+m}$*

Proof. We prove this by induction on n. So let

$$S_m = \{n \in \mathbb{N} | a^n \cdot a^m = a^{n+m}\}$$

then we have:

$0 \in S_m$. Then $a^0 \cdot a^m = 1 \cdot a^m = a^m = a^{0+m}$ proving that $0 \in S_m$.

$n \in S_m \Rightarrow n + 1 \in S_m$. We have

$$\begin{array}{rcl} a^{n+1} \cdot a^m & = & (a \cdot a^n) \cdot a^m \\ & = & a \cdot (a^n \cdot a^m) \\ & \underset{m \in S}{=} & a \cdot a^{n+m} \\ & = & a^{(n+m)+1} \\ & = & a^{(n+1)+m} \end{array}$$

proving that $n + 1 \in \mathcal{S}_m$ $\square$

Proposition 11.264. *Let F be a field then for $n \in \mathbb{N}_0$ we have*

$$\begin{array}{rcl} 1^n & = & 1 \\ (-1)^n & = & -1 \text{ or } 1 \\ (-1)^n & = & \begin{cases} -1 \text{ if } n \text{ is odd} \\ 1 \text{ if } m \text{ is even} \end{cases} \end{array}$$

Further if $n \in \mathbb{N}$ then $0^n = 0$

Proof.

1. We use induction to prove this, so let

$$S = \{n \in \mathbb{N}_0 | 1^n = 1\}$$

then we have:

$0 \in S$. As $1^0 = 1$ by definition we have $0 \in S$

$n \in S \Rightarrow n + 1 \in S$. As $1^{n+1} = 1 \cdot 1^n \underset{n \in S}{=} 1 \cdot 1 = 1$ proving that $n + 1 \in S$

2. Again we use induction, so let

$$S = \{n \in \mathbb{N}_0 | (-1)^n \in \{-1, 1\}\}$$

then we have:

$0 \in S$. By definition $(-1)^0 = 1$ proving that $0 \in S$

$n \in S \Rightarrow n + 1 \in S$. As $n \in S$ we have either:

$(-1)^n = 1$. Then $(-1)^{n+1} = (-1) \cdot (-1)^n = (-1) \cdot 1 = -1$

$(-1)^n = -1$. Then $(-1)^{n+1} = (-1) \cdot (-1)^n = (-1) \cdot (-1) \underset{\text{[proposition: 11.258]}}{=} 1$

proving that $n+1 \in S$.

3. If n is even then there exists a $m \in \mathbb{N}_0$ such that $n = 2 \cdot m = m + m$. Now for $(-1)^m$ we have by (2) either:

   $(-1)^m = 1$. Then $(-1)^n = (-1)^{m+m} \underset{[\text{theorem: } 11.263]}{=} (-1)^m \cdot (-1)^m = 1 \cdot 1 = 1$

   $(-1)^m = -1$. Then

   $$(-1)^n = (-1)^{m+m} \underset{[\text{theorem: } 11.263]}{=} (-1)^m \cdot (-1)^m = (-1) \cdot (-1) \underset{[\text{theorem: } 11.258]}{=} 1$$

   hence

   $$(-1)^n = 1$$

4. If n is odd then there exist a $m \in \mathbb{N}_0$ such that $n = 2 \cdot m + 1$ then we have

   $$(-1)^n = (-1)^{2 \cdot m + 1} = (-1) \cdot (-1)^{2 \cdot m} \underset{(3)}{=} (-1) \cdot 1 = -1$$

5. If $n \in \mathbb{N}$ then $n - 1 \in \mathbb{N}_0$ so that

   $$0^n = 0^{(n-1)+1} = 0 \cdot 0^{n-1} = 0$$

□

Theorem 11.265. *Let $\langle F, +, \cdot \rangle$ be a field and $n \in \mathbb{N}_0$ then we have that*

1. $((-1)^n)^{-1} = (-1)^n$
2. $\forall m \in \mathbb{N}_0$ *we have* $(-1)^{n+2 \cdot m} = (-1)^n$

Proof.

1. $(-1)^n = 1 \cdot (-1)^n = (((-1)^n)^{-1} \cdot (-1)^n) \cdot (-1)^n = ((-1)^n)^{-1}((-1)^m \cdot (-1)^m) = ((-1)^n)^{-1}$
2. $(-1)^{n+2 \cdot m} \underset{[\text{proposition: } 11.263]}{=} (-1)^n \cdot (-1)^{2 \cdot m} = (-1)^n$. □

All the above work is done so that we can define a mapping of $\{1, -1\} \subseteq \mathbb{Z}$ to F which we will use in the definition of skew symmetric multilinear mappings.

Definition 11.266. *Let $\langle F, +, \cdot \rangle$ be a field then we define*

$$\odot: \{1, -1\} \times F \rightarrow F \text{ by } z \odot f = \begin{cases} f \text{ if } z = 1 \\ -f \text{ if } z = -1 \end{cases}$$

where $\{-1, 1\} \subseteq \mathbb{Z}$.

Theorem 11.267. *Let $n \in \mathbb{N}_0$ then*

$$\forall f \in F \text{ we have } \underbrace{(-1)^n}_{\in \{1, -1\} \subseteq \mathbb{Z}} \odot f = \underbrace{(-1)^n}_{\in \{1, -1\} \in F} \cdot f$$

Proof. We use induction to prove this, so let $f \in F$ and define

$$S_f = \{n \in \mathbb{N}_0 | (-1)^n \odot f = (-1)^n \cdot f\}$$

then we have:

$0 \in S$. As $(-1)^0 \odot f = 1 \odot f \underset{\text{def}}{=} f = 1 \cdot f \underset{[\text{definition: } 11.262]}{=} (-1)^0 \cdot f$ proving that $0 \in S$

$n \in S \Rightarrow n + 1 \in S$. Let $n \in S$ then we have either:

n is even. Then $n + 1$ is odd

$$\begin{aligned} (-1)^{n+1} \odot f &= (-1) \odot f \\ &\underset{\text{def}}{=} -f \\ &\underset{[\text{theorem: } 11.260]}{=} (-1) \cdot f \\ &\underset{[\text{theorem: } 11.264]}{=} (-1)^{n+1} \cdot f \end{aligned}$$

proving that $n+1\in S$ in this case.

***n* is odd.** Then $n+1$ is even

$$\begin{array}{rcl}(-1)^{n+1}\odot f & = & 1\odot f\\ & \underset{\text{def}}{=} & f\\ & = & 1\cdot f\\ & \underset{\text{[theorem: 11.264]}}{=} & (-1)^{n+1}\cdot f\end{array}$$

proving that $n+1\in S$ □

From now on, to avoid excessive notation we use . instead of $\odot$ and rely on context to figure out which operator is used. More specific if $\sigma\in P_n$ then $\text{sign}(\sigma)\in\{-1,1\}\subseteq\mathbb{Z}$ so that if $f\in F$ $\text{sign}(\sigma)\cdot f$ is actually $\text{sign}(\alpha)\odot f$.

Definition 11.268. *Let X,Y be vector spaces over a field F, $m\in\mathbb{N}$ then a n-linear mapping $L\in\text{Hom}^n(X;Y)$ [see definition: 11.246] is **symmetric** if $\forall\sigma\in P_n$ we have*

$$\sigma L=L$$

[see definition: 11.219].

More useful than symmetric n-linear mappings are skew-linear mappings.

Definition 11.269. *Let X,Y be vector spaces over a field F, $n\in\mathbb{N}$ then a n-linear mapping $L\in\text{Hom}^n(X;Y)$ [see definition: 11.246] is **skew-symmetric** if $\forall\sigma\in P_n$ we have*

$$\sigma L=\text{sign}(\sigma)\cdot L$$

[see definition: 11.219].

A trivial example of a mapping that is at the same time symmetric and skew-symmetric is the following.

Example 11.270. Let X,Y be vector space over a field F then if $L\in\text{Hom}^1(X;Y)$ L is symmetric and skew-symmetric.

Proof. If $\sigma\in P_1$ then $\sigma=\text{Id}_{\{1\}}$ so that $\text{sign}(\sigma)=1$ and $\sigma L(x)=L(x\circ\sigma)=L$ proving that

$$\sigma L=L=1\cdot L=\text{sign}(\sigma)\cdot L$$ □

If $L\in\text{Hom}^n(X;Y)$ is symmetric then $\sigma L=L\in\text{Hom}^n(X;Y)$ and if L is skew-symmetric we have $\sigma L=\text{sign}(\sigma)\cdot L\in\text{Hom}^n(X;Y)$. Actually this is a general result valid for every n-linear mapping.

Theorem 11.271. *Let X,Y be vector spaces over a field F, $n\in\mathbb{N}$ and $L\in\text{Hom}^n(X;Y)$ a n-linear mapping then*

$$\forall\sigma\in P_n \textit{ we have } \sigma L\in\text{Hom}^n(X;Y)$$

Proof. Let $i\in\{1,\ldots,\}$, $\alpha\in F$ and $u,v\in X_i$. Then if

$$\begin{array}{rcl}z & = & \left(x_1,\ldots\underbrace{u+\alpha\cdot v}_{i}\ldots,x_n\right)\in X^n\\ r & = & \left(x_1,\ldots\underbrace{u}_{i}\ldots,x_n\right)\in X^n\\ s & = & \left(x_1,\ldots\underbrace{v}_{i}\ldots,x_n\right)\in X^n\end{array}$$

Let $k=\sigma^{-1}(i)$ then if $j\neq k$ we must have $\sigma(j)\neq i$ [otherwise if $\sigma(j)=i$ we have as σ is bijective that $j=\sigma^{-1}(i)=k$ a contradiction]. So $\forall j\in\{1,\ldots,n\}\setminus\{k\}$ $\sigma(j)\neq i$ and we have:

$$(z\circ\sigma)_j = x(\sigma(j))=(r\circ\sigma)_j \tag{11.148}$$

$$(z\circ\sigma)_j = x(\sigma(j))=(s\circ\sigma)_j \tag{11.149}$$

Then

$$\begin{array}{rcl}
(\sigma L)\left(x_1,\ldots\underbrace{u+\alpha\cdot v}_{i}\ldots,x_n\right) & = & \\
(\sigma L)(z) & = & \\
L(z\circ\sigma) & \underset{\text{[theorem: 11.241]}}{=} & \\
L\left((z\circ\sigma)_1,\ldots\underbrace{u+\alpha\cdot v}_{\sigma^{-1}(i)}\ldots,(z\circ\sigma)_n\right) & = & \\
L\left((z\circ\sigma)_1,\ldots\underbrace{u}_{\sigma^{-1}(i)}\ldots,(z\circ\sigma)_n\right)+\alpha\cdot L\left((z\circ\sigma)_1,\ldots\underbrace{v}_{\sigma^{-1}(i)}\ldots,(z\circ\sigma)_n\right) & \underset{\text{[eqs: 11.148,11.149]}}{=} & \\
L\left((r\circ\sigma)_1,\ldots\underbrace{u}_{\sigma^{-1}(i)}\ldots,(r\circ\sigma)_n\right)+\alpha\cdot L\left((s\circ\sigma)_1,\ldots\underbrace{v}_{\sigma^{-1}(i)}\ldots,(s\circ\sigma)_n\right) & \underset{\text{[theorem: 11.241]}}{=} & \\
L(r\circ\sigma)+\alpha\cdot L(s\circ\sigma) & = & \\
(\sigma L)(r)+\alpha\cdot(\sigma l)(s) & = & \\
(\sigma L)(x_1,\ldots,x_{i-1},u,x_{i+1},\ldots,x_n)+\alpha\cdot(\sigma L)(x_1,\ldots,x_{i-1},v,x_{i+1},\ldots,x_n) & &
\end{array}$$

proving that σL is multilinear. □

The next theorem shows we can construct a new skew-symmetric from a n-linear mapping. First we need a little lemma:

Lemma 11.272. *Let $n\in\mathbb{N}$ and $\rho\in P_n$ then*

$$T_\rho: P_n\rightarrow P_n \text{ defined by } T_\rho(\sigma)=\rho\circ\sigma$$

is a bijection.

Proof.

injectivity. If $T_\rho(\sigma)=T_\rho(\sigma')$ then

$$\begin{array}{rcl}
\rho\circ\sigma=\rho\circ\sigma' & \Rightarrow & \rho^{-1}\circ(\rho\circ\sigma)=\rho^{-1}\circ(\rho\circ\sigma')\\
& \Rightarrow & (\rho^{-1}\circ\rho)\circ\sigma=(\rho^{-1}\circ\rho)\circ\sigma'\\
& \Rightarrow & \mathrm{Id}_{\{1,\ldots,n\}}\circ\sigma-\mathrm{Id}_{\{1,\ldots,n\}}\circ\sigma'\\
& \Rightarrow & \sigma=\sigma'
\end{array}$$

proving injectivity.

surjectivity. Let $\sigma\in P_n$ then $\rho^{-1}\circ\sigma\in P_n$ [see theorem: 11.208] and

$$T_\rho(\rho^{-1}\circ\sigma)=\rho\circ(\rho^{-1}\circ\sigma)=(\rho\circ\rho^{-1})\circ\sigma=\mathrm{Id}_{\{1,\ldots,n\}}\circ\sigma=\sigma$$

proving surjectivity. □

Theorem 11.273. *Let $n\in\mathbb{N}$, X,Y be vector spaces over a field F and a $L\in\mathrm{Hom}^n(X;Y)$ a n-linear mapping then*

$$\sum_{\sigma\in P_n}\mathrm{sign}(\sigma)\cdot\sigma L$$

is a skew-symmetric mapping:

***Note* 11.274.** *The sum is well defined as P_n is finite, $\forall\sigma\in P_n$ $\sigma L\in\mathrm{Hom}^n(X;Y)\Rightarrow\mathrm{sign}(\sigma)\cdot\sigma L\in\mathrm{Hom}^n(X;Y)\subseteq Y^{x^n}$ and Y^{X^n} is a vector space [see theorems: 11.207, 11.271 and 11.252].*

Proof. Using [theorem: 11.271] we have $\forall \sigma \in P_n$ that $\operatorname{sign}(\sigma) \cdot \sigma L \in \operatorname{Hom}^n(X;Y)$ so that by [theorem: 11.78]

$$\sum_{\sigma \in P_n} \operatorname{sign}(\sigma) \cdot \sigma L \in \operatorname{Hom}^n(X;Y) \text{ or } \sum_{\sigma \in P_n} \operatorname{sign}(\sigma) \cdot \sigma L \text{ is } n\text{-linear}$$

Next we have to prove that $\sum_{\sigma \in P_n} \operatorname{sign}(\sigma) \cdot \sigma L$ is skew-symmetric. So let $\rho \in P_n$ then we have for $x \in X^n$ that

$$\begin{aligned}
\left(\rho\left(\sum_{\sigma \in P_n} \operatorname{sign}(\sigma) \cdot \sigma L\right)\right)(x) &= \left(\sum_{\sigma \in P_n} \operatorname{sign}(\sigma) \cdot \sigma L\right)(x \circ \rho) \\
&= \sum_{\sigma \in P_n} (\operatorname{sign}(\sigma) \cdot \sigma L)(x \circ \rho) \\
&= \sum_{\sigma \in P_n} \operatorname{sign}(\sigma) \cdot (\sigma L)(x \circ \rho) \\
&= \sum_{\sigma \in P_n} \operatorname{sign}(\sigma) \cdot L((x \circ \rho) \circ \sigma) \\
&= \sum_{\sigma \in P_n} \operatorname{sign}(\sigma) \cdot L(x \circ (\rho \circ \sigma)) \\
&= \sum_{\sigma \in P_n} \operatorname{sign}(\sigma) \cdot ((\rho \circ \sigma) L)(x) \\
&\underset{\operatorname{sign}(\rho) \cdot \operatorname{sign}(\rho) = 1}{=} \sum_{\sigma \in P_n} (\operatorname{sign}(\rho) \cdot \operatorname{sign}(\rho)) \cdot \operatorname{sign}(\sigma) \cdot ((\rho \circ \sigma) L)(x) \\
&= \sum_{\sigma \in P_n} \operatorname{sign}(\rho) \cdot ((\operatorname{sign}(\rho) \cdot \operatorname{sign}(\sigma)) \cdot ((\rho \circ \sigma) L)(x)) \\
&= \operatorname{sign}(\rho) \cdot \sum_{\sigma \in P_n} (\operatorname{sign}(\rho) \cdot \operatorname{sign}(\sigma)) \cdot ((\rho \circ \sigma) L)(x) \\
&\underset{[\text{theorem: } 11.235]}{=} \operatorname{sign}(\rho) \cdot \sum_{\sigma \in P_n} \operatorname{sign}(\rho \circ \sigma) \cdot ((\rho \circ \sigma) L)(x) \\
&\underset{[\text{lemma: } 11.272]}{=} \operatorname{sign}(\rho) \cdot \sum_{\sigma \in P_n} \operatorname{sign}(T_\rho(\sigma)) \cdot (T_\rho(\sigma) L)(x) \\
&\underset{[\text{theorem: } 11.36]}{=} \operatorname{sign}(\rho) \cdot \sum_{\sigma \in P_n} \operatorname{sign}(\sigma) \cdot (\sigma L)(x)
\end{aligned}$$

proving that

$$\rho\left(\sum_{\sigma \in P_n} \operatorname{sign}(\sigma) \cdot \sigma L\right) = \operatorname{sign}(\rho) \cdot \sum_{\sigma \in P_n} \operatorname{sign}(\sigma) \cdot \sigma L$$

□

Theorem 11.275. *Let X, Y be vector spaces over a field F with characteristic zero [see definition: 10.47], $n \in \mathbb{N}$ and $L \in \operatorname{Hom}^n(X;Y)$ then the following are equivalent:*

1. *L is skew-symmetric.*
2. *For every $x \in X^n$ for which $\exists i, j \in \{1, \ldots, n\}$ with $i \neq j$ such that $x_i = x_j$ we have $L(x) = 0$.*
3. *For every $x \in X^n$ such that $\{x_i\}_{i \in \{1, \ldots, n\}} \subseteq X$ is linear dependent we have $L(x) = 0$.*

Proof.

1 ⇒ 2. Let $x \in X^n$ such that $\exists i, j \in \{1, \ldots, n\}$ with $i \neq j$ and $x_i = x_j \Rightarrow x(i) = x(j)$. Then we have for $k \in \{1, \ldots, n\}$ that

$$\begin{aligned}
(x \circ (i \leftrightarrow j_n))(k) &= x\left(\left(i \underset{n}{\leftrightarrow} j\right)(k)\right) \\
&= \begin{cases} x(i) \text{ if } k = j \\ x(j) \text{ if } k = i \\ x(k) \text{ if } k \in \{1, \ldots, n\} \setminus \{i, j\} \end{cases} \\
&\underset{x(i) = x(j)}{=} \begin{cases} x(k) \text{ if } k = j \\ x(k) \text{ if } k = i \\ x(k) \text{ if } k \in \{1, \ldots, n\} \setminus \{i, j\} \end{cases} \\
&= x(k)
\end{aligned}$$

proving that

$$x \circ (i \underset{n}{\leftrightarrow} j) = x$$

Then

$$\begin{array}{rcl} L(x) & = & L(x \circ (i \underset{n}{\leftrightarrow} j)) \\ & \underset{L \text{ is skew-symmetric}}{=} & \operatorname{sign}((i \underset{n}{\leftrightarrow} j)) \cdot L(x) \\ & \underset{[\text{theorem: } 11.235]}{=} & (-1) \cdot L(x) \\ & = & -L(x) \end{array}$$

so that $L(x) + L(x) = 0$ or $(1+1) \cdot L(x) = 0$, as F has characteristic zero it follows that $L(x) = 0$.

$2 \Rightarrow 1$. Assume that (2) hold then we prove the following

$$\forall x \in X^n \text{ we have } \forall i, j \in \{1, \ldots, n\} \text{ with } i \neq j \text{ that } L(x \circ (i \underset{n}{\leftrightarrow} j)) = -L(x) \tag{11.150}$$

Proof. Let $x \in X^n$, $i, j \in \{1, \ldots, n\}$ with $i \neq j$. Then

$$\begin{array}{r} 0 = \\ L\left(x_1, \ldots \underbrace{\boldsymbol{x_i} + \boldsymbol{x_j}}_{i} \ldots \underbrace{\boldsymbol{x_i} + \boldsymbol{x_j}}_{j} \ldots, x_n\right) = \\ L\left(x_1, \ldots \underbrace{\boldsymbol{x_i} + \boldsymbol{x_j}}_{i} \ldots \underbrace{\boldsymbol{x_i}}_{j} \ldots, x_n\right) + L\left(x_1, \ldots \underbrace{\boldsymbol{x_i} + \boldsymbol{x_j}}_{i} \ldots \underbrace{\boldsymbol{x_j}}_{j} \ldots, x_n\right) = \\ L\left(x_1, \ldots \underbrace{\boldsymbol{x_i}}_{i} \ldots \underbrace{\boldsymbol{x_i}}_{j} \ldots, x_n\right) + L\left(x_1, \ldots \underbrace{\boldsymbol{x_j}}_{i} \ldots \underbrace{\boldsymbol{x_i}}_{j} \ldots, x_n\right) + L\left(x_1, \ldots \underbrace{\boldsymbol{x_i}}_{i} \ldots \underbrace{\boldsymbol{x_j}}_{j} \ldots, x_n\right) + L\left(x_1, \ldots \underbrace{\boldsymbol{x_j}}_{i} \ldots \underbrace{\boldsymbol{x_j}}_{j} \ldots, x_n\right) = \\ 0 + L\left(x_1, \ldots \underbrace{\boldsymbol{x_j}}_{i} \ldots \underbrace{\boldsymbol{x_i}}_{j} \ldots, x_n\right) + L\left(x_1, \ldots \underbrace{\boldsymbol{x_i}}_{i} \ldots \underbrace{\boldsymbol{x_j}}_{j} \ldots, x_n\right) + 0 = \\ = \\ L(x \circ (i \underset{n}{\leftrightarrow} j)) + L(x) \end{array}$$

proving that

$$L(x \circ (i \underset{n}{\leftrightarrow} j)) = -L(x)$$

□

Next we prove that

$$\text{If } m \in \mathbb{N} \ \{\text{if } (i_k \leftrightarrow j_k)\}_{k \in \{,\ldots,m\}} \text{ is a family of strict transformations then } ((i_1 \underset{n}{\leftrightarrow} j_1) \circ \cdots \circ (i_m \underset{n}{\leftrightarrow} j_m))L = (-1)^m L \tag{11.151}$$

Proof. We proceed by induction, so let

$S = \{m \in \mathbb{N} | \text{if } \{(i_k \leftrightarrow j_k)\}_{k \in \{,\ldots,m\}}$ is a family of **strict** transpositions then $((i_1 \underset{n}{\leftrightarrow} j_1) \circ \cdots \circ (i_m \underset{n}{\leftrightarrow} j_m))L = (-1)^m L\}$ then we have:

$1 \in S$.

$$\begin{array}{rcl} ((i_1 \underset{n}{\leftrightarrow} j_1) \circ \cdots \circ (i_1 \underset{n}{\leftrightarrow} j_1))L & = & (i_1 \underset{n}{\leftrightarrow} j_1)L \\ & \underset{[\text{eq: } 11.150]}{=} & -L \\ & = & (-1)^1 \cdot L \end{array}$$

proving that $1 \in S$.

$\boldsymbol{m\in S\Rightarrow m+1\in S}$. We have

$$\begin{array}{rc}
\left(\left(i_1\underset{n}{\leftrightarrow}j_1\right)\circ\cdots\circ\left(i_{m+1}\underset{n}{\leftrightarrow}j_{m+1}\right)\right)L & = \\
\left(\left(\left(i_1\underset{n}{\leftrightarrow}j_1\right)\circ\cdots\circ,\left(i_m\underset{n}{\leftrightarrow}j_m\right)\right)\circ\left(i_{m+1}\leftrightarrow\right)\right)L & \underset{\text{[theorem: 11.222]}}{=} \\
\left(\left(i_1\underset{n}{\leftrightarrow}j_1\right)\circ\cdots\circ,\left(i_m\underset{n}{\leftrightarrow}j_m\right)\right)\left(\left(i_{m+1}\leftrightarrow\right)L\right) & \underset{\text{[eq: 11.150]}}{=} \\
\left(\left(i_1\underset{n}{\leftrightarrow}j_1\right)\circ\cdots\circ,\left(i_m\underset{n}{\leftrightarrow}j_m\right)\right)(-L) & = \\
\left(\left(i_1\underset{n}{\leftrightarrow}j_1\right)\circ\cdots\circ,\left(i_m\underset{n}{\leftrightarrow}j_m\right)\right)((-1)\cdot L) & \underset{\text{[proposition: 11.223]}}{=} \\
(-1)\cdot\left(\left(\left(i_1\underset{n}{\leftrightarrow}j_1\right)\circ\cdots\circ,\left(i_m\underset{n}{\leftrightarrow}j_m\right)\right)L\right) & \underset{m\in S}{=} \\
(-1)\cdot((-1)^m\cdot L) & = \\
((-1)\cdot(-1)^m)\cdot L & = \\
(-1)^{m+1}\cdot L &
\end{array}$$

proving that $m+1\in S$. □

Let $\sigma\in P_n$ then by [theorem: 11.234] there exists a $\left\{\left(i_k\underset{n}{\leftrightarrow}j_k\right)\right\}_{k\in\{1,\ldots,m\}}$ of **strict** transpositions such that

$$\sigma=\left(i_1\underset{n}{\leftrightarrow}j_1\right)\circ\cdots\circ(i_m\leftrightarrow j_m)\text{ and }\operatorname{sign}(\sigma)=(-1)^m$$

So

$$\sigma L=\left(\left(i_1\underset{n}{\leftrightarrow}j_1\right)\circ\cdots\circ(i_m\leftrightarrow j_m)\right)L\underset{\text{[eq: 11.151]}}{=}(-1)^m\cdot L=\operatorname{sign}(\sigma)\cdot L$$

2 ⇒ 3. As $\{x_i\}_{i\in\{1,\ldots,n\}}\subseteq X$ is linear dependent there exist by [theorem: 11.123] a $k\in\{1,...,n\}$ and a $\{\alpha_i\}_{i\in\{1,\ldots,n\}\setminus\{k\}}\subseteq F$ such that

$$x_k=\sum_{i\in\{1,\ldots,n\}\setminus\{k\}}\alpha_i\cdot x_i \tag{11.152}$$

If $i\in\{1,\ldots,n\}\setminus\{k\}\Rightarrow i\neq k$ we have

$$\Bigg(x_1,\ldots\underbrace{x_i}_{k}\ldots,x_n\Bigg)_k=x_i$$

so that by (2) $L\Bigg(x_1,\ldots\underbrace{x_i}_{k}\ldots,x_n\Bigg)=0$. Hence

$$\forall i\in\{1,\ldots,n\}\setminus\{k\}\text{ we have }L\Bigg(x_1,\ldots\underbrace{x_i}_{k}\ldots,x_n\Bigg)=0 \tag{11.153}$$

Finally

$$\begin{array}{rcl}
L(x) & = & L\Bigg(x_1,\ldots\underbrace{x_k}_{k}\ldots,x_n\Bigg) \\
& \underset{\text{[eq: 11.152]}}{=} & L\Bigg(x_1,\ldots\underbrace{\sum_{i\in\{1,\ldots,n\}\setminus\{k\}}\alpha_i\cdot x_i}_{k}\ldots,x_n\Bigg) \\
& \underset{\text{[theorems: 11.250, 11.180]}}{=} & \sum_{i\in\{1,\ldots,n\}\setminus\{k\}}\alpha_i\cdot L\Bigg(x_1,\ldots\underbrace{x_i}_{k}\ldots,x_n\Bigg) \\
& \underset{\text{[eq: 11.153]}}{=} & 0
\end{array}$$

3 ⇒ 2. If $x=(x_1,\ldots,x_n)\in X^n$ is such that $\exists i,j\in\{1,\ldots,n\}$ where $i\neq j$ and $x_i=x_j$ then by [theorem: 11.124] $\{x_i\}_{i\in\{1,\ldots,n\}}$ is linear dependent, hence by (3) $L(x)=0$ □

Theorem 11.276. *Let X,Y be vector spaces over a field F with characteristic zero [see definition: 10.47], $n\in\mathbb{N}$ and $L\in\operatorname{Hom}^n(X;Y)$ a skew-symmetric n-linear mapping then if $x=(x_1,\ldots,x_n)\in X^n$ is such that $\{x_i|i\in\{1,\ldots,n\}\}$ is a linear dependent set then $L(x)=0$*

Proof. Let $x=(x_1,\ldots,x_n)\in X^n$ such that $W=\{x_i|i\in\{1,\ldots,n\}\}$ is linear dependent, then as W is finite and linear dependent there exists by [theorem: 11.112] a $u\in W$ and a $\{\alpha_w\}_{w\in W\setminus\{u\}}\subseteq F$ such that

$$u=\sum_{w\in W\setminus\{u\}}\alpha_w\cdot w \tag{11.154}$$

As $u \in W = \{x_i | i \in \{1, \ldots, n\}\}$ there exists a $k \in \{1, \ldots, n\}$ such that $x_k = u$. Let $w \in W \setminus \{u\}$ then $\exists i \in \{1, \ldots, n\}$ such that $w = x_i$ and $i \neq k$ [for if $i = k$ then $x_i = x_k = u$] hence by [theorem: 11.275] $L(x_{1,\ldots}, x_{k-1}, w, x_{k+1}, \ldots, x_n) = 0$. So

$$\exists k \in \{1, \ldots, n\} \text{ so that } u = x_k \text{ and } \forall w \in W \setminus \{u\} \;\; L\left(x_1, \ldots \underbrace{w}_{k} \ldots, x_n\right) = 0 \tag{11.155}$$

Finally

$$\begin{aligned}
L(x) &= L\left(x_1, \ldots \underbrace{x_k}_{k} \ldots, x_n\right)\\
&\underset{\text{[eq: 11.154]}}{=} L\left(x_1, \ldots \underbrace{u}_{k} \ldots, x_n\right)\\
&\underset{\text{[eq: 11.155]}}{=} L\left(x_1, \ldots \underbrace{\sum_{w \in W \setminus \{u\}} \alpha_w \cdot w}_{k} \ldots, x_n\right)\\
&\underset{\text{[theorems: 11.250, 11.180]}}{=} \sum_{w \in W \setminus \{u\}} \alpha_w \cdot L\left(x_1, \ldots \underbrace{w}_{k} \ldots, x_n\right)\\
&\underset{\text{[eq: 11.153]}}{=} 0
\end{aligned}$$

□

Theorem 11.277. *Let X, Y be vector spaces over a field F with characteristic zero, $n \in \mathbb{N}$ and $L: X^n \rightarrow Y$ a skew-symmetric n-linear mapping. If X is finite dimensional with $\dim(X) < n$ then*

$$\forall x \in X \text{ we have } L(x) = 0$$

in another words

$$L = C_0$$

Proof. For $n \in \mathbb{N}$ we have either:

$n = 1$. Then, as $\dim(X) < n$, it follows that $\dim(X) = 0$ so that by [example: 11.147] $X = \{0\}$. Hence if $L \in \mathrm{Hom}^1(X; Y)$ we have for $x \in X^1$ that $x = (0)$ so that $L(x) = L(0) \underset{\text{[theorem: 11.243]}}{=} 0$ proving the theorem.

$1 < n$. Let $x = (x_1, \ldots, x_n) \in X^n$ defined by the function $x: \{1, \ldots, n\} \rightarrow X^n$ then we have two cases to consider:

***x* is injective.** $x: \{1, \ldots, n\} \rightarrow \{x_i | i \in \{1, \ldots, n\}\}$ is a bijection, so we have by [definition: 11.79] that

$$\{x_i | i \in \{1, \ldots, n\}\} = \{x_1, \ldots, x_n\} \tag{11.156}$$

As X is finite dimensional with $\dim(X) < n$ there exist a $B \subseteq X$ with $\mathrm{card}(B) < n$ and $X = \mathrm{span}(B)$. Assume that $\{x_i | i \in \{1, \ldots, n\}\}$ is linear independent then using Steinitz lemma [see lemma: 11.120] together with [eq: 11.156] we have that $n \leqslant \mathrm{card}(B)$ which, as $\mathrm{card}(B) < n$, leads to the contradiction $n < n$. Hence we must have that $\{x_i | i \in \{1, \ldots, n\}\}$ is linear dependent. So as L is skew-symmetric it follows from [theorem: 11.276] that

$$L(x) = 0$$

***x* is not injective.** Then there exist $i, j \in \{1, \ldots, n\}$ with $x_i = x_j$ such that $i \neq j$. So by [theorem: 11.276] we have

$$L(x) = 0$$

So in all cases we have $\forall x \in X$ that $L(x) = 0$. □

Theorem 11.278. *Let $n \in \mathbb{N}$, X, Y vector spaces over the field F, $E = \{e_1, \ldots, e_n\} \subseteq X$ a basis for X and $L \in \mathrm{Hom}^n(X; Y)$ such that L is skew-symmetric then $\forall x \in X$ we have:*

$$L(x_1, \ldots, x_n) = \sum_{\sigma \in P_n} \left(\prod_{j \in \{1, \ldots, n\}} \alpha_j^{\sigma(j)}\right) \cdot (\sigma L)(e_1, \ldots, e_n)$$

where $\forall i \in \{1, \ldots, n\}$ $\{\alpha_j^i\}_{j \in \{1, \ldots, n\}}$ is the ***unique family*** *satisfying $x_i = \sum_{j \in \{1, \ldots, n\}} \alpha_i^j \cdot e_j$ [which exists by [theorem: 11.137]*

Proof. As $P_n = \{\sigma \in \{1, \ldots, n\}^{\{1, \ldots, n\}} | \sigma: \{1, \ldots, n\} \rightarrow \{1, \ldots, n\} \text{ is a bijection}\}$ we have

$$P_n \subseteq \{1, \ldots, n\}^{\{1, \ldots, n\}} \underset{\text{[definition: 6.82]}}{=} \{1, \ldots, n\}^n.$$

Take $\sigma \in \{1, \ldots, n\}^n \setminus P_n$ then we have that $\sigma: \{1, \ldots, n\} \rightarrow \{1, \ldots, n\}$ is not a bijection. For $\sigma(\{1, \ldots, n\})$ we have two cases to consider:

$\sigma(\{1, \ldots, n\}) = \{1, \ldots, n\}$. Then σ can not be injective [otherwise $\sigma: \{1, \ldots, n\} \rightarrow \{1, \ldots, n\}$ is a bijection].

$\sigma(\{1, \ldots, n\}) \neq \{1, \ldots, n\}$. Then as $\sigma(\{1, \ldots, n\}) \subseteq \{1, \ldots, n\}$ we have $\sigma(\{1, \ldots, n\}) \subset \{1, \ldots, n\}$ so that by [theorem: 10.89]

$$\operatorname{card}(\sigma(\{1, \ldots, n\})) < \operatorname{card}(\{1, \ldots, n\}) = n$$

Assume that σ is injective then $\sigma: \{1, \ldots, n\} \rightarrow \sigma(\{1, \ldots, n\})$ is a bijection so that $\{1, \ldots, n\} \approx \sigma(\{1, \ldots, n\})$ or $\operatorname{card}(\sigma(\{1, \ldots, n\})) = n$ contradicting $\operatorname{card}(\sigma(\{1, \ldots, n\})) < n$. So σ is not injective.

As σ is not injective, there exists $i, j \in \{1, \ldots, n\}$ with $i \neq j$ such that $\sigma(i) = \sigma(j)$ or $e_{\sigma(i)} = e_{\sigma(j)}$. Hence by [theorem: 11.275 (2)] we have $L(e_{\sigma(1)}, \ldots, e_{\sigma(n)}) = 0$. To summarize

$$\text{If } \sigma \in \{1, \ldots, n\} \setminus P_n \text{ then } L(e_{\sigma(1)}, \ldots, e_{\sigma(n)}) = 0 \tag{11.157}$$

Now if $x \in X$ we have

$$\begin{aligned}
L(x_1, \ldots, x_n) &\underset{[\text{theorem: } 11.256]}{=} \sum_{\sigma \in \{1, \ldots, n\}^n} \left(\prod_{j \in \{1, \ldots, n\}} \alpha_j^{\sigma_j} \right) \cdot L(e_{\sigma_1}, \ldots, e_{\sigma_n}) \\
&= \sum_{\sigma \in \{1, \ldots, n\}^n} \left(\prod_{j \in \{1, \ldots, n\}} \alpha_j^{\sigma(j)} \right) \cdot L(e_{\sigma(1)}, \ldots, e_{\sigma(n)}) \\
&= \sum_{\sigma \in \{1, \ldots, n\}^n \setminus P_n} \left(\prod_{j \in \{1, \ldots, n\}} \alpha_j^{\sigma(j)} \right) \cdot L(e_{\sigma(1)}, \ldots, e_{\sigma(n)}) + \sum_{\sigma \in P_n} \left(\prod_{j \in \{1, \ldots, n\}} \alpha_j^{\sigma(j)} \right) \cdot L(e_{\sigma(1)}, \ldots, e_{\sigma(n)}) \\
&= \sum_{\sigma \in \{1, \ldots, n\}^n \setminus P_n} \left(\prod_{j \in \{1, \ldots, n\}} \alpha_j^{\sigma(j)} \right) \cdot 0 + \sum_{k \in P_n} \left(\prod_{j \in \{1, \ldots, n\}} \alpha_j^{\sigma(j)} \right) \cdot L(e_{\sigma(1)}, \ldots, e_{\sigma(n)}) \\
&= \sum_{\sigma \in P_n} \left(\prod_{j \in \{1, \ldots, n\}} \alpha_j^{\sigma(j)} \right) \cdot L(e_{\sigma(1)}, \ldots, e_{\sigma(n)}) \\
&= \sum_{\sigma \in P_n} \left(\prod_{j \in \{1, \ldots, n\}} \alpha_j^{\sigma(j)} \right) \cdot (\sigma L)(e_1, \ldots, e_n)
\end{aligned}$$

□

Corollary 11.279. *Let $n \in \mathbb{N}$, X, Y vector spaces over the field F, $E = \{e_1, \ldots, e_n\} \subseteq X$ a basis for X and $L \in \operatorname{Hom}^n(X; Y)$ a L a skew-symmetric n-linear mapping then $\forall x \in X$ we have*

$$L(x_1, \ldots, x_n) = \left(\sum_{\sigma \in P_n} \operatorname{sign}(\sigma) \cdot \prod_{j \in \{1, \ldots, n\}} \alpha_j^{\sigma(j)} \right) \cdot L(e_1, \ldots, e_n)$$

where $\forall i \in \{1, \ldots, n\}$ $\{\alpha_j^i\}_{j \in \{1, \ldots, n\}}$ is the **unique family** satisfying $x_i = \sum_{j \in \{1, \ldots, n\}} \alpha_i^j \cdot e_j$ [which exists by [theorem: 11.137].

Proof. By [theorem: 11.278] we have

$$\begin{aligned}
L(x_1, \ldots, x_n) &= \sum_{\sigma \in P_n} \left(\prod_{j \in \{1, \ldots, n\}} \alpha_j^{\sigma(j)} \right) \cdot (\sigma L)(e_1, \ldots, e_n) \\
&\underset{L \text{ is skew-symmetric}}{=} \sum_{\sigma \in P_n} \left(\prod_{j \in \{1, \ldots, n\}} \alpha_j^{\sigma(j)} \right) \cdot (\operatorname{sign}(\sigma) \cdot L(e_1, \ldots, e_n)) \\
&= \sum_{\sigma \in P_n} \operatorname{sign}(\sigma) \cdot \left(\prod_{j \in \{1, \ldots, n\}} \alpha_j^{\sigma(j)} \right) \cdot L(e_1, \ldots, e_n)) \\
&\underset{[\text{theorem: } 11.75]}{=} \left(\sum_{\sigma \in P_n} \operatorname{sign}(\sigma) \cdot \left(\prod_{j \in \{1, \ldots, n\}} \alpha_j^{\sigma(j)} \right) \right) \cdot L(e_1, \ldots, e_n)
\end{aligned}$$

□

The above theorem proves that a skew-symmetric n-linear depends only on its values at $(e_1, \ldots, e_n)$ as is expressed in the following corollary.

Corollary 11.280. *Let $n\in\mathbb{N}$, X,Y vector spaces over the field F, $E=\{e_1,\ldots,e_n\}\subseteq X$ a basis for X then we have:*

1. *If $L_1,L_2\in\operatorname{Hom}^n(X;Y)$ are skew-symmetric n-linear mappings such that*
$$L_1(e_1,\ldots,e_n)=L_2(e_1,\ldots,e_n)$$
then
$$L_1=L_2$$
2. *If $L\in\operatorname{Hom}^n(X;Y)$ is a skew-symmetric n-linear mapping such that $L(e_1,\ldots,e_n)=0$ then*
$$L=C_0$$

Proof.

1. Using [theorem: 11.279] we have for $x\in X$
$$\begin{array}{rcl} L_1(x) & = & L_1(x_1,\ldots,x_n)\\ & = & \left(\sum_{\sigma\in P_n}\left(\prod_{j\in\{1,\ldots,m\}}\alpha_j^{\sigma(j)}\right)\cdot\operatorname{sign}(\sigma)\right)\cdot L_1(e_1,\ldots,e_n)\\ & = & \left(\sum_{\sigma\in P_n}\left(\prod_{j\in\{1,\ldots,m\}}\alpha_j^{\sigma(j)}\right)\cdot\operatorname{sign}(\sigma)\right)\cdot L_2(e_1,\ldots,e_n)\\ & = & L_2(x_1,\ldots,x_n)\\ & = & L_2(x)\end{array}$$
where $\forall i\in\{1,\ldots,n\}$ $\{\alpha_j^i\}_{j\in\{1,\ldots,n\}}$ is the **unique family** satisfying $x_i=\sum_{j\in\{1,\ldots,n\}}\alpha_i^j\cdot e_j$ [which exist by [theorem: 11.137]. Proving that
$$L_1=L_2$$
2. Using [theorem: 11.279] we have for $x\in X$ that
$$\begin{array}{rcl} L(x) & = & L(x_1,\ldots,x_n)\\ & = & \left(\sum_{\sigma\in P_n}\left(\prod_{j\in\{1,\ldots,m\}}\alpha_j^{\sigma(j)}\right)\cdot\operatorname{sign}(\sigma)\right)\cdot L_1(e_1,\ldots,e_n)\\ & = & \left(\sum_{\sigma\in P_n}\left(\prod_{j\in\{1,\ldots,m\}}\alpha_j^{\sigma(j)}\right)\cdot\operatorname{sign}(\sigma)\right)\cdot 0\\ & = & 0\end{array}$$
where $\forall i\in\{1,\ldots,n\}$ $\{\alpha_j^i\}_{j\in\{1,\ldots,n\}}$ is the **unique family** satisfying $x_i=\sum_{j\in\{1,\ldots,n\}}\alpha_i^j\cdot e_j$ [which exists by [theorem: 11.137]. So that
$$L=C_0$$ □

We are now ready to define determinant functions.

Definition 11.281. (determinant function) *Let $n\in\mathbb{N}$, X a n-dimensional vector space over a field with characteristic zero then a skew-symmetric mapping $\Delta\in\operatorname{Hom}^n(X;F)$ is called a **determinant function** or **determinant mapping**.*

Example 11.282. Let $n\in\mathbb{N}$, X a n-dimensional vector space over a field with characteristic zero then $C_0: X^n\rightarrow F$ defined by $C_0(x)=0$ is a determinant function. This determinant function is called the trivial determinant function.

Proof. If $i\in\{1,\ldots,n\}$ $u,v\in X_i$ then
$$\begin{array}{rcl} C_0\left(x_1,\ldots\underbrace{u+v}_{i}\ldots,x_n\right) & = & \\ 0 & = & \\ 0+0 & = & \\ C_0\left(x_1,\ldots\underbrace{u}_{i}\ldots,x_n\right)+C_0\left(x_1,\ldots\underbrace{v}_{i}\ldots,x_n\right) & = & \end{array}$$

and

$$\begin{aligned} C_0\Big(x_1,\ldots\underbrace{\alpha\cdot u}_{i}\ldots,x_n\Big) &= 0 \\ &= \alpha\cdot 0 \\ &= \alpha\cdot C_0\Big(x_1,\ldots\underbrace{u}_{i}\ldots,x_n\Big) \end{aligned}$$

So

$$C_0\in \mathrm{Hom}^n(X;Y)$$

Further if $\sigma\in P_n$ then for $x\in X$ we have

$$(\sigma C_0)(x)=C_0(x\circ\sigma)=0=\mathrm{sign}(\sigma)\cdot 0=\mathrm{sign}(\sigma)\cdot C_0(x)$$ □

The following theorem shows that there exist non trivial determinant function.

Theorem 11.283. *Let $n\in\mathbb{N}$, X a n-dimensional vector space over a field F of characteristic zero and $E=\{e_1,\ldots,e_n\}\subseteq X$ a basis for X then there exists a determinant function $\Delta\in\mathrm{Hom}\,(X^n;F)$ such that $\Delta(e_1,\ldots,e_n)=1$. This proves that there exist a non trivial determinant function.*

Proof. Let $i\in\{1,\ldots,n\}$ and define $f_i: X\rightarrow F$ by $f_i(x)=\alpha_i$ where $\{\alpha_i\}_{i\in\{1,\ldots,n\}}\subseteq F$ is the **unique** family such that $x=\sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot e_i$ [see theorem: 11.137]. Then we have that

$$x=\sum_{i\in\{1,\ldots,n\}} f_i(x)\cdot e_i$$

Further if $x,y\in X$ and $\alpha\in F$ then we have

$$\begin{aligned} x+y &\underset{\text{def } f_i}{=} \sum_{i\in\{1,\ldots,n\}} f_i(x)\cdot e_i+\sum_{i\in\{1,\ldots,n\}} f_i(y)\cdot e_i \\ &\underset{[\text{theorem: } 11.38]}{=} \sum_{i\in\{1,\ldots,n\}} (f_i(x)\cdot e_i+f_i(y)\cdot e_i) \\ &= \sum_{i\in\{1,\ldots,n\}} (f_i(x)+f_i(y))\cdot e_i \\ x+y &\underset{\text{def } f_i}{=} \sum_{i\in\{1,\ldots,n\}} f_i(x+y)\cdot e_i \end{aligned}$$

so that by uniqueness we have

$$\forall i\in\{1,\ldots,n\} \text{ that } f_i(x+y)=f_i(x)+f_i(y) \tag{11.158}$$

Likewise

$$\begin{aligned} \alpha\cdot x &\underset{\text{def } f_i}{=} \alpha\cdot\sum_{i\in\{1,\ldots,n\}} f_i(x)\cdot e_i \\ &\underset{[\text{theorem: } 11.75]}{=} \sum_{i\in\{1,\ldots,n\}} \alpha\cdot f_i(x)\cdot e_i \\ &= \sum_{i\in\{1,\ldots,n\}} (\alpha\cdot f_i(x))\cdot e_i \\ \alpha\cdot x &\underset{\text{def } f_i}{=} \sum_{i\in\{1,\ldots,n\}} f_i(\alpha\cdot x) \end{aligned}$$

so by uniqueness we have

$$\forall i\in\{1,\ldots,n\}\ f_i(\alpha\cdot x)=\alpha\cdot f_i(x) \tag{11.159}$$

Let $j\in\{1,\ldots,n\}$ then we have by [theorem: 11.157] that $e_j=\sum_{i\in\{1,\ldots,n\}}\delta_{i,j}\cdot e_j$, which as $e_j=\sum_{i\in\{1,\ldots,n\}} f_i(e_j)\cdot e_i$, proving by uniqueness

$$\forall i\in\{1,\ldots,n\}\ f_i(e_j)=\delta_{i,j} \tag{11.160}$$

Define

$$\Psi: X^n\rightarrow F \text{ by } \Psi(x_1,\ldots,x_n)=\prod_{i\in\{1,\ldots,n\}} f_i(x_i)$$

then we have for $i \in \{1, \ldots, n\}$, $u, v \in X_i$ and $\alpha \in F$ that for

$$y = \left(x_1, \ldots \underbrace{u+v}_{i} \ldots, x_n\right)$$

$$z = \left(x_1, \ldots \underbrace{u}_{i} \ldots, x_n\right)$$

$$r = \left(x_1, \ldots \underbrace{v}_{i} \ldots, x_n\right)$$

$$s = \left(x_1, \ldots \underbrace{\alpha \cdot u}_{i} \ldots, x_n\right)$$

$$\begin{aligned}
\Psi\left(x_1, \ldots \underbrace{u+v}_{i} \ldots, x_n\right) &= \prod_{j \in \{1, \ldots, n\}} y_j \\
&= \left(\prod_{j \in \{i\}} y_j\right) \cdot \prod_{j \in \{1, \ldots, n\} \setminus \{i\}} y_j \Bigg) \\
&= y_i \cdot \prod_{j \in \{1, \ldots, n\} \setminus \{i\}} y_j \\
&= (u+v) \cdot \prod_{j \in \{1, \ldots, n\} \setminus \{i\}} y_j \\
&= u \cdot \prod_{j \in \{1, \ldots, n\} \setminus \{i\}} y_j + v \cdot \prod_{j \in \{1, \ldots, n\} \setminus \{i\}} y_j \\
&= z_i \cdot \prod_{j \in \{1, \ldots, n\} \setminus \{i\}} z_j + r_i \cdot \prod_{j \in \{1, \ldots, n\} \setminus \{i\}} r_j \\
&= \prod_{j \in \{1, \ldots, n\}} z_j + \prod_{j \in \{1, \ldots, n\}} r_j \\
&= L\left(x_1, \ldots \underbrace{u}_{i} \ldots, x_n\right) + L\left(x_1, \ldots \underbrace{v}_{i} \ldots, x_n\right)
\end{aligned}$$

and

$$\begin{aligned}
\Psi\left(x_1, \ldots \underbrace{\alpha \cdot u}_{i} \ldots, x_n\right) &= \prod_{j \in \{1, \ldots, n\}} s_j \\
&= \left(\prod_{j \in \{i\}} s_j\right) \cdot \prod_{j \in \{1, \ldots, n\} \setminus \{i\}} s_j \Bigg) \\
&= s_j \cdot \prod_{j \in \{1, \ldots, n\} \setminus \{i\}} s_j \\
&= (\alpha \cdot u) \cdot \prod_{j \in \{1, \ldots, n\} \setminus \{i\}} s_j \\
&= \alpha \cdot \left(u \cdot \prod_{j \in \{1, \ldots, n\} \setminus \{i\}} s_j\right) \\
&= \alpha \cdot \left(z_i \cdot \prod_{j \in \{1, \ldots, n\} \setminus \{i\}} z_j\right) \\
&= \alpha \cdot \prod_{j \in \{1, \ldots, n\}} z_j \\
&= \alpha \cdot L\left(x_1, \ldots \underbrace{u}_{i} \ldots, x_n\right)
\end{aligned}$$

So we have that

$$\Psi \in \mathrm{Hom}^n(X; F) \text{ or } \Psi \text{ is } n\text{-linear}$$

Using [theorem: 11.273] we create the following skew symmetric n-linear mapping

$$\Delta: X^n \rightarrow F \text{ by } \Delta = \sum_{\sigma \in P_n} \operatorname{sign}(\sigma) \cdot \sigma\Psi \tag{11.161}$$

which is by definition a determinant function. Let $\sigma \in P_n$ then we have either:

$\sigma = \mathrm{Id}_{\{1,\ldots,n\}}$. Then

$$\sigma\Psi(e_1,\ldots,e_n) = \Psi(e_1,\ldots,e_n) = \prod_{i \in \{1,\ldots,n\}} f_i(e_i) \underset{\forall i \in I f_i(i) = \delta_{i,i} = 1}{=} 1$$

$\sigma \in P_n \setminus \{\mathrm{Id}_{\{1,\ldots,n\}}\}$. Then there exists a $i \in \{1,\ldots,n\}$ such that $i \neq \mathrm{Id}_{\{1,\ldots,n\}}(i) = \sigma(i)$ so that $f_i(e_{\sigma(i)}) \underset{[\text{eq: } 11.160]}{=} \delta_{i,\sigma(i)} = 0$ then

$$\begin{aligned}
\sigma\Psi(e_1,\ldots,e_n) &= \Psi(e_{\sigma(1)},\ldots,e_{\sigma(n)}) \\
&= \prod_{j \in \{1,\ldots,n\}} f_j(e_{\sigma(j)}) \\
&= \left(\prod_{j \in \{i\}} f_j(e_{\sigma(j)})\right) \cdot \prod_{j \in \{1,\ldots,n\} \setminus \{i\}} f_j(e_{\sigma(j)}) \\
&= f_i(e_{\sigma(i)}) \cdot \prod_{j \in \{1,\ldots,n\} \setminus \{i\}} f_j(e_{\sigma(j)}) \\
&= 0 \cdot \prod_{j \in \{1,\ldots,n\} \setminus \{i\}} f_j(e_{\sigma(j)}) \\
&= 0
\end{aligned}$$

So that

$$\sigma\Psi(e_1,\ldots,e_n) = \begin{cases} 1 \text{ if } \sigma = \mathrm{Id}_{\{1,\ldots,n\}} \\ 0 \text{ if } \sigma \in P_n \setminus \{\mathrm{Id}_{\{1,\ldots,n\}}\} \end{cases} \tag{11.162}$$

so that

$$\begin{aligned}
\Delta(e_1,\ldots,e_n) &\underset{[\text{eq: } 11.161]}{=} \\
\sum_{\sigma \in P_n} \operatorname{sign}(\sigma) \cdot (\sigma\Psi)(e_1,\ldots,e_n) &\underset{[\text{theorem: } 11.43]}{=} \\
\sum_{\sigma \in P_n \setminus \{\mathrm{Id}_{\{1,\ldots,n\}}\}} \operatorname{sign}(\sigma) \cdot (\sigma\Psi)(e_1,\ldots,e_n) + \sum_{\sigma \in \{\mathrm{Id}_{\{1,\ldots,n\}}\}} \operatorname{sign}(\sigma) \cdot (\sigma\Psi)(e_1,\ldots,e_n) &\underset{[\text{eq: } 11.162]}{=} \\
\sum_{\sigma \in P_n \setminus \{\mathrm{Id}_{\{1,\ldots,n\}}\}} \operatorname{sign}(\sigma) \cdot 0 + \sum_{\sigma \in \{\mathrm{Id}_{\{1,\ldots,n\}}\}} \operatorname{sign}(\sigma) \cdot (\sigma\Psi)(e_1,\ldots,e_n) &= \\
\sum_{\sigma \in \{\mathrm{Id}_{\{1,\ldots,n\}}\}} \operatorname{sign}(\sigma) \cdot (\sigma\Psi)(e_1,\ldots,e_n) &\underset{[\text{theorem: } 11.34]}{=} \\
\operatorname{sign}(\mathrm{Id}_{\{1,\ldots,n\}}) \cdot (\mathrm{Id}_{\{1,\ldots,n\}}\Psi)(e_1,\ldots,e_n) &\underset{[\text{eq: } 11.162]}{=} \\
1 \cdot 1 &\underset{[\text{eq: } 11.162]}{=} \\
1 &
\end{aligned}$$

proving that

$$\Delta(e_1,\ldots,e_n) = 1$$

□

We show now that every skew-symmetric n-linear function can be written as the scalar product of a determinant function and a vector in the target domain.

Theorem 11.284. *Let $n \in \mathbb{N}$, X a vector space over a field F of characteristic zero, Y a vector space over the same field F and Δ a non zero determinant function. Then for every skew-symmetric $L \in \operatorname{Hom}^n(X;Y)$ there exist a* ***unique*** *$y \in Y$ such that*

$$\forall x \in X \text{ we have } L(x) = \Delta(x) \cdot y$$

Proof. Let $E = \{e_1,\ldots,e_n\} \subseteq X$ be a basis of X. By [theorem: 11.279]

$$\Delta(x_1,\ldots,x_n) = \left(\sum_{\sigma \in P_n} \operatorname{sign}(\sigma) \cdot \prod_{j \in \{1,\ldots,n\}} \alpha_j^{\alpha_j}\right) \cdot \Delta(e_1,\ldots,e_n)$$

so as $\Delta \neq C_0$ we must have that

$$\Delta(e_1, \ldots, e_n) \neq 0$$

Define

$$\{\alpha_i\}_{i \in \{1, \ldots, n\}} \subseteq F \setminus \{0\} \text{ by } \alpha_i = \begin{cases} \Delta(e_1, \ldots, e_n)^{-1} \text{ if } i = 1 \\ 1 \text{ if } i \in \{2, \ldots, n\} \end{cases}$$

then by [theorem: 11.138] we have that

$$h: \{1, \ldots, n\} \rightarrow \{\alpha_i \cdot e_i | i \in \{1, \ldots, n\}\} \text{ defined by } h(i) = \alpha_i \cdot e_i = \begin{cases} \Delta(e_1, \ldots, e_n)^{-1} \cdot e_1 \text{ if } i = 1 \\ e_i \text{ if } i \in \{2, \ldots, n\} \end{cases}$$

is a bijection defining the finite set

$$H = \{h_1, \ldots, h_n\} = \{\Delta(e_1, \ldots, e_n)^{-1} \cdot e_1, e_2, \ldots, e_n\}$$

and that

$$H = \{\Delta(e_1, \ldots, e_n)^{-1} \cdot e_1, e_2, \ldots, e_n\} \text{ is a basis of } X \tag{11.163}$$

Now

$$\begin{aligned} \Delta(h_1, \ldots, h_n) &= \Delta(\Delta(e_1, \ldots, e_n)^{-1} \cdot e_1, e_2, \ldots, e_n) \\ &\underset{\Delta \in \mathrm{Hom}^n(X;Y)}{=} \Delta(e_1, \ldots, e_n)^{-1} \cdot \Delta(e_1, \ldots, e_n) \\ &= 1 \end{aligned} \tag{11.164}$$

Set

$$y = L(h_1, \ldots, h_n) \tag{11.165}$$

and define

$$K: X^n \rightarrow Y \text{ by } K(x_1, \ldots, x_n) = \Delta(x_1, \ldots, x_n) \cdot y \tag{11.166}$$

Then for $i \in \{1, \ldots, n\}$, $u, v \in X$ we have that

$$\begin{aligned} K\left(x_1, \ldots \underbrace{u+v}_{i} \ldots, x_n\right) &= \\ \Delta\left(x_1, \ldots \underbrace{u+v}_{i} \ldots, x_n\right) \cdot y &= \\ \left(\Delta\left(x_1, \ldots \underbrace{u}_{i} \ldots, x_n\right) + \Delta\left(x_1, \ldots \underbrace{v}_{i} \ldots, x_n\right)\right) \cdot y &= \\ \Delta\left(x_1, \ldots \underbrace{u}_{i} \ldots, x_n\right) \cdot y + \Delta\left(x_1, \ldots \underbrace{v}_{i} \ldots, x_n\right) \cdot y &= \\ K\left(x_1, \ldots \underbrace{u}_{i} \ldots, x_n\right) + K\left(x_1, \ldots \underbrace{v}_{i} \ldots, x_n\right) & \end{aligned}$$

and

$$\begin{aligned} K\left(x_1, \ldots \underbrace{\alpha \cdot u}_{i} \ldots, x_n\right) &= \Delta\left(x_1, \ldots \underbrace{\alpha \cdot u}_{i} \ldots, x_n\right) \cdot y \\ &= \left(\alpha \cdot \Delta\left(x_1, \ldots \underbrace{u}_{i} \ldots, x_n\right)\right) \cdot y \\ &= \alpha \cdot \left(\Delta\left(x_1, \ldots \underbrace{u}_{i} \ldots, x_n\right) \cdot y\right) \\ &= \alpha \cdot K\left(x_1, \ldots \underbrace{u}_{i} \ldots, x_n\right) \end{aligned}$$

which proves that

$$K \in \mathrm{Hom}^n(X; Y)$$

Further if $\sigma \in P_n$ then if $x \in X^n$ we have

$$(\sigma K)(x) = K(x \circ \sigma) = \Delta(x \circ \sigma) \cdot y = (\mathrm{sign}(\sigma) \cdot \Delta(x)) \cdot y = \mathrm{sign}(\sigma) \cdot (\Delta(x) \cdot y) = \mathrm{sign}(\sigma) \cdot K(x)$$

proving that

$$K \text{ is a skew-symmetric linear function}$$

As $K(h_1, \ldots, h_n) = \Delta(h_1, \ldots, h_n) \cdot y \underset{[\text{eq: } 11.164]}{=} y = L(h_1, \ldots, h_n)$ we have, as $\{h_1, \ldots, h_n\}$ is a basis, by [theorem: 11.280] that $L = K$ or using the definition of K [see eq: 11.166] we have

$$\forall x \in X \text{ we have } L(x) = \Delta(x) \cdot y$$

Now for uniqueness, assume that there exists a $y' \in Y$ such that $\forall x \in X$ we have $L(x) = \Delta(x) \cdot y'$. Then $\Delta(f_1, \ldots, f_n) \cdot y = \Delta(f_1, \ldots, f_n) \cdot y' \underset{[\text{eq: } 11.164]}{\Rightarrow} 1 \cdot y = 1 \cdot y'$ proving that $y = y'$. □

Corollary 11.285. *Let $n \in \mathbb{N}$, X a n-dimensional vector space over a field F of characteristic zero. Let Δ be a non zero determinant function then if Δ' is another determinant function there exist a $\alpha \in F$ such that*

$$\Delta' = \alpha \cdot \Delta$$

Proof. If Δ' is another determinant function then we have either:

$\Delta = C_0$. Take $\alpha = 0$ then

$$\Delta' = C_0 = 0 \cdot \Delta$$

$\Delta \neq C_0$. As Δ is a determinant function, hence a skew-symmetric n-linear mapping with range F, we have by the previous theorem [theorem: 11.284] that there exist a $\alpha \in F$ such that

$$\Delta' = \Delta \cdot \alpha = \alpha \cdot \Delta$$

□

Definition 11.286. *Let $n \in \mathbb{N}$, X a set, $y \in X$, $i \in \{1, \ldots, n\}$ and $x \in X^n$ then*

$$(y, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \in X^n$$

is defined by

$$(y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_k = \begin{cases} y \text{ if } k = 1 \\ x_{k-1} \text{ if } k \in \{2, \ldots, i\} \\ x_k \text{ if } k \in \{i+1, \ldots, n\} \end{cases}$$

Lemma 11.287. *Let $n \in \mathbb{N}$, X a set, $i \in \{1, \ldots, n\}$and $x \in X^n$ then*

$$(x_i, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) = x \circ \left(i \underset{n}{\rightsquigarrow} 1\right)$$

Proof. For $i \in \{1, \ldots, n\}$ we have either:

$i = 1$. Then by [definition: 11.236] we have that $\left(1 \underset{n}{\rightsquigarrow} i\right) = \left(1 \underset{n}{\rightsquigarrow} 1\right) = \mathrm{Id}_{\{1, \ldots, n\}}$. Then we have for $j \in \{1, \ldots, n\}$

$$\begin{aligned}
(x_i, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_j &= \begin{cases} x_i \text{ if } j = 1 \\ x_{i-1} \text{ if } j \in \{2, \ldots, i\} \\ x_k \text{ if } j \in \{i+1, \ldots, n\} \end{cases} \\
&\underset{i=1}{=} \begin{cases} x_1 \text{ if } j = 1 \\ x_{i-1} \text{ if } j \in \{2, \ldots, 1\} \\ x_k \text{ if } j \in \{2, \ldots, n\} \end{cases} \\
&= \begin{cases} x_1 \text{ if } j = 1 \\ x_{i-1} \text{ if } j \in \{2, \ldots, 1\} \\ x_k \text{ if } j \in \{2, \ldots, n\} \end{cases} \\
&\underset{\{2, \ldots, 1\} = \varnothing}{=} \begin{cases} x_1 \text{ if } j = 1 \\ x_k \text{ if } j \in \{2, \ldots, n\} \end{cases} \\
&= x_j \\
&= (x \circ \mathrm{Id}_{\{1, \ldots, n\}})_j \\
&= \left(x \circ \left(i \underset{n}{\rightsquigarrow} 1\right)\right)(j)
\end{aligned}$$

proving that

$$(x_i, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) = x \circ \left(i \underset{n}{\rightsquigarrow} 1\right)$$

$\boldsymbol{i \in \{2, \ldots, n\}}$**.** Then $1 < i$ and for $j \in \{1, \ldots, n\}$ we have either:

$\boldsymbol{j = 1}$**.** Then

$$\begin{aligned}(x_i, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_j &= (x_i, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_1 \\ &= x_i \\ &= x(i) \\ &\underset{j=1 \wedge [\text{definition: } 11.236(3)]}{=} x\left(\left(i \underset{n}{\rightsquigarrow} 1\right)(j)\right) \\ &= \left(x \circ \left(i \underset{n}{\rightsquigarrow} 1\right)\right)(j)\end{aligned}$$

$\boldsymbol{j \in \{2, \ldots, i\}}$**.** Then

$$\begin{aligned}(x_i, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_j &= x_{j-1} \\ &= x(j-1) \\ &\underset{1 < j \leqslant i \wedge [\text{definition: } 11.236(3)]}{=} x\left(\left(i \underset{n}{\rightsquigarrow} 1\right)(j)\right) \\ &= \left(x \circ \left(i \underset{n}{\rightsquigarrow} 1\right)\right)(j)\end{aligned}$$

$\boldsymbol{j \in \{i+1, \ldots, n\}}$**.** Then

$$\begin{aligned}(x_i, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_j &= x_j \\ &= x(j) \\ &\underset{i < i+1 \leqslant j \leqslant n \wedge [\text{definition: } 11.236(3)]}{=} x\left(\left(i \underset{n}{\rightsquigarrow} 1\right)(j)\right) \\ &= \left(x \circ \left(i \underset{n}{\rightsquigarrow} 1\right)\right)(j)\end{aligned}$$

proving that

$$(x_i, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) = x \circ \left(i \underset{n}{\rightsquigarrow} 1\right)$$

□

Definition 11.288. *Let $n \in \mathbb{N}$, X a n-dimensional vector space over a field F with characteristic zero and $\Delta \in \mathrm{Hom}^n(X; F)$ a determinant function then we define:*

$$\underline{\Delta}: X \times X^n \rightarrow X \text{ where } \underline{\Delta}(y, (x_1, \ldots, x_n)) = \sum_{i \in \{1, \ldots, n\}} (-1)^{i-1} \cdot \Delta(y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \cdot x_i$$

where $(y, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$ is defined by [see the previous definition: 11.286]

$$(y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_k = \begin{cases} y \text{ if } k = 1 \\ x_{k-1} \text{ if } k \in \{2, \ldots, i\} \\ x_k \text{ if } k \in \{i+1, \ldots, n\} \end{cases}$$

To calculate $\underline{\Delta}(y, (x_1, \ldots, x_n))$ we need the following lemma.
TODO check this

Lemma 11.289. *Let $n \in \mathbb{N}$, X a vector space over a field F, $i, j \in \{1, \ldots, n\}$, $y \in X$ and $x = (x_1, \ldots, x_n) \in X^n$ then for $z = (y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$ we have:*

1. *If $j + 1 \leqslant i$ then* $\left(z_1, \ldots \underbrace{x_i}_{j+1} \ldots, z_n\right) \circ \left(j + 1 \underset{n}{\rightsquigarrow} i\right) = (y, x_1, \ldots, x_{j-1}, x_j, \ldots, x_n)$
2. *If $i + 1 \leqslant j$ then* $\left(z_1, \ldots \underbrace{x_i}_{j} \ldots, z_n\right) \circ \left(j \underset{n}{\rightsquigarrow} i + 1\right) = (y, x_1, \ldots, x_{j-1}, x_j, \ldots, x_n)$

Proof. The prove is simple but rather elaborated because we have to check so many cases.

1. Let $k \in \{1, \ldots, n\}$ and take $a_k = \left(z_1, \ldots \underbrace{x_i}_{j+1} \ldots, z_n\right)\left(\left(j + 1 \underset{n}{\rightsquigarrow} i\right)(k)\right)$ then using the definition of $\underset{n}{\rightsquigarrow}$ [see definition: 11.236] we must look at the following cases for $j + 1, i$

   $\boldsymbol{j + 1 = i}$**.** Then by [definition: 11.236(1)] we have that $\left(j + 1 \underset{n}{\rightsquigarrow} i\right) = \mathrm{Id}_{\{1, \ldots, n\}}$ so that

$$a_k = \left(z_1, \ldots \underbrace{x_i}_{j+1} \ldots, z_n\right)(k) \tag{11.167}$$

now for $k \in \{1, \ldots, n\}$ we have to look at the following cases:

$k = 1$. Then

$$\begin{aligned} a_k &\underset{\text{[eq: 11.167]}}{=} \Big(z_1, \ldots \underbrace{x_i}_{j+1} \ldots, z_n\Big)_k \\ &\underset{k=1<j+1}{=} z_k \\ &= (y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_k \\ &\underset{\text{[definition: 11.286]} \wedge k=1}{=} y \\ &\underset{\text{[definition: 11.286]} \wedge k=1}{=} (y, x_1, \ldots, x_{j-1}, x_j, \ldots, x_n)_k \end{aligned}$$

$1 < k < j$. Then

$$\begin{aligned} a_k &\underset{\text{[eq: 11.167]}}{=} \Big(z_1, \ldots \underbrace{x_i}_{j+1} \ldots, z_n\Big)_k \\ &\underset{k<j<j+1}{=} z_k \\ &= (y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_k \\ &\underset{\text{[definition: 11.286]} \wedge 1<k<j<j+1=i}{=} x_{k-1} \\ &\underset{\text{[definition: 11.286]} \wedge 1<k<j}{=} (y, x_1, \ldots, x_{j-1}, x_j, \ldots, x_n)_k \end{aligned}$$

$1 < k = j + 1$. Then

$$\begin{aligned} a_k &\underset{\text{[eq: 11.167]}}{=} \Big(z_1, \ldots \underbrace{x_i}_{j+1} \ldots, z_n\Big)_k \\ &\underset{k=j+1}{=} x_i \\ &\underset{k=j+1=i}{=} x_k \\ &\underset{\text{[definition: 11.286]} \wedge k=j+1}{=} (y, x_1, \ldots, x_{j-1}, x_j, \ldots, x_n)_k \end{aligned}$$

$j + 1 < k \leqslant n$. Then

$$\begin{aligned} a_k &\underset{\text{[eq: 11.167]}}{=} \Big(z_1, \ldots \underbrace{x_i}_{j+1} \ldots, z_n\Big)_k \\ &\underset{j+1<k}{=} z_k \\ &= (y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_k \\ &\underset{\text{[definition: 11.286]} \wedge i=j+1<k \Rightarrow i+1 \leqslant k}{=} x_k \\ &\underset{\text{[definition: 11.286]} \wedge j+1<k}{=} (y, x_1, \ldots, x_{j-1}, x_j, \ldots, x_n)_k \end{aligned}$$

$j + 1 < i$. Then [definition: 11.236(3)] applies and we have to look at the following cases:

$1 = k$. Then $1 = k < j + 1$ so that $\left(j + 1 \underset{n}{\rightsquigarrow} i\right)(k) \underset{\text{[definition: 11.236(2)]}}{=} k$ hence

$$\begin{aligned} a_k &= \Big(z_1, \ldots \underbrace{x_i}_{j+1} \ldots, z_n\Big)_k \\ &\underset{k=1\leqslant j<j+1}{=} z_k \\ &\underset{k=1}{=} z_1 \\ &= (y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_1 \\ &= y \\ &\underset{\text{[definition: 11.286]} \wedge k=1}{=} (y, x_1, \ldots, x_{j-1}, x_j, \ldots, x_n)_k \end{aligned}$$

$1<k<j+1$. Then $\left(j+1 \underset{n}{\leadsto} i\right)(k) \underset{\text{[definition: 11.236(2)]}}{=} k$ hence

$$\begin{array}{rcl} a_k & = & \left(z_1,\ldots \underbrace{x_i}_{j+1} \ldots, z_n\right)_k \\ & \underset{k=1\leqslant j<j+1}{=} & z_k \\ & = & (y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_k \\ & \underset{\text{[definition: 11.286]}\wedge 1<k<j+1\leqslant i}{=} & x_{k-1} \\ & \underset{\text{[definition: 11.286]}\wedge 1<k<j+1\Rightarrow 1<k\leqslant j}{=} & (y, x_1, \ldots, x_{j-1}, x_j, \ldots, x_n)_k \end{array}$$

$j+1\leqslant k<i$. Then $\left(j+1 \underset{n}{\leadsto} i\right)(k) \underset{\text{[definition: 11.236(2)]}}{=} k+1$ hence

$$\begin{array}{rcl} a_k & = & \left(z_1,\ldots \underbrace{x_i}_{j+1} \ldots, z_n\right)_{k+1} \\ & \underset{j+1\leqslant k<k+1}{=} & z_{k+1} \\ & = & (y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_{k+1} \\ & \underset{\text{[definition: 11.286]}\wedge 1<k+1\leqslant i}{=} & x_{(k+1)-1} \\ & = & x_k \\ & \underset{\text{[definition: 11.286]}\wedge j+1\leqslant k}{=} & (y, x_1, \ldots, x_{j-1}, x_j, \ldots, x_n)_k \end{array}$$

$k=i$. Then $\left(j+1 \underset{n}{\leadsto} i\right)(k) \underset{\text{[definition: 11.236(2)]}}{=} j+1$ hence

$$\begin{array}{rcl} a_k & = & \left(z_1,\ldots \underbrace{x_i}_{j+1} \ldots, z_n\right)_{j+1} \\ & = & x_i \\ & \underset{i=k}{=} & x_k \\ & \underset{\text{[definition: 11.286]}j<j+1<i=k}{=} & (y, x_1, \ldots, x_{j-1}, x_j, \ldots, x_n)_k \end{array}$$

$i<k\leqslant n$. Then $\left(j+1 \underset{n}{\leadsto} i\right)(k) \underset{\text{[definition: 11.236(2)]}}{=} k$ hence

$$\begin{array}{rcl} a_k & = & \left(z_1,\ldots \underbrace{x_i}_{j+1} \ldots, z_n\right)_k \\ & = & z_k \\ & = & (y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_k \\ & \underset{\text{[definition: 11.286]}\wedge i<k}{=} & x_k \\ & \underset{\text{[definition: 11.286]}\wedge j+1<i<k}{=} & (y, x_1, \ldots, x_{j-1}, x_j, \ldots, x_n)_k \end{array}$$

so in all cases we have $a_k = (y, x_1, \ldots, x_{j-1}, x_j, \ldots, x_n)_k$ proving that

$$\left(z_1,\ldots \underbrace{x_i}_{j+1} \ldots, z_n\right) \circ \left(j+1 \underset{n}{\leadsto} i\right) = (y, x_1, \ldots, x_{j-1}, x_j, \ldots, x_n)$$

2. Let $k\in\{1,\ldots,n\}$ and take $b_k = \left(z_1,\ldots \underbrace{x_i}_{j} \ldots, z_n\right)\left(\left(j \underset{n}{\leadsto} i+1\right)(k)\right)$ then using the definition of $\underset{n}{\leadsto}$ [see definition: 11.236] we must look at the following cases for $j, i+1$:

$i+1=j$. Then by [definition: 11.236] we have that $\left(j \underset{n}{\leadsto} i+1\right) = \mathrm{Id}_{\{1,\ldots,n\}}$ so that

$$b_k = \left(z_1,\ldots \underbrace{x_i}_{j} \ldots, z_n\right)_k \tag{11.168}$$

Now for $k \in \{1, \ldots, n\}$ we have to look at the following cases:

$k = 1$. Then

$$\begin{aligned} b_k &\underset{\text{[eq: 11.168]}}{=} \left(z_1, \ldots \underbrace{x_i}_{j} \ldots, z_n\right)_k \\ &\underset{k=1<i+1=j}{=} z_1 \\ &= (y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_1 \\ &\underset{\text{[definition: 11.286]}\wedge k=1}{=} y \\ &\underset{\text{[definition: 11.286]}\wedge k=1}{=} (y, x_1, \ldots, x_{j-1}, x_j, \ldots, x_n)_k \end{aligned}$$

$1 < k < i$. Then

$$\begin{aligned} b_k &\underset{\text{[eq: 11.168]}}{=} \left(z_1, \ldots \underbrace{x_i}_{j} \ldots, z_n\right)_k \\ &\underset{k<i<i+1=j}{=} z_k \\ &= (y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_k \\ &\underset{\text{[definition: 11.286]}\wedge 1<k<i}{=} x_{k-1} \\ &\underset{\text{[definition: 11.286]}\wedge 1<k<i<i+1<j}{=} (y, x_1, \ldots, x_{j-1}, x_j, \ldots, x_n)_k \end{aligned}$$

$1 < k = i$. Then

$$\begin{aligned} b_k &\underset{\text{[eq: 11.168]}}{=} \left(z_1, \ldots \underbrace{x_i}_{j} \ldots, z_n\right)_k \\ &\underset{k=i<i+1=j}{=} z_k \\ &= (y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_k \\ &\underset{\text{[definition: 11.286]}\wedge 1<k=i}{=} x_{k-1} \\ &\underset{\text{[definition: 11.286]}\wedge k=i<i+1}{=} (y, x_1, \ldots, x_{j-1}, x_j, \ldots, x_n)_k \end{aligned}$$

$k = i + 1$. Then

$$\begin{aligned} b_k &\underset{\text{[eq: 11.168]}}{=} \left(z_1, \ldots \underbrace{x_i}_{j} \ldots, z_n\right)_k \\ &\underset{k=i+1=j}{=} x_i \\ &\underset{k=i+1}{=} x_{k-1} \\ &\underset{\text{[definition: 11.286]}\wedge 1<k=i+1=j}{=} (y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_k \end{aligned}$$

$i + 1 < k \leqslant n$. Then

$$\begin{aligned} b_k &\underset{\text{[eq: 11.168]}}{=} \left(z_1, \ldots \underbrace{x_i}_{j} \ldots, z_n\right)_k \\ &\underset{j=i+1<k}{=} z_k \\ &\underset{\text{[definition: 11.286]}\wedge j=i+1<k\Rightarrow j+1\leqslant k}{=} (y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_k \end{aligned}$$

$i + 1 < j$. Then by [definition: 11.236(3)] we have look at the following cases:

$k = 1$. Then $\left(j \underset{n}{\rightsquigarrow} i+1\right)(k) \underset{\text{[definition: 11.236(3)]}}{=} k$ hence

$$\begin{aligned} b_k &= \left(z_1, \ldots \underbrace{x_i}_{j} \ldots, z_n\right)_k \\ &\underset{k=1<i+1<j}{=} z_k \\ &= (y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_k \end{aligned}$$

$$\underset{[\text{definition: } 11.286]\wedge k=1}{=} y$$
$$\underset{[\text{definition: } 11.286]\wedge k=1}{=} (y, x_1, \ldots, x_{j-1}, x_j, \ldots, x_n)_k$$

$1 < k < i+1$. Then $(j \underset{n}{\rightsquigarrow} i+1)(k) \underset{[\text{definition: } 11.236(3)]}{=} k$ hence

$$\begin{aligned}
b_k &= \Big(z_1, \ldots \underbrace{x_i}_{j} \ldots, z_n\Big)_k \\
&\underset{k<i+1<j}{=} z_k \\
&= (y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_k \\
&\underset{[\text{definition: } 11.286]\wedge 1<k<i+1 \Rightarrow 1<k\leqslant i}{=} x_{k-1} \\
&\underset{[\text{definition: } 11.286]\wedge 1<k<i+1<j}{=} (y, x_1, \ldots, x_{j-1}, x_j, \ldots, x_n)_k
\end{aligned}$$

$1 < k = i+1$. Then $(j \underset{n}{\rightsquigarrow} i+1)(k) \underset{[\text{definition: } 11.236(3)]}{=} j$ hence

$$\begin{aligned}
b_k &\underset{11.236}{=} \Big(z_1, \ldots \underbrace{x_i}_{j} \ldots, z_n\Big)_j \\
&= x_i \\
&\underset{k=i+1}{=} x_{k-1} \\
&\underset{[\text{definition: } 11.286]\wedge 1<k=i+1<j}{=} (y, x_1, \ldots, x_{j-1}, x_j, \ldots, x_n)_k
\end{aligned}$$

$i+1 < k \leqslant j$. Then $(j \underset{n}{\rightsquigarrow} i+1)(k) \underset{[\text{definition: } 11.236(3)]}{=} k-1$ hence

$$\begin{aligned}
b_k &\underset{11.236}{=} (z_1, \ldots, z_{j-1}, x_i, z_{j+1}, \ldots, z_n)_{k-1} \\
&\underset{k\leqslant j \Rightarrow k-1<j}{=} z_{k-1} \\
&= (y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_{k-1} \\
&\underset{[\text{definition: } 11.286]\wedge i+1<k \Rightarrow i+1\leqslant k-1}{=} x_{k-1} \\
&\underset{[\text{definition: } 11.286]\wedge 1<i+1<k\leqslant j}{=} (y, x_1, \ldots, x_{j-1}, x_j, \ldots, x_n)_k
\end{aligned}$$

$j < k \leqslant n$. Then $(j \underset{n}{\rightsquigarrow} i+1)(k) \underset{[\text{definition: } 11.236(3)]}{=} k$ hence

$$\begin{aligned}
b_k &\underset{11.236}{=} \Big(z_1, \ldots \underbrace{x_i}_{j} \ldots, z_n\Big)_k \\
&\underset{j<k}{=} z_k \\
&= (y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_k \\
&\underset{[\text{definition: } 11.286]\wedge i+1<j<k}{=} x_k \\
&\underset{[\text{definition: } 11.286]\wedge j<k \Rightarrow j+1\leqslant k}{=} (y, x_1, \ldots, x_{j-1}, x_j, \ldots, x_n)_k
\end{aligned}$$

So in all cases we have $b_k = (y, x_1, \ldots, x_{j-1}, x_j, \ldots, x_n)_k$ proving that

$$\Big(z_1, \ldots \underbrace{x_i}_{j} \ldots, z_n\Big) \circ (j \underset{n}{\rightsquigarrow} i+1) = (y, x_1, \ldots, x_{j-1}, x_j, \ldots, x_n)$$ □

TODO check this

Theorem 11.290. *Let $n \in \mathbb{N}$, X a finite dimensional vector space over a field of characteristic zero with $\dim(X) = n$, $y \in X$, $(x_1, \ldots, x_n) \in X^n$ and $\Delta \in \operatorname{Hom}^n(X; F)$ a determinant function then*

$$\underline{\Delta}(y, (x_1, \ldots, x_n)) = \Delta(x_1, \ldots, x_n) \cdot y$$

or using the definition of $\underline{\Delta}: X \times X^n \rightarrow X$ we have [see definition: 11.288]

$$\sum_{i \in \{1, \ldots, n\}} (-1)^{i-1} \cdot \Delta(y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \cdot x_i = \Delta(x_1, \ldots, x_n) \cdot y$$

Proof. Let $x=(x_1,\ldots,x_n)\in X^n$ then we have for $\{x_i\}_{i\in\{1,\ldots,n\}}\subseteq X$ the following cases to consider:

$\{x_i\}_{i\in\{1,\ldots,n\}}$ is linear dependent. Then by [theorem: 11.275] we have that

$$\Delta(x_1,\ldots,x_n)=0 \tag{11.169}$$

So if we prove that $\underline{\Delta}(y,(x_1,\ldots,x_n))=0$ we are done for this case. Consider the following cases for $n\in\mathbb{N}$.

$n=1$. As $\{x_i\}_{i\in\{1\}}$ is linear dependent there exists a $\{\alpha_i\}_{i\in\{1\}}$ not all zero [so $\alpha_1\neq 0$] such that $0=\sum_{i\in\{1\}}\alpha_i\cdot x_i=\alpha_1\cdot x_1$. So, as $\alpha_1\neq 0$, we have that $x_1=0$, hence

$$\Delta(y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)\cdot x_1=\Delta(y)\cdot 0=0$$

So

$$\begin{aligned}\underline{\Delta}(y,(x_1,\ldots,x_n)) &= \sum_{i\in\{1,\ldots,1\}}(-1)^{i-1}\cdot\Delta(y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)\cdot x_i\\ &= (-1)^0\cdot\Delta(y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)\cdot x_1\\ &= 0\end{aligned}$$

$n=2$. As $\{x_i\}_{i\in\{1,2\}}$ is linear dependent it follows by [theorem: 11.123] that there exists a $k\in\{1,2\}$ and a $\{\alpha_i\}_{i\in\{1,\ldots,2\}\setminus\{k\}}$ such that $x_k=\sum_{i\in\{1,2\}\setminus\{k\}}\alpha_i\cdot x_i$. For $k\in\{1,.2\}$ we have either:

$k=1$. Then $x_1=\sum_{i\in\{2\}}\alpha_i\cdot x_i=\alpha_2\cdot x_2$ so that

$$\begin{aligned}\underline{\Delta}(y,(x_1,\ldots,x_n)) &= \sum_{i\in\{1,\ldots,2\}}(-1)^{i-1}\cdot\Delta(y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)\cdot x_i\\ &= (-1)^{1-1}\cdot\Delta(y,x_2)\cdot x_1+(-1)^{2-1}\cdot\Delta(y,x_1)\cdot x_2\\ &= \Delta(y,x_2)\cdot x_1-\Delta(y,x_1)\cdot x_2\\ &= \Delta(y,x_2)\cdot(\alpha_2\cdot x_2)-\Delta(y,\alpha_2\cdot x_2)\cdot x_2\\ &= \alpha_2\cdot(\Delta(y,x_2)\cdot x_2)-\alpha_2\cdot\Delta(y,x_2)\cdot x_2\\ &= 0\end{aligned}$$

$k=2$. Then $x_2=\sum_{i\in\{1\}}\alpha_i\cdot x_i=\alpha_1\cdot x_1$ so that

$$\begin{aligned}\underline{\Delta}(y,(x_1,\ldots,x_n)) &= \sum_{i\in\{1,\ldots,2\}}(-1)^{i-1}\cdot\Delta(y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)\cdot x_i\\ &= (-1)^{1-1}\cdot\Delta(y,x_2)\cdot x_1+(-1)^{2-1}\cdot\Delta(y,x_1)\cdot x_2\\ &= \Delta(y,x_2)\cdot x_1-\Delta(y,x_1)\cdot x_2\\ &= \Delta(y,\alpha_1\cdot x_1)\cdot x_1-\Delta(y,x_1)\cdot(\alpha_1\cdot x_1)\\ &= \alpha_1\cdot(\Delta(y,x_1)\cdot x_1)-\alpha_1\cdot\Delta(y,x_1)\cdot x_1\\ &= 0\end{aligned}$$

$2<n$. As $\{x_i\}_{i\in\{1,\ldots,n\}}$ is linear dependent there exists by [theorem: 11.123] a $k\in\{1,\ldots,n\}$ and a $\{\alpha_i\}_{i\in\{1,\ldots,n\}\setminus\{k\}}\subseteq F$ such that

$$x_k=\sum_{i\in\{1,\ldots,n\}\setminus\{k\}}\alpha_i\cdot x_i \tag{11.170}$$

So

$$\begin{aligned}\underline{\Delta}(y,(x_1,\ldots,x_n)) &= \sum_{i\in\{1,\ldots,n\}}((-1)^{i-1}\cdot\Delta(y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n))\cdot x_i\\ &= A+B \end{aligned}\tag{11.171}$$

where

$$A = \sum_{i\in\{1,\ldots,n\}\setminus\{k\}}(-1)^{i-1}\cdot\Delta(y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)\cdot x_i \tag{11.172}$$

$$B = \sum_{i\in\{k\}}(-1)^{i-1}\cdot\Delta(y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)\cdot x_i \tag{11.173}$$

Now

$$\begin{aligned}B &= \sum_{i\in\{k\}} (-1)^{i-1}\cdot\Delta(y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)\cdot x_i\\ &= (-1)^{k-1}\cdot\Delta(y,x_1,\ldots,x_{k-1},x_{k+1},\ldots,x_n)\cdot x_k\\ &= (-1)\cdot((-1)^k\cdot\Delta(y,x_1,\ldots,x_{k-1},x_{k+1},\ldots,x_n)\cdot x_k)\end{aligned} \tag{11.174}$$

For $i\in\{1,\ldots,n\}\setminus\{k\}$ we have the following cases:

$\boldsymbol{k<i}$**.** Let

$$z=(y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)$$

then

$$\begin{aligned}z_{k+1} &= (y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_{k+1}\\ &\underset{[\text{definition: } 11.286]\wedge 1<k+1\leqslant i}{=} x_{(k+1)-1}\\ &= x_k\\ &\underset{[\text{eq: } 11.170]}{=} \sum_{j\in\{1,\ldots,n\}\setminus\{k\}} \alpha_j\cdot x_j\end{aligned}$$

So that

$$\begin{aligned}\Delta(y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n) &=\\ \Delta(z) &=\\ \Delta\left(z_1,\ldots\underbrace{\sum_{j\in\{1,\ldots,n\}\setminus\{k\}}\alpha_j\cdot x_j}_{k+1}\ldots,z_n\right) &=\\ \sum_{j\in\{1,\ldots,n\}\setminus\{k\}}\alpha_j\cdot\Delta\left(z_1,\ldots\underbrace{x_j}_{k+1}\ldots,z_n\right)\end{aligned} \tag{11.175}$$

Now for $j\in\{1,\ldots,n\}\setminus\{k\}$ consider the following cases for i,j:

$\boldsymbol{j<i}$**.** Then we have

$$\begin{aligned}\left(z_1,\ldots\underbrace{x_j}_{k+1}\ldots,z_n\right)_{j+1} &\underset{j\neq k\Rightarrow j+1\neq k+1}{=}\\ z_{j+1} &=\\ (y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_{j+1} &\underset{[\text{definition: } 11.286]\wedge 1<j+1\leqslant i}{=}\\ x_j &=\\ \left(z_1,\ldots\underbrace{x_j}_{k+1}\ldots,z_n\right)_{k+1}\end{aligned}$$

So we have $j\neq k\Rightarrow j+1\neq k+1$ and

$$\left(z_1,\ldots\underbrace{x_j}_{k+1}\ldots,z_n\right)_{k+1}=(y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_{j+1}$$

which by [theorem: 11.275] proves that

$$\Delta\left(z_1,\ldots\underbrace{x_j}_{k+1}\ldots,z_n\right)=0 \tag{11.176}$$

$\boldsymbol{i<j}$**.** Then we have

$$\begin{aligned}\left(z_1,\ldots\underbrace{x_j}_{k+1}\ldots,z_n\right)_{j} &\underset{k<i\Rightarrow k+1\leqslant i<j\Rightarrow k+1\neq j}{=}\\ z_j &=\\ (y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_j &\underset{[\text{definition: } 11.286]\wedge i<j\Rightarrow i+1\leqslant j}{=}\\ x_j &=\\ \left(z_1,\ldots\underbrace{x_j}_{k+1}\ldots,z_n\right)_{k+1}\end{aligned}$$

So we have $k<i\Rightarrow k+1\leqslant i<j\Rightarrow k+1\neq j$ and

$$\left(z_1,\ldots\underbrace{x_j}_{k+1}\ldots,z_n\right)_j=\left(z_1,\ldots\underbrace{x_j}_{k+1}\ldots,z_n\right)_{k+1}$$

which by [theorem: 11.275] proves that

$$\Delta\left(z_1,\ldots\underbrace{x_j}_{k+1}\ldots,z_n\right)=0 \tag{11.177}$$

***i* = *j*.** As $k<i\Rightarrow k+1\leqslant i$ and $z=(y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)$ we have by [theorem: 11.289] that

$$\left(z_1,\ldots\underbrace{x_i}_{k+1}\ldots,z_n\right)\circ\left(k+1\underset{n}{\rightsquigarrow}i\right)=(y,x_1,\ldots,x_{k-1},x_{k+1},\ldots,x_n)$$

Hence

$$\begin{array}{rcl}\Delta(y,x_1,\ldots,x_{k-1},x_{k+1},\ldots,x_n) & = & \\ \Delta\left(\left(z_1,\ldots\underbrace{x_i}_{k+1}\ldots,z_n\right)\circ\left(k+1\underset{n}{\rightsquigarrow}i\right)\right) & = & \\ \left(\left(k+1\underset{n}{\rightsquigarrow}i\right)\Delta\right)\left(z_1,\ldots\underbrace{x_i}_{k+1}\ldots,z_n\right) & \underset{\text{[theorem: 11.239]}}{=} & \\ (-1)^{|k+1-i|}\cdot\Delta\left(z_1,\ldots\underbrace{x_i}_{k+1}\ldots,z_n\right) & \underset{k<i\Rightarrow k+1\leqslant i}{=} & \\ (-1)^{i-k-1}\cdot\Delta\left(z_1,\ldots\underbrace{x_i}_{k+1}\ldots,z_n\right) & & \end{array}$$

which by [ref: 11.265] gives

$$\Delta\left(z_1,\ldots\underbrace{x_i}_{k+1}\ldots,z_n\right)=(-1)^{i-k-1}\cdot\Delta(y,x_1,\ldots,x_{k-1},x_{k+1},\ldots,x_n) \tag{11.178}$$

Further, using the fact that $\{1,\ldots,n\}\setminus\{k\}$ is the disjoint union of

$$\{1,\ldots,i-1\}\setminus\{k\},\ \{i\}\setminus\{k\}\text{ and }\{i+1,\ldots,n\}\setminus\{k\}$$

we have

$$\begin{array}{c}\sum_{j\in\{1,\ldots,n\}\setminus\{k\}}\alpha_j\cdot\Delta\left(z_1,\ldots\underbrace{x_j}_{k+1}\ldots,z_n\right)\underset{\text{[theorem: 11.43]}}{=}\\ A_1+A_2+A_2\end{array} \tag{11.179}$$

where

$$\begin{array}{rcl}A_1 & = & \sum_{j\in\{1,\ldots,i-1\}\setminus\{k\}}\alpha_j\cdot\Delta\left(z_1,\ldots\underbrace{x_j}_{k+1}\ldots,z_n\right)\\ & \underset{\text{[eq: 11.176]}}{=} & 0\\ A_2 & = & \sum_{j\in\{i\}\setminus\{k\}}\alpha_j\cdot\Delta\left(z_1,\ldots\underbrace{x_j}_{k+1}\ldots,z_n\right)\\ & \underset{i\neq k\Rightarrow\{i\}\setminus\{k\}=\{i\}}{=} & \alpha_i\cdot\Delta\left(z_1,\ldots\underbrace{x_j}_{k+1}\ldots,z_n\right)\\ & \underset{\text{[eq: 11.178]}}{=} & (\alpha_i\cdot(-1)^{i-k-1}\cdot\Delta(y,x_1,\ldots,x_{k-1},x_{k+1},\ldots,x_n))\\ A_3 & = & \sum_{j\in\{i+1,\ldots,n\}\setminus\{k\}}\alpha_j\cdot\Delta\left(z_1,\ldots\underbrace{x_j}_{k+1}\ldots,z_n\right)\\ & \underset{\text{[eq: 11.177]}}{=} & 0\end{array}$$

proving by [eqs 11.175, 11.179] and the above that that

$$\begin{array}{r}\Delta(y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n) \; = \\ \alpha_i\cdot(-1)^{i-k-1}\cdot\Delta(y,x_1,\ldots,x_{k-1},x_{k+1},\ldots,x_n)\end{array} \tag{11.180}$$

***i* < *k*.** Let

$$z=(y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)$$

then

$$\begin{aligned} z_k &= (y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_k \\ &\underset{[\text{definition: } 11.286]\wedge i<k\Rightarrow i+1\leqslant k}{=} x_k \\ &\underset{[\text{eq: } 11.170]}{=} \sum_{j\in\{1,\ldots,n\}\setminus\{k\}}\alpha_j\cdot x_j\end{aligned}$$

So that

$$\begin{array}{r}\Delta(y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n) \; = \\ \Delta(z) \; = \\ \Delta\left(z_1,\ldots\underbrace{\sum_{j\in\{1,\ldots,n\}\setminus\{k\}}\alpha_j\cdot x_j}_{k}\ldots,z_n\right) \; = \\ \sum_{j\in\{1,\ldots,n\}\setminus\{k\}}\alpha_j\cdot\Delta\left(z_1,\ldots\underbrace{x_j}_{k}\ldots,z_n\right)\end{array} \tag{11.181}$$

Now for $j\in\{1,\ldots,n\}\setminus\{k\}$ we have the following cases:

***j* < *i*.** Then we have:

$$\begin{array}{rc}\left(z_1,\ldots\underbrace{x_j}_{k}\ldots,z_n\right)_{j+1} & \underset{j<i\Rightarrow j+1\leqslant i<k}{=} \\ z_{j+1} & = \\ (y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_{j+1} & \underset{[\text{definition: } 11.286]\wedge j<i\Rightarrow 1<j+1\leqslant i}{=} \\ x_{(j+1)-1} & = \\ x_j & \underset{j\neq k}{=} \\ \left(z_1,\ldots\underbrace{x_j}_{k}\ldots,z_n\right)_k & \end{array}$$

So we have $j<i\Rightarrow j+1\leqslant i<k\Rightarrow j+1\neq k$ and

$$\left(z_1,\ldots\underbrace{x_j}_{k}\ldots,z_n\right)_{j+1}=\left(z_1,\ldots\underbrace{x_j}_{k}\ldots,z_n\right)_k$$

which by [theorem: 11.275] proves that

$$\Delta\left(z_1,\ldots\underbrace{x_j}_{k}\ldots,z_n\right)=0 \tag{11.182}$$

***i* < *j*.** Then we have:

$$\begin{array}{rc}\left(z_1,\ldots\underbrace{x_j}_{k}\ldots,z_n\right)_j & \underset{j\neq k}{=} \\ z_j & = \\ (y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n) & \underset{[\text{definition: } 11.286]\wedge i<j\Rightarrow i+1\leqslant j}{=} \\ x_j & = \\ \left(z_1,\ldots\underbrace{x_j}_{k}\ldots,z_n\right)_k & \end{array}$$

So we have $j \neq k$ and

$$\left(z_1, \ldots \underbrace{x_j}_{k} \ldots, z_n\right)_j = \left(z_1, \ldots \underbrace{x_j}_{k} \ldots, z_n\right)_k$$

which by [theorem: 11.275] proves that

$$\Delta\left(z_1, \ldots \underbrace{x_j}_{k} \ldots, z_n\right) = 0 \tag{11.183}$$

***i* = *j*.** As As $i < k \Rightarrow i + 1 \leqslant k$ and $z = (y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$ we have by [theorem: 11.289] that

$$\left(z_1, \ldots \underbrace{x_i}_{k} \ldots, z_n\right) \circ \left(k \underset{n}{\rightsquigarrow} i + 1\right) = (y, x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n)$$

hence

$$\begin{aligned}
\Delta(y, x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n) &= \\
\Delta\left(\left(z_1, \ldots \underbrace{x_i}_{k} \ldots, z_n\right) \circ \left(k \underset{n}{\rightsquigarrow} i + 1\right)\right) &= \\
\left(\left(k \underset{n}{\rightsquigarrow} i + 1\right)\Delta\right)\left(z_1, \ldots \underbrace{x_i}_{k} \ldots, z_n\right) &= \\
\operatorname{sign}\left(\left(k \underset{n}{\rightsquigarrow} i + 1\right)\right) \cdot \Delta\left(z_1, \ldots \underbrace{x_i}_{k} \ldots, z_n\right) &\underset{\text{[theorem: 11.239]}}{=} \\
(-1)^{|k-i-1|} \cdot \Delta\left(z_1, \ldots \underbrace{x_i}_{k} \ldots, z_n\right) &\underset{i<k \Rightarrow i+1 \leqslant k}{=} \\
(-1)^{k-i-1} \cdot \Delta\left(z_1, \ldots \underbrace{x_i}_{k} \ldots, z_n\right) &
\end{aligned}$$

which by [theorem: 11.265] gives TODO

$$\Delta\left(z_1, \ldots \underbrace{x_j}_{k} \ldots, z_n\right) = (-1)^{k-i-1} \cdot \Delta(y, x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n) \tag{11.184}$$

So that as $\{1, \ldots, n\} \setminus \{k\}$ is the disjoint union of

$$\{1, \ldots, i-1\} \setminus \{k\}, \{i\} \setminus \{k\} \text{ and } \{i+1, \ldots, n\} \setminus \{k\}$$

we have

$$\begin{aligned}
\sum_{j \in \{1, \ldots, n\} \setminus \{k\}} \alpha_j \cdot \Delta\left(z_1, \ldots \underbrace{x_j}_{k} \ldots, z_n\right) &\underset{\text{[theorem: 11.43]}}{=} \\
A_1 + A_2 + A_2 &
\end{aligned} \tag{11.185}$$

Where

$$\begin{aligned}
A_1 &= \sum_{j \in \{1, \ldots, i-1\} \setminus \{k\}} \alpha_j \cdot \Delta\left(z_1, \ldots \underbrace{x_j}_{k} \ldots, z_n\right) \\
&\underset{\text{[eq: 11.182]}}{=} 0 \\
A_2 &= \sum_{j \in \{i\} \setminus \{k\}} \alpha_j \cdot \Delta\left(z_1, \ldots \underbrace{x_j}_{k} \ldots, z_n\right) \\
&\underset{i \neq k}{=} \alpha_i \cdot \Delta\left(z_1, \ldots \underbrace{x_i}_{k} \ldots, z_n\right) \\
&\underset{\text{[eq: 11.184]}}{=} \alpha_i \cdot (-1)^{k-i-1} \cdot \Delta(y, x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n) \\
A_3 &= \sum_{j \in \{i+1, \ldots, n\} \setminus \{k\}} \alpha_j \cdot \Delta\left(z_1, \ldots \underbrace{x_j}_{k} \ldots, x_n\right) \\
&\underset{\text{[eq: 11.183]}}{=} 0
\end{aligned}$$

proving by [eqs: 11.181, 11.185] and the above that

$$\Delta(y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) = \alpha_i \cdot (-1)^{k-i-1} \cdot \Delta(y, x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n) \tag{11.186}$$

Further as $\{1, \ldots, n\} \setminus k$ is the disjoint union of $\{1, \ldots, k-1\}$ and $\{k+1, \ldots, n\}$ we have

$$\begin{aligned} A &\underset{\text{[eq: 11.172]}}{=} \sum_{i \in \{1, \ldots, n\} \setminus \{k\}} (-1)^{i-1} \cdot \Delta(y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \cdot x_i \\ &\underset{\text{[theorem: 11.43]}}{=} B_1 + B_2 \end{aligned}$$

where

$$\begin{aligned} B_1 &= \sum_{i \in \{1, \ldots, k-1\}} (-1)^{i-1} \cdot \Delta(y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \cdot x_i \\ &\underset{\text{[eq: 11.186]}}{=} \sum_{i \in \{1, \ldots, k-1\}} (-1)^{i-1} \cdot \alpha_i \cdot (-1)^{k-i-1} \cdot \Delta(y, x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n) \cdot x_i \\ &= \sum_{i \in \{1, \ldots, k-1\}} (-1)^{k-2} \cdot \alpha_i \cdot \Delta(y, x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n) \cdot x_i \\ &= (-1)^{k-2} \cdot \Delta(y, x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n) \cdot \sum_{i \in \{1, \ldots, k-1\}} \alpha_i \cdot x_i \\ &= (-1)^2 \cdot (-1)^{k-2} \cdot \Delta(y, x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n) \cdot \sum_{i \in \{1, \ldots, k-1\}} \alpha_i \cdot x_i \\ &= (-1)^k \cdot \Delta(y, x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n) \cdot \sum_{i \in \{1, \ldots, k-1\}} \alpha_i \cdot x_i \\ B_2 &= \sum_{i \in \{k+1, \ldots, n\}} (-1)^{i-1} \cdot \Delta(y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \cdot x_i \\ &\underset{\text{[eq: 11.178]}}{=} \sum_{i \in \{k+1, \ldots, n\}} (-1)^{i-1} \cdot \alpha_i \cdot (-1)^{i-k-1} \cdot \Delta(y, x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n) \cdot x_i \\ &= \sum_{i \in \{k+1, \ldots, n\}} (-1)^{2 \cdot i-k-2} \cdot \alpha_i \cdot \Delta(y, x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n) \cdot x_i \\ &= (-1)^{2 \cdot i-k-2} \cdot \Delta(y, x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n) \cdot \sum_{i \in \{k+1, \ldots, n\}} \alpha_i \cdot x_i \\ &= (-1)^{2 \cdot k} \cdot (-1)^{2 \cdot i-k-2} \cdot \Delta(y, x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n) \cdot \sum_{i \in \{k+1, \ldots, n\}} \alpha_i \cdot x_i \\ &= (-1)^{2 \cdot i-2+k} \cdot \Delta(y, x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n) \cdot \sum_{i \in \{k+1, \ldots, n\}} \alpha_i \cdot x_i \\ &= (-1)^k \cdot \Delta(y, x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n) \cdot \sum_{i \in \{k+1, \ldots, n\}} \alpha_i \cdot x_i \end{aligned}$$

So, as $x_k = \sum_{i \in \{1, \ldots, n\} \setminus \{k\}} \alpha_i \cdot x_i = \sum_{i \in \{1, \ldots, k-1\}} \alpha_i \cdot x_i + \sum_{i \in \{k+1, \ldots, n\}} \alpha_i \cdot x_i$, we have that

$$\begin{aligned} A &= (-1)^k \cdot \Delta(y, x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n) \\ &\underset{\text{[eq: 11.174]}}{=} (-1) \cdot B \end{aligned}$$

So that $A + B = 0$ proving by [eq: 11.171] that

$$\underline{\Delta}(y, (x_1, \ldots, x_n)) = 0$$

So in all cases we have $\underline{\Delta}(y, (x_1, \ldots, x_n)) = 0$ which by [eq: 11.169] proves that

$$\underline{\Delta}(y, (x_1, \ldots, x_n)) = 0 \cdot y \underset{\text{[eq: 11.169]}}{=} \Delta(x_1, \ldots, x_n) \cdot y$$

$\{x_i\}_{i \in \{1, \ldots, n\}}$ is linear independent. Then by [corollary: 11.150] we have that

$$\{x_i | i \in \{1, \ldots, n\}\} \text{ is a basis of } X$$

So if $y \in X$ there exists by [theorem: 11.137] a $\{\alpha_i\}_{i \in \{1,\ldots,n\}} \subseteq F$ such that

$$y = \sum_{i \in \{1,\ldots,n\}} \alpha_i \cdot x_i \tag{11.187}$$

By [definition: 11.288]

$$\underline{\Delta}(y, (x_1, \ldots, x_n)) = \sum_{i \in \{1,\ldots,n\}} (-1)^{i-1} \cdot \Delta(y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \cdot x_i \tag{11.188}$$

Let $i \in \{1, \ldots, n\}$ then we have as Δ is n-linear that

$$\Delta(y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) = \sum_{j \in \{1,\ldots,n\}} \alpha_j \cdot \Delta(x_j, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \tag{11.189}$$

For i, j we have the following cases:

***i* < *j*.** Then

$$\begin{array}{rcl} (x_j, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_j & \underset{[\text{theorem: } 11.286] \wedge i < j \Rightarrow i+1 \leqslant j}{=} & x_j \\ & \underset{[\text{theorem: } 11.286]}{=} & (x_j, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_1 \end{array}$$

which as $1 \leqslant i < j \Rightarrow j \neq 1$ proves by [theorem: 11.275] that

$$\Delta(x_j, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) = 0 \tag{11.190}$$

***j* < *i*.** Then

$$\begin{array}{rcl} (x_j, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_{j+1} & \underset{[\text{theorem: } 11.286] \wedge j < i \Rightarrow 1 < j+1 \leqslant i}{=} & x_{(j+1)-1} \\ & = & x_j \\ & \underset{[\text{theorem: } 11.286]}{=} & (x_j, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_1 \end{array}$$

which as $1 < j + 1$ proves by [theorem: 11.275] that

$$\Delta(x_j, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) = 0 \tag{11.191}$$

***i* = *j*.** Let

$$A = (x_j, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \circ \left(1 \underset{n}{\rightsquigarrow} j\right) \tag{11.192}$$

then we have the following cases to consider:

***j* = 1.** Then $\left(1 \underset{n}{\rightsquigarrow} j\right) = \left(1 \underset{n}{\rightsquigarrow} 1\right) \underset{[\text{definition: } 11.236]}{=} \mathrm{Id}_{\{1,\ldots,n\}}$ so that $\forall k \in \{1, \ldots, n\}$

$$\begin{array}{rcl} A_k & = & (x_j, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots x_n)_k \\ & \underset{[\text{theorem: } 11.286]}{=} & \begin{cases} x_j \text{ if } k = 1 \\ x_{k-1} \text{ if } k \in \{2, \ldots, j\} \\ x_k \text{ if } k \in \{j+1, \ldots, n\} \end{cases} \\ & \underset{j=1 \wedge \{2,\ldots,j\} = \varnothing}{=} & \begin{cases} x_1 \text{ if } k = 1 \\ x_k \text{ if } k \in \{2, \ldots, n\} \end{cases} \\ & = & x_k \end{array}$$

so that

$$A = (x_1, \ldots, x_n)$$

***j* ∈ {2, …, *n*}.** Then $1 < j$ we have the following cases to consider (helped by [definition: 11.236]) for $k \in \{1, \ldots n\}$:

1 ⩽ *k* < *j*. Then

$$\begin{array}{rcl} A_k & = & (x_j, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_{(1 \underset{n}{\rightsquigarrow} j)(k)} \\ & \underset{[\text{definition: } 11.236]}{=} & (x_j, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_{k+1} \\ & \underset{[\text{theorem: } 11.286] \wedge 1 \leqslant k < j \Rightarrow 1 < k+1 \leqslant j}{=} & x_{(k+1)-1} \\ & = & x_k \end{array}$$

$k=j$. Then

$$\begin{aligned} A_k &= (x_j, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_{(1 \underset{n}{\rightsquigarrow} j)(k)} \\ &\underset{\text{[definition: 11.236]}}{=} (x_j, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_1 \\ &\underset{\text{[theorem: 11.286]}}{=} x_j \\ &\underset{k=j}{=} x_k \end{aligned}$$

$j<k\leqslant n$. Then

$$\begin{aligned} A_k &= (x_j, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_{(1 \underset{n}{\rightsquigarrow} j)(k)} \\ &\underset{\text{[definition: 11.236]}}{=} (x_j, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_k \\ &\underset{\text{[theorem: 11.286]} \wedge j<k \Rightarrow j+1 \leqslant k}{=} x_k \end{aligned}$$

so that

$$A = (x_1, \ldots, x_n)$$

So in all cases $A=(x_1,\ldots,x_n)$ proving that

$$(x_j, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \circ \left(1 \underset{n}{\rightsquigarrow} j\right) = (x_1, \ldots, x_n) \tag{11.193}$$

So we have that

$$\begin{aligned} \Delta(x_1, \ldots, x_n) &= \Delta\left((x_j, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \circ \left(1 \underset{n}{\rightsquigarrow} j\right)\right) \\ &= \left(\left(1 \underset{n}{\rightsquigarrow} j\right)\Delta\right)(x_j, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \\ &= \operatorname{sign}\left(\left(1 \underset{n}{\rightsquigarrow} j\right)\right) \cdot \Delta(x_j, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \\ &\underset{\text{[theorem: 11.239]}}{=} (-1)^{j-1} \cdot \Delta(x_j, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \\ &\underset{i=j}{=} (-1)^{i-1} \cdot \Delta(x_j, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \end{aligned}$$

proving that

$$\Delta(x_j, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) = (-1)^{i-1} \cdot \Delta(x_1, \ldots, x_n) \tag{11.194}$$

Now as $\{1,\ldots,n\}$ i the disjoint union of $\{1,\ldots,i-1\}$, $\{x_i\}$ and $\{x_{i+1},\ldots,x_n\}$ we have

$$\sum_{j \in \{1, \ldots, n\}} \alpha_j \cdot \Delta(x_j, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \underset{\text{[theorem: 11.43]}}{=} C_1 + C_2 + C_3$$

where

$$\begin{aligned} C_1 &= \sum_{j \in \{1, \ldots, i-1\}} \alpha_j \cdot \Delta(x_j, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \\ &\underset{\text{[eq: 11.191]}}{=} 0 \\ C_2 &= \sum_{j \in \{i\}} \alpha_j \cdot \Delta(x_j, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \\ &= \alpha_i \cdot \Delta(x_i, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \\ &\underset{\text{[eq: 11.195]}}{=} (-1)^{i-1} \cdot \alpha_i \cdot \Delta(x_1, \ldots, x_n) \\ C_3 &= \sum_{j \in \{i+1, \ldots, n\}} \alpha_j \cdot \Delta(x_j, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \\ &\underset{\text{[eq: 11.190]}}{=} 0 \end{aligned}$$

So that

$$\sum_{j \in \{1, \ldots, n\}} \alpha_j \cdot \Delta(x_j, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) = (-1)^{i-1} \cdot \alpha_i \cdot \Delta(x_1, \ldots, x_n)$$

Substituting the above in [eq: 11.189] proves

$$\Delta(y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \cdot x_i = \alpha_i \cdot (-1)^{i-1} \cdot \Delta(x_1, \ldots, x_n) \tag{11.195}$$

So

$$\begin{aligned}\underline{\Delta}(y,(x_1,\ldots,x_n)) &\underset{\text{def}}{=} \sum_{i\in\{1,\ldots,n\}} (-1)^{i-1}\cdot\Delta(y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)\cdot x_i\\ &\underset{\text{[eq: 11.195]}}{=} \sum_{i\in\{1,\ldots,n\}} (-1)^{i-1}\cdot\alpha_i\cdot(-1)^{i-1}\cdot\Delta(x_1,\ldots,x_n)\cdot x_i\\ &= \Delta(x_1,\ldots,x_n)\cdot\sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot x_i\\ &\underset{\text{[eq: 11.187}}{=} \Delta(x_1,\ldots,x_n)\cdot y\end{aligned}$$

proving that

$$\underline{\Delta}(y,(x_1,\ldots,x_n)) = \Delta(x_1,\ldots,x_n)\cdot y$$

So in all the case we have proved that

$$\underline{\Delta}(y,(x_1,\ldots,x_n)) = \Delta(x_1,\ldots,x_n)\cdot y$$ □

We show no how given a determinant function and a linear mapping we can create a new determinant function.

Definition 11.291. *Let $n\in\mathbb{N}$, X a vector space over a field F such that $\operatorname{dom}(X)=n$, $\Delta\in\operatorname{Hom}^n(X;F)$ a non trivial determinant function and $L\in\operatorname{Hom}(X,X)$ [a linear transformation] then we define*

$$\Delta_L: X^n\rightarrow F \text{ by } \Delta_L(x_1,\ldots,x_n)=\Delta(L(x_1),\ldots,L(x_1))$$

It turns out that Δ_L is also a determinant function.

Theorem 11.292. *Let $n\in\mathbb{N}$, X a vector space over a field F with characteristic zero such that $\operatorname{dom}(X)=n$, $\Delta\in\operatorname{Hom}^n(X;F)$ a non trivial determinant function and $L\in\operatorname{Hom}(X,X)$ then*

$$\Delta_L \text{ is a determinant function}$$

Proof. First we proof that Δ_L is n-linear. Let $i\in\{1,\ldots,n\}$, $\alpha\in F$ and $u,v\in X$ then we have:

$$\begin{aligned}\Delta_L\left(x_1,\ldots\underbrace{u+\alpha\cdot v}_{i}\ldots,x_n\right) &=\\ \Delta\left(L(x_1),\ldots\underbrace{L(u+\alpha\cdot v)}_{i}\ldots,L(x_n)\right) &=\\ \Delta\left(L(x_1),\ldots\underbrace{L(u)+\alpha\cdot L(v)}_{i}\ldots,L(x_n)\right) &=\\ \Delta\left(L(x_1),\ldots\underbrace{L(u)}_{i}\ldots,L(x_n)\right)+\alpha\cdot\Delta\left(L(x_1),\ldots\underbrace{L(v)}_{i}\ldots,L(x_n)\right) &=\\ \Delta_L\left(x_1,\ldots\underbrace{u}_{i}\ldots,x_n\right)+\alpha\cdot\Delta_L\left(x_1,\ldots\underbrace{v}_{i}\ldots,x_n\right)&\end{aligned}$$

proving that

$$\Delta_L\in\operatorname{Hom}^n(X;F) \text{ or } \Delta_L \text{ is } n\text{-linear}$$

As for skew-symmetry. Let $(x_1,\ldots,x_n)\in X^n$ such that $\exists i,j\in\{1,\ldots,n\}$ with $i\neq j$ and $x_i=x_j$ then $L(x_i)=L(x_j)$, hence

$$\begin{aligned}\Delta_L(x_1,\ldots,x_n) &= \Delta(L(x_1),\ldots,L(x_n))\\ &\underset{\text{[theorem: 11.275]}}{=} 0\end{aligned}$$

which by [theorem: 11.275] proves that

$$\Delta_L \text{ is skew-symmetric}$$ □

Note 11.293. It is not true that Δ_L is a non trivial determinant function if Δ is a non trivial determinant function. For example, if $L=C_0: X\rightarrow X$ is defined by $C_0(x)=0$ then for $(x_1,\ldots,x_n)\in X^n$ we have $\Delta_L(x_1,\ldots,x_n)=\Delta(L(x_1),\ldots,L(x_n))=L(0,\ldots,0)=0$, even if Δ is non trivial.

Theorem 11.294. *Let $n \in \mathbb{N}$, X a finite dimensional vector space over a field F with $\dim(X) = n$ and $L \in \mathrm{Hom}(X,X)$ then there exist a $\alpha \in F$ such that for **every** non trivial determinant function $\Delta \in \mathrm{Hom}^n(X;F)$ we have*

$$\Delta_L = \alpha \cdot \Delta$$

Proof. Let Δ be the non trivial that exists by [theorem: 11.283], for Δ_L we have two possibilities:

$\boldsymbol{\Delta_L = C_0}$**.** Then if we take $\alpha = 0$ we have that

$$\Delta_L = C_0 = 0 \cdot \Delta$$

$\boldsymbol{\Delta_L \neq C_0}$**.** Then by [theorem: 11.292] Δ_L is also a determinant function,so by [theorem: 11.285] there exists a $\alpha \in F$ such that

$$\Delta_L = \alpha \cdot \Delta$$

So for L there exist a $\alpha \in F$ such that

$$\Delta_L = \alpha \cdot \Delta \tag{11.196}$$

proving existence. Let now Δ' be another non trivial determent function then by [theorem: 11.285] there exist a $\lambda \in F$ such that

$$\Delta' = \lambda \cdot \Delta \tag{11.197}$$

Let $(x_1, \ldots, x_n) \in X^n$ then

$$\begin{aligned}\Delta'_L(x_1, \ldots, x_n) &= \Delta'(L(x_1), \ldots, L(x_n)) \\ &\underset{[\text{eq: } 11.197]}{=} \lambda \cdot \Delta(L(x_1), \ldots, L(x_n)) \\ &= \lambda \cdot \Delta_L(x_1, \ldots, x_n)\end{aligned}$$

So

$$\begin{aligned}\Delta'_L &= \lambda \cdot \Delta_L \\ &\underset{[\text{eq: } 11.196]}{=} \lambda \cdot (\alpha \cdot \Delta) \\ &= \alpha \cdot (\lambda \cdot \Delta) \\ &\underset{[\text{eq: } 11.197]}{=} \alpha \cdot \Delta'\end{aligned}$$

proving that

$$\Delta'_L = \alpha \cdot \Delta_L$$ □

The above theorem ensures that the following definition of the determinant of a linear function makes sense.

Definition 11.295. *Let $n \in \mathbb{N}$, X a vector space over a field F with characteristic zero such that $\dim(X) = n$ and $L \in \mathrm{Hom}(X,X)$ then we define the function*

$$\det: \mathrm{Hom}^n(X;Y) \to F$$

where $\det(L) \in F$ is the scalar such that for every non trivial determinant function Δ we have

$$\Delta_L = \det(L) \cdot \Delta$$

Theorem 11.296. *Let $n \in \mathbb{N}$, X a vector space over a field F with characteristic zero such that $\dim(X) = n$, $E = \{e_1, \ldots, e_n\}$ basis for X and $L \in \mathrm{Hom}(X,X)$ then*

$$\det(L) = \Delta_L(e_1, \ldots, e_n) = \Delta(L(e_1), \ldots, L(e_n))$$

where $\Delta: X^n \to F$ is the determinant function such that $\Delta(e_1, \ldots, e_n) = 1$ [see 11.283 for existence]

Proof. This follows from

$$\Delta(L(e_1), \ldots, L(e_n)) = \Delta_L(e_1, \ldots, e_n) \underset{\text{def}}{=} \det(L) \cdot \Delta(e_1, \ldots, e_n) = \det(L) \cdot 1 = \det(L)$$ □

Example 11.297. Let $n \in \mathbb{N}$, X a vector space over a field F with characteristic zero such that $\dim(X) = n$ and $\lambda \in F$ then we have for

$$\lambda \cdot \mathrm{Id}_X: X \to X \text{ defined by } (\lambda \cdot \mathrm{Id}_X)(x) = \lambda \cdot x$$

that

$$\det(\lambda \cdot \mathrm{Id}_X) = \lambda^n$$

or in the case of $\lambda = 1$ that

$$\det(\mathrm{Id}_X) = 1$$

Proof. Let Δ be a non trivial determinant function then

$$\begin{aligned}\Delta_{\lambda \cdot \mathrm{Id}_X}(x_1, \ldots, x_n) &= \Delta(\lambda \cdot \mathrm{Id}_X(x_1), \ldots, \lambda \cdot \mathrm{Id}_X(x_n)) \\ &= \Delta(\lambda \cdot x_1, \ldots, \lambda \cdot x_n) \\ &\underset{[\text{theorem: } 11.254]}{=} \lambda^n \cdot \Delta(x_1, \ldots, x_n)\end{aligned}$$

proving that $\Delta_{\lambda \cdot \mathrm{Id}_X} = \lambda^n \cdot \Delta$, hence

$$\det(\lambda \cdot \mathrm{Id}_x) = \lambda^n$$

□

Example 11.298. If X is a one dimensional vector space with a basis $\{e_1\}$ and $L \in \mathrm{Hom}(X, X)$ then

$$\det(L) = \beta$$

where β is defined by

$$L(e_1) = \beta \cdot e_1$$

Proof. If $x = (x_1) \in X^1$ then as $x_1 \in X$ there exists a $\alpha \in F$ such that $x_1 = \alpha \cdot e_1$. Hence $L(x_1) = L(\alpha \cdot e_1) = \alpha \cdot L(e_1)$. As $L(e_1) \in X$ there exists a $\beta \in F$ such that $L(e_1) = \beta \cdot e_1$. So we have

$$\Delta_L(x_1) = \Delta(L(x_1)) = \Delta((\alpha \cdot \beta) \cdot e_1) = \beta \cdot \Delta(\alpha \cdot e_1) = \beta \cdot \Delta(x_1)$$

proving that

$$\det(L) = \beta s$$

□

Theorem 11.299. *Let $n \in \mathbb{N}$, X a vector space over a field F with characteristic zero such that $\dim(X) = n$ then we have:*

1. $\det(\mathrm{Id}_X) = 1$
2. *Let $L \in \mathrm{Hom}(X, X)$ then L is a linear isomorphism* $\Leftrightarrow \det(L) \neq 0$
3. *If $L_1, L_2 \in \mathrm{Hom}(X, X)$ then* $\det(L_1 \circ L_2) = \det(L_1) \cdot \det(L_2)$
4. *If $L \in \mathrm{Hom}(X, X)$ is a linear isomorphism then* $\det(L^{-1}) = (\det(L))^{-1}$

Proof.

1. Let $(x_1, \ldots, x_n) \in X^n$ and Δ a non trivial determinant function then we have

$$\Delta_{\mathrm{Id}_X}(x_1, \ldots, x_n) = \Delta(\mathrm{Id}_X(x_1), \ldots, \mathrm{Id}_X(x_n)) = \Delta(x_1, \ldots, x_n)$$

   proving that $\Delta_{\mathrm{Id}_X} = \Delta = 1 \cdot \Delta$. So that $\det(\mathrm{Id}) = 1$.

2. As $\dim(X) = n$ there exists a finite basis $E = \{e_1, \ldots, e_n\}$ of X.

   ⇒. As $e: \{1, \ldots, n\} \to E$ is a bijection and $L: X \to X$ is a bijection we have that $L \circ e$ is injective so that $L \circ e: \{1, \ldots, n\} \to L(e(\{1, \ldots, n\})) = L(E)$ is a bijection. Hence $\mathrm{card}(L(E)) = n$ and $L(E) = \{(L \circ e)_1, \ldots, (L \circ e)_n\} = \{L(e_1), \ldots, L(e_n)\}$. As L is injective and E is linear independent it follows from [theorem: 11.187] that $L(E) = \{L(e_1), \ldots, L(e_n)\}$ is linear independent. So applying [theorem: 11.149] it follows that

$$L(E) = \{L(e_1), \ldots, L(e_n)\} \text{ is a basis for } X$$

   Using [theorem: 11.283] there exist a non trivial determinant function Δ such that

$$\Delta(L(e_1), \ldots, L(e_n)) = 1$$

   So

$$1 = \Delta(L(e_1), \ldots, L(e_n)) = \Delta_L(e_1, \ldots, e_n) = \det(L) \cdot \Delta(e_1, \ldots, e_n)$$

   If now $\det(L) = 0$ then we have $1 = 0$, so we must have that

$$\det(L) \neq 0$$

⇐. Using [theorem: 11.187] there exist a non trivial determinant function Δ such that

$$\Delta(e_1,\ldots,e_n)=1$$

then

$$\begin{array}{rcl}\Delta(L(e_1),\ldots,L(e_n)) & = & \Delta_L(e_1,\ldots,e_n)\\ & = & \det(L)\cdot\Delta(e_1,\ldots,e_n)\\ & = & \det(L)\end{array}$$

Assume that $\{L(e_i)\}_{i\in\{1,\ldots,n\}}\subseteq X$ is linear dependent then it follows from [theorem: 11.275] that $\Delta(L(e_1),\ldots,L(e_n))=0$, so that $\det(L)=0$, contradicting $\det(L)\neq 0$. Hence we have that $\{L(e_i)\}_{i\in\{1,\ldots,n\}}$ is linear independent. Let $x\in\ker(L)$ then $L(x)=0$ and as $E=\{e_1,\ldots,e_n\}$ is a basis we have that there exist a $\{\alpha_i\}_{i\in\{1,\ldots,n\}}\subseteq F$ such that $x=\sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot e_i$ so that

$$\begin{array}{rcl}0 & = & L(x)\\ & = & L\left(\sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot e_i\right)\\ & = & \sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot L(e_i)\end{array}$$

which as $\{L(e_i)\}_{i\in\{1,\ldots,n\}}$ is linear independent proves that $\forall i\in\{1,\ldots,n\}$ $\alpha_i=0$. So that $x=\sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot e_i=0$. Hence we have that $\ker(L)=\{0\}$ which by [theorem: 11.185] proves that

$$L \text{ is injective}$$

As $\dim(X)$ is finite and $L: X\rightarrow X$ is injective it follows from [theorem: 11.189] that

$$L: X\rightarrow X \text{ is a linear isomorphism}$$

3. First we have if $(x_1,\ldots,x_n)\in X^n$ and Δ a non trivial determinant function:

$$\begin{array}{rcl}\Delta_{L_1\circ L_2}(x_1,\ldots,x_n) & = & \Delta((L_1\circ L_2)(x_1),\ldots,(L_1\circ L_2)(x))\\ & = & \Delta(L_1(L_2(x)),\ldots,L_1(L_2(x)))\\ & = & \Delta_{L_1}(L_2(x_1),\ldots,L_2(x_n))\\ & = & \det(L_1)\cdot\Delta(L_2(x_1),\ldots,L_2(x_n))\\ & = & \det(L_1)\cdot\Delta_{L_2}(x_1,\ldots,x_n)\\ & = & \det(L_1)\cdot(\det(L_2)\cdot\Delta(x_1,\ldots,x_n))\\ & = & (\det(L_1)\cdot\det(L_2))\cdot\Delta(x_1,\ldots,x_n)\end{array}$$

so that $\Delta_{L_1\circ L_2}=(\det(L_1)\cdot\det(L_2))\cdot\Delta$ proving that

$$\det(\Delta_{L_1\circ L_2})=\det(L_1)\cdot\det(L_2)$$

4. As L is a linear isomorphism we have that

$$\begin{array}{rcl}1 & \underset{(1)}{=} & \det(\mathrm{Id}_X)\\ & = & \det(L\circ L^{-1})\\ & \underset{(3)}{=} & \det(L)\cdot\det(L^{-1})\end{array}$$

which as by (2) $\det(L)\neq 0$ proves that

$$\det(L^{-1})=(\det(L))^{-1}$$

□

Corollary 11.300. *Let $n\in\mathbb{N}$, X a vector space over a field F with characteristic zero such that $\dim(X)=n$, $m\in\mathbb{N}$ and $\{L_i\}_{i\in\{1,\ldots,m\}}\subseteq\mathrm{Hom}(X,X)$ then*

$$\det(L_1\circ\cdots\circ L_n)\underset{\text{def}}{=}\det\left(\prod_{i=1}^m L_i\right)=\prod_{i\in\{1,\ldots,m\}}\det(L_i)$$

Proof. We prove this by induction on m so let

$$S=\left\{m\in\mathbb{N}|\text{If }\{L_i\}_{i\in\{1,\ldots,m\}}\subseteq\operatorname{Hom}(X,X)\text{ then }\det\left(\prod_{i=1}^{m}L_i\right)=\prod_{i\in\{1,\ldots,m\}}\det(L_i)\right\}$$

then we have:

$1\in S$. If $\{L_i\}_{i\in\{1\}}\subseteq\operatorname{Hom}(X,X)$ then we have

$$\det\left(\prod_{i=1}^{1}L_i\right)=\det(L_1)=\prod_{i\in\{1\}}\det(L_i)$$

proving that $1\in S$.

$m\in S\Rightarrow m+1\in S$. Let $\{L_i\}_{i\in\{1,\ldots,m+1\}}\subseteq\operatorname{Hom}(X,X)$ then we have

$$\begin{aligned}\det\left(\prod_{i=1}^{m+1}L_i\right) &= \det\left(\left(\prod_{i=1}^{m}L_i\right)\circ L_{m+1}\right)\\ &\underset{[\text{theorem: }11.299]}{=} \det\left(\prod_{i=1}^{m}L_i\right)\circ\det(L_{m+1})\\ &\underset{m\in S}{=} \left(\prod_{i\in\{1,\ldots,m\}}\det(L_i)\right)\cdot\det(L_{m+1})\\ &= \prod_{i\in\{1,\ldots,m+1\}}\det(L_i)\end{aligned}$$

proving that $m+1\in S$. □

Definition 11.301. *Let $n\in\mathbb{N}$, X a vector space over a field F with characteristic zero such that $\dim(X)=n$, Δ a non trivial determinant function and $L\in\operatorname{Hom}(X,X)$ then we define*

$$\overline{\Delta L}: X^n\rightarrow\operatorname{Hom}(X,X)\text{ where }\overline{\Delta L}(x_1,\ldots,x_n): X\rightarrow X\text{ is defined by}$$

$$\overline{\Delta L}(x_1,\ldots,x_n)(y)=\sum_{i\in\{1,\ldots,n\}}(-1)^{i-1}\cdot\Delta(y,L(x_1),\ldots,L(x_{i-1}),L(x_{i+1}),\ldots,L(x_n))\cdot x_i$$

Proof. Of course we must prove that $\overline{\Delta L}(x_1,\ldots,x_n)\in\operatorname{Hom}(X,X)$. So let $\alpha\in F$, $u,v\in X$ then we have

$$\begin{aligned}&\overline{\Delta L}(x_1,\ldots,x_n)(u+v) =\\ &\sum_{i\in\{1,\ldots,n\}}(-1)^{i-1}\cdot\Delta(u+v,L(x_1),\ldots,L(x_{i-1}),L(x_{i+1}),\ldots,L(x_n))\cdot x_i =\\ &\sum_{i\in\{1,\ldots,n\}}(-1)^{i-1}\cdot(\Delta(u,L(x_1),\ldots,L(x_{i-1}),L(x_{i+1}),\ldots,L(x_n))+\Delta(v,L(x_1),\ldots,L(x_{i-1}),L(x_{i+1}),\ldots,L(x_n))) \underset{[\text{theorem: }11.38]}{=}\\ &\sum_{i\in\{1,\ldots,n\}}(-1)^{i-1}\cdot\Delta(u,L(x_1),\ldots,L(x_{i-1}),L(x_{i+1}),\ldots,L(x_n))\cdot x_i+\sum_{i\in\{1,\ldots,n\}}(-1)^{i}\cdot\Delta(v,L(x_1),\ldots,L(x_{i-1}),L(x_{i+1}),\ldots,\\ &L(x_n))\cdot x_i =\\ &\overline{\Delta L}(x_1,\ldots,x_n)(u)+\overline{\Delta L}(x_1,\ldots,x_n)(v)\end{aligned}$$

and

$$\begin{aligned}&\overline{\Delta L}(x_1,\ldots,x_n)(\alpha\cdot u) =\\ &\sum_{i\in\{1,\ldots,n\}}(-1)^{i-1}\cdot\Delta(\alpha\cdot u,L(x_1),\ldots,L(x_{i-1}),L(x_{i+1}),\ldots,L(x_n))\cdot x_i =\\ &\sum_{i\in\{1,\ldots,n\}}(-1)^{i-1}\cdot\alpha\cdot\Delta(u,L(x_1),\ldots,L(x_{i-1}),L(x_{i+1}),\ldots,L(x_n))\cdot x_i \underset{[\text{theorem: }11.75]}{=}\\ &\alpha\cdot\sum_{i\in\{1,\ldots,n\}}(-1)^{i-1}\cdot\Delta(u,L(x_1),\ldots,L(x_{i-1}),L(x_{i+1}),\ldots,L(x_n))\cdot x_i =\\ &\alpha\cdot\overline{\Delta L}(x_1,\ldots,x_n)(u)\end{aligned}$$

□

Lemma 11.302. *Let $n\in\mathbb{N}$, X a set, $x=(x_1,\ldots,x_n)\in X^n$, $i,j\in\{1,\ldots,n\}$ such that $i\neq j$ and $x_i=x_j$ and $t\in X$ then we have*

1. *If $j<i$ then*
$$(t,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)\circ\left(i\underset{n}{\rightsquigarrow}j+1\right)=(t,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)$$
2. *If $i<j$ then*
$$(t,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)\circ\left(i+1\underset{n}{\rightsquigarrow}j\right)=(t,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)$$

Proof.

1. Let $j<i$. Define
$$A=(t,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)\circ\left(i\underset{n}{\rightsquigarrow}j+1\right) \tag{11.198}$$
then we have as $j+1\leqslant i$ the following cases to consider for i:

   ***j*+1=*i*.** Then by [definition: 11.236(1)] we have that $\left(i\underset{n}{\rightsquigarrow}j+1\right)=\mathrm{Id}_{\{1,\ldots,n\}}$ so that
$$A=(t,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n) \tag{11.199}$$
   Now for $k\in\{1,\ldots,n\}$ we have either (motivated by [definition: 11.286] the following possible cases

   ***k*=1.** Then
$$\begin{aligned}A_k &\underset{[\text{eq: } 11.199]}{=} (t,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_k\\ &\underset{[\text{definition: } 11.286]\wedge k=1}{=} t\\ &\underset{[\text{definition: } 11.286]\wedge k=1}{=} (t,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_k\end{aligned}$$

   2⩽*k*⩽*j*. Then
$$\begin{aligned}A_k &\underset{[\text{eq: } 11.199]}{=} (t,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_k\\ &\underset{[\text{definition: } 11.286]\wedge 2\leqslant k<j}{=} x_{k-1}\\ &\underset{[\text{definition: } 11.286]\wedge 2\leqslant k\leqslant j=i-1<i}{=} (t,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_k\end{aligned}$$

   ***k*=*j*+1.** Then
$$\begin{aligned}A_k &\underset{[\text{eq: } 11.199]}{=} (t,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_k\\ &\underset{[\text{definition: } 11.286]\wedge i=j+1=k}{=} x_k\\ &\underset{k=j+1=i}{=} x_i\\ &\underset{x_i=x_j}{=} x_j\\ &\underset{j=k-1}{=} x_{k-1}\\ &\underset{[\text{definition: } 11.286]i=j+1=k}{=} (t,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_k\end{aligned}$$

   ***j*+1<*k*.** Then
$$\begin{aligned}A_k &\underset{[\text{eq: } 11.199]}{=} (t,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_k\\ &\underset{[\text{definition: } 11.286]j+1<k}{=} x_k\\ &\underset{[\text{definition: } 11.286]i=j+1<k\Rightarrow i+1\leqslant k}{=} (t,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_k\end{aligned}$$

   ***j*+1<*i*.** Then we have motivated by [definition: 11.236(3)] the following possible cases for k:

   ***k*=1.** Then
$$\begin{aligned}A_k &= (t,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_{(i\underset{n}{\rightsquigarrow}j+1)(k)}\\ &\underset{[\text{definition: } 11.236(3)]\wedge k=1<j+1}{=} (t,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_k\\ &\underset{[\text{definition: } 11.286]\wedge k=1}{=} t\\ &\underset{[\text{definition: } 11.286]\wedge k=1}{=} (t,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_k\end{aligned}$$

$\boldsymbol{k \neq 1 \wedge k < j+1.}$ Then

$$\begin{array}{rcl} A_k & = & (t, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_{(i \underset{n}{\rightsquigarrow} j+1)(k)} \\ & \underset{\text{[definition: 11.236(3)]} \wedge k < j+1}{=} & (t, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_k \\ & \underset{\text{[definition: 11.286]} \wedge 1 < k < j+1 \Rightarrow 1 < k \leqslant j}{=} & x_{k-1} \\ & \underset{\text{[definition: 11.286]} \wedge 1 < k < j+1 = i}{=} & (t, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_k \end{array}$$

$\boldsymbol{k \neq 1 \wedge k = j+1.}$ Then

$$\begin{array}{rcl} A_k & = & (t, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_{(i \underset{n}{\rightsquigarrow} j+1)(k)} \\ & \underset{\text{[definition: 11.236(3)]} \wedge k = j+1}{=} & (t, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_i \\ & \underset{\text{[definition: 11.286]} \wedge j+1 < i}{=} & x_i \\ & \underset{x_i = x_j}{=} & x_j \\ & \underset{j = k-1}{=} & x_{k-1} \\ & \underset{1 < k = j+1 < i}{=} & (t, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_k \end{array}$$

$\boldsymbol{k \neq 1 \wedge j+1 < k \leqslant i.}$ Then

$$\begin{array}{rcl} A_k & = & (t, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_{(i \underset{n}{\rightsquigarrow} j+1)(k)} \\ & \underset{\text{[definition: 11.236(3)]} \wedge j+1 < k \leqslant i}{=} & (t, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_{k-1} \\ & \underset{\text{[definition: 11.286]} \wedge j+1 < k \Rightarrow j+1 \leqslant k-1}{=} & x_{k-1} \\ & \underset{\text{[definition: 11.286]} \wedge 1 < k \leqslant i}{=} & (t, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_k \end{array}$$

$\boldsymbol{k \neq 1 \wedge i < k.}$ Then

$$\begin{array}{rcl} A_k & = & (t, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_{(i \underset{n}{\rightsquigarrow} j+1)(k)} \\ & \underset{\text{[definition: 11.236(3)]} \wedge i < k \leqslant n}{=} & (t, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_k \\ & \underset{\text{[definition: 11.286]} \wedge j+1 < i < k}{=} & x_k \\ & \underset{\text{[definition: 11.286]} \wedge i < k \Rightarrow i+1 \leqslant k}{=} & (t, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_k \end{array}$$

So in all possible cases we have

$$(t, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_{(i \underset{n}{\rightsquigarrow} j+1)(k)} = (t, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_k$$

proving that we have

$$(t, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \circ \left(i \underset{n}{\rightsquigarrow} j+1\right) = (t, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$$

2. Let $i < j$. Define

$$B = (t, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \circ \left(i+1 \underset{n}{\rightsquigarrow} j\right) \tag{11.200}$$

then we have as $i+1 \leqslant j$ the following cases to consider for j:

$\boldsymbol{i+1 = j.}$ Then by [definition: 11.236(1)] we have that $\left(i \underset{n}{\rightsquigarrow} j+1\right) = \mathrm{Id}_{\{1,\ldots,n\}}$ so that

$$B = (t, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \tag{11.201}$$

Now for $k \in \{1, \ldots, n\}$ we have either (motivated by [definition: 11.286] following cases:

$\boldsymbol{k = 1.}$ Then

$$\begin{array}{rcl} B_k & \underset{\text{[eq: 11.201]}}{=} & (t, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_k \\ & \underset{\text{[definition: 11.286]} \wedge k = 1}{=} & t \\ & \underset{\text{[definition: 11.286]} \wedge k = 1}{=} & (t, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_k \end{array}$$

$1<k\leqslant i$. Then

$$\begin{aligned} B_k &\underset{\text{[eq: 11.201]}}{=} (t,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_k \\ &\underset{\text{[definition: 11.286]}\wedge 1<k\leqslant i<i+1=j}{=} x_{k-1} \\ &\underset{\text{[definition: 11.286]}\wedge 1<k\leqslant i}{=} (t,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_k \end{aligned}$$

$k=i+1$. Then

$$\begin{aligned} B_k &\underset{\text{[eq: 11.201]}}{=} (t,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_k \\ &\underset{\text{[definition: 11.286]}\wedge 1<k=i+1=j}{=} x_{k-1} \\ &\underset{k-1=i}{=} x_i \\ &\underset{x_i=x_j}{=} x_j \\ &\underset{j=i+1=k}{=} x_k \\ &\underset{\text{[definition: 11.286]}\wedge k=i+1}{=} (t,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_k \end{aligned}$$

$i+1<k$. Then

$$\begin{aligned} B_k &\underset{\text{[eq: 11.201]}}{=} (t,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_k \\ &\underset{\text{[definition: 11.286]}\wedge j=i+1<k\Rightarrow j+1\leqslant k}{=} x_k \\ &\underset{\text{[definition: 11.286]}\wedge i+1<k}{=} (t,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_k \end{aligned}$$

$i+1<j$. Then we have motivated by [definition: 11.236(2)] the following possible cases for k:

$k=1$. Then

$$\begin{aligned} B_k &= (t,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_{(i+1 \underset{n}{\rightsquigarrow} j)(k)} \\ &\underset{\text{[theorem: 11.236(2)]}\wedge 1=k<i+1}{=} (t,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_k \\ &\underset{\text{[definition: 11.286]}\wedge k=1}{=} t \\ &\underset{\text{[definition: 11.286]}\wedge k=1}{=} (t,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_k \end{aligned}$$

$k\neq 1\wedge k<i+1$. Then

$$\begin{aligned} B_k &= (t,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_{(i+1 \underset{n}{\rightsquigarrow} j)(k)} \\ &\underset{\text{[definition: 11.236(2)]}\wedge 1<k<i+1}{=} (t,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_k \\ &\underset{\text{[definition: 11.286]}\wedge 1<k<i+1<j}{=} x_{k-1} \\ &\underset{\text{[definition: 11.286]}\wedge 1<k<i+1\Rightarrow 1<k\leqslant i}{=} (t,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_k \end{aligned}$$

$k\neq 1\wedge i+1\leqslant k<j$. Then

$$\begin{aligned} B_k &= (t,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_{(i+1 \underset{n}{\rightsquigarrow} j)(k)} \\ &\underset{\text{[definition: 11.236(2)]}\wedge i+1\leqslant k<j}{=} (t,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_{k+1} \\ &\underset{\text{[definition: 11.286]}\wedge 1<k<j\Rightarrow 1<k+1\leqslant j}{=} x_{(k+1)-1} \\ &= x_k \\ &\underset{\text{[definition: 11.286]}\wedge i+1\leqslant k}{=} (t,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_k \end{aligned}$$

$\boldsymbol{k \neq 1 \wedge k = j}$**.** Then

$$\begin{array}{lll} B_k & = & (t, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_{(i+1 \underset{n}{\rightsquigarrow} j)(k)} \\ & \underset{[\text{definition: } 11.236(2)] \wedge k=j}{=} & (t, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_{i+1} \\ & \underset{[\text{definition: } 11.286] \wedge 1<i+1<j}{=} & x_{(i+1)-1} \\ & = & x_i \\ & \underset{x_i = x_j}{=} & x_j \\ & \underset{j=k}{=} & x_k \\ & \underset{[\text{definition: } 11.286] \wedge i<i+1<j=k}{=} & (t, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_k \end{array}$$

$\boldsymbol{k \neq 1 \wedge j < k}$**.** Then

$$\begin{array}{lll} B_k & = & (t, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_{(i+1 \underset{n}{\rightsquigarrow} j)(k)} \\ & \underset{[\text{definition: } 11.236] \wedge j<k}{=} & (t, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_k \\ & \underset{[\text{definition: } 11.286] \wedge j<k \rightarrow j+1 \leqslant k}{=} & x_k \\ & \underset{[\text{definition: } 11.286] \wedge i+1<j<k}{=} & (t, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_k \end{array}$$

So in all possible cases we have

$$(t, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_{(i+1 \underset{n}{\rightsquigarrow} j)(k)} = (t, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_k$$

proving that we have

$$(t, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \circ \left(i+1 \underset{n}{\rightsquigarrow} j\right) = (t, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$$ □

The above lemma will be used to prove the following theorem that will be used to define the adjoint of a linear transformation. TODO check this

Theorem 11.303. *Let $n \in \mathbb{N}$, X a vector space over a field of characteristic zero with $\dim(X) = n$, Δ a non trivial determinant function, $L \in \operatorname{Hom}(X, X)$ then*

$$\overline{\Delta L} \in \operatorname{Hom}^n(X; \operatorname{Hom}(X, X)) \text{ and } \overline{\Delta L} \text{ is skew-symmetric}$$

Proof. Given $x \in X^n$, $t \in X$ then we have by definition

$$\begin{aligned} \overline{\Delta L}(x)(t) &= \sum_{j \in \{1, \ldots, n\}} (-1)^{j-1} \cdot \Delta(t, L(x_1), \ldots, L(x_{j-1}), L(x_{j+1}), \ldots, L(x_n)) \cdot x_j \\ &= \sum_{j \in \{1, \ldots, n\}} (-1)^{j-1} \cdot A_j(t, x) \end{aligned} \tag{11.202}$$

where

$$A_j(t, x) = \Delta(t, L(x_1), \ldots, L(x_{j-1}), L(x_{j+1}), \ldots, L(x_n)) \cdot x_j \tag{11.203}$$

Let $x = (x_1, \ldots, x_n) \in X^n$, $i \in \{1, \ldots, n\}$, $\alpha \in F$ and $u, v, t \in X$. Define

$$\begin{aligned} y &= \left(x_1, \ldots, \underbrace{u + \alpha \cdot v}_{i} \ldots, x_n\right) \\ r &= \left(x_1, \ldots \underbrace{u}_{i} \ldots, x_n\right) \\ s &= \left(x_1, \ldots \underbrace{v}_{i} \ldots, x_n\right) \end{aligned}$$

Let $j \in \{1, \ldots, n\}$ and define

$$\begin{array}{rcl} y^j & = & (t, L(y_1), \ldots, L(y_{j-1}), L(y_{j+1}), \ldots, L(y_n)) \\ r^j & = & (t, L(r_1), \ldots, L(r_{j-1}), L(r_{j+1}), \ldots, L(r_n)) \\ s^j & = & (t, L(s_1), \ldots, L(s_{j-1}), L(s_{j+1}), \ldots, L(s_n)) \end{array}$$

then

$$A_j(t, y) = \Delta(y^j) \cdot y_j \wedge A_j(t, r) = \Delta(r^j) \cdot r_j \wedge A_j(t, s) = \Delta(s^j) \cdot s_j \tag{11.204}$$

Consider the following possible cases for $j \in \{1, \ldots, n\}$:

$\boldsymbol{j = i}$**.** Then for $k \in \{1, \ldots, n\}$ we have either:

$\boldsymbol{k = 1}$**.** Then

$$\begin{array}{rcl} (y^j)_k & = & (t, L(y_1), \ldots, L(y_{j-1}), L(y_{j+1}), \ldots, L(y_n))_1 \\ & \underset{[\text{definition: } 11.286] \wedge k=1}{=} & t \\ (r^j)_k & = & (t, L(r_1), \ldots, L(r_{j-1}), L(r_{j+1}), \ldots, L(r_n))_1 \\ & \underset{[\text{definition: } 11.286] \wedge k=1}{=} & t \\ (s^j)_k & = & (t, L(s_1), \ldots, L(s_{j-1}), L(s_{j+1}), \ldots, L(s_n))_1 \\ & \underset{[\text{definition: } 11.286] \wedge k=1}{=} & t \end{array}$$

$\boldsymbol{2 \leqslant k \leqslant j}$**.** Then

$$\begin{array}{rcl} (y^j)_k & = & (t, L(y_1), \ldots, L(y_{j-1}), L(y_{j+1}), \ldots, L(y_n))_k \\ & \underset{[\text{definition: } 11.286] \wedge 2 \leqslant k \leqslant j}{=} & L(y_{k-1}) \\ & \underset{k-1 < k \leqslant j = i \Rightarrow k-1 \neq i}{=} & L(x_{k-1}) \\ (r^j)_k & = & (t, L(r_1), \ldots, L(r_{j-1}), L(r_{j+1}), \ldots, L(r_n))_k \\ & \underset{[\text{definition: } 11.286] \wedge 2 \leqslant k \leqslant j}{=} & L(r_{k-1}) \\ & \underset{k-1 < k \leqslant j = i \Rightarrow k-1 \neq i}{=} & L(x_{k-1}) \\ (s^j)_k & = & (t, L(s_1), \ldots, L(s_{j-1}), L(s_{j+1}), \ldots, L(s_n))_k \\ & \underset{[\text{definition: } 11.286] \wedge 2 \leqslant k \leqslant j}{=} & L(s_{k-1}) \\ & \underset{k-1 < k \leqslant j = i \Rightarrow k-1 \neq i}{=} & L(x_{k-1}) \end{array}$$

$\boldsymbol{j + 1 \leqslant k}$**.** Then

$$\begin{array}{rcl} (y^j)_k & = & (t, L(y_1), \ldots, L(y_{j-1}), L(y_{j+1}), \ldots, L(y_n))_k \\ & \underset{[\text{definition: } 11.286] \wedge j+1 \leqslant k}{=} & L(y_k) \\ & \underset{i = j < j+1 \leqslant k \Rightarrow i \neq k}{=} & L(x_k) \\ (r^j)_k & = & (t, L(r_1), \ldots, L(r_{j-1}), L(r_{j+1}), \ldots, L(r_n))_k \\ & \underset{[\text{definition: } 11.286] \wedge j+1 \leqslant k}{=} & L(r_k) \\ & \underset{i = j < j+1 \leqslant k \Rightarrow i \neq k}{=} & L(x_k) \\ (s^j)_k & = & (t, L(s_1), \ldots, L(s_{j-1}), L(s_{j+1}), \ldots, L(s_n))_k \\ & \underset{[\text{definition: } 11.286] \wedge j+1 \leqslant k}{=} & L(s_k) \\ & \underset{i = j < j+1 \leqslant k \Rightarrow i \neq k}{=} & L(x_k) \end{array}$$

So $\forall k \in \{1, \ldots, n\}$ we have $(y^j)_k = (r^j)_k = (s^j)_k$ or

$$y^j = r^j = s^j \tag{11.205}$$

Hence

$$\begin{array}{rcl} A_j(t,y) & \underset{[\text{eq: } 11.204]}{=} & \Delta(y^j)\cdot y_j \\ & \underset{i=j}{=} & \Delta(y^j)\cdot y_i \\ & = & \Delta(y^j)\cdot(u+\alpha\cdot v) \\ & = & \Delta(y^j)\cdot u+\alpha\cdot\Delta(y^j)\cdot v \\ & \underset{[\text{eq: } 11.205]}{=} & \Delta(r^j)\cdot u+\alpha\cdot\Delta(s^j)\cdot v \\ & = & \Delta(r^j)\cdot r_i+\alpha\cdot\Delta(s^j)\cdot s_i \\ & \underset{i=j}{=} & \Delta(r^j)\cdot r_j+\alpha\cdot\Delta(s^j)\cdot s_j \\ & = & A_j(t,r)+A_j(t,s) \end{array}$$

proving that in this case

$$A_j(t,y)=A_j(t,r)+A_j(t,s) \tag{11.206}$$

j < i. Then we have

$$\begin{array}{rcl} (y^j)_i & = & (t,L(y_1),\ldots,L(y_{j-1}),L(y_{j+1}),\ldots,L(y_n))_i \\ & \underset{[\text{definition: } 11.286]\wedge 1<j+1\leqslant i}{=} & L(y_i) \\ & = & L(u+\alpha\cdot v) \\ & = & L(u)+\alpha\cdot L(v) \end{array}$$

hence $y^j=\left((y^j)_1,\ldots\underbrace{L(u)+\alpha\cdot L(v)}_{i}\ldots,(y^i)_n\right)$ so that

$$\begin{array}{rl} \Delta(y^j) & = \\ \Delta\left((y^j)_1,\ldots\underbrace{L(u)+\alpha\cdot L(v)}_{i}\ldots,(y^i)_n\right) & = \\ \Delta\left((y^j)_1,\ldots\underbrace{L(u)}_{i}\ldots,(y^i)_n\right)+\alpha\cdot\Delta\left((y^j)_1,\ldots\underbrace{L(v)}_{i}\ldots,(y^i)_n\right) & \end{array} \tag{11.207}$$

Further we have for $k\in\{1,\ldots,n\}$ either:

k = 1. Then

$$\begin{array}{rc} \left((y^j)_1\ldots\underbrace{\boldsymbol{L(u)}}_{i}\ldots,(y^i)_n\right)_k & \underset{1\leqslant j<i\Rightarrow i\neq 1}{=} \\ (y^j)_1 & = \\ (t,L(y_1),\ldots,L(y_{j-1}),L(y_{j+1}),\ldots,L(y_n))_1 & \underset{[\text{definition: } 11.286]\wedge k=1}{=} \\ t & \end{array}$$

and

$$\begin{array}{rcl} (r^j)_k & = & (t,L(r_1),\ldots,L(r_{j-1}),L(r_{j+1}),\ldots,L(r_n)) \\ & \underset{[\text{definition: } 11.286]\wedge k=1}{=} & t \end{array}$$

and

$$\begin{array}{rc} \left((y^j)_1,\ldots\underbrace{\boldsymbol{L(v)}}_{i}\ldots,(y^i)_n\right)_k & \underset{1\leqslant j<i\Rightarrow i\neq 1}{=} \\ (y^j)_1 & = \\ (t,L(y_1),\ldots,L(y_{j-1}),L(y_{j+1}),\ldots,L(y_n))_1 & \underset{[\text{definition: } 11.286]\wedge k=1}{=} \\ t & \end{array}$$

and

$$\begin{array}{rcl}(s^j)_k & = & (t, L(s_1), \ldots, L(s_{j-1}), L(s_{j+1}), \ldots, L(s_n)) \\ & \underset{[\text{definition: } 11.286] \wedge k=1}{=} & t\end{array}$$

Hence we have in this case that

$$\begin{array}{rcl}\left((y^j)_1, \ldots \underbrace{L(u)}_{i} \ldots, (y^i)_n\right)_k & = & (r^j)_k \\ \left((y^j)_1, \ldots \underbrace{L(v)}_{i} \ldots, (y^i)_n\right)_k & = & (s^j)_k\end{array}$$

$2 \leqslant k \leqslant j$. Then

$$\begin{array}{rcl}\left((y^j)_1, \ldots \underbrace{\boldsymbol{L(u)}}_{i} \ldots, (y^i)_n\right)_k & \underset{k \leqslant j < i \Rightarrow k \neq i}{=} & \\ (y^j)_k & = & \\ (t, L(y_1), \ldots, L(y_{j-1}), L(y_{j+1}), \ldots, L(y_n))_k & \underset{[\text{definition: } 11.286]}{=} & \\ L(y_{k-1}) & \underset{k-1 < k \leqslant j < i \Rightarrow k \neq i}{=} & \\ L(x_{k-1}) & & \end{array}$$

and

$$\begin{array}{rcl}\left((y^j)_1, \ldots \underbrace{\boldsymbol{L(v)}}_{i} \ldots, (y^i)_n\right)_k & \underset{k \leqslant j < i \Rightarrow k \neq i}{=} & \\ (y^j)_k & = & \\ (t, L(y_1), \ldots, L(y_{j-1}), L(y_{j+1}), \ldots, L(y_n))_k & \underset{[\text{definition: } 11.286]}{=} & \\ L(y_{k-1}) & \underset{k-1 < k \leqslant j < i \Rightarrow k \neq i}{=} & \\ L(x_{k-1}) & & \end{array}$$

and

$$\begin{array}{rcl}(r^j)_k & = & (t, L(r_1), \ldots, L(r_{j-1}), L(r_{j+1}), \ldots, L(r_n))_k \\ & \underset{[\text{definition: } 11.286]}{=} & L(r_{k-1}) \\ & \underset{k-1 < k \leqslant j < i \Rightarrow k \neq i}{=} & L(x_{k-1}) \\ (s^j)_k & = & (t, L(s_1), \ldots, L(s_{j-1}), L(s_{j+1}), \ldots, L(s_n))_k \\ & \underset{[\text{definition: } 11.286]}{=} & L(s_{k-1}) \\ & \underset{k-1 < k \leqslant j < i \Rightarrow k \neq i}{=} & L(x_{k-1})\end{array}$$

Hence we have that

$$\begin{array}{rcl}\left((y^j)_1, \ldots \underbrace{L(u)}_{i} \ldots, (y^i)_n\right)_k & = & (r^j)_k \\ \left((y^j)_1, \ldots \underbrace{L(v)}_{i} \ldots, (y^i)_n\right)_k & = & (s^j)_k\end{array}$$

$\boldsymbol{j+1\leqslant k<i.}$ Then

$$\begin{array}{rcl}\left((y^j)_1,\ldots\underbrace{\boldsymbol{L(u)}}_{i}\ldots,(y^i)_n\right)_k & \underset{k<i\Rightarrow k\neq i}{=} & \\ (y^j)_k & = & \\ (t,L(y_1),\ldots,L(y_{j-1}),L(y_{j+1}),\ldots,L(y_n))_k & \underset{[\text{definition: } 11.286]}{=} & \\ L(y_k) & \underset{k<i\Rightarrow k\neq i}{=} & \\ L(x_k) & & \end{array}$$

and

$$\begin{array}{rcl}\left((y^j)_1,\ldots\underbrace{\boldsymbol{L(v)}}_{i}\ldots,(y^i)_n\right)_k & \underset{k<i\Rightarrow k\neq i}{=} & \\ (y^j)_k & = & \\ (t,L(y_1),\ldots,L(y_{j-1}),L(y_{j+1}),\ldots,L(y_n))_k & \underset{[\text{definition: } 11.286]}{=} & \\ L(y_k) & \underset{k<i\Rightarrow k\neq i}{=} & \\ L(x_k) & & \end{array}$$

and

$$\begin{array}{rcl}(r^j)_k & = & (t,L(r_1),\ldots,L(r_{j-1}),L(r_{j+1}),\ldots,L(r_n))_k\\ & \underset{[\text{definition: } 11.286]}{=} & L(r_k)\\ & \underset{k<i\Rightarrow k\neq i}{=} & L(x_k)\\ (s^j)_k & = & (t,L(s_1),\ldots,L(s_{j-1}),L(s_{j+1}),\ldots,L(s_n))_k\\ & \underset{[\text{definition: } 11.286]}{=} & L(s_k)\\ & \underset{k<i\Rightarrow k\neq i}{=} & L(x_k)\end{array}$$

Hence we have that

$$\begin{array}{rcl}\left((y^j)_1,\ldots\underbrace{L(u)}_{i}\ldots,(y^i)_n\right)_k & = & (r^j)_k\\ \left((y^j)_1,\ldots\underbrace{L(v)}_{i}\ldots,(y^i)_n\right)_k & = & (s^j)_k\end{array}$$

$\boldsymbol{k=i.}$ Then

$$\begin{array}{rcl}\left((y^j)_1,\ldots\underbrace{\boldsymbol{L(u)}}_{i}\ldots,(y^i)_n\right)_k & \underset{k=i}{=} & L(u)\\ \left((y^j)_1,\ldots\underbrace{\boldsymbol{L(v)}}_{i}\ldots,(y^i)_n\right)_k & \underset{k=i}{=} & L(v)\end{array}$$

and

$$\begin{array}{rcl}(r^j)_k & = & (t,L(r_1),\ldots,L(r_{j-1}),L(r_{j+1}),\ldots,L(r_n))_k\\ & \underset{[\text{definition: } 11.286]\wedge j<i=k\Rightarrow j+1\leqslant k}{=} & L(r_k)\\ & \underset{i=k}{=} & L(u)\\ (s^j)_k & = & (t,L(s_1),\ldots,L(s_{j-1}),L(s_{j+1}),\ldots,L(s_n))_k\\ & \underset{[\text{definition: } 11.286]\wedge j<i=k\Rightarrow j+1\leqslant k}{=} & L(s_k)\\ & = & L(v)\end{array}$$

Hence we have that

$$\begin{array}{rcl}\left((y^j)_1,\ldots\underbrace{L(u)}_{i}\ldots,(y^i)_n\right)_k & = & (r^j)_k\\ \left((y^j)_1,\ldots\underbrace{L(v)}_{i}\ldots,(y^i)_n\right)_k & = & (s^j)_k\end{array}$$

$\boldsymbol{i+1\leqslant k.}$ Then

$$\begin{array}{rl}\left((y^j)_1,\ldots\underbrace{\boldsymbol{L(u)}}_{i}\ldots,(y^i)_n\right)_k & \underset{i<i+1\leqslant k\Rightarrow k\neq i}{=}\\ (y^j)_k & =\\ (t,L(y_1),\ldots,L(y_{j-1}),L(y_{j+1}),\ldots,L(y_n))_k & \underset{[\text{definition: } 11.286]}{=}\\ L(y_k) & \underset{i<i+1\leqslant k\Rightarrow k\neq i}{=}\\ L(x_k) & \end{array}$$

and

$$\begin{array}{rl}\left((y^j)_1,\ldots\underbrace{\boldsymbol{L(v)}}_{i}\ldots,(y^i)_n\right)_k & \underset{i<i+1\leqslant k\Rightarrow k\neq i}{=}\\ (y^j)_k & =\\ (t,L(y_1),\ldots,L(y_{j-1}),L(y_{j+1}),\ldots,L(y_n))_k & \underset{[\text{definition: } 11.286]}{=}\\ L(y_k) & \underset{i<i+1\leqslant k\Rightarrow k\neq i}{=}\\ L(x_k) & \end{array}$$

and

$$\begin{array}{rcl}(r^j)_k & = & (t,L(r_1),\ldots,L(r_{j-1}),L(r_{j+1}),\ldots,L(r_n))_k\\ & \underset{[\text{definition: } 11.286]}{=} & L(r_k)\\ & \underset{i<i+1\leqslant k\Rightarrow k\neq i}{=} & L(x_k)\\ (s^j)_k & = & (t,L(s_1),\ldots,L(s_{j-1}),L(s_{j+1}),\ldots,L(s_n))_k\\ & \underset{[\text{definition: } 11.286]}{=} & L(s_k)\\ & \underset{i<i+1\leqslant k\Rightarrow k\neq i}{=} & L(x_k)\end{array}$$

Hence we have that

$$\begin{array}{rcl}\left((y^j)_1,\ldots\underbrace{L(u)}_{i}\ldots,(y^i)_n\right)_k & = & (r^j)_k\\ \left((y^j)_1,\ldots\underbrace{L(v)}_{i}\ldots,(y^i)_n\right)_k & = & (s^j)_k\end{array}$$

So we have proved that $\forall k\in\{1,\ldots,n\}$

$$\begin{array}{rcl}\left((y^j)_1,\ldots\underbrace{L(u)}_{i}\ldots,(y^i)_n\right)_k & = & (r^j)_k\\ \left((y^j)_1,\ldots\underbrace{L(v)}_{i}\ldots,(y^i)_n\right)_k & = & (s^j)_k\end{array}$$

or

$$\begin{aligned}\left((y^j)_1,\ldots\underbrace{L(u)}_{i}\ldots,(y^i)_n\right) &= r^j\\ \left((y^j)_1,\ldots\underbrace{L(v)}_{i}\ldots,(y^i)_n\right) &= s^j\end{aligned}$$

which combined with [eq: 11.207] proves that

$$\Delta(y^j)=\Delta(r^j)+\alpha\cdot\Delta(s^j)$$

So that

$$\begin{aligned}A_j(t,y) &= \Delta(y^j)\cdot y_j\\ &= (\Delta(r^j)+\alpha\cdot\Delta(s^j))\cdot y_j\\ &= \Delta(r^j)\cdot y_j+\alpha\cdot\Delta(s^j)\cdot y_j\\ &\underset{j<i\Rightarrow i\neq j}{=} \Delta(r^j)\cdot r_j+\alpha\cdot\Delta(s^j)\cdot s_j\\ &= A_j(t,r)+\alpha\cdot A_j(t,s)\end{aligned}$$

or summarized

$$A_j(t,y)=A_j(t,r)+\alpha\cdot A_j(t,s) \tag{11.208}$$

***i* < *j*.** Then we have

$$\begin{aligned}(y^j)_{i+1} &= (t,L(y_1),\ldots,L(y_{j-1}),L(y_{j+1}),\ldots,L(y_n))_{i+1}\\ &\underset{[\text{definition: } 11.286]\wedge i<j\Rightarrow 1<i+1\leqslant j}{=} L(y_{(i+1)-1})\\ &= L(y_i)\\ &= L(u+\alpha\cdot v)\\ &= L(u)+\alpha\cdot L(v)\end{aligned}$$

hence $y^j=\left((y^j)_1,\ldots\underbrace{L(u)+\alpha\cdot L(v)}_{i+1}\ldots,(y^i)_n\right)$ so that

$$\begin{aligned}\Delta(y^j) &=\\ \Delta((y^j)_1,\ldots,(y^j)_i,L(u)+\alpha\cdot L(v),(y^j)_{i+2},\ldots,(y^i)_n) &=\\ \Delta\left((y^j)_1,\ldots\underbrace{L(u)}_{i+1}\ldots,(y^i)_n\right)+\alpha\cdot\Delta\left((y^j)_1,\ldots\underbrace{\alpha\cdot L(v)}_{i+1}\ldots,(y^i)_n\right) &=\end{aligned} \tag{11.209}$$

Further we have for $k\in\{1,\ldots,n\}$ either:

***k* = 1.** Then

$$\begin{aligned}\left((y^j)_1,\ldots\underbrace{\boldsymbol{L(u)}}_{i+1}\ldots,(y^i)_n\right)_k &\underset{k=1<i+1\Rightarrow k\neq i+1}{=}\\ (y^j)_1 &=\\ (t,L(y_1),\ldots,L(y_{j-1}),L(y_{j+1}),\ldots,L(y_n))_1 &\underset{[\text{definition: } 11.286]}{=}\\ t\end{aligned}$$

and

$$\begin{aligned}\left((y^j)_1,\ldots\underbrace{\boldsymbol{L(v)}}_{i+1}\ldots,(y^i)_n\right)_k &\underset{k=1<i+1\Rightarrow k\neq i+1}{=}\\ (y^j)_1 &=\\ (t,L(y_1),\ldots,L(y_{j-1}),L(y_{j+1}),\ldots,L(y_n))_1 &\underset{[\text{definition: } 11.286]}{=}\\ t\end{aligned}$$

and

$$\begin{array}{rcl} (r^j)_k & = & (t, L(r_1), \ldots, L(r_{j-1}), L(r_{j+1}), \ldots, L(r_n))_k \\ & \underset{[\text{definition: } 11.286] \wedge k=1}{=} & t \\ (s^j)_k & = & (t, L(s_1), \ldots, L(s_{j-1}), L(s_{j+1}), \ldots, L(s_n))_k \\ & \underset{[\text{definition: } 11.286] \wedge k=1}{=} & t \end{array}$$

Hence we have

$$\begin{array}{rcl} \left((y^j)_1, \ldots \underbrace{\boldsymbol{L(u)}}_{i+1} \ldots, (y^i)_n\right)_k & = & (r^j)_k \\ \left((y^j)_1, \ldots \underbrace{\boldsymbol{L(v)}}_{i+1} \ldots, (y^i)_n\right)_k & = & (s^j)_k \end{array}$$

$2 \leqslant k \leqslant i$. Then

$$\begin{array}{rc} \left((y^j)_1, \ldots \underbrace{\boldsymbol{L(u)}}_{i+1} \ldots, (y^i)_n\right)_k & \underset{k \leqslant i < i+1 \Rightarrow k \neq i+1}{=} \\ (y^j)_k & = \\ (t, L(y_1), \ldots, L(y_{j-1}), L(y_{j+1}), \ldots, L(y_n))_k & \underset{[\text{definition: } 11.286]}{=} \\ L(y_{k-1}) & \underset{k \leqslant i \Rightarrow k-1 \neq i}{=} \\ L(x_{k-1}) & \end{array}$$

and

$$\begin{array}{rc} \left((y^j)_1, \ldots \underbrace{\boldsymbol{L(v)}}_{i+1} \ldots, (y^i)_n\right)_k & \underset{k \leqslant i < i+1 \Rightarrow k \neq i+1}{=} \\ (y^j)_k & = \\ (t, L(y_1), \ldots, L(y_{j-1}), L(y_{j+1}), \ldots, L(y_n))_k & \underset{[\text{definition: } 11.286]}{=} \\ L(y_{k-1}) & \underset{k \leqslant i \Rightarrow k-1 \neq i}{=} \\ L(x_{k-1}) & \end{array}$$

and

$$\begin{array}{rcl} (r^j)_k & = & (t, L(r_1), \ldots, L(r_{j-1}), L(r_{j+1}), \ldots, L(r_n))_k \\ & \underset{[\text{definition: } 11.286]}{=} & L(r_{k-1}) \\ & \underset{k \leqslant i \Rightarrow k-1 \neq i}{=} & L(x_{k-1}) \\ (s^j)_k & = & (t, L(s_1), \ldots, L(s_{j-1}), L(s_{j+1}), \ldots, L(s_n))_k \\ & \underset{[\text{definition: } 11.286]}{=} & L(s_{k-1}) \\ & \underset{k \leqslant i \Rightarrow k-1 \neq i}{=} & L(x_{k-1}) \end{array}$$

Hence we have

$$\begin{array}{rcl} \left((y^j)_1, \ldots \underbrace{\boldsymbol{L(u)}}_{i+1} \ldots, (y^i)_n\right)_k & = & (r^j)_k \\ \left((y^j)_1, \ldots \underbrace{\boldsymbol{L(v)}}_{i+1} \ldots, (y^i)_n\right)_k & = & (s^j)_k \end{array}$$

$\boldsymbol{k = i+1.}$ Then

$$\left((y^j)_1,\ldots\underbrace{\boldsymbol{L(u)}}_{i+1}\ldots,(y^i)_n\right)_k \underset{k=i+1}{=} L(u)$$
$$\left((y^j)_1,\ldots\underbrace{\boldsymbol{L(v)}}_{i+1}\ldots,(y^i)_n\right)_k \underset{k=i+1}{=} L(v)$$

and

$$\begin{array}{rcl} (r^j)_k & = & (t, L(r_1), \ldots, L(r_{j-1}), L(r_{j+1}), \ldots, L(r_n))_k \\ & \underset{[\text{definition: } 11.286] \wedge i<j \Rightarrow 2 \leqslant i+1=k \leqslant j}{=} & L(r_{k-1}) \\ & \underset{k=i+1}{=} & L(r_i) \\ & = & L(u) \\ (s^j)_k & = & (t, L(s_1), \ldots, L(s_{j-1}), L(s_{j+1}), \ldots, L(s_n))_k \\ & \underset{[\text{definition: } 11.286] \wedge i<j \Rightarrow 2 \leqslant i+1=k \leqslant j}{=} & L(s_{k-1}) \\ & \underset{k=i+1}{=} & L(s_i) \\ & = & L(v) \end{array}$$

Hence we have

$$\begin{array}{rcl} ((y^j)_1, \ldots, (y^j)_i, \boldsymbol{L(u)}, (y^j)_{i+2}, \ldots, (y^i)_n)_k & = & (r^j)_k \\ ((y^j)_1, \ldots, (y^j)_i, \boldsymbol{L(v)}, (y^j)_{i+2}, \ldots, (y^i)_n)_k & = & (s^j)_k \end{array}$$

$\boldsymbol{i+1 < k \leqslant j.}$ Then

$$\begin{array}{rc} ((y^j)_1, \ldots, (y^j)_i, \boldsymbol{L(u)}, (y^j)_{i+2}, \ldots, (y^i)_n)_k & \underset{i+1<k \Rightarrow k \neq i+1}{=} \\ (y^j)_k & = \\ (t, L(y_1), \ldots, L(y_{j-1}), L(y_{j+1}), \ldots, L(y_n))_k & \underset{[\text{definition: } 11.286]}{=} \\ L(y_{k-1}) & \underset{i+1<k \Rightarrow i \neq k-1}{=} \\ L(x_{k-1}) & \end{array}$$

and

$$\begin{array}{rc} \left((y^j)_1,\ldots\underbrace{\boldsymbol{L(v)}}_{i+1}\ldots,(y^i)_n\right)_k & \underset{i+1<k \Rightarrow k \neq i+1}{=} \\ (y^j)_k & = \\ (t, L(y_1), \ldots, L(y_{j-1}), L(y_{j+1}), \ldots, L(y_n))_k & \underset{[\text{definition: } 11.286]}{=} \\ L(y_{k-1}) & \underset{i+1<k \Rightarrow i \neq k-1}{=} \\ L(x_{k-1}) & \end{array}$$

and

$$\begin{array}{rcl} (r^j)_k & = & (t, L(r_1), \ldots, L(r_{j-1}), L(r_{j+1}), \ldots, L(r_n))_k \\ & \underset{[\text{definition: } 11.286]}{=} & L(r_{k-1}) \\ & \underset{i+1<k \Rightarrow i \neq k-1}{=} & L(x_{k-1}) \\ (s^j)_k & = & (t, L(s_1), \ldots, L(s_{j-1}), L(s_{j+1}), \ldots, L(s_n))_k \\ & \underset{[\text{definition: } 11.286]}{=} & L(s_{k-1}) \\ & \underset{i+1<k \Rightarrow i \neq k-1}{=} & L(x_{k-1}) \end{array}$$

Hence we have

$$\left((y^j)_1, \ldots \underbrace{\boldsymbol{L(u)}}_{i+1} \ldots, (y^i)_n\right)_k = (r^j)_k$$
$$\left((y^j)_1, \ldots \underbrace{\boldsymbol{L(v)}}_{i+1} \ldots, (y^i)_n\right)_k = (s^j)_k$$

$\boldsymbol{j+1 \leqslant k}$**.** Then

$$\begin{aligned}\left((y^j)_1, \ldots \underbrace{\boldsymbol{L(u)}}_{i+1} \ldots, (y^i)_n\right)_k &\underset{i<j<j+1\leqslant k \Rightarrow i+1\neq k}{=} \\ (y^j)_k &= \\ (t, L(y_1), \ldots, L(y_{j-1}), L(y_{j+1}), \ldots, L(y_n))_k &\underset{[\text{definition: } 11.286]}{=} \\ L(y_k) &\underset{i<j<j+1\leqslant k \Rightarrow i\neq k}{=} \\ L(x_k)\end{aligned}$$

and

$$\begin{aligned}\left((y^j)_1, \ldots \underbrace{\boldsymbol{L(v)}}_{i+1} \ldots, (y^i)_n\right)_k &\underset{i<j<j+1\leqslant k \Rightarrow i+1\neq k}{=} \\ (y^j)_k &= \\ (t, L(y_1), \ldots, L(y_{j-1}), L(y_{j+1}), \ldots, L(y_n))_k &\underset{[\text{definition: } 11.286]}{=} \\ L(y_k) &\underset{i<j<j+1\leqslant k \Rightarrow i\neq k}{=} \\ L(x_k)\end{aligned}$$

and

$$\begin{aligned}(r^j)_k &= (t, L(r_1), \ldots, L(r_{j-1}), L(r_{j+1}), \ldots, L(r_n))_k \\ &\underset{[\text{definition: } 11.286]}{=} L(r_k) \\ &\underset{i<j<j+1\leqslant k \Rightarrow i\neq k}{=} L(x_k) \\ (s^j)_k &= (t, L(s_1), \ldots, L(s_{j-1}), L(s_{j+1}), \ldots, L(s_n))_k \\ &\underset{[\text{definition: } 11.286]}{=} L(s_k) \\ &\underset{i<j<j+1\leqslant k \Rightarrow i\neq k}{=} L(x_k)\end{aligned}$$

Hence we have

$$\left((y^j)_1, \ldots \underbrace{\boldsymbol{L(u)}}_{i+1} \ldots, (y^i)_n\right)_k = (r^j)_k$$
$$\left((y^j)_1, \ldots \underbrace{\boldsymbol{L(v)}}_{i+1} \ldots, (y^i)_n\right)_k = (s^j)_k$$

So we have proved that $\forall k \in \{1, \ldots, n\}$

$$\left((y^j)_1, \ldots \underbrace{L(u)}_{i+1} \ldots, (y^i)_n\right)_k = (r^j)_k$$
$$\left((y^j)_1, \ldots \underbrace{L(v)}_{i+1} \ldots, (y^i)_n\right)_k = (s^j)_k$$

or

$$\begin{pmatrix} (y^j)_1,\ldots \underbrace{L(u)}_{i+1}\ldots,(y^i)_n \end{pmatrix} = r^j$$
$$\begin{pmatrix} (y^j)_1,\ldots \underbrace{L(v)}_{i+1}\ldots,(y^i)_n \end{pmatrix} = s^j$$

which combined with [eq: 11.209] proves that

$$\Delta(y^j) = \Delta(r^j) + \alpha \cdot \Delta(s^j)$$

So that

$$\begin{array}{rcl} A_j(t,y) & = & \Delta(y^j) \cdot y_j \\ & = & (\Delta(r^j) + \alpha \cdot \Delta(s^j)) \cdot y_j \\ & = & \Delta(r^j) \cdot y_j + \alpha \cdot \Delta(s^j) \cdot y_j \\ & \underset{j<i \Rightarrow i \neq j}{=} & \Delta(r^j) \cdot r_j + \alpha \cdot \Delta(s^j) \cdot s_j \\ & = & A_j(t,r) + \alpha \cdot A_j(t,s) \end{array}$$

or summarized

$$A_j(t,y) = A_j(t,r) + \alpha \cdot A_j(t,s) \tag{11.210}$$

By [eqs: 11.208, 11.208 and11.210] we have that

$$\forall j \in \{1,\ldots,n\}\ A_j(t,y) = A_j(t,r) + \alpha \cdot A_j(t,s)$$

Next

$$\begin{array}{rcl} \overline{\Delta L}(y)(t) & \underset{[\text{theorem: } 11.202]}{=} & \sum_{j\in\{1,\ldots,n\}} (-1)^{j-1} \cdot A_j(t,y) \\ & \underset{[\text{theorem: } 11.210]}{=} & \sum_{j\in\{1,\ldots,n\}} (-1)^{j-1} \cdot (A_j(t,r) + \alpha \cdot A_j(t,s)) \\ & \underset{[\text{theorem: } 11.38]}{=} & \sum_{j\in\{1,\ldots,n\}} (-1)^{j-1} \cdot A_j(t,r) + \alpha \cdot \sum_{j\in\{1,\ldots,n\}} (-1)^{j-1} \cdot A_j(t,s) \\ & = & \overline{\Delta L}(r)(t) + \alpha \cdot \overline{\Delta L}(s)(t) \end{array}$$

proving, as $t \in X$ was chosen arbitrary, that $\overline{\Delta L}(y) = \overline{\Delta L}(r) + \alpha \cdot \overline{\Delta L}(s)$ or

$$\overline{\Delta L}(x_1,\ldots,x_{i-1},u+\alpha\cdot v,x_{i+1},\ldots,x_n) = \overline{\Delta L}(x_1,\ldots,x_{i-1},u,x_{i+1},\ldots,x_n) + \alpha \cdot \overline{\Delta L}(x_1,\ldots,x_{i-1},v,x_{i+1},\ldots,x_n)$$

This finally proves that

$$\overline{\Delta L} \in \operatorname{Hom}^n(X;\operatorname{Hom}(X,X))$$

Next we have to prove that $\overline{\Delta L}$ is skew-symmetric. Let $t \in X$ and $x = (x_1,\ldots,x_n) \in X^n$ such that $\exists (k,l) \in \{1,\ldots,n\}$ satisfying $k \neq l$ and $x_k = x_l$. We may always assume that $k < l$ [otherwise exchange k and l]. Consider now the cases for $i, j \in \{1,\ldots,n\}$

$j \neq k \wedge j \neq l$. Then we have that following sub cases to consider for j:

$j < k$. Then we have $j+1 \leqslant k < l$ hence

$$\begin{array}{rcl} (t,L(x_1),\ldots,L(x_{j-1}),L(x_{j+1}),\ldots,L(x_n))_k & \underset{[\text{definition: } 11.286] \wedge j+1 \leqslant k}{=} & \\ L(x_k) & \underset{x_k = x_l}{=} & \\ L(x_l) & \underset{[\text{definition: } 11.286] \wedge j+1 < l}{=} & \\ (t,L(x_1),\ldots,L(x_{j-1}),L(x_{j+1}),\ldots,L(x_n))_l & & \end{array}$$

which as $k \neq l$ proves by [theorem: 11.275] that

$$\Delta(t,L(x_1),\ldots,L(x_{j-1}),L(x_{j+1}),\ldots,L(x_n)) = 0$$

$k<j<l$. Then we have $1<k+1\leqslant j$ and $j+1\leqslant l$ so that

$$\begin{array}{rcl}(t,L(x_1),\ldots,L(x_{j-1}),L(x_{j+1}),\ldots,L(x_n))_{k+1} & \underset{[\text{definition: } 11.286]\wedge 2\leqslant k+1\leqslant j}{=} & \\ L(x_{(k+1)-1}) & = & \\ L(x_k) & \underset{x_k=x_l}{=} & \\ L(x_l) & \underset{[\text{definition: } 11.286]\wedge j+1\leqslant l}{=} & \\ (t,L(x_1),\ldots,L(x_{j-1}),L(x_{j+1}),\ldots,L(x_n))_l & & \end{array}$$

which as $k<j<l\Rightarrow k+1\leqslant j<l\Rightarrow k+1\neq l$ proves by [theorem: 11.275] that

$$\Delta(t,L(x_1),\ldots,L(x_{j-1}),L(x_{j+1}),\ldots,L(x_n))=0$$

$l<j$. Then we have $1<k+1\leqslant j$ and $1<l+1\leqslant j$ so that

$$\begin{array}{rcl}(t,L(x_1),\ldots,L(x_{j-1}),L(x_{j+1}),\ldots,L(x_n))_{k+1} & \underset{[\text{definition: } 11.286]\wedge k+1\leqslant j}{=} & \\ L(x_{(k+1)-1}) & = & \\ L(x_k) & \underset{x_k=x_l}{=} & \\ L(x_l) & = & \\ L(x_{(l+1)-1}) & \underset{[\text{definition: } 11.286]2\leqslant l+1<j}{=} & \\ (t,L(x_1),\ldots,L(x_{j-1}),L(x_{j+1}),\ldots,L(x_n))_{k+1} & & \end{array}$$

which as $k\neq l\Rightarrow k+1\neq l+1$ proves by [theorem: 11.275] that

$$\Delta(t,L(x_1),\ldots,L(x_{j-1}),L(x_{j+1}),\ldots,L(x_n))=0$$

So in all cases we have

$$\Delta(t,L(x_1),\ldots,L(x_{j-1}),L(x_{j+1}),\ldots,L(x_n))=0 \tag{11.211}$$

$j=k\vee j=l$. For $k<l$ we have by [lemma: 11.302] and $L(x_k)=L(x_l)$ that

$$\begin{array}{rl}(t,L(x_1),\ldots,L(x_{k-1}),L(x_{k+1}),\ldots,L(x_n))\circ\left(k\underset{n}{\rightsquigarrow}l+1\right) & = \\ (t,L(x_1),\ldots,L(x_{l-1}),L(x_{l+1}),\ldots,L(x_n)) & \end{array}$$

so that

$$\begin{array}{rcl}\Delta(t,L(x_1),\ldots,L(x_{l-1}),L(x_{l+1}),\ldots,L(x_n)) & = & \\ \Delta\big((t,L(x_1),\ldots,L(x_{k-1}),L(x_{k+1}),\ldots,L(x_n))\circ\left(k\underset{n}{\rightsquigarrow}l+1\right)\big) & = & \\ \big(\left(k\underset{n}{\rightsquigarrow}l+1\right)\Delta\big)(t,L(x_1),\ldots,L(x_{k-1}),L(x_{k+1}),\ldots,L(x_n)) & = & \\ \operatorname{sign}\big(\left(k\underset{n}{\rightsquigarrow}l+1\right)\big)\cdot\Delta(t,L(x_1),\ldots,L(x_{k-1}),L(x_{k+1}),\ldots,L(x_n)) & \underset{[\text{theorem: } 11.239]}{=} & \\ (-1)^{l+1-k}\cdot\Delta(t,L(x_1),\ldots,L(x_{k-1}),L(x_{k+1}),\ldots,L(x_n)) & & \end{array}$$

So by [proposition: 11.265] we have that

$$\begin{array}{rl}\Delta(t,L(x_1),\ldots,L(x_{k-1}),L(x_{k+1}),\ldots,L(x_n)) & = \\ (-1)^{l+1-k}\cdot\Delta(t,L(x_1),\ldots,L(x_{l-1}),L(x_{l+1}),\ldots,L(x_n)) & \end{array} \tag{11.212}$$

Now

$$\begin{array}{rcl}\overline{\Delta L}(x_1,\ldots,x_n)(t) & = & \\ \sum_{j\in\{1,\ldots,n\}}(-1)^j\cdot\Delta(t,L(x_1),\ldots,L(x_{j-1}),L(x_{j+1}),\ldots,L(x_n))\cdot x_j & \underset{[\text{theorem: } 11.43]}{=} & \\ A+B+C & & \end{array} \tag{11.213}$$

where

$$\begin{array}{rcl}
A & = & \sum_{j\in\{1,\ldots,n\}\setminus\{k.l\}} (-1)^{j-1}\cdot\Delta(t,L(x_1),\ldots,L(x_{j-1}),L(x_{j+1}),\ldots,L(x_n))\cdot x_j \\
 & \underset{[\text{theorem: } 11.211]}{=} & \sum_{j\in\{1,\ldots,n\}\setminus\{k.l\}} (-1)^{j-1}\cdot 0\cdot x_j \\
 & = & 0 \\
B & = & \sum_{j\in\{k\}} (-1)^{j-1}\cdot\Delta(t,L(x_1),\ldots,L(x_{j-1}),L(x_{j+1}),\ldots,L(x_n))\cdot x_j \\
 & & (-1)^{k-1}\cdot\Delta(t,L(x_1),\ldots,L(x_{k-1}),L(x_{k+1}),\ldots,L(x_n))\cdot x_k \\
 & \underset{[\text{eq: } 11.212]}{=} & (-1)^{k-1}\cdot(-1)^{l+1-k}\cdot\Delta(t,L(x_1),\ldots,L(x_{l-1}),L(x_{l+1}),\ldots,L(x_n))\cdot x_k \\
 & = & (-1)^{l}\cdot\Delta(t,L(x_1),\ldots,L(x_{l-1}),L(x_{l+1}),\ldots,L(x_n))\cdot x_k \\
 & = & (-1)\cdot(-1)^{l-1}\cdot\Delta(t,L(x_1),\ldots,L(x_{l-1}),L(x_{l+1}),\ldots,L(x_n))\cdot x_k \\
 & \underset{x_k=x_l}{=} & (-1)\cdot(-1)^{l-1}\cdot\Delta(t,L(x_1),\ldots,L(x_{l-1}),L(x_{l+1}),\ldots,L(x_n))\cdot x_l \\
C & = & \sum_{j\in\{l\}} (-1)^{j-1}\cdot\Delta(t,L(x_1),\ldots,L(x_{j-1}),L(x_{j+1}),\ldots,L(x_n))\cdot x_j \\
 & = & (-1)^{l-1}\cdot\Delta(t,L(x_1),\ldots,L(x_{l-1}),L(x_{l+1}),\ldots,L(x_n))\cdot x_l \\
 & = & -B
\end{array}$$

so that using [eq: 11.213] $\overline{\Delta L}(x_1,\ldots,x_n)(t)=0$, which, as $t\in X$ was chosen arbitrary, that

$$\overline{\Delta L}(x_1,\ldots,x_n)=C_0 \text{ [where } C_0 \text{ is the neutral element in } \mathrm{Hom}(X,X)\text{]}$$

$$\overline{\Delta L} \text{ is skew-symmetric}$$

□

The above theorem allows us to define the adjoint of a linear map.

Definition 11.304. *Let $n\in\mathbb{N}$, X a vector space over a field F with characteristic zero such that $\dim(X)=n$, $L\in\mathrm{Hom}(X,X)$ a linear transformation. then the function **adjoint** is defined by*

$$\mathrm{adjoint}\colon \mathrm{Hom}(X,X)\rightarrow\mathrm{Hom}(X,X)$$

where $\mathrm{adjoint}(L)\in\mathrm{Hom}(X,X)$ is such that for every non trivial determinant function Δ we have $\forall x\in X^n$ that

$$\overline{\Delta L}(x)=\Delta(x)\cdot\mathrm{adjoint}(L)$$

Proof. Of course w must prove that given a $L\in\mathrm{Hom}(X,X)$ there exist a $y\in\mathrm{Hom}(X,X)$ such that $\overline{\Delta L}(x)=\Delta(x)\cdot\mathrm{adjoint}(L)$ and that y is independent of the chosen Δ.

existence. Let $L\in\mathrm{Hom}(X,X)$ then by the previous theorem [theorem: 11.303]

$$\overline{\Delta L}\in\mathrm{Hom}^n(X;\mathrm{Hom}(X,X))$$

is a skew-symmetric n-linear mapping. Using [theorem: 11.284] to get a $y\in\mathrm{Hom}(X,X)$ such that

$$\forall x\in X^n \text{ we have } \overline{\Delta L}(x)=\Delta(x)\cdot y$$

uniqueness. Let Δ' another non trivial determinant function and $y'\in\mathrm{Hom}(X,X)$ such that

$$\forall x\in X^n \text{ we have } \overline{\Delta' L}(x)=\Delta'(x)\cdot y'$$

Using [theorem: 11.285] there exists a $\lambda\in F$ such that

$$\Delta'=\lambda\cdot\Delta$$

Then we have $\forall x \in X^n, t \in X$ that

$$\begin{array}{rcl}(\Delta'(x)\cdot y')(t) & = & (\overline{\Delta' L}(x))(t)\\ & = & \sum_{j\in\{1,\ldots,n\}} (-1)^{j-1}\cdot \Delta'(t, L(x_1),\ldots,L(x_{j-1}),L(x_{j+1}),\ldots,L(x_n))\cdot x_j\\ & = & \sum_{j\in\{1,\ldots,n\}} (-1)^{j-1}\cdot \lambda\cdot\Delta(t, L(x_1),\ldots,L(x_{j-1}),L(x_{j+1}),\ldots,L(x_n))\cdot x_j\\ & = & \lambda\cdot\sum_{j\in\{1,\ldots,n\}} (-1)^{j-1}\cdot \Delta(t, L(x_1),\ldots,L(x_{j-1}),L(x_{j+1}),\ldots,L(x_n))\cdot x_j\\ & = & \lambda\cdot(\overline{\Delta L}(x)(t))\\ & = & (\lambda\cdot(\Delta(x)\cdot y))(t)\\ & = & ((\lambda\cdot\Delta)(x)\cdot y)(t)\\ & = & (\Delta'(x)\cdot y)(t)\end{array}$$

which, as $t \in X$ is arbitrary chosen, proves that

$$\forall x \in X^n \text{ we have } \Delta'(x)\cdot y' = \Delta'(x)\cdot y$$

As Δ' is non trivial there exists a $x \in X$ such that $\Delta'(x) \neq 0$ which combined with the above proves that

$$y = y'$$

□

Example 11.305. If X is a one dimensional space with basis $\{e_1\}$ then if $L \in \mathrm{Hom}(X, X)$ we have

1. $\mathrm{adjoint}(L) = \mathrm{Id}_X$
2. $\mathrm{adjoint}(L) \circ L = \det(L)\cdot \mathrm{Id}_X$
3. $L \circ \mathrm{adjoint}(L) = \det(L)\cdot\mathrm{Id}_X$

Proof. If $(x_1) \in X^1$ then there exists a $\alpha \in F$ such that $x_1 = \alpha\cdot e_1$, further by [example: 11.298] we have that $L(e_1) = \det(L)\cdot e_1$ and if $t \in X$ then there exists a $\beta \in F$ such that $t = \beta \cdot e_1$. Let Δ be a determinant function then

$$\begin{array}{rcl}\overline{\Delta L}(x_1)(t) & = & \sum_{i\in\{1\}} (-1)^{i-1}\cdot\Delta(t)\cdot x_i\\ & = & \Delta(t)\cdot x_1\\ & = & \Delta(\beta\cdot e_1)\cdot(\alpha\cdot e_1)\\ & = & \alpha\cdot\beta\cdot\Delta(e_1)\cdot e_1\\ & = & \Delta(\alpha\cdot e_1)\cdot\beta\cdot e_1\\ & = & \Delta(x_1)\cdot t\\ & = & \Delta(x_1)\cdot\mathrm{Id}_X(t)\\ & = & (\Delta(x_1)\cdot\mathrm{Id}_X)(t)\end{array}$$

so that $\overline{\Delta L}(x_1) = \Delta(x_1)\cdot\mathrm{Id}_X$, hence by definition

$$\mathrm{adjoint}(l) = \mathrm{Id}_X.$$

Next we have

1. $\mathrm{adjoint}(L(x)) = \mathrm{Id}_X(L(x)) = L(\alpha\cdot e_1) = \alpha\cdot L(e_1) = \alpha\cdot\det(L)\cdot e_1 = \det(L)\cdot(\alpha\cdot e_1) = \det(L)\cdot x = \det(L)\cdot\mathrm{Id}_X(x)$ proving that

$$\mathrm{adjoint}(L)\circ L = \det(L)\circ L$$

2. $L(\mathrm{adjoint}(x)) = L(\mathrm{Id}_X(x)) = L(x) = \alpha\cdot L(e_1) = a\cdot\det(L)\cdot e_1 = \det(L)\cdot(\alpha\cdot e_1) = \det(L)\cdot x = \det(L)\cdot\mathrm{Id}_X(x)$ proving that

$$L\circ\mathrm{adjoint}(L) = \det(L)\cdot\mathrm{Id}_X$$

□

Actually the above example points to a more general result as is expressed in the following theorem.

Lemma 11.306. *Let $n \in \mathbb{N}$, X a vector space over a field F of characteristic zero such that $\dim(X) = n$ and $L \in \mathrm{Hom}(X, X)$ then we have*

1. $\mathrm{adjoint}(L)\circ L = \det(L)\cdot\mathrm{Id}_X$

2. $L \circ \mathrm{adjoint}(L) = \det(L) \cdot \mathrm{Id}_X$

Proof. Let $\{e_1,\ldots,e_n\}$ be a basis for X and let Δ be the non trivial determinant function such that $\Delta(e_1,\ldots,e_n)=1$ [see theorem: 11.283].

1. Let $x \in X$ then we have

$$\begin{aligned}
(\mathrm{adjoint}(L)\circ L)(x) &= \\
\mathrm{adjoint}(L)(L(x)) &= \\
(1\cdot \mathrm{adjoint}(L))(L(x)) &= \\
(\Delta(e_1,\ldots,e_n)\cdot \mathrm{adjoint}(L))(L(x)) &\underset{[\text{definition: } 11.304]}{=} \\
\overline{\Delta L}(e_1,\ldots,e_n)(L(x)) &\underset{[\text{definition: } 11.301]}{=} \\
\sum_{i\in\{1,\ldots,n\}} (-1)^{i-1}\cdot \Delta(L(x),L(e_1),\ldots,L(e_{i-1}),L(e_{i+1}),\ldots,L(e_n))\cdot e_i &= \\
\sum_{i\in\{1,\ldots,n\}} (-1)^{i-1}\cdot \Delta_L(x,e_1,\ldots,e_{i-1},e_{i+1},\ldots,e_n)\cdot e_i &\underset{[\text{theor., def: } 11.292, 11.288]}{=} \\
\underline{\Delta_L}(x,(e_1,\ldots,e_n)) &\underset{[\text{theorem: } 11.290]}{=} \\
\Delta_L(e_1,\ldots,e_n)\cdot x &\underset{[\text{theorem: } 11.295]}{=} \\
(\det(L)\cdot\Delta(e_1,\ldots,e_n))\cdot x &= \\
\det(L)\cdot x &= \\
\det(L)\cdot \mathrm{Id}_X(x)
\end{aligned}$$

proving that

$$\mathrm{adjoint}(L)\circ L = \det(L)\cdot \mathrm{Id}_X$$

2. Let $x \in X$ then we have

$$\begin{aligned}
(L\circ \mathrm{adjoint}(L))(x) &= \\
L(\mathrm{adjoint}(L)(x)) &= \\
L(1\cdot \mathrm{adjoint}(L)(x)) &= \\
L(\Delta(e_1,\ldots,e_n)\cdot \mathrm{adjoint}(L)(x)) &\underset{[\text{theorem: } 11.304]}{=} \\
L(\overline{\Delta L}(e_1,\ldots,e_n)(x)) &\underset{[\text{definition: } 11.301]}{=} \\
L\left(\sum_{i\in\{1,\ldots,n\}} (-1)^{i-1}\cdot \Delta(x,L(e_1),\ldots,L(e_{i-1}),L(e_{i+1}),\ldots,L(e_n))\cdot e_i\right) &= \\
\sum_{i\in\{1,\ldots,n\}} (-1)^{i-1}\cdot \Delta(x,L(e_1),\ldots,L(e_{i-1}),L(e_{i+1}),\ldots,L(e_n))\cdot L(e_i) &\underset{[\text{def: } 11.288]}{=} \\
\underline{\Delta}(x,(L(e_1),\ldots,L(e_n))) &\underset{[\text{theorem: } 11.290]}{=} \\
\Delta(L(e_1),\ldots,L(e_n))\cdot x &= \\
\Delta_L(e_1,\ldots,e_n)\cdot x &\underset{[\text{theorem: } 11.295]}{=} \\
(\det(L)\cdot\Delta(e_1,\ldots,e_n))\cdot x &= \\
\det(L)\cdot x &= \\
\det(L)\cdot \mathrm{Id}_X(x)
\end{aligned}$$

proving that

$$L\circ \mathrm{adjoint}(L) = \det(L)\cdot \mathrm{Id}_X$$ □

The above lemma leads to a theorem that allows us to calculate the inverse of a linear mapping.

Theorem 11.307. *Let $n\in\mathbb{N}$, X a vector space over a field F of characteristic zero such that $\dim(X)=n$ and $L\in \mathrm{Hom}(X,X)$ then we have:*

1. *L is a isomorphism* $\Leftrightarrow \det(L)\neq 0$
2. *If L is a isomorphism then* $L^{-1}=(\det(L))^{-1}\cdot \mathrm{adjoint}(L)$

Proof.

1. This is already proved in [theorem: 11.299].
2. Using the previous lemma [lemma: 11.306] we have

$$
\begin{array}{rcl}
L\circ \operatorname{adjoint}(L)=\det(L)\cdot \mathrm{Id}_X & \Rightarrow & \operatorname{adjoint}(L)=L^{-1}\circ(\det(L)\cdot \mathrm{Id}_X)\\
& \underset{\text{[theorem: 11.171]}}{\Rightarrow} & \operatorname{adjoint}(L)=\det(L)\cdot L^{-1}\circ \mathrm{Id}_X\\
& \Rightarrow & \operatorname{adjoint}(L)=\det(L)\cdot L^{-1}\\
& \underset{\det(L)\neq 0}{\Rightarrow} & (\det(L))^{-1}\cdot \operatorname{adjoint}(L)=L^{-1}
\end{array}
$$

□

11.8 Matrices

In finite dimensional vector spaces a linear mapping can be represented by a matrix. Composition of linear mappings becomes then multiplication of matrices. Further the adjoint and determinant of a linear mapping can be translated to the adjoint and determinant of matrices. One disadvantage of this approach is that the matrix representation of a linear mapping is dependent on the chosen basis and that this approach works only in finite dimensional vector spaces. However one big advantage is that all calculations can then be done using essential multiplication and addition in a field (in most cases the real or complex numbers). This is the main reason for the existence of this section.

11.8.1 Definition and properties

Definition 11.308. *Let F be a field, $n,m\in\mathbb{N}$ then a $n\times m$ **matrix** in F is the graph of a mapping from $\{1,\ldots,n\}\times\{1,\ldots,m\}$ to F. The set of all $n\times m$ matrices is called $\mathcal{M}_{n,m}(F)$ or $\mathcal{M}_{n,m}$ if F is clear from the context. Hence*

$$\mathcal{M}_{n,m}(F)=F^{\{1,\ldots,n\}\times\{1,\ldots,m\}}\underset{\text{[definition: 2.30]}}{=}\{M|M:\{1,\ldots,n\}\times\{1,\ldots,m\}\rightarrow F\}$$

*As $1\times n$ matrix is called a **row vector** and a $n\times 1$ matrix is called a* column vector. *Let $M\in\mathcal{M}_{n,m}(F)$ and $i,j\in\{1,\ldots,n\}\times\{1,\ldots,m\}$ then we use the following notation*

$$M(i,j)\underset{notation}{=}M_{i,j}$$

Another notation for $M\in\mathcal{M}_{n,m}(F)$ is

$$M=\begin{pmatrix} M_{1,1} & \ldots & M_{1,m}\\ \vdots & \ddots & \vdots\\ M_{n,1} & \ldots & M_{n,m}\end{pmatrix}$$

For a row vector $R\in\mathcal{M}_{1,n}(F)$ we use the notation

$$R=\begin{pmatrix} R_1 & \ldots & Rn\end{pmatrix}$$

and for a column vector $C\in\mathcal{M}_{n,1}$ we use the notation

$$C=\begin{pmatrix} C_1\\ \vdots\\ C_n\end{pmatrix}$$

*In essence a $n\times m$ matrix can be said to be a collection of n $1\times m$ row vectors or a collection of m $n\times 1$ column vectors. Elements of $\mathcal{M}_{n,n}$ [matrices where the number of column vectors is equal to the number of row vectors] are called **square matrices**.*

Example 11.309. Let $n,m\in\mathbb{N}$, F be a field then $E\in\mathcal{M}_{n,m}(F)$ is defined by

$$\forall(i,j)\in\{1,\ldots,n\}\times\{1,\ldots,m\} \text{ we have } E_{i,j}=\delta_{i,j} \text{ [see definition: 11.156]}$$

this matrix is called the **identity matrix**. So

$$E=\begin{pmatrix} 1 & 0 & \ldots & 0 & 1\\ 0 & 1 & \ddots & 0 & 0\\ \vdots & 0 & \ddots & 0 & 0\\ 0 & 0 & \ldots & 0 & 1\end{pmatrix}$$

Example 11.310. Let $n, m \in \mathbb{N}$, F a field with neutral element 0 then $0 \in \mathcal{M}_{n,m}(F)$ is defined by

$$\forall (i,j) \in \{1,\ldots,n\} \times \{1,\ldots,m\} \text{ we have } 0_{i,j} = 0$$

this matrix is called the null matrix. SO

$$0 = \begin{pmatrix} 0 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & 0 \end{pmatrix}$$

We define now the different arithmetic operations on matrices.

Definition 11.311. *Let $n, m \in \mathbb{N}$ then we define*

1. **(sum)** $+ : \mathcal{M}_{n,m}(F) \times \mathcal{M}_{n,m}(F) \to \mathcal{M}_{n,m}(F)$ *is defined by*

$$A + B \in \mathcal{M}_{n,m}(F) \text{ where } \forall (i,j) \in \{1,\ldots,n\} \times \{1,\ldots,m\}\ (A+B)_{i,j} = A_{i,j} + B_{i,j}$$

2. **(scalar product)** $\cdot : F \times \mathcal{M}_{n,m}(F) \to \mathcal{M}_{n,m}(F)$ *is defined by*

$$\alpha \cdot A \in \mathcal{M}_{n,m}(F) \text{ where } \forall (i,j) \in \{1,\ldots,n\} \times \{1,\ldots,m\}\ (\alpha \cdot A)_{i,j} = \alpha \cdot A_{i,j}$$

3. **(inner product)** *If $k \in \mathbb{N}$ then we define* $\cdot : \mathcal{M}_{n,m}(F) \times \mathcal{M}_{m,k}(F) \to \mathcal{M}_{n,k}(F)$ *by*

$$A \cdot B \in \mathcal{M}_{n,k}(F) \text{ where } (A \cdot B)_{i,j} = \sum_{r \in \{1,\ldots,m\}} A_{i,r} \cdot B_{r,j}$$

4. **(null matrix)** $0 \in \mathcal{M}_{n,m}(F)$ *is defined by* $\forall (i,j) \in \{1,\ldots,n\} \times \{1,\ldots,m\}\ 0_{i,j} = 0$
5. **(identity matrix)** $E \in \mathcal{M}_{n,m}(F)$ *is defined by* $\forall (i,j) \in \{1,\ldots,n\} \times \{1,\ldots,m\}\ E_{i,j} = \delta_{i,j}$
6. **(additive inverse)** *If $M \in \mathcal{M}_{n,m}(F)$ then $-M$ is defined by* $\forall (i,j) \in \{1,\ldots,n\} \times \{1,\ldots,m\}\ (-M)_{i,j} = -M_{i,j}$

Example 11.312. Then we have

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 4 & 2 \end{pmatrix} + \begin{pmatrix} 0 & 2 & 6 \\ 4 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 4 & 9 \\ 8 & 5 & 3 \end{pmatrix}$$

$$10 \cdot \begin{pmatrix} 1 & 2 & 3 \\ 4 & 4 & 2 \end{pmatrix} = \begin{pmatrix} 10 & 20 & 30 \\ 40 & 40 & 20 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 4 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ 0 & 3 \\ 2 & 0 \end{pmatrix} = \begin{pmatrix} 7 & 8 \\ 8 & 20 \end{pmatrix}$$

Theorem 11.313. *Let $n, m \in \mathbb{N}$ and F a field then, $\langle \mathcal{M}_{n,m}(F), +, \cdot \rangle$ is a vector space over the field F [using the operations in [definition: 11.311]]*

Proof. First we prove the group axioms:

associativity. Let $A, B, C \in \mathcal{M}_{n,m}(F)$ then for $(i,j) \in \{1,\ldots,n\} \times \{1,\ldots,m\}$ we have

$$\begin{aligned} ((A+B)+C)_{i,j} &= (A+B)_{i,j} + C_{i,j} \\ &= (A_{i,j} + B_{i,j}) + C_{i,j} \\ &= A_{i,j} + (B_{i,j} + C_{i,j}) \\ &= A_{i,j} + (B+C)_{i.j} \\ &- (A + (B+C))_{i,j} \end{aligned}$$

proving that

$$(A+B)+C = A+(B+C)$$

commutativity. Let $A, B \in \mathcal{M}_{n,m}(F)$ then for $(i,j) \in \{1,\ldots,n\} \times \{1,\ldots,m\}$ we have

$$\begin{aligned}(A+B)_{i,j} &= A_{i,j}+B_{i,j}\\ &= B_{i,j}+A_{i,j}\\ &= (B+A)_{i,j}\end{aligned}$$

proving that

$$A+B=B+A$$

neutral element. Let $A \in \mathcal{M}_{n,m}(F)$ then for $(i,j) \in \{1,\ldots,n\} \times \{1,\ldots,m\}$ we have

$$(A+0)_{i,j}=A_{i,j}+0_{i,j}=A_{i,j}+0=A_{i,j}$$

proving that

$$0+A \underset{\text{commutativity}}{=} A+0=A$$

inverse element. Let $A \in \mathcal{M}_{n,m}(F)$ then for $(i,j) \in \{1,\ldots,n\} \times \{1,\ldots,m\}$ we have

$$(A+(-A))_{i,j}=A_{i,j}+(-A)_{i,j}=A_{i,j}+(-A_{i,j})=0=0_{i,j}$$

proving that

$$(-A)+A \underset{\text{commutativity}}{=} A+(-A)=0$$

Next we check the rest of the vector space axioms:

1. Let $A, B \in \mathcal{M}_{n,m}(F)$, $\alpha \in F$ then for $(i,j) \in \{1,\ldots,n\} \times \{1,\ldots,m\}$ we have

$$\begin{aligned}(\alpha\cdot(A+B))_{i,j} &= \alpha\cdot(A+B)_{i,j}\\ &= \alpha\cdot(A_{i,j}+B_{i,j})\\ &= \alpha\cdot A_{i,j}+\alpha\cdot B_{i,j}\\ &= (\alpha\cdot A)_{i,j}+(\alpha\cdot B)_{i,j}\\ &= (\alpha\cdot A+\alpha\cdot B)_{i,j}\end{aligned}$$

   proving that

$$\alpha\cdot(A+B)=\alpha\cdot A+\alpha\cdot B$$

2. Let $A \in \mathcal{M}_{n,m}(F)$, $\alpha, \beta \in F$ then for $(i,j) \in \{1,\ldots,n\} \times \{1,\ldots,m\}$ we have

$$\begin{aligned}((\alpha+\beta)\cdot A)_{i,j} &= (\alpha+\beta)\cdot A_{i,j}\\ &= \alpha\cdot A_{i,j}+\beta\cdot A_{i.j}\\ &= (\alpha\cdot A)_{i,j}+(\beta\cdot A)_{i,j}\\ &= (\alpha\cdot A+\beta\cdot A)_{i,j}\end{aligned}$$

   proving that

$$(\alpha+\beta)\cdot A=\alpha\cdot A+\beta\cdot A$$

3. Let $A \in \mathcal{M}_{n,m}(F)$, $\alpha, \beta \in F$ then for $(i,j) \in \{1,\ldots,n\} \times \{1,\ldots,m\}$ we have

$$\begin{aligned}((\alpha\cdot\beta)\cdot A)_{i,j} &= (\alpha\cdot\beta)\cdot A_{i,j}\\ &= \alpha\cdot(\beta\cdot A_{i,j})\\ &= \alpha\cdot(\beta\cdot A)_{i,j}\\ &= (\alpha\cdot(\beta\cdot A))_{i,j}\end{aligned}$$

   proving that

$$(\alpha\cdot\beta)\cdot A=\alpha\cdot(\beta\cdot A)$$

4. Let $A \in \mathcal{M}_{n,m}(F)$ then for $(i,j) \in \{1,\ldots,n\} \times \{1,\ldots,m\}$ we have

$$(1\cdot A)_{i,j}=1\cdot A_{i,j}=A_{i,j}=A_{i,j}$$

proving that

$$1 \cdot A = A \qquad \square$$

Theorem 11.314. *Let $n \in \mathbb{N}$ and F a field then $\langle \mathcal{M}_{n,n}(F), \cdot \rangle$ is a semi-group with as neutral element the identity matrix E.*

Proof.

associativity. Let $A, B, C \in \mathcal{M}_{n,m}(F)$ then for $(i, j) \in \{1, \ldots, n\} \times \{1, \ldots, m\}$ we have

$$\begin{aligned}
((A \cdot B) \cdot C)_{i,j} &= \sum_{k \in \{1, \ldots, n\}} (A \cdot B)_{i,k} \cdot C_{k,j} \\
&= \sum_{k \in \{1, \ldots, n\}} \left(\sum_{l \in \{1, \ldots, n\}} A_{i,l} \cdot B_{l,k} \right) \cdot C_{k,j} \\
&\underset{\text{[theorem: 11.75]}}{=} \sum_{k \in \{1, \ldots, n\}} \left(\sum_{l \in \{1, \ldots, n\}} (A_{i,l} \cdot B_{l,k}) \cdot C_{k,j} \right) \\
&\underset{\text{[theorem: 11.41]}}{=} \sum_{l \in \{1, \ldots, n\}} \left(\sum_{k \in \{1, \ldots, n\}} (A_{i,l} \cdot B_{l,k}) \cdot C_{k,j} \right) \\
&= \sum_{l \in \{1, \ldots, n\}} \left(\sum_{k \in \{1, \ldots, n\}} A_{i,l} \cdot (B_{l,k}) \cdot C_{k,j} \right) \\
&\underset{\text{[theorem: 11.75]}}{=} \sum_{l \in \{1, \ldots, n\}} A_{i,l} \cdot \left(\sum_{k \in \{1, \ldots, n\}} (B_{l,k}) \cdot C_{k,j} \right) \\
&= \sum_{l \in \{1, \ldots, n\}} A_{i,l} \cdot (B \cdot C)_{l,j} \\
&= (A \cdot (B \cdot C))_{i,j}
\end{aligned}$$

proving that

$$(A \cdot B) \cdot C = A \cdot (B \cdot C)$$

neutral element. Let $A \in \mathcal{M}_{n,m}(F)$ then for $(i, j) \in \{1, \ldots, n\} \times \{1, \ldots, m\}$ we have

$$\begin{aligned}
(A \cdot E)_{i,j} &= \sum_{k \in \{1, \ldots, n\}} A_{i,k} \cdot E_{k,j} \\
&= \sum_{k \in \{1, \ldots, n\}} A_{i,k} \cdot \delta_{k,j} \\
&= A_{i,j} \\
(E \cdot A)_{i,j} &= \sum_{k \in \{1, \ldots, n\}} E_{i,k} \cdot A_{k,j} \\
&= \sum_{k \in \{1, \ldots, n\}} \delta_{i,k} \cdot A_{k,j} \\
&= A_{i,j}
\end{aligned}$$

proving that

$$A \cdot E = A = E \cdot A \qquad \square$$

Note 11.315. That the inner product is not commutative as the following example shows:

$$\begin{aligned}
\begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} \cdot \begin{pmatrix} 4 & 1 \\ 2 & 1 \end{pmatrix} &= \begin{pmatrix} 10 & 4 \\ 16 & 6 \end{pmatrix} \\
\begin{pmatrix} 4 & 1 \\ 2 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} &= \begin{pmatrix} 6 & 16 \\ 4 & 10 \end{pmatrix}
\end{aligned}$$

Also there exist non zero matrices that do not have a inverse, for example

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

and we have

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & a \\ 0 & c \end{pmatrix} \neq \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

so that $\left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$ has no inverse.

We define now the transpose of a matrix, which is essential the operation that interchange the column vectors and row vectors of a matrix.

Definition 11.316. *Let $n,m\in\mathbb{N}$ and $M\in\mathcal{M}_{n,m}(F)$ then $M^T\in\mathcal{M}_{m,n}(F)$ is defined by*

$$\forall (i,j)\in\{1,\ldots,m\}\times\{1,\ldots,n\}\ (M^T)_{i,j}=M_{j,i}$$

The following definitions allows use to extract the rows and columns from a matrix and introduces the important concept of row (column) rank of a matrix that eventually leads to the important concept of the rank from a matrix.

Definition 11.317. *Let F be a field, $n,m\in\mathbb{N}$ then we define:*

1. $\mathrm{row}\colon\mathcal{M}_{n,m}(F)\times\{1,\ldots,n\}\to F^m$ *where* $\mathrm{row}(M,i)\in F^m$ *is defined by* $(\mathrm{row}(M,i))_j=M_{i,j}\ \forall j\in\{1,\ldots,m\}$. *In other words this function extract the i-the row from M.*
2. $\mathrm{col}\colon\mathcal{M}_{n,m}(F)\times\{1,\ldots,m\}\to F^n$ *where* $\mathrm{col}(M,i)\in F^n$ *is defined by* $(\mathrm{col}(M,i))_j=M_{j,i}\ \forall j\in\{1,\ldots,n\}$. *In other words this function extracts the i-the column from M.*
3. *Let $M\in\mathcal{M}_{n,m}(F)$ then the set of rows of M is noted as* $\mathrm{rows}(M)$, *hence*

$$\mathrm{rows}(M)=\{\mathrm{row}(M,i)|i\in\{1,\ldots,n\}\}\subseteq F^m$$

4. *Let $M\in\mathcal{M}_{n,m}(F)$ then the set of columns of M is noted as* $\mathrm{cols}(M)$, *hence:*

$$\mathrm{cols}(M)=\{\mathrm{col}(M,i)|i\in\{1,\ldots,m\}\}\subseteq F^n$$

As F^n, F^m are vector spaces over F with $\dim(F^n)=n$, $\dim(F^m)=m$ [see example: 11.159], and, given $M\in\mathcal{M}_{n,m}(F)$, $\mathrm{span}(\mathrm{rows}(M))$ and $\mathrm{span}(\mathrm{cols}(M))$ are sub spaces of F^m and F^n, we have that $\dim(\mathrm{span}(\mathrm{rows}(M)))$ and $\dim(\mathrm{span}(\mathrm{cols}(M)))$ are finite with $\dim(\mathrm{span}(\mathrm{rows}(M)))\leqslant m$ and $\dim(\mathrm{span}(\mathrm{cols}(M)))\leqslant n$ [see theorems: 11.105, 11.152]. This allows us to define:

1. *If $M\in\mathcal{M}_{n,m}(F)$ then the row rank of M noted as* $\mathrm{rrank}(M)$ *is defined by*

$$\mathrm{rrank}(M)=\dim(\mathrm{span}(\mathrm{rows}(M)))\leqslant m$$

2. *If $M\in\mathcal{M}_{n,m}(F)$ then the column rank of M noted as* $\mathrm{crank}(M)$ *is defined by*

$$\mathrm{crank}(M)=\dim(\mathrm{span}(\mathrm{cols}(M)))\leqslant n$$

It turns out that the column rank and row rank of a matrix are equal, which leads to the definition of the rank of a matrix.

Theorem 11.318. *Let F a field, $n,m\subset\mathbb{N}$ and $M\in\mathcal{M}_{n,m}(F)$ then*

$$\mathrm{rrank}(M)=\mathrm{crank}(M)$$

Proof. Let

$$r=\mathrm{rrank}(M)=\dim(\mathrm{span}(\mathrm{rows}(M)))\leqslant m \text{ and } c=\mathrm{crank}(M)=\dim(\mathrm{span}(\mathrm{cols}(M)))\leqslant n$$

A $r=\dim(\mathrm{span}(\mathrm{rows}(M)))$ there exist a basis $E=\{e_1,\ldots,e_r\}\subseteq\mathrm{span}(\mathrm{rows}(M))\subseteq F^n$. Let $i\in\{1,\ldots,n\}$ then there exist a $\{\lambda_{i,j}\}_{j\in\{1,\ldots,r\}}\subseteq F$ such that

$$\mathrm{row}(M,i)=\sum_{j\in\{1,\ldots,r\}}\lambda_{i,j}\cdot e_j \tag{11.214}$$

now let $k\in\{1,\ldots,m\}$ then

$$\begin{aligned}\operatorname{col}(M,k)_i &= M_{i,k}\\ &= \operatorname{row}(M,i)_k\\ &\underset{[\text{eq: } 11.214]}{=} \left(\sum_{j\in\{1,\ldots,r\}} \lambda_{i,j}\cdot e_j\right)_k\\ &\underset{[\text{theorem: } 11.47]}{=} \sum_{j\in\{1,\ldots,r\}} \lambda_{i,j}\cdot (e_j)_k\\ &= \sum_{j\in\{1,\ldots,r\}} (e_j)_k\cdot\lambda_{i,j}\end{aligned}\tag{11.215}$$

Define now $\{f_i\}_{\in\{1,\ldots,r\}}\subseteq F^n$ by $(f_i)_l=\lambda_{l,i}$ $\forall l\in\{1,\ldots,n\}$ then by [eq: 11.215] we have

$$\operatorname{col}(M,k)_i=\sum_{j\in\{1,\ldots,r\}} (e_j)_k\cdot (f_j)_i$$

so that

$$\operatorname{col}(M,k)=\sum_{j\in\{1,\ldots,r\}} (e_j)_k\cdot f_j$$

proving using [theorem: 11.102] that $\operatorname{col}(M,k)\in\operatorname{span}(\{f_i|i\in\{1,\ldots,r\}\})$. Hence

$$\operatorname{cols}(M)\subseteq\operatorname{span}(\{f_i|i\in\{1,\ldots,r\}\})$$

so that by [theorems: 11.104, 11.107]

$$\operatorname{span}(\operatorname{cols}(M))\subseteq\operatorname{span}(\{f_i|i\in\{1,\ldots,r\}\})$$

Using [theorem: 11.152]

$$\dim(\operatorname{span}(\operatorname{cols}(M)))\leqslant\dim(\operatorname{span}(\{f_i|i\in\{1,\ldots,r\}\}))$$

As $\operatorname{span}(\{f_i|i\in\{1,\ldots,r\}\})$ is a vector space we have by [theorem: 11.148] that

$$\dim(\operatorname{span}(\{f_i|i\in\{1,\ldots,r\}\}))\leqslant r=\operatorname{rrank}(M)$$

giving

$$\operatorname{crank}(M)=\dim(\operatorname{span}(\operatorname{cols}(M)))\leqslant\operatorname{rrank}(M)\tag{11.216}$$

Now we use the same reasoning to prove the opposite equation. As $c=\dim(\operatorname{span}(\operatorname{cols}(M)))$ there exist a basis $G=\{g_1,\ldots,g_c\}$ of $\operatorname{span}(\operatorname{cols}(M))$. Let $i\in\{1,\ldots,m\}$ then there exists a $\{\alpha_{i,j}\}_{j\in\{1,\ldots,c\}}\subseteq F$ such that

$$\operatorname{col}(M,i)=\sum_{j\in\{1,\ldots,c\}}\alpha_{i,j}\cdot g_j\tag{11.217}$$

Now let $k\in\{1,\ldots,n\}$ then we have

$$\begin{aligned}(\operatorname{row}(M,k))_i &= M_{k,i}\\ &= (\operatorname{col}(M,i))_k\\ &\underset{[\text{eq: } 11.217]}{=} \left(\sum_{j\in\{1,\ldots,c\}} \alpha_{i,j}\cdot g_j\right)_k\\ &\underset{[\text{theorem: } 11.47]}{=} \sum_{j\in\{1,\ldots,c\}} \alpha_{i,j}\cdot (g_j)_k\\ &= \sum_{j\in\{1,\ldots,c\}} (g_j)_k\cdot\alpha_{i,j}\end{aligned}\tag{11.218}$$

Define now $\{h_i\}_{i\in\{1,\ldots,c\}}\subseteq F^m$ by $(h_i)_l=\alpha_{l,i}$ $\forall l\in\{1,\ldots,m\}$ then by [eq: 11.218] we have

$$(\operatorname{row}(M,k))_i=\sum_{j\in\{1,\ldots,c\}}(g_j)_k\cdot(h_j)_i$$

so that

$$\operatorname{row}(M,k)=\sum_{j\in\{1,\ldots,c\}}(g_j)_k\cdot h_j$$

proving, using [theorem: 11.102], that $\operatorname{row}(M,k)\in\operatorname{span}(\{h_i|i\in\{1,\ldots,c\}\})$. Hence

$$\operatorname{rows}(M)\subseteq\operatorname{span}(\{h_i|i\in\{1,\ldots,c\}\})$$

so that by that by [theorems: 11.104, 11.107]

$$\operatorname{span}(\operatorname{rows}(M))\subseteq\operatorname{span}(\{h_i|i\in\{1,\ldots,c\}\})$$

Using [theorem: 11.152] we have:

$$\dim(\operatorname{span}(\operatorname{rows}(M)))\leqslant\dim(\operatorname{span}(\{h_i|i\in\{1,\ldots,c\}\}))$$

By [theorem: 11.148] it follows that

$$\dim(\operatorname{span}(\{h_i|i\in\{1,\ldots,c\}\}))\leqslant c.$$

So we have

$$\operatorname{rrank}(M)=\dim(\operatorname{span}(\operatorname{rows}(M)))\leqslant c=\operatorname{crank}(M)$$

combining the above with [eq: 11.216] proves

$$\operatorname{rrank}(M)=\operatorname{crank}(M)$$

□

The above theorem let us define the rank of a matrix.

Definition 11.319. *Let $n,m\in\mathbb{N}$, F a field and $M\in\mathcal{M}_{n,m}(F)$ then*

$$\operatorname{rank}(M)=\operatorname{rrank}(M)\underset{\text{[theorem: 11.318]}}{=}\operatorname{crank}(M)$$

Note: As by [definition: 11.317] $\operatorname{rrank}(M)\leqslant m$ *and* $\operatorname{crank}(M)\leqslant n$ *it follows that*

$$\operatorname{rank}(M)\leqslant\min(n,m)$$

11.8.2 Matrices and linear mappings

First note that a linear mapping between finite dimensional spaces is uniquely determined by its values on the basis vectors of the domain.

Theorem 11.320. *Let $n,m\in\mathbb{N}$, X,Y finite dimensional vector spaces over a field F with bases $\{e_1,\ldots,e_n\}\subseteq X,\{f_1,\ldots,f_m\}\subseteq F$ then if $L_1,L_2\in\operatorname{Hom}(X,Y)$ with $\forall i\in\{1,\ldots,n\}$ $L_1(e_i)=L_2(e_i)$ then $L_1=L_2$.*

Proof. Let $x\in X$ then there exists a unique family $\{x_i\}_{i\in\{1,\ldots,n\}}\subseteq F$ such that

$$x=\sum_{i\in\{1,\ldots,n\}}x_i$$

Then

$$\begin{aligned}L_1(x) &= L_1\left(\sum_{i\in\{1,\ldots,n\}}x_i\right)\\ &= \sum_{i\in\{1,\ldots,n\}}x_i\cdot L_1(e_i)\\ &= \sum_{i\in\{1,\ldots,n\}}x_i\cdot L_2(e_i)\\ &= L_2\left(\sum_{i\in\{1,\ldots,n\}}x\right)\\ &= L_2(x)\end{aligned}$$

proving that $L_1=L_2$. □

We show now how to associate a matrix with a linear mapping between finite dimensional spaces.

Definition 11.321. *Let $n,m\in\mathbb{N}$, X,Y finite dimensional vector spaces over a field F with basis $\{e_1,\ldots,e_n\}\subseteq X,\{f_1,\ldots,f_n\}\subseteq Y$ and $L\in\operatorname{Hom}(X,Y)$. Then we define $\mathcal{M}(L;E,F)\in\mathcal{M}_{m,n}(F)$ to be the **unique** matrix such that*

$$\forall i\in\{1,\ldots,n\}\ L(e_i)=\sum_{j\in\{1,\ldots,m\}}\mathcal{M}(L;E,F)_{j,i}\cdot f_j$$

[theorem: 11.137 ensures uniqueness and existence]. $\mathcal{M}(L;E,F)$ is called the matrix of L for the basis E and F.

Note 11.322. The matrix $M(L;E,F)$ depends clearly not only on L but also on the basis used. If it is clear from the context which basis are used then we use the notation $\mathcal{M}(L)$.

We show now how the matrix of a linear mapping can be used to calculate the result of applying the linear mapping.

Theorem 11.323. *Let $n,m\in\mathbb{N}$, X,Y finite dimensional vector spaces over a field F with bases $\{e_1,\ldots,e_n\}\subseteq X,\{f_1,\ldots,f_m\}\subseteq F$ and $L\in\operatorname{Hom}(X,Y)$ then we have:*

1. *If $x\in X$ then by [theorem: 11.137] there exists **unique** $\{x_i\}_{i\in\{1,\ldots,n\}}\subseteq F$ and $\{L(x)\}_{i\in\{1,\ldots,m\}}$ such that $x=\sum_{i\in[1,\ldots,n]} x_i\cdot e_i$ and $L(x)=\sum_{i\in\{1,\ldots,m\}} L(x)_i\cdot f_i$. Then we have*

$$\forall j\in\{1,\ldots,m\}\ L(x)_j=\sum_{i\in\{1,\ldots,n\}}\mathcal{M}(L;E,F)_{j,i}\cdot x_i$$

   so if we define

$$X=\begin{pmatrix} x_1\\ \vdots\\ x_n\end{pmatrix}\in\mathcal{M}_{n,1}(F) \text{ the column vector uniquely defined by the components of } x$$

   and

$$Y=\begin{pmatrix} L(x)_1 & \ldots & L(x)_m\end{pmatrix} \text{ the row vector uniquely defined by the components of } L(x)$$

   then we have

$$Y=\mathcal{M}(L;E,F)\cdot X$$

   reducing applying a linear map to multiplication of matrices, involving only operations in F.

2. $\operatorname{rank}(L)=\operatorname{rank}(\mathcal{M}(L;E,F))$

Proof.

1. We have

$$\begin{aligned}
\sum_{j\in\{1,\ldots,m\}} L(x)_j\cdot f_j &= L(x)\\
&= L\left(\sum_{i\in\{1,\ldots,n\}} x_i\cdot e_i\right)\\
&\underset{[\text{theorem: } 11.180]}{=} \sum_{i\in\{1,\ldots,n\}} x_i\cdot L(e_i)\\
&\underset{\text{def}}{=} \sum_{i\in\{1,\ldots,n\}} x_i\cdot\left(\sum_{j\in\{1,\ldots,m\}}\mathcal{M}(L;E,F)_{j,i}\cdot f_j\right)\\
&\underset{[\text{theorem: } 11.75}{=} \sum_{i\in\{1,\ldots,n\}}\left(\sum_{j\in\{1,\ldots,m\}} x_i\cdot(\mathcal{M}(L;E,F)_{j,i}\cdot f_j)\right)\\
&= \sum_{i\in\{1,\ldots,n\}}\left(\sum_{j\in\{1,\ldots,m\}} (x_i\cdot\mathcal{M}(L;E,F)_{j,i})\cdot f_j\right)\\
&\underset{[\text{theorem: } 11.41]}{=} \sum_{j\in\{1,\ldots,m\}}\left(\sum_{i\in\{1,\ldots,n\}} (x_i\cdot\mathcal{M}(L;E,F)_{j,i})\cdot f_j\right)\\
&\underset{[\text{theorem: } 11.75}{=} \sum_{j\in\{1,\ldots,m\}}\left(\sum_{i\in\{1,\ldots,n\}} x_i\cdot\mathcal{M}(L;E,F)_{j,i}\right)f_j\\
&= \sum_{j\in\{1,\ldots,m\}}\left(\sum_{i\in\{1,\ldots,n\}}\mathcal{M}(L;E,F)_{j,i}\cdot x_i\right)f_j
\end{aligned}$$

which as $\{f_i|i\in\{1,\ldots,m\}\}$ is a basis proves by the uniqueness of the expansion proves that

$$\forall j\in\{1,\ldots,m\}\; L(x)_j=\sum_{i\in\{1,\ldots,n\}}\mathcal{M}(L;E,F)_{j,i}\cdot x_i$$

2. Let $M=\mathcal{M}(L;E,F)\in\mathcal{M}_{m,n}(F)$. Let $y\in L(X)$ then there exists a $x\in X$ such that $y=L(x)$. As $\{e_1,\ldots,e_n\}$ is a basis of X we have $x=\sum_{i\in\{1,\ldots,n\}}x_i\cdot e_i$ so that

$$y=L\left(\sum_{i\in\{1,\ldots,n\}}x_i\cdot e_i\right)\underset{\text{[theorem: 11.180]}}{=}\sum_{i\in\{1,\ldots,n\}}x_i\cdot L(e_i)$$

proving by [theorem: 11.97] that $y\in\operatorname{span}(\{L(e_i)|i\in\{1,\ldots,n\}\})$, hence

$$L(X)\subseteq\operatorname{span}(\{L(e_i)|i\in\{1,\ldots,n\}\})\tag{11.219}$$

If $y\in\operatorname{span}(\{L(e_i)|i\in\{1,\ldots,n\}\})$ then by [theorem: 11.97] there exists a $\{\alpha_i\}_{i\in\{1,\ldots,n\}}\subseteq F$ such that

$$y=\sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot L(e_i)\underset{\text{[theorem: 11.180]}}{=}L\left(\sum_{i\in[1,\ldots,n]}\alpha_i\cdot e_i\right)\in L(X)$$

proving that $\operatorname{span}(\{L(e_i)|i\in\{1,\ldots,n\}\})\subseteq L(X)$. Combining this with [eq: 11.219] gives

$$L(X)=\operatorname{span}(\{L(e_i)|i\in\{1,\ldots,n\}\})\tag{11.220}$$

Using [theorem: 11.141] there exist a basis

$$B\subseteq\{L(e_i)|i\in\{1,\ldots,n\}\}\text{ for }L(X)$$

Using the definition of the rank of a linear map we have that

$$\operatorname{rank}(L)\underset{\text{def}}{=}\dim(L(X))=\operatorname{card}(B)$$

By [theorem: 11.90] there exist a bijection $\beta:\{1,\ldots,\operatorname{rank}(L)\}\rightarrow J\subseteq\{1,\ldots,n\}$ such that

$$B=\{L(e_{\beta(1)},\ldots,L(e_{\beta(\operatorname{rank}(L))}))\}\text{ where }B\text{ is a basis of }L(X)\tag{11.221}$$

Define now

$$B_c=\{\operatorname{col}(M,\beta(i))|i\in\{1,\ldots,\operatorname{rank}(L)\}\}\subseteq\operatorname{cols}(M)$$

then by [theorem: 11.104] we have that

$$\operatorname{span}(B_c)\subseteq\operatorname{span}(\operatorname{cols}(M))\tag{11.222}$$

For the opposite inclusion, let $c\in\operatorname{span}(\operatorname{cols}(M))=\operatorname{span}(\operatorname{col}(M,i)|i\in\{1,...,n\})$ then by [theorem: 11.97] there exist a $\{\lambda_i\}_{i\in\{1,\ldots,n\}}\subseteq F$ such that

$$c=\sum_{i\in\{1,\ldots,n\}}\lambda_i\cdot\operatorname{col}(M,i)$$

Let $k\in\{1,\ldots,m\}$ then we have

$$\begin{aligned}c_k&=\left(\sum_{i\in\{1,\ldots,n\}}\lambda_i\cdot\operatorname{col}(M,i)\right)_k\\&\underset{\text{[theorem: 11.47]}}{=}\sum_{i\in\{1,\ldots,n\}}\lambda_i\cdot(\operatorname{col}(M,i))_k\\&\underset{\text{[definition: 11.317]}}{=}\sum_{i\in\{1,\ldots,n\}}\lambda_i\cdot M_{k,i}\end{aligned}\tag{11.223}$$

Consider now the sum $\sum_{i\in\{1,\ldots,n\}}\lambda_i\cdot L(e_i)$, then we have by [theorem: 11.97] that

$$\sum_{i\in\{1,\ldots,n\}}\lambda_i\cdot L(e_i)\in\operatorname{span}(\{L(e_i)|i\in\{1,\ldots,n\}\})\underset{\text{[eq: 11.220]}}{=}L(X)$$

By the above, [eq: 11.221] and [theorem: 11.137] there exist a $\{\gamma_i\}_{i\in\{1,\ldots,\mathrm{rank}(L)\}}\subseteq F$ such that

$$\begin{aligned}
\sum_{i\in\{1,\ldots,n\}}\lambda_i\cdot L(e_i) &= \sum_{j\in\{1,\ldots,\mathrm{rank}(L)\}}\gamma_j\cdot L(e_{\beta(j)})\\
&\underset{[\text{definition: } 11.321]}{=} \sum_{j\in\{1,\ldots,\mathrm{rank}(L)\}}\gamma_j\cdot\left(\sum_{i\in\{1,\ldots,m\}}\mathcal{M}(L;E,F)_{i,\beta(j)}\cdot f_i\right)\\
&= \sum_{j\in\{1,\ldots,\mathrm{rank}(L)\}}\gamma_j\cdot\left(\sum_{i\in\{1,\ldots,m\}}M_{i,\beta(j)}\cdot f_i\right)\\
&\underset{[\text{theorem: } 11.75]}{=} \sum_{j\in\{1,\ldots,\mathrm{rank}(L)\}}\left(\sum_{i\in\{1,\ldots,m\}}\gamma_j\cdot(M_{i,\beta(j)}\cdot f_i)\right)\\
&\underset{[\text{theorem: } 11.41]}{=} \sum_{i\in\{1,\ldots,m\}}\left(\sum_{j\in\{1,\ldots,\mathrm{rank}(L)\}}\gamma_j\cdot(M_{i,\beta(j)}\cdot f_i)\right)\\
&= \sum_{i\in\{1,\ldots,m\}}\left(\sum_{j\in\{1,\ldots,\mathrm{rank}(L)\}}(\gamma_j\cdot M_{i,\beta(j)})\cdot f_i\right)\\
&\underset{[\text{theorem: } 11.75]}{=} \sum_{i\in\{1,\ldots,m\}}\left(\sum_{j\in\{1,\ldots,\mathrm{rank}(L)\}}\gamma_j\cdot M_{i,\beta(j)}\right)\cdot f_i\\
\sum_{i\in\{1,\ldots,n\}}\lambda_i\cdot L(e_i) &\underset{[\text{definition: } 11.321]}{=} \sum_{j\in\{1,\ldots,n\}}\lambda_j\cdot\left(\sum_{i\in\{1,\ldots,m\}}\mathcal{M}(L;E,F)_{i,j}\cdot f_i\right)\\
&= \sum_{j\in\{1,\ldots,n\}}\lambda_j\cdot\left(\sum_{i\in\{1,\ldots,m\}}M_{i,j}\cdot f_i\right)\\
&\underset{[\text{theorem: } 11.75]}{=} \sum_{j\in\{1,\ldots,n\}}\left(\sum_{i\in\{1,\ldots,m\}}\lambda_j\cdot(M_{i,j}\cdot f_i)\right)\\
&\underset{[\text{theorem: } 11.41]}{=} \sum_{i\in\{1,\ldots,m\}}\left(\sum_{j\in\{1,\ldots,n\}}\lambda_j\cdot(M_{i,j}\cdot f_i)\right)\\
&= \sum_{i\in\{1,\ldots,m\}}\left(\sum_{j\in\{1,\ldots,n\}}(\lambda_j\cdot M_{i,j})\cdot f_i\right)\\
&\underset{[\text{theorem: } 11.75]}{=} \sum_{i\in\{1,\ldots,m\}}\left(\sum_{j\in\{1,\ldots,n\}}\lambda_j\cdot M_{i,j}\right)\cdot f_i
\end{aligned}$$

Using the uniqueness of the expansion of a vector in a basis we have from the above that

$$\forall i\in\{1,\ldots,n\}\text{ we have }\sum_{j\in\{1,\ldots,\mathrm{rank}(L)\}}\gamma_j\cdot M_{i,\beta(j)}=\sum_{j\in\{1,\ldots,n\}}\lambda_j\cdot M_{i,j}$$

and substituting this in [eq: 11.223] proves that $\forall k\in\{1,\ldots,m\}$

$$\begin{aligned}
c_k &= \sum_{j\in\{1,\ldots,\mathrm{rank}(L)\}}\gamma_j\cdot M_{k,\beta(j)}\\
&\underset{[\text{definition: } 11.317]}{=} \sum_{j\in\{1,\ldots,\mathrm{rank}(L)\}}\gamma_j\cdot\mathrm{col}(M,\beta(j))_k\\
&\underset{[\text{theorem: } 11.47]}{=} \left(\sum_{j\in\{1,\ldots,\mathrm{rank}(L)\}}\gamma_j\cdot\mathrm{col}(M,\beta(j))\right)_k
\end{aligned}$$

proving that

$$c=\sum_{j\in\{1,\ldots,\mathrm{rank}(L)\}}\gamma_j\cdot\mathrm{col}(M,\beta(j))\in\mathrm{span}(B_c)$$

Hence $\operatorname{span}(\operatorname{cols}(M)) \subseteq \operatorname{span}(B_c)$ which together with [eq: 11.222] gives $\operatorname{span}(\operatorname{cols}(M)) = \operatorname{span}(B_c)$ or using the definition of B_c.

$$\operatorname{span}\{\operatorname{col}(M, \beta(i)) | i \in \{1, \ldots, \operatorname{rank}(L)\}\} = \operatorname{span}(\operatorname{cols}(M)) \tag{11.224}$$

Next we prove that $\{\operatorname{col}(M, \beta(i))\}_{i \in \{1, \ldots, \operatorname{rank}(L)\}}$ is linear independent. Let $\{\alpha_i\}_{i \in \{1, \ldots, \operatorname{rank}(L)\}} \subseteq F$ such that

$$\sum_{i \in \{1, \ldots, \operatorname{rank}(L)\}} \alpha_i \cdot \operatorname{col}(M, \beta(i)) = 0$$

then $\forall k \in \{1, \ldots, m\}$ we have

$$\begin{aligned}
0 &= \left(\sum_{i \in \{1, \ldots, \operatorname{rank}(L)\}} \alpha_i \cdot \operatorname{col}(M, \beta(i)) \right)_k \\
&\underset{[\text{theorem: } 11.47]}{=} \sum_{i \in \{1, \ldots, \operatorname{rank}(L)\}} \alpha_i \cdot \operatorname{col}(M, \beta(i))_k \\
&\underset{[\text{definition: } 11.317]}{=} \sum_{i \in \{1, \ldots, \operatorname{rank}(L)\}} \alpha_i \cdot M_{k, \beta(i)}
\end{aligned} \tag{11.225}$$

Now

$$\begin{aligned}
\sum_{i \in \{1, \ldots, \operatorname{rank}(L)\}} \alpha_i \cdot L(e_{\beta(i)}) &\underset{[\text{definition: } 11.321]}{=} \\
\sum_{i \in \{1, \ldots, \operatorname{rank}(L)\}} \alpha_i \cdot \left(\sum_{j \in \{1, \ldots, m\}} \mathcal{M}(L; E, F)_{j, \beta(i)} \cdot f_j \right) &= \\
\sum_{i \in \{1, \ldots, \operatorname{rank}(L)\}} \alpha_i \cdot \left(\sum_{j \in \{1, \ldots, m\}} M_{j, \beta(i)} \cdot f_j \right) &\underset{[\text{theorem: } 11.75]}{=} \\
\sum_{i \in \{1, \ldots, \operatorname{rank}(L)\}} \left(\sum_{j \in \{1, \ldots, m\}} \alpha_i \cdot (M_{j, \beta(i)} \cdot f_j) \right) &\underset{[\text{theorem: } 11.41]}{=} \\
\sum_{j \in \{1, \ldots, m\}} \left(\sum_{i \in \{1, \ldots, \operatorname{rank}(L)\}} \alpha_i \cdot (M_{j, \beta(i)} \cdot f_j) \right) &= \\
\sum_{j \in \{1, \ldots, m\}} \left(\sum_{i \in \{1, \ldots, \operatorname{rank}(L)\}} (\alpha_i \cdot M_{j, \beta(i)}) \cdot f_j \right) &\underset{[\text{theorem: } 11.75]}{=} \\
\sum_{j \in \{1, \ldots, m\}} \left(\sum_{i \in \{1, \ldots, \operatorname{rank}(L)\}} \alpha_i \cdot M_{j, \beta(i)} \right) \cdot f_j &\underset{[\text{eq: } 11.225]}{=} \\
\sum_{j \in \{1, \ldots, m\}} 0 \cdot f_j &= \\
0 &
\end{aligned} \tag{11.226}$$

Now by [eq: 11.221] $\{L(e_{\beta(1)}), \ldots, L(e_{\beta(\operatorname{rank}(L))})\}$ is a basis, hence linear independent, so by [theorem: 11.115]

$$\forall i \in \{1, \ldots, m\} \text{ we have } \alpha_i = 0$$

So $\{\operatorname{col}(M, \beta(i))\}_{i \in \{1, \ldots, \operatorname{rank}(L)\}}$ is linear independent or by [theorem: 11.129]

$$\{\operatorname{col}(M, \beta(i)) | i \in \{1, \ldots, \operatorname{rank}(L)\}\} \text{ is linear independent}$$

Combining this with [eq: 11.224] proves that

$$\{\operatorname{col}(M, \beta(i)) | i \in \{1, \ldots, \operatorname{rank}(L)\}\} \text{ is a basis for } \operatorname{span}(\operatorname{cols}(M))$$

which by [theorem: 11.150] proves that

$$\dim(\operatorname{span}(\operatorname{cols}(M))) = \operatorname{rank}(L)$$

So

$$\operatorname{rank}(M) \underset{\text{def}}{=} \dim(\operatorname{span}(\operatorname{cols}(M))) = \operatorname{rank}(L)$$

proving, as $M = \mathcal{M}(L; E, F)$, that

$$\operatorname{rank}(\mathcal{M}(L; E, F)) = \operatorname{rank}(L)$$

□

Next we show that the mapping that associate a matrix with a linear mapping is linear.

Theorem 11.324. *Let $n, m \in \mathbb{N}$, X, Y finite dimensional vector spaces over a field $\mathcal{F}$ with bases $E = \{e_1, \ldots, e_n\}$,$F = \{f_1, \ldots, f_m\}$ for X, Y then*

$$\mathcal{M}(E, F)\colon \mathrm{Hom}(X, Y) \rightarrow \mathcal{M}_{m,n}(\mathcal{F}) \text{ defined by } \mathcal{M}(E, F)(L) = M(L; E, F)$$

satisfies:

1. $\forall \alpha \in \mathcal{F}$ *and* $L \in \mathrm{Hom}(X, Y)$ *we have* $\mathcal{M}(E, F)(\alpha \cdot L) = \alpha \cdot \mathcal{M}(E, F)(L)$
2. $\forall K, L \in \mathrm{Hom}(X, Y)$ *we have* $\mathcal{M}(E, F)(K + L) = \mathcal{M}(E, F)(K) + \mathcal{M}(E, F)(L)$
3. $\mathcal{M}(E, F)(C_0) = 0$
4. $\mathcal{M}(E, E)(\mathrm{Id}_X) = E$ *[be aware the last E is the identity matrix and the first two E's specifies the basis in X]*

in other words

$$\mathcal{M}(E, G) \in \mathrm{Hom}(\mathrm{Hom}(X, Y), \mathcal{M}_{m,n}(F))$$

Proof.

1. Let $i \in \{1, \ldots, n\}$ then

$$\begin{aligned} \sum_{j \in \{1, \ldots, m\}} \mathcal{M}(\alpha \cdot L; E, F)_{j,i} \cdot f_j &= (\alpha \cdot L)(e_i) \\ &= \alpha \cdot L(e_i) \\ &= \alpha \cdot \sum_{j \in \{1, \ldots, m\}} \mathcal{M}(L; E, F)_{j,i} \cdot f_j \\ &= \sum_{j \in \{1, \ldots, m\}} (\alpha \cdot \mathcal{M}(L; E, F)_{j,i}) \cdot f_j \end{aligned}$$

   proving by the uniqueness of the expansion that $\forall i, j \in \{1, \ldots, m\}$

$$\mathcal{M}(\alpha \cdot L; E, F)_{j,i} = \alpha \cdot \mathcal{M}(L; E, F)_{j,i}$$

   or

$$\mathcal{M}(E, F)(\alpha \cdot L) = \mathcal{M}(\alpha \cdot L; E, F) = \alpha \cdot \mathcal{M}(L; E, F) = \mathcal{M}(E, F)(L)$$

2. Let $i \in \{1, \ldots, n\}$ then

$$\begin{aligned} \sum_{j \in \{1, \ldots, m\}} \mathcal{M}(K + L; E, F)_{j,i} \cdot f_j &= \\ (K + L)(e_i) &= \\ K(e_i) + L(e_i) &= \\ \sum_{j \in \{1, \ldots, m\}} \mathcal{M}(K; E, F)_{j,i} \cdot f_j + \sum_{j \in \{1, \ldots, m\}} \mathcal{M}(L; E, F)_{j,i} \cdot f_j &= \\ \sum_{j \in \{1, \ldots, m\}} (\mathcal{M}(K; E, F)_{j,i} \cdot f_j + \mathcal{M}(L; E, F)_{j,i} \cdot f_j) &= \\ \sum_{j \in \{1, \ldots, m\}} (\mathcal{M}(K; E, F)_{j,i} + \mathcal{M}(L; E, F)_{j,i}) \cdot f_j & \end{aligned}$$

   proving by the uniqueness of the expansion that $\forall j \in \{1, \ldots, m\}$ $\forall i, j \in \{1, \ldots, m\}$

$$\mathcal{M}(K + L; E, F)_{j,i} = \mathcal{M}(K; E, F)_{j,i} + \mathcal{M}(L; E, F)_{j,i}$$

   or

$$\mathcal{M}(E, F)(K + L) = \mathcal{M}(K + L; E, F) = \mathcal{M}(K; E, F) + \mathcal{M}(L; E, F) = \mathcal{M}(E, F)(K) + \mathcal{M}(E, F)(L)$$

3. Let $i \in \{1, \ldots, n\}$ then

$$\begin{aligned} \sum_{j \in \{1, \ldots, m\}} 0 \cdot f_j &= 0 \\ &= C_0(e_i) \\ &= \sum_{j \in \{1, \ldots, m\}} \mathcal{M}(C_0; E, F)_{j,i} \cdot f_j \end{aligned}$$

proving by the uniqueness of the expansion that $\forall i,j\in\{1,\ldots,m\}$

$$\mathcal{M}(C_0;E,F)_{j,i}=0$$

or

$$\mathcal{M}(E,F)(C_0)=\mathcal{M}(C_0;E,F)=0$$

4. Let $i\in\{1,\ldots,n\}$ then

$$\begin{aligned}\sum_{j\in\{1,\ldots,n\}}\delta_{j,i}\cdot e_j &= e_i\\ &= \mathrm{Id}_X(e_i)\\ &= \sum_{j\in\{1,\ldots,n\}}\mathcal{M}(\mathrm{Id}_X;E,E)_{j,i}\cdot f_j\end{aligned}$$

proving by the uniqueness of the expansion that $\forall i,j\in\{1,\ldots,n\}$

$$\mathcal{M}(\mathrm{Id}_X;E,E)_{j,i}=\delta_{j,i}=E_{j,i}$$

or

$$\mathcal{M}(E,E)(\mathrm{Id}_X)=\mathcal{M}(\mathrm{Id}_X;E,E)=E$$ □

Next we prove that $\mathcal{M}(E,F)$ is actually a bijection so that $\mathrm{Hom}(X,Y)$ and $\mathcal{M}_{m,n}(F)$ are isomorphic.

Theorem 11.325. *Let $n,m\in\mathbb{N}$, X,Y finite dimensional vector spaces over a field $\mathcal{F}$ with bases $E=\{e_1,\ldots,e_n\}$,$F=\{f_1,\ldots,f_m\}$ for X, Y then*

$$\mathcal{M}(E,F)\colon\mathrm{Hom}(X,Y)\rightarrow\mathcal{M}_{m,n}(F)\text{ where }\mathcal{M}(E,F)(L)=\mathcal{M}(L;E,F)$$

is a bijection. Further if $M\in\mathcal{M}_{m,n}(F)$ then $\mathcal{M}^{-1}(E,F)(M)$ is defined by

$$\mathcal{M}^{-1}(E,F)(M)(x)=\sum_{i\in\{1,\ldots,m\}}\left(\sum_{j\in\{1,\ldots,n\}}M_{i,j}\cdot x_j\right)\cdot f_i$$

where $\{x_i\}_{i\in\{1,\ldots,n\}}\subseteq F$ is the unique family such that $x=\sum_{i\in\{1,\ldots,n\}}x_i\cdot e_i$

Proof.

injectivity. Let $L_1,L_2\in\mathrm{Hom}(X,Y)$ such that $\mathcal{M}(E,F)(L_1)=\mathcal{M}(E,F)(L_2)$ then we have

$$\mathcal{M}(L_1;E,F)=\mathcal{M}(L_2;E,F) \tag{11.227}$$

Let $x\in X$ then there exists a $\{x_i\}_{i\in\{1,\ldots,n\}}\subseteq F$ such that $x=\sum_{i\in\{1,\ldots,n\}}x_i\cdot e_i$, So

$$\begin{aligned}L_1(x) &= L_1\left(\sum_{i\in\{1,\ldots,n\}}x_i\cdot e_i\right)\\ &\underset{[\text{theorem: }11.180]}{=} \sum_{i\in\{1,\ldots,n\}}x_i\cdot L_1(e_i)\\ &= \sum_{i\in\{1,\ldots,n\}}x_i\cdot\left(\sum_{j\in\{1,\ldots,m\}}\mathcal{M}(L_1;E,F)_{j,i}\cdot f_j\right)\\ &\underset{[\text{eq: }11.227]}{=} \sum_{i\in\{1,\ldots,n\}}x_i\cdot\left(\sum_{j\in\{1,\ldots,m\}}\mathcal{M}(L_2;E,F)_{j,i}\cdot f_j\right)\\ &= \sum_{i\in\{1,\ldots,n\}}x_i\cdot L_2(e_i)\\ &\underset{[\text{theorem: }11.180]}{=} L_2\left(\sum_{i\in\{1,\ldots,n\}}x_i\cdot e_i\right)\\ &= L_2(x)\end{aligned}$$

proving that $L_1 = L_2$.

surjectivity. Let $M \in \mathcal{M}_{m,n}(F)$ then define $L: X \to Y$ by

$$L(x) = \sum_{i \in \{1,\ldots,m\}} \left(\sum_{j \in \{1,\ldots,n\}} M_{i,j} \cdot x_j \right) \cdot f_i \text{ where } x = \sum_{i \in \{1,\ldots,n\}} x_i \cdot e_i$$

then we have for $x, y \in X$ and $\alpha \in F$ that

$$x + y = \sum_{i \in \{1,\ldots,n\}} x_i \cdot e_i + \sum_{i \in \{1,\ldots,n\}} y_i \cdot e_i = \sum_{i \in \{1,\ldots,n\}} (x_i + y_i) \cdot e_i$$

and

$$\alpha \cdot x = \alpha \cdot \sum_{i \in \{1,\ldots,n\}} x_i \cdot e_i = \sum_{i \in \{1,\ldots,n\}} (\alpha \cdot x_i) \cdot e_i$$

So that

$$\begin{aligned} L(x+y) &= \\ \sum_{i \in \{1,\ldots,m\}} \left(\sum_{j \in \{1,\ldots,n\}} M_{i,j} \cdot (x_j + y_j) \right) \cdot f_i &= \\ \sum_{i \in \{1,\ldots,m\}} \left(\sum_{j \in \{1,\ldots,n\}} M_{i,j} \cdot x_j + \sum_{j \in \{1,\ldots,n\}} M_{i,j} \cdot y_j \right) &= \\ \sum_{i \in \{1,\ldots,m\}} \left(\sum_{j \in \{1,\ldots,n\}} M_{i,j} \cdot x_j \right) + \sum_{i \in \{1,\ldots,m\}} \left(\sum_{j \in \{1,\ldots,n\}} M_{i,j} \cdot y_j \right) &= \\ L(x) + L(y) & \end{aligned}$$

and

$$\begin{aligned} L(\alpha \cdot x) &= \sum_{i \in \{1,\ldots,m\}} \left(\sum_{j \in \{1,\ldots,n\}} M_{i,j} \cdot (\alpha \cdot x_j) \right) \cdot f_i \\ &= \sum_{i \in \{1,\ldots,m\}} \left(\alpha \cdot \sum_{j \in \{1,\ldots,n\}} M_{i,j} \cdot x_j \right) \cdot f_i \\ &= \alpha \cdot \sum_{i \in \{1,\ldots,m\}} \left(\sum_{j \in \{1,\ldots,n\}} M_{i,j} \cdot x_j \right) \cdot f_i \\ &= \alpha \cdot L(x) \end{aligned}$$

proving that

$$L \in \mathrm{Hom}(X, Y)$$

Let $i \in \{1, \ldots, n\}$ then

$$e_i = \sum_{j \in \{1,\ldots,m\}} \delta_{j,i} \cdot e_j \text{ so that } (e_i)_j = \delta_{j,i}$$

$$\begin{aligned} \sum_{k \in \{1,\ldots,m\}} \mathcal{M}(L; E, F)_{k,i} \cdot f_k &= L(e_i) \\ &= \sum_{k \in \{1,\ldots,m\}} \left(\sum_{j \in \{1,\ldots,n\}} M_{k,j} \cdot \delta_{j,i} \right) \cdot f_k \\ &= \sum_{k \in \{1,\ldots,m\}} M_{k,i} \cdot f_k \end{aligned}$$

So by uniqueness of the expansion in a basis we have that $\forall i \in \{1, \ldots, n\}$, $\forall k \in \{1, \ldots, m\}$

$$\mathcal{M}(L; E, F)_{k,i} = M_{k,i}$$

hence

$$\mathcal{M}(E, F)(L) = M$$

proving surjectivity and

$$\mathcal{M}(E,F)^{-1}(M)=L$$

□

Corollary 11.326. *Let $n,m\in\mathbb{N}$, X,Y finite dimensional vector spaces over a field $\mathcal{F}$ with bases $E=\{e_1,\ldots,e_n\}$,$F=\{f_1,\ldots,f_m\}$ for X,Y then*

$$\mathcal{M}(E,F)\colon \operatorname{Hom}(X,Y)\rightarrow\mathcal{M}_{m,n}(\mathcal{F}) \text{ defined by } \mathcal{M}(E,F)(L)=M(L;E,F)$$

is a isomorphism. Hence $\operatorname{Hom}(X,Y)$ *is isomorphic with* $\mathcal{M}_{m,n}(\mathcal{F})$

Proof. This follows from [theorem: 11.324] and [theorem: 11.325]. □

Example 11.327. Let $\mathcal{F}$ be a field, $n,m\in\mathbb{N}$ and $E=\{e_1,\ldots,e_n\}\subseteq\mathcal{F}^n$, $F=\{f_1,\ldots,f_m\}\subseteq\mathcal{F}^m$ defined by $(e_i)_j=\delta^n_{i,j}$ and $\{f_i\}_j=\delta^m_{i,j}$ the canonical bases for the vector spaces $\mathcal{F}^n$, $\mathcal{F}^m$ over $\mathcal{F}$ then for $M\in\mathcal{M}_{m,n}(\mathcal{F})$ we have for $x\in\mathcal{F}^n$ that $\forall i\in\{1,\ldots,m\}$

$$(\mathcal{M}(E,F)^{-1}(M)(x))_i=\sum_{j\in\{1,\ldots,n\}} M_{i,j}\cdot x_j$$

or if we define $X=\begin{pmatrix} x_1\\ \vdots\\ x_n\end{pmatrix}$ then

$$\mathcal{M}(E,F)^{-1}(M)(x)=M\cdot X$$

Proof. Let $x\in\mathcal{F}^n$ then we have

$$x_i=\sum_{j\in\{1,\ldots,n\}} x_j\cdot\delta^n_{j,i}=\sum_{j\in\{1,\ldots,n\}} x_j\cdot(e_j)_i=\left(\sum_{j\in\{1,\ldots,n\}} x_j\cdot e_j\right)_i$$

so that

$$x=\sum_{j\in\{1,\ldots,n\}} x_j\cdot e_j$$

hence

$$\begin{aligned}(\mathcal{M}(E,F)^{-1}(M)(x))_k &\underset{[\text{theorem: } 11.325]}{=} \left(\sum_{i\in\{1,\ldots,m\}}\left(\sum_{j\in\{1,\ldots,n\}} M_{i,j}\cdot x_j\right)\cdot f_i\right)_k\\ &\underset{[\text{theorem: } 11.47]}{=} \sum_{i\in\{1,\ldots,m\}}\left(\sum_{j\in\{1,\ldots,n\}} M_{i,j}\cdot x_j\right)\cdot(f_i)_k\\ &= \sum_{i\in\{1,\ldots,m\}}\left(\sum_{j\in\{1,\ldots,n\}} M_{i,j}\cdot x_j\right)\cdot\delta^m_{i,k}\\ &= \sum_{j\in\{1,\ldots,n\}} M_{k,j}\cdot x_j\end{aligned}$$

□

We have the following relation between the composition of linear mappings and the product of matrices.

Theorem 11.328. *Let $n,m,r\in\mathbb{N}$, X,Y,Z finite dimensional vector spaces over a field $\mathcal{F}$ with bases $E=\{e_1,\ldots,e_n\}$,$F=\{f_1,\ldots,f_m\}$,$G=\{g_1,\ldots,g_r\}$ for X,Y,Z then we have for $L_1\in\operatorname{Hom}(X,Y)$ and $L_2\in\operatorname{Hom}(Y,Z)$ that*

$$\mathcal{M}(L_2\circ L_1;E,G)=\mathcal{M}(L_2;F,G)\cdot\mathcal{M}(L_1;E,F)$$

or in other words

$$\mathcal{M}(E,G)(L_2\circ L_1)=\mathcal{M}(F,G)(L_2)\cdot\mathcal{M}(E,F)(L_1)$$

Proof. Let $i\in\{1,\ldots,n\}$ then we have

$$\begin{aligned}\sum_{k\in\{1,\ldots,r\}}\mathcal{M}(L_2\circ L_1;E,G)_{k,i}\cdot g_k &=\\ (L_2\circ L_1)(e_i) &=\\ L_2(L_1(e_i)) &=\end{aligned}$$

$$
\begin{aligned}
L_2\left(\sum_{j\in\{1,\ldots,m\}} \mathcal{M}(L_1;E,F)_{j,i}\cdot f_j\right) &\underset{[\text{theorem: } 11.180]}{=}\\
\sum_{j\in\{1,\ldots,m\}} \mathcal{M}(L_1;F,G)_{j,i}\cdot L_2(f_j) &=\\
\sum_{j\in\{1,\ldots,m\}} \mathcal{M}(L_1;E,F)_{j,i}\cdot\left(\sum_{k\in\{1,\ldots,r\}} \mathcal{M}(L_2;F,G)_{k,j}\cdot g_k\right) &\underset{[\text{theorem: } 11.75]}{=}\\
\sum_{j\in\{1,\ldots,m\}}\left(\sum_{k\in\{1,\ldots,r\}} \mathcal{M}(L_1;E,F)_{j,i}\cdot(\mathcal{M}(L_2;F,G)_{k,j}\cdot g_k)\right) &\underset{[\text{theorem: } 11.41]}{=}\\
\sum_{k\in\{1,\ldots,r\}}\left(\sum_{j\in\{1,\ldots,m\}} (\mathcal{M}(L_2;F,G)_{k,j}\cdot\mathcal{M}(L_1;E,F)_{j,i})\cdot g_k\right) &\underset{[\text{theorem: } 11.75]}{=}\\
\sum_{k\in\{1,\ldots,r\}} (\mathcal{M}(L_2;F.G)\cdot\mathcal{M}(L_1;E,F))_{k,i}\cdot g_k &
\end{aligned}
$$

By the uniqueness of the expansion in a basis it follow then that $\forall i\in\{1,\ldots,n\}$ and $\forall k\in\{1,\ldots r\}$ that

$$\mathcal{M}(L_2\circ L_1;E,G)_{k,i}=(\mathcal{M}(L_2;F.G)\cdot\mathcal{M}(L_1;E,F))_{k,i}$$

proving

$$\mathcal{M}(L_2\circ L_1;E,G)=(\mathcal{M}(L_2;F.G)\cdot\mathcal{M}(L_1;E,F))$$ □

11.8.3 Inverse, Determinant and Adjoint of matrices

For linear transformations we have the inverse of a linear transformation based on the composition operator $\circ$. Not every linear transformation has a inverse but we can use the determinant of a linear transformation to check if a linear transformation is invertible. Finally we can use the determinant and the adjoint of a linear transformation to calculate the inverse of a invertible linear transformation. Having just proved that the composition of linear mappings can be translated in the product of their associated matrices, it seems reasonable to define the inverse of a matrix in terms of the product of matrices and introduce the concept of the determinant and adjoint of matrices. It will also turn out that the determinant and adjoint of a linear transformation are related to the determinant and adjoint of a linear mapping.

Definition 11.329. *Let $n\in\mathbb{N}$, F a field and $M\in\mathcal{M}_{n,n}(F)$ then $N\in\mathcal{M}_{n,n}(F)$ is a **inverse** of M if*

$$N\cdot M=E=M\cdot N$$

It turns out that a matrix can have only one inverse.

Theorem 11.330. *Let $n\in\mathbb{N}$, F a field and $M\in\mathcal{M}_{n,n}$ then if N_1 and N_2 are inverses of M we have that $N_1=N_2$*

Proof. Let N_1,N_2 be inverses of M then we have

$$N_1=N_1\cdot E=N_1\cdot(M\cdot N_2)=(N_1\cdot M)\cdot N_2=E\cdot N_2=N_2$$ □

The above leads to the following definition

Definition 11.331. *Let $n\in\mathbb{N}$ and F a field then $M\in\mathcal{M}_{n,n}(F)$ is **invertible** if there exist a $N\in\mathcal{M}_{n,n}(F)$ that is the inverse of M. This **unique** inverse is noted as M^{-1}, so we have that*

$$M^{-1}\cdot M=E=M\cdot M^{-1}$$

The inverse of a inverse is the invertible matrix itself.

Theorem 11.332. *Let $n\in\mathbb{N}$, F a field and $M\in\mathcal{M}_{n,n}(F)$ a invertible matrix then M^{-1} is also invertible and $(M^{-1})^{-1}=M$*

Proof. This can be proved by a easy calculation:

$$M^{-1}\cdot M=E=M\cdot M^{-1}$$

hence M is the inverse of M^{-1}, so that $(M^{-1})^{-1}=M$. □

Theorem 11.333. *Let $n\in\mathbb{N}$, X be a vector space over a field , $F=\{f_1,\ldots,f_n\}$ a basis for X and $L\in\operatorname{Hom}(X,X)$ then we have*

$$L \text{ has a inverse } L^{-1} \text{ [or equivalently by [theorem: 2.71] } L \text{ is a isomorphism]}$$
$$\Updownarrow$$
$$\mathcal{M}(L;F,F) \text{ is invertible}$$

Further if L^{-1} exists then

$$\mathcal{M}(L^{-1};F,F)=\mathcal{M}(L,F,F)^{-1}$$

Proof.

⇒. We have for L^{-1} that

$$\begin{aligned}
E &\underset{\text{[theorem: 11.324 (4)]}}{=} \mathcal{M}[\mathrm{Id}_X;F,F]\\
&= \mathcal{M}[L\circ L^{-1};F,F]\\
&\underset{\text{[theorem: 11.328]}}{=} \mathcal{M}(L;F,F)\cdot\mathcal{M}(L^{-1};F,F)\\
E &\underset{\text{[theorem: 11.324 (4)]}}{=} \mathcal{M}[\mathrm{Id}_X;F,F]\\
&= \mathcal{M}[L^{-1}\circ L;F,F]\\
&\underset{\text{[theorem: 11.328]}}{=} \mathcal{M}(L^{-1};F,F)\cdot\mathcal{M}(L;F,F)
\end{aligned}$$

proving that $\mathcal{M}(L;F,F)$ has a inverse and that

$$\mathcal{M}(L;F,F)^{-1}=\mathcal{M}(L^{-1};F,F)$$

⇐. Define $K\underset{\text{[theorem: 11.325]}}{=}\mathcal{M}(F,F)^{-1}(\mathcal{M}(L;F,F)^{-1})$ then we have

$$\begin{aligned}
K\circ L &\underset{\text{[theorem: 11.325]}}{=}\\
\mathcal{M}(F,F)^{-1}(\mathcal{M}(F,F)(K\circ L)) &\underset{\text{[theorem:11.328]}}{=}\\
\mathcal{M}(F,F)^{-1}(\mathcal{M}(F,F)(K)\cdot\mathcal{M}(F,F)(L)) &=\\
\mathcal{M}(F,F)^{-1}(\mathcal{M}(F,F)(\mathcal{M}(F,F)^{-1}(\mathcal{M}(L;F,F)^{-1}))\cdot\mathcal{M}(F,F)(L)) &=\\
\mathcal{M}(F,F)^{-1}(\mathcal{M}(L;F,F)^{-1}\cdot\mathcal{M}(F,F)(L)) &=\\
\mathcal{M}(F,F)^{-1}(\mathcal{M}(L;F,F)^{-1}\cdot\mathcal{M}(L;F,F)) &=\\
\mathcal{M}(F,F)^{-1}(E) &\underset{\text{[theorem: 11.324]}}{=}\\
\mathrm{Id}_X &
\end{aligned}$$

and

$$\begin{aligned}
L\circ K &\underset{\text{[theorem: 11.325]}}{=}\\
\mathcal{M}(F,F)^{-1}(\mathcal{M}(F,F)(L\circ K)) &\underset{\text{[theorem:11.328]}}{=}\\
\mathcal{M}(F,F)^{-1}(\mathcal{M}(F,F)(L)\cdot\mathcal{M}(F,F)(K)) &=\\
\mathcal{M}(F,F)^{-1}(\mathcal{M}(F,F)(L)\cdot\mathcal{M}(F,F)(\mathcal{M}(F,F)^{-1}(\mathcal{M}(L;F,F)^{-1}))) &=\\
\mathcal{M}(F,F)^{-1}(\mathcal{M}(F,F)(L)\cdot\mathcal{M}(L;F,F)^{-1}\cdot) &=\\
\mathcal{M}(F,F)^{-1}(\mathcal{M}(L;F,F)\cdot\mathcal{M}(L;F,F)^{-1}) &=\\
\mathcal{M}(F,F)^{-1}(E) &\underset{\text{[theorem: 11.324]}}{=}\\
\mathrm{Id}_X &
\end{aligned}$$

which proves that $K\circ L=\mathrm{Id}_X=L\circ K$ so that

$$L \text{ is invertible with inverse } K=\mathcal{M}(F,F)^{-1}(\mathcal{M}(F,F)(L))$$ □

We are now ready to define the determinant of a matrix.

Definition 11.334. (Determinant) *Let $n\in\mathbb{N}$, F a field and $M\in\mathcal{M}_{n,n}(F)$ a square matrix then*

$$\det(M)=\sum_{\sigma\in P_n}\operatorname{sign}(\sigma)\cdot\left(\prod_{i\in\{1,\ldots,n\}}M_{i,\sigma(i)}\right)$$

Another notation for the determinant is $|M|$ so using this notation we have

$$|M|=\sum_{\sigma\in P_n}\operatorname{sign}(\sigma)\cdot\left(\prod_{i\in\{1,\ldots,n\}}M_{i,\sigma(i)}\right)$$

Example 11.335. If $n=1$ then $M\in\mathcal{M}_{1,1}(F)$ is of the form $\left(\, M_{1,1}\,\right)$ and $P_1=\{\mathrm{Id}_{\{1\}}\}$ so that

$$\det(M)=\sum_{\sigma\in\{\mathrm{Id}_{\{1\}}\}}\operatorname{sing}(\sigma)\cdot\left(\prod_{i\in\{1\}}M_{i,\sigma(i)}\right)=\operatorname{sign}(\mathrm{Id}_{\{1\}})\left(\prod_{i\in\{1\}}M_{i,\mathrm{Id}_{\{1\}}(i)}\right)=M_{11}$$

Definition 11.336. (diagonal matrix) *Let $n,m\in\mathbb{N}$ and F a field then $M\in\mathcal{M}_{n,m}$ is a **diagonal matrix** if $\forall(i,j)\in\{1,\ldots,n\}\cdot\{1,\ldots,m\}$ $M_{i,j}=\delta_{i,j}\cdot M_{i,i}$*

One nice property for diagonal square matrices is that the determinant is easy to calculate.

Theorem 11.337. *Let $n\in\mathbb{N}$, F a field, $M\in\mathcal{M}_{n,n}(F)$ a diagonal matrix then*

$$\det(M)=\prod_{i\in\{1,\ldots,n\}}M_{i,i}$$

Proof. First if $\sigma\in P_n$ is such that $\sigma\neq\mathrm{Id}_{\{1,\ldots,n\}}$ then $\exists k\in\{1,\ldots,n\}$ such that $k\neq\sigma(k)$. Hence

$$\begin{aligned}\prod_{i\in\{1,\ldots,n\}}M_{i,\sigma(i)} &= \left(\prod_{i\in\{1,\ldots,n\}\setminus\{k\}}M_{i,j}\right)\cdot\left(\prod_{i\in\{k\}}M_{i\sigma(i)}\right)\\ &= \left(\prod_{i\in\{1,\ldots,n\}\setminus\{k\}}M_{i,j}\right)\cdot M_{k,\sigma(k)}\\ &= \left(\prod_{i\in\{1,\ldots,n\}\setminus\{k\}}M_{i,j}\right)\cdot M_{k,k}\cdot\delta_{k,\sigma(k)}\\ &= \left(\prod_{i\in\{1,\ldots,n\}\setminus\{k\}}M_{i,j}\right)\cdot M_{k,k}\cdot 0\\ &= 0\end{aligned}$$

So that

$$\begin{aligned}\det(M) &=\\ \sum_{\sigma\in P_n}\left(\operatorname{sign}(\sigma)\cdot\prod_{i\in\{1,\ldots,n\}}M_{i,\sigma(i)}\right) &=\\ \sum_{\sigma\in P_n\setminus\{\mathrm{Id}_{\{1,\ldots,n\}}\}}\left(\operatorname{sign}(\sigma)\cdot\prod_{i\in\{1,\ldots,n\}}M_{i,\sigma(i)}\right)+\sum_{\sigma\in\{\mathrm{Id}_{\{1,\ldots,n\}}\}}\left(\operatorname{sign}(\sigma)\cdot\prod_{i\in\{1,\ldots,n\}}M_{i,\sigma(i)}\right) &=\\ \sum_{\sigma\in P_n\setminus\{\mathrm{Id}_{\{1,\ldots,n\}}\}}\operatorname{sign}(\sigma)\cdot 0+\sum_{\sigma\in\{\mathrm{Id}_{\{1,\ldots,n\}}\}}\left(\operatorname{sign}(\sigma)\cdot\prod_{i\in\{1,\ldots,n\}}M_{i,\sigma(i)}\right) &=\\ \operatorname{sign}(\mathrm{Id}_{\{1,\ldots,n\}})\cdot\prod_{i\in\{1,\ldots,n\}}M_{i,\mathrm{Id}_{\{1,\ldots,n\}}(i)} &=\\ \prod_{i\in\{1,\ldots,n\}}M_{i,i} &=\end{aligned}$$

□

The determinant is invariant under the transpose operation on a matrix.

Theorem 11.338. *Let $n\in\mathbb{N}$, F a field and $M\in\mathcal{M}_{n,n}(F)$ then*

$$\det(M)=\det(M^T)$$

Proof. First define $(.)^{-1}: P_n\rightarrow P_n$ by $(.)^{-1}(\sigma)=\sigma^{-1}$ then we have:

injectivity. If $(.)^{-1}(\sigma)=(.)^{-1}(\rho)$ then

$$\begin{aligned}\sigma^{-1}=\rho^{-1} &\Rightarrow \sigma\circ\sigma^{-1}=\sigma\circ\rho^{-1}\\ &\Rightarrow \mathrm{Id}_{\{1,\ldots,n\}}=\sigma\circ\rho^{-1}\\ &\Rightarrow \mathrm{Id}_{\{1,\ldots,n\}}\circ\rho=(\sigma\circ\rho^{-1})\circ\rho\\ &\Rightarrow \rho=\sigma\end{aligned}$$

surjectivity. If $\sigma\in P_n$ then $(.)^{-1}(\sigma^{-1})=(\sigma^{-1})^{-1}=\sigma$

so that

$$(.)^{-1}: P_n\rightarrow P_n \text{ is a bijection} \tag{11.228}$$

First note that $\langle F,\cdot\rangle$ is a Abelian semi group

$$\prod_{i\in\{1,\ldots,n\}} M_{\sigma(i),i}=\prod_{i\in\{1,\ldots,n\}} M_{\sigma(i),\sigma(\sigma^{-1}(i))}\underset{[\text{theorem: } 11.36]}{=}\prod_{i\in\{1,\ldots,n\}} M_{i,\sigma^{-1}(i)} \tag{11.229}$$

So that

$$\begin{aligned}\det(M^T) &= \sum_{\sigma\in P_n}\mathrm{sign}(\sigma)\cdot\left(\prod_{i\in\{1,\ldots,n\}}(M^T)_{i,\sigma(i)}\right)\\ &= \sum_{\sigma\in P_n}\mathrm{sign}(\sigma)\cdot\left(\prod_{i\in\{1,\ldots,n\}}M_{\sigma(i),i}\right)\\ &\underset{[\text{eq: } 11.229]}{=} \sum_{\sigma\in P_n}\mathrm{sign}(\sigma)\cdot\left(\prod_{i\in\{1,\ldots,n\}}M_{i,\sigma^{-1}(i)}\right)\\ &\underset{[\text{theorem: } 11.235]}{=} \sum_{\sigma\in P_n}\mathrm{sign}(\sigma^{-1})\cdot\left(\prod_{i\in\{1,\ldots,n\}}M_{i,\sigma^{-1}(i)}\right)\\ &= \sum_{\sigma\in P_n}\mathrm{sign}((.)^{-1}(\sigma))\cdot\left(\prod_{i\in\{1,\ldots,n\}}M_{i,(.)^{-1}(\sigma)(i)}\right)\\ &\underset{[\text{eq: } 11.228 \text{ and theorem: } 11.36]}{=} \sum_{\sigma\in P_n}\mathrm{sign}(\sigma)\cdot\left(\prod_{i\in\{1,\ldots,n\}}M_{i,\sigma(i)}\right)\\ &= \det(M)\end{aligned}$$

□

Theorem 11.339. *Let $n\in\mathbb{N}$, X a vector space over a field F with characteristic zero with basis $E=\{e_1,\ldots,e_n\}\subseteq X$ and $L\in\mathrm{Hom}(X,X)$ then we have:*

$$\det(L)=\det(\mathcal{M}(L;E,E))=\det(\mathcal{M}(E,E)(L))$$

Proof. Using [theorem: 11.283] there exists a determinant function Δ in F^n such that $\Delta(e_1,\ldots,e_n)=1$. Then we have

$$\Delta_L(e,\ldots,e_n)\underset{[\text{theorem: } 11.295]}{=}\det(L)\cdot\Delta(e_1,\ldots,e_n)=\det(L) \tag{11.230}$$

Let $M = \mathcal{M}(L, E, E)$ then we have

$$\begin{aligned}
\det(L) &\underset{[\text{eq: } 11.230]}{=} \Delta_L(e, \ldots, e_n) \\
&\underset{\text{def}}{=} \Delta(L(e_1), \ldots, L(e_n)) \\
&= \Delta\left(\sum_{i \in \{1, \ldots, n\}} M_{i,1} \cdot e_i, \ldots, \sum_{i \in \{1, \ldots, n\}} M_{i,n} \cdot e_i\right) \\
&\underset{[\text{theorem: } 11.278]}{=} \sum_{\sigma \in P_n} \left(\prod_{i \in \{1, \ldots, n\}} M_{\sigma(i),i}\right) \cdot (\sigma\Delta)(e_1, \ldots, e_n) \\
&\underset{\Delta \text{ is skew-symmetry}}{=} \sum_{\sigma \in P_n} \left(\prod_{i \in \{1, \ldots, n\}} M_{\sigma(i),i}\right) \cdot (\mathrm{sign}(\sigma) \cdot \Delta(e_1, \ldots, e_n)) \\
&= \sum_{\sigma \in P_n} \mathrm{sign}(\sigma) \cdot \left(\prod_{i \in \{1, \ldots, n\}} M_{\sigma(i),i}\right) \\
&= \det(M^T) \\
&\underset{[\text{theorem: } 11.338]}{=} \det(M)
\end{aligned}$$

□

Corollary 11.340. *Let $n \in \mathbb{N}$, X a vector space over a field F with characteristic zero with basis $E = \{e_1, \ldots, e_n\} \subseteq X$ and $L \in \mathrm{Hom}(X, X)$ then*

$$L \textit{ is a linear isomorphism}$$

$$\Updownarrow$$

$$\det(\mathcal{M}(L; E, E)) \neq 0$$

Proof. We have

⇒. If L is a linear isomorphism it follows from [theorem: 11.299] that $\det(L) \neq 0$. So that $\det(\mathcal{M}(L; E, E)) \underset{[\text{theorem: } 11.339]}{=} \det(L) \neq 0$.

⇐. As $\det(L) \underset{[\text{theorem: } 11.339]}{=} \det(\mathcal{M}(L; E, E)) \neq 0$ it follows from [theorem: 11.299] that L is a linear isomorphism. □

We use now the above theorem to prove the following:

Theorem 11.341. *Let $n \in \mathbb{N}$ and $\mathcal{F}$ a field of characteristic zero then we have:*

1. *If $M_1, M_2 \in \mathcal{M}_{n,n}(\mathcal{F})$ then $\det(M_1 \cdot M_2) = \det(M_1) \cdot \det(M_2)$*
2. $\det(E) = 1$
3. *If M has a inverse M^{-1} then $\det(M) \neq 0$ and $\det(M^{-1}) = (\det(M))^{-1}$*

Proof. Let $\mathcal{F}^n$ be the vector space over $\mathcal{F}$ with the canonical basis $F = \{e_1, \ldots, e_n\}$ defined by $\{e_i\}_k = \delta_{i,k}$ for $k \in \{1, \ldots, n\}$ [see: theorem: 11.159].

1. Take then $L_1 = \mathcal{M}[F, F]^{-1}(M_1)$ and $L_2 = \mathcal{M}(F, F)^{-1}(M_2)$ then we have:

$$\mathcal{M}(F, F)(L_1) = M_1 \text{ and } \mathcal{M}(F, F)(L_2) = M_2 \tag{11.231}$$

$$\begin{aligned}
\mathcal{M}(F, F)(L_1 \circ L_2) &\underset{[\text{theorem: } 11.328]}{=} \mathcal{M}(F, F)(L_1) \cdot \mathcal{M}(F, F)(L_2) \\
&= \mathcal{M}(F, F)(\mathcal{M}[F, F]^{-1}(M_1)) \cdot \mathcal{M}(F, F)(\mathcal{M}[F, F]^{-1}(M_2)) \\
&= M_1 \cdot M_2
\end{aligned} \tag{11.232}$$

hence:

$$\begin{aligned}
\det(M_1 \cdot M_2) &\underset{[\text{eq: } 11.232]}{=} \det(\mathcal{M}(F,F)(L_1 \circ L_2)) \\
&\underset{[\text{theorem: } 11.339]}{=} \det(L_1 \circ L_2) \\
&\underset{[\text{theorem: } 11.299]}{=} \det(L_1) \cdot \det(L_2) \\
&\underset{[\text{theorem: } 11.339]}{=} \det(\mathcal{M}(F,F)(L_1)) \cdot \det(\mathcal{M}(F,F)(L_2)) \\
&\underset{[\text{eq: } 11.231]}{=} \det(M_1) \cdot \det(M_2)
\end{aligned}$$

2. As $\mathcal{M}(F,F)(\mathrm{Id}_{\mathcal{F}^n}) \underset{[\text{theorem: } 11.324]}{=} E$ we have

$$\det(E) = \det(\mathcal{M}(F,F)(\mathrm{Id})) \underset{[\text{theorem: } 11.339]}{=} \det(\mathrm{Id}_{\mathcal{F}^n}) \underset{[\text{theorem: } 11.299]}{=} 1$$

3. As $E = M \cdot M^{-1}$ we have $1 \underset{(2)}{=} \det(E) = \det(M \cdot M^{-1}) = \det(M) \cdot \det(M^{-1})$ hence $\det(M) \neq 0$ and $(\det(M))^{-1} = \det(M^{-1})$. □

Theorem 11.342. *Let $n \in \mathbb{N}$, F a field of characteristic zero, $M \in \mathcal{M}_{n,n}(F)$, $E = \{e_1, \ldots, e_n\}$ the canonical basis for F^n [see theorem: 11.159] and Δ a determinant function in F^n such that $\Delta(e_1, \ldots, e_n) = 1$ then*

$$\det(M) = \Delta(\mathrm{col}(M,1), \ldots, \mathrm{col}(M,n)) = \Delta(\mathrm{row}(M,1), \ldots, \mathrm{row}(M,n))$$

Proof. Let $M \in \mathcal{M}_{n,n}(F)$ and take $L = \mathcal{M}(E,E)^{-1}(M) \in \mathrm{Hom}(F^n, F^n)$ then for $k \in \{1, \ldots, n\}$ we have that

$$\begin{aligned}
(L(e_i))_k &= (\mathcal{M}(E,E)^{-1}(M)(e_i))_k \\
&\underset{[\text{example: } 11.327}{=} \sum_{j \in \{1,\ldots,n\}} M_{k,j} \cdot (e_i)_j \\
&= \sum_{j \in \{1,\ldots,n\}} M_{k,j} \cdot \delta_{j,i} \\
&= M_{k,i} \\
&= \mathrm{col}(M,i)_k
\end{aligned} \tag{11.233}$$

So

$$\begin{aligned}
\Delta(\mathrm{col}(M,1), \ldots, \mathrm{col}(M,n)) &\underset{[\text{eq: } 11.233]}{=} \Delta(L(e_1), \ldots, L(e_n)) \\
&= \Delta_L(e_1, \ldots, e_n) \\
&\underset{[\text{definition: } 11.295]}{=} \det(L) \cdot \Delta(e_1, \ldots, e_n) \\
&= \det(L) \cdot 1 \\
&= \det(L) \\
&\underset{[\text{theorem: } 11.339]}{=} \det(\mathcal{M}(E,E)(L)) \\
&= \det(M)
\end{aligned}$$

proving that

$$\det(M) = \Delta(\mathrm{col}(M,1), \ldots, \mathrm{col}(M,n)) \tag{11.234}$$

For the rows note that $\forall i, j \in \{1, \ldots, n\}$ we have

$$\mathrm{col}(M^T, i)_j = (M^T)_{j,i} = M_{i,j} = \mathrm{row}(M,i)_j$$

so that

$$\mathrm{col}(M^T, i) = \mathrm{row}(M, i) \tag{11.235}$$

and

$$\begin{aligned}
\det(M) &\underset{[\text{theorem: } 11.338]]}{=} \det(M^T) \\
&\underset{[\text{eq: } 11.234]}{=} \Delta(\mathrm{col}(M^T,1), \ldots, \mathrm{col}(M^T,n)) \\
&\underset{[\text{eq: } 11.235]}{=} \Delta(\mathrm{row}(M,1), \ldots, \mathrm{row}(M,n))
\end{aligned}$$

giving finally

$$\Delta(\text{row}(M,1),\ldots,\text{row}(M,n)) = \det(M)$$ □

The above can be used to determine what happens to the determinant of a matrix if we permute rows or columns of the matrix.

Definition 11.343. *Let $n\in\mathbb{N}$, F a field, $M\in\mathcal{M}_{n,n}(F)$ and $\sigma\in P_n$ then:*

1. *$M_\sigma\in\mathcal{M}_{n,n}(F)$ is defined by $\forall i,j\in\{1,\ldots,n\}$ $(M_\sigma)_{i,j}=M_{\sigma(i),j}$ [essential we permute the rows of the matrix).*
2. *$M^\sigma\in\mathcal{M}_{n,n}(F)$ is defined by $\forall i,j\in\{1,\ldots,n\}$ $(M^\sigma)_{i,j}=M_{i,\sigma(j)}$ [essential we permute columns of the matrix].*

Theorem 11.344. *Let $n\in\mathbb{N}$, F a field, $M\in\mathcal{M}_{n,n}(F)$ and $\sigma\in P_n$ then:*

1. $\det(M_\sigma)=\text{sign}(\sigma)\cdot\det(M)$
2. $\det(M^\sigma)=\text{sign}(\sigma)\cdot\det(M)$

Proof. Let $E=\{e_1,\ldots,e_n\}$ be the canonical basis for F^n [see: theorem: 11.159] then by [theorem: 11.283] there exists a determinant function Δ in F^n such that

$$\Delta(e_1,\ldots,e_n)=1$$

1. $\forall i,j\in\{1,\ldots,n\}$ we have $\text{row}(M_\sigma,i)_j=(M_\sigma)_{i,j}=M_{\sigma(i),j}=\text{row}(M,\sigma(i))_j$ proving that

$$\text{row}(M_\sigma,i)=\text{row}(M,\sigma(i)) \tag{11.236}$$

so that

$$\begin{aligned}\det(M_\sigma) &\underset{[\text{theorem: }11.342]}{=} \Delta(\text{row}(M_\sigma,1),\ldots,\text{row}(M_\sigma,n))\\ &\underset{[\text{eq: }11.236]}{=} \Delta(\text{row}(M,\sigma(1)),\ldots,\text{row}(M,\sigma(i)))\\ &\underset{\text{skew}-\text{symmetry}}{=} \text{sign}(\sigma)\cdot\Delta(\text{row}(M,1),\ldots,\text{row}(M,n))\\ &\underset{[\text{theorem: }11.342]}{=} \text{sign}(\sigma)\cdot\det(M)\end{aligned}$$

2. $\forall i,j\in\{1,\ldots,n\}$ we have $\text{col}(M^\sigma,i)_j=(M^\sigma)_{j,i}=M_{j,\sigma(i)}=\text{col}(M,\sigma(i))_j$ proving that

$$\text{col}(M_\sigma,i)=\text{col}(M,\sigma(i)) \tag{11.237}$$

so that

$$\begin{aligned}\det(M^\sigma) &\underset{[\text{theorem: }11.342]}{=} \Delta(\text{col}(M_\sigma,1),\ldots,\text{col}(M_\sigma,n))\\ &\underset{[\text{eq: }11.236]}{=} \Delta(\text{col}(M,\sigma(1)),\ldots,\text{col}(M,\sigma(n)))\\ &\underset{\text{skew}-\text{symmetry}}{=} \text{sign}(\sigma)\cdot\Delta(\text{col}(M,1),\ldots,\text{col}(M,n))\\ &\underset{[\text{theorem: }11.342]}{=} \text{sign}(\sigma)\cdot\det(M)\end{aligned}$$

□

The above allows us to construct a test to determine if the determinant of a matrix is zero.

Corollary 11.345. *Let $n\in\mathbb{N}$, F a field with characteristic zero and $M\in\mathcal{M}_{n,n}(F)$ then we have*

$$\begin{gathered}\det(M)=0\\ \Updownarrow\\ \text{rank}(M)<n\end{gathered}$$

Proof.

⇒. We prove this by contradiction, so assume that $n\leqslant\text{rank}(M)$ then, as by [definition: 11.319] $\text{rank}(M)\leqslant n$, we must have that $\text{rank}(M)=n\underset{[\text{definitions:}11.317,11.319]}{=}\dim(\text{span}(\text{cols}(M)))$. Using [theorem: 11.151] it follows that $\text{cols}(M)=\{\text{col}(M,i)|i\in\{1,\ldots,n\}\}$ is a basis for $\text{span}(\text{cols}(M))\subseteq F^n$, hence $\{\text{col}(M,i)|i\in\{1,\ldots,n\}\}$ is linear independent. Using [theorem: 11.149] and the fact that $\text{card}(\text{col}(M,i)|i\in\{1,\ldots,n\})=\dim(\text{span}(\text{cols}(M)))=n$ it follows that

$$\{\text{col}(M,i)|i\in\{1,\ldots,n\}\}\text{ is a basis for }F^n$$

By [theorem: 11.283] there exists a determinant function Δ in F^n such that

$$\Delta(\text{cols}(M,1),\ldots,\text{cols}(M,n))=1$$

Using [theorem: 11.342] we have $\det(M)=\Delta(\text{cols}(M,1),...,\text{cols}(M,n))$ so that $\det(M)=1$ contradicting the hypothesis $\det(M)=0$. Hence we must have that $\text{rank}(M)<n$.

⇐. Let $\{e_1,...,e_n\}$ be the canonical basis for F^n, then by [theorems: 11.283,11.342] there exists a determinant function in F^n such that

$$\det(M)=\Delta(\text{col}(M,1),\ldots,\text{col}(M,n)) \tag{11.238}$$

By the hypothesis we have $\dim(\text{span}(\text{cols}(M)))=\text{rank}(M)<n$. Define

$$\text{col}(M)\colon\{1,\ldots,n\}\rightarrow\text{cols}(M) \text{ by } \text{col}(M)(i)=\text{col}(M,i)$$

then we have two possibilities:

col(M) is injective. Then as

$$\text{col}(M)=\{\text{col}(M)(i)|i\in\{1,\ldots,n\}\}=\{\text{col}(M,i)|i\in\{1,\ldots,n\}\}=\text{cols}(M)$$

we have that $\text{col}(M)\colon\{1,\ldots,n\}\rightarrow\text{cols}(M)$ is a bijection so that $\text{card}(\text{cols}(M))=n$. If $\text{cols}(M)$ is linear independent then it is a basis for $\text{span}(\text{cols}(M))$ so that $\dim(\text{cols}(M))=n$, contradicting $\dim(\text{span}(\text{cols}(M)))=\text{rank}(M)<n$. Hence we must have that $\text{cols}(M)=\{\text{col}(M,i)|i\in\{1,\ldots,n\}\}$ is linear dependent. So by [theorem: 11.125] $\{\text{col}(M,i)\}_{i\in\{1,\ldots,n\}}$ is linear dependent, hence by [theorem: 11.275] we have that $\Delta(\text{col}(M,1),\ldots,\text{col}(M,n))=0$, proving by [eq: 11.238] that

$$\det(M)=0$$

col(M) is not injective. Then there exists $i,j\in\{1,\ldots,n\}$ with $i\neq j$ and $\text{col}(M,i)=\text{col}(M,j)$. So that by [theorem: 11.275] we have that $\Delta(\text{col}(M,1),\ldots,\text{col}(M,n))=0$, proving by [eq: 11.238] that

$$\det(M)=0$$

□

We will now develop the necessary tools to calculate the inverse of a matrix (if the matrix is invertible).

Definition 11.346. *Let $n\in\mathbb{N}\setminus\{1\}$, X a set and $a\in X$ then we define*

$$[+a]\colon X^{n-1}\rightarrow X^n \textit{ by } [+a](x) \textit{ where } \forall i\in\{1,\ldots,n\}\ ([+a](x))_i=\begin{cases} a \textit{ if } i=1\\ x_{i-1} \textit{ if } i\in\{2,\ldots,n\}\end{cases}$$

In other words $[+a](x_1,\ldots,x_{n-1})=(a,x_1,\ldots,x_n)$

Lemma 11.347. *Let $n\in\mathbb{N}\setminus\{1\}$, F a field and $0\in F$ the additive neutral element then*

$$[+0]\in\text{Hom}(F^{n-1},F^n)$$

Proof. Let $\alpha\in F$, $x,y\in F^{n-1}$ then we have for $i\in\{1,\ldots,n\}$

$$\begin{aligned}[+0](x+y) &= \begin{cases} 0 \text{ if } i=1\\ (x+y)_{i-1} \text{ if } i\in\{2,\ldots,n\}\end{cases}\\ &= \begin{cases} 0+0 \text{ if } i=1\\ x_{i-1}+y_{i-1} \text{ if } i\in\{2,\ldots,n\}\end{cases}\\ &= [+0](x)+[+0](y)\end{aligned}$$

and

$$[+0](\alpha\cdot x)=\begin{cases} 0 \text{ if } i=1\\ (\alpha\cdot x)_{i-1} \text{ if } i\in\{2,\ldots,n+1\}\end{cases}=\begin{cases} \alpha\cdot 0 \text{ if } i=1\\ \alpha\cdot x_{i-1} \text{ if } i\in\{2,\ldots,n+1\}\end{cases}=\alpha\cdot([+0](x))$$

□

We use the above now to generate a new determinant function from a existing determinant function.

Theorem 11.348. *Let $n\in\mathbb{N}$, F a field with characteristic zero, $E=\{e_1,\ldots,e_n\}$ the canonical basis for F^n [see: theorem: 11.159] then if*

$$\Delta\colon(F^n)^n\rightarrow F$$

is a determinant function and we define

$$\Delta^-: (F^{n-1})^{n-1} \to F \textit{ where } \Delta^-(x_1, \ldots, x_{n-1}) = \Delta(e_1, [+0](x_1), \ldots, [+0](x_{n-1}))$$

we have that

1. *Δ^- is a determinant function.*
2. *If $\Delta(e_1, \ldots, e_n) = 1$ then $\Delta^-(e_1, \ldots, e_{n-1}) = 1$*

Proof.

1. First we prove multilinearity. Let $x, y \in F^{n-1}$ and $\alpha \in F$ then for

$$\left(a_1, \ldots \underbrace{x + \alpha \cdot y}_{i} \ldots, a_{n-1}\right) \in (F^{n-1})^{n-1}$$

we have

$$\begin{aligned}
\Delta^-\left(a_1, \ldots \underbrace{x + \alpha \cdot y,}_{i} \ldots, a_{n-1}\right) &= \\
\Delta\left(e_1, [+0](a_1), \ldots \underbrace{[+0](x + \alpha \cdot y)}_{i} \ldots, [+0](a_n)\right) &= \\
\Delta\left(e_1, [+0](a_1), \ldots \underbrace{[+0](x) + \alpha \cdot [+0](y)}_{i} \ldots, [+0](a_n)\right) &= \\
\Delta\left(e_1, [+0](a_1), \ldots \underbrace{[+0](x)}_{i} \ldots, [+0](a_n)\right) + \alpha \cdot \Delta\left(e_1, [+0](a_1), \ldots \underbrace{[+0](y)}_{i} \ldots, [+0](a_n)\right) &= \\
\Delta^-\left(a_1, \ldots \underbrace{x}_{i} \ldots, a_{n-1}\right) + \alpha \cdot \Delta^-\left(a_1, \ldots \underbrace{y}_{i} \ldots, a_{n-1}\right)
\end{aligned}$$

Next we prove skew-symmetry. Let $\{a_i\}_{i \in \{1, \ldots, n-1\}} \subseteq F^{n-1}$ such that there exists a $i \neq j \in \{1, \ldots, n-1\}$ with $a_i = a_j$ then $\forall k \in \{1, \ldots, n\}$ we have

$$([+0](a_i))_k = \begin{cases} 0 \text{ if k=1} \\ (a_i)_{k-1} \text{ if } k \in \{2, \ldots, n\} \end{cases} = \begin{cases} 0 \text{ if k=1} \\ (a_j)_{k-1} \text{ if } k \in \{2, \ldots, n\} \end{cases} = ([+0](a_j))_k$$

proving that

$$[+0](a_i) = [+0](a_j)$$

Hence we have

$$0 = \Delta(e_1, [+0](a_1), \ldots, [+0](a_n)) = \Delta^-(a_1, \ldots, a_{i-1})$$

proving by [theorem: 11.275] that

$$\Delta^- \text{ is skew-symmetric}$$

2. Assume that $\Delta(e_1, \ldots, e_n) = 1$. Let $i, j \in \{1, \ldots, n-1\}$ then

$$\begin{aligned}
([+0](e_i))_j &= \begin{cases} 0 \text{ if } j = 1 \\ (e_i)_{j-1} \text{ if } i \in \{2, \ldots, n\} \end{cases} \\
&= \begin{cases} 0 \text{ if } j = 1 \\ \delta_{i, j-1} \text{ if } i \in \{2, \ldots, n\} \end{cases} \\
&\underset{i+1=j \Rightarrow i=j-1}{=} \delta_{i+1, j} \\
&= (e_{i+1})_j
\end{aligned}$$

proving that $[+0](e_i) = e_{i+1}$ so that

$$\Delta^-(e_1, \ldots, e_{n-1}) = \Delta(e_1, [+0](e_1), \ldots, [+0](e_{n-1})) = \Delta(e_1, e_2, \ldots, e_n) = 1$$

□

It will be useful to decrease the size of a matrix, for this we introduce the $[n \boxplus m]$ function that removes a row and a column from a matrix.

Definition 11.349. *Let $n \in \mathbb{N} \setminus \{1\} = \{2, \ldots, \infty\}$, F a field and $i, j \in \{1, \ldots, n-1\}$ then we define:*

$$[i \boxplus j]: \mathcal{M}_{n,n}(F) \rightarrow \mathcal{M}_{n-1,n-1}(F) \text{ where } ([i \boxplus j](M))_{k,l} = \begin{cases} M_{k,l} \text{ if } 1 \leqslant k < i \wedge 1 \leqslant l < j \\ M_{k+1,l} \text{ if } i \leqslant k \leqslant n-1 \wedge 1 \leqslant l < j \\ M_{k,l+1} \text{ if } 1 \leqslant k < i \wedge j \leqslant l \leqslant n-1 \\ M_{k+1,l+1} \text{ if } i \leqslant k \leqslant n-1 \wedge j \leqslant l \leqslant n-1 \end{cases}$$

In other words $[i \boxplus j](M)$ is the matrix that you get if you remove the i-the row and the j-the column of M.

Example 11.350.

$$[2 \boxplus 3]\begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 10 & 20 & 30 & 40 \\ 50 & 60 & 70 & 80 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 4 \\ 10 & 20 & 40 \\ 50 & 60 & 80 \end{pmatrix}$$

We show not that removing the i-the row and j-the column of the transpose of a matrix is the same as transposing the result of removing the j-the row and i-the column from the matrix.

Theorem 11.351. *Let $n \in \mathbb{N} \setminus \{1\}$, F a field, $i, j \in \{1, \ldots, n\}$ and $M \in \mathcal{M}_{n,n}(F)$ then*

$$[i \boxplus j](M^T) = ([j \boxplus i]M)^T$$

Proof. This is easterly proved by considering all the possible cases, so if $k, l \in \{1, \ldots, n-1\}$ then we have either:

$\boldsymbol{1 \leqslant k < i \wedge 1 \leqslant l < j}$**.** Then

$$\begin{aligned} ([i \boxplus j](M^T))_{k,l} &= (M^T)_{k,l} \\ &= M_{l.k} \\ &= ([j \boxplus i]M)_{l.k} \\ &= (([j \boxplus i]M)^T)_{k,l} \end{aligned}$$

$\boldsymbol{i \leqslant k \leqslant n-1 \wedge 1 \leqslant l < j}$**.** Then

$$\begin{aligned} ([i \boxplus j](M^T))_{k,l} &= (M^T)_{k+1,l} \\ &= M_{l,k+1} \\ &= ([j \boxplus i]M)_{l,k} \\ &= (([j \boxplus i]M)^T)_{k,l} \end{aligned}$$

$\boldsymbol{1 \leqslant k < i \wedge j \leqslant l \leqslant n-1}$**.** Then

$$\begin{aligned} ([i \boxplus j](M^T))_{k,l} &= (M^T)_{k,l+1} \\ &= M_{l+1,k} \\ &= ([j \boxplus i]M)_{l,k} \\ &= (([j \boxplus i]M)^T)_{k,l} \end{aligned}$$

$\boldsymbol{i \leqslant k \leqslant n-1 \wedge j \leqslant l \leqslant n-1}$**.** Then

$$\begin{aligned} ([i \boxplus j](M^T))_{k,l} &= (M^T)_{k+1.l+1} \\ &= M_{l+1,k+1} \\ &= ([j \boxplus i]M)_{l,k} \\ &= (([j \boxplus i]M)^T)_{k,l} \end{aligned}$$

so that

$$[i \boxplus j](M^T) = ([j \boxplus i]M)^T$$

□

Removing a row and column of a matrix reduces the number of rows and columns with one, allowing us later in this book to express the calculation of the determinant of a matrix by calculating the determinant of a sub matrix. We can also reduce the size of a matrix by removing all the rows or columns after a specified row or specified column. This is the purpose of the following definition.

Definition 11.352. *Let $n\in\mathbb{N}\setminus\{1\}$, F a field and $M\in\mathcal{M}_{n,n}(F)$ then we define*

1. *If $m\in\{2,\ldots,n\}$ then $[<m](M)\in\mathcal{M}_{m-1,m-1}$ by $\forall k,l\in\{1,\ldots,m-1\}$*
$$([<m](M))_{k.l}=M_{k.l}$$
2. *If $m\in\{1,\ldots,n-1\}$ then $[>m](M)\in\mathcal{M}_{n-m,n-m}(F)$ by $\forall k,l\in\{1,\ldots,n-m\}$*
$$([>m](M))_{k,l}=M_{k+m,l+m}$$

We have now the following properties for $[<m]$ and $[>m]$.

Theorem 11.353. *Let $n\in\mathbb{N}\setminus\{1\}$, F a field then we have for $M\in\mathcal{M}_{n,n}(F)$*

1. $[1\boxplus 1](M)=[>1](M)$
2. $[n\boxplus n](M)=[<n](M)$
3. *If $m_1\in\{2,\ldots,n\}$ and $m_2\in\{2,\ldots,m_1-1\}$ then $[<m_2]([<m_1](M))=[<m_2](M)$*
4. *If $m_1\in\{1,\ldots,n-1\}$ and $m_2\in\{1,\ldots,n-m_1-1\}$ then $[>m_2]([>m_1](M))=[>(m_1+m_2)](M)$*
5. *If $m\in\{2,\ldots,n\}$ then $([<m](M))^T=[<m](M^T)$*
6. *If $m\in\{1,\ldots,n-1\}$ then $([>m](M))^T=[>m](M^T)$*

Proof.

1. $\forall i,j\in\{1,\ldots,n-1\}$ we have
$$\begin{aligned}([1\boxplus 1](M))_{i,j} &\underset{1\leqslant i,j\leqslant n-1\wedge[\text{definition: } 11.349]}{=} M_{i+1,j+1}\\ &= ([>1]M)_{i,j}\end{aligned}$$
so we have $[1\boxplus 1](M)=[>1]M$
2. $\forall i,j\in\{1,\ldots,n-1\}$ we have
$$\begin{aligned}([n\boxplus n](M))_{i,j} &\underset{1\leqslant i,j\leqslant n-1<n\wedge[\text{definition: } 11.349]}{=} M_{i,j}\\ &\underset{1\leqslant i,j\leqslant n-1}{=} ([<n](M))_{i,j}\end{aligned}$$
so we have $[n\boxplus n](M)=[<n](M)$
3. If $i,j\in\{1,\ldots,m_2-1\}$ then we have
$$\begin{aligned}([<m_2]([<m_1](M)))_{i,j} &\underset{1\leqslant i,j\leqslant m_2-1}{=} ([<m_1](M))_{i,j}\\ &\underset{1\leqslant i,j\leqslant m_2-1<m_2\leqslant m_1-1}{=} M_{i,j}\\ &\underset{1\leqslant i,j\leqslant m_2-1}{=} ([<m_2]M)_{i,j}\end{aligned}$$
so we have $[<m_2]([<m_1](M))=[<m_2](M)$
4. First note that
$$\begin{aligned}[>m_1](M)\in\mathcal{M}_{n-m_1,n-m_1}(F) &\Rightarrow [>m_2]([>m_1](M))\in\mathcal{M}_{n-m_1-m_2,n-m_1-m_2}(F)\\ &\Rightarrow [>m_2]([>m_1](M))\in\mathcal{M}_{n-(m_1+m_2),n-(m_1+m_2)}(F)\end{aligned}$$
and
$$[>(m_1+m_2)](M)\in\mathcal{M}_{n-(m_1+m_2),n-(m_1+m_2)}(F)$$
So that
$$[>(m_1+m_2)](M),\ [>m_2]([>m_1](M))\in\mathcal{M}_{n-(m_1+m_2),n-(m_1+m_2)}(F)$$
Next $\forall i,j\in\{1,\ldots,n-(m_1+m_2)\}$ we have
$$\begin{aligned}([>m_2]([>m_1](M)))_{i,j} &= ([>m_1](M))_{i+m_2,j+m_2}\\ &= M_{i+m_2+m_1,j+m_2+m_1}\\ &= M_{i+(m_1+m_2),j+(m_1+m_2)}\\ &= ([>(m_1+m_2)](M))_{i,j}\end{aligned}$$

proving

$$[>m_2]([>m_1](M)) = [>(m_1+m_2)](M)$$

5. For $\forall i,j \in \{1,\ldots,m-1\}$ we have

$$\begin{array}{rcl}(([<m](M))^T)_{i,j} &=& ([<m](M))_{j,i}\\ &=& M_{j,i}\\ &=& (M^T)_{i,j}\\ &=& ([<m](M^T))_{i,j}\end{array}$$

so

$$([<m](M))^T = [<m](M^T)$$

6. For $\forall i,j \in \{1,\ldots,n-m\}$ we have

$$\begin{array}{rcl}(([>m](M))^T)_{i,j} &=& ([>m](M))_{j,i}\\ &=& M_{j+m,i+m}\\ &=& (M^T)_{i+m,j+m}\\ &=& ([>m](M^T))_{i,j}\end{array}$$

so

$$([>m](M))^T = [>m](M^T)$$

□

We want now to calculate the determinant of a $n \times n$ matrix in terms of a $(n-1)\times(n-1)$ sub matrix. More specific we want to prove that

$$\left|\begin{array}{cccccccc} m_{1,1} & \ldots & m_{1,i-1} & 0 & \ldots & m_{1,i+1} & \ldots & m_{1,n}\\ & & & \vdots & & & & \vdots\\ m_{j-1,1} & & & 0 & & & & \vdots\\ 0 & 0 & \ldots & 1 & 0 & \ldots & 0 & 0\\ m_{j+1,1} & & & 0 & & & & \vdots\\ \vdots & & & \vdots & & & & \vdots\\ m_{n,1} & \ldots & \ldots & 0 & \ldots & \ldots & & m_{n,n} \end{array}\right|$$

is equal to

$$(-1)^{i+j}\cdot\left|\begin{array}{cccccc} m_{1,1} & \ldots & m_{1,i-1} & m_{1,i+1} & \ldots & m_{1,n}\\ \vdots & & \vdots & \vdots & & \vdots\\ m_{j-1,1} & \ldots & m_{j-1,i-1} & m_{j-1,i+1} & \ldots & m_{j-1,n}\\ m_{j+1,1} & \ldots & m_{j+1,i-1} & m_{j+1}, i+1 & \ldots & m_{j+1,n}\\ \vdots & & \vdots & \vdots & & \vdots\\ m_{n,1} & \ldots & m_{n,i-1} & m_{n,i+1} & \ldots & m_{n,n} \end{array}\right|$$

First we prove that

$$\left|\begin{array}{cccc} 1 & 0 & \ldots & 0\\ 0 & m_{2,2} & \ldots & m_{2,n}\\ \vdots & \vdots & \ddots & \vdots\\ 0 & m_{n,2} & \ldots & m_{n,n} \end{array}\right| = \left|\begin{array}{ccc} m_{2,2} & \ldots & m_{2,n}\\ \vdots & \ddots & \vdots\\ m_{n,2} & \ldots & m_{n,n} \end{array}\right|$$

Lemma 11.354. *Let $n \in \mathbb{N}\setminus\{1\}$ and $M \in M_{n,n}(F)$ such that $\forall i \in \{1,\ldots,n\}$*

$$\mathrm{row}(M,1)_i = \delta_{i,1} = \mathrm{col}(M,1)_i$$

then

$$\det(M) = \det([1 \boxplus 1](M))$$

Proof. Let $\{e_1,\ldots,e_n)$ be the canonical basis of F^n. Let $i \in \{1,\ldots,n\}$ then we have either:

$\boldsymbol{i=1.}$ Then we have $\forall j \in \{1,\ldots,n\}$ that $\mathrm{row}(M,1)_j = \delta_{i,1} = (e_1)_j$ so that

$$\mathrm{row}(M,1) = e_1 \tag{11.239}$$

$i \in \{2, \ldots, n\}$. Then for $j \in \{1, \ldots, n\}$ we have either:

$j = 1$. Then

$$\begin{aligned}([+0](\mathrm{row}([1 \boxplus 1](M), i)))_j &= ([+0](\mathrm{row}([1 \boxplus 1](M), i)))_1 \\ &\underset{[\text{definition: } 11.346]}{=} 0 \\ &\underset{i+1 \neq 1}{=} \delta_{i+1,1} \\ &= \mathrm{col}(M, 1)_{i+1} \\ &= M_{i+1,1} \\ &= \mathrm{row}(M, i+1)_1 \\ &= \mathrm{row}(M, i+1)_j\end{aligned}$$

$j \in \{2, \ldots, n\}$. Then $1 \leqslant j - 1 \leqslant n \wedge 1 \leqslant i \leqslant n$

$$\begin{aligned}([+0](\mathrm{row}([1 \boxplus 1](M), i)))_j &\underset{[\text{definition: } 11.346] \wedge j \in \{2, \ldots, n\}}{=} \mathrm{row}([1 \boxplus 1](M), i)_{j-1} \\ &\underset{[\text{definition: } 11.349]}{=} ([1 \boxplus 1](M))_{i,j-1} \\ &= M_{i+1,j} \\ &= \mathrm{row}(M, i+1)_j\end{aligned}$$

So we have that

$$\mathrm{row}(M, i+1) = [+0](\mathrm{row}([1 \boxplus 1](M), i)) \tag{11.240}$$

Let Δ be a determinant function such that $\Delta(e_1, \ldots, e_n) = 1$ [see theorem: 11.283] then we have

$$\begin{aligned}\det(M) &\underset{[\text{theorem: } 11.342]}{=} \Delta(\mathrm{row}(M, 1), \mathrm{row}(M, 2), \ldots, \mathrm{row}(M, n)) \\ &\underset{[\text{eq: } 11.239]}{=} \Delta(e_1, \mathrm{row}(M, 2), \ldots, \mathrm{row}(M, n)) \\ &\underset{[\text{eq: } 11.240]}{=} \Delta(e_1, [+0](\mathrm{row}([1 \boxplus 1](M), 1)), \ldots, [+0](\mathrm{row}([1 \boxplus 1](M), n-1))) \\ &\underset{[\text{definition: } 11.348]}{=} \Delta^-(\mathrm{row}([1 \boxplus 1](M), 1), \ldots, \mathrm{row}([1 \boxplus 1](M), n-1))\end{aligned} \tag{11.241}$$

As Δ is a determinant function with $\Delta(e_1, \ldots, e_n) = 1$ we have by [theorem: 11.348] that

$$\Delta^- \text{ is a determinant function and } \Delta^-(e_1, \ldots, e_{n-1}) = 1$$

which as $\{e_1, \ldots, e_{n-1}\}$ is the canonical basis of F^{n-1} proves by [theorem: 11.342] that

$$\Delta^-(\mathrm{row}([1 \boxplus 1](M), 1), \ldots, \mathrm{row}([1 \boxplus](M), n-1)) = \det([1 \boxplus 1](M))$$

Substituting the above in [eq: 11.241] proves then

$$\det(M) = \det([1 \boxplus 1](M))$$

□

We extend now the above lemma to a more general case:

$$\begin{vmatrix} 0 & \ldots & 0 & 1 & 0 & \ldots & 0 \\ m_{2,1} & & m_{2,j-1} & 0 & m_{2,j+1} & \ldots & m_{2,n} \\ \vdots & & & \vdots & & & \vdots \\ m_{n,1} & \ldots & m_{n,j-1} & 0 & m_{n,j+1} & \ldots & m_{n,n} \end{vmatrix}$$

is equal to

$$(-1)^{1+j} \cdot \begin{vmatrix} m_{2,1} & \ldots & m_{2,n} \\ \vdots & \ddots & \vdots \\ m_{2,n} & \ldots & m_{n,n} \end{vmatrix}$$

Theorem 11.355. *Let $n \in \mathbb{N} \setminus \{1\}$, F a field with characteristic zero and $M \in \mathcal{M}_{n,n}(F)$ satisfying $\exists j \in \{1, \ldots, n\}$ such that $\forall i \in \{1, \ldots, n\}$ we have $\mathrm{row}(M, 1)_i = \delta_{i,j}$ and $\mathrm{col}(M, j)_i = \delta_{i,1}$ then*

$$\det(M) = (-1)^{1+j} \cdot \det([1 \boxplus j](M))$$

Proof. For j we have the following cases:

$\boldsymbol{j=1}$. Then $\forall i\in\{1,\ldots,n\}$ we have $\mathrm{row}(M,1)_i=\delta_{i,1}=\mathrm{col}(M,1)_i$ so that the conditions of the previous lemma [lemma: 11.354] are satisfied. Hence $\det(M)=\det([1\boxplus 1](M))$ which, as $j=1$ proves

$$\det(M)=(-1)^{1+j}\cdot\det([1\boxplus j](M))$$

$\boldsymbol{j\in\{2,\ldots,n\}}$. Define M' by moving the j-the column before the first column [see definition: 11.236], in other words M' is defined by

$$\forall k,l\in\{1,\ldots,n\}\; M'_{k,l}=M_{k,(j\underset{n}{\rightsquigarrow}1)(l)}$$

then using [definition: 11.343] we have that

$$M'=M^{(j\underset{n}{\rightsquigarrow}1)} \tag{11.242}$$

Further $\forall k\in\{1,\ldots n\}$ we have

$$\begin{aligned}
\mathrm{row}(M',1)_k &= M'_{1,k}\\
&= M_{1,(j\underset{n}{\rightsquigarrow}1)(k)}\\
&\underset{1<j\wedge[\text{definition: }11.236(3)]}{=} \begin{cases} M_{1,k} \text{ if } 1\leqslant k<1\\ M_{1,j} \text{ if } k=1\\ M_{1,k-1} \text{ if } 1<k\leqslant j\\ M_{1,k} \text{ if } j<k\leqslant n\end{cases}\\
&= \begin{cases} M_{1,j} \text{ if } k=1\\ M_{1,k-1} \text{ if } 1<k\leqslant j\\ M_{1,k} \text{ if } j<k\leqslant n\end{cases}\\
&= \begin{cases} \mathrm{col}(M,j)_1 \text{ if } k=1\\ \mathrm{row}(M,1)_{k-1} \text{ if } 1<k\leqslant j\\ \mathrm{row}(M,1)_k \text{ if } j<k\leqslant n\end{cases}\\
&= \begin{cases} \delta_{k,1} \text{ if } k=1\\ \delta_{k-1,j} \text{ if } 1<k\leqslant j\\ \delta_{k,1} \text{ if } j<k\leqslant n\end{cases}\\
&= \begin{cases} 1 \text{ if } k=1\\ \delta_{k-1,j} \text{ if } 1<k\leqslant j\\ \delta_{k,j} \text{ if } j<k\leqslant n\end{cases}\\
&\underset{k\leqslant j\Rightarrow k-1<j\Rightarrow k-1\neq j}{=} \begin{cases} 1 \text{ if } k=1\\ 0 \text{ if } 1<k\leqslant j\\ \delta_{k,j} \text{ if } j<k\leqslant n\end{cases}\\
&\underset{j<k\Rightarrow j\neq k}{=} \begin{cases} 1 \text{ if } k=1\\ 0 \text{ if } 1<k\leqslant j\\ 0 \text{ if } j<k\leqslant n\end{cases}\\
&= \delta_{k,1}
\end{aligned}$$

and we have also

$$\begin{aligned}
\mathrm{col}(M',1)_k &= M'_{k,1}\\
&= M_{k,(j\underset{n}{\rightsquigarrow}1)(1)}\\
&\underset{1=1\wedge[\text{definition: }11.236(1)]}{=} M_{k,j}\\
&= \mathrm{col}(M,j)_k\\
&= \delta_{k,1}
\end{aligned}$$

Hence M' satisfies the conditions of the previous lemma [lemma: 11.354] giving us

$$\det(M')=\det([1\boxplus 1](M')) \tag{11.243}$$

Now if $k,l\in\{1,\ldots,n-1\}$ then $1\leqslant k,l\leqslant n-1$ it follows from [definition: 11.349] that

$$([1\boxplus 1](M'))_{k,l}=M'_{k+1,l+1} \tag{11.244}$$

Take $k\in\{1,\ldots,n-1\}$ then for $l\in\{1,\ldots,n-1\}$ we have the following possible cases:

$1\leqslant l<j$. Then

$$\begin{array}{lll} M'_{k+1,l+1} & = & M_{k+1,(j\underset{n}{\rightsquigarrow}1)(l+1)} \\ & \underset{l<j\Rightarrow 1<l+1\leqslant j\wedge[\text{definition: } 11.236(3)]}{=} & M_{k+1,l} \\ & \underset{1\leqslant k\leqslant n-1, l<j\wedge[\text{definition: } 11.349]}{=} & ([1\boxplus j](M))_{k,l} \end{array}$$

$j\leqslant l\leqslant n-1$. Then

$$\begin{array}{lll} M'_{k+1,l+1} & = & M_{k+1,(j\underset{n}{\rightsquigarrow}1)(l+1)} \\ & \underset{j\leqslant l\Rightarrow j<l+1\leqslant n\wedge[\text{definition: } 11.236(3)]}{=} & M_{k+1,l+1} \\ & \underset{1\leqslant k\leqslant n-1\wedge j\leqslant l\leqslant n-1\wedge[\text{definition: } 11.349]}{=} & ([1\boxplus j](M))_{k,l} \end{array}$$

So in all cases we gave $M'_{k+1,l+1}=([1\boxplus j](M))_{k,l}$, combining this with [eq: 11.244] results in $([1\boxplus 1](M'))_{k,l}=([1\boxplus j](M))_{k,l}$. Hence we have $[1\boxplus 1](M')=[1\boxplus j](M)$. Substituting this result in [eq: 11.243] gives

$$\det(M')=\det([1\boxplus j](M)) \tag{11.245}$$

Further

$$\begin{array}{lll} \det(M') & \underset{[\text{eq: } 11.242]}{=} & \det\left(M^{(j\underset{n}{\rightsquigarrow}1)}\right) \\ & \underset{[\text{theorem: } 11.344]}{=} & \operatorname{sign}\left(\left(j\underset{n}{\rightsquigarrow}1\right)\right)\cdot\det(M) \\ & \underset{[\text{theorem: } 11.239]}{=} & (-1)^{j-1}\cdot\det(M) \end{array}$$

Hence by [theorem: 11.265]

$$\det(M)=(-1)^{j+1}\cdot\det(M')\underset{[\text{eq: } 11.245]}{=}(-1)^{1+j}\cdot\det([1\boxplus j](M))$$

proving the lemma. □

Finally we prove that

$$\begin{vmatrix} m_{1,1} & \ldots & m_{1,i-1} & 0 & \ldots & m_{1,i+1} & \ldots & m_{1,n} \\ & & & \vdots & & & & \vdots \\ m_{j-1,1} & & & 0 & & & & \vdots \\ 0 & 0 & \ldots & 1 & 0 & \ldots & 0 & 0 \\ m_{j+1,1} & & & 0 & & & & \vdots \\ \vdots & & & \vdots & & & & \vdots \\ m_{n,1} & \ldots & \ldots & 0 & \ldots & \ldots & & m_{n,n} \end{vmatrix}$$

is equal to

$$(-1)^{i+j}\cdot\begin{vmatrix} m_{1,1} & \ldots & m_{1,i-1} & m_{1,i+1} & \ldots & m_{1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ m_{j-1,1} & \ldots & m_{j-1,i-1} & m_{j-1,i+1} & \ldots & m_{j-1,n} \\ m_{j+1,1} & \ldots & m_{j+1,i-1} & m_{j+1},i+1 & \ldots & m_{j+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ m_{n,1} & \ldots & m_{n,i-1} & m_{n,i+1} & \ldots & m_{n,n} \end{vmatrix}$$

Theorem 11.356. *Let $n\in\mathbb{N}\setminus\{1\}$, F a field with characteristic zero and $M\in\mathcal{M}_{n,n}(F)$ satisfying $\exists i,j\in\{1,\ldots,n\}$ such that $\forall k\in\{1,\ldots,n\}$ we have $\operatorname{row}(M,i)_k=\delta_{j,k}$ and $\operatorname{col}(M,j)_k=\delta_{i,k}$ then*

$$\det(M)=(-1)^{i+j}\cdot\det([i\boxplus j](M))$$

Proof. For $i \in \{1, \ldots, n\}$ we have the following cases to consider:

$i = 1$. Then $\forall k \in \{1, \ldots, n\}$ we have $\text{row}(M, 1)_k = \delta_{j,k}$ and $\text{col}(M, j)_k = \delta_{i,k}$ so that the conditions for the previous lemma [lemma: 11.355] are satisfied, hence we have that

$$\det(M) = (-1)^{1+j} \cdot \det([1 \boxplus j])$$

which as $i = 1$ proves that

$$\det(M) = (-1)^{i+j} \cdot \det([i \boxplus j](M))$$

$i \in \{2, \ldots, n\}$. We move now the i-the row before the first row, so define M' by

$$M' = M_{(i \underset{n}{\rightsquigarrow} 1)} \tag{11.246}$$

or using [definition: 11.343] we have

$$\forall k, l \in \{1, \ldots, n\} \text{ that } M'_{k,l} = M_{(i \underset{n}{\rightsquigarrow} 1)(k),l} \tag{11.247}$$

Let $k \in \{1, \ldots, n\}$ then we have

$$\begin{aligned}
\text{row}(M', 1)_k &= M'_{1,k} \\
&= M_{(i \rightsquigarrow 1_n)(1),k} \\
&\underset{1=1<k \wedge [\text{definition: } 11.236(3)]}{=} M_{i,k} \\
&= \text{row}(M, i)_k \\
&= \delta_{j,k}
\end{aligned}$$

and

$$\begin{aligned}
\text{col}(M', j)_k &= M'_{k,j} \\
&= M_{(i \rightsquigarrow 1_n)(k),j} \\
&\underset{[\text{definition: } 11.236(3)]}{=} \begin{cases} M_{k,j} \text{ if } 1 \leqslant k < 1 \\ M_{i,j} \text{ if } k = 1 \\ M_{k-1,j} \text{ if } 1 \leqslant k \leqslant i \\ M_{k,j} \text{ if } i < k \leqslant \text{nn} \end{cases} \\
&= \begin{cases} M_{i,j} \text{ if } k = 1 \\ M_{k-1,j} \text{ if } 1 \leqslant k \leqslant i \\ M_{k,j} \text{ if } i < k \leqslant n \end{cases} \\
&= \begin{cases} \text{row}(M, i)_j \text{ if } k = 1 \\ \text{col}(M, j)_{k-1} \text{ if } 1 \leqslant k \leqslant i \\ \text{col}(M, j)_k \text{ if } i < k \leqslant n \end{cases} \\
&= \begin{cases} \delta_{j,j} \text{ if } k = 1 \\ \delta_{i,k-1} \text{ if } 1 \leqslant k \leqslant i \\ \delta_{i,k} \text{ if } i < k \leqslant n \end{cases} \\
&= \begin{cases} 1 \text{ if } k = 1 \\ \delta_{i,k-1} \text{ if } 1 \leqslant k \leqslant i \\ \delta_{i,k} \text{ if } i < k \leqslant n \end{cases} \\
&\underset{k \leqslant i \Rightarrow k-1 < i \Rightarrow k-1 \neq i}{=} \begin{cases} 1 \text{ if } k = 1 \\ 0 \text{ if } 1 \leqslant k \leqslant i \\ \delta_{i,k} \text{ if } i < k \leqslant n \end{cases} \\
&\underset{i < k \Rightarrow i \neq k}{=} \begin{cases} 1 \text{ if } k = 1 \\ 0 \text{ if } 1 \leqslant k \leqslant i \\ 0 \text{ if } i < k \leqslant n \end{cases} \\
&= \delta_{1,k}
\end{aligned}$$

So the conditions for the previous lemma [lemma: 11.355] are satisfied for M', hence

$$\det(M') = (-1)^{1+j} \cdot \det([1 \boxplus j](M')) \tag{11.248}$$

Now if $k, l \in \{1, \ldots, n-1\}$ we have the following possibilities for k, l:

$1 \leqslant k < i \wedge 1 \leqslant l < j$. Then

$$\begin{aligned} ([1 \boxplus j](M'))_{k,l} &\underset{1 \leqslant k \leqslant n-1 \wedge 1 \leqslant l < j \wedge [\text{definition: } 11.349]}{=} M'_{k+1,l} \\ &= M_{(i \underset{n}{\rightsquigarrow} 1)(k+1),l} \\ &\underset{1 \leqslant k < i \Rightarrow 1 < k+1 \leqslant i \wedge [\text{theorem: } 11.236(3)]}{=} M_{k,l} \\ &\underset{1 \leqslant k \leqslant i \wedge 1 \leqslant l < j \wedge [\text{definition: } 11.349]}{=} ([i \boxplus j](M))_{k.l} \end{aligned}$$

$i \leqslant k \leqslant n-1 \wedge 1 \leqslant l < j$. Then

$$\begin{aligned} ([1 \boxplus j](M'))_{k,l} &\underset{1 \leqslant k \leqslant n-1 \wedge 1 \leqslant l < j \wedge [\text{definition: } 11.349]}{=} M'_{k+1.l} \\ &= M_{(i \underset{n}{\rightsquigarrow} 1)(k+1),l} \\ &\underset{i \leqslant k < n-1 \Rightarrow i < k+1 \leqslant n \wedge [\text{theorem: } 11.236(3)]}{=} M_{k+1,l} \\ &\underset{i \leqslant k \leqslant n-1 \wedge 1 \leqslant l < j \wedge [\text{definition: } 11.349]}{=} ([i \boxplus j](M))_{k,l} \end{aligned}$$

$1 \leqslant k < i \wedge j \leqslant l < n-1$. Then

$$\begin{aligned} ([1 \boxplus j](M'))_{k,l} &\underset{1 \leqslant k \leqslant n-1 \wedge j \leqslant l < n-1 \wedge [\text{definition: } 11.349]}{=} M'_{k+1,l+1} \\ &= M_{(i \underset{n}{\rightsquigarrow} 1)(k+1),l+1} \\ &\underset{1 \leqslant k < i \Rightarrow 1 < k+1 \leqslant n \wedge [\text{theorem: } 11.236(3)]}{=} M_{k,l+1} \\ &\underset{1 \leqslant k < i \wedge j \leqslant l < n-1 \wedge [\text{definition: } 11.349]}{=} ([i \boxplus j](M))_{k,l} \end{aligned}$$

$i \leqslant k \leqslant n-1 \wedge j \leqslant l \leqslant n-1$. Then

$$\begin{aligned} ([1 \boxplus j](M'))_{k,l} &\underset{1 \leqslant k \leqslant n-1 \wedge j \leqslant l < n-1 \wedge [\text{definition: } 11.349]}{=} M'_{k+1,l+1} \\ &= M_{(i \underset{n}{\rightsquigarrow} 1)(k+1),l+1} \\ &\underset{i \leqslant k < n-1 \Rightarrow i < k+1 \leqslant n \wedge [\text{theorem: } 11.236(3)]}{=} M_{k+1,l+1} \\ &\underset{i \leqslant k \leqslant n-1 \wedge j \leqslant l < n-1 \wedge [\text{definition: } 11.349]}{=} ([i \boxplus j](M))_{k,l} \end{aligned}$$

So we have proved that $\forall k, l \in \{1, \ldots, n-1\}$ $([1 \boxplus j](M'))_{k.l} = ([i \boxplus j](M))_{k,l}$ giving

$$[1 \boxplus j](M') = [i \boxplus j](M)$$

Substituting the above in [eq: 11.248] gives then

$$\det(M') = (-1)^{1+j} \cdot \det([i \boxplus j](M)) \tag{11.249}$$

Now

$$\begin{aligned} \det(M') &\underset{[\text{eq: } 11.246]}{=} \det\big(M_{(i \underset{n}{\rightsquigarrow} 1)}\big) \\ &\underset{[\text{theorem: } 11.344]}{=} \operatorname{sign}\big(\big(i \underset{n}{\rightsquigarrow} 1\big)\big) \cdot \det(M) \\ &\underset{[\text{theorem: } 11.239]}{=} (-1)^{i-1} \cdot \det(M) \end{aligned}$$

Substituting the above in [eq: 11.249] gives

$$(-1)^{i-1} \cdot \det(M) = (-1)^{1+j} \cdot \det([i \boxplus j](M)).$$

Multiplying both sides by $(-1)^{i-1}$ and using the fact that $(-1)^{i-1}\cdot(-1)^{i-1}=1$ and $(-1)^{i-1}\cdot(-1)^{j+1}=(-1)^{i+j}$ we have finally that

$$\det(M)=(-1)^{i+j}\cdot\det([i\boxplus j](M))$$

□

For linear transformations we can calculate the inverse of a linear transformation using the determinant and adjoint of a linear transformation. We have introduced the determinant of a matrix and shows its relation with the determinant of the associated linear transformation. Now we do the same for the adjoint, so we will define the adjoint of a matrix and show its relation with the associated linear transformation.

Definition 11.357. *Let $n,m\in\mathbb{N}$, F a field and $M\in\mathcal{M}_{n,m}(F)$ then for $i,j\in\{1,\ldots,n\}$ $[i\oplus j](M)$ is defined by*

$$\forall k,l\in\{1,\ldots,n\} \text{ we have } ([i\oplus j](M))_{k,l}=\begin{cases} M_{k,l} \text{ if } l\in\{1,\ldots,n\}\setminus\{i\}\\ \delta_{k,j} \text{ if } l=i\end{cases}$$

In other words we replace the i-the column by the column $\begin{pmatrix}\delta_{1,j}\\ \ldots\\ \delta_{n,j}\end{pmatrix}$

Example 11.358.

$$[2\oplus 3]\left(\begin{pmatrix}1&2&3&4\\5&6&7&8\\10&20&30&40\\11&12&13&14\end{pmatrix}\right)=\begin{pmatrix}1&0&3&4\\5&0&7&8\\10&1&30&40\\11&0&13&14\end{pmatrix}$$

Definition 11.359. (adjoint) *Let $n\in\mathbb{N}$, F a field and $M\in\mathcal{M}_{n,n}(F)$ then we define*

$$\text{adjoint}:\mathcal{M}_{n,n}(F)\rightarrow\mathcal{M}_{n,n}(F) \text{ where } \forall i,j\in\{1,\ldots,n\}\ (\text{adjoint}(M))_{i,j}=\det([i\oplus j](M))$$

The relation between the adjoint of a linear transformation and its matrix is expressed in the following theorem.

Theorem 11.360. *Let $n\in\mathbb{N}$, X a finite dimensional vector space over a field with characteristic zero, $E=\{e_1,\ldots,e_n\}$ a basis of X and $L\in\text{Hom}(X,X)$ then*

$$\mathcal{M}(E,E)(\text{adjoint}(L))=\text{adjoint}(\mathcal{M}(E,E)(L))$$

or in other words

$$\mathcal{M}(\text{adjoint}(L);E,E)=\text{adjoint}(\mathcal{M}(L;E,E))$$

So the matrix of the adjoint of a linear transformation is the adjoint of the matrix of the linear transformation.

Proof. Using [theorem: 11.283] there exists a determinant function in X such that

$$\Delta(e_1,\ldots,e_n)=1$$

Now using the definition of the matrix of a linear map we have for $i\in\{1,\ldots,n\}$

$$\begin{aligned}
\sum_{i\in\{1,\ldots,n\}}\mathcal{M}(\text{adjoint}(L);E,E)_{i,j}\cdot e_i &\underset{\text{def}}{=}\\
\text{adjoint}(L)(e_j)\\
1\cdot\text{adjoint}(L)(e_j) &=\\
\Delta(e_1,\ldots,e_n)\cdot\text{adjoint}(L)(e_j) &\underset{[\text{definition: } 11.304]}{=}\\
\overline{\Delta L}(e_1,\ldots,e_n)(e_j) &\underset{[\text{definition: } 11.301]}{=}\\
\sum_{i\in\{1,\ldots,n\}}(-1)^{i-1}\cdot\Delta(e_j,L(e_1),\ldots,L(e_{i-1}),L(e_{i+1}),\ldots,L(e_n))\cdot e_i
\end{aligned}$$

From the uniqueness of expanding in a basis it follows that $\forall i,j\in\{1,\ldots,n\}$ we have

$$\mathcal{M}(\text{adjoint}(L);E,E)_{i,j}=(-1)^{i-1}\cdot\Delta(e_j,L(e_1),\ldots,L(e_{i-1}),L(e_{i+1}),\ldots,L(e_n)) \tag{11.250}$$

Given $i,j\in\{1,\ldots,n\}$ define

$$x^{(i,j)}\in X^n \text{ by } \forall k\in\{1,\ldots,n\}\ x_k^{(i,j)}=\begin{cases} e_j \text{ if } k=i\\ L(e_k) \text{ if } k\in\{1,\ldots,n\}\setminus\{i\}\end{cases} \tag{11.251}$$

Then we have for $i,j\in\{1,\ldots,n\}$ and $k\in\{1,\ldots,n\}$ that

$$\begin{aligned}(e_j,L(e_1),\ldots,L(e_{i-1}),L(e_{i+1}),\ldots,L(e_n))_k &\underset{\text{[definition: 11.286]}}{=}\\ \begin{cases} e_j \text{ if } k=1\\ L(e_{k-1}) \text{ if } k\in\{2,\ldots,\boldsymbol{i}\}\\ L(e_k) \text{ if } k\in\{\boldsymbol{i+1},\ldots,n\}\end{cases}\\ \begin{cases} x_i^{(i,j)} \text{ if } k=1\\ x_{k-1}^{(i,j)} \text{ if } k\in\{2,\ldots,i\}\\ x_k^{(i,j)} \text{ if } k\in\{i+1,\ldots,n\}\end{cases} &=\\ \left(x_i^{(i,j)},x_1^{(i,j)},\ldots,x_{i-1}^{(i,j)},x_{i+1}^{(i,j)},\ldots,x_n^{(i,j)}\right)_k &\underset{\text{[theorem: 11.287]}}{=}\\ \left(x^{(i,j)}_{(i\underset{n}{\rightsquigarrow}1)(1)},\ldots,x^{(i,j)}_{(i\underset{n}{\rightsquigarrow}1)(n)}\right)\end{aligned}$$

So substituting the above in [eq: 11.250] gives $\forall i,j\in\{1,\ldots,n\}$ that

$$\begin{aligned}\mathcal{M}(\operatorname{adjoint}(L);E,E)_{i,j} &= (-1)^{i-1}\cdot\Delta\left(x^{(i,j)}_{(i\underset{n}{\rightsquigarrow}1)(1)},\ldots,x^{(i,j)}_{(i\underset{n}{\rightsquigarrow}1)(n)}\right)\\ &= (-1)^{i-1}\cdot\left(\left(i\underset{n}{\rightsquigarrow}1\right)\Delta\right)\left(x_1^{(i,j)},\ldots,x_n^{(i,j)}\right)\\ &= (-1)^{i-1}\cdot\operatorname{sign}\left(\left(i\underset{n}{\rightsquigarrow}1\right)\right)\cdot\Delta\left(x_1^{(i,j)},\ldots,x_n^{(i,j)}\right)\\ &\underset{\text{[theorem: 11.239]}}{=} (-1)^{i-1}\cdot(-1)^{i-1}\cdot\Delta\left(x_1^{(i,j)},\ldots,x_n^{(i,j)}\right)\\ &= \Delta\left(x_1^{(i,j)},\ldots,x_n^{(i,j)}\right)\end{aligned} \tag{11.252}$$

Let $i,j\in\{1,\ldots,n\}$ and define $M^{(i,j)}\in\mathcal{M}_{n,n}(F)$ by

$$M^{(i,j)}=[i\oplus j](\mathcal{M}(E,E)(L)) \tag{11.253}$$

so that by [definition: 11.357]

$$\forall r,s\in\{1,\ldots,n\} \text{ we have } (M^{(i,j)})_{r,s}=\begin{cases}(\mathcal{M}(E,E)(L))_{r,s} \text{ if } s\in\{1,\ldots,n\}\setminus\{i\}\\ \delta_{r,j} \text{ if } s=i\end{cases} \tag{11.254}$$

Let $L^{(i,j)}: X\rightarrow X$ be defined by $L^{(i,j)}=\mathcal{M}(E,E)^{-1}(M^{(i,j)})$ so that $\mathcal{M}(E,E)(L^{(i,j)})=M^{(i,j)}$. Then we have

$$\forall r\in\{1,\ldots,n\} \text{ we have } L^{(i,j)}(e_r)=\sum_{s\in\{1,\ldots,n\}}(M^{(i,j)})_{s,r}\cdot e_s \tag{11.255}$$

For $r\in\{1,\ldots,n\}$ we have either:

$\boldsymbol{r=i.}$ Then

$$\begin{aligned}L^{(i,j)}(e_r) &= L^{(i,j)}(e_i)\\ &\underset{\text{[eq: 11.255]}}{=} \sum_{s\in\{1,\ldots,n\}}(M^{(i,j)})_{s,i}\cdot e_s\\ &\underset{\text{[eq: 11.254]}}{=} \sum_{s\in\{1,\ldots,n\}}\delta_{s,j}\cdot e_s\\ &= e_j\\ &\underset{r=i\wedge\text{[eq: 11.251]}}{=} x_i^{(i,j)}\\ &= x_r^{(i,j)}\end{aligned}$$

$\boldsymbol{r \neq i}$**.** Then

$$
\begin{array}{rcl}
L^{(i,j)}(e_r) & \underset{[\text{eq: } 11.255]}{=} & \sum_{s\in\{1,\ldots,n\}} (M^{(i,j)})_{s,r}\cdot e_s \\
& \underset{[\text{eq: } 11.254]}{=} & \sum_{s\in\{1,\ldots,n\}} (\mathcal{M}(E,E)(L))_{s,r}\cdot e_s \\
& = & \sum_{s\in\{1,\ldots,n\}} (\mathcal{M}(L;E,E)(L))_{s,r}\cdot e_s \\
& = & L(e_r) \\
& \underset{r\neq u\wedge[\text{eq: } 11.251]}{=} & x_r^{i,j}
\end{array}
$$

So we have proved that

$$\forall r\in\{1,\ldots,n\} \text{ we have } L^{(i,j)}(e_r) = x_r^{(i,j)} \tag{11.256}$$

Substituting the above in [eq: 11.252] results in $\forall i,\in\{1,\ldots,n\}$

$$
\begin{array}{rcl}
\mathcal{M}(\text{adjoint}(L);E,E)_{i,j} & = & \Delta(L^{(i,j)}(e_1),\ldots,L^{(i,j)}(e_n)) \\
& = & \Delta_{L^{(i,j)}}(e_1,\ldots,e_n) \\
& \underset{[\text{definition: } 11.295]}{=} & \det(L^{(i,j)})\cdot\Delta(e_1,\ldots,e_n) \\
& = & \det(L^{(i,j)}) \\
& \underset{[\text{theorem: } 11.339]}{=} & \det(\mathcal{M}[E,E](L^{(i,j)})) \\
& = & \det(\mathcal{M}[E,E](\mathcal{M}[E,E]^{-1}(M^{(i,j)}))) \\
& = & \det(M^{(i,j)}) \\
& \underset{[\text{eq: } 11.253]}{=} & \det([i\oplus j](\mathcal{M}[E,E](L))) \\
& \underset{[\text{definition: } 11.359]}{=} & (\text{adjoint}(\mathcal{M}[E,E])(L))_{i,j}
\end{array}
$$

proving finally that

$$\mathcal{M}(\text{adjoint}(L);E,E) = \text{adjoint}(\mathcal{M}[E,E])(L)$$

□

The following theorem is a reformulation of [theorem: 11.306] in terms of matrices instead of linear transformations.

Theorem 11.361. *Let $n\in\mathbb{N}$, X a finite dimensional vector space over a field with characteristic zero, $E=\{e_1,\ldots,e_n\}$ a basis of X and $L\in\operatorname{Hom}(X,X)$ then we have:*

1. $\mathcal{M}(E,E)(L)\cdot\text{adjoint}(\mathcal{M}(E,E)(L)) = \det(\mathcal{M}(E,E)(L))\cdot E$
2. $\text{adjoint}(\mathcal{M}(E,E)(L))\cdot\mathcal{M}(E,E)(L) = \det(\mathcal{M}(E,E)(L))\cdot E$

or taking in account that $\mathcal{M}(E,E)(L)=\mathcal{M}(L;E,E)$ that:

1. $\mathcal{M}(L;E,E)\cdot\text{adjoint}(\mathcal{M}(L;E,E)) = \det(\mathcal{M}(L;E,E))\cdot E$
2. $\text{adjoint}(\mathcal{M}(L;E,E))\cdot\mathcal{M}(L;E,E) = \det(\mathcal{M}(L;E,E))\cdot E$

or if we write out the matrix product:

1. $\forall i,j\in\{1,\ldots,n\}$ *we have*

$$\sum_{k\in\{1,\ldots,n\}} \mathcal{M}(L;E,E)_{i,k}\cdot\text{adjoint}(\mathcal{M}(L;E,E))_{k,j} \;=\; \det(\mathcal{M}(L;E,E))\cdot\delta_{i,j}$$

2. $\forall i,j\in\{1,\ldots,n\}$ *we have*

$$\sum_{k\in\{1,\ldots,n\}} \text{adjoint}(\mathcal{M}(L;E,E))_{i,k}\cdot\mathcal{M}(L;E,E)_{k,j} \;=\; \det(\mathcal{M}(L;E,E))\cdot\delta_{i,j}$$

Proof.

1. We have

$$\begin{aligned}\det(\mathcal{M}(E,E)(L))\cdot E &\underset{[\text{theorem: } 11.324]}{=} \det(\mathcal{M}(E,E)(L))\cdot\mathcal{M}(E,E)(\mathrm{Id}_X)\\ &\underset{[\text{theorem: } 11.324]}{=} \mathcal{M}(E,E)(\det(\mathcal{M}(E,E)(L))\cdot\mathrm{Id}_X)\\ &\underset{[\text{theorem: } 11.339]}{=} \mathcal{M}(E,E)(\det(L)\cdot\mathrm{Id}_X)\\ &\underset{[\text{theorem: } 11.306]}{=} \mathcal{M}(E,E)(L\circ\mathrm{adjoint}(L))\\ &\underset{[\text{theorem: } 11.328]}{=} \mathcal{M}(E,E)(L)\cdot\mathcal{M}(E,E)(\mathrm{adjoint}(L))\end{aligned}$$

2. We have

$$\begin{aligned}\det(\mathcal{M}(E,E)(L))\cdot E &\underset{[\text{theorem: } 11.324]}{=} \det(\mathcal{M}(E,E)(L))\cdot\mathcal{M}(E,E)(\mathrm{Id}_X)\\ &\underset{[\text{theorem: } 11.324]}{=} \mathcal{M}(E,E)(\det(\mathcal{M}(E,E)(L))\cdot\mathrm{Id}_X)\\ &\underset{[\text{theorem: } 11.339]}{=} \mathcal{M}(E,E)(\det(L)\cdot\mathrm{Id}_X)\\ &\underset{[\text{theorem: } 11.306]}{=} \mathcal{M}(E,E)(\mathrm{adjoint}(L)\circ L)\\ &\underset{[\text{theorem: } 11.328]}{=} \mathcal{M}(E,E)(\mathrm{adjoint}(L))\cdot\mathcal{M}(E,E)(L)\end{aligned}$$

□

Next we look at some properties of the determinant and adjoint of a matrix that helps us to calculate the determinant and adjoint of a matrix and in the end the inverse of invertible matrices.

Theorem 11.362. *Let $n\in\mathbb{N}$, F a field with characteristic zero and $M\in\mathcal{M}_{n,n}(F)$ then we have:*

1.

$$M\cdot\mathrm{adjoint}(M)=\det(M)\cdot E,$$

hence $\forall i,j\in\{1,\ldots,n\}$

$$\sum_{k\in\{1,\ldots,n\}} M_{i,k}\cdot\mathrm{adjoint}(M)_{k,j}=\det(M)\cdot\delta_{i,j}$$

2.

$$\mathrm{adjoint}(M)\cdot M=\det(M)\cdot E$$

hence $\forall i,j\in\{1,\ldots,n\}$

$$\sum_{k\in\{1,\ldots,n\}} \mathrm{adjoint}(M)_{i,k}\cdot M_{k,j}=\det(M)\cdot\delta_{i,j}$$

3. *If* $n>1$ *then* $\forall i,j\in\{1,\ldots,n\}$ *we have*

$$\mathrm{adjoint}\,(M)_{i,j}=(-1)^{i+j}\cdot\det([j\boxplus i](M))$$

4. *If* $n>1$ *then we have* $\forall j\in\{1,\ldots,n\}$

$$\det(M)=\sum_{i\in\{1,\ldots,n\}}(-1)^{i+j}\cdot M_{i,j}\cdot\det([i\boxplus j](M))$$

we call this the expansion of the determinant with respect to the j-the column.

5. *If* $n>1$ *then we have* $\forall j\in\{1,\ldots,n\}$

$$\det(M)=\sum_{i\in\{1,\ldots,n\}}(-1)^{i+j}\cdot M_{j,i}\cdot\det([j\boxplus i](M))$$

we call this the expansion of the determinant with respect to the j-the row.

6. $\det(E)=1$

7. *If* $M\in\mathcal{M}_{1,1}(F)$ *then* $\det(M)=M_{1,1}$ *and* $\mathrm{adjoint}(M)=(1)$

Proof.

1. Take $L=\mathcal{M}(E,E)^{-1}(M)$ so that $M=\mathcal{M}(E,E)(L)$ then we have that

$$\begin{array}{rcl} M\cdot\mathrm{adjoint}(M) & = & \mathcal{M}(E,E)(L)\cdot\mathrm{adjoint}(\mathcal{M}(E,E)(L)) \\ & \underset{[\text{theorem: } 11.361]}{=} & \det(\mathcal{M}(E,E)(L))\cdot E \\ & = & \det(M)\cdot E \end{array}$$

2. Take $L=\mathcal{M}(E,E)^{-1}(M)$ so that $M=\mathcal{M}(E,E)(L)$ then we have that

$$\begin{array}{rcl} \mathrm{adjoint}(M)\cdot M & = & \mathrm{adjoint}(\mathcal{M}(E,E)(L))\cdot\mathcal{M}(E,E)(L) \\ & \underset{[\text{theorem: } 11.361]}{=} & \det(\mathcal{M}(E,E)(L))\cdot E \\ & = & \det(M)\cdot E \end{array}$$

3. First if $i,j\in\{1,\ldots,n\}$ then we have

$$\begin{array}{rcl} \det([i\oplus j](M)) & \underset{[\text{theorem: } 11.338]}{=} & \det(([i\oplus j](M))^T) \\ & = & \det(([i\oplus j](M))^T)\cdot\delta_{i,i} \\ & \underset{(1)}{=} & \sum_{k\in\{1,\ldots,n\}} (([i\oplus j](M))^T)_{i,k}\cdot\mathrm{adjoint}(([i\oplus j](M))^T)_{k,i} \\ & = & \sum_{k\in\{1,\ldots,n\}} ([i\oplus j](M))_{k,i}\cdot\mathrm{adjoint}(([i\oplus j](M))^T)_{k,i} \\ & \underset{[\text{definition: } 11.357]}{=} & \sum_{k\in\{1,\ldots,n\}} \delta_{k,j}\cdot\mathrm{adjoint}(([i\oplus j](M))^T)_{k,i} \\ & = & \mathrm{adjoint}(([i\oplus j](M))^T)_{j,i} \\ & \underset{[\text{definition: } 11.359]}{=} & \det([j\oplus i](([i\oplus j](M))^T)) \end{array}$$

giving

$$\det([i\oplus j](M))=\det([j\oplus i](([i\oplus j](M))^T)) \tag{11.257}$$

Then $\forall k\in\{1,\ldots,n\}$ we have

$$\begin{array}{rcl} \mathrm{row}([j\oplus i](([i\oplus j](M))^T),i)_k & = & ([j\oplus i](([i\oplus j](M))^T))_{i,k} \\ & \underset{[\text{definition: } 11.357]}{=} & \begin{cases} (([i\oplus j](M))^T)_{i,k} \text{ if } k\in\{1,\ldots,n\}\setminus\{j\} \\ \delta_{i,i} \text{ if } k=j \end{cases} \\ & = & \begin{cases} ([i\oplus j](M))_{k,i} \text{ if } k\in\{1,\ldots,n\}\setminus\{j\} \\ 1 \text{ if } k=j \end{cases} \\ & \underset{[\text{definition: } 11.357]}{=} & \begin{cases} \delta_{k,j} \text{ if } k\in\{1,\ldots,n\}\setminus\{j\} \\ 1 \text{ if } k=j \end{cases} \\ & = & \begin{cases} 0 \text{ if } k\in\{1,\ldots,n\}\setminus\{j\} \\ 1 \text{ if } k=j \end{cases} \\ & = & \delta_{k,j} \end{array}$$

giving

$$\mathrm{row}([j\oplus i](([i\oplus j](M))^T),i)_k=\delta_{k,j} \tag{11.258}$$

Further

$$\begin{array}{rcl} \mathrm{col}([j\oplus i](([i\oplus j](M))^T),j)_k & = & ([j\oplus i](([i\oplus j](M))^T))_{k,j} \\ & \underset{[\text{definition: } 11.357]}{=} & \begin{cases} (([i\oplus j](M))^T)_{k,j} \text{ if } j\in\{1,\ldots,n\}\setminus\{j\} \\ \delta_{k,i} \text{ if } j=j \end{cases} \\ & = & \delta_{k,i} \end{array}$$

giving

$$\mathrm{col}([j\oplus i](([i\oplus j](M))^T),j)_k=\delta_{k,i} \tag{11.259}$$

From [eqs: 11.258, 11.259] it follows that the conditions of [theorem: 11.356] are satisfied, hence

$$\det([j\oplus i](([i\oplus j](M))^T))=(-1)^{i+j}\cdot\det((i\boxplus j)([j\oplus i](([i\oplus j](M))^T)))$$

which combined with [eq: 11.257] results in

$$\det([i\oplus j](M))=(-1)^{i+j}\cdot\det((i\boxplus j)([j\oplus i](([i\oplus j](M))^T))) \tag{11.260}$$

Let $k,l\in\{1,\ldots,n-1\}$ then we have for $((i\boxplus j)([j\oplus i](([i\oplus j](M))^T)))_{k.l}$ the following cases to consider for k,l:

$1\leqslant k<i\wedge 1\leqslant l<j$. Then

$$\begin{aligned}
((i\boxplus j)([j\oplus i](([i\oplus j](M))^T)))_{k.l} &\underset{[\text{definition: } 11.349]}{=}\\
([j\oplus i](([i\oplus j](M))^T))_{k,l} &\underset{l<j\Rightarrow l\neq j\wedge[\text{definition: } 11.357]}{=}\\
(([i\oplus j](M))^T)_{k,l} &=\\
([i\oplus j](M))_{l.k} &\underset{k<i\Rightarrow k\neq i\wedge[\text{definition: } 11.357]}{=}\\
M_{l,k} &=\\
(M^T)_{k,l} &\underset{[\text{definition: } 11.349]}{=}\\
([i\boxplus j](M^T))_{k,l}
\end{aligned}$$

$i\leqslant k\leqslant n-1\wedge 1\leqslant l<j$. Then

$$\begin{aligned}
((i\boxplus j)([j\oplus i](([i\oplus j](M))^T)))_{k.l} &\underset{[\text{definition: } 11.349]}{=}\\
([j\oplus i](([i\oplus j](M))^T))_{k+1,l} &\underset{l<j\Rightarrow l\neq j\wedge[\text{definition: } 11.357]}{=}\\
(([i\oplus j](M))^T)_{k+1,l} &=\\
([i\oplus j](M))_{l,k+1} &\underset{i\leqslant k\Rightarrow i\neq k+1\wedge[\text{definition: } 11.357]}{=}\\
M_{l,k+1} &=\\
(M^T)_{k+1,l} &\underset{[\text{definition: } 11.349]}{=}\\
([i\boxplus j](M^T))_{k,l}
\end{aligned}$$

$1\leqslant k<i\wedge j\leqslant l\leqslant n-1$. Then

$$\begin{aligned}
((i\boxplus j)([j\oplus i](([i\oplus j](M))^T)))_{k.l} &\underset{[\text{definition: } 11.349]}{=}\\
([j\oplus i](([i\oplus j](M))^T))_{k,l+1} &\underset{j\leqslant l\Rightarrow l+1\neq j\wedge[\text{definition: } 11.357]}{=}\\
(([i\oplus j](M))^T)_{k,l+1} &=\\
([i\oplus j](M))_{l+1,k} &\underset{k<i\Rightarrow k\neq i\wedge[\text{definition: } 11.357]}{=}\\
M_{l+1,k} &=\\
(M^T)_{k,l+1} &\underset{[\text{definition: } 11.349]}{=}\\
([i\boxplus j](M^T))_{k,l}
\end{aligned}$$

$i\leqslant k\leqslant n-1\wedge j\leqslant l\leqslant n-1$. Then

$$\begin{aligned}
((i\boxplus j)([j\oplus i](([i\oplus j](M))^T)))_{k.l} &\underset{[\text{definition: } 11.349]}{=}\\
([j\oplus i](([i\oplus j](M))^T))_{k+1,l+1} &\underset{j\leqslant l\Rightarrow j\neq l+1\wedge[\text{definition: } 11.357]}{=}\\
(([i\oplus j](M))^T)_{k+1,l+1} &=\\
([i\oplus j](M))_{l+1,k+1} &\underset{i\leqslant k\Rightarrow i\neq k+1\wedge[\text{definition: } 11.357]}{=}\\
M_{l+1,k+1} &=\\
(M^T)_{k+1,l+1} &\underset{[\text{definition: } 11.349]}{=}\\
([i\boxplus j](M^T))_{k,l}
\end{aligned}$$

The above proves that

$$(i \boxplus j)([j \oplus i]((([i \oplus j](M))^T)) = [i \boxplus j](M^T)$$

combining the above with [eq: 11.260] gives

$$\begin{array}{rcl} \det([i \oplus j](M)) & = & (-1)^{i+j} \cdot \det([i \boxplus j](M^T)) \\ & \underset{\text{[theorem: 11.351]}}{=} & (-1)^{i+j} \cdot \det(([j \boxplus i](M))^T) \end{array} \tag{11.261}$$

Finally

$$\begin{array}{rcl} \operatorname{adjoint}(M)_{i,j} & \underset{\text{[definition: 11.359]}}{=} & \det([i \oplus j](M)) \\ & \underset{\text{[eq: 11.261]}}{=} & (-1)^{i+j} \cdot \det(([j \boxplus i](M))^T) \\ & \underset{\text{[theorem: 11.338]}}{=} & (-1)^{i+j} \cdot \det([j \boxplus i]M) \end{array}$$

4. Let $i \in \{1, \ldots, n\}$ then we have

$$\begin{array}{rcl} \det(M) & = & \det(M) \cdot \delta_{i,i} \\ & \underset{(2)}{=} & \sum_{k \in \{1, \ldots, n\}} \operatorname{adjoint}(M)_{i,k} \cdot M_{k,i} \\ & \underset{(3)}{=} & \sum_{k \in \{1, \ldots, n\}} (-1)^{i+k} \cdot \det([k \boxplus i](M)) \cdot M_{k,i} \\ & = & \sum_{k \in \{1, \ldots, n\}} (-1)^{i+k} \cdot M_{k,i} \cdot \det([k \boxplus i](M)) \end{array}$$

5. Let $i \in \{1, \ldots, n\}$ then we have

$$\begin{array}{rcl} \det(M) & = & \det(M) \cdot \delta_{i,i} \\ & \underset{(1)}{=} & \sum_{k \in \{1, \ldots, n\}} M_{i,k} \cdot \operatorname{adjoint}(M)_{k,i} \\ & \underset{(3)}{=} & \sum_{k \in \{1, \ldots, n\}} M_{i,k} \cdot (-1)^{i+k} \cdot \det([i \boxplus k](M)) \\ & = & \sum_{k \in \{1, \ldots, n\}} (-1)^{i+k} \cdot M_{i,k} \cdot \det([i \boxplus k](M)) \end{array}$$

6. This follows from [theorem: 11.362]
7. Using [theorem: 11.335] we have for $M \in \mathcal{M}_{1,1}(F)$ that

$$\det(M) = M_{1,1}$$

Further

$$\operatorname{adjoint}(M) = \det\big([1 \oplus 1]\big(\ M_{1,1}\ \big)\big) = \det((\ 1\ )) = 1$$

□

We can use (4),(7) from the above to calculate the determinant of a matrix as show in the following example:

Example 11.363. Using [theorem: 11.362 (4)] we have

$$\begin{array}{rcl} \left|\begin{array}{ccc} -2 & 2 & 3 \\ -1 & 1 & 3 \\ 2 & 0 & 1 \end{array}\right| & = & \\ (-1)^{1+1} \cdot M_{1,1} \cdot \det([1 \boxplus 1]M) + (-1)^{1+2} \cdot M_{2,1} \cdot \det([2 \boxplus 1] \cdot M) + (-1)^{1+3} \cdot M_{3,1} \cdot \det([3 \boxplus 1]M) & = & \\ -2 \cdot \left|\begin{array}{cc} 1 & 3 \\ 0 & 1 \end{array}\right| + (-1) \cdot (-1) \cdot \left|\begin{array}{cc} 2 & 3 \\ 0 & 1 \end{array}\right| + 2 \cdot \left|\begin{array}{cc} 2 & 3 \\ 1 & 3 \end{array}\right| & = & \\ -2 \cdot (1 \cdot |\ 1\ | + (-1) \cdot 0 \cdot |\ 3\ |) + (2 \cdot |\ 1\ | + (-1) \cdot 0 \cdot |\ 3\ |) + 2 \cdot (2 \cdot |\ 3\ | + (-1) \cdot 1 \cdot |\ 3\ |) & = & \\ -2 + 2 + 2 \cdot 3 & = & \\ 6 & & \end{array}$$

To simplify calculation of the adjoint of a matrix we introduce the co-factor matrix

Definition 11.364. *Let $n \in \mathbb{N} \setminus \{1\}$, F a field with characteristic zero and $M \in \mathcal{M}_{n,n}(F)$ then the co-factor matrix noted by* $\operatorname{cofactor}(M)$ *is defined by*

$$\operatorname{cofactor}(M) \in \mathcal{M}_{n,n}(M) \textit{ where } \forall i, j \in \{1, \ldots, n\} \operatorname{cofactor}(M)_{i,j} = (-1)^{i+j} \cdot \det([i \boxplus j](M))$$

The use of the co-factor matrix is shown in the following theorem:

Theorem 11.365. *Let $n \in \mathbb{N} \setminus \{1\}$, F a field with characteristic zero and $M \in \mathcal{M}_{n,n}(F)$ then*

$$\operatorname{adjoint}(M) = (\operatorname{cofactor}(M))^T$$

Proof. Let $i, j \in \{1, \ldots, n\}$ then

$$((\operatorname{cofactor}(M))^T)_{i,j} = \operatorname{cofactor}(M)_{j,i} = (-1)^{j+i} \cdot \det([j \boxplus i](M)) \underset{[\text{theorem: } 11.362(3}{=} \operatorname{adjoint}(M)_{i,j}$$

proving that

$$\operatorname{adjoint}(M) = (\operatorname{cofactor}(M))^T$$ □

Example 11.366.

$$\begin{aligned}\operatorname{cofactor}\left(\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}\right) &= \begin{pmatrix} 1 \cdot \begin{vmatrix} 5 & 6 \\ 8 & 9 \end{vmatrix} & -1 \cdot \begin{vmatrix} 4 & 6 \\ 7 & 9 \end{vmatrix} & 1 \cdot \begin{vmatrix} 4 & 5 \\ 7 & 8 \end{vmatrix} \\ -1 \cdot \begin{vmatrix} 2 & 3 \\ 8 & 9 \end{vmatrix} & 1 \cdot \begin{vmatrix} 1 & 3 \\ 7 & 9 \end{vmatrix} & -1 \cdot \begin{vmatrix} 1 & 2 \\ 7 & 8 \end{vmatrix} \\ 1 \cdot \begin{vmatrix} 2 & 3 \\ 5 & 6 \end{vmatrix} & -1 \cdot \begin{vmatrix} 1 & 3 \\ 4 & 6 \end{vmatrix} & 1 \cdot \begin{vmatrix} 1 & 2 \\ 4 & 5 \end{vmatrix} \end{pmatrix} \\ &= \begin{pmatrix} -3 & 6 & -3 \\ 6 & -12 & 6 \\ -3 & 6 & -3 \end{pmatrix}\end{aligned}$$

so that

$$\begin{aligned}\operatorname{adjoint}\left(\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}\right) &= \begin{pmatrix} -3 & 6 & -3 \\ 6 & -12 & 6 \\ -3 & 6 & -3 \end{pmatrix}^T \\ &= \begin{pmatrix} -3 & 6 & -3 \\ 6 & -12 & 6 \\ -3 & 6 & -3 \end{pmatrix}\end{aligned}$$

Sometimes there exists shortcuts for calculating the determinant of a matrix.

Corollary 11.367. *Let $n \in \mathbb{N} \setminus \{1\}$ then for $M \in \mathcal{M}_{n,n}(F)$ we have:*

1. *If $\exists m \in \mathbb{N}$ with $1 \leqslant m < n$ such that $\forall i \in \{1, \ldots, m\}$ and $\forall j \in \{1, \ldots, n\}$ we have*

$$\operatorname{col}(M, i)_j = \delta_{i,j}$$

   then

$$\det(M) = \det([>m](M))$$

2. *If $\exists m \in \mathbb{N}$ with $1 < m \leqslant n$ such that $\forall i \in \{m, \ldots, n\}$ we have*

$$\operatorname{col}(M, i)_j = \delta_{i,j}$$

   then

$$\det(M) = \det([<m](M))$$

3. *If $\exists m \in \mathbb{N}$ with $1 \leqslant m < n$ such that $\forall i \in \{1, \ldots, m\}$ and $\forall j \in \{1, \ldots, n\}$ we have*

$$\operatorname{row}(M, i)_j = \delta_{i,j}$$

   then

$$\det(M) = \det([>m](M))$$

4. *If* $\exists m \in \mathbb{N}$ *with* $1 < m \leqslant n$ *such that* $\forall i \in \{m, \ldots, n\}$ *and* $\forall j \in \{1, \ldots, n\}$ *we have*

$$\mathrm{row}(M, i)_j = \delta_{i,j}$$

then

$$\det(M) = \det([<m](M))$$

Proof.

1. *We prove this by induction, so let*

$S_n = \{m \in \mathbb{N} | \text{If } m < n \text{ and } M \in \mathcal{M}_{n,n}(F) \text{ satisfies } \forall i \in \{1, \ldots, m\}, \forall j \in \{1, \ldots, n\} \ \mathrm{col}(M, i)_j = \delta_{i,j} \text{ then } \det(M) = \det([>m](M))\}$

then we have:

$1 \in S_n$. *Let* $M \in \mathcal{M}_{n,n}(F)$ *be such that* $\forall i \in \{1\}, \forall j \in \{1, \ldots, n\}$ $\mathrm{col}(M, i)_j = \delta_{j,i}$*, hence* $\mathrm{col}(M, 1) = \delta_{1,i}$*, then we have*

$$\begin{aligned}
\det(M) &\underset{[\text{theorem: } 11.362(4)]}{=} \sum_{i \in \{1, \ldots, n\}} (-1)^{i+1} \cdot M_{i,1} \cdot \det([i \boxplus 1](M)) \\
&= \sum_{i \in \{1, \ldots, n\}} (-1)^{i+1} \cdot \delta_{i,1} \cdot \det([i \boxplus 1](M)) \\
&= (-1)^2 \cdot \det([1 \boxplus 1](M)) \\
&= \det([1 \boxplus 1](M)) \\
&\underset{[\text{theorem: } 11.353]}{=} \det([>1](M))
\end{aligned}$$

proving that $1 \in S_n$.

$m \in S_n \Rightarrow m+1 \in S_n$. *Let* $m+1 < n$ *then we have*

$$1 < n - m \Rightarrow 1 \leqslant n - m - 1 \Rightarrow 1 \in \{1, \ldots, n - m - 1\} \tag{11.262}$$

Further let $M \in \mathcal{M}_{n,n}(F)$ *be such that* $\forall i \in \{1, \ldots, m+1\}, \forall j \in \{1, \ldots, n\}$ *we have* $\mathrm{col}(M, i)_j = \delta_{i,j}$*. As* $m \in S_n$*,* $m < m+1 < n$ *and* $\forall i \in \{1, \ldots, m\}, \forall j \in \{1, \ldots, n\}$ $\mathrm{col}(M, i)_j = \delta_{i,j}$ *we have*

$$\det(M) = \det([>m](M)) \tag{11.263}$$

If $j \in \{1, \ldots, n - m\}$ *then*

$$\begin{aligned}
\mathrm{col}([>m](M), 1)_j &= ([>m](M))_{j,1} \\
&\underset{[\text{definition: } 11.352]}{=} M_{j+m, 1+m} \\
&= \mathrm{col}(1+m, j+m) \\
&= \delta_{1+m, j+m} \\
&= \delta_{j,1}
\end{aligned} \tag{11.264}$$

Next

$$\begin{aligned}
\det(M) &\underset{[\text{eq: } 11.263]}{=} \\
\det([>m](M)) &\underset{[\text{theorem: } 11.362(4)]}{=} \\
\sum_{i \in \{1, \ldots, n\}} (-1)^{i+1} \cdot ([>m](M))_{i,1} \cdot \det([i \boxplus 1]([>m])(M)) &= \\
\sum_{i \in \{1, \ldots, n\}} (-1)^{i+1} \cdot \mathrm{col}([>m](M), 1)_j \cdot \det([i \boxplus 1]([>m])(M)) &\underset{[\text{eq: } 11.264]}{=} \\
\sum_{i \in \{1, \ldots, n\}} (-1)^{i+1} \cdot \delta_{j,1} \cdot \det([i \boxplus 1]([>m])(M)) &= \\
(-1)^2 \cdot \det([1 \boxplus 1]([>m])(M)) &\underset{[\text{theorem: } 11.353]}{=} \\
\det([>1]([>m](M))) &\underset{[\text{eq:} 11.262][\text{theor:} 11.353]}{=} \\
\det([>m+1](M))
\end{aligned}$$

proving that $m+1 \in S_n$.

Mathematical induction proves then that $S_n = \mathbb{N}$*. So if* $1 \leqslant m < n$ *and* $M \in \mathcal{M}_{n,n}(F)$ *satisfies* $\forall i \in \{1, \ldots, m\}, \forall j \in \{1, \ldots, n\}$ $\mathrm{col}(M, i)_j = \delta_{i,j}$ *then we have, as* $m \in \mathbb{N} = S_n$*, that* $\det(M) = \det([>m](M))$.

2. *We prove this by induction, so let*

$S = \{n \in \{2, \ldots, \infty\} | \text{If } M \in \mathcal{M}_{n,n}(F) \text{ satisfies } \exists m \in \{2, \ldots, n\} \text{ such that } \forall i \in \{m, \ldots, n\}, \forall j \in \{1, \ldots, n\} \text{ we have } \operatorname{col}(M, i)_j = \delta_{i,j} \text{ then } \det(M) = \det([<m](M))\}$

then we have:

$2 \in S$. *Let $M \in \mathcal{M}_{2,2}(F)$ satisfying $\exists m \in \{2, \ldots, 2\}$ such that $\forall i \in \{m, \ldots, 2\}$ and $\forall j \in \{1, \ldots, 2\}$ we have $\operatorname{col}(M, i)_j = \delta_{i,j}$ then $\forall j \in \{1, \ldots, 2\}$ $\operatorname{col}(M, 2) = \delta_{j,2}$. So we have*

$$\begin{aligned}
\det(M) &\underset{\text{[theorem: 11.362(4)]}}{=} \sum_{i \in \{1, \ldots, 2\}} (-1)^{i+2} \cdot M_{i,2} \cdot \det([i \boxplus 2](M)) \\
&= \sum_{i \in \{1, \ldots, 2\}} (-1)^{i+2} \cdot \operatorname{col}(M, 2)_i \cdot \det([i \boxplus 2](M)) \\
&= \sum_{i \in \{1, \ldots, n\}} (-1)^{i+2} \cdot \delta_{2,i} \cdot \det([i \boxplus 2](M)) \\
&= (-1)^{2+2} \cdot \det([2 \boxplus 2](M)) \\
&= \det([2 \boxplus 2](M)) \\
&\underset{\text{[theorem: 11.353]} \wedge n=2}{=} \det([<2](M))
\end{aligned}$$

proving that $2 \in S$.

$n \in S \Rightarrow n+1 \in S$. *Let $M \in \mathcal{M}_{n+1,n+1}(F)$ such that $\exists m \in \{2, \ldots, n+1\}$ such that $\forall i \in \{m, \ldots, n+1\}$ and $\forall j \in \{1, \ldots, n+1\}$ we have $\operatorname{col}(M, i)_j = \delta_{i,j}$. As*

$$[n+1 \boxplus n+1](M) \underset{\text{[theorem: 11.353]}}{=} [<n+1](M) \tag{11.265}$$

it follows that $\forall i \in \{m, \ldots, n\}$ and $\forall j \in \{1, \ldots, n\}$ we have

$$\begin{aligned}
\operatorname{col}([n+1 \boxplus n+1](M), i)_j &= ([n+1 \boxplus n+1](M))_{j,i} \\
&\underset{i<n+1 \wedge j<n+1 \wedge \text{[definition: 11.349]}}{=} M_{j,i} \\
&= \operatorname{col}(M, i)_j \\
&= \delta_{i,j}
\end{aligned}$$

which as $[n+1 \boxplus n+1](M) \in \mathcal{M}_{n,n}(F)$ and $n \in S$ proves that

$$\det([n+1 \boxplus n+1](M)) = \det([<m]([n+1 \boxplus n+1](M)))$$

substituting [eq: 11.265] in the last term from the above gives

$$\det([n+1 \boxplus n+1](M)) = \det([<m]([<n+1](M)))$$

As $m \in \{2, \ldots, n+1\}$ it follows from [theorem: 11.353] that

$$[<m]([<n+1](M)) = [<m](M)$$

so that

$$\det([n+1 \oplus n+1](M)) = \det([<m](M)) \tag{11.266}$$

Further we have

$$\begin{aligned}
\det(M) &\underset{\text{[theorem: 11.362(4)]}}{=} \sum_{i \in \{1, \ldots, n\}} (-1)^{i+(n+1)} \cdot M_{i,n+1} \cdot \det([i \boxplus n+1](M)) \\
&= \sum_{i \in \{1, \ldots, n\}} (-1)^{i+(n+1)} \cdot \operatorname{col}(M, n+1)_i \cdot \det([i \boxplus n+1](M)) \\
&= \sum_{i \in \{1, \ldots, n\}} (-1)^{i+(n+1)} \cdot \delta_{i,n+1} \cdot \det([i \boxplus n+1](M)) \\
&= (-1)^{(n+1)+(n+1)} \cdot \det([n+1 \boxplus n+1](M)) \\
&= \det([n+1 \boxplus n+1](M)) \\
&\underset{\text{[eq: 11.266]}}{=} \det([<m](M))
\end{aligned}$$

proving that $n+1\in S$.

3. *As* $\forall i\in\{1,\ldots,m\}$ *and* $\forall j\in\{1,\ldots,n\}$ *we have* $\delta_{i,j}=\mathrm{row}(M,i)_j=M_{i,j}=(M^T)_{j,i}=\mathrm{col}(M^T,i)_j$ *we can use (1) to get*

$$\det(M^T)=\det([>m](M^T)) \tag{11.267}$$

hence

$$\begin{array}{rcl}\det(M) & \underset{[\text{theorem: } 11.338]}{=} & \det(M^T)\\ & \underset{[\text{eq: } 11.267]}{=} & \det([>m](M^T))\\ & \underset{[\text{theorem: } 11.353]}{=} & \det(([>m](M))^T)\\ & \underset{[\text{theorem: } 11.338]}{=} & \det([<m](M))\end{array}$$

4. *As* $\forall i\in\{m,\ldots,n\}$ *and* $\forall j\in\{1,\ldots,n\}$ *we have* $\delta_{i,j}=\mathrm{row}(M,i)_j=M_{i,j}=(M^T)_{j,i}=\mathrm{col}(M^T,i)_j$ *we can use (2) to get*

$$\det(M^T)=\det([<m](M^T)) \tag{11.268}$$

hence

$$\begin{array}{rcl}\det(M) & \underset{[\text{theorem: } 11.338]}{=} & \det(M^T)\\ & \underset{[\text{eq: } 11.268]}{=} & \det([<m](M^T))\\ & \underset{[\text{theorem: } 11.353]}{=} & \det(([<m](M))^T)\\ & \underset{[\text{theorem: } 11.338]}{=} & \det([>m](M))\end{array}$$

$\square$

The next theorem shows sufficient and necessary conditions for a matrix to be invertible and how to calculate the inverse of the matrix.

Theorem 11.368. *Let* $n\in\mathbb{N}$, $M\in\mathcal{M}_{n,n}(F)$ *then we have*

$$\begin{array}{c}\mathrm{rank}(M)=m\\ \Updownarrow\\ \det(M)\neq 0\\ \Updownarrow\\ \exists M^{-1}\in\mathcal{M}_{n,n}(F) \text{ such that } M\cdot M^{-1}=E=M^{-1}\cdot M \text{ [in other word } M \text{ is invertible]}\end{array}$$

Further if M *is invariable then*

$$M^{-1}=(\det(M))^{-1}\cdot\mathrm{adjoint}(M)$$

Proof. We have from [theorem: 11.345] that

$$\mathrm{rank}(M)=m\Leftrightarrow\det(M)=0$$

If $\det(M)\neq 0$ then $(\det(M))^{-1}$ exists and we have

$$\begin{array}{rcl}((\det(M))^{-1}\cdot\mathrm{adjoint}(M))\cdot M & = & (\det(M))^{-1}\cdot(\mathrm{adjoint}(M)\cdot M)\\ & \underset{[\text{theorem: } 11.362(2)]}{=} & (\det(M))^{-1}\cdot(\det(M)\cdot E)\\ & = & E\\ M\cdot((\det(M))^{-1}\cdot\mathrm{adjoint}(M)) & = & (\det(M))^{-1}\cdot(M\cdot\mathrm{adjoint}(M))\\ & \underset{[\text{theorem: } 11.362(1)]}{=} & (\det(M))^{-1}\cdot(\det(M)\cdot E)\\ & = & E\end{array}$$

proving

$$\det(M)\neq 0\Rightarrow M^{-1}\cdot M=E=M\cdot M^{-1} \text{ where } M^{-1}=(\det(M))^{-1}\cdot\mathrm{adjoint}(M)$$

Finally if there exists $M^{-1}\in\mathcal{M}_{n,n}(F)$ such that $M^{-1}\cdot M=E=M\cdot M^{-1}$ then

$$1=\det(E)=\det(M\cdot M^{-1})\underset{[\text{theorem: } 11.341]}{=}\det(M)\cdot\det(M^{-1})$$

hence we must have that $\det(M) \neq 0$. So we have

$$\text{If } \exists M^{-1} \in \mathcal{M}_{n,n}(F) \text{ such that } M^{-1} \cdot M = E = M \cdot M^{-1} \text{ then } \det(M) \neq 0$$

□

11.9 Nonsingular transformations

The purpose of this section is to prove that every invertible (nonsingular) transformation on a vector space can always be written as a composition of a **limited** set of linear transformations called the **elementary transformations**. So statements about a general nonsingular transformations can be proved by proving this statement for the elementary transformations (as long as these statements stay valid under composition of transformations).

Definition 11.369. *Let $n \in \mathbb{N}$, X a finite dimensional vector space over a field F with $\dim(X) = n$ then a linear transformation $L \in \operatorname{Hom}(X, X)$ is nonsingular if $\det(L) \neq 0$. Using [theorem: 11.299(2)] this is equivalent with saying that $L: X \rightarrow X$ is a linear isomorphism. The set of nonsingular transformations is noted as $\mathrm{GL}(X)$ so*

$$\mathrm{GL}(X) = \{L \in \operatorname{Hom}(X, X) | \det(L) \neq 0\}$$

We show now that $\mathrm{GL}(X)$ together with the composition forms a group

Theorem 11.370. *Let $n \in \mathbb{N}$, X a finite dimensional vector space over a filed F with $\dim(X) = n$ then*

$$\langle \mathrm{GL}(X), \circ \rangle \textit{ is a sub semi-group of } \langle \operatorname{Hom}(X, X), \circ \rangle \textit{ [see theorem: 11.173}$$

Furthermore

$$\langle \mathrm{GL}(X), \circ \rangle \textit{ is actually a group where } L^{-1} \textit{ is the inverse of } L \in \mathrm{GL}(X)$$

Proof. If $L_1, L_2 \in \mathrm{GL}(X)$ then $\det(L_1) \neq 0 \wedge \det(L_2) \neq 0$ so that

$$\operatorname{delt}(L_1 \circ L_2) \underset{[\text{theorem: } 11.299]}{=} \det(L_1 \circ L_2) = \det(L_1) \cdot \det(L_2) \neq 0$$

proving that

$$\forall L_1.L_2 \in \mathrm{GL}(X) \text{ we have } L_1 \circ L_2 \in \mathrm{GL}(X)$$

Further as $\det(\mathrm{Id}_X) \underset{[\text{theorem: } 11.299]}{=} 1 \neq 0$ we have

$$\mathrm{Id}_X \in \mathrm{GL}(X) \text{ where } \mathrm{Id}_X \text{ is the neutral element in } \langle \operatorname{Hom}(X, X) \rangle$$

So by [definition: 4.12]

$$\mathrm{GL}(X) \text{ is a sub-semi-group of } \langle \operatorname{Hom}(X, X), \circ \rangle$$

Using [theorem: 4.14] that

$$\langle \mathrm{GL}(X), \circ \rangle \text{ is a semi-group}$$

Finally if $L \in \mathrm{GL}(X)$ then L is a linear isomorphism [see theorem: 11.299(2)], so that L^{-1} by [theorem: 11.169] L^{-1} is a linear isomorphism, hence $L^{-1} \in \mathrm{GL}(X)$. Which as $L \circ L^{-1} = \mathrm{Id}_X = L^{-1} \circ L$ proves that

$$\langle \mathrm{GL}(X), \circ \rangle \text{ is a group}$$

□

As $\langle \mathrm{GL}(X), \circ \rangle$ is a group, hence a semi-group we can use [remark: 11.2] to define the finite product of elements in $\mathrm{GL}(X)$ defined by the 'product' operator $\circ$.

We will now introduce a special class of nonsingular operators called elementary transformations.

Definition 11.371. *Let $n \in \mathbb{N}$, X a finite dimensional vector space with $\dim(X) = n$ over a field F and $E = \{e_1, \ldots, e_n\}$ a basis for X then we define:*

1. *Let $i \in \{1, \ldots, n\}$ and $\alpha \in F$ define*

$$M^n_{[i,\alpha]} \in \mathcal{M}_{n,n}(F) \textit{ by } \forall k, l \in \{1, \ldots, n\} \ (M^n_{[i,\alpha]})_{k,l} = \begin{cases} \delta_{k,l} \textit{ if } l \in \{1, \ldots, n\} \backslash \{i\} \\ \alpha \cdot \delta_{k,i} \textit{ if } l = i \end{cases}$$

and

$$\sigma^n_{[i,\alpha]} \in \operatorname{Hom}(X, X) \textit{ by } \sigma^n_{[i,\alpha]} = \mathcal{M}(E, E)^{-1}(M^n_{[i,\alpha]})$$

So that $\forall k \in \{1, \ldots, n\} \setminus \{i\}$ *we have*

$$\begin{aligned}\sigma^n_{[i,\alpha]}(e_k) &= \sum_{j\in\{1,\ldots,n\}} (M^n_{[i,\alpha]})_{j,k} \cdot e_j \\ &= \sum_{j\in\{1,\ldots,n\}} \delta_{j,k} \cdot e_j \\ &= e_k\end{aligned}$$

and

$$\begin{aligned}\sigma^n_{[i,\alpha]}(e_i) &= \sum_{j\in\{1,\ldots,n\}} (M^n_{[i,\alpha]})_{j,i} \cdot e_j \\ &= \sum_{j\in\{1,\ldots,n\}} \alpha \cdot \delta_{j,i} \cdot e_j \\ &= \alpha \cdot e_i\end{aligned}$$

or more compact

$$\forall k \in \{1, \ldots, n\} \text{ we have } \sigma^n_{[i,\alpha]}(e_k) = \begin{cases} e_k \text{ if } k \in \{1, \ldots, n\} \setminus \{i\} \\ \alpha \cdot e_i \text{ if } k = i \end{cases}$$

2. *Let* $i, j \in \{1, \ldots, n\}$ *with* $i \neq j$ *and define*

$$N^n_{[i,j]} \in \mathcal{M}_{n,n}(F) \text{ by } \forall k, l \in \{1, \ldots, n\}\ (N^n_{[i,j]})_{k,l} = \begin{cases} \delta_{j,k} \text{ if } l = i \\ \delta_{i,k} \text{ if } l = j \\ \delta_{k,l} \text{ if } l \in \{1, \ldots, n\} \setminus \{i, j\} \end{cases}$$

and

$$\tau^n_{[i,j]} \in \mathrm{Hom}(X, X) \text{ by } \tau^n_{[i,j]} = \mathcal{M}(E, E)^{-1}(N^n_{[i,j]})$$

So that $\forall k \in \{1, \ldots, n\} \setminus \{i, j\}$ *we have*

$$\begin{aligned}\tau^n_{[i,j]}(e_k) &= \sum_{l\in\{1,\ldots,n\}} (N^n_{[i,j]})_{l,k} \cdot e_l \\ &= \sum_{l\in\{1,\ldots,n\}} \delta_{l,k} \cdot e_l \\ &= e_k\end{aligned}$$

and

$$\begin{aligned}\tau^n_{[i,j]}(e_i) &= \sum_{l\in\{1,\ldots,n\}} (N^n_{[i,j]})_{l,i} \cdot e_l \\ &= \sum_{l\in\{1,\ldots,n\}} \delta_{j,l} \cdot e_l \\ &= e_j\end{aligned}$$

and

$$\begin{aligned}\tau^n_{[i,j]}(e_j) &= \sum_{l\in\{1,\ldots,n\}} (N^n_{[i,j]})_{l,j} \cdot e_l \\ &= \sum_{l\in\{1,\ldots,n\}} \delta_{i,l} \cdot e_l \\ &= e_i\end{aligned}$$

proving that

$$\forall k \in \{1, \ldots, n\}\ \tau^n_{[i,j]}(e_k) = \begin{cases} e_i \text{ if } k = j \\ e_j \text{ if } k = i \\ e_k \text{ if } k \in \{1, \ldots, n\} \setminus \{i, j\} \end{cases}$$

3. *Let* $i, j \in \{1, \ldots, n\}$ *with* $i \neq j$, $\alpha \in F$ *and define*

$$O^n_{[i,j,\alpha]} \in \mathcal{M}_{n,n}(F) \text{ by } \forall k, l \in \{1, \ldots, n\}\ (O^n_{[i,j,\alpha]})_{k,l} = \begin{cases} \delta_{k.i} + \alpha \cdot \delta_{j,k} \text{ if } l = i \\ \delta_{k,l} \text{ if } l \in \{1, \ldots, n\} \setminus \{i\} \end{cases}$$

and

$$\beta^n_{[i,j,\alpha]} \in \operatorname{Hom}(X,X) \textit{ by } \beta^n_{[a,i,j]} = \mathcal{M}[E,E]^{-1}(O^n_{[i,j,\alpha]})$$

So that $\forall k \in \{1,\ldots,n\} \setminus \{i\}$ *we have*

$$\begin{aligned}\beta^n_{[i,j,\alpha]}(e_k) &= \sum_{l\in\{1,\ldots,n\}} (O^n_{[i,j,\alpha]})_{l,k} \cdot e_l \\ &= \sum_{l\in\{1,\ldots,n\}} \delta_{l,k} \cdot e_l \\ &= e_k\end{aligned}$$

and

$$\begin{aligned}\beta^n_{[i,j,\alpha]}(e_i) &= \sum_{l\in\{1,\ldots,n\}} (O^n_{[i,j,\alpha]})_{l,i} \cdot e_l \\ &= \sum_{l\in\{1,\ldots,n\}} (\delta_{l,i} + \alpha \cdot \delta_{j,l}) \cdot e_l \\ &= \sum_{l\in\{1,\ldots,n\}} \delta_{l,i} \cdot e_l + \sum_{l\in\{1,\ldots,n\}} \alpha \cdot \delta_{j,l} \cdot e_l \\ &= e_i + \alpha \cdot e_j\end{aligned}$$

or in other words

$$\forall k \in \{1,\ldots,n\} \textit{ we have } \beta^n_{[i,j,\alpha]}(e_k) = \begin{cases} e_k \textit{ if } k \in \{1,\ldots,n\} \setminus \{i\} \\ e_i + \alpha \cdot e_j \textit{ if } k = i \end{cases}$$

The above transformations have the following properties

Theorem 11.372. *Let* $n \in \mathbb{N}$, X *a finite dimensional vector space with* $\dim(X) = n$ *over a field* F *with characteristic zero then we have:*

1. $\forall i \in \{1,\ldots,n\}$, $\forall \alpha \in F$ *we have* $\det(\sigma^n_{[i,\alpha]}) = \alpha$. *Hence by [theorem: 11.299]*
$$\sigma^n_{[i,\alpha]} \textit{ is a linear isomorphism} \Leftrightarrow \alpha \neq 0$$
2. $\forall i,j \in \{1,\ldots,n\}$ *with* $i \neq j$ *we have*
$$\det(\tau^n_{[i,j]}) = -1$$
so that by [theorem: 11.299] $\tau^n_{[i,j]}$ *is a linear isomorphism.*
3. $\forall i,j \in \{1,\ldots,n\}$ *with* $i \neq j$ *and* $\forall \alpha \in F$ *we have that* $\det(\beta^n_{[i,j,\alpha]}) = 1$ *so that by [theorem: 11.299]* $\beta^n_{[i,j,\alpha]}$ *is a linear isomorphism.*
4. $\forall i,j \in \{1,\ldots,n\}$ *with* $i \neq j$ *we have* $\tau^n_{[i,j]} \circ \tau^n_{[i,j]} = \mathrm{Id}_{\{1,\ldots,n\}}$ *so that* $(\tau^n_{[i,j]})^{-1} = \tau^n_{[i,j]}$

Proof. Let $\{e_1,\ldots,e_n\}$ be a basis for X and let $\Delta: X^n \rightarrow F$ be the determinant function such that $\Delta(e_1,\ldots,e_n) = 1$ [see theorem: 11.283] then we have:

1.
$$\begin{aligned}\det(\sigma^n_{[i,\alpha]}) &\underset{\text{[theorem: 11.296]}}{=} \Delta(\sigma^n_{[i,\alpha]}(e_1),\ldots,\sigma^n_{[i,\alpha]}(e_n)) \\ &= \Delta\left(\sigma^n_{[i,\alpha]}(e_1),\ldots \underbrace{\sigma^n_{[i,\alpha]}(e_i)}_{i} \ldots, \sigma^n_{[i,\alpha]}(e_n)\right) \\ &= \Delta\left(e_1,\ldots \underbrace{\alpha \cdot e_i}_{i} \ldots, e_n\right) \\ &= \alpha \cdot \Delta(e_1,\ldots,e_n) \\ &= \alpha \cdot 1 \\ &= \alpha\end{aligned}$$

2. Let $k \in \{1,\ldots,n\}$ then
$$\tau^n_{[i,j]}(e_k) \underset{\text{[definition: 11.371]}}{=} \begin{cases} e_i \text{ if } k = j \\ e_j \text{ if } k = i \\ e_k \text{ if } k \in \{1,\ldots,n\} \setminus \{i,j\} \end{cases} \underset{\text{[definition: 11.212]}}{=} e_{(i \underset{n}{\leftrightarrow} j)(k)}$$

so that

$$\begin{array}{rcl}\det(\tau^n_{[i,j]}) & \underset{[\text{theorem: } 11.296]}{=} & \Delta(\tau^n_{[i,j]}(e_1),\ldots,\tau^n_{[i,j]}(e_n)) \\ & = & \Delta\left(e_{(i \underset{n}{\leftrightarrow} j)(1)},\ldots,e_{(i \underset{n}{\leftrightarrow} j)(n)}\right) \\ & = & \left((i \underset{n}{\leftrightarrow} j)\Delta\right)(e_1,\ldots,e_n) \\ & \underset{[\text{theorem: } 11.269]}{=} & \operatorname{sign}\left(i \underset{n}{\leftrightarrow} j\right)\cdot\Delta(e_1,\ldots,e_n) \\ & \underset{i\neq j\wedge[\text{theorem:} 11.235]}{=} & (-1)\cdot\Delta(e_1,\ldots,e_n) \\ & = & -1\end{array}$$

3. We have

$$\begin{array}{rc}\det(\beta^n_{[i,j,\alpha]}) & \underset{[\text{theorem: } 11.296]}{=} \\ \Delta(\beta^n_{[i,j,\alpha]}(e_1),\ldots,\beta^n_{[i,j,\alpha]}(e_n)) & = \\ \Delta\left(\beta^n_{[i,j,\alpha]}(e_1)\ldots\underbrace{\beta^n_{[i,j,\alpha]}(e_i)}_{i}\ldots,\beta^n_{[i,j,\alpha]}(e_n)\right) & \underset{[\text{theorem: } 11.296]}{=} \\ \Delta\left(e_1,\ldots\underbrace{e_i+\alpha\cdot e_j}_{i}\ldots,e_n\right) & = \\ \Delta\left(e_1,\ldots\underbrace{e_i}_{i}\ldots,e_n\right)+\alpha\cdot\Delta\left(e_1,\ldots\underbrace{e_j}_{i}\ldots,e_n\right) & = \\ \Delta(e_1,\ldots,e_n)+\alpha\cdot\Delta\left(e_1,\ldots\underbrace{e_j}_{i}\ldots,e_n\right) & \underset{i\neq j\wedge[\text{theorem: } 11.275]}{=} \\ 1+\alpha\cdot 0 & \\ 1 & \end{array}$$

4. Let $k\in\{1,\ldots,n\}$ then we have

$$\begin{array}{rcl}\tau^n_{[i,j]}(\tau^n_{[i,j]}(e_k)) & = & \begin{cases}\tau^n_{[i,j]}(e_j) \text{ if } k=i \\ \tau^n_{[i,j]}(e_i) \text{ if } k=j \\ \tau^n_{[i,j]}(e_k) \text{ if } k\in\{1,\ldots,n\}\setminus\{i,j\}\end{cases} \\ & = & \begin{cases}e_i \text{ if } k=i \\ e_j \text{ if } k=j \\ e_k \text{ if } k\in\{1,\ldots,n\}\setminus\{i,j\}\end{cases} \\ & = & e_k\end{array}$$

Let $x\in X$ then there exists a $\{x_i\}_{i\in\{1,\ldots,n\}}\subseteq F$ such that $x=\sum_{i\in\{1,\ldots,n\}} x_i\cdot e_i$ so that

$$(\tau^n_{[i,j]}\circ\tau^n_{[i,j]})(x)=\sum_{i\in\{1,\ldots,n\}} x_i\cdot(\tau^n_{[i,j]}\circ\tau^n_{[i,j]})(e_i)=\sum_{i\in\{1,\ldots,n\}} x_i\cdot e_i=x$$

proving that

$$\tau^n_{[i,j]}\circ\tau^n_{[i,j]}=\mathrm{Id}_{\{1,\ldots,n\}}$$

□

We are now ready to define elementary transformations:

Definition 11.373. *Let $n\in\mathbb{N}$, X a finite dimensional vector space with $\dim(X)=n$ over a field F with characteristic zero then a nonsingular transformation $T\in\mathrm{Hom}(X,X)$ is a* ***elementary transformation in X*** *if T is either of the form*

$$\begin{array}{rcl}T & = & \mathrm{Id}_X \\ T & = & \sigma^n_{[i,\alpha]} \textit{ where } i\in\{1,\ldots,n\}\wedge\alpha\in F \\ T & = & \tau^n_{[i,j]} \textit{ where } i,j\in\{1,\ldots,n\} \textit{ with } i\neq j \\ T & = & \beta^n_{[i,j,\alpha]} \textit{ where } i,j\in\{1,\ldots,n\} \textit{ with } i\neq j \textit{ and } \alpha\in F\end{array}$$

The set of all the elementary transformations is noted by $\mathrm{Elem}(X)$, hence

$$\mathrm{Elem}(X) \;=\; \{T\in\mathrm{Hom}(X,X)|T \textit{ is a elementary transformation}\}$$

We say that a linear transformation $L\in\mathrm{Hom}(X,Y)$ *is* ***composited of elementary transformations in X*** *if there exists a family* $\{T_i\}_{i\in\{1,\ldots,m\}}\subseteq\mathrm{Elem}(X)$*,* $m\in\mathbb{N}$ *such that*

$$L=\prod_{i=1}^{n} T_1$$

Lemma 11.374. *Let* X *be a finite dimensional vector space over a field* F *with characteristic zero with* $\dim(X)=n+1$*,* $n\in\mathbb{N}$*,* $E=\{e_1,\ldots,e_{n+1}\}$ *a basis for* X*,* $m\in\{1,\ldots,n\}$ *and* $\{\alpha_i\}_{i\in\{1,\ldots,m\}}\subseteq F$ *then we have:*

1. $\forall k\in\{1,\ldots,n\}$ *we have*

$$\left(\prod_{i=1}^{m}\beta^{n+1}_{[n+1,i,\alpha_i]}\right)(e_k)=e_k$$

2.

$$\left(\prod_{i=1}^{m}\beta^{n+1}_{[n+1,i,\alpha_i]}\right)(e_{n+1})=\sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot e_i+e_{n+1}$$

In other words $\forall k\in\{1,\ldots,n+1\}$ *we have*

$$\left(\prod_{i=1}^{m}\beta^{n+1}_{[n+1,i,\alpha_i]}\right)(e_k)=\begin{cases} e_k \text{ if } k\in\{1,\ldots,n\}\\ \sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot e_i+e_{n+1} \text{ if } k=n+1\end{cases}$$

Proof.

1. We prove this by induction, so given $k\in\{1,\ldots,n\}$ let

$$S_{n,k}=\left\{m\in\mathbb{N}|\text{If } m\leqslant n \text{ then } \forall\{\alpha_i\}_{i\in\{1,\ldots,m\}} \text{ we have } \left(\prod_{i=1}^{m}\beta^{n+1}_{[n+1,i,\alpha_i]}\right)(e_k)=e_k\right\}$$

then:

$\boldsymbol{1\in S_{n,k}}$**.** Let $\{\alpha_i\}_{i\in\{1,\ldots,1\}}\subseteq F$ then

$$\left(\prod_{i=1}^{1}\beta^{n+1}_{[n+1,i,\alpha_i]}\right)(e_k)=\beta^{n+1}_{[n+1,1,\alpha_1]}(e_k)\underset{k\neq n+1\wedge[\text{definition: } 11.371]}{=}e_k$$

proving that $1\in S_{n,k}$.

$\boldsymbol{m\in S_{n,k}\Rightarrow m+1\in S_{n,k}}$**.** Assume that $m+1\leqslant n$ and let $\{\alpha_i\}_{i\in\{1,\ldots,m+1\}}\subseteq F$ then we have:

$$\begin{aligned}\left(\prod_{i=1}^{m+1}\beta^{n+1}_{[n+1,i,\alpha_i]}\right)(e_k) &= \left(\prod_{i=1}^{m}\beta^{n+1}_{[n+1,i,\alpha_i]}\circ\beta^{n+1}_{[n+1,m+1,\alpha_{m+1}]}\right)(e_k)\\ &= \left(\prod_{i=1}^{m}\beta^{n+1}_{[n+1,i,\alpha_i]}\right)(\beta^{n+1}_{[n+1,m+1,\alpha_{m+1}]}(e_k))\\ &\underset{k\neq n+1[\text{theorem: } 11.371]}{=} \left(\prod_{i=1}^{m}\beta^{n+1}_{[n+1,i,\alpha_i]}\right)(e_k)\\ &\underset{m\in S_{n,k}}{=} e_k\end{aligned}$$

proving that $m+1\in S_{n,k}$.

2. We prove this by induction, so let

$$S_n=\left\{m\in\mathbb{N}|\text{If } m\leqslant n \text{ then } \forall\{\alpha_i\}_{i\in\{1,\ldots,m\}}\subseteq F \text{ we have } \left(\prod_{i=1}^{m}\beta^{n+1}_{[n+1,i,\alpha_i]}\right)(e_{n+1})=\sum_{i\in\{1,\ldots,m\}}\alpha_i\cdot e_i+e_{n+1}\right\}$$

then we have:

$\boldsymbol{1\in S_n}$**.** If $\{\alpha_i\}_{i\in\{1\}}$ then

$$\begin{aligned}\left(\prod_{i=1}^{1}\beta^{n+1}_{[n+1,i,\alpha_i]}\right)(e_{n+1}) &= \beta^{n+1}_{[n+1,1,\alpha_1]}(e_{n+1})\\ &\underset{k=n+1[\text{definition: } 11.371]}{=} e_{n+1}+\alpha_1\cdot e_1\\ &= \sum_{i\in\{1\}}\alpha_i\cdot e_i+e_{n+1}\end{aligned}$$

proving that $1 \in S_n$.

$m \in S_n \Rightarrow m+1 \in S_n$. Assume that $m+1 \leqslant n$ and let $\{\alpha_i\}_{i \in \{1,\ldots,m+1\}} \subseteq F$ then we have

$$
\begin{aligned}
\left(\prod_{i=1}^{m+1} \beta^{n+1}_{[n+1,i,\alpha_i]}\right)(e_{n+1}) &= \\
\left(\prod_{i=1}^{m} \beta^{n+1}_{[n+1,i,\alpha_i]} \circ \beta^{n+1}_{[n+1,m+1,\alpha_{m+1}]}\right)(e_{n+1}) &= \\
\prod_{i=1}^{m} \beta^{n+1}_{[n+1,i,\alpha_i]}(\beta^{n+1}_{[n+1,m+1,\alpha_{m+1}]}(e_{n+1})) &\underset{k=n+1 \wedge [\text{definition: } 11.371]}{=} \\
\left(\prod_{i=1}^{m} \beta^{n+1}_{[n+1,i,\alpha_i]}\right)(e_{n+1} + \alpha_{m+1} \cdot e_{m+1}) &= \\
\left(\prod_{i=1}^{m} \beta^{n+1}_{[n+1,i,\alpha_i]}\right)(e_{n+1}) + \alpha_{m+1} \cdot \left(\prod_{i=1}^{m} \beta^{n+1}_{[n+1,i,\alpha_i]}\right)(e_{m+1}) &\underset{m+1 \leqslant n < n+1 \text{ and } (1)}{=} \\
\left(\prod_{i=1}^{m} \beta^{n+1}_{[n+1,i,\alpha_i]}\right)(e_{n+1}) + \alpha_{m+1} \cdot e_{m+1} &\underset{m \in S_n}{=} \\
\sum_{i=1}^{m} \alpha_i \cdot e_i + e_{n+1} + \alpha_{m+1} \cdot e_{m+1} &= \\
\sum_{i=1}^{m+1} \alpha_i \cdot e_i + e_{n+1}
\end{aligned}
$$

proving that $n+1 \in S_n$. □

Lemma 11.375. *Let X a finite dimensional vector space with $\dim(X) = n+1$ over a field F with characteristic zero, $E = \{e_1, \ldots, e_{n+1}\}$ a basis for X, $m \in \{1, \ldots, n\}$ and $\{\alpha_i\}_{i \in \{1,\ldots,m\}} \subseteq F$ then we have:*

1. *$\forall k \in \{m+1, \ldots, n+1\}$ we have*
$$\left(\prod_{i=1}^{m} \beta^{n+1}_{[i,n+1,\alpha_i]}\right)(e_k) = e_k$$
2. *$\forall k \in \{1, \ldots, m\}$ we have*
$$\left(\prod_{i=1}^{m} \beta^{n+1}_{[i,n+1,\alpha_i]}\right)(e_k) = e_k + \alpha_k \cdot e_{n+1}$$

In other words we have $\forall k \in \{1, \ldots, n+1\}$ that

$$\left(\prod_{i=1}^{m} \beta^{n+1}_{[i,n+1,\alpha_i]}\right)(e_k) = \begin{cases} e_k + \alpha_k \cdot e_{n+1} \text{ if } k \in \{1, \ldots, m\} \\ e_k \text{ if } k \in \{m+1, \ldots, n+1\} \end{cases}$$

Proof.

1. We use induction on m for this proof, so let

$$S_n = \left\{ m \in \mathbb{N} \middle| \text{If } m \leqslant n \text{ then } \forall \{\alpha_i\}_{i \in \{1,\ldots,m\}} \text{ and } \forall k \in \{m+1, \ldots, n+1\} \text{ we have } \left(\prod_{i=1}^{m} \beta^{n+1}_{[i,n+1,\alpha_i]}\right)(e_k) = e_k \right\}$$

then we have:

$1 \in S_n$. Let $\{\alpha_i\}_{i \in \{1\}} \subseteq F$ then for $k \in \{1+1, \ldots, n+1\} \Rightarrow k \neq 1$ we have

$$\left(\prod_{i=1}^{1} \beta^{n+1}_{[i,n+1,\alpha_i]}\right)(e_k) = \beta^{n+1}_{[1,n+1,\alpha_1]}(e_k) \underset{k \neq 1 [\text{definition: } 11.371]}{=} e_k$$

proving that $1 \in S_n$.

$m\in S_n\Rightarrow m+1\in S_n$. Assume that $m+1\leqslant n$ and let $\{\alpha_i\}_{i\in\{1,\ldots,m+1\}}\subseteq F$ then we have for $k\in\{(m+1)+1,\ldots,n+1\}\subseteq\{m+1,\ldots,n+1\}\Rightarrow k\neq m+1$ and $k\in\{m+1,\ldots,n+1\}$. Further

$$\begin{aligned}\left(\prod_{i=1}^{m+1}\beta^{n+1}_{[i,n+1,\alpha_i]}\right)(e_k) &=\\ \left(\prod_{i=1}^{m}\beta^{n+1}_{[i,n+1,\alpha_i]}\circ\beta^{n+1}_{[m+1,n+1,\alpha_{m+1}]}\right)(e_k)\\ \left(\prod_{i=1}^{m}\beta^{n+1}_{[i,n+1,\alpha_i]}\right)(\beta^{n+1}_{[m+1,n+1,\alpha_{m+1}]}(e_k)) &\underset{k\neq 1m+1\wedge[\text{definition: }11.371]}{=}\\ \left(\prod_{i=1}^{m}\beta^{n+1}_{[i,n+1,\alpha_i]}\right)(e_k) &\underset{m\in S_n\wedge k\in\{m+1,\ldots,n+1\}}{=}\\ e_k\end{aligned}$$

proving that $m+1\in S_n$.

2. This is proved by induction, so let

$$S_n=\left\{m\in\mathbb{N}|\text{If } m\leqslant n \text{ then } \forall\{\alpha_i\}_{i\in\{1,\ldots,n\}}\subseteq F \text{ and } \forall k\in\{1,\ldots,m\} \text{ we have } \left(\prod_{i=1}^{m}\beta^{n+1}_{[i,n+1,\alpha_i]}\right)(e_k)=e_k+\alpha_k\cdot e_{n+1}\right\}$$

then we have:

$1\in S$. Let $\{\alpha_i\}_{i\in\{1\}}\subseteq F$ then for $k\in\{1\}\Rightarrow k=1$ we have

$$\begin{aligned}\left(\prod_{i=1}^{1}\beta^{n+1}_{[i,n+1,\alpha_i]}\right)(e_1) &= (\beta^{n+1}_{[1,n+1,\alpha_1]})(e_1)\\ &\underset{k=1\wedge[\text{definition: }11.371]}{=} e_1+\alpha_1\cdot e_{n+1}\\ &\underset{k=1}{=} e_k+\alpha_k\cdot e_{n+1}\end{aligned}$$

proving that $1\in S_n$.

$m\in S\Rightarrow m+1\in S$. Assume that $m+1\leqslant n$ and let $\{\alpha_i\}_{i\in\{1,\ldots,m+1\}}\subseteq F$ then for $k\in\{1,\ldots,m+1\}$ we have either:

$k=m+1$. Then

$$\begin{aligned}\left(\prod_{i=1}^{m+1}\beta^{n+1}_{[i,n+1,\alpha_i]}\right)(e_k) &=\\ \left(\prod_{i=1}^{m+1}\beta^{n+1}_{[i,n+1,\alpha_i]}\right)(e_{m+1}) &=\\ \left(\prod_{i=1}^{m}\beta^{n+1}_{[i,n+1,\alpha_i]}\circ\beta^{n+1}_{[m+1,n+1,\alpha_{m+1}]}\right)(e_{m+1}) &=\\ \left(\prod_{i=1}^{m}\beta^{n+1}_{[i,n+1,\alpha_i]}\right)(\beta^{n+1}_{[m+1,n+1,\alpha_{m+1}]}(e_{m+1})) &\underset{[\text{definition: }11.371]}{=}\\ \left(\prod_{i=1}^{m}\beta^{n+1}_{[i,n+1,\alpha_i]}\right)(e_{m+1}+\alpha_{m+1}\cdot e_{n+1}) &=\\ \left(\prod_{i=1}^{m}\beta^{n+1}_{[i,n+1,\alpha_i]}\right)(e_{m+1})+\alpha_{m+1}\cdot\left(\prod_{i=1}^{m}\beta^{n+1}_{[i,n+1,\alpha_i]}\right)(e_{n+1}) &\underset{(1)\wedge m+1,n+1\in\{m+1,\ldots n+1\}}{=}\\ e_{m+1}+\alpha_{m+1}\cdot e_{n+1} &\underset{m+1=k}{=}\\ e_k+\alpha_k\cdot e_{n+1}\end{aligned}$$

$\boldsymbol{k \in \{1,\ldots,m\}}$**.** Then

$$\left(\prod_{i=1}^{m+1} \beta^{n+1}_{[i,n+1,\alpha_i]}\right)(e_k) = \left(\prod_{i=1}^{m} \beta^{n+1}_{[i,n+1,\alpha_i]} \circ \beta^{n+1}_{[m+1,n+1,\alpha_{m+1}]}\right)(e_k) = \left(\prod_{i=1}^{m} \beta^{n+1}_{[i,n+1,\alpha_i]}\right)(\beta^{n+1}_{[m+1,n+1,\alpha_{m+1}]}(e_k)) \underset{k\neq m+1\wedge[\text{definition: } 11.371]}{=} \left(\prod_{i=1}^{m} \beta^{n+1}_{[i,n+1,\alpha_i]}\right)(e_k) \underset{k\in\{1,\ldots,m\}\wedge m\in S_n}{=} e_k+\alpha_k\cdot e_{n+1}$$

So in all cases we have $(\prod_{i=1}^{m+1}\beta^{n+1}_{[i,n+1,\alpha_i]})(e_k)=e_k+\alpha_k\cdot e_{n+1}$ proving that

$$m+1\in S_n.$$

□

We have the following trivial lemma about transformations that are compositions of elementary transformations.

Lemma 11.376. *Let $n\in\mathbb{N}$, X a finite dimensional vector space with $\dim(X)=n$ over a field F with characteristic zero then we have:*

1. *If $T\in\mathrm{GL}(X)$ is a elementary transformation and $L\in\mathrm{GL}(X)$ is composed of elementary transformations then $L\circ T$ and $T\circ L$ are composed of linear transformations*
2. *If $L_1,L_2\in\mathrm{GL}(X)$ are composed of elementary transformations then $L_1\circ L_2$ is composed of elementary transformations.*

Proof.

1. As L is composed of elementary transformations there exists a $m\in\mathbb{N}$ and $\{E_i\}_{i\in\{1,\ldots,m\}}\subseteq\mathrm{Elem}(X)$ such that $L=\prod_{i=1}^m E_i$. Define now

$$\{S_i\}_{i\in\{1,\ldots,m+1\}}\subseteq\mathrm{Elem}(X) \text{ by } S_i=\begin{cases} T \text{ if } i=1\\ E_{i-1} \text{ if } i\in\{2,\ldots,m+1\}\end{cases}$$

and

$$\{T_i\}_{i\in\{1,\ldots,m+1\}}\subseteq\mathrm{Elem}(X) \text{ by } T_i=\begin{cases} E_i \text{ if } i\in\{1,\ldots,m\}\\ T \text{ if } i=m+1\end{cases}$$

Then we have

$$\begin{aligned}\prod_{i=1}^{m+1} S_i &\underset{[\text{definition: } 11.12]}{=} \prod_{i\in 0}^{m} S_{i+1}\\ &\underset{[\text{theorem: } 11.22]}{=} S_{0+1}\circ\prod_{i=1}^m S_{i+1}\\ &= S_1\circ\prod_{i=1}^m S_{i+1}\\ &= T\circ\prod_{i=1}^m E_{(i-1)+1}\\ &= T\circ\prod_{i=1}^m E_i\\ &= T\circ L\end{aligned}$$

and

$$\begin{aligned}\prod_{i=1}^{m+1} T_i &\underset{[[\text{theorem: } 11.15]}{=} \left(\prod_{i=1}^m T_i\right)\circ T_{m+1}\\ &= \left(\prod_{i=1}^m E_i\right)\circ T\\ &= L\circ T\end{aligned}$$

proving that $L \circ T$ and $T \circ L$ are composed of elementary transformations.

2. If L_1, L_2 are composed of elementary transformations then there exists $k, l \in \mathbb{N}$ and families of elementary transformations $\{T_i\}_{i \in \{1, \ldots, k\}} \subseteq \mathrm{Elem}(X)$, $\{S_i\}_{i \in \{1, \ldots, l\}} \subseteq \mathrm{Elem}(X)$ such that

$$L_1 = \prod_{i=1}^{k} T_i \text{ and } L_2 = \prod_{i=1}^{l} S_i$$

Define then the family $\{E_i\}_{i \in \{1, \ldots, k+l\}} \subseteq \mathrm{GL}(X)$ by

$$\forall i \in \{1, \ldots, k+l\}\ E_i = \begin{cases} T_i \text{ if } i \in \{1, \ldots, k\} \\ S_{i-k} \text{ if } i \in \{k+1, \ldots, k+l\} \end{cases}$$

then we have

$$\begin{aligned} L_1 \circ L_2 &= \prod_{i=1}^{k} T_i \circ \prod_{i=1}^{l} S_i \\ &= \prod_{i=1}^{k} T_i \circ \prod_{i=1}^{l} S_i \\ &\underset{[\text{theorem: } 11.14]}{=} \prod_{i=1}^{k} T_i \circ \prod_{i=k+1}^{k+l} S_{i-k} \\ &= \prod_{i=1}^{k} E_i \circ \prod_{i=k+1}^{k+l} E_i \\ &\underset{[\text{theorem: } 11.23]}{=} \prod_{i=1}^{k+l} E_i \end{aligned}$$

proving that $L_1 \circ L_2$ is composed of elementary transpositions. □

To use mathematical induction on the dimension of a vector space we introduce the following definition.

Definition 11.377. *Let $n \in \mathbb{N}$, X a finite dimensional vector space over a field $\mathcal{F}$ with basis $E = \{e_1, \ldots, e_{n+1}\}$, $Y = \mathrm{span}(e_1, \ldots, e_n)$ the vector space with basis $F = \{e_1, \ldots, e_n\}$ and $L \in \mathrm{Hom}(Y, Y)$ then we define $M_{[n,L]} \in \mathcal{M}_{n+1,n+1}(F)$ by*

$$\forall k, l \in \{1, \ldots, n+1\} \text{ we have } (M_{[n+1,L]})_{k,l} = \begin{cases} \mathcal{M}(L; F, F)_{k,l} \text{ if } (k,l) \in \{1, \ldots, n\} \times \{1, \ldots, n\} \\ \delta_{k,l} \text{ if } k = n+1 \vee l = n+1 \end{cases}$$

and define $L^{[n]} \in \mathrm{Hom}(X, X)$ by

$$L^{[n+1]} = \mathcal{M}(E, E)^{-1}(M_{[n+1,L]})$$

then we have $\forall k \in \{1, \ldots, n\}$ that

$$\begin{aligned} L^{[n+1]}(e_k) &= \sum_{i \in \{1, \ldots, n+1\}} (M_{[n+1,L]})_{i,k} \cdot e_i \\ &= \sum_{i \in \{1, \ldots, n\}} (M_{[n+1,L]})_{i,k} \cdot e_i + \sum_{i \in \{n+1\}} (M_{[n+1,L]})_{i,k} \cdot e_i \\ &= \sum_{i \in \{1, \ldots, n\}} (M_{[n+1,L]})_{i,k} \cdot e_i + (M_{[n+1,L]})_{n+1,k} \cdot e_{n+1} \\ &= \sum_{i \in \{1, \ldots, n\}} \mathcal{M}(L; F, F)_{i,k} \cdot e_i + \delta_{n+1,k} \cdot e_{n+1} \\ &= L(e_k) + \delta_{n+1,k} \cdot e_{n+1} \\ &\underset{k \neq n+1}{=} L(e_k) + 0 \\ &\underset{k \neq n}{=} L(e_k) \end{aligned}$$

and

$$\begin{array}{rcl} L^{[n+1]}(e_{n+1}) & = & \sum_{i \in \{1,\ldots,n+1\}} (M_{[n+1,L]})_{i,n+1} \cdot e_i \\ & = & \sum_{i \in \{1,\ldots,n+1\}} \delta_{i,n+1} \cdot e_i \\ & = & e_{n+1} \end{array}$$

proving that

$$\forall k \in \{1,\ldots,n\} \text{ we have } L^{[n+1]}(e_k) = \begin{cases} L(e_k) \text{ if } k \in \{1,\ldots,n\} \\ e_{n+1} \text{ if } k = n+1 \end{cases}$$

Lemma 11.378. *Let $n \in \mathbb{N}$, X a finite dimensional vector space over a field $\mathcal{F}$ with basis $E = \{e_1, \ldots, e_{n+1}\}$, $Y = \operatorname{span}(\{e_1, \ldots, e_n\})$ a n-dimensional vector space with basis $F = \{e_1, \ldots, e_n\}$ then we have:*

1. $(\mathrm{Id}_Y)^{[n+1]} = \mathrm{Id}_X$
2. $\forall i \in \{1, \ldots, n\}$ *and* $\alpha \in F$ *we have* $(\sigma^n_{[i,\alpha]})^{[n+1]} = \sigma^{n+1}_{[i,\alpha]}$
3. $\forall i, j \in \{1, \ldots, n\}$ *with* $i \neq j$ *we have* $(\tau^n_{[i,j]})^{[n+1]} = \tau^{n+1}_{[i,j]}$
4. $\forall i, j \in \{1, \ldots, n\}$ *with* $i \neq j$ *and* $\alpha \in F$ *then* $(\beta^n_{[i,j,\alpha]})^{[n+1]} = \beta^{n+1}_{[i,j,\alpha]}$

in other words if $T \in \mathrm{Elem}(X)$ is a elementary transformation in Y then $T^{[n+1]}$ is a elementary transformation in X.

Proof.

1. Let $k \in \{1, \ldots, n+1\}$ then

$$\begin{array}{rcl} (\mathrm{Id}_X)^{[n+1]}(e_k) & = & \begin{cases} \mathrm{Id}_X(e_k) \text{ if } k \in \{1,\ldots,n\} \\ e_{n+1} \text{ if } k = n+1 \end{cases} \\ & = & \begin{cases} e_k \text{ if } k \in \{1,\ldots,n\} \\ e_{n+1} \text{ if } k = n+1 \end{cases} \\ & = & e_k \\ & = & \mathrm{Id}_Y(e_k) \end{array}$$

   so that by [theorem: 11.320] $(\mathrm{Id}_X)^{[n+1]} = \mathrm{Id}_Y$.

2. Let $k \in \{1, \ldots, n+1\}$ then

$$\begin{array}{rcl} (\sigma^n_{[i,\alpha]})^{[n+1]}(e_k) & = & \begin{cases} \sigma^n_{[i,\alpha]}(e_k) \text{ if } k \in \{1,\ldots,n\} \\ e_{n+1} \text{ if } k = n+1 \end{cases} \\ & \underset{[\text{definition: } 11.371]}{=} & \begin{cases} \alpha \cdot e_i \text{ if } k \in \{1,\ldots,n\} \wedge k = i \\ e_k \text{ if } k \in \{1,\ldots,n\} \setminus \{i\} \\ e_{n+1} \text{ if } k = n+1 \end{cases} \\ & \underset{\{1,\ldots,n\}\setminus\{i\} \subseteq \{1,\ldots,n+1\}\setminus\{i\}}{=} & \begin{cases} \sigma^{n+1}_{[i,\alpha]}(e_k) \text{ if } k \in \{1,\ldots,n\} \\ \sigma^{n+1}_{[i,\alpha]}(e_k) \text{ if } k \in \{1,\ldots,n\} \setminus \{i\} \\ e_{n+1} \text{ if } k = n+1 \end{cases} \\ & \underset{i \neq n+1}{=} & \begin{cases} \sigma^{n+1}_{[i,\alpha]}(e_k) \text{ if } k \in \{1,\ldots,n\} \\ \sigma^{n+1}_{[i,\alpha]}(e_k) \text{ if } k \in \{1,\ldots,n\} \setminus \{i\} \\ \sigma^{n+1}_{[i,\alpha]}(e_k) \text{ if } k = n+1 \end{cases} \\ & = & \sigma^{n+1}_{[i,\alpha]}(e_k) \end{array}$$

so that by [theorem: 11.320] $\sigma^n_{[i,\alpha]} = \sigma^{n+1}_{[i,\alpha]}$.

3. Let $k \in \{1, \ldots, n+1\}$ then

$$\begin{aligned}
(\tau^n_{[i,j]})^{[n+1]}(e_k) &= \begin{cases} \tau^n_{[i,j]}(e_k) \text{ if } k \in \{1, \ldots, n\} \\ e_{n+1} \text{ if } k = n+1 \end{cases} \\
&\underset{[\text{definition: } 11.371]}{=} \begin{cases} e_i \text{ if } k \in \{1, \ldots, n\} \wedge k = j \\ e_j \text{ if } k \in \{1, \ldots, n\} \wedge k = i \\ e_k \text{ if } k \in \{1, \ldots, n\} \setminus \{i, j\} \\ e_{n+1} \text{ if } k = n+1 \end{cases} \\
&\underset{\{1,\ldots,n\}\setminus\{i,j\} \subseteq \{1,\ldots,n+1\}\setminus\{i,j\}}{=} \begin{cases} \tau^{n+1}_{[i,j]}(e_k) \text{ if } k \in \{1, \ldots, n\} \wedge k = j \\ \tau^{n+1}_{[i,j]}(e_k) \text{ if } k \in \{1, \ldots, n\} \wedge k = i \\ \tau^{n+1}_{[i,j]}(e_k) \text{ if } k \in \{1, \ldots, n\} \setminus \{i, j\} \\ e_{n+1} \text{ if } k = n+1 \end{cases} \\
&\underset{i,j \neq n_{+1}}{=} \begin{cases} \tau^{n+1}_{[i,j]}(e_k) \text{ if } k \in \{1, \ldots, n\} \wedge k = j \\ \tau^{n+1}_{[i,j]}(e_k) \text{ if } k \in \{1, \ldots, n\} \wedge k = i \\ \tau^{n+1}_{[i,j]}(e_k) \text{ if } k \in \{1, \ldots, n\} \setminus \{i, j\} \\ \tau^{n+1}_{[i,j]}(e_k) \text{ if } k = n+1 \end{cases} \\
&= \tau^{n+1}_{[i,j]}(e_k)
\end{aligned}$$

proving by [theorem: 11.320] $\tau^n_{[i,j]} = \tau^{n+1}_{[i,j]}$.

4. Let $k \in \{1, \ldots, n+1\}$ then

$$\begin{aligned}
(\beta^n_{[i,j,\alpha]})^{[n+1]}(e_k) &= \begin{cases} \beta^n_{[i,j,\alpha]}(x_k) \text{ if } k \in \{1, \ldots n\} \\ e_{n+1} \text{ if } k = n+1 \end{cases} \\
&\underset{[\text{definition: } 11.371]}{=} \begin{cases} e_i + \alpha \cdot e_j \text{ if } k \in \{1, \ldots, n\} \wedge k = i \\ e_k \text{ if } k \in \{1, \ldots, n\} \setminus \{i\} \\ e_{n+1} \text{ if } k = n+1 \end{cases} \\
&= \begin{cases} \beta^{n+1}_{[i,j,\alpha]}(e_k) \text{ if } k \in \{1, \ldots, n\} \wedge k = i \\ e_k \text{ if } k \in \{1, \ldots, n\} \setminus \{i\} \\ e_{n+1} \text{ if } k = n+1 \end{cases} \\
&\underset{\{1,\ldots,n\}\setminus\{i\} \subseteq \{1,\ldots,n+1\}\setminus\{i\}}{=} \begin{cases} \beta^{n+1}_{[i,j,\alpha]}(e_k) \text{ if } k \in \{1, \ldots, n\} \wedge k = i \\ \beta^{n+1}_{[i,j,\alpha]}(e_k) \text{ if } k \in \{1, \ldots, n\} \setminus \{i\} \\ e_{n+1} \text{ if } k = n+1 \end{cases} \\
&\underset{i \neq n+1}{=} \begin{cases} \beta^{n+1}_{[i,j,\alpha]}(e_k) \text{ if } k \in \{1, \ldots, n\} \wedge k = i \\ \beta^{n+1}_{[i,j,\alpha]}(e_k) \text{ if } k \in \{1, \ldots, n\} \setminus \{i\} \\ \beta^{n+1}_{[i,j,\alpha]}(e_k) \text{ if } k = n+1 \end{cases} \\
&= \beta^{n+1}_{[i,j,\alpha]}(e_k)
\end{aligned}$$

proving by [theorem: 11.320] $\beta^n_{[i,j,\alpha]} = \beta^{n+1}_{[i,j,\alpha]}$. □

Extending a elementary transformation to a higher dimension also applies to nonsingular transformations that are composed of elementary transformation.

Lemma 11.379. *Let $n \in \mathbb{N}$, X a finite dimensional vector space over a field F with basis $E = \{e_1, \ldots, e_{n+1}\}$, $Y = \operatorname{span}(\{e_{1,\ldots,e_n}\})$ a n-dimensional vector space with basis $F = \{e_1, \ldots, e_n\}$ then if $K, L \in \mathrm{GL}(Y)$ we have that*

$$(K \circ L)^{[n+1]} = K^{[n+1]} \circ L^{[n+1]}$$

Proof. Let $k \in \{1, \ldots, n+1\}$ then we have either:

$\boldsymbol{k \in \{1, \ldots, n\}}$**.** Then

$$\begin{array}{rcl}
(K \circ L)^{[n+1]}(e_k) & = & (K \circ L)(e_k) \\
& = & K(L(e_k)) \\
& = & K\left(\sum_{i \in \{1, \ldots, n\}} \mathcal{M}(L; F, F)_{i,k} \cdot e_i\right) \\
& = & \sum_{i \in \{1, \ldots, n\}} \mathcal{M}(L; F, F)_{i,k} \cdot K(e_i) \\
& = & \sum_{i \in \{1, \ldots, n\}} \mathcal{M}(L; F, F)_{i,k} \cdot K^{[n+1]}(e_i) \\
& = & K^{[n+1]}\left(\sum_{i \in \{1, \ldots, n\}} \mathcal{M}(L; F, F)_{i,k} \cdot e_i\right) \\
& = & K^{[n+1]}\left(\sum_{i \in \{1, \ldots, n\}} \mathcal{M}(L; F, F)_{i,k} \cdot e_i + 0 \cdot e_{n+1}\right) \\
& \underset{k \neq n+1}{=} & K^{[n+1]}\left(\sum_{i \in \{1, \ldots, n\}} \mathcal{M}(L; F, F)_{i,k} \cdot e_i + \delta_{n+1,k} \cdot e_{n+1}\right) \\
& \underset{\text{[see: 11.377[}}{=} & K^{[n+1]}\left(\sum_{i \in \{1, \ldots, n\}} \left(M_{[n+1,L]_{i,k}}\right) \cdot e_i + (M_{[n+1,L]})_{n+1,k} \cdot e_{n+1}\right) \\
& = & K^{[n+1]}\left(\sum_{i \in \{1, \ldots, n+1\}} \left(M_{[n+1,L]_{i,k}}\right) \cdot e_i\right) \\
& = & K^{[n+1]}(L^{[n+1]}(e_k)) \\
& = & (K^{[n+1]} \circ L^{[n+1]})(e_k)
\end{array}$$

$\boldsymbol{k = n+1}$**.** Then

$$\begin{array}{rcl}
(K \circ L)^{[n+1]}(e_k) & = & e_{n+1} \\
& = & L^{[n+1]}(e_k) \\
& \underset{L^{[n+1]}(e_k) = e_{n+1}}{=} & K^{[n+1]}(L^{[n+1]}(e_k)) \\
& = & (K^{[n+1]} \circ L^{[n+1]})(e_k)
\end{array}$$

So we have $\forall k \in \{1, \ldots, n+1\}$ that $(K \circ L)^{[n+1]}(e_k) = (K^{[n+1]} \circ L^{[n+1]})(e_k)$ which by [theorem: 11.320] proves that

$$(K \circ L)^{[n+1]} = K^{[n+1]} \circ L^{[n+1]}$$

□

Lemma 11.380. *Let $n \in \mathbb{N}$, X a finite dimensional vector space over a field F with basis $E = \{e_1, \ldots, e_{n+1}\}$, $Y = \operatorname{span}(\{e_1, \ldots, e_n\})$ a n-dimensional vector space with basis $F = \{e_1, \ldots, e_n\}$ then if $m \in \mathbb{N}, \{L_i\}_{i \in \{1, \ldots, m\}} \subseteq \mathrm{GL}(Y)$ we have that*

$$\left(\prod_{i=1}^{m} L_i\right)^{[n+1]} = \prod_{i=1}^{m} (L_i)^{[n+1]}$$

Proof. We prove this by induction, so let

$$S=\left\{m\in\mathbb{N}|\text{If } \{L_i\}_{i\in\{1,\ldots,m\}}\subseteq \mathrm{GL}(Y) \text{ then } \left(\prod_{i=1}^m L_i\right)^{[n+1]}=\prod_{i=1}^m (L_i)^{[n+1]}\right\}$$

then we have:

$1\in S$. Let $\{L_i\}_{i\in\{1,\ldots,1\}}\subseteq \mathrm{GL}(Y)$ then $(\prod_{i=1}^1 L_i)^{[n+1]}=(L_1)^{[n+1]}=\prod_{i=1}^1 (L_i)^{[n+1]}$

$m\in S\Rightarrow m+1\in S$. Let $\{L_i\}_{i\in\{1,\ldots,m+1\}}\subseteq \mathrm{GL}(X)$ then we have

$$\begin{aligned}\left(\prod_{i=1}^{m+1} L_i\right)^{[n+1]} &= \left(\left(\prod_{i=1}^m L_i\right)\circ L_{m+1}\right)^{[n+1]}\\ &\underset{[\text{lemma: } 11.379]}{=} \left(\prod_{i=1}^m L_i\right)^{[n+1]}\circ (L_{m+1})^{[n+1]}\\ &\underset{n\in S}{=} \left(\prod_{i=1}^m (L_i)^{[n+1]}\right)\circ (L_{m+1})^{[n+1]}\\ &= \prod_{i=1}^{m+1} (L_i)^{[n+1]}\end{aligned}$$

proving that $m+1\in S$. □

Corollary 11.381. *Let $n\in\mathbb{N}$, X a finite dimensional vector space over a field F with basis $E=\{e_1,\ldots,e_{n+1}\}$, $Y=\mathrm{span}(\{e_1,\ldots,e_n\})$ a n-dimensional vector space with basis $F=\{e_1,\ldots,e_n\}$ then if $L\in\mathrm{GL}(Y)$ is composed of elementary transformations in Y it follows that $L^{[n+1]}$ is composed of elementary transformations in X.*

Proof. As $L\in\mathrm{GL}(Y)$ is composed of linear transformations there exists a $m\in\mathbb{N}$ and a family $\{T_i\}_{i\in\{1,\ldots,m\}}\subseteq\mathrm{Elem}(Y)$ of elementary transformations such that

$$L=\prod_{i=1}^m T_i$$

Using [lemma: 11.378] we have $\forall i\in\{1,\ldots,m\}$ that $(L_i)^{[n+1]}\in\mathrm{Elem}(X)$. So as

$$L^{[n+1]}\underset{[\text{theorem: } 11.380]}{=}\prod_{i=1}^m (T_i)^{[n+1]}$$

we have that $L^{[n+1]}$ is composed of elementary transformations in X. □

We are now ready to prove that every nonsingular transformation is composed of elementary transformations.

Lemma 11.382. *Let $n\in\mathbb{N}$, X a finite dimensional vector space over a filed F with characteristic zero, $L\in\mathrm{GL}(X)$ a nonsingular transformation then L is composed of elementary transformations. In other words there exists a $m\in\mathbb{N}$ and $\{T_i\}_{i\in\{1,\ldots,m\}}\subseteq\mathrm{Elem}(X)$ such that*

$$L=\prod_{i=1}^m T_i\underset{\text{def}}{=}T_1\circ\cdots\circ T_m$$

Proof. We prove this by induction on the dimension n of X, so let

$$S=\{n\in\mathbb{N}|\text{If } L\in\mathrm{GL}(X) \text{ where } X \text{ is a } n\text{-dimensional vector space over } F \text{ then } L \text{ is composed of elementary transformations in } X\}$$

then we have:

$1\in S$. As X is one dimensional it has a base $\{e_i\}_{i\in\{1,\ldots,1\}}$ then, if $L\in\mathrm{GL}(X)$ we have $L(e_1)=\alpha\cdot e_1$, so $L=\sigma^1_{[1,\alpha]}\in\mathrm{Elem}(X)$. As $L=\sigma^1_{[1,\alpha]}=\prod_{i=1}^1\sigma^1_{[i,\alpha]}$ it follows that $1\in S$.

$n\in S\Rightarrow n+1\in S$. Let $L\in\mathrm{GL}(X)$ where $\dim(X)=n+1$ is a $(n+1)$-dimensional vector space over F and $E=\{e_1,\ldots,e_{n+1}\}$ a basis for X. Let $M=\mathcal{M}(L;E,E)\in\mathcal{M}_{n+1,n+1}(F)$ be the matrix associated with the nonsingular linear transformation. Then

$$\forall i\in\{1,\ldots,n\} \text{ we have } L(e_i)=\sum_{k\in\{1,\ldots,n+1\}} M_{k,i}\cdot e_k \tag{11.269}$$

As $L\in \mathrm{GL}(X)$ we have that $\det(L)\neq 0$ hence

$$\begin{array}{rcl} 0 & \neq & \det(L) \\ & \underset{[\text{theorem: } 11.339]}{=} & \det(M) \\ & \underset{[\text{theorem: } 11.362]}{=} & \displaystyle\sum_{i\in\{1,\ldots,n+1\}} (-1)^{i+(n+1)}\cdot M_{i,n+1}\cdot \det([i\boxplus n+1](M)) \end{array}$$

If $\forall i\in\{1,\ldots,n+1\}$ $\det([i\boxplus n+1](M))=0$ then the above leads to the contradiction $0\neq 0$, hence

$$\exists i_0\in\{1,\ldots,n+1\} \text{ such that } \det([i_0\boxplus n+1](M))\neq 0 \tag{11.270}$$

For i_0 we have now two possibilities to consider:

$\boldsymbol{i_0=n+1}$**.** Then by taking $L_1=L$ and $T=\mathrm{Id}_X$ [so that T is a elementary transformation] $L=T\circ L_1$ and $\det([n+1,n+1](\mathcal{M}(L_{1;E,E})))\neq 0$ resulting in

$$\exists T\in \mathrm{Elem}(X) \text{ such that } L=T\circ L_1 \text{ and } \det([n+1,n+1](\mathcal{M}(L_1;E,E)))\neq 0 \tag{11.271}$$

$\boldsymbol{i_0\in\{1,\ldots,n\}}$**.** Take $k,l\in\{1,\ldots,n\}$ and consider $([i_0\boxplus n+1](M))_{k,l}$ then for k we have either:

$\boldsymbol{1\leqslant k<i_0}$**.** Then $([i_0\boxplus n+1](M))_{k,l} \underset{1\leqslant l<n+1\wedge[\text{definition: } 11.349]}{=} M_{k,l}$

$\boldsymbol{i_0\leqslant k\leqslant n}$**.** Then $([i_0\boxplus n+1](M))_{k,l} \underset{1\leqslant l<n+1\wedge[\text{definition: } 11.349]}{=} M_{k+1,l}$

proving that

$$\forall k,l\in\{1,\ldots,n\} \text{ we have } ([i_0\boxplus n+1](M))_{k,l}=\begin{cases} M_{k,l} \text{ if } 1\leqslant k<i_0 \\ M_{k+1,l} \text{ if } i_0\leqslant k\leqslant n \end{cases} \tag{11.272}$$

Use [definition: 11.343] to define $M'\in\mathcal{M}_{n,n}(F)$ by

$$M'=([i_0\boxplus n+1](M))_{(n\underset{\text{ni}}{\rightsquigarrow} i_0)} \text{ so that } \forall k,l\in\{1,\ldots,n\}\ M'_{k,l}=([i_0\boxplus n+1](M))_{(n\rightsquigarrow i_0)(k),l}$$

then we have for $i_0\in\{1,\ldots,n\}$ the following cases to consider:

$\boldsymbol{i_0=n}$**.** Then $\forall k,n\in\{1,\ldots,n\}$ we have

$$\begin{array}{rcl} M'_{k,l} & = & ([i_0\boxplus n+1](M))_{(n\rightsquigarrow i_0)(k),l} \\ & \underset{[\text{definition: } 11.236(1)]}{=} & ([i_0\boxplus n+1](M))_{k,l} \\ & \underset{[\text{eq: } 11.272]}{=} & \begin{cases} M_{k,l} \text{ if } 1\leqslant k<i_0 \\ M_{k+1,l} \text{ if } i_0\leqslant k\leqslant n \end{cases} \\ & \underset{i_0=n}{=} & \begin{cases} M_{k,l} \text{ if } 1\leqslant k<n \\ M_{k+1,l} \text{ if } k=i_0 \end{cases} \\ & \underset{i_0=n}{=} & \begin{cases} M_{k,l} \text{ if } k\in\{1,\ldots,n\}\setminus\{i_0\} \\ M_{k+1,l} \text{ if } k=i_0 \end{cases} \end{array}$$

$\boldsymbol{i_0\in\{1,\ldots,n-1\}}$**.** Then we have

$$\begin{array}{rcl} M'_{k,l} & = & ([i_0\boxplus n+1](M))_{(n\rightsquigarrow i_0)(k),l} \\ & \underset{i_0<n\wedge[\text{definition: } 11.236(3)]}{=} & \begin{cases} ([i_0\boxplus n+1](M))_{k,l} \text{ if } 1\leqslant k<i_0 \\ ([i_0\boxplus n+1](M))_{n,l} \text{ if } k=i_0 \\ ([i_0\boxplus n+1](M))_{k-1,l} \text{ if } i_o<k\leqslant n \\ ([i_0\boxplus n+1](M))_{k,l} \text{ if } n<k\leqslant n \end{cases} \\ & \underset{n<k\leqslant n \text{ is impossible}}{=} & \begin{cases} ([i_0\boxplus n+1](M))_{k,l} \text{ if } 1\leqslant k<i_0 \\ ([i_0\boxplus n+1](M))_{n,l} \text{ if } k=i_0 \\ ([i_0\boxplus n+1](M))_{k-1,l} \text{ if } i_o<k\leqslant n \end{cases} \\ & \underset{k<i_0\wedge l\leqslant n<n+1\wedge[\text{definition: } 11.349]}{=} & \begin{cases} M_{k,l} \text{ if } 1\leqslant k<i_0 \\ ([i_0\boxplus n+1](M))_{n,l} \text{ if } k=i_0 \\ ([i_0\boxplus n+1](M))_{k-1,l} \text{ if } i_o<k\leqslant n \end{cases} \end{array}$$

$$\underset{i_0<n\wedge l\leqslant n<n+1\wedge[\text{definition: }11.349]}{=} \begin{cases} M_{k,l} \text{ if } 1\leqslant k<i_0 \\ M_{n+1,l} \text{ if } k=i_0 \\ ([i_0\boxplus n+1](M))_{k-1,l} \text{ if } i_o<k\leqslant n \end{cases}$$

$$\underset{l\leqslant n<n+1\wedge[\text{definition: }11.349]}{=} \begin{cases} M_{k,l} \text{ if } 1\leqslant k<i_0 \\ M_{n+1,l} \text{ if } k=i_0 \\ M_{(k-1)+1,l} \text{ if } i_0<k\leqslant n \end{cases}$$

$$= \begin{cases} M_{k,l} \text{ if } k\in\{1,\ldots,n\}\setminus\{i_0\} \\ M_{k+1,l} \text{ if } k=i_0 \end{cases}$$

So in all cases we have

$$\forall k,l\in\{1,\ldots,n\} \text{ we have } M'_{k,l}=\begin{cases} M_{k,l} \text{ if } k\in\{1,\ldots,n\}\setminus\{i_0\} \\ M_{k+1,l} \text{ if } k=i_0 \end{cases} \tag{11.273}$$

Further

$$\begin{aligned} \det(M') &= \det\big(([i_0\boxplus n+1](M))_{(n\underset{\text{ni}}{\rightsquigarrow} i_0)}\big) \\ &\underset{[\text{theorem: }11.344]}{=} \operatorname{sign}\big((n\underset{\text{ni}}{\rightsquigarrow} i_0)\big)\cdot\det(([i_0\boxplus n+1](M))) \\ &\neq_{[\text{eq:}11.270]} 0 \end{aligned}$$

proving that

$$\det(M')\neq 0 \tag{11.274}$$

Define

$$L_1=\tau^{n+1}_{[n+1,i_0]}\circ L \tag{11.275}$$

then we have $\forall i\in\{1,\ldots,\}$ that

$$\begin{aligned} \sum_{k\in\{1,\ldots,n+1\}} \mathcal{M}(L_1,E,E)_{k,i}\cdot e_k &= \\ L_1(e_i) &= \\ \tau^{n+1}_{[n+1,i_0]}(L(e_i)) & \\ \tau^{n+1}_{[n+1,i_0]}\left(\sum_{k\in\{1,\ldots,n+1\}} M_{k,i}\cdot e_k\right) &= \\ \sum_{k\in\{1,\ldots,n+1\}} M_{k,i}\cdot\tau^{n+1}_{[n+1,i_0]}(e_k) &= \\ \sum_{k\in\{1,\ldots,n\}\setminus\{n+1,i_0\}} M_{k,i}\cdot\tau^{n+1}_{[n+1,i_0]}(e_k)+\sum_{k\in\{n+1\}} M_{k,i}\cdot\tau^{n+1}_{[n+1,i_0]}(e_k)+\sum_{k\in\{i_0\}} M_{k,i}\cdot\tau^{n+1}_{[n+1,i_0]}(e_k) &= \\ \sum_{k\in\{1,\ldots,n\}\setminus\{n+1,i_0\}} M_{k,i}\cdot\tau^{n+1}_{[n+1,i_0]}(e_k)+M_{n+1,i}\cdot\tau^{n+1}_{[n+1,i_0]}(e_{n+1})+M_{i_0,i}\cdot\tau^{n+1}_{[n+1,i_0]}(e_{i_0}) &= \\ \sum_{k\in\{1,\ldots,n\}\setminus\{i_0\}} M_{k,i}\cdot\tau^{n+1}_{[n+1,i_0]}(e_k)+M_{n+1,i}\cdot e_{i_0}+M_{i_0,i}\cdot e_{n+1} &= \\ \sum_{k\in\{1,\ldots,n\}\setminus\{i_0\}} M_{k,i}\cdot e_k+M_{n+1,i}\cdot e_{i_0}+M_{i_0,i}\cdot e_{n+1} & \end{aligned}$$

which as the expansion in a basis is unique proves that

$$\forall k,l\in\{1,\ldots,n+1\}\ (\mathcal{M}(L_1;E,E))_{k,l}=\begin{cases} M_{k,l} \text{ if } k\in\{1,\ldots,n\}\setminus\{i_0\} \\ M_{n+1,l} \text{ if } k=i_0 \\ M_{i_0,l} \text{ if } k=n+1 \end{cases} \tag{11.276}$$

Now $\forall k, l \in \{1, \ldots, n\}$ we have

$$
\begin{aligned}
([n+1 \boxplus n+1](\mathcal{M}(L_1; E, E)))_{k.l} &\underset{k,l \leqslant n+1}{=} \mathcal{M}(L_1; E, E)_{k,l}\\
&\underset{\text{[eq: 11.276]}}{=} \begin{cases} M_{k,l} \text{ if } k \in \{1, \ldots, n\} \setminus \{i_0\} \\ M_{n+1,l} \text{ if } k = i_0 \\ M_{i_0,l} \text{ if } k = n+1 \end{cases}\\
&\underset{k,l<n+1}{=} \begin{cases} M_{k,l} \text{ if } k \in \{1, \ldots, n\} \setminus \{i_0\} \\ M_{n+1,l} \text{ if } k = i_0 \end{cases}\\
&\underset{\text{[eq: 11.273]}}{=} M'_{k,l}
\end{aligned}
$$

so that $[n+1 \boxplus n+1](\mathcal{M}(L_1; E, E)) = M'$, hence by [eq: 11.274] that

$$\det([n+1 \boxplus n+1](\mathcal{M}(L_1; E, E))) \neq 0 \tag{11.277}$$

Next we have

$$\tau^{n+1}_{[n+1,i_0]} \circ L_1 \underset{\text{[eq: 11.275]}}{=} \tau^{n+1}_{[n+1,i_0]} \circ (\tau^{n+1}_{[n+1,i_0]} \circ L) \underset{\text{[theorem: 11.372]}}{=} \mathrm{Id}_X \circ L = L$$

So if we take $T = \tau^{n+1}_{[n+1,i_0]} \in \mathrm{Elem}(X)$ we have together with [eq: 11.277] that

$$\exists T \in \mathrm{Elem}(X) \text{ such that } L = T \circ L_1 \text{ and } \det([n+1, n+1](\mathcal{M}(L_1; E, E))) \neq 0 \tag{11.278}$$

So in all cases we have by [eqs: 11.271, 11.278] that

$$\exists T \in \mathrm{Elem}(X) \text{ such that } L = T \circ L_1 \text{ and } \det([n+1, n+1](\mathcal{M}(L_1; E, E))) \neq 0 \tag{11.279}$$

Let

$$N = \mathcal{M}(L_1; E, E) \tag{11.280}$$

Take now $Y = \mathrm{span}(\{e_1, \ldots, e_n\})$ so that $F = \{e_1, \ldots, e_n\}$ is a basis for Y. Define now $L_2 \in \mathrm{Hom}(Y, Y)$ by

$$L_2 = \mathcal{M}(F, F)^{-1}([n+1, n+1](N)) \tag{11.281}$$

so that $\det(L_2) \underset{\text{[theorem: 11.339]}}{=} \det([n+1, n+1](N)) = \det([n+1, n+1](\mathcal{M}(L_{1;E,E}))) \neq 0$ proving that $L_2 \in \mathrm{GL}(Y)$. Hence as $n \in S$ we have that

$$L_2 \text{ is composed of elementary transformations in } Y \tag{11.282}$$

Using the above and [corollary: 11.381] it follows that

$$L_3 = L_2^{[n+1]} \text{ is composed of elementary transformations in } X \tag{11.283}$$

Define now

$$L_4 = \prod_{i=1}^{n} \beta^{n+1}_{[n+1,i,N_{i,n+1}]} \tag{11.284}$$

then as L_4 and L_3 are composed of linear transformations it follows from [lemma: 11.376] that

$$L_5 = L_4 \circ L_3 \text{ is composed of linear transformations} \tag{11.285}$$

If $i \in \{1, \ldots, n\}$ then we have

$$
\begin{aligned}
L_5(e_i) &= \left(\prod_{k=1}^{n} \beta^{n+1}_{[n+1,k,N_{k,n+1}]}\right)(L_3(e_i))\\
&= \left(\prod_{k=1}^{n} \beta^{n+1}_{[n+1,k,N_{k,n+1}]}\right)(L_2^{[n+1]}(e_i))\\
&\underset{\text{[definition: 11.377]}}{=} \left(\prod_{k=1}^{n} \beta^{n+1}_{[n+1,k,N_{k,n+1}]}\right)(L_2(e_i))\\
&= \left(\prod_{k=1}^{n} \beta^{n+1}_{[n+1,k,N_{k,n+1}]}\right)\left(\sum_{j=\{1,\ldots,n\}} ([n+1, n+1]N)_{j,i} \cdot e_j\right)
\end{aligned}
$$

$$\underset{i,j\leqslant n+1\wedge[\text{definition: }11.349]}{=} \left(\prod_{k=1}^{n} \beta^{n+1}_{[n+1,k,N_{k,n+1}]}\right)\left(\sum_{j=\{1,\ldots,n\}} N_{j,i}\cdot e_j\right)$$
$$= \sum_{j=\{1,\ldots,n\}} N_{j,i}\cdot\left(\prod_{k=1}^{n} \beta^{n+1}_{[n+1,k,N_{k,n+1}]}\right)(e_j)$$
$$\underset{[\text{theorem: }11.374]}{=} \sum_{j=\{1,\ldots,n\}} N_{j,i}\cdot e_j$$

and

$$\begin{aligned} L_5(e_{n+1}) &= \left(\prod_{k=1}^{n} \beta^{n+1}_{[n+1,k,N_{k,n+1}]}\right)(L_3(e_i)) \\ &= \left(\prod_{k=1}^{n} \beta^{n+1}_{[n+1,k,N_{k,n+1}]}\right)(L_2^{[n+1]}(e_{n+1})) \\ &\underset{[\text{definition: }11.377]}{=} \left(\prod_{k=1}^{n} \beta^{n+1}_{[n+1,k,N_{k,n+1}]}\right)(e_{n+1}) \\ &\underset{[\text{theorem: }11.374]}{=} \sum_{j\in\{1,\ldots,n\}} N_{j,n+1}\cdot e_j + e_{n+1} \end{aligned}$$

proving that

$$\forall i\in\{1,\ldots,n+1\}\ L_5(e_i) = \begin{cases} \sum_{j\in\{1,\ldots,n\}} N_{j,i}\cdot e_j \text{ if } i\in\{1,\ldots,n\} \\ \sum_{j\in\{1,\ldots,n\}} N_{j,n+1}\cdot e_j + e_{n+1} \text{ if } i=n+1 \end{cases} \tag{11.286}$$

Take $\alpha\in F$ [later we will choose a value for α] then as $\sigma^{n+1}_{[n+1,\alpha]}\in \text{Elem}(X)$ and L_5 is composed of elementary transformations [see eq: 11.285] we have by [lemma: 11.376] that

$$L_6 = \sigma^{n+1}_{[n+1,\alpha]}\circ L_5 \text{ is composed of elementary transformations in } X \tag{11.287}$$

For $i\in\{1,\ldots,n\}$ we have

$$\begin{aligned} L_6(E_i) &= \sigma^{n+1}_{[n+1,\alpha]}(L_5(e_i)) \\ &\underset{[\text{eq: }11.286]}{=} \sigma^{n+1}_{[n+1,\alpha]}\left(\sum_{j\in\{1,\ldots,n\}} N_{j,i}\cdot e_j\right) \\ &= \sum_{j\in\{1,\ldots,n\}} N_{j,i}\cdot\sigma^{n+1}_{[n+1,\alpha]}(e_j) \\ &\underset{j\in\{1,\ldots,n\}\Rightarrow j\neq n+1[\text{definition: }11.371]}{=} \sum_{j\in\{1,\ldots,n\}} N_{j,i}\cdot e_j \end{aligned}$$

and

$$\begin{aligned} L_6(e_{n+1}) &= \sigma^{n+1}_{[n+1,\alpha]}(L_5(e_i)) \\ &\underset{[\text{eq: }11.286]}{=} \sigma^{n+1}_{[n+1,\alpha]}\left(\sum_{j\in\{1,\ldots,n\}} N_{j,n+1}\cdot e_j + e_{n+1}\right) \\ &= \sigma^{n+1}_{[n+1,\alpha]}\left(\sum_{j\in\{1,\ldots,n\}} N_{j,n+1}\cdot e_j\right) + \sigma^{n+1}_{[n+1,\alpha]}(e_{n+1}) \\ &\underset{[\text{definition: }11.371]}{=} \sigma^{n+1}_{[n+1,\alpha]}\left(\sum_{j\in\{1,\ldots,n\}} N_{j,n+1}\cdot e_j\right) + \alpha\cdot e_{n+1} \\ &= \sum_{j\in\{1,\ldots,n\}} N_{j,n+1}\cdot\sigma^{n+1}_{[n+1,\alpha]}(e_j) + \alpha\cdot e_{n+1} \\ &\underset{[\text{definition: }11.371]}{=} \sum_{j\in\{1,\ldots,n\}} N_{j,n+1}\cdot e_j + \alpha\cdot e_{n+1} \end{aligned}$$

proving that

$$\forall i \in \{1, \ldots, n+1\} \text{ we have } L_6(e_i) = \begin{cases} \sum_{j \in \{1, \ldots, n\}} N_{j,i} \cdot e_j \text{ if } i \in \{1, \ldots, n\} \\ \sum_{j \in \{1, \ldots, n\}} N_{j,n+1} \cdot e_j + \alpha \cdot e_{n+1} \text{ if } n+1 \end{cases} \tag{11.288}$$

Next given $B \in \mathcal{M}_{n+1,1}(F)$ [we will specify the content of B later] define

$$L_7 = \left(\prod_{i \in \{1, \ldots, n\}} \beta^{n+1}_{[i,n+1,B_i]} \right) \circ L_6$$

Which as L_6 and $\prod_{i \in \{1, \ldots, n\}} \beta^{n+1}_{[i,n+1,B_i]}$ are composed of elementary transformations proves that

$$L_7 \text{ is composed of linear transformations.} \tag{11.289}$$

Further if $i \in \{1, \ldots, n\}$ we have

$$\begin{aligned} L_7(e_i) &= \left(\prod_{i \in \{1, \ldots, n\}} \beta^{n+1}_{[i,n+1,B_i]} \right) (L_6(e_i)) \\ &\underset{[\text{eq: } 11.288]}{=} \left(\prod_{i \in \{1, \ldots, n\}} \beta^{n+1}_{[i,n+1,B_i]} \right) \left(\sum_{j \in \{1, \ldots, n\}} N_{j,i} \cdot e_j \right) \\ &= \sum_{j \in \{1, \ldots, n\}} N_{j,i} \cdot \left(\prod_{i \in \{1, \ldots, n\}} \beta^{n+1}_{[i,n+1,B_i]} \right) (e_j) \\ &\underset{[\text{theorem: } 11.375]}{=} \sum_{j \in \{1, \ldots, n\}} N_{j,i} \cdot (e_j + B_j \cdot e_{n+1}) \\ &= \sum_{j \in \{1, \ldots, n\}} N_{j,i} \cdot e_j + \sum_{j \in \{1, \ldots, n\}} (N_{j,i} \cdot B_j) \cdot e_{n+1} \\ &= \sum_{j \in \{1, \ldots, n\}} N_{j,i} \cdot e_j + \left(\sum_{j \in \{1, \ldots, n\}} N_{j,i} \cdot B_j \right) \cdot e_{n+1} \end{aligned}$$

and

$$\begin{aligned} L_7(e_{n+1}) &= \\ \left(\prod_{i \in \{1, \ldots, n\}} \beta^{n+1}_{[i,n+1,B_i]} \right) (L_6(e_{n+1})) &\underset{[\text{eq: } 11.288]}{=} \\ \left(\prod_{i \in \{1, \ldots, n\}} \beta^{n+1}_{[i,n+1,B_i]} \right) \left(\sum_{j \in \{1, \ldots, n\}} N_{j,n+1} \cdot e_j + \alpha \cdot e_{n+1} \right) &= \\ \left(\prod_{i \in \{1, \ldots, n\}} \beta^{n+1}_{[i,n+1,B_i]} \right) \left(\sum_{j \in \{1, \ldots, n\}} N_{j,n+1} \cdot e_j \right) + \left(\prod_{i \in \{1, \ldots, n\}} \beta^{n+1}_{[i,n+1,B_i]} \right) (\alpha \cdot e_{n+1}) &= \\ \sum_{j \in \{1, \ldots, n\}} N_{j,n+1} \cdot \left(\prod_{i \in \{1, \ldots, n\}} \beta^{n+1}_{[i,n+1,B_i]} \right) (e_j) + \alpha \cdot \left(\prod_{i \in \{1, \ldots, n\}} \beta^{n+1}_{[i,n+1,B_i]} \right) (e_{n+1}) &\underset{[\text{theorem: } 11.375]}{=} \\ \sum_{j \in \{1, \ldots, n\}} N_{j,n+1} \cdot (e_j + B_j \cdot e_{n+1}) + \alpha \cdot e_{n+1} &= \\ \sum_{j \in \{1, \ldots, n\}} N_{j,n+1} \cdot e_j + \left(\sum_{j \in \{1, \ldots, n\}} N_{j,n+1} \cdot B_j \right) \cdot e_{n+1} + \alpha \cdot e_{n+1} &= \\ \sum_{j \in \{1, \ldots, n\}} N_{j,n+1} \cdot e_j + \left(\alpha + \sum_{j \in \{1, \ldots, n\}} N_{j,n+1} \cdot B_j \right) \cdot e_{n+1} &= \end{aligned}$$

proving that $\forall i \in \{1, \ldots, n\}$

$$L_7(e_i) = \begin{cases} \sum_{j \in \{1, \ldots, n\}} N_{j,i} \cdot e_j + \left(\sum_{j \in \{1, \ldots, n\}} N_{j,i} \cdot B_j\right) \cdot e_{n+1} \text{ if } i \in \{1, \ldots, n\} \\ \sum_{j \in \{1, \ldots, n\}} N_{j,n+1} \cdot e_j + \left(\alpha + \sum_{j \in \{1, \ldots, n\}} N_{j,n+1} \cdot B_j\right) \cdot e_{n+1} \text{ if i=n+1} \end{cases} \tag{11.290}$$

As

$$\begin{array}{rcl} \det(([n+1 \boxplus n+1](N))^T) & \underset{[\text{theorem: } 11.338]}{=} & \det(([n+1 \boxplus n+1](N))) \\ & \underset{[\text{eq: } 11.280]}{=} & \det(([n+1 \boxplus n+1](\mathcal{M}(L_1; E, E)))) \\ & \neq_{[\text{eq:} 11.279]} & 0 \end{array}$$

$([n+1 \boxplus n+1](N))^T$ has a inverse, so we can define

$$K = (([n+1 \boxplus n+1](N))^T)^{-1}$$

Then we have $\forall i, j \in \{1, \ldots, n\}$ we have

$$\begin{array}{rcl} \delta_{i,j} & = & ((([n+1 \boxplus n+1](N))^T) \cdot K)_{i,j} \\ & = & \sum_{k \in \{1, \ldots, n\}} (([n+1 \boxplus n+1](N))^T)_{i,k} \cdot K_{k,j} \\ & = & \sum_{k \in \{1, \ldots, n\}} ([n+1 \boxplus n+1](N))_{k,i} \cdot K_{k,j} \\ & \underset{[\text{theorem: } 11.349]}{=} & \sum_{k \in \{1, \ldots, n\}} N_{k,i} \cdot K_{k,j} \end{array} \tag{11.291}$$

Now we choose the value of $B \in \mathcal{M}_{n+1,1}(F)$ to be

$$\forall i \in \{1, \ldots, n+1\}\ B_i = \begin{cases} \sum_{k \in \{1, \ldots, n\}} K_{i,k} \cdot N_{n+1,k} \text{ if } i \in \{1, \ldots, n\} \\ 1 \text{ if } i = n+1 \end{cases}$$

then we have $\forall i \in \{1, \ldots, n\}$ that

$$\begin{array}{rcl} \sum_{j \in \{1, \ldots, n\}} N_{j,i} \cdot B_j & = & \sum_{j \in \{1, \ldots, n\}} N_{j,i} \cdot \left(\sum_{k \in \{1, \ldots, n\}} K_{j,k} \cdot N_{n+1,k}\right) \\ & = & \sum_{j \in \{1, \ldots, n\}} \left(\sum_{k \in \{1, \ldots, n\}} N_{j,i} \cdot K_{j,k} \cdot N_{n+1,k}\right) \\ & \underset{[\text{theorem: } 11.45]}{=} & \sum_{k \in \{1, \ldots, n\}} \left(\sum_{j \in \{1, \ldots, n\}} N_{j,i} \cdot K_{j,k} \cdot N_{n+1,k}\right) \\ & = & \sum_{k \in \{1, \ldots, n\}} \left(\sum_{j \in \{1, \ldots, n\}} N_{j,i} \cdot K_{j,k}\right) \cdot N_{n+1,k} \\ & \underset{[\text{eq: } 11.291}{=} & \sum_{k \in \{1, \ldots, n\}} \delta_{i,k} \cdot N_{n+1,k} \\ & = & N_{n+1,i} \end{array}$$

Substituting the above result in [eq: 11.290] results in $\forall i \in \{1, \ldots, n+1\}$

$$L_7(e_i) = \begin{cases} \sum_{j \in \{1, \ldots, n\}} N_{j,i} \cdot e_j + N_{n+1,i} \cdot e_{n+1} \text{ if } i \in \{1, \ldots, n\} \\ \sum_{j \in \{1, \ldots, n\}} N_{j,n+1} \cdot e_j + \left(\alpha + \sum_{j \in \{1, \ldots, n\}} N_{j,n+1} \cdot B_j\right) \cdot e_{n+1} \text{ if i=n+1} \end{cases}$$

or

$$L_7(e_i) = \begin{cases} \sum_{j \in \{1, \ldots, n+1\}} N_{j,i} \cdot e_j \text{ if } i \in \{1, \ldots, n\} \\ \sum_{j \in \{1, \ldots, n\}} N_{j,n+1} \cdot e_j + \left(\alpha + \sum_{j \in \{1, \ldots, n\}} N_{j,n+1} \cdot B_j\right) \cdot e_{n+1} \text{ if } i = n+1 \end{cases}$$

Next choose α to be $N_{n+1,n+1} - \sum_{j\in\{1,\ldots,n\}} N_{j,n+1}\cdot B_j$ so that after substituting this in the above gives

$$\begin{aligned}
L_7(e_i) &= \begin{cases} \sum_{j\in\{1,\ldots,n+1\}} N_{j,i}\cdot e_j \text{ if } i\in\{1,\ldots,n\} \\ \sum_{j\in\{1,\ldots,n\}} N_{j,n+1}\cdot e_j + N_{n+1,n+1}\cdot e_{n+1} \text{ if } i=n+1 \end{cases} \\
&= \begin{cases} \sum_{j\in\{1,\ldots,n+1\}} N_{j,i}\cdot e_j \text{ if } i\in\{1,\ldots,n\} \\ \sum_{j\in\{1,\ldots,n+1\}} N_{j,n+1}\cdot e_j \text{ if } i=n+1 \end{cases} \\
&= \sum_{j\in\{1,\ldots,n+1\}} N_{j,i}\cdot e_j \\
&\underset{\text{[eq: 11.280]}}{=} \sum_{j\in\{1,\ldots,n+1\}} \mathcal{M}(L_1;E,E)_{j,i}\cdot e_j \\
&= L_1(e_i)
\end{aligned}$$

So that $\forall i\in\{1,\ldots,n+1\}$ we have $L_7(e_i)=L_1(e_i)$ proving by [theorem: 11.320] that $L_7=L_1$, hence L_1 is composed of elementary transformations [see eq: 11.289]. As $L \underset{\text{[eq: 11.279]}}{=} T\circ L_1$ and $T\in\mathrm{Elem}(X)$ we have by [lemma: 11.376] that L is composed of elementary transformations. Proving

$$n+1\in S$$

Mathematical induction proves then that $\mathbb{N}=S=\{n\in\mathbb{N}|\text{If } L\in\mathrm{GL}(X) \text{ where } X \text{ is a } n\text{-dimensional vector space over } F \text{ then } L \text{ is composed of elementary transformations in } X\}$ proving the theorem. □

In the proof of the above theorem it is not guaranteed that all the elementary transformations in the composition are non singular, the following theorem prove that we may assume that they are all nonsingular.

Theorem 11.383. *Let $n\in\mathbb{N}$, X a finite dimensional vector space over a field F with characteristic zero, $L\in\mathrm{GL}(X)$ a nonsingular transformation then L is composed of non singular elementary transformations. In other words there exists a $m\in\mathbb{N}$ and $\{T_i\}_{i\in\{1,\ldots,m\}}\subseteq\mathrm{Elem}(X)$ such that $\forall i\in\{1,\ldots,m\}$ $\det(T_i)\neq 0$ [or equivalent T_i is nonsingular]*

$$L=\prod_{i=1}^{m} T_i \underset{\text{def}}{=} T_1\circ\cdots\circ T_m$$

Proof. *Let $L\in\mathrm{GL}(X)$, using the previous lemma [lemma: 11.382] there exist a $\{T_i\}_{i\in\{1,\ldots,m\}}\subseteq\mathrm{Elem}(X)$ such that*

$$L=\prod_{i=1}^{m} T_i$$

Assume that $\exists k\in\{1,\ldots,m\}$ such that $\det(L_k)=0$ then we have

$$\det(L)=\det\left(\prod_{i=1}^{m} T_i\right) \underset{\text{[theorem: 11.300]}}{=} \prod_{i\in\{1,\ldots,m\}} \det(L_i) = \left(\prod_{i\in\{1,\ldots,m\}\setminus\{k\}} \det(L_i)\right)\cdot\det(L_k)=0$$

contradicting $L\in\mathrm{GL}(X)\Rightarrow\det(L)\neq 0$. Hence $\forall i\in\{1,\ldots,m\}$ we have $\det(L_i)\neq 0$. □

Chapter 12
Internal Direct Sum

We will now extend the sum of two subsets of a vector space [see definition: 11.194] to a sum of a finite family of subsets of a vector space.

Definition 12.1. *Let V be a vector space over a field F, I a finite set and $\{V_i\}_{i\in I}\subseteq\mathcal{P}(V)$ [a finite family of subsets of V] then the sum of $\{V_i\}_{i\in I}\subseteq\mathcal{P}(V)$ noted as $\sum_{i\in I}V_i$ is defined as*

$$\sum_{i\in I}V_i=\left\{v\in V\,|\,\exists\{v_i\}_{i\in I}\subseteq V \text{ such that } \forall i\in I\ v_i\in V_i \text{ and } v=\sum_{i\in I}w_i\right\}\subseteq V$$

Example 12.2. Let V be a vector space over a field F, $\{V_i\}_{i\in\varnothing}\subseteq\mathcal{P}(V)$ a empty family of sub spaces [see example: 2.103]

Proof. Use [example: 2.103] to get a $\{v_i\}_{i\in\varnothing}\subseteq V$ then we have $0\underset{[\text{definition: } 11.32]}{=}\sum_{i\in\varnothing}v_i$ and vacuously $\forall i\in\varnothing\ v_i\in V_i$, proving that $0\in\sum_{i\in\varnothing}V_i$. Further if $v\in\sum_{i\in\varnothing}V_i$ then there exist a $\{v_i\}_{i\in\varnothing}\subseteq V$ such that $v=\sum_{i\in\varnothing}v_i\underset{[\text{definition: } 11.32]}{=}0$ proving that $\sum_{i\in\varnothing}V_i\subseteq\{0\}$, Hence we have

$$\sum_{i\in\varnothing}V_i=\{0\}$$

□

Some of the properties of finite sums of elements of the vector space transfer to the finite sum of a family of non empty subsets.

Theorem 12.3. *Let $\langle V,+,\cdot\rangle$ be a vector space over a field F then we have for a finite set I and a finite family $\{V_i\}_{i\in I}\subseteq\mathcal{P}(V)$ of subspaces of V that*

1. *If $I=\{k\}$ then we have $\sum_{i\in\{k\}}V_i=V_k$*
2. *If $\sigma\colon J\rightarrow I$ is a bijection then we have $\sum_{i\in J}V_{\sigma(i)}=\sum_{i\in I}V_i$*
3. *If $k\in I$ then $\sum_{i\in I}V_i=\left(\sum_{i\in I\setminus\{k\}}V_i\right)+V_k$*
4. *If $I=I_1\bigcup I_2$ with $I_1\bigcap I_2=\varnothing$ then $\sum_{i\in I}V_i=\sum_{i\in\{1,2\}}\left(\sum_{j\in I_i}V_j\right)$*
5. *If $I=\bigcup_{i\in\{1,\ldots,m\}}I_i$ with $\forall i,j\in I$ if $i\neq j$ then $I_i\bigcap I_j=\varnothing$ hen*

$$\sum_{i\in I}V_i=\sum_{i\in\{1,\ldots,m\}}\left(\sum_{j\in I_i}V_j\right)$$

Proof.

1. Let $v\in\sum_{i\in\{k\}}V_i$ then $\exists\{v_i\}_{i\in\{k\}}$ such that $\forall i\in\{k\}$ we have $v_i\in V_i$ and

$$v=\sum_{i\in\{k\}}v_i\underset{[\text{theorem: } 11.34]}{=}v_k\in V_k$$

proving that

$$\sum_{i\in\{k\}}V_i\subseteq V_k \tag{12.1}$$

Let $v\in V_k$ and define $\{v_i\}_{i\in\{k\}}\subseteq V$ by $v_k=v$ then $\forall i\in\{k\}$ we have $v_i=v\in V_k$ and

$$v=v_k\underset{[\text{theorem: } 11.34]}{=}\sum_{i\in\{k\}}v_i$$

proving that $v \in \sum_{i \in \{k\}} v_i$. Hence $V_k \subseteq \sum_{i \in \{k\}} V_i$ which combined with [eq: 12.1]

$$\sum_{i \in \{k\}} V_i = V_k$$

2. Let $v \in \sum_{i \in I} V_i$ then there exist a $\{v_i\}_{i \in I} \subseteq V$ with $\forall i \in I$ $v_i \in V_i$ and $v = \sum_{i \in I} v_i$. Define now

$$\{w_i\}_{i \in J} \subseteq V \text{ by } w_i = v_{\sigma(i)}$$

then $\forall i \in J$ we have $w_i = v_{\sigma(i)} \in V_{\sigma(i)}$ and

$$\sum_{i \in J} w_i = \sum_{i \in J} v_{\sigma(i)} \underset{[\text{theorem: } 11.36]}{=} \sum_{i \in I} v_i = v$$

hence $v \in \sum_{i \in J} V_{\beta(i)}$ proving that

$$\sum_{i \in I} V_i \subseteq \sum_{i \in J} V_{\beta(i)} \tag{12.2}$$

If $v \in \sum_{i \in J} V_{\beta(i)}$ then there exist a $\{w_i\}_{i \in J} \subseteq V$ such that $\forall i \in J$ we have $w_i \in V_{\beta(i)}$ and $v = \sum_{i \in J} w_i$. Define now

$$\{v_i\}_{i \in I} \subseteq V \text{ by } v_i = w_{\sigma^{-1}(i)}$$

then $\forall i \in I$ we have $v_i = v_{\sigma(\sigma^{-1}(i))} \in V_{\sigma(\sigma^{-1}(i))} = V_i$ and

$$\sum_{i \in I} v_i = \sum_{i \in I} w_{\sigma^{-1}(i)} \underset{[\text{theorem: } 11.36]}{=} \sum_{i \in J} w_i = v$$

proving that $v \in \sum_{i \in I} V_i$. Hence $\sum_{i \in J} V_{\beta(i)} \subseteq \sum_{i \in J} V_{\beta(i)}$ which combined with [eq: 12.2]

$$\sum_{i \in I} V_i = \sum_{i \in J} V_{\beta(i)}$$

3. If $v \in \sum_{i \in I} V_i$ then $\exists \{v_i\}_{i \in I}$ such that $\forall i \in I$ $v_i \in V_i$ and $v = \sum_{i \in I} v_i$. Let $w = \sum_{i \in I \setminus \{k\}} v_i$ then as trivially $\forall i \in I \setminus \{k\}$ $v_i \in V_i$ we have that

$$w \in \sum_{i \in I \setminus \{k\}} V_i$$

and

$$v = \sum_{i \in I} v_i \underset{[\text{theorem: } 11.43]}{=} \sum_{i \in I \setminus \{k\}} v_i + \sum_{i \in \{k\}} v_i \underset{[\text{theorem: } 11.34]}{=} \sum_{i \in I \setminus \{k\}} v_i + v_k = w + v_k$$

so, as $v_k \in V_k$, it follows that $v \in \left(\sum_{i \bowtie I \setminus \{k\}} V_i\right) + V_k$. Hence

$$\sum_{i \in I} V_i \subseteq \left(\sum_{i \bowtie I \setminus \{k\}} V_i\right) + V_k \tag{12.3}$$

On the other hand if $v \in \left(\sum_{i \in I \setminus \{k\}} V_i\right) + V_k$ then there exists a $u \in \sum_{i \bowtie I \setminus \{k\}} V_i$ and $w \in V_k$ such that $v = u + v$. As $u \in \sum_{i \bowtie I \setminus \{k\}} V$ there exist a $\{v_i\}_{i \in I \setminus \{k\}} \subseteq V$ such that $\forall i \in I \setminus \{k\}$ $v_i \in V_i$ and $u = \sum_{i \in I \setminus \{k\}} v_i$. Define now

$$\{w_i\}_{i \in I} \subseteq V \text{ by } w_i = \begin{cases} w \text{ if } i = k \\ v_i \text{ if } i \in I \setminus \{k\} \end{cases}$$

then we have

$$\sum_{i \in I} w_i \underset{[\text{theorem: } 11.43]}{=} \sum_{i \in I \setminus \{k\}} w_i + \sum_{i \in k} w_i \underset{[\text{theorem: } 11.34]}{=} \sum_{i \in I \setminus k} w_i + w_k = \sum_{i \in I \setminus k} v_i + w = u + w$$

which, as $\forall i \in I$ $w_i = \left\{ \begin{smallmatrix} w \in V_k \text{ if } i = k \\ v_i \in V_i \text{ if } i \in I \setminus \{k\} \end{smallmatrix} \right. \in V_i$, results in $v \in \sum_{i \in I} V_i$ proving that

$$\left(\sum_{i \in I \setminus \{k\}} V_i\right) + V_k \subseteq \sum_{i \in I} V_i$$

which combined with [eq: 12.3] proves that

$$\sum_{i \in I} V_i = \left(\sum_{i \bowtie I \setminus \{k\}} V_i\right) + V_k$$

4. If $v\in\sum_{i\in I}V_i$ then there exist a $\{v_i\}_{i\in I}\subseteq V$ such that $\forall i\in I$ we have $v_i\in V_i$ and $v=\sum_{i\in I}v_i$. Define
$$\{w_i\}_{i\in\{1.2\}}\subseteq V \text{ by } w_i=\sum_{j\in I_i}v_j$$
If $i\in\{1,2\}$ then $\forall j\in I_i\subseteq I$ $v_j\in V_j$ so we have
$$\forall i\in\{1,2\}\text{ we have that } w_i\in\sum_{j\in I_i}V_j$$
further we have
$$\begin{aligned}\sum_{i\in\{1,,2\}}w_i &= \sum_{i\in\{1,2\}}\left(\sum_{j\in I_i}v_j\right)\\ &\underset{[\text{theorem: } 11.44]}{=} \sum_{i\in I_1\bigcup I_2}v_i\\ &= v\end{aligned}$$
proving that $v\in\sum_{i\in\{1,2\}}(\sum_{j\in I_i}v_j)$. Hence
$$\sum_{i\in I}V_i\subseteq\sum_{i\in\{1,2\}}\left(\sum_{j\in I_i}v_j\right)\tag{12.4}$$
If $v\in\sum_{i\in\{1,2\}}(\sum_{j\in I_i}V_j)$ then there exists a $\{v_i\}_{i\in\{1,.2\}}\subseteq V$ with $\forall i\in\{1,2\}$ $v_i\in\sum_{j\in I_i}V_j$ such that $v=\sum_{i\in\{1,2\}}v_i$. Let $i\in\{1.2\}$ then there exist a $\{v_{i,j}\}_{j\in I_i}\subseteq V$ with $\forall j\in I_j$ $v_{i,j}\in V_j$ and $v_i=\sum_{j\in I_i}v_{i,j}$. We can then, as $I_1\bigcap I_2\neq\varnothing$, define
$$\{w_i\}_{i\in I}\subseteq V \text{ by } w_i=\begin{cases}v_{1,i} \text{ if } i\in I_1\\ v_{2,i} \text{ if } i\in I_2\end{cases}$$
then we have $\forall i\in I=I_1\bigcup I_2$ that
$$w_i=\begin{cases}v_{1,i}\in V_i \text{ if } i\in I_1\\ v_{2,i}\in V_2 \text{ if } i\in I_2\end{cases}\in V_i$$
and
$$\begin{aligned}\sum_{i\in I}w_i &\underset{[\text{theorem: } 11.44]}{=} \sum_{i\in I_1}w_i+\sum_{i\in I_2}w_i\\ &= \sum_{i\in I_1}v_{1,i}+\sum_{i\in I_2}v_{2,i}\\ &= v_1+v_2\\ &= \sum_{i\in\{1,2\}}v_i\\ &= v\end{aligned}$$
proving that $v\in\sum_{i\in I}V_i$. So $\sum_{i\in\{1,2\}}(\sum_{j\in I_i}v_j)\subseteq\sum_{i\in I}V_i$ which combined with [eq: 12.4] gives
$$\sum_{i\in I}V_i=\sum_{i\in\{1,2\}}\left(\sum_{j\in I_i}v_j\right)$$

5. This is proved by induction, so let
$$S=\left\{n\in\mathbb{N}|\text{If } I=\bigcup_{i\in\{1,\ldots,m\}}I_i \text{ with } \forall i,j\in I \text{ if } i\neq j \text{ then } I_i\bigcap I_j=\varnothing, \text{ then } \sum_{i\in I}V_i=\sum_{i\in\{1,\ldots,m\}}\left(\sum_{j\in I_i}V_i\right)\right\}$$
then we have:

$1\in S$. If $I=\bigcup_{i\in\{1\}}I_i$ [so we have automatically $\forall i,j\in I$ if $i\neq j$ then $I_i\bigcap I_j=\varnothing$] then
$$\sum_{i\in\{1\}}\left(\sum_{j\in I_i}V_j\right)\underset{(1)}{=}\sum_{j\in I_1}V_i\underset{I=I_1}{=}\sum_{j\in I}V_i$$
proving that $1\in S$.

$n\in S\Rightarrow n+1\in S$. If $I=\bigcup_{i\in\{1,\ldots,m+1\}}I_i$ with $\forall i,j\in I$ if $i\neq j$ then $I_i\bigcap I_j=\varnothing$. Let $J_1=\bigcup_{i\in\{1,\ldots,m\}}I_i$ and $J_2=I_{m+1}$ then
$$J_1\bigcap J_2=\left(\bigcup_{i\in\{1,\ldots,m\}}I_i\right)\bigcap J_2=\bigcup_{i\in\{1,\ldots,m\}}(I_i\bigcap I_{m+1})=\varnothing$$

and $I = J_1 \bigcup I_2$. Hence we have

$$\begin{aligned}
\sum_{i\in I} V_i &\underset{(4)}{=} \sum_{i\in\{1,2\}}\left(\sum_{j\in J_i} V_j\right)\\
&\underset{(3)}{=} \left(\sum_{i\in\{1\}}\left(\sum_{j\in J_i} V_j\right)\right) + \sum_{j\in J_2} V_j\\
&\underset{(1)}{=} \sum_{j\in J_1} V_j + \sum_{j\in J_2} V_j\\
&= \sum_{j\in J_1} V_j + \sum_{j\in I_{m+1}} V_j\\
&= \left(\sum_{j\in\bigcup_{i\in\{1,\ldots,m\}} I_i} V_j\right) + \sum_{j\in I_{m+1}} V_j\\
&\underset{n\in S}{=} \sum_{i\in\{1,\ldots,m\}}\left(\sum_{j\in I_i} V_j\right) + \sum_{j\in I_{m+1}} V_j\\
&\underset{(3)}{=} \sum_{i\in\{1,\ldots,m+1\}}\left(\sum_{i\in I_i} V_j\right)
\end{aligned}$$

proving that $m+1\in S$. □

We consider now the case where we have a family of sub-spaces of a vector space.

Theorem 12.4. *If V is a vector space over a field F and $\{V_i\}_{i\in I}\subseteq\mathcal{P}(V)$ such that $\forall i\in I$ V_i is a sub space of V*

1. *$\sum_{i\in I} V_i$ is a sub space of V*
2. *If $J\subseteq I$ then $\sum_{i\in J} V_i \subseteq \sum_{i\in I} V_i$*
3. *If $k\in I$ then $V_k\subseteq\sum_{i\in I} V_i$*

then

Proof.

1. As $\forall i\in I$ V_i is a sub space of V we have that $0\in V_i$, so as $\sum_{i\in I} 0 = 0$ it follows that $0\in\sum_{i\in I} V_i$ proving that
$$\varnothing\neq\sum_{i\in I} V_i$$
Let $\alpha\in F$ and $x,y\in\sum_{i\in I} V_i$ then there exists a $\{x_i\}_{i\in I}\subseteq V$ and $\{y_i\}_{i\in I}\subseteq V$ such that $\forall i\in I$ $x_i,y_i\in V_i$ and $x=\sum_{i\in I} x_i$, $y=\sum_{i\in I} y_i$. As $\forall i\in I$ V_i is a subspace we have that $\alpha\cdot x_i + y_i\in V_i$ and
$$\alpha\cdot x + y = \alpha\cdot\sum_{i\in I} x_i + \sum_{i\in I} y_i \underset{[\text{theorems: } 11.75,\ 11.17]}{=} \sum_{i\in I}(\alpha\cdot x_i + y_i)$$
proving that $\alpha\cdot x + y\in\sum_{i\in I} V_i$.
2. If $v\in\sum_{i\in J} V_i$ then there exist a $\{v_i\}_{i\in J}\subseteq V$ such that $\forall i\in J$ $v_i\in V_i$ and $v=\sum_{i\in J} v_i$. Define then
$$\{w_i\}_{i\in I}\subseteq V \text{ by } w_i = \begin{cases} 0 \text{ if } i\in I\setminus J\\ v_i \text{ if } i\in J\end{cases}$$
then $\forall i\in I$ we have $w_i\in V_i$ [as $0\in V_i$] and
$$\begin{aligned}
\sum_{i\in I} w_i &\underset{[\text{theorem: } 11.43]}{=} \sum_{i\in I\setminus J} w_i + \sum_{i\in J} w_i\\
&= \sum_{i\in I\setminus J} 0 + \sum_{i\in J} v_i\\
&= v
\end{aligned}$$
Hence $v\in\sum_{i\in I} V_i$ proving
$$\sum_{i\in J} V_i \subseteq \sum_{i\in I} V_i$$

3. Let $k\in I$ and $v\in V_k$, define

$$\{v_i\}_{i\in I} \text{ by } v_i=\begin{cases} v \text{ if } i=k\\ 0 \text{ if } i\in I\setminus\{k\}\end{cases}$$

then, as $V_i's$ are sub-spaces of V hence containing 0 we have

$$\forall i\in I\setminus\{k\} \text{ that } v_i=\begin{cases} v\in V_k \text{ if } i=k\\ 0\in V_i \text{ if } i\in I\setminus\{k\}\end{cases}\in V_i$$

Further

$$\sum_{i\in I} v_i \underset{\text{[theorem: 11.43]}}{=} \sum_{i\in I\setminus\{k\}} v_i+\sum_{i\in\{k\}} v_i \underset{\text{[theorem:11.34]}}{=} \sum_{i\in I\setminus\{k\}} v_i+v_k=\sum_{i\in I\setminus\{k\}} 0+v=v$$

proving that $v\in\sum_{i\in I} v_i$. Hence

$$V_k\subseteq\sum_{i\in I} V_i$$

□

If $v\in\sum_{i\in I} V_i$ then we can write v as $\sum_{i\in I} v_i$ where $v_i\in V_i$, however this expansion is not unique, which is not very useful (compare this with the uniqueness of the expansion of a vector in a basis). To solve this we extend the concept of a direct sum of two subspaces [see definition: 11.195] to a direct sum of a finite family.

Definition 12.5. (Internal Direct Sum) *If V is a vector space over a field then*

$$V=\bigoplus_{i\in I} V_i$$

iff

1. *I is a finite set*
2. *$\{V_i\}_{i\in I}\subseteq\mathcal{P}(V)$ satisfies $\forall i\in I$ V_i is a sub space of V*
3. *If $v\in V$ then there exist a **unique** family $\{v_i\}_{i\in I}\subseteq V$ such that $\forall i\in I$ $v_i\in V_i$ and $v=\sum_{i\in I} v_i$.*

Example 12.6. Let V be a vector space over a field F, $\{V_i\}_{i\in\varnothing}\subseteq\mathcal{P}(V)$ a empty family of sub-spaces of V [see example: 2.103] then

$$\bigoplus_{i\in\varnothing} V_i=\{0\}$$

Proof. If $v\in\oplus_{i\in\varnothing}V_i$ then there exist a $\{v_i\}_{i\in\varnothing}\subseteq V$ such that $v=\sum_{i\in\varnothing} v_i \underset{\text{[definition: 11.32]}}{=} 0$ proving that

$$\bigoplus_{i\in\varnothing} V_i\subseteq\{0\},$$

Use [example: 2.103] to get a $\{v_i\}_{i\in\varnothing}\subseteq V$ then we have $0 \underset{\text{[definition: 11.32]}}{=} \sum_{i\in\varnothing} v_i$ and vacuously $\forall i\in\varnothing$ $v_i\in V_i$. If $\{w_i\}_{i\in\varnothing}\subseteq V$ is such that $0=\sum_{i\in\varnothing} w_i$ then we have vacuously $v_i=w_i$ hence $\{v_i\}_{i\in\varnothing}=\{w_i\}_{i\in\varnothing}$ proving uniqueness. Hence $0\in\bigoplus_{i\in\varnothing} V_i$ so that

$$\bigoplus_{i\in\varnothing} V_i=\{0\}$$

□

Example 12.7. If V is a vector space then and W a sub-space then $W=\bigoplus_{i\in\{k\}} V_i$ where $V_k=W$.

Proof. $V_1=W$ is a subspace of V and if $v\in V$ we have for $\{v_i\}_{i\in\{1\}}\subseteq V$ that $v=\sum_{i\in\{1\}} v_i$ and if $\{w_i\}_{i\in\{1\}}$ such that $v=\sum_{i\in\{1\}} w_i$ then $w_i=v=v_i$ $\forall i\in\{1\}$. □

We can rephrase the above definition in another way as follows

Theorem 12.8. *If V is a vector space over a field, I a finite set, $\{V_i\}_{i\in I}\subseteq\mathcal{P}(V)$ a finite family of sub-spaces of V then*

$$V=\bigoplus_{i\in I} V_i$$

$$\Updownarrow$$

$$V=\sum_{i\in I} V_i \textit{ and } \forall k\in I \textit{ we have } V_k\bigcap\left(\sum_{i\in I\setminus\{k\}} V_i\right)=\{0\}$$

Proof.

⇒. Let $v\in V$ then, as $V=\bigoplus_{i\in I} V_i$, there exist a **unique** $\{v_i\}_{i\in I}\subseteq V$ such that $\forall i\in I$ we have $v_i\in V_i$ and $v=\sum_{i\in I} v_i$, proving $V\subseteq \sum_{i\in I} V_i$. As by [definition: 12.1] $\sum_{i\in I} V_i\subseteq V$ we have

$$V=\sum_{i\in I} V_i$$

Let $k\in I$. If $v\in V_k\bigcap\left(\sum_{i\in I\setminus\{k\}} V_i\right)$ then $v\in V_k$ and $v\in\sum_{i\in I\setminus\{k\}} V_i$. So there exists a $\{v_i\}_{i\in I\setminus\{k\}}\subseteq V$ such that $\forall i\in I\setminus\{k\}$ $v_i\in V_i$ and $v=\sum_{i\in I\setminus\{k\}} V_i$. Define

$$\{w_i\}_{i\in I}\subseteq V \text{ by } w_i=\begin{cases} -v \text{ if } i=k \\ v_i \text{ if } i\in I\setminus\{k\}\end{cases}\in V_i$$

then

$$\forall i\in I \text{ we have } w_i\in\begin{cases} -v\in V_k \text{ if } i=k \\ v_i\in V_i \text{ if } i\in I\setminus\{k\}\end{cases}\in V_i$$

and

$$\begin{aligned}\sum_{i\in I} w_i &\underset{\text{[theorem: 11.43]}}{=} \sum_{i\in I\setminus\{k\}} w_i+\sum_{i\in\{k\}} w_i \\ &\underset{\text{[theorem: 11.34]}}{=} \sum_{i\in I\setminus\{k\}} w_i+w_k \\ &= \sum_{i\in I\setminus\{k\}} v_i+(-v) \\ &= v+(-v) \\ &= 0\end{aligned}$$

Further for $\{u_i\}_{i\in I}\subseteq V$ defined by $u_i=0$ we have $\forall i\in I$ $u_i=0\in V_i$ and $0\underset{\text{[theorem: 11.37]}}{=}\sum_{i\in I} u_i$, so, as $0\in V$ and $V=\bigoplus_{i\in I} V_i$, we have by **uniqueness** that $\forall i\in I$ $u_i=w_i$, hence $-v=w_k=u_k=0$ proving that $v=0$. As $v\in V_k\bigcap\left(\sum_{i\in I\setminus\{k\}} V_i\right)$ was chosen arbitrary we conclude that

$$V_k\bigcap\left(\sum_{i\in I\setminus\{k\}} V_i\right)\subseteq\{0\}$$

By [theorem: 12.4] $\sum_{i\in I\setminus\{k\}} V_i$ is a sub-space of V and by the hypothesis V_k is a subspace of V so we have $0\in V_k\bigcap\left(\sum_{i\in I\setminus\{\}} V_i\right)$ and it follows that

$$V_k\bigcap\left(\sum_{i\in I\setminus\{k\}} V_i\right)=\{0\}$$

⇐. Let $v\in V$ then as $V=\sum_{i\in I} V_i$ there exists a $\{v_i\}_{i\in I}\subseteq V$ with $\forall i\in I$ $v_i\in V_i$ and $v=\sum_{i\in I} v_i$. Let $\{w_i\}_{i\in I}\subseteq V$ such that $\forall i\in I$ $w_i\in V_i$ and $v=\sum_{i\in I} v_i$. Assume that $\{v_i\}_{i\in I}\neq\{w_i\}_{i\in I}$ then there exist a $k\in I$ such that $v_k\neq w_k$ and we have

$$\begin{aligned}0 &= v+(-v) \\ &= \sum_{i\in I} v_i+\left(-\sum_{i\in I} w_i\right) \\ &\underset{\text{[theorems: 11.38, 11.40]}}{=} \sum_{i\in I}(v_i-w_i) \\ &= \sum_{i\in I\setminus\{k\}}(v_i-w_i)+\sum_{i\in\{k\}}(v_i-w_i) \\ &= \sum_{i\in I\setminus\{k\}}(v_i-w_i)+(v_k-w_k)\end{aligned}$$

so that

$$w_k-v_k=\sum_{i\in I\setminus\{k\}}(v_i-w_i)$$

Now $w_k - v_k \in V_k$ and $\forall i \in I \setminus \{k\}$ $v_i - w_i \in V_i$ so that $w_k - v_k \in V_k \bigcap \sum_{i \in I \setminus \{k\}} V_i$, hence by the hypothesis we have that $w_k - v_k = 0$ contradicting $v_k \neq w_k$. So we must have that $\{v_i\}_{i \in I} = \{w_i\}_{i \in I}$. Proving that

$$V = \bigoplus_{i \in I} V_i$$

□

Theorem 12.9. *Let V be a vector space over a field F, I_1, I_2 finite sets such that $I_1 \bigcap I_2 = \varnothing$, U, W sub-spaces of V, $\{V_i\}_{i \in I_1 \bigcup I_2} \subseteq \mathcal{P}(V)$ a finite family of sub-spaces of V such that*

$$U = \bigoplus_{i \in I_1} V_i,\ W = \bigoplus_{i \in I_2} V_i \text{ and } V = U \oplus W$$

then

$$V = \bigoplus_{i \in I_1 \bigcup I_2} V_i$$

Proof. As $U = \bigoplus_{i \in I_1} V_i$, $W = \bigoplus_{i \in I_2} V_i$ we have by definition that

$$\forall i \in I_1\ V_i \text{ is a sub-space of } U, V \text{ and } \forall i \in I_2\ V_i \text{ is a sub-space of } W, W \tag{12.5}$$

Let $v \in V$ then there exists **unique** $u \in U$ and a $w \in W$ such that

$$v = u + w \tag{12.6}$$

As $U = \bigoplus_{i \in I_1} U_i$ and $W = \bigoplus_{i \in I_2} W_i$ there exists **unique** $\{u_i\}_{i \in I_1} \subseteq U \subseteq V$ and $\{w_i\}_{i \in I_2} \subseteq W \subseteq V$ such that

$$\forall i \in I_1\, u_i \in V_i \wedge \forall i \in I_2\ w_i \in V_i \wedge u = \sum_{i \in I_1} u_i \wedge w = \sum_{i \in I_2} w_i \tag{12.7}$$

As $I_1 \bigcap I_2 = \varnothing$ we can define

$$\{v_i\}_{i \in I_1 \bigcup I_2} \subseteq V \text{ by } v_i = \begin{cases} u_i \text{ if } i \in I_1 \\ w_i \text{ if } i \in I_2 \end{cases} \in V_i$$

then

$$\begin{aligned} \sum_{i \in I_1 \bigcup I_2} v_i &\underset{[\text{theorem: } 11.43]}{=} \sum_{i \in I_1} v_i + \sum_{i \in I_2} v_i \\ &= \sum_{i \in I_1} u_i + \sum_{i \in I_2} w_i \\ &= u + w \\ &\underset{[\text{eq: } 12.6]}{=} v \end{aligned}$$

proving that

$$v = \sum_{i \in I_1 \bigcup I_2} v_i \text{ and } \forall i \in I_1 \bigcup I_2\ v_i = \begin{cases} u_i \in V_i \text{ if } i \in I_1 \\ w_i \in V_i \text{ if } i \in I_2 \end{cases} \in V_i$$

Let $\{z_i\}_{i \in I_1 \bigcup I_2} \subseteq V$ such that $\forall i \in I_1 \bigcup I_2$ $z_i \in V_i$ and $v = \sum_{i \in I_1 \bigcup I_2} z_i$. Define then u', w' by

$$u' = \sum_{i \in I_1} z_i \text{ and } w' = \sum_{i \in I_2} z_i \tag{12.8}$$

then we have that

$$u' + w' = \sum_{i \in I_1} z_i + \sum_{i \in I_2} z_i \underset{[\text{theorem: } 11.43]}{=} \sum_{i \in I_1 \bigcup I_2} z_i = v$$

As $\forall i \in I_1$ V_i is a sub-space of U, $\forall i \in I_2$ is a sub-space of W [see: 12.5] we have by [theorem: 12.4] that $u' \in U$ and $w' \in W$ so

$$u' + w' = v \underset{[\text{eq: } 12.6]}{=} u + w \wedge u, u' \in U \wedge w, w' \in W$$

Hence as $V = U \oplus W$ we have by the uniqueness of the decomposition that

$$u = u' \wedge w = w' \tag{12.9}$$

Combining [eqs: 12.7, 12.8, 12.9] we have

$$u = \sum_{i \in I_1} u_i = u' = \sum_{i \in I_1} z_i \wedge w = \sum_{i \in I_2} w_i = w' = \sum_{i \in I_2} z_i$$

Now as $U = \bigoplus_{i \in I_1} V_i$, $W = \bigoplus_{i \in I_2} V_i$ we have by the above and uniqueness that

$$\forall i \in I_1\ z_i = u_i \text{ and } \forall i \in I_2\ z_i = w_i$$

hence we have $\forall i \in I_1 \bigcup I_2$ that

$$v_i = \begin{cases} u_i \text{ if } i \in I_1 \\ w_i \text{ if } i \in I_2 \end{cases} = \begin{cases} z_i \text{ if } i \in I_1 \\ z_i \text{ if } i \in I_2 \end{cases} = z_i$$

proving that the decomposition of $v \in V$ is unique. Hence

$$V = \bigoplus_{i \in I_1 \bigcup I_2} V_i$$

□

Corollary 12.10. *Let V be a vector space over a field F, I a finite set, $k \notin I$, U a sub-space of V $\{V_i\}_{i \in I \bigcup \{k\}} \subseteq \mathcal{P}(V)$ a finite family of sub-spaces of V such that*

$$\forall i \in I\ V_i \textit{ is a sub-space of } U$$

and

$$U = \bigoplus_{i \in I} V_i \textit{ and } V = U \bigoplus V_k$$

then

$$V = \bigoplus_{i \in I \bigcup \{k\}} V_i$$

Proof. Take $I_1 = I$ and $I_2 = \{k\}$ then we have as $k \notin I$ that $I_1 \bigcap I_2 = \varnothing$. Further $\forall i \in I_1\ V_i$ is a sub-space of U and $\forall i \in I_2\ V_i = V_k$ is a sub-space of V_k. Further as $V_k \underset{\text{[example: 12.7]}}{=} \bigoplus_{i \in \{k\}} W_i$ and by the hypothesis $U = \bigoplus_{i \in I} V_i$ and $V = U \bigoplus V_k$ it follows from [theorem: 12.9] that

$$V = \bigoplus_{i \in I \bigcup \{k\}} V_i$$

□

Theorem 12.11. *Let V be a vector space over a field F, I a finite set and $\{V_i\}_{i \in I} \subseteq \mathcal{P}(V)$ a family of sub-spaces of V such that*

$$V = \bigoplus_{i \in I} V_i$$

then we have:

1. *If $J \subseteq I$ we have that*

$$\sum_{i \in J} V_i = \bigoplus_{i \in J} V_i$$

2. *If $I = I_1 \bigcup I_2$ with $I_1 \bigcap I_2 = \varnothing$ then*

$$\sum_{i \in I_1} V_i = \bigoplus_{i \in I_1} V_i \textit{ and } \sum_{i \in I_2} V_i = \bigoplus_{i \in I_2} V_i$$

   and

$$V = \left(\sum_{i \in I_1} V_i\right) \oplus \left(\sum_{i \in I_2} V_i\right)$$

3. *If $k \in I$ then*

$$\sum_{i \in I \setminus \{k\}} V_i = \bigoplus_{i \in I \setminus \{k\}} V_i$$

   and

$$V = \left(\sum_{i \in I \setminus \{k\}} V_i\right) \oplus V_k$$

***Note* 12.12.** *By [theorem: 12.4] we have that $\sum_{i \in J} V_i$, $\sum_{i \in I_1} V_i$ and $\sum_{i \in I_2} V_i$ are sub-spaces of V as is required for (1) and (2).*

Proof.

1. Let $j \in \sum_{i \in J} V_i$ then there exists a $\{v_i\}_{i \in J} \subseteq V$ such that $\forall i \in J\ v_i \in V_i$ and $v = \sum_{i \in J} v$. Assume that there exists also a $\{w_i\}_{i \in J} \subseteq V$ such that $\forall i \in J\ w_i \in V_i$ and $v = \sum_{i \in J} w_i$. Define then

$$\{v_i'\}_{i \in I} \subseteq V \text{ by } v_i' = \begin{cases} 0 \text{ if } i \in I \setminus J \\ v_i \text{ if } i \in J \end{cases} \text{ and } \{w_i'\}_{i \in I} \subseteq V \text{ by } w_i' = \begin{cases} 0 \text{ if } i \in I \setminus J \\ w_i \text{ if } i \in J \end{cases}$$

then we have $\forall i\in I$

$$v'_i=\left\{\begin{array}{l}0\in V_i \text{ if } i\in I\setminus J\\ u_i\in V_i \text{ if } i\in J\end{array}\right.\in V_i \text{ and } w'_i=\left\{\begin{array}{l}0\in V_i \text{ if } i\in I\setminus J\\ w_i\in V_i \text{ if } i\in J\end{array}\right.\in V_i \tag{12.10}$$

Further

$$\begin{aligned}\sum_{i\in I} v'_i &\underset{[\text{theorem: } 11.43]}{=} \sum_{i\in I\setminus J} v'_i+\sum_{i\in J} v'_i\\ &= \sum_{i\in I\setminus J} 0+\sum_{i\in J} v_i\\ &\underset{[\text{theorem: } 11.37]}{=} 0+v\\ &= v \qquad (12.11)\\ \sum_{i\in I} w'_i &\underset{[\text{theorem: } 11.43]}{=} \sum_{i\in I\setminus J} w'_i+\sum_{i\in J} w'_i\\ &= \sum_{i\in I\setminus J} 0+\sum_{i\in J} w_i\\ &\underset{[\text{theorem: } 11.37]}{=} 0+v\\ &= v \qquad (12.12)\end{aligned}$$

As $V=\bigoplus_{i\in I} V_i$ we have by [eqs: 12.10, 12.11 and 12.12] that $\forall i\in I$ we have $v'_i=w'_i$. Hence

$$\forall j\in J \text{ we have } v_i=v'_i=w'_i=w_i$$

proving that

$$\sum_{i\in J} V_i=\bigoplus_{i\in J} V_i$$

2. As $I=I_1\bigcup I_2$ we have I_1, I_2 so that by (1) we have

$$\sum_{i\in I_1} V_i=\bigoplus_{i\in I_1} V_i \text{ and } \sum_{i\in I_2} V_i=\bigoplus_{i\in I_2} V_i \tag{12.13}$$

As $V=\bigoplus_{i\in I} V_i$ we have by [theorem: 12.8] that $V=\sum_{i\in I} V_i$, further by [theorem: 12.3 (4)] $\sum_{i\in I} V_i=(\sum_{i\in I_1} V_i)+(\sum_{i\in I_2} V_i)$ so that

$$V=\left(\sum_{i\in I_1} V_i\right)+\left(\sum_{i\in I_2} V_i\right) \tag{12.14}$$

Let $v\in V$ then by the above there exist

$$v^1\in\sum_{i\in I_1} V_i,\ v^2\in\sum_{i\in I_2} V_i \text{ such that } v=v^1+v^2 \tag{12.15}$$

hence there exists $\{v_i^1\}_{i\in I_1}\subseteq V$ and $\{v_i^2\}_{i\in I_2}\subseteq V$ such that

$$\forall i\in I_1\ v_i^1\in V_i,\ \forall i\in I_2\ v_i^2\in V_i,\ v^1=\sum_{i\in I_1} v_i^1 \text{ and } v^2=\sum_{i\in I_2} v_i^2 \tag{12.16}$$

Assume that there exist also

$$w^1\in\sum_{i\in I_1} V_i,\ w^2\in\sum_{i\in I_2} V_i \text{ such that } v=w^1+w^2 \tag{12.17}$$

then there exists $\{w_i^1\}_{i\in I_1}\subseteq V$ and $\{w_i^2\}_{i\in I_2}\subseteq V$ such that

$$\forall i\in I_1\ w_i^1\in V_i,\ \forall i\in I_2\ w_i^2\in V_i,\ w^1=\sum_{i\in I_1} w_i^1 \text{ and } w^2=\sum_{i\in I_2} w_i^2 \tag{12.18}$$

As $I=I_1\bigcup I_2$ and $I_1\bigcap I_2=\varnothing$ we can define

$$\{v_i\}_{i\in I}\subseteq V \text{ by } v_i=\left\{\begin{array}{l}v_i^1 \text{ by } i\in I_1\\ v_i^2 \text{ by } i\in I_2\end{array}\right. \text{ and } \{w_i\}_{i\in I}\subseteq V \text{ by } w_i==\left\{\begin{array}{l}w_i^1 \text{ by } i\in I_1\\ w_i^2 \text{ by } i\in I_2\end{array}\right.$$

then we have by [eqs: 12.16 and 12.18]

$$\forall i\in I\ v_i=\left\{\begin{array}{l}v_i^1\in V_i \text{ if } i\in I_1\\ v_i^1\in V_i \text{ if } i\in I_2\end{array}\right.\in V_i \text{ and } w_i=\left\{\begin{array}{l}w_i^1\in V_i \text{ if } i\in I_1\\ w_i^1\in V_i \text{ if } i\in I_2\end{array}\right.\in V_i \tag{12.19}$$

Further

$$\sum_{i\in I} v_i \underset{[\text{theorem: } 11.43]}{=} \sum_{i\in I_1} v_i + \sum_{i\in I_2} v_i = \sum_{i\in I_1} v_i^1 + \sum_{i\in I_2} v_i^2 \underset{[\text{eq: } 12.16]}{=} v^1 + v^2 \underset{[\text{eq: } 12.15]}{=} v \tag{12.20}$$

$$\sum_{i\in I} w_i \underset{[\text{theorem: } 11.43]}{=} \sum_{i\in I_1} w_i + \sum_{i\in I_2} w_i = \sum_{i\in I_1} w_i^1 + \sum_{i\in I_2} w_i^2 \underset{[\text{eq: } 12.18]}{=} w^1 + w^2 \underset{[\text{eq: } 12.17]}{=} v \tag{12.21}$$

As $V = \bigoplus_{i\in I} V_i$ we have by [eqs: 12.19, 12.20 and 12.21] that

$$\forall i \in I \; v_i = w_i \tag{12.22}$$

Next

$$v^1 \underset{[\text{eq: } 12.16]}{=} \sum_{i\in I_1} v_i^1 = \sum_{i\in I_1} v_i \underset{[\text{eq: } 12.22]}{=} \sum_{i\in I_1} w_i = \sum_{i\in I_1} w_i^1 \underset{[\text{eq: } 12.18]}{=} w^1$$

$$v^2 \underset{[\text{eq: } 12.16]}{=} \sum_{i\in I_2} v_i^2 = \sum_{i\in I_2} v_i \underset{[\text{eq: } 12.22]}{=} \sum_{i\in I_2} w_i = \sum_{i\in I_2} w_i^2 \underset{[\text{eq: } 12.18]}{=} w^2$$

proving that

$$v = v^1 + v^2 = w^1 + w^2 \wedge v^1 = w^1 \wedge v^2 = w^2$$

So that we have, taking in account [eqs: 12.15, 12.18] and [definition: 11.195] that

$$V = \left(\sum_{i\in I_1} V_i\right) \oplus \left(\sum_{i\in I_2} V_i\right)$$

□

Theorem 12.13. *If V is a vector space over a field, I finite, $\{V_i\}_{i\in I} \subseteq \mathcal{P}(V)$ a finite family of sub spaces of V and $\sigma: J \rightarrow I$ a bijection such that $V = \bigoplus_{i\in J} V_{\sigma(i)}$ then $V = \bigoplus_{i\in I} V_i$.*

Proof. As $V = \bigoplus_{i\in J} V_{\sigma(i)}$ we have by [corollary: 12.8] that

$$V = \sum_{i\in J} V_{\sigma(i)} \tag{12.23}$$

and

$$\forall k \in I \text{ we have } V_{\sigma(k)} \bigcap \sum_{i \in I \setminus \{k\}} V_{\sigma(i)} = \{0\} \tag{12.24}$$

Now using [theorem: 12.3] we have that $\sum_{i \in J} V_{\sigma(i)} = \sum_{i \in I} V_i$ so that [eq: 12.23] becomes

$$V = \sum_{i \in I} V_i \tag{12.25}$$

Let $l \in I$ and take $k = \sigma^{-1}(l) \in J$ so that $l = \sigma(k)$. Let $i \in \sigma(J \setminus \{k\})$ then $\exists j \in J$ with $j \neq k$ such that $i = \sigma(j)$. Assume that $i = l$ then $\sigma(j) = i = l = \sigma(k) \underset{\sigma \text{ is injective}}{\Rightarrow} j = k$ contradicting $j \neq k$, hence $i \neq l$ so that $i \in \sigma(J) \setminus \{l\}$, proving that

$$\sigma(J \setminus \{k\}) \subseteq \sigma(J) \setminus \{l\}$$

On the other hand if $i \in \sigma(J) \setminus \{l\}$ then $\exists j \in J$ such that $i = \sigma(j)$ and $i \neq l$, assume that $j = k$ then $i = \sigma(j) = \sigma(k) = l$ contradicting $i \neq l$, hence $j \neq k$ so that $i \in \sigma(J \setminus \{k\})$, proving, together with the above, that

$$\sigma(J) \setminus \{l\} = \sigma(J \setminus \{k\})$$

As $\sigma: J \rightarrow I$ is a bijection we have that $\sigma(J) = I$ so that $\sigma(J \setminus \{k\}) = I \setminus \{l\}$ hence

$$\sigma_{|J \setminus \{k\}}: J \setminus \{k\} \rightarrow \sigma(J \setminus \{k\}) = I \setminus \{l\} \text{ is a bijection}$$

Applying then [theorem: 12.3] we have that

$$\sum_{i \in J \setminus \{k\}} V_{\sigma(i)} = \sum_{i \in I \setminus \{k\}} V_i$$

so that

$$V_l \bigcap \sum_{i \in I \setminus \{k\}} V_i = V_{\sigma(k)} \bigcap \sum_{i \in J \setminus \{k\}} V_{\sigma(i)} \underset{[\text{eq: } 12.24]}{=} \{0\}$$

Combining this with [eq: 12.25] proves, using [theorem: 12.8], that

$$V = \bigoplus_{i \in I} V_i$$

□

The following exsmple shows that the direct interal sum defined in this section is indeed a extension of the internal direct sum in [definition: 11.195].

Example 12.14. Let $\{V_i\}_{i \in \{1, \ldots, 2\}} \subseteq \mathcal{P}(V)$ be a family of two sub-spaces of a vector space over a field F then

$$V = \bigoplus_{i \in \{1,2\}} V_i$$

then

$$V = V_1 \oplus V_2$$

Proof. If $V = \bigoplus_{i \in \{1,2\}} V_i$ then we have

$$V \underset{[\text{theorem: } 12.11]}{=} \left(\sum_{i \in \{1\}} \right) \oplus V_2 \underset{[\text{theorem: } 12.3]}{=} V_1 \oplus V_2$$

□

Theorem 12.15. *Let V be a vector space over a field F, $n \in \mathbb{N}$ finite such that*

$$V = \bigoplus_{i \in \{1, \ldots, n\}} V_i$$

then V is finite dimensional and

$$\dim(V) = \sum_{i \in \{1, \ldots, n\}} \dim(V_i)$$

Proof. W use mathematical induction for the proof, so let

$$S = \left\{ n \in \mathbb{N} | \text{If } V = \bigoplus_{i \in \{1, \ldots, n\}} V_i \text{ where } V_i \text{ are finite dimensional then } V \text{ is finite dimensional and } \dim(V) = \sum_{i \in \{1, \ldots, n\}} \dim(V_i) \right\}$$

then we have:

$1\in S$. If $V=\bigoplus_{i\in\{1\}}V_i$ then by [example: 12.7] $V=V_1$, hence V is finite dimensional and

$$\dim(V)=\dim(V_1)=\sum_{i\in\{1\}}\dim(V_i)$$

proving that $1\in S$

$n\in S\Rightarrow n+1\in S$. Assume that $V=\bigoplus_{i\in\{1,\ldots n+1\}}V_i$ where $\forall i\in\{1,\ldots,n+1\}$ V_i is a finite dimensional sub-space of F then we have by [theorem: 12.11] that

$$\sum_{i\in\{1,\ldots,n\}}V_i=\bigoplus_{i\in\{1,\ldots,n\}}V_i \text{ and } V=\left(\bigoplus_{i\in\{1,\ldots,n\}}V_i\right)\oplus V_{n+1} \tag{12.26}$$

As $n\in S$ $\sum_{i\in\{1,\ldots,n\}}V_i$ is finite dimensional and

$$\dim\left(\bigoplus_{i\in\{1,\ldots,n\}}V_i\right)=\sum_{i\in\{1,\ldots,n\}}\dim(V_i) \tag{12.27}$$

As V_{n+1} is also finite dimensional we can apply [theorem: 11.197] resulting in

$$V=\left(\bigoplus_{i\in\{1,\ldots,n\}}V_i\right)\oplus V_{n+1} \text{ is finite dimensional}$$

and

$$\dim\left(\left(\bigoplus_{i\in\{1,\ldots,n\}}V_i\right)\oplus V_{n+1}\right)=\dim\left(\bigoplus_{i\in\{1,\ldots,n\}}V_i\right)+\dim(V_{n+1})$$

substituting the above in [eq: 12.27] gives

$$\dim(V)=\sum_{i\in\{1,\ldots,n\}}\dim(V_i)+\dim(V_{n+1})=\sum_{i\in\{1,\ldots,n+1\}}\dim(V_i)$$

proving that $n+1\in S$. □

Theorem 12.16. *Let V,W be vector spaces over a field F such that $V=\bigoplus_{i\in I}V_i$ and $L:V\rightarrow W$ a linear isomorphism then*

$$W=\bigoplus_{i\in I}L(V_i)$$

Proof. Let $w\in W$ then as L is bijective there exist a $v\in V$ such that $L(v)=W$. As $V=\bigoplus_{i\in I}V_i$ there exist a $\{v_i\}_{i\in I}\subseteq V$ such that $\forall i\in I$ $v_i\in V_i$ and $v=\sum_{i\in I}v_i$. Define then

$$\{w_i\}_{i\in I}\subseteq W \text{ by } w_i=L(v_i)$$

then $\forall i\in I$ $w_i=L(v_i)\in L(W_i)$ and $w=L(v)=L(\sum_{i\in I}v_i)=\sum_{i\in I}L(v_i)=\sum_{i\in I}w_i$. Let $\{w_i'\}_{i\in I}\subseteq W$ be such that $\forall i\in I$ $w_i'\in L(V_i)$ and $w=\sum_{i\in I}w_i'$. Define

$$\{v_i'\}_{i\in I}\subseteq V \text{ by } L^{-1}(w_i')$$

then we have

$$\begin{array}{rcl} v & = & L^{-1}(w) \\ & = & L^{-1}\left(\sum_{i\in I}w_i'\right) \\ & \underset{L^{-1}\text{ is linear}}{=} & \sum_{i\in I}L^{-1}(w_i') \\ & = & \sum_{i\in I}v_i' \end{array}$$

As $V=\bigoplus_{i\in I}V_i$ we have that $\forall i\in I$ $v_i'=v_i$ so that $w_i=L(v_i)=L(v_i')=L(L^{-1}(w_i'))=w_i'$. Hence we have that

$$W=\bigoplus_{i\in I}L(V_i)$$

□

Chapter 13

Tensor product of vector spaces

Multilinear mappings are difficult to handle, the idea of the tensor product is to focus all attention a single pair of a vector space P and a fixed multilinear mapping between $\prod_{i\in\{1,\ldots,n\}} V_i$ and P such that every multilinear mapping between $\prod_{i\in\{1,\ldots,n\}} V_i$ and a vector space can be written as the composition of the fixed multilinear mapping and a linear mapping.

Definition 13.1. **(Tensor Product)** *Let $n\in\mathbb{N}$, $\{V_i\}_{i\in\{1,\ldots,n\}}$ a finite family of vector spaces over a field F then* ***a tensor product of $V_1,\ldots,V_n$*** *is a pair $\langle P,\nu\rangle$ where*

1. *P is a vector space over the field F*
2. *$\nu\in\mathrm{Hom}(V_1,\ldots,V_n;P)$ [see definition: 11.242]*
3. *$\mathrm{span}\left(\nu\left(\prod_{i\in\{1,\ldots,n\}} V_i\right)\right)=P$*
4. **(Universal Factorization property)** *If U is any vector space over F and $\varphi\in\mathrm{Hom}(V_1,\ldots,V_n;U)$ then there exist a $h\in\mathrm{Hom}(P,U)$ such that $\varphi=h\circ\nu$*

It turns out that the map h such that $\varphi=h\circ\nu$ is actually unique. Hence every multilinear mapping $\varphi\in\mathrm{Hom}(V_1,\ldots,V_n;U)$ can be represented uniquely by a linear mapping $h\in\mathrm{Hom}(P,U)$ such thast $\varphi=h\circ\nu$.

Theorem 13.2. *Let $n\in\mathbb{N}$, $\{V_i\}_{i\in\{1,\ldots,n\}}$ a finite family of vector spaces over a field F, U a vector space over the same field F and $\langle P,\nu\rangle$ a tensor product then $\forall\varphi\in\mathrm{Hom}(V_1,\ldots,V_n;U)$ there exist a* ***unique*** *linear mapping $h\in\mathrm{Hom}(P,U)$ such that $\varphi=h\circ\nu$*

Proof. As $\langle P,\nu\rangle$ is a tensor product P is a vector space, $\nu\in\mathrm{Hom}(V_1,\ldots,V_n;P)$, $P=\mathrm{span}\left(\nu\left(\prod_{i\in\{1,\ldots,n\}} V_i\right)\right)$. Let $\varphi\in\mathrm{Hom}(V_1,\ldots,V_n;U)$ then by definition there exist a $h\in\mathrm{Hom}(P,U)$ such that $\varphi=h\circ\nu$. Assume that there exist a $k\in\mathrm{Hom}(P,U)$ such that $\varphi=k\circ\nu$. Let $x\in P$ then, as $P=\mathrm{span}\left(\nu\left(\prod_{i\in\{1,\ldots,n\}} V_i\right)\right)$, there exist by [theorem: 11.102] a $\{y_i\}_{i\in\{1,\ldots,n\}}\subseteq\nu\left(\prod_{i\in\{1,\ldots,n\}} V_i\right)$ and a $\{\alpha_i\}_{i\in\{1,\ldots,n\}}\subseteq F$ such that

$$x=\sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot y_i \tag{13.1}$$

If $i\in\{1,\ldots,n\}$ then as $y_i\in\nu\left(\prod_{i\in\{1,\ldots,n\}} V_i\right)$ there exist a $z\in\prod_{i\in\{1,\ldots,n\}} V_i$ such that $y_i=\nu(z)$. Hence $h(y_i)=h(\nu(z))=(h\circ\nu)(z)=\varphi(z)=(k\circ\nu)(z)=k(\nu(z))=k(y_i)$ proving that

$$\forall i\in\{1,\ldots,n\}\text{ we have }h(y_i)=k(y_i) \tag{13.2}$$

So we have

$$\begin{aligned}
h(x) &\underset{[\text{eq: }13.1]}{=} h\left(\sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot y_i\right)\\
&= \sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot h(y_i)\\
&\underset{[\text{eq: }13.2]}{=} \sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot k(y_i)\\
&= k\left(\sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot y_i\right)\\
&= k(x)
\end{aligned}$$

Which, as $x\in P$ was chosen arbitrary, proves that

$$h=k$$

□

We show now that given a finite family of vector spaces $\{V_i\}_{i\in\{1,\ldots,n\}}$ there exist a least one tensor product. First we introduce the concept of a free vector space which is a vector space created from a given set and field.

Definition 13.3. *Let X be a non empty set and $\langle F,+,\cdot\rangle$ a field then $\mathcal{F}(X,F)$ is defined by*

$$\mathcal{F}(X,F)=\{f\in F^X \mid f^{-1}(F\setminus\{0\}) \text{ is finite}\}\subseteq F^X$$

In other words $\mathcal{F}(X,F)$ is the set of graphs of functions from X to F such that the functions are non zero for a finite number of elements in F.

Next we define the operators needed for a vector space.

Definition 13.4. *Let X be a non empty set and a field $\langle F,+,\cdot\rangle$ then we define:*

1. *If $f,g\in\mathcal{F}(X,F)$ then $f+g$ is defined by $(f+g)(x)=f(x)+g(x)$*
2. *If $f\in\mathcal{F}(X,F)$ and $\alpha\in F$ then $(\alpha\cdot f)(x)=\alpha\cdot f(x)$*
3. *$\forall a\in X$ we define $\delta_a: X\rightarrow F$ by $\delta_a(x)=\begin{cases}1 & \text{if } x=a\\ 0 & \text{if } x\in X\setminus\{a\}\end{cases}$*

Finally we define the free vector space.

Theorem 13.5. *Let X be a set and $\langle F,+,\cdot\rangle$ then $\langle\mathcal{F}(X,F),+,\cdot\rangle$ [see the above definition] is a vector space over the field $\langle F,+,\cdot\rangle$ where the neutral element is C_0. This vector space is called the **free vector space of X**.*

Proof. Let $f,g\in\mathcal{F}(X,F)$. If $x\in(f+g)^{-1}(F\setminus\{0\})$ then we have $(f+g)(x)\in F\setminus\{0\}$ so that $f(x)+g(x)\neq 0$. So we must have $f(x)\neq 0\vee g(x)\neq 0$ [otherwise $(f+g)(x)=f(x)+g(x)=0$], hence $x\in f^{-1}(F\setminus\{0\})\bigcup g^{-1}(F\setminus\{0\})$, proving that

$$(f+g)^{-1}(F\setminus\{0\})\subseteq f^{-1}(F\setminus\{0\})\bigcup g^{-1}(F\setminus\{0\})$$

As $f,g\in\mathcal{F}(X,F)$ we have that $f^{-1}(F\setminus\{0\}), g^{-1}(F\setminus\{0\})$ are finite hence using [theorems: 10.89, 6.33] we have that $(f+g)^{-1}(F\setminus\{0\})$ is a finite set so that $f+g\in\mathcal{F}(X,F)$ and

$$+: \mathcal{F}(X,F)\times\mathcal{F}(X,F)\rightarrow\mathcal{F}(X,F) \text{ is well defined}$$

Let $\alpha\in F$ and $f\in\mathcal{F}(X,F)$. If $x\in(\alpha\cdot f)^{-1}(F\setminus\{0\})$ then $\alpha\cdot f(x)\in F\setminus\{0\}$ so that $\alpha\cdot f(x)\neq 0$ hence we must have that $f(x)\neq 0$ or $x\in f^{-1}(F\setminus\{0\})$ proving $(\alpha\cdot f)^{-1}(F\setminus\{0\})\subseteq f^{-1}(F\setminus\{0\})$ [a finite set], so by [theorem: 10.89] $f^{-1}(F\setminus\{0\})$ is finite or $\alpha\cdot f\in\mathcal{F}(X,F)$, proving that

$$\cdot: F\times\mathcal{F}(X,F)\rightarrow\mathcal{F}(X,F) \text{ is well defined}$$

Further for $C_0: X\rightarrow F$ defined by $C_0(x)=0$ we have trivially that $(C_0)^{-1}(F\setminus\{0\})=\varnothing$ [which is finite] hence

$$C_0\in\mathcal{F}(X,F)$$

Next we check the axioms of a vector space:

1. $\langle\mathcal{F}(X,F),+\rangle$ is a Abelian group:

   associativity. If $f,g,h\in\mathcal{F}(X,F)$ then we have $\forall x\in X$ that

$$\begin{aligned}(f+(g+h))(x) &= f(x)+(g+h)(x)\\ &= f(x)+(g(x)+h(x))\\ &= (f(x)+g(x))+h(x)\\ &= (f+g)(x)+h(x)\\ &= ((f+g)+h)(x)\end{aligned}$$

   proving that $f+(g+h)=(f+g)+h$

   commutatitivity. If $f,g\in\mathcal{F}(X,F)$ then we have $\forall x\in X$ that

$$(f+g)(x)=f(x)+g(x)=g(x)+f(x)=(g+f)(x)$$

   proving that

$$f+g=g+f$$

neutral element. If $f\in\mathcal{F}(X,F)$ then we have $\forall x\in X$ that

$$(f+C_0)(x)=f(x)+C_0(x)=f(x)+0=f(x)=0+f(x)=C_0(x)+f(x)=(C_0+f)(x)$$

proving that

$$f+C_0=f=C_0+f$$

inverse element. Let $f\in\mathcal{F}(X,F)$ then define $(-f):X\to F$ by $(-f)(x)=-f(x)$. Let $x\in(-f)^{-1}(F\setminus\{0\})$ then $(-f)(x)\neq 0\Rightarrow -f(x)\neq 0\Rightarrow f(x)\neq 0$, proving that $x\in f^{-1}(F\setminus\{0\})$. Hence $(-f)^{-1}(F\setminus\{0\})\subseteq f^{-1}(F\setminus\{0\})$ [a finite set], proving that $(-f)^{-1}(F\setminus\{0\})$ is finite so that $-f\in\mathcal{F}(X,F)$. Next $\forall x\in X$ we have

$$\begin{aligned}(f+(-f))(x) &= f(x)+(-f)(x)\\ &= f(x)+(-(f(x)))\\ &= 0\\ &= C_0(x)\end{aligned}$$

proving that $(-f)+f\underset{\text{commutativity}}{=}f+(-f)=C_0$

2. Additional vector space axioms:

   a. If $\alpha\in F$ and $f,g\in\mathcal{F}(X,F)$ we have $\forall x\in X$ that

$$\begin{aligned}(\alpha\cdot(f+g))(x) &= \alpha\cdot(f+g)(x)\\ &= \alpha\cdot(f(x)+g(x))\\ &= \alpha\cdot f(x)+\alpha\cdot g(x)\\ &= (\alpha\cdot f)(x)+(\alpha\cdot g)(x)\\ &= (\alpha\cdot f+\alpha\cdot g)(x)\end{aligned}$$

   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 we have $\forall x\in X$ that

$$\begin{aligned}((\alpha+\beta)\cdot f)(x) &= (\alpha+\beta)\cdot f(x)\\ &= \alpha\cdot f(x)+\beta\cdot f(x)\\ &= (\alpha\cdot f)(x)+(\beta\cdot f)(x)\\ &= (\alpha\cdot f+\beta\cdot f)(x)\\ &= (\alpha\cdot f+\beta\cdot f)(x)\end{aligned}$$

   proving 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 we have $\forall x\in X$ that

$$\begin{aligned}((\alpha\cdot\beta)\cdot f)(x) &= (\alpha\cdot\beta)\cdot f(x)\\ &= \alpha\cdot(\beta\cdot f(x))\\ &= \alpha\cdot(\beta\cdot f)(x)\\ &= (\alpha\cdot(\beta\cdot f))(x)\end{aligned}$$

   proving that

$$(\alpha\cdot\beta)\cdot f=\alpha\cdot(\beta\cdot f)$$

   d. If $f\in\mathcal{F}(X,Y)$ then $\forall x\in X$ we have

$$(1\cdot f)(x)=1\cdot f(x)=f(x)$$

   proving that $1\cdot f=f$ □

We prove now that $\{\delta_a|a\in X\}$ is a basis for for $\mathcal{F}(X,F)$ and that X can be embedded in $\mathcal{F}(X,F)$ using the injection $\delta:X\to\mathcal{F}(X,F)$.

Theorem 13.6. *Let X be a non empty set and $\langle F,+,\cdot\rangle$ a field then*

1. $\delta:X\to\mathcal{F}(X,F)$ defined by $\delta(a)=\delta_a$ is a injection, hence X is embedded in $\mathcal{F}(X,F)$ and

$$\delta:X\to\delta(X)\textit{ is a bijection}$$

2. $\delta(X) = \{\delta(a) | a \in X\} = \{\delta_a | a \in X\}$ *is a basis of* $\mathcal{F}(X, F)$

Proof.

1. First we must ensure that $\forall a \in X$ $\delta_a \in \mathcal{F}(X, F)$, so let $x \in (\delta_a)^{-1}(F \setminus \{0\})$ then we must have that $(\delta_a)(x) \neq 0$. As $(\delta_a)(x) = \left\{ \begin{smallmatrix} 1 \text{ if } x = a \\ 0 \text{ if } x \in X \setminus \{a\} \end{smallmatrix} \right.$, we must have that $x = a$, hence $(\delta_a)^{-1}(F \setminus \{0\}) \subseteq \{a\}$ a finite set. Hence

$$\forall a \in X \text{ we have } \delta_a \in \mathcal{F}(X, F)$$

and the following function is well defined

$$\delta: X \to \mathcal{F}(X, F) \text{ where } \delta(a) = \delta_a$$

If $\delta(a) = \delta(b)$ then $\delta_a = \delta_b$. Assume that $a \neq b$ then $\delta_a(b) \underset{b \in X \setminus \{a\}}{=} 0$ and $\delta_b(b) = 1$, as $1 \neq 0$ in a field, we have $\delta_a(b) \neq \delta_b(b)$ contradicting $\delta_a = \delta_b$, hence $a = b$. Hence

$$\delta \text{ is injective}$$

2. Define

$$B = \{\delta_a | a \in X\} \subseteq \mathcal{F}(X, F)$$

Let $I \subseteq B$, I finite and $\{\alpha_i\}_{i \in I} \subseteq F$ such that $\sum_{i \in I} \alpha_i \cdot i = 0$. Let $j \in I \subseteq B$ then $\exists a \in X$ such that $j = \delta(a)$. If $i \in I \setminus \{j\} \subseteq B$ then $\exists b \in X$ such that $i = \delta(b)$ and as $i \neq j$ we must have $a \neq b$ hence $i(a) = \delta(b)(a) = \delta_b(a) = 0$. So we have

$$\begin{aligned} 0 &= \left(\sum_{i \in I} \alpha_i \cdot i\right)(a) \\ &= \sum_{i \in I} \alpha_i \cdot i(a) \\ &= \sum_{i \in I \setminus \{j\}} \alpha_i \cdot i(a) + \sum_{i \in \{j\}} \alpha_i \cdot i(a) \\ &= \sum_{i \in I \setminus \{j\}} \alpha_i \cdot 0 + \alpha_j \cdot j(a) \\ &= \alpha_j \cdot j(a) \\ &= \alpha_j \cdot \delta(a)(a) \\ &= \alpha_j \cdot \delta_a(a) \\ &= \alpha_j \cdot 1 \\ &= \alpha_j \end{aligned}$$

So we have proved that

$$\forall j \in I \ \alpha_j = 0$$

from which it follows that

$$B \text{ is linear independent} \tag{13.3}$$

Next if $f \in \mathcal{F}(X, F)$ then $f^{-1}(F \setminus \{0\})$ is finite and we have two cases to consider:

$\boldsymbol{f^{-1}(F \setminus \{0\}) = \varnothing}$**.** Then $f = C_0$ and, as X is non empty, there exist a $a \in X$. Then we can write f as $f = C_0 = 0 \cdot \delta_a$. Hence if we define $I = \{\delta_a\} \subseteq B$ and $\{\alpha_i\}_{i \in \{\delta_a\}} \subseteq F$ by $\alpha_{\delta_a} = 0$ we have that $f = 0 \cdot \delta_a = \sum_{i \in \{\delta_a\}} \alpha_i \cdot i$. Hence

$$f \in \operatorname{span}(B)$$

$\boldsymbol{f^{-1}(F \setminus \{0\}) \neq \varnothing}$**.** Then $I = \{\delta_a | a \in f^{-1}(F \setminus \{0\})\} \subseteq B$ is finite [see theorem: 6.44] . As δ is injective we have that

$$\delta: f^{-1}(F \setminus \{0\}) \to \delta(f^{-1}(F \setminus \{0\})) = \{\delta_a | a \in f^{-1}(F \setminus \{0\})\} \text{ is a bijection}$$

Define now

$$\{\alpha_i\}_{i \in I} \text{ by } \alpha_i = f(\delta^{-1}(i))$$

then we have for $x \in X$ either

$\boldsymbol{f(x) = 0}$**.** Let $i \in I$ then $\exists a \in f^{-1}(F \setminus \{0\})$ such that $i = \delta_a$, as $f(a) \neq 0$ we can not have that $a = x$ [because then $f(a) = f(x) = 0$], hence $i(x) = \delta_a(x) = 0$. In other words

$$\forall i \in I \text{ we have } i(x) = 0$$

So

$$\begin{aligned}\left(\sum_{i\in I}\alpha_i\cdot i\right)(x) &= \sum_{i\in I}\alpha_i\cdot i(x)\\ &= \sum_{i\in I}\alpha_i\cdot 0\\ &= 0\\ &= f(x)\end{aligned}$$

$\boldsymbol{f(x)\neq 0.}$ As $f(x)\neq 0$ we have that $x\in f^{-1}(F\setminus\{0\})$ so that $\delta_x\in I$. Let $i\in I\setminus\{\delta_x\}$ then $\exists a\in f^{-1}(F\setminus\{0\})$ with $a\neq x$ such that $i=\delta_a$, hence $i(x)=\delta_a(x)\underset{x\neq a}{=}0$. So

$$\begin{aligned}\left(\sum_{i\in I}\alpha_i\cdot i\right)(x) &= \sum_{i\in I}\alpha_i\cdot i(x)\\ &= \sum_{i\in I\setminus\{\delta_x\}}\alpha_i\cdot i(x)+\sum_{i\in\{\delta_x\}}\alpha_i\cdot i(x)\\ &= \sum_{i\in I\setminus\{\delta_x\}}\alpha_i\cdot i(x)+\alpha_{\delta(x)}\cdot\delta_x(x)\\ &= \sum_{i\in I\setminus\{\delta_{a_0},\delta_x\}}\alpha_i\cdot 0+f(\delta^{-1}(\delta_x))\cdot\delta_x(x)\\ &= f(\delta^{-1}(\delta_x))\cdot\delta_x(x)\\ &= f(x)\cdot 1\\ &= f(x)\end{aligned}$$

So in all cases we have $(\sum_{i\in I}\alpha_i\cdot i)(x)=f(x)$ proving that $f=\sum_{i\in I}\alpha_i\cdot i$, so, as $I\subseteq B$, we have

$$f\in\operatorname{span}(B)$$

So we have proved that $\forall f\in\mathcal{F}(X,F)\ f\in\operatorname{span}(B)\subseteq\mathcal{F}(X,F)$ proving that $\mathcal{F}(X,F)\subseteq\operatorname{span}(B)\subseteq\mathcal{F}(X,F)$, hence $\mathcal{F}(X,F)=\operatorname{span}(B)$ which as B is linear independent [see eq: 13.3] proves that

$$B=\delta(X)\text{ is a basis of }\mathcal{F}(X,F)$$

□

Giving a family $\{V_i\}_{i\in\{1,\ldots,n\}}$ of vector spaces over a field F, we can construct the free vector space $\langle\mathcal{F}(\prod_{i\in\{1,\ldots,\}}V_i),+,\cdot\rangle$ with basis $B=\delta(\prod_{i\in\{1,\ldots,n\}}V_i)$. We want now to create a new vector space and a surjective mapping from $\mathcal{F}(\prod_{i\in\{1,\ldots,\}}V_i)$ to this new vector space such that all elements of the form

$$\delta_{(v_1,\ldots,v_{i-1},x+\alpha\cdot y,v_{i+1},\ldots,v_n)}-\delta_{(v_1,\ldots,v_{i-1},x,v_{i+1},\ldots,v_n)}-\alpha\cdot\delta_{(v_1,\ldots,v_{i-1},y,v_{i+1},\ldots,v_n)}$$

are mapped on the neutral element of this newe vector space. The composition of δ and this new mapping will be a multilinear mapping from $\prod_{i\in\{1,\ldots,\}}V_i$ to this new space. The standard way to do this is by creating a factor space based on a sub-space so that the projection mapping maps this sub-space on the neutral element in the factor space [see theorems: 11.73, 11.74].

Definition 13.7. *Let $n\in\mathbb{N}$, $\{V_i\}_{i\in\{1,\ldots,n\}}$ a family of vector spaces over a field F and $\langle\mathcal{F}(\prod_{i\in\{1,\ldots,\}}V_i),+,\cdot\rangle$ the free vector space of $\prod_{i\in\{1,\ldots,n\}}V_i$ then*

$$\mathcal{N}\left(\prod_{i\in\{1,\ldots,n\}}V_i,F\right)\subseteq\mathcal{F}\left(\prod_{i\in\{1,\ldots,\}}V_i\right)$$

is defined by

$$\mathcal{N}\left(\prod_{i\in\{1,\ldots,n\}}V_i,F\right)=\bigcup_{i\in\{1,\ldots,n\}}\mathcal{N}_i$$

where

$$\mathcal{N}_i=\left\{\delta_{[i\rightarrow x+\alpha\cdot y]v}-\delta_{[i\rightarrow x]v}-\alpha\cdot\delta_{[i\rightarrow y]v}|v\in\prod_{i\in\{1,\ldots,n\}\setminus\{i\}}V_i,\ x,y\in V_i,\ \alpha\in F\right\}$$

and for $i \in \{1, \ldots, n\}$, $x \in \prod_{i \in \{1, \ldots, n\} \setminus \{i\}} V_i$ *and* $y \in V_i$ *we have*

$$\forall k \in \{1, \ldots, n\} \text{ we have } ([i \rightarrow y]x)_k = \begin{cases} y \text{ if } k = i \\ x_i \text{ if } k \in \{1, \ldots, n\} \setminus \{i\} \end{cases}$$

in other words

$$[i \rightarrow y]x = \{x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n\}$$

Definition 13.8. *Let* $n \in \mathbb{N}, \{V_i\}_{i \in \{1, \ldots, n\}}$ *a family of vector spaces over a field* F *nd* $\langle \mathcal{F}(\prod_{i \in \{1, \ldots, \}} V_i), +, \cdot \rangle$ *the free vector space of* $\prod_{i \in \{1, \ldots, n\}} V_i$ *then* $V_1 \odot \cdots \odot V_n$ *is the factor space [see theorem: 11.73]*

$$\mathcal{F}\left(\prod_{i \in \{1, \ldots, \}} V_i\right) / \operatorname{span}\left(\mathcal{N}\left(\prod_{i \in \{1, \ldots, n\}} V_i, F\right)\right)$$

giving the vector space

$$\langle V_1 \odot \cdots \odot V_n, +, \cdot \rangle = \left\langle \mathcal{F}\left(\prod_{i \in \{1, \ldots, \}} V_i\right) / \operatorname{span}\left(\mathcal{N}\left(\prod_{i \in \{1, \ldots, n\}} V_i, F\right)\right), +, \cdot \right\rangle$$

over the field F

Note 13.9. *In the above definition we must use the span of* $\mathcal{N}(\prod_{i \in \{1, \ldots, n\}} V_i, F)$ *because* $\mathcal{N}(\prod_{i \in \{1, \ldots, n\}} V_i, F)$ *is not a sub-space of* $\mathcal{F}(\prod_{i \in \{1, \ldots, \}} V_i)$.

Definition 13.10. *Let* $n \in \mathbb{N}, \{V_i\}_{i \in \{1, \ldots, n\}}$ *a family of vector spaces over a field* F *then we define*

$$\odot: \prod_{i \in \{1, \ldots, n\}} V_i \rightarrow V_1 \odot \cdots \odot V_n \text{ by } \odot = \pi \circ \delta$$

where:

1. $\delta: \prod_{i \in \{1, \ldots, n\}} V_i \rightarrow \mathcal{F}(\prod_{i \in \{1, \ldots, \}} V_i, F)$ *is defined in [theorem: 13.6] by*

$$\delta(x_1, \ldots, x_n) = \delta_{(x_1, \ldots, x_n)} \text{ and } \delta_{(x_1, \ldots, x_n)}(r_1, \ldots, r_n) = \begin{cases} 1 \text{ if } (x_1, \ldots, x_n) = (r_1, \ldots, r_n) \\ 0 \text{ if } (\Pi_{i \in \{1, \ldots, n\}} V_i) \setminus (x_1, \ldots, x_n) \end{cases}$$

2. $\pi: \mathcal{F}(\prod_{i \in \{1, \ldots, \}} V_i, F) \rightarrow \mathcal{F}(\prod_{i \in \{1, \ldots, \}} V_i, F) / \operatorname{span}(\mathcal{N}(\prod_{i \in \{1, \ldots, n\}} V_i, F))$ *is defined in [theorem: 11.74] by*

$$\pi(x) = \left\{ y | x - y \in \operatorname{span}\left(\mathcal{N}\left(\prod_{i \in \{1, \ldots, n\}} V_i, F\right)\right) \right\}$$

If $(v_1, \ldots, v_n) \in \prod_{i \in \{1, \ldots, n\}} V_i$ *then we note* $\odot(v_1, \ldots, v_n)$ *as* $v_1 \odot \cdots \odot v_n$ *and call this the tensor product of the vectors* $v_1, \ldots, v_n$.

Note 13.11. We use the word **tensor product** for two different things. Given a family $\{V_i\}_{i \in \{1, \ldots, n\}}$ of vector spaces then we have a tensor product of $V_1, \ldots, V_n$ [see definition: 13.1] or the tensor product of vectors in the above definition.

The following theorem will be needed to prove the universal factorization property of a tensor product.

Theorem 13.12. (Universal property) *Let* X, Z *be vector spaces over a field* F, Y *a sub-space of* X *and* $L \in \operatorname{Hom}(X, Z)$ *such that* $Y \subseteq \ker(L)$ *then there exist a* ***unique*** $h \in \operatorname{Hom}(X / Y, Z)$ *such that* $L = h \circ \pi_Y$. *Here* X / Y *and* π_Y *are defined in [theorems: 11.73, 11.74].*

Proof. Let $x, y \in X$ such that $\sim_Y[x] = \sim_Y[y]$ then $x - y \in Y$ hence, as $Y \subseteq \ker(L)$ we have $L(x - y) = 0$, so, as L is linear, $L(x) - L(y) = 0 \Rightarrow L(x) = L(y)$. This allows use to define the function

$$h: X / Y \rightarrow Z \text{ where } h(\sim[x]) = L(x)$$

Let $\sim_Y[x], \sim_Y[y] \in X / Y$ and $\alpha \in F$ then

$$h(\sim_Y[x] + \sim_Y[y]) = h(\sim_Y[x + y]) = L(x + y) = L(x) + L(y) = h(\sim_Y[x]) + h(\sim_Y[y])$$

and

$$h(\alpha \cdot \sim_Y[x]) = h(\sim_Y[\alpha \cdot x]) = L(\alpha \cdot x) = \alpha \cdot L(x) = \alpha \cdot h(\sim_Y[x])$$

proving that

$$h \in \operatorname{Hom}(X / Y, Z)$$

Further let $x \in X$ then $(h \circ \pi_Y)(x) = h(\pi_Y(x)) = h(\sim_Y[x]) = L(x)$ so that

$$L = h \circ \pi_Y$$

Finally if $k \in \operatorname{Hom}(X/Y, Z)$ is such that $L = k \circ \pi_Y$ then if $\sim_Y[x] \in X/Y$ we have

$$h(\sim_Y[x]) = L(x) = k(\pi_Y(x)) = k(\sim_Y[x])$$

so that

$$h = k$$

□

We are now ready to prove that for every family $\{V_i\}_{i \in \{1,\ldots,n\}}$ of vector spaces there exists a tensor product of $V_1, \ldots, V_n$.

Theorem 13.13. *Let $n \in \mathbb{N}, \{V_i\}_{i \in \{1,\ldots,n\}}$ a family of vector spaces over a field F then*

$$\langle V_1 \odot \cdots \odot V_n, \odot \rangle$$

is a tensor product [see definition: 13.1] of $\{V_i\}_{i \in \{1,\ldots,n\}}$. Using [definition: 13.1] we must prove that:

1. *$V_1 \odot \cdots \odot V_n$ is a vector space over F*
2. *$\odot \in \operatorname{Hom}(V_1, \ldots, V_n; V_1 \odot \cdots \odot V_n)$*
3. *$\operatorname{span}\left(\odot\left(\prod_{i \in \{1,\ldots,n\}} V_i\right)\right) = V_1 \odot \cdots \odot V_n$*
4. ***(Universal factorization property)** If U is a vector space over F and $\varphi \in \operatorname{Hom}(V_1,\ldots,V_n;U)$ then there exist a $h \in \operatorname{Hom}(V_1 \odot \cdots \odot V_n, U)$ such that $\varphi = h \circ \odot$.*

   Note* 13.14.** *Using [theorem: 13.2] h is **unique

Proof.

1. By [definition: 13.8] $\langle V_1 \odot \cdots \odot V_n, +, \cdot \rangle$ is a vector space over the field F
2. Let $i \in \{1, \ldots, n\}$, $\alpha \in F$, $x, y \in V_i$ and $v \in \prod_{i \in \{1,\ldots,n\} \setminus \{i\}} V_i$ and take

$$(v_1, \ldots, v_{i-1}, x + \alpha \cdot y, v_{i+1}, \ldots, v_n) = [i \rightarrow x + \alpha \cdot y]v$$
$$(v_1, \ldots, v_{i-1}, x, v_{i+1}, \ldots, v_n) = [i \rightarrow x]v$$
$$(v_1, \ldots, v_{i-1}, y, v_{i+1}, \ldots, v_n) = [i \rightarrow y]v$$

then

$$\delta_{[i \rightarrow x + \alpha \cdot y]v} - \delta_{[i \rightarrow x]v} - \alpha \cdot \delta_{[i \rightarrow y]v} \in \mathcal{N}\left[\prod_{i \in \{1,\ldots,n\}} V_i, F\right] \subseteq \operatorname{span}\left(\mathcal{N}\left[\prod_{i \in \{1,\ldots,n\}} V_i, F\right]\right)$$

so that by [theorem: 11.74]

$$\pi(\delta_{[i \rightarrow x + \alpha \cdot y]v} - \delta_{[i \rightarrow x]v} - \alpha \cdot \delta_{[i \rightarrow y]v}) = 0 \tag{13.4}$$

As by [theorem: 11.74] π is linear we have that

$$0 = \pi(\delta_{[i \rightarrow x + \alpha \cdot y]v} - \delta_{[i \rightarrow x]v} - \alpha \cdot \delta_{[i \rightarrow y]v}) = \pi(\delta_{[i \rightarrow x + \alpha \cdot y]v}) - \pi(\delta_{[i \rightarrow x]v}) - \alpha \cdot \pi(\delta_{[i \rightarrow y]v})$$

proving that

$$\pi(\delta_{[i \rightarrow x + \alpha \cdot y]v}) = \pi(\delta_{[i \rightarrow x]v}) + \alpha \cdot \pi(\delta_{[i \rightarrow y]v})$$

hence

$$\pi(\delta([i \rightarrow x + \alpha \cdot y]v)) = \pi(\delta([i \rightarrow x]v)) + \alpha \cdot \pi(\delta([y]v))$$

which as $\odot = \pi \circ \delta$ gives

$$\odot([i \rightarrow x + \alpha \cdot y]v) = \odot([i \rightarrow x]v) + \alpha \cdot \odot([i \rightarrow y]v)$$

or

$$\begin{aligned} \odot(v_1, \ldots, v_{i-1}, x + \alpha \cdot y, v_{i+1} \ldots v_n) &= \\ \odot(v_1, \ldots, v_{i-1}, x, v_{i+1} \ldots v_n) + \alpha \cdot \odot(v_1, \ldots, v_{i-1}, y, v_{i+1} \ldots v_n) \end{aligned}$$

proving by [theorem: 11.244] that

$$\odot \text{ is multilinear hence } \odot\in\mathrm{Hom}(V_1,\ldots,V_n;V_1\odot\cdots\odot V_n)$$

3. As $\odot: \prod_{i\in\{1,\ldots,n\}} V_i \rightarrow V_1\odot\cdots\odot V_n$ we have already that $\odot\left(\prod_{i\in\{1,\ldots,n\}} V_i\right)\subseteq V_1\odot\cdots\odot V_n$, so that by [theorems: 11.104, 11.106] that

$$\mathrm{span}\left(\odot\left(\prod_{i\in\{1,\ldots,n\}} V_i\right)\right)\subseteq V_1\odot\cdots\odot V_n \tag{13.5}$$

For the opposite inclusion, let $x\in V_1\odot\cdots\odot V_n$ then as π is a surjection [see theorem: 11.74] there exists a $y\in\mathcal{F}\left(\prod_{i\in\{1,\ldots,\}} V_i\right)$ such that $x=\pi(y)$. By [theorem: 13.6] the set $\left\{\delta(a)|a\in\prod_{i\in\{1,\ldots,n\}} V\right\}$ is a basis of $\mathcal{F}\left(\prod_{i\in\{1,\ldots,\}} V_i\right)$, hence there exist a finite $I\subseteq\left\{\delta(a)|a\in\prod_{i\in\{1,\ldots,n\}} V\right\}$ and $\{\alpha_i\}_{i\in I}\subseteq F$ such that $y=\sum_{i\in I}\alpha_i\cdot i$. So

$$x=\pi(y)=\pi\left(\sum_{i\in I}\alpha_i\cdot i\right)\underset{\text{linearity of }\pi}{=}\sum_{i\in I}\alpha_i\cdot\pi(i) \tag{13.6}$$

Now if $i\in I\subseteq\left\{\delta(a)|a\in\prod_{i\in\{1,\ldots,n\}} V\right\}$ then $\exists a\in\prod_{i\in\{1,\ldots,n\}} V_i$ such that $i=\delta(a)$, hence $\pi(i)=\pi(\delta(a))=\odot(a)\in\odot\left(\prod_{i\in\{1,\ldots,\}}\right)$. Hence by [eq: 13.6] it follows that $x\in\mathrm{span}\left(\odot\left(\prod_{i\in\{1,\ldots,n\}} V_i\right)\right)$. So we conclude that $V_1\odot\cdots\odot V_n\subseteq\mathrm{span}\left(\odot\left(\prod_{i\in\{1,\ldots,n\}} V_i\right)\right)$, which combined with [eq: 13.5] gives

$$\mathrm{span}\left(\odot\left(\prod_{i\in\{1,\ldots,n\}} V_i\right)\right)=V_1\odot\cdots\odot V_n$$

4. Using [theorem: 13.6] we have that

$$\delta\left(\prod_{i\in\{1,\ldots,n\}} V_i\right)=\left\{\delta(a)|a\in\prod_{i\in\{1,\ldots,n\}} V_i\right\}\text{ is a basis for }\mathcal{F}\left(\prod_{i\in\{1,\ldots,\}} V_i\right)$$

and

$$\delta:\prod_{i\in\{1,\ldots,n\}} V_i\rightarrow\delta\left(\prod_{i\in\{1,\ldots,n\}} V_i\right)=\left\{\delta(a)|a\in\prod_{i\in\{1,\ldots,n\}} V_i\right\}\text{ is a bijection}$$

Let $x\in\mathcal{F}\left(\prod_{i\in\{1,\ldots,n\}} V_i\right)$ then there exist a finite $I\subseteq\left\{\delta(a)|a\in\prod_{i\in\{1,\ldots,n\}} V_i\right\}$ and a $\{\alpha_i\}_{i\in I}\subseteq F$ such that $x=\sum_{i\in I}\alpha_i\cdot i$. Assume that there exists another finite $J\subseteq\left\{\delta(a)|a\in\prod_{i\in\{1,\ldots,n\}} V_i\right\}$ and $\{\beta_i\}_{i\in J}\subseteq J$ such that $x=\sum_{i\in J}\beta_j\cdot i$. Then by [theorem: 11.134] we have

$$\forall i\in I\setminus J\text{ that }\alpha_i=0,\ \forall J\setminus I\text{ that }\beta_i=0\text{ and }\forall i\in I\bigcap J\text{ that }\alpha_i=\beta_i \tag{13.7}$$

Hence

$$\begin{aligned}
\sum_{i\in I}\alpha_i\cdot\varphi(\delta^{-1}(i)) &= \sum_{i\in I\setminus J}\alpha_i\cdot\varphi(\delta^{-1}(i))+\sum_{i\in I\bigcap J}\alpha_i\cdot\varphi(\delta^{-1}(i))\\
&\underset{[\text{eq: }13.7]}{=} \sum_{i\in I\setminus J}0\cdot\varphi(\delta^{-1}(i))+\sum_{i\in I\bigcap J}\beta_i\cdot\varphi(\delta^{-1}(i))\\
&\underset{[\text{eq: }13.7]}{=} \sum_{i\in I\setminus J}\beta_i\cdot\varphi(\delta^{-1}(i))+\sum_{i\in I\bigcap J}\beta_i\cdot\varphi(\delta^{-1}(i))\\
&= \sum_{i\in I}\beta_i\cdot\varphi(\delta^{-1}(i))
\end{aligned}$$

The above ensures that the following function is well defined

$$\gamma:\mathcal{F}\left(\prod_{i\in\{1,\ldots,\}} V_i,F\right)\rightarrow U\text{ by }\gamma(x)=\sum_{i\in I}\alpha_i\cdot\varphi(\delta^{-1}(i))\text{where }x=\sum_{i\in I}\alpha_i\cdot i,\ I\text{ is finite}$$

If $x\in\prod_{i\in\{1,\ldots,\}} V_i$ then if we define $\{\alpha_i\}_{i\in\{\delta(x)\}}\subseteq F$ by $\alpha_{\delta(x)}=1$ we have

$$\sum_{i\in\{\delta(x)\}}\alpha_i\cdot i=\alpha_{\delta(x)}\cdot\delta(x)=1\cdot\delta(x)=\delta(x)$$

so that

$$(\gamma\circ\delta)(x)=\gamma(\delta(x))=\sum_{i\in\{\delta(x)\}}\alpha_i\cdot\varphi(\delta^{-1}(i))=\alpha_{\delta(x)}\cdot\varphi(\delta^{-1}(\delta(x)))=1\cdot\varphi(x)=\varphi(x)$$

proving that

$$\gamma\circ\delta=\varphi \tag{13.8}$$

Let $x,y\in\mathcal{F}\left(\prod_{i\in\{1,\ldots,\}}V_i,F\right)$ and $\alpha\in F$ then there exists finite $I,J\subseteq\left\{\delta(a)|a\in\prod_{i\in\{1,\ldots,n\}}V_i\right\}$, $\{\alpha_i\}_{i\in I}\subseteq F$ and $\{\beta_i\}_{i\in J}\subseteq F$ with $x=\sum_{i\in I}\alpha_i\cdot i$ and $y=\sum_{i\in J}\beta_i\cdot i$. Define then

$$\{\gamma_i\}_{i\in I}\subseteq F\text{ by }\gamma_i=\begin{cases}\alpha_i\text{ if }i\in I\setminus J\\ \alpha\cdot\beta_i\text{ if }j\in J\setminus I\\ \alpha_i+\alpha\cdot\beta_i\text{ if }i\in I\bigcap J\end{cases}$$

then we have

$$\begin{aligned}\sum_{i\in I\bigcup J}\gamma_i\cdot i &= \sum_{i\in I\setminus J}\gamma_i\cdot i+\sum_{i\in J\setminus I}\gamma_i\cdot i+\sum_{i\in I\bigcap J}\gamma_i\cdot i\\ &= \sum_{i\in I\setminus J}\alpha_i\cdot i+\sum_{i\in J\setminus I}\alpha\cdot\beta_i\cdot i+\sum_{i\in I\bigcap J}(\alpha_i+\alpha\cdot\beta_i)\cdot i\\ &= \sum_{i\in I\setminus J}\alpha_i\cdot i+\alpha\cdot\sum_{i\in J\setminus I}\beta_i\cdot i+\sum_{i\in I\bigcap J}\alpha_i\cdot i+\alpha\cdot\sum_{i\in I\bigcap J}\beta_i\cdot i\\ &= \left(\sum_{i\in I\setminus J}\alpha_i\cdot i+\sum_{i\in I\bigcap J}\alpha_i\cdot i\right)+\alpha\cdot\left(\sum_{i\in J\setminus I}\beta_i\cdot i+\sum_{i\in I\bigcap J}\beta_i\cdot i\right)\\ &= \sum_{i\in I\bigcup J}\alpha_i\cdot i+\sum_{i\in I\bigcup J}\beta_i\cdot i\\ &= x+\alpha\cdot y\end{aligned}$$

so that by definitiob of γ

$$\begin{aligned}\gamma(x+\alpha\cdot y) &=\\ \sum_{i\in I\bigcup J}\gamma_i\cdot\varphi(\delta^{-1}(i))\\ \sum_{i\in I\setminus J}\gamma_i\cdot\varphi(\delta^{-1}(i))+\sum_{i\in J\setminus I}\gamma_i\cdot\varphi(\delta^{-1}(i))+\sum_{i\in I\bigcap J}\gamma_i\cdot\varphi(\delta^{-1}(i)) &=\\ \sum_{i\in I\setminus J}\alpha_i\cdot\varphi(\delta^{-1}(i))+\sum_{i\in J\setminus I}\alpha\cdot\beta_i\cdot\varphi(\delta^{-1}(i))+\sum_{i\in I\bigcap J}(\alpha_i+\alpha\cdot\beta_i)\cdot\varphi(\delta^{-1}(i)) &=\\ \underbrace{\sum_{i\in I\setminus J}\alpha_i\cdot\varphi(\delta^{-1}(i))}_{(1)}+\underbrace{\alpha\cdot\sum_{i\in J\setminus I}\beta_i\cdot\varphi(\delta^{-1}(i))}_{(2)}+\underbrace{\sum_{i\in I\bigcap J}\alpha_i\cdot\varphi(\delta^{-1}(i))}_{(1)}+\underbrace{\alpha\cdot\sum_{i\in I\bigcap J}\beta_i\cdot\varphi(\delta^{-1}(i))}_{(2)} &=\\ \underbrace{\sum_{i\in I}\alpha_i\cdot\varphi(\delta^{-1}(i))}_{(1)}+\underbrace{\alpha\cdot\sum_{i\in I}\beta_i\cdot\varphi(\delta^{-1}(i))}_{(2)} &=\\ \gamma(x)+\alpha\cdot\gamma(y)\end{aligned}$$

proving that γ is linear, hence

$$\gamma\in\operatorname{Hom}\left(\mathcal{F}\left(\prod_{i\in\{1,\ldots,n\}}V_i,F\right),U\right) \tag{13.9}$$

Let $x\in\mathcal{N}\left(\prod_{i\in\{1,\ldots,n\}}V_i,F\right)\subseteq\mathcal{F}(\prod_{i\in I}V_i,F)$ then by [definition: 13.7] there exist a $i\in\{1,\ldots,n\}$, $v\in\prod_{i\in\{1,\ldots,n\}\setminus\{i\}}V_i$, $x,y\in V_i$ and $\alpha\in F$ such that

$$x=\delta_{[i\to x+\alpha\cdot y]v}-\delta_{[i\to x]v}-\alpha\cdot\delta_{[i\to y]v}$$

Then

$$\begin{array}{rl}\gamma(x) & = \\ \gamma(\delta_{[i\rightarrow x+\alpha\cdot y]v}-\delta_{[i\rightarrow x]v}-\alpha\cdot\delta_{[i\rightarrow y]v}) & \underset{[\text{eq: } 13.9]}{=} \\ \gamma(\delta_{[i\rightarrow x+\alpha\cdot y]v})-\gamma(\delta_{[i\rightarrow x]v})-\alpha\cdot\gamma(\delta_{[i\rightarrow y]v}) & = \\ \gamma(\delta([i\rightarrow x+\alpha\cdot y]v))-\gamma(\delta([i\rightarrow x]v))-\alpha\cdot\gamma(\delta([i\rightarrow y]v)) & \underset{[\text{eq: } 13.8]}{=} \\ \varphi([i\rightarrow x+\alpha\cdot y]v)-\varphi([i\rightarrow x]v)-\alpha\cdot\varphi([i\rightarrow y]v) & = \\ \varphi(v_1,\ldots,v_{i-1},x+\alpha\cdot y,v_{i+1},\ldots,v_n)-\varphi(v_1,\ldots,v_{i-1},x,v_{i+1},\ldots,v_n)-\alpha\cdot\varphi(v_1,\ldots,v_{i-1},y,v_{i+1},\ldots,v_n) & \underset{\varphi\in\text{Hom}(V_1,\ldots,V_n;U)}{=} \\ 0 & = \end{array}$$

so that $\mathcal{N}\left(\prod_{i\in\{1,\ldots,n\}}V_i,F\right)\subseteq\ker(\gamma)$. By [theorem: 11.183] $\ker(\gamma)$ is a sub-space, hence using [theorems: 11.104, 11.106] we have

$$\text{span}\left(\mathcal{N}\left(\prod_{i\in\{1,\ldots,n\}}V_i,F\right)\right)\subseteq\ker(\gamma) \tag{13.10}$$

Finally [eqs: 13.9 and 13.10] allows us to apply [theorem: 13.12] to find a unique

$$h\in\text{Hom}\left(\mathcal{F}\left(\prod_{i\in\{1,\ldots,n\}}V_i,\right)/\text{span}\left(\mathcal{N}\left(\prod_{i\in\{1,\ldots,n\}}V_i,F\right)\right),U\right)$$

such that

$$\gamma=h\circ\pi$$

As $\mathcal{F}\left(\prod_{i\in\{1,\ldots,n\}}V_i,\right)/\text{span}\left(\mathcal{N}\left(\prod_{i\in\{1,\ldots,n\}}V_i,F\right)\right)\underset{\text{def}}{=}V_1\odot\cdots\odot V_n$ we have

$$\exists!h\in\text{Hom}(V_1\odot\ldots\odot V_n,U)\text{ with }\gamma=h\circ\pi$$

Using then [eq: 13.8] we have that $\varphi=\gamma\circ\delta=(h\circ\pi)\circ\delta=h\circ(\pi\circ\rho)\underset{[\text{definition: } 13.10]}{=}h\circ\odot$ giving:

$$\varphi=h\circ\odot\text{ where }h\in\text{Hom}(V_1\odot\cdots\odot V_n,U)$$

□

So given a family $\{V_I\}_{i\in\{1,\ldots,n\}}$ of vector spaces over a field we have proved that there exist a tensor product of $V_1,\ldots,V_n$. Next we prove that this vector space is uniquely determined within a canonical linear isomorphism.

Theorem 13.15. *Let $n\in\mathbb{N}$, $\{V_i\}_{i\in\{1,\ldots,n\}}$ a finite family of vector space over a field F then if $\langle P,\nu\rangle$ and $\langle Q,\mu\rangle$ are tensor products of $V_1,\ldots,V_n$ then there exists a **unique** linear isomorphism*

$$k\colon P\rightarrow Q$$

such that

$$k\circ\nu=\mu$$

Proof. As $\mu\in\text{Hom}(V_1,\ldots,V_n;Q)$ we have, as $\langle P,\nu\rangle$ is a tensor product of $V_1,\ldots,V_n$, that there exist a $k\in\text{Hom}(P,Q)$ such that

$$k\circ\nu=\mu \tag{13.11}$$

Likewise, as $\nu\in\text{Hom}(V_1,\ldots,V_n;P)$ we have, as $\langle Q,\mu\rangle$ is a tensor product of $V_1,\ldots,V_n$, that there exist a $h\in\text{Hom}(Q,P)$ such that

$$h\circ\mu=\nu \tag{13.12}$$

Next we have

$$\mu\underset{[\text{eq: } 13.11]}{=}k\circ\nu\underset{[\text{eq: } 13.12]}{=}k\circ(h\circ\mu)=(k\circ h)\circ\mu\text{ and }\nu\underset{[\text{eq: } 13.12]}{=}h\circ\mu\underset{[\text{eq: } 13.11]}{=}h\circ(k\circ\nu)=(h\circ k)\circ\nu \tag{13.13}$$

If $y\in P$ then as $P=\text{span}\left(\nu\left(\prod_{i\in\{1,\ldots,n\}}V_i\right)\right)$ there exist by [theorem: 11.96] a $\{\alpha_i\}_{i\in\{1,\ldots,k\}}\subseteq F$ and $\{y_i\}_{i\in\{1,\ldots,k\}}\subseteq\nu\left(\prod_{i\in\{1,\ldots,n\}}V_i\right)$ such that $y=\sum_{i\in\{1,\ldots,k\}}\alpha_i\cdot y_i$. As $\{y_i\}_{i\in\{1,\ldots,k\}}\subseteq\nu\left(\prod_{i\in\{1,\ldots,n\}}V_i\right)$ we have by a consequence of the Axiom of Choice [see theorem: 3.130] that there exist a $\{x_i\}_{i\in\{1,\ldots,k\}}\subseteq\prod_{i\in\{1,\ldots,n\}}V_i$ such that $\forall i\in\{1,\ldots,k\}$ $\nu(x_i)=y_i$. Let

$$x=\sum_{i\in\{1,\ldots,k\}}\alpha_i\cdot x_i\in\prod_{i\in\{1,\ldots,n\}}V_i\text{ [a vector space by [theorem: 11.68]]}$$

then we have

$$\nu(x) = \nu\left(\sum_{i \in \{1,\ldots,\}} \alpha_i \cdot x_i\right) = \sum_{i \in \{1,\ldots,k\}} \alpha_i \cdot \nu(x_i) = \sum_{i \in \{1,\ldots,k\}} \alpha_i \cdot y_i = y$$

or

$$\forall y \in P \text{ there exist a } x \in \prod_{i \in \{1,\ldots,n\}} V_i \text{ such that } \nu(x) = y \tag{13.14}$$

Now if $y \in P$ we have

$$\begin{aligned} (h \circ k)(y) &\underset{[\text{eq: } 13.14]}{=} (h \circ k)(\nu(x)) \\ &= ((h \circ k) \circ \nu)(x) \\ &\underset{[\text{eq: } 13.13]}{=} \nu(x) \\ &\underset{[\text{eq: } 13.14]}{=} y \\ &= \mathrm{Id}_P(y) \end{aligned}$$

proving that $\forall y \in P$ we have $(h \circ k)(y) = \mathrm{Id}_P(y)$ or

$$h \circ k = \mathrm{Id}_P \tag{13.15}$$

If $y \in Q$ then as $Q = \mathrm{span}\left(\mu\left(\prod_{i \in \{1,\ldots,n\}} V_i\right)\right)$ there exist by [theorem: 11.96] a $\{\alpha_i\}_{i \in \{1,\ldots,k\}} \subseteq F$ and $\{y_i\}_{i \in \{1,\ldots,k\}} \subseteq \mu\left(\prod_{i \in \{1,\ldots,n\}} V_i\right)$ such that $y = \sum_{i \in \{1,\ldots,k\}} \alpha_i \cdot y_i$. As $\{y_i\}_{i \in \{1,\ldots,k\}} \subseteq \mu\left(\prod_{i \in \{1,\ldots,n\}} V_i\right)$ we have by a consequence of the Axiom of Choice [see theorem: 3.130] that there exist a $\{x_i\}_{i \in \{1,\ldots,k\}} \subseteq \prod_{i \in \{1,\ldots,n\}} V_i$ such that $\forall i \in \{1,\ldots,k\}$ $\mu(x_i) = y_i$. Let

$$x = \sum_{i \in \{1,\ldots,k\}} \alpha_i \cdot x_i \in \prod_{i \in \{1,\ldots,n\}} V_i \text{ [a vector space by [theorem: 11.68]]}$$

then we have

$$\mu(x) = \mu\left(\sum_{i \in \{1,\ldots,\}} \alpha_i \cdot x_i\right) = \sum_{i \in \{1,\ldots,k\}} \alpha_i \cdot \mu(x_i) = \sum_{i \in \{1,\ldots,k\}} \alpha_i \cdot y_i = y$$

or

$$\forall y \in Q \text{ there exist a } x \in \prod_{i \in \{1,\ldots,n\}} V_i \text{ such that } \mu(x) = y \tag{13.16}$$

Now

$$\begin{aligned} (k \circ h)(y) &\underset{[\text{eq: } 13.16]}{=} (k \circ h)(\mu(x)) \\ &= ((k \circ h) \circ \mu)(x) \\ &\underset{[\text{eq: } 13.13]}{=} \mu(x) \\ &\underset{[\text{eq: } 13.16]}{=} y \\ &= \mathrm{Id}_Q(y) \end{aligned}$$

proving that $\forall y \in P$ we have $(h \circ k)(y) = \mathrm{Id}_P(y)$ or

$$k \circ h = \mathrm{Id}_Q \tag{13.17}$$

From [eqs: 13.15, 13.17] it follows, using [theorem: 2.71], that

$$k: Q \rightarrow P \text{ is a bijection}$$

and as k is linear we have together with [eq: 13.11] that

$$k \text{ is a linear isomorphism and } k \circ \nu = \mu$$

Next we must prove that $k: Q \rightarrow P$ is unique. So assume that $k' \in \mathrm{Hom}(P, Q)$ such that $k' \circ \nu = \mu$, then by [eq: 13.11]

$$k' \circ \nu = k \circ \nu \tag{13.18}$$

Let $y \in P$ then by [eq: 13.14] there exists a $x \in \prod_{i \in \{1,\ldots,n\}} V_i$ such that $y = \nu(x)$ hence

$$k'(y) = k'(\nu(x)) = (k' \circ \nu)(x) \underset{[\text{eq: } 13.18]}{=} (k \circ \nu)(x) = k(\nu(x)) = k(y)$$

proving that

$$k = k'$$

□

Corollary 13.16. *Let $n\in\mathbb{N}$, $\{V_i\}_{i\in\{1,\ldots,n\}}$ a finite family of vector space over a field F and $\langle v, Q\rangle$ a tensor product of $V_1,\ldots,V_n$ then*

1. *There exist a **unique** linear isomorphism*
$$k: V_1\odot\cdots\odot V_n\to Q$$
such that
$$k\circ\odot=\nu$$
2. *There exist a **unique** linear isomorphism*
$$h: Q\to V_1\odot\cdots\odot V_n$$
such that
$$h\circ\nu=\odot$$

Proof. Using [theorem: 13.13] $\langle V_1\odot\cdots\odot V_n,\odot\rangle$ is a tensor product of $V_1,\ldots,V_n$.

1. By [theorem: 13.15] there exist a **unique** linear isomorphism
$$k: V_1\odot\cdots\odot V_n\to Q$$
such that
$$k\circ\odot=\nu$$
2. By [theorem: 13.15] there exist a **unique** linear isomorphism
$$h: Q\to V_1\odot\cdots\odot V_n$$
such that
$$h\circ\nu=\odot$$
□

Theorem 13.17. *Let V be a vector space over a field F and $\{V_i\}_{i\in\{1\}}$ the family defined by $V_1=V$ then if we define*
$$\mu: \prod_{i\in\{1\}} V_i\to V \text{ by } \mu((v))=v$$
then we have:

1. *$\langle V,\mu\rangle$ is a tensor product of $V_1,\ldots,V_1$*
2. *If $\langle U,\nu\rangle$ is another tensor product of $V_1,\ldots,V_1$ then there exist a linear isomorphism*
$$k: V\to U$$
such that
$$\nu: \prod_{i\in\{1\}} V_i\to U \text{ is defined by } \nu((v))=k(v)$$
3. *If W is a sub-space of V and $\{W_i\}_{i\in\{1\}}\subseteq\mathcal{P}(V)$ is defined by $W_i=W$ then*
$$\left\langle k(W),\nu_{|\prod_{i\in\{1\}}W_i}\right\rangle \text{ is a tensor produce and } k(W) \text{ is a sub-space of } U$$
4. *If $\{W_i\}_{i\in\{1,\ldots,n\}}$ is a family of sub-spaces of V so that $V=\bigoplus_{i\in\{1,\ldots,n\}}W_i$ then we have, if we define for $i\in\{1,\ldots,n\}$ $\{W_j^i\}_{j\in\{1\}}$ by $W_1^i=W_i$ that*
$$U=\bigoplus_{i\in\{1,\ldots,n\}} k(W_i)$$
and $\forall i\in\{1,\ldots,n\}$
$$\left\langle k(W_i),\otimes_{|\prod_{j\in\{1\}}W_j^i}\right\rangle \text{ is a tensor product of } W_1^i,\ldots,W_1^i \text{ and } k(W_i) \text{ is a sub-space of } U$$

Proof.

1. If $u,v\in V_1,\alpha\in F$ then trivially $\mu((u+\alpha\cdot v))=u+\alpha\cdot v=\mu((u))+\alpha\cdot\mu((v))$, so we have
$$\mu\in\mathrm{Hom}(V_1,\ldots,V_1;V) \tag{13.19}$$

Let $v\in V$ then for $(v)\in\prod_{i\in\{1\}}V_i$ we have $\mu((v))=v$ so that $V\subseteq\mu\left(\prod_{i\in\{1\}}V_i\right)$ which, as $\mu\left(\prod_{i\in\{1\}}V_i\right)\subseteq V$, proves that $V=\mu\left(\prod_{i\in\{1\}}V_i\right)$. Using [theorem: 11.106] we have then

$$V=\operatorname{span}\left(\mu\left(\prod_{i\in\{1,\ldots,1\}}V_i\right)\right) \tag{13.20}$$

Let $\varphi\in\operatorname{Hom}(V_1,\ldots,V_1;X)$ and define

$$h:V\rightarrow X \text{ by } h(v)=\varphi((v))$$

then for $u,v\in V$ and $\alpha\in F$ we have

$$h(u+\alpha\cdot v)=\varphi((u+\alpha\cdot v))=\varphi((u))+\alpha\cdot\varphi((v))=h(u)+\alpha\cdot h(v)$$

proving that

$$h\in\operatorname{Hom}(V,X) \tag{13.21}$$

If $x=(v)\in\prod_{i\in\{1\}}V_i$ then

$$(h\circ\mu)(x)=h(\mu((v)))=h(v)=\varphi((v))=\varphi(x)$$

resulting in

$$h\circ\mu=\varphi \tag{13.22}$$

By [eqs: 13.19, 13.20, 13.21 and 13.22] we have by the definition of the tensor product that

$$\langle V,\mu\rangle \text{ is a tensor product of } \{V_i\}_{i\in\{1\}} \tag{13.23}$$

2. If $\langle U,\nu\rangle$ is another tensor product of $\{V_i\}_{i\in\{1\}}$ then by [theorem: 13.15] there exist a **unique** linear isomorphism $k:V\rightarrow U$ such that $k\circ\mu=\nu$. Let $(v)\in\prod_{i\in\{1\}}V_i$ then

$$\nu((v))=k(\mu((v)))=k(v)$$

so that

$$\nu:\prod_{i\in\{1\}}V_i\rightarrow U \text{ is defined by } \nu((v))=k(v) \tag{13.24}$$

3. As W is a sub-space of V we have by [theorem: 11.183] and the fact that $k:V\rightarrow U$ is a isomorphism, hence a linear mapping, that

$$k(W) \text{ is a sub-space of } U \tag{13.25}$$

If $v\in\nu\left(\prod_{i\in\{1\}}W_i\right)$ then there exist a $(w)\in\prod_{i\in\{1\}}W_i\subseteq\prod_{i\in\{1\}}V_i$ such that

$$v=\nu((w))\underset{[\text{eq: }13.24]}{=}k(w)\in k(W)$$

proving that

$$\nu\left(\prod_{i\in\{1\}}W_i\right)\subseteq k(W), \tag{13.26}$$

Using [theorem: 11.104] on the above gives

$$\operatorname{span}\left(\nu\left(\prod_{i\in\{1\}}W_i\right)\right)\subseteq\operatorname{span}(k(W))\underset{[\text{eq: }13.25]\wedge[\text{theorem: }11.106]}{=}k(W) \tag{13.27}$$

If $v\in k(W)$ then there exist a $w\in W\Rightarrow(w)\in\prod_{i\in\{1\}}W_i\subseteq\prod_{i\in\{1\}}V_i$ such that $v=k(w)$, hence by [eq: 13.24] $v=k(w)=\nu((w))\in\nu\left(\prod_{i\in\{1\}}W_i\right)$. Hence

$$k(W)\subseteq\nu\left(\prod_{i\in\{1\}}W_i\right)\subseteq\operatorname{span}\left(\nu\left(\prod_{i\in\{1\}}W_i\right)\right)=\operatorname{span}\left(\nu_{|\prod_{i\in\{1\}}W_i}\left(\prod_{i\in\{1\}}W_i\right)\right) \tag{13.28}$$

which combined with [eq: 13.27] results in

$$k(W)=\operatorname{span}\left(\nu_{|\prod_{i\in\{1\}}W_i}\left(\prod_{i\in\{1\}}W_i\right)\right) \tag{13.29}$$

If $u,v\in W_1=W$ and $\alpha\in F$ then

$$\nu_{|\prod_{i\in\{1\}}W_i}((u+\alpha\cdot v))=\nu((u+\alpha\cdot v))=\nu((u))+\alpha\cdot\nu((v))=\nu_{|\prod_{i\in\{1\}}}(u)+\alpha\cdot\nu_{|\prod_{i\in\{1\}}}(v)$$

proving that [using [eq: 13.26]

$$\nu_{|\prod_{i\in\{1\}}W_i}\in \mathrm{Hom}(W_1,\ldots,W_1;k(W)) \tag{13.30}$$

If $\varphi\in\mathrm{Hom}(W_1,\ldots,W_1;Y)$ then by [theorem: 11.255] there exist a $\psi\in\mathrm{Hom}(V_1,\ldots,V_1;Y)$ such that $\psi_{|\prod_{i\in\{1\}}W_i}=\varphi$. As $\langle U,\nu\rangle$ is a tensor product of $V_1,\ldots,V_1$ there exist a $h:U\rightarrow Y$ such that $h\circ\nu=\psi$, define $f:k(W)\rightarrow Y$ by $f=h_{|k(W)}$. Let $x\in\prod_{i\in\{1\}}W_i$ then

$$\begin{array}{rcl}\left(f\circ\nu_{|\psi_{|\prod_{i\in\{1\}}W_i}}\right)(x) & = & f\left(\nu_{|\psi_{|\prod_{i\in\{1\}}W_i}}(x)\right)\\ & = & f(\nu(x))\\ & = & h_{|k(W)}(\nu(x))\\ & \underset{[\text{eq: }13.26]}{=} & h(\nu(x))\\ & = & \psi(x)\\ & \underset{x\in\prod_{i\in\{1\}}W_i}{=} & \psi_{|\prod_{i\in\{1\}}W_i}(x)\\ & = & \varphi(x)\end{array}$$

proving that

$$\text{If } \varphi\in\mathrm{Hom}(W_1,\ldots,W_1) \text{ then } f\circ\nu_{|\prod_{i\in\{i\}}W_i}=\varphi \tag{13.31}$$

Finally [eqs: 13.25, 13.29, 13.30 and 13.31] proves that

$$\left\langle k(W),\nu_{|\prod_{i\in\{1\}}W_i}\right\rangle \text{ is a tensor space of } W_1,\ldots,W_1 \text{ and } k(W) \text{ is a sub-space of } U$$

4. First using (3) we have $\forall i\in\{1,\ldots,n\}$

$$\left\langle k(W_i),\otimes_{|\prod_{j\in\{1\}}W_j^i}\right\rangle \text{ is a tensor product of } W_1^i,\ldots,W_1^i \text{ and } k(W_i) \text{ is a sub-space of } U$$

Further, as $V=\bigoplus_{i\in\{1,\ldots,n\}}W_i$ and $k:V\rightarrow U$ is a isomorphism, we have by [theorem: 12.16] that

$$U=\bigoplus_{i\in\{1,\ldots,n\}}k(W_i)$$ □

If $\{V_i\}_{i\in\{1,\ldots,n\}}$ is a family of vector spaces over a field and $\{W_i\}_{i\in\{1,\ldots,n\}}$ is a family of sub-spaces then a tensor product of W_1,..., W_n is in general not a sub-space of a tensor product of V_1,..., V_n. However given a tensor product of V_1,..., V_n we can find a tensor product that is a sub-space of the given tensor product. This was proved in the previous theorem for the case $n=1$ however it is also valid in the general case.

Theorem 13.18. *Let $n\in\mathbb{N}$, $\{V_i\}_{i\in\{1,\ldots,n\}}$ a family of vector spaces over a field F, $\{W_i\}_{i\in\{1,\ldots n\}}$ a family such that $\forall i\in\{1,\ldots,n\}$ W_i is a sub-space of V_i and $\langle V,\nu\rangle$ a tensor product of $V_1,\ldots,V_n$. Define*

$$W=\mathrm{span}\left(\nu\left(\prod_{i\in\{1,\ldots,n\}}W_i\right)\right)$$

then

$$W \text{ is a sub-space of } V \text{ [hence a vector space]}$$

and

$$\left\langle W,\nu_{|\prod_{i\in\{1,\ldots,n\}}W_i}\right\rangle \text{ is a tensor product of } W_1,\ldots,W_n$$

Proof. By [theorem: 2.139] $\prod_{i\in\{1,\ldots,n\}}W_i\subseteq\prod_{i\in\{1,\ldots,n\}}V_i$ so that

$$\nu\left(\prod_{i\in\{1,\ldots,n\}}W_i\right)\subseteq\nu\left(\prod_{i\in\{1,\ldots,n\}}V_i\right)\subseteq\mathrm{span}\left(\nu\left(\prod_{i\in\{1,\ldots,n\}}V_i\right)\right)=V$$

so that by [theorems: 11.105, 11.57]

$$W\underset{\text{definition}}{=}\mathrm{span}\left(\nu\left(\prod_{i\in\{1,\ldots,n\}}W_i\right)\right) \text{ is a vector space that is a sub-space of } V$$

If $i\in\{1,\ldots,n\}$, $x\in\prod_{j\in\{1,\ldots,n\}\setminus\{i\}} W_i$, $u,v\in X_i$ and $\alpha\in F$ then we have

$$\begin{aligned}\nu_{|\prod_{i\in\{1,\ldots,n\}}W_i}(x_1,\ldots,x_{i-1},u+\alpha\cdot v,x_{i+1},\ldots,x_n) &= \\ \nu(x_1,\ldots,x_{i-1},u+\alpha\cdot v,x_{i+1},\ldots,x_n) &= \\ \nu(x_1,\ldots,x_{i-1},u,x_{i+1},\ldots,x_n)+\alpha\cdot\nu(x_1,\ldots,x_{i-1},v,x_{i+1},\ldots,x_n) &= \\ \nu_{|\prod_{i\in\{1,\ldots,n\}}W_i}(x_1,\ldots,x_{i-1},u,x_{i+1},\ldots,x_n)+\alpha\cdot\nu_{|\prod_{i\in\{1,\ldots,n\}}W_i}(x_1,\ldots,x_{i-1},v,x_{i+1},\ldots,x_n)\end{aligned}$$

proving that

$$\nu_{|\prod_{i\in\{1,\ldots,n\}}W_i}\in\mathrm{Hom}(W_1,\ldots,W_n;W) \tag{13.32}$$

Next

$$\mathrm{span}\left(\nu_{|\prod_{i\in\{1,\ldots,n\}}W_i}\left(\prod_{i\in\{1,\ldots,n\}}W_i\right)\right)=\mathrm{span}\left(\nu\left(\prod_{i\in\{1,\ldots,n\}}W_i\right)\right)\underset{\text{def}}{=}W \tag{13.33}$$

Let U be a vector space over the field F and $\varphi\in\mathrm{Hom}(W_1,\ldots,W_n;U)$ then by [theorem: 11.255] there exist a $\psi\in\mathrm{Hom}(V_1,\ldots,V_n;U)$ such that $\psi_{|\prod_{i\in 1,\ldots,n}W_i}=\varphi$. As $\langle V,\nu\rangle$ is a tensor product there exist a $k\in\mathrm{Hom}(V,U)$ such that $\psi=k\circ\nu$. Define $h=k_{|W}$ then we have $\forall x\in\prod_{i\in\{1,\ldots,n\}}W_i$ that

$$\begin{array}{rcl}\left(h\circ\nu_{|\prod_{i\in\{1,\ldots,n\}}W_i}\right)(x) & = & h(\nu(x))\\ & \underset{\nu(\prod_{i\in\{1,\ldots,n\}}W_i)\subseteq W}{=} & h_{|W}(\nu(x))\\ & = & k(\nu(x))\\ & = & \psi(x)\\ & \underset{x\in\prod_{i\in\{1,\ldots,n\}}W_i}{=} & \psi_{|\prod_{i\in\{1,\ldots,n\}}W_i}(x)\\ & = & \varphi(x)\end{array}$$

proving that

$$\varphi=h\circ\nu_{|\prod_{i\in\{1,\ldots,n\}}W_i} \tag{13.34}$$

From [eqs: 13.32, 13.33 and 13.34] we conclude that

$$\left\langle W,\nu_{|\prod_{i\in\{1,\ldots,n\}}W_i}\right\rangle \text{ is a tensor product of } W_1,\ldots,W_n$$

Finally

As $\prod_{i\in\{1,\ldots,n\}}W_i\subseteq\prod_{i\in\{1,\ldots,n\}}V_i$ we have

$$\nu_{|\prod_{i\in\{1,\ldots,n\}}W_i}\left(\prod_{i\in\{1,\ldots,n\}}W_i\right)=\nu\left(\prod_{i\in\{1,\ldots,n\}}W_i\right)\subseteq\nu\left(\prod_{i\in\{1,\ldots,n\}}V_i\right)$$

so that

$$W=\mathrm{span}\left(\nu_{|\prod_{i\in\{1,\ldots,n\}}W_i}\left(\prod_{i\in\{1,\ldots,n\}}W_i\right)\right)\subseteq\mathrm{span}\left(\nu\left(\prod_{i\in\{1,\ldots,n\}}V_i\right)\right)=V$$

□

To simplify notation we use the following notation convention.

Convention 13.19. If $\{V_i\}_{i\in\{1,\ldots,n\}}$ is a family of vector spaces over a field then a tensor product is a pair $\langle V,\mu\rangle$ where V is a vector space and $\mu\in\mathrm{Hom}(V_1,\ldots,V_n;V)$. We have now two conventions:

Focus on the vector space of a vector product.

The notation

$$V_1\otimes\cdots\otimes V_n$$

means that $V_1\otimes\cdots\otimes V_n$ is a vector space and there exist a (unspecified) multilinear mapping $\mu\in\mathrm{Hom}(V_1,\ldots,V_n;V_1\otimes\cdots\otimes V_n)$ such that $\langle V_1\otimes\cdots\otimes V_n,\mu\rangle$ is a tensor product of $V_1,\ldots,V_n$. More specific it means that:

1. $V_1\otimes\cdots\otimes V_n$ is a a vector space over the same field F as (1)
2. There exist a $\mu\in\mathrm{Hom}(V_1,\ldots,V_n;V_1\otimes\cdots\otimes V_n)$ such that $\langle V_1\otimes\cdots\otimes V_n,\mu\rangle$ is a tensor product of $V_1,\ldots,V_n$ or more in detail:

   a. $V_1\otimes\cdots\otimes V_n=\mathrm{span}\left(\mu\left(\prod_{i\in\{1,\ldots,n\}}V_i\right)\right)$

b. If U is a vector space and $\varphi \in \text{Hom}(V_1, \ldots, V_n; U)$ then there exist a

$$h: \text{Hom}(V_1 \otimes \cdots \otimes V_n, U)$$

such that

$$\varphi = h \circ \mu$$

If $n = 2$ then we use the shorter notation

$$V_1 \otimes V_2 \text{ for } V_1 \otimes \cdots \otimes V_n$$

We slightly abuse our definition and says that $V_1 \otimes \cdots \otimes V_n$ is a tensor product of $V_1, \ldots, V_n$ when we actually should say that there exist a $\mu \in \text{Hom}(V_1, \ldots, V_n; V_1 \otimes \cdots \otimes V_n)$ such that $\langle V_1 \otimes \ldots \otimes V_n, \mu \rangle$ is a tensor product of $V_1, \ldots, V_n$. In many cases we also use the notation $\otimes$ for the assumed multilinear mapping that makes $V_1 \otimes \cdots \otimes V_n$ part of a tensor product.

One vector space and more then one multilinear mapping.

Then we use $V_1 \otimes \cdots \otimes V_n$ for the vector space and different symbols like $\otimes$, $\circledast$, $\otimes_1$, $\otimes_i$...for the multilinear mapping. Examples

$$\langle V_1 \otimes \cdots \otimes V_n, \otimes \rangle$$

$$\langle V_1 \otimes \cdots \otimes V_n, \circledast \rangle$$

$$\langle V_1 \otimes \cdots \otimes V_n, \otimes_i \rangle$$

$$\ldots$$

More then one vector space and multilinear mapping.

Then we use different symbols $\otimes$, $\circledast$, $\odot$, $\otimes_i$ ... as shown in the following examples

$$\langle V_1 \otimes \cdots \otimes V_n, \otimes \rangle$$

$$\langle V_1 \circledast \ldots \circledast V_n, \circledast \rangle$$

$$\langle V_1 \otimes_i \cdots \otimes_i V_n, \otimes_i \rangle$$

Using the above convention we can rephrase [theorems: 13.13, 13.15, 13.17 and 13.18].

Theorem 13.20. *Let $\{V_1\}_{i \in \{1, \ldots, n\}}$ be a family of vector spaces over a field then there exists a tensor product*

$$V_1 \otimes \cdots \otimes V_n$$

Proof. See [theorem: 13.13] and [convention: 13.19]. □

Theorem 13.21. *Let $\{V_i\}_{i \in \{1, \ldots, n\}}$ be a family of vector spaces over a field F then for*

$$\langle V_1 \otimes \cdots \otimes V_n, \otimes \rangle, \ \langle V_1 \otimes \cdots \otimes V_n, \circledast \rangle \text{ two tensor products based on } V_1, \ldots, V_n$$

then there exists a unique linear isomorphism

$$k: V_1 \otimes \cdots \otimes V_n \rightarrow V_1 \circledast \ldots \circledast V_n$$

such that

$$k \circ \otimes = \circledast$$

Proof. See [theorem: 13.15] and [convention: 13.19]. □

Theorem 13.22. *Let V be a vector space over a field F [defining the family of vector spaces $\{V_i\}_{i \in \{1\}}$ where $V_1 = V$] and $\langle V_1 \otimes \cdots \otimes V_1, \otimes \rangle$ a tensor product of $V_1, \ldots, V_1$. If we define*

$$\circledast: \prod_{i \in \{1\}} V_i \rightarrow V \text{ by } \circledast((v)) = v$$

then we have:

1. *$\langle V, \circledast \rangle$ is a tensor product of $V_1, \ldots, V_1$ in other words $V_1 \circledast \ldots \circledast V_1 = V$*
2. *If $\langle V_1 \otimes \cdots \otimes V_n, \otimes \rangle$ is another tensor product of $V_1, \ldots, V_1$ then here exists a isomorphism*

$$k: V \rightarrow V_1 \otimes \cdots \otimes V_1$$

such that

$$\otimes: \prod_{i \in \{1\}} V_i \rightarrow V_1 \otimes \cdots \otimes V_1 \text{ we have } \otimes((v)) = k(v)$$

3. *If W is a sub-space of V then for $\{W_i\}_{i\in\{1\}}\subseteq\mathcal{P}(V)$ defined by $W_1=W$ we have that there exist a tensor product $\langle W_1\otimes\cdots\otimes W_1, \otimes_{|\prod_{i\in\{1\}}W_i}\rangle$ of $W_1,\ldots,W_1$ such that*
   a. $W_1\otimes\cdots\otimes W_1=k(W)$
   b. *$W_1\otimes\cdots\otimes W_1$ is a sub-space of V*
4. *If $V=\bigoplus W_i$ then, if we define $\forall i\in\{1,\ldots,n\}$ $\{W_j^i\}_{j\in\{1\}}$ by $W_1^i=W_i$, we have that*

$$\forall i\in\{1,\ldots,n\} \text{ there exist a tensor product } W_1^i\otimes\cdots\otimes W_1^i \text{ of } W_1^i,\ldots,W_1^i$$

such that

$$W_1^i\otimes\cdots\otimes W_1^i \text{ is a subspace of } V_1\otimes\cdots\otimes V_1$$

and

$$V_1\otimes\cdots\otimes V_1=\bigoplus_{i\in\{1,\ldots,n\}} W_1^i\otimes\cdots\otimes W_1^i$$

Proof. See [theorem: 13.17] and [convention: 13.19] □

Theorem 13.23. *Let $n\in\mathbb{N}$, $\{V_i\}_{i\in\{1,\ldots,n\}}$ a family of vector spaces over a field F, $\{W_i)_{i\in\{1,\ldots,n\}}$ a family such that $\forall i\in\{1,\ldots,n\}$ W_i is a sub-space of V_i. If $\langle V_1\otimes\cdots\otimes V_n,\otimes\rangle$ is a tensor product of $V_1,\ldots,V_n$ then if we define $W_1\otimes\cdots\otimes W_n=\operatorname{span}\left(\otimes\left(\prod_{i\in\{1,\ldots,n\}}V_i\right)\right)$ we have that*

$$\left\langle W_1\otimes\cdots\otimes W_n,\otimes_{|\prod_{i\in\{1,\ldots,n\}}W_i}\right\rangle \text{ is a tensor product of } W_1,\ldots,W_n$$

and

$$W_1\otimes\cdots\otimes W_n\subseteq V_1\otimes\cdots\otimes V_n$$

Proof. See [theorem: 13.18] and [convention: 13.19] □

We prove now some useful theorems about tensor products

Theorem 13.24. *Let $n\in\mathbb{N}$, $\{V_i\}_{i\in\{1,\ldots,n\}}$ a finite family of vector space over a field F, a tensor product $V_1\otimes\cdots\otimes V_n$ and $(v_1,\ldots,v_n)\in\prod_{i\in\{1,\ldots,n\}}V_i$ then we have*

$$v_1\otimes\cdots\otimes v_n=0\Leftrightarrow\exists i\in\{1,\ldots,n\}\ v_i=0$$

Proof.

⇒. Let $v_1\otimes\cdots\otimes v_n=0$. Assume $\forall i\in\{1,\ldots,n\}$ we have $v_i\neq 0$ then by [theorem: 11.177] there exist a $L_i\in\operatorname{Hom}(V_i,F)$ such that $L_i(v_i)=1$ then by [theorem: 11.245] $\prod_{i\in\{1,\ldots,n\}}L_i\in\operatorname{Hom}(V_1,\ldots,V_n;F)$. As $\langle V_1\otimes\cdots\otimes V_n,\otimes\rangle$ is a tensor product there exist a $h\in\operatorname{Hom}(P,F)$ such that $\prod_{i\in\{1,\ldots,n\}}L_i=h\circ\otimes$. Then

$$\begin{array}{rcl} 0 & \underset{\mu v_1\otimes\cdots\otimes v_n=0}{=} & h(v_1\otimes\cdots\otimes v_n)\\ & = & h(\otimes(v_1,\ldots,v_n))\\ & = & (h\circ\otimes)(v_1,\ldots,v_n)\\ & = & \left(\prod_{i\in\{1,\ldots,n\}}L_i\right)(v_1,\ldots,v_n)\\ & = & \prod_{i\in\{1,\ldots,n\}}L_i(v_i)\\ & \underset{\text{[theorem: 11.37]}}{=} & 1 \end{array}$$

giving the contradiction $0=1$. Hence we must have that $\exists i\in\{1,\ldots,n\}$ such that $v_i=0$.

⇐. This follows from [theorem: 11.243] and the fact that $\otimes\in\operatorname{Hom}(V_1,\ldots,V_n;V_1\otimes\cdots\otimes V_n)$ □

Lemma 13.25. *Let $\{V_i\}_{i\in\{1,2\}}$ family of two vector spaces over a field f with a tensor product $\langle V_1\otimes V_2,\otimes\rangle$. Take $z\in V_1\otimes V_2$ and define*

$$K_z=\left\{k\in\mathbb{N}|\exists\{u_i\}_{i\in\{1,\ldots,k\}}\subseteq V_1,\{v_i\}_{i\in\{1,\ldots,k\}}\subseteq V \text{ such that } z=\sum_{i\in\{1,\ldots,k\}}u_i\otimes v_i\right\}$$

then

$$K_z \neq 0$$

so that by [theorem: 10.19]

$$k_z = \min(K_z) \text{ exist}$$

Then for every $\{u_i\}_{i\in\{1,\ldots,k_z\}} \subseteq V_1, \{v_i\}_{i\in\{1,\ldots,k_z\}} \subseteq V$ *such that*

$$z = \sum_{i\in\{1,\ldots,k_z\}} u_i \otimes v_i$$

we have then that

$$\{u_i\}_{i\in\{1,\ldots,k_z\}} \text{ and } \{v_i\}_{i\in\{1,\ldots,k_z\}} \text{ are linear independent}$$

Proof. Let $z \in V_1 \otimes V_1$. As by definition $V_1 \otimes V_2 = \operatorname{span}\left(\otimes\left(\prod_{i\in\{1,\ldots,n\}} V_i\right)\right)$ there exists by [theorem: 11.102] a $\{x_i\}_{i\in\{1,\ldots,n\}} \subseteq \otimes\left(\prod_{i\in\{1,\ldots,n\}} V_i\right)$ and $\{\alpha_i\}_{i\in\{1,\ldots,n\}} \subseteq F$ such that $z = \sum_{i\in\{1,\ldots,n\}} \alpha_i \cdot x_i$. As $\{x_i\}_{i\in\{1,\ldots,n\}} \subseteq \otimes\left(\prod_{i\in\{1,\ldots,n\}} V_i\right)$ their exist by [theorem: 3.130] a $\{(u_i', v_i)\}_{i\in\{1,\ldots,n\}} \subseteq \prod_{i\in\{1,\ldots,n\}} V_i$ such that $\forall i \in \{1,\ldots,.n\}$ we have $x_i = \otimes(u_i', v_i)$. Define $\{u_i\}_{i\in\{1,\ldots,n\}} \subseteq V_1$ by $u_i = \alpha_i \cdot u_i'$ then we have

$$\begin{aligned}
\sum_{i\in\{1,\ldots,n\}} u_i \otimes v_i &= \sum_{i\in\{1,\ldots,n\}} (\alpha_i \cdot u_i') \otimes v_i \\
&= \sum_{i\in\{1,\ldots,n\}} \alpha_i \cdot (u_i' \otimes v_i) \\
&= \sum_{i\in\{1,\ldots,n\}} \alpha_i \cdot \otimes(u_i', v_i) \\
&= \sum_{i\in\{1,\ldots,n\}} \alpha_i \cdot x_i \\
&= z
\end{aligned}$$

proving that $n \in K_z$ or

$$K_z \neq 0$$

So by [theorem: 10.19] $k_z = \min(K_z)$ exist. Let $\{u_i\}_{i\in\{1,\ldots,k_z\}} \subseteq V_1$ and $\{v_i\}_{i\in\{1,\ldots,k_z\}} \subseteq V_2$ such that $z = \sum_{i\in\{1,\ldots,k_z\}} u_i \otimes v_i$ then we have:

1. Assume that $\{u_i\}_{i\in\{1,\ldots,k_z\}}$ is linear dependent then by]theorem: 11.123] there exists a $l \in \{1,\ldots,k_z\}$ and $\{\alpha_i\}_{i\in\{1,\ldots,k_z\}\setminus\{l\}} \subseteq F$ such that $u_l = \sum_{i\in\{1,\ldots,k_z\}\setminus\{l\}} \alpha_i \cdot u_i$. Define now

   $$\{w_i\}_{i\in\{1,\ldots,k_z\}\setminus\{l\}} \subseteq V_2 \text{ by } w_i = v_i + \alpha_i \cdot v_l$$

   then we have

   $$\begin{aligned}
   \sum_{i\in\{1,\ldots,k_z\}\setminus\{l\}} u_i \otimes w_i &= \sum_{i\in\{1,\ldots,k_z\}\setminus\{l\}} u_i \otimes (v_i + \alpha_i \cdot v_l) \\
   &= \sum_{i\in\{1,\ldots,k_z\}\setminus\{l\}} (u_i \otimes v_i + \alpha_i \cdot (u_i \otimes v_l)) \\
   &= \sum_{i\in\{1,\ldots,k_z\}\setminus\{l\}} u_i \otimes v_i + \sum_{i\in\{1,\ldots,k_z\}\setminus\{l\}} \alpha_i \cdot (u_i \otimes v_l) \\
   &= \sum_{i\in\{1,\ldots,k_z\}\setminus\{l\}} u_i \otimes v_i + \sum_{i\in\{1,\ldots,k_z\}\setminus\{l\}} ((\alpha_i \cdot u_i) \otimes v_l) \\
   &= \sum_{i\in\{1,\ldots,k_z\}\setminus\{l\}} u_i \otimes v_i + \left(\sum_{i\in\{1,\ldots,k_z\}\setminus\{l\}} \alpha_i \cdot u_i\right) \otimes v_l \\
   &= \sum_{i\in\{1,\ldots,k_z\}\setminus\{l\}} u_i \otimes v_i + u_l \otimes v_l \\
   &= \sum_{i\in\{1,\ldots,k_z\}} u_i \otimes v_i \\
   &= z
   \end{aligned}$$

Using [corollary: 10.92] we have that $\operatorname{card}(\{1,\ldots,k_z\}\backslash\{l\}) = k_z - 1$ so that there exist a bijection $\beta: \{1,\ldots,k_z-1\} \rightarrow \{1,\ldots,k_z\}\backslash\{l\}$ hence if we define

$$\{u_i')_{i\in\{1,\ldots,k_z-1\}} \subseteq V_1 \text{ by } u_i' = u_{\beta(i)} \text{ and } \{v_i')_{i\in\{1,\ldots,k_z-1\}} \subseteq V_2 \text{ by } v_i' = w_{\beta(i)}$$

then we have

$$\sum_{i\in\{1,\ldots,k_z-1\}} u_i' \otimes v_i' = \sum_{i\in\{1,\ldots,k_z-1\}} v_{\beta(i)} \otimes w_{\beta(i)} \underset{[\text{theorem: } 11.36]}{=} \sum_{i\in\{1,\ldots,k_z\}\backslash\{l\}} u_i \otimes w_i = z$$

so that $k_z - 1 \in K_z$ contradicting $k_z = \min(K_z)$. Hence $\{u_i\}_{i\in\{1,\ldots,k_z\}}$ is not lineard dependent, proving that $\{u_i\}_{i\in\{1,\ldots,k_z\}}$ is linear dependent.

2. Assume that $\{v_i\}_{i\in\{1,\ldots,k_z\}}$ is linear dependent then by]theorem: 11.123] there exists a $l \in \{1,\ldots,k_z\}$ and $\{\alpha_i\}_{i\in\{1,\ldots,k_z\}\backslash\{l\}} \subseteq F$ such that $u_l = \sum_{i\in\{1,\ldots,k_z\}\backslash\{l\}} \alpha_i \cdot v_i$. Define now

$$\{w_i\}_{i\in\{1,\ldots,k_z\}\backslash\{l\}} \subseteq V_1 \text{ by } w_i = u_i + \alpha_i \cdot u_l$$

then we have

$$\begin{aligned}
\sum_{i\in\{1,\ldots,k_z\}\backslash\{l\}} w_i \otimes v_i &= \sum_{i\in\{1,\ldots,k_z\}\backslash\{l\}} (u_i + \alpha_i \cdot u_l) \otimes v_i \\
&= \sum_{i\in\{1,\ldots,k_z\}\backslash\{l\}} (u_i \otimes v_i + \alpha_i \cdot (u_l \otimes v_i)) \\
&= \sum_{i\in\{1,\ldots,k_z\}\backslash\{l\}} u_i \otimes v_i + \sum_{i\in\{1,\ldots,k_z\}\backslash\{l\}} \alpha_i \cdot (u_l \otimes v_i) \\
&= \sum_{i\in\{1,\ldots,k_z\}\backslash\{l\}} u_i \otimes v_i + \sum_{i\in\{1,\ldots,k_z\}\backslash\{l\}} (u_l \otimes (\alpha_i \cdot v_i)) \\
&= \sum_{i\in\{1,\ldots,k_z\}\backslash\{l\}} u_i \otimes v_i + u_l \otimes \left(\sum_{i\in\{1,\ldots,k_z\}\backslash\{l\}} \alpha_i \cdot v_i \right) \\
&= \sum_{i\in\{1,\ldots,k_z\}\backslash\{l\}} u_i \otimes v_i + u_l \otimes v_l \\
&= \sum_{i\in\{1,\ldots,k_z\}} u_i \otimes v_i \\
&= z
\end{aligned}$$

Using [corollary: 10.92] we have that $\operatorname{card}(\{1,\ldots,k_z\}\backslash\{l\}) = k_z - 1$ so that there exist a bijection $\beta: \{1,\ldots,k_z-1\} \rightarrow \{1,\ldots,k_z\}\backslash\{l\}$ hence if we define

$$\{u_i')_{i\in\{1,\ldots,k_z-1\}} \subseteq V_1 \text{ by } u_i' = w_{\beta(i)} \text{ and } \{v_i')_{i\in\{1,\ldots,k_z-1\}} \subseteq V_2 \text{ by } v_i' = v_{\beta(i)}$$

then we have

$$\sum_{i\in\{1,\ldots,k_z-1\}} u_i' \otimes v_i' = \sum_{i\in\{1,\ldots,k_z-1\}} w_{\beta(i)} \otimes v_{\beta(i)} \underset{[\text{theorem: } 11.36]}{=} \sum_{i\in\{1,\ldots,k_z\}\backslash\{l\}} w_i \otimes v_i = z$$

so that $k_z - 1 \in K_z$ contradicting $k_z = \min(K_z)$. Hence $\{u_i\}_{i\in\{1,\ldots,k_z\}}$ is not lineard dependent, proving that $\{u_i\}_{i\in\{1,\ldots,k_z\}}$ is linear dependent. □

The following theorems will be important for induction arguments.

Theorem 13.26. *Let $n \in \mathbb{N}$ and $\{V_i\}_{i\in\{1,\ldots,n+1\}}$ be a family of vector spaces over F,*

$$\langle V_1 \otimes_n \ldots \otimes_n V_n, \otimes_n \rangle \text{ a tensor product of } V_1, \ldots, V_n$$

and

$$\langle (V_1 \otimes_n \cdots \otimes_n V_n) \otimes V_{n+1}, \otimes \rangle \text{ a tensor product of } V_1 \otimes_n \cdots \otimes_n V_n, V_{n+1}$$

then there exist a tensor product

$$\langle V_1 \otimes_{n+1} \cdots \otimes_{n+1} V_{n+1}, \otimes_{n+1} \rangle \text{ of } V_1, \ldots, V_{n+1}$$

such that

$$V_1 \otimes_{n+1} \ldots \otimes_{n+1} V_{n+1} = (V_1 \otimes_n \cdots \otimes_n V_n) \otimes V_{n+1}$$

Proof. Let $V_1 \otimes_{n+1} \ldots \otimes_{n+1} V_{n+1}$ be the vector space $(V_1 \otimes_n \cdots \otimes_n V_n) \otimes V_{n+1}$, so we have by definition

$$V_1 \otimes_{n+1} \ldots \otimes_{n+1} V_{n+1} = (V_1 \otimes_n \cdots \otimes_n V_n) \otimes V_{n+1}$$

Define now

$$\otimes_{n+1}: \prod_{i \in \{1, \ldots, n+1\}} V_i \rightarrow (V_1 \otimes_n \cdots \otimes_n V_n) \otimes V_{n+1} \text{ where } \otimes_{n+1} \text{ is defined by} \tag{13.35}$$

$$\otimes_{n+1}(v_1, \ldots, v_{n+1}) = \otimes(\otimes_n(v_1, \ldots, v_n), v_{n+1})$$

Let $i \in \{1, \ldots, n+1\}$ and take $\{v_j\}_{j \in \{1, \ldots, n+1\} \setminus \{i\}}$ $\alpha \in F$, $u, v \in V_i$ then we have for

$$\otimes_{n+1}(v_1, \ldots, v_i, u + \alpha \cdot v, v_{i+1}, \ldots, v_n)$$

either:

$\boldsymbol{i \in \{1, \ldots, n\}}$**.** Then

$$\begin{aligned} \otimes_{n+1}(v_1, \ldots, v_i, u + \alpha \cdot v, v_{i+1}, \ldots, v_n) &= \\ \otimes(\otimes_n(v_1, \ldots, v_i, u + \alpha \cdot v, v_{i+1}, \ldots, v_n), v_{n+1}) &= \\ \otimes(\otimes_n(v_1, \ldots, v_i, u, v_{i+1}, \ldots, v_n) + \alpha \cdot (v_1, \ldots, v_i, v, v_{i+1}, \ldots, v_n), v_{n+1}) &= \\ \otimes(\otimes_n(v_1, \ldots, v_i, u, v_{i+1}, \ldots, v_n), v_{n+1}) + \alpha \cdot \otimes(\otimes_n(v_1, \ldots, v_i, v, v_{i+1}, \ldots, v_n), v_{n+1}) &= \\ \otimes_{n+1}(v_1, \ldots, v_i, u, v_{i+1}, \ldots, v_n) + \alpha \cdot \otimes_{n+1}(v_1, \ldots, v_i, v, v_{i+1}, \ldots, v_n) & \end{aligned}$$

$\boldsymbol{i = n+1}$**.** Then

$$\begin{aligned} \otimes_{n+1}(v_1, \ldots, v_i, u + \alpha \cdot v, v_{i+1}, \ldots, v_n) &= \\ \otimes(\otimes_n(v_1, \ldots, v_n), u + \alpha \cdot v) &= \\ \otimes(\otimes_n(v_1, \ldots, v_n), u) + \alpha \cdot \otimes(\otimes_n(v_1, \ldots, v_n), v) &= \\ \otimes_{n+1}(v_1, \ldots, v_i, u, v_{i+1}, \ldots, v_n) + \alpha \cdot \otimes_{n+1}(v_1, \ldots, v_i, v, v_{i+1}, \ldots, v_n) & \end{aligned}$$

which proves that

$$\otimes_{n+1} \in \mathrm{Hom}(V_1, \ldots, V_{n+1}; (V_1 \otimes_n \cdots \otimes_n V_n) \otimes V_{n+1}) \tag{13.36}$$

If $v \in \otimes\left(\left(\otimes\left(\prod_{i \in \{1, \ldots, n\}} V_i\right)\right) \cdot V_{n+1}\right)$ then there exists a $v_1 \in \otimes\left(\prod_{i \in \{1, \ldots, n\}} V_1\right)$ and $v_2 \in V_{n+1}$ such that $v = \otimes(v_1, v_2)$. As $v_1 \in \otimes\left(\prod_{i \in \{1, \ldots, b\}} V_1\right)$ there exist a $(v_{1,1}, \ldots, v_{1,n}) \in \prod_{i \in \{1, \ldots, n\}} V_1$ such that $v_1 = \otimes_n(v_{1,1}, \ldots, v_{1,n})$. So

$$v = \otimes(v_1, v_2) = \otimes(\otimes_n(v_{1,1}, \ldots, v_{1,n}), v_2) \underset{\text{def}}{=} \otimes_{n+1}(v_{1,1}, \ldots, v_{1,n}, v_2) \in \otimes_{n+1}\left(\prod_{i \in \{1, \ldots, n+1\}} V_i\right)$$

Hence we have that

$$\otimes\left(\left(\otimes_n\left(\prod_{i \in \{1, \ldots, n\}} V_i\right)\right) \cdot V_{n+1}\right) \subseteq \otimes_{n+1}\left(\prod_{i \in \{1, \ldots, n+1\}} V_i\right) \tag{13.37}$$

If $v \in \otimes((V_1 \otimes_n \cdots \otimes_n V_n), V_{n+1})$ then there exists a $v_1 \in V_1 \otimes_n \cdots \otimes_n V_n = \mathrm{span}\left(\otimes_n\left(\prod_{i \in \{1, \ldots, n\}} V_i\right)\right)$ and $v_2 \in V_{n+1}$ such that $v = \otimes(v_1, v_2)$. As $v_1 \in \mathrm{span}\left(\otimes_n\left(\prod_{i \in \{1, \ldots, n\}} V_i\right)\right)$ there exist a finite $I \subseteq \prod_{i \in \{1, \ldots, n\}} V_i$ and a $\{\alpha_u\}_{u \in I} \subseteq F$ such that $v_1 = \sum_{u \in I} \alpha_u \cdot u$. So

$$\begin{aligned} v &= \otimes(v_1, v_2) \\ &= \otimes\left(\sum_{u \in I} \alpha_u \cdot u, v_2\right) \\ &= \sum_{u \in I} \alpha_\alpha \cdot \otimes(u, v_2) \\ &\subseteq \mathrm{span}\left(\otimes_{n+1}\left(\prod_{i \in \{1, \ldots, n+1\}} V_i\right)\right) \text{ [see eq: 13.37]} \end{aligned}$$

proving

$$\otimes((V_1 \otimes_n \cdots \otimes_n V_n), V_{n+1}) \subseteq \operatorname{span}\left(\otimes_{n+1}\left(\prod_{i \in \{1,\ldots,n+1\}} V_i\right)\right)$$

Using [theorems: 11.104, 11.107] we have

$$(V_1 \otimes_n \cdots \otimes_n V_n) \otimes V_{n+1} = \operatorname{span}(\otimes((V_1 \otimes_n \cdots \otimes_n V_n) \cdot V_{n+1})) \subseteq \operatorname{span}\left(\otimes_{n+1}\left(\prod_{i \in \{1,\ldots,n+1\}} V_i\right)\right) \tag{13.38}$$

By the definition of $\otimes_{n+1}$ [see eq: 13.35]

$$\otimes_{n+1}\left(\prod_{i \in \{1,\ldots,n+1\}} V_i\right) \subseteq (V_1 \otimes_n \cdots \otimes_n V_n) \otimes V_{n+1}$$

hence by [theorems: 11.104]

$$\operatorname{span}\left(\otimes_{n+1}\left(\prod_{i \in \{1,\ldots,n+1\}} V_i\right)\right) \subseteq \operatorname{span}((V_1 \otimes_n \cdots \otimes_n V_n) \otimes V_{n+1}) = (V_1 \otimes_n \cdots \otimes_n V_n) \otimes V_{n+1}$$

Combining the above with [eq: 13.38] gives

$$\operatorname{span}\left(\otimes_{n+1}\left(\prod_{i \in \{1,\ldots,n+1\}} V_i\right)\right) = (V_1 \otimes_n \cdots \otimes_n V_n) \otimes V_{n+1} \tag{13.39}$$

Let $\varphi \in \operatorname{Hom}(V_1, \ldots, V_{n+1}; U)$. Given $v \in V_{n+1}$ define

$$\varphi_v: \prod_{i \in \{1,\ldots,n\}} V_i \to U \text{ by } \varphi_v(v_1, \ldots, v_n) = \varphi(v_1, \ldots, v_n, v)$$

then φ_v is multilinear because φ is multilinear. As $\langle V_1 \otimes_n \cdots \otimes_n V_n, \otimes_n \rangle$ is a tensor product there exist a $k_v \in \operatorname{Hom}(V_1 \otimes_n \cdots \otimes_n V_n, U)$ such that

$$k_v \circ \otimes_n = \varphi_v \tag{13.40}$$

Let $u, v \in V_{n+1}$, $\alpha \in F$ and $w \in \otimes\left(\prod_{i \in \{1,\ldots,n\}} V_i\right)$ then there exist a $(v_1, \ldots, v_n)$ such that $w = \otimes_n(v_1, \ldots, v_n)$. Further

$$\begin{aligned} k_{u+\alpha\cdot v}(w) &= k_{u+\alpha\cdot v}(\otimes(v_1, \ldots, v_n)) \\ &= \varphi_{u+\alpha\cdot v}(v_1, \ldots, v_n) \\ &= \varphi(v_1, \ldots, v_n, u + \alpha \cdot v) \\ &= \varphi(v_1, \ldots, v_n, u) + \alpha \cdot \varphi(v_1, \ldots, v_n, v) \\ &= \varphi_u(v_1, \ldots, v_n) + \alpha \cdot \varphi_v(v_1, \ldots, v_n) \\ &= k_u(\otimes_n(v_1, \ldots, v_n)) + \alpha \cdot k_v(\otimes_n(v_1, \ldots, v_n)) \\ &= k_u(w) + \alpha \cdot k_v(w) \end{aligned} \tag{13.41}$$

Let $w \in \operatorname{span}\left(\otimes\left(\prod_{i \in \{1,\ldots,n\}} V_i\right)\right)$ then there exists a finite set $I \subseteq \otimes\left(\prod_{i \in \{1,\ldots,n\}} V_i\right)$ and $\{\alpha_i\}_{i \in I} \subseteq F$ such that

$$w = \sum_{z \in I} \alpha_z \cdot z$$

Then

$$\begin{aligned} k_{u+\alpha\cdot v}(w) &= k_{u+\alpha\cdot v}\left(\sum_{z \in I} \alpha_z \cdot z\right) \\ &= \sum_{z \in I} \alpha_z \cdot k_{u+\alpha\cdot v}(z) \\ &= \sum_{z \in I} \alpha_z \cdot (k_u(z) + \alpha \cdot k_v(z)) \\ &= \sum_{z \in I} \alpha_z \cdot k_u(z) + \alpha \cdot \sum_{z \in I} \alpha_z \cdot k_v(z) \\ &= k_u(w) + \alpha \cdot k_v(w) \end{aligned}$$

proving that

$$\forall u, v \in V_{n+1}, \forall \alpha \in F \text{ we have } k_{u+\alpha \cdot v} = k_u + \alpha \cdot k_v \tag{13.42}$$

Define

$$\psi: (V_1 \otimes_n \cdots \otimes_n V_n) \cdot V_{n+1} \rightarrow U \text{ by } \psi(u, v) = k_v(u) \tag{13.43}$$

then we have for $x, y \in V_1 \otimes_n \cdots \otimes_n V_n$, $\alpha \in F$ and $v \in V_{n+1}$

$$\psi(x + \alpha \cdot y, v) = k_v(x + \alpha \cdot y) = k_v(x) + \alpha \cdot k_v(y) = \psi(x) + \alpha \cdot \psi(y)$$

and for $v \in V_1 \otimes_n \cdots \otimes_n V_n$, $x, y \in V_{n+1}$, $\alpha \in F$ that

$$\psi(v, x + \alpha \cdot y) = k_{x+\alpha \cdot y}(v) \underset{[\text{eq: } 13.42]}{=} k_x(v) + \alpha \cdot k_v(v) = \psi(v, x) + \alpha \cdot \psi(v, y)$$

From the above it follows that

$$\psi \in \operatorname{Hom}(V_1 \otimes_n \cdots \otimes_n V_n, V_{n+1}; U) \tag{13.44}$$

As $\langle (V_1 \otimes_n \cdots \otimes_n V_n) \otimes V_{n+1}, \otimes \rangle$ is a tensor product there exist a $h \in \operatorname{Hom}((V_1 \otimes_n \cdots \otimes_n V_n) \otimes V_{n+1}, U)$ such that

$$h \circ \otimes = \psi \tag{13.45}$$

Let $(v_1, \ldots, v_{n+1}) \in \prod_{i \in \{1, \ldots, n+1\}} V_i$ then we have

$$\begin{array}{rcl} (h \circ \otimes_{n+1})(v_1, \ldots, v_{n+1}) & = & h(\otimes_{n+1}(v_1, \ldots, v_{n+1})) \\ & = & h(\otimes(\otimes_n(v_1, \ldots, v_n), v_{n+1})) \\ & \underset{[\text{eq: } 13.45]}{=} & \psi(\otimes_n(v_1, \ldots, v_n), v_{n+1}) \\ & \underset{[\text{eq: } 13.43]}{=} & k_{v_{n+1}}(\otimes_n(v_1, \ldots, v_n)) \\ & \underset{[\text{eq: } 13.40]}{=} & \varphi_{v_{n+1}}(v_1, \ldots, v_n) \\ & = & \varphi(v_1, \ldots, v_n, v_{n+1}) \end{array}$$

which proves that

$$h \circ \otimes_{n+1} = \varphi \tag{13.46}$$

Finally from [eqs: 13.39, 13.36 and 13.46] it follows that

$$\langle (V_1 \otimes_n \cdots \otimes_n V_n) \otimes V_{n+1}, \otimes_{n+1} \rangle \text{ is a tensor product of } V_1, \ldots, V_{n+1}$$ □

Theorem 13.27. *Let $n \in \mathbb{N}$ and $\{V_i\}_{i \in \{1, \ldots, n+1\}}$ be a family of vector spaces over F. If*

$$\langle V_1 \otimes_{n+1} \cdots \otimes_{n+1} V_{n+1}, \otimes_{n+1} \rangle \textit{ and } \langle V_1 \otimes_n \ldots \otimes_n V_n, \otimes_n \rangle \textit{ are tensor products}$$

then there exist a tensor product

$$\langle (V_1 \otimes_n \ldots \otimes_n V_n) \otimes V_{n+1}, \otimes \rangle \textit{ of } V_1 \otimes_n \ldots \otimes_n V_n \textit{ and } V_{n+1}$$

such that

$$V_1 \otimes_{n+1} \ldots \otimes_{n+1} V_{n+1} = (V_1 \otimes_n \ldots \otimes_n V_n) \otimes V_{n+1}$$

Proof. Take the family $\{W_i\}_{i \in \{1,2\}}$ by $W_1 = V_1 \otimes_n \cdots \otimes_n V_n$ and $W_2 = V_{n+1}$ of vector spaces. Using [theorem: 13.13] there exist a tensor product $\langle P, \nu \rangle$ of $V_1 \otimes_n \cdots \otimes_n V_n$ and V_{n+1}, so

$$\nu \in \operatorname{Hom}\left(\prod_{i \in \{1,.2\}} W_i; P \right) \tag{13.47}$$

$$P = \operatorname{span}\left(\nu\left(\prod_{i \in \{1,2\}} W_i \right) \right) \tag{13.48}$$

$$\forall \varphi \in \operatorname{Hom}\left(\prod_{i \in \{1,.2\}} W_i; U \right) \text{ there exist a } h \in \operatorname{Hom}(P, U) \text{ such that } \varphi = h \circ \nu \tag{13.49}$$

Let $w \in V_{n+1}$ and define

$$\beta_w: \prod_{i \in \{1,\ldots,n\}} V_i \rightarrow V_1 \otimes_{n+1} \cdots \otimes_{n+1} V_{n+1} \text{ by } \beta_w(v_1,\ldots,v_n) = \otimes_{n+1}(v_1,\ldots,v_n,w)$$

then for $i \in \{1,\ldots,n\}$ and $\alpha \in F$, $u,v \in V_i$ we have

$$\begin{array}{rl} \beta_w(v_1,\ldots,v_{i-1},u+\alpha\cdot v,v_{i+1},v_{i+1},\ldots,v_n) & = \\ \otimes_{n+1}(v_1,\ldots,v_{i-1},u+\alpha\cdot v,v_{i+1},\ldots,v_n,w) & = \\ \otimes_{n+1}(v_1,\ldots,v_{i-1},u,v_{i+1},\ldots,v_n,w)+\alpha\cdot\otimes_{n+1}(v_1,\ldots,v_{i-1},v,v_{i+1},\ldots,v_n,w) & = \\ \beta_w(v_1,\ldots,v_{i-1},u,v_{i+1},v_{i+1},\ldots,v_n)+\alpha\cdot\beta_w(v_1,\ldots,v_{i-1},v,v_{i+1},v_{i+1},\ldots,v_n) & \end{array}$$

proving that

$$\forall w \in V_{n+1} \text{ we have } \beta_w \in \mathrm{Hom}(V_1,\ldots,V_n;V_1 \otimes_{n+1} \cdots \otimes_{n+1} V_{n+1}) \tag{13.50}$$

Further for $\alpha \in F$ and $u,v \in V_{n+1}$ we have

$$\begin{array}{rcl} \beta_{u+\alpha\cdot v}(v_1,\ldots,v_n) & = & \otimes_{n+1}(v_1,\ldots,v_n,\mu+\alpha\cdot v) \\ & = & \otimes_{n+1}(v_1,\ldots,v_n,\mu)+\alpha\cdot\otimes_{n+1}(v_1,\ldots,v_n,v) \\ & = & \beta_u(v_1,\ldots,v_n)+\alpha\cdot\beta_v(v_1,\ldots,v_n) \end{array}$$

proving that

$$\beta_{u+\alpha\cdot v} = \beta_u + \alpha\cdot\beta_v \tag{13.51}$$

As $V_1 \otimes_n \cdots \otimes_n V_n$ is a tensor product of $\{V_i\}_{i\in\{1,\ldots,n\}}$ we have from [eq: 13.50] that $\forall w \in V_{n+1}$ there exist a

$$g_w \in \mathrm{Hom}(V_1 \otimes_n \cdots \otimes_n V_n, V_1 \otimes_{n+1} \cdots \otimes_{n+1} V_{n+1}) \text{ such that } \beta_w = g_w \circ \otimes_n \tag{13.52}$$

Hence $\forall (v_1,\ldots,v_n) \in \prod_{i\in\{1,\ldots,n\}} V_i$, $\forall w \in V_{n+1}$ we have

$$g_w(\otimes_n(v_1,\ldots,v_n)) = \beta_w(v_1,\ldots,v_n) \underset{[\text{eq: } 13.50]}{=} \otimes_{n+1}(v_1,\ldots,v_n,w) \tag{13.53}$$

Define now

$$\varphi: \prod_{i\in\{1,\ldots,n+1\}} V_i \rightarrow P \text{ by } \varphi(v_1,\ldots,v_{n+1}) = \nu(\otimes_n(v_1,\ldots,v_n),v_{n+1}) \tag{13.54}$$

then we have for $i \in \{1,\ldots,n+1\}$, $u,v \in V_i$ and $\alpha \in F$ either:

$\boldsymbol{i \in \{1,\ldots,n\}}$.

$$\begin{array}{rl} \varphi(v_1,\ldots,v_{i-1},u+\alpha\cdot v,v_{i+1},\ldots,v_n) & = \\ \nu(\otimes_n(v_1,\ldots,v_{i-1},u+\alpha\cdot v,v_{i+1},\ldots,v_n),v_{n+1}) & = \\ \nu(\otimes_n(v_1,\ldots,v_{i-1},u,v_{i+1},\ldots,v_n)+\alpha\cdot\otimes_n(v_1,\ldots,v_{i-1},v,v_{i+1},\ldots,v_n),v_{n+1}) & = \\ \nu(\otimes_n(v_1,\ldots,v_{i-1},u,v_{i+1},\ldots,v_n),v_{n+1})+\alpha\cdot\nu(\otimes_n(v_1,\ldots,v_{i-1},v,v_{i+1},\ldots,v_n),v_{n+1}) & = \\ \varphi(v_1,\ldots,v_{i-1},u,v_{i+1},\ldots,v_n)+\alpha\cdot\varphi(v_1,\ldots,v_{i-1},v,v_{i+1},\ldots,v_n) & \end{array}$$

$\boldsymbol{i = n+1}$.

$$\begin{array}{rl} \varphi(v_1,\ldots,v_{i-1},u+\alpha\cdot v,v_{i+1},\ldots,v_n) & = \\ \nu(\otimes_n(v_1,\ldots,v_n),u+\alpha\cdot v) & = \\ \nu(\otimes_n(v_1,\ldots,v_n),u)+\alpha\cdot\nu(\otimes_n(v_1,\ldots,v_n),v) & = \\ \varphi(v_1,\ldots,v_{i-1},u,v_{i+1},\ldots,v_n)+\alpha\cdot\varphi(v_1,\ldots,v_{i-1},v,v_{i+1},\ldots,v_n) & \end{array}$$

proving that

$$\varphi \in \mathrm{Hom}(V_1,\ldots,V_{n+1},P) \tag{13.55}$$

As $V_1 \otimes_{n+1} \ldots \otimes_{n+1} V_{n+1}$ is a tensor product of $V_1,\ldots,V_{n+1}$ there exist a $h \in \mathrm{Hom}(V_1 \otimes_{n+1} \ldots \otimes_{n+1} V_{n+1}, P)$ such that $\varphi = h \circ \otimes_{n+1}$ or $\forall (v_1,\ldots,v_{n+1}) \in \prod_{i\in\{1,\ldots,n+1\}} V_i$ we have

$$h(\otimes_{n+1}(v_1,\ldots,v_{n+1})) = \nu(\otimes_n(v_1,\ldots,v_n),v_{n+1}) \tag{13.56}$$

Define now

$$\psi: \prod_{i \in \{1,2\}} W_i \rightarrow V_1 \otimes_{n+1} \cdots \otimes_{n+1} V_{n+1} \text{ by } \psi(z,v) \underset{[\text{eq: } 13.52]}{=} g_v(z) \tag{13.57}$$

Then if $i \in \{1,2\}$ we have for $\alpha \in F$ and $u, v \in W_i$ and either:

***i* = 1.** Then, as $W_i = V_1 \otimes_n \cdots \otimes_n V_n$, we have $u, v \in V_1 \otimes_n \cdots \otimes_n V_n$, so that given $v_2 \in V_{n+1}$

$$\begin{array}{rcl} \psi(u + \alpha \cdot v, v_2) & = & g_{v_2}(u + \alpha \cdot v) \\ & \underset{[\text{eq: } 13.52]}{=} & g_{v_2}(u) + \alpha \cdot g_{v_2}(v) \\ & = & \psi(u, v_2) + \alpha \cdot \psi(v, v_2) \end{array}$$

***i* = 2.** Then $W_2 = V_{n+1}$ and $u, v \in V_{n+1}$, let $v_1 \in V_1 \otimes_n \cdots \otimes_n V_n$. As $V_1 \otimes_n \cdots \otimes_n V_n = \text{span}\left(\otimes_n\left(\prod_{i \in \{1,\ldots,n\}} V_i\right)\right)$ there exists by [theorem: 11.102] a $\{u_i\}_{i \in \{1,\ldots,l\}} \subseteq \otimes_n\left(\prod_{i \in \{1,\ldots,n\}} V_i\right)$ and a $\{\alpha_i\}_{i \in \{1,\ldots,l\}} \subseteq F$ such that $v_1 = \sum_{i \in \{1,\ldots,l\}} \alpha_i \cdot u_i$. If $i \in \{1, \ldots, l\}$ then for u_i there exists a $x \in \prod_{i \in \{1,\ldots,n\}} V_i$ such that $u_i = \otimes_n(x)$, hence

$$\begin{array}{rcl} g_{u+\alpha \cdot v}(u_i) & = & g_{u+\alpha \cdot v}(\otimes_n(x)) \\ & \underset{[\text{eq: } 13.52]}{=} & \beta_{u+\alpha \cdot v}(x) \\ & \underset{[\text{eq: } 13.51]}{=} & \beta_u(x) + \alpha \cdot \beta_v(x) \\ & \underset{[\text{eq: } 13.52]}{=} & g_u(\otimes_n(x)) + \alpha \cdot g_v(\otimes_n(x)) \\ & = & g_u(u_i) + \alpha \cdot g_v(u_i) \end{array}$$

So that

$$\begin{array}{rcl} \psi(v_1, u + \alpha \cdot v) & = & g_{u+\alpha \cdot v}(v_1) \\ & = & g_{u+\alpha \cdot v}\left(\sum_{i \in \{1,\ldots,l\}} \alpha_i \cdot u_i\right) \\ & \underset{[\text{eq: } 13.52]}{=} & \sum_{i \in \{1,\ldots,l\}} \alpha_i \cdot g_{u+\alpha \cdot v}(u_i) \\ & = & \sum_{i \in \{1,\ldots,l\}} \alpha_i \cdot (g_u(u_i) + \alpha \cdot g_v(u_i)) \\ & = & \sum_{i \in \{1,\ldots,l\}} \alpha_i \cdot g_u(u_i) + \alpha \cdot \sum_{i \in \{1,\ldots,l\}} \alpha_i \cdot g_v(u_i) \\ & \underset{[\text{eq: } 13.52]}{=} & g_u\left(\sum_{i \in \{1,\ldots,l\}} \alpha_i \cdot u_i\right) + \alpha \cdot g_v\left(\sum_{i \in \{1,\ldots,l\}} \alpha_i \cdot u_i\right) \\ & = & g_u(v_1) + \alpha \cdot g_v(v_1) \\ & = & \psi(v_1, u) + \alpha \cdot \psi(v_1, v) \end{array}$$

proving that ψ is multilinear or

$$\psi \in \text{Hom}(W_1, W_2; V_1 \otimes_{n+1} \cdots \otimes_{n+1} V_{n+1}) = \text{Hom}(V_1 \otimes_n \cdots \otimes_n V_n, V_{n+1}; V_1 \otimes_{n+1} \cdots \otimes_{n+1} V_{n+1}) \tag{13.58}$$

By [eq: 13.49] we have that there exist a

$$k \in \text{Hom}(P, V_1 \otimes_{n+1} \cdots \otimes_{n+1} V_{n+1}) \text{ such that } \psi = k \circ \nu \tag{13.59}$$

Let $(v_1, \ldots, v_n) \in \prod_{i \in \{1,\ldots,n\}} V_i$ and $v \in V_{n+1}$ then as $\otimes_n \in \text{Hom}(V_1, \ldots, V_n; V_1 \otimes_n \cdots \otimes_n V_n)$ we have that $\otimes_n(v_1, \ldots, v_n) \in V_1 \otimes_n \cdots \otimes_n V_n$ so that $\nu(\otimes_n(v_1, \ldots, v_n), v) \in P$ is in the domain of k and we have

$$\begin{array}{rcl} k(\nu(\otimes_n(v_1, \ldots, v_n), v)) & \underset{[\text{eq: } 13.59]}{=} & \psi(\otimes_n(v_1, \ldots, v_n), v) \\ & \underset{[\text{eq: } 13.57]}{=} & g_v(\otimes_n(v_1, \ldots, v_n)) \\ & \underset{[\text{eq: } 13.53]}{=} & \otimes_{n+1}(v_1, \ldots, v_n, v) \end{array} \tag{13.60}$$

As $h\in\operatorname{Hom}(V_1\otimes_{n+1}\ldots\otimes_{n+1}V_{n+1},P)$ and $k\in\operatorname{Hom}(P,V_1\otimes_{n+1}\cdots\otimes_{n+1}V_{n+1})$ it follows that

$$k\circ h\in\operatorname{Hom}(V_1\otimes_{n+1}\cdots\otimes_{n+1}V_{n+1},V_1\otimes_{n+1}\cdots\otimes_{n+1}V_{n+1})$$

If $z\in\otimes_{n+1}\left(\prod_{i\in\{1,\ldots,n+1\}}V_i\right)$ then there exist a $(v_1,\ldots,v_{n+1})$ such that $z=\otimes_{n+1}(v_1,\ldots,v_{n+1})$ and

$$\begin{array}{rcl}(k\circ h)(z) & = & (k\circ h)(\otimes_{n+1}(v_1,\ldots,v_{n+1}))\\ & = & k(h(\otimes_{n+1}(v_1,\ldots,v_{n+1})))\\ & \underset{[\text{eq: }13.56]}{=} & k(\nu(\otimes_n(v_1,\ldots,v_n),v_{n+1}))\\ & \underset{[\text{eq: }13.60]}{=} & \otimes_{n+1}(v_1,\ldots,v_{n+1})\\ & = & z\end{array} \tag{13.61}$$

Next if $z\in V_1\otimes_{n+1}\cdots\otimes_{n+1}V_{n+1}=\operatorname{span}\left(\otimes_{n+1}\left(\prod_{i\in\{1,\ldots,n+1\}}V_i\right)\right)$ there exist a finite $I\subseteq\otimes_{n+1}\left(\prod_{i\in\{1,\ldots,n+1\}}V_i\right)$ and a $\{\alpha_i\}_{i\in I}\subseteq F$ such that $z=\sum_{i\in I}\alpha_i\cdot i$. Hence

$$(k\circ h)(z)=(k\circ h)\left(\sum_{i\in I}\alpha_i\cdot i\right)=\sum_{i\in I}\alpha_i\cdot(k\circ h)(i)\underset{[\text{eq: }13.61]}{=}\sum_{i\in I}\alpha_i\cdot i=z$$

proving that

$$k\circ h=\operatorname{Id}_{V_1\otimes\cdots\otimes V_{n+1}} \tag{13.62}$$

Again as $h\in\operatorname{Hom}(V_1\otimes_{n+1}\ldots\otimes_{n+1}V_{n+1},P)$ and $k\in\operatorname{Hom}(P,V_1\otimes_{n+1}\cdots\otimes_{n+1}V_{n+1})$ it follows that

$$h\circ k\in\operatorname{Hom}(P,P) \tag{13.63}$$

Let $v=(v_1,\ldots,v_n)\in\prod_{i\in\{1,\ldots,n+1\}}v_i$, $w\in V_{n+1}$ then we have

$$\begin{array}{rcl}h(k(\nu(\otimes_n(v),w))) & = & h(k(\nu(\otimes_n(v_1,\ldots,v_n),w)))\\ & \underset{[\text{eq: }13.60]}{=} & h(\otimes_{n+1}(v_1,\ldots,v_n,w))\\ & \underset{[\text{eq: }13.56]}{=} & \nu(\otimes_n(v_1,\ldots,v_n),w)\\ & = & \nu(\otimes_n(v),w)\end{array} \tag{13.64}$$

Let $z\in V_1\otimes_n\cdots\otimes_n V_n=\operatorname{span}\left(\otimes_n\left(\prod_{i\in\{1,\ldots,n\}}V_i\right)\right)$ then by [theorem: 11.102] there exist a $\{z_i\}_{i\in\{1,\ldots,l\}}\subseteq\prod_{i\in\{1,\ldots,n\}}V_i$ and a $\{\alpha_i\}_{i\in\{1,\ldots,l\}}\subseteq F$ such that $z=\sum_{i\in\{1,\ldots,l\}}\alpha_i\cdot z_i$. Using [theorem: 3.130] there exist a $\{y_i\}_{i\in\{1,\ldots,l\}}\subseteq\prod_{i\in\{1,\ldots,n\}}V_i$ such that $\forall i\in\{1,\ldots,l\}$ $\otimes_n(y_i)=z_i$. So given $w\in V_{n+1}$ we have

$$\begin{array}{rcl}h(k(\nu(z,w))) & = & h\left(k\left(\nu\left(\sum_{i\in\{1,\ldots,l\}}\alpha_i\cdot z_i,w\right)\right)\right)\\ & \underset{\nu\text{ is multilinear}}{=} & h\left(k\left(\sum_{i\in\{1,\ldots,l\}}\alpha_i\cdot\nu(z_i,w)\right)\right)\\ & = & \sum_{i\in\{1,\ldots,l\}}\alpha_i\cdot h(k(\nu(\otimes_n(y_i),w)))\\ & \underset{[\text{eq: }13.64]}{=} & \sum_{i\in\{1,\ldots,l\}}\alpha_i\cdot\nu(\otimes_n(y_i),w)\\ & = & \sum_{i\in\{1,\ldots,l\}}\alpha_i\cdot\nu(z_i,w)\\ & \underset{\nu\text{ is multilinear}}{=} & \nu\left(\sum_{i\in\{1,\ldots,l\}}\alpha_i\cdot z_i,w\right)\\ & = & \nu(z,w)\end{array}$$

proving that

$$\forall z\in V_1\otimes_n\cdots\otimes_n V_n\text{ and }w\in V_{n+1}\text{ that }h(k(\nu(z,w)))=\nu(z,w) \tag{13.65}$$

Let $z\in\prod_{i\in\{1,2\}} W_i$ then $z=(z_1,z_2)$ with $z_1\in W_1=V_1\otimes_n\cdots\otimes_n V_n$ and $z_2\in V_{n+1}$ then

$$h(k(\nu(z)))=h(k(\nu(z_1,z_2)))\underset{\text{[eq: 13.65]}}{=}\nu(z,w)=\nu(z)\tag{13.66}$$

If $z\in P=\operatorname{span}\left(\nu\left(\prod_{i\in\{1,.2\}} W_i\right)\right)$ then by [theorem: 11.102] there exists a $\{z_i\}_{i\in\{1,\ldots,l\}}\subseteq\nu\left(\prod_{i\in\{1,2\}} W\right)_i$ and a $\{\alpha_i\}_{i\in\{1,\ldots,l\}}\subseteq F$ such that $z=\sum_{i\in[1,\ldots,l]}\alpha_i\cdot z_i$. Using [theorem: 3.130] there exist a $\{y_i\}_{i\in\{1,\ldots,l\}}\subseteq\prod_{i\in\{1,.2\}} W_i$ such that $\forall i\in\{1,\ldots,l\}$ we have $z_i=\nu(y_i)$.

$$\begin{array}{rcl}(h\circ k)(z) & = & h(k(z))\\ & = & h\left(k\left(\sum_{i\in\{1,\ldots,l\}}\alpha_i\cdot z_i\right)\right)\\ & = & \sum_{i\in\{1,\ldots,l\}}\alpha_i\cdot h(k(z_i))\\ & = & \sum_{i\in\{1,\ldots,l\}}\alpha_i\cdot h(k(\nu(y_i)))\\ & \underset{\text{[eq: 13.66]}}{=} & \sum_{i\in\{1,\ldots,l\}}\alpha_i\cdot\nu(y_i)\\ & = & \sum_{i\in\{1,\ldots,l\}}\alpha_i\cdot z_i\\ & = & z\end{array}$$

proving that

$$h\circ k=\mathrm{Id}_P\tag{13.67}$$

Now we are ready to consruct the tensor product, starting with the multilinear mapping. Define [see eq: 13.47, 13.59]

$$\otimes:\prod_{i\in\{1,2\}} W_i\rightarrow V_1\otimes_{n+1}\cdots\otimes_{n+1}V_{n+1}\text{ by }\otimes(z,v)=k(\nu(z,v))\tag{13.68}$$

If $i\in\{1,2\}$ we have for $\alpha\in F$ and $u,v\in W_i$ either:

$i=1$. Then as $W_1=V_1\otimes_n\cdots\otimes_n V_n$ we have $u,v\in V_1\otimes_n\cdots\otimes_n V_n$ and, as ν is multilinear and k linear, we have for $v_2\in V_{n+1}$

$$\begin{array}{rcl}\otimes(u+\alpha\cdot v,v_2) & = & k(\nu(u+\alpha\cdot v,v_2))\\ & \underset{\nu\text{ is multilinear}}{=} & k(\nu(u,v_2)+\alpha\cdot\nu(v,v_2))\\ & \underset{k\text{ is linear}}{=} & k(\nu(u,v_2))+\alpha\cdot k(\nu(v,v_2))\\ & = & \otimes(u,v_2)+\alpha\cdot\otimes(v,v_2)\end{array}$$

$i=2$. Then as $W_i=V_{n+1}$ we have $u,v\in V_{n+1}$ and, as ν is multilinear and k linear, that for $v_1\in V_1\otimes_n\cdots\otimes_n V_n$

$$\begin{array}{rcl}\otimes(v_1,u+\alpha\cdot v) & = & k(\nu(v_1,u+\alpha\cdot v))\\ & \underset{\nu\text{ is multilinear}}{=} & k(\nu(v_1,u)+\alpha\cdot\nu(v_1,v))\\ & = & k(\nu(v_1,u))+\alpha\cdot\nu(v_1,v)\\ & = & \otimes(v_1,u)+\alpha\cdot\otimes(v_1,v)\end{array}$$

proving that

$$\otimes\in\mathrm{Hom}(W_1,W_2;V_1\otimes_{n+1}\cdots\otimes_{n+1}V_{n+1})=\mathrm{Hom}(V_1\otimes_n\cdots\otimes_n V_n,V_{n+1};V_1\otimes_{n+1}\cdots\otimes_{n+1}V_{n+1})\tag{13.69}$$

By the definition of $\otimes$ [see: 13.68] we have $\otimes\left(\prod_{i\in\{1,2\}} W_i\right)\subseteq V_1\otimes_{n+1}\cdots\otimes_{n+1}V_{n+1}$ proving by [theorems: 11.104, 11.106] that

$$\operatorname{span}\left(\otimes\left(\prod_{i\in\{1,2\}} W_i\right)\right)\subseteq V_1\otimes_{n+1}\cdots\otimes_{n+1}V_{n+1}\tag{13.70}$$

If $v\in V_1\otimes_{n+1}\cdots\otimes_{n+1}V_{n+1}=\operatorname{span}\left(\otimes_{n+1}\left(\prod_{i\in\{1,\ldots,n+1\}} V_i\right)\right)$ then there exists $\{v_i\}_{i\in\{1,\ldots,l\}}\subseteq\otimes_{n+1}\left(\prod_{i\in\{1,\ldots,n+1\}} V_i\right)$ and $\{\alpha_i\}_{i\in\{1,\ldots,l\}}\subseteq F$ such that

$$v=\sum_{i\in\{1,\ldots,l\}}\alpha_i\cdot v_i$$

If $i\in\{1,\ldots,l\}$ then as $v_i\in\otimes_{n+1}\left(\prod_{i\in\{1,\ldots,n+1\}} V_i\right)$ there exists a $(u_1,\ldots,u_{n+1})\in\prod_{i\in\{1,\ldots,n+1\}} V_i$ such that

$$\begin{array}{rcl} v_i & = & \otimes_{n+1}(u_1,\ldots,u_{n+1}) \\ & \underset{[\text{eq: }13.62]}{=} & k(h(\otimes_{n+1}(u_1,\ldots,u_{n+1}))) \\ & \underset{[\text{eq: }13.56]}{=} & k(\nu(\otimes_n(u_1,\ldots,u_n),u_{n+1})) \\ & \underset{[\text{eq: }13.68]}{=} & \otimes(\otimes_n(u_1,\ldots,u_n),u_{n+1}) \\ & \in & \otimes\left(\prod_{i\in\{1,.2\}} W_i\right) \end{array}$$

which proves that $v\in\operatorname{span}\left(\otimes\left(\prod_{i\in\{1,2\}} W_i\right)\right)$. Hence $V_1\otimes_{n+1}\cdots\otimes_{n+1}V_{n+1}\subseteq\operatorname{span}\left(\otimes\left(\prod_{i\in\{1,2\}} W_i\right)\right)$ and combining this with [eq: 13.70] gives

$$V_1\otimes_{n+1}\cdots\otimes_{n+1}V_{n+1}=\operatorname{span}\left(\otimes\left(\prod_{i\in\{1,2\}} W_i\right)\right) \tag{13.71}$$

Let $\theta\in\operatorname{Hom}(W_1,W_2;U)=\operatorname{Hom}(V_1\otimes_n\cdots\otimes_n V_n,V_{n+1};U)$ then by [eq: 13.49] there exists a

$$f\in\operatorname{Hom}(P,U)\text{ such that }\theta=f\circ\nu \tag{13.72}$$

as $f\in\operatorname{Hom}(P,U)$ and $h\in\operatorname{Hom}(V_1\otimes_{n+1}\ldots\otimes_{n+1}V_{n+1},P)$ we have that

$$g=f\circ h\in\operatorname{Hom}(V_1\otimes_{n+1}\cdots\otimes_{n+1}V_{n+1},U) \tag{13.73}$$

Further if $z=(z_1,z_2)\in\prod_{i\in\{1,.2\}} W_i$ then we have

$$\begin{array}{rcl} (g\circ\otimes)(z) & = & g(\otimes(z_1,z_2)) \\ & \underset{[\text{eq: }13.68]}{=} & g(k(\nu(z_1,z_2))) \\ & \underset{[\text{eq: }13.73]}{=} & f(h(k(\nu(z_1,z_2)))) \\ & \underset{[\text{eq: }13.67]}{=} & f(\nu(z_1,z_2)) \\ & \underset{[\text{eq: }13.72]}{=} & \theta(z_1,z_2) \\ & = & \theta(z) \end{array}$$

proving that

$$\forall\theta\in\operatorname{Hom}(W_1,W_2;U)\text{ we have }\exists g\in\operatorname{Hom}(V_1\otimes_{n+1}\cdots\otimes_{n+1}V_{n+1},U)\text{ such that }\theta=g\circ\otimes \tag{13.74}$$

To conclude using [eqs: 13.68, 13.69, 13.71, 13.74] it follows that

$$\langle V_1\otimes_{n+1}\cdots\otimes_{n+1}V_{n+1},\otimes\rangle\text{ is a tensor product of }W_1,W_2$$

or using [convention: 13.19]

$$(V_1\otimes_n\ldots\otimes_n V_n)\otimes V_{n+1}\text{ is a tensor product and }V_1\otimes_{n+1}\ldots\otimes_{n+1}V_{n+1}=(V_1\otimes_n\ldots\otimes_n V_n)\otimes V_{n+1}$$ □

Corollary 13.28. *Let $n\in\mathbb{N}$, $\{V_i\}_{i\in\{1,\ldots,n+1\}}$ be family of vector spaces over a field F and $\langle V_1\otimes_{n+1},\ldots,\otimes_{n+1}V_{n+1},\otimes_{n+1}\rangle$ a tensor product of $V_1,\ldots,V_{n+1}$ then there exist a tensor product of $\langle V_1\otimes_n\ldots\otimes_n V_n,\otimes_n\rangle$ of $V_1,\ldots,V_n$ and a tensor product $\langle(V_1\otimes_n\ldots\otimes_n V_n)\otimes V_{n+1},\otimes\rangle$ of $V_1\otimes_n\ldots\otimes_n V_n,V_{n+1}$ such that*

$$V_1\otimes_{n+1}\cdots\otimes_{n+1}V_{n+1}=(V_1\otimes_n\cdots\otimes_n V_n)\otimes V_{n+1}$$

Proof. By [theorem: 13.20] there exist a tensor product $V_1\otimes\cdots\otimes V_n$ of $V_1,\ldots,V_n$, hence using the previous theorem [theorem: 13.27] there exist also a tensor product $\langle(V_1\otimes_n\cdots\otimes_n V_n)\otimes V_{n+1},\otimes\rangle$ of $V_1\otimes_n\ldots\otimes_n V_n.V_{n+1}$ such that

$$V_1\otimes_{n+1}\cdots\otimes_{n+1}V_{n+1}=(V_1\otimes_n\cdots\otimes_n V_n)\otimes V_{n+1}$$ □

Lemma 13.29. *Let $n\in\mathbb{N}$, $\{X_i\}_{i\in\{1,\ldots,n\}}$ a finite family of finite sets then $\prod_{i\in\{1,\ldots,n\}} X_i$ is finite and $\operatorname{card}\left(\prod_{i\in\{1,\ldots,n\}} X_i\right)=\prod_{i\in\{1,\ldots,n\}}\operatorname{card}(X_i)$*

Proof. We use induction to proof this, so let

$$S=\left\{n\in\mathbb{N}|\text{If }\{X_i\}_{i\in\{1,\ldots,n\}}\text{ is a family of finite sets then }\prod_{i\in\{1,\ldots,n\}}X_i\text{ is finite and card}\left(\prod_{i\in\{1,\ldots,n\}}X_i\right)=\prod_{i\in\{1,\ldots,n\}}\text{card}(X_i)\right\}$$

then we have:

$1\in S$. Let $\{X_i\}_{i\in\{1\}}$ is such that X_1 is finite. By [example: 2.137] we have that $\prod_{i\in\{1\}}X_i$ is bijective with X_1, hence $\prod_{i\in\{1\}}X_i$ is finite and $\text{card}\left(\prod_{i\in\{1\}}X_i\right)=\text{card}(X_1)=\prod_{i\in\{1\}}\text{card}(X_i)$. so $1\in S$

$n\in S\Rightarrow n+1\in S$. Let $\{X_i\}_{i\in\{1,\ldots,n+1\}}$ be family of finite sets. As $n\in S$ we have that

$$\prod_{i\in\{1,\ldots,n\}}X_i\text{ is finite and card}\left(\prod_{i\in\{1,\ldots,n\}}X_i\right)=\prod_{i\in\{1,\ldots,n\}}\text{card}(X_i)\tag{13.75}$$

Using [theorem: 10.90] and the fact that X_{n+1} is finite we have that $\left(\prod_{i\in\{1,\ldots,n\}}X_i\right)\times X_{n+1}$ is finite and $\text{card}\left(\left(\prod_{i\in\{1,\ldots,n\}}X_i\right)\times X_{n+1}\right)=\text{card}\left(\prod_{i\in\{1,\ldots,n\}}X_i\right)\cdot\text{card}(X_{n+1})\underset{\text{[eq: 13.75]}}{=}\prod_{i\in\{1,\ldots,n+1\}}\text{card}(X_i)$

By [theorem: 6.80] $\prod_{i\in\{1,\ldots,n+1\}}X$ is bijective with $\left(\prod_{i\in\{1,\ldots,n\}}X_i\right)\times X_{n+1}$ so that $\left(\prod_{i\in\{1,\ldots,n\}}X_i\right)\times X_{n+1}$ is finite and $\text{card}\left(\left(\prod_{i\in\{1,\ldots,n\}}X_i\right)\times X_{n+1}\right)=\prod_{i\in\{1,\ldots,n+1\}}\text{card}(X_i)$ proving that $n+1\in S$. □

We introduce now the concept of a multi-index

Definition 13.30. *Let $k\in\mathbb{N}$, $\{n_i\}_{i\in\{1,\ldots,k\}}\subseteq\mathbb{N}$ be finite family of natural numbers then set of multi-indexes named $\Gamma(n_1,\ldots,n_k)$ is defined by*

$$\Gamma(n_1,\ldots,n_k)=\prod_{i\in\{1,\ldots,k\}}\{1,\ldots,n_i\}$$

Using the previous lemma [lemma: 13.29] we have that that

$$\text{card}(\Gamma(n_1,\ldots,n_k))=\prod_{i\in\{1,\ldots,k\}}n_i$$

Example 13.31. $\Gamma(2,3)=\{(1,1),(1,2),(1,3),(2,1),(2,2),(2,3)\}$

Theorem 13.32. *Let $k\in\mathbb{N}$, $\{V_i\}_{i\in\{1,\ldots,k\}}$ be a finite family of vector space over a field F, V a vector space over the same field, $\varphi\in\text{Hom}(V_1,\ldots,V_k;V)$ a multilinear mapping, $\{n_i\}_{i\in\{1,\ldots,k\}}\subseteq\mathbb{N}$ and $\{w_{i,j}\}_{j\in\{1,\ldots,n_i\}}\subseteq V_i$, $i\in\{1,\ldots,k\}$ then*

$$\varphi\left(\sum_{j\in\{1,\ldots,n_1\}}w_{1,j},\ldots,\sum_{j\in\{1,\ldots,n_k\}}w_{k,j}\right)=\sum_{\gamma\in\Gamma(n_1,\ldots,n_k)}\varphi(w_{1,\gamma(1)},\ldots,w_{k,\gamma(k)})$$

Proof. We prove this by induction on k, so let

$$S=\left\{k\in\mathbb{N}|\text{If }\varphi\in\text{Hom}(V_1,\ldots,V_k;V),\{n_i\}_{i\in\{1,\ldots,k\}}\subseteq N,i\in\{1,\ldots,m\},\{w_{i,j}\}_{j\in\{1,\ldots,n_i\}}\text{ then }\varphi\left(\sum_{j\in\{1,\ldots,n_1\}}w_{1,j},\ldots,\sum_{j\in\{1,\ldots,n_k\}}w_{k,j}\right)=\sum_{\gamma\in\Gamma(n_1,\ldots,n_k)}\varphi(w_{1,\gamma(1)},\ldots,w_{k,\gamma(k)})\right\}$$

then we have:

$1\in S$. Define $\tau:\{1,\ldots,n_1\}\rightarrow\Gamma(n_1,\ldots,n_1)=\prod_{i\in\{1\}}\{1,\ldots,n_i\}$ where $\tau(i)$ is defined by $\tau(i)=(i)\Rightarrow\tau(i)_1=\tau(i)(1)=i$ then we have:

- **injectivity.** If $\tau(i)=\tau(j)$ then $i=\tau(i)(1)=\tau(j)(1)=j$
- **surjectivity.** If $\gamma\in\Gamma(n_1,\ldots,n_1)$ then $\gamma:\{1\}\rightarrow\{1,\ldots,n_1\}$. Take $i=\gamma(1)\in\{1,\ldots,n_1\}$ then $\forall j\in\{1\}$ we have $\tau(i)(j)=i=\gamma(1)=\gamma(j)$ proving that $\tau(i)=\gamma$.

proving that

$$\tau:\{1,\ldots,n_1\}\rightarrow\Gamma(n_1,\ldots,n_1)\text{ is a bijection}$$

Let $\{w_{1,j}\}_{j\in\{1,\ldots,n_1\}}\subseteq V_1$ and $\varphi\in\mathrm{Hom}(V_1,\ldots,V_1;V)$ then

$$\varphi\left(\sum_{j\in\{1,\ldots,n_1\}} w_{1,j}\right) \underset{\varphi \text{ is multilinear}}{=} \sum_{j\in\{1,\ldots,n_1\}} \varphi(w_{1,j}) = \sum_{j\in\{1,\ldots,n_1\}} \varphi(w_{1,\tau(j)(1)}) \underset{[\text{theorem: } 11.36]}{=} \sum_{\gamma\in\Gamma(n_1,\ldots,n_1)} \varphi(w_{1,\gamma(1)})$$

proving that $1\in S$.

***k* ∈ *S* ⇒ *k* + 1 ∈ *S*.** Let $\varphi\in\mathrm{Hom}(V_1,\ldots,V_{k+1};V)$, $\{n_i\}_{i\in\{1,\ldots,k+1\}}\subseteq\mathbb{N}$, $\{w_{i,j}\}_{j\in\{1,\ldots,n_i\}}$ where $i\in\{1,\ldots,k+1\}$. Define

$$\psi: \prod_{i\in\{1,\ldots,k\}} V_i\rightarrow V \text{ by } \psi\left(v_1,\ldots,v_k, \sum_{j\in\{1,\ldots,n_{k+1}\}} w_{k+1,j}\right)$$

then trivially $\psi\in\mathrm{Hom}(V_1,\ldots,V_k;V)$. Then we have

$$\begin{aligned}
\varphi\left(\sum_{j\in\{1,\ldots,n_1\}} w_{1,j},\ldots,\sum_{j\in\{1,\ldots,n_{k+1}\}} w_{k+1,j}\right) &= \\
\psi\left(\sum_{j\in\{1,\ldots,n_1\}} w_{1,j},\ldots,\sum_{j\in\{1,\ldots,n_k\}} w_{k,j}\right) &\underset{k\in S}{=} \\
\sum_{\gamma\in\Gamma(n_1,\ldots,n_k)} \psi(w_{1,\gamma(1)},\ldots,w_{k,\gamma(k)}) &= \\
\sum_{\gamma\in\Gamma(n_1,\ldots,n_k)} \varphi\left(w_{1,\gamma(1)},\ldots,w_{k,\gamma(k)},\sum_{j\in\{1,\ldots,n_{k+1}\}} w_{k+1,j}\right) &= \\
\sum_{\gamma\in\Gamma(n_1,\ldots,n_{k+1})}\left(\sum_{j\in\{1,\ldots,n_{k+1}\}} \varphi(w_{1,\gamma(1)},\ldots,w_{k,\gamma(k)},w_{k+1,j})\right) &\underset{[\text{theorem: } 11.45]}{=} \\
\sum_{(\gamma,j)\in\Gamma(n_1,\ldots,n_k)\times\{1,\ldots,n_{k+1}\}} \varphi(w_{1,\gamma(1)},\ldots,w_{k,\gamma(k)},w_{k+1,j}) &
\end{aligned}\tag{13.76}$$

Define now

$$\chi:\Gamma(n_1,\ldots,n_k)\times\{1,\ldots,n_{k+1}\}\rightarrow\Gamma(n_1,\ldots,n_{k+1})$$

where

$$\chi(\gamma,i) \text{ is defined by } \chi(\gamma,i)(j)=\begin{cases} i \text{ if } j=k+1\\ \gamma(j) \text{ if } j\in\{1,\ldots,k\}\end{cases}$$

then we have:

injectivity. If $\chi(\gamma,r)=\chi(\zeta,s)$ then $\forall j\in\{1,\ldots,k\}$ we have

$$\gamma(j)=\chi(\gamma,r)(j)=\chi(\zeta,s)(j)=\zeta(j)$$

proving that $\gamma=\zeta$. Further $r=\chi(\gamma,r)(k+1)=\chi(\zeta,s)(k+1)=s$ so that we have that $(\gamma,r)=(\zeta,s)$.

surjectivity. If $\gamma\in\Gamma(n_1,\ldots,n_{k+1})$ take then $\zeta=\gamma_{|\{1,\ldots,k\}}$ and $r=\gamma(k+1)$ then we have $\chi(\zeta,r)(j)=\begin{cases} r \text{ if } j=k+1\\ \zeta(j) \text{ if } j\in\{1,\ldots,k\}\end{cases}=\begin{cases} \gamma(k+1) \text{ if } j\in k+1\\ \gamma(j) \text{ if } j\in\{1,\ldots,k\}\end{cases}=\gamma(j)$. Hence $\chi(\zeta,r)=\gamma$ proving surjectivity.

Hence $\chi:\Gamma(n_1,\ldots,n_k)\times\{1,\ldots,n_{k+1}\}\rightarrow\Gamma(n_1,\ldots,n_{k+1})$ is a bijection. So

$$\begin{aligned}
\sum_{(\gamma,j)\in\Gamma(n_1,\ldots,n_k)\times\{1,\ldots,n_{k+1}\}} \varphi(w_{1,\gamma(1)},\ldots,w_{k,\gamma(k)},w_{k+1,j}) &= \\
\sum_{(\gamma,j)\in\Gamma(n_1,\ldots,n_k)\times\{1,\ldots,n_{k+1}\}} \varphi(w_{1,\chi(\gamma,j)(1)},\ldots,w_{k,\chi(\gamma,j)(k)},w_{k+1,\chi(\gamma,j)(k+1)}) &\underset{[\text{theorem: } 11.36]}{=} \\
\sum_{\gamma\in\Gamma(n_1,\ldots,n_{k+1})} \varphi(w_{1,\gamma(1)},\ldots,w_{k+1,\gamma(k+1)}) &
\end{aligned}$$

which combined with [eq: 13.76] gives

$$\varphi\left(\sum_{j\in\{1,\ldots,n_1\}} w_{1,j},\ldots,\sum_{j\in\{1,\ldots,n_{k+1}\}} w_{k+1,j}\right)=\sum_{\gamma\in\Gamma(n_1,\ldots,n_{k+1})}\varphi(w_{1,\gamma(1)},\ldots,w_{k+1,\gamma(k+1)})$$

proving that $k+1\in S$. □

Lemma 13.33. *Let $k\in\mathbb{N}$, $\{V_i\}_{i\in\{1,\ldots,k\}}$ a finite family of vector spaces over a field F, $\langle V_1\otimes\cdots\otimes V_n,\otimes\rangle$ a tensor product of $V_1,\ldots,V_n$, $\{n_i\}_{i\in\{1,\ldots,k\}}\subseteq\mathbb{N}$ and $\forall i\in\{1,\ldots,k\}$ a family $\{W_{i,j}\}_{j\in\{1,\ldots,n_i\}}$ a satisfying*

$$\forall j\in\{1,\ldots,n_i\}\ W_{i,j}\text{ is a sub-space of }V_i$$

and

$$V_i=\sum_{j\in\{1,\ldots,n_i\}} W_{i,j}\ [\text{see definition } 12.1]$$

then

$$V_1\otimes\cdots\otimes V_k=\sum_{\gamma\in\Gamma(n_1,\ldots,n_k)} W_{1,\gamma(1)}\otimes\cdots\otimes W_{k,\gamma(k)}$$

where $\forall\gamma\in\Gamma(n_1,\ldots,n_k)$

$$W_{1,\gamma(1)}\otimes\cdots\otimes W_{k,\gamma(k)}=\operatorname{span}\left(\otimes\left(\prod_{i\in\{1,\ldots,k\}} W_{i,\gamma(i)}\right)\right)$$

and

$$\left\langle W_{1,\gamma(1)}\otimes\cdots\otimes W_{k,\gamma(k)},\otimes_{|\prod_{j\in\{1,\ldots,k\}}W_{i,\gamma(i)}}\right\rangle\text{ is a tensor product of }W_{1,\gamma(1)},\ldots,W_{k,\gamma(k)}$$

with

$$W_{1,\gamma(1)}\otimes\cdots\otimes W_{k,\gamma(k)}\subseteq V_1\otimes\cdots\otimes V_n$$

[see theorem: 13.18].

Proof. Let $v=(v_1,\ldots,v_k)\in\prod_{i\in\{1,\ldots,k\}}V_i$ then by the hypothesis we have $\forall i\in\{1,\ldots,k\}$ that there exist a $\{w_{i,j}\}_{j\in\{1,\ldots,n_i\}}\subseteq W_i$ such that $v_i=\sum_{j\in\{1,\ldots,n_i\}}w_{i,j}$. As $\otimes$ is multilinear we have

$$\begin{aligned}\otimes(v_1,\ldots,v_k) &= \otimes\left(\sum_{j\in\{1,\ldots,n_1\}} w_{1,j},\ldots,\sum_{j\in\{1,\ldots,n_k\}} w_{k,j}\right)\\ &\underset{[\text{theorem: } 13.32]}{=} \sum_{\gamma\in\Gamma n_1,\ldots,n_k}\otimes(w_{1,\gamma(1)},\ldots,w_{k,\gamma(k)})\end{aligned}\tag{13.77}$$

As $\forall i\in\{1,\ldots,k\}$, $\forall j\in\{1,\ldots,n_i\}$ we have $w_{i,\gamma(i)}\in W_{i,\gamma(i)}$ we have

$$\begin{aligned}\otimes(w_{1,\gamma(1)},\ldots,w_{k,\gamma(k)}) &\in \otimes\left(\prod_{i\in\{1,\ldots,k\}} W_{i,\gamma(i)}\right)\\ &\subseteq \operatorname{span}\left(\otimes\left(\prod_{i\in\{1,\ldots,k\}} W_{i,\gamma(i)}\right)\right)\\ &= W_{1,\gamma(1)}\otimes\cdots\otimes W_{k,\gamma(k)}\end{aligned}$$

So using the above, [eq: 13.77], the definition of the sum of sets [see definition 12.1] it follows that

$$\otimes(v_1,\ldots,v_k)\in\sum_{\gamma\in\Gamma(n_1,\ldots,n_k)} W_{1,\gamma(1)}\otimes\cdots\otimes W_{k,\gamma(k)}$$

or

$$\otimes\left(\prod_{i\in\{1,\ldots,k\}} V_i\right)\subseteq\sum_{\gamma\in\Gamma(n_1,\ldots,n_k)} W_{1,\gamma(1)}\otimes\cdots\otimes W_{k,\gamma(k)}$$

which using [theorem: 11.104] gives

$$V_1\otimes\cdots\otimes V_k=\operatorname{span}\left(\otimes\left(\prod_{i\in\{1,\ldots,k\}} V_i\right)\right)\subseteq\operatorname{span}\left(\sum_{\gamma\in\Gamma(n_1,\ldots,n_k)} W_{1,\gamma(1)}\otimes\cdots\otimes W_{k,\gamma(k)}\right)$$

By [theorem: 13.18] $W_{1,\gamma(1)}\otimes\cdots\otimes W_{k,\gamma(k)}$ is a sub-space of $V=V_1\otimes\cdots\otimes V_k$ so that by [theorem: 12.4] $\sum_{\gamma\in\Gamma(n_1,\ldots,n_k)} W_{1,\gamma(1)}\otimes\cdots\otimes W_{k,\gamma(k)}$ is a sub-space of $V_1\otimes\cdots\otimes V_n$. Hence by [theorems: 11.104, 11.106]

$$\sum_{\gamma\in\Gamma(n_1,\ldots,n_k)} W_{1,\gamma(1)}\otimes\cdots\otimes W_{k,\gamma(k)}=\operatorname{span}\left(\sum_{\gamma\in\Gamma(n_1,\ldots,n_k)} W_{1,\gamma(1)}\otimes\cdots\otimes W_{k,\gamma(k)}\right)\subseteq V_1\otimes\cdots\otimes V_k$$

Combining the last two results proves then

$$\sum_{\gamma\in\Gamma(n_1,\ldots,n_k)} W_{1,\gamma(1)}\otimes\cdots\otimes W_{k,\gamma(k)}=V_1\otimes\cdots\otimes V_k$$

□

Lemma 13.34. *Let $\{V_i\}_{i\in\{1,\ldots,2\}}$ be a family of two vector spaces over a field and $\langle V_1\otimes V_2,\otimes\rangle$ a tensor product of V_1,V_2 then*

1. *If $V_1=W_1\oplus W_2$ [W_1,W_2 sub-spaces of V_1] then we have that*

$$V_1\otimes V_2=(W_1\otimes V_2)\oplus(W_2\otimes V_2)$$

where we have used [theorem: 13.18] to define

$$W_1\otimes V_2=\operatorname{span}(\otimes(W_1\cdot V_2))\subseteq V_1\otimes V_2,\ W_2\otimes V_2=\operatorname{span}(\otimes(W_2\cdot V_2))\subseteq V_1\otimes V_2$$

so that

$$\langle W_1\otimes V_2,\otimes_{|W_1\cdot V_2}\rangle \text{ and } \langle W_2\otimes V_2,\otimes_{|W_2\cdot V_2}\rangle \text{ are tensor products of } W_1,V_2 \text{ and } W_2,V_2$$

2. *If $V_2=W_1\oplus W_2$ [W_1,W_2 sub-spaces of V_2] then we have that*

$$V_1\otimes V_2=(V_1\otimes W_1)\oplus(V_2\otimes W_2)$$

where we have used [theorem: 13.18] to define

$$V_1\otimes W_1=\operatorname{span}(\otimes(V_1\cdot V_1))\subseteq V_1\otimes V_2,\ V_1\otimes W_2=\operatorname{span}(\otimes(V_1\cdot W_2))\subseteq V_1\otimes V_2$$

so that

$$\langle V_1\otimes W_1,\otimes_{|V_1\cdot W_1}\rangle \text{ and } \langle V_1\otimes W_2,\otimes_{|V_1\cdot W_2}\rangle \text{ are tensor products of } V_1,W_1 \text{ and } V_1,W_2$$

Here we use [definition: 6.78] for $W_1\cdot V_2$, $W_2\cdot V_2$, $V_1\cdot W_1$ and $V_1\cdot W_2$

Proof.

1. If $v=(v_1,v_2)\in V_1\times V_2$ then as $V_1=W_1\oplus W_2$ there exists **unique** $w_1\in W_1$, $w_2\in W_2$ such that $v_1=w_1+w_2$. Hence $\otimes(v)=\otimes(w_1+w_2,v_2)=\otimes(w_1,v_2)+\otimes(w_2,v_2)$ where

$$\otimes(w_1,v_2)\in\otimes(W_1\cdot V_2)\subseteq\operatorname{span}(\otimes(W_1\cdot V_2))=W_1\otimes V_2$$

and

$$\otimes(w_2,v_2)\in\otimes(W_2\cdot V_2)\subseteq\operatorname{span}(\otimes(W_2\cdot V_2))=W_2\otimes V_2$$

hence $\otimes(v)\in W_1\otimes V_2+W_2\otimes V_2$. So $\otimes(V_1\cdot V_2)\subseteq W_1\otimes V_2+W_2\otimes V_2$ which using [theorem: 11.104] gives

$$V_1\otimes V_2=\operatorname{span}(\otimes(V_1\cdot V_2))\subseteq\operatorname{span}(W_1\otimes V_2+W_2\otimes V_2)\tag{13.78}$$

As $W_1\otimes V_2$ and $W_2\otimes V_2$ are sub spaces of $V_1\otimes V_2$ [see theorem: 13.18] we have by [theorem: 12.4] that $W_1\otimes V_2+W_2\otimes V_2$ is a sub-space of $V_1\otimes V_2$, hence using [theorems: 11.104, 11.106] we have

$$W_1\otimes V_2+W_2\otimes V_2=\operatorname{span}(W_1\otimes V_2+W_2\otimes V_2)\subseteq V_1\otimes V_2$$

combining this with [eq: 13.78] gives then

$$W_1\otimes V_2+W_2\otimes V_2=V_1\otimes V_2\tag{13.79}$$

Assume now that there exist a $x\in(W_1\otimes V_2)\bigcap(W_2\otimes V_2)$ with $x\neq 0$. Using [theorem: 13.25] there exists $\{w_{1,i}\}_{i\in\{1,\ldots,k_x\}}\subseteq W_1$, $\{v_{1,i}\}_{i\in\{1,\ldots,k_x\}}\subseteq V_2$, $\{w_{2,i}\}_{i\in\{1,\ldots,l_x\}}\subseteq W_2$, $\{v_{2,i}\}_{i\in\{1,\ldots,l_x\}}\subseteq V_2$ such that

$$x=\sum_{i\in\{1,\ldots,k_x\}} w_{1,i}\otimes v_{1,i}\tag{13.80}$$

$$x=\sum_{i\in\{1,\ldots,l_x\}} w_{2,i}\otimes v_{2,i}\tag{13.81}$$

and $\{w_{1,i}\}_{i\in\{1,\ldots,k_x\}}\subseteq W_1$, $\{v_{1,i}\}_{i\in\{1,\ldots,k_x\}}\subseteq V_2$, $\{w_{2,i}\}_{i\in\{1,\ldots,l_x\}}\subseteq W_2$, $\{v_{2,i}\}_{i\in\{1,\ldots,l_x\}}\subseteq V_2$ are linear independent. Using [theorem: 11.179] there exist a $L\in\operatorname{Hom}(V_2,F)$ such that $L(v_{1,1})=1$ and $\forall i\in\{2,\ldots,k_x\}$ $L(v_{1,i})=0$. Let now $K\in\operatorname{Hom}(V_1,F)$ then

$$K\cdot L: V_1\times V_2\to F \text{ by } (K\cdot L)(v_1,v_2)=K(v_1)\cdot L(v_2)$$

then from the linearity of K, L it follows that $K \cdot L \in \mathrm{Hom}(V_1, V_2; F)$. As $\langle V_1 \otimes V_2, \otimes\rangle$ is a tensor product of V_1, V_2 there exist a $h \in \mathrm{Hom}(V_1 \otimes V_2, F)$ such that

$$K \cdot L = h \circ \otimes \tag{13.82}$$

$$\begin{aligned}
h(x) &\underset{[\text{eq: } 13.80]}{=} h\left(\sum_{i \in \{1, \ldots, k_x\}} w_{1,i} \otimes v_{1,i}\right)\\
&\underset{h \text{ is linear}}{=} \sum_{i \in \{1, \ldots, k_x\}} h(w_{1,i} \otimes v_{1,i})\\
&= \sum_{i \in \{1, \ldots, k_x\}} h(\otimes(w_{1,i}, v_{1,i}))\\
&\underset{[\text{eq: } 13.82]}{=} \sum_{i \in \{1, \ldots, k_x\}} (K \cdot L)(w_{1,i}, v_{1,i})\\
&= \sum_{i \in \{1, \ldots, k_x\}} K(w_{1,i}) \cdot L(v_{1,i})\\
&\underset{[\text{theorem: } 11.43]}{=} \sum_{i \in \{2, \ldots, k_x\}} K(w_{1,i}) \cdot L(v_{1,i}) + \sum_{i \in \{1\}} K(w_{1,i}) \cdot L(v_{1,i})\\
&\underset{[\text{theorem: } 11.34]}{=} \sum_{i \in \{2, \ldots, k_x\}} K(w_{1,i}) \cdot L(v_{1,i}) + K(w_{1,1}) \cdot L(v_{1,1})\\
&= \sum_{i \in \{2, \ldots, k_x\}} 0 \cdot L(v_{1,i}) + K(w_{1,1}) \cdot 1\\
&\underset{[\text{theorem: } 11.37]}{=} K(w_{1,1})
\end{aligned}$$

but we have also

$$\begin{aligned}
h(x) &\underset{[\text{theorem: } 13.81]}{=} h\left(\sum_{i \in \{1, \ldots, l_x\}} w_{2,i} \otimes v_{2,i}\right)\\
&= \sum_{i \in \{1, \ldots, l_x\}} h(w_{2,i} \otimes v_{2,i})\\
&= \sum_{i \in \{1, \ldots, l_x\}} h(\otimes(w_{2,i}, v_{2,i}))\\
&\underset{[\text{eq: } 13.82]}{=} \sum_{i \in \{1, \ldots, l_x\}} (K \cdot L)((w_{2,i}, v_{2,i}))\\
&= \sum_{i \in \{1, \ldots, l_x\}} K(w_{2,i}) \cdot L(v_{2,i})\\
&= \sum_{i \in \{1, \ldots, l_x\}} L(v_{2,i}) \cdot K(w_{2,i})\\
&\underset{K \text{ is linear}}{=} K\left(\sum_{i \in \{1, \ldots, l_x\}} L(v_{2,i}) \cdot w_{2,i}\right)
\end{aligned}$$

so we have

$$\forall K \in \mathrm{Hom}(V_1, F) \text{ we have } K(w_{1,1}) = K\left(\sum_{i \in \{1, \ldots, l_x\}} L(v_{2,i}) \cdot w_{2,i}\right) \tag{13.83}$$

So we can apply [theorem: 11.178] giving that $w_{1,1} = \sum_{i \in \{1, \ldots, l_x\}} L(v_{2,i}) \cdot w_{2,i}$. As $w_{1,1} \in W_1$ and $\sum_{i \in \{1, \ldots, l_x\}} L(v_{2,i}) \cdot w_{2,i} \in W_2$ [as W_2 is a sub-space of V_1] it follows that

$$w_{1,1} \in W_1 \bigcap W_2 \underset{V_1 = W_1 \oplus W_2}{=} \{0\}$$

proving that $w_{1,1} = 0$. However this means by [theorem: 11.124] that $\{w_{1,i}\}_{i \in \{1, \ldots, k_x\}}$ is linear dependent contradicting the fact that $\{w_{1,i}\}_{i \in \{1, \ldots, k_x\}}$ is linear independent. Hence the assumption $x \neq 0$ is wrong and we must have that $x = 0$ or $(W_1 \otimes V_2) \bigcap (W_2 \otimes V_2) \subseteq \{0\}$. This together with fact that by $W_1 \otimes V_2$, $W_2 \otimes V_2$ are vector spaces so that $0 \in (W_1 \otimes V_2) \bigcap (W_2 \otimes V_2)$ proves that

$$(W_1 \otimes V_2) \bigcap (W_2 \otimes V_2) = \{0\}$$

which together with [eq: 13.78] proves that

$$V_1 \otimes V_2 = (W_1 \otimes V_2) \oplus (W_2 \otimes V_2)$$

2. The proof is simular to the proof of (1). If $v = (v_1, v_2) \in V_1 \cdot V_2$ then as $V_2 = W_1 \oplus W_2$ there exists **unique** $w_1 \in W_1$, $w_2 \in W_2$ such that $v_2 = w_1 + w_2$. Hence $\otimes(v) = \otimes(v_1, w_1 + w_2) = \otimes(v_1, w_1) + \otimes(v_1, w_2)$ where

$$\otimes(v_1, w_1) \in \otimes(V_1 \cdot W_1) \subseteq \mathrm{span}(\otimes(V_1 \cdot W_1)) = V_1 \otimes W_1$$

and

$$\otimes(v_1, w_2) \in \otimes(V_1 \cdot W_2) \subseteq \mathrm{span}(\otimes(V_1 \cdot W_2)) = V_1 \otimes W_2$$

hence $\otimes(v) \in V_1 \otimes W_1 + V_1 \otimes W_2$. So $\otimes(V_1 \cdot V_2) \subseteq V_1 \otimes W_1 + V_1 \otimes W_2$ which using [theorem: 11.104] gives

$$V_1 \otimes V_2 = \mathrm{span}(\otimes(V_1 \cdot V_2)) \subseteq \mathrm{span}(V_1 \otimes W_1 + V_1 \otimes W_2) \tag{13.84}$$

As $V_1 \otimes W_1$ and $V_1 \otimes W_2$ are sub spaces of $V_1 \otimes V_2$ [see theorem: 13.18] we have by [theorem: 12.4] that $V_1 \otimes W_1 + V_1 \otimes W_2$ is a sub-space of $V_1 \otimes V_2$, hence using [theorems: 11.104, 11.106] we have

$$V_1 \otimes W_1 + V_1 \otimes W_2 = \mathrm{span}(V_1 \otimes W_1 + V_1 \otimes W_2) \subseteq V_1 \otimes V_2$$

combining this with [eq: 13.84] gives then

$$W_1 \otimes V_2 + W_2 \otimes V_2 = V_1 \otimes V_2 \tag{13.85}$$

Assume now that there exist a $x \in (V_1 \otimes W_1) \bigcap (V_1 \otimes W_2)$ with $x \neq 0$. Using [theorem: 13.25] there exists $\{v_{1,i}\}_{i \in \{1, \ldots, k_x\}} \subseteq V_1$, $\{w_{1,i}\}_{i \in \{1, \ldots, k_x\}} \subseteq W_1$, $\{v_{2,i}\}_{i \in \{1, \ldots, l_x\}} \subseteq V_1$, $\{W_{2,i}\}_{i \in \{1, \ldots, l_x\}} \subseteq W_2$ such that

$$x = \sum_{i \in \{1, \ldots, k_x\}} v_{1,i} \otimes w_{1,i} \tag{13.86}$$

$$x = \sum_{i \in \{1, \ldots, l_x\}} v_{2,i} \otimes w_{2,i} \tag{13.87}$$

and $\{v_{1,i}\}_{i \in \{1, \ldots, k_x\}} \subseteq V_1$, $\{w_{1,i}\}_{i \in \{1, \ldots, k_x\}} \subseteq W_1$, $\{v_{2,i}\}_{i \in \{1, \ldots, l_x\}} \subseteq V_1$, $\{w_{2,i}\}_{i \in \{1, \ldots, l_x\}} \subseteq W_2$ are linear independent. Using [theorem: 11.179] there exist a $L \in \mathrm{Hom}(V_1, F)$ such that $L(v_{1,1}) = 1$ and $\forall i \in \{2, \ldots, k_x\}$ $L(v_{1,i}) = 0$. Let now $K \in \mathrm{Hom}(V_2, F)$ then

$$L \cdot K: V_1 \cdot V_2 \rightarrow F \text{ by } (L \cdot K)(v_1, v_2) = L(v_1) \cdot K(v_2)$$

then from the linearity of L, K it follows that $L \cdot K \in \mathrm{Hom}(V_1, V_2; F)$. As $\langle V_1 \otimes V_2, \otimes \rangle$ is a tensor product of V_1, V_2 there exist a $h \in \mathrm{Hom}(V_1 \otimes V_2, F)$ such that

$$L \cdot K = h \circ \otimes \tag{13.88}$$

$$\begin{aligned}
h(x) &\underset{\text{[eq: 13.86]}}{=} h\left(\sum_{i \in \{1, \ldots, k_x\}} v_{1,i} \otimes w_{1,i}\right) \\
&\underset{h \text{ is linear}}{=} \sum_{i \in \{1, \ldots, k_x\}} h(v_{1,i} \otimes w_{1,i}) \\
&= \sum_{i \in \{1, \ldots, k_x\}} h(\otimes(v_{1,i}, w_{1,i})) \\
&\underset{\text{[eq: 13.88]}}{=} \sum_{i \in \{1, \ldots, k_x\}} (L \cdot K)(v_{1,i}, w_{1,i}) \\
&= \sum_{i \in \{1, \ldots, k_x\}} L(v_{1,i}) \cdot K(w_{1,i}) \\
&\underset{\text{[theorem: 11.43]}}{=} \sum_{i \in \{2, \ldots, k_x\}} L(v_{1,i}) \cdot K(w_{1,i}) + \sum_{i \in \{1\}} L(v_{1,i}) \cdot K(w_{1,i}) \\
&\underset{\text{[theorem: 11.34]}}{=} \sum_{i \in \{2, \ldots, k_x\}} L(v_{1,i}) \cdot K(w_{1,i}) + L(v_{1,1}) \cdot K(w_{1,1}) \\
&= \sum_{i \in \{2, \ldots, k_x\}} 0 \cdot L(v_{1,i}) + K(w_{1,1}) \cdot 1 \\
&\underset{\text{[theorem: 11.37]}}{=} K(w_{1,1})
\end{aligned}$$

but we have also

$$\begin{aligned}
h(x) &\underset{[\text{theorem: } 13.87]}{=} h\left(\sum_{i\in\{1,\ldots,l_x\}} v_{2,i}\otimes w_{2,i}\right)\\
&= \sum_{i\in\{1,\ldots,l_x\}} h(v_{2,i}\otimes w_{2,i})\\
&= \sum_{i\in\{1,\ldots,l_x\}} h(\otimes(v_{2,i},w_{2,i}))\\
&\underset{[\text{eq: } 13.88]}{=} \sum_{i\in\{1,\ldots,l_x\}} (L\cdot K)((v_{2,i},w_{2,i}))\\
&= \sum_{i\in\{1,\ldots,l_x\}} L(v_{2,i})\cdot L(w_{2,i})\\
&= \sum_{i\in\{1,\ldots,l_x\}} L(v_{2,i})\cdot K(w_{2,i})\\
&\underset{K \text{ is linear}}{=} K\left(\sum_{i\in\{1,\ldots,l_x\}} L(v_{2,i})\cdot w_{2,i}\right)
\end{aligned}$$

so we have

$$\forall K\in \mathrm{Hom}(V_1,F) \text{ we have } K(w_{1,1}) = K\left(\sum_{i\in\{1,\ldots,l_x\}} L(v_{2,i})\cdot w_{2,i}\right) \tag{13.89}$$

So we can apply [theorem: 11.178] giving that $w_{1,1}=\sum_{i\in\{1,\ldots,l_x\}} L(v_{2,i})\cdot w_{2,i}$. As $w_{1,1}\in W_1$ and $\sum_{i\in\{1,\ldots,l_x\}} L(v_{2,i})\cdot w_{2,i}\in W_2$ [as W_2 is a sub-space of V_1] it follows that

$$w_{1,1}\in W_1\bigcap W_2 \underset{V_1=W_1\oplus W_2}{=} = \{0\}$$

proving that $w_{1,1}=0$. However this means by [theorem: 11.124] that $\{w_{1,i}\}_{i\in\{1,\ldots,k_x\}}$ is linear dependent contradicting the fact that $\{w_{1,i}\}_{i\in\{1,\ldots,k_x\}}$ is linear independent. Hence the assumption $x\neq 0$ is wrong and we must have that $x=0$ or $(V_1\otimes W_1)\bigcap (V_1\otimes W_2)\subseteq \{0\}$. This together with fact that by $V_1\otimes W_1$, $V_1\otimes W_2$ are vector spaces so that $0\in (V_1\otimes W_1)\bigcap (V_1\otimes W_2)$ proves that

$$(V_1\otimes W_1)\bigcap (V_1\otimes W_2) = \{0\}$$

which together with [eq: 13.78] proves that

$$V_1\otimes V_2 = (V_1\otimes W_1)\oplus (V_1\otimes W_2)$$

□

Lemma 13.35. *Let $n\in\mathbb{N}$, $\{V_i\}_{i\in\{1,2\}}$ be vector spaces over a field F then*

1. *If $V_1=\bigoplus_{i\in\{1,\ldots,n\}} W_i$ then if $\langle V_1\otimes V_2,\otimes\rangle$ is a tensor product of V_1, V_2 we have*

$$V_1\otimes V_2 = \bigoplus_{i\in\{1,\ldots,n\}} W_i\otimes V_2$$

where $\forall i\in\{1,\ldots,n\}$ $W_i\otimes V_2=\mathrm{span}(\otimes(W_i\cdot V_2))$ so that $\langle W_i\otimes V_2,\otimes_{|W_i\cdot V_2}\rangle$ is a tensor product of W_i, V_2 [see theorem: 13.18]

2. *If $V_2=\bigoplus_{i\in\{1,\ldots,n\}} W_i$ then if $\langle V_1\otimes V_2,\otimes\rangle$ is a tensor product of V_1, V_2 we have*

$$V_1\otimes V_2 = \bigoplus_{i\in\{1,\ldots,n\}} V_1\otimes W_i$$

where $\forall i\in\{1,\ldots,n\}$ $V_1\otimes V_2=\mathrm{span}(\otimes(V_1\cdot W_i))$ so that $\langle V_1\otimes W_i,\otimes_{|V_1\cdot W_i}\rangle$ is a tensor product of V, W_i [see theorem: 13.18]

Proof.

1. We proceed by mathematical induction, so let

$$S=\left\{n\in\mathbb{N}|\text{If } V_1,V_2 \text{ are vector spaces over a field } F \text{ such that } V_1=\bigoplus_{i\in\{1,\ldots,n\}} W_i \text{ then } V_1\otimes V_2=\bigoplus_{i\in\{1,\ldots,n\}} W_i\otimes V_2\right\}$$

then we have that:

$1\in S$. Then if $V_1=\sum^{\oplus}_{i\in\{1\}} W_i$ we have by [example: 12.7] that $V_1=W_1$ and

$$V_1\otimes V_2=W_1\otimes V_2 \underset{\text{[example: 12.7]}}{=} \sum^{\oplus}_{i\in\{1\}} W_i\otimes V_2$$

proving that $1\in S$.

$n\in S\Rightarrow n+1\in S$. If $V_1=\bigoplus_{i\in\{1,\ldots,n+1\}} W_i$ then we have by [theorem: 12.11] that

$$\sum_{i\in\{1,\ldots n\}} W_i=\bigoplus_{i\in\{1,\ldots,n\}} W_i \text{ and } V_1=\left(\sum_{i\in\{1,\ldots n\}} W_i\right)\oplus W_{n+1}$$

or, if we take $V_1'=\sum_{i\in\{1,\ldots n\}} W_i$, that

$$V_1'=\bigoplus_{i\in\{1,\ldots,n\}} W_i \text{ and } V_1=V_1'\oplus W_{n+1} \tag{13.90}$$

By [lemma: 13.34] we have then

$$V_1\otimes V_2=(V_1'\otimes V_2)\oplus(W_{n+1}\otimes V_2) \tag{13.91}$$

Further as $V_1'=\bigoplus_{i\in\{1,\ldots,n\}} V_i$ and $n\in S$ we have that

$$V_1'\otimes V_2=\bigoplus_{i\in\{1,\ldots,n\}} W_i\otimes V_2 \tag{13.92}$$

Using then [theorem: 12.9] on [eqs: 13.91 and 13.92] results in

$$V_1\otimes V_2=\bigoplus_{i\in\{1,\ldots,n+1\}} W_i\otimes V_2$$

proving that $n+1\in S$

2. We proceed by mathematical induction, so let

$$S=\left\{n\in\mathbb{N}|\text{If } V_1,V_2 \text{ are vector spaces over a field } F \text{ such that } V_2=\bigoplus_{i\in\{1,\ldots,n\}} W_i \text{ then } V_1\otimes V_2=\bigoplus_{i\in\{1,\ldots,n\}} V_1\otimes W_i\right\}$$

then we have that:

$1\in S$. Then if $V_2=\sum^{\oplus}_{i\in\{1\}} W_i$ we have by [example: 12.7] that $V_2=W_1$ and

$$V_1\otimes V_2=V_1\otimes W_1 \underset{\text{[example: 12.7]}}{=} \bigoplus_{i\in\{1\}} V_i\otimes W_i$$

proving that $1\in S$.

$n\in S\Rightarrow n+1\in S$. If $V_2=\bigoplus_{i\in\{1,\ldots,n+1\}} W_i$ then we have by [theorem: 12.11] that

$$\sum_{i\in\{1,\ldots n\}} W_i=\bigoplus_{i\in\{1,\ldots,n\}} W_i \text{ and } V_2=\left(\sum_{i\in\{1,\ldots n\}} W_i\right)\oplus W_{n+1}$$

or, if we take $V_{12}'=\sum_{i\in\{1,\ldots n\}} W_i$, that

$$V_2'=\bigoplus_{i\in\{1,\ldots,n\}} W_i \text{ and } V_2=V_2'\oplus W_{n+1} \tag{13.93}$$

By [lemma: 13.34] we have then

$$V_1\otimes V_2=(V_1\otimes V_2')\oplus(V_1\otimes W_{n+1}) \tag{13.94}$$

Further as $V_2'=\bigoplus_{i\in\{1,\ldots,n\}} V_i$ and $n\in S$ we have that

$$V_1\otimes V_2'=\bigoplus_{i\in\{1,\ldots,n\}} V_1\otimes W_i \tag{13.95}$$

Using then [theorem: 12.9] on [eqs: 13.94 and 13.95] results in

$$V_1\otimes V_2=\bigoplus_{i\in\{1,\ldots,n+1\}} V_1\otimes W_i$$

proving that $n+1\in S$ □

Corollary 13.36. *Let $n\in\mathbb{N}$, $\{V_i\}_{i\in\{1,2\}}$ be vector spaces over a field F and I a finite set then*

1. *If $V_1=\bigoplus_{i\in I}W_i$ then if $\langle V_1\otimes V_2,\otimes\rangle$ is a tensor product of V_1, V_2 we have*

$$V_1\otimes V_2=\bigoplus_{i\in I}W_i\otimes V_2$$

where $\forall i\in I$ $W_i\otimes V_2=\operatorname{span}(\otimes(W_i\cdot V_2))$ so that $\langle W_i\otimes V_2,\otimes_{|W_i\cdot V_2}\rangle$ is a tensor product of W_i,V_2 [see theorem: 13.18]

2. *If $V_2=\bigoplus_{i\in I}W_i$ then if $\langle V_1\otimes V_2,\otimes\rangle$ is a tensor product of V_1, V_2 we have*

$$V_1\otimes V_2=\bigoplus_{i\in I}V_1\otimes W_i$$

where $\forall i\in I$ $V_1\otimes V_2=\operatorname{span}(\otimes(V_1\cdot W_i))$ so that $\langle V_1\otimes W_i,\otimes_{|V_1\cdot W_i}\rangle$ is a tensor product of V,W_i [see theorem: 13.18]

Proof. As I is finite there exist a $n\in\mathbb{N}$ and a bijection $\beta: I\rightarrow\{1,\ldots,n\}$

1. Define $\{U_i\}_{i\in\{1,\ldots,n\}}$ by $U_i=W_{\beta(i)}$ then we have

$$\bigoplus_{i\in\{1,\ldots,n\}}U_i=\bigoplus_{i\in\{1,\ldots,n\}}W_{\beta(i)}\underset{\text{[theorem: 12.13]}}{=}\bigoplus_{i\in I}W_i=V_1$$

Using the previous lemma [lemma: 13.34] we have then that

$$V_1\otimes V_2=\bigoplus_{i\in\{1,\ldots,n\}}U_i\otimes V_2=\bigoplus_{i\in\{1,\ldots,n\}}W_{\beta(i)}\otimes V_2\underset{\text{[theorem: 12.13]}}{=}\bigoplus_{i\in I}W_i\otimes V_2$$

and $\forall i\in\{1,\ldots,n\}$ $U_i\otimes V_2=\operatorname{span}(\otimes(U_i\cdot V_2))$. Hence we have

$$V_1\otimes V_2=\bigoplus_{i\in I}W_i\otimes V_2$$

and if $i\in I$ we have, as $W_i=W_{\beta(\beta^{-1}(i))}=U_{\beta^{-1}(i)}$, that

$$W_i\otimes V_2=U_{\beta^{-1}(i)}=\operatorname{span}(\otimes(U_{\beta^{-1}(i)}\cdot V_2))=\operatorname{span}(\otimes(W_i\cdot V_2))$$

2. Define $\{U_i\}_{i\in\{1,\ldots,n\}}$ by $U_i=W_{\beta(i)}$ then we have

$$\bigoplus_{i\in\{1,\ldots,n\}}U_i=\bigoplus_{i\in\{1,\ldots,n\}}W_{\beta(i)}\underset{\text{[theorem: 12.13]}}{=}\bigoplus_{i\in I}W_i=V_2$$

Using the previous lemma [lemma: 13.34] we have then that

$$V_1\otimes V_2=\bigoplus_{i\in\{1,\ldots,n\}}V_1\otimes U_i=\bigoplus_{i\in\{1,\ldots,n\}}V_1\otimes W_{\beta(i)}\underset{\text{[theorem: 12.13]}}{=}\bigoplus_{i\in I}V_1\otimes W_i$$

and $\forall i\in\{1,\ldots,n\}$ $V_1\otimes U_i=\operatorname{span}(\otimes(V_1\cdot U_i))$. Hence we have

$$V_1\otimes V_2=\bigoplus_{i\in I}V_1\otimes W_i$$

and if $i\in I$ we have, as $W_i=W_{\beta(\beta^{-1}(i))}=U_{\beta^{-1}(i)}$, that

$$V_1\otimes W_i=U_{\beta^{-1}(i)}=\operatorname{span}(\otimes(V_1\cdot U_{\beta^{-1}(i)}))=\operatorname{span}(\otimes(V_1\cdot W_i))$$ □

Lemma 13.37. *Let V,W be two vector spaces over a field such that*

$$V=\bigoplus_{i\in\{1,\ldots,k\}}V_i\text{ and }W=\bigoplus_{i\in\{1,\ldots,l\}}W_i$$

then if $\langle V \otimes W, \otimes\rangle$ *is a tensor product of* V, W *we have*

$$V \otimes W = \bigoplus_{\gamma \in \Gamma(k,l)} V_{\gamma(1)} \otimes W_{\gamma(2)}$$

where as usual $\forall \gamma \in \Gamma(k,l)$ $V_{\gamma(1)} \otimes W_{\gamma(2)} = \operatorname{span}(\otimes(V_{\gamma(1)} \cdot W_{\gamma(2)}))$ *so that* $\langle V_{\gamma(1)} \otimes W_{\gamma(2)}, \otimes_{|V_{\gamma(1)} \cdot W_{\gamma(2)}}\rangle$ *is a tensor product of* $V_{\gamma(1)} \cdot W_{\gamma(2)}$ *[see theorem: 13.18]*

Proof. We use induction to prove this. Let V be a vector space over a field F such that $V = \bigoplus_{i \in \{1,\ldots,k\}} V_i$. Define

$$S_V = \left\{ l \in \mathbb{N} | \text{If } W \text{ is a vector space over } F \text{ with } W = \bigoplus_{i \in \{1,\ldots,l\}} W_i \text{ then } V \otimes W = \bigoplus_{\gamma \in \Gamma(k,l)} V_{\gamma(1)} \otimes W_{\gamma(2)} \right\}$$

then we have:

$\boldsymbol{1 \in S_V}$**.** If $W = \bigoplus_{i \in \{1\}} W_i$ then we have by [example: 12.7] that $W = W_1$ so we can use [lemma: 13.35] to get

$$V \otimes W = \bigoplus_{i \in \{1,\ldots,k\}} V_i \otimes W_1 \tag{13.96}$$

Define

$$\tau: \{1,\ldots,k\} \to \Gamma(k,1) \text{ where } \tau(i) \text{ is defined by } \tau(i) = (i,1) \text{ so that } \tau(i)(1) = i \wedge \tau(i)(2) = 1$$

then we have:

injectivity. If $\tau(i) = \tau(j)$ then $i = \gamma(i)(1) = \gamma(j)(1) = j$ proving injectivity.

surjectivity. If $\gamma \in \Gamma(k,1)$ then $\gamma(1) \in \{1,\ldots,k\}$ and $\gamma(2) \in \{1\} \Rightarrow \gamma(2) = 1$, hence if we take $i = \gamma(1)$ we have $\tau(i)(1) = i = \gamma(1) \wedge \tau(i)(2) = 1 = \gamma(2)$ proving that $\tau(i) = \gamma$.

so we conclude that

$$\tau: \{1,\ldots,k\} \to \Gamma(k,1) \text{ is a bijection with } \tau(i)(1) = i \wedge \tau(i)(2) = 1 \tag{13.97}$$

Next

$$\begin{aligned} \bigoplus_{i \in \{1,\ldots,k\}} V_i \otimes W_1 &= \bigoplus_{i \in \{1,\ldots,k\}} V_{\tau(i)(1)} \otimes W_{\tau(i)(2)} \\ &\underset{[\text{theorem: } 12.13]}{=} \bigoplus_{\gamma \in \Gamma(k,1)} V_{\gamma(1)} \otimes W_{\gamma(2)} \end{aligned}$$

which together with [eq: 13.96] proves that $1 \in S_V$.

$\boldsymbol{l \in S \Rightarrow l+1 \in S_V}$**.** If $W = \bigoplus_{i \in \{1,\ldots,l+1\}} W_i$ then we have by [theorem: 12.11] that

$$\sum_{i \in \{1,\ldots,l\}} W_i = \bigoplus_{i \in \{1,\ldots,l\}} W_i \text{ and } W = \left(\sum_{i \in \{1,\ldots,l\}} W_i \right) \oplus W_{l+1}$$

or

$$W' = \bigoplus_{i \in \{1,\ldots,l\}} W_i \text{ and } W = W' \oplus W_{l+1} \text{ where } W' = \sum_{i \in \{1,\ldots,l\}} W_i \tag{13.98}$$

Hence using [lemma: 13.34]

$$V \otimes W = (V \otimes W') \oplus (V \otimes W_{l+1}) \tag{13.99}$$

As $l \in S$ and $W' \underset{[\text{eq: } 13.98]}{=} \bigoplus_{i \in \{1,\ldots,l\}} W_i$ we have that

$$V \otimes W' = \bigoplus_{\gamma \in \Gamma(k,l)} V_{\gamma(1)} \otimes W_{\gamma(2)} \tag{13.100}$$

Further, as $V = \bigoplus_{i \in 1,\ldots,k} V_i$, we have by [lemma: 13.35] that

$$V \otimes W_{l+1} = \bigoplus_{i \in \{1,\ldots,k\}} V_i \otimes W_{l+1} \tag{13.101}$$

Define

$$\chi: \{1, \ldots, k\} \to \{1, \ldots, k\} \cdot \{l+1\} \text{ where } \chi(i) = (i, l+1) \Rightarrow \chi(i)(1) = i \wedge \chi(i)(2) = l+1$$

then we have:

injectivity. If $\chi(i) = \chi(j)$ then $i = \chi(i)(1) = \chi(j)(1) = j$ proving injectivity.

surjectivity. Let $\gamma \in \{1, \ldots, k\} \cdot \{l+1\}$ and take $i = \gamma(1)$ then we have $\chi(i)(1) = i = \gamma(1)$ and $\chi(i)(2) = l+1 = \gamma(2)$ [as $\gamma(2) \in \{l+1\}$] so that $\chi(i) = \gamma$

So

$$\chi: \{1, \ldots, k\} \to \{1, \ldots, k\} \cdot \{l+1\} \text{ where } \chi(i)(1) = i \wedge \chi(i)(2) = l+1 \text{ is a bijection} \tag{13.102}$$

Using this bijection we have

$$\begin{aligned} V \otimes W_{l+1} &\underset{\text{[eq: 13.101]}}{=} \bigoplus_{i \in \{1, \ldots, k\}} V_i \otimes W_{l+1} \\ &= \bigoplus_{i \in \{1, \ldots, k\}} V_i \otimes W_{l+1} \\ &= \bigoplus_{i \in \{1, \ldots, k\}} V_{\chi(i)(1)} \otimes W_{\chi(i)(2)} \\ &\underset{\text{[eq: 13.102] and [theorem: 12.13]}}{=} \bigoplus_{\gamma \in \{1, \ldots, k\} \cdot \{l+1\}} V_{\gamma(1)} \otimes W_{\gamma(2)} \end{aligned}$$

so that

$$V \otimes W_{l+1} = \bigoplus_{\gamma \in \{1, \ldots, k\} \cdot \{l+1\}} V_{\gamma(1)} \otimes W_{\gamma(2)} \tag{13.103}$$

Now

$$\begin{aligned} \Gamma(k, l+1) &= \{1, \ldots, k\} \cdot \{1, \ldots, l+1\} \\ &= \{1, \ldots, k\} \cdot \left(\{1, \ldots, l\} \bigcup \{l+1\}\right) \\ &\underset{\text{[theorem: 6.79]}}{=} (\{1, \ldots, k\} \cdot \{1, \ldots, l\}) \bigcup (\{1, \ldots, k\} \cdot \{l+1\}) \\ &= \Gamma(k, l) \bigcup (\{1, \ldots, k\} \cdot \{l+1\}) \end{aligned}$$

and

$$\begin{aligned} \Gamma(k, l) \bigcap (\{1, \ldots, k\} \cdot \{l+1\}) &= (\{1, \ldots, k\} \cdot \{1, \ldots, l\}) \bigcap (\{1, \ldots, k\} \cdot \{l+1\}) \\ &\underset{\text{[theorem: 6.79]}}{=} \left(\{1, \ldots, k\} \bigcap \{1, \ldots, k\}\right) \cdot \left(\{1, \ldots l\} \bigcap \{l+1\}\right) \\ &= \left(\{1, \ldots, k\} \bigcap \{1, \ldots, k\}\right) \cdot \varnothing \\ &\underset{\text{[theorem: 2.136]}}{=} \varnothing \end{aligned}$$

So we have

$$\begin{aligned} \Gamma(k, l+1) &= \Gamma(k, l) \bigcup (\{1, \ldots, k\} \cdot \{l+1\}) \wedge \Gamma(k, l) \bigcap (\{1, \ldots, k\} \cdot \{l+1\}) = \varnothing \\ V \otimes W' &\underset{\text{[eq: 13.100]}}{=} \bigoplus_{\gamma \in \Gamma(k, l)} V_{\gamma(1)} \otimes W_{\gamma(2)} \\ V \otimes W_{l+1} &\underset{\text{[eq: 13.103]}}{=} \bigoplus_{\gamma \in \{1, \ldots, k\} \cdot \{l+1\}} V_{\gamma(1)} \otimes W_{\gamma(2)} \\ V \otimes W &\underset{\text{[eq: 13.99]}}{=} (V \otimes W') \oplus (V \otimes W_{l+1}) \end{aligned}$$

which allows us to use [theorem: 12.9] to get

$$V \otimes W = \bigoplus_{\gamma \in \Gamma(k, l+1)} V_{\gamma(1)} \otimes W_{\gamma(2)}$$

which proves that

$$l+1\in S$$

□

Theorem 13.38. *Let $n\in\mathbb{N}_0$, $\{V_i\}_{i\in\{1,\ldots,n\}}$ be vector spaces over a field V such that $\forall i\in\{1,\ldots,n\}$ $\{W_{i,j}\}_{j\in\{1,\ldots,k_i\}}\subseteq\mathcal{P}(V_i)$ is a family of sub-spaces of V_i such that $V_i=\bigoplus_{j\in\{1,\ldots,k_i\}}W_{i,j}$ then we have for a tensor product $\langle V_1\otimes\cdots\otimes V_n,\otimes\rangle$ that*

$$V_1\otimes\cdots\otimes V_n=\bigoplus_{\gamma\in\Gamma(k_1,\ldots,k_n)}(W_{1,\gamma(1)}\otimes\cdots\otimes W_{n,\gamma(n)})$$

where

$$W_{1,\gamma(1)}\otimes\cdots\otimes W_{n,\gamma(n)}\text{ is a tensor product of }W_{1,\gamma(1)},\ldots,W_{n,\gamma(n)}$$

Proof. We prove this by induction on n, so let

$S=\left\{n\in\mathbb{N}|\{V_i\}_{i\in\{1,\ldots,n\}}\text{ are vector spaces, }\{W_{i,j}\}_{j\in\{1,\ldots,k_i\}}\subseteq\mathcal{P}(V_i),\ V_i=\bigoplus_{j\in\{1,\ldots,k_i\}}W_{i,j},\ i\in\{1,\ldots,n\}\text{ and }\langle V,\otimes\cdots\otimes V_n,\otimes\rangle\text{ a tensor product of }V_1,\ldots,V_n\text{ then }V_1\otimes\cdots\otimes V_n=\bigoplus_{\gamma\in\Gamma(k_1,\ldots,k_n)}W_{1,\gamma(1)}\otimes\cdots\otimes W_{n,\gamma(n)}\text{ where }W_{1,\gamma(1)}\otimes\cdots\otimes W_{n,\gamma(n)}\text{ is a tensor product of }W_{1,\gamma(1)},\ldots,W_{n,\gamma(n)}\right\}$

then we have:

$1\in S$. Let $\langle V_1\otimes\cdots\otimes V_1,\otimes\rangle$ be a tensor product of $V_1,\ldots,V_1$ and $\{W_{1,j}\}_{j\in\{1,\ldots,n_1\}}$ a family of sub-spaces of V_1 such that $V_1=\bigoplus_{j\in\{1,\ldots,k_1\}}W_{1,j}$. Using [theorem: 13.22] there exist for every $j\in\{1,\ldots,k_1\}$ a tensor product $W_{1,j}\otimes\cdots\otimes W_{1,j}$ of $W_{1,j},\ldots,W_{1,j}$ such that

$$V_1\otimes\cdots\otimes V_1=\bigoplus_{j\in\{1,\ldots,k_1\}}W_{1,j}\otimes\cdots\otimes W_{1,j}$$

Define now

$$\kappa:\{1,\ldots,k_1\}\rightarrow\Gamma(k_1)\text{ by }\kappa(i)(1)=i$$

then we have:

injectivity. If $i,j\in\{1,\ldots,k_i\}$ satisfies $\kappa(i)=\kappa(j)$ then $i=\kappa(i)(1)=\kappa(j)(1)=j$ proving injectivity.

surjectivity. Let $\gamma\in\Gamma(k_1)$ then for $i=\gamma(1)$ we have $\kappa(i)(1)=i=\gamma(1)$ proving that $\kappa(i)=\gamma$ hence surjectivity.

proving that

$$\kappa:\{1,\ldots,k\}\rightarrow\Gamma(k_1)\text{ where }\kappa(i)(1)=i$$

So

$$\begin{aligned}V_1\otimes\cdots\otimes V_1 &= \bigoplus_{j\in\{1,\ldots,k_1\}}W_{1,j}\otimes\cdots\otimes W_{1,j}\\ &= \bigoplus_{j\in\{1,\ldots,k_1\}}W_{1,\kappa(j)(1)}\otimes\cdots\otimes W_{1,\kappa(j)(1)}\\ &\underset{[\text{theorem: }12.13]}{=} \bigoplus_{\gamma\in\Gamma(k_1)}W_{1,\gamma(1)}\otimes\cdots\otimes W_{1,\gamma(1)}\end{aligned}$$

proving that $1\in S$.

$n\in S\Rightarrow n+1\in S$. Let $\langle V_1\otimes\cdots\otimes V_{n+1},\otimes_{n+1}\rangle$ be a tensor product of $V_1,\ldots,V_n$ and $\forall i\in\{1,\ldots,n+1\}$ $\{W_{i,j}\}_{j\in\{1,\ldots,k_{n+1}\}}$ a family of sub-spaces of V_i such that $V_i=\bigoplus_{j\in\{1,\ldots,k_{n+1}\}}W_{i,j}$. Using [theorem: 13.28] we have that there exist a tensor product $V_1\otimes\cdots\otimes V_n$ of $V_1,\ldots,V_n$ and a tensor product $(V_1\otimes\cdots\otimes V_n)\circ V_{n+1}$ of $V_1\otimes\cdots\otimes V_n$, V_{n+1} such that

$$V_1\otimes\cdots\otimes V_{n+1}=(V_1\otimes\cdots\otimes V_n)\otimes V_{n+1}\tag{13.104}$$

As $n\in S$ there exists $\forall\gamma\in\Gamma(k_1,\ldots,k_n)$ a tensor product $W_{1,\gamma(1)}\otimes\cdots\otimes W_{n,\gamma(n)}$ of $W_{1,\gamma(1)}\otimes\cdots\otimes W_{n,\gamma(n)}$ such that

$$V_1\otimes\cdots\otimes V_n=\bigoplus_{\gamma\in\Gamma(k_1,\ldots,k_n)}W_{1,\gamma(1)}\otimes\cdots\otimes W_{n,\gamma(n)}\tag{13.105}$$

So we have

$$\begin{array}{rc} V_1 \otimes \cdots \otimes V_{n+1} & \underset{[\text{eq: } 13.104]}{=} \\ (V_1 \otimes \cdots \otimes V_n) \otimes V_{n+1} & \underset{[\text{eq: } 13.105]}{=} \\ \left(\bigoplus_{\gamma \in \Gamma(k_1, \ldots, k_n)} W_{1,\gamma(1)} \otimes \cdots \otimes W_{n,\gamma(n)} \right) \oplus V_{n+1} & = \\ \left(\bigoplus_{\gamma \in \Gamma(k_1, \ldots, k_n)} W_{1,\gamma(1)} \otimes \cdots \otimes W_{n,\gamma(n)} \right) \otimes \left(\bigoplus_{j \in \{1, \ldots, k_{n+1}\}} W_{n+1,j} \right) & \underset{[\text{theorem: } 13.35]}{=} \\ \bigoplus_{j \in \{1, \ldots, k_{n+1}\}} \left(\left(\bigoplus_{\gamma \in \Gamma(k_1, \ldots, k_n)} W_{1,\gamma(1)} \otimes \cdots \otimes W_{n,\gamma(n)} \right) \otimes W_{n+1,j} \right) & \underset{[\text{theorem: } 13.36]}{=} \\ \bigoplus_{j \in \{1, \ldots, k_{n+1}\}} \left(\bigoplus_{\gamma \in \Gamma(k_1, \ldots, k_n)} (W_{1,\gamma(1)} \otimes \cdots \otimes W_{n,\gamma(n)}) \otimes W_{n+1,j} \right) & \end{array} \tag{13.106}$$

Given $j \in \{1, \ldots, k_{n+1}\}$ define

$$\Gamma^j(k_1, \ldots, k_{n+1}) = \{\gamma \in \Gamma(k_1, \ldots, k_{n+1}) | \gamma(n+1) = j\} \subseteq \Gamma(k_1, \ldots, k_{n+1})$$

If $\gamma \in \Gamma(k_1, \ldots, k_{n+1})$ then clearly $\gamma(n+1) \in \{1, \ldots, k_{n+1}\}$ so that $\gamma \in \Gamma^{\gamma(n+1)}(k_1, \ldots, k_{n+1})$ proving that $\gamma \in \bigcup_{j \in \{1, \ldots, k+1\}} \Gamma^j(k_1, \ldots, k_{n+1})$. Hence

$$\Gamma(k_1, \ldots, k_{n+1}) = \bigcup_{j \in \{k_1, \ldots, k_{n+1}\}} \Gamma^j(k_1, \ldots, k_{n+1})$$

Let $i, j \in \{1, \ldots, k_{n+1}\}$ with $i \neq j$ then if $\gamma \in \Gamma^i(k_1, \ldots, k_{n+1}) \bigcap \Gamma^j(k_1, \ldots, k_{n+1})$ we would have the contradiction $\gamma(n+1) = i \neq j = \gamma(n+1)$ hence $\Gamma^i(k_1, \ldots, k_{n+1}) \bigcap \Gamma^j(k_1, \ldots, k_{n+1}) = \varnothing$. So

$$\Gamma(k_1, \ldots, k_{n+1}) \text{ is the disjoint union of } \Gamma^j(k_1, \ldots, k_{n+1}) \; j \in \{1, \ldots, k_{n+1}\} \tag{13.107}$$

Further if we define

$$\alpha_j: \Gamma^j(k_1, \ldots, k_{n+1}) \rightarrow \Gamma(k_1, \ldots, k_n) \text{ by } \alpha(\gamma) = \gamma_{|\{1, \ldots, n\}} \text{ [hence } \alpha(\gamma)(i) = \gamma(i) \, \forall i \in \{1, \ldots, n\}]$$

then we have:

injectivity. If $\alpha_j(\gamma) = \alpha_j(\delta)$ then $\forall i \in \{1, \ldots, n\}$ $\gamma(i) = \alpha_j(\gamma)(i) = \alpha_j(\delta)(i) = \delta(i)$ and $\gamma(n+1) = j = \delta(n+1)$, proving that $\gamma = \delta$.

surjectivity. If $\gamma \in \Gamma(k_1, \ldots, k_n)$ define then $\gamma' \in \Gamma^j(k_1, \ldots, k_{n+1})$ by $\gamma'(i) = \left\{ \begin{array}{l} j \text{ if } i = n+1 \\ \gamma(i) \text{ if } i \in \{1, \ldots, n+1\} \end{array} \right.$ then $\alpha_j(\gamma') = \gamma$.

so we have that

$$\alpha_j: \Gamma^j(k_1, \ldots, k_{n+1}) \rightarrow \Gamma(k_1, \ldots, k_n) \text{ is a bijection with } \forall i \in \{1, \ldots, n\} \, \alpha_j(\gamma)(i) = \gamma(i) \tag{13.108}$$

So we have by [theorem: 12.3] and [eq: 13.107] that

$$\begin{array}{rc} \bigoplus_{\gamma \in \Gamma(k_1, \ldots, k_{n+1})} (W_{1,\gamma(1)} \otimes \cdots \otimes W_{n,\gamma(n)}) \otimes W_{n+1,\gamma(n+1)} & = \\ \bigotimes_{i \in \{1, \ldots, k_{n+1}\}} \left(\bigoplus_{\gamma \in \Gamma^j(k_1, \ldots, k_{n+1})} (W_{1,\gamma(1)} \otimes \cdots \otimes W_{n,\gamma(n)}) \otimes W_{n+1,\gamma(n+1)} \right) & \underset{\gamma(n+1) = j}{=} \\ \bigotimes_{i \in \{1, \ldots, k_{n+1}\}} \left(\bigoplus_{\gamma \in \Gamma^j(k_1, \ldots, k_{n+1})} (W_{1,\gamma(1)} \otimes \cdots \otimes W_{n,\gamma(n)}) \otimes W_{n+1,j} \right) & \underset{[\text{eq: } 13.108]}{=} \\ \bigotimes_{i \in \{1, \ldots, k_{n+1}\}} \left(\bigoplus_{\gamma \in \Gamma^j(k_1, \ldots, k_{n+1})} (W_{1,\alpha_j(\gamma)(1)} \otimes \cdots \otimes W_{n,\alpha_j(\gamma)(n)}) \otimes W_{n+1,j} \right) & \underset{[\text{theorem: } 12.13]}{=} \\ \bigotimes_{i \in \{1, \ldots, k_{n+1}\}} \left(\bigoplus_{\gamma \in \Gamma(k_1, \ldots, k_n)} (W_{1,\gamma(1)} \otimes \cdots \otimes W_{n,\gamma(n)}) \otimes W_{n+1,j} \right) & \underset{[\text{eq: } 13.106]}{=} \\ V_1 \otimes \cdots \otimes V_{n+1} & \end{array}$$

or summarized

$$V_1 \otimes \cdots \otimes V_{n+1} = \bigoplus_{\gamma \in \Gamma(k_1, \ldots, k_{n+1})} (W_{1,\gamma(1)} \otimes \cdots \otimes W_{n,\gamma(n)}) \otimes W_{n+1,\gamma(n+1)} \tag{13.109}$$

Applying [theorem: 13.26] we have then that

$$(W_{1,\gamma(1)} \otimes \cdots \otimes W_{n,\gamma(n)}) \otimes W_{n+1,\gamma(n+1)} = W_{1,\gamma(1)} \otimes \cdots \otimes W_{n+1,\gamma(n+1)}$$

a tensor product of $W_{1,\gamma(1)}, \ldots, W_{n+1,\gamma(n+2)}$. Hence we have

$$V_1 \otimes \cdots \otimes V_{n+1} = \bigoplus_{\gamma \in \Gamma(k_1, \ldots, k_{n+1})} W_{1,\gamma(1)} \otimes \cdots \otimes W_{n+1,\gamma(n+1)}$$

proving that $n+1 \in S$. □

The above theorem can be used to construct a basis of a tensor product out of the bases of the factors. First we consider the one dimensional case.

Lemma 13.39. *Let $n \in \mathbb{N}$, $\{V_i\}_{i \in \{1, \ldots, m\}}$ a family of one-dimensional vector spaces over a field F and $\forall i \in \{1, \ldots, n\}$ $\{e_i\}$ a basis for V_i [so that $e_i \neq 0$ and $V_i = \operatorname{span}(\{e_i\})$] then for every tensor product $\langle V_1 \otimes \cdots \otimes V_n. \otimes \rangle$ of $V_1, \ldots, V_n$ we have*

$$V_1 \otimes \cdots \otimes V_n = \operatorname{span}(\{\otimes(e_1, \ldots, e_n)\}) \underset{\text{notation}}{=} \operatorname{span}(\{e_1 \otimes \cdots \otimes e_n\})$$

Proof. Define

$$P = \operatorname{span}(\{\otimes(e_1, \ldots, e_n)\})$$

and

$$\nu: \prod_{i \in \{1, \ldots, n\}} V_i = \prod_{i \in \{1, \ldots, n\}} \operatorname{span}(\{e_i\}) \rightarrow P \text{ by } \nu(v_1, \ldots, v_n) = \left(\prod_{i \in \{1, \ldots, n\}} \lambda_i\right) \cdot \otimes(e_1, \ldots, e_n)$$

whete $\forall i \in \{1, \ldots, n\}$ $v_i = \lambda_i \cdot e_i$ [the unique expansion of a vector in a given basis]. Then we proof that

$$\langle P, \nu \rangle \text{ is a tensor product of } V_1, \ldots, V_n \tag{13.110}$$

Proof.

1. By [theorem: 11.105] $P = \operatorname{span}(\{\otimes(e_1, \ldots, e_n)\})$ is a vector space.
2. Let $i \in \{1, \ldots, n\}$ $u, v \in V_i$, $\alpha \in F$ then for $(v_1, \ldots, v_{i-1}, u + \alpha \cdot v_{i+1}, \ldots, v_n)$ we have that there exists a unique $\{\lambda_j\}_{j \in \{1, \ldots, n\}} \subseteq F$ such that $v_i = \lambda_i \cdot e_i$ and a $\alpha_u, \alpha_\nu \in F$ such that $u = \alpha_u \cdot e_i$ and $v = \alpha_v \cdot e_i$. Define then

$$\{\beta_j\}_{j \in \{1, \ldots, n\}} \subseteq F \text{ by } \beta_j = \begin{cases} \alpha_u + \alpha \cdot \alpha_v \text{ if } j = i \\ \lambda_j \text{ if } j \in \{1, \ldots, n\} \setminus \{i\} \end{cases}$$

$$\{\zeta_j\}_{j \in \{1, \ldots, n\}} \subseteq F \text{ by } \zeta_j = \begin{cases} \alpha_u \text{ if } j = i \\ \lambda_j \text{ if } j \in \{1, \ldots, n\} \setminus \{i\} \end{cases}$$

$$\{\gamma_j\}_{j \in \{1, \ldots, n\}} \subseteq F \text{ by } \gamma_j = \begin{cases} \alpha_u \text{ if } j = i \\ \lambda_j \text{ if } j \in \{1, \ldots, n\} \setminus \{i\} \end{cases}$$

so that $\forall j \in \{1, \ldots, n\}$ we have

$$\beta_j \cdot e_j = \begin{cases} \alpha_u \cdot e_i + \alpha \cdot \alpha_v \cdot e_i \text{ if } j = i \\ \lambda_j \cdot e_j \text{ if } j \in \{1, \ldots, n\} \setminus \{i\} \end{cases} = \begin{cases} u + \alpha \cdot v \text{ if } j = i \\ v_j \text{ if } j \in \{1, \ldots, n\} \setminus \{i\} \end{cases}$$

$$\zeta_j \cdot e_j = \begin{cases} \alpha_u \cdot e_i \text{ if } j = i \\ \lambda_j \cdot e_j \text{ if } j \in \{1, \ldots, n\} \setminus \{i\} \end{cases} = \begin{cases} u \text{ if } j = i \\ v_j \text{ if } j \in \{1, \ldots, n\} \setminus \{i\} \end{cases}$$

$$\gamma_j \cdot e_j = \begin{cases} \alpha_v \cdot e_i \text{ if } j = i \\ \lambda_j \cdot e_j \text{ if } j \in \{1, \ldots, n\} \setminus \{i\} \end{cases} = \begin{cases} v \text{ if } j = i \\ v_j \text{ if } j \in \{1, \ldots, n\} \setminus \{i\} \end{cases}$$

Hence

$$\begin{aligned}
\nu(v_1,\ldots,v_{i-1},u+\alpha\cdot v,v_{i+1},\ldots,v_n) &= \\
\left(\prod_{j\in\{1,\ldots,n\}}\beta_j\right)\cdot\otimes(e_1,\ldots,e_n) &\underset{[\text{theorem: } 11.43]}{=} \\
\left(\left(\prod_{j\in[1,\ldots,n]\setminus\{i\}}\beta_j\right)\cdot\left(\prod_{j\in[1,\ldots,n]\setminus\{i\}}\beta_j\right)\right)\cdot\otimes(e_1,\ldots,e_n) &\underset{[\text{theorem: } 11.34]}{=} \\
\left(\left(\prod_{j\in[1,\ldots,n]\setminus\{i\}}\beta_j\right)\cdot\beta_i\right)\cdot\otimes(e_1,\ldots,e_n) &= \\
\left(\left(\prod_{j\in[1,\ldots,n]\setminus\{i\}}\lambda_j\right)\cdot(\alpha_u+\alpha\cdot\alpha_v)\right)\cdot\otimes(e_1,\ldots,e_n) &= \\
\left(\left(\prod_{j\in[1,\ldots,n]\setminus\{i\}}\lambda_j\right)\cdot\alpha_u\right)\cdot\otimes(e_1,\ldots,e_n)+\alpha\cdot\left(\left(\prod_{j\in[1,\ldots,n]\setminus\{i\}}\lambda_j\right)\cdot\alpha_v\right)\cdot\otimes(e_1,\ldots,e_n) &= \\
\left(\left(\prod_{j\in[1,\ldots,n]\setminus\{i\}}\zeta_j\right)\cdot\zeta_i\right)\cdot\otimes(e_1,\ldots,e_n)+\alpha\cdot\left(\left(\prod_{j\in[1,\ldots,n]\setminus\{i\}}\gamma_j\right)\cdot\gamma_i\right)\cdot\otimes(e_1,\ldots,e_n) &\underset{[\text{theorem: } 11.34]}{=} \\
\left(\left(\prod_{j\in[1,\ldots,n]\setminus\{i\}}\zeta_j\right)\cdot\left(\prod_{j\in\{i\}}\zeta_i\right)\cdot\otimes(e_1,\ldots,e_n)+\alpha\cdot\left(\left(\prod_{j\in[1,\ldots,n]\setminus\{i\}}\gamma_j\right)\cdot\left(\prod_{j\in\{i\}}\gamma_j\right)\right)\cdot\otimes(e_1,\ldots,e_n) &\underset{[\text{theorem: } 11.43]}{=} \\
\left(\prod_{j\in[1,\ldots,n]}\zeta_j\right)\cdot\otimes(e_1,\ldots,e_n)+\alpha\cdot\left(\prod_{j\in[1,\ldots,n]}\gamma_j\right)\cdot\otimes(e_1,\ldots,e_n) &= \\
\nu(v_1,\ldots,v_{i-1},u,v_{i+1},\ldots,v_n)+\alpha\cdot(v_1,\ldots,v_{i-1},v,v_{i+1},\ldots,v_n)
\end{aligned}$$

proving that

$$\nu\in\mathrm{Hom}(V_1,\ldots,V_n;P)$$

3. As by definition of ν we have that $\nu\left(\prod_{i\in\{1,\ldots,n\}}V_i\right)\subseteq P$ it follows by [theorems: 11.104, 11.106] that

$$\mathrm{span}\left(\nu\left(\prod_{i\in\{1,\ldots,n\}}V_i\right)\right)\subseteq P \tag{13.111}$$

Further if $v\in P$ then there exists a $\lambda\in F$ such that $v=\lambda\cdot\otimes(e_1,\ldots,e_n)$, take then $\{\lambda_i\}_{i\in\{1,\ldots,n\}}\subseteq F$ by $\lambda_i=\left\{\begin{smallmatrix}\lambda \text{ if } i=1\\ 1 \text{ if } i\in\{2,\ldots,n\}\end{smallmatrix}\right.$ then we have for $(\lambda_1\cdot e_1,\ldots,\lambda_n\cdot e_n)$ that

$$\begin{aligned}
\nu(\lambda_1\cdot e_1,\ldots,\lambda_n\cdot e_n) &= \left(\prod_{i\in\{1,\ldots,n\}}\lambda_i\right)\cdot\otimes(e_1,\ldots,e_n)\\
&\underset{[\text{theorem: } 11.43]}{=} \left(\left(\prod_{i\in\{2,\ldots,n\}}\lambda_i\right)\cdot\left(\prod_{i\in\{1\}}\lambda_i\right)\right)\cdot\otimes(e_1,\ldots,e_n)\\
&\underset{[\text{theorem: } 11.34]}{=} \left(\left(\prod_{i\in\{2,\ldots,n\}}1\right)\cdot\lambda\right)\cdot\otimes(e_1,\ldots,e_n)\\
&\underset{[\text{theorem: } 11.37]}{=} (1\cdot\lambda)\cdot\otimes(e_1,\ldots,e_n)\\
&= \lambda\cdot\otimes(e_1,\ldots,e_n)\\
&= v
\end{aligned}$$

proving that $P\subseteq\nu\left(P=\prod_{i\in\{1,\ldots,n\}}V_i\right)\subseteq\mathrm{span}\left(\nu\left(P=\prod_{i\in\{1,\ldots,n\}}V_i\right)\right)$. Combining this with [eq: 13.111] results in

$$P=\mathrm{span}\left(\nu\left(\prod_{i\in\{1,\ldots,n\}}V_i\right)\right)$$

4. Let $\varphi \in \operatorname{Hom}(V_1, \ldots, V_n; P) = \operatorname{Hom}(\operatorname{span}(\{e_1\}), \ldots, \operatorname{span}(\{e_n\}); U)$. Define

$$h: P = \operatorname{span}(\{\otimes(e_1, \ldots, e_n)\}) \rightarrow U \text{ by } h(v) = \lambda \cdot \varphi(e_1, \ldots, e_n) \text{ where } v = \lambda \cdot \otimes(e_1, \ldots, e_n)$$

If $u, v \in \operatorname{span}(\{\otimes(e_1,\ldots,e_n)\})$ and $\alpha \in F$ then there exists a $\lambda_u, \lambda_v \in F$ such that $u = \lambda_u \cdot \otimes(e_1,\ldots,e_n)$, $v = \lambda_v \cdot \otimes(e_1,\ldots,e_n)$ then we have

$$\begin{aligned} h(u + \alpha \cdot v) &= h(\lambda_u \cdot \otimes(e_1, \ldots, e_n) + \alpha \cdot \lambda_v \otimes (e_1, \ldots, e_n)) \\ &= h((\lambda_u + \alpha \cdot \lambda_v) \cdot \otimes(e_1, \ldots, e_n)) \\ &= (\lambda_u + \alpha \cdot \lambda_v) \cdot \varphi(e_1, \ldots, e_n) \\ &= \lambda_u \cdot \varphi(e_1, \ldots, e_n) + \alpha \cdot \lambda_v \cdot \varphi(e_1, \ldots, e_n) \\ &= h(u) + \alpha \cdot h(v) \end{aligned}$$

proving that

$$h \in \operatorname{Hom}(P, U)$$

Let $(v_1, \ldots, v_n) \in \prod_{i \in \{1, \ldots, n\}} V_i = \prod_{i \in \{1, \ldots, n\}} \operatorname{span}(\{e_i\})$ then there exists a $\{\lambda_i\}_{i \in \{1, \ldots, n\}} \subseteq F$ such that $(v_1, \ldots, v_n) = (\lambda_1 \cdot e_1, \ldots, \lambda_n \cdot e_m)$. So

$$\begin{aligned} (h \circ \nu)(v_1, \ldots, v_n) &= h(\nu(v_1, \ldots v_n)) \\ &= h(\nu(\lambda_1 \cdot e_1, \ldots \lambda_n \cdot e_n)) \\ &= h\left(\left(\prod_{i \in \{1, \ldots, n\}} \lambda_i\right) \cdot \otimes(e_1, \ldots, e_n)\right) \\ &= \left(\prod_{i \in \{1, \ldots, n\}} \lambda_i\right) \cdot \varphi(e_1, \ldots, e_n) \\ &\underset{[\text{theorem: } 11.254]}{=} \varphi(\lambda_1 \cdot e_1, \ldots, \lambda_n \cdot e_n) \\ &= \varphi(v_1, \ldots, v_n) \end{aligned}$$

proving that

$$h \circ \nu = \varphi$$

Finally (1),(2),(3) and (4) completes the proof of [eq: 13.110]. □

As $\langle V_1 \otimes \cdots \otimes V_n, \otimes \rangle$ is a tensor product we have that

$$V_1 \otimes \cdots \otimes V_n = \operatorname{span}\left(\otimes\left(\prod_{i \in \{1, \ldots, n\}} V_i\right)\right)$$

Hence as $\otimes(\{e_1, \ldots, e_n\}) \subseteq \otimes\left(\prod_{i \in \{1, \ldots, n\}} V_i\right)$ it follows from [theorem: 11.104] that

$$P = \operatorname{span}(\{\otimes(e_1, \ldots, e_n)\}) \subseteq \operatorname{span}\left(\otimes\left(\prod_{i \in \{1, \ldots, n\}} V_i\right)\right) = V_1 \otimes \cdots \otimes V_n \tag{13.112}$$

Further by [theorem: 13.24] and the fact that $\forall i \in \{1, \ldots, n\}$ $e_i \neq 0$ it follows that $\otimes(e_1, \ldots, e_n) \neq 0$ hence

$$P = \operatorname{span}(\{\otimes(e_1, \ldots, e_n)\}) \text{ is one-dimensional}$$

As $\langle V_1 \otimes, \ldots, \otimes V_n, \otimes \rangle$ is also a tensor product of $V_1, \ldots, V_n$ we have by [theorem: 13.15] that $V_1 \otimes \cdots \otimes V_n$ is linear isomorphic with P, hence by [theorem: 11.191] $V_1 \otimes \cdots \otimes V_n$ is also one dimensional so that there exists a $v_0 \in V_1 \otimes \cdots \otimes V_n$ such that

$$V_1 \otimes \cdots \otimes V_n = \operatorname{span}(v_0)$$

As $\otimes(e_1, \ldots, e_n) \in \otimes\left(\prod_{i \in \{1, \ldots, n\}} V_i\right) \subseteq \operatorname{span}\left(\otimes\left(\prod_{i \in \{1, \ldots, n\}} V_i\right)\right) = V_1 \otimes \cdots \otimes V_n$ there exist a $\lambda \in F$ such that $\otimes(e_1, \ldots, e_n) = \lambda \cdot v_0$. As $\otimes(e_1, \ldots, e_n) \neq 0$ we must have that $\lambda \neq 0$ hence

$$v_0 = \lambda^{-1} \cdot \otimes(e_1, \ldots, e_n)$$

Let $v \in V_1 \otimes \cdots \otimes V_n$ then there exists a $\gamma \in F$ such that $v = \gamma \cdot v_0 = (\gamma \cdot \lambda^{-1}) \cdot \otimes(e_1, \ldots, e_n)$, proving that $v \in \operatorname{span}(\{\otimes(e_1, \ldots, e_n)\})$. Hence $V_1 \otimes \cdots \otimes V_n \subseteq \operatorname{span}(\{\otimes(e_1, \ldots, e_n)\})$ so combining this with [eq: 13.112] proves that

$$V_1 \otimes \cdots \otimes V_n = \operatorname{span}(\{\otimes(v_1, \ldots, v_n)\})$$ □

Theorem 13.40. *Let $n \in \mathbb{N}$, $\{V_i\}_{i \in \{1,\ldots,n\}}$ a finite family of finite vector spaces over a field F with $\forall i \in \{1,\ldots,n\}$ bases $\{e_{i,1},\ldots,e_{i,k_i}\}$ of V_i [where $k_i = \dim(V_i)$] then*

$$\{\otimes(e_{1,\gamma(1)},\ldots,e_{n,\gamma(n)}) | \gamma \in \Gamma(k_1,\ldots,k_n)\} = \{e_{1,\gamma(1)} \otimes \cdots \otimes e_{n,\gamma(n)} | \gamma \in \Gamma(k_1,\ldots,k_n)\}$$

is a basis of $V_1 \otimes \cdots \otimes V_n$. Hence by [theorem: 13.30] we have that

$$\dim(V_1 \otimes \cdots \otimes V_n) = \prod_{i \in \{1,\ldots,n\}} k_i$$

Proof. Let $i \in \{1,\ldots,n\}$ and $v \in V_i$ then by [theorem: 11.137] there exists a **unique** $\{\alpha_j\}_{j \in \{1,\ldots,k_i\}} \subseteq F$ such that $v = \sum_{j \in \{1,\ldots,k_i\}} \alpha_j \cdot e_{i,j}$. As $\forall j \in \{1,\ldots,k_i\}$ $\alpha_j \cdot e_j \in \operatorname{span}(\{e_j\})$ it follows from the definition of a direct sum [see definition: 12.5] that

$$V_i = \bigoplus_{j \in \{1,\ldots,k_i\}} \operatorname{span}(\{e_{i,j}\})$$

So we have that

$$\begin{aligned} V_1 \otimes \cdots \otimes V_n &\underset{\text{[theorem: 13.38]}}{=} \bigoplus_{\gamma \in \Gamma(k_1,\ldots,k_n)} (\operatorname{span}(\{e_{i,\gamma(i)}\}) \otimes \cdots \otimes \operatorname{span}(\{e_{n,\gamma(n)}\})) \\ &\underset{\text{[theorem: 13.39]}}{=} \bigoplus_{\gamma \in \Gamma(k_1,\ldots,k_n)} \operatorname{span}(\otimes(e_1,\ldots,e_n)) \end{aligned}$$

Let $v \in V_1 \otimes \cdots \otimes V_n$ then by the definition of a direct sum [see definition: 12.5] there exist a **unique** $\{\alpha_i\}_{i \in \Gamma(k_1,\ldots,k_n)}$ such that $v = \sum_{\Gamma(k_1,\ldots,k_n)} \alpha_i \cdot \otimes(e_1,\ldots,e_n)$. By [theorem: 11.136] it follows then that

$$\{\otimes(e_{1,\gamma(1)},\ldots,e_{n,\gamma(n)}) | \gamma \in \Gamma(k_1,\ldots,k_n)\} \text{ is a basis for } V_1 \otimes \cdots \otimes V_n$$ □

Once we have a basis of a vector space we can consider the expansion of its elements in the basis, this leads to the idea of the components of a tensor product.

Definition 13.41. *Let $n \in \mathbb{N}$, $\{V_i\}_{i \in \{1,\ldots,n\}}$ a finite family of finite vector spaces over a field F with $\forall i \in \{1,\ldots,n\}$ bases $\{e_{i,1},\ldots,e_{i,k_i}\}$ of V_i [where $k_i = \dim(V_i)$] then by [theorem: 13.40] we have that basis*

$$\{e_{1,\gamma(1)} \otimes \cdots \otimes e_{n,\gamma(n)} | \gamma \in \Gamma(k_1,\ldots,k_n)\}$$

*Let $T \in V_1 \otimes \cdots \otimes V_n$ then there exists **unique** $\{T_\gamma\}_{\gamma \in \Gamma(k_1,\ldots,k_n)} \subseteq F$ such that*

$$T = \sum_{\gamma \in \Gamma(k_1,\ldots,k_n)} T_\gamma \cdot (e_{1,\gamma(1)} \otimes \cdots \otimes e_{n,\gamma(n)})$$

*The $\prod_{i \in \{1,\ldots,n\}} k_i$ T_γ are called the **components** of T. As $\Gamma(k_1,\ldots,k_n) = \prod_{i \in \{1,\ldots,n\}} \{1,\ldots,k_i\}$ we can write a $\gamma \in \Gamma(k_1,\ldots,k_n)$ as $\gamma = (\gamma_1,\ldots,\gamma_n)$. Hence we also use the notation*

$$\{T_{\gamma_1,\ldots,\gamma_n}\}_{\gamma \in \Gamma(k_1,\ldots,k_n)}$$

for the components so that

$$T = \sum_{\gamma \in \Gamma(k_1,\ldots,k_n)} T_{\gamma_1,\ldots,\gamma_n}$$

We introduct now the concept of p-q tensors.

Definition 13.42. *Let V be a vector space over a field, $p, q \in \mathbb{N}_0$ then a p-contravariant, q-covariant tensor product noted as V^p_q is defined by*

1. *If $p \neq 0 \vee q \neq 0$ then*

$$V^p_q = V_1 \otimes \cdots \otimes V_{p+q} \text{ where } \forall i \in \{1,\ldots p+q\}\ V_i = \begin{cases} V \text{ if } i \in \{1,\ldots,p\} \\ V^* \text{ if } i \in \{p+1,\ldots,p+q\} \end{cases}$$

   where $V^* = \operatorname{Hom}(V,F)$ is the dual space of V [see definition: 11.176].

2. *If $p = q = 0$ then we defined*

$$V^p_q = F$$

Another notation used is

$$V_q^p = \underbrace{V \otimes \cdots \otimes V}_{p} \otimes \underbrace{V^* \otimes \cdots \otimes V^*}_{q}$$

elements of V_q^p *are called* p*-contravariant,* q*-covariant tensors or in short* p*-*q*-tensors.*

Chapter 14

Topology

14.1 Topological spaces

Definition 14.1. (Topological Space) *Let X be a set then $\mathcal{T} \subseteq \mathcal{P}(X)$ is a* ***topology*** *on X if*

1. $X \in \mathcal{T}$
2. $\varnothing \in \mathcal{T}$
3. $\forall U, V \in \mathcal{T}$ *we have* $U \bigcap V \in \mathcal{T}$
4. *If* $\{U_i\}_{i \in I} \subseteq \mathcal{T}$ *is a family of sets in* $\mathcal{T}$ *then* $\bigcup_{i \in I} U_i \in \mathcal{T}$

Elements of $\mathcal{T}$ are called open sets and the pair $\langle X, \mathcal{T} \rangle$ is called a topological space.

The most trivial (and least usable) example of a topological space is $\langle \varnothing, \{\varnothing\} \rangle$.

Example 14.2. $\langle \varnothing, \{\varnothing\} \rangle$ is a topological space. Further the only topology on $\varnothing$ is $\{\varnothing\}$

Proof. As $\varnothing \subseteq \varnothing$ we have that $\varnothing \in \mathcal{P}(\varnothing)$ hence $\mathcal{T} = \{\varnothing\} \subseteq \mathcal{P}(\varnothing)$. Further

1. $X = \varnothing \in \{\varnothing\} = \mathcal{T}$
2. $\varnothing \in \{\varnothing\} = \mathcal{T}$
3. If $U, V \in \mathcal{T} = \{\varnothing\}$ then $U = \varnothing = V$ so that $U \bigcap V = \varnothing \in \{\varnothing\} = \mathcal{T}$
4. If $\{U_i\}_{i \in I} \subseteq \mathcal{T} = \{\varnothing\}$ then we have $\forall i \in I$ that $U_i = \varnothing$ hence $\bigcup_{i \in I} U_i = \varnothing \in \{\varnothing\} = \mathcal{T}$

Further if $\mathcal{T}$ is a topology on $\varnothing$ we have $\varnothing \in \mathcal{T}$ so that $\{\varnothing\} \subseteq \mathcal{T}$ and if $U \in \mathcal{T}$ then $U \subseteq \varnothing \Rightarrow U = \varnothing$ so that $\mathcal{T} \subseteq \{\varnothing\}$. Hence $\mathcal{T} = \{\varnothing\}$. □

Example 14.3. If X is a set then $\langle X, \{X, \varnothing\} \rangle$ is a topological space

Proof. For $\mathcal{T} = \{X, \varnothing\}$ we have:

1. $X \in \{X, \varnothing\} = \mathcal{T}$
2. $\varnothing \in \{X, \varnothing\} = \mathcal{T}$
3. If $U, V \in \mathcal{T} = \{X, \{0\}\}$ then we have either:

   $\boldsymbol{U = X \wedge V = X}$**.** Then $U \bigcap V = X \in \{\varnothing, X\} = \mathcal{T}$

   $\boldsymbol{U = \varnothing \wedge V = X}$**.** Then $U \bigcap V = \varnothing \in \{\varnothing, X\} = \mathcal{T}$

   $\boldsymbol{U = X \wedge V = \varnothing}$**.** Then $U \bigcap V = \varnothing \in \{\varnothing, X\} = \mathcal{T}$

   $\boldsymbol{U = \varnothing \wedge V = \varnothing}$**.** Then $U \bigcap V = \varnothing \in \{\varnothing, X\} = \mathcal{T}$

4. If $\{U_i\}_{i \in I} \subseteq \mathcal{T}$ then we have either:

   $\boldsymbol{\forall i \in I}$ **we have** $\boldsymbol{U_i = \varnothing}$**.** Then $\bigcup_{i \in I} U_i = \varnothing \in \{\varnothing, X\} = \mathcal{T}$

   $\boldsymbol{\exists i \in I}$ **with** $\boldsymbol{U_i = X}$**.** Then $\bigcup_{i \in I} U_i = X \in \{\varnothing, X\} = \mathcal{T}$ □

Example 14.4. (Discrete Topology) If X is a set then $\langle X, \mathcal{P}(X) \rangle$ is a topological space. $\mathcal{P}(X)$ is called the **discrete topology** on X.

Proof. For $\mathcal{T} \in \mathcal{P}(X)$ we have:

1. As $X \subseteq X$ we have $X \in \mathcal{P}(X) = \mathcal{T}$
2. As $\varnothing \subseteq X$ we have $\varnothing \in \mathcal{P}(X) = \mathcal{T}$

3. If $U, V \subseteq \mathcal{P}(X)$ then $U, V \subseteq X$, hence by [theorem: 1.25] $U \bigcap V \subseteq X$ so that $U \bigcap V \in \mathcal{P}(X)$
4. If $\{U_i\}_{i \in I} \subseteq \mathcal{T} = \mathcal{P}(X)$ then $\forall i \in I$ $U_i \in \mathcal{P}(X) \Rightarrow U_i \subseteq X$, hence by [theorem: 2.125] $\bigcup_{i \in I} U_i \subseteq X$ proving that $\bigcup_{i \in I} U_i \in \mathcal{P}(X) = \mathcal{T}$ $\square$

The last two examples shows that given a set X there could exist two different topologies topologies on X. In the examples both $\{\varnothing, X\}$ and $\mathcal{P}(X)$ are topologies on X. So if we say that a set in X is open we have to specify which topology is used.

Definition 14.5. *Let X be a set and $\mathcal{T}_1$, $\mathcal{T}_2$ two topologies on X then we say that $\mathcal{T}_1$ **is finer then** $\mathcal{T}_2$ iff $\mathcal{T}_2 \subseteq \mathcal{T}_1$. In other words $\mathcal{T}_1$ is finer then $\mathcal{T}_2$ if every open set in $\mathcal{T}_2$ is also a open set in $\mathcal{T}_2$. Clearly if $\mathcal{T}_1$ is finer then $\mathcal{T}_2$ and $\mathcal{T}_2$ is finer then $\mathcal{T}_1$ then $\mathcal{T}_1 = \mathcal{T}_2$.*

Using mathematical induction it follows that every **finite** intersection of open sets is open.

Theorem 14.6. *Let $\langle X, \mathcal{T} \rangle$ be a topological space, $I \neq \varnothing$ a finite set and $\{U_i\}_{i \in I} \subseteq \mathcal{T}$ a finite family of open sets then $\bigcap_{i \in I} U_i \in \mathcal{T}$ [or in other words $\bigcap_{i \in I} U_i$ is open].*

Proof. We use induction to prove this, so define

$$\mathcal{S} = \left\{ n \in \mathbb{N} | \text{If } I \text{ is finite with } \operatorname{card}(I) = n \text{ and } \{U_i\}_{i \in I} \subseteq \mathcal{T} \text{ then } \bigcap_{i \in I} U_i \in \mathcal{T} \right\}$$

then we have:

$1 \in S$. Then $I = \{i\}$ so that $\bigcap_{j \in \{i\}} U_j = U_i \in \mathcal{T}$ proving that $1 \in S$.

$n \in S \Rightarrow n + 1 \in S$. Let I be a finite set with $\operatorname{card}(I) = n + 1$ then there exists a bijection

$$\beta: \{1, \ldots, n_{+1}\} \rightarrow I$$

so that by [theorem: 2.89]

$$\beta_{|\{1, \ldots, n\}}: \{1, \ldots, n\} \rightarrow I \setminus \{\beta(n+1)\} \text{ is a bijection}$$

or $\operatorname{card}(I \setminus \{n+1\}) = n$. As $n \in S$ it follows that $\bigcap_{i \in I \setminus \{n+1\}} U_i \in \mathcal{T}$, hence, as $U_{n+1} \in \mathcal{T}$, we have

$$\bigcap_{i \in I} U_i = \left(\bigcap_{i \in I \setminus \{n+1\}} U_i \right) \bigcap U_{n+1} \in \mathcal{T}$$

proving that $n + 1 \in S$. $\square$

Theorem 14.7. *Let $\langle X, \mathcal{T} \rangle$ be a topological space then we have for $U \subseteq X$*

$$U \in \mathcal{T} \Leftrightarrow \forall x \in U \text{ there exist a } V \in \mathcal{T} \text{ such that } x \in V \subseteq U$$

Proof.

$\Rightarrow$. If $U \in \mathcal{T}$ then we have either:

$U = \varnothing$. Then $\forall x \in U = \varnothing$ we have vacuously that there exist a $V \in \mathcal{T}$ such that $x \in V \subseteq U$.

$U \neq \varnothing$. Then $\forall x \in U$ we have $x \in U \subseteq U$ where $U \in \mathcal{T}$.

$\Leftarrow$. By the hypothesis we have that $\forall x \in X$ there exist a $V \in \mathcal{T}$ such that $x \in V \subseteq U$ this defines by [theorem: 3.104] a family $\{V_x\}_{x \in U} \subseteq \mathcal{T}$ such that $\forall x \in U$ $x \in V_x \subseteq U$. By [theorem: 2.125] it follows that $\bigcup_{x \in U} V_x \subseteq U$. Further if $x \in U$ then $x \in V_x \subseteq \bigcup_{\in U} V_x$, proving that $U \subseteq \bigcup_{x \in U} V_x$. Hence

$$U = \bigcup_{x \in U} V_x \in \mathcal{T}$$ $\square$

Given a topology on a set then we can induce a topology on a subset of this set.

Definition 14.8. (Subspace Topology) *Let $\langle X, \mathcal{T} \rangle$ be a topological space and $Y \subseteq X$ then*

$$\langle Y, \mathcal{T}_{|Y} \rangle \text{ where } \mathcal{T}_{|Y} = \left\{ U \bigcap Y | U \in \mathcal{T} \right\}$$

*is a topology called the **subspace topology on Y induced by $\mathcal{T}$.***

Note 14.9. *$\mathcal{T}_{|X} = \{U \bigcap X | U \in \mathcal{T}\} \underset{U \in \mathcal{T} \Rightarrow U \subseteq X}{=} \{U | U \in \mathcal{T}\} = \mathcal{T}$. In other words the subspace topology on X induced by $\mathcal{T}$ is the topology of X itself.*

Proof. Of course we must prove that $\mathcal{T}_{|Y}$ is indeed a topology on Y.

1. As $Y \subseteq X$ and $X \in \mathcal{T}$ it follows that $Y = X \bigcap Y \in \mathcal{T}_{|Y}$.
2. $\varnothing = \varnothing \bigcap Y \in \mathcal{T}_{|Y}$ because $\varnothing \in \mathcal{T}$.
3. If $V_1, V_2 \in \mathcal{T}_{|Y}$ then there exists $U_1, U_2 \in \mathcal{T}$ such that $V_1 = U_1 \bigcap Y$ and $V_2 = U_2 \bigcap Y$. Then we have
$$V_1 \bigcap V_2 = (U_1 \bigcap Y) \bigcap (U_2 \bigcap Y) = (U_1 \bigcap U_2) \bigcap Y \in \mathcal{T}_Y \text{ because } U_1 \bigcap U_2 \in \mathcal{T}$$
4. Let $\{V_i\}_{i \in I} \subseteq \mathcal{T}$ then there exists a $\{U_i\}_{i \in I}$ such that $\forall i \in I\ V_i = U_i \bigcap Y$, hence
$$\bigcup_{i \in I} V_i = \bigcup_{i \in I} (U_i \bigcap Y) \underset{[\text{theorem: } 2.128}{=} \left(\bigcup_{i \in I} U_i\right) \bigcap Y \in \mathcal{T} \text{ because } \bigcup_{i \in I} U_i \in \mathcal{T}$$
□

The subspace topology of a subspace topology is again a subspace topology.

Theorem 14.10. *Let $\langle X, \mathcal{T} \rangle$ be a topological space, $Z \subseteq Y \subseteq X$ then $\mathcal{T}_{|Z} = (\mathcal{T}_{|Y})_{|Z}$*

Proof. First we have

$$\begin{array}{rcl} U \in (\mathcal{T}_{|Y})_{|Z} & \Rightarrow & \exists V \in \mathcal{T}_{|Y} \text{ such that } U = V \bigcap Z \\ & \Rightarrow & \exists W \in \mathcal{T} \text{ such that } V = W \bigcap Y \\ & \Rightarrow & U = (W \bigcap Y) \bigcap Z = W \bigcap (Z \bigcap Y) \\ & \underset{Z \subseteq Y}{\Rightarrow} & U = W \bigcap Z \\ & \Rightarrow & U \in \mathcal{T}_{|Z} \end{array}$$

Second we have

$$\begin{array}{rcl} U \in \mathcal{T}_{|Z} & \Rightarrow & \exists W \in \mathcal{T} \text{ such that } U = W \bigcap Z \\ & \underset{Z \subseteq Y}{\Rightarrow} & U = W \bigcap (Y \bigcap Z) = (W \bigcap Y) \bigcap Z \\ & \underset{W \bigcap Y \in \mathcal{T}_{|Y}}{\Rightarrow} & U \in (\mathcal{T}_{|Y})_{|Z} \end{array}$$

□

A subspace topology is in general not a subset of the topology that induced it, however if the subset defining the subspace topology is open this is the case.

Theorem 14.11. *Let $\langle X, \mathcal{T} \rangle$ be a topological space and $Y \subseteq \mathcal{T}$ then*

$$Y \text{ is open} \Leftrightarrow \mathcal{T}_{|Y} \subseteq \mathcal{T}$$

Proof.

⇒. Let $V \in \mathcal{T}_{|Y}$ then there exist a $U \in \mathcal{T}$ such that $V = U \bigcap Y$, as $Y, U \in \mathcal{T}$ we have that $V = U \bigcap Y \in \mathcal{T}$ proving that $\mathcal{T}_{|Y} \subseteq \mathcal{T}$.

⇐. As $\mathcal{T}_{|Y}$ is a topology on Y we have that $Y \in \mathcal{T}_{|Y}$ hence using $\mathcal{T}_{|Y} \subseteq \mathcal{T}$ we have $Y \in \mathcal{T}$ proving that Y is open. □

Given a subset of a topological space we can find the largest open set that is contained in the set, this is the idea of the interior of a set.

Definition 14.12. (Interior) *Let $\langle X, \mathcal{T} \rangle$ be a topological space and $A \subseteq X$ then the **interior** of A noted as A° is defined by*

$$A^0 = \bigcup_{U \in \{V \in \mathcal{T} | V \subseteq A\}} U$$

***Note* 14.13.** *If $U \in \{V \in \mathcal{T} | V \subseteq A\}$ then $U \in \mathcal{T}$ and $U \subseteq A$ so that $A^\circ = \bigcup_{U \in \{V \in \mathcal{T} | V \subseteq A\}} U \in \mathcal{T}$ and $A^0 \subseteq A$.*

It turns out that the interior of a set is the biggest open subset of this set.

Theorem 14.14. *Let $\langle X, \mathcal{T} \rangle$ be a topological space, $A \subseteq X$ and U a open set such that $U \subseteq A$ then $U \subseteq A^\circ$.*

Proof. If U is open and $U\subseteq A$ then $U\in\{V\in\mathcal{T}|V\subseteq A\}$ so that $U\subseteq\bigcup_{V\in\{V\in\mathcal{T}|V\subseteq A\}}V=A^{\circ}$ □

Theorem 14.15. *Let $\langle X,\mathcal{T}\rangle$ be a topological space and $U\subseteq X$ then*

$$U \text{ is open} \Leftrightarrow U=U^{\circ}$$

Proof.

⇒. If U is open then, as $U\subseteq U$, we have by the previous theorem [theorem: 14.14] that $U\subseteq U^{\circ}$ and by [definition: 14.12] $U^{\circ}\subseteq U$ Hence

$$U=U^{\circ}$$

⇐. As by [definition: 14.12] U° is open it follows from $U=U^{\circ}$ that U is open. □

14.1.1 Closed Sets

Definition 14.16. *Let $\langle X,\mathcal{T}\rangle$ be a topological space then the set of closed sets noted by $\mathcal{T}^{\mathcal{C}}$ is defined by*

$$\mathcal{T}^{\mathcal{C}}=\{A\in\mathcal{P}(X)|X\setminus A\in\mathcal{T}\}\subseteq\mathcal{P}(X)$$

*A element of $\mathcal{T}^{\mathcal{C}}$ is called a **closed** set.*

We have the following condition for closed sets in a subspace topology.

Theorem 14.17. *Let $\langle X,\mathcal{T}\rangle$ be a topological space and $Y\subseteq X$ then*

$$(\mathcal{T}_{|Y})^{\mathcal{C}}=\left\{A\bigcap Y|A\in\mathcal{T}^{\mathcal{C}}\right\}$$

In other words $C\subseteq Y$ is closed in $\mathcal{T}_{|Y}$ if there exist a closed set C' in $\mathcal{T}$ such that $C=C'\bigcap Y$

Proof. If $C\in(\mathcal{T}_{|Y})^{\mathcal{C}}$ then $Y\setminus C\in\mathcal{T}_{|Y}$ hence $\exists U\in\mathcal{T}$ such that $Y\setminus C=U\bigcap Y$. Hence

$$\begin{array}{rcl} C & = & Y\setminus(Y\setminus C)\\ & = & Y\setminus\left(U\bigcap Y\right)\\ & \underset{\text{[theorem: 1.31]}}{=} & (Y\setminus U)\bigcup(Y\setminus Y)\\ & = & Y\setminus U\\ & \underset{Y\subseteq X}{=} & \left(X\bigcap Y\right)\setminus U\\ & = & Y\bigcap(X\setminus U)\end{array}$$

which as $X\setminus U\in\mathcal{T}^{C}$ proves that $C\in\{A\bigcap Y|A\in\mathcal{T}^{\mathcal{C}}\}$. Hence

$$(\mathcal{T}_{|Y})^{\mathcal{C}}\subseteq\left\{A\bigcap Y|A\in\mathcal{T}^{\mathcal{C}}\right\} \tag{14.1}$$

If $C\in\{A\bigcap Y|A\in\mathcal{T}^{\mathcal{C}}\}$ then there exist a $C'\in\mathcal{T}^{\mathcal{C}}$ such that $C=C'\bigcap Y$ then we have

$$Y\setminus C=Y\setminus\left(C'\bigcap Y\right)\underset{\text{[theorem: 1.31]}}{=}(Y\setminus C')\bigcup(Y\setminus Y)=(Y\setminus C')=\left(X\bigcap Y\right)\setminus C'=(X\setminus C')\bigcap Y\in\mathcal{T}_{|Y}$$

proving that $C\in(\mathcal{T}_{|Y})^{\mathcal{C}}$. Hence $\{A\bigcap Y|A\in\mathcal{T}^{\mathcal{C}}\}\subseteq(\mathcal{T}_{|Y})^{\mathcal{C}}$ which combined with [eq: 14.1] results in

$$(\mathcal{T}_{|Y})^{\mathcal{C}}=\left\{A\bigcap Y|A\in\mathcal{T}^{\mathcal{C}}\right\}$$ □

Example 14.18. Let $\langle X,\mathcal{T}\rangle$ be a topological space then $\varnothing$ and X are open en closed sets

Proof. By definition of a topology $\varnothing,X$ are open sets. Further, as $X\setminus\varnothing=X\in\mathcal{T}$ and $X\setminus X=\varnothing\in\mathcal{T}$ we have that $\varnothing$ and X are also closed sets. □

The above example shows that a set can be at the same time open and closed. You could also have open sets that are not closed and closed sets that are not open. The following example shows that you can also have sets that are neither open or closed.

Example 14.19. Take the topological space $\langle\{1,2,3\},\{\varnothing,\{1,2,3\}\}\rangle$ [see example: 14.3] then $\{1\},\{2\},\{3\},\{1,2\},\{1,3\},\{2,3\}$ are neither open or closed.

Proof. Using the topology $\mathcal{T}=\{\varnothing,\{1,2,3\}\}$ it is clear that $\mathcal{T}$ does not contains $\{1\},\{2\},\{3\},\{1,2\},\{1,3\},\{2,3\}$ hence $\{1\},\{2\},\{3\}$, $\{1,2\},\{1,3\},\{2,3\}$ are not open. Further we have

$$\begin{array}{rcl} \{1,2,3\}\setminus\{1\} & = & \{2,3\}\notin\mathcal{T} \\ \{1,2,3\}\setminus\{2\} & = & \{1,3\}\notin\mathcal{T} \\ \{1,2,3\}\setminus\{3\} & = & \{1,2\}\notin\mathcal{T} \\ \{1,2,3\}\setminus\{1,2\} & = & \{3\}\notin\mathcal{T} \\ \{1,2,3\}\setminus\{1,3\} & = & \{2\}\notin\mathcal{T} \\ \{1,2,3\}\setminus\{2,3\} & = & \{1\}\notin\mathcal{T} \end{array}$$

so that $\{1\},\{2\},\{3\},\{1,2\},\{1,3\},\{2,3\}$ are not closed. □

The above two example shows that we can have sets that are open and closed, sets that are only open, sets that are only closed and sets that are neither open or closed. So open and closed are not exclusive properties. However they are complementary as the following theorem shows.

Theorem 14.20. *Let X be a set and $\mathcal{T}\subseteq\mathcal{P}(X)$ and*

$$\mathcal{T}^{\mathcal{C}}=\{A\in\mathcal{P}(X)|X\setminus A\in\mathcal{T}\}$$

then

$$\mathcal{T}=\{A\in\mathcal{P}(X)|X\setminus A\in\mathcal{T}^{\mathcal{C}}\}$$

Proof. If $A\in\{A\in\mathcal{P}(X)|X\setminus A\in\mathcal{T}^{\mathcal{C}}\}$ then $A\in\mathcal{P}(X)\Rightarrow A\subseteq X$ and $X\setminus A\in\mathcal{T}^{\mathcal{C}}$ so that $A\underset{\text{[theorem: 1.26]}}{=}X\setminus(X\setminus A)\in\mathcal{T}$ proving

$$\{A\in\mathcal{P}(X)|X\setminus A\in\mathcal{T}^{\mathcal{C}}\}\subseteq\mathcal{T} \tag{14.2}$$

If $A\in\mathcal{T}$ then as $\mathcal{T}\subseteq\mathcal{P}(X)$ we have $X\setminus(X\setminus A)\underset{\text{[theorem: 1.26]}}{=}A\in\mathcal{T}$ so that $X\setminus A\in\mathcal{T}^{\mathcal{C}}$ proving that $A\in\{A\in\mathcal{P}(X)|X\setminus A\in\mathcal{T}^{\mathcal{C}}\}$. Hence $\mathcal{T}\subseteq\{A\in\mathcal{P}(X)|X\setminus A\in\mathcal{T}^{\mathcal{C}}\}$ which combined with [eq: 14.2] proves that

$$\mathcal{T}=\{A\in\mathcal{P}(X)|X\setminus A\in\mathcal{T}^{\mathcal{C}}\}$$ □

Theorem 14.21. *Let $\langle X,\mathcal{T}\rangle$ be a topological space then the set of closed sets $\mathcal{T}^{\mathcal{C}}$ satisfies the following:*

1. $\emptyset\in\mathcal{T}^{\mathcal{C}}$
2. $X\in\mathcal{T}^{\mathcal{C}}$
3. *If $U\in\mathcal{T}$ then $X\setminus U\in\mathcal{T}^{\mathcal{C}}$*
4. *If $A,B\in\mathcal{T}^{\mathcal{C}}$ then $A\bigcup B\in\mathcal{T}^{\mathcal{C}}$*
5. *If I is finite and $\{A_i\}_{i\in I}\subseteq\mathcal{T}^{\mathcal{C}}$ is a finite family of closed sets then $\bigcup_{i\in I}A_i\in\mathcal{T}^{\mathcal{C}}$*
6. *If $\{A_i\}_{i\in I}$ is a family in $\mathcal{T}^{\mathcal{C}}$ (a family of closed sets) then $\bigcap_{i\in I}A_i\in\mathcal{T}^{\mathcal{C}}$*

Furthermore when $\mathcal{C}\subseteq\mathcal{P}(X)$ is a set of subsets of X satisfying 1,2,3 and 5 of the above then $\mathcal{T}=\{X\setminus A|A\in\mathcal{C}\}$ is a topology on X with the set of closed sets $\mathcal{T}^{\mathcal{C}}=\mathcal{C}$.

Proof.

1. $X\setminus\emptyset=X\in\mathcal{T}\Rightarrow\emptyset\in\mathcal{T}^{\mathcal{C}}$
2. $X\setminus X=\emptyset\in\mathcal{T}\Rightarrow X\in\mathcal{T}^{\mathcal{C}}$
3. This follows from [theorem: 14.20].
4. If $A,B\in\mathcal{T}^{\mathcal{C}}$ then $X\setminus(A\bigcup B)\underset{\text{[theorem: 1.31]}}{=}(X\setminus A)\bigcap(X\setminus B)\in\mathcal{T}$ [as $X\setminus A,X\setminus B\in\mathcal{T}$] so that $A\bigcup B\in\mathcal{T}^{\mathcal{C}}$.
5. If $\{A_i\}_{i\in I}\subseteq\mathcal{T}^{\mathcal{C}}$ then $\forall i\in I$ we have that $X\setminus A_i$ is open. As I is finite we have by [theorem: 14.6] that $\bigcap_{i\in I}(X\setminus A_i)$ is open, hence $X\setminus(\bigcap_{i\in I}(X\setminus A_i))$ is closed. Finally

$$X\setminus\left(\bigcap_{i\in I}(X\setminus A_i)\right)\underset{\text{[theorem: 2.130]}}{=}\bigcup_{i\in I}(X\setminus(X\setminus A_i))\underset{\text{[theorem: 1.26]}}{=}\bigcup_{i\in I}A_i$$

6. If $\{A_i\}_{i\in I}$ is a family in $\mathcal{T}^{\mathcal{C}}$ then $X\setminus(\bigcap_{i\in I} A_i) \underset{[\text{theorem: } 2.130]}{=} \bigcup_{i\in I}(X\setminus A_i)\in\mathcal{T}$ [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 5 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 $X\setminus U\in\mathcal{C}\wedge X\setminus V\in\mathcal{C}$ so that by (4) $(X\setminus U)\bigcup(X\setminus V)\in\mathcal{C}$. As by [theorem:1.31] $X\setminus(U\bigcap V)=(X\setminus U)\bigcup(X\setminus V)$ it follows that $U\bigcap V\in\mathcal{T}$.
4. If $\{U_i\}_{i\in I}\subseteq\mathcal{T}$ then $\forall i\in I$ we have $X\setminus U_i\in\mathcal{C}$ so that by (6) $\bigcap_{i\in I}(X\setminus U_i)\in\mathcal{C}$. As by [theorem: 2.130] $X\setminus(\bigcup_{i\in I}U_i)=\bigcap_{i\in I}(X\setminus U_i)$ it follows that $\bigcup_{i\in I}U_i\in\mathcal{T}$.

From the above it follows then that

$$\mathcal{T}\text{ is a topology}$$

If $A\in\mathcal{T}^{\mathcal{C}}$ then $X\setminus A\in\mathcal{T}$ so $X\setminus(X\setminus A)\in\mathcal{C}$ which as $A\underset{[\text{theorem: } 1.26]}{=}X\setminus(X\setminus A)$ proves that $A\in\mathcal{C}$, hence

$$\mathcal{T}^{\mathcal{C}}\subseteq\mathcal{C}$$

Further if $A\in\mathcal{C}$ then $X\setminus A\in\mathcal{T}$ so $A\in\mathcal{T}^{\mathcal{C}}$ proving that $\mathcal{C}\subseteq\mathcal{T}^{\mathcal{C}}$ which combined with the above proves that

$$\mathcal{T}^{\mathcal{C}}=\mathcal{C}$$

□

Just as we have defined the interior of a set as the largest open set containing in the set we can define the closure of a set as the smallest closed set containing the set.

Definition 14.22. (Closure of a Set) *Let $\langle X,\mathcal{T}\rangle$ be a topological space and $A\subseteq X$ then the **closure** of A noted by $\overline{A}$ is defined by*

$$\overline{A}=\bigcap_{C\in\{C\in\mathcal{T}^{\mathcal{C}}|A\subseteq C\}}C$$

***Note* 14.23.** *If $C\in\{C\in\mathcal{T}^{\mathcal{C}}|A\subseteq C\}$ then $A\subseteq C$ and $C\in\mathcal{T}^{\mathcal{C}}$ hence by [theorem: 2.125] we have*

$$A\subseteq\overline{A}$$

further by [theorem: 14.21] we have that

$$\overline{A}\in\mathcal{T}^{\mathcal{C}}\text{ or }\overline{A}\text{ is closed}$$

Next we show that the closure of a set is the smallest closed set containing the set.

Theorem 14.24. *Let $\langle X,\mathcal{T}\rangle$ be a topological space and $A\subseteq X$ then if C is closed set such that $A\subseteq C$ it follows that $\overline{A}\subseteq C$.*

Proof. If C is closed set such that $A\subseteq C$ then $C\in\{C\in\mathcal{T}^{\mathcal{C}}|A\subseteq C\}$ so that $\bigcap_{B\in\{C\in\mathcal{T}^{\mathcal{C}}|A\subseteq C\}}B\subseteq C$ □

It turns out that a closed set is a set that is equal to its closure.

Theorem 14.25. *Let $\langle X,\mathcal{T}\rangle$ be a topological space and $A\subseteq X$ then*

$$A\text{ is closed}\Leftrightarrow A=\overline{A}$$

Proof.

⇒. If A is closed then as $A\subseteq A$ we have by the previous theorem [theorem: 14.24] that $\overline{A}\subseteq A$ and by [definition: 14.22] $A\subseteq\overline{A}$. Hence

$$A=\overline{A}$$

⇐. As by [definition: 14.22] $\overline{A}$ is closed it follows from $A=\overline{A}$ that A is closed. □

Definition 14.26. *If $\langle X,\mathcal{T}\rangle$ is a topological space and $A\subseteq X$ then $x\in X$ is a **limit point** or **accumulation point** of A if $\forall U\in\mathcal{T}$ with $x\in U$ we have $(A\setminus\{x\})\bigcap U\neq\varnothing$. The set of all limit points of A is called the **derived** set of A and noted by A' hence*

$$A'=\{x\in X|x\text{ is a limit point of }A\}$$

Theorem 14.27. *Let $\langle X, \mathcal{T}\rangle$ be a topological space and $A\subseteq X$ then*

$$\overline{A}=A\bigcup A'$$

Proof. Let $x\in\overline{A}$ then we have either:

- $\boldsymbol{x\in A}$**.** Then trivially $x\in A\bigcup A'$
- $\boldsymbol{x\notin A}$**.** Assume that $x\notin A'$ then there exist a $U\in\mathcal{T}$ with $x\in U$ such that $(A\setminus\{x\})\bigcap U=\varnothing$. As $x\notin A$ we have that $A\setminus\{x\}=A$ hence $A\bigcap U=\varnothing$. So $\forall a\in A$ we have that $a\notin U\Rightarrow a\in X\setminus U$ proving that $A\subseteq X\setminus U$ which, as $X\setminus U$ is closed, gives by [theorem: 14.24] that $\overline{A}\subseteq X\setminus U$. Hence if $x\in\overline{A}$ we have $x\notin U$ contradicting $x\in U$. So the assumption is wrong and we must have that $x\in A'\subseteq A\bigcup A'$.

As in all cases we have $x\in A\bigcup A'$ it follows that

$$\overline{A}\subseteq A\bigcup A' \tag{14.3}$$

Let $x\in A\bigcup A'$ then we have either:

- $\boldsymbol{x\in A}$**.** Then, as by [definition: 14.22] $A\subseteq\overline{A}$, we have $x\in\overline{A}$.
- $\boldsymbol{x\notin A}$**.** Then $A\setminus\{x\}=A$ and as $x\in A\bigcup \mathrm{XA}'$ we must have $x\in A'$. Assume that $x\notin\overline{A}$ then $x\in X\setminus\overline{A}$ a open set, so as $x\in A'$ we have by [definition: 14.26] that

  $$\varnothing\neq(A\setminus\{x\})\bigcap(X\setminus\overline{A})\underset{A\setminus\{x\}=A}{=}A\bigcap(X\setminus\overline{A})\subseteq\overline{A}\bigcap(X\setminus\overline{A})=\varnothing$$

  a contradiction. Hence we must have that $x\in\overline{A}$.

As in all cases $x\in\overline{A}$ it follows that $A\bigcup A'\subseteq\overline{A}$ which combined with [eq: 14.3] proves

$$\overline{A}=A\bigcup A'$$

□

Corollary 14.28. *Let $\langle X, \mathcal{T}\rangle$ be a topological space and $A\subseteq X$ then we have*

$$\overline{A}=\left\{x\in X|\forall U\in\mathcal{T}\text{ with }x\in U\text{ we have }A\bigcap U\neq\varnothing\right\}$$

Proof. Let $x\in\overline{A}$ then we have either:

- $\boldsymbol{x\in A}$**.** Then if $U\in\mathcal{T}$ with $x\in U$ we have $x\in A\bigcap U\Rightarrow A\bigcap U\neq\varnothing$. Hence

  $$x\in\left\{x\in X|\forall U\in\mathcal{T}\text{ with }x\in U\text{ we have }A\bigcap U\neq\varnothing\right\}$$

- $\boldsymbol{x\notin A}$**.** As $\overline{A}\underset{[\text{theorem: }14.27]}{=}A\bigcup A'$ we must have that $x\in A'$, then $\forall U\in\mathcal{T}$ with $x\in U$ we have by [definition: 14.26] that $\varnothing\neq(A\setminus\{x\})\bigcap U\subseteq A\bigcap U$. Hence

  $$x\in\left\{x\in X|\forall U\in\mathcal{T}\text{ with }x\in U\text{ we have }A\bigcap U\neq\varnothing\right\}$$

So we have that

$$\overline{A}\subseteq\left\{x\in X|\forall U\in\mathcal{T}\text{ with }x\in U\text{ we have }A\bigcap U\neq\varnothing\right\} \tag{14.4}$$

Let $x\in\{x\in X|\forall U\in\mathcal{T}\text{ with }x\in U\text{ we have }A\bigcap U\neq\varnothing\}$. Assume that $x\notin\overline{A}$ then $x\in X\setminus\overline{A}\in\mathcal{T}$ hence $\varnothing\neq A\bigcap(X\setminus\overline{A})\subseteq\overline{A}\bigcap(X\setminus\overline{A})=\varnothing$ a contradiction. So we must have that $x\in\overline{A}$ proving that $\{x\in X|\forall U\in\mathcal{T}\text{ with }x\in U\text{ we have }A\bigcap U\neq\varnothing\}\subseteq\overline{A}$, which combined with [eq: 14.4] proves

$$\overline{A}=\left\{x\in X|\forall U\in\mathcal{T}\text{ with }x\in U\text{ we have }A\bigcap U\neq\varnothing\right\}$$

□

Corollary 14.29. *Let $\langle X, \mathcal{T}\rangle$ be a topological space and $A\subseteq X$ then*

$$A\text{ is closed}\Leftrightarrow A=\left\{x\in X|\forall U\in\mathcal{T}\text{ with }x\in U\text{ we have }A\bigcap U\neq\varnothing\right\}$$

Proof. This follows from [theorem: 14.25] and the previous theorem [theorem: 14.28]. □

Corollary 14.30. *Let $\langle X, \mathcal{T}\rangle$ be a topological space and $A\subseteq B\subseteq X$ then*

$$\overline{A}\subseteq\overline{B}$$

Proof. Let $x\in\overline{A}$ then for every $U\in\mathcal{T}$ with $x\in U$ we have that $A\bigcap U\neq\varnothing$ so that as $A\bigcap U\subseteq B\bigcap U$ we have $B\bigcap U\neq\varnothing$ proving $x\in\overline{B}$. Hence $\overline{A}\subseteq\overline{B}$. □

14.1.2 Basis of a topological space

When dealing with a vector space we can prove many statements by limiting ourselves to a basis of a vector space because every vector can be written as a linear combination of the basis vectors. It turns out that we can do something similar with topological spaces.

Definition 14.31. *Let $\langle X, \mathcal{T}\rangle$ be a topological space then $\mathcal{B} \subseteq \mathcal{P}(X)$ is a **basis for** $\mathcal{T}$ if $\mathcal{B} \subseteq \mathcal{T}$ and $\forall U \in \mathcal{T}$ there exists a $\{B_i\}_{i\in I} \subseteq \mathcal{B}$ such that $U = \bigcup_{i\in I} B_i$. In other words $\mathcal{B}$ is a basis for $\mathcal{T}$ if $\mathcal{B}$ consist of open sets and every open set is the union of sets in $\mathcal{B}$.*

Theorem 14.32. *Let $\langle X, \mathcal{T}\rangle$ be a topological space and $\mathcal{B}$ a basis for $\mathcal{T}$ then*

$$U \in \mathcal{T} \Leftrightarrow \forall x \in U \text{ there exist a } B \in \mathcal{B} \text{ such that } x \in B \subseteq U$$

Proof.

- **⇒.** As U is open and $\mathcal{B}$ is a basis there exist a $\{B_i\}_{i\in I} \subseteq \mathcal{B}$ such that $U = \bigcup_{i\in I} B_i$. Hence if $x \in U$ there exist a $i \in I$ such that $x \in B_i \subseteq \bigcup_{i\in I} B_i = U$ which as $B_i \in \mathcal{B}$ proves our assertion.
- **⇐.** Assume that $\forall x \in U$ there exist a $B \in \mathcal{B}$ such that $x \in B \subseteq U$ then by a consequence of the Axiom of Choice [see theorem: 3.104] there exist a $\{B_x\}_{x\in U} \subseteq \mathcal{B}$ such that

  $$\forall x \in U \text{ we have } \quad x \in B_x \subseteq U$$

  Hence using [theorem: 2.125] it follow that $\bigcup_{x\in U} B_x \subseteq U$, further if $x \in U$ then $x \in B_x$ so that $x \in \bigcup_{y\in U} B_y$, proving that $U \subseteq \bigcup_{y\in U} B_y$. Hence

  $$U = \bigcup_{x\in U} B_x \text{ where } \{B_x\}_{x\in U} \subseteq \mathcal{B}$$

  Finally as $\forall x \in U$ $B_x \in \mathcal{B} \subseteq \mathcal{T} \Rightarrow B_x \in \mathcal{T}$ it follows that $U = \bigcup_{x\in U} B_x \in \mathcal{T}$, □

Given a basis of a topology it is easy to construct a basis of a sub-topology.

Theorem 14.33. *Let $\langle X, \mathcal{T}\rangle$ be a topological space, $\mathcal{B}$ a basis for $\mathcal{T}$ and $A \subseteq X$ then*

$$\mathcal{B}_{|A} = \left\{B \bigcap A | B \in \mathcal{B}\right\} \text{ is a basis for } \mathcal{T}_{|A} = \left\{U \bigcap A | U \in \mathcal{T}\right\}$$

Proof. If $B \in \mathcal{B}_A$ then there exist a $B' \in \mathcal{B}$ such that $B = B' \bigcap A$, hence, as $\mathcal{B} \subseteq \mathcal{T} \Rightarrow B' \in \mathcal{T}$, we have that $B \in \mathcal{T}_{|A}$ proving that

$$\mathcal{B}_{|A} \subseteq \mathcal{T}_{|A}$$

If $U \in \mathcal{T}_{|A}$ then there exists a $U' \in \mathcal{T}$ such that $U = U' \bigcap A$. As $\mathcal{B}$ is a basis of $\mathcal{T}$ there exist a $\{B_i'\}_{i\in I} \subseteq \mathcal{B}$ such that $U' = \bigcup_{i\in I} B_i'$. Then

$$U = U' \bigcap A = \left(\bigcup_{i\in I} B_i'\right) \bigcap A \underset{\text{[theorem: 2.128]}}{=} \bigcup_{i\in I} \left(B_i' \bigcap A\right)$$

where $\{B_i' \bigcap A\}_{i\in I} \subseteq \mathcal{B}_{|A}$. □

Theorem 14.34. *Let $\langle X, \mathcal{T}\rangle$ be a topological space and let $\mathcal{B} \subseteq \mathcal{T}$ then*

$$\mathcal{B} \text{ is a basis for } \mathcal{T} \Leftrightarrow \forall U \in \mathcal{T} \text{ we have } \forall x \in U \text{ that there exist a } B \in \mathcal{B} \text{ such that } x \in B \subseteq U$$

Proof.

- **⇒.** Let $U \in \mathcal{T}$ then, as $\mathcal{B}$ is a basis for $\mathcal{T}$ we have by [theorem: 14.32] that $\forall x \in U$ there exist a $B \in \mathcal{B}$ such that $x \in B \subseteq U$.
- **⇐.** If $U \in \mathcal{T}$ then $\forall x \in U$ there exists a $B \in \mathcal{B} \subseteq \mathcal{T}$ such that $x \in B \subseteq U$, hence using [theorem: 3.104] there exist a $\{B_x\}_{x\in U} \subseteq \mathcal{B}$ such that $\forall x \in B_x$ $x \in \mathcal{B}_x \subseteq U$. As $\forall x \in U$ $B_x \subseteq U$ we have $\bigcup_{x\in U} B_x \subseteq U$, further if $x \in U$ then $x \in B_x \in \mathcal{B}$ so that $x \in \bigcup_{y\in U} B_y$. Hence

  $$U = \bigcup_{x\in U} B_x \text{ where } \{B_x\}_{x\in U} \subseteq \mathcal{B}$$

  Proving that $\mathcal{B}$ is a basis for $\mathcal{T}$. □

Corollary 14.35. *Let $\langle X, \mathcal{T}\rangle$ be a topological space, $\mathcal{B}$ a basis for $\mathcal{T}$ and $\mathcal{A}$ such that $\mathcal{B} \subseteq \mathcal{A} \subseteq \mathcal{T}$ then $\mathcal{A}$ is a basis of $\mathcal{T}$*

Proof. Let $U \in \mathcal{T}$ and $x \in U$ then as $\mathcal{B}$ is a basis of $\mathcal{T}$ there exist by [theorem: 14.34] a $B \in \mathcal{B}$ such that $x \in B \subseteq U$. As $\mathcal{B} \subseteq \mathcal{A}$ we have that $B \in \mathcal{A}$ proving by [theorem: 14.34] and the fact that $\mathcal{A} \subseteq \mathcal{T}$ that $\mathcal{A}$ is a basis for $\mathcal{T}$. □

Theorem 14.36. *LetX be a set and $\mathcal{T}_1$, $\mathcal{T}_2$ two topologies on X with basis $\mathcal{B}_1$, $\mathcal{B}_2$ then*

$$\mathcal{T}_2 \text{ is finer then } \mathcal{T}_1$$
$$\Updownarrow$$
$$\forall x \in X,\ \forall B \in \mathcal{B}_1 \text{ with } x \in B \text{ there exist a } B' \in \mathcal{B}_2 \text{ such that } x \in B' \subseteq B$$

Proof.

⇒. Let $x \in X$ and take $B \in \mathcal{B}_1$ such that $x \in B$. As $B \in \mathcal{B}_1 \subseteq \mathcal{T}_1 \subseteq \mathcal{T}_2 \Rightarrow B \in \mathcal{T}_2$ and $\mathcal{B}_2$ is a basis for $\mathcal{T}_2$ we have by [theorem: 14.32] that there exist a $B' \in \mathcal{B}_2$ such that $x \in B' \subseteq B$.

⇐. Let $U \in \mathcal{T}_1$ and take $x \in U$ then by [theorem: 14.32] and the fact that $\mathcal{B}_1$ is a basis for $\mathcal{T}_1$ it follows that there exist a $B \in \mathcal{B}_1$ such that $x \in B \subseteq U$. Using the hypothesis there exist a $B' \in \mathcal{B}_2$ such that $x \in B' \subseteq B \subseteq U$. Applying then [theorem: 14.32] again proves that $U \in \mathcal{T}_2$. Hence we have $\mathcal{T}_1 \subseteq \mathcal{T}_2$ or

$$\mathcal{T}_2 \text{ is finer then } \mathcal{T}_1$$ □

Corollary 14.37. *Let $\langle X, \mathcal{T} \rangle$ be a topological space, $\mathcal{B}$ a basis for $\mathcal{T}$ and $A \subseteq X$ then*

$$A \text{ is closed} \Leftrightarrow \forall x \in X \setminus A \text{ there exists a } B \in \mathcal{B} \text{ such that } x \in B \subseteq X \setminus A$$

Proof.

⇒. As A is closed we have by definition $X \setminus A \in \mathcal{T}$. Hence if $x \in X \setminus A$ we have by the previous theorem [theorem: 14.32] a $B \in \mathcal{B}$ such that $x \in B \subseteq X \setminus A$.

⇐. If $x \in X \setminus A$ then $\exists B \in \mathcal{B}$ such that $x \in B \subseteq X \setminus A$ so by the previous theorem [theorem: 14.32] $X \setminus A$ is open, hence A is closed. □

Theorem 14.38. *Let $\langle X, \mathcal{T} \rangle$ be a topological space and B a basis for $\mathcal{T}$ then:*

1. *$\forall x \in X$ there exist a $B \in \mathcal{B}$ such that $x \in B$.*
2. *$\forall B_1, B_2 \in \mathcal{B}$ we have $\forall x \in B_1 \bigcap B_2$ that $\exists B \in \mathcal{B}$ such that $x \in B \subseteq B_1 \bigcap B_2$*

Proof.

1. As $X \in \mathcal{T}$ it follows from [theorem: 14.32] that for $x \in X$ there exist a $B \in \mathcal{B}$ such that $x \in B \in \mathcal{B}$.
2. If $x \in B_1 \bigcap B_2$ then as B_1, $B_2 \in \mathcal{B} \subseteq \mathcal{T}$ we have that $B_1 \bigcap B_2 \in \mathcal{T}$, hence as $\mathcal{B}$ is a basis for $\mathcal{T}$ there exist by [theorem: 14.32] a $B \in \mathcal{B}$ such that $x \in B \subseteq B_1 \bigcap B_2$. □

The above show the necessary conditions that a basis must satisfies, the following shows that any set of subset fulfilling the above conditions can be the basis of a topology.

Theorem 14.39. *Let X be a set, $\mathcal{B} \subseteq \mathcal{P}(X)$ a set of subsets of X satisfying:*

1. *$\forall x \in X$ there exist a $B \in \mathcal{B}$ such that $x \in B$*
2. *$\forall B_1, B_2 \in \mathcal{B}$ we have $\forall x \in B_1 \bigcap B_2$ that there exist a $B \in \mathcal{B}$ such that $x \in B \subseteq B_1 \bigcap B_2$*

then

$$\mathcal{T}[\mathcal{B}] = \{U \in \mathcal{P}(X) | \forall x \in U \text{ there exist a } B \in \mathcal{B} \text{ such that } x \in B \subseteq U\} \text{ is a topology}$$

and

$$\mathcal{B} \text{ is a basis for } \mathcal{T}[\mathcal{B}]$$

*We call $\mathcal{B}$ the **generating basis** for $\mathcal{T}[\mathcal{B}]$ and $\mathcal{T}[\mathcal{B}]$ the **topology generated by** $\mathcal{B}$.*

Proof.

a) $\forall x \in \varnothing$ we have vacuously that there exist a $B \in \mathcal{B}$ such that $x \in B \subseteq \varnothing$, hence $\varnothing \in \mathcal{T}[\mathcal{B}]$.

b) Let $x \in X$ then by (1) there exists a $B \in \mathcal{B} \subseteq \mathcal{P}(X)$ such that $x \in B \subseteq X$ hence $X \in \mathcal{T}[\mathcal{B}]$.

c) Let $U_1, U_2 \in \mathcal{T}[\mathcal{B}]$. If $x \in U_1 \bigcap U_2$ then $x \in U_1 \wedge x \in U_2$, so there exists $B_1, B_2 \in \mathcal{B}$ such that $x \in B_1 \subseteq U_1$ and $x \in B_2 \subseteq U_2$, so $x \in B_1 \bigcap B_2 \subseteq U_1 \bigcap U_2$. By (2) there exist a $B \in \mathcal{B}$ such that $x \in B \subseteq B_1 \bigcap B_2 \subseteq U_1 \bigcap U_2$ proving that $U_1 \bigcap U_2 \in \mathcal{T}[\mathcal{B}]$

d) Let $\{U_i\}_{i\in I}\subseteq\mathcal{T}[\mathcal{B}]$, if $x\in\bigcup_{i\in I}U_i$ then there exist a $i\in I$ such that $x\in U_i$, hence there exist a $B\in\mathcal{B}$ such that $x\in B\subseteq U_i$ which, as $U_i\subseteq\bigcup_{i\in I}U_i$ proves that $\bigcup_{i\in I}U_i\in\mathcal{T}[\mathcal{B}]$

From (a),(b),(c) and (d) it follows that

$$\mathcal{T}[\mathcal{B}] \text{ is a topology on } X$$

If $B\in\mathcal{B}$ then $\forall x\in B$ we have $x\in B\subseteq B$ which proves that $B\in\mathcal{T}[\mathcal{B}]$, hence

$$\mathcal{B}\subseteq\mathcal{T}[\mathcal{B}]$$

Further if $U\in\mathcal{T}[\mathcal{B}]$ then if $x\in U$ there exist by definition of $\mathcal{T}[\mathcal{B}]$ a $B\in\mathcal{B}$ such that $x\in B\subseteq U$. By [theorem: 14.34] it follows that

$$\mathcal{B} \text{ is a basis of } \mathcal{T}[\mathcal{B}]$$

□

The above theorem shows how a subset of $\mathcal{P}(X)$ satisfying condition (1) and (2) can be used to generate a topology on X. What about a general subset of $\mathcal{P}(X)$ (without any conditions), can we construct a topology from it. It tuns out that the answer is yes. Given a $S\subseteq\mathcal{P}(X)$ we can create a $\mathcal{B}\subseteq\mathcal{P}(X)$ that satisfies (1) and (2) and then generate a topology from $\mathcal{B}$.

Theorem 14.40. *Let X be a set and $\mathcal{S}\subseteq\mathcal{P}(X)$ then if we define $\mathcal{B}[\mathcal{S}]$ by*

$$\mathcal{B}[\mathcal{S}]=\left\{B\in\mathcal{P}(X)|\exists\{S_i\}_{i\in I}\subseteq\mathcal{S},\ I\text{ finite and } I\neq\varnothing \text{ such that } B=\bigcap_{i\in I}S_i\right\}\bigcup\{X\}$$

we have that:

1. *$\forall x\in X$ there exist a $B\in\mathcal{B}[\mathcal{S}]$ such that $x\in B$*
2. *$\forall B_1,B_2\in\mathcal{B}[\mathcal{S}]$ we have $\forall x\in B_1\bigcap B_2$ there exist a $B\in\mathcal{B}[\mathcal{S}]$ such that $x\in B\subseteq B_1\bigcap B_2$*
3. *$\mathcal{S}\subseteq\mathcal{B}[\mathcal{S}]$*

hence by [theorem: 14.39] we have that

$$\mathcal{T}[\mathcal{B}[\mathcal{S}]]=\{U\in\mathcal{P}(X)|\forall x\in U \text{ there exist a } B\in\mathcal{B}[\mathcal{S}] \text{ such that } x\in B\subseteq U\}$$

is a topology on X. This topology is called the ***topology generated by the sub-basis $\mathcal{S}$.***

Proof.

1. As $\mathcal{B}[\mathcal{S}]=\{B\in\mathcal{P}(X)|\exists\{S_i\}_{i\in I}\subseteq\mathcal{S},\ I$ finite and $I\neq\varnothing$ such that $B[\mathcal{S}]=\bigcap_{i\in I}S_i\}\bigcup\{X\}$ we have that $X\in\mathcal{B}[\mathcal{S}]$. Hence

$$\forall x\in X \text{ we have } x\in X\subseteq X \text{ where } X\in\mathcal{B}[\mathcal{S}]$$

2. Let $x\in B_1\bigcap B_2$ where $B_1,B_2\in\mathcal{B}[\mathcal{S}]$ then there exists finite non empty families $\{S_i\}_{i\in I_1}\subseteq\mathcal{S}$, $\{T_i\}_{i\in I_2}\subseteq\mathcal{S}$ such that $B_1=\bigcap_{i\in I_1}S_1$ and $B_2=\bigcap_{i\in I_2}T_i$. Define

$$\{R_i\}_{i\in(I_1\times\{0\})\bigcup(I_2\times\{1\})}\subseteq\mathcal{S} \text{ by } R_{(i,j)}=\begin{cases}S_i \text{ if } j=0\\ T_i \text{ if } j=1\end{cases}$$

then we have

$$\begin{aligned}
y\in\bigcap_{i\in(I_1\times\{0\})\bigcup(I_2\times\{1\})}R_i &\Leftrightarrow \forall(i,j)\in(I_1\times\{0\})\bigcup(I_2\times\{1\}) \text{ we have } y\in R_{(i,j)}\\
&\Leftrightarrow [\forall(i,j)\in I_1\times\{0\} \text{ we have } y\in R_{(i,j)}]\wedge[\forall(i,j)\in I_2\times\{1\} \text{ we have } y\in R_{(i,j)}]\\
&\Leftrightarrow [\forall i\in I_1 \text{ we have } y\in R_{(i,0)}]\wedge[\forall i\in I_2 \text{ we have } y\in R_{(i,1)}]\\
&\Leftrightarrow [\forall i\in I_1 \text{ we have } y\in S_i]\wedge[\forall i\in I_2 \text{ we have } y\in T_i]\\
&\Leftrightarrow y\in\bigcap_{i\in I_1}S_i\wedge y\in\bigcap_{i\in I_2}T_i\\
&\Leftrightarrow y\in B_1\bigcap B_2
\end{aligned}$$

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 $x\in B\subseteq B_1\bigcap B_2$.

3. If $B\in\mathcal{S}$ define $\{S_i\}_{i\in\{1\}}\subseteq\mathcal{S}$ with $S_1=B$ then $B=S_1=\bigcap_{i\in\{1\}}S_i$ so that $B\in\mathcal{B}[\mathcal{S}]$. □

Given a family of a topological spaces we can generate a topology for the products of the spaces. It turns out that we have two trivial choices: the box topology and the product topology. In the finite case these topologies are the same.

Definition 14.41. *Let I be a non empty set, $\{\langle X_i, \mathcal{T}_i\rangle\}_{i\in I}$ be a family of of topological spaces then the **box topology** noted by $\mathcal{T}_{\text{box}}$ is the topology on $\prod_{i\in I} X_i$ generated by the basis*

$$\mathcal{A} = \left\{\prod_{i\in I} U_i | \{U_i\}_{i\in I} \textit{ is a family such that } \forall i \in I\ U_i \in \mathcal{T}_i\right\}$$

or

$$\mathcal{T}_{\text{box}} = \mathcal{T}[\mathcal{A}]$$

Proof. Of course we must ensure that $\mathcal{A}$ satisfies the conditions specified in [theorem: 14.39]. First if $B \in \mathcal{A}$ then there exist a $\{U_i\}_{i\in I}$ satisfying $\forall i \in I$ we have $U_i \in \mathcal{T}_i \subseteq \mathcal{P}(X_i) \Rightarrow U_i \subseteq X_i$ such that $B = \prod_{i\in I} U_i$. Hence using]theorem: 2.139] we have that $B = \prod_{i\in I} U_i \subseteq \prod_{i\in I} X_i$, proving that $B \in \mathcal{P}(\prod_{i\in I} X_i)$ or

$$\mathcal{A} \subseteq \mathcal{P}\left(\prod_{i\in I} X_i\right)$$

Further we have:

1. As $\forall i \in I\ \ \prod_{i\in I} X_i \subseteq \bigcup_{i\in I} X_i \Rightarrow \prod_{i\in I} X_i \in \mathcal{P}(\prod_{i\in I} X_i)$ and $\forall i \in I\ X_i \in \mathcal{T}_i$ it follows that

$$\prod_{i\in I} X_i \in \mathcal{A}$$

Hence $\forall x \in \prod_{i\in I} X_i$ we have for $B = \prod_{i\in I} X_i \in \mathcal{A}$ that $x \in B \in \mathcal{A}$.

2. Let $B_1, B_2 \in \mathcal{A}$ then there exist $\{U_i\}_{i\in I}, \{V_i\}_{i\in I}$ satisfying $\forall i \in I\ U_i, V_i \in \mathcal{T}_i$ such that $B_1 = \prod_{i\in I} U_i$ and $B_2 = \prod_{i\in I} V_i$ then $\forall i \in I$ we have $U_i \bigcap V_i \in \mathcal{T}_i$ so that

$$B_1 \bigcap B_2 = \left(\prod_{i\in I} U_i\right) \bigcap \left(\prod_{i\in I} V_i\right) \underset{\text{[theorem: 2.140]}}{=} \prod_{i\in I} (U_i \bigcap V_i) \in \mathcal{A}$$

hence $\forall x \in B_1 \bigcap B_2$ we have $x \in B_1 \bigcap B_2 \subseteq B_1 \bigcap B_2$ where $B_1 \bigcap B_2 \in \mathcal{A}$.

By [theorem: 14.39] it follows that $\mathcal{T}[\mathcal{A}]$ is a topology on $\prod_{i\in I} X_i$. □

We can construct the box topology based on the bases of $\mathcal{T}_i$.

Theorem 14.42. *Let I be a non empty set, $\{\langle X_i, \mathcal{T}_i\rangle\}_{i\in I}$ be a family of of topological spaces with $\forall i \in I\ \mathcal{B}_i$ a basis for $\mathcal{T}_i$ then*

$$\mathcal{B} = \left\{\prod_{i\in I} B_i | \{B_i\}_{i\in I} \textit{ is a family such that } \forall i \in I\ B_i \in \mathcal{B}_i\right\}$$

is a basis of $\mathcal{T}_{\text{box}}$ or

$$\mathcal{T}_{\text{box}} = \mathcal{T}[\mathcal{B}]$$

Proof. Using [theorem: 14.41] we have

$$\mathcal{T}_{\text{box}} = \mathcal{T}[\mathcal{A}] \text{ where } \mathcal{A} = \left\{\prod_{i\in I} U_i | \{U_i\}_{i\in I} \text{ is a family such that } \forall i \in I\ U_i \in \mathcal{T}_i\right\}$$

So if $U \in \mathcal{T}_{\text{box}}$ and $x \in U$ then by [theorem: 14.39] there exist a $\{U_i\}_{i\in I}$ satisfying $\forall i \in I\ U_i \in \mathcal{T}_i$ such that

$$x \in \prod_{i\in I} U_i \subseteq U \text{ hence } \forall i \in I \text{ we have } x_i \in U_i$$

Let $i \in I$ then as $\mathcal{B}_i$ is a basis of $\mathcal{T}_i$ there exist by [theorem: 14.32] a $B \in \mathcal{B}_i$ such that $x_i \in B \subseteq U_i$. So if we define $\mathcal{A}_i = \{B \in \mathcal{B}_i | x_i \in B \subseteq U_i\} \subseteq \mathcal{B}_i$ then $\mathcal{A}_i \neq \varnothing$. Using the Axiom of Choice [see theorem: 3.127] there exist a function $B: I \rightarrow \bigcup_{i\in I} \mathcal{A}_i$ such that $\forall i \in I\ B_i \in \mathcal{A}_i \subseteq \mathcal{B}_i$. This defines a family $\{B_i\}_{i\in I}$ such that $\forall i \in I\ x_i \in B_i \subseteq U_i$ and thus by [theorem: 2.139] $x \in \prod_{i\in I} B_i \subseteq \prod_{i\in I} U_i = U$ where $\prod_{i\in I} B_i \in \mathcal{B}$. So by [theorem: 14.34] $\mathcal{B}$ is a basis of $\mathcal{T}_{\text{box}}$. □

The second topology that we can define on a product of sets is the product topology.

Definition 14.43. *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 subbase*

$$\mathcal{S} = \{\pi_i^{-1}(V) | i \in I \wedge V \in \mathcal{T}_i\}$$

where $\forall i\in I$

$$\pi_i\colon \prod_{j\in I} X_j \rightarrow X_i \text{ is the projection map defined by } \pi_i(x)=x_i$$

or in other terms

$$\mathcal{T}_{\text{product}}=\mathcal{T}[\mathcal{B}[\mathcal{S}]]$$

[see theorem: 14.40].

***Note* 14.44.** *As* $\forall i\in I$ *we have* $\pi_i^{-1}(X_i)=X$ *it follows that* $X\in S$ *so the basis for* $\mathcal{T}_{\text{poduct}}$ *is by [theorem: 14.40]*

$$\mathcal{B}[S]=\left\{B\in\mathcal{P}\left(\prod_{i\in I} X_i\right)|\exists\{S_i\}_{i\in J}\subseteq\mathcal{S},\ J \text{ finite and } J\neq\varnothing \text{ such that } B=\bigcap_{i\in J} S_i\right\}$$

Theorem 14.45. *Let* I *be a non empty set,* $\{\langle X_i,\mathcal{T}_i\rangle\}_{i\in I}$ *be a family of topological spaces then if we define*

$$\mathcal{B}=\left\{\prod_{i\in I} U_i|\{U_i\}_{i\in I} \text{ is such that there exist a finite } J\subseteq I \text{ with } \forall i\in J\ U_i\in\mathcal{T}_i \text{ and } \forall i\in I\setminus J\ U_i=X_i\right\}$$

we have

$$\mathcal{B}=\mathcal{B}[\mathcal{S}]$$

where [see definition: 14.43]

$$\mathcal{B}[\mathcal{S}]=\left\{B\in\mathcal{P}\left(\prod_{i\in I} X_i\right)|\exists\{S_i\}_{i\in J}\subseteq\mathcal{S},\ J \text{ finite and } J\neq\varnothing \text{ such that } B=\bigcap_{i\in J} S_i\right\}$$

and

$$\mathcal{S}=\{\pi_i^{-1}(V)|i\in I\wedge V\in\mathcal{T}_i\}$$

So

$$\mathcal{T}_{\text{product}}\underset{\text{definition}}{=}\mathcal{T}[\mathcal{B}[\mathcal{S}]]=\mathcal{T}[\mathcal{B}]$$

Proof. Let $B\in\mathcal{B}[S]$. Then there must exist a $\{S_i\}_{i\in J}\subseteq\mathcal{S}$, $\varnothing\neq J$ finite such that $B=\bigcap_{i\in J}S_i$. Let $i\in J$ then as $S_i\in\mathcal{S}$ we have $\exists k_i\in I$ and a $U_{k_i}\in\mathcal{T}_{k_i}$ such that $S_i=\pi_{k_i}^{-1}(U_{k_i})$. Define

$$\forall i\in J\ \{B_{k_i,k}\}_{k\in I} \text{ by } B_{k_i,k}=\begin{cases} U_{k_i}\in\mathcal{T}_k \text{ if } k_i=k\\ X_k\in\mathcal{T}_k \text{ if } k\in I\setminus\{k_i\}\end{cases}\in\mathcal{T}_k$$

Let $x\in S_i=\pi_{k_i}^{-1}(U_{k_i})$ then $\forall k\in I$ we have either:

$\boldsymbol{k=k_i}$**.** Then $\pi_k(x)=\pi_{k_i}(x)\in U_{k_i}=B_{k_i,k}$

$\boldsymbol{k\in I\setminus\{k_i\}}$**.** Then $\pi_k(x)\in X_k=B_{k_i,k}$

proving that

$$S_i\subseteq\prod_{k\in I} B_{k_i,k} \tag{14.5}$$

Further if $x\in\prod_{k\in I}B_{k_i,k}$ then $\pi_{k_i}(x)\in B_{k_i,k_i}=U_{k_i}$ so that $x\in\pi_{k_i}^{-1}(U_{k_i})=S_i$, hence $\prod_{k\in I}B_{k_i,k}\subseteq\mathcal{S}_i$. Combining this with [eq: 14.5] gives

$$S_i=\prod_{k\in I} B_{k_i,k}$$

Then

$$\begin{aligned} B &= \bigcap_{i\in J} S_i\\ &= \bigcap_{i\in J}\left(\prod_{k\in I} B_{k_i,k}\right)\\ &\underset{[\text{theorem: } 2.141]}{=} \prod_{k\in I}\left(\bigcap_{i\in J} B_{k_i,k}\right)\\ &= \prod_{k\in I} C_k \end{aligned} \tag{14.6}$$

where $\{C_k\}_{k\in I}$ is defined by $C_k=\bigcap_{i\in J}B_{k_i,k}$. Let $k\in I$ then as $\varnothing\neq J$ is finite and $B_{k_i,k}\in\mathcal{T}_k$ it follows by [theorem: 14.6] that $C_k\in\mathcal{T}_k$, hence

$$\forall k\in I \text{ we have } C_k\in\mathcal{T}_k \tag{14.7}$$

Define $K=\{k_i|i\in J\}\subseteq I$ which is finite as J is finite [see theorem: 6.43], further as $J\neq\varnothing$ $K\neq\varnothing$. If $k\in I\setminus K$ then $\forall i\in J$ we have $k\neq k_i$ so that $B_{k_i,k}=X_k$ hence $C_k=\bigcap_{i\in J}X_k=X_k$. To conclude we have

$$B=\prod_{k\in I}C_k \text{ and } \forall k\in I\ C_k\in\mathcal{T}_k \text{ and } \forall k\in I\setminus K\ C_k=X_k \text{ where } K \text{ is finite and non empty}$$

So by definition of $\mathcal{B}$ we have $B\in\mathcal{B}$. Hence

$$\mathcal{B}(\mathcal{S})\subseteq\mathcal{B} \tag{14.8}$$

Let $B\in\mathcal{B}$. Then $B=\prod_{i\in I}U_i$ where $\{U_i\}_{i\in I}$ is a family such that there exist a finite non empty $J\subseteq I$ such that

$$\forall i\in J\ U_i\in\mathcal{T}_i \text{ and } \forall i\in I\setminus J \text{ we have } U_i=X_i \tag{14.9}$$

Define

$$\{U_{i,j}\}_{(i,j)\in I\times J} \text{ by } U_{i,j}=\begin{cases} U_i \text{ if } i=j \\ X_i \text{ if } (i,j)\in I\times J\setminus\{(i,j)\in I\times J|i=j\}\end{cases}=\begin{cases} U_i \text{ if } i=j \\ X_i \text{ if } i\neq j\end{cases} \tag{14.10}$$

Let $i\in I$. If $x\in U_i$ then $\forall j\in J$ we have either $i=j$ then $x\in U_i=U_{i,j}$ or $i\neq j$ so that $x\in X_i=U_{i,j}$ proving that $x\in\bigcap_{j\in J}U_{i,j}$. Hence

$$U_i\subseteq\bigcap_{j\in J}U_{i,j} \tag{14.11}$$

If $x\in\bigcap_{j\in J}U_{i,j}$ then $\forall j\in J$ $x\in U_{i,j}$. For i we have either:

- $\boldsymbol{i\in I\setminus J}$**.** Then $\forall j\in J$ we have $i\neq j$ so that $x\in U_{i,j}\underset{\text{[eq: 14.10]}}{=}X_i\underset{\text{[eq: 14.9]}}{=}U_i$
- $\boldsymbol{i\in J}$**.** As $x\in\bigcap_{j\in J}U_{i,j}$ we have $x\in U_{i,i}\underset{\text{[eq: 14.10]}}{=}U_i$

Hence $\bigcap_{j\in J}U_{i,j}\subseteq U_i$ which by [eq: 14.11] proves that $U_i=\bigcap_{j\in J}U_{i,j}$. Summarized

$$\forall i\in I \text{ we have } U_i=\bigcap_{j\in J}U_{i,j} \tag{14.12}$$

Next

$$\begin{aligned} B &= \prod_{i\in I}U_i \\ &\underset{\text{[eq: 14.12]}}{=} \prod_{i\in I}\left(\bigcap_{j\in J}U_{i,j}\right) \\ &\underset{\text{[theorem: 2.141]}}{=} \bigcap_{j\in J}\left(\prod_{i\in I}U_{i,j}\right) \end{aligned} \tag{14.13}$$

Let $j\in J$. If $x\in\prod_{i\in I}U_{i,j}$ then $\pi_j(x)\in U_{j,j}\underset{\text{[eq: 14.10]}}{=}U_j$ so that $x\in\pi_j^{-1}(U_j)$ proving that

$$\prod_{i\in I}U_{i,j}\subseteq\pi_j^{-1}(U_j) \tag{14.14}$$

Let $x\in\pi_j^{-1}(U_j)$ then for $i\in I$ we have either:

- $\boldsymbol{i=j}$**.** Then $\pi_i(x)\underset{i=j}{=}\pi_j(x)\in U_j\underset{\text{[eq: 14.10]}}{=}U_{j,j}=U_{i,j}$
- $\boldsymbol{i\neq j}$**.** Then $\pi_i(x)\in X_i\underset{\text{[eq: 14.10]}}{=}U_{i,j}$

so that in all cases $\pi_i(x)\in U_{i,j}$ proving that $x\in\prod_{i\in I}U_{i,j}$. Hence $\pi_j^{-1}(U_j)\subseteq\prod_{i\in I}U_{i,j}$ which combined with [eq: 14.14] proves that

$$\prod_{i\in I}U_{i,j}=\pi_j^{-1}(U_j)$$

Substituting this in [eq: 14.13] gives

$$B=\bigcap_{j\in J}\pi_j^{-1}(U_j)$$

Further $\forall j \in J$ $U_j \in \mathcal{T}_j$ so that by definition of S $\pi_j^{-1}(U_j) \in S$ which together with the above and the fact that J is finite proves that $B \in \mathcal{B}[\mathcal{S}]$. Hence $\mathcal{B} \subseteq \mathcal{B}[S]$ which combined with [eq: 14.8] gives

$$\mathcal{B} = \mathcal{B}[\mathcal{S}]$$

proving the theorem. □

Corollary 14.46. *Let I be a non empty set, $\{\langle X_i, \mathcal{T}_i\rangle\}_{i \in I}$ be a family of topological spaces then*

$$\mathcal{B} = \left\{ \bigcap_{i \in J} \pi_i^{-1}(U_i) | \{U_i\}_{i \in J} \text{ is a family such that } J \text{ is finite non empty and } \forall j \in U_j \in \mathcal{T}_j \right\}$$

is basis for the product topology $\mathcal{T}_{\text{product}}$ on $\prod_{i \in I} X_i$.

Proof. Using the definition of the product topology [definition: 14.43] we have

$$\mathcal{T}_{\text{product}} = \mathcal{T}[\mathcal{B}[\mathcal{S}]] \text{ where } \mathcal{S} = \{\pi_i^{-1}(V) | i \in I \wedge V \in \mathcal{T}_i\}$$

Let $B \in \mathcal{B}$ then there exist a finite non empty J and $\{U_i\}_{i \in J}$ satisfying $\forall j \in J$ $U_j \in \mathcal{T}_j$ such that $B = \bigcap_{i \in J} \pi_i^{-1}(U_i)$. If $i \in J$ then $\pi_i^{-1}(U_i) \in \mathcal{S} \subseteq \mathcal{B}[\mathcal{S}] \subseteq \mathcal{T}_{\text{product}}$ so that $\pi_i^{-1}(U_i) \in \mathcal{T}_{\text{product}}$. As J is finite non empty it follows from [theorem: 14.6] that $B = \bigcap_{i \in J} \pi_i^{-1}(U_i) \in \mathcal{T}_{\text{product}}$. Hence we have

$$\mathcal{B} \subseteq \mathcal{T}_{\text{product}} \tag{14.15}$$

Using [theorem: 14.45] we have that

$\mathcal{A} = \left\{ \prod_{i \in I} U_i | \{U_i\}_{i \in I} \text{ such that there exist a finite non empty } J \subseteq I \text{ with } \forall i \in J\ U_i \in \mathcal{T}_i \text{ and } \forall i \in I \setminus J\ U_i = X_i \right\}$ is a basis for the product topology $\mathcal{T}_{\text{product}}$ on $\prod_{i \in I} X_i$

If $B \in \mathcal{A}$ then $B = \prod_{i \in I} U_i$ where for $\{U_i\}_{i \in I}$ there exist a finite non empty $J \subseteq I$ such that $\forall i \in J$ $U_i \in \mathcal{T}_i$ and $\forall I \setminus J$ we have $U_i = X_i$. Then we have:

1. If $x \in \prod_{i \in I} U_i$ then $\forall i \in I$ we have that $\pi_i(x) \in U_i$ or $x \in \pi_i^{-1}(U_i)$ proving that $x \in \bigcap_{i \in I} \pi_i^{-1}(U_i)$ proving that

$$\prod_{i \in I} U_i \subseteq \bigcap_{i \in I} \pi_i^{-1}(U_i)$$

2. If $x \in \bigcap_{i \in I} \pi_i^{-1}(U_i)$ then $\forall i \in I$ we have either:

   $i \in J$. Then $x \in \pi_i^{-1}(U_i)$ so that $\pi_i(x) \in U_i$

   $i \in I \setminus J$. Then $U_i = X_i$ and $\pi_i(x) \in X_i = U_i$

   so that $x \in \prod_{i \in I} U_i$. Hence we have

$$\bigcap_{i \in I} \pi_i^{-1}(U_i) \subseteq \prod_{i \in I} U_i$$

hence

$$B = \prod_{i \in I} U_i = \bigcap_{i \in I} \pi_i^{-1}(U_i)$$

proving that $B \in \mathcal{B}$. Hence we have that $\mathcal{A} \subseteq \mathcal{B} \subseteq \mathcal{T}_{\text{product}}$ which as $\mathcal{A}$ is a basis for $\mathcal{T}_{\text{product}}$ proves by [theorem: 14.35] that $\mathcal{B}$ is a basis for $\mathcal{T}_{\text{product}}$. □

Using the above corollary it is easy to prove that in the finite case the box topology is equivalent with the product topology.

Theorem 14.47. *Let $\{\langle X_i, \mathcal{T}_i\rangle\}_{i \in I}$ be a family of topological spaces then*

1. *$\mathcal{T}_{\text{product}} \subseteq \mathcal{T}_{\text{box}}$ [in other words the box topology is finer then the product topology]*
2. *If I is finite then $\mathcal{T}_{\text{box}} = \mathcal{T}_{\text{product}}$*

Proof. Consider the bases $\mathcal{B}_{\text{product}}$, $\mathcal{B}_{\text{box}}$ for the product and box topologies [see: 14.41 and 14.45]

$$\mathcal{B}_{\text{box}} = \left\{\prod_{i\in I} U_i | \{U_i\}_{i\in I} \text{ is a family such that } \forall i\in I\; U_i\in\mathcal{T}_i\right\}$$

$$\mathcal{B}_{\text{product}} = \left\{\prod_{i\in I} U_i | \{U_i\}_{i\in I} \text{ is such that there exist a finite } J\subseteq I \text{ with } \forall i\in J\; U_i\in\mathcal{T}_i \text{ and } \forall i\in I\setminus J\; U_i = X_i\right\}$$

1. Let $B\in\mathcal{B}_{\text{product}}$ then $B=\prod_{i\in I}U_i$ where $\{U_i\}_{i\in I}$ is such that there exist a finite $J\subseteq I$ with $\forall i\in J\; U_i\in\mathcal{T}_i$ and $\forall i\in I\setminus J\; U_i = X_i$. As $\forall i\in I\setminus J\; U_i = X_i\in\mathcal{T}_i$ we have $\forall i\in I$ that $U_i\in\mathcal{T}_i$ it follows that $B\in\mathcal{B}_{\text{box}}$. Hence

$$\mathcal{B}_{\text{product}}\subseteq\mathcal{B}_{\text{box}} \tag{14.16}$$

   If $x\in X$ and $B\in\mathcal{B}_{\text{product}}$ satisfies $x\in B$ then as $\mathcal{B}_{\text{product}}\subseteq\mathcal{B}_{\text{box}}$ we have also $B\in\mathcal{B}_{\text{box}}$ hence $x\in B\subseteq B$ where $B\in\mathcal{B}_{\text{box}}$. So by [theorem: 14.36] we have that $\mathcal{T}_{\text{box}}$ is finer then $\mathcal{T}_{\text{product}}$ hence $\mathcal{T}_{\text{product}}\subseteq\mathcal{T}_{\text{box}}$.

2. Assume that I is finite. Take $B\in\mathcal{B}_{\text{box}}$ then $B=\prod_{i\in I}U_i\subseteq\prod_{i\in I}X_i$ is such that $\forall i\in I\; U_i\in\mathcal{T}_i$ which as I is finite and $I\setminus I=\varnothing$ means that $B\in\mathcal{B}_{\text{product}}$. Hence $\mathcal{B}_{\text{box}}\subseteq\mathcal{B}_{\text{product}}$ so combining this with [eq: 14.16] gives

$$\mathcal{B}_{\text{product}}=\mathcal{B}_{\text{box}}$$

   hence

$$\mathcal{T}_{\text{box}}=\mathcal{B}[\mathcal{B}_{\text{box}}]=\mathcal{B}[\mathcal{B}_{\text{product}}]=\mathcal{T}_{\text{product}}$$

□

Theorem 14.48. *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 $(\mathcal{T}_i)_{|A_i}$ then the subspace topology $(\mathcal{T}_{\text{product}})_{|\prod_{i\in I}A_i}$ for $\prod_{i\in I}A_i$ is the product topology of $\{\langle A_i,(\mathcal{T}_i)_{|A_i}\rangle\}_{i\in I}$.*

Proof. Let $\mathcal{B}_{\text{product}}$ be the basis of $\mathcal{T}_{\text{product}}$ then by [theorem: 14.33]

$$(\mathcal{B}_{\text{product}})_{|\prod_{i\in I}A_i} = \left\{B\bigcap\prod_{i\in I}A_i | B\in\mathcal{B}_{\text{product}}\right\} \text{ is a basis for } (\mathcal{T}_{\text{product}})_{|\prod_{i\in I}A_i}$$

further let $\mathcal{T}^{\text{product}}$ be the product topology of $\{\langle A_i,(\mathcal{T}_i)_{|A_i}\rangle\}_{i\in I}$ and $\mathcal{B}^{\text{product}}$ the basis van $\mathcal{T}^{\text{product}}$.

1. If $B\in(\mathcal{B}_{\text{product}})_{|\prod_{i\in I}A_i}$ then by [definition: 14.45] $B=(\prod_{i\in I}U_i)\bigcap(\prod_{i\in I}A_i)$ where $U_i=X_i$ except for a finite subset J of I where $U_i\in\mathcal{T}$. Hence $B=\prod_{i\in I}(U_i\bigcap A_i)$ where for $i\in I\setminus J\; U_i\bigcap A_i=X_i\bigcap A_i=A_i$ and for $i\in J\;\; U_i\bigcap A_i\in(\mathcal{T}_i)_{|A_i}$. So by [definition: 14.45] it follows that $B\in\mathcal{B}^{\text{product}}$ the basis of $\mathcal{T}^{\text{product}}$ [the product topology of $\{\langle A_i,(\mathcal{T}_i)_{|A_i}\rangle\}_{i\in I}$]. Hence we have

$$(\mathcal{B}_{\text{product}})_{|\prod_{i\in I}A_i}\subseteq\mathcal{B}^{\text{product}}$$

2. If $B\in\mathcal{B}^{\text{product}}$ then by [definition: 14.45] $B=\prod_{i\in I}U_i$ where $U_i=A_i=X_i\bigcap A_i$ except for a finite $J\subseteq I$ where $U_i\in(\mathcal{T}_i)_{A_i}\Rightarrow U_i=W_i\bigcap A_i$ where $W_i\in\mathcal{T}_i$. Then if we define $\{V_i\}_{i\in I}$ by $V_i=\left\{\begin{smallmatrix} X_i \text{ if } i\in I\setminus J\\ W_i \text{ if } i\in J\end{smallmatrix}\right.$ then $B=\prod_{i\in I}(V_i\bigcap A_i)\underset{\text{[theorem: 2.140]}}{=}(\prod_{i\in I}V_i)\bigcap(\prod_{i\in I}A_i)$. By [definition: 14.45] $\prod_{i\in I}V_i\in\mathcal{B}_{\text{product}}$ so that $B\in(\mathcal{B}_{\text{product}})_{|\prod_{i\in I}A_i}$. Hence we have

$$\mathcal{B}^{\text{product}}\subseteq(\mathcal{B}_{\text{product}})_{|\prod_{i\in I}A_i}$$

Using (1) and (2) it follows that $(\mathcal{B}_{\text{product}})_{|\prod_{i\in I}A_i}=\mathcal{B}^{\text{product}}$ proving that

$$(\mathcal{T}_{\text{product}})_{|\prod_{i\in I}A_i}=\mathcal{T}^{\text{product}} \text{ the product topology of } \{\langle A_i,\mathcal{T}_i\rangle\}_{i\in I}$$

□

14.1.3 Dense sets

Definition 14.49. (Dense set) *Let $\langle X,\mathcal{T}\rangle$ be a topological space then $A\subseteq X$ is a **dense subset of X** if $\overline{A}=X$ [see definition: 14.22 for $\overline{A}$].*

Theorem 14.50. *Let $\langle X,\mathcal{T}\rangle$ be a topological space and $A\subseteq X$ then*

$$A \text{ is a dense subset of } X$$
$$\Updownarrow$$
$$\forall x\in X \text{ we have } \forall U\in\mathcal{T} \text{ with } x\in U \text{ that there } \exists a\in A \text{ such that } a\in U.$$

Proof.

- **⇒.** If A is dense then $X=\overline{A}\underset{[\text{theorem: 14.28}]}{=}\{x\in X|\forall U\in\mathcal{T}\text{ with }x\in U\text{ we have }U\bigcap A\neq\varnothing\}$. So if $x\in X$ then $\forall U\in\mathcal{T}$ with $x\in U$ we have $U\bigcap A\neq\varnothing\Rightarrow\exists a\in A$ such that $a\in U$.
- **⇐.** If $x\in X$ then if $U\in\mathcal{T}$ with $x\in U$ there exist by the hypothesis a $x\in A$ such that $a\in U$, hence $a\in U\bigcap A$ proving that $U\bigcap A\neq\varnothing$. So by [theorem: 14.28] $x\in\overline{A}$. Hence $X\subseteq\overline{A}$ which as $\overline{A}\subseteq X$ proves that $X=\overline{A}$ or that A is dense in X. □

We can rephrase the previous theorem in terms of a basis of a topology.

Theorem 14.51. *Let $\langle X,\mathcal{T}\rangle$ be a topological space with basis $\mathcal{B}$ and $A\subseteq X$ then*

$$A\text{ is dense in }X$$
$$\Updownarrow$$
$$\forall x\in X\text{ we have }\forall B\in\mathcal{B}\text{ with }x\in B\text{ there exist a }a\in A\text{ such that }a\in B$$

Proof.

- **⇒.** Assume that A is dense in X. Let $B\in\mathcal{B}$ be such that $x\in B$ then as $\mathcal{B}\subseteq\mathcal{T}$ we have by the previous theorem [theorem: 14.50] that there exists a $a\in A$ such that $a\in B$.
- **⇐.** Let $x\in X$ then if $U\in\mathcal{T}$ satisfies $x\in U$ we have, as $\mathcal{B}$ is a basis for $\mathcal{T}$, by [theorem: 14.32] that $x\in B\subseteq U$. Using the hypothesis there exist a $a\in A$ such that $a\in B\subseteq U$. Hence using the previous theorem [theorem: 14.50] it follows that A is dense in X. □

Theorem 14.52. *Let $\langle X,\mathcal{T}\rangle$ be a topological space and $A\subseteq X$ then*

$$A^\circ=\varnothing\Leftrightarrow X\setminus A\text{ is dense in }X$$

Proof.

- **⇒.** Assume that $A^\circ=\varnothing$. Let $x\in X$ and $U\in\mathcal{T}$ such that $x\in U$. Assume that $U\subseteq A$ then by [definition: 14.12] $x\in A^\circ$ contradicting $A^\circ\neq\varnothing$, hence we must have that $A\nsubseteq U$. So there exist a $y\in U$ such that $y\notin A\Rightarrow y\in X\setminus A$ proving that $y\in U\bigcap(X\setminus A)$ or $U\bigcap(X\setminus A)\neq\varnothing$. By [theorem: 14.28] it follows that $x\in\overline{X\setminus A}$, hence $X\subseteq\overline{X\setminus A}$ which as $\overline{X\setminus A}\subseteq X$ proves that $X=\overline{X\setminus A}$. Hence $X\setminus A$ is dense in X.
- **⇐.** Assume that $x\in A^\circ$ then as A° is open and by the hypothesis $X\setminus A$ dense there exist by [theorem: 14.50] a $a\in X\setminus A$ such that $a\in A^\circ$, hence $\varnothing\neq A^\circ\bigcap(X\setminus A)\subseteq A\bigcap(X\setminus A)=\varnothing$, leading the the contradiction $\varnothing\neq\varnothing$, Hence the assumption is wrong and we must have that
$$A^\circ=\varnothing$$ □

Definition 14.53. *A topological space $\langle X,\mathcal{T}\rangle$ is a Baire space if for every $\{A_i\}_{i\in\mathbb{N}}\subseteq\mathcal{P}(X)$ satisfying $\forall i\in\mathbb{N}$ A_i is closed and $(A_i)^\circ=\varnothing$ we have that $(\bigcup_{i\in\mathbb{N}}A_i)^\circ=\varnothing$. In other words the union of any family of closed sets with empty interior has also a empty interior.*

The most trivial and least usefully example of a Baire set is $\langle\varnothing,\{\varnothing\}\rangle$

Example 14.54. Let $X=\varnothing$ then by [theorem: 14.2] the only topology on X is $\{\varnothing\}$. The topological $\langle\varnothing,\{\varnothing\}\rangle$ is a Baire space.

Proof. If $\{A_i\}_{i\in\mathbb{N}}\subseteq\mathcal{P}(X)=\{\varnothing\}$ then $\forall i\in\mathbb{N}$ $A_i=\varnothing$ so that so that $(\bigcup_{i\in\mathbb{N}}A_i)^\circ\subseteq\bigcup_{i\in\mathbb{N}}A_i=\varnothing$. □

Theorem 14.55. *Let $\langle X,\mathcal{T}\rangle$ be a topological space then*

$$\langle X,\mathcal{T}\rangle\text{ is a Baire space}$$
$$\Updownarrow$$
$$\text{For every family }\{U_i\}_{i\in\mathbb{N}}\subseteq\mathcal{T}\text{ such that }\forall i\in\mathbb{N}\ U_i\text{ is open en dense in }X\text{ we have that }\bigcap_{i\in\mathbb{N}}U_i\text{ is dense in }X$$

Proof.

- **⇒.** Let $\{U_i\}_{i\in\mathbb{N}}$ be a family of open sets that are dense in X. Define $\{A_i\}_{i\in\mathbb{N}}$ by $A_i=X\setminus U_i$ then $\forall i\in\mathbb{N}$ A_i is closed and as $X\setminus A_i=X\setminus(X\setminus U_i)=U_i$ is dense in X we have by [theorem: 14.52] that $(A_i)^\circ=\varnothing$. Hence by the definition of a Baire space we have
$$\left(\bigcup_{i\in\mathbb{N}}A_i\right)^\circ=\varnothing$$

Using [theorem: 14.52] again we have that

$$X \setminus \left(\bigcup_{i \in \mathbb{N}} A_i \right) \text{ is dense in } X$$

As $X \setminus (\bigcup_{i \in \mathbb{N}} A_i) \underset{[\text{theorem: } 2.130]}{=} \bigcap_{i \in \mathbb{N}} (X \setminus A_i) = \bigcap_{i \in \mathbb{N}} (X \setminus (X \setminus U_i)) = \bigcap_{i \in \mathbb{N}} U_i$ proving by the above that

$$\bigcap_{i \in \mathbb{N}} U_i \text{ is dense in } X$$

⇐. Let $\{A_i\}_{i \in \mathbb{N}}$ be a family of closed sets with $\forall i \in \mathbb{N}\ (A_i)^\circ = \varnothing$. Define $\{U_i\}_{i \in \mathbb{N}}$ by $U_i = X \setminus A_i$ then $\forall i \in \mathbb{N}$ we have that U_i is open and as $(X \setminus U_i)^\circ = (X \setminus (X \setminus A_i))^\circ = (A_i)^\circ = \varnothing$ we have by [theorem: 14.52] that $U_i = X \setminus (X \setminus U_i)$ is dense in X. By the hypothesis we have then that $\bigcap_{i \in \mathbb{N}} U_i$ is dense in X. As $\bigcap_{i \in \mathbb{N}} U_i = \bigcap_{i \in \mathbb{N}} (X \setminus A_i) \underset{[\text{theorem: } 2.130]}{=} X \setminus (\bigcup_{i \in \mathbb{N}} A_i)$ it follows that

$$X \setminus \left(\bigcup_{i \in \mathbb{N}} A_i \right) \text{ is dense in } X$$

Using [theorem: 14.52] again it follows that

$$\left(\bigcup_{i \in \mathbb{N}} A_i \right)^\circ = \varnothing$$

Hence by definition $\langle X, d \rangle$ is a Baire space. □

14.2 Metric spaces

A metric space is a set with a concept of a distance that allows us to define a topology.

Definition 14.56. (Pseudo Metric Space) *A pseudo metric space $\langle X, d \rangle$ is a **non empty** set X together with a function*

$$d: X \times X \to \mathbb{R}$$

such that

1. $\forall x \in X\ d(x, x) = 0$
2. $\forall x, y \in X$ *we have* $d(x, y) = d(y, x)$
3. $\forall x, y, z \in X$ *we have* $d(x, z) \leqslant d(x, y) + d(y, z)$

*the function $d: X \times X \to \mathbb{R}$ is called a **pseudo metric**.*

Definition 14.57. (Metric Space) *A **pseudo metric space** $\langle X, d \rangle$ is called a **metric space** if we have $\forall x, y \in X$ we have $d(x, y) = 0 \Leftrightarrow x = y$. In other words*

$$d: X \times X \to \mathbb{R}$$

satisfies:

1. $\forall x, y \in X$ *we have* $d(x, y) = d(y, x)$
2. $\forall x, y, z \in X$ *we have* $d(x, z) \leqslant d(x, y) + d(y, z)$
3. $\forall x, y \in X$ *we have* $d(x, y) = 0 \Leftrightarrow x = y$ *[hence* $\forall x \in X\ d(x, x) = 0$*]*

the function $d: X \times X \to \mathbb{R}$ is called a metric.

Theorem 14.58. *If $\langle X, d \rangle$ is a pseudo metric space then $\forall x, y \in X$ we have $0 \leqslant d(x, y)$*

Proof. If $x, y \in X$ then $0 = d(x, x) \leqslant d(x, y) + d(y, x) = d(x, y) + d(x, y) = 2 \cdot d(x, y)$ hence

$$0 \leqslant d(x, y)$$

□

To define a topology we need the concept of balls.

Definition 14.59. (Open Ball) *Let $\langle X, d \rangle$ be a metric space, $\varepsilon \in \mathbb{R}^+$ and $x \in X$ then a **open ball centered around** x noted by $B_d(x, \varepsilon)$ is defined by*

$$B_d(x, \varepsilon) = \{y \in X | d(x, y) < \varepsilon\} \subseteq X$$

Definition 14.60. (Closed Ball) *Let $\langle X, d\rangle$ be a metric space, $\varepsilon \in \mathbb{R}^+$ and $x \in X$ then a **closed ball centered around** x noted by $\overline{B_d(x,\varepsilon)}$ is defined by*

$$\overline{B_d(x,\varepsilon)} = \{y \in X | d(x,y) \leqslant \varepsilon\} \subseteq X$$

We proceed now to define topology based on a pseudo metric.

Lemma 14.61. *Let $\langle X, d\rangle$ be a pseudo metric space then if $x_1, x_2 \in X$ and $\varepsilon_1, \varepsilon_2 \in \mathbb{R}^+$ we have that*

$$\forall x \in B_d(x_1,\varepsilon_1) \bigcap B_d(x_2,\varepsilon_2) \text{ there } \exists \varepsilon \in \mathbb{R}^+ \text{ such that } x \in B_d(x,\varepsilon) \subseteq B_d(x_1,\varepsilon_1) \bigcap B_d(x_2,\varepsilon_2)$$

Proof. Let $x \in B_d(x_1,\varepsilon_1) \bigcap B_d(x_2,\varepsilon_2)$ then $d(x,x_1) < \varepsilon_1$ and $d(x,x_2) < \varepsilon_2$. Define

$$\varepsilon = \min(\varepsilon_1 - d(x_1,x), \varepsilon_2 - d(x_2,x))$$

then as $0 < \varepsilon_1 - d(x_1,x) \wedge 0 < \varepsilon_2 - d(x_2,x)$ we have that

$$\varepsilon \in \mathbb{R}^+$$

Let $y \in B_d(x,\varepsilon)$ then $d(x,y) < \varepsilon$ so that

$$d(x_1,y) \leqslant d(x_1,x) + d(x,y) < d(x_1,y) + \varepsilon \leqslant d(x_1,x) + \varepsilon_1 - d(x_1,x) = \varepsilon_1$$

and

$$d(x_2,y) \leqslant d(x_2,x) + d(x,y) < d(x_2,y) + \varepsilon < d(x_2,x) + \varepsilon_2 - d(x_2,x) = \varepsilon_2$$

so that $y \in B_d(x_1,\varepsilon_1) \wedge y \in B_d(x_2,\varepsilon_2)$ or $y \in B_d(x_1,\varepsilon_1) \bigcap B_d(x_2,\varepsilon_2)$, Hence

$$B_d(x,\varepsilon) \subseteq B_d(x_1,\varepsilon_1) \bigcap B_d(x_2,\varepsilon_2)$$

Finally as $d(x,x) = 0 < \varepsilon$ we have that

$$x \in B_d(x,\varepsilon)$$

□

Corollary 14.62. *Let $\langle X, d\rangle$ be a pseudo metric space, $x \in X$ and $\varepsilon \in \mathbb{R}^+$ then*

$$\forall y \in B_d(x,\varepsilon) \text{ there exist a } \delta \in \mathbb{R}^+ \text{ such that } y \in B_d(y,\delta) \subseteq B_d(x,\varepsilon)$$

Proof. As $y \in B_d(x,\varepsilon) = B_d(x,\varepsilon) \bigcap B_d(x,\varepsilon)$ we can use the previous lemma [lemma: 14.61] to find a $\delta \in \mathbb{R}^+$ such that $B_d(y,\delta) \subseteq B_d(x,\varepsilon) \bigcap B_d(x,\varepsilon) = B_d(x,\varepsilon)$. □

Next we show that open balls can be used to define a topology on a metric space.

Theorem 14.63. *Let $\langle X, d\rangle$ be a pseudo metric space then if we define $\mathcal{B}_d$ by*

$$\mathcal{B}_d = \{B_d(x,\varepsilon) | x \in X \wedge \varepsilon \in \mathbb{R}^+\}$$

we have:

1. *$\forall x \in X$ there exists a $B \in \mathcal{B}_d$ such that $x \in B$.*
2. *$\forall B_1, B_2 \in \mathcal{B}_d$ we have $\forall x \in B_1 \bigcap B_2$ that there exist a $B \in \mathcal{B}_d$ such that $x \in B \subseteq B_1 \bigcap B_2$*

so that by [theorem: 14.39]

$$\mathcal{T}_d = \mathcal{T}[\mathcal{B}_d] \underset{\text{def}}{=} \{U \in \mathcal{P}(X) | \forall x \in U \text{ there exist a } B \in \mathcal{B}_d \text{ such that } x \in B \subseteq U\}$$

is a topology with $\mathcal{B}_d$ as a basis.

Proof.

1. $\forall x \in X$ we have as $d(x,x) = 0 < 1$ that $x \in B_d(x,1) \in \mathcal{B}_d$.
2. If $B_1, B_2 \in \mathcal{B}_d$ then there exist a $x_1, x_2 \in X$ and $\varepsilon_1, \varepsilon_2 \in \mathbb{R}^+$ such that $B_1 = B_d(x_1,\varepsilon_1)$ and $B_2 = B_d(x_2,\varepsilon_2)$. So if $x \in B_1 \bigcap B_2$ we have $x \in B_d(x_1,\varepsilon_1) \bigcap B_d(x_2,\varepsilon_2)$ hence by [lemma: 14.61] there exist a $\varepsilon \in \mathbb{R}^+$ such that for $B = B_d(x.\varepsilon) \in \mathcal{B}_d$ we have

$$x \in B = B_d(x.\varepsilon) \subseteq B_d(x_1,\varepsilon_1) \bigcap B_d(x_2,\varepsilon_2) = B_1 \bigcap B_2$$

□

Corollary 14.64. *Let $\langle X, d\rangle$ be a pseudo metric space that*

$$U \in \mathcal{T}_d \Leftrightarrow \forall x \in U \text{ there exist a } \varepsilon \in \mathbb{R}^+ \text{such that} \, x \in B_d(x,\varepsilon) \subseteq U$$

Proof.

⇒. 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 [corollary: 14.62] there exists a $\delta > 0$ such that $x \in B_d(x, \delta) \subseteq B_d(y, \varepsilon) \subseteq U$.

⇐. If for every $x \in U$ there exists a $\delta > 0$ such that $x \in B_d(x, \delta) \subseteq U$ we have as $B_d(x, \delta) \in \mathcal{B}_d$ that $U \in \mathcal{T}_d$ [using theorem: 14.32]. □

The reason why $\overline{B_d(x, \varepsilon)}$ is called a closed ball because it is closed using the metric topology.

Theorem 14.65. *Let $\langle X, d \rangle$ be a pseudo metric space with the metric topology $\mathcal{T}_d$, $x \in X$ and $\varepsilon \in \mathbb{R}^+$ then $\overline{B_d(x, \varepsilon)}$ is closed.*

Proof. Let $y \in X \setminus \overline{B_d(x, \varepsilon)}$ then we must have that $\varepsilon < d(x, y)$. Take $\delta = d(x, y) - \varepsilon > 0$ then if $z \in B_d(y, \delta)$ we have $d(y, z) < \delta = d(x, y) - \varepsilon$. Assume that $d(x, z) \leqslant \varepsilon$ then we have

$$d(x, y) \leqslant d(x, z) + d(y, z) < \varepsilon + d(x, y) - \varepsilon < d(x, y)$$

leading to the contradiction $d(x, y) < d(x, y)$. Hence the assumption is wrong and we must have $\varepsilon < d(x, z)$ so that $z \in X \setminus \overline{B_d(x, \varepsilon)}$ proving that $y \in B_d(y, \delta) \subseteq X \setminus \overline{B_d(x, \varepsilon)}$. So $X \setminus \overline{B_d(x, \varepsilon)}$ is open from which it follows that $\overline{B_d(x, \varepsilon)}$ is closed. □

If we have a metric space $\langle X, d \rangle$ with the metric topology $\mathcal{T}_d$ then the subspace topology has a very simple form as is show in the next theorem:

Theorem 14.66. *Let $\langle X, d \rangle$ be a pseudo metric space with the metric topology $\mathcal{T}_d$ and $A \subseteq X$ then for the restricted function $d_{|A \times A}$: $A \times A \to R$ we have*

1. *$\langle A, d_{|A \times A} \rangle$ is a pseudo metric space.*
2. *If $\langle X, d \rangle$ is a metric space then $\langle A, d_{|A \times A} \rangle$ is a metric space.*
3. *$\mathcal{T}_{d_{|A \times A}} = (\mathcal{T}_d)_{|A}$ [the induced subspace topology on A is the metric topology of the restricted metric topology].*

Proof.

1. We have:
   a. $\forall x \in A \quad d_{|A \times A}(x, x) = d(x, x) = 0$
   b. $\forall x, y \in A \quad d_{|A \times A}(x, y) = d(x, y) = d(y, x) = d_{|A \times A}(y, x)$
   c. $\forall x, y, z \in A \quad d_{|A \times A}(x, z) = d(x, z) \leqslant d(x, y) + d(y, z) = d_{|A \times A}(x, y) + d_{|A \times A}(y, z)$
2. Let $x, y \in A$ such that $d_{|A \times A}(x, y) = 0$ then $d(x, y) = 0$ hence $x = y$.
3. Let $x \in A$ and $\varepsilon \in \mathbb{R}^+$ then we have

$$\begin{aligned} y \in B_{d_{|A \times A}}(x, \varepsilon) &\Leftrightarrow y \in A \wedge d_{|A \times A}(x, y) < \varepsilon \\ &\Leftrightarrow y \in A \wedge d(x, y) < \varepsilon \\ &\underset{A \subseteq X \wedge x \in A}{\Leftrightarrow} y \in A \wedge x \in X \wedge d(x, y) < \varepsilon \\ &\Leftrightarrow y \in A \wedge x \in B_d(x, y) \\ &\Leftrightarrow y \in B_d(x, y) \bigcap A \end{aligned}$$

proving that

$$\forall x \in A \text{ and } \forall \varepsilon \in \mathbb{R}^+ \; B_{d_{|A \times A}}(x, \varepsilon) = B_d(x, \varepsilon) \bigcap A \tag{14.17}$$

Let $U \in (\mathcal{T}_d)_{|A}$ then there exist a $V \in \mathcal{T}_d$ such that $U = V \bigcap A$. Let $x \in U$ then $x \in V$ so that that by [corollary: 14.64] there exist a $\varepsilon \in \mathbb{R}^+$ such that $x \in B_d(x, \varepsilon) \subseteq V$, hence $x \in B_d(x, \varepsilon) \bigcap A \subseteq V \bigcap A = U$ or using [eq: 14.17] $x \in B_{d_{|A \times A}}(x, \varepsilon) \subseteq U$. So using [corollary: 14.64] again it follows that $U \in \mathcal{T}_{d_{|A \times A}}$ giving

$$(\mathcal{T}_d)_{|A} \subseteq \mathcal{T}_{d_{|A \times A}} \tag{14.18}$$

Let $U \in \mathcal{T}_{d_{|A \times A}}$ and $x \in U$ then as $\mathcal{T}_{d_{|A \times A}}$ is a topology on A we have $U \subseteq A$, hence $x \in A$. By [corollary: 14.64] there exist a $\varepsilon \in \mathbb{R}^+$ such that $x \in B_{d_{|A \times A}}(x, \varepsilon) \subseteq U \underset{\text{[eq: 14.17]}}{\Rightarrow} x \in B_d(x, \varepsilon) \bigcap A \subseteq U$. As $B_d(x, \varepsilon) \bigcap A \in (\mathcal{B}_d)_{|A}$ a basis for $(\mathcal{T}_d)_{|A}$ [see theorem: 14.33] it follows from [theorem: 14.32] that $U \in (\mathcal{T}_d)_{|A}$. Hence $\mathcal{T}_{d_{|A \times A}} \subseteq (\mathcal{T}_d)_{|A}$ which combined with [eq: 14.18] result in

$$(\mathcal{T}_d)_{|A} = \mathcal{T}_{d_{|A \times A}}$$

□

Definition 14.67. *Two pseudo metrics d_1, d_2 on a set X are equivalent if $\mathcal{T}_{d_1} = \mathcal{T}_{d_2}$. In other words two pseudo metrics are equivalent iff there generated topologies are the same.*

Theorem 14.68. *Let d_1, d_2 be two pseudo metrics on a set X then*

$$\mathcal{T}_{d_2} \text{ is a finer then } \mathcal{T}_{d_1}$$

$$\Updownarrow$$

$$\forall x \in X, \forall \varepsilon \in \mathbb{R}^+ \text{ there exist s } \delta \in \mathbb{R}^+ \text{ such that } x \in B_{d_2}(x, \delta) \subseteq B_{d_1}(x, \varepsilon)$$

Proof.

⇒. If $x \in X$ and $\varepsilon \in \mathbb{R}^+$ we have for $B_{d_1}(x, \varepsilon) \in \mathcal{B}_{d_1}$ using [theorem: 14.36] that there exist a $y \in X$ and a $\delta' \in \mathbb{R}^+$ such that $x \in B_{d_2}(y, \delta') \subseteq B_{d_1}(x, \varepsilon)$. Hence using [theorem: 14.62] there exist a $\delta \in \mathbb{R}^+$ such that $x \in B_{d_2}(x, \delta) \subseteq B_{d_2}(y, \delta') \subseteq B_{d_1}(x, \varepsilon)$.

⇐. Let $B \in \mathcal{B}_{d_1}$ then $\exists x \in X$, $\varepsilon \in \mathbb{R}^+$ such that $B = B_{d_1}(x, \varepsilon)$, using the hypothesis there exists a $\delta \in \mathbb{R}^+$ such that $x \in B_{d_2}(x, \delta) \subseteq B_{d_1}(x, \varepsilon)$. As $B_{d_2}(x, \varepsilon) \in \mathcal{B}_{d_2}$ it follows from [theorem: 14.36] that $\mathcal{T}_{d_2}$ is finer then $\mathcal{T}_{d_1}$. □

Definition 14.69. *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 a **isometry** if $\forall x, y \in X$ we have*

$$d_Y(\varphi(x), \varphi(y)) = d_X(x, y)$$

*If in addition $\varphi: X \to Y$ is **bijective** then we say that $\varphi: X \to Y$ is a **isometric isomorphism**.*
*If between two pseudo metric spaces there exist a **isometric isomorphism** then we say that these pseudo metric spaces are **isometric**. We note this relation as*

$$X \approx Y$$

*In essence two metric spaces are **isometric** if there exist a bijection between these spaces that conserves distances.*

Note that a isometry between a metric space [not just pseudo metric space] and a pseudo metric space is automatically injective.

Theorem 14.70. *Let $\langle X, d_X \rangle$ be a metric space, $\langle Y, d_Y \rangle$ a pseudo metric space and*

$$\varphi: X \to Y \text{ a isometry}$$

then

$$\varphi \text{ is injective}$$

Hence if in addition $\varphi: X \to Y$ is surjective it follows that $\varphi: X \to Y$ is a bijection so making $\varphi: X \to Y$ a isometric isomorphism.

Proof. If for $x, y \in X$ we have that $\varphi(x) = \varphi(y)$ then $d_X(x, y) = d_Y(\varphi(x), \varphi(y)) = 0$. As d_X is a metric it follows that $x = y$ proving injectivity. □

Theorem 14.71. *Let $\langle X, d_X \rangle$, $\langle Y, d_Y \rangle$ be two pseudo metric spaces and $\varphi: X \to Y$ a **isometric isomorphism** then*

$$\varphi^{-1}: Y \to X$$

is a isometry.

Proof. By [theorem: 2.72] $\varphi^{-1}: Y \to X$ is a function. Further if $x, y \in Y$ then $d_Y(x, y) = d_Y(\varphi(\varphi^{-1}(x)), \varphi(\varphi^{-1}(y))) \underset{\varphi \text{ is a isometry}}{=} d_X(\varphi^{-1}(x), d_x(\varphi^{-1}(y)))$. □

Theorem 14.72. *Let $\langle X, d_X \rangle$, $\langle Y, d_Y \rangle$, $\langle Z, d_Z \rangle$ be pseudo metric spaces and $\varphi: X \to Y$, $\psi: Y \to Z$ isometries then*

$$\psi \circ \varphi: X \to Z \text{ is a isometry}$$

If in addition φ and ψ are isometric isomorphisms then

$$\psi \circ \varphi: X \to Z \text{ is a isometric isomorphism}$$

Proof. First we have

$$d_Z((\psi \circ \varphi)(x), (\psi \circ \varphi)(y)) = d_Z(\psi(\varphi(x)), \psi(\varphi(y))) = d_Y(\varphi(x), \varphi(y)) = d_X(x, y)$$

proving that $\psi \circ \varphi : X \to Z$ is a isometry. Further if φ, ψ isometric isomorphisms, hence bijective, then by [theorem: 2.75] $\psi \circ \varphi$ is a bijection, so $\psi \circ \varphi$ is a isometric isomorphism. □

Theorem 14.73. *Let $\langle X, d_X \rangle$, $\langle Y, d_Y \rangle$ be two pseudo metric spaces and $\varphi : X \to Y$ a isometric isomorphism then we have:*

1. *$\forall x \in X$, $\varepsilon \in \mathbb{R}^+$ we have $\varphi(B_{d_X}(x, \varepsilon)) = B_{d_Y}(\varphi(x), \varepsilon)$*
2. *$\forall x \in Y$, $\varepsilon \in \mathbb{R}^+$ we have $\varphi^{-1}(B_{d_Y}(y, \varepsilon)) = B_{d_X}(\varphi^{-1}(y), \varepsilon)$*

Proof.

1. If $y \in \varphi(B_{d_X}(x, \varepsilon))$ then there exist a $z \in B_{d_X}(x, \varepsilon)$ such that $y = \varphi(z)$, hence $d_X(x, z) < \varepsilon$ or as φ is a isometry

$$d_Y(\varphi(x), y) = d_Y(\varphi(x), \varphi(z)) = d_X(x, z) < \varepsilon$$

proving that $y \in B_{d_Y}(\varphi(x), \varepsilon)$. On the other hand if $y \in B_{d_Y}(\varphi(x), \varepsilon)$ then $d_Y(\varphi(x), y) < \varepsilon$, hence

$$d_X(x, \varphi^{-1}(y)) = d_Y(\varphi(x), \varphi(\varphi^{-1}(y))) = d_Y(\varphi(x), y) < \varepsilon$$

so that $\varphi^{-1}(y) \in B_{d_X}(x, \varepsilon)$.

2. As $\psi = \varphi^{-1} : Y \to X$ is also a isometry we have

$$B_{d_X}(\varphi^{-1}(y), \varepsilon) = B_{d_X}(\psi(y), \varepsilon) \underset{(1)}{=} \psi(B_{d_Y}(y, \varepsilon)) = \varphi^{-1}(B_{d_Y}(y, \varepsilon))$$

□

Theorem 14.74. *Let $\langle X, d_x \rangle$, $\langle Y, d_Y \rangle$ be two pseudo metric spaces and $\varphi : X \to Y$ a isometric isomorphism. Then*

1. *$\mathcal{T}_{d_X} = \{\varphi^{-1}(V) | V \in \mathcal{T}_{d_Y}\} = \{U \subseteq X | \varphi(U) \in \mathcal{T}_{d_Y}\}$*
2. *$\mathcal{T}_{d_Y} = \{\varphi(U) | U \in \mathcal{T}_{d_X}\} = \{V \subseteq Y | \varphi^{-1}(V) \in \mathcal{T}_{d_X}\}$*

Proof.

1. Let $U \in \{\varphi^{-1}(V) | V \in \mathcal{T}_{d_Y}\}$ then there exist a $V \in \mathcal{T}_{d_Y}$ such that $U = \varphi^{-1}(V)$, hence $\varphi(U) = V \in \mathcal{T}_{d_Y}$ proving that

$$\{\varphi^{-1}(V) | V \in \mathcal{T}_{d_Y}\} \subseteq \{U \subseteq X | \varphi(U) \in \mathcal{T}_{d_Y}\} \tag{14.19}$$

If $U \in \{U \subseteq X | \varphi(U) \in \mathcal{T}_{d_Y}\}$ then $\varphi(U) \in \mathcal{T}_{d_Y}$ so, as $U = \varphi^{-1}(\varphi(U))$, we have $U \in \{\varphi^{-1}(V) | V \in \mathcal{T}_{d_Y}\}$, proving that $\{U \subseteq X | \varphi(U) \in \mathcal{T}_{d_Y}\} \subseteq \{\varphi^{-1}(V) | V \in \mathcal{T}_{d_Y}\}$. Combining this with [eq: 14.19] proves that

$$\{\varphi^{-1}(V) | V \in \mathcal{T}_{d_Y}\} = \{U \subseteq X | \varphi(U) \in \mathcal{T}_{d_Y}\} \tag{14.20}$$

Let $U \in \mathcal{T}_{d_X}$ and take $y \in \varphi(U)$ then $\exists x \in U$ such that $y = \varphi(x)$. Using [theorem: 14.64] there exist a $\varepsilon \in \mathbb{R}^+$ such that $x \in B_{d_X}(x, \varepsilon) \subseteq U$ hence

$$y = \varphi(x) \in \varphi(B_{d_X}(x, \varepsilon)) \underset{\text{[theorem: 14.73]}}{=} B_{d_Y}(\varphi(x), \varepsilon) \subseteq \varphi(U)$$

Hence, using [theorem: 14.64] it follows that $\varphi(U) \in \mathcal{T}_{d_Y}$ giving $U \in \{U \subseteq X | \varphi(U) \in \mathcal{T}_{d_Y}\}$. So we conclude that

$$\mathcal{T}_{d_X} \subseteq \{U \subseteq X | \varphi(U) \in \mathcal{T}_{d_Y}\} \tag{14.21}$$

Let $U \in \{\varphi^{-1}(V) | V \in \mathcal{T}_{d_Y}\}$ and $x \in U$ then there exist a $V \in \mathcal{T}_{d_Y}$ such that $x \in U = \varphi^{-1}(V) \Rightarrow \varphi(x) \in V$. As V is open there exist by [theorem: 14.64] a $\varepsilon \in \mathbb{R}^+$ such that

$$\varphi(x) \in B_{d_Y}(\varphi(x), \varepsilon) \subseteq V$$

Hence

$$x = \varphi^{-1}(\varphi(x)) \in \varphi^{-1}(B_{d_Y}(\varphi(x), \varepsilon)) \subseteq \varphi^{-1}(V) = U$$

As $\varphi^{-1}(B_{d_Y}(\varphi(x), \varepsilon)) \underset{\text{[theorem: 14.73]}}{=} B_{d_X}(\varphi^{-1}(\varphi(x)), \varepsilon) = B_{d_X}(x, \varepsilon)$ it follows from the above that

$$x \in B_{d_X}(x, \varepsilon) \subseteq U$$

So by [theorem: 14.64] we have $U \in \mathcal{T}_{d_X}$. Hence $\{\varphi^{-1}(V) | V \in \mathcal{T}_{d_Y}\} \subseteq \mathcal{T}_{d_X}$ and combining this with [eq: 14.21] gives

$$\mathcal{T}_{d_X} \subseteq \{U \subseteq X | \varphi(U) \in \mathcal{T}_{d_Y}\} \underset{\text{[eq: 14.20]}}{=} \{\varphi^{-1}(V) | V \in \mathcal{T}_{d_Y}\} \subseteq \mathcal{T}_{d_X}$$

or

$$\mathcal{T}_{d_X} = \{U \subseteq X | \varphi(U) \in \mathcal{T}_{d_Y}\} = \{\varphi^{-1}(V) | V \in \mathcal{T}_{d_Y}\}$$

2. As $\psi: Y \to X$ where $\psi = \varphi^{-1}$ is a isometry it follows from (1) that

$$\begin{aligned}
\mathcal{T}_{d_Y} &= \{U \subseteq Y | \psi(U) \in \mathcal{T}_{d_X}\} \\
&= \{U \subseteq Y | \varphi^{-1}(U) \in \mathcal{T}_{d_X}\} \\
\mathcal{T}_{d_Y} &= \{\psi^{-1}(V) | V \in \mathcal{T}_{d_X}\} \\
&= \{\varphi(V) | V \in \mathcal{T}_{d_X}\}
\end{aligned}$$

hence

$$\mathcal{T}_{d_Y} = \{U \subseteq Y | \varphi^{-1}(U) \in \mathcal{T}_{d_X}\} = \{\varphi(V) | V \in \mathcal{T}_{d_X}\}$$

□

Definition 14.75. *Let $\langle X, d\rangle$ be a pseudo metric space and $A \subseteq X$ then A* ***is bounded*** *if*

$$\exists M \in \mathbb{R}^+ \text{ such that } \forall x, y \in A \text{ we have } d(x, y) \leqslant M$$

If A is bounded and ***non empty*** *then using the fact that $\langle \mathbb{R}, \leqslant\rangle$ is conditional complete [see theorem: 10.18] it follows that*

$$\sup(\{d(x, y) | x, y \in A\}) \text{ exist for } A$$

The ***diameter of A*** *noted as $\operatorname{diam}(A)$ is then defined by*

$$\operatorname{diam}(A) = \sup(\{d(x, y) | x, y \in A\})$$

Example 14.76. Let $\langle X, d\rangle$ be a metric space then $\varnothing$ is bounded.

Proof. $\forall x, y \in \varnothing$ we have vacuously that $d(x, y) \leqslant 1$ □

Theorem 14.77. *Let $\langle X, d\rangle$ a pseudo metric space and $U \subseteq X$ a bounded set then every subset of U is also bounded.*

Proof. As U is bounded there exist a $M \in \mathbb{R}^+$ such that $\forall x, y \in U$ we have $d(x, y) \leqslant M$. Let $V \subseteq U$ then $\forall x, y \in V$ we have as $V \subseteq U$ that $x, y \in U$ hence $d(x, y) \leqslant M$, proving that V is bounded. □

Theorem 14.78. *Let $\langle X, d\rangle$ be a pseudo metric space then $\forall x \in X$, $\forall \varepsilon \in \mathbb{R}^+$ we have that*

$$B_d(x, \varepsilon) \text{ and } \overline{B_d(x,\varepsilon)} \text{ are bounded}$$

and

$$\operatorname{diam}(B_d(x, \varepsilon)) \leqslant 2 \cdot \varepsilon \wedge \operatorname{diam}(\overline{B_d(x, \varepsilon)}) \leqslant 2 \cdot \varepsilon$$

Proof.

1. Let $y, z \in B_d(x, \varepsilon)$ then $d(y, z) \leqslant d(y, x) + d(x, z) < \varepsilon + \varepsilon = 2 \cdot \varepsilon$ proving that $B_d(x, \varepsilon)$ is bounded and $\operatorname{diam}(B_d(x, \varepsilon)) = \sup(\{d(x, y) | x, y \in B_d(x, \varepsilon)\}) \leqslant 2 \cdot \varepsilon$.

2. Let $y, z \in \overline{B_d(x, \varepsilon)}$ then $d(y, z) \leqslant d(y, x) + d(x, z) \leqslant \varepsilon + \varepsilon = 2 \cdot \varepsilon$ proving that $\overline{B_d(x, \varepsilon)}$ is bounded and $\operatorname{diam}(\overline{B_d(x, \varepsilon)}) = \sup(\{d(x, y) | x, y \in \overline{B_d(x, \varepsilon)}\}) \leqslant 2 \cdot \varepsilon$. □

Theorem 14.79. *Let $\{\langle X_i, d_i\rangle\}_{i \in I}$ be a* ***finite*** *family of pseudo metric spaces then for*

$$d_{\max}: \left(\prod_{i \in I} X_i\right) \times \left(\prod_{i \in I} X_i\right) \to \mathbb{R} \text{ defined by } d_{\max}(x, y) = \max(\{d_i(\pi_i(x), \pi_i(y)) | i \in I\})$$

where $\forall i \in I$ π_i is the projection map defined by

$$\pi_i: \prod_{j \in I} X_j \to X_i \text{ is defined by } \pi_i(x) = x(i) \underset{\text{note}}{=} x_i$$

we have:

1. *$\langle \prod_{i \in I} X_i, d_{\max}\rangle$ is a pseudo metric space.*

2. *If* $\forall i\in I$ $\langle X_i, d_i\rangle$ *is a metric space then* $\langle \prod_{i\in I} X_i, d_{\max}\rangle$ *is a metric space*
3. $\mathcal{T}_{d_{\max}} = \mathcal{T}_{\text{box}} = \mathcal{T}_{\text{product}}$ *[the metric topology is equal to the box and product topology of* $\prod_{i\in I} X_i$*.*

***Note* 14.80.** *As I is finite* $\max(\{d_i(\pi_i(x), \pi_i(y))|i\in I\})$ *exist by [theorem: 6.48].*

Proof.

1. Let $x, y, z\in \prod_{i\in I} X_i$ then
   a. $d_{\max}(x, x) = \max(\{d_i(\pi_i(x), \pi_i(x))|i\in I\}) = \max(\{0\}) = 0$
   b. $d_{\max}(x, y) = \max(\{d_i(\pi_i(x).\pi_i(y))\}|i\in I) = \max(\{d_i(\pi_i(y).\pi_i(x))|i\in I\}) = d(y, x)$
   c. We have

$$\begin{array}{rcl} d_{\max}(x,z) & = & \max(\{d_i(\pi_i(x), \pi_i(z))|i\in I\}) \\ & \leqslant_{[\text{theorem: } 3.66]} & \max(\{d_i(\pi_i(x), \pi_i(y)) + d(\pi_i(y), \pi_i(z))|i\in I\}) \\ & \leqslant_{[\text{theorem: } 3.66]} & \max(\{\max(\{d_i(\pi_i(x), \pi_i(y))|i\in I\}) + d(\pi_i(y), \pi_i(z))|i\in I\}) \\ & \underset{[\text{theorem: } 10.26]}{=} & \max(\{d_i(\pi_i(x), \pi_i(y))|i\in I\}) + \max(\{d(\pi_i(y), \pi_i(z))|i\in I\}) \\ & = & d_{\max}(x,y) + d_{\max}(y,z) \end{array}$$

2. Let $d(x, y) = 0$ then as $\forall i\in I$ we have

$$0\leqslant d_i(\pi_i(x), \pi_i(y)) \leqslant \max(\{d_i(\pi_i(x), \pi_i(y))|i\in I\} = d(x, y) = 0$$

   so that $d_i(\pi_i(x), \pi_i(y)) = 0$. As d_i is a metric it follows that $\pi_i(x) = \pi_i(y) \Rightarrow x(i) = y(i)$ proving finally that $x = y$.

3. As I is finite we have by [theorem: 14.47] that

$$\mathcal{T}_{\text{box}} = \mathcal{T}_{\text{product}} \tag{14.22}$$

   The bases of $\mathcal{T}_d$ and $\mathcal{T}_{\text{box}}$ are [see definitions: 14.63, 14.42]

$$\mathcal{B}_{d_{\max}} = \{B_{d_{\max}}(x, \varepsilon)|x\in X\wedge \varepsilon\in \mathbb{R}^+\}$$

   and

$$\mathcal{B}_{\text{box}} = \left\{\prod_{i\in I} B_i|\{B_i\}_{i\in I} \text{ such that } \forall i\in I B_i\in \mathcal{B}_{d_i}\right\}$$

   Let $x\in \prod_{i\in I} X_i$ and $B\in \mathcal{B}_{d_{\max}}$ such that $x\in B$. Then there exists a $y\in X$, $\varepsilon\in \mathbb{R}^+$ such that $x\in B_{d_{\max}}(y, \varepsilon)\subseteq B$ or using [theorem: 14.62] there exist a $\delta\in \mathbb{R}^+$ such that

$$x\in B_{d_{\max}}(x, \delta)\subseteq B_{d_{\max}}(y, \varepsilon)\subseteq B$$

   Take $\prod_{i\in I} B_{d_i}(\pi_i(x), \delta)\in \mathcal{B}_{\text{box}}$ then, as $\forall i\in I$ we have $d_i(\pi_i(x), \pi_i(x_i)) = 0 < \delta$, it follows that $\pi_i(x)\in B_{d_i}(\pi_i(x), \delta)$, proving

$$x\in \prod_{i\in I} B_{d_i}(\pi_i(x), \delta)$$

   Let $z\in \prod_{i\in I} B_{d_i}(\pi_i(x), \delta)$ then $\forall i\in I$ we have $\pi_i(z)\in B_{d_i}(\pi_i(x), \delta)$ so that $d_i(\pi_i(x), \pi_i(z)) < \delta$. Hence $d_{\max}(x, z) = \max(\{d_i(\pi_i(x), \pi_i(z))|i\in I\}) < \delta$ which proves that $z\in B_{d_{\max}}(x, \delta)$, giving

$$\prod_{i\in I} B_{d_i}(\pi_i(x), \delta)\subseteq B_{d_{\max}}(x, \delta)\subseteq B$$

   So if we take $B' = \prod_{i\in I} B_{d_i}(\pi_i(x), \delta)\in \mathcal{B}_{\text{box}}$ then we have

$$\forall x\in \prod_{i\in I} X_i \text{ and } \forall B\in \mathcal{B}_{d_{\max}} \text{ there exist a } B'\in \mathcal{B}_{\text{box}} \text{ such that } x\in B'\subseteq B \tag{14.23}$$

   Let $x\in \prod_{i\in I} X_i$ and $B\in \mathcal{B}_{\text{box}}$ such that $x\in B$. So $B = \prod_{i\in I} B_i$ where $\forall i\in I$ $B_i\in \mathcal{B}_{d_i}\subseteq \mathcal{T}_{d_i}$ and $\pi_i(x)\in B_i$. Using [theorem: 14.64] there exist a $\delta_i\in \mathbb{R}^+$ such that $\pi_i(x)\in B_{d_i}(\pi_i(x), \delta_i)\subseteq B_i$. Take $\delta = \min(\{\delta_i|i\in I\})\in \mathbb{R}^+$ then for $z\in B_{d_{\max}}(x, \delta)$ we have

$$\forall i\in I\ d_i(\pi_i(x), \pi_i(z)) \leqslant \max(\{d_i(\pi_i(x), \pi_i(x))|i\in I\}) = d(x, z) < \delta\leqslant \delta_i$$

so that $\forall i\in I$ $\pi_i(z)\in B_{d_i}(\pi_i(x),\pi_i(z))\subseteq B_i$. So we have that $z\in\prod_{i\in I}B_i$ proving that $\mathcal{B}_{d_{\max}}(x,\delta)\subseteq\prod_{i\in I}B_i=B$. So if we take $B'=B_{d_{\max}}(x,\delta)$ then we have proved that

$$\forall x\in\prod_{i\in I}X_i \text{ and } \forall B\in\mathcal{B}_{\text{box}} \text{ there exist a } B'\in\mathcal{B}_{d_{\max}} \text{ such that } x\in B'\subseteq B \tag{14.24}$$

Combining now [eqs: 14.23, 14.24] with [theorem: 14.36] proves

$$\mathcal{T}_{d_{\max}}=\mathcal{T}_{\text{box}}$$ □

For metric spaces we have a simple test to determine if a set is dense in a topological space.

Theorem 14.81. *Let $\langle X,d\rangle$ a pseudo metric space and $A\subseteq X$ then we have*

$$A \text{ is dense in } X \text{ [using the metric topology } \mathcal{T}_d]$$
$$\Updownarrow$$
$$\forall x\in X \text{ we have } \forall\varepsilon\in\mathbb{R}^+ \text{ that there exist a } a\in A \text{ such that } d(x,a)<\varepsilon$$

Proof.

- **⇒.** Let $x\in X$ then $\forall\varepsilon\in\mathbb{R}^+$ $x\in B_d(x,\varepsilon)$ so that by [theorem: 14.51] there exist a $a\in A$ such that $a\in B_d(x,\varepsilon)$, hence $d(x,a)<\varepsilon$.
- **⇐.** Let $x\in X$ and let $B\in\mathcal{B}_d$ such that $x\in B$. As $B\in\mathcal{B}_d$ there exists a $y\in X$ and a $\varepsilon\in\mathbb{R}^+$ such that $x\in B=B_d(y,\varepsilon)$. By [theorem: 14.62] there exist a $\delta\in\mathbb{R}^+$ such that

  $$x\in B_d(x,\delta)\subseteq B_d(y,\varepsilon)=B.$$

  By the hypotheses there exist a $a\in A$ such that $d(x,a)<\delta$ or $a\in B_d(x,\delta)\subseteq B$. So

  $$\forall x\in X \text{ we have } \forall B\in\mathcal{B}_d \text{ with } x\in B \text{ there exist a } a\in A \text{ such that } a\in B$$

  which by [theorem: 14.51] implies that A is dense in X. □

14.3 Normed space

The idea of a (pseudo) normed space is that of a real or complex vector space where the concept of a length of a vector is defined.

Notation 14.82. *from now on $\mathbb{K}$ represent either the field $\mathbb{R}$ or the field $\mathbb{C}$*

Definition 14.83. (Pseudo Norm) *A **pseudo normed** space noted as $\langle X,\|\|\rangle$ is a real [or complex] vector space X together with a function $\|\|: X\rightarrow\mathbb{R}$ called the **norm** such that*

1. *$\forall x\in X$ we have $0\leqslant\|x\|$*
2. *$\forall x\in X$, $\forall\alpha\in\mathbb{K}$ we have $\|\alpha\cdot x\|=|\alpha|\cdot\|x\|$*
3. *$\forall x,y\in X$ we have $\|x+y\|\leqslant\|x\|+\|y\|$*

Theorem 14.84. *Let $\langle X,\|\|\rangle$ be a pseudo normed space then $\|0\|=0$*

Proof. We have $\|0\|=\|0\cdot 0\|=|0|\cdot\|0\|=0\cdot\|0\|=0$ □

Definition 14.85. (Norm) *A pseudo normed space $\langle X,\|\|\rangle$ is a normed space if we have also*

$$\forall x\in X \text{ with } \|x\|=0 \text{ we have } x=0$$

Theorem 14.86. *Let $\langle X,\|\|\rangle$ be a pseudo normed space and $\{x_i\}_{i\in I}\subseteq X$, I finite then*

$$\left\|\sum_{i\in I}x_i\right\|\leqslant\sum_{i\in I}\|x_I\|$$

Proof. We first prove this for $I=\{1,\ldots,n\}$, so let

$$\mathcal{S}=\left\{n\in\mathbb{N}_0|\forall\{x_i\}_{i\in\{1,\ldots,n\}}\subseteq X \text{ we have } \left\|\sum_{i\in\{1,\ldots,n\}} x_i\right\|\leqslant\sum_{i\in\{1,\ldots,n\}}\|x_I\|\right\}$$

then we have:

$0\in S$. As $\{1,\ldots,0\}=\varnothing$ we have

$$\left\|\sum_{i\in\{1,\ldots,0\}} x_i\right\|=\left\|\sum_{i\in\varnothing} x_i\right\|=\|0\|=0=\sum_{i\in\varnothing}\|x_i\|=\sum_{i\in\{1,\ldots,0\}}\|x_i\|$$

hence $0\in S$.

$n\in S\Rightarrow n+1\in S$. Let $\{x_i\}_{i\in\{1,\ldots,n+1\}}\subseteq X$ then we have

$$\begin{aligned}
\left\|\sum_{i\in\{1,\ldots,n+1\}} x_i\right\| &\underset{[\text{theorem: } 11.43]}{=} \left\|\sum_{i\in\{1,\ldots,n\}} x_i+\sum_{i\in\{n+1\}} x_i\right\|\\
&\underset{[\text{theorem: } 11.34]}{=} \left\|\sum_{i\in\{1,\ldots,n\}} x_i+x_{n+1}\right\|\\
&\leqslant \left\|\sum_{i\in\{1,\ldots,n\}} x_i\right\|+\|x_{n+1}\|\\
&\leqslant_{n\in S} \sum_{i\in\{1,\ldots,n\}}\|x_i\|+\|x_{n+1}\|\\
&\underset{[\text{theorem: } 11.34]}{=} \sum_{i\in\{1,\ldots,n\}}\|x_i\|+\sum_{i\in\{n+1\}}\|x_{n+1}\|\\
&\underset{[\text{theorem: } 11.43]}{=} \sum_{i\in\{1,\ldots,n+1\}}\|x_i\|
\end{aligned}$$

proving that $n+1\in S$.

Using mathematical induction proves that

$$\forall n\in\mathbb{N}_0 \text{ we have if } \{x_i\}_{i\in\{1,\ldots,n\}}\subseteq X \text{ that } \left\|\sum_{i\in\{1,\ldots,n\}} x_i\right\|\leqslant\sum_{i\in\{1,\ldots,n\}}\|x_i\| \tag{14.25}$$

Let now I be a finite set and $\{x_i\}_{i\in I}\subseteq X$ then there exist a bijection $\beta:\{1,\ldots,n\}\rightarrow I$ so that we have

$$\begin{aligned}
\left\|\sum_{i\in I} x_i\right\| &\underset{[\text{theorem: } 11.36]}{=} \left\|\sum_{i\in\{1,\ldots,n\}} x_{\beta(i)}\right\|\\
&\leqslant_{[\text{eq: } 14.25]} \sum_{i\in\{1,\ldots,n\}}\|x_{\beta(i)}\|\\
&\underset{[\text{theorem: } 11.36]}{=} \sum_{i\in I}\|x_i\|
\end{aligned}$$

□

Theorem 14.87. *Let $\langle X,\|\|\rangle$ be a pseudo normed space then $\forall x,y\in X$ we have*

$$\|x\|-\|y\|\leqslant\|x+y\|\wedge\|y\|-\|x\|\leqslant\|x+y\|\wedge|\|x\|-\|y\||\leqslant\|x+y\|$$

Proof. We have $\|x\|\leqslant\|x+y-y\|\leqslant\|x+y\|+\|y\|$ and $\|y\|\leqslant\|y+x-x\|\leqslant\|x+y\|+\|x\|$ so that

$$\|x\|-\|y\|\leqslant\|x+y\| \text{ and } \|y\|-\|x\|\leqslant\|x+y\|$$

For $\|x\|-\|y\|$ we have either

$0\leqslant\|x\|-\|y\|$. Then $|\|x\|-\|y\||=\|x\|-\|y\|\leqslant\|x+y\|$

$\|x\|-\|y\|<0$. Then $|\|x\|-\|y\||=\|y\|-\|x\|\leqslant\|x+y\|$ □

Given a normed space we can define a metric and thus also a topology.

Theorem 14.88. *Let $\langle X, \|\|\rangle$ be a pseudo normed space and define*

$$d_{\|\|}: X \times X \to X \text{ by } d_{\|\|}(x, y) = \|x - y\|$$

then

$$\langle X, d_{\|\|}\rangle \text{ is a pseudo metric space}$$

Further if $\langle X, \|\|\rangle$ is a normed space then $\langle X, d_{\|\|}\rangle$ is a metric space.

Proof. If $\langle X, \|\|\rangle$ is a pseudo normed space then we have for $x, y, z \in X$

1. $d_{\|\|}(x, y) = \|x - y\| \geqslant 0$
2. $d_{\|\|}(x, y) = \|x - y\| = \|(-1) \cdot (y - x)\| = |-1| \cdot \|y - x\| = d_{\|\|}(y, x)$
3. $d_{\|\|}(x, z) = \|x - z\| = \|x - y + y - z\| \leqslant \|x - y\| + \|y - z\| = d_{\|\|}(x, y) + d_{\|\|}(y, z)$

If in addition if $\langle X, \|\|\rangle$ is a normed space then if $d_{\|\|}(x, y) = 0$ we have $\|x - y\| = 0$ or as $\|\|$ is a norm that $x - y = 0 \Rightarrow x = y$. Hence $d_{\|\|}$ is a metric. □

As a pseudo metric defines a topology and a pseudo norm defines a pseudo metric a pseudo norm can be used to define a topology.

Definition 14.89. *Let $\langle X, \|\|\rangle$ be pseudo normed space, $x \in X$ and $\varepsilon \in \mathbb{R}^+$ then the open ball $B_{\|\|}(x, \varepsilon)$ in the pseudo normed space is defined by*

$$B_{\|\|}(x, \varepsilon) = \{y \in X \,|\, \|x - y\| < \varepsilon\}$$

the set of all open balls in the pseudo normed space is noted by $\mathcal{B}_{\|\|}$ hence

$$\mathcal{B}_{\|\|} = \{B_{\|\|}(x, \varepsilon) | x \in X \wedge \varepsilon \in \mathbb{R}^+\}$$

***Note* 14.90.** *As $d_{\|\|}(x, y) = \|x - y\|$ we have that*

$$B_{d_{\|\|}}(x, \varepsilon) = \{y \in X \,|\, d_{\|\|}(x, y) < \varepsilon\} = \{y \in X \,|\, \|x - y\| < \varepsilon\} = B_{\|\|}(x, \varepsilon)$$

and further

$$\mathcal{B}_{d_{\|\|}} = \{B_{d_{\|\|}}(x, \varepsilon) | x \in X \wedge \varepsilon \in \mathbb{R}^+\} = \{B_{\|\|}(x, \varepsilon) | x \in X \wedge \varepsilon \in \mathbb{R}^+\} = \mathcal{B}_{\|\|}$$

Theorem 14.91. *Let $\langle X, \|\|_X\rangle$ be a pseudo normed space then we have:*

1. *If $\varepsilon \in \mathbb{R}^+$ then $B_{\|\|}(0, \varepsilon) = \varepsilon \cdot B_{\|\|}(0, 1)$*
2. *If $x \in X$ and $\varepsilon \in \mathbb{R}^+$ then $B_{\|\|}(x, \varepsilon) = x + B_{\|\|}(0, \varepsilon)$*

Proof.

1. If $x \in B_{\|\|}(0, \varepsilon)$ then $\|x\| < \varepsilon$, define $x' = \frac{1}{\varepsilon} \cdot x$ so that $x = \varepsilon \cdot x'$ and

   $$\|x'\| = \left\|\frac{1}{\varepsilon} \cdot x\right\| = \frac{1}{\varepsilon}\|x\| < \frac{1}{\varepsilon} \cdot \varepsilon = 1$$

   proving that $x \in \varepsilon \cdot B_{\|\|}(0, 1)$. Further if $x \in \varepsilon \cdot B_{\|\|}(0, 1)$ then $x = \varepsilon \cdot x'$ where $x' \in B_{\|\|}(0, 1)$ so that $\|x'\| < 1$ or $\|x\| = \varepsilon \cdot \|x\| < \varepsilon \cdot 1 = \varepsilon$ proving that $x \in B_{\|\|}(0, \varepsilon)$. Hence we conclude that

   $$B_{\|\|}(0, \varepsilon) = \varepsilon \cdot B_{\|\|}(0, 1)$$

2. $y \in B_{\|\|}(x, \varepsilon) \Leftrightarrow \|y - x\| < \varepsilon \Leftrightarrow y - x \in B_{\|\|}(0, \varepsilon) \Leftrightarrow y = x + B_{\|\|}(0, \varepsilon)$ □

Definition 14.92. *Let $\langle X, \|\|\rangle$ be a pseudo normed space then the norm topology $\mathcal{T}_{\|\|}$ on X is defined by*

$$\mathcal{T}_{\|\|} = \mathcal{T}_{d_{\|\|}}$$

where

$$d_{\|\|}: X \times X \to X \text{ is defined by } d_{\|\|}(x, y) = \|x - y\|$$

So

$$\mathcal{T}_{\|\|} = \mathcal{T}[\mathcal{B}_{d_{\|}}]$$

where

$$\mathcal{B}_{d_{\|\|}} = \{B_{d_{\|\|}}(x,\varepsilon)|x\in X\wedge\varepsilon\in\mathbb{R}^+\} = \{B_{\|\|}(x,\varepsilon)|x\in X\wedge\varepsilon\in\mathbb{R}^+\} \underset{\text{def}}{=} \mathcal{B}_{\|\|}$$

Using the above definition we have the equivalent version of [theorem: 14.64].

Theorem 14.93. *Let $\langle X, \|\|_X\rangle$ be a pseudo normed space then we have*

$$U\in\mathcal{T}_{\|\|} \Leftrightarrow \forall x\in U\ \exists\varepsilon\in\mathbb{R}^+ \text{ such that } x\in B_{\|\|}(x,\varepsilon)\subseteq U$$

Proof. This follows from [theorem: 14.64] and the fact that $\mathcal{T}_{\|\|}=\mathcal{T}_{d_{\|\|}}$ and $B_{\|\|}(x,\varepsilon)=B_{d_{\|\|}}(x,\varepsilon)$. □

Giving a finite family of normed spaces we can create a norm for the product of these spaces.

Theorem 14.94. *Let $\langle X_i, \|\|_i\rangle_{i\in I}$, I finite be a finite family of normed spaces then*

$$\|\|_{\max}: \prod_{i\in I} X_i \rightarrow \mathbb{R} \text{ defined by } \|x\|_{\max} = \max(\{\|\pi_i(x)\|_i|i\in I\})$$

*is a norm called the **maximum norm**. Further the generated topology $\mathcal{T}_{\|\|_{\max}}$ is the same as the product topology or*

$$\mathcal{T}_{\|\|_{\max}} = \mathcal{T}_{\text{product}}$$

Proof. First we prove that $\|\|_{\max}$ is a norm.

1. Let $x\in\prod_{i\in I}X_i$ then as $\forall i\in I\ 0\leqslant\|x_i\|_i$ we have $0\leqslant\max(\{\|x_i\|_i|i\in I\})=\|x\|_{\max}$.
2. Let $x\in\prod_{i\in I}X_i$ and $\alpha\in\mathbb{K}$ then

$$\begin{aligned}
\|\alpha\cdot x\|_{\max} &= \max(\{\|\pi_i(\alpha\cdot x)\|_i|i\in I\})\\
&\underset{[\text{example: } 11.166]}{=} \max(\{\|\alpha\cdot\pi_i(x)\|_i|i\in I\})\\
&= \max(\{|\alpha|\cdot\|\pi_i(x)\|_i|i\in I\})\\
&\underset{[\text{theorem: } 10.27]}{=} |\alpha|\cdot\max(\{\|\pi_i(x)\|_i|i\in I\})\\
&= |\alpha|\cdot\|x\|_{\max}
\end{aligned}$$

3. Let $x,y\in\prod_{i\in I}X_i$ then

$$\begin{aligned}
\|x+y\|_{\max} &= \max(\{\|\pi_i(x+y)\|_i|i\in I\})\\
&\underset{[\text{example: } 11.166]}{=} \max(\{\|\pi_i(x)+\pi_i(y)\|_i\in I\})\\
&\leqslant_{[\text{theorem: } 3.66]\wedge\|\|_i \text{ is a norm}} \max(\{\|\pi_i(x)\|_i+\|\pi_i(y)\|_i|i\in I\})\\
&\underset{[\text{theorem: } 10.25]}{=} \max(\{\|\pi_i(x)\|_i|i\in I\})+\max(\{\|\pi_i(y)\|_i|i\in I\})\\
&= \|x\|_{\max}+\|y\|_{\max}
\end{aligned}$$

4. If $\|x\|_{\max}=0$ then $\forall i\in I$ we have $0\leqslant\|\pi_i(x)\|_i\leqslant\|x\|_{\max}=0$, hence $\forall i\in x_i=\pi_i(x)=0$ so that $x=0$.

Finally $\forall i\in I$ and $x,y\in X$ we have that

$$\begin{aligned}
d_{\|\|_{\max}}(x,y) &= \|x-y\|_{\max}\\
&= \max(\{\|\pi_i(x-y)\|_i|i\in I\})\\
&\underset{[\text{example: } 11.166]}{=} \max(\{\|\pi_i(x)-\pi_i(y)\|_i|i\in I\})\\
&= \max(\{d_{\|\|_i}(x,y)|i\in I\})\\
&\underset{[\text{theorem: } 14.79]}{=} d_{\max}(x,y)
\end{aligned}$$

so that we have by [theorem: 14.79] that

$$\mathcal{T}_{d_{\|\|_{\max}}} = \mathcal{T}_{d_{\max}} = \mathcal{T}_{\text{box}} = \mathcal{T}_{\text{product}}$$

□

Example 14.95. $\langle\mathbb{C},\|\rangle$ is a complex normed space where $\|$ is the complex norm [see definition: 10.82]

Proof. This was proved in [theorem: 10.83] □

Example 14.96. $\langle\mathbb{R},\|\rangle$ is a real normed space where $\|$ is defined by

$$\|: \mathbb{R}\times\mathbb{R}\rightarrow\mathbb{R} \text{ by } |x| = \begin{cases} x \text{ if } 0\leqslant x \\ -x \text{ if } x<0 \end{cases}$$

Note 14.97. If $0\leqslant x$ then $x=|x|\Rightarrow x<|x|$ and if $x<0$ then $x<0\leqslant-|x|$ so $\forall x\in X$ we have $x<|x|$.

Proof. We have:

1. If $x\in\mathbb{R}$ then we have either:

   $\boldsymbol{0\leqslant x}$. Then $0\leqslant x=|x|$

   $\boldsymbol{x<0}$. Then $0<-x=|x|$

   so in all cases we have $0\leqslant|x|$

2. Let $\alpha\in\mathbb{R}$ and $x\in\mathbb{R}$ then we have either:

   $\boldsymbol{0\leqslant\alpha\wedge 0\leqslant x}$. Then $0\leqslant\alpha\cdot x$ so that $|\alpha\cdot x|=\alpha\cdot x=|\alpha|\cdot|x|$.

   $\boldsymbol{\alpha<0\wedge 0\leqslant x}$. Then $\alpha\cdot x\leqslant 0$ so that $|\alpha\cdot x|=-(\alpha\cdot x)=(-\alpha)\cdot x=|\alpha|\cdot|x|$.

   $\boldsymbol{0\leqslant\alpha\wedge x<0}$. Then $\alpha\cdot x\leqslant 0$ so that $|\alpha\cdot x|=-(\alpha\cdot x)=\alpha\cdot(-x)=|\alpha|\cdot|x|$.

   $\boldsymbol{\alpha<0\wedge x<0}$. Then $0<\alpha\cdot x$ so that $|\alpha\cdot x|=(-\alpha)\cdot(-x)=|\alpha|\cdot|x|$.

   so in all cases we have

   $$|\alpha\cdot x|=|\alpha|\cdot|x|$$

3. Let $x,y\in\mathbb{R}$ then we have for $x+y$ either:

   $\boldsymbol{0\leqslant x+y}$. Then $|x+y|=x+y\leqslant|x|+|y|$

   $\boldsymbol{x+y<0}$. Then $|x+y|=-(x+y)=(-x)+(-y)\leqslant|-x|+|-y|\underset{(2)}{=}|-1|\cdot|x|+|-1|\cdot|y|=|x|+|y|$,

   so in all cases we have

   $$|x+y|\leqslant|x|+|y|$$

4. If $|x|=0$ then we have for $x\in\mathbb{R}$ either:

   $\boldsymbol{0\leqslant x}$. Then $x=|x|=0$

   $\boldsymbol{x<0}$. Then $-x=|x|=0\Rightarrow x=0$ □

For $\langle\mathbb{R},\|\rangle$ we have the following equivalences for the basis of the normed topology based on the absolute value norm.

Theorem 14.98. *Given the normed space $\langle\mathbb{R},\|\rangle$ we have for the basis $\mathcal{B}_{||}$ of the normed topology $\mathcal{T}_{||}$ that*

$$\begin{aligned}\mathcal{B}_{||} &\underset{\text{def}}{=} \{B_{||}(x,\varepsilon)|x\in\mathbb{R}\wedge\varepsilon\in\mathbb{R}^+\} \\ &= \{]x-\varepsilon,x+\varepsilon[|x\in\mathbb{R}\wedge\varepsilon\in\mathbb{R}^+\} \\ &= \{]a,b[|a,b\in\mathbb{R}\wedge a<b\}\end{aligned}$$

Proof. First we have for $x\in\mathbb{R}$ and $\varepsilon\in\mathbb{R}^+$ that

$$\begin{aligned}y\in B_{||}(x,\varepsilon) &\Leftrightarrow |x-y|<\varepsilon \\ &\Leftrightarrow x-y<\varepsilon\wedge y-x<\varepsilon \\ &\Leftrightarrow x-\varepsilon<y\wedge y<x+\varepsilon \\ &\Leftrightarrow y\in]x-\varepsilon,x+\varepsilon[\end{aligned}$$

proving that

$$B_{||}(x,\varepsilon)=]x-\varepsilon,x+\varepsilon[$$

From this it follows that

$$\{B_{||}(x,\varepsilon)|x\in\mathbb{R}\wedge\varepsilon\in\mathbb{R}^+\}=\{]x-\varepsilon,x+\varepsilon[|x\in\mathbb{R}\wedge\varepsilon\in\mathbb{R}^+\}$$

If $B\in\{]x-\varepsilon,x+\varepsilon[|x\in\mathbb{R}\wedge\varepsilon\in\mathbb{R}^+\}$ then there exist a $x\in\mathbb{R}$ and a $\varepsilon\in\mathbb{R}^+$ such that $B=]x-\varepsilon,x+\varepsilon[$, which, as $x-\varepsilon,x+\varepsilon\in\mathbb{R}\wedge x-\varepsilon<x+\varepsilon$, proves that $B\in\{]a,b[|a,b\in\mathbb{R}\wedge a<b\}$. Hence

$$\{]x-\varepsilon,x+\varepsilon[|x\in\mathbb{R}\wedge\varepsilon\in\mathbb{R}^+\}\subseteq\{]a,b[|a,b\in\mathbb{R}\wedge a<b\} \tag{14.26}$$

If $B\in\{]a,b[|a,b\in\mathbb{R}\wedge a<b\}$ then there exist a $a,b\in\mathbb{R}$ with $a<b$ such that $B=]a,b[$. Take then $\varepsilon=(b-a)/2$ and $x=a+\varepsilon$ then we have

$$\begin{aligned}
y\in]x-\varepsilon,x+\varepsilon[&\Leftrightarrow x-\varepsilon<y\wedge y<x+\varepsilon\\
&\Leftrightarrow (a+\varepsilon)-\varepsilon<y\wedge y<(a+\varepsilon)+\varepsilon\\
&\Leftrightarrow a<y\wedge y<a+2\cdot\varepsilon\\
&\Leftrightarrow a<y\wedge y<a+b-a\\
&\Leftrightarrow a<y\wedge y<b\\
&\Leftrightarrow y\in]a,b[
\end{aligned}$$

proving that $B=]x-\varepsilon,x+\varepsilon[\in\{]x-\varepsilon,x+\varepsilon[|x\in\mathbb{R}\wedge\varepsilon\in\mathbb{R}^+\}$. Hence $\{]a,b[|a,b\in\mathbb{R}\wedge a<b\}\subseteq\{]x-\varepsilon,x+\varepsilon[|x\in\mathbb{R}\wedge\varepsilon\in\mathbb{R}^+\}$, combining this with [eq: 14.26] results in

$$\{]x-\varepsilon,x+\varepsilon[|x\in\mathbb{R}\wedge\varepsilon\in\mathbb{R}^+\}=\{]a,b[|a,b\in\mathbb{R}\wedge a<b\}$$ □

Theorem 14.99. *Let $\langle\mathbb{R},\mathcal{T}_{||}\rangle$ be the topological space generated by the norm $||$, $a,b\in\mathbb{R}$ then the following intervals are open in $\mathcal{T}_{||}$:*

$$\begin{gathered}]a,b[\\ ]-\infty,a[\\ ]a,\infty[\\ ]-\infty,\infty[\end{gathered}$$

the following intervals are closed:

$$\begin{gathered}[a,b]\\ [a,\infty[\\ ]-\infty,a]\\ ]-\infty,\infty[\end{gathered}$$

and the following intervals are neither open or closed:

$$\begin{gathered}[a,b[\ \ \textit{where } a<b\\ ]a,b]\ \ \textit{where } a<b\end{gathered}$$

Proof. Let $a,b\in\mathbb{R}$ then we have:

$]\boldsymbol{a},\boldsymbol{b}[\in\boldsymbol{\mathcal{T}_{||}}$**.** For a,b we have either:

$\boldsymbol{a}\leqslant\boldsymbol{b}$**.** Then by [theorem: 14.98] $]a,b[\in\mathcal{B}_{||}\subseteq\mathcal{T}_{||}$

$\boldsymbol{b}\leqslant\boldsymbol{a}$**.** Then $]a,b[=\varnothing\in\mathcal{T}_{||}$

$]\boldsymbol{-\infty},\boldsymbol{a}[\in\boldsymbol{\mathcal{T}_{||}}$**.** If $x\in]-\infty,a[$ then for $]x-1,a[\in\mathcal{B}_{||}$ we have $x\in]x-1,a[\subseteq]-\infty,a[$ proving by [theorem: 14.32] that $]-\infty,a[\in\mathcal{T}_{||}$.

$]\boldsymbol{a},\boldsymbol{\infty}[$**.** If $x\in]a,\infty[$ then for $]a,x+1[\in\mathcal{B}_{||}$ we have $x\in]a,x+1[\subseteq]a,\infty[$ proving by [theorem: 14.32] that $]-\infty,a[\in\mathcal{T}_{||}$.

$]\boldsymbol{-\infty},\boldsymbol{\infty}[\in\boldsymbol{\mathcal{T}_{||}}$**.** $]-\infty,\infty[$ is open and closed.

$[\boldsymbol{a},\boldsymbol{b}]$ **is closed.** We have

$$\begin{aligned}
x\in\mathbb{R}\setminus[a,b] &\Leftrightarrow x\in\mathbb{R}\wedge\neg(a\leqslant x\wedge x\leqslant b)\\
&\Leftrightarrow x\in\mathbb{R}\wedge(x<a\vee b<x)\\
&\Leftrightarrow (x\in\mathbb{R}\wedge x<a)\vee(x\in\mathbb{R}\wedge b<x)\\
&\Leftrightarrow x\in]-\infty,a[\bigcup]b,\infty[
\end{aligned}$$

proving that $\mathbb{R}\backslash[a,b]=]-\infty,a[\bigcup]b,\infty[\in\mathcal{T}_{||}$ hence $[a,b]$ is closed.

$[a,\infty[$ is closed. We have

$$\begin{aligned}x\in\mathbb{R}\backslash[a,\infty[&\Leftrightarrow x\in\mathbb{R}\wedge\neg(a\leqslant x)\\&\Leftrightarrow x\in\mathbb{R}\wedge x<a\\&\Leftrightarrow x\in]-\infty,a[\end{aligned}$$

proving that $\mathbb{R}\backslash[a,\infty[=]-\infty,a[\in\mathcal{T}_{||}$ so that $[a,\infty[$ is closed.

$]-\infty,a]$ is closed. We have

$$\begin{aligned}x\in\mathbb{R}\backslash]-\infty,a] &\Leftrightarrow x\in\mathbb{R}\wedge\neg(x\leqslant a)\\&\Leftrightarrow x\in\mathbb{R}\wedge a<x\\&\Leftrightarrow x\in]a,\infty[\end{aligned}$$

proving that $\mathbb{R}\backslash]-\infty,a[=]a,\infty[\in\mathcal{T}_{||}$ so that $]-\infty,a]$ is closed.

Assume that $a<b$ then we have

$[a,b[$ is neither open/closed. Assume that $[a,b[$ is open then as $a\leqslant a<b$ we have $a\in[a,b[$, so by [theorem: 14.32] there exist a $]x,y[\in\mathcal{B}_{||}\underset{\text{[theorem: 14.98]}}{=}\{]x,y[|x,y\in\mathbb{R}\wedge x<y\}$ such that $a\in]x,y[\subseteq[a,b[$. Hence $x<a<y$, using [theorem: 10.32] there exist a $z\in\mathbb{R}$ such that $x<z<a<y\Rightarrow z\in]x,y[\subseteq[a,b[$ hence $a\leqslant z$ contradicting $z<a$, hence

$$[a,b[\text{ is not open}$$

Assume that $\mathbb{R}\backslash[a,b[$ is open then as $b\in\mathbb{R}\backslash[a,b[$ we have by [theorem: 14.32] that there exist a $]x,y[\in\mathcal{B}_{||}\underset{\text{[theorem: 14.98]}}{=}\{]x,y[|x,y\in\mathbb{R}\wedge x<y\}$ such that $b\in]x,y[\subseteq\mathbb{R}\backslash[a,b[$ so that $x<b<y$. As also $a<b$ we have that $x\leqslant\max(a,x)<b<y$, using [theorem: 10.32] there exists $z\in\mathbb{R}$ such that

$$x\leqslant\max(a,x)<z<b<y \tag{14.27}$$

so that $z\in]x,y[\subseteq\mathbb{R}\backslash[a,b[$. Hence we have $z<a\vee b\leqslant z$, as $a\leqslant\max(a,x)<z$ we can't have $z<a$ so we must have that $b\leqslant z$ contradicting $z<b$ [see eq: 14.27] hence the assumption is wrong and $\mathbb{R}\backslash[a,b]$ is not open proving that

$$[a,b[\text{ is not closed}$$

$]a,b]$ is neither open/closed. Assume that $]a,b]$ is open then as $a<b\leqslant b$ we have $b\in[a,b[$, so by [theorem: 14.32] there exist a $]x,y[\in\mathcal{B}_{||}\underset{\text{[theorem: 14.98]}}{=}\{]x,y[|x,y\in\mathbb{R}\wedge x<y\}$ such that $b\in]x,y[\subseteq]a,b]$. Hence $x<b<y$, using [theorem: 10.32] there exist a $z\in\mathbb{R}$ such that $x<b<z<y\Rightarrow z\in]x,y[\subseteq]a,b]$ giving $z\leqslant b$ contradicting $b<z$, hence

$$[a,b[\text{ is not open}$$

Assume that $\mathbb{R}\backslash]a,b]$ is open then as $a\in\mathbb{R}\backslash]a,b]$ we have by [theorem: 14.32] that there exist a $]x,y[\in\mathcal{B}_{||}\underset{\text{[theorem: 14.98]}}{=}\{]x,y[|x,y\in\mathbb{R}\wedge x<y\}$ such that $a\in]x,y[\subseteq\mathbb{R}\backslash]a,b]$ so that $x<a<y$. As also $a<b$ we have that $x<a<\min(b,y)\leqslant y$, using [theorem: 10.32] there exists a $z\in\mathbb{R}$ such that

$$x<a<z<\min(b,y)\leqslant y \tag{14.28}$$

so that $z\in]x,y[\subseteq\mathbb{R}\backslash]a,b]$ hence we have $z\leqslant a\vee b<z$, as $z<\min(b,y)<b$ we can't have $b<z$ so we must have $z\leqslant a$ contradicting $a<z$ [see eq: 14.28]. So the assumption is wrong and $\mathbb{R}\backslash]a,b]$ is not open proving that

$$]a,b]\text{ is not closed}$$

□

Corollary 14.100. *Let $\langle\mathbb{R},\mathcal{T}_{||}\rangle$ be the topological space generated by the norm $||$ and I a* ***non empty open generalized interval*** *then I is one of the following sets*

$$\begin{array}{rl}]a,b[& a,b\in\mathbb{R}\text{ where }a<b\text{ [otherwise }I=\varnothing]\\ ]-\infty,a[& a\in\mathbb{R}\\ ]a,\infty[& a\in\mathbb{R}\\ ]-\infty,\infty[& \end{array}$$

Proof. Let I be a **non empty** generalized interval then by [theorem: 3.135] I is of the following forms, where $a, b \in \mathbb{R}$

$\boldsymbol{]a, b[\wedge a < b}$**.** Then by the previous theorem [theorem: 14.99]

$$I =]a, b[\text{ is open}$$

$\boldsymbol{]a, b] \wedge a \leqslant b}$**.** Then by the previous theorem [theorem: 14.99]

$$I =]a, b[\text{ is not open}$$

$\boldsymbol{[a, b[\wedge a \leqslant b}$**.** Then by the previous theorem [theorem: 14.99]

$$I = [a, b[\text{ is not open}$$

$\boldsymbol{[a, b] \wedge a \leqslant b}$**.** As $[a, b] = I \neq \varnothing$ we must have $a \leqslant b$. Assume that $[a, b]$ is open then as $a \leqslant a \leqslant b$ we have $a \in [a, b]$, so by [theorem: 14.32] there exist a $]x, y[\in \mathcal{B}_{||} \underset{\text{[theorem: 14.98]}}{=} \{]x, y[|x, y \in \mathbb{R} \wedge x < y\}$ such that $a \in]x, y[\subseteq [a, b]$, hence $x < a < y$ and using [theorem: 10.32] there exists a $z \in \mathbb{R}$ such that $x < z < a < y$. So $z \in]x, y[\subseteq [a, b]$ resulting in $a \leqslant z$ contradicting $z < a$. Hence the assumption is wrong and we have that

$$I = [a, b] \text{ is not open}$$

$\boldsymbol{[a, \infty[}$**.** Assume that $[a, \infty[$ is open then as $a \in [a, \infty[$ there exist by [theorem: 14.32] a $]x, y[\in \mathcal{B}_{||} \underset{\text{[theorem: 14.98]}}{=} \{]x, y[|x, y \in \mathbb{R} \wedge x < y\}$ such that $a \in]x, y[\subseteq [a, \infty[$, hence $x < a < y$. Using [theorem: 10.32] there exist a $z \in \mathbb{R}$ such that $x < z < a < y$ hence $z \in]x, y[\subseteq [a, \infty[$ giving $a \leqslant z$ which contradict $z < a$, hence the assumption is wrong and

$$I = [a, \infty[\text{ is not open}$$

$\boldsymbol{]a, \infty[}$**.** Then by the previous theorem [theorem: 14.99]

$$I =]a, \infty[\text{ is open}$$

$\boldsymbol{]-\infty, a]}$**.** Assume that $]-\infty, a]$ is open then as $a \in]-\infty, a]$ there exist by [theorem: 14.32] a $]x, y[\in \mathcal{B}_{||} \underset{\text{[theorem: 14.98]}}{=} \{]x, y[|x, y \in \mathbb{R} \wedge x < y\}$ such that $a \in]x, y[\subseteq]-\infty, a]$, hence $x < a < y$. Using [theorem: 10.32] there exist a $z \in \mathbb{R}$ such that $x < a < z < y$ hence $z \in]x, y[\subseteq]-\infty, a]$ giving $z \leqslant a$ which contradict $a < z$, hence the assumption is wrong and

$$I =]-\infty, a] \text{ is not open}$$

$\boldsymbol{]-\infty, a[}$**.** Then by the previous theorem [theorem: 14.99]

$$I =]-\infty, a[\text{ is open}$$

$\boldsymbol{]-\infty, \infty[}$**.** Then by the previous theorem [theorem: 14.99]

$$I =]-\infty, \infty[\text{ is open}$$ □

Corollary 14.101. *Let $\langle \mathbb{R}, \mathcal{T}_{||} \rangle$ be the topological space generated by the norm $||$ and I a* ***non empty closed generalized interval*** *then I is one of the following sets.*

$$\begin{array}{rl} [a, b] & \textit{where } a, b \in \mathbb{R} \textit{ with } a \leqslant b \textit{ [otherwise } I = \varnothing \textit{]} \\ [a, \infty[& \textit{where } a \in \mathbb{R} \\ ]-\infty, a] & \textit{where } a \in \mathbb{R} \end{array}$$

Proof. Let I be a **non empty** generalized interval then by [theorem: 3.135] I is of the following forms, where $a, b \in \mathbb{R}$

$\boldsymbol{]a, b[\wedge a < b}$**.** Assume that $]a, b[$ is closed then $\mathbb{R} \backslash]a, b[$ is open. As $a \notin]a, b[$ hence $a \in \mathbb{R} \backslash]a, b[$, we have by [theorems: 14.32, 14.98] that there exists $r, s \in \mathbb{R}$ such that $a \in]r, s[\subseteq \mathbb{R} \backslash]a, b[$ so that $r < a < s$ and thus $a < \min(s, b)$. Using [theorem: 10.32] there exist a $x \in \mathbb{R}$ such that $r < a < x < \min(s, b)$ so that $r < x < s \wedge a < x < b$ hence $x \in]r, s[\bigcap]a, b[$ contradicting $]r, s[\subseteq \mathbb{R} \backslash]a, b[$. Hence the assumption is wrong and $]a, b[$ is not closed.

$\boldsymbol{]a, b] \wedge a \leqslant b}$**.** Then by the previous theorem [theorem: 14.99]

$$I =]a, b[\text{ is not closed}$$

$[a,b[\wedge a\leqslant b$. Then by the previous theorem [theorem: 14.99]

$$I=[a,b[\text{ is not closed}$$

$[a,b]\wedge a\leqslant b$. Then by the previous theorem [theorem: 14.99] $I=[a,b]$ is closed.

$[a,\infty[$. Then by the previous theorem [theorem: 14.99] $I=[a,\infty[$ is closed.

$]a,\infty[$. Assume that $]a,\infty[$ is closed then $\mathbb{R}\backslash]a,\infty[$ is open. As $a\notin]a,\infty[$ we have $a\in\mathbb{R}\backslash]a.\infty[$ so that by [theorems: 14.32, 14.98] there exists $r,s\in\mathbb{R}$ with $a\in]r,s[\subseteq\mathbb{R}\backslash]a,\infty[$. Hence $r<a<s$ so that by [theorem: 10.32] there exist a $x\in\mathbb{R}$ such that $r<a<x<s$, hence $x\in]r,s[\bigcap]a,\infty[$ contradicting $]r,s[\subseteq\mathbb{R}\backslash]a,\infty[$. So the assumption is wrong and $]a,\infty[$ is not closed.

$]-\infty,a]$. Then by the previous theorem [theorem: 14.99] $I=]-\infty,a]$ is closed.

$]-\infty,a[$. Assume that $]-\infty,a[$ is closed then $\mathbb{R}\backslash]-\infty,a[$ is open. As $a\notin]-\infty,a[$ we have $a\in\mathbb{R}\backslash]-\infty.a[$ so that by [theorems: 14.32, 14.98] there exist $r,s\in\mathbb{R}$ with $a\in]r,s[\subseteq\mathbb{R}\backslash]-\infty,a[$. Hence $r<a<s$ so that by [theorem: 10.32] there exist a $x\in\mathbb{R}$ such that $r<x<a<s$, hence $x\in]r,s[\bigcap]-\infty,a[$ contradicting $]r,s[\subseteq\mathbb{R}\backslash]-\infty,a[$. So the assumption is wrong and $]-\infty,a[$ is not closed.

$]-\infty,\infty[$. Then by the previous theorem [theorem: 14.99]

$$I=]-\infty,\infty[\text{ is closed}$$

□

Note 14.102. The condition that $I\neq\varnothing$ in the above theorem is essential because if $I=\varnothing$ then I is open and closed and either of the form:

$$\begin{array}{lll}]a,b[& \text{where} & b\leqslant a \\ [a,b] & \text{where} & b<a \\ [a,b[& \text{where} & b<a \\ ]a,b] & \text{where} & b<a \end{array}$$

so that for the empty interval there are more forms then stated in the previous two corollaries.
We prove now that every open set in $\langle\mathbb{R}.\mathcal{T}_{||}\rangle$ is a countable union of open intervals.

Theorem 14.103. *Given $\langle\mathbb{R},\mathcal{T}_{||}\rangle$ the set of real numbers endowed with the topology generated by the absolute value norm and $U\in\mathcal{T}_{||}$ a open set then there exists a $\mathcal{U}\subseteq\mathcal{T}_{||}$ where $\mathcal{U}$ is countable and $\forall I\in\mathcal{U}$ we have that I is a open generalized interval such that*

$$U=\bigcup_{I\in\mathcal{U}} I$$

Proof. Define the relation $\sim\in U\times U$ by $x\sim y$ iff $]\min(x,y),\max(x,y)[\subseteq U$ then we have

reflexivity. $\forall x\in U$ $]\min(x,x),\max(x,x)[=]x,x[=\varnothing\subseteq U$ so that $x\sim x$

symmetry. If $x\sim y$ then $]\min(x,y),\max(x,y)[\subseteq U$ so that

$$]\min(y,x),\max(y,x)[=]\min(x,y),\max(x,y)]\subseteq U$$

giving $y\sim x$.

transitivity. If $x\sim y$ and $y\sim z$ then we have

$$]\min(x,y),\max(x,y)[\subseteq U \text{ and }]\min(y,z),\max(y,z)[\subseteq U \tag{14.29}$$

now for $x,y,z\in U$ we have either:

$x\leqslant y\leqslant z$. Then by [eq: 14.29] we have $]x,y[\subseteq U\wedge]y,z[\subseteq U$ and as $y\in U$

$$]\min(x,z),\max(x,)[=]x,z[=]x,y[\bigcup\{y\}\bigcup]y,z[\subseteq U$$

proving that $x\sim z$.

$x\leqslant z\leqslant y$. Then by [eq: 14.29] we have $]x,y[\subseteq U\wedge]z,y[\subseteq U$

$$]\min(x,z),\max(x,z)[=]x,z[\subseteq]x,y[\subseteq U$$

proving that $x\sim z$.

$y \leqslant x \leqslant z$. Then by [eq: 14.29] we have $]y,x[\subseteq U \wedge]y,z[\subseteq U$

$$]\min(x,z),\max(x,z)[=]x,z[\subseteq]y,z[\subseteq U$$

proving that $x \sim z$.

$y \leqslant z \leqslant x$. Then by [eq: 14.29] we have $]y,x[\subseteq U \wedge]y,z[\subseteq U$

$$]\min(x,z),\max(x,z)[=]z,x[\subseteq]y,x[\subseteq U$$

proving that $x \sim z$.

$z \leqslant x \leqslant y$. Then by [eq: 14.29] we have $]x,y[\subseteq U \wedge]z,y[\subseteq U$

$$]\min(x,z).\max(x,z)[=]z,x[\subseteq]z,y[\subseteq U$$

proving that $x \sim z$.

$z \leqslant y \leqslant x$. Then by [eq: 14.29] we have $]y,x[\subseteq U \wedge]z,y[\subseteq U$ and as $y \in U$ we have that

$$]\min(x,z),\max(x,z)[=]z,x[=]z,y[\bigcup \{y\}\bigcup]y,x[\subseteq U$$

proving that $x \sim z$.

So in all cases we have $x \sim z$.

Hence we have that

$$\sim \text{ is a equivalence relation}$$

Define then $\mathcal{U}=\{\sim[x]|x\in U\}=I/\sim$ the set of equivalence classes defined by $\sim$ then we have using [theorem: 3.18] that

$$U=\bigcup_{I\in\mathcal{U}} I \text{ and } \forall I,J\in\mathcal{U} \text{ with } I\neq J \;\; I\bigcap J=\varnothing \tag{14.30}$$

Let $I\in\mathcal{U}$ then there exist a $x\in U$ such that $I=\sim[x]$. Let $y\in\sim[x]\subseteq U$ then by [theorem: 14.32] there exist a $]a,b[\in\mathcal{B}_{||} \underset{\text{[theorem: 14.98]}}{=} \{]x,y[|x,y\in\mathbb{R}\wedge x<y\}$ such that $y\in]a,b[\subseteq U$ then $\forall z\in]a,b[$ we have either:

$y \leqslant z$. Then $a<y\leqslant z<b$ so that $]\min(y,z),\max(y,z)[=]y,z[\subseteq]a,b[\subseteq U$

$z < y$. Then $a<z<y<b$ so that $]\min(y,z),\max(y,z)[=]z,y[\subseteq]a,b[\subseteq U$

proving that $y\sim z$. From $y\in\sim[x]$ we have $y\sim x$ so by transitivity and symmetry that $z\sim x$ hence $z\in\sim[x]=I$ proving that $]a,b[\subseteq I$ or as $y\in]a,b[$ that $y\in]a,b[\subseteq I$. To summarize

$$\forall y\in I=\sim[x] \text{ there exist a }]a,b[\in\mathcal{B}_{||} \text{ such that } y\in]a,b[\subseteq I$$

proving by [definition: 14.31] that

$$I=\sim[x] \text{ is open} \tag{14.31}$$

Let $r,s\in I=\sim[x]$ with $r<s$ then $r\sim x\wedge s\sim x$ so that by symmetry and transitivity we have $s\sim r$ so that $]r,s[=]\min(r,s),\max(r,s)[\subseteq U$. If $t\in]r,s[\Rightarrow r<t<s$ then

$$]\min(t,r),\max(t,r)[=]r,t[\subseteq]r,s[\subseteq U$$

proving that $t\sim r$, as $r\sim x$ it follows that $t\sim x$ so that $t\in\sim[x]=I$. Hence

$$\forall r,s\in I \text{ with } r<s \text{ we have that }]r,s[\subseteq I$$

which by [definition: 3.133] proves that

$$I \text{ is a generalized interval} \tag{14.32}$$

As by [eq: 14.31] I is open it follows from [corollary: 14.100] that

$$\forall I\in\mathcal{U} \text{ we have } I=]a,b[\text{ or } I=]a,\infty[\text{ or } I=]-\infty,a[\text{ or } I=]-\infty,\infty[\text{ where } a,b\in\mathbb{R}$$

Now if $I\in\mathcal{U}$ then $\exists x\in U$ so that $I=\sim[x]\neq\varnothing$, by the above we have the following cases to consider:

$I=]a,b[$. Then as $I\neq\varnothing$ we have $a<b$ so using the density of $\mathbb{Q}$ in $\mathbb{R}$ [see theorem: 10.32] there exist a $q\in\mathbb{Q}$ such that $a<q<b$, hence $q\in]a,b[=I$ so that

$$\mathbb{Q}\bigcap I\neq\varnothing$$

$I =]-\infty, a[$. Then as $a-1<a$ we have by the density of $\mathbb{Q}$ in $\mathbb{R}$ [see theorem: 10.32] that there exist a $q\in\mathbb{Q}$ such that $a-1<q<a$ hence $q\in]-\infty,a[=I$ proving

$$\mathbb{Q}\bigcap I\neq\varnothing$$

$I =]a, \infty[$. Then as $a<a+1$ we have by the density of $\mathbb{Q}$ in $\mathbb{R}$ [see theorem: 10.32] that there exist a $q\in\mathbb{Q}$ such that $a<q<a+1$ hence $q\in]a,\infty[=I$ proving

$$\mathbb{Q}\bigcap I\neq\varnothing$$

$I =]-\infty, \infty[$. Then as $0\in\mathbb{Q}\subseteq\mathbb{R}=]-\infty,\infty[$ we have

$$\mathbb{Q}\bigcap I\neq\varnothing$$

So we have

$$\forall I\in\mathcal{U} \text{ we have that } \mathbb{Q}\bigcap I\neq\varnothing \text{ so that } \left\{\mathbb{Q}\bigcap I\right\}_{I\in\mathcal{U}}\subseteq\mathcal{P}(\mathbb{R}) \text{ is a family of non empty subsets of } \mathbb{R}$$

Using a consequence of the Axiom of Choice [see theorem 3.103] there exist a function

$$f:\mathcal{U}\rightarrow\bigcup_{I\in\mathcal{U}}\left(\mathbb{Q}\bigcap I\right) \text{ such that } f(I)\in\mathbb{Q}\bigcap I$$

Let $I,J\in\mathcal{U}$ such that $f(I)=f(J)$. As $f(I)\in\mathbb{Q}\bigcap I$ and $f(J)\in\mathbb{Q}\bigcap J$ we have as $f(I)=f(J)$ that $f(I)\in(\mathbb{Q}\bigcap I)\bigcap(\mathbb{Q}\bigcap J)=\mathbb{Q}\bigcap(I\bigcap J)$ so that $I\bigcap J\neq\varnothing$. Then by [eq: 14.30] we must have $I=J$ which proves as $\bigcup_{I\in\mathcal{U}}(\mathbb{Q}\bigcap I)\subseteq\mathbb{Q}$

$$f:\mathcal{U}\rightarrow\mathbb{Q} \text{ is injective}$$

Now as by [theorem: 10.5] $\mathbb{Q}$ is denumerable we have by [theorem: 6.69] that

$$\mathcal{U} \text{ is countable and by [eqs: 14.30, 14.31,14.32] } U=\bigcup_{I\in\mathcal{U}} I \text{ and } \forall I\in\mathcal{U}\ I \text{ is a open generalized interval}$$ □

Corollary 14.104. *Given $\langle\mathbb{R},\mathcal{T}_{||}\rangle$ the set of reals endowed with the topology generated by the absolute value norm and $U\in\mathcal{T}_{||}$ a open set then there exist a countable set $\mathcal{I}$ and a family*

$$\{]a_i,b_i[\}_{i\in\mathcal{I}}\subseteq\mathcal{T}_{||} \textit{ such that } U=\bigcup_{i\in\mathcal{I}}]a_i,b_i[$$

Proof. Using the previous theorem [theorem: 14.103] there exists a countable set $\mathcal{U}\subseteq\mathcal{T}_{||}$ such that $\forall I\in\mathcal{U}\ I$ is open generalized interval and

$$U=\bigcup_{I\in\mathcal{U}} I \tag{14.33}$$

So let $I\in\mathcal{U}$ then by [corollary: 14.100] I is either

$I =]a, b[$. Define then $\mathcal{I}_I=\{1\}$ and $\{]a_{(I,i)},b_{(I,i)}[\}_{i\in\mathcal{I}_I}$ by $a_1=a\wedge b_1=b$ so that

$$\mathcal{I}_i \text{ is countable [because it is finite] and } I=\bigcup_{i\in\mathcal{I}_I}]a_{(I,i)},b_{(I,i)}[$$

$I =]-\infty, a[$. Take $\mathcal{I}_I=\mathbb{N}_0$ and $\{]a_{(I,i)},b_{(I,i)}[\}_{i\in\mathcal{I}_I}=\{]a_{(I,i)},b_{(I,i)}[\}_{i\in\mathbb{N}_0}$ by $a_{(I,i)}=-i\wedge b_{(I,i)}=a$. Then $\forall i\in\mathcal{I}_I$ we have $\forall x\in]a_{(I,i)},b_{(I,i)}[=]-i,a[\Rightarrow -i<x<a\Rightarrow x\in]-\infty,a[=I$ proving that

$$\bigcup_{i\in\mathcal{I}_I}]a_{(I,i)},b_{(I,i)}[\subseteq I \tag{14.34}$$

Further if $x\in I$ then $x<a$ and by a consequence of the Archimedean Property of the real numbers [see theorem: 10.30] there exist a $n\in\mathbb{N}_0=\mathcal{I}_I$ such that $-n<x$ hence $x\in]-n,a[=]a_n,b_n[\subseteq\bigcup_{i\in\mathcal{I}_I}]a_{(I,i)},b_{(I,i)}[$. Proving that $I\subseteq\bigcup_{i\in\mathcal{I}_I}]a_{(I,i)},b_{(I,i)}[$ which combined with 14.34] proves that

$$\mathcal{I}_I \text{ is countable and } I=\bigcup_{i\in\mathcal{I}_I}]a_{(I,i)},b_{(I,i)}[$$

$I =]a, \infty[$. Take $\mathcal{I}_I=\mathbb{N}_0$ and $\{]a_{(I,i)},b_{(I,i)}[\}_{i\in\mathcal{I}_I}=\{]a_{(I,i)},b_{(I,i)}[\}_{i\in\mathbb{N}_0}$ by $a_{(I,i)}=a\wedge b_{(I,i)}=n$. Then $\forall i\in\mathcal{I}_I$ we have $\forall x\in]a_{(I,i)},b_{(I,i)}[=]a,i[\Rightarrow a<x<i\Rightarrow x\in]a,\infty[$ proving that

$$\bigcup_{i\in\mathcal{I}_I}]a_{(I,i)},b_{(I,i)}[\subseteq I \tag{14.35}$$

Further if $x \in I$ then $a < x$ then by a consequence of the Archimedean Property of the real numbers [see theorem: 10.30] there exist a $n \in \mathbb{N}_0 = \mathcal{I}_I$ such that $x < n$ hence $x \in]a, n[=]a_n, b_n[\subseteq \bigcup_{i \in \mathcal{I}_I}]a_{(I,i)}, b_{(I,i)}[$. Proving that $I \subseteq \bigcup_{i \in \mathcal{I}_I}]a_{(I,i)}, b_{(I,i)}[$ which combined with 14.35] proves that

$$\mathcal{I}_I \text{ is countable and } I = \bigcup_{i \in \mathcal{I}_I}]a_{(I,i)}, b_{(I,i)}[$$

$\boldsymbol{I =]-\infty, \infty[}$**.** Then $I = \mathbb{R}$. Take $\mathcal{I}_I = \mathbb{N}_0$ and $\{]a_{(I,i)}, b_{(I,i)}[\}_{i \in \mathcal{I}_I} = \{]a_{(I,i)}, b_{(I,i)}[\}_{i \in \mathbb{N}_0}$ by $a_{(I,i)} = -i \wedge b_{(I,i)} = i$. Then $\forall i \in \mathcal{I}_I$ we have $]a_{(I,i)}, b_{(I,i)}[\subseteq \mathbb{R} = I$ so that

$$\bigcup_{i \in \mathcal{I}_I}]a_{(I,i)}, b_{(I,i)}[\subseteq I \tag{14.36}$$

Further if $x \in \mathbb{R}$ then by [theorem: 10.30] there exist a $n \in \mathbb{N}_0 = \mathcal{I}_I$ such that $|x| < n$ hence $x < |x| < n$ and $-x < |x| < n \Rightarrow -n < x$ proving that $-n < x < n \Rightarrow x \in]-n, n[=]a_n, b_n[\subseteq \bigcup_{i \in \mathcal{I}_I}]a_{(I,i)}, b_{(I,i)}[$. Hence $I \subseteq \bigcup_{i \in \mathcal{I}_I}]a_{(I,i)}, b_{(I,i)}[$ which combined with [eq: 14.36] proves that

$$\mathcal{I}_I \text{ is countable and } I = \bigcup_{i \in \mathcal{I}_I}]a_{(I,i)}, b_{(I,i)}[$$

So we have that

$$\forall I \in \mathcal{U} \text{ there exists a countable } \mathcal{I}_I \text{ and } \{]a_{(I,i)}, b_{(I,i)}[\}_{i \in \mathcal{I}_i} \subseteq \mathcal{T}_{||} \text{ so that } I = \bigcup_{i \in \mathcal{I}_I}]a_{(I,i)}, b_{(I,i)}[\tag{14.37}$$

Now we have

$$\begin{aligned}
x \in \bigcup_{r \in \bigcup_{I \in \mathcal{U}} \{I\} \times \mathcal{I}_I}]a_r, b_r[\quad &\Rightarrow \quad x \in \bigcup_{r \in \bigcup_{I \in \mathcal{U}} \{I\} \times \mathcal{I}_I}]a_{(r_1,r_2)}, b_{(r_1,r_2)}[\\
&\Rightarrow \quad \exists r \in \bigcup_{I \in \mathcal{U}} \{I\} \times \mathcal{I}_I \text{ such that } x \in]a_{(r_1,r_2)}, b_{(r_1,r_2)}[\\
&\Rightarrow \quad \exists I \in \mathcal{U} \text{ such that } r \in \{I\} \times \mathcal{T}_I \text{ such that } x \in]a_{(r_1,r_2)}, b_{(r_1,r_2)}[\\
&\underset{\text{[eq: 14.37]}}{\Rightarrow} \quad x \in I \\
&\Rightarrow \quad x \in \bigcup_{I \in \mathcal{U}} I
\end{aligned}$$

proving that

$$\bigcup_{r \in \bigcup_{I \in \mathcal{U}} \{I\} \times \mathcal{I}_I}]a_r, b_r[\subseteq \bigcup_{I \in \mathcal{U}} I \tag{14.38}$$

and

$$\begin{aligned}
x \in \bigcup_{I \in \mathcal{U}} I \quad &\Rightarrow \quad \exists I \in \mathcal{U} \text{ such that } x \in I \\
&\underset{\text{[eq: 14.37]}}{\Rightarrow} \quad \exists i \in \mathcal{I}_I \text{ such that } x \in]a_{(I,i)}, b_{(I,i)}[\\
&\Rightarrow \quad x \in]a_{(r_1,r_2)}, b_{(r_1,r_2)}[\text{ where } r = (I,i) \in \{I\} \times \mathcal{T}_I \\
&\Rightarrow \quad x \in]a_r, b_r[\text{ where } r = (I,i) \in \{I\} \times \mathcal{T}_I \subseteq \bigcup_{I \in \mathcal{U}} \{I\} \times \mathcal{I}_I \\
&\Rightarrow \quad x \in \bigcup_{r \in \bigcup_{I \in \mathcal{U}} \{I\} \times \mathcal{I}_I}]a_r, b_r[
\end{aligned}$$

proving that $\bigcup_{I \in \mathcal{U}} I \subseteq \bigcup_{r \in \bigcup_{I \in \mathcal{U}} \{I\} \times \mathcal{I}_I}]a_r, b_r[$. Combining this with [eqs: 14.33, 14.38] results in

$$U = \bigcup_{I \in \mathcal{U}} I = \bigcup_{r \in \bigcup_{I \in \mathcal{U}} \{I\} \times \mathcal{I}_I}]a_r, b_r[\tag{14.39}$$

Now using [theorem: 6.71] together with [eq: 14.37] we have $\forall I \in \mathcal{U}$ that $\{I\} \times \mathcal{I}_I$ is countable. As $\mathcal{U}$ is countable it follows from [theorem: 6.70] that $\bigcup_{I \in \mathcal{U}} \{I\} \times \mathcal{I}_I$ is countable. Hence if we take $\mathcal{I} = \bigcup_{I \in \mathcal{U}} \{I\} \times \mathcal{I}_I$ we have that

$$\mathcal{I} \text{ is countable and } U \underset{\text{[eq: 14.39]}}{=} \bigcup_{r \in \mathcal{I}}]a_r, b_r[\qquad \square$$

Theorem 14.105. *Given $\langle \mathbb{R}, \mathcal{T}_{||} \rangle$ the set of real numbers endowed with the topology generated by the absolute value norm and $S \subseteq \mathbb{R}$ a non empty closed set then we have:*

1. *If* $\sup(S)$ *exist then* $\sup(S) \in S$

2. *If* $\inf(S)$ *exist then* $\inf(S) \in S$

Proof.

1. Assume that $\sup(S) \notin S$ then $\sup(S) \in X \setminus S \in \mathcal{T}_{||}$, hence by [theorems: 14.93, 14.98] there exist a $\varepsilon \in \mathbb{R}^+$ such that $\sup(S) \in]\sup(S) - \varepsilon, \sup(S) + \varepsilon[\subseteq X \setminus S$. As $\sup(S) - \varepsilon < \sup(S)$ there exist a $s \in S$ such that $\sup(S) - \varepsilon < s \leqslant \sup(S) < \sup(S) + \varepsilon$ so that $s \in X \setminus S$ contradicting $s \in S$. Hence we must have $\sup(S) \in S$.

2. Assume that $\inf(S) \notin S$ then $\inf(S) \in X \setminus S \in \mathcal{T}_{||}$, hence by [theorems: 14.93, 14.98] then there exist a $\varepsilon \in \mathbb{R}^+$ such that $\inf(S) \in]\inf(S) - \varepsilon, \inf(S) + \varepsilon[\subseteq X \setminus S$. As $\inf(S) < \inf(S) + \varepsilon$ there exist a $s \in S$ such that $\inf(S) - \varepsilon < \inf(S) \leqslant s < \inf(S) + \varepsilon$ so that $s \in X \setminus S$ contradicting $s \in S$. Hence we must have $\inf(S) \in S$. □

The concept of a bounded set [see definition: 14.75] in a metric space becomes simpler if the metric space happens to be a normed space.

Theorem 14.106. *Let* $\langle X, \|\|\rangle$ *be a pseudo normed space,* $A \subseteq X$ *then*

$$A \text{ is bounded} \Leftrightarrow \exists M \in \mathbb{R}^+ \text{ such that } \forall x \in X \text{ we have } \|x\| \leqslant M$$

Proof.

⇒. If A is a bounded set then we have either:

$A = \varnothing$. Then $\forall x \in A$ we have vacuously that $\|x\| \leqslant 1$.

$A \neq \varnothing$. Then $\exists x_0 \in A$. As A is bounded there exist a $N \in \mathbb{R}$ such that $\forall x, y \in A$ we have $\|x - y\| = d_{\|\|}(x, y) \leqslant N$. Take $M = N + \|x_0\|$ then if $z \in A$ we have $\|z\| = \|z - x_0 + x_0\| \leqslant \|z - x_0\| + \|x_0\| \leqslant N + \|x_0\| = M$

⇐. Let $\mathbb{N} \in \mathbb{R}$ such that $\forall x \in A$ we have $\|x\| \leqslant N$ and set $M = 2 \cdot N$. Then $\forall x, y \in A$ we have $d_{\|\|}(x, y) = \|x - y\| \leqslant \|x\| + \|y\| = N + N = 2 \cdot N = M$ □

Example 14.107. Let $\langle \mathbb{R}, ||\rangle$ be the normed space of the real numbers then we have:

1. $\forall a, b \in \mathbb{R}$ with $a \leqslant b$ we have that
   a. $[a, b]$ is bounded
   b. $[a, b[$ is bounded
   c. $]a, b]$ is bounded
   d. $]a, b[$ is bounded
2. $\forall a \in \mathbb{R}$ we have that
   a. $[a, \infty[$ is not bounded
   b. $]a, \infty[$ is not bounded
   c. $]-\infty, a]$ is not bounded
   d. $]-\infty, a[$ is not bounded

Proof.

1. Let $M = \max(|a|, |b|)$ then
   a. Then for $x \in [a, b]$ we have either:
      $0 \leqslant x$. Then $|x| = x \leqslant b \leqslant |b| \leqslant \max(|a|, |b|) = M$
      $x < 0$. Then $|x| = -x \leqslant -a \leqslant |a| \leqslant \max(|a|, |b|) = M$
   b. Then for $x \in [a, b[$ we have either:
      $0 \leqslant x$. Then $|x| = x < b \leqslant |b| \leqslant \max(|a|, |b|) = M$
      $x < 0$. Then $|x| = -x \leqslant -a \leqslant |a| \leqslant \max(|a|, |b|) = M$
   c. Then for $x \in]a, b]$ we have either:
      $0 \leqslant x$. Then $|x| = x \leqslant b \leqslant |b| \leqslant \max(|a|, |b|) = M$

$x<0$. Then $|x|=-x<-a\leqslant|a|\leqslant\max(|a|,|b|)=M$

d. Then for $x\in]a,b[$ we have either:

$0\leqslant x$. Then $|x|=x<b\leqslant|b|\leqslant\max(|a|,|b|)=M$

$x<0$. Then $|x|=-x<-a\leqslant|a|\leqslant\max(|a|,|b|)=M$

2.

a. Assume that $[a,\infty[$ is bounded then there exist a M such that $\forall x\geqslant a$ we have $|x|\leqslant M$. As $a\leqslant|a|$ it follows that $|a|\leqslant M$ hence $a<|a|\leqslant M<M+1$ so that $|M+1|=M+1\leqslant M$, a contradiction. Hence $[a,\infty[$ is not bounded.

b. Assume that $]a,\infty[$ is bounded then there exist a M such that $\forall x>a$ we have $|x|\leqslant M$. As $a\leqslant|a|<|a|+1$ it follows that $|a|+1\leqslant M$ hence $a<|a|+1\leqslant M<M+1$ so that $|M+1|=M+1\leqslant M$, a contradiction. Hence $]a,\infty[$ is not bounded.

c. Assume that $]-\infty,a]$ is bounded then there exist a M such that $\forall x\leqslant a$ we have that $|x|\leqslant M$. As $-|a|\leqslant a$ it follows that $|a|=|-|a||\leqslant M<M+1$ so that $-(M+1)<-|a|\leqslant a$, hence $M+1=|-(M+1)|\leqslant M$, a contradiction. So $]-\infty,a]$ is not bounded.

d. Assume that $]-\infty,a[$ is bounded then there exist a M such that $\forall x<a$ we have that $|x|\leqslant M$. As $-|a|-1<-|a|\leqslant a$ it follows that $|a|+1=|-|a|-1|\leqslant M<M+1$ so that $-(M+1)<-|a|-1<-|a|\leqslant a$, hence $M+1=|-(M+1)|\leqslant M$, a contradiction. So $]-\infty,a]$ is not bounded. $\square$

Theorem 14.108. *Let $\langle X,\|\|\rangle$ be a pseudo normed space with norm topology $\mathcal{T}_{\|\|}$ and $A\subseteq X$ then for $\|\|_{|A}$ defined by*

$$\|\|_{|A|}: A\rightarrow\mathbb{R} \text{ defined by } \|x\|_{|A}=\|x\|$$

we have

1. *$\langle A,\|\|_{|A}\rangle$ is a pseudo normed space.*
2. *If $\langle A,\|\|\rangle$ is a normed space $\langle A,\|\|_{|A}\rangle$ is also a normed space*
3. *$\mathcal{T}_{\|\|_{|A}}=(\mathcal{T}_{\|\|})_{|A}$*

Proof.

1. We have:

   a. $\forall x\in A$ we have $0\leqslant\|x\|=\|x\|_{|A}$

   b. $\forall x\in A$ we have $\forall\alpha\in\mathbb{R}$ [or $\alpha\in\mathbb{C}$] that $\|\alpha\cdot x\|_{|A}=|\alpha|\|x\|=|a|\cdot\|x\|_{|A}$

   c. $\forall x,y\in A$ we have $\|x+y\|_{|A}=\|x+y\|\leqslant\|x\|+\|y\|\leqslant\|x\|_{|A}+\|y\|_{|A}$

2. Let $x\in A$ such that $\|x\|_{|A}=0$ then $\|x\|=0$ hence as $\|\|$ is a norm we have $x=0$.

3. Let $x,y\in A$ then $(d_{\|\|})_{|A\times A}(x,y)=d_{\|\|}(x,y)=\|x-y\|=\|x-y\|_{|A}=d_{\|\|_{|A}}(x,y)$ proving that

$$(d_{\|\|})_{|A}=d_{\|\|_{|A}} \tag{14.40}$$

   So we have

$$\mathcal{T}_{\|\|_{|A}}\underset{\text{def}}{=}\mathcal{T}_{d_{\|\|_{|A}}}\underset{[\text{eq: } 14.40]}{=}\mathcal{T}_{(d_{\|\|})_{|A\times A}}\underset{[\text{theorem: } 14.66]}{=}(\mathcal{T}_{d_{\|\|}})_{|A}\underset{\text{def}}{=}(\mathcal{T}_{\|\|})_{|A}$$ $\square$

Theorem 14.109. *Let $\langle X,\|\|\rangle$ be a pseudo normed space, $M\subseteq X$ then $\forall\alpha\in\mathbb{K}$ with $\alpha\neq 0$ and $\forall x\in X$ we have:*

1. *$\overline{\alpha\cdot M}=\alpha\cdot\overline{M}$*
2. *If M is closed in $\mathcal{T}_{\|\|}$ then $\alpha\cdot M$ is closed in $\mathcal{T}_{\|\|}$*
3. *If M is open in $\mathcal{T}_{\|\|}$ then $\alpha\cdot M$ is open in $\mathcal{T}_{\|\|}$*
4. *If M is open in $\mathcal{T}_{\|\|}$ then $x+M$ is open in $\mathcal{T}_{\|\|}$*
5. *If M is open in $\mathcal{T}_{\|\|}$ and $A\subseteq X$ then $A+M$ is open in $\mathcal{T}_{|M}$*
6. *If $x\in X$ then $x+\overline{M}\subseteq\overline{x+M}$*
7. *$\overline{M}-\overline{M}\subseteq\overline{M-M}$*

See [definition: 11.69] for the definition of $\alpha\cdot M$, $x+M$, $A+M$ and $M-M$

Proof.

1. Let $x\in\overline{\alpha\cdot M}$ then as $\alpha\neq 0$ we can define $z=1/\alpha\cdot x$ so that $x=\alpha\cdot z$. Assume that U is a open set such that $z\in U$ then by [theorem: 14.64] there exist a $\delta\in\mathbb{R}^+$ such that
$$z\in B_{\|\|}(z,\delta)\subseteq U \tag{14.41}$$
Further we have trivially $x\in B_{\|\|}(x,|\alpha|\cdot\delta)$ which is open so by [theorem: 14.28] and $x\in\overline{\alpha\cdot M}$ we have that $B_{\|\|}(x,|\alpha|\cdot\delta)\bigcap(\alpha\cdot M)\neq\varnothing$, hence there exist a $m\in M$ such that $\alpha\cdot m\in B_{\|\|}(x,|\alpha|\cdot\delta)$, from which it follows that $\|x-\alpha\cdot m\|<|\alpha|\cdot\delta$ or as $x=\alpha\cdot z$ that $|\alpha|\|z-m\|=\|\alpha\cdot z-\alpha\cdot m\|=\|x-\alpha\cdot m\|<|\alpha|\cdot\delta$ from which we conclude, as $|\alpha|\neq 0$, that $\|z-m\|<\delta$ or that $m\in B_{\|\|}(z,\delta)\subseteq_{[\text{eq: }14.41]}U$. Hence $U\bigcap M\neq\varnothing$ which, as $z\in U$, means by [theorem: 14.28], that $z\in\overline{M}$ so that $x=\alpha\cdot z\in\alpha\cdot\overline{M}$. So we have proved that
$$\overline{\alpha\cdot M}\subseteq\alpha\cdot\overline{M} \tag{14.42}$$
For the opposite inclusion, assume that $x\in\alpha\cdot\overline{M}$. If U is a open set such that $x\in U$ then by [theorem: 14.64] there exist a $\delta\in\mathbb{R}^+$ such that
$$x\in B_{\|\|}(x,\delta)\subseteq U \tag{14.43}$$
As $x\in\alpha\cdot\overline{M}$ there exist a $z\in\overline{M}$ such that $x=\alpha\cdot z$, further we have trivially $z\in B_{\|\|}(z,\delta/|\alpha|)$ which is a open set so that by [theorem: 14.28] we have that $M\bigcap B_{\|\|}(z,\delta/|\alpha|\}\neq\varnothing$, hence there exist a $m\in M$ such that $m\in B_{\|\|}(z,\delta/|\alpha|\}\Rightarrow\|z-m\|<\delta/|\alpha|$. So $\|x-\alpha\cdot m\|=\|\alpha\cdot z-\alpha\cdot m\|=|\alpha|\cdot\|z-m\|<|\alpha|\cdot(\delta/|\alpha|)=\delta$ proving that $\alpha\cdot m\in B_{\|\|}(x,\delta)\subseteq U$ hence, as $\alpha\cdot m\in\alpha\cdot M$ we have that $U\bigcap(a\cdot M)\neq\varnothing$, proving by [theorem: 14.28] that $x\in\overline{\alpha\cdot M}$. So $\alpha\cdot\overline{M}\subseteq\overline{\alpha\cdot M}$ which combined with [eq: 14.42] gives
$$\overline{\alpha\cdot M}=\alpha\cdot\overline{M}$$

2. If A is closed then by [theorem: 14.25] $A=\overline{A}$, hence $\alpha\cdot A=\alpha\cdot\overline{A}\underset{(1)}{=}\overline{\alpha\cdot A}$ proving by [theorem: 14.25] that $\alpha\cdot A$ is closed.

3. If $x\in\alpha\cdot M$ then $x=\alpha\cdot m$ where $m\in M$. As M is open there exists by [theorem: 14.64] a $\delta\in\mathbb{R}^+$ such that $m\in B_{\|\|}(m,\delta)\subseteq M$. If $z\in B_{\|\|}(x,|\alpha|\cdot\delta)$ then $\|z-x\|<|\alpha|\cdot\delta$ so that
$$\|m-(1/\alpha)\cdot z\|=\|(1/\alpha)\cdot x-(1/\alpha)\cdot z\|=(1/|\alpha|)\cdot\|x-z\|<(1/|\alpha|)\cdot|\alpha|\cdot\delta=\delta$$
proving that $(1/\alpha)\cdot z\in B_{\|\|}(m,\delta)\subseteq M\Rightarrow z\in\alpha\cdot M$. Hence $x\in B_{\|\|}(x,|\alpha|\cdot\delta)\subseteq\alpha\cdot M$ proving by [theorem: 14.64] that $\alpha\cdot M$ is open.

4. Let $z\in x+M$ then $z-x\in M$ a open set, hence there exist by [theorem: 14.64] a $\delta\in\mathbb{R}^+$ such that
$$z-x\in B_{\|\|}(z-x,\delta)\subseteq M$$
Take $y\in B_{\|\|}(z,\delta)$ then $\|z-y\|<\delta\Rightarrow\|(y-x)-(z-x)\|=\|y-z\|=\|z-y\|<\delta$ proving that $y-x\in\mathcal{B}_{\|\|}(z-x,\delta)\subseteq M\Rightarrow y\in x+M$. Hence $z\in B_{\|\|}(z,\delta)\subseteq x+M$ proving by [theorem: 14.32] that $x+M$ is open.

5. This follows from (4) as $A+M=\bigcup_{x\in A}(x+M)$ that is a union of open sets, which is open by definition.

6. Let $z\in x+\overline{M}$ then $z=x+y$ where $y\in\overline{M}$. If U is a open set such that $z\in U$ then by [theorem: 14.64] there exist a $\delta\in\mathbb{R}^+$ such that
$$z\in B_{\|\|}(z,\delta)\subseteq U$$
As $y\in B_{\|\|}(y,\delta)\in\mathcal{T}_{\|\|}$ and $y\in\overline{M}$ it follows from [theorem: 14.28] that $B_{\|\|}(y,\delta)\bigcap M\neq\varnothing$, hence there exist a $y_1\in B_{\|\|}(y,\delta)\bigcap M$. Take $z_1=x+y_1$ then $z_1\in x+M$ and $\|z-z_1\|=\|x+y-x-y_1\|=\|y-y_1\|<\delta$ so that $z_1\in B_{\|\|}(z,\delta)\subseteq U$. Hence $z_1\in U\bigcap(x+M)$ so that $U\bigcap(x+M)\neq\varnothing$ proving by [theorem: 14.28] that $z\in\overline{x+M}$. Hence
$$x+\overline{M}\subseteq\overline{x+M}$$

7. If $x\in\overline{M}-\overline{M}$ then $x=y_1-y_2$ where $y_1,y_2\in\overline{M}$. Let U be a open set such that $x\in U$ then by [theorem: 14.64] there exist a $\delta\in\mathbb{R}^+$ such that
$$x\in B_{\|\|}(x,\delta)\subseteq U \tag{14.44}$$
As $y_1,y_2\in\overline{M}$ we have, as $y_1\in B_{\|\|}(y_1,\delta/2)\in\mathcal{T}_{\|\|}$, $y_2\in B_{\|\|}(y_2,\delta/2)\in\mathcal{T}_{\|\|}$, by [theorem: 14.28] that there exists a $z_1\in B_{\|\|}(y_1,\delta/2)\bigcap M$ and a $z_2\in B_{\|\|}(y_2,\delta/2)\bigcap M$. Take $z=z_1-z_2\in M-M$ then
$$\|x-z\|=\|(y_1-y_2)-(z_1-z_2)\|=\|(y_1-z_1)+(z_2-y_2)\|\leqslant\|y_1-z_1\|+\|y_2-z_2\|<\delta$$

proving that $z\in B_{\|\|}(x,\delta)$. As $z\in M-M$ it follows that $\varnothing\neq B_{\|\|}(x,\delta)\bigcap(M-M\subseteq U\bigcap(M-M$. So applying [theorem: 14.28] proves then that $\overline{M-M}\subseteq\overline{M}-\overline{M}$. □

Theorem 14.110. *Let $\langle X,d\rangle$ a pseudo normed space and $A\subseteq X$ then we have*

$$A \text{ is dense in } X \text{ [using the norm topology } \mathcal{T}_{\|\|}]$$

$$\Updownarrow$$

$$\forall x\in X \text{ we have } \forall\varepsilon\in\mathbb{R}^+ \text{ there exist a } a\in A \text{ such that } \|x-a\|<\varepsilon$$

Proof.

⇒. Let $x\in X$ then $\forall\varepsilon\in\mathbb{R}^+$ such that $x\in B_{d_{\|\|}}(x,\varepsilon)$ we have by [theorem: 14.51] that $\exists a\in A$ such that $a\in B_{d_{\|\|}}(x,\varepsilon)$, hence $\|x-a\|<\varepsilon$.

⇐. Let $x\in X$ and let $B\in\mathcal{B}_{d_{\|\|}}$ such that $x\in B$. As $B\in\mathcal{B}_{d_{\|\|}}$ there exists a $y\in X$ and a $\varepsilon\in\mathbb{R}^+$ such that $x\in B=B_{d_{\|\|}}(y,\varepsilon)$. By [theorem: 14.62] there exist a $\delta\in\mathbb{R}^+$ such that $x\in B_{d_{\|\|}}(x,\delta)\subseteq B_{d_{\|\|}}(y,\varepsilon)=B$. By the hypothesis there exist a $a\in A$ such that $\|x-a\|<\delta$ or $a\in B_{d_{\|\|}}(x,\delta)\subseteq B_{d_{\|\|}}(y,\varepsilon)=B$. So

$$\forall x\in X \text{ we have } \forall B\in\mathcal{B}_{d_{\|\|}} \text{ with } x\in B \text{ that there exist a } a\in A \text{ such that } a\in B$$

Applying [theorem: 14.51] on the above proves that A is dense in X. □

Theorem 14.111. *Let $\|\|_1$, $\|\|_2$ two norms on a vector space X and $\mathcal{T}_{\|\|_1}$, $\mathcal{T}_{\|\|_2}$ the associated topologies then $\mathcal{T}_{\|\|_2}$ is finer then $\mathcal{T}_{\|\|_1}$*

$$\mathcal{T}_{\|\|_1} \text{ is finer then } \mathcal{T}_{\|\|_2}\Leftrightarrow\exists M\in\mathbb{R}^+ \text{ such that } \forall x\in X \text{ we have } \|x\|_1\leqslant M\cdot\|x\|_2$$

Proof.

⇒. As $\mathcal{T}_{\|\|_1}$ is finer then $\mathcal{T}_{\|\|_2}$ we have by [theorem: 14.68] there exist a $\delta\in\mathbb{R}^+$ such that

$$0\in B_{\|\|_2}(0,\delta)\subseteq B_{\|\|_1}(0,1) \tag{14.45}$$

Take $M=2/\delta$. Let $x\in X$ then we have either:

$x=0$. Then $\|x\|_1=0\leqslant 0=M\cdot 0=M\cdot\|x\|_2$

$x\neq 0$. Take $y=(1/(M\cdot\|x\|_2))\cdot x$ [as $\|\|_2$ is a norm $\|x\|_2\neq 0$] then we have

$$\|y\|_2=|1/(M\cdot\|x\|_2)|\cdot\|x\|_2=1/M=\delta/2<\delta$$

so that by [eq: 14.45] $y\in B_{\|\|_1}(0,1)$, hence $\|y\|_1<1$ or as

$$1>\|y\|_1=\|(1/(M\cdot\|x\|_2))\cdot x\|_1=|1/(M\cdot\|x\|_2)|\cdot\|x\|_1=\|x\|_1/(M\cdot\|x\|_2)$$

so that $\|x\|_1/(M\cdot\|x\|_2)<1$ proving that

$$\|x\|_1\leqslant M\cdot\|x\|_2$$

hence

$$\forall x\in X \text{ we have } \|x\|_1\leqslant M\cdot\|x\|_2$$

⇐. Let $\varepsilon\in\mathbb{R}^+$ and $x\in B_{\|\|_1}(x,\varepsilon)$. As $M\in\mathbb{R}^+$ we can tan take $\delta=\varepsilon/M\in\mathbb{R}^+$ then if $y\in B_{\|\|_2}(x,\delta)$ we have $\|x-y\|_2<\delta$ and by the hypothesis

$$\|x-y\|_1\leqslant M\cdot\|x-y\|_2<M\cdot\delta=M\cdot(\varepsilon/M)=\varepsilon$$

so that $y\in B_{\|\|_1}(x,\varepsilon_1)$. Hence $B_{\|\|_2}(x,\delta)\subseteq B_{\|\|_1}(x,\varepsilon)$ which by [theorem: 14.68] proves that

$$\mathcal{T}_{\|\|_2} \text{ is finer then } \mathcal{T}_{\|\|_1}$$ □

Definition 14.112. *Let X be a real or complex vector space then two pseudo norms are equivalent if*

$$\mathcal{T}_{\|\|_1}=\mathcal{T}_{\|\|_2}$$

Theorem 14.113. *Let X be a real or complex vector space and*

$$\mathcal{N}(X)=\{\|\|\in\mathbb{R}^{X\times X}\,|\,\|\| \text{ is a pseudo norm}\}$$

then if we define $\approx\subseteq\mathcal{N}(X)\times\mathcal{N}(X)$ *by* $\|\|_1\approx\|\|_2$ *if* $\|\|_1$ *is equivalent with* $\|\|_3$ *we have that*

$$\approx \text{ is a equivalence relation}$$

Proof. The proof is trivial because set equality is a equivalence relation

reflexivity. If $\|\|$ is a pseudo norm on X then clearly $\mathcal{T}_{\|\|}=\mathcal{T}_{\|\|}$ so that $\|\|\approx\|\|$

symmetry. If $\|\|_1\approx\|\|_2$ then $\mathcal{T}_{\|\|_1}=\mathcal{T}_{\|\|_2}\Rightarrow\mathcal{T}_{\|\|_2}=\mathcal{T}_{\|\|_1}$ so that $\|\|_2\approx\|\|_1$

transitivity. If $\|\|_1\approx\|\|_1$ and $\|\|_2\approx\|\|_3$ then $\mathcal{T}_{\|\|_1}=\mathcal{T}_{\|\|_2}\wedge\mathcal{T}_{\|\|_2}\approx\mathcal{T}_{\|\|_3}$, hence $\mathcal{T}_{\|\|_1}=\mathcal{T}_{\|\|_3}$ proving that $\|\|_1\approx\|\|_3$ □

Theorem 14.114. *Let* X *be a real or complex vector space and* $\|\|_1$, $\|\|_2$ *pseudo norms on* X *then*

$$\|\|_1 \text{ is equivalent with } \|\|_2 \Leftrightarrow \exists M_1,M_2\in\mathbb{R}^+ \text{ such that } \forall x\in X\ M_1\cdot\|x\|_1\leqslant\|x\|_2\leqslant M_2\cdot\|x\|_1$$

Proof. Let $\mathcal{T}_{\|\|_1}$, $\mathcal{T}_{\|\|_2}$ be the topologies generated by $\|\|_1$, $\|\|_2$ then we have

⇒. Then as $\mathcal{T}_{\|\|_1}=\mathcal{T}_{\|\|_2}$ we have that $\mathcal{T}_{\|\|_2}$ is finer then $\mathcal{T}_{\|\|_1}$ and $\mathcal{T}_{\|\|_1}$ is finer then $\mathcal{T}_{\|\|_2}$ so by [theorem: 14.111] there exists $M_1',M_2\in\mathbb{R}^+$ such that $\forall x\in X$ $\|x\|_1\leqslant M_1'\cdot\|x\|_2$ and $\|x\|_2\leqslant M_2\cdot\|x\|_1$ hence if we take $M_1=1/M_1'$ then we have

$$M_1\cdot\|x\|_1\leqslant\|x\|_2\leqslant M_2\cdot\|x\|_1$$

⇐. Let $x\in X$ then as $M_1\cdot\|x\|_1\leqslant\|x\|_2\leqslant M_2\cdot\|x\|_1$ we have $\|x\|_1\leqslant(1\setminus M_1)\cdot\|x\|_2$ and $\|x\|_2\leqslant M_2\cdot\|x\|_1$ so that by [theorem: 14.111] we have that $\mathcal{T}_{\|\|_1}$ is finer then $\mathcal{T}_{\|\|_2}$ and $\mathcal{T}_{\|\|_2}$ is finer then $\mathcal{T}_{\|\|_1}$. So $\mathcal{T}_{\|\|_1}=\mathcal{T}_{\|\|_2}$ or $\|\|_1$ is equivalent with $\|\|_2$. □

Example 14.115. In the real vector space $\mathbb{R}$ all norms are equivalent.

Proof. Note that by [example: 14.96]

$$|| \text{ is a norm on } \mathbb{R}$$

Let $\|\|$ be a norm on $\mathbb{R}$ then given $x\in X$ we have $\|x\|=\|1\cdot x\|=|x|\cdot\|1\|$, so if we take $N=M=\|1\|\in\mathbb{R}^+$ [because $1\neq 0$ and $\|\|$ is a norm] we have

$$N\cdot|x|\leqslant N\cdot|x|=\|1\|\cdot|x|=\|x\|=\|1\|\cdot|x|=M\cdot|x|\leqslant M\cdot|x|$$

so that $\|\|$ is equivalent with $||$. Hence every norm on $\mathbb{R}$ is equivalent with $||$. As norm equivalence is a equivalence relation [see theorem: 14.113] it follows that all norms on $\mathbb{R}$ are equivalent. □

Example 14.116. In the complex vector space $\mathbb{C}$ all norms are equivalent.

Proof. Note that by [example: 14.96]

$$|| \text{ is a norm on } \mathbb{C}$$

Let $\|\|$ be a norm on $\mathbb{C}$ then given $x\in X$ we have $\|x\|=\|1\cdot x\|=|x|\cdot\|1\|$, so if we take $N=M=\|1\|\in\mathbb{R}^+$ [because $1\neq 0$ and $\|\|$ is a norm] we have

$$N\cdot|x|\leqslant N\cdot|x|=\|1\|\cdot|x|=\|x\|=\|1\|\cdot|x|=M\cdot|x|\leqslant M\cdot|x|$$

so that $\|\|$ is equivalent with $||$. Hence every norm on $\mathbb{C}$ is equivalent with $||$. As norm equivalence is a equivalence relation [see theorem: 14.113] it follows that all norms on $\mathbb{C}$ are equivalent. □

Definition 14.117. *Let* $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$ *be pseudo normed spaces then* $\varphi: X\rightarrow Y$ *is a* ***linear isometry*** *if*

1. $\varphi: X\rightarrow Y$ *is a linear mapping*
2. $\forall x\in X$ *we have* $\|x\|_X=\|\varphi(x)\|_Y$

If in addition $\varphi: X\rightarrow Y$ *is a* ***bijection*** *then*

$$\varphi: X\rightarrow Y \textbf{ is a linear isometric isomorphism}$$

If between $\langle X,\|\|_X\rangle$ *and* $\langle Y,\|\|_Y\rangle$ *there exist a linear isometric isomorphism then we says that* $\langle X,\|\|_X\rangle$ *and* $\langle Y,\|\|_Y\rangle$ ***are isometric***. *We note this as*

$$X\approx Y$$

Theorem 14.118. *Let $\langle X,\|\|_X\rangle$ a normed space, $\langle Y,\|\|_Y\rangle$ a pseudo normed spaces and $\varphi: X \rightarrow Y$ a **linear isometry** then φ is **injective**. Hence if $\varphi: X \rightarrow Y$ is also surjective then $\varphi: X \rightarrow Y$ is a bijection and thus a **linear isometric isomorphism**.*

Proof. Let $x, y \in X$ such that $\varphi(x) = \varphi(y)$ then

$$\|x-y\|_X = \|\varphi(x-y)\|_Y = \|\varphi(x)-\varphi(y)\|_Y = \|0\|_Y = 0$$

which as $\|\|_X$ is a norm proves that $x - y = 0$ hence $x = y$. □

Theorem 14.119. *Let $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$ be pseudo normed spaces and $\varphi: X \rightarrow Y$ a linear isometric isomorphism then $\varphi^{-1}: Y \rightarrow X$ is a linear isometric isomorphism.*

Proof. Because of [theorem: 11.169] we have that $\varphi^{-1}: Y \rightarrow X$ is a linear isomorphism. Further if $y \in Y$ then $\|y\|_Y = \|\varphi(\varphi^{-1}(y))\|_Y = \|\varphi^{-1}(y)\|_X$. Hence $\varphi^{-1}: Y \rightarrow X$ is a linear isometry. □

Theorem 14.120. *Let $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$ and $\langle Z,\|\|_Z\rangle$ be pseudo metric spaces and $\varphi: X \rightarrow Y$, $\psi: Y \rightarrow Z$ linear isometries then*

$$\psi \circ \varphi: X \rightarrow Z \textit{ is a } \text{linear isometry}$$

If in addition φ, ψ are linear isometric isomorphisms then

$$\psi \circ \varphi: X \rightarrow Z \textit{ is a linear isometric isomorphism}$$

Proof. Using [theorem: 11.170] $\psi \circ \varphi: X \rightarrow Z$ is a linear mapping. Further given $x \in X$ we have

$$\|(\psi \circ \varphi)(x)\|_Z = \|\psi(\varphi(x))\|_Z = \|\varphi(x)\|_Y = \|x\|_X$$

hence

$$\psi \circ \varphi: X \rightarrow Z \text{ is a linear isometry}$$

Further if φ, ψ are linear isometric isomorphisms,then they are bijection, hence by [theorem: 2.75] $\psi \circ \varphi$ is a bijection, so that $\psi \circ \varphi$ is a linear isometric isomorphism. □

We have the following relation between a linear isometry in a normed space and a linear isometry in its associated metric space.

Theorem 14.121. *Let $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$ be normed spaces with associated metric spaces and $\varphi: X \rightarrow Y$ a linear isometry then*

$$\forall x, y \in X \textit{ we have } d_{\|\|_X}(x,y) = d_{\|\|_Y}(\varphi(x),\varphi(y)) \Leftrightarrow \forall x \in X \textit{ we have } \|x\|_X = \|\varphi(x)\|_Y$$

In other words if φ is a linear mapping then

$$\varphi \textit{ is a isometry } \Leftrightarrow \varphi \textit{ is a linear isometry}$$

and

$$\varphi \textit{ is a isometric isomorphism} \Leftrightarrow \varphi \textit{ is a linear isometric isomorphism}$$

Proof.

⇒. Let $x \in X$ then we have

$$\|x\|_X = \|x-0\|_X = d_{\|\|_X}(x,0) = d_{\|\|_Y}(\varphi(x),\varphi(0)) = \|\varphi(x)-\varphi(y)\|_Y = \|\varphi(x-0)\|_Y = \|\varphi(x)\|_Y$$

⇐. Let $x, y \in X$ then we have

$$d_{\|\|_X}(x,y) = \|x-y\|_X = \|\varphi(x-y)\|_Y = \|\varphi(x)-\varphi(y)\|_Y = d_{\|\|_Y}(\varphi(x),\varphi(y))$$ □

Theorem 14.122. *Let $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$ be normed spaces and $\varphi: X \rightarrow Y$ a linear isometric isomorphism then*

$$\mathcal{T}_{\|\|_X} = \{\varphi^{-1}(V) | V \in \mathcal{T}_{\|\|_Y}\} = \{U \subseteq X | \varphi(U) \in \mathcal{T}_{\|\|_Y}\}$$

and

$$\mathcal{T}_{\|\|_Y} = \{V \subseteq Y | \varphi^{-1}(V) \in \mathcal{T} \in \mathcal{T}_{\|\|_X}\} = \{\varphi(U) | U \in \mathcal{T}_{\|\|_X}\}$$

Proof. This follows from [theorem: 14.74] and [theorem: 14.121] □

TODO We skip the inner product for later

14.4 Continuous functions

14.4.1 Continuous and open functions

Definition 14.123. *Let $\langle X, \mathcal{T}_X\rangle$, $\langle Y, \mathcal{T}_Y\rangle$ be topological spaces, $x \in X$ then a function*

$$f: X \to Y \text{ is } \textbf{continuous at } x \text{ if } \forall V \in \mathcal{T}_Y \text{ with } f(x) \in V \text{ there exist a } U \in \mathcal{T}_X \text{ such that } x \in U \wedge f(U) \subseteq V$$

Note 14.124. Be aware that the notation of continuity does depends on the function but also on the chosen topologies.

Theorem 14.125. *Let $\langle X, \mathcal{T}_X\rangle$, $\langle Y, \mathcal{T}_Y\rangle$ be topological spaces, $\mathcal{B}_X$ a topological basis for $\mathcal{T}_X$, $\mathcal{B}_Y$ a topological basis for $\mathcal{T}_Y$, $x \in X$ and $f: X \to Y$ a function then we have the following equivalences:*

1. *f is continuous at x*
2. *$\forall V \in \mathcal{B}_Y$ with $f(x) \in V$ there exist a $U \in \mathcal{T}_X$ such that $x \in U$ and $f(U) \subseteq V$*
3. *$\forall V \in \mathcal{B}_Y$ with $f(x) \in V$ there exist a $U \in \mathcal{B}_X$ such that $x \in U$ and $f(U) \subseteq V$*

Proof.

1 ⇒ 2. This is trivial as $\mathcal{B}_Y \in \mathcal{T}_Y$.

2 ⇒ 3. Let $V \in \mathcal{B}_Y$ with $f(x) \in V$ then by (2) there exist a $U' \in \mathcal{T}_X$ such that $x \in U' \wedge f(U') \subseteq V$, by [theorem: 14.32] there exist a $U \in \mathcal{B}_X$ with $x \in U \subseteq U'$ so that $f(U) \subseteq f(U') \subseteq V$.

3 ⇒ 1. Let $V \in \mathcal{T}_Y$ with $f(x) \in V$ then by [theorem: 14.32] there exist a $W \in \mathcal{B}_Y$ such that $f(x) \in W \subseteq V$. Hence by (3) there exist a $U \in \mathcal{B}_X$ such that $x \in U$ and $f(U) \subseteq W \subseteq V$ proving that f is continuous at x. □

If we work with metric or normed spaces then we can use the classical ε-δ definition of functions continuous at a point.

Theorem 14.126. *Let $\langle X, d_X\rangle$, $\langle Y, d_Y\rangle$ be metric spaces, $x \in X$ and $f: X \to Y$ then we have*

$$f: X \to Y \text{ is continuous at } x \text{ using the topologies } \mathcal{T}_{d_X}, \mathcal{T}_{d_Y}$$

$$\Updownarrow$$

$$\forall \varepsilon \in \mathbb{R}^+ \ \exists \delta \in \mathbb{R}^+ \text{ such that } \forall z \in X \text{ satisfying } d_X(x, z) < \delta \text{ we have } d_Y(f(x), f(z)) < \varepsilon$$

Proof.

⇒. Let $\varepsilon \in \mathbb{R}^+$ then $f(x) \in B_{d_Y}(f(x), \varepsilon) \in \mathcal{B}_{d_Y}$, hence by [theorem: 14.125] there exists a $B \in \mathcal{B}_{d_X}$ such that $f(x) \in B$ and $f(B) \subseteq B_{d_Y}(f(x), \varepsilon)$, hence there exist a $y \in X$ and a $\delta' \in \mathbb{R}^+$ such that $x \in B_{d_X}(y, \delta') \in \mathcal{B}_{d_X}$ and $f(B_{d_X}(y, \delta')) \subseteq B_{d_Y}(f(x), \varepsilon)$. Using [theorem: 14.62] there exist a $\delta \in \mathbb{R}^+$ such that $x \in B_{d_X}(x, \delta) \subseteq B_{d_X}(y, \delta')$. So we have

$$f(B_{d_X}(x, \delta)) \subseteq f(B_{d_X}(y, \delta')) \subseteq B_{d_Y}(f(x), \varepsilon)$$

So if $z \in X$ satisfies $d_X(x, y) < \delta$ then $z \in B_{d_X}(x, \delta)$ so that $f(z) \in B_{d_Y}(f(x), \varepsilon)$ or $d_Y(f(x), f(z)) < \varepsilon$.

⇐. Let $V \in \mathcal{T}_Y$ such that $f(x) \in V$ then by [theorem: 14.64] there exist a $\varepsilon \in \mathbb{R}^+$ such that $f(x) \in B_{d_Y}(f(x), \varepsilon) \subseteq V$. By the hypothesis there exist a $\delta \in \mathbb{R}^+$ such that

$$\forall z \in X \text{ with } d_X(x, z) < \delta \text{ we have } d_Y(f(x), f(z)) < \varepsilon$$

If $y \in f(B_{d_X}(x, \delta))$ then there exist a $z \in X$ such that $d_X(x, z) < \delta$ and $y = f(x)$ hence using the above we have that $d_Y(f(x), f(z)) < \varepsilon$ or $y = f(z) \in B_{d_Y}(f(x), \varepsilon) \subseteq V$. As $y \in f(B_{d_X}(x, \delta))$ was chosen arbitrary it follows that

$$f(B_d(x, y)) \subseteq V \text{ where } x \in B_{d_X}(x, \delta) \in \mathcal{T}_X$$

So we have that $\forall V \in \mathcal{T}_Y$ with $f(x) \in V$ we found a $B_{d_X}(x, \delta) \in \mathcal{T}_X$ such that $f(B_d(x, y)) \subseteq V$ proving that f is continuous at x. □

Theorem 14.127. *Let $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$ be pseudo normed spaces, $x \in X$ and $f: X \to Y$ then we have*

$$f: X \to Y \text{ is continuous at } x \text{ using the topologies } \mathcal{T}_{\|\|_X}, \mathcal{T}_{\|d\|_Y}$$

$$\Updownarrow$$

$$\forall \varepsilon \in \mathbb{R}^+ \ \exists \delta \in \mathbb{R}^+ \text{ such that } \forall y \in X \text{ satisfying } \|x - y\|_X < \delta \text{ we have } \|f(x) - f(y)\|_Y < \varepsilon$$

Proof.

⇒. Let $\varepsilon \in \mathbb{R}^+$ then by [theorem: 14.126] there exist a $\delta \in \mathbb{R}^+$ such that $\forall y \in X$ with $d_X(x, y) < \delta$ we have $d_Y(f(x), f(y)) < \varepsilon$. Hence if $\|x - y\|_X < \delta$ then $d_X(x, y) < \delta$ so that $\|f(x) - f(y)\|_Y = d_Y(f(x), f(y)) < \varepsilon$.

⇐. Let $\varepsilon \in \mathbb{R}^+$ then by the hypothesis there exist a $\delta \in \mathbb{R}^+$ such that for all $y \in X$ with $\|x - y\|_X < \delta$ we have $\|f(x) - f(y)\|_Y < \varepsilon$. So if $y \in X$ satisfies $d_{\|\|_X}(x, y) < \delta$ then $\|x - y\|_Y < \delta$ and we have $d_{\|\|_Y}(f(x), f(y)) = \|f(x) - f(y)\|_Y < \varepsilon$. Hence by [theorem: 14.126] it follows that f is continuous at x. □

The above works also for the subspace topology.

Corollary 14.128. *Let $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$ be pseudo normed spaces, $U \subseteq X$, $x \in U$ and $f: U \to Y$ then using the subspace topology $(\mathcal{T}_{\|\|_X})_{|U}$ on X and $\mathcal{T}_{\|\|_Y}$ on Y we have:*

$$f: U \to Y \text{ is continuous at } x$$

$$\Updownarrow$$

$$\forall \varepsilon \in \mathbb{R}^+ \ \exists \delta \in \mathbb{R}^+ \text{ such that } \forall y \in U \text{ satisfying } \|x - y\|_X < \delta \text{ we have } \|f(x) - f(y)\|_Y < \varepsilon$$

Proof. Note that using [theorem: 14.108] the norm for the subspace topology $(\|\|_X)_{|U \times U}$ is defined by $\forall x \in U$ $(\|\|_X)_{|U \times U}(x) = \|x\|_X$.

⇒. Let $\varepsilon \in \mathbb{R}^+$ then by [theorem: 14.127] there exist a $\delta \in \mathbb{R}^+$ such that $\forall y \in U$ statisfying $(\|x - y\|_X)_{|U \times U} < \delta$ we have $\|(x) - f(y)\|_Y < \varepsilon$. If $y \in U$ satisfies $\|x - y\|_X < \delta$ then $(\|x - y\|_X)_{|U \times U} < \delta$ so that $\|f(x) - f(y)\|_Y < \varepsilon$.

⇐. Let $\varepsilon \in \mathbb{R}^+$ then there exist a $\delta \in \mathbb{R}^+$ such that $\forall y \in U$ with $\|x - y\|_X < \delta$ we have $\|f(x) - f(y)\|_Y < \varepsilon$. Hence if $(\|x - y\|_X)_{|U \times U} < \delta$ we have $\|x - y\|_X < \delta$ so that $\|f(x) - f(y)\|_Y < \varepsilon$, so that, using [theorem: 14.127], $f: U \to Y$ is continuous. □

Definition 14.129. *Let $\langle X, \mathcal{T}_X\rangle$, $\langle Y, \mathcal{T}_Y\rangle$ be topological spaces then a function $f: X \to Y$ is* ***continuous*** *if $\forall x \in X$ f is continuous at x. The set of all the graphs of continuous functions from X to Y is noted as $\mathcal{C}(X, Y)$. So*

$$\mathcal{C}(X, Y) = \{f \in Y^X | f: X \to Y \text{ is continuous}\}$$

Theorem 14.130. *Let $\langle X, \mathcal{T}_X\rangle$, $\langle Y, \mathcal{T}_Y\rangle$ be topological spaces and $f: X \to Y$ a function then*

$$\forall x \in f \text{ is continuous at } x \Leftrightarrow \forall V \in \mathcal{T}_Y \text{ we have } f^{-1}(V) \in \mathcal{T}_X$$

or using the above definition:

$$f \text{ is continuous} \Leftrightarrow \forall V \in \mathcal{T}_Y \text{ we have } f^{-1}(V) \in \mathcal{T}_X$$

Proof.

⇒. Let $V \in \mathcal{T}_Y$ then $\forall x \in f^{-1}(V) \subseteq X$ we have $f(x) \in V$ so that by continuity at x there exist a $U \in \mathcal{T}_X$ such that $x \in U \wedge f(U) \subseteq V$. Hence by a consequence of the axiom of choice [see: 3.104] there exist a $\{U_x\}_{x \in f^{-1}(V)} \subseteq \mathcal{T}_X$ such that $\forall x \in f^{-1}(V)$ $x \in U_x$ and $f(U_x) \subseteq V$. So $U_x \subseteq f^{-1}(f(U_x)) \subseteq f^{-1}(V)$ so that

$$\bigcup_{x \in f^{-1}(V)} U_x \subseteq f^{-1}(V)$$

Further if $x \in f^{-1}(V)$ then $x \in U_x \subseteq \bigcup_{x \in f^{-1}(V)} U_x$ so that $f^{-1}(V) \subseteq \bigcup_{x \in f^{-1}(V)} U_x$ or combining this with the above that

$$f^{-1}(V) = \bigcup_{x \in f^{-1}(V)} U_x \text{ where } \{U_x\}_{x \in f^{-1}(V)} \subseteq \mathcal{T}_X$$

so that by definition

$$f^{-1}(V) \in \mathcal{T}_X$$

⇐. Let $x \in X$ and $V \in \mathcal{T}_Y$ such that $f(x) \in V$ then $x \in f^{-1}(V)$. By the hypothesis we have $f^{-1}(V) \in \mathcal{T}_X$, so if we take $U = f^{-1}(V)$ then $U \in \mathcal{T}_X$ and $x \in U$. Further by [theorem: 2.25] we have $f(f^{-1}(V)) \subseteq V$ so that $f(U) = f(f^{-1}(V)) \subseteq V$ proving that f is continuous at x. □

Example 14.131. Let $\langle X, \mathcal{T}_X\rangle$, $\langle Y, \mathcal{T}_Y\rangle$ be topological spaces, $y \in Y$ then $C_y: X \to Y$ defined by $C_y(x) = y$ is continuous.

Proof. If $V \in \mathcal{T}_X$ then we have either:

$\boldsymbol{y \in V}$**.** Then $f^{-1}(V) = X \in \mathcal{T}_X$

$y \notin V$. Then $f^{-1}(V) = \varnothing \in \mathcal{T}_X$

So using [theorem: 14.130] it follows that C_y is continuous. □

Example 14.132. Let $\langle X, \mathcal{T}_X \rangle$ be a topological space then $\mathrm{Id}_X: X \to X$ is continuous.

Proof. Let $V \in \mathcal{T}_X$ then $(\mathrm{Id}_X)^{-1}(V) = V \in \mathcal{T}_X$ proving by [theorem: 14.130] continuity. □

Theorem 14.133. *Let $\langle X, \mathcal{T}_X \rangle$, $\langle Y, \mathcal{T}_Y \rangle$ be topological spaces, $A \subseteq X$, $x \in A$ and $f: X \to Y$ a function that is continuous at x then*

$$f_{|A}: A \to Y$$

is continuous at x using the subspace topology $(\mathcal{T}_X)_{|A} = \{A \bigcap U | U \in \mathcal{T}_X\}$ on A and the topology $\mathcal{T}_Y$ on Y.

Proof. Let $V \in \mathcal{T}_Y$ such that $f_{|A}(x) \in V$. Then $f(x) \in V$ and by continuity of f at X there exist a $U \in \mathcal{T}_X$ with $x \in U$ and $f(U) \subseteq V$. As $x \in A$ we have that $x \in A \bigcap U \in (\mathcal{T}_X)_{|A}$ and $f_{|A}(A \bigcap U) \underset{\text{[theorem: 2.84]}}{=} f(A \bigcap U) \subseteq f(U) \subseteq V$ proving that $f_{|A}$ is continuous at x. □

The opposite is true if A is open showing that continuity at a point is a local property.

Theorem 14.134. *Let $\langle X, \mathcal{T}_X \rangle$, $\langle Y, \mathcal{T}_Y \rangle$ be topological spaces, $U \in \mathcal{T}_X$, $x \in U$ and $f: X \to Y$ a function such that $f_{|U}: U \to Y$ is continuous at x then f is continuous at x.*

Proof. Let $V \in \mathcal{T}_Y$ such that $f(x) \in V$ then as $f_{|U}$ is continuous at x there exist a $W' \in (\mathcal{T}_X)_{|U}$ such that $x \in W'$ and $f_{|U}(W') \subseteq V \Rightarrow f(W') \subseteq V$. As $W' \in (\mathcal{T}_X)_{|U}$ there exist a $W \in \mathcal{T}_X$ such that $W' = W \bigcap U$, as $U \in \mathcal{T}_X$ it follows that $W' \in \mathcal{T}_X$. So we found a $W' \in \mathcal{T}_X$ such that $x \in W'$ and $f(W') \subseteq V$ proving that f is continuous at x. □

Theorem 14.135. *Let $\langle X, \mathcal{T}_X \rangle$, $\langle Y, \mathcal{T}_Y \rangle$ be topological spaces, $A \subseteq X$ and $f: X \to Y$ a continuous function then*

$$f_{|A}: A \to Y$$

is a continuous function using the subspace topology $(\mathcal{T}_X)_{|A} = \{A \bigcap U | U \in \mathcal{T}_X\}$ on A and the topology $\mathcal{T}_Y$ on Y.

Proof. Let $V \in \mathcal{T}_Y$ then we have $(f_{|A})^{-1}(V) \underset{\text{[theorem: 2.84]}}{=} A \bigcap f^{-1}(V)$. As f is continuous $f^{-1}(V) \in \mathcal{T}_X$ so that $A \bigcap f^{-1}(V) \in (\mathcal{T}_X)_{|A}$ proving that $(f_{|A})^{-1}(V) \in (f_{|A})^{-1}(V)$, hence by [theorem: 14.130] f is continuous. □

Theorem 14.136. *Let $\langle X, \mathcal{T}_X \rangle$, $\langle Y, \mathcal{T}_Y \rangle$ be topological spaces then the following are equivalent:*

1. *f is continuous*
2. *$\forall A \subseteq X$ we have $f(\overline{A}) \subseteq \overline{f(A)}$*
3. *$\forall F \subseteq Y$ F closed we have that $f^{-1}(F)$ is closed*

Proof.

1 ⇒ 2. Let $y \in f(\overline{A})$ then there exist a $x \in \overline{A}$ such that $y = f(x)$. Let $V \in \mathcal{T}_Y$ such that $y \in V$ then, as f is continuous, there exist a $U \in \mathcal{T}_X$ such that $x \in U$ and $f(U) \subseteq V$. Given that $x \in \overline{A}$ we have by [theorem: 14.29] that $A \bigcap U \neq 0$ so that $\varnothing \neq f(A \bigcap U)$. By [theorem: 2.93] $f(A \bigcap U) \subseteq f(A) \bigcap f(U)$ so that $\varnothing \neq f(A) \bigcap f(U) \subseteq f(A) \bigcap V$ proving that $y \in \overline{f(A)}$. As $y \in f(\overline{A})$ was chosen arbitrary it follows that

$$f(\overline{A}) \subseteq \overline{f(V)}$$

2 ⇒ 3. Let $A \subseteq Y$ be closed then by (2)

$$f(\overline{f^{-1}(A)}) \subseteq \overline{f(f^{-1}(A))} \tag{14.46}$$

Further by [theorem: 2.25] we have that $f(f^{-1}(A)) \subseteq A$ so that by [corollary: 14.30] we have $\overline{f(f^{-1}(A))} \subseteq \overline{A} \underset{\text{[theorem: 14.25]}}{=} A$ or combining this with [eq: 14.46] gives

$$f(\overline{f^{-1}(A)}) \subseteq A$$

Hence by [theorem: 2.92] we have that $f^{-1}(f(\overline{f^{-1}(A)})) \subseteq f^{-1}(A)$ and by [theorem: 2.25] $\overline{f^{-1}(A)} \subseteq f^{-1}(f(\overline{f^{-1}(A)}))$ so that $\overline{f^{-1}(A)} \subseteq f^{-1}(A)$. As by [theorem: 14.27] $f^{-1}(A) \subseteq \overline{f^{-1}(A)}$ it follows that $f^{-1}(A) = \overline{f^{-1}(A)}$. Hence by [theorem: 14.25] that

$$f^{-1}(A) \text{ is closed}$$

3 ⇒ 1. If $V \in \mathcal{T}_Y$ then by [theorem: 14.20] $Y \setminus V$ is closed so by (3) we have that $f^{-1}(Y \setminus V)$ is closed so that $X \setminus f^{-1}(Y \setminus V) \in \mathcal{T}_X$. Finally by [theorem: 2.92] we have that

$$f^{-1}(V) = f^{-1}(Y \setminus (Y \setminus V)) = X \setminus f^{-1}(Y \setminus V) \in \mathcal{T}_X$$

proving that for every $V \in \mathcal{T}_Y$ we have $f^{-1}(V) \in \mathcal{T}_X$. Hence by [theorem: 14.130] $f: X \rightarrow Y$ is continuous. □

If a function is continuous then the preimage of a open set is open, we can consider also functions so that the image of a open set is open. This is the idea of a open function.

Definition 14.137. (Open Function) *Let $\langle X, \mathcal{T}_X \rangle$, $\langle Y, \mathcal{T}_Y \rangle$ be topological spaces then a function*

$$f: X \rightarrow Y \textbf{\textit{ is open}}$$

if $\forall U \in \mathcal{T}_X$ we have that $f(U) \in \mathcal{T}_Y$

The composition of two continuous or open functions is again a continuous or open function.

Theorem 14.138. *Let $\langle X, \mathcal{T}_X \rangle$, $\langle Y, \mathcal{T}_Y \rangle$ and $\langle Z, \mathcal{T}_Z \rangle$ be topological spaces then we have:*

1. *If $f: X \rightarrow Y$ and $g: Y \rightarrow Z$ are continuous functions then $g \circ f: X \rightarrow Z$ is continuous.*
2. *If $f: X \rightarrow Y$ and $g: Y \rightarrow Z$ are open functions then $g \circ f: X \rightarrow Z$ is open*

Proof.

1. Let $V \in \mathcal{T}_Z$ then, as g is continuous we have by [theorem: 14.130] that $g^{-1}(V) \in \mathcal{T}_Y$, applying [theorem: 14.130] again results in $f^{-1}(g^{-1}(V)) \in \mathcal{T}_X$. As

$$(g \circ f)^{-1}(V) \underset{\text{[theorem: 2.23]}}{=} f^{-1}(g^{-1}(V)) \in \mathcal{T}_X$$

we have by [theorem: 14.130] that $g \circ f$ is continuous.

2. Let $U \in \mathcal{T}_Z$ then, as f is open, $f(U) \in \mathcal{T}_Y$, hence as g is open we have that

$$(g \circ f)(U) \underset{\text{[theorem: 2.23]}}{=} g(f(U)) \in \mathcal{T}_Z$$

proving that $g \circ f$ is open. □

One of the reasons to use the product topology on a product of topological spaces is that the projection map is open and continuous.

Theorem 14.139. *Let $\{\langle X_i, \mathcal{T}_i \rangle\}_{i \in I}$ be a family of topological spaces and $\langle \prod_{i \in I} X_i, \mathcal{T}_{\text{product}} \rangle$ be the product topological space [see definition: 14.43] then $\forall i \in I$ we have that the projection map*

$$\pi_i: \prod_{j \in I} X_j \rightarrow X_i \textit{ defined by } \pi_i(x) = x_i$$

is open and continuous.

Proof. Let $S = \{\pi_i^{-1}(U) | i \in I \wedge U \in \mathcal{T}_i\}$ then by [theorem: 14.40] we have

$$S \subseteq \mathcal{B}[\mathcal{S}] \subseteq \mathcal{T}[\mathcal{B}[S]] \underset{\text{[definition: 14.43]}}{=} \mathcal{T}_{\text{product}}$$

So if $i \in I$ then for $U \in \mathcal{T}_i$ we have $\pi_i^{-1}(U) \in \mathcal{S} \subseteq \mathcal{T}_{\text{product}}$ so that

$$\forall i \in I \text{ we have } \pi_i \text{ is continuous} \tag{14.47}$$

Let $U \in \mathcal{T}_{\text{product}}$ and let $t \in \pi_i(U)$ then there exist a $x \in U$ so that $t = \pi_i(x) = x_i$. By [theorem: 14.45] we have that

$$\mathcal{B}[S] = \left\{ \prod_{i \in I} U_i | \{U_i\}_{i \in I} \subseteq \mathcal{P}\left(\prod_{i \in I} X_i \right) \text{ is such that there exist a finite } J \subseteq I \text{ with } \forall i \in J\ U_i \in \mathcal{T}_i \text{ and } \forall i \in I \setminus J\ U_i = X_i \right\}$$

Hence there exist J finite and a $\{U_i\}_{i \in I}$ satisfying $U_i \in \mathcal{T}_i$ if $i \in J$ and $U_i = X_i \in \mathcal{T}_i$ if $i \in I \setminus J$ such that

$$x \in \prod_{i \in I} U_i \subseteq U$$

so $t = \pi_i(x) = x_i \in U_i \subseteq \pi_i(\prod_{i \in I} U_i) \subseteq \pi_i(U)$. To summarize we have

$$\forall t \in \pi_i(U) \text{ there exist a } U_i \in \mathcal{T}_i \text{ such that } t \in U_i \subseteq \pi_i(U)$$

which by [theorem: 14.7] proves that $\pi_i(U) \in \mathcal{T}_i$. Hence

$$\forall i \in I \text{ we have that } \pi_i \text{ is open}$$

□

Corollary 14.140. *Let $\langle X, \mathcal{T}_X\rangle$ be a topological space, $\{\langle X_i, \mathcal{T}_i\rangle\}_{i \in I}$ be a family of topological spaces, $\langle \prod_{i \in I} X_i, \mathcal{T}_{\text{product}}\rangle$ be the product topological space [see definition: 14.43] and*

$$f: X \rightarrow \prod_{i \in I} X_i \text{ a function}$$

then

$$f \text{ is continuous} \Leftrightarrow \forall i \in I \; f_i = \pi_i \circ f \text{ is continuous}$$

Proof.

⇒. Let $i \in I$ then as f is continuous and by [theorem: 14.139] π_i is continuous we have by [theorem: 14.138] that $\pi_i \circ f$ is continuous.

⇐. Using [definitions: 14.43] we have that the basis for $\mathcal{T}_{\text{product}}$ is

$$\mathcal{B}_{\text{product}} = \left\{ B \in \mathcal{P}(X) | \exists \{S_i\}_{i \in J} \subseteq \mathcal{S},\ J \text{ finite and } J \neq \varnothing \text{ such that } B = \bigcap_{i \in J} S_i \right\}$$

where

$$S = \{\pi_i^{-1}(U) | i \in I \wedge U \in \mathcal{T}_i\}$$

Let $x \in X$ and $V \in \mathcal{T}_{\text{product}}$ such that $f(x) \in V$ then by [theorem: 14.32] there exists a finite J, $\{S_i\}_{i \in J} \subseteq S$ such that

$$f(x) \in \bigcap_{i \in J} S_i \subseteq V \text{ and } \bigcap_{i \in J} S_i \in \mathcal{B}_{\text{product}} \subseteq \mathcal{T}_{\text{product}} \tag{14.48}$$

Let $i \in J$ then $\exists i_j \in I$ and a $U_j \in \mathcal{T}_{i_j}$ such that $S_i = \pi_{i_j}^{-1}(U_j)$ and as $f(x) \in \bigcap_{i \in J} S_i \subseteq V$ we have also $f(x) \in S_i$. Hence we have

$$x \in f^{-1}(S_i) = f^{-1}(\pi_{i_j}^{-1}(U_j)) = (\pi_{i_j} \circ f)^{-1}(U_j) = (f_{i_j})^{-1}(U_j)$$

As f_{i_j} is continuous and $U_j \in \mathcal{T}_j$ we have that $f_{i_j}^{-1}(U_j) \in \mathcal{T}_X$ so, as J is finite, it follows that, if we take $U = \bigcap_{i \in J} f^{-1}(S_i)$, that

$$U = \bigcap_{i \in J} f^{-1}(S_i) = \bigcap_{i \in J} (f_{i_j})^{-1}(U_j) \in \mathcal{T}_X$$

and as $x \in f^{-1}(S_i)\ \forall i \in J$ that

$$x \in \bigcap_{i \in J} f^{-1}(S_i) = U$$

Further

$$\begin{aligned} f(U) &= f\left(\bigcap_{i \in J} f^{-1}(S_i)\right) \\ &\subseteq_{[\text{theorem: } 2.134]} \bigcap_{i \in J} f(f^{-1}(S_i)) \\ &\subseteq_{[\text{theorem: } 2.25]} \bigcap_{i \in J} S_i \\ &\subseteq_{[\text{eq: } 14.48]} V \end{aligned}$$

Hence we have proved that for $\forall x \in X$ we have $\forall V \in \mathcal{T}_{\text{product}}$ with $x \in V$ there exist a $U \in \mathcal{T}_X$ such that $x \in U$ and $f(U) \subseteq V$ which proves that f is continuous at x. Hence we have that f is continuous. □

Theorem 14.141. *Let $\langle X_1, \mathcal{T}_1\rangle$, $\langle X_2, \mathcal{T}_2\rangle$, $\langle Y, \mathcal{T}_Y\rangle$ be topological spaces, $\left\langle \prod_{i \in \{1,2\}} X_i, \mathcal{T}_{\text{product}}\right\rangle$ the topological space based on the product of X_1, X_2 and*

$$f: \prod_{i \in \{1,2\}} X_i \rightarrow Y \text{ a continuous function}$$

then we have

1. *$\forall x \in X_1$ $f_1(x): X_2 \rightarrow Y$ defined by $f_1(x)(y) = f(x, y)$ is continuous.*
2. *$\forall x \in X_2$ $f_2(x): X_1 \rightarrow Y$ defined by $f_2(x)(y) = f(y, x)$ is continuous.*

Proof. First as $\{1,2\}$ is finite we have that $\mathcal{T}_{\text{product}} \underset{[\text{theorem: } 14.47]}{=} \mathcal{T}_{\text{box}}$ so that the basis for $\mathcal{T}_{\text{product}}$ is

$$\mathcal{B} \underset{[\text{definition: } 14.41]}{=} \left\{ \prod_{i \in I} U_i | \{U_i\}_{i \in I} \subseteq \mathcal{P}\left(\prod_{i \in I} X_i \right) \text{ is a family such that } \forall i \in I \; U_i \in \mathcal{T}_i \right\} \tag{14.49}$$

1. Let $x \in X_1$ and let $V \in \mathcal{T}_Y$ then by continuity of f we have that $f^{-1}(V) \in \mathcal{T}_{\text{product}}$. Take now $y \in f_1(x)^{-1}(V)$ then $f(x,y) = f_1(x)(y) \in V$ so that $(x,y) \in f^{-1}(V)$ and by [eq: 14.49] there exists a $U_1 \in \mathcal{T}_1$, $U_2 \in \mathcal{T}_2$ such that $(x,y) \in \prod_{i \in \{1,2\}} U_i \subseteq f^{-1}(V)$. Hence $x \in U_1$ and $y \in U_2$. If $t \in U_2$ then $(x,t) \in \prod_{i \in \{1,2\}} U_i \subseteq f^{-1}(V)$ hence $f_1(x)(t) = f(x,t) \in V$ or $t \in f_1(x)^{-1}(V)$ proving that $y \in U_2 \subseteq f_1(x)^{-1}(V)$ which by [theorem: 14.7] proves that $f_1(x)^{-1}(V)$ is open. Hence by [theorem: 14.130]

$$f_1(x) \text{ is continuous}$$

2. Let $x \in X_2$ and let $V \in \mathcal{T}_Y$ then by continuity of f we have that $f^{-1}(V) \in \mathcal{T}_{\text{product}}$. Take now $y \in f_2(x)^{-1}(V)$ then $f(y,x) = f_2(x)(y) \in V$ so that $(y,x) \in f^{-1}(V)$ and by [eq: 14.49] there exists a $U_1 \in \mathcal{T}_1$, $U_2 \in \mathcal{T}_2$ such that $(y,x) \in \prod_{i \in \{1,2\}} U_i \subseteq f^{-1}(V)$. Hence $x \in U_2$ and $y \in U_1$. If $t \in U_1$ then $(t,x) \in \prod_{i \in \{1,2\}} U_i \subseteq f^{-1}(V)$ hence $f_2(x)(t) = f(t,x) \in V$ or $t \in f_2(x)^{-1}(V)$ proving that $y \in U_1 \subseteq f_2(x)^{-1}(V)$ which by [theorem: 14.7] proves that $f_2(x)^{-1}(V)$ is open. Hence by [theorem: 14.130]

$$f_2(x) \text{ is continuous}$$

□

Theorem 14.142. *Let $\langle X, \|\| \rangle$ be a pseudo normed space then*

$$\cdot : \mathbb{K} \times X \rightarrow X \text{ defined by } \cdot(\alpha, x) = \alpha \cdot x$$

is continuous [using the product topology based on $\mathcal{T}_{||}, \mathcal{T}_{\|\|}$ and $\mathcal{T}_{\|\|}$]

Proof. Using [theorem: 14.94] the product topology on $\mathcal{T}_{\text{product}} = \mathcal{T}_{\|\|_{\max}}$ where for $(\alpha, x) \in \mathbb{K} \times X$ $\|(\alpha, x)\|_{\max} = \max(\{|a|, \|x\|\})$. Let $(\alpha, x) \in \mathbb{K} \times X$ and take $\varepsilon \in \mathbb{R}^+$. If $(\beta, y) \in \mathbb{K} \times X$ is such that $\|x - y\| \leqslant \delta$ and $|\alpha - \beta| < \delta$ then we have

$$\begin{aligned}
\|\cdot(\alpha, x) - \cdot(\beta, y)\| &= \|\alpha \cdot x - \beta \cdot y\| \\
&= \|\alpha \cdot x - \beta \cdot x + \beta \cdot x - \beta \cdot y\| \\
&= \|(\alpha - \beta) \cdot x + \beta \cdot (x - y)\| \\
&\leqslant |\alpha - \beta| \cdot \|x\| + |\beta| \cdot \|x - y\| \\
&= |\alpha - \beta| \cdot \|x\| + |\beta - \alpha + \alpha| \cdot \|x - y\| \\
&\leqslant |\alpha - \beta| \cdot \|x\| + (|\beta - \alpha| + |\alpha|) \cdot \|x - y\| \\
&= |\alpha - \beta| \cdot \|x\| + |\beta - \alpha| \cdot \|x - y\| + |\alpha| \cdot \|x - y\| \\
&< \delta \cdot \|x\| + \delta \cdot \delta + |\alpha| \cdot \delta \\
&= \delta \cdot (\|x\| + \delta + |\alpha|)
\end{aligned}$$

Take now $\delta = \min(1, \varepsilon / (\|x\| + |\alpha| + 1)) > 0$ then we have that

$$\begin{aligned}
\delta \cdot (\|x\| + \delta + |\alpha|) &<_{\delta \leqslant 1} \delta \cdot (\|x\| + 1 + |\alpha|) \\
&< (\varepsilon / (\|x\| + |\alpha| + 1)) \cdot (\|x\| + 1 + |\alpha|) \\
&= \varepsilon
\end{aligned}$$

So we have proved that for $\delta = \min(1, \varepsilon / (\|x\| + |\alpha| + 1))$ and $(\beta, y) \in \mathbb{K} \times X$ such that $\|x - y\| \leqslant \delta$ and $|\alpha - \beta| < \delta$ that

$$\|\cdot(\alpha, x) - \cdot(\beta, y)\| < \varepsilon$$

So if we take $(\beta, y) \in \mathbb{K} \times X$ such that $\|(\alpha, x) - (\beta, y)\|_{\max} < \delta$ then $|\alpha - \beta| < \delta$ and $\|x - y\| < \delta$ resulting in $\|\cdot(\alpha, x) - \cdot(\beta, y)\| < \varepsilon$. Applying then [theorem: 14.127] proves that $\cdot$ is continuous at $(\alpha, x) \in \mathbb{K} \times X$ which as (α, x) was chosen arbitrary proves continuity. □

Corollary 14.143. *Let $\langle X, \|\| \rangle$ be a pseudo normed space over the field $\mathbb{K}$ then we have:*

1. *$\forall \alpha \in \mathbb{K}$ the function $\mu_\alpha : X \rightarrow X$ defined by $\mu_\alpha(x) = \alpha \cdot x$ is continuous*
2. *$\forall x \in \mathbb{K}$ the function $\nu_x : \mathbb{K} \rightarrow X$ defined by $\nu_x(\alpha) = \alpha \cdot x$ is continuous*

Proof. This follows from [theorem: 14.141] and [theorem: 14.142]. □

Theorem 14.144. *Let $\langle X, \|\|_X \rangle$, $\langle Y, \|\|_Y \rangle$ be pseudo normed spaces then we have:*

1. *If $f: X \rightarrow Y$, $g: X \rightarrow Y$ are continuous functions at $x \in X$ then*
$$f + g: X \rightarrow Y \text{ where } (f + g)(x) = f(x) + g(x)$$
is continuous at x. So if f, g are continuous then $f + g$ is continuous.
2. *If $\alpha \in \mathbb{K}$ and $f: X \rightarrow Y$ is a continuous function at x then*
$$\alpha \cdot f: X \rightarrow Y \text{ where } (\alpha \cdot f)(x) = \alpha \cdot f(x)$$
is continuous at x. So if f are continuous then $\alpha \cdot f$ is continuous.

Proof. We use the ε-δ-definition of continuity [see theorem: 14.127] to proof this theorem. So take $x \in X$ and $\varepsilon \in \mathbb{R}^+$.

1. As f, g are continuous at x there exists $\delta_1, \delta_2 \in \mathbb{R}^+$ such that $\forall x' \in X$ such that $\|x - x'\|_X < \delta_1$ we have $\|f(x) - f(x')\|_Y < \frac{\varepsilon}{2}$ and $\forall x' \in X$ such that $\|x - x'\|_X < \delta_2$ we have $\|g(x) - g(x')\|_Y < \frac{\varepsilon}{2}$. So if $\|x - x'\|_X < \delta$ then
$$\begin{array}{rcl} \|(f+g)(x) - (f+g)(x')\|_Y & = & \|f(x) + g(x) - f(x') - g(x')\|_Y \\ & \leqslant & \|f(x) - f(x')\|_Y + \|g(x) - g(x')\|_Y \\ & < & \dfrac{\varepsilon}{2} + \dfrac{\varepsilon}{2} \\ & = & \varepsilon \end{array}$$
proving that $f + g$ is continuous at x.
2. As f is continuous at x there exist a $\delta \in \mathbb{R}^+$ such that if $\|x - x'\|_X < \delta$ then
$$\|f(x) - f(x')\|_Y \leqslant \varepsilon / (|\alpha| + 1)$$
then we have
$$\|(\alpha \cdot f)(x) - (\alpha \cdot f)(x')\|_Y = \|\alpha \cdot (f(x) - f(x'))\|_Y = |\alpha| \cdot \|f(x) - f(x')\|_Y < \alpha \cdot (\varepsilon / (|\alpha| + 1)) < \varepsilon$$
proving that $\alpha \cdot f$ is continuous at x. □

Using induction we can extend the above to a finite family of functions.

Theorem 14.145. *Let $n \in \mathbb{N}$, $\langle X, \|\|_X \rangle$, $\langle Y, \|\|_Y \rangle$ be pseudo normed spaces and $\{f_i\}_{i \in \{1, \ldots, n\}} \subseteq \mathcal{C}(X, Y)$ a family of continuous functions then*
$$\sum_{i=1}^n f_i: X \rightarrow Y \text{ defined by } \left(\sum_{i=1}^n f_i\right)(x) = \sum_{i=1}^n f_i(x)$$
is continuous.

Proof. Let
$$S = \left\{n \in \mathbb{N} | \text{If } \{f_i\}_{i \in \{1, \ldots, n\}} \subseteq \mathcal{C}(X, Y) \text{ then } \sum_{i=1}^n f_i \in \mathcal{C}(X, Y)\right\}$$
then we have:

$1 \in S$. If $\{f_i\}_{i \in \{1\}} \subseteq \mathcal{C}(X, Y)$ then $\forall x \in X$ we have $(\sum_{i=1}^1 f_i)(x) = \sum_{i=1}^1 f_i(x) = f_1(x)$, proving that $\sum_{i=1}^1 f_i = f_1 \in \mathcal{C}(X, Y)$. Hence
$$1 \in S$$

$n \in S \Rightarrow n + 1 \in S$. Let $\{f_i\}_{i \in \{1, \ldots, n+1\}} \subseteq \mathcal{C}(X, Y)$ then we have $\forall x \in X$ that
$$\begin{array}{rcl} \left(\displaystyle\sum_{i=1}^{n+1} f_i\right)(x) & = & \displaystyle\sum_{i=1}^{n+1} f_i(x) \\ & = & \left(\displaystyle\sum_{i=1}^n f_i(x)\right) + f_{n+1}(x) \\ & = & \left(\displaystyle\sum_{i=1}^n f_i\right)(x) + f_{n+1}(x) \\ & = & \left(\left(\displaystyle\sum_{i=1}^n f_i\right) + f_{n+1}\right)(x) \end{array}$$

proving that $\sum_{i=1}^{n+1} f_i = (\sum_{i=1}^{n} f_i) + f_{n+1}$. As $n \in S$ $\sum_{i=1}^{n} f_i$ is continuous, so by the continuity of f_{n+1} we have by [theorem: 14.144] that $\sum_{i=1}^{n+1} f_i$ is continuous. Hence $n+1 \in S$. $\square$

Theorem 14.146. *Let $\langle X, \|\|\rangle$ be a pseudo normed space and $f: X \to \mathbb{K}$, $g: X \to \mathbb{K}$ functions continuous at x then*

$$f \cdot g: X \to \mathbb{K} \text{ defined by } (f \cdot g)(x) = f(x) \cdot g(x)$$

is continuous at x. Hence if f, g are continuous functions then $f \cdot g$ is a continuous function.

Proof. We use the ε-δ-definition of continuity [see theorem: 14.127] to proof this theorem. So take $x \in X$ and $\varepsilon \in \mathbb{R}^+$. Then as f is continuous at x there exist a $\delta_1 \in \mathbb{R}^+$ such that $\forall x' \in X$ with $\|x - x'\| < \varepsilon$ we have $|f(x) - f(x')| < 1$. Then

$$|f(x')| \leqslant |f(x') - f(x) + f(x)| \leqslant |f(x)| + |f(x) - f(x')| < |f(x)| + 1 \tag{14.50}$$

Using continuity again there exist a $\delta_2 \in \mathbb{R}^+$ such that $\forall x \in X$ with $\|x - x'\| < \delta_2$ we have

$$|f(x) - f(x')| < \frac{\varepsilon}{2 \cdot (1 + |g(x)|)} \tag{14.51}$$

As g is also continuous at x there exists a $\delta_2 \in \mathbb{R}^+$ such that $\forall x' \in X$ with $\|x - x'\| < \delta_3$

$$|g(x) - g(x')| < \frac{\varepsilon}{2 \cdot (1 + |f(x)|)} \tag{14.52}$$

Take $\delta = \min(\delta_1, \delta_2, \delta_3) \in \mathbb{R}^+$ then

$$\begin{array}{rcl}
|(f \cdot g)(x) - f(x') \cdot g(x')| & = & |f(x) \cdot g(x) - f(x') \cdot g(x')| \\
& = & |f(x) \cdot g(x) - f(x') \cdot g(x) + f(x') \cdot g(x) - f(x') \cdot g(x')| \\
& \leqslant & |f(x) \cdot g(x) - f(x') \cdot g(x)| + |f(x') \cdot g(x) - f(x') \cdot g(x')| \\
& = & |(f(x) - f(x')) \cdot g(x)| + |f(x') \cdot (g(x) - g(x'))| \\
& = & |g(x)| \cdot |f(x) - f(x')| + |f(x')| \cdot |g(x) - g(x')| \\
& <_{[\text{eq: } 14.50]} & |g(x)| \cdot |f(x) - f(x')| + (|f(x)| + 1) \cdot |g(x) - g(x')| \\
& <_{[\text{eqs: } 14.51, 14.52]} & |g(x)| \cdot \dfrac{\varepsilon}{2 \cdot (1 + |g(x)|)} + (|f(x)| + 1) \cdot \dfrac{\varepsilon}{2 \cdot (1 + |f(x)|)} \\
& < & \dfrac{\varepsilon}{2} + \dfrac{\varepsilon}{2} \\
& = & \varepsilon
\end{array}$$

proving that $f \cdot g$ is continuous at x. $\square$

Using mathematical induction we can then prove the following corollary

Corollary 14.147. *Let $n \in \mathbb{N}_0$, $\langle X, \|\|\rangle$ a pseudo normed space and $f: X \to \mathbb{K}$ a function continuous at x then*

$$f^n: X \to \mathbb{K} \text{ defined by } f^n(x) = (f(x))^n$$

is continuous at x. So if f is continuous then f^n is also continuous.

Proof. Take

$$S = \{n \in \mathbb{N}_0 | f^n \text{ is continuous at } x\}$$

then we have:

$0 \in S$. As $\forall x \in X$ $(f^0)(x) = (f(x))^0 = 1$ we have that $f^0 = \mathcal{C}_1$ which is continuous by [example: 14.131]. Hence $0 \in S$

$n \in S \Rightarrow n+1 \in S$. As $\forall x \in X$ $f^{n+1}(x) = (f(x))^{n+1} = (f(x))^n \cdot f(x) = (f^n \cdot f)(x)$ we have that $f^{n+1} = f^n \cdot f$. From $n \in S$ we have that f^n is continuous at x so that by the continuity of f at x and [theorem: 14.146] $f^{n+1} = f^n \cdot f$ is continuous at x. Hence $n+1 \in S$ $\square$

Using the identity function Id_X in the above corollary proves then that the power function itself is continuous.

Corollary 14.148. *Let $n \in \mathbb{N}_0$ then for $\langle \mathbb{K}, ||\rangle$ we have that the function*

$$.^n: \mathbb{K} \to \mathbb{K} \text{ defined by } (.^n)(x) = x^n$$

is continuous.

Proof. $\forall x \in X$ we have $(.^n)(x) = x^n = (\mathrm{Id}_X(x))^n = (\mathrm{Id}_X)^n(x)$ which proves that $\cdot^n = (\mathrm{Id}_X)^n$. So we can apply [example: 14.132] and the previous corollary [corollary: 14.147]. $\square$

Theorem 14.149. *Let $n \in \mathbb{N}$ and $\langle \mathbb{R}^n, \|\|_n \rangle$ the normed space with the maximum norm $\|\|_n$ defined by $\|x\|_n = \max(\{|\pi_i(x)| | i \in \{1,\ldots,n\}\})$ and $\|\|$ another norm on $\mathbb{R}^n$ then*

$$\|\|: \mathbb{R}^n \to \mathbb{R} \text{ defined by } \|\|(x) = \|x\|$$

is continuous [using the topological spaces $\langle \mathbb{R}^n, \|\|_n \rangle$ and $\langle \mathbb{R}, \|\rangle$

Proof. Let $E = \{e_1, \ldots, e_n\}$ be the canonical basis of $\mathbb{R}^n$ defined by $(e_i)_j = \delta_{i,j}$ [see: 11.159] then $\forall x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ we have $x = \sum_{i \in \{1,\ldots,n\}} x_i \cdot e_i$. Hence

$$\begin{aligned}
\|x\| &= \left\| \sum_{i \in \{1,\ldots,n\}} x_i \cdot e_i \right\| \\
&\leqslant_{[\text{theorem: } 14.86]} \sum_{i \in \{1,\ldots,n\}} \|x_i \cdot e_i\| \\
&= \sum_{i \in \{1,\ldots,n\}} |x_i| \cdot \|e_i\| \\
&\leqslant \sum_{i \in \{1,\ldots,n\}} \max\{|x_i| | i \in \{1,\ldots,n\}\} \cdot \|e_i\| \\
&= \sum_{i \in \{1,\ldots,n\}} \max\{|\pi_i(x)| | i \in \{1,\ldots,n\}\} \cdot \|e_i\| \\
&= \sum_{i \in \{1,\ldots,n\}} \|x\|_n \cdot \|e_i\| \\
&= \|x\|_n \cdot \sum_{i \in \{1,\ldots,n\}} \|e_i\| \\
&= \|x\|_n \cdot A \text{ where } A = \sum_{i \in \{1,\ldots,n\}} \|e_i\|
\end{aligned}$$

hence we have that

$$\forall x \in \mathbb{R}^n \text{ we have } \|x\| \leqslant \|x\| \cdot A$$

Let $x \in \mathbb{R}^n$ and $\varepsilon \in \mathbb{R}^+$ then if we take $\delta = \frac{\varepsilon}{A+1}$ then if $\|x - y\|_n \leqslant \delta$ we have

$$\|x - y\| \leqslant A \cdot \|x - y\|_n < A \cdot \delta = A \cdot \frac{\varepsilon}{1 + A} < \varepsilon$$

proving that $\|\|$ is continuous [see theorem: 14.127]. $\square$

14.4.2 Uniform and Lipschitz continuity

We have a stronger form of continuity in the case of normed spaces called uniform continuity.

Definition 14.150. *Let $\langle X, d_X \rangle$, $\langle Y, d_Y \rangle$ be topological spaces, then a function $f: X \to Y$ is **uniform continuous** if $\forall \varepsilon \in \mathbb{R}^+$ there exists a $\delta \in \mathbb{R}^+$ such that $\forall x, y \in X$ with $d_X(x, y) < \delta$ we have $d_Y(f(x), f(y)) < \varepsilon$*

Theorem 14.151. *Let $\langle X, \|\|_X \rangle$, $\langle Y, \|\|_Y \rangle$ be topological spaces, then a function $f: X \to Y$ is **uniform continuous** if $\forall \varepsilon \in \mathbb{R}^+$ there exists a $\delta \in \mathbb{R}^+$ such that $\forall x, y \in X$ with $\|x - y\|_X < \delta$ we have $\|f(x) - f(y)\|_Y < \varepsilon$*

Proof. This follows from the fact that $d_X(x, y) = \|x - y\|_Y$ $\square$

The difference with normal continuity at x is that there exist a δ **which can depend** on x such that $\forall y \in Y$ with $\|x - y\|_X$ we have $\|f(x) - f(y)\|_Y < \varepsilon$. However uniform continuity implies continuity as is mentioned in the following theorem.

Theorem 14.152. *Let $\langle X, \|\|_X \rangle$, $\langle Y, \|\|_Y \rangle$ be pseudo normed spaces and $f: X \to Y$ a uniform continuous function then f is continuous.*

Proof. Let $x \in X$ and $\varepsilon \in \mathbb{R}^+$ then there exist a $\delta \in \mathbb{R}^+$such that for $x, y \in X$ with $\|x - y\|_X < \delta$ we have $\|f(x) - f(y)\|_Y < \varepsilon$. Hence if $x \in X$ then for every $y \in X$ such that $\|x - y\|_X < \delta$ we have $\|f(x) - f(y)\|_Y < \varepsilon$ proving that f is continuous at x. As $x \in X$ was chosen arbitrary it follows that f is continuous. □

Example 14.153. Let $\langle X, \|\|\rangle$ be a pseudo normed space then

$$\|\|: X \to \mathbb{R} \text{ where } \|\|(x) = \|x\|$$

is **uniform continuous** [hence by [theorem: 14.150] continuous] using the topological spaces $\langle X, \mathcal{T}_{\|\|}\rangle$ and $\langle \mathbb{R}, \mathcal{T}_{\mathbb{R}_{||}}\rangle$

Proof. Let $x \in X$ then if $\varepsilon \in \mathbb{R}^+$ we have for $\delta = \varepsilon \in \mathbb{R}^+$ that $\forall x, y \in X$ such that $\|x - y\| < \delta = \varepsilon$

$$|\|\|(x) - \|\|(y)| = |\|x\| - \|y\|| \leqslant_{[\text{theorem: } 14.87]} \|x - y\| < \delta = \varepsilon$$

proving by [theorem: 14.127] that $\|\|$ is uniform continuous at x. □

Example 14.154. Given the topological spaces $\langle \mathbb{C}, \mathcal{T}_{||}\rangle$ and $\langle \mathbb{R}, \mathcal{T}_{||}\rangle$ [see examples: 14.96 and 14.95] then

1. Re: $\mathbb{C} \to \mathbb{R}$ is uniform continuous [hence by [theorem: 14.152] continuous]
2. Img: $\mathbb{C} \to \mathbb{R}$ is uniform continuous [hence by [theorem: 14.152] continuous]

Proof. Let $\varepsilon \in \mathbb{R}^+$, take $\delta = \varepsilon \in \mathbb{R}^+$ then $\forall x, y \in \mathbb{C}$ such that $|x - y| < \delta = \varepsilon$ we have

1. $|\text{Re}(x) - \text{Re}(y)| \underset{[\text{theorem: } 10.78]}{=} |\text{Re}(x - y)| \leqslant_{[\text{theorem: } 10.83]} |x - y| < \delta = \varepsilon$ proving that Re is uniform continuous at x.
2. $|\text{Img}(x) - \text{Img}(y)| \underset{[\text{theorem: } 10.78]}{=} |\text{Img}(x - y)| \leqslant_{[\text{theorem: } 10.83]} |x - y| < \delta = \varepsilon$ proving that Img is uniform continuous at x. □

Theorem 14.155. *Let $\langle X, \|\|\rangle$ be a pseudo normed space, $\mathcal{T}_{\|\|}$ the norm topology on X and $\mathcal{T}_{\text{product}}$ the product topology on $X \times X$ then $+: X \times X \to X$ is uniform continuous [hence by [theorem: 14.152] continuous] where $+(x, y) = x + y$.*

Proof. Using [theorem: 14.94] the product topology on $\mathcal{T}_{\text{product}} = \mathcal{T}_{\|\|_{\max}}$ where for $x = (x_1, x_2) \in X \times X$ $\|x\|_{\max} = \|(x_1, x_2)\|_{\max} = \max(\{\|x_i\| | i \in \{1, 2\}\})$. Let $\varepsilon \in \mathbb{R}^+$ and take $\delta = \frac{\varepsilon}{2}$ then $\forall x = (x_1, x_2), y = (y_1, y_2) \in X \times X$ with $\|x - y\|_{\max} < \delta$ we have $\max(\{\|x_1 - y_1\|, \|x_2 - y_2\|\}) < \delta$ so that

$$\|x_1 - y_1\| < \delta \wedge \|x_2 - y_2\| < \delta$$

Next

$$\begin{aligned}
\|+(x) - (+(y))\| &= \|(x_1 + x_2) - (y_1 + y_2)\| \\
&= \|(x_1 - y_1) + (x_2 - y_2)\| \\
&\leqslant \|x_1 - y_1\| + \|x_2 - y_2\| \\
&< \delta + \delta \\
&= \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\
&= \varepsilon
\end{aligned}$$

which proves that $+: X \times X \to X$ is uniform continuous. □

Definition 14.156. *Let $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$ be pseudo normed spaces then a function is **Lipschitz continuous** if $\forall x, y \in X$ we have $\|f(x) - f(y)\|_Y \leqslant \|x - y\|_X$*

It turns out that Lipschitz continuous is a stronger condition then uniform continuity.

Theorem 14.157. *Let $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$ be pseudo normed spaces and $f: X \to Y$ a **Lipschitz continuous function** then f is **uniform continuous** and by [theorem: 14.152] f is continuous.*

Proof. Let $\varepsilon \in \mathbb{R}^+$ and take $\delta = \varepsilon \in \mathbb{R}^+$ then if $x, y \in X$ satisfies $\|x - y\|_X < \delta$ we have

$$\|f(x) - f(y)\|_Y \leqslant \|x - y\|_X < \delta = \varepsilon$$

proving uniform continuity. □

Theorem 14.158. *Let $\langle X,\|\|\rangle$ be a pseudo normed space then $\forall x\in X$ we have that*

$$\delta_x: X\rightarrow X \text{ defined by } \delta_x(y)=x+y$$

is Lipschitz continuous [hence by [theorem: 14.157] absolute continuous and uniform continuous]

Proof. Let $x\in X$ and $y,z\in X$ then

$$\|\delta_x(y)-\delta_x(z)\|=\|x+y-(x+z)\|=\|y-z\|$$ □

An example of a Lipschitz continuous function is the set distance function to a non empty set.

Definition 14.159. *Let $\langle X,\|\|\rangle$ be a pseudo normed space and $\varnothing\neq S\subseteq X$ a non empty set then the* ***set distance function*** *noted by δ_S is defined by*

$$\delta_S: X\rightarrow\mathbb{R} \text{ where } \delta_S(x)=\inf(\{\|x-y\|\mid y\in S\})$$

***Note* 14.160.** *$\forall d\in\{\|x-y\|\mid y\in S\}$ we have $0\leqslant d$, hence 0 is a lower bound of $\{\|x-y\|\mid y\in S\}$,* further *$S\neq 0\Rightarrow\{\|x-y\|\mid y\in S\}\neq 0$. So using the conditional completeness of $\mathbb{R}$ [see theorem: 10.18] it follows that $\inf(\{\|x-y\|\mid y\in S\})$ exists.*
Further if $x\in S$ then $0\leqslant\inf(\{\|x-y\|\mid y\in S\})\leqslant 0$ proving that

$$\forall x\in S \text{ we have } \delta_S(x)=0$$

Theorem 14.161. *Let $\langle X,\|\|\rangle$ be a pseudo normed space $\varnothing\neq S\subseteq X$ then $\delta_S: X\rightarrow\mathbb{R}$ is Lipschitz continuous [hence uniform continuous and thus continuous [see theorems: 14.152, 14.157]*

Proof. Let $x,y\in X$ and $\varepsilon\in\mathbb{R}^+$ then as $\delta_S(x)=\inf(\{\|x-z\|\mid z\in S\})<\inf(\{\|x-z\|\mid z\in S\})+\varepsilon$ there exist a $z_\varepsilon\in S$ such that $\delta_S(x)\leqslant\|x-z_\varepsilon\|<\delta_S(x)+\varepsilon$ hence

$$-\delta_S(x)<\varepsilon-\|x-z_\varepsilon\|.$$

Further, as $z_\varepsilon\in S$, we have $\delta_S(y)=\inf(\{\|y-z\|\mid z\in S\})\leqslant\|y-z_\varepsilon\|$ hence we have that

$$\begin{aligned}\delta_S(y)-\delta_S(x) &< \|y-z_\varepsilon\|-\|x-z_\varepsilon\|+\varepsilon\\ &= \|y-x+x-z_\varepsilon\|-\|x-z_\varepsilon\|+\varepsilon\\ &\leqslant \|y-x\|+\|x-z_\varepsilon\|-\|x-z_\varepsilon\|+\varepsilon\\ &= \|y-x\|+\varepsilon\end{aligned}$$

So we have

$$\forall\varepsilon\in\mathbb{R}^+ \text{ we have } \delta_S(y)-\delta_S(x)<\|y-x\|+\varepsilon$$

which by [theorem: 10.31] proves that $\delta_S(y)-\delta_S(x)\leqslant\|y-x\|$. Interchanging x and y gives $-(\delta_S(y)-\delta_S(x))=\delta_S(x)-\delta_S(y)\leqslant\|x-y\|=\|y-x\|$, hence we have

$$|\delta_S(y)-\delta_S(x)|\leqslant\|y-x\|$$

proving Lipschitz continuity. □

14.4.3 Homeomorphism

Ring-, group, order- and linear- isomorphisms allows use to identify two rings, groups, partial ordered sets and vector spaces. We can do the same thing for topological spaces.

Definition 14.162. *Let $\langle X,\mathcal{T}_X\rangle$, $\langle Y,\mathcal{T}_Y\rangle$ be two topological spaces then a* ***homeomorphism*** *between $\langle X,\mathcal{T}_X\rangle$ and $\langle Y,\mathcal{T}\rangle$ is a* ***bijection*** *$f: X\rightarrow Y$ such that $f: X\rightarrow Y$ and $f^{-1}: Y\rightarrow X$ are* ***continuous****. If there exist a homeomorphism between two topological space then these spaces are called* ***homeomorphic*** *spaces.*

We have trivially the following equivalent definition of a homeomorphism

Theorem 14.163. *Let $\langle X,\mathcal{T}_X\rangle$, $\langle Y,\mathcal{T}_Y\rangle$ be two topological spaces and $f: X\rightarrow Y$ a bijection then*

$$f \text{ is a homeomorphism} \Leftrightarrow f \text{ is continuous and open}$$

Proof.

- **⇒.** As f is a homeomorphism we have that $f: X \rightarrow Y$ is continuous and $f^{-1}: Y \rightarrow X$ is continuous. So if $U \in \mathcal{T}_X$ we have $(f^{-1})^{-1}(U) \in \mathcal{T}_Y$ which, as $(f^{-1})^{-1}(U) \underset{\text{[theorem: 2.68]}}{=} f(U)$, proves that $f(U) \in \mathcal{T}_Y$. Hence $f: X \rightarrow Y$ is continuous and open.
- **⇐.** Let $U \in \mathcal{T}_X$ then as $f: X \rightarrow Y$ is open we have $f(U) \in \mathcal{T}_Y$, which as $(f^{-1})^{-1}(U) \underset{\text{[theorem: 2.68]}}{=} f(U)$ proves that $(f^{-1})^{-1}(U) \in \mathcal{T}_Y$. So $f^{-1}: Y \rightarrow X$ is continuous, as by the hypothesis $f: X \rightarrow Y$ is also continuous, it follows that $f: X \rightarrow Y$ is a homeomorphism. □

We have trivially that the the inverse of a homeomorphism is a homeomorphism.

Theorem 14.164. *Let $\langle X, \mathcal{T}_X \rangle$, $\langle Y, \mathcal{T}_Y \rangle$ be two topological spaces and $f: X \rightarrow Y$ a homeomorphism then $f^{-1}: Y \rightarrow X$ is a homeomorphism.*

Proof. This follows from the definition and [theorems: 2.72 and 2.73]. □

Theorem 14.165. *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 homeomorphisms then $g \circ f: X \rightarrow Z$ is a homeomorphism.*

Proof. This follows from the definition and [theorems: 2.75, 14.138]. □

Theorem 14.166. *Let $\langle X, \mathcal{T}_X \rangle$, $\langle Y, \mathcal{T}_Y \rangle$ be two topological spaces, $\varnothing \neq A \subseteq X$ a non empty subset of X, $f: X \rightarrow Y$ a homeomorphism then $f_{|A}: A \rightarrow f(A)$ is a homeomorphism [using the subspace topologies $(\mathcal{T}_X)_{|A}$ and $(\mathcal{T}_Y)_{|f(A)}$*

Proof. First, by [theorem: 2.88]

$$f_{|A}: A \rightarrow f(A) \text{ is a bijection} \tag{14.53}$$

Further if $V \in (\mathcal{T}_Y)_{|f(A)}$ then there exist a $V' \in \mathcal{T}_Y$ such that $V = V' \bigcap f(A)$. So

$$\begin{aligned}
(f_{|A})^{-1}(V) &= (f_{|A})^{-1}\left(V' \bigcap f(A)\right) \\
&\underset{\text{[theorem: 2.84]}}{=} A \bigcap f^{-1}\left(V' \bigcap f(A)\right) \\
&\underset{\text{[theorem: 2.93]}}{=} A \bigcap \left(f^{-1}(V') \bigcap f^{-1}(f(A))\right) \\
&\underset{\text{[theorem: 2.55]}}{=} A \bigcap \left(f^{-1}(V') \bigcap A\right) \\
&= f^{-1}(V') \bigcap A
\end{aligned}$$

As f is continuous and V' open we have that $f^{-1}(V') \in \mathcal{T}_X$ hence $f^{-1}(V') \bigcap A \in (\mathcal{T}_X)_{|A}$ proving that $(f_{|A})^{-1}(V) \in (\mathcal{T}_X)_{|A}$, Hence

$$f_{|A}: A \rightarrow f(A) \text{ is continuous} \tag{14.54}$$

Let $U \in (\mathcal{T}_X)_{|A}$ then there exist a $U' \in \mathcal{T}_X$ such that $U = U' \bigcap A$ then

$$f_{|A}(U) = f_{|A}\left(U' \bigcap A\right) \underset{U' \bigcap A \subseteq U}{=} f\left(U' \bigcap A\right) \underset{\text{[theorem: 2.93]} \wedge f_{|A} \text{ is bijective}}{=} f(U') \bigcap f(A)$$

By [theorem: 14.163] f is open so that $f(U') \in \mathcal{T}_Y$ hence $f(U') \bigcap f(A) \in (\mathcal{T}_Y)_{|f(A)}$ proving that $f_{|A}(U) \in (\mathcal{T}_Y)_{|f(A)}$. Hence we have that

$$f_{|A}: A \rightarrow f(A) \text{ is open} \tag{14.55}$$

Finally using [theorem: 14.163] on [eqs: 14.53, 14.54 and 14.55] proves that

$$f_{|A}: A \rightarrow f(A) \text{ is a homeomorphism using the topologies } (\mathcal{T}_X)_{|A} \text{ and } (\mathcal{T}_Y)_{|f(A)}$$ □

Every linear isometric isomorphism [see definition: 14.121] is automatically a homeomorphism.

Theorem 14.167. *Let $\langle X, \|\|_X \rangle$, $\langle Y, \|\|_Y \rangle$ be pseudo normed spaces over $\mathbb{K}$ and $f: X \rightarrow Y$ a llinear isometric isomorphism then $f: X \rightarrow Y$ is a homeomorphism using the topologies $\mathcal{T}_{\|\|_X}$, $\mathcal{T}_{\|\|_Y}$. In other words isometric spaces are homeomorphic.*

Proof. As $f: X \rightarrow Y$ is a linear isometric isomorphism it follows that

$$f: X \rightarrow Y \text{ is a bijection}$$

Further as $\forall x, y \in X$ we have $\|f(x) - f(y)\|_Y = \|f(x - y)\|_Y = \|x - y\|_X$ it follows that $f: X \rightarrow Y$ is Lipschitz continuous, hence continuous. Further by [theorem: 14.119] $f^{-1}: Y \rightarrow X$ is also a linear isometry hence continuous. So $f: X \rightarrow Y$ and $f^{-1}: Y \rightarrow X$ are continuous which proves that $f: X \rightarrow Y$ is homeomorphism. □

Theorem 14.168. *Let $\langle X, \mathcal{T}\rangle$ be a topological space. Y a set and $f: X \rightarrow Y$ a bijection. Then we have:*

1. *$\mathcal{T}_{f,\mathcal{T}} = \{f(U) | U \in \mathcal{T}\}$ forms a topology on Y*
2. *$f: X \rightarrow Y$ is a homeomorphism using the topologies $\mathcal{T}$ and $\mathcal{T}_{f,\mathcal{T}}$*
3. *If $g: X \rightarrow Y$ satisfies $f = g \circ h$ where $h: X \rightarrow X$ is a homeomorphism using the topologies $\mathcal{T}$ and $\mathcal{T}$ then g is a homeomorphism and $\mathcal{T}_{f,\mathcal{T}} = \mathcal{T}_{g,\mathcal{T}}$*

Proof.

1.
   a. As $f(\varnothing) = \varnothing$ it follows that
   $$\varnothing \in \mathcal{T}_{f,\mathcal{T}}$$
   b. As $f: X \rightarrow Y$ is a bijection we have $f(X) = Y$ hence as $X \in \mathcal{T}$
   $$Y \in \mathcal{T}_{f,\mathcal{T}}$$
   c. If $\{V_i\}_{i \in I} \subseteq \mathcal{T}_{f,\mathcal{T}}$ then $\forall i \in I$ there exist a $U \in \mathcal{T}$ such that $f(U) = V_i$, hence by a consequence of the Axiom of Choice [see theorem: 3.104] there exist a $\{U_i\}_{i \in I} \subseteq \mathcal{T}$ such that $\forall i \in I$ $f(U_i) = V_i$. Hence
   $$\bigcup_{i \in I} V_i = \bigcup_{i \in I} f(U_i) \underset{[\text{theorem: } 2.134]}{=} f\left(\bigcup_{i \in I} U_i\right)$$
   which, as $\bigcup_{i \in I} U_i \in \mathcal{T}$, proves that
   $$\bigcup_{i \in I} V_i \in \mathcal{T}_{f,\mathcal{T}}$$
   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
   $$V_1 \bigcap V_2 = f(U_1) \bigcap f(U_2) \underset{[\text{theorem: } 2.93]}{=} f\left(U_1 \bigcap U_2\right)$$
   which, as $U_1 \bigcap U_2 \in \mathcal{T}$, proves that
   $$V_1 \bigcap V_2 \in \mathcal{T}_{f,\mathcal{T}}$$

2. If $U \in \mathcal{T}$ then by definition $f(U) \in \mathcal{T}_{f,\mathcal{T}}$ proving that
   $$f \text{ is open}$$
   Further if $V \in \mathcal{T}_{f,\mathcal{T}}$ then there exist a $U \in \mathcal{T}$ so that $V = f(U)$, so
   $$f^{-1}(V) = f^{-1}(f(U)) \underset{[\text{theorem: } 2.55]}{=} U \in \mathcal{T}$$
   proving that
   $$f \text{ is continuous}$$
   Hence as $f: X \rightarrow Y$ is a bijection we have that
   $$f: X \rightarrow Y \text{ is s homeomorphism}$$

3. First as as $h: X \rightarrow X$ is a homeomorphism we have by [theorem: 14.164] that $h^{-1}: X \rightarrow X$ is a homeomorphism. Further
   $$g = g \circ \mathrm{Id}_X = g \circ (h \circ h^{-1}) = (g \circ h) \circ h^{-1} = f \circ h^{-1} \tag{14.56}$$
   so, as f is a homeomorphism by (2) and h^{-1} is a homeomorphism, we have by [theorem: 14.165] that
   $$g: X \rightarrow Y \text{ is a homeomorphism}$$
   Next if $V \in \mathcal{T}_{g,\mathcal{T}}$ then there exist a $U \in \mathcal{T}$ such that $V = g(U)$, as
   $$V = g(U) \underset{[\text{eq: } 14.56]}{=} (f \circ h^{-1})(U) = f((h^{-1})(U)) \in \mathcal{T}_{f,\mathcal{T}} \text{ [as } h^{-1} \text{ is open } (h^{-1})(U) \in \mathcal{T}] \tag{14.57}$$
   So
   $$\mathcal{T}_{g,\mathcal{T}} \subseteq \mathcal{T}_{f,\mathcal{T}} \tag{14.58}$$
   If $V \in \mathcal{T}_{f,\mathcal{T}}$ then there exist a $U \in \mathcal{T}$ such that $V = f(U)$, as
   $$f(U) = (g \circ h)(U) = g(h(U)) \in \mathcal{T}_{g,\mathcal{T}} \text{ [as } h \text{ is open } h(U) \in \mathcal{T}]$$

Proving that $\mathcal{T}_{f,\mathcal{T}} \subseteq \mathcal{T}_{g,\mathcal{T}}$, combining this with [eq: 14.58] gives finally

$$\mathcal{T}_{f,\mathcal{T}} = \mathcal{T}_{g,\mathcal{T}}$$

□

Theorem 14.169. *Let X be a set, $\langle Y, \mathcal{T}\rangle$ a topological space and $f: X \rightarrow Y$ a function then*

$$\mathcal{T}_{f,\mathcal{T}}^{-1} = \{f^{-1}(U) | U \in \mathcal{T}\} \text{ is a topology on } X$$

This topology is called the ***inverse induced topology of $\mathcal{T}$ by f***. *Further $f: X \rightarrow Y$ is continuous using the topologies $\mathcal{T}_{f,\mathcal{T}}^{-1}$ and $\mathcal{T}$.*

Proof. First we have

1. As $\varnothing = f^{-1}(\varnothing)$ we have
$$\varnothing \in \mathcal{T}_{f,\mathcal{T}}^{-1}$$
2. As $X = f^{-1}(Y)$ we have
$$X \in \mathcal{T}_{f,\mathcal{T}}^{-1}$$
3. Let $\{V_i\}_{i \in I} \subseteq \mathcal{T}_{f,\mathcal{T}}^{-1}$ then $\forall i \in I$ there exist a $U \in \mathcal{T}$ such that $f^{-1}(U) = V_i$, hence by a consequence of the Axiom of Choice [see theorem: 3.104] there exist a $\{U_i\}_{i \in I} \subseteq \mathcal{T}$ such that $\forall i \in I$ $f^{-1}(U_i) = V_i$. Hence
$$\bigcup_{i \in I} V_i = \bigcup_{i \in I} f^{-1}(U_i) \underset{[\text{theorem: } 2.134]}{=} f^{-1}\left(\bigcup_{i \in I} U_i\right) \in \mathcal{T}_{f,\mathcal{T}}^{-1} \text{ [as } \bigcup_{i \in I} U_i \in \mathcal{T}]$$
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
$$V_1 \bigcap V_2 = f^{-1}(U_1) \bigcap f^{-1}(U_2) \underset{[\text{theorem: } 2.93]}{=} f^{-1}(U_1 \bigcap U_2) \text{ [as } U_1 \bigcap U_2 \in \mathcal{T}]$$

□

Theorem 14.170. *Let $\langle X, \|\|_X\rangle$ be a finite dimensional pseudo normed vector space with $\dim(X) = n$ then we have given a basis $E = \{e_1, \ldots, e_n\}$ for E that*

1. *$\|\|_{\langle E\rangle}: \mathbb{K}^n \rightarrow \mathbb{R}$ defined by $\|(x_1, \ldots, x_n)\|_{\langle E\rangle} = \left\|\sum_{i \in \{1, \ldots, .n\}} x_i \cdot e_i\right\|_X$ is a pseudo norm on $\mathbb{K}^n$. Further if $\|\|_X$ is a norm then $\|\|_{\langle E\rangle}$ is a norm.*
2. *$\varphi: \mathbb{K}^n \rightarrow X$ defined by $\varphi(x_1, \ldots, x_n) = \sum_{i \in \{1, \ldots, .n\}} x_i \cdot e_i$ is a linear isometric isomorphism using the norms $\|\|_{\langle E\rangle}$ and $\|\|_X$*

Hence by [theorem: 14.167] φ is a homeomorphism so that $\mathbb{K}^n$ is homeomorphic with X.

Proof.

1. To prove that $\|\|_{\langle E\rangle}$ is a norm
   a. $\forall x = (x_1, \ldots, x_n) \in \mathbb{K}^n$ we have $\|x\|_{\langle E\rangle} = \left\|\sum_{i \in \{1, \ldots, .n\}} x_i \cdot e_i\right\|_X \geqslant 0$
   b. If $\alpha \in \mathbb{K}$, $x = (x_1, \ldots, x_n) \in \mathbb{K}^n$ we have
$$\begin{aligned}
\|\alpha \cdot x\|_{\langle E\rangle} &= \|(\alpha \cdot x_1, \ldots, \alpha_n \cdot x_n)\|_{\langle E\rangle} \\
&= \left\|\sum_{i \in \{1, \ldots, .n\}} (\alpha \cdot x_i) \cdot e_i\right\|_X \\
&= \left\|\sum_{i \in \{1, \ldots, .n\}} \alpha \cdot (x_i \cdot e_i)\right\|_X \\
&= \left\|\alpha \cdot \sum_{i \in \{1, \ldots, .n\}} x_i \cdot e_i\right\|_X \\
&= |\alpha| \cdot \left\|\sum_{i \in \{1, \ldots, .n\}} x_i \cdot e_i\right\|_X \\
&= |\alpha| \cdot \|x\|_{\langle E\rangle}
\end{aligned}$$

c. Let $x=(x_1,\ldots,x_n),\ y=(y_1,\ldots,y_n)\in\mathbb{K}^n$ then we have

$$\begin{array}{rcl}\|x+y\|_{\langle E\rangle} & = & \|(x_1+y_1,\ldots,x_n+y_n)\|_{\langle E\rangle}\\ & = & \left\|\sum_{i\in\{1,\ldots,.n\}}(x_i+y_i)\right\|_X\\ & = & \left\|\sum_{i\in\{1,\ldots,.n\}}x_i+\sum_{i\in\{1,\ldots,.n\}}y_i\right\|_X\\ & \leqslant & \left\|\sum_{i\in\{1,\ldots,.n\}}x_i\right\|_X+\left\|\sum_{i\in\{1,\ldots,.n\}}y_i\right\|_X\\ & = & \|x\|_{\langle E\rangle}+\|y\|_{\langle E\rangle}\end{array}$$

Assume that $\|\|_X$ is a norm. If $x=(x_1,\ldots,x_n)\in\mathbb{K}^n$ satisfies $\|x\|_{\langle E\rangle}=0$ then

$$\left\|\sum_{i\in\{1,\ldots,.n\}}x_i\cdot e_i\right\|_X=\|x\|_{\langle E\rangle}=0.$$

So as, $\|\|_X$ is a norm, $0=\sum_{i\in\{1,\ldots,.n\}}x_i\cdot e_i$. As $\{e_1,...,e_n\}$ is a basis we have by linear independency that $\forall i\in\{1,...,n\}\ x_i=0$ or $x=0$.

2. For $\varphi:\mathbb{K}^n\rightarrow X$ we have:

   injectivity. Let $\varphi(x)=\varphi(y)$ then $\sum_{i\in\{1,\ldots,n\}}x_i\cdot e_i=\sum_{i\in\{1,\ldots,n\}}y_i\cdot e_i$, hence, as $\{e_1,\ldots,e_n\}$ is a basis we have by [theorem: 11.137] that $\forall i\in\{1,\ldots,n\}\ x_i=y_i$. Hence $x=y$.

   surjectivity. If $y\in X$ then by [theorem: 11.137] there exists a $\{y_i\}_{i\in\{,\ldots,\ldots,n\}}\subseteq\mathbb{K}$ such that $y=\sum_{i\in\{1,\ldots,n\}}y_i\cdot e_i$. So if we take $x=(y_1,\ldots,y_n)$ then $\varphi(x)=y$.

From the above it follows that

$$\varphi:\mathbb{K}^n\rightarrow X\text{ is a bijection}$$

Let $x=(x_1,\ldots,x_n),\ y=(y_1,\ldots,y_n)\in\mathbb{K}^n$ and $\alpha\in\mathbb{K}$ then we have

$$\begin{array}{rcl}\varphi(x+\alpha\cdot y) & = & \sum_{i\in\{1,\ldots,n\}}(x_i+\alpha\cdot y_i)\cdot e_i\\ & = & \sum_{i\in\{1,\ldots,n\}}x_i\cdot e_i+\alpha\cdot\sum_{i\in\{1,\ldots,n\}}y_i\cdot e_i\\ & = & \varphi(x)+\alpha\cdot\varphi(y)\end{array}$$

proving that

$$\varphi\text{ is linear}$$

Finally if $x=(x_1,\ldots,x_n)\in\mathbb{K}^n$ then we have

$$\|\varphi(x)\|_X=\left\|\sum_{i\in\{1,\ldots,n\}}x_i\cdot e_i\right\|_X=\|x\|_{\langle E\rangle}$$

proving that

$$\varphi:\mathbb{K}^n\rightarrow X\text{ is a linear isometric isomorphism}$$

□

14.5 Linear mappings and continuity

Definition 14.171. *Let $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$ be normed spaces then the set of linear continuous function graphs is noted as $L(X,Y)$. So*

$$L(X,Y)=\{L\in\mathrm{Hom}(X,Y)|L:X\rightarrow Y\textit{ is continuous using the topologies }\mathcal{T}_{\|\|_X}\textit{ and }\mathcal{T}_{\|\|_Y}\}$$

Example 14.172. Let $\langle X,\|\|\rangle$ be a normed space then $\mathrm{Id}_X\in L(X,X)$

Proof. This follows from [theorem: 11.173] and [example: 14.132]. $\square$

Theorem 14.173. *Let $\langle X, \|\|_X\rangle$ and $\langle Y, \|\|_Y\rangle$ be normed spaces over $\mathbb{K}$ then $L(X,Y)$ is a subspace of the vector space* $\operatorname{Hom}(X,Y)$ *[see theorem: 11.175]. Hence by [theorem: 11.57] $L(X,Y)$ is a vector space.*

Proof. Let $L_1, L_2 \in L(X,Y)$ and $\alpha \in \mathbb{K}$ then by [theorem: 14.144] $L_1 + \alpha \cdot L_2 \in L(X,Y)$. Further as C_0 is continuous [see theorem: 14.131] we have that $C_0 \in L(X,Y)$ hence $L(X,Y) = \varnothing$. $\square$

We examine now the conditions for a linear mapping to be continuous in the normed topologies.

Theorem 14.174. *Let $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$ be normed spaces and $L \in \operatorname{Hom}(X,Y)$ then we have the following equivalences:*

1. *L is continuous [in other words $L \in L(X,Y)$]*
2. *L is continuous at $0 \in X$*
3. *$\exists M \in \mathbb{R}_0^+$ such that $\forall x \in X$ with $\|x\|_X = 1$ we have $\|L(x)\|_Y \leqslant M$*
4. *$\exists M \in \mathbb{R}_0^+$ such that $\forall x \in X$ we have $\|L(x)\|_Y \leqslant M \cdot \|x\|_X$*

Proof.

1 ⇒ 2. This follows from the definition of continuity.

2 ⇒ 3. As L is continuous at 0 we have by [theorem: 14.127] that there exist a $\delta \in \mathbb{R}^+$ such that $\forall x \in X$ satisfying $\|x\|_X = \|x - 0\|_X < \delta$ we have $\|L(x)\|_Y = \|L(x) - L(0)\|_Y < 1$. As $0 < \delta$ there exist by [theorem: 10.30] a $\delta' \in \mathbb{R}$ such that $0 < \delta' < \delta$. If $x \in X$ such that $\|x\|_X = 1$ then $\|\delta' \cdot x\|_X = |\delta'| \cdot \|x\|_X = \delta' \cdot 1 = \delta' < \delta$ so that $\delta' \cdot \|L(x)\|_Y = \|\delta' \cdot L(x)\|_Y = \|L(\delta' \cdot x)\|_Y < 1$, hence $\|L(x)\|_Y < \frac{1}{\delta'}$. So if we take $M = \frac{1}{\delta'} \in \mathbb{R}^+ \subseteq \mathbb{R}_0^+$ then $\forall x \in X$ with $\|x\|_X = 1$ we have $\|L(x)\|_Y < M \Rightarrow \|L(x)\|_Y \leqslant M$.

3 ⇒ 4. By (3) there exist a $M \in \mathbb{R}_0^+$ such that $\forall x \in X$ with $\|x\|_X = 1$ we have $\|L(x)\|_Y \leqslant M$. Let $x \in X$ then we have either:

$\|x\|_X = 0$. Then as $\|\|_X$ is a norm $x = 0$ hence $\|L(x)\|_X = \|L(0)\|_Y = \|0\|_Y = 0 = M \cdot \|x\|_X$

$\|x\|_X \neq 0$. Then we have $\left\|\frac{1}{\|x\|_X} \cdot x\right\|_X = \left|\frac{1}{\|x\|_X}\right| \cdot \|x\|_x = \frac{\|x\|_X}{\|x\|_X} = 1$ so that

$$\frac{1}{\|x\|_X} \cdot \|L(x)\|_Y = \left\|\frac{1}{\|x\|_X} \cdot L(x)\right\|_Y = \left\|L\left(\frac{1}{\|x\|_X} \cdot x\right)\right\|_Y \leqslant M$$

Hence

$$\|L(x)\|_Y \leqslant M \cdot \|x\|_X$$

4 ⇒ 1. Let $M \in \mathbb{R}_0^+$ such that $\forall x \in X$ we have $\|L(x)\|_Y \leqslant M \cdot \|x\|_X$. Let $x \in X$ and $\varepsilon \in \mathbb{R}^+$ then, if we take $\delta = \frac{\varepsilon}{M+1}$, we have for $y \in X$ with $\|x - y\|_X < \delta$ that

$$\|L(x) - L(y)\|_X = \|L(x - y)\|_Y \leqslant M \cdot \|x - y\|_X \leqslant M \cdot \frac{\varepsilon}{M+1} < \varepsilon$$

So L is continuous at x and as $x \in X$ was chosen arbitrary it follows that f is continuous. $\square$

A consequence of the above is the following corollary.

Corollary 14.175. *Let $n, m \in \mathbb{N}$ and $\langle \mathbb{K}^n, \|\|_n\rangle$, $\langle \mathbb{K}^m, \|\|\rangle$ where*

$$\|x\|_n = \max(\{|\pi_i(x)| | i \in \{1, \ldots, n\}\}),\ \|x\|_m = \max(\{|\pi_i(x)| | i \in \{1, \ldots, m\}\})$$

are the maximum norms then every linear function $L: \mathbb{K}^n \rightarrow \mathbb{K}^m$ is continuous.

Proof. Let $\{e_i\}_{i \in \{1, \ldots, n\}}$ be the canonical basis of $\mathbb{K}^n$ [see 11.159] then $\forall x = (x_1, \ldots, x_n) \in \mathbb{K}^n$ we have $x = \sum_{i \in \{1, \ldots, n\}} x_i \cdot e_i = \sum_{i \in \{1, \ldots, n\}} \pi_i(x) \cdot e_i$ where $\pi_i(x) = x_i$. So that

$$\|L(x)\|_m = \left\|\sum_{i \in \{1, \ldots, n\}} \pi_i(x) \cdot L(e_i)\right\|_m \leqslant \sum_{i \in \{1, \ldots, n\}} |\pi_i(x)| \cdot \|L(e_i)\|_m \leqslant M' \cdot \sum_{i \in \{1, \ldots, n\}} |\pi_i(x)|$$

where $M' = \max(\{\|L(e_i)\|_m | i \in \{1, \ldots, n\}\})$. As $\forall i \in \{1, \ldots, n\}$ we have

$$|\pi_i(x)| \leqslant \max(\{|\pi_i(x)| | i \in \{1, \ldots, n\}\}) = \|x\|_n$$

it follows that $\sum_{i \in \{1,\ldots,n\}} |\pi_i(x)| \leqslant \sum_{i \in \{1,\ldots,n\}} \|x\|_m = n \cdot \|x\|_m$. So if $M = n \cdot M'$ then

$$\|L(x)\|_m \leqslant M \cdot \|x\|_m$$

By the previous theorem [theorem: 14.174] it follows then that

$$L: \mathbb{K}^n \rightarrow \mathbb{K}^m \text{ is continuous}$$

□

Corollary 14.176. *Let $n, m \in \mathbb{N}$ and $\langle \mathbb{R}^n, \|\|_n \rangle$ where $\|\|_n$ is the maximum norm then every linear isomorphism $L: \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a homeomorphism*

Proof. This follows directly from the previous corollary [corollary: 14.175]. □

Theorem 14.177. *Let $\langle X, \|\|_X \rangle$, $\langle Y, \|\|_Y \rangle$ be normed spaces and $L \in L(X, Y)$ then for $A_{L.s}$ and $A_{L,r}$ defined by*

$$A_{l,s} = \{M \in \mathbb{R}_0^+ | \forall x \in X \textit{ with } \|x\|_X = 1 \textit{ we have } \|L(x)\|_Y \leqslant M\}$$

and

$$A_{L,r} = \{M \in \mathbb{R}_0^+ | \forall x \in X \textit{ we have } \|L(x)\|_Y \leqslant M \cdot \|x\|_X\}$$

we have

1. $A_{L,s} = A_{L,r}$
2. $\inf(A_{L,r})$ *exist and* $0 \leqslant \inf(A_{L,r})$ *[or by (1)* $\inf(A_{L,r})$ *exist and* $0 \leqslant \inf(A_{L,r})$*)*

Proof.

1. Let $M \in A_{L,s}$ then $\forall x \in X$ we have either:

   $\|x\|_X = 0$. Then as $\|\|_X$ is a norm $x = 0$ and thus

   $$\|L(x)\|_Y = \|L(0)\|_Y = \|0\|_Y = 0 \leqslant M \cdot 0 = M \cdot \|x\|_X$$

   $\|x\|_X \neq 0$. Then $\left\| \frac{1}{\|x\|_X} \cdot x \right\|_X = \frac{1}{\|x\|_X} \cdot \|x\|_X = 1$ so tha,t as $M \in A_{L,s}$, we have

   $$\frac{1}{\|x\|_X} \cdot \|L(x)\|_Y = \left\| \frac{1}{\|x\|_X} \cdot L(x) \right\|_Y = \left\| L\left(\frac{1}{\|x\|_X} \cdot x \right) \right\|_Y \leqslant M$$

   hence

   $$\|L(x)\|_X \leqslant M \cdot \|x\|_X$$

   So as in all cases $\|L(x)\|_X \leqslant M \cdot \|x\|_X$ for $x \in X$, hence it follows that $M \in A_{L,r}$ proving that

   $$A_{L,s} \subseteq A_{L,r} \tag{14.59}$$

   If $M \in A_{L,r}$ then if $x \in X$ with $\|x\|_X = 1$ we have $\|L(x)\|_X \leqslant M \cdot \|x\|_X = M \cdot 1 = M$ so that $M \in A_{L,s}$, hence $A_{L,r} \subseteq A_{L,s}$. Combining this with [eq: 14.59] proves that

   $$A_{L,s} = A_{L,r}$$

2. As $L \in L(X, Y)$ we have by [theorem: 14.174] that there exist a $M \in \mathbb{R}_0^+$ such that $\forall x \in X$ with $\|x\|_X = 1$ $\|L(x)\|_Y \leqslant M$ proving that $M \in A_{L,r}$. Hence we have

   $$A_{L,r} \neq \varnothing$$

   Further, as $A_{L,r} \subseteq \mathbb{R}_0^+$, it follows that $A_{L,r}$ is bounded below by 0, so as $\mathbb{R}$ is conditional complete [see theorem: 10.18] we have that

   $$\inf(A_{L,r}) \text{ exist and } 0 \leqslant \inf(A_{L,r})$$

□

The above theorem allows the following definition.

Definition 14.178. (Operator Norm) *Let , $\langle Y, \|\|_Y \rangle$ be normed spaces and $L \in L(X, Y)$ then the **operator norm**, noted as $\|\|_{L(X,Y)}$ is defined by*

$$\|L\|_{L(X,Y)} = \inf(A_{L,s}) \underset{\text{[theorem: 14.177]}}{=} \inf(A_{L,r})$$

where

$$A_{L,s} = \{M \in \mathbb{R}_0^+ | \forall x \in X \textit{ with } \|x\|_X = 1 \textit{ we have } \|L(x)\|_Y \leqslant M\}$$

and

$$A_{L,r} = \{M \in \mathbb{R}_0^+ | \forall x \in X \text{ we have } \|L(x)\|_Y \leqslant M \cdot \|x\|_X\}$$

By [theorem: 14.173] $L(X,Y)$ is a vector space, we prove now that $\|\|_{L(X,Y)}$ is a norm on $L(X,Y)$.

Theorem 14.179. *Let $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$ be normed spaces over a field $\mathbb{K}$ then we have that:*

1. *$\forall x \in X$ we have $\|L(x)\|_Y \leqslant \|L\|_{L(X,Y)} \cdot \|x\|_Y$*
2. *$\|C_0\|_{L(X,Y)} = 0$ where $C_0: X \to Y$ is defined by $C_0(x) = 0$ [the neutral element in $L(X,Y)$*
3. *$\|\|_{L(X,Y)}$ is a norm on $L(X,Y)$ making $\langle L(X,Y), \|\|_{L(X,Y)}\rangle$ a normed space.*

Proof.

1. Let $x \in X$ then we have either:

   $\|x\|_X = 0$. Then
   $$\|L(x)\|_Y = \|L(0)\|_Y = \|0\|_Y = 0 = \|L\|_{L(X,Y)} \cdot 0 = \|L\|_{L(X,Y)} \cdot \|x\|_X$$

   $\|x\|_X \neq 0$. Assume that $\|L\|_{L(X,Y)} < \frac{1}{\|x\|_X} \cdot \|L(x)\|_Y$ then, as $\|L\|_{L(X,Y)} = \inf(A_{L,r})$, there exist a
   $$M \in A_{L,r} = \{M \in \mathbb{R}_0^+ | \forall x \in X \text{ we have } \|L(x)\|_Y \leqslant M \cdot \|x\|_X\}$$
   such that $\|L\|_{L(X,Y)} \leqslant M < \frac{1}{\|x\|_X} \cdot \|L(x)\|_Y$ so that $M \cdot \|x\|_X < \|L(x)\|_Y$ contradicting the fact that $M \in A_{L,r} \Rightarrow \|L(x)\|_Y \leqslant M \cdot \|x\|_X$. So we must have that $\frac{1}{\|x\|_X} \cdot \|L(x)\|_Y \leqslant \|L\|_{L(X,Y)}$ or
   $$\|L(x)\|_Y \leqslant \|L\|_{L(X,Y)} \cdot \|x\|_X$$

   So in all cases we have
   $$\|L(x)\|_Y \leqslant \|L\|_{L(X,Y)} \cdot \|x\|_X$$

2. $\forall x \in X$ we have $\|C_0(x)\|_Y = \|0\|_Y = 0 \leqslant 0 \cdot \|x\|_X$ proving that $0 \in A_{C_0,r}$, hence
   $$0 \leqslant \inf(A_{C_0,r}) \leqslant 0$$
   resulting in
   $$\|C_0\|_{L(X,Y)} = \inf(A_{C_0,r}) = 0$$

3. Let $L \in L(X,Y)$ then by [theorem: 14.177] we have
   $$\forall L \in L(X,Y) \text{ that } 0 \leqslant \|L\|_{L(X,Y)} \tag{14.60}$$
   Further if $L_1, L_2 \in L(X,Y)$ then for $x \in X$ we have by (1) that
   $$\|L_1(x)\|_Y \leqslant \|L\|_{L(X,Y)} \cdot \|x\|_X \text{ and } \|L_2(x)\|_Y \leqslant \|L_2\|_{L(X,Y)} \cdot \|x\|_X \tag{14.61}$$
   Further
   $$\begin{array}{rcl} \|(L_1 + L_2)(x)\|_Y & = & \|L_1(x) + L_2(x)\|_Y \\ & \leqslant & \|L_1(x)\|_Y + \|L_2(x)\|_Y \\ & \leqslant_{[\text{eq: } 14.61]} & \|L\|_{L(X,Y)} \cdot \|x\|_X + \|L_2\|_{L(X,Y)} \cdot \|x\|_x \\ & = & (\|L\|_{L(X,Y)} + \|L\|_{L(X,Y)}) \cdot \|x\|_X \end{array}$$
   proving that $\|L\|_{L(X,Y)} + \|L\|_{L(X,Y)} \in A_{(L_1+L_2),r}$ so that
   $$\|L_1 + L_2\|_{L(X,Y)} = \inf(A_{(L_1+L_2),r}) \leqslant \|L\|_{L(X,Y)} + \|L\|_{L(X,Y)} \tag{14.62}$$
   Let $\alpha \in \mathbb{K}$ and $L \in L(X,Y)$. By (1) we have $\forall x \in X$ that $\|L(x)\|_Y \leqslant \|L\|_{L(X,Y)} \cdot \|x\|_X$ then
   $$\|(\alpha \cdot L)(x)\|_X = \|\alpha \cdot L(x)\|_X = |\alpha| \cdot \|L(x)\|_Y \leqslant |\alpha| \cdot \|L\|_{L(X,Y)} \cdot \|x\|_X$$
   proving that $|\alpha| \cdot \|L\|_{L(X,Y)} \in A_{\alpha \cdot L,r}$. Hence
   $$\|\alpha \cdot L\|_{L(X,Y)} = \inf(A_{\alpha \cdot L,r}) \leqslant |\alpha| \cdot \|L\|_{L(X,Y)} \tag{14.63}$$

For the opposite inequality we have two cases for $|\alpha|$ to consider:

$|\alpha| = 0$. Then, as $||$ is a norm in $\mathbb{K}$, we have $\alpha = 0$ so that $\alpha \cdot L = C_0$, hence

$$\|\alpha \cdot L\|_{L(X,Y)} = \|C_0\|_{L(X,Y)} \underset{(2)}{=} 0 = |\alpha| \cdot \|L\|_{L(X,Y)}$$

or

$$|\alpha| \cdot \|L\|_{L(X,Y)} = 0 = \|\alpha \cdot L\|_{L(X,Y)} \Rightarrow |\alpha| \cdot \|L\|_{L(X,Y)} \leqslant \|\alpha \cdot L\|_{L(X,Y)}$$

$|\alpha| \neq 0$. As $\forall x \in X$ we have $\|(\alpha \cdot L)(x)\|_Y = \|\alpha \cdot L(x)\|_Y = |\alpha| \cdot \|L(x)\|_Y \leqslant |$ so that

$$\|L(x)\|_Y = \frac{\|(\alpha \cdot L)(x)\|_Y}{|\alpha|} \leqslant \frac{\|\alpha \cdot L\|_{L(X,Y)} \cdot \|x\|_X}{|a|}$$

proving that $\frac{\|\alpha \cdot L\|_{L(X,Y)}}{|\alpha|} \in A_{L,r}$, hence $\|L\|_{L(X,Y)} = \inf(A_{L,r}) \leqslant \frac{\|\alpha \cdot L\|_{L(X,Y)}}{|\alpha|}$. So

$$|\alpha| \cdot \|L\|_{L(X,Y)} \leqslant \|\alpha \cdot L\|_{L(X,Y)}$$

So in all cases we have $|\alpha| \cdot \|L\|_{L(X,Y)} \leqslant \|\alpha \cdot L\|_{L(X,Y)}$ which combined with [eq: 14.63] results in

$$\|\alpha \cdot L\|_{L(X,Y)} = |\alpha| \cdot \|L\|_{L(X,Y)} \tag{14.64}$$

Using [eqs: 14.60, 14.62 and 14.64] it follows that

$$\|\|_{L(X,Y)} \text{ is a pseudo norm}$$

To prove that $\|\|_{L(X,Y)}$ is a norm let $L \in L(X,Y)$ such that $\|L\|_{L(X,Y)}$. Assume that $L \neq C_o$ [the neutral element in $L(X,Y)$] then there exist a $x \in X$ such that $L(x) \neq C_o(x) = 0$. As $\|\|_Y$ is a norm we have $\|L(x)\|_Y \neq 0$, but then as $0 \neq \|L(x)\|_Y \leqslant_{(1)} \|L\|_{L(X,Y)} \cdot \|x\|_X = 0$ resulting in a contradiction. So the assumption is wrong and we must have that $L = C_0$ which proves that

$$\|\|_{L(X,Y)} \text{ is a norm on } L(X,Y)$$

□

Up to now the operator norm $\|\|_{L(\mathrm{XY})}$ is defined using a infinum, the following theorem shows that we can also define the operator norm as a supremum.

Theorem 14.180. *Let $\langle X \|\|_X \rangle$, $\langle Y \|\|_Y \rangle$ be normed spaces where X is not trivial [there exists a $x \in X$ with $x \neq 0$], $L \in L(X,Y)$ and define*

$$A_{L,t} = \{\|L(x)\|_Y | x \in X \text{ with } \|x\|_X = 1\}$$

and

$$A_{L,u} = \{\|L(x)\|_Y | x \in X \text{ with } \|x\|_X \leqslant 1\}$$

then

$$\|L\|_{L(X,Y)} = \sup(A_{L,t}) = \sup(A_{L,u})$$

Proof. If $M \in A_{L,t}$ then $\exists x \in X$ with $\|x\|_X = 1$ such that $M = \|L(x)\|_Y$ which as $\|x\|_X = 1 \leqslant 1$ proves that $M \in A_{L,u}$. Hence

$$A_{L,t} \subseteq A_{L,u} \tag{14.65}$$

As $\exists x_0 \in X$ with $x_0 \neq 0$ we have as $\|\|_X$ is a norm that $\left\|\frac{1}{\|x_0\|} \cdot x_0\right\|_X = \frac{1}{\|x_0\|_X} \cdot \|x_0\|_X = 1$ so that $\|L(x_0)\|_Y \in A_{L,t}, A_{L,u}$ proving that

$$\varnothing \neq A_{L.t} \wedge \varnothing = A_{L,u} \tag{14.66}$$

Now if $M \in A_{L,u}$ then $\exists x \in X$ with $\|x\|_X \leqslant 1$ such that $M = \|L(x)\|_Y$, using [theorem: 14.179] we have that $\|L(x)\|_Y \leqslant \|L\|_{L(X,Y)} \cdot \|x\|_X \leqslant \|L\|_{L(X,Y)} \cdot 1 = \|L\|_{L(X,Y)}$, proving that $\|L\|_{L(X,Y)}$ is a upper bound for $A_{L,u}$ and by [eq: 14.65] $A_{L,u}$. Hence by [theorem: 10.18] $\sup(A_{L,t})$, $\sup(A_{L,u})$ exists and $\sup(A_{L,t}), \sup(A_{L,u}) \leqslant \|L\|_{L(X,Y)}$. Further by by [eq: 14.65] and [theorem: 3.74] we have that $\sup(A_{L,t}) \leqslant \sup(A_{L,u})$. Summarized we have that

$$\sup(A_{L,t}) \leqslant \sup(A_{L,u}) \leqslant \|L\|_{L(X,Y)}$$

Now $\forall x \in X$ with $\|x\|_X = 1$ we have that $\|L(x)\|_Y \in A_{L,t}$ so that $\|L(x)\|_Y \leqslant \sup(A_{L,t})$ or using the definition of $A_{L,s}$ [see: 14.178] that $\sup(A_{L,t}) \in A_{L,s}$. Hence

$$\|L\|_{L(X,Y)} = \inf(A_{L,s}) \leqslant \sup(A_{L,t}) \leqslant \sup(A_{L,u}) \leqslant \|L\|_{L(X,Y)}$$

proving that

$$\sup(A_{L,t}) = \sup(A_{L,u}) = \|L\|_{L(X,Y)}$$

□

To summarize the the above theorems about $L(X,Y)$ we have

Theorem 14.181. *Let $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$ be normed spaces then*

$$\langle L(X,Y), \|\|_{L(X,Y)}\rangle$$

is a normed space where

$$\forall L \in L(X,Y) \text{ we have } \|L\|_{L(X,Y)} = \inf(A_{L,s}) = \inf(A_{L,r}) = \sup(A_{L,t}) = \sup(A_{L,u})$$

and

$$\begin{array}{rcl} A_{L,s} &=& \{M \in \mathbb{R}_0^+ | \forall x \in X \text{ with } \|x\|_X = 1 \text{ we have } \|L(x)\|_Y \leqslant M\} \\ A_{L,r} &=& \{M \in \mathbb{R}_0^+ | \forall x \in X \text{ we have } \|L(x)\|_Y \leqslant M \cdot \|x\|_X\} \\ A_{L,t} &=& \{\|L(x)\|_Y | x \in X \text{ with } \|x\|_X = 1\} \\ A_{L,u} &=& \{\|L(x)\|_Y | x \in X \text{ with } \|x\|_X \leqslant 1\} \end{array}$$

Proof. This follows from [theorems: 14.173, 14.179, 14.180]. □

Example 14.182. Let $\langle X, \|\|\rangle$ be a normed space then $\mathrm{Id}_X \in L(X,X)$

Proof. From [example: 14.182] $\mathrm{Id}_X \in L(X,X)$. Further for $x \in X$ we have $\|\mathrm{Id}_X(x)\| = \|x\|$ which, using [theorem: 14.181], proves that $\|\mathrm{Id}_X\|_{L(X,X)} \leqslant 1$. □

Example 14.183. Let $\langle X, \|\|\rangle$ be a normed space, $\langle \mathbb{K}, ||\rangle$ be normed space of real or complex numbers and $x \in X$ then if we define

$$(\cdot x): \mathbb{K} \to X \text{ by } (\cdot x)(r) = r \cdot x$$

we have that

$$(\cdot x) \in L(\mathbb{K}, X)$$

with

$$\|(\cdot x)\|_{L(\mathbb{K},X)} = \|x\|$$

Proof. Let $r, s \in \mathbb{K}$ and $\alpha \in \mathbb{K}$ then we have

$$(\cdot x)(r + \alpha \cdot s) = (r + \alpha \cdot s) \cdot x = r \cdot x + (\alpha \cdot s) \cdot x = r \cdot x + \alpha \cdot (s \cdot x) = (\cdot x)(r) + \alpha \cdot (\cdot x)(s)$$

proving that

$$(\cdot x) \in \mathrm{Hom}(\mathbb{K}, X)$$

Let $r \in \mathbb{K}$ be such that $|r| = 1$ then $\|(\cdot x)(r)\| = \|r \cdot x\| = |r| \cdot \|x\| = \|x\|$ so that by [theorem: 14.174]

$$(\cdot x) \in L(\mathbb{K}, X)$$

Further

$$A_{(\cdot x),t} = \{\|(\cdot x)(r)\| | |r| = 1\} = \{\|x\|\}$$

so that

$$\|(\cdot x)\|_{L(\mathbb{K},X)} \underset{[\text{theorem: } 14.181]}{=} \sup(A_{(\cdot x),t}) = \sup(\{\|x\|\}) = \|x\|$$ □

Example 14.184. Let $n \in \mathbb{N}$, $\{\langle X_i, \|\|_i\rangle\}_{i \in \{1,\ldots,n\}}$ be a family of normed spaces and $\left\langle \prod_{i \in \{1,\ldots,n\}} X_i, \|\|_{\max}\right\rangle$ the product space with the maximum norm [see theorem: 14.94] then

$$\forall i \in \{1,\ldots,n\} \ \pi_i \in L\left(\prod_{j \in \{1,\ldots,n\}} X_j, X_i\right) \text{ and } \|\pi_i\|_{L(\prod_{j \in \{1,\ldots,n\}} X_j, X_i)} \leqslant 1$$

Proof. By [theorem: 11.166] we have

$$\forall i \in \{1,\ldots,n\} \ \pi_i: \prod_{j \in \{1,\ldots,n\}} X_j \to X_i \text{ is a linear mapping}$$

Further if $x \in \prod_{j \in \{1,\ldots,n\}} X_i$ then

$$\|\pi_i(x)\|_i \leqslant \max(\{\|\pi_i(x)\| | i \in \{1,\ldots,n\}\}) = \|x\|_{\max} \leqslant 1 \cdot \|x\|_{\max}$$

so that by [theorem: 14.174] π_i is continuous. Further $1 \in A_{L,r}$ so that $\|\pi_i\|_{L(\prod_{j \in \{1,\ldots,n\}} X_j, X_i)} \leqslant 1$. Hence we have

$$\forall i \in \{1,\ldots,n\} \ \pi_i \in L\left(\prod_{j \in \{1,\ldots,n\}} X_j, X_i\right) \text{ and } \|\pi_i\|_{L(\prod_{j \in \{1,\ldots,n\}} X_j, X_i)} \leqslant 1$$

□

Theorem 14.185. *Let $\langle X, \|\|_X \rangle$, $\langle Y, \|\|_Y \rangle$ and $\langle Z, \|\|_Z \rangle$ be normed spaces over $\mathbb{K}$, $L_1 \in L(X,Y)$, $L_2 \in L(Y,Z)$ then $L_2 \circ L_1 \in L(X,Z)$ and $\|L_2 \circ L_1\|_{L(X,Z)} \leqslant \|L_1\|_{L(X,Y)} \cdot \|L_2\|_{L(Y,Z)}$*

Proof. By [theorems: 11.170 and 14.138] it follows that

$$L_2 \circ L_1 \in L(X,Z)$$

Further $\forall x \in X$ we have

$$\|(L_2 \circ L_1)(x)\|_Z = \|L_2(L_1(x))\|_Z \leqslant \|L_2\|_{L(Y,Z)} \cdot \|L_1(x)\|_Y \leqslant \|L_1\|_{L(X,Y)} \cdot \|L_2\|_{L(Y,Z)} \cdot \|x\|_X$$

Hence $\|L_1\|_{L(X,Y)} \cdot \|L_2\|_{L(Y,Z)} \in A_{L_2 \circ L_1, r}$ so that

$$\|L_2 \circ L_1\|_{L(X,Z)} = \inf(A_{L_2 \circ L_1, r}) \leqslant \|L_1\|_{L(X,Y)} \cdot \|L_2\|_{L(Y,Z)}$$

□

For open linear mappings we have:

Theorem 14.186. *Let $\langle X, \|\|_X \rangle$, $\langle Y, \|\|_Y \rangle$ be normed spaces over $\mathbb{K}$ then for a linear mapping*

$$L: X \rightarrow Y$$

we have

$$L \text{ is open} \Leftrightarrow \exists \delta \in \mathbb{R}^+ \text{ such that } B_{\|\|_Y}(0,\delta) \subseteq L(B_{\|\|_X}(0,1))$$

Proof.

⇒. As $B_{\|\|_X}(0,1) \in \mathcal{T}_{\|\|_X}$ and L is open it follows that $L(B_{\|\|_X}(0,1)) \in \mathcal{T}_{\|\|_Y}$, further $0 \in B_{\|\|_X}(0,1) \Rightarrow 0 = L(0) \in L(B_{\|\|_X}(0,1))$. So by [theorem: 14.93] there exist a $\delta \in \mathbb{R}^+$ such that

$$0 \in B_{\|\|_Y}(0,\delta) \subseteq L(B_{\|\|_X}(0,1)).$$

⇐. Let $U \in \mathcal{T}_{\|\|_X}$ and let $y \in L(U)$ then there exist a $x \in U$ such that $y = L(x)$, by [theorem: 14.93] there exists a $\delta' \in \mathbb{R}^+$ such that

$$x \in B_{\|\|_X}(x,\delta') \subseteq U \Rightarrow L(x) \in L(B_{\|\|_X}(x,\delta')) \subseteq L(U).$$

By the hypothesis there exist a $\delta'' \in \mathbb{R}^+$ such that $0 \in B_{\|\|_Y}(0,\delta'') \subseteq L(B_{\|\|_X}(0,1))$. Then we have for $\delta = \delta'' \cdot \delta'$ that

$$\begin{aligned}
z \in B_{\|\|_Y}(y,\delta) &\Rightarrow z \in B_{\|\|_Y}(y,\delta'' \cdot \delta') \\
&\Rightarrow z \in B_{\|\|_Y}(L(x),\delta'' \cdot \delta') \\
&\Rightarrow \|z - L(x)\|_Y < \delta'' \cdot \delta' \\
&\Rightarrow \frac{1}{\delta'} \cdot \|z - L(x)\|_Y < \delta'' \\
&\Rightarrow \left\|\frac{1}{\delta'} \cdot (z - L(x))\right\|_Y < \delta'' \\
&\Rightarrow \frac{1}{\delta'} \cdot (z - L(x)) \in B_{\|\|_Y}(0,\delta'') \subseteq L(B_{\|\|_X}(0,1)) \\
&\Rightarrow \exists x_0 \in X \text{ such that } \|x_0\|_X < 1 \text{ and } L(x_0) = \frac{1}{\delta'} \cdot (z - L(x)) \\
&\Rightarrow z = \delta' \cdot L(x_0) + L(x) \wedge \|x_0\|_X < 1 \\
&\Rightarrow z = L(\delta' \cdot x_0 + x) \wedge \|x_0\|_X < 1 \\
&\Rightarrow z = L(\delta' \cdot x_0 + x) \wedge \|\delta' \cdot x_0\|_X < \delta' \\
&\Rightarrow z = L(\delta' \cdot x_0 + x) \wedge \|(\delta' \cdot x_0 + x) - x\|_X < \delta' \\
&\Rightarrow z = L(\delta' \cdot x + x) \wedge \delta' \cdot x_0 + x \in B_{\|\|_X}(x,\delta') \\
&\Rightarrow z \in L(B_{\|\|_X}(x,\delta')) \subseteq L(U)
\end{aligned}$$

proving that $B_{\|\|_Y}(y,\delta) \subseteq L(U)$. In other words

$$\forall y \in L(U) \text{ there exist a } \delta \in \mathbb{R}^+ \text{ such that } y \in B_{\|\|_Y}(y,\delta) \subseteq L(U)$$

So by [theorem: 14.93] it follows that $L(U) \in \mathcal{T}_{\|\|_Y}$ proving that L is open. □

14.6 Multilinear mappings and continuity

We prove now that the set of continuous multilinear mappings is a sub space of the set of multilinear mappings.

Theorem 14.187. *Let $n \in \mathbb{N}$, $\{\langle X_i, \|\|_{i \in \{1,\ldots,n\}}\rangle\}$ a family of normed vector spaces, $\langle Y, \|\|_Y\rangle$ a normed vector space over $\mathbb{K}$ and $L \in \operatorname{Hom}(X_1, \ldots, X_n; Y)$ then we have the following equivalences:*

1. *L is continuous [using the product topology on $\prod_{i \in \{1,\ldots,n\}} X_i$ [see theorem: 14.94] and the topology $\mathcal{T}_{\|\|_Y}$ on Y].*
2. *L is continuous at 0.*
3. *$\exists M \in \mathbb{R}_0^+$ such that $\forall x \in \prod_{i \in \{1,\ldots,n\}} X_i$ satisfying $\forall i \in \{1, \ldots, n\}$ $\|x_i\|_i = 1$ we have that $\|L(x)\|_Y \leqslant M$.*
4. *$\exists M \in \mathbb{R}_0^+$ such that $\forall x \in \prod_{i \in \{1,\ldots,n\}} X_i$ we have $\|L(x)\|_Y \leqslant M \cdot \prod_{i \in \{1,\ldots,n\}} \|x_i\|_i$.*

Proof. First by [theorem: 14.94] we have $\mathcal{T}_{\text{product}} = \mathcal{T}_{\|\|_{\max}}$ where

$$\|x\|_{\max} = \max(\{\|\pi_i(x)\|_i | i \in \{1, \ldots, n\}\}) = \max(\{\|x_i\|_i | i \in \{1, \ldots, n\}\})$$

Secondly we have:

1 ⇒ 2. This is trivial.

2 ⇒ 3. Using [theorem: 14.127] there exist a $\delta \in \mathbb{R}^+$ such that $\forall y = (y_1, \ldots, y_n) \in \prod_{i \in \{1,\ldots,n\}} X_i$ with $\|y\|_{\max} < \delta$ we have $\|L(y)\|_Y = \|L(y) - L(0)\|_Y < 1$. Let $x \in \prod_{i \in \{1,\ldots,n\}} X_i$ such that $\forall i \in \{1, \ldots, n\}$ $\|x_i\|_i = 1$ then $\|x\|_{\max} = \max(\{\|x_i\|_i | i \in \{1, \ldots, n\}\}) = 1$ so

$$\left\|\frac{\delta}{2} \cdot x\right\|_{\max} = \frac{\delta}{2} \cdot \|x\|_{\max} = \frac{\delta}{2} < \delta$$

and we must have that $\left\|L\left(\frac{\delta}{2} \cdot x\right)\right\|_Y < 1$. So if we take $M = \left(\frac{2}{\delta}\right)^n$ then

$$\begin{aligned}
\frac{1}{M} \cdot \|L(x)\|_Y &= \left(\frac{\delta}{2}\right)^n \cdot \|L(x)\|_Y \\
&= \left\|\left(\frac{\delta}{2}\right)^n \cdot L(x)\right\|_Y \\
&= \left\|\left(\frac{\delta}{2}\right)^n \cdot L(x_1, \ldots, x_n)\right\|_Y \\
&\underset{[\text{theorem: } 11.254]}{=} \left\|L\left(\frac{\delta}{2} \cdot x_1, \ldots, \frac{\delta}{2} \cdot x_n\right)\right\|_Y \\
&= \left\|L\left(\frac{\delta}{2} \cdot (x_1, \ldots, x_n)\right)\right\|_Y \\
&= \left\|L\left(\frac{\delta}{2} \cdot x\right)\right\|_Y \\
&< 1
\end{aligned}$$

proving that

$$\|L(x)\|_Y < M$$

3 ⇒ 4. Let $M \in \mathbb{R}_0^+$ be such that $\forall y \in \prod_{i \in \{1,\ldots,n\}} X_i$ with $\forall i \in \{1, \ldots, n\}$ $\|y_i\|_i = 1$ we have $\|L(y)\|_Y \leqslant M$. Then given $x \in \prod_{i \in \{1,\ldots,n\}} \|x_i\|_i$ we have two cases possible:

$\exists i \in \{1, \ldots, n\}$ with $x_i = 0$. Then by [theorem: 11.243] $L(x) = 0$ so that

$$\|L(x)\|_Y = \|0\|_Y = 0 \leqslant 0 = M \cdot 0 = M \cdot \prod_{i \in \{1,\ldots,n\}} \|x_i\|_i$$

$\forall i \in \{1, \ldots, n\}$ we have $x_i \neq 0$. Define then $y \in \prod_{i \in \{1, \ldots, n\}} X_i$ by $y_i = \frac{1}{\|x_i\|} \cdot x_i$ then

$$\|y_i\|_i = \left\| \frac{1}{\|x\|_i} \cdot x_i \right\|_i = \frac{1}{\|x_i\|_i} \cdot \|x_i\|_i = 1.$$

Further

$$\begin{array}{rcl} \frac{1}{\prod_{i \in \{1, \ldots, n\}} \|x_i\|_i} \cdot \|L(x_1, \ldots, x_n)\|_Y & = & \\ \left\| \frac{1}{\prod_{i \in \{1, \ldots, n\}} \|x_i\|_i} \cdot L(x_1, \ldots, x_n) \right\|_Y & \underset{[\text{theorem: } 11.254]}{=} & \\ \left\| L\left(\frac{1}{\|x\|_1} \cdot x_1, \ldots, \frac{1}{\|x_n\|_n} \cdot x_n \right) \right\|_Y & = & \\ \|L(y_1, \ldots, y_n)\|_Y & \leqslant & M \end{array}$$

so that

$$\|L(x_1, \ldots, x_n)\|_Y \leqslant M \cdot \prod_{i \in \{1, \ldots, n\}} \|x_i\|_i$$

hence in all cases we have

$$\|L(x_1, \ldots, x_n)\|_Y \leqslant M \cdot \prod_{i \in \{1, \ldots, n\}} \|x_i\|_i$$

$4 \Rightarrow 1$. By the hypothesis there exist a $M \in \mathbb{R}_0^+$ such that

$$\forall x \in \prod_{i \in \{1, \ldots, n\}} X_i \text{ we have } \|L(x)\|_Y \leqslant M \cdot \prod_{i \in \{1, \ldots, n\}} \|x_i\|_i$$

Let $x \in \prod_{i \in \{1, \ldots, n\}} X_i$ and $\varepsilon \in \mathbb{R}^+$ take then $\delta = \min\left(1, \frac{\varepsilon}{M+1}\right)$ then we have if $\|x - y\|_{\max} < \delta$ that $\forall i \in \{1,...,n\}$ $\|x_i - y_i\|_i < \delta$ so that

$$\begin{array}{rcl} \|L(x) - L(y)\|_Y & = & \|L(x - y)\|_Y \\ & \leqslant & M \cdot \prod_{i \in \{1, \ldots, n\}} \|x_i - y_i\|_i \\ & = & M \cdot \left(\prod_{i \in \{1, \ldots, n-1\}} \|x_i - y_i\|_i \right) \cdot \|x_n - y_i\|_n \\ & < & M \cdot 1 \cdot \|x_n - y_n\|_n \\ & < & M \cdot \frac{\varepsilon}{M+1} \\ & < & \varepsilon \end{array}$$

proving that L is continuous at x. As $x \in \prod_{i \in \{1, \ldots, n\}} X_i$ was chosen arbitrary it follows that L is continuous. □

Example 14.188. Let $n \in \mathbb{N}$, $\langle \mathbb{K}, || \rangle$ be the complex [real] vector space over $\mathbb{K}$ then we have:

1. $\forall x = (x_1, \ldots, x_n) \in \mathbb{K}^n$ we have $\left|\prod_{i \in \{1, \ldots, n\}} x_i\right| = \prod_{i \in \{1, \ldots, n\}} |x_i|$
2. $g: \mathbb{K}^n \rightarrow \mathbb{K}$ defined by $g(x_1, \ldots, x_n) = \prod_{i \in \{1, \ldots, n\}} x_i$ is multilinear and continuous [using the norms $||_n = \max(\{|x_i| | i \in \{1, \ldots, n\}\})$ and $||$.
3. Let $\langle X, \|\|_X \rangle$ be a normed space and $z \in X$ then

$$h_z: \mathbb{K}^n \rightarrow X \text{ defined by } h_z(x) = z \cdot g(x) = \left(\prod_{i \in \{1, \ldots, n\}} x_i \right) \cdot z$$

   is multilinear and continuous [using the norms $||_n = \max(\{|x_i| | i \in \{1, \ldots, n\}\})$ and $\|\|_X$.

Proof.

1. We prove this by induction, so let

$$S = \left\{ n \in \mathbb{N} | \text{If } (x_1, \ldots, x_n) \in \mathbb{K}^n \text{ then } \left| \prod_{i \in \{1, \ldots, n\}} x_i \right| = \prod_{i \in \{1, \ldots, n\}} |x_i| \right\}$$

then we have:

$1\in S$. If $(x_1,\ldots,x_1)\in\mathbb{K}^1$ then $\left|\prod_{i\in\{1\}} x_i\right|=|x_1|=\prod_{i\in\{1\}}|x_i|$ proving that $1\in S$.

$n\in S\Rightarrow n+1\in S$. Let $(x_1,...,x_{n+1})\in\mathbb{K}^{n+1}$ then $(x_1,...,x_n)\in\mathbb{K}^n$ so that, as $n\in S$, we have $\left|\prod_{i\in\{1,\ldots,n\}} x_i\right|=\prod_{i\in\{1,\ldots,n\}}|x_i|$. So

$$\begin{aligned}\left|\prod_{i\in\{1,\ldots,n+1\}} x_i\right| &= \left|\left(\prod_{i\in\{1,\ldots,n\}} x_i\right)\cdot x_{n+1}\right|\\ &= \left|\prod_{i\in\{1,\ldots,n\}} x_i\right|\cdot|x_{n+1}|\\ &= \left(\prod_{i\in\{1,\ldots,n\}}|x_i|\right)\cdot|x_{n+1}|\\ &= \prod_{i\in\{1,\ldots,n+1\}}|x_i|\end{aligned}$$

proving that $n+1\in S$.

2. First using [theorem: 11.247] it follows that

$$g:\mathbb{K}^n\rightarrow\mathbb{K} \text{ is multi linear}$$

Secondly let $(x_1,\ldots,x_n)\in\mathbb{K}^n$ then we have

$$|g(x_1,\ldots,x_n)|=\left|\prod_{i\in\{1,\ldots,n\}} x_i\right| \underset{(1)}{=} \prod_{i\in\{1,\ldots,n\}}|x_i|=1\cdot\prod_{i\in\{1,\ldots,n\}}|x_i|$$

which by [theorem: 14.187] proves that

$$g:\mathbb{K}^n\rightarrow\mathbb{K} \text{ is continuous}$$

3. Let $\alpha,x,y\in\mathbb{K}$, $i\in\{1,\ldots,n\}$ then for $\left(x_1,\ldots\underbrace{x+\beta\cdot y}_{i}\ldots,x_n\right)\in\mathbb{K}^n$ we have:

$$\begin{aligned}h_z\left(x_1,\ldots\underbrace{x+\beta\cdot y}_{i}\ldots,x_n\right) &=\\ g\left(x_1,\ldots\underbrace{x+\beta\cdot y}_{i}\ldots,x_n\right)\cdot z &\underset{(2)}{=}\\ \left(g\left(x_1,\ldots\underbrace{x}_{i}\ldots,x_n\right)+\beta\cdot g\left(x_1,\ldots\underbrace{y}_{i}\ldots,x_n\right)\right)\cdot z &=\\ h_z\left(x_1,\ldots\underbrace{x}_{i}\ldots,x_n\right)+\beta\cdot h_z\left(x_1,\ldots\underbrace{y}_{i}\ldots,x_n\right)\end{aligned}$$

proving that

$$h_z:\mathbb{K}^n\rightarrow X \text{ is multilinear}$$

Further we have

$$\begin{aligned}\|h_z(x_1,\ldots,x_n)\|_X &= \|g(x_1,\ldots,x_n)\cdot z\|\\ &= |g(x_1,\ldots,x_n)|\cdot\|z\|_X\\ &\leqslant_{(2)} \|z\|_X\cdot\prod_{i\in\{1,\ldots,n\}}|x_i|\end{aligned}$$

proving that

$$h_z:\mathbb{K}^n\rightarrow X \text{ is continuous}$$

□

Definition 14.189. *Let $n\in\mathbb{N}$, $\{\langle X_i,\|\|_i\rangle\}_{i\in\{1,\ldots,n\}}$ be a family of normed spaces and $\langle Y,\|\|_Y\rangle$ a normed space then theset of the graphs of multilinear continuous mappins between $\prod_{i\in\{1,\ldots,n\}} X_i$ and Y is noted as $L(X_1\ldots X_n;Y)$ hence*

$$L(X_1\ldots X_n;Y)=\{L\in\mathrm{Hom}(X_1\ldots X_n;Y)|L \textit{ is continuous}\}$$

Theorem 14.190. *Let $n\in\mathbb{N}$, $\{\langle X_i,\|\|_i\rangle\}_{i\in\{1,\ldots,n\}}$ be a family of normed spaces and $\langle Y,\|\|_Y\rangle$ a normed space then $L(X_1,\ldots,X_n;Y)$ is a subspace of $\operatorname{Hom}(X_1\ldots X_n;Y)$ [and thus by [theorem: 11.57] a vector space].*

Proof. Let $C_0\colon \prod_{i\in\{1,\ldots,n\}} X_i\rightarrow Y$ be defined by $C_0(x)=0$. By [theorem: 11.252]

$$C_0\in\operatorname{Hom}(X_1\ldots X_n;Y) \tag{14.67}$$

Further as $\forall x\in\prod_{i\in\{1,\ldots,n\}} X_i$ we have that $\|C_0(x)\|_Y=\|0\|_Y=0\leqslant 1\cdot\prod_{i\in\{1,\ldots,n\}}\|x_i\|_i$ it follows by [theorem: 14.187] that

$$C_0 \text{ is continuous} \tag{14.68}$$

Hence by [eqs: 14.67, 14.68] $C_0\in L(X_1,\ldots,X_n;Y)$ proving that

$$L(X_1\ldots X_n;Y)\neq\varnothing$$

Further if $\alpha\in\mathbb{K}$ and L_1, $L_2\in L(X_1\ldots X_n;Y)$ then by [theorem: 11.252]

$$L_1+\alpha\cdot L_2\in\operatorname{Hom}(X_1\ldots X_n;Y) \tag{14.69}$$

Also using [theorem: 14.187] we have that there exists M_1, $M_2\in\mathbb{R}^+$ such that $\forall x\in\prod_{i\in\{1,\ldots,n\}} X_i$ $\|L_1(x)\|_Y\leqslant M_1\cdot\prod_{i\in\{1,\ldots,n\}}\|x_i\|_i$ and $\|L_2(x)\|_Y\leqslant M_2\cdot\prod_{i\in\{1,\ldots,n\}}\|x_i\|_i$. Hence

$$\begin{aligned}
\|(L_1+\alpha\cdot L_2)(x)\|_Y &= \|L_1(x)+\alpha\cdot L_2(x)\|_Y\\
&\leqslant \|L_1(x)\|_Y+\|\alpha\cdot L_2(x)\|_Y\\
&= \|L_1(x)\|_Y+|\alpha|\cdot\|L_2(x)\|_Y\\
&\leqslant M_1\cdot\prod_{i\in\{1,\ldots,n\}}\|x_i\|_i+|\alpha|\cdot M_2\cdot\prod_{i\in\{1,\ldots,n\}}\|x_i\|_i\\
&= (M_1+|\alpha|\cdot M_2)\cdot\prod_{i\in\{1,\ldots,n\}}\|x_i\|_i
\end{aligned}$$

proving by [theorem: 14.187] that

$$L_1+\alpha\cdot L_2 \text{ is continuous} \tag{14.70}$$

Combining [eqs: 14.69, 14.70] proves

$$L_1+\alpha\cdot L_2\in L(X_1,\ldots,X_n;Y)$$ □

Next we proceed to create a norm for the vector space $L(X_1,\ldots,X_n;Y)$.

Theorem 14.191. *Let $n\in\mathbb{N}$, $\{\langle X_i,\|\|_i\rangle\}_{i\in\{1,\ldots,n\}}$ be a finite family of normed vector spaces, $\langle Y,\|\|_Y\rangle$ a normed spaces and $L\in L(X_1\ldots X_n)$ and define $M_{L,s}$ and $M_{L,r}$ by*

$$M_{L,s}=\left\{M\in\mathbb{R}_0^+|\forall x\in\prod_{i\in\{1,\ldots,n\}} X_i\vdash\forall i\in\{1,\ldots,n\}\text{ we have }\|x_i\|=1\text{ then we have }\|L(x)\|_Y\leqslant M\right\}$$

and

$$M_{L,r}=\left\{M\in\mathbb{R}_0^+|\forall x\in\prod_{i\in\{1,\ldots,n\}} X_i\text{ we have }\|L(x)\|_Y\leqslant M\cdot\prod_{i\in\{1,\ldots,n\}}\|x_i\|_i\right\}$$

then we have:

1. *$M_{L,s}=M_{L,r}$*
2. *$\inf(M_{L,s})$ exists [hence by (1) $\inf(M_{L,r})$ exist and $\inf(M_{L,s})=\inf(M_{L,r})$*

Proof.

1. First we prove that $M_{L,s}=M_{L,r}$. Let $M\in M_{L,s}$ then if $x=(x_1,\ldots,x_n)\in\prod_{i\in\{1,\ldots,n\}} X_i$ we have to look at the following possibilities for x:

   $\exists i\in\{1,\ldots,n\}$ with $\|x_i\|_i=0$. Then as $\|\|_i$ is a norm it follows that $x_i=0$. So by [theorem: 11.243] we have $L(x)=0$, hence

   $$\|L(x)\|_Y=\|0\|_Y=0=M\cdot 0\underset{[\text{theorem: 11.51}]}{=}M\cdot\prod_{i\in\{1,\ldots,n\}}\|x_i\|_i$$

$\forall i\in\{1,\ldots,n\}$ we have $\|x_i\|_i\neq 0$. Then we can define $\{y_i\}_{i\in I}$ by $y_i=\frac{1}{\|x_i\|}\cdot x_i$ so that $\|y_i\|=1$. Hence we have as $M\in M_{L,s}$

$$\begin{array}{rcl}\frac{1}{\prod_{i\in\{1,\ldots,n\}}\|x_i\|_i}\cdot\|L(x)\|_Y & = & \frac{1}{\prod_{i\in\{1,\ldots,n\}}\|x_i\|_i}\cdot\|L(x_1,\ldots,x_n)\|_Y\\ & = & \left\|\frac{1}{\prod_{i\in\{1,\ldots,n\}}\|x_i\|_i}\cdot L(x_1,\ldots,x_n)\right\|_Y\\ & \underset{[\text{theorem: } 11.254]}{=} & \left\|L\left(\frac{1}{\|x_1\|_1}\cdot x_1,\ldots,\frac{1}{\|x_n\|_n}\cdot x_n\right)\right\|_Y\\ & = & \|L(y_1,\ldots,y_n)\|_Y\\ & \leqslant & M\end{array}$$

hence we have

$$\|L(x)\|_Y\leqslant M\cdot\prod_{i\in\{1,\ldots,n\}}\|x_i\|_i$$

So in all cases we have

$$\|L(x)\|_Y\leqslant M\cdot\prod_{i\in\{1,\ldots,n\}}\|x_i\|_i$$

proving that $M\in M_{L,r}$ or that

$$M_{L,s}\subseteq M_{L,r} \tag{14.71}$$

For the opposite inclusion let $M\in M_{L,r}$ then for $x\in\prod_{i\in\{1,\ldots,n\}}X_i$ satisfying $\forall i\in\{1,\ldots,n\}$ $\|x_i\|_i=1$ we have

$$\|L(x)\|_Y\leqslant M\cdot\prod_{i\in\{1,\ldots,n\}}\|x_i\|_i=M\cdot 1=M$$

which proves that $M\in M_{L,s}$, hence $M_{L,r}\subseteq M_{L,s}$ which combined with [eq: 14.71] results in

$$M_{L,s}=M_{L.r}$$

2. Using [theorem: 14.187] there exists a $M\in\mathbb{R}^+$ such that $\forall x\in\prod_{i\in\{1,\ldots,n\}}X_i$ we have $\|L(x)\|_Y\leqslant M\cdot\prod_{i\in\{1,\ldots,n\}}\|x_i\|_i$, so $M\in M_{L,r}$ proving that $M_{L,r}\neq 0$. As $\forall M\in M_{L,r}$ $0\leqslant M$ and the real numbers are conditional complete [see theorem: 10.18] it follows that

$$\inf(M_{L,s})\text{ exists and }0\leqslant\inf(M_{L,s})$$

□

The above theorem ensures that the following definition makes sense.

Definition 14.192. *Let $n\in\mathbb{N}$, $\{\langle X_i,\|\|_i\rangle\}_{i\in\{1,\ldots,n\}}$ be a finite family of normed vector spaces over the field $\mathbb{K}$, $\langle Y,\|\|_Y\rangle$ is a normed spaces over $\mathbb{K}$ and $L\in L(X_1\ldots X_n;Y)$ then the operator norm of L noted as $\|L\|_{L(X,\ldots X_n;Y)}$ is defined as*

$$\|L\|_{L(X_1\ldots X_n;Y)}=\inf(M_{L,r})\underset{[\text{theorem: } 14.191]}{=}\inf(M_{L,s})$$

It turns out that $\|L\|_{L(X_1,\ldots,X_n;Y)}$ is a norm on $L(X_1,\ldots,X_n;Y)$.

Theorem 14.193. *Let $n\in\mathbb{N}$, $\{\langle X_i,\|\|_i\rangle\}_{i\in\{1,\ldots,n\}}$ be a finite family of normed vector spaces over the field $\mathbb{K}$ and $\langle Y,\|\|_Y\rangle$ a normed spaces over $\mathbb{K}$ then we have*

1. *If $L\in L(X_1\ldots X_n;Y)$ then $\forall x\in\prod_{i\in\{1,\ldots,n\}}X_i$ we have*

$$\|L(x)\|_Y\leqslant\|L\|_{L(X_1\ldots X_n;Y)}\cdot\prod_{i\in\{1,\ldots,n\}}\|x_i\|_i$$

2. *If $L\in L(X_1\ldots X_n;Y)$ then $\forall M\in\mathbb{R}^+$ such that $\forall x\in\prod_{i\in\{1,\ldots,n\}}X_i$*

$$\|L(x)\|_Y\leqslant M\cdot\prod_{i\in\{1,\ldots,n\}}\|x_i\|_i$$

we have

$$\|L\|_{L(X_1\ldots X_n;Y)}\leqslant M$$

3. *If* $L \in L(X_1 \ldots X_n; Y)$ *then* $\forall M \in \mathbb{R}^+$ *such that* $\|L(x)\|_Y \leqslant M$ $\forall x \in \prod_{i \in \{1, \ldots, n\}} X_i$ *with* $\forall i \in \{1, \ldots, n\}$ $\|x_i\|_i = 1$ *we have*

$$\|L\|_{L(X_1 \ldots X_n; Y)} \leqslant M$$

4. $\|\|_{L(X_1, \ldots, X_n; Y)}: L(X_1 \ldots X_n; Y) \rightarrow \mathbb{R}$ *defined by* $\|\|_{L(X_1 \ldots X_n; Y)}(L) = \|L\|_{L(X_1 \ldots X_n; Y)}$ *is a norm on* $L(X_1 \ldots X_n; Y)$*. In other words* $\langle L(X_1 \ldots X_n; Y), \|\|_{L(X_1 \ldots X_n; Y)} \rangle$ *is a normed space.*

Proof.

1. Let $x \in \prod_{i \in \{1, \ldots, n\}} X_i$ then we have the following possibilities for x

   $\exists i \in \{1, \ldots, n\}$ with $\|x_i\|_i = 0$. Then by [theorem: 11.243] $L(x) = 0$ so that

$$\begin{array}{rcl} \|L(x)\|_Y & = & \|0\|_Y \\ & = & 0 \\ & = & \|L\|_{L(X_1 \ldots X_n; Y)} \cdot 0 \\ & \underset{\text{[theorem: 11.51]}}{=} & \|L\|_{L(X_1 \ldots X_n; Y)} \cdot \prod_{i \in \{1, \ldots, n\}} \|x_i\|_i \end{array}$$

   $\forall i \in \{1, \ldots, n\}$ we have $\|x_i\|_i \neq 0$. Then by [theorem: 11.50] $0 < \prod_{i \in \{1, \ldots, n\}} \|x_i\|_i$. Assume that $\|L\|_{L(X_1, \ldots, X_n; Y)} \cdot \prod_{i \in \{1, \ldots, n\}} \|x_i\|_i < \|L(x)\|_Y$ then

$$\|L\|_{L(X_1, \ldots, X_n; Y)} < \frac{1}{\prod_{i \in \{1, \ldots, n\}} \|x_i\|_i} \cdot \|L(x)\|_Y$$

   As $\|L\|_{L(X_1, \ldots, X_n; Y)} = \inf(M_{L,r})$ there exist a $M \in M_{L,r}$ such that

$$\|L\|_{L(X_1 \ldots X_n; Y)} \leqslant M < \frac{1}{\prod_{i \in \{1, \ldots, n\}} \|x_i\|_i} \cdot \|L(x)\|_Y$$

   hence $M \cdot \prod_{i \in \{1, \ldots, n\}} \|x_i\|_i < \|L(x)\|_Y$ contradicting $M \in M_{L,r}$. Hence we must have that

$$\|L(x)\|_Y \leqslant \|L\|_{L(X_1 \ldots X_n; Y)} \cdot \prod_{i \in \{1, \ldots, n\}} \|x_i\|_i$$

   So in all cases we have

$$\|L(x)\|_Y \leqslant \|L\|_{L(X_1 \ldots X_n; Y)} \cdot \prod_{i \in \{1, \ldots, n\}} \|x_i\|_i$$

2. This follows from

$$\begin{array}{rl} \|L\|_{L(X_1 \ldots X_n; Y)} & = \\ \inf(M_{L,r}) & = \\ \inf\left(\left\{M \in \mathbb{R}_+ | \forall x \in \prod_{i \in \{1, \ldots, n\}} X_i \text{ we have } \|L(x)\|_Y \leqslant M \cdot \prod_{i \in \{1, \ldots, n\}} \|x_i\|_i\right\}\right) & = \end{array}$$

3. This follows from

$$\begin{array}{rl} \|L\|_{L(X_1 \ldots X_n; Y)} & = \\ \inf(M_{L,s}) & = \\ \inf\left(M \in \mathbb{R}_+ | \forall x \in \prod_{i \in \{1, \ldots, n\}} X_i \vdash \forall i \in \{1, \ldots, n\} \|x_i\| = 1 \text{ we have } \|L(x)\|_Y \leqslant M\right) & \end{array}$$

4. Let $L_1, L_2 \in L(X_1 \ldots X_n; Y)$ then by [theorem: 14.190] we have $L_1 + L_2 \in L(X_1 \ldots X_n; Y)$. Further we have for $x \in \prod_{i \in \{1, \ldots, n\}} X_i$ with $\forall i \in \{1, \ldots, n\}$ $\|x_i\|_i = 1$ that

$$\begin{array}{rl} \|(L_1 + L_2)(x)\|_Y & = \\ \|L_1(x) + L_2(x)\|_Y & \leqslant \\ \|L_1(x)\|_Y + \|L_2(x)\|_Y & \leqslant_{(1)} \\ \|L_1\|_{L(X_1, \ldots, X_n; Y)} \cdot \prod_{i \in \{1, \ldots, n\}} \|x_i\|_i + \|L_2\|_{L(X_1, \ldots, X_n; Y)} \cdot \prod_{i \in \{1, \ldots, n\}} \|x_i\|_i & = \\ (\|L_1\|_{L(X_1, \ldots, X_n; Y)} + \|L_2\|_{L(X_1, \ldots, X_n; Y)}) \cdot \prod_{i \in \{1, \ldots, n\}} \|x_i\|_i & \end{array}$$

which by (2) proves that

$$\|L_1+L_2\|_{|L(X_1\ldots X_n;Y)} \leqslant \|L_1\|_{L(X_1\ldots X_n;Y)}+\|L_2\|_{L(X_1\ldots X_n;Y)} \tag{14.72}$$

For C_0 we have by [theorem: 14.190] that $C_0 \in L(X_1 \ldots X_n; Y)$. Further for $x \in \prod_{i\in\{1,\ldots,n\}} X_i$ we have

$$\|C_0(x)\|_Y = \|0\|_Y = 0 = 0 \cdot \prod_{i\in\{1,\ldots,n\}} \|x_i\|_i$$

so that by (1)

$$\|C_0\|_{L(X_1\ldots X_n;Y)} = 0 \tag{14.73}$$

Assume that for $L \in L(X_1 \ldots X_n; Y)$ we have that $\|L\|_{L(X_1\ldots X_n;Y)} = 0$ then by (1) we have $\forall x \in \prod_{i\in\{1,\ldots,n\}} X_i$ that $0 \leqslant \|L(x)\|_Y \leqslant \|L\|_{L(X_1,\ldots,X_n;Y)} \cdot \prod_{i\in\{1,\ldots,n\}} \|x_i\|_i = 0$. Hence $\|L\|_Y = 0$ which as $\|\|_Y$ is a norm proves that $L = C_0$ [the neutral element of $L(X_1 \ldots X_n; Y)$. So we have

$$L = C_0 \Leftrightarrow \|L\|_{L(X_1\ldots X_n;Y)} = 0 \tag{14.74}$$

If $L \in L(X_1 \ldots X_n; Y)$ and $\alpha \in \mathbb{K}$ then by [theorem: 14.190] $\alpha \cdot L \in L(X_1 \ldots X_n; Y)$. Let $x \in \prod_{i\in\{1,\ldots,n\}} X_i$ then we have

$$\|(\alpha \cdot L)(x)\|_Y = \|\alpha \cdot L(x)\|_Y = |\alpha| \cdot \|L(x)\|_Y \leqslant |\alpha| \cdot \|L\|_{L(X_1\ldots X_n;Y)} \prod_{i\in\{1,\ldots n\}} \|x_i\|_i$$

which proves by (1) that

$$\|\alpha \cdot L\|_{L(X_1\ldots X_n;Y)} \leqslant |\alpha| \cdot \|L\|_{L(X_1\ldots X_n;Y)} \tag{14.75}$$

Now for the opposite inequality for $\alpha \in \mathbb{K}$ we have either:

$\boldsymbol{\alpha = 0}$. Then

$$\begin{array}{rcl} |\alpha| \cdot \|L\|_{L(X_1,\ldots X_n;Y)} & = & 0 \\ & \underset{[\text{eq: } 14.73]}{=} & \|C_0\|_{L(X_1\ldots X_n;Y)} \\ & = & \|(0 \cdot L)\|_{L(X_1\ldots X_n;Y)} \\ & = & \|(\alpha \cdot L)\|_{L(X_1,\ldots X_n;Y)} \end{array}$$

$\boldsymbol{\alpha \neq 0}$. Assume that $\|\alpha \cdot L\|_{L(X_1\ldots X_n;Y)} < |\alpha| \cdot \|L\|_{L(X_1\ldots X_n;Y)}$. Then as we have $\|\alpha \cdot L\|_{L(X_1\ldots X_n;Y)} = \inf(M_{\alpha\cdot L,r})$ it follows that there exist a $M \in M_{\alpha \cdot L,r}$ such that

$$\|\alpha \cdot L\|_{L(X_1\ldots X_n;Y)} \leqslant M < |\alpha| \cdot \|L\|_{L(X_1\ldots X_n;Y)} \tag{14.76}$$

If $x \in \prod_{i\in\{1,\ldots,n\}} X_i$ we have, as $M \in M_{\alpha\cdot L,r}$ that

$$|\alpha| \cdot \|L(x)\|_Y = \|\alpha \cdot L(x)\| = \|(\alpha \cdot L)(x)\|_Y \leqslant M \cdot \prod_{i\in\{1,\ldots,n\}} \|x_i\|_i$$

so that

$$\|L(x)\|_Y \leqslant \frac{M}{|\alpha|} \cdot \prod_{i\in\{1,\ldots,n\}} \|x_i\|_i$$

proving that $\frac{M}{|\alpha|} \in M_{L,r}$ hence

$$\|L\|_{L(X_1\ldots X_n;Y)} = \inf(M_{L,r}) \leqslant \frac{M}{|a|} <_{[\text{eq: } 14.76]} \frac{|\alpha|}{|\alpha|} \cdot \|L\|_{L(X_1\ldots X_n;Y)}$$

leading to the contradiction $\|L\|_{L(X_1\ldots X_n;Y)} > \|L\|_{L(X_1\ldots X_n;Y)}$. Hence we must have that $\|\alpha \cdot L\|_{L(X_1\ldots X_n;Y)} \leqslant |\alpha| \cdot \|L\|_{L(X_1\ldots X_n;Y)}$.

So in both cases we have $\|\alpha \cdot L\|_{L(X_1\ldots X_n;Y)} \leqslant |\alpha| \cdot \|L\|_{L(X_1\ldots X_n;Y)}$ which combined with [eq: 14.75] proves that

$$\|\alpha \cdot L\|_{L(X_1\ldots X_n;Y)} = |\alpha| \cdot \|L\|_{L(X_1\ldots X_n;Y)} \tag{14.77}$$

Finally using [eqs: 14.72, 14.74 and 14.77] it follows that

$$\|\|_{L(X_1\ldots X_n;Y)} \text{ is a norm in } L(X_1 \ldots X_n; Y)$$ □

If all $\langle X_i, \|\|_i\rangle$ are the same normed space then we use the following definition.

Definition 14.194. *Let $n \in \mathbb{N}$ and $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$ be normed spaces over $\mathbb{K}$ then we define*

$$L^n(X;Y) = \begin{cases} L(X,Y) \text{ if } n=1 \\ L\Big(\underbrace{X\ldots X}_{n};Y\Big) \text{ if } 1<n \end{cases}$$

and for $L \in L^n(X;Y)$

$$\|L\|_{L^n(X;Y)} = \begin{cases} \|L\|_{L(X,Y)} \text{ if } n=1 \\ \|L\|_{L\left(\underbrace{X\ldots X}_{n};Y\right)} \text{ if } 1<n \end{cases}$$

Theorem 14.195. *Let $n \in \mathbb{N}$, $\langle X, \|\|\rangle$ a normed space and $L \in L^n(\mathbb{K}^n;X)$ then*

$$\|L\|_{L^n(\mathbb{K}^n;X)} \leqslant \left\|L\Big(\underbrace{1,\ldots,1}_{n}\Big)\right\|$$

Proof. If $x \in \mathbb{K}^n$ then

$$\begin{array}{rcl} \|L(x_1,\ldots,x_n)\| & = & \|L(x_1\cdot 1,\ldots,x_n\cdot 1)\| \\ & \underset{[\text{theorem: } 11.254]}{=} & \left\|\left(\prod_{i=1}^{n} x_i\right)\cdot L\Big(\underbrace{1,\ldots,1}_{n}\Big)\right\| \\ & = & \left\|L\Big(\underbrace{1,\ldots,1}_{n}\Big)\right\|\cdot\left|\prod_{i=1}^{n} x_i\right| \\ & = & \left\|L\Big(\underbrace{1,\ldots,1}_{n}\Big)\right\|\cdot\prod_{i=1}^{n}|x_i| \end{array}$$

so that $\left\|L\Big(\underbrace{1,\ldots,1}_{n}\Big)\right\| \in M_{L,r}$. Hence we have

$$\|L\|_{L^n(\mathbb{K}^n;X)} = \inf(M_{L,r}) \leqslant \left\|L\Big(\underbrace{1,\ldots,1}_{n}\Big)\right\|$$

□

Theorem 14.196. *Let $n \in \mathbb{N}$, $\{\langle X_i, \|\|_i\rangle\}_{i\in\{1,\ldots,n\}}$ be a finite family of normed vector spaces, $\langle Y, \|\|_Y\rangle$, $\langle Z, \|\|_Z\rangle$ normed spaces, $L \in L(X_1\ldots X_n;Y)$ and $K \in L(Y,Z)$ then*

1. *$K\circ L \in L(X_1\ldots X_n;Y)$*
2. *$\|K\circ L\|_{L(X_1\ldots X_n;Z)} \leqslant \|L\|_{L(X_1\ldots X_n;Y)}\cdot\|K\|_{L(Y,Z)}$*

Proof.

1. From [theorem: 11.253] it follows that $K\circ L \in \operatorname{Hom}(X_1\ldots X_n;Y)$, Further as the composition of continuous mappings is continuous [see theorem: 14.138] we have that $K\circ L$ is continuous. Hence

$$K\circ L \in L(X_1\ldots X_n;Y)$$

2. Let $x \in \prod_{i\in\{1,\ldots,n\}} X_i$ then we have

$$\begin{array}{rcl} \|(K\circ L)(x)\|_Z & = & \|K(L(x))\|_Z \\ & \leqslant_{[\text{theorem: } 14.179]} & \|K\|_{L(Y,Z)}\cdot\|L(x)\|_Y \\ & \leqslant_{[\text{theorem: } 14.193]} & \|K\|_{L(Y,Z)}\cdot\|L\|_{L(X_1,\ldots,X_n;Y)}\cdot\prod_{i\in\{1,\ldots,n\}}\|x_i\|_i \end{array}$$

proving by [theorem: 14.193] that

$$\|K\circ L\|_{L(X_1\ldots X_n;Y)} \leqslant \|L\|_{L(X_1\ldots X_n;Y)}\cdot\|K\|_{L(Y,Z)}$$

□

We show now that the composition of linear mappings is a multilinear mapping itself.

Theorem 14.197. *Let $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$ and $\langle Z,\|\|_Z\rangle$ be normed spaces over $\mathbb{K}$ then if we define*

$$(\circ): L(Y,Z)\times L(X,Y)\rightarrow L(X,Z) \text{ defined by } (\circ)(L_2,L_1)=L_2\circ L_1$$

then we have that

1. $(\circ)\in L(L(Y,Z),L(X,Y);L(X,Z))$ *[in other words $(\circ)$ is multilinear and continuous]*
2. $\|(\circ)\|_{L(L(Y,Z),L(X,Y);L(X,Z))}\leqslant 1$

Proof. Let $\alpha\in\mathbb{K}$. Then we have for $L_1,L_2\in L(Y,Z)$ and $L\in L(X,Y)$ that $\forall x\in X$ we have

$$\begin{array}{rcl}(\circ)(L_1+\alpha\cdot L_2,L)(x) & = & ((L_1+\alpha\cdot L_2)\circ L)(x)\\ & = & (L_1+\alpha\cdot L_2)(L(x))\\ & = & L_1(L(x))+\alpha\cdot L_2(L(x))\\ & = & (L_1\circ L)(x)+\alpha\cdot(L_2\circ L)(x)\\ & = & ((L_1\circ L)+\alpha\cdot(L_2\circ L))(x)\\ & = & (\circ(L_1,L)+\alpha\cdot\circ(L_2,L))(x)\end{array}$$

proving that

$$\circ(L_1+\alpha\cdot L_2,L)=\circ(L_1,L)+\alpha\cdot\circ(L_2,L)$$

Further if $L\in L(Y,Z)$ and $L_1,L_2\in L(X,Y)$ then $\forall x\in X$ we have

$$\begin{array}{rcl}\circ(L,L_1+\alpha\cdot L_2)(x) & = & (L\circ(L_1+\alpha\cdot L_2))(x)\\ & = & L((L_1+\alpha\cdot L_2)(x))\\ & = & L(L_1(x)+\alpha\cdot L_2(x))\\ & = & L(L_1(x))+\alpha\cdot L(L_2(x))\\ & = & (L\circ L_1)(x)+\alpha\cdot(L\circ L_2)(x)\\ & = & ((L\circ L_1)+\alpha\cdot(L\circ L_2))(x)\\ & = & ((\circ)(L,L_1)+\alpha\cdot(\circ)(L,L_2))(x)\end{array}$$

proving that

$$(\circ)(L,L_1+\alpha\cdot L_2)=(\circ)(L,L_1)+\alpha\cdot(\circ)(L,L_2)$$

Hence $\circ$ is multilinear or

$$(\circ)\in\mathrm{Hom}(L(Y,Z),L(X,Y);L(X,Z))$$

For continuity note that for $L_1\in L(Y,Z)$, $L_2\in L(X,Y)$ and $x\in X$ we have

$$\begin{array}{rcl}\|(\circ)(L_1,L_2)(x)\|_Z & = & \|(L_1\circ L_2)(x)\|_Z\\ & = & \|L_1(L_2(x))\|_Z\\ & \leqslant_{[\text{theorem: }14.179]} & \|L_1\|_{L(Y,Z)}\cdot\|L_2(x)\|_Y\\ & \leqslant_{[\text{theorem: }14.179]} & \|L_1\|_{L(Y,Z)}\cdot\|L_2\|_{L(X,Y)}\cdot\|x\|_x\end{array}$$

which proves by [theorem: 14.193 that

$$\|(\circ)(L_1,L_2)\|_{L(X,Z)}\leqslant\|L_1\|_{L(Y,Z)}\cdot\|L_2\|_{L(X,Y)}=1\cdot\|L_1\|_{L(Y,Z)}\cdot\|L_2\|_{L(X,Y)}$$

so that by [theorem: 14.187] $(\circ)$ is continuous. Hence

$$(\circ)\in L(L(Y,Z),L(X,Y);L(X,Z))$$

and by [theorem 14.193] that

$$\|(\circ)\|_{L(L(Y,Z),L(X,Y);L(X,Z))}\leqslant 1.$$

□

Theorem 14.198. *Let $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$ and $\langle Z,\|\|_Z\rangle$ be normed spaces over $\mathbb{K}$ then we can construct a linear isometric isomorphism*

$$\varphi: L(X,Y;Z)\rightarrow L(X,L(Y,Z))$$

Likewise if we take $\psi=\varphi^{-1}$ *then*

$$\psi: L(X,L(Y,Z))\rightarrow L(X,Y;Z) \text{ is a linear isometric isomorphism}$$

In other words $L(X,L(Y,Z))$ *and* $L(X,Y;Z)$ *are isometric [see definition: 14.117] or*

$$L(X,L(Y,Z))\approx L(X,Y;Z)$$

Further we have $\forall (x,y)\in X\cdot Y$ *that*

$$\forall L\in L(X,Y;Z)\ \varphi(L)(x)(y)=L(X,Y)$$

and

$$\forall L\in L(X,(L(Y,Z)))\ \text{that}\ \psi(L)(x,y)=L(x)(y)$$

Proof. Let $L\in L(X,Y;Z)$ and given $x\in X$ define

$$L_x: Y\rightarrow Z \text{ by } L_x(y)=L(x,y) \tag{14.78}$$

If $r,s\in Y$ and $\alpha\in\mathbb{K}$ then

$$L_x(r+\alpha\cdot s)=L(x,r+\alpha\cdot s)\underset{\text{multilinearity}}{=}L(x,r)+\alpha\cdot L(x,s)=L_x(r)+\alpha\cdot L_x(s)$$

proving that

$$L_x\in \mathrm{Hom}(Y,Z) \tag{14.79}$$

Further for $r\in Y$ we have

$$\|L_x(r)\|_Z=\|L(x,r)\|_Z\leqslant_{[\text{theorem: 14.193}]}\|L\|_{L(X,Y;Z)}\cdot\|x\|_X\cdot\|r\|_Y$$

proving by [theorem: 14.174] and [definition: 14.178] that

$$\forall x\in X\ L_x\in L(Y,Z) \text{ and } \|L_x\|_{L(Y,Z)}\leqslant\|L\|_{L(X,Y;Z)}\cdot\|x\|_X \tag{14.80}$$

The above allows us, given $L\in L(X,Y;Z)$, define

$$\varphi_L: X\rightarrow L(Y,Z) \text{ by } \varphi_L(x)=L_x \tag{14.81}$$

Then if $x,y\in X$ and $\alpha\in\mathbb{K}$ we have for $r\in Y$ that

$$\begin{array}{rcl}(\varphi_L(x+\alpha\cdot y))(r) & = & L_{x+\alpha\cdot y}(r)\\ & \underset{[\text{eq: 14.78}]}{=} & L(x+\alpha\cdot y,r)\\ & \underset{\text{multilinearity}}{=} & L(x,r)+\alpha\cdot L(y,r)\\ & \underset{[\text{eq: 14.78}]}{=} & L_x(r)+\alpha\cdot L_y(r)\\ & = & (\varphi_L(x))(r)+\alpha\cdot(\varphi_L(y))(r)\\ & = & (\varphi_L(x)+\alpha\cdot\varphi_L(y))(r)\end{array}$$

proving that $\varphi_L(x+\alpha\cdot y)=\varphi_L(x)+\alpha\cdot\varphi_L(y)$. Hence

$$\varphi_L\in\mathrm{Hom}(X,L(Y,Z)) \tag{14.82}$$

For $x\in X$ we have

$$\|\varphi_L(x)\|_{L(Y,Z)}\underset{[\text{eq: 14.81}]}{=}\|L_x\|_{L(Y,Z)}\leqslant_{[\text{eq: 14.80}]}\|L\|_{L(X,Y;Z)}\cdot\|x\|_X$$

proving by [theorem: 14.174] and [definition: 14.178] that

$$\varphi_L\in L(X,L(Y,Z)) \text{ and } \|\varphi_L\|_{L(X,L(Y,Z))}\leqslant\|L\|_{L(X,Y;Z)} \tag{14.83}$$

The above let us define

$$\varphi: L(X,Y;Z)\rightarrow L(X,L(Y,Z)) \text{ by } \varphi(L)=\varphi_L \tag{14.84}$$

Let $L_1,L_2\in L(X,Y;Z)$ and $\alpha\in\mathbb{K}$ then for $x\in X$ we have $\forall r\in Y$ that

$$\begin{array}{rcl}(L_1+\alpha\cdot L_2)_x(r) & \underset{[\text{eq: 14.78}]}{=} & (L_1+\alpha\cdot L_2)(x,r)\\ & = & L_1(x,r)+\alpha\cdot L_2(x,r)\\ & = & (L_1)_x(r)+\alpha\cdot(L_2)_x(r)\\ & = & ((L_1)_x+\alpha\cdot(L_2)_x)(r)\end{array}$$

proving that $(L_1+\alpha\cdot L_2)_x=(L_1)_x+\alpha\cdot(L_2)_x$. Hence

$$\varphi_{L_1+\alpha\cdot L_2}(x)=(L_1+\alpha\cdot L_2)_x=(L_1)_x+\alpha\cdot(L_2)_x=\varphi_{L_1}(x)+\alpha\cdot\varphi_{L_2}(x)=(\varphi_{L_1}+\alpha\cdot\varphi_{L_2})(x)$$

proving that $\varphi_{L_1+\alpha\cdot L_2}=\varphi_{L_1}+\alpha\cdot\varphi_{L_2}$. So

$$\varphi(L_1+\alpha\cdot L_2)=\varphi_{L_1+\alpha\cdot L_2}=\varphi_{L_1}+\alpha\cdot\varphi_{L_2}=\varphi(L_1)+\alpha\cdot\varphi(L_2)$$

proving that

$$\varphi\in\mathrm{Hom}(L(X,Y;Z),L(X,L(Y,Z))) \tag{14.85}$$

Further giving $L\in L(X,Y;Z)$ we have

$$\|\varphi(L)\|_{L(X,L(Y,Z))}=\|\varphi_L\|_{L(X,L(Y,Z))}\leqslant_{[\text{eq: }14.83]}\|L\|_{L(X,Y;Z)}$$

giving by [theorem: 14.174]

$$\varphi\in L((L(X,Y;Z),L(X,L(Y,Z)))) \text{ and } \|\varphi(L)\|_{L(X,L(Y,Z))}\leqslant\|L\|_{L(X,Y;Z)} \tag{14.86}$$

Next we will define the inverse mapping of φ. Let $L\in L(X,L(Y,Z))$ and define

$$\psi_L: X\cdot Y\rightarrow Z \text{ by } \psi_L(x,y)=(L(x))(y) \tag{14.87}$$

Let $x,y\in X$, $r\in Y$ and $\alpha\in\mathbb{K}$ then we have

$$\begin{array}{rcl}\psi_L(x+\alpha\cdot y,r) & \underset{[\text{eq: }14.87]}{=} & (L(x+\alpha\cdot y))(r)\\ & \underset{L\in L(X,L(Y,Z))}{=} & (L(x)+\alpha\cdot L(y))(r)\\ & = & (L(x))(r)+\alpha\cdot(L(y))(r)\\ & \underset{[\text{eq: }14.87]}{=} & \psi_L(x,r)+\alpha\cdot\psi_L(y,r)\end{array}$$

and for $x\in X$, $r,s\in Y$ and $\alpha\in\mathbb{K}$ we have

$$\begin{array}{rcl}\psi_L(x,r+\alpha\cdot s) & \underset{[\text{eq: }14.87]}{=} & (L(x))(r+\alpha\cdot s)\\ & \underset{L(x)\in L(Y,Z)}{=} & (L(x))(r)+\alpha\cdot(L(x))(s)\\ & \underset{[\text{eq: }14.87]}{=} & \psi_L(x,r)+\alpha\cdot\psi_L(x,s)\end{array}$$

proving that

$$\psi_L\in\mathrm{Hom}(X,Y;Z)$$

If $(x,y)\in X\cdot Y$ then

$$\|\psi_L(x,y)\|_Z=\|(L(x))(y)\|_Z\leqslant_{L(x)\in L(Y,Z)}\|L(x)\|_{L(Y,Z)}\cdot\|y\|_Y\leqslant\|L\|_{L(X,L(Y,Z))}\cdot\|x\|_X\cdot\|y\|_Y$$

proving by [theorems: 14.187, 14.193] that

$$\psi_L\in L(X,Y;Z) \text{ and } \|\psi_L\|_{L(X,Y;Z)}\leqslant\|L\|_{L(X,L(Y,Z))} \tag{14.88}$$

The above allows use to define

$$\psi: L(X,L(Y,Z))\rightarrow L(X,Y;Z) \text{ by } \psi(L)=\psi_L \tag{14.89}$$

where by [eq: 14.88]

$$\forall L\in L(X,L(Y,Z))\ \|\psi(L)\|_{L(X,Y;Z)}=\|\psi_L\|_{L(X,Y;Z)}\leqslant\|L\|_{L(X,L(Y,Z))} \tag{14.90}$$

Let $L\in L(X,L(Y,Z))$ and $x\in X$, $y\in Y$ then we have

$$\begin{array}{rcl}(((\varphi\circ\psi)(L))(x))(y) & = & ((\varphi(\psi(L)))(x))(y)\\ & \underset{[\text{eq: }14.84]}{=} & ((\varphi_{\psi(L)})(x))(y)\\ & \underset{[\text{eq: }14.81]}{=} & (\psi(L))_x(y)\\ & \underset{[\text{eq: }14.78]}{=} & \psi(L)(x,y)\\ & \underset{[\text{eq: }14.89]}{=} & \psi_L(x,y)\\ & \underset{[\text{eq: }14.87]}{=} & (L(x))(y)\end{array}$$

so that $((\varphi\circ\psi)(L))(x)=L(x)$ from which it follows that $(\varphi\circ\psi)(L)=L$. Hence we have that

$$\varphi\circ\psi=\mathrm{Id}_{L(X,L(Y,Z))} \tag{14.91}$$

On the other hand if $L\in L(X,Y;Z)$ and $(x,y)\in X\cdot Y$ then we have

$$\begin{array}{rcl}((\psi\circ\varphi)(L))(x,y) & = & (\psi(\varphi(L)))(x,y)\\ & \underset{[\text{eq: } 14.89]}{=} & \psi_{\varphi(L)}(x,y)\\ & \underset{[\text{eq: } 14.87]}{=} & ((\varphi(L))(x))(y)\\ & \underset{[\text{eq: } 14.84]}{=} & ((\varphi_L)(x))(y)\\ & \underset{[\text{eq: } 14.81]}{=} & L_x(y)\\ & \underset{[\text{eq: } 14.78]}{=} & L(x,y)\end{array}$$

proving that $(\psi\circ\varphi)(L)=L$, hence

$$\psi\circ\varphi=\mathrm{Id}_{L(X,Y;Z)} \tag{14.92}$$

From [eqs: 14.91, 14.92] it follows by [theorem: 2.71] that $\varphi: L(X,Y;Z)\rightarrow L(X,L(Y,Z))$ is a bijection. Combining this with]eq: 14.85] proves that

$$\varphi \text{ is a linear isomorphism with } \psi=\varphi^{-1} \tag{14.93}$$

Finally we have for $L\in L(X,Y;Z)$ that

$$\begin{array}{rcl}\|L\|_{L(X,Y;Z)} & = & \|\mathrm{Id}_{L(X,Y;Z)}(L)\|_{L(X,Y;Z)}\\ & \underset{[\text{eq: } 14.92]}{=} & \|(\psi\circ\varphi)(L)\|_{L(X,Y;Z)}\\ & = & \|\psi(\varphi(L))\|_{L(X,Y;Z)}\\ & \leqslant_{[\text{eq: } 14.90]} & \|\varphi(L)\|_{L(X,L(Y,Z))}\end{array}$$

which combined with [eq: 14.86] proves that $\|\varphi(L)\|_{L(X,Y;Z)}=\|L\|_{L(X,Y;Z)}$. This together with [eq: 14.93] proves that

$$\varphi: L(X,Y;Z)\rightarrow L(X,L(Y,Z)) \text{ is a linear isometric isomorphism}$$

As $\psi=\varphi^{-1}$ we have by [theorem: 14.119] that

$$\psi: L(X,L(Y,Z))\rightarrow L(X,Y;Z) \text{ is a linear isometric isomorphism}$$

Finally if $x\in X$, $y\in Y$ then we have for $L\in L(X,Y;Z)$ that

$$\varphi(L)(x)(y)\underset{[\text{eq: } 14.84]}{=}\varphi_L(x)(y)\underset{[\text{eq: } 14.81]}{=}L_x(y)\underset{[\text{eq: } 14.78]}{=}L(x,y)$$

so that

$$\forall L\in L(X,L(Y,Z)) \text{ we have } \varphi(L)(x)(y)=L(x,y) \tag{14.94}$$

Further if $L\in L(X,L(Y,Z))$ then

$$\begin{array}{rcl}L(x)(y) & = & ((\varphi\circ\psi)(L))(x)(y)\\ & = & (\varphi(\psi(L))(x))(y)\\ & \underset{[\text{eq: } 14.94]}{=} & \psi(L)(x,y)\end{array}$$

so that

$$\forall L\in L(X,L(Y,Z)) \text{ we have } \psi(L)(x,y)=(L(x))(y)$$

□

If $\langle X,\|\|_X\rangle$ and $\langle Y,\|\|_Y\rangle$ are the same in the above then we have

$$L(X,L(X,Z))\approx L(X,X;Y)\underset{[\text{definition: } 14.194]}{=}L^2(X;Z)$$

actually we can extend this to arbitrary $n\in\mathbb{N}$ as is show in the next theorem. This result will be very important if we define higher order derivatives.

Theorem 14.199. *Let $n\in\mathbb{N}$, $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$ be normed spaces over $\mathbb{K}$ then we can construct a* ***unique*** *linear isometric isomorphism*

$$\varphi_{[X,Y,n]}: L^{n+1}(X;Y)\rightarrow L(X,L^n(X;Y))$$

and if we take $\psi_{[X,Y,n]}=\varphi^{-1}_{[X,Y,n]}$ we have the linear isometric isomorphism

$$\psi_{[X,Y,n]}: L(X,L^n(X;Y))\rightarrow L^{n+1}(X;Y)$$

such that for every $(x_1,\ldots,x_{n+1})\in X^n$

$$\forall L\in L^{n+1}(X;Y)\ \ \varphi_{[X,Y,n]}(L)(x_1)(x_2,\ldots,x_{n+1})=L(x_1,\ldots,x_{n+1})$$

and

$$\forall L\in L(X,L^n(X;Y))\ \ \psi_{[X,Y,n]}(L)(x_1,\ldots,x_{n+1})=L(x_1)(x_2,\ldots,x_{n+1})$$

Further if there exist a $\varphi': L^{n+1}(X;Y)\rightarrow L(X,L^n(X;Y))$ such that $\forall(x_1,\ldots,x_{n+1})\in X^n$ $\varphi(L)(x_1)(x_2,\ldots,x_{n+1})=L(x_1,\ldots,x_{n+1})$ then

$$\varphi_{[X,Y,n]}=\varphi'$$

and if there exist a $\psi': L(X,L^n(X;Y))\rightarrow L^{n+1}(X;Y)$ such that $\forall(x_1,\ldots,x_{n+1})\in X^n$ $\psi(L)(x_1,\ldots,x_{n+1})=L(x_1)(x_2,\ldots,x_{n+1})$ then

$$\psi_{[X,Y,n]}=\psi'$$

In other words using [definition: 14.194] it follows that

$$L(X,L^n(X;Y))\approx L^{n+1}(X;Y)$$

so we can identify $L(X,L^n(X;Y))$ with $L^{n+1}(X;Y)$ using a ***unique linear isometric isomorphism****.*

Proof. For $n\in\mathbb{N}$ we have either:

$n=1$. Then by [theorem: 14.198] we have that there exist linear isometric isomorphisms

$$\varphi L(X,X;Y)\rightarrow L(X,L(X,Y))$$

$$\psi: L(X,L(X,Y))\rightarrow L(X,X;Y)$$

and for $(x_1,x_2)\in X\cdot X=X^2$ that

$$\forall L\in L(X\cdot X;Y) \text{ we have } \varphi(L)(x_1)(x_2)=L(x_1,x_2)$$

and

$$\forall L\in L(X,L(X,Y)) \text{ we have } \psi(L)(x_1,x_2)=L(x_1)(x_2)$$

Which, as by [definition: 14.194] $L(X,Y)=L^1(X,Y)$ and $L(X,X;Y)=L^2(X,Y)$, proves the first part of the theorem for the case $n=1$.

$1<n$. To prove the theorem for this case we use a slightly modified version of the proof of the previous theorem. Let $L\in L^{n+1}(X;Z)$ and $x\in X$ define

$$L_x: X^n\rightarrow Y \text{ by } L_x(x_1,\ldots,x_n)=L(x,x_1,\ldots,x_n) \tag{14.95}$$

Given $r,s\in X$ and $\alpha\in\mathbb{K}$ then we have for $i\in\{1,\ldots,n\}$ that

$$\begin{aligned}
L_x\Big(x_1,\ldots\underbrace{r+\alpha\cdot s}_{i}\ldots,x_n\Big) &=\\
L\Big(x,x_1,\ldots\underbrace{r+\alpha\cdot s}_{i+1}\ldots,x_n\Big) &\underset{\text{multilinearity}}{=}\\
L\Big(x,x_1,\ldots\underbrace{r}_{i+1}\ldots,x_n\Big)+\alpha\cdot L\Big(x,x_1,\ldots\underbrace{s}_{i+1}\ldots,x_n\Big) &=\\
L_x\Big(x_1,\ldots\underbrace{r}_{i}\ldots,x_n\Big)+\alpha\cdot L_x\Big(x_1,\ldots\underbrace{s}_{i}\ldots,x_n\Big)
\end{aligned}$$

proving that

$$L_x\in\mathrm{Hom}^n(X;Y)\underset{\text{def}}{=}\mathrm{Hom}\Big(\underbrace{X,\ldots,X}_{n};Y\Big)$$

Further for $r \in X^n$ we have if we define $(y_1, \ldots, y_{n+1}) \in X^{n+1}$ by $y_i = \begin{cases} x \text{ if } i = 1 \\ r_{i-1} \text{ if } i \in \{2, \ldots, n+1\} \end{cases}$

$$\begin{array}{rcl} \|L_x(r_1, \ldots, r_n)\|_Y & = & \|L(x, r_1, \ldots, r_n)\|_Y \\ & = & \|L(y_1, \ldots, y_{n+1})\|_Y \\ & \leqslant_{[\text{theorem: } 14.193]} & \|L\|_{L\left(\underbrace{X, \ldots, X}_{n+1}; Y\right)} \cdot \prod_{i \in \{1, \ldots, n+1\}} \|y_i\|_X \\ & \underset{\text{def}}{=} & \|L\|_{L^{n+1}(X;Y)} \cdot \prod_{i \in \{1, \ldots, n+1\}} \|y_i\|_X \\ & = & \|L\|_{L^{n+1}(X;Y)} \cdot \|y_1\|_X \cdot \prod_{i \in \{2, \ldots, n+1\}} \|y_i\|_X \\ & = & \|L\|_{L^{n+1}(X;Y)} \cdot \|x\|_X \cdot \prod_{i \in \{2, \ldots, n+1\}} \|r_{i-1}\|_X \\ & = & (\|L\|_{L^{n+1}(X;Y)} \cdot \|x\|_X) \cdot \prod_{i \in \{1, \ldots, n\}} \|r_i\|_X \end{array}$$

proving by [theorem: 14.187] and [definition: 14.178] that

$$\forall x \in X \; L_x \in L^n(X;Y) \text{ and } \|L_x\|_{L^n(X;Y)} \leqslant \|L\|_{L^{n+1}(X;Y)} \cdot \|x\|_X \tag{14.96}$$

The above allows us, given $L \in L^{n+1}(X;Y)$, to define

$$\varphi_L: X \to L^n(X;Y) \text{ by } \varphi_L(x) = L_x \tag{14.97}$$

Then if $x, y \in X$ and $\alpha \in \mathbb{K}$ we have for $r \in X^n$ that

$$\begin{array}{rcl} (\varphi_L(x + \alpha \cdot y))(r) & \underset{[\text{eq: } 14.97]}{=} & L_{x+\alpha \cdot y}(r_1, \ldots, r_n) \\ & \underset{[\text{eq: } 14.95]}{=} & L(x + \alpha \cdot y, r_1, \ldots, r_n) \\ & = & L(x, r_1, \ldots, r_n) + \alpha \cdot L(y, r_1, \ldots, r_n) \\ & \underset{[\text{eq: } 14.95]}{=} & L_x(r_1, \ldots, r_n) + \alpha \cdot L_y(r_1, \ldots, r_n) \\ & \underset{[\text{eq: } 14.97]}{=} & \varphi_L(x)(r) + \alpha \cdot \varphi_L(y)(r) \\ & = & (\varphi_L(x) + \alpha \cdot \varphi_L(y))(r) \end{array}$$

proving that $\varphi_L(x + \alpha \cdot y) = \varphi_L(x) + \alpha \cdot \varphi_L(y)$. Hence

$$\varphi_L \in \mathrm{Hom}(X, L^n(X;Y)) \tag{14.98}$$

For $x \in X$ we have

$$\|\varphi_L(x)\|_{L^n(X;Y)} \underset{[\text{eq: } 14.97]}{=} \|L_x\|_{L^n(X;Y)} \leqslant_{[\text{eq: } 14.96]} \|L\|_{L^{n+1}(X;Y)} \cdot \|x\|_X$$

proving by [theorem: 14.174] and [definition: 14.178] that

$$\varphi_L \in L(X, L^n(X;Y)) \text{ and } \|\varphi_L\|_{L(X, L^n(X;Y))} \leqslant \|L\|_{L^{n+1}(X;Y)} \tag{14.99}$$

The above let us define

$$\varphi: L^{n+1}(X;Y) \to L(X, L^n(X;Y)) \text{ by } \varphi(L) = \varphi_L \tag{14.100}$$

Let $L_1, L_2 \in L^{n+1}(X;Y)$ and $\alpha \in \mathbb{K}$ then for $x \in X$ and $\forall r \in X^n$ we have

$$\begin{array}{rcl} (L_1 + \alpha \cdot L_2)_x(r) & = & (L_1 + \alpha \cdot L_2)(x, r_1, \ldots, r_n) \\ & = & L_1(x, r_1, \ldots, r_n) + \alpha \cdot L_2(x, r_1, \ldots, r_n) \\ & = & (L_1)_x(r_1, \ldots, r_n) + \alpha \cdot (L_2)_x(r_1, \ldots, r_n) \\ & = & ((L_1)_x + \alpha \cdot (L_2)_x)(r_1, \ldots, r_n) \\ & = & ((L_1)_x + \alpha \cdot (L_2)_x)(r) \end{array}$$

proving that $(L_1 + \alpha \cdot L_2)_x = (L_1)_x + \alpha \cdot (L_2)_x$. Hence

$$\varphi_{L_1 + \alpha \cdot L_2}(x) = (L_1 + \alpha \cdot L_2)_x = (L_1)_x + \alpha \cdot (L_2)_x = \varphi_{L_1}(x) + \alpha \cdot \varphi_{L_2}(x)$$

so

$$\varphi_{L_1+\alpha\cdot L_2}=\varphi_{L_1}+\alpha\cdot\varphi_{L_2}$$

proving by [eq: 14.100]

$$\varphi(L_1+\alpha\cdot L_2)\underset{[\text{eq: }14.100]}{=}\varphi(L_1)+\alpha\cdot\varphi(L_2)$$

so that

$$\varphi\in\mathrm{Hom}(L^{n+1}(X;Y),L(X,L^n(X;Y)))\tag{14.101}$$

Further given $L\in L^{n+1}(X;Y)$ we have

$$\|\varphi(L)\|_{L(X,L^n(X;Y))}\underset{[\text{eq: }14.100]}{=}\|\varphi_L\|_{L(X,L^n(X;Y))}\leqslant_{[\text{eq: }14.99]}\|L\|_{L^{n+1}(X;Y)}$$

giving by [theorem: 14.174]

$$\varphi\in L(L^{n+1}(X;Y),L(X,L^n(X;Y)))\text{ and }\|\varphi(L)\|_{L(X,L^n(X;Y))}\leqslant\|L\|_{L^{n+1}(X;Y)}\tag{14.102}$$

Next we must define the inverse mapping of φ. Let $L\in L(X,L^n(X;Y))$ and define

$$\psi_L\colon X^{n+1}\to Y\text{ by }\psi_L(x_1,\ldots,x_{n+1})=L(x_1)(x_2,\ldots,x_{n+1})\tag{14.103}$$

Let $i\in\{1,\ldots,n+1\}$, $y,z\in X$, $x\in\prod_{j\in\{1,\ldots,n+1\}\setminus\{i\}}X$ and $\alpha\in\mathbb{K}$ then for i we have either:

$i=1$. Then

$$\begin{aligned}\psi_L\Bigl(x_1,\ldots\underbrace{y+\alpha\cdot z}_{i}\ldots,x_{n+1}\Bigr)&\underset{i=1}{=}\\ \psi_L(y+\alpha\cdot z,x_2,\ldots,x_{n+1})&\underset{[\text{eq: }14.103]}{=}\\ L(y+\alpha\cdot z)(x_2,\ldots,x_{n+1})&\underset{L\in L(X,L^n(X;Y))}{=}\\ (L(y)+\alpha\cdot L(z))(x_2,\ldots,x_{n+1})&=\\ (L(y))(x_2,\ldots,x_{n+1})+\alpha\cdot(L(z))(x_2,\ldots,x_{n+1})&\underset{[\text{eq: }14.103]}{=}\\ \psi_L(y,x_2,\ldots,x_{n+1})+\alpha\cdot\psi_L(z,x_2,\ldots,x_{n+1})&\underset{i=1}{=}\\ \psi_L\Bigl(x_1,\ldots\underbrace{y}_{i}\ldots,x_{n+1}\Bigr)+\alpha\cdot\psi_L\Bigl(x_1,\ldots\underbrace{z}_{i}\ldots,x_{n+1}\Bigr)&\end{aligned}$$

$i\in\{2,\ldots,n+1\}$. Then

$$\begin{aligned}\psi_L\Bigl(x_1,\ldots\underbrace{y+\alpha\cdot z}_{i}\ldots,x_{n+1}\Bigr)&\underset{[\text{eq: }14.103]\wedge i\neq 1}{=}\\ (L(x_1))\Bigl(x_2,\ldots\underbrace{y+\alpha\cdot z}_{i-1}\ldots,x_{n+1}\Bigr)&\underset{L(x_1)\in L^n(X;Y)}{=}\\ (L(x_1))\Bigl(x_2,\ldots\underbrace{y}_{i-1}\ldots,x_{n+1}\Bigr)+\alpha\cdot(L(x_1))\Bigl(x_2,\ldots\underbrace{z}_{i-1}\ldots,x_{n+1}\Bigr)&\underset{[\text{eq: }14.103]}{=}\\ \psi_L\Bigl(x_1,\ldots\underbrace{y}_{i}\ldots,x_{n+1}\Bigr)+\alpha\cdot\psi_L\Bigl(x_1,\ldots\underbrace{z}_{i}\ldots,x_{n+1}\Bigr)&\end{aligned}$$

proving that

$$\psi_L\in\mathrm{Hom}^{n+1}(X;Y)\tag{14.104}$$

Further if $x\in X^{n+1}$ then

$$\begin{aligned}\|\psi_L(x_1,\ldots,x_{n+1})\|_Y&\underset{[\text{eq: }14.103]}{=}&&\|L(x_1)(x_2,\ldots,x_{n+1})\|_Y\\ &\leqslant_{L(x_1)\in L^n(X;Y)}&&\|L(x_1)\|_{L^n(X;Y)}\cdot\prod_{i\in\{2,\ldots,n+1\}}\|x_i\|_X\\ &\leqslant_{L\in}&&\|L\|_{L(X,L^n(X;Y))}\cdot\|x_1\|_X\cdot\prod_{i\in\{2,\ldots,n+1\}}\|x_i\|_X\\ &=&&\|L\|_{L(X,L^n(X;Y))}\cdot\prod_{i\in\{1,\ldots,n+1\}}\|x_i\|_X\end{aligned}$$

proving by [theorems: 14.187, 14.193] that

$$\psi_L \in L^{n+1}(X;Y) \text{ and } \|\psi_L\|_{L^{n+1}(X;Y)} \leqslant \|L\|_{L(X,L^n(X;Y))} \tag{14.105}$$

The above allows us to define

$$\psi: L(X, L^n(X;Y)) \to L^{n+1}(X;Y) \text{ by } \psi(L) = \psi_L \tag{14.106}$$

where by [eq: 14.105]

$$\forall L \in L(X, L^n(X;Y)) \text{ we have } \|\psi(L)\|_{L^{n+1}(X;Y)} \leqslant \|L\|_{L(X,L^n(X;Y))} \tag{14.107}$$

Let $L \in L(X, L^n(X,Y))$ then for $x \in X$ and $r \in X^n$ we have

$$\begin{array}{rcl}(((\varphi\circ\psi)(L))(x))(r) & = & (((\varphi\circ\psi)(L))(x))(r_1,\ldots,r_n)\\ & = & (((\varphi(\psi(L)))(x)))(r_1,\ldots,r_n)\\ & \underset{\text{[eq: 14.100}}{=} & (\varphi_{\psi(L)}(x))(r_1,\ldots,r_n)\\ & \underset{\text{[eq: 14.97]}}{=} & (\psi(L))_x(r_1,\ldots,r_n)\\ & \underset{\text{[eq: 14.95]}}{=} & (\psi(L))(x,r_1,\ldots,r_n)\\ & \underset{\text{[eq: 14.106]}}{=} & \psi_L(x,r_1\ldots,r_n)\\ & \underset{\text{[eq: 14.103]}}{=} & (L(x))(r_1,\ldots,r_n)\\ & = & (L(x))(r)\end{array}$$

from which it follows that $((\varphi\circ\psi)(L))(x) = L(x)$, hence $(\varphi\circ\psi)(L) = L$ proving that

$$\varphi\circ\psi = \mathrm{Id}_{L(X,L^n(X;Y))} \tag{14.108}$$

On the other hand if $L \in L^{n+1}(X;Y)$ and $x \in X^{n+1}$ then we have

$$\begin{array}{rcl}((\psi\circ\varphi)(L))(x) & = & ((\psi\circ\varphi)(L))(x_1,\ldots,x_{n+1})\\ & = & (\psi(\varphi(L)))(x_1,\ldots,x_{n+1})\\ & \underset{\text{[eq: 14.106]}}{=} & \psi_{\varphi(L)}(x_1,\ldots,x_{n+1})\\ & \underset{\text{[eq: 14.103]}}{=} & ((\varphi(L))(x_1))(x_2,\ldots,x_n)\\ & \underset{\text{[eq: 14.100]}}{=} & (\varphi_L(x_1))(x_2,\ldots,x_{n+1})\\ & \underset{\text{[eq: 14.97]}}{=} & L_{x_1}(x_2,\ldots,x_{n+1})\\ & \underset{\text{[eq: 14.95]}}{=} & L(x_1,x_2,\ldots,x_{n+1})\\ & = & L(x)\end{array}$$

proving that $(\psi\circ\varphi)(L) = L$ or

$$\psi\circ\varphi = \mathrm{Id}_{L^{n+1}(X;Y)} \tag{14.109}$$

From [eqs: 14.108, 14.109] it follows by [theorem: 2.71] that

$$\varphi: L^{n+1}(X;Y) \to L(X, L^n(X;Y))$$

is a bijection. Combining this with]eq: 14.101] proves that

$$\varphi \text{ is a linear isomorphism} \tag{14.110}$$

Finally we have for $L \in L^{n+1}(X;Y)$ that

$$\begin{array}{rcl}\|L\|_{L^{n+1}(X;Y)} & = & \|\mathrm{Id}_{L^{n+1}(X;Y)}(L)\|_{L^{n+1}(X;Y)}\\ & \underset{\text{[eq: 14.109]}}{=} & \|(\psi\circ\varphi)(L)\|_{L^{n+1}(X;Y)}\\ & = & \|\psi(\varphi(L))\|_{L^{n+1}(X;Y)}\\ & \leqslant_{\text{[eq: 14.107]}} & \|\varphi(L)\|_{L(X,L^n(X;Y))}\end{array}$$

which combined with [eq: 14.102] proves that $\|\varphi(L)\|_{L(X,L^n(X;Y))} = \|L\|_{L^{n+1}(X;Y)}$. This together with [eq: 14.110] proves that

$$\varphi: L^{n+1}(X;Y) \to L(X, L^n(X,Y)) \text{ is a linear isometric isomorphism}$$

and as $\psi = \varphi^{-1}$ we have by [theorem: 14.119] that

$$\psi: L(X, L^n(X, Y)) \rightarrow L^{n+1}(X; Y) \text{ is a linear isometric isomorphism}$$

Let $(x_1, \ldots, x_{n+1}) \in X^{n+1}$. Take $L \in L^{n+1}(X; Y)$ then we have

$$\begin{aligned}\varphi(L)(x_1)(x_2, \ldots, x_{n+1}) &\underset{[\text{eq: } 14.100]}{=} \varphi_L(x_2, \ldots, x_{n+1}) \\ &\underset{[\text{eq: } 14.97]}{=} L_{x_1}(x_2, \ldots, x_{n+1}) \\ &\underset{[\text{eq: } 14.95]}{=} L(x_1, \ldots, x_{n+1})\end{aligned}$$

proving that

$$\forall L \in L^{n+1}(X; Y) \text{ we have } \varphi(L)(x_1)(x_2, \ldots, x_{n+1}) = L(x_1, \ldots, x_{n+1}) \tag{14.111}$$

furhter if $L \in L(X, L^n(X, Y))$ then

$$\begin{aligned}L(x_1)(x_2, \ldots, x_{n+1}) &= ((\varphi \circ \psi)(L))(x_1)(x_2, \ldots, x_{n+1}) \\ &= (\varphi(\psi(L))(x_1))(x_2, \ldots, x_{n+1}) \\ &\underset{[\text{eq: } 14.111]}{=} \psi(L)(x_1, x_2, \ldots, x_{n+1})\end{aligned}$$

proving that

$$\forall L \in L(X, L^n(X, Y)) \text{ we have } \psi(L)(x_1, x_2, \ldots, x_{n+1}) = L(x_1)(x_2, \ldots, x_{n+1})$$

proving the first part of the theorem for the case $1 < n$.

As for uniqueness if $\varphi': L^{n+1}(X; Y) \rightarrow L(X, L^n(X, Y))$ satisfies

$$\varphi(L)(x_1)(x_2, \ldots, x_{n+1}) = L(x_1, \ldots, x_{n+1})$$

then we have $\forall L \in L^{n+1}(X; Y)$ $\forall x \in X$ that $\forall y \in X^n$

$$\varphi(L)(x)(y) = L(x, y_1, \ldots, y_n) = \varphi(L)(x)(y)$$

so that $\varphi'(L)(x) = \varphi(L)(x)$ or $\varphi'(L) = \varphi(L)$ proving that

$$\varphi' = \varphi$$

Likewise if $\psi': L(X, L^n(X, Y)) \rightarrow L^{n+1}(X; Y)$ satisfies

$$\psi'(L)(x_1, x_2, \ldots, x_{n+1}) = L(x_1)(x_2, \ldots, x_{n+1})$$

then we have $\forall L \in L(X, L^n(X; Y))$ $\forall x \in X^{n+1}$ that $\forall y \in X^n$

$$\psi'(L)(x_1, \ldots, x_{n+1}) = L(x_1)(x_2, \ldots, x_{n+1}) = \psi(L)(x_1, \ldots, x_{n+1})$$

so that $\psi'(L) = \psi(L)$ proving that

$$\psi' = \psi$$

Taking $\varphi_{[X,Y,n]} = \varphi$ and $\psi_{[X,Y,n]} = \psi$ completes then the proof if the theorem. □

14.7 Separation

Definition 14.200. *A topological space $\langle X, \mathcal{T}\rangle$ is **Hausdorff** if $\forall x, y \in X$ with $x \neq y$ there exists $U, V \in \mathcal{T}$ such that $x \in U$, $y \in V$ and $U \bigcap V = \varnothing$. In other words a topological space is **Hausdorff** if two different points can be separated by disjoint open sets.*

It turns out that all metric and normed spaces are Hausdorff

Theorem 14.201. *If $\langle X, d\rangle$ is a metric space then $\langle X, \mathcal{T}_d\rangle$ is Hausdorff [using the metric topology $\mathcal{T}_d$]*

***Note* 14.202.** *If $\langle X, \|\|\rangle$ is a normed space then $\langle X, \mathcal{T}_{\|\|}\rangle \underset{[\text{definition: } 14.92]}{=} \langle X, \mathcal{T}_{d_{\|\|}}\rangle$ so that every normed space is also Hausdorff.*

Proof. Let $x, y \in X$ such that $x \neq y$ then as d is a metric we have that $0 < d(x, y)$ [for if $d(x, y) = 0$ then $x = y$]. Take $\varepsilon = d(x, y) > 0$ then $x \in B_d\left(x, \frac{\varepsilon}{2}\right)$ and $y \in B_d\left(y, \frac{\varepsilon}{2}\right)$. Assume that

$$z \in B_d\left(x, \frac{\varepsilon}{2}\right) \bigcap B_d\left(y, \frac{\varepsilon}{2}\right)$$

then $\varepsilon = d(x,y) \leqslant d(x,z) + d(z,y) = d(x,z) + d(y,z) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$ leading to the contradiction $\varepsilon < \varepsilon$. Hence the assumption is wrong so that

$$B_d\left(x, \frac{\varepsilon}{2}\right) \bigcap B_d\left(y, \frac{\varepsilon}{2}\right) = \varnothing \text{ and } x \in B_d\left(x, \frac{\varepsilon}{2}\right) \in \mathcal{T}_d,\ y \in B_d\left(y, \frac{\varepsilon}{2}\right) \in \mathcal{T}_d$$

□

One benefit of a Hausdorff spaces is that every finite set is closed in a Hausdorff space.

Theorem 14.203. *If $\langle X, \mathcal{T} \rangle$ is a Hausdorff topological space then every finite set is closed.*

Proof. Let $x \in X$ then if $y \in X \setminus \{x\}$ we have $y \neq x$ so by the Hausdorff condition there exists $U, V \in \mathcal{T}$ with $x \in U$, $y \in V$ and $U \bigcap V = \varnothing$. Hence $x \notin V$ so that $y \in V \subseteq X \setminus \{x\}$ proving by [theorem: 14.7] that $X \setminus \{x\}$ is open so that $\{x\}$ is closed. Hence we have that

$$\forall x \in X \text{ we have that } \{x\} \text{ is closed} \tag{14.112}$$

Now if $A \subseteq X$ is finite then we have either:

***$A = \varnothing$*.** Then A is trivially closed.

***$A \neq \varnothing$*.** Then by [theorem: 10.85] there exist a bijection $\beta: \{1, \ldots, n\} \rightarrow A$ so that

$$A = \beta(\{1, \ldots, n\}) = \beta\left(\bigcup_{i \in \{1, \ldots, n\}} \{i\}\right) = \bigcup_{i \in \{1, \ldots, n\}} \beta(\{i\}) = \bigcup_{i \in \{1, \ldots, n\}} \{\beta(i)\}$$

proving by [eq: 14.112] and [theorem: 14.21] that A is closed. □

Definition 14.204. *A topological space $\langle X, \mathcal{T} \rangle$ is **regular** if for every closed set A and $x \in X \setminus A$ there exists $U, V \in \mathcal{T}$ such that $x \in U$, $A \subseteq V$ and $U \bigcap V = \varnothing$. In other words a topological space is **regular** if every closed set and a point outside the closed set can be separated by disjoint open sets.*

All metric and normed spaces are not only Hausdorff but also regular.

Theorem 14.205. *If $\langle X, d \rangle$ is a metric space then $\langle X, \mathcal{T}_d \rangle$ is regular [using the metric topology $\mathcal{T}_d$]*

***Note* 14.206.** *If $\langle X, \|\| \rangle$ is a normed space then $\langle X, \mathcal{T}_{\|\|} \rangle \underset{\text{[definition: 14.92]}}{=} \langle X, \mathcal{T}_{d_{\|\|}} \rangle$ so that every normed space is also regular.*

Proof. Let A be a closed set and $x \in X \setminus A$. As A is closed $X \setminus A$ is open, hence by [theorem: 14.64] there exist a $\delta \in \mathbb{R}^+$ such that

$$x \in B_d(x, \delta) \subseteq X \setminus A \text{ hence } B_d(x, \delta) \bigcap A = \varnothing \tag{14.113}$$

Given $a \in A$ assume that $z \in B_d\left(x, \frac{\delta}{2}\right) \bigcap B_d(a, \delta_2)$ then $d(z, x) < \frac{\delta}{2} \wedge d(z, a) < \frac{\delta}{2}$ then

$$d(x, a) \leqslant d(x, z) + d(z, a) < \frac{\delta}{2} + \frac{\delta}{2} = \delta$$

so that $a \in B_d(x, \delta)$, hence $a \in A \bigcap B_d(x, \delta) = \varnothing$ contradicting [eq: 14.113]. Hence it follows that

$$B_d\left(x, \frac{\delta}{2}\right) \bigcap B_d\left(a, \frac{\delta}{2}\right) = \varnothing \tag{14.114}$$

Take $V = \bigcup_{a \in A} B_d\left(a, \frac{\delta}{2}\right) \in \mathcal{T}_d$ and $U = B_d\left(x, \frac{\delta}{2}\right) \in \mathcal{T}_d$ then clearly

$$A \subseteq V \text{ and } x \in U$$

Further

$$V \bigcap U = \left(\bigcup_{a \in A} B_d\left(a, \frac{\delta}{2}\right)\right) \bigcap B_d\left(x, \frac{\delta}{2}\right) = \bigcup_{a \in A} \left(B_d\left(a, \frac{\delta}{2}\right) \bigcap B_d\left(x, \frac{\delta}{2}\right)\right) \underset{\text{[eq: 14.114]}}{=} \varnothing$$

□

Theorem 14.207. *Let $\langle X, \mathcal{T} \rangle$ be a **regular** topological space, U a open set, $x \in U$ then there exist a $V \in \mathcal{T}$ such that $x \in V$ and $\bar{V} \subseteq U$.*

Proof. Take $x \in U$ then as $X \setminus U$ is a closed set and $x \in U \underset{\text{[theorem: 1.26]}}{=} X \setminus (X \setminus U)$ there exist a $V, W \in \mathcal{T}$ such that $x \in V$, $X \setminus U \subseteq W$ and $V \bigcap W = \varnothing$. As $V \bigcap W = \varnothing$ we have that $V \subseteq X \setminus W$, as $X \setminus W$ is closed it follows from [theorem: 14.24] that $\bar{V} \subseteq X \setminus W$. As $X \setminus U \subseteq W$ we have $X \setminus W \subseteq X \setminus (X \setminus U) = U$ so that

$$\bar{V} \subseteq U$$

□

Corollary 14.208. *Let $\langle X, \mathcal{T}\rangle$ be a regular topological space, A a closed set, $x \in X \setminus A$ and U a open set with $x \in U$ then there exist a open set $V \in \mathcal{T}$ such that*

$$x \in V,\ \bar{V} \bigcap A = \varnothing \text{ and } \bar{V} \subseteq U$$

Proof. As A is a closed set and $x \notin A$ there exists by regularity two open sets $V_1, V_2 \in \mathcal{T}$ such that

$$x \in V_1,\ A \subseteq V_2 \text{ and } V_1 \bigcap V_2 = \varnothing$$

As $x \in V_1 \bigcap U \in \mathcal{T}$ there exist by [theorem: 14.207] a open set $V \in \mathcal{T}$ such that

$$x \in V \text{ and } \bar{V} \subseteq V_1 \bigcap U \subseteq U$$

so that

$$\bar{V} \bigcap A \subseteq (V_1 \bigcap U) \bigcap A \subseteq V_1 \bigcap A \subseteq V_1 \bigcap V_2 = \varnothing$$

Hence we have

$$V \in \mathcal{T},\ x \in V,\ \bar{V} \subseteq U \text{ and } \bar{V} \bigcap A = \varnothing$$

□

Definition 14.209. *A topological space $\langle X, \mathcal{T}\rangle$ is **normal** if for every pair of distinct closed sets A, B there exist $U, V \in \mathcal{T}$ such that $A \subseteq U$, $B \subseteq V$ and $U \bigcap V$. In other words a topological space is normed if every disjoint pair of closed sets can be separated by a disjoint pair of open sets.*

Theorem 14.210. *Let $\langle X, \mathcal{T}\rangle$ be a topological space then we have:*

1. *If $\langle X, \mathcal{T}\rangle$ is a normal space such that $\forall x \in X$ $\{x\}$ is closed then $\langle X, \mathcal{T}\rangle$ is regular*
2. *If $\langle X, \mathcal{T}\rangle$ is a regular space such that $\forall x \in X$ $\{x\}$ is closed then $\langle X, \mathcal{T}\rangle$ is Hausdorff.*

Proof.

1. Let A be a closed set and $x \in X \setminus A$ then by the hypothesis $\{x\}$ is closed, so by normality there exists $U, V \in \mathcal{T}$ such that $A \subseteq U$, $x \in \{x\} \subseteq V$ and $U \bigcap V = \varnothing$. Hence $\langle X, \mathcal{T}\rangle$ is regular.
2. Let $x, y \in X$ such that $x \neq y$ then $\{x\}$ and $\{y\}$ are closed and $\{x\} \bigcap \{y\} = \varnothing$ so that by normality there exists $U, V \in \mathcal{T}$ such that $x \in \{x\} \subseteq U$, $y \in \{y\} \subseteq V$ and $U \bigcap V = \varnothing$. Hence $\langle X, \mathcal{T}\rangle$ is Hausdorff. □

Definition 14.211. *Let $\langle X, \mathcal{T}\rangle$ be a topological space and $x \in X$ then $A \subseteq X$ is a **neighborhood of x** iff there exist a open set $U \in \mathcal{T}$ such that $x \in U \subseteq A$. If A itself is open then A is called a **open neighborhood of** x.*

Definition 14.212. *Let $\langle X, T\rangle$ 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 exist a $N \in \mathcal{N}$ such that $x \in N \subseteq A$.*

Definition 14.213. *A topological space $\langle X, \mathcal{T}\rangle$ is **first countable** if every element of X has a **countable** fundamental system of neighborhoods.*

Theorem 14.214. *If $\langle X, d\rangle$ is a metric space then $\langle X, \mathcal{T}_d\rangle$ is **first countable** [using the metric topology $\mathcal{T}_d$].*

***Note* 14.215.** *If $\langle X, \|\|\rangle$ is a normed space then $\langle X, \mathcal{T}_{\|\|}\rangle \underset{[definition:\ 14.92]}{=} \langle X, \mathcal{T}_{d_{\|\|}}\rangle$ so that every normed space is also first countable.*

Proof. Given $x \in X$ define $\mathcal{N}_x = \{B_d(x, \frac{1}{n}) | n \in \mathbb{N}\}$ which is countable [see theorem: 6.69]. If A is a neighborhood of x then there exist a $U \in \mathcal{T}$ such that $x \in U \subseteq A$. Using [theorem: 14.64] there exist a $\delta \in \mathbb{R}^+$ such that $x \in B_d(x, \delta) \subseteq U$. As $0 < \delta$ there exist a $n \in \mathbb{N}$ such that $0 < \frac{1}{n} < \delta$ [see theorem: 10.30] hence we have that $x \in B_d(x, \frac{1}{n}) \subseteq B_d(x, \delta) \subseteq U \subseteq A$ which as $B_d(x, \frac{1}{n}) \in \mathcal{N}_x$ proves first countability. □

Definition 14.216. *A topological space $\langle X, \mathcal{T}\rangle$ is **second countable** if $\mathcal{T}$ has a **countable basis**.*

Theorem 14.217. *Let $\langle X, \mathcal{T}\rangle$ be a **second countable** topological space then $\langle X, \mathcal{T}\rangle$ 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 $x \in B \subseteq U \subseteq A$ [see theorem: 14.34]. So $\mathcal{B}$ is a fundamentally system of neighborhoods of every element of x. □

14.8 Compact Spaces

Definition 14.218. *A topological space $\langle X,\mathcal{T}\rangle$ is* ***compact*** *if for every $\{U_i\}_{i\in I}\subseteq\mathcal{T}$ satisfying $X=\bigcup_{i\in I}U_i$ there exist a finite $J\subseteq I$ such that $X=\bigcup_{i\in J}U_i$. In other words a topological space is compact if every covering of the space by a family of open sets contains a finite subfamily of sets that covers the space.*

In addition to compact topological spaces we have also compact subsets of a topological space.

Definition 14.219. *Let $\langle X,\mathcal{T}\rangle$ be a topological space and $C\subseteq X$ then C is a* ***compact subset*** *if $\langle X,\mathcal{T}_{|C}\rangle$ is a* ***compact topological space*** *where $\mathcal{T}_{|C}=\{U\bigcap C|U\in\mathcal{T}\}$ is the subspace topology on C [see definition: 14.8]*

We show now that a compact topological spaces is a topological space that is its own compact subspace.

Theorem 14.220. *Let $\langle X,\mathcal{T}\rangle$ be a topological space then we have*

$$\langle X,\mathcal{T}\rangle \text{ is compact} \Leftrightarrow X \text{ is a compact subset of } \langle X,\mathcal{T}\rangle$$

Proof. First note that $\mathcal{T}_{|X} \underset{\text{[definition: 14.8]}}{=} X$ so that $\langle X,\mathcal{T}\rangle=\langle X,\mathcal{T}_{|X}\rangle$

- **⇒.** If $\langle X,\mathcal{T}\rangle$ is compact then $\langle X,\mathcal{T}_{|X}\rangle$ is compact so that X is a compact subset of $\langle X,\mathcal{T}\rangle$.
- **⇐.** If X is a compact subset of $\langle X,\mathcal{T}\rangle$ then $\langle X,\mathcal{T}_{|X}\rangle$ is compact so that $\langle X,\mathcal{T}\rangle$ is compact. □

A alternative definition of a compact subset is given in the following theorem.

Theorem 14.221. *Let $\langle X,\mathcal{T}\rangle$ be a topological space and $C\subseteq X$ then*

$$C \text{ is compact}$$
$$\Updownarrow$$
$$\text{For every } \{U_i\}_{i\in I}\subseteq\mathcal{T} \text{ satisfying } C\subseteq\bigcup_{i\in I}U_i \text{ there exist a finite } J\subseteq I \text{ such that } C\subseteq\bigcup_{i\in J}U_i$$

Proof.

- **⇒.** 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 $C=\bigcup_{i\in J}(U_i\bigcap C)\subseteq\bigcup_{i\in J}U_i$.
- **⇐.** 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 $C=\bigcup_{i\in I}(U_i\bigcap C)\subseteq\bigcup_{i\in I}U_i$. Hence by the hypothesis there exists a finite $J\subseteq I$ such that $C\subseteq\bigcup_{i\in J}U_i\Rightarrow C=C\bigcap(\bigcup_{i\in J}U_i)=\bigcup_{i\in J}(U_i\bigcap C)=\bigcup_{i\in J}V_i$ proving that $\langle C,\mathcal{T}_{|C}\rangle$ is a compact topological space. Hence by definition C is a topological subset of $\langle X,\mathcal{T}\rangle$. □

We can also express compactness using a basis for the topology.

Theorem 14.222. *Let $\langle X,\mathcal{T}\rangle$ be a topological space with a basis $\mathcal{B}$ and $C\subseteq X$ then*

$$C \text{ is a compact subset} \Leftrightarrow \forall\{B_i\}_{i\in I}\subseteq\mathcal{B} \text{ with } C\subseteq\bigcup_{i\in I}B_i \text{ there exist a finite } J \text{ we have } C\subseteq\bigcup_{i\in J}B_i$$

Proof.

- **⇒.** This follows from $\mathcal{B}\subseteq\mathcal{T}$ and [theorem: 14.221].
- **⇐.** Let $\{U_i\}_{i\in I}\subseteq\mathcal{T}$ be a family of open sets such that $C\subseteq\bigcup_{i\in I}U_i$. Then $\forall c\in C$ there exist a $i_c\in I$ such that $c\in U_{i_c}$ as $\mathcal{B}$ is a basis it follows from [theorem: 14.34] that there exists a B_c such that $c\in B_c\subseteq U_{i_c}$. So $C\subseteq\bigcup_{c\in C}B_c$ and by the hypothesis there exists a finite set $F\subseteq C$ such that $C\subseteq\bigcup_{c\in F}B_c$. Define then the finite set $J=\{i_c|c\in F\}\subseteq I$ then if $c\in C$ there exist a $c\in F$ such that $c\in B_c\subseteq U_{i_c}$ where $i_c\in J$ proving that
$$C\subseteq\bigcup_{i\in J}U_i$$
□

Example 14.223. If $\langle X,\mathcal{T}\rangle$ is a topological space then $\varnothing$ is a compact set.

Proof. If $\{U_i\}_{i\in I}$ is a non empty family of open sets covering $\emptyset$ then we have for I either

$I=\emptyset$. then $J=\emptyset=I$ is finite and $\varnothing=\bigcup_{i\in\varnothing}U_i=\bigcup_{i\in J}U_i$ so $\{U_i\}_{i\in J}$ is a finite covering of $\emptyset$

$I\neq\emptyset$. then $\exists i\in I$ so if we take $J=\{i\}\subseteq I$ then J is finite and $\emptyset\subseteq E_i=\bigcup_{j\in\{i\}}U_i$ □

Example 14.224. Let $\langle X,\mathcal{T}\rangle$ be a topological space, $x\in X$ then $\{x\}$ is a compact subset of X.

Proof. Let $\{U_i\}_{i\in I}\subseteq\mathcal{T}$ be such that $\{x\}\subseteq\bigcup_{i\in I}U_i$ then $x\in\bigcup_{i\in I}U_i$ so there exist a $i\in I$ such that $x\in U_i$. Hence $\{x\}\subseteq U_i=\bigcup_{j\in\{i\}}U_j$ which as $\{i\}$ is finite and $\{i\}\subseteq I$ proves that $\{x\}$ is compact. □

Theorem 14.225. *Let $\langle X,\mathcal{T}\rangle$ be a topological space, $n\in\mathbb{N}$ and $\{C_i\}_{i\in\{1,\ldots,n\}}$ a finite family of compact subsets of X then $\bigcup_{i\in\{1,\ldots,n\}}C_i$ is a compact subset.*

Proof. Let $\bigcup_{i\in I}U_i$ be a open cover of $\bigcup_{i\in\{1,\ldots,n\}}C_i$ then $\forall i\in\{1,\ldots,n\}$ we have

$$C_i\subseteq\bigcup_{i\in\{1,\ldots,n\}}C_i\subseteq\bigcup_{j\in I}U_j\underset{[\text{theorem: }14.221]}{\Rightarrow}\exists\text{finite } I_i\subseteq I \text{ such that } C_i\subseteq\bigcup_{j\in I_i}U_j$$

so that

$$\bigcup_{i\in\{1,\ldots,n\}}C_i\subseteq\bigcup_{i\in\{1,\ldots,n\}}\left(\bigcup_{j\in I_i}U_j\right)$$

If now $x\in\bigcup_{i\in\{1,\ldots,.n\}}(\bigcup_{j\in I_i}U_i)$ then there exist a $i\in\{1,\ldots,n\}$ such that $x\in\bigcup_{j\in I_i}U_j$, hence there exists a $j\in I_i\subseteq\bigcup_{k\in\{1,\ldots,n\}}I_k$ with $x\in U_j$, proving that $x\in\bigcup_{j\in\bigcup_{k\in\{1,\ldots,n\}}I_k}U_j$. So we have

$$\bigcup_{i\in\{1,\ldots,.n\}}\left(\bigcup_{j\in I_i}U_i\right)\subseteq\bigcup_{j\in\bigcup_{k\in\{1,\ldots,n\}}I_k}U_j$$

hence

$$\bigcup_{i\in\{1,\ldots,n\}}C_i\subseteq\bigcup_{j\in\bigcup_{k\in\{1,\ldots,n\}}I_k}U_j$$

where $\bigcup_{k\in\{1,\ldots,n\}}I_k\subseteq I$ and by [theorem: 6.35] $\bigcup_{k\in\{1,\ldots,n\}}I_k$ is finite. □

Theorem 14.226. *Let $\langle X,\mathcal{T}\rangle$ be a topological space, $A\subseteq X$ and $C\subseteq A$ then*

$$C \text{ is compact in } \langle A,\mathcal{T}_{|A}\rangle\Leftrightarrow C \text{ is compact in } \langle X,\mathcal{T}\rangle$$

Proof.

$\Rightarrow$. If $\{U_i\}_{i\in I}\subseteq\mathcal{T}$ satisfies $C\subseteq\bigcup_{i\in I}U_i$ then $C\underset{C\subseteq A}{=}C\bigcap A\subseteq\bigcup_{i\in I}(U_i\bigcap A)$ where $\forall i\in I$ $U_i\bigcap A\in\mathcal{T}_{|A}$. As C is compact in $\langle A,\mathcal{T}_{|A}\rangle$, it follows from [theorem: 14.221] that there exist a finite $J\subseteq I$ so that $C\subseteq\bigcup_{i\in J}(U_i\bigcap A)\subseteq\bigcup_{i\in J}U_i$ proving by [theorem: 14.221] that C is compact in $\langle X,\mathcal{T}\rangle$.

$\Leftarrow$. If $\{V_i\}_{i\in I}\subseteq\mathcal{T}_{|A}$ satisfies $C\subseteq\bigcup_{i\in I}V_i$ then $\forall i\in I$ we have that $\exists U_i\in\mathcal{T}$ such that $V_i=U_i\bigcap A\subseteq U_i$. Hence $C\subseteq\bigcup_{i\in I}U_i$ which as C is compact in $\langle X,\mathcal{T}\rangle$ proves by [theorem: 14.221] that there exist a $J\subseteq I$ such that $C\subseteq\bigcup_{i\in J}U_i$. So

$$C=C\bigcap A\subseteq\left(\bigcup_{i\in J}U_i\right)\bigcap A=\bigcup_{i\in J}(U_i\bigcap A)=\bigcup_{i\in J}V_i$$

proving by [theorem: 14.221] that C is compact in $\langle A,\mathcal{T}_{|A}\rangle$ □

In pseudo metric [or normed spaces] every compact set must be bounded.

Theorem 14.227. *Let $\langle X,d\rangle$ be a pseudo metric space and $C\subseteq X$ a compact subset then C is bounded.*

Proof. Let C be a compact subset of X then for $\{B_d(c,1)\}_{c\in C}\subseteq\mathcal{T}_d$ we have $C\subseteq\bigcup_{c\in C}B_d(c,1)$. As C is compact in $\langle X,\mathcal{T}_d\rangle$ we have by [theorem: 14.221] that there exist a finite set $J\subseteq C$ such that

$$C\subseteq\bigcup_{c\in J}B_d(c,1)\tag{14.115}$$

We have now two cases for J:

$J=\varnothing$. Then $C\subseteq\bigcup_{c\in J}B_d(c,1)=\varnothing$ hence by [example: 14.76] C is bounded

$J\neq\varnothing$. Then $\{d(c,t)|(c,t)\in J\times J\}\neq\varnothing$ and by [theorems: 6.40, 6.45] $\{d(c,t)|(c,t)\in J\times J\}$ is finite, so that by [theorem: 6.48] $N=\max(\{d(c,t)|(c,t)\in J\times J\})$ exists. Take then

$$M=N+2$$

Let $x,y\in C$ then by [eq: 14.115] there exists $c_x,c_y\in J$ such that $x\in B_d(c_x,1)$, $y\in B_d(c_y,1)$, hence

$$\begin{array}{rcl} d(x,y) & \leqslant & d(x,c_x)+d(c_x,y)\\ & \leqslant & d(x,c_x)+d(c_x,c_y)+d(c_y,y)\\ & < & 1+N+1\\ & = & M \end{array}$$

proving that C is bounded. □

Corollary 14.228. *Let $\langle X,\|\|\rangle$ be a pseudo normed space and $C\subseteq X$ a compact subset then there exist a $M\in\mathbb{R}^+$ such that $\forall c\in C$ we have $\|c\|\leqslant M$*

Proof. This follows from the previous theorem [theorem: 14.227] and [theorem: 14.106] □

Continuous maps preserves compactness.

Theorem 14.229. *Let $\langle X,\mathcal{T}_X\rangle$, $\langle Y,\mathcal{T}_Y\rangle$ be topological spaces, $f:X\rightarrow Y$ a continuous function and C a compact set in $\langle X,\mathcal{T}_X\rangle$ then $f(C)$ is compact in $\langle Y,\mathcal{T}_Y\rangle$.*

Proof. Let $\{V_i\}_{i\in I}$ be a family of open sets in Y so that $f(C)\subseteq\bigcup_{i\in I}V_i$. Let $x\in C$ then we have $f(x)\in f(C)\subseteq\bigcup_{i\in I}V_i$ and thus $\exists i\in I$ such that $f(x)\in V_i$. By continuity of f there exists a U_i open in X with $x\in U_i$ and $f(x)\in f(U_i)\subseteq V_i$, hence $x\in\bigcup_{i\in I}U_i$. As $x\in C$ was chosen arbitrary it follows that

$$C\subseteq\bigcup_{i\in I}V_i$$

By compactness of C there exist a finite $J\subseteq I$ with

$$C\subseteq\bigcup_{i\in J}U_i\Rightarrow f(C)\subseteq f\left(\bigcup_{i\in J}U_i\right)=\bigcup_{i\in J}f(U_i)\subseteq\bigcup_{i\in J}V_i$$

proving that $f(C)$ is compact. □

The following two theorems shows the relation between compactness and closedness.

Theorem 14.230. *Let $\langle X,\mathcal{T}\rangle$ be a Hausdorff topological space then every compact subset is closed.*

Proof. Let C be a compact subset of X. Take $y\in X\setminus C$ then $\forall c\in C$ we have by the Hausdorff condition that there exists $U_c,V_c\in\mathcal{T}$ such that $c\in U_c$, $y\in V_c$ and $\varnothing\neq U_c\bigcap V_c$. So $C\subseteq\bigcup_{c\in C}U_c$, hence, as C is compact, there exist a finite $F\subseteq C$ satisfying $C\subseteq\bigcup_{c\in F}U_c$. As $\forall c\in F\subseteq C$ we have $y\in V_c$ it follows that $y\in\bigcap_{c\in F}V_c$ which is open because F is finite. Further

$$\begin{array}{rcl} \left(\bigcap_{c\in F}V_c\right)\bigcap C & \subseteq & \left(\bigcap_{c\in F}V_c\right)\bigcap\left(\bigcup_{x\in F}U_x\right)\\ & = & \bigcup_{x\in F}\left(\left(\bigcap_{c\in F}V_c\right)\bigcap U_x\right)\\ & \underset{x\in F\Rightarrow\bigcap_{c\in F}V_c\subseteq V_x}{\subseteq} & \bigcup_{x\in F}(V_x\bigcap U_x)\\ & = & \bigcup_{x\in F}\varnothing\\ & = & \varnothing \end{array}$$

So for every $y\in X\setminus C$ we have found a open set $\bigcap_{c\in F}V_c$ such that $y\in\bigcap_{c\in F}V_c\subseteq X\setminus C$ proving by [theorem: 14.7] that $X\setminus C$ is open or that C is closed. □

Theorem 14.231. *Let $\langle X,\mathcal{T}\rangle$ be a topological space, C a compact subset of X, $F\subseteq C$ a closed subset of C then C is compact.*

Proof. Let $\{U_i\}_{i\in I}\subseteq\mathcal{T}$ be a family of open sets such that $F\subseteq\bigcup_{i\in I}U_i$. As F is closed we have that $X\setminus F$ is open, using [theorem: 1.68] there exist a k such that $k\notin I$. Define then

$$\{V_i\}_{i\in I\bigcup\{k\}}\subseteq\mathcal{T} \text{ by } V_i=\begin{cases} U_i \text{ if } i\in I\\ X\setminus F \text{ if } i=k\end{cases}$$

Then for $x\in C$ we have either:

- **$x\in F$.** Then $\exists i\in I\subseteq I\bigcup\{k\}$ such that $x\in U_i=V_i$ hence $x\in\bigcup_{i\in I\bigcup\{k\}}V_i$
- **$x\notin F$.** Then $x\in X\setminus F=V_k$ such that $x\in\bigcup_{i\in I\bigcup\{k\}}V_i$

this proves that $C\subseteq\bigcup_{i\in I\bigcup\{k\}}V_i$. As C is compact there exist a finite set $J\subseteq I$ such that $C\subseteq\bigcup_{i\in J}V_i$ which as $F\subseteq C$ gives

$$F\subseteq\bigcup_{i\in J}V_i$$

Let $x\in F$ then there exist a $i\in J$ such that $x\in V_i$. Assume that $i=k$ then $x\in V_k=X\setminus F$ contradicting $x\in F$, hence $i\in J\setminus\{k\}$ and $x\in V_i=U_i$. So $F\subseteq\bigcup_{i\in J\setminus\{k\}}U_i$ which as $J\setminus\{k\}$ is finite proves that F is compact. □

We have seen that for a Hausdorff space a compact set is closed and bounded. We show now that the opposite is true for the real numbers.

Theorem 14.232. (Heine Borel) *Let $a,b\in\mathbb{R}$ then $[a,b]$ is a compact subset of $\mathbb{R}$ [using the norm topology $\mathcal{T}_{||}$ generated by the absolute value $||$].*

Proof. First if $b<a$ then $[a,b]=\varnothing$ which is compact by [example: 14.223] hence we have only to prove the remaining case $a\leqslant b$. Let $\{U_i\}_{i\in I}\subseteq\mathcal{T}$ be family of open sets such that $[a,b]\subseteq\bigcup_{i\in I}U_i$. Define

$$\mathcal{G}=\left\{x\in\mathbb{R}|a\leqslant x \text{ and there exist a finite } J\subseteq I \text{ such that } [a,x]\subseteq\bigcup_{i\in J}U_i\right\}$$

If $x\in\mathcal{G}$ then there exist a finite $J\subseteq I$ such that $[a,x]\subseteq\bigcup_{i\in J}U_i$, let $y\in[a,x]$ then

$$a\leqslant y \text{ and } [a,y]\subseteq[a,x]\subseteq[a,x]\subseteq\bigcup_{i\in J}U_i$$

proving that $y\in\mathcal{G}$. Hence we have proved that

$$\forall x\in\mathcal{G}\ [a,x]\subseteq\mathcal{G} \tag{14.116}$$

As $a\in[a,b]\subseteq\bigcup_{i\in I}U_i$ there exist a $k\in I$ such that $a\in U_k$, by [theorems: 14.34, 14.98] there exist a $\delta\in\mathbb{R}^+$ such that $]a-\delta,a+\delta[\subseteq U_k=\bigcup_{i\in\{k\}}U_i$. Hence $\forall y\in[a,a+\delta[$ we have $y\in U_k=\bigcup_{i\in\{k\}}U_i$ which as $\{k\}$ is finite, $\{k\}\subseteq I$ and $a\leqslant y$ proves that $y\in\mathcal{G}$. So we have

$$\exists\delta\in\mathbb{R}^+ \text{ such that } [a,a+\delta[\subseteq\mathcal{G} \text{ and as } a\in[a,a+\delta[\ \mathcal{G}\neq\varnothing \tag{14.117}$$

For $\mathcal{G}$ we have now two possible cases:

- **$\forall u\in\mathbb{R}\ \exists x\in\mathcal{G}$ such that $u\leqslant x$.** In particular there exist a $x\in\mathcal{G}$ such that $b\leqslant x$ so that by [eq: 14.116] $[a,x]\subseteq\mathcal{G}$ which as $b\in[a,x]$ [because $a\leqslant b\leqslant x$] proves that

  $$b\in\mathcal{G}$$

- **$\exists u\in\mathbb{R}\ \forall x\in\mathcal{G}$ we have $x\leqslant u$.** Then $\mathcal{G}$ is bounded above by u hence, as by [eq: 14.117] $\mathcal{G}\neq\varnothing$, we have by the conditional completeness of $\mathbb{R}$ [see theorem: 10.18] that

  $$s=\sup(\mathcal{G}) \text{ exist}$$

  Assume that $s\leqslant b$. Then as $a+\frac{\delta}{2}\in[a,a+\delta[\subseteq_{\text{[eq: 14.117]}}\mathcal{G}$ it follows that $a<a+\frac{\delta}{2}\leqslant s$ proving that

  $$a<s\leqslant b$$

  hence $s\in[a,b]\subseteq\bigcup_{i\in I}U_i$. So there exist a $l\in I$ such that $s\in U_l$, hence by [theorems: 14.34, 14.98] there exist a $\varepsilon\in\mathbb{R}^+$ such that

  $$s\in]s-\varepsilon,s+\varepsilon[\subseteq U_l\Rightarrow s\in\left[s-\frac{\varepsilon}{2},s+\frac{\varepsilon}{2}\right]\subseteq]s-\varepsilon,s+\varepsilon[\subseteq U_l$$

  As $s-\frac{\varepsilon}{2}<s=\sup(\mathcal{G})$ there exist a $c\in\mathcal{G}$ with

  $$s-\frac{\varepsilon}{2}<c\leqslant s. \tag{14.118}$$

As $c\in\mathcal{G}$ there exist a finite $J\subseteq I$ such that $[a,c]\subseteq\bigcup_{i\in J}U_i$ and $a\leqslant c$. Hence

$$[a,c]\bigcup\left[s-\frac{\varepsilon}{2},s+\frac{\varepsilon}{2}\right]\subseteq\bigcup_{i\in J\bigcup\{l\}}U_i \tag{14.119}$$

If $y\in\left[a,s+\frac{\varepsilon}{2}\right]$ then $a\leqslant y\leqslant s+\frac{\varepsilon}{2}$ and for y we have either:

- $\boldsymbol{y\leqslant c}$**.** Then $y\in[a.c]\subseteq[a,c]\bigcup\left[s-\frac{\varepsilon}{2},s+\frac{\varepsilon}{2}\right]$
- $\boldsymbol{c<y}$**.** Then by [eq: 14.118] $s-\frac{\varepsilon}{2}<c<y\leqslant s+\frac{\varepsilon}{2}$ proving that $y\in\left[s-\frac{\varepsilon}{2},s+\frac{\varepsilon}{2}\right]\subseteq[a,c]\bigcup\left[s-\frac{\varepsilon}{2},s+\frac{\varepsilon}{2}\right]$

proving using [eq: 14.119] that

$$\left[a,s+\frac{\varepsilon}{2}\right]\subseteq\bigcup_{i\in J\bigcup\{l\}}U_i$$

which as $J\bigcup\{l\}$ is finite, $J\bigcup\{l\}\subseteq I$ and $c\in\mathcal{G}\Rightarrow a\leqslant c\leqslant s<s+\frac{\varepsilon}{2}$ proves that $s+\frac{\varepsilon}{2}\in\mathcal{G}$. Hence as $s=\sup(\mathcal{G})$ we must have that $s+\frac{\varepsilon}{2}\leqslant s$ a contradiction. Hence the assumption is wrong and we must have that

$$b<s$$

As $s=\sup(\mathcal{G})$ there exist a $d\in\mathcal{G}$ such that $b<d\leqslant s$, using [eq: 14.116] we have then that $[a,d]\in\mathcal{G}$. From $a\leqslant b<d$ it follows that $b\in[a,d]$ proving that

$$b\in\mathcal{G}$$

So in all cases we have $b\in\mathcal{G}$ which means by the definition of $\mathcal{G}$ that there exist a finite $J\subseteq I$ such that $[a,b]\subseteq\bigcup_{i\in J}U_i$. Hence $[a,b]$ is compact. □

Corollary 14.233. *If $C\subseteq\mathbb{R}$ then using the topology $\mathcal{T}_{||}$*

$$C \textit{ is compact} \Leftrightarrow C \textit{ is closed and bounded}$$

Proof.

- **⇒.** Using [theorem: 14.227] C is bounded, further by [theorem: 14.201] $\mathcal{T}_{||_{\max}}$ is Hausdorff, so using [theorem: 14.230] C is closed.
- **⇐.** As C is bounded there exist a $M\in\mathbb{R}^+$ such that $\forall x\in C$ $|x|\leqslant M$ so we have $x\leqslant M$ and $-x\leqslant M\Rightarrow -M\leqslant x$ proving that $C\subseteq[-M,M]$. By [theorem: 14.232] $[-M,M]$ is compact which as C is closed proves by [theorem: 14.231] that C is compact. □

Theorem 14.234. (Extrem Value Theorem) *Let $\langle X,\mathcal{T}\rangle$ be a topological space, $f:X\rightarrow\mathbb{R}$ a continuous function [using the topology $\mathcal{T}_{||}$ on $\mathbb{R}$] and $C\subseteq X$ a compact subset of X then there exist a $M\in\mathbb{R}^+$ such that*

$$\forall x\in C \textit{ we have } |f(x)|\leqslant M$$

Proof. As C is compact and f continuous we have by [theorem: 14.229] that $f(C)$ is a compact subset of $\mathbb{R}$, hence using [corollary: 14.233] $f(C)$ is bounded. So by [theorem:14.228] there exist a $M\in\mathbb{R}^+$ such that $\forall y\in f(C)$ we have $|f(y)|\leqslant M$, hence if $x\in C\Rightarrow f(x)\in f(C)$ we have that $|f(x)|\leqslant M$. □

Theorem 14.235. (Extreme Value Theorem) *Let $\langle X,\mathcal{T}\rangle$ be a topological space, $\varnothing\neq C\subseteq X$ a non empty compact subset of X and $f:X\rightarrow\mathbb{R}$ a continuous function [using the topology $\mathcal{T}_{||}$ on $\mathbb{R}$] then there exists $m,M\in C$ such that*

$$\forall x\in C \textit{ we have } f(m)\leqslant f(x)\leqslant f(M)$$

Proof. As C is compact and f continuous we have by [theorem: 14.229] that

$$f(C) \text{ is compact}$$

As by [theorem: 14.201] $\langle\mathbb{R},\mathcal{T}_{||}\rangle$ is Hausdorff and $f(C)$ is compact it follows from [theorem: 14.230] that

$$f(C) \text{ is closed}$$

Using [theorem:14.228] there exist a $N\in\mathbb{R}^+$ such that $\forall x\in f(C)$ we have $|x|\leqslant N\Rightarrow -N\leqslant x\leqslant N$ proving that $f(C)$ is bounded above and below. As further $f(C)\neq\varnothing$ we have using the fact that $\mathbb{R}$ is conditional complete [see theorem: 10.18] that $\sup(f(C))$ and $\inf(f(C))$ exist. As $f(C)$ is closed and non empty we have by [theorem: 14.105] that $\sup(f(C))\in f(C)$ and $\inf(f(C))\in f(C)$. Hence there exists $m,M\in C$ such that $\sup(f(C))=f(M)$ and $\inf(f(C))=f(m)$. So we have $\forall x\in C$ that

$$f(m)=\inf(f(C))\leqslant f(x)\leqslant\sup(C)=f(M)$$

□

Theorem 14.236. *Let $\langle X,\mathcal{T}\rangle$ be a **compact Hausdorff space** then it is **regular** and **normal***

Proof. Let A be a closed set and $x\in X\setminus A$. As A is closed we have we have by [theorem: 14.231] that

$$A \text{ is compact}$$

then by the Hausdorff condition we have $\forall a\in A$ that there exists $U_a,V_a\in\mathcal{T}$ such that $a\in U_a$, $x\subseteq V_a$ and $U_a\bigcap V_a=\varnothing$. So $A\subseteq\bigcup_{a\in A}U_a$ which as A is compact means that there exist a finite $I\subseteq A$ such that $A\subseteq\bigcup_{a\in I}U_a$. Take now $U=\bigcup_{a\in I}U_a\in\mathcal{T}$ and $V=\bigcap_{a\in I}V_a\in\mathcal{T}$ [as I is finite] then we have $A\subseteq U$, $x\in V$ and

$$U\bigcap V=\left(\bigcup_{a\in I}U_a\right)\bigcap\left(\bigcap_{b\in I}V_b\right)=\bigcup_{a\in I}\left(U_a\bigcap\left(\bigcap_{b\in I}V_b\right)\right)\underset{U_a\bigcap\left(\bigcap_{b\in I}V_b\right)\subseteq U_a\bigcap V_a=\varnothing}{=}\varnothing$$

proving that

$$\langle X,\mathcal{T}\rangle \text{ is regular} \tag{14.120}$$

Let A,B be closed sets such that $A\bigcap B=\varnothing$. Then $\forall a\in A$ we have as $a\in X\setminus B$ by [eq: 14.120] that there exists $U_a,V_a\in\mathcal{T}$ with $a\in U_a$, $B\subseteq V_a$ and $U_a\bigcap V_a=\varnothing$. As by [theorem: 14.231] A is compact and $A\subseteq\bigcup_{a\in A}U_a$ there exist a finite $I\subseteq A$ such that $A\subseteq\bigcup_{a\in I}U_a$. So if we take $U=\bigcup_{a\in I}U_a\in\mathcal{T}$ and $V=\bigcap_{a\in I}V_a\in\mathcal{T}$ [as I is finite] then we have $A\subseteq U$ and $B\subseteq V$ and

$$U\bigcap V=\left(\bigcup_{a\in I}U_a\right)\bigcap\left(\bigcap_{b\in I}V_b\right)=\bigcup_{a\in I}\left(U_a\bigcap\left(\bigcap_{b\in I}V_b\right)\right)\underset{U_a\bigcap\left(\bigcap_{b\in I}V_b\right)\subseteq U_a\bigcap V_a=\varnothing}{=}\varnothing$$

proving that

$$\langle X,\mathcal{T}\rangle \text{ is normal}$$

□

A weaker version of compactness is sequential compactness.

Definition 14.237. (Limit Point Compact Set) *Let $\langle X,\mathcal{T}\rangle$ be a topological space then a subset $C\subseteq X$ is **limit point compact** if every **infinite** $A\subseteq K$ has a limit point [see definition: 14.26]*

Theorem 14.238. *Every compact subspace of a topological space $\langle X,\mathcal{T}\rangle$ is limit point compact.*

Proof. We proceed by contradiction. Let C be a compact subset and let $A\subseteq C$ be a infinite subset of C that does not have a limit point. Then by the definition of a limit point [see definition: 14.26] we have $\forall x\in X$ that there exist a $U_x\in\mathcal{T}$ with $x\in U_x$ so that $U_x\bigcap(A\setminus\{x\})=\varnothing$. Let $y\in U_x\bigcap A$ then if $y\neq x$ we have $y\in(U_x\bigcap A)\setminus\{x\}=U_x\bigcap(A\setminus\{x\})=\varnothing$ a contradiction, hence we have that $y=x\in\{x\}$, proving that

$$U_x\bigcap A\subseteq\{x\} \tag{14.121}$$

As $C\subseteq\bigcup_{x\in C}U_x$ we have, as C is compact, that there exist a finite $F\subseteq C$ such that $C\subseteq\bigcup_{x\in F}U_x$. Further from $A\subseteq C$ we have

$$A=C\bigcap A\subseteq\left(\bigcup_{x\in F}U_x\right)\bigcap A=\bigcup_{x\in F}(U_x\bigcap A)\subseteq_{[\text{eq: }14.121]}\bigcup_{x\in F}\{x\}=F$$

proving that A is finite contradicting the fact that we have taken A to be infinite. So we must have that every infinite subset of C has a limit point. □

Definition 14.239. (Compact Class) *Let X be a set then $C\subseteq\mathcal{P}(X)$ a set of subsets of X is a **compact class** if for any sequence $\{K_i\}_{i\in\mathbb{N}}\subseteq C$ with $\bigcap_{i\in\mathbb{N}}K_i=\varnothing$ there exist a $N\in\mathbb{N}$ such that $\bigcap_{i\in\{1,\ldots,N\}}K_i=\varnothing$.*

Theorem 14.240. *Let $\langle X,\mathcal{T}\rangle$ be a Hausdorff topological space and C a set of compact subsets of X then C is a compact class.*

Proof. We prove this by contradiction. Let $\{K_i\}_{i\in\mathbb{N}}\subseteq C$ be a sequence of subsets of C such that $\bigcap_{i\in\mathbb{N}}K_i=\varnothing$. Assume that $\forall n\in\mathbb{N}$ we have $\bigcap_{i\in\{1,\ldots,n\}}K_i\neq\varnothing$. Define

$$\{E_i\}_{i\in\mathbb{N}}\subseteq\mathcal{P}(X)\text{ by }E_i=\bigcap_{j\in\{1,\ldots,n\}}K_j \tag{14.122}$$

so that

$$E_n\subseteq E_m\text{ for }n,m\in\mathbb{N}\text{ with }m\leqslant n \tag{14.123}$$

Let $j\in\mathbb{N}$ then K_i is compact hence, as $\mathcal{T}$ is Hausdorff it follows from [theorem: 14.230] that K_j is closed, hence for $i\in\mathbb{N}$ we have by [theorem: 14.21] that $E_i=\bigcap_{j\in\{1,\ldots,n\}} K_j$ is closed, hence $X\setminus E_i$ is open proving that

$$\{X\setminus E_i\}_{i\in\mathbb{N}} \text{ is a sequence of open sets in other words } \{X\setminus E_i\}_{i\in\mathbb{N}}\subseteq\mathcal{T}$$

Further as $\forall i\in\mathbb{N}$ we have that $E_i=\bigcap_{j\in\{1,\ldots,i\}} K_i\subseteq K_i$ it follows from [theorem: 1.26] that $X\setminus K_i\subseteq X\setminus E_i$. Let $x\in K_1$ then as $\bigcap_{i\in\mathbb{N}} K_i=\varnothing$ there exist a $n\in\mathbb{N}$ such that $x\notin K_n$ hence $x\in X\setminus K_n\subseteq X\setminus E_n$ proving that $K_1\subseteq\bigcup_{n\in\mathbb{N}}(X\setminus E_n)$. As K_1 is compact there exist a finite $I\subseteq\mathbb{N}$ such that

$$K_1\subseteq\bigcup_{i\in I}(X\setminus E_n) \tag{14.124}$$

Take $m=\max(I)$ then as $\forall i\in I$ we have $i\leqslant m$ so that by [eq: 14.123] $E_m\subseteq E_i\Rightarrow X\setminus E_i\subseteq X\setminus E_m$ which using [eq: 14.124] proves that

$$K_1\subseteq X\setminus E_m \tag{14.125}$$

Hence

$$E_m \underset{\text{[eq: 14.122]}}{=} \bigcap_{i\in\{1,\ldots,m\}} K_i\subseteq K_1\subseteq_{\text{[eq: 14.125]}} X\setminus E_m$$

from which it follows that $E_m=\varnothing$ [for if $y\in E_m$ then $y\in X\setminus E_m$ so that $y\notin E_m$]. As $E_m=\bigcap_{i\in\{1,\ldots,m\}} K_i$ we have $\varnothing=\bigcap_{i\in\{1,\ldots,m\}} K_i$ which contradicts the assumption that $\forall n\in\mathbb{N}$ we have $\bigcap_{i\in\{1,\ldots,n\}} K_i\neq\varnothing$. So the assumption is wrong and we must have that there exist a $N\in\mathbb{N}$ such that $\bigcap_{i\in\{1,\ldots,N\}} K_i=\varnothing$ from which it follows that C is a compact class □

Theorem 14.241. *Let $\langle X,d_X\rangle$, $\langle Y,d_Y\rangle$, K a non empty compact subset of X and $f:K\rightarrow Y$ a continuous function then f is uniform continuous [see definition: 14.150].*

Proof. Let $\varepsilon\in\mathbb{R}^+$. Take $x\in K$ then by the continuity of f there exist by [theorem: 14.126] a $\delta_x\in\mathbb{R}^+$ [depending on x] such that for every $y\in K$ with $(d_X)_{|K\times K}(x,y)<\delta_x$ we have $d_Y(f(x),f(y))<\frac{\varepsilon}{2}$. As $d_X(x,y)=(d_X)_{|K\times K}(x,y)<\delta_x$ we have $x\in B_{d_X}(x,\delta_x)$ so that

$$K\subseteq\bigcup_{x\in K} B_{d_X}\left(x,\frac{\delta_x}{2}\right), \tag{14.126}$$

hence, as K is compact, there exist a finite $I\subseteq K$ such that $K\subseteq\bigcup_{x\in I} B_{d_X}\left(x,\frac{\delta_x}{2}\right)$. As $K\neq\varnothing$ and $\left\{\frac{\delta_x}{2}|x\in I\right\}$ is finite [see theorem: 6.45] it follows that

$$\delta=\min\left(\left\{\frac{\delta_x}{2}|x\in I\right\}\right)\text{exist} \tag{14.127}$$

Let now $x,y\in K$ be such that $(d_X)_{|K\times K}(x,y)<\delta\Rightarrow d_X(x,y)<\delta$ then, as $x\in K$, there exist by [eq: 14.126] a $z\in I$ such that

$$x\in B_{d_X}\left(z,\frac{\delta_z}{2}\right)\text{ or }(d_X)_{|K\times K}(x,z)=d_X(x,z)<\frac{\delta_z}{2} \tag{14.128}$$

Then

$$(d_X)_{|K\times K}(z,y)=d_X(z,y)\leqslant d_X(z,x)+d_X(x,y)<\frac{\delta_z}{2}+\delta\leqslant_{\text{[eq: 14.127]}}\frac{\delta_z}{2}+\frac{\delta_z}{2}=\delta_z$$

so that

$$d_Y(f(z),f(y))<\frac{\varepsilon}{2} \tag{14.129}$$

As $(d_X)_{|K\times K}(x,z)<_{\text{[eq: 14.128]}}\frac{\delta_z}{2}$ we have

$$d_Y(f(x),f(z))<\frac{\varepsilon}{2} \tag{14.130}$$

Finally we have

$$d_Y(f(x),f(y))\leqslant d_Y(f(x),f(z))+d_Y(f(z),f(y))<_{\text{[eqs: 14.129, 14.130]}}\frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon$$

proving uniform continuity. □

14.8.1 Product of Compact sets

The purpose of this section is to prove that the product of compact sets is compact, this is know as Tychonoff's theorem. The best way to do this is to develop the framework of filter bases.

14.8.1.1 Filter bases

Definition 14.242. (Filter Base) *Let X be a set then a a non empty family $\mathfrak{U}=\{A_\alpha\}_{\alpha\in I}\subseteq\mathcal{P}(X)$ of a subsets of X is a **filter base in** X if*

1. *$\forall\alpha\in I$ we have $A_\alpha\neq\varnothing$*
2. *$\forall\alpha,\beta\in I$ there exist a $\gamma\in I$ such that $A_\gamma\subseteq A_\alpha\bigcap B_\beta$*

Using mathematical induction we can extend (2) of the above definition.

Theorem 14.243. *Let X be a set and $\mathfrak{U}=\{A_\alpha\}_{\alpha\in I}\subseteq\mathcal{P}(X)$ a filter base then for a finite non empty $J\subseteq I$ there exist a $\gamma\in I$ such that $A_\gamma\subseteq\bigcap_{\alpha\in J}A_\alpha$*

Proof. Define

$$S=\left\{n\in\mathbb{N}|\text{If } J\subseteq I \text{ with } \operatorname{card}(J)=n \text{ then there exist a } \gamma\in I \text{ such that } A_\gamma\subseteq\bigcap_{\alpha\in J}A_\alpha\right\}$$

then we have:

$1\in S$. If $J\subseteq I$ with $\operatorname{card}(J)=1$ then there exist a $\gamma\in I$ such that $J=\{\gamma\}$. As $A_\gamma\subseteq A_\gamma=\bigcup_{\alpha\in J}A_\alpha$ it follows that $1\in S$.

$n\in S\Rightarrow n+1\in S$. Let $J\subseteq I$ with $\operatorname{card}(J)=n+1$ then there exist a bijection

$$\beta:\{1,\ldots,n+1\}\to J$$

By [theorem: 2.89] $\beta_{|\{1,\ldots,n\}}:\{1,\ldots,n\}\to J\setminus\{\beta(n+1)\}$ is a bijection so that $\operatorname{card}(J\setminus\beta(n+1))=n$. As $n\in S$ there exist a $\zeta\in J\setminus\{\beta(n+1)\}$ such that

$$A_\zeta\subseteq\bigcap_{\alpha\in J\setminus\{\beta(n+1)\}}A_\alpha$$

By definition of a filter base there exist a $\gamma\in I$ such that

$$A_\gamma\subseteq A_\zeta\bigcap A_{\beta(n+1)}\subseteq\left(\bigcap_{\alpha\in J\setminus\{b(n+1)\}}A_\gamma\right)\bigcap A_{\beta(n+1)}=\bigcap_{\alpha\in J}A_\alpha$$

proving that $n+1\in S$.

By mathematical induction it follows that $S=\mathbb{N}$. Hence if $J\subseteq I$ is finite we have as $\operatorname{card}(J)\in\mathbb{N}$ that there exist a $\gamma\in I$ such that $A_\gamma\subseteq\bigcap_{\alpha\in I}A_\alpha$. □

Example 14.244. Let X be a set, $\varnothing\neq A\subseteq X$ then $\mathfrak{U}=\{A_\alpha\}_{\alpha\in\{A\}}\subseteq\mathcal{P}(X)$ defined by $A_\alpha=\alpha$ is a filter base in X

Proof. First as $\{A\}\neq\oslash$ $\{A_\alpha\}_{\alpha\in\{A\}}$ is a non empty family of subsets in X. Further

1. $\forall\alpha\in\{A\}$ we have $A_\alpha=\alpha=A\neq\varnothing$.
2. If $\alpha,\beta\in\{A\}$ then $\alpha=\beta=A$ so for $\gamma=A$ we have $A_\gamma=A\subseteq A\bigcap A=A_\alpha\bigcap A_\beta$ □

Definition 14.245. *Let $\langle X,\mathcal{T}\rangle$ be a topological space, $x\in X$ then $\mathfrak{U}(x)=\{U\}_{U\in\{U\in\mathcal{T}|x\in U\}}\subseteq\mathcal{P}(X)$ [a family of open sets containing x] is called a **neighborhood filter base of** x.*

The name used in the above definition suggest that $\mathfrak{U}(x)$ is a filter base in X which is proved in the following theorem.

Theorem 14.246. *Let $\langle X,\mathcal{T}\rangle$ be a topological space then $\mathfrak{U}(x)$ is a filter base in X*

Proof. Let $x\in X$ then as $X\in\mathcal{T}$ we have that $X\in\{U\in\mathcal{T}|x\in U\}$ so that $\mathfrak{U}(x)$ is a non empty family. Further we have

1. If $U\in\{U\in\mathcal{T}|x\in U\}$ then $x\in U$ so that $U\neq\varnothing$,
2. If $U,V\in\{U\in\mathcal{T}|x\in U\}$ then $x\in U\bigcap V\in\mathcal{T}$ so that for $W=U\bigcap V\in\{U\in\mathcal{T}|x\in U\}$ we have

$$W\subseteq U\bigcap V$$

proving by [definition: 14.242] that $\mathfrak{U}(x)$ is a filter base in X. □

Theorem 14.247. *Let X be a set and let $\mathfrak{U}=\{A_\alpha\}_{\alpha\in I}\subseteq\mathcal{P}(X)$ and $\mathfrak{W}=\{B_\alpha\}_{\alpha\in J}\subseteq\mathcal{P}(X)$ filter bases then we have:*

1. *$\mathfrak{U}\bigcup\mathfrak{W}$ is a filter base in X where $\mathfrak{U}\bigcup\mathfrak{W}=\{A_\alpha\bigcup B_\beta\}_{(\alpha,\beta)\in I\times J}$*

2. *If* $\forall(\alpha,\beta)\in I\times J$ $A_\alpha\bigcap B_\beta\neq\varnothing$ *then* $\mathfrak{A}\bigcap\mathfrak{W}$ *is a filter base where* $\mathfrak{A}\bigcap\mathfrak{W}=\{A_\alpha\bigcap B_\beta\}_{(\alpha,\beta)\in I\times J}$

Proof. First as $I\neq\varnothing$ and $J\neq\varnothing$ we have that $I\times J\neq\varnothing$, further we have:

$\mathfrak{U}\bigcup\mathfrak{W}$ is a filter base. Because

1. $\forall(\alpha,\beta)\in I\times J$ we have $A_\alpha\neq\varnothing\wedge B_\alpha\neq\varnothing$ so that $A_\alpha\bigcup B_\alpha\neq\varnothing$
2. $\forall(\alpha_1,\beta_1),(\alpha_2,\beta_2)\in I\times J$ there exists a $\gamma_1\in I$ and a $\gamma_2\in J$ such that $A_{\gamma_1}\subseteq A_{\alpha_1}\bigcap A_{\alpha_2}$ and $B_{\gamma_2}\subseteq B_{\beta_1}\bigcap B_{\beta_2}$, hence for $(\gamma_1,\gamma_2)\in I\times J$ we have

$$A_{\gamma_1}\bigcup B_{\gamma_2}\subseteq(A_{\alpha_1}\bigcap A_{\alpha_2})\bigcup(B_{\beta_1}\bigcap B_{\beta_2}) \tag{14.131}$$

Further

$$\begin{aligned} A_{\alpha_1}\bigcap A_{\alpha_2} &\subseteq A_{\alpha_1}\subseteq A_{\alpha_1}\bigcup B_{\beta_1}\\ A_{\alpha_1}\bigcap A_{\alpha_2} &\subseteq A_{a_2}\subseteq A_{\alpha_2}\bigcup B_{\beta_2}\end{aligned}$$

so that

$$A_{\alpha_1}\bigcap A_{\alpha_2}\subseteq(A_{\alpha_1}\bigcup B_{\beta_1})\bigcap(A_{\alpha_2}\bigcup B_{\beta_2}) \tag{14.132}$$

and

$$\begin{aligned} B_{\beta_1}\bigcap B_{\beta_2} &\subseteq B_{\beta_1}\subseteq A_{\alpha_1}\bigcup B_{\beta_1}\\ B_{\beta_1}\bigcap B_{\beta_2} &\subseteq B_{\beta_2}\subseteq A_{\alpha_2}\bigcup B_{\beta_2}\end{aligned}$$

so that

$$B_{\beta_1}\bigcap B_{\beta_2}\subseteq(A_{\alpha_1}\bigcup B_{\beta_1})\bigcap(A_{\alpha_2}\bigcup B_{\beta_2}) \tag{14.133}$$

Combining [eqs: 14.131, 14.132 and 14.133] gives finally

$$A_{\gamma_1}\bigcup B_{\gamma_2}\subseteq(A_{\alpha_1}\bigcup B_{\beta_1})\bigcap(A_{\alpha_2}\bigcup B_{\beta_2})$$

$\mathfrak{U}\bigcap\mathfrak{W}$ is a filter base. Because

1. $\forall(\alpha,\beta)\in I\times J$ we have $A_\alpha\bigcap B_\beta\neq\varnothing$ by the definition of $\mathfrak{U}\bigcap\mathfrak{W}$
2. $\forall(\alpha_1,\beta_1),(\alpha_2,\beta_2)\in I\times J$ there exists a $\gamma_1\in I$ and a $\gamma_2\in J$ such that $A_{\gamma_1}\subseteq A_{\alpha_1}\bigcap A_{\alpha_2}$ and $B_{\gamma_2}\subseteq B_{\beta_1}\bigcap B_{\beta_2}$, hence for $(\gamma_1,\gamma_2)\in I\times J$ we have

$$A_{\gamma_1}\bigcap B_{\gamma_2}\subseteq(A_{\alpha_1}\bigcap A_{\alpha_2})\bigcap(B_{\beta_1}\bigcap B_{\beta_2})=(A_{\alpha_1}\bigcap B_{\beta_1})\bigcap(A_{\alpha_2}\bigcap B_{\beta_2})$$ □

Theorem 14.248. *Let X be a set and $\mathfrak{U}=\{A_\alpha\}_{\alpha\in I}\subseteq\mathcal{P}(X)$ a filter base in X then it has the* ***finite intersection property****, meaning that every* ***finite*** *$J\subseteq I$ we have $\bigcap_{\alpha\in J}A_\alpha\neq\varnothing$*

Proof. Using [theorem: 14.243] there exist a $\gamma\in I$ such that $A_\gamma\subseteq\bigcap_{\alpha\in J}A_\alpha$, as by the definition of a filter base $A_\gamma\neq\varnothing$ it follows that $\bigcap_{\alpha\in J}A_\alpha\neq\varnothing$. □

Definition 14.249. *Let $\langle X,\mathcal{T}\rangle$ be a topological space, $x\in X$ and $\mathfrak{U}=\{A_\alpha\}_{\alpha\in I}\subseteq\mathcal{P}(X)$ a filter base in X then we say that:*

1. *$\mathfrak{U}$* ***converges to*** *x noted as $\mathfrak{U}\rightarrow x$ if $\forall U\in\mathcal{T}$ with $x\in U$ there exist a $\alpha\in I$ such that $A_\alpha\subseteq U$*
2. *$\mathfrak{U}$* ***accumulates at*** *x noted as $\mathfrak{U}\succ x$ if $\forall U\in\mathcal{T}$ with $x\in U$ we have that $\forall\alpha\in I$ $A_\alpha\bigcap U\neq\varnothing$*

Theorem 14.250. *Let $\langle X,\mathcal{T}\rangle$ be a topological space then*

$\mathcal{T}$ is Hausdorff $\Leftrightarrow$ each filter base that converges converges to exactly one point

Proof.

$\Rightarrow$. Assume that $\mathcal{T}$ is Hausdorff and $\mathfrak{U}=\{A_\alpha\}_{\alpha\in I}\subseteq\mathcal{P}(X)$ be such that for $x,y\in X$ $\mathfrak{U}\rightarrow x$ and $\mathfrak{U}\rightarrow y$. If $x\neq y$ then as $\mathcal{T}$ is Hausdorff there exist $U,V\in T$ such that $x\in U$, $y\in V$ and $U\bigcap V=\varnothing$. From $\mathfrak{U}\rightarrow x$ and $\mathfrak{U}\rightarrow y$ it follows that there exists $\alpha,\beta\in I$ such that $A_\alpha\subseteq U$ and $A_\beta\subseteq V$. Then there exist a $\gamma\in I$ such that $\varnothing\neq A_\gamma\subseteq A_\alpha\bigcap A_\beta\subseteq U\bigcap V$, hence $\varnothing\neq U\bigcap V$ contradicting $U\bigcap V=\varnothing$. So we must have that $x=y$ proving uniqueness.

⇐. Assume that $\mathcal{T}$ is not Hausdorff then $\exists x, y \in X$ with $x \neq y$ such that $\forall U, V \in \mathcal{T}$ with $x \in U \wedge y \in V$ we have $U \bigcap V \neq \varnothing$. We construct now a filter base that converges to x and y. By [theorem: 14.247] and [definition: 14.245] we have that $\mathfrak{W} = \mathfrak{U}(x) \bigcap \mathfrak{U}(y)$ is a filter base in X where

$$\mathfrak{U}(x) = \{U\}_{U \in \{U \in \mathcal{T} | x \in U\}} \text{ and } \mathfrak{U}(y) = \{U\}_{U \in \{U \in \mathcal{T}\}}$$

If $U \in \mathcal{T}$ with $x \in U$ then as $y \in X \in \mathcal{T}$ we have that $(U, X) \in \{U \in \mathcal{T} | x \in U\} \times \{V \in \mathcal{T} | y \in V\}$ so that $U \bigcap X = U \subseteq U$ proving that

$$\mathfrak{W} \rightarrow x$$

Further if $V \in \mathcal{T}$ with $y \in V$ then as $x \in X \in \mathcal{T}$ we have that $(X, V) \in \{U \in \mathcal{T} | x \in U\} \times \{V \in \mathcal{T} | y \in V\}$ so that $X \bigcap V = V \subseteq V$ proving that

$$\mathfrak{W} \rightarrow y$$

By the hypothesis it follows then that $x = y$ contradicting $x \neq y$. So the assumption that $\mathcal{T}$ is not Hausdorff is wrong proving that $\mathcal{T}$ is Hausdorff. □

Theorem 14.251. *Let $\langle X, \mathcal{T}\rangle$ be a topological space, $x \in X$ and $\mathfrak{U} = \{A_\alpha\}_{\alpha \in I} \subseteq \mathcal{P}(X)$ a filter base in X then*

$$\mathfrak{U} \succ x \Leftrightarrow x \in \bigcap_{\alpha \in I} \overline{A_\alpha}$$

Proof.

⇒. If $\mathfrak{U} \succ x$ then $\forall U \in \mathcal{T}$ with $x \in U$ we have $\forall \alpha \in I$ $A_\alpha \bigcap U \neq \varnothing$, so that by [theorem: 14.28] $\forall \alpha \in I$ $x \in \overline{A_\alpha}$ hence $x \in \bigcap_{\alpha \in I} \overline{A_\alpha}$.

⇐. Let $U \in \mathcal{T}$ with $x \in U$ then as $x \in \bigcap_{\alpha \in I} \overline{A_i}$ we have that $\forall \alpha \in I$ $x \in \overline{A_\alpha}$ so that by [theorem: 14.28] $A_\alpha \bigcap U \neq \varnothing$. Hence $\mathfrak{U} \succ x$. □

Example 14.252. Let $\langle X . \mathcal{T}\rangle$ be a topological space and $x \in X$ then $\mathfrak{U}(x) \rightarrow x$

Proof. By [definition: 14.245] $\mathfrak{U}(x) = \{U\}_{U \in \{U \in \mathcal{T} | x \in U\}}$. So if $U \in \mathcal{T}$ with $x \in U$ then $U \in \{U \in \mathcal{T} | x \in U\}$ and $U \subseteq U$ proving that $\mathfrak{U}(x) \rightarrow x$. □

Definition 14.253. *Let X be a set and $\mathfrak{U} = \{A_\alpha\}_{\alpha \in I} \subseteq \mathcal{P}(X)$, $\mathfrak{W} = \{B_\alpha\}_{\alpha \in J} \subseteq \mathcal{P}(X)$ two filter bases in X then $\mathfrak{W}$ is **subordinate** to $\mathfrak{U}$ noted as $\mathfrak{W} \gg \mathfrak{U}$ or $\mathfrak{U} \ll \mathfrak{W}$ iff*

$$\forall \alpha \in I \text{ we have that } \exists \beta \in J \text{ such that } B_\beta \subseteq A_\alpha$$

Definition 14.254. *Let X be a set and $\mathfrak{U} = \{A_\alpha\}_{\alpha \in I} \subseteq \mathcal{P}(X)$, $\mathfrak{W} = \{B_\alpha\}_{\alpha \in J} \subseteq \mathcal{P}(X)$ two filter bases in X then $\mathfrak{U} \sqsubseteq \mathfrak{W}$ iff*

$$\forall \alpha \in I \text{ there exist a } \beta \in J \text{ such that } A_\alpha = B_\beta$$

Theorem 14.255. *Let X be a set then we have:*

1. *If $\mathfrak{U} = \{A_\alpha\}_{\alpha \in I} \subseteq \mathcal{P}(X)$, $\mathfrak{W} = \{B_\alpha\}_{\alpha \in J} \subseteq \mathcal{P}(X)$ are two filter bases in X such that $\mathfrak{U} \sqsubseteq \mathfrak{W}$ then $\mathfrak{W} \gg \mathfrak{U}$*
2. *If $\mathfrak{U} = \{A_\alpha\}_{\alpha \in I} \subseteq \mathcal{P}(X)$, $\mathfrak{W} = \{B_\alpha\}_{\alpha \in J} \subseteq \mathcal{P}(X)$ are two filter bases in X such that $\mathfrak{W} \gg \mathfrak{U}$ then $\forall \beta \in J$ and $\forall \alpha \in I$ we have that $A_\alpha \bigcap B_\beta \neq \varnothing$*
3. *If $\langle X, \mathcal{T}\rangle$ is a topological space, $x \in X$ and $\mathfrak{U} = \{A_\alpha\}_{\alpha \in I} \subseteq \mathcal{P}(X)$ a filter base in X*
$$\mathfrak{U} \rightarrow x \Leftrightarrow \mathfrak{U} \geqslant \mathfrak{U}(x)$$
4. *If $\mathfrak{U}_1 = \{A_\alpha\}_{\alpha \in I} \subseteq \mathcal{P}(X)$, $\mathfrak{U}_2 = \{B_\alpha\}_{\sigma \in J} \subseteq \mathcal{P}(X)$ and $\mathfrak{U}_3 = \{C_\alpha\}_{\alpha \in K} \subseteq \mathcal{P}(X)$ are filter bases in X such that $\mathfrak{U}_1 \gg \mathfrak{U}_2$ and $\mathfrak{U}_2 \gg \mathfrak{U}_3$ then $\mathfrak{U}_1 \gg \mathfrak{U}_3$*
5. *If $\mathfrak{U}$ is a filter base in X then $\mathfrak{U} \gg \mathfrak{U}$*

Proof.

1. Let $\alpha \in I$ then as $\mathfrak{U} \sqsubseteq \mathfrak{W}$ there exist a $\beta \in J$ such that $A_\alpha = B_\beta$ so that $B_\beta \subseteq A_\alpha$ proving that $\mathfrak{W} \gg \mathfrak{U}$.
2. We proceed by contradiction. So assume that $\exists \beta \in J$ and $\exists \alpha \in I$ such that $A_\alpha \bigcap B_\beta = \varnothing$. As $\mathfrak{W} \gg \mathfrak{U}$ there exist a $\gamma \in J$ such that $B_\gamma \subseteq A_\alpha$ and by the definition of a filter base there exist a $\delta \in J$ such that $\varnothing \neq B_\delta \subseteq B_\gamma \bigcap B_\beta \subseteq A_\alpha \bigcap B_\beta$ proving that $A_\alpha \bigcap B_\beta \neq \varnothing$ a contradiction. Hence the assumption is wrong and we have $\forall \beta \in J$ and $\forall \alpha \in I$ $A_\alpha \bigcap B_\beta \neq \varnothing$.

3. By definition $\mathfrak{U}(x) = \{U\}_{U \in \{U \in \mathcal{T} | x \in U\}} \subseteq \mathcal{P}(X)$ and

$$\begin{aligned} \mathfrak{U} \rightarrow x &\Leftrightarrow \forall U \in \mathcal{T} \text{ with } x \in U\ \exists \alpha \in I \text{ such that } A_\alpha \subseteq U \\ &\Leftrightarrow \forall U \in \{U \in \mathcal{T} | x \in U\}\ \exists \alpha \in I \text{ such that } A_\alpha \subseteq U \\ &\Leftrightarrow \mathfrak{U} \gg \mathfrak{U}(x) \end{aligned}$$

4. Let $\gamma \in K$ then as $\mathfrak{U}_2 \geqslant \mathfrak{U}_3$ there exist a $\beta \in J$ such that $B_\beta \subseteq C_\gamma$, as $\mathfrak{U}_1 \gg \mathfrak{U}_2$ there exist a $\alpha \in I$ such that $A_\alpha \subseteq B_\beta$. Hence $A_\alpha \subseteq C_\gamma$ proving that $\mathfrak{U}_1 \geqslant \mathfrak{U}_3$.

5. If $\mathfrak{U} = \{A_i\}_{i \in I} \subseteq \mathcal{P}(X)$ is a filter base in X then $\forall i \in I$ we have $A_i = A_i \Rightarrow A_i \subseteq A_i$, hence $\mathfrak{U} \gg \mathfrak{U}$. □

Theorem 14.256. *Let $\langle X, \mathcal{T}\rangle$ be a topological space and $\mathfrak{U} = \{A_\alpha\}_{\alpha \in I} \subseteq \mathcal{P}(X)$, $\mathfrak{W} = \{B_\alpha\}_{\alpha \in J} \subseteq \mathcal{P}(X)$ be filter bases in X then we have for $x, y \in X$ that*

1. *If $\mathfrak{U} \rightarrow x$ then $\mathfrak{U} \succ x$*
2. *If X is Hausdorff, $\mathfrak{U} \rightarrow x$ and $\mathcal{U} \succ y$ then $x = y$*
3. *If $\mathfrak{W} \gg \mathfrak{U}$ then we have:*
   a. *If $\mathfrak{U} \rightarrow x$ then $\mathfrak{W} \rightarrow x$*
   b. *If $\mathfrak{W} \succ x$ then $\mathfrak{U} \succ x$*

Proof.

1. Let $U \in \mathcal{T}$ with $x \in U$ then as $\mathfrak{U} \rightarrow x$ there exist a $\alpha \in I$ such that $A_\alpha \subseteq U$. Given $\beta \in I$ then as there exist a $\gamma \in I$ with $\varnothing \neq A_\gamma \subseteq A_\alpha \bigcap A_\beta \subseteq U \bigcap A_\beta \Rightarrow U \bigcap A_\beta \neq \varnothing$ proving that $\mathfrak{U} \succ x$.

2. Assume that $x \neq y$ then as X is Hausdorff there exists $U, V \in \mathcal{T}$ with $x \in U \wedge y \in V$ such that $\bigcup \bigcap V = \varnothing$. As $\mathfrak{U} \rightarrow x$ there exist a $\alpha \in I$ such that $A_\alpha \subseteq U$ so that

$$A_\alpha \bigcap V \subseteq U \bigcap V = \varnothing$$

As $\mathfrak{U} \rightarrow y$ we have that $A_\alpha \bigcap V \neq \varnothing$ contradicting the above hence we must have $x = y$.

3. Let $\mathfrak{W} \gg \mathfrak{U}$ then we have:
   a. Let $U \in \mathcal{T}$ with $x \in U$ then as $\mathfrak{U} \rightarrow x$ there exist a $\alpha \in I$ such that $A_\alpha \subseteq U$. As $\mathfrak{M} \gg \mathfrak{U}$ there exist a $\beta \in J$ such that $B_\beta \subseteq A_\alpha \subseteq U$ proving that $\mathfrak{W} \rightarrow x$
   b. Let $U \in \mathcal{T}$ with $x \in U$ then as $\mathfrak{W} \succ x$ we have $\forall \beta \in J\ U \bigcap B_\beta \neq \varnothing$. Let $\alpha \in I$ then as $\mathfrak{W} \gg \mathfrak{U}$ there exist a $\gamma \in J$ such that $B_\gamma \subseteq A_\alpha$ hence $\varnothing \neq U \bigcap B_\gamma \subseteq U \bigcap A_\alpha$ proving that $\mathfrak{A} \succ x$. □

Definition 14.257. *Let $\langle X, \mathcal{T}\rangle$ be a topological space, $A \subseteq X$ and $\mathfrak{U} = \{A_\alpha\}_{\alpha \in I} \subseteq \mathcal{P}(X)$ a filter base then we say that $\mathfrak{U}$ is* **on** *A iff $\forall \alpha \in I$ we have $A_\alpha \subseteq A$.*

Theorem 14.258. *Let $\langle X, \mathcal{T}\rangle$ be a topological space, $A \subseteq X$ then*

$$x \in \overline{A} \Leftrightarrow \textit{there exist a filter base } \mathfrak{U} = \{A_\alpha\}_{\alpha \in I} \subseteq \mathcal{P}(X) \textbf{ on } A \textit{ with } \mathfrak{U} \rightarrow x$$

Proof.

⇒. Let $x \in \overline{A}$ then by [theorem: 14.28] $\forall U \in \mathcal{T}$ with $x \in U$ we have $A \bigcap U \neq \varnothing$. As $\varnothing \neq A \bigcap U \subseteq A \Rightarrow A \neq \varnothing$ we have using [example: 14.244] that $\mathfrak{U} = \{A_\alpha\}_{\alpha \in \{A\}}$ where $A_\alpha = A$ is a filter base in X, further by [theorem: 14.246] $\mathfrak{U}(x) = \{U\}_{U \in \{U \in \mathcal{T} | x \in U\}}$ is also a filter base in X. Define $\mathfrak{W} = \{A_\alpha \bigcap U\}_{(\alpha, U) \in \{A\} \times \{U \in \mathcal{T} | x \in U\}}$. As $\forall (\alpha, U) \in \{A\} \times \{U \in \mathcal{T} | x \in U\}$ we have that $\varnothing \neq A \bigcap U \subseteq U$ it follows from [theorem: 14.247] that

$$\mathfrak{W} \text{ is a filter base } \textbf{on } A$$

Further as $\forall U \in \mathcal{T}$ with $x \in U$ we have that for $(A, U) \in \{A\} \times \{U \in \mathcal{T} | x \in U\}$ $A \bigcap U \subseteq U$ it follows that

$$\mathfrak{W} \rightarrow x$$

⇐. By the hypothesis there exist a fileter base $\mathfrak{U} = \{A_\alpha\}_{\alpha \in I} \subseteq \mathcal{P}(X)$ **on** A with $\mathfrak{U} \rightarrow x$. Then if $U \in \mathcal{T}$ with $x \in U$ there exist a $\alpha \in I$ such that $A_\alpha \subseteq U$ [because $\mathfrak{U} \rightarrow x$], further, as $\mathfrak{U}$ is a filter base **on** A, $A_\alpha \subseteq A$. So it follows that $\varnothing \neq A_\alpha = A_\alpha \bigcap A \subseteq U \bigcap A$ which by [theorem: 14.28] proves that $x \in \overline{A}$. □

Theorem 14.259. *Let $f: X \rightarrow Y$ be a function between sets X and Y. If $\mathfrak{U} = \{A_\alpha\}_{\alpha \in I} \subseteq \mathcal{P}(X)$ is a filter base in X then if we define $f[\mathfrak{U}]$ by*

$$f[\mathfrak{U}] = \{f(A_\alpha)\}_{\alpha \in I} \subseteq \mathcal{P}(Y)$$

we have that

$$f[\mathfrak{U}] \text{ is a filter base in } Y$$

Proof. As $\mathfrak{U}$ is a filter base we have $I \neq \varnothing$ and

1. $\forall \alpha \in I$ $A_\alpha \neq \varnothing$ so that

$$f(A_\alpha) \neq \varnothing$$

2. If $\alpha, \beta \in I$ then there exist a $\gamma \in I$ such that $A_\gamma \subseteq A_\alpha \bigcap A_\beta$ so that using [theorem: 2.93]

$$f(A_\gamma) \subseteq f\left(A_\alpha \bigcap A_\beta\right) \subseteq f(A_\alpha) \bigcap f(A_\beta)$$

proving that $f[\mathfrak{U}] = \{f(A_\alpha)\}_{\alpha \in I} \subseteq \mathcal{P}(Y)$ is a filter base of Y. □

Filter bases allows us to have a equivalent definition of a continuous function.

Theorem 14.260. *Let $\langle X, \mathcal{T} \rangle$, $\langle Y, \mathcal{S} \rangle$ be topological spaces and $f: X \rightarrow Y$ is a function then*

$$f \text{ is continuous at } x \Leftrightarrow f[\mathfrak{U}(x)] \rightarrow f(x)$$

where

$$f[\mathfrak{U}(x)] = \{f(U)\}_{U \in \{U \in \mathcal{T} | x \in U\}} \text{ is a filter base by [theorem: 14.259]}$$

Proof. First note that $\mathfrak{U}(x) = \{U\}_{U \in \{U \in \mathcal{T} | x \in U\}} \subseteq \mathcal{P}(X)$ so that

$$f[\mathfrak{U}(x)] = \{f(U)\}_{U \in \{U \in \mathcal{T} | x \in U\}}$$

⇒. Let $V \in \mathcal{S}$ such that $f(x) \in V$ then as f is continuous at x there exist a $U \in T$ with $x \in U$ such that $f(U) \subseteq V$. In other words $\exists U \in \{U \in \mathcal{T} | x \in U\}$ such that $f(U) \subseteq V$ which by [definition: 14.249] proves that

$$f[\mathfrak{U}(x)] \rightarrow f(x)$$

⇐. Let $V \in \mathcal{S}$ with $f(x) \in V$ then as $f[\mathfrak{U}(x)] \rightarrow f(x)$ there exist a $U \in \{U \in \mathcal{T} | x \in U\}$ such that

$$f(U) \subseteq V$$

proving that f is continuous at x. □

Theorem 14.261. *Let $\langle X, \mathcal{T} \rangle$, $\langle Y, \mathcal{S} \rangle$ be topological spaces and $f: X \rightarrow Y$ a function then*

$$f \text{ is continuous on } X \Leftrightarrow \forall x \in X \text{ and for every filter base } \mathfrak{U} \text{ in } X \text{ with } \mathfrak{U} \rightarrow x \text{ we have } f[\mathfrak{U}] \rightarrow f(x)$$

Proof.

⇒. Let $x \in X$ and $\mathfrak{U} = \{A_\alpha\}_{\alpha \in I} \subseteq \mathcal{P}(X)$ such that $\mathfrak{U} \rightarrow x$ then by [theorem: 14.255] $\mathfrak{U} \gg \mathfrak{U}(x)$. So if $U \in \{U \in \mathcal{T} | x \in U\}$ by [definition: 14.253] there exist a $\alpha \in I$ such that $A_\alpha \subseteq U$, hence $f(A_\alpha) \subseteq f(U)$. So by [definition: 14.253]

$$f[\mathfrak{U}] \gg f[\mathfrak{U}(x)] \tag{14.134}$$

As f is continuous, hence continuous at x, we have by the previous theorem [theorem: 14.260] that

$$f[\mathfrak{U}(x)] \rightarrow f(x) \tag{14.135}$$

Using [theorem: 14.256] together with [eqs: 14.134, 14.135] proves that

$$f[\mathfrak{U}] \rightarrow f(x) \tag{14.136}$$

⇐. Take $A \subseteq X$ and let $y \in f(\overline{A})$ then $\exists x \in \overline{A}$ such that $y = f(x)$. Using [theorem: 14.258] there exist a filter base $\mathfrak{U} = \{A_\alpha\}_{\alpha \in I} \subseteq \mathcal{P}(X)$ **on** A such that $\mathfrak{U} \rightarrow x$. As $\mathfrak{U}$ is a filter base **on** A we have that $\forall \alpha \in I$ $A_\alpha \subseteq A$ so that $f(A_\alpha) \subseteq f(A)$ proving that $f[\mathfrak{U}]$ is **on** $f(A)$. This together with $f[\mathfrak{U}] \rightarrow f(x)$ [by the hypothesis] gives by [theorem: 14.258] that $y = f(x) \in \overline{f(A)}$ proving that

$$f(\overline{A}) \subseteq \overline{f(A)}$$

Finally using [theorem: 14.136] on the above proves that f is continuous. □

Theorem 14.262. *Let $\{\langle X_i, \mathcal{T}_i\rangle\}_{i\in I}$ be a family of topological spaces, $\langle X, \mathcal{T}_{\text{product}}\rangle$ the topological space defined by $X = \prod_{i\in I} X_i$ together with the product topology $\mathcal{T}_{\text{product}}$ [see theorem: 14.43], $x \in X$ and $\mathfrak{U} = \{A_\alpha\}_{\alpha\in J} \subseteq \prod_{i\in I} X_i$ a filter base in X and $x \in X$ then*

$$\mathfrak{U} \to x \Leftrightarrow \forall i \in I \text{ we have for } \pi_i: X \to X_i \text{ the projection function that } \pi_i[\mathfrak{U}] \to \pi_i(x) = x_i$$

Proof.

⇒. Let $i \in I$ as $\mathfrak{U} \to x$ we have, as $\pi_i: \prod_{i\in I} X_i \to X_i$ is continuous [see theorem: 14.139], by the previous theorem [theorem: 14.261] that $\pi_i[\mathfrak{U}] \to \pi_i(x) = x_i$.

⇐. Let $x \in \prod_{i\in I} X_i$ and $U \in \mathcal{T}_{\text{product}}$ such that $x \in U$. By the definition of the product topology [see theorem: 14.43] there exist a finite J and $\{S_i\}_{i\in J} \subseteq \{\pi_i^{-1}(V) | i \in I \wedge U \in \mathcal{T}_i\}$ such that

$$x \in \bigcap_{j\in J} S_j \subseteq U \tag{14.137}$$

Let $j \in J$ then $\exists i_j \in I$ and a $U_j \in \mathcal{T}_{i_j}$ such that $S_j = \pi_{i_j}^{-1}(U_j)$, so $\pi_{i_j}(S_j) \subseteq U_j$, hence $\pi_{i_j}(x) \in \pi_{i_j}(S_j) \subseteq U_j \in \mathcal{T}_{i_j}$. By the hypothesis $\pi_{i_j}[\mathfrak{U}] \to \pi_{i_j}(x)$ so there exist a $\alpha_j \in I$ such that $\pi_{i_j}(A_{\alpha_j}) \subseteq U_j$ hence $A_{\alpha_j} \subseteq \pi_{i_j}^{-1}(U_j) = S_j$. So $\bigcap_{j\in J} A_{\alpha_j} \subseteq \bigcap_{j\in J} S_j \subseteq U$. As J is finite it follows from [theorem: 14.243] that $\exists \gamma \in I$ such that

$$A_\gamma \subseteq \bigcap_{j\in J} A_{\alpha_j} \subseteq U$$

proving that

$$\mathfrak{U} \to x$$ □

Definition 14.263. *Let X a set then a filter base $\mathfrak{U}$ in X is called **maximal** or a **ultra filter** if for every filter base $\mathfrak{W}$ in X with $\mathfrak{W} \gg \mathfrak{U}$ we have that $\mathfrak{U} \gg \mathfrak{W}$.*

Theorem 14.264. *Let X a set then and $\mathfrak{U} = \{A_\alpha\}_{\alpha\in I} \subseteq \mathcal{P}(X)$ a filter base in X then*

$$\mathfrak{U} \text{ is maximal}$$

$$\Updownarrow$$

$$\forall A \subseteq X \text{ we have } (\exists \beta \in I \text{ with } A_\beta \subseteq A \wedge \forall \alpha \in I\ A_\alpha \nsubseteq X \setminus A\ ) \vee (\forall \alpha \in I A_\alpha \nsubseteq A \wedge \exists \beta \in I \text{ with } A_\beta \subseteq X \setminus A)$$

In other words

$$\mathfrak{U} \text{ is maximal} \Leftrightarrow \forall A \subseteq X \text{ only one of the sets } A, X \setminus A \text{ contains a member of } \mathfrak{U}$$

Proof.

⇒. Let $A \subseteq X$ and assume that there exist $\alpha, \beta \in I$ such that $A_\alpha \subseteq A$ and $A_\beta \subseteq X \setminus A$ then $A_\alpha \bigcap A_\beta \subseteq A \bigcap (X \setminus A) = \varnothing$ proving that $A_\alpha \bigcap A_\beta = \varnothing$. By definition of a filter base there exist a $\gamma \in I$ such that $\varnothing \neq A_\gamma \subseteq A_\alpha \bigcap A_\beta = \varnothing$ leading to the contradiction $\varnothing \neq \varnothing$. Hence

$$\neg(\exists \alpha, \beta \in I \text{ such that } A_\alpha \subseteq A \wedge A_\beta \subseteq X \setminus A) \tag{14.138}$$

We have now two possibilities to consider for $A \subseteq X$:

$\forall \alpha \in I A_\alpha \nsubseteq A$. Then $\forall \alpha \in I$ there exist a $x \in A_\alpha$ so that $x \notin A$, hence $(X \setminus A) \bigcap A_\alpha \neq \varnothing$. Further if $\alpha, \beta \in I$ then as $\mathfrak{U}$ is a filter base there exist a $\gamma \in I$ such that $A_\gamma \subseteq A_\alpha \bigcap A_\beta$ so that $(X \setminus A) \bigcap A_\gamma \subseteq (X \setminus A) \bigcap (A_\alpha \bigcap A_\beta) = ((X \setminus A) \bigcap A_\alpha) \bigcap ((A \setminus A) \bigcap A_\beta)$. Hence we have that

$$\mathfrak{W} = \left\{(X \setminus A) \bigcap A_\alpha\right\}_{\alpha\in I} \subseteq \mathcal{P}(X) \text{ is a filter base in X}$$

Further as $\forall \alpha \in I\ (X \setminus A) \bigcap A_\alpha \subseteq A_\alpha$ we have by [definition: 14.253] that $\mathfrak{W} \gg \mathfrak{U}$, as $\mathfrak{U}$ is maximal it follows that $\mathfrak{U} \gg \mathfrak{W}$. Hence $\forall \alpha \in I$ there exist a $\beta \in I$ such that $A_\beta \subseteq (X \setminus A) \bigcap A_\alpha \subseteq X \setminus A$. So in this case we have proved that

$$\forall \alpha \in I\ A_\alpha \nsubseteq A \wedge \exists \beta \in I \text{ with } A_\beta \subseteq X \setminus A$$

$\exists \beta \in I$ with $A_\beta \subseteq A$. Then by [eq: 14.138] we must have that $\forall \alpha \in I$ that $A_\alpha \nsubseteq X \setminus A$ proving that in this case

$$\exists \beta \in I \text{ with } A_\beta \subseteq A \wedge \forall \alpha \in I\ A_\alpha \nsubseteq X \setminus A$$

⇐. Let $\mathfrak{U}=\{A_\alpha\}_{\alpha\in I}\subseteq\mathcal{P}(X)$ be a filter bases such that

$$\forall A\subseteq X \text{ one and only one of the two sets } A, X\setminus A \text{ contains a member of } \mathfrak{U} \tag{14.139}$$

Let then $\mathfrak{W}=\{B_\beta\}_{\beta\in J}$ such that $\mathfrak{W}\gg\mathfrak{U}$. If $\beta\in J$ then for B_β we have by [eq: 14.139] that there exist a $\alpha\in I$ such that either $A_\alpha\subseteq B_\beta$ or $A_\alpha\subseteq X\setminus B_\beta$. The last case would mean that $A_\alpha\bigcap B_\beta=\varnothing$ contradicting [theorem: 14.255 (2)], so we must have that $A_\alpha\subseteq B_\beta$ proving that $\mathfrak{U}\gg\mathfrak{W}$ and thus maximality of $\mathfrak{U}$. □

Theorem 14.265. *Let X be a set, $\mathfrak{U}=\{A_\alpha\}_{\alpha\in I}\subseteq\mathcal{P}(X)$ a filter base in X then there exist a maximal filter base $\mathfrak{W}$ in X such that $\mathfrak{W}\gg\mathfrak{U}$*

Proof. We use a extension of Zorn's lemma [see theorem: 3.128] in this proof. To use this we must first define a pre-order [see definition: 3.25]. Let

$$\mathcal{A}=\{\mathfrak{W}|\mathfrak{W} \text{ is a filter base in } X \text{ with } \mathfrak{W}\gg\mathfrak{U}\}$$

Using [theorem: 14.255] we have that $\mathfrak{U}\gg\mathfrak{U}$ so that

$$\mathfrak{U}\in\mathcal{A} \text{ hence } \mathcal{A}\neq\varnothing$$

Define now the following relation $\leqslant$ on $\mathcal{A}$

$$\leqslant\subseteq\mathcal{A}\times\mathcal{A} \text{ by } \mathfrak{V}\leqslant\mathfrak{W} \text{ iff } \mathfrak{W}\gg\mathfrak{V}$$

then we have:

reflexivity. Let $\mathfrak{W}\in\mathcal{A}$ then by then by [theorem: 14.255] we have $\mathfrak{W}\leqslant\mathfrak{W}$.

transitivity. Assume that $\mathfrak{W}_1\leqslant\mathfrak{W}_2\wedge\mathfrak{W}_2\leqslant\mathfrak{W}_3$ then $\mathfrak{W}_2\gg\mathfrak{W}_1\wedge\mathfrak{W}_3\gg\mathfrak{W}_2$ hence by [theorem: 14.255] we have $\mathfrak{W}_3\gg\mathfrak{W}_1$ so that $\mathfrak{W}_1\leqslant\mathfrak{W}_3$.

Hence we have that

$$\langle\mathcal{A},\leqslant\rangle \text{ is a pre-ordered set}$$

Let $\mathcal{C}\subseteq\mathcal{A}$ be any **chain** in $\langle\mathcal{A},\leqslant\rangle$ then we must consider two cases for $\mathcal{C}$

$\mathcal{C}=\varnothing$. Then we have $\forall\mathfrak{C}\in\mathcal{C}$ that $\mathfrak{C}\leqslant\mathfrak{U}$ is satisfied vacuously proving that $\mathfrak{U}$ is a upper bound of $\mathcal{C}$.

$\mathcal{C}\neq\varnothing$. Given $\mathfrak{C}=\{C_\alpha\}_{\alpha\in J}\in\mathcal{C}$ define $S_{\mathfrak{C}}=\{C_\alpha|\alpha\in J\}\neq\varnothing$ [as $J\neq\varnothing$] and define

$$\mathbb{B}=\bigcup_{\mathfrak{C}\in\mathcal{C}}S_{\mathfrak{C}}\neq\varnothing \text{ [as } \mathcal{C}\neq\varnothing] \text{ and } \mathfrak{B}=\{B\}_{B\in\mathbb{B}}\subseteq\mathcal{P}(X) \tag{14.140}$$

then we have

$$\forall\mathfrak{C}=\{C_\alpha\}_{\alpha\in J}\in\mathcal{C} \text{ we have } \forall\alpha\in J \text{ that } C_\alpha\in\mathbb{B} \tag{14.141}$$

Next we have

1. $\forall B\in\mathbb{B}$ there exist a $\mathfrak{C}=\{C_\alpha\}_{\alpha\in J}\in\mathcal{C}$ such that $B\in S_{\mathfrak{C}}$, hence $\exists\alpha\in J$ such that $B=C_\alpha\neq\varnothing$ [as $\mathfrak{C}$ is a filter base]. Hence we have

$$\forall B\in\mathbb{B} \text{ that } B\neq\varnothing \tag{14.142}$$

2. Let $B_1,B_2\in\mathbb{B}$ then there exists $\mathfrak{C}_1=\{C^1_\alpha\}_{\alpha\in J_1}\in\mathcal{C}$, $\mathfrak{C}_2=\{C^2_\alpha\}_{\alpha\in J_2}\in\mathcal{C}$ such that $\exists\gamma_1\in J_1$, $\exists\gamma_2\in J_2$ with $B_1=C^1_{\gamma_1}$, $B_2=C^2_{\gamma_2}$. Now as $\mathcal{C}$ is a chain we have the following possible cases:

   $\mathfrak{C}_1\leqslant\mathfrak{C}_2$. Then $\mathfrak{C}_2\gg\mathfrak{C}_1$, hence there exist a $\gamma\in J_2$ such that $C^2_\gamma\subseteq C^1_{\gamma_1}=B_1$. As $\mathfrak{C}_2$ is a filter base there exist a $\beta\in J_2$ such that $C^2_\beta\subseteq C^2_\gamma\bigcap C^2_{\gamma_2}\subseteq B_1\bigcap B_2$. Take $B_3=C^2_\beta\in\mathbb{B}$ then $B_3\subseteq B_1\bigcap B_2$.

   $\mathfrak{C}_2\leqslant\mathfrak{C}_1$. Then $\mathfrak{C}_1\gg\mathfrak{C}_2$, hence there exist a $\gamma\in J_1$ such that $C^1_\gamma\subseteq C^2_{\gamma_1}=B_2$. As $\mathfrak{C}_1$ is a filter base there exist a $\beta\in J_1$ such that $C^1_\beta\subseteq C^1_{\gamma_1}\bigcap C^1_\gamma\subseteq B_1\bigcap B_2$. Take $B_3=C^1_\beta\in\mathbb{B}$ then $B_3\subseteq B_1\bigcap B_2$.

   So in all cases we have

$$\forall B_1,B_2\in\mathbb{B} \text{ there exist a } B_3\in\mathbb{B} \text{ such that } B_3\subseteq B_1\bigcap B_2 \tag{14.143}$$

Combining [eqs: 14.140, 14.142 and 14.143] proves that

$$\mathfrak{B} \text{ is a filter base}$$

As $\mathcal{C}\neq\varnothing$ there exist a $\mathfrak{C}=\{C_\alpha\}_{\alpha\in J}\in\mathcal{C}\subseteq\mathcal{A}$ so that $\mathfrak{C}\gg\mathfrak{U}$, hence if $\alpha\in I$ then there exist a $\gamma\in J$ such that $C_\gamma\subseteq A_\alpha$. By [eq: 14.141] we have that $C_\gamma\in\mathbb{B}$ proving that $\mathfrak{B}\gg\mathfrak{U}$, hence

$$\mathfrak{B}\in\mathcal{A}$$

Further if $\mathfrak{D} = \{D_\alpha\}_{\alpha \in K} \in \mathcal{C}$ then $\forall \alpha \in K$ $D_\alpha \subseteq D_\alpha$ hence, as by [eq: 14.141] $D_\alpha \in \mathbb{B}$, we have that $\mathfrak{B} \gg \mathfrak{D}$ proving that $\mathfrak{D} \leqslant \mathfrak{B}$. Hence

$$\mathfrak{B} \text{ is a upper bound of } \mathfrak{C}$$

So in all cases we have that $\mathfrak{C}$ has a upper bound. In other words we have proved that every chain in $\mathcal{A}$ has a upper bound. So using a variant of Zorn's lemma [theorem: 3.128] there exist a $\mathfrak{M} \in \mathcal{A}$ such that $\forall \mathfrak{N} \in \mathcal{A}$ with $\mathfrak{M} \leqslant \mathfrak{N}$ we have $\mathfrak{N} \leqslant \mathfrak{M}$. Hence

$$\forall \mathfrak{N} \in \mathcal{A} \text{ with } \mathfrak{N} \gg \mathfrak{M} \text{ we have } \mathfrak{M} \gg \mathfrak{N} \text{ and as } \mathfrak{M} \in \mathcal{A} \text{ we have also } \mathfrak{M} \gg \mathfrak{U} \tag{14.144}$$

Now if $\mathfrak{O}$ is a filter base in X with $\mathfrak{O} \gg \mathfrak{W}$ we have as $\mathfrak{W} \geqslant \mathfrak{U}$ by [theorem: 14.255] that $\mathfrak{O} \gg \mathfrak{U}$ so that $\mathfrak{O} \in \mathcal{A}$, hence by [eq: 14.144] it follows that $\mathfrak{M} \gg \mathfrak{O}$. Hence

$$\mathfrak{M} \text{ is a ultra filter or maximal filter base with } \mathfrak{M} \gg \mathfrak{U}$$

□

Theorem 14.266. *Let $\langle X, \mathcal{T} \rangle$ be a topological space and $\mathfrak{U} = \{A_i\}_{i \in I} \subseteq \mathcal{P}(X)$ a maximal filter then $\forall x \in X$*

$$\mathfrak{U} \succ x \Leftrightarrow \mathfrak{U} \rightarrow x$$

Proof.

- **$\Rightarrow$.** Let $U \in \mathcal{T}$ with $x \in U$. Using [theorem: 14.264] there exist a $\alpha \in I$ such that either $A_\alpha \subseteq U$ or $A_\alpha \subseteq X \setminus U$. Now as $\mathfrak{U} \succ x$ we must have that $A_\alpha \bigcap U \neq \varnothing$ contradicting $A_\alpha \subseteq X \setminus U$ hence we must have that $A_\alpha \subseteq U$, so $\mathfrak{U} \rightarrow x$.
- **$\Leftarrow$.** This follows from [theorem: 14.256]. □

Theorem 14.267. *Let X, Y be sets and $f: X \rightarrow Y$ a function then if $\mathfrak{U} = \{A_i\}_{i \in I} \subseteq \mathcal{P}(X)$ is a maximal filter base in X then $f[\mathfrak{U}]$ is a maximal filter base in Y.*

Proof. Using [theorem: 14.259] $f[\mathfrak{U}] = \{f(A_i)\}_{i \in I}$ is a filter base in Y. To prove maximality we use [theorem: 14.264]. So let $A \subseteq Y$ and consider $f^{-1}(A)$ then by the maximality of $\mathfrak{U}$ we have either:

- **$\exists \beta \in I$ with $A_\beta \subseteq f^{-1}(A) \wedge \forall \alpha \in I$ $A_\alpha \nsubseteq X \setminus f^{-1}(A)$.** Then $f(A_\beta) \subseteq A$ and $\forall \alpha \in I$ there exist a $x \in A_\alpha$ such that $x \notin X \setminus f^{-1}(A) \Rightarrow x \in f^{-1}(A) \Rightarrow f(x) \in A \Rightarrow f(x) \notin Y \setminus A$ so that $f(A_\alpha) \nsubseteq Y \setminus A$.
- **$\forall \alpha \in I$ $A_\alpha \nsubseteq f^{-1}(A) \wedge \exists \beta \in I$ with $A_\beta \subseteq X \setminus f^{-1}(A)$.** Then $\forall \alpha$ we have that $\exists x \in A_\alpha$ such that $x \notin f^{-1}(A) \Rightarrow f(x) \notin A$ proving that $f(A_\alpha) \nsubseteq A$. Further $\forall x \in A_\beta$ we have $x \in X \setminus f^{-1}(A) \Rightarrow x \notin f^{-1}(A) \Rightarrow f(x) \notin A$ proving that $f(A_\beta) \subseteq Y \setminus A$.

Using [theorem: 14.264] it follows then that $f[\mathfrak{U}]$ is a maximal filter base in Y. □

14.8.1.2 Tychonoff's theorem

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 14.268. *Let $\langle X, \mathcal{T} \rangle$ be a topological space then the following are equivalent*

1. *X is compact*
2. *If $\{A_\alpha\}_{\alpha \in I} \subseteq \mathcal{T}^c$ is a non empty family of closed sets in X satisfying $\bigcap_{\alpha \in I} A_\alpha = \varnothing$ then there exist a non empty finite $J \subseteq I$ such that $\bigcap_{\alpha \in J} A_\alpha = \varnothing$.*
3. *If $\{A_\alpha\}_{\alpha \in I} \subseteq \mathcal{T}^c$ is a non empty family of closed sets in X such that $\forall J \subseteq I$ J finite and non empty we have $\bigcap_{\alpha \in J} A_\alpha \neq \varnothing$ then $\bigcap_{\alpha \in I} A_\alpha \neq \varnothing$.*
4. *If $\mathfrak{U}$ is a filter base in X then there exist a $x \in X$ such that $\mathfrak{U} \succ x$*
5. *If $\mathfrak{U}$ is a maximal filter base in X then there exist a $x \in X$ such that $\mathfrak{U} \rightarrow x$*

Proof.

$1 \Rightarrow 2$. Let $\{A_\alpha\}_{\alpha \in I} \subseteq \mathcal{T}^c$ be a family of closed sets in X such that $\bigcap_{\alpha \in I} A_\alpha = \varnothing$ then

$$X = X \setminus \varnothing = X \setminus \left(\bigcap_{\alpha \in I} A_\alpha \right) \underset{[\text{theorem: } 2.130]}{=} \bigcup_{\alpha \in I} (X \setminus A_\alpha)$$

As $\forall \alpha \in I$ $X \setminus A_\alpha \in \mathcal{T}$ and X is compact there exist a finite $J \subseteq I$ such that

$$X = \bigcup_{\alpha \in J} (X \setminus A_\alpha)$$

hence

$$\varnothing = X \setminus X = X \setminus \left(\bigcup_{\alpha \in J} (X \setminus A_\alpha) \right) \underset{[\text{theorem: } 2.130]}{=} \bigcap_{\alpha \in J} (X \setminus (X \setminus A_\alpha)) = \bigcap_{\alpha \in J} A_\alpha$$

2 ⇒ 1. Let $\{U_\alpha\}_{\alpha \in I} \subseteq \mathcal{T}$ be a family of open sets such that $X = \bigcup_{\alpha \in I} U_\alpha$ then

$$\varnothing = X \setminus X = X \setminus \left(\bigcup_{\alpha \in I} U_\alpha \right) \underset{[\text{theorem: } 2.130]}{=} \bigcap_{\alpha \in I} (X \setminus U_\alpha)$$

As $\forall \alpha \in I$ we have that $X \setminus U_\alpha$ is closed we have by the hypothesis that there exist a non empty finite $J \subseteq I$ such that $\bigcap_{\alpha \in J} (X \setminus U_\alpha) = \varnothing$. Hence we have

$$X = X \setminus \varnothing = X \setminus \left(\bigcap_{\alpha \in J} (X \setminus U_\alpha) \right) \underset{[\text{theorem: } 2.130]}{=} \bigcup_{\alpha \in J} (X \setminus (X \setminus U_\alpha)) = \bigcup_{\alpha \in J} U_\alpha$$

proving that X is compact.

2 ⇒ 3. Let $\{A_\alpha\}_{\alpha \in I} \subseteq \mathcal{T}^c$ be a family of closed sets in X such that $\forall J \subseteq I$ J finite and non empty we have $\bigcap_{\alpha \in J} A_\alpha \neq \varnothing$. Assume that $\bigcap_{\alpha \in I} A_\alpha = \varnothing$ then by (2) there exist a finite non empty $J \subseteq I$ such that $\bigcap_{\alpha \in J} A_\alpha = \varnothing$ contradiction the fact that $\bigcap_{\alpha \in J} A_\alpha \neq \varnothing$. Hence we must have that

$$\bigcap_{\alpha \in I} A_\alpha \neq \varnothing$$

3 ⇒ 2. Let $\{A_\alpha\}_{\alpha \in I} \subseteq \mathcal{T}^c$ be a family of closed sets in X such that $\bigcap_{\alpha \in I} A_\alpha = \varnothing$. Assume that $\forall J \subseteq I$ finite non empty we have $\bigcap_{\alpha \in J} A_\alpha \neq \varnothing$ then by (3) $\bigcap_{\alpha \in I} A_\alpha \neq \varnothing$ contradicting $\bigcap_{\alpha \in I} A_\alpha = \varnothing$. Hence there must exist a finite non empty $J \subseteq I$ such that $\bigcap_{\alpha \in J} A_\alpha = \varnothing$, proving (3).

3 ⇒ 4. Let $\mathfrak{U} = \{A_\alpha\}_{\alpha \in I} \subseteq \mathcal{P}(X)$ be a filter base, then for every finite $J \subseteq I$ we have by [theorem: 14.243] that there exist a $\gamma \in I$ such that $\varnothing \neq A_\gamma \subseteq \bigcap_{\alpha \in J} A_\alpha$ proving that $\varnothing \neq \bigcap_{\alpha \in J} A_\alpha$ hence, as $\forall \alpha \in J$ we have $A_\alpha \subseteq \overline{A_\alpha}$, $\varnothing \neq \bigcap_{\alpha \in J} \overline{A_\alpha}$. Using (3) it follows then that $\varnothing \neq \bigcap_{\alpha \in I} \overline{A_\alpha}$, hence there exist a $x \in X$ such that $x \in \bigcap_{\alpha \in I} \overline{A_a}$. Applying [theorem: 14.251] proves that

$$\mathfrak{U} \succ x$$

4 ⇒ 5. Let $\mathfrak{U}$ be a maximal filter base then by (4) there exist a $x \in X$ such that $\mathfrak{U} \succ x$, by [theorem: 14.266] we have then that

$$\mathfrak{U} \rightarrow x$$

5 ⇒ 3. Let $\{F_\alpha\}_{\alpha \in I} \subseteq \mathcal{T}^c$ a family of closed sets in X such that for every finite non empty $J \subseteq I$ we have $\bigcap_{\alpha \in J} F_\alpha \neq \varnothing$. Take $K = \{\bigcap_{\alpha \in J} F_\alpha | J \subseteq I,\ J \text{ finite non empty}\}$ and define the family

$$\mathfrak{U} = \{A\}_{A \in K} \subseteq \mathcal{P}(X)$$

We prove now that $\mathfrak{U}$ is a filter base. First, as $I \neq \varnothing$ there exist a $i \in I$, as $\{i\}$ is finite and $\{i\} \subseteq I$ it follows that $\bigcap_{\alpha \in \{i\}} F_\alpha \in K$ so that

$$K \neq \varnothing \tag{14.145}$$

Second, if $A \in K$ then there exist a finite $J \subseteq I$ such that $A = \bigcap_{\alpha \in K} F_\alpha$, as $\bigcap_{\alpha \in K} F_\alpha \neq \varnothing$ it follows that $A \neq \varnothing$. So we have proves that

$$\forall A \in K \text{ we have } A \neq \varnothing \tag{14.146}$$

Third, let $A_1, A_2 \in K$ then there exist finite non empty $J_1, J_2 \subseteq I$ such that $A_1 = \bigcap_{\alpha \in J_1} F_\alpha$, $A_2 = \bigcap_{\beta \in J_2} F_\beta$. If $x \in A_1 \bigcap A_2 = (\bigcap_{\alpha \in J_1} F_\alpha) \bigcap (\bigcap_{\beta \in J_2} F_\beta)$ then if $\gamma \in J_1 \bigcup J_2$ we have either or both $\gamma \in J_1 \Rightarrow x \in F_\gamma$ or $\gamma \in J_2 \Rightarrow x \in F_\gamma$ so that $x \in \bigcap_{\gamma \in J_1 \bigcup J_2} F_\gamma$, hence

$$A_1 \bigcap A_2 \subseteq \bigcap_{\gamma \in J_1 \bigcup J_2} F_\gamma \tag{14.147}$$

If $x \in \bigcap_{\gamma \in J_1 \bigcup J_2} F_\gamma$ then $\forall \alpha \in J_1$ we have $\alpha \in J_1 \bigcup J_2$ so that $x \in F_\alpha$, hence $x \in \bigcap_{\alpha \in J_1} F_\alpha = A_1$ and $\forall \beta \in J_2$ we have $\beta \in J_1 \bigcup J_2$ so that $x \in F_\beta$, hence $x \in \bigcap_{\alpha \in J_2} F_\beta = A_2$. So $x \in A_1 \bigcap A_2$ proving that $\bigcap_{\gamma \in J_1 \bigcup J_2} F_\gamma \subseteq A_1 \bigcap A_2$, combining this with [eq: 14.147] proves that

$$A_1 \bigcap A_2 = \bigcap_{\gamma \in J_1 \bigcup J_2} F_\gamma$$

As $J_1 \bigcup J_2$ is finite non empty it follows that $A_3 = \bigcap_{\gamma \in J_1 \bigcup J_2} F_\gamma \in K$. So we have proved that

$$\forall A_1, A_2 \in K \text{ there exist a } A_3 \in K \text{ such that } A_3 = A_1 \bigcap A_2 \in K \quad (14.148)$$

From [eqs: 14.145, 14.146, 14.148] it follows that

$$\mathfrak{U} = \{A\}_{A \in K} \subseteq \mathcal{P}(X) \text{ is a filter base in } X$$

Using [theorem: 14.265] there exist a maximal filter base $\mathfrak{M}$ in X such that $\mathfrak{M} \gg \mathfrak{U}$. By (5) there exist a $x \in X$ such that $\mathfrak{M} \to x$, using [theorem: 14.256] it follows that $\mathfrak{M} \succ x$, finally as $\mathfrak{M} \gg \mathfrak{U}$ it follows from [theorem: 14.256] that

$$\mathfrak{U} \succ x$$

By [theorem: 14.251] we have then $x \in \bigcap_{A \in K} \overline{A}$ hence $\forall A \in K$ we have $x \in \overline{A}$. For every $\alpha \in I$ $\{\alpha\} \subseteq I$ and $\{\alpha\}$ is finite so that $F_\alpha = \bigcap_{\alpha \in \{a\}} F_a \in K$, hence $x \in \overline{F_\alpha} \underset{F_\alpha \text{ is closed}}{=} F_a$. So we have proved that

$$x \in \bigcap_{\alpha \in I} F_\alpha \Rightarrow \bigcap_{\alpha \in I} F_\alpha \neq \varnothing \qquad \square$$

We are now ready for Tychonoff's theorem.

Theorem 14.269. *Let $\{\langle X_i, \mathcal{T}_i\rangle\}_{i \in I}$ be a family of non empty topological spaces then for the topological space $\langle \prod_{i \in I} X_i, \mathcal{T}_{\text{product}}\rangle$ based on the product topology we have*

$$\prod_{i \in I} X_i \textit{ is compact} \Leftrightarrow \forall i \in I\ X_i \textit{ is compact}$$

Proof.

⇒. Let $i \in I$ then by [theorem: 14.139] π_i: $\prod_{i \in I} X_i \to X_i$ is continuous, hence using [theorem: 14.229] it follows that $\pi_i(\prod_{i \in I} X_i)$ is compact. Further as $\forall i \in I$ $X_i \neq \varnothing$ we have by [theorem: 3.105] π_i: $\prod_{i \in I} X_i \to X_i$ is surjective so that $\pi_i(\prod_{i \in I} X_i) = X_i$. Hence we have that X_i is compact.

⇐. Let $\mathfrak{U}$ be a maximal filter base in $\prod_{i \in I} X_i$. Given $i \in I$ we have by [theorem: 14.267] that $\pi_i[\mathfrak{U}]$ is a maximal filter base in X_i, as X_i is compact it follows from [theorem: 14.268 (5)] that $\exists x_i \in X_i$ such that $\pi_i[\mathfrak{U}] \to x_i$. This defines a $x \in \prod_{i \in I} X_i$ so that $\pi_i[\mathfrak{U}] \to x_i = \pi_i(x)$. Using [theorem: 14.262] it follows that $\mathfrak{U} \to x$ so by [theorem: 14.268 (5)] again we have that $\prod_{i \in I} X_i$ is compact. □

Corollary 14.270. *Let $\{\langle X_i, \mathcal{T}_i\rangle\}_{i \in I}$ be a family of topological spaces then if $\forall i \in I$ we have $\varnothing \neq C_i \subseteq X_i$ we have that*

$$\prod_{i \in I} C_i \textit{ is compact} \Leftrightarrow \forall i \in I\ C_i \text{ is compact}$$

Proof. Note that the subspace topology of $\prod_{i \in I} C_i$ is equal to the product topology of the subspace topologies [see theorem 14.48]. As by definition a subset is compact is if it is compact in the subspace topology we can use Tychonoff's [theorem: 14.269] to prove this theorem. □

Note 14.271. If in the above corollary there exist a $i \in I$ such that $C_i = \varnothing$ then $\prod_{i \in I} C_i = \varnothing$ hence by [example: 14.223] compact even if the rest of the C_i's are not compact. Hence

14.8.1.3 Consequences of Tychonoff's Theorem

Theorem 14.272. *Let $\langle \mathbb{R}^n, \mathcal{T}_{\|\|_n}\rangle$ be the topological space of tupples of real numbers where the norm $\|\|_n$ is defined by $\|x\|_n = \max(\{|\pi_i(x)||i \in \{1, \ldots, n\}\})$ then for $\{a_i\}_{i \in \{1, \ldots, n\}} \subseteq \mathbb{R}$ and $\{b_i\}_{i \in \{1, \ldots, n\}} \subseteq \mathbb{R}$ with $\forall i \in \{1, \ldots, n\}$ $a_i \leqslant b_i$ we have that*

$$\prod_{i \in \{1, \ldots, n\}} [a_i, b_i] \textit{ is compact}$$

Proof. First using [theorem: 14.94] we have that $\mathcal{T}_{\|\|_n}$ is the product topology based on the product of the normed spaces $\langle \mathbb{R}, \|\|\rangle$. Now by Heine Borel for the real numbers [see theorem: 14.232] we have that $\forall i \in \{1, \ldots, n\}$ $[a_i, b_i]$ is compact, so using Tychonoff's theorem [theorem: 14.270] it follows that $\prod_{i \in \{1, \ldots, n\}} [a_i, b_i]$ is compact. □

Theorem 14.273. (Heine-Borel) *Let $\langle \mathbb{R}^n, \mathcal{T}_{\|\|_n}\rangle$ be the topological space of the tupples of real numbers based on the maximum norm and $C \subseteq \mathbb{R}^n$ then we have*

$$C \textit{ is compact} \Leftrightarrow C \textit{ is bounded and closed}$$

Proof.

$\Rightarrow$. Using [theorem: 14.227] C is bounded, further by [theorem: 14.201] $\mathcal{T}_{\|\|_n}$ is Hausdorff, so using [theorem: 14.230] C is closed.

$\Leftarrow$. If C is empty then C is compact by [example: 14.223] so we may assume that $C \neq 0$. Hence their exist a $c \in C$. As C is bounded there exist a $M \in \mathbb{R}^+$ such that

$$\forall x \in C \text{ we have } \max(\{|\pi_i(x-c)| | i \in \{1,\ldots,n\}\}) = \|x-c\|_{\max} \leqslant M$$

So given $x \in C$ and $i \in \{1,\ldots,n\}$ we have that $|x_i - c_i| = |\pi_i(x) - \pi_i(c)| = |\pi_i(x-c)| \leqslant M$ or $x_i \in [c_i - M, c_i + M]$, hence we have that $x \in \prod_{i \in \{1,\ldots,n\}} [c_i - M, c_i + M]$ proving that

$$C \subseteq \prod_{i \in \{1,\ldots,n\}} [c_i - M, c_i + M]$$

Using the previous theorem [theorem: 14.272] $\prod_{i \in \{1,\ldots,n\}} [c_i - M, c_i + M]$ is compact and as C is closed we have by [theorem: 14.231] that C is compact. □

A important consequence of the Heine-Borel theorem is that all norms on $\mathbb{R}^n$ are equivalent [see definition: 14.112 and theorem: 14.114].

Theorem 14.274. *let $n \in \mathbb{N}$ then all norms on $\mathbb{R}^n$ are equivalent.*

Proof. Let $\|\|_n$ be the maximum norm defined as $\|x\|_n = \max(\{\pi_i(x) | i \in \{1,\ldots,n\}\})$ and let $\|\|$ be another norm defined on $\mathbb{R}^n$. Using [theorem: 14.149] it follows that

$$\|\|: \mathbb{R}^n \rightarrow \mathbb{R} \text{ is continuous using the norms } \|\|_n \text{ and } || \tag{14.149}$$

Further using [example: 14.153] it follows that

$$\|\|_n: \mathbb{R}^n \rightarrow \mathbb{R} \text{ is continuous using the norms } \|\|_n \text{ and } || \tag{14.150}$$

Define now

$$S = \{x \in \mathbb{R}^n | \|x\|_n = 1\}$$

so that

$$S = \|\|_n^{-1}(\{1\})$$

As $\langle \mathbb{R}, || \rangle$ is Hausdorff [see theorem: 14.201] we have by [theorem: 14.203] that $\{1\}$ is closed. Using the continuity of $\|\|_n$ [see eq: 14.150] it follows from [theorem: 14.136] that $\|\|_n^{-1}(\{1\})$ is closed, hence

$$S \text{ is closed} \tag{14.151}$$

Further if $x, y \in S$ then $\|x - y\|_n \leqslant \|x\|_n + \|y\|_n \leqslant 1 + 1 = 2$ which proves that

$$S \text{ is bounded} \tag{14.152}$$

So [eqs: 14.151, 14.152] allows us to use Heine-Borel [theorem: 14.273] to prove that

$$S \text{ is compact}$$

Using the above and the fact that $\|\|$ is continuous [eq: 14.149] allows us to use the extreme value theorem [theorem: 14.235], hence there exists $x_1, x_2 \in \mathbb{R}^n$ such that $\forall x \in S$ we have $\|x_1\| \leqslant \|x\| \leqslant \|x_2\|$. Define $M_1 = \|x_1\|$ and $M_2 = \|x_2\|$ then as $\|x_1\|_n = 1 \Rightarrow x_1 \neq 0$ and $\|x_2\|_n = 1 \Rightarrow x_2 \neq 0$ we have

$$\exists \alpha, \beta \in \mathbb{R}^+ \text{ such that } \forall x \in \mathbb{R}^n \text{ with } \|x\|_n = 1 \text{ we have } M_1 \leqslant \|x\| \leqslant M_2 \text{ where } M_1, M_2 \in \mathbb{R}^+ \tag{14.153}$$

Let $x \in X$ then we have either

$x = 0$. Then $\|x\|_n = 0 = \|x\|$ so that $M_1 \cdot \|x\|_n = 0 = \|x\| = 0 = M_2 \cdot \|x\|_n$ hence

$$M_1 \cdot \|x\|_n \leqslant \|x\| \leqslant M_2 \cdot \|x\|_n$$

$x \neq 0$. Then $\|x\|_n \neq 0$ so that $\left\|\frac{1}{\|x\|_n} \cdot x\right\|_n = \frac{1}{\|x\|_n} \cdot \|x\|_n = 1$ so that by [eq: 14.153]

$$M_1 \leqslant \left\|\frac{1}{\|x\|_n} \cdot x\right\| \leqslant M_2 \Rightarrow M_1 \leqslant \frac{1}{\|x\|_n} \cdot \|x\| \leqslant M_2$$

giving

$$M_1 \cdot \|x\|_n \leqslant \|x\| \leqslant M_2 \cdot \|x\|_n$$

so in all cases we have $M_1 \cdot \|x\|_n \leqslant \|x\| \leqslant M_2 \cdot \|x\|_n$. Applying then [theorem: 14.114] proves that

$$\|\|_n \text{ and } \|\| \text{ are equivalent norms}$$

As $\|\|$ was chosen arbitrary we have proved that every norm in $\mathbb{R}^n$ is equivalent with $\|\|_n$. Hence by [theorem: 14.113] all the norms in $\mathbb{R}^n$ are equivalent. □

Corollary 14.275. *Let X be a finite dimensional vector space over $\mathbb{R}$ then all norms on X are equivalent*

Proof. Let $\|\|_1$, $\|\|_2$ two norms on X then by [theorem: 14.170] there exist two norms $\|\|_{(1)}$ and $\|\|_{(2)}$ on $\mathbb{R}^n$ and a **isometric** isomorphism (depending only on the basis chosen in X)

$$\varphi: \mathbb{R}^n \rightarrow X \text{ so that } \forall x \in \mathbb{R}^n \text{ we have } \|\varphi(x)\|_1 = \|x\|_{(1)} \wedge \|\varphi(x)\|_2 = \|x\|_{(2)}$$

By the previous theorem [theorem: 14.274] $\|\|_{(1)}$ and $\|\|_{(2)}$ are equivalent, hence by [theorem: 14.114] there exists $M_1, M_2 \in \mathbb{R}^+$ such that

$$\forall x \in \mathbb{R}^n \text{ we have } M_1 \cdot \|x\|_{(1)} \leqslant \|x\|_{(2)} \leqslant M_2 \cdot \|x\|_{(1)}$$

Let $x \in X$ then $x = \varphi(\varphi^{-1}(x))$ so that

$$\|x\|_1 = \|\varphi(\varphi^{-1}(x))\|_1 = \|\varphi^{-1}(x)\|_{(1)} \text{ and } \|x\|_2 = \|\varphi(\varphi^{-1}(x))\|_2 = \|\varphi^{-1}(x)\|_{(2)}$$

hence we have

$$M_1 \cdot \|x\|_1 = M_1 \cdot \|\varphi^{-1}(x)\|_{(1)} \leqslant \|\varphi^{-1}(x)\|_{(2)} = \|x\|_2$$

and

$$\|x\|_2 = \|\varphi^{-1}(x)\|_{(2)} \leqslant M_2 \cdot \|\varphi^{-1}(x)\|_{(1)} = M_2 \cdot \|x\|_1$$

giving

$$M_1 \cdot \|x\|_1 \leqslant \|x\|_2 \leqslant M_2 \cdot \|x\|_1$$

proving by [theorem: 14.114] that $\|\|_1$ and $\|\|_2$ are equivalent. □

Theorem 14.276. *Let $\langle X, \|\| \rangle$ be a normed space over $\mathbb{R}$, $n \in \mathbb{N}$, $L: \mathbb{R}^n \rightarrow X$ a linear isomorphism and $C \subseteq X$ then*

$$C \textit{ is compact} \Leftrightarrow C \textit{ is closed and bounded}$$

Proof. Define

$$\|\|^*: \mathbb{R}^n \rightarrow \mathbb{R} \text{ by } \|x\|^* = \|L(x)\|$$

then we have

1. If $x \in \mathbb{R}^n$ then $0 \leqslant \|L(x)\| = \|x\|^*$
2. If $\alpha \in \mathbb{R}$, $x \in \mathbb{R}^n$ then
$$\|\alpha \cdot x\|^* = \|L(\alpha \cdot x)\| = \|\alpha \cdot L(x)\| = |\alpha| \cdot \|L(x)\| = |\alpha| \cdot \|x\|^*$$
3. If $x, y \in \mathbb{R}^n$ then
$$\|x + y\|^* = \|L(x + y)\| = \|L(x) + L(y)\| \leqslant \|L(x)\| + \|L(y)\| = \|x\|^* + \|y\|^*$$
4. If $\|x\|^* = 0$ then $\|L(x)\| = 0$ which as $\|\|$ is a norm means that $L(x) = 0$, hence as L is a isomorphism it follows that $x = 0$.

So we have proved that

$$\|\|^* \text{ is a norm on } \mathbb{R}^n \text{ and } L \text{ is a linear isometric isomorphism between } \langle \mathbb{R}^n, \|\|^* \rangle \text{ and } \langle X, \|\| \rangle \tag{14.154}$$

Using [theorem: 14.167] it follows that

$$L: \mathbb{R}^n \rightarrow X \text{ is a homeomorphism using the topologies } \mathcal{T}_{\|\|^*} \text{ on } \mathbb{R}^n \text{ and } \mathcal{T}_{\|\|} \text{ on } X$$

As by [theorem: 14.274] all norms are equivalent we have that $\mathcal{T}_{\|\|_n} = \mathcal{T}_{\|\|^*}$ where $\|\|_n$ is the maximum norm on $\mathbb{R}^n$. Hence we have that

$$L: \mathbb{R}^n \rightarrow X \text{ is a homeomorphism using the topologies } \mathcal{T}_{\|\|_n} \text{ on } \mathbb{R}^n \text{ and } \mathcal{T}_{\|\|} \text{ on } X \tag{14.155}$$

Further as $\|\|^*$ and $\|\|_n$ are equivalent [see theorem: 14.274] there exist by [theorem: 14.114] a $N_1, N_2 \in \mathbb{R}^+$ such that

$$\forall x \in \mathbb{R}^n \ \|x\|^* \leqslant N_1 \cdot \|x\|_n \text{ and } \|x\|_n \leqslant N_2 \cdot \|x\|^\star \tag{14.156}$$

Assume now that $C \subseteq X$ is compact. Then, as $K = L^{-1}: X \rightarrow \mathbb{R}^n$ is a continuous mapping between $\langle X, \mathcal{T}_{\|\|}\rangle$ and $\langle \mathbb{R}^n, \mathcal{T}_{\|\|_n}\rangle$, we have using [theorem: 14.229] that $K(C)$ is compact in $\langle \mathbb{R}^n, \mathcal{T}_{\|\|_n}\rangle$, hence using Heine-Borel [see theorem: 14.273] we have that

$$K(C) \text{ is closed and bounded in } \langle \mathbb{R}^n, \|\|_n\rangle$$

As K is continuous it follows from [theorem: 14.136] we have that $K^{-1}(K(C))$ is closed, which as $K = L^{-1}$ is a bijection, proves by [theorem: 2.55] that $K^{-1}(K(C)) = C$, hence we have that

$$C \text{ is closed in } \langle X, \|\|\rangle$$

Further as $K(C)$ is bounded there exist a $M \in \mathbb{R}^+$ such that $\forall x, y \in K(C)$ we have $\|x - y\|_n \leqslant M$. Let $x, y \in C$ then

$$\|x - y\| = \|L(L^{-1}(x - y))\| = \|L(K(x - y))\| = \|K(x - y)\|^* \leqslant_{[\text{eq: } 14.156]} N_1 \cdot \|K(x - y)\|_n$$

as $K(x), K(y) \in K(C)$ we have $\|K(x - y)\|_n = \|K(x) - K(y)\|_n \leqslant M$, hence $\|x - y\| \leqslant N_1 \cdot M$. So we have proved that

$$C \text{ is bounded in } \langle X, \|\|\rangle$$

Summarized we have proved that

$$C \text{ is compact} \Rightarrow C \text{ is closed and bounded} \tag{14.157}$$

For the opposite implication assume that C is closed and bounded. As by [eq: 14.155] L is a continuous map between $\langle \mathbb{R}^n, \mathcal{T}_{\|\|_n}\rangle$ and $\langle X, \mathcal{T}_{\|\|}\rangle$ it follows from [theorem: 14.136] that

$$L^{-1}(C) \text{ is closed in } \mathcal{T}_{\|\|_n} \tag{14.158}$$

Further as C is bounded there exist a $M \in \mathbb{R}^+$ such that $\forall x, y \in C$ we have $\|x - y\| \leqslant M$. Now if $x, y \in L^{-1}(C)$ then as $L(x), L(y) \in C$ we have

$$\|x - y\|_n \leqslant_{[\text{eq: } 14.156]} N_2 \cdot \|x - y\|^* = N_2 \cdot \|L(x - y)\| = N_2 \cdot \|L(x) - L(y)\| \leqslant N_2 \cdot M$$

proving that

$$L^{-1}(C) \text{ is bounded in } \langle \mathbb{R}^n, \|\|_n\rangle \tag{14.159}$$

Using Heine-Borel [see theorem: 14.273] on [eqs: 14.158, 14.159] proves that $L^{-1}(C)$ is compact in $\mathcal{T}_{\|\|_n}$. As by [eq: 14.155] $L: \mathbb{R}^n \rightarrow X$ is continuous [using $\mathcal{T}_{\|\|_n}$ and $\mathcal{T}_{\|\|}$ on $\mathbb{R}^n$ and X] it follows by [theorem: 14.229] that $C \underset{[\text{theorem: } 2.55]}{=} L(L^{-1}(C))$ is compact. Hence we have proved that

$$C \text{ is closed and bounded} \Rightarrow C \,\text{is compact} \tag{14.160}$$

Combining [eqs: 14.157, 14.160] proves that

$$C \text{ is compact} \Leftrightarrow C \text{ is closed and bounded}$$

□

A application of the above theorem is the following

Corollary 14.277. *Let $\langle \mathbb{C}, |\,|\rangle$ be the complex space equipped with the absolute value norm $|\,|$ and $C \subseteq \mathbb{C}$ then*

$$C \textit{ is compact} \Leftrightarrow C \textit{ is closed and bounded}$$

Proof. Using [example: 11.168] there exist a linear isomorphism between $\mathbb{R}^2$ and $\mathbb{C}$, hence applying the previous theorem [see 14.276] proves the theorem. □

More general we have

Theorem 14.278. *Let $\langle X, \|\|\rangle$ be a finite dimensional normed space over $\mathbb{R}$ and $C \subseteq X$ then*

$$C \textit{ is compact} \Leftrightarrow C \textit{ is closed and bounded}$$

Proof. Using [theorem: 14.170] there exist a isomorphism between X and $\mathbb{R}^n$ where $n = \dim(X)$, hence applying the previous theorem [see 14.276] proves the theorem. □

Theorem 14.279. *Let $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$ be finite dimensional spaces over $\mathbb{R}$ then*

$$\operatorname{Hom}(X,Y)=L(X,Y)$$

in other words every linear mapping is continuous.

Proof. Using [theorem: 14.170] there exist linear isomorphisms $\varphi_1: \mathbb{R}^n \rightarrow X$, $\varphi_2: \mathbb{R}^m \rightarrow Y$ and norms $\|\|_{\langle n\rangle}$, $\|\|_{\langle m\rangle}$ on $\mathbb{R}^n$, $\mathbb{R}^m$ such that

$$\varphi_1: \mathbb{R}^n \rightarrow X \text{ is a homeomorphism using the norms } \|\|_{\langle n\rangle}, \|\|_X \text{ on } \mathbb{R}^n, X \tag{14.161}$$

$$\varphi_2: \mathbb{R}^n \rightarrow Y \text{ is a homeomorphism using the norms } \|\|_{\langle m\rangle}, \|\|_Y \text{ on } \mathbb{R}^n, Y \tag{14.162}$$

Let $L \in \operatorname{Hom}(X,Y)$ and define

$$L': \mathbb{R}^n \rightarrow \mathbb{R}^m \text{ by } L' = \varphi_2^{-1} \circ L \circ \varphi_1$$

Then as φ_1, φ_2^{-1} and L are linear mappings it follows from [theorem: 11.170] that

$$L': \mathbb{R}^n \rightarrow \mathbb{R}^m \text{ is a linear mapping}$$

Using [theorem: 14.175]

$$L': R^n \rightarrow R^m \text{ is linear continuous using the maximum norms } \|\|_n, \|\|_m$$

As by [theorem: 14.274] it follows that $\mathcal{T}_{\|\|_{\langle n\rangle}} = \mathcal{T}_{\|\|_n}$ and $T_{\|\|_{\langle m\rangle}} = \mathcal{T}_m$ proving that

$$L': R^n \rightarrow R^m \text{ is linear continuous using the norms } \|\|_{\langle n\rangle}, \|\|_{\langle m\rangle} \tag{14.163}$$

Using [theorem: 14.185] on [eqs: 14.161, 14.162 and 14.163] that

$$\varphi_2 \circ L' \circ \varphi_1^{-1} \in L(X,Y)$$

Finally, as $\varphi_2 \circ L' \circ \varphi_1^{-1} = \varphi_2 \circ \varphi_2^{-1} \circ L \circ \varphi_1 \circ \varphi_1^{-1} = \operatorname{Id}_Y \circ L \circ \operatorname{Id}_X = L$, it follows from the above that

$$L \in L(X,Y)$$

□

14.9 Convergence

Convention 14.280. For simplicity we use the following convention if $N \in \mathbb{N}_0$

$$\begin{array}{lll} \forall n \geqslant N & \text{is a shorthand for} & \forall n \in \mathbb{N}_0 \text{ with } n \geqslant N \\ \forall n > N & \text{is a shorthand for} & \forall n \in \mathbb{N}_0 \text{ with } n > N \end{array}$$

14.9.1 Sequences and limits

Definition 14.281. *Let X be a set, $k \in \mathbb{N}_0$ then a family $\{x_i\}_{i \in \{k,\ldots,\infty\}} \subseteq X$ is called a* ***sequence in X****.*

A special kind of sequence in a partial ordered set are increasing or decreasing sequences.

Definition 14.282. *Let $\langle X, \leqslant\rangle$ be a partial ordered set, $k \in \mathbb{N}_0$ and $\{x_i\}_{i \in \{k,\ldots,\infty\}} \subseteq X$ a sequence in X then we say:*

1. *$\{x_i\}_{i \in \{k,\ldots,\infty\}}$ is* ***increasing*** *if $\forall i \in \{k,\ldots,\infty\}$ $x_i \leqslant x_{i+1}$.*
2. *$\{x_i\}_{i \in \{k,\ldots,\infty\}}$ is* ***strictly increasing*** *if $\forall i \in \{k,\ldots,\infty\}$ $x_i < x_{i+1}$.*
3. *$\{x_i\}_{i \in \{k,\ldots,\infty\}}$ is* ***decreasing*** *if $\forall i \in \{k,\ldots,\infty\}$ $x_{i+1} \leqslant x_i$.*
4. *$\{x_i\}_{i \in \{k,\ldots,\infty\}}$ is* ***strictly decreasing*** *if $\forall i \in \{k,\ldots,\infty\}$ $x_{i+1} < x_i$.*

Using mathematical induction on the above definition gives the following theorem.

Theorem 14.283. *Let $\langle X, \leqslant\rangle$ be a partial ordered set, $k \in \mathbb{N}_0$ and $\{x_i\}_{i \in \{k,\ldots,\infty\}} \subseteq X$ a sequence in X then*

1. *If $\{x_i\}_{i \in \{k,\ldots,\infty\}}$ is* ***increasing*** *then $\forall n, m \in \{k,\ldots,\infty\}$ with $n \leqslant m$ we have $x_n \leqslant x_m$.*
2. *If $\{x_i\}_{i \in \{k,\ldots,\infty\}}$ is* ***strictly increasing*** *then $\forall n, m \in \{k,\ldots,\infty\}$ with $n < m$ we have $x_n < x_m$.*

3. *If* $\{x_i\}_{i\in\{k,\ldots,\infty\}}$ *is* ***decreasing*** *then* $\forall n,m\in\{k,\ldots,\infty\}$ *with* $n\leqslant m$ *we have* $x_m\leqslant x_n$.
4. *If* $\{x_i\}_{i\in\{k,\ldots,\infty\}}$ *is* ***strictly decreasing*** *then* $\forall n,m\in\{k,\ldots,\infty\}$ *with* $n<m$ *we have* $x_m<x_n$.

Proof.

1. Let $n\in\{k,\ldots,n\}$ and take

$$S_n=\{m\in\mathbb{N}_0|x_n\leqslant x_{n+m}\}$$

then we have:

$0\in S_n$. As $x_n=x_{n+0}$ it follows that $0\in S_n$

$m\in S_n\Rightarrow m+1\in S_n$. As $m\in S_n$ we have $x_n\leqslant x_{n+m}$ which as $x_{n+m}\leqslant x_{(n+m)+1}$ proves that $x_n\leqslant x_{(n+m)+1}$, hence $m+1\in S_n$.

Using mathematical induction $S_n=\mathbb{N}_0$ so if $n\in\{k,\ldots,\infty\}$ and $m\in\{k,\ldots,\infty\}$ with $n\leqslant m$ then $m-n\in\mathbb{N}_0=S_n$ so that $x_n\leqslant x_{n+(m-n)}=x_m$.

2. Let $n\in\{k,\ldots,n\}$ and take

$$S_n=\{m\in\mathbb{N}|x_n<x_{n+m}\}$$

then we have:

$1\in S_n$. As $x_n<x_{n+1}$ it follows that $1\in S_n$

$m\in S_n\Rightarrow m+1\in S_n$. As $m\in S_n$ we have $x_n<x_{n+m}$ which as $x_{n+m}<x_{(n+m)+1}$ proves that $x_n<x_{(n+m)+1}$, hence $m+1\in S_n$.

Using mathematical induction $S_n=\mathbb{N}$ so if $n\in\{k,\ldots,\infty\}$ and $m\in\{k,\ldots,\infty\}$ with $n<m$ then $m-n\in\mathbb{N}=S_n$ so that $x_n<x_{n+(m-n)}=x_m$.

3. Let $n\in\{k,\ldots,n\}$ and take

$$S_n=\{m\in\mathbb{N}_0|x_{n+m}\leqslant x_n\}$$

then we have:

$0\in S_n$. As $x_n=x_{n+0}$ it follows that $0\in S_n$

$m\in S_n\Rightarrow m+1\in S_n$. As $m\in S_n$ we have $x_{n+m}\leqslant x_n$ which as $x_{(n+m)+1}\leqslant x_m$ proves that $x_{(n+m)+1}\leqslant x_n$, hence $m+1\in S_n$.

Using mathematical induction $S_n=\mathbb{N}_0$ so if $n\in\{k,\ldots,\infty\}$ and $m\in\{k,\ldots,\infty\}$ with $n\leqslant m$ then $m-n\in\mathbb{N}_0=S_n$ so that $x_m=x_{n+(m-n)}\leqslant x_n$.

4. Let $n\in\{k,\ldots,n\}$ and take

$$S_n=\{m\in\mathbb{N}|x_{n+m}<x_n\}$$

then we have:

$1\in S_n$. As $x_{n+1}<x_n$ it follows that $0\in S_n$

$m\in S_n\Rightarrow m+1\in S_n$. As $m\in S_n$ we have $x_{n+m}<x_n$ which as $x_{(n+m)+1}<x_m$ proves that $x_{(n+m)+1}<x_n$, hence $m+1\in S_n$.

Using mathematical induction $S_n=\mathbb{N}$ so if $n\in\{k,\ldots,\infty\}$ and $m\in\{k,\ldots,\infty\}$ with $n<m$ then $m-n\in\mathbb{N}=S_n$ so that $x_m=x_{n+(m-n)}<x_n$. □

Corollary 14.284. *Let* $k\in\mathbb{N}_0$, $\{A_i\}_{i\in\{k,\ldots,\infty\}}$ *be a sequence of sets then we have*

1. *If* $\forall i\in\{k,\ldots,\infty\}$ $A_i\subseteq A_{i+1}$ *then* $\forall n,m\in\{k,\ldots,\infty\}$ *with* $n\leqslant m$ *we have* $A_n\subseteq A_m$
2. *If* $\forall i\in\{k,\ldots,\infty\}$ $A_{i+1}\subseteq A_i$ *then* $\forall n,m\in\{k,\ldots,\infty\}$ *with* $n\leqslant m$ *we have* $A_m\subseteq A_n$

Proof. This follows from [theorem: 3.33] and [theorem: 14.283]. □

We introduce now the concept of convergence of a sequence.

Definition 14.285. *Let* $\langle X,\mathcal{T}\rangle$ *be a topological space,* $k\in\mathbb{N}_0$, $x\in X$ *and* $\{x_i\}_{i\in\{k,\ldots,\infty\}}\subseteq X$ *a sequence then we say that* $\{x_i\}_{i\in\{k,\ldots,\infty\}}$ *is* ***convergent to x*** *if* $\forall U\in\mathcal{T}$ *with* $x\in X$ *there exist a* $N\in\{k,\ldots,\infty\}$ *such that* $\forall n\geqslant N$ *we have* $x_n\in U$. *We call* $\{x_i\}_{i\in\{k,\ldots,\infty\}}$ *a* ***convergent sequence***.

We show now that convergence is independent of the start point of the domain of the family.

Theorem 14.286. *Let $\langle X, \mathcal{T}\rangle$ be a topological space, $k \in \mathbb{N}_0$, $x \in X$ and $\{x_i\}_{i \in \{k, \ldots, \infty\}} \subseteq X$ a sequence then we have:*

1. *$\forall m \in \{k, \ldots, \infty\}$ we have $\{x_i\}_{i \in \{k, \ldots, \infty\}}$ converges to $x \Leftrightarrow \{x_i\}_{i \in \{m, \ldots, \infty\}}$ converges to x*
2. *$\forall m \in \mathbb{N}_0$ if $\{x_i\}_{i \in \{k, \ldots, \infty\}}$ converges to x then $\{x_{i+m}\}_{i \in \{k, \ldots, \infty\}}$ converges to x*

Proof.

1. We have

   ⇒. Let $U \in \mathcal{T}$ such that $x \in U$ then, as $\{x_i\}_{i \in \{k, \ldots, \infty\}}$ converges to x, there exist a $N \in \{k, \ldots, \infty\}$ such that $\forall n \geqslant N$ we have $x_n \in U$. Hence if we take $M = \max(N, m)$ then $M \in \{m, \ldots, \infty\}$and we have $\forall n \geqslant M$ that $x_n \in U$. So $\{x_i\}_{i \in \{m, \ldots, \infty\}}$ converges to x.

   ⇐. Let $U \in \mathcal{T}$ such that $x \in U$ then as $\{x_i\}_{i \in \{m, \ldots, \infty\}}$ converges to x there exist a $N \in \{m, \ldots, \infty\}$ such that $\forall n \geqslant M$ we have $x_n \in U$. As $m \in \{k, \ldots, \infty\}$ we have that $N \in \{k, \ldots, \infty\}$ hence $\{x_i\}_{i \in \{k, \ldots, \infty\}}$ converges to x.

2. Let $U \in \mathcal{T}$ such that $x \in U$ then as $\{x_i\}_{i \in \{k, \ldots, \infty\}}$ converges to x there exist a $N \in \{k, \ldots, \infty\}$ such that $\forall n \geqslant N$ we have $x_n \in U$. As $N \leqslant n \leqslant n + m$ we have that $x_{n+m} \in U$ proving that $\{x_{n+m}\}_{n \in \{k, \ldots, \infty\}}$ converges to x. □

Theorem 14.287. *Let $\langle X, \mathcal{T}\rangle$ be a Hausdorff topological space, $k \in \mathbb{N}_0$, $x, y \in X$ and $\{x_i\}_{i \in \{k, \ldots, \infty\}} \subseteq X$ a sequence that is* ***convergent to x and y*** *then* $\boldsymbol{x = y}$.

Proof. Assume that $x \neq y$ then as $\mathcal{T}$ is Hausdorff there exist a $U, V \in \mathcal{T}$ such that $x \in U$, $y \in V$ and $U \bigcap V = \varnothing$. By the definition of convergence there exist a $N_x, N_y \in \{k, \ldots, \infty\}$ such that $\forall n \geqslant N_x$ $x_n \in U$ and $\forall n \geqslant N_n$ we have $x_n \in V$. So for $N = \max(N_x, N_y)$ we have $x_N \in U \bigcap V$ contradicting $U \bigcap V = \varnothing$. Hence we must have

$$x = y$$

□

Definition 14.288. *Let $\langle X, \mathcal{T}\rangle$ be a Hausdorff topological space, $k \in \mathbb{N}_0$ and $\{x_i\}_{i \in \{k, \ldots, \infty\}} \subseteq X$ a sequence that is convergent to the* ***unique*** *$x \in X$ then we note this x as*

$$\lim_{n \to \infty} x_n$$

and say that x is the ***limit*** *of $\{x_i\}_{i \in \{k, \ldots, \infty\}}$. So if we say that*

$$\lim_{n \to \infty} x_n = x$$

then we mean that $\{x_n\}_{n \in \{k, \ldots, \infty\}}$ is convergent to $x \in X$.

We have the following equivalent definition for the limit in a metric space.

Theorem 14.289. *Let $\langle X, d\rangle$ be a metric space [which by [theorem: 14.201 is Hausdorff], $k \in \mathbb{N}_0$, $x \in X$ and $\{x_i\}_{i \in \{k, \ldots, \infty\}} \subseteq X$ then*

$$\lim_{n \to \infty} x_n = x \Leftrightarrow \forall \varepsilon \in \mathbb{R}^+ \ \exists N \in \{k, \ldots, \infty\} \text{ such that } \forall n \in \mathbb{N} \text{ with } N \leqslant n \ d(x, x_n) < \varepsilon$$

Further if $\langle X, \|\|\rangle$ is a normed space we have as $d(x, x_n) = \|x_n - x\|$ that

$$\lim_{n \to \infty} x_n = x \Leftrightarrow \forall \varepsilon \in \mathbb{R}^+ \ \exists N \in \{k, \ldots, \infty\} \text{ such that } \forall n \in \mathbb{N} \text{ with } N \leqslant n \ \|x_n - x\| < \varepsilon$$

Proof. If $\langle X, d\rangle$ is a metric space then we have

⇒. Let $\varepsilon \in \mathbb{R}^+$ then as $x \in B_d(x, \varepsilon) \in \mathcal{T}$ we have from $\lim_{n \to \infty} x_n = x$ that there exist a $N \in \{k, \ldots, \infty\}$ such that $\forall n \geqslant N$ we have $x_n \in B_d(x, \varepsilon)$, hence $d(x, x_n) < \varepsilon$.

⇐. Let $U \in \mathcal{T}$ be such that $x \in U$ then by [theorem: 14.64] there exist a $\varepsilon \in \mathbb{R}^+$ such that $x \in B_d(x, \varepsilon) \subseteq U$. By the hypothesis there exist a $N \in \{k, \ldots, \infty\}$ such that $\forall n \geqslant N$ we have $d(x, x_n) < \varepsilon$, hence $x \in B_d(x, \varepsilon) \subseteq U$, proving that $\lim_{n \to \infty} x_n = x$.

If $\langle X, \|\|\rangle$ is a normed space then we have

⇒. Let $\varepsilon \in \mathbb{R}^+$ then as $x \in B_{\|\|}(x, \varepsilon) \in \mathcal{T}$ we have from $\lim_{n \to \infty} x_n = x$ that there exist a $N \in \{k, \ldots, \infty\}$ such that $\forall n \geqslant N$ we have $x_n \in B_{\|\|}(x, \varepsilon)$, hence $\|x_n - x\| < \varepsilon$.

⇐. Let $U \in \mathcal{T}$ be such that $x \in U$ then by [theorem: 14.93] there exist a $\varepsilon \in \mathbb{R}^+$ such that $x \in B_{\|\|}(x, \varepsilon) \subseteq U$. By the hypothesis there exist a $N \in \{k, \ldots, \infty\}$ such that $\forall n \geqslant N$ we have $\|x_n - x\| < \varepsilon$, hence $x \in B_{\|\|}(x, \varepsilon) \subseteq U$, proving that $\lim_{n \to \infty} x_n = x$. □

We can reformulate the above theorem to a limit to 0 in $\mathbb{R}$.

Theorem 14.290. *If $\langle X, d\rangle$ is a metric space, $k\in\mathbb{N}_0$, $x\in X$ and $\{x_n\}_{n\in\{k,\ldots,\infty\}}\subseteq X$ a sequence then we have*

$$\lim_{n\rightarrow\infty} x_n = x \text{ [defined in } \langle X,d\rangle] \Leftrightarrow \lim_{n\rightarrow\infty} d(x,x_n)=0 \text{ [defined in } \langle\mathbb{R},\|\rangle]$$

If $\langle X,\|\|\rangle$ is a normed space, $k\in\mathbb{N}_0$, $x\in X$ and $\{x_n\}_{n\in\{k,\ldots,\infty\}}\subseteq X$ a sequence then we have

$$\lim_{n\rightarrow\infty} x_n = x \text{ [defined in } \langle X,\|\|\rangle] \Leftrightarrow \lim_{n\rightarrow\infty} \|x_n - x\|=0 \text{ [defined in } \langle\mathbb{R},\|\rangle]$$

Proof. If $\langle X, d\rangle$ is a metric space then we have

- **⇒.** Let $\varepsilon\in\mathbb{R}^+$ then, as $\lim_{n\rightarrow\infty} x_n = x$, there exist a $N\in\{k,\ldots,\infty\}$ such that $\forall n\geqslant N$ we have $d(x,x_n)<\varepsilon$. Hence $|d(x,x_n)-0|=|d(x,x_n)|=d(x,x_n)<\varepsilon$ proving that $\{d(x,x_n)\}_{n\in\{k,\ldots,\infty\}}$ converges to 0 or that

$$\lim_{n\rightarrow\infty} d(x,x_n)=0$$

- **⇐.** Let $\varepsilon\in\mathbb{R}^+$ then there exist a $N\in\{k,\ldots,\infty\}$ such that $\forall n\geqslant N$ we have

$$d(x,x_n)=|d(x,x_n)|=|d(x,x_n)-0|<\varepsilon$$

  proving that

$$\lim_{n\rightarrow\infty} x_n = x$$

If $\langle X,\|\|\rangle$ is a normed space then we have

- **⇒.** Let $\varepsilon\in\mathbb{R}^+$ then, as $\lim_{n\rightarrow\infty} x_n = x$, there exist a $N\in\{k,\ldots,\infty\}$ such that $\forall n\geqslant N$ $\|x_n-x\|<\varepsilon$. Hence $|\|x_n-x\|-0|=|\|x_n-x\||=\|x_n-x\|<\varepsilon$ proving that $\{\|x_n-x\|\}_{n\in\{k,\ldots,\infty\}}$ converges to 0 or that

$$\lim_{n\rightarrow\infty} \|x_n-x\|=0$$

- **⇐.** Let $\varepsilon\in\mathbb{R}^+$ then there exist a $N\in\{k,\ldots,\infty\}$ such that $\forall n\geqslant N$ we have

$$\|x_n-x\|=|\|x_n-x\||=|\|x_n-x\|-0|<\varepsilon$$

  proving that

$$\lim_{n\rightarrow\infty} x_n = x$$

□

Example 14.291. If $\langle X,\|\|\rangle$ is a normed space, $k\in\mathbb{N}_0$, $x\in X$ then we have for $\{x_i\}_{i\in\{k\}}\subseteq X$ defined by $x_i=x$ that

$$\lim_{n\rightarrow\infty} x_n = x$$

Proof. Let $\varepsilon\in\mathbb{R}^+$ and take $N=k$ then we have $\forall n\geqslant N$ that

$$\|x_n-x\|=\|x-x\|=0<\varepsilon$$

proving that $\lim_{n\rightarrow\infty} x_n = x$. □

Example 14.292. If $b\in\mathbb{R}$ with $0\leqslant b<1$ then $\{b^n\}_{n\in\mathbb{N}}$ converges to 0.

Proof. We have to consider two cases for b

- $\boldsymbol{b=0}$**.** Then $\forall n\in\mathbb{N}$ we have $b^n=0$ so that by [example: 14.291] $\lim_{n\rightarrow\infty} b^n=0$.
- $\boldsymbol{0<b}$**.** Let $\varepsilon\in\mathbb{R}^+$ then by [theorem: 10.63] there exist a $N\in\mathbb{N}$ such that $\forall n\geqslant N$ we have $0<b^n<\varepsilon$ hence $|b^n-0|=|b^n|=b^n<\varepsilon$, proving that $\lim_{n\rightarrow\infty} b^n=0$. □

Example 14.293. If $a\in\mathbb{R}$, $k\in\mathbb{N}_0$ with $0<a+k$ so that $\left\{\frac{1}{a+n}\right\}_{n\in\{k,\ldots,\infty\}}\subseteq\mathbb{R}$ is well defined then

$$\lim_{n\rightarrow\infty}\frac{1}{a+n}=0$$

Proof. Let $\varepsilon\in\mathbb{R}^+$ then by [theorem: 10.30] there exist a $N\in\mathbb{N}$ such that $\max\left(\frac{1}{\varepsilon}-a,-a\right)<N$ so if $N\leqslant n$ then we have $\frac{1}{\varepsilon}-a<N\leqslant n\Rightarrow \frac{1}{\varepsilon}<n+a\Rightarrow\frac{1}{a+n}<\varepsilon$ and $-a<n\Rightarrow 0<a+n$. Hence $0<\frac{1}{a+n}<\varepsilon$, so that

$$\left|\frac{1}{a+n}-0\right|=\left|\frac{1}{a+n}\right|=\frac{1}{a+n}<\varepsilon$$

proving

$$\lim_{n\rightarrow\infty}\frac{1}{a+n}=0$$

□

Example 14.294. Let $k\in\mathbb{N}_0$, $a,b\in\mathbb{R}$ with $0<b+k$ so that $\left\{\frac{a+n}{b+n}\right\}_{n\in\{k,\ldots,\infty\}}$ is well defined then

$$\lim_{n\rightarrow\infty}\frac{a+n}{b+n}=1$$

Proof. Let $\varepsilon\in\mathbb{R}^+$ then by the Archimedan property of the real numbers [see theorem: 10.30] there exist a $M\in\{k,\ldots,\infty\}$ such that $\frac{|a-b|}{\varepsilon}-b<M$. Take $N=\max(k,M)$ then for $\forall n\geqslant N$ we have

$$\frac{|a-b|}{\varepsilon}-b<n\Rightarrow\frac{|a-b|}{\varepsilon}<n+b\underset{0<b+k<b+n}{\Rightarrow}\frac{|a-b|}{b+n}<\varepsilon\Rightarrow\left|\frac{a-b}{b+n}\right|<\varepsilon$$

Next

$$\left|\frac{a+n}{b+n}-1\right|=\left|\frac{a+n-(b+n)}{b+n}\right|=\left|\frac{a-b}{b+n}\right|<\varepsilon$$

proving that

$$\lim_{n\rightarrow\infty}\frac{a+n}{b+n}=1$$

□

14.9.2 Properties of the limit

Theorem 14.295. *Let $k\in\mathbb{N}_0$, $\langle X,\|\|\rangle$ be a normed space, $\{s_i\}_{i\in\{k,\ldots,\infty\}}\subseteq\mathbb{R}$ such that $\lim_{i\rightarrow\infty}s_i=0$ exist and $\{x_i\}_{i\in\{k,\ldots,\infty\}}\subseteq X$ satisfies that $\forall i\in\{k,\ldots,\infty\}$ $\|x_i\|<s_i$ then $\lim_{i\rightarrow\infty}x_i=0$ exists.*

Proof. Let $\varepsilon\in\mathbb{R}^+$ then there exist a $N\in\{k,\ldots,\infty\}$ such that $\forall n\geqslant N$ $|s_n|=|s_n-0|<\varepsilon$. Hence if $n\geqslant N$ we have $\|x_n-0\|=\|x\|\leqslant s_n<\varepsilon$ proving that

$$\lim_{i\rightarrow\infty}x_i=0\text{ exists}$$

□

Theorem 14.296. *Let $\langle X,\|\|\rangle$ be a normed space, $k\in\mathbb{N}_0$ and $\{x_n\}_{n\in\{k,\ldots,\infty\}}\subseteq X$ a sequence such that*

$$\lim_{n\rightarrow\infty}x_n=x$$

then for $a\in X$ we have for $\{x_n+a\}_{n\in\{k,\ldots,\infty\}}\subseteq X$ that

$$\lim_{n\rightarrow\infty}(x_n+a)=x+a$$

In other words

$$\lim_{n\rightarrow\infty}(x_n+a)=\lim_{n\rightarrow\infty}x_n+a$$

Proof. As $\lim_{n\rightarrow\infty}x_n=x$ there exist a $N\in\{k,\ldots,\infty\}$ such that $\forall n\geqslant N$ we have $\|x_n-x\|<\varepsilon$. Hence we have

$$\|(x_n+a)-(x+a)\|=\|x_n-x\|<\varepsilon$$

proving that

$$\lim_{n\rightarrow\infty}(x_n+a)=x+a$$

□

Theorem 14.297. *Let $\langle X,\|\|\rangle$ be a normed space, $k,l\in\mathbb{N}_0$ and $\{x_n\}_{n\in\{k,\ldots,\infty\}}\subseteq X$, $\{y_n\}_{n\in\{k,\ldots,\infty\}}\subseteq Y$ sequences such that*

$$\lim_{n\rightarrow\infty}x_n=x\text{ and }\lim_{n\rightarrow\infty}y_n=y$$

then for $\{x_n + y_n\}_{n\in\{\max(k,l),\ldots,\infty\}} \subseteq X$ *we have*

$$\lim_{n\rightarrow\infty} (x_n + y_n) = x + y$$

In other words

$$\lim_{n\rightarrow\infty} (x_n + y_n) = \lim_{n\rightarrow\infty} x_n + \lim_{n\rightarrow\infty} y_n$$

Proof. As $\lim_{n\rightarrow\infty} x_n = x$ and $\lim_{n\rightarrow\infty} y_n = y$ there exists $N_x \in \{k,\ldots,\infty\}$, $N_y \in \{l,\ldots,\infty\}$ such that $\forall n \geqslant N_1$ we have $\|x_n - x\| < \frac{\varepsilon}{2}$ and $\forall n \in N_2$ we have $\|y_n - y\| < \frac{\varepsilon}{2}$. Hence if we take $N = \max(N_1, N_2)$ then $\forall n \geqslant N$ we have

$$\|(x_n + y_n) - (x + y)\| \leqslant \|x_n - x\| + \|y_n - y\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

proving that

$$\lim_{n\rightarrow\infty} (x_n + y_n) = x + y$$

□

Theorem 14.298. *Let* $\langle X, \|\|\rangle$ *be a normed space,* $k \in \mathbb{N}_0$ *and* $\{x_n\}_{n\in\{k,\ldots,\infty\}} \subseteq X$ *a sequence such that*

$$\lim_{n\rightarrow\infty} x_n = x$$

then given $\alpha \in \mathbb{K}$ *we have for the sequence* $\{\alpha \cdot x_n\}_{n\in\{k,\ldots,\infty\}} \subseteq X$ *that*

$$\lim_{n\rightarrow\infty} (\alpha \cdot x_n) = \alpha \cdot x$$

In other words

$$\lim_{n\rightarrow\infty} (\alpha \cdot x_n) = \alpha \cdot \lim_{n\rightarrow\infty} x_n$$

Proof. For $\alpha \in \mathbb{K}$ we have either:

$\boldsymbol{\alpha = 0}$**.** Then $\forall n \in \{k,\ldots,\infty\}$ we have $\alpha \cdot x_n = 0$ so that by [example: 14.291]

$$\lim_{n\rightarrow\infty} (\alpha \cdot x_n) = 0 = 0 \cdot x = \alpha \cdot x$$

$\boldsymbol{\alpha \neq 0}$**.** As $\lim_{n\rightarrow\infty} x_n = x$ there exist a $N \in \{k,\ldots,\infty\}$ such that $\forall n \geqslant N$ we have $\|x_n - x\| < \frac{\varepsilon}{|\alpha|}$. Hence we have

$$\|(\alpha \cdot x_n) - (\alpha \cdot x)\| = \|\alpha(x_n - x)\| = |\alpha| \cdot \|x_n - x\| < |\alpha| \cdot \frac{\varepsilon}{|\alpha|} = \varepsilon$$

proving that

$$\lim_{n\rightarrow\infty} (\alpha \cdot x_n) = \alpha \cdot x$$

□

Theorem 14.299. *Let* $k \in \mathbb{N}_0$ *and* $\{z_n\}_{n\in\{k,\ldots,\infty\}} \subseteq \mathbb{C}$ *a sequence of complex numbers then if* $||$ *is the absolute value norm on* $\mathbb{R}$ *and* $||_c$ *is the canonical norm on* $\mathbb{C}$ *[see definition: 10.82] then we have that*

$$\{z_n\}_{n\in\{k,\ldots,\infty\}} \text{ converges in } \langle\mathbb{C}, ||_c\}$$

$$\Updownarrow$$

$$\{\mathrm{Re}(z_n)\}_{n\in\{k,\ldots,\infty\}} \text{ and } \{\mathrm{Img}(z_n)\}_{n\in\{k,\ldots,\infty\}} \text{ converges in } \langle\mathbb{R}, ||\rangle$$

Further if $\{z_n\}_{n\in\{k,\ldots,\infty\}}$ *converges then*

$$\lim_{n\rightarrow\infty} z_n = \lim_{n\rightarrow\infty} \mathrm{Re}(z_n) + \lim_{n\rightarrow\infty} \mathrm{Img}(z_n)$$

Proof. We have

$\Rightarrow$**.** As $\{z_n\}_{n\in\{k,\ldots,\infty\}}$ converges there exist a $z \in \mathbb{C}$ such that $\forall \varepsilon \in \mathbb{R}^+$ there exist a $N \in \{k,\ldots,n\}$ such that

$$\forall n \geqslant N \text{ we have } |z_n - z|_c < \varepsilon$$

Now using [theorem: 10.83] we have that

$$|\mathrm{Re}(z_n) - \mathrm{Re}(z)| = |\mathrm{Re}(z_n - z)| \leqslant |z_n - z|_c < \varepsilon$$

$$|\mathrm{Img}(z_n) - \mathrm{Img}(z)| = |\mathrm{Img}(z_n - z)| < |z_n - z|_c < \varepsilon$$

proving that $\{\mathrm{Re}(z_n)\}_{n\in\{k,\ldots,\infty\}}$ converges to $\mathrm{Re}(x)$ and $\{\mathrm{Img}(z_n)\}_{n\in\{k,\ldots,\infty\}}$ converges to $\mathrm{Img}(z)$, Hence

$$\lim_{n\to\infty} z_n = \lim_{n\to\infty} \mathrm{Re}(z_n) + \lim_{n\to\infty} \mathrm{Img}(z_n)$$

⇐. As $\{\mathrm{Re}(z_n)\}_{n\in\{k,\ldots,\infty\}}$ and $\{\mathrm{Img}(z_n)\}_{n\in\{k,\ldots,\infty\}}$ converges there exists $x,y\in\mathbb{R}$ such that $\forall\varepsilon\in\mathbb{R}^+$ there exists $N_1,N_2\in\{k,\ldots,\infty\}$ such that $\forall n\geqslant N_1$ $|\mathrm{Re}(z_n)-x|<\frac{\varepsilon}{2}$ and $\forall m\geqslant N_2$ $|\mathrm{Img}(z_n)-y|<\frac{\varepsilon}{2}$. Hence if we take $z=x+i\cdot y$ and $m\geqslant\max(N_1,N_2)$ then we have that

$$|z_n-z|_c \leqslant_{[\text{theorem: } 10.83]} |\mathrm{Re}(z_n-z)|+|\mathrm{Img}(z_n-z)| = |\mathrm{Re}(z_n)-x|+|\mathrm{Img}(z_n)-y| < \frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon$$

proving that $\{z_n\}_{n\in\{k,\ldots,\infty\}}$ converges to z. □

Theorem 14.300. *Let $k\in\mathbb{N}_0$, $\{x_n\}_{i\in\{n,\ldots,\infty\}}\subseteq\mathbb{C}$ be a convergent sequence of complex numbers then $\{\overline{x_n}\}_{n\in\{k,\ldots,\infty\}}$ is convergent and $\lim_{n\to\infty}\overline{x_n}=\overline{\lim_{n\to\infty}x_n}$*

Proof. Let $x=\lim_{n\to\infty}x_n$ then for $\varepsilon\in\mathbb{R}^+$ there exist a $N\in\{k,\ldots,\infty\}$ such that $\forall n\geqslant N$ we have $|x_n-x|<\varepsilon$. Now $|\overline{x_n}-\overline{x}| \underset{[\text{theorem: } 10.81]}{=} |\overline{x_n-x}| \underset{[\text{theorem: } 10.83]}{=} |x_n-x|<\varepsilon$ proving that

$$\lim_{n\to\infty}\overline{x_n}=\overline{x}=\overline{\lim_{n\to\infty}x_n}$$ □

Theorem 14.301. *Let $k\in\mathbb{N}_0$, $\langle X,\|\|\rangle$ be a normed space and $\{x_n\}_{n\in\{k,\ldots,\infty\}}\subseteq X$ a convergent sequence such that $\lim_{n\to\infty}x_n=x$ then we have for $A\in\mathbb{R}$:*

1. *If $\exists M\in\{k,\ldots,\infty\}$ such that $\forall n\geqslant M$ $\|x_n\|\leqslant A$ then $\|x\|\leqslant A$.*
2. *If $\exists M\in\{k,\ldots,\infty\}$ such that $\forall n\geqslant M$ $A\leqslant\|x_n\|$ then $A\leqslant\|x\|$.*

Proof.

1. We proceed by contradiction. So assume that $A<\|x\|$ then $\varepsilon=\|x\|-A\in\mathbb{R}^+$ hence, as $\lim_{n\to\infty}x_n=x$, there exist a $N\in\{k,\ldots,\infty\}$ such that $\forall n\geqslant N$ $\|x_n-x\|<\varepsilon$ or as $N\leqslant\max(N,M)$

$$\|x\|\leqslant\|x-x_{\max(N,M)}\|+\|x_{\max(N,M)}\|<\varepsilon+\|x_{\max(N,M)}\|\leqslant\varepsilon+A=\|x\|-A+A=\|x\|$$

leading to the contradiction $\|x\|<\|x\|$. Hence we must have $\|x\|\leqslant A$.

2. We proceed by contradiction. So assume that $\|x\|<A$ then $\varepsilon=A-\|x\|\in\mathbb{R}^+$ hence, as $\lim_{n\to\infty}x_n=x$, there exist a $N\in\{k,\ldots,\infty\}$ such that $\forall n\geqslant N$ $\|x_n-x\|<\varepsilon$ or as $N\leqslant\max(N,M)$

$$A\leqslant\|x_{\max(M,N)}\|\leqslant\|x_{\max(M,N)}-x\|+\|x\|<\varepsilon+\|x\|=A-\|x\|+\|x\|=A$$

leading to the contradiction $A<A$. Hence we must have $A\leqslant\|x\|$. □

Theorem 14.302. *Let $\langle X,\|\|\rangle$ be a normed space, $k\in\mathbb{N}_0$ and $\{x_n\}_{n\in\{k,\ldots,\infty\}}\subseteq X$ a convergent sequence then there exist a $K\in\mathbb{R}^+$ such that $\forall n\in\{k,\ldots,\infty\}$ we have $\|x_n\|\leqslant K$.*

Proof. Let $x=\lim_{n\to\infty}x_i$ then as $1\in\mathbb{R}^+$ there exist a $N\in\{k,\ldots,\infty\}$ such that $\forall n\geqslant N$ we have $\|x_n\|=\|x_n-x+x\|\leqslant\|x_n-x\|+\|x\|<1+\|x\|$. Hence if we define

$$K=\max(1+\|x\|,\max(\{x_i|i\in\{k,\ldots,N\}\}))$$

we have

$$\forall n\in\{k,\ldots,\infty\} \text{ that } \|x_n\|\leqslant K$$ □

Theorem 14.303. *Let $k\in\mathbb{N}_0$, $\langle X,\|\|\rangle$ a normed space, $\{x_n\}_{n\in\{k,\ldots,\infty\}}\subseteq X$ a **convergent** sequence in X and $\{s_n\}_{n\in\{k,\ldots,\infty\}}\subseteq\mathbb{R}$ a **convergent** sequence such that $\forall n\in\{k,\ldots,\infty\}$ we have $\|x_n\|\leqslant s_n$ then $\left\|\lim_{n\to\infty}x_n\right\|\leqslant\lim_{n\to\infty}s_n$*

Proof. Let $x=\lim_{n\to\infty}x_n$ and $s=\lim_{n-<\infty}s_n=s$. As $\forall n\in\{k,\ldots,\infty\}$ we have $0\leqslant\|x_n\|<s_n$ it follows from [theorem: 14.301] that

$$0\leqslant s$$

Assume that $s < \|x\|$ then $\varepsilon = \|x\| - s \in \mathbb{R}^+$. Using the definition of a limit there exists N_1, $N_2 \in \{k, \ldots, \infty\}$ such that $\forall n \geqslant N_1$ $\|x_n - x\| < \frac{\varepsilon}{2}$ and $\forall m \geqslant N_2$ $|s_n - s| < \frac{\varepsilon}{2}$. Take $N = \max(N_1, N_2)$ then we have

$$\|x\| \leqslant \|x - x_N\| + \|x_N\| < \frac{\varepsilon}{2} + \|x_N\| \leqslant \frac{\varepsilon}{2} + s_N \leqslant \frac{\varepsilon}{2} + |s_N| \leqslant \frac{\varepsilon}{2} + |s_N - s| + |s| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} + |s| = \varepsilon + s = \|x\|$$

leading to the contradiction $\|x\| < \|x\|$. Hence we must have that $\|x\| \leqslant s$. □

Theorem 14.304. *Let $k \in \mathbb{N}_0$, $\langle X, \|\|\rangle$ a normed space, $\{x_n\}_{n \in \{k, \ldots, \infty\}} \subseteq X$, $\{y_n\}_{n \in \{k, \ldots \infty\}} \subseteq X$ convergent sequences such that $\forall n \in \{k, \ldots, \infty\}$ $\|x_n\| \leqslant \|y_n\|$ then $\left\|\lim_{n\to\infty} x_n\right\| \leqslant \left\|\lim_{n\to\infty} y_n\right\|$*

Proof. Let $x = \lim_{n\to\infty} x_n$, $y = \lim_{n\to\infty} y_n$ and assume that $\|y\| < \|x\|$ so that $\varepsilon = \|x\| - \|y\| \in \mathbb{R}^+$. Find $N_1, N_2 \in \{k, \ldots \infty\}$ such that $\forall n \geqslant N_1$ $\|x_n - x\| < \frac{\varepsilon}{2}$ and $\forall m > N_2$ $\|y_m - y\| < \frac{\varepsilon}{2}$. Take $N = \max(N_1, N_2)$ then we have:

$$\begin{aligned}
\|x\| &\leqslant \|x_N - x\| + \|x_N\| \\
&\leqslant \|x_N - x\| + \|y_N\| \\
&\leqslant \|x_N - x\| + \|y_N - y\| + \|y\| \\
&< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} + \|y\| \\
&= \varepsilon + \|y\| \\
&= \|x\|
\end{aligned}$$

giving the contradiction $\|x\| < \|x\|$. So the assumption is wrong and we must have that $\|x\| \leqslant \|y\|$. □

Theorem 14.305. *Let $k \in \mathbb{N}_0$, $\{x_n\}_{n \in \{k, \ldots, \infty\}} \subseteq [0, \infty[$ a convergent sequence of non negative numbers then $0 \leqslant \lim_{n\to\infty} x_n$*

Proof. Assume that $x = \lim_{i\to\infty} x_i < 0$ then $\varepsilon = -x \in \mathbb{R}^+$. By the definition of a limit there exists a $N \in \{k, \ldots, \infty\}$ such that $|x_N - x| < \frac{\varepsilon}{2}$. As $0 \leqslant x_N \wedge x < 0 \Rightarrow 0 < -x$ 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 14.306. *Let $\{x_n\}_{n \in \{k, \ldots, \infty\}}, \{y_n\}_{n \in \{k, \ldots n\}} \subseteq \mathbb{R}$ be convergent sequences with $\forall n \in \{k, \ldots, \infty\}$ $x_n \leqslant y_n$ then $\lim_{n\to\infty} x_n \leqslant \lim_{n\to\infty} y_n$*

Proof. Using the assumption we have that $\{y_i - x_i\}_{i \in \{k, \ldots, \infty\}} \subseteq [0, \infty[$ and using [theorems: 14.297, 14.298] it follows that $\lim_{n\to\infty} (y_n - x_n) = \lim_{n\to\infty} y_n - \lim_{n\to\infty} x_n$. From [theorem: 14.305] we have that $0 \leqslant \lim_{n\to\infty} (y_n - x_n)$ so that $\lim_{n\to\infty} x_n \leqslant \lim_{n-<\infty} x_n$. □

There is a relation between convergence and continuity of a function.

Theorem 14.307. *If $\langle X, \mathcal{T}_X\rangle$, $\langle Y, \mathcal{T}_Y\rangle$ and $f: X \to Y$ a continuous function at x then*

$$\forall \{x_n\}_{n \in \{k, \ldots, \infty\}} \subseteq X \text{ with } \lim_{n\to\infty} x_n = x \text{ we have that } \lim_{n\to\infty} f_n(x) = f(x)$$

Proof. As f is continuous at x we have for $V \in \mathcal{T}_Y$ with $f(x) \in V$ that there exist a $U \in \mathcal{T}_X$ such that $x \in U$ and $f(x) \in f(U) \subseteq V$. If $\{x_n\}_{n \in \{k, \ldots, n\}} \subseteq X$ satisfies $\lim_{n\to\infty} x_n = x$ then there exist a $N \in \{k, \ldots, \infty\}$ such that $\forall n \geqslant N$ we have $x_n \in U$, hence as $f(U) \subseteq V$ we have $f(x_n) \in V$. So we have

$$\lim_{n\to\infty} f(x_n) = f(x)$$ □

For metric and normed spaces we have also the opposite of the above implication.

Theorem 14.308. *If $\langle X, d_x\rangle$ and $\langle Y, d_Y\rangle$ are metric spaces, $f: X \to Y$ a function then we have*

$$f \text{ is continuous at } x \Leftrightarrow \forall \{x_n\}_{n \in \mathbb{N}} \subseteq X \text{ with } \lim_{n\to\infty} x_n = x \text{ we have } \lim_{n\to\infty} f(x_n) = f(x)$$

***Note* 14.309.** *As a normed space is a metric space the theorem is also applicable to normed spaces.*

Proof.

⇒. This follows from the previous theorem [theorem: 14.307].

⇐. Assume that f is not continuous at x then by [theorem: 14.126] there exist a $\varepsilon > 0$ such that $\forall \delta \in \mathbb{R}^+$ there exist a $y \in X$ with $d_X(x, y) < \delta$ such that $\varepsilon \leqslant d_Y(f(x), f(y))$. Hence $\forall n \in \mathbb{N}_0$ there exist a $x_n \in X$ with $d_X(x, x_n) < \frac{1}{n}$ such that $\varepsilon \leqslant d_Y(f(x), f(x_n))$. Take $\zeta \in \mathbb{R}^+$ then by the Archimedean property [see theorem: 10.29] there exist a $N \in \mathbb{N}$ such that $\frac{1}{N} < \zeta$, hence $\forall n \geqslant N$ we have

$$d_X(x, x_n) < \frac{1}{n} \leqslant \frac{1}{N} < \zeta$$

proving that

$$\lim_{n \to \infty} x_n = x$$

Hence by the hypothesis we have that

$$\lim_{n \to \infty} f(x_n) = x$$

So there exist a $M \in \mathbb{N}$ such that $d_Y(f(x), f(x_M)) < \varepsilon$ contradicting the fact that $\forall n \in \mathbb{N}\ \varepsilon \leqslant d_Y(f(x), f(x_n))$. So the assumption is wrong and we must have that f is continuous at x. □

Corollary 14.310. *Let $k \in \mathbb{N}_0$, $\langle X, \|\|\rangle$ a normed space and $\{x_i\}_{i \in \{k, \ldots, \infty\}} \subseteq X$ a sequence such that $\lim_{n \to \infty} x_n$ exist then $\lim_{n \to \infty} \|x_n\|$ exist and $\lim_{n \to \infty} \|x_n\| = \left\| \lim_{n \to \infty} x_n \right\|$*

Proof. Using [example: 14.153] $\|\|: X \to \mathbb{R}$ defined by $\|\|(x) = \|x\|$ is continuous. So as $\lim_{n \to \infty} x_n$ converges to x we have by the previous theorem [theorem: 14.308] that $\lim_{n \to \infty} \|\|(x_n)$ converges to $\|\|(x) = \|x\|$. In other words we have $\lim_{n \to \infty} \|x_n\| = \left\| \lim_{n \to \infty} x_n \right\|$, □

14.9.3 Sequences of real numbers

In this section we always assume that we work with the normed space $\langle \mathbb{R}, || \rangle$. The basic property that sets $\mathbb{R}$ apart from other normed spaces is the fact that $\langle \mathbb{R}, \leqslant \rangle$ is conditional complete.

For the real numbers there is a nice relation between the supremum and infinum and increasing or decreasing sequences.

Theorem 14.311. *Let $k \in \mathbb{N}_0$ and $\{x_n\}_{n \in \{k, \ldots, \infty\}} \subseteq \mathbb{R}$ then we have*

1. *If $\{x_n\}_{n \in \{k, \ldots, \infty\}}$ is increasing then we have*

$$\sup(\{x_n | n \in \{k, \ldots, \infty\}\}) \text{ exists} \Leftrightarrow \lim_{n \to \infty} x_n \text{ exists}$$

further if the supremum or the limit exists we have that

$$\lim_{n \to \infty} x_n = \sup(\{x_n | n \in \{k, \ldots, \infty\}\})$$

2. *If $\{x_n\}_{n \in \{k, \ldots, \infty\}}$ is decreasing then we have*

$$\inf(\{x_n | n \in \{k, \ldots, \infty\}\}) \text{ exists} \Leftrightarrow \lim_{n \to \infty} x_n \text{ exists}$$

further if the supremum or the limit exists we have that

$$\lim_{n \to \infty} x_n = \inf(\{x_n | n \in \{k, \ldots, \infty\}\})$$

Proof.

1. We have

⇒. Let $x = \sup(\{x_n | n \in \{k, \ldots, \infty\}\})$ and take $\varepsilon \in \mathbb{R}^+$ then $x - \varepsilon < x$ so that there exist a $N \in \{k, \ldots, \infty\}$ such that $x - \varepsilon < x_N$. By [theorem: 14.283] we have $\forall n \leqslant N$ that $x - \varepsilon < x_N \leqslant x_n \leqslant x < x + \varepsilon$, hence $x - x_n < \varepsilon \wedge x_n - x < \varepsilon \Rightarrow |x_n - x| < \varepsilon$, proving that

$$\lim_{n \to \infty} x_n \text{ exists and } \lim_{n \to \infty} x_n = x = \sup(\{x_n | n \in \{k, \ldots, \infty\}\})$$

⇐. Let $x = \lim_{n \to \infty} x_n$ then as $1 \in \mathbb{R}^+$ there exist a $N \in \{k, \ldots, \infty\}$ such that $\forall n \geqslant N$ we have $|x_n - x| < 1 \Rightarrow x_n - x < 1 \Rightarrow x_n < x + 1$. So $\forall n \in \{k, \ldots, \infty\}$ we have that $x_n \leqslant \max(\{x_i | i \in \{1, \ldots, N\} \bigcup \{x + 1\}|\})$ or in other words $\{x_n | n \in \{k, \ldots, \infty\}\}$ is bounded above. Using the conditional completeness of the real numbers [see theorem: 10.18] it follows that

$$\sup(\{x_n | n \in \{k, \ldots, \infty\}\}) \text{ exist}$$

2. We have

⇒. Let $x = \inf(\{x_n | n \in \{k, \ldots, \infty\}\})$ and take $\varepsilon \in \mathbb{R}^+$ then $x < x + \varepsilon$ so that there exist a $N \in \{k, \ldots, \infty\}$ such that $x_N < x + \varepsilon$. By [theorem: 14.283] we have $\forall n \leqslant N$ that $x - \varepsilon < x \leqslant x_n \leqslant x_N < x + \varepsilon$, hence $x - x_n < \varepsilon \wedge x_n - x < \varepsilon \Rightarrow |x_n - x| < \varepsilon$, proving that

$$\lim_{n\to\infty} x_n \text{ exists and } \lim_{n\to\infty} x_n = x = \inf(\{x_n | n \in \{k, \ldots, \infty\}\})$$

⇐. Let $x = \lim_{n\to\infty} x_n$ then as $1 \in \mathbb{R}^+$ there exist a $N \in \{k, \ldots, \infty\}$ such that $\forall n \geqslant N$ we have $|x_n - x| < 1 \Rightarrow x - x_n < 1 \Rightarrow x - 1 < x_n$. So $\forall n \in \{k, \ldots, \infty\}$ we have that $\min(\{x_i | i \in \{1, \ldots, N\} \bigcup \{x - 1\}|\}) \leqslant x_n$ or in other words $\{x_n | n \in \{k, \ldots, \infty\}\}$ is bounded below. Using the conditional completeness of the real numbers [see theorem: 10.18] it follows that

$$\inf(\{x_n | n \in \{k, \ldots, \infty\}\}) \text{ exist}$$ □

Corollary 14.312. *Let $k \in \mathbb{N}_0$ and $\{x_n\}_{n\in\{k,\ldots,\infty\}} \subseteq \mathbb{R}$ then we have*

1. *If $\{x_n\}_{n\in\{k,\ldots,\infty\}}$ is increasing and there exist a $M \in \mathbb{R}$ and a $N \in \{k, \ldots, \infty\}$ such that $\forall n \geqslant N$ $x_n \leqslant M$ then $\lim_{n\to\infty} x_n$ exist and $\lim_{n\to\infty} x_n \leqslant M$.*
2. *If $\{x_n\}_{n\in\{k,\ldots,\infty\}}$ is decreasing and there exist a $M \in \mathbb{R}$ and a $N \in \{k, \ldots, \infty\}$ such that $\forall n \geqslant N$ $M \leqslant x_n$ then $\lim_{n\to\infty} x_n$ exist and $M \leqslant \lim_{n\to\infty} x_n$.*

Proof.

1. As $\forall n \geqslant N$ $x_n \leqslant M$ we have that $\emptyset \neq \{x_n | n \in \{N, \ldots, \infty\}\} \leqslant M$ it follows from the conditional completeness of $\mathbb{R}$ [see theorem: 10.18] that $\sup(\{x_n \in \{N, \ldots, \infty\}\})$ exist and

$$\sup(\{x_n | n \in \{N, \ldots, \infty\}\}) \leqslant M.$$

Hence by [theorem: 14.311] $\{x_n\}_{n\in\{N,\ldots,\infty\}}$ converges to $\sup(\{x_n \in \{N, \ldots, \infty\}\})$. Which using [theorem: 14.286] proves that $\{x_n\}_{n\in\{k,\ldots,\infty\}}$ converges to $\sup(\{x_n \in \{N, \ldots, \infty\}\})$. In other words

$$\lim_{n\to\infty} x_n \text{ exist and } \lim_{n\to\infty} x_n \leqslant M$$

2. As $\forall n \geqslant N$ $M \leqslant x_n$ we have that $\emptyset \neq \{x_n | n \in \{N, \ldots, \infty\}\} \geqslant M$ it follows from the conditional completeness of $\mathbb{R}$ [see theorem: 10.18] that $\inf(\{x_n \in \{N, \ldots, \infty\}\})$ exist and

$$M \leqslant \inf(\{x_n | n \in \{N, \ldots, \infty\}\}).$$

Hence by [theorem: 14.311] $\{x_n\}_{n\in\{N,\ldots,\infty\}}$ converges to $\inf(\{x_n \in \{N, \ldots, \infty\}\})$. Which using [theorem: 14.286] proves that $\{x_n\}_{n\in\{k,\ldots,\infty\}}$ converges to $\inf(\{x_n \in \{N, \ldots, \infty\}\})$. In other words

$$\lim_{n\to\infty} x_n \text{ exist and } M \leqslant \lim_{n\to\infty} x_n$$ □

We introduce now a extension of the concept of a limit.

Definition 14.313. *Let $k \in \mathbb{N}_0$, $\{x_n\}_{n\in\{k,\ldots,\infty\}} \subseteq \mathbb{R}$ a sequence of real numbers then the limit superior and the limit inferior noted as $\limsup_{n\to\infty} x_i$ and $\liminf_{n\to\infty} x_i$ are defined as follows:*

1. *$\limsup_{n\to\infty} x_n$ exist if*
   a. *$\forall n \in \{k, \ldots, \infty\}$ $\sup(\{x_i | i \in \{n, \ldots, \infty\}\})$ exist*
   b. *$\lim_{n\to\infty} \sup(\{x_i | i \in \{n, \ldots, \infty\}\})$ exist*

   If $\limsup_{n\to\infty} x_n$ exist then $\limsup_{n\to\infty} x_n = \lim_{m\to\infty} \sup(\{x_i | i \in \{n, \ldots, \infty\}\})$
2. *$\liminf_{n\to\infty} x_n$ exist if*
   a. *$\forall n \in \{k, \ldots, \infty\}$ $\inf(\{x_i | i \in \{n, \ldots, \infty\}\})$ exist*
   b. *$\lim_{n\to\infty} \inf(\{x_i | i \in \{n, \ldots, \infty\}\})$ exist*

If $\liminf_{n\rightarrow\infty} x_n$ *exist then* $\liminf_{n\rightarrow\infty} x_n = \lim_{m\rightarrow\infty} \inf(\{x_i|i\in\{n,\ldots,\infty\}\})$

Theorem 14.314. *Let* $k\in\mathbb{N}_0, \{x_i\}_{i\in\{k,\ldots,\infty\}}\subseteq\mathbb{R}$ *such that* $\liminf_{n\rightarrow\infty} x_i$ *and* $\limsup_{n\rightarrow\infty} x_i$ *exists then*

$$\liminf_{n\rightarrow\infty} x_i \leqslant \limsup_{n\rightarrow\infty} x_i$$

Proof. Let $n\in\{k,\ldots,\infty\}$ then we have

$$\inf(\{x_i|i\in\{n,\ldots,\infty\}\})\leqslant x_n\leqslant\sup(\{x_i|i\in\{n,\ldots,\infty\}\})$$

So that by [theorem: 14.306]

$$\liminf_{n\rightarrow\infty} x_i \underset{\text{def}}{=} \lim_{n\rightarrow\infty}\inf(\{x_i|i\in\{n,\ldots,\infty\}\})\leqslant\lim_{n\rightarrow\infty}\sup(\{x_i|i\in\{n,\ldots,\infty\}\})\underset{\text{def}}{=}\limsup_{n\rightarrow\infty} x_i$$

□

We have the following equivalent definition of the lim sup and lim inf of a sequence.

Theorem 14.315. *Let* $k\in\mathbb{N}$, $\{x_n\}_{n\in\{k,\ldots,\infty\}}\subseteq\mathbb{R}$ *a sequence of real numbers then*

1.

$$\limsup_{n\rightarrow\infty} x_n \textit{ exist}$$

$$\Updownarrow$$

$$\forall n\in\{k,\ldots,\infty\}\ \sup(\{x_i|i\in\{n,\ldots,\infty\}\}) \textit{ exist}$$

$$\text{and}$$

$$\inf(\{\sup(\{x_i|i\in\{n,\ldots.,\infty\}\})|n\in\{k,\ldots,\infty\}\}) \textit{ exist}$$

Further if $\limsup_{n\rightarrow\infty} x_n$ *exist then*

$$\limsup_{n\rightarrow\infty} x_n = \inf(\{\sup(\{x_i|i\in\{n,\ldots,\infty\}\})|n\in\{k,\ldots,\infty\}\})$$

2.

$$\liminf_{n\rightarrow\infty} x_n \textit{ exist}$$

$$\Updownarrow$$

$$\forall n\in\{k,\ldots,\infty\}\ \inf(\{x_i|i\in\{n,\ldots,\infty\}\}) \textit{ exist}$$

$$\text{and}$$

$$\sup(\{\inf(\{x_i|i\in\{n,\ldots.,\infty\}\})|n\in\{k,\ldots,\infty\}\}) \textit{ exist}$$

Further if $\liminf_{n\rightarrow\infty} x_n$ *exist then*

$$\liminf_{n\rightarrow\infty} x_n = \sup(\{\inf(\{x_i|i\in\{n,\ldots,\infty\}\})|n\in\{k,\ldots,\infty\}\})$$

Proof.

1. Assume that $\forall n\in\{k,\ldots,\infty\}\ \sup(\{x_i|i\in\{k,\ldots,\infty\}\})$ exist. Then $\forall n\in\{k,\ldots,\infty\}$ we have $\{x_i|i\in\{n+1,\ldots,\infty\}\}\subseteq\{x_i|i\in\{n,\ldots,\infty\}\}$ so that by [theorem: 3.74]

$$\sup(\{x_i|i\in\{n+1,\ldots,\infty\}\})\leqslant\sup(\{x_i|i\in\{n,\ldots,\infty\}\})$$

   proving that $\{\sup(\{x_i|i\in\{n,\ldots,\infty\}\})\}_{n\in\{k,\ldots,\infty\}}$ is a decreasing sequence. The rest of the proof follows then from [definition: 14.313] and [theorem: 14.311].

2. Assume that $\forall n\in\{k,\ldots,\infty\}\ \inf(\{x_i|i\in\{k,\ldots,\infty\}\})$ exist. Then $\forall n\in\{k,\ldots,\infty\}$ we have $\{x_i|i\in\{n+1,\ldots,\infty\}\}\subseteq\{x_i|i\in\{n,\ldots,\infty\}\}$ so that by [theorem: 3.74]

$$\inf(\{x_i|i\in\{n,\ldots,\infty\}\})\leqslant\inf(\{x_i|i\in\{n+1,\ldots,\infty\}\})$$

   proving that $\{\inf(\{x_i|i\in\{n,\ldots,\infty\}\})\}_{n\in\{k,\ldots,\infty\}}$ is a increasing sequence. The rest of the proof follows then from [definition: 14.313] and [theorem: 14.311].

□

We have the following relation between the limit of a sequence and the limit inferior and the limit superior of a sequence.

Theorem 14.316. *Let $k \in \mathbb{N}_0$, $\{x_i\}_{i \in \{k,\ldots,\infty\}} \subseteq \mathbb{R}$ then we have that*

$$\lim_{n\to\infty} x_n \text{ exist} \Leftrightarrow \liminf_{n\to\infty} x_n \text{ and } \limsup_{n\to\infty} x_n \text{ exists and } \liminf_{n\to\infty} x_n = \limsup_{n\to\infty} x_n$$

Further if $\lim_{n\to\infty} x_i$ then

$$\lim_{n\to\infty} x_n = \liminf_{n\to\infty} x_i = \limsup_{n\to\infty} x_n$$

Proof.

⇒. As $\lim_{n\to\infty} x_i$ exist we have by [theorem: 14.302] that there exist a $K \in \mathbb{R}^+$ such that $\forall i \in \{k,\ldots,\infty\}$ $|x_i| \leqslant K \Rightarrow -x_i, x_i \leqslant K \Rightarrow -K \leqslant x_i \leqslant K$. Hence $\forall n \in \{k,\ldots,\infty\}$ $\{x_i | i \in \{n,\ldots,\infty\}\}$ is bounded below by $-K$ and above by K. So by the fact that $\mathbb{R}$ is conditional complete [see theorem: 10.18] it follows that

$$\sup(\{x_i | \in \{n,\ldots,\infty\}\}) \text{ and } \inf(\{x_i | i \in \{n,\ldots,\infty\}\}) \text{ exists}$$

and

$$\inf(\{x_i | i \in \{n,\ldots,\infty\}\}) \leqslant x_n \leqslant K \text{ and } -K \leqslant x_n \leqslant \sup(\{x_i | i \in \{n,\ldots,\infty\}\})$$

Using the above we have that $\{\inf(\{x_i | i \in \{n,\ldots,\infty\}\}) | n \in \{k,\ldots,\infty\}\}$ is bounded above by K and $\{\sup(\{x_i | i \in \{n,\ldots,\infty\}\}) | n \in \{k,\ldots,\infty\}\}$ is bounded below by $-K$. Using [theorem: 10.18] again it follows that $\sup(\{\inf(\{x_i | i \in \{n,\ldots,\infty\}\}) | n \in \{k,\ldots,\infty\}\})$ and $\inf(\{\sup(\{x_i | i \in \{n,\ldots,\infty\}\}) | n \in \{k,\ldots,\infty\}\})$ exists. Using then [theorem: 14.315] it follows that

$$\liminf_{n\to\infty} x_i \text{ and } \limsup_{n\to\infty} x_i \text{ exist} \tag{14.164}$$

and

$$\liminf_{n\to\infty} x_n = \sup(\{\inf(\{x_i | i \in \{n,\ldots,\infty\}\}) | n \in \{k,\ldots,\infty\}\}) \tag{14.165}$$

$$\limsup_{n\to\infty} x_n = \inf(\{\sup(\{x_i | i \in \{n,\ldots,\infty\}\}) | n \in \{k,\ldots,\infty\}\}) \tag{14.166}$$

Let $x = \lim_{n\to\infty} x_n$. Then giving $\varepsilon \in \mathbb{R}^+$ there exist a $n \in \{k,\ldots,\infty\}$ such that $\forall i \in \{n,\ldots,\infty\}$ we have $|x_i - x| < \varepsilon \Rightarrow x - \varepsilon < x_i < x + \varepsilon$. Hence $x - \varepsilon \leqslant \inf(\{x_i | i \in \{n,\ldots,\infty\}\})$ or $x \leqslant \inf(\{x_i | i \in \{n,\ldots,\infty\}\}) + \varepsilon$ and $\sup(\{x_i | i \in \{n,\ldots,\infty\}\}) \leqslant x + \varepsilon$. Hence using [theorem: 10.31] we have that

$$x \leqslant \inf(\{x_i | i \in \{n,\ldots,\infty\}\} \text{ and } \sup(\{x_i | i \in \{n,\ldots,\infty\}\}) \leqslant x$$

hence

$$x \leqslant \sup(\{\inf(\{x_i | i \in \{n,\ldots,\infty\}\}) | n \in \{k,\ldots,\infty\}\}) \underset{[\text{eq: } 14.165]}{=} \liminf_{n\to\infty} x_n$$

and

$$\limsup_{n\to\infty} x_n \underset{[\text{eq: } 14.166]}{=} \inf(\{\sup(\{x_i | i \in \{n,\ldots,\infty\}\}) | n \in \{k,\ldots,\infty\}\}) \leqslant x$$

which as by [theorem: 14.314] $\liminf_{n\to\infty} x_n \leqslant \limsup_{n\to\infty} x_n$ proves that

$$x \leqslant \liminf_{n\to\infty} x_n \leqslant \limsup_{n\to\infty} x_n \leqslant x$$

so that

$$\lim_{n\to\infty} x_n = x = \liminf_{n\to\infty} x_n = \limsup_{n\to\infty} x_n$$

⇐. By the hypothesis $\liminf_{n\to\infty} x_n$ and $\limsup_{n\to\infty} x_n$ exist and $x = \liminf_{n\to\infty} x_n = \limsup_{n\to\infty} x_n$. Let $\varepsilon \in \mathbb{R}^+$. Then $x - \varepsilon < x = \liminf_{n\to\infty} x_n \underset{[\text{theorem: } 14.315]}{=} \sup(\{\inf(\{x_i | i \in \{n,\ldots,\infty\}\}) | n \in \{k,\ldots,\infty\}\})$ so there exist a $N \in \{k,\ldots,\infty\}$ such that $x - \varepsilon < \inf(\{x_i | i \in \{N,\ldots,\infty\}\})$. In other words

$$\forall i \geqslant N \text{ we have } x - \varepsilon < x_i \tag{14.167}$$

Further $\inf(\{\sup(\{x_i | i \in \{n,\ldots,\infty\}\}) | n \in \{k,\ldots,\infty\}\}) = \limsup_{n\to\infty} x_n = x < x + \varepsilon$ so there exist a $M \in \{k,\ldots,\infty\}$ such that $\sup(\{x_i | i \in \{M,\ldots,\infty\}\}) < x + \varepsilon$. In other words

$$\forall i \geqslant M \text{ we have } x_i < x + 1 \tag{14.168}$$

Let $i \geqslant \max(M, N)$ then by [eqs: 14.167 and 14.168] we have $x - \varepsilon < x_i < x + \varepsilon \Rightarrow |x_i - x| < \varepsilon$ which as $\varepsilon \in \mathbb{R}^+$ was chosen arbitrary proves that

$$\lim_{n\rightarrow\infty} x_n \text{ exists and } \lim_{n\rightarrow\infty} x_n = x$$

□

Theorem 14.317. *Let $k \in \mathbb{N}$, $\{x_n\}_{n\in\{k,\ldots,\infty\}} \subseteq \mathbb{R}$ a sequence of real numbers then we have*

1. *If $\liminf_{n\rightarrow\infty} x_n$ exists then $\limsup_{n\rightarrow\infty} (-x_n)$ exists and $\liminf_{n\rightarrow\infty} x_n = -\limsup_{n\rightarrow\infty} (-x_n)$*
2. *If $\limsup_{n\rightarrow\infty} x_n$ exists then $\liminf_{n\rightarrow\infty} (-x_n)$ exists and $\limsup_{n\rightarrow\infty} x_n = -\liminf_{n\rightarrow\infty} (-x_n)$*
3. *If $\limsup_{n\rightarrow\infty} x_n$ exists and $\alpha \in [0, \infty[$ then $\limsup_{n\rightarrow\infty} (\alpha \cdot x_n)$ exists and*
$$\limsup_{n\rightarrow\infty} (\alpha \cdot x_n) = \alpha \cdot \limsup_{n\rightarrow\infty} x_i$$
4. *If $\liminf_{n\rightarrow\infty} x_n$ exists and $\alpha \in [0, \infty[$ then $\liminf_{n\rightarrow\infty} (\alpha \cdot x_n)$ exists and $\liminf_{n\rightarrow\infty} (\alpha \cdot x_n) = \alpha \cdot \liminf_{n\rightarrow\infty} x_n$*
5. *If $x \in \mathbb{R}$ and $\liminf_{n\rightarrow\infty} x_n$ exists then for $\{x_n + x\}_{n\in\{k,\ldots,\infty\}}$ $\liminf_{n\rightarrow\infty} (x_n + x)$ exists and*
$$\liminf_{n\rightarrow\infty} (x_n + x) = \liminf_{n\rightarrow\infty} x_n + x$$
6. *If $x \in \mathbb{R}$ and $\limsup_{n\rightarrow\infty} x_n$ exists then for $\{x_n + x\}_{n\in\{k,\ldots,\infty\}}$ $\limsup_{n\rightarrow\infty} (x_n + x)$ exists and*
$$\limsup_{n\rightarrow\infty} (x_n + x) = \limsup_{n\rightarrow\infty} x_n + x$$

Proof.

1. As $\liminf_{n\rightarrow\infty} x_i$ exists we have
$$\forall n \in \{k, \ldots, \infty\} \text{ that } \inf(\{x_i | i \in \{n, \ldots, \infty\}\}) \text{ exists} \tag{14.169}$$
Using [theorem: 10.21] we have then that $\sup(\{-x_i | i \in \{n, \ldots, \infty\}\})$ exists and
$$\forall n \in \mathbb{N}_0 \in \{k, \ldots, \infty\} \sup(\{-x_i | i \in \{n, \ldots, \infty\}\}) = -\inf(\{x_i | i \in \{n, \ldots, \infty\}\}) \tag{14.170}$$
Again using the existence of $\liminf_{n\rightarrow\infty} x_i$ we have the existence of $\lim_{n\rightarrow\infty} \inf(\{x_i | i \in \{n, \ldots, \infty\}\})$ and further we have
$$\begin{aligned} \liminf_{n\rightarrow\infty} x_n &= \lim_{n\rightarrow\infty} \inf(\{x_i | i \in \{n, \ldots, \infty\}\}) \\ &\underset{\text{[eq: 14.170]}}{=} \lim_{n\rightarrow\infty} (-\sup(\{-x_i | i \in \{n, \ldots, \infty\} \\ &\underset{\text{[theorem: 14.298]}}{=} -\lim_{n\rightarrow\infty} (\sup(\{-x_i | i \in \{n, \ldots, \infty\}\})) \\ &\underset{\text{def}}{=} -\limsup_{n\rightarrow\infty} (-x_n) \end{aligned}$$
2. As $\limsup_{n\rightarrow\infty} x_i$ exists we have
$$\forall n \in \{k, \ldots, \infty\} \sup(\{x_i | i \in \{n, \ldots, \infty\}\}) \text{ exists} \tag{14.171}$$
Using [theorem: 10.21] we have then that $\inf(\{-x_i | i \in \{n, \ldots, \infty\}\})$ exists and
$$\forall n \in \mathbb{N}_0 \in \{k, \ldots, \infty\} \inf(\{-x_i | i \in \{n, \ldots, \infty\}\}) = -\sup(\{x_i | i \in \{n, \ldots, \infty\}\}) \tag{14.172}$$
Again using the existence of $\limsup_{n\rightarrow\infty} x_i$ we have the existence of $\lim_{n\rightarrow\infty} \sup(\{x_i | i \in \{n, \ldots, \infty\}\})$, further we have
$$\begin{aligned} \limsup_{n\rightarrow\infty} x_n &= \lim_{n\rightarrow\infty} \sup(\{x_i | i \in \{n, \ldots, \infty\}\}) \\ &\underset{\text{[eq: 14.172]}}{=} \lim_{n\rightarrow\infty} (-\inf(\{-x_i | i \in \{n, \ldots, \infty\} \\ &\underset{\text{[theorem: 14.298]}}{=} -\lim_{n\rightarrow\infty} (\inf(\{-x_i | i \in \{n, \ldots, \infty\}\})) \\ &\underset{\text{def}}{=} -\liminf_{n\rightarrow\infty} (-x_n) \end{aligned}$$

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 [theorem: 10.22] 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 $\lim_{n\rightarrow\infty} \sup(\{x_i | i \in \{n, \ldots, \infty\}\})$ exists we have by [theorem: 14.298] that $\lim_{n\rightarrow\infty} (\alpha \cdot \sup(\{x_i | i \in \{n, \ldots, \infty\}\}))$ exists and

$$\begin{aligned}
\alpha \cdot \limsup_{n\rightarrow\infty} x_n &= \alpha \cdot \lim_{n\rightarrow\infty} \sup(\{\alpha \cdot x_i | i \in \{n, \ldots, \infty\}\}) \\
&\underset{[\text{theorem: } 14.298]}{=} \lim_{n\rightarrow\infty} (\alpha \cdot \sup(\{x_i | i \in \{n, \ldots, \infty\}\})) \\
&= \lim_{n\rightarrow\infty} (\sup(\{\alpha \cdot x_i | i \in \{n, \ldots, \infty\}\})) \\
&= \limsup_{n\rightarrow\infty} (\alpha \cdot x_n)
\end{aligned}$$

4. As $\liminf_{n\rightarrow\infty} x_i$ exists we have

$$\forall n \in \{k, \ldots, \infty\} \inf(\{x_i | i \in \{n, \ldots, \infty\}\}) \text{ exists}$$

Using [theorem: 10.22] 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 $\lim_{n\rightarrow\infty} \inf(\{x_i | i \in \{n, \ldots, \infty\}\})$ exists we have by [theorem: 14.298] that $\lim_{n\rightarrow\infty} (\alpha \cdot \inf(\{x_i | i \in \{n, \ldots, \infty\}\}))$ exists and

$$\begin{aligned}
\alpha \cdot \liminf_{n\rightarrow\infty} x_n &= \alpha \cdot \lim_{n\rightarrow\infty} \inf(\{\alpha \cdot x_i | i \in \{n, \ldots, \infty\}\}) \\
&\underset{[\text{theorem: } 14.298]}{=} \lim_{n\rightarrow\infty} (\alpha \cdot \inf(\{x_i | i \in \{n, \ldots, \infty\}\})) \\
&= \lim_{n\rightarrow\infty} (\inf(\{\alpha \cdot x_i | i \in \{n, \ldots, \infty\}\})) \\
&= \liminf_{n\rightarrow\infty} (\alpha \cdot x_n)
\end{aligned}$$

5. As $\liminf_{n\rightarrow\infty} x_i$ exists we have

$$\forall n \in \{k, \ldots, \infty\} \ \inf(\{x_i | i \in \{n, \ldots, \infty\}\}) \text{ exist}$$

Using [theorem: 10.24] 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

$$\begin{aligned}
\liminf_{n\rightarrow\infty} x_n &= \lim_{n\rightarrow\infty} (\inf(\{x_i + x | i \in \{n, \ldots, \infty\}\})) \\
&= \lim_{n\rightarrow\infty} (\inf(\{x_i | i \in \{n, \ldots, \infty\}\}) + x) \\
&\underset{[\text{theorem: } 14.296]}{=} \lim_{n\rightarrow\infty} \inf(\{x_i | i \in \{n, \ldots, \infty\}\}) + x \\
&= \liminf_{n\rightarrow\infty} x_i + x
\end{aligned}$$

6. As $\limsup_{n\rightarrow\infty} x_i$ exists we have

$$\forall n \in \{k, \ldots, \infty\} \ \sup(\{x_i | i \in \{n, \ldots, \infty\}\})$$

Using [theorem: 10.24] 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

$$\begin{aligned}
\limsup_{n\rightarrow\infty} x_n &= \lim_{n\rightarrow\infty} (\sup(\{x_i + x | i \in \{n, \ldots, \infty\}\})) \\
&= \lim_{n\rightarrow\infty} (\sup(\{x_i | i \in \{n, \ldots, \infty\}\}) + x) \\
&\underset{[\text{theorem: } 14.296]}{=} \lim_{n\rightarrow\infty} \sup(\{x_i | i \in \{n, \ldots, \infty\}\}) + x \\
&= \limsup_{n\rightarrow\infty} x_i + x
\end{aligned}$$

$\square$

14.10 Complete spaces

Next we define a necessary condition for a sequence to be convergent in a metric or normed space.

Definition 14.318. (Cauchy Condition) *Let $\langle X, d\rangle$ be a metric space, $k\in\mathbb{N}_0$ then a sequence $\{x_n\}_{n\in\{k,\ldots,\infty\}}\subseteq X$ is **Cauchy** if*

$$\forall\varepsilon\in\mathbb{R}^+ \text{ there exists a } N\in\{k,\ldots,\infty\} \text{ such that } \forall n,m\in\mathbb{N} \text{ with } N\leqslant n,m \text{ we have } d(x_n,x_m)<\varepsilon$$

For a normed space the above definition becomes

Definition 14.319. (Cauchy condition) *Let $\langle X, \|\|\rangle$ be a metric space, $k\in\mathbb{N}_0$ then a sequence $\{x_n\}_{n\in\{k,\ldots,\infty\}}\subseteq X$ is **Cauchy** if*

$$\forall\varepsilon\in\mathbb{R}^+ \text{ there exists a } N\in\{k,\ldots,\infty\} \text{ such that } \forall n,m\in\mathbb{N}_0 \text{ with } N\leqslant n,m \text{ we have } \|x_n-x_m\|<\varepsilon$$

Convergence and Cauchy are topological constructs as is show in the following theorem.

Theorem 14.320. *Let X be a set with two equivalent norms $\|\|_1$ and $\|\|_2$ then for a sequence $\{x_n\}_{n\in\{k,\ldots,\infty\}}\subseteq X$ we have*

1. *$\{x_n\}_{n\in\{k,\ldots,\infty\}}$ converges to x using $\|\|_1 \Leftrightarrow \{x_n\}_{n\in\{k,\ldots,\infty\}}$ converges to x using $\|\|_2$*
2. *$\{x_n\}_{n\in\{k,\ldots,\infty\}}$ is Cauchy using $\|\|_1 \Leftrightarrow \{x_n\}_{n\in\{k,\ldots,\infty\}}$ is Cauchy using $\|\|_2$*

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 we have by [theorem: 14.114] there exists $M_1,M_2\in\mathbb{R}^+$ 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. Let $\varepsilon>0$ then by convergence in $\|\|_1$ there exists a $N\in\{k,\ldots,\infty\}$ such that $\forall n\geqslant N$ we have $\|x-x_n\|_1<\frac{\varepsilon}{M_2}$ so that $\|x-x_n\|_2\leqslant M_2\cdot\|x-x_n\|_1<M_2\cdot\frac{\varepsilon}{M_2}=\varepsilon$
2. 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 $\|x_n-x_m\|_1<\frac{\varepsilon}{M_2}$ and thus $\|x_n-x_m\|_2\leqslant M_2\cdot\|x_n-x_m\|_1<M_2\cdot\frac{\varepsilon}{M_2}=\varepsilon$ □

A Cauchy sequence in a normed space is bounded.

Theorem 14.321. *Let $\langle X, \|\|\rangle$ be a normed space, $k\in\mathbb{N}$ and $\{x_i\}_{i\in\{k,\ldots,\infty\}}\subseteq X$ a Cauchy sequence then there exists a $K\in\mathbb{R}^+$ such that $\forall i\in\{k,\ldots,\infty\}$ we have $\|x_i\|\leqslant K$*

Proof. As $1\in\mathbb{R}^+$ then exist a $N\in\{k,...,\infty\}$ such that $\forall n,m\in\mathbb{N}_0$ with $N\leqslant n,m$ we have $\|x_n-x_m\|<1$. Let $M=\max(\max(\{\|x_i\||i\in\{k,..., N\}\}),\|x_N\|+1)$ then for $i\in\{k,\ldots,\infty\}$ we have either:

$\boldsymbol{i\in\{k,\ldots,N\}}$**.** Then $\|x_i\|\leqslant\max(\{\|x_i\||i\in\{k,\ldots,N\}\})\leqslant M$

$\boldsymbol{N<i}$**.** Then $\|x_i\|=\|x_i-x_N+x_N\|\leqslant\|x_i-x_N\|+\|x_N\|<1+\|x_N\|\leqslant M$

proving that $\forall i\in\{k,\ldots,N\}$ we have $\|x_i\|\leqslant M$. □

We prove now that the Cauchy condition is indeed necessary for convergence.

Theorem 14.322. *If $\langle X, d\rangle$ is a metric space then every convergent sequence is Cauchy. Likewise if $\langle X, \|\|\rangle$ is a normed space then every convergent sequence is Cauchy.*

Proof. Let $\langle X, d\rangle$ be a metric space and let $\{x_n\}_{n\in\{k,\ldots,\infty\}}\subseteq X$ be a sequence that converges to x then for $\varepsilon\in\mathbb{R}^+$ there exist a $N\in\{k,\ldots,\infty\}$ such that $\forall n\geqslant N$ we have $d(x,x_n)<\frac{\varepsilon}{2}$. So if $n,m\in\mathbb{N}$ with $n,m\geqslant N$ then we have

$$d(x_n,x_m)\leqslant d(x_n,x)+d(x,x_m)<\frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon$$

proving that $\{x_n\}_{n\in\{k,\ldots,\infty\}}$ is Cauchy. Likewise let $\langle X, \|\|\rangle$ be a normed space and let $\{x_n\}_{n\in\{k,\ldots,\infty\}}\subseteq X$ be a sequence that converges to x then for $\varepsilon\in\mathbb{R}^+$ there exist a $N\in\{k,\ldots,\infty\}$ such that $\forall n\geqslant N$ we have $\|x_n-x\|<\frac{\varepsilon}{2}$. So if $n,m\in\mathbb{N}$ with $n,m\geqslant N$ then we have

$$\|x_m-x_n\|=\|x_m-x+x-x_n\|\leqslant\|x_m-x\|+\|x_n-x\|<\frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon$$

proving that $\{x_n\}_{n\in\{k,\ldots,\infty\}}$ is Cauchy. □

Example 14.323. Take $]0,\infty[$ and define $d:]0,\infty[\times]0,\infty[\rightarrow\mathbb{R}$ by $d(x,y)=|x-y|$ then

$$\langle]0,\infty[\times]0,\infty[,d\rangle \text{ is a metric space}$$

and further $\{\frac{1}{n}\}_{n\in\mathbb{N}}$ is Cauchy but does not converges

Proof. As $\langle\mathbb{R},d_{||}\rangle$ is a metric space and $d=(d_{||})_{|]0,\infty[\times]0,\infty[}$ it follows from [theorem: 14.66] that $\langle]0,\infty[,d\rangle$ is a metric space. As by [example: 14.293] $\{\frac{1}{n}\}_{n\in\mathbb{N}}$ converges to 0 in $\langle\mathbb{R},d_{||}\rangle$ it is Cauchy in $\langle\mathbb{R},d_{||}\rangle$. Hence $\forall\varepsilon\in\mathbb{R}^+$ there exist a $N\in\mathbb{R}^+$ such that $\forall n,m\in\{N,\ldots,\infty\}$ we have $d(x_n,x_m)=d_{||}(x_n,x_m)<\varepsilon$ proving that $\{\frac{1}{n}\}_{n\in\mathbb{N}}$ is Cauchy in $\langle]0,\infty[,d\rangle$. However as $\{\frac{1}{n}\}_{n\in\mathbb{N}}$ converges to 0 in $\langle\mathbb{R},d_{||}\rangle$ and $0\notin]0,\infty[$ it follows that $\{\frac{1}{n}\}_{n\in\mathbb{N}}$ does not converges in $\langle]0,\infty[,d\rangle$. □

The above example proves that although every convergent sequence is Cauchy the opposite is not always true. So if we want that Cauchy is a sufficient condition for converging we must state this as a extra property for the metric or the norm.

Definition 14.324. *A metric space $\langle X,d\rangle$ or a normed space $\langle X,\|\|\rangle$ is* ***complete*** *if every Cauchy sequence is convergent. A complete normed space is called a* ***Banach*** *space.*

Theorem 14.325. *Let $n\in\mathbb{N}$, $\{\langle X_i,\|\|_i\rangle\}_{i\in\{1,\ldots,n\}}$ a family of Banach spaces then if $\|\|_{\max}$ is the maximum norm [see theorem: 14.94] then*

$$\left\langle\prod_{i\in\{1,\ldots,n\}}X_i,\|\|_{\max}\right\rangle \textit{ is a Banach space}$$

Proof. Using [theorem: 14.94] we have that $\langle\prod_{i\in\{1,\ldots,n\}}X_i,\|\|_{\max}\rangle$ is a normed space. Let $\{x_i\}_{i\in\{k,\ldots,\infty\}}\subseteq\prod_{i\in\{1,\ldots,n\}}X_i$ be a Cauchy sequence. Take $\varepsilon\in\mathbb{R}^+$ then there exists a $N\in\{k,\ldots,\infty\}$ such that $\forall k,l\geqslant N$ we have $\max(\{\|\pi_i(x_k-x_l)\|_i|i\in\{1,\ldots,n\}\})=\|x_k-x_l\|_{\max}<\varepsilon$. Hence $\forall i\in\{1,\ldots,n\}$ we have $\|\pi_i(x_k)-\pi_i(x_l)\|_i=\|\pi_i(x_k-x_l)\|_i<\varepsilon$ which as $\langle X_i,\|\|_i\rangle$ is Banach proves that there exist a $y_i\in X_i$ such that $\lim_{k\rightarrow\infty}\pi_i(x_k)=y_i$. In other words given $\varepsilon\in\mathbb{R}^+$ there exist a $N_i\in\{k,\ldots,\infty\}$ such that $\forall k\geqslant N_i$ we have $\|\pi_i(x_k)-y_i\|_i<\varepsilon$. Take $y=(y_1,\ldots,y_n)\in\prod_{i\in\{1,\ldots,n\}}X_i$ then we have that for $k\geqslant\max(\{N_i|i\in\{1,\ldots,n\}\})$ that

$$\|x_k-y\|_{\max}=\max(\{\pi_i(x_k-y)|i\in\{1,\ldots,n\}\})=\max(\{\pi_i(x_k)-y_i|i\in\{1,\ldots,n\}\})<\varepsilon$$

proving that $\{x_i\}_{i\in\{k,\ldots,\infty\}}$ converges. □

Next we prove that in every metric space the limit point of $\{x_n|n\in\{k,\ldots,\infty\}\}$ where $\{x_n\}_{n\in\{k,\ldots,\infty\}}$ is a Cauchy sequence. To prove this we first need a little lemma.

Lemma 14.326. *Let $k\in\mathbb{N}_0$, $\langle X,d\rangle$ be a metric space, $\{x_n\}_{n\in\{k,\ldots,\infty\}}\subseteq X$ a sequence and x a limit point of $\{x_n|n\in\{k,\ldots,\infty\}\}$ then $\forall\varepsilon\in\mathbb{R}^+$, $\forall N\in\{k,\ldots,\infty\}$ there exist a $m>N$ such that $d(x,x_m)<\varepsilon$*

Proof. Let $\varepsilon\in\mathbb{R}^+$ and define

$$A_\varepsilon=\{n\in\{k,\ldots,\infty\}|0<d(x,x_i)<\varepsilon\}\subseteq\{k,\ldots,\infty\}$$

Assume that A_ε is finite. As x is a limit point of $\{x_n|n\in\{k,\ldots,\infty\}\}$ and $x\in B_d(x,\varepsilon)$ a open set we have that $B_d(x,\varepsilon)\bigcap((\{x_n|n\in\{k,\ldots,\infty\}\})\setminus\{x\})\neq\varnothing$, hence there exists a $i\in\{k,\ldots,\infty\}$ such that

$$x_i\neq x\Rightarrow 0<d(x_i,x) \text{ and } d(x,x_i)<\varepsilon$$

proving that

$$A_\varepsilon\neq\varnothing$$

Define $B_\varepsilon=\{d(x,x_n)|n\in A_\varepsilon\}\underset{A_\varepsilon\neq\varnothing}{\Rightarrow}B_\varepsilon\neq\varnothing$ and $f: A_\varepsilon\rightarrow B_\varepsilon$ by $f(n)=d(x,x_n)$ is a surjection, as A_ε is assumed to be finite it follows from [theorem: 6.43] that

$$\varnothing\neq B_\varepsilon=\{d(x,x_n)|n\in A_\varepsilon\} \text{ is finite}$$

Hence using [theorem: 6.48]

$$\delta=\min(B_\varepsilon) \text{ exists}$$

As $\delta\in B_\varepsilon$ there exist a $n\in A_\varepsilon$ such that $\delta=d(x,x_n)$ hence

$$0<\delta=d(x,x_n)<\varepsilon$$

Again using the fact that x is a limit point of $\{x_n|n\in\{k,\ldots,\infty\}\}$ and $x\in B_d(x,\delta)$ there exist a $j\in\{k,\ldots,\infty\}$ such that

$$x_j = x \Rightarrow 0 < d(x,x_j) \text{ and } d(x,x_j) < \delta < \varepsilon$$

hence $j\in A_\varepsilon$ so that $d(x,x_j)\in B_\varepsilon$, as $\delta=\min(B_\varepsilon)$ we have $\delta\leqslant d(x,x_j)<\delta$ leading to the contradiction $\delta<\delta$. So the assumption that A_ε is finite is false proving that

$$A_\varepsilon \text{ is infinite}$$

Assume now that there exist a $N\in\{k,\ldots,\infty\}$ such that $\forall n>N$ we have $\varepsilon\leqslant d(x,x_n)$ then $n\notin A_\varepsilon$ hence $A_\varepsilon\subseteq\{k,\ldots,N\}\Rightarrow A_\varepsilon$ is finite which contradicts the fact that A_ε is infinite. So the assumption is wrong and we have that

$$\forall N\in\{k,\ldots,\infty\} \text{ there exist a } n>N \text{ such that } d(x,x_n)<\varepsilon$$

proving the theorem. □

Theorem 14.327. *Let $k\in\mathbb{N}_0$, $\langle X,d\rangle$ a metric space and $\{x_n\}_{n\in\{k,\ldots,\infty\}}\subseteq X$ a Cauchy sequence then if x and y are limit points of $\{x_n|n\in\{k,\ldots,\infty\}\}$ we have that $x=y$.*

Proof. Let x,y be limit points of $\{x_n|n\in\{k,\ldots,\infty\}\}$ and assume that $x\neq y$. As $x\neq y$ we have that $\varepsilon=d(x,y)\in\mathbb{R}^+$. By the Cauchy condition there exist a $N\in\{k,\ldots,\infty\}$ such that $\forall i,j\geqslant N$ we have $d(x_i,x_j)<\frac{\varepsilon}{3}$. Using the previous lemma [lemma: 14.326] there exists $n,m>N$ such that $d(x,x_n)<\frac{\varepsilon}{3}$ and $d(y,x_m)<\frac{\varepsilon}{3}$. Hence

$$\varepsilon=d(x,y)\leqslant d(x,x_n)+d(x_n,y)\leqslant d(x,x_n)+d(x_n,x_m)+d(x_m,y)<\frac{\varepsilon}{3}+\frac{\varepsilon}{3}+\frac{\varepsilon}{3}=\varepsilon$$ □

14.10.1 Examples of complete spaces

First we prove that every compact metric space is complete, to do this we first need a lemma about limit points and sequences.

Theorem 14.328. *Every compact metric space $\langle X,d\rangle$ is complete.*

Proof. Let $\{x_n\}_{n\in\{k,\ldots,\infty\}}\subseteq X$ be a Cauchy sequence in X and consider $S=\{x_n|n\in\{k,\ldots,\infty\}\}$ then for S we must consider the following cases:

***S* is finite.** Assume that $\forall N\in\{k,\ldots,\infty\}$ there exists $i_0,j_0\geqslant N$ with $x_i\neq x_j$. Define

$$A=\{d(x,y)|(x,y)\in S\times S\wedge x\neq y\}\subseteq\{d(x,y)|(x,y)\in S\}$$

Now $S\times S$ is finite by [theorem: 6.40] and $f: S\times S\rightarrow\{d(x,y)|(x,y)\in S\}$ defined by $f(x,y)=d(x,y)$ is a surjection so that by [theorem: 6.43] $\{d(x,y)|(x,y)\in S\}$ is finite and thus that A is finite. Further by the assumption there exist $i,j\geqslant k$ such that $x_i\neq x_j$ proving that $d(x_i,x_j)\in A$ and thus that $A\neq\varnothing$. So $\varepsilon=\min(A)$ exist and $0<\varepsilon$ [for if $d\in A\Rightarrow\exists x,y\in S$ with $x\neq y\Rightarrow d=d(x,y)>0$]. As $\{x_n\}_{n\in\{k,\ldots,\infty\}}$ is Cauchy there exist a $N\in\{k,\ldots,\infty\}$ such that $\forall i,j\geqslant N$ we have $d(x_i,x_j)<\varepsilon$. By the assumption there exist a $i_0,j_0\geqslant N$ such that $x_{i_0}\neq x_{j_0}\Rightarrow d(x_{i_0},x_{j_0})\in A$, hence $\varepsilon=\min(S)\leqslant d(x_{i_0},y_{i_0})<\varepsilon$ a contradiction. So the assumption must be false meaning that

$$\exists N\in\{k,\ldots,\infty\} \text{ such that } \forall i,j\geqslant N\; x_i=x_j\Rightarrow d(x_i,x_j)=0$$

or taking $j=N$ and $x=x_N$ we have that

$$\forall i\geqslant N\; d(x_i,x)=0$$

proving as $\forall\varepsilon\in\mathbb{R}^+\; 0<\varepsilon$ that

$$\lim_{n\rightarrow\infty}x_n=x$$

***S* is infinite.** Then as $S\subseteq X$ a compact space we have using [theorem: 14.238] and [definition: 14.237] that there exist a limit point for S which because of [theorem: 14.327] is unique. Let's call this unique limit point of S x. Take $\varepsilon>0$ 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 [lemma: 14.326] there exist a $n>N$ such that $d(x,x_n)<\frac{\varepsilon}{2}$. Hence if $m\geqslant N$ we have that

$$d(x,x_m)\leqslant d(x,x_n)+d(x_n,x_m)<\frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon$$ □

Theorem 14.329. *Let $\langle X, d\rangle$ be a complete space and $A\subseteq X$ where A is closed then $\langle A, d_{|A}\rangle$ is complete.*

Proof. Let $\{x_n\}_{n\in\{k,\ldots,\infty\}}\subseteq A$ be a Cauchy sequence in $\langle A, d_{|A}\rangle$ then given $\varepsilon\in\mathbb{R}^+$ there exist a $N\in\{k,\ldots,\infty\}$ such that $\forall n,m\geqslant N$ we have that $d(x_n,x_m)=d_{|A}(x_n,x_m)<\varepsilon$. So $\{x_n\}_{n\in\{k,\ldots,\infty\}}$ is Cauchy in X a complete space, hence $\{x_n\}_{n\in\{k,\ldots,\infty\}}$ converges to $x\in X$ or

$$\forall\varepsilon\in\mathbb{R}^+\,\exists N\in\{k,\ldots,\infty\}\text{ such that } d(x_n,x)<\varepsilon \tag{14.173}$$

Let $U\in\mathcal{T}_d$ such that $x\in A$ then there exist a $\delta\in\mathbb{R}^+$ such that $x\in B_d(x,\delta)\subseteq U$. By [eq: 14.173] there exist a $N\in\{k,\ldots,\infty\}$ such that $d(x_N,x)<\delta$ hence $x_N\in B_d(x,\varepsilon)\subseteq U$ proving as $x_N\in A$ that $A\bigcap U\neq\varnothing$. So using [theorem: 14.29] it follows that

$$x\in A$$

Hence if $\zeta\in\mathbb{R}^+$ we have by [eq: 14.173] that there exist a $N\in\{k,\ldots,\infty\}$ such that $\forall n\geqslant N$ $d(x,x_n)<\varepsilon\underset{x_n,x\in A}{\Rightarrow}d_{|A}(x,x_n)=d(x_n,x)<\zeta$ proving that $\{x_n\}_{n\in\{k,\ldots,\infty\}}$ converges in A. □

A very important example of a complete normed space is the set of real numbers with the absolute value norm.

Theorem 14.330. *The normed space $\langle\mathbb{R},\|\rangle$ is a Banach space.*

Proof. Let $\{x_n\}_{n\in\{k,\ldots,\infty\}}\subseteq\mathbb{R}$ be a Cauchy sequence in $\mathbb{R}$ then by [theorem: 14.321] there exist a $K\in\mathbb{R}^+$ such that $\forall n\in\{k,\ldots,\infty\}$ $|x_n|\leqslant K$ so that

$$\{x_n\}_{n\in\{k,\ldots,\infty\}}\subseteq[-K,K] \tag{14.174}$$

By [theorem: 14.232] is a compact subset of $\mathbb{R}$, hence by definition $[-K,K]$ is a compact space using the subspace topology which by [theorem: 14.108] is generated by the norm $||_{|[-K,K]}$. By the previous theorem [theorem: 14.328] it follows that $\langle[-K,K],||_{|[-K,K]}\rangle$ is complete. Given $\varepsilon\in\mathbb{R}^+$ we have as $\{x_n\}_{n\in\{k,\ldots,\infty\}}$ is Cauchy a $N\in\{k,\ldots,\infty\}$ such that $\forall i,j\geqslant N$ we have $|x_i-x_j|<\varepsilon\underset{[\text{eq: }14.174]}{\Rightarrow}|x_i-x_j|_{|[-K,K]}<\varepsilon$ proving that $\{x_n\}_{n\in\{k,\ldots,\infty\}}$ is Cauchy in $\langle[-K,K],||_{|[-K,K]}\rangle$. Hence there exist a $x\in[-K,K]$ such that $\lim_{n\to\infty}x_n=x$ [using $||_{|[-K,K]}$]. Hence given $\varepsilon\in\mathbb{R}^+$ there exist a $N\in\{k,\ldots,\infty\}$ such that $\forall n\geqslant N$ $|x_n-x|_{|[-K,K]}<\varepsilon\underset{[\text{eq: }14.174]}{\Rightarrow}|x_n-x|<\varepsilon$ proving that

$$\{x_n\}_{n\in\{k,\ldots,\infty\}}\text{ converges to }x\text{ using }\langle\mathbb{R},\|\rangle$$

□

Corollary 14.331. *Let $\|\|$ be a norm on $\mathbb{R}^n$ then $\langle\mathbb{R}^n,\|\|\rangle$ is a Banach space.*

Proof. Let $\|\|_n$ be the maximum norm defined by $\|x\|_n=\max(\{|\pi_i(x)||i\in\{1,\ldots,n\}\})$ then as $\langle\mathbb{R},\|\rangle$ is a Banach space by the previous theorem [theorem: 14.330] we have by [theorem: 14.325] that

$$\langle\mathbb{R}^n,\|\|_n\rangle\text{ is a Banach space} \tag{14.175}$$

Now using [theorem: 14.274]

$$\|\|_n\text{ is equivalent with }\|\| \tag{14.176}$$

Let $\{x_i\}_{i\in\{k,\ldots,\infty\}}\subseteq\mathbb{R}^n$ be a Cauchy sequence using $\|\|$ then by [theorem: 14.320] and [eq: 14.176] it follows that $\{x_i\}_{i\in\{k,\ldots,\infty\}}$ is a Cauchy sequence, hence by [eq: 14.175] $\{x_i\}_{i\in\{k,\ldots,\infty\}}$ converges. Finally using [theorem: 14.320] and [eq: 14.176] it follows that

$$\{x_i\}_{i\in\{k,\ldots,\infty\}}\text{ converges in }\langle\mathbb{R}^n,\|\|\rangle$$

proving that $\langle\mathbb{R}^n,\|\|\rangle$ is a Banach space. □

Corollary 14.332. *Every finite dimensional normed vector space over $\mathbb{R}$ is a Banach space*

Proof. Let $\langle X,\|\|_X\rangle$ be a n-dimensional normed vector space then using [theorem: 14.170] there exist a norm $\|\|$ on $\mathbb{R}^n$ and a linear isometric isomorphism $\varphi: X\to\mathbb{R}^n$ using the norms $\|\|_X$ and $\|\|$. Let $\{x_i\}_{i\in\{k,\ldots,\infty\}}\subseteq X$ be a Cauchy sequence in $\langle X,\|\|_X\rangle$. Then given $\varepsilon\in\mathbb{R}^+$ there exist a $N\in\{k,\ldots,\infty\}$ such that $\forall i,j\geqslant N$ $\|x_i-x_j\|_X<\varepsilon$. Consider now the sequence $\{\varphi(x_i)\}_{i\in\{k,\ldots,\infty\}}\subseteq\mathbb{R}^n$ then $\|\varphi(x_i)-\varphi(x_j)\|=\|\varphi(x_i-x_j)\|=\|x_i-x_j\|_X<\varepsilon$ proving that $\{\varphi(x_i)\}_{i\in\{k,\ldots,\infty\circ\}}$ is a Cauchy sequence in $\langle\mathbb{R}^n,\|\|\rangle$. As $\langle\mathbb{R}^n,\|\|\rangle$ is a Banach space by [theorem: 14.331] so that $\{\varphi(x_i)\}_{i\in\{k,\ldots,\infty\}}$ converges to a $y\in\mathbb{R}^n$. Hence given $\varepsilon\in\mathbb{R}^+$ there exist a $N\in\{k,\ldots,\infty\}$ such that $\forall m\geqslant N$ we have $\|\varphi(x_m)-y\|<\varepsilon$. Take $x=\varphi^{-1}(y)$ then we have $\|x_n-x\|_X=\|\varphi(x_n-x)\|=\|\varphi(x_n)-\varphi(\varphi^{-1}(y))\|=\|\varphi(x_n)-x\|<\varepsilon$ proving that $\{x_i\}_{i\in\{k,\ldots,\infty\}}$ converges to x. Hence $\langle X,\|\|\rangle$ is a Banach space. □

Corollary 14.333. *The normed space $\langle\mathbb{C},\|\rangle$ is a Banach space [where $\langle\mathbb{C},\|\rangle$ is defined in [example: 14.95]*

Proof. To avoid any confusion between the norm on $\mathbb{R}$ and $\mathbb{C}$ we use $||$ for the absolute value norm on $\mathbb{R}$ and $||_c$ for the canonical norm on $\mathbb{C}$. Let $\{z_n\}_{n\in\{k,\ldots,\infty\}}\subseteq\mathbb{C}$ be a Cauchy sequence in $\langle\mathbb{C},||_c\rangle$ then $\forall n\in\{k,\ldots,\infty\}$ we have that $z_n=x_n+i\cdot y_n$ where $x_n,y_n\in\mathbb{R}$. Now $\forall n,m\in\{k,\ldots,\infty\}$ we have by [theorem: 10.83 (3)] that

$$|x_m-x_n|=|\mathrm{Re}(z_m-z_n)|<|z_m-z_n|_c \text{ and } |y_m-y_n|_c=|\mathrm{Img}(z_m-z_n)|<|z_m-z_n|_c \tag{14.177}$$

By the Cauchy condition there exist a $N\in\{k,\ldots,\infty\}$ such that $\forall n\geqslant N$ we have that $|z_n-z_m|_c<\varepsilon$, combining this with [eqs: 14.177] gives $|x_m-x_n|<\varepsilon\wedge|y_m-y_n|<\varepsilon$. Hence

$$\{x_n\}_{n\in\{k,\ldots,\infty\}},\{y_n\}_{n\in\{k,\ldots,\infty\}} \text{ are Cauchy in } \langle\mathbb{R},||\rangle$$

As by [theorem: 14.330] $\langle\mathbb{R},||\rangle$ is complete there exists $x,y\in\mathbb{R}$ such that

$$\lim_{n\to\infty}x_n=x\wedge\lim_{n\to\infty}y_n=y$$

Hence using [theorem: 14.299] it follows that $\{z_n\}_{n\in\{k,\ldots,\infty\}}$ converges to z. □

Theorem 14.334. *Let $\langle X,\|\|_X\rangle$ be a normed space, $\langle Y,\|\|_Y\rangle$ a Banach space then the normed space $\langle L(X,Y),\|\|_{L(X,Y)}\rangle$ [see theorem: 14.179] is a Banach space.*

Proof. Let $\{L_n\}_{n\in\{k,\ldots,\infty\}}\subseteq L(X,Y)$ be a Cauchy sequence. Let $x\in X$ then given $\varepsilon\in\mathbb{R}^+$ there exist a $N\in\{k,\ldots,\infty\}$ such that $\forall n,m\geqslant N$ we have $\|L_n-L_m\|_{L(X,Y)}<\frac{\varepsilon}{\|x\|+1}$. Hence

$$\|L_n(x)-L_m(x)\|_Y=\|(L_n-L_m)(x)\|_Y\leqslant_{[\text{theorem: }14.179}\|L_n-L_m\|_{L(X,Y)}\cdot\|x\|_X<\frac{\varepsilon}{\|x\|+1}\cdot\|x\|<\varepsilon$$

proving that

$$\{L_n(x)\}_{n\in\{k,\ldots,\infty\}} \text{ is a Cauchy sequence in } \langle Y,\|\|_Y\rangle \tag{14.178}$$

As $\langle Y,\|\|_Y\rangle$ is a Banach space $\lim_{n\to\infty}L_n(x)$ exist, allowing us to define

$$L:X\to Y \text{ by } L(x)=\lim_{n\to\infty}L_n(x) \tag{14.179}$$

Let $x,y\in X$ and $\alpha\in\mathbb{K}$ then we have $\{L_n(x+\alpha\cdot y)\}_{n\in\{k,\ldots,\infty\}}=\{L_n(x)+\alpha\cdot L_n(y)\}_{n\in\{k,\ldots,\infty\}}$. Using [theorems: 14.298, 14.297] $\{L_n(x+\alpha\cdot y)\}_{n\in\{k,\ldots,\infty\}}$ converges and

$$\lim_{n\to\infty}L_n(x+\alpha\cdot y)=\lim_{n\to\infty}L_n(x)+\alpha\cdot\lim_{n\to\infty}L_n(y)$$

proving that

$$L(x+\alpha\cdot y)=L(x)+\alpha\cdot L(y)$$

or that

$$L\in\mathrm{Hom}(X,Y) \tag{14.180}$$

As $\{L_n\}_{n\in\{k,\ldots,\infty\}}$ is Cauchy there exist by [theorem: 14.321] a $K\in\mathbb{R}^+$ such that $\forall n\in\{k,\ldots,\infty\}$ we have $\|L_n\|_{L(X,Y)}\leqslant K$. Hence $\forall n\in\{k,\ldots,\infty\}$ we have for $x\in X$

$$\|L_n(x)\|_Y\leqslant_{[\text{theorem: }14.179}\|L_n\|_{L(X,Y)}\cdot\|x\|_X\leqslant K\cdot\|x\|_X$$

which by [theorem: 14.301] proves

$$\|L(x)\|_Y=\left\|\lim_{n\to\infty}L_n(x)\right\|_Y\leqslant K\cdot\|x\|_X \tag{14.181}$$

Hence using [eq: 14.180] and [theorem: 14.174] it follows that

$$L\in L(X,Y) \tag{14.182}$$

Let $\varepsilon\in\mathbb{R}^+$ then for $x\in X$ with $\|x\|_X=1$ we have, as $L(x)=\lim_{n\to\infty}L(x)$, that there exist a $N_x\in\{k,\ldots,\infty\}$ such that

$$\forall n\geqslant N_x \text{ we have } \|L_n(x)-L(x)\|<\frac{\varepsilon}{4}$$

As $\{L_n\}_{n\in\{k,\ldots,\infty\}}$ is Cauchy there exist a $N\in\{k,\ldots,\infty\}$ such that $\forall n,m\geqslant N$ we have $\|L_n-L_m\|_{L(X,Y)}<\frac{\varepsilon}{4}$ hence

$$\|L_n(x)-L_m(x)\|_Y=\|(L_n-L_m)(x)\|\leqslant\|L_n-L_m\|_{L(X,Y)}\cdot\|x\|_X=\|L_n-L_m\|_{L(X,Y)}<\frac{\varepsilon}{4}$$

Let $n \geqslant N$ then we have

$$\begin{array}{rcl}\|(L_n - L)(x)\|_Y & = & \|L_n(x) - L(x)\|_Y\\ & \leqslant & \|L_n(x) - L_{\max(N,N_x)} + L_{\max(N,N_x)} - L(x)\|_Y\\ & \leqslant & \|L_n(x) - L_{\max(N,N_x)}\|_Y + \|L_{\max(N,N_x) - L(x)}\|_Y\\ & < & \frac{\varepsilon}{4} + \frac{\varepsilon}{4}\\ & = & \frac{\varepsilon}{2}\end{array}$$

Hence using [theorem: 14.178] it follows that $\|L_n - L\|_{L(X,Y)} < \frac{\varepsilon}{2} < \varepsilon$ proving that

$$\{L_n\}_{n \in \{k,\ldots,\infty\}} \text{ converges to } L \text{ in } \langle L(X,Y), \|\|_{L(X,Y)}\rangle$$

which implies that $\langle L(X,Y), \|\|_{L(X,Y)}\rangle$ is complete and thus a Banach space. □

We have a similar theorem for multilinear mappings.

Theorem 14.335. *Let $n \in \mathbb{N}$, $\{\langle X_i, \|\|_i\rangle\}_{i \in \{1,\ldots,n\}}$ a sequence of normed space, $\langle Y, \|\|_Y\rangle$ a Banach space then the normed space $\langle L(X_1,\ldots,X_n;Y), \|\|_{L(X_1,\ldots,X_n;Y)}\rangle$ [see theorem: 14.193] is a Banach space.*

Proof. Let $\{L_i\}_{i \in \{k,\ldots,\infty\}} \subseteq L(X_1,\ldots,X_n;Y)$ be a Cauchy sequence. Let $x = (x_1,\ldots,x_n) \in \prod_{l \in \{1,\ldots,n\}} X_l$ then given $\varepsilon \in \mathbb{R}^+$ we have that $\exists N \in \{k,\ldots,\infty\}$ such that $\forall i,j \geqslant N$ we have

$$\|L_i - L_j\|_{L(X_1,\ldots,X_n;Y)} < \frac{\varepsilon}{1 + \prod_{l \in \{1,\ldots,n\}} \|x_l\|_l}$$

So that

$$\begin{array}{rcl}\|L_i(x) - L_j(y)\|_Y & = & \|(L_i - L_j)(x)\|_Y\\ & \leqslant & \|L_i - L_j\|_{L(X_1,\ldots,X_n;Y)} \cdot \prod_{l \in \{1,\ldots,n\}} \|x_l\|_l\\ & < & \frac{\varepsilon}{1 + \prod_{l \in \{1,\ldots,n\}} \|x_l\|_l} \cdot \prod_{l \in \{1,\ldots,n\}} \|x_l\|_l\\ & < & \varepsilon\end{array}$$

which proves that

$$\{L_i(x)\}_{i \in \{1,\ldots,n\}} \text{ is a Cauchy sequence in } \langle Y, \|\|_Y\rangle$$

As $\langle Y, \|\|_Y\rangle$ is a Banach space it follows that $\lim_{i \to \infty} L_i(x)$ exist, this allows to define the function

$$L: \prod_{i \in \{1,\ldots,n\}} X_i \to Y \text{ where } L(x) = \lim_{i \to \infty} L_i(x) \tag{14.183}$$

Let $\alpha \in \mathbb{K}, i \in \{1,\ldots,n\}$, $x, y \in X_i$ and $\left(x_1,\ldots \underbrace{x + \alpha \cdot y}_{i} \ldots, x_n\right) \in \prod_{j \in \{1,\ldots,n\} \setminus \{i\}} X_i$ then we have that

$$\left\{L_l\left(x_1,\ldots \underbrace{x + \alpha \cdot y}_{i} \ldots, x_n\right)\right\}_{l \in \{k,\ldots,\infty\}} =$$

$$\left\{L_l\left(x_1,\ldots \underbrace{x}_{i} \ldots, x_n\right) + \alpha \cdot L_l\left(x_1,\ldots \underbrace{y}_{i} \ldots, x_n\right)\right\}_{l \in \{k,\ldots,\infty\}}$$

Using [theorems: 14.298, 14.297] $\left\{L_l\left(x_1,\ldots \underbrace{x + \alpha \cdot y}_{i} \ldots, x_n\right)\right\}_{l \in \{k,\ldots,\infty\}}$ converges and

$$\lim_{l \to \infty} L_l\left(x_1,\ldots \underbrace{x + \alpha \cdot y}_{i} \ldots, x_n\right) =$$

$$\lim_{l \to \infty} L_l\left(x_1,\ldots \underbrace{x}_{i} \ldots, x_n\right) + \alpha \cdot \lim_{l \to \infty} L_l\left(x_1,\ldots \underbrace{y}_{i} \ldots, x_n\right)$$

proving that

$$L\in\mathrm{Hom}(X_1,\ldots,X_n;Y) \tag{14.184}$$

As $\{L_i\}_{i\in\{k,\ldots,\infty\}}$ is Cauchy there exist by [theorem: 14.321] a $K\in\mathbb{R}^+$ such that $\forall i\in\{k,\ldots,\infty\}$ we have $\|L_i\|_{L(X,Y)}\leqslant K$. Hence $\forall i\in\{k,\ldots,\infty\}$ we have for $x=(x_1,\ldots,x_n)\in\prod_{j\in\{1,\ldots,n\}}X_j$ that

$$\|L_i(x)\|_Y\leqslant\|L_i\|_{L(X_1,\ldots,X_n;Y)}\cdot\prod_{j\in\{1,\ldots,n\}}\|x_j\|_j\leqslant K\cdot\prod_{j\in\{1,\ldots,n\}}\|x_j\|_j$$

which by [theorem: 14.301] proves

$$\|L(x)\|_Y=\left\|\lim_{i\rightarrow\infty}L_i(x)\right\|_Y\leqslant K\cdot\prod_{j\in\{1,\ldots,n\}}\|x_j\|_j$$

Hence using [eq: 14.184] and [theorem: 14.187] it follows that

$$L\in L(X_1,\ldots,X_n;Y) \tag{14.185}$$

Let $\varepsilon\in\mathbb{R}^+$ then for $x=(x_1,\ldots,x_n)\in\prod_{i\in\{1,\ldots,n\}}X_i$ with $\forall i\in\{1,\ldots,n\}$ $\|x_i\|_i=1$ we have by [eq: 14.183] that there exist a $N_x\in\{k,\ldots,\infty\}$ such that

$$\forall n\geqslant N_x \text{ we have } \|L_n(x)-L(x)\|_Y<\frac{\varepsilon}{4}$$

Further as $\{L_i\}_{i\in\{k,\ldots,\infty\}}$ is Cauchy there exist a $N\in\{k,\ldots,\infty\}$ such that $\forall m,l\geqslant N$ we have $\|L_m-L_l\|_{L(X_1,\ldots,X_n;Y)}<\frac{\varepsilon}{4}$ so that

$$\begin{aligned}\|L_m(x)-L_l(x)\|_Y &= \|(L_m-L_l)(x)\|_Y\\ &\leqslant \|L_m-L_l\|_{L(X_1,\ldots,X_n;Y)}\cdot\prod_{i\in\{1,\ldots,n\}}\|x_i\|_i\\ &= \|L_m-L_l\|_{L(X_1,\ldots,X_n;Y)}\\ &< \frac{\varepsilon}{4}\end{aligned}$$

Hence if $m\geqslant N$ then we have

$$\begin{aligned}\|L_m(x)-L(x)\|_Y &= \|L_m(x)-L_{\max(N,N_x)}(x)+L_{\max(N,N_x)}(x)-L(x)\|_Y\\ &\leqslant \|L_m(x)-L_{\max(N,N_x)}(x)\|_Y+\|L_{\max(N,N_x)}(x)-L(x)\|_Y\\ &< \frac{\varepsilon}{4}+\frac{\varepsilon}{4}\\ &= \frac{\varepsilon}{2}\end{aligned}$$

which by [theorem: 14.193] proves that $\|L_m-L\|_{L(X_1,\ldots,X_n;Y)}\leqslant\frac{\varepsilon}{2}<\varepsilon$. So $\{L_i\}_{i\in\{k,\ldots,\infty\}}$ converges to L in $\langle L(X_1,\ldots,X_n;Y),\|\|_{L(X_1,\ldots,X_n;Y)}\rangle$ proving that $\langle L(X_1,\ldots,X_n;Y),\|\|_{L(X_1,\ldots,X_n;Y)}\rangle$ is a Banach space. □

Corollary 14.336. *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.*

Proof. For $n\in\mathbb{N}$ we have either

$n=1$. By [definition: 14.194] $\langle L^1(X;Y),\|\|_{L^1(X;Y)}\rangle=\langle L(X,Y),\|\|_{L(X,Y)}\rangle$ which is Banach by the [theorem: 14.334].

$1<n$. Then by [definition: 14.194]

$$\langle L^n(X;Y),\|\|_{L^n(X;Y)}\rangle=\left\langle L^n\left(\underbrace{X,\ldots,X}_{n};Y\right),\|\|_{L^n\left(\underbrace{X,\ldots,X}_{n};Y\right)}\right\rangle$$

which is Banach by [theorem: 14.335]. □

14.10.2 Uniform convergence and Bounded functions

Definition 14.337. *Let $\varnothing\neq X$ be a set, $\langle Y,\|\|\rangle$ a normed space then a function*

$$f:X\rightarrow Y \textit{ is bounded} \text{ if } f(X) \textit{ is bounded}$$

In other words, using [theorem: 14.106], $f: X \rightarrow Y$ is bounded if there exist a $M \in \mathbb{R}^+$ such that $\forall x \in X$ we have $\|f(x)\| \leqslant M$. The set of all bounded functions between X and Y is noted by $\mathcal{B}(X, Y)$ hence

$$\mathcal{B}(X, Y) = \{f \in Y^X | f: X \rightarrow Y \text{ is bounded}\} \subseteq Y^X$$

Theorem 14.338. *Let $\varnothing \neq X$ be a set, $\langle Y, \|\|\rangle$ a normed space over the field F then $\langle \mathcal{B}(X, Y), +, \cdot\rangle$ is a sub-space of $\langle Y^X, +, \cdot\rangle$ [see theorem: 11.64]. Hence by [theorem: 11.57] $\langle \mathcal{B}(X, Y), +, \cdot\rangle$ is a vector space.*

Proof. For the neutral element C_0 in $\langle Y^X, +, \cdot\rangle$ we have $\forall x \in X$ that $\|C_0(x)\| = \|0\| = 0 < 1$ so that

$$C_0 \in \mathcal{B}(X, Y) \Rightarrow \varnothing \neq \mathcal{B}(X, Y)$$

Further if $f, g \in \mathcal{B}(X, Y)$ and $\alpha \in F$ then there exists $M_1, M_2 \in \mathbb{R}^+$ such that $\forall x \in X$ we have $\|f(x)\| \leqslant M_1 \wedge \|g(x)\| \leqslant M_2$. Take $M = \max(M_1, M_2, |\alpha| \cdot M_2 + 1)$ then we have

$$\|(f + \alpha \cdot g)(x)\| = \|f(x) + \alpha \cdot g(x)\| \leqslant \|f(x)\| + \|\alpha \cdot g(x)\| = \|f(x)\| + |\alpha| \cdot \|g(x)\| \leqslant M_1 + |\alpha| \cdot M_2 \leqslant M$$

proving that $f + \alpha \cdot g \in \mathcal{B}(X, Y)$. □

We define now a norm on $\mathcal{B}(X, Y)$

Theorem 14.339. *Let $\varnothing \neq X$ be a set, $\langle Y, \|\|\rangle$ a normed space over F then $\|\|_\infty$ defined by*

$$\|\|_\infty: \mathcal{B}(X, Y) \rightarrow \mathbb{R} \text{ by } \|\|_\infty = \sup(\{\|f(y)\| | y \in X\})$$

is well defined and $\|\|_\infty$ is a norm on $\mathcal{B}(X, Y)$ making $\langle \mathcal{B}(X, Y), \|\|_\infty\rangle$ a normed space. $\|\|_\infty$ is called the supremum norm.

Proof. As $X \neq 0$ there exist a $x \in X$ so that $\varnothing \neq \{\|f(y)\| | y \in X\}$. If $f \in \mathcal{B}(X, Y)$ then there exist a $M \in \mathbb{R}^+$ such that $\forall x \in X$ we have $\|f(x)\| \leqslant M$, hence $\{\|f(y)\| | y \in X\}$ is bounded above. As $\mathbb{R}$ is conditional complete [see theorem: 10.18] $\sup(\{\|f(y)\| | y \in X\})$ exist. Now we have

1. $\forall f \in \mathcal{B}(X, Y)$ we have as $0 \leqslant \|f(x)\|$ $\forall x \in X$ that $0 \leqslant \sup(\|f(y)\| | y \in X)$.
2. $\forall f \in \mathcal{B}(X, Y)$ and $\forall \alpha \in F$ then

   $$\{\|(\alpha \cdot f)(y)\| | y \in X\} = \{|\alpha| \cdot \|f(y)\| | y \in X\} = |\alpha| \cdot \{\|f(y)\| | y \in X\}$$

   so that

   $$\|\alpha \cdot f\|_\infty = \sup(|\alpha| \cdot \{\|f(y)\| | y \in X\}) \underset{[\text{theorem: } 10.22]}{=} |\alpha| \cdot \sup(\{\|f(y)\| | y \in X\}) = |\alpha| \cdot \|f\|_\infty$$

3. If $f, g \in \mathcal{B}(X, Y)$ then

   $$\begin{array}{rcl} \|f + g\|_\infty & = & \sup(\{\|(f + g)(y)\| | y \in X\}) \\ & = & \sup(\{\|f(y) + g(y)\| | y \in X\}) \\ & \leqslant_{[\text{theorem: } 3.75]} & \sup(\{\|f(y)\| + \|g(y)\| | y \in X\}) \\ & \leqslant_{[\text{theorem: } 3.75]} & \sup(\{\|f(y)\| | y \in X\} + \{\|g(y)\| | y \in X\}) \\ & \underset{[\text{theorem: } 10.23]}{=} & \sup(\{\|f(x)\| | x \in X\}) + \sup(\{\|g(x)\| | x \in X\}) \\ & \leqslant & \|f\|_\infty + \|g\|_\infty \end{array}$$

4. If $\|f\|_\infty = 0$ then $\sup(\{\|f(x)\| | x \in X\}) = 0$ so that $\forall x \in X$ $0 \leqslant \|f(x)\| \leqslant 0 \Rightarrow \|f(x)\| = 0$ proving, as $\|\|$ is a norm, that $f(x) = 0 = C_0(x)$. Hence $f = C_0$. □

We introduce now a stronger version of convergence for a sequence of functions to a normed space.

Definition 14.340. *Let $k \in \mathbb{N}_0, \varnothing \neq X$ be a set, $\langle Y, \|\|\rangle$ a normed space and $\{f_n\}_{n \in \{k, \ldots, \infty\}} \subseteq Y^X$ then $\{f_n\}_{n \in \{k, \ldots, \infty\}}$* ***converges uniformly*** *to a $f \in Y^X$ if $\forall \varepsilon \in \mathbb{R}$ there exist a $N \in \{k, \ldots, \infty\}$ such that $\forall n \geqslant N$ we have $\forall x \in X$ that*

$$\|f_n(x) - f(x)\| < \varepsilon$$

Note 14.341. *The N in the definition is not dependent on the x which would be the case in pointwise convergence. Hence uniform convergence is stronger then pointwise convergence.*

Theorem 14.342. *Let $\langle X, \|\|_X\rangle$ be a normed space, $\langle Y, \|\|_Y\rangle$ a normed space then if $\{f_n\}_{n\in\{k,\ldots,\infty\}} \subseteq \mathcal{B}(X,Y)$ is a sequence of continuous functions from X to Y uniformly converging to a function $f: X \rightarrow Y$ then f is continuous.*

Proof. Let $x \in X$ and $\varepsilon \in \mathbb{R}^+$ then by uniform convergence that there exist a $N \in \{k, \ldots, \infty\}$ such that $\forall n \geqslant N$ we have

$$\forall x \in X \quad \|f_n(x) - f(x)\|_Y < \frac{\varepsilon}{3}$$

As f_N is continuous we have that there exist a $\delta_N \in \mathbb{R}^+$ 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 we have

$$\begin{aligned}
\|f(x) - f(y)\|_Y &= \|f(x) - f_N(x) + f_N(x) - f_N(y) + f_N(y) - f(y)\|_Y \\
&\leqslant \|f(x) - f_N(x)\|_Y + \|f_N(x) - f_N(y)\|_Y + \|f_N(y) - f(y)\|_Y \\
&< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} \\
&= \varepsilon
\end{aligned}$$

proving by [theorem: 14.127] that f is continuous. □

It turns out that convergence in the supremum norm $\|\|_\infty$ is equal to uniform convergence.

Theorem 14.343. *Let $X \neq \varnothing$ and $\langle Y, \|\|\rangle$ a normed space and $\{f_n\}_{n\in\{k,\ldots,\infty\}} \subseteq \mathcal{B}(X,Y)$ then*

$$\lim_{n\rightarrow\infty} f_n = f \text{ using } \|\|_\infty \Leftrightarrow \{f_n\}_{n\in\{k,\ldots,\infty\}} \text{ converges uniformly to } f$$

Proof.

⇒. Given $\varepsilon \in \mathbb{R}^+$ there exist a $N \in \{k, \ldots, \infty\}$ such that $\forall n \geqslant N$

$$\sup(\{\|(f_n - f)(x)\| | x \in X\}) = \|f_n - f\|_\infty < \varepsilon$$

hence $\forall x \in X$ we have

$$\|f_n(x) - f(x)\| \leqslant \sup(\{\|(f_n - f)(x)\| | x \in X\}) < \varepsilon$$

proving uniform convergence.

⇐. Given $\varepsilon > 0$ there exists a $N \in \mathbb{N}$ such that if $n \geqslant N$ we have $\forall x \in X$ that $\|f(x) - f(x)\| < \varepsilon$ proving that ε is a upper bound of $\{\|f_n(x) - f(x)\| | x \in X\}$ so that

$$\|f_n - f\|_\infty = \sup(\{\|f_n(x) - f(x)\| | x \in X\}) < \varepsilon$$

proving that $\lim_{n\rightarrow\infty} f_n = f$ □

Theorem 14.344. *Let $X \neq \varnothing$ and $\langle Y, \|\|\rangle$ a Banach space then $\langle \mathcal{B}(X,Y), \|\|_\infty\rangle$ [see theorem: 14.339] is a Banach space.*

Proof. Let $\{f_n\}_{n\in\{k,\ldots,\infty\}} \subseteq \mathcal{B}(X,Y)$ be a Cauchy sequence in $\mathcal{B}(X,Y)$. Let $\varepsilon \in \mathbb{R}^+$ then there exists a $N \in \{k, \ldots, \infty\}$ such that $\forall n, m \geqslant N$ we have

$$\sup(\|f_n(x) - f_m(x)\| x \in X) = \|f_n(x) - f_m(x)\|_\infty < \varepsilon$$

Hence given $x \in X$ we have $\|f_n(x) - f_m(x)\| \leqslant \sup(\|f_n(x) - f_m(x)\| x \in X) < \varepsilon$ proving that $\{f_n(x)\}_{\in\{k,\ldots\infty\}}$ is a Cauchy sequence, hence as Y is complete there exist a unique $\lim_{n\rightarrow\infty} f_n(x)$. This allows use to define a function

$$f: X \rightarrow Y \text{ by } f(x) = \lim_{n\rightarrow\infty} f_n(x) \tag{14.186}$$

Next we prove that f is bounded so that $f \in \mathcal{B}(X,Y)$. Let $x \in X$ then as $1 \in \mathbb{R}^+$ we have by the definition of the limit that there exist a $N_x \in \{k, \ldots, \infty\}$ [depending on x] such that $\forall n \geqslant N_x$ we have $\|f_n(x) - f(x)\| < 1$ so that

$$\|f(x)\| \leqslant \|f_n(x) - f(x)\| + \|f_n(x)\| < 1 + \|f_n(x)\| \leqslant 1 + \sup(\{\|f_n(x)\| | x \in X\}) = 1 + \|f_n\|_\infty$$

hence

$$\forall x \in X \ \exists N_x \in \mathbb{R}^+ \text{ such that } \forall n \geqslant N_x \ \|f(x)\| < 1 + \|f_n\|_\infty \tag{14.187}$$

Further as $\{f_n\}_{n\in\{k,\ldots,\infty\}}$ is Cauchy in $\mathcal{B}(X,Y)$ and $1\in\mathbb{R}^+$ there exist a $N\in\{k,\ldots,\infty\}$ such that $\forall m\geqslant N\Rightarrow m,N\geqslant N$ we have $\|f_m-f_N\|_\infty<1$, hence $\|f_m\|_\infty\leqslant\|f_N\|_\infty+\|f_m-f_N\|_\infty<\|f_N\|_\infty+1$. Summarized

$$\exists N\in\{k,\ldots,\infty\} \text{ such that } \forall m\geqslant N \text{ we have } \|f_m\|_\infty<\|f_N\|_\infty+1 \tag{14.188}$$

Let $x\in X$ and take $M_x=\max(N,N_x)$ then we have

$$\begin{array}{rcl} \|f(x)\| & \leqslant_{[\text{eq: } 14.187]} & 1+\|f_{M_x}\|_\infty \\ & \leqslant_{[\text{eq: } 14.188]} & 1+1+\|f_N\|_\infty \\ & \leqslant & M \end{array}$$

where $M=2+\|f_N\|_\infty$ independent of x, proving that $\|f\|_\infty=\sup(\{\|f(x)\||x\in X\})\leqslant M$. Hence we conclude that f is bounded or

$$f\in\mathcal{B}(X,Y) \tag{14.189}$$

Next we must prove that $\lim_{n\to\infty} f_n=f$ using the norm $\|\|_\infty$. Let $\varepsilon\in\mathbb{R}^+$ then for $x\in X$ we have using [eq: 14.186] that there exists a $N_x\in\{k,\ldots,\infty\}$ such that $\forall n\geqslant N_x$ we have $\|f(x)-f_n(x)\|<\frac{\varepsilon}{2}$. Further as $\{f_n\}_{n\in\{k,\ldots,\infty\}}$ is Cauchy [using the norm $\|\|_\infty$] there exist a $N\in\{k,\ldots,\infty\}$ such that $\forall n,m\geqslant N$ we have $\|f_n-f_m\|<\frac{\varepsilon}{2}$. Let $x\in X$ then for $n\geqslant N$ we have $\max(n,N_x)\geqslant N,N_x$ so that

$$\begin{array}{rcl} \|(f_n-f)(x)\| & = & \|f_n(x)-f(x)\| \\ & = & \|f_n(x)-f_{\max(N,N_x)}(x)+f_{\max(N,N_x)}(x)-f(x)\| \\ & \leqslant & \|f_n(x)-f_{\max(N,N_x)}(x)\|+\|f_{\max(N,N_x)}(x)-f(x)\| \\ & \leqslant & \|(f_n-f_{\max(N,N_x)})(x)\|+\|f_{\max(N,N_x)}(x)-f(x)\| \\ & \leqslant & \sup(\{\|(f_n-f_{\max(N,N_x)})(x)\||x\in X|\})+\|f_{\max(N,N_x)}(x)-f(x)\| \\ & = & \|f_n-f_{\max(N,N_x)}\|_\infty+\|f_{\max(N,N_x)}(x)-f(x)\| \\ & <_{N,N_x\leqslant\max(N,N_x)} & \dfrac{\varepsilon}{2}+\dfrac{\varepsilon}{2} \\ & = & \varepsilon \end{array}$$

which as $x\in X$ was chosen arbitrary proves that $\|f_n-f\|_\infty=\sup(\{\|(f_n-f)(x)\||x\in X\})<\varepsilon$. Hence it follows that

$$\lim_{n\to\infty} f_n\to f \text{ using the norm } \|\|_\infty$$

□

14.10.3 Series

Series are a extension of finite sums to infinite sums. Note that in contrast with finite sums infinite sums do not have in general the properties of commutativity and associativity.

Definition 14.345. (Series) *Let $k\in\mathbb{N}_0$, $\langle X,\|\|\rangle$ be a normed space and $\{x_i\}_{i\in\{k,\ldots,\infty\}}\subseteq X$ a sequence such that $\lim_{n\to\infty}(\sum_{i=k}^\infty x_i)$ exist then we say that the **series $\sum_{i=k}^\infty x_i$ converges** and note the limit as $\sum_{i=k}^\infty x_i$, hence*

$$\sum_{i=k}^\infty x_i=\lim_{n\to\infty}\left(\sum_{i=k}^n x_i\right)$$

Remark 14.346. To avoid excessive notation instead of saying '$\{x\}_{i\in\{k,\ldots,\infty\}}$ is a sequence such that the series $\sum_{i=k}^\infty x_i$ converges' we say just 'the series $\sum_{i=k}^\infty x_i$ converges' where the sequence $\{x_i\}i\in\{k,\ldots,\infty\}$ is implicit assumed.

A example of a series in $\mathbb{R}$ is the power series, to prove this we first need a little lemma.

Lemma 14.347. *Let $\lambda\in\mathbb{R}\setminus\{1\}$ then if $k\in\mathbb{N}$ and $n\in\mathbb{N}_0$ we have that*

$$\sum_{i=1}^k \lambda^{n+(i-1)}=\lambda^n\cdot\frac{1-\lambda^k}{1-\lambda}$$

In particular we have

1. *If $n=0$ then $\sum_{i=1}^k\lambda^{(i-1)}=\frac{1-\lambda^k}{1-\lambda}$*
2. *If $n=1$ then $\sum_{i=1}^k\lambda^i=\lambda\cdot\frac{1-\lambda^k}{1-\lambda}$*

Proof. We prove this by induction so let

$$S=\left\{k\in\mathbb{N}|\sum_{i=1}^{k}\lambda^{n+(i-1)}=\lambda^{n}\cdot\frac{1-\lambda^{k}}{1-\gamma}\right\}$$

then we have:

$1\in S$. 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 $1\in S$

$k\in S\Rightarrow k+1\in S$. 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 $k+1\in S$ □

Theorem 14.348. *Let $\lambda\in\mathbb{R}$ with $0\leqslant\lambda<1$ then $\sum_{i=1}^{\infty}\lambda^{i}$ converges and $\sum_{i=1}^{\infty}\lambda^{i}=\frac{\lambda}{1-\lambda}$*

Proof. Let $k\in\mathbb{N}$ then we have that

$$\begin{aligned}\left|\sum_{i=1}^{k}\lambda^{i}-\frac{\lambda}{1-\lambda}\right| &= \left|\sum_{i=1}^{k}\lambda\cdot\lambda^{i-1}-\frac{\lambda}{1-\lambda}\right|\\ &= \left|\lambda\cdot\sum_{i=1}^{k}\lambda^{i-1}-\frac{\lambda}{1-\gamma}\right|\\ &\underset{[\text{theorem: }14.347]}{=} \left|\lambda\cdot\frac{1-\lambda^{k}}{1-\gamma}-\frac{\lambda}{1-\lambda}\right|\\ &= \left|\frac{\lambda-\lambda^{k+1}-\lambda}{1-\lambda}\right|\\ &= \left|\frac{-\lambda^{k+1}}{1-\lambda}\right|\\ &= \frac{\lambda^{k+1}}{1-\lambda}\\ &= \frac{\lambda}{1-\lambda}\cdot\lambda^{k}\end{aligned}$$

Using [example: 14.292] we have that $\lim_{k\to\infty}\lambda^{k}=0$, hence given $\varepsilon\in\mathbb{R}^{+}$ there exist a $N\in\{1,\ldots,\infty\}$ such that $\forall k\geqslant N$ $\lambda^{k}=|\lambda^{k}-0|<\varepsilon\cdot\frac{1-\lambda}{1+\lambda}$ so that $\frac{\lambda}{1-\lambda}\cdot\lambda^{k}<\frac{\lambda}{1-\lambda}\cdot\frac{1-\lambda}{\lambda+1}\cdot\varepsilon=\frac{\lambda}{1+\lambda}\cdot\varepsilon<\varepsilon$, hence

$$\left|\sum_{i=1}^{k}\lambda^{i}-\frac{\lambda}{1-\lambda}\right|<\varepsilon$$

proving that

$$\lim_{k\to\infty}\sum_{i=1}^{k}\lambda^{i}=\frac{\lambda}{1-\lambda}$$ □

If a series is convergent then the terms must tend to zero as is show in the following theorem.

Theorem 14.349. *Let $k\in\mathbb{N}_0$, $\langle X,\|\|\rangle$ be a normed space then if $\sum_{i=k}^{\infty}x_i$ converges then $\lim_{i\rightarrow\infty}x_i=0$*

Proof. Let $\varepsilon\in\mathbb{R}^+$. As $\lim_{n\rightarrow\infty}\sum_{i=k}^{n}x_i$ exist we have that $\{\sum_{i=k}^{n}x_i\}_{n\in\{k,\ldots,\infty\}}$ is Cauchy so there exist a $N\in\{k,\ldots,\infty\}$ such that $n,m\geqslant\mathbb{N}$ we have that $\|\sum_{i=k}^{n}x_i-\sum_{i=k}^{m}x_i\|<\varepsilon$. In particular we have $\forall n\geqslant N\Rightarrow n+1\geqslant N$ that $\|x_n-0\|=\|x_n\|=\|\sum_{i=k}^{n+1}x_i-\sum_{i=k}^{n}x_i\|<\varepsilon$. Hence

$$\lim_{n\rightarrow\infty}x_n=0$$

□

In a Banach space we have the following sufficient and necessary condition for convergence of a series.

Theorem 14.350. *Let $k\in\mathbb{N}_0$, $\langle X,\|\|\rangle$ be a Banach space then*

$$\sum_{i=k}^{\infty}x_i \text{ converges} \Leftrightarrow \forall\varepsilon\in\mathbb{R}^+\ \exists N\in\{k,\ldots,\infty\} \text{ such that } \forall n,m\geqslant N \text{ with } n\leqslant m \text{ we have } \left\|\sum_{i=n}^{m}x_i\right\|<\varepsilon$$

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 [theorem: 14.322], 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 $\|\sum_{i=n}^{m}x_i\|\underset{11.23}{=}\|\sum_{i=k}^{m}x_i-\sum_{i=k}^{n-1}x_i\|=\|s_m-s_{n-1}\|<\varepsilon$.
- **⇐.** 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 $\sum_{i=n+1}^{m}x^i\underset{11.23}{=}\sum_{i=k}^{m}x^i-\sum_{i=k}^{n}x^i=s_m-s_n$ 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 14.351. *Let $k,l\in\mathbb{N}_0$, $\langle X,\|\|\rangle$ a normed space and $\{x_i\}_{i\in\{k+l,\ldots,\infty\}}\subseteq X$ a sequence in X then we have*

$$\sum_{i=k}^{\infty}x_{i+l} \text{ converges} \Leftrightarrow \sum_{i=k+l}^{\infty}x_i \text{ converges}$$

Further if $\sum_{i=k}^{\infty}x_{i+l}$ or $\sum_{i=k+l}^{\infty}x_i$ converges then

$$\sum_{i=k}^{\infty}x_{i+l}=\sum_{i=k+l}^{\infty}x_i$$

Proof.

- **⇒.** Let $\varepsilon\in\mathbb{R}^+$. As $\sum_{i=k}^{\infty}x_{i+l}$ converges there exist a $x\in X$ such that there exist a $N'\in\{k,\ldots,\infty\}$ such that $\forall n\geqslant N'$ we have that $\|\sum_{i=k}^{n}x_{i+l}-x\|<\varepsilon$. Hence if we take $N=N'+l\in\{k+l,\ldots,\infty\}$ then if $n\geqslant N$ we have that $n\geqslant N'+l\Rightarrow n-l\geqslant N'$ so that
$$\left\|\sum_{i=k+l}^{n}x_i-x\right\|=\left\|\sum_{i=k}^{n-l}x_{i+l}-x\right\|<\varepsilon.$$
This proves that $\sum_{i=k+l}^{\infty}x_i$ converges to x or
$$\sum_{i=k+l}^{\infty}x_i=x=\sum_{i=k}^{\infty}x_{i+l}$$
- **⇐.** Let $\varepsilon\in\mathbb{R}^+$. As $\sum_{i=k+l}^{\infty}x_i$ converges there exist a $x\in X$ and a $N'\in\{k+l,\ldots,\infty\}$ such that $\forall n\geqslant N'$ we have $\|\sum_{i=k+l}^{n}x_i-x\|<\varepsilon$. Take $N=N'-l\in\{k,\ldots,\infty\}$ then $\forall n\geqslant N$ we have $n+l\geqslant N'$ so that
$$\left\|\sum_{i=k}^{n}x_{i+l}-x\right\|=\left\|\sum_{i=k+l}^{n+l}x_i-x\right\|<\varepsilon$$
proving that $\sum_{i=k}^{\infty}x_{i+l}$ converges to x or
$$\sum_{i=k}^{\infty}x_{i+l}=x=\sum_{i=k+l}^{\infty}x_i$$
□

Theorem 14.352. *Let $k\in\mathbb{N}_0$, $\langle X,\|\|\rangle$ a normed space, $m\in\{k+1,\ldots,\infty\}$ and $\{x_i\}_{i\in\{k,\ldots,n\}}\subseteq X$ then*

$$\sum_{i=k}^{\infty}x_i \text{ converges} \Leftrightarrow \sum_{i=m}^{\infty}x_i \text{ converges}$$

Further if $\sum_{i=k}^{\infty} x_i$ *or* $\sum_{i=m}^{\infty} x_i$ *converges then*

$$\sum_{i=k}^{\infty} x_i = \sum_{i=k}^{m-1} x_i + \sum_{i=m}^{\infty} x_i$$

Proof.

⇒. Let $\varepsilon \in \mathbb{R}^+$. As $\sum_{i=k}^{\infty} x_i$ converges there exist a $x \in X$ and a $N \in \{k, \ldots, \infty\}$ such that $\forall n \geqslant N$ we have $\|\sum_{i=k}^{n} x_i - x\| < \varepsilon$, hence we have for $n \geqslant \max(m, N)$ that

$$\left\|\sum_{i=m}^{n} x_i - \left(x - \sum_{i=k}^{m-1} x_i\right)\right\| = \left\|\sum_{i=m}^{n} x_i + \sum_{i=k}^{m-1} x_i - x\right\| \underset{[\text{theorem: } 11.23]}{=} \left\|\sum_{i=k}^{n} x_i - x\right\| < \varepsilon$$

proving that $\sum_{i=m}^{\infty} x_i$ converges and that $\sum_{i=m}^{\infty} x_i = x - \sum_{i=k}^{m-1} x_i = \sum_{i=k}^{\infty} x_i - \sum_{i=k}^{m-1} x_i$. Hence

$$\sum_{i=k}^{\infty} x_i = \sum_{i=k}^{m-1} x_i + \sum_{i=m}^{\infty} x_i$$

⇐. Let $\varepsilon \in \mathbb{R}^+$. As $\sum_{i=m}^{\infty} x_i$ converges there exist a $x \in X$ and a $N \in \{m, \ldots, \infty\}$ such that $\forall n \geqslant N$ we have $\|\sum_{i=m}^{n} x_i - x\| < \varepsilon$. Then we have

$$\left\|\sum_{i=k}^{n} x_i - \left(\sum_{i=k}^{m-1} x_i + x\right)\right\| = \left\|\sum_{i=k}^{n} x_i - \sum_{i=k}^{m-1} x_i - x\right\| = \left\|\sum_{i=m}^{n} x_i - x\right\| < \varepsilon$$

proving that $\sum_{i=k}^{\infty} x_i$ converges and that

$$\sum_{i=k}^{\infty} x_i = \sum_{i=k}^{m-1} x_i + x = \sum_{i=k}^{m-1} x_i + \sum_{i=m}^{\infty} x_i$$

□

Theorem 14.353. *Let* $\langle X, \|\|\rangle$ *be a normed space and* $\sum_{i=k}^{\infty} x_i$ *a convergent space [so that by [theorem: 14.352]* $\forall m \in \{k, \ldots, \infty\}$ $\sum_{i=m}^{\infty} x_i$ *converges and thus exist] then*

$$\lim_{m \rightarrow \infty} \left(\sum_{i=m}^{\infty} x_i\right) = 0$$

Proof. Let $\varepsilon \in \mathbb{R}^+$. As $\sum_{i=k}^{\infty} x_i$ converges there exist a $N \in \{, \ldots, \infty\}$ so that $\forall n \geqslant N$ we have $\|\sum_{i=1}^{n} x_i - x\| < \varepsilon$ where $x = \sum_{i=k}^{\infty} x_i$. Let $m \geqslant N + 1 \Rightarrow m - 1 \geqslant N$ then

$$\left\|\sum_{i=m}^{\infty} x_i - 0\right\| = \left\|\sum_{i=m}^{\infty} x_i\right\| \underset{[\text{theorem: } 14.352]}{=} \left\|\sum_{i=k}^{\infty} x_i - \sum_{i=k}^{m-1} x\right\| = \left\|\sum_{i=k}^{m-1} x_i - x\right\| < \varepsilon$$

proving that $\{\sum_{i=m}^{\infty} x_i\}_{m \in \{k, \ldots, \infty\}}$ converges to 0, Hence

$$\lim_{m \rightarrow \infty} \left(\sum_{i=m}^{\infty} x_i\right) = 0$$

□

The sum and scalar product of series is a series.

Theorem 14.354. *Let* $\langle X, \|\|\rangle$ *be a normed space then we have:*

1. *If* $\sum_{i=k}^{\infty} x_i$ *and* $\sum_{i=k}^{\infty} y_i$ *are convergent series then* $\sum_{i=k}^{\infty} (x_i + y_i)$ *is convergent and*

$$\sum_{i=k}^{\infty} (x_i + y_i) = \sum_{i=k}^{\infty} x_i + \sum_{i=k}^{\infty} y_i$$

2. *If* $\sum_{i=k}^{\infty} x_i$ *is convergent and* $\alpha \in \mathbb{K}$ *then* $\sum_{i=k}^{\infty} (\alpha \cdot x_i)$ *is convergent and*

$$\sum_{i=k}^{\infty} (\alpha \cdot x_i) = \alpha \cdot \sum_{i=k}^{\infty} x_i$$

Proof.

1. As $\sum_{i=k}^{\infty} x_i$ and $\sum_{i=k}^{\infty} y_i$ are convergent

$$\lim_{n \rightarrow \infty} \sum_{i=k}^{\infty} x_i \text{ and } \lim_{n \rightarrow \infty} \sum_{i=k}^{n} y_i \text{ exists}$$

Hence using [theorem: 14.297] $\lim_{n\rightarrow\infty}\sum_{i=k}^{\infty}(x_i+y_i)=\lim_{n\rightarrow\infty}\left(\sum_{i=k}^{n}x_i+\sum_{i=k}^{n}y_i\right)$ exists and is equal to $\lim_{n\rightarrow\infty}\sum_{i=k}^{n}x_i+\lim_{n\rightarrow\infty}\sum_{i=k}^{n}y_i$. Hence by definition $\sum_{i=k}^{\infty}(x_i+y_i)$ is convergent and

$$\sum_{i=k}^{\infty}(x_i+y_i)=\sum_{i=k}^{\infty}x_i+\sum_{i=k}^{\infty}y_i$$

2. As $\sum_{i=k}^{\infty}x_i$ is convergent $\lim_{n\rightarrow\infty}\sum_{i=k}^{n}x_i$ exist, so that by [theorem: 14.298] $\lim_{n\rightarrow\infty}\left(\sum_{i=k}^{n}\alpha\cdot x_i\right)=\lim_{n\rightarrow\infty}\alpha\cdot\left(\sum_{i=k}^{n}x_i\right)$ exist and is equal to $\alpha\cdot\lim_{n\rightarrow\infty}\sum_{i=k}^{n}x_i$. Hence by definition $\sum_{i=k}^{\infty}(\alpha\cdot x_i)$ is convergent and

$$\sum_{i=k}^{\infty}(\alpha\cdot x_i)=\alpha\cdot\sum_{i=k}^{\infty}x_i$$

□

Corollary 14.355. *Let $\langle\mathbb{C},||\rangle$ be the normed space of complex numbers with the norm $||$ and $\{z_i\}_{i\in\{k,\ldots,\infty\}}\subseteq\mathbb{C}$ then*

$$\sum_{i=k}^{\infty}x_i \text{ converges} \Leftrightarrow \sum_{i=k}^{\infty}\mathrm{Img}(z_i) \text{ and } \sum_{i=k}^{\infty}\mathrm{Re}(z_i) \text{ converges}$$

Further if $\sum_{i=k}^{\infty}z_i$ converges then

$$\mathrm{Img}\left(\sum_{i=k}^{\infty}z_i\right)=\sum_{i=k}^{\infty}\mathrm{Img}(z_i) \text{ and } \mathrm{Re}\left(\sum_{i=k}^{\infty}z_i\right)=\sum_{i=k}^{\infty}\mathrm{Re}(z_i)$$

so that

$$\sum_{i=k}^{\infty}z_i=\sum_{i=k}^{\infty}\mathrm{Re}(z_i)+i\cdot\sum_{i=k}^{\infty}\mathrm{Img}(z_i)$$

Proof.

⇒. Assume that $\sum_{i=k}^{\infty}z_i$ converges to z then given $\varepsilon>0$ there exists a $N\in\{k,\ldots,\infty\}$ such that $\forall n\in\{N,\ldots,\infty\}$ $\left|\sum_{i=k}^{n}z_i-z\right|<\varepsilon$. Now

$$\begin{aligned}\left|\left(\sum_{i=k}^{n}\mathrm{Re}(z_i)\right)-\mathrm{Re}(z)\right| &= \left|\mathrm{Re}\left(\sum_{i=k}^{n}z_i-z\right)\right|\\ &\leqslant_{[\text{theorem: } 10.83]} \left|\sum_{i=k}^{\infty}z_i-z\right|\\ &< \varepsilon\end{aligned}$$

proving that

$$\sum_{i=k}^{\infty}\mathrm{Re}(z_i) \text{ converges to } \mathrm{Re}(z)=\mathrm{Re}\left(\sum_{i=k}^{\infty}x_i\right)$$

Similar

$$\begin{aligned}\left|\left(\sum_{i=k}^{n}\mathrm{Img}(z_i)\right)-\mathrm{Img}(z)\right| &= \left|\mathrm{Img}\left(\sum_{i=k}^{n}z_i-z\right)\right|\\ &\leqslant_{[\text{theorem: } 10.83]} \left|\sum_{i=k}^{\infty}z_i-z\right|\\ &< \varepsilon\end{aligned}$$

proving that

$$\sum_{i=k}^{\infty}\mathrm{Img}(z_i) \text{ converges to } \mathrm{Img}(z)=\mathrm{Img}\left(\sum_{i=k}^{\infty}x_i\right)$$

⇐. Assume that $\sum_{i=k}^{\infty}\mathrm{Re}(z_i)$ and $\sum_{i=k}^{\infty}\mathrm{Img}(z_i)$ converges. As $\forall i\in\{k,\ldots,\infty\}$ $z_i=\mathrm{Re}(z_i)+i\cdot\mathrm{Img}(z_i)$ it follows from the previous theorem [see theorem: 14.354]] that $\sum_{i=k}^{\infty}z_i$ converges. □

14.10.3.1 Series of non negative numbers

Series are in general not commutative, however series of non negative real numbers (and later absolute convergent series) are commutative. First we look at some equivalences of convergence of series of non negative real numbers

Theorem 14.356. *Let $k\in\mathbb{N}_0$, $\{x_i\}_{i\in\{k,\ldots,\infty\}}\subseteq[0,\infty[$ be a sequence of non negative real numbers then we have the following equivalences:*

1. *$\sum_{i=k}^{\infty}x_i=x$ [in other words $\sum_{i=k}^{\infty}x_i$ converges to x]*

2. $\sup\left(\left\{\sum_{i=k}^{n} x_i | n \in \{k, \ldots, \infty\}\right\}\right)$ *exists and* $x = \sup\left(\left\{\sum_{i=k}^{n} x_i | n \in \{k, \ldots, \infty\}\right\}\right)$
3. $\sup\left(\left\{\sum_{i \in K} x_i | K \in \mathcal{P}(\{k, \ldots, \infty\}) \wedge K \text{ is finite }\right\}\right)$ *exists and*

$$x = \sup\left(\left\{\sum_{i \in K} x_i | K \in \mathcal{P}(\{k, \ldots, \infty\}) \wedge K \textit{ is finite }\right\}\right)$$

4. $\forall \varepsilon \in \mathbb{R}^+$ *there exist a* $N \in \{k, \ldots, \infty\}$ *such that* $\forall m \in \mathbb{N}$ *and* $n \geqslant N$ *we have* $\sum_{i=n+1}^{n+m} x_i < \varepsilon$
5. $\forall \varepsilon \in \mathbb{R}^+$ *there exist a finite* $K \subseteq \{k, \ldots, \infty\}$ *such that* $\forall H \subseteq \{k, \ldots, \infty\}$ *with* H *finite and* $K \bigcap H = \varnothing$ *we have* $\sum_{i \in H} x_i < \varepsilon$

Proof. First if $s \in \left\{\sum_{i=k}^{n} x_i | n \in \{k, \ldots, \infty\}\right\}$ then $\exists n \in \{k, \ldots, \infty\}$ so that

$$s = \sum_{i=k}^{n} x_i \underset{[\text{theorem: } 11.35]}{=} \sum_{i \in \{k, \ldots, n\}} x_i \in \left\{\sum_{i \in K} x_i | K \subseteq \{k, \ldots, \infty\} \wedge K \text{ is finite}\right\}$$

proving that

$$\left\{\sum_{i=k}^{n} x_i | n \in \{k, \ldots, \infty\}\right\} \subseteq \left\{\sum_{i \in K} x_i | K \subseteq \{k, \ldots, \infty\} \wedge K \text{ is finite}\right\} \tag{14.190}$$

Next

1 ⇔ 2. Let $n \in \{k, \ldots, \infty\}$ then $\sum_{i=k}^{n} x_i \leqslant \sum_{i=k}^{n} x_i + x_{n+1} = \sum_{i=k}^{n+1} x_i$, hence $\left\{\sum_{i=k}^{n}\right\}_{n \in \{k, \ldots, \infty\}}$ is a increasing sequence. So

$$\sum_{i=k}^{\infty} x_i \text{ exists and } \sum_{i=k}^{\infty} x_i = x \quad \underset{\text{definition}}{\Leftrightarrow}$$

$$\lim_{n\to\infty} \sum_{i=k}^{n} x_i \text{ exists } \lim_{n\to\infty} \sum_{i=k}^{n} x_i = x \quad \underset{[\text{theorem: } 14.311]}{\Leftrightarrow}$$

$$\sup\left(\left\{\sum_{i=k}^{n} x_i | n \in \{k, \ldots, \infty\}\right\}\right) \text{ exist and } \sup\left(\left\{\sum_{i=k}^{n} x_i | n \in \{k, \ldots, \infty\}\right\}\right) = x$$

2 ⇒ 3. Let $s \in \left\{\sum_{i \in K} x_i | K \in \mathcal{P}(\{k, \ldots, \infty\}) \wedge K \text{ is finite }\right\}$ then there exists a finite $K \subseteq \{k, \ldots, \infty\}$ so that $s = \sum_{i \in K} x_i$. Take $n = \max(K)$ then $K \subseteq \{k, \ldots, m\}$ so that

$$\sum_{i \in K} x_i \leqslant \sum_{i \in K} x_i + \sum_{i \in \{k, \ldots, m\} \setminus K} x_i = \sum_{i \in \{k, \ldots, m\}} x_i = \sum_{i=k}^{m} x_i \leqslant \sup\left(\left\{\sum_{i=k}^{n} x_i | n \in \{k, \ldots, \infty\}\right\}\right)$$

hence $\left\{\sum_{i \in K} x_i | K \in \mathcal{P}(\{k, \ldots, \infty\}) \wedge K \text{ is finite }\right\}$ is bounded above by $\sup\left(\left\{\sum_{i=k}^{n} x_i | n \in \{k, \ldots, \infty\}\right\}\right)$ so that by [theorem: 10.18]

$$\sup\left(\left\{\sum_{i \in K} x_i | K \in \mathcal{P}(\{k, \ldots, \infty\}) \wedge K \text{ is finite }\right\}\right) \text{ exists} \tag{14.191}$$

and

$$\sup\left\{\sum_{i \in K} x_i | K \in \mathcal{P}(\{k, \ldots, \infty\}) \wedge K \text{ is finite }\right\} \leqslant \sup\left(\left\{\sum_{i=k}^{n} x_i | n \in \{k, \ldots, \infty\}\right\}\right) \tag{14.192}$$

Combining [eq: 14.190] with [theorem: 3.74] gives

$$\sup\left(\left\{\sum_{i=k}^{n} x_i | n \in \{k, \ldots, \infty\}\right\}\right) \leqslant \sup\left\{\sum_{i \in K} x_i | K \in \mathcal{P}(\{k, \ldots, \infty\}) \wedge K \text{ is finite }\right\}$$

which combined with [eq: 14.192] proves that

$$\sup\left(\left\{\sum_{i=k}^{n} x_i | n \in \{k, \ldots, \infty\}\right\}\right) = \sup\left(\left\{\sum_{i \in K} x_i | K \in \mathcal{P}(\{k, \ldots, \infty\}) \wedge K \text{ is finite }\right\}\right)$$

3 ⇒ 2. Let $s \in \left\{\sum_{i=k}^{n} x_i | n \in \{k, ..., \infty\}\right\}$ then from [eq: 14.190] $s \in \left\{\sum_{i \in K} x_i | K \in \mathcal{P}(\{k, ..., \infty\}) \wedge K \text{ is finite}\right\}$ so that $s \leqslant \sup\left(\left\{\sum_{i \in K} x_i | K \in \mathcal{P}(\{k, \ldots, \infty\}) \wedge K \text{ is finite}\right\}\right)$ hence by [theorem: 10.18]

$$\sup\left(\left\{\sum_{i=k}^{n} x_i | n \in \{k, \ldots, \infty\}\right\}\right) \text{ exists}$$

further as $\{\sum_{i=k}^{n} x_i\}_{n\in\{k,\ldots,\infty\}}$ is increasing we have by [theorem: 14.311] that

$$\sup\left(\left\{\sum_{i=k}^{n} x_i | n\in\{k,\ldots,\infty\}\right\}\right)=\lim_{n\to\infty}\sum_{i=k}^{n} x_i=\sum_{i=k}^{\infty} x_i$$

1 ⇒ 4. As $\sum_{i=k}^{\infty} x_i$ converges $\lim_{n\to\infty}\sum_{i=k}^{n} x_i$ exists so that by [theorem: 14.322] $\{\sum_{i=k}^{n} x_i\}_{n\in\{k,\ldots,\infty\}}$ is Cauchy. Hence given $\varepsilon\in\mathbb{R}^+$ there exist a $N\in\{k,\ldots,\infty\}$ such that $\forall n,m\geqslant N$ we have $|\sum_{i=k}^{n} x_i-\sum_{i=k}^{m} x_i|<\varepsilon$. So if $m\in\mathbb{N}$ and $n\geqslant N$ then we have $N+m$, $n\geqslant N$ hence

$$\sum_{i=n+1}^{n+m} x_i=\left|\sum_{i=n+1}^{n+m} x_i\right|=\left|\sum_{i=k}^{n+m} x_i-\sum_{i=k}^{n} x_i\right|<\varepsilon$$

4 ⇒ 1. Let $\varepsilon\in\mathbb{R}^+$ then by the hypothesis there exist a $N\in\{k,\ldots,\infty\}$ such that $\forall m\in\mathbb{N}$ and $\forall n\geqslant m$ that $\sum_{i=n+1}^{n+m} x_i<\varepsilon$. Take r, $s\geqslant N$ then without loosing generality we my assume that $s\leqslant r$. If $r=s$ then $|\sum_{i=k}^{r} x_i-\sum_{i=k}^{s} x_i|=0<\varepsilon$ and if $s<r$ then

$$\left|\sum_{i=k}^{r} x_i-\sum_{i=k}^{s} x_i\right|=\left|\sum_{i=s+1}^{r} x_i\right|=\sum_{i=s+1}^{r} x_i=\sum_{i=s+1}^{(r-s)+s} x_i<_{0<r-s}\varepsilon$$

proving that $\{\sum_{i=k}^{n} x_i\}_{n\in\{k,\ldots,\infty\}}$ is a Cauchy sequence. As by [theorem: 14.330] $\langle\mathbb{R},||\rangle$ is a Banach space hence complete it follows that $\lim_{n\to\infty}\sum_{i=k}^{n} x_i$ exists or that

$$\sum_{i=1}^{\infty} x_i \text{ converges}$$

4 ⇒ 5. Let $\varepsilon\in\mathbb{R}^+$ then by the hypothesis there exist a $N\in\{k,\ldots,\infty\}$ such that $\forall m\in\mathbb{N}$ and $\forall n\geqslant m$ that $\sum_{i=n+1}^{n+m} x_i<\varepsilon$. Take $K=\{k,\ldots,N\}$ then if $H\subseteq\{k,\ldots,\infty\}$ is finite and $K\bigcap H=\varnothing$ we must have that $H\subseteq\{N+1,\ldots,\infty\}$. We must now consider two cases for H:

- $\boldsymbol{H=\varnothing}$. Then $\sum_{i\in H} x_i=0<\varepsilon$
- $\boldsymbol{H\neq\varnothing}$. Then $N<\min(H)\leqslant\max(H)$ so that $m=\max(H)-N>0$ and

  $$H\subseteq\{N+1,\ldots,\max(H)\}=\{N+1,\ldots,N+m\}$$

  so that

  $$\sum_{i\in H} x_i\leqslant\sum_{i\in H} x_i+\sum_{i\in\{N+1,\ldots,N+m\}\setminus H} x_i=\sum_{i\in\{N+1,\ldots,N+m\}} x_i<\varepsilon$$

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 $\sum_{i\in H} x_i<\varepsilon$. Take $K'=K\bigcup\{k\}$ then $K'\neq\varnothing$ so that $N=\max(K')$ exists. Then $\forall n\geqslant N$ and $m\in\mathbb{N}$ we have

$$K\bigcap\{n+1,\ldots,n+m\}\subseteq K'\bigcap\{n+1,\ldots,n+m\}\underset{\max(K')\leqslant N\leqslant n<n+1}{=}\varnothing$$

hence for $H=\{n+1,\ldots,n+m\}$ we have $K\bigcap H=\varnothing$ so that

$$\sum_{i=n+1}^{n+m} x_i=\sum_{i\in\{n+1,\ldots,n+m\}} x_i=\sum_{i\in H} x_i<\varepsilon$$ □

Corollary 14.357. *Let $k\in\mathbb{N}_0$, $\{x_i\}_{i\in\{k,\ldots,\infty\}}\subseteq[0,\infty[$ be a sequence of non negative real numbers such that $\exists l\in\{k,\ldots,n\}$ with $0<x_l$ then if $\sum_{i=k}^{\infty} x_i$ converges we have $0<\sum_{i=k}^{\infty} x_i$.*

Proof. By the previous theorem [theorem: 14.356] it follows that $\sup(\{\sum_{i=k}^{n} x_i|n\in\{k,\ldots,\infty\}\}$ exists and $\sup(\{\sum_{i=k}^{n} x_i|n\in\{k,\ldots,\infty\}\}=\sum_{i=k}^{n} x_i$. As $0<x_l\leqslant\sup(\{\sum_{i=k}^{n} x_i|n\in\{k,\ldots,\infty\}\}$ it follows that $0<\sum_{i=k}^{\infty} x_i$. □

We are now ready to prove commutativity of series of positive numbers.

Theorem 14.358. *Let $k\in\mathbb{N}_0$, $\{x_i\}_{i\in\{k,\ldots,\infty\}}\subseteq[0,\infty[$ be a sequence of non negative real numbers and $\beta:\{k,\ldots,\infty\}\to\{k,\ldots,\infty\}$ a bijection then we have*

$$\sum_{i=k}^{\infty} x_i \text{ converges} \Leftrightarrow \sum_{i=k}^{\infty} x_{\beta(i)} \text{ converges}$$

Further if $\sum_{i=k}^{\infty} x_i$ *converges then* $\sum_{i=k}^{\infty} x_i = \sum_{i=k}^{\infty} x_{\beta(i)}$

Proof. Let $s \in \{\sum_{i \in K} x_i | K \in \mathcal{P}(\{k, \ldots, \infty\}) \wedge K \text{ is finite}\}$ then there exist a finite $K \subseteq \{k, \ldots, \infty\}$ such that $s = \sum_{i \in K} x_i$. Take $L = \beta^{-1}(K)$ then as $\beta: \{k, \ldots, \infty\} \rightarrow \{k, \ldots, \infty\}$ is a bijection

$$\beta_{|L}: L \rightarrow \beta(L) = \beta(\beta^{-1}(K)) = K \text{ is a bijection}$$

Hence

$$s = \sum_{i \in K} x_i \underset{\text{[theorem: 11.36]}}{=} \sum_{i \in L} x_{\beta_{|L}(i)} = \sum_{i \in L} x_{\beta(i)}$$

As K is finite L must be finite [otherwise as β is a bijection $K = \beta(\beta^{-1}(K)) = \beta(L)$ is infinite], hence $s = \sum_{i \in L} x_{\beta(i)} \in \{\sum_{i \in K} x_{\beta(i)} | K \in \mathcal{P}(\{k, \ldots, \infty\}) \wedge K \text{ is finite}\}$ proving that

$$\left\{\sum_{i \in K} x_i | K \in \mathcal{P}(\{k, \ldots, \infty\}) \wedge K \text{ is finite}\right\} \subseteq \left\{\sum_{i \in K} x_{\beta(i)} | K \in \mathcal{P}(\{k, \ldots, \infty\}) \wedge K \text{ is finite}\right\} \tag{14.193}$$

Let $s \in \{\sum_{i \in K} x_{\beta(i)} | K \in \mathcal{P}(\{k, \ldots, \infty\}) \wedge K \text{ is finite}\}$ then there exist a finite $K \subseteq \{k, \ldots, \infty\}$ such that $s = \sum_{i \in K} x_{\beta(i)}$. As $\beta: \{k, \ldots, \infty\} \rightarrow \{k, \ldots, \infty\}$ we have that $\beta_{|K}: K \rightarrow \beta(K)$ is a bijection. Hence we have

$$s = \sum_{i \in K} x_{\beta(i)} = \sum_{i \in K} x_{\beta_{|K}(i)} \underset{\text{[theorem: 11.36]}}{=} \sum_{i \in \beta(K)} x_i$$

Which as $\beta(K)$ is finite [for K is finite and β is a bijection] means that $s = \sum_{i \in \beta(K)} x_i \in \{\sum_{i \in K} x_i | K \in \mathcal{P}(\{k, \ldots, \infty\}) \wedge K \text{ is finite}\}$. So that

$$\left\{\sum_{i \in K} x_{\beta(i)} | K \in \mathcal{P}(\{k, \ldots, \infty\}) \wedge K \text{ is finite}\right\} \subseteq \left\{\sum_{i \in K} x_i | K \in \mathcal{P}(\{k, \ldots, \infty\}) \wedge K \text{ is finite}\right\}$$

which combined with [eq; 14.193] prove that

$$\left\{\sum_{i \in K} x_i | K \in \mathcal{P}(\{k, \ldots, \infty\}) \wedge K \text{ is finite}\right\} = \left\{\sum_{i \in K} x_{\beta(i)} | K \in \mathcal{P}(\{k, \ldots, \infty\}) \wedge K \text{ is finite}\right\} \tag{14.194}$$

Now we have:

⇒. As $\sum_{i=k}^{\infty} x_i$ converges it follows from [theorem: 14.356] that

$$\sup\left(\left\{\sum_{i \in K} x_i | K \in \mathcal{P}(\{k, \ldots, \infty\}) \wedge K \text{ is finite}\right\}\right) \text{ exists}$$

so that by [eq: 14.194] $\sup(\{\sum_{i \in K} x_{\beta(i)} | K \in \mathcal{P}(\{k, \ldots, \infty\}) \wedge K \text{ is finite}\})$ exists, hence using [theorem: 14.356] again it follows that

$$\sum_{i=k}^{\infty} x_{\beta(i)} \text{ converges}$$

Further we have

$$\begin{aligned}
\sum_{i=k}^{\infty} x_i &\underset{\text{[theorem: 14.356]}}{=} \sup\left(\left\{\sum_{i \in K} x_i | K \in \mathcal{P}(\{k, \ldots, \infty\}) \wedge K \text{ is finite}\right\}\right) \\
&\underset{\text{[eq: 14.194]}}{=} \sup\left(\left\{\sum_{i \in K} x_{\beta(i)} | K \in \mathcal{P}(\{k, \ldots, \infty\}) \wedge K \text{ is finite}\right\}\right) \\
&\underset{\text{[theorem: 14.356]}}{=} \sum_{i=k}^{\infty} x_{\beta(i)}
\end{aligned}$$

⇐. As $\sum_{i=k}^{\infty} x_{\beta(i)}$ converges it follows from [theorem: 14.356] that

$$\sup\left(\left\{\sum_{i \in K} x_{\beta(i)} | K \in \mathcal{P}(\{k, \ldots, \infty\}) \wedge K \text{ is finite}\right\}\right) \text{ exists}$$

so that by [eq: 14.194] $\sup(\{\sum_{i \in K} x_i | K \in \mathcal{P}(\{k, \ldots, \infty\}) \wedge K \text{ is finite}\})$ exists, hence using [theorem: 14.356] again it follows that

$$\sum_{i=k}^{\infty} x_i \text{ converges}$$

□

Just as commutativity in finite sums allows us to define a more general form of a sum over elements of a finite set [see definition: 11.32] we can use the above theorem to define a infinite sum over a countable set of non negative numbers.

Theorem 14.359. *Let I be a denumerable set [see definition: 6.24], $\{x_i\}_{i\in I}\subseteq [0,\infty[$ a denumerable family of non negative numbers. If there exist a $\beta:\mathbb{N}_0\rightarrow I$ such that $\sum_{i=0}^{\infty} x_{\beta(i)}=x$ exist then for every bijection $\alpha:\mathbb{N}_0\rightarrow I$ we have that $\sum_{i=0}^{\infty} x_{\alpha(i)}=x$.*

Proof. As $\beta:\mathbb{N}_0\rightarrow I$ is a bijection we have that $\beta^{-1}:I\rightarrow\mathbb{N}_0$ is a bijection so that $\beta^{-1}\circ\alpha:\mathbb{N}_0\rightarrow\mathbb{N}_0$ is a bijection. Hence, as $\sum_{i=0}^{\infty} x_{\beta(i)}=x$, it follows from [theorem: 14.358] that $\sum_{i=0}^{\infty} x_{\beta((\beta^{-1}\circ\alpha)(x))}=x$. As $\forall i\in\mathbb{N}_0$ we have $x_{\beta((\beta^{-1}\circ\alpha)(i))}=x_{\beta(\beta^{-1}(\alpha(i)))}=x_{\alpha(i)}$ it follows that $\sum_{i=0}^{\infty} x_{\alpha(i)}=x$. □

The above theorem ensures that the following definition makes set

Definition 14.360. *Let I be a denumerable set [see definition: 6.24], $\{x_i\}_{i\in I}\subseteq[0,\infty[$ a denumerable family of non negative numbers then we say that*

$$\sum_{i\in I} x_i \textit{ converges} \text{ to } x \textit{ or } \sum_{i\in I} x_i = x$$

if there exists a bijection $\beta:\mathbb{N}_0\rightarrow[0,\infty[$ and a $x\in\mathbb{R}$ such that $\sum_{i=0}^{\infty} x_{\beta(i)}$ converges to x.

As a countable family is either a finite family or a denumerable family of non negative real number we can combine definitions [definition: 11.32] and [definition: 14.359] to define the sum of a countable family of non negative real numbers.

Definition 14.361. *Let I be a countable set [so by [definition: 6.25] I is finite or denumerable] and $\{x_i\}_{i\in I}\subseteq[0,\infty[$ then we say that $\sum_{i\in I} x_i$ converges if either*

$$I \textit{ is finite and then } \sum_{i\in I} x_i=\sum_{i=0}^{n} x_{\beta(i)} \textit{ where } \beta:\{0,\ldots,n\}\rightarrow I \textit{ is a bijection [see definition: 11.32]}$$

or

$$I \textit{ is denumerable and there exist a } x\in\mathbb{R} \textit{ such that } \sum_{i\in I} x_i=x \textit{ [see definition: 14.359]}$$

***Note* 14.362.** *In the finite case $\sum_{i\in I} x_i$ is guaranteed to be defined, however in the infinite case convergence is needed to ensure that $\sum_{i\in I} x_i$ exist.*

Example 14.363. If $\{x_i\}_{i\in\varnothing}\subseteq[0,\infty[$ then $\sum_{i\in\varnothing} x_i$ converges to 0

Proof. As $\varnothing$ is finite this follows from [definition: 11.32]. □

Theorem 14.364. *Let I,J be countable sets, $I\subseteq J$ and $\{x_i\}_{i\in J}\subseteq[0,\infty[$ such that $\sum_{i\in J} x_i$ converges then $\sum_{i\in I} x_i$ converges and $\sum_{i\in I} x_i\leqslant\sum_{i\in J} x_i$.*

Proof. For J we have either:

***J* is finite.** Then as $I\subseteq J$ I is also finite so that $\sum_{i\in I} x_i, \sum_{i\in J} x_i$ are defined by [theorem: 11.32] hence

$$\sum_{i\in I} x_i \leqslant \sum_{i\in I} x_i + \sum_{i\in J\setminus I} x_i \underset{\text{[theorem: 11.43]}}{=} \sum_{i\in J} x_i$$

***J* is denumerable.** By definition there exist a bijection $\beta:\mathbb{N}_0\rightarrow J$ such that $\sum_{i=0}^{\infty} x_{\beta(i)}$ converges to a real number x and $\sum_{i\in J} x_i = x=\sum_{i=0}^{\infty} x_{\beta(i)}$. For I we have now either:

***I* is finite.** As $\beta:\mathbb{N}_0\rightarrow J$ is a bijection and I is finite we have that $\beta^{-1}(I)\underset{\text{[theorem: 2.68]}}{=}(\beta^{-1})(I)$ is finite [see theorem: 6.44] so that $\beta_{|\beta^{-1}(I)}:\beta^{-1}(I)\rightarrow\beta(\beta^{-1}(I))=I$ is a bijection. Hence we have

$$\sum_{i\in I} x_i \underset{\text{[theorem: 11.36]}}{=} \sum_{i\in\beta^{-1}(I)} x_{\beta_{|\beta^{-1}(I)}(i)} = \sum_{i\in\beta^{-1}(I)} x_{\beta(i)}$$

As $\sum_{i\in I} x_i=\sum_{i\in\beta^{-1}(I)} x_{\beta(i)}\in\sup\left(\{\sum_{i\in K} x_{\beta(i)}|K\in\mathcal{P}(\mathbb{N}_0)\wedge K \text{ is finite}\}\right)$ we have that

$$\sum_{i\in I} x_i \leqslant \sup\left(\left\{\sum_{i\in K} x_{\beta(i)}|K\subseteq\mathcal{P}(\mathbb{N}_0)\wedge K \text{ is finite}\right\}\right) \underset{\text{[theorem: 14.356]}}{=} \sum_{i=0}^{\infty} x_{\beta(i)} = \sum_{i\in J} x_i$$

hence

$$\sum_{i\in I} x_i \leqslant \sum_{i\in J} x_i$$

***I* is denumerable.** Then there exist a a bijection $\alpha: \mathbb{N}_0 \rightarrow I \subseteq J$. Take a finite $K \subseteq \mathbb{N}_0$ then $\alpha(K)$ is finite and $\alpha(K) \subseteq J$ so that $\beta^{-1}(\alpha(K)) \underset{[\text{theorem: } 2.68]}{=} (\beta^{-1})(\alpha(K))$ is finite [see theorem: 6.44]. Define then the following bijections:

$$\alpha_{|K}: K \rightarrow \alpha(K) \text{ and } \beta_{|\beta^{-1}(\alpha(K))}: \beta^{-1}(\alpha(K)) \rightarrow \beta(\beta^{-1}(\alpha(K))) = \alpha(K)$$

Then we have

$$\begin{aligned}
\sum_{i\in K} x_{\alpha(i)} & \underset{[\text{theorem: } 11.36]}{=} \sum_{i\in\alpha(K)} x_i \\
& \underset{[\text{theorem: } 11.36]}{=} \sum_{i\in\beta^{-1}(\alpha(K))} x_{\beta(i)} \\
& \leqslant \sup\left(\left\{\sum_{i\in L} x_{\beta(i)} | L \in \mathcal{P}(\mathbb{N}_0) \wedge L \text{ is finite}\right\}\right) \\
& \underset{[\text{theorem: } 14.356]}{=} \sum_{i=0}^{\infty} x_{\beta(i)} \\
& = \sum_{i\in J} x_i
\end{aligned}$$

Hence $\{\sum_{i\in K} x_{\alpha(i)} | K \in \mathcal{P}(\mathbb{N}_0) \wedge K \text{ is finite}\}$ is bounded above by $\sum_{i\in J} x_i$ and trivially non empty. So by [theorem: 10.18] $\sup(\{\sum_{i\in K} x_{\alpha(i)} | K \in \mathcal{P}(\mathbb{N}_0) \wedge K \text{ is finite}\})$ exist and $\sup(\{\sum_{i\in K} x_{\alpha(i)} | K \in \mathcal{P}(\mathbb{N}_0) \wedge K \text{ is finite}\}) \leqslant \sum_{i\in J} x_i$. So using [theorem: 14.356] it follows that $\sum_{i=0}^{\infty} x_{\alpha(i)}$ exist and

$$\sum_{i=0}^{\infty} x_{\alpha(i)} = \sup\left(\left\{\sum_{i\in K} x_{\alpha(i)} | K \in \mathcal{P}(\mathbb{N}_0) \wedge K \text{ is finite}\right\}\right) \leqslant \sum_{i\in J} x_i$$

Finally using [definition: 14.359] we have that $\sum_{i\in I} x_i = \sum_{i=0}^{\infty} x_{\alpha(i)}$ so that

$$\sum_{i\in I} x_i \leqslant \sum_{i\in J} x_i$$

□

For countable families we have also the following equivalent definition of convergence.

Theorem 14.365. *Let I be a countable set and $\{x_i\}_{i\in I} \subseteq [0,\infty[$ a countable family of non negative real numbers then*

$$\sum_{i\in I} x_i \textit{ converges} \Leftrightarrow \sup\left(\left\{\sum_{i\in K} x_i | K \in \mathcal{P}(I) \wedge K \textit{ is finite}\right\}\right) \textit{ exist}$$

Further if $\sum_{i\in I} x_i$ converges then

$$\sum_{i\in I} x_i = \sup\left(\left\{\sum_{i\in K} x_i | K \in \mathcal{P}(I) \wedge K \textit{ is finite}\right\}\right)$$

Proof. As I is countable we have two possibilities to consider:

***I* is finite.** Then for a finite $K \subseteq I$ we have by [theorem: 14.364] that $\sum_{i\in K} x_i \leqslant \sum_{i\in I} x_i$ so that $\{\sum_{i\in K} x_i | K \in \mathcal{P}(I) \wedge K \text{ is finite}\}$ is bounded above by $\sum_{i\in I} x_i$. Hence by [theorem: 10.18] $\sup(\{\sum_{i\in K} x_i | K \in \mathcal{P}(I) \wedge K \text{ is finite}\})$ exist and

$$\sup\left(\left\{\sum_{i\in K} x_i | K \in \mathcal{P}(I) \wedge K \text{ is finite}\right\}\right) \leqslant \sum_{i\in I} x_i \tag{14.195}$$

As $\sum_{i\in I} x_i$ converges by definition if I is finite we have

$$\sum_{i\in I} x_i \text{ converges} \Leftrightarrow \sup\left(\left\{\sum_{i\in K} x_i | K \in \mathcal{P}(I) \wedge K \text{ is finite}\right\}\right) \text{ exist}$$

As I is finite we have clearly $\sum_{i \in I} x_i \in \{\sum_{i \in K} x_i | K \in \mathcal{P}(I) \wedge K \text{ is finite}\}$ so that

$$\sum_{i \in I} x_i \leqslant \sup\left(\left\{\sum_{i \in K} x_i | K \in \mathcal{P}(I) \wedge K \text{ is finite}\right\}\right)$$

which combined with [eq: 14.195] gives

$$\sum_{i \in I} x_i = \sup\left(\left\{\sum_{i \in K} x_i | K \in \mathcal{P}(I) \wedge K \text{ is finite}\right\}\right)$$

***I* is denumerable.** Then there exist a bijection $\beta: \mathbb{N}_0 \rightarrow I$. Let $s \in \{\sum_{i \in K} x_{\beta(i)} | K \in \mathcal{P}(\mathbb{N}_0) \wedge K \text{ is finite}\}$ then there exist a finite $K \subseteq \mathbb{N}_0$ such that $s = \sum_{i \in K} x_{\beta(i)}$. As $\beta_{|K}: K \rightarrow \beta(K)$ is a bijection we have

$$s = \sum_{i \in K} x_{\beta(i)} = \sum_{i \in K} x_{\beta_{|K}(i)} \underset{[\text{theorem: } 11.36]}{=} \sum_{i \in \beta(K)} x_i$$

which as $\beta(K)$ is finite [see theorem: 6.44] and $\beta(K) \subseteq I$ proves that $s \in \{\sum_{i \in K} x_i | K \in \mathcal{P}(I) \wedge K \text{ is finite}\}$ hence

$$\left\{\sum_{i \in K} x_{\beta(i)} | K \in \mathcal{P}(\mathbb{N}_0) \wedge K \text{ is finite}\right\} \subseteq \left\{\sum_{i \in K} x_i | K \in \mathcal{P}(I) \wedge K \text{ is finite}\right\} \tag{14.196}$$

Let $s \in \{\sum_{i \in K} x_i | K \in \mathcal{P}(I) \wedge K \text{ is finite}\}$ then there exist a finite $K \subseteq I$ such that $s = \sum_{i \in K} x_i$. As $\beta_{|\beta^{-1}(K)}: \beta^{-1}(K) \rightarrow \beta(\beta^{-1}(K)) = K$ is a bijection we have

$$s = \sum_{i \in K} x_i \underset{[\text{theorem: } 11.36]}{=} \sum_{i \in \beta^{-1}(K)} x_{\beta_{|\beta^{-1}(K)}(i)} = \sum_{i \in \beta^{-1}(K)} x_{\beta(i)}$$

which as $\beta^{-1}(K) \underset{[\text{theorem: } 2.68]}{=} (\beta^{-1})(K)$ is finite by [theorem: 6.44] and $\beta^{-1}(K) \subseteq \mathbb{N}_0$ proves that $s \in \{\sum_{i \in K} x_{\beta(i)} | K \in \mathcal{P}(\mathbb{N}_0) \wedge K$ is finite$\}$. Hence

$$\left\{\sum_{i \in K} x_i | K \in \mathcal{P}(I) \wedge K \text{ is finite}\right\} \subseteq \left\{\sum_{i \in K} x_{\beta(i)} | K \in \mathcal{P}(\mathbb{N}_0) \wedge K \text{ is finite}\right\}$$

which combined with [eq: 14.196] proves that

$$\left\{\sum_{i \in K} x_{\beta(i)} | K \in \mathcal{P}(\mathbb{N}_0) \wedge K \text{ is finite}\right\} = \left\{\sum_{i \in K} x_i | K \in \mathcal{P}(I) \wedge K \text{ is finite}\right\} \tag{14.197}$$

Finally we have

$$\begin{aligned}
\sum_{i \in I} x_i \text{ converges to } x & \underset{[\text{definition: } 14.359]}{\Leftrightarrow} \\
\sum_{i=0}^{\infty} x_{\beta(i)} \text{ converges to } x & \underset{[\text{theorem: } 14.356]}{\Leftrightarrow} \\
\sup\left(\left\{\sum_{i \in K} x_{\beta(i)} | K \in \mathcal{P}(\mathbb{N}_0) \wedge K \text{ is finite}\right\}\right) \text{ exists and is equal to } x & \underset{[\text{eq: } 14.197]}{\Leftrightarrow} \\
\sup\left(\left\{\sum_{i \in K} x_i | K \in \mathcal{P}(I) \wedge K \text{ is finite}\right\}\right) \text{ exists and is equal to } x &
\end{aligned}$$

□

A trivial example of the use of the above theorem is the following

Example 14.366. Let I be a countable set and $\{x_i\}_{i \in I} \subseteq [0, \infty[$ such that $\forall i \in I$ $x_i = 0$ 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 $\sum_{i \in K} x_i = 0$ so that $\sup(\{\sum_{i \in K} x_i | K \subseteq I \wedge K \text{ is finite}\}) = 0$ which by the previous theorem proves the example. □

Convergent series of non negative numbers are associative, to prove this we need first a little lemma.

Lemma 14.367. *Let $n \in \mathbb{N}_0$ and $\{A_i\}_{i \in \{0, \ldots, n\}} \subseteq \mathcal{P}(\mathbb{R}) \setminus \{\varnothing\}$ a family of non empty sets of real numbers then if there exist a $s \in \mathbb{R}$ such that $\forall (x_0, \ldots, x_n) \in \prod_{i \in \{0, \ldots, n\}} A_i$ $\sum_{i=0}^n x_i \leqslant s$ it follows that $\forall i \in \{0, \ldots, n\}$ $\sup(A_i)$ exist and $\sum_{i=0}^n \sup(A_i) \leqslant s$*

Proof. We prove this by induction so let

$$S = \left\{ n \in \mathbb{N}_0 | \text{If } \{A_i\}_{i \in \{0,\ldots,n\}} \subseteq \mathcal{P}(\mathbb{R}) \setminus \{\varnothing\} \text{ satisfies that for a } s \in \mathbb{R} \text{ with } \forall x \in \prod_{i \in \{0,\ldots,n\}} A_i \text{ we have } \sum_{i=0}^{n} x_i \leqslant s \text{ then we have that } \forall i \in \{0,\ldots,n\} \sup(A_i) \text{ exist and } \sum_{i=0}^{n} \sup(A_i) \leqslant s \right\}$$

then we have:

$0 \in S$. Let $y \in A_0$ then $y = (x) \in \prod_{i \in \{0\}} A_i$ so that $x = \sum_{i=0}^{0} y_0 \leqslant s$ proving that A_0 is bounded above by s, hence, as also $A_0 \neq \varnothing$, it follows from the conditional completeness of $\mathbb{R}$ [see theorem: 10.18] that $\sup(A_0)$ exists and $\sup(A_0) \leqslant s \Rightarrow \sum_{i=0}^{0} \sup(A_i) = \sup(A_0) \leqslant s$. Hence $0 \in S$.

$n \in A \Rightarrow n+1 \in S$. Let $\{A_i\}_{i \in \{0,\ldots,n+1\}} \subseteq \mathcal{P}(\mathbb{R}) \setminus \{\varnothing\}$ and let $s \in \mathbb{R}$ be such that $\forall x \in \prod_{i \in \{0,\ldots,n+1\}} A_i$ we have $\sum_{i=0}^{n+1} x_i \leqslant s$. Let $u \in A_{n+1}$ then we have $\forall x \in \prod_{i \in \{0,\ldots,n\}} A_i$ that $y = (x_0, \ldots, x_n, u) \in \prod_{i \in \{0,\ldots,n+1\}} A_i$ so that $u + \sum_{i=0}^{n} x_i = \sum_{i=0}^{n+1} y_i \leqslant s$, hence

$$\forall x \in \prod_{i \in \{0,\ldots,n\}} A_i \text{ we have } \sum_{i=0}^{n} x_i \leqslant s - u$$

So as $n \in S$ we have that

$$\forall i \in \{0,\ldots,n\} \sup(A_i) \text{ exist and } \sum_{i=0}^{n} \sup(A_i) \leqslant s - u \tag{14.198}$$

From the above it follows that for $u \in A_{n+1}$ $u \leqslant s - \sum_{i=0}^{n} \sup(A)$ proving that A_{n+1} is bounded above by $s - \sum_{i=0}^{n} \sup(A_i)$, as further $\varnothing \neq A_{n+1}$ it follows form [theorem: 10.18] that

$$\sup(A_{n+1}) \text{ exist and } \sup(A_{n+1}) \leqslant s - \sum_{i=0}^{n} \sup(A_i) \tag{14.199}$$

So by [eqs: 14.198, 14.199] we have

$$\forall i \in \{0,\ldots,n+1\} \sup(A_i) \text{ exist and } \sum_{i=0}^{n+1} \sup(A_i) = \sup(A_{n+1}) + \sum_{i=0}^{n} \sup(A_i) \leqslant s$$

proving that $n+1 \in S$. □

Lemma 14.368. *Let I be a countable set, $\{x_i\}_{i \in I} \subseteq [0, \infty[$ such that $\sum_{i \in I} x_i$ converges, $n \in \mathbb{N}_0$ and $\{N_i\}_{i \in \{0,\ldots,n\}} \subseteq \mathcal{P}(I)$ such that $\forall i, j \in \{0,\ldots,n\}$ with $i \neq j$ we have $N_i \bigcap N_j = \varnothing$ and $I = \bigcup_{i \in \{0,\ldots,n\}} N_i$ then $\forall i \in \{0,\ldots,n\}$ we have that*

$$\sum_{j \in N_i} x_j \textit{ converges}$$

and

$$\sum_{i \in I} x_i = \sum_{i=0}^{n} \left(\sum_{j \in N_i} x_j \right)$$

Proof. As $\forall i \in \{0,\ldots,n\}$ $N_i \subseteq I$ and $\sum_{i \in I} x_i$ converges it follows from [theorem: 14.364] that

$$\forall i \in \{0,\ldots,n\} \sum_{j \in N_i} x_j \text{ converges} \tag{14.200}$$

As $\sum_{i\in I} x_i$ converges it follows from [theorem: 14.365] that $\sup\left(\left\{\sum_{i\in K} x_i | K\subseteq I\wedge K \text{ finite}\right\}\right)$ exist and

$$\sum_{i\in I} x_i = \sup\left(\left\{\sum_{i\in K} x_i | K\subseteq I\wedge K \text{ finite}\right\}\right) \tag{14.201}$$

Let $\varepsilon\in\mathbb{R}^+$ then, as $\sum_{i\in I} x_i - \varepsilon < \sup\left(\left\{\sum_{i\in K} x_i | K\subseteq I\wedge K \text{ finite}\right\}\right)$, there exist a finite $K\subseteq I$ such that $\sum_{i\in I} x_i - \varepsilon < \sum_{i\in K} x_i$ or

$$\sum_{i\in I} x_i < \sum_{i\in K} x_i + \varepsilon \tag{14.202}$$

Now given $i\in\{0,\ldots,n\}$ we have that $N_i\bigcap K$ is finite and as $N_i\bigcap K\subseteq N_i$ it follows from [theorem: 14.364] that

$$\sum_{j\in N_i\bigcap K} x_j \leqslant \sum_{j\in N_i} x_j \tag{14.203}$$

Further as $K\subseteq I$

$$K = I\bigcap K = \left(\bigcup_{i\in\{0,\ldots,n\}} N_i\right)\bigcap K = \bigcup_{i\in\{0,\ldots,n\}} (N_i\bigcap K)$$

and $\forall i,j\in\{0,\ldots,n\}$ with $i\neq j$ we have

$$(N_i\bigcap K)\bigcap (N_j\bigcap K) = (N_i\bigcap N_j)\bigcap K = \varnothing\bigcap K = \varnothing$$

so using [theorem: 11.43] we have

$$\begin{aligned}\sum_{i\in K} x_i &\underset{\text{[theorem: 11.43]}}{=} \sum_{i=0}^{n}\left(\sum_{j\in N_i\bigcap K} x_j\right)\\ &\leqslant_{\text{[eq: 14.203]}} \sum_{i=0}^{n}\left(\sum_{j\in N_i} x_j\right)\end{aligned}$$

Combining the above with [eq: 14.202] results in $\sum_{i\in I} x_i < \sum_{i=0}^{n}\left(\sum_{j\in N_i} x_j\right) + \varepsilon$ which as $\varepsilon\in\mathbb{R}^+$ was chosen arbitrary allows us to use [theorem: 10.31] giving

$$\sum_{i\in I} x_i \leqslant \sum_{i=0}^{n}\left(\sum_{j\in N_i} x_j\right) \tag{14.204}$$

Take now $s\in\prod_{i\in\{0,\ldots,n\}}\left\{\sum_{j\in K} x_j | K\subseteq N_i\wedge K \text{ finite}\right\}$ then there exist a family of finite sets $\{K_i\}_{i\in\{0,\ldots,n\}}$ with $\forall i\in\{0,\ldots,n\}$ $K_i\subseteq N_i$ [so the sets are finite and pairwise disjoint] and $s_i = \sum_{j\in K} x_j$. Hence

$$\sum_{i=0}^{n} s_i = \sum_{i=0}^{n}\left(\sum_{j\in K_i} x_j\right) \underset{\text{[theorem: 11.43]}}{=} \sum_{j\in\bigcup_{i\in\{0,\ldots,n\}} K_i} x_i \tag{14.205}$$

As $\bigcup_{i\in\{0,\ldots,n\}} K_i \subseteq \bigcup_{i\in\{0,\ldots,n\}} N_i = I$ we have by [theorem: 14.364] as $\sum_{i\in I} x_i$ is convergent that $\sum_{j\in\bigcup_{i\in\{0,\ldots,n\}} K_i} x_i \leqslant \sum_{i\in I} x_i$, hence by [eq: 14.205]

$$\sum_{i=0}^{n} s_i \leqslant \sum_{i\in I} x_i$$

So we can use the previous lemma [lemma: 14.367] proving that

$$\forall i\in\{0,\ldots,n\} \sup\left(\sum_{j\in K} x_j | K\subseteq N_i\wedge K \text{ finite}\right) \text{ exist and } \sum_{i=0}^{n} \sup\left(\sum_{j\in K} x_j | K\subseteq N_i\wedge K \text{ finite}\right) \leqslant \sum_{i\in I} x_i$$

Using [theorem: 14.365] we have $\forall i\in\{0,\ldots,n\}$ $\sum_{j\in N_i} x_j = \sup\left(\sum_{j\in K} x_j | K\subseteq N_i\wedge K \text{ finite}\right)$ which combined with the above proves that

$$\sum_{i=0}^{n}\left(\sum_{j\in N_i} x_j\right) \leqslant \sum_{i\in I} x_i$$

which combined with [eq: 14.204] proves

$$\sum_{i\in I} x_i = \sum_{i=0}^{n}\left(\sum_{j\in N_i} x_j\right)$$

□

Lemma 14.369. *Let I be a countable set, $\{x_i\}_{i\in I}\subseteq[0,\infty[$ such that $\sum_{i\in I}x_i$ converges, $\{N_i\}_{i\in\mathbb{N}_0}\subseteq\mathcal{P}(I)$ be such that $\forall i,j\in\mathbb{N}_0$ with $i\neq j$ $N_i\bigcap N_j=\varnothing$ and $I=\bigcup_{i\in\mathbb{N}_0}N_i$ then we have that*

$$\forall i\in\mathbb{N}_0 \sum_{j\in N_i} x_j \text{ converges}$$

and

$$\sum_{i=0}^{\infty}\left(\sum_{j\in N_i}x_j\right) \text{ converges and } \sum_{i=0}^{\infty}\left(\sum_{j\in N_i}x_j\right)=\sum_{i\in I}x_i$$

Proof. First as $\forall i\in\mathbb{N}_0$ $N_i\subseteq I$ we have by [theorem: 14.364] that

$$\forall i\in\mathbb{N}_0 \sum_{j\in N_i}x_j \text{ converges} \tag{14.206}$$

Let $n\in\mathbb{N}_0$ and define $K_n=\bigcup_{i\in\{0,\ldots,n\}}N_i\subseteq I$ then by [theorem: 14.364] again we have that

$$\forall n\in\mathbb{N}_0 \sum_{i\in K_n}x_i \text{ converges and } \sum_{i\in K_n}x_i\leqslant\sum_{i\in I}x_i \tag{14.207}$$

Further as I is countable it follows by [theorem: 6.68] that $\forall n\in\mathbb{N}_0$ K_n is countable and as $\{N_i\}_{i\in\{0,\ldots,n\}}$ is pairwise disjoint it follows from the previous lemma [lemma: 14.368] that

$$\sum_{i=0}^{n}\left(\sum_{j\in N_i}x_j\right)\underset{[\text{lemma: }14.368]}{=}\sum_{i\in K_n}x_i\leqslant_{[\text{eq: }14.207]}\sum_{i\in I}x_i$$

so that $\{\sum_{i=0}^{n}(\sum_{j\in N_i}x_j)|n\in\mathbb{N}_0\}$ is bounded above by $\sum_{i\in I}x_i$, hence we have by [theorem 10.18] that $\sup(\{\sum_{i=0}^{n}(\sum_{j\in N_i}x_j)|n\in\mathbb{N}_0\})$ exist and

$$\sup\left(\left\{\sum_{i=0}^{n}\left(\sum_{j\in N_i}x_j\right)|n\in\mathbb{N}_0\right\}\right)\leqslant\sum_{i\in I}x_i$$

Using then [theorem: 14.356] we have that

$$\sum_{i=0}^{\infty}\left(\sum_{j\in N_i}x_j\right) \text{ converges and } \sum_{i=0}^{\infty}\left(\sum_{j\in N_i}x_j\right)\leqslant\sum_{i\in I}x_i \tag{14.208}$$

For the opposite equality, let $\varepsilon\in\mathbb{R}^+$ then as $\sum_{i\in I}x_i-\varepsilon<\sum_{i\in I}x_i$ and

$$\sum_{i\in I}x_i\underset{[\text{theorem: }14.365]}{=}\sup\left(\left\{\sum_{i\in K}x_i|K\subseteq I\wedge K\text{ is finite}\right\}\right)$$

there exist a finite $K\subseteq I$ such that $\sum_{i\in I}x_i-\varepsilon<\sum_{i\in K}x_i$ or

$$\sum_{i\in I}x_i<\sum_{i\in K}x_i+\varepsilon \tag{14.209}$$

Take now $J=\{i\in\mathbb{N}_0|K\bigcap N_i\neq\varnothing\}\subseteq\mathbb{N}_0$ so that by [theorem: 6.68] J is countable. If $x\in K$ then as $K=K\bigcap I=K\bigcap(\bigcup_{i\in\mathbb{N}_0}N_i)=\bigcup_{i\in\mathbb{N}_0}(K\bigcap N_i)$ there exist a $i\in\mathbb{N}_0$ such that $x\in K\bigcap N_i\Rightarrow K\bigcap N_i\neq\varnothing$ hence $i\in J$ proving that $x\in\bigcup_{i\in J}(K\bigcap N_i)$. Hence $K\subseteq\bigcup_{i\in J}(K\bigcap N_i)$ which as trivially $\bigcup_{i\in J}(K\bigcap N_i)\subseteq K$ proves that

$$K=\bigcup_{i\in J}\left(K\bigcap N_i\right) \tag{14.210}$$

We have now two cases to consider for J:

$J=\varnothing$. Then by [eq: 14.210] $K=\varnothing$ hence $\sum_{i\in K}x_i=0$ so that by [eq: 14.209] $\sum_{i\in I}x_i<\varepsilon$ which by [theorem: 10.31] proves that

$$\sum_{i\in I}x_i\leqslant 0\leqslant\sum_{i=0}^{\infty}\left(\sum_{j\in N_i}x_j\right)$$

$J\neq\varnothing$. Assume that J is not finite then as $\{K\bigcap N_i\}_{i\in J}$ is a family of pairwise disjoint non empty sets if follows from [theorem: 6.47] that $K=\bigcup_{i\in J}(K\bigcap N_i)$ is infinite, contradicting the fact that K was chosen to be finite. Hence the assumption is wrong and J is finite and non empty so that by [theorem: 6.48]

$$m=\max(J)\text{ exist}$$

Let $x\in\bigcup_{i\in\{0,\ldots,m\}}(K\bigcap N_i)$ then $\exists i\in\{0,\ldots,m\}$ such that $x\in K\bigcap N_i\Rightarrow K\bigcap N_i\neq\varnothing$ so that $i\in J$ proving that $x\in\bigcup_{i\in J}(K\bigcap N_i)\underset{\text{[eq: 14.93]}}{=}K$ proving that

$$\bigcup_{i\in\{0,\ldots,m\}}(K\bigcap N_i)\subseteq K \tag{14.211}$$

On the other hand if $x\in K$ then by [eq: 14.210] there exist a $i\in J$ such that $x\in K\bigcap N_i$ which as $J\subseteq\{0,\ldots,\max(J)\}=\{0,\ldots,m\}$ proves that $x\in\bigcup_{i\in\{0,\ldots,m\}}(K\bigcap N_i)$ or $K\subseteq\bigcup_{i\in\{0,\ldots,m\}}(K\bigcap N_i)$, combining this with [eq: 14.211] gives

$$K=\bigcup_{i\in\{0,\ldots,m\}}(K\bigcap N_i)$$

As $\{K\bigcap N_i\}_{i\in J}$ is pairwise disjoint we have by the above and [theorem: 11.43] that

$$\begin{aligned}
\sum_{i\in K}x_i &\underset{\text{[theorem: 11.43]}}{=} \sum_{i=0}^{m}\left(\sum_{j\in K\bigcap N_i}x_j\right)\\
&\leqslant_{\text{[theorem: 14.364]}} \sum_{i=0}^{m}\left(\sum_{j\in N_i}x_j\right)\\
&\leqslant \sup\left(\left\{\sum_{i=0}^{m}\left(\sum_{j\in N_i}x_j\right)\middle|m\in\mathbb{N}_o\right\}\right)\\
&\underset{\text{[theorem: 14.356}}{=} \sum_{i=0}^{\infty}\left(\sum_{j\in N_i}x_j\right)
\end{aligned}$$

So we have

$$\sum_{i\in I}x_i<_{\text{[eq: 14.209]}}\sum_{i\in K}x_i+\varepsilon\leqslant\sum_{i=0}^{\infty}\left(\sum_{j\in N_i}x_j\right)+\varepsilon$$

so that by [theorem: 10.31] we have

$$\sum_{i\in I}x_i\leqslant\sum_{i=0}^{\infty}\left(\sum_{j\in N_i}x_j\right)$$

So in all cases we have $\sum_{i\in I}x_i\leqslant\sum_{i=0}^{\infty}(\sum_{j\in N_i}x_j)$ which combined with [eq: 14.208] it follows that

$$\sum_{i\in I}x_i=\sum_{i=0}^{\infty}\left(\sum_{j\in N_i}x_j\right)$$

□

Finally we are ready to prove the general case of countable families.

Theorem 14.370. *Let I,K be countable sets, $I\neq\varnothing$, $\{K_i\}_{i\in I}$ a countable family of countable sets such that $K=\bigcup_{i\in I}K_i$ and $\forall i,j\in I$ with $i\neq j$ $K_i\bigcap K_j=\varnothing$ and $\{x_i\}_{i\in K}\subseteq[0,\infty[$ a countable family of non negative numbers such that $\sum_{i\in K}x_i$ converges then*

$$\forall i\in I\ \sum_{j\in K_i}x_j\text{ converges and }\sum_{i\in K}x_i=\sum_{i\in I}\left(\sum_{j\in K_i}x_j\right)$$

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 $\beta:\{0,\ldots,n\}\rightarrow I$. Define then

$$\{L_i\}_{i\in\{0,\ldots,n\}}\text{ by }L_i=K_{\beta(i)}$$

then we have

$$K=\bigcup_{i\in I}K_i\underset{\text{[theorem: 2.117]}}{=}\bigcup_{i\in\{0,\ldots,n\}}K_{\beta(i)}=\bigcup_{i\in\{0,\ldots,n\}}L_i$$

Further if $i,j\in\{0,\ldots,n\}$ with $i\neq j$ then as β is a bijection $\beta(i)\neq\beta(j)$ so that $L_i\bigcap L_j=K_{\beta(i)}\bigcap K_{\beta(j)}=\varnothing$. So we can use [lemma: 14.368] giving

$$\forall i\in\{0,\ldots,n\}\ \sum_{j\in L_i}x_j\text{ converges and }\sum_{i=0}^{n}\left(\sum_{j\in L_i}x_j\right)=\sum_{i\in K}x_i \tag{14.212}$$

Finally $\forall i\in I$ we have that $K_i = K_{\beta(\beta^{-1}(i))} = L_{\beta^{-1}(i)}$ so that

$$\sum_{j\in K_i} x_j = \sum_{j\in L_{\beta^{-1}(i)}} x_j \text{ converges}$$

and

$$\begin{aligned}\sum_{i\in K} x_i &\underset{[\text{eq: } 14.212]}{=} \sum_{i=0}^{n}\left(\sum_{j\in L_i} x_j\right)\\ &= \sum_{i=0}^{n}\left(\sum_{j\in K_{\beta(i)}} x_j\right)\\ &\underset{[\text{definition: } 11.32]}{=} \sum_{i\in I}\left(\sum_{j\in K_i} x_j\right)\end{aligned}$$

***I* is denumerable.** Then there exist a bijection $\beta\colon \mathbb{N}_0 \rightarrow I$. Define then

$$\{L_i\}_{i\in\mathbb{N}_0} \text{ by } L_i = K_{\beta(i)}$$

then we have

$$K = \bigcup_{i\in I} K_i \underset{[\text{theorem: } 2.117]}{=} \bigcup_{i\in\mathbb{N}_0} K_{\beta(i)} = \bigcup_{i\in\mathbb{N}_0} L_i$$

Further if $i, j\in\mathbb{N}_0$ with $i\neq j$ then as β is a bijection $\beta(i)\neq\beta(j)$ so that $L_i\bigcap L_j = K_{\beta(i)}\bigcap K_{\beta(j)} = \varnothing$. So we can use [lemma: 14.369] giving

$$\forall i\in\mathbb{N}_0 \sum_{j\in L_i} x_j \text{ converges and } \sum_{i=0}^{\infty}\left(\sum_{j\in L_i} x_j\right) = \sum_{i\in K} x_i \tag{14.213}$$

Finally $\forall i\in I$ we have that $K_i = K_{\beta(\beta^{-1}(i))} = L_{\beta^{-1}(i)}$ so that

$$\sum_{j\in K_i} x_j = \sum_{j\in L_{\beta^{-1}(i)}} x_j \text{ converges}$$

and

$$\begin{aligned}\sum_{i\in K} x_i &\underset{[\text{eq: } 14.213]}{=} \sum_{i=0}^{\infty}\left(\sum_{j\in L_j} x_j\right)\\ &= \sum_{i=0}^{\infty}\left(\sum_{j\in K_{\beta(i)}} x_j\right)\\ &\underset{[\text{definition: } 14.360]}{=} \sum_{i\in I}\left(\sum_{j\in K_i} x_j\right)\end{aligned}$$

□

14.10.3.2 Absolute Convergent Series

Definition 14.371. (Absolute Convergence) *Let $\langle X, \|\|\rangle$ be a normed space and $\{x_i\}_{i\in\{k,\ldots,\infty\}}\subseteq X$ a sequence then $\sum_{i=k}^{\infty} x_i$ is **absolute convergent** if $\sum_{i=k}^{\infty}\|x_i\|$ is convergent.*

The following theorem shows that in a Banach space every series where the norm of its terms is dominated by the terms of a convergent series is convergent.

Theorem 14.372. (Dominant Convergence) *Let $k\in\mathbb{N}_0$, $\langle X, \|\|\rangle$ be a Banach space, $\{s_i\}_{i\in\{k,\ldots,\infty\}}\subseteq [0,\infty[$ such that $\sum_{i=k}^{\infty} s_i$ is convergent and $\{x_i\}_{i\in\{k,\ldots,\infty\}}\subseteq X$ satisfies $\forall i\in\{k,\ldots\infty\}$ $\|x_i\|\leqslant s_i$ then*

$$\sum_{i=k}^{\infty} x_i \text{ converges}$$

and

$$\left\|\sum_{i=k}^{\infty} x_i\right\| \leqslant \sum_{i=k}^{\infty} s_i$$

Proof. Let $\varepsilon \in \mathbb{R}^+$ then as $\sum_{i=k}^{\infty} s_i$ converges we have that $\{\sum_{i=k}^{n} s_i\}_{n \in \{k,\ldots,\infty\}}$ convergence, hence by [theorem: 14.322] there exist $N \in \{k, \ldots, \infty\}$ such that $\forall n, m \geqslant N$ we have

$$\left|\sum_{i=k}^{n} s_i - \sum_{i=k}^{m} s_i\right| < \varepsilon$$

For n, m we have the following possible cases to consider:

$\boldsymbol{n = m.}$ Then $\|\sum_{i=k}^{n} x_i - \sum_{i=k}^{m} x_i\| = \|\sum_{i=k}^{n} x_i - \sum_{i=k}^{n} x_i\| = \|0\| = 0 < \varepsilon$

$\boldsymbol{n < m.}$ Then

$$\begin{array}{rcl} \left\|\sum_{i=k}^{n} x_i - \sum_{i=k}^{m} x_i\right\| & = & \left\|\sum_{i=k}^{m} x_i - \sum_{i=k}^{n} x_i\right\| \\ & = & \left\|\sum_{i=n+1}^{m} x_i\right\| \\ & \leqslant_{[\text{theorem: } 14.86]} & \sum_{i=n+1}^{m} \|x_i\| \\ & \leqslant & \sum_{i=n+1}^{m} s_i \\ & = & \left|\sum_{i=n+1}^{m} s\right| \\ & = & \left|\sum_{i=k}^{m} s_i - \sum_{i=k}^{n} s_i\right| \\ & < & \varepsilon \end{array}$$

$\boldsymbol{m < n.}$

$$\begin{array}{rcl} \left\|\sum_{i=k}^{n} x_i - \sum_{i=k}^{m} x_i\right\| & = & \left\|\sum_{i=n+1}^{m} x_i\right\| \\ & \leqslant_{[\text{theorem: } 14.86]} & \sum_{i=n+1}^{m} \|x_i\| \\ & \leqslant & \sum_{i=n+1}^{m} s_i \\ & = & \left|\sum_{i=n+1}^{m} s\right| \\ & = & \left|\sum_{i=k}^{m} s_i - \sum_{i=k}^{n} s_i\right| \\ & < & \varepsilon \end{array}$$

Hence in all cases $\|\sum_{i=k}^{n} x_i - \sum_{i=k}^{m} x_i\| < \varepsilon$ proving that $\{\sum_{i=k}^{n} x_i\}_{n \in \{k,\ldots,\infty\}}$ is Cauchy. As $\langle X, \|\|\rangle$ is a Banach space $\lim_{n \rightarrow \infty} x_n$ exist so that by definition

$$\sum_{i=k}^{\infty} x_i \text{ converges}$$

Further if $n \in \{k, \ldots, \infty\}$ then we have

$$\left\|\sum_{i=k}^{n} x_i\right\| \leqslant \sum_{i=k}^{n} \|x_i\| \leqslant \sum_{i=k}^{n} s_i$$

Using [theorem: 14.304] it follows that

$$\lim_{n \rightarrow \infty}\left\|\sum_{i=k}^{n} x_i\right\| \leqslant \lim_{n \rightarrow \infty} \sum_{i=k}^{n} s_n = \sum_{i=k}^{\infty} s_i$$

As by [theorem: 14.310] $\lim_{n\rightarrow\infty} \|\sum_{i=k}^{n} x_i\| = \left\|\lim_{n\rightarrow\infty} \sum_{i=k}^{n} x_i\right\| \underset{\text{def}}{=} \|\sum_{i=k}^{\infty} x_i\|$ it follows that

$$\left\|\sum_{i=k}^{\infty} x_i\right\| \leqslant \sum_{i=k}^{\infty} s_i$$

□

We have a variant of the dominant convergence theorem where the inequality is strict, this variant will be used in the proof of the open mapping theorem.

Corollary 14.373. *Let $k \in \mathbb{N}_0$, $\langle X, \|\|\rangle$ be a Banach space, $\{s_i\}_{i\in\{k,\ldots,\infty\}} \subseteq [0,\infty[$ such that $\sum_{i=k}^{\infty} s_i$ is convergent and $\{x_i\}_{i\in\{k,\ldots,\infty\}} \subseteq X$ a sequence in X such that:*

1. $\forall i \in \{k, \ldots \infty\}$ $\|x_i\| \leqslant s_i$
2. $\exists n_0 \in \{k, \ldots, \infty\}$ *such that* $\|x_{n_0}\| < s_{n_o}$

then we have

$$\sum_{i=k}^{\infty} x_i \textit{ converges}$$

and

$$\left\|\sum_{i=k}^{\infty} x_i\right\| < \sum_{i=k}^{\infty} s_i$$

Proof. Using [theorem: 14.352] it follows that

$$\sum_{i=n_0+1}^{\infty} s_i \text{ converges and } \sum_{i=k}^{\infty} s_i = \sum_{i=k}^{n_0} s_i + \sum_{i=n_0+1}^{\infty} s_i \tag{14.214}$$

As $\forall i \in \{n_0 + 1, \ldots \infty\}$ $\|x_i\| \leqslant s_i$ we can use dominant convergence [theorem: 14.372] that

$$\sum_{i=n_0+1}^{\infty} x_i \text{ converges and } \left\|\sum_{i=n_0+1}^{\infty} x_i\right\| \leqslant \sum_{i=n_0+1}^{\infty} s_i \tag{14.215}$$

So using [theorem: 14.352] we have that

$$\sum_{i=k}^{\infty} x_i \text{ converges and } \sum_{i=k}^{\infty} x_i = \sum_{i=k}^{n_0} x_i + \sum_{i=n_0+1}^{\infty} x_i \tag{14.216}$$

We have also that

$$\begin{aligned}
\left\|\sum_{i=k}^{n_0} x_i\right\| &\leqslant \sum_{i=k}^{n_0} \|x_i\| \\
&= \sum_{i=k}^{n_0-1} \|x_i\| + \|x_{n_0}\| \\
&\leqslant \sum_{i=k}^{n_0-1} s_i + \|x_{n_0}\| \\
&< \sum_{i=k}^{n_0-1} s_i + s_{n_0} \\
&= \sum_{i=k}^{n_0} s_i
\end{aligned}$$

proving that

$$\left\|\sum_{i=k}^{n_0} x_i\right\| < \sum_{i=k}^{n_0} s_o \tag{14.217}$$

So we have that $\sum_{i=k}^{\infty} x_i$ converges and

$$\begin{aligned}\left\|\sum_{i=k}^{\infty} x_i\right\| &\underset{[\text{eq: } 14.216]}{=} \left\|\sum_{i=k}^{n_0} x_i + \sum_{i=n_0+1}^{\infty} x_i\right\| \\ &\leqslant \left\|\sum_{i=k}^{n_0} x_i\right\| + \left\|\sum_{i=n_0+1}^{\infty} x_i\right\| \\ &\leqslant_{[\text{eq: } 14.215]} \left\|\sum_{i=k}^{n_0} x_i\right\| + \sum_{i=n_0+1}^{\infty} s_i \\ &<_{[\text{eq: } 14.217]} \sum_{i=k}^{n_0} s_i + \sum_{i=n_0+1}^{\infty} s_i \\ &\underset{[\text{eq: } 14.214]}{=} \sum_{i=k}^{\infty} s_i\end{aligned}$$

proving

$$\left\|\sum_{i=k}^{\infty} x_i\right\| < \sum_{i=k}^{\infty} s_i$$

□

A consequence of the dominant convergence theorem is that every absolute convergent serie is convergent.

Corollary 14.374. *Let $k \in \mathbb{N}_0$, $\langle X, \|\|\rangle$ be a Banach space and $\sum_{i=k}^{\infty} x_i$ a **absolute convergent series** then $\sum_{i=k}^{\infty} x_i$ converges and $\|\sum_{i=k}^{\infty} x_i\| \leqslant \sum_{i=k}^{\infty} \|x_i\|$*

Proof. As $\sum_{i=k}^{\infty} x_i$ is absolute convergent we have by definition that $\sum_{i=k}^{\infty} \|x_i\|$ converges and as trivially $\forall i \in \{k, \ldots, \infty\}$ $\|x_i\| \leqslant \|x_i\|$ it follows from the previous theorem [theorem: 14.372] that $\sum_{i=k}^{\infty} x_i$ converges and that $\|\sum_{i=k}^{\infty} x_i\| \leqslant \sum_{i=k}^{\infty} \|x_i\|$. □

Theorem 14.375. *Let $k \in \mathbb{N}_0$, $\langle X, \|\|\rangle$ be a normed space and $\{x_i\}_{i \in \{k, \ldots, \infty\}} \subseteq X$ a sequence such that there exist a convergent series $\sum_{i=k}^{\infty} s_i$ with $\forall i \in \{k, \ldots, \infty\}$ $\|x_i\| \leqslant s_i$ then $\sum_{i=k}^{\infty} \|x_i\|$ converges [using the normed space $\langle \mathbb{R}, \|\rangle$].*

Proof. As $\{\|x_i\|\}_{i \in \{k, \ldots, \infty\}} \subseteq \mathbb{R}$ is a sequence in the Banach space $\langle \mathbb{R}, \|\rangle$ [see theorem: 14.330] it follows from [theorem: 14.372] that $\sum_{i=k}^{\infty} \|x_i\|$ converges. □

The following theorem show that only in a Banach space convergence and absolute convergence is the same.

Theorem 14.376. *Let $\langle X, \|\|\rangle$ be a normed space then*

$$\langle X, \|\|\rangle \textit{ is Banach} \Leftrightarrow \text{every absolute convergent series converges}$$

Proof.

- **⇒.** This follows from [corollary: 14.374]
- **⇐.** Let $\{x_n\}_{n \in \{k, \ldots, \infty\}} \subseteq X$ be a Cauchy sequence. Then $\forall n \in \mathbb{N}$ there exist a $N'_n \in \{k, \ldots, \infty\}$ such that $\forall r, s \geqslant N'_n$ we have $\|x_r - x_s\| < \frac{1}{2^{n+1}}$ define $N_n = \max(N'_n, n)$ then $N_n + 1, N_n \geqslant N'_n$ so that

  $$\|x_{N_n+1} - x_{N_n}\| < \frac{1}{2^{n+1}}$$

  Hence we have found a sequence $\{x_{N_n}\}_{n \in \mathbb{N}} \subseteq X$ such that

  $$\forall n \in \mathbb{N}\ \|x_{N_n+1} - x_{N_n}\| < \frac{1}{2^{n+1}} \text{ and } N_n \geqslant n \tag{14.218}$$

  Define

  $$\{y_i\}_{i \in \mathbb{N}} \text{ by } y_i = \begin{cases} x_{N_1} \text{ if } i = 1 \\ x_{N_i} - x_{N_{i-1}} \text{ if } i \in \{2, \ldots, \infty\} \end{cases}$$

  then we have

  $$\forall i \in \{2, \ldots, \infty\}\ \|y_i\| = \|x_{N_i} - x_{N_{i-1}}\| < \frac{1}{2^{(i-1)+1}} = \frac{1}{2^i} \tag{14.219}$$

Using [theorem: 14.348] $\sum_{i=1}^{\infty} \frac{1}{2^i}$ is convergent so that by [theorem: 14.352] $\sum_{i=2}^{\infty} \frac{1}{2^i}$ converges. Using then [theorem: 14.375] on [eq: 14.219] proves that $\sum_{i=2}^{\infty} \|y_i\|$ converges, further by [theorem: 14.352] we have that $\sum_{i=1}^{\infty} \|y_i\|$ converges. Hence $\{y_i\}_{i\in\mathbb{N}}$ is absolute convergent. By the hypothesis every absolute convergent series converges hence we have that

$$\sum_{i=1}^{\infty} y_i \text{ converges or } \lim_{n\rightarrow\infty} \sum_{i=1}^{n} y_i \text{ exist} \tag{14.220}$$

Let $n\in\mathbb{N}$ then we have that

$$\begin{aligned}
\sum_{i=1}^{n} y_i &= \begin{cases} y_1 \text{ if } n=1 \\ y_1+\sum_{i=2}^{n} y_i \text{ if } 1<n \end{cases}\\
&= \begin{cases} x_{N_1} \text{ if } n=1 \\ x_{N_1}+\sum_{i=2}^{n} (x_{N_i}-x_{N_{i-1}}) \text{ if } 1<n \end{cases}\\
&\underset{\text{[theorem: 11.21]}}{=} \begin{cases} x_{N_1} \text{ if } n=1 \\ x_{N_1}+x_{N_n}-x_{N_1} \text{ if 1<n} \end{cases}\\
&= x_{N_n}
\end{aligned}$$

Combining the above with [eq: 14.220] proves that

$$\{x_{N_i}\}_{i\in\mathbb{N}} \text{ converges to a limit } x \tag{14.221}$$

Take $\varepsilon\in\mathbb{R}^+$ then as $\{x_n\}_{n\in\{k,\ldots,\infty\}}$ is Cauchy there exist a $M_1\in\{k,\ldots,\infty\}$ such that $\forall n,m\geqslant M_1$ we have

$$\|x_n-x_m\|<\frac{\varepsilon}{2}$$

Further as $\lim_{n\rightarrow\infty} x_{N_n}=x$ there exist a $M_2\in\mathbb{N}$ such that $\forall n\geqslant M_2$ we have

$$\|x_{N_n}-x\|<\frac{\varepsilon}{2}$$

Let $N=\max(M_1,M_2)$ then if $n\geqslant N$ we have $n\geqslant M_1 \underset{\text{[eq: 14.218]}}{\Rightarrow} N_n\geqslant n\geqslant M_1$ and $n\geqslant M_2$ so that

$$\|x_n-x\|=\|x_n-x_{N_n}+x_{N_n}-x\|\leqslant\|x_n-x_{N_n}\|+\|x_{N_n}-x\|<\frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon$$

proving that $\{x_n\}_{n\in\{k,\ldots,\infty\}}$ converges to x. As $\{x_n\}_{n\in\{k,\ldots,\infty\}}$ was a arbitrary Cauchy sequence it follows that $\langle X,\|\|\rangle$ is a Banach space. □

We have the following equivalent definitions for absolute convergent series.

Theorem 14.377. *Let $k\in\mathbb{N}_0$, $\langle X,\|\|\rangle$ a normed space and $\{x_i\}_{i\in\{k,\ldots,\infty\}}\subseteq X$ a sequence then the following are equivalent:*

1. *$\sum_{i=k}^{\infty} x_i$ is absolute convergent*
2. *$\forall\varepsilon\in\mathbb{R}^+$ there exist a $N\in\{k,\ldots,\infty\}$ such that $\forall m\in\mathbb{N}$ and $\forall n\geqslant N$ we have $\sum_{i=n+1}^{n+m}\|x_i\|<\varepsilon$*
3. *$\forall\varepsilon\in\mathbb{R}^+$ there exist a finite $K\subseteq\{k,\ldots,\infty\}$ such that $\forall H\subseteq\{k,\ldots,\infty\}$ with H finite and $K\bigcap H=\varnothing$ we have $\sum_{i\in H}\|x_i\|<\varepsilon$*

Proof. As $\sum_{i=k}^{\infty}\|x_i\|$ is a series of non negative numbers and absolute convergences of $\sum_{i=k}^{\infty} x_i$ is the same as convergence of $\sum_{i=k}^{\infty}\|x_i\|$ the proof follows from [theorem: 14.356]. □

Lemma 14.378. *Let $k\in\mathbb{N}_0$, $\langle X,\|\|\rangle$ be a Banach space and $\sum_{i=k}^{\infty} x_i$ a absolute convergent series then $\forall\varepsilon\in\mathbb{R}^+$ there exist a finite $K\subseteq\{k,\ldots,\infty\}$ such that*

1. *$\forall L\subseteq\{k,\ldots,\infty\}$ with L finite and $K\bigcap L=\varnothing$ we have $\sum_{i\in L}\|x_i\|<\varepsilon$*
2. *$\forall H\subseteq\{k,\ldots,\infty\}$ with H finite and $K\subseteq H$ we have $\|\sum_{i=k}^{\infty} x_i-\sum_{i\in H} x_i\|<\varepsilon$*

Proof. Let $\varepsilon\in\mathbb{R}^+$. As by [theorem: 14.376] absolute convergence implies convergence we have that $\lim_{n\rightarrow\infty}\sum_{i=k}^{n} x_i$ converges to $\sum_{i=k}^{\infty} x_i$, hence there exist a $N\in\{k,\ldots,\infty\}$ so that

$$\forall n\geqslant N \text{ we have } \left\|\sum_{i=k}^{n} x_i-\sum_{i=k}^{\infty} x_i\right\|<\frac{\varepsilon}{2} \tag{14.222}$$

Using absolute convergence and the previous theorem [see theorem: 14.377] there exist a finite $K'\subseteq\{k,\ldots,\infty\}$ such that for every finite $L\subseteq\{k,\ldots,\infty\}$ with $K'\bigcap L=\varnothing$ we have that that $\sum_{i\in L}\|x_i\|<\frac{\varepsilon}{2}$. Take $K=\{k,\ldots,N\}\bigcup K'$ then we have $\{k,\ldots,N\}\subseteq K\subseteq\{k,\ldots,\infty\}$ and for every finite $L\subseteq\{k,\ldots,\infty\}$ with $K\bigcap L=\varnothing$ we have, as $K'\subseteq L$, that $K'\bigcap L\neq\varnothing$ so that $\sum_{i\in L}\|x_i\|<\frac{\varepsilon}{2}$. So we have the existence of a $K\subseteq\{k,\ldots,\infty\}$ such that

$$\{k,\ldots,N\}\subseteq K\wedge\forall L\subseteq\{k,\ldots,\infty\}\text{ with } L \text{ finite and } K\bigcap L=\varnothing \text{ we have }\sum_{i\in L}\|x_i\|<\frac{\varepsilon}{2}\tag{14.223}$$

proving (1).

Take $H\subseteq\{k,\ldots,\infty\}$ with $K\subseteq H\wedge H$ finite. Let $m\underset{\text{is finite}}{=}\max(H)$ so that $H\subseteq\{k,\ldots,m\}$ then

$$\begin{aligned}\left\|\sum_{i=k}^{\infty}x_i-\sum_{i\in H}x_i\right\| &= \left\|\sum_{i\in H}x_i-\sum_{i=k}^{\infty}x_i\right\|\\ &= \left\|\sum_{i\in\{k,\ldots,m\}}x_i-\sum_{i\in\{k,\ldots,m\}\setminus H}x_i-\sum_{i=k}^{\infty}x_i\right\|\\ &= \left\|\sum_{i=k}^{m}x_i-\sum_{i\in\{k,\ldots,m\}\setminus H}x_i-\sum_{i=k}^{\infty}x_i\right\|\\ &\leqslant \left\|\sum_{i=k}^{m}x_i-\sum_{i=k}^{\infty}x_i\right\|+\left\|\sum_{i\in\{k,\ldots,m\}\setminus H}x_i\right\|\end{aligned}\tag{14.224}$$

As $\{k,\ldots,m\}\setminus H\subseteq\{k,\ldots,\infty\}$ and

$$K\bigcap(\{k,\ldots,m\}\setminus H)=K\bigcap(\{k,\ldots,m\}\bigcap H^c)=(K\bigcap H^c)\bigcap\{k,\ldots,m\}\underset{K\subseteq H}{=}\varnothing\bigcap\{k,\ldots,m\}=\varnothing$$

so that by [eq: 14.223] we have

$$\left\|\sum_{i\in\{k,\ldots,m\}\setminus H}x_i\right\|<\frac{\varepsilon}{2}\tag{14.225}$$

As $\{k,\ldots,N\}\subseteq K\subseteq H$ we have that $N\leqslant\max(H)=m$ so that by [eq: 14.222] we have

$$\left\|\sum_{i=k}^{m}x_i-\sum_{i=k}^{\infty}x_i\right\|<\frac{\varepsilon}{2}\tag{14.226}$$

combining [eqs: 14.224, 14.225 and 14.226] proves that

$$\left\|\sum_{i=k}^{\infty}x_i-\sum_{i\in H}x_i\right\|<\varepsilon$$

and (2). □

Theorem 14.379. *Let $k\in\mathbb{N}_0$, $\langle X,\|\|\rangle$ a Banach space and $\sum_{i=k}^{\infty}x_i$ a absolute convergent series then we have:*

1. *If $\beta:\{k,\ldots,\infty\}\rightarrow\{k,\ldots,\infty\}$ is a bijection then $\sum_{i=k}^{\infty}x_{\beta(i)}$ is absolute convergent [and by [theorem: 14.376] it follows that $\sum_{i=k}^{\infty}x_{\beta(i)}$ is also convergent]*
2. *If $\alpha,\beta:\{k,\ldots,\infty\}\rightarrow\{k,\ldots,\infty\}$ are bijections then $\sum_{i=k}^{\infty}x_{\alpha(i)}=\sum_{i=k}^{\infty}x_{\beta(i)}$*
3. *If $\beta:\{k,\ldots,\infty\}\rightarrow\{k,\ldots,\infty\}$ is a bijection then $\sum_{i=k}^{\infty}x_i=\sum_{i=k}^{\infty}x_{\beta(i)}$*

Proof.

1. As $\sum_{i=k}^{\infty}x_i$ is absolute convergent $\sum_{i=k}^{\infty}\|x_i\|$ is a convergent series of non negative real numbers, hence using [theorem: 14.358] $\sum_{i=k}^{\infty}\|x_{\beta(i)}\|$ is also convergent proving that $\sum_{i=k}^{\infty}x_{\beta(i)}$ is absolute convergent.
2. Let $s=\sum_{i=k}^{\infty}x_{\alpha(i)}$ and $t=\sum_{i=k}^{\infty}x_{\beta(i)}$. Assume that $s\neq t$ then $\|s-t\|\in\mathbb{R}^+$. Using [lemma: 14.378] there exists finite K_α, $K_\beta\subseteq\{k,\ldots,\infty\}$ such that for every finite H_α,H_β with $K_\alpha\subseteq H_\alpha$, $K_\beta\subseteq H_\beta$ we have

$$\left\|\sum_{i\in H_\alpha}x_{\alpha(i)}-s\right\|<\frac{\|s-t\|}{4}\wedge\left\|\sum_{i\in H_\beta}x_{\beta(i)}-t\right\|<\frac{\|s-t\|}{4}\tag{14.227}$$

Define now $P_{\alpha\beta}=K_\alpha\bigcup\alpha^{-1}(\beta(K_\beta))$ and $P_{\beta\alpha}=K_\beta\bigcup\beta^{-1}(\alpha(K_\alpha))$. As α and β are bijections and K_α and K_β are finite it follows from [theorems: 2.68, 6.44 and 6.33] that $P_{\alpha\beta}$ and $P_{\beta\alpha}$ are finite and $K_\alpha\subseteq P_{ab}$ and $K_\beta\subseteq P_{\beta\alpha}$ so that by [eq: 14.227]

$$\left\|\sum_{i\in P_{\alpha\beta}} x_{\alpha(i)}-s\right\|<\frac{\|s-t\|}{4}\wedge\left\|\sum_{i\in P_{\beta\alpha}} x_{\beta(i)}-t\right\|<\frac{\|s-t\|}{4} \tag{14.228}$$

Further as $\alpha:\{k,\ldots,\infty\}\rightarrow\{k,\ldots,\infty\}$ and $\beta:\{k,\ldots,\infty\}\rightarrow\{k,\ldots,\infty\}$ are bijections we have that

$$\alpha_{|P_{\alpha\beta}}:P_{\alpha\beta}\rightarrow\alpha(P_{\alpha\beta})=\alpha(K_\alpha)\bigcup\beta(K_\beta)\text{ and }\beta_{|P_{\beta\alpha}}:P_{\beta\alpha}\rightarrow\beta(P_{\beta\alpha})=\beta(K_\alpha)\bigcup\alpha(K_\alpha)$$

are bijections so that we can apply [theorem: 11.36]

$$\begin{aligned}
\sum_{i\in P_{\alpha\beta}} x_{\alpha(i)} &= \sum_{i\in P_{\alpha\beta}} x_{\alpha_{|P_{\alpha\beta}}(i)}\\
&\underset{\text{[theorem: 11.36]}}{=} \sum_{i\in\alpha(P_{\alpha\beta})} x_i\\
&= \sum_{i\in\alpha(K_\alpha)\bigcup\beta(K_\beta)} x_i\\
&= \sum_{i\in\beta(K_\beta)\bigcup\alpha(K_\alpha)} x_i\\
&= \sum_{i\in\beta(P_{\beta\alpha})} x_i\\
&\underset{\text{[theorem: 11.36]}}{=} \sum_{i\in P_{\beta\alpha}} x_{\beta_{|P_{\alpha\beta}}(i)}\\
&= \sum_{i\in P_{\beta\alpha}} x_{\beta(i)}
\end{aligned}$$

proving that

$$\sum_{i\in P_{\alpha\beta}} x_{\alpha(i)}=\sum_{i\in P_{\beta\alpha}} x_{\beta(i)} \tag{14.229}$$

Now

$$\begin{aligned}
\|s-t\| &\underset{\text{[eq: 14.229]}}{=} \left\|s-\sum_{i\in P_{\alpha\beta}} x_{\alpha(i)}+\sum_{i\in P_{\beta\alpha}} x_{\beta(i)}-t\right\|\\
&\leqslant \left\|s-\sum_{i\in P_{\alpha\beta}} x_{\alpha(i)}\right\|+\left\|\sum_{i\in P_{\beta\alpha}} x_{\beta(i)}-t\right\|\\
&<_{\text{[eq: 14.228]}} \frac{\|s-t\|}{4}+\frac{\|s-t\|}{4}\\
&< \frac{\|s-t\|}{2}\\
&< \|s-t\|
\end{aligned}$$

a contradiction. So the assumption is incorrect and we must have $s=t$ or

$$\sum_{i=k}^{\infty} x_{\alpha(i)}=\sum_{i=k}^{\infty} x_{\beta(i)}$$

3. As $\mathrm{Id}_{\{k,\ldots,\infty\}}:\{k,\ldots,\infty\}\rightarrow\{k,\ldots,\infty\}$ is a bijection we have

$$\sum_{i=k}^{\infty} x_{\alpha(i)}\underset{(2)}{=}\sum_{i=k}^{\infty} x_{\mathrm{Id}_{\{k,\ldots,\infty\}}(i)}=\sum_{i=k}^{\infty} x_i$$

□

14.10.4 Properties of complete spaces.

Definition 14.380. *Let $\langle X,d\rangle$ be a metric space then a function $f:X\rightarrow X$ is a* ***contraction*** *if there exist a $\lambda\in[0,1[$ such that $\forall x,y\in X$ we have $d(f(x),f(y))\leqslant\lambda\cdot d(x,y)$*

Theorem 14.381. *Let $\langle X, d\rangle$ be a metric space and $f: X \to X$ a contraction then f is continuous.*

Proof. As $f: X \to X$ is a contraction there exist a $\lambda \in [0, 1[$ such that

$$\forall x, y \in X\ d(f(x), f(y)) \leqslant \lambda \cdot d(x, y)$$

Let $x \in X$ and $\varepsilon \in \mathbb{R}^+$. If $y \in X$ is such that $d(x, y) < \varepsilon$ then we have

$$d(f(x), f(y)) \leqslant \lambda \cdot d(x, y) < d(x, y) < \varepsilon$$

which by [theorem: 14.126] proves that f is continuous at x. Hence as x was chosen arbitrary $f: X \to X$ is continuous. □

Contractions are special because they guarantee the existence of a fixed point, that is a point that is mapped on itself by the contraction, this is proved in the Banach Fixed Point theorem.

Theorem 14.382. (Banach Fixed Point Theorem) *Let $\langle X, d\rangle$ be a **complete** metric space and $f: X \to X$ a contraction then there exists a **unique** $x_0 \in X$ such that $f(x_0) = x_0$. Further if we define for $n \in \mathbb{N}$ $f^{(n)}$ recursively by*

$$\begin{array}{rcl} f^{(1)} & = & f \\ f^{(n+1)} & = & f \circ f^{(n)} \end{array}$$

then $\forall x \in X$ we have

$$\lim_{n \to \infty} f^{(n)}(x) = x_0$$

Proof. As $f: X \to X$ is a contraction there exist a $\lambda \in [0, 1[$ such that

$$\forall x, y \in X\ d(f(x), f(y)) \leqslant \lambda \cdot d(x, y)$$

As a metric space is non empty there exist a $x \in X$. We prove now by induction that

$$\forall n \in \mathbb{N} \text{ we have } d(f^{(n)}(x), f^{(n+1)}(x)) \leqslant \lambda^n \cdot d(x, f(x)) \tag{14.230}$$

Proof. Define

$$S = \{n \in \mathbb{N} | d(f^{(n)}(x), f^{(n+1)}(x)) \leqslant \lambda^n \cdot d(x, f(x))\}$$

then we have:

$1 \in S$. As

$$d(f^{(1)}(x), f^{(1+1)}(x)) = d(f(x), (f \circ f)(x)) = d(f(x), f(f(x))) \leqslant \lambda \cdot d(x, f(x)) = \lambda^1 \cdot d(x, f(x))$$

it follows that $1 \in S$

$n \in S \Rightarrow n + 1 \in S$. We have

$$\begin{array}{rcl} d(f^{(n+1)}(x), f^{((n+1)+1)}(x)) & = & d((f \circ f^{(n)})(x), (f \circ f^{(n+1)})(x)) \\ & = & d(f(f^{(n)}(x)), f(f^{(n+1)}(x))) \\ & \leqslant & \lambda \cdot d(f^{(n)}(x), f^{(n+1)}(x)) \\ & \leqslant_{n \in S} & \lambda \cdot \lambda^n \cdot d(x, f(x)) \\ & = & \lambda^{n+1} \cdot d(x, f(x)) \end{array}$$

proving that $n + 1 \in S$.

Mathematical induction proves [eq: 14.230]. □

Next we prove by induction that

$$\forall n, k \in \mathbb{N} \text{ we have } d(f^{(n)}(x), f^{(n+k)}(x)) \leqslant \left(\sum_{i=1}^{k} \lambda^{n+(i-1)}\right) \cdot d(x, f(x)) \tag{14.231}$$

Proof. Given $n \in \mathbb{N}$ define

$$T_n = \left\{k \in \mathbb{N} | d(f^{(n)}(x), f^{(n+k)}(x)) \leqslant \left(\sum_{i=1}^{k} \lambda^{n+(i-1)}\right) \cdot d(x, f(x))\right\}$$

then we have:

$1 \in T_n$. We have by [eq: 14.230] that

$$\begin{aligned} d(f^{(n)}(x), f^{(n+1)}(x)) &\leqslant \lambda^n \cdot d(x, f(x)) \\ &= \lambda^{n+(1-1)} \cdot d(x, f(x)) \\ &= \left(\sum_{i=1}^{1} \lambda^{n+(i-1)}\right) \cdot d(x, f(x)) \end{aligned}$$

proving that $1 \in T_n$.

$k \in T_n \Rightarrow k+1 \in T_n$. Then we have:

$$\begin{aligned} d(f^{(n)}(x), f^{(n+(k+1))}(x)) &\leqslant \\ d(f^{(n)}(x), f^{(n+k)}(x)) + d(f^{(n+k)}(x), f^{(n+(k+1))}(x)) &\leqslant_{k \in T_n} \\ \left(\sum_{i=1}^{k} \lambda^{n+(i-1)}\right) \cdot d(x, f(x)) + d(f^{(n+k)}(x), f^{(n+(k+1))}(x)) &\leqslant_{[\text{eq: } 14.230]} \\ \left(\sum_{i=1}^{k} \lambda^{n+(i-1)}\right) \cdot d(x, f(x)) + \lambda^{n+k} \cdot d(x, f(x)) &= \\ \left(\sum_{i=1}^{k} \lambda^{n+(i-1)} + \lambda^{n+k}\right) \cdot d(x, f(x)) &= \\ \left(\sum_{i=1}^{k} \lambda^{n+(i-1)} + \lambda^{n+((k+1)-1)}\right) \cdot d(x, f(x)) &= \\ \left(\sum_{i=1}^{k+1} \lambda^{n+(i-1)}\right) \cdot d(x, f(x)) & \end{aligned}$$

proving that $k+1 \in T_n$.

Mathematical induction proves [eq: 14.231] □

As $\lambda \in [0,1[$ we have that $\lambda \neq 1$ so that for $k \in \mathbb{N}$

$$\sum_{i=1}^{k} \lambda^{n+(i-1)} \underset{[\text{lemma: } 14.347]}{=} \lambda^n \cdot \frac{1-\lambda^k}{1-\lambda}$$

which combined with [eq: 14.231] proves that

$$\forall k, n \in \mathbb{N} \text{ that } d(f^{(n)}(x), f^{(n+k)}(x)) \leqslant \lambda^n \cdot \frac{1-\lambda^k}{1-\lambda} \cdot d(x, f(x)) < \frac{\lambda^n}{1-\lambda} \cdot d(x, f(x)) \tag{14.232}$$

Let $\varepsilon \in \mathbb{R}^+$ then as $0 \leqslant \lambda < 1$ we have by [theorem: 14.292] that there exist a $N \in \mathbb{N}$ such that $\forall n \geqslant N$ $\lambda^n = |\lambda^n - 0| < \frac{\varepsilon}{\frac{1}{1-\lambda} \cdot d(x, f(x))}$. If n, $m \geqslant N$ then we may always assume that $m \geqslant n$ so that

$$\begin{aligned} d(f^{(n)}(x), f^{(m)}(x)) &= d(f^{(n)}(x), f^{(n+(m-n))}(x)) \\ &<_{[\text{eq: } 14.232]} \frac{\lambda^n}{1-\lambda} \cdot d(x, f(x)) \\ &< \frac{\varepsilon}{\frac{1}{1-\lambda} \cdot d(x, f(x))} \cdot \frac{1}{1-\lambda} \cdot d(x, f(x)) \\ &= \varepsilon \end{aligned}$$

This proves that $\{f^{(n)}(x)\}_{n \in \{1, \ldots, \infty\}}$ is Cauchy, as $\langle X, d \rangle$ is complete it follows that

$$\lim_{n \rightarrow \infty} f^{(n)}(x) = x_0 \text{ exist} \tag{14.233}$$

As by [theorem: 14.381] $f: X \rightarrow X$ is continuous it follows from the above and [theorem: 14.308] that

$$\lim_{n \rightarrow \infty} f^{(n+1)}(x) = \lim_{n \rightarrow \infty} f(f^{(n)}(x)) = f\left(\lim_{n \rightarrow \infty} f^{(n)}(x)\right) = f(x_0) \tag{14.234}$$

So given $\varepsilon \in \mathbb{R}^+$ there exist by [eq: 14.233] a $N_1 \in \mathbb{N}$ such that $\forall n \geqslant N$ we have

$$d(f^{(n)}(x), x_0) < \frac{\varepsilon}{2}$$

and by [eq: 14.234] there exist a $N_2 \in \mathbb{N}$ such that $\forall n \geqslant N_2$ we have

$$d(f^{(n+1)}(x), f(x_0)) < \frac{\varepsilon}{2}$$

Hence if $n \geqslant \max(N_1, N_2 + 1)$ then we have

$$\begin{array}{lll} d(f(x_0), x_0) & \leqslant & d(f(x_0), f^{(n)}(x)) + d(f^{(n)}(x), x_0) \\ & = & d(f^{(n)}(x), f(x_0)) + d(f^{(n)}(x), x_0) \\ & <_{[\text{eq: } 14.233]} & d(f^{(n)}(x), f(x_0)) + \frac{\varepsilon}{2} \\ & = & d(f^{((n-1)+1)}(x), f(x_0)) + \frac{\varepsilon}{2} \\ & <_{n > N_2+1 \Rightarrow n-1 \geqslant N_2} & \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\ & = & \varepsilon \end{array}$$

As $\varepsilon \in \mathbb{R}^+$ was chosen arbitrary we can use [theorem: 10.31] on the above giving $d(f(x_0), x_0) = 0$, hence

$$f(x_0) = x_0 \text{ where } \lim_{n \rightarrow \infty} f^{(n)}(x) = x_0$$

Finally to prove uniqueness, assume that there exist a $x_1 \in X$ with $x_0 \neq x_1$ such that $f(x_1) = x_1$. As $\lambda < 1$ and $0 < d(x_0, x_1)$ it follows that $\lambda \cdot d(x_1, x_0) < 1 \cdot d(x_1, x_0) = d(x_1, x_o)$, further we have

$$0 < d(x_1, x_0) = d(f(x_1), f(x_0)) \leqslant \lambda \cdot d(x_1, x_0) < d(x_1, x_0)$$

giving the contradiction $d(x_1, x_0) < d(x_1, x_0)$. So our assumption must be wrong proving that x_0 is unique. □

Lemma 14.383. *Let $\langle X, d\rangle$ be a complete metric space and let $\{A_i\}_{i \in \mathbb{N}}$ be a sequence of non empty bounded closed sets such that $\forall i \in \mathbb{N}$ we have $A_{i+1} \subseteq A_i$ and $\lim_{n \rightarrow \infty} \operatorname{diam}(A_i) = 0$ then we have that*

$$\bigcap_{i \in \mathbb{N}} A_i \neq \varnothing$$

Proof. As $\{A_i\}_{i \in \mathbb{N}}$ is a family of non empty sets it follows from the Axiom of Choice [see theorem: 3.103] there exist a function x: $\mathbb{N} \rightarrow \bigcup_{i \in \mathbb{N}} A_i$ such that $\forall i \in \mathbb{N}$ $x(i) \in A_i$. This defines a family

$$\{x_i\}_{i \in \mathbb{N}} \text{ such that } \forall i \in \mathbb{N}\ x_i \in A_i$$

Let $\varepsilon \in \mathbb{R}^+$ then as $\lim_{n \rightarrow \infty} \operatorname{diam}(A_i) = 0$ there exist a $N \in \mathbb{N}$ such that $\forall n \geqslant N$ we have

$$\operatorname{diam}(A_n) = |\operatorname{diam}(A_n) - 0| < \varepsilon$$

If $n, m \geqslant N$ then as $\{A_n\}_{n \in \mathbb{N}}$ is decreasing it follows from [theorem: 14.284] that $A_n \subseteq A_N \wedge A_m \subseteq A_N$ so that $x_n \in A_N \wedge x_m \in A_N$, hence

$$d(x_n, x_m) \leqslant \sup(\{d(x, y) | x, y \in A_N\}) = \operatorname{diam}(A_N) < \varepsilon$$

which proves that

$$\{x_i\}_{i \in \mathbb{N}} \text{ is a Cauchy sequence}$$

As $\langle X, d\rangle$ is complete it follows that

$$y = \lim_{i \rightarrow \infty} x_i \text{ exists}$$

Let $n \in \mathbb{N}$, $\varepsilon \in \mathbb{R}^+$. Give $U \in \mathcal{T}$ a open set such that $y \in U$ then by [theorem: 14.64] there exist a $\delta \in \mathbb{R}^+$ such that

$$y \in B_d(y, \delta) \subseteq U$$

As $y = \lim_{i \rightarrow \infty} x_i$ there exist a $M \in \mathbb{N}$ such that $\forall m \geqslant M$ we have

$$d(x_m, y) < \delta \text{ or } x_m \in B_d(y, \delta) \subseteq U$$

If $K = \max(n, M)$ then $n \leqslant K$ so that $x_K \in A_K \subseteq A_n$, $M \leqslant K \Rightarrow x_K \in U$ proving that $x_K \in A_n \bigcap U$ or $A_n \bigcap U \neq \varnothing$. Using [theorem: 14.28] we conclude then that $y \in \overline{A_n} =_{A_n \text{ is closed} \wedge [\text{theorem: } 14.25]} A_n$. So we have proved that $\forall n \in \mathbb{N}$ $y \in A_n$ or $y \in \bigcap_{n \in \mathbb{N}} A_n$, hence

$$\bigcap_{n \in \mathbb{N}} A_n \neq \varnothing$$

□

Theorem 14.384. (Baire Category Theorem)

1. *If $\langle X, \mathcal{T}\rangle$ is a compact Hausdorff topological space then X is a Baire space.*
2. *If $\langle X, d\rangle$ is a complete metric space then X is a Baire space*

See [definition: 14.53] for the definition of a Baire space.

Proof. Let $\{A_n\}_{n\in\mathbb{N}}$ be a countable family of sets in X such that $\forall n\in\mathbb{N}$ A_n is closed and $(A_n)^\circ=\varnothing$ then we must prove that $(\bigcup_{n\in\mathbb{N}} A_n)^\circ=\varnothing$

1. As $\langle X, \mathcal{T}\rangle$ is compact and Hausdorff it follows from [theorem: 14.236] that

$$\langle X, \mathcal{T}\rangle \text{ is regular} \tag{14.235}$$

Let $U_0\in\mathcal{T}$ be a open set such that $U_0\neq\varnothing$. Then we use recursion to construct a family of non empty open sets $\{U_n\}_{n\in\mathbb{N}}\subseteq\mathcal{T}$ such that

$$\forall n\in\mathbb{N} \text{ we have } \varnothing\neq U_n,\ \overline{U}_n\subseteq U_{n-1} \text{ and } \overline{U_n}\bigcap A_n=\varnothing \tag{14.236}$$

Proof. As $(A_1)^0=\varnothing$ we have that $U_0\nsubseteq A_1$ [for if $U_0\subseteq A_1$ then $\varnothing\neq U_0\subseteq(A_1)^\circ=\varnothing$ a contradiction] so there exist a $y\in U_0$ such that $y\notin A_1$. Using the fact that A_1 is closed and X is regular we can use [theorem: 14.208] to find a $U_1\in\mathcal{T}$ such that

$$y\in U_1\Rightarrow U_1\neq\varnothing,\ \overline{U_1}\subseteq U_0 \text{ and } \overline{U_1}\bigcap A_1=\varnothing$$

For the recursion step. Let U_{n-1} a non empty open set, as $(A_n)^\circ=\varnothing$, we have that $U_{n-1}\nsubseteq A_n$ [for if $U_{n-1}\subseteq A_n$ then $\varnothing\neq U_{n-1}\subseteq(A_n)^\circ=\varnothing$ a contradiction], hence there exist a $y\in U_{n-1}$ such that $y\notin A_n$. Using [theorem: 14.208] there exist a $U_n\in\mathcal{T}$ such that

$$y\in U_n\Rightarrow U_n\neq\varnothing,\ \overline{U_n}\subseteq U_{n-1} \text{ and } \overline{U_n}\bigcap A_n=\varnothing$$ □

By construction we have $\forall n\in\mathbb{N}$ that $\overline{U_{n+1}}\subseteq U_n\subseteq\overline{U_n}$. Let I be a finite non empty subset of $\mathbb{N}_0$ then $m=\max(I)$ exist, so $\forall n\in I$ $n\leqslant m$ and we have by [theorem: 14.284] $\overline{U_m}\subseteq\overline{U_n}$ proving that $\overline{U_m}\subseteq\bigcap_{n\in I}\overline{U_n}$. As $\varnothing\neq U_m\subseteq\overline{U_m}$ it follows that $\bigcap_{n\in I}\overline{U_n}\neq\varnothing$ proving that $\{\overline{U_n}\}_{n\in\mathbb{N}}$ has the finite intersection property. As X is compact it follows from [theorem: 14.268]

$$\bigcap_{n\in\mathbb{N}}\overline{U_n}\neq\varnothing \text{ or there exist a } x\in\bigcap_{n\in\mathbb{N}}\overline{U_n}$$

As $1\in\mathbb{N}$ we have $x\in\overline{U_1}\subseteq U_0$, further, as $\forall n\in\mathbb{N}$ $x\in\overline{U_n}\underset{\overline{U_n}\bigcap A_n=\varnothing}{\Rightarrow}x\notin A_n$, it follows that $x\notin\bigcup_{n\in\mathbb{N}}A_n$. Hence we have proved that for every non empty open set U_0 $U_0\nsubseteq\bigcup_{i\in\mathbb{N}}A_i$ which, as $(\bigcup_{n\in\mathbb{N}}A_n)^\circ\subseteq\bigcup_{n\in\mathbb{N}}A_n$, proves that $(\bigcup_{n\in\mathbb{N}}A_n)^\circ=\varnothing$. Hence $\langle X, \mathcal{T}\rangle$ is a Baire space.

2. As $\langle X, d\rangle$ is a metric space it follows from [theorem: 14.205] that

$$\langle X, d\rangle \text{ is regular} \tag{14.237}$$

Let $U_0\in\mathcal{T}$ be a open set such that $U_0\neq\varnothing$. Then we use recursion to construct a family of non empty open sets $\{U_n\}_{n\in\mathbb{N}}\subseteq\mathcal{T}$ such that

$$\forall n\in\mathbb{N} \text{ we have } \varnothing\neq U_n,\ \overline{U}_n\subseteq U_{n-1},\ \overline{U_n}\bigcap A_n=\varnothing,\ U_n \text{ bounded and } \operatorname{diam}(U_n)\leqslant\frac{1}{n} \tag{14.238}$$

Proof. As $(A_1)^0=\varnothing$ we have that $U_0\nsubseteq A_1$ [for if $U_0\subseteq A_1$ then $\varnothing\neq U_0\subseteq(A_1)^\circ=\varnothing$ a contradiction] so there exist a $y\in U_0$ such that $y\notin A_1$. Using the fact that A_1 is closed and X is regular we can use [theorem: 14.208] to find a $V_1\in\mathcal{T}$ such that

$$y\in V_1\Rightarrow V_1\neq\varnothing,\ \overline{V_1}\subseteq U_0 \text{ and } \overline{V_1}\bigcap A_1=\varnothing$$

Take $U_1=V_1\bigcap B_d(y,\frac{1}{4})\ni y$ then

$$U_1\subseteq V_1\Rightarrow\overline{U_1}\bigcap A_1\subseteq\overline{V_1}\bigcap A_1=\varnothing \text{ and } \overline{U_1}\subseteq\overline{V_1}\subseteq U_0$$

$$U_1\subseteq B_d\left(y,\frac{1}{2}\right)\Rightarrow\overline{U_1}\subseteq\overline{B_d\left(y,\frac{1}{2}\right)}.$$

Using [example: 14.78] we have that $\overline{B_d(y,\frac{1}{2})}$ is bounded with $\operatorname{diam}(\overline{B_d(y,\frac{1}{2})})\leqslant 2\cdot\frac{1}{2}$, so, as $\overline{U_1}\subseteq\overline{B_d(y,\frac{1}{2})}$, it follows that $\overline{U_1}$ is bounded with $\operatorname{diam}(\overline{U_1})\leqslant 1$. Summarized

$$\varnothing\neq U_1,\ \overline{U_1}\subseteq U_0,\ \overline{U_1}\bigcap A_1=\varnothing,\ \overline{U_1} \text{ is bounded and } \operatorname{diam}(\overline{U_1})\leqslant 1$$

For the recursion step, for U_{n-1} a non empty open set, as $(A_n)^\circ = \varnothing$, we have that $U_{n-1} \nsubseteq A_n$ [for if $U_{n-1} \subseteq A_n$ then $\varnothing \neq U_{n-1} \subseteq (A_n)^\circ = \varnothing$ a contradiction], hence there exist a $y \in U_{n-1}$ such that $y \notin A_n$. Using [theorem: 14.208] there exist a $V_n \in \mathcal{T}$ such that

$$y \in V_n \Rightarrow V_n \neq \varnothing,\ \overline{V_n} \subseteq U_{n-1} \text{ and } \overline{V_n} \bigcap A_n = \varnothing$$

Take $U_n = V_n \bigcap B_d\left(y, \frac{1}{2 \cdot n}\right) \ni y$ then

$$U_n \subseteq V_n \Rightarrow \overline{U_n} \bigcap A_n \subseteq \overline{V_n} \bigcap A_n = \varnothing \text{ and } \overline{U_n} \subseteq \overline{V_0} \subseteq U_{n-1}$$

$$U_n \subseteq B_d\left(y, \frac{1}{2 \cdot n}\right) \Rightarrow \overline{U_n} \subseteq \overline{B_d\left(y, \frac{1}{2 \cdot n}\right)}$$

By [example: 14.78] we have that $\overline{B_d\left(y, \frac{1}{2 \cdot n}\right)}$ is bounded with $\operatorname{diam}\left(\overline{B_d\left(y, \frac{1}{2 \cdot n}\right)}\right) \leqslant \frac{2}{2 \cdot n} = \frac{1}{n}$, so, as $\overline{U_n} \subseteq \overline{B_d\left(y, \frac{1}{2 \cdot n}\right)}$, it follows that $\overline{U_n}$ is bounded and $\operatorname{diam}(\overline{U_n}) < \frac{1}{n}$. Summarized

$$\varnothing \neq U_n,\ \overline{U_n} \subseteq U_{n-1},\ \overline{U_n} \bigcap A_n = \varnothing,\ \overline{U_n} \text{ is bounded and } \operatorname{diam}(\overline{U_1}) \leqslant \frac{1}{n}$$

□

Let $\varepsilon \in \mathbb{R}^+$ then by [theorem: 10.30] there exist a $N \in \mathbb{N}$ such that $0 < \frac{1}{N} < \varepsilon$ so if $n \geqslant N$ we have

$$|\operatorname{diam}(\overline{U_n}) - 0| = \operatorname{diam}(\overline{U_n}) \leqslant \frac{1}{n} \leqslant \frac{1}{N} < \varepsilon$$

proving that

$$\lim_{n \to \infty} \operatorname{diam}(\overline{U_n}) = 0$$

Hence as $\forall n \in \mathbb{N}$ $\overline{U_{n+1}} \subseteq \overline{U_n}$ and $\varnothing \neq U_n \subseteq \overline{U_n}$ it follows, using [lemma: 14.383], that

$$\bigcap_{n \in \mathbb{N}} \overline{U_n} \neq \varnothing \text{ or there exist a } x \in \bigcap_{n \in \mathbb{N}} \overline{U_n}$$

As $1 \in \mathbb{N}$ we have $x \in \overline{U_1} \subseteq U_0$, further, as $\forall n \in \mathbb{N}$ $x \in \overline{U_n} \underset{\overline{U_n} \bigcap A_n = \varnothing}{\Rightarrow} x \notin A_n$, it follows that $x \notin \bigcup_{n \in \mathbb{N}} A_n$. Hence we have proved that for every non empty open set U_0 $U_0 \nsubseteq \bigcup_{i \in \mathbb{N}} A_i$ which, as $\left(\bigcup_{n \in \mathbb{N}} A_n\right)^\circ \subseteq \bigcup_{n \in \mathbb{N}} A_n$, proves that $\left(\bigcup_{n \in \mathbb{N}} A_n\right)^\circ = \varnothing$. Hence $\langle X, \mathcal{T} \rangle$ is a Baire space. □

Lemma 14.385. *Let $\langle X, \|\|_X \rangle$, $\langle Y, \|\|_Y \rangle$ be Banach spaces and $L: X \to Y$ a surjective linear function then we have:*

$$L \text{ is open} \Leftrightarrow \exists \varepsilon \in \mathbb{R}^+ \text{ such that } B_{\|\|_Y}(0, \varepsilon) \subseteq L(B_{\|\|_X}(0, 1))$$

Proof.

⇒. As L is open and $B_{\|\|_X}(0,1)$ is a open set we have that $L(B_{\|\|_X}(0,1))$ is a open set, further as $L(0) = 0$ $0 \in L(B_{\|\|_X}(0,1))$. Hence by [theorem: 14.93] there exist a $\varepsilon \in \mathbb{R}^+$ such that $B_{\|\|_Y}(0, \varepsilon) \subseteq L(B_{\|\|_X}(0,1))$.

⇐. By the hypothesis there exist a $\varepsilon \in \mathbb{R}^+$ such that

$$B_{\|\|_Y}(0, \varepsilon) \subseteq L(B_{\|\|_X}(0,1)) \tag{14.239}$$

Let $r \in \mathbb{R}^{+1}$. We use now r to scale [eq: 14.239]. Let $y \in B_{\|\|_Y}(0, r \cdot \varepsilon)$ then $\|y\|_Y < r \cdot \varepsilon$ so that $\left\|\frac{1}{r} \cdot y\right\|_Y = \frac{1}{r} \cdot \|y\|_Y < \frac{1}{r} \cdot r \cdot \varepsilon = \varepsilon$ so that $\frac{1}{r} \cdot y \in B_{\|\|_Y}(0, \varepsilon)$. Hence by [eq: 14.239] there exist a $x \in B_{\|\|_X}(0,1)$ such that $\frac{1}{r} \cdot y = L(x)$ or $y = r \cdot L(x) = L(r \cdot x)$. As $\|r \cdot x\|_X = r \cdot \|x\|_x < r \cdot 1 = r$ we have $r \cdot x \in B_{\|\|_X}(0, r)$ so that $y = L(r \cdot x) \in L(B_{\|\|_X}(0, r))$. Hence we have that

$$B_{\|\|_Y}(0, r \cdot \varepsilon) \subseteq L(B_{\|\|_X}(0, r)) \tag{14.240}$$

Let $b \in Y$. As L is a surjection there exist a $a \in X$ such that $b = L(a)$. Let $y \in B_{\|\|_Y}(b, r \cdot \varepsilon)$ then $\|y - L(a)\| = \|y - b\|_Y < r \cdot \varepsilon$, hence $y - L(a) \in B_{\|\|_Y}(0, r \cdot \varepsilon)$ so that by [eq: 14.240] we have that $y - L(a) \in L(B_{\|\|_X}(0, r))$. Hence there exist a $x \in B_{\|\|_X}(0, r)$ such that $y - L(a) = L(x)$ or $y = L(x) + L(a) = L(x + a)$. Further $\|(x + a) - a\|_X = \|x\|_X < r$ so that $x + a \in B_{\|\|_X}(a, r)$ proving that $y \in L(B_{\|\|_X}(a, r))$ or

$$B_{\|\|_Y}(b, r \cdot \varepsilon) \subseteq L(B_{\|\|_X}(a, r))$$

Summarized we have

$$\forall b \in Y,\ \forall r \in \mathbb{R}^+ \text{ we have } B_{\|\|_Y}(b, r \cdot \varepsilon) \subseteq L(B_{\|\|_X}(a, r)) \text{ where } b = L(a) \tag{14.241}$$

Let U be a open set in X. Given $y\in L(U)$ there exist a $x\in U$ such that $y=L(x)$. As U is open there exist by [theorem: 14.93] a $\delta\in\mathbb{R}^+$ such that $x\in B_{\|\|_X}(x,\delta)\subseteq U$. Take then in [eq: 14.241] $r=\delta$ and $a=x\Rightarrow b=L(a)=L(x)=y$ then we have

$$y\in B_{\|\|_Y}(y,\delta\cdot\varepsilon)\subseteq L(B_{\|\|_X}(x,\delta))\subseteq L(U)$$

proving by [theorem: 14.93] that $L(U)$ is open. Hence $L: X\rightarrow Y$ is a open mapping □

Lemma 14.386. *Let $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$ be Banach spaces and $L: X\rightarrow Y$ a surjective continuous linear function then $\exists\lambda\in\mathbb{R}^+$ such that $B_{\|\|_Y}(0,\lambda)\subseteq L(B_{\|\|_X}(0,1))$.*

Proof. First by the Baire Category theorem [see theorem: 14.384] we have that

$$\langle Y,\|\|_Y\rangle \text{ is a Baire Space}$$

Let $x\in X$ then by the Archimedean property of $\mathbb{R}$ [see theorem: 10.30] there exist a $n\in\mathbb{N}$ such that $\|x\|<n$ hence $x\in B_{\|\|_X}(0,n)\subseteq\bigcup_{n\in\mathbb{N}}B_{\|\|_X}(0,n)$. Hence $X\subseteq\bigcup_{n\in\mathbb{N}}B_{\|\|_X}(0,n)$ so that

$$Y\underset{L\text{ is surjective}}{=}L(X)\subseteq L\left(\bigcup_{n\in\mathbb{N}}B_{\|\|_X}(0,n)\right)\underset{\text{[theorem: 2.134]}}{=}\bigcup_{n\in\mathbb{N}}L(B_{\|\|_X}(0,n))\subseteq\bigcup_{n\in\mathbb{N}}\overline{L(B_{\|\|_X}(0,n))}\subseteq Y$$

proving that

$$Y=\bigcup_{n\in\mathbb{N}}\overline{L(B_{\|\|_X}(0,n))} \tag{14.242}$$

If now $\forall n\in\mathbb{N}$ $(\overline{L(B_{\|\|_X}(0,n))})^\circ=\varnothing$ then, as Y is a Baire Space, we have that $(\bigcup_{n\in\mathbb{N}}\overline{L(B_{\|\|_X}(0,n))})^\circ=\varnothing$ hence $Y^\circ=\varnothing$, as Y is open we have by [theorem: 14.15] that $Y=Y^\circ$, so that $Y=\varnothing$ contradicting $0\in Y$. Hence there exist a $n_0\in\mathbb{N}$ such that

$$(\overline{L(B_{\|\|_X}(0,n_0))})^\circ\neq\varnothing$$

Now

$$\begin{aligned}\overline{L(B_{\|\|_X}(0,n_0))}&\underset{\text{[theorem: 14.91]}}{=}\overline{L(n_0\cdot B_{\|\|_x}(0,1))}\\&\underset{\text{[theorem: 11.161]}}{=}\overline{n_0\cdot L(B_{\|\|_X}(0,1))}\\&\underset{\text{[theorem: 14.109]}}{=}n_0\cdot\overline{L(B_{\|\|_X}(0,1))}\end{aligned}$$

so that we have

$$\varnothing\neq(n_0\cdot\overline{L(B_{\|\|_X}(0,1))})^\circ$$

Hence there exist a $v_0'\in Y$ such that $v_0'\in(n_0\cdot\overline{L(B_{\|\|_X}(0,1))})^\circ$, using [theorem: 14.93] there exist a $r'\in\mathbb{R}^+$ such that $v_0'\in B_{\|\|_Y}(v_0',r')\subseteq(n_0\cdot\overline{L(B_{\|\|_X}(0,1))})^\circ\subseteq n_0\cdot\overline{L(B_{\|\|_X}(0,1))}$. Take $r=\frac{r'}{n_0}$ and $v_0=\frac{1}{n_0}\cdot v_0'$ then if $y\in B_{\|\|_Y}(v_0,r)$ we have

$$\|v_0-y\|_Y<r=\frac{r'}{n_0}$$

As

$$\frac{1}{n_0}\cdot\|v_0'-n_0\cdot y\|_Y=\left\|\frac{1}{n_0}\cdot(v_0'-n_0\cdot y)\right\|_Y=\|v_0-y\|_Y$$

it follows that $\frac{1}{n_0}\cdot\|v_0'-n_0\cdot y\|_Y<\frac{r'}{n_0}$ so that $\|v_0'-n_0\cdot y\|_Y<r'$ hence

$$n_0\cdot y\in B_{\|\|_Y}(v_0',r')\subseteq n_0\cdot\overline{L(B_{\|\|_X}(0,1))}\Rightarrow y\in\overline{L(B_{\|\|_X}(0,1))}$$

This proves that $B_{\|\|_Y}(v_0,r)\subseteq\overline{L(B_{\|\|_X}(0,1))}$ or if we take $\delta=\frac{r}{4}$ this results in

$$\exists v_0\in Y,\ \delta\in\mathbb{R}^+ \text{ such that } v_0\in B_{\|\|_Y}(v_0,4\cdot\delta)\subseteq\overline{L(B_{\|\|_X}(0,1))} \tag{14.243}$$

As $v_0\in\overline{L(B_{\|\|_X}(0,1))}$ we have by [theorem: 14.28] that $B_{\|\|_Y}(v_0,2\cdot\delta)\bigcap L(B_{\|\|_X}(0,1))\neq\varnothing$. So there exist a $v_1\in B_{\|\|_Y}(v_0,2\cdot\delta)\bigcap L(B_{\|\|_X}(0,1))$ or

$$\|v_1-v_0\|_Y<2\cdot\delta \text{ and there exist a } u_1\in B_{\|\|_X}(0,1)\Rightarrow\|u_1\|_X<1 \text{ such that } v_1=L(u_1) \tag{14.244}$$

If $x\in B_{\|\|_Y}(v_1,2\cdot\delta)$ then we have

$$\|x-v_0\|_Y\leqslant\|x-v_1\|_Y+\|v_1-v_0\|<2\cdot\delta+2\cdot\delta=4\cdot\delta$$

so that $x\in B_{\|\|_Y}(v_0,4\cdot\delta)$. Hence

$$B_{\|\|_Y}(v_1,2\cdot\delta)\subseteq B_{\|\|_Y}(v_0,4\cdot\delta)\subseteq\overline{L(B_{\|\|_X}(0,1))}\tag{14.245}$$

Let $v\in B_{\|\|_Y}(0,\delta)$ then $\|(2\cdot v+v_1)-v_1\|_Y=\|2\cdot v\|_Y=2\cdot\|v\|_Y<2\cdot\delta$ so that $2\cdot v+v_1\in B_{\|\|_Y}(v_1,2\cdot\delta)$ or

$$v+\frac{1}{2}\cdot v_1=\frac{1}{2}\cdot(2\cdot v+v_1)\in\frac{1}{2}\cdot B_{\|\|_Y}(v_1,2\cdot\delta)\subseteq_{[\text{eq: }14.245]}\frac{1}{2}\cdot\overline{L(B_{\|\|_X}(0,1))}$$

so that

$$v\in-\frac{1}{2}\cdot v_1+\frac{1}{2}\cdot\overline{L(B_{\|\|_X}(0,1))}$$

As

$$\begin{aligned}\frac{1}{2}\cdot\overline{L(B_{\|\|_X}(0,1))}&\underset{[\text{theorem: }14.109]}{=}\overline{\frac{1}{2}\cdot L(B_{\|\|_X}(0,1))}\\&\underset{[\text{theorem: }11.161]}{=}\overline{L\left(\frac{1}{2}\cdot B_{\|\|_X}(0,1)\right)}\\&\underset{[\text{theorem: }14.91]}{=}\overline{L\left(B_{\|\|_X}\left(0,\frac{1}{2}\right)\right)}\end{aligned}$$

it follows, taking account that $v_1\underset{[\text{eq: }14.244]}{=}L(u_1)$, that

$$v\in-\frac{1}{2}\cdot L(u_1)+\overline{L\left(B_{\|\|_X}\left(0,\frac{1}{2}\right)\right)}$$

Now

$$\begin{aligned}-\frac{1}{2}\cdot L(u_1)+\overline{L\left(B_{\|\|_X}\left(0,\frac{1}{2}\right)\right)}&\subseteq_{[\text{theorem: }14.109]}\overline{-\frac{1}{2}\cdot L(u_1)+L\left(B_{\|\|_X}\left(0,\frac{1}{2}\right)\right)}\\&\underset{[\text{theorem: }11.161]}{=}\overline{L\left(-\frac{1}{2}\cdot u_1\right)+L\left(B_{\|\|_X}\left(0,\frac{1}{2}\right)\right)}\\&\underset{[\text{theorem: }11.161]}{=}\overline{L\left(-\frac{1}{2}\cdot u_1+B_{\|\|_X}\left(0,\frac{1}{2}\right)\right)}\end{aligned}$$

so that

$$v\in\overline{L\left(-\frac{1}{2}\cdot u_1+B_{\|\|_X}\left(0,\frac{1}{2}\right)\right)}$$

If $x\in-\frac{1}{2}\cdot u_1-B_{\|\|_X}\left(0,\frac{1}{2}\right)$ then there exist a $y\in B_{\|\|_X}\left(0,\frac{1}{2}\right)$ such that $x=-\frac{1}{2}\cdot u_1+y$, hence

$$\|x\|_X=\left\|-\frac{1}{2}\cdot u_1\right\|_X+\|y\|_X<\left\|\frac{1}{2}\cdot u_1\right\|_X+\frac{1}{2}<_{[\text{eq: }14.244]}\frac{1}{2}+\frac{1}{2}=1$$

proving that $-\frac{1}{2}\cdot u_1-B_{\|\|_X}\left(0,\frac{1}{2}\right)\subseteq B_{\|\|_X}(0,1)$ so that we have

$$v\in\overline{L(B_{\|\|_X}(0,1))}$$

Therefore as $v\in B_{\|\|_Y}(0,\delta)$ was chosen arbitrary it follows that

$$B_{\|\|_Y}(0,\delta)\subseteq\overline{L(B_{\|\|_X}(0,1))}\tag{14.246}$$

Let $n\in\mathbb{N}$ then we have that

$$\begin{aligned}B_{\|\|_Y}\left(0,\left(\frac{1}{2}\right)^n\cdot\delta\right)&\underset{[\text{theorem: }14.91]}{=}\left(\frac{1}{2}\right)^n\cdot B_{\|\|_Y}(0,\delta)\\&\subseteq_{[\text{eq: }14.246]}\left(\frac{1}{2}\right)^n\cdot\overline{L(B_{\|\|_X}(0,1))}\\&\underset{[\text{theorem: }14.109]}{=}\overline{\left(\frac{1}{2}\right)^n\cdot L(B_{\|\|_X}(0,1))}\\&\underset{[\text{theorem: }14.91]}{=}\overline{L\left(B_{\|\|_X}\left(0,\left(\frac{1}{2}\right)^n\right)\right)}\end{aligned}$$

proving that

$$B_{\|\|_Y}\left(0,\left(\frac{1}{2}\right)^n\cdot\delta\right)\subseteq\overline{L\left(B_{\|\|_X}\left(0,\left(\frac{1}{2}\right)^n\right)\right)}\tag{14.247}$$

Let $y\in B_{\|\|_Y}\left(0,\frac{\delta}{2}\right)$. Then using recursion we construct a sequence $\{x_n\}_{n\in\mathbb{N}}\subseteq X$ such that

$$\forall n\in\mathbb{N}\text{ we have }\left\|y-\sum_{i=1}^n L(x_i)\right\|_Y<\left(\frac{1}{2}\right)^{n+1}\cdot\delta\text{ and }\|x_n\|_X<\left(\frac{1}{2}\right)^n\tag{14.248}$$

Proof.

Case 1. As $y\in B_{\|\|_Y}\left(0,\frac{\delta}{2}\right)=B_{\|\|_Y}\left(0,\left(\frac{\delta}{2}\right)^1\right)$ we have by [eq: 14.247] that

$$y\in\overline{L\left(B_{\|\|_X}\left(0,\left(\frac{1}{2}\right)^1\right)\right)}=\overline{L\left(B_{\|\|_X}\left(0,\frac{1}{2}\right)\right)}.$$

Hence as $y\in B_{\|\|_Y}\left(y,\frac{\delta}{4}\right)\in\mathcal{T}_{\|\|_Y}$ we have by [theorem: 14.28] that

$$B_{\|\|_Y}\left(y,\frac{\delta}{4}\right)\bigcap L\left(B_{\|\|_X}\left(0,\frac{1}{2}\right)\right)\neq\varnothing$$

hence there exist a $x_1\in B_{\|\|_X}(0,\frac{1}{2})$ such that $\|y-L(x_1)\|_Y<\frac{\delta}{4}$ or

$$\left\|y-\sum_{i=1}^1 L(x_i)\right\|_Y<\left(\frac{1}{2}\right)^{1+1}\cdot\delta\text{ and }\|x_1\|_X<\left(\frac{1}{2}\right)^1$$

Case n+1. As $\|y-\sum_{i=1}^n L(x_i)\|_Y<(\frac{1}{2})^{n+1}\cdot\delta$ we have $y-\sum_{i=1}^n L(x_i)\in B_{\|\|_Y}(0,(\frac{1}{2})^{n+1}\cdot\delta)$ hence using [eq: 14.247] we have that $y-\sum_{i=1}^n L(x_i)\in\overline{L(B_{\|\|_X}(0,(\frac{1}{2})^{n+1}))}$. As

$$y-\sum_{i=1}^n L(x_i)\in B_{\|\|_Y}\left(y-\sum_{i=1}^n L(x_i),\left(\frac{1}{2}\right)^{(n+1)+1}\cdot\delta\right)\in\mathcal{T}_{\|\|_Y}$$

it follows from [theorem: 14.28] that

$$B_{\|\|_Y}\left(y-\sum_{i=1}^n L(x_i),\left(\frac{1}{2}\right)^{(n+1)+1}\cdot\delta\right)\bigcap\overline{L\left(B_{\|\|_X}\left(0,\left(\frac{1}{2}\right)^{n+1}\right)\right)}\neq\varnothing$$

Hence there exist a $x_{n+1}\in B_{\|\|_X}(0,(\frac{1}{2})^{n+1})$ such that

$$\left\|y-\sum_{i=1}^n L(x_i)-L(x_{n+1})\right\|_Y<\left(\frac{1}{2}\right)^{(n+1)+1}\cdot\delta$$

or

$$\left\|y-\sum_{i=1}^{n+1} L(x_i)\right\|_Y<\left(\frac{1}{2}\right)^{(n+1)+1}\cdot\delta\text{ and }\|x_{n+1}\|_X<\left(\frac{1}{2}\right)^{n+1}$$

□

As $\forall n\in\mathbb{N}$ $\|x_n\|_X<(\frac{1}{2})^n$ [see eq: 14.248], $\sum_{i=1}^\infty(\frac{1}{2})^n\underset{\text{[theorem: 14.348]}}{=}\frac{1/2}{1-1/2}=1$ [see theorem: 14.348], $\langle X,\|\|_X\rangle$ is a Banach space it follows from the dominant convergence [theorem: 14.373] that

$$x=\sum_{i=1}^\infty x_i\text{ exist and }\|x\|<1\text{ so that }x\in B_{\|\|_X}(0,1)\tag{14.249}$$

Further if $\varepsilon\in\mathbb{R}^+$ then by [theorem: 10.63] there exist a $N\in\mathbb{N}$ such that $(\frac{1}{2})^N<\frac{\varepsilon}{\delta}$. Let $n\geqslant N$ then we have

$$\left\|\sum_{i=1}^n L(x_i)-y\right\|_Y<\left(\frac{1}{2}\right)^{n+1}\cdot\delta<_{\text{[theorem: 10.61]}}\left(\frac{1}{2}\right)^N\cdot\delta<\frac{\varepsilon}{\delta}\cdot\delta=\varepsilon$$

proving that

$$\sum_{i=1}^\infty L(x_i)=y\tag{14.250}$$

Now as L is continuous we have using [theorem: 14.307] that

$$\begin{aligned} L(x) &\underset{\text{[eq: 14.249]}}{=} L\left(\sum_{i=1}^{\infty} x_i\right) \\ &= L\left(\lim_{n\rightarrow\infty}\sum_{i=1}^{n} x_i\right) \\ &\underset{\text{[theorem: 14.307]}}{=} \lim_{n\rightarrow\infty} L\left(\sum_{i=1}^{n} x_i\right) \\ &= \lim_{n\rightarrow\infty}\sum_{i=1}^{n} L(x_i) \\ &= \sum_{i=1}^{\infty} L(x_i) \\ &\underset{\text{[eq: 14.250]}}{=} y \end{aligned}$$

Which, as $x\in B_{\|\|_X}(0,1)$ [see eq: 14.249], proves that $y\in L(B_{\|\|_X}(0,1))$, hence as $y\in B_{\|\|_Y}\left(0,\frac{\delta}{2}\right)$ was chosen arbitrary this proves that $B_{\|\|_Y}\left(0,\frac{\delta}{2}\right)\subseteq L(B_{\|\|_X}(0,1))$ or if we take $\lambda=\frac{\delta}{2}$ that

$$B_{\|\|_Y}(0,\lambda)\subseteq L(B_{\|\|_X}(0,1))$$

□

We are now finally ready to proof the Open Mapping Theorem stating that every continuous linear surjective mapping between two Banach spaces is open.

Theorem 14.387. (Open Mapping Theorem) *Let $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$ be Banach spaces and $L: X\rightarrow Y$ a surjective continuous linear function then $L: X\rightarrow Y$ is a **open mapping**.*

Proof. Using [lemma: 14.386] there exist a $\lambda\in\mathbb{R}^+$ such that

$$B_{\|\|_Y}(0,\lambda)\subseteq L(B_{\|\|_X}(0,1))$$

Applying then [lemma: 14.385] proves that L is open. □

Corollary 14.388. *Let $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$ be Banach spaces and $L: X\rightarrow Y$ a continuous linear isomorphism then $L: X\rightarrow Y$ is a homeomorphism.*

Proof. As $L: X\rightarrow Y$ is a isomorphism it is surjective. Hence as L is continuous and linear if follows from [theorem: 14.387] that $L: X\rightarrow Y$ is a open mapping. Hence if U is a open set in X we have for $L^{-1}: Y\rightarrow X$ that $(L^{-1})^{-1}(U) \underset{\text{[theorem: 2.68]}}{=} L(U)$ is open. So that $L^{-1}: Y\rightarrow X$ is continuous which by the continuity of $L: X\rightarrow Y$ proves that $L: X\rightarrow Y$ is a homeomorphism. □

14.11 Connected Sets

Definition 14.389. *A topological space $\langle X,\mathcal{T}\rangle$ is **connected** if $\forall U_1,U_2\in\mathcal{T}$ with $U_1\bigcap U_2=\varnothing$ and $X=U_1\bigcup U_2$ we have that either $U_1=\varnothing$ or $U_2=\varnothing$. In other words a topological space is connected if it is not the union of two disconnected non empty sets.*

Definition 14.390. *A topological space is **disconnected** if it is not connected or in other words there exists $U_1,U_2\in\mathcal{T}$ with $U_1\neq\varnothing$, $U_2\neq\varnothing$, $U_1\bigcap U_2=\varnothing$ and $X=U_1\bigcup U_2$.*

Lemma 14.391. *Let A,B,C be classes such that $A\bigcap B=\varnothing$ and $C=A\bigcup B$ then we have:*

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. If $x\in A$ then, as $A\bigcap B=\emptyset$, we must have $x\notin B$ so that $x\in A\setminus B$ hence $A\subseteq A\setminus B\subseteq A$ so that $A\setminus B=A$. □

Theorem 14.392. *Let $\langle X,\mathcal{T}\rangle$ be a topological space then the following are equivalent:*

1. *$\varnothing$ and X are the only subsets of X that are open and closed.*
2. *X is connected*
3. *For every pair of closed sets A_1, A_2 with $A_1\bigcap A_2=\varnothing$ and $X=A_1\bigcup A_2$ we have either $A_1=\varnothing$ or $A_2=\varnothing$.*

Proof.

1 ⇒ 2. Assume that U_1, U_2 are **disjoint open** sets with $X=U_1\bigcup U_2$ then by [lemma: 14.391] $U_2\underset{\text{[lemma: 14.391]}}{=}X\setminus U_1$ so that U_2 is open and closed. According to (1) we must then have that either:

$U_2=\varnothing$. Then $U_1=\varnothing\vee U_2=\varnothing$

$U_2=X$. Then $U_1\underset{\text{[lemma: 14.391]}}{=}X\setminus U_2=X\setminus X=\varnothing$ so that $U_1=\varnothing\vee U_2=\varnothing$

2 ⇒ 3. Assume that we have two **disjoint closed** sets A_1, A_2 with $X=A_1\bigcup A_2$ then $A_1\underset{\text{[lemma: 14.391]}}{=}X\setminus A_2$ is open and $A_2\underset{\text{[lemma: 14.391]}}{=}X\setminus A_1$ is open, so by the connectnes of X we must have either $A_1=\emptyset$ or $A_2=\emptyset$.

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 14.393. *Let $\langle X,\mathcal{T}\rangle$ be a topological space and $A\subseteq X$ then A is connected if $\langle A,\mathcal{T}_{|A}\rangle$ is connected where $\mathcal{T}_{|A}=\{U\bigcap A|U\in\mathcal{T}\}$ is the subspace topology on A.*

Theorem 14.394. *Let $\langle X,\mathcal{T}_X\rangle$ be connected topological space, $\langle Y,\mathcal{T}_Y\rangle$ a topological space and*

$$f: X\rightarrow Y \text{ a continuous function}$$

then

$$f(X) \text{ is a connected subset of } Y$$

Proof. We proves this by contradiction. So assume that $f(X)$ is not connected then there exist open sets $V_1, V_2\in\mathcal{T}_Y$ with $V_1\bigcap f(X)\neq\varnothing$, $V_2\bigcap f(X)\neq\varnothing$, $(V\bigcap f(X))\bigcap(V\bigcap f(X))=\varnothing$ and $(V_1\bigcap f(X))\bigcup(V_2\bigcap f(X))=f(X)$. As f is continuous we have by [theorem: 14.130] that

$$U_1=f^{-1}(V_1)\in\mathcal{T}_X \text{ and } U_2=f^{-1}(V_2)\in\mathcal{T}_Y \tag{14.251}$$

As $V_1\bigcap f(X)$, $V_2\bigcap f(X)\neq\varnothing$ there exists a $y_1\in V_1\bigcap f(X)$ and a $y_2\in V_2\bigcap f(X)$. Hence there exist $x_1, x_2\in X$ such that $f(x_1)=y_1\in V_1$ and $f(x_2)=y_2\in V_2$ so that $x_1\in f^{-1}(V_1)=U_1$ and $x_2\in f^{-1}(V_2)=U_2$, hence

$$U_1\neq\varnothing \text{ and } U_2\neq\varnothing \tag{14.252}$$

Let $x\in X$ then $f(x)\in f(X)=(V_1\bigcap f(X))\bigcup(V_2\bigcap f(X))=V_1\bigcup V_2$ so that

$$x\in f^{-1}\left(V_1\bigcup V_2\right)\underset{\text{[theorem: 2.93]}}{=}f^{-1}(V_1)\bigcup f^{-1}(V_2)=U_1\bigcup U_2\subseteq X$$

proving

$$X=U_1\bigcup U_2 \tag{14.253}$$

Further if $x\in U_1\bigcap U_2=f^{-1}(V_1)\bigcap f^{-1}(V_2)$ then $f(x)\in V_1\bigcap V_2$ which as $f(x)\in f(X)$ results in $f(x)\in(V_1\bigcap f(X))\bigcap(V_2\bigcap f(X))=\varnothing$ a contradiction hence we must have that

$$U_1\bigcap U_2=\varnothing \tag{14.254}$$

Combining [eqs: 14.251, 14.252 and 14.254] we conclude that X is not connected contradicting the fact that X was assumed to be connected. Hence the assumption that $f(X)$ is not connected is wrong and we must have that $f(X)$ is connected. □

A very important example of a connected subset of $\mathbb{R}$ is the closed interval (later we will show that every generalized interval is connected).

Lemma 14.395. *Let $\langle\mathbb{R},\|\rangle$ be the normed space of real numbers then $\forall a,b\in\mathbb{R}$ with $a\leqslant b$ we have that $[a,b]$ is connected.*

Proof. Note that $[a,b]=\left[\frac{a+b}{2}-\frac{b-a}{2},\frac{a+b}{2}+\frac{b-x}{2}\right]=\overline{B\left(\frac{a+b}{2},\frac{a-b}{2}\right)}$ which is a closed set in $\mathcal{T}_{||}$ by [theorem: 14.65]. Let $A\subseteq[a,b]$ be a closed set in the subspace topology $(\mathcal{T}_{||})_{|[a,b]}$ then there exist by [theorem: 14.17] a closed set A' in $\mathcal{T}_{||}$ such that $A=A'\bigcap[a,b]$ which is then closed in $\mathcal{T}_{||}$. Hence it follows that

$$\text{every closed set in } (\mathcal{T}_{||})_{|[a,b]} \text{ is closed in } \mathcal{T}_{||} \tag{14.255}$$

We proceed now by contradiction. So assume that $[a,b]$ is not connected. Then by [theorem: 14.392] there exists sets $A_1,A_2\subseteq[a,b]$ such that A_1,A_2 are closed in $(\mathcal{T}_{||})_{|[a,b]}$, hence also closed in $\mathcal{T}_{||}$, $A_1\neq\varnothing$, $A_2\neq 0$, $A_1\bigcap A_2=\varnothing$ and $[a,b]=A_1\bigcup A_2$. We can always assume that $b\in A_2$ [otherwise switch A_1 and A_2]. As $A_1\subseteq A_1\bigcup A_2=[a,b]$ it follows that A_1 is bounded above by b which together with $A_1\neq\varnothing$ allows us to use the conditional completeness of $\mathbb{R}$ [see theorem: 10.18] proving that

$$c=\sup(A_1) \text{ exist}$$

Let $U\in\mathcal{T}_{||}$ such that $c\in U$ then there exist a $\varepsilon\in\mathbb{R}^+$ with $c\in]c-\varepsilon,c+\varepsilon[=B_{||}(c,\varepsilon)\subseteq U$. As $c-\varepsilon<c=\sup(A_1)$ there exist a $c'\in A_1$ such that $c-\varepsilon<c'\leqslant c<c+\varepsilon$ so that

$$c'\in]c-\varepsilon,c+\varepsilon[\bigcap A_1=B_{||}(c,\varepsilon)\bigcap A_1\subseteq U\bigcap A_1\Rightarrow U\bigcap A_1\neq\varnothing \tag{14.256}$$

proving by [theorem: 14.29] and the fact that A_1 is closed that

$$c\in A_1 \tag{14.257}$$

As $c\in A_1\subseteq[a,b]$ we have $a\leqslant c<b$ or

$$]c,b]\subseteq[a,b] \tag{14.258}$$

Assume that $x\in]c,b]\bigcap A_1$ then $c<x$ and $x\leqslant\sup(A_1)=c$ giving the contradiction $x<c$, hence $]c,b]\bigcap A_1=\varnothing$. Combining this with [eq: 14.258] results in

$$]c,b]\subseteq[a,b]\setminus A_1=(A_1\bigcup A_2)\setminus A_1\underset{[\text{lemma: } 14.391]}{=}A_2$$

As $c\in A_1$, $b\in A_2$, $A_1\bigcap A_2=\varnothing$ we have that $c\neq b$ which as b is a upper bound of A_1 proves

$$c<b\Rightarrow 0<b-c$$

Let $U\in\mathcal{T}_{|||}$ such that $c\in U$ then there exist a $\varepsilon\in\mathbb{R}^+$ such that $c\in]c-\varepsilon,c+\varepsilon[\subseteq U$. Take $\delta=\min(b-c,\varepsilon)>0$ then we have $c<c+\frac{\delta}{2}<c+\delta\leqslant c+b-c=b$ so that $c+\frac{\delta}{2}\in]c,b]\subseteq A_2$ and as $c-\varepsilon<c<c+\frac{\delta}{2}\leqslant c+\frac{\varepsilon}{2}<c+\varepsilon$ we have that $c+\frac{\delta}{2}\in]c-\varepsilon,c+\varepsilon[\bigcap A_2\subseteq U\bigcap A_2$ proving that $U\bigcap A_2\neq\varnothing$. Using [theorem: 14.29] it follows that $c\in A_2$ which combined with [eq: 14.257] gives $c\in A_1\bigcap A_2$ contradicting $A_1\bigcap A_2\underset{[\text{eq: } 14.256]}{=}\varnothing$. Hence the assumption that $[a,b]$ is not connected is wrong so that we must have that $[a,b]$ is connected. □

Actually it turns out that the only connected sets in $[a,b]$ are generalized intervals [see definition: 3.133 and theorem: 3.136].

Theorem 14.396. *Let $\langle\mathbb{R},\|\rangle$ be the normed spaces of real numbers and $I\subseteq\mathbb{R}$ then we have*

$$I \text{ is connected} \Leftrightarrow I \text{ is a generalized interval}$$

Proof.

⇒. We prove this by contradiction. So assume that I is a connected set but not a generalized interval. As $\langle\mathbb{R},\leqslant\rangle$ is conditional complete we have by [theorem: 3.134] that there exists $x,y\in I$ such that $]x,y[\nsubseteq I$. Hence there exist a $z\in]x,y[$ such that $z\in\mathbb{R}\setminus I$. Now as $I\subseteq\mathbb{R}=]-\infty,z[\bigcup\{z\}\bigcup]z,\infty[$ and $z\notin I$ it follows that $I\subseteq]-\infty,z[\bigcup]z,\infty[$ so that

$$I\subseteq(]-\infty,z[\bigcap I)\bigcup(]z,\infty[\bigcap I) \tag{14.259}$$

We have also

$$(]-\infty,z[\bigcap I)\bigcap(]z,\infty[\bigcap I)=I\bigcap(]-\infty,z[\bigcap]z,\infty[)=I\bigcap\varnothing=\varnothing \tag{14.260}$$

As $x,y\in I$ and $z\in]x,y[\Rightarrow x<z<y\Rightarrow x\in]-\infty,z[\bigcap I\wedge y\in]z,\infty[\bigcap I$ it follows that

$$]-\infty,z[\bigcap I\neq\varnothing \text{ and }]z,\infty[\bigcap I\neq\varnothing \tag{14.261}$$

Finally using [theorem: 14.99] we have that $]-\infty,z[\in\mathcal{T}_{||}$ and $]z,\infty[\in\mathcal{T}_{||}$ so that

$$]-\infty,z[\bigcap I\in(\mathcal{T}_{||})_{|I} \text{ and }]z,\infty[\bigcap I\in(\mathcal{T}_{||})_{|I} \tag{14.262}$$

Hence using [eqs: 14.259, 14.260, 14.261 and 14.262] it follows that I is not connected contradicting the hypothesis that I is connected. So our assumption that I is not a generalized interval is wrong, proving that is indeed a generalized interval.

$\Leftarrow$. We prove this also by contradiction. So assume that I is a generalized interval that is not connected. Then there exists open sets $U_1,U_2\in\mathcal{T}_{||}$ such that $(U_1\bigcap I)\neq\varnothing=(U_2\bigcap I)$, $I=(U_1\bigcap I)\bigcup(U_2\bigcap I)$ and $(U_1\bigcap I)\bigcap(U_2\bigcap I)=\varnothing$. Then there exists $x\in U_1\bigcap I$ and a $y\in U_2\bigcap I$. As $(U_1\bigcap I)\bigcap(U_2\bigcap I)=\varnothing$ it follows that $x\neq y$ and we can always assume that $x<y$ [if $y<x$ interchange x and y]. As I is a generalized interval we have by [theorem: 3.133] that $[x,y]\subseteq I$. So that

$$\begin{aligned}
[x,y] &= [x,y]\bigcap I\\
&= [x,y]\bigcap((U_1\bigcap I)\bigcup(U_2\bigcap I))\\
&= ([x,y]\bigcap U_1\bigcap I)\bigcup([x,y]\bigcap U_2\bigcap I)\\
&= ([x,y]\bigcap U_1)\bigcup([x,y]\bigcap U_2)
\end{aligned}$$

or

$$[x,y]=([x,y]\bigcap U1)\bigcup([x,y]\bigcap U_2) \tag{14.263}$$

Further as $[x,y]\subseteq I$ we have $([x,y]\bigcap U1)\bigcap([x,y]\bigcap U_2)\subseteq(I\bigcap U_1)\bigcap(I\bigcap U_2)=\varnothing$ proving that

$$([x,y]\bigcap U1)\bigcap([x,y]\bigcap U_2)=\varnothing \tag{14.264}$$

Further as $x\in U_1\bigcap I\Rightarrow x\in U_1 \underset{x\in[x,y]}{\Rightarrow} x\in U_1\bigcap[x,y]$ and $y\in U_2\bigcap I\Rightarrow y\in U_2 \underset{y\in[x,y]}{\Rightarrow} y\in U_2\bigcap[x,y]$ we have

$$U_1\bigcap[x,y]\neq\varnothing \text{ and } U_2\bigcap[x,y]\neq\varnothing \tag{14.265}$$

Finally by definition we have that

$$U_1\bigcap[x,y]\in(\mathcal{T}_{||})_{|[x,y]} \text{ and } U_2\bigcap[x,y]\in(\mathcal{T}_{||})_{|[x,y]} \tag{14.266}$$

From [eqs: 14.263, 14.264, 14.265 and [eq: 14.266] it follows that $[x,y]$ is not connected contradicting the fact that by [lemma: 14.395] $[x,y]$ is connected. Hence the assumption that I is not connected is wrong so that I must be connected. $\square$

A important consequence of the fact that generalized intervals in $\mathbb{R}$ are connected is the Intermediate Value Theorem. First we need a little lemma proving that the image of a closed interval of a continuous real function on the interval is again a closed interval.

Lemma 14.397. *Let $\langle\mathbb{R},||\rangle$ be the normed spaces of real numbers, $a,b\in\mathbb{R}$ with $a\leqslant b$ and*

$$f:[a,b]\rightarrow\mathbb{R} \textit{ a continuous function}$$

then there exist $c,d\in\mathbb{R}$ such that

$$f([a,b])=[c,d]$$

So there exists $x,y\in[a,b]$ such that $c=f(x)$ and $d=f(y)$ and $\forall z\in[a,b]$ $f(x)\leqslant f(z)\leqslant f(y)$

Proof. Using Heine Borell [theorem: 14.232] it follows that $[a,b]$ is a compact set. As $f:[a,b]\rightarrow\mathbb{R}$ is continuous it follows from [theorem: 14.229] that

$$f([a,b]) \text{ is compact}$$

As compact sets in $\mathbb{R}$ are closed and bounded [see theorem: 14.233] it follows that

$$f([a,b]) \text{ is closed and bounded}$$

As $[a,b]$ is connected [see lemma: 14.395] we have, using [theorem: 14.394] by the continuity of f that

$$f([a,b]) \text{ is connected}$$

Hence using the previous theorem [theorem: 14.396] it follows that

$$f([a,b]) \text{ is a generalized interval}$$

As $f([a, b])$ is closed and not empty it follows from [theorem: 14.101] that

$$f([a, b]) \text{ is either } [c, d],\ ]-\infty, c] \text{ or } [c, \infty[$$

By [example: 14.107] $]-\infty, c]$ and $[c, \infty[$ are unbounded while $[c, d]$ is bounded, hence we have that

$$f([a, b]) = [c, d]$$

□

Theorem 14.398. (Intermediate Value Theorem) *Let $\langle \mathbb{R}, \| \rangle$ be the normed space of real numbers, $a, b \in \mathbb{R}$ with $a \leqslant b$ and $f: [a, b] \to \mathbb{R}$ a continuous function then we have:*

1. *If $y \in \mathbb{R}$ satisfies $f(a) < y < f(b)$ then there exist a $x \in [a, b]$ such that $y = f(x)$.*
2. *If $y \in \mathbb{R}$ satisfies $f(b) < y < f(a)$ then there exist a $x \in [a, b]$ such that $y = f(x)$.*

Proof. Using [lemma: 14.397] we have that $f([a, b]) = [c, d]$, hence $f(a), f(b) \in [c, d]$. Let $y \in \mathbb{R}$ then we have:

$\boldsymbol{f(a) < y < f(b)}$**.** Then $y \in [f(a), f(b)]$ and $f(a) < f(b)$ where $f(a), f(b) \in [c, d]$. As $[c, d]$ is a generalized interval it follows from [theorem: 3.133] that $[f(a), f(b)] \subseteq [c, d] = f([a, b])$. Hence $y \in f([a, b])$ so there exist a $x \in [a, b]$ such that $y = f(x)$.

$\boldsymbol{f(b) < y < f(a)}$**.** Then $y \in [f(b), f(a)]$ and $f(b) < f(a)$ where $f(a), f(b) \in [c, d]$. As $[c, d]$ is a generalized interval it follows from [theorem: 3.133] that $[f(b), f(a)] \subseteq [c, d] = f([a, b])$. Hence $y \in f([a, b])$ so there exist a $x \in [a, b]$ such that $y = f(x)$. □

Chapter 15

Integration in Banach spaces

There are many ways to define integrals of functions, Riemann integrals, Lebesgue Integrals, Gauge Integrats, etc ... In this chapter we define a extension of the classic Riemann Darboux integral to functions whose range is a Banach space. Lebesque and Gauge Integrals (like Henstock integrals) are more general however to prove later the Fundamental Theorem of Calculus later we must define integrals before we introduce the concept of differentation. For this reason we define Riemann integrals here and more general forms of integration in later chapters.

First we introduce the concept of partitioning a closed interval in sub intervals with no gaps. The way to do this is by specifying a ordered family of boundary point for the sub intervals. For example the interval $[1,12]$ can be partitioned in the following way:

$$[1,12]=[1,4]\bigcup[4,8]\bigcup[8,12]$$

by the family $\{t_i\}_{i\in\{1,\ldots,4\}}$ where

$$\begin{aligned} t_1 &= 1\\ t_2 &= 4\\ t_3 &= 8\\ t_4 &= 12\end{aligned}$$

We use underscores to show the index of the boundary points of the sub intervals. Hence we use the following notation.

$$[1,12]=[1_1,4_2]\bigcup[4_2,8_3]\bigcup[8_3,12_4]$$

This is the idea of the following definition.

Definition 15.1. *Let $a,b\in\mathbb{R}$ with $a<b$ then a family $\{t_i\}_{i\in\{1,\ldots,n\}}\subseteq\mathbb{R}$ is a **partition of $[a,b]$** if $n\in\mathbb{N}\setminus\{1\}$, $t_1=a$, $t_n=b$ and $\forall i\in\{1,\ldots,n-1\}$ we have $t_i<t_{i+1}$.*

Note 15.2. The strict inequality $a<b$ is needed for if $a=b$ we would have the contradiction $a=t_1<t_2=b$.

Lemma 15.3. *If $n\in\mathbb{N}\setminus\{1\}$ and $\{t_i\}_{i\in\{1,\ldots,n\}}\subseteq\mathbb{R}$ such that $\forall i\in\{1,\ldots,n-1\}$ $t_i<t_{i+1}$ then we have*

1. *$\forall i,j\in\{1,\ldots,n\}$ with $i<j$ that $t_i<t_j$*
2. $\min(\{t_i|i\in\{1,\ldots,n\}\})=t_1$
3. $\max(\{t_i|i\in\{1,\ldots,n\}\})=t_n$

Proof.

1. We use induction to prove this, so given $i\in\{1,\ldots,n-1\}$ define

   $$S_{n,i}=\{k\in\mathbb{N}|\text{If } i+k\leqslant n \text{ then } t_i<t_{i+k}\}$$

   then we have:

   $1\in S_{n,i}$. If $i+1\leqslant n$ then $i\in\{1,\ldots,n-1\}$ and by definition $i_i<t_{i+1}$ so that

   $k\in S_{n,i}\Rightarrow k+1\in S_{n,i}$. If $i+(k+1)\leqslant n$ then $i+k\leqslant n-1$ so that $i+k\in\{1,\ldots,n-1\}$, hence $t_{i+k}<i_{i+(k+1)}$ by definition. As $k\in S_{n,i}$ and $i+k\leqslant i+(k+1)\leqslant n$ we have $t_i<t_{i+k}$ so that $t_i<t_{i+(k+1)}$. Proving that $k+1\in S_{n,i}$.

   Using induction we have that $\mathcal{S}_{n,i}=\mathbb{N}$. So if $i,j\in\{1,\ldots,n\}$ with $i<j$ then $k=j-i\in\mathbb{N}\in\mathcal{S}_{n,i}$ and $i+k=j\leqslant n$ so that $t_i<t_{i+k}=t_j$.
2. If $x\in\{t_i|i\in\{1,\ldots,n\}\}$ then there exist a $i\in\{1,\ldots,n\}$ such that $x=t_i$. As $1\leqslant i$ it follows then from (1) that $t_1\leqslant x$ which, as $t_1\in\{t_i|i\in\{1,\ldots,n\}\}$, proves that $t_1=\min(\{t_i|i\in\{1,\ldots,n\}\})$.
3. If $x\in\{t_i|i\in\{1,\ldots,n\}\}$ then there exist a $i\in\{1,\ldots,n\}$ such that $x=t_i$. As $i\leqslant n$ it follows then from (1) that $x\leqslant t_n$ which, as $t_n\in\{t_i|i\in\{1,\ldots,n\}\}$, proves that $t_n=\max(\{t_i|i\in\{1,\ldots,n\}\})$. □

Theorem 15.4. *Let $a, b \in \mathbb{R}$ with $a < b$ and $\{t_i\}_{i \in \{1,\ldots,n\}} \subseteq \mathbb{R}$ a partition of $[a, b]$ then we have:*

1. $\forall i, j \in \{1, \ldots, n\}$ *with* $i < j$ *we have* $t_i < t_j$
2. $\forall i \in \{1, \ldots, n\}$ *we have* $t_i \in [a, b]$
3. *If* $t_i = t_j$ *then* $i = j$
4. *If* $i \in \{1, \ldots, n-1\}$ *and* $k \in \{1, \ldots, n\}$ *such that* $t_k \in [t_i, t_{i+1}]$ *then* $k = i$ *or* $k = i+1$
5. $\forall i \in \{1, \ldots, n-1\}$ *we have* $\forall k \in \{1, \ldots, n\}$ *we have* $t_k \notin]t_i, t_{i+1}[$

Proof.

1. This follows from [lemma: 15.3].
2. If $i \in \{1, \ldots, n\}$ then we have either:

   $i = 1$. Then we have $t_i = a \in [a, b]$

   $i = n$. Then we have $t_i = b \in [a, b]$

   $i \in \{2, \ldots, n-1\}$. Then we have by (1) $a = t_1 < t_i < t_n = b$ so that $t_i \in [a, b]$
3. If $t_i = t_j$ then if $i \neq j$ we have either $i < j$ or $j < i$ we have by (1) that $t_i < t_j$ or $t_j < t_i$ contradicting $t_i = t_k$. So we must have $i = j$.
4. 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 $k = i+1$.
5. Assume that $t_k \in]t_i, t_{i+1}[$ then $t_i < t_k < t_{i+1}$. For k we have either:

   $k \leqslant i$. Then by (1) we have $t_k \leqslant t_i$ contradicting $t_i < t_k$.

   $i < k$. Then $i+1 \leqslant k$ so that by (1) we have $t_{i+1} \leqslant t_k$ contradicting $t_k < t_{i+1}$. □

The reason for the name partition of a interval is that is can be used to divide the interval in a union of sub intervals.

Theorem 15.5. *Let $a, b \in \mathbb{R}$ with $a < b$ and $\{t_i\}_{i \in \{1,\ldots,n\}} \subseteq \mathbb{R}$ a partition of $[a, b]$ then we have that*

1. $[a, b[= \bigcup_{i \in \{1,\ldots,n-1\}} [t_i, t_{i+1}[$ *and* $\forall i, j \in \{1, \ldots, n-1\}$ *with* $i \neq j$ *we have* $[t_i, t_{i+1}[\bigcap [t_j, t_{j+1}[= \varnothing$
2. $[a, b] = \bigcup_{i \in \{1,\ldots,n-1\}} [t_i, t_{i+1}]$

Proof.

1. Let $i \in \{1, \ldots, n-1\}$ then by [theorem: 15.4] we have $t_i \in [a, b]$ so that $a \leqslant t_i < t_{i+1} \leqslant t_n = b$. If $x \in [t_i, t_{i+1}[$ then $t_i \leqslant x < t_{i+1}$ so that $a \leqslant x < b$ or $x \in [a, b[$. Hence $[t_i, t_{i+1}[\subseteq [a, b[$ so that

$$\bigcup_{i \in \{1,\ldots,n-1\}} [t_i, t_{i+1}[\subseteq [a, b[\tag{15.1}$$

   For the opposite inequality, let $x \in [a, b[$ then $a \leqslant x < b$. Define $M_x = \{i \in \{1, \ldots, n-1\} | t_i \leqslant x\}$ then, as $M_x \subseteq \{1, \ldots, n-1\}$ M_x is finite, further as $t_1 = a \leqslant x$ we have $1 \in M_x$ so that $M_x \neq \varnothing$. Hence $m = \max(M_x)$ exist. As $m \in M_x$ we have that $t_m \leqslant x$. Now for m we have either:

   $m = n-1$. Then $m+1 = n$ and as $x < b$ we have $t_m \leqslant x < b = t_{m+1}$ so that

$$x \in [t_m, t_{m+1}[\subseteq \bigcup_{i \in \{1,\ldots,n-1\}} [t_i, t_{i+1}[$$

   $m < n-1$. If $t_{m+1} \leqslant x$ then, as $m+1 \in \{1,\ldots,n-1\}$ we have $m+1 \in M_x$ so that $m \leqslant \max(M_x) = m$ leading to the contradiction $m < m$. Hence we must have $x < t_{m+1}$ so that $t_m \leqslant x < t_{m+1}$ or

$$x \in [t_m, t_{m+1}[\subseteq \bigcup_{i \in \{1,\ldots,n-1\}} [t_i, t_{i+1}[$$

   So we have that $[a, b[\subseteq \bigcup_{i \in \{1,\ldots,n-1\}} [t_i, t_{i+1}[$ which combined with [eq: 15.1] proves

$$[a, b] = \bigcup_{i \in \{1,\ldots,n-1\}} [t_i, t_{i+1}[$$

Let $i,j\in\{1,\ldots,n-1\}$ with $i\neq j$ then we my assume that without loosing generality that $i<j$ [if not exchange i,j]. Let $x\in[t_i,t_{i+1}[\bigcap[t_j,t_{j+1}[$ then $t_i\leqslant x<t_{i+1}$ and $t_j\leqslant x<t_{j+1}$. As $i<j\Rightarrow t+1\leqslant j$ so that $t_{i+1}\leqslant t_j$, hence $x<t_{i+1}\leqslant t_j\leqslant x$ giving the contradiction $x<x$. So we must have that

$$[t_i,t_{i+1}[\bigcap[t_i,t_{j+1}[=\varnothing$$

2. Let $i\in\{1,\ldots,n-1\}$ then by [theorem: 15.4] we have $t_i\in[a,b]$ so that $a\leqslant t_i<t_{i+1}\leqslant t_n=b$. If $x\in[t_i,t_{i+1}]$ then $t_i\leqslant x\leqslant t_{i+1}$ so that $a\leqslant x\leqslant b$ or $x\in[a,b]$. Hence $[t_i,t_{i+1}]\subseteq[a,b]$ so that

$$\bigcup_{i\in\{1,\ldots,n-1\}}[t_i,t_{i+1}]\subseteq[a,b] \tag{15.2}$$

For the opposite inequality, let $x\in[a,b]$ then $a\leqslant x\leqslant b$. Define $M_x=\{i\in\{1,\ldots,n-1\}|t_i\leqslant x\}$ then, as $M_x\subseteq\{1,\ldots,n-1\}$ M_x is finite, further as $t_1=a\leqslant x$ we have $1\in M_x$ so that $M_x\neq\varnothing$. Hence $m=\max(M_x)$ exist. As $m\in M_x$ we have that $t_m\leqslant x$. Now for m we have either:

$\boldsymbol{m=n-1.}$ Then $m+1=n$ and as $x\leqslant b$ we have $t_m\leqslant x\leqslant b=t_{m+1}$ so that

$$x\in[t_m,t_{m+1}]\subseteq\bigcup_{i\in\{1,\ldots,n-1\}}[t_i,t_{i+1}]$$

$\boldsymbol{m<n-1.}$ If $t_{m+1}\leqslant x$ then, as $m+1\in\{1,\ldots,n-1\}$ we have $m+1\in M_x$ so that $m\leqslant\max(M_x)=m$ leading to the contradiction $m<m$. Hence we must have $x<t_{m+1}$ so that $t_m\leqslant x<t_{m+1}$ or

$$x\in[t_m,t_{m+1}[\subseteq[t_m,t_{m+1}]\subseteq\bigcup_{i\in\{1,\ldots,n-1\}}[t_i,t_{i+1}]$$

So we have that $[a,b]\subseteq\bigcup_{i\in\{1,\ldots,n-1\}}[t_i,t_{i+1}]$ which combined with [eq: 15.2] proves

$$[a,b]=\bigcup_{i\in\{1,\ldots,n-1\}}[t_i,t_{i+1}]$$ □

If we have two partitions of a interval then we can always find a partition that is 'composed' than the two partitions. With finer we mean that the boundary points of this partition is based on the boundary points of the two given partitions.

$\mathcal{P}_1=[1_1,4_2]\bigcup[4_2,8_3]\bigcup[8_3,12_4]$, boundary points $\{1_1,4_2,8_3,12_4\}$
$\mathcal{P}_2=[1_1,2_2]\bigcup[2_2,5_3]\bigcup[5_3,6_4]\bigcup[6_4,12_5]$, boundary points $\{1_1,2_2,5_3,6_4,,12_5\}$

then the following partition can be based on the boundary points of $\mathcal{P}_1$ and $\mathcal{P}_2$

$\mathcal{P}=[1_1,2_2]\bigcup[2_2,4_3]\bigcup[4_3,5_4]\bigcup[5_4,6_5]\bigcup[6_5.8_6]\bigcup[8_6,12_7]$, boundary points

$$\{1_1,2_2,4_3,5_4,6_5,8_6,12_7\}=\{1_1,4_2,8_3,12_4\}\bigcup\{1_1,2_2,5_3,6_4,12_5\}$$

It will turn out that there only exist a such partition that satisfies this condition about the boundary points. Althoug intuitively such a partition exist and is unique, proving this is rather elaborate. This is done in the next two lemmas. This partition $\mathcal{P}$ based on $\mathcal{P}_1$ and $\mathcal{P}_2$ is noted as $\mathcal{P}_1\square\mathcal{P}_2$

Lemma 15.6. *Let $a,b\in\mathbb{R}$ wiFth $a<b$ and $\mathcal{P}_1=\{t_i^{(1)}\}_{i\in\{1,\ldots,n_1\}}\subseteq\mathbb{R}$, $\mathcal{P}_2=\{t_i^{(2)}\}_{i\in\{1,\ldots,n_2\}}\subseteq\mathbb{R}$ two partitions of $[a,b]$ then there exist a partition $\mathcal{P}=\{t_i\}_{i\in\{1,\ldots,n\}}\subseteq\mathbb{R}$ such that*

$$\{t_i|i\in\{1,\ldots,n\}\}=\{t_i^{(1)}|i\in\{1,\ldots,n_1\}\}\bigcup\{t_i^{(2)}|i\in\{1,\ldots,n_2\}\}$$

Proof. Define $P=\{t_i^{(1)}|i\in\{1,\ldots,n_1\}\}\bigcup\{t_i^{(2)}|i\in\{1,\ldots,n_2\}\}$ then as $\{t_i^{(1)}|i\in\{1,\ldots,n_1\}\}$, $\{t_i^{(2)}|i\in\{1,\ldots,n_2\}\}$ are finite [using [theorem: 6.45] we have by [theorem: 6.33] that P is finite. Hence there exist a $n\in\mathbb{N}$ and a bijection $\beta:\{1,\ldots,n\}\rightarrow P$ which defines a family $\{\beta(i)\}_{i\in\{1,\ldots,n\}}\subseteq\mathbb{R}$. Using [theorem: 6.54] there exist a bijection $\alpha:\{1,\ldots,n\}\rightarrow\{1,\ldots,n\}$ such that $\forall i\in\{1,\ldots,n-1\}$ we have $\beta(\alpha(i))\leqslant\beta(\alpha(i+1))$. If $\beta(\alpha(i))=\beta(\alpha(i+1))\underset{\beta\text{ is injective}}{\Rightarrow}\alpha(i)=\alpha(i+1)\underset{\alpha\text{ is injective}}{\Rightarrow}i=i+1$ a contradiction, so we must have that $\beta(\alpha(i))<\beta(\alpha(i+1))$. Hence if we define

$$\{t_i\}_{i\in\{1,\ldots,n\}}\subseteq\mathbb{R}\text{ by }t_i=(\beta\circ\alpha)(i)=\beta(\alpha(i))$$

Then

$$\forall i \in \{1, \ldots, n-1\}\ t_i < t_{i+1} \tag{15.3}$$

If $t \in \{t_i | i \in \{1, \ldots, n\}\}$ then there exist a a $i \in \{1, \ldots, n\}$ such that $t = t_i = \beta(\alpha(i)) \in P$. Futher if $x \in P$ then as $\beta \circ \alpha: \{1, \ldots, n\} \rightarrow P$ is a bijection there exist a $i \in \{1, \ldots, n\}$ such that $x = (\beta \circ \alpha)(i) = t_i \in \{t_i | i \in \{1, \ldots, n\}\}$. Hence we have

$$\{t_i | i \in \{1, \ldots, n\}\} = P = \left\{t_i^{(1)} | i \in \{1, \ldots, n_1\}\right\} \bigcup \left\{t_i^{(2)} | i \in \{1, \ldots, n_2\}\right\} \tag{15.4}$$

Using the above we have

$$\begin{aligned}
t_1 &\underset{[\text{lemma: } 15.3]}{=} \\
\min\left(\{t_i | i \in \{1, \ldots, n\}\}\right) &\underset{[\text{eq: } 15.4]}{=} \\
\min\left(\left\{t_i^{(1)} | i \in \{1, \ldots, n_1\}\right\} \bigcup \left\{t_i^{(2)} | i \in \{1, \ldots, n_2\}\right\}\right) &\underset{[\text{theorem: } 3.64]}{=} \\
\min\left(\min\left(\left\{t_i^{(1)} | i \in \{1, \ldots, n_1\}\right\}\right), \min\left(\left\{t_i^{(2)} | i \in \{1, \ldots, n_2\}\right\}\right)\right) &\underset{[\text{lemma: } 15.3]}{=} \\
\min\left(\left\{t_1^{(1)}, t_1^{(2)}\right\}\right) &= \\
\min\left(\{a\}\right) &= \\
a & \\
t_n &\underset{[\text{lemma: } 15.3]}{=} \\
\max\left(\{t_i | i \in \{1, \ldots, n\}\}\right) &\underset{[\text{eq: } 15.4]}{=} \\
\max\left(\left\{t_i^{(1)} | i \in \{1, \ldots, n_1\}\right\} \bigcup \left\{t_i^{(2)} | i \in \{1, \ldots, n_2\}\right\}\right) &\underset{[\text{theorem: } 3.64]}{=} \\
\max\left(\max\left(\left\{t_i^{(1)} | i \in \{1, \ldots, n_1\}\right\}\right), \max\left(\left\{t_i^{(2)} | i \in \{1, \ldots, n_2\}\right\}\right)\right) &\underset{[\text{lemma: } 15.3]}{=} \\
\max\left(\left\{t_{n_1}^{(1)}, t_{n_2}^{(2)}\right\}\right) &= \\
\max\left(\{b\}\right) &=
\end{aligned}$$

so that $a = t_1$ and $b = t_n$ which combined with [eqs: 15.3] proves that

$$\{t_i\}_{i \in \{1, \ldots, n\}} \subseteq \mathbb{R} \text{ is a partition of } [a, b]$$

□

Lemma 15.7. *Let $a, b \in \mathbb{R}$ with $a < b$, $\mathcal{P}_1 = \left\{t_i^{(1)}\right\}_{i \in \{1, \ldots, n_1\}} \subseteq \mathbb{R}$, $\mathcal{P}_2 = \left\{t_i^{(2)}\right\}_{i \in \{1, \ldots, n_2\}} \subseteq \mathbb{R}$ two partitions of $[a, b]$ then if $\mathcal{P} = \{t_i\}_{i \in \{1, \ldots, n\}} \subseteq \mathbb{R}$ and $Q = \{s_i\}_{i \in \{1, \ldots, m\}} \subseteq \mathbb{R}$ are two partitions of $[a, b]$ such that*

$$\{t_i | i \in \{1, \ldots, n\}\} = \left\{t_i^{(1)}\right\}_{i \in \{1, \ldots, n_1\}} \bigcup \left\{t_i^{(2)} | i \in \{1, \ldots, n_2\}\right\} = \{s_i | i \in \{1, \ldots, m\}\}$$

then $\mathcal{P} = Q$.

Proof. Define the functions

$$s: \{1, \ldots, n\} \rightarrow \{s_i | i \in \{1, \ldots, n\}\} \text{ by } s(i) = s_i \text{ and } t: \{1, \ldots, m\} \rightarrow \{t_i | i \in \{1, \ldots, m\}\} \text{ by } t(i) = t_i$$

then these functions are bijective [injectivity follows from [see theorem: 15.4 (3)] and surjectivity from the definition of the functions]. As $\{s_i | i \in \{1, \ldots, n\}\} = \{t_i | i \in \{1, \ldots, m\}\}$ it follows that we have the bijection

$$t^{-1} \circ s: \{1, \ldots, n\} \rightarrow \{1, \ldots, m\}$$

from which it follows that

$$n = m$$

Next we use induction to prove that the above functions are equal. So define

$$S = \{k \in \mathbb{N} | \text{If } k \in \{1, \ldots, n\} \text{ then } \forall i \in \{1, \ldots, k\} \text{ we have } s_i = t_i\}$$

then we have:

$\boldsymbol{1 \in S}$**.** This follows from the fact that $s_1 = a = t_1$.

$\boldsymbol{k\in S\Rightarrow k+1\in S.}$ Assume that $s_{k+1}\neq t_{k+1}$ then we have either:

$\boldsymbol{s_{k+1}<t_{k+1}}$**.** As $\{s_i|i\in\{1,\ldots,n\}\}=\{t_i|i\in\{1,\ldots,n\}\}$ there exist a $l\in\{1,\ldots,n\}$ such that

$$s_{k+1}=t_l \tag{15.5}$$

So that $t_l=s_{k+1}<t_{k+1}$ from which it follows that $l<k+1$ [otherwise $t_{k+1}\leqslant t_l$], hence we have that $l\leqslant k$. As $k\in S$ we have that

$$s_l=t_l$$

Hence using [eq: 15.5] on the above it follows that $s_{k+1}=s_l$ from which it follows by [theorem: 15.4 (3)] that $k+1=l$ so that we have by [eq: 15.5] that $s_{k+1}=t_{k+1}$ contradicting $s_{k+1}<t_{k+1}$.

$\boldsymbol{t_{k+1}<s_{k+1}}$**.** As $\{s_i|i\in\{1,\ldots,n\}\}=\{t_i|i\in\{1,\ldots,n\}\}$ there exist a $l\in\{1,\ldots,n\}$ such that

$$t_{k+1}=s_l \tag{15.6}$$

So that $s_l<t_{k+1}<s_{k+1}$ from which it follows that $l<k+1$ [otherwise $s_{k+1}\leqslant s_l$], hence we have that $l\leqslant k$. As $k\in S$ we have that

$$s_l=t_l$$

Hence using [eq: 15.6] it follows that $t_{k+1}=t_k$, using [theorem: 15.4 (3)] it follows that $k+1=l$ and using [eq: 15.6] this results in $s_{k+1}=t_{k+1}$ which contradicts $t_{k+1}<s_{k+1}$.

As we have a contradiction in both cases the assumption must be wrong, hence we have that $s_{k+1}=t_{k+1}$, proving that $k+1\in S$. □

Combing the two above lemmas gives the following theorem.

Theorem 15.8. *Let $a,b\in\mathbb{R}$ with $a<b$ and $\mathcal{P}_1=\{t_i^{(1)}\}_{i\in\{1,\ldots,n_1\}}\subseteq\mathbb{R}$, $\mathcal{P}_2=\{t_i^{(2)}\}_{i\in\{1,\ldots,n_2\}}\subseteq\mathbb{R}$ two partitions of $[a,b]$ then there exist a **unique** partition $\mathcal{P}=\{t_i\}_{i\in\{1,\ldots,n\}}\subseteq\mathbb{R}$ 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 **unique** partition is noted as $\mathcal{P}_1\square\mathcal{P}_2$*

Proof. This follows from [lemma: 15.6] and [lemma: 15.7]. □

Consider the following two partitions of $[1,12]$

$\mathcal{P}_1=[1_1,4_2]\bigcup[4_2,8_3]\bigcup[8_3,12_4]$, boundary points $\{1_1,4_2,8_3,12_4\}$
$\mathcal{P}_2=[1_1,2_2]\bigcup[2_2,5_3]\bigcup[5_3,6_4]\bigcup[6_4,12_5]$, boundary points $\{1_1,2_2,5_3,6_4,12_5\}$

and the 'combined' partition

$\mathcal{P}_1\square\mathcal{P}_2=[1_1,2_2]\bigcup[2_2,4_3]\bigcup[4_3,5_4]\bigcup[5_4,6_5]\bigcup[6_5.8_6]\bigcup[8_6,12_7]$, boundary points

$$\{1_1,2_2,4_3,5_4,6_5,8_6,12_7\}=\{1_1,4_2,8_3,12_4\}\bigcup\{1_1,2_2,5_3,6_4,12_5\}$$

Note that for every sub interval I in $\mathcal{P}_1\square\mathcal{P}_2$ we can find a sub interval in $\mathcal{P}_1$ that contains I and a sub interval in $\mathcal{P}_2$ that contains I. Actually we can find a function $I_{\mathcal{P}_1}$ that maps the index of the first boundary point of I to the index of the first boundary point of the containing sub interval in $\mathcal{P}_1$ and a function $I_{\mathcal{P}_2}$ that map the inde of the first boundary point of I to index of the first boundary point of the containing sub interval in $\mathcal{P}_2$.

For $\mathcal{P}_1\square\mathcal{P}_2$ and $\mathcal{P}_1$ we have:

$$\begin{array}{rcl}[1_1,2_2],[2_2,4_3]\subseteq[1_1,4_2] & \Rightarrow & I_{\mathcal{P}_1}(1)=1\\ & & I_{\mathcal{P}_1}(2)=1\\ [4_3,5_4],[5_4,6_5],[6_5.8_6]\subseteq[4_2,8_3] & \Rightarrow & I_{\mathcal{P}_1}(3)=2\\ & & I_{\mathcal{P}_1}(4)=2\\ & & I_{\mathcal{P}_1}(5)=2\\ [8_6,12_7]\subseteq[8_3,12_4] & \Rightarrow & I_{\mathcal{P}_1}(6)=3\end{array}$$

For $\mathcal{P}_1\square\mathcal{P}_2$ and $\mathcal{P}_1$ we have:

$$\begin{array}{rcl} [1_1,2_2]\subseteq[1_1,2_2] & \Rightarrow & I_{\mathcal{P}_2}(1)=1\\ [2_2,4_3],[4_3,5_4]\subseteq[2_2,5_3] & \Rightarrow & I_{\mathcal{P}_2}(2)=2\\ & & I_{\mathcal{P}_2}(3)=2\\ [5_4,6_5]\subseteq[5_3,6_4] & \Rightarrow & I_{\mathcal{P}_2}(4)=3\\ [6_5,8_6],[8_6,12_7]\subseteq[6_4,12_5] & \Rightarrow & I_{\mathcal{P}_2}(5)=4\\ & & I_{\mathcal{P}_2}(6)=4 \end{array}$$

Hence we have that

$$I_{\mathcal{P}_1}:\{1,\ldots,7-1\}\to\{1,\ldots,4-1\} \text{ is surjective}$$

$$I_{\mathcal{P}_3}:\{1,\ldots,7-1\}\to\{1,\ldots,5-1\} \text{ is surjective}$$

Further we have

$$\begin{array}{rcl} I_{\mathcal{P}_1}^{-1}(\{1\})=\{1\} & = & \{1,\ldots,1\}\\ I_{\mathcal{P}_1}^{-1}(\{2\})=\{3,4,5\} & = & \{3,\ldots,5\}\\ I_{\mathcal{P}_1}^{-1}(\{3\})=\{6\} & = & \{6,\ldots,6\}\\ & & \\ I_{\mathcal{P}_2}^{-1}(\{1\})=\{1\} & = & \{1,\ldots,1\}\\ I_{\mathcal{P}_2}^{-1}(\{2\})=\{2,3\} & = & \{2,\ldots,3\}\\ I_{\mathcal{P}_2}^{-1}(\{3\})=\{4\} & = & \{4,\ldots,4\}\\ I_{\mathcal{P}_2}^{-1}(\{4\})=\{5,6\} & = & \{5,\ldots,6\} \end{array}$$

so that

$$\begin{array}{rl} I_{\mathcal{P}_1}^{-1}(\{1\}),I_{\mathcal{P}_1}^{-1}(\{2\}),I_{\mathcal{P}_1}^{-1}(\{3\}) & \text{are pairwise disjoint}\\ I_{\mathcal{P}_2}^{-1}(\{1\}),I_{\mathcal{P}_2}^{-1}(\{2\}),I_{\mathcal{P}_2}^{-1}(\{3\}),I_{\mathcal{P}_2}^{-1}(\{4\}), & \text{are pairwise disjoint} \end{array}$$

The exact proof of the above in general is done in the following two lemmas.

Lemma 15.9. *Let $a,b\in\mathbb{R}$ with $a<b$ and $\mathcal{P}_1=\{t_i^{(1)}\}_{i\in\{1,\ldots,n_1\}}\subseteq\mathbb{R}$, $\mathcal{P}_2=\{t_i^{(2)}\}_{i\in\{1,\ldots,n_2\}}\subseteq\mathbb{R}$ two partitions of $[a,b]$, let $\mathcal{P}_1\square\mathcal{P}_2=\{t_i\}_{i\in\{1,\ldots,n\}}\subseteq\mathbb{R}$ then we have that*

$$\forall i\in\{1,\ldots,n-1\} \text{ there exist a } \textbf{\textit{unique}}\ i_{\mathcal{P}_1}\in\{1,\ldots,n_1-1\} \text{ and a } \textbf{\textit{unique}}\ i_{\mathcal{P}_2}\in\{1,\ldots,n_2-1\}$$

such that

$$[t_i,t_{i+1}]\subseteq\left[t_{i_{\mathcal{P}_1}}^{(1)},t_{i_{\mathcal{P}_1}+1}^{(1)}\right] \text{ and } [t_i,t_{i+1}]\subseteq\left[t_{i_{\mathcal{P}_1}}^{(2)},t_{i_{\mathcal{P}_1}+1}^{(2)}\right]$$

This allows us to define two functions

$$I_{\mathcal{P}_1}:\{1,\ldots,n-1\}\to\{1,\ldots,n_1\} \text{ by } I_{\mathcal{P}_1}(i)=i_{\mathcal{P}_1}$$

$$I_{\mathcal{P}_2}:\{1,\ldots,n-1\}\to\{1,\ldots,n_2\} \text{ by } I_{\mathcal{P}_2}(i)=i_{\mathcal{P}_2}$$

such that

$$[t_i,t_{i+1}]\subseteq\left[t_{I_{\mathcal{P}_1}(i)}^{(1)},t_{I_{\mathcal{P}_1}(i)+1}^{(1)}\right] \text{ and } [t_i,t_{i+1}]\subseteq\left[t_{I_{\mathcal{P}_1}(i)}^{(2)},t_{I_{\mathcal{P}_1}(i)+1}^{(2)}\right]$$

Proof. We prove this for $i_{\mathcal{P}_1}$ [the proof for $i_{\mathcal{P}_2}$ is exact the same as this proof by exchanging 1 with 2]. Let $i\in\{1,\ldots,n-1\}$ then we have

$$t_i\neq t_n=b=t_n^{(1)}=t_n^{(2)} \text{ so that } t_i\neq t_n^{(1)} \text{ and } t_i\neq t_n^{(2)}$$

Define $B=\{j\in\{1,\ldots,n_1\}|t_j^{(1)}\leqslant t_i\}\subseteq\{1,\ldots,n_1\}$ then as $t_1^{(1)}=a\leqslant t_i$ we have $1\in B$ so that $B\neq\varnothing$. As B is finite $m=\max(B)\in\{1,\ldots,n_1\}$ exist. If $m=n_1$ then $b=t_{n_1}^{(1)}=t_m^{(1)}\leqslant t_i$ which as $t_i\neq b$ is a contradiction. Hence we must have that $m\neq n_1$ so that

$$m\in\{1,\ldots,n_1-1\} \tag{15.7}$$

As $m = \max(B)$ we have $m+1 \notin B$ so that $t_i < t_{m+1}^{(1)}$ and, as $m \in B$, we have $t_m^{(1)} \leqslant t_i$ or

$$t_m^{(1)} \leqslant t_i < t_{m+1}^{(1)} \tag{15.8}$$

By [theorem: 15.8] there exist a $l \in \{1, \ldots, n\}$ such that

$$t_{m+1}^{(1)} = t_l \tag{15.9}$$

Assume that $t_{m+1}^{(1)} < t_{i+1}$, then by the above and [eq: 15.8] we have $t_i < t_l < t_{i+1}$ so that $t_l \in]t_i, t_{i+1}[$ which is impossible by [theorem: 15.4 (5)]. Hence the assumption is wrong and we must have $t_{i+1} \leqslant t_{m+1}^{(1)}$. Combining this with [eq: 15.8] gives $t_m^{(1)} \leqslant t_i < t_{i+1} \leqslant t_{m+1}^{(1)}$ or $[t_i, t_{i+1}] \subseteq [t_m^{(1)}, t_{m+1}^{(1)}]$. So if we take $i_{\mathcal{P}_1} = m$ then we have

$$i_{\mathcal{P}_1} \in \{1, \ldots, n_1 - 1\} \text{ and } [t_i, t_{i+1}] \subseteq [t_{i_{\mathcal{P}_1}}^{(1)}, t_{i_{\mathcal{P}_1}+1}^{(1)}]$$

This proves existence, next we prove uniqueness. So assume that there exist a $k \in \{1, \ldots, n_1 - 1\}$ such that $[t_i, t_{i+1}] \subseteq [t_k^{(1)}, t_{k+1}^{(1)}]$. Assume that $k \neq i_{\mathcal{P}_1}$ then we have two cases to consider:

$\boldsymbol{k < i_{\mathcal{P}_1}}$**.** Then $t_k^{(1)} < t_{i_{\mathcal{P}_1}}^{(1)} \leqslant t_i < t_{i+1} \leqslant t_{k+1}^{(1)}$ so that $t_{i_{\mathcal{P}_1}}^{(1)} \in]t_k^{(1)}, t_{k+1}^{(1)}[$ which is impossible by [theorem: 15.4 (5)].

$\boldsymbol{i_{\mathcal{P}_1} < k}$**.** Then $t_{i_{\mathcal{P}_1}}^{(1)} < t_k^{(1)} \leqslant t_i < t_{i+1} \leqslant t_{i_{\mathcal{P}_1}}^{(1)}$ so that $t_k^{(1)} \in]t_{i_{\mathcal{P}_1}}^{(1)}, t_{i_{\mathcal{P}_1}+1}^{(i)}[$ which is impossible by [theorem: 15.4 (5)].

So as in all cases we have a contradiction the assumption must be wrong. Hence $k = i_{\mathcal{P}_1}$ which proves uniqueness. □

Lemma 15.10. *Let $a, b \in \mathbb{R}$ with $a < b$ and $\mathcal{P}_1 = \{t_i^{(1)}\}_{i \in \{1, \ldots, n_1\}} \subseteq \mathbb{R}$, $\mathcal{P}_2 = \{t_i^{(2)}\}_{i \in \{1, \ldots, n_2\}} \subseteq \mathbb{R}$ two partitions of $[a, b]$, let $\mathcal{P}_1 \square \mathcal{P}_2 = \{t_i\}_{i \in \{1, \ldots, n\}} \subseteq \mathbb{R}$ then for the functions*

$$I_{\mathcal{P}_1}: \{1, \ldots, n-1\} \to \{1, \ldots, n_1 - 1\}$$

$$I_{\mathcal{P}_2}: \{1, \ldots, n-1\} \to \{1, \ldots, n_2 - 1\}$$

functions such that

$$\forall i \in \{1, \ldots, n-1\} \text{ we have } [t_i, t_{i+1}] \subseteq [t_{I_{\mathcal{P}_1}(i)}^{(1)}, t_{I_{\mathcal{P}_1}(i)+1}^{(1)}] \text{ and } [t_{I_{\mathcal{P}_2}(i)}^{(2)}, t_{I_{\mathcal{P}_2}(i)+1}^{(2)}]$$

[who exist by [lemma: 15.9]] we have that:

1.

   a. $I_{\mathcal{P}_1}: \{1, \ldots, n-1\} \to \{1, \ldots, n_1 - 1\}$ *is surjective*

   b. $I_{\mathcal{P}_2}: \{1, \ldots, n-1\} \to \{1, \ldots, n_{12} - 1\}$ *is surjective*

   and $I_{\mathcal{P}_2}: \{1, \ldots, n-1\} \to \{1, \ldots, n_2 - 1\}$ *are surjective.*

2.

   a. $\forall i \in \{1, \ldots, n_1 - 1\}$ $I_{\mathcal{P}_1}^{-1}(\{i\}) = \{m_i^{(1)}, \ldots, M_i^{(1)}\}$ *where* $t_{m_i^{(1)}} = t_i^{(1)}$ *and* $t_{M_i^{(1)}+1} = t_{i+1}^{(1)}$

   b. $\forall i \in \{1, \ldots, n_2 - 1\}$ $I_{\mathcal{P}_2}^{-1}(\{i\}) = \{m_i^{(2)}, \ldots, M_i^{(2)}\}$ *where* $t_{m_i^{(2)}} = t_i^{(2)}$ *and* $t_{M_i^{(2)}+1} = t_{i+1}^{(2)}$

3.

   a. $\forall i, j \in \{1, \ldots, n_1 - 1\}$ *with* $i \neq j$ *we have* $I_{\mathcal{P}_1}^{-1}(\{i\}) \bigcap I_{\mathcal{P}_1}^{-1}(\{j\}) = \varnothing$

   b. $\forall i, j \in \{1, \ldots, n_2 - 1\}$ *with* $i \neq j$ *we have* $I_{\mathcal{P}_2}^{-1}(\{i\}) \bigcap I_{\mathcal{P}_2}^{-1}(\{j\}) = \varnothing$

Proof.

1.

   a. Let $k \in \{1, \ldots, n_1 - 1\}$ then for $t_k^{(1)}$ there exist by [theorem: 15.8] there exist a $i \in \{1, \ldots, n\}$ such that $t_k^{(1)} = t_i$. As $k < n_1$ we have that $t_k^{(1)} < t_n^{(1)} = b$ so that $t_i \neq b$ proving that $i \neq n$ or

$$i \in \{1, \ldots, n-1\} \text{ and } t_i = t_k^{(1)} \tag{15.10}$$

Assume that $t_{k+1}^{(1)} < t_{i+1}$. By [theorem: 15.8] there exist a $j \in \{1, \ldots, n\}$ such that $t_j = t_{k+1}^{(1)}$. Then we have $t_i = t_k^{(1)} < t_{k+1}^{(1)} = t_j = t_{k+1}^{(1)} < t_{i+1}$ or $t_j \in]t_I, t_{I+1}[$, which is impossible by [theorem: 15.4 (5)], so that assumption must be wrong and we must have $t_{i+1} \leqslant t_{k+1}^{(1)}$, combining this with [eq: 15.10] results in $t_k^{(1)} = t_i < t_{i+1} \leqslant t_{k+1}^{(1)}$ or $[t_i, t_{i+1}] \subseteq [t_k^{(1)}, t_{k+1}^{(1)}]$. This proves that $k = I_{\mathcal{P}_1}(i)$ and thus surjectivity.

b. This is similar as the proof of (1.a) by replacing 1 by 2 in the proof.

2.

a. Let $i \in \{1, \ldots, n_1 - 1\}$ then, as $I_{\mathcal{P}_1}$ is surjective we have that $\varnothing = I_{\mathcal{P}_1}^{-1}(\{i\}) \subseteq \{1, \ldots, n-1\}$. Hence then minimum and maximum exist, take

$$m_i^{(1)} = \min\left(I_{\mathcal{P}_1}^{-1}(\{i\})\right) \text{ and } M_i^{(1)} = \max\left(I_{\mathcal{P}_1}^{-1}(\{i\})\right) \tag{15.11}$$

By the definition of $\min, \max$ we have

$$I_{\mathcal{P}_1}^{-1}(\{i\}) \subseteq \left\{m_i^{(1)}, \ldots, M_i^{(1)}\right\} \text{ and } m_i^{(1)}, M_i^{(1)} \in I_{\mathcal{P}_1}^{-1}(\{i\}) \tag{15.12}$$

If $k \in \left\{m_i^{(1)}, \ldots, M_i^{(1)}\right\}$ then $m_i^{(1)} \leqslant k \leqslant M_i^{(1)} \Rightarrow k + 1 \leqslant M_i^{(1)}$ so that

$$t_{m_i^{(1)}} \leqslant t_k \text{ and } t_{k+1} \leqslant t_{M_i^{(1)}+1} \tag{15.13}$$

As we have $m_i^{(1)}, M_i^{(1)} \in I_{\mathcal{P}_1}^{-1}(\{i\})$ it follows that $I_{\mathcal{P}_1}\left(m_i^{(1)}\right) = i = I_{\mathcal{P}_1}\left(M_i^{(1)}\right)$ so that $\left[t_{m_i^{(1)}}, t_{m_i^{(1)}+1}\right] \subseteq \left[t_i^{(1)}, t_{i+1}^{(1)}\right]$ and $\left[t_{M_i^{(1)}}, t_{M_i^{(1)}+1}\right] \subseteq \left[t_i^{(1)}, t_{i+1}^{(1)}\right]$ so that

$$t_i^{(1)} \leqslant t_{m_i^{(1)}} \leqslant t_k < t_{k+1} \leqslant t_{M_i^{(1)}+1} \leqslant t_{i+1}^{(1)} \text{ and } t_i^{(1)} \leqslant t_{M_i^{(1)}} < t_{M_i^{(1)}+1} \leqslant t_{i+1}^{(1)} \tag{15.14}$$

or $[t_k, t_{k+1}] \subseteq \left[t_i^{(1)}, t_{i+1}^{(1)}\right]$, hence $I_{\mathcal{P}_1}(k) = i$ or $k \in I_{\mathcal{P}_1}^{-1}(\{i\})$. So we have proved that $\left\{m_i^{(1)}, \ldots, M_i^{(1)}\right\} \subseteq I_{\mathcal{P}_1}^{-1}(\{i\})$, combining this with [eq: 15.12] gives

$$I_{\mathcal{P}_1}^{-1}(\{i\}) = \left\{m_i^{(1)}, \ldots, M_i^{(1)}\right\} \tag{15.15}$$

Assume that $t_i^{(1)} < t_{m_i^{(1)}}$. By [theorem: 15.8] there exist a $l \in \{1, \ldots, n\}$ such that

$$t_l = t_i^{(1)}$$

giving $t_l < t_{m_i^{(1)}}$, resulting in

$$l < m_i^{(1)}, \tag{15.16}$$

hence $l + 1 < m_i^{(1)} + 1$. From this it follows that

$$t_i^{(1)} = t_l < t_{l+1} < t_{m^{(1)}+1} \leqslant_{[\text{eq: } 15.14]} t_{i+1}^{(1)}$$

So that $[t_l, t_{l+1}] \subseteq \left[t_i^{(1)}, t_{i+1}^{(1)}\right]$ proving that $I_{\mathcal{P}_1}(l) = i$ or $l \in I_{\mathcal{P}_1}^{-1}(\{i\}) = \left\{m_i^{(1)}, \ldots, M_i^{(1)}\right\}$, hence

$$m_i^{(1)} \leqslant l <_{[\text{eq: } 15.16]} m_i^{(1)}$$

a contradiction. From this we conclude that the assumption is wrong, so $t_{m_i^{(1)}} \leqslant t_i^{(1)}$, combining this with [eq: 15.14] proves that

$$t_{m_i^{(1)}} = t_i^{(1)} \tag{15.17}$$

Assume that $t_{M_i^{(1)}+1} < t_{i+1}^{(1)}$. By [theorem: 15.8] there exist a $k \in \{1, \ldots, n\}$ such that

$$t_k = t_{i+1}^{(1)}$$

giving $t_{M_k^{(1)}+1} < t_k$ so that resulting in $M_k^{(1)} + 1 < k$ or

$$M_k^{(1)} < k - 1 \tag{15.18}$$

so that

$$t_i^{(1)} \leqslant_{[\text{eq: } 15.14]} t_{M_k^{(1)}} < t_{k-1} < t_k = t_{i+1}^{(1)}$$

So that $[t_{k-1},t_k]\subseteq[t_i^{(1)},t_{i+1}^{(1)}]$ proving that $I_{\mathcal{P}_1}(k-1)=i$ or $k-1\in I_{\mathcal{P}_1}^{-1}(\{i\})=\{m_i^{(1)},\ldots,M_i^{(1)}\}$ hence $k-1\leqslant M_i^{(1)}<_{[\text{eq: }15.18]}k-1$ a contradiction. Hence the assumption must be wrong and we have $t_{i+1}^{(1)}\leqslant t_{M_{i+1}^{(1)}}$, combining this with [eq: 15.14] gives

$$t_{M_{i+1}^{(1)}}=t_{i+1}^{(1)} \tag{15.19}$$

Finall (2.a) is proved by [eqs: 15.15, 15.17 and 15.19].

b. This is similar as the proof of (2.a) by replacing 1 by 2 in the proof.

3.

a. Let $i,j\in\{1,\ldots,n_1-1\}$ with $i\neq j$. Assume that $k\in I_{\mathcal{P}_1}^{-1}(\{i\})\bigcap I_{\mathcal{P}_1}^{-1}(\{j\})$ then $I_{\mathcal{P}_1}(\{k\})=i$ and $I_{\mathcal{P}_1}(\{k\})=j$ so that by the defition of a function we have $i=j$ a contradiction. Hence $I_{\mathcal{P}_1}^{-1}(\{i\})\bigcap I_{\mathcal{P}_1}^{-1}(\{j\})=\varnothing$.

b. This is similar as the proof of (3.a) by replacing 1 by 2 in the proof. □

To be able to define a integral of a function on a closed interval we need a limiting proess where the length of the intervals in the partitions become smaller and smaller so that more details of the function are used in the calculation of the integral. To quantify this we introduce a norm of a partition. Beware this norm is unrelated to norms on normed spaces.

Definition 15.11. *Let $a,b\in\mathbb{R}$ with $a<b$ and $\mathcal{P}=\{t_i\}_{i\in\{1,\ldots,n\}}\subseteq\mathbb{R}$ a partition of $[a,b]$ then the **norm** of the partition noted as $\mu(\mathcal{P})$ is defined by*

$$\mu(\mathcal{P})=\max\left(\{|t_{i+1}-t_i||i\in\{1,\ldots,n-1\}\}\right)\underset{t_i<t_{i+1}}{=}\max\left(\{t_{i+1}-t_i|i\in\{1,\ldots,n-1\}\}\right)$$

Definition 15.12. *Let $a,b\in\mathbb{R}$ with $a<b$ and $\mathcal{P}=\{t_i\}_{i\in\{1,\ldots,n\}}\subseteq\mathbb{R}$ a partition of $[a,b]$ then a **tag** on $\mathcal{P}$ is a family $\{s_i\}_{i\in\{1,\ldots,n-1\}}\subseteq\mathbb{R}$ such that $\forall i\in\{1,\ldots,n-1\}$ we have $s_i\in[t_i,t_{i+1}]$.*
*A **tagged partition** on $[a,b]$ is a pair of a partition on the partition. So a tagged partition is a pair*

$$\langle\{t_i\}_{i\in\{1,\ldots,n\}}\subseteq\mathbb{R},\{s_i\}_{i\in\{1,\ldots,n-1\}}\subseteq\mathbb{R}\rangle$$

such that:

1. $n\in\mathbb{N}\setminus\{1\}$
2. $\forall i\in\{1,\ldots,n-1\}$ *we have* $t_i<t_{i+1}$ *and* $s_i\in[t_i,t_{i+1}]$
3. $t_1=a$ *and* $t_2=b$

The norm of a tagged partition is the norm of the partition of the tagged partition in othere words if

$$\langle\{t_i\}_{i\in\{1,\ldots,n\}}\subseteq\mathbb{R},\{s_i\}_{i\in\{1,\ldots,n-1\}}\subseteq\mathbb{R}\rangle$$

is a tagged partition then

$$\mu(\langle\{t_i\}_{i\in\{1,\ldots,n\}}\subseteq\mathbb{R},\{s_i\}_{i\in\{1,\ldots,n-1\}}\subseteq\mathbb{R}\rangle)=\mu(\{t_i\}_{i\in\{1,\ldots,n\}}\subseteq\mathbb{R})$$

Theorem 15.13. *Let $a,b\in\mathbb{R}^+$ with $a<b$, $\varepsilon\in\mathbb{R}^+$ then there exist a tagged partition $\mathbb{P}$ on $[a,b]$ such that $\mu(\mathbb{P})<\varepsilon$*

Proof. Using [theorem: 10.30] there exist a $N\in\mathbb{N}$ such that $\frac{1}{N}<\frac{\varepsilon}{b-a}$ and take $\delta=\frac{b-a}{N}$. Then we have $\frac{b-a}{N}=(b-a)\cdot\frac{1}{N}<(b-a)\cdot\frac{\varepsilon}{b-a}=\varepsilon$ so that

$$\delta<\varepsilon \tag{15.20}$$

Define now

$$\{t_i\}_{i\in\{1,\ldots,N+1\}}\text{ by }t_i=a+(i-1)\cdot\delta\text{ and }\{s_i\}_{i\in\{1,\ldots,N+1\}}\text{ by }s_i=t_i$$

then $\forall i\in\{1,\ldots,(N+1)-1\}$ we have

$$s_i\in[t_i,t_{i+1}]$$

and

$$t_i=a+(i-1)\cdot\delta<a+(i-1)\cdot\delta+\delta=a+((i+1)-1)\cdot\delta=t_{i+1}$$

and

$$t_{i+1}-t_i=a+((i+1)-1)\cdot\delta-(a+(i-1)\cdot\delta)=\delta<\varepsilon$$

so that

$$\mu(\{t_i\}_{i\in\{1,\ldots,N+1\}}) = \max(\{t_{i+1}-t_i | i\in\{1,\ldots,(N+1)-1\}\}) = \max(\{\delta\}) = \delta < \varepsilon$$

Further

$$t_1 = a+(1-1)\cdot\delta = a \text{ and } t_{N+1} = a+((N+1)-1)\cdot\delta = a+N\cdot\delta = a+\frac{b-a}{N}\cdot N = b$$

Hence

$$\mathbb{P} = \langle\{t_i\}_{i\in\{1,\ldots,N+1\}}, \{s_i\}_{i\in\{1,\ldots,N+1\}}\rangle$$

is a tagged partition on $[a,b]$ with $\mu(\mathbb{P}) < \varepsilon$. $\square$

We are now ready to define a Riemann sum that will be used to aproximate the integral of a function.

Definition 15.14. *Let $a,b\in\mathbb{R}$ with $a<b$, $\mathbb{P} = \langle\{t_i\}_{i\in\{1,\ldots,n\}} \subseteq \mathbb{R}, \{s_i\}_{i\in\{1,\ldots,n-1\}} \subseteq \mathbb{R}\rangle$ a tagged partition on $[a,b]$, $\langle X, \|\|\rangle$ a normed space and $f:[a,b]\to X$ a function then the **Riemann sum of f using the tagged partition** $\mathbb{P}$ noted ans $\mathcal{S}(f,\mathbb{P})$ is defined by*

$$\mathcal{S}(f,\mathbb{P}) = \sum_{i=1}^{n-1}(t_{i+1}-t_i)\cdot f(s_i)$$

Theorem 15.15. *Let $a,b\in\mathbb{R}$ with $a<b$, $\mathbb{P} = \langle\{t_i\}_{i\in\{1,\ldots,n\}} \subseteq \mathbb{R}, \{s_i\}_{i\in\{1,\ldots,n-1\}} \subseteq \mathbb{R}\rangle$ a tagged partition on $[a,b]$, $\alpha\in\mathbb{K}$ and $f:[a,b]\to X$, $g:[a,b]\to X$ functions then*

$$\mathcal{S}(\alpha\cdot f+g,\mathbb{P}) = \alpha\cdot\mathcal{S}(f,\mathbb{P}) + \mathcal{S}(g,\mathbb{P})$$

Proof. We have

$$\begin{aligned}
\mathcal{S}(\alpha\cdot f+g,\mathbb{P}) &= \sum_{i=1}^{n-1}(t_{i+1}-t_i)\cdot(\alpha\cdot f+g)(s_i)\\
&= \sum_{i=1}^{n-1}(t_{i+1}-t_i)\cdot(\alpha\cdot f(s_i)+g(s_i))\\
&= \sum_{i=1}^{n-1}(\alpha\cdot(t_{i+1}-t_i)\cdot f(s_i)+(t_{i+1}-t_i)\cdot g(s_i))\\
&= \alpha\cdot\sum_{i=1}^{n-1}(t_{i+1}-t_i)\cdot f(s_i)+\sum_{i=1}^{n-1}(t_{i+1}-t_i)\cdot g(s_i)\\
&= \alpha\cdot\mathcal{S}(f,\mathbb{P})+\mathcal{S}(g,\mathbb{P})
\end{aligned}$$

$\square$

Theorem 15.16. *Let $a,b\in\mathbb{R}$ with $a<b$, $\mathbb{P} = \langle\{t_i\}_{i\in\{1,\ldots,n\}} \subseteq \mathbb{R}, \{s_i\}_{i\in\{1,\ldots,n-1\}} \subseteq \mathbb{R}\rangle$ a taaged partition on $[a,b]$. $\langle\mathbb{R},\|\rangle$ the normed space of real numbers and $f:[a,b]\to\mathbb{R}$, $g:[a,b]\to\mathbb{R}$ functions 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 $\mathcal{S}(f,\mathbb{P}) = \sum_{i=1}^{n-1}(t_{i+1}-t_i)\cdot f(s_i) = \sum_{i=1}^{n-1}(t_{i+1}-t_i)\cdot g(s_i) = \mathcal{S}(g,\mathbb{P})$. $\square$

Lemma 15.17. *Let $a,b\in\mathbb{R}$ with $a<b$,*

$$\mathbb{P}_1 = \left\langle\left\{t_i^{(1)}\right\}_{i\in\{1,\ldots,n_1\}} \subseteq \mathbb{R}, \left\{s_i^{(1)}\right\}_{i\in\{1,\ldots,n_1-1\}}\right\rangle$$

$$\mathbb{P}_2 = \left\langle\left\{t_i^{(2)}\right\}_{i\in\{1,\ldots,n_2\}} \subseteq \mathbb{R}, \left\{s_i^{(2)}\right\}_{i\in\{1,\ldots,n_2-1\}}\right\rangle$$

tagged partitions on $[a,b]$, $\langle X,\|\|\rangle$ a normed space and

$$f:[a.b]\to X \text{ a function}$$

then for

$$\mathcal{P} = \{t_i\}_{i\in\{1,\ldots,n\}} \text{ defined by } \mathcal{P} = \left\{t_i^{(1)}\right\}_{i\in\{1,\ldots,n_1\}} \square \left\{t_i^{(1)}\right\}_{i\in\{1,\ldots,n_2}$$

we have:

1. $S(f,\mathbb{P}_1)=\sum_{i=1}^{n-1}(t_{i+1}-t_i)\cdot f\big(s^{(1)}_{I_{\mathcal{P}_1}(i)}\big)$
2. $S(f,\mathbb{P}_2)=\sum_{i=1}^{n-1}(t_{i+1}-t_i)\cdot f\big(s^{(2)}_{I_{\mathcal{P}_2}(i)}\big)$
3. $S(f,\mathbb{P}_1)-S(f,\mathbb{P}_2)=\sum_{i=1}^{n-1}(t_{i+1}-t_i)\cdot\big(f\big(s^{(1)}_{I_{\mathcal{P}_1}(i)}\big)-f\big(s^{(2)}_{I_{\mathcal{P}_2}(i)}\big)\big)$

where

$$I_{\mathcal{P}_1}:\{1,\ldots,n-1\}\rightarrow\{1,\ldots,n_1-1\} \text{ and } I_{\mathcal{P}_2}:\{1,\ldots,n-1\}\rightarrow\{1,\ldots,n_2-1\}$$

are the functions defined in lemmas [lemma: 15.9] and [lemma: 15.10].

Proof.

1. As $\{1,\ldots,n_1-1\}=\bigcup_{k\in\{1,\ldots,n_1-1\}}\{k\}$ we have

$$\begin{aligned}\{1,\ldots,n-1\} &= I_{\mathcal{P}_1}^{-1}(\{1,\ldots,n_1-1\})\\ &= I_{\mathcal{P}_1}^{-1}\left(\bigcup_{k\in\{1,\ldots,n_1-1\}}\{k\}\right)\\ &\underset{[\text{theorem: }2.134]}{=} \bigcup_{k\in\{1,\ldots,n_1-1\}} I_{\mathcal{P}_1}^{-1}(\{k\})\end{aligned}$$

and by [lemmma: 15.10] we have

$$\forall i,j\in\{1,\ldots,n_1-1\}\text{ with } i\neq j \text{ we have } I_{\mathcal{P}_1}^{-1}(\{i\})\bigcap I_{\mathcal{P}_1}^{-1}(\{j\})=\varnothing$$

allowing us to apply [theorem: 11.43] in the following.

$$\begin{aligned}\sum_{i=1}^{n-1}(t_{i+1}-t_i)\cdot f\big(s^{(1)}_{I_{\mathcal{P}_1}(i)}\big) &= \sum_{i\in\{1,\ldots,n-1\}}(t_{i+1}-t_i)\cdot f\big(s^{(1)}_{I_{\mathcal{P}_1}(i)}\big)\\ &= \sum_{i\in\bigcup_{k\in\{1,\ldots,n_1-1\}}I_{\mathcal{P}_1}^{-1}(\{k\})}(t_{i+1}-t_i)\cdot f\big(s^{(1)}_{I_{\mathcal{P}_1}(i)}\big)\\ &\underset{[\text{theorem: }11.43]}{=} \sum_{k=1}^{n_1-1}\left(\sum_{i\in I_{\mathcal{P}_1}^{-1}(\{k\})}(t_{i+1}-t_i)\cdot f\big(s^{(1)}_{I_{\mathcal{P}_1}(i)}\big)\right)\\ &= \sum_{k=1}^{n_1-1}\left(\sum_{i\in I_{\mathcal{P}_1}^{-1}(\{k\})}(t_{i+1}-t_i)\cdot f\big(s^{(1)}_k\big)\right)\\ &= \sum_{k=1}^{n_1-1}\left(\sum_{i\in I_{\mathcal{P}_1}^{-1}(\{k\})}(t_{i+1}-t_i)\right)\cdot f\big(s^{(1)}_k\big)\\ &\underset{[\text{lemma: }15.10]}{=} \sum_{k=1}^{n_1-1}\left(\sum_{i\in\{m^{(1)}_k,\ldots,M^{(1)}_k\}}(t_{i+1}-t_i)\right)\cdot f\big(s^{(1)}_k\big)\\ &= \sum_{k=1}^{n_1-1}\left(\sum_{i\in m^{(1)}_k}^{M^{(1)}_k}(t_{i+1}-t_i)\right)\cdot f\big(s^{(1)}_k\big)\\ &\underset{[\text{theprem: }11.21]}{=} \sum_{k=1}^{n_1-1}\big(t_{M^{(1)}_k+1}-t_{m^{(1)}_k}\big)\cdot f\big(s^{(1)}_k\big)\\ &\underset{[\text{lemma: }15.10]}{=} \sum_{k=1}^{n_1-1}\big(t^{(1)}_{k+1}-t^{(1)}_k\big)\cdot f\big(s^{(1)}_k\big)\\ &= \mathcal{S}(f,\mathbb{P}_1)\end{aligned}$$

2. The proof is similar to the proof of (1) just replace 1 with 2.

3. We have that

$$\begin{aligned} S(f,\mathbb{P}_1)-S(f,\mathbb{P}_2) & \underset{(1,2)}{=} \sum_{i=1}^{n-1} (t_{i+1}-t_i)\cdot f\big(s^{(1)}_{I_{\mathcal{P}_1}(i)}\big)-\sum_{i=1}^{n-1} (t_{i+1}-t_i)\cdot f\big(s^{(2)}_{I_{\mathcal{P}_2}(i)}\big) \\ & = \sum_{i=1}^{n-1} (t_{i+1}-t_i)\cdot \big(f\big(s^{(1)}_{I_{\mathcal{P}_1}(i)}\big)-f\big(s^{(2)}_{I_{\mathcal{P}_2}(i)}\big)\big) \end{aligned}$$

□

Lemma 15.18. *Let $a,b\in\mathbb{R}$ with $a<b$, $\langle X,\|\|\rangle$ a normed space and*

$$f:[a.b]\to X \text{ a continuous function}$$

then for every $\varepsilon\in\mathbb{R}^+$ there exist a $\delta\in\mathbb{R}^+$ such that for every two tagged partitions $\mathbb{P}_1$, $\mathbb{P}_2$ on $[a,b]$ with $\mu(\mathbb{P}_1)<\delta$ and $\mu(\mathbb{P}_2)<\delta$ we have $\|\mathcal{S}(f,\mathbb{P}_1)-\mathcal{S}(f,\mathbb{P}_2)\|<\varepsilon$

Proof. As $[a,b]$ is compact [see theorem: 14.232] and f is continuous we can use [theorem: 14.241] that f is uniform continuous on $[a,b]$. So there exist a $\delta'\in\mathbb{R}^+$ such that

$$\forall x,x'\in[a,b] \text{ with } |x-x'|< \text{ we have } \|f(x)-f(x')\|<\frac{\varepsilon}{b-a} \tag{15.21}$$

Take $\delta=\frac{\delta'}{2}$ and let

$$\mathbb{P}_1=\big\langle \{t^{(1)}_i\}_{i\in\{1,\ldots,n_1\}}\subseteq\mathbb{R}, \{s^{(1)}_i\}_{i\in\{1,\ldots,n_1-1\}}\big\rangle$$

$$\mathbb{P}_2=\big\langle \{t^{(2)}_i\}_{i\in\{1,\ldots,n_2\}}\subseteq\mathbb{R}, \{s^{(2)}_i\}_{i\in\{1,\ldots,n_2-1\}}\big\rangle$$

be two tagged partitions on $[a,b]$ such that $\mu(\mathbb{P}_1)<\delta$ and $\mu(\mathbb{P}_2)<\delta$. Consider

$$\mathcal{P}=\{t_i\}_{i\in\{1,\ldots,n\}} \text{ defined by } \mathcal{P}=\{t^{(1)}_i\}_{i\in\{1,\ldots,n_1\}} \square \{t^{(1)}_i\}_{i\in\{1,\ldots,n_2}$$

then we have by [lemma: 15.17] that

$$\|\mathcal{S}(f,\mathbb{P}_1)-\mathcal{S}(f,\mathbb{P}_2)\|=\left\|\sum_{i=1}^{n-1}(t_{i+1}-t_i)\cdot\big(f\big(s^{(1)}_{I_{\mathcal{P}_1}(i)}\big)-f\big(s^{(2)}_{I_{\mathcal{P}_2}(i)}\big)\big)\right\| \tag{15.22}$$

Let $i\in\{1,\ldots,n-1\}$ then by [lemma: 15.9] $t_i\in[t_i,t_{i+1}]\subseteq\Big[t^{(1)}_{I_{\mathcal{P}_1}(i)},t^{(1)}_{I_{\mathcal{P}_1}(i)+1}\Big]\bigcap\Big[t^{(2)}_{I_{\mathcal{P}_2}(i)},t^{(2)}_{I_{\mathcal{P}_2}(i)+1}\Big]$ so that

$$t^{(1)}_{I_{\mathcal{P}_1}(i)}\leqslant t_i\leqslant t^{(1)}_{I_{\mathcal{P}_1}(i)+1} \text{ and } t^{(2)}_{I_{\mathcal{P}_2}(i)}\leqslant t_i\leqslant t^{(2)}_{I_{\mathcal{P}_2}(i)+1}$$

Further we have $s^{(1)}_{I_{\mathcal{P}_1}(i)}\in\Big[t^{(1)}_{I_{\mathcal{P}_1}(i)},t^{(1)}_{I_{\mathcal{P}_1}(i)+1}\Big]$ and $s^{(2)}_{I_{\mathcal{P}_2}(i)}\in\Big[t^{(2)}_{I_{\mathcal{P}_2}(i)},t^{(2)}_{I_{\mathcal{P}_2}(i)+1}\Big]$ so that

$$t^{(1)}_{I_{\mathcal{P}_1}(i)}\leqslant s^{(1)}_{I_{\mathcal{P}_1}(i)}\leqslant t^{(1)}_{I_{\mathcal{P}_1}(i)+1} \text{ and } t^{(2)}_{I_{\mathcal{P}_2}(i)}\leqslant s^{(2)}_{I_{\mathcal{P}_2}(i)}\leqslant t^{(2)}_{I_{\mathcal{P}_2}(i)+1}$$

So $s^{(1)}_{I_{\mathcal{P}_1}(i)}-t_i\leqslant t^{(1)}_{I_{\mathcal{P}_1}(i)+1}-t^{(1)}_{I_{\mathcal{P}_1}(i)}\leqslant\mu(\mathbb{P}_1)<\delta$ and $t_i-s^{(1)}_{I_{\mathcal{P}_1}(i)}\leqslant t^{(1)}_{I_{\mathcal{P}_1}(i)+1}-t^{(1)}_{I_{\mathcal{P}_1}(i)}\leqslant\mu(\mathbb{P}_1)<\delta$ proving that

$$\big|s^{(1)}_{I_{\mathcal{P}_1}(i)}-t_i\big|<\delta$$

Also we have $s^{(2)}_{I_{\mathcal{P}_2}(i)}-t_i\leqslant t^{(2)}_{I_{\mathcal{P}_2}(i)+1}-t^{(2)}_{I_{\mathcal{P}_2}(i)}\leqslant\mu(\mathbb{P}_2)<\delta$ and $t_i-s^{(2)}_{I_{\mathcal{P}_2}(i)}\leqslant t^{(2)}_{I_{\mathcal{P}_2}(i)+1}-t^{(2)}_{I_{\mathcal{P}_2}(i)}\leqslant\mu(\mathbb{P}_2)<\delta$ proving

$$\big|s^{(2)}_{I_{\mathcal{P}_2}(i)}-t_i\big|<\delta$$

So that

$$\big|s^{(1)}_{I_{\mathcal{P}_1}(i)}-s^{(2)}_{I_{\mathcal{P}_2}(i)}\big|\leqslant\big|s^{(1)}_{I_{\mathcal{P}_1}(i)}-t_i\big|+\big|t_i-s^{(2)}_{I_{\mathcal{P}_2}(i)}\big|<\delta+\delta=2\cdot\delta=\delta'$$

Hence by [eq: 15.21] it follows that

$$\big\|f\big(s^{(1)}_{I_{\mathcal{P}_1}(i)}\big)-f\big(s^{(2)}_{I_{\mathcal{P}_2}(i)}\big)\big\|<\frac{\varepsilon}{b-a} \tag{15.23}$$

Now

$$\begin{array}{rl}
\left\|\sum_{i=1}^{n-1} (t_{i+1}-t_i)\cdot\left(f\left(s^{(1)}_{I_{\mathcal{P}_1}(i)}\right)-f\left(s^{(2)}_{I_{\mathcal{P}_2}(i)}\right)\right)\right\| & \leqslant \\
\sum_{i=1}^{n-1} \left\|(t_{i+1}-t_i)\cdot\left(f\left(s^{(1)}_{I_{\mathcal{P}_1}(i)}\right)-f\left(s^{(2)}_{I_{\mathcal{P}_2}(i)}\right)\right)\right\| & = \\
\sum_{i=1}^{n-1} \left(|t_{i+1}-t_i|\cdot\left\|\left(f\left(s^{(1)}_{I_{\mathcal{P}_1}(i)}\right)-f\left(s^{(2)}_{I_{\mathcal{P}_2}(i)}\right)\right)\right\|\right) & <_{\text{[eq: 15.23]}} \\
\sum_{i=1}^{n-1} |t_{i+1}-t_i|\cdot\frac{\varepsilon}{b-a} & = \\
\frac{\varepsilon}{b-a}\cdot\sum_{i=1}^{n-1} |t_{i+1}-t_i| & =_{t_i<t_{i+1}} \\
\frac{\varepsilon}{b-a}\cdot\sum_{i=1}^{n-1} (t_{i+1}-t_i) & =_{\text{[theorem: 11.21]}} \\
\frac{\varepsilon}{b-a}\cdot(t_n-t_1) & = \\
\frac{\varepsilon}{b-a}\cdot(b-a) & = \\
\varepsilon & =
\end{array}$$

□

Lemma 15.19. *Let $a,b\in\mathbb{R}$ with $a<b$, $\langle X,\|\|\rangle$ a normed space and $f:[a,b]\rightarrow X$ a function and $I_1,I_2\in X$ satisfying:*

1. *$\forall\varepsilon\in\mathbb{R}^+$ there exist a $\delta\in\mathbb{R}^+$ such that for every tagged partition $\mathbb{P}$ on $[a,b]$ with $\mu(\mathbb{P})<\delta$ we have $\|I_1-\mathcal{S}(f,\mathbb{P})\|<\varepsilon$*
2. *$\forall\varepsilon\in\mathbb{R}^+$ there exist a $\delta\in\mathbb{R}^+$ such that for every tagged partition $\mathbb{P}$ on $[a,b]$ with $\mu(\mathbb{P})<\delta$ we have $\|I_2-\mathcal{S}(f,\mathbb{P})\|<\varepsilon$*

then we have that $I_1=I_2$.

Proof. Assume that $I_1\neq I_2$ then we have $\varepsilon=\|I_1-I_2\|\in\mathbb{R}^+$. By (1),(2) there exist $\delta_1,\delta_2\in\mathbb{R}^+$ such that for every tagged partition $\mathbb{P}_1$ with $\mu(\mathbb{P}_1)<\delta_1$ and for every tagged partition $\mathbb{P}_2$ with $\mu(\mathbb{P}_2)<\delta_2$ we have $\|I_1-\mathcal{S}(f,\mathbb{P}_1)\|<\frac{\varepsilon}{2}$ and $\|I_2-\mathcal{S}(f,\mathbb{P}_2)\|<\frac{\varepsilon}{2}$. Using [theorem: 15.13] there exist a partition $\mathbb{P}$ such that $\mu(\mathbb{P})<\min(\delta_1,\delta_2)$, so that we have

$$\|I_1-\mathcal{S}(f,\mathbb{P})\|<\frac{\varepsilon}{2}\text{ and }\|I_1-\mathcal{S}(f,\mathbb{P})\|<\frac{\varepsilon}{2}$$

so that

$$\varepsilon=\|I_1-I_2\|\leqslant\|I_1-\mathcal{S}(f,\mathbb{P})\|+\|\mathcal{S}(f,\mathbb{P})-I_2\|<\frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon$$

which is a contradiction. Hence the assumption that $I_1\neq I_2$ is wrong and we must have $I_1=I_2$.

□

We are now ready for the definition of a Riemann Integrable function and the Riemann Integral of this function.

Definition 15.20. *Let $a,b\in\mathbb{R}$ with $a<b$, $\langle X,\|\|\rangle$ then a function $f:[a,b]\rightarrow X$ is* ***Riemann Integrable*** *if there exist a $I\in X$ such that $\forall\varepsilon\in\mathbb{R}$ there exist a $\delta\in\mathbb{R}^+$ such that for every tagged partition $\mathbb{P}$ on $[a,b]$ with $\mu(\mathbb{P})<\varepsilon$ we have $\|I-\mathcal{S}(f,\mathbb{P})\|<\varepsilon$. If f is Riemann integrable then by [lemma: 15.19] I is unique. We call this I [if it exist] the* ***Riemann integral*** *of f and not this as $\int_a^b f$. The set of graphs of Riemann integrable function is noted by $\mathcal{R}[[a,b],X]$ in other words:*

$$\mathcal{R}([a,b],X)=\{f\in X^{[a,b]}|f:[a,b]\rightarrow X\text{ is Riemann integrable}\}$$

The integral $\int_a^b$ is then a function from $\mathcal{R}[[a,b],X]$ to X defined by

$$\int_a^b:\mathcal{R}[[a,b],X]\rightarrow X\text{ where }\left(\int_a^b\right)(f)=\int_a^b f$$

Note 15.21. Another notation that you find in the literature is $\int_a^b f(x)\mathrm{dx}$ instead of $\int_a^b f$. One benefit of this notation is if there is a expression to calulate the value of a function based on a expression. For example instead of using

$$f:\mathbb{R}\to\mathbb{R} \text{ is defined by } f(x)=\frac{\cos(x)+1}{x+1}$$

and then refering to the integral using

$$\int_a^b f$$

we can simple use

$$\int_a^b \frac{\cos(x)+1}{x+1} dx$$

Also if we have a function $f:[a,b]\times Y\to\mathbb{R}$ by $f(t,x)=\cos(t)\cdot x$ and then defining for a $x\in Y$

$$g_x:[a,b]\to\mathbb{R} \text{ by } g_x(t)=f(t,x)=\cos(t).x$$

so that we can talk about $\int_a^b g_x$ we can simplify this by

$$\int_a^b \cos(t)\cdot x\, dt$$

Example 15.22. Let $a,b\in\mathbb{R}$ such that $a<b$, $\langle X,\|\|\rangle$ a normed space, $x\in X$ then

$$C_x:[a,b]\to X \text{ defined by } C_x(i)=x$$

is Riemann integrable and

$$\int_a^b C_x=(b-a)\cdot x$$

Proof. Let $\varepsilon\in\mathbb{R}^+$ and let $\mathbb{P}=\langle\{t_i\}_{i\in\{1,\ldots,n\}},\{s_i\}_{i\in\{1,\ldots,n-1\}}\rangle$ be any tagged partition on $[a,b]$ with $\mu(\mathbb{P})<1$ then we have:

$$\begin{aligned}
\mathcal{S}(f,\mathbb{P}) &= \sum_{i=1}^{n-1}(t_{i+1}-t_i)\cdot C_x(s_i)\\
&= \sum_{i=1}^{n-1}(t_{i+1}-t_i)\cdot x\\
&= \left(\sum_{i=1}^{n-1}(t_{i+1}-t_i)\right)\cdot x\\
&\underset{[\text{theorem: } 11.10}{=} (t_n-t_1)\cdot x\\
&= (b-a)\cdot x
\end{aligned}$$

so that

$$\|S(f,\mathbb{P})-(b-a)\cdot x\|=0<\varepsilon$$

Hence C_x is Riemann integrable and $\int_a^b C_x=(b-a)\cdot x$

□

Lemma 15.23. *Let $a,b\in\mathbb{R}$ with $a<b$ then there exist a sequence $\{\mathbb{P}_i\}_{i\in\mathbb{N}}$ of tagged partitions on $[a,b]$ such that*

$$\lim_{i\to\infty}\mu(\mathbb{P}_i)=0$$

Proof. Let $k\in\mathbb{N}$ and define

$$\{t^{(k)}\}_{i\in\{1,\ldots,k+1\}} \text{ by } t_i^{(k)}=a+(i-1)\cdot\frac{b-a}{k} \text{ and } \{s_i^{(k)}\}_{i\in\{1,\ldots,k\}} \text{ by } s_i^{(k)}=t_i^{(k)}$$

allowing us to define

$$\mathbb{P}=\left\langle\{t^{(k)}\}_{i\in\{1,\ldots,k+1\}},\{s_i^{(k)}\}_{i\in\{1,\ldots,k\}}\right\rangle$$

Let $i\in\{1,\ldots,(k+1)-1\}$ then we have

$$t_i^{(k)}=a+(i-1)\cdot\frac{b-a}{k}<a+(i-1)\cdot\frac{b-a}{k}+\frac{b-a}{k}=a+((i+1)-i)\cdot\frac{b-a}{k}=t_{i+1}^{(k)}$$

and

$$t_{i+1}^{(k)} - t_i^{(k)} = a + ((i+1) - i) \cdot \frac{b-a}{k} - \left(a + (i-1) \cdot \frac{b-a}{k}\right) = \frac{b-a}{k}$$

and

$$s_i^{(k)} \subseteq \left[t_{i+1}^{(k)} - t_i^{(k)}\right]$$

So that

$$\mu(\{t^{(k)}\}_{i \in \{1,\ldots,k+1\}}) = \max\left(\{t_{i+1}^{(k)} - t_i^{(k)} | i \in \{\{1,\ldots,(k+1)-1\}\}\}\right) = \max\left(\left\{\frac{b-a}{k}\right\}\right) = \frac{b-a}{k}$$

Hence $\forall k \in \mathbb{N}$ we have that $\mathbb{P}_k$ is a tagged partition on $[a,b]$ and

$$\lim_{k\to\infty} \mu(\mathbb{P}_k) = \lim_{k\to\infty} \frac{b-a}{k} \underset{\text{[theorem: 14.298]}}{=} (b-a) \cdot \lim_{k\to\infty} \frac{1}{k} \underset{\text{[example: 14.293]}}{=} 0$$ □

Theorem 15.24. *Let $a, b \in \mathbb{R}$ with $a < b$, $\langle X, \|\|\rangle$ a normed space and $f: [a,b] \to X$ a* ***continuous*** *function then the following are equivalent:*

1. *f is Riemann integrable with integral $\int_a^b f$*
2. *There exist a $I \in X$ such that for every sequence of tagged partitions $\{\mathbb{P}_i\}_{i\in\mathbb{N}}$ with*

$$\lim_{i\to\infty} \mu(\mathbb{P}_i) = 0$$

we have

$$\lim_{i\to\infty} \mathcal{S}(f, \mathbb{P}_i) = I$$

Furthermore we have $I = \int_a^b f$

Proof.

1 ⇒ 2. Let $\varepsilon \in \mathbb{R}^+$ then as f is Riemann integrable with integral $\int_a^b f$ there exist a $\delta \in \mathbb{R}^+$ such that for any tagged partition on $[a,b]$ with $\mu(\mathbb{P}) < \delta$ we have $\left\|\mathcal{S}(f,\mathbb{P}) - \int_a^b f\right\| < \varepsilon$. Let $\{\mathbb{P}_i\}_{i\in\mathbb{N}}$ be a familly of tagged partitions on $[a,b]$ such that $\lim_{i\to\infty} \mu(\mathbb{P}_i) = 0$. Then there exist a $N \in \mathbb{N}$ such that $\forall n \geqslant N$ we have $\mu(\mathbb{P}_n) = |\mu(\mathbb{P}_n) - 0| < \delta$ so that $\left\|\mathcal{S}(f,\mathbb{P}_n) - \int_a^b f\right\| < \varepsilon$ which proves that

$$\lim_{n\to\infty} S(f, \mathbb{P}_n) = \int_a^b f$$

2 → 1. Using [lemma: 15.231] there exists a sequence $\{\mathbb{P}_i\}_{i\in\mathbb{N}}$ of tagged partitions on $[a,b]$ such that $\lim_{i\to\infty} \mu(\mathbb{P}_i) = 0$. Then by the hypothese we have that there exist a $I \in X$ such that

$$\lim_{i\to\infty} \mathcal{S}(f, \mathbb{P}_i) = I$$

Let $\varepsilon \in \mathbb{R}^+$ then there exist a $N_1 \in \mathbb{N}$ such that

$$\forall n \geqslant N_1 \text{ we have } \|\mathcal{S}(f, \mathbb{P}_n) - I\| < \frac{\varepsilon}{2}$$

Using [lemma: 15.18] there exist a $\delta \in \mathbb{R}^+$ such that for every two partitions $\mathbb{P}'$, $\mathbb{P}''$ on $[a,b]$ with $\mu(\mathbb{P}') < \delta$ and $\mu(\mathbb{P}'') < \delta$ we have

$$\|\mathcal{S}(f, \mathbb{P}') - \mathcal{S}(f, \mathbb{P}'')\| < \frac{\varepsilon}{2} \tag{15.24}$$

As $\lim_{i\to\infty} \mu(\mathbb{P}_i)$ there exist a $N_2 \in \mathbb{N}$ such that $\forall n \geqslant N_2$ we have that

$$\mu(\mathbb{P}_n) = |\mu(\mathbb{P}_n) - 0| < \delta \tag{15.25}$$

Take $N = \max(N_1, N_2)$. Let $\mathbb{P}$ be a partition on $[a,b]$ with $\mu(\mathbb{P}) < \delta$ then

$$\begin{array}{rcl} \|\mathcal{S}(f,\mathbb{P}) - I\| & = & \|\mathcal{S}(f,\mathbb{P}) - \mathcal{S}(f,\mathbb{P}_N) + \mathcal{S}(f,\mathbb{P}_N) - I\| \\ & \leqslant & \|\mathcal{S}(f,\mathbb{P}) - \mathcal{S}(f,\mathbb{P}_N)\| + \|\mathcal{S}(f,\mathbb{P}_N) - I\| \\ & <_{N > N_2 \wedge \text{[eqs: 15.24, 15.25]}} & \frac{\varepsilon}{2} + \|\mathcal{S}(f,\mathbb{P}_N) - I\| \\ & <_{N > N_1} & \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\ & = & \varepsilon \end{array}$$

proving that f is Riemann integrable and by uniqueness [see lemma: 15.19] that $\int_a^b f = I$. □

We show now that every continuous function from $[a,b]$ to a Banach space is Riemann integrable.

Theorem 15.25. *Let $a,b\in\mathbb{R}$ with $a<b$, $\langle X,\|\|\rangle$ a Banach space and $f:[a,b]\rightarrow X$ a continous function then f is Riemann integrable. In other words we have $\mathcal{C}([a,b],X)\subseteq\mathcal{R}([a,b],X)$.*

Proof. Let $\{\mathbb{P}_i\}_{i\in\mathbb{N}}$ be a sequence of tagged partitions on $[a,b]$ with $\lim_{i\rightarrow\infty}\mu(\mathbb{P}_i)=0$. Let $\varepsilon\in\mathbb{R}^+$ then by [lemma: 15.18] there exist a $\delta\in\mathbb{R}^+$ such that for any two tagged partitions $\mathbb{P}$, $\mathbb{P}$ with $\mu(\mathbb{P})<\delta$ amd $\mu(\mathbb{P}')<\delta$ we have

$$\|\mathcal{S}(f,\mathbb{P})-\mathcal{S}(f,\mathbb{P}')\|<\varepsilon \tag{15.26}$$

As $\lim_{i\rightarrow\infty}\mu(\mathbb{P}_i)=0$ there exist a $N\in\mathbb{N}$ such that $\forall n\geqslant N$ we have $\mu(\mathbb{P}_n)=|\mu(\mathbb{P}_n)-0|<\delta$. So if $n,m\geqslant\mathbb{N}$ we have $\mu(\mathbb{P}_n)<\delta$ and $\mu(\mathbb{P}_m)<\delta$, so that by [eq: 15.26] $\|\mathcal{S}(f,\mathbb{P}_n)-\mathcal{S}(f,\mathbb{P}_m)\|<\varepsilon$. This proves that $\{\mathcal{S}(f,\mathbb{P}_i)\}_{i\in\mathbb{N}}$ is a Cauchy sequence in X. As X is a Banach space

$$\lim_{i\rightarrow\infty}S(f,\mathbb{P}_i)\text{ exist}$$

Hence by [theorem: 15.24] it follows that f is Riemann integrable. □

Theorem 15.26. *Let $a,b\in\mathbb{R}$ with $a<b$, $\langle\mathbb{R},||\rangle$ the normed space of the real numbers and*

$$f,g\in\mathcal{C}([a,b],\mathbb{R})\text{ [so that by [theorem: 15.25] } f,g \text{ are Riemann integrable] such that } \forall x\in[a,b]\text{ we have }f(x)\leqslant g(x)$$

then we have

$$\int_a^b f\leqslant\int_a^b g$$

Proof. Using [lemma: 15.231] there exists $\{\mathbb{P}_i\}_{i\in\mathbb{N}}$ such that $\lim_{i\rightarrow\infty}\mu(\mathbb{P}_i)=0$. Using [theorem: 15.16] we have

$$\forall i\in\mathbb{N}\ S(f,\mathbb{P}_i)\leqslant S(g,\mathbb{P}_i) \tag{15.27}$$

Further we have using [theorem: 15.24] that

$$\begin{aligned}\int_a^b f &\underset{\text{[theorem: 15.24]}}{=} \lim_{i\rightarrow\infty}\mathcal{S}(f,\mathbb{P}_i)\\ &\leqslant_{\text{[eq: 15.27] and [theorem: 14.306]}} \lim_{i\rightarrow\infty}\mathcal{S}(g,\mathbb{P}_i)\\ &\underset{\text{[theorem: 15.24]}}{=} \int_a^b g\end{aligned}$$

□

Up to now we have always assumed that $a<b$ in the definition of the Riemann integral, we remove now this restriction by extending the definition of the Riemann integral.

Definition 15.27. *Let $a,b\in\mathbb{R}$ with $a\leqslant b$, $\langle X,\|\|\rangle$ a normed space and $f:[a,b]\rightarrow X$ is Riemann integrable with integral $\int_a^b f$ if*

1. *If $a<b$ then we use [definition: 15.20]*
2. *If $a=b$ then f is Riemann integrable and $\int_a^b f=0$*

The set of Riemann integrable function graphs between is noted again by $\mathcal{R}([a,b],X)$ so that

$$\mathcal{R}([a,b],X)=\begin{cases}X^{[a,b]}\text{ if } a=b\\ \{f\in X^{[a,b]}|f:[a,b]\rightarrow X\text{ is Riemann integrable as defined in [definition: 15.20]}\}\end{cases}$$

We show now that $\int_a^b$ is a linear mapping.

Theorem 15.28. *Let $a,b\in\mathbb{R}$ with $a\leqslant b$, $\langle X,\|\|\rangle$ a normed space then*

$$\int_a^b:\mathcal{R}([a,b],X)\rightarrow X\text{ is a linear mapping}$$

In other words If $\alpha\in\mathbb{K}$ and $f,g\in\mathcal{R}([a,b],X)$ then we have that $\alpha\cdot f+g\in\mathcal{R}([a,b],X)$ and

$$\int_a^b(\alpha\cdot f+g)=\alpha\cdot\int_a^b f+\int_a^b g$$

Proof. We have to consider the following cases for a, b.

$\boldsymbol{a=b}$. Then by definition $\alpha \cdot f + g \in \mathcal{R}([a,b],X)$ and

$$\int_a^b (\alpha \cdot f + g) = 0 = 0 + 0 = \alpha \cdot \int_a^b f + \int_a^b g$$

$\boldsymbol{a<b}$. Let $\varepsilon \in \mathbb{R}^+$ then as f, g are integrable there exists $\delta_1, \delta_2 \in \mathbb{R}^+$ such that for every tagged partition $\mathbb{P}'$ on $[a,b]$ with $\mu(\mathbb{P}') < \delta_1$ we have

$$\left\| S(f, \mathbb{P}') - \int_a^b f \right\| < \frac{\varepsilon}{2 \cdot (1 + |\alpha|)} \tag{15.28}$$

and for every tagged partition $\mathbb{P}''$ on $[a,b]$ with $\mu(\mathbb{P}'') < \delta_2$ we have

$$\left\| S(f, \mathbb{P}'') - \int_a^b f \right\| < \frac{\varepsilon}{2} \tag{15.29}$$

Let now $\mathbb{P}$ be a tagged partition on $[a,b]$ such that $\mu(\mathbb{P}) < \min(\delta_1, \delta_2)$ then we have

$$\begin{array}{rcl} \left\| \mathcal{S}(\alpha \cdot f + g, \mathbb{P}) - \alpha \cdot \int_a^b f - \int_a^b g \right\| & \underset{\text{[theorem: 15.15]}}{=} & \\ \left\| \alpha \cdot \mathcal{S}(f, \mathbb{P}) + \mathcal{S}(g, \mathbb{P}) - \alpha \cdot \int_a^b f - \int_a^b g \right\| & \leqslant & \\ \left\| \alpha \cdot \mathcal{S}(f, \mathbb{P}) - \alpha \cdot \int_a^b f \right\| + \left\| \mathcal{S}(g, \mathbb{P}) - \int_a^b g \right\| & = & \\ |\alpha| \cdot \left\| \mathcal{S}(f, \mathbb{P}) - \int_a^b f \right\| + \left\| \mathcal{S}(g, \mathbb{P}) - \int_a^b g \right\| & <_{\text{[eqs: 15.28,15.29]}} & \\ |\alpha| \cdot \frac{\varepsilon}{2 \cdot (1 + |\alpha|)} + \frac{\varepsilon}{2} & < & \\ \frac{\varepsilon}{2} + \frac{\varepsilon}{2} & = & \\ \varepsilon & & \end{array}$$

Which proves that $\alpha \cdot f + g$ is Riemann integrable and that

$$\int_a^b (\alpha \cdot f + g) = \alpha \cdot \int_a^b f + \int_a^b g$$

□

Theorem 15.29. *Let $a, b \in \mathbb{R}$ with $a \leqslant b$, $\langle X, \|\|\rangle$ a Banach space and $f: [a,b] \rightarrow X$ a function then we have:*

1. *If $f \in \mathcal{R}([a,b],X)$ then $\forall L \in L(X, \mathbb{R})$ we have that $L \circ f$ is integrable and $\int_a^b (L \circ f) = L\left(\int_a^b f\right)$*
2. *If f is ontinuous [so that by [theorem: 15.25] f is Riemann integrable] then $\|f\|: [a.b] \rightarrow \mathbb{R}$ defined by $\|f\|(x) = \|f(x)\|$ is Riemann integrable and $\left\|\int_a^b f\right\| \leqslant \int_a^b \|f\|$*
3. *If $x \in X$ then $C_x: [a,b] \rightarrow X$ by $C_x(i) = x$ is Riemann integrable and $\int_a^b C_x = (b-a) \cdot x$*
4. *If f is continuous [hence by [theorem: 15.25] Riemann integrable] and $m \in \mathbb{R}$ such that $\forall x \in [a,b]$ we have $\|f(x)\| \leqslant m$ then*

$$\left\| \int_a^b f \right\| \leqslant m \cdot (b-a)$$

Proof. We have to look at the following cases for a, b:

$\boldsymbol{a=b}$. Then we have:

1. By definition $L \circ f$ is Riemann integrable and

$$\int_a^b (L \circ f) = 0 = L(0) = L\left(\int_a^b f\right)$$

2. By definition $\|f\|$ is Riemann integrable and

$$\int_a^b \|f\| = 0 = \|0\| = \left\| \int_a^b f \right\|$$

3. By definition C_x is Riemann integrable and

$$\int_a^b C_x = 0 = 0 \cdot x = (b-a) \cdot x$$

4. By definition f is Riemann integrable and

$$\left\| \int_a^b f \right\| = \|0\| = m \cdot 0 = m \cdot (b-a)$$

***a* < *b*.**

1. Let $\mathbb{P} = \langle \{t_i\}_{i \in \{1, \ldots, n\}}, \{s_i\}_{i \in \{1, \ldots, n-1\}}) \rangle$ be any tagged partition on $[a, b]$ then we have:

$$\begin{aligned} \mathcal{S}(L \circ f, \mathbb{P}) &= \sum_{i=1}^{n-1} (t_{i+1} - t_i) \cdot (L \circ f)(s_i) \\ &= \sum_{i=1}^{n-1} (t_{i+1} - t_i) \cdot L(f(s_i)) \\ &\underset{L \in L(X, \mathbb{R})}{=} L\left(\sum_{i=1}^{n-1} (t_{i+1} - t()) \cdot f(s_i) \right) \\ &= L(\mathcal{S}(f, \mathbb{P})) \end{aligned} \tag{15.30}$$

Let $\varepsilon \in \mathbb{R}^+$. As $L \in L(X, \mathbb{R})$ L is continuous, hence there exist a $\delta' \in \mathbb{R}^+$ such that if $\|x\| = \|x - 0\| < \delta'$ then

$$|L(x)| < \varepsilon \tag{15.31}$$

As f is Riemann integrable there exist a $\delta \in \mathbb{R}^+$ such that for every tagged partition $\mathbb{P}$ on $[a, b]$ with $\mu(\mathbb{P}) < \delta$ we have $\left\| \mathcal{S}(f, \mathbb{P}) - \int_a^b f \right\| < \delta'$. So

$$\begin{aligned} \left| \mathcal{S}(L \circ f, \mathbb{P}) - L\left(\int_a^b f \right) \right| &\underset{\text{[eq: 15.30]}}{=} \left| L(\mathcal{S}(f, \mathbb{P})) - L\left(\int_a^b f \right) \right| \\ &= \left| L\left(\mathcal{S}(f, \mathbb{P}) - \int_a^b f \right) \right| \\ &<_{\text{[eq: 15.31]}} \varepsilon \end{aligned}$$

proving that $L \circ f$ is Riemann integrable and that

$$\int_a^b (L \circ f) = L\left(\int_a^b f \right)$$

2. By [example: 14.149] $\|\|: X \to \mathbb{R}$ is continuous, hence as $\|f\| = \|\| \circ f$, it follows from [theorem: 14.138] it follows that

$$\|f\|: [a, b] \to \mathbb{R} \text{ is continue}$$

Hence using [theorem: 15.25] it follows that

$$\|f\| \text{ is Riemann integrable}$$

Let $\mathbb{P} = \langle \{t_i\}_{i \in \{1, \ldots, n\}}, \{s_i\}_{i \in \{1, \ldots, n-1\}}) \rangle$ be any tagged partition on $[a, b]$ then we have that

$$\begin{aligned} \|\mathcal{S}(f, \mathbb{P})\| &= \left\| \sum_{i=1}^{n-1} (t_{i+1} - t_i) \cdot f(s_i) \right\| \\ &\leqslant \sum_{i=1}^{n-1} \|(t_{i+1} - t_i) \cdot f(s_i)\| \\ &= \sum_{i=1}^{n-1} |t_{i+1} - t_i| \cdot \|f(s_i)\| \\ &\underset{t_i < t_{i+1}}{=} \sum_{i=1}^{n-1} (t_{i+1} - t_i) \cdot \|f(s_i)\| \\ &= \mathcal{S}(\|f\|, \mathbb{P}) \end{aligned} \tag{15.32}$$

Let $\{\mathbb{P}_i\}_{i\in\mathbb{N}}$ be a sequence of tagged paritition s on $[a,b]$ such that $\lim_{i\rightarrow\infty}\mu(\mathbb{P}_i)=0$. As $\|f\|$ is continuous and Riemann integrable it follows from [theorem: 15.24] that

$$\lim_{i\rightarrow\infty}\mathcal{S}(\|f\|,\mathbb{P}_i) \text{ exist and } \lim_{i\rightarrow\infty}\mathcal{S}(\|f\|,\mathbb{P}_i)=\int_a^b\|f\| \tag{15.33}$$

As also f is continuous and Riemann integrable it follows from [theorem: 15.24] that

$$\lim_{i\rightarrow\infty}\mathcal{S}(f,\mathbb{P}_i) \text{ exist and } \lim_{i\rightarrow\infty}\mathcal{S}(f,\mathbb{P}_i)=\int_a^b f \tag{15.34}$$

By [eq: 15.32] we have that $\forall i\in\mathbb{N}$ that $\|S(f,\mathbb{P}_i)\|\leqslant S(\|f\|,\mathbb{P}_i)$ so that we can apply [theorem: 14.303] together with [eqs: 15.33, 15.34] gives

$$\left\|\int_a^b f\right\|=\left\|\lim_{i\rightarrow\infty}\mathcal{S}(f,\mathbb{P}_i)\right\|\leqslant\lim_{i\rightarrow\infty}\mathcal{S}(\|f\|,\mathbb{P}_i)=\int_a^b\|f\|$$

3. This follows from [example: 15.22].
4. Let $\mathbb{P}=\langle\{t_i\}_{i\in\{1,\ldots,n\}},\{s_i\}_{i\in\{1,\ldots,n-1\}})\rangle$ be any tagged partition on $[a,b]$. then we have:

$$\begin{aligned}
\|\mathcal{S}(f,\mathbb{P})\| &= \left\|\sum_{i=1}^{n-1}(t_{i+1}-t_i)\cdot f(s_i)\right\| \\
&\leqslant \sum_{i=1}^{n-1}\|(t_{i+1}-t_i)\cdot f(s_i)\| \\
&= \sum_{i=1}^{n-1}|t_{i+1}-t_i|\cdot\|f(s_i)\| \\
&\underset{t_i<t_i}{=} \sum_{i=1}^{n-1}(t_{i+1}-t_i)\cdot\|f(s_I)\| \\
&\leqslant \sum_{i=1}^{n-1}((t_{i+1}-t_i)\cdot m) \\
&= \left(\sum_{i=1}^{n-1}(t_{i+1}-t_i)\right)\cdot m \\
&\underset{\text{[theorem: 11.10]}}{=} (t_n-t_1)\cdot m \\
&= (b-a)\cdot m
\end{aligned}$$

proving that

$$\|\mathcal{S}(f,\mathbb{P})\|\leqslant(b-a)\cdot m \tag{15.35}$$

Let $\{\mathbb{P}_i\}_{i\in\mathbb{N}}$ be a sequence of tagged partitions on $[a,b]$ with $\lim_{i\rightarrow\infty}\mu(\mathbb{P}_i)$ then as f is continuous we have that

$$\lim_{i\rightarrow\infty}\mathcal{S}(f,\mathbb{P}_i)=\int_a^b f$$

exist. Using the above together with [eq: 15.35] allows us to apply [theorem: 14.301] giving

$$\left\|\int_a^b f\right\|=\left\|\lim_{i\rightarrow\infty}\mathcal{S}(f,\mathbb{P}_i)\right\|\leqslant(b-a)\cdot m$$

□

Lemma 15.30. *Let $a,b\in\mathbb{R}$ with $a\leqslant b$, $c\in]a,b[$, $\mathbb{P}_1=\left\langle\left\{t_i^{(1)}\right\}_{i\in\{1,\ldots,n_1\}},\left\{s_i^{(1)}\right\}_{i\in\{1,\ldots,n_1-1\}}\right\rangle$ a tagged partition on $[a,c]$ and $\mathbb{P}_2=\left\langle\left\{t_i^{(2)}\right\}_{i\in\{1,\ldots,n_2\}},\left\{s_i^{(2)}\right\}_{i\in\{1,\ldots,n_2-1\}}\right\rangle$ a tagged partition on $[c,b]$ then for $\mathbb{P}_1\bigsqcup\mathbb{P}_2$ defined by*

$$\mathbb{P}_1\bigsqcup\mathbb{P}_2=\langle\{t_i\}_{i\in\{1,\ldots,n_1+n_2-1\}},\{s_i\}_{i\in\{1,\ldots,n_1+n_2-2\}}\rangle$$

where

$$\forall i\in\{1,\ldots,n_1+n_2-1\} \text{ we have } t_i=\begin{cases} t_i^{(1)} \text{ if } i\in\{1,\ldots,n_1-1\} \\ t_{i-n_1+1}^{(2)} \text{ if } i\in\{n_1,\ldots,n_1+n_2-1\}\end{cases}$$

and

$$\forall i\in\{1,\ldots,n_1+n_2\} \text{ we have } s_i=\begin{cases} s_i^{(1)} \text{ if } i\in\{1,\ldots,n_1-1\}\\ s_{i-n_1+1}^{(2)} \text{ if } i\in\{n_1,\ldots,n_1+n_2-2\}\end{cases}$$

then we have:

1. $\mathbb{P}_1\bigsqcup\mathbb{P}_2$ *is a tagged partition on* $[a,b]$
2. $\mu(\mathbb{P}_1\bigsqcup\mathbb{P}_2)=\max(\mu(\mathbb{P}_1),\mu(\mathbb{P}_2))$
3. *If* $\langle X,\|\|\rangle$ *is a normed space and* $f:[a,b]\rightarrow X$ *a function then*

$$\mathcal{S}\big(f,\mathbb{P}_1\bigsqcup\mathbb{P}_2\big)=\mathcal{S}[f_{|[a,c},\mathbb{P}_1]+\mathcal{S}[f_{[c,b]},\mathbb{P}_2]$$

Proof.

1. First we have that

$$t_1=t_1^{(1)}=a \tag{15.36}$$

Second, using [definition: 15.12] $1<n_2\Rightarrow 0\leqslant n_2-1$ we have $n_1=n_1\leqslant n_1+n_2-1$, so that

$$t_{n_1+n_2-1}=t_{(n_1+n_2-1)-n_1+1}^{(2)}=t_{n_2}^{(2)}=b \tag{15.37}$$

Third for $i\in\{1,\ldots,n_1+n_2-2\}$ we have either:

$\boldsymbol{i\in\{1,\ldots,n_1-2\}}$**.** Then

$$t_i=t_i^{(1)}<t_{i+1}^{(1)}\underset{i<n_1-2\leqslant n_1-1}{=}t_{i+1}$$

and

$$s_i=s_i^{(1)}\in\big[t_i^{(1)},t_{i+1}^{(1)}\big]\underset{i<n_1-2\leqslant n_1-1}{=}[t_i,t_{i+1}]$$

$\boldsymbol{i=n_1-1}$**.** Then

$$t_i=t_{n_1-1}=t_{n_1-1}^{(1)}<t_{n_1}^{(1)}=c=t_1^{(2)}=t_{n_1-n_1+1}^{(2)}=t_{n_1}=t_{i+1}$$

and as

$$s_{n_1-1}=s_{n_1-1}^{(1)}\in\big[t_{n_1-1}^{(1)},t_{n_1}^{(1)}\big]=\big[t_{n_1-1},t_{n_1}^{(1)}\big]=[t_{n_1-1},c]=\big[t_{n_1-1},t_1^{(2)}\big]=[t_{n_1-1},t_{n_1}]$$

we have

$$s_i\in[t_i,t_{i+1}]$$

$\boldsymbol{i\in\{n_1,\ldots,n_1+n_2-2\}}$**.** Then

$$t_i=t_{i-n_1+1}^{(2)}<t_{(i+1)-n_1+1}^{(2)}=t_{i+1}$$

and

$$s_i=s_{i-n_1+1}^{(2)}\in\big[t_{i-n_1+1}^{(2)},t_{(i+1)-n_1+1}^{(2)}\big]=[t_i,t_{i+1}]$$

Hence in all cases we have

$$t_i<t_{i+1} \text{ and } s_i\in[t_i,t_{i+1}] \tag{15.38}$$

So using [eqs: 15.36, 15.37 and 15.38] it follows that

$$\mathbb{P}_1\bigsqcup\mathbb{P}_2 \text{ is a tagged partition on } [a,b]$$

2. Note that

$$\begin{aligned}
\{t_{i+1}-t_i|i\in\{1,\ldots,n_1+n_2-2\}\} &=\\
\{t_{i+1}-t_i|i\in\{1,\ldots,n_1-2\}\}\bigcup\{t_{i+1}-t_i|i\in\{n_1-1\}\}\bigcup\{t_{i+1}-t_i|i\in\{n_1,\ldots,n_1+n_2-2\}\} &=\\
\big\{t_{i+1}^{(1)}-t_i^{(1)}|i\in\{1,\ldots,n_1-2\}\big\}\bigcup\big\{t_1^{(2)}-t_{n_1-1}^{(1)}\big\}\bigcup\big\{t_{(i+1)-n_1+1}^{(2)}-t_{i-n_1+1}^{(2)}|i\in\{n_1,\ldots,n_1+n_2-2\}\big\} &=\\
\big\{t_{i+1}^{(1)}-t_i^{(1)}|i\in\{1,\ldots,n_1-2\}\big\}\bigcup\big\{c-t_{n_1-1}^{(1)}\big\}\bigcup\big\{t_{(i+1)-n_1+1}^{(2)}-t_{i-n_1+1}^{(2)}|i\in\{n_1,\ldots,n_1+n_2-2\}\big\} &=\\
\big\{t_{i+1}^{(1)}-t_i^{(1)}|i\in\{1,\ldots,n_1-2\}\big\}\bigcup\big\{t_{n_1}^{(1)}-t_{n_1-1}^{(1)}\big\}\bigcup\big\{t_{(i+1)-n_1+1}^{(2)}-t_{i-n_1+1}^{(2)}|i\in\{n_1,\ldots,n_1+n_2-2\}\big\} &=\\
\big\{t_{i+1}^{(1)}-t_i^{(1)}|i\in\{1,\ldots,n_1-1\}\big\}\bigcup\big\{t_{(i+1)-n_1+1}^{(2)}-t_{i-n_1+1}^{(2)}|i\in\{n_1,\ldots,n_1+n_2-2\}\big\} &=\\
\big\{t_{i+1}^{(1)}-t_i^{(1)}|i\in\{1,\ldots,n_1-1\}\big\}\bigcup\big\{t_{i+1}^{(2)}-t_i^{(2)}|i\in\{1,\ldots,n_2-1\}\big\} &
\end{aligned}$$

so that

$$\begin{aligned}\mu(\mathbb{P}_1 \bigsqcup \mathbb{P}_2) &= \\ \max(\{t_{i+1}-t_i|i\in\{1,\ldots,n_1+n_2-2\}\}) &= \\ \max\left(\{t_{i+1}^{(1)}-t_i^{(1)}|i\in\{1,\ldots,n_1-1\}\}\bigcup\{t_{i+1}^{(2)}-t_i^{(2)}|i\in\{1,\ldots,n_2-1\}\}\right) &= \\ \max\left(\max\left(\{t_{i+1}^{(1)}-t_i^{(1)}|i\in\{1,\ldots,n_1-1\}\right),\max\left(\{t_{i+1}^{(2)}-t_i^{(2)}|i\in\{1,\ldots,n_2-1\}\}\right)\right) &\underset{[\text{theor:}3.64]}{=} \\ \max(\mu(\mathbb{P}_1),\mu(\mathbb{P}_2))\end{aligned}$$

3. We have

$$\begin{aligned}\mathcal{S}(f,\mathbb{P}) &= \\ \sum_{i=1}^{n_1+n_2-2}(t_{i+1}-t_i)\cdot f(s_i) &= \\ \sum_{i=1}^{n_1-2}(t_{i+1}-t_i)\cdot f(s_i)+\sum_{i=n_1-1}^{n_1-1}(t_{i+1}-t_i)\cdot f(s_i)+\sum_{i=n_1}^{n_1+n_2-2}(t_{i+1}-t_i)\cdot f(s_i) &= \\ \sum_{i=1}^{n_1-2}\left(t_{i+1}^{(1)}-t_i^{(1)}\right)\cdot f\left(s_i^{(1)}\right)+(t_{n_1}-t_{n_1-1})\cdot f(s_{n_1-1})+\sum_{i=n_1}^{n_1+n_2-2}\left(t_{(i+1)-n_1+1}^{(2)}-t_{i-n_1+1}^{(2)}\right)\cdot f\left(s_{i-n_1+1}^{(2)}\right) &= \\ \sum_{i=1}^{n_1-2}\left(t_{i+1}^{(1)}-t_i^{(1)}\right)\cdot f\left(s_i^{(1)}\right)+\left(t_1^{(2)}-t_{n_1-1}^{(1)}\right)\cdot f\left(s_{n_1-1}^{(1)}\right)+\sum_{i=1}^{n_2-1}\left(t_{i+1}^{(2)}-t_i^{(2)}\right)\cdot f\left(s_i^{(2)}\right) &= \\ \sum_{i=1}^{n_1-2}\left(t_{i+1}^{(1)}-t_i^{(1)}\right)\cdot f\left(s_i^{(1)}\right)+\left(c-t_{n_1-1}^{(1)}\right)\cdot f\left(s_{n_1-1}^{(1)}\right)+\sum_{i=1}^{n_2-1}\left(t_{i+1}^{(2)}-t_i^{(2)}\right)\cdot f\left(s_i^{(2)}\right) &= \\ \sum_{i=1}^{n_1-2}\left(t_{i+1}^{(1)}-t_i^{(1)}\right)\cdot f\left(s_i^{(1)}\right)+\left(t_{n_1}^{(1)}-t_{n_1-1}^{(1)}\right)\cdot f\left(s_{n_1-1}^{(1)}\right)+\sum_{i=1}^{n_2-1}\left(t_{i+1}^{(2)}-t_i^{(2)}\right)\cdot f\left(s_i^{(2)}\right) &= \\ \sum_{i=1}^{n_1-1}\left(t_{i+1}^{(1)}-t_i^{(1)}\right)\cdot f\left(s_i^{(1)}\right)+\sum_{i=1}^{n_2-1}\left(t_{i+1}^{(2)}-t_i^{(2)}\right)\cdot f\left(s_i^{(2)}\right) &= \\ \sum_{i=1}^{n_1-1}\left(t_{i+1}^{(1)}-t_i^{(1)}\right)\cdot f_{|[a,c]}\left(s_i^{(1)}\right)+\sum_{i=1}^{n_2-1}\left(t_{i+1}^{(2)}-t_i^{(2)}\right)\cdot f_{|[c,b]}\left(s_i^{(2)}\right) &= \\ \mathcal{S}(f_{|[a,c]},\mathbb{P}_1)+\mathcal{S}(f_{|[c,b]},\mathbb{P}_2)\end{aligned}$$

□

We prove now a theorem that will be critical to prove later the Fundamental Theorem of Calculus.

Theorem 15.31. *Let $a,b\in\mathbb{R}$ with $a\leqslant b$, $c\in[a,b]$, $\langle X,\|\|\rangle$ a Banach space and $f\in\mathcal{C}([a,b],X)$ [so that by [theorem: 15.25 is Riemann integrable] then*

$$f_{|a,c}:[a,c]\to X,\ f_{|[c,b]}:[c,b]\to X \text{ are Riemann integrable}$$

and

$$\int_a^b f=\int_a^c f_{|[a,c]}+\int_c^b f_{|[c,b]}$$

Note 15.32. *To simplify notation we use the convention that if $f\in\mathcal{C}([a,b],X)$ and $x\in[a,b]$ that $\int_a^x f$ is the same as $\int_a^x f_{|[a,x]}$ and $\int_x^b f$ is the same as $\int_x^b f_{|[x,b]}$. So that we can write*

$$\int_a^b f=\int_a^c f+\int_c^b f$$

where the theorem ensures that $\int_a^c f$, $\int_c^b f$ exist and are well defined.

Proof. We have the following cases to consider for a,b,c:

$\boldsymbol{a=b}$**.** Then by definition we have that f, $f_{|[a,c]}$, $f_{|[c,b]}$ are Riemann integrable and

$$\int_a^b f=0=0+0=\int_a^c f_{|[a,c]}+\int_c^b f_{|[c,b]}$$

$\boldsymbol{a<b\wedge a=c}$. Then $f_{|[a,c]}$ is Riemann integrable by definition, $[c,b]=[a,b]$ and

$$\int_a^b f = 0+\int_a^b f = 0+\int_c^b f_{|c,b} = \int_a^c f_{|[a,c]} + \int_c^b f_{|c,b}$$

$\boldsymbol{a<b\wedge b=c}$. Then $f_{|[b,c]}$ is Riemann integrable by definition, $[a,c]=[a,b]$ and

$$\int_a^b f = \int_a^c f_{|[a,c]} = \int_a^c f_{|[a,c]} + 0 = \int_a^c f_{|[a,c]} + \int_c^b f_{|[c,b]}$$

$\boldsymbol{a<c<b}$. As $f:[a,b]\to X$ is continuous it follows from [theorem: 14.133] that $f_{|[a,c]}:[a,c]\to X$ and $f_{|[c,b]}:[c,b]\to X$ are continuous. Using [theorem: 15.25] we conclude that $f_{|[a,c]}$ and $f_{|[c,b]}$ are Riemann integrable. By [lemma: 15.231] there exists a sequence of tagged partitions $\{\mathbb{P}_i^{(1)}\}_{i\in\mathbb{N}}$ on $[a,c]$ and a sequence of tagged partitions $\{\mathbb{P}_i^{(2)}\}_{i\in\mathbb{N}}$ on $[c,b]$ such that

$$\lim_{i\to\infty}\mu(\mathbb{P}_i^{(1)})=0 \text{ and } \lim_{i\to\infty}\mu(\mathbb{P}_i^{(2)})=0$$

Let $\varepsilon\in\mathbb{R}^+$ then there exists $N_1,N_2\in\mathbb{N}$ such that for $n\geqslant N_1$ we have

$$\mu(\mathbb{P}_n^1)=|\mu(\mathbb{P}_n^{(1)})-0|<\varepsilon \text{ and } \mu(\mathbb{P}_n^{(2)})=|\mu(\mathbb{P}_n^{(2)})-0|<\varepsilon$$

So if $n\geqslant\max(N_1,N_2)$

$$|\mu(\mathbb{P}_n^{(1)}\bigsqcup\mathbb{P}_n^{(2)})-0|=\mu(\mathbb{P}_n^{(1)}\bigsqcup\mathbb{P}_n^{(2)})=\max(\mu(\mathbb{P}_n^{(1)}),\max(\mu(\mathbb{P}_n^{(2)})))<\varepsilon$$

which proves that

$$\lim_{i\to\infty}\mu(\mathbb{P}_n^{(1)}\bigsqcup\mathbb{P}_n^{(2)})=0$$

Now using [theorem: 15.24] and the fact that f, $f_{|[a,c]}$ and $f_{[c,b]}$ are Riemann integrable and continuous, it follows that

$$\lim_{i\to\infty}\mathcal{S}(f,\mathbb{P}_i^{(1)}\bigsqcup\mathbb{P}_i^{(2)}),\ \lim_{i\to\infty}\mathcal{S}(f_{|[a,c]},\mathbb{P}_i^{(1)}) \text{ and } \lim_{i\to\infty}\mathcal{S}(f_{|[c,b]},\mathbb{P}_i^{(2)}) \text{ exists} \tag{15.39}$$

and

$$\int_a^b f=\lim_{i\to\infty}\mathcal{S}(f,\mathbb{P}_i^{(1)}\bigsqcup\mathbb{P}_i^{(2)}),\int_a^c f_{|[a,c]}=\lim_{i\to\infty}\mathcal{S}(f_{|[a,c]},\mathbb{P}_i^{(1)}),\int_c^b f_{|[c,b]}=\lim_{i\to\infty}\mathcal{S}(f_{|[c,b]},\mathbb{P}_i^{(2)}) \tag{15.40}$$

Further we have $\forall i\in\mathbb{N}$ that

$$\mathcal{S}(f,\mathbb{P}_i^{(1)}\bigcup\mathbb{P}_i^{(2)}) \underset{[\text{lemma: } 15.30]}{=} \mathcal{S}(f_{|[a,b]},\mathbb{P}_i^{(1)})+\mathcal{S}(f_{|[c,b]},\mathbb{P}_i^{(2)}) \tag{15.41}$$

Finally using [theorem: 14.297] together with [eqs: 15.39, 15.40 and 15.41] proves that

$$\int_a^b f = \int_a^c f_{|[a,c]} + \int_c^b f_{|[c,b]}$$

□

Chapter 16
Differentiation in Normed spaces

16.1 Fréchet differentiability

In this subsection we introduce the Fréchet differential of a function, which is a extension of the derivative of a function. To avoid evaluating a function outside it's domain the following definition is introduced.

Definition 16.1. *Let $\langle X, \|\|\rangle$ be a normed space, $U \subseteq X$ and $x \in U$ then*

$$U_x = \{h \in X | x + h \in U\}$$

***Note* 16.2.** *Note as $x + 0 = x \in U$ we have*

$$0 \in U_x$$

The following is a equivalent definition for U_x.

Theorem 16.3. *Let $\langle X, \|\|\rangle$ be a normed space, $U \subseteq X$ and $x \in U$ then*

$$U_x = -x + U = U - x$$

where

$$-x + U = \{-x + u | i \in U\} \text{ and } U - x = \{u - x | u \in U\}$$

Proof. We have

$$\begin{array}{rcl} h \in U_x & \Leftrightarrow & x + h \in U \\ & \Leftrightarrow & \exists u \in U \text{ with } x + h = u \\ & \Leftrightarrow & \exists u \in U \text{ with } h = (-x) + u \\ & \Leftrightarrow & h \in (-x) + U \\ h \in (-x) + U & \Leftrightarrow & \exists u \in U \text{ with } h = (-x) + u \\ & \Leftrightarrow & \exists u \in U \text{ with } h = u + (-x) \\ & \Leftrightarrow & h \in U + (-x) \end{array}$$

□

Theorem 16.4. *Let $\langle X, \|\|\rangle$ be a normed space, $x \in X$ and $\delta \in \mathbb{R}^+$ then $B_{\|\|}(0, \delta) = B_{\|\|}(x, \delta)_x$.*

Proof.

$$\begin{array}{rcl} x \in B_{\|\|}(x, \delta)_x & \Leftrightarrow & x + h \in B_{\|\|}(x, \delta) \\ & \Leftrightarrow & x \in X \wedge \|(x + h) - x\| < \delta \\ & \Leftrightarrow & x \in X \wedge \|h\| < \delta \\ & \Leftrightarrow & h \in B_{\|\|}(0, \delta) \end{array}$$

□

16.1.1 Limit of a function

First we introduce the concept of the limit of a function at a point.

Definition 16.5. *Let $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$ be normed spaces, $U \subseteq X$, $x \in X$ and $f: U \to Y$ a function then we say that* ***f converges to y at x*** *if $\forall \varepsilon \in \mathbb{R}^+$ there exist a $\delta \in \mathbb{R}^+$ such that $\forall z \in U$ with $0 < \|x - z\|_X < \delta$ we have $\|f(z) - y\|_Y < \varepsilon$.*

Note 16.6. As $0<\|x-z\|_X$ we have that $x\neq z$, so in the above definition we never evaluate f at x, hence x must not be part of the domain f.

We show now that the y in the above is unique in a normed space (not in pseudo normed spaces).

Lemma 16.7. *Let $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$ be normed spaces, $U\subseteq X$, $x\in X$ and $f:U\rightarrow Y$ a function such that f converges to y_1 and y_2 at x then $y_1=y_2$.*

Proof. Assume that $y_1\neq y_2$ then $\|y_1-y_2\|_Y\in\mathbb{R}^+$. As f converges to y_1 and y_2 at x there exist $\delta_1,\delta_2\in\mathbb{R}^+$ such that $\forall z\in U$ with $0<\|x-z\|_X<\delta_1$ we have $\|f(z)-y_1\|_Y<\frac{\|y_1-y_2\|_Y}{2}$ and $\forall z\in U$ with $\|x-z\|_X<\delta_2$ we have $\|f(z)-y_2\|_Y<\frac{\|y_1-y_2\|_Y}{2}$. So if $z\in U$ satisfies $0<\|x-z\|_X<\min(\delta_1,\delta_2)$ then

$$\|y_1-y_2\|_Y\leqslant\|y_1-f(z)\|_Y+\|f(z)-y_2\|_Y<\frac{\|y_1-y_2\|_Y}{2}+\frac{\|y_1-y_2\|_Y}{2}=\|y_1-y_2\|_Y$$

giving the contradiction $\|y_1-y_2\|_Y<\|y_1-y_2\|_Y$. So we must have that $y_1=y_2$. □

The above lemma ensure that the following definition is well defined.

Definition 16.8. *Let $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$ be normed spaces, $U\subseteq X$, $x\in X$ and $f:U\rightarrow Y$ a function so that f converges to $y\in Y$ at x then this **unique** y [see lemma: 16.7] is noted as*

$$\lim_{\substack{z\rightarrow x\\ U}} f(z)$$

In other words if we say that

$$\lim_{\substack{z\rightarrow x\\ U}} f(z)=y \text{ exists [or converges]}$$

then $\forall\varepsilon\in\mathbb{R}^+$ there exist a $\delta\in\mathbb{R}^+$ such that $\forall z\in U$ with $0<\|x-z\|_X<\delta$ we have $\|f(z)-y\|_Y<\varepsilon$.

We have the following characterization for continuity **at a point in its domain**.

Theorem 16.9. *Let $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$ be normed spaces, $U\subseteq X$, $x\in U$ and $f:U\rightarrow Y$ a function then*

$$f \text{ is continuous at } x\Leftrightarrow\lim_{\substack{z\rightarrow x\\ U}} f(z)=f(x)$$

Proof. This follows from [definition: 16.8] and [theorem: 14.128] □

16.1.2 Classical derivative of a function

In this subsection we introduce the calculus concepts of differentiability and derivative. First we need a little lemma to ensure that the derivative of a function is unique.

Lemma 16.10. *Let $\langle\mathbb{K},||\rangle$ be the normed space of real or complex numbers, $U\subseteq\mathbb{K}$ a open set, $x\in U$, $\langle Y,\|\|\rangle$ a normed space and $f:U\rightarrow Y$ a function. Assume that there exists y_1, y_2 such that*

$$\forall\varepsilon\in\mathbb{R}^+\ \exists\delta_1\in\mathbb{R}^+ \text{ such that } \forall h\in U_x \text{ with } 0<|h|<\delta_1 \text{ we have } \left\|\frac{f(x+h)-f(x)}{h}-y_1\right\|<\varepsilon$$

$$\forall\varepsilon\in\mathbb{R}^+\ \exists\delta_2\in\mathbb{R}^+ \text{ such that } \forall h\in U_x \text{ with } 0<|h|<\delta_2 \text{ we have } \left\|\frac{f(x+h)-f(x)}{h}-y_2\right\|<\varepsilon$$

then

$$y_1=y_2$$

Proof. As $U_x\underset{\text{[eq: 16.3]}}{=}(-x)+U$ is open [see theorem: 14.109] and $0\in U_x$ there exist a $\rho\in\mathbb{R}^+$ such that $0\in B_{\|\|_X}(0,\rho)\subseteq U_x$. Assume that $y_1\neq y_2$ then $\|y_1-y_2\|\in\mathbb{R}^+$ hence there exists $\delta_1,\delta_2\in\mathbb{R}^+$ such that

$$\forall\varepsilon\in\mathbb{R}^+\ \exists\delta_1\in\mathbb{R}^+ \text{ such that } \forall h\in U_x \text{ with } 0<|h|<\delta_1 \text{ we have } \left\|\frac{f(x+h)-f(x)}{h}-y_1\right\|<\frac{\|y_1-y_2\|}{2}$$

$$\forall\varepsilon\in\mathbb{R}^+\ \exists\delta_2\in\mathbb{R}^+ \text{ such that } \forall h\in U_x \text{ with } 0<|h|<\delta_2 \text{ we have } \left\|\frac{f(x+h)-f(x)}{h}-y_2\right\|<\frac{\|y_1-y_2\|}{2}$$

Take $h_0 = \frac{\min(\rho, \delta_1, \delta_2)}{2} \in \mathbb{R}^+$ then $h_0 \in B_{\|\|_X}(0, \rho) \subseteq U_x$ and $0 < |h_0| < \delta_1, \delta_2$ so that

$$\begin{aligned}\|y_1 - y_2\| &= \left\|\frac{f(x+h_0) - f(x)}{h_0} - y_2 - \left(\frac{f(x+h_0) - f(x)}{h_0} - y_1\right)\right\| \\ &\leqslant \left\|\frac{f(x+h_0) - f(x)}{h_0} - y_2\right\| + \left\|\frac{f(x+h_0) - f(x)}{h_0} - y_1\right\| \\ &< \frac{\|y_2 - y_1\|}{2} + \frac{\|y_2 - y_1\|}{2} \\ &= \|y_1 - y_2\|\end{aligned}$$

which is a contradiction. So the assumption $y_1 \neq y_2$ is wrong and we must have $y_1 = y_2$. □

Note 16.11. The above proof of [lemma: 16.10] requires that U is a open set. Hence, as [lemma: 16.10] is used in the definition below of the derivative to ensure uniqueness of the derivative at a point, the domain of functions with a derivative must be open.

Definition 16.12. *Let $\langle \mathbb{K}, \| \rangle$ be the normed space of real or complex numbers, a open set $U \subseteq \mathbb{K}$, $x \in U$, $\langle Y, \|\|\rangle$ a normed space and $f: U \to Y$ a function. Then* ***f has a derivative at x*** *if there exist a $y \in Y$ such that*

$$\forall \varepsilon \in \mathbb{R}^+ \ \exists \delta \in \mathbb{R}^+ \text{ so that } \forall h \in U_x \text{ with } 0 < |h| < \delta \ \left\|\frac{f(x+h) - f(x)}{h} - y\right\| < \varepsilon$$

Using [lemma: 16.10] this y is unique and it is called the ***derivative of f at x*** *and noted by f'_x. In other words if f has a derivative f'_x then we have:*

$$\forall \varepsilon \in \mathbb{R}^+ \ \exists \delta \in \mathbb{R}^+ \text{ such that } \forall h \in U_x \text{ with } 0 < |h| < \delta \text{ we have } \left\|\frac{f(x+h) - f(x)}{h} - f'(x)\right\| < \varepsilon$$

The above is what is called the $\varepsilon - \delta$ definition of the derivative of a function. We use this because it is easier to construct exact proofs about derivatives without having to rely on limits in the proofs. However there is a equivalent definition (used in many books) based on limits.

Theorem 16.13. *Let $\langle \mathbb{K}, \| \rangle$ be the normed space of real [complex numbers], a open set $U \subseteq \mathbb{K}$, $x \in U$, $\langle Y, \|\|\rangle$ a normed space and $f: U \to Y$ a function then we have*

$$f \text{ has a derivative } f'(x) \text{ at } x$$

$$\Updownarrow$$

$$\lim_{h \underset{U_x \setminus \{0\}}{\to} 0} \frac{f(x+h) - f(x)}{h} = f'(x) \text{ exist}$$

***Note* 16.14.** *Here* $\lim_{h \underset{U_x \setminus \{0\}}{\to} 0} \frac{f(x+h) - f(x)}{h}$ *is a shorthand for* $\lim_{h \underset{U_x \setminus \{0\}}{\to} 0} \Delta(h)$ *where*

$$\Delta(h): U_x \setminus \{0\} \to Y \text{ is defined by } \Delta(h) = \frac{f(x+h) - f(x)}{h}$$

Proof.

⇒. Let $\varepsilon \in \mathbb{R}^+$. As f has a derivative f'_x at x there exist a $\delta \in \mathbb{R}^+$ such that $\forall h \in U_x$ with $0 < |h| < \delta$ we have

$$\left\|\frac{f(x+h) - f(x)}{h} - f'_x\right\| < \varepsilon \tag{16.1}$$

Define now

$$\Delta: U_x \setminus \{0\} \to Y \text{ by } \Delta(h) = \frac{f(x+h) - f(x)}{h}$$

then if $h \in U_x \setminus \{0\}$ and $0 < |h| = |0 - h| < \delta$ we have

$$\|\Delta(h) - f'_x\| = \left\|\frac{f(x+h) - f(x)}{h} - f'_x\right\| <_{[\text{eq: } 16.1]} \varepsilon$$

proving that

$$\lim_{h \underset{U_x \setminus \{0\}}{\to}} \Delta(h) = f'_x$$

⇐. Let $\varepsilon \in \mathbb{R}^+$. As for

$$\Delta: U_x \setminus \{0\} \rightarrow Y \text{ by } \Delta(h) = \frac{f(x+h) - f(x)}{h}$$

we have that

$$\lim_{h \underset{U_x \setminus \{0\}}{\rightarrow}} \Delta(h) = L \text{ exists}$$

there exist a $\delta \in \mathbb{R}^+$ such that if $h \in U_x$ and $0 < |0 - h| < \delta$ that

$$\|\Delta(h) - L\| = \left\| \frac{f(x+h) - f(x)}{h} - L \right\| < \varepsilon$$

Hence if $h \in U_x$ with $0 < |h| < \delta$ we have that $\left\| \frac{f(x+h) - f(x)}{h} - L \right\| < \varepsilon$ proving that f has a derivative $f'_x = L$. □

Example 16.15. Let $\langle \mathbb{K}, \| \rangle$ be the normed space of real or complex numbers, $\langle X, \|\| \rangle$ a normed space, $x \in X$ then

$$(*{\cdot}x): \mathbb{K} \rightarrow X \text{ defined by } (*{\cdot}x)(t) = t \cdot x$$

has a derivative at every $t \in \mathbb{R}$ with $(*{\cdot}x)'_x = x$

Proof. Let $\varepsilon \in \mathbb{R}^+$ then for $0 < |h| < 1$ we have

$$\left\| \frac{(*, \cdot x)(t+h) - (*, \cdot x)(t)}{h} - x \right\| = \left\| \frac{(t+h) \cdot x - t \cdot x}{h} - x \right\| = \left\| \frac{h \cdot x}{h} - x \right\| = 0 < \varepsilon$$

so that $(*{\cdot}x)$ has x as its derivative at t. □

Note 16.16. We use f'_x to note the derivative of f at x. Many books use instead the notation $f'_{|x}$ or $f'(x)$. The problem with $f'_{|x}$ is that it can be confused with the restriction of a function f' to x and $f'(x)$ can also be confused with the value of a function f' at x. The derivative is however not a function but a value of Y. If f has a dervative at every x in U then we can define a function

$$f': U \rightarrow Y \text{ by } f'(x) = f'_x$$

So using the notation convention of this book we can make a distinction between the derivative at a point, which is a value of Y, and the derivative **function** which only exist if a derivative exist at every point in U. If the derivative function exist then by its definition $f'(x) = f'_x$.

If we want to extend the definition of a derivative of a function on $U \subseteq \mathbb{K}$ to a function whose domain is a open subset of a normed space we run in the following problem. If $\langle X, \|\|_X \rangle$, $\langle Y, \|\|_Y \rangle$ are normed spaces, $U \subseteq X$ a open set, $x \in U$ and $f: U \rightarrow Y$ a function then as $f(x+h) - f(x) \in Y$ and $h \in X$ the quotient $\frac{f(x+h) - f(x)}{h}$ is not defined so that $\left\| \frac{f(x+h) - f(x)}{h} - f'(x) \right\|_Y$ is ill-defined. To solve this, note that in the case where $\langle X, \|\|_X \rangle = \langle \mathbb{K}, \| \rangle$ we can multiply both sides of $\left\| \frac{f(x+h) - f(x)}{h} - f'(x) \right\| < \varepsilon$ by $|h|$ to get

$$|h| \cdot \left\| \frac{f(x+h) - f(x)}{h} - f'(x) \right\| < \varepsilon \cdot |h|$$

Then as

$$\begin{aligned} |h| \cdot \left\| \frac{f(x+h) - f(x)}{h} - f'(x) \right\| &= \left\| h \cdot \left(\frac{f(x+h) - f(x)}{h} - f'(x) \right) \right\| \\ &= \|f(x+h) - f(x) - f'(x) \cdot h\| \end{aligned}$$

we have

$$\|f(x+h) - f(x) - f'(x) \cdot h\| < \varepsilon \cdot |h|$$

Define $L: \mathbb{K} \rightarrow Y$ by $L(h) = f'(x) \cdot h$ then by [example: 14.183] we have that $L \in L(\mathbb{K}, Y)$ hence we can rewrite the above as

$$\|f(x+h) - f(x) - L(h)\| < \varepsilon \cdot |h|$$

This can extended to more general cases where the normed spaces is $\langle X, \|\|_X \rangle$ instead of $\langle \mathbb{K}, \| \rangle$ if we change the above expression in

$$\|f(x+h) - f(x) - L(h)\|_Y < \varepsilon \cdot \|h\|_X$$

This is the idea behind Fréchet differentiability of a function which is covered in the next subsection.

16.1.3 Fréchet differential of a function

We are now ready to define the Fréchet differentiability of a function. First we need a lemma that ensures that the differential of a function [if it exist] is unique.

Lemma 16.17. *Let $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$ be normed spaces, U a open set in X , $x\in U$, $f:U\to Y$ a function and $L_1,L_2\in L(X,Y)$ such that*

$$\forall\varepsilon\in\mathbb{R}^+\ \exists\delta\in\mathbb{R}^+ \text{ such that } \forall h\in U_x \text{ with } \|h\|_X<\delta \text{ we have } \|f(x+h)-f(x)-L_1(h)\|_Y\leqslant\varepsilon\cdot\|h\|_X$$

and

$$\forall\varepsilon\in\mathbb{R}^+\ \exists\delta\in\mathbb{R}^+ \text{ such that } \forall h\in U_x \text{ with } \|h\|_X<\delta \text{ we have } \|f(x+h)-f(x)-L_2(h)\|_Y\leqslant\varepsilon\cdot\|h\|_X$$

then we have that

$$L_1=L_2$$

Proof. Assume that $L_1\neq L_2$ then there exist a $h_0\in X$ such that $L_1(h_0)\neq L_2(h_0)$ hence we have that

$$\varepsilon=\|L_1(h_0)-L_2(h_0)\|_Y\in\mathbb{R}^+$$

Now using the hypothesis there exists $\delta_1,\delta_2\in\mathbb{R}^+$ such that

$$\forall k\in U_x \text{ with } \|k\|_X<\delta_1 \text{ we have } \|f(x+k)-f(x)-L_1(k)\|_Y\leqslant\frac{\varepsilon}{2\cdot(\|h_0\|_X+1)}\cdot\|k\|$$

and

$$\forall k\in U_x \text{ with } \|k\|_X<\delta_2 \text{ we have } \|f(x+k)-f(x)-L_2(k)\|_Y\leqslant\frac{\varepsilon}{2\cdot(\|h_0\|_X+1)}\cdot\|k\|$$

As $U_x \underset{[\text{eq: } 16.3]}{=} (-x)+U$ is open [see theorem: 14.109] and $0\in U_x$ there exist a $\delta_3\in\mathbb{R}^+$ such that $0\in B_{\|\|_X}(0,\delta_3)\subseteq U_x$. Take $\delta=\min(\delta_1,\delta_2,\delta_3)$ and define $h_1=\frac{\delta}{(\|h\|_X+1)}\cdot h_0$ then

$$\|h_1\|_X=\left\|\frac{\delta}{(\|h_0\|_X+1)}\cdot h_0\right\|_X=\delta\cdot\frac{\|h_0\|_X}{\|h_0\|_X+1}<\delta\leqslant\delta_1,\delta_2,\delta_3$$

so that $h_1\in U_x$ and

$$\begin{aligned}
\frac{\delta}{\|h_0\|_X+1}\cdot\|L_1(h_0)-L_2(h_0)\|_Y &= \left\|\frac{\delta}{\|h_0\|_X+1}\cdot L_1(h_0)-\frac{\delta}{\|h_0\|_X+1}\cdot L_2(h_0)\right\|_Y\\
&= \left\|L_1\left(\frac{\delta}{\|h_0\|_X+1}\cdot h_0\right)\right\|_Y+\left\|L_2\left(\frac{\delta}{\|h_0\|_X+1}\cdot h_0\right)\right\|_Y\\
&= \|L_1(h_1)-L_2(h_1)\|_Y\\
&= \|f(x+h_1)-f(x)-L_2(h_1)-(f(x+h_1)-f(x)-L_1(h_1))\|_Y\\
&\leqslant \|f(x+h_1)-f(x)-L_2(h_1)\|_Y+\|f(x+h_1)-f(x)-L_1(h_1)\|_Y\\
&\leqslant \frac{\varepsilon}{2\cdot(\|h_0\|_X+1)}\cdot\|h_1\|_X+\frac{\varepsilon}{2\cdot(\|h_0\|_X+1)}\cdot\|h_1\|_X\\
&= \frac{\varepsilon}{\|h_0\|_X+1}\cdot\|h_1\|_X\\
&= \frac{\varepsilon}{\|h_0\|_X+1}\cdot\left\|\frac{\delta}{(\|h_0\|_X+1)}\cdot h_0\right\|_X\\
&= \frac{\varepsilon}{\|h_0\|_X+1}\cdot\frac{\delta}{(\|h_0\|_X+1)}\cdot\|h_0\|_X
\end{aligned}$$

so that

$$\|L_1(h_0)-L_2(h_0)\|_Y\leqslant\frac{\varepsilon}{\|h_0\|_X+1}\cdot\|h_0\|_X<\varepsilon=\|L_1(h)-L_2(h)\|_Y$$

a contradiction. So we must have

$$L_1=L_2$$

□

Definition 16.18. (Fréchet differentiability) *Let $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$ be normed spaces, a open set $U\subseteq X$ and $x\in U$ then a function*

$$f:U\to Y$$

is Fréchet differentiable at x if there exist a $L\in L(X,Y)$ such that $\forall\varepsilon\in\mathbb{R}^+$ there exist a $\delta\in\mathbb{R}^+$ such that $\forall h\in U_x$ with $\|h\|_X<\delta$ we have

$$\|f(x+h)-f(x)-L(h)\|_Y\leqslant\varepsilon\cdot\|h\|_X$$

*If f is Fréchet differentiable then by [lemma: 16.17] the above L is **unique** and is called the **Fréchet differential of f at x** and noted by $D_x f$. In other words f is Fréchet differentiable at x if there exist a*

$$D_{\,x}f\in L(X,Y)$$

such that

$$\forall\varepsilon\in\mathbb{R}^+\ \exists\delta\in\mathbb{R}^+ \text{ such that } \forall h\in U_x \text{ with } \|h\|_X \text{ we have } \|f(x+h)-f(x)-D_{\,x}f(h)\|_Y\leqslant\varepsilon\cdot\|h\|_X$$

Note 16.19. Note the uniqueness of the Fréchet differential is guaranteed by [lemma: 16.17] whose proof depends on the crucial fact that U is open. Hence the requirement that U must be open in the previous definition.

Definition 16.20. *Let $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$ be normed spaces, $U\subseteq X$ a open set, $A\subseteq U$ then a function $f:U\rightarrow Y$ is **Fréchet differentiable on A** if $\forall x\in A$ we have that f is Fréchet differentiable at x.*

Note 16.21. First note that D_xf is a element of $L(X,Y)$ hence a continuous linear function from X to so that for $h\in X$ we have that $D_xf(h)\in Y$. In many texts about differentiation the notation $Df(x)$ is used for the Fréchet diffferential at x this again suggest that the differential at x is the value of a function Df that is not yet defined. If however f is differentiable on a subset A of U then we can define the function

$$Df: A\rightarrow L(X,Y) \text{ by } Df(x)=D_xf$$

that maps every point of A to a continuous linear mapping between X and Y. In this book we want to keep these two things separate in notation. Hence we use $Df(x)$ only if f is Fréchet differentiable on a set and not if f is Fréchet at a single point. If Df is defined on A and $x\in A$ then $Df(x)\in L(X,Y)$ so that given a $h\in X$ we have that $(Df(x))(h)\in Y$, to simplify notation we note this as $Df(x)(h)$. This is part of a more general notation covention used where the function application associates to the left for example $((f(x))(y))(z)$ is noted as $f(x)(y)(z)$.

Example 16.22. Let $\langle X,\|\|\rangle$ be a normed space then $\forall x\in X$ we have that $\mathrm{Id}_X: X\rightarrow X$ is Fréchet differentiable at x and $D_x\mathrm{Id}_X=\mathrm{Id}_X$. So if we define

$$D\mathrm{Id}_X: X\rightarrow L(X,X) \text{ by } D\mathrm{Id}_X(x)=D_{\,x}\mathrm{Id}_X$$

then

$$D\mathrm{Id}_X=C_{\mathrm{Id}_x}$$

Proof. Let $\varepsilon\in\mathbb{R}^+$ then we have for $h\in X=X_x$ that

$$\|\mathrm{Id}_X(x+h)-\mathrm{Id}_X(x)-\mathrm{Id}_X(h)\|=\|x+h-x-h\|=\|0\|=0<\varepsilon\cdot\|h\|$$ □

Although Fréchet differentiability and the differential is defined using norms, the definition is actually dependent on the topology generated by the norms. This is expressed in the following theorem.

Theorem 16.23. *Let X,Y be vector spaces over $\mathbb{K}$ with equivalent norms $\|\|_{X_1},\|\|_{X_2}$ on X and $\|\|_{Y_1}$, $\|\|_{Y_2}$ on Y, $U\subseteq X$ a open set [using the topology $\mathcal{T}_{\|\|_{X_1}}\underset{\text{norms are equivalent}}{=}\mathcal{T}_{\|\|_{X_2}}$], $x\in X$ and*

$$f: X\rightarrow Y \text{ is a Fréchet differentiable function with differential } D_xf \text{ using } \|\|_{X_1} \text{ and } \|\|_{Y_1}$$

then

$$f: X\rightarrow Y \text{ is a Fréchet differentiable function with differential } D_xf \text{ using } \|\|_{X_2} \text{ and } \|\|_{Y_2}$$

Proof. As the norms are equivalent we have by [theorem: 14.114] there exists $\alpha_X,\alpha_Y,\beta_X,\beta_Y\in\mathbb{R}^+$ such that $\forall x\in X$ and $\forall y\in Y$ we have

$$\alpha_X\cdot\|x\|_{X_2}\leqslant\|x\|_{X_1}\leqslant\beta_X\cdot\|x\|_{X_2} \text{ and } \alpha_Y\cdot\|y\|_{Y_1}\leqslant\|y\|_{Y_2}\leqslant\beta_Y\cdot\|y\|_{Y_1}$$

Let $\varepsilon\in\mathbb{R}^+$ then there exist a $\delta'\in\mathbb{R}^+$ such that $\forall h\in U_x$ with $\|h\|_{X_1}<\delta'$ we have

$$\|f(x+h)-f(x)-D_{\,x}f\|_{Y_1}\leqslant\frac{\varepsilon}{\beta_Y\cdot\beta_X}\cdot\|h\|_{X_1}$$

Take $\delta = \frac{\delta'}{\beta_X}$ then $\forall h \in U_x$ with $\|h\|_{X_2} < \delta$ we have $\|h\|_{X_1} \leqslant \beta_X \cdot \|h\|_{X_2} < \delta \cdot \beta_X = \delta'$ so that from the above we have

$$\begin{aligned} \|f(x+h) - f(x) - D_x f\|_{Y_1} &\leqslant \frac{\varepsilon}{\beta_Y \cdot \beta_X} \cdot \|h\|_{X_1} \\ &\leqslant \frac{\varepsilon}{\beta_Y \cdot \beta_X} \cdot \beta_X \cdot \|h\|_{X_2} \\ &= \frac{\varepsilon}{\beta_Y} \cdot \|h\|_{X_2} \end{aligned}$$

so that

$$\|f(x+h) - f(x) - D_x f(h)\|_{Y_2} \leqslant \beta_Y \cdot \|f(x+h) - f(x) - D_x f(h)\|_{Y_1} \leqslant \beta_Y \cdot \frac{\varepsilon}{\beta_Y} \cdot \|h\|_{X_2} = \varepsilon \cdot \|h\|_{X_2}$$

This together with the fact $D_x f$ is also a continuous linear function between X and Y using the norms $\|\|_{X_2}$, $\|\|_{Y_2}$ [because the norms are equivalent] proves that f is Fréchet differentiable at x with differential $D_x f$ using the norms $\|\|_{X_2}$, $\|\|_{Y_2}$. □

Differentiability of a function at a point is a local property it only depends on the function values for a arbitrary open neighborhood of this point not on the domain of the function.

Theorem 16.24. *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, $V \subseteq U$ a open subset of U such that $x \in V$ then we have*

$$f \text{ is Fréchet differentiable at } x \text{ with differential } D_x f$$

$$\Updownarrow$$

$$f_{|V} \text{ is Fréchet differentiable at } x \text{ with differential } D_x f_{|V}$$

further if f or $f_{|V}$ is Fréchet differentiable at x then $D_x f = D_x f_{|V}$.

Proof.

⇒. Let $\varepsilon \in \mathbb{R}^+$. As f is Fréchet differentiable at x with differential $D_x f$ there exist a $\delta \in \mathbb{R}^+$ such that $\forall h \in U_x$ with $\|h\|_X < \delta$ we have $\|f(x+h) - f(x) - D_x f\|_Y \leqslant \varepsilon \cdot \|h\|_X$. Let $h \in V_x$ with $\|h\|_X < \delta$ then $x + h \in V \subseteq U$, so that $h \in U_x$. Hence we have

$$\|f_{|V}(x+h) - f_{|V}(x) - D_x f(h)\|_Y \underset{x+h, x \in V}{=} \|f(x+h) - f(x) - D_x f(h)\|_Y \leqslant \varepsilon \cdot \|h\|_X$$

proving that $f_{|V}$ is Fréchet differentiable with differential $D_x f$. Hence $D_x f = D_x f_{|V}$.

⇐. Let $\varepsilon \in \mathbb{R}^+$. As $x + 0 = x \in V$ we have that $0 \in V_x$ which is open [see theorem: 14.109] so that there exists a $\delta_1 \in \mathbb{R}^+$ such that $0 \in B_{\|\|_X}(0, \delta_1) \subseteq V_x$. Further, as $f_{|V}$ is Fréchet differentiable at x with differential $D_x f_V$, there exist a $\delta_2 \in \mathbb{R}^+$ such that $\forall h \in V_x$ with $\|h\|_X < \delta_2$ we have $\|f_{|V}(x+h) - f_{|V}(x) - D_x f_{|V}(h)\|_Y \leqslant \varepsilon \cdot \|h\|_X$. Take $\delta = \min(\delta_1, \delta_2)$ then if $h \in U_x$ with $\|h\|_X < \delta$ we have $\|h\|_X < \delta_1$ so that $h \in B_{\|\|_X}(0, \delta_1) \subseteq V_x$ and $\|h\|_X < \delta_2$, hence we have

$$\|f(x+h) - f(x) - D_x f_{|V}(h)\|_Y \underset{x, x+h \in V}{=} \|f_{|V}(x+h) - f_{|V}(x) - D_x f_{|V}(h)\|_Y \leqslant \varepsilon \cdot \|h\|_X$$

This proves that f is Fréchet differentiable with differential $D_x f_{|V}$. Hence $D_x f = D_x f_{|V}$. □

Corollary 16.25. *Let $\langle X, \|\|_X \rangle$, $\langle Y, \|\|_Y \rangle$ be normed spaces, V_1, V_2 open sets in X, $x \in V_1 \bigcap V_2$, $f: V_1 \rightarrow Y$, $g: V_2 \rightarrow Y$ such that there exist a open set W with $x \in W \subseteq V_1 \bigcap V_2$ and $\forall y \in W$ we have $f(y) = g(y)$. Then if f is Fréchet differentiable at x it follows that g is also Fréchet differentiable at x and $D_x f = D_x g$.*

Proof. As $\forall y \in W$ we have that $f(y) = g(y)$ it follows that $f_{|W} = g_{|W}$. As f is Fréchet differentiable at x it follows from [theorem: 16.24] that $f_{|W}$ is Fréchet differentiable at x with $D_x f = D_x f_{|W} \underset{f_{|W} = g_{|W}}{=} D g_{|W}(x)$. Hence using [theorem: 16.24] again it follows that g is Fréchet differentiable at x with $D_x g = D_x g_{|W} = D_x f$. □

Example 16.26. Let $\langle X, \|\|_X \rangle$, $\langle Y, \|\|_Y \rangle$, $y \in Y$ U a open set in X then $\forall x \in U$ we have that

$$C_y: U \rightarrow Y \text{ defined by } C_y(x) = y$$

is Fréchet differentiable at x and $D_x C_y = C_0$ where $C_0: X \rightarrow Y$ is defined by $C_0(z) = 0$. AS C_0 is the neutral element in $L(X, Y)$ we can also note C_0 as 0 [not to be confused with $0 \in Y$ or $0 \in X$]

Proof. Let $x\in U$ and $\varepsilon\in\mathbb{R}^+$ then for $h\in U_x$

$$\|C_y(x+h)-C_y(x)-C_0(h)\|_Y=\|y-y-0\|_Y=0<\varepsilon\cdot\|h\|_X$$

□

Example 16.27. Let $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$, $L\in L(X,Y)$ then $\forall x\in X$ we have that

$$L \text{ is Fréchet differentiable at } x \text{ with differential } D_xL=L$$

Proof. Let $x\in X$ and $\varepsilon\in\mathbb{R}^+$ then we have

$$\|L(x+h)-L(x)-L(h)\|_Y=\|L(x)+L(h)-L(x)-L(h)\|_Y=\|0\|_Y=0<\varepsilon\cdot\|h\|$$

□

We look now at some alternative definitions of Fréchet differentiability. First we need a definition of a ε-mapping.

Definition 16.28. *Let $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$ be normed spaces, $U\subseteq X$ a open set and $0\in U$ then a **ε-mapping** is a function $\varepsilon: U\rightarrow Y$ which is continuous at $0\in U$ [using the subspace topology on U] and satisfies $\varepsilon(0)=0$.*

Theorem 16.29. *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 Fréchet differentiable at x*
2. *There exist a $L\in L(X,Y)$ and a ε-mapping $\varepsilon: U_x\rightarrow Y$ such that $\forall h\in U_x$ we have*
$$f(x+h)-f(x)-L(h)=\|h\|_X\cdot\varepsilon(h)$$
3. *There exist a $L\in L(X,Y)$ such that $\forall\varepsilon\in\mathbb{R}^+$ $\exists\delta\in\mathbb{R}^+$ such that $\forall h\in U_x$ with $0<\|h\|_X<\delta$ we have*
$$\frac{\|f(x+h)-f(x)-L(h)\|_Y}{\|h\|_X}<\varepsilon$$
4. *There exist a $L\in L(X,Y)$ such that*
$$\lim_{\substack{y\rightarrow x\\ U\setminus\{x\}}}\frac{\|f(y)-f(x)-L(y-x)\|_Y}{\|y-x\|_X}=0$$
5. *There exist a $L\in L(X,Y)$ such that $\forall\{h_i\}_{i\in\mathbb{N}}\subseteq U_x$ with $\forall i\in\mathbb{N}$ $0<\|h_i\|_X$ and $\lim_{i\rightarrow\infty}h_i=0$ we have*
$$\lim_{n\rightarrow\infty}\frac{\|f(x+h_i)-f(x)-L(h_i)\|_Y}{\|h_i\|_X}=0$$

further if either 1,2,3,4,5 is true then $L=D_xf$.

Proof.

1 ⇒ 2. Let $L=D_xf\in L(X,Y)$ and define

$$\varepsilon: U_x\rightarrow Y \text{ by } \varepsilon(h)=\begin{cases}0 \text{ if } h=0\\ \frac{f(x+h)-f(x)-L(h)}{\|h\|_X} \text{ if } h\in U_x\setminus\{0\}\end{cases}$$

then we have $\forall h\in U_x$ that

$$\begin{aligned}\|h\|_X\cdot\varepsilon(h) &= \begin{cases}0 \text{ if } h=0\\ f(x+h)-f(x)-L(h) \text{ if } h\in U_x\setminus\{0\}\end{cases}\\ &= \begin{cases}f(x+0)-f(x)-L(0) \text{ if } h=0\\ f(x+h)-f(x)-L(h) \text{ if } h\in U_x\setminus\{0\}\end{cases}\\ &= f(x+h)-f(x)-L(h)\end{aligned}$$

proving that

$$\forall h\in U_x \text{ we have } f(x+h)-f(x)-L(h)=\|h\|_X\cdot\varepsilon(h)$$

Rest us to prove that ε is continuous at 0. Let $\zeta\in\mathbb{R}^+$. As f is Fréchet differentiable there exist a $\delta\in\mathbb{R}^+$ such that $\forall h\in U_x$ with $\|h\|_X<\delta$ we have

$$\|f(x+h)-f(x)-D_xf(h)\|_Y<\zeta\cdot\|h\|_X$$

Hence we have

$$\begin{array}{rcl}\|\varepsilon(h)-\varepsilon(0)\|_Y & = & \|\varepsilon(h)\|_Y\\ & = & \left\|\begin{cases}\|0\|_Y<\zeta \text{ if } h=0\\ \frac{\|f(x+h)-f(x)-Df(x)(h)\|_Y}{\|h\|_X}<\zeta \text{ if } h\in U_x\setminus\{0\}\end{cases}\right\|_Y\\ & < & \zeta\end{array}$$

which by [theorem: 14.128] proves that ε is continuous at 0 in the subspace topology of U_x.

2 ⇒ 3. Let $L\in L(X,Y)$ and $\varepsilon: U_x\rightarrow Y$ a ε-mapping such that $\forall h\in U_x$

$$f(x+h)-f(x)-L(h)=\|h\|_X\cdot\varepsilon(h)$$

Take $\zeta\in\mathbb{R}^+$. As $\varepsilon: U_x\rightarrow Y$ is continuous we can use [theorem: 14.128] to find a $\delta\in\mathbb{R}^+$ such that $\forall h\in U_x$ with $\|h\|_X=\|h-0\|_X<\delta$ we have $\|\varepsilon(h)\|_Y=\|\varepsilon(h)-\varepsilon(0)\|_Y<\zeta$. Hence if $0<\|h\|_X<\delta$ then

$$\frac{\|f(x+h)-f(x)-L(h)\|_Y}{\|h\|_X}=\frac{\|h\|_X\cdot\varepsilon(h)}{\|h\|_X}=\varepsilon(h)<\zeta$$

3 ⇒ 1. Let $L\in L(X,Y)$ be such that $\forall\varepsilon\in\mathbb{R}^+$ there exist a $\delta\in\mathbb{R}^+$ so that $\forall h\in U_x$ with $0<\|h\|_X<\delta$ we have

$$\frac{\|f(x+h)-f(x)-L(h)\|_Y}{\|h\|_X}<\varepsilon$$

Now $\forall h\in U_x$ with $\|h\|_X<\delta$ we have two cases to consider:

$\|h\|_X=0$. Then $h=0$ and

$$\|f(x+h)-f(x)-L(h)\|_Y=\|f(x+0)-f(x)-L(0)\|_Y=\|0\|_Y=\varepsilon\cdot 0=\varepsilon\cdot\|h\|_X$$

$0<\|h\|_X$. Then

$$\|f(x+h)-f(x)-L(h)\|_Y=\|h\|_X\cdot\frac{\|f(x+h)-f(x)-L(h)\|_Y}{\|h\|_X}<\|h\|_X\cdot\varepsilon=\varepsilon\cdot\|h\|_x$$

so that in all cases we have

$$\|f(x+h)-f(x)-L(h)\|_Y\leqslant\varepsilon\cdot\|h\|_X$$

Hence f is Fréchet differentiable at x with differential L or $L=D_x f$.

3 ⇒ 4. Let $L\in L(X,Y)$ be such that $\forall\varepsilon\in\mathbb{R}^+$ there exist a $\delta\in\mathbb{R}^+$ so that $\forall h\in U_x$ with $0<\|h\|_X<\delta$ we have

$$\frac{\|f(x+h)-f(x)-L(h)\|_Y}{\|h\|_X}<\varepsilon \tag{16.2}$$

Define then

$$F: U\setminus\{x\}\rightarrow\mathbb{R} \text{ by } F(y)=\frac{\|f(y)-f(x)-L(y-x)\|_Y}{\|y-x\|_X}$$

Let $y\in U\setminus\{x\}\Rightarrow y-x\in U_x$ with $0<\|y-x\|<\delta$ then we have

$$|F(y)-0|=\frac{\|f(y)-f(x)-L(y-x)\|_Y}{\|y-x\|_X}=\frac{\|f(x+(y-x))-f(x)-L(y-x)\|_Y}{\|(y-x)\|_X}<_{\text{[eq 16.2]}}\varepsilon$$

proving that

$$\lim_{\substack{y\rightarrow x\\ U_x}}\frac{\|f(y)-f(x)-L(y-x)\|_Y}{\|y-x\|_X}\underset{\text{notation}}{=}\lim_{\substack{y\rightarrow x\\ U\setminus\{x\}}}F(y)=0$$

4 ⇒ 3. Let $L\in L(X,Y)$ be such that for $F: U\setminus\{0\}\rightarrow Y$ defined by $F(y)=\frac{\|f(y)-f(x)-L(y-x)\|_Y}{\|y-x\|_X}$ we have

$$\lim_{\substack{y\rightarrow x\\ U\setminus\{x\}}}F(y)=0$$

Take $\varepsilon\in\mathbb{R}^+$ then as $\lim_{\substack{y\rightarrow x\\ U\setminus\{x\}}}F(y)=0$ there exist a $\delta\in\mathbb{R}^+$ such that $\forall y\in U\setminus\{x\}$ with $0<\|x-y\|_X<\delta$ we have that

$$\frac{\|f(y)-f(x)-L(y-x)\|_Y}{\|y-x\|_X}=|F(y)-0|<\varepsilon$$

Hence if $h\in U_x$ with $0<\|h\|_X<\delta$ we have, if we take $y=x+h$, that $0<\|y-x\|_X<\delta$ and $y\in U\setminus\{x\}$ so that

$$\frac{\|f(x+h)-f(x)-L(h)\|_Y}{\|h\|_X}=\frac{\|f(y)-f(x)-L(y-x)\|_Y}{\|y-x\|_X}<\varepsilon$$

3 ⇒ 5. Let $\{h_i\}_{i\in\mathbb{N}}\subseteq U_x$ be a sequence with $\forall i\in\mathbb{N}$ $0<\|h_i\|_X$ such that $\lim_{i\rightarrow\infty} h_i=0$. Take $\varepsilon\in\mathbb{R}^+$ then by (3) there exist a $L\in L(X,Y)$ and a $\delta\in\mathbb{R}^+$ such that $\forall h\in U_x$ with $0<\|h\|_X<\delta$ we have

$$\frac{\|f(x+h)-f(x)-L(h)\|_Y}{\|h\|_X}<\varepsilon$$

As $\lim_{i\rightarrow\infty} h_i=0$ there exist a $N\in\mathbb{N}$ such that $\forall n\geqslant N$ we have that $0<\|h_n\|_X=\|h_n-0\|_X<\delta$, so that

$$\left|\frac{\|f(x+h_n)-f(x)-L(h_n)\|_Y}{\|h_n\|_X}-0\right|=\frac{\|f(x+h_n)-f(x)-L(h_n)\|_Y}{\|h_n\|_X}<\varepsilon$$

This proves that

$$\lim_{i\rightarrow\infty}\frac{\|f(x+h_i)-f(x)-L(h_i)\|_Y}{\|h_i\|_X}=0$$

5 ⇒ 3. Assume that (3) is false. Then given a $L\in L(X,Y)$ there exist a $\varepsilon_0\in\mathbb{R}^+$ such that $\forall\delta\in\mathbb{R}^+$ there exist a $h\in U_x$ with $0<\|h\|_X<\delta$ such that

$$\frac{\|f(x+h)-f(x)-L(h)\|_Y}{\|h\|_X}>\varepsilon_0$$

Using [theorem: 3.104] it follows that $\forall n\in\mathbb{N}$ we have, as $\frac{1}{n}\in\mathbb{R}^+$, that there exist a $h_n\in U_x$ with $0<\|h_n\|_X<\frac{1}{n}$ such that

$$\frac{\|f(x+h_n)-f(x)-L(h_n)\|_Y}{\|h_n\|_X}>\varepsilon_0 \tag{16.3}$$

Now using [example: 14.293] together with [theorem: 14.295] it follows that

$$\lim_{n\rightarrow\infty} h_n=0$$

Hence by (5) we have

$$\lim_{n\rightarrow\infty}\frac{\|f(x+h_n)-f(x)-L(h_n)\|_Y}{\|h_n\|_X}=0$$

So there exist a $N\in\mathbb{N}$ such that

$$\frac{\|f(x+h_N)-f(x)-L(h_N)\|_Y}{\|h_N\|_X}<\varepsilon_0$$

contradicting [eq: 16.3]. Hence the assumption is wrong and (3) follows. □

Using the fact that differentiability is a local property we can restate the above theorem as follows.

Corollary 16.30. *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 Fréchet differentiable at x*
2. *There exist $L\in L(X,Y)$, a $\delta\in\mathbb{R}^+$ such that $B_{\|\|_X}(x,\delta)\subseteq U$ and a ε-mapping $\varepsilon: B_{\|\|_X}(0,\delta)\rightarrow Y$ such that*

$$\forall h\in B_{\|\|_X}(0,\delta)\;\; f(x+h)-f(x)-L(h)=\|h\|_X\cdot\varepsilon(h)$$

Further if (1) or (2) is true then $D_x f=L$

Proof.

1 ⇒ 2. As U is open and $x\in U$ there exist a $\delta\in\mathbb{R}^+$ such that $x\in B_{\|\|_X}(x,\delta)\subseteq U$. Using the locality of differentiability [see theorem: 16.24] it follows that $f_{|B_{\|\|x}(x,\delta)}$ is Fréchet differentiable at x and $D_x f=D_x f_{|B_{\|\|_X}(x,\delta)}$. Using the previous theorem [theorem: 16.29] there exist a ε-mapping

$$\varepsilon:(B_{\|\|_X}(x,\delta))_x\underset{\text{[theorem: 16.4]}}{=}B_{\|\|_X}(0,\delta)\rightarrow Y$$

such that $\forall h\in(B_{\|\|_X}(x,\delta))_x\underset{\text{[theorem: 16.4]}}{=}B_{\|\|_X}(0,\delta)$ we have

$$\begin{aligned} f(x+h)-f(x)-D_x f(h) &= f_{|B_{\|\|_X}(x,\delta)}(x+h)-f_{|B_{\|\|_X}(x.\delta)}(x)-D_x f_{|B_{\|\|_X}(x,\delta)}(h)\\ &= \|h\|_X\cdot\varepsilon(h)\end{aligned}$$

2 ⇒ 1. As $\varepsilon: B_{\|\|_X}(0,\delta) \underset{[\text{theorem: } 16.4]}{=} (B_{\|\|_X}(x,\delta))_x \rightarrow Y$ is a ε-mapping such that

$$\forall h \in B_{\|\|_X}(0,\delta) \underset{[\text{theorem: } 16.4]}{=} (B_{\|\|_X}(x,\delta))_x$$

we have

$$f_{|B_{\|\|_X}(x,\delta)}(x+h) - f_{|B_{\|\|_X}(x,\delta)}(x) - L(h) = f(x+h) - f(x) - L(h) = \|h\|_X \cdot \varepsilon(h)$$

it follows from [theorem: 16.29] that $f_{|B_{\|\|_X}(x,\delta)}$ is Fréchet differentiable at x with $L = D_x f_{|B_{\|\|_X}(x,\delta)}$. Using the locality of differentiability [see theorem: 16.24] it follows that f is differentiable at x with $D_x f = D_x f_{|B_{\|\|_X}(x,\delta)} = L$.

□

A function that is Fréchet differentiable at x is continuous at x as proved in the next theorem.

Theorem 16.31. *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 Fréchet differentiable at x then f is continuous at x.*

Proof. Let $\varepsilon \in \mathbb{R}^+$. As f is Fréchet differentiable at x there exist a $\delta_1 \in \mathbb{R}^+$ such that $\forall h \in U_x$ with $\|h\|_X < \delta$ we have

$$\|f(x+h) - f(x) - D_x f(h)\|_Y \leqslant \frac{\varepsilon}{2} \cdot \|h\|_X \tag{16.4}$$

As $D_x f \in L(X,Y)$ we have [see theorem: 14.179] that

$$\forall h \in X \;\; \|D_x f(h)\|_Y \leqslant \|D_x f\|_{L(X,Y)} \cdot \|h\|_X \tag{16.5}$$

Define now $\delta = \min\left(\delta_1, 1, \frac{\varepsilon}{2 \cdot (\|D_x f\|_{L(X,Y)} + 1)}\right) \in \mathbb{R}^+$ then for every $y \in U$ with $\|y - x\|_X < \delta$ we have $y - x \in U_x$ [as $(y-x) + x = y \in U$] and $\|y - x\|_X \leqslant \delta_1, 1, \frac{\varepsilon}{2 \cdot \|D_x f\|_{L(X,Y)} + 1}$. Using [eq: 16.4] we have

$$\begin{aligned} \|f(y) - f(x) - D_x f(y-x)\|_Y &= \|f(x + (y-x)) - f(x) - D_x f(y-x)\|_Y \\ &\leqslant \frac{\varepsilon}{2} \cdot \|y - x\|_X \\ &\leqslant \frac{\varepsilon}{2} \end{aligned}$$

and by [eq: 16.5] it follows that

$$\|D_x f(y-x)\|_Y \leqslant \|D_x f\|_{L(X,Y)} \cdot \|y - x\|_X \leqslant \|D_x f\|_{L(X,Y)} \cdot \frac{\varepsilon}{2 \cdot (\|D_x f\|_{L(X,Y)} + 1)} < \frac{\varepsilon}{2}$$

Hence we have $\forall y \in U$ with $\|y - x\|_X < \delta$ that

$$\begin{aligned} \|f(y) - f(x)\|_Y &= \|f(y) - f(x) - D_x f(y-x) + D_x f(y-x)\|_Y \\ &\leqslant \|f(y) - f(x) - D_x f(y-x)\|_Y + \|D_x f(y-x)\|_Y \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\ &= \varepsilon \end{aligned}$$

So by [theorem: 14.128] it follows that f is continuous at x. □

We look now at the relation between the Fréchet differential of a function and the derivative of a function.

Theorem 16.32. *Let $\langle \mathbb{K}, \|\rangle$ be the normed space of real or complex numbers, $\langle Y, \|\|\rangle$ a normed space, $U \subseteq \mathbb{K}$ a open set, $x \in U$ and $f: U \rightarrow X$ a function then we have*

$$f \text{ has a derivative } f'_x \text{ at } x$$

$$\Updownarrow$$

$$f \text{ is Fréchet differentiable at } x \text{ with Fréchet differential } D_x f$$

Further if f'_x or $D_x f$ exist then

$$f'_x = D_x f(1)$$

and

$$\forall h \in \mathbb{K} \text{ we have } D_x f(h) = h \cdot f'_x$$

Proof.

⇒. Define $L:\mathbb{K}\rightarrow Y$ by $L(h)=h\cdot f'(x)$ then by [example: 14.183]

$$L\in L(\mathbb{K},Y)$$

Let $\varepsilon\in\mathbb{R}^+$ then by [definition: 16.12] there exist a $\delta\in\mathbb{R}^+$ such that for $0<|h|<\delta$ we have

$$\left\|\frac{f(x+h)-f(x)}{h}-f'_x\right\|<\varepsilon \tag{16.6}$$

Let $h\in U_x$ with $|h|<\delta$ then we have either:

$|h|=0$. Then $h=0$ so that

$$\|f(x+h)-f(x)-L(h)\|=\|f(x)-f(x)-0\cdot f'_x\|=\|0\|=0=\varepsilon\cdot|0|=\varepsilon\cdot|h|$$

$0<|h|$. Then $h\neq 0$ so that

$$\begin{aligned}
\|f(x+h)-f(x)-L(h)\| &= \|f(x+h)-f(x)-h\cdot f'_x\|\\
&= \left\|h\cdot\frac{f(x+h)-f(x)}{h}-h\cdot f'_x\right\|\\
&= \left\|h\cdot\left(\frac{f(x+h)-f(x)}{h}-f'_x\right)\right\|\\
&= |h|\cdot\left\|\frac{f(x+h)-f(x)}{h}-f'_x\right\|\\
&<_{[\text{eq: } 16.6]} \varepsilon\cdot|h|
\end{aligned}$$

So in all cases we have

$$\|f(x+h)-f(x)-L(h)\|\leqslant\varepsilon\cdot|h|$$

proving that f is Fréchet differentiable at x with the Fréchet differential L, hence D_xf is defined by

$$D_xf(h)=h\cdot f'_x \text{ so that } f'_x=D_xf(1)$$

⇐. Take $f'_x=D_xf(1)$ and let $\varepsilon\in\mathbb{R}^+$. Then as f is Fréchet differentiable at x with Fréchet differential D_xf there exist a $\delta\in\mathbb{R}^+$ such that

$$\|f(x+h)-f(x)-D_xf(h)\|\leqslant\frac{\varepsilon}{2}\cdot|h| \tag{16.7}$$

Let $h\in U_x$ be such that $0<|h|<\delta$ then we have

$$\begin{aligned}
\left\|\frac{f(x+h)-f(x)}{h}-f'_x\right\| &= \left\|\frac{f(x+h)-f(x)}{h}-D_xf(1)\right\|\\
&= \left\|\frac{f(x+h)-f(x)-h\cdot D_xf(1)}{h}\right\|\\
&= \left\|\frac{f(x+h)-f(x)-D_xf(h\cdot 1)}{h}\right\|\\
&= \left\|\frac{f(x+h)-f(x)-D_xf(h)}{h}\right\|\\
&= \frac{1}{|h|}\cdot\|f(x+h)-f(x)-D_xf(h)\|\\
&\leqslant_{[\text{eq: } 16.7]} \frac{1}{|h|}\cdot\frac{\varepsilon}{2}\cdot|h|\\
&= \frac{\varepsilon}{2}\\
&< \varepsilon
\end{aligned}$$

Proving that f has a derivative $f'_x=D_xf(1)$. □

We prove now that the existence of a derivative is a local property.

Corollary 16.33. *Let $\langle\mathbb{K},\|\rangle$ be the normed space of real (or complex numbers), $\langle X,\|\|\rangle$ a normed space, U,V two open sets in $\mathbb{K}$, $x\in U\bigcap V$ and $f:U\to X$, $g:V\to X$ two functions such that there exist a open set W with $x\in W\subseteq U\bigcap V$ and $\forall y\in W$ $f(y)=g(y)$. Then if f has a derivative f'_x at x it follows that g has also a derivative and that $f'_x=g'_x$.*

***Note* 16.34.** *In the special case of $V\subseteq U$ and $g=f_{|V}$ this reduces to the following. If f has a derivative f'_x at x then $f_{|V}$ has a derivative at x and $f'_x=(f_{|V})'_x$.*

Proof. As f is has a derivative at x it follows from [theorem: 16.32] that f is Fréchet differentiable at x and $f'_x=D_xf(1)$. Using [corollary: 16.25] it follows that g is Fréchet differentiable at x and $D_xg=D_xf$. Hence using [theorem: 16.32] again it follows that g has a derivative at x and $g'_x=D_xg(1)=D_xf(1)=f'_x$. □

16.1.4 Properties of the Fréchet differential

Theorem 16.35. *Let $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$ be normed spaces, $U\subseteq X$ a open set, $x\in U$ then we have:*

1. *If $f:U\to Y$ is Fréchet differentiable at x and $g:U\to Y$ is Fréchet differentiable at x then*

$$f+g:U\to Y \text{ is Fréchet differentiable with Fréchet differential } D_x(f+g)=D_xf+D_xg$$

2. *If $f:U\to Y$ is Fréchet differentiable at x and $\alpha\in\mathbb{K}$ then*

$$\alpha\cdot f:U\to Y \text{ is Fréchet differentiable at } x \text{ with Fréchet differential } D_x(\alpha\cdot f)=\alpha\cdot D_xf$$

Proof.

1. Let $\varepsilon\in\mathbb{R}^+$. As f,g are Fréchet differentiable at x there exist δ_1,δ_2 such that

$$\forall h\in U_x \text{ with } \|h\|_x<\delta_1 \text{ we have } \|f(x+h)-f(x)-D_xf(h)\|_Y\leqslant\frac{\varepsilon}{2}\cdot\|h\|_X$$

$$\forall h\in U_x \text{ with } \|h\|_x<\delta_2 \text{ we have } \|g(x+h)-g(x)-D_xg(h)\|_Y\leqslant\frac{\varepsilon}{2}\cdot\|h\|_X$$

As D_xf, $D_xg\in L(X,Y)$ we have by [theorem: 14.173] that $D_xf+D_xg\in L(X,Y)$. Further if $h\in U_x$ with $\|h\|_X<\min(\delta_1,\delta_2)$ then we have

$$\begin{aligned}\|(f+g)(x+h)-(f+g)(x)-(D_xf+D_Xg)(h)\|_Y &=\\ \|f(x+h)+g(x+h)-f(x)-g(x)-(D_xf(h)+D_xg(h))\|_Y &\leqslant\\ \|f(x+h)-f(x)-D_xf(h)\|_Y+\|g(x+h)-g(x)-D_xg(h)\|_Y &\leqslant\\ \frac{\varepsilon}{2}\cdot\|h\|_X+\frac{\varepsilon}{2}\cdot\|h\|_Y &=\\ \varepsilon\cdot\|h\|_X\end{aligned}$$

proving that $f+g$ is Fréchet differentiable at x with Fréchet differential

$$D_x(f+g)=D_xf+D_xg$$

2. Let $\varepsilon\in\mathbb{R}^+$. As f is Fréchet differentiable at x there exist a $\delta\in\mathbb{R}^+$ such that

$$\forall h\in U_x \text{ with } \|h\|_x<\delta_1 \text{ we have } \|f(x+h)-f(x)-D_xf(h)\|_Y\leqslant\frac{\varepsilon}{|\alpha|+1}\cdot\|h\|_X$$

As $D_xf\in L(X,Y)$ it follows from [theorem: 14.173] that $\alpha\cdot D_xf\in L(X,Y)$. Further if $h\in U_x$ with $\|h\|_X<\delta$ then we have

$$\begin{aligned}\|(\alpha\cdot f)(x+h)-(\alpha\cdot f)(x)-(\alpha\cdot D_xf)(h)\|_Y &=\\ \|\alpha\cdot f(x+h)-\alpha\cdot f(x)-\alpha\cdot D_xf(h)\|_Y &=\\ \|\alpha\cdot(f(x+h)-f(x)-D_xf(h))\|_Y &=\\ |\alpha|\cdot\|f(x+h)-f(x)-D_xf(h)\|_Y &\leqslant\\ |\alpha|\cdot\frac{\varepsilon}{|\alpha|+1}\cdot\|h\|_X &<\\ \varepsilon\cdot\|h\|_X\end{aligned}$$

proving that $\alpha\cdot f$ is Fréchet differentiable at x with Fréchet differential $\alpha\cdot D_xf$. □

We introduce now the most important rule for Fréchet differentiation.

Theorem 16.36. (Chain Rule) *Let $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$, $\langle Z,\|\|_Z\rangle$ be normed spaces, $U\subseteq X$ a open set in X, $x\in U$, $V\subseteq Y$ a open set in Y, $f:U\to Y$, $g:V\to Z$ functions such that $f(U)\subseteq V$, f is Fréchet differentiable at x and g is Fréchet differentiable at $f(x)$ then*

$$g\circ f:U\to Z \text{ is Fréchet differentiable at } x$$

and

$$D_x(g\circ f)=D_{f(x)}g\circ D_x f$$

Proof. As f is Fréchet differentiable at x we have by [theorem: 16.29] that there exist a ε-mapping $\varepsilon_f:U_x\to Y$ such that

$$\forall h\in U_x \text{ we have } f(x+h)-f(x)-D_x f(h)=\|h\|_X\cdot\varepsilon_f(h) \tag{16.8}$$

Further as g is Fréchet differentiable at $f(x)$ we can use [theorem: 16.29] again, giving a ε-mapping $\varepsilon_g:V_{f(x)}\to Z$ satisfying

$$\forall k\in V_{f(x)} \text{ we have } g(f(x)+k)-g(f(x))-D_{f(x)}g(k)=\|k\|_Y\cdot\varepsilon_g(k) \tag{16.9}$$

Let $h\in U_x$. As $(f(x+h)-f(x))+f(x)=f(x+h)\in f(U)\subseteq V$ we have that $f(x+h)-f(x)\in V_{f(x)}$, combining this with [eq: 16.9] gives

$$\begin{array}{rl}
\|f(x+h)-f(x)\|_Y\cdot\varepsilon_g(f(x+h)-f(x)) & \underset{[\text{eq: } 16.9]}{=} \\
g(f(x)+(f(x+h)-f(x)))-g(f(x))-D_{f(x)}g(f(x+h)-f(x)) & = \\
g(f(x+h))-g(f(x))-D_{f(x)}g(f(x+h)-f(x)) & \underset{[\text{eq: } 16.8]}{=} \\
g(f(x+h))-g(f(x))-D_{f(x)}g(\|h\|_X\cdot\varepsilon_f(h)+D_x f(h)) & = \\
g(f(x+h))-g(f(x))-D_{f(x)}g(\|h\|_X\cdot\varepsilon_f(h))-D_{f(x)}\,g(D_x f(h)) & = \\
(g\circ f)(x+h)-(g\circ f)(x)-D_{f(x)}g(\|h\|_X\cdot\varepsilon_f(h))-(D_{f(x)}\,g\circ D_x f)(h) &
\end{array}$$

proving that

$$\begin{array}{rl}
(g\circ f)(x+h)-(g\circ f)(x)-(D_{f(x)}\,g\circ D_x f)(h) & = \\
\|f(x+h)-f(x)\|_Y\cdot\varepsilon_g(f(x+h)-f(x))+D_{f(x)}\,g(\|h\|_X\cdot\varepsilon_f(h)) & = \\
\|f(x+h)-f(x)\|_Y\cdot\varepsilon_g(f(x+h)-f(x))+\|h\|_X\cdot D_{f(x)}g(\varepsilon_f(h)) &
\end{array} \tag{16.10}$$

The next logical step is to find a ε-mapping $\zeta:U_x\to Y$ such that

$$\zeta(h)\cdot\|h\|_x=\|f(x+h)-f(x)\|_Y\cdot\varepsilon_g(f(x+h)-f(x))+\|h\|_X\cdot D_{f(x)}g(\varepsilon_f(h))$$

because we can then use [theorem: 16.29] to finish te proof. So define

$$\zeta:U_x\to Y \text{ by } \begin{cases} \frac{\|f(x+h)-f(x)\|_Y\cdot\varepsilon_g(f(x+h)-f(x))}{\|h\|_X}+D_{f(x)}g(\varepsilon_f(h)) \text{ if } h\in U_x\setminus\{0\} \\ 0 \text{ if } h=0 \end{cases}$$

Then for $h\in U_x$ we have either:

$h=0$. Then

$$\begin{array}{rl}
\|f(x+h)-f(x)\|_Y\cdot\varepsilon_g(f(x+h)-f(x))+\|h\|_X\cdot D_{f(x)}g(\varepsilon_f(h)) & = \\
\|f(x+0)-f(x)\|_Y\cdot\varepsilon_g(f(x+0)-f(x))+\|0\|_X\cdot D_{f(x)}g(\varepsilon_f(0)) & = \\
0 & = \\
0\cdot\zeta(0) & = \\
\|h\|_X\cdot\zeta(h) &
\end{array}$$

$h\neq 0$. Then

$$\begin{array}{rl}
\|f(x+h)-f(x)\|_Y\cdot\varepsilon_g(f(x+h)-f(x))+\|h\|_X\cdot D_{f(x)}g(\varepsilon_f(h)) & = \\
\|h\|_X\cdot\frac{\|f(x+h)-f(x)\|_Y\cdot\varepsilon_g(f(x+h)-f(x))+\|h\|_X\cdot D_{f(x)}g(\varepsilon_f(h))}{\|h\|_X} & = \\
\|h\|_X\cdot\left(\frac{\|f(x+h)-f(x)\|_Y\cdot\varepsilon_g(f(x+h)-f(x))}{\|h\|_X}+D_{f(x)}g(\varepsilon_f(h))\right) & = \\
\|h\|_X\cdot\zeta(h) &
\end{array}$$

Combining the above with [eq: 16.10] gives us the desired

$$(g\circ f)(x+h)-(g\circ f)(x)-(D_{f(x)}g\circ D_x f)(h)=\|h\|_X\cdot\zeta(h) \tag{16.11}$$

Next we must prove that ζ is continuous at 0. Let $\varepsilon\in\mathbb{R}^+$. As ε_f is a ε-mapping, it is continuous at 0. So by [theorem: 14.128] there exist a $\delta_f\in\mathbb{R}^+$ such that if $h\in U_x$ with $\|h\|_X<\delta_f$ then

$$\|\varepsilon_f(h)\|_Y<1 \tag{16.12}$$

So if additional $0<\|h\|_X$ we have

$$\begin{aligned}\frac{\|f(x+h)-f(x)\|_Y}{\|h\|_X} &\underset{[\text{eq: }16.8]}{=} \frac{\|\|h\|_X\cdot\varepsilon_f(h)+D_x f(h)\|_Y}{\|h\|_X}\\ &\leqslant \frac{\|\|h\|_X\cdot\varepsilon_f(h)\|_Y+\|D_x f(h)\|_Y}{\|h\|_X}\\ &= \frac{\|h\|_X\cdot\|\varepsilon_f(h)\|_Y+\|D_x f(h)\|_Y}{\|h\|_X}\\ &\leqslant \frac{\|h\|_X\cdot\|\varepsilon_f(h)\|_Y+\|D_x f\|_{L(X,Y)}\cdot\|h\|_X}{\|h\|_X}\\ &= \|\varepsilon_f(h)\|_Y+\|D_x f\|_{L(X,Y)}\\ &<_{[\text{eq: }16.12]} 1+\|D_x f\|_{L(X,Y)}\end{aligned}$$

hence

$$\forall h\in U_x \text{ with } 0<\|h\|_X<\delta_f \text{ we have that } \frac{\|f(x+h)-f(x)\|_Y}{\|h\|_X}<1+\|D_x f\|_{L(X,Y)} \tag{16.13}$$

As ε_g is a ε-mapping it is continuous at 0. So by [theorem: 14.128] there exist a $\delta_g\in\mathbb{R}^+$ such that

$$\forall k\in V_{f(x)} \text{ with } \|k\|_Y<\delta_g \text{ we have } \|\varepsilon_g(k)\|_Z<\frac{\varepsilon}{2\cdot(1+\|D_x f\|_{L(X,Y)})} \tag{16.14}$$

As f is Fréchet differentiable at x we have by [theorem: 16.31] that f is continuous at x, hence there exist by [theorem: 14.128] a $\delta_1\in\mathbb{R}^+$ such that $\forall y\in U$ with $\|y-x\|_X<\delta_1$ we have $\|f(y)-f(x)\|_Y<\delta_g$. So if $h\in U_x$ with $\|h\|_X<\delta_1$ we have $x+h\in U$ and $\|(x+h)-x\|_X=\|h\|_X<\delta_1$ so that $\|f(x+h)-f(x)\|_Y<\delta_g$ which by [eq: 16.14] results in $\|\varepsilon_g(f(x+h)-f(x))\|_Z<\frac{\varepsilon}{2\cdot(1+\|D_x f\|_{L(X,Y)})}$. Summarized

$$\forall h\in U_x \text{ with } \|h\|_X<\delta_1 \text{ we have } \|\varepsilon_g(f(x+h)-f(x))\|_Z<\frac{\varepsilon}{2\cdot(1+\|D_x f\|_{L(X,Y)})}$$

Combining this with [eq: 16.13] we have $\forall h\in U_x$ with $0<\|h\|_X<\min(\delta_1,\delta_f)$ that

$$\begin{aligned}\frac{\|f(x+h)-f(x)\|_Y}{\|h\|_X}\cdot\|\varepsilon_g(f(x+h)-f(x))\|_Z &\leqslant\\ (1+\|D_x f\|_{L(X,Y)})\cdot\|\varepsilon_g(f(x+h)-f(x))\|_Z &<\\ (1+\|D_x f\|_{L(X,Y)})\cdot\frac{\varepsilon}{2\cdot(1+\|D_x f\|_{L(X,Y)})} &=\\ \frac{\varepsilon}{2}&\end{aligned}$$

or

$$\forall h\in U_x \text{ with } 0<\|h\|_X<\min(\delta_1,\delta_f)\ \|\varepsilon_g(f(x+h)-f(x))\|_Z\cdot\frac{\|f(x+h)-f(x)\|_Y}{\|h\|_X}<\frac{\varepsilon}{2} \tag{16.15}$$

As $\varepsilon_f: U_x\rightarrow Y$ is a ε-mapping is is continuous at 0, so by [theorem: 14.128] there exist a $\delta_2\in\mathbb{R}^+$ such that $\forall h\in U_x$ with $\|h\|_X<\delta_2$ we have $\|\varepsilon_f(h)\|_Y<\frac{\varepsilon}{2\cdot(1+\|_{f(x)}g\|_{L(Y,Z)})}$. Hence

$$\begin{aligned}\|D_{f(x)}g(\varepsilon_f(h))\|_Z &\leqslant \|D_{f(x)}g\|_{L(Y,Z)}\cdot\|\varepsilon_f(h)\|_Y\\ &\leqslant \|D_{f(x)}g\|_{L(Y,Z)}\cdot\frac{\varepsilon}{2\cdot(1+\|D_{f(x)}g\|_{L(Y,Z)})}\\ &< \frac{\varepsilon}{2}\end{aligned}$$

In other words

$$\forall h \in U_x \text{ with } \|h\|_X < \delta_2 \text{ we have } \|D_{f(x)} g(\varepsilon_f(h))\|_Z < \frac{\varepsilon}{2} \tag{16.16}$$

Take now $\delta = \min(\delta_f, \delta_1, \delta_2)$ and $h \in U_x$ with $\|h\|_X = \|h - 0\|_X < \delta$ then we have either:

$\boldsymbol{h = 0}$. Then $\|\zeta(h) - 0\|_Z = \|\zeta(0)\|_Z = \|0\|_Z < \varepsilon$

$\boldsymbol{h \neq 0}$. Then $0 < \|h\|_X < \delta_f, \delta_1, \delta_2$ so that

$$\begin{aligned}
\|\zeta(h) - 0\|_Z &= \|\zeta(h)\|_Z \\
&= \left\| \frac{\|f(x+h) - f(x)\|_Y \cdot \varepsilon_g(f(x+h) - f(x))}{\|h\|_X} + D_{f(x)} g(\varepsilon_f(h)) \right\|_Z \\
&\leqslant \left\| \frac{\|f(x+h) - f(x)\|_Y \cdot \varepsilon_g(f(x+h) - f(x))}{\|h\|_X} \right\|_Z + \|D_{f(x)} g(\varepsilon_f(h))\|_Z \\
&= \frac{\|f(x+h) - f(x)\|_Y}{\|h\|_X} \cdot \|\varepsilon_g(f(x+h) - f(x))\|_Z + \|D_{f(x)} g(\varepsilon_f(h))\|_Z \\
&<_{[\text{eq: } 16.15]} \frac{\varepsilon}{2} + \|D_{f(x)} g(\varepsilon_f(h))\|_Z \\
&<_{[\text{eq: } 16.16]} \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\
&= \varepsilon
\end{aligned}$$

So we have in all cases that $\|\zeta(h) - 0\|_Z < \varepsilon$ which as $\varepsilon \in \mathbb{R}^+$ was chosen arbitrary proves that ζ is continuous at 0. As also $\zeta(0) = 0$ it follows by definition that $\zeta: U_x \to Z$ is a ε-mapping. Hence as by [eq: 16.11] $(g \circ f)(x+h) - (g \circ f)(x) - (D_{f(x)} g \circ D_x f)(h) = \|h\|_X \cdot \zeta(h)$ it follows from [theorem: 16.29] that

$$g \circ f \text{ is Fréchet differentiable with differential } D_x(g \circ f) = D_{f(x)} g \circ D_x f$$

□

Corollary 16.37. *Let $\langle X, \|\|\rangle$ be a normed space, $\langle \mathbb{K}, \|\rangle$ the normed space of real or complex numbers, $U \subseteq \mathbb{K}$ a open set in $\mathbb{K}$, $x \in U$ $V \subseteq \mathbb{K}$ a open set in $\mathbb{K}$, $f: U \to \mathbb{K}$ a function and $g: V \to X$ functions such that $f(U) \subseteq V$, f has a derivative f'_x at x and g has a derivative $g'_{f(x)}$ at $f(x)$. Then $g \circ f: U \to Y$ has a derivative at x and*

$$(g \circ f)'_x = f'_x \cdot g'_{f(x)}$$

Proof. Using [theorem: 16.32] it follows that f is Fréchet differentiable at x and g is Fréchet differentiable at $f(x)$. Using the Chain Rule [theorem: 16.36] we have that $g \circ f$ is Fréchet differentiable at x. Hence, using [theorem: 16.32], we have that $g \circ f$ has a derivative at x. Further we have

$$\begin{aligned}
(g \circ f)'_x &\underset{[\text{theorem: } 16.32]}{=} (D_x(g \circ f) \circ D_x f)(1) \\
&\underset{[\text{theorem: } 16.36]}{=} (D_{f(x)} g \circ D_x f)(1) \\
&= D_{f(x)} g(D_x f(1)) \\
&\underset{[\text{theorem: } 16.32]}{=} D_{f(x)} g(f'_x) \\
&\underset{[\text{theorem: } 16.32]}{=} f'_x \cdot g'_{f(x)}
\end{aligned}$$

□

Corollary 16.38. *Let $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle, \langle Z, \|\|_Z\rangle$ be normed spaces, $U \subseteq X$ a open set in X, $x \in U$, $f: U \to Y$ a function that is Fréchet differentiable at x and $L \in L(Y, Z)$ then*

$$L \circ f: U \to X \textit{ is Fréchet differentiable at } x$$

and

$$D_x(L \circ f) = L \circ D_x f$$

Proof. Using [example: 16.27] L is Fréchet differentiable at $f(x)$ and $D_{f(x)} L = L$. Hence using [theorem: 16.36] $L \circ f$ is Fréchet differentiable at x and

$$D_x(L \circ f) = D_{f(x)} L \circ D_x f = L \circ D_x f$$

□

16.1.5 Partial differentials

Definition 16.39. *Let $n\in\mathbb{N}\setminus\{1\}$, $\{X_i\}_{i\in[1,\ldots,n]}$ a finite family of sets, $x\in\prod_{i\in\{1,\ldots,n\}}X_i$, $i\in\{1,\ldots,n\}$, $t\in X_i$ then $I^{[i,x]}\colon X_i\rightarrow\prod_{i\in\{1,\ldots,n\}}X_i$ is defined by*

$$I^{[i,x]}(t)=\left(x_1,\ldots\underbrace{t}_{i}\ldots,x_n\right)$$

or more formally

$$(I^{[i,x]}(t))_j=\begin{cases} t \text{ if } j=i\\ x_j \text{ if } j\in\{1,\ldots,n\}\setminus\{i\}\end{cases}$$

***Note* 16.40.** *$I^{[i,x]}$ is not a function because we do not specify the domain, codomain. This is*

Lemma 16.41. *Let $n\in\mathbb{N}\setminus\{1\}$, $\{\langle X_i,\|\|_i\rangle\}_{i\in[1,\ldots,n]}$ be a finite family of normed spaces, $\left\langle\prod_{j\in\{1,\ldots,n\}}X_j,\|\|_{\max}\right\rangle$ the product space equipped with the maximum norm [see theorem: 14.94] then we have $\forall i\in\{1,\ldots,n\}$ that:*

1. *Then $I^{[i,0}\in L\left(X_i,\prod_{j\in\{1,\ldots,n\}}X_i\right)$ and $\forall t\in X_i$ we have $\|t\|_i=\|I^{[i,0]}(t)\|_{\max}$*
2. *If $x=(x_1,\ldots,x_n)\in\prod_{j\in\{1,\ldots,n\}}X_j$ then $I^{[i,x]}(x_i)=x$*
3. *If $x=(x_1,\ldots,x_n)\in\prod_{j\in\{1,\ldots,n\}}X_j$ then $x=\sum_{i=1}^n I^{[i,0]}(x_i)$*
4. *If $x=(x_1,\ldots,x_n)\in\prod_{j\in\{1,\ldots,n\}}X_j$, U a open set in X_i then $(I^{[i,x]})_{|U}$ is Fréchet differentiable on U with $\forall t\in U$ $D_t(I^{[i,x]})_{|U}=I^{[i,0]}$.*
5. *If $x=(x_1,\ldots,x_n)\in\prod_{j\in\{1,\ldots,n\}}X_j$, U a open set in X_i then $(I^{[i,x]})_{|U}$ is continuous.*
6. *$\forall i,k\in\{1,\ldots,n\}$, $\forall t\in X_i$ we have $(I^{[i,0]}(t))_k=\delta_{i,k}\cdot t$*
7. *If $x=(x_1,\ldots,x_n)\in\prod_{j\in\{1,\ldots,n\}}X_j$ then $\forall i\in\{1,\ldots,n\}$, $t\in X_i$ we have $I^{[i,x]}(t)=x+I^{[i,0]}(t)$*

Proof.

1. Let $r,s\in X_i$, $\alpha\in\mathbb{K}$ then

$$\begin{aligned} I^{[i,0]}(r+\alpha\cdot s) &= \left(0\ldots\underbrace{r+\alpha\cdot s}_{i}\ldots0\right)\\ &= \left(0\ldots\underbrace{r}_{i}\ldots0\right)+\alpha\cdot\left(0\ldots\underbrace{s}_{i}\ldots0\right)\\ &= I^{[i,0]}(r)+\alpha\cdot I^{[i,0]}(s)\\ &= L(r)+\alpha\cdot L(s)\end{aligned}$$

   proving that

$$L\in\mathrm{Hom}\left(X_i,\prod_{j\in\{1,\ldots,n\}}X_j\right)$$

   Further if $t\in X_i$ then for $j\in\{1,\ldots,n\}$ we have

$$\pi_j(I^{[i,0]}(t))=\pi_j\left(\left(0,\ldots\underbrace{t}_{j}\right)\ldots,0\right)=\begin{cases} t \text{ if } j=i\\ 0_j \text{ if } j\in\{1,\ldots,n\}\setminus\{i\}\end{cases}\tag{16.17}$$

   so that

$$\begin{aligned}\|I^{[i,0]}(t)\|_{\max} &= \max\left(\{\|\pi_j(I^{[i,0]}(t))\|_j|j\in\{1,\ldots,n\}\}\right)\\ &= \max\left(\{\|\pi_j(I^{[i,0]}(t))\|_j|j\in\{1,\ldots,n\}\setminus\{i\}\}\bigcup\{\|\pi_j(I^{[i,0]}(t))\|_j|j\in\{i\}\}\right)\\ &= \max\left(\{\|0_j\|_j\}\bigcup\{\|t\|_i\}\right)\\ &= \max\left(\{0,\|t\|_i\}\right)\\ &= \|t\|_i\end{aligned}$$

   proving that

$$\|I^{[i,0]}(t)\|_{\max}=\|t\|_i=1\cdot\|t\|_i$$

Hence by [theorem: 14.174] it follows that

$$I^{[i,0]} \in L\left(X_i, \prod_{j\in\{1,\ldots,n\}} X_j\right)$$

2. This follows from $I^{[x,i]}(x) = \left(x_1, \ldots \underbrace{x_i}_{i} \ldots, x_n\right) = x$

3. Let $x \in \prod_{j\in\{1,\ldots,n\}} X_j$ and $j \in \{1, \ldots, n\}$ then

$$\begin{aligned}\left(\sum_{i=1}^{n} I^{[i,0]}(x_i)\right)_j &= \left(\sum_{i\in\{1,\ldots,n\}} I^{[i,0]}(x_i)\right)_j \\ &= \pi_j\left(\sum_{i\in\{1,\ldots,n\}} I^{[i,0]}(x_i)\right) \\ &= \sum_{i\in\{1,\ldots,n\}} \pi_j(I^{[i,0]}(x_i)) \\ &= \sum_{i\in\{1,\ldots,n\}\setminus\{j\}} \pi_j(I^{[i,0]}(x_i)) + \sum_{i\in\{j\}} \pi_j(I^{[i,0]}(x_i)) \\ &= 0 + x_j \\ &= x_j\end{aligned}$$

proving that

$$x = \sum_{i=1}^{n} I^{[i,0]}(x_i)$$

4. Let $x \in \prod_{j\in\{1,\ldots,n\}} X_j$ and $t \in U$. Let $\varepsilon \in \mathbb{R}^+$ then we have $\forall h \in X_i = (X_i)_x$ that

$$\begin{aligned}\|(I^{[i,x]})_{|U}(t+h) - (I^{[i,x]})_{|U}(t) - I^{[i,0]}(h)\|_{\max} &= \\ \|I^{[i,x]}(t+h) - I^{[i,x]}(t) - I^{[i,0]}(h)\|_{\max} &= \\ \|(x_1, \ldots, x_{i-1}, t+h, x_i, \ldots, x_n) - (x_1, \ldots, x_{i-1}, t, x_{i+1}, \ldots, x_n) - I^{[i,0]}(h)\|_{\max} &= \\ \|(0_1, \ldots, 0_{i-1}, h, 0_{i+1}, \ldots, 0_n) - I^{[i,0]}(h)\|_{\max} &= \\ \|(0_1, \ldots, 0_{i-1}, h, 0_{i+1}, \ldots, 0_n) - (0_1, \ldots, 0_{i-1}, h, 0_{i+1}, \ldots, 0_n)\|_{\max} &= \\ \|(0_1, \ldots, 0_n)\|_{\max} &= \\ 0 &< \\ \varepsilon \cdot \|h\|_i\end{aligned}$$

which together with (1) proves that

$$(I^{[i,x]})_{|U} \text{ is Fréchet differentiable at } t \text{ with differential } D_t I^{[i,x]} = I^{[i,0]}$$

5. Let $x \in \prod_{j\in\{1,\ldots,n\}} X_j$ then by (4) we have that $(I^{[i,x]})_{|U}$ is Fréchet differentiable at t and using [theorem: 16.31] it follows then that $I^{[i,x]}$ is continuous at t. Hence $I^{[i,x]}$ is continuous mapping between X_i and $\prod_{j\in\{1,\ldots,n\}} X_j$.

6. Let $i \in \{1, \ldots, n\}$ and $t \in X_i$ then we have for $k \in \{1, \ldots, n\}$

$$(I^{[i,0]}(t))_k = \left(0, \ldots \underbrace{t}_{i} \ldots 0\right)_k = \delta_{i,k} \cdot t$$

□

Definition 16.42. *Let $n \in \mathbb{N} \setminus \{1\}$, $\{X_i\}_{i\in[1,\ldots,n]}$ be a finite family of sets, $U \subseteq \prod_{j\in\{1,\ldots,n\}} X_j$, $x \in U$, Y a set, $i \in \{1, \ldots, n\}$ and $f: U \rightarrow Y$ a function then we define*

$$U_{[i,x]} = (I^{[i,x]})^{-1}(U) = \{t \in X_i | I^{[i,x]}(t) \in U\} = \{t \in X_i | (x_1 \ldots \underbrace{t}_{} \ldots x_{n_i})\}$$

and

$$f^{[i,x]}: \{t \in X_i | I^{[i,x]}(t) \in U\} \rightarrow Y \text{ by } f^{[i,x]} = f \circ (I^{[i,x]})_{|U_{[i,x]}}$$

So $\forall t \in U_{[i,x]}$ *we have*

$$f^{[i,x]}(t) = (f \circ I^{[i,x]})(t) = f(I^{[i,x]}(t)) = f\left(x_1, \ldots \underbrace{t}_{i} \ldots, x_n\right)$$

We are now ready to define the partial differential of a multi-parameter function.

Definition 16.43. *Let* $n \in \mathbb{N} \setminus \{1\}$, $\{\langle X_i, \|\|_i\rangle\}_{i \in \{1,\ldots,n\}}$ *a finite family of normed spaces,* $\langle Y, \|\|\rangle$ *a normed space,* $\left\langle \prod_{j \in \{1,\ldots,n\}} X_j, \|\|_{\max}\right\rangle$ *the product space equipped with the maximum norm [see theorem: 14.94],* $U \subseteq \prod_{j \in \{1,\ldots,n\}} X_j$ *a open set,* $x \in U$, $i \in \{1,\ldots,n\}$ *and*

$$f: U \rightarrow Y \text{ a function}$$

then

$$f \text{ is } i\text{-partial differentiable at } x \text{ with } i\text{-partial differential } D_{x,i}f$$

if

$$f^{[i,x]}: U_{[i,x]} \rightarrow Y \text{ is Fréchet differentiable at } x_i \text{ and } D_{x,i}f = D_{x_i}f^{[i,x]} \in X_i$$

***Note* 16.44.** *Using [lemma: 16.41]* $I^{[i,x]}(x_i) = x \in U$ *and* $I^{[i,x]}$ *is continuous, hence*

$$U_{[i,x]} = (I^{[i,x]})^{-1}(U)$$

is open and $x_i \in U_{[i,x]}$, *which is required for the Fréchet differentiation definition.*

It turns out that partial differentiability is a local property.

Theorem 16.45. *Let* $n \in \mathbb{N} \setminus \{1\}$, $\{\langle X_i, \|\|_i\rangle\}_{i \in \{1,\ldots,n\}}$ *a finite family of normed spaces,* $\langle Y, \|\|\rangle$ *a normed space,* $\left\langle \prod_{j \in \{1,\ldots,n\}} X_j, \|\|_{\max}\right\rangle$ *the product space equipped with the maximum norm [see theorem: 14.94],* $U \subseteq \prod_{j \in \{1,\ldots,n\}} X_j$ *a open set,* $x = (x_1, \ldots, x_n) \in U$, $i \in \{1, \ldots, n\}$ *and*

$$f: U \rightarrow Y \text{ a function}$$

then for V *a open set with* $x \in V \subseteq U$ *we have*

$$f \text{ is } i\text{-partial differentiable at } x$$

$$\Updownarrow$$

$$f_{|V} \text{ is } i\text{-partial differentiable at } x$$

Further if f *or* $f_{|V}$ *is* i*-partial differentiable at* x *then* $D_{x,i}f = D_{x,i}f_{|V}$.

Proof. For $t \in V_{[i,x]}$ we have that

$$\begin{aligned} (f_{|V})^{[i,x]}(t) &= f_{|V}\left(x_1 \ldots \underbrace{t}_{i} \ldots x_n\right) \\ &= f\left(x_1 \ldots \underbrace{t}_{i} \ldots x_n\right) \\ &= f^{[i,x]}(t) \end{aligned}$$

so that

$$(f_{|V})^{[i,x]} = (f^{[i,x]})_{|V_{[i,x]}} \tag{16.18}$$

Further we have

$$x_i \in V_{[i,x]} = (I^{[i,x]})^{-1}(V) \subseteq (I^{[i,x]})^{-1}(U) = U_{[i,x]}$$

Hence we have:

⇒. If f is i-partial differentiable at x with i-partial differential $D_{x,i}f$ then by definition $f^{[i,x]}$ is Fréchet differentiable at x_i and $D_{x,i}f = D_{x_i}f^{[i,x]}$. Using [theorem: 16.24] it follows that $(f^{[i,x]})_{|V_{[i,x]}}$ is Fréchet differentiable at x_i and $D_{x_i}(f^{[i,x]})_{|V_{[i,x]}} = D_{x_i}f^{[i,x]}$. Hence using [eq: 16.18] $(f_{|V})^{[i,x]}$ is Fréchet differentiable at x_i with $D_{x_i}(f_{|V})^{[i,x]} = D_{x_i}f^{[i,x]}$ or using the definition of partial differentiation

$$f_{|V} \text{ is } i\text{-partial differentiable at } x \text{ and } D_{x,i}f_{|V} = D_{x,i}f$$

⇐. If $f_{|V}$ is i-partial differentiable at x with i-partial differential $D_{x,i}f_{|V}$ then by definition $(f_{|V})^{[i,x]}$ is Fréchet differentiable at x_i and $D_{x,i}(f_{|V})=D_{x_i}(f_{|V})^{[i,x]}$. Using [eq: 16.18] we have that $(f^{[i,x]})_{|V_{[i,x]}}$ is Fréchet differentiable at x_i and

$$D_{x,i}f_{|V}=D_{x_i}(f_{|V})^{[i,x]}=D_{x_i}(f^{[i,x]})_{|V_{[i,x]}}$$

Using [theorem: 16.24] again it follows that $f^{[i,x]}$ is Fréchet differentiable at x_i with $D_{x_i}(f^{[i,x]})_{|V_{[i,x]}}=D_{x_i}f^{[i,x]}$. Hence by definition

$$f \text{ is partial differentiable at } x \text{ and } D_{x,i}f_{|V}=D_{x,i}f$$

□

Corollary 16.46. *Let $n\in\mathbb{N}\setminus\{1\}$, $\{\langle X_i,\|\|_i\rangle\}_{i\in\{1,\ldots,n\}}$ a finite family of normed spaces, $\langle Y,\|\|\rangle$ a normed space, $\left\langle\prod_{j\in\{1,\ldots,n\}}X_j,\|\|_{\max}\right\rangle$ the product space equipped with the maximum norm [see theorem: 14.94], $U,V\subseteq\prod_{j\in\{1,\ldots,n\}}X_j$ open sets, $x=(x_1,\ldots,x_n)\in U\bigcap V$, $i\in\{1,\ldots,n\}$ and*

$$f: U\rightarrow Y,\ g: V\rightarrow Y$$

two functions such that there exist a open set $W\subseteq U\bigcap V$ with $x\in W$ and $\forall y\in W$ that $f(y)=g(y)$. Then if f is i-partial differentiable at x we have that g is i-partial differentiable at x and $D_{x,i}f=D_{i,y}g$.

Proof. As $\forall y\in W$ we have that $f(y)=g(y)$ it follows that $f_{|W}=g_{|W}$. As f is i-partial differentiable at x it follows from [theorem: 16.45] that $f_{|W}$ is i-partial differentiable at x with $D_{x,i}f=D_{x,i}f_{|W}\underset{f_{|W}=g_{|W}}{=}Dg_{|W}(x)$. Hence using [theorem: 16.45] again it follows that g is i-partial differentiable at x with $D_{x,i}g=D_{x,i}g_{|W}=D_{x,i}f$. □

Fréchet differentiability of a multi-parameter function implies partial differentiability.

Theorem 16.47. *Let $n\in\mathbb{N}\setminus\{1\}$, $\{\langle X_i,\|\|_i\rangle\}_{i\in\{1,\ldots,n\}}$ a finite family of normed spaces, $\langle Y,\|\|\rangle$ a normed space, $\left\langle\prod_{j\in\{1,\ldots,n\}}X_j,\|\|_{\max}\right\rangle$ the product space equipped with the maximum norm [see theorem: 14.94], $U\subseteq\prod_{j\in\{1,\ldots,n\}}X_j$ a open set, $x\in U$ and $f: U\rightarrow Y$ a function that is Fréchet differentiable at x then $\forall i\in\{1,\ldots,n\}$ f is i-partial differentiable at x with i-partial differential*

$$D_{x,i}f=D_xf\circ I^{[i,0]}$$

Further we have

$$D_xf=\sum_{i=1}^n(D_{x,i}f\circ\pi_i)$$

In other words $\forall h\in\prod_{i\in\{1,\ldots,n\}}X_i$ we have

$$D_xf(h)=\sum_{i=1}^n D_{x,i}f(h_i)$$

Proof. Using [lemma: 16.41] together with [theorem: 16.24] it follows that

$$(I^{[i,x]})_{|U_{[i,x]}}: U_{[i,x]}\rightarrow\prod_{j\in\{1,\ldots,n\}}X_j \text{ is Fréchet differentiable at } x_i$$

with

$$D_{x_i}(I^{[i,x]})_{|U_{[i,x]}}\underset{\text{[lemma: 16.41]}}{=}I^{[i,0]} \tag{16.19}$$

Further, as $f: U\rightarrow Y$ is Fréchet differentiable at $x=I^{[i,x]}(x_i)=(I^{[i,x]})_{|U_{[i,x]}}(x_i)$ and

$$(I^{[i,x]})_{|U_{[i,x]}}((I^{[i,x]})^{-1}(U))\subseteq U$$

we can use the Chain Rule [theorem: 16.36] so that

$$f^{[i,x]}=f\circ(I^{[i,x]})_{|U_{[i,x]}} \text{ is differentiable at } x_i$$

with

$$\begin{aligned}
D_{x_i}f^{[i,x]} &= D_{x_i}(f\circ(I^{[i,x]})_{|U_{[i,x]}})\\
&= D_{(I^{[i,x]})_{|U_{[i,x]}}(x_i)}f\circ D_{x_i}(I^{[i,x]})_{|U_{[i,x]}}\\
&= D_xf\circ D_{x_i}(I^{[i,x]})_{|U_{[i,x]}}\\
&\underset{\text{[eq: 16.19]}}{=} D_xf\circ I^{[i,0]}
\end{aligned}$$

So that by definition

$$f \text{ is } i\text{-partial differentiable at } x \text{ with } D_{x,i}f(x)=D_xf\circ I^{[i,0]} \tag{16.20}$$

Further if $h \in \prod_{j \in \{1, \ldots, n\}} X_j$ we have

$$\begin{aligned} D_x f(h) &\underset{[\text{lemma: } 16.41]}{=} D_x f\left(\sum_{i=1}^{n} I^{[i,0]}(h_i)\right) \\ &= \sum_{i=1}^{n} D_x f(I^{[i,0]}(h_i)) \\ &= \sum_{i=1}^{n} (D_x f \circ I^{[i,0]})(h_i) \\ &\underset{[\text{eq: } 16.20]}{=} \sum_{i=1}^{n} D_{x,i} f(h_i) \\ &= \sum_{i=1}^{n} D_{x,i} f(\pi_i(h)) \\ &= \sum_{i=1}^{n} (D_{x,i} f \circ \pi_i)(h) \\ &= \left(\sum_{i=1}^{n} (D_{x,i} f \circ \pi_i)\right)(h) \end{aligned}$$

so that we have

$$D_x f = \sum_{j=1}^{n} (D_{x,i} f \circ \pi_i)$$

□

Example 16.48. Let $n \in \mathbb{N} \setminus \{1\}$, $\{\langle X_i, \|\|_i \rangle\}_{i \in \{1, \ldots, n\}}$ a finite family of normed spaces, $\langle Y, \|\| \rangle$ a normed space, $\left\langle \prod_{j \in \{1, \ldots, n\}} X_j, \|\|_{\max} \right\rangle$ the product space equipped with the maximum norm [see theorem: 14.94], $U \subseteq \prod_{j \in \{1, \ldots, n\}} X_j$ a open set, $x \in U$ and $y \in Y$ then $\forall i \in \{1, \ldots, n\}$ we have that the constant function

$$C_y: U \rightarrow Y \text{ defined by } C_y(x) = y$$

has a i-partial differential at x with

$$D_{y,i} C_i = C_0$$

where

$$C_0: X_i \rightarrow Y \text{is defined by } C_0(h) = 0$$

Proof. By [example: 16.26] C_y is Fréchet differentiable at x with $D_x C_y = C_0$ where

$$C_0: X \rightarrow Y \text{ is defined by } X_0(x) = 0$$

Hence using [theorem: 16.47] we have $\forall i \in \{1, \ldots, n\}$ that C_y has a i-partial derivative with

$$D_{x,i} C_y = C_0 \circ I^{[i,0]} = C_0$$

[because $\forall h \in X_i$ $(C_0 \circ I^{[i,0]})(h) = C_0(I^{[i,0]}(h)) = 0$]. □

Note 16.49. Although Fréchet differentiability of a multi parameter function implies that all the partial differentials exist the opposite is not true as the following example proves.

Example 16.50. Define

$$f: \mathbb{R}^2 \rightarrow \mathbb{R} \text{ by } f(x,y) = \begin{cases} \frac{x \cdot y}{x^2 + y^2} \text{ if } (x,y) \in \mathbb{R}^2 \setminus \{(0,0)\} \\ 0 \text{ if } (x,y) = (0,0) \end{cases}$$

then f is not Fréchet differentiable at $(0,0)$ however f is 1-partial and 2-partial differentiable at $(0,0)$

Proof. Let $\varepsilon = \frac{1}{4}$, $\delta \in \mathbb{R}^+$ and consider $z = \left(\frac{\delta}{2}, \frac{\delta}{2}\right)$ then we have

$$\|z\|_{\max} = \max\left(\left\{\left|\frac{\delta}{2}\right|, \left|\frac{\delta}{2}\right|\right\}\right) = \frac{\delta}{2} < \delta$$

However

$$
\begin{array}{rcl}
|f(z)-f(0,0)| & \underset{(\delta,\delta)\neq(0,0)}{=} & \left|\dfrac{\frac{\delta}{2}\cdot\frac{\delta}{2}}{\left(\frac{\delta}{2}\right)^2+\left(\frac{\delta}{2}\right)^2}-0\right| \\
& = & \left|\dfrac{1}{2}\right| \\
& > & \varepsilon
\end{array}
$$

which proves that f is not continuous at $(0,0)$. So f can not be Fréchet differentiable at $(0,0)$ because Fréchet differentiable functions are continuous [see theorem: 16.31]. However we have that

$$f^{[1,(0,0)]}(t)=f(t,0)=\begin{cases}\frac{t\cdot 0}{t^2+0^2} \text{ if } (t,0)\neq(0,0)\\ 0 \text{ if } (t,0)=0\end{cases}=0=C_0(t)$$

$$f^{[2,(0,0)]}(t)=f(t,0)=\begin{cases}\frac{0\cdot t}{0^2+t^2} \text{ if } (t,0)\neq(0,0)\\ 0 \text{ if } (t,0)=0\end{cases}=0=C_0(t)$$

proving that

$$f^{[1,[(0,0)]]}=C_0=f^{[2,[(0,0)]]}$$

By [example: 16.26] $f^{[1,[(0,0)]]}$, $f^{[2,[(0,0)]]}$ are Fréchet differentiable at 0 with $D_0 f^{[1,[(0,0)]]}=C_0=D_0 f^{[2,[(0,0)]]}$. Hence f is is 1-partial and 2-partial differentiable at $(0,0)$ with

$$D_{(0,0),1}f=C_0=D_{(0,0),2}f$$

□

As the partial differential is based on the Fréchet differential of $f\circ(I^{[i,x]})_{|(I^{[i,x]})^{-1}(U)}$ many properties of the Fréchet differential also applies.

Theorem 16.51. *Let $n\in\mathbb{N}\setminus\{1\}$, $\{\langle X_i,\|\|_i\rangle\}_{i\in\{1,\ldots,n\}}$ a finite family of normed spaces, $\left\langle\prod_{j\in\{1,\ldots,n\}}X_j,\|\|_{\max}\right\rangle$ the product space equipped with the maximum norm [see theorem: 14.94], $\langle Y,\|\|\rangle$ a normed space, $U\subseteq\prod_{i\in\{1,\ldots,n\}}X_i$ a open set, $x\in U$ and $i\in\{1,\ldots,n\}$ then we have:*

1. *If $f:U\rightarrow Y$ and $g:U\rightarrow Y$ are functions that are i-partial differentiable at x then*

$$f+g:U\rightarrow Y$$

is i-partial differentiable at x with

$$D_{x,i}(f+g)=D_{x,i}f+D_{x,i}g$$

2. *If $\alpha\in\mathbb{K}$ and $f:U\rightarrow Y$ is a function that is i-partial differentiable at x then*

$$\alpha\cdot f:U\rightarrow Y$$

is i-partial differentiable at x with

$$D_{x,i}(\alpha\cdot f)=\alpha\cdot D_{x,i}f$$

3. *If $f:X\rightarrow Y$ is such that $f\circ I^{[i,x]}\in L(X_i,Y)$ then f is i-partial differentiable and*

$$D_{x,i}f=f\circ I^{[i,x]}$$

Proof.

1. As f,g are i-partial differentiable at x we have by definition that $f\circ(I^{[i,x]})_{|U_{[i,x]}}$ and $(g\circ I^{[i,x]})_{|U_{[i,x]}}$ are Fréchet differentiable at x_i with

$$D_{x,i}f=D_{x_i}(f\circ(I^{[i,x]})_{|U_{[i,x]}}) \text{ and } D_{x,i}g=D_{x_i}(g\circ(I^{[i,x]})_{|U_{[i,x]}}) \tag{16.21}$$

Using [theorem: 16.35] we have that

$$f\circ(I^{[i,x]})_{|U_{[i,x]}})+g\circ(I^{[i,x]})_{|U_{[i,x]}} \text{ is Fréchet differentiable at } x_i$$

with

$$D_{x_i}(f\circ(I^{[i,x]})_{|U_{[i,x]}}+g\circ(I^{[i,x]})_{|U_{[i,x]}})=D_{x_i}(f\circ(I^{[i,x]})_{|U_{[i,x]}})+D_{x_i}(g\circ(I^{[i,x]})_{|U_{[i,x]}})$$

As $(f+g)\circ(I^{[i,x]})_{|U_{[i,x]}}=f\circ(I^{[i,x]})_{|U_{[i,x]}}+g\circ(I^{[i,x]})_{|U_{[i,x]}}$ it follows that $f+g$ is Fréchet differentiable at x_i with

$$D_{x_i}((f+g)\circ(I^{[i,x]})_{|U_{[i,x]}})=D_{x_i}(f\circ(I^{[i,x]})_{|U_{[i,x]}})+D_{x_i}(g\circ(I^{[i,x]})_{|U_{[i,x]}})\underset{[\text{eq: }16.21]}{=}D_{x,i}f+D_{x,i}g$$

Hence by definition

$$f+g \text{ is } i\text{-partial differentiable with } D_{x,i}(f+g)(x)=D_{x,i}f(x)+D_{x,i}g$$

2. As f is i-partial differentiable we have by definition that $f\circ(I^{[i,x]})_{|U_{[i,x]}}$ is Fréchet differentiable at x_i with

$$D_{x,i}f=D_{x_i}(f\circ(I^{[i,x]})_{|U_{[i,x]}}) \tag{16.22}$$

Using [theorem: 16.35] we have that

$$\alpha\cdot(f\circ I^{[i,x]}) \text{ is Fréchet differentiable at } x_i$$

with

$$D_{x_i}(\alpha\cdot(f\circ(I^{[i,x]})_{|U_{[i,x]}}))=\alpha\cdot D_{x_i}(f\circ(I^{[i,x]})_{|U_{[i,x]}})$$

As $(\alpha\cdot f)\circ(I^{[i,x]})_{|U_{[i,x]}}=\alpha\cdot(f\circ(I^{[i,x]})_{|U_{[i,x]}})$ it follows that $\alpha\cdot f$ is Fréchet differentiable at x_i with

$$D_{x_i}((\alpha f)\circ(I^{[i,x]})_{|U_{[i,x]}})=\alpha\cdot D_{x_i}(f\circ(I^{[i,x]})_{|U_{[i,x]}})\underset{[\text{eq: }16.22]}{=}\alpha\cdot D_{x,i}f$$

Hence by definition

$$\alpha\cdot f \text{ is } i\text{-partial differentiable with } D_{x,i}(\alpha\cdot f)=\alpha\cdot D_{x,i}f$$

3. As $f\circ I^{[i,x]}\in L(X_i,Y)$ we have by [example: 16.27] that $f\circ I^{[i,x]}$ is Fréchet differentiable at x_i with $D_{x_i}(f\circ I^{[i,x]})=f\circ I^{[i,x]}$. Hence by definition we have that

$$f \text{ is } i\text{-partial differentiable at } x \text{ with } D_{x,i}f(x)=f\circ I^{[i,x]}$$

□

We have the following version of the chain rule for partial differentiation.

Theorem 16.52. *Let $n\in\mathbb{N}\setminus\{1\}$, $\{\langle X_i,\|\|_i\rangle\}_{i\in\{1,\ldots,n\}}$ a finite family of vector spaces, $\left\langle\prod_{i\in\{1,\ldots,n\}}X_i,\|\|_{\max}\right\rangle$ the product space equipped with the maximum norm [see theorem: 14.94], $\langle Y,\|\|_Y\rangle$, $\langle Z,\|\|_Z\rangle$ normed spaces, $U\subseteq\prod_{i\in\{1,\ldots,n\}}X_i$ a open set in $\prod_{i\in\{1,\ldots,n\}}X_i$, $V\subseteq Y$ a open set in Y, $x\in X$,*

$$f:U\rightarrow Y \text{ a function differentiable at } x \text{ such that } f(U)\subseteq V$$

and

$$g:V\rightarrow Z \text{ a function differentiable at } f(x)$$

then

$$\forall i\in\{1,\ldots,n\}\ g\circ f \text{ is } i\text{-partial differentiable at } x$$

and

$$D_{x,i}(g\circ f)=D_{f(x)}g\circ D_{x,i}f$$

Proof. Using the Chain Rule [theorem: 16.36] it follows that

$$g\circ f:U\rightarrow Z$$

is differentiable at x with

$$D_x(g\circ f)=D_{f(x)}g\circ D_xf \tag{16.23}$$

Hence, using [theorem: 16.47] it follows that $g\circ f$ is i-partial differentiable at x with i-partial differential

$$D_{x,i}(g\circ f)=D_x(g\circ f)\circ I^{[i,0]}$$

Substituting [eq: 16.23] in the above gives

$$\begin{aligned}D_{x,i}(g\circ f) &= (D_{f(x)}g\circ D_xf)\circ I^{[i,0]}\\ &= D_{f(x)}g\circ(D_xf\circ I^{[i,0]})\\ &\underset{[\text{theorem: }16.47]}{=} D_{f(x)}g\circ D_{x,i}f\end{aligned}$$

proving that

$$D_{x,i}(g\circ f)=D_{f(x)}g\circ D_{x,i}f \tag{16.24}$$ □

Next we define the calculus partial derivative of a multi-parameter function.

Definition 16.53. *Let $n\in\mathbb{N}\setminus\{1\}$, $\langle\mathbb{K}^n,\|\|_n\rangle$ the vector space $\mathbb{K}^n$ equipped with the maximum norm $\|\|_n$, $\langle X,\|\|\rangle$ a normed space, $U\subseteq\mathbb{K}^n$ a open set in $\mathbb{K}^n$, $x\in U$, $i\in\{1,\ldots,n\}$ and*

$$f:U\rightarrow X \text{ a function}$$

*then f has a **i-partial derivative at x** noted by $\partial_{x,i}f(x)\in X$ if $f\circ(I^{[i,x]})_{|U_{[i,x]}}$ has a derivative at x_i, the **i-partial derivative at x** is defined by*

$$\partial_{x,i}f(x)=(f\circ(I^{[i,x]})_{|U_{[i,x]}})'_{x_i}\in X$$

Using [definitions: 16.39 and 16.12] this is equivalent with

$$\forall\varepsilon\in\mathbb{R}^+ \text{ there exist a } \delta\in\mathbb{R}^+ \text{ such that } \forall h\in(U_{[i,x]})_{x_i} \text{ with } 0<|h|<\delta \text{ we have}$$

$$\begin{aligned}\left\|\frac{f\left(x_1,\ldots\underbrace{x_i+h}_{i}\ldots,x_n\right)-f(x_1,\ldots,x_n)}{h}-\partial_{x,i}f\right\| &=\\ \left\|\frac{f(I^{[i,x]}(x+h))-f(I^{[i,x]}(x))}{h}-\partial_{x,i}f\right\| &=\\ \left\|\frac{(f\circ I^{[i,x]})(x+h)-(f\circ I^{[i,x]})(x)}{h}-\partial_{x,i}f(x)\right\| &< \varepsilon\end{aligned}$$

We have the following equivalent definition for the partial derivative of a function.

Theorem 16.54. *Let $n\in\mathbb{N}\setminus\{1\}$, $\langle\mathbb{K}^n,\|\|_n\rangle$ the vector space $\mathbb{K}^n$ equipped with the maximum norm $\|\|_n$, $\langle X,\|\|\rangle$ a normed space, $U\subseteq\mathbb{K}^n$ a open set in $\mathbb{K}^n$, $x\in U$, $i\in\{1,\ldots,n\}$ and*

$$f:U\rightarrow X \text{ a function}$$

then

$$f \text{ has a } i\text{-partial derivative at } x\Leftrightarrow f \text{ has a } i\text{-partial differential at } x$$

Further if f has a i-partial derivative or differential at x then

$$\partial_{x,}f(x)=D_{x,i}f(1)$$

and

$$\forall h\in\mathbb{K} \text{ we have } D_{x,i}f(h)=h\cdot\partial_{x,i}f$$

Proof.

⇒. As f has a i-partial derivative at x $f\circ(I^{[i,x]})_{|U_{[i,x]}}$ has a derivative at x_i. Using [theorem: 16.32] it follows that $f\circ(I^{[i,x]})_{|U_{[i,x]}}$ is Fréchet differentiable at x_i, hence by definition

$$f \text{ is } i\text{-partial differentiable at } x$$

Further by [theorem: 16.32] we have also

$$(f\circ(I^{[i,x]})_{|U_{[i,x]}})'_{x_i}=D_{x_i}(f\circ(I^{[i,x]})_{|U_{[i,x]}})(1)$$

and

$$\forall h\in\mathbb{K}\ D_{x_i}(f\circ(I^{[i,x]})_{|U_{[i,x]}})=h\cdot(f\circ(I^{[i,x]})_{|U_{[i,x]}})'_{x_i}$$

As $\partial_{x,i}f\underset{\text{def}}{=}(f\circ(I^{[i,x]})_{|U_{[i,x]}})'_{x_i}$ and $D_{x,i}f\underset{\text{def}}{=}D_{x_i}(f\circ(I^{[i,x]})_{|U_{[i,x]}})$ we have

$$\partial_{x,i}f=D_{x,i}f(1) \text{ and } \forall h\in\mathbb{K} \text{ that } D_{x,i}f(h)=h\cdot\partial_{x,i}f$$

⇐. As f has a i-partial differential at x $f\circ(I^{[i,x]})_{|U_{[i,x]}}$ is Fréchet differentiable at x_i. Using [theorem: 16.32] it follows that $f\circ(I^{[i,x]})_{|U_{[i,x]}}$ has a derivative at x_i and thus by definition that f has a i-partial derivative at x. □

Using the above definition it is easy to prove that partial derivatives are local.

Corollary 16.55. *Let $n\in\mathbb{N}\setminus\{1\}$, $\langle\mathbb{K}^n,\|\|_n\rangle$ the vector space $\mathbb{K}^n$ equipped with the maximum norm $\|\|_n$, $\langle X,\|\|\rangle$ a normed space, U, $V\subseteq\mathbb{K}^n$ open sets in $\mathbb{K}^n$, $x\in U\bigcap V$, $i\in\{1,\ldots,n\}$ and*

$$f{:}U\rightarrow Y \text{ and } g{:}V\rightarrow Y$$

functions such that there exist a open set W with $x\in W\subseteq U\bigcap V$ and $\forall y\in W$ $f(y)=g(y)$ Then if f has a i-partial derivative at x g has also a i-partial derivative at x and $\partial_{x,i}f=\partial_{x,i}g$.

Proof. As f has a i-partial derivative at x it follows from [theorem: 16.54] that f has a i-partial differential at x and $\partial_{x,i}f=D_{x,i}f(1)$. As partial differentiation is local [see theorem: 16.46] it follows that g has a i-partial differential at x and $D_{x,i}f=D_{x,i}g$. Hence using [theorem: 16.54] again g has a i-partial derivative at x and $\partial_{x,i}g=D_{x,i}g(1)=D_{x,i}f(1)=\partial_{x,i}g$. □

We can rephrase [theorem: 16.47] as follows for partial derivatives.

Theorem 16.56. *Let $n\in\mathbb{N}\setminus\{1\}$, $\langle\mathbb{K}^n,\|\|_n\rangle$ the vector space $\mathbb{K}^n$ equipped with the maximum norm $\|\|_n$, $\langle X,\|\|\rangle$ a normed space, $U\subseteq\mathbb{K}^n$ a open set in $\mathbb{K}^n$, $x\in U$ and*

$$f{:}U\rightarrow X \text{ a function that is Fréchet differentiable at } x$$

then we have:

1. *$\forall i\in\{1,\ldots,n\}$ f has a i-partial derivative $\partial_{x,i}f$ at x and $D_{x,i}f\in L(\mathbb{K},X)$ is defined by*

$$D_{x,i}f(h)=h\cdot\partial_{x,i}f$$

so that

$$\partial_{x,i}f=D_{x,i}f(1)$$

2. *Let $\{e_1,\ldots,e_n\}$ be the canonical basis on $\mathbb{K}^n$ defined by $(e_i)_j=\delta_{i,j}$ [see theorem: 11.159] then*

$$\forall i\in\{1,\ldots,n\} \text{we have } \partial_{x,i}f=D_xf(e_i)$$

and

$$\forall h\in\mathbb{K}^n \text{ we have } D_xf(h)=\sum_{i=1}^n h_i\cdot\partial_{x,i}f$$

If we use matrix notation where

$$h=\left(\begin{array}{c}h_1\\ \vdots\\ h_n\end{array}\right)\in\mathcal{M}_{n,1} \text{ is a column vector}$$

and

$$\partial_xf(x)=\left(\begin{array}{ccc}\partial_{x,1}f & \ldots & \partial_{x,n}f\end{array}\right) \text{ is a row vector}$$

then it follows that we can write this using matrix notation as

$$D_xf(h)=\partial_xf\cdot h$$

Proof.

1. As f is Fréchet differentiable at x it follows from [theorem: 16.47] that $\forall i\in\{1,\ldots,n\}$ is i-partial differentiable at x. Using the previous theorem [theorem: 16.54] we conclude that f has a i-partial derivative and

$$D_{x,i}f(h)=h\cdot\partial_{x,i}f$$

and

$$\partial_{x,i}f=D_{x,i}f(1)$$

2. We have for $i,j\in\{1,\ldots,n\}$ that

$$(I^{[i,0]}(1))_j=\left(0,\ldots,0,\underbrace{1}_{i},0,\ldots,0\right)_j=\delta_{i,j}=(e_i)_j$$

or

$$I^{[i,0]}(1) = e_i \tag{16.25}$$

so that

$$\begin{aligned}\partial_{x,i} f &\underset{[\text{theorem: } 16.54]}{=} D_{x,i} f(1)\\ &\underset{[\text{theorem: } 16.47]}{=} (D_x f \circ I^{[i,0]})(1)\\ &= D_x f(I^{[i,0]}(1))\\ &\underset{[\text{eq: } 16.25]}{=} D_x f(e_i)\end{aligned}$$

proving that

$$\forall i \in \{1, \ldots, n\}\ \partial_{x,i} f = D_{x,i} f(e_i)$$

Further give $h \in \mathbb{K}^n$ we have

$$\begin{aligned}D_x f(h) &\underset{[\text{theorem: } 16.47]}{=} \sum_{i=1}^{n} D_{x,i} f(h_i)\\ &\underset{(1)}{=} \sum_{i=1}^{n} h_i \cdot \partial_{x,i} f\end{aligned}$$

□

After examining the differentiation of a multi-parameter function we look now at differentiation of vector functions [functions that maps to a product of normed spaces].

Definition 16.57. *Let $n \in \mathbb{N} \backslash \{1\}$, X a set, $\{X_i\}_{i \in \{1, \ldots, n\}}$ a finite family of sets and $\{f_i\}_{i \in \{1, \ldots, n\}}$ a family such that $f_i \in X_i^X$ then we define the following function*

$$(f_1, \ldots, f_n)_*: X \rightarrow \prod_{i \in \{1, \ldots, n\}} X_i \text{ by } (f_1, \ldots, f_n)_*(x) = (f_1(x), \ldots, f_n(x))$$

Note 16.58. We use the notation $(f_1, \ldots, f_n)_*$ so that we don't confuse the above function with the tuple $(f_1, \ldots, f_n)$ which is a element of $\prod_{i \in \{1, \ldots, n\}} X_i^X$ and not a function from X to $\prod_{i \in \{1, \ldots, n\}} X_i$.

Theorem 16.59. *Let $n \in \mathbb{N} \backslash \{1\}$, $\langle X, \|\|\rangle$ a normed space, $\{\langle X_i, \|\|_i\rangle\}_{i \in \{1, \ldots, n\}}$ a finite family of normed spaces, $\langle \prod_{i \in \{1, \ldots, n\}} X_i, \|\|_{\max}\rangle$ the product space equipped with the maximum norm [see theorem: 14.94] and $\{L_i\}_{i \in \{1, \ldots, n\}}$ a family such that $\forall i \in \{1,\ldots,n\}\ L_i \in L(X, X_i)$ then*

$$(L_1, \ldots, L_n)_* \in L\left(X, \prod_{i \in \{1, \ldots, n\}} X_i\right)$$

and

$$\|(L_1, \ldots, L_n)_*\|_{L(X, \prod_{i \in \{1, \ldots, n\}} X_i)} \leqslant \max\left(\{\|L_i\|_{L(X, X_i)} | i \in \{1, \ldots, n\}\}\right)$$

Proof. Let $x, y \in X$ and $\alpha \in \mathbb{K}$ then $\forall i \in \{1, \ldots, n\}$ we have

$$\begin{aligned}((L_1, \ldots, L_n)_*(x + \alpha \cdot y))_i &= (L_1(x + \alpha \cdot y), \ldots, L_i(x + \alpha \cdot y))_i\\ &= L_i(x + \alpha \cdot y)\\ &\underset{L_i \in L(X, X_i)}{=} L_i(x) + \alpha \cdot L_i(y)\\ &= (L_1(x), \ldots, l_n(x))_i + \alpha \cdot (L_1(y), \ldots, L_n(y))_i\\ &= ((L_1, \ldots, L_n)_*(x))_i + (\alpha \cdot (L_1, \ldots, L_n)_*(y))_i\\ &= ((L_1, \ldots, L_n)_*(x) + \alpha \cdot (L_1, \ldots, L_n)_*(y))_i\end{aligned}$$

proving that $(L_1, \ldots, L_n)_*(x + \alpha \cdot y) = (L_1, \ldots, L_n)_*(x) + \alpha \cdot (L_1, \ldots, L_n)_*(y)$. Hence

$$(L_1, \ldots, L_n)_* \in \operatorname{Hom}\left(X, \prod_{i \in \{1, \ldots, n\}} X_i\right)$$

Further if $x \in X$ then we have

$$\begin{aligned} \|(L_1,\ldots,L_n)_*(x)\|_{\max} &= \|(L_1(x),\ldots,L_n(x))\|_{\max} \\ &= \max(\{\|L_i(x)\|_i | i \in \{1,\ldots,n\}\}) \\ &\leqslant \max(\{\|L_i\|_{L(X,X_i)} \cdot \|x\| | i \in \{1,\ldots,n\}\}) \\ &\leqslant_{[\text{theorem: } 10.27]} \max(\{\|L_i\|_{L(X,X_i)} | i \in \{1,\ldots,n\}\}) \cdot \|x\| \end{aligned}$$

proving by [theorem: 14.174] that

$$(L_1,\ldots,L_n)_\star \in L\left(X, \prod_{i \in \{1,\ldots,n\}} X_i\right)$$

and by [theorem: 14.181]

$$\|(L_1,\ldots,L_n)_*\|_{L(X,\prod_{i\in\{1,\ldots,n\}} X_i)} \leqslant \max(\{\|L_i\|_{L(X,X_i)} | i \in \{1,\ldots,n\}\})$$

□

Theorem 16.60. *Let $n \in \mathbb{N} \setminus \{1\}$, X a set and $\{X_i\}_{i \in \{1,\ldots,n\}}$ a family of sets and*

$$f: X \to \prod_{i \in \{1,\ldots,n\}} X_i \text{ a function}$$

then

$$f = (\pi_1 \circ f, \ldots, \pi_n \circ f)_*$$

Proof. Let $x \in X$ then we have $\forall i \in \{1,\ldots,n\}$ that

$$(f(x))_i = \pi_i(f(x)) = (\pi_i \circ f)(x)$$

so that

$$f(x) = ((\pi_i \circ f)(x), \ldots, (\pi_n \circ f)(x)) = (\pi_1 \circ f, \ldots, \pi_n \circ f)_*(x)$$

proving that

$$f = (\pi_1 \circ f, \ldots, \pi_n \circ f)_*$$

□

Theorem 16.61. *Let $n \in \mathbb{N}_0 \setminus \{1\}$, $\langle X, \|\|\rangle$ be a normed space, $\{\langle X_i, \|\|_i\rangle\}_{i \in \{1,\ldots,n\}}$ a family of normed spaces, $\langle \prod_{i \in \{1,\ldots,n\}} X_i, \|\|_{\max}\rangle$ the normed product space equipped with the maximum norm [see theorem: 14.94], $U \subseteq X$ a open set in X, $x \in U$ and*

$$f: U \to \prod_{i \in \{1,\ldots,n\}} X_i \text{ a function}$$

then we have

$$f \text{ is Fréchet differentiable at } x \Leftrightarrow \forall i \in \{1,\ldots,n\}\ \pi_i \circ f \text{ is Fréchet differentiable at } x$$

Further if f is Fréchet differentiable at x then

$$\forall i \in \{1,\ldots,n\}\ \pi_i \circ D_x f = D_x(\pi_i \circ f)$$

or in other words

$$D_x f = (D_x(\pi_1 \circ f), \ldots, D_x(\pi_n \circ f))_*$$

Proof.

⇒. Let $\varepsilon \in \mathbb{R}^+$. As f is differentiable at x there exist a $\delta \in \mathbb{R}^+$ such that $\forall h \in U_x$ with $\|h\| < \delta$ we have

$$\|f(x+h) - f(x) - D_x f(h)\|_{\max} \leqslant \varepsilon \cdot \|h\| \tag{16.26}$$

Let $i \in \{1,\ldots,n\}$ then we have $\forall h \in U_x$ with $\|h\| < \delta$

$$\begin{aligned} \|(\pi_i \circ f)(x+h) - (\pi_i \circ f)(x) - (\pi_i \circ D_x f)(h)\|_i &= \\ \|(\pi_i(f(x+h))) - \pi_i(f(x)) - \pi_i(D_x f(h))\|_i &= \\ \|\pi_i(f(x+h) - f(x) - D_x f(h))\|_i &\leqslant \\ \max(\{\|\pi_i(f(x+h) - f(x) - D_x f(h))\|_i | i \in \{1,\ldots,n\}\}) &\underset{\text{def}}{=} \\ \|f(x+h) - f(x) - D_x f(h)\|_{\max} &\leqslant_{[\text{eq: } 16.26]} \\ \varepsilon \cdot \|h\| & \end{aligned}$$

proving that

$$\pi_i \circ f \text{ is differentiable at } x \text{ with } D_x(\pi_i \circ f) = \pi_i \circ D_x f$$

⇐. Let $\varepsilon \in \mathbb{R}^+$. Let $i \in \{1, \ldots, n\}$ then as $\pi_i \circ f$ is differentiable at x there exist a $\delta_i \in \mathbb{R}^+$ so that $\forall h \in U_x$ with $\|h\| < \delta_i$ we have $\|(\pi_i \circ f)(x+h) - (\pi_i \circ f)(x) - D_x(\pi_i \circ f)(h)\|_i \leqslant \varepsilon \cdot \|h\|$ so that

$$\max(\{\|(\pi_i \circ f)(x+h) - (\pi_i \circ f)(x) - D_x(\pi_i \circ f)(h)\|_i | i \in \{1, \ldots, n\}\}) \leqslant \varepsilon \cdot \|h\| \tag{16.27}$$

If $h \in U_x$ satisfies $\|h\| < \max(\{\delta_i | i \in \{1, \ldots, n\}\})$ then we have

$$\begin{aligned}
\|f(x+h) - f(x) - (D_x(\pi_1 \circ f), \ldots, D_x(\pi_n \circ f))_*(h)\|_{\max} &= \\
\max(\{\pi_i(f(x+h) - f(x) - (D_x(\pi_1 \circ f), \ldots, D_x(\pi_n \circ f))_*(h)) | i \in \{1, \ldots, n\}\}) &= \\
\max(\{\pi_i(f(x+h) - f(x) - (D_x(\pi_1 \circ f)(h), \ldots, D_x(\pi_n \circ f)(h))) | i \in \{1, \ldots, n\}\}) &= \\
\max(\{\pi_i(f(x+h)) - \pi_i(f(x)) - \pi_i((D_x(\pi_1 \circ f)(h), \ldots, D_x(\pi_n \circ f)(h))) | i \in \{1, \ldots, n\}\}) &= \\
\max(\{\pi_i(f(x+h)) - \pi_i(f(x)) - D_x(\pi_i \circ f)(h) i \in \{1, \ldots, n\}\}) &= \\
\max(\{(\pi_i \circ f)(x+h) - (\pi_i \circ f)(x) - D_x(\pi_i \circ f)(h) i \in \{1, \ldots, n\}\}) &\leqslant \\
\varepsilon \cdot \|h\|
\end{aligned}$$

proving, as by [theorem: 16.59] $(D_x(\pi_1 \circ f), \ldots, D_x(\pi_n \circ f))_* \in L\left(X, \prod_{i \in \{1, \ldots, n\}} X_i\right)$, that

$$f \text{ is Fréchet differentiable at } x \text{ with } D_x f = (D_x(\pi_1 \circ f), \ldots, D_x(\pi_n \circ f))_*$$

□

Corollary 16.62. *Let $n \in \mathbb{N}_0 \setminus \{1\}$, $\langle X, \|\|_X \rangle$ be a normed space, $\{\langle Y_i, \|\|_i \rangle\}_{i \in \{1, \ldots, n\}}$ a family of normed spaces, $\left\langle \prod_{i \in \{1, \ldots, n\}} Y_i, \|\|_{\max} \right\rangle$ the normed spaces equipped with the maximum norm [see theorem: 14.94], $\langle Z, \|\|_Z \rangle$ a normed space, $U \subseteq X$ a open set in X, $x \in U$, $V \subseteq \prod_{i \in \{1, \ldots, n\}} Y_i$ a open set in $\prod_{i \in \{1, \ldots, n\}} Y_i$ and*

$$f: U \rightarrow \prod_{i \in \{1, \ldots, n\}} Y_i \text{ and } g: V \rightarrow Z$$

functions such that

$$f(U) \subseteq V, f \text{ Fréchet differentiable at } x \text{ and } g \text{ Fréchet differentiable at } f(x)$$

[so that by the Chain Rule [see theorem: 16.36] $g \circ f$ is differentiable at x] then we have

$$D_x(g \circ f) = \sum_{i=1}^{n} (D_{f(x),i} g \circ D_x(\pi_i \circ f))$$

Proof. Let $h \in X$ then

$$\begin{aligned}
D_x(g \circ f)(h) &\underset{[\text{theorem: } 16.36]}{=} (D_{f(x)}\, g \circ D_x\, f)(h) \\
&= D_{f(x)}\, g(D_x f(h)) \\
&\underset{[\text{theorem: } 16.47]}{=} \sum_{i=1}^{n} D_{f(x),i} g((D_x\, f(h))_i) \\
&= \sum_{i=1}^{n} D_{f(x),i} g(\pi_i(D_x f(h))) \\
&= \sum_{i=1}^{n} D_{f(x),i} g((\pi_i \circ D_x f)(h)) \\
&\underset{[\text{theorem: } 16.61]}{=} \sum_{i=1}^{n} D_{f(x),i} g((D_x(\pi_i \circ f))(h)) \\
&= \sum_{i=1}^{n} (D_{f(x),i} g \circ D_x(\pi_i \circ f))(h) \\
&= \left(\sum_{i=1}^{n} (D_{f(x),i} g \circ D_x(\pi_i \circ f))\right)(h)
\end{aligned}$$

proving that

$$D_x(g \circ f) = \sum_{i=1}^{n} (D_{f(x),i} g \circ D_x(\pi_i \circ f))$$

□

We can now look at the differentiability of functions between $\mathbb{K}^n$ and $\mathbb{K}^m$.

Theorem 16.63. (Jacobian matrix) *Let $n,m\in\mathbb{N}\setminus\{1\}$, $\langle\mathbb{K}^n,\|\|_n\rangle$, $\langle\mathbb{K}^m,\|\|_m\rangle$ two normed spaces [based on the maximum norms) over the field $\mathbb{K}$, $U\subseteq\mathbb{K}^n$ a open set in $\mathbb{K}^n$, $x\in U$ and*

$$f:U\rightarrow\mathbb{K}^m \text{ a function that is Fréchet differentiable at } x$$

then $\forall i\in\{1,\ldots,n\}$, $\forall j\in\{1,\ldots,m\}$ we have that $\pi_j\circ f:U\rightarrow\mathbb{K}$ has a i-partial derivative and

$$\forall h\in\mathbb{K}^n \text{ we have } (D_xf(h))_j \underset{\text{def}}{=} \pi_j(D_xf(h))=\sum_{i=1}^n \partial_{x,i}(\pi_j\circ f)\cdot h_i$$

If we define the following matrices

$$h=\begin{pmatrix} h_1\\ \vdots\\ h_n\end{pmatrix}\in\mathcal{M}_{n,1}(\mathbb{K})$$

and

$$D_xf(h)=\begin{pmatrix} (D_xf(h))_1\\ \vdots\\ (D_xf(h))_m\end{pmatrix}\in\mathcal{M}_{m,1}(\mathbb{K})$$

and

$$\left[\begin{array}{c}\frac{\partial(f_1,\ldots,f_m)}{\partial(1,\ldots,n)}\end{array}\right]_x=\begin{pmatrix} \partial_{x,1}(\pi_1\circ f) & \ldots & \partial_{x,n}(\pi_1\circ f)\\ \vdots & \ddots & \vdots\\ \partial_{x,1}(\pi_m\circ f) & \ldots & \partial_{x,n}(\pi_m\circ f)\end{pmatrix}\in\mathcal{M}_{m,n}(\mathbb{K})$$

or

$$\forall i\in\{1,\ldots,m\},\forall j\in\{1,\ldots,n\} \text{ we have } \left(\left[\begin{array}{c}\frac{\partial(f_1,\ldots,f_m)}{\partial(1,\ldots,n)}\end{array}\right]_x\right)_{i,j}=\partial_{x,j}(\pi_i\circ f)$$

then we can express the Fréchet differential of f at x as the following matrix product

$$D_xf(h)=\left[\begin{array}{c}\frac{\partial(f_1,\ldots,f_m)}{\partial(1,\ldots,n)}\end{array}\right]_x\cdot h$$

The matrix $\left[\begin{array}{c}\frac{\partial(f_1,\ldots,f_m)}{\partial(1,\ldots,n)}\end{array}\right]_x$ is called the ***Jacobian matrix of f at x****. Further if $E=\{e_1,\ldots,e_n\}$ defined by $(e_i)_j=\delta_{i,j}$ is the canonical basis of $\mathbb{K}^n$ and $F=\{f_1,\ldots,f_m\}$ defined by $(f_i)_j=\delta_{i,j}$ is the canonical basis of $\mathbb{K}^m$ then*

$$\left[\begin{array}{c}\frac{\partial(f_1,\ldots,f_m)}{\partial(1,\ldots,n)}\end{array}\right]_x=\mathcal{M}(D_xf;E,F) \text{ [see definition: 11.321]}$$

Proof. Let $i\in\{1,\ldots,n\}$ and $j\in\{1,\ldots,m\}$. As f is Fréchet differentiable at x it follows from [theorem: 16.61] that $\pi_j\circ f:U\rightarrow\mathbb{K}$ is Fréchet differentiable at x with

$$D_x(\pi_j\circ f)=\pi_j\circ D_xf \tag{16.28}$$

Hence using [theorem: 16.56] $\pi_j\circ f$ has a partial derivative at x with

$$D_x(\pi_j\circ f)(h)=\sum_{i=1}^n h_i\cdot\partial_{x,i}(\pi_j\circ f) \tag{16.29}$$

Further for $h\in\mathbb{K}^n$ we have

$$\begin{array}{rcl} (D_xf(h))_j & = & \pi_j(D_xf(h))\\ & = & (\pi_j\circ D_xf)(h)\\ & \underset{\text{[eq: 16.28]}}{=} & D_x(\pi_j\circ f)(h)\\ & \underset{\text{[eq: 16.29]}}{=} & \sum_{i=1}^n h_i\cdot\partial_{x,i}(\pi_j\circ f)\\ & \underset{\partial_{x,i}(\pi_j\circ f)\in\mathbb{K}}{=} & \sum_{i=1}^n \partial_{x,i}(\pi_j\circ f)\cdot h_i \end{array}$$

proving that

$$(D_xf(h))_j=\sum_{i=1}^n \partial_{x,i}(\pi_j\circ f)\cdot h_i \tag{16.30}$$

Next for the matrix notation we have for $j \in \{1, \ldots, m\}$ that

$$\begin{aligned}\left(\left[\frac{\partial(f_1,\ldots,f_m)}{\partial(1,\ldots,n)}\right]_x \cdot h\right)_j &= \begin{pmatrix}\partial_{x,1}(\pi_1\circ f) & \ldots & \partial_{x,n}(\pi_1\circ f)\\ \vdots & \ddots & \vdots\\ \partial_{x,1}(\pi_m\circ f) & \ldots & \partial_{x,n}(\pi_m\circ f)\end{pmatrix}\cdot\begin{pmatrix}h_1\\ \vdots\\ h_n\end{pmatrix}\\ &= \sum_{i=1}^n \left(\left[\frac{\partial(f_1,\ldots,f_n)}{\partial(1,\ldots,n)}\right]_x\right)_{i,j}\cdot h_i\\ &= \sum_{i=1}^n \partial_{x,i}(\pi_j\circ f)\cdot h_i\\ &\underset{[\text{eq: } 16.30]}{=} (D_x f(h))_j\\ &= \begin{pmatrix}(D_x f(h))_1\\ \vdots\\ (D_x f(h))_m\end{pmatrix}_j\end{aligned}$$

so that

$$\left[\frac{\partial(f_1,\ldots,f_n)}{\partial(1,\ldots,n)}\right]_x \cdot h = D_x f(h)$$

Finally if $i \in \{1, \ldots, n\}$, $j \in \{1, \ldots, m\}$ that we have by [definition: 11.321] that for $i \in \{1, \ldots, n\}$ and

$$\begin{aligned}(D_x f(e_i))_j &= \left(\sum_{k=1}^m \mathcal{M}(D_x f; E, F)_{k,i}\cdot f_k\right)_j\\ &= \sum_{k=1}^m \mathcal{M}(D_x f; E, F)_{k,i}\cdot (f_k)_j\\ &= \sum_{k=1}^m \mathcal{M}(D_x f; E, F)_{k,i}\cdot \delta_{k,j}\\ &= \mathcal{M}(D_x f; E, F)_{j,i}\end{aligned}$$

Further we have

$$\begin{aligned}(D_x f(e_i))_j &\underset{[\text{eq: } 16.30]}{=} \sum_{k=1}^n \partial_{x,k}(\pi_j\circ f)\cdot (e_i)_k\\ &= \sum_{k=1}^n \partial_{x,k}(\pi_j\circ f)\cdot \delta_{i,k}\\ &= \partial_{x,i}(\pi_j\circ f)\\ &= \left(\left[\frac{\partial(f_1,\ldots,f_m)}{\partial(1,\ldots,n)}\right]_x\right)_{j,i}\end{aligned}$$

so that

$$\mathcal{M}(D_x f; E, F)_{j,i} = \left(\left[\frac{\partial(f_1,\ldots,f_m)}{\partial(1,\ldots,n)}\right]_x\right)_{j,i}$$

or

$$\mathcal{M}(D_x f; E, F) = \left[\frac{\partial(f_1,\ldots,f_m)}{\partial(1,\ldots,n)}\right]_x$$

□

Lemma 16.64. *Let $\langle X_1, \|\|_1\rangle$, $\langle X_2, \|\|_2\rangle$ and $\langle Y, \|\|_Y\rangle$ be normed spaces and $\langle X_1\cdot X_2, \|\|_{\max}\rangle$ the normed space of the product $X_1\cdot X_2$ with the maximum norm $\|\|_{\max}$, $L \in L(X_1, X_2; Y)$ [a bi-linear continuous mapping], $x = (x_1, x_2) \in X_1\cdot X_2$ then L is differentiable at x and*

$$D_x L = D_{(x_1,x_2)} L = L(x_1, *) + L(*, x_2)$$

where

$$L(x_1, *) : X_1\cdot X_2 \rightarrow Y \text{ is defined by } L(x_1, \star)(h) = L(x_1, *)(h_1, h_2) = L(x_1, h_2)$$

and

$$L(*, x_2) : X_1\cdot X_2 \rightarrow Y \text{ is defined by } L(\star, x_2)(h) = L(*, x_2)(h_1, h_2) = L(h_1, x_2)$$

Proof. First we prove that $L(x_1,*)+L(*,x_2)$ is linear. So let $(h_1,h_2),(g_1,g_2)\in X_1\cdot X_2$ and $\alpha\in\mathbb{K}$ then we have

$$\begin{array}{rl}(L(x_1,*)+L(*,x_2))((h_1,h_2)+\alpha\cdot(g_1,g_2)) & = \\ (L(x_1,*)+L(*,x_2))(h_1+\alpha\cdot g_1,h_2+\alpha\cdot g_2) & = \\ L(x_1,h_2+\alpha\cdot g_2)+L(h_1+\alpha\cdot g_1,x_2) & = \\ L(x_1,h_2)+\alpha\cdot L(x_1,g_2)+L(h_1,x_2)+\alpha\cdot L(g_1,x_2) & = \\ L(x_1,h_2)+L(h_1,x_2)+\alpha\cdot(L(x_1,g_2)+L(g_1,x_2)) & = \\ (L(x_1,*)+L(*,x_2))(h_1,h_2)+\alpha\cdot((L(x_1,*)+L(*,x_2))(g_1,g_2)) & \end{array}$$

Hence we have that

$$L(x_1,*)+L(*,x_2)\in\mathrm{Hom}(X_1,X_2;Y) \tag{16.31}$$

Next we prove that $L(x_1,*)+L(*,x_2)$ is continuous. So let $(h_1,h_2)\in X_1\cdot X_2$ with

$$\|h_1\|_1=1=\|h_2\|_2 \tag{16.32}$$

then we have

$$\begin{array}{rl}\|(L(x_1,*)+L(*,x_2))(h_1,h_2)\|_Y & = \\ \|L(x_1,h_2)+L(h_1,x_2)\|_Y & \leqslant \\ \|L(x_1,h_2)\|_Y+\|L(h_1,x_2)\|_Y & \leqslant_{[\text{theorem: }14.193]} \\ \|L\|_{L(X_1,X_2;Y)}\cdot\|x_1\|_1\cdot\|h_2\|_2+\|L\|_{L(X_1,X_2;Y)}\cdot\|x_2\|_2\cdot\|h_2\|_1 & =_{[\text{eq: }16.32]} \\ \|L\|_{L(X_1,X_2;Y)}\cdot\|x_1\|_1+\|L\|_{L(X_1,x_2;Y)}\cdot\|x_2\|_2 & \end{array}$$

proving by [theorem: 14.187] that $L(x_1,*)+L(*,x_2)$ is continuous, combining this with [eq: 16.31] results in

$$L(x_1,*)+L(*,x_2)\in L(X_1,X_2;Y) \tag{16.33}$$

Finally Let $\varepsilon\in\mathbb{R}^+$ and take $\delta=\frac{\varepsilon}{\|L\|_{L(X_1,X_2;Y)}+1}\in\mathbb{R}^+$ then, if $(h_1,h_2)\in X_1\cdot X_2=(X_1\cdot X_2)_x$ and $\|(h_1,h_2)\|_{\max}<\delta$, we have

$$\|h_1\|_1,\|h_2\|_2\leqslant\max(\{\|h_i\|_i|i\in\{1,2\}\})=\|(h_1,h_2)\|_{\max}<\delta \tag{16.34}$$

and

$$\begin{array}{rl}\|L((x_1,x_2)+(h_1,h_2))-L(x_1,x_2)-(L(x_1,*)+L(*,x_2))(h_1,h_2)\|_Y & = \\ \|L(x_1+h_1,x_2+h_2)-L(x_1,x_2)-L(x_1,h_2)-L(h_1,x_2)\|_Y & = \\ \|L(x_1,x_2+h_2)+L(h_1,x_2+h_2)-L(x_1,x_2)-L(x_1,h_2)-L(h_1,x_2)\|_Y & = \\ \left\|\underbrace{L(x_1,x_2)}_{1}+\underbrace{L(x_1,h_2)}_{2}+\underbrace{L(h_1,x_2)}_{3}+L(h_1,h_2)-\underbrace{L(x_1,x_2)}_{1}-\underbrace{L(x_1,h_2)}_{2}-\underbrace{L(h_1,x_2)}_{3}\right\|_Y & = \\ \|L(h_1,h_2)\|_Y & \leqslant_{[\text{theorem: }14.193]} \\ \|L\|_{L(X_1,X_2;Y)}\cdot\|h_1\|_1\cdot\|h_2\|_2 & <_{[\text{eq: }16.34]} \\ \|L\|_{L(X_1,X_2;Y)}\cdot\dfrac{\varepsilon}{\|L\|_{L(X_1,x_2;Y)}+1}\cdot\|h_2\|_2 & < \\ \varepsilon\cdot\|h_2\|_2 & \leqslant_{[\text{eq: }16.34]} \\ \varepsilon\cdot\|(h_1,h_2)\|_{\max} & \end{array}$$

proving that

$$L\text{ is Fréchet differentiable at }x\text{ with }D_xL=L(x_1,*)+L(*,x_2)$$ □

We use the above lemma to prove that the product of differentiable functions is differentiable.

Corollary 16.65. *Let $\langle X,\|\|_X\rangle$ be a normed space, $\langle\mathbb{K},||\rangle$, $U\subseteq\mathbb{K}$ open with $x\in U$ and*

$$f:U\rightarrow\mathbb{K}\textit{ and }g:U\rightarrow\mathbb{K}\textit{ functions that are Fréchet differentiable at }x$$

then

$$g\cdot f:U\rightarrow\mathbb{K}\textit{ defined by }(g\cdot f)(x)=g(x)\cdot f(x)$$

is Fréchet differentiable at x with

$$D_x(g\cdot f)=f(x)\cdot D_xg+g(x)\cdot D_xf$$

where $f(x)\cdot D_xg+g(x)\cdot D_xf$ is defined by

$$(f(x)\cdot D_xg+g(x)\cdot D_xf)(k)=f(x)\cdot D_x g(k)+g(x)\cdot D_xg(k)$$

Proof. Define the function

$$L:\mathbb{K}^2\rightarrow\mathbb{K}\text{ by }L(x,y)=x\cdot y$$

Let $x,y,z,\alpha\in\mathbb{K}$ then we have

$$\begin{array}{rcl}L(x,y+\alpha\cdot z) & = & x\cdot(y+\alpha\cdot z)=x\cdot y+\alpha\cdot(y\cdot z)=L(x,y)+\alpha\cdot L(y,z)\\ L(x+\alpha\cdot y,z) & = & (x+\alpha\cdot y)\cdot z=x\cdot z+\alpha\cdot(y\cdot z)=L(x,z)+\alpha\cdot(L(y,z))\end{array}$$

proving that L is multilinear or

$$L\in\operatorname{Hom}\left(\underbrace{\mathbb{K},\ldots,\mathbb{K}}_{2};Y\right)\tag{16.35}$$

Further if $h=(h_1,h_2)\in\mathbb{K}^2$ with $|h_1|=|h_2|=1$ then we have

$$|L(h)|=|L(h_1,h_2)|=|h_1\cdot h_2|=|h_1|\cdot|h_2|=1$$

proving by [theorem: 14.187] that L is continuous, hence combined with [eq: 16.35] it follows that $L\in L\left(\underbrace{\mathbb{K},\ldots,\mathbb{K}}_{2};Y\right)$. Using the previous lemma [lemma: 16.64] it follows that

$$\forall y=(y_1,y_2)\in\mathbb{K}^2\ L\text{ is differentiable at }y\text{ with }D_{(y_1,y_2)}L=L(y_1,*)+L(*,y_2)\tag{16.36}$$

Define

$$h:U\rightarrow\mathbb{K}^2\text{ by }h(x)=(g(x),f(x))$$

then $\pi_1\circ h=g$ and $\pi_2\circ h=g$. As f, g are Fréchet differentiable at x it follows that $\pi_1\circ h$, $\pi_2\circ h$ are Fréchet differentiable at x with $D_x(\pi_1\circ h)=D_xg$ and $D_x(\pi_2\circ h)=D_xf$. So we can use [theorem: 16.61] giving

$$h\text{ is Fréchet differentiable at }x\text{ with }\pi_1\circ D_xh=D_xg\text{ and }\pi_2\circ D_xh=D_xf\wedge D_xh\tag{16.37}$$

Now $\forall y\in U$ we have

$$(L\circ h)(x)=(L(h(x)))=L(g(x),f(x))=g(x)\cdot f(x)=(g\cdot f)(x)$$

proving that

$$L\circ h=g\cdot f$$

Using the Chain Rule [see theorem: 16.36] on [eqs: 16.36 and 16.37] we have that $L\circ h$ is Fréchet differentiable at x so that

$$g\cdot f\text{ is Fréchet differentiable at }x$$

and $\forall k\in X$ we have

$$\begin{array}{rcl}D_x(g\cdot f)(k) & = & D_x(L\circ h)(k)\\ & = & (D_{h(x)}L\circ D_xh)(k)\\ & = & D_{h(x)}L(D_xh(k))\\ & = & D_{(g(x),f(x))}L(D_xh(k))\\ & \underset{[\text{theorem: }16.61]}{=} & D_{(g(x),f(x))}L((D_x(\pi_1\circ h),D_x(\pi_2\circ h))_*(k))\\ & \underset{[\text{eq: }16.37]}{=} & D_{(g(x),f(x))}L((D_xg,D_xf)_*(k))\\ & = & D_{(g(x),f(x))}L(D_xg(k),D_xf(k))\\ & \underset{[\text{eq: }16.36]}{=} & (L(g(x),*)+L(*,f(x)))(D_xg(k),D_xf(k))\\ & = & L(g(x),D_xf(k))+L(D_xg(k),f(x))\\ & = & g(x)\cdot D_xf(k)+(D_xg(k))\cdot f(x)\\ & = & g(x)\cdot D_xf(k)+f(x)\cdot D_xg(k)\\ & = & (g(x)\cdot D_xf)(k)+(f(x)\cdot D_xg)(k)\\ & = & (g(x)\cdot D_xf+f(x)\cdot D_xg)(k)\end{array}$$

proving that

$$D_x(g\cdot f)=g(x)\cdot D_x f+f(x)\cdot D_x g$$ □

If we apply the above for $X=\mathbb{K}$ we retrieve the calculus product rule of derivatives.

Corollary 16.66. *Let $\langle\mathbb{K},\|\rangle$ be the normed space of real (complex) numbers, $U\subseteq\mathbb{K}$ a open set, $x\in U$ and $f:U\to\mathbb{K}$ and $g:U\to\mathbb{K}$ are functions that have a derivative at x then $g\cdot f$ has a derivative at x and*

$$(g\cdot f)'_x=g(x)\cdot f'_x+f(x)\cdot g'_x$$

Proof. Using [theorem: 16.32] we have that f,g are Fréchet differentiable at x and

$$f'_x=D_x f(1) \text{ and } g'_x=D_x g(1)$$

By the previous corollary [corollary: 16.65] it follows that

$$g\cdot f \text{ is Fréchet differentiable at } x \text{ and } D_x(g\cdot f)=g(x)\cdot D_x f+f(x)\cdot D_x g \tag{16.38}$$

We can use [theorem: 16.32] again, giving that

$$g\cdot f \text{ has a derivative at } x$$

and

$$\begin{array}{rcl}(g\cdot f)'_x & = & D_x(g\cdot f)(1)\\ & \underset{[\text{eq: } 16.38]}{=} & (g(x)\cdot D_x f+f(x)\cdot D_x g)(1)\\ & = & g(x)\cdot D_x f(1)+f(x)\cdot D_x g(1)\\ & = & g(x)\cdot f'_x+f(x)\cdot g'_x\end{array}$$

□

Theorem 16.67. *Let $\langle\mathbb{K},\|\rangle$ be the normed space of real (complex) numbers, $x\in\mathbb{K}\setminus\{0\}$ [a open set] we have that*

$$\left(\frac{1}{*}\right):\mathbb{K}\setminus\{0\}\to\mathbb{K} \text{ defined by } \left(\frac{1}{*}\right)(x)=\frac{1}{x}$$

has a derivative at x with

$$\left(\frac{1}{*}\right)'_x=-\frac{1}{x^2}$$

Proof. Let $\varepsilon\in\mathbb{R}^+$. Then we have for $h\in(\mathbb{K}\setminus\{0\})_x$ with

$$\begin{array}{rcl}\frac{\frac{1}{x+h}-\frac{1}{x}}{h}-\left(-\frac{1}{x^2}\right) & = & \frac{\frac{1}{x+h}-\frac{1}{x}}{h}+\frac{1}{x^2}\\ & = & \frac{\frac{x-x-h}{(x+h)\cdot x}}{h}+\frac{1}{x^2}\\ & = & \frac{\frac{-h}{(x+h)\cdot x}}{h}+\frac{1}{x^2}\\ & = & \frac{-1}{(x+h)\cdot x}+\frac{1}{x^2}\\ & = & \frac{-x+(x+h)}{(x+h)\cdot x^2}\\ & = & \frac{h}{x^2\cdot(x+h)}\end{array} \tag{16.39}$$

Take now $\delta=\min\left(\frac{\varepsilon}{2}\cdot|x|^3,\frac{|x|}{2}\right)\in\mathbb{R}^+$ [as $x\neq 0$] then we have for $h\in(\mathbb{K}\setminus\{0\})_x$ with $0<|h|<\delta$ that

$$|h|<\frac{|x|}{2}<|x|\Rightarrow 0<|x|-|h| \text{ and } \frac{|x|}{2}=|x|-\frac{|x|}{2}<|x|-|h|\Rightarrow\frac{1}{|x|-|h|}<\frac{2}{|x|} \tag{16.40}$$

so that

$$|x| = |x+h-h| \leqslant |x+h| + |h| \Rightarrow |x| - |h| < |x+h| \underset{\text{[eq: 16.40]}}{\Rightarrow} \frac{1}{|x+h|} < \frac{1}{|x|-|h|} < \frac{2}{|x|} \tag{16.41}$$

and

$$|h| < \frac{\varepsilon}{2} \cdot |x|^3 \tag{16.42}$$

Hence we have

$$\begin{array}{rcl} \left| \frac{\frac{1}{x+h} - \frac{1}{x}}{h} - \left(-\frac{1}{x^2}\right) \right| & \underset{\text{[eq: 16.39]}}{=} & \left| \frac{h}{x^2 \cdot (x+h)} \right| \\ & = & \frac{|h|}{|x|^2 \cdot |x+h|} \\ & <_{\text{[eq: 16.41]}} & \frac{|h| \cdot 2}{|x|^2 \cdot |x|} \\ & = & \frac{|h| \cdot 2}{|x|^3} \\ & <_{\text{[eq: 16.42]}} & \frac{\varepsilon \cdot |x|^3}{2 \cdot |x|^3} \\ & = & \varepsilon \end{array}$$

which proves that

$$\left(\frac{1}{*}\right) \text{ has a derivative at } x \text{ and } \left(\frac{1}{*}\right)'_x = -\frac{1}{x^2}$$

□

Theorem 16.68. *We have*

1. *If $n \in \mathbb{N}$ then $\forall x \in \mathbb{K}$*

$$(\cdot)^n : \mathbb{K} \rightarrow \mathbb{K} \textit{ defined by } (\cdot)^n(x) = x^n$$

   has a derivative at x and

$$((\cdot)^n)'_x = n \cdot (\cdot)^{n-1}(x)$$

   or in a shorter notation

$$(x^n)'_x = n \cdot x^{n-1}$$

2. *If $z \in \{-n | n \in \mathbb{N}\}$ then $\forall x \in \mathbb{K} \setminus \{0\}$ [a open subset of $\mathbb{K}$] then*

$$(\cdot) : \mathbb{K} \setminus \{0\} \rightarrow \mathbb{K} \textit{ defined by } (\cdot)^z(x) = \frac{1}{x^{(-z)}}$$

   has a derivative at x and

$$((\cdot)^z)'_x = z \cdot (\cdot)^{z-1}$$

   or in a shorter notation

$$(x^{-n})'_x = -n \cdot x^{-(n+1)} = -n \cdot x^{-n-1}$$

Proof.

1. We use induction to prove this by induction so given $x \in \mathbb{K}$ define

$$S = \{n \in \mathbb{N} | (\cdot)^n \text{ is calculus differentiable at } x \text{ and } ((\cdot)^n)'_x = n \cdot (\cdot)^{n-1}(x)\}$$

   then we have:

   $1 \in S$. If $n = 1$ then $(\cdot)^n = (\cdot)^1 = \mathrm{Id}_{\mathbb{K}}$ so that by [theorem: 16.22] $(\cdot)^1$ is Fréchet differentiable at x with $D(\cdot)^1(x) = \mathrm{Id}_{\mathbb{K}}$. Using [theorem: 16.32] it follows that

$$(\cdot)^1 \text{ has a derivative at } x \text{ and}$$

   and

$$((\cdot)^1)'_x = D_x(\cdot)^1(1) = \mathrm{Id}_{\mathbb{K}}(1) = 1 = 1 \cdot (\cdot)^0(x) = 1 \cdot (\cdot)^{1-1}(x)$$

proving that $1 \in S$.

$\boldsymbol{n \in S \Rightarrow n+1 \in S.}$ Given $y \in X$ we have

$$(\cdot)^{n+1}(x) = x^{n+1} = x \cdot x^n = \mathrm{Id}_X(x) \cdot (\cdot)^n(x)$$

so that

$$(\cdot)^{n+1} = \mathrm{Id}_{\mathbb{K}} \cdot (\cdot)^n$$

As $n \in S$ we have that $(\cdot)^n$ has a derivative at x with

$$((\cdot)^n)'_x = n \cdot (\cdot)^{n-1}(x).$$

So $(\cdot)^{n+1}$ is the product of two functions that have a derivative at x, hence by [corollary: 16.66] we have that

$$(\cdot)^{n+1} \text{ has a derivative at } x$$

and

$$\begin{array}{rcl} ((\cdot)^{n+1})'_x &=& (\mathrm{Id}_{\mathbb{K}} \cdot (\cdot)^n)'_x \\ &=& \mathrm{Id}_{\mathbb{K}}(x) \cdot ((\cdot)^n)'_x + (\cdot)^n(x) \cdot (\mathrm{Id}_{\mathbb{K}})'_x \\ &=& x \cdot n \cdot x^{n-1} + x^n \cdot 1 \\ &=& n \cdot x^n + 1 \cdot x^n \\ &=& (n+1) \cdot x^n \\ &=& (n+1) \cdot x^{(n+1)-1} \\ &=& (n+1) \cdot (\cdot)^{(n+1)-1}(x) \end{array}$$

proving that $n+1 \in S$.

2. Let $z = -n \in \{-n | n \in \mathbb{N}\}$ then for every $y \in \mathbb{K} \setminus \{0\}$ we have

$$\left(\left(\frac{1}{*}\right) \circ (\cdot)^n\right)(y) = \left(\frac{1}{*}\right)((\cdot)^n(y)) = \left(\frac{1}{*}\right)(y^n) = \frac{1}{y^n} = \frac{1}{y^{-z}} = (\cdot)^z(y) \tag{16.43}$$

So $(\cdot)^z$ is the composition of $\left(\frac{1}{*}\right)$ and $(\cdot)^n$ where these functions have a derivative at x and

$$((\cdot)^n)'_x = n \cdot x^{n-1} \text{ and } \left(\frac{1}{\cdot}\right)'_x \underset{[\text{eq: } 16.67]}{=} -\frac{1}{x^2} \tag{16.44}$$

Hence using [theorem: 16.37] we have that $(\cdot)^z$ has a derivative at x and

$$\begin{array}{rcl} ((\cdot)^z)'_x &=& \left(\left(\frac{1}{*}\right) \circ (\cdot)^n\right)'_x \\ &\underset{[\text{theorem: } 16.37]}{=}& ((\cdot)^n)'_x \cdot \left(\frac{1}{*}\right)'_{(\cdot)^n(x)} \\ &\underset{[\text{eq: } 16.44]}{=}& n \cdot x^{n-1} \cdot \frac{-1}{((\cdot)^n(x))^2} \\ &=& n \cdot x^{n-1} \cdot -\frac{1}{x^{2 \cdot n}} \\ &=& -n \cdot x^{n-1} \cdot \frac{1}{x^{n-1} \cdot x^{2 \cdot n - (n-1)}} \\ &=& -n \cdot \frac{1}{x^{n+1}} \\ &=& z \cdot \frac{1}{x^{n+1}} \\ &=& z \cdot (\cdot)^{-(n+1)}(x) \\ &=& z \cdot (\cdot)^{z-1}(x) \end{array}$$

$\square$

16.2 Higher order differentials

16.2.1 Linear mappings to linear mappings

The idea of higher order differentiation looks simple, just differentiate the function that maps a point to it's differential, but it actually rather complex. If a function $f: U \rightarrow Y$ is Fréchet differentiable on U then $\forall x \in U$ $D_x f \in L(X, Y)$ exist so that we can define the function

$$Df: U \rightarrow L(X, Y) \text{ by } (Df)(x) = D_x f$$

and ask if that function is Fréchet differentiable at $x \in U$. If so then $D_x(Df) \in L(X, L(X, Y))$ is defined to be the second order differential of f noted as $D_x^2 f$. Proceeding in this way we can define the third order differential $D_x^3 f$at x and so on. There is however a problem, if we look at the types of the differentials we have

$$\begin{array}{rcl} D_x f & \in & L(X, Y) \\ D_x^2 f & \in & L(X, L(X, Y)) \\ D_x^3 f & \in & L(X, L(X, L(X, Y))) \\ & \dots & \end{array}$$

So we have to learn to work with sets like $L(X, L(X, Y))$, $L(X, L(X, Y)) \dots$. Evaluating the higher order differentials is laso rather elaborated because we have to use

$$\begin{array}{rcl} (Df(x))(y) & \in & Y \\ ((D^2 f(x))(y_1))(y_2) & \in & Y \\ (((D^3 f(x))(y_1))(y_2))(y_3) & \in & Y \\ & \dots & \end{array}$$

It would be nicer if the higher order differentials are multi-linear functions so that we have

$$\begin{array}{rcl} D_x f(y) & \in & Y \\ D_x^2 f(y_1, y_2) & \in & Y \\ D_x^3 f(y_1, y_2, y_3) & \in & Y \\ & \dots & \\ D_x^n f(y_1, \dots, y_n) & \in & Y \end{array}$$

To be able to do this we must do the following identifications

$$\begin{array}{rcl} L(X, L(X, Y)) & \text{is identified with} & L^2(X; Y) \\ L(X, L(X, L(X, Y))) & \text{is identified with} & L^3(X; Y) \\ & \dots & \end{array}$$

where the identification maintains linearity, continuity and the norm. This is the focus of this section. First we need some recursive definitions.

Definition 16.69. *Let $n \in \mathbb{N}$, $\langle X, \|\|_X \rangle$, $\langle Y, \|\|_Y \rangle$ normed spaces then we define $L_n(X; Y)$ recursively as follows.*

$$L_n(X; Y) = \begin{cases} L(X, Y) \text{ if } n = 1 \\ L(X, L_{n-1}(X; Y)) \text{ if } 1 < n \end{cases}$$

Example 16.70. Using the above definition we have

$$\begin{array}{rcl} L_1(X; Y) & = & L(X, Y) \\ L_2(X; Y) & = & L(X, L_1(X; Y)) \\ & = & L(X, L(X, Y)) \\ L_3(X; Y) & = & L(X, L_2(X; Y)) \\ & = & L(X, L(X, L(X, Y))) \\ \dots & \dots & \dots \\ L_n(X; Y) & = & L(X, L(X, \dots (L(X, Y)))) \end{array}$$

Lemma 16.71. *Let $n, m \in \mathbb{N}$, $\langle X, \|\|_X \rangle$, $\langle Y, \|\|_Y \rangle$ normed spaces then*

$$L_{n+m}(X;Y) = L_m(X; L_n(X;Y))$$

Proof. We prove this by induction, so define

$$S_n = \{m \in \mathbb{N} | \text{If } \langle X, \|\| \rangle \text{ is a normed space then } L_{n+m}(X;Y) = L_m(X; L_n(X;Y))\}$$

then we have:

$1 \in S_n$. As $L_{n+1}(X;Y) \underset{\text{def}}{=} L(X, L_n(X;Y)) \underset{\text{def}}{=} L_1(X; L_n(X;Y))$ it follows that $1 \in S_n$.

$m \in S_n \Rightarrow m+1 \in S_n$. We have

$$\begin{array}{rcl} L_{n+(m+1)}(X;Y) & = & L_{(n+m)+1}(X;Y) \\ & = & L(X, L_{n+m}(X;Y)) \\ & = \underset{m \in S_n}{=} & L(X, L_m(X; L_n(X;Y))) \\ & \underset{\text{def}}{=} & L_{m+1}(X, L_n(X;Y)) \end{array}$$

proving that $m+1 \in S$. □

Corollary 16.72. *Let $n \in \mathbb{N} \setminus \{1\}$, $\langle X, \|\|_X \rangle$, $\langle Y, \|\|_Y \rangle$ normed spaces then*

$$L_n(X;Y) = L_{n-1}(X; L_1(X;Y)) \underset{L_1(X;Y) = L(X,Y)}{=} L_{n-1}(X; L(X,Y))$$

Proof. We have

$$\begin{array}{rcl} L_n(X;Y) & = & L_{(n-1)+1}(X;Y) \\ & \underset{n-1 \in \mathbb{N} \wedge [\text{lemma: } 16.71]}{=} & L_{n-1}(X; L_1(X;Y)) \end{array}$$

□

We want now to show that $L_n(X;Y)$ is isometric to $L^n(X;Y)$, first we show how elements of $L_n(X;Y)$ can act on tuples of elements of X.

Definition 16.73. *Let $n \in \mathbb{N}$, $\langle X, \|\|_X \rangle$, $\langle Y, \|\|_Y \rangle$ normed spaces, $L \in L_n(X;Y)$, $m \in \{1, \ldots, n\}$ and $x = (x_1, \ldots, x_m) \in X^m$ then we define $L(x_1 : \ldots : x_m)$ recursively as*

$$L(x_1 : \ldots : x_m) = \begin{cases} L(x_1) \text{ if } m = 1 \\ (L(x_1))(x_2 : \ldots : x_m) \text{ if } 1 < m \leqslant n \end{cases}$$

***Note* 16.74.** *This is well defined as for $1 < m \leqslant n$ we have that $1 < n$ so that*

$$L \in L_n(X;Y) = L(X, L_{n-1}(X;Y)) \text{ hence } L(x_1) \in L_{n-1}(X;Y)$$

Example 16.75. Let $n = 4$ then for $L \in L_4(X;Y) = L(X, L_3(X;Y))$ we have

$$\begin{array}{rcl} L(x_1 : \ldots : x_4) & = & (L(x_1))(x_2 : \ldots x_4) \text{ where } L(x_1) \in L_3(X;Y) = L(X, L_2(X;Y)) \\ & = & ((L(x_1))(x_2))(x_3 : \ldots : x_4) \text{ where } (L(x_1))(x_2) \in L_2(X;Y) = L(X, L_1(X;Y)) \\ & = & (((L(x_1))(x_2))(x_3))(x_4 : \ldots : x_4) \text{ where } ((L(x_1))(x_2))(x_3) \in L_1(X;Y) = L(X,Y) \\ & = & (((L(x_1))(x_2))(x_3))(x_4) \in Y \\ L(x_1 : \ldots : x_3) & = & (L(x_1))(x_2 : \ldots : x_3) \text{ where } L(x_1) \in L_3(X;Y) = L(x, L_2(X;Y)) \\ & = & ((L(x_1))(x_2))(x_3 : \ldots : x_3) \text{ where } (L(x_1))(x_2) \in L_2(X;Y) = L(X, L_1(X;Y)) \\ & = & ((L(x_1))(x_2))(x_3) \in L_1(X;Y) \end{array}$$

The above example suggests the following theorem that also ensures that [definition: 16.73] is valid.

Lemma 16.76. *Let $n \in \mathbb{N}$, $\langle X, \|\|_X \rangle$, $\langle Y, \|\|_Y \rangle$ normed spaces, $L \in L_n(X;Y)$, $m \in \{1, \ldots, n\}$ and $x = (x_1, \ldots, x_m) \in X^m$ then we have*

$$L(x_1 : \ldots : x_m) \in \begin{cases} Y \text{ if } m = n \\ L_{n-m}(X;Y) \text{ if } m < n \end{cases}$$

Proof. We use induction to prove this so define

$$S=\left\{n\in\mathbb{N}|\text{If } m\in\{1,\ldots,n\} \text{ then } \forall L\in L_n(X;Y) \text{ we have } \forall(x_1,\ldots,x_m)\in X^m \text{ that } L(x_1{:}\ldots{:}x_m)\in\left\{\begin{array}{l} Y \text{ if } m=n\\ L_{n-m}(X;Y) \text{ if } m<n\end{array}\right.\right\}$$

then we have:

$1\in S$. If $m\in\{1,...,1\}=\{1\}$ then $m=1=n$. Let $L\in L_1(X;Y)=L(X,Y)$ then we have for every $(x_1)\in X^1$ that $L(x_1{:}...{:}x_1)=L(x_1)\in Y$. Proving that $1\in S$.

$n\in S\Rightarrow n+1\in S$. Let $L\in L_{n+1}(X;Y)$ then for $m\in\{1,\ldots,n+1\}$ we have for m either:

$m=1$. As $L_{n+1}(X;Y)=L(X,L_n(X;Y))$ and $L(x_1{:}\ldots{:}x_m)=L(x_1{:}\ldots{:}x_1)=L(x_1)$ it follows that

$$L(x_1{:}\ldots{:}x_m)\in L_n(X;Y)=L_{(n+1)-m}(X;Y)$$

$m=n+1$. Given $(x_1,\ldots,x_{n+1})\in X^m=X^{n+1}$ we have, as $L_{n+1}(X;Y)=L(x,L_n(X;Y))$, that $L(x_1)\in L_n(X;Y)$. As $n\in S$, $(x_2,\ldots,x_{n+1})\in X^n$ it follows that

$$(L(x_1))(x_2{:}\ldots{:}x_{n+1})\in Y$$

hence

$$L(x_1{:}\ldots{:}x_m)\underset{m=n+1}{=}L(x_1{:}\ldots{:}x_{n+1})=(L(x_1))(x_2{:}\ldots{:}x_{n+1})\in Y$$

proving as $m\nless n+1$ that

$$L(x_1{:}\ldots{:}x_m)\in\left\{\begin{array}{l} Y \text{ if } m=n+1\\ L_{(n+1)-m}(X;Y) \text{ if } m<n+1\end{array}\right.$$

$m\in\{2,\ldots,n\}$. Given $(x_1,\ldots,x_m)\in X^m$ we have, as $L_{n+1}(X;Y)=L(X,L_n(X;Y))$, that $L(x_1)\in L_n(X;Y)$. As $n\in S$, $(x_2,\ldots,x_m)\in X^{m-1}$ and $1\leqslant m-1<n$ it follows that

$$(L(x_1))(x_2{:}\ldots{:}x_m)\in L_{n-(m-1)}(X;Y)=L_{(n+1)-m}(X;Y)$$

hence

$$L(x_1{:}\ldots{:}x_m)=(L(x_1))(x_2{:}\ldots{:}x_m)\in L_{(n+1)-m(X;Y)}$$

proving as $1<m<n+1$

$$L(x_1{:}\ldots{:}x_m)\in\left\{\begin{array}{l} Y \text{ if } m=n+1\\ L_{(n+1)-m}(X;Y) \text{ if } m<n+1\end{array}\right.$$

So in all cases we have

$$\forall(x_1,\ldots,x_m) \text{ that } L(x_1{:}\ldots{:}x_m)\in\left\{\begin{array}{l} Y \text{ if } m=n+1\\ L_{(n+1)-m}(X;Y) \text{ if } m<n+1\end{array}\right.$$

proving that $n+1\in S$. □

In case $X=\mathbb{K}$ we have a simpler expression for $L(x_1{:}\ldots{:}x_n)$.

Theorem 16.77. *Let $n\in\mathbb{N}$, $\langle X,\|\|\rangle$ a normed space, $L\in L_n(\mathbb{K};X)$ and $x=(x_1,\ldots,x_n)\in\mathbb{K}^n$ then we have*

$$L(x_1{:}\ldots{:}x_n)=\left(\prod_{i=1}^n x_i\right)\cdot L\Big(\underbrace{1{:}\ldots{:}1}_{n}\Big)$$

Proof. We prove this by induction, so define

$$S=\left\{n\in\mathbb{N}|\text{If } L\in L_n(\mathbb{K};X) \text{ and } (x_1,\ldots,x_n)\in\mathbb{K}^n \text{ then } L(x_1{:}\ldots{:}x_n)=\left(\prod_{i=1}^n x_i\right)\cdot L\Big(\underbrace{1{:}\ldots{:}1}_{n}\Big)\right\}$$

then we have:

$1\in S$. If $L\in L_1(\mathbb{K};X)=L(\mathbb{K},X)$ then for $(x_1)\in\mathbb{K}^1$ we have

$$L(x_1{:}\ldots{:}x_1)=L(x_1)=L(x_1\cdot 1)\underset{L\in L(\mathbb{K},X)}{=}x_1\cdot L(1)=\left(\prod_{i=1}^1 x_i\right)\cdot L\Big(\underbrace{1{:}\ldots{:}1}_{1}\Big)$$

proving that $1\in S$.

$n\in S\Rightarrow n+1\in S$. If $L\in L_{n+1}(\mathbb{K};X)=L(\mathbb{K},L_n(\mathbb{K};X))$ then we have for $(x_1,\ldots,x_{n+1})\in\mathbb{K}^n$ that $L(x_1)\in L_n(\mathbb{K};X)$. So as $n\in S$ and $(x_2,\ldots,x_{n+1})\in\mathbb{K}^n$ we have

$$\begin{aligned}
L(x_1)(x_2:\ldots:x_{n+1}) &= \left(\prod_{i=2}^{n+1} x_i\right)\cdot(L(x_1))\Bigl(\underbrace{1:\ldots:1}_{n}\Bigr)\\
&\underset{L\in L(X,L_n(\mathbb{K};X))}{=} \left(\prod_{i=2}^{n+1} x_i\right)\cdot x_1\cdot(L(1))\Bigl(\underbrace{1:\ldots:1}_{n}\Bigr)\\
&= \left(\prod_{I=1}^{n+1} x_i\right)\cdot(L(1))\Bigl(\underbrace{1:\ldots:1}_{n}\Bigr)\\
&= \left(\prod_{i=1}^{n+1} x_i\right)\cdot L\Bigl(\underbrace{1:\ldots:1}_{n+1}\Bigr)
\end{aligned}$$

proving that

$$L(x_1:\ldots:x_{n+1})=L(x_1)(x_2:\ldots:x_{n+1})=L\Bigl(\underbrace{1:\ldots:1}_{n+1}\Bigr)\cdot\prod_{i=1}^{n} x_i$$

hence

$$n+1\in S$$

□

Lemma 16.78. *Let $n\in\mathbb{N}\setminus\{1\}$, $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$ normed spaces and $L\in L_n(X;Y)$ then for $(x_1,\ldots,x_n)\in X^n$ we have*

$$L(x_1:\ldots:x_n)=(L(x_1:\ldots:x_{n-1}))(x_n)$$

Proof. We use induction to prove this, so define

$$S=\{n\in\{2,\ldots,n\}|\text{If } L\in L_n(X;Y) \text{ then for } (x_1,\ldots,x_n)\in X^n \text{ we have } L(x_1:\ldots:x_n)=(L(x_1:\ldots:x_{n-1}))(x_n)\}$$

then we have:

$2\in S$. If $L\in L_2(X;Y)=L(X,L_1(X;Y))=L(X,L(X,Y))$ then for $(x_1,x_2)\in X^2$ we have

$$L(x_1:\ldots:x_2)=(L(x_1))(x_2:\ldots:x_2)=(L(x_1))(x_2)=(L(x_1:\ldots:x_1))(x_2)=(L(x_1:\ldots:x_{2-1}))(x_2)$$

proving that $2\in S$.

$n\in S\Rightarrow n+1\in S$. Let $L\in L_{n+1}(X;Y)$ then for $(x_1,\ldots,x_{n+1})\in X^{n+1}$ we have, as $L_{n+1}(X;Y)=L(X,L_n(X;Y))$, that $L(x_1)\in L_n(X;Y)$. As $n\in S$ and $(x_2,\ldots,x_n)\in X^n$ we have that

$$(L(x_1))(x_2:\ldots:x_{n+1})=((L(x_1))(x_2:\ldots:x_n))(x_{n+1})=(L(x_1:\ldots:x_n))(x_{n+1})$$

so that

$$L(x_1:\ldots:x_{n+1})=(L(x_1))(x_2:\ldots:x_{n+1})=(L(x_1:\ldots:x_n))(x_{n+1})$$

proving that

$$n+1\in S$$

□

Lemma 16.79. *Let $n\in\mathbb{N}$, $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$ normed spaces, $\alpha\in\mathbb{K}$, L_1, $L_2\in L_n(X;Y)$ then $\forall(x_1,\ldots,x_n)\in X^n$ we have*

$$(L_1+\alpha\cdot L_2)(x_1:\ldots:x_n)=L_1(x_1:\ldots:x_n)+\alpha\cdot L_2(x_1:\ldots:x_n)$$

Proof. The proof is by induction, so let

$$S=\{n\in\mathbb{N}|\forall\alpha\in\mathbb{K},\ \forall L_1,L_2\in L_n(X;Y)\text{ we have }\forall(x_1,\ldots,x_n)\in X^n\text{ that }(L_1+\alpha\cdot L_2)(x_1:\ldots:x_n)=L_1(x_1:\ldots:x_n)+\alpha\cdot L_2(x_1:\ldots:x_n)\}$$

then we have:

$1\in S$. Take $\alpha\in\mathbb{K}$, $L_1,L_2\in L_1(X;Y)=L(X,Y)$ then we have for $(x_1)\in X^1$ that

$$\begin{array}{lcl}(L_1+\alpha\cdot L_2)(x_1:\ldots:x_1) & = & (L_1+\alpha\cdot L_2)(x_1)\\ & \underset{\text{pointwise definition}}{=} & L_1(x)+\alpha\cdot L_2(x_1)\\ & = & L_1(x_1:\ldots:x_1)+\alpha\cdot L_2(x_1:\ldots:x_1)\end{array}$$

proving that $1\in S$

$n\in S\Rightarrow n+1\in S$. Take $\alpha\in\mathbb{K}$, $L_1,L_2\in L_{n+1}(X;Y)=L(X,L_n(X;Y))$ so that

$$(L_1+\alpha\cdot L_2)(x_1)\underset{\text{pointwise definition}}{=}L_1(x_1)+\alpha\cdot L_2(x_1)$$

hence

$$\begin{array}{rcl}(L_1+\alpha\cdot L_2)(x_1:\ldots:x_n) & = & ((L_1+\alpha\cdot L_2)(x_1))(x_2:\ldots:x_n)\\ & = & (L_1(x_1)+\alpha\cdot L_2(x_1))(x_2:\ldots:x_n)\end{array}\tag{16.45}$$

As $L_1(x_1),L_2(x_1)\in L_n(X;Y)$ and $n\in S$ it follows that

$$\begin{array}{rcl}(L_1(x_1)+\alpha\cdot L_2(x_1))(x_2:\ldots:x_n) & = & (L_1(x_1))(x_2:\ldots:x_n)+\alpha\cdot(L_2(x_1))(x_2:\ldots:x_n)\\ & = & L_1(x_1:\ldots:x_n)+\alpha\cdot L_2(x_1:\ldots:x_n)\end{array}$$

which combined with [eq: 16.45] proves that

$$(L_1+\alpha\cdot L_2)(x_1:\ldots:x_n)=L_1(x_1:\ldots:x_n)+\alpha\cdot L_2(x_1:\ldots:x_n)$$

Hence we have that

$$n+1\in S$$

□

Lemma 16.80. *Let $n\in\mathbb{N}$, $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$ normed spaces and $L\in L_n(X;Y)$, $i\in\{1,\ldots,n\}$ then for $i\in\{1,\ldots,n\}$ we have for $\left(x_1,\ldots\underbrace{r+\alpha\cdot t}_{i}\ldots,x_n\right)\in X^n$ that*

$$L\left(x_1:\ldots\underbrace{r+\alpha\cdot t}_{i}\ldots:x_n\right)=L\left(x_1:\ldots\underbrace{r}_{i}\ldots:x_n\right)+\alpha\cdot L\left(x_1:\ldots\underbrace{t}_{i}\ldots:x_n\right)$$

Proof. We use mathematical induction for the proof, so let

$S=\left\{n\in\mathbb{N}|\text{If } L\in L_n(X;Y) \text{ then } \forall i\in\{1,\ldots,n\} \text{ we have for } \left(x_1,\ldots\underbrace{r+\alpha\cdot t}_{i}\ldots,x_n\right)\in X^n \text{ that } L\left(x_1,\ldots\underbrace{r+\alpha\cdot t}_{i}\ldots,x_n\right)=L\left(x_1,\ldots\underbrace{r}_{i}\ldots,x_n\right)+\alpha\cdot L\left(x_1,\ldots\underbrace{t}_{i}\ldots,x_n\right)\right\}$

then we have:

$1\in S$. Let $L\in L_1(X;Y)=L(X,Y)$ then if $i\in\{1,\ldots,1\}=\{1\}$ we have $i=1$ and given

$$(r+\alpha\cdot t)=\left(x_1,\ldots\underbrace{r+\alpha\cdot t}_{1}\ldots,x_1\right)\in X^1$$

that

$$\begin{array}{rl}L\left(x_1,\ldots\underbrace{r+\alpha\cdot t}_{i}\ldots:x_n\right) & =\\ L(r+\alpha\cdot t) & \underset{L\in L(X,Y)}{=}\\ L(y)+\alpha\cdot L(z) & =\\ L\left(x_1,\ldots\underbrace{r}_{i}\ldots:x_n\right)+\alpha\cdot L\left(x_1,\ldots\underbrace{t}_{i}\ldots:x_n\right) & \end{array}$$

proving that $1\in S$.

$n\in S\Rightarrow n+1\in S$. Let $L\in L_{n+1}(X;Y)=L(X,L_n(X;Y))$ and take $i\in\{1,\ldots,n+1\}$ then we have either:

$i=1$. Then for $\left(x_1,\ldots\underbrace{r+\alpha\cdot t}_{1}\ldots,x_{n+1}\right)\in X^{n+1}$ we have

$$\begin{aligned}
L\left(x_1,\ldots\underbrace{r+\alpha\cdot t}_{1}\ldots:x_{n+1}\right) &= \\
L(r+\alpha\cdot t:x_2:\ldots:x_{n+1}) &= \\
(L(r+\alpha\cdot t))(x_2:\ldots:x_{n+1}) &\underset{L\in L(X,L_n(X;Y))}{=} \\
(L(r)+\alpha\cdot L(t))(x_2:\ldots:x_{n+1}) &\underset{\text{[lemma: 16.79]}}{=} \\
(L(r))(x_2:\ldots:x_{n+1})+\alpha\cdot L(t)(x_2:\ldots:x_{n+1}) &= \\
L(r:x_2:\ldots:x_{n+1})+\alpha\cdot L(t:x_2:\ldots:x_{n+1}) &= \\
L\left(x_1:\ldots\underbrace{r}_{i}\ldots:x_{n+1}\right)+\alpha\cdot L\left(x_1:\ldots\underbrace{s}_{i}\ldots:x_{n+1}\right)
\end{aligned}$$

proving that in this case that

$$n+1\in S$$

$i\in\{2,\ldots,n+1\}$. Then for $\left(x_1\ldots\underbrace{r+\alpha\cdot t}_{1}\ldots x_{n+1}\right)\in X^{n+1}$ we have $\left(x_2\ldots\underbrace{r+\alpha\cdot t}_{i-1}\ldots x_{n+1}\right)\in X^n$ which as $L(x_1)\in L_n(X;Y)$ and $n\in S$ results in

$$\begin{aligned}
L\left(x_1,\ldots\underbrace{r+\alpha\cdot t}_{i}\ldots,x_{n+1}\right) &= \\
(L(x_1))\left(x_2:\ldots\underbrace{r+\alpha\cdot t}_{i-1}\ldots:x_n\right) & \\
(L(x_1))\left(x_2:\ldots\underbrace{r}_{i-1}\ldots:x_n\right)+\alpha\cdot(L(x_1))\left(x_2:\ldots\underbrace{s}_{i-1}\ldots:x_n\right) &= \\
L\left(x_1:\ldots\underbrace{r}_{i}\ldots:x_n\right)+\alpha\cdot L(x_1:\ldots\underbrace{t}_{i}\ldots:x_n) &=
\end{aligned}$$

so that in this case we have also

$$n+1\in S$$

So in all cases we have

$$n+1\in S$$

□

Lemma 16.81. *Let $n\in\mathbb{N}$, $\langle X,\|\|\rangle$, $\langle Y,\|\|_Y\rangle$ normed spaces, $L\in L_n(X;Y)$ and $(x_1,\ldots,x_n)\in X^n$ [so that by [lemma: 16.76] $L(x_1:\ldots:x_n)\in Y$] then*

$$\|L(x_1:\ldots:x_n)\|_Y\leqslant\|L\|_{L_n(X;Y)}\cdot\prod_{i=1}^{n}\|x_i\|_X$$

Proof. We prove this by induction, so define

$$S=\left\{n\in\mathbb{N}|\text{If } L\in L_n(X;Y)\text{ then }\forall(x_1,\ldots,x_n)\in X^n\text{ we have }\|L(x_1:\ldots:x_n)\|_Y\leqslant\|L\|_{L_n(X;Y)}\cdot\prod_{i=1}^{n}\|x_i\|_X\right\}$$

then we have:

$1\in S$. If $L\in L_1(X;Y)=L(X,Y)$ then $\forall(x_1)\in X^1$ we have

$$\|L(x_1:\ldots:x_1)\|_Y=\|L(x_1)\|_Y\leqslant\|L\|_{L(X,Y)}\cdot\|x_1\|_X=\|L\|_{L_1(X;Y)}\cdot\|x_1\|_X=\|L\|_{L_1(X;Y)}\cdot\prod_{i=1}^{1}\|x_i\|$$

proving that $1\in S$.

$n\in S\Rightarrow n+1\in S$. Let $L\in L_{n+1}(X;Y)=L(X,L_n(X;Y))$ and $(x_1,\ldots,x_{n+1})\in X^n$ then, as $L(x_1)\in L_n(X;Y)$, $(x_2,\ldots,x_{n+1})\in X^n$ and $n\in S$, it follows that

$$\|L(x_1)(x_2:\ldots:x_{n+1})\|_Y\leqslant\|L(x_1)\|_{L_n(X;Y)}\cdot\prod_{i=1}^{n}\|x_{i+1}\|_X=\|L(x_i)\|_{L_n(X;Y)}\cdot\prod_{i=2}^{n+1}\|x_i\|_X \tag{16.46}$$

Next we have

$$\begin{array}{rcl}\|L(x_1\colon\ldots\colon x_{n+1})\|_Y & = & \|(L(x_1))(x_2\colon\ldots\colon x_{n+1})\|_Y\\ & \leqslant_{[\text{eq: } 16.46]} & \|L(x_i)\|_{L_n(X;Y)}\cdot\prod_{i=2}^{n+1}\|x_i\|_X\\ & \leqslant & \|L\|_{L_{n+1}(X;Y)}\cdot\|x_1\|_X\cdot\prod_{i=2}^{n+1}\|x_i\|_X\\ & \leqslant & \|L\|_{L_{n+1}(X;Y)}\cdot\prod_{i=1}^{n+1}\|x_i\|_X\end{array}$$

proving that

$$n+1\in S$$

□

Lemma 16.82. *Let $n\in\mathbb{N}$, $\langle X,\|\|\rangle$, $\langle Y,\|\|_Y\rangle$ normed spaces, $L_1,L_2\in L_n(X;Y)$ such that $\forall(x_1,\ldots,x_n)\in X^n$ we have $L_1(x_1\colon\ldots\colon x_n)=L_2(x_1\colon\ldots\colon x_n)$ then $L_1=L_2$.*

Proof. We use induction to prove this, so define

$$S=\{n\in\mathbb{N}|\text{If } L_1,L_2\in L_n(X;Y) \text{ satisfies } \forall(x_1,\ldots,x_n)\ L_1(x_1\colon\ldots\colon x_n)=L_2(x_1\colon\ldots\colon x_n)\text{ then } L_1=L_2\}$$

then we have:

$1\in S$. Let $L_1,L_2\in L_1(X;Y)=L(X,Y)$ be such that $\forall(x_1)\in X^1$ we have $L_1(x_1\colon\ldots\colon x_1)=L_2(x_1\colon\ldots\colon x_1)$. Then $\forall x\in X$ we have for $y=(x)\in X^1$ [so that $y_1=x$] that

$$L_1(x)=L_1(y_1)=L_1(y_1\colon\ldots\colon y_1)=L_2(y_1\colon\ldots\colon y_1)=L_2(y_1)=L_2(x)$$

proving that $L_1=L_2$, hence $1\in S$.

$n\in S\Rightarrow n+1\in S$. Let $L_1,L_2\in L_{n+1}(X;Y)=L(X,L_n(X;Y))$ such that

$$\forall x=(x_1,\ldots,x_{n+1})\in X^{n+1}\ L_1(x_1\colon\ldots\colon x_{n+1})=L_2(x_1\colon\ldots\colon x_{n+1})$$

Let $y\in X$ then $L_1(y),L_2(y)\in L_n(X;Y)$ and $\forall(x_1,\ldots,x_n)\in X^n$ we have

$$L_1(y)(x_1\colon\ldots\colon x_n)=L_1(y\colon x_1\colon\ldots\colon x_n)=L_2(y\colon x_1\colon\ldots\colon x_n)=L_2(y)(x_1\colon\ldots\colon x_n)$$

which as $n\in S$ proves that $L_1(y)=L_2(y)$. As y is chosen arbitrary we conclude that $L_1=L_2$ proving

$$n+1\in S$$

□

Lemma 16.83. *Let $n\in\mathbb{N}$, $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$ normed spaces an $L\in L_n(X;Y)$. If $\forall(x_1,\ldots,x_n)\in X^n$ we have that $\|L(x_1\colon\ldots\colon x_n)\|_Y\leqslant M\cdot\prod_{i=1}^n\|x_i\|_X$ then it follows that*

$$\|L\|_{L_n(X;Y)}\leqslant M$$

Proof. We prove this by induction, so let

$$S=\left\{n\in\mathbb{N}|\text{If } L\in L_n(X;Y) \text{ satisfies } \forall(x_1,\ldots,x_n)\in X^n\ \|L(x_1\colon\ldots\colon x_n)\|_Y\leqslant M\cdot\prod_{i=1}^n\|x_i\|_X \text{ then } \|L\|_{L_n(X;Y)}\leqslant M\right\}$$

then we have:

$1\in S$. Let $L\in L_1(X;Y)=L(X,Y)$ such that $\forall(x_1)\in X^n$ we have $\|L(x_1\colon\ldots\colon x_1)\|_Y\leqslant M\cdot\prod_{i=1}^1\|x_i\|_X$. Then given $y\in X$ we have for $x=(y)\in X^1$ that

$$\|L(y)\|_Y=\|L(x_1)\|_Y=\|L(x_1\colon\ldots\colon x_1)\|_Y\leqslant M\cdot\prod_{i=1}^1\|x_i\|_X=M\cdot\|y\|_X$$

so that by [definition: 14.178] $\|L\|_{L_1(X;Y)}=\|L\|_{L(X,Y)}\leqslant M$ proving that $1\in S$.

$n\in S\Rightarrow n+1\in S$. Let $L\in L_{n+1}(X;Y)=L(X,L_n(X;Y))$ such that $\forall(x_1,\ldots,x_{n+1})\in X^{n+1}$ we have

$$\|L(x_1\colon\ldots\colon x_n)\|_Y\leqslant M\cdot\prod_{i=1}^{n+1}\|x_i\|_X$$

Fix $x \in X$ then $\forall (y_1, \ldots, y_n) \in X^n$ we have

$$\|(L(x))(y_1 \colon \ldots \colon y_n)\|_Y = \|L(x \colon y_1 \colon \ldots \colon y_n)\|_Y \leqslant M \cdot \|x\|_X \cdot \prod_{i=1}^{n} \|y_i\|_X = (M \cdot \|x\|_X) \cdot \prod_{i=1}^{n} \|y_i\|_X$$

As $n \in S$ and $L(x) \in L_n(X;Y)$ it follows that $\|L(x)\|_{L_n(X;Y)} \leqslant M \cdot \|x\|_X$. Hence by [definition: 14.178] it follows that $\|L\|_{L_{n+1}(X;Y)} \leqslant M$ proving

$$n+1 \in S$$

□

Lemma 16.84. *Let $n \in \mathbb{N}$, $\langle X, \|\|_X \rangle$, $\langle Y, \|\|_Y \rangle$ normed spaces and $L \in L^n(X;Y)$ then there exist a $K \in L_n(X;Y)$ such that $\forall (x_1,...,x_n) \in X^n$ we have $K(x_1 \colon \ldots \colon x_n) = L(x_1, \ldots, x_n)$*

Proof. We use induction to prove this so let

$$S = \{n \in \mathbb{N} | \text{If } L \in L^n(X;Y) \text{ then there exist a } K \in L_n(X;Y) \text{ such that } \forall (x_1, \ldots, x_n) \in X^n \text{ we have } K(x_1 \colon \ldots \colon x_n) = L(x_1, \ldots, x_n)\}$$

then we have:

$1 \in S$. If $L \in L^1(X;Y)$ then as $L^1(X;Y) = L(X,Y) = L_1(\text{XlY})$ we have that $L \in L_1(X;Y)$ so if we take $K = L$ then we have $K \in L_1(X;Y)$ and for $(x_1) \in X^1$ we have $K(x_1 \colon \ldots \colon x_1) = L(x_1) = L(x_1, \ldots, x_1)$. So we have that $1 \in S$.

$n \in S \Rightarrow n+1 \in S$. Let $L \in L^{n+1}(X;Y)$. Take $x \in X$ and define

$$L_x \colon X^n \to Y \text{ by } L_x(x_1, \ldots, x_n) = L(x, x_1, \ldots, x_n)$$

If $i \in \{1, \ldots, n\}$ and $\left(x_1, \ldots \underbrace{r + \alpha \cdot t}_{i} \ldots, x_n\right) \in X^n$ then

$$\begin{array}{rcl} L_x\left(x_1, \ldots \underbrace{r + \alpha \cdot t}_{i} \ldots, x_n\right) & = & \\ L\left(x, x_1, \ldots \underbrace{r + \alpha \cdot t}_{i+1} \ldots, x_n\right) & =_{[\text{lemma: } 16.80]} & \\ L\left(x, x_1, \ldots \underbrace{r}_{i+1} \ldots, x_n\right) + \alpha \cdot L\left(x, x_1, \ldots \underbrace{t}_{i+1} \ldots, x_n\right) & = & \\ L_x\left(x_1, \ldots \underbrace{r}_{i} \ldots, x_n\right) + \alpha \cdot L_x\left(x_1, \ldots \underbrace{s}_{i} \ldots, x_n\right) & & \end{array}$$

proving that

$$L_x \in \text{Hom}^n(X;Y) \tag{16.47}$$

Let $(x_1, \ldots, x_n) \in X^n$ then for $(y_1, \ldots, y_{n+1})$ defined by $y_i = \begin{cases} x \text{ if } i = 1 \\ x_{i-1} \text{ if } i \in \{2, \ldots, n+1\} \end{cases}$

$$\begin{array}{rcl} \|L_x(x_1, \ldots, x_n)\|_Y & = & L(x, x_1, \ldots, x_n) \\ & = & L(y_1, \ldots, y_{n+1}) \\ & \leqslant_{L \in L^{n+1}(X;Y)} & \|L\|_{L^{n+1}(X;Y)} \cdot \prod_{i=1}^{n+1} \|y_i\|_X \\ & = & \|L\|_{L^{n+1}(X;Y)} \cdot \|y_1\|_X \cdot \prod_{i=2}^{n+1} \|y_i\|_X \\ & = & \|L\|_{L^{n+1}(X;Y)} \cdot \|x\|_X \cdot \prod_{i=2}^{n+1} \|x_{i-1}\|_X \\ & = & (\|x\|_X \cdot \|L\|_{L^{n+1}(X;Y)}) \cdot \prod_{i=1}^{n} \|x_i\|_X \end{array}$$

proving by [theorem: 14.187] and [eq: 16.47] that L_x is continuous. Hence

$$L_x \in L^n(X;Y) \tag{16.48}$$

As $n \in S$ there exist a $K_x \in L_n(X;Y)$ such that $\forall (x_1, \ldots, x_n) \in X^n$ we have

$$K_x(x_1: \ldots : x_n) = L_x(x_1, \ldots, x_n) = L(x, x_1, \ldots, x_n) \tag{16.49}$$

This allows us to define

$$K: X \rightarrow L_n(X;Y) \text{ where } K(x) = K_x$$

then by [eq: 16.49] we have that for $(x_1, \ldots, x_{n+1}) \in X^{n+1}$

$$K(x_1)(x_2: \ldots : x_{n+1}) = K_{x_1}(x_2: \ldots : x_1) = L(x_1, \ldots, x_{n+1}) \tag{16.50}$$

If $x, y \in X$ and $\alpha \in \mathbb{K}$ then we have $\forall (x_1, \ldots, x_n) \in X^n$ that

$$\begin{aligned} K(x + \alpha \cdot y)(x_1: \ldots : x_n) &\underset{\text{[eq: 16.50]}}{=} L(x + \alpha \cdot y, x_1, \ldots, x_n) \\ &\underset{L \in L^{n+1}(X;Y)}{=} L(x, x_1, \ldots, x_n) + \alpha \cdot L(y, x_1, \ldots, x_n) \\ &\underset{\text{[eq: 16.50]}}{=} K(x)(x_1: \ldots : x_n) + \alpha \cdot K(y)(x_1: \ldots : x_n) \\ &\underset{\text{[lemma: 16.79]}}{=} (K(x) + \alpha \cdot K(y))(x_1: \ldots : x_n) \end{aligned}$$

which by [lemma: 16.82] proves that

$$K(x + \alpha \cdot y) = K(x) + \alpha \cdot K(y)$$

Hence we have that

$$K \in \mathrm{Hom}(X, L_n(X;Y)) \tag{16.51}$$

Let $x \in X$ then we have for $(y_1, \ldots, y_n) \in X^n$ that

$$\begin{aligned} \|K(x)(y_1: \ldots : y_n)\|_Y &= \|K_x(y_1: \ldots : y_n)\|_Y \\ &\underset{\text{[eq: 16.50]}}{=} \|L(x, y_1, \ldots, y_n)\|_Y \\ &\leqslant_{L \in L^n(X;Y)} \|L\|_{L^{n+1}(X;Y)} \cdot \|x\|_X \cdot \prod_{i=1}^{n} \|y_i\|_X \\ &= (\|L\|_{L^{n+1}(X;Y)} \cdot \|x\|_X) \cdot \prod_{i=1}^{n} \|y_i\|_X \end{aligned}$$

which as $K(x) \in L_n(X;Y)$ proves using [lemma: 16.83] that

$$\|K(x)\|_{L_n(X;Y)} \leqslant \|L\|_{L^{n+1}(X;Y)} \cdot \|x\|_X$$

Hence using [eq: 16.51] together with [theorem: 14.174] proves that K is continuous or

$$K \in L(X, L_n(X;Y)) = L_{n+1}(X;Y)$$

further

$$K(x_1: \ldots : x_{n+1}) = K(x_1)(x_2: \ldots : x_{n+1}) \underset{\text{[eq: 16.50]}}{=} L(x_1, \ldots, x_{n+1})$$

The above proves that

$$n + 1 \in S$$

□

We are now ready to define the identification mapping between $L_n(X;Y)$ and $L^n(X;Y)$.

Theorem 16.85. *Let $n \in \mathbb{N}$, $\langle X, \|\|_X \rangle$, $\langle Y, \|\|_Y \rangle$ normed spaces, then*

$$\mathcal{I}_{n,X,Y}: L_n(X;Y) \rightarrow L^n(X;Y) \textit{where } \mathcal{I}_{n,X,Y}(L): X^n \rightarrow Y \textit{ is defined by}$$

$$(\mathcal{I}_{n,X,Y}(L))(x_1, \ldots, x_n) = L(x_1: \ldots : x_n)$$

is a linear isometric isomorphism. In other words

1. *$\forall L \in L_n(X;Y)$ $\mathcal{I}_{n,X,Y}(L) \in L^n(X;Y)$ so that $\mathcal{I}_{n,X,Y}: L_n(X;Y) \rightarrow L^n(X;Y)$ is a function*
2. *$\mathcal{I}_{n,X,Y}: L_n(X;Y) \rightarrow L^n(X;Y)$ is a bijection.*

3. *$\forall L_1, L_2 \in L_n(X;Y)$ and $\alpha \in \mathbb{K}$ we have $\mathcal{I}_{n,X,Y}(L_1 + \alpha \cdot L_2) = \mathcal{I}_{n,X,Y}(L_1) + \alpha \cdot \mathcal{I}_{n,X,Y}(L_2)$*
4. *$\forall L \in L_n(X:Y)$ we have $\|L\|_{L_n(X;Y)} = \|\mathcal{I}_{n,X,Y}(L)\|_{L^n(X;Y)}$*

***Note* 16.86.** *If $n=1$ then as $\forall (x_1) \in X^1$ we have $\mathcal{I}_{1,X,Y}(L)(x_1) = L(x_1 : \ldots : x_1) = L(x_1)$ and $L^1(X;Y) = L(X,Y) = L_1(X;Y)$ it follows that*

$$\mathcal{I}_{1,X,Y} = \mathrm{Id}_{L(X,Y)}$$

***Note* 16.87.** *If $K \in L^n(X;Y)$ then we have $K = \mathcal{I}_{n,X,Y}((\mathcal{I}_{n,X,Y})^{-1}(K))$ so that for $x \in X^n$ we have*

$$K(x_1, \ldots, x_n) = \mathcal{I}_{n,X,Y}((\mathcal{I}_{n,X,Y})^{-1}(K))(x_1, \ldots, x_n) = (\mathcal{I}_{n,X,Y})^{-1}(K)(x_1 : \ldots : x_n)$$

proving that

$$\forall x \in X^n \ (\mathcal{I}_{n,X,Y})^{-1}(K)(x_1 : \ldots : x_n) = K(x_1, \ldots, x_n)$$

Proof. We have

1. Let $L \in L_n(X;Y)$ then for $i \in \{1, \ldots, n\}$ and $\left(x_1, \ldots \underbrace{r + \alpha \cdot t}_{i} \ldots, x_n\right) \in X^n$ we have

$$\begin{array}{rcl}
\mathcal{I}_{n,X,Y}(L)\left(x_1, \ldots \underbrace{r + \alpha \cdot t}_{i} \ldots, x_n\right) & = & \\
L\left(x_1 : \ldots \underbrace{r + \alpha \cdot t}_{i} \ldots : x_n\right) & =_{[\text{lemma: } 16.80]} & \\
L\left(x_1 : \ldots \underbrace{r}_{i} \ldots : x_n\right) + \alpha \cdot L\left(x_1 : \ldots \underbrace{t}_{i} \ldots : x_n\right) & = & \\
\mathcal{I}_{n,X,Y}(L)\left(x_1, \ldots \underbrace{r}_{i} \ldots, x_n\right) + \alpha \cdot \mathcal{I}_{n,X,Y}(L)\left(x_1, \ldots \underbrace{s}_{i} \ldots, x_n\right) & &
\end{array}$$

   proving that

$$\mathcal{I}_{n,X,Y}(L) \in \mathrm{Hom}^n(X;Y) \text{ [see definition: 11.246]} \tag{16.52}$$

   Further if $(x_1, \ldots, x_n) \in X^n$ then we have:

$$\begin{array}{rcl}
\|\mathcal{I}_{n,X,Y}(L)(x_1, \ldots, x_n)\|_Y & = & \|L(x_1 : \ldots : x_n)\|_Y \\
& \leqslant_{[\text{lemma: } 16.81]} & \|L\|_{L_n(X;Y)} \cdot \prod_{i=1}^{n} \|x_i\|
\end{array}$$

   Using then [theorems: 14.187 and 14.193] it follows that

$$\mathcal{I}_{n,X,Y}(L) : X^n \rightarrow Y \text{ is continuous and } \|\mathcal{I}_{n,X,Y}(L)\|_{L^n(X;Y)} \leqslant \|L\|_{L_n(X;Y)} \tag{16.53}$$

   Combining the above with [eq: 16.52] proves that

$$\mathcal{I}_{n,X,Y}(L) \in L^n(X;Y)$$

   so that

$$\mathcal{I}_{n,X,Y} : L_n(X;Y) \rightarrow L^n(X;Y) \text{ is a well defined function}$$

2. To prove bijectivity we have to prove:

   injectivity. Let $L_1, L_2 \in L_n(X;Y)$be such that $\mathcal{I}_{n,X,Y}(L_1) = \mathcal{I}_{n,X,Y}(L_2)$. Then $\forall (x_1, \ldots, x_n) \in X^n$ we have

$$L_1(x_1 : \ldots : x_n) = \mathcal{I}_{n,X,Y}(L_1)(x_1, \ldots, x_n) = \mathcal{I}_{n,X,Y}(L_2)(x_1, \ldots, x_n) = L_2(x_1 : \ldots : x_n)$$

   Using [lemma: 16.82] it follows that $L_1 = L_2$ proving injectivity.

   surjectivity. Let $L \in L^n(X;Y)$ then using [lemma: 16.84] there exist a $K \in L_n(X;Y)$ such that $\forall (x_1, \ldots, x_n) \in X^n$ we have $K(x_1 : \ldots : x_n) = L(x_1, \ldots, x_n)$. Hence

$$\mathcal{I}_{n,X,Y}(K)(x_1, \ldots, x_n) = K(x_1 : \ldots : x_n) = L(x_1, \ldots, x_n)$$

   proving that $\mathcal{I}_{n,X,Y}(K) = L$ and thus surjectivity.

3. Let $L_1, L_2 \in L_n(X;Y)$ and $\alpha \in \mathbb{K}$ then $\forall (x_1, \ldots, x_n) \in X^n$ we have

$$\begin{aligned} \mathcal{I}_{n,X,Y}(L_1 + \alpha \cdot L_2)(x_1, \ldots, x_n) &= \\ (L_1 + \alpha \cdot L_2)(x_1 \colon \ldots \colon x_n) &\underset{\text{[lemma: 16.79]}}{=} \\ L_1(x_1 \colon \ldots \colon x_n) + \alpha \cdot L_2(x_1 \colon \ldots \colon x_n) &= \\ \mathcal{I}_{n,X,Y}(L_1)(x_1, \ldots, x_n) + \alpha \cdot \mathcal{I}_{n,X,Y}(L_2)(x_1, \ldots, x_n) &= \\ (\mathcal{I}_{n,X,Y}(L_1) + \alpha \cdot \mathcal{I}_{n,X,Y}(L_2))(x_1, \ldots, x_n) &= \end{aligned}$$

proving that

$$\mathcal{I}_{n,X,Y}(L_1 + \alpha \cdot L_2) = \mathcal{I}_{n,X,Y}(L_1) + \alpha \cdot \mathcal{I}_{n,X,Y}(L_2)$$

4. Let $L \in L_n(X;Y)$ then by [eq: 16.53] we have that

$$\|\mathcal{I}_{n,X,Y}(L)\|_{L^n(X;Y)} \leqslant \|L\|_{L_n(X;Y)} \tag{16.54}$$

For the opposite inequality, let $(x_1, \ldots, x_n) \in X^n$ then

$$\|L(x_1 \colon \ldots \colon x_n)\|_Y = \|\mathcal{I}_{n,X,Y}(L)(x_1, \ldots, x_n)\|_Y \leqslant \|\mathcal{I}_{n,X,Y}(L)\|_{L^n(X;Y)} \cdot \prod_{i \in 1}^{n} \|x_i\|_X$$

so that by [lemma: 16.83] it follows that

$$\|L\|_{L_n(X;Y)} \leqslant \|\mathcal{I}_{n,X,Y}(L)\|_{L^n(X;Y)}$$

which combined with [eq: 16.54] proves that

$$\|L\|_{L_n(X;Y)} = \|\mathcal{I}_{n,X,Y}(L)\|_{L^n(X;Y)}$$

□

16.2.2 Higher order Fréchet differentiation

16.2.2.1 Definition of higher order differentials

We are now ready to define higher order Fréchet differentiation and higher order derivation.

Definition 16.88. *Let $n \in \mathbb{N}$, $\langle X, \|\|_X \rangle$ and $\langle Y, \|\|_Y \rangle$ be normed spaces, U a open set in X, $x \in U$ and $f: U \rightarrow Y$ a function then f is n-times Fréchet differentiable at x with differential $D_x^{[n]} f(x) \in L_n(X;Y)$ if we have for n*

$n = 1$. *then f must be Fréchet differentiable at x and $D^{[1]} f$ is defined by*

$$D_x^{[1]} f = D_x f \in L(X, Y) = L_1(X;Y)$$

$1 < n$. *then there must exist a open set V in X with $x \in V \subseteq U$ such that $\forall y \in V$ f is $(n-1)$-times differentiable at y and the function*

$$D^{[n-1]} f: V \rightarrow L_{n-1}(X;Y) \text{ defined by } D^{[n-1]} f(y) = D_y^{[n-1]} f$$

is Fréchet differentiable at x and $D_x^{[n]} f$ is defined by

$$D_x^{[n]} f = D_x(D^{[n-1]} f) \in L(X, L_{n-1}(X;Y)) = L_n(X;Y)$$

***Note* 16.89.** *In the case $1 < n$ we must ensure that $D_x(D^{[n-1]})$ does not depended on the choice of V to be valid definition. This is ensured by the locality of differentials [see corollary: 16.25].*

Just as the Fréchet differentiation is local [see theorem: 16.24] higher order Fréchet differentiation is local.

Theorem 16.90. *Let $n \in \mathbb{N}$, $\langle X, \|\|_X \rangle$ and $\langle Y, \|\|_Y \rangle$ be normed spaces, U a open set in X, $x \in U$, V a open set with $x \in V \subseteq U$ and f: $U \rightarrow Y$ a function then we have*

$$f \text{ is } n\text{-times differentiable at } x$$
$$\Updownarrow$$
$$f_{|V} \text{ is } n\text{-times differentiable at } x$$

Further if f or $f_{|V}$ is n-times differentiable at x then $D^{[n]}_x f_{|V} = D^{[n]}_x f$.

Proof.

⇒. We prove this by induction so let

$S = \{n \in \mathbb{N} | \text{If } f \text{ is } n\text{-times differentiable at } x \in U \text{ then for every open } V \text{ with } x \in V \subseteq U \text{ we have that } f_{|V} \text{ is } n\text{-times differentiable at } x \text{ and } D^{[n]}_x f_{|V} = D^{[n]}_x f\}$

then we have:

$1 \in S$. If f is 1-times differentiable at $x \in U$ and V a open set with $x \in V \subseteq U$ then by definition f is Fréchet differentiable at x and $D^{[1]}_x f = D_x f$. Using [theorem: 16.24] $f_{|V}$ is Fréchet differentiable at x and $D_x f_{|V} = D_x f$. Hence $f_{|V}$ is 1-times differentiable at x and $D^{[1]}_x f_{|V} = D_x f_{|V} = D_x f = D^1_x f$. From this it follows that $1 \in S$.

$n \in S \Rightarrow n+1 \in S$. If f is $(n+1)$-times differentiable at x and V is a open set with $x \in V \subseteq U$. Then there exist a open set W with $x \in W \subseteq U$ such that $\forall y \in W$ we have that f is n-times differentiable at y and the function

$$D^{[n]} f: W \rightarrow L_n(X; Y) \text{ defined by } D^{[n]} f(y) = D^{[n]}_y f \tag{16.55}$$

is Fréchet differentiable at x and

$$D^{[n+1]}_x f = D_x(D^{[n]} f) \tag{16.56}$$

Take $V' = W \bigcap V$ then V' is open, $x \in V' \subseteq V \subseteq U$ and $\forall y \in V'$ we have, as $y \in W$, that f is n-times differentiable at y. As $n \in S$ it follows that $f_{|V}$ is n-times differentiable at y with

$$D^{[n]}_y f_{|V} = D^{[n]}_y f \tag{16.57}$$

So we can define the function

$$D^{[n]} f_{|V}: V' \rightarrow L_n(X; Y) \text{ where } D^{[n]} f_{|V}(y) = D^{[n]}_y f_{|V} \tag{16.58}$$

which, as $\forall y \in V'$ we have $D^{[n]} f_{|V}(y) = D^{[n]}_y f_{|V} \underset{[\text{eq: } 16.57]}{=} D^{[n]}_y f = D^{[n]} f(y)$, proves that

$$D^{[n]} f_{|V} = (D^{[n]} f)_{|V'} \tag{16.59}$$

As $D^{[n]} f$ is Fréchet differentiable at x it follows from [theorem: 16.24] that $(D^{[n]} f)_{|V'}$ is Fréchet differentiable at x and

$$D_x((D^{[n]} f)_{|V'}) \underset{[\text{theorem: } 16.24]}{=} D_x(D^{[n]} f) \underset{[\text{eq: } 16.56]}{=} D^{[n+1]}_x f \tag{16.60}$$

So, as $D^{[n]} f_{|V} \underset{[\text{eq: } 16.59]}{=} (D^{[n]} f)_{|V'}$, it follows that $D^{[n]} f_{|V}$ is Fréchet differentiable at x and

$$D_x(D^{[n]} f_{|V}) = D_x(D^{[n]} f)_{|V'} \underset{[\text{eq: } 16.60]}{=} D^{[n+1]}_x f$$

Hence by definition we have that

$$f_{|V} \text{ is } (n+1)\text{-times differentiable at } x \text{ with } D^{[n+1]}_x f_{|V} = D^{[n+1]}_x f$$

proving that

$$n+1 \in S$$

⇐. We prove this by induction on n, so let

$S = \{n \in \mathbb{N} | \text{If } f: U \rightarrow Y \text{ is a function and there exist a open } V \text{ with } x \in V \subseteq U \text{ such that } f_{|V}: V \rightarrow Y \text{ is } n\text{-times differentiable at } x \text{ then } f: U \rightarrow Y \text{ is } n\text{-times differentiable at } x \text{ and } D^{[n]}_x f = D^{[n]}_x f_{|V}\}$

then we have:

$1 \in S$. If $f_{|V}$ is 1-times differentiable at $x \in V \subseteq U$ then by definition $f_{|V}$ is Fréchet differentiable at x and $D^{[1]}_x f_{|V} = D_x f_{|V}$. Using [theorem: 16.24] it follows that f is Fréchet differentiable at x and $D_x f = D_x f_{|V}$. Hence f is 1-times differentiable at x and $D^{[1]}_x f(x) = D^{[1]}_x f_{|V}$. This proves that $1 \in S$.

$n\in S\Rightarrow n+1\in S$. If V is a open set with $x\in V\subseteq U$ such that $f_{|V}:V\rightarrow Y$ is $[n+1]$-times differentiable at x then there exist a open set W with $x\in W\subseteq V$ such that $\forall y\in W\subseteq V$ $f_{|V}$ is n-times differentiable at x and the function

$$D^{[n]}f_{|V}:W\rightarrow L_n(X;Y) \text{ defined by } D^{[n]}f_{|V}(x)=D^{[n]}_x f_{|V}$$

is Fréchet differentiable at x and

$$D^{[n+1]}_x f_{|V}=D_x(D^{[n]}f_{|V}) \tag{16.61}$$

As $n\in S$ we have $\forall y\in W$ that f is n-times differentiable at y and that

$$D^{[n]}_y f=D^{[n]}_y f_{|V} \tag{16.62}$$

Hence for the function defined by

$$D^{[n]}f:W\rightarrow L_n(X;Y) \text{ where } D^{[n]}f(y)=D^{[n]}_y f$$

we have $\forall y\in Y$ $D^{[n]}f(y)=D^{[n]}_y f \underset{\text{[eq: 16.62]}}{=} D^{[n]}_y f_{|V}=D^{[n]}f_{|V}(y)$ proving that

$$D^{[n]}f=D^{[n]}f_{|V}.$$

Hence $D^n f$ is Fréchet differentiable at x and

$$D^{[n+1]}_x f \underset{\text{def}}{=} D_x(D^{[n]}f)=D_x(D^{[n]}f_{|V}) \underset{\text{[eq: 16.61]}}{=} D^{[n+1]}_x f_{|V}$$

proving

$$n+1\in S$$

□

Definition 16.91. *Let $n\in\mathbb{N}$, $\langle X,\|\|_X\rangle,\langle Y,\|\|_Y\rangle$ be normed spaces, U a open set in X then a function $f:U\rightarrow X$ is **n-times differentiable on U** if $\forall x\in U$ we have that f is n-times differentiable at x.*

Note that for n-times differentiability at a point we need to specify a subset W with $x\in W\subseteq$ otherwise we have not a function $D^{[n-1]}f$ that we can differentiate. For n-times differentiability on a set we can use this set itself.

Theorem 16.92. *Let $n\in\mathbb{N}\setminus\{1\}$, $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$ be normed spaces, U a open set in X then we have the following equivalences:*

1. *f is n-times Fréchet differentiable on U*
2. *For n we have*

    $n=1$. *Then $\forall x\in U$ f is Fréchet differentiable at x with $D^{[1]}_x f=D_x f$*

    $1<n$. *Then $\forall x\in U$ f is $(n-1)$-times Fréchet differentiable at x and the function*

    $$D^{[n-1]}f:U\rightarrow L_n(X;Y) \text{ defined by } (D^{[n-1]}f)(x)=D^{[n-1]}_x f$$

    is Fréchet differentiable on U. Further $D^{[n]}_x f=D_x(D^{[n-1]}f)$ $\forall x\in U$.

Proof.

$1\Rightarrow 2$. For $n\in\mathbb{N}$ we must consider two cases:

$n=1$. Then as $\forall x\in U$ is f 1-times differentiable at x it follows from [definition: 16.88] that f is Fréchet differentiable at x and $D^{[1]}_x f=D_x f$.

$1<n$. Let $x\in U$ then there exist by [definition: 16.88] a open set V_x with $x\in V_x\subseteq U$ such that $\forall y\in V_x$ f is $(n-1)$-times differentiable at y and for the function

$$D^{[n-1]}f_{V_x}:V_x\rightarrow L_{n-1}(X;Y) \text{ defined by } D^{[n-1]}f_{V_x}(x)=D^{[n-1]}_x f$$

we have that

$$D^{[n-1]}f_{V_x} \text{ is Fréchet differentiable at } x \text{ with } D^{[n]}_x f=D_x(D^{[n-1]}f_{V_x}) \tag{16.63}$$

Let $x\in U$ then as $x\in V_x\subseteq U$ it follows that f is $(n-1)$-times differentiable at x so we can define the function

$$D^{[n-1]}f:U\rightarrow L_{n-1}(X;Y) \text{ by } D^{[n-1]}f(x)=D^{[n-1]}_x f$$

Let $x \in U$ then $\forall y \in V_x$ we have $D^{[n-1]} f_{V_x}(x) = D_x^{[n-1]} f = D^{[n-1]} f(x)$ proving that

$$D^{[n-1]} f_{V_x} = (D^{[n-1]} f)_{|V_x} \tag{16.64}$$

So $(D^{[n-1]} f)_{|V_x}$ is Fréchet differentiable at x [because by [eq: 16.63] $D^{[n-1]} f_{V_x}$ is Fréchet differentiable at x], hence by [theorem: 16.24] $D^{[n-1]} f$ is Fréchet differentiable at x, which as $x \in U$ is chosen arbitrary means that $D^{[n-1]} f$ is Fréchet differentiable on U, and $\forall x \in U$

$$D_x^{[n]} f \underset{\text{[eq: 16.63]}}{=} D_x(D^{[n-1]} f_{V_x}) \underset{\text{[eq: 16.64]}}{=} D_x(D^{[n-1]} f)_{|V_x} \underset{\text{theorem: 16.24]}}{=} D_x(D^{[n-1]} f).$$

2 ⇒ 1. For $n \in \mathbb{N}$ we must consider two cases:

$n = 1$. As $\forall x \in U$ f is Fréchet differentiable at x it follows from [definition: 16.88] that f is 1-times differentiable at x, hence f is 1-times differentiable on U.

$1 < n$. Let $x \in U$ then by taking $V = U$, so that $x \in V \subseteq U$, we have by the hypothesis combined with [definition: 16.88] that f is n-times differentiable at x. Hence as $x \in U$ was chosen arbitrary f is n-times differentiable on U.

□

Definition 16.93. *Let $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$ be normed spaces, U a open set in X then*

$$f: U \rightarrow Y \text{ is } \infty\text{-times differentiable at } x \in U \text{ if } \forall n \in \mathbb{N}\ f \text{ is } n\text{-times differentiable at } x$$

and

$$f: U \rightarrow Y \text{ is } \infty\text{-times differentiable on } U \text{ if } \forall n \in \mathbb{N}\ f \text{ is } n\text{-times differentiable on } U$$

16.2.2.2 Higher order differentials as multi-linear mappings

The problem with $D_x^{[n]} f$ is that it is a element of $L_n(X; Y)$ so that for example

$$D_x^{[3]} f \in L(X, L(X, L(X, Y)))$$

and we have to deal with expressions like $\big(\big(D_x^{[3]} f(h_1)\big)(h_2)\big)(h_3)$. Lucky we can use [theorem: 16.85] to turn $D^{[n]} f(x)$ in a a multilinear mapping.

Definition 16.94. *Let $n \in \mathbb{N}$, $\langle X, \|\|_X\rangle$ and $\langle Y, \|\|_Y\rangle$ be normed spaces, U a open set in X, $x \in U$ and $f: U \rightarrow Y$ a function that is n-times differentiable at x then*

$$D_x^n f = \mathcal{I}_{n,X,Y}\big(D_x^{[n]} f\big) \in L^n(X; Y)$$

Note 16.95. Using [theorem: 16.85] we have that $D_x^1 f = D_x^{[1]} f = D_x f$ and for $n \in \mathbb{N}$

$$D_x^n f(h_1, \ldots, h_n) = D_x^{[n]} f[h_1: \ldots : h_n].$$

We use in general $D_x^{[n]} f$ because the proofs are simpler, however if we have to calculate the n-the differential we can use $D_x^n f$.

16.2.2.3 Differentiable classes

Definition 16.96. *Let $n \in \mathbb{N}_0 \bigcup \{\infty\}$, $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$ normed spaces, U a open set in X and $f: U \rightarrow Y$ a function then f is of class C^n if for*

$n = 0$. *$f: U \rightarrow Y$ is a continuous function.*

$n \in \mathbb{N}$. *f is n-times differentiable on U and $D^{[n]} f: U \rightarrow Y$ defined by $D^{[n]} f(x) = D_x^{[n]} f$ is a continuous function.*

$n = \infty$. *If $\forall n \in \mathbb{N}_0$ we have that f is of class C^n*

Theorem 16.97. *Let $n \in \mathbb{N}_0 \bigcup \{\infty\}$, $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$ normed spaces, U a open set in X, $f: U \rightarrow Y$ a function of class C^n and V a open set with $V \subseteq U$ then $f_{|V}: V \rightarrow Y$ is of class C^n. Further if $n \in \mathbb{N}$ then we have for the functions*

$$D^{[n]} f: U \rightarrow L_n(X; Y) \text{ defined by } D^{[n]} f(x) = D_x^{[n]} f$$

$$D^{[n]} f_{|V}: V \rightarrow L_n(X; Y) \text{ defined by } D^{[n]} f_{|V}(x) = D_x^{[n]} f_{|V}$$

that

$$D^{[n]}f_{|V}=(D^{[n]}f)_{|V}$$

Proof. For n we have either:

$\boldsymbol{n=0}$**.** Then by definition $f:U\rightarrow Y$ is a continuous function and using [theorem: 14.135] it follows that $f_{|V}:V\rightarrow Y$ is a continuous function.

$\boldsymbol{n\in\mathbb{N}}$**.** Then $\forall x\in U$ f is n-times differentiable at x and

$$D^{[n]}f:U\rightarrow L_n(X;Y) \text{ defined by } D^{[n]}f(x)=D_x^{[n]}f$$

is continuous. Using [theorem: 16.90] it follows that $\forall x\in V$ $f_{|V}$ is n-times differentiable at x with $D_x^{[n]}f_{|V}=D_x^{[n]}f$. Hence if we define

$$D^{[n]}f_{|V}:V\rightarrow Y \text{ by } D^{[n]}f_{|V}(x)=D_x^{[n]}f_{|V}$$

then

$$D^{[n]}f_{|V}=(D^{[n]}f)_{|V}$$

So, using [theorem: 14.135], $D^{[n]}f_{|V}$ is continuous. So by definition $f_{|V}$ is of class C^n.

$\boldsymbol{n=\infty}$**.** Then $\forall n\in\mathbb{N}$ f is of class C^n and by the case $n\in\mathbb{N}$ $f_{|V}$ is of class C^n, hence $f_{|V}$ is of class C^n.

□

Every ∞-times differentiable function is of class C^∞ as is proved in the following theorem.

Theorem 16.98. *Let $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$ be normed spaces, U a open set in X and $f:U\rightarrow Y$ a function then*

$$f \text{ is } \infty\text{-times differentiable on } U$$

$$\Updownarrow$$

$$f \text{ is of class } C^\infty$$

Proof.

$\Rightarrow$**.** As $\forall x\in U$ f is ∞-times differentiable it is 1-times differentiable at x, hence by definition it is Fréchet differentiable at x, applying [theorem: 16.31] proves that f is continuous at x. Hence f is a continuous function so that

$$f \text{ is of class } C^0$$

Let $n\in\mathbb{N}$ then as $\forall x\in U$ f is ∞-times differentiable at x it is n-times and $(n+1)$-times differentiable at x, from the $(n+1)$-times differentiability it follows that

$$D^{[n]}f:U\rightarrow Y \text{ defined by } D^{[n]}f(x)=D_x^{[n]}f$$

is Fréchet differentiable at x, hence by [theorem: 16.31] it follows that $D^{[n]}f$ is continuous at x. So $D^{[n]}f$ is a continuous function which proves that

$$f \text{ is of class } C^n$$

$\Leftarrow$**.** This follows from the definition from C^∞. □

Next we examine the relation between function of class C^1 and derivatives, first we need a little lemma.

Lemma 16.99. *Let $\langle\mathbb{K},\|\rangle$ be the normed space of the real (complex) numbers, $\langle X,\|\|\rangle$ a normed space over $\mathbb{K}$, U a open set in $\mathbb{K}$ and*

$$f:U\rightarrow L(\mathbb{K},X) \text{ a function}$$

then if we define

$$f(*)(1):U\rightarrow X \text{ by } (f(*)(1))(x)=f(x)(1)$$

then we have

$$f \text{ is continuous} \Leftrightarrow f(*)(1) \text{ is continuous}$$

Proof.

⇒. Let $x\in U$ and $\varepsilon\in\mathbb{R}^+$. As f is continuous it is continuous at x, hence there exist a $\delta\in\mathbb{R}^+$ such that if $y\in U$ with $|y-x|<\delta$ we have $\|f(y)-f(x)\|_{L(\mathbb{K},X)}<\varepsilon$. Then we have

$$\begin{aligned}\|(f(*)(1))(y)-(f(\star)(1))(x)\| &= \|f(y)(1)-f(x)(1)\|\\ &= \|(f(x)-f(y))(1)\|\\ &\leqslant \|f(x)-f(y)\|_{L(\mathbb{K},X)}\cdot|1|\\ &= \|f(x)-f(y)\|_{L(\mathbb{K},X)}\\ &< \varepsilon\end{aligned}$$

proving continuity of $f(*)(1)$ at x. As x was chosen arbitrary it follows that $f(*)(1)$ is continuous.

⇐. Let $x\in U$ and $\varepsilon\in\mathbb{R}^+$ then, as $f(\star)(1)$ is continuous, $f(*)(1)$ is continuous at x. Hence there exist a $\delta\in\mathbb{R}^+$ such that for $y\in U$ with $|y-x|<\delta$ we have

$$\|(f(*)(1))(y)-(f(\star)(1))(x)\|<\varepsilon$$

then we have for $r\in\mathbb{K}$ that

$$\begin{aligned}\|(f(y)-f(x))(r)\| &= \|(f(y)-f(x))(r\cdot 1)\|\\ &= \|r\cdot(f(y)-f(x))(1)\|\\ &= |r|\cdot\|(f(y)-f(x))(1)\|\\ &= |r|\cdot\|f(y)(1)-f(x)(1)\|\\ &= |r|\cdot\|(f(*)(1))(y)-(f(*)(1))(x)\|\\ &< \varepsilon\cdot|r|\end{aligned}$$

so that by [theorem: 14.181] we have $\|f(y)-f(x)\|_{L(\mathbb{K},X)|}<\varepsilon$, hence f is continuous. As x was chosen arbitrary it follows that $f(*)(1)$ is continuous. □

Theorem 16.100. *Let $\langle\mathbb{K},\|\rangle$ be the normed space of the real (complex) numbers, $\langle X,\|\|\rangle$ a normed space, U a open set in $\mathbb{K}$ and f: $U\rightarrow X$ a function then the following are equivalent:*

1. *f is of class C^1*
2. *$\forall x\in U$ f'_x exist at x and the function*

   $$f'{:}\,U\rightarrow X \text{ defined by } f'(x)=f'_x$$

   is continuous.

Further if (1) or (2) is true then $f'=(D^{[1]}f)(\star)(1)$ where

$$D^{[1]}f{:}\,U\rightarrow L(\mathbb{K},X) \text{ is defined by } D^{[1]}f(x)=D^{[1]}_x f=D_x f$$

Proof.

1 ⇒ 2. As f is of class C^1 we have that $\forall x\in U$ f is 1-times differentiable at x and the function

$$D^{[1]}f{:}\,U\rightarrow L(\mathbb{K},X) \text{ defined by } (D^{[1]}f)(x)=D^{[1]}_x f=D_x f$$

is continuous. Let $x\in U$ then as f is 1-times differentiable at x f is Fréchet differentiable at x. Using [theorem: 16.32] it follows that f has a derivative at x with $D_x f(1)=f'_x$. Hence if we define

$$f'{:}\,U\rightarrow X \text{ by } (f')(x)=f'_x$$

then we have for $x\in U$ that

$$((D^{[1]}f)(\star)(1))(x)=((D^{[1]}f)(x))(1)=D^{[1]}_x f(1)=f'_x=f'(x)$$

so that $(D^{[1]}f)(*)(1)=f'$. Further, as $D^{[1]}f$ is continuous, it follows from [lemma: 16.99] that $(D^{[1]}f)(*)(1)$ is continuous which implies that f' is continuous.

2 ⇒ 1. Let $x\in U$ then as f has a derivative at x it follows from [theorem: 16.32] that f is Fréchet differentiable at x with $D_x f(1)=f'_x$. Hence f is 1-times differentiable at x with $D^{[1]}_x f(1)=D_x f(1)=f'_x$. So we can define the function

$$D^{[1]}f{:}\,U\rightarrow X \text{ by } D^{[1]}f(x)=D^{[1]}_x f$$

As for $x \in U$ $((D^{[1]}f)(*)(1))(x) = (D^{[1]}f)(x)(1) = D_x^{[1]}f(1) = f'_x$ it follows that $f' = (D^{[1]}f)(*)(1)$, so that by [lemma: 16.99] and the fact that f' is continuous it follows that $D^{[1]}f$ is continuous. Hence we conclude that f is of class C^1. □

16.2.2.4 Higher order derivatives

Just as we have higher order differentiation we have also higher order derivatives.

Definition 16.101. *Let $n \in \mathbb{N}$, $\langle \mathbb{K}, \| \rangle$ be the normed space of real (complex) numbers based on the norm $||$, $\langle X, \|\| \rangle$ a normed space, U a open set in $\mathbb{K}$, $x \in U$ and $f: U \to X$ a function then f has a n-the derivative $f^{(n)} \in Y$ at x if for*

- $\boldsymbol{n=1}$**.** *f has a derivative at x and we define $f_x^{(1)}$ by $f_x^{(1)} = f'_x$*
- $\boldsymbol{1<n}$**.** *there exist a open set V with $x \in V \subseteq U$ such that $\forall y \in V$ we have that f has a $(n-1)$-the derivative $f_y^{(n-1)}$ and the function*

$$f^{(n-1)}: V \to X \text{ defined by } f^{(n-1)}(y) = f_y^{(n-1)}$$

has a derivative at x. Then the n-the derivative $f_x^{(n)}$ at x is then defined to be

$$f_x^{(n)} = (f^{(n-1)})'_x$$

Let's examine now the relation between n-times differentiability and the existence of the n-the derivative. First we need a little lemma.

Lemma 16.102. *Let $n \in \mathbb{N}$, $\langle X, \|\| \rangle$ be a normed space, $\langle \mathbb{K}, \| \rangle$ the normed space of real (complex) numbers based on the canonical norm $||$ and $L \in L_n(\mathbb{K}; X)$ then*

$$\|L\|_{L_n(\mathbb{K};X)} \leqslant \left\| L\left(\underbrace{1:\ldots:1}_{n}\right)\right\|$$

Proof. As usually we prove this by induction on n, so define

$$S = \left\{n \in \mathbb{N} | \text{If } L \in L_n(\mathbb{K}; X) \text{ then } \|L\|_{L_n(\mathbb{K};Y)} \leqslant \left\| L\left(\underbrace{1:\ldots:1}_{n}\right)\right\|\right\}$$

then we have:

$\boldsymbol{1 \in S}$**.** If $L \in L_1(\mathbb{K}; X) = L(\mathbb{K}, X)$ then if $x \in X^1$ satisfies $|x| = 1$ we have

$$\|L(x)\| = \|L(x \cdot 1)\| = \|x \cdot L(1)\| = |x| \cdot \|L(1)\| = \left\| L\left(\underbrace{1:\ldots:1}_{1}\right)\right\| \tag{16.65}$$

Further

$$\begin{aligned} \|L\|_{L_1(\mathbb{K};X)} &= \|L\|_{L(\mathbb{K},X)} \\ &\underset{\text{[theorem: 14.181]}}{=} \inf\left(\{M \in \mathbb{R}^+ | \forall x \in \mathbb{K} \text{ with } |x| = 1 \text{ we have } \|L(x)\| \leqslant M\}\right) \\ &\leqslant_{\text{[eq: 16.65]}} \left\| L\left(\underbrace{1:\ldots:1}_{1}\right)\right\| \end{aligned}$$

proving that $1 \in S$.

$\boldsymbol{n \in S \Rightarrow n+1 \in S}$**.** If $L \in L_{n+1}(\mathbb{K}; X) = L(\mathbb{K}, L_n(\mathbb{K}; X))$ then $L(1) \in L_n(\mathbb{K}; X)$. As $n \in S$ we have that

$$\|L(1)\|_{L_n(\mathbb{K};X)} \leqslant \left\| L(1)\left(\underbrace{1:\ldots:1}_{n}\right)\right\| = \left\| L\left(\underbrace{1:\ldots:1}_{n+1}\right)\right\| \tag{16.66}$$

If $x \in \mathbb{K}$ such that $|x| = 1$ then as $L \in L(\mathbb{K}, L_n(\mathbb{K}; X))$ we have $L(x) \in L_n(\mathbb{K}; X)$ so that

$$\begin{aligned} \|L(x)\|_{L_n(\mathbb{K};X)} &= \|L(x \cdot 1)\|_{L_n(\mathbb{K};X)} \\ &= \|x \cdot L(1)\|_{L_n(\mathbb{K};X)} \\ &= |x| \cdot \|L(1)\|_{L_n(\mathbb{K};X)} \\ &= \|L(1)\|_{L_n(\mathbb{K};X)} \\ &\leqslant_{\text{[eq: 16.66]}} \left\| L\left(\underbrace{1:\ldots:1}_{n+1}\right)\right\| \end{aligned} \tag{16.67}$$

Further we have

$$\begin{array}{rl}\|L\|_{L_{n+1}(\mathbb{K};X)} & = \\ \|L\|_{L(\mathbb{K},L_n(\mathbb{K};X))} & =_{\text{[theorem: 14.181]}} \\ \inf(\{M\in\mathbb{R}^+|\forall x\in\mathbb{K} \text{ with } |x|=1 \text{ we have } \|L(x)\|_{L_n(\mathbb{K};X)}\leqslant M\}) & \leqslant_{\text{[eq: 16.67}} \\ \left\|L\left(\underbrace{1:\ldots:1}_{n+1}\right)\right\| & \end{array}$$

Hence we have

$$n+1\in S$$

□

Theorem 16.103. *Let $n\in\mathbb{N}$, $\langle\mathbb{K},\|\rangle$ be the normed space of real (complex) numbers based on the norm $||$, $\langle X,\|\|\rangle$ a normed space, U a open set in $\mathbb{K}$, $x\in U$ and $f:U\to X$ a function then*

$$f \text{ is } n\text{-times differentiable at } x$$

$$\Updownarrow$$

$$f \text{ has a } n\text{-the derivative at } x$$

Further if f is n-times differentiable at x or f has a n-the derivative at x then

$$f_x^{(n)}=D_x^{[n]}f\left(\underbrace{1:\ldots:1}_{n}\right)\underset{\text{[definition: 16.94]}}{=}D_x^n f\left(\underbrace{1,\ldots,1}_{n}\right)$$

Proof.

⇒. We prove this by induction so define

$$S=\left\{n\in\mathbb{N}|\text{If } f \text{ is } n\text{-times differentiable at } x \text{ then } f \text{ has a } n\text{-the derivative at } x \text{ and } f_x^{(n)}=D_x^{[n]}f\left(\underbrace{1:\ldots:1}_{n}\right)\right\}$$

then we have:

$n=1$. If f is 1-times differentiable at x then by definition f is Fréchet differentiable at x and $D_x^{[1]}f=D_xf$. Using [theorem: 16.32], f has a derivative at x and

$$f_x'=D_xf(1)=D_xf\left(\underbrace{1:\ldots:1}_{1}\right)=D_x^{[1]}f\left(\underbrace{1:\ldots:1}_{n}\right)$$

Hence by definition f has a 1-the derivative at x with $f_x^{(1)}=f_x'=D_x^{[1]}f\left(\underbrace{1:\ldots:1}_{n}\right)$ proving that $1\in S$.

$1<n$. If f is $(n+1)$-times differentiable at x then there exist a open set V with $x\in V\subseteq U$ such that $\forall y\in V$ f is n-times differentiable at x and the function

$$D^{[n]}f:V\to L_n(\mathbb{K};X) \text{ defined by } D^{[n]}f(y)=D_y^{[n]}f$$

is Fréchet differentiable at x. Let $\varepsilon\in\mathbb{R}^+$ then by the Fréchet differentiability at x there exist a $\delta\in\mathbb{R}^+$ such that $\forall h\in V_x$ with $|h|<\delta$ we have

$$\|D^{[n]}f(x+h)-D^{[n]}f(x)-(D_x(D^{[n]}f))(h)\|_{L_n(X;Y)}\leqslant\frac{\varepsilon}{2}\cdot|h|<\varepsilon\cdot|h| \tag{16.68}$$

Further we have

$$D_x^{[n+1]}f=D_x(D^{[n]}f)\in L(\mathbb{K},L_n(\mathbb{K},X))=L_{n+1}(\mathbb{K};X) \tag{16.69}$$

As $n\in S$ we have, as $\forall y\in V$ is f n-times differentiable at y, that

$$f \text{ has a } n\text{-the derivative at } y \text{ with } f_y^{(n)}=D_y^{[n]}f\left(\underbrace{1:\ldots:1}_{n}\right) \tag{16.70}$$

So we can define the function

$$f^{(n)}:V\to X \text{ by } f^{(n)}(y)=f_y^{(n)}$$

Then for $h \in V_x$ with $0 < |h| < \delta$ we have

$$\begin{aligned}
\left\| f^{(n)}(x+h) - f^{(n)}(x) - h \cdot D_x^{[n+1]} f\left(\underbrace{1:\ldots:1}_{n+1}\right)\right\| &= \\
\left\| f^{(n)}(x+h) - f^{(n)}(x) - h \cdot \left(D_x^{[n+1]} f(1)\right)\left(\underbrace{1:\ldots:1}_{n}\right)\right\| &= \\
\left\| f^{(n)}(x+h) - f^{(n)}(x) - \left(D_x^{[n+1]} f(h)\right)\left(\underbrace{1:\ldots:1}_{n}\right)\right\| & \\
\left\| f^{(n)}_{x+h} - f^{(n)}_x - \left(D_x^{[n+1]} f(h)\right)\left(\underbrace{1:\ldots:1}_{n}\right)\right\| &\underset{[\text{eq: } 16.70]}{=} \\
\left\| D^n_{x+h} f\left(\underbrace{1:\ldots:1}_{n}\right) - D^n_x f\left(\underbrace{1:\ldots:1}_{n}\right) - D_x^{[n+1]} f(h)\left(\underbrace{1:\ldots:1}_{n}\right)\right\| &\underset{[\text{lemma: } 16.79]}{=} \\
\left\| \left(D^n_{x+h} f - D^n_x f - D_x^{[n+1]} f(h)\right)\left(\underbrace{1:\ldots:1}_{n}\right)\right\| &\leqslant \\
\left\| D^n_{x+h} f - D^n_x f - D_x^{[n+1]} f(h)\right\|_{L_n(\mathbb{K};X)} \cdot \prod_{i=1}^{n} |1| &= \\
\left\| D^n_{x+h} f - D^n_x f - D_x^{[n+1]} f(h)\right\|_{L_n(\mathbb{K};X)} &= \\
\left\| D^n f(x+h) - D^n f(x) - D_x^{[n+1]} f(h)\right\|_{L_n(\mathbb{K};X)} &\underset{[\text{eq: } 16.69]}{=} \\
\left\| D^n f(x+h) - D^n f(x) - D_x(D^{[n]} f)(h)\right\|_{L_n(\mathbb{K};X)} &<_{[\text{eq: } 16.68]} \\
\varepsilon \cdot |h| &
\end{aligned}$$

hence we have

$$\begin{aligned}
\left\| \frac{f^{(n)}(x+h) - f^{(n)}(x)}{h} - D_x^{[n+1]} f\left(\underbrace{1:\ldots:1}_{n+1}\right)\right\| &= \\
\left\| \frac{f^{(n)}(x+h) - f^{(n)}(x) - h \cdot D_x^{[n+1]} f\left(\underbrace{1:\ldots:1}_{n+1}\right)}{h}\right\| &= \\
\frac{\left\| f^{(n)}(x+h) - f^{(n)}(x) - h \cdot D_x^{[n+1]} f\left(\underbrace{1:\ldots:1}_{n+1}\right)\right\|}{|h|} &< \\
\varepsilon &
\end{aligned}$$

proving that $f^{(n)}: V \to Y$ has a derivative at x with

$$(f^{(n)})'_x = D_x^{[n+1]} f\left(\underbrace{1:\ldots:1}_{n+1}\right)$$

So by definition f has a $(n+1)$-the derivative $f_x^{(n+1)} = (f^{(n)})'_x = D_x^{[n+1]} f\left(\underbrace{1:\ldots:1}_{n+1}\right)$ proving that

$$n+1 \in S$$

⇐. We use induction to prove this, so define

$$S = \left\{ n \in \mathbb{N} | \text{If } f \text{ has a } n\text{-the derivative at } x \text{ then } f \text{ is } n\text{-times differentiable at } x \text{ with } f_x^{(n)} = D_x^{[n]} f\left(\underbrace{1:\ldots:1}_{n}\right)\right\}$$

then we have:

$1\in S$. If f has a 1-the derivative at x then f has a derivative at x, hence by [theorem: 16.32] f is Fréchet differentiable at x [hence 1-times differentiable at x] and

$$f_x^{(1)}=f'_x=D_x f(1)=D_x^{[1]}\Big(\underbrace{1:,\ldots,:1}_{1}\Big)$$

proving that $1\in S$.

$n\in S\Rightarrow n+1\in S$. If f has a $(n+1)$-the derivative $f^{(n+1)}(x)$ at x then, as $1<n+1$, there exist a open set V with $x\in V\subseteq U$ such that $\forall y\in V$ f has a n-the derivative $f_y^{(n)}$ at y and that

$$f^{(n)}: V\to Y \text{ defined by } f^{(n)}(y)=f_y^{(n)}$$

has a derivative at x and

$$f_x^{(n+1)}=(f^{(n)})'_x \tag{16.71}$$

Let $\varepsilon\in\mathbb{R}^+$ then by the above there exist a $\delta\in\mathbb{R}^+$ such that $\forall h\in V_x$ with $0<|h|<\delta$ we have

$$\left\|\frac{f^{(n)}(x+h)-f^{(n)}(x)}{h}-f_x^{(n+1)}\right\|<\varepsilon \tag{16.72}$$

As $n\in S$ we have that f is n-times differentiable at every $y\in V$ and $f_y^{(n)}=D_y^{[n]}f\Big(\underbrace{1:\ldots:1}_{n}\Big)$. So for

$$D^{[n]}f: V\to L_n(X;Y) \text{ defined by } D^{[n]}f(y)=D_y^{[n]}f$$

we have $\forall y\in V$ that $D^{[n]}f(y)\Big(\underbrace{1:\ldots:1}_{n}\Big)=f_y^{(n)}$. Hence

$$\left\|\frac{D^{[n]}f(x+h)\Big(\underbrace{1:\ldots:1}_{n}\Big)-D^{[n]}f(x)\Big(\underbrace{1:\ldots:1}_{n}\Big)-h\cdot f_x^{(n+1)}}{h}\right\| = \left\|\frac{f^{(n)}(x+h)-f^{(n)}(x)}{h}-f_x^{(n+1)}\right\| <_{[\text{eq: } 16.72]} \varepsilon$$

After multiplying by $|h|$ and taking in account that

$$\left\|D^{[n]}f(x+0)\Big(\underbrace{1,\ldots,1}_{n}\Big)-D^{[n]}f(x)\Big(\underbrace{1,\ldots,1}_{n}\Big)-0\cdot f_x^{(n+1)}\right\|=\|0\|=0\leqslant\varepsilon\cdot|0|$$

it follows that $\forall h\in V_x$ with $|h|<\delta$ we have

$$\left\|D^{[n]}f(x+h)\Big(\underbrace{1:\ldots:1}_{n}\Big)-D^{[n]}f(x)\Big(\underbrace{1:\ldots:1}_{n}\Big)-h\cdot f_x^{(n+1)}\right\|\leqslant\varepsilon\cdot|h| \tag{16.73}$$

Define

$$L'_x:\mathbb{K}^n\to Y \text{ by } L'_x(k)=f_x^{(n+1)}\cdot\prod_{i=1}^n k_i$$

then

$$L'_x\Big(\underbrace{1,\ldots,1}_{n}\Big)=f^{(n+1)}(x) \tag{16.74}$$

and by [example: 14.188] we have that $L'_x\in L^n(\mathbb{K};X)$, define

$$L_x=\mathcal{I}_{n,\mathbb{K},X}^{-1}(L'_x)\in L_n(\mathbb{K};X) \text{ so that } h\cdot L_x\in L_n(\mathbb{K};X)$$

then we have by [theorem: 16.85] that

$$L_x\Big(\underbrace{1:\ldots:1}_{n}\Big)=L'_x\Big(\underbrace{1,\ldots,1}_{n}\Big) \underset{[\text{eq: } 16.74]}{=} f_x^{[n+1]} \tag{16.75}$$

Define

$$K_x\colon \mathbb{K} \rightarrow L^n(\mathbb{K};Y) \text{ by } K_x(h) = h \cdot L_x \tag{16.76}$$

then by [example: 14.183]

$$K_x \in L(\mathbb{K}, L_n(\mathbb{K};X)) = L_{n+1}(\mathbb{K};X)$$

Further

$$\begin{aligned}
\|D^{[n]}f(x+h) - D^{[n]}f(x) - K_x(h)\|_{L_n(\mathbb{K};X)} &\leqslant_{[\text{lemma: } 16.102]} \\
\left\|(D^{[n]}f(x+h) - D^{[n]}f(x) - K_x(h))\left(\underbrace{1:\ldots:1}_{n}\right)\right\| &\underset{[\text{lemma: } 16.79]}{=} \\
\left\|D^{[n]}f(x+h)\left(\underbrace{1:\ldots:1}_{n}\right) - D^{[n]}f(x)\left(\underbrace{1:\ldots:1}_{n}\right) - K_x(h)\left(\underbrace{1:\ldots:1}_{n}\right)\right\| &\underset{[\text{eq: } 16.76]}{=} \\
\left\|D^{[n]}f(x+h)\left(\underbrace{1:\ldots:1}_{n}\right) - D^{[n]}f(x)\left(\underbrace{1:\ldots:1}_{n}\right) - h \cdot L_x\left(\underbrace{1:\ldots:1}_{n}\right)\right\| &\underset{[\text{eq: } 16.75]}{=} \\
\left\|D^{[n]}f(x+h)\left(\underbrace{1:\ldots:1}_{n}\right) - D^{[n]}f(x)\left(\underbrace{1:\ldots:1}_{n}\right) - h \cdot f_x^{[n+1]}\right\| &<_{[\text{eq: } 16.73]} \\
\varepsilon \cdot |h|
\end{aligned}$$

proving that $D^{[n]}f$ is Fréchet differentiable at x and that $D_x(D^{[n]}f) = K_x$. Hence

$$f \text{ is } (n+1)\text{-times differentiable at } x \text{ with } D_x^{[n+1]}f = K_x$$

Further we have

$$\begin{aligned}
D_x^{[n+1]}f\left(\underbrace{1:\ldots:1}_{n+1}\right) &= K_x\left(\underbrace{1:\ldots:1}_{n+1}\right) \\
&= K_x(1)\left(\underbrace{1:\ldots:1}_{n}\right) \\
&\underset{[\text{eq: } 16.76]}{=} 1 \cdot L_x\left(\underbrace{1:\ldots:1}_{n}\right) \\
&= L_x\left(\underbrace{1:\ldots:1}_{n}\right) \\
&\underset{[\text{eq: } 16.75]}{=} f_x^{(n+1)}
\end{aligned}$$

so that we have

$$n+1 \in S$$

□

Definition 16.104. *Let $n \in \mathbb{N}$, $\langle \mathbb{K}, || \rangle$ be the normed space of real (complex) numbers based on the norm $||$, $\langle X, \|\| \rangle$ a normed space, U a open set in $\mathbb{K}$ then $f\colon U \rightarrow X$ has a n-the derivative on U if $\forall x \in U$ f has a n-the derivative at x.*

Theorem 16.105. *Let $n \in \mathbb{N}$, $\langle \mathbb{K}, || \rangle$ be the normed space of real (complex) numbers based on the norm $||$, $\langle X, \|\| \rangle$ a normed space, U a open set in X and $f\colon U \rightarrow X$ a function then the following are equivalent:*

1. *f has a n-the derivative on U.*
2. *For n we have:*

   ***n* = 1.** *$\forall x \in U$ f has a derivative at x and $f^{(1)}(x) = f'_x$.*

   1 < *n*. *$\forall x \in U$ f has a $(n-1)$-the derivative at x and the function*

   $$f^{(n-1)}\colon U \rightarrow Y \text{ defined by } (f^{(n-1)})(x) = f_x^{(n-1)}$$

   has a derivative at x with $f_x^{(n)} = (f^{(n-1)})'_x$ for every $x \in U$.

Proof.

1 ⇒ 2. For $n \in \mathbb{N}$ we must consider two cases:

***n* = 1.** Then, as $\forall x \in U$ f has a 1-derivative at x, it follows from [definition: 16.101] that f has a derivative at x with $f_x^{(1)} = f'_x$.

$1<n$. Let $x\in U$ then there exist by [definition: 16.101] a open set V_x with $x\in V_x\subseteq U$ such that $\forall y\in V_x$ f has a $(n-1)$-the derivative at y and the function

$$f_{V_x}^{(n-1)}: V_x\rightarrow Y \text{ defined by } \left(f_{V_x}^{(n-1)}\right)(y)=f_y^{(n-1)}$$

has a derivative at x with $f_x^{(n)}=\left(f_{V_x}^{(n-1)}\right)'_x$. If $x\in U$ then as $x\in V_x\subseteq U$ it follows that f has a $(n-1)$-the derivative at x and we can define

$$f^{(n-1)}: U\rightarrow Y \text{ by } (f^{(n-1)})(x)=f_x^{(n-1)}$$

Let $y\in V_x$ then $(f^{(n-1)})_{|V_x}(y)=f^{(n-1)}(y)=f_y^{(n-1)}=f_{V_x}^{(n-1)}(y)$ so that $(f^{(n-1)})_{|V_x}=f_{V_x}^{(n-1)}$, hence $(f^{(n-1)})_{|V_x}$ has a derivative at x. So by [theorem: 16.33] $f^{(n-1)}$ has a derivative at x and $(f^{(n-1)})'_x=((f^{(n-1)})_{|V_x})'_x=\left(f_{V_x}^{(n-1)}\right)'_x=f_x^{(n)}$.

$2\Rightarrow 1$. For $n\in\mathbb{N}$ we must consider two cases:

$n=1$. As $\forall x\in U$ f has a derivative at x it follows by [definition: 16.101] that f has a 1-the derivative at x, hence f has a 1-the derivative on U.

$1<n$. Let $x\in U$ then by taking $V=U$ so that $x\in V\subseteq U$ we have by the hypothesis combined with [definition: 16.101] that f has a n-the derivative at x. Hence as $x\in U$ was chosen arbitrary f has a n-the derivative on U. □

16.2.2.5 Properties of higher order differentiation

Theorem 16.106. *Let $n\in\mathbb{N}$, $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$ normed spaces, U a open set in X, $x\in U$ and*

$$f: U\rightarrow Y$$

a function that is n-times Fréchet differentiable at x then $\forall m\in\{1,\ldots,n\}$ f is m-times differentiable at x.

Proof. We proceed by induction so define

$$S_n=\{k\in\mathbb{N}_0|\text{If } k<n \text{ then } f \text{ is } (n-k)\text{-times Fréchet differentiable at } x\}$$

then we have:

$0\in S_n$. Then $n-k=n-0=n$ which as f is n-times Fréchet differentiable at x prove that $0\in S_n$.

$k\in S_n\Rightarrow k+1\in S_n$. If $k+1<n$ we have $k<n$ so as $k\in S_n$ it follows that f is $(n-k)$-times differentiable at x. As $k+1<n\Rightarrow 1<n-k$ there exist a open V such that $x\in V\subseteq U$ and $\forall y\in V$ f is $((n-k)-1)$-times differentiable at y, in particular f is $((n-k)-1)$-times differentiable at x. As $n-(k+1)=(n-k)-1$ it follows that $k+1\in S_n$.

By mathematical induction it follows that $S_n=\mathbb{N}_0$, hence if $m\in\{1,\ldots,n\}$ then $k=n-m\in\mathbb{N}_0=S_n$ and $k=n-m<n$ so that f is $(n-k)$-times differentiable at x or as $n-k=n-(n-m)=m$ f is m-times differentiable at x. □

We have a similar theorem for differentiable classes.

Theorem 16.107. *Let $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$ be two normed spaces, a open set $U\subseteq X$, $n\in\mathbb{N}_0$ and*

$$f: U\rightarrow Y \text{ a function of class } C^n$$

then $\forall m\in\{0,\ldots,n\}$ we have that f is of class C^m.

Proof. For n we have either:

$n=0$. Then $\{0,\ldots,n\}=\{0\}$, hence if $m\in\{0,\ldots,n\}$ we have $m=n=0$ which as f is of class C^n proves that f is of class C^m.

$0<n$. Then for $m\in\{0,\ldots,n\}$ we have either:

$m=n$. Then as f is of class C^n we have that f is of class C^m.

$m<n$. Then for m we have either:

$m=0$. Let $x\in U$ then, as f is of class C^n, it follows that f is n-times differentiable at x. Hence using [theorem: 16.106] we have that f is 1-times differentiable at x or by definition f is Fréchet differentiable at x. Applying then [theorem: 16.31] proves that f is continuous at x. Hence f is continuous proving that f is of class $C^0=C^m$.

$1\leqslant m$. As $1\leqslant m<n$ we have that $m+1\in\{1,\ldots,n\}$ hence it follows from [theorem: 16.106] that f is $(m+1)$-times differentiable on U. Hence by [theorem: 16.92] it follows that f is m-times differentiable on U and

$$D^{[m]}f: U\rightarrow L_m(X;Y) \text{ defined by } D^{[m]}f(y)=D_y^{[m]}f$$

is Fréchet differentiable on U. Further using [theorem: 16.31] it follows that $D^{[m]}f$ is continuous. Hence it follows that f is of class C^m. □

TODO checked this file up to here, check the rest note $U_{[i,x]}=(I^{[i,x]})^{-1}(U)$ and $f^{[i,x]}=f\circ(I^{[i,x]})_{U_{[i,x]}}$

Theorem 16.108. *Let $n,m\in\mathbb{N}$, $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$, U a open set in X and $f:U\to Y$ a function of class C^{n+m} [hence by [theorem: 16.107] is also of class n] so that the following functions are well defined*

$$D^{[n+m]}f:U\to L_{n+m}(X;Y) \text{ where } D^{[n+m]}f(x)=D_x^{[n+m]}f$$

$$D^{[n]}f:U\to L_n(X;Y) \text{ where } D^{[n]}f(x)=D_x^{[n]}f$$

then $D^{[n]}f$ is of class C^m and for

$$D^{[m]}(D^{[n]}f):U\to L_m(X;L_n(X;Y)) \underset{[\text{lemma: }16.71]}{=} L_{n+m}(X;Y) \text{ where } D^{[m]}(D^{[n]}f)(x)=D_x^{[m]}(D^{[n]}f)$$

we have $D^{[m]}(D^{[n]}f)=D^{[n+m]}f$.

Proof. We prove this by induction on m, so let $n\in\mathbb{N}$

$S_n=\left\{m\in\mathbb{N}|\text{If } f:U\to Y \text{ is of class } C^{n+m} \text{ then } D^{[n]}f:U\to L_n(X;Y) \text{ defined by } D^{[n]}f(x)=D_x^{[n]}f \text{ is of class } C^m \text{ and for } D^{[n+m]}f:U\to L_{n+m}(X;Y) \text{ defined by } D^{[n+m]}f(x)=D_x^{[n+m]}f \text{ we have } D^{[n+m]}f=D^{[m]}(D^{[n]}f) \text{ where } D^{[m]}(D^{[n]}f):U\to L_m(X;L_n(X;Y)) \underset{[\text{lemma: }16.71]}{=} L_{n+m}(X;Y) \text{ is defined by } D^{[m]}(D^{[n]}f)(x)=D_x^{[m]}(D^{[n]}f)\right\}$

then we have:

$1\in S_n$. Let f be of class C^{n+1} then by [theorem: 16.107] f is also of class C^m and the following functions are well defined and continuous.

$$D^{[n+1]}f:U\to L_{n+1}(X;Y) \text{ where } D^{[n+1]}f(x)=D_x^{[n+1]}f$$

$$D^{[n]}f:U\to L_n(X;Y) \text{ where } D^{[n]}f(x)=D_x^{[n]}f$$

Further $\forall x\in U$ we have that $D^{[n]}f$ is Fréchet differentiable at x [hence 1-times differentiable at x] and $D_x^{[n+1]}f=D_x(D^{[n]}f)=D_x^{[1]}(D^{[n]}f)$ so that

$$D^{[1]}(D^{[n]}f):U\to L_{n+1}(X;Y) \text{ where } D^{[1]}(D^{[n]}f)(x)=D_x(D^{[n]}f)$$

is well defined. Let $x\in U$ then $D^{[1]}(D^{[n]}f)(x)=D_x(D^{[n]}f)=D_x^{[n+1]}f=D^{[n+1]}f(x)$ proving that

$$D^{[n+1]}f=D^{[1]}(D^{[n]}f)$$

As $D^{[n+1]}f$ is continuous it follows that $D^{[1]}(D^{[n]}f)$ is continuous proving that

$$D^{[n]}f \text{ is of class } C^1$$

So we must conclude that $1\in S_n$.

$m\in S_n\Rightarrow m\in S_n$. If f is of class $C^{n+(m+1)}$ then by definition f is $(n+m+1)$-times differentiable on U and the function

$$D^{[n+m+1]}f:U\to L_{n+m+1}(X;Y) \text{ defined by } D^{[n+m+1]}f(x)=D_x^{[n+m+1]}f$$

is continuous. As f is $(n+m+1)$-differentiable on U and $1<n+m+1$ we have by [theorem: 16.92] that f is $(n+m)$-times differentiable on U and that

$$D^{[n+m]}f:U\to L_n(X;Y) \text{ defined by } D^{[n+m]}f(x)=D_x^{[n+m]}f$$

is Fréchet differentiable on U with

$$\forall x\in U\; D_x^{[n+m+1]}f=D_x(D^{[n+m]}f) \tag{16.77}$$

Further by [theorem: 16.31] we have that $D^{[n+m]}f$ is continuous so that f is of class C^{n+m}. As $m\in S_n$ we have that

$$D^{[n]}f:U\to L_n(X;Y) \text{ defined by } D^{[n]}f(x)=D_x^{[n]}f$$

is of class C^m and for the function

$$D^{[m]}(D^{[n]}f):U\to L_m(X;L_n(X;Y))=L_{n+m}(X;Y) \text{ where } D^{[m]}(D^{[n]}f)(x)=D_x^{[m]}(D^{[n]}f)$$

we have that

$$D^{[n+m]}f = D^{[m]}(D^{[n]}f) \tag{16.78}$$

As $D^{[n+m]}f$ is Fréchet differentiable on U it follows from the above that $D^{[m]}(D^{[n]}f)$ is Fréchet differentiable on U so that

$$D^{[n]}f \text{ is } (m+1)\text{-times differentiable on } U \tag{16.79}$$

and

$$\forall x \in U \; D_x^{[m+1]}(D^{[n]}f) = D_x(D^{[m]}(D^{[n]}f)) \tag{16.80}$$

So for $x \in U$ we have

$$\begin{aligned} D_x^{[m+1]}(D^{[n]}f) &\underset{[\text{eq: } 16.80]}{=} D_x(D^{[m]}(D^{[n]}f)) \\ &\underset{[\text{eq: } 16.78]}{=} D_x(D^{[n+m]}f) \\ &\underset{[\text{eq: } 16.77]}{=} D_x^{[n+m+1]}f \\ &= D^{[n+m+1]}f(x) \end{aligned} \tag{16.81}$$

so that for

$$D^{[m+1]}(D^{[n]}f): U \to L_{m+1}(X; L_n(X;Y)) = L_{n+m+1}(X;Y) \text{ where } D^{[m+1]}(D^{[n]}f)(x) = D_x^{[m+1]}(D^{[n]}f)$$

we have that $D^{[m+1]}(D^{[n]}f)(x) = D_x^{[m+1]}(D^{[n]}f) \underset{[\text{eq: } 16.81]}{=} D^{[n+m+1]}f(x)$. Hence we have

$$D^{[n+m+1]}f = D^{[m+1]}(D^{[n]}f)$$

As $D^{[n+m+1]}f$ is continuous it follows that $D^{[m+1]}(D^{[n]}f)$ is continuous. The above together with [eq: 16.79] proves that

$$D^{[n]} \text{ is of class } C^{m+1}$$

This proves that

$$n+1 \in S_n$$

□

The following theorem will be essential in proofs by induction.

Theorem 16.109. *Let $\langle X, \|\|_X \rangle$, $\langle Y, \|\|_Y \rangle$ be two normed spaces, a open set $U \subseteq X$, $x \in U$, $n \in \mathbb{N} \setminus \{1\}$ then the following are equivalent:*

1. *f is n-times Fréchet differentiable at x*
2. *There exit a open set V with $x \in V \subseteq U$ such that $\forall y \in V$ f is 1-times Fréchet differentiable at y and*

   $$D^{[1]}f: V \to L_1(X;Y) = L(X,Y) \text{ defined by } D^{[1]}f(y) = D_y^{[1]}f$$

   is $(n-1)$-times Fréchet differentiable at x. Further we have $D_x^{[n]}f = D_x^{[n-1]}(D^{[1]}f)$.

***Note* 16.110.** *As $\forall y \in V$ $D_y^{[1]}f = D_y f$ we can say that the following are equivalent*

1. *f is n-times Fréchet differentiable at x*
2. *There exit a open set V with $x \in V \subseteq U$ such that $\forall y \in V$ f is 1-times Fréchet differentiable at y and*

   $$Df: V \to L_1(X;Y) = L(X,Y) \text{ defined by } Df(y) = D_y f$$

   is $(n-1)$-times Fréchet differentiable at x. Further we have $D_x^{[n]}f = D_x^{[n-1]}(Df)$.

Proof.

1 ⇒ 2. We prove this by induction on n, so define

$S = \{n \in \{2, \ldots, \infty\} |$If $f: U \to Y$ is n-times Fréchet differentiable at $x \in U$, Uopen in X then there exist a open V with $x \in V \subseteq U$ such that $\forall y \in V$ f is 1-times Fréchet differentiable at y and $D^{[1]}f: V \to L_1(X;Y)$ defined by $(D^{[1]}f)(y) = D^{[1]}f(x)$ is $(n-1)$-times Fréchet differentiable at x. Further we have $D_x^{[n]}f = D_x^{[n-1]}(D^{[1]}f)\}$

then we have:

$2\in S$. If $f:U\rightarrow Y$ is 2-times Fréchet differentiable at $x\in U$, U open in X then there exist a open set $V\subseteq X$ with $x\in V\subseteq U$ such that $\forall y\in V$ f is 1-times Fréchet differentiable at y and

$$D^{[1]}f: V\rightarrow L_1(X;Y) \text{ defined by } D^{[1]}f(y)=D^{[1]}_y f$$

is Fréchet differentiable at x and

$$D^{[2]}_x f=D_x(D^{[1]}f)$$

Hence we have by definition that $D^{[1]}f: V\rightarrow L_1(X;Y)$ is 1-times Fréchet differentiable at x and that $D^{[2]}f(x)=D^{[1]}_x(D^{[1]}f)$ proving that $2\in S$.

$n\in S\Rightarrow n+1\in S$. If f is $(n+1)$-times Fréchet differentiable at x it follows that there exist a open V in X with $x\in V\subseteq U$ such that $\forall y\in V$ f is n-times Fréchet differentiable and the function

$$D^{[n]}f: V\rightarrow L_n(X;Y) \text{ defined by } D^{[n]}f(y)=D^{[n]}_y \text{ is Fréchet differentiable at } x \tag{16.82}$$

and

$$D^{[n+1]}_x f=D_x(D^{[n]}f) \tag{16.83}$$

As $n\in S$ we have that $\forall y\in V$ there exist a open set W_y with $y\in W_y\subseteq V$ such that $\forall z\in W_y$ f is 1-times differentiable at y and the function

$$D^{[1]}_{W_y}f: W_y\rightarrow L_1(X;Y) \text{ defined by } D^{[1]}_{W_y}f(z)=D^{[1]}_z f \tag{16.84}$$

is $[n-1]$-times differentiable at y and

$$D^n_y f=D^{[n-1]}_y\big(D^{[1]}_{W_y}f\big) \tag{16.85}$$

As $\forall y\in V$ we have $y\in W_y\subseteq V$ it follows that

$$V=\bigcup_{y\in V} W_y$$

As $\forall y\in V$ we have $y\in W_y$ so that f is 1-times differentiable at y, we can define the function

$$D^{[1]}f: V\rightarrow L_n(X;Y) \text{ where } D^{[1]}f(y)=D^{[1]}_y f$$

Then, as $\forall z\in W_y$ we have $D^{[1]}_{W_y}f(z)\underset{\text{[eq: 16.84]}}{=}D^{[1]}_y f=D^{[1]}f(y)$, it follows that

$$\forall y\in V\; D^{[1]}_{W_y}f=(D^{[1]}f)_{|W_y}$$

Hence, as $D^{[1]}_{W_y}f$ is $(n-1)$-times differentiable at y with $D^{[n-1]}_y\big(D^{[1]}_{W_y}f\big)\underset{\text{[eq: 16.85]}}{=}D^{[n]}_y f$, it follows that $(D^{[1]}f)_{|W_y}$ is $(n-1)$-times differentiable at y and

$$D^{[n-1]}_y(D^{[1]}f)_{|W_y}=D^{[n]}_y f$$

Using the locality of higher order differentiation [see theorem: 16.90] it follows that $\forall y\in V\, D^{[1]}f$ is $(n-1)$-times differentiable at y and $D^{[n-1]}_y(D^{[1]}f)=D^{[n]}_y f$. Hence if we define

$$D^{[n-1]}(D^1 f): V\rightarrow L_{n-1}(X;L_1(X;Y))\underset{\text{[corollary: 16.72]}}{=}L_n(X;Y) \text{ by } (D^{[n-1]}(D^1 f))(y)=D^{[n-1]}_y(D^{[1]}f)$$

it follows that $D^{[n-1]}(D^1 f)=D^{[n]}f$. So using [eqs: 16.82 and 16.83] it follows that $D^{[n-1]}(D^1 f)$ is Fréchet differentiable at x and

$$D^{[n+1]}_x f=D_x(D^{[n-1]}(D^{[1]}f)).$$

Hence we have by definition that $D^1 f$ is n-times differentiable at x with $D^{n+1}_x f=D^{[n]}_x(D^{[1]}f)$. This proves as $(n+1)-1=n$ that

$$n+1\in S$$

$2\Rightarrow 1$. We use recursion to prove this. So define

$S=\{n\in\{2,\ldots,\infty\}|$If for $f:U\rightarrow Y$ there exist a open V with $x\in V\subseteq U$ such that $\forall y\in V$ f is 1-times differentiable at y and $D^{[1]}f: V\rightarrow L_1(X;Y)$ defined by $D^{[1]}f(y)=D^{[1]}_y f$ is $(n-1)$-times differentiable at x then f is n-times differentiable at x and $D^{[n-1]}_x(D^{[1]}f)=D^{[n]}_x f\}$

then we have:

$2\in S$. Let V be a open set with $x\in V\subseteq U$ such that $\forall y\in V$ we have that f is 1-times differentiable at y and

$$D^{[1]}f: V\rightarrow L_1(X;Y) \text{ defined by } D^{[1]}f(y)=D^{[1]}_y f$$

is $2-1=1$-times differentiable at x. Then $D^{[1]}f$ is Fréchet differentiable at x and $D^{[1]}_x(D^{[1]}f)=D_x(D^{[1]}f)$ hence by definition f is 2-times differentiable at x and

$$D^{[2]}_x f=D_x(D^{[1]}f)=D^{[1]}_x(D^{[1]}f)(x)$$

So we have that $2\in S$.

$n\in S\Rightarrow n+1\in S$. First as $n\in S\subseteq\{2,\ldots,\infty\}$ it follows that $1<n$. Let V be a open set such that $x\in V\subseteq U$ and $\forall y\in V$ f is 1-times differentiable at y and

$$D^{[1]}f: V\rightarrow L_1(X;Y) \text{ defined by } (D^{[1]}f)(y)=D^{[1]}f(y)$$

is n-times differentiable at x. Then, as $1<n$, there exist a open W with $x\in W\subseteq V$ such that $\forall y\in W$ $D^{[1]}f$ is $(n-1)$-times differentiable at y and

$$D^{[n-1]}(D^{[1]}f): W\rightarrow L(X,L_{n-1}(X;Y)) \underset{[\text{corollary: } 16.72]}{=} L_n(X;Y) \text{ defined by } D^{[n-1]}(D^{[1]}f)(y)=D^{[n-1]}_y(D^{[1]}f)$$

is Fréchet differentiable at x and

$$D^{[n]}_x(D^{[1]}f)=D_x(D^{[n-1]}(D^{[1]}f)) \tag{16.86}$$

As $\forall y\in W\subseteq V$ f is 1-times differentiable and $D^{[1]}f$ is $(n-1)$-times differentiable at y we have, as $n\in S$, that f is n-times differentiable at y and

$$D^{[n]}_y f=D^{[n-1]}_y(D^{[1]}f) \tag{16.87}$$

Hence we have for

$$D^{[n]}f: W\rightarrow L_n(X;Y) \text{ defined by } D^{[n]}f(y)=D^{[n]}_y f$$

that $\forall y\in W$ $D^{[n]}f(y)=D^{[n]}_y f \underset{[\text{eq: } 16.87]}{=} D^{[n-1]}_y(D^{[1]}f)=D^{[n-1]}(D^{[1]}f)(y)$ proving that

$$D^{[n]}f=D^{[n-1]}(D^{[1]}f) \tag{16.88}$$

so that $D^{[n]}f$ if Fréchet differentiable at x with

$$D_x(D^{[n]}f) \underset{[\text{eq: } 16.88]}{=} D_x(D^{[n-1]}(D^{[1]}f)) \underset{[\text{eq: } 16.86]}{=} D^{[n]}_x(D^{[1]}f)$$

Hence we have by definition that f is $(n+1)$-times differentiable at x with

$$D^{[n+1]}_x f=D_x(D^{[n]}f)=D^{[n]}(D^{[1]}f)(x)$$

which proves that $n+1\in S$. □

Corollary 16.111. *Let $n\in\mathbb{N}$, $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$ be two normed spaces, a open set $U\subseteq X$ and $f: U\rightarrow Y$ a function then we have*

$$f \text{ is of class } C^n$$
$$\Updownarrow$$
$$\forall x\in U\ f \text{ is 1-times differentiable at } x \text{ and } D^{[1]}f: U\rightarrow L_1(X;Y) \text{ defined by } D^{[1]}f(x)=D^{[1]}_x f \text{ is of class } C^{[n-1]}$$

Proof.

$\Rightarrow$. For $n\in\mathbb{N}$ we have two cases to consider:

$n=1$. As f is of class C^1 we have that $\forall x\in U$ f is 1-times differentiable at x and

$$D^{[1]}f: U\rightarrow L_1(X;Y) \text{ defined by } D^{[1]}f(x)=D^{[1]}_x f$$

is continuous. Hence $D^{[1]}f$ is of class $C^0=C^{1-1}$.

$1<n$. As f is of class C^n we have that $\forall x\in U$ f is n-times differentiable at x and

$$D^{[n]}f: U\rightarrow L_n(X;Y) \text{ defined by } D^{[n]}f(x)=D^{[n]}_x f$$

is continuous. Using the previous theorem [theorem: 16.109] it follows that $\forall x\in U$ there exist a open set V_x with $x\in V_x\subseteq U$ such that $\forall y\in V_x$ f is 1-times differentiable at y and that

$$D^{[1]}_{V_x}f: V_x\rightarrow L_n(X;Y) \text{ defined by } D^{[1]}_{V_x}f(y)=D^{[1]}_y f \tag{16.89}$$

is $(n-1)$-times differentiable at x with $D^{[n-1]}_x\big(D^{[1]}_{V_x}f\big)=D^{[n]}_x f$. Hence as $x\in V_x\subseteq U$ we have that

$$\forall x\in U f \text{ is 1-times differentiable at } x$$

allowing us to define

$$D^{[1]}f: U\rightarrow L_n(X;Y) \text{ by } D^{[1]}f(y)=D^{[1]}_y f \tag{16.90}$$

From [eq: 16.89, 16.90] it follows then that $\forall y\in V_x$ $D^{[1]}_{V_x}f(y)=D^{[1]}f(y)$ so that

$$\forall x\in U \text{ we have } D^{[1]}_{V_x}f=(D^{[1]}f)_{|V_x}$$

As $D^{[1]}_{V_x}f$ is $(n-1)$-times differentiable at x with $D^{[n-1]}_x\big(D^{[1]}_{V_x}f\big)=D^{[n]}_x f$ it follows from the above that $(D^{[1]}f)_{|V_x}$ is $(n-1)$-times differentiable at x and

$$D^{[n-1]}_x(D^{[1]}f)_{|V_x}=D^{[n]}_x f$$

Hence using [theorem: 16.90] we have that

$$D^{[1]}f \text{ is } (n-1)\text{-times differentiable with } D^{[n-1]}_x(D^{[1]}f)=D^{[n]}_x f\ \forall x\in U$$

So for

$$D^{[n-1]}(D^{[1]}f): U\rightarrow L_n(X;Y) \text{ defined by } D^{[n-1]}(D^{[1]}f)(x)=D^{[n-1]}_x(D^{[1]}f)$$

we have that $D^{[n-1]}(D^{[1]}f)=D^n f$. Hence, as $D^n f$ is continuous, $D^{[n-1]}(D^{[1]}f)$ is continuous proving that

$$D^1 f \text{ is of class } C^{n-1}$$

$\Leftarrow$. For $n\in\mathbb{N}$ we have two possible cases:

$n=1$. As $\forall x\in U$ f is 1-times differentiable at x and

$$D^{[1]}f: U\rightarrow L_1[X;Y]$$

is of class C^0 it follows that $D^{[1]}f$ is continuous. Hence f is of class C^1.

$1<n$. By the hypothesis we have that $\forall x\in U$ f is 1-times differentiable at x and

$$D^{[1]}f: U\rightarrow L_n(X;Y) \text{ defined by } D^{[1]}f\,(x)=D^{[1]}_x f$$

is of class C^{n-1}. So $\forall x\in U$ $D^{[1]}f$ is $(n-1)$-times differentiable at x and

$$D^{[n-1]}(D^{[1]}f): U\rightarrow L_n(X;Y) \text{ defined by } D^{[n-1]}(D^{[1]}f)(x)=D^{[n-1]}_x(D^{[1]}f)$$

is continuous. Using [theorem: 16.109] it follows then that $\forall x\in U$ f is n-times differentiable at x and $D^n_x f=D^{[n-1]}_x(D^{[1]}f)$. This proves that for

$$D^{[n]}f: U\rightarrow L_n(X;Y) \text{ defined by } D^{[n]}f(x)=D^{[n]}_x f$$

we have $D^{[n]}f=D^{[n-1]}(D^{[1]}f)$. As $D^{[n-1]}(D^{[1]}f)$ is continuous it follows that $D^{[n]}f$ is continuous hence we have that f is of class $C^{[n]}$. □

Example 16.112. Let $\langle\mathbb{K},\|\rangle$ be the normed space of real (complex) numbers, $\langle X,\|\|\rangle$ a normed space over $\mathbb{K}$, U a open set in $\mathbb{K}$, $x,y\in X$ then $\varphi: U\rightarrow X$ where $\varphi(t)=t\cdot x+y$ is of class C^∞ and $\forall t\in U$ $D^{[1]}_t f=L$ where $L:\mathbb{K}\rightarrow X$ is defined by $L(t)=t\cdot x$. So by [theorem: 16.103] we have that φ has a derivative for every $t\in U$ with $\varphi'_t=D^{[1]}_t(1)=L(1)=x$.

Proof. Define $L:\mathbb{K}\rightarrow X$ by $L(t)=t\cdot x+y$ then by [example: 14.183] it follows that $L\in L(\mathbb{K},X)$. Let $t\in U$ then given $\varepsilon\in\mathbb{R}^+$ we have for $h\in U_x$ with $|h|<1$ that

$$\|\varphi(t+h)-\varphi(t)-L(h)\|=\|(t+h)\cdot x+y-(t\cdot x+y)-h\cdot x\|=\|h\cdot x-h\cdot x\|=\|0\|\leqslant\varepsilon\cdot|h|$$

proving that φ is Fréchet differentiable at t with $D_t\varphi=L$. Hence φ is 1-times differentiable on U with $D^{[1]}_1\varphi=L$. So if we define

$$D^{[1]}\varphi: U\rightarrow L_1(X;Y) \text{ by } D^{[1]}\varphi(t)=D^{[1]}_t\varphi$$

then $D^{[1]}\varphi = C_L$ a constant function. As a constant function is continuous it follows that φ is of class C^1. Further as a constant function is of class C^∞ [see example: 16.115] if follows that $D^1\varphi$ is of class C^∞ hence by [theorem: 16.111] it follows that φ is of class C^∞. □

Just as the Fréchet differential is linear [see theorem: 16.35] we have the same for higher order differentials and derivatives.

Theorem 16.113. *Let $n \in \mathbb{N}$, $\langle X, \|\|_X \rangle$, $\langle Y, \|\|_Y \rangle$ normed spaces, U a open set in X then we have:*

1. *If $f: U \to Y$, $g: U \to Y$ are functions such that f, g are n-times differentiable at $x \in U$ then $f + g$ is n-times differentiable at x and $D_x^{[n]}(f + g) = D_x^{[n]} f + D_x^{[n]} g$.*
2. *If $k \in \mathbb{N}$ and $\{f_i: U \to Y\}_{i \in \{1, \ldots, k\}}$ is a family of functions such that $\forall i \in \{1, \ldots, k\}$ f^i is n-times differentiable at $x \in U$ then $\sum_{i=1}^k f_i$ is n-times differentiable at and*
$$D_x^n \left(\sum_{i=1}^k f_i \right) = \sum_{i=1}^k D_x^{[n]} f$$
3. *If $\alpha \in \mathbb{K}$, $f: U \to Y$ is a function that is n-times differentiable at x then $\alpha \cdot f$ is n-times differentiable at x and $D_x^{[n]}(\alpha \cdot f) = \alpha \cdot D_x^{[n]} f$.*
4. *If $f: U \to Y$, $g: U \to Y$ are functions of class C^n then $f + g$ is of class C^n.*
5. *If $k \in \mathbb{N}$ and $\{f_i: U \to Y\}_{i \in \{1, \ldots, k\}}$ is a family of functions such that $\forall i \in \{1, \ldots, k\}$ f^i is of class C^n then $\sum_{i=1}^k f_i$ is of class C^n.*
6. *If $\alpha \in \mathbb{K}$, $f: U \to Y$ is of class C^n then $\alpha \cdot f$ is of class C^n.*

Proof.

1. We use induction on n to prove this, so let

   $S = \{n \in \mathbb{N} | \text{If } f, g: U \to Y$ are n-times differentiable at $x \in U$ where U is open in X then $f + g$ is n-times differentiable at x and $D_x^{[n]}(f + g) = D_x^{[n]} f + D_x^{[n]} g\}$

   then we have:

   $1 \in S$. If $f, g: U \to Y$ are 1-times differentiable at x then f, g are Fréchet differentiable at $x \in U$ and $D_x^{[1]} f = D_x f$, $D_x^{[1]} g = D_x g$. Using [theorem: 16.35] it follows that $f + g$ is Fréchet differentiable at x [hence 1-times differentiable at x] and
   $$D_x^{[1]}(f + g) = D_x(f + g) = D_x f + D_x g = D_x^{[1]} f + D_x^{[1]} g$$
   Hence we have $1 \in S$.

   $n \in S \Rightarrow n + 1 \in S$. If f, g are $(n+1)$-times differentiable at x then there exist open sets V_1, V_2 with $x \in V_1 \subseteq U$, $x \in V_2 \subseteq U$ such that $\forall y \in V_1$ f is n-times differentiable at y and $\forall y \in V_2$ g is n-times differentiable at x. Further we have that
   $$D^{[n]} f: V_1 \to L_n(X; Y) \text{ defined by } D^{[n]} f(y) = D_y^{[n]} f$$
   and
   $$D^{[n]} g: V_2 \to L_n(X; Y) \text{ defined by } D^{[n]} g(y) = D_y^n g$$
   are differentiable at x with $D_x^{[n+1]} f = D_x(D^{[n]} f)$ and $D_x^{[n+1]} g = D_x(D^{[n]} g)$. Take $V = V_1 \bigcap V_2$ then if $y \in V \subseteq U$ we have $\forall y \in V_1 \bigcap V_2$ that f, g are n-times differentiable at y. As $n \in S$ it follows that $f + g$ is n-times differentiable and $D_y^{[n]}(f + g) = D_y^{[n]} f + D_y^{[n]} g$. This allows us to define
   $$D^{[n]}(f + g): V \to L_n(X; Y) \text{ where } D^{[n]}(f + g)(y) = D_y^{[n]}(f + g)$$
   and $\forall y \in V$ we have that
   $$\begin{aligned} (D^{[n]}(f + g))(y) &= D_y^{[n]}(f + g) \\ &= D_y^{[n]} f + D_y^{[n]} g \\ &= D^{[n]} f(y) + D^{[n]} g(y) \\ &= (D^{[n]} f)_{|V}(y) + (D^{[n]} g)_{|V}(y) \end{aligned}$$
   proving that
   $$D^{[n]}(f + g) = (D^{[n]} f)_{|V} + (D^n g)_{|V} \tag{16.91}$$

As $D^{[n]}f$, $D^{[n]}g$ are Fréchet differentiable at x and $x \in V_1 \bigcap V_2 = V$ it follows from [theorem: 16.24] that $(D^{[n]}f)_{|V}$, $(D^{[n]}g)_{|V}$ are Fréchet differentiable at x with

$$D_x(D^{[n]}f)_{|V} = D_x(D^{[n]}f) \text{ and } D_x(D^{[n]}g)_{|V} = D_x(D^{[n]}g)$$

Hence using [theorem: 16.35] $(D^{[n]}f)_{|V} + (D^n g)_{|V}$ is Fréchet differentiable at x with

$$\begin{aligned} D_x((D^{[n]}f)_{|V} + (D^n g)_{|V}) &= D_x(D^{[n]}f)_{|V} + D_x(D^{[n]}g)_{|V} \\ &= D_x(D^{[n]}f) + D_x(D^{[n]}f) \end{aligned} \tag{16.92}$$

Which as $D^{[n]}(f+g) \underset{[\text{eq } 16.91]}{=} (D^{[n]}f)_{|V} + (D^n g)_{|V}$ proves that $f+g$ is $(n+1)$-times differentiable at x with

$$\begin{aligned} D_x^{[n+1]}(f+g) &= D_x(D^{[n]}(f+g)) \\ &= D_x((D^{[n]}f)_{|V} + (D^n g)_{|V}) \\ &= D_x(D^{[n]}f) + D_x(D^{[n]}f) \\ &= D_x^{[n+1]}f + D_x^{[n+1]}g \end{aligned}$$

proving that

$$n+1 \in S$$

2. We prove this by induction on k, so define

$S = \left\{ k \in \mathbb{N} | \text{If } \{f_i | U \rightarrow Y\}_{i \in \{1,\ldots,k\}} \text{ is a family of } n\text{-times differentiable functions at } x \text{ then } \sum_{i=1}^{k} f_i \text{ is } n\text{-times differentiable at } x \right.$
$\left. \text{with } D_x^{[n]}\left(\sum_{i=1}^{k} f_i\right) = \sum_{i=1}^{k} D_x^{[n]} f \right)$

then we have:

$1 \in S$. If $\{f_i | U \rightarrow Y\}_{i \in \{1\}}$ is a family of n-times differentiable at x then as $\sum_{i=1}^{1} f_i = f_1$ it follows that $\sum_{i=1}^{1} f_i$ is n-times differentiable at x and

$$D_z^{[n]}\left(\sum_{i=1}^{1} f_i\right) = D_x^{[n]} f_1 = \sum_{i=1}^{1} D_x^{[n]} f_i$$

proving that $1 \in S$.

$k \in S \Rightarrow k+1 \in S$. Let $\{f_i: U \rightarrow Y\}_{i \in \{1,\ldots,k+1\}}$ be a family of n-times differentiable functions at x then we have as $k \in S$ that $\sum_{i=1}^{k} f_i$ is n-times differentiable at x with $D_x^{[n]}(\sum_{i=1}^{k} f_i) = \sum_{i=1}^{k} D_x^{[n]} f_i$. As also f_{k+1} is n-times differentiable at x it follows from (1) that $\sum_{i=1}^{k+1} f_i = \sum_{i=1}^{k} f_i + f_{k+1}$ is n-times differentiable at x with

$$\begin{aligned} D_x^{[n]}\left(\sum_{i=1}^{k+1} f_i\right) &= D_x^{[n]}\left(\sum_{i=1}^{k} f_i + f_{k+1}\right) \\ &\underset{(1)}{=} D_x^{[n]}\left(\sum_{i=1}^{k} f_i\right) + D_x^{[n]} f_{k+1} \\ &\underset{n \in S}{=} \sum_{i=1}^{k} D_x^{[n]} f_i + D_x^{[n]} f_{k+1} \\ &= \sum_{i=1}^{k+1} D_x^{[n]} f_i \end{aligned}$$

proving that

$$n+1 \in S$$

3. We use induction on n, so let

$S = \{ n \in \mathbb{N} | \text{If } f: U \rightarrow Y \text{ is } n\text{-times differentiable at } x \in U \text{where } U \text{ is open in } X \text{ then } \alpha \cdot f \text{ is } n\text{-times differentiable at } x \text{ and}$
$D_x^{[n]}(\alpha \cdot f) = \alpha \cdot D_x^{[n]} f \}$

then we have:

$1\in S$. If $f:U\rightarrow Y$ is 1-times differentiable at x then f is Fréchet differentiable at x and $D_x^{[1]}f=D_xf$. Using [theorem: 16.35] it follows that $\alpha\cdot f$ is Fréchet differentiable at x and $D_x(\alpha\cdot f)=\alpha\cdot D_xf$. Hence $\alpha\cdot f$ is 1-times differentiable and

$$D_x^1(\alpha\cdot f)=D_x(\alpha\cdot f)=\alpha\cdot D_xf=\alpha\cdot D_x^{[1]}f$$

proving that $1\in S$.

$n\in S\Rightarrow n+1\in S$. If $f:U\rightarrow Y$ is $(n+1)$-times differentiable at x then, as $1<n+1$, there exist a open set V with $x\in V\subseteq U$ such that $\forall y\in V$ f is n-times differentiable at y and

$$D^nf:V\rightarrow L_n(X;Y) \text{ defined by } D^{[n]}f(y)=D_y^{[n]}f$$

is Fréchet differentiable at x, further $D_x^{[n+1]}f$ is defined to be $D_x(D^{[n]}f)$. As $n\in\mathbb{N}$ it follows that $\forall y\in V$ $\alpha\cdot f$ is n-times differentiable at y and

$$D_y^{[n]}(\alpha\cdot f)=\alpha\cdot D_y^{[n]}f \tag{16.93}$$

So if we define

$$D^{[n]}(\alpha\cdot f):V\rightarrow L_n(X;Y) \text{ by } D^n(\alpha\cdot f)(y)=D_y^{[n]}(\alpha\cdot f)$$

we have $\forall y\in V$ that $D^n(\alpha\cdot f)(y)=D_y^{[n]}(\alpha\cdot f)\underset{\text{[eq: 16.93]}}{=}\alpha\cdot D_y^{[n]}f=\alpha\cdot D^{[n]}f(y)$ so that

$$D^{[n]}(\alpha\cdot f)=\alpha\cdot D_y^{[n]}f \tag{16.94}$$

As D^nf is Fréchet differentiable at x it follows form [theorem: 16.35] that $\alpha\cdot D^{[n]}f$ is Fréchet differentiable at x with $D_x(\alpha\cdot D^{[n]}f)=\alpha\cdot D_x(D^{[n]}f)=\alpha\cdot D_x^{[n+1]}f$. Hence by [eq: 16.94] $D^{[n]}(\alpha\cdot f)$ is Fréchet differentiable at x with

$$D_x(D^{[n]}(\alpha\cdot f))=\alpha\cdot D_x^{[n+1]}f$$

proving that $\alpha\cdot f$ is $(n+1)$-times differentiable at x with

$$D_x^{[n+1]}(\alpha\cdot f)\underset{\text{def}}{=}D_x(D^{[n]}(\alpha\cdot f))=\alpha\cdot D_x^{[n+1]}f$$

So we conclude that

$$n+1\in S$$

4. If $f:U\rightarrow Y$ and $g:U\rightarrow Y$ are of class C^n then we have two cases to consider:

$n=0$. Then f,g are continuous functions, applying [theorem: 14.144] proves that $f+g$ is continuous. Hence $f+g$ is of class C^0.

$1<n$. Then $\forall y\in U$ f, g are n-times differentiable at y and

$$D^{[n]}\,f:U\rightarrow L_n(X;Y) \text{ defined by } D^{[n]}f(y)=D_y^{[n]}f$$

$$D^{[n]}\,g:U\rightarrow L_n(X;Y) \text{ defined by } D^{[n]}g(y)=D_y^{[n]}g$$

are continuous functions. Using (1) it follows that $\forall y\in U$ $f+g$ is n-times differentiable at y and that $D_y^{[n]}(f+g)=D_y^{[n]}f+D_y^{[n]}g$. So if we define

$$D^{[n]}(f+g):U\rightarrow L_n(X;Y) \text{ by } D^{[n]}(f+g)(y)=D_y^{[n]}(f+g)$$

it follows that $\forall y\in U$

$$\begin{aligned} D^{[n]}(f+g)(y) &= D_y^{[n]}(f+g)\\ &= D_y^{[n]}f+D_y^{[n]}g\\ &= D^{[n]}f(y)+D^{[n]}g(y)\\ &= (D^{[n]}f+D^{[n]}g)(y)\end{aligned}$$

so that

$$D^{[n]}(f+g)=D^{[n]}f+D^{[n]}g$$

which by [theorem: 14.144 is continuous. Hence $f+g$ is of class C^n.

5. If $\{f_i: U \rightarrow Y\}_{i \in \{1, \ldots, k\}}$ is a family of functions of class C^n then for n we have either:

   $n = 0$. Then we have that $\forall i \in \{1, \ldots, n\}$ f_i is continuous. Using [theorem: 14.145] we can conclude that $\sum_{i=1}^n f_i$ is continuous hence of class C^0.

   $1 < n$. As $\forall i \in \{1, \ldots, k\}$ f_i is of class C^n it follows that $\forall y \in U$ f_i is n-times differentiable and that

   $$D^{[n]} f_i: U \rightarrow L_n(X; Y) \text{ defined by } D^{[n]} f(y) = D_y^{[n]} f$$

   is continuous. Using (3) follows that $\sum_{i=1}^k f_i$ is n-times differentiable at $y \in U$ and $D_y^{[n]}(\sum_{i=1}^k f_i) = \sum_{i=1}^k D_y^{[n]} f_i$. So if we define

   $$D^{[n]}\left(\sum_{i=1}^k f_i\right): U \rightarrow L_n(X; Y) \text{ by } D^{[n]}\left(\sum_{i=1}^k f_i\right)(y) = D_y^{[n]}\left(\sum_{i=1}^k f_i\right)$$

   then $\forall y \in U$

   $$D^n\left(\sum_{i=1}^k f_i\right)(y) = D_y^{[n]}\left(\sum_{i=1}^k f_i\right) = \sum_{i=1}^k D_y^{[n]} f_i = \sum_{i=1}^k D^{[n]} f_i(y) = \left(\sum_{i=1}^k D^{[n]} f_i\right)(y)$$

   proving that

   $$D^{[n]}\left(\sum_{i=1}^k f_i\right) = \sum_{i=1}^k D^{[n]} f_i$$

   which is continuous by [theorem: 14.145]. Hence it follows that $\sum_{i=1}^k f_i$ is of class C^n.

6. As f is of class C^n it follows that $\forall y \in U$ f is n-times differentiable at y and that the function

   $$D^{[n]} f: U \rightarrow L_n(X; Y) \text{ defined by } D^{[n]} f(y) = D_y^{[n]} f$$

   is continuous. Using (3) we have that $\forall y \in U$ $(\alpha \cdot f)$ is n-times differentiable at y and $D^{[n]}(\alpha \cdot f)(y) = \alpha \cdot D^{[n]} f(y)$. Hence if we define

   $$D^{[n]}(\alpha \cdot f): U \rightarrow L_n(X; Y) \text{ by } D^{[n]}(\alpha \cdot f)(y) = D_y^{[n]}(\alpha \cdot f)$$

   then

   $$D^{[n]}(\alpha \cdot f) = \alpha \cdot D^{[n]} f$$

   Hence, as by [theorem: 16.35] $\alpha \cdot D^{[n]} f$ is continuous, it follows that $D^{[n]}(\alpha \cdot f)$ is continuous, proving that $\alpha \cdot f$ is of class C^n. □

Corollary 16.114. *Let $n \in \mathbb{N}$ $\langle \mathbb{K}, \| \rangle$ the normed space of real (complex) numbers using the norm $||$, $\langle X, \|\|\rangle$ a normed space, U a open set in $\mathbb{K}$ and $x \in U$ then we have:*

1. *If $f: U \rightarrow Y$, $g: U \rightarrow Y$ have a n-the derivative at x then $f + g$ has a n-the derivative at x and*

   $$(f + g)_x^{(n)} = f_x^{(n)} + g_x^{(n)}$$

2. *If $\alpha \in \mathbb{K}$ and $f: U \rightarrow Y$ has a n-the derivative at x then $\alpha \cdot f$ has a n-the derivative x and $(\alpha \cdot f)_x^{(n)} = \alpha \cdot f_x^{(n)}$.*

Proof.

1. If $f: U \rightarrow Y$, $g: U \rightarrow Y$ have a n-the derivative at x then by [theorem: 16.103] f, g are n-times differentiable at x and $f_x^{(n)} = D^{[n]} f(x)\left(\underbrace{1: \ldots 1}_{n}\right)$, $g_x^{(n)} = D^{[n]} g(x)\left(\underbrace{1: \ldots 1}_{n}\right)$. Using [theorem: 16.113] $f + g$ is n-times differentiable at x with $D_x^{[n]}(f + g) = D_x^{[n]} f + D_x^{[n]} g$. Hence by [theorem: 16.103] $f + g$ has a n-the derivative at x with

   $$\begin{aligned} (f + g)_x^{(n)} &= D_x^{[n]}(f + g)\left(\underbrace{1: \ldots : 1}_{n}\right) \\ &= \left(D_x^{[n]} f + D_x^{[n]} g\right)(1: \ldots : 1) \\ &\underset{[\text{lemma: } 16.79]}{=} D_x^{[n]} f\left(\underbrace{1: \ldots : 1}_{n}\right) + D_x^{[n]} g\left(\underbrace{1: \ldots : 1}_{n}\right) \\ &= f_x^{[n]} + g_x^{[n]} \end{aligned}$$

2. As f has a n-the derivative at x it follows by [theorem: 16.103] that f is n-times differentiable at x and that $f_x^{(n)} = D_x^{[n]} f\Big(\underbrace{1: \ldots :1}_{n}\Big)$.

   Using [theorem: 16.113] it follows that $\alpha \cdot f$ is n-times differentiable at x with $D_x^{[n]}(\alpha \cdot f) = \alpha \cdot D_x^{[n]} f$. Hence using [theorem: 16.103] again $\alpha \cdot f$ has a n-the derivative at x with

$$\begin{aligned} (\alpha \cdot f)_x^{(n)} &= D_x^{[n]}(\alpha \cdot f)\Big(\underbrace{1: \ldots :1}_{n}\Big) \\ &= \left(\alpha \cdot D_x^{[n]} f\right)\Big(\underbrace{1: \ldots :1}_{n}\Big) \\ &\underset{\text{[lemma: 16.79]}}{=} \alpha \cdot D_x^{[n]} f\Big(\underbrace{1: \ldots :1}_{n}\Big) \\ &= \alpha \cdot f_x^{[n]} \end{aligned}$$

□

16.2.2.6 Examples of ∞-times differentiable functions

We show now some example of ∞-times differentiable functions or equivalent functions of class C^∞ [see theorem: 16.98].

Example 16.115. Let $\langle X, \|\|_X \rangle$, $\langle Y, \|\|_Y \rangle$ be normed spaces, U a open set in X and $y \in Y$ then the constant function $C_y: U \rightarrow Y$ defined by $C_y(x) = y$ is ∞-times differentiable for every $x \in U$ [hence by [theorem: 16.98] C_y is of class C^∞]. Furthermore $\forall n \in \mathbb{N}$, $\forall x \in U$ we have that $D_x^{[n]} C_y = 0_n$ where $0_n \in L_n(X; Y)$ is the neutral element in $L_n(X; Y)$.

Proof. We prove this by induction on n, so define

$$S = \left\{n \in \mathbb{N} | C_y: U \rightarrow Y \text{ is } n\text{-times differentiable with } \forall x \in U\ D_x^{[n]} C_y = 0_n\right\}$$

then we have:

$1 \in S$. Let $x \in U$, $\varepsilon \in \mathbb{R}^+$ then for $h \in U_x$ with $|h| < 1$ we have that

$$\|C_y(x+h) - C_y(h) - 0_1(h)\|_Y = \|y - y - 0\|_Y = \|0\|_Y = 0 \leqslant \varepsilon \cdot \|h\|_X$$

proving that C_y is Fréchet differentiable at x with $D_x f = 0_1$. Hence by definition f is 1-times differentiable with $D_x^{[1]} f = 0_1$ proving that $1 \in S$.

$n \in S \Rightarrow n+1 \in S$. As $n \in S$ it follows that $\forall x \in U$ C_0 is n-times differentiable with

$$D_x^{[n]} C_y = 0_n \in L_n(X; Y). \tag{16.95}$$

Allowing us to define

$$D^{[n]} C_y: U \rightarrow L_n(X; Y) \text{ by } D^{[n]} C_y(x) = D_x^{[n]} C_y = 0_n$$

Let $\varepsilon \in \mathbb{R}^+$ then for $h \in U_x$ with $|h| < 1$ we have

$$\begin{aligned} \|(D^{[n]} C_y)(x+h) - (D^{[n]} C_n)(x) - 0_{n+1}(h)\|_{L_n(X;Y)} &\underset{\text{[eq: 16.95]}}{=} \|0_n - 0_n - 0_{n+1}(h)\|_{L_n(X;Y)} \\ &= \|0_n\|_{L_n(X;y)} \\ &= 0 \\ &\leqslant \varepsilon \cdot \|h\|_x \end{aligned}$$

proving that $D^{[n]} C_y$ is Fréchet differentiable at x with $D_x(D^{[n]} C_y) = 0_{n+1}$. Hence f is $(n+1)$-times differentiable at x with $D_x^{[n+1]} C_y = 0_{n+1}$. □

Example 16.116. Let $\langle X, \|\|_X \rangle$, $\langle Y, \|\|_Y \rangle$ be normed spaces and $L \in L(X, Y)$ then we have:

1. L is ∞-times differentiable [hence by [theorem: 16.98] L is of class C^∞]
2. $\forall n \in \mathbb{N}$ we have $\forall x \in X$ that

$$D_x^{[n]} L = \begin{cases} L \text{ if } n = 1 \\ 0_n \text{ if } 1 < n \end{cases}$$

Proof. For $n\in\mathbb{N}$ we have either:

$n=1$. Then by [example: 16.27] we have that L is Fréchet differentiable at $x\in X$ with $D_x f=L$. So by definition L is 1-times differentiable at x and $D^{[1]}f(x)=L$.

$1<n$. We proceed by induction, so let

$$S=\left\{n\in\{2,\ldots,\infty\}|L \text{ is n-times differentiable at } x\in X \text{ with } D_x^{[n]}L=0_n\right\}$$

then we have:

$2\in S$. As L is 1-times differentiable with $D^{[1]}L(x)=L$ for every $x\in X$ it follows that for

$$D^{[1]}L\colon X\rightarrow L_1(X;Y) \text{ defined by } D^{[1]}L(x)=D_x^{[1]}L$$

we have $D^{[1]}L=C_L\in L(X,L(X,Y))=L_2(X;Y)$. Hence using [example: 16.115] it follows that $D^{[1]}L$ is 1-times differentiable (hence Fréchet differentiable) at x where $D_x(D^{[1]}L)=D_x^{[1]}(D^{[1]}L)$ is the neutral element in $L_1(X,L_1(X;Y))=L_2(X;Y)$, or in other words $D_x(D^{[1]}L)=0_2$. So L is 2-times differentiable at x with $D^{[2]}L(x)=0_2$ proving that $2\in S$.

$n\in S\Rightarrow n+1\in S$. As $n\in S$ it follows that $\forall x\in X$ L is n-times differentiable at x and that $D^{[n]}L(x)=0_n$. Hence if we define

$$D^{[n]}L\colon X\rightarrow L_n(X;Y) \text{ by } D^{[n]}L(x)=D_x^{[n]}L(x)$$

it follows that $D^{[n]}L=C_{0_n}$. Using [example: 16.115] it follows then that $\forall x\in X$ $D^{[n]}L$ is 1-times differentiable (hence Fréchet differentiable) at x and that $D_x(D^{[n]}L)=D_x^{[1]}(D^{[n]}L)$ is the neutral element of $L_1(X;L_n(X;Y))=L_{n+1}(X;Y)$. Hence L is $(n+1)$-times differentiable at x and $D_x^{[n+1]}L=D_x(D^{[n]}L)=0_{n+1}$ proving that

$$n+1\in S$$

□

Example 16.117. Let $k\in\mathbb{N}\setminus\{1\}$, $\{\langle X_i,\|\|_i\rangle\}_{i\in\{1,\ldots,k\}}$ a finite family of normed spaces, $\langle Y,\|\|\rangle$ a normed space, $i\in\{1,\ldots,k\}$, $x\in\prod_{j\in\{1,\ldots,k\}}X_j$ then

$$I^{[i,x]}\colon X_i\rightarrow\prod_{j\in\{1,\ldots,k\}}X_j \text{ defined by } I^{[i,x]}(t)=\left(x_1,\ldots\underbrace{t}_{i}\ldots,x_n\right) \text{ [see definition: 16.39]}$$

is ∞-times differentiable at $t\in X_i$ [hence of class C^∞] Further $\forall n\in\mathbb{N}$

$$D_x^{[1]}I^{[i,x]}(x)=\begin{cases}I^{[i,0]} \text{ if n=1}\\ 0_n \text{ if } 1<n\end{cases}$$

Proof. Let $x\in\prod_{j\in\{1,\ldots,k\}}X_j$ and $i\in\{1,\ldots,k\}$. If $t\in X_i$ then

$$\begin{aligned}\left(I^{[i,0]}+C_{\left(x_1,\ldots\underbrace{0}_{i}\ldots,x_n\right)}\right)(t) &= I^{[i,0]}(t)+C_{(x_1,\ldots,x_{i-1},0,x_{i+1},\ldots,x_n)}(t)\\ &= \left(0,\ldots,\underbrace{t}_{i},\ldots,0_i\right)+(x_1,\ldots\underbrace{0}\ldots,x_n)\\ &= \left(x_1,\ldots\underbrace{0}_{i}\ldots,x_n\right)\\ &= I^{[i,x]}(t)\end{aligned}$$

so that

$$I^{[i,x]}=I^{[i,0]}+C_{\left(x_1,\ldots\underbrace{0}_{i}\ldots,x_n\right)}$$

By [lemma: 16.41] $I^{[i,0]}\in L\left(X_i,\prod_{j\in\{1,\ldots,k\}}X_j\right)$ so that by [examples: 16.116 and 16.115] $I^{[i,0]}$ and $C_{\left(x_1,\ldots\underbrace{0}_{i}\ldots,x_n\right)}$ are ∞-times differentiable at $t\in X_i$. Using [theorem: 16.113] proves then that

$$I^{[i,x]} \text{ is } \infty\text{-times differentiable at } t$$

and $\forall n\in\mathbb{N}$ we have

$$D_t^{[n]}I^{[i,x]}=D_t^{[n]}I^{[i,0]}+D_t^{[n]}C_{\left(x_1,\ldots\underbrace{0}_{i}\ldots,x_n\right)}\underset{\text{[example: 16.115]}}{=}D_t^{[n]}I^{[i,0]} \tag{16.96}$$

If $n\in\mathbb{N}$ then we have either:

$n=1$. Then by [example: 16.116] $D^{[1]}I^{[i,0]}(t)=I^{[i,0]}$

$1<n$. Then by [example: 16.116] $D^{[1]}I^{[i,0]}(t)=0_n$ □

Example 16.118. Let $\langle X,\|\|\rangle$ be a normed space, $x\in X$ then for the translation function τ_x defined by

$$\tau_x\colon X\rightarrow X \text{ where } \tau_x(y)=x+y$$

we have that:

1. τ_x is ∞-times differentiable at $y\in Y$ [hence of class C^∞]
2. $\forall n\in\mathbb{N}$, $\forall y\in X$ we have
$$D^{[n]}_y\tau_x=\begin{cases}\mathrm{Id}_X \text{ if } n=1\\ 0_n \text{ if } 1<n\end{cases}$$

Proof. Let $x\in X$ then $\forall y\in X$ we have $(\mathrm{Id}_X+C_x)(y)=\mathrm{Id}_X(y)+C_x(y)=y+x=\tau_x(y)$ proving that

$$\tau_x=\mathrm{Id}_X+C_x$$

As $\mathrm{Id}_X\in L(X,X)$ [see example: 14.172] we have by [example: 16.116] that Id_X is ∞-times differentiable, further by [example: 16.115] we have that C_x is ∞-times differentiable. Using [theorem: 16.113] it follows that τ_x is ∞-times differentiable and

$$D^{[n]}_y\tau_x=D^{[n]}_y\mathrm{Id}_X+D^{[n]}_yC_x\underset{\text{[example: 16.115]}}{=}D^{[n]}_y\mathrm{Id}_X+0_n=D^{[n]}_y\mathrm{Id}_X$$

Hence using [example: 16.116] it follows that

$$D^{[n]}_y\tau_x=\begin{cases}\mathrm{Id}_X \text{ if } n=1\\ 0_n \text{ if } 1<n\end{cases}$$ □

Example 16.119. Let $k\in\mathbb{N}$, $\{\langle X_i,\|\|_i\rangle\}_{i\in\{1,\ldots,k\}}$ be a finite family of normed spaces, $i\in\{1,\ldots,k\}$ then the projection map $\pi_i\colon\prod_{j\in\{1,\ldots,k\}}X_j\rightarrow X_i$ is ∞-times differentiable at $t\in X_i$ [hence of class C^∞] and $\forall n\in\mathbb{N}$ we have

$$D^{[n]}_t\pi_i=\begin{cases}\pi_i \text{ if } n=1\\ 0_n \text{ if } 1<n\end{cases}$$

Proof. As by [theorem: 14.184] $\pi_i\in L\left(\prod_{j\in\{1,\ldots,k\}}X_j,X_i\right)$ the proof follows from [example: 16.116]. □

16.2.2.7 Properties of higher order differentiation

Theorem 16.120. *Let $\langle X_1,\|\|_1\rangle$, $\langle X_2,\|\|_2\rangle$ and $\langle Y,\|\|\rangle$ be normed spaces, $\langle X_1\cdot X_2,\|\|_{\max}\rangle$ the normed space based on $X_1\cdot X_2$ using the maximum norm and $L\in L(X_1,X_2;Y)$ be a bi-linear mapping then $\forall x=(x_1,x_2)\in X_1\cdot X_2$ we have*

1. *L is ∞-times differentiable at x [hence L is of class C^∞]*
2. *$D^{[1]}_xL\underset{x=(x_1,x_2)}{=}D^{[1]}_{(x_1,x_2)}L=L(x_1,*)+L(*,x_2)$ where*
$$L(x_1,*)+L(*,x_2)\colon X_1\cdot X_2\rightarrow Y \text{ is defined by } (L(x_1,*)+L(*,x_2))(r,s)=L_1(x_1,s)+L(r,x_2)$$
3. *$D^{(2)}_xL\underset{x=(x_1,x_2)}{=}D^{[2]}_{(x_1,x_2)}L=D^{[1]}L$ where*
$$D^{[1]}L\colon X_1\cdot X_2\rightarrow L_2(X_1\cdot X_2,L(X_1\cdot X_2;Y)) \text{ is defined by } (D^{[1]}L)(x_1,x_2)=D^{[1]}_{(x_1,x_2)}L$$
4. *If $n\in\{3,\ldots.\infty\}$ then $D^{[n]}L(x_1,x_2)=0_n$*

Proof. For $n\in\mathbb{N}$ we have either:

$n=1$. Using [lemma: 16.64] L is Fréchet differentiable at (x_1,x_2) with

$$D_{(x_1,x_2)}L=L(x_1,*)+L(*,x_2)$$

hence L is 1-times differentiable and $D^{[1]}_{(x_1,x_2)}L=L(x_1,*)+L(*,x_2)$

$n=2$. Define

$$D^{[1]}L: X_1 \cdot X_2 \rightarrow L(X_1 \cdot X_2, Y) \text{ by } D^{[1]}L(x_1, x_2) = D^{[1]}_{(x_1,x_2)}L = L(x_1, *) + L(*, x_2)$$

If $(x_1, x_2), (y_1, y_2) \in X_1 \cdot X_2,, \alpha \in \mathbb{K}$ then for every $(r, s) \in X_1 \cdot X_2$ we have

$$\begin{array}{rcl} D^{[1]}L((x_1,x_2)+\beta\cdot(y_1,y_2))(r,s) & = & D^{[1]}L(x_1+\beta\cdot y_1, x_2+\beta\cdot y_2)(r,s) \\ & = & (L(x_1+\beta\cdot y_1, *)+L(*, x_2+\beta\cdot y_2))(r,s) \\ & = & L(x_1+\beta\cdot y_1, s)+L(r, x_2+\beta\cdot y_2) \\ & = & L(x_1,s)+\beta\cdot L(y_1,s)+L(r,x_2)+\beta\cdot L(r,y_2) \\ & = & (L(x_1,s)+L(r,x_2))+\beta\cdot(L(y_1,s)+L(r,y_2)) \\ & = & D^{[1]}L(x_1,x_2)(r,s)+\beta\cdot D^{[1]}L(y_1,y_2)(r,s) \\ & = & (D^{[1]}L(x_1,x_2)+D^{[1]}L(y_1,Y_2))(r,s) \end{array}$$

proving that

$$D^{[1]}L((x_1,x_2)+\beta\cdot(y_1,y_2)) = D^{[1]}L(x_1,x_2)+D^{[1]}L(y_1,Y_2)$$

Hence we have that

$$D^{[1]}L \text{ is linear} \tag{16.97}$$

Further we have that $\forall (x_1, x_2), (r, s) \in X_1 \cdot X_2$

$$\begin{array}{rl} \|D^{[1]}L(x_1,x_2)(r,s)\| & = \\ \|L(x_1,s)+L(r,x_2)\| & \leqslant \\ \|L(x_1,s)\|+\|L(r,x_2)\| & \leqslant_{[\text{theorem: } 14.187]} \\ \|L\|_{L(X_1,X_2;Y)}\cdot\|x\|_1\cdot\|s\|_2+\|L\|_{L(X_1,X_2;Y)}\cdot\|r\|_1\cdot\|x_2\|_2 & \leqslant \\ \|L\|_{L(X_1,X_2;Y)}\cdot(\|x\|_1\cdot\|s\|_2+\|r\|_1\cdot\|x_2\|_2) & \leqslant \\ \|L\|_{L(X_1,X_2;Y)}\cdot(\|x\|_1\cdot\max(\|r\|_1,\|s\|_2)+\|x_2\|\cdot\max(\|r\|_1,\|s\|_2)) & = \\ \|L\|_{L(X_1,X_2;Y)}\cdot(\|x_1\|_1+\|x_2\|_2)\cdot\max(\|r\|_1,\|s\|_2) & = \\ \|L\|_{L(X_1,X_2;Y)}\cdot(\|x_1\|_1+\|x_2\|_2)\cdot\|(r,s)\|_{\max} & \leqslant \\ \|L\|_{L(X_1,X_2;Y)}\cdot 2\cdot\max(\|x_1\|_1,\|x_2\|_2)\cdot\|(r,s)\|_{\max} & = \\ (2\cdot\|L\|_{L(X_1,X_2;Y)}\cdot\|(x_1,x_2)\|_{\max})\cdot\|(r,s)\|_{\max} & \end{array}$$

proving, as $D^{[1]}L(x_1, x_2) \in L(X_1 \cdot X_2, Y)$, by [definition: 14.178] that

$$\|D^{[1]}L(x_1,x_2)\|_{L(X_1\cdot X_2,Y)} \leqslant 2\cdot\|L\|_{L(X_1,X_2;Y)}\cdot\|(x_1,x_2)\|_{\max}$$

so that by [theorem: 14.187] it follows that

$$D^{[1]}L \in L(X_1 \cdot X_2, L(X_1 \cdot X_2, Y)) = L_2(X_1 \cdot X_2; Y) \tag{16.98}$$

Using [example: 16.27] it follows that $D^{[1]}L$ is differentiable at $(x_1, x_2) \in X_1 \cdot X_2$ with $D_{(x_1,x_2)}(D^{[1]}L) = D^{[1]}L$, hence it follows that

$$L \text{ is 2-times differentiable at } (x_1,x_2) \text{ and } D^{[2]}_{(x_1,x_2)}L = D^{[1]}L \tag{16.99}$$

$n \in \{3, \ldots, n\}$. We use induction on n to proceed, so define

$$S = \left\{n \in \{3, \ldots, \infty\} | L \text{ is } n\text{-times differentiable at } (x_1,x_2) \in X_1 \cdot X_2 \text{ and } D^{[n]}_{(x_1,x_2)}L = 0_n\right\}$$

then we have:

$3 \in S$. Define

$$D^{[2]}f: X_1 \cdot X_2 \rightarrow L_2(X_1 \cdot X_2; Y) \text{ by } D^{[2]}f(x_1,x_2) = D^{[2]}_{(x_1,x_2)}f$$

then by [eq: 16.99] $D^{[2]}f = C_{D^{[1]}f}$. Hence using [example: 16.26] it follows that $D^{[2]}f$ is differentiable at $(x_1, x_2) \in X_1 \cdot X_2$ with $D_{(x_1,x_2)}(D^{[2]}f) = 0_3$. So by definition L is 3-times differentiable at (x_1, x_2) with $D^{[3]}_{(x_1,x_2)}L = 0_3$.

$n \in S \Rightarrow n+1 \in S$. As $n \in S$ it follows that $\forall (x_1, x_2) \in X_1 \cdot X_2$ L is n-times differentiable at x and that $D^{[n]}_{(x_1,x_2)}L(x) = 0_n$. Hence if we define

$$D^{[n]}L: X_1 \cdot X_2 \rightarrow L_n(X_1 \cdot X_2; Y) \text{ by } (D^{[n]}L)(x_1,x_2) = D^{[n]}_{(x_1,x_2)}L$$

it follows that $D^{[n]}L = C_{0_n}$. Using [example: 16.115] it follows then that $\forall (x_1, x_2) \in X_1 \cdot X_2$ $D^{[n]}L$ is 1-times differentiable (hence Fréchet differentiable) at x and that $D_{(x_1,x_2)}(D^{[n]}L) = D^{[1]}_{(x_1,x_2)}(D^{[n]}L)$ is the neutral element of $L_1(X_1 \cdot X_2; L_n(X_1 \cdot X_2; Y)) = L_{n+1}(X_1 \cdot X_2; Y)$. Hence L is $(n+1)$-times differentiable at x and $D^{[n+1]}_{(x_1,x_n)} L = D_{(x_1,x_2)}(D^{[n]}L) = 0_{n+1}$ proving that $n+1 \in S$. □

16.2.2.8 The Chain Rule for Higher order Differentiation

Lemma 16.121. *Let $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$, $\langle Z, \|\|_Z\rangle$ be normed spaces, $L \in L(Y, Z)$ then if we define*

$$(L \circ *): L(X, Y) \rightarrow L(X, Z) \text{ by } (L \circ *)(T) = L \circ T$$

then we have that

$$(L \circ *) \in L(L(X, Y), L(X, Z))$$

Proof. First note that if $L \in L(Y, Z)$ and $T \in L(X, Y)$ then by [theorem: 14.185] $L \circ T \in L(X, Z)$ so that $(* \circ L)$ is indeed a function between $L(X, Y)$ and $L(X, Z)$. Second we proof linearity. Let $T_1, T_2 \in L(X, Y)$, $\alpha \in \mathbb{K}$ then for $z \in X$ we have

$$\begin{aligned}
((L \circ *)(T_1 + \alpha \cdot T_2)(z) &= (L \circ (T_1 + \alpha \cdot T_2))(z) \\
&= L((T_1 + \alpha \cdot T_2)(z)) \\
&= L(T_1(z) + \alpha \cdot T_2(z)) \\
&= L(T_1(z)) + \alpha \cdot L(T_2(z)) \\
&= (L \circ T_1)(z) + \alpha \cdot (L \circ T_2)(z) \\
&= ((L \circ T_1) + \alpha \cdot (L \circ T_2))(z) \\
&= ((L \circ *)(T_1) + \alpha \cdot (L \circ *)(T_2))(z)
\end{aligned}$$

proving that $(L \circ *)(T_1 + \alpha \cdot T_2) = (L \circ *)(T_1) + \alpha \cdot (L \circ *)(T_2)$ or

$$(L \circ *) \in \mathrm{Hom}(L(X, Y), L(X, Z))$$

As for continuity. Let $T \in L(X, Y)$ then as $L \in L(Y, Z)$ we have by [theorem: 14.185] that

$$\|(L \circ *)(T)\|_{L(X,Z)} = \|L \circ T\|_{L(X,Z)} \leqslant \|L\|_{L(Y,Z)} \cdot \|T\|_{L(X,Y)}$$

which proves by [theorem: 14.174] that

$$(L \circ *) \in L(L(X, Y), L(X, Z))$$

□

Lemma 16.122. *Let $n \in \mathbb{N}$, $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$, $\langle Z, \|\|_Z\rangle$ normed spaces, U a open set in X, $x \in U$ $f: U \rightarrow Y$ is a function that is n-times differentiable at x then $L \circ f$ is n-times differentiable at x.*

Proof. We use induction to prove this, so define

$S = \{n \in \mathbb{N} |$If $\langle X, \|\|_X\rangle, \langle Y, \|\|_Y\rangle, \langle Z, \|\|_Z\rangle$ are normed spaces, U a open set in X, $x \in U$, $f: U \rightarrow Y$ n-times differentiable at x and $L \in L(Y, Z)$ then $L \circ f$ is n-times differentiable at $x\}$

then we have:

$1 \in S$. If f is 1-times differentiable at x then f is Fréchet differentiable at x. Hence if $L \in L(Y, Z)$ we have by [corollary: 16.38] that $L \circ f$ is Fréchet differentiable at x, hence by definition 1-times differentiable at x, which proves that $1 \in S$.

$n \in S \Rightarrow n+1 \in S$. Let $f: U \rightarrow Y$ be $(n+1)$ differentiable at x and let $L \in L(Y, Z)$. Using the fact that $n+1 \in \mathbb{N} \setminus \{1\}$ and [theorem: 16.109] there exist a open set V with $x \in V \subseteq U$ such that $\forall y \in V$ f is 1-times differentiable at y [hence Fréchet differentiable at y] and

$$D^{[1]}f: V \rightarrow L(X, Y) \text{ defined by } D^{[1]}f(y) = D^{[1]}_y f \underset{\text{def}}{=} D_y f \tag{16.100}$$

is n-times differentiable at x. As f is Fréchet differentiable at $y \in U$ we can use [corollary: 16.38] to prove that $L \circ f$ is Fréchet differentiable at y and $D_y(L \circ f) = L \circ D_y f$, so by definition

$$L \circ f \text{ is 1-times differentiable at } y \in U \text{ and } D^{[1]}_y(L \circ f) = L \circ D^{[1]}_y f \tag{16.101}$$

Define then

$$D^{[1]}(L \circ f): U \rightarrow L(X, L(X, Z)) \text{ by } D^{[1]}(L \circ f)(y) = D^{[1]}_y(L \circ f) \tag{16.102}$$

Take now $(L\circ *)\in L(L(X,Y),L(X,Z))$ defined by $(L\circ *)(T)=L\circ T$ [see lemma: 16.121]. Let $y\in V$ then

$$\begin{aligned}((L\circ *)\circ D^1 f)(y) &= (L\circ *)(D^{[1]}f(y))\\ &= (L\circ *)\left(D_y^{[1]}f\right)\\ &= L\circ D_y^{[1]}f\\ &\underset{[\text{eq: }16.101]}{=} D_y^{[1]}(L\circ f)\\ &\underset{[\text{eq: }16.102]}{=} D^{[1]}(L\circ f)(y)\end{aligned}$$

which proves that

$$((L\circ *)\circ D^1 f)=D^{[1]}(L\circ f) \tag{16.103}$$

As $D^{[1]}f$ is n-times differentiable at x, $(L\circ *)\in L(L(X,Y),L(X,Z))$ and $n\in S$ it follows that $(L\circ *)\circ D^{[1]}f$ is n-times differentiable at x, hence, using [eq: 16.103], it follows that $D^{[1]}(L\circ f)$ is n-times differentiable at x. Using [eq: 16.101] together with [theorem: 16.109] proves that $L\circ f$ is $(n+1)$-times differentiable at x. So

$$n+1\in S$$ □

We use a similar proof for the following lemma.

Lemma 16.123. *Let $n\in\mathbb{N}$, $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$, $\langle Z,\|\|_Z\rangle$ normed spaces, U a open set in X, $x\in U$ $f:U\rightarrow Y$ is a function that is of class C^n, $L\in L(Y,Z)$ then $L\circ f$ is of class C^n.*

Proof. For $n\in\mathbb{N}_0$ we have either:

$n=0$. As f is of class C^0, f is continuous, hence, as L is continuous, $L\circ f$ is continuous so that $L\circ f$ is of class C^0.

$n\in\mathbb{N}$. We use induction to prove this, so define

$S=\{n\in\mathbb{N}_0|$If $\langle X,\|\|_X\rangle,\langle Y,\|\|_Y\rangle,\langle Z,\|\|_Z\rangle$ are normed spaces, U a open set in X, $f:U\rightarrow Y$ is of class C^n and $L\in L(Y,Z)$ then $L\circ f$ is of class $C^n\}$

then we have:

$1\in S$. If f is of class C^1 we have that $\forall x\in U$ f is 1-times differentiable at x [hence f is Fréchet differentiable at x] and

$$D^{[1]}f:U\rightarrow L(X,Y) \text{ defined by } D^{[1]}f(x)=D_x^{[1]}f=D_x f$$

is continuous. Let $x\in U$ then as f is Fréchet differentiable at x it follows from [corollary: 16.38] that $L\circ f$ is Fréchet differentiable at x [hence 1-times differentiable at x] and

$$D_x^{[1]}(L\circ f)=D_x(L\circ f)=L\circ D_x f=L\circ D_x^{[1]}f \tag{16.104}$$

Define

$$D^{[1]}(L\circ f):U\rightarrow L(X,Z) \text{ by } D^{[1]}(L\circ f)(x)=D_x^{[1[}(L\circ f)\underset{[\text{eq: }16.104]}{=}L\circ D_x^{[1]}f \tag{16.105}$$

Further

$$((L\circ *)\circ D^{[1]}f)(x)=(L\circ *)(D^{[1]}f(x))=L\circ D^{[1]}f(x)=L\circ D_x^{[1]}f\underset{[\text{eq: }16.105]}{=}D^{[1]}(L\circ f)(x)$$

so that that

$$(L\circ *)\circ D^{[1]}f=D^{[1]}(L\circ f).$$

As $D^{[1]}f$ is continuous and $(L\circ *)$ is continuous [see lemma: 16.121] it follows that $D^{[1]}(L\circ f)$ is continuous. Hence we have that $L\circ f$ is of class C^1 proving that $1\in S$.

$n\in S\Rightarrow n+1\in S$. If f is of class C^{n+1} then by [theorem: 16.111] it follows that $\forall x\in U$ f is 1-times differentiable at x [hence Fréchet differentiable at x] and

$$D^{[1]}f:U\rightarrow L(X,Y) \text{ defined by } (D^{[1]}f)(x)=D^{[1]}f(x) \text{ is of class } C^n$$

As $\forall x\in U$ f is Fréchet differentiable at x it follows from [corollary: 16.38] that $L\circ f$ is Fréchet differentiable at x [hence 1-times differentiable at x] and

$$D_x^{[1]}(L\circ f)=D_x(L\circ f)=L\circ D_x f=L\circ D_x^{[1]}f \tag{16.106}$$

Define

$$D^{[1]}(L\circ f): U\rightarrow L(X,Z) \text{ by } D^{[1]}(L\circ f)(x)=D^{[1]}_x(L\circ f) \tag{16.107}$$

then we have

$$\begin{array}{rcl} ((L\circ *)\circ D^{[1]}f)(x) & = & (L\circ *)(D^{[1]}f)(x)\\ & = & L\circ D^{[1]}f(x)\\ & = & L\circ D^{[1]}_x f\\ & \underset{[\text{eq: }16.106]}{=} & D^{[1]}_x(L\circ f)\\ & \underset{[\text{eq: }16.107]}{=} & D^{[1]}(L\circ f)(x) \end{array}$$

so that

$$(L\circ *)\circ D^{[1]}f=D^{[1]}(L\circ f) \tag{16.108}$$

As $D^1 f\in L(X,L(X,Y))$ is of class C^n, $(L\circ *)\in L(L(X,Y),L(X,Z))$ and $n\in S$ it follows that $(L\circ *)\circ D^{[1]}f$ is of class C^n, hence, using [eq 16.108] $D^{[1]}(L\circ f)$ is of class $C^{[n]}$. Applying then [theorem: 16.111] it follows that $L\circ f$ is of class C^{n+1} proving that

$$n+1\in S$$

□

Lemma 16.124. *Let $n\in\mathbb{N}$, $k\in\mathbb{N}$, $\langle X,\|\|\rangle$ a normed space, $\{\langle Y_i,\|\|_i\rangle\}_{i\in\{1,\ldots,k\}}$ a finite family of normed spaces then if we define [see lemma: 16.59]*

$$\mathcal{K}: \prod_{i\in\{1,\ldots,k\}} L(X,Y_i)\rightarrow L\left(X,\prod_{i\in\{1,\ldots,k\}} Y_i\right) \text{ by } \mathcal{K}((L_1,\ldots,L_n))=(L_1,\ldots,L_k)_*$$

we have

$$\mathcal{K}\in L\left(\prod_{i\in\{1,\ldots,k\}} L(X,Y_i), L\left(X,\prod_{i\in\{1,\ldots,k\}} Y_i\right)\right)$$

where $\langle\prod_{i\in\{1,\ldots,k\}} L(X,Y_i),\|\|_{\max}\rangle$ is the normed space with the maximum norm $\|\|_{\max}$ defined to be $\|(L_1,\ldots,L_k)\|_{\max}=\max(\{\|L_i\|_{L(X,Y_i)}|i\in\{1,\ldots,k\}\})$

Proof. Let $\alpha\in\mathbb{K}$, $(L_1,\ldots,L_k)$, $(K_1,\ldots,K_k)\in\prod_{i\in\{1,\ldots,k\}} L(X,Y_i)$ then we have for $x\in X$ that

$$\begin{array}{rcl} \mathcal{K}((L_1,\ldots,L_k)+\alpha\cdot(K_1\ldots,K_k))(x) & = & \mathcal{K}((L_1+\alpha\cdot K_1,\ldots,L_k+\alpha\cdot K_k))(x)\\ & = & (L_1+\alpha\cdot K_1,\ldots,L_k+\alpha\cdot K_k)_*(x)\\ & = & ((L_1+\alpha\cdot K_1)(x),\ldots,(L_k+\alpha\cdot K_x)(x))\\ & = & (L_1(x)+\alpha\cdot K_1(x),\ldots,L_2(x)+\alpha\cdot K_2(x))\\ & = & (L_1(x),\ldots,L_k(x))+\alpha\cdot(K_1(x),\ldots,K_k(x))\\ & = & (L_1,\ldots,L_k)_*(x)+\alpha\cdot(K_1,\ldots,K_k)_*(x)\\ & = & \mathcal{K}((L_1,\ldots,L_k))(x)+\alpha\cdot\mathcal{K}((K_1,\ldots,K_k))(x)\\ & = & (\mathcal{K}((L_1,\ldots,L_k))+\alpha\cdot\mathcal{K}((K_1,\ldots,K_k)))(x) \end{array}$$

proving that $\mathcal{K}((L_1,\ldots,L_k)+\alpha\cdot(K_1\ldots,K_k))=\mathcal{K}((L_1,\ldots,L_k))+\alpha\cdot\mathcal{K}((K_1,\ldots,K_k))$. So we have

$$\mathcal{K}\in\mathrm{Hom}\left(\prod_{i\in\{1,\ldots,k\}} L(X,Y_i), L\left(X,\prod_{i\in\{1,\ldots,k\}} Y_i\right)\right)$$

Let $(L_1,\ldots,L_k)\in\prod_{i\in\{1,\ldots,k\}} L(X,Y_i)$ then we have that

$$\begin{array}{rcl} \|\mathcal{K}((L_1,\ldots,L_k))\|_{L(X,\prod_{i\in\{1,\ldots,k\}}Y_i)} & = & \|(L_1,\ldots,L_k)_*\|_{L(X,\prod_{i\in\{1,\ldots,k\}}Y_i)}\\ & \leqslant_{[\text{lemma: }16.59]} & \max(\{\|L_i\|_{L(X,Y_i)}|i\in\{1,\ldots,k\}\})\\ & = & \|(L_1,\ldots,L_k)\|_{\max} \end{array}$$

proving that

$$\mathcal{K}\in L\left(\prod_{i\in\{1,\ldots,k\}} L(X,Y_i), L\left(X,\prod_{i\in\{1,\ldots,k\}} Y_i\right)\right)$$

□

Theorem 16.125. *Let $n\in\mathbb{N}$, $k\in\mathbb{N}$, $\langle X,\|\|\rangle$ a normed space, $\{\langle Y_i,\|\|_i\rangle\}_{i\in\{1,\ldots,k\}}$ a finite family of normed spaces, $U\subseteq X$, $x\in U$ and $f:U\rightarrow\prod_{i\in\{1,\ldots,n\}}Y_i$ then we have*

$$f \text{ is } n\text{-times differentiable at } x$$
$$\Updownarrow$$
$$\forall i\in\{1,\ldots,k\} \text{ we have } \pi_i\circ f \text{ is } n\text{-times differentiable at } x$$

Proof.

⇒. Using [example: 14.184] it follows that $\forall i\in\{1,\ldots,k\}$ $\pi_i\in L(X,Y_i)$. So, as f is n-times differentiable at x, it follows from [lemma: 16.122] that $\pi_i\circ f$ is n-times differentiable at x.

⇐. We use recursion to prove this, so let

$$S=\left\{n\in\mathbb{N}|\text{Let } \langle X,\|\|\rangle \text{ be a normed space, } \{\langle Y_i\rangle,\|\|_i\}_{i\in\{1,\ldots,k\}} \text{ be normed spaces , } x\in U \text{ a open set in } X \text{ and } f:U\rightarrow\prod_{j\in\{1,\ldots,n\}}Y_j\right.$$

a function such that $\forall i\in\{1,\ldots,k\}$ $\pi_i\circ f$ is n-times differentiable at x then f is n-times differentiable at $x\}$

then we have:

$1\in S$. As $\forall i\in\{1,\ldots,k\}$ $\pi_i\circ f$ is 1-times differentiable at x it follows that $\pi_i\circ f$ is Fréchet differentiable at x. Hence using [theorem: 16.61] f is Fréchet differentiable at x proving that f is 1-times differentiable at x. So we conclude that $1\in S$.

$n\in S\Rightarrow n+1\in S$. Let $i\in\{1,\ldots,k\}$ then as $\pi_i\circ f$ is $(n+1)$-times differentiable it follows form [theorem: 16.109] that there exist a open set V_i with $x\in V_i\subseteq U$ such that $\forall y\in V_i$ $\pi_i\circ f$ is 1-times differentiable at y and

$$D^{[1]}_{V_i}(\pi_i\circ f):V_i\rightarrow L(X,Y_i) \text{ defined by } D^{[1]}_{V_i}(\pi_i\circ f)(y)=D^{[1]}_y(\pi_i\circ f)$$

is n-times differentiable at x. Using the locality of differentiability [see theorem: 16.90] it follows that for $V=\bigcap_{j\in\{1,\ldots,k\}}V_j\subseteq U$ $x\in V$ and

$$D^{[1]}(\pi_i\circ f):V\rightarrow L(X,Y_i) \text{ defined by } D^{[1]}(\pi_i\circ f)(y)=D^{[1]}_y(\pi_i\circ f)$$

is n-times differentiable at x. Now for

$$(D^{[1]}(\pi_1\circ f),\ldots,D^{[1]}(\pi_k\circ f))_*:V\rightarrow\prod_{i\in\{1,\ldots,k\}}L(X,Y_i) \tag{16.109}$$

defined by [see definition: 16.57]

$$(D^{[1]}(\pi_1\circ f),\ldots,D^{[1]}(\pi_k\circ f))_*(y)=(D^{[1]}(\pi_1\circ f)(y),\ldots,D^{[1]}(\pi_k\circ f)(y))$$

we have for $y\in V$, $i\in\{1,\ldots,k\}$ that

$$\begin{aligned}(\pi_i\circ(D^{[1]}(\pi_1\circ f),\ldots,D^{[1]}(\pi_k\circ f))_*)(y) &=\\ \pi_i((D^{[1]}(\pi_1\circ f),\ldots,D^{[1]}(\pi_k\circ f))_*(y)) &=\\ \pi_i((D^{[1]}(\pi_1\circ f))(y),\ldots,(D^{[1]}(\pi_k\circ f))(y)) &=\\ (D^{[1]}(\pi_i\circ f))(y)\end{aligned}$$

proving that $\pi_i\circ(D^{[1]}(\pi_1\circ f),\ldots,D^{[1]}(\pi_k\circ f))_*=D^{[1]}(\pi_i\circ f)$. Hence, as $D^{[1]}(\pi_i\circ f)$ is n-times differentiable at x, we have proved that $\forall i\in\{1,\ldots,k\}$

$$\pi_i\circ(D^{[1]}(\pi_1\circ f),\ldots,D^{[1]}(\pi_k\circ f))_*:V\rightarrow L(X,Y_i)$$

is n-times differentiable at x. As $n\in S$ it follows that

$$(D^{[1]}(\pi_1\circ f),\ldots,D^{[1]}(\pi_k\circ f))_* \text{ is } n\text{-times differentiable at } x \tag{16.110}$$

As $\forall i\in\{1,\ldots,k\}$ $\pi_i\circ f$ is 1-times differentiable at $y\in V$ it follows that $\pi_i\circ f$ is Fréchet differentiable at y, hence using [theorem: 16.61] f is Fréchet differentiable at y and

$$D^{[1]}_y f=D_yf(y)=(D_y(\pi_1\circ f),\ldots,D_y(\pi_k\circ f))_*\in L\left(X,\prod_{i\in\{1,\ldots,k\}}Y_i\right) \tag{16.111}$$

allowing use to define

$$D^{[1]}f: V \to L\left(X, \prod_{i \in \{1,\ldots,k\}} Y_i\right) \text{ by } D^{[1]}f(y) = D_y^{[1]}f$$

Using $\mathcal{K}: \prod_{i \in \{1,\ldots,k\}} L(X, Y_i) \to L\left(X, \prod_{i \in \{1,\ldots,k\}} Y_i\right)$ defined in [lemma: 16.124] we have that for $y \in V$

$$\begin{aligned}
(\mathcal{K} \circ (D^{[1]}(\pi_1 \circ f), \ldots, D^{[1]}(\pi_k \circ f))_*)(y) &= \\
\mathcal{K}((D^{[1]}(\pi_1 \circ f), \ldots, D^{[1]}(\pi_k \circ f))_*(y)) &= \\
\mathcal{K}((D^{[1]}(\pi_1 \circ f)(y), \ldots, D^{[1]}(\pi_k \circ f)(y))) &= \\
((D^{[1]}(\pi_1 \circ f)(y), \ldots, D^{[1]}(\pi_k \circ f)(y)))_* &= \\
\left(\left(D_y^{[1]}(\pi_1 \circ f), \ldots, D_y^{[1]}(\pi_k \circ f)\right)\right)_* &= \\
((D_y(\pi_1 \circ f), \ldots, D_y(\pi_k \circ f)))_* &\underset{\text{[eq: 16.111}}{=} \\
D^{[1]}f(y) &= \\
(D^{[1]}f)(y)
\end{aligned}$$

which proves that

$$D^{[1]} f = \mathcal{K} \circ (D^{[1]}(\pi_1 \circ f), \ldots, D^{[1]}(\pi_k \circ f))_*$$

As $(D^{[1]}(\pi_1 \circ f), \ldots, D^{[1]}(\pi_k \circ f))_*$ is n-times differentiable at x by [eq: 16.110], it follows that $D^{[1]}f$ is n-times differentiable at x. Hence using [theorem: 16.109] it follows that

$$f \text{ is } (n+1)\text{-times differentiable at } x$$

proving that

$$n + 1 \in S$$

□

We can use a similar proof for differential classes.

Theorem 16.126. *Let $n \in \mathbb{N}_0$, $k \in \mathbb{N}$, $\langle X, \|\|\rangle$ a normed space, $\{\langle Y_i, \|\|_i\rangle\}_{i \in \{1,\ldots,k\}}$ a finite family of normed spaces, $U \subseteq X$, $x \in U$ and $f: U \to \prod_{i \in \{1,\ldots,n\}} Y_i$ a function then we have*

$$f \textit{ is } \text{of class } C^n$$
$$\Updownarrow$$
$$\forall i \in \{1,\ldots,k\} \textit{ we have } \pi_i \circ f \textit{ is of class } C^n$$

Proof.

⇒. Using [example: 14.184] it follows that $\forall i \in \{1,\ldots,k\}$ $\pi_i \in L(X, Y_i)$. So, as f is of class C^n, it follows from [lemma: 16.123] that $\pi_1 \circ f$ is of class C^n.

⇐. If $n = 0$ then $\forall i \in \{1,\ldots,k\}$ $\pi_i \circ f$ is of class C^0 so that $\pi_i \circ f$ is continuous, using [theorem: 14.140] it follows that f is continuous proving that f is of class C^0. Hence we just have to prove the remaining cases $n \in \mathbb{N}$. We use induction to prove this for $n \in \mathbb{N}$, so define

$S = \Big\{ n \in \mathbb{N} |$Let $\langle X, \|\|\rangle$ be a normed space, $\{\langle Y_i\rangle, \|\|_i\}_{i \in \{1,\ldots,k\}}$ be normed spaces , U a open set in X and $f: U \to \prod_{j \in \{1,\ldots,n\}} Y_j$ a function such that $\forall i \in \{1,\ldots,k\}$ $\pi_i \circ f$ is of class C^n then f is of class $C^n\}$

then we have:

$1 \in S$. Let $i \in \{1,\ldots,k\}$ then as $\pi_i \circ f$ is of class C^1 it follows that $\forall x \in U$ $\pi_i \circ f$ is 1-times differentiable at x [hence Fréchet differentiable at x] and the function

$$D^{[1]}(\pi_i \circ f): U \to L(X, Y_i) \text{ by } D^{[1]}(\pi_i \circ f)(x) = D_x^{[1]}(\pi_i \circ f) = D(\pi_i \circ f)(x) \tag{16.112}$$

is continuous. By [theorem: 16.61] it follows then that for $x \in U$ f is Fréchet differentiable at x and $D_x f = (D_x(\pi_1 \circ f), \ldots, D_x(\pi_k \circ f))_*$, hence f is 1-times differentiable at x and

$$D_x^{[1]} f = \left(D_x^{[1]}(\pi_1 \circ f), \ldots, D_x^{[1]}(\pi_k \circ f)\right)_* \in L\left(X, \prod_{i \in \{1,\ldots,k\}} Y_i\right) \tag{16.113}$$

Take now the following function

$$D^{[1]}f: U \rightarrow L\left(X, \prod_{i \in \{1, \ldots, k\}} Y_i\right) \text{ defined by } D^{[1]}f(x) = D^{[1]}_x f \tag{16.114}$$

Using $\mathcal{K}: \prod_{i \in \{1, \ldots, k\}} L(X, Y_i) \rightarrow L\left(X, \prod_{i \in \{1, \ldots, k\}} Y_i\right)$ defined in [lemma: 16.124] we have that for $x \in U$ that

$$\begin{array}{rcl} (\mathcal{K} \circ (D^{[1]}(\pi_1 \circ f), \ldots, D^{[1]}(\pi_k \circ f))_*)(x) & = & \\ \mathcal{K}((D^{[1]}(\pi_1 \circ f), \ldots, D^{[1]}(\pi_k \circ f))_*(x)) & = & \\ \mathcal{K}((D^{[1]}(\pi_1 \circ f)(x), \ldots, D^{[1]}(\pi_k \circ f)(x))) & = & \\ \mathcal{K}\left(\left(D^{[1]}_x(\pi_1 \circ f), \ldots, D^{[1]}_x(\pi_k \circ f)\right)\right) & & \\ \left(D^{[1]}_x(\pi_1 \circ f), \ldots, D^{[1]}_x(\pi_k \circ f)\right)_* & \underset{[\text{eq: } 16.113]}{=} & \\ D^{[1]}f(x) & \underset{[\text{eq: } 16.114]}{=} & \\ (D^{[1]}f)(x) & & \end{array}$$

proving that

$$D^{[1]}f = \mathcal{K} \circ (D^{[1]}(\pi_1 \circ f), \ldots, D^{[1]}(\pi_k \circ f))_* \tag{16.115}$$

Let $i \in \{1, \ldots, n\}$ then $\forall x \in U$ we have

$$\begin{array}{rcl} (\pi_i \circ (D^{[1]}(\pi_1 \circ f), \ldots, D^{[1]}(\pi_k \circ f))_*)(x) & = & \\ \pi_i((D^{[1]}(\pi_1 \circ f)(x), \ldots, D^{[1]}(\pi_k \circ f)(x))) & = & \\ (D^{[1]}(\pi_i \circ f))(x) & = & \end{array}$$

proving that

$$\pi_i \circ (D^{[1]}(\pi_1 \circ f), \ldots, D^{[1]}(\pi_k \circ f))_* = D^{[1]}(\pi_i \circ f)$$

Hence as $D^{[1]}(\pi_i \circ f)$ is continuous it follows from [theorem: 14.140] that

$$(D^{[1]}(\pi_1 \circ f), \ldots, D^{[1]}(\pi_k \circ f))_*$$

is continuous. Using the fact that $\mathcal{K}$ is also continuous [see lemma: 16.124] we conclude from [eq: 16.115] that $D^{[1]}f$ is continuous. Hence f is of class C^1 so that

$$1 \in S$$

$\boldsymbol{n \in S \Rightarrow n+1 \in S}$**.** Let $i \in \{1, \ldots, k\}$ then as $\pi_i \circ f$ is of class C^{n+1} it follows that $\forall x \in U$ we have that $\pi_i \circ f$ is $(n+1)$-times differentiable at x and that the function

$$D^{[n+1]}(\pi_i \circ f): U \rightarrow L_{n+1}(X; Y_i) \text{ by } D^{[n+1]}(\pi_i \circ f)(x) = D^{[n+1]}_x(\pi_i \circ f)$$

is continuous. As $\pi_i \circ f$ is of class C^{n+1} it follows from [theorem: 16.111] that $\forall x \in U$ $\pi_1 \circ f$ is 1-times differentiable at x [hence Fréchet differentiable at x] and that the function

$$D^{[1]}(\pi_i \circ f): U \rightarrow L(X, Y_i) \text{ by } D^{[1]}(\pi_i \circ f)(x) = D^{[1]}_x(\pi_i \circ f) = D_x(\pi_i \circ f)$$

is of class C^n. Consider now the function [see definition: 16.57]

$$(D^{[1]}(\pi_1 \circ f), \ldots, D^{[1]}(\pi_k \circ f))_*: U \rightarrow \prod_{i \in \{1, \ldots, k\}} L(X, Y_i)$$

defined by

$$(D^{[1]}(\pi_1 \circ f), \ldots, D^{[1]}(\pi_k \circ f))_*(x) = (D^{[1]}(\pi_1 \circ f)(x), \ldots, D^{[1]}(\pi_k \circ f)(x))$$

then we have $\forall i \in \{1, \ldots, k\}$ that

$$\begin{array}{rcl} (\pi_i \circ (D^{[1]}(\pi_1 \circ f), \ldots, D^{[1]}(\pi_k \circ f))_*)(x) & = & \\ \pi_i((D^{[1]}(\pi_1 \circ f), \ldots, D^{[1]}(\pi_k \circ f))_*(x)) & = & \\ \pi_i(((D^{[1]}(\pi_1 \circ f))(x), \ldots, (D^{[1]}(\pi_k \circ f))(x))) & = & \\ (D^{[1]}(\pi_i \circ f))(x) & & \end{array}$$

proving that

$$\pi_i \circ (D^{[1]}(\pi_1 \circ f), \ldots, D^{[1]}(\pi_k \circ f))_* = D^{[1]}(\pi_i \circ f)$$

so that $\pi_i \circ (D^{[1]}(\pi_1 \circ f), \ldots, D^{[1]}(\pi_k \circ f))_*$ is of class C^n. As $n \in S$ it follows that

$$(D^{[1]}(\pi_1 \circ f), \ldots, D^{[1]}(\pi_k \circ f))_* \text{ is of class } C^n \tag{16.116}$$

Let $x \in U$ then as $\forall i \in \{1, \ldots, k\}$ $(\pi_i \circ f)$ is Fréchet differentiable at x it follows from [theorem: 16.61] that f is Fréchet differentiable at x and $D_x f = (D_x(\pi_1 \circ f), \ldots, D_x(\pi_k \circ f))_*$. Translating this to 1-times differentiability we have that f is 1-times differentiable at x and

$$D_x^{[1]} f = \left(D_x^{[1]}(\pi_1 \circ f), \ldots, D_x^{[1]}(\pi_k \circ f)\right)_* \in L\left(X, \prod_{i \in \{1, \ldots, k\}} L(X, Y_i)\right) \tag{16.117}$$

Consider the function

$$D^{[1]} f: U \rightarrow L\left(X, \prod_{i \in \{1, \ldots, k\}} Y_i\right) \text{ defined by } D^{[1]} f(x) = D_x^{[1]} f \tag{16.118}$$

Using $\mathcal{K}: \prod_{i \in \{1, \ldots, k\}} L(X, Y_i) \rightarrow L\left(X, \prod_{i \in \{1, \ldots, k\}} Y_i\right)$ defined in [lemma: 16.124] we have that for $x \in U$

$$\begin{aligned}
(\mathcal{K} \circ (D^{[1]}(\pi_1 \circ f), \ldots, D^{[1]}(\pi_k \circ f))_*)(x) &= \\
\mathcal{K}((D^{[1]}(\pi_1 \circ f), \ldots, D^{[1]}(\pi_k \circ f))_*(x)) &= \\
\mathcal{K}((D^{[1]}(\pi_1 \circ f)(x), \ldots, D^{[1]}(\pi_k \circ f)(x))) &= \\
(D^{[1]}(\pi_1 \circ f)(x), \ldots, D^{[1]}(\pi_k \circ f)(x))_* &= \\
\left(D_x^{[1]}(\pi_1 \circ f), \ldots, D_x^{[1]}(\pi_k \circ f)\right)_* &\underset{[\text{eq: } 16.117]}{=} \\
D_x^{[1]} f &\underset{[\text{eq: } 16.118]}{=} \\
(D^{[1]} f)(x)
\end{aligned}$$

proving that

$$D^{[1]} f = \mathcal{K} \circ (D^{[1]}(\pi_1 \circ f), \ldots, D^{[1]}(\pi_k \circ f))_* \tag{16.119}$$

As by [eq: 16.116] $(D^{[1]}(\pi_1 \circ f), \ldots, D^{[1]}(\pi_k \circ f))_*$ is of class C^n we can use [lemma: 16.123] on the above proving that $D^{[1]} f$ is of class C^n. Applying then [theorem: 16.111] results in the fact that f is of class C^{n+1}. Hence

$$n + 1 \in S$$

□

Lemma 16.127. *Let $\langle X, \|\|_X \rangle$, $\langle Y, \|\|_Y \rangle$, $\langle Z, \|\|_Z \rangle$ be normed spaces then*

$$(\circ): L(Y, Z) \cdot L(X, Y) \rightarrow L(X, Z) \text{ defined by } (\circ)(L, K) = L \circ K$$

is ∞-times differentiable on X. Further using [theorem: 16.98] $(\circ)$ is of class C^∞.

Proof. Using [theorem: 14.197] we have that $(\circ) \in L((Y, Z), L(X, Y); L(X, Z))$ so that by [lemma: 16.120] $(\circ)$ is ∞-times differentiable at every $x \in X$. □

Theorem 16.128. (General Chain Rule) *Let $n \in \mathbb{N}$, $\langle X, \|\|_X \rangle$, $\langle Y, \|\|_Y \rangle$, $\langle Z, \|\|_Z \rangle$ be normed spaces, U a open set in X, $x \in U$, V a open set in Y and $f: U \rightarrow Y$, $g: V \rightarrow Z$ functions such that $f(U) \subseteq V$, f is n-times differentiable on x and g is n-times differentiable on $f(x)$ then*

$$g \circ f: U \rightarrow Y \text{ is } n\text{-times differentiable at } x$$

Proof. We prove this by induction on n, so define

$S = \{n \in \mathbb{N} | \text{If } \langle X, \|\|_X \rangle, \langle Y, \|\|_Y \rangle, \langle Z, \|\|_Z \rangle$ are normed spaces, U a open set in X, $x \in U$, V a open set in Y, $f: U \rightarrow Y$, $g: V \rightarrow Z$ such that $f(U) \subseteq V$, f is n-times differentiable at x and g is n-times differentiable at $f(x)$ then $g \circ f$ is n-times differentiable at $x\}$

then we have:

$1 \in S$. As f is 1-times differentiable at x and g is 1-times differentiable at $f(x)$ we have by definition that f is Fréchet differentiable at x and g is Fréchet differentiable at $f(x)$. So using [theorem: 16.36] it follows that $g \circ f$ is Fréchet differentiable at x. Hence $g \circ f$ is 1-times differentiable at x.

$\boldsymbol{n \in S \Rightarrow n+1 \in S.}$ As f is $(n+1)$-times differentiable, it follows from [theorem: 16.106] that

$$f \text{ is } n\text{-times differentiable at } x \tag{16.120}$$

and using [theorem: 16.109] that there exist a open V_1 in X with $x \in V_1 \subseteq U$ such that $\forall y \in V_1$ we have that f is 1-times differentiable at y and that the function

$$D^{[1]}_{V_1} f: V_1 \rightarrow L(X, Y) \text{ defined by } D^{[1]}_{V_1} f(y) = D^{[1]}_y f \underset{\text{def}}{=} D_y f \text{ is } n\text{-times differentiable at } x$$

Likewise, as g is $(n+1)$-times differentiable at $f(x)$ there exist a open set V_2 with $f(x) \in V_2 \subseteq V$ such that $\forall z \in V_2$ we have that g is 1-times differentiable at z and that the function

$$D^{[1]} g: V_2 \rightarrow L(Y, Z) \text{ where } D^{[1]} g(z) = D^{[1]}_z g \underset{\text{def}}{=} D_x g \text{ is } n\text{-times differentiable at } f(x) \tag{16.121}$$

As f is Fréchet differentiable at x [because f is 1-times differentiable at x] f is by [theorem: 16.31] continuous at x, hence as $f(x) \in V_2$ there exist a open U' with $x \in U'$ such that $f(U') \subseteq V_2$. Take $W = V_1 \bigcap U'$ then $x \in W \subseteq U$ and $f(W) \subseteq f(U') \subseteq V_2$. So for the function

$$D^{[1]} f: W \rightarrow L(X, Y) \text{ defined by } D^{[1]} f(y) = D^{[1]}_y f = D_y f$$

we have that $D^{[1]} f = (D^{[1]}_{V_1})_{|W}$ so that by [theorem: 16.90]

$$D^{[1]} f \text{ is } n\text{-times differentiable at } x \tag{16.122}$$

As $\forall y \in W$ $y \in V_1$, $f(y) \in f(W) \subseteq V_2$, f is 1-times differentiable at y [hence Fréchet differentiable at y] and g is 1-times differentiable at $f(y)$ [hence Fréchet differentiable at $f(y)$] it follows from [theorem: 16.36] that $g \circ f$ is Fréchet differentiable at y [hence 1-times differentiable at y] with

$$D^{[1]}_y (g \circ f) = D_y (g \circ f) = D_{f(y)} g \circ D_y f = D^{[1]}_{f(y)} g \circ D^{[1]}_y f$$

This allows us to define

$$D^{[1]}(g \circ f): W \rightarrow L(X, Z) \text{ by } D^{[1]}(g \circ f)(y) = D^{[1]}_y (g \circ f) = D^{[1]}_{f(y)} g \circ D^{[1]}_y f \tag{16.123}$$

As $f(W) \subseteq V_2$ we can define

$$h: W \rightarrow L(Y, Z) \cdot L(X, Y) \text{ by } h(y) = \left(D^{[1]}_{f(y)} g, D^{[1]}_y f\right)$$

then we have for $y \in W$ that

$$(\pi_1 \circ h)(y) \;=\; \pi_1(h(y)) = \pi_1\left(\left(D^{[1]}_{f(y)} g, D^{[1]}_y f\right)\right) = D^{[1]}_{f(y)} = \left(D^{[1]} g\right)(f(y)) = ((D^{[1]} g) \circ f)(y)$$

and

$$(\pi_2 \circ h)(y) = \pi_2(h(y)) = \pi_2\left(\left(D^{[1]}_{f(y)} g, D^{[1]}_y f\right)\right) = D^{[1]}_y f = (D^{[1]} f(y)$$

proving that

$$\pi_1 \circ h = (D^{[1]} g) \circ f \text{ and } \pi_2 \circ h = D^{[1]} f$$

By [eq: 16.120, 16.122] $D^{[1]} f$ is n-times differentiable at $f(x)$ and f is n-times differentiable at x so as $n \in S$ it follows that $(D^{[1]} g) \circ f$ is n-times differentiable at x, further by [eqs: 16.122] $\pi_2 \circ h$ is continuous. Hence $\pi_1 \circ h$ and $\pi_2 \circ h$ are n-times differentiable at x. So using [theorem: 16.125] it follows that

$$h \text{ is } n\text{-times differentiable at } x \tag{16.124}$$

Let $y \in W$ then

$$\begin{aligned} ((\circ) \circ h)(y) &= (\circ)(h(y)) \\ &= \left(\circ\left(\left(D^{[1]}_{f(y)} g, D^{[1]}_y f\right)\right)\right) \\ &= D^1_{f(y)} g \circ D^{[1]}_y f \\ &\underset{\text{[eq: 16.123]}}{=} D^{[1]}(g \circ f)(y) \end{aligned}$$

proving that

$$D^{[1]}(g \circ f) = (\circ) \circ h$$

As by [lemma: 16.127] $(\circ)$ is n-times differentiable on $L(Y, Z) \times L(X, Y)$ and h is n-times differentiable on x it follows that $(\circ) \circ h$ is n-times differentiable at x. Hence, as $D^{[1]}(g \circ f) = (\circ) \circ h$, it follows that $D^{[1]}(g \circ f)$ is n-times differentiable at x. So using [theorem: 16.109] it follows that $g \circ f$ is $(n+1)$-times differentiable at x proving that

$$n + 1 \in S$$

□

We can use the same proof technique as above to prove that the composition of functions of class C^n is of class C^n.

Theorem 16.129. *Let $n\in\mathbb{N}_0$, $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$, $\langle Z,\|\|_Z\rangle$ be normed spaces, U a open set in X, $x\in U$, V a open set in Y and $f:U\rightarrow Y$, $g:V\rightarrow Z$ functions such that $f(U)\subseteq V$, f is of class C^n and g is of class C^n then $g\circ f$ is of class C^n.*

Proof. If $n=0$ then as f,g are of class C^0 f,g are continuous functions it follows from [theorem: 14.138] that $g\circ f$ is a continuous function, hence $g\circ f$ is of class C^0. This proves the theorem for $n=0$ so we must only prove the theorem for the case $n\in\mathbb{N}$. We proceed by induction so define

$S=\{n\in\mathbb{N}|$If $\langle X,\|\|_X\rangle,\langle Y,\|\|_Y\rangle,\langle Z,\|\|_Z\rangle$ are normed spaces, U a open set in X, $x\in U$, V a open set in Y, $f:U\rightarrow Y$, $g:V\rightarrow Z$ such that $f(U)\subseteq V$ and f,g are of class C^n then $g\circ f$ is of class $C^n\}$

then we have:

$1\in S$. As f is of class C^1 we have that $\forall x\in U$ f is 1-times differentiable at x [hence Fréchet differentiable at x] and the function

$$D^{[1]}f:U\rightarrow L(X,Y) \text{ defined by } D^{[1]}f(x)=D^{[1]}_x f=D_x f \text{ is continuous} \tag{16.125}$$

As g is of class C^1 we have that $\forall y\in V$ g is 1-times differentiable at y [hence Fréchet differentiable at y[and the function

$$D^{[1]}g:V\rightarrow L(Y,Z) \text{ defined by } D^{[1]}g(x)=D^{[1]}_x g=D_x g \text{ is continuous} \tag{16.126}$$

Using the chain rule [see theorem: 16.36] it follows that $g\circ f$ is Fréchet differentiable at x and $D_x(g\circ f)=D_{f(x)}\,g\circ D_x f$. Hence $g\circ f$ is 1-times differentiable at x and

$$D^{[1]}_x(g\circ f)=D_x(g\circ f)=D_{f(x)}\,g\circ D_x f=D^{[1]}_{f(x)}g\circ D^{[1]}_x f \tag{16.127}$$

Consider the following functions

$$D^{[1]}(g\circ f):U\rightarrow L(X,Z) \text{ by } D^{[1]}(g\circ f)(x)=D^{[1]}_x(g\circ f) \tag{16.128}$$

$$h:U\rightarrow L(Y,Z)\cdot L(X,Y) \text{ by } h(x)=\left(D^{[1]}_{f(x)}g,D^{[1]}_x f\right)$$

then we have for $x\in U$ that

$$\begin{aligned}
(\pi_1\circ h)(x) &= \pi_1(h(x))\\
&= \pi_1\left(D^{[1]}_{f(x)}g,D^{[1]}_x f\right)\\
&= D^{[1]}_{f(x)}g\\
&= D^{[1]}g(f(x))\\
&= ((D^{[1]}g)\circ f)(x)\\
(\pi_2\circ h)(x) &= \pi_2(h(x))\\
&= \pi_2\left(D^{[1]}_{f(x)}g,D^{[1]}_y f\right)\\
&= D^{[1]}_x f\\
&= D^{[1]}f(x)
\end{aligned}$$

proving that

$$\pi_1\circ h=(D^{[1]}g)\circ f \text{ and } \pi_2\circ h=D^{[1]}f \tag{16.129}$$

Now as $\forall x\in U$ f is Fréchet differentiable at x, it follows from [theorem: 16.31] that f is a continuous function, further by [eq: 16.126] $D^{[1]}g$ is continuous, so using [theorem: 14.138] $\pi_1\circ h$ is continuous. Further by [eq: 16.125] $D^{[1]}f$ is also continuous, hence $\pi_2\circ h$ is continuous. So it follows from [theorem: 14.140] that

$$h \text{ is continuous} \tag{16.130}$$

Now given $x\in U$ we have

$$\begin{aligned}
((\circ)\circ h)(x) &= (\circ)(h(x))\\
&= (\circ)\left(D^{[1]}_{f(x)}g,D^{[1]}_x f\right)\\
&= D^{[1]}_{f(x)}g\circ D^{[1]}_x f\\
&\underset{[\text{eq: } 16.127]}{=} D^{[1]}(g\circ f)(x)
\end{aligned}$$

proving that

$$(\circ)\circ h=D^{[1}(g\circ f) \tag{16.131}$$

By [lemma: 16.127] $(\circ)$ is of class C^0 hence continuous, by [eq: 16.130] h is continuous so that $(\circ)\circ h$ is continuous or using the above that $D^{[1]}(g\circ f)$ is continuous, hence $g\circ f$ is of class C^1. This prove that $1\in S$.

$\boldsymbol{n\in S\Rightarrow n+1\in S.}$ As f is of class C^{n+1} it follows from [theorem: 16.107] that

$$f\text{ is of class }C^n \tag{16.132}$$

Further using [theorem: 16.111] it follows that $\forall x\in U$ f is 1-times differentiable at x [Fréchet differentiable at x] and that

$$D^{[1]}f\colon U\rightarrow L(X,Y)\text{ defined by }D^{[1]}f(x)=D_x^{[1]}f=D_xf\text{ is of class }C^{[n]} \tag{16.133}$$

As g is of class C^{n+1} we can use [theorem: 16.107] again so that $\forall y\in V$ g is 1-times differentiable at y and that

$$D^{[1]}g\colon V\rightarrow L(Y,Z)\text{ defined by }D^{[1]}g(y)=D_y^{[1]}g=D_1g\text{ is of class }C^n \tag{16.134}$$

Let $x\in U$ then $f(x)\in f(U)\subseteq V$, as f is Fréchet differentiable at x and g is Fréchet differentiable at $f(x)$ it follows from the chain rule [theorem: 16.36] that $g\circ f$ is Fréchet differentiable at x and that $D_x(g\circ f)=D_{f(x)}\,g\circ D_xf$. Hence $g\circ f$ is 1-times differentiable at x and

$$D_x^{[1]}(g\circ f)=D_{f(x)}^{[1]}\,g\circ D_x^{[1]}f \tag{16.135}$$

Consider now the functions

$$D^{[1]}(g\circ f)\colon U\rightarrow L(X,Z)\text{ defined by }D^{[1]}(g\circ f)(x)=D_x^{[1]}(g\circ f) \tag{16.136}$$

$$h\colon U\rightarrow L(Y,Z)\cdot L(X,Y)\text{ defined by }h(x)=\left(D_{f(x)}^{[1]}g,D_x^{[1]}f\right)$$

then we have for $x\in U$

$$\begin{aligned}
(\pi_1\circ h)(x) &= \pi_1(h(x))\\
&= \pi_1\left(\left(D_{f(x)}^{[1]}g,D_x^{[1]}f\right)\right)\\
&= D_{f(x)}^{[1]}g\\
&= D^{[1]}g(f(x))\\
&= ((D^{[1]}g)\circ f)(x)\\
(\pi_2\circ h)(x) &= \pi_2(h(x))\\
&= \pi_2\left(\left(D_{f(x)}^{[1]}g,D_x^{[1]}f\right)\right)\\
&= D_x^{[1]}f\\
&= D^{[1]}f(x)\\
&= (D^{[1]}f)(x)
\end{aligned}$$

which proves that

$$\pi_1\circ h=(D^{[1]}g)\circ f\text{ and }\pi_2\circ h=D^{[1]}f \tag{16.137}$$

By [eqs: 16.134, 16.132] we have that $(D^{[1]}g)\circ f$, f are of class C^n so, as $n\in S$ it follows that $\pi_1\circ h$ is of class C^n, further by [eq: 16.133] $\pi_2\circ h$ is of class C^n. Using then [lemma: 16.126] it follows that

$$h\text{ is of class }C^n \tag{16.138}$$

Let $x\in U$ then we have

$$\begin{aligned}
((\circ)\circ h)(x) &= (\circ)(h(x))\\
&= (\circ)\left(\left(D_{f(z)}^{[1]}g,D_x^{[1]}f\right)\right)\\
&= D_{f(x)}^{[1]}g\circ D_x^{[1]}f\\
&\underset{\text{[eq: 16.135]}}{=} D_x^{[1]}(g\circ f)\\
&= (D^{[1]}(g\circ f))(x)
\end{aligned}$$

proving that

$$D^{[1]}(g\circ f)=(\circ)\circ h$$

Now using [lemma: 16.127] ($\circ$) is of class C^n and by [eq: 16.138] h is of class C^n, hence, as $n \in S$, it follows that $D^{[1]}(g \circ f)$ is of class C^n. Using [theorem: 16.111] it follows then that $g \circ f$ is of class C^{n+1} which proves that

$$n + 1 \in S$$

□

Theorem 16.130. *Let $n \in \mathbb{N}$, $\langle X, \|\|\rangle$ be a normed space, $\langle \mathbb{K}, ||\rangle$ be the normed space of real (complex) numbers, U a open set in X, $x \in U$ and $f: U \to X$, $g: U \to X$ functions that are n-times differentiable at x then*

$$f \cdot g: U \to \mathbb{K}$$

is n-times differentiable at x.

Proof. Define

$$\varphi: U \to \mathbb{K} \cdot \mathbb{K} \text{ by } \varphi(x) = (f(x), g(x))$$

then we have for $x \in U$ that

$$\begin{array}{rcl} (\pi_1 \circ \varphi)(x) & = & \pi_1(\varphi(x)) \\ & = & \pi_1((f(x), g(x))) \\ & = & f(x) \\ (\pi_2 \circ \varphi)(x) & = & \pi_2(\varphi(x)) \\ & = & \pi_2((f(x), g(x))) \\ & = & g(x) \end{array}$$

proving that

$$\pi_1 \circ \varphi = f \text{ and } \pi_2 \circ \varphi = g$$

Using [theorem: 16.125] it follows that

$$\varphi \text{ is } n\text{-times differentiable at } x$$

Consider the function

$$(\cdot): \mathbb{K} \times \mathbb{K} \to \mathbb{K} \text{ by } (\cdot)(r, s) = r \cdot s$$

then we have for $r, s, t, \alpha \in \mathbb{K}$ that

$$(\cdot)(r + \alpha \cdot s, t) = (r + \alpha \cdot s) \cdot t = r \cdot t + \alpha \cdot (s \cdot t) = (\cdot)(r, t) + \alpha \cdot (\cdot)(s, t)$$

and

$$(\cdot)(r, s + \alpha \cdot t) = r \cdot (s + \alpha \cdot t) = r \cdot s + \alpha \cdot (r \cdot t) = (\cdot)(r, s) + \alpha \cdot (\cdot)(r, t)$$

proving that

$$(\cdot) \in \mathrm{Hom}(\mathbb{K}, \mathbb{K}; \mathbb{K})$$

Further we have for $r, s \in \mathbb{K}$ that

$$|(\cdot)(r, s)| = |r \cdot s| = |r| \cdot |t|$$

proving by [theorem: 14.187] that $(\cdot) \in L(\mathbb{K}, \mathbb{K}; \mathbb{K})$. Hence by [theorem: 16.120]

$$(\cdot) \text{ is } \infty\text{-times differentiable on } \mathbb{K}$$

Let $x \in U$ then

$$\begin{array}{rcl} ((\cdot) \circ \varphi)(x) & = & (\cdot)(\varphi(x)) \\ & = & (\cdot)(f(x), g(x)) \\ & = & f(x) \cdot g(x) \\ & = & (f \cdot g)(x) \end{array}$$

proving that $(\cdot) \circ \varphi = f \cdot g$. So using the generalized chain rule [theorem: 16.128] it follows that

$$f \cdot g \text{ is } n\text{-times differentiable at } x$$

□

Theorem 16.131. *Let $n \in \mathbb{N}_0$, $\langle X, \|\|\rangle$ be a normed space, $\langle \mathbb{K}, ||\rangle$ be the normed space of real (complex) numbers, U a open set in X and $f: U \to X$, $g: U \to X$ functions that are of class C^n then*

$$f \cdot g: U \to \mathbb{K}$$

is of class C^n.

16.3 Intermediate value and main value theorems

16.3.1 Derivatives extremums, concavity, and convexity

Definition 16.132. *Let* $\langle X, \|\|\rangle$ *be a normed space,* $\langle \mathbb{R}, ||\rangle$ *be the normed space of real numbers,* U *a open set in* X *and* $f: U \rightarrow \mathbb{R}$ *then*

1. $x \in U$ *is a* ***local weak minimum of f*** *if there exist a* $\delta \in \mathbb{R}^+$ *such that* $B_{\|\|}(x, \delta) \subseteq U$ *and* $\forall y \in B_{\|\|}(x, \delta)$ $f(x) \leqslant f(y)$.
2. $x \in U$ *s a* ***local weak maximum of f*** *if there exist a* $\delta \in \mathbb{R}^+$ *such that* $B_{\|\|}(x, \delta) \subseteq U$ *and* $\forall y \in B_{\|\|}(x, \delta)$ $f(y) \leqslant f(x)$.

A ***local weak minimum*** *or* ***local weak maximum*** *of* f *is called a* ***local extremum****.*

Theorem 16.133. *Let* $\langle \mathbb{R}, ||\rangle$ *be the normed space of real numbers,* U *a open set in* $\mathbb{R}$ *and*

$$f: U \rightarrow \mathbb{R}$$

a function with a local extremum x *then if* f *has a derivative at* x *it follows that* $f'(x) = 0$.

Proof. If x is a local extremum of f then we have either:

x is a local weak minimum. Then there exist a $\delta \in \mathbb{R}^+$ such that $]x - \delta, x + \delta[\subseteq U$ and $\forall y \in]x - \delta, x + \delta[$ we have $f(x) \leqslant f(y)$. Assume that $f'(x) \neq 0$ then we have either:

$0 < f'(x)$. As $f'(x) \in \mathbb{R}^+$ there exist a $\delta_{f'(x)} \in \mathbb{R}^+$ such that

$$\forall h \in U_x \text{ with } 0 < |h| < \delta_{f'(x)} \text{ we have } \left|\frac{f(x+h) - f(x)}{h} - f'(x)\right| < f'(z) \tag{16.139}$$

Let $k = \frac{\min(\delta, \delta_{f'(x)})}{2} \in \mathbb{R}^+$ then $0 < k = |-k| = k < \delta, \delta_{f'(x)}$ so that $x - \delta < x - k < x + \delta$ or $x - k \in]x - \delta, x + \delta[$ so that

$$f(x) \leqslant f(x-k) \Rightarrow 0 \leqslant f(x-k) - f(x) \Rightarrow 0 \leqslant -\left(\frac{f(x-k) - f(x)}{-k}\right) \tag{16.140}$$

As $x - k \in]x - \delta, x + \delta[\subseteq U$ we have that $-k \in U_x$ so that by [eq: 16.139]

$$\left|\frac{f(x-k) - f(x)}{-k} - f'(x)\right| < f'(x)$$

combining this with [eq: 16.140]

$$f'(x) \leqslant f'(x) - \frac{f(x-k) - f(x)}{-k} \leqslant \left|\frac{f(x-k) - f(x)}{-k} - f'(x)\right| < f'(x)$$

leading to the contradiction $f'(x) < f'(x)$.

$f'(x) < 0$. As $-f'(x) \in \mathbb{R}^+$ there exist $\delta_{-f'(x)} \in \mathbb{R}^+$ such that we have

$$\forall h \in U_x \text{ with } 0 < |h| < \delta_{-f'(x)} \text{ we have } \left|\frac{f(x+h) - f(x)}{h} - f'(x)\right| < -f'(x) \tag{16.141}$$

Let $k = \frac{\min(\delta, \delta_{-f'(x)})}{2} \in \mathbb{R}^+$ then $0 < k = |k| = k < \delta, \delta_{-f'(x)}$ so that $x - \delta < x < x + k < x + \delta$ or $x + k \in]x - \delta, x + \delta[$ so that

$$f(x) \leqslant f(x+k) \Rightarrow 0 \leqslant f(x+k) - f(x) \Rightarrow 0 \leqslant \frac{f(x+k) - f(x)}{k} \tag{16.142}$$

As $x + k \in]x - \delta, x + \delta[\subseteq U$ we have that $k \in U_x$ so that by [eq: 16.141]

$$\left|\frac{f(x+k) - f(x)}{k} - f'(x)\right| < -f'(x)$$

Hence using [eq: 16.142]

$$-f'(x) \leqslant -f'(x) + \frac{f(x+k) - f(x)}{k} \leqslant \left|\frac{f(x+k) - f(x)}{k} - f'(x)\right| < -f'(x)$$

leading to the contradiction $-f'(x) < -f'(x)$.

As in all case we have a contradiction the assumption is wrong, hence we must have that

$$f'(x) = 0.$$

***x* is a local weak maximum.** Then there exist a $\delta \in \mathbb{R}^+$ such that $]x-\delta, x+\delta[\subseteq U$ and $\forall y \in]x-\delta, x+\delta[$ we have $f(y) \leqslant f(x)$. Assume that $f'(x) \neq 0$ then we have either:

$0 < f'(x)$. As $f'(x) \in \mathbb{R}^+$ there exist a $\delta_{f'(x)} \in \mathbb{R}^+$ such that

$$\forall h \in U_x \text{ with } 0 < |h| < \delta_{f'(x)} \text{ we have } \left|\frac{f(x+h) - f(x)}{h} - f'(x)\right| < f'(x) \tag{16.143}$$

Let $k = \frac{\min(\delta, \delta_{f'(x)})}{2} \in \mathbb{R}^+$ then $0 < k = |k| = k < \delta, \delta_{f'(x)}$ so that $x - \delta < x < x + k < x + \delta$ or $x + k \in]x - \delta, x + \delta[$ so that

$$f(x+k) \leqslant f(x) \Rightarrow 0 \leqslant -(f(x+k) - f(x)) \Rightarrow 0 \leqslant -\frac{f(x+k) - f(x)}{k} \tag{16.144}$$

As $x + k \in]x - \delta, x + \delta[\subseteq U$ we have that $k \in U_x$ so that by [eq: 16.143]

$$\left|\frac{f(x+k) - f(x)}{k} - f'(x)\right| < f'(x)$$

Hence using [eq: 16.144]

$$f'(x) \leqslant f'(x) - \frac{f(x+k) - f(x)}{k} \leqslant \left|\frac{f(x+k) - f(x)}{k} - f'(x)\right| < f'(x)$$

leading to the contradiction $f'(x) < f'(x)$.

$f'(x) < 0$. As $-f'(x) \in \mathbb{R}^+$ there exist a $\delta_{-f'(x)} \in \mathbb{R}^+$ such that

$$\forall h \in U_x \text{ with } 0 < |h| < \delta_{-f'(x)} \text{ so that } \left|\frac{f(x+h) - f(x)}{h} - f'(x)\right| < -f'(x) \tag{16.145}$$

Let $k = \frac{\min(\delta, \delta_{f'(x)})}{2}$ then $0 < k = |-k| = k < \delta, \delta_{-f'(x)}$ so that $x - \delta < x - k < x < x + \delta$ or $x - k \in]x - \delta, x + \delta[$ so that

$$f(x-k) \leqslant f(x) \Rightarrow 0 \leqslant -(f(x-k) - f(x)) \Rightarrow 0 \leqslant \frac{f(x-k) - f(x)}{-k} \tag{16.146}$$

As $x - k \in]x - \delta, x + \delta[\subseteq U$ we have that $-k \in U_x$ so that by [eq: 16.145]

$$\left|\frac{f(x-k) - f(x)}{-k} - f'(x)\right| < -f'(x)$$

Hence using [eq: 16.146]

$$-f'(x) \leqslant -f'(x) + \frac{f(x-k) - f(x)}{-k} \leqslant \left|\frac{f(x-k) - f(x)}{-k} - f'(x)\right| < -f'(x)$$

leading to the contradiction that $-f'(x) < -f'(x)$

As in all case we have a contradiction the assumption is wrong, hence we must have that

$$f'(x) = 0.$$

□

Theorem 16.134. (Rolle's theorem) *Let $\langle\mathbb{R}, \|\rangle$ be the normed space of real numbers, $a, b \in \mathbb{R}$ with $a < b$ and $f \in \mathcal{C}([a, b], \mathbb{R})$ [see definition: 14.129] such that $f(a) = f(b)$ and $f_{|]a,b[}$ has a derivative at every $x \in]a, b[$ [or equivalent $f_{|]a,b[}$ is Fréchet differentiable on $]a, b[$] then there exist a $\zeta \in]a, b[$ such that $(f_{|]a,b[})'_\zeta = 0$.*

Proof. Take $k = f(a) = f(b)$. As f is continuous on $[a, b]$ we can use [theorem: 14.397] to find $c, d \in \mathbb{R}$ such that

$$f([a, b]) = [c, d]$$

Hence as $a \in [a, b]$ it follows that $k = f(a) \in [c, d]$ so that $c \leqslant k \leqslant d$ and we have the following cases to consider:

***c* = *d* = *k*.** Then $\forall x \in [a, b]$ we have $f(x) \in [k, k] = \{k\}$ so that $f_{|]a,b[} = C_k$. As $a < b$ there exist a $\zeta \in]a, b[$ and we have for ζ

$$(f_{|]a,b[})'_\zeta \underset{[\text{theorem: } 16.32]}{=} D_\zeta C_k(1) \underset{[\text{theorem: } 16.26]}{=} C_0(1) = 0$$

$c<k$. As $f([a,b])=[c,d]$ there exist a $\zeta\in[a,b]$ such that $f(\zeta)=c<k=f(a)=f(b)$ so that $\zeta\neq a,b$, from which it follows that $\zeta\in]a,b[$. As $]a,b[$ is open there exist a $\delta\in\mathbb{R}^+$ such that $\zeta\in B_{||}(\zeta,\delta)\subseteq]a,b[$. Further $\forall x\in B_{||}(\zeta,\delta)$ we have $f(x)\in f([a,b])=[c,d]$ so that $f(\zeta)=c\leqslant f(x)$ which proves that ζ is a local weak minimum of $f_{|]a,b[}$. Hence using [theorem: 16.133] it follows that $(f_{]a,b[})'_\zeta=0$.

$k<d$. As $f([a,b])=[c,d]$ there exist a $\zeta\in[a,b]$ such that $f(\zeta)=d>k=f(a)=f(b)$ so that $\zeta\neq a,b$ from which it follows that $\zeta\in]a,b[$. As $]a,b[$ is open there exist a $\delta\in\mathbb{R}^+$ such that $\zeta\in B_{||}(\zeta,\delta)\subseteq]a,b[$. Further $\forall x\in B_{||}(\zeta,\delta)$ we have $f(x)\in f([a,b])=[c,d]$ so that $f(x)\leqslant d=f(\zeta)$ which proves that ζ is a local weak maximum of $f_{|]a,b[}$. Hence using [theorem: 16.133] it follows that $(f_{]a,b[})'_\zeta=0$. □

Theorem 16.135. (Lagrange's Mean Value theorem) *Let $\langle\mathbb{R},\|\rangle$ be the normed space of real numbers, $a,b\in\mathbb{R}$ with $a<b$ and $f\in\mathcal{C}([a,b],\mathbb{R})$ such that $f_{|]a,b[}$ has a derivative at every $x\in]a,b[$ [or equivalent $f_{|]a,b[}$ is Fréchet differentiable at $]a,b[$] then there exist a $\xi\in]a,b[$ such that*

$$f(b)-f(a)=(f_{|]a,b[})'_\zeta\cdot(b-a)$$

Proof. *Define*

$$g:[a,b]\rightarrow\mathbb{R} \text{ by } g(x)=f(x)-\left(f(a)+\frac{f(b)-f(a)}{b-a}\cdot(x-a)\right)$$

then by [examples: 14.131, 14.132] and [theorems: 14.144, 14.133] it follows that

$$g\in\mathcal{C}([a,b],\mathbb{R})$$

Further using [examples: 16.26, 16.22] and [theorems: 16.32, 16.35, 16.24] it follows that

$$\forall x\in]a,b[\ g \text{ is Fréchet differentiable at } x$$

Now

$$\begin{aligned} g(a) &= f(a)-\left(f(a)+\frac{f(b)-f(a)}{b-a}\cdot(a-a)\right)=f(a)-f(a)=0\\ g(b) &= f(b)-\left(f(a)+\frac{f(b)-f(a)}{b-a}\cdot(b-a)\right)=f(b)-f(a)+f(b)-f(a)=0\end{aligned}$$

so we can apply Rolle's theorem [theorem: 16.134] to find a $\zeta\in]a,b[$ such that

$$(g_{|]a,b[})'_\zeta=0$$

As

$$(g_{|]a,b[})'_\zeta = (f_{|]a,b[})'_\zeta-\frac{f(b)-f(a)}{b-a}$$

it follows that

$$f(b)-f(a)=(f_{|]a,b[})'_\zeta\cdot(b-a)$$ □

Corollary 16.136. *Let $\langle\mathbb{R},\|\rangle$ be the normed space of real numbers, U a open set in $\mathbb{R}$, $a,b\in\mathbb{R}$ with $a<b$ such that $[a,b]\subseteq U$ and f: $U\rightarrow\mathbb{R}$ a function that has a derivative at every $x\in U$ [hence is Fréchet differentiable at every $x\in U$] then there exist a $\xi\in]a,b[$ such that*

$$f(b)-f(a)=f'_\zeta\cdot(b-a)$$

Proof. Let $x\in U$ then, as f has a derivative at every $x\in U$, it follows from [theorem: 16.32] that f is Fréchet differentiable at x, hence continuous at x [see theorem: 16.31] proving that f is continuous. Hence using [theorem: 14.135]

$$f_{|[a,b]}:[a,b]\rightarrow\mathbb{R} \text{ is continuous or } f_{|[a.b]}\in\mathcal{C}([a,b],\mathbb{R})$$

Further as $]a,b[\subseteq[a,b]\subseteq U$ it follows from the locality of the differential [see theorem: 16.24] it follows that $\forall x\in]a,b[$ we have that $f_{|]a,b[}$ is Fréchet differentiable at x and $D_x(f_{]a,b[})=D_xf$. As $f_{|]a,b[}=(f_{|[a,b]|})_{|]a,b[}$ it follows that

$$\forall x\in]a,b[\text{ we have that } (f_{|[a,b]})_{|]a,b[} \text{ is Fréchet differentiable at } x \text{ and } D_x((f_{|[a,b]|})_{|]a,b[})=D_xf$$

Using then [theorem: 16.32] it follows that

$$\forall x\in]a,b[\text{ we have that } (f_{|[a,b]})_{|]a,b[} \text{ has a derivative at } x \text{ and } ((f_{|[a,b]|})_{|]a,b[})'_x=f'_x \tag{16.147}$$

By Lagrange's theorem [theorem: 16.135] there exist a $\zeta\in]a,b[$ such that

$$f_{|[a,b]}(b)-f_{|[a,b]}(a)=((f_{|[a,b]|})_{|]a,b[})'_\zeta\cdot(b-a) \tag{16.148}$$

Hence

$$\begin{aligned} f(b)-f(a) & \underset{a,b\in[a,b]}{=} f_{|[a,b]}(b)-f_{|[a,b]}(a) \\ & \underset{[\text{eq: } 16.148]}{=} ((f_{|[a,b]|})_{|]a,b[})'_\zeta\cdot(b-a) \\ & \underset{[\text{eq: } 16.147]}{=} f'_\zeta\cdot(b-a) \end{aligned}$$

□

Corollary 16.137. *Let $\langle\mathbb{R},\|\rangle$ be the normed space of the real numbers, U a open set in $\mathbb{R}$, I a generalized interval [see definition: 3.133] such that $I\subseteq U$ and*

$$f:U\rightarrow\mathbb{R}$$

a function that is differentiable on U then we have:

1. *If $\forall x\in I$ $0<f'_x$ then f is strictly increasing on I*
2. *If $\forall x\in I$ $0\leqslant f'_x$ then f is increasing on I.*
3. *If $\forall x\in I$ $f'_x<0$ then f is strictly decreasing on I.*
4. *If $\forall x\in I$ $f'_x\leqslant 0$ then f is decreasing on I.*

Proof. First as f is differentiable on U we have by [theorem: 16.31] that f is continuous on U. If $x,y\in I$ with $x<y$ then we have, as I is a generalized interval] that $[x,y]\subseteq I\subseteq U$ so that by [theorem: 14.133] $f_{|[x,y]}:[x,y]\rightarrow\mathbb{R}$ is a continuous function. Further by [theorem: 16.24] it follows that $f_{|]x,y[}$ is differentiable on $]x,y[$. So we can apply Lagrange's theorem [theorem: 16.135] to find a $\zeta\in]x,y[$ such that $f(y)-f(x)=(f_{|]a,b[})'_\zeta\cdot(y-x)\underset{[\text{theorem: } 16.33]}{=} f'_\zeta\cdot(y-x)$. Summarized

$$\forall x,y\in I \text{ with } x<y \text{ there exist a } \zeta\in]x,y[\text{ such that } f(y)-f(x)<f'_\zeta\cdot(y-x) \tag{16.149}$$

1. Let $x,y\in I$ with $x<y$. As $\forall y\in I$ $0<f'(y)$ and $[x,y]\subseteq I$, we have that $0<f'_\zeta$, so that
$$f(y)-f(x)\underset{[\text{eq: } 16.149]}{=} f'_\zeta\cdot(y-x)>0$$
proving that $f(y)>f(x)$.
2. Let $x,y\in I$ with $x<y$. As $\forall y\in I$ $0\leqslant f'(y)$ and $[x,y]\subseteq I$, we have that $0\leqslant f'_\zeta$, so that
$$f(y)-f(x)\underset{[\text{eq: } 16.149]}{=} f'_\zeta\cdot(y-x)\geqslant 0$$
proving that $f(y)\geqslant f(x)$.
3. Let $x,y\in I$ with $x<y$. As $\forall y\in I$ $f'(y)<0$ and $[x,y]\subseteq I$, we have that $f'_\zeta<0$, so that
$$f(y)-f(x)\underset{[\text{eq: } 16.149]}{=} f'_\zeta\cdot(y-x)<0$$
proving that $f(y)<f(x)$.
4. Let $x,y\in I$ with $x<y$. As $\forall y\in I$ $f'(y)\leqslant 0$ and $[x,y]\subseteq I$, we have that $f'_\zeta\leqslant 0$, so that
$$f(y)-f(x)\underset{[\text{eq: } 16.149]}{=} f'_\zeta\cdot(y-x)\leqslant 0$$
proving that $f(y)\leqslant f(x)$.

□

Definition 16.138. *Let $\langle\mathbb{R},\|\rangle$ be the normed space of the real numbers, U a open set in $\mathbb{R}$, $a,b\in\mathbb{R}$ with $a<b$ and $[a,b]\subseteq U$ and*

$$f:U\rightarrow\mathbb{R}$$

a function then we say:

1. *f is concave on $[a,b]$ if $\forall x\in[a,b]$ we have*
$$f(x)\leqslant\frac{f(a)\cdot(b-x)+f(b)\cdot(x-a)}{b-a}=f(a)+\frac{f(b)-f(a)}{b-a}\cdot(x-a)$$
2. *f is convex on $[a,b]$ if $\forall x\in[a,b]$ we have*
$$f(a)+\frac{f(b)-f(a)}{b-a}\cdot(x-a)=\frac{f(a)\cdot(b-x)+f(b)\cdot(x-a)}{b-a}\leqslant f(x)$$

In other words a function is concave on a interval if the function curve lies below the line connecting $(a, f(a))$ and $(b, f(b))$ and convex on a interval if the function curve lies above the line connecting $(a, f(a))$ and $(b, f(b))$.

Lemma 16.139. *Let $\langle \mathbb{R}, \|\| \rangle$ be the normed space of the real numbers, U a open set in $\mathbb{R}$, $f: U \rightarrow \mathbb{R}$ a function, $a, b \in \mathbb{R}$ with $a < b$ and $[a, b] \subseteq U$ then we have:*

1. *$\forall \alpha, \beta \in [0, 1]$ with $\alpha + \beta = 1$ we have $\alpha \cdot a + \beta \cdot b \in [a, b]$*
2. *If $x \in [a, b]$ then for $\alpha = \frac{x-a}{b-a}$, $\beta = \frac{b-x}{b-a}$ we have:*
   a. *$\alpha, \beta \in [0, 1]$*
   b. *$\alpha + \beta = 1$*
   c. *$x = \alpha \cdot a + \beta \cdot b$*
   d. *$\frac{f(a) \cdot (b-x) + f(b) \cdot (x-a)}{b-a} = \alpha \cdot f(a) + \beta \cdot f(b)$*
3. *If $\forall \alpha, \beta \in [0, 1]$ with $\alpha + \beta = 1$ $f(\alpha \cdot a + \beta \cdot b) \geqslant \alpha \cdot f(a) + \beta \cdot f(b)$ then f is convex on $[a, b]$*
4. *If $\forall \alpha, \beta \in [0, 1]$ with $\alpha + \beta = 1$ $f(\alpha \cdot a + \beta \cdot b) \leqslant \alpha \cdot f(a) + \beta \cdot f(b)$ then f is concave on $[a, b]$*

Proof.

1. We have as $\beta \leqslant 1$ and $0 < b - a$ we have

$$\begin{aligned} \beta \cdot (b-a) \leqslant (b-a) & \Rightarrow & \beta \cdot (b-a) + a \leqslant b \\ & \underset{\alpha+\beta=1}{\Rightarrow} & \beta \cdot (b-a) + (\alpha+\beta) \cdot a \leqslant b \\ & \Rightarrow & \alpha \cdot a + \beta \cdot b \leqslant b \end{aligned}$$

   Further as $a - b < 0$ and $\alpha \leqslant 1$ we have

$$\begin{aligned} a - b \leqslant \alpha \cdot (a-b) & \Rightarrow & a \leqslant \alpha \cdot (a-b) + b \\ & \Rightarrow & a \leqslant \alpha \cdot (a-b) + (\alpha+\beta) \cdot b \\ & \Rightarrow & a \leqslant \alpha \cdot a + \beta \cdot b \end{aligned}$$

   proving that

$$\alpha \cdot a + \beta \cdot b \in [a, b]$$

2. If $x \in [a, b] \Rightarrow a \leqslant x \leqslant b$ so that $0 \leqslant \frac{x-a}{b-a} \leqslant 1$ and $0 \leqslant \frac{b-x}{b-a} \leqslant 1$. So if we take $\alpha = \frac{x-\alpha}{b-a}$, $\beta = \frac{b-x}{b-a}$ then

$$\alpha, \beta \in [0, 1] \text{ and } \alpha + \beta = \frac{x-a}{b-a} + \frac{b-x}{b-a} = \frac{b-a}{b-a} = 1 \tag{16.150}$$

   Further we have

$$\begin{aligned} \alpha \cdot a + \beta \cdot b &= \frac{(b-x) \cdot a + (x-a) \cdot b}{b-a} \\ &= \frac{b \cdot a - x \cdot a + x \cdot b - a \cdot b}{b-a} \\ &= \frac{x \cdot (b-a)}{b-a} \\ &= x \end{aligned} \tag{16.151}$$

   Finally

$$\begin{aligned} \frac{f(a) \cdot (b-x) + f(b) \cdot (x-a)}{b-a} & \underset{[\text{eq: } 16.151]}{=} & \frac{f(a) \cdot (b - (\alpha \cdot a + \beta \cdot b)) + f(b) \cdot ((\alpha \cdot a + \beta \cdot b) - a)}{b-a} \\ & = & \frac{f(a) \cdot (b \cdot (1-\beta) - \alpha \cdot a) + f(b) \cdot (a \cdot (\alpha - 1) + \beta \cdot b)}{b-a} \\ & \underset{\alpha+\beta=1}{=} & \frac{f(a) \cdot (b \cdot \alpha - \alpha \cdot a) + f(b) \cdot (a \cdot (-\beta) + \beta \cdot b)}{b-a} \\ & = & \frac{\alpha \cdot f(a) \cdot (b-a) + \beta \cdot f(b) \cdot (b-a)}{b-a} \\ & = & \alpha \cdot f(a) + \beta \cdot f(b) \end{aligned} \tag{16.152}$$

3. Take $\alpha = \frac{x-a}{b-a}$, $\beta = \frac{b-x}{b-a}$ then by (2) we have that $\alpha, \beta \in [0,1]$, $\alpha + \beta = 1$ and $\alpha \cdot a + \beta \cdot b = x$ so that

$$\begin{array}{rcl} f(x) & = & f(\alpha \cdot a + \beta \cdot b) \\ & \geqslant & \alpha \cdot f(a) + \beta \cdot f(b) \\ & \underset{(2)}{=} & \dfrac{f(a) \cdot (b-x) + f(b) \cdot (x-a)}{b-a} \end{array}$$

proving that f is convex.

4. Take $\alpha = \frac{x-a}{b-a}$, $\beta = \frac{b-x}{b-a}$ then by (2) we have that $\alpha, \beta \in [0,1]$, $\alpha + \beta = 1$ and $\alpha \cdot a + \beta \cdot b = x$ so that

$$\begin{array}{rcl} f(x) & = & f(\alpha \cdot a + \beta \cdot b) \\ & \leqslant & \alpha \cdot f(a) + \beta \cdot f(b) \\ & \underset{(2)}{=} & \dfrac{f(a) \cdot (b-x) + f(b) \cdot (x-a)}{b-a} \end{array}$$

proving that f is concave. □

The next theorem shows the relation between the second derivative and convexity and concavity.

Theorem 16.140. *Let $\langle \mathbb{R}, \| \rangle$ be the normed space of real numbers, U a open set in $\mathbb{R}$, $a, b \in \mathbb{R}$ with $a < b$ and $[a,b] \subseteq U$, $f: U \rightarrow \mathbb{R}$ a function such that f has a 2-the derivative on U then*

1. *If $\forall x \in]a,b[\ 0 \leqslant f_x^{(2)}$ then f is concave on $[a,b]$.*
2. *If $\forall x \in]a,b[\ f_x^{(2)} \leqslant 0$ then f is convex on $[a,b]$.*

Proof. Let $x \in [a,b]$ then we have either

$\boldsymbol{x = a}$**.** Then $\frac{f(a) \cdot (b-x) + f(b) \cdot (x-a)}{b-a} = f(a) = f(x)$. As this result is independent of the sign of $f''(x)$ we have

$$\begin{cases} \frac{f(a) \cdot (b-x) + f(b) \cdot (x-a)}{b-a} \geqslant f(x) \text{ if } \forall x \in]a,b[\ f^{(2)} \geqslant 0 \\ \frac{f(a) \cdot (b-x) + f(b) \cdot (x-a)}{b-a} \leqslant f(x) \text{ if } \forall x \in]a,b[\ f^{(2)}(x) \leqslant 0 \end{cases}$$

$\boldsymbol{x = b}$**.** Then $\frac{f(a) \cdot (b-x) + f(b) \cdot (x-a)}{b-a} = f(b) = f(x)$. As this result is independent of the sign of $f''(x)$ we have

$$\begin{cases} \frac{f(a) \cdot (b-x) + f(b) \cdot (x-a)}{b-a} \geqslant f(x) \text{ if } f^{(2)}(x) \geqslant 0 \\ \frac{f(a) \cdot (b-x) + f(b) \cdot (x-a)}{b-a} \leqslant f(x) \text{ if } f^{(2)}(x) \leqslant 0 \end{cases}$$

$\boldsymbol{x \in]a,b[}$**.** Then

$$\frac{x-a}{b-a} + \frac{b-x}{b-a} = \frac{x-a+b-x}{b-a} = 1$$

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

$$\begin{array}{rl} \dfrac{f(a) \cdot (b-x) + f(b) \cdot (x-a)}{b-a} - f(x) & = \\ \dfrac{b-x}{b-a} \cdot f(a) + \dfrac{x-a}{b-a} \cdot f(b) - \dfrac{x-a}{b-a} \cdot f(x) - \dfrac{b-x}{b-a} \cdot f(x) & = \\ \dfrac{(x-a)}{b-a} \cdot (f(b) - f(x)) - \dfrac{b-x}{b-a} \cdot (f(x) - f(a)) & = \\ \dfrac{(x-a) \cdot (b-x)}{b-a} \cdot \dfrac{f(b) - f(x)}{b-x} - \dfrac{(b-x) \cdot (x-a)}{b-a} \cdot \dfrac{f(x) - f(a)}{x-a} & = \\ \dfrac{(x-a) \cdot (b-x)}{b-a} \left[\dfrac{f(b) - f(x)}{b-x} - \dfrac{f(x) - f(a)}{x-a} \right] & \end{array} \tag{16.153}$$

As $\forall x \in U$ f has a 2-the derivative it follows from [theorem: 16.105] that f has a derivative at every $x \in U$ and the function

$$f': U \rightarrow \mathbb{R} \text{ defined by } f'(x) = f'_x \text{ has a derivative at every } x \in U \text{ and } \forall x \in U\ f_x^{(2)} = (f')'_x \tag{16.154}$$

As $[a,x],[x,b]\subseteq[a,b]\subseteq U$ we can apply Lagrange's theorem [corollary: 16.136] on $f_{|[a,b]}$ to find $y_1\in]a,x]$, $y_2\in]x,b[$ such that

$$f(x)-f(a)=f'_{y_1}\cdot(x-a) \text{ and } f(b)-f(x)=f'_{y_2}\cdot(b-x) \tag{16.155}$$

Combining [eq: 16.153] and [eq: 16.155] gives

$$\frac{f(a)\cdot(b-x)+f(b)\cdot(x-a)}{b-a}-f(x)=\frac{(x-a)\cdot(b-x)}{b-a}\cdot(f'(y_2)-f'(y_1)) \tag{16.156}$$

As $y_1\in]a,x[,\ y_2\in]x,b[$ we have $a<y_1<x<y_2<b$ that $y_1<y_2$ and $[y_1,y_2]\subseteq[a,b]\subseteq U$. Hence using [eq: 16.154] we can apply Lagrange's theorem [theorem: 16.136] on f' to find a $z\in]y_1,y_2[$ such that

$$f'(y_2)-f'(y_1)=(f^{(1)})'_z\cdot(y_2-y_1)=f^{(2)}_z\cdot(y_2-y_1)$$

substituting this in [eq: 16.156] gives

$$\frac{f(a)\cdot(b-x)+f(b)\cdot(x-a)}{b-a}-f(x)=\frac{(x-a)\cdot(b-x)}{b-a}\cdot f^{(2)}(z)$$

Now as $x\in[a,b]$ we have that $\frac{(x-a)\cdot(b-x)}{b-a}\geqslant 0$ so that

$$\begin{cases}\frac{f(a)\cdot(b-x)+f(b)\cdot(x-a)}{b-a}\geqslant f(x) \text{ if } \forall x\in]a,b[\ f^{(2)}(x)\geqslant 0\\ \frac{f(a)\cdot(b-x)+f(b)\cdot(x-a)}{b-a}\leqslant f(x) \text{ if } \forall x\in]a,b[\ f^{(2)}(x)\leqslant 0\end{cases}$$

So in all cases we have $\begin{cases}\frac{f(a)\cdot(b-x)+f(b)\cdot(x-a)}{b-a}\geqslant f(x) \text{ if } \forall x\in]a,b[\ f^{(2)}(x)\geqslant 0\\ \frac{f(a)\cdot(b-x)+f(b)\cdot(x-a)}{b-a}\leqslant f(x) \text{ if } \forall x\in]a,b[\ f^{(2)}(x)\leqslant 0\end{cases}$ proving that

1. If $\forall x\in]a,b[\ 0\leqslant f^{(2)}(x)$ then f is concave on $[a,b]$
2. If $\forall x\in]a,b[\ f^{(2)}(x)\leqslant 0$ then f is convex on $[a,b]$ $\square$

We show now how the concept of convexity or concavity on a interval allows us to determine if we have a minimum or maximum.

Theorem 16.141. *Let $\langle\mathbb{R},\|\rangle$ be the normed space of real numbers, U a open set in X, $f:U\rightarrow\mathbb{R}$ a function, x_0 a local extremum, such that such that f has a 2-the derivative on U then we have:*

1. *If there exist a open set V with $x_0\in V\subseteq U$ and $\forall x\in V\ f^{(2)}(x)\geqslant 0$ then x_0 is a local weak minimum.*
2. *If there exist a open set V with $x_0\in V\subseteq U$ and $\forall x\in V\ f^{(2)}(x)\leqslant 0$ then x_0 is a local weak maximum..*

Proof. As x_0 is a local extremum it follows from [theorem: 16.133] that

As f has a 2-the derivative on U it follows from [theorem: 16.105] that $\forall x\in U$ f has a 1-the derivative at x and the function

$$f':U\rightarrow X \text{ defined by } f'(x)=f'_x$$

has a derivative at every $x\in U$ exist and $f^{(2)}_x=(f')'_x$.

1. As V is open there exists a $\delta>0$ such that $]x_0-\delta,x_0+\delta[\subseteq V$ hence

   $$\forall x\in]x_0-\delta,x_0+\delta[\text{ we have } f^{(2)}(x)\geqslant 0$$

   Using [corollary: 16.137] it follows that f' is increasing on $]x_0-\delta,x_0+\delta[$, hence we have

   $$\forall x\in]x_0-\delta,x_0]\ f'(x)\leqslant f'(x_0)=0 \text{ and } \forall x\in[x_0,x_0+\delta[\ 0=f'(x_0)\leqslant f'(x)$$

   So using [corollary: 16.137] f is decreasing on $]x_0-\delta,x_0]\Rightarrow\forall x\in]x_0-\delta,x_0]$ we have $f(x)\geqslant f(x_0)$ and f is increasing on $[x_0,x_0+\delta[\Rightarrow\forall x\in[x_0,x_o+\delta[$ we have $f(x_0)\leqslant f(x)$. Hence

   $$\forall x\in]x_0-\delta,x_0+\delta[\text{ we have } f(x_0)\leqslant f(x) \text{ or } x_0 \text{ is a local weak minimum.}$$

2. As V is open there exists a $\delta>0$ such that $]x_0-\delta,x_0+\delta[\subseteq V$ hence

   $$\forall x\in]x_0-\delta,x_0+\delta[\text{ we have } f^{(2)}(x)\leqslant 0$$

So using [corollary: 16.137] it follows that f' is decreasing on $]x_0-\delta,x_0+\delta[$, hence we have

$$\forall x\in]x_0-\delta,x_0]\;0=f'(x_0)\leqslant f'(x)\text{ and }\forall x\in[x_0,x_0+\delta[\;f'(x)\leqslant f'(x_0)=0$$

So using [corollary: 16.137] f is increasing on $]x_0-\delta,x_0]\Rightarrow\forall x\in]x_0-\delta,x_0]$ we have $f(x_0)\geqslant f(x)$ and f is decreasing on $[x_0,x_0+\delta[\Rightarrow\forall x\in[x_0,x_o+\delta[$ we have $f(x)\leqslant f(x_0)$. Hence

$$\forall x\in]x_0-\delta,x_0+\delta[\text{ we have }f(x)\leqslant f(x_0)\text{ or }x_0\text{ is a local weak maximum.}$$ □

Actually we can extend the above to a maximum and minimum on a interval instead of a local weak minimum or local weak maximum.

Theorem 16.142. *Let $\langle\mathbb{R},\|\rangle$ be the normed space of real numbers, U a open set in X, $a,b\in U$ with $a\leqslant b$, $[a,b]\subseteq U$, $f:U\rightarrow\mathbb{R}$ a function such that f has a 2-the derivative on X and that there exist a $x_0\in[a,b]$ with $f'(x_0)=0$ then we have:*

1. *If $\forall x\in U\;0\leqslant f^{(2)}(x)$ then $\forall x\in[a,b]\;f(x_0)\leqslant f(x)$ [in other words $f(x_0)$ is the minimum of $f([a,b])$.*
2. *If $\forall x\in U\;0\leqslant f^{(2)}(x)$ then $\forall x\in[a,b]\;f(x_0)\leqslant f(x)$ [in other words $f(x_0)$ is the maximum of $f([a,b])$.*

Proof. If $a=b$ then $[a,b]=\{x_0\}$ and the theorem is trivial true. So we must only prove the theorem for the case $a<b$.

1. From the previous theorem [theorem:16.141] it follows that x_0 is a local weak minimum of f hence there exist a $\delta\in\mathbb{R}^+$ such that

$$\forall x\in]x_0-\delta,x_0+\delta[\text{ we have }f(x_0)\leqslant f(x)\tag{16.157}$$

Assume that there exist a $y\in[a,b]$ such that $f(y)<f(x_0)$ then $y\neq x_0$, so we have either:

***$y<x_0$*.** Then $\max(x_0-\delta,y)<x_0$ so there exist a $z\in\mathbb{R}$ such that $x_0-\delta,y<z<x_0$ hence by [eq: 16.157] we have

$$f(x_0)\leqslant f(z)\tag{16.158}$$

Further using [theorem: 16.140] f is concave on $[y,x_0]$ so that

$$f(z)\leqslant\frac{f(y)\cdot(x_0-z)+f(x_0)\cdot(z-y)}{x_0-y}\tag{16.159}$$

As $0<x_0-z$ and by assumption $f(y)<f(x_0)$ it follows that

$$f(y)\cdot(x_0-z)<f(x_0)\cdot(x_0-z)\tag{16.160}$$

Hence we have

$$\begin{array}{rcl} f(x_0) & = & \dfrac{f(x_0)\cdot(x_0-y)}{x_0-y}\\ & = & \dfrac{f(x_0)\cdot(x_0-z+z-y)}{x_0-y}\\ & = & \dfrac{f(x_0)\cdot(x_0-z)+f(x_0)\cdot(z-y)}{x_0-y}\\ & >_{[\text{eq: }16.160]} & \dfrac{f(y)\cdot(x_0-z)+f(x_0)\cdot(z-y)}{x_0-y}\\ & \geqslant_{[\text{eq: }16.159]} & f(z)\\ & \geqslant_{[\text{eq: }16.158]} & f(x_0)\end{array}$$

leading to the contradiction $f(x_0)>f(x_0)$.

***$x_0<y$*.** Then $x_0<\min(y,x_0+\delta)$ so there exist a $z\in\mathbb{R}$ such that $x_0<z<y,x_0+\delta$ hence by [eq: 16.157] we have

$$f(x_0)\leqslant f(z)\tag{16.161}$$

Further using [theorem: 16.140] f is concave on $[x_0,y]$ so that

$$f(z)\leqslant\frac{f(x_0)\cdot(y-z)+f(y)\cdot(z-x_0)}{y-x_0}\tag{16.162}$$

As $0<z-x_0$ and $f(y)<f(x_0)$ [by assumption] we have

$$f(y)\cdot(z-x_0)<f(x_0)\cdot(z-x_0)\tag{16.163}$$

Hence we have

$$\begin{aligned} f(x_0) &= \frac{f(x_0)\cdot(y-x_0)}{(y-x_0)} \\ &= \frac{f(x_0)\cdot(y-z+z-x_0)}{(y-x_0)} \\ &= \frac{f(x_0)\cdot(y-z)+f(x_0)\cdot(z-x_0)}{(y-x_0)} \\ &>_{[\text{eq: } 16.163]} \frac{f(x_0)\cdot(y-z)+f(y)\cdot(z-x_0)}{(y-x_0)} \\ &\geqslant_{[\text{eq: } 16.162]} f(z) \\ &\geqslant_{[\text{eq: } 16.161]} f(x_0) \end{aligned}$$

leading to the contradiction $f(x_0) > f(x_0)$.

As in all cases we have a contradiction the assumption must be false, so it follows that

$$\forall x \in [a,b] \text{ we have } f(x_0) \leqslant f(x)$$

2. From the previous theorem [theorem:16.141] it follows that x_0 is a local weak maximum of f hence there exist a $\delta \in \mathbb{R}^+$ such that

$$\forall x \in]x_0-\delta, x_0+\delta[\text{ we have } f(x) \leqslant f(x_0) \tag{16.164}$$

Assume that there exist a $y \in [a,b]$ such that $f(x_0) < f(y)$ then we can not have $y = x_0$, hence we must consider the following cases:

$\boldsymbol{y < x_0}$**.** Then $\max(x_0-\delta, y) < x_0$ so there exist a $z \in \mathbb{R}$ with $x_0-\delta, y < z < x_0$. Applying [eq: 16.164] gives

$$f(z) \leqslant f(x_0) \tag{16.165}$$

Further using [theorem: 16.140] f is convex on $[y, x_0]$ so that

$$\frac{f(y)\cdot(x_0-z)+f(x_0)\cdot(z-y)}{x_0-y} \leqslant f(z) \tag{16.166}$$

As $0 < x_0 - z$ and by the assumption $f(x_0) < f(y)$ it follows that

$$f(x_0)\cdot(x_0-z) < f(y)\cdot(x_0-z) \tag{16.167}$$

Hence we have

$$\begin{aligned} f(x_0) &= \frac{f(x_0)\cdot(x_0-y)}{x_0-y} \\ &= \frac{f(x_0)\cdot(x_0-z+z-y)}{x_0-y} \\ &= \frac{f(x_0)\cdot(x_0-z)+f(x_0)\cdot(z-y)}{x_0-y} \\ &<_{[\text{eq: } 16.167]} \frac{f(y)\cdot(x_0-z)+f(x_0)\cdot(z-y)}{x_0-y} \\ &\leqslant_{[\text{eq: } 16.166]} f(z) \\ &\leqslant_{[\text{eq: } 16.165]} f(x_0) \end{aligned}$$

leading to the contradiction $f(x_0) < f(x_0)$.

$\boldsymbol{x_0 < y}$**.** Then $x_0 < \min(y, x_0+\delta)$ so there exist a $z \in \mathbb{R}$ with $x_0 < z < y, x_0+\delta$. Applying [eq: 16.164] gives

$$f(z) \leqslant f(x_0) \tag{16.168}$$

Further using [theorem: 16.140] f is convex on $[x_0, y]$ so that

$$\frac{f(x_0)\cdot(y-z)+f(y)\cdot(z-x_0)}{y-x_0} \leqslant f(z) \tag{16.169}$$

As $0 < z - x_0$ and by assumption $f(x_0) < f(y)$ it follows that

$$f(x_0)\cdot(z-x_0) < f(y)\cdot(z-x_0) \tag{16.170}$$

Hence we have

$$\begin{aligned} f(x_0) &= \frac{f(x_0)\cdot(y-x_0)}{(y-x_0)} \\ &= \frac{f(x_0)\cdot(y-z+z-x_0)}{(y-x_0)} \\ &= \frac{f(x_0)\cdot(y-z)+f(x_0)\cdot(z-x_0)}{(y-x_0)} \\ &<_{[\text{eq: } 16.170]} \frac{f(x_0)\cdot(y-z)+f(y)\cdot(z-x_0)}{(y-x_0)} \\ &\leqslant_{[\text{eq: } 16.169]} f(z) \\ &\leqslant_{[\text{eq: } 16.168]} f(x_0) \end{aligned}$$

leading to the contradiction $f(x_0)<f(x_0)$.

As in all cases we have a contradiction the assumption must be false, so it follows that

$$\forall x\in[a,b] \text{ we have } f(x)\leqslant f(x_0)$$

□

16.3.2 Derivatives on a closed interval

Definition 16.143. *Let $\langle\mathbb{R},\|\|\rangle$ be the real normed space with the canonical norm, $a,b\in\mathbb{R}$ with $a<b$, $\langle X,\|\|\rangle$ a normed space and f: $[a,b]\rightarrow X$ a function then we define the left and right derivative of f at a point as follows:*

1. *If $x\in[a,b[$ then f has a right derivative at x noted as $f'_{+,x}\in X$ if $\forall\varepsilon\in\mathbb{R}^+$ there exist a $\delta\in\mathbb{R}^+$ such that $\forall h\in\mathbb{R}$ with $0<h<\delta$ and $x+h\in[a,b]$ we have*
$$\left\|\frac{f(x+h)-f(x)}{h}-f'_{+,x}\right\|<\varepsilon$$
2. *If $x\in]a,b]$ then f has a left derivative at x noted as $f'_{-,x}\in X$ if $\forall\varepsilon\in\mathbb{R}^+$ there exist a $\delta\in\mathbb{R}^+$ such that $\forall h\in\mathbb{R}$ with $0<h<\delta$ and $x-h\in[a,b]$ we have*
$$\left\|\frac{f(x-h)-f(x)}{-h}-f'_{-,x}\right\|<\varepsilon$$

Proof. For this definition to make sense we must of course prove that the left and right derivative is unique.

1. Assume that d_1,d_2 are right derivatives at $x\in[a,b[$ with $d_1\neq d_2$. Take $\varepsilon=\|d_2-d_1\|\in\mathbb{R}^+$ then there exists $\delta_1,\delta_2\in\mathbb{R}^+$ such that if $0<h<\delta_1$ and $x+h\in[a,b]$ we have
$$\left\|\frac{f(x+h)-f(x)}{h}-d_1\right\|<\frac{\varepsilon}{2}$$
and if $0<h<\delta_2$ and $x+h\in[a,b]$ we have
$$\left\|\frac{f(x+h)-f(x)}{h}-d_2\right\|<\frac{\varepsilon}{2}$$
As $0<\min(\delta_1,\delta_2)$ there exist a $h\in\mathbb{R}$ with $0<h<\min(\delta_1,\delta_2)$ and we have
$$\begin{aligned} \varepsilon &= \|d_2-d_1\| \\ &= \left\|\frac{f(x+h)-f(x)}{h}-d_1-\left(\frac{f(x+h)-f(x)}{h}-d_2\right)\right\| \\ &\leqslant \left\|\frac{f(x+h)-f(x)}{h}-d_1\right\|+\left\|\frac{f(x+h)-f(x)}{h}-d_2\right\| \\ &< \frac{\varepsilon}{2}+\frac{\varepsilon}{2} \\ &= \varepsilon \end{aligned}$$
leading to the contradiction $\varepsilon<\varepsilon$. Hence the assumptions must be false and we must have $d_1=d_2$.
2. Assume that d_1,d_2 are left derivatives at $x\in[a,b[$ with $d_1\neq d_2$. Take $\varepsilon=\|d_2-d_1\|\in\mathbb{R}^+$ then there exists $\delta_1,\delta_2\in\mathbb{R}^+$ such that if $0<h<\delta_1$ and $x-h\in[a,b]$ we have
$$\left\|\frac{f(x-h)-f(x)}{-h}-d_1\right\|<\frac{\varepsilon}{2}$$

and if $0<h<\delta_2$ and $x-h\in[a,b]$ we have

$$\left\|\frac{f(x-h)-f(x)}{-h}-d_2\right\|<\frac{\varepsilon}{2}$$

As $0<\min(\delta_1,\delta_2)$ there exist a $h\in\mathbb{R}$ with $0<h<\min(\delta_1,\delta_2)$ and we have

$$\begin{array}{rcl}\varepsilon & = & \|d_2-d_1\| \\ & = & \left\|\frac{f(x-h)-f(x)}{-h}-d_1-\left(\frac{f(x-h)-f(x)}{-h}-d_2\right)\right\| \\ & \leqslant & \left\|\frac{f(x-h)-f(x)}{-h}-d_1\right\|+\left\|\frac{f(x-h)-f(x)}{-h}-d_2\right\| \\ & < & \frac{\varepsilon}{2}+\frac{\varepsilon}{2} \\ & = & \varepsilon\end{array}$$

leading to the contradiction $\varepsilon<\varepsilon$. Hence the assumptions must be false and we must have $d_1=d_2$. $\square$

Theorem 16.144. *Let $\langle\mathbb{R},\|\rangle$ be the real normed space with the canonical norm, $a,b\in\mathbb{R}$ with $a<b$ $\langle X,\|\|\rangle$ a normed space then we have:*

1. *If $f:[a,b]\rightarrow X$ and $g:[a,b]\rightarrow X$ has a right derivative at $x\in[a,b[$ then $f+g:[a,b]\rightarrow X$ has a right derivative at x and $(f+g)'_{+,x}=f'_{+,x}+g'_{+,x}$*
2. *If $f:[a,b]\rightarrow X$ and $g:[a,b]\rightarrow X$ has a left derivative at $x\in]a,b]$ then $f+g:[a,b]\rightarrow X$ has a left derivative at x and $(f+g)'_{-,x}=f'_{-,x}+g'_{-,x}$*
3. *If $\alpha\in\mathbb{R}$ and $f:[a,b]\rightarrow X$ has a right derivative at $x\in[a,b[$ then $\alpha\cdot f:[a,b]\rightarrow X$ has a right derivative at x and $(\alpha\cdot f)'_{+,x}=\alpha\cdot f'_{+,x}$.*
4. *If $\alpha\in\mathbb{R}$ and $f:[a,b]\rightarrow X$ has a left derivative at $x\in]a,b]$ then $\alpha\cdot f:[a,b]\rightarrow X$ has a left derivative at x and $(\alpha\cdot f)'_{-,x}=\alpha\cdot f'_{+-x}$.*

Proof.

1. Let $\varepsilon\in\mathbb{R}^+$. Then there exist $\delta_1,\delta_2\in\mathbb{R}^+$ such that if $0<h<\delta_1$ and $a+h\in[a,b]$ then

$$\left\|\frac{f(a+h)-f(a)}{h}-f'_{+,a}\right\|<\frac{\varepsilon}{2}$$

and if $0<h<\delta_2$ and $a+h\in[a,b]$ then

$$\left\|\frac{g(a+h)-g(a)}{h}-g'_{+,a}\right\|<\frac{\varepsilon}{2}$$

Take $\delta=\min(\delta_1,\delta)$ then we have

$$\begin{array}{rl}\left\|\frac{(f+g)(a+h)-(f+g)(a)}{h}-(f'_{+,a}+g'_{+,a})\right\| & = \\ \left\|\frac{f(a+h)+g(a+h)-f(a)-g(a)}{h}-f'_{+,a}-g'_{+,a}\right\| & = \\ \left\|\frac{f(a+h)-f(a)}{h}-f'_{+,a}+\frac{g(a+h)-g(a)}{h}-g'_{+,a}\right\| & \leqslant \\ \left\|\frac{f(a+h)-f(a)}{h}-f'_{+,a}\right\|+\left\|\frac{g(a+h)-g(a)}{h}-g'_{+,a}\right\| & < \\ \frac{\varepsilon}{2}+\frac{\varepsilon}{2} & = \\ \varepsilon & \end{array}$$

proving that $f+g$ has a right derivative at a and $(f+g)'_{+,a}=f'_{+,a}+g'_{+,a}$.

2. Let $\varepsilon\in\mathbb{R}^+$. Then there exist $\delta_1,\delta_2\in\mathbb{R}^+$ such that if $0<h<\delta_1$ and $b-h\in[a,b]$ then

$$\left\|\frac{f(b-h)-f(b)}{-h}-f'_{-,a}\right\|<\frac{\varepsilon}{2}$$

and if $0<h<\delta_2$ and $b-h\in[a,b]$ then

$$\left\|\frac{g(b-h)-g(b)}{-h}-g'_{-,b}\right\|<\frac{\varepsilon}{2}$$

Take $\delta = \min(\delta_1, \delta)$ then we have

$$\begin{aligned}
\left\|\frac{(f+g)(b-h)-(f+g)(b)}{-h} - (f'_{-,b} + g'_{-,a})\right\| &= \\
\left\|\frac{f(b-h)+g(b-h)-f(b)-g(b)}{-h} - f'_{-,b} - g'_{-,b}\right\| &= \\
\left\|\frac{f(b-h)-f(b)}{-h} - f'_{-,a} + \frac{g(b-h)-g(b)}{-h} - g'_{-,b}\right\| &\leqslant \\
\left\|\frac{f(b-h)-f(b)}{-h} - f'_{-,b}\right\| + \left\|\frac{g(b-h)-g(b)}{-h} - g'_{-,b}\right\| &< \\
\frac{\varepsilon}{2} + \frac{\varepsilon}{2} &= \\
\varepsilon &
\end{aligned}$$

proving that $f+g$ has a left derivative at a and $(f+g)'_{-,b} = f'_{-,b} + g'_{-,b}$.

3. Let $\varepsilon \in \mathbb{R}^+$. Then there exist a $\delta \in \mathbb{R}^+$ such that if $0 < h < \delta$ and $a+h \in [a,b]$ then

$$\left\|\frac{f(a+h)-f(a)}{h} - f'_{+,a}\right\| < \frac{\varepsilon}{|\alpha|+1}$$

then we have

$$\begin{aligned}
\left\|\frac{(\alpha\cdot f)(a+h)-(\alpha\cdot f)}{h} - \alpha\cdot f'_{+,a}\right\| &= \left\|\alpha\cdot\left(\frac{f(a+h)-f(a)}{h} - f'_{+,a}\right)\right\| \\
&= |\alpha|\cdot\left\|\frac{f(a+h)-f(a)}{h} - f'_{+,a}\right\| \\
&< |\alpha|\cdot\frac{\varepsilon}{|\alpha|+1} \\
&< \varepsilon
\end{aligned}$$

proving that $\alpha\cdot f$ has a right derivative at a and $(\alpha\cdot f)'_{+.a} = \alpha\cdot f'_{+,a}$.

4. Let $\varepsilon \in \mathbb{R}^+$. Then there exist a $\delta \in \mathbb{R}^+$ such that if $0 < h < \delta$ and $b-h \in [a,b]$ then

$$\left\|\frac{f(b-h)-f(b)}{-h} - f'_{-,a}\right\| < \frac{\varepsilon}{|\alpha|+1}$$

then we have

$$\begin{aligned}
\left\|\frac{(\alpha\cdot f)(b-h)-(\alpha\cdot f)}{-h} - \alpha\cdot f'_{-,b}\right\| &= \left\|\alpha\cdot\left(\frac{f(b-h)-f(b)}{-h} - f'_{-,b}\right)\right\| \\
&= |\alpha|\cdot\left\|\frac{f(b-h)-f(b)}{-h} - f'_{-,b}\right\| \\
&< |\alpha|\cdot\frac{\varepsilon}{|\alpha|+1} \\
&< \varepsilon
\end{aligned}$$

proving that $\alpha\cdot f$ has a left derivative at b and $(\alpha\cdot f)'_{-.b} = \alpha\cdot f'_{-,b}$. □

We have the following relation between left and right derivatives and derivatives.

Theorem 16.145. *Let $\langle\mathbb{R},\|\rangle$ be the real normed space with the canonical norm, $a,b\in\mathbb{R}$ with $a<b$, $\langle X,\|\|\rangle$ a normed space and $f:[a,b]\to X$ a function then*

1. *If $x \in]a,b[$ and f has a right derivative $f'_{+,x}$ and left derivative $f'_{-,x}$ with $f'_{+,x} = f'_{-,x}$ then $f_{|]a,b[}$ has a derivative at x with $(f_{|]a,b[})' = f'_{+,x} = f'_{-,x}$*
2. *If $x \in]a,b[$ and $f_{|]a,b[}$ has a derivative at x then f has a left and right derivative at x with $(f_{|]a,b[})'_x = f'_{+,x} = f'_{-,x}$.*
3. *If $[a,b]\subseteq U$ and $f: U\to X$ is a function such that $\forall x\in U$ f has a derivative f'_x then*
   a. *$f_{|[a,b]}$ has a right derivative at a with $(f_{|[a,b]})'_{+,a} = f'_a$.*

b. *$f_{|[a,b]}$ has a left derivative at b with $(f_{|[a,b]})'_{+,b} = f'_b$.*

c. *We have $\forall x \in]a,b[$ that $f_{[a,b]}$ has a left and right derivative at x with*

$$(f_{|[a,b]})'_{+,x} = f'_x = (f_{|[a,b]})'_{-,x}$$

Proof.

1. Let $d = f'_{+,x} = f'_{-,x}$ and let $\varepsilon \in \mathbb{R}^+$ then there exist $\delta_1, \delta_2 \in \mathbb{R}^+$ such that if $0 < h < \delta_1$ and $x + h \in [a,b]$ we have
$$\left\| \frac{f(x+h) - f(x)}{h} - d \right\| < \varepsilon \tag{16.171}$$
and if $0 < h < \delta_2$ and $x - h \in [a,b]$ we have
$$\left\| \frac{f(x-h) - f(x)}{-h} - d \right\| < \varepsilon \tag{16.172}$$
So if $0 < |h| < \min(\delta_1, \delta_2)$ with $h \in]a,b[_x$ [or $x + h \in]a,b[$] we have either:
$0 < h$. Then as $0 < |h| < \min(\delta_1, \delta_2) < \delta_1$ we have $0 < h < \delta_1$ and from $h \in]a,b[_x$ that $x + h \in]a,b[\subseteq [a,b]$, hence
$$\begin{aligned}\left\| \frac{f_{|]a,b[}(x+h) - f_{]a,b[}(x)}{h} - d \right\| &\underset{x,x+h \in]a,b[}{=} \left\| \frac{f(x+h) - f(x)}{h} - d \right\| \\ &<_{[\text{eq: } 16.171]} \varepsilon\end{aligned}$$
$h < 0$. Then $-h = |h|$ so that from $h \in]a,b[_x$ we have $x - |h| \in]a,b[$, further we have $0 < |h| < \delta_2$, hence
$$\begin{aligned}\left\| \frac{f_{|]a,b[}(x+h) - f_{]a,b[}(x)}{h} - d \right\| &\underset{x,x+h \in]a,b[}{=} \left\| \frac{f(x+h) - f(x)}{h} - d \right\| \\ &= \left\| \frac{f(x-|h|) - f(x)}{-|h|} - d \right\| \\ &<_{[\text{eq:} 16.172]} \varepsilon\end{aligned}$$
so in all cases we have
$$\left\| \frac{f_{|]a,b[}(x+h) - f_{]a,b[}(x)}{h} - d \right\| < \varepsilon$$
proving that $f_{|]a,b[}$ has a derivative at x and $(f_{|]a,b[})'_{+,x} = d = f'_{+,x} = f'_{-,x}$.

2. Let $\varepsilon \in \mathbb{R}^+$ then, as $f_{|]a,b[}$ has a derivative at x, there exist a $\delta_1 \in \mathbb{R}^+$ such that $\forall h \in]a,b[_x$ [or equivalently $x + h \in]a,b[$ with $0 < |h| < \delta_1$ we have
$$\left\| \frac{f(x+h) - f(x)}{h} \right\| \underset{x \in, x+h \in]a,b[}{=} \left\| \frac{f_{|]a,b[}(x+h) - f_{|]a,b[}(x)}{h} - (f_{|]a,b[})'_x \right\| < \varepsilon \tag{16.173}$$
As $x \in]a,b[$ a open set there exist a $\delta_2 \in \mathbb{R}^+$ such that $x \in]x - \delta_2, x + \delta_2[\subseteq]a,b[$. Take now $\delta = \min(\delta_1, \delta_2)$ then if $0 < h < \delta \leqslant \delta_1, \delta_2$ we have $h < \delta_2$ so that $x - \delta_2 < x - h < x < x + h < x + \delta_2$ or $x - h, x + h \in]x - \delta_2, x + \delta_2[\subseteq]a,b[$ so that $h, -h \in]a,b[$ and $0 < |h|$, $|-h| < \delta_1$. Using then [eq: 16.173] it follows that
$$\left\| \frac{f(x+h) - f(x)}{h} - (f_{|]a,b[})'_x \right\| < \varepsilon \text{ and } \left\| \frac{f(x-h) - f(x)}{-h} - (f_{|]a,b[})'_x \right\| < \varepsilon$$
proving that f has a left and right derivative at x and
$$(f_{|]a,b[})'_x = f'_{+,x} = f'_{-,x}$$

3. Let $x \in [a,b]$ then as $[a,b] \subseteq U$ it follows that f has a derivative at x, hence if $\varepsilon \in \mathbb{R}^+$ there exist a $\delta_{\varepsilon,x} \in \mathbb{R}^+$ such that
$$\forall h \in U_x \text{ with } 0 < |h| < \delta_{\varepsilon,x} \text{ we have } \left\| \frac{f(x+h) - f(x)}{h} - f'_x \right\| < \varepsilon \tag{16.174}$$

   a. Let $\varepsilon \in \mathbb{R}^+$ then $\forall h \in \mathbb{R}$ with $0 < h < \delta_{\varepsilon,a}$ and $a + h \in [a,b]$ we have $0 < |h| < \delta_{\varepsilon,a}$ and $a + h \in U \Rightarrow h \in U_a$ so that
$$\left\| \frac{f_{|[a,b]}(a+h) - f_{|[a,b]}(a)}{h} - f'_a \right\| = \left\| \frac{f(a+h) - f(a)}{h} - f'_a \right\| <_{[\text{eq: } 16.174]} \varepsilon$$

proving that $(f_{|[a,b]})'_{+,a} = f'_a$.

b. Let $\varepsilon \in \mathbb{R}^+$ then $\forall h \in \mathbb{R}$ with $0 < h < \delta_{\varepsilon,b}$ and $b - h \in [a,b]$ we have $0 < |-h| < \delta_{\varepsilon,b}$ and $b - h \in U \Rightarrow -h \in U_b$ so that

$$\left\| \frac{f_{|[a,b]}(b-h) - f_{|[a,b]}(b)}{-h} - f'_b \right\| = \left\| \frac{f(b-h) - f(b)}{-h} - f'_b \right\| <_{[\text{eq: } 16.174]} \varepsilon$$

proving that $(f_{|[a,b]})'_{-,b} = f'_b$.

c. Let $\varepsilon \in \mathbb{R}^+$ then $\forall h \in \mathbb{R}$ with $0 < h < \delta_\varepsilon$ and $x + h \in [a,b]$ we have $0 < |h| < \delta$ and $x + h \in U \Rightarrow h \in U_x$ so that

$$\left\| \frac{f_{|[a,b]}(x+h) - f_{|[a,b]}(x)}{h} - f'_x \right\| = \left\| \frac{f(x+h) - f(x)}{h} - f'_x \right\| <_{[\text{eq: } 16.174]} \varepsilon$$

proving that $(f_{|[a,b]})'_{+,x} = f'_x$. Let $\varepsilon \in \mathbb{R}^+$ then $\forall h \in \mathbb{R}$ with $0 < h < \delta_\varepsilon$ and $x - h \in [a,b]$ we have $0 < |-h| < \delta$ and $x - h \in U \Rightarrow -h \in U_x$ so that

$$\left\| \frac{f_{|[a,b]}(x-h) - f_{|]a,b[}(x)}{-h} - f'_x \right\| = \left\| \frac{f(x-h) - f(x)}{-h} - f'_x \right\| <_{[\text{eq: } 16.174]} \varepsilon$$

proving that $(f_{|[a,b]})'_{-x} = f'_x$.

□

We use the above theorem for the following definition to define the extend the concept of derivatives to a **closed interval** $[a,b]$.

Definition 16.146. *Let $\langle \mathbb{R}, \| \rangle$ be the normed space of the real numbers, $\langle X, \|\|\| \rangle$ a normed space, $a, b \in \mathbb{R}$ with $a < b$ and $f: [a,b] \to X$ a function, $x \in [a,b]$ then f has a derivative at x noted as f'_x if for x we have*

***x*=*a*.** *f has a right derivative at x*

***x* = *b*.** *f has a left derivative at x*

$x \in]a,b[$. *$f_{|]a,b[}$ has a derivative at x using [definition: 16.12].*

the derivative is then defined by

$$f'_x = \begin{cases} f'_{+,a} \text{ if } x = a \\ f'_{-,b} \text{ if } x = b \\ (f_{|]a,b[})'_x \text{ if } x \in]a,b[\end{cases}$$

*We say that **f has a derivate on [a,b]** if $\forall x \in [a,b]$ f has a derivative at x [in the above sense]. If **f has a derivative on [a,b]** then the function f' is defined to be*

$$f': [a,b] \to X \text{ where } f'(x) = f'_x$$

*In other words **f has a derivative on [a,b]** if*

1. *$\forall x \in]a,b[$ we have that $f_{|]a,b[}$ has a derivative at x.*
2. *f has a right derivative at a and a left derivative at b.*

If f has a derivative on $[a,b]$ then the function f' is defined to be

$$f': [a,b] \to \mathbb{R} \text{ by } f'(x) = \begin{cases} f'_{+,a} \text{ if } x = a \\ f'_{-,b} \text{ if } x = b \\ (f_{|]a,b[})'_x \text{ if } x \in]a,b[\end{cases}$$

***Note* 16.147.** *Be careful, in this definition the word derivative has two meanings, the derivative at $x \in [a,b]$ (defined here) and the derivative of $f_{|]a,b[}$ (defined in [definition: 16.12]). The same happens for f'_x which is defined in this definition and $f'_{|]a,b[}$ which is defined in [definition: 16.12]. The two definitions happens to coincide on $]a,b[$ but for a, b [definition: 16.12] does not work and we use the right/left derivative. In general if we look carefully at the context there is not a problem and we have avoided introducing excessive notation.*

Example 16.148. Let $\langle \mathbb{R}, \| \rangle$ be the normed space of the real numbers, $\langle X, \|\|\| \rangle$ a normed space, $a, b \in \mathbb{R}$ with $a < b$, $y \in [a,b]$ then

$$C_y: [a,b] \to X \text{ defined by } C_y(r) = x$$

has a derivative on $[a,b]$ and $\forall x \in [a,b]$ $(C_y)'(x) = 0$ or in other words $(C_y)' = C_0$

Proof. Let $x\in[a,b]$ then we have:

$x=a$. Let $\varepsilon\in\mathbb{R}^+$. If $0<h<1$ with $a+h\in[a,b]$ we have

$$\left\|\frac{C_y(a+h)-C_y(a)}{h}-0\right\|=\left\|\frac{x-x}{h}-0\right\|=0<\varepsilon$$

which proves that C_y has a right derivative at a with $(C_y)'_{+,a}=0$.

$x=b$. Let $\varepsilon\in\mathbb{R}^+$. If $0<h<1$ with $b-h\in[a,b]$ we have

$$\left\|\frac{C_y(b-h)-C_y(b)}{-h}-0\right\|=\left\|\frac{x-x}{-h}-0\right\|=0<\varepsilon$$

which proves that C_y has a left derivative at a with $(C_y)'_{-,b}=0$.

$x\in]a,b[$. Let $\varepsilon\in\mathbb{R}^+$. If $h\in[a,b]_x$ with $0<|h|<1$ then we have

$$\left\|\frac{C_y(x+h)-C_y(x)}{h}-0\right\|=\left\|\frac{y-y}{h}-0\right\|=0<\varepsilon$$

which proves that $(C_y)_{|]a,b[}$ has a derivative at x with $((C_y)_{|]a,b[})'_x=0$.

Proving that C_y has a derivative on $[a,b]$ with $\forall x\in[a,b]$ $(C_y)'(x)=0$. □

Theorem 16.149. *Let $\langle\mathbb{R},\|\rangle$ be the normed space of the real numbers, $\langle X,\|\|\rangle$ a normed space, $a,b\in\mathbb{R}$ with $a<b$, $x\in[a,b]$ then:*

1. *If $f:[a,b]\rightarrow X$ and $g:[a,b]\rightarrow X$ has a derivative at x then $f+g$ has a derivative at x and $(f+g)'_x=f'_x+g'_x$.*
2. *If $\alpha\in\mathbb{R}$ and $f:[a,b]\rightarrow X$ has a derivative at x then $\alpha\cdot f$ has a derivative at x and $(\alpha\cdot f)'_x=\alpha\cdot f'_x$.*

Proof.

1. For x we have three cases to consider:

   $x=a$. Then f,g have a right derivative at a and $f'_x=f'_{+,a}$, $g'_x=g'_{+,a}$. Using [theorem: 16.144] it follows that $f+g$ has a right derivative at a and $(f+g)'_{+,a}=f'_{+,a}+g'_{+,a}$. Hence by definition $f+g$ has a derivative at x and $(f+g)'_x=f'_x+g'_x$.

   $x=b$. Then f,g have a left derivative at b and $f'_x=f'_{-,b}$, $g'_x=g'_{-,b}$. Using [theorem: 16.144] it follows that $f+g$ has a left derivative at b and $(f+g)'_{-,b}=f'_{-,b}+g'_{-,b}$. Hence by definition $f+g$ has a derivative at x and $(f+g)'_x=f'_x+g'_x$.

   $x\in]a,b[$. Then $f_{|]a,b[}$, $g_{|]a,b[}$ has a derivative at x, so by [theorem: 16.149] $(f+g)_{|]a,b[}=f_{|]a,b[}+g_{|]a,b[}$ has a derivative at x with $((f+g)_{|]a,b[})'_x=(f_{|]a,b[}+g_{|]a,b[})'_x=(f_{|]a,b[})'_x+(g_{|]a,b[})'_x$. Hence by definition $f+g$ has a derivative at x with $(f+g)'_x=f'_x+g'_x$.

   Hence in all cases we have that $f+g$ has a derivative at x and $(f+g)'_x=f'_x+g'_x$.

2. For x we have three cases to consider:

   $x=a$. Then f has a right derivative at a and $f'_x=f'_{+,a}$. Using [theorem: 16.144] it follows that $\alpha\cdot f$ has a right derivative at a and $(\alpha\cdot f)'_{+,a}=\alpha\cdot f'_{+,a}$. Hence by definition $\alpha\cdot f$ has a derivative at x and $(\alpha\cdot f)'_x=\alpha\cdot f'_x$.

   $x=b$. Then f has a left derivative at b and $f'_x=f'_{-,b}$. Using [theorem: 16.144] it follows that $\alpha\cdot f$ has a left derivative at b and $(\alpha\cdot f)'_{-,b}=\alpha\cdot f'_{-,b}$. Hence by definition $\alpha\cdot f$ has a derivative at x and $(\alpha\cdot f)'_x=\alpha\cdot f'_x$.

   $x\in]a,b[$. Then $f_{|]a,b[}$ has a derivative at x, so so by [theorem: 16.149] $\alpha\cdot f_{|]a,b[}$ and $(\alpha\cdot f_{|]a,b[})'_x=\alpha\cdot(f_{|]a,b[})'_x$. Hence by definition $\alpha\cdot f$ has a derivative at x and $(\alpha\cdot f)'_x=\alpha\cdot f'_x$.

   So in all cases we have that $\alpha\cdot f$ has a derivative at x and $(\alpha\cdot f)'_x=\alpha\cdot f'_x$. □

Theorem 16.150. *Let $\langle\mathbb{R},\|\rangle$ be the normed space of the real numbers, $\langle X,\|\|\rangle$ a normed space, $a,b\in\mathbb{R}$ and a function $f:[a,b]\rightarrow\mathbb{R}$ then the following are equivalent:*

1. *There exist a open set U with $[a,b]\subseteq U$ and a function $f^U:U\rightarrow X$ with $(f^U)_{|[a,b]}=f$ such that $\forall x\in U$ $(f^U)'_x$ exist and that the function*

   $$(f^U)':U\rightarrow Y \text{ defined by } (f^U)'(x)=(f^U)'_x$$

   is continuous.
2. *f has a derivative on $[a,b]$ and f' is continuous.*

Further if (1) or (2) is valid then $\forall x\in[a,b]$ $(f^U)'_x=f'$ so that $f'=((f^U)')_{|[a,b]}$

Proof.

1 ⇒ 2. Using the previous theorem [theorem: 16.145 (3)] it follows that $f=(f^U)_{|[a,b]}$ has a right derivative at a, a left derivative at b and $\forall x\in]a,b[$ f has a left and right derivative at x, further we have

$$
\begin{aligned}
f'_{+,a} &= (f^U)'_a &(16.175)\\
f'_{-,b} &= (f^U)'_b &(16.176)\\
f'_{+.x}=f'_{-,x} &= (f^U)'_x \text{ where } x\in]a,b[&(16.177)
\end{aligned}
$$

Let $x\in[a,b]\subseteq U$ then, as $(f^U)'_x$ exist, we have by [theorem: 16.33] that $((f^U)_{|]a,b[})'_x$ exists and $((f^U)_{|]a,b[})'_x=(f^U)'_x$. Further as $f_{|]a,b[}\underset{f=(f^U)_{|[a.b]}}{=}((f^U)_{|[a,.b]})_{|]a,b[}\underset{\text{[theorem: 2.84]}}{=}(f^U)_{|]a,b[}$ we have that $(f_{|]a,b[})'_x$ exist and

$$(f_{|]a,b[})'_x=((f^U)_{|]a,b[})'_x=(f^U)'_x\underset{\text{[eq: 16.177]}}{=}f'_{+,x}=f'_{-,x} \tag{16.178}$$

Hence f has a derivative on $[a,b]$ and $\forall x\in[a,b]$ we have

$$f'(x)=\begin{cases} f'_{+,a} \text{ if } x=a\\ f'_{-,b} \text{ if } x=b\\ (f_{|]a,b[})'_x \text{ if } x\in]a,b[\end{cases}\underset{\text{[eqs: 16.175, 16.176,16.178]}}{=}\begin{cases}(f^U)'_a \text{ if } x=a\\ (f^U)'_b \text{ if } x=b\\ (f^U)'_x \text{ if } x\in]a,b[\end{cases}=(f^U)'_x=(f^{U\prime})(x)$$

proving that

$$f'=((f^U)')_{[a,b]} \tag{16.179}$$

Hence as $(f^U)'$ is continuous it follows from [theorem: 14.135] that $((f^U)')_{[a,b]}:[a,b]\to\mathbb{R}$ is continuous, hence by [eq: 16.179] f' is continuous.

2 ⇒ 1. As f has a derivative on $[a,b]$ we have that $f'_+(a)$, $f'_-(a)$ exists and $\forall x\in[a,b]$ $f_{|]a,b[}$ has a derivative at x. Take $U=]a-1,b+1[$ so that $[a,b]\subseteq U$ and define

$$f^U:U\to X \text{ by } f^U(x)=\begin{cases} f(x) \text{ if } x\in[a,b]\\ f(a)+f'_{+,a}\cdot(x-a) \text{ if } x\in]a-1,a[\\ f(b)+f'_{-,b}\cdot(x-b) \text{ if } x\in]b,b+1[\end{cases} \tag{16.180}$$

then we have

$$(f^U)_{|[a,b]}=f \tag{16.181}$$

Let $x\in]a-1,b+1[$ then we have the following cases to consider:

$\boldsymbol{x\in]a-1,a[}$**.** Let $\varepsilon\in\mathbb{R}^+$. Take $\delta=a-x\in\mathbb{R}^+$ then if $h\in]a-1,b+1[_x$ and $0<|h|<\delta$ we have $x+h\in]a-1,b+1[\Rightarrow a-1<x+h$ and $h<|h|<a-x$ so that $a-1<x+h<a$. Then we have

$$
\begin{aligned}
\left\|\frac{f^U(x+h)-f(x)}{h}-f'_{+,a}\right\| &=\\
\left\|\frac{(f(a)+f'_{+,a}\cdot(x+h-a))-(f(a)+f'_{+,a}\cdot(x-a))}{h}-f'_{+,a}\right\| &=\\
\left\|\frac{\underbrace{f(a)}_{1}+\underbrace{f'_{+,a}\cdot x}_{2}+f'_{+,a}\cdot h-\underbrace{f'_{+,a}\cdot a}_{3}-\underbrace{f(a)}_{1}-\underbrace{f'_{+,a}\cdot x}_{2}+\underbrace{f'_{+,a}\cdot a}_{3}}{h}-f'_{+,a}\right\| &=\\
\left\|\frac{f'_{+,a}\cdot h}{h}-f'_{+,a}\right\| &=\\
0 &<\\
\varepsilon
\end{aligned}
$$

proving that

$$(f^U)'_x \text{ exist and } (f^U)'_x=f'_{+,a}$$

$\boldsymbol{x=a}$**.** Let $\varepsilon\in\mathbb{R}^+$. As f has a right derivative at a it follows by the definition of $f'_{+,a}$ that there exist a $\delta_1\in\mathbb{R}^+$ such that if $0<h<\delta_1$ and $a+h\in[a,b]$ we have

$$\left\|\frac{f(a+h)-f(a)}{h}-f'_{+,a}\right\|<\varepsilon \tag{16.182}$$

Take $\delta = \min(\delta_1, b-a) \in \mathbb{R}^+$. Let $h \in]a-1, b+1[_a \Rightarrow a+h \in]a-1, b+1[$ such that $0 < |h| < \delta$, then we have either:

$h < 0$. Then $a+h < a$ so that $a+h \in]a-1, a[$ hence

$$\begin{aligned}\left\|\frac{f^U(a+h)-f^U(a)}{h} - f'_{+,a}\right\| &= \left\|\frac{f(a)+f'_{+,a}\cdot(a+h-a)-f(a)}{h} - f'_{+,a}\right\| \\ &= \left\|\frac{f'_+(a)\cdot h}{h} - f'_{+,a}\right\| \\ &= 0 \\ &< \varepsilon\end{aligned}$$

$0 < h$. Then $0 < h < \delta \leqslant b-a$ so that $a < a+h < b$ proving that $a+h \in]a,b[\subseteq [a,b]$ so that

$$\begin{aligned}\left\|\frac{f^U(a+h)-f^U(a)}{h} - f'_{+,a}\right\| &= \left\|\frac{f(a+h)-f(a)}{h} - f'_{+,a}\right\| \\ &<_{0<h<\delta_1 \wedge [\text{eq: } 16.182]} \varepsilon\end{aligned}$$

So in all cases $\left\|\frac{f^U(a+h)-f^U(a)}{h} - f'_{+,a}\right\| < \varepsilon$ proving that

$$(f^U)'_x \text{ exist and } (f^U)'_x = f'_{+,a}$$

$x \in]a,b[$. Let $\varepsilon \in \mathbb{R}^+$. As $f_{|]a,b[}$ has a derivative at x there exist a $\delta_1 \in \mathbb{R}^+$ such that if $h \in]a,b[_x$ and $0 < |h| < \delta_1$ then

$$\left\|\frac{f_{|]a,b[}(x+h) - f_{|]a,b[}(x)}{h} - (f_{|]a,b[})'_x\right\| < \varepsilon \tag{16.183}$$

Take $\delta = \min(\delta_1, x-a, b-x) \in \mathbb{R}^+$ then if $0 < |h| < \delta$ we have $h < |h| < b-x \Rightarrow x+h < b$ and $-h \leqslant |-h| = |h| < x-a \Rightarrow a < x+h$ so that $x+h \in]a,b[\subseteq [a,b]$ hence we have

$$\begin{aligned}&\left\|\frac{f^U(x+h)-f^U(x)}{h} - (f_{|]a,b[})'_x\right\| &&=_{x,x+h\in]a,b[\subseteq[a,b]} \\ &\left\|\frac{f(x+h)-f(x)}{h} - (f_{|]a,b[})'_x\right\| &&=_{x,x+h\in]a,b[} \\ &\left\|\frac{f_{|]a,b[}(x+h)-f_{|]a,b[}(x)}{h} - (f_{|]a,b[})'_x\right\| &&<_{|h|<\delta_1\wedge[\text{eq: } 16.183]} \\ &\varepsilon\end{aligned}$$

proving that

$$(f^U)'_x \text{ exist and } (f^U)'_x = (f_{|]a,b[})'_x$$

$x = b$. Let $\varepsilon \in \mathbb{R}^+$. As f has a left derivative at b it follows by the definition of $f'_{-,b}$ that there exist a $\delta_1 \in \mathbb{R}^+$ such that if $0 < h < \delta_1$ and $b-h \in [a,b]$ we have

$$\left\|\frac{f(b-h)-f(b)}{-h} - f'_{-,b}\right\| < \varepsilon \tag{16.184}$$

Take $\delta = \min(\delta_1, b-a) \in \mathbb{R}^+$. Let $h \in]a-1, b+1[_b \Rightarrow b+h \in]a-1, b+1[$ such that $0 < |h| < \delta, \delta_1$, then we have either:

$h < 0$. Then $-h = |h| < b-a \Rightarrow a < b+h < b \Rightarrow b+h \in [a,b]$ hence

$$\begin{aligned}&\left\|\frac{f^U(b+h)-f^U(b)}{h} - f'_{-,b}\right\| &&=_{b,b+h\in[a.b]} \\ &\left\|\frac{f(b+h)-f(b)}{h} - f'_-(b)\right\| &&=_{-|h|=h} \\ &\left\|\frac{f(b-|h|)-f(b)}{-|h|} - f'_-(b)\right\| &&<_{|h|<\delta_1\wedge[\text{eq: } 16.184]} \\ &\varepsilon\end{aligned}$$

$0<h$. then $b<b+h<b+1$ so that $b+h\in]b,b+1[$ hence

$$\begin{array}{rl}\left\|\frac{f^U(b+h)-f^U(b)}{h}-f'_{-,b}\right\| & = \\ \left\|\frac{f(b)+f'_{-,b}\cdot(b+h-b)-f(b)}{h}-f'_{-,b}\right\| & = \\ \left\|\frac{f'_{-,b}\cdot h}{h}-f'_{-,b}\right\| & = \\ 0 & < \\ \varepsilon & \end{array}$$

So in all cases we have $\left\|\frac{f^U(b+h)-f^U(b)}{h}-f'_{-,b}\right\|<\varepsilon$ proving that

$$(f^U)'_x \text{ exist and } (f^U)'_x=f'_{-,b}$$

$x\in]b,b+1[$. Let $\varepsilon\in\mathbb{R}^+$. Take $\delta=x-b\in\mathbb{R}^+$ then if $h\in]a-1,b+1[_x$ and $0<|h|<\delta$ we have $x+h\in]a-1,b+1[$ and $-h\leqslant|h|<x-b\Rightarrow b<x+h$ so that $x+h\in]b,b+1[$. So that

$$\begin{array}{rl}\left\|\frac{f^U(b+h)-f^U(b)}{h}-f'_{-,b}\right\| & = \\ \left\|\frac{f(b)+f'_{-,b}\cdot(x+h-b)-(f(b)+f'_{-,b}\cdot(x-b))}{h}-f'_{-,b}\right\| & = \\ \left\|\frac{\underbrace{f(b)}_{1}+\underbrace{f'_{-,b}\cdot x}_{2}+f'_{-,b}\cdot h-\underbrace{f'_{-,b}\cdot b}_{3}-\underbrace{f(b)}_{1}-\underbrace{f'_{-,b}\cdot x}_{2}+\underbrace{f'_{-,b}\cdot b}_{3}}{h}-f'_{-,b}\right\| & = \\ \left\|\frac{f'_{-,b}\cdot h}{h}-f'_{-,b}\right\| & = \\ 0 & < \\ \varepsilon & \end{array}$$

which proves that

$$(f^U)'_x \text{ exists and } (f^U)'_x=f'_{-,b}$$

To summarize we have $\forall x\in]a-1,b+1[$ that

$$(f^U)'_x \text{ exist and } (f^U)'_x=\left\{\begin{array}{l}f'_{+,a} \text{ if } x\in]a-1,a]\\ f'_{-,b} \text{ if } x\in[b,b+1[\\ (f_{|]a,b[})'_x \text{ if } x\in]a,b[\end{array}\right. \underset{\text{def}}{=} f'(x) \tag{16.185}$$

Next we have to prove that the function

$$(f^U)'\colon U\rightarrow X \text{ defined by } (f^U)'(x)=(f^U)'_x$$

is continuous. Let $x\in U=]a-1,b+1[$ then we have either:

$x\in]a-1,a[$. Let $\varepsilon\in\mathbb{R}^+$. Take $\delta=\min(a-x,x-(a-1))\in\mathbb{R}^+$ then for $y\in]a-1,b+1[$ with $|x-y|<\delta$ we have $x-y\leqslant|x-y|<\delta<x-(a-1)\wedge y-x\leqslant|x-y|<a-x$ so that $a-1<y\wedge y<a$ or $y\in]a-1,a[$. Hence

$$\|(f^U)'(y)-(f^U)'(x)\|=\|(f^U)'_y-(f^U)'_x\| \underset{[\text{eq: } 16.185]}{=} \|f'_{+,a}-f'_{+,a}\|=0<\varepsilon$$

proving that $(f^U)'$ is continuous at x in this case.

$x=a$. Let $\varepsilon\in\mathbb{R}^+$. As f' is continuous at a there exist a $\delta_1\in\mathbb{R}^+$ such that $\forall y\in[a,b]$ with $|y-a|<\delta_1$ we have $\|f'(y)-f'(a)\|<\varepsilon$. Take $\delta=\min(\delta_1,b-a)\in\mathbb{R}^+$ then for $y\in]a-1,b+1[$ with $|a-y|<\delta$ we have either:

$y\leqslant a$. Then $y\in]a-1,a]$ so that

$$\|(f^U)'(y)-(f^U)'(x)\|=\|(f^U)'_y-(f^U)'_a\| \underset{[\text{eq: } 16.185]}{=} \|f'_{+,a}-f'_{+,a}\|=0<\varepsilon$$

$a < y$. Then as $y - a \leqslant |y - a| < b - a \Rightarrow y < b$ we have $y \in]a, b[$ so that

$$\|(f^U)'(y) - (f^U)'(x)\| = \|(f^U)'_y - (f^U)'_a\| \underset{[\text{eq: } 16.185]}{=} \|f'(y) - f'(a)\| < \varepsilon$$

so in all cases we have $\|(f^U)'(y) - (f^U)'(x)\| < \varepsilon$, proving that $(f^U)'$ is continuous at x in this case.

$x \in]a, b[$. Let $\varepsilon \in \mathbb{R}^+$. As f' is continuous at x there exist a $\delta_1 \in \mathbb{R}^+$ such that $\forall y \in [a, b]$ with $|y - x| < \delta_1$ we have $\|f'(y) - f'(x)\| < \varepsilon$. Take $\delta = \min(\delta_1, x - a, b - x)$ then for $y \in]a - 1, b + 1[$ with $|x - y| < \delta$ we have $x - y < x - a \wedge y - x < b - x$ so that $\in]a, b[$ and $|y - x| < \delta_1$. Hence we have

$$\begin{aligned} \|(f^U)'(y) - (f^U)'(x)\| &= \|(f^U)'_y - (f^U)'_x\| \\ &\underset{[\text{eq: } 16.185]}{=} \|f'(y) - f'(x)\| \\ &< \varepsilon \end{aligned}$$

proving that $(f^U)'$ is continuous at x in this case.

$x = b$. Let $\varepsilon \in \mathbb{R}^+$. As f' is continuous at b there exist a $\delta_1 \in \mathbb{R}^+$ such that $\forall y \in [a, b]$ with $|y - b| < \delta_1$ we have $\|f'(y) - f'(b)\| < \varepsilon$. Take $\delta = \min(\delta_1, b - a)$ then for $y \in]a - 1, b + 1[$ with $|y - b| < \delta$ we have the following cases to consider:

$b \leqslant y$. Then we have $y \in [b, b - 1[$ so that

$$\|(f^U)'(y) - (f^U)'(x)\| = \|(f^U)'_y - (f^U)'_b\| \underset{[\text{eq: } 16.185]}{=} \|f'_{-,b} - f'_{-,b}\| = 0 < \varepsilon$$

$y < b$. Then, as $|y - b| < \delta$, we have $b - y \leqslant |y - b| < b - a$ so that $y \in]a, b[$ and $|y - b| < \delta_1$. Hence

$$\|(f^U)'(y) - (f^U)'(x)\| = \|(f^U)'_y - (f^U)'_b\| \underset{[\text{eq: } 16.185]}{=} \|f'(y) - f'(b)\| < \varepsilon$$

so in all cases we have $\|(f^U)'(y) - (f^U)'(x)\| < \varepsilon$ proving that $(f^U)'$ is continuous at x in this case.

$x \in]b, b + 1[$. Let $\varepsilon \in \mathbb{R}^+$. Take $\delta = \min(b + 1 - x, x - b) \in \mathbb{R}^+$ then for $y \in]a - 1, b + 1[$ with $|y - x| < \delta$ we have $y - x \leqslant |y - x| < b + 1 - x \wedge x - y \leqslant |y - x| < x - b$ so that $y \in]b, b + 1[$. Hence

$$\|(f^U)'(y) - (f^U)'(x)\| = \|(f^U)'_y - (f^U)'_x\| \underset{[\text{eq: } 16.185]}{=} \|f'_{-,b} - f'_{-,b}\| = 0 < \varepsilon$$

proving that $(f^U)'$ is continuous at x in this case.

So we have proved that $\forall x \in [a - 1, b + 1]$ $(f^U)'$ is continuous at x,hence $(f^U)'$ is continuous. □

Corollary 16.151. *Let $\langle \mathbb{R}, \|\| \rangle$ be the normed space of the real numbers, $\langle X, \|\|\| \rangle$ a normed space, $a, b \in \mathbb{R}$ with $a < b$ and $f: [a, b] \rightarrow \mathbb{R}$ a function then the following are equivalent:*

1. *There exist a open set U in $\mathbb{R}$ with $[a, b] \subseteq U$ and a function $f^U: U \rightarrow X$ of class C^1 such that $(f^U)_{|[a,b]} = f$.*
2. *f has a derivative on $[a, b]$ and $f': [a.b] \rightarrow X$ is continuous.*

Further if (1) or (2) is valid then $\forall x \in [a, b]$ $(f^U)'_x = f'$ so that $f' = ((f^U)')_{|[a,b]}$

Proof. This follows by applying [theorem: 16.100] on the previous theorem [theorem: 16.150] □

16.3.3 Fundamental theorem of Calculus

We examine now the relation between the Riemann integral and the derivative of a function.

Theorem 16.152. (Fundamental Theorem of Calculus (I)) *Let $\langle \mathbb{R}, \|\| \rangle$ be the normed space of the real numbers, $\langle X, \|\|\| \rangle$ be a Banach space, $a, b \in \mathbb{R}$ with $a < b$ and $f: [a, b] \rightarrow X$ a continuous function [so that by [theorem: 15.31] $\forall x \in [a, b]$ $\int_a^x f_{|[a,x]}$ exist] then for the function*

$$F: [a, b] \rightarrow X \text{ defined by } F(x) = \int_a^x f_{|[a,x]}$$

we have that F has a derivative on $[a, b]$ and $F' = f$

Proof. As $x \in [a,b]$ we have to consider the following cases:

$\boldsymbol{x=a}$**.** Take $\varepsilon \in \mathbb{R}^+$. As f is continuous at a there exist a $\delta \in \mathbb{R}^+$ such that if $y \in [a,b]$ with $|y-a|<\delta$ it follows that $\|f(y)-f(a)\|<\frac{\varepsilon}{2}$. So if $0<h<\delta$ and $h \in [a,b]_a \Rightarrow a+h \in [a,b]$ we have $\forall y \in [a,a+h]$ that $|y-a| = y-a < a+h-a = h < \delta$ so that $\|f(y)-f(a)\|<\frac{\varepsilon}{2}$. Hence, as $f(y)-f(a) = f(y) - C_{f(a)}(y) = (f - C_{f(a)})(y)$, we have

$$\forall y \in [a,a+h] \text{ we have } \|(f-C_{f(a)})(y)\| = \|f(y)-f(a)\| < \frac{\varepsilon}{2} \tag{16.186}$$

Next

$$\begin{aligned}
\left\|\frac{F(a+h)-F(a)}{h} - f(a)\right\| &= \left\|\frac{\int_a^{a+h} f - \int_a^a f}{h} - f(a)\right\| \\
&\underset{[\text{definition: } 15.27]}{=} \left\|\frac{\int_a^{a+h} f}{h} - f(a)\right\| \\
&= \left\|\frac{\left(\int_a^{a+h} f\right) - h\cdot f(a)}{h}\right\| \\
&\underset{[\text{example: } 15.22]}{=} \left\|\frac{\left(\int_a^{a+h} f\right) - \int_a^{a+h} C_{f(a)}}{h}\right\| \\
&\underset{[\text{theorem: } 15.28]}{=} \left\|\frac{\int_a^{a+h} (f - C_{f(a)})}{h}\right\| \\
&\underset{0<h}{=} \frac{\left\|\int_a^{a+h} (f - C_{f(a)})\right\|}{h} \\
&\leqslant_{[\text{theorem: } 15.29] \wedge [\text{eq: } 16.186]} \frac{(a+h-a)\cdot\frac{\varepsilon}{2}}{h} \\
&= \frac{\varepsilon}{2} \\
&< \varepsilon
\end{aligned}$$

proving that F has a right derivative at a and

$$F'_+(a) = f(a)$$

$\boldsymbol{x=b}$**.** Take $\varepsilon \in \mathbb{R}^+$. As f is continuous at b there exist a $\delta \in \mathbb{R}^+$ such that if $y \in [a,b]$ with $|y-b|<\delta$ it follows that $\|f(y)-f(b)\|<\frac{\varepsilon}{2}$. Hence if $0<h<\delta$ we have for $-h \in [a,b]_b \Rightarrow b-h \in [a,b]$ that $\forall y \in [b-h,b]$ $|y-b| = b-y \leqslant b+h-b = h < \delta$ so that $\|f(y)-f(b)\|<\frac{\varepsilon}{2}$. Hence, as $f(y)-f(b) = f(y) - C_{f(b)}(y) = (f - C_{f(b)})(y)$, we have

$$\forall y \in [b-h,b] \text{ we have } \|(f-C_{f(b)})(y)\| < \frac{\varepsilon}{2} \tag{16.187}$$

Next

$$\begin{aligned}
\left\|\frac{F(b-h)-F(b)}{-h} - f(b)\right\| &= \\
\left\|\frac{\left(\int_a^{b-h} f\right) - \int_a^b f}{-h} - f(b)\right\| &\underset{[\text{theorem: } 15.31]}{=} \\
\left\|\frac{\left(\int_a^{b-h} f\right) - \left(\left(\int_a^{b-h} f\right) + \int_{b-h}^b f\right)}{-h} - f(b)\right\| &= \\
\left\|\frac{-\int_{b-h}^b f}{-h} - f(b)\right\| &= \\
\left\|\frac{\left(\int_{b-h}^b f\right) - h\cdot f(b)}{h}\right\| &\underset{[\text{example: } 15.22]}{=} \\
\left\|\frac{\left(\int_{b-h}^b f\right) - \left(\int_{b-h}^b C_{f(b)}\right)}{h}\right\| &\underset{[\text{theorem: } 15.28]}{=}
\end{aligned}$$

$$\begin{aligned}
\left\|\frac{\int_{b-h}^{b}(f-C_{f(b)})}{h}\right\| & \underset{0<h}{=} \\
\frac{\left\|\int_{b-h}^{b}(f-C_{f(b)})\right\|}{h} & \leqslant_{[\text{theorem: } 15.29]\wedge[\text{eq: } 16.187]} \\
\frac{(b-(b-h))}{h}\cdot\frac{\varepsilon}{2} & = \\
\frac{\varepsilon}{2} & < \\
\varepsilon
\end{aligned}$$

proving that F has a left derivative at b and

$$F'_-(b) = f(b)$$

***x* ∈]*a*, *b*[.** Let $\varepsilon \in \mathbb{R}^+$. As f is continuous at x there exist a $\delta \in \mathbb{R}^+$ such that $\forall h \in [a, b]$ with $|y-x| < \delta$ we have $\|f(y)-f(x)\| < \frac{\varepsilon}{2}$. If h satisfies $0 < |h| < \delta$ and $h \in]a, b[_x \Rightarrow x+h \in]a, b[$ then we have the following cases to consider:

0 < *h*. If $y \in [x, x+h]$ then $|y-x| = y-x \leqslant x+h-x = h = |h| < \delta$ so that $\|f(y)-f(x)\| < \frac{\varepsilon}{2}$. Hence, as $f(y)-f(x) = f(y)-C_{f(x)}(y) = (f-C_{f(x)})(y)$. It follows that

$$\forall y \in [x, x+h] \text{ we have } \|(f-C_{f(x)})(y)\| < \frac{\varepsilon}{2} \tag{16.188}$$

Next

$$\begin{aligned}
\left\|\frac{F_{|]a,b[}(x+h)-F_{|]a,b[}(x)}{h}-f(x)\right\| & \underset{x+h,x\in]a,b[}{=} \\
\left\|\frac{F(x+h)-F(x)}{h}-f(x)\right\| & = \\
\left\|\frac{\left(\int_a^{x+h} f\right)-\int_a^x f}{h}-f(x)\right\| & \underset{[\text{theorem: } 15.31]}{=} \\
\left\|\frac{\left(\int_a^x f\right)+\left(\int_x^{x+h} f\right)-\int_a^x f}{h}-f(x)\right\| & = \\
\left\|\frac{\int_x^{x+h} f}{h}-f(x)\right\| & = \\
\left\|\frac{\left(\int_x^{x+h} f\right)-h\cdot f(x)}{h}\right\| & \underset{[\text{example: } 15.22]}{=} \\
\left\|\frac{\left(\int_x^{x+h} f\right)-\int_x^{x+h} C_{f(x)}}{h}\right\| & \underset{[\text{theorem: } 15.28]}{=} \\
\left\|\frac{\int_x^{x+h}(f-C_{f(x)})}{h}\right\| & \underset{0<h}{=} \\
\frac{\left\|\int_x^{x+h}(f-C_{f(x)})\right\|}{h} & \leqslant_{[\text{theorem: } 15.29]\wedge[\text{eq: } 16.188]} \\
\frac{(x+h-x)}{h}\cdot\frac{\varepsilon}{2} & = \\
\frac{\varepsilon}{2} & < \\
\varepsilon
\end{aligned}$$

***h* < 0.** In this case $h = -|h|$. If $y \in [x+h, x] = [x-|h|, x]$ then

$$|y-x| = x-y < x-(x-|h|) = |h| < \delta$$

and thus $\|f(y)-f(x)\|<\frac{\varepsilon}{2}$. Hence, as $f(y)-f(x)=f(y)-C_{f(x)}(y)=(f-C_{f(x)})(y)$. It follows that

$$\forall y\in[x+h,x] \text{ we have } \|(f-C_{f(x)})(y)\|<\frac{\varepsilon}{2} \tag{16.189}$$

Next

$$\begin{array}{rcl}
\left\|\frac{F_{|]a,b[}(x+h)-F_{|]a,b[}(x)}{h}-f(x)\right\| & \underset{x+h,x\in]a,b[}{=} & \\
\left\|\frac{F(x+h)-F(x)}{h}-f(x)\right\| & = & \\
\left\|\frac{F(x-|h|)-F(x)}{-|h|}-f(x)\right\| & = & \\
\left\|\frac{\left(\int_a^{x-|h|}f\right)-\int_a^x f}{-|h|}-f(x)\right\| & \underset{[\text{theorem: } 15.31]}{=} & \\
\left\|\frac{\left(\int_a^{x-|h|}f\right)-\left(\int_a^{x-|h|}f\right)-\int_{x-|h|}^x f}{-|h|}\right\| & = & \\
\left\|\frac{-\int_{x-|h|}^x f}{-|h|}-f(x)\right\| & = & \\
\left\|\frac{\int_{x-|h|}^x f}{|h|}-f(x)\right\| & = & \\
\left\|\frac{\left(\int_{x-|h|}^x f\right)-|h|\cdot f(x)}{|h|}\right\| & \underset{[\text{example: } 15.22]}{=} & \\
\left\|\frac{\left(\int_{x-|h|}^x f\right)-\int_{x-|h|}^x C_{f(x)}}{|h|}\right\| & \underset{[\text{theorem: } 15.28]}{=} & \\
\left\|\frac{\int_{x-|h|}^x (f-C_{f(x)})}{|h|}\right\| & \underset{0<|h|}{=} & \\
\frac{\left\|\int_{x-|h|}^x (f-C_{f(x)})\right\|}{|h|} & \leqslant_{[\text{theorem: } 15.29]\wedge[\text{eq: } 16.189]} & \\
\frac{x-(x-|h|)}{|h|}\cdot\frac{\varepsilon}{2} & = & \\
\frac{\varepsilon}{2} & < & \\
\varepsilon & &
\end{array}$$

Hence in all cases we have $\left\|\frac{F_{|]a,b[}(x+h)-F_{|]a,b[}(x)}{h}-f(x)\right\|<\varepsilon$ proving that $F_{|]a,b[}$ has a derivative on $]a,b[$ and $\forall x\in]a,b[$

$$(F_{|]a,b[})'(x)=f(x)$$

So by [definition: 16.146] we have that

$$F \text{ has a derivative on } [a,b] \text{ and } F'=f$$ □

We use now the above theorem to prove the Fundamental Theorem of Calculus.

Theorem 16.153. (Fundamental Theorem of Calculus (II)) *Let $\langle\mathbb{R},\|\rangle$ be the normed space of real numbers, $\langle X,\|\|\rangle$ be a Banach space, $a,b\in\mathbb{R}$ with $a<b$ and $f:[a,b]\rightarrow X$ a function such that f has a derivative on $[a,b]$ and f' is continuous then*

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

Proof. As f has a derivative on $[a, b]$, $f'\colon [a, b] \rightarrow X$ is continuous and $\langle X, \|\|\rangle$ is a Banach space, it follows from [theorem: 15.31] that $\forall x \in [a, b]$ $\int_a^x f'$ is well defined. This allows us to define

$$g\colon [a, b] \rightarrow X \text{ by } g(x) = f(a) + \int_a^x f'$$

So if we define

$$F\colon [a, b] \rightarrow X \text{ by } F(x) = \int_a^x f'$$

then we have that

$$g = C_{f(a)} + F$$

By [example: 16.148] $C_{f(a)}$ has a derivative on $[a, b]$ with $(C_{f(a)}) = C_0$. using [theorem: 16.152] F has a derivative on $[a, b]$ with $F' = f'$. Using then the linearity of derivation [see theorem: 16.149] it follows that g has a derivative on $[a, b]$ with

$$g' = (C_{f(a)})' + F' = C_0 + f' = f' \tag{16.190}$$

Note also that $g(a) = C_{f(a)} + f(a) = f(a) + \int_a^a f' \underset{[\text{definition: } 15.27]}{=} f(a) + 0 = f(a)$ or

$$f(a) = g(a) \tag{16.191}$$

Define now the function φ by

$$\varphi\colon [a, b] \rightarrow X \text{ where } \varphi(x) = \|f(x) - g(x)\|$$

Let $x \in]a, b[$ then as $]a, b[$ is open there exist a $\delta \in \mathbb{R}^+$ such that $B_{||}(x, \delta) \subseteq]a, b[$. If $0 < |h| < \delta$ then $|x + h - x| = |h| < \delta$ so that $x + h \in B_{||}(x, \delta) \subseteq]a, b[$. Then we have

$$\begin{aligned}
\left|\frac{\varphi(x+h) - \varphi(x)}{h}\right| &= \left|\frac{\|f(x+h) - g(x+h)\| - \|f(x) - g(x)\|}{h}\right| \\
&\leqslant_{\langle \text{smart-ref}|\text{normed absolute value norm differences}\rangle} \frac{\|f(x+h) - g(x+h) + (g(x) - f(x))\|}{|h|} \\
&= \left\|\frac{(f(x+h)) - f(x) - (g(x+h) - g(x))}{h}\right\| \\
&\underset{f' = g' [\text{see } 16.190]}{=} \left\|\frac{f(x+h) - f(x)}{h} - f'(x) - \left(\frac{g(x+h) - g(x)}{h} - g'(x)\right)\right\| \\
&\leqslant \left\|\frac{f(x+h) - f(x)}{h} - f'(x)\right\| + \left\|\frac{g(x+h) - g(x)}{h} - g'(x)\right\|
\end{aligned}$$

Summarized, if $0 < |h| < \delta$ then

$$\left|\frac{\varphi(x+h) - \varphi(x)}{h}\right| \leqslant \left\|\frac{f(x+h) - f(x)}{h} - f'(x)\right\| + \left\|\frac{g(x+h) - g(x)}{h} - g'(x)\right\| \tag{16.192}$$

Let $\varepsilon \in \mathbb{R}^+$. As f has a derivative on $[a, b]$ we have by definition that $f_{|]a,b[}$ has a derivative at x. Hence there exist a $\delta_1 \in \mathbb{R}^+$ such that if $0 < |h| < \delta_1$ and $h \in]a, b[_x \Rightarrow x, x + h \in]a, b[$ that

$$\left\|\frac{f(x+h) - f(x)}{h} - f'(x)\right\| = \left\|\frac{f_{|]a,b[}(x+h) - f_{|]a,b[}(x)}{h} - (f_{|]a,b[})'(x)\right\| < \frac{\varepsilon}{2}$$

Likewise as g has a derivative on $[a, b]$ we have by definition that $g_{|]a,b[}$ has a derivative at x. Hence there exist a $\delta_2 \in \mathbb{R}^+$ such that if $0 < |h| < \delta_2$ and $h \in]a, b[_x \Rightarrow x, x + h \in]a, b[$ that

$$\left\|\frac{g(x+h) - g(x)}{h} - g'(x)\right\| = \left\|\frac{g_{|]a,b[}(x+h) - g_{|]a,b[}(x)}{h} - (g_{|]a,b[})'(x)\right\| < \frac{\varepsilon}{2}$$

So if $0 < |h| < \min(\delta, \delta_1, \delta_2)$ and $h \in]a, b[_x$ we have using the above and [eq: 16.192] that

$$\left|\frac{\varphi_{|]a,b[}(x+h) - \varphi_{|]a,b[}(x)}{h} - 0\right| \underset{x, x+h \in]a,b[}{=} \left|\frac{\varphi(x+h) - \varphi(x)}{h} - 0\right| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

which has a derivative at x with $\varphi'_x = 0$. Hence

$$\forall x \in]a, b[\ \varphi \text{ has a derivative at } x \text{ and } (\varphi_{|]a,b[|})'_x = 0$$

Let $x, y \in [a, b]$ then we can use Lagrange's theorem [see theorem: 16.135] to find a $\zeta \in]a, b[$ with

$$\varphi(x) - \varphi(y) = (\varphi_{|]a,b[})'_{\zeta} \cdot (x - y) = 0 \cdot (x - y) = 0$$

This prove that $\forall x \in [a, b]$ $\varphi(x) = \varphi(a) = \|f(a) - g(a)\| \underset{[\text{eq: } 16.191]}{=} 0$. Hence $\forall x \in [a, b]$ we have

$$\|f(x) - g(x)\| = \varphi(x) = 0$$

proving that $f(x) = g(x)$ so that $f(b) = g(b) = f(a) + \int_a^b f'$. Hence

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

□

16.3.4 Mean Value Theorems

Mean Value Theorems put a maximum on the variation of function values based on the variation of the function arguments. In essence they state that $\|f(x) - f(y)\|_Y \leqslant C \cdot \|x - y\|_X$ where C is a constant depending on the differential of f. A first example of a Mean Value Theorem is Lagrange's Mean Value theorem [see theorem: 16.135] (where we actually have a equality instead of a inequality).

Lemma 16.154. *Let $\langle X, \|\|\rangle$ be a normed vector space, $\langle \mathbb{R}, \|\rangle$ the normed spaces of real numbers, $x, y \in X$ then we have for the function*

$$\sigma: \mathbb{R} \rightarrow X \textit{ defined by } \sigma(t) = x + t \cdot (y - x)$$

that σ is of class C^1. Further $\forall t \in \mathbb{R}$ we have that $D_t\sigma \in L(\mathbb{R}, X)$ is defined by

$$D_t\sigma: \mathbb{R} \rightarrow X \textit{ where } D_t\sigma(h) = h \cdot (y - x)$$

so that

$$\sigma'_t \underset{[\textit{theorem: } 16.32]}{=} D_t\sigma(1) = y - x$$

Proof. Define

$$L: \mathbb{R} \rightarrow X \text{ by } L(h) = h \cdot (y - x)$$

then by [theorem: 14.183] $L \in L(X, \mathbb{R})$. Let $t \in \mathbb{R}$, $\varepsilon \in \mathbb{R}^+$ then if $h \in \mathbb{R}_x$ with $0 < |h| < 1$ we have that

$$\begin{array}{rcl} \|\sigma(t + h) - \sigma(t) - L(h)\| & = & \|x + (t + h) \cdot (y - x) - (x + t \cdot (y - x)) - h \cdot (y - x)\| \\ & = & \|h \cdot (y - x) - h \cdot (y - x)\| \\ & = & 0 \\ & \leqslant & \varepsilon \cdot |h| \end{array}$$

proving that σ is differentiable at t with $D_t\sigma = L$. Hence

$$\forall t \in \mathbb{R} \ \sigma \text{ is differentiable at } t \text{ where } D_t f: \mathbb{R} \rightarrow X \text{ is defined by } D_t\sigma(h) = h \cdot (y - x) \tag{16.193}$$

Using [theorem: 16.31] it follows that

$$\sigma \text{ is a continuous function}$$

Further for the function

$$D^{[1]}\sigma: \mathbb{R} \rightarrow L_1(\mathbb{R}; X) \text{ defined by } D^{[1]}\sigma(t) = D_t\sigma$$

we have, as $\forall t \in \mathbb{R}$ $D_t\sigma = L$, that $D^{[1]}\sigma = C_L$ a constant function. Hence using [theorem: 14.131] $D^{[1]}\sigma$ is continuous, which combined with [eq: 16.193] proves that

$$\sigma \text{ is of class } C^1$$

□

Theorem 16.155. (Mean Value Theorem (I)) *Let $\langle X, \|\|\rangle$ be a normed vector space, U a open set U in X, $\langle \mathbb{R}, \|\rangle$ the normed spaces of real numbers and $f: U \rightarrow \mathbb{R}$ of class C^1. If $x, y \in U$ such that $\forall t \in [0, 1]$ $(1 - t) \cdot x + t \cdot y \in U$ then there exist a $\zeta \in [0, 1]$ such that*

$$f(y) - f(x) = D_z f(y - x) \textit{ where } z = x + \zeta \cdot (x - y)$$

Proof. Define

$$\sigma: \mathbb{R} \rightarrow X \text{ by } \sigma(t) = x + t \cdot (y - x) = x \cdot (1 - t) + t \cdot y$$

Take $W=\sigma^{-1}(U)$ then as σ is continuous [see lemma: 16.154] it follows that W is open. Let $t\in[0,1]$ then $\sigma(t)=(1-t)\cdot x+y\in U$ so that $\sigma([0,1])\subseteq U$ hence

$$[0,1]\subseteq\sigma^{-1}(U)=W$$

Then for

$$\sigma_{|W}: W\rightarrow X$$

we have that $\sigma_{|W}(W)=\sigma(W)=\sigma(\sigma^{-1}(U))\subseteq U$. Hence the following function is well defined:

$$f\circ\sigma_{|W}: W\rightarrow\mathbb{R}$$

Let $t\in W$ then by [see lemma: 16.154] σ is differentiable at x with $D_t\sigma\in L(\mathbb{R},X)$ defined by

$$D_t\sigma(h)=h\cdot(y-x).$$

Hence using [theorem: 16.24] it follows that $\sigma_{|W}$ is differentiable a t with $D_t\sigma_{|W}=D_t\sigma$. Further, as f is of class C^1, f is differentiable at $\sigma(t)$, hence using the chain rule [theorem: 16.36] $f\circ\sigma_{|W}$ is differentiable at t and

$$D_t(f\circ\sigma_{|W})=D_{\sigma_{|W}(t)}f\circ D_t\sigma\underset{t\in W}{=}D_{\sigma(t)}f\circ D_t\sigma \tag{16.194}$$

Define now g by

$$g:[0,1]\rightarrow\mathbb{R}\text{ by }g=(f\circ\sigma_{|W})_{|[0,1]}$$

As f is of class C^1, hence of class C^0 [see theorem: 16.107], it follows that f is continuous. Further σ is continuous, hence by [theorem: 14.135] $\sigma_{|w}$ is continuous so that $f\circ\sigma_{|W}$ is continuous. Using [theorem: 14.135] again we have that

$$g\text{ is continuous.} \tag{16.195}$$

Further for $t\in]0,1[$ we have $g_{|]0,1]}(t)\underset{t\in]0,1[}{=}g(t)\underset{t\in]0,1[\subseteq[0,1]}{=}(f\circ\sigma_{|W})(t)\underset{t\in]0,1]}{=}(f\circ\sigma_{|W})_{]0,1[}$ so that $g_{|]0,1[}=(f\circ\sigma_{|W})_{]0,1[}$. As $]0,1[\subseteq W$ and $f\circ\sigma_{|W}$ is differentiable at $t\in W$ it follows from [theorem: 16.24] that $(f\circ\sigma_{|W})_{]0,1[}$ is differentiable at t with

$$D_t(f\circ\sigma_{|W})_{]0,1[}=D_t(f\circ\sigma_{|W})\underset{\text{[eq: 16.194]}}{=}D_{\sigma(t)}f\circ D_t\sigma.$$

Hence

$$\forall t\in]0,1[\ g_{|]0,1[}\text{ is differentiable at }t\text{ with }D_t(g_{|]0,1[})=D_{\sigma(t)}f\circ D_t\sigma \tag{16.196}$$

Because of [eqs: 16.195 and 16.196] we can use Lagrange's theorem [theorem: 16.135] to find a $\zeta\in]0,1[$ so that

$$g(1)-g(0)=(g_{]0,1[})'_\zeta\cdot(1-0)=(g_{]0,1[})'_\zeta \tag{16.197}$$

Take $z=x+\zeta\cdot(y-x)=\sigma(\zeta)$ then we have

$$\begin{array}{rcl}
f(y)-f(x) & = & f(x+1\cdot(y-x))-f(x+0\cdot(y-x))\\
 & = & f(\sigma(1))-f(\sigma(0))\\
 & \underset{1,0\in[0,1]}{=} & f(\sigma_{|W}(1))-f(\sigma_{|W}(0))\\
 & = & (f\circ\sigma_{|W})(1)-(f\circ\sigma_{|w})(0)\\
 & \underset{1,0\in[0,1]}{=} & (f\circ\sigma_{|W})_{|[0,1]}(1)-(f\circ\sigma_{|w})_{|[0,1]}(0)\\
 & = & g(1)-g(0)\\
 & \underset{\text{[eq: 16.197]}}{=} & (g_{]0,1[})'_\zeta\\
 & \underset{\text{[theorem: 16.32]}}{=} & D_\zeta(g_{|]0,1[})(1)\\
 & \underset{\text{[eq: 16.196]}}{=} & (D_{\sigma(\zeta)}f\circ D_\zeta\sigma)(1)\\
 & = & D_{\sigma(\zeta)}f(D_\zeta\sigma(1))\\
 & \underset{\text{[lemma: 16.154]}}{=} & D_{\sigma(\zeta)}f(y-x)\\
 & \underset{z=\sigma(\zeta)}{=} & D_zf(y-x)
\end{array}$$

□

We generalize now the Mean Value Theorem to arbitrary normed spaces, first we need the concept of convex sets.

Definition 16.156. *Let X be a vector space over $\mathbb{K}$ then $C \subseteq X$ is convex if $\forall x, y \in C$ we have $\forall t \in [0,1]$ that $x + t \cdot (y - x) = (1 - t) \cdot x + t \cdot y \in C$*

Theorem 16.157. *Let $\langle X, \|\|\rangle$ be a normed space, $x \in X$ and $\delta \in \mathbb{R}^+$ then $B_{\|\|}(x, \delta)$ and $\overline{B_{\|\|}(x, \delta)}$ are convex.*

Proof. Let $x \in X$, $\delta \in \mathbb{R}^+$, $y, z \in X$ and $t \in [0,1]$ then we have

$$\begin{array}{rcl} \|(1-t)\cdot y + t\cdot z - x\| & = & \|(1-t)\cdot y + t\cdot z - x + t\cdot x - t\cdot x\| \\ & = & \|(1-t)\cdot y - x\cdot(1-t) + t\cdot z - t\cdot x\| \\ & = & \|(1-t)\cdot(y-x) + t\cdot(z-x)\| \\ & \leqslant & \|(1-t)\cdot(y-x)\| + \|t\cdot(z-x)\| \\ & = & |1-t|\cdot\|y-x\| + |t|\cdot\|z-x\| \\ & \underset{0\leqslant 1-t,t}{=} & (1-t)\cdot\|y-x\| + t\cdot\|z-x\| \end{array} \quad (16.198)$$

Hence if $y, z \in B_{\|\|}(x, \delta)$ then

$$\begin{array}{rcl} \|(1-t)\cdot y + t\cdot z - x\| & \leqslant_{[\text{eq: } 16.198]} & (1-t)\cdot\|y-x\| + t\cdot\|z-x\| \\ & < & (1-t)\cdot\delta + t\cdot\delta \\ & = & \delta \end{array}$$

proving that

$$(1-t)\cdot y + t\cdot z \in B_{\|\|}(x, \delta)$$

Further if $y, z \in \overline{B_{\|\|}(x, \delta)}$ then

$$\begin{array}{rcl} \|(1-t)\cdot y + t\cdot z - x\| & \leqslant_{[\text{eq: } 16.198]} & (1-t)\cdot\|y-x\| + t\cdot\|z-x\| \\ & \leqslant & (1-t)\cdot\delta + t\cdot\delta \\ & = & \delta \end{array}$$

so that

$$(1-t)\cdot y + t\cdot z \in \overline{B_{\|\|}(x, \delta)}$$

□

Theorem 16.158. (Mean Value Theorem (II)) *Let $\langle X, \|\|_X\rangle$ a normed space, $\langle Y, \|\|_Y\rangle$ a Banach space, U a open set in X and $f: U \rightarrow Y$ a function of class C^1. Assume that there is a* ***convex subset*** *$C \subseteq U$ and a $k \in \mathbb{R}$ such that $\forall x \in C$ $\|D_x f\|_{L(X,Y)} \leqslant k$ then*

$$\forall x, y \in C \text{ we have } \|f(y) - f(x)\|_Y \leqslant k \cdot \|y - x\|_X$$

Proof. Let $x, y \in C$. Define the function

$$\sigma: \mathbb{R} \rightarrow X \text{ by } \sigma(t) = x + t\cdot(y - x)$$

then if $t \in [0,1]$ we have $\sigma(t) = x + t\cdot(y - x) = x\cdot(1 - t) + t\cdot y \in C$ [as C is convex] proving that

$$\sigma([0,1]) \subseteq C \quad (16.199)$$

Hence $\sigma([0,1]) \subseteq C \subseteq U$ or $[0,1] \subseteq \sigma^{-1}(U)$. As σ is continuous [see lemma: 16.154] it follows that $\sigma^{-1}(U)$ is open. Take $W = \sigma^{-1}(U)$ then we have

$$W \text{ is open, } [0,1] \subseteq W,\ \sigma(W) \subseteq U$$

Using [see lemma: 16.154] σ is of class C^1 so that by [theorem: 16.97]

$$\sigma_{|W}: W \rightarrow X$$

is of class C^1 and if we define

$$D^{[1]}\sigma: \mathbb{R} \rightarrow L_1(\mathbb{R}; Y) = L(\mathbb{R}, Y) \text{ by } D^{[1]}\sigma(x) = D^{[1]}_x\sigma$$

$$D^{[1]}\sigma_{|W}: \sigma^{-1}(U) \rightarrow L_1(\mathbb{R}; Y) = L(\mathbb{R}, Y) \text{ by } D^{[1]}\sigma_{|W}(x) = D^{[1]}_x\sigma_{|W}$$

then

$$D^{[1]}\sigma_{|W} = (D^{[1]}\sigma)_{|W} \tag{16.200}$$

As $\sigma_{|W}(W) = \sigma(W)$ it follows that the following function is well defined

$$g = f \circ \sigma_{|W}: W \rightarrow Y$$

and as f, $\sigma_{|W}$ are of class C^1 it follows that

$$g \text{ is of class } C^1 \tag{16.201}$$

Let $t \in W$ then as g is 1-times differentiable at t [hence Fréchet differentiable at t] it follows from [theorem: 16.32] that g has a derivative at t and

$$\begin{array}{rcl} g'_t & = & D_t g(1) \\ & = & D_t(f \circ \sigma_{|W})(1) \\ & \underset{[\text{theorem: } 16.36]}{=} & (D_{\alpha_{|W}(t)} f \circ D_t \sigma_{|W})(1) \\ & = & \left(D_{\sigma_{|W}(t)} f \circ D_t^{[1]} \sigma_{|W}\right)(1) \\ & \underset{[\text{eq: } 16.200] \wedge t \in W}{=} & \left(D_{\sigma(t)} f \circ D_t^{[1]} \sigma\right)(1) \\ & = & D_{\sigma(t)} f(D_t \sigma)(1) \\ & \underset{[\text{lemma: } 16.154]}{=} & D_{\sigma(t)} f(y - x) \end{array} \tag{16.202}$$

As g is of class C^1 it follows from [corollary: 16.151] that

$$h = g_{|[0,1]}: [0,1] \rightarrow Y$$

has a derivative on $[0,1]$, $h': [0,1] \rightarrow Y$ is continuous and $\forall t \in [0,1] \quad g'_t = h'_t$. So if $t \in [0,1]$ then we have

$$\begin{array}{rcl} \|h'_t\|_{|Y} & = & \|g'_t\|_Y \\ & \underset{[\text{eq: } 16.202]}{=} & \|D_{\sigma(t)} f(y - x)\| \\ & \leqslant & \|D_{\sigma(t)} f\|_{L(X,Y)} \cdot \|y - x\|_X \end{array}$$

By the hypothesis we have $\forall x \in C$ that $\|D_{\sigma(t)} f\|_{L(X,Y)} \leqslant k$, hence, as $\sigma(t) \in \sigma([0,1]) \subseteq C$ [see eq: 16.199] it follow from the above that

$$\|h'(t)\| = \|h'_t\|_Y \leqslant k \cdot \|y - x\|_X \tag{16.203}$$

Finally we have

$$\begin{array}{rcl} \|f(y) - f(x)\|_Y & = & \|f(x + 1 \cdot (x - y)) - f(x + 0 \cdot (x - y))\|_Y \\ & = & \|f(\sigma(1)) - f(\sigma(0))\|_Y \\ & 0,1 \in [0,1] \subseteq W & \|f(\sigma_{|W}(1)) - f(\sigma_{|W}(0))\|_Y \\ & = & \left\|\left(f \circ \sigma_{|W}\right)(1) - (f \circ \sigma_{|W}(0))\right\|_Y \\ & = & \|g(1) - g(0)\| \\ & \underset{0,1 \in [0,1]}{=} & \|h(1) - h(0)\| \end{array}$$

As h has a derivative on $[0,1]$ and $h': [0,1] \rightarrow X$ is continuous, we can apply the fundamental the0rem of calculus [theorem: 16.153] to get $h(1) - h(0) = \int_0^1 h'$. Combining this with the above gives

$$\begin{array}{rcl} \|f(y) - f(x)\|_Y & = & \left\|\int_0^1 h'\right\| \\ & \leqslant_{[\text{eq: } 16.203] \wedge [\text{theorem: } 15.29]} & (1 - 0) \cdot k \cdot \|y - x\| \\ & = & k \cdot \|y - x\| \end{array}$$

which completes the proof of the theorem. □

Definition 16.159. *Let X be a vector space, $a,b\in X$ then the* ***line segment connecting a to b*** *noted as $L_{a,b}$ is defined to be the set*

$$L_{a,b}=\{t\cdot a+(1-t)\cdot b|t\in[0,1]\}=\{b-t\cdot(a-b)|t\in[0,1]\}\subseteq X$$

***Note* 16.160.** *As $a=1\cdot a+(1-1)\cdot b$ and $b=0\cdot a+(1-0)\cdot b$ it follows that*

Example 16.161. $[0,1]=L_{0,1}$ in $\mathbb{R}$

Proof. If $x\in L_{0,1}$ then there exist a $t\in[0,1]$ that $x=t\cdot 0+(1-t)\cdot 1=1-t$. As $0\leqslant t$ it follows that $x=1-t\leqslant 1$, further as $t\leqslant 1\Rightarrow -1\leqslant -t\Rightarrow x=0=1-1\leqslant 1-t=x$. proving that $x\in[0,1]$. Hence

$$L_{0,1}\subseteq[0,1] \tag{16.204}$$

On the other hand if $x\in[0,1]$ then $0\leqslant x\leqslant 1\Rightarrow 0\leqslant 1-x\leqslant 1$, take $t=(1-x)$ then $t\in[0,1]$ and

$$t\cdot 0+(1-t)\cdot 1=(1-t)=(1-(1-x))=x$$

proving $x\in L_{0,1}$. Hence $[0,1]\subseteq L_{0,1}$ which combined with [eq: 16.204] gives

$$L_{0,1}=[0,1]$$

□

Lemma 16.162. *Let $\langle X,\|\|\rangle$ be a normed space and $a,b\in X$ then $L_{a,b}$ is convex, compact and bounded.*

Proof. Let $x,y\in L_{a,b}$, then there exists $t_x,t_y\in[0,1]$ such that $x=t_x\cdot a+(1-t_x)\cdot b$ and $y=t_y\cdot a+(1-t_y)\cdot b$. Let $t\in[0,1]$ then we have

$$\begin{aligned}(1-t)\cdot x+t\cdot y &= (1-t)\cdot(t_x\cdot a+(1-t_x)\cdot b)+t\cdot(t_y\cdot a+(1-t_y)\cdot b)\\ &= (1-t)\cdot t_x\cdot a+(1-t)\cdot(1-t_x)\cdot b+t\cdot t_y\cdot a+t\cdot(1-t_y)\cdot b\\ &= ((1-t)\cdot t_x+t\cdot t_y)\cdot a+((1-t)\cdot(1-t_x)+t\cdot(1-t_y))\cdot b\\ &= (t_x+t\cdot(t_y-t_x))\cdot a+(1-t_x-t\cdot 1+t\cdot t_x+t\cdot 1-t\cdot t_y)\cdot b\\ &= (t_x+t\cdot(t_y-t_x))\cdot a+(1-(t_x+t\cdot(t_y-t_x)))\cdot b\\ &= s\cdot a+(1-s)\cdot b \text{ where } s=t_x+t\cdot(t_y-t_x)\end{aligned}$$

Now as $0\leqslant t,t_x,t_y\leqslant 1$ we have that

$$t_x\cdot t\leqslant t_x\Rightarrow 0\leqslant t_x-t_x\cdot t\Rightarrow 0\leqslant t_x-t\cdot t_x+t\cdot t_y=t_x+t\cdot(t_y-t_x)\leqslant t_x+(t_y-t_x)=t_y\leqslant 1$$

proving that $s\in[0,1]$ and thus that $x+t\cdot(y-x)\in L_{a,b}$. Hence we have that

$$L_{a,b} \text{ is convex}$$

Define $\sigma:\mathbb{R}\rightarrow X$ by $\sigma(t)=t\cdot a+(1-t)\cdot b$. then we have

$$L_{a,b}=\sigma([0,1])$$

Given $\varepsilon\in\mathbb{R}^+$, $s\in\mathbb{R}$ take $\delta=\frac{\varepsilon}{\|a-b\|_Y+1}\in\mathbb{R}^+$ then if $|s-t|<\delta$ and $t\in\mathbb{R}$ we have

$$\begin{aligned}\|\sigma(t)-\sigma(s)\| &= \|t\cdot a+(1-t)\cdot b-s\cdot a-(1-s)\cdot b\|\\ &= \|(t-s)\cdot a+(s-t)\cdot b\|\\ &= |t-s|\cdot\|a-b\|\\ &< \delta\cdot\|a-b\|\\ &< \varepsilon\end{aligned}$$

proving that

$$\sigma \text{ is continuous}$$

As $[0,1]$ is compact [see theorem: 14.232] and σ is continuous it follows from [theorem: 14.229] that

$$L_{a,b}=\sigma([0,1]) \text{ is compact}$$

Applying then [theorem: 14.227] proves that

$$L_{a,b} \text{ is bounded}$$

□

Theorem 16.163. (Mean Value Theorem (III)) *Let $\langle X,\|\|_X\rangle$ be a normed, $\langle Y,\|\|_Y\rangle$ a Banach space, U a open set in X, $x,y\in X$ such that $L_{x,y}\in U$ and $f:U\rightarrow Y$ a function of class C^1 then*

$$\sup\left(\{\|D_\zeta f\|_{L(X,Y)}|\zeta\in L_{x,y}\}\right)$$

exists and

$$\|f(y)-f(x)\|_Y\leqslant\sup\left(\{\|D_\zeta f\|_{L(X,Y)}|\zeta\in L_{x,y}\}\right)\cdot\|y-x\|_X$$

Proof. As f is of class C^1 the function

$$Df:U\rightarrow L(X,Y)$$

is well defined and is continuous, further by the previous lemma [lemma: 16.162] $L_{x,y}$ is compact. So we can use [theorem: 14.229] to prove that $Df(L_{x,y})$ is compact, applying then [theorem: 14.227] it follows that $Df(L_{x,y})$ is bounded. So there exist a $M\in\mathbb{R}^+$ such that $\forall L_1,L_2\in Df(L_{x,y})$ we have that $\|L_1-L_2\|_{L(X,Y)}\leqslant M$. Let $\zeta\in L_{x,y}\subseteq U$ then we have

$$\begin{array}{rcl}\|Df(\zeta)\|_{L(X,Y)} & = & \|Df(\zeta)-Df(x)+Df(x)\|_{L(X,Y)}\\ & \leqslant & \|Df(\zeta)-Df(x)\|_{L(X,Y)}+\|Df(x)\|_{L(X,Y)}\\ & \leqslant & M+\|Df(x)\|_{L(X,Y)}\end{array}$$

proving that $\{\|D_\zeta f\|_{L(X,Y)}|\zeta\in L_{x,y}\}=\{\|Df(\zeta)\||\zeta\in L_{x,y}\}$ is bounded above by $M+\|Df(x)\|_{L(X,Y)}$. As $\mathbb{R}$ is conditionally complete [see theorem: 10.18] and $\{\|D_\zeta f\|_{L(X,Y)}|\zeta\in L_{x,y}\}\neq\varnothing$ [because $0\in L_{x,y}\Rightarrow L_{x,y}\neq\varnothing$] it follows that

$$\sup\left(\{\|D_\zeta f\|_{L(X,Y)}|\zeta\in L_{x,y}\}\right)\text{ exist}$$

As $L_{x,y}$ is convex [see lemma: 16.162], $L_{x,y}\subseteq U$ and $\forall\zeta\in L_{x,y}$

$$\|D_\zeta f\|_{L(X,Y)}\leqslant\sup\left(\{\|D_\zeta f\|_{L(X,Y)}|\zeta\in L_{x,y}\}\right)$$

it follows from [theorem: 16.158] that

$$\|f(y)-f(x)\|_{L(X,Y)}\leqslant\sup\left(\{\|D_\zeta f\|_{L(X,Y)}|\zeta\in L_{x,y}\}\right)\cdot\|y-x\|_X$$

□

Corollary 16.164. *Let $\langle X,\|\|_X\rangle$ be a normed space, $\langle Y,\|\|_Y\rangle$ a Banach space, U a open set in X, $x\in U$, $\delta\in\mathbb{R}^+$ such that $B_{\|\|_X}(x,\delta)\subseteq U$, $f:U\rightarrow Y$ a function of class C^1 then $\forall y,z\in B_{\|\|_X}(x,\delta)$ we have that*

$$\sup\left(\{\|D_{y+t\cdot(z-y)}f\|_{L(X,Y)}|t\in[0,1]\}\right)$$

exist and

$$\|f(z)-f(y)\|_Y\leqslant\sup\left(\{\|D_{y+t\cdot(z-y)}f\|_{L(X,Y)}|t\in[0,1]\}\right)\cdot\|z-y\|_X$$

Proof. Let $y,z\in B_{\|\|_X}(x,\delta)$ then as open balls are convex [see theorem: 16.157] we have that $L_{z,y}=\{y+t\cdot(z-y)|t\in[0,1]\}\subseteq B_{\|\|_X}(x,\delta)\subseteq U$. So by the previous theorem [theorem: 16.163]

$$\sup\left(\{\|D_\zeta f\|_{L(X,Y)}|\zeta\in L_{z,y}\}\right)$$

exists and

$$\|f(z)-f(y)\|_Y\leqslant\sup\left(\{\|D_\zeta f\|_{L(X,Y)}|\zeta\in L_{x,y}\}\right)\cdot\|z-y\|_X$$

Now

$$\begin{array}{rcl}r\in\{\|D_\zeta f\|_{L(X,Y)}|\zeta\in L_{z,y}\} & \Leftrightarrow & \exists\zeta\in L_{z,y}\text{ with }r=\|D_\zeta f\|_{L(X,Y)}\\ & \Leftrightarrow & \exists t\in[0,1]\text{ with }\zeta=y+t\cdot(z-y)\wedge r=\|D_\zeta f\|_{L(X,Y)}\\ & \Leftrightarrow & \exists t\in[0,1]\text{ with }r=\|D_{y+t\cdot(z-y)}\|_{L(X,Y)}\\ & \Leftrightarrow & r\in\{\|D_{y+t\cdot(z-y)}f\|_{L(X,Y)}|t\in[0,1]\}\end{array}$$

So $\{\|D_\zeta f\|_{L(X,Y)}|\zeta\in L_{z,y}\}=\{\|D_{y+t\cdot(z-y)}f\|_{L(X,Y)}|t\in[0,1]\}$ proving that

$$\{\|D_{y+t\cdot(z-y)}f\|_{L(X,Y)}|t\in[0,1]\}$$

exist and

$$\|f(z)-f(y)\|_Y \leqslant \sup(\{\|D_{y+t\cdot(z-y)}f\|_{L(X,Y)}|t\in[0,1]\})\cdot\|z-y\|$$

□

Theorem 16.165. (Mean Value Theorem (IV)) *Let $\langle X,\|\|_X\rangle$ be a normed space, $\langle Y,\|\|_Y\rangle$ a Banach space, U a open set in X, a, $b\in X$ such that $L_{a,b}\subseteq U$ and $f:U\rightarrow Y$ a function of class C^1 then $\forall x\in U$ we have*

$$\|f(a)-f(b)-D_xf(a-b)\|_Y \leqslant \sup(\{\|D_\zeta f-D_xf\|_{L(X,Y)}|\zeta\in L_{a,b}\})\cdot\|a-b\|_X$$

Proof. Let $x\in U$ and define

$$g:U\rightarrow Y \text{ by } g(\zeta)=f(\zeta)-D_xf(\zeta) \tag{16.205}$$

Then, as f is of class C^1 and $D_xf\in L(X,Y)$ is of class C^∞ [using example: 16.116], it follows from [theorem: 16.113] that g is of class C^1. Hence we have that g has a differential at $\zeta\in U$ with

$$D_\zeta g=D_\zeta f-D_\zeta(D_xf) \underset{\text{[example: 16.27]}}{=} D_\zeta g=D_\zeta f-D_xf \tag{16.206}$$

By the Mean Value Theorem (III) [theorem: 16.163] it follows that

$$\|g(a)-g(b)\|_Y \leqslant \sup(\{\|D_\zeta g\|_{|L(X,Y)}|\zeta\in L_{a,b}|\})\cdot\|a-b\|$$

Hence

$$\begin{aligned}
\|f(a)-f(b)-D_xf(a-b)\|_Y &= \|f(a)-f(b)-D_xf(a)-D_cf(b)\|_Y\\
&\underset{\text{[eq: 16.205]}}{} \|g(a)-g(b)\|_{|Y|}\\
&\leqslant \sup(\{\|D_\zeta g\|_{|L(X,Y)}|\zeta\in L_{a,b}|\})\cdot\|a-b\|_X\\
&\underset{\text{[eq: 16.206]}}{=} \sup(\{\|D_\varsigma f-D_xf\|_{|L(X,Y)}|\zeta\in L_{a,b}\})\cdot\|a-b\|_X
\end{aligned}$$

□

To prove the last Mean Value Theorem (from Dieudonné) we first need a little lemma.

Lemma 16.166. *Let $\langle X,\|\|\rangle$ be a normed space, $a,b\in\mathbb{R}$ with $a<b$, $x\in[a,b[$, $f:[a,b[\rightarrow X$ a function that has a derivative at x then given $\varepsilon\in\mathbb{R}^+$ there exist a $\lambda\in]x,b]$ such that $\forall\zeta\in[x,\lambda[$ we have*

$$\|f(\zeta)-f(x)-f'_x\cdot(\zeta-x)\|\leqslant\varepsilon\cdot(\zeta-x)$$

Proof. Let $\varepsilon\in\mathbb{R}^+$. As $x\in[a,b[$ we have either:

$\boldsymbol{x=a}$**.** Then by [definition: 16.146] f has a right derivative at $a=x$ and $f'_x=f'_{+,x}$.

$\boldsymbol{x\in]a,b[}$**.** Then by [definition: 16.146] $f_{|]a,b[}$ has a derivative at x and $f'_x=(f_{|]a,b[})'_x$. Using [theorem: 16.145] it follows that f has a right derivative at x and $f'_x\underset{\text{def}}{=}f'_{|]a,b[}=f'_{+,x}$

So in all cases we have that

$$f \text{ has a right derivative at } x \text{ and } f'_x=f'_{+,x} \tag{16.207}$$

So there exist a $\delta\in\mathbb{R}^+$ such that if $0<h<\delta$ and $x+h\in[a,b]$ we have

$$\left\|\frac{f(x+h)-f(x)}{h}-f'_{+,x}\right\|<\varepsilon \tag{16.208}$$

Take $\lambda=\min(x+\delta,b)$ so that, as $x<x+\delta\wedge x<b$, $\lambda\in]x,b]$ then for $\zeta\in[x,\lambda[\Rightarrow\zeta\leqslant\zeta<\lambda$ we have either:

$\boldsymbol{\zeta=x}$**.** Then

$$\begin{aligned}
\|f(\zeta)-f(x)-f'_x\cdot(\zeta-x)\| &\underset{\zeta=x}{=} \|f(\zeta)-f(\zeta)-f'_{+,x}\cdot(\zeta-\zeta)\|\\
&= \|0\|\\
&= 0\\
&\leqslant \varepsilon\cdot(\zeta-x)
\end{aligned}$$

$\boldsymbol{x<\zeta}$**.** Then $0<\zeta-x<\lambda-x=\delta$ and $x+(\zeta-x)=\zeta\in[a,b]$ [because $a\leqslant x<\zeta\wedge\zeta<\lambda\leqslant b$] so that we can use [eq: 16.208] to get

$$\left\|\frac{f(\zeta)-f(x)-f'_{+,x}}{\zeta-x}\right\|=\left\|\frac{f(x+(\zeta-x))-f(x)}{\zeta-x}-f'_{+,x}\right\|<\varepsilon \tag{16.209}$$

hence

$$\begin{array}{rcl} \|f(\zeta)-f(x)-f'_x\cdot(\zeta-x)\| & \underset{[\text{eq: } 16.207]}{=} & \|f(\zeta)-f(x)-f'_{+,x}\cdot(\zeta-x)\| \\ & = & \left\|(\zeta-x)\cdot\frac{f(\zeta)-f(x)-f'_{+,x}}{\zeta-x}\right\| \\ & = & |\zeta-x|\cdot\left\|\frac{f(\zeta)-f(x)-f'_{+,x}}{\zeta-x}\right\| \\ & \underset{x<\zeta}{=} & (\zeta-x)\cdot\left\|\frac{f(\zeta)-f(x)-f'_{+,x}}{\zeta-x}\right\| \\ & <_{[\text{eq: } 16.209]} & \varepsilon\cdot(\zeta-x) \end{array}$$

Hence in both cases we have that

$$\|f(\zeta)-f(x)-f'_x\cdot(\zeta-x)\|\leqslant\zeta\cdot(\zeta-x)$$

□

Theorem 16.167. (Main Value Theorem (V)) *Let $a,b\in\mathbb{R}$ with $a<b$, $\langle X,\|\|\rangle$ a normed space, $f:[a,b]\rightarrow X$, $\varphi:[a,b]\rightarrow\mathbb{R}$ continuous functions such that there exist a denumerable set $E\subseteq[a,b]$ such that $\forall x\in[a,b]\backslash E$ f and φ has a derivative f'_x and φ'_x at x with $\|f'_x\|\leqslant\varphi'_x$ then $\|f(b)-f(a)\|\leqslant\varphi(b)-\varphi(a)$.*

Proof. As E is denumerable there exist a bijection $\rho:\mathbb{N}\rightarrow E$. Define for $\beta\in[a,b]$ $\xi(\beta)\in[0,\infty[$ as follows

$$\xi(\beta)=\sum_{n\in\{i\in\mathbb{N}|\rho(i)<\beta\}}\frac{1}{2^n}\underset{[\text{definition: } 11.32]}{=}\begin{cases} 0 \text{ if } \{i\in\mathbb{N}|\rho(i)<\beta\}=\varnothing \\ \sum_{n\in\{i\in\mathbb{N}|\rho(i)<\beta\}}\frac{1}{2^n} \text{ if } \{i\in\mathbb{N}|\rho(i)<\beta\}=\varnothing \end{cases}$$

Note that $\{i\in\mathbb{N}|\rho(i)<\beta\}$ is finite because by [theorem: 10.30] there exist a $n\in\mathbb{N}$ such that $\beta\leqslant n$, hence $\{i\in\mathbb{N}|\rho(i)<\beta\}\subseteq\{1,\ldots,n\}$, so the sum $\sum_{n\in\{i\in\mathbb{N}|\rho(i)<\beta\}}\frac{1}{2^n}$ is well defined. Further as $\rho(\mathbb{N})=E\subseteq[a,b]$ we have $\forall i\in\mathbb{N}$ that $a\leqslant\rho(i)$ so that $\rho(i)\nless a$, hence $\{i\in\mathbb{N}|\rho(i)<a\}=\varnothing$. So we have

$$\xi(a)=0 \tag{16.210}$$

Define for $\varepsilon\in\mathbb{R}^+$ A_ε by

$$A_\varepsilon=\{\beta\in[a,b]|\forall\gamma\in[a,\beta[\vDash\|f(\gamma)-f(a)\|\leqslant\varphi(\gamma)-\varphi(a)+\varepsilon\cdot(\gamma-a)+\varepsilon\cdot\xi(\gamma)\}\subseteq[a,b]$$

Then for $\gamma=a$ we have $\|f(a)-f(a)\|=0=\varphi(a)-\varphi(a)$, $\varepsilon\cdot(a-a)=0$ and $\xi(a)\underset{[\text{eq: } 16.210]}{=}0$ we have

$$\|f(\gamma)-f(a)\|=0=\varphi(\gamma)-\varphi(a)+\varepsilon\cdot(\gamma-a)+\varepsilon\cdot\xi(\gamma)$$

proving that

$$a\in A_\varepsilon \text{ so that } A_\varepsilon\neq\varnothing \tag{16.211}$$

If $\beta\in A_\varepsilon$ then if $\xi\in[a,\beta]$ then if $\lambda\in[a,\xi[$ we have that $\gamma\in[a,\beta[$ so that

$$\|f(\gamma)-f(a)\|\leqslant\varphi(\gamma)-\varphi(a)+\varepsilon\cdot(\gamma-a)+\varepsilon\cdot\xi(\gamma)$$

which proves that $\zeta\in A_\varepsilon$. Hence

$$\forall\beta\in A_\varepsilon \text{ we have } [a,\beta]\subseteq A_\varepsilon \tag{16.212}$$

As $\varnothing\neq A_\varepsilon$ and A_ε is bounded above by b [because $A_\varepsilon\subseteq[a,b]$] it follows from the conditional completeness of the real numbers [see theorem: 10.18] that $\sup(A_\varepsilon)$ exist. As $A_\varepsilon\subseteq[a,b]$ we have $\forall\zeta\in A_\varepsilon$ that $a\leqslant\zeta\leqslant b$ so that $a\leqslant\sup(A_\varepsilon)\leqslant b$. So if we take $\sigma=\sup(A_\varepsilon)$ then we have

$$\sigma=\sup(A_\varepsilon) \text{ and } \sigma\in[a,b] \tag{16.213}$$

If $\zeta\in[a,\sigma[$ then $\zeta<\sigma=\sup(A_\varepsilon)$ so that there exist a $\beta\in A_\varepsilon$ such that $a\leqslant\zeta<\beta\leqslant\sup(A_\varepsilon)=\sigma$ so that $\zeta\in[a,\beta[\subseteq[a,\beta]$, using [eq: 16.212] $[a,\beta]\subseteq A_\varepsilon$ from which it follows that $\zeta\in A_\varepsilon$. Hence

$$[a,\sigma[\subseteq A_\varepsilon$$

Further as $\beta \in A_\varepsilon$ and $\zeta \in [a, \beta[$ it follows that $\|f(\zeta) - f(a)\| \leqslant \varphi(\zeta) - \varphi(a) + \varepsilon \cdot (\zeta - a) + \varepsilon \cdot \xi(\gamma)$ so, as $\zeta \in [a, \sigma[$ was chosen arbitrary, it follows that $\sigma \in A_\varepsilon$. Hence we have

$$[a, \sigma] \subseteq A_\varepsilon \tag{16.214}$$

Assume that $x \in A_\varepsilon \setminus [a, \sigma]$ then, as $A_\varepsilon \subseteq [a, b]$ so that $a \leqslant x$, we must have $\sigma < x$ which contradicts the fact that $x \leqslant \sup(A_\varepsilon) = \sigma$. Hence $A_\varepsilon \setminus [a, \sigma] = \varnothing$, so if $x \in A_\varepsilon$ then we must have $x \in [a, \sigma]$ proving that $A_\varepsilon \subseteq [a, \sigma]$ which combined with [eq: 16.214] results in

$$[a, \sigma] = A_\varepsilon \tag{16.215}$$

Take $\alpha, \beta \in [a, b]$ with $\alpha < \beta$. If $n \in \{i \in \mathbb{N} | \rho(i) < \alpha\}$ then we have $\rho(n) < \alpha < \beta$ so that $n \in \{i \in \mathbb{N} | \rho(i) < \beta\}$ proving that $\{i \in \mathbb{N} | \rho(i) < \alpha\} \subseteq \{i \in \mathbb{N} | \rho(i) < \beta\}$. We consider now the following cases for α:

$\{i \in \mathbb{N} | \rho(i) < \alpha\} = \varnothing$. Then $\xi(\alpha) = 0 \leqslant \xi(\beta)$

$\{i \in \mathbb{N} | \rho(i) < \alpha\} \neq \varnothing$. Then as $\{i \in \mathbb{N} | \rho(i) < \alpha\} \subseteq \{i \in \mathbb{N} | \rho(i) < \beta\}$ so that $\{i \in \mathbb{N} | \rho(i) < \beta\} \neq \varnothing$ and

$$\begin{aligned}
\xi(\beta) &= \sum_{n \in \{i \in \mathbb{N} | \rho(i) < \beta\}} \frac{1}{2^n} \\
&= \sum_{n \in \{i \in \mathbb{N} | \rho(i) < \alpha\}} \frac{1}{2^n} + \sum_{n \in \{i \in \mathbb{N} | \rho(i) < \beta\} \setminus \{i \in \mathbb{N} | \rho(i) < \alpha\}} \frac{1}{2^n} \\
&\geqslant \sum_{n \in \{i \in \mathbb{N} | \rho(i) < \alpha\}} \frac{1}{2^n} \\
&= \xi(\alpha)
\end{aligned}$$

proving in all cases that $\xi(\alpha) \leqslant \xi(\beta)$. Hence

$$\forall \alpha, \beta \in [a, b] \text{ with } \alpha < \beta \text{ we have } 0 \leqslant \xi(\alpha) \leqslant \xi(\beta) \tag{16.216}$$

As $\sigma \in A_\varepsilon \subseteq [a, b]$ we have that $a \leqslant \sigma$ and we have then two cases to consider:

$a = \sigma$. Then

$$\begin{aligned}
\|f(\sigma) - f(a)\| &= \|f(\sigma) - f(\sigma)\| \\
&= 0 \\
&\leqslant \varepsilon \cdot \xi(\sigma) \\
&\leqslant \varphi(\sigma) - \varphi(\sigma) + \varepsilon \cdot (\sigma - \sigma) + \varepsilon \cdot \xi(\sigma) \\
&= \|f(\sigma) - f(a)\| \leqslant \varphi(\sigma) - \varphi(a) + \varepsilon \cdot (\sigma - a) + \varepsilon \cdot \xi(\sigma)
\end{aligned}$$

$a < \sigma$. Choose $\zeta \in \mathbb{R}^+$. By the continuity of φ at σ there exist a $\delta_\varphi \in \mathbb{R}^+$ such that

$$\forall \alpha \in]\sigma - \delta_\varphi, a + \delta_\varphi[\bigcap [a, b] \text{ we have } |\varphi(\alpha) - \varphi(\sigma)| < \frac{\zeta}{2}$$

further using the continuity of f at σ there exist a $\delta_f \in \mathbb{R}^+$ such that

$$\forall \alpha \in]\sigma - \delta_f, \sigma + \delta_f[\bigcap [a, b] \text{ we have } \|f(\alpha) - f(\sigma)\| < \frac{\zeta}{2}$$

Take $\delta = \min(\delta_\varphi, \delta_f, \sigma - a) \in \mathbb{R}^+$ [because $a < \sigma$] then as $\sigma - \delta < \sigma$ there exist a $\beta \in \mathbb{R}$ such that $\sigma - \delta < \beta < \sigma < \sigma + \delta$. As $\delta \leqslant \sigma - a$ we have that $a \leqslant \sigma - \delta < \beta$, further, as $\sigma \in A_\varepsilon \subseteq [a, b]$, it follows that $\beta < \sigma \leqslant b$. Hence we conclude that $a \leqslant \beta \leqslant b$ and $\sigma - \delta < \beta < \sigma + \delta$, which, as $\delta_f, \delta_\varphi \leqslant \delta$, proves that $\beta \in]\sigma - \delta_\varphi, \sigma + \delta_\varphi[\bigcap]\delta - \delta_f, \sigma + \delta_f[\bigcap [a, b]$. Hence we have that

$$|\varphi(\beta) - \varphi(\sigma)| < \frac{\zeta}{2} \text{ and } \|f(\beta) - f(\sigma)\| < \frac{\zeta}{2} \tag{16.217}$$

Next

$$\begin{aligned}
\varphi(\beta) - \varphi(a) &= \varphi(\beta) - \varphi(a) = \varphi(\sigma) - \varphi(a) + \varphi(\beta) - \varphi(\sigma) \\
&\leqslant \varphi(\sigma) - \varphi(a) + |\varphi(\beta) - \varphi(\sigma)| \\
&<_{\text{[eq: 16.217}} \varphi(\sigma) - \varphi(a) + \frac{\zeta}{2}
\end{aligned} \tag{16.218}$$

Further

$$\begin{array}{rcl}\|f(\sigma)-f(a)\| & \leqslant & \|f(\beta)-f(a)+f(\sigma)-f(\beta)\| \\ & \leqslant & \|f(\beta)-f(a)\|+\|f(\sigma)-f(\beta)\| \\ & <_{[\text{eq: }16.217]} & \|f(\beta)-f(a)\|+\dfrac{\zeta}{2} \\ & \leqslant_{\sigma\in A_{\varepsilon}\wedge a\leqslant b<\sigma} & \varphi(\beta)-\varphi(a)+\varepsilon\cdot(\beta-a)+\varepsilon\cdot\xi(\beta)+\dfrac{\zeta}{2} \\ & <_{[\text{eq: }16.218]} & \varphi(\sigma)-\varphi(a)+\dfrac{\zeta}{2}+\varepsilon\cdot(\beta-a)+\varepsilon\cdot\xi(\beta)+\dfrac{\zeta}{2} \\ & = & \varphi(\sigma)-\varphi(a)+\varepsilon\cdot(\beta-a)+\varepsilon\cdot\xi(\beta)+\zeta \\ & \leqslant_{[\text{eq: }16.216]\wedge\beta<\sigma} & \varphi(\sigma)-\varphi(a)+\varepsilon\cdot(\sigma-a)+\varepsilon\cdot\xi(\sigma)+\zeta\end{array}$$

As ζ was chosen arbitrary it follows from [theorem: 10.31] that

$$\|f(\sigma)-f(a)\|\leqslant\varphi(\sigma)-\varphi(a)+\varepsilon\cdot(\sigma-a)+\varepsilon\cdot\xi(\sigma)$$

So in all possible cases for σ we have

$$\|f(\sigma)-f(a)\|\leqslant\varphi(\sigma)-\varphi(a)+\varepsilon\cdot(\sigma-a)+\varepsilon\cdot\xi(\sigma) \tag{16.219}$$

Suppose now that $\sigma<b$ then we have for σ the following possibilities in relation to E:

$\boldsymbol{\sigma\notin E}$. Then φ,f have a derivative at σ as $a\leqslant\sigma<b$ it follows from [lemma: 16.166] then there exists $\lambda_{\varphi},\lambda_f\in]\sigma,b]$ such that $\forall\zeta\in[\sigma,\lambda_{\varphi}[$ $|\varphi(\zeta)-\varphi(\sigma)-\varphi'_{\sigma}\cdot(\zeta-\sigma)|\leqslant\frac{\varepsilon}{2}\cdot(\zeta-\sigma)$ and for $\zeta\in[\sigma,\lambda_f]$ that $\|f(\sigma)-f(\sigma)-f'_{\sigma}\cdot(\zeta-\sigma)\|\leqslant\frac{\varepsilon}{2}\cdot(\zeta-\sigma)$. So if $\zeta\in[\sigma,\min(\lambda_{\varphi},\lambda_f)[$ we have

$$\begin{array}{rcl}\varphi'_{\sigma}\cdot(\zeta-\sigma) & = & \varphi'_{\sigma}\cdot(\zeta-\sigma)-(\varphi(\zeta)-\varphi(\sigma))+\varphi(\zeta)-\varphi(\sigma) \\ & \leqslant & |\varphi'_{\sigma}\cdot(\zeta-\sigma)-(\varphi(\zeta)-\varphi(\sigma))|+\varphi(\zeta)-\varphi(\sigma) \\ & \leqslant & \dfrac{\varepsilon}{2}\cdot(\zeta-\sigma)+\varphi(\zeta)-\varphi(\sigma)\end{array} \tag{16.220}$$

and

$$\begin{array}{rcl}\|f(\zeta)-f(\sigma)\| & = & \|f(\zeta)-f(\sigma)-f'_{\sigma}\cdot(\zeta-\sigma)+f'_{\sigma}\cdot(\zeta-\sigma)\| \\ & \leqslant & \|f(\zeta)-f(\sigma)-f'_{\sigma}\cdot(\zeta-\sigma)\|+\|f'_{\sigma}\cdot(\zeta-\sigma)\| \\ & \leqslant & \dfrac{\varepsilon}{2}\cdot(\zeta-\sigma)+\|f'_{\sigma}\cdot(\zeta-\sigma)\| \\ & = & \dfrac{\varepsilon}{2}\cdot(\zeta-\sigma)+\|f'_{\sigma}\|\cdot|\zeta-\sigma| \\ & \underset{\zeta\in[\sigma,\min(\lambda_{\varphi},\lambda_f)[\Rightarrow\sigma\leqslant\zeta}{=} & \dfrac{\varepsilon}{2}\cdot(\zeta-\sigma)+\|f'_{\sigma}\|\cdot(\zeta-\sigma) \\ & \leqslant_{\text{hypothesis}} & \dfrac{\varepsilon}{2}\cdot(\zeta-\sigma)+\varphi'_{\sigma}\cdot(\zeta-\sigma) \\ & \leqslant_{[\text{eq:}16.220]} & \dfrac{\varepsilon}{2}\cdot(\zeta-\sigma)+\dfrac{\varepsilon}{2}\cdot(\zeta-\sigma)+\varphi(\zeta)-\varphi(\sigma) \\ & = & \varphi(\zeta)-\varphi(\sigma)+\varepsilon\cdot(\zeta-\sigma)\end{array}$$

proving

$$\|f(\zeta)-f(\sigma)\|\leqslant\varphi(\zeta)-\varphi(\sigma)+\varepsilon\cdot(\zeta-\sigma) \tag{16.221}$$

Next

$$\begin{array}{rcl}\|f(\zeta)-f(a)\| & = & \|f(\zeta)-f(\sigma)+f(\sigma)-f(a)\| \\ & \leqslant & \|f(\zeta)-f(\sigma)\|+\|f(\sigma)-f(a)\| \\ & \leqslant_{[\text{eq: }16.221]} & \varphi(\zeta)-\varphi(\sigma)+\varepsilon\cdot(\zeta-\sigma)+\|f(\sigma)-f(a)\| \\ & \leqslant_{[\text{eq: }16.219]} & \varphi(\zeta)-\underbrace{\varphi(\sigma)}_{1}+\underbrace{\varepsilon\cdot(\zeta-\sigma)}_{2}+\underbrace{\varphi(\sigma)}_{1}-\varphi(a)+\underbrace{\varepsilon\cdot(\sigma-a)}_{2}+\varepsilon\cdot\xi(\sigma) \\ & = & \varphi(\zeta)-\varphi(a)+\varepsilon\cdot(\zeta-a)+\varepsilon\cdot\xi(\sigma) \\ & \leqslant_{\sigma<\zeta\wedge[\text{eq: }16.216]} & \varphi(\zeta)-\varphi(a)+\varepsilon\cdot(\zeta-a)+\varepsilon\cdot\xi(\zeta)\end{array}$$

So we have

$$\forall \zeta \in [\sigma, \min(\lambda_\varphi, \lambda_f)[\text{ we have } \|f(\zeta) - f(a)\| \leqslant \varphi(\zeta) - \varphi(a) + \varepsilon \cdot (\zeta - a) + \varepsilon \cdot \xi(\zeta)$$

As $\sigma \in A_\varepsilon$ we have also

$$\forall \zeta \in [a, \sigma[\quad \|f(\zeta) - f(a)\| \leqslant \varphi(\zeta) - \varphi(a) + \varepsilon \cdot (\zeta - a) + \varepsilon \cdot \xi(\zeta)$$

so that

$$\forall \zeta \in [a, \min(\lambda_\varphi, \lambda_f)[\; \|f(\zeta) - f(a)\| \leqslant \varphi(\zeta) - \varphi(a) + \varepsilon \cdot (\zeta - a) + \varepsilon \cdot \xi(\zeta)$$

proving that $\min(\lambda_\varphi, \lambda_f) \in A_\varepsilon$ so that $\min(\lambda_\varphi, \lambda_f) \leqslant \sup(A_\varepsilon) = \sigma$. As $\lambda_\varphi, \lambda_f \in]\sigma, b]$ we have $\sigma < \min(\lambda_\varphi, \lambda_f)$ leading to the contradiction $\sigma < \sigma$. Hence this case leads to a contradiction.

$\sigma \in E$. Then there exist a $m \in \mathbb{N}$ such that $\sigma = \rho(m)$. As f, φ are continuous at σ there exists $\gamma_f, \gamma_\varphi \in \mathbb{R}^+$ such that

$$\forall \alpha \in]\sigma - \gamma_\varphi, \sigma + \gamma_\varphi[\bigcap [a, b] \text{ we have } |\varphi(\alpha) - \varphi(\sigma)| < \frac{\varepsilon}{2} \cdot \frac{1}{2^m}$$

and

$$\forall \alpha \in]\sigma - \gamma_f, \sigma + \gamma_f[\bigcap [a, b] \text{ we have } \|f(\alpha) - f(\sigma)\| < \frac{\varepsilon}{2} \cdot \frac{1}{2^m}$$

As we assumed that $\sigma < b$ it follows that $\mu = \min(b - \sigma, \gamma_\varphi, \gamma_f) \in \mathbb{R}^+$ and if $\zeta \in [\sigma, \sigma + \mu[$ then $a \leqslant \sigma M = \zeta < \sigma + \mu \leqslant \sigma + (b - \sigma)$, $\sigma + \gamma_\varphi, \sigma + \gamma_f$ so that $\zeta \in]\sigma - \gamma_\varphi, \sigma + \gamma_\varphi[\bigcap]\sigma - \gamma_f, \sigma + \gamma_f[\bigcap [a, b]$ hence we have

$$\forall \zeta \in [\sigma, \sigma + \mu[\text{ we have } |\varphi(\zeta) - \varphi(\sigma)| < \frac{\varepsilon}{2} \cdot \frac{1}{2^m} \wedge \|f(\zeta) - f(\sigma)\| \leqslant \frac{\varepsilon}{2} \cdot \frac{1}{2^m} \tag{16.222}$$

Next

$$\begin{aligned}
\|f(\zeta) - f(a)\| &= \|f(\zeta) - f(\sigma) + f(\sigma) - f(a)\| \\
&\leqslant \|f(\zeta) - f(\sigma)\| + \|f(\sigma) - f(a)\| \\
&\leqslant_{[\text{eq: } 16.222} \frac{\varepsilon}{2} \cdot \frac{1}{2^m} + \|f(\sigma) - f(a)\| \\
&\leqslant_{[\text{eq: } 16.219]} \frac{\varepsilon}{2} \cdot \frac{1}{2^m} + \varphi(\sigma) - \varphi(a) + \varepsilon \cdot (\sigma - a) + \varepsilon \cdot \xi(\sigma) \\
&= \frac{\varepsilon}{2} \cdot \frac{1}{2^m} + \varphi(\zeta) - \varphi(\zeta) + \varphi(\sigma) - \varphi(a) + \varepsilon \cdot (\sigma - a) + \varepsilon \cdot \xi(\sigma) \\
&\leqslant \frac{\varepsilon}{2} \cdot \frac{1}{2^m} + |\varphi(\sigma) - \varphi(\zeta)| + \varphi(\zeta) - \varphi(a) + \varepsilon \cdot (\sigma - a) + \varepsilon \cdot \xi(\sigma) \\
&\leqslant_{[\text{eq: } 16.222} \frac{\varepsilon}{2} \cdot \frac{1}{2^m} + \frac{\varepsilon}{2} \cdot \frac{1}{2^m} + \varphi(\zeta) - \varphi(a) + \varepsilon \cdot (\sigma - a) + \varepsilon \cdot \xi(\sigma) \\
&= \varphi(\zeta) - \varphi(a) + \varepsilon \cdot \frac{1}{2^m} + \varepsilon \cdot (\sigma - a) + \varepsilon \cdot \xi(\sigma) \\
&= \varphi(\zeta) - \varphi(a) + \varepsilon \cdot \frac{1}{2^m} + \varepsilon \cdot (\sigma - a) + \varepsilon \cdot \sum_{n \in \{i \in \mathbb{N} | \rho(i) < \sigma\}} \frac{1}{2^m} \\
&\underset{\sigma = \rho(m)}{=} \varphi(\zeta) - \varphi(a) + \varepsilon \cdot (\sigma - a) + \varepsilon \cdot \sum_{n \in \{i \in \mathbb{N} | \rho(i) \leqslant \sigma\}} \frac{1}{2^m} \\
&\leqslant_{\sigma < \zeta} \varphi(\zeta) - \varphi(a) + \varepsilon \cdot (\sigma - a) + \varepsilon \cdot \sum_{n \in \{i \in \mathbb{N} | \rho(i) < \zeta\}} \frac{1}{2^m} \\
&= \varphi(\zeta) - \varphi(a) + \varepsilon \cdot (\sigma - a) + \varepsilon \cdot \xi(\zeta)
\end{aligned}$$

proving that

$$\forall \zeta \in [\sigma, \sigma + \mu[\; \|f(\zeta) - f(a)\| \leqslant \varphi(\zeta) - \varphi(a) + \varepsilon \cdot (\sigma - a) + \varepsilon \cdot \xi(\zeta)$$

As $\sigma \in A_\varepsilon$ we have also

$$\forall \zeta \in [a, \sigma[\quad \|f(\zeta) - f(a)\| \leqslant \varphi(\zeta) - \varphi(a) + \varepsilon \cdot (\zeta - a) + \varepsilon \cdot \xi(\zeta)$$

so that

$$\forall \zeta \in [\sigma, \sigma + \mu[\quad \|f(\zeta) - f(a)\| \leqslant \varphi(\zeta) - \varphi(a) + \varepsilon \cdot (\zeta - a) + \varepsilon \cdot \xi(\zeta)$$

proving that $\sigma + \mu \in A_\varepsilon$. Hence $\sigma < \sigma + \mu \leqslant \sup(A_\varepsilon) \leqslant \sigma$ giving the contradiction $\sigma < \sigma$.

As in all cases we have a contradiction the assumption $\sigma < b$ must be false, hence we must have $\sigma = b$, then by [eq: 16.219] we have that

$$\|f(b) - f(a)\| \leqslant \varphi(b) - \varphi(a) + \varepsilon \cdot (b - a) + \varepsilon \cdot \xi(b)$$

As $\varepsilon \in \mathbb{R}^+$ was chosen arbitrary it follows from [theorem: 10.31] that

$$\|f(b) - f(a)\| \leqslant \varphi(b) - \varphi(a)$$

□

Corollary 16.168. (Main Value Theorem (VI)) *Let $a, b \in \mathbb{R}$ with $a < b$, $\langle X, \|\|\rangle$ a normed space, $f: [a, b] \rightarrow X$, $\varphi: [a, b] \rightarrow \mathbb{R}$ continuous functions such that $\forall x \in [a, b]$ f and φ has a derivative f'_x and φ'_x at x with $\|f'_x\| \leqslant \varphi'_x$ then $\|f(b) - f(a)\| \leqslant \varphi(b) - \varphi(a)$.*

Proof. Let $E = \left\{a + (b - a) \cdot \frac{1}{n} | n \in \mathbb{N}_0\right\}$ then as $\forall n \in \mathbb{N}$ we have that $0 < \frac{1}{n} < 1 \underset{0 < b - a}{\Rightarrow} 0 < (b - a) \cdot \frac{1}{n} < (b - a) \Rightarrow a < a + (b - a) \cdot \frac{1}{n} < b$ it follows that $E \subseteq [a, b]$. Define $\rho: \mathbb{N} \rightarrow E$ by $\rho = a + (b - a) \cdot \frac{1}{n}$ then clearly ρ is surjective. Further if $\rho(n) = \rho(n')$ then $a + (b - a) \cdot \frac{1}{n} = a + (b - a) \cdot \frac{1}{n\prime} \underset{a < b}{\Rightarrow} \frac{1}{n} = \frac{1}{n'}$ so that $n = n'$, which proves that ρ is injective. Hence we have a bijection ρ between $\mathbb{N}$ and E so that E is denumerable. As $\forall x \in [a, b]$ f'_x, φ'_x exists with $\|f'_x\| \leqslant \varphi'_x$ the same applies for $[a, b] \setminus E$, so applying [theorem: 16.167] proves that

$$\|f(b) - f(a)\| \leqslant \varphi(b) - \varphi(a)$$

□

Corollary 16.169. (Main value Theorem (VII)) *Let $a, b \in \mathbb{R}$ with $a < b$, $\langle Y, \|\|\rangle$ a normed space, $f: [a, b] \rightarrow Y$ a continuous function such that there exist a denumerable set $E \subseteq [a, b]$ so that $\forall x \in [a, b] \setminus E$ f'_x exists and $\|f'_x\| \leqslant M$ then $\|f(b) - f(a)\| \leqslant M \cdot (b - a)$*

Proof. Define $\varphi: \mathbb{R} \rightarrow \mathbb{R}$ by $\varphi(x) = M \cdot (x - a)$ then by [examples: 16.148,16.148] and [theorems: 16.150, 16.149] it follows that $\forall t \in [a, b] \setminus E$ φ has a derivative at t with $\varphi'_t = x$. So using [theorem: 16.167] it follows that:

$$\|f(b) - f(a)\| \leqslant \varphi(b) - \varphi(a) = M \cdot (b - a) - M \cdot (a - a) = M \cdot (b - a)$$

□

Corollary 16.170. (Main Value Theorem (VIII)) *Let $a, b \in \mathbb{R}$ with $a < b$, $\langle Y, \|\|\rangle$ a normed space, $f: [a, b] \rightarrow Y$ a continuous function such that $\forall x \in [a, b]$ $f'(x)$ exists and $\|f'_x\| \leqslant M$ then $\|f(b) - f(a)\| \leqslant M \cdot (b - a)$.*

Proof. Let $E = \left\{a + (b - a) \cdot \frac{1}{n} | n \in \mathbb{N}_0\right\}$ then as $\forall n \in \mathbb{N}$ we have that $0 < \frac{1}{n} < 1 \underset{0 < b - a}{\Rightarrow} 0 < (b - a) \cdot \frac{1}{n} < (b - a) \Rightarrow a < a + (b - a) \cdot \frac{1}{n} < b$ it follows that $E \subseteq [a, b]$. Define $\rho: \mathbb{N} \rightarrow E$ by $\rho = a + (b - a) \cdot \frac{1}{n}$ then clearly ρ is surjective. Further if $\rho(n) = \rho(n')$ then $a + (b - a) \cdot \frac{1}{n} = a + (b - a) \cdot \frac{1}{n\prime} \underset{a < b}{\Rightarrow} \frac{1}{n} = \frac{1}{n'}$ so that $n = n'$, which proves that ρ is injective. Hence we have a bijection ρ between $\mathbb{N}$ and E so that E is denumerable. As $\forall x \in [a, b]$ we have that f'_x exist and $\|f'_x\| \leqslant M$ the same applies to $[a, b] \setminus Y$. So by the previous corollary [corollary: 16.169] it follows that $\|f(b) - f(a)\| \leqslant M \cdot (b - a)$. □

16.4 Symmetry of Higher Order Differentials

We set out now to proof that $D_x^{[n]} f(h_1 : \ldots : h_n) = D_x^{[n]} f(v_{\sigma(1)} : \ldots : v_{\sigma(n)})$ where $\sigma \in P_n$ or, using the multilinear representation of the higher order differential, that $D_x^n f(v_1, \ldots, v_n) = D_x^n f(v_{\sigma(1)}, \ldots, v_{\sigma(n)})$. In other words $D_x^n f$ is a symmetric multilinear function. The proof is rather elaborated, so to simplify it we divide the proof in different lemma'a. First we prove symmetry for the case $n = 2$.

Lemma 16.171. *Let $\langle X, \|\|_X\rangle$ be a normed space, $\langle Y, \|\|_Y\rangle$ a Banach space, U a open set in X and $f: U \rightarrow Y$ a function of class C^2 then $\forall x \in X$ we have $\forall u, v \in X$ that*

$$\left(D_x^{[2]} f\right)(v)(w) = \left(D_x^{[2]} f\right)(w)(v)$$

or for the function defined by

$$D^{[2]} f: U \rightarrow L_2(X; Y) \text{ where } D^{[2]} f(x) = D_x^{[2]} f$$

we have

$$(D^{[2]} f(x))(v)(w) = (D^{[2]} f(x))(w)(v)$$

Proof. Let $x \in U$ then as U is open there exist a $r \in \mathbb{R}^+$ such that

$$x \in B_{\|\|_X}(x, r) \subseteq U$$

Let $v,w\in X$ such that $\|v\|_X<\frac{r}{4}$, $\|w\|_X<\frac{r}{4}$ and define

$$\varphi_{v,w}:]-2,2[\rightarrow X \text{ by } \varphi_{v,w}(t)=x+t\cdot v \text{ and } \psi_{v,w}:]-2,2[\rightarrow X \text{ by } \psi_{v,w}(t)=x+w+t\cdot v \tag{16.223}$$

If $t\in]-2,2[$ then we have $\|\varphi_{v,w}(t)-x\|_X=\|t\cdot v\|_X=|t|\cdot\|v\|_x\leqslant 2\cdot\frac{r}{4}<r$ and $\|\psi_{v,w}(t)-x\|_X=\|w+t\cdot v\|_X\leqslant\|w\|_X+|t|\cdot\|v\|_X\leqslant\frac{r}{4}+2\cdot\frac{r}{4}<r$ proving that

$$\varphi_{v,w}(]-2,2[)\subseteq U\wedge\psi_{v,w}(]-2,2[)\subseteq U \tag{16.224}$$

The above allows us to define

$$g_{v,w}:]-2,2[\rightarrow Y \text{ by } g_{v,w}=f\circ\psi_{v,w}-f\circ\varphi_{v,w} \tag{16.225}$$

As by [example: 16.112] $\psi_{v,w},\varphi_{v,w}$ are of class C^∞ and f is of class C^2 it follows from [theorems: 16.129, 16.113] it follows that $g_{v,w}$ is of class C^2. Hence by [theorem: 16.107] $g_{v,w}$ is of class C^1. So, as $L_{0,1}\underset{\text{[example: 16.161]}}{=}[0,1]\subseteq]-2,2[$ and $0\in]-2,2[$, we can use the second mean value theorem (IV) [theorem: 16.165] to get

$$\|g_{v,w}(0)-g_{v,w}(1)-D_0g_{v,w}(0-1)\|_Y\leqslant\sup(\{\|D_\zeta g_{v,w}-D_0g_{v,w}\|_{L(\mathbb{K},Y)}|\zeta\in L_{0,1}\})\cdot|0-1|$$

As $L_{0,1}=[0,1]$ and $|0-1|=1$ we have

$$\|g_{v,w}(0)-g_{v,w}(1)-D_0g_{v,w}(0-1)\|_Y\leqslant\sup(\{\|D_\zeta g_{v,w}-D_0g_{v,w}\|_{L(\mathbb{K},Y)}|\zeta\in[0,1]\}) \tag{16.226}$$

As $D_0g_{v,w}(1)\underset{\text{[theorem: 16.32]}}{=}(g_{v,w})'_0$ we have that

$$\begin{array}{rcl}\|g_{v,w}(0)-g_{v,w}(1)-D_0g_{v,w}(0-1)\|_Y & = & \|g_{v,w}(1)-g_{v,w}(0)-D_0g_{v,w}(1-0)\|_Y\\ & = & \|g_{v,w}(1)-g_{v,w}(0)-(g_{v,w})'_0\|_Y\end{array}$$

so that by [eq: 16.226]

$$\|g_{v,w}(1)-g_{v,w}(0)-(g_{v,w})'_0\|_Y\leqslant\sup(\{\|D_\zeta g_{v,w}-D_0g_{v,w}\|_{L(\mathbb{K},Y)}|\zeta\in[0,1]\}) \tag{16.227}$$

Now given $h\in\mathbb{R}$ we have

$$\begin{array}{rcl}\|(D_\zeta g_{v,w}-D_0g_{v,w})(h)\|_Y & = & \|D_\zeta g_{v,w}(h)-D_0g_{v,w}(h)\|_Y\\ & \underset{\text{[theorem: 16.32]}}{=} & \|h\cdot(g_{v,w})'_\zeta-h\cdot(g_{v,w})'_0\|_Y\\ & = & \|(g_{v,w})'_\zeta-(g_{v,w})'_0\|_Y\cdot|h|\end{array}$$

proving by the definition of the norm of a linear continuous mapping [see theorem: 14.181] that $\|D_\zeta g_{v,w}-D_0g_{v,w}\|_{L(\mathbb{K},Y)}\leqslant\|(g_{v,w})'_\zeta-(g_{v,w})'_0\|_Y$, hence by [eq: 16.227] it follows that

$$\|g_{v,w}(1)-g_{v,w}(0)-(g_{v,w})'_0\|_Y\leqslant\sup(\{\|(g_{v,w})'_\zeta-(g_{v,w})'_0\|_Y|\zeta\in[0,1]\}) \tag{16.228}$$

Let $\zeta\in[0,1]$ then we have using the chain rule that

$$\begin{array}{rcl}(g_{v,w})'_\zeta & \underset{\text{[theorem: 16.32]}}{=} & D_\zeta g_{v,w}(1)\\ & = & D_\zeta(f\circ\psi_{v,w}-f\circ\varphi_{v,w})(1)\\ & \underset{\text{[theorem: 16.35]}}{=} & D_\zeta(f\circ\psi_{v,w})(1)-\psi_{v,w}D_\zeta(f\circ\varphi_{v,w})(1)\\ & \underset{\text{[theorem: 16.36]}}{=} & (D_{\psi_{v,w}(\zeta)}f\circ D_\zeta\psi_{v,w})(1)-(D_{\varphi_{v,w}(\zeta)}f\circ D_\zeta\varphi_{v,w})(1)\\ & = & D_{\psi_{v,w}(\zeta)}f(D_\zeta\psi_{v,w}(1))-D_{\varphi_{v,w}(\zeta)}f(D_\zeta\varphi_{v,w}(1))\\ & \underset{\text{[example: 16.112]}}{=} & D_{\psi_{v,w}(\zeta)}f(v)-D_{\varphi_{v,w}(\zeta)}f(v)\\ & = & D_{x+\zeta\cdot v+w}f(v)-D_{x+\zeta\cdot v}f(v)\end{array}$$

so that

$$\forall\zeta\in[0,1]\ (g_{v,w})'_\zeta=D_{x+\zeta\cdot v+w}f(v)-D_{x+\zeta\cdot v}f(v) \tag{16.229}$$

As f is C^2 we have that the function

$$Df:U\rightarrow L_1(X;Y)=L(X,Y) \text{ defined by } Df(x)=D_xf$$

is Fréchet differentiable on x and $D_x(Df) = D_x^{[2]}f$. Then given $\varepsilon \in \mathbb{R}^+$ there exist a $\delta \in \mathbb{R}^+$ such that if $h \in U_x$ and $\|x\| < \delta$ then

$$\left\|Df(x+h) - Df(x) - D_x^{[2]}f(h)\right\|_{L(X,Y)} \leqslant \varepsilon \cdot \|h\|_X \tag{16.230}$$

Let $v, w \in X$ with $\|v\|_X < \min\left(\frac{\delta}{4}, \frac{r}{4}\right)$, $\|w\|_X \leqslant \min\left(\frac{\delta}{4}, \frac{r}{4}\right)$ then $\forall \zeta \in [0,1]$ we have

$$\|\zeta \cdot v + w\|_X \leqslant |\zeta| \cdot \|v\|_X + \|w\|_X \leqslant 1 \cdot \frac{\delta}{4} + \frac{\delta}{4} = \frac{\delta}{2} < \delta \text{ and } \|\zeta \cdot v\|_X = |\zeta| \cdot \|v\|_X \leqslant \frac{\delta}{4} < \delta$$

As $\|v\|_X, \|w\|_X < \frac{r}{4}$ we have that

$$\|x + \zeta \cdot v + w - x\|_X \leqslant |\zeta| \cdot \|v\|_X + \|w\|_X \leqslant \frac{r}{4} + \frac{r}{4} = r \text{ and } \|x + \zeta \cdot v - x\|_X \leqslant |\zeta| \cdot \|v\|_X < r$$

so that $x + \zeta \cdot v + w, x + \zeta \cdot v \in B_{\|\|_X}(x, r) \subseteq U$ or

$$\zeta \cdot v + w, \zeta \cdot v \in U_x$$

Using the above we have that

$$\left\|Df(x + \zeta \cdot v + w) - Df(x) - D_x^{[2]}f(\zeta \cdot v + w)\right\|_{L(X,Y)} \leqslant \varepsilon \cdot \|\zeta \cdot v + w\|_X \tag{16.231}$$

$$\left\|Df(x + \zeta \cdot v) - Df(x) - D_x^{[2]}f(\zeta \cdot v)\right\|_{L(X,Y)} \leqslant \varepsilon \cdot \|\zeta \cdot v\|_X \tag{16.232}$$

Also

$$D_x^{[2]}f(\zeta \cdot v + w)(v) - D_x^{[2]}f(\zeta \cdot v)(v) \underset{D_x^2 f \in L(X, L(X,Y))}{=} D_x^{[2]}(w)(v) \tag{16.233}$$

Now for $\zeta \in [0,1]$

$$\begin{array}{rl}
\left\|(g_{v,w})'_\zeta - D_x^{[2]}f(w)(v)\right\|_Y & \underset{\text{[eq: 16.229]}}{=} \\
\left\|D_{x+\zeta\cdot v+w}f(v) - D_{x+\zeta\cdot v}f(v) - D_x^{[2]}f(w)(v)\right\|_Y & \underset{\text{def of } Df}{=} \\
\left\|Df(x+\zeta\cdot v+w)(v) - Df(x+\zeta\cdot v)(v) - D_x^{[2]}f(w)(v)\right\|_Y & \underset{\text{[eq: 16.233}}{=} \\
\left\|\underbrace{Df(x+\zeta\cdot v+w)(v)}_{1} - \underbrace{Df(x+\zeta\cdot v)(v)}_{2} - \left(\underbrace{D_x^{[2]}f(\zeta\cdot v+w)(v)}_{3} - \underbrace{D_x^{[2]}f(\zeta\cdot v)(v)}_{4}\right)\right\|_Y & = \\
\left\|\underbrace{Df(x+\zeta\cdot v+w)(v)}_{1} - Df(x)(v) - \underbrace{D_x^{[2]}f(\zeta\cdot v+w)(v)}_{3} - \left(\underbrace{Df(x+\zeta\cdot v)(v)}_{2} - Df(x)(v) - \underbrace{D_x^{[2]}f(\zeta\cdot v)(v)}_{4}\right)\right\|_Y & \leqslant \\
\left\|Df(x+\zeta\cdot v+w)(v) - Df(x)(v) - D_x^{[2]}f(\zeta\cdot v+w)(v)\right\|_Y + \left\|Df(x+\zeta\cdot v)(v) - Df(x)(v) - D_x^{[2]}f(\zeta\cdot v)(v)\right\|_Y & \leqslant \\
\left\|Df(x+\zeta\cdot v+w) - Df(x) - D_x^{[2]}f(\zeta\cdot v+w)\right\|_{L(X,Y)} \cdot \|v\|_X + \left\|Df(x+\zeta\cdot v)(v) - Df(x) - D_x^{[2]}f(\zeta\cdot v)\right\|_Y \cdot \|v\|_X & \leqslant_{\text{[eqs: 16.231, 16.232}} \\
\varepsilon \cdot (\|\zeta\cdot v+w\|_X + \|\zeta\cdot v\|_X) \cdot \|v\|_X & \leqslant \\
\varepsilon \cdot \|v_x\| \cdot (|\zeta| \cdot \|v\|_x + \|w\|_X + |\zeta| \cdot \|v\|_X) & \leqslant_{\zeta\in[0,1]} \\
\varepsilon \cdot \|v_x\| \cdot (\|v\|_x + \|w\|_{X+}\|v\|_X) & \leqslant \\
\varepsilon \cdot \|v_x\| \cdot (\|v\|_x + \|w\|_X + \|w\|_X + \|v\|_X) & = \\
2 \cdot \varepsilon \cdot \|v\|_X \cdot (\|v\|_X + \|w\|_X) &
\end{array}$$

proving that

$$\forall \zeta \in [0,1] \text{ we have } \left\|(g_{v,w})'_\zeta - D_x^{[2]}f(w)(v)\right\|_Y \leqslant 2 \cdot \varepsilon \cdot \|v\|_X \cdot (\|v\|_X + \|w\|_X) \tag{16.234}$$

Further for $t \in [0,1]$ we have:

$$\begin{array}{rl}
\|(g_{v,w})'_t - (g_{v,w})'_0\|_Y & = \\
\left\|(g_{v,w})'_t - D_x^{[2]}f(w)(v) - \left((g_{v,w})'_0 - D_x^{[2]}f(w)(v)\right)\right\|_Y & \leqslant \\
\left\|(g_{v,w})'_t - D_x^{[2]}f(w)(v)\right\|_Y + \left\|(g_{v,w})'_0 - D_x^{[2]}f(w)(v)\right\|_Y & \leqslant_{t,0\in[0,1]\wedge\text{[eq: 16.234]}} \\
2 \cdot \varepsilon \cdot \|v\|_X \cdot (\|v\|_X + \|w\|_X) + 2 \cdot \varepsilon \cdot \|v\|_X \cdot (\|v\|_X + \|w\|_X) & = \\
4 \cdot \varepsilon \cdot \|v\|_X \cdot (\|v\|_X + \|w\|_X) &
\end{array}$$

proving

$$\|(g_{v,w})'_t - (g_{v,w})'_0\|_Y \leqslant 4\cdot\varepsilon\cdot\|v\|_X\cdot(\|v\|_X+\|w\|_X) \tag{16.235}$$

As $g_{v,w}$ is of class C^1, it follows from [theorem: 16.100] that the function

$$g'_{v,w}:]-2,2[\rightarrow Y \text{ defined by } g'_{v,w}(\zeta)=(g_{v,w})'_\zeta$$

is continuous. Define

$$h:]-2,2[\rightarrow\mathbb{R} \text{ by } h(\zeta)=\|g'_{v,w}(\zeta)-g'_{v,w}(0)\|_Y=(\|\|\circ(g'_{v,w}-C_{g'_{v,w}(0)}))(\zeta)$$

then as $g'_{v,w}$ is continuous we have by [example: 14.131] and [theorems: 14.149, 14.138] that h is continuous. Then as $[0,1]$ is compact and $[0,1]\subseteq]-2,2[$ it follows from the extreme value theorem [theorem: 14.235] that there exist a $\xi_0\in[0,1]$ such that $\forall t\in[0,1]$ $h(t)\leqslant h(\xi_0)$. Hence we have $h(\xi_0)\leqslant\sup(h[0,1])\leqslant h(\xi_0)$ or

$$\|g'_{v,w}(\xi_0)-g'_{v,w}(0)\|_Y=h(\xi_0)=\sup(h([0,1])) \tag{16.236}$$

As $h([0,1])=\{h(\zeta)|\zeta\in[0,1]\}=\{\|g'_{v,w}(\zeta)-g'_{v,w}(0)\|_Y|\zeta\in[0,1]\}$ it follows that

$$h(\zeta_0)=\sup(\{\|g'_{v,w}(\zeta)-g'_{v,w}(0)\|_Y|\zeta\in[0,1]\})=\sup(\{\|(g_{v,w})'_\zeta-(g_{v,w})'_0\|_Y|\zeta\in[0,1]\}) \tag{16.237}$$

Next

$$\begin{array}{rl}
\left\|g_{v,w}(1)-g_{v,w}(0)-D^{[2]}_x f(w)(v)\right\|_Y & = \\
\left\|g_{v,w}(1)-g_{v,w}(0)-(g_{v,w})'_0+(g_{v,w})'_0-(g_{v,w})'_{\xi_0}+(g_{v,w})'_{\xi_0}-D^{[2]}_x f(w)(v)\right\|_Y & \leqslant \\
\|g_{v,w}(1)-g_{v,w}(0)-(g_{v,w})'_0\|_Y+\|(g_{v,w})'_0-(g_{v,w})'_{\xi_0}\|_Y+\left\|(g_{v,w})'_{\xi_0}-D^{[2]}_x f(w)(v)\right\|_Y & \leqslant_{\text{[eq: 16.228]}} \\
\sup(\{\|(g_{v,w})'_\zeta-(g_{v,w})'_0\|_Y|\zeta\in[0,1]\})+\|(g_{v,w})'_0-(g_{v,w})'_{\xi_0}\|_Y+\left\|(g_{v,w})'_{\xi_0}-D^{[2]}_x f(w)(v)\right\|_Y & =_{\text{[eq: 16.237,16.236]}} \\
\|g'_{v,w}(\xi_0)-g'_{v,w}(0)\|_Y+\|(g_{v,w})'_0-(g_{v,w})'_{\xi_0}\|_Y+\left\|(g_{v,w})'_{\xi_0}-D^{[2]}_x f(w)(v)\right\|_Y & = \\
\|(g_{v,w})'_{\xi_9}-(g_{v,w})'_0\|_Y+\|(g_{v,w})'_0-(g_{v,w})'_{\xi_0}\|_Y+\left\|(g_{v,w})'_{\xi_0}-D^{[2]}_x f(w)(v)\right\|_Y & = \\
2\cdot\|(g_{v,w})'_{\xi_0}-(g_{v,w})'_0\|_Y+\left\|(g_{v,w})'_{\xi_0}-D^{[2]}_x f(w)(v)\right\|_Y & \leqslant_{\text{[eq: 16.235]}} \\
8\cdot\varepsilon\cdot\|v\|_X\cdot(\|v\|_X+\|w\|_X)+\left\|(g_{v,w})'_{\xi_0}-D^{[2]}_x f(w)(v)\right\|_Y & \leqslant_{\text{[eq: 16.234]}} \\
8\cdot\varepsilon\cdot\|v\|_X\cdot(\|v\|_X+\|w\|_X)+2\cdot\varepsilon\cdot\|v\|_X\cdot(\|v\|_X+\|w\|_X) & = \\
10\cdot\varepsilon\cdot\|v\|_X\cdot(\|v\|_X+\|w\|_X) &
\end{array}$$

proving that $\forall v,w\in X$ with $\|v\|_X<\min\left(\frac{\delta}{4},\frac{r}{4}\right)$, $\|w\|_X<\min\left(\frac{\delta}{4},\frac{r}{4}\right)$ we have

$$\left\|g_{v,w}(1)-g_{v,w}(0)-D^{[2]}_x f(w)(v)\right\|_Y\leqslant 10\cdot\varepsilon\cdot\|v\|_X\cdot(\|v\|_X+\|w\|_X) \tag{16.238}$$

As for $v'=w$ and w' we have also $\|v'\|_X<\min\left(\frac{\delta}{4},\frac{r}{4}\right)$, $\|w'\|_X<\min\left(\frac{\delta}{4},\frac{r}{4}\right)$ it follows that

$$\left\|g_{v',w'}(1)-g_{v',w'}(0)-D^{[2]}_x f(w')(v')\right\|_Y\leqslant 10\cdot\varepsilon\cdot\|v'\|_X\cdot(\|v'\|_X+\|w'\|_X)$$

or

$$\left\|g_{w,v}(1)-g_{w,v}(0)-D^{[2]}_x f(v)(w)\right\|_Y\leqslant 10\cdot\varepsilon\cdot\|w\|_X\cdot(\|v\|_X+\|w\|_X) \tag{16.239}$$

Now

$$\begin{array}{rcl}
g_{v,w}(1)-g_{v,w}(0) & = & (f\circ\psi_{v,w})(1)-(f\circ\varphi_{v,w})(1)-((f\circ\psi_{v,w})(0)-(f\circ\varphi_{v,w})(0)) \\
& = & f(\psi_{v,w}(1))-f(\varphi_{v,w}(1))-f(\psi_{v,w}(0))+f(\varphi_{v,w}(0)) \\
& = & f(x+v\cdot 1+w)-f(x+v\cdot 1)-f(x+v\cdot 0+w)+f(x+v\cdot 0) \\
& = & f(x+v+w)-f(x+v)-f(x+w)+f(x) \\
& = & f(x+w+v)-f(x+w)-f(x+v)+f(x) \\
& = & f(x+w\cdot 1+v)-f(x+w\cdot 1)-f(x+w\cdot 0+v)+f(x+w\cdot 0) \\
& = & f(\psi_{w,v}(1))-f(\varphi_{w,v}(1))-f(\psi_{w,v}(0))+f(\varphi_{w,v}(0)) \\
& = & (f\circ\psi_{w,v})(1)-(f\circ\varphi_{w,v})(1)-((f\circ\psi_{w,v})(0)-(f\circ\varphi_{w,v})(0)) \\
& = & g_{w.v}(1)-g_{w,v}(0)
\end{array}$$

substituting in [eq: 16.239] gives

$$\left\| g_{v,w}(1) - g_{v,w}(0) - D_x^{[2]} f(v)(w) \right\|_Y \leqslant 10 \cdot \varepsilon \cdot \|w\|_X \cdot (\|v\|_X + \|w\|_X) \tag{16.240}$$

Next

$$\begin{array}{rl} \left\| D_x^{[2]} f(v)(w) - D_x^{[2]} f(w)(v) \right\|_Y & = \\ \left\| g_{v,w}(1) - g_{v,w}(0) - D_x^{[2]} f(w)(v) - \left(g_{v,w}(1) - g_{v,w}(0) - D_x^{[2]} f(v)(w) \right) \right\|_Y & \leqslant \\ \left\| g_{v,w}(1) - g_{v,w}(0) - D_x^{[2]} f(w)(v) \right\|_Y + \left\| g_{v,w}(1) - g_{v,w}(0) - D_x^{[2]} f(v)(w) \right\|_Y & \leqslant_{[\text{eq: } 16.238, 16.240]} \\ 10 \cdot \varepsilon \cdot \|v\|_X \cdot (\|v\|_X + \|w\|_X) + 10 \cdot \varepsilon \cdot \|w\|_X \cdot (\|v\|_X + \|w\|_X) & = \\ \varepsilon \cdot [10 \cdot \|v\|_X \cdot (\|v\|_X + \|w\|_X) + 10 \cdot \|w\|_X \cdot (\|v\|_X + \|w\|_X)] & \end{array}$$

As $\varepsilon \in \mathbb{R}^+$ was chosen arbitrary it follows from [theorem: 10.31] that $\left\| D_x^{[2]} f(v)(w) - D_x^{[2]} f(w)(v) \right\|_Y \leqslant 0$ so that $\left\| D_x^{[2]} f(v)(w) - D_x^{[2]} f(w)(v) \right\|_Y = 0$ proving that $D_x^{[2]} f(v)(w) = D_x^{[2]} f(w)(v)$. Hence

$$\forall v, w \in X \text{ with } \|v\|_X < \min\left(\frac{\delta}{4}, \frac{r}{4}\right), \|w\|_X < \min\left(\frac{\delta}{4}, \frac{r}{4}\right) \text{ we have } D_x^{[2]} f(v)(w) = D_x^{[2]} f(w)(v)$$

Let $u, v \in X$ then for $v' = \frac{\min\left(\frac{\delta}{4}, \frac{r}{4}\right)}{\|v\|_X + 1} \cdot v$, $w' = \frac{\min\left(\frac{\delta}{4}, \frac{r}{4}\right)}{\|w\|_X + 1} \cdot w$ we have $\|v'\|_X < \min\left(\frac{\delta}{4}, \frac{r}{4}\right), \|w'\|_X < \min\left(\frac{\delta}{4}, \frac{r}{4}\right)$ so that

$$\begin{array}{rcl} \frac{\min\left(\frac{\delta}{4}, \frac{r}{4}\right)}{\|v\|_X + 1} \cdot \frac{\min\left(\frac{\delta}{4}, \frac{r}{4}\right)}{\|w\|_X + 1} \cdot D_x^{[2]} f(v)(w) & = & D_x^{[2]} f\left(\frac{\min\left(\frac{\delta}{4}, \frac{r}{4}\right)}{\|v\|_X + 1} \cdot v \right)\left(\frac{\min\left(\frac{\delta}{4}, \frac{r}{4}\right)}{\|w\|_X + 1} \cdot w \right) \\ & = & D_x^{[2]} f(v')(w') \\ & = & D_x^{[2]} f(w')(v') \\ & = & D_x^{[2]} f\left(\frac{\min\left(\frac{\delta}{4}, \frac{r}{4}\right)}{\|w\|_X + 1} \cdot w \right)\left(\frac{\min\left(\frac{\delta}{4}, \frac{r}{4}\right)}{\|v\|_X + 1} \cdot v \right) \\ & = & \frac{\min\left(\frac{\delta}{4}, \frac{r}{4}\right)}{\|v\|_X + 1} \cdot \frac{\min\left(\frac{\delta}{4}, \frac{r}{4}\right)}{\|w\|_X + 1} \cdot D_x^{[2]} f(w)(v) \end{array}$$

proving finally that

$$D_x^{[2]} f(v)(w) = D_x^{[2]} f(w)(v) \text{ for any } v, w \in X.$$

□

The following theorem shows that if a differential, as a member of $L_n(X; Y)$, needs only to be symmetric for a transposition [see definition: 11.212].

Lemma 16.172. *Let $n \in \mathbb{N} \setminus \{1\}$, $\langle X, \|\|_X \rangle$, $\langle Y, \|\|_Y \rangle$ be normed spaces, $L \in L_n(X; Y)$, $(x_1, \ldots, x_n) \in X^n$ then if $\forall i \in \{1, \ldots, n-1\}$ $L(x_1 : \ldots : x_n) = L(x_{(i \underset{n}{\leftrightarrow} i+1)(1)} : \ldots : x_{(i \underset{n}{\leftrightarrow} i+1)(n)})$ it follows that $\forall \sigma \in P_n$ $L(x_1 : \ldots : x_n) = L(x_{\sigma(1)} : \ldots : x_{\sigma(n)})$*

Proof. Using [theorem: 16.85] there exist a linear isometric isomorphism

$$\mathcal{I}_{n,X,Y} : L_n(X; Y) \rightarrow L^n(X; Y)$$

such that $\forall (x_1, \ldots, x_n) \subseteq X^n$ $(\mathcal{I}_{n,X,Y}(L))(x_1, \ldots, x_n) = L(x_1 : \ldots : x_n)$. Take $K = \mathcal{I}_{n,X,Y}(L)$ then $\forall i \in \{1, \ldots, n-1\}$ we have

$$\begin{array}{rcl} \left(\left(i \underset{n}{\longleftrightarrow} i+1 \right) K \right)(x_1, \ldots, x_n) & = & \left(K(x_{(i \underset{n}{\leftrightarrow} i+1)(1)}, \ldots, x_{(i \underset{n}{\leftrightarrow} i+1)(n)}) \right) \\ & = & L(x_{(i \underset{n}{\leftrightarrow} i+1)(1)} : \ldots : x_{(i \underset{n}{\leftrightarrow} i+1)(n)}) \\ & = & L(x_1 : \ldots : x_n) \\ & = & K(x_1, \ldots, x_n) \end{array}$$

proving that $\left(i \underset{n}{\longleftrightarrow} i+1\right)K = K$. Hence if $\sigma \in P_n$ we have by [corollary: 11.226] that $\sigma K = K$. Hence

$$\begin{aligned} L(x_{\sigma(1)}:\ldots:x_{\sigma(n)}) &= K(x_{\sigma(1)},\ldots,x_{\sigma(n)}) \\ &= (\sigma K)(x_1,\ldots,x_n) \\ &= K(x_1,\ldots,x_n) \\ &= L(x_1:\ldots:x_n) \end{aligned}$$

□

Next we define the evaluation operator and prove that is linear and continuous hence of class C^∞.

Definition 16.173. *Let $n \in \mathbb{N}$, $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$ be normed spaces and $(x_1,\ldots,x_n) \in X^n$ then the evaluation operator $\mathrm{ev}_{x_1,\ldots,x_n}$ is defined by*

$$\mathrm{ev}_{x_1,\ldots,x_n}: L_n(X;Y) \rightarrow Y \text{ where } \mathrm{ev}_{x_1,\ldots,x_n}(L) = L(x_1:\ldots:x_n)$$

Lemma 16.174. *Let $n \in \mathbb{N}$, $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$ be normed spaces and $(x_1,\ldots,x_n) \in X^n$ then*

$$\mathrm{ev}_{x_1,\ldots,x_n} \in L(L_n(X;Y), Y)$$

Hence by [theorem: 16.116] $\mathrm{ev}_{x_1,\ldots,x_n}$ is of class C^∞ and $\forall L \in L_n(X;Y)$

$$D_L^{[n]}\mathrm{ev}_{x_1,\ldots,x_n} = \begin{cases} \mathrm{ev}_{x_1,\ldots,x_n} \text{ if } n=1 \\ 0_n \text{ if } 1<n \end{cases}$$

Proof. Let $\beta \in \mathbb{K}$, $L_1, L_2 \in L_n(X;Y)$ then

$$\begin{aligned} \mathrm{ev}_{x_1,\ldots,x_n}(L_1 + \beta \cdot L_2) &= (L_1 + \beta \cdot L_2)(x_1:\ldots:x_n) \\ &\underset{[\text{theorem: } 16.79]}{=} L_1(x_1:\ldots:x_n) + \beta \cdot L_2(x_1:\ldots:x_n) \\ &= \mathrm{ev}_{x_1,\ldots,x_n}(L_1) + \beta \cdot \mathrm{ev}_{x_1,\ldots,x_n}(L_2) \end{aligned}$$

proving that

$$\mathrm{ev}_{x_1,\ldots,x_n} \in \mathrm{Hom}(L_n(X;Y), Y)$$

Further if $L \in L_n(X;Y)$ then

$$\begin{aligned} \|\mathrm{ev}_{x_1,\ldots,x_n}(L)\|_Y &= \|L(x_1:\ldots:x_n)\|_Y \\ &\leqslant_{[\text{lemma: } 16.81]} \|L\|_{L_n(X;Y)} \cdot \prod_{i=1}^n \|x_i\|_X \\ &= \left(\prod_{i=1}^n \|x_i\|_X\right) \cdot \|L\|_{L_n(X;Y)} \end{aligned}$$

proving by [theorem: 14.174] that $\mathrm{ev}_{x_1,\ldots,x_n}$ is continuous hence

$$\mathrm{ev}_{v_1,\ldots,v_n} \in L(L_n(X;Y), Y)$$ □

Lemma 16.175. *Let $n \in \mathbb{N}$, $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$ be normed spaces, U a open set in X, $(x_1,\ldots,x_n) \in X^n$ and $f: Y \rightarrow Y$ a function of class C^{n+1} then for*

$$g: U \rightarrow Y \text{ defined by } g(y) = D_y^{[n]} f(x_1:\ldots:x_n)$$

we have that g is of class C^1 and $\forall y \in U$ we have $\forall h \in X$ that $D_y g(h) = D_y^{[n+1]} f(h: x_1:\ldots:x_n)$

Proof. As f is of class C^{n+1}, hence $(n+1)$-times differentiable on U, the function

$$D^{[n]} f: U \rightarrow L_n(X;Y) \text{ defined by } D^{[n]} f(y) = D_y^{[n]} f$$

is differentiable on U with $\forall y\in U$ $D_y^{[n+1]}f=D_y(D^{[n]}f)$. Let $y\in U$ then we have

$$\begin{array}{rcl} g(y) & = & D_y^{[n]}f(x_1:\ldots:x_n)\\ & = & \mathrm{ev}_{x_1,\ldots x_n}\big(D_y^{[n]}f\big)\\ & = & \mathrm{ev}_{x_1,\ldots,x_n}(D^{[n]}f(y))\\ & = & (\mathrm{ev}_{x_1,\ldots,x_n}\circ D^{[n]}f)(y)\end{array}$$

proving that

$$g=(\mathrm{ev}_{x_1,\ldots,x_n}\circ D^{[n]}f)$$

As f is of class C^{n+1} it follows that f is of class C^n [see theorem: 16.107] hence $D^{[n]}f$ must be continuous which as $D^{[n]}f$ is differentiable on U proves that $D^{[n]}f$ is of class C^1. Given that by [lemma: 16.174] $\mathrm{ev}_{x_1,\ldots,x_n}$ is of class C^∞ it follows from the chain rule [theorem: 16.129]

$$g=\mathrm{ev}_{x_1,\ldots,x_n}\circ D^{[n]}f \text{ is of class } C^1 \tag{16.241}$$

Further

$$\begin{array}{rcl} D_y g & = & D_y(\mathrm{ev}_{x_1,\ldots,x_n}\circ D^{[n]}f)\\ & \underset{[\text{theorem: 16.36}]}{=} & D_{D^{[n]}f(y)}\mathrm{ev}_{x_1,\ldots,x_n}\circ D_y(D^{[n]}f)\\ & \underset{[\text{lemma: 16.174}]}{=} & \mathrm{ev}_{x_1,\ldots,x_n}\circ D_y(D^{[n]}f)\\ & = & \mathrm{ev}_{x_1,\ldots,x_n}\circ D_y^{[n+1]}f\end{array}$$

proving that

$$D_y g=\mathrm{ev}_{x_1,\ldots,x_n}\circ D_y^{[n+1]}f$$

Hence if $h\in X$ we have

$$\begin{array}{rcl} D_y g(h) & = & \big(\mathrm{ev}_{x_1,\ldots,x_n}\circ D_y^{[n+1]}f\big)(h)\\ & = & \mathrm{ev}_{x_1,\ldots,x_n}\big(D_y^{[n+1]}f(h)\big)\\ & = & D_y^{[n+1]}f(h)(x_1:\ldots:x_n)\\ & = & D_y^{[n+1]}f(h:x_1:\ldots:x_n)\end{array}$$

The lemma is then proved by the above and [eq: 16.241]. □

Lemma 16.176. *Let $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$ be normed spaces then*

$$\mathrm{ev}\colon L(X,Y)\cdot X\rightarrow Y \text{ defined by } \mathrm{ev}(L,x)=L(x)$$

is of class C^∞. Further $\forall L\in L(X,Y)$ and $x\in X$ we have

$$D_{(L,x)}^{[1]}\mathrm{ev}=D_{(L,x)}\mathrm{ev}=\mathrm{ev}(L,*)+\mathrm{ev}(*,x)$$

[see theorem: 16.120].

Proof. Using [theorem: 16.120] it is enough to prove that ev is a bi-linear continuous function or in other words $\mathrm{ev}\in L(L(X,Y),X;Y)$. Let $L_1,L_2\in L(X,Y)$, $x\in X$ and $\alpha\in\mathbb{K}$ then

$$\begin{array}{rcl} \mathrm{ev}(L_1+\alpha\cdot L_2,x) & = & (L_1+\alpha\cdot L_2)(x)\\ & = & L_1(x)+\alpha\cdot L_2(x)\\ & = & \mathrm{ev}(L_1,x)+\alpha\cdot\mathrm{ev}(L_2,x)\end{array}$$

and for $L\in L(X,Y)$, $x,y\in X$ and $\alpha\in\mathbb{K}$

$$\begin{array}{rcl} \mathrm{ev}(L,x+\alpha\cdot y) & = & L(x+\alpha\cdot y)\\ & = & L(x)+\alpha\cdot L(y)\\ & = & \mathrm{ev}(L,x)+\alpha\cdot\mathrm{ev}(L,y)\end{array}$$

proving that

$$\mathrm{ev}\in \mathrm{Hom}(L(X,Y),X;Y)$$

Let $(L,x)\in L(X,Y)\cdot X$ then

$$\begin{array}{rcl}\|\mathrm{ev}(L,x)\|_Y & = & \|L(x)\|_Y\\ & \leqslant & \|L\|_{L(X,Y)}\cdot\|x\|_X\end{array}$$

proving by [theorem: 14.174] that

$$\mathrm{ev}\in L(L(X,Y),X;Y)$$

□

Lemma 16.177. *Let $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$ be normed spaces, U a open set in X, $x\in U$, $n\in\mathbb{N}\setminus\{1\}$ and*

$$f:U\rightarrow Y \text{ is a } n\text{-times differentiable at } x$$

then for $h\in X$ we have that there exist a open set V with $x\in V\subseteq U$

$$Df(*)(h):U\rightarrow Y \text{ defined by } (Df(*)(h))(y)=Df(y)(h)$$

is $(n-1)$-times differentiable at x. Further $\forall k,h,x\in X$ we have

$$(D_x(Df(*)(h)))(k) \;=\; \left(\left(D_x^{[2]}f\right)(k)\right)(h)$$

Proof. As f is n-times differentiable at x and $1<n$ we have by [theorem: 16.109] that there exist a open set V such that $x\in V\subseteq U$ and $\forall y\in V$ we have that f is 1-times differentiable at y and the function

$$Df:V\rightarrow L(X,Y) \text{ defined by } (Df)(y)=Df(y) \tag{16.242}$$

is $(n-1)$-times differentiable. Given $v\in X$ define

$$\varphi_v:U\rightarrow L(X,Y)\cdot X \text{ by } \varphi_v(x)=(Df(x),v) \tag{16.243}$$

Then we have for $y\in X$

$$\begin{array}{rcl}(\pi_1\circ\varphi_v)(y) & = & \pi_1(\varphi_v(y))\\ & = & \pi_1((Df(x),v))\\ & = & Df(x)\\ & = & (Df)(x)\\ (\pi_2\circ\varphi_v)(y) & = & \pi_2(\varphi_v(y))\\ & = & \pi_2((Df(x),v))\\ & = & v\end{array}$$

proving that

$$\pi_1\circ\varphi_v=Df \text{ and } \pi_2\circ\varphi_v=C_v \tag{16.244}$$

As $\pi_1\circ\varphi_v=Df$ is $(n-1)$-times differentiable at x and $\pi_2\circ\varphi_v=C$ is ∞-times differentiable at x [see theorem: 16.115] it follows by [theorem: 16.125] that

$$\varphi_v \text{ is } (n-1)\text{-times differentiable at } x \tag{16.245}$$

As $1<n$ it follows that $1\in\{1,\ldots,n-1\}$ so that by [theorem: 16.106] φ_v is 1-times differentiable at x, hence using [theorem: 16.61] it follows that

$$\pi_1\circ D_x\varphi_v=D_x(\pi_1\circ\varphi_v)\underset{[\text{eq: }16.244]}{=}D_x(Df)=D_x^{[2]}f$$

and

$$\pi_2\circ D_x\varphi_v=D_x(\pi_2\circ\varphi_v)\underset{[\text{eq: }16.244]}{=}D_xC_v\underset{[\text{theorem: }16.115]}{=}C_{1,0}$$

So that

$$D_x\varphi_v=\left(D_x^{[2]}f,C_{1,0}\right)_* \tag{16.246}$$

Given $h\in X$ we can, as $\forall y\in V$ f is 1-times differentiable at y, define

$$Df(*)(h):U\rightarrow Y \text{ by } (Df(*)(h))(y)=(Df(y))(h)$$

Let $h\in X$ then we have for $x\in U$

$$\begin{array}{rcl}(Df(*)(h))(x) & = & Df(x)(h)\\ & = & \mathrm{ev}(Df(x),h)\\ & = & \mathrm{ev}(\varphi_h(x))\\ & = & (\mathrm{ev}\circ\varphi_h)(x)\end{array}$$

proving that

$$Df(*)(h)=\mathrm{ev}\circ\varphi_h \tag{16.247}$$

As ev is of ∞-times differentiable by [lemma: 16.176] and φ_v is by [eq: 16.245] $(n-1)$-times differentiable at x it follows that

$$Df(*)(h) \text{ is } (n-1)\text{-times differentiable at } x$$

Given that $1<n$ it follows that $1\in\{1,\ldots,n-1\}$ so that by [theorem: 16.106] $Df(*)(h)$ is 1-times differentiable at x then for $k\in Z$ we have

$$\begin{array}{rcl}(D_x(Df(*)(h)))(k) & \underset{[\text{eq: } 16.247]}{=} & (D_x(\mathrm{ev}\circ\varphi_h))(k)\\ & \underset{[\text{theorem: } 16.36]}{=} & (D_{\varphi_h(x)}\mathrm{ev}\circ D_x\varphi_h)(k)\\ & = & (D_{(Df(x),h)}\circ D_x\varphi_h)(k)\\ & \underset{[\text{lemma: } 16.176]}{=} & ((\mathrm{ev}(D^{[1]}f(x),*)+\mathrm{ev}(*,h))\circ D\varphi_h(x))(k)\\ & \underset{[\text{eq: } 16.246]}{=} & \left((\mathrm{ev}(D^{[1]}f(x),*)+\mathrm{ev}(*,h))\circ\left(D_x^{[2]}f,C_{1,0}\right)_*\right)(k)\\ & = & (\mathrm{ev}(D^{[1]}f(x),*)+\mathrm{ev}(*,h))\left(\left(D_x^{[2]}f,C_{1,0}\right)_*(k)\right)\\ & = & (\mathrm{ev}(D^{[1]}f(x),*)+\mathrm{ev}(*,h))\left(D_x^{[2]}f(k),C_{1,0}(k)\right)\\ & = & (\mathrm{ev}(D^{[1]}f(x),*)+\mathrm{ev}(*,h))\left(D_x^{[2]}f(k),0\right)\\ & = & \mathrm{ev}(D^{[1]}f(x),0)+\mathrm{ev}\left(D_x^{[2]}f(k),h\right)\\ & = & D^{[1]}f(x)(0)+\left(\left(D_x^{[2]}f\right)(k)\right)(h)\\ & = & \left(D_x^{[2]}f(k)\right)(h)\end{array}$$

$\square$

Lemma 16.178. *Let $n\in\mathbb{N}$ with $2<n$, $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$ normed spaces, U a open set in X, $(x_1,\ldots,x_{n-2})\in X^n$ and $f:U\rightarrow Y$ a function of class C^n then the function*

$$g:U\rightarrow Y \text{ defined by } g(x)=D_x^{[n-2]}f(x_1,\ldots,x_{n-2})$$

is of class C^2 and for $u,v\in X$ we have $D_x^{[2]}g(u)(v)=D_x^{[n]}f(v:u:x_1:\ldots:x_{n-2})$

Proof. As f is of class C^n and $n-1,n-2\in\{1,\ldots,n\}$ it follows from [theorem: 16.107] that f is of class C^{n-2} and C^{n-1}, hence f is $(n-2)$-times differentiable and $(n-1)$-times differentiable on U so that the following functions are well defined

$$D^{[n-1]}f:U\rightarrow L_{n-1}(X;Y) \text{ defined by } D^{[n-1]}f(y)=D_y^{[n-1]}f$$

$$D^{[n-2]}f:U\rightarrow L_{n-2}(X;Y) \text{ defined by } D^{[n-2]}f(y)=D_y^{[n-2]}f$$

and continuous. Let $y\in U$ then we have

$$\begin{array}{rcl}g(y) & = & D_y^{[n-2]}f(x_1:\ldots:x_{n-2})\\ & = & D^{[n-2]}f(y)(x_1:\ldots:x_{n-2})\\ & = & \mathrm{ev}_{x_1,\ldots,x_{n-2}}(D^{[n-2]}f(y))\\ & = & (\mathrm{ev}_{x_1,\ldots,x_{n-2}}\circ D^{[n-2]}f)(y)\end{array}$$

proving that

$$g = \mathrm{ev}_{x_1,\ldots,x_{n-2}} \circ D^{[n-2]} f \tag{16.248}$$

As $n = (n-2) + 2$ it follows from [theorem: 16.108] that $D^{[n-2]} f$ is of class C^2 and $D^{[2]}(D^{[n-2]} f) = D^{[n]} f$ where

$$D^{[n]} f: U \to L_n(X;Y) \text{ is defined by } D^{[n]} f(x) = D_x^{[n]} f$$

$$D^{[2]}(D^{[n-2]} f): U \to L_2(X; L_{n-2}(X;Y)) = L_n(X;Y) \text{ defined by } D^{[2]}(D^{[n-2]} f)(x) = D_x^{[2]}(D^{[n-2]} f)$$

As $D^{[n-2]} f$ is of class C^2 and $\mathrm{ev}_{x_1,\ldots,x_{n-2}}$ is of class C^∞ [see lemma: 16.174] it follows from the chain rule [theorem: 16.129] that

$$g \text{ is of class } C^2 \text{ and by [theorem: 16.107] also of class } C^1$$

Take

$$Dg: U \to L(X,Y) \text{ defined by } Dg(y) = D_y g$$

then for $y \in U$ we have:

$$\begin{array}{rcl} Dg(y) & = & D_y g \\ & \underset{\text{[theorem: 16.36]}}{=} & D_{f(y)} \mathrm{ev}_{x_1,\ldots,x_{n-2}} \circ D_y(D^{[n-2]} f) \\ & \underset{\text{[lemma: 16.174]}}{=} & \mathrm{ev}_{x_1,\ldots,x_{n-2}} \circ D_y(D^{[n-2]} f) \\ & = & \mathrm{ev}_{x_1,\ldots,x_{n-2}} \circ D_y^{[n-1]} f \end{array}$$

Let $u \in X$ then we have

$$\begin{array}{rcl} Dg(y)(u) & = & \left(\mathrm{ev}_{x_1,\ldots,x_{n-2}} \circ D_y^{[n-1]} f\right)(u) \\ & = & \mathrm{ev}_{x_1,\ldots,x_{n-2}}\left(D_y^{[n-1]} f)u)\right) \\ & = & D_y^{[n-1]} f(u)(x_1: \ldots : x_{n-2}) \\ & = & D_y^{[n-1]} f(u: x_1: \ldots : x_{n-2}) \\ & = & \mathrm{ev}_{u,x_1,\ldots,x_{n-2}}\left(D_y^{[n-1]} f\right) \\ & = & \mathrm{ev}_{u,x_1,\ldots,x_{n-2}}(D^{[n-1]} f(y)) \\ & = & (\mathrm{ev}_{u,x_1,\ldots,x_{n-2}} \circ D^{[n-1]} f)(y) \end{array} \tag{16.249}$$

Define now

$$Dg(*)(u): U \to Y \text{ by } (Dg(*)(u))(y) = Dg(y)(u)$$

then we have $\forall y \in U \ (Dg(*)(u))(y) = Dg(y)(u) \underset{\text{[eq: 16.249]}}{=} (\mathrm{ev}_{u,x_1,\ldots,x_{n-2}} \circ D^{[n-1]})(y)$ so that

$$Dg(\star)(u) = \mathrm{ev}_{u,x_1,\ldots,x_{n-2}} \circ D^{[n-1]} f \tag{16.250}$$

As $D^{[n-1]} f$ is of class C^1 [see theorem: 16.108] and $\mathrm{ev}_{x_1,\ldots,x_{n-2}}$ is of class C^∞ [see lemma: 16.174], it follows from the chain rule [theorem: 16.129] and the above that

$$Dg(\star)(u) \text{ is of class } C^1 \text{ hence differentiable on } U$$

and for $y \in U$ that

$$\begin{array}{rcl} D_y(Dg(*)(u)) & = & D_y(\mathrm{ev}_{u,x_1,\ldots,x_{n-2}} \circ D^{[n-1]} f) \\ & \underset{\text{[theorem: 16.36]}}{=} & D_{D^{[n-1]} f(y)} \mathrm{ev}_{u,x_1,\ldots,x_{n-2}} \circ D_y(D^{[n-1]} f) \\ & \underset{\text{[lemma: 16.174]}}{=} & \mathrm{ev}_{u,x_1,\ldots,x_{n-2}} \circ D_y(D^{[n-1]} f) \\ & = & \mathrm{ev}_{u,x_1,\ldots,x_{n-2}} \circ D_y^{[n]} f \end{array} \tag{16.251}$$

Let additional $v\in X$ then

$$\begin{aligned} D_y^{[2]}g(v)(u) &\underset{[\text{lemma: } 16.177]}{=} (D_y(Dg(*)(u)))(v) \\ &= \left(\mathrm{ev}_{u,x_1,\ldots,x_{n-2}}\circ D_y^{[n]}f\right)(v) \\ &= \mathrm{ev}_{u,x_1,\ldots,x_{n-2}}\left(D_y^{[n]}f(v)\right) \\ &= D_y^{[n]}f(v)(u:x_1:\ldots x_{n-2}) \\ &= D_y^{[n]}f(v:u:x_1:\ldots:x_{n-2}) \end{aligned}$$

proving the last part of the lemma. □

Lemma 16.179. *Let $n\in\mathbb{N}\setminus\{1\}$, $\langle X,\|\|_X\rangle$ a normed space, $\langle Y,\|\|_Y\rangle$ a Banach space, U a open set in X, $f:U\rightarrow Y$ a function of class C^n then $\forall x\in U$ and $\forall i\in\{1,\ldots,n-1\}$ we have for $(u_1,\ldots,u_n)\in X^n$ that*

$$D_x^n f(u_1:\ldots:u_n)=D_x^n f(u_{(i\underset{n}{\leftrightarrow}i+1)}:\ldots:u_{(i\underset{n}{\leftrightarrow}i+1)})$$

Proof. We prove this by induction on n, so let

$S=\left\{n\in\{2,\ldots,\infty\}|\text{If } f:U\rightarrow Y \text{ is of class } C^n \text{ then } \forall x\in U\wedge\forall i\in\{1,\ldots,n-1\} \text{ we have that for } (u_1,\ldots,u_n)\in X^n\ D_x^n f(u_1:\ldots:u_n)=D_x^n f(u_{(i\underset{n}{\leftrightarrow}i+1)}:\ldots:u_{(i\underset{n}{\leftrightarrow}i+1)})\right\}$

then we have:

$2\in S$. Let $f:U\rightarrow Y$ be a function of class C^2, $x\in U$ then for $i\in\{1,\ldots,2-1\}=\{1\}\Rightarrow i=1$ and $(u_1,u_2)\in X^2$ we have

$$\begin{aligned} D_x^{[2]}f(u_1:u_2) &= D_x^{[2]}f(u_1)(u_2) \\ &\underset{[\text{theorem: } 16.171]}{=} D_x^{[2]}f(u_2)(u_1) \\ &= D_x^{[2]}f(u_2:u_1) \end{aligned}$$

proving that $2\in S$.

$n\in S\Rightarrow n+1\in S$. Let $f:U\rightarrow Y$ be a function of class C^2, $x\in U$, $(u_1,\ldots,u_{n+1})\in X^{n+1}$ then for $i\in\{1,\ldots,(n+1)-1\}=\{1,\ldots,n\}$ we have the following cases to consider:

$i=1$. Then $\left(i\underset{n+1}{\leftrightarrow}i+1\right)=\left(1\underset{n+1}{\leftrightarrow}2\right)$ so that $\left(i\underset{n+1}{\leftrightarrow}i+1\right)(1)=2$, $\left(i\underset{n+1}{\leftrightarrow}i+1\right)(2)=1$ and $\forall k\in\{3,\ldots,n+1\}$ $\left(i\underset{n+1}{\leftrightarrow}i+1\right)(k)=k$. Define

$$g:U\rightarrow Y \text{ by } g(y)=D_y^{[(n+1)-2]}f(u_3:\ldots:u_{n+1})$$

then by the previous lemma [lemma: 16.178] that g is of class C^2 and

$$\begin{aligned} D_x^{[n+1]}f(u_1:u_2:\ldots:u_{n+1}) &\underset{[\text{lemma: } 16.178]}{=} \\ D_x^{[2]}g(u_1)(u_2) &\underset{[\text{theorem: } 16.171]}{=} \\ D_x^{[2]}g(u_2)(u_1) &\underset{[\text{lemma: } 16.178]}{=} \\ D_x^{[n+1]}f(u_2:u_1:\ldots:u_{n+1}) &= \\ D_x^{[n+1]}f\left(u_{(i\underset{n+1}{\leftrightarrow}i+1)(1)}:u_{(i\underset{n+1}{\leftrightarrow}i+1)}:\ldots:u_{(i\underset{n+1}{\leftrightarrow}i+1)}\right) & \end{aligned}$$

proving that in this case $n+1\in S$.

$i\in\{2,\ldots,n\}$. Then $\forall k\in\{1,\ldots,n-1\}$ we have

$$\begin{aligned} \left(i-1\underset{n}{\leftrightarrow}i\right)(k)+1 &= 1+\begin{cases} i-1 \text{ if } k=i \\ i \text{ if } k=i-1 \\ k \text{ if } k\neq i-1,i \end{cases} \\ &= \begin{cases} i \text{ if } k=i \\ i+1 \text{ if } k=i-1 \\ k+1 \text{ if } k\neq i-1.i \end{cases} \end{aligned} \tag{16.252}$$

further

$$\left(i \underset{n+1}{\leftrightarrow} i+1\right)(k+1) = \begin{cases} i \text{ if } k+1=i+1 \\ i+1 \text{ if } k+1=i \\ k+1 \text{ if } k+1 \neq i+1, i \end{cases} = \begin{cases} i \text{ if } k=i \\ i+1 \text{ if } k=i-1 \\ k+1 \text{ if } k \neq i-1, i \end{cases} \underset{[\text{eq: } 16.252]}{=} \left(i-1 \underset{n}{\leftrightarrow} i\right)(k)+1$$

so that

$$\forall k \in \{1,\ldots,n-1\} \text{ we have } \left(i-1 \underset{n}{\leftrightarrow} i\right)(k)+1 = \left(i \underset{n+1}{\leftrightarrow} i+1\right)(k+1) \tag{16.253}$$

Let $y \in U$ then, as f is of class C^{n+1}, we have also that f is of class C^n [theorem: 16.107] , so we have if we define $(y_1,\ldots,y_n)$ by $y_i = u_{i+1}$ that

$$\begin{aligned} D_y^{[n]} f(u_2 \colon \ldots \colon u_{n+1}) &= D_y^{[n]} f(y_1 \colon \ldots \colon y_n) \\ &\underset{n \in S}{=} D_y^{[n]} f\left(y_{(i-1 \underset{n}{\leftrightarrow} i)(1)} \colon \ldots \colon y_{(i-1 \underset{n}{\leftrightarrow} i)(n)}\right) \\ &= D_y^{[n]} f\left(u_{(i-1 \underset{n}{\leftrightarrow} i)(1)+1} \colon \ldots \colon u_{(i-1 \underset{n}{\leftrightarrow} i)(n)+1}\right) \\ &\underset{[\text{eq: } 16.253]}{=} D_y^{[n]} f\left(u_{(i \underset{n+1}{\leftrightarrow} i+1)(2)} \colon \ldots \colon u_{(i \underset{n+1}{\leftrightarrow} i+1)(n+1)}\right) \end{aligned} \tag{16.254}$$

Define

$$g \colon U \to Y \text{ by } g(y) = D_y^{[n]} f(u_2 \colon \ldots \colon u_{n+1})$$

and

$$h \colon U \to Y \text{ by } h(y) = D_y^{[n]} f\left(u_{(i \underset{n+1}{\leftrightarrow} i+1)(2)} \colon \ldots \colon u_{(i \underset{n+1}{\leftrightarrow} i+1)(n+1)}\right)$$

then we have by [eq: 16.254] that

$$g = h \tag{16.255}$$

As $1 < i < i+1$ we have that $\left(i \underset{n+1}{\leftrightarrow} i+1\right)(1) = 1$ so that

$$\begin{aligned} D_x^{[n+1]} f(u_1 \colon \ldots \colon u_{n+1}) &= D_x^{[n+1]} f\left(u_{(i \underset{n+1}{\leftrightarrow} i+1)(1)} \colon u_2 \colon \ldots \colon u_{n+1}\right) \\ &\underset{[\text{lemma: } 16.175]}{=} D_y g\left(u_{(i \underset{n+1}{\leftrightarrow} i+1)(1)}\right) \\ &\underset{[\text{eq: } 16.255]}{=} D_y h\left(u_{(i \underset{n+1}{\leftrightarrow} i+1)(1)}\right) \\ &\underset{[\text{lemma: } 16.175]}{=} D_x^{[n+1]} f\left(u_{(i \underset{n+1}{\leftrightarrow} i+1)(1)} \colon \ldots \colon u_{(i \underset{n+1}{\leftrightarrow} i+1)(n+1)}\right) \end{aligned}$$

proving that in this case we have also $n+1 \in S$.

Mathematical induction proves then the theorem. □

Having done all the heavy lifting with the previous lemma's we can now prove that higher order differentials are symmetric in there arguments.

Theorem 16.180. *Let $n \in \mathbb{N}$, $\langle X, \|\|_X \rangle$ a normed space, $\langle Y, \|\|_Y \rangle$ a Banach space, U a open set and $f \colon U \to Y$ a function of class C^n then $\forall x \in U$, $\sigma \in P_n$ and $(x_1,\ldots,x_n) \in X^n$ we have*

$$D_x^{[n]} f(x_1 \colon \ldots \colon x_n) = D_x^{[n]} f(x_{\sigma(1)} \colon \ldots \colon x_{\sigma(n)})$$

***Note* 16.181.** *As $D_x^n f(y_1,\ldots,y_n) = D_x^{[n]}(y_1 \colon \ldots \colon y_n)$ we have also*

$$D_x^n f(x_1,\ldots \colon x_n) = D_x^n f(x_{\sigma(1)},\ldots,x_{\sigma(n)})$$

***Note* 16.182.** *If we define*

$$D^{[n]}f:U\rightarrow L_n(X;Y) \text{ by } D^{[n]}f(x)=D^{[n]}_x f$$

then we have

$$D^{[n]}f(x)(x_1:\ldots:x_n)=D^{[n]}f(x)(x_{\sigma(1)}:\ldots:x_{\sigma(n)})$$

Proof. For $n\in\mathbb{N}$ we have two possibilities to consider:

$\boldsymbol{n=1}$**.** The $P_n=\{\mathrm{Id}_{\{1\}}\}$ so that we have trivially

$$D_x^{[1]}f\left(\underbrace{x_1:\ldots:x_1}_{1}\right)=D_x^{[1]}f\left(\underbrace{x_{\sigma(1)}:\ldots:x_{\sigma(1)}}_{1}\right)$$

$\boldsymbol{n\in\mathbb{N}\setminus\{1\}}$**.** $\forall i\in\{1,\ldots,n-1\}$ we have by [lemma: 16.179] that

$$D_x^{[n]}f(x_1:\ldots:x_n)=D_x^{[n]}f(x_{\sigma(1)}:\ldots:x_{\sigma(n)})$$

Let $\sigma\in P_n$ then by the above and [lemma: 16.172] it follows that

$$D_x^{[n]}f(x_1:\ldots:x_n)=D_x^{[n]}f(x_{(\sigma 1)}:\ldots:x_{\sigma(n)})$$

□

16.5 Higher Order Partial Differentiation

16.5.1 Linear mappings to linear mappings

In order to define higher order partial derivation we must extend the definition of $L_n(X;Y)$ [see definition: 16.69]

Definition 16.183. *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 we define $L_n(X_1...X_n;Y)$ as follows*

$$L_n(X_1\ldots X_n;Y)=\begin{cases} L(X_1,Y) \text{ if } n=1\\ L(X_1,L_{n-1}(X_2\ldots X_n)) \text{ if } 1<n\end{cases}$$

Example 16.184.

$$\begin{aligned} L_4(X_1\ldots X_4;Y) &= L(X_1,L_3(X_2\ldots X_4;Y))\\ &= L(X_1,L(X_2,L_2(X_3\ldots X_4;Y)))\\ &= L(X_1,L(X_2,L(X_3,L_1(X_4\ldots X_4;Y))))\\ &= L(X_1,L(X_2,L(X_3,L(X_4,Y))))\end{aligned}$$

Further we have

$$\begin{aligned} L_2(X_1\ldots X_2;L_2(X_3\ldots X_4;Y)) &= L(X_1,L_1(X_2\ldots X_2;L_2(X_3\ldots X_4;Y)))\\ &= L(X_1,L(X_2,L_2(X_3\ldots X_4;Y)))\\ &= L(X_1,L(X_2,L(X_3,L_1(X_4\ldots X_4;Y))))\\ &= L(X_1,L(X_2,L(X_3,L(X_4,Y))))\\ &= L_4(X_1\ldots X_4;Y)\end{aligned}$$

The above example suggest the following two lemmas.

Lemma 16.185. *Let $n\in\mathbb{N}$ and $\{\langle X_i,\|\|_i\rangle\}_{i\in n+1}$ be a finite family of normed spaces then we have*

$$L_{n+1}(X_1\ldots X_{n+1};Y)=L_n(X_1\ldots X_n;L(X_{n+1},Y))$$

Proof. We prove this by induction so let

$S=\{n\in\mathbb{N}|\text{If } \{X_i,\|\|_i\}_{i\in\{1,\ldots,n+1\}} \text{ is a family of normed spaces and } \langle Y,\|\|_Y\rangle \text{ a normed space then } L_{n+1}(X_1\ldots X_{n+1};Y)=L_n(X_1\ldots X_n;(L(X_{n+1},Y))\}$

then we have:

$1\in S$. If $\{\langle X_i,\|\|_i\rangle\}_{i\in\{1,\ldots,2\}}$ is a family of normed spaces and $\langle Y,\|\|_Y\rangle$ a normed space then

$$L_2(X_1\ldots X_2;Y)=L(X_1,L_1(X_2\ldots X_2;Y))=L_1(X_1,L(X_2,Y))=L_1(X_1\ldots X_1;L(X_2,Y))$$

proving that $1\in S$.

$n\in S\Rightarrow n+1\in S$. If $\{\langle X_i,\|\|_i\rangle\}_{i\in\{1,\ldots,(n+1)+1\}}$ is a family of normed spaces and $\langle Y,\|\|_Y\rangle$ a normed space. Define $\{Y_i\}_{i\in\{1,\ldots,n+1\}}$ by $Y_i=X_{i+1}$ then we have

$$\begin{array}{rcl} L_{(n+1)+1}(X_1\ldots X_{(n+1)+1};Y) & \underset{\text{def}}{=} & L(X_1,L_{n+1}(X_2\ldots X_{(n+1)+1};Y)) \\ & = & L(X_1,L_{n+1}(Y_1\ldots Y_{n+1};Y)) \\ & \underset{n\in S}{=} & L(X_1,L_n(Y_1,\ldots,Y_n;L(Y_{n+1},Y))) \\ & = & L(X_1,L_n(X_2\ldots X_{n+1});L(X_{(n+1)+1},Y)) \\ & \underset{\text{def}}{=} & L_{n+1}(X_1\ldots X_{n+1};L(X_{(n+1)+1},Y)) \end{array}$$

proving that

$$n+1\in S$$

□

Lemma 16.186. *Let $n,m\in\mathbb{N}$, $\{\langle X_i,\|\|_i\rangle\}_{i\in\{1,\ldots,n+m\}}$ a finite family of normed spaces, $\langle Y,\|\|_Y\rangle$ a normed space then*

$$L_{n+m}(X_1\ldots X_{n+m})=L_n(X_1\ldots X_n;L_m(X_{n+1}\ldots X_{n+m};Y))$$

Proof. For $n\in\mathbb{N}$ we have two cases to consider:

$n=1$. Then

$$\begin{array}{rcl} L_{1+m}(X_1,\ldots,X_{1+m};Y) & = & L(X_1,L_m(X_2\ldots X_{1+m};Y)) \\ & = & L_1(X_1\ldots X_1;L_m(X_2\ldots X_{1+m};Y)) \end{array}$$

proving the lemma for this case.

$1<n$. We prove this case by induction on m, so let

$S_n=\{m\in\mathbb{N}|\text{If }\{\langle X_i,\|\|_i\rangle\}_{i\in\{1,\ldots,n+m\}}\text{ is a family of normed spaces then }L_{n+m}(X_1\ldots X_{n+m};Y)=L_n(X_1\ldots X_n;L_m(X_{n+1}\ldots X_{n+m};Y))\}$

then we have:

$1\in S_n$. If $\{\langle X_i,\|\|_i\rangle\}_{i\in\{1,\ldots,n+1\}}$ is a family of normed spaces then

$$L_{n+1}(X_1\ldots X_{n+1};Y)=L(X_1;L_n(X_2\ldots X_{n+1};Y))$$

proving that $1\in S_n$.

$m\in S_n\Rightarrow m+1\in S_n$. Let $\{\langle X_i,\|\|_i\rangle\}_{i\in\{1,\ldots,n+(m+1)\}}$ be a family of normed spaces. Define $\{Y_i\}_{i\in\{1,\ldots,n+m\}}$ by $Y_i=X_{i+1}$ then we have

$$\begin{array}{rc} L_{n+(m+1)}(X_1\ldots X_{n+(m+1)};Y) & = \\ L(X_1,L_{n+m}(X_2\ldots X_{n+m+1};Y)) & = \\ L(X_1,L_{n+m}(Y_1\ldots Y_{n+m};Y)) & \underset{m\in S_n}{=} \\ L(X_1,L_n(Y_1\ldots Y_n;L_m(Y_{n+1}\ldots Y_{n+m};Y))) & \underset{1<n\wedge[\text{lemma: }16.185]}{=} \\ L(X_1,L_{n-1}(Y_1\ldots Y_{n-1};L(Y_n,L_m(Y_{n+1}\ldots Y_{n+m};Y)))) & \underset{\text{def}}{=} \\ L(X_1,L_{n-1}(Y_1\ldots Y_{n-1};L_{m+1}(Y_n\ldots Y_{n+m};Y))) & = \\ L(X_1,L_{n-1}(X_2,\ldots,X_n;L_{m+1}(X_{n+1}\ldots X_{n+m+1};Y))) & \underset{\text{def}}{=} \\ L_n(X_1\ldots X_n;L_{m+1}(X_{n+1}\ldots X_{n+(m+1)};Y)) & \end{array}$$

proves that $n+1$

Mathematical induction proves the lemma for the case $1 < n$. □

We extend now the definition of $L(x_1 : \ldots : x_n)$ [see definition: 16.73] to the more general case.

Definition 16.187. *Let $n \in \mathbb{N}$, $\{\langle X_i, \|\|_i\rangle\}_{i \in \{1,\ldots,n+m\}}$ a finite family of normed spaces, $\langle Y, \|\|_Y\rangle$ a normed space, $m \in \mathbb{N}$ with $m \leqslant n$ then if $L \in L_n(X_1 \ldots X_n; Y)$ and $(x_1, \ldots, x_m) \in \prod_{i \in \{1,\ldots,m\}} X_i$ then we define*

$$L(x_1 : \ldots : x_m) = \begin{cases} L(x_1) \text{ if } m = 1 \\ (L(x_1))(x_2 : \ldots x_m) \text{ if } 1 < m \end{cases}$$

Lemma 16.188. *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, $m \in \mathbb{N}$ with $m \leqslant n$ then if $L \in L_n(X_1 \ldots X_n; Y)$ and $(x_1, \ldots, x_m) \in \prod_{i \in \{1,\ldots,m\}} X_i$ then*

$$L(x_1 : \ldots : x_m) \in \begin{cases} Y \text{ if } m = n \\ L_{n-m}(X_{m+1} \ldots X_n; Y) \text{ if } m < n \end{cases}$$

Proof. We prove this by induction, so let

$$S = \left\{ n \in \mathbb{N} | \text{If } \{\langle X_i, \|\|_i\rangle\}_{i \in \{1,\ldots,n\}} \text{ is a family of normed space, } L \in L_n(X_1 \ldots X_n; Y) \text{ then for } m \in \{1, \ldots, n\} \text{ and } (x_1, \ldots, x_m) \in \prod_{i \in \{1,\ldots,m\}} X_i \right.$$
$$\left. \text{we have } L(x_1 : \ldots : x_m) \in \left\{ \begin{array}{l} Y \text{ if } m = n \\ L_{n-m}(X_{m+1} \ldots X_n; Y) \text{ if } m < n \end{array} \right\} \right.$$

then we have:

$1 \in S$. If $m \in \{1, \ldots, 1\} = \{1\}$ then for $L \in L_1(X_1 \ldots X_1; Y) = L(X_1, Y)$ we have for $(x_1) \in \prod_{i \in \{1\}} X_i$ that $L(x_1 : \ldots : x_1) = L(x_1) \in Y$ which as $m = 1 = n$ proves that $1 \in S$.

$n \in S \Rightarrow n + 1 \in S$. Let $L \in L_{n+1}(X_1 \ldots X_{n+1}; Y)$ and $m \in \{1, \ldots, n+1\}$ then we must consider the following cases:

$m = 1$. Let $(x_1) \subseteq X^1$ then, as $L_{n+1}(X_1 \ldots X_{n+1}; Y) = L(X_1, L_n(X_2 \ldots X_{n+1}; Y))$, we have

$$L(x_1) \in L_n(X_2 \ldots X_{n+1}; Y) \underset{m=1}{=} L_{(n+1)-m}(X_{m+1} \ldots X_{n+1})$$

$m = n + 1$. Let $(x_1, \ldots, x_{n+1}) \in \prod_{i \in \{1,\ldots,n+1\}} X_i$ then, as

$$L_{n+1}(X_1 \ldots X_{n+1}; Y) = L(X_1, L_n(X_2 \ldots X_{n+1}; Y))$$

we have

$$L(x_1) \in L_n(X_2 \ldots X_{n+1}; Y)$$

Define $\{\langle Y_i, \|\|_{Y,i}\rangle\}_{i \in \{1,\ldots,n\}}$ by $\langle Y_i, \|_{Y,i}\|\rangle = \langle X_{i+1}, \|\|_{i+1}\rangle$ and $(y_1, \ldots, y_n) \in \prod_{i \in \{1,\ldots,n\}} Y_i$ by $y_i = x_{i+1}$ then $L(x_1) \in L_n(Y_1 \ldots Y_n; Y)$. As $n \in S$ we have that $L(x_1)(y_1 : \ldots : y_n) \in Y$, so $L(x_1)(x_2 : \ldots : x_{n+1}) \in Y$. Hence

$$L(x_1 : \ldots : x_m) = L(x_1 : \ldots : x_{n+1}) = L(x_1)(x_2 : \ldots : x_{n+1}) \in Y$$

$m \in \{2, \ldots, n\}$. Let $(x_1, \ldots, x_m) \in \prod_{i \in \{1,\ldots,m\}} X_i$ then as

$$L_{n+1}(X_1 \ldots X_{n+1}; Y) = L(X_1, L_n(X_2 \ldots X_{n+1}; Y))$$

we have that

$$L(x_1) \in L(X_2 \ldots X_{n+1}; Y)$$

Define $\{\langle Y_i, \|\|_{Y,i}\rangle\}_{i \in \{1,\ldots,n\}}$ by $\langle Y_i, \|_{Y,i}\|\rangle = \langle X_{i+1}, \|\|_{i+1}\rangle$ and $(y_1, \ldots, y_{m-1}) \in \prod_{i \in \{1,\ldots,n\}} Y_i$ by $y_i = x_{i+1}$ then

$$L(x_1) \in L(Y_1 \ldots Y_n; Y)$$

so, as $1 \leqslant m - 1 < n$ and $n \in S$ it follows that

$$L(x_1)(y_1 : \ldots : y_{m-1}) \in L_{n-(m-1)}(Y_{(m-1)+1} \ldots Y_n; Y)$$

hence

$$L(x_1)(x_2:\ldots:x_m)\in L_{(n+1)-m}(X_{m+1}\ldots X_{n+1};Y)$$

so that

$$L(x_1:\ldots x_m)=L(x_1)(x_2:\ldots:x_m)\in L_{(n+1)-m}(X_{m+1}\ldots X_{n+1};Y)$$

So we have proved that

$$L(x_1:\ldots:x_m)\in\begin{cases}Y \text{ if } m=n+1\\ L_{(n+1)-m}(X_{m+1}\ldots X_{n+1};Y) \text{ if } m<n\end{cases}$$

from which it follows that

$$n+1\in S$$

□

Lemma 16.189. *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, $\alpha\in\mathbb{K}$, $L_1,L_2\in L_n(X_1\ldots X_n;Y)$ then $\forall(x_1,\ldots,x_n)\in\prod_{i\in\{1,\ldots,n\}}X_i$ we have*

$$(L_1+\alpha\cdot L_2)(x_1:\ldots:x_n)=L_1(x_1:\ldots:x_n)+\alpha\cdot L_2(x_1:\ldots:x_n)$$

Proof. The proof is by induction, so let

$$S=\left\{n\in\mathbb{N}|\forall\alpha\in\mathbb{K},\forall L_1,L_2\in L_n(X;Y)\text{ we have }\forall(x_1,\ldots,x_n)\in\prod_{i\in\{1,\ldots,n\}}X_i\text{ that }(L_1+\alpha\cdot L_2)(x_1:\ldots:x_n)=L_1(x_1:\ldots:x_n)+\alpha\cdot L_2(x_1:\ldots:x_n)\right\}$$

then we have:

$1\in S$. Take $\alpha\in\mathbb{K}$, $L_1,L_2\in L_1(X_1\ldots X_1;Y)=L(X,Y)$ then we have for $(x_1)\in\prod_{i\in\{1\}}X_i$ that

$$\begin{array}{rcl}(L_1+\alpha\cdot L_2)(x_1:\ldots:x_1) & = & (L_1+\alpha\cdot L_2)(x_1)\\ & \underset{\text{pointwise definition}}{=} & L_1(x)+\alpha\cdot L_2(x_1)\\ & = & L_1(x_1:\ldots:x_1)+\alpha\cdot L_2(x_1:\ldots:x_1)\end{array}$$

proving that $1\in S$

$n\in S\Rightarrow n+1\in S$. Take $\alpha\in\mathbb{K}$, $L_1,L_2\in L_{n+1}(X_1\ldots X_{n+1};Y)=L(X_1,L_n(X_2\ldots X_{n+1};Y))$ and $(x_1,\ldots,x_n)\in\prod_{i\in\{1,\ldots,n\}}X_i$ then

$$(L_1+\alpha\cdot L_2)(x_1)\underset{\text{pointwise definition}}{=}L_1(x_1)+\alpha\cdot L_2(x_1)$$

hence

$$\begin{array}{rcl}(L_1+\alpha\cdot L_2)(x_1:\ldots:x_n) & = & ((L_1+\alpha\cdot L_2)(x_1))(x_2:\ldots:x_n)\\ & = & (L_1(x_1)+\alpha\cdot L_2(x_1))(x_2:\ldots:x_n)\end{array}\tag{16.256}$$

As $L_1(x_1),L_2(x_1)\in L_n(X_1\ldots X_n;Y)$ and $n\in S$ it follows that

$$\begin{array}{rcl}(L_1(x_1)+\alpha\cdot L_2(x_1))(x_2:\ldots:x_n) & = & (L_1(x_1))(x_2:\ldots:x_n)+\alpha\cdot(L_2(x_1))(x_2:\ldots:x_n)\\ & = & L_1(x_1:\ldots:x_n)+\alpha\cdot L_2(x_1:\ldots:x_n)\end{array}$$

which combined with [eq: 16.256] proves that

$$(L_1+\alpha\cdot L_2)(x_1:\ldots:x_n)=L_1(x_1:\ldots:x_n)+\alpha\cdot L_2(x_1:\ldots:x_n)$$

Hence we have that

$$n+1\in S$$

□

Lemma 16.190. *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, $L\in L_n(X_1\ldots X_n;Y)$ and $i\in\{1,\ldots,n\}$ we have for*

$$\Big(x_1,\ldots\underbrace{r+\alpha\cdot t}_{i}\ldots,x_n\Big)\in\prod_{i\in\{1,\ldots,n\}}X_i$$

that

$$L\Big(x_1,\ldots\underbrace{r+\alpha\cdot t}_{i}\ldots,x_n\Big)=L\Big(x_1,\ldots\underbrace{r}_{i}\ldots,x_n\Big)+\alpha\cdot L\Big(x_1,\ldots\underbrace{t}_{i}\ldots,x_n\Big)$$

Proof. We use mathematical induction on n for the proof, so let

$S=\left\{n\in\mathbb{N}|\text{If } L\in L_n(X_1\ldots X_n;Y) \text{ then } \forall i\in\{1,\ldots,n\} \text{ we have for } \left(x_1,\ldots\underbrace{r+\alpha\cdot t}_{i}\ldots,x_n\right)\in\prod_{i\in\{1,\ldots,n\}}X_i \text{ that } L\left(x_1{:}\ldots\underbrace{r+\alpha\cdot t}_{i}\ldots{:}x_n\right)=L\left(x_1{:}\ldots\underbrace{r}_{i}\ldots{:}x_n\right)+\alpha\cdot L\left(x_1,\ldots\underbrace{t}_{i}\ldots,x_n\right)\right\}$

then we have:

$1\in S$. Let $L\in L_1(X_1\ldots X_1;Y)=L(X_1,Y)$ then if $i\in\{1,\ldots,1\}=\{1\}$ we have $i=1$ and given

$$(r+\alpha\cdot t)=\left(x_1,\ldots\underbrace{r+\alpha\cdot t}_{i}\ldots,x_n\right)\in X^1$$

that

$$\begin{aligned}
L\left(x_1{:}\ldots\underbrace{r+\alpha\cdot t}_{i}\ldots{:}x_n\right) &=\\
L(r+\alpha\cdot t) &\underset{L\in L(X_1,Y)}{=}\\
L(y)+\alpha\cdot L(z) &=\\
L\left(x_1,\ldots\underbrace{r}_{i}\ldots{:}x_n\right)+\alpha\cdot L\left(x_1{:}\ldots\underbrace{s}_{i}\ldots{:}x_n\right)
\end{aligned}$$

proving that $1\in S$.

$n\in S\Rightarrow n+1\in S$. Let $L\in L_{n+1}(X_1\ldots X_{n+1};Y)=L(X_1,L_n(X_2\ldots X_{n+1};Y))$ and take $i\in\{1,\ldots,n+1\}$ then we have either:

$i=1$. Then for $\left(x_1,\ldots\underbrace{r+\alpha\cdot t}_{i}\ldots,x_{n+1}\right)=(r+\alpha\cdot t,x_2,\ldots,x_{n+1})$ we have

$$\begin{aligned}
L\left(x_1,\ldots\underbrace{r+\alpha\cdot t}_{i}\ldots{:}x_{n+1}\right) &=\\
L(r+\alpha\cdot t{:}\,x_2{:}\ldots{:}x_{n+1}) &=\\
(L(r+\alpha\cdot t))(x_2{:}\ldots{:}x_{n+1}) &=\\
(L(r)+\alpha\cdot L(t))(x_2{:}\ldots{:}x_{n+1}) &\underset{[\text{lemma: } 16.79]}{=}\\
(L(r))(x_2{:}\ldots{:}x_{n+1})+\alpha\cdot L(t)(x_2{:}\ldots{:}x_{n+1}) &=\\
L(r{:}\,x_2{:}\ldots{:}x_{n+1})+\alpha\cdot L(t{:}\,x_2{:}\ldots{:}x_{n+1}) &=\\
L\left(x_1{:}\ldots\underbrace{r}_{i}\ldots{:}x_{n+1}\right)+\alpha\cdot L\left(x_1{:}\ldots\underbrace{t}_{i}\ldots{:}x_{n+1}\right)
\end{aligned}$$

proving

$$n+1\in S$$

$i\in\{2,\ldots,n+1\}$. Then for $\left(x_1,\ldots\underbrace{r+\alpha\cdot t}_{i}\ldots,x_{n+1}\right)\in\prod_{i\in\{1,\ldots,n\}}X_i$ we have $\left(x_2,\ldots\underbrace{r+\alpha\cdot t}_{i-1}\ldots,x_{n+1}\right)\in\prod_{i\in\{2,\ldots,n+1\}}X_i$ which as $L(x_1)\in L_n(X_2\ldots X_{n+1};Y)$ and $n\in S$ results in

$$\begin{aligned}
L\left(x_1,\ldots\underbrace{r+\alpha\cdot t}_{i}\ldots,x_{n+1}\right) &=\\
(L(x_1))\left(x_2{:}\ldots\underbrace{r+\alpha\cdot t}_{i-1}\ldots{:}x_n\right) &\\
(L(x_1))\left(x_2{:}\ldots\underbrace{r}_{i-1}\ldots{:}x_n\right)+\alpha\cdot(L(x_1))\left(x_2{:}\ldots\underbrace{t}_{i-1}\ldots{:}x_n\right) &=\\
L\left(x_1{:}\ldots\underbrace{r}_{i}\ldots{:}x_n\right)+\alpha\cdot L(x_1{:}\ldots\underbrace{t}_{i}\ldots{:}x_n) &=
\end{aligned}$$

so that in this case we have also

$$n+1\in S$$

So in all cases we have

$$n+1\in S$$

□

Lemma 16.191. *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, $L\in L_n(X_1\ldots X_n;Y)$ and $(x_1,\ldots,x_n)\in\prod_{i\in\{1,\ldots,n\}}X_i$ then we have*

$$\|L(x_1:\ldots:x_n)\|_Y\leqslant\|L\|_{L_n(X_1\ldots X_n;Y)}\cdot\prod_{i=1}^n\|x_i\|_i$$

Proof. We prove this by induction, so define

$$S=\left\{n\in\mathbb{N}|\text{If } L\in L_n(X_1\ldots\text{Xn};Y)\text{ then }\forall(x_1,\ldots,x_n)\in\prod_{i\in\{1,\ldots,n\}}X_i\text{ we have }\|L(x_1:\ldots:x_n)\|_Y\leqslant\|L\|_{L_n(X;Y)}\cdot\prod_{i=1}^n\|x_i\|_i\right\}$$

then we have:

$1\in S$. If $L\in L_1(X_1\ldots X_1;Y)=L(X_1,Y)$ then $\forall(x_1)\in\prod_{i\in\{1\}}X_i$ we have

$$\begin{aligned}\|L(x_1:\ldots:x_1)\|_Y &= \|L(x_1)\|_Y\\ &\leqslant \|L\|_{L(X_1,Y)}\cdot\|x_1\|_1\\ &= \|L\|_{L_1(X_1\ldots X_1;Y)}\cdot\|x_1\|_1\\ &= \|L\|_{L_1(X\ldots X_1;Y)}\cdot\prod_{i=1}^1\|x_i\|_i\end{aligned}$$

proving that $1\in S$.

$n\in S\Rightarrow n+1\in S$. Let $L\in L_{n+1}(X_1\ldots X_{n+1};Y)=L(X_1,L_n(X_2\ldots X_{n+1};Y))$ and $(x_1,\ldots,x_{n+1})\in\prod_{i\in\{1,\ldots,n+1\}}X_i$ then, as $L(x_1)\in L_n(X_1\ldots X_1;Y)$, $(x_2,\ldots,x_{n+1})\in\prod_{i\in\{2,\ldots,n+1\}}X_i$ and $n\in S$, we have

$$\begin{aligned}\|L(x_1)(x_2:\ldots:x_{n+1})\|_Y &\leqslant \|L(x_1)\|_{L_n(X_2\ldots X_{n+1};Y)}\cdot\prod_{i=1}^n\|x_{i+1}\|_{i+1}\\ &= \|L(x_i)\|_{L_n(X_2\ldots X_{n+1};Y)}\cdot\prod_{i=2}^{n+1}\|x_i\|_X\end{aligned}\tag{16.257}$$

Next we have

$$\begin{aligned}\|L(x_1:\ldots:x_{n+1})\|_Y &= \|(L(x_1))(x_2:\ldots:x_{n+1})\|_Y\\ &\leqslant_{[\text{eq: 16.257}]} \|L(x_i)\|_{L_n(X_2\ldots X_{n+1};Y)}\cdot\prod_{i=2}^{n+1}\|x_i\|_X\\ &\leqslant \|L\|_{L_{n+1}(X_1\ldots X_{n+1};Y)}\cdot\|x_1\|_X\cdot\prod_{i=2}^{n+1}\|x_i\|_X\\ &\leqslant \|L\|_{L_{n+1}(X_1\ldots X_{n+1};Y)}\cdot\prod_{i=1}^{n+1}\|x_i\|_X\end{aligned}$$

proving that

$$n+1\in S$$

□

Lemma 16.192. *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, $L_1,L_2\in L_n(X_1\ldots X_n;Y)$ such that $\forall(x_1,\ldots,x_n)\in\prod_{i\in\{1,\ldots,n\}}X_i$ we have*

$$L_1(x_1:\ldots:x_n)=L_2(x_1:\ldots x_n)$$

then

$$L_1=L_2$$

Proof. We use induction to prove this, so define

$$S = \{n \in \mathbb{N} | \text{If } L_1, L_2 \in L_n(X_1 \ldots X_n; Y) \text{ satisfies } \forall(x_1, \ldots, x_n)\ L_1(x_1 : \ldots : x_n) = L_2(x_1 : \ldots : x_n) \text{ then } L_1 = L_2\}$$

then we have:

$1 \in S$. Let $L_1, L_2 \in L_1(X_1; Y) = L(X_1, Y)$ satisfies that $\forall(x_1) \in X^1$ we have that $L_1(x_1 : \ldots : x_1) = L_2(x_1 : \ldots : x_1)$. Then $\forall x \in X$ we have for $y = (x) \in X^1$ [so that $y_1 = x$] that

$$L_1(x) = L_1(y_1) = L_1(y_1 : \ldots : y_1) = L_2(y_1 : \ldots : y_1) = L_2(y_1) = L_2(x)$$

proving that $L_1 = L_2$. Hence $1 \in S$.

$n \in S \Rightarrow n + 1 \in S$. Let $L_1, L_2 \in L_{n+1}(X_1 \ldots X_{n+1}; Y) = L(X_1, L_n(X_2 \ldots X_{n+1}; Y))$ such that

$$\forall x = (x_1, \ldots, x_{n+1}) \in \prod_{i \in \{1, \ldots, n+1\}} X_i\ L_1(x_1 : \ldots : x_{n+1}) = L_2(x_1 : \ldots : x_{n+1})$$

Let $y \in X$ then $L_1(y), L_2(y) \in L_n(X; Y)$ and $\forall(x_1, \ldots, x_n) \in X^n$ we have

$$L_1(y)(x_1 : \ldots : x_n) = L_1(y : x_1 : \ldots : x_n) = L_2(y : x_1 : \ldots : x_n) = L_2(y)(x_1 : \ldots : x_n)$$

which as $n \in S$ proves that $L_1(y) = L_2(y)$. As y is chosen arbitrary we conclude that $L_1 = L_2$. proving

$$n + 1 \in S$$

□

Lemma 16.193. *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. If $\forall(x_1, \ldots, x_n) \in \prod_{i \in \{1, \ldots, n\}} X_i$ we have*

$$\|L(x_1 : \ldots : x)\|_Y \leqslant M \cdot \prod_{i=1}^{n} \|x_i\|_i$$

then it follows that

$$\|L\|_{L_n(X_1 \ldots X_n; Y)} \leqslant M$$

Proof. We prove this by induction, so let

$$S = \left\{ n \in \mathbb{N} | \text{If } L \in L_n(X; Y) \text{ satisfies } \forall(x_1, \ldots, x_n) \in X^n\ \|L(x_1 : \ldots : x_n)\|_Y \leqslant M \cdot \prod_{i=1}^{n} \|x_i\|_X \text{ then } \|L\|_{L_n(X_1 \ldots X_n; Y)} \leqslant M \right\}$$

then we have:

$1 \in S$. Let $L \in L_1(X_1 \ldots X_1; Y) = L(X_1, Y)$ such that $\forall(x_1) \in X^n$ we have $\|L(x_1 : \ldots : x_1)\|_Y \leqslant M \cdot \prod_{i=1}^{1} \|x_i\|_X$. Then given $y \in X$ we have for $x = (y) \in \prod_{i \in \{1\}} X_i$ that

$$\|L(y)\|_Y = \|L(x_1)\|_Y = \|L(x_1 : \ldots : x_1)\|_Y \leqslant M \cdot \prod_{i=1}^{1} \|x_i\|_i = M \cdot \|y\|_X$$

so that by [definition: 14.178] $\|L\|_{L_1(X;Y)} = \|L\|_{L(X,Y)} \leqslant M$ proving that $1 \in S$.

$n \in S \Rightarrow n + 1 \in S$. Let $L \in L_{n+1}(X_1 \ldots X_{n+1}; Y) = L(X_1, L_n(X_2 \ldots X_n; Y))$ such that $\forall(x_1, \ldots, x_{n+1}) \in X^n$ we have

$$\|L(x_1 : \ldots : x_n)\|_Y \leqslant M \cdot \prod_{i=1}^{n+1} \|x_i\|_i$$

Fix $x_1 \in X$ then $\forall(x_2, \ldots, x_{n+1}) \in \prod_{i \in \{2, \ldots, n+1\}} X_i$ we have

$$\|(L(x_1))(x_2 : \ldots : x_{n+1})\|_Y = \|L(x_1 : x_2 : \ldots : x_{n+1})\|_Y \leqslant M \cdot \|x\|_1 \cdot \prod_{i=2}^{n+1} \|x_i\|_i = (M \cdot \|x_1\|_1) \cdot \prod_{i=2}^{n+1} \|x_i\|_i$$

As $n \in S$ and $L(x) \in L_n(X_2 \ldots X_{n+1}; Y)$ it follows that $\|L(x)\|_{L_n(X_2 \ldots X_{n+1}; Y)} \leqslant M \cdot \|x_1\|_1$. Hence by [definition: 14.178] it follows that $\|L\|_{L_{n+1}(X_1 \ldots X_{n+1}; Y)} \leqslant M$ proving

$$n + 1 \in S$$

□

Lemma 16.194. *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, $L \in L(X_1\ldots X_n; Y)$ then there exist a $K \in L_n(X_1\ldots X_n; L)$ such that $\forall (x_1,\ldots,x_n) \in \prod_{i\in\{1,\ldots,n\}} X_i$*

$$K(x_1:\ldots:x_n) = L(x_1,\ldots,x_n)$$

Proof. We use induction to prove this, so let

$$S = \left\{n \in \mathbb{N} | \text{If } L \in L(X_1\ldots X_n; Y) \text{ then there exist a } K \in L_n(X_1\ldots X_n; Y) \text{ such that } \forall (x_1,\ldots,x_n) \in \prod_{i\in\{1,\ldots,n\}} X_i \text{ we have } K(x_1:\ldots: x_n) = L(x_1,\ldots,x_n)\right\}$$

then we have:

$1 \in S$. If $L \in L(X_1\ldots X_1; Y)$ then as $L_1(X_1\ldots X_1; Y) = L(X_1, Y) = L_n(X_1\ldots X_1; Y)$ we have that $L \in L_n(X_1\ldots X_1, Y)$ so if we take $K = L$ then we have $K \in L_1(X_1\ldots X_1, Y)$ and for $(x_1) \in \prod_{i\in\{1\}} X_i$ we have $K(x_1:\ldots:x_1) = L(x_1) = L(x_1,\ldots,x_1)$. So we have that $1 \in S$.

$n \in S \Rightarrow n+1 \in S$. Let $L \in L(X_1\ldots X_{n+1}; Y)$. Take $x \in X$ and define

$$L_x: \prod_{i\in\{2,\ldots,n+1\}} X_i \to Y \text{ by } L_x(x_1,\ldots,x_n) = L(x, x_1,\ldots,x_n)$$

If $i \in \{1,\ldots,n\}$ and $\left(x_1,\ldots \underbrace{r+\alpha\cdot t}_{i} \ldots, x_n\right) \in \prod_{i\in\{2,\ldots,n+1\}} X_i$ then

$$\begin{aligned}
L_x\left(x_1,\ldots \underbrace{r+\alpha\cdot t}_{i} \ldots, x_n\right) &= \\
L\left(x, x_1,\ldots \underbrace{r+\alpha\cdot t}_{i+1} \ldots, x_n\right) &= \\
L\left(x, x_1,\ldots \underbrace{r}_{i+1} \ldots, x_n\right) + \alpha\cdot L\left(x, x_1,\ldots \underbrace{t}_{i+1} \ldots, x_n\right) &= \\
L_x\left(x_1,\ldots \underbrace{r}_{i} \ldots, x_n\right) + \alpha\cdot L_x\left(x_1,\ldots \underbrace{t}_{i} \ldots, x_n\right)
\end{aligned}$$

proving that

$$L_x \in \mathrm{Hom}(X_2\ldots X_{n+1}; Y) \tag{16.258}$$

Let $(x_2,\ldots,x_n) \in \prod_{i\in\{2,\ldots,n+1\}} X_i$ then for $(y_1,\ldots,y_{n+1})$ where $y_i = \left\{\begin{smallmatrix} x \text{ if } i=1 \\ x_{i-1} \text{ if } i\in\{2,\ldots,n+1\}\end{smallmatrix}\right.$ we have

$$\begin{aligned}
\|L_x(x_1,\ldots,x_n)\|_Y &= L(x, x_1,\ldots,x_n) \\
&= L(y_1,\ldots,y_{n+1}) \\
&\leqslant_{L\in L(X_1\ldots X_{n+1};Y)} \|L\|_{L(X_1\ldots X_{n+1};Y)} \cdot \prod_{i=1}^{n+1} \|y_i\|_i \\
&= \|L\|_{L(X_1\ldots X_{n+1};Y)} \cdot \|y_1\|_1 \cdot \prod_{i=2}^{n+1} \|y_i\|_i \\
&= \|L\|_{L(X_1\ldots X_{n+1};Y)} \cdot \|x\|_1 \cdot \prod_{i=2}^{n+1} \|x_{i-1}\|_i \\
&= (\|x\|_X \cdot \|L\|_{L^{n+1}(X;Y)}) \cdot \prod_{i=1}^{n} \|x_i\|_{i+1}
\end{aligned}$$

proving by [theorem: 14.187] and [eq: 16.258] that L_x is continuous. Hence

$$L_x \in L(X_2\ldots X_{n+1}; Y) \tag{16.259}$$

As $n \in S$ there exist a $K_x \in L_n(X_2\ldots X_{n+1}; Y)$ such that $\forall (x_1,\ldots,x_n) \in \prod_{i\in\{2,\ldots,n+1\}} X_i$ we have

$$K_x(x_1:\ldots:x_n) = L_x(x_1,\ldots,x_n) = L(x, x_1,\ldots,x_n) \tag{16.260}$$

This allows us to define

$$K: X_1 \rightarrow L_n(X_2 \ldots . X_{n+1}; Y) \text{ where } K(x) = K_x$$

then by [eq: 16.260] we have that for $(x_1, \ldots, x_{n+1}) \in \prod_{i \in \{1, \ldots, n+1\}} X_i$

$$K(x_1)(x_2: \ldots : x_{n+1}) = K_{x_1}(x_2: \ldots : x_1) = L(x_1, \ldots, x_{n+1}) \tag{16.261}$$

If $x, y \in X_1$ and $\alpha \in \mathbb{K}$ then we have $\forall (x_1, \ldots, x_n) \in \prod_{i \in \{2, \ldots, n+1\}} X_i$ that

$$\begin{aligned} K(x + \alpha \cdot y)(x_1: \ldots : x_n) &\underset{[\text{eq: } 16.261]}{=} L(x + \alpha \cdot y, x_1, \ldots, x_n) \\ &\underset{L \in L(X_1 \ldots X_{n+1}; Y)}{=} L(x, x_1, \ldots, x_n) + \alpha \cdot L(y, x_1, \ldots, x_n) \\ &\underset{[\text{eq: } 16.261]}{=} K(x)(x_1: \ldots : x_n) + \alpha \cdot K(y)(x_1: \ldots : x_n) \\ &\underset{[\text{lemma: } 16.79]}{=} (K(x) + \alpha \cdot K(y))(x_1: \ldots : x_n) \end{aligned}$$

which by [lemma: 16.82] proves that

$$K(x + \alpha \cdot y) = K(x) + \alpha \cdot K(y)$$

Hence we have that

$$K \in \mathrm{Hom}(X, L_n(X_1 \ldots X_{n+1}; Y)) \tag{16.262}$$

Let $x_1 \in X_1$ then we have for $(x_2, \ldots, x_{n+1}) \in \prod_{i \in \{2, \ldots, n+1\}} X_i$ that

$$\begin{aligned} \|K(x_1)(x_2: \ldots : x_{n+1})\|_Y &= \|K_{x_1}(x_2: \ldots : x_{n+1})\|_Y \\ &\underset{[\text{eq: } 16.261]}{=} \|L(x_1, x_2, \ldots, x_{n+1})\|_Y \\ &\leqslant_{L \in L(X_1 \ldots X_{n+1}; Y)} \|L\|_{L(X_1 \ldots X_{n+1}; Y)} \cdot \prod_{i=1}^{n+1} \|x_i\|_i \\ &= \|L\|_{L(X_1 \ldots X_{n+1}; Y)} \cdot \|x_1\|_1 \cdot \prod_{i=2}^{n+1} \|x_i\|_i \\ &= (\|L\|_{L(X_1 \ldots X_{n+1}; Y)} \cdot \|x_1\|_1) \cdot \prod_{i=2}^{n+1} \|x_i\|_i \end{aligned}$$

which as $K(x_1) \in L_n(X_2 \ldots X_n; Y)$ proves using [lemma: 16.193] that

$$\|K(x_1)\|_{L_n(X_2 \ldots X_{n+1}; Y)} \leqslant \|L\|_{L(X_1 \ldots X_{n+1}; Y)} \cdot \|x_1\|_X$$

Hence using [eq: 16.262] together with [theorem: 14.174] proves that K is continuous or

$$K \in L(X_1, L_n(X_2 \ldots X_{n+1}; Y)) = L_{n+1}(X_1 \ldots X_{n+1}; Y)$$

further

$$K(x_1: \ldots : x_{n+1}) = K(x_1)(x_2: \ldots : x_{n+1}) \underset{[\text{eq: } 16.261]}{=} L(x_1, \ldots, x_{n+1})$$

The above proves that

$$n + 1 \in S$$

□

Theorem 16.195. *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*

$$\mathcal{I}_{X_1 \ldots X_n, Y}: L_n(X_1 \ldots X_n) \rightarrow L(X_1 \ldots X_n; Y) \text{ where } \mathcal{I}_{X_1 \ldots X_n, Y}(L): \prod_{i \in \{1, \ldots, n\}} X \rightarrow Y \text{ is defined by}$$

$$(\mathcal{I}_{X_1 \ldots X_n, Y}(L))(x_1, \ldots, x_m) = L(x_1: \ldots x_m)$$

is a linear isometric isomorphism. In other words

1. $\forall L \in L_n(X_1 \ldots X_n; Y)$ $\mathcal{I}_{X_1 \ldots X_n, Y}(L) \in L(X_1 \ldots X_n; Y)$ *so that*

$$\mathcal{I}_{X_1 \ldots X_n, Y}: L_n(X_1 \ldots X_n; Y) \rightarrow L(X_1 \ldots X_n; Y)$$

is a function

2. $\mathcal{I}_{X_1 \ldots X_n,Y}: L_n(X_1 \ldots X_n;Y) \rightarrow L(X_1 \ldots X_n;Y)$ *is a bijection.*
3. *For* $L \in L(X_1 \ldots X_n;Y)$ *we have* $\forall (x_1,\ldots,x_n) \in \prod_{i \in \{1,\ldots,n\}} X_i$ *that*
$$(\mathcal{I}_{X_1 \ldots X_n,Y})^{-1}(L)(x_1: \ldots :x_n) = L(x_1,\ldots,x_n)$$
4. $\forall L_1, L_2 \in L_n(X_1 \ldots X_n;Y)$ *and* $\alpha \in \mathbb{K}$ *we have*
$$\mathcal{I}_{X_1 \ldots X_n,Y}(L_1 + \alpha \cdot L_2) = \mathcal{I}_{X_1 \ldots X_n,Y}(L_1) + \alpha \cdot \mathcal{I}_{X_1 \ldots X_n,Y}(L_2)$$
5. $\forall L \in L_n(X_1 \ldots X_n:Y)$ *we have* $\|L\|_{L_n(X_1 \ldots X_n;Y)} = \|\mathcal{I}_{X_1 \ldots X_n,Y}(L)\|_{L(X_1 \ldots X_n;Y)}$

***Note* 16.196.** *Using [theorem: 14.119] it follows that*
$$(\mathcal{I}_{X_1 \ldots X_n,Y})^{-1}: L(X_1 \ldots X_n;Y) \rightarrow L_n(X_1 \ldots X_n;Y)$$
is a linear isometric isomorphism. In other words:

1. $(\mathcal{I}_{X_1 \ldots X_n,Y})^{-1}$ *is a bijection.*
2. $\forall L_1, L_2 \in L(X_1 \ldots X_n;Y)$ *and* $\alpha \in \mathbb{K}$ *we have*
$$(\mathcal{I}_{X_1 \ldots X_n,Y})^{-1}(L_1 + \alpha \cdot L_2) = (\mathcal{I}_{X_1 \ldots X_n,Y})^{-1}(L_1) + \alpha \cdot (\mathcal{I}_{X_1 \ldots X_n,Y})^{-1}(L_2)$$
3. $\forall L \in L(X_1,\ldots,X_n)$ *we have* $\|L\|_{L(X_1 \ldots X_n;Y)} = \|(\mathcal{I}_{X_1 \ldots X_n,Y})^{-1}(L)\|_{L_n(X_1 \ldots X_n;Y)}$

Proof.

1. Let $L \in L_n(X_1 \ldots X_n;Y)$ then for $i \in \{1,\ldots,n\}$ and $\left(x_1,\ldots \underbrace{r + \alpha \cdot r}_{i} \ldots, x_n\right) \in \prod_{i \in \{1,\ldots,n\}} X_i$ we have
$$\begin{aligned}
\mathcal{I}_{X_1 \ldots X_n,Y}(L)\left(x_1,\ldots \underbrace{r + \alpha \cdot t}_{i} \ldots, x_n\right) &= \\
L\left(x_1: \ldots \underbrace{r + \alpha \cdot t}_{i} \ldots :x_n\right) &\underset{\text{[lemma: 16.190]}}{=} \\
L\left(x_1: \ldots \underbrace{r}_{i} \ldots :x_n\right) + \alpha \cdot L\left(x_1: \ldots \underbrace{t}_{i} \ldots :x_n\right) &= \\
\mathcal{I}_{X_1 \ldots X_n,Y}(L)\left(x_1,\ldots \underbrace{r}_{i} \ldots x_n\right) + \alpha \cdot \mathcal{I}_{X_1 \ldots X_n,Y}(L)\left(x_1,\ldots \underbrace{s}_{i} \ldots x_n\right) &
\end{aligned}$$
proving that
$$\mathcal{I}_{X_1 \ldots X_n,Y}(L) \in \mathrm{Hom}(X_1,\ldots,X_n;Y) \tag{16.263}$$
Further if $(x_1,\ldots,x_n) \in \prod_{i \in \{1,\ldots,n\}} X_i$ then we have:
$$\begin{aligned}
\|\mathcal{I}_{X_1 \ldots X_n,Y}(L)(x_1,\ldots,x_n)\|_Y &= \|L(x_1: \ldots :x_n)\|_Y \\
&\leqslant_{\text{[lemma: 16.191]}} \|L\|_{L_n(X_1 \ldots X_n;Y)} \cdot \prod_{i=1}^{n} \|x_i\|_i
\end{aligned}$$
Using [theorems: 14.187 and 14.193] it follows that
$$\mathcal{I}_{X_1 \ldots X_n,Y}(L) \text{ is continuous and } \|\mathcal{I}_{X_1 \ldots X_n,Y}(L)\|_{L(X_1,\ldots,X_n;Y)} \leqslant \|L\|_{L_n(X_1 \ldots X_n;Y)} \tag{16.264}$$
Combining the above with [eq: 16.263] we have
$$\|L\|_{L_n(X_1 \ldots X_n;Y)} \in L(X_1,\ldots,X_n;Y)$$
2. For bijectivity we have to prove:

   injectivity. If $L_1, L_2 \in L_n(X_1 \ldots X_n;Y)$ is such that $\mathcal{I}_{X_1 \ldots X_n,Y}(L_1) = \mathcal{I}_{X_1 \ldots X_n,Y}(L_2)$ then we have $\forall (x_1,\ldots,x_n) \in \prod_{i \in \{1,\ldots,n\}} X_i$ that
$$L_1(x_1: \ldots :x_n) = \mathcal{I}_{X_1 \ldots X_n,Y}(L_1) = \mathcal{I}_{X_1 \ldots X_n,Y}(L_2) = L_2(x_1: \ldots :x_n)$$

Hence by [lemma: 16.192] it follows that $L_1 = L_2$.

surjectivity. Let $L \in L(X_1 \ldots X_n; Y)$ then using [lemma: 16.194] there exist a $K \in L_n(X_1 \ldots X_n; Y)$ such that $\forall (x_1, \ldots, x_n) \in \prod_{i \in \{1, \ldots, n\}} X_i$ we have

$$K(x_1 : \ldots : x_n) = L(x_1, \ldots, x_n)$$

Hence

$$\mathcal{I}_{X_1 \ldots X_n, Y}(K)(x_1, \ldots, x_n) = K(x_1 : \ldots : x_n) = L(x_1, \ldots, x_n)$$

which proves that

$$\mathcal{I}_{X_1 \ldots X_n, Y}(K) = L$$

so that $\mathcal{I}_{X_1 \ldots X_n, Y}$ is surjective.

3. Let $L \in L(X_1 \ldots X_n; Y)$ and take $K = (\mathcal{I}_{X_1 \ldots X_n, Y})^{-1}(L) \in L_n(X_1 \ldots X_n; Y)$ then we have that $L = \mathcal{I}_{X_1 \ldots X_n, Y}(K)$ so that $K(x_1 : \ldots x_n) = \mathcal{I}_{X_1 \ldots X_n, Y}(K)(x_1, \ldots, x_n) = L(x_1, \ldots, x_n)$.
4. Let $L_1, L_2 \in L_n(X_1 \ldots X_n; Y)$ and $\alpha \in \mathbb{K}$ then $\forall (x_1, \ldots, x_n) \in \prod_{i \in \{1, \ldots, n\}} X_i$ then we have:

$$\begin{array}{rcl} \mathcal{I}_{X_1 \ldots X_n, Y}(L_1 + \alpha \cdot L_2)(x_1, \ldots, x_n) & = & \\ (L_1 + \alpha \cdot L_2)(x_1 : \ldots : x_n) & \underset{[\text{lemma: } 16.189]}{=} & \\ L_1(x_1 : \ldots : x_n) + \alpha \cdot L_2(x_1 : \ldots : x_n) & = & \\ \mathcal{I}_{X_1 \ldots X_n, Y}(L_1)(x_1, \ldots, x_n) + \alpha \cdot \mathcal{I}_{X_1 \ldots X_n, Y}(L_2)(x_1, \ldots, x_n) & = & \\ (\mathcal{I}_{X_1 \ldots X_n, Y}(L_1) + \alpha \cdot \mathcal{I}_{X_1 \ldots X_n, Y}(L_2))(x_1, \ldots, x_n) & & \end{array}$$

proving that

$$\mathcal{I}_{X_1 \ldots X_n, Y}(L_1 + \alpha \cdot L_2) = \mathcal{I}_{X_1 \ldots X_n, Y}(L_1) + \alpha \cdot \mathcal{I}_{X_1 \ldots X_n, Y}(L_2)$$

5. Let $L \in L_n(X_1 \ldots X_n; Y)$ then by [eq: 16.264] we have

$$\|\mathcal{I}_{X_1 \ldots X_n, Y}(L)\|_{L(X_1 \ldots X_n; Y)} \leqslant \|L\|_{L_n(X_1 \ldots X_n; Y)} \tag{16.265}$$

For the opposite inequality, let $(x_1, \ldots, x_n) \in \prod_{i \in \{1, \ldots, n\}} X_i$ then

$$\begin{array}{rcl} \|L(x_1 : \ldots : x_n)\|_Y & = & \|\mathcal{I}_{X_1 \ldots X_n, Y}(L)(x_1, \ldots, x_n)\|_Y \\ & \leqslant & \|\mathcal{I}_{X_1 \ldots X_n, Y}(L)\|_{L(X_1 \ldots X_n; Y)} \cdot \prod_{i=1}^{n} \|x_i\|_i \end{array}$$

so that by [lemma: 16.193]

$$\|L\|_{L_n(X_1 \ldots X_n; Y)} \leqslant \|\mathcal{I}_{X_1 \ldots X_n, Y}(L)\|_{L(X_1 \ldots X_n; Y)}$$

which combined with [eq: 16.265] proves that

$$\|L\|_{L_n(X_1 \ldots X_n; Y)} = \|\mathcal{I}_{X_1 \ldots X_n, Y}(L)\|_{L(X_1 \ldots X_n; Y)}$$

□

16.5.2 Higher Order Partial Differentiation

Definition 16.197. *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 space, $\langle X, \|\|_{\max} \rangle$ the normed space based on the product $X = \prod_{i \in \{1, \ldots, n\}} X_i$ with the maximum norm $\|\|_{\max}$ [see definition: 14.94], U a open set in X, $x \in U$, $m \in \mathbb{N}$, $\{i_j\}_{j \in \{1, \ldots, m\}} \subseteq \{1, \ldots, n\}$ and $f: U \rightarrow Y$ a function. Then f has a partial differential $D^{[m]}_{x, i_m \ldots i_1} f$ of order m **at x** if for m we have*

$m = 1$. *f is i_1-partial differentiable at x [see definition: 16.43] and*

$$D^{[1]}_{x, i_1 \ldots i_1} f = D_{x, i_1} f \in L(X_i, Y) = L_1(X_1 \ldots X_1; Y)$$

$1 < m$. *There exist a open set V with $x \in V \subseteq U$ such that $\forall x \in V$ we have that f has a partial derivative $D^{[m-1]}_{x, i_{m-1} \ldots i_1} f$ at x and that*

$$D^{[m-1]}_{i_{m-1} \ldots i_1} f: V \rightarrow L_{m-1}(X_{i_m} \ldots X_{i_1}; Y) \text{ defined by } D^{[m-1]}_{i_{m-1} \ldots i_1} f(x) = D^{[m-1]}_{x, i_{m-1} \ldots i_m}$$

is i_m-partial differentiable at x. $D^{[m]}_{x,i_m\ldots i_1}f$ is then be defined by

$$D^{[m]}_{x,i_m\ldots i_1}f = D_{x,i_m}(D^{m-1}_{x,i_{m-1}\ldots i_1}) \in L(X_{i_m}, L_{m-1}(X_{m-1}\ldots X_1;Y)) = L(X_{i_m}\ldots X_{i_1};Y)$$

Note 16.198. *In case $1<m$ we have to ensure that $D^{[m]}_{x,i_m\ldots i_1}f$ is independent of the choice of X. Luckily this is ensured by the fact that partial differentiability is a local property [see theorem: 16.45].*

Definition 16.199. *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 space, $\langle X, \|\|_{\max}\rangle$ the normed space based on the product $X=\prod_{i\in\{1,\ldots,n\}} X_i$ with the maximum norm $\|\|_{\max}$ [see definition: 14.94], U a open set in X, $x\in U$, $m\in\mathbb{N}$, $\{i_j\}_{j\in\{1,\ldots,m\}}\subseteq\{1,\ldots,n\}$ and $f: U\rightarrow Y$ a function then f has a **$(i_m\ldots i_1)$-partial differential $D^{[m]}_{i_m\ldots i_1}f$ of order m on U** if $\forall x\in U$ f has partial differential $D^{[m]}_{x,i_m\ldots i_1}f$ of order m at x.*

Just as Fréchet differentiability implies partial differentiability, higher order differentiability implies higher order partial differentiability. To prove this we first need some lemmas.

Next we prove that 1-times differentiability implies the existence of partial differentials of order 1, which is rather trivial as Fréchet differentiability implies partial differentiation.

Lemma 16.200. *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 space, $\langle X, \|\|_{\max}\rangle$ the normed space based on the product $X=\prod_{i\in\{1,\ldots,n\}} X_i$ with the maximum norm $\|\|_{\max}$ [see definition: 14.94], U a open set in X, $x\in U$ and $f: U\rightarrow Y$ a function that is 1-times differentiable at x then $\forall i\in\{1,\ldots,n\}$ f has a partial derivative $D^{[1]}_{x,i\ldots i}f$ at x of order 1. Further*

$$D^{[1]}_{x,i\ldots i}f = D^{[1]}_x f\circ I^{[i,o]}$$

and

$$D^{[1]}_x f = \sum_{i=1}^n D^{[1]}_{x,i\ldots i}f\circ\pi_i$$

Proof. Let $x\in U$ then as f is 1-times differentiable at x it follows that f is Fréchet differentiable at x. Then by [theorem: 16.47] f has a i-partial differential $D_{x,i}f$ at x $\forall i\in\{1,\ldots,n\}$ and

$$D_{x,i}f = D_x f\circ I^{[i,0]},\ D_x f = \sum_{i=1}^n D_{x,i}f\circ\pi_i$$

Hence by definition we have that f has a partial derivative $D^{[1]}_{x,i}f$ of order 1 at x and $D^{[1]}_{x,i\ldots i}f = D_{x,i}f$ so that

$$D^{[1]}_{x,i\ldots i}f = D^{[1]}_x f\circ I^{[i,0]},\ D^{[1]}_x f = D^{[1]}_x f = \sum_{i=1}^n {D^{[1]}}_{x,i\ldots i}f\circ\pi_i$$ □

For the general case we need first a little lemma.

Lemma 16.201. *Let $n\in\mathbb{N}\setminus\{1\}$, $m\in\mathbb{N}$, $\{\langle X_i, \|\|_i\rangle\}_{i\in\{1,\ldots,n\}}$ a finite family of normed spaces, $\langle Y, \|\|_Y\rangle$ a normed space, $\langle X, \|\|_{\max}\rangle$ the normed space based on the product $X=\prod_{i\in\{1,\ldots,n\}} X_i$ with the maximum norm $\|\|_{\max}$ [see definition: 14.94] and $\{i_j\}_{j\in\{1,\ldots,m\}}\subseteq\{1,\ldots, n\}$ then there exist a*

$$\Phi\in L(L_m(X;Y), L_m(X_{i_m}\ldots X_{i_1};Y))$$

such that for every $L\in L_m(X_{i_m}\ldots X_{i_1};Y)$

$$\forall(h_1,\ldots,h_m)\in\prod_{j\in\{1,\ldots,m\}} X_{i_{m-j+1}}$$

we have

$$\Phi(L)(h_1:\ldots:h_m) = L(I^{[i_m,0]}(h_1):\ldots:I^{[i_1,0]}(h_m))$$

Proof. Let $L\in L_m(X;Y)$

$$\Psi_L: \prod_{j\in\{1,\ldots,m\}} X_{m-j+1}\rightarrow Y \text{ by } \Psi_L(h_1,\ldots,h_m) = L(I^{[i_m,0]}(h_1):\ldots:I^{[i_1,0]}(h_m))\in Y \tag{16.266}$$

where $X=\prod_{i\in\{1,\ldots,n\}} X_i$ so that $I^{[m-j+1,0]}(h_j)\in X$. Let $j\in\{1,\ldots,m\}$ and

$$\left(h_1,\ldots\underbrace{k+\alpha\cdot l}_{j}\ldots,h_m\right)\in\prod_{j\in\{1,\ldots,m\}} X_{i_{m-j+1}}$$

where $k,l\in X_{i_{m-j+1}}$ and $k\in\mathbb{K}$ then we have

$$\begin{array}{rcl}
\Psi_L\left(h_1,\ldots\underbrace{k+\alpha\cdot l}_{j}\ldots,h_m\right) & = & \\
L\left(I^{[i_m,0]}(h_1):\ldots\underbrace{I^{[i_{m-j+1},0]}(k+\alpha\cdot l)}_{j}\ldots:I^{[i_1,0]}(h_m)\right) & \underset{\text{[theorem: 16.41]}}{=} & \\
L\left(I^{[i_m,0]}(h_1):\ldots\underbrace{I^{[i_{m-j+1},0]}(k)+\alpha\cdot I^{[i_{m-j+1},0]}(l)}_{j}\ldots:I^{[i_1,0]}(h_m)\right) & \underset{\text{[lemma: 16.80]}}{=} & \\
L\left(I^{[i_m,0]}(h_1):\ldots\underbrace{I^{[i_{m-j+1},0]}(k)}_{j}\ldots:I^{[i_1,0]}(h_m)\right)+\alpha\cdot L\left(I^{[i_m,0]}(h_1):\ldots\underbrace{I^{[i_{m-j+1},0]}(l)}_{j}\ldots:I^{[i_1,0]}(h_m)\right) & = & \\
\Psi_L\left(h_1,\ldots\underbrace{k}_{j}\ldots,h_m\right)+\alpha\cdot\Psi_L\left(h_1,\ldots\underbrace{l}_{j}\ldots,h_m\right) & &
\end{array}$$

proving that

$$\Psi_L\in\mathrm{Hom}(X_{i_m}\ldots X_{i_1};Y) \tag{16.267}$$

Further if $(h_1,\ldots,h_m)\in\prod_{j\in\{1,\ldots,m\}} X_{i_{m-j+1}}$ then we have

$$\begin{array}{rcl}
\|\Psi_L(h_1,\ldots,h_m)\|_Y & = & \|L(I^{[i_m,0]}(h_1):\ldots:I^{[i_1,0]}(h_m))\|_Y\\
& \leqslant_{\text{[lemma: 16.81]}} & \|L\|_{L^m(X;Y)}\cdot\prod_{j=1}^{m}\|I^{[i_{m-j+1}]}(h_j)\|_{\max}\\
& \underset{\text{[theorem: 16.41]}}{=} & \|L\|_{L^m(X;Y)}\cdot\prod_{j=1}^{m}\|(h_j)\|_{i_{m-j+1}}
\end{array}$$

proving by [theorems: 14.187, 14.181] and [eq: 16.267] that

$$\Psi_L\in L(X_{i_m}\ldots X_{i_1};Y)\text{ and }\|\Psi_L\|_{L(X_{i_m}\ldots X_{i_1};Y)}\leqslant\|L\|_{L^m(X;Y)} \tag{16.268}$$

Define now the function

$$\Psi: L_m(X;Y)\rightarrow L(X_{i_m}\ldots X_{i_1};Y)\text{ by }\Psi(L)=\Psi_L \tag{16.269}$$

Let $\alpha\in\mathbb{K}$, $L_1,L_2\in L_m(X;Y)$ and $(h_1,\ldots,h_m)\in\prod_{j\in\{1,\ldots,m\}} X_{i_{m-j+1}}$ then we have

$$\begin{array}{rc}
(\Psi(L_1+\alpha\cdot L_2))(h_1:\ldots h_m) & \underset{\text{[eq: 16.269]}}{=}\\
\Psi_{L_1+\alpha\cdot L_2}(h_1:\ldots:h_m) & \underset{\text{[eq: 16.266]}}{=}\\
(L_1+\alpha\cdot L_2)(I^{[i_m,0]}(h_1):\ldots:I^{[i_1,0]}(h_m)) & \underset{\text{[lemma: 16.79]}}{=}\\
L_1((I^{[i_m,0]}(h_1):\ldots:I^{[i_1,0]}(h_m)))+\alpha\cdot L_2((I^{[i_m,0]}(h_1):\ldots:I^{[i_1,0]}(h_m))) & \underset{\text{[eq: 16.266]}}{=}\\
\Psi_{L_1}(h_1:\ldots:h_m)+\alpha\cdot\Psi_{L_2}(h_1:\ldots:h_m) & =\\
(\Psi_{L_1}+\alpha\cdot\Psi_{L_2})(h_1:\ldots:h_m) & \underset{\text{[eq: 16.269]}}{=}\\
(\Psi(L_1)+\alpha\cdot\Psi(L_2))(h_1:\ldots:h_m) &
\end{array}$$

proving by [lemma: 16.192] that

$$\Psi(L_1+\alpha\cdot L_2)=\Psi(L_1)+\alpha\cdot\Psi(L_2) \tag{16.270}$$

Using [theorem: 16.195] we can define the function

$$\Phi: L_m(X;Y)\rightarrow L_m(X_{i_m}\ldots X_{i_1};Y)\text{ by }\Phi=(\mathcal{I}_{X_{i_m}\ldots X_{i_1},Y})^{-1}\circ\Psi \tag{16.271}$$

then we have

$$\begin{array}{rcl}\Phi(L_1+\alpha\cdot L_2) & \underset{[\text{eq: } 16.271]}{=} & ((\mathcal{I}_{X_{i_m}\ldots X_{i_1},Y})^{-1}\circ\Psi)(L_1+\alpha\cdot L_2)\\ & = & (\mathcal{I}_{X_{i_m}\ldots X_{i_1},Y})^{-1}(\Psi(L_1+\alpha\cdot L_2))\\ & \underset{[\text{eq: } 16.270]}{=} & (\mathcal{I}_{X_{i_m}\ldots X_{i_1},Y})^{-1}(\Psi(L_1)+\alpha\cdot\Psi(L_2))\\ & \underset{[\text{theorem: } 16.195]}{=} & (\mathcal{I}_{X_{i_m}\ldots X_{i_1},Y})^{-1}(\Psi(L_1))+\alpha\cdot(\mathcal{I}_{X_{i_m}\ldots X_{i_1},Y})^{-1}(\Psi(L_2))\\ & = & ((\mathcal{I}_{X_1\ldots X_n,Y})^{-1}\circ\Psi)(L_1)+\alpha\cdot((\mathcal{I}_{X_1\ldots X_n,Y})^{-1}\circ\Psi)(L_2)\\ & = & \Phi(L_1)+\alpha\cdot\Phi(L_2)\end{array}$$

proving that

$$\Phi\in\text{Hom}(L_m(X;Y),L_m(X_{i_m}\ldots X_{i_1};Y)) \tag{16.272}$$

If $L\in L_m(X;Y)$ then

$$\begin{array}{rcl}\|\Phi(L)\|_{L_m(X_{i_m}:\ldots:X_{i_1};Y)} & \underset{[\text{eq: } 16.271]}{=} & \|((\mathcal{I}_{X_{i_m}\ldots X_{i_1},Y})^{-1}\circ\Psi)(L)\|_{L_m(X_{i_m}:\ldots:X_{i_1};Y)}\\ & = & \|(\mathcal{I}_{X_{i_m}\ldots X_{i_1},Y})^{-1}(\Psi(L))\|_{L_m(X_{i_m}:\ldots:X_{i_1};Y)}\\ & \underset{[\text{theorem: } 16.195]}{=} & \|\Psi(L)\|_{L(X_{i_m}\ldots X_{i_1};Y)}\\ & \underset{[\text{eq: } 16.269]}{=} & \|\Psi_L\|_{L(X_{i_m}\ldots X_{i_1};Y)}\\ & \leqslant_{[\text{eq: } 16.268]} & \|L\|_{L^m(X;Y)}\end{array}$$

proving by [theorem: 14.174] together with [eq: 16.272] that

$$\Phi\in L(L_m(X;Y),L_m(X_{i_m}\ldots X_{i_1};Y)) \tag{16.273}$$

Finally if $L\in L_m(X;Y)$ and $(h_1,\ldots,h_m)\in\prod_{j\in\{1,\ldots,m\}}X_{m-j+1}$ then we have

$$\begin{array}{rcl}\Phi(L)(h_1:\ldots:h_m) & \underset{[\text{eq: } 16.271]}{=} & (((\mathcal{I}_{X_{i_m}\ldots X_{i_1},Y})^{-1}\circ\Psi)(L))(h_1:\ldots:h_m)\\ & = & ((\mathcal{I}_{X_{i_m}\ldots X_{i_1},Y})^{-1}\Psi(L))(h_1:\ldots:h_m)\\ & \underset{[\text{theorem: } 16.195]}{=} & \Psi(L)(h_1,\ldots,h_m)\\ & \underset{[\text{eq: } 16.269]}{=} & \Psi_L(h_1,\ldots,h_m)\\ & [\text{eq: } 16.266] & L(I^{[i_m,0]}(h_1):\ldots:I^{[i_1,0]}(m))\end{array}$$

proving

$$\Phi(L)(h_1:\ldots:h_m)=L(I^{[i_m,0]}(h_1)\ldots I^{[i_1,0]}(m))$$

□

Theorem 16.202. *Let $n\in\mathbb{N}\setminus\{1\}$, $m\in\mathbb{N}$, $\{\langle X_i,\|\|_i\rangle\}_{i\in\{1,\ldots,n\}}$ a finite family of normed spaces, $\langle Y,\|\|_Y\rangle$ a normed space, $\langle X,\|\|_{\max}\rangle$ the normed space based on the product $X=\prod_{i\in\{1,\ldots,n\}}X_i$ with the maximum norm $\|\|_{\max}$ [see definition: 14.94], U a open set in X, $x\in U$ and $f:U\rightarrow Y$ a function that is m-times differentiable at x then $\forall\{i_j\}_{j\in\{1,\ldots,m\}}\subseteq\{1,\ldots,n\}$ f has a partial derivative $D^{[m]}_{x,i_m,\ldots,i_1}f$ of order m at x. Further $\forall h=(h_1,\ldots,h_n)\in\prod_{j\in\{1,\ldots,m\}}X_{i_{m-j+1}}$ we have*

$$D^{[m]}_{x,i_m\ldots i_1}f(h_1:\ldots:h_m)=D^{[m]}_xf(I^{[i_m,0]}(h_1):\ldots:I^{[i_1,0]}(h_m))$$

Proof. We prove this by induction, so let

$$S=\left\{m\in\mathbb{N}|\text{If } U \text{ is a open set in } X \text{ and } f:U\rightarrow Y \text{ is a function that is } m\text{-times differentiable at } x\in U \text{ then } \forall\{i_j\}_{j\in\{1,\ldots,m\}}\subseteq\{1,\ldots,n\} \text{ we have that } f \text{ has } D^{[m]}_{x,i_m\ldots i_1}f \text{ exist and } \forall(h_1,\ldots,h_m)\in\prod_{j\in\{1,\ldots,m\}}X_{i_{m-j+1}} \text{ that } D^{[m]}_{x,i_m\ldots i_1}f(h_1:\ldots:h_m)=D^{[m]}_xf(I^{[i_m,0]}(h_1):\ldots:I^{[i_1,0]}(h_m))\right\}$$

then we have:

$1\in S$. Using [lemma: 16.200] it follows that $D^{[1]}_{x,i_1,\ldots,i_1}f$ exist and

$$D^{[1]}_{x,i\ldots i}f=D^{[1]}_xf\circ I^{[i,0]} \tag{16.274}$$

If $h\in\prod_{j\in\{1\}}X_{i_j}$ then

$$\begin{array}{rcl}D^{[1]}_{x,i_1\ldots i_1}f(h_1:\ldots:h_1) & = & D^{[1]}_{x,i_1\ldots i_1}f(h_1)\\ & \underset{\text{[eq: 16.274]}}{=} & \left(D^{[1]}_x f\circ I^{[i,0]}\right)(h_1)\\ & = & D^{[1]}_x f(I^{[i,0]}(h_1))\\ & = & D^{[1]}_x f(I^{[i,0]}(h_1)\ldots I^{[i,0]}(h_1))\end{array}$$

Hence

$$1\in S$$

$\boldsymbol{m\in S\Rightarrow m+1\in S.}$ Assume that f is $(m+1)$-times differentiable at x and take $\{i_j\}_{j\in\{1,\ldots,m+1\}}\subseteq\{1,\ldots,n\}$. As f is $(m+1)$-times differentiable at x there exist a open set V with $x\in V\subseteq U$ such that $\forall y\in V$ f is m-times differentiable at y and the function

$$D^{[m]}f: V\rightarrow L_m(X;Y) \text{ defined by } D^{[m]}f=D^{[m]}_x f$$

is Fréchet differentiable at x with by definition

$$D^{[m+1]}_x f=D_x(D^{[m]}f) \tag{16.275}$$

As $m\in S$ it follows that for $y\in V$ $D^{[m]}_{y,i_m\ldots i_1}f$ exists and for $(h_1,\ldots,h_m)\in\prod_{j\in\{1,\ldots,n\}}X_{m-j+1}$ we have

$$D^{[m]}_{y,i_m\ldots i_1}f(h_1:\ldots:h_m)=D^{[m]}_y f(I^{[i_m,0]}(h_1):\ldots:I^{(i_1,0)}(h_1)) \tag{16.276}$$

As $\forall y\in V$ $D^{[m]}_{y,i_m\ldots i_1}f$ exist we can define

$$D^{[m]}_{y,i_m\ldots i_1}f: V\rightarrow L_m(X_{i_m}\ldots X_{i_1};Y) \text{ by } \left(D^{[m]}_{i_m\ldots i_1}f\right)(y)=D^{[m]}_{y,i_m\ldots i_1}f \tag{16.277}$$

Using [lemma: 16.201] there exist a $\Phi\in L(L_m(X;Y),L_m(X_{i_m}\ldots X_{i_2};Y))$ such that $\forall L\in L_m(X;Y)$ and $(h_1,\ldots,h_m)\in\prod_{j\in\{1,\ldots,n\}}X_{m-j+1}$ we have

$$\Phi(L)(h_1:\ldots:h_m)=L(I^{[i_m,0]}(h_1):\ldots:I^{[i_1,0]}(h_m)) \tag{16.278}$$

Let $y\in V$ and $(h_1,\ldots,h_m)\in\prod_{j\in\{1,\ldots,n\}}X_{m-j+1}$ then as $D^{[m]}f(y)=D^{[m]}_y f\in L_m(X;Y)$ we have that

$$\begin{array}{rcl}\Phi(D^{[m]}f(y))(h_1:\ldots:h_m) & = & (D^{[m]}f(y))(I^{[i_m,0]}(h_1):\ldots:I^{[i_1,0]}(h_m))\\ & = & D^{[m]}_y f(I^{[i_m,0]}(h_1):\ldots:I^{[i_1,0]}(h_m))\\ & \underset{\text{[eq: 16.276]}}{=} & D^{[m]}_{y,i_m\ldots i_1}f(h_1:\ldots:h_m)\\ & \underset{\text{[eq: 16.277]}}{=} & \left(D^{[m]}_{i_m\ldots i_1}f(y)\right)(h_1:\ldots:h_m)\end{array}$$

proving by [lemma: 16.192] that $(\Phi\circ D^{[m]}f)(y)=\Phi(D^{[m]}f(y))=D^{[m]}_{i_m\ldots i_1}f(y)$ so that

$$\Phi\circ D^{[m]}f=D^{[m]}_{i_m\ldots i_1}f \tag{16.279}$$

As $D^{[m]}f$ is Fréchet differentiable at x and $\Phi\in L(L_m(X;Y),L_m(X_{i_m}\ldots X_{i_2};Y))$ hence ∞-times differentiable on $L_m(X;Y)$ [see 16.116] it follows from the chain rule [see theorem: 16.36] that $\Phi\circ D^{[m]}f$ is Fréchet differentiable at x, hence $D^{[m]}_{i_m\ldots i_1}f$ is Fréchet differentiable at x. Further

$$\begin{array}{rcl}D_x\left(D^{[m]}_{i_m\ldots i_1}f\right) & \underset{\text{[eq: 16.279]}}{=} & D_x(\Phi\circ D^{[m]}f)\\ & \underset{\text{[theorem: 16.36]}}{=} & D_{D^{[m]}f(x)}\Phi\circ D_x(D^{[m]}f)\\ & \underset{\text{[theorem: 16.27]}}{=} & \Phi\circ D_x(D^{[m]}f)\\ & = & \Phi\circ D^{[m+1]}f\end{array}$$

proving that

$$D_x\left(D^{[m]}_{i_m\ldots i_1}f\right)=\Phi\circ D^{[m+1]}f \tag{16.280}$$

As by [theorem: 16.47] Fréchet differentiability implies partial differentiability it follows that $D^{[m]}_{i_m\ldots i_1}f$ has a i_{m+1}-partial differential at x and

$$D_{x,i_{m+1}}\big(D^{[m]}_{i_m\ldots i_1}f\big) \underset{\text{[theorem: 16.47]}}{=} D_x\big(D^{[m]}_{i_m\ldots i_1}f\big)\circ I^{i_{m+1},0} \underset{\text{[eq: 16.280]}}{=} \Phi\circ D^{[m+1]}_x f\circ I^{i_{m+1},0}$$

Hence by definition

$$f \text{ has a partial differential } D^{[m+1]}_{x,i_{m+1}\ldots i_1}f \text{ of order } m+1 \text{ at } x \tag{16.281}$$

and

$$D^{[m+1]}_{x,i_{m+1}\ldots i_1}f \underset{\text{def}}{=} D_{x,i_{m+1}}\big(D^{[m]}_{i_m\ldots i_1}f\big) = \Phi\circ D^{[m+1]}_x f\circ I^{i_{m+1},0} \tag{16.282}$$

Further, as $D^{[m+1]}_x f\in L_{m+1}(X;Y)=L(X,L_m(X;Y))$,

$$D^{[m+1]}_x f(I^{[i_{m+1},0]}(h_1))\in L_m(X;Y) \tag{16.283}$$

So for $(h_1,\ldots,h_{m+1})\in\prod_{j\in\{1,\ldots,m+1\}} X_{i_{(m+1)-j+1}}$ we have

$$\begin{aligned}
D^{[m+1]}_{x,i_{m+1}\ldots,i_1}f(h_1:\ldots:h_{m+1}) &\underset{\text{[eq: 16.282]}}{=} \\
\big(\Phi\circ D^{[m+1]}_x f\circ I^{[i_{m+1},0]}\big)(h_1:\ldots:h_{m+1}) &= \\
\big(\big(\Phi\circ D^{[m+1]}_x f\circ I^{[i_{m+1},0]}\big)(h_1)\big)(h_2:\ldots:h_{m+1}) &= \\
\big(\Phi\big(D^{[m+1]}_x f(I^{[i_{m+1},0]}(h_1))\big)\big)(h_2:\ldots:h_{m+1}) &\underset{\text{[eqs: 16.278, 16.283]}}{=} \\
\big(D^{[m+1]}_x f(I^{[i_{m+1},0]}(h_1))\big)(I^{[i_m,0]}(h_2):\ldots:I^{[i_1,0]}(h_{m+1})) &= \\
D^{[m+1]}_x f(I^{[m+1,0]}(h_1):I^{[i_m,0]}(h_2):\ldots:I^{[i_1,0]}(h_{m+1})) &= \\
D^{[m+1]}_x f(I^{[m+1,0]}(h_1):\ldots:I^{[i_1,0]}(h_{m+1}))
\end{aligned}$$

proving that

$$D^{[m+1]}_{x,i_{m+1}\ldots,i_1}f(h_1:\ldots:h_{m+1}) = D^{[m+1]}_x f(I^{[m+1,0]}(h_1):\ldots:I^{[i_1,0]}(h_{m+1}))$$

which together with [eq: 16.281] proves

$$m+1\in S$$

□

Theorem 16.203. *Let $n\in\mathbb{N}\setminus\{1\}$, $m\in\mathbb{N}$, $\{\langle X_i,\|\|_i\rangle\}_{i\in\{1,\ldots,n\}}$ a finite family of normed spaces, $\langle Y,\|\|_Y\rangle$ a normed space, $\langle X,\|\|_{\max}\rangle$ the normed space based on the product $X=\prod_{i\in\{1,\ldots,n\}}X_i$ with the maximum norm $\|\|_{\max}$ [see definition: 14.94], U a open set in X and*

$$f: U\rightarrow Y \text{ a function of class } C^m$$

then $\forall\{i_j\}_{j\in\{1,\ldots,m\}}\subseteq\{1,\ldots,n\}$ we have that $\forall x\in U$

$$f \text{ has a partial differential } D^{[m]}_{x,i_m\ldots i_1}f \text{ of order } m \text{ at } x$$

and

$$D^{[m]}_{i_m\ldots i_1}f: U\rightarrow L_m(X_{i_m}\ldots X_{i_1};Y) \text{ by } \big(D^{[m]}_{i_m\ldots i_1}f\big)(x) = D^{[m]}_{x,i_m\ldots i_1}f$$

is continuous.

Proof. Let $x\in X$, as f is of class C^m f is m-times differentiable at x. Hence, using the previous theorem [theorem: 16.202] f has a partial derivative $D^{[m]}_{x,i_m\ldots i_1}f$ of order m at x so we can define the function

$$D^{[m]}_{i_m\ldots i_1}f: U\rightarrow L_m(X_{i_m}\ldots X_{i_1};Y) \text{ by } D^{[m]}_{i_m\ldots i_1}f(x) = D^{[m]}_{x,i_m\ldots i_1}f$$

As f is m-times differentiable on U we can define the function

$$D^{[m]}f: U\rightarrow L_m(X;Y) \text{ by } D^{[m]}f(x) = D^{[m]}_x f$$

Further for every $(h_1,\ldots,h_m)\in\prod_{j\in\{1,\ldots,m\}}X_{m-j+1}$ we have

$$D^{[m]}_{x,i_m\ldots i_1}f(h_1:\ldots:h_m) = D^{[m]}_x f(I^{[i_m,0]}(h_1):\ldots:I^{[i_1,0]}(h_m))$$

so that

$$D^{[m]}_{i_m\ldots i_1}f(x)(h_1:\ldots:h_m)=D^{[m]}f(x)(I^{[i_m,0]}(h_1):\ldots:I^{[i_1,0]}(h_m)) \tag{16.284}$$

Now using [lemma: 16.201] there exist a $\Phi\in L(L_m(X;Y),L_m(X_{i_m}\ldots X_{i_1};Y))$ such that

$$\Phi(D^{[m]}f(x))(h_1:\ldots:h_m)=D^{[m]}f(x)(I^{[i_m,0]}(h_1):\ldots:I^{[i_1,0]}(h_m))=D^{[m]}_{i_m\ldots i_1}f(x)(h_1:\ldots:h_m)$$

proving by [lemma: 16.192] that $\Phi(D^{[m]}f(x))=D^{[m]}_{i_m\ldots i_1}f(x)$. Hence as $(\Phi\circ D^{[m]}f)(x)=(D^{[m]}f(x))$ it follows that

$$D^{[m]}_{i_m\ldots i_1}f=\Phi\circ D^{[m]}f$$

As $\Phi\in L(L_m(X;Y),L_m(X_{i_m}\ldots X_{i_1};Y))$ and f is of class C^m we have that Φ and $D^{[m]}f$ are continuous, so $\Phi\circ D^{[m]}f$ is continuous proving that $D^{[m]}_{i_m\ldots i_1}f$ is continuous. □

Next we show that the order in which you do the partial differentiation does not matter.

Theorem 16.204. *Let $n\in\mathbb{N}\setminus\{1\}$, $m\in\mathbb{N}$, $\{\langle X_i,\|\|_i\rangle\}_{i\in\{1,\ldots,n\}}$ a finite family of normed spaces, $\langle Y,\|\|_Y\rangle$ a **Banach** space, $\langle X,\|\|_{\max}\rangle$ the normed space based on the product $X=\prod_{i\in\{1,\ldots,n\}}X_i$ with the maximum norm $\|\|_{\max}$ [see definition: 14.94], U a open set in X, $x\in U$ and $f:U\rightarrow Y$ a function that is m-times differentiable at x, $\{i_j\}_{j\in\{1,\ldots,m\}}\subseteq\{1,\ldots,n\}$ and $(h_1,\ldots,h_m)\in\prod_{j\in\{1,\ldots,m\}}X_{m-j+1}$ then $\forall\sigma\in P_n$ we have*

$$D^{[m]}_{x,i_m,\ldots i_1}f(h_1:\ldots h_m)=D^{[m]}_{i_{\sigma(m)}\ldots i_{\sigma(1)}}f(h_{\sigma(1)},\ldots,h_{\sigma(m)})$$

Note 16.205. *Existence of the partial differentials of order m at x is guaranteed by [theorem: 16.202]*

Proof. We have

$$\begin{aligned}D^{[m]}_{x,i_{\sigma(m)},\ldots i_{\sigma(1)}}f(h_{\sigma(1)}:\ldots h_{\sigma(m)}) &\underset{\text{[theorem: 16.202]}}{=}\\ D^{[m]}_x f(I^{i_{[\sigma(m)],0}}(h_{\sigma(1)}):\ldots:I^{[i_{\sigma(1)},0]}(h_{\sigma(m)})) &\underset{\text{[theorem: 16.180]}}{=}\\ D^{[m]}_x(I^{[i_m,0]}(h_1):\ldots:I^{[i_1,0]}(h_1)) &\underset{\text{[theorem: 16.202]}}{=}\\ D^{[m]}_{x,i_m\ldots i_1}f(h_1:\ldots:h_m)\end{aligned}$$

□

It was show in [example: 16.50] that the existence of partial differentials at a point does not mean that that the function is differentiable at the point. However if the partial derivatives exist on a open neighborhood of the point and are continuous then the function is differentiable at the point. To prove this we need first some lemma's.

Definition 16.206. *Let $n\in\mathbb{N}$, $\{\langle X_i,\|\|_i\rangle\}_{i\in\{1,\ldots,n\}}$ a finite family of normed spaces, $\langle X,\|\|_{\max}\rangle$ the normed space based on the product $X=\prod_{i\in\{1,\ldots,n\}}X_i$ with the maximum norm $\|\|_{\max}$ [see definition: 14.94] then given $h=(h_1,\ldots,h_n)\in X$ and $i\in\{0,\ldots,n\}$ we define*

$$h^{\{i\}}=\begin{cases}\sum_{j=1}^n I^{[j,0]}(h_j) & \text{if } i\in\{1,\ldots,n\}\\ 0 & \text{if } i=0\end{cases}$$

so that

$$h^{\{i\}}=\begin{cases}\left(h_1,\ldots,h_i,\underbrace{0\ldots,0}_{n-i}\right) & \text{if } i\in\{1,\ldots,n\}\\ \left(\underbrace{0,,\ldots,,0}_{n}\right) & \text{if } 0\end{cases}$$

Lemma 16.207. *Let $n\in\mathbb{N}$, $\{\langle X_i,\|\|_i\rangle\}_{i\in\{1,\ldots,n\}}$ a finite family of normed spaces, $\langle X,\|\|_{\max}\rangle$ the normed space based on the product $X=\prod_{i\in\{1,\ldots,n\}}X_i$ with the maximum norm $\|\|_{\max}$ [see definition: 14.94] then given $h\in X$ we have*

1. $h^{\{0\}}=0$

2. *$h^{\{n\}}=h$*
3. *$\forall i\in\{1,\ldots,n\}$ we have $\forall k\in\{1,\ldots,n\}$ that $(h^{\{i\}})_k=\begin{cases}0 \text{ if } i<k\\ h_j \text{ if } k\leqslant i\end{cases}$*
4. *Let $x,h\in\prod_{i\in\{1,\ldots,n\}}X_i$ then $\forall i\in\{1,\ldots,n\}$ we have*
   a. *$I^{[i,x+h^{\{i\}}]}(x_i+h_i)=x+h^{\{i\}}$*
   b. *$I^{[i,x+h^{\{i\}}]}(x_i)=x+h^{\{i-1\}}$*
5. *$\forall i\in\{0,\ldots,n\}$ we have $\|h^{\{i\}}\|_{\max}\leqslant\|h\|_{\max}$*

Proof.

1. This follows from the definition.
2. $h^{\{n\}}\underset{n\in\mathbb{N}\Rightarrow 0\neq n}{=}\sum_{j=1}^n I^{[j,0]}h_j\underset{[\text{theorem: }16.41]}{=}h$
3. Let $i\in\{1,\ldots,n\}$ then Then for $k\in\{1,\ldots,n\}$ we have either:

   k ⩽ i. Then

$$\begin{aligned}(h^{\langle i\rangle})_k &= \left(\sum_{j=1}^{i} I^{[j,0]}(h_j)\right)_k\\ &= \sum_{j=1}^{i}(I^{[j,0]}(h_j))_k\\ &\underset{[\text{lemma: }16.41]}{=} \sum_{j=1}^{i}\delta_{j,k}\cdot h_j\\ &= h_k\end{aligned}$$

   i < k. Then

$$\begin{aligned}(h^{\langle i\rangle})_k &= \left(\sum_{j=1}^{i} I^{[j,0]}(h_j)\right)_k\\ &= \sum_{j=1}^{i}(I^{[j,0]}(h_j))_k\\ &\underset{[\text{lemma: }16.41]}{=} \sum_{j=1}^{i}\delta_{j,k}\cdot h_j\\ &\underset{\forall j\in\{1,\ldots,n\}\ k\neq j}{=} 0\end{aligned}$$

   so that

$$(h^{\{i\}})_k=\begin{cases}0 \text{ if } i<k\\ h_j \text{ if } k\leqslant i\end{cases}.$$

4. Let $x,h\in\prod_{i\in\{1,\ldots,n\}}X_i$ and $i\in\{1,\ldots,n\}$ then we have:
   a. Then we have for $k\in\{1,\ldots,n\}$ either:

      k = i. Then

$$\begin{aligned}\left(I^{[i,x+h^{\{i\}}]}(x_i+h_i)\right)_k &\underset{k=i}{=} \left(I^{[i,x+h^{\{i\}}]}(x_i+h_i)\right)_i\\ &= x_i+h_i\\ &= x_k+h_k\\ &\underset{(3)\wedge k\leqslant i}{=} x_k+(h^{\{i\}})_k\\ &= (x+h^{\{i\}})_k\end{aligned}$$

$k \neq i$. Then

$$\left(I^{[i,x+h^{\{i\}}]}(x_i+h_i)\right)_k = (x+h^{\{i\}})_k$$

proving that

$$I^{[i,x+h^{\{i\}}]}(x_i+h_i) = x+h^{\{i\}}$$

b. Then we have for $k \in \{1,\ldots,n\}$ either:

$k = i$.

$$\begin{aligned}\left(I^{[i,x+h^{\{i\}}]}(x_i)\right)_k &\underset{k=i}{=} x_i \\ &= x_i+0 \\ &= x_k+0 \\ &\underset{i-1<i=k}{=} x_k+(h^{\{i-1\}})_k \\ &= (x+h^{\{i-1\}})_k\end{aligned}$$

$k < i$.

$$\begin{aligned}\left(I^{[i,x+h^{\{i\}}]}(x_i)\right)_k &= (x+h^{\{i\}})_k \\ &= x_k+(h^{\{i\}})_k \\ &\underset{(3)\wedge k\leqslant i}{=} x_k+h_k \\ &\underset{(3)\wedge k\leqslant i-1}{=} x_k+(h^{\{i-1\}})_k \\ &= (x+h^{\{i-1\}})_k\end{aligned}$$

$i < k$.

$$\begin{aligned}\left(I^{[i,x+h^{\{i\}}]}(x_i)\right)_k &= (x+h^{\{i\}})_k \\ &= x_k+(h^{\{i\}})_k \\ &\underset{(3)\wedge i<k}{=} x_k+0 \\ &\underset{(3)\wedge i-1<i<k}{=} x_k+(h^{\{i-1\}})_k \\ &= (x+h^{\{i-1\}})_k\end{aligned}$$

proving that

$$I^{[i,x+h^{\{i\}}]}(x_i) = x+h^{\{i-1\}}$$

5. Given $i \in \{0,\ldots,n\}$ then we have either:

$i = 0$. Then

$$\|h^{\langle i\rangle}\|_{\max} = \|h^{\langle 0\rangle}\|_{\max} = \|0\|_{\max} \leqslant \|h\|_{\max}$$

$i \in \{1,\ldots,n\}$. Then for $k \in \{1,\ldots,n\}$ we have

$k \in \{1,\ldots,i\}$. Then

$$\begin{aligned}(h^{\langle i\rangle})_k &= \left(\sum_{j=1}^{i} I^{[j,0]}(h_j)\right)_k \\ &= \sum_{j=1}^{i} (I^{[j,0]}(h_j))_k \\ &\underset{[\text{lemma: } 16.41]}{=} \sum_{j=1}^{i} \delta_{j,k}\cdot h_j \\ &= h_k\end{aligned}$$

so that $\|(h^{\langle i\rangle})_k\|_k = \|h_k\|_k$

$\boldsymbol{k\in\{i+1,\ldots,n\}}$**.** Then

$$\begin{aligned}(h^{\langle i\rangle})_k &= \left(\sum_{j=1}^{i} I^{[j,0]}(h_j)\right)_k\\ &= \sum_{j=1}^{i} (I^{[j,0]}(h_j))_k\\ &\underset{[\text{lemma: } 16.41]}{=} \sum_{j=1}^{i} \delta_{j,k}\cdot h_j\\ &\underset{\forall j\in\{1,\ldots,n\}\ k\neq j}{=} 0\end{aligned}$$

so that $\|(h^{\langle i\rangle})_k\|_k = \|0\|_k \leqslant \|h_k\|_k$

proving that $\forall k\in\{1,\ldots,n\}\ \|(h^{\langle i\rangle})_k\|_k \leqslant \|h_k\|_k$. Hence

$$\|h^{\langle i\rangle}\|_{\max} = \max(\{\|h^{\langle i\rangle}\|_k | k\in\{1,\ldots,n\}\}) \leqslant \max(\{\|h_k\|_k | k\in\{1,\ldots n\}\}) = \|h\|_{\max}$$

So we have that

$$\forall i\in\{0,\ldots,n\}\ \|h^{\langle i\rangle}\|_{\max} \leqslant \|h\|_{\max} \text{ for every } h\in X$$

□

Lemma 16.208. *Let $n\in\mathbb{N}\setminus\{1\}$, $m\in\mathbb{N}$, $\{\langle X_i,\|\|_i\rangle\}_{i\in\{1,\ldots,n\}}$ a finite family of normed spaces, $\langle Y,\|\|_Y\rangle$ a **Banach** space, $\langle X,\|\|_{\max}\rangle$ the normed space over $\mathbb{R}$ based on the product $X=\prod_{i\in\{1,\ldots,n\}} X_i$ with the maximum norm $\|\|_{\max}$ [see definition: 14.94], U a open set in X, $x\in U$, $\delta\in\mathbb{R}^+$ with $B_{\|\|_{\max}}(x,\delta)$, $i\in\{1,\ldots,n\}$ and a function*

$$f: U\rightarrow Y$$

such that $\forall y\in U$ f has a i-partial derivative $D_{y,i}f$ at y and the function

$$D_i f: U\rightarrow L(X_i;Y) \text{ defined by } D_i f(y) = D_{y,i} f$$

is continuous. Then if $u,v\in B_{\|\|_i}(x_i,\delta)$ we have that

$$\sup(\{\|D_i(I^{[i,x]}(u+t\cdot(v-u)))\|_{L(X_i,Y)} | t\in[0,1]\})$$

exists and

$$\|f(I^{[i,x]}(u)) - f(I^{[i,x]}(v))\|_Y \leqslant \sup(\{\|D_i(I^{[i,x]}(u+t\cdot(v-u)))\|_{L(X_i,Y)} | t\in[0,1]\})$$

Proof. We have

$$\begin{aligned} t\in(I^{i,[x]})^{-1}(B_{\|\|_{\max}}(x,\delta)) &\Leftrightarrow I^{[i,x]}(t)\in B_{\|\|_{\max}}(x,\delta)\\ &\Leftrightarrow \|I^{[i,x]}(t)-x\|_{\max}<\delta\\ &\Leftrightarrow \left\|\left(x_1,\ldots\underbrace{t}_{i}\ldots,x_n\right)-(x_1,\ldots,x_n)\right\|_{\max}<\delta\\ &\Leftrightarrow \|(0,\ldots\underbrace{t-x_i}\ldots,x_n)\|_{\max}<\delta\\ &\Leftrightarrow \max\left(\left\{\|0\|_1,\ldots\underbrace{\|t-x_i\|_i}\ldots\|0\|_n\right\}\right)<\delta\\ &\Leftrightarrow \|t-x\|_i<\delta\\ &\Leftrightarrow t\in B_{\|\|_i}(x_i,\delta)\end{aligned}$$

so that

$$B_{\|\|_i}(x,\delta) = (I^{i,[x]})^{-1}(B_{\|\|_{\max}}(x,\delta)) \subseteq (I^{i,[x]})^{-1}(U) \text{ and } I^{[i,x]}(B_{\|\|_i}(x,\delta))\subseteq U \quad (16.285)$$

The above allows us to define the function $h = f \circ (I^{[i,x]})_{|B_{\|\|_i}(x_i,\delta)}$ so that

$$h: B_{\|\|_i}(x_i,\delta) \rightarrow Y \text{ is defined by } h(t) = f(I^{[i,x]}(t)) \tag{16.286}$$

Let $y \in B_{\|\|_i}(x_i,\delta)$ and define $z = I^{[i,x]}(y)$ then by [eq: 16.285] $z \in U$, hence z has a i-partial derivative at z. Then by definition the function

$$f \circ (I^{[i,z]})_{|(I^{[i,z]})^{-1}(U)}: (I^{[i,z]})^{-1}(U) \rightarrow Y$$

is Fréchet differentiable at z_i and $D_i f(z) = D_{z,i} f = D_{z_i}(f \circ (I^{[i,z]})_{|(I^{[i,z]})^{-1}(U)})$. Let $s \in X_i$ then $I^{[i,z]}(s) = \left(z_1, \ldots \underbrace{s}_{i} \ldots z_n\right) \underset{z = I^{[i,x]}(y)}{=} \left(I^{[i,x]}(x_1) \ldots \underbrace{s}_{i} \ldots I^{[i,x]}(x_n)\right) = (x_1, \ldots \underbrace{s} \ldots x_{n_i}) = I^{[i,x]}(s)$ proving that $I^{[i,z]} = I^{[i,x]}$, further $z_i = (I^{[i,x]}(y))_i = \left(x_1, \ldots \underbrace{y}_{i} \ldots, x_n\right) = y$. So we have that the function

$$f^{[i,x]}: f \circ (I^{[i,x]})_{(I^{[i,x]})^{-1}(U)}: (I^{[i,x]})^{-1}(U) \rightarrow Y$$

is Fréchet differentiable at y and $D_i f(I^{[i,x]}(y)) = D_{I^{[i,x]}(y),i} f = D_y(f \circ (I^{[i,x]})_{|(I^{[i,x]})^{-1}(U)}) = D_y f^{[i,x]}$. Let $t \in B_{\|\|_i}(x_i,\delta)$ then

$$\begin{array}{rcl} h(t) & = & (f \circ (I^{[i,x]})_{|B_{\|\|_i}(x_i,\delta)})(t) \\ & = & f((I^{[i,x]})_{|B_{\|\|_i}(x_i,\delta)}(t)) \\ & = & f(I^{[i,x]}(t)) \\ & \underset{B_{\|\|_i}(x_i,t) \subseteq (I^{[i,x]})^{-1}(U)}{=} & f((I^{[i,x]})_{|(I^{[i,x]})^{-1}(U)}(t)) \\ & = & (f \circ (I^{[i,x]})_{|(I^{[i,x]})^{-1}(U)})(t) \\ & = & f^{[i,x]}(t) \end{array}$$

proving that

$$h = (f^{[i,x]})_{|B_{\|\|_i}(x_i,\delta)}$$

So using [theorem: 16.24] it follows that $\forall y \in B_{\|\|_i}(x,\delta)$ h is Fréchet differentiable at y and

$$D_y h = D_y(f^{[i,x]})_{|B_{\|\|_i}(x_i,\delta)} = D_y f^{[i,x]} = D_{I^{[i,x]},i} f = D_i f(I^{[i,x]}(y)) \tag{16.287}$$

Hence if we define

$$Dh: B_{\|\|_i}(x,\delta) \rightarrow Dh(y) = D_y h = D_i f(I^{[i,x]}(y))$$

then $Dh = D_i f \circ I^{[i,x]}$. As by the hypothesis $D_i f$ is continuous and $I^{[i,x]}$ is continuous by [lemma: 16.41] it follows that Dh is continuous. So we have proved that

$$h \text{ is of class } C^1$$

Using the mean value theorem [corollary: 16.164] we have for $u, v \in B_{\|\|_i}(x,\delta)$ that

$$\sup\left(\{\|D_{u+t\cdot(v-u)} h\|_{L(X_i,Y)} | t \in [0,1]\}\right)$$

exist and

$$\|h(u) - h(v)\|_Y \leqslant \|u - v\|_i \cdot \sup\left(\{\|D_{u+t\cdot(v-u)} h\|_{L(X_i,Y)} | t \in [0,1]\}\right)$$

Using [eqs: 16.286, 16.287] we have that

$$\sup\left(\{\|D_i f(I^{[i,x]}(u + t\cdot(v-u)))\|_{L(X_i,Y)} | t \in [0,1]\}\right)$$

exist and

$$\|f(I^{[i,x]}(u)) - f(I^{[i,x]})(v)\|_Y \leqslant \|u - v\|_i \cdot \sup\left(\{\|D_i f(I^{[i,x]}(u + t\cdot(v-u)))\|_{L(X_i,Y)} | t \in [0,1]\}\right)$$ □

Theorem 16.209. *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* ***Banach*** *space, $\langle X, \|\|_{\max} \rangle$ the normed space over $\mathbb{R}$ based on the product $X = \prod_{i \in \{1,\ldots,n\}} X_i$ with the maximum norm $\|\|_{\max}$ [see definition: 14.94], U a open set in X, $f: U \rightarrow Y$ a function such that $\forall x \in U$ and $\forall i \in \{1,\ldots,n\}$ f has a i-partial derivative $D_{x,i} f$ at x and the function*

$$D_i f: U \rightarrow L(X_i; Y) \text{ defined by } D_i f(y) = D_{y,i} f$$

is continuous. Then f is of class C^1 and

$$D_x f = \sum_{i=1}^{n} (D_{x,i} f \circ \pi_i)$$

So if we define

$$D f: U \rightarrow L(X;Y): (D f)(x) = D_x f$$

then

$$D f(x) = \sum_{i=1}^{n} (D_i f(x) \circ \pi_i)$$

Proof. By the hypothesis we have that

$$\forall i \in \{1, \ldots, n\} \text{ that } D_i f: U \rightarrow L(X_i, Y) \text{ defined by } D_i f(x) = D_{x,i} f \text{ is continuous}$$

Let $x \in U$ then as U is open there exist a $\delta \in \mathbb{R}^+$ such that

$$x \in B_{\|\|_{\max}}(x, \delta_1) \subseteq U$$

Define

$$K: X \rightarrow Y \text{ by } K(x) = \sum_{j=1}^{n} D_j f(x)(y_j) \tag{16.288}$$

We proceed now to prove that K is linear and continuous. Let $y, z \in X$ then we have

$$\begin{aligned} K(y + \alpha \cdot z) &= \sum_{j=1}^{n} D_j f(x)(y + \alpha \cdot z)_j \\ &= \sum_{j=1}^{n} D_i f(x)(y_i + \alpha \cdot z_i) \\ &= \sum_{j=1}^{n} (D_j f(x)(y_i) + \alpha \cdot D_i f(x)(z_j)) \\ &= \sum_{i=1}^{n} D_j f(x)(y_j) + \alpha \cdot \sum_{j=1}^{n} D_j f(x)(z_j) \\ &= K(y) + \alpha \cdot K(z) \end{aligned}$$

proving that

$$L \in \text{Hom}(X, Y) \tag{16.289}$$

Further for every $y \in X$ we have

$$\begin{aligned} \|K(y)\|_Y &= \left\| \sum_{j=1}^{n} D_j f(x)(y_j) \right\|_Y \\ &\leqslant \sum_{j=1}^{n} \|D_j f(x)(y_j)\|_Y \\ &\leqslant \sum_{j=1}^{n} \|D_j f(x)\|_{L(X_i,Y)} \cdot \|y_j\| \\ &\leqslant \sum_{j=1}^{n} \|D_j f(x)\|_{L(X_i,Y)} \cdot \|y\|_{\max} \\ &\leqslant \left(\sum_{j=1}^{n} \|D_j f(x)\|_{L(X_j,Y)} \right) \cdot \|y\|_{\max} \end{aligned}$$

So using [theorem: 14.174] it follows that

$$K \in L(X, Y) \tag{16.290}$$

Let $i\in\{1,\ldots,n\}$. By [theorems: 16.116, 16.24] K is ∞-times differentiable on $B_{\|\|_{\max}}(x,\delta_1)$ and $\forall y\in B_{\|\|_{\max}}(x,\delta)$ $D_y K=K$. Using [theorem: 16.47] we have that

$$\forall y\in B_{\|\|_{\max}}(x,\delta) \text{ we have that } D_{y,i}K \text{ exists and } D_{y,i}K=K\circ I^{[i,0]}$$

Now for $k\in X_i$ we have that

$$\begin{aligned}(K\circ I^{[i,0]})(k) &= L((i\to 0)(k))\\ &= \sum_{j=1}^n D_j f(x)((i\to 0)(k))_j\\ &\underset{[\text{lemma: }16.41]}{=} \sum_{j=1}^n D_j f(x)\cdot\delta_{i,j}(k)\\ &= D_i f(x)(k)\end{aligned}$$

proving that $K\circ I^{[i,0]}=D_i f(x)$. Hence we have

$$\forall y\in B_{\|\|_{\max}}(x,\delta) \text{ we have that } D_{y,i}K \text{ exists and } D_{y,i}K=D_i f(x) \tag{16.291}$$

Further using [theorem: 16.45] we have that

$$\forall y\in B_{\|\|_{\max}}(x,\delta)\ D_{y,i}(f_{|B_{\|\|_{\max}}(x,\delta)}) \text{ exist and } D_{y,i}(f_{|B_{\|\|_{\max}}(x,\delta)})=D_{y,i}f=D_i f(y) \tag{16.292}$$

Also, using [example: 16.48] it follows that

$$\forall y\in B_{\|\|_{\max}}(x,\delta)\ D_{y,i}C_{f(x)} \text{ exist and } D_{y,i}C_{f(x)}=C_0 \tag{16.293}$$

$$\forall\in B_{\|\|_{\max}}(x,\delta)\ D_{y,i}\left(\sum_{j=1}^n D_i f(x)\cdot x_i\right) \text{ exist and } D_{y,i}C_{\sum_{j=1}^n D_i f(x)\cdot x_i}=C_0 \tag{16.294}$$

Define

$$L: B_{\|\|_{\max}}(x,\delta)\to Y \text{ by } L(x)=\sum_{j=1}^n D_j f(x)(y_j-x_j)=\left(\sum_{j=1}^n D_j f(x)(y_j)\right)-\sum_{j=1}^n D_j f(x)(x_j)$$

then

$$L=K-C_{\sum_{j=1}^n D_j f(x)(x_j)} \tag{16.295}$$

Using [theorem: 16.51] together with [eqs: 16.291,16.294] it follows that

$$\forall y\in B_{\|\|_{\max}}(x,\delta)\ D_{y,x}L \text{ exists and } D_{y,i}L=D_i f(x) \tag{16.296}$$

Define

$$g: B_{\|\|_{\max}}(x,\delta)\to Y \text{ by } g(y)=f(y)-f(x)-\sum_{j=1}^n D_j f(x)(y_j-x_j) \tag{16.297}$$

then $g=f_{|B_{\|\|_{\max}}(x,\delta)}-C_{f(x)}-L$. Hence by [eqs: 16.292, 16.294, 16.296] and [theorem: 16.51] it follows that

$$\forall y\in B_{\|\|_{\max}}(x,\delta)\, D_{y,i}g \text{ exist and } D_{y,i}g=D_{y,i}f-D_i f(x) \tag{16.298}$$

or if we define

$$D_i g: B_{\|\|_{\max}}(x,\delta)\to L(X_i,Y) \text{ by } D_i g(y)=D_{y,i}g$$

then

$$\forall y\in B_{\|\|_{\max}}(x,\delta)\ D_i g(y)=D_i f(y)-D_i f(x) \tag{16.299}$$

So $D_i g=(D_i f)_{|B_{\|\|_{\max}}(x,\delta)}-C_{D_i f(x)}$ and, as $D_i f$ is continuous it follows from [theorem: 14.135] that

$$D_i g \text{ is continuous} \tag{16.300}$$

Further we have

$$g(x)\underset{[\text{eq: }16.297]}{=} f(x)-f(x)-\sum_{j=1}^n D_j f(x_j-x_j)=0 \tag{16.301}$$

Take $\varepsilon \in \mathbb{R}^+$. As $D_i f$ is continuous at x there exist a $\delta_{i,2} \in \mathbb{R}^+$ such that $\forall y \in U$ with $\|y - x\|_{\max} < \delta_{i,2}$ we have $\|D_i f(y) - D_i f(x)\|_{L(X_i,Y)} < \frac{\varepsilon}{n}$. Define $\delta_2 = \min(\delta_1, \min(\{\delta_{2,j} | j \in \{1,\ldots,n\}\})) \in \mathbb{R}^+$ then we have [using 16.299] that

$$\forall i \in \{1,\ldots,n\} \ \forall y \in B_{\|\|_{\max}}(x,\delta_2) \text{ we have } \|D_i g(y)\|_{L(X_i,Y)} < \frac{\varepsilon}{n} \tag{16.302}$$

Let $h \in B_{\|\|_{\max}}(0,\delta_2)$ then $\forall i \in \{0,\ldots,n\}$ we have that

$$\|(x + h^{\{i\}}) - x\|_{\max} = \|h^{\{i\}}\|_{\max} \leqslant_{[\text{lemma: } 16.207]} \|h\| < \delta_2$$

so that

$$\forall i \in \{0,\ldots,n\} \ x + h^{\{i\}} \in B_{\|\|_{\max}}(x,\delta_2) \text{ and } \|D_i g(x + h^{\{i\}})\|_{L(X_i,Y)} < \frac{\varepsilon}{n} \tag{16.303}$$

Next

$$\begin{aligned} g(x+h) &\underset{[\text{lemma: } 16.207]}{=} g(x + h^{\{n\}}) \\ &\underset{[\text{eq: } 16.301]}{=} g(x + h^{\{n\}}) - g(0) \\ &\underset{[\text{lemma: } 16.207]}{=} g(x + h^{\{n\}}) - g(x + h^{\{0\}}) \\ &\underset{[\text{theorem: } 11.21]}{=} \sum_{i=1}^{n} (g(x + h^{\{i\}}) - g(x + h^{\{i-1\}})) \end{aligned}$$

so that

$$\|g(x+h)\|_Y \leqslant \sum_{i=1}^{n} \|g(x + h^{\{i\}}) - g(x + h^{\{i-1\}})\| \tag{16.304}$$

Let $i \in \{1,\ldots,n\}$ then as $\|x_i + h_1 - x_i\|_i = \|h_i\|_i < \delta_2$ so that $x_i + h_i, x_i \in B_{\|\|_i}(x_i,\delta)$. Using this together with [eqs: 16.298, 16.300, 16.303] allows use [lemma: 16.208] giving

$$\left\|g\left(I^{[i,x+h^{\{i\}}]}(x_i + h_i)\right) - g\left(I^{[i,x+h^{\{i\}}]}(x_i)\right)\right\|_Y \leqslant \|h_i\|_i \cdot \sup\left(\left\{\left\|D_i g\left(I^{[i,x+h^{\{i\}}]}(x_i + h_i - t \cdot h_i)\right)\right\| | t \in [0,1]\right\}\right)$$

As $I^{[i,x+h^{\{i\}}]}(x + h_i) \underset{[\text{lemma: } 16.207]}{=} x + h^{\{i\}}$ and $I^{[i,x+h^{\{i\}}]}(x_i) = x + h^{\{i-1\}}$ it follows from [eq: 16.303] that

$$\|g((+h^{\{i\}})) - g(x + h^{\{i-1\}})\|_Y \leqslant \|h_i\| \cdot \sup\left(\left\{\left\|D_i g\left(I^{[i,x+h^{\{i\}}]}(x_i + h_i - t \cdot h_i)\right)\right\|_{L(X_i,Y)} | t \in [0,1]\right\}\right) \tag{16.305}$$

As for $I^{[i,x+h^{\{i\}}]}(x_i + h_i - t \cdot h_i)$ we have that for $k \in \{1,\ldots,n\}$ that

***k* = *i*.**

$$\begin{aligned} \left\|\left(I^{[i,x+h^{\{i\}}]}(x_i + h_i - t \cdot h_i) - x\right)_k\right\|_k &= \left\|\left(I^{[i,x+h^{\{i\}}]}(x_i + h_i - t \cdot h_i) - x\right)_k\right\|_k \\ &= \|x_i + h_i - t \cdot h_i - x_i\|_i \\ &= \|h_i \cdot (1-t)\|_i \\ &\underset{0 \leqslant 1-t}{=} (1-t) \cdot \|h\|_i \\ &< \|h_i\|_i \\ &\leqslant \|h\|_{\max} \\ &< \delta_2 \end{aligned}$$

***k* ≠ *i*.**

$$\begin{aligned} \left\|\left(I^{[i,x+h^{\{i\}}]}(x_i + h_i - t \cdot h_i) - x\right)_k\right\|_k &= \|(x + h^{\{i\}})_k - x_k\|_k \\ &= \|x_k + (h^{\{i\}})_k - x_k\|_k \\ &= \|(h^{\{i\}})_k\|_k \\ &\leqslant \|h^{\{i\}}\|_{\max} \\ &\leqslant_{[\text{lemma: } 16.207]} \|h\|_{\max} \\ &< \delta_2 \end{aligned}$$

so that

$$\begin{array}{rcl}\left\|I^{[i,x+h^{\{i\}}]}(x_i+h_i-t\cdot h_i)\right\|_{\max} & = & \max\left(\left\{\left\|\left(I^{[i,x+h^{\{i\}}]}(x_i+h_i-t\cdot h_i)-x\right)_k\right\|_k | k\in\{1,\ldots,n\}\right\}\right)\\ & < & \delta_2\end{array}$$

Hence $I^{[i,x+h^{\{i\}}]}(x_i+h_i-t\cdot h_i)\in B_{\|\|_{\max}}(x,\delta_2)$ so that by [eq: 16.302] that

$$\left\|D_i g\left(I^{[i,x+h^{\{i\}}]}(x_i+h_i-t\cdot h_i)\right)\right\|_{L(X_i,Y)}<\frac{\varepsilon}{n}$$

which by [eq: 16.305] proves that

$$\|g((+h^{\{i\}}))-g(x+h^{\{i-1\}})\|_Y<\|h_i\|_i\cdot\frac{\varepsilon}{n}\leqslant\|h\|_{\max}\cdot\frac{\varepsilon}{n}$$

Substituting the above in [eq: 16.304] results then in

$$\|g(x+h)\|_Y\leqslant\sum_{i=1}^n\|g(x+h^{\{i\}})-g(x+h^{\{i-1\}})\|<\sum_{i=1}^n\frac{\varepsilon}{n}\cdot\|h\|_{\max}=\varepsilon\cdot\|h\|_{\max}\tag{16.306}$$

Further

$$\begin{array}{rcl}g(x+h) & = & f(x+h)-f(x)-\sum_{i=1}^n D_i f(x)((x_i+h_i)-x_i)\\ & = & f(x+h)-f(x)-\sum_{i=1}^n D_i f(x)(h_i)\end{array}$$

so that

$$\left\|f(x+h)-f(x)-\sum_{i=1}^n D_i f(h_i)\right\|_Y\leqslant\varepsilon\cdot\|h\|_{\max}$$

which proves that

$$f\text{ is Fréchet differentiable at }x\text{ and }D_x f(h)=\sum_{i=1}^n D_i f(x)(h_i)$$

As $x\in U$ was chosen arbitrary, f is Fréchet differentiable on U the function

$$Df: U\rightarrow L(X,Y)\text{ by }Df(y)=D_y f$$

is well defined and

$$Df(x)=\sum_{i=1}^n D_i f(x)\circ\pi_i$$

Finally we have to prove that Df is continuous. Let $x\in U$ and $\varepsilon\in\mathbb{R}^+$ as $D_i f$ is continuous there exist a $\delta\in\mathbb{R}^+$ such that $\|D_i f(x)-D_i(f)\|_{L(X,Y)}<\frac{\varepsilon}{n}$. Then we have

$$\begin{array}{rcl}\|(Df(x)-Df(y))(h)\|_Y & = & \|Df(x)(h)-Dg(y)(h)\|_Y\\ & = & \left\|\sum_{i=1}^n D_i f(x)(h)-\sum_{i=1}^n D_i f(y)(h)\right\|_Y\\ & = & \left\|\sum_{i=1}^n(D_i f(x)(h)-D_i f(y)(h))\right\|_Y\\ & = & \left\|\sum_{i=1}^n(D_i f(x)-D_i f(y))(h)\right\|_Y\\ & \leqslant & \sum_{i=1}^n\|(D_i f(x)-D_i f(y))(h)\|_Y\\ & = & \left(\sum_{i=1}^n\|D_i f(x)-D_i f(y)\|_{L(X,Y)}\cdot\|h\|_{\max}\right)\\ & \leqslant & \sum_{i=1}^n\left(\frac{\varepsilon}{n}\cdot\|h\|_{\max}\right)\\ & = & \varepsilon\cdot\|h\|_{\max}\end{array}$$

so that by [theorem: 14.181] $\|Df(x) - Df(y)\|_{L(X,Y)} < \varepsilon$. Hence Df is continuous at x and as x was chosen arbitrary we have that Df is continuous. So we have finally that

$$f \text{ is of class } C^1$$ □

Combining the above theorem with [theorem: 16.203] gives the following equivalence.

Corollary 16.210. *Let $n \in \mathbb{N} \setminus \{1\}$, $\{\langle X_i, \|\|_i\rangle\}_{i \in \{1,\ldots,n\}}$ a finite family of normed space, $\langle Y, \|\|_Y\rangle$ a* ***Banach*** *space, $\langle X, \|\|_{\max}\rangle$ the normed space over $\mathbb{R}$ based on the product $X = \prod_{i \in \{1,\ldots,n\}} X_i$ with the maximum norm $\|\|_{\max}$ [see definition: 14.94], U a open set in X, $f: U \rightarrow Y$ a function then the following are equivalent:*

1. *f is of class C^1*
2. *$\forall x \in X$ we have that $\forall i \in \{1,\ldots,n\}$ f has a i-partial derivative $D_{x,i}f$ at x and the function*

   $$D_i f: U \rightarrow L(X_i; Y) \text{ defined by } D_i f(y) = D_{y,i} f$$

   is continuous.

Proof.

⇒. Using [theorem: 16.203] for the case $m = 1$ proves that $\forall i \in \{1,\ldots,n\}$ f has a i-partial derivative $D_{x,i}f$ at x and the function

$$D_i f: U \rightarrow L(X_i; Y) \text{ defined by } D_i f(y) = D_{y,i} f$$

is continuous.

⇐. This is proved in the previous theorem [theorem: 16.209]. □

16.5.3 Higher order derivatives

Definition 16.211. *Let $n \in \mathbb{N} \setminus \{1\}$, $m \in \mathbb{N}$ be the vector space $\langle \mathbb{K}^n, \|\|_n\rangle$ the normed space with the maximum norm $\|\|_n$, $\langle Y, \|\|_Y\rangle$ a normed space, U a open set in $\mathbb{K}^n$, $x \in U$, $\{i_j\}_{j \in \{1,\ldots,m\}} \subseteq \{1,\ldots,n\}$ and $f: U \rightarrow Y$ then f has a* ***partial derivative*** *$\partial^m_{x,i_m \ldots i_1} f$ of* ***order m*** *at x if for $m \in \mathbb{N}$ we have either:*

m = 1. *$\partial_{x,i_1} f \in \mathbb{K}$ exist [see definition: 16.53].*

1 < m. *If there exist a open set V with $x \in V \subseteq U$ such that $\forall y \in V$ $\partial_{y,i_{m-1}\ldots i_1} f$ exist and the function*

$$\partial_{i_{m-1}\ldots i_1} f: U \rightarrow \mathbb{K} \text{ defined by } \partial_{i_{m-1}\ldots i_1} f(y) = \partial_{y,i_{m-1}\ldots i_1} f$$

has a i_m-partial derivative at x and $\partial_{x,i_m \ldots i_1} f = \partial_{x,i_m}(\partial_{i_{m-1}\ldots i_1} f)$.

***Note* 16.212.** *[theorem: 16.55] ensures that the above definition is independent of the chosen V.*

Definition 16.213. *Let $n \in \mathbb{N} \setminus \{1\}$, $m \in \mathbb{N}$ be the vector space $\langle \mathbb{K}^n, \|\|_n\rangle$ the normed space with the maximum norm $\|\|_n$, $\langle Y, \|\|_Y\rangle$ a normed space, U a open set in $\mathbb{K}^n$, $\{i_j\}_{j \in \{1,\ldots,m\}} \subseteq \{1,\ldots,n\}$ and $f: U \rightarrow Y$ then f has a* ***$(i_m \ldots i_1)$-partial derivative*** *$\partial^m_{i_m \ldots i_1} f$ of* ***order m on U*** *if $\forall x \in U$ f has a partial derivative $\partial^m_{x,i_m \ldots i_1} f$ at x.*

Just as there is a relation between partial derivatives and partial differentials [see theorem: 16.54] the same applies to partial derivatives and differentials of higher order.

Theorem 16.214. *Let $n \in \mathbb{N} \setminus \{1\}$, $m \in \mathbb{N}$ be the vector space $\langle \mathbb{K}^n, \|\|_n\rangle$ the normed space with the maximum norm $\|\|_n$, $\langle Y, \|\|_Y\rangle$ a normed space, U a open set in $\mathbb{K}^n$, $x \in U$, $\{i_j\}_{j \in \{1,\ldots,m\}} \subseteq \{1,\ldots,n\}$ and $f: U \rightarrow Y$ a function that has a partial differential $D^{[m]}_{x,i_m \ldots i_1} f$ of order m at x then f has a partial derivative $\partial^m_{x,i_m \ldots i_1} f$ of order m and $\partial^m_{x,i_m \ldots i_1} f = D^{[m]}_{x,i_m \ldots i_1} f\left(\underbrace{1:\ldots:1}_{m}\right)$.*

Proof. We prove this by induction so let

$$\mathcal{S} = \left\{m \in \mathbb{N} | \text{If } \{i_j\}_{j \in \{1,\ldots,m\}} \subseteq \{1,\ldots,n\},\ U \subseteq \mathbb{R}^n \text{ a open set and } f: U \rightarrow Y \text{ a function such that } D^m_{x,i_m,\ldots,i_n} f \text{ exist then } \partial^m_{x,i_m,\ldots,i_1} f \text{ exist with } \partial^m_{x,i_m,\ldots,i_1} f = D^m_{x,i_m,\ldots,i_1} f\left(\underbrace{1:\ldots:1}_{m}\right)\right\}$$

then we have:

$1\in S$. Given $\{i_j\}_{j\in\{1\}}\subseteq\{1,\ldots,n\}$, let $f:U\rightarrow Y$ be a function such that $D^{[1]}_{x,i_1\ldots i_1}f$ exist then by definition $D_{x,i_1}f$ exist and $D_{x,i_1}f=D^{[1]}_{x,i_1\ldots i_1}f$. Hence using [theorem: 16.54] $\partial_{x,i_1}f$ exist and $\partial_{x,i_1}f=D_{x,i_1}f(1)$. So by definition $\partial_{x,i_1\ldots i_1}f\underset{\text{def}}{=}\partial_{x,i_1}f=D^{[1]}_{x,i_1\ldots i_1}f(1)$ proving that $1\in S$.

$m\in S\Rightarrow m+1\in S$. Let $\{i_j\}_{j\in\{1,\ldots,m+1\}}\subseteq\{1,\ldots,n\}$ and $f:U\rightarrow Y$ a function such that $D^{[m+1]}_{x,i_{m+1}\ldots i_1}f$ exist. Then there exist a open set V with $x\in V\subseteq U$ such that $\forall y\in V\ D^{[m]}_{x,i_m\ldots i_1}f$ exist and the function

$$D^{[m]}_{i_m\ldots i_1}f:V\rightarrow L_m(\mathbb{K};Y)\text{ defined by }D^{[m]}_{i_m\ldots i_1}f(y)=D^{[m]}_{y,i_m\ldots i_1}f$$

has a i_{m+1}-partial derivative at x and $D^{[m+1]}_{x,i_{m+1}\ldots i_1}f=D_{x,i_m}\left(D^{[m]}_{i_m\ldots i_1}f\right)$. As $m\in S$ we have that $\forall y\in V\ \partial^m_{y,i_m\ldots i_1}f$ exist and $\partial^m_{y,i_m\ldots i_1}f=D^{[m]}_{y,i_m\ldots i_1}f\left(\underbrace{1:\ldots:1}_{m}\right)$. Define

$$\partial^m_{i_m\ldots i_1}f:V\rightarrow Y\text{ by }\partial^m_{i_m\ldots i_1}f(y)=\partial^m_{y,i_m\ldots i_1}f$$

Consider the evaluation operator defined by [see definition: 16.173

$$\mathrm{ev}_{\underbrace{1\ldots1}_{m}}:L_m(\mathbb{K};Y)\rightarrow Y\text{ where }\mathrm{ev}_{\underbrace{1\ldots1}_{m}}(L)=L\left(\underbrace{1:\ldots:1}_{m}\right)$$

then

$$\begin{aligned}
D^{[m]}_{y,i_m\ldots i_1}f\left(\underbrace{1:\ldots:1}_{m}\right) &= v_{\underbrace{1\ldots1}_{m}}\left(D^{[m]}_{y,i_m\ldots i_1}f\right)\\
&= v_{\underbrace{1\ldots1}_{m}}\left(D^{[m]}_{i_m\ldots i_1}f(y)\right)\\
&= \left(\mathrm{ev}_{\underbrace{1\ldots1}_{m}}\circ D^{[m]}_{i_m\ldots i_1}f\right)(y)\\
&= \left(\mathrm{ev}_{\underbrace{1\ldots1}_{n}}\circ D^{[m]}_{i_m\ldots i_1}f\right)(y)
\end{aligned}$$

So if $y\in V$ then

$$\partial_{i_m\ldots i_1}f(y)=\partial_{y,i_m\ldots i_1}f=D^{[m]}_{y,i_m\ldots i_1}f\left(\underbrace{1:\ldots:1}_{n}\right)=\left(v_{\underbrace{1\ldots1}_{n}}\circ D^{[m]}_{i_m\ldots i_1}f\right)(y)$$

proving that

$$\partial_{i_m\ldots i_1}f=\mathrm{ev}_{\underbrace{1\ldots1}_{n}}\circ D^{[m]}_{i_m\ldots i_1}f \tag{16.307}$$

As $D^{[m]}_{i_m\ldots i_1}f$ has a i_{m+1}-partial differential at x, and by [lemma: 16.174] $\mathrm{ev}_{\underbrace{1\ldots1}_{n}}$ is differentiable on $L_m(\mathbb{K};Y)$ it follows from the chain rule for partial differentiation [theorem: 16.52] that $\mathrm{ev}_{\underbrace{1\ldots1}_{n}}\circ D^{[m]}_{i_m\ldots i_1}f$ has a i_{m+1}-partial derivative at x and

$$\begin{aligned}
D_{x,i_{m+1}}\left(\mathrm{ev}_{\underbrace{1\ldots1}_{n}}\circ D^{[m]}_{i_m\ldots i_1}f\right) &= D_{D_{i_m\ldots i_1}(x)}\mathrm{ev}_{\underbrace{1\ldots1}_{n}}\circ D_{x,i_{i_{m+1}}}\left(D^{[m]}_{i_m\ldots i_1}f\right)\\
&\underset{\text{[lemma: 16.174]}}{=} \mathrm{ev}_{\underbrace{1\ldots1}_{m}}\circ D_{x,i_{i_{m+1}}}\left(D^{[m]}_{i_m\ldots i_1}f\right)\\
&= \mathrm{ev}_{\underbrace{1\ldots1}_{m}}\circ D^{[m+1]}_{x,i_{m+1}\ldots i_1}f\\
&= D^{[m+1]}_{x,i_{m+1}\ldots i_1}f\left(\underbrace{1:\ldots:1}_{n}\right)
\end{aligned}$$

combining this with [eq: 16.307] proves that $\partial_{i_m\ldots i_1}f$ has a i_{m+1}-partial derivative at x and

$$D_{x,i_{m+1}}(\partial_{i_m\ldots i_1}f)=D^{[m+1]}_{x,i_{m+1}\ldots i_1}f\left(\underbrace{1:\ldots:1}_{n}\right) \tag{16.308}$$

Using [theorem: 16.54] it follows that $\partial_{i_m \ldots i_1} f$ has a i_{m+1} partial derivative at x and

$$\begin{array}{rcl}\partial_{x,i_{m+1}\ldots i_1} f & = & \partial_{x,i_{m+1}}(\partial_{i_m\ldots i_1} f)\\ & = & D_{x,i_{m+1}}(\partial_{i_m\ldots i_1} f)(1)\\ & \underset{\text{[eq: 16.308]}}{=} & D^{[m+1]}_{x,i_{m+1}\ldots i_1} f\left(\underbrace{1:\ldots:1}_{m}\right)(1)\\ & \underset{\text{[lemma: 16.78]}}{=} & D^{[m+1]}_{x,i_{m+1}\ldots i_1} f\left(\underbrace{1:\ldots:1}_{m+1}\right)\end{array}$$

proving that

$$m+1\in S$$

□

So we have proved that higher order partial differentiation implies higher order derivation, next we prove the opposite.

Definition 16.215. *Let $m\in\mathbb{N}$, $\langle\mathbb{K}^m,\|\|_m\rangle$ the normed space with the maximum norm $\|\|_m$, $\langle Y,\|\|\rangle$ a normed space and $a\in Y$ then we define*

$$h_{m,a}:\mathbb{K}^m\rightarrow\mathbb{K} \text{ by } h_{m,a}(x)=\left(\prod_{i=1}^m x_i\right)\cdot a$$

Using [example: 14.188] it follows that $h_{m,a}\in L^m(\mathbb{K},\mathbb{K})$ so that we can use

$$\mathcal{I}_{m,\mathbb{K},Y}: L_m(\mathbb{K};Y)\rightarrow L^m(\mathbb{K};Y) \text{ defined in [theorem: 16.85]}$$

to finally define $\mathcal{P}[m,a]$ by

$$\mathcal{P}[m,x]=(\mathcal{I}_{m,\mathbb{K},Y})^{-1}(h_{m,a})\in L_m(\mathbb{K};Y)$$

***Note* 16.216.** *If $x=(x_1,\ldots,x_m)\in\mathbb{K}^n$ then we have*

$$\mathcal{P}[m,a](x_1:\ldots:x_1)=(\mathcal{I}_{m,\mathbb{K},Y})^{-1}(h_{m,a})(x_1:\ldots:x_m)\underset{\text{[theorem: 16.85]}}{=} h_{m,a}(x_1,\ldots,x_m)=\left(\prod_{i=1}^m x_i\right)\cdot a$$

so that

$$\mathcal{P}[m,a]\left(\underbrace{1:\ldots:1}_{m}\right)=a$$

Lemma 16.217. *Let $m\in\mathbb{N}$, $\langle\mathbb{K}^m,\|\|_m\rangle$ the normed space with the maximum norm $\|\|_m$, $\langle Y,\|\|\rangle$ a normed space then if we define*

$$\mathcal{P}[m]: Y\rightarrow L_m(\mathbb{K};Y) \text{ by } \mathcal{P}[m](y)=\mathcal{P}[m,y]$$

then

$$\mathcal{P}[m]\in L(Y;L_m(\mathbb{K};Y))$$

so that by [example: 16.116] $\mathcal{P}[m]$ is ∞-times differentiable on Y and $\forall y\in Y$

$$D_y\mathcal{P}[m]=D^{[1]}_y\mathcal{P}[m]=\mathcal{P}[m]$$

Proof. First we prove linearity, so let $\alpha\in\mathbb{K}$ and $x,y\in Y$ then we have for $h\in\mathbb{K}^m$ that

$$\begin{array}{rcl}(\mathcal{P}[m](x+\alpha\cdot y))(h_1:\ldots:h_m) & = & \mathcal{P}[m,x+\alpha\cdot y](h_1:\ldots:h_m)\\ & = & \left(\prod_{i=1}^m h_i\right)\cdot(x+\alpha\cdot y)\\ & = & \left(\prod_{i=1}^m h_i\right)\cdot x+\alpha\cdot\left(\left(\prod_{i=1}^m h_i\right)\cdot y\right)\\ & = & (\mathcal{P}[m](x))(h_1:\ldots:h_m)+(\alpha\cdot\mathcal{P}[m](y))(h_1:\ldots:h_m)\\ & \underset{\text{[lemma: 16.79]}}{=} & (P[m](x)+\alpha\cdot\mathcal{P}[m](y))(h_1:\ldots:h_m)\end{array}$$

proving by [lemma: 16.82] that $\mathcal{P}[m](x+\alpha\cdot y)=P[m](x)+\alpha\cdot\mathcal{P}[m](y)$. Hence

$$\mathcal{P}[m]\in\operatorname{Hom}(Y;L_m(\mathbb{K};Y))$$

Further we have

$$\begin{aligned}\|(\mathcal{P}[m](x))(h_1{:}\ldots{:}h_m)\|_Y &= \|\mathcal{P}[m,x](h_1\ldots h_m)\|_Y\\ &= \left\|\left(\prod_{i=1}^m h_i\right)\cdot x\right\|_Y\\ &= \left|\prod_{i=1}^m h_i\right|\cdot\|x\|_Y\\ &= \|x\|_Y\cdot\prod_{i=1}^m |h_i|\end{aligned}$$

so that by [lemma: 16.83] we have $\|\mathcal{P}[m](x)\|_{L_m(\mathbb{K};Y)}\leqslant\|x\|_Y$. Hence by [theorem: 14.174] $\mathcal{P}[m]$ is continuous proving that

$$\mathcal{P}[m]\in L(Y,L_n(\mathbb{K};Y))$$

□

Theorem 16.218. *Let $n\in\mathbb{N}\setminus\{1\}$, $m\in\mathbb{N}$ be the vector space $\langle\mathbb{K}^n,\|\|_n\rangle$ the normed space with the maximum norm $\|\|_n$, $\langle Y,\|\|_Y\rangle$ a normed space, U a open set in $\mathbb{K}^n$, $x\in U$, $\{i_j\}_{j\in\{1,\ldots,m\}}\subseteq\{1,\ldots,n\}$ and $f{:}U\rightarrow Y$ has a partial derivative $\partial^m_{x,i_m\ldots i_1}$ or order m at then f has a partial differential $D^{[m]}_{x,i_m\ldots i_1}f$ of order m at x and $D^{[m]}_{x,i_m\ldots i_1}f=\mathcal{P}[m,\partial_{x,i_m\ldots i_1}f]$.*

Proof. We prove this by induction, so define

$S=\{m\in\mathbb{N}|\text{If }\{i_j\}_{j\in\{1,\ldots,m\}}\subseteq\{1,\ldots,n\},\ U\subseteq\mathbb{K}^n$ a open set, $x\in U$ and $f{:}U\rightarrow Y$a function such that $\partial^m_{x,i_m\ldots i_1}f$ exists the $D^{[m]}_{x,i_m\ldots i_1}f$ exist and $D^{[m]}_{x,i_m\ldots i_1}f=\mathcal{P}[m,\partial_{x,i_m\ldots i_1}f]\}$

then we have:

$1\in S$. Given $\{i_j\}_{j\in\{1\}}\subseteq\{1,\ldots,n\}$, let $f{:}U\rightarrow Y$ be a function such that $\partial^{[1]}_{x,i_1\ldots\partial_{i_1}}f$ exist. Then by definition $\partial_{x,i_1}f$ exist and $\partial_{x,i_1}f=\partial^1_{x,i_1\ldots i_1}f$. Hence using [theorem: 16.54] $D_{x,i_1}f$ exist and is defined by

$$D_{x,i_1}f{:}\mathbb{K}\rightarrow Y\text{ by }D_{x,i_1}f(h)=h\cdot\partial_{x,i_1}f$$

Define $(h_1,\ldots,h_1)\in\mathbb{K}^1$ by $h_1=h$ then, as by definition $D_{x,i_1}f(h)=D^{[1]}_{x,i_1\ldots i_1}f\left(\underbrace{h{:}\ldots{:}h}_{1}\right)$ and $\partial_{x,i_1}f=\partial^1_{x,i_1\ldots i_1}f$, it follows that $D^{[1]}_{x,i_1\ldots i_1}f$ exist and $D^{[1]}_{x,i_1\ldots i_1}f=\mathcal{P}\big[1,\partial^{[1]}_{x,i_1\ldots i_1}f\big]$. So we must conclude that $1\in S$.

$m\in S\Rightarrow m+1\in S$. Let $\{i_j\}_{j\in\{1,\ldots,n\}}\subseteq\{1,\ldots,n\}$ and $f{:}U\rightarrow Y$ a function such that $\partial^{[m+1]}_{x,i_{m+1}\ldots i_1}f$ exist. Then by definition there exist a open set V with $x\in V\subseteq U$ such that $\forall y\in V\ \partial^m_{y,i_m\ldots i_1}f$ exist and the function

$$\partial^m_{x,y_1\ldots y_m}f{:}V\rightarrow L_m(\mathbb{K};Y)\text{ defined by }\partial^m_{i_m\ldots i_1}f(y)=\partial^m_{y,i_m\ldots i_1}f$$

has a i_{m+1}-derivative at x and

$$\partial^{m+1}_{x,i_{m+1}\ldots i_1}f=\partial_{x,i_{m+1}}(\partial^m_{x,i_m\ldots i_1}f)\tag{16.309}$$

As $m\in S$ we have that $\forall y\in V\ D^{[m]}_{y,i_m\ldots i_1}f$ exist and

$$\begin{aligned}D^{[m]}_{y,i_m\ldots i_1}f &= \mathcal{P}\big[m,\partial^{[m]}_{y,i_m\ldots i_1}f\big]\\ &= \mathcal{P}\big[m,\partial^{[m]}_{i_m\ldots i_1}f(y)\big]\\ &\underset{[\text{lemma: }16.217]}{=} \mathcal{P}[m]\big(\partial^{[m]}_{i_m\ldots i_1}f(y)\big)\\ &= \big(\mathcal{P}[m]\circ\partial^{[m]}_{i_m\ldots i_1}f\big)(y)\end{aligned}$$

So if we define

$$D^{[m]}_{i_m\ldots i_1}f{:}V\rightarrow L_m(\mathbb{K};Y)\text{ by }D^{[m]}_{i_m\ldots i_1}f(y)=D^{[m]}_{y,i_m\ldots i_1}f$$

then $D^{[m]}_{i_m\ldots i_1}f(y)=D^{[m]}_{y,i_m\ldots i_1}f=\big(\mathcal{P}[m]\circ\partial^{[m]}_{i_m\ldots i_1}f\big)(y)$, proving that

$$D^{[m]}_{i_m\ldots i_1}f=\mathcal{P}[m]\circ\partial^{[m]}_{i_m\ldots i_1}f\tag{16.310}$$

As $\partial^m_{x,y_1\ldots y_m} f$ has a i_{m+1}-partial derivative at x we can use [theorem: 16.54] to prove that $\partial^m_{x,y_1\ldots y_m} f$ has a i_{m+1}-partial differential at x and that the function $D_{x,i_{m+1}}(\partial^m_{i_m\ldots i_1} f) \in L(\mathbb{K}, Y)$ is defined by

$$D_{x,i_{m+1}}(\partial^m_{y_m\ldots y_1} f): \mathbb{K} \to Y \text{ where } D_{x,i_{m+1}}(\partial^m_{i_m\ldots i_1} f)(h) = h \cdot \partial_{x,i_{m+1}}(\partial^m_{i_m\ldots i_1} f) \tag{16.311}$$

As $\mathcal{P}[m]$ is Fréchet differentiable on Y [see lemma: 16.217] and $\partial^m_{x,y_1\ldots y_m} f$ has a i_{m+1}-partial differential at x we can apply the chain rule for partial differentials [see theorem: 16.52] resulting in the fact that $\mathcal{P}[m] \circ \partial^{[m]}_{i_m\ldots i_1} f$ has a i_{m+1}-partial differential at x and

$$\begin{aligned} D_{x,i_{m+1}}(\mathcal{P}[m] \circ \partial^{[m]}_{i_m\ldots i_1} f) &= D_{\partial^{[m]}_{i_m\ldots i_1} f(x)} \mathcal{P}[m] \circ D_{x,i_{m+1}} \partial^{[m]}_{i_m\ldots i_1} f \\ &\underset{[\text{lemma: } 16.217]}{=} \mathcal{P}[m] \circ D_{x,i_{m+1}} \partial^{[m]}_{i_m\ldots i_1} f \end{aligned}$$

Using [eq: 16.310] and the above we get that $D^{[m]}_{i_m\ldots i_1} f$ has a i_{m+1}-partial differential at x and

$$D_{x,i_{m+1}}(D^{[m]}_{i_m\ldots i_1} f) = \mathcal{P}[m] \circ D_{x,i_{m+1}} \partial^{[m]}_{i_m\ldots i_1} f$$

Hence by definition f has a partial differential $D^{[m+1]}_{x,i_{m+}\ldots i_1} f$ and

$$D^{[m+1]}_{x,i_{m+1}\ldots i_1} f = \mathcal{P}[m] \circ D_{x,i_{m+1}} \partial^{[m]}_{i_m\ldots i_1} f \tag{16.312}$$

Now given $h \in \mathbb{K}^{m+1}$ we have

$$\begin{aligned} D^{[m+1]}_{x,i_{m+1}\ldots i_1} f(h_1 : \ldots : h_{m+1}) &\underset{[\text{eq: } 16.312]}{=} \\ (\mathcal{P}[m] \circ D_{x,i_{m+1}} \partial^{[m]}_{i_m\ldots i_1} f)(h_1 : \ldots : h_{m+1}) &= \\ ((\mathcal{P}[m] \circ D_{x,i_{m+1}} \partial^{[m]}_{i_m\ldots i_1} f)(h_1))(h_2 : \ldots : h_{m+1}) &= \\ (\mathcal{P}[m](D_{x,i_{m+1}}(\partial^{[m]}_{i_m\ldots i_1} f)(h_1)))(h_2 : \ldots : h_{m+1}) &\underset{[\text{eq: } 16.311]}{=} \\ (\mathcal{P}[m](h_1 \cdot \partial_{x,i_{m+1}}(\partial^m_{i_m\ldots i_1} f)))(h_2 : \ldots : h_{m+1}) &\underset{\mathcal{P}[m] \in L(Y; L_m(\mathbb{K};Y))}{=} \\ h_1 \cdot (\mathcal{P}[m](h_1 \cdot \partial_{x,i_{m+1}}(\partial^m_{i_m\ldots i_1} f)))(h_2 : \ldots h_{m+1}) &= \\ \left(h_1 \cdot \prod_{i=1}^m h_{i+1}\right) \cdot \partial_{x,i_{m+1}}(\partial^m_{i_m\ldots i_1} f) &= \\ \left(\prod_{i=1}^{m+1} h_i\right) \cdot \partial_{x,i_{m+1}}(\partial^m_{i_m\ldots i_1} f) &\underset{[\text{eq: } 16.309]}{=} \\ \left(\prod_{i=1}^{m+1} h_i\right) \cdot \partial_{x,i_{m+1}\ldots i_1} f &= \\ \mathcal{P}[m+1, \partial_{x,i_{m+1}\ldots i_1} f](h_1 : \ldots : h_{m+1}) & \end{aligned}$$

Hence by [lemma: 16.82] we have

$$D^{[m+1]}_{x,i_{m+1}\ldots i_1} f = \mathcal{P}[m+1, \partial_{x,i_{m+1}\ldots i_1} f]$$

proving

$$m+1 \in S$$

□

We can now combine theorems [theorem: 16.214] and [theorem: 16.218] to find the following equivalences.

Theorem 16.219. *Let $n \in \mathbb{N} \setminus \{1\}$, $m \in \mathbb{N}$ be the vector space $\langle \mathbb{K}^n, \|\|_n \rangle$ the normed space with the maximum norm $\|\|_n$, $\langle Y, \|\|_Y \rangle$ a normed space, U a open set in $\mathbb{K}^n$, $x \in U$, $\{i_j\}_{j \in \{1,\ldots,m\}} \subseteq \{1, \ldots, n\}$ and $f: U \to Y$ a function then we have*

$$f \textit{ has a partial differential } D^{[m]}_{x,i_m\ldots i_1} f \textit{ at } x$$

$$\Updownarrow$$

$$f \textit{ has a partial} \text{ derivate } \partial^m_{x,i_m\ldots i_1} f$$

Further if f has a partial differential or derivative at x then

$$\partial^m_{x,i_m\ldots i_1} f = D^{[m]}_{x,i_m\ldots i_1} f(\underbrace{1 : \ldots : 1}_{m})$$

and

$$D^{[m]}_{x,i_m\ldots i_1}f\in L_m(\mathbb{K};Y) \text{ is defined by } D^{[m]}_{x,i_m\ldots i_1}f=\mathcal{P}[m,\partial^m_{x,i_m\ldots i_1}]$$

so that for $h\in\mathbb{K}^m$ we have

$$D^{[m]}_{x,i_m\ldots i_1}f(h_1:\ldots h_m)=\left(\prod_{i=1}^m h_i\right)\cdot\partial^m_{x,i_m\ldots i_1}f$$

Proof. This follows directly from [theorems: 16.214 and 16.218]. □

Theorem 16.220. *Let $n\in\mathbb{N}\setminus\{1\}$, $m\in\mathbb{N}$ be the vector space $\langle\mathbb{K}^n,\|\|_n\rangle$ the normed space with the maximum norm $\|\|_n$, $\langle Y,\|\|_Y\rangle$ a normed space, U a open set in $\mathbb{K}^n$, $\{i_j\}_{j\in\{1,\ldots,m\}}\subseteq\{1,\ldots,n\}$ and $f:U\to Y$ a function then we have the following equivalences:*

1. *f has a $(i_m\ldots i_1)$-partial differential on U and*

$$D^{[m]}_{i_m\ldots i_1}f:U\to L_m(\mathbb{K};Y) \text{ defined by } D^{[m]}_{i_m\ldots i_1}f(x)=D^{[m]}_{x,i_m\ldots i_1}f$$

   is continuous.

2. *f has a $(i_m\ldots i_1)$-partial derivative on U and*

$$\partial_{i_m\ldots i_1}f:U\to L_m(\mathbb{K};Y) \text{ defined by } \partial_{i_m\ldots i_1}f(x)=\partial_{x,i_m\ldots i_1}f$$

   is continuous.

Proof.

⇒. Using the previous theorem [theorem: 16.219] it follows that f has a $(i_m\ldots i_1)$-partial derivative on U so we must only prove continuity. Let $x\in U$ then we have by [theorem: 16.219] that

$$\begin{array}{rcl}\partial^m_{i_1\ldots i_m}f(x)=\partial^m_{x,i_m\ldots i_1}f & = & D^{[m]}_{x,i_m\ldots i_1}f\Big(\underbrace{1:\ldots:1}_{m}\Big)\\ & \underset{[\text{definition: } 16.173]}{=} & \mathrm{ev}_{\underbrace{1\ldots1}_{m}}\big(D^{[m]}_{x,i_m\ldots i_1}f\big)\\ & = & \mathrm{ev}_{\underbrace{1\ldots1}_{m}}\big(D^{[m]}_{i_m\ldots i_1}f(x)\big)\\ & = & \Big(\mathrm{ev}_{\underbrace{1\ldots1}_{m}}\circ D^{[m]}_{i_m\ldots i_1}f\Big)(x)\end{array}$$

proving that $\partial^m_{i_1\ldots i_m}f=\mathrm{ev}_{\underbrace{1\ldots1}_{m}}\circ D^{[m]}_{i_m\ldots i_1}f$. As by [lemma: 16.174] $\mathrm{ev}_{\underbrace{1\ldots1}_{m}}$ is continuous and $D^{[m]}_{i_m\ldots i_1}f$ is continuous it follows that $\partial^m_{i_1\ldots i_m}f$ is continuous.

⇐. Using the previous theorem [theorem: 16.219] it follows that f has a $(i_m\ldots i_1)$-partial differential on U so we must only prove continuity. Let $x\in U$ then we have by [theorem: 16.219] that

$$\begin{array}{rcl}D^{[m]}_{i_m\ldots i_1}f(x) & = & D^{[m]}_{x,i_m\ldots i_1}f\\ & = & \mathcal{P}[m,\partial_{x,\partial_m\ldots\partial_1}f]\\ & = & \mathcal{P}[m](\partial_{x,\partial_m\ldots\partial_1}f)\\ & = & \mathcal{P}[m](\partial_{\partial_m\ldots\partial_1}f(x))\\ & = & (\mathcal{P}[m]\circ\partial_{\partial_m\ldots\partial_1}f)(x)\end{array}$$

proving that $D^{[m]}_{i_m\ldots i_1}f=\mathcal{P}[m]\circ\partial_{\partial_m\ldots\partial_1}f$. As by [lemma: 16.217] is continuous and $\partial_{\partial_m\ldots\partial_1}f$ is continuous it follows that $D^{[m]}_{i_m\ldots i_1}f$ is continuous. □

Theorem 16.221. *Let $n\in\mathbb{N}\setminus\{1\}$, $m\in\mathbb{N}$, $\langle\mathbb{K}^n,\|\|\rangle$ be the vector space $\mathbb{K}^n$ with the maximum norm, $\langle Y,\|\|_Y\rangle$ a normed space, U a open set in $\mathbb{K}^n$, $f:U\to Y$ a function that is m-times differentiable then $\forall\{i_j\}_{j\in\{1,\ldots,m\}}\subseteq\{1,\ldots,n\}$ $\partial^m_{i_m,\ldots,i_1}f$ exists and $\forall x\in U$ we have*

$$\partial^m_{i_m\ldots i_1}f=D^{[m]}_{x,i_m\ldots i_1}f\Big(\underbrace{1:\ldots:1}_{n}\Big)=D^{[m]}_xf(e_{i_m}:\ldots:e_{i_1})$$

where $\{e_i\}_{i\in\{1,\ldots,n\}}$ is the canonical basis on $\mathbb{K}^n$ defined by $(e_j)_j=\delta_{i,j}$.

Proof. Let $\{i_j\}_{j\in\{1,\ldots,m\}}\subseteq\{1,\ldots,n\}$. Using [theorem: 16.202] it follows that $D^{[m]}_{x,i_m\ldots i_1}f$ exist and for $h\in\mathbb{K}^m$

$$D^{[m]}_{x,i_m\ldots i_1}f(h_1:\ldots:h_n)=D^{[m]}_x f(I^{[i_m,0]}(h_1):\ldots:I^{[i_1,0]}(h_1))$$

Next by [theorem: 16.219] $\partial^m_{x,i_m\ldots i_1}f$ exist and

$$\partial_{x.i_m\ldots i_1}f=D^{[m]}_{x,i_m\ldots i_1}f\left(\underbrace{1:\ldots:1}_{n}\right)$$

so that

$$\begin{aligned}\partial_{x.i_m\ldots i_1}f &= D^{[m]}_x f\left(\underbrace{I^{[i_m,0]}(1):\ldots:I^{[i_1,0]}(1)}_{n}\right)\\ &\underset{[\text{theorem: }16.41]}{=} D^{[m]}_x f(\delta_{i_m,1}:\ldots:\delta_{i_1,m})\\ &= D^{[m]}_x f(e_{i_m}:\ldots:e_{i_1})\end{aligned}$$

□

Theorem 16.222. *Let $n\in\mathbb{N}\setminus\{1\}$, $m\in\mathbb{N}$, $\langle\mathbb{K}^n,\|\|\rangle$ be the vector space $\mathbb{K}^n$ with the maximum norm, $\langle Y,\|\|_Y\rangle$ a Banach space, a open set $U\subseteq\mathbb{K}^n$, $f:U\rightarrow Y$ a function that is of class C^m then $\forall x\in U$ and $\forall\{i_j\}_{j\in\{1,\ldots,m\}}\subseteq\{1,\ldots,n\}$ $\partial_{x,i_m\ldots i_1}f$ exist and the function*

$$\partial_{i_m\ldots i_1}f:U\rightarrow Y \text{ defined by } \partial_{i_m\ldots i_1}f(x)=\partial_{x,i_m\ldots i_1}f$$

is continuous. Further for every permutation $\sigma\in P_m$ we have

$$\partial_{i_{\sigma(m)}\ldots i_{\sigma(1)}}f(x)=\partial_{i_m\ldots i_1}f(x)$$

Proof. Using [theorem: 16.203] it follows that $\forall\{i_j\}_{j\in\{1,\ldots,m\}}\subseteq\{1,\ldots,n\}$ and $\forall x\in U$ $D_{x,i_m,\ldots,i_1}f$ exist and the function

$$D_{i_m\ldots,i_1}f:U\rightarrow L_m(X_i;Y) \text{ defined by } D_{i_m\ldots i_1}f(x)=D_{x,i_m\ldots i_1}f$$

is continuous. Hence by [theorem: 16.220] it follows that $\forall x\in U$ $\partial_{x,i_m\ldots i_1}f$ exist and the

$$\partial_{i_m\ldots i_1}f:U\rightarrow Y \text{ defined by } \partial_{i_m\ldots i_1}f(x)=\partial_{x,i_m\ldots i_1}f$$

is continuous. Further let $\sigma\in P_n$ then we have

$$\begin{aligned}\partial_{i_m\ldots i_1}f(x) &= \partial_{x,i_m\ldots i_1}f(x)\\ &\underset{[\text{theorem: }16.214]}{=} D^{[m]}_{x,i_m\ldots i_1}f\left(\underbrace{1:\ldots:1}_{n}\right)\\ &\underset{[\text{theorem: }16.204]}{=} D^{[m]}_{x,i_{\sigma(m)}\ldots i_{\sigma(1)}}f\left(\underbrace{1:\ldots:1}_{n}\right)\\ &= \partial_{x,i_{\sigma(m)}\ldots i_{\sigma(1)}}f(x)\end{aligned}$$

□

16.6 Inverse Function Theorem

The purpose of this section is to prove the inverse function theorem that gives a sufficient conditions for a function between Banach spaces to be invertible. First we introduce some concepts needed to prove this theorem. We start with toplinear isomorphism's.

Definition 16.223. *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. *$\cdot:\mathbb{K}\times X\rightarrow X$ defined by $(\alpha,x)\rightarrow\cdot(\alpha,x)=\alpha\cdot x$ is continuous*

Here we use the product topology on $X\times X$ based on the topology $\mathcal{T}$ of X and the product topology on $\mathbb{K}\times X$ based on the absolute norm topology $\mathcal{T}_{||}$ of $\mathbb{K}$ and the topology $\mathcal{T}$ of X.

Example 16.224. If $\langle X, \|\|\rangle$ is a normed space over $\mathbb{K}$ then $\langle X, \mathcal{T}_{\|\|}\rangle$ is a topological vector space.

Proof. This follows from [theorems: 14.142 and 14.155] □

Definition 16.225. *Let X, Y be topological vector spaces then a is a **toplinear isomorphism** is a function $L: X \rightarrow Y$ satisfying*

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*

In other words a toplinear isomorphism is linear homeomorphism.

***Note* 16.226.** *Using [theorem: 11.169] L^{-1} is also a continuous linear isomorphism, hence we have that $L^{-1} \in L(Y, X)$ if L is a toplinear isomorphism.*

Theorem 16.227. *Let $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$ be normed spaces then $\forall L \in L(X, Y)$ we have that L is a toplinear isomorphism if and only if L is a bijection and L^{-1} is continuous.*

Proof. This follows trivial from the definition of $L(X, Y)$ and a toplinear isomorphism. □

We can use a consequence of the open mapping theorem [see corollary: 14.388] to formulate the following corollary.

Theorem 16.228. *If $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$ are Banach spaces. If $L: X \rightarrow Y$ is a continuous isomorphism [in other words $L \in L(X, Y)$ and L is a bijection] then L is a toplinear isomorphism.*

Proof. By [corollary: 14.388] L is a homeomorphism so that L^{-1} is continuous so that by [theorem: 16.227] L is a toplinear isomorphism. □

Definition 16.229. *If $\langle X, \|\|\rangle$ is a normed space then $\mathcal{GL}(X)$ is defined by*

$$\mathcal{GL}(X) = \{X \in L(X, X) | X \text{ is a toplinear isomorphism}\} \subseteq L(X, X)$$

The set of toplinear isomorphisms on a Banach space forms a group under composition as stated in the next theorem.

Theorem 16.230. *If $\langle X, \|\|\rangle$ is a normed space then $\mathcal{GL}(X)$ is a group with the composition $\circ$ as the operator and the identity function Id_X [see definition: 2.47] as neutral element and for every $L \in \mathcal{GL}(X)$ L^{-1} as inverse of L.*

Proof. If $L_1, L_2 \in \mathcal{GL}(X)$ then by definition $L_1, L_2 \in L(X, Y)$ and L_1, L_2 are toplinear isomorphism. Hence by [theorem: 16.227 L_1, L_2 are bijections and L_1^{-1}, L_2^{-1} are continuous. As the composition of bijections is a bijection [see theorem: 2.75] $L_1 \circ L_2$ is bijective, further by [theorem: 14.185] we have $L_1 \circ L_2 \in L(X, Y)$. As $(L_1 \circ L_2)^{-1} \underset{[\text{theorem: } 2.75]}{=} L_2^{-1} \circ L_1^{-1}$ we have by [theorem: 14.138] that $(L_1 \circ L_2)^{-1}$ is continuous. Hence $L_1 \circ L_2 \in \mathcal{GL}(X)$ so that

$$\circ: \mathcal{GL}(X) \times \mathcal{GL}(X) \rightarrow \mathcal{GL}(X)$$

is well defined. Next we have:

identity. As Id_X is a bijection [see theorem: 2.64], $\mathrm{Id}_X \in L(X, X)$ [see theorem: 14.172] and $(\mathrm{Id}_X)^{-1} = \mathrm{Id}_X \Rightarrow \mathrm{Id}_X$ is continuous it follows that $\mathrm{Id}_X \in \mathcal{GL}(X)$. Finally we have trivially $\forall L \in L(X, X)$ $\mathrm{Id}_X \circ L = L = L \circ \mathrm{Id}_X$.

inverse element. If $L \in \mathcal{GL}(X)$ then L is a linear isomorphism and L^{-1} is continuous. As by [theorem: 11.169] L^{-1} is a linear isomorphism it follows that $L^{-1} \in L(X, X)$, further $(L^{-1})^{-1} = L$ so that $(L^{-1})^{-1}$ is continuous. proving that $L^{-1} \in \mathcal{GL}(X)$. Finally $L \circ L^{-1} = \mathrm{Id}_X = L^{-1} \circ L$.

associativity. This is trivial as composition of function is associative. □

We will need the following theorem about limits in $L(X, X)$ and composition

Theorem 16.231. *Let $\langle X, \|\|\rangle$ be a normed space, $L \in L(X, X)$ and $\{K_n\}_{n \in \mathbb{N}} \subseteq L(X, X)$ a countable family such that $\lim_{n \rightarrow \infty} K_n$ exist. Then we have*

$$\lim_{n \rightarrow \infty} (L \circ K_n) \text{ and } \lim_{n \rightarrow \infty} (L \circ K_n) = L \circ \left(\lim_{n \rightarrow \infty} K_n \right)$$

and

$$\lim_{n \rightarrow \infty} (K_n \circ L) \text{ exist and } \lim_{n \rightarrow \infty} (K_n \circ L) = \left(\lim_{n \rightarrow \infty} K_n \right) \circ L$$

***Note* 16.232.** *[theorem: 14.185] ensures that $L \circ K_n,\ K_n \circ L \in L(X, X)$ fore every $n \in \mathbb{N}$.*

Proof. As $\lim_{n \rightarrow \infty} K_n$ exist there exist a $K = \lim_{n \rightarrow \infty} K_n$ such that $\forall \varepsilon \in \mathbb{R}^+$ there exist a $N \in \mathbb{N}$ such that $\forall n \geqslant N$ we have

$$\|K_n - K\|_{L(X,X)} < \frac{\varepsilon}{1 + \|L\|_{L(X,X)}}$$

Further we have

$$L \circ K_n - L \circ K \underset{[\text{theorem: } 14.197]}{=} L \circ (K_n - K) \text{ and } K_n \circ L - K \circ L \underset{[\text{theorem: } 14.197]}{=} (K_n - K) \circ L \tag{16.313}$$

Next if $n \geqslant N$ then we have

$$\begin{aligned}
\|L \circ K_n - L \circ K\|_{L(X,Y)} &\underset{[\text{eq: } 16.313]}{=} \|L \circ (K_n - K)\|_{L(X,X)} \\
&\leqslant_{[\text{theorem: } 14.185]} \|L\|_{L(X,X)} \cdot \|K_n - K\|_{L(X,X_n)} \\
&< \|L\|_{L(X,X)} \cdot \frac{\varepsilon}{1 + \|L\|_{L(X,X)}} \\
&< \varepsilon \\
\|K_n \circ L - K \circ L\|_{L(X,X)} &\underset{[\text{eq: } 16.313]}{=} \|(K_n - K) \circ L\|_{L(X,X)} \\
&\leqslant_{[\text{theorem: } 14.185]} \|K_n - K\|_{L(X,X_n)} \cdot \|L\|_{L(X,X)} \\
&< \frac{\varepsilon}{1 + \|L\|_{L(X,X)}} \cdot \|L\|_{L(X,X)} \\
&< \varepsilon
\end{aligned}$$

This proves that

$$\lim_{n \rightarrow \infty} (L \circ K_n) = L \circ K = L \circ \left(\lim_{n \rightarrow \infty} K_n \right)$$

and

$$\lim_{n \rightarrow \infty} (K_n \circ L) = K \circ L = \left(\lim_{n \rightarrow \infty} K_n \right)$$

□

Definition 16.233. *Let $\langle X, \|\|\rangle$ be a normed space and $H \in L(X, X)$, $n \in \mathbb{N}$ then $H^n: X \rightarrow X$ is defined by*

$$H^n = \begin{cases} H \text{ if } n = 1 \\ H \circ H^{n-1} \text{ if } 1 < n \end{cases}$$

Theorem 16.234. *Let $\langle X, \|\|\rangle$ be a normed space, $n \in \mathbb{N}$ and $H \in L(X, X)$ then we have:*

1. $H^n \circ H = H^{n+1}$
2. $\forall \alpha \in \mathbb{K}$ *we have* $(\alpha \cdot H)^n = (\alpha)^n \cdot H^n$
3. $H^n \in L(X, X)$ *and* $\|H^n\|_{L(X,X)} \leqslant \|H\|^n_{L(X,Y)}$
4. $\sum_{i=1}^n H^i \in L(X, X)$ *and* $\|\sum_{i=1}^n H^i\|_{L(X,X)} \leqslant \sum_{i=1}^n \|H\|^i_{L(X,X)}$

Proof.

1. We prove this by induction so let

   $$S = \{n \in \mathbb{N} | H^n \circ H = H^{n+1}\}$$

   we have:

   $1 \in S$. As $H^1 \circ H = H \circ H = H \circ H^1 = H^{1+1}$ it follows that $1 \in S$

$\boldsymbol{n\in S\Rightarrow n+1\in S}$**.** We have

$$\begin{array}{rcl} H^{n+1}\circ H & = & (H^1\circ H^n)\circ H\\ & \underset{\circ\text{ is associative}}{=} & H^1\circ(H^n\circ H)\\ & \underset{n\in S}{=} & H^1\circ H^{n+1}\\ & = & H\circ H^{n+1}\\ & = & H^{(n+1)+1}\end{array}$$

proving that $n+1\in S$.

2. This is easily proved by induction, so let

$$S=\{n\in\mathbb{N}|(\alpha\cdot H)^n=\alpha^n\cdot H^n\}$$

then we have:

$\boldsymbol{1\in S}$**.** Let $x\in X$ then $(\alpha\cdot H)^1(x)=(\alpha\cdot H)(x)=\alpha\cdot H(x)=\alpha^1\cdot H^1(x)=(\alpha\cdot H^1)(x)$ proving that $(\alpha\cdot H)^1=\alpha^1\cdot H^1$. Hence $1\in S$.

$\boldsymbol{n\in S\Rightarrow n+1\in S}$**.** Let $x\in X$ then we have as $1<n+1$

$$\begin{array}{rcl} (\alpha\cdot H)^{n+1}(x) & = & ((\alpha\cdot H)\circ(\alpha\cdot H)^n)(x)\\ & = & (\alpha\cdot H)((\alpha\cdot H)^n(x))\\ & \underset{n\in S}{=} & (\alpha\cdot H)((\alpha^n\cdot H^n)(x))\\ & = & (\alpha\cdot H)(\alpha^n\cdot H^n(x))\\ & = & \alpha\cdot H(\alpha^n\cdot H^n(x))\\ & = & \alpha\cdot\alpha^n\cdot H(H^n(x))\\ & = & \alpha^{n+1}\cdot((H\circ H^n))(x)\\ & = & \alpha^{n+1}\cdot H^{n+1}(x)\\ & = & (\alpha^{n+1}\cdot H^{n+1})(x)\end{array}$$

proving that $(\alpha\cdot H)^{n+1}=\alpha^{n+1}\cdot H^{n+1}$. Hence

$$n+1\in S$$

3. We prove this by induction, so let

$$S=\{n\in\mathbb{N}|H^n\in L(X,X)\text{ and }\|H^n\|_{L(X,Y)}\leqslant\|H\|^n_{L(X,X)}\}$$

then we have:

$\boldsymbol{1\in S}$**.** As $H^1=H\in L(X,X)$ and $\|H^1\|_{L(X,X)}=\|H\|_{L(X,X)}=\|H\|^1_{L(X,X)}$ it follows that $1\in S$.

$\boldsymbol{n\in S\Rightarrow n+1\in S}$**.** Let $\alpha\in\mathbb{K}$, $x,y\in X$ then we have

$$\begin{array}{rcl} H^{n+1}(x+\beta\cdot y) & = & (H\circ H^n)(x+\beta\cdot y)\\ & = & H(H^n(x+\beta\cdot y))\\ & \underset{n\in S}{=} & H(H^n(x)+\beta\cdot H^n(y))\\ & = & H^{n+1}(x)+\beta\cdot H^{n+1}(y)\end{array}$$

proving that

$$H^{n+1}\in\operatorname{Hom}(X,X)$$

Further we have

$$\begin{array}{rcl} \|H^{n+1}(x)\| & = & \|(H\circ H^n)(x)\|\\ & = & \|H(H^n(x))\|\\ & \leqslant & \|H\|_{L(X,Y)}\cdot\|H^n(x)\|\\ & \leqslant_{n\in S\Rightarrow H^n\in L(X<X)} & \|H\|_{L(X,Y)}\cdot\|H^n\|_{L(X,X)}\cdot\|x\|\\ & \leqslant_{n\in S} & \|H\|_{L(X,Y)}\cdot\|H\|^n_{L(X,X)}\cdot\|x\|\\ & = & \|H\|^{n+1}_{L(X,X)}\cdot\|x\|\end{array}$$

so that by [theorem: 14.174] $H^{n+1}\in L(X,X)$ and by [theorem: 14.181] that $\|H_{n+1}\|_{L(X,X)}\leqslant\|H\|^{n+1}_{L(X,X)}$. Hence we have

$$n+1\in S$$

4. Let $\alpha\in\mathbb{K}$ and $x,y\in X$

$$\begin{aligned}\left(\sum_{i=1}^{n} H^i\right)(x+\beta\cdot y) &\underset{\text{def}}{=} \sum_{i=1}^{n} H^i(x+\beta\cdot y)\\ &\underset{(1)}{=} \sum_{i=1}^{n} (H^i(x)+\beta\cdot H^i(y))\\ &= \sum_{i=1}^{n} H^i(x)+\beta\cdot\sum_{i=1}^{n} H^i(y)\\ &= \left(\sum_{i=1}^{n} H^i\right)(x)+\beta\cdot\left(\sum_{i=1}^{n} H^i\right)(y)\end{aligned}$$

proving that

$$\sum_{i=1}^{n} H^i\in\operatorname{Hom}(X,X)$$

Further

$$\begin{aligned}\left\|\left(\sum_{i=1}^{n} H^i\right)(x)\right\| &= \left\|\sum_{i=1}^{n} H^i(x)\right\|\\ &\leqslant \sum_{i=1}^{n} \|H^i(x)\|\\ &\leqslant_{(1)} \sum_{i=1}^{n} (\|H\|^i_{L(X,X)}\cdot\|x\|)\\ &= \left(\sum_{i=1}^{n} \|H\|^i_{L(X,Y)}\right)\cdot\|x\|\end{aligned}$$

so that by [theorem: 14.174] $\sum_{i=1}^n H^i\in L(X,X)$ and by [theorem: 14.181] that $\|\sum_{i=1}^n H^i\|_{L(X,X)}\leqslant\sum_{i=1}^n\|H\|^i_{L(X,Y)}$.

□

Lemma 16.235. *Let $\langle X,\|\|\rangle$ be a normed space, $k\in\mathbb{N}$ and $H\in L(X,X)$ such that $\sum_{i=k}^{\infty}H^i$ converges [using the norm $\|\|_{L(X,X)}$] then $\forall x\in X$ we have that $\sum_{i=k}^{\infty}H^i(x)$ converges and $(\sum_{i=k}^{\infty}H^i)(x)=\sum_{i=k}^{\infty}H^i(x)$*

Proof. Let $x\in X$ and $\varepsilon\in\mathbb{R}^+$ then there exist a $N\in\mathbb{N}$ such that for $n\geqslant N$ we have that

$$\left\|\sum_{i=k}^{n} H^i-\sum_{i=k}^{\infty} H^i\right\|_{L(X,X)}<\frac{\varepsilon}{1+\|x\|}$$

then we have

$$\begin{aligned}\left\|\sum_{i=k}^{n} H^i(x)-\left(\sum_{i=k}^{\infty} H^i\right)(x)\right\| &= \left\|\left(\sum_{i=k}^{n} H^i-\sum_{i=k}^{\infty} H^i\right)(x)\right\|\\ &\leqslant \left\|\sum_{i=k}^{n} H^i-\sum_{i=k}^{\infty} H^i\right\|_{L(X,X)}\cdot\|x\|\\ &< \frac{\varepsilon}{1+\|x\|}\cdot\|x\|\\ &< \varepsilon\end{aligned}$$

proving that $\lim_{n\to\infty}\sum_{i=k}^n H^i(x)$ converges and

$$\lim_{n\to\infty}\sum_{i=k}^{n} H^i(x)=\left(\sum_{i=k}^{\infty} H^i\right)(x)$$

so that

$$\sum_{i=k}^{\infty} H^i(x) = \left(\sum_{i=k}^{\infty} H^i\right)(x)$$

□

Lemma 16.236. *Let $\langle X, \|\|\rangle$ be a Banach space and $H \in B_{\|\|_{L(X,X)}}(0,1)$ then*

$$\sum_{i=1}^{\infty} H^i \text{ converges}$$

and

$$\mathrm{Id}_X - H \in \mathcal{G}L(X) \text{ and } (\mathrm{Id}_X - H)^{-1} = \mathrm{Id}_X + \sum_{i=1}^{\infty} H^i$$

Note 16.237. In this lemma and the rest of this subsection the neutral element of $L(X,X)$ is noted as 0. Context can be used to determine if 0 is the neutral element of $\mathbb{K}$, X or $L(X,X)$.

Proof. As $H \in B_{\|\|_{L(X,X)}}(0,1)$ we have that $\|H\|_{L(X,Y)} < 1$ so that by [example: 14.348]

$$\sum_{i=1}^{\infty} \|H\|_{L(X,X)}^i \text{ exist and } \sum_{i=1}^{\infty} \|H\|_{L(X,X)}^i = \frac{\|H\|_{L(X,X)}}{1 - \|H\|_{L(X,X)}}$$

As X is a Banach space it follows from [theorem: 14.334] is a Banach space. Further for $i \in \mathbb{N}$ we have $\|H^i\|_{L(X,Y)} \leqslant_{[\text{theorem: } 16.234]} \|H\|_{L(X,X)}^i$. So we can use dominant convergence [see 14.372] to prove that

$$\sum_{i=1}^{\infty} H^i \text{ exist, } \sum_{i=1}^{\infty} H^i \in L(X,X) \text{ and } \left\|\sum_{i=1}^{\infty} H^i\right\|_{L(X,X)} \leqslant \sum_{i=1}^{\infty} \|H\|_{L(X,X)}^i = \frac{\|H\|_{L(X,X)}}{1 - \|H\|_{L(X,X)}} \tag{16.314}$$

From the above and the fact that $\mathrm{Id}_X \in L(X,X)$ [see 14.172]

$$\mathrm{Id}_X + \sum_{i=1}^{\infty} H^i \in L(X,X) \tag{16.315}$$

Further as $\sum_{i=1}^{\infty} H^i$ converges it follows from [theorem: 14.349] that

$$\lim_{n \to \infty} H^i = 0 \tag{16.316}$$

and

$$\begin{aligned}
\left\|\mathrm{Id}_X + \sum_{i=1}^{\infty} H^i\right\|_{L(X,X)} &\leqslant \|\mathrm{Id}_X\|_{L(X,Y)} + \left\|\sum_{i=1}^{\infty} H^i\right\|_{L(X,X)} \\
&\leqslant_{[\text{example: } 14.182]} 1 + \left\|\sum_{i=1}^{\infty} H^i\right\|_{L(X,X)} \\
&\leqslant_{[\text{eq: } 16.314]} 1 + \frac{\|H\|_{L(X,X)}}{1 - \|H\|_{L(X,X)}} \\
&= \frac{1 - \|H\|_{L(X,X)} + \|H\|_{L(X.X)}}{1 - \|H\|_{L(X,X)}} \\
&= \frac{1}{1 - \|H\|_{L(X,X)}}
\end{aligned} \tag{16.317}$$

Next we prove by induction on n that

$$(\mathrm{Id}_X - H) \circ \left(\mathrm{Id}_X + \sum_{i=1}^{n} H^i\right) = \mathrm{Id}_X - H^{n+1} \tag{16.318}$$

Proof. Let

$$S = \left\{n \in \mathbb{N} | (\mathrm{Id}_X - H) \circ \left(\mathrm{Id}_X + \sum_{i=1}^{n} H^i\right) = \mathrm{Id}_X - H^{n+1}\right\}$$

then we have:

$1\in S$. Let $x\in X$ then we have

$$
\begin{array}{rcl}
\left((\mathrm{Id}_X-H)\circ\left(\mathrm{Id}_X+\sum_{i=1}^{1}H^i\right)\right)(x) & = & (\mathrm{Id}_X-H)\left(\left(\mathrm{Id}_X+\sum_{i=1}^{1}H^i\right)(x)\right)\\
& = & (\mathrm{Id}_X-H)\left(\mathrm{Id}_X(x)+\left(\sum_{i=1}^{1}H^i\right)(x)\right)\\
& = & (\mathrm{Id}_X-H)(x+H(x))\\
& = & \mathrm{Id}_X(x+H(x))-H(x+H(x))\\
& = & x+H(x)-H(x)-H(H(x))\\
& = & x-(H\circ H)(x)\\
& = & x-H^2(x)\\
& = & \mathrm{Id}_X(x)-H^2(x)\\
& = & (\mathrm{Id}_X-H^2)(x)
\end{array}
$$

proving that $((\mathrm{Id}_X-H)\circ(\mathrm{Id}_X+\sum_{i=1}^{1}H^i))$. So it follows that $1\in S$.

$n\in S\Rightarrow n+1\in S$. Let $x\in X$ then

$$
\begin{array}{rl}
\left((\mathrm{Id}_X-H)\circ\left(\mathrm{Id}_X+\sum_{i=1}^{n+1}H^i\right)\right)(x) & =\\
(\mathrm{Id}_X-H)\left(\left(\mathrm{Id}_X+\sum_{i=1}^{n+1}H^i\right)(x)\right) & =\\
(\mathrm{Id}_X-H)\left(\mathrm{Id}_X(x)+\left(\sum_{i=1}^{n+1}H^i\right)(x)\right) & =\\
(\mathrm{Id}_X-H)\left(\mathrm{Id}_X(x)+\sum_{i=1}^{n+1}H^i(x)\right) & =\\
(\mathrm{Id}_X-H)\left(\mathrm{Id}_X(x)+\sum_{i=1}^{n}H^i(x)+H^{n+1}(x)\right) & =\\
(\mathrm{Id}_X-H)\left(\mathrm{Id}_X(x)+\sum_{i=1}^{n}H^i(x)\right)+(\mathrm{Id}_X-H)(H^{n+1}(x)) & =\\
(\mathrm{Id}_X-H)\left(\left(\mathrm{Id}_X+\sum_{i=1}^{n}H^i\right)(x)\right)+(\mathrm{Id}_X-H)(H^{n+1}(x)) & =\\
\left((\mathrm{Id}_X-H)\circ\left(\mathrm{Id}_X+\sum_{i=1}^{n}H^i\right)\right)(x)+\mathrm{Id}_X(H^{n+1}(x))-H(H^{n+1}(x)) & \underset{n\in S}{=}\\
(\mathrm{Id}_X-H^{n+1})(x)+H^{n+1}(x)-(H\circ H^{n+1})(x) & =\\
\mathrm{Id}_X(x)-H^{(n+1)+1}(x)(x) & =\\
(\mathrm{Id}_X-H^{(n+1)+1})(x) &
\end{array}
$$

proving that $((\mathrm{Id}_X-H)\circ(\mathrm{Id}_X+\sum_{i=1}^{n+1}H^i))=(\mathrm{Id}_X-H^{(n+1)+1})$. Hence

$$n+1\in S$$

Mathematical induction completes the proof of [eq: 16.318]. Next using induction we prove by induction that

$$\left(\mathrm{Id}_X+\sum_{i=1}^{n}H^i\right)\circ(\mathrm{Id}_X-H)=\mathrm{Id}_X-H^{n+1}\tag{16.319}$$

So define

$$S=\left\{n\in\mathbb{N}|\left(\mathrm{Id}_X+\sum_{i=1}^{n}H^i\right)\circ(\mathrm{Id}_X-H)=\mathrm{Id}_X-H^{n+1}\right\}$$

then we have:

$1\in S$. Let $x\in X$ then

$$\begin{array}{rcl}\left(\left(\mathrm{Id}_X+\sum_{i=1}^{1} H^i\right)\circ(\mathrm{Id}_X-H)\right)(x) & = & ((\mathrm{Id}_X+H)\circ(\mathrm{Id}_X-H))(x)\\ & = & (\mathrm{Id}_X+H)((\mathrm{Id}_X-H)(x))\\ & = & (\mathrm{Id}_X+H)(\mathrm{Id}_X(x)-H(x))\\ & = & \mathrm{Id}_X(\mathrm{Id}_X(x)-H(x))+H(\mathrm{Id}_X(x)-H(x))\\ & = & \mathrm{Id}_X(x)-H(x)+H(\mathrm{Id}_X(x))-H(H(x))\\ & = & \mathrm{Id}_X(x)-H(x)+H(x)-(H\circ H)(x)\\ & = & \mathrm{Id}_X(x)-H^2(x)\\ & = & (\mathrm{Id}_X-H^2)(x)\end{array}$$

proving that $(\mathrm{Id}_X+\sum_{i=1}^{1} H^i)\circ(\mathrm{Id}_X-H)=\mathrm{Id}_X-H^{1+1}$. Hence $1\in S$.

$n\in S\Rightarrow n+1$. Let $x\in X$ then

$$\begin{array}{rc}\left(\left(\mathrm{Id}_X+\sum_{i=1}^{n+1} H^i\right)\circ(\mathrm{Id}_X-H)\right)(x) & =\\ \left(\mathrm{Id}_X+\sum_{i=1}^{n+1} H^i\right)((\mathrm{Id}_X-H)(x)) & =\\ \left(\mathrm{Id}_X+\sum_{i=1}^{n} H^i+H^{n+1}\right)((\mathrm{Id}_X-H)(x)) & =\\ \left(\mathrm{Id}_X+\sum_{i=1}^{n} H^i\right)((\mathrm{Id}_X-H)(x))+H^{n+1}((\mathrm{Id}_X-H)(x)) & =\\ \left(\left(\mathrm{Id}_X+\sum_{i=1}^{n} H^i\right)\circ(\mathrm{Id}_X-H)\right)(x)+H^{n+1}(x-H(x)) & \underset{n\in S}{=}\\ (\mathrm{Id}_X-H^{n+1})(x)+H^{m+1}(x)-H^{n+1}(H(x)) & =\\ \mathrm{Id}_X(x)-(H^{n+1}\circ H)(x) & \underset{[\text{theorem: } 16.234]}{=}\\ \mathrm{Id}_X(x)-H^{(n+1)+1}(x) & =\\ (\mathrm{Id}_X-H^{(n+1)+1})(x) & \end{array}$$

proving that $(\mathrm{Id}_X+\sum_{i=1}^{n+1} H^i)\circ(\mathrm{Id}_X-H)-\mathrm{Id}_X-H^{(n+1)+1}$. Hence

$$n+1\in S$$

Mathematical induction proves then [eq: 16.319]. □

To summarize [eqs: 16.318,16.319] we have

$$(\mathrm{Id}_X-H)\circ\left(\mathrm{Id}_X+\sum_{i=1}^{n} H^i\right)=\mathrm{Id}_X-H^{n+1}=\left(\mathrm{Id}_X+\sum_{i=1}^{n+1} H^i\right)\circ(\mathrm{Id}_X-H) \tag{16.320}$$

Let $\varepsilon\in\mathbb{R}^+$ then as $\lim_{n\rightarrow\infty} H^i=0$ [see eq: 16.316] there exist a $N\in\mathbb{N}$ such that for $n>N$

$$\|H^n\|_{L(X.X)}<\varepsilon$$

then

$$\begin{array}{rcl}\left\|(\mathrm{Id}_X-H)\circ\left(\mathrm{Id}_X+\sum_{i=1}^{n} H^i\right)-\mathrm{Id}_X\right\|_{L(X,X)} & \underset{[\text{eq: } 16.320]}{=} & \|\mathrm{Id}_X-H^{n+1}-\mathrm{Id}_X\|_{L(X,Y)}\\ & = & \|H^{n+1}\|_{L(X,Y)}\\ & <_{n+1>n\geqslant N} & \varepsilon\end{array}$$

and

$$\left\|\left(\mathrm{Id}_X+\sum_{i=1}^{n} H^i\right)\circ(\mathrm{Id}_X-H)-\mathrm{Id}_X\right\|_{L(X,X)} \underset{[\text{eq: } 16.320]}{=} \|\mathrm{Id}_X-H^{n+1}-\mathrm{Id}_X\|_{L(X,Y)} = \|H^{n+1}\|_{L(X,Y)} \underset{n+1>n\geqslant N}{<} \varepsilon$$

So that

$$\lim_{n\rightarrow\infty}\left(\left(\mathrm{Id}_X+\sum_{i=1}^{n} H^i\right)\circ(\mathrm{Id}_X-H)\right)=\mathrm{Id}_X=\lim_{n\rightarrow\infty}\left((\mathrm{Id}_X-H)\circ\left(\mathrm{Id}_X+\sum_{i=1}^{n} H^i\right)\right) \tag{16.321}$$

Next we have

$$\begin{aligned}
\lim_{n\rightarrow\infty}\left(\left(\mathrm{Id}_X+\sum_{i=1}^{n} H^i\right)\circ(\mathrm{Id}_X-H)\right) &\underset{[\text{theorem: } 16.231]}{=} \left(\lim_{n\rightarrow\infty}\left(\mathrm{Id}_X+\sum_{i=1}^{n} H^i\right)\right)\circ(\mathrm{Id}_X-H)\\
&\underset{[\text{theorem: } 14.296]}{=} \left(\mathrm{Id}_X+\lim_{n\rightarrow\infty}\sum_{i=1}^{n} H^i\right)\circ(\mathrm{Id}_X-H)\\
&= \left(\mathrm{Id}_X+\sum_{i=1}^{n} H^i\right)\circ(\mathrm{Id}_X-H)
\end{aligned}$$

and

$$\begin{aligned}
\lim_{n\rightarrow\infty}\left((\mathrm{Id}_X-H)\circ\left(\mathrm{Id}_X+\sum_{i=1}^{n} H^i\right)\right) &\underset{[\text{theorem: } 16.231]}{=} (\mathrm{Id}_X-H)\circ\left(\lim_{n\rightarrow\infty}\left(\mathrm{Id}_X+\sum_{i=1}^{n} H^i\right)\right)\\
&\underset{[\text{theorem: } 14.296]}{=} (\mathrm{Id}_X-H)\circ\left(\mathrm{Id}_X+\lim_{n\rightarrow\infty}\left(\sum_{i=1}^{n} H^i\right)\right)\\
&= (\mathrm{Id}_X-H)\circ\left(\mathrm{Id}_X+\sum_{i=1}^{\infty} H^i\right)
\end{aligned}$$

which combined with [eq: 16.321] proves that

$$(\mathrm{Id}_X-H)\circ\left(\mathrm{Id}_X+\sum_{i=1}^{\infty} H^i\right)=\mathrm{Id}_X=\left(\mathrm{Id}_X+\sum_{i=1}^{\infty} H^i\right)\circ(\mathrm{Id}_X-H)$$

hence

$$\mathrm{Id}_X-H \text{ is a bijection and } (\mathrm{Id}_X-H)^{-1}=\mathrm{Id}_X+\sum_{i=1}^{\infty} H^i\in L(X,X)$$

Finally by [theorem: 16.228] it follows that Id_X-H is a toplinear isomorphism or that

$$\mathrm{Id}_X-H\in\mathcal{GL}(X)$$

□

Corollary 16.238. *Let $\langle X,\|\|\rangle$ be a Banach space, $H\in B_{\|\|_{L(X,X)}}(0,1)$ then $\sum_{i=1}^{\infty}(-1)^i\cdot H^i$ converges, $\mathrm{Id}_X+H\in\mathcal{GL}(X)$ and $(\mathrm{Id}_X+H)^{-1}=\mathrm{Id}_X+\sum_{i=1}^{\infty}(-1)^i\cdot H^i$.*

Proof. If $H\in B_{\|\|_{L(X,X)}}(0,1)$ then $\|-H\|_{L(X,X)}=\|(-1)\cdot H\|_{L(X,Y)}=\|H\|_{L(X,X)}<1$. Hence by the previous theorem we have that

$$\mathrm{Id}_X+H=\mathrm{Id}_X-(-H)\in\mathcal{GL}(X)$$

and

$$\sum_{i=1}^{\infty}(-1)^i\cdot H^i \underset{[\text{theorem: } 16.234]}{=} \sum_{i=1}^{\infty}(-H)^i \text{ exist}$$

and

$$(\mathrm{Id}_X+H)^{-1}=(\mathrm{Id}_X-(-H))^{-1}=\mathrm{Id}_X+\sum_{i=1}^{\infty}(-H)^i \underset{[\text{theorem: } 16.234]}{=} \mathrm{Id}_X+\sum_{i=1}^{\infty}(-1)^n\cdot H^i$$

□

Theorem 16.239. *Let $\langle X, \|\|\rangle$ be a Banach space then we have:*

1. *$\mathcal{G}L(X)$ is a open set in $L(X, X)$*

2. *The function*

$$\tau: \mathcal{G}L(X) \rightarrow L(X, X) \text{ defined by } \tau(L) = L^{-1}$$

is Fréchet differentiable on $\mathcal{G}L(X)$ and $\forall L \in \mathcal{G}L(X)$ we have that

$$D_L\tau \in L(L(X, X), L(X, X)) \text{ is defined by } D_L\tau(H) = -L^{-1} \circ H \circ L$$

3. *The function τ is of class C^∞.*

Proof.

1. Let $L \in \mathcal{G}L(X) \subseteq L(X, X)$ then for any $H \in L(X, X)$ we have

$$\begin{array}{rcl} H & = & L + (H - L) \\ & = & L \circ \mathrm{Id}_X + \mathrm{Id}_X \circ (H - L) \\ & = & L \circ \mathrm{Id}_X + (L \circ L^{-1}) \circ (H - L) \\ & = & L \circ \mathrm{Id}_X + L \circ (L^{-1} \circ (H - L)) \\ & \underset{[\text{theorem: } 14.197]}{=} & L \circ (\mathrm{Id}_X + L^{-1} \circ (H - L)) \end{array} \tag{16.322}$$

Take $\delta_L = \frac{1}{\|L^{-1}\|_{L(X,Y)} + 1} \in \mathbb{R}^+$ then if $\|H - L\|_{L(X,Y)} < \delta_L$ we have that

$$\begin{array}{rcl} \|L^{-1} \circ (H - L)\|_{L(X,X)} & \leqslant_{[\text{theorem: } 14.185]} & \|L^{-1}\|_{L(X,X)} \cdot \|H - L\|_{L(X,X)} \\ & < & \|L^{-1}\|_{L(X,X)} \cdot \delta_L \\ & = & \frac{\|L^{-1}\|_{L(X,X)}}{\|L^{-1}\|_{L(X,X)} + 1} \\ & < & 1 \end{array}$$

Hence $\|L^{-1} \circ (H - L)\|_{L(X,X)} \in B_{\|\|_{L(X,X)}}(0, 1)$ so that by [corollary: 16.238]

$$\mathrm{Id}_X + L^{-1} \circ (H - L) \in \mathcal{G}L(X)$$

and as $L \in \mathcal{G}L(X)$ we have by [theorem: 16.230] that $L \circ (\mathrm{Id}_X + L^{-1} \circ (H - L)) \in \mathcal{G}L(X)$, so by [eq: 16.322] $H \in \mathrm{GL}(X)$. In other words if $H \in B_{\|\|_{L(X,X)}}(L, \delta_L)$ then $H \in \mathcal{G}L(X)$ so that

$$\forall L \in \mathcal{G}L(X) \text{ we have } L \in B_{\|\|_{L(X,X)}}(L, \delta_L) \subseteq \mathcal{G}L(X) \tag{16.323}$$

which proves that

$$\mathcal{G}L(X) \text{ is a open set}$$

2. If $H \in B_{\|\|_{L(X,X)}}(\mathrm{Id}_X, 1)$ then $\|H - \mathrm{Id}_X\|_{L(X,X)} < 1$ so that by [corollary: 16.238]

$$H = \mathrm{Id}_X + (H - \mathrm{Id}_X) \in \mathcal{G}L(X)$$

so that

$$B_{\|\|_{L(X,X)}}(\mathrm{Id}_X, 1) \subseteq \mathcal{G}L(X) \tag{16.324}$$

To prove differentiability we will rely on [corollary: 16.30], hence the need for a ε-mapping [see definition: 16.28]. Let $H \in B_{\|\|_{L(X,X)}}(0, 1)$ then by [corollary: 16.238] $\sum_{i=1}^{\infty} (-1)^i \cdot H^i$ converges, so applying [theorem: 14.352] we have that

$$\sum_{i=2}^{\infty} (-1)^i \cdot H^i \in L(X, X) \text{ converges and } \sum_{i=1}^{\infty} (-1)^i H^i = -H + \sum_{i=2}^{\infty} (-1)^i \cdot H^i \tag{16.325}$$

As $0, \sum_{i=2}^{\infty} (-1)^i \cdot H^i \in L(X, X)$ it follows that the following function is well defined

$$\varepsilon\colon B_{\|\|_{L(X,X)}}(0,1)\rightarrow L(X,Y) \text{ by } \varepsilon(H)=\begin{cases}0 \text{ if } H=0\\ \frac{1}{\|H\|_{L(X,X)}}\cdot\sum_{i=2}^{\infty}(-1)^i\cdot H^i \text{ if } H\neq 0\end{cases} \tag{16.326}$$

Let $x\in X$ then we have

$$\begin{aligned}\|(\varepsilon(H))(x)\| &= \begin{cases}\|0(x)\| \text{ if } H=0\\ \left\|\frac{1}{\|\|_{L(X,X)}}\cdot\left(\sum_{i=2}^{\infty}(-1)\cdot H^i\right)(x)\right\| \text{ if } H\neq 0\end{cases}\\ &= \begin{cases}0 \text{ if } H=0\\ \frac{1}{\|H\|_{L(X,X)}}\cdot\|\left(\sum_{i=2}^{\infty}(-1)\cdot H^i\right)(x)\| \text{ if } H\neq 0\end{cases}\\ &\leqslant \begin{cases}0\cdot\|x\| \text{ if } H=0\\ \frac{1}{\|H\|_{L(X,X)}}\cdot\|\sum_{i=2}^{\infty}(-1)\cdot H^i\|_{L(X,X)}\cdot\|x\| \text{ if } H\neq 0\end{cases}\\ &= \|x\|\cdot\begin{cases}0\cdot\|x\| \text{ if } H=0\\ \frac{1}{\|H\|_{L(X,X)}}\cdot\|\sum_{i=2}^{\infty}(-1)\cdot H^i\|_{L(X,X)} \text{ if } H\neq 0\end{cases}\end{aligned}$$

proving by [theorem: 14.181] that

$$\|\varepsilon(H)\|_{L(X,Y)}\leqslant\begin{cases}0\cdot\|x\| \text{ if } H=0\\ \frac{1}{\|H\|_{L(X,X)}}\cdot\|\sum_{i=2}^{\infty}(-1)\cdot H^i\|_{L(X,X)} \text{ if } H\neq 0\end{cases} \tag{16.327}$$

For $H\in B_{\|\|_{L(X,X)}}(0,1)$ we have $\|H\|_{L(X,X)}<1$ so that by [example: 14.348] $\sum_{i=1}^{\infty}\|H\|_{L(X,X)}^i=\frac{\|H\|_{L(X,X)}}{1-\|H\|_{L(X,X)}}$. Hence we have by [theorem: 14.352] that

$$\begin{aligned}\sum_{i=2}^{\infty}\|H\|_{L(X,X)}^i &\underset{\text{[eq: 16.325]}}{=} \left(\sum_{i=1}^{\infty}\|H\|_{L(X,X)}^i\right)-\|H\|_{L(X,X)}\\ &= \frac{\|H\|_{L(X,X)}}{1-\|H\|_{L(X,X)}}-\|H\|_{L(X,X)}\\ &= \frac{\|H\|_{L(X,X)}-\|H\|_{L(X,X)}\cdot(1-\|H\|_{L(X,X)})}{1-\|H\|_{L(X,X)}}\\ &= \frac{\|H\|_{L(X,X)}^2}{1-\|H\|_{L(X,X)}}\end{aligned} \tag{16.328}$$

As by [theorem: 16.234] we have $\|(-1)^iH^i\|_{L(X,X)}=\|H^i\|_{L(X,X)}\leqslant\|H\|_{L(X,X)}^i$ we can apply the dominant convergence theorem [theorem: 14.372] proving that $\|\sum_{i=2}^{\infty}(-1)^i\cdot H^i\|_{L(X,X)}$ converges and $\|\sum_{i=2}^{\infty}(-1)^i\cdot H^i\|_{L(X,X)}\leqslant\frac{\|H\|_{L(X,X)}^2}{1-\|H\|_{L(X,X)}}$. Substituting this result in [eq: 16.327] gives

$$\begin{aligned}\|\varepsilon(H)\|_{L(X,X)} &\leqslant \begin{cases}0\cdot\|x\| \text{ if } H=0\\ \frac{1}{\|H\|_{L(X,X)}}\cdot\frac{\|H\|_{L(X,X)}^2}{1-\|H\|_{L(X,X)}} \text{ if } H\neq 0\end{cases}\\ &= \begin{cases}0 \text{ if } H=0\\ \frac{\|H\|_{L(X,X)}}{1-\|H\|_{L(X,X)}} \text{ if } H\neq 0\end{cases}\end{aligned}$$

so that

$$\forall H\in B_{\|\|_{L(X,X)}}(0,1) \text{ we have } \|\varepsilon(H)\|_{L(X,X)}\leqslant\frac{\|H\|_{L(X,X)}}{1-\|H\|_{L(X,X)}} \tag{16.329}$$

Let $\zeta \in \mathbb{R}^+$ then if $\|H\|_{L(X,X)} < \min\left(\frac{1}{2}, \frac{\zeta}{2}\right) < 1$ it follows that

$$\begin{array}{rcl} \|\varepsilon(H) - \varepsilon(0)\|_{L(X,X)} & = & \|\varepsilon(H)\|_{L(X,Y)} \\ & \leqslant_{[\text{eq: } 16.329]} & \dfrac{\|H\|_{L(X,X)}}{1 - \|H\|_{L(X,X)}} \\ & < & \dfrac{\|H\|_{L(X,X)}}{1 - \frac{1}{2}} \\ & < & \dfrac{\zeta}{2} \cdot \dfrac{2}{1} \\ & = & \zeta \end{array}$$

which proves that

$$\varepsilon \text{ is continuous at } 0 \tag{16.330}$$

or using [definition: 16.28]

$$\varepsilon \text{ is a } \varepsilon\text{-mapping} \tag{16.331}$$

Consider $L \in \mathcal{G}L(X)$ then we have the following cases to consider:

$\boldsymbol{L = \mathrm{Id}_X}$**.** If $H \in B_{\|\|_{L(X,X)}}(0,1)$ then

$$\begin{array}{rcl} \tau(\mathrm{Id}_X + H) - \tau(\mathrm{Id}_X) & = & (\mathrm{Id}_X + H)^{-1} - (\mathrm{Id}_X)^{-1} \\ & = & (\mathrm{Id}_X + H)^{-1} - \mathrm{Id}_X \\ & \underset{[\text{corollary: } 16.238]}{=} & \mathrm{Id}_X + \sum_{i=1}^{\infty} (-1)^i \cdot H^i - \mathrm{Id}_X \\ & = & \sum_{i=1}^{\infty} (-1)^i \cdot H^i \\ & \underset{[\text{theorem: } 14.352]}{=} & -H + \sum_{i=2}^{\infty} (-1)^i \cdot H^i \end{array}$$

As $\tau(\mathrm{Id}_X + 0) - \tau(\mathrm{Id}_X) - 0 = 0$ we have using the above that

$$\begin{array}{rl} \tau(\mathrm{Id}_X + H) - \tau(\mathrm{Id}_X) - (-H) & = \\ \left\{\begin{array}{l} 0 \text{ if } H = 0 \\ \sum_{i=2}^{\infty} (-1)^i \cdot H^i \text{ if } H \neq 0 \end{array}\right. & = \\ \left\{\begin{array}{l} 0 \cdot H \text{ if } H = 0 \\ \|H\|_{L(X,X)} \cdot \left(\frac{1}{\|H\|_{L(X,X)}} \cdot \sum_{i=2}^{\infty} (-1)^i \cdot H^i\right) \text{ if } H \neq 0 \end{array}\right. & = \\ \|H\|_{L(X,X)} \cdot \varepsilon(H) & \end{array} \tag{16.332}$$

As $-\mathrm{Id}_X \circ H \circ \mathrm{Id}_X = -H$ we have from the above that

$$\tau(\mathrm{Id}_X + H) - \tau(\mathrm{Id}_X) - (-\mathrm{Id}_X \circ H \circ \mathrm{Id}_X) = \|H\|_{L(X,X)} \cdot \varepsilon(H) \tag{16.333}$$

Note that $-\mathrm{Id}_{L(L(X,X),L(X,X))} \in L(L(X,X), L(X,X))$ and

$$-\mathrm{Id}_{L(L(X,X),L(X,X))}(H) = -H = -\mathrm{Id}_X \circ H \circ \mathrm{Id}_X$$

So that $\forall H \in B_{\|\|_X}(0,1)$ we have

$$\tau(\mathrm{Id}_X + H) - \tau(\mathrm{Id}_X) - (-\mathrm{Id}_{L(L(X,X),L(X,X))})(H) = \|H\|_{L(X,X)} \cdot \varepsilon(H) \tag{16.334}$$

Using the above, the fact that ε is a ε-mapping [see eq: 16.331] and $B_{\|\|_{L(X,X)}}(\mathrm{Id}_X, 1) \subseteq \mathcal{G}L(X)$ [see eq: 16.324] we can apply [corollary: 16.30] resulting in

$$D_{\mathrm{Id}_X}\tau \text{ exist and } D_{\mathrm{Id}_X}\tau = -\mathrm{Id}_{L(L(X,X),L(X,X))}$$

where $D_{\mathrm{Id}_X}\tau \in L(L(X,X), L(X,X))$ is defined by

$$D_{\mathrm{Id}_X}\tau(H) = -\mathrm{Id}_{L(L(X,X),L(X,X))}(H) = -H = -\mathrm{Id}_X \circ H \circ \mathrm{Id}_X$$

$\boldsymbol{L \neq \mathrm{Id}_X}$. Take $\delta = \min\left(\frac{1}{1 + \|L^{-1}\|_{L(X,X)}}, 1\right)$ then if $H \in B_{\|\|_{L X,X}}(0, \delta)$ we have

$$\begin{array}{rcl} \|H \circ L^{-1}\|_{L(X,X)} & \leqslant_{[\text{theorem: } 14.185]} & \|H\|_{L(X,X)} \cdot \|L^{-1}\|_{L(X,X)} \\ & < & \delta \cdot \|L^{-1}\|_{L(X,X)} \\ & \leqslant & \frac{1}{1 + \|L^{-1}\|_{L(X,X)}} \cdot \|L^{-1}\|_{L(X,X)} \\ & < & 1 \end{array}$$

Hence using [corollary: 16.238] we get

$$\mathrm{Id}_X + H \circ L^{-1} \in \mathcal{GL}(X) \tag{16.335}$$

As $L \in \mathcal{GL}(X)$ we have that $(\mathrm{Id}_X + H \circ L^{-1}) \circ L \in \mathcal{GL}(X)$, further we have that $(\mathrm{Id}_X + H \circ L^{-1}) \circ L \underset{[\text{theorem: } 14.197]}{=} \mathrm{Id}_X \circ L + (H \circ L^{-1}) \circ L = \mathrm{Id}_X \circ L + H = L + H$, hence it follows that

$$L + H = (\mathrm{Id}_X + H \circ L^{-1}) \circ L \in \mathcal{GL}(X) \tag{16.336}$$

If $H \in B_{\|\|_{L(X,X)}}(L, \delta)$ then $\|H - L\|_{L(X,X)} < \delta$ so that $L - H \in B_{\|\|_{L(X,X)}}(0, \delta)$ so that by [eq: 16.336] $L = (L - H) + H \in \mathcal{GL}(X)$ proving that

$$B_{\|\|_{L(X,X)}}(L, \delta) \subseteq \mathcal{GL}(X) \tag{16.337}$$

If $H \in B_{\|\|_{L(X,X)}}(0, \delta)$ then by [theorem: 14.185] we have

$$\begin{array}{rcl} \|H \circ L^{-1}\|_{L(X,X)} & \leqslant & \|H\|_{L(X,X)} \cdot \|L^{-1}\|_{L(X,X)} \\ & < & \delta \cdot \|L^{-1}\|_{L(X,X)} \\ & \leqslant & \frac{1}{1 + \|L^{-1}\|_{L(X,X)}} \cdot \|L^{-1}\|_{L(X,X)} \\ & < & 1 \end{array}$$

so that

$$H \circ L^{-1} \in B_{\|\|_{L(X,X)}}(0, 1)$$

proving, as $\mathrm{dom}(\varepsilon) \underset{[\text{eq: } 16.326]}{=} B_{\|\|_{L(X,X)}}(0, 1)$, that

$$\forall H \in B_{\|\|_{L(X,X)}}(0, \delta) \text{ we have that } \varepsilon(H \circ L^{-1}) \text{ is defined}$$

and by inserting $H \circ L^{-1}$ for H in [eq: 16.332]

$$\tau(\mathrm{Id}_X + H \circ L^{-1}) - \tau(\mathrm{Id}_X) = -H \circ L + \|H \circ L^{-1}\|_{L(X,X)} \cdot \varepsilon(H \circ L^{-1}) \tag{16.338}$$

Hence

$$\begin{array}{rcl} \tau(L + H) - \tau(H) & = & \\ (L + H)^{-1} - L^{-1} & \underset{[\text{eq: } 16.336]}{=} & \\ ((\mathrm{Id}_X + H \circ L^{-1}) \circ L)^{-1} - L^{-1} & = & \\ L^{-1} \circ (\mathrm{Id}_X + H \circ L^{-1})^{-1} - L^{-1} & = & \\ L^{-1} \circ (\mathrm{Id}_X + H \circ L^{-1})^{-1} - L^{-1} \circ \mathrm{Id}_X & \underset{[\text{theorem: } 14.197]}{=} & \\ L^{-1} \circ ((\mathrm{Id}_X + H \circ L^{-1})^{-1} - \mathrm{Id}_X) & = & \\ L^{-1} \circ (\tau(\mathrm{Id}_X + H \circ L^{-1}) - \mathrm{Id}_X) & = & \\ L^{-1} \circ (\tau(\mathrm{Id}_X + H \circ L^{-1}) - \mathrm{Id}_X^{-1}) & = & \\ L^{-1} \circ (\tau(\mathrm{Id}_X + H \circ L^{-1}) - \tau(\mathrm{Id}_X)) & \underset{[\text{eq: } 16.338]}{=} & \\ L^{-1} \circ (-H \circ L^{-1} + \|H \circ L^{-1}\|_{L(X,X)} \cdot \varepsilon(H \circ L^{-1})) & \underset{[\text{theorem: } 14.197]}{=} & \\ -L \circ H \circ L^{-1} + \|H \circ L^{-1}\|_{L(X,X)} \cdot L^{-1} \circ \varepsilon(H \circ L^{-1}) & & \end{array}$$

proving that

$$\tau(L + H) - \tau(H) = -L \circ H \circ L^{-1} + \|H \circ L^{-1}\|_{L(X,X)} \cdot L^{-1} \circ \varepsilon(H \circ L^{-1}) \tag{16.339}$$

As $0\in L(X,X)$ and by]theorem: 14.185] $\frac{\|H\circ L^{-1}\|_{L(X,X)}}{\|H\|_{L(X,X)}}\cdot L^{-1}\circ\varepsilon(H\circ L^{-1})\in L(X,X)$ we can define the following function

$$\varepsilon'\colon B_{\|\|_{L(X,X)}}(0,\delta)\rightarrow L(X,X) \text{ by } \varepsilon'(H)=\begin{cases} 0 \text{ if } H=0\\ \frac{\|H\circ L^{-1}\|_{L(X,X)}}{\|H\|_{L(X,X)}}\cdot L^{-1}\circ\varepsilon(H\circ L^{-1}) \text{ if } H\neq 0\end{cases}$$

then $\forall H\in B_{\|\|_{L(X,X)}}(0,\delta)$

$$\|H\|_{L(X,X)}\cdot\varepsilon'(H)=\|H\circ L^{-1}\|_{L(X,X)}\cdot L^{-1}\circ\varepsilon(H\circ L^{-1}) \tag{16.340}$$

Further for $H\in B_{\|\|_{L(X,X)}}(0,\delta)$ and $x\in X$ we have

$$\begin{aligned}
\|\varepsilon'(H)(x)\| &= \\
\left\|\begin{cases} 0(x) \text{ if } H=0\\ \frac{\|H\circ L^{-1}\|_{L(X,X)}}{\|H\|_{L(X,X)}}\cdot(L^{-1}\circ\varepsilon(H\circ L^{-1}))(x) \text{ if } H\neq 0\end{cases}\right\| &= \\
\left\|\begin{cases} 0 \text{ if } H=0\\ \frac{\|H\circ L^{-1}\|_{L(X,X)}}{\|H\|_{L(X,X)}}\cdot(L^{-1}\circ\varepsilon(H\circ L^{-1}))(x) \text{ if } H\neq 0\end{cases}\right\| & \\
\begin{cases} \|x\|\cdot 0 \text{ if } H=0\\ \frac{\|H\circ L^{-1}\|_{L(X,X)}}{\|H\|_{L(X,X)}}\cdot\|(L^{-1}\circ\varepsilon(H\circ L^{-1}))(x)\| \text{ if } H\neq 0\end{cases} &\leqslant \\
\begin{cases} \|x\|\cdot 0 \text{ if } H=0\\ \frac{\|H\circ L^{-1}\|_{L(X,X)}}{\|H\|_{L(X,X)}}\cdot\|(L^{-1}\circ\varepsilon(H\circ L^{-1}))\|_{L(X,X)}\cdot\|x\| \text{ if } H\neq 0\end{cases} &= \\
\|x\|\cdot\begin{cases} 0 \text{ if } H=0\\ \frac{\|H\circ L^{-1}\|_{L(X,X)}}{\|H\|_{L(X,X)}}\cdot\|(L^{-1}\circ\varepsilon(H\circ L^{-1}))\|_{L(X,X)} \text{ if } H\neq 0\end{cases} &\leqslant_{[14.185]} \\
\|x\|\cdot\begin{cases} 0 \text{ if } H=0\\ \frac{\|H\circ L^{-1}\|_{L(X,X)}\cdot}{\|H\|_{L(X,X)}}\|\cdot L^{-1}\|_{L(X,X)}\cdot\|\varepsilon(H\circ L^{-1})\|_{L(X,X)} \text{ if } H\neq 0\end{cases} &\leqslant_{[14.185]} \\
\|x\|\cdot\begin{cases} 0 \text{ if } H=0\\ \frac{\|H\|_{L(X,X)}\cdot}{\|H\|_{L(X,X)}}\cdot\|L^{-1}\|^2_{L(X,X)}\cdot\|\varepsilon(H\circ L^{-1})\|_{L(X,X)} \text{ if } H\neq 0\end{cases} &= \\
\|x\|\cdot\begin{cases} 0 \text{ if } H=0\\ \|L^{-1}\|^2_{L(X,X)}\cdot\|\varepsilon(H\circ L^{-1})\|_{L(X,X)} \text{ if } H\neq 0\end{cases} &\leqslant \\
\|x\|\cdot\|L^{-1}\|^2_{L(X,X)}\cdot\|\varepsilon(H\circ L^{-1})\|_{L(X,X)} &
\end{aligned}$$

proving by [theorem: 14.181] that

$$\|\varepsilon'(H)\|_{L(X,X)}\leqslant\|L^{-1}\|^2_{L(X,X)}\cdot\|\varepsilon(H\circ L^{-1})\|_{L(X,X)} \tag{16.341}$$

Let $\zeta\in\mathbb{R}^+$ then as ε is continue at 0 [see 16.330] there exist a $\delta_1\in\mathbb{R}^+$ such that if $\|H\|_{L(X,X)}<\delta_1$ then

$$\|\varepsilon(H)\|_{L(X,X)}<\frac{\zeta}{\|L^{-1}\|^2_{L(X,X)}+1}$$

Take $\delta_2=\frac{\delta_1}{\|L^{-1}\|_{L(X,X)}+1}$ then if $\|H\|_{L(X,X)}<\delta_2$ we have

$$\|H\circ L^{-1}\|_{L(X,X)}\leqslant_{[14.185]}\|H\|_{L(X,X)}\cdot\|L^{-1}\|_{L(X,X)}<\frac{\delta_1}{\|L^{-1}\|_{L(X,X)}+1}\cdot\|L^{-1}\|_{L(X,X)}<\delta_1$$

so that

$$\|\varepsilon(H\circ L^{-1})\|_{L(X,X)}<\frac{\zeta}{\|L^{-1}\|^2_{L(X,X)}+1}$$

Substituting the above in [eq: 16.341] we have for $\|H\|_{L(X,Y)}<\min(\delta,\delta_2)$ that

$$\|\varepsilon'(H)-\varepsilon'(0)\|=\|\varepsilon'(H)\|_{L(X,X)}\leqslant\|L^{-1}\|^2_{L(X,X)}\cdot\frac{\zeta}{\|L^{-1}\|^2_{L(X,X)}+1}<\zeta$$

which proves that ε' is continuous at 0. Hence

$$\varepsilon' \text{ is a } \varepsilon\text{-mapping at } 0 \tag{16.342}$$

Define now

$$\psi: L(X,X)\rightarrow L(X,X) \text{ by } \psi(H)=-L^{-1}\circ H\circ L$$

Then for $H_1,H_2\in L(X,X)$ and $\alpha\in\mathbb{K}$ we have

$$\begin{array}{rcl}\psi(H_1+\alpha\cdot H_2) & = & -L^{-1}\circ(H_1+\alpha\cdot H_2)\circ L\\ & \underset{\text{associativity}}{=} & -L^{-1}\circ((H_1+\alpha\cdot H_2)\circ L)\\ & \underset{[\text{theorem: }14.197]}{=} & -L^{-1}\circ(H_1\circ L+\alpha\cdot H_2\circ L)\\ & \underset{[\text{theorem: }14.197]}{=} & -L^{-1}\circ H_1\circ L+\alpha\cdot(-L^{-1}\circ H_2\circ L)\\ & = & \psi(H_1)+\alpha\cdot\psi(H_2)\end{array}$$

proving that

$$\psi\in\operatorname{Hom}(L(X,X),L(X,X))$$

Further for $H\in L(X,X)$ we have

$$\begin{array}{rcl}\|\psi(H)\|_{L(X,X)} & = & \|-L\circ H\circ L\|_{L(X,X)}\\ & \leqslant_{[\text{theorem: }14.185]} & \|-L\circ H\|_{L(X,X)}\cdot\|L\|_{L(X,X)}\\ & \leqslant & \|L\|_{L(X,X)}\cdot\|H\|_{L(X,X)}\cdot\|L\|_{L(X,X)}\\ & = & \|L\|^2_{L(X,X)}\cdot\|H\|_{L(X,X)}\end{array}$$

proving by [theorem: 14.181] that

$$\psi\in L(L(X,X),L(X,X)) \text{ and } \|\psi\|_{L(L(X,X),L(X,X))}\leqslant\|L\|^2_{L(X,X)} \tag{16.343}$$

Next

$$\begin{array}{rcl}\tau(L+H)-\tau(L)-\psi(H) & = & \tau(L+H)-\tau(L)-(-L\circ H\circ L)\\ & \underset{[\text{eq: }16.339]}{=} & \|H\circ L^{-1}\|_{L(X,X)}\cdot L^{-1}\circ\varepsilon(H\circ L^{-1})\\ & \underset{[\text{eq: }16.340]}{=} & \|H\|_{L(X,X)}\cdot\varepsilon'(H)\end{array}$$

Finally using [eqs: 16.337, 16.342 and the above allows us to] apply [corollary: 16.30] proving that τ is Fréchet differentiable at L and

$$D_L\tau=\psi$$

so that $D_L\tau\in L(L(X,X),L(X,X))$ is defined by $D_L\tau(H)=\psi(H)=-L\circ H\circ L$.

So in all cases we have that $\forall L\in\mathcal{GL}(X)$ τ is Fréchet differentiable at L and $D_L\tau\in L(L(X,X),L(X,X))$ is defined by $D_L\tau(H)=-L\circ H\circ L$.

3. Using (2) we have that $\forall L\in\mathcal{GL}(X)$ τ has a Fréchet differential $D_L\tau$ at L where $D_L\tau\in L(L(X,X),L(X,X))$ is defined by $D_L\tau(H)=-L\circ H\circ L$. Hence we can define

$$D^{[1]}\tau:\mathcal{GL}(X)\rightarrow L(L(X,X),L(X,X)) \text{ by } D^{[1]}\tau(L)=D_L\tau$$

where

$$D_L\tau \text{ is defined by } D_L(\tau)(H)=-L\circ H\circ L \tag{16.344}$$

We proceed now to show that $D^{[1]}\tau$ is the composition of ∞-times differentiable mappings and τ. Define

$$\chi:\mathcal{GL}(X)\rightarrow\mathcal{GL}(X)\cdot\mathcal{GL}(X) \text{ by } \chi(L)=(L,L)$$

then $\forall L\in\mathcal{GL}(X)$ we have

$$(\pi_1\circ\chi)(L)=\pi_1(\chi(L))=\pi_1((L,L))=L=\operatorname{Id}_{\mathcal{GL}(X)}(L)$$

and

$$(\pi_2 \circ \chi)(L) = \pi_2(\chi(L)) = \pi_2((L, L)) = L = \mathrm{Id}_{\mathcal{GL}(X)}(L)$$

so that $\pi_1 \circ \chi = \mathrm{Id}_{\mathcal{GL}(X)}$, $\pi_2 \circ \chi = \mathrm{Id}_{\mathcal{GL}(X)}$. As $\mathrm{Id}_{\mathcal{GL}(X)}$ is of class C^∞ [see theorem: 16.115], it follows from [theorem: 16.61] that

$$\mathcal{X} \text{ is of class } C^\infty \tag{16.345}$$

Given $L, K \in L(X, X)$ define

$$\varphi_{L,K}: L(X, X) \rightarrow L(X, X) \text{ by } \varphi_{L,K}(H) = -L \circ H \circ K \tag{16.346}$$

Let $H_1, H_2 \in L(X, X)$, $\alpha \in \mathbb{K}$ then we have

$$\begin{array}{rcl} \varphi_{L,K}(H_1 + \alpha \cdot H_2) & = & -L \circ (H_1 + \alpha \cdot H_2) \circ K \\ & \underset{[\text{theorem: } 14.197]}{=} & -L \circ (H_1 \circ K + \alpha \cdot H_2 \circ K) \\ & \underset{[\text{theorem: } 14.197]}{=} & -L \circ H_1 \circ K + \alpha \cdot (-L \circ H_2 \circ K) \\ & = & \varphi_{L,K}(H_1) + \alpha \cdot \varphi_{L,K}(H_2) \end{array}$$

proving that

$$\varphi_{L,K} \in \mathrm{Hom}(L(X, X), L(X, X))$$

Further if $H \in L(X, X)$ then

$$\begin{array}{rcl} \|\varphi_{L,K}(H)\|_{L(X,X)} & = & \|-L \circ H \circ K\|_{L(X,X)} \\ & = & \|-L \circ (H \circ K)\|_{L(X,X)} \\ & \leqslant_{[\text{theorem: } 14.185]} & \|L\|_{L(X<X)} \cdot \|H \circ K\|_{L(X,X)} \\ & \leqslant_{[\text{theorem: } 14.185]} & \|L\|_{L(X<X)} \cdot \|H\|_{L(X,X)} \cdot \|K\|_{L(X,X)} \\ & = & (\|L\|_{L(X,X)} \cdot \|K\|_{L(X,X)}) \cdot \|H\|_{L(X,X)} \end{array}$$

hence, using [theorems: 14.174 and 14.181] it follows that

$$\varphi_{L,K} \in L(L(X, X), L(X, X)) \wedge \|\varphi_{L,K}\|_{L(L(X,X),L(X,X))} \leqslant \|L\|_{L(X,X)} \cdot \|K\|_{L(X,X)} \tag{16.347}$$

So we can define

$$\varphi: L(X, X) \cdot L(X, X) \rightarrow L(L(X, X), L(X, X)) \text{ by } \varphi(L, K) = \varphi_{L,K} \tag{16.348}$$

Then we have for $L, L_1, L_2 \in L(X, X)$ and $\alpha \in \mathbb{K}$ that

$$\begin{array}{rcl} \varphi(L, L_1 + \alpha \cdot L_2)(H) & \underset{[\text{eq: } 16.348}{=} & \varphi_{L, L_1 + \alpha \cdot L_2}(H) \\ & \underset{[\text{eq: } 16.346]}{=} & -L \circ H \circ (L_1 + \alpha \cdot L_2) \\ & \underset{[\text{theorem: } 14.197]}{=} & -L \circ (H \circ L_1 + \alpha \cdot H \circ L_2) \\ & \underset{[\text{theorem: } 14.197]}{=} & -L \circ H \circ L_1 + \alpha \cdot (-L \circ H \circ L_2) \\ & = & \varphi(L, L_1)(H) + \alpha \cdot \varphi(L, L_2)(H) \\ & = & (\varphi(L, L_1) + \alpha \cdot \varphi(L, L_2))(H) \\ \varphi(L_1 + \alpha \cdot L_2, L)(H) & = & \varphi_{L_1 + \alpha \cdot L_2, L}(H) \\ & \underset{[\text{eq: } 16.346]}{=} & -(L_1 + \alpha \cdot L_2) \circ H \circ L \\ & = & (-(L_1 + \alpha \cdot L_2) \circ H) \circ L \\ & \underset{[\text{theorem: } 14.197]}{=} & (-L_1 \circ H + \alpha \cdot (-L_2 \circ H)) \circ L \\ & \underset{[\text{theorem: } 14.197]}{=} & -L_1 \circ H \circ L + \alpha \cdot (-L_2 \circ H \circ L) \\ & = & \varphi(L_1, L)(H) + \alpha \cdot \varphi(L_2, L)(H) \\ & = & (\varphi(L_1, L) + \alpha \cdot \varphi(L_2, L))(H) \end{array}$$

which proves that

$$\varphi(L, L_1 + \alpha \cdot L_2) = \varphi(L_1, L) + \alpha \cdot \varphi(L_2, L) \text{ and } \varphi(L_1 + \alpha \cdot L_2, L) = \varphi(L_1, L) + \alpha \cdot \varphi(L_2, L)$$

Hence

$$\varphi \in \mathrm{Hom}(L(X,X), L(X,X); L(L(X,X), L(X,X)))$$

Further if $(L,K) \in L(X,X) \cdot L(X,X)$ then we have

$$\begin{aligned} \|\varphi(L,K)\|_{|L(L(X,X),L(X,X))} &\underset{[\text{eq: } 16.348]}{=} \|\varphi_{L,K}\|_{L(L(X,X),L(X,X))} \\ &\leqslant_{[\text{eq: } 16.347]} \|L\|_{L(X,X)} \cdot \|K\|_{L(X,X)} \end{aligned}$$

which by [theorem: 14.187] proves that

$$\varphi \in L(L(X,X), L(X,X); L(L(X,X), L(X,X)))$$

Hence by [theorem: 16.120] we have that

$$\varphi \text{ is of class } C^\infty \tag{16.349}$$

Take $L \in \mathcal{G}L(X)$ then for $H \in L(X,X)$ we have

$$\begin{aligned} ((\varphi \circ \chi \circ \tau)(L))(H) &= (\varphi(\chi(\tau(L))))(H) \\ &= (\varphi(\chi(L^{-1})))(H) \\ &= (\varphi(L^{-1}, L^{-1}))(H) \\ &\underset{[\text{eq: } 16.348}{=} \varphi_{L^{-1},L^{-1}}(H) \\ &\underset{[\text{eq: } 16.346]}{=} -L \circ H \circ L \\ &\underset{[\text{eq: } 16.344]}{=} D_L\tau(H) \end{aligned}$$

proving that $\forall L \in \mathcal{G}L(X)$ $(\varphi \circ \chi \circ \tau)(L) = D_L\tau = D^{[1]}\tau(L)$. Hence we have

$$D^{[1]}\tau = \varphi \circ \chi \circ \tau \tag{16.350}$$

We use now induction for the final part of the proof for (3). So define

$$S = \{n \in \mathbb{N}_0 | \tau \text{ is of class } C^n\}$$

then we have:

$0 \in S$. As $\forall L \in \mathcal{G}L(X)$ τ is Fréchet differentiable at L, it follows by [theorem: 16.31] that τ is continuous. Hence τ is of class C^0 proving that $0 \in S$.

$n \in S \Rightarrow n+1 \in S$. By [eqs: 16.345, 16.349] φ, χ are of class C^∞ and, as $n \in S$, τ is of class C^n. So we can use the chain rule [see theorem: 16.129] to prove that $\varphi \circ \chi \circ \tau$ is of class C^n. Hence by [eq: 16.350] $D^{[1]}\tau$ is of class C^n, which by [theorem: 16.111] means that τ is of class C^{n+1}. So that

$$n+1 \in S$$

By mathematical induction $S = \mathbb{N}$ so that τ is of class C^∞. □

We introduce now the concept of differentiable functions who have a differentiable inverse.

Definition 16.240. *Let $\langle X, \|\|_X \rangle$, $\langle Y, \|\|_Y \rangle$ be normed spaces, U a open set in X, V a open set in Y and $f: U \rightarrow V$ a function then we say that*

1. *f is a **diffeomorphism** iff*
   a. *$f: U \rightarrow V$ is a bijection*
   b. *$f: U \rightarrow Y$ is differentiable on U*
   c. *$f^{-1}: V \rightarrow X$ is differentiable on V*
2. *f is a **diffeomorphism of class C^n** where $n \in \mathbb{N}_0 \bigcup \{\infty\}$ iff*
   a. *$f: U \rightarrow V$ is a bijection*
   b. *$f: U \rightarrow Y$ is of class C^n*
   c. *$f^{-1}: V \rightarrow X$ is of class C^n*

Example 16.241. Let $\langle X, \|\|_X \rangle$, $\langle Y, \|\|_Y \rangle$ be normed spaces then a homeomorphism $f: X \rightarrow Y$ is of class C^0

Proof. This follows trivially from the definitions of a homeomorphism [definition: 14.162] and diffeomorphism of class C^0 [definition: 16.240]. □

Example 16.242. Let $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$ be Banach spaces and $L: X\rightarrow Y$ a continuous linear isomorphism [in other words $L\in L(X,Y)$ and L is a bijection] then L is a diffeomorphism of class C^∞.

Proof. First using [theorem: 14.388] is a homeomorphism so that L, L^{-1} are continuous. Taking in account [theorem: 11.169] we have also that L^{-1} is linear. So $L\in L(X,Y)$ and $L^{-1}\in L(Y,X)$. Finally by [example: 16.116] it follows that L,L^{-1} are of class C^∞ proving that L is a diffeomorphism of class C^∞. □

Theorem 16.243. *Let $n\in\mathbb{N}_0$, $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$ be normed spaces, U a open set in X, V a open set in Y and $f: U\rightarrow V$ a diffeomorphism of class C^n then for every open set $W\subseteq U$ we have that*

$$f_{|W}: W\rightarrow f(W)$$

is a diffeomorphism of class C^n.

Proof. As f is a diffeomorphism of class C^n $f: U\rightarrow Y$ is of class C^n and $f^{-1}: V\rightarrow X$ is of class C^n. Using [theorem: 16.97] it follows that $f_{|W}: W\rightarrow Y$ is a of class C^n. Further as $f^{-1}: V\rightarrow Y$ is of class C^n it is of class C^0 hence continuous. So $f(W) \underset{\text{[theorem: 2.68]}}{=} (f^{-1})^{-1}(W)$ is open in the subspace topology of V hence open in the topology of X [see theorem: 14.11]. Applying then [theorem: 16.97] it follows that $(f^{-1})_{|f(W)}: f(W)\rightarrow X$ is of class C^n. Finally by [theorem: 2.90] $(f^{-1})_{|f(W)} = (f_{|W})^{-1}$ so that $(f_{|W})^{-1}$ is of class C^n. Hence $f_{|W}: W\rightarrow Y$ is a diffeomorphism of class C^n. □

Theorem 16.244. *Let $n\in\mathbb{N}_0$, $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$ be normed spaces, U a open set in X, V a open set in Y and $f: U\rightarrow V$ then we have*

f is a diffeomorphism [of class C^n]

$$\Updownarrow$$

$f: U\rightarrow Y$ is a differentiable on U [or of class C^n] and there exist a function $g: V\rightarrow U$ that is a diffeomorphism [of class C^n] and $f\circ g=\mathrm{Id}_Y\wedge g\circ f=\mathrm{Id}_X$

Proof.

- **⇒.** If f is a diffeomorphism [of class C^n] then $f: U\rightarrow Y$ is differentiable on U [or of class C^n] and $f^{-1}: V\rightarrow X$ is differentiable [or of class C^n]. Take $g=f^{-1}$ then g is differentiable on V [or of class C^n], $f\circ g=f\circ f^{-1}=\mathrm{Id}_Y$ and $g\circ f=f^{-1}\circ f=\mathrm{Id}_X$.
- **⇐.** As $f\circ g=\mathrm{Id}_Y\wedge g\circ f=\mathrm{Id}_X$ it follows from [theorems: 2.71, 2.73] that $g=f^{-1}$ proving that f is a diffeomorphism [or of class C^n]. □

We examine now what happens when diffeomorphism's are composed.

Theorem 16.245. *Let $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$ be normed spaces, U a open set in X, V a open set in Y and $f: U\rightarrow V$ a diffeomorphism and $g=f^{-1}: V\rightarrow U$ then $\forall x\in U$ we have that*

$$D_{f(x)}g\circ D_xf=\mathrm{Id}_X \text{ and } D_xf\circ D_{f(x)}g=\mathrm{Id}_Y$$

so that

$$D_{f(x)}g=(D_xf)^{-1}$$

Further $D_{f(x)}g$ and D_xf are toplinear isomorphisms [see definition: 16.228]

Proof. As $\mathrm{Id}_X\in L(X,X)$ and $\mathrm{Id}_Y\in L(Y,Y)$ it follows from [example: 16.116] that Id_X, Id_Y are Fréchet differentiable on X, Y with $\forall x\in X$ that $D_x\mathrm{Id}_X=\mathrm{Id}_X$ and $D_{f(x)}\mathrm{Id}_Y=\mathrm{Id}_Y$. Further using the chain rule [see theorem: 16.36] we have

$$\mathrm{Id}_X=D_x\mathrm{Id}_X=D_x(f^{-1}\circ f)=D_x(g\circ f)=D_{f(x)}g\circ D_xf$$

and

$$\mathrm{Id}_Y=D_{f(x)}\mathrm{Id}_Y=D_{f(x)}(f\circ f^{-1})=D_{f^{-1}(f(x))}f\circ D_{f(x)}f^{-1}=D_xf\circ D_{f(x)}g$$

Further as $D_{f(x)}g\circ D_xf=\mathrm{Idx}$ and $D_xf\circ D_{f(x)}g=\mathrm{Id}_Y$ it follows from [theorem: 2.71, 2.73] that $D_{f(x)}g$, D_xf are bijections and $(D_xf)^{-1}=D_{f(x)}g\in L(Y,X)$, $(D_{f(x)}g)^{-1}=D_xf\in L(X,Y)$ proving that $(D_xf)^{-1},(D_{f(x)}g)^{-1}$ are continuous. Applying then [theorem: 16.227] proves that $D_{f(x)}g$ and D_xf are toplinear isomorphisms. □

Composition of diffeomorphism's is again a diffeomorphism as is show in the next theorem.

Theorem 16.246. *Let $n \in \mathbb{N}_0$, $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$, $\langle Z, \|\|_Z\rangle$ be normed spaces, U a open set in X, V a open set in Y, W a open set in X, $f: U \to V$ a diffeomorphism [of class C^n] and $g: V \to W$ a diffeomorphism of class C^n then $g \circ f: V \to W$ is a diffeomorphism [of class C^n].*

Proof. Using [theorem: 2.75] it follows that $g \circ f$ is a bijection and $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$. If f, g are diffeomorphism's then f is differentiable on U, f^{-1}, g are differentiable on V and g^{-1} is differentiable on W. Using the chain rule [theorem: 16.36] $g \circ f$ is differentiable on U and $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$ is differentiable on W, hence $g \circ f$ is a diffeomorphism. Likewise if f, g are diffeomorphism's of class C^n then f, g, f^{-1}, g^{-1} are of class C^n, hence, using the chain rule [theorem: 16.129], $g \circ f$ and $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$ are of class C^n. From this it follows that $g \circ f$ are differentiable of class C^n. □

We weaken now the concept of a diffeomorphism to that of a local diffeomorphism.

Definition 16.247. *Let $n \in \mathbb{N}_0$, $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$ be normed spaces, U a open set in X, $x \in U$ and $f: U \to Y$ a function then* ***f is a local diffeomorphism [of class C^n] at x*** *if there exist a open set U_x in X with $x \in U_x \subseteq U$ and a open set $V_{f(x)}$ in Y with $f(x) \in V_{f(x)}$ such that*

$$f_{|U_x}: U_x \to V_{f(x)}$$

is a diffeomorphism [of class C^n].

We are ow ready to state and prove the inverse function theorem.

Theorem 16.248. (Inverse Function Theorem) *Let $n \in \mathbb{N}$, $\langle X, \|\|_X\rangle$, $\langle Y, \|\|_Y\rangle$ be Banach spaces, U a open set in X, $x_0 \in U$, f: $U \to Y$ a function of class C^n such that $D_{x_0} f \in L_n(X, Y)$ is a bijection then f is a local diffeomorphism of class C^n at x_0. In other words there exist a open set U_{x_0} in X with $x_0 \in U_{x_0} \subseteq U$ and a open set $V_{f(x_0)}$ in Y with $f(x_0) \in V_{f(x_0)}$ such that*

$$f_{|U_{x_0}}: U_{x_0} \to V_{f(x_0)} \text{ is a diffeomorphism of class } C^n$$

Proof. The proof is very long and will span several pages, so we split it up in different propositions and conjectures. First note that by the definition of Fréchet differential [see definition: 16.18] that $D_{x_0} f \in L(X, X)$, as by the hypothesis $D_{x_o} f$ is a bijection and X, Y are bijections it follows from [theorem: 16.228] that

$$D_{x_0} f \text{ is a toplinear isomorphism} \tag{16.351}$$

Next we prove a limited version of the Inverse Function Theorem and extend it later to the more general case.

Proposition 16.249. *Let $n \in \mathbb{N}$, $\langle X, \|\|_X\rangle$ be a Banach space and U a open set with $0 \in U$ and $f: U \to X$ a function of class C^n such that $D_0 f = \mathrm{Id}_X$ and $f(0) = 0$ then f is a local diffeomorphism of class C^nat 0.*

Proof. Define

$$T = \mathrm{Id}_X - f: U \to X \tag{16.352}$$

Then

$$T(0) = \mathrm{Id}_X(0) - f(0) = 0$$

By [example: 16.116] $\mathrm{Id}_X \in L(X, X)$ is of class C^∞ hence by [theorem: 16.107] of class C^n. So using [theorem: 16.113] it follows that

$$T \text{ is of class } C^n \tag{16.353}$$

and using [theorem: 16.107] again it follows that T is of class C^1 hence we have that

$$DT: U \to L(X, X) \text{ defined by } DT(x) = D_x T \text{ is continuous} \tag{16.354}$$

and $\forall x \in U$ we have

$$DT(x) = D_x T = D_x \mathrm{Id}_X - D_x f = \mathrm{Id}_X - Df(x) \tag{16.355}$$

so that

$$DT(0) = D_0 f = \mathrm{Id}_x - \mathrm{Id}_X = 0 \in L(X, X) \tag{16.356}$$

As DT is continuous at 0 [by eq: 16.354] there exist a $\rho_1 \in \mathbb{R}^+$ such that $\forall x \in U$ with $\|x\| < \rho_1$ we have $\|DT(x)\|_{L(X,X)} = \|DT(x) - 0\|_{L(X,X)} \underset{\text{[eq: 16.356]}}{=} \|DT(x) - DT(0)\|_{L(X,X)} < \frac{1}{2}$. Further as $0 \in U$ there exist a $\rho_2 \in \mathbb{R}^+$ such that $0 \in B_{\|\|_X}(0, \delta_2) \subseteq U$. Take $\rho_3 = \min(\rho_1, \rho_2) \in \mathbb{R}^+$ then if $x \in B_{\|\|_X}(0, \rho_3)$ we have $\|x\| < \delta_2 \Rightarrow x \in U$ and $\|x\| \leqslant \delta_1 \Rightarrow \|Df(x)\|_X < \frac{1}{2}$. Take $\rho = \frac{\rho_3}{2}$ then we have

$$\overline{B_{\|\|_X}(0, \rho)} \subseteq B_{\|\|_X}(0, \rho_3) \subseteq U \text{ and } \forall x \in \overline{B_{\|\|_X}(0, \rho)} \text{ we have } \|DT(x)\|_{L(X,X)} < \frac{1}{2} < 1 \tag{16.357}$$

If $x \in B_{\|\|_X}(0, \rho)$ then by the above $DT(x) \in B_{\|\|_{L(X,X)}}(0, 1)$ so that by [lemma: 16.236] $\mathrm{Id}_X - DT(x) \in \mathcal{GL}(X)$ and using [eq: 16.355] that $D_x f \in \mathcal{GL}(X)$. Summarized we have

$$\forall x \in B_{\|\|_X}(0, \rho) \;\; D_x f \in \mathcal{GL}(X) \text{ [hence by [theorem: 16.230] } (D_x f)^{-1} \in \mathcal{GL}(X) \subseteq L(X, X)] \tag{16.358}$$

As $\overline{B_{\|\|_X}(0, \rho)}$ is convex [see theorem: 16.157] we can use [eq: 16.357] to apply the mean value theorem [theorem: 16.158] giving

$$\forall x \in \overline{B_{\|\|_X}(0, \rho)} \text{ we have } \|T(x)\| = \|T(x) - T(0)\| \leqslant \frac{1}{2} \cdot \|x - 0\| = \frac{1}{2} \cdot \|x\| \tag{16.359}$$

To finish the proposition we prove the following conjectures.

Conjecture 16.250. *If* $0 < \rho' < \rho$ *then* $\forall y \in \overline{B_{\|\|_X}\left(0, \frac{\rho'}{2}\right)}$ *there exist a* ***uniquely determined*** $x \in \overline{B_{\|\|_X}(0, \rho')}$ *such that* $y = f(x)$.

Proof. Let $y \in \overline{B_{\|\|_X}\left(0, \frac{\rho'}{2}\right)}$ and define

$$T_y: U \rightarrow X \text{ by } T_y(x) \underset{\text{[eq: 16.352]}}{=} y + T(x) = y + x - f(x)$$

in other words

$$T_y = C_y + T$$

Hence as C_y is of class C^∞ [see example: 16.115, 16.116 and T is of class C^n [see eq: 16.353] it follows from [theorem: 16.113] that

$$T_y \text{ is of class } C^n$$

Let $x \in \overline{B_{\|\|_X}(0, \rho')} \subseteq \overline{B_{\|\|_X}(0, \rho)} \subseteq_{\text{[eq: 16.357]}} U$ then we have

$$\|T_y(x)\| = \|y + T(x)\| \leqslant \|y\| + \|T(x)\| <_{\text{[eq: 16.359]}} \|y\| + \frac{1}{2} \cdot \|x\| < \frac{\rho'}{2} + \frac{1'}{2} \cdot \|x\| \leqslant \frac{\rho'}{2} + \frac{\rho'}{2} = \rho'$$

proving that

$$\forall x \in \overline{B_{\|\|_X}(0, \rho')} \text{ we have } \|T_y(x)\| \leqslant \rho' \text{ so that } T_y(x) \in B_{\|\|_X}(0, \rho') \tag{16.360}$$

Define

$$DT_y: U \rightarrow L(X, X) \text{ by } DT_y(x) = D_x T_y = D_x C_y + D_x T = 0 + D_x T = D_x T \tag{16.361}$$

Using [eq: 16.357] on the above results

$$\forall x \in \overline{B_{\|\|_X}(0, \rho')} \subseteq \overline{B_{\|\|_X}(0, \rho)} \text{ we have } \|DT_y\|_{L(X,X)} = \|D_x T\|_{L(X,X)} < \frac{1}{2} \tag{16.362}$$

As X is a Banach space and $\overline{B_{\|\|_X}(0, \rho')}$ is convex we can apply the mean value theorem [theorem: 16.158] on the above to get

$$\forall x_1, x_2 \in \overline{B_{\|\|_X}(0, \rho')} \text{ we have } \|T_y(x_1) - T_y(x_2)\| \leqslant \frac{1}{2} \cdot \|x_1 - x_2\| \tag{16.363}$$

Using [eq: 16.360] it follows that the following function is well defined

$$S_y = (T_y)_{|\overline{B_{\|\|_X}(0,\rho')}}: \overline{B_{\|\|_X}(0, \rho')} \rightarrow \overline{B_{\|\|_X}(0, \rho')}$$

and that this function is a contraction [see definition: 14.380]. As $\overline{B_{\|\|_X}(0, \rho')}$ is closed and X is a Banach space it follows from [theorem: 14.329] that $\overline{B_{\|\|_X}(0, \rho')}$ is a Banach space. So we can apply the Banach Fixed Point theorem [see theorem: 14.382] giving

$$\text{There exist a } \textbf{unique } x \in \overline{B_{\|\|_X}(0, \rho')} \text{ such that } T_y(x) = x \tag{16.364}$$

From the definition of T_y we have that $f(x) = y + x - T_y(x) \underset{\text{[eq: 16.364]}}{=} y + x - x = y$ proving that

$$y = f(x)$$

As for uniqueness assume that there exist a $x' \in \overline{B_{\|\|_X}(0, \rho')}$ such that $y = f(x')$ then we have that $T_y(x') = y + x' - f(x') = y + x' - y = x'$ so that by [eq: 16.364] $x = x'$. □

Conjecture 16.251. *For the open set*

$$U_\rho = B_{\|\|_X}\left(0, \frac{\rho}{4}\right) \bigcap f^{-1}\left(B_{\|\|_X}\left(0, \frac{\rho}{4}\right)\right)$$

we have that

$$0 \in U_\rho$$

and

$$f_{|U_\rho}: U_\rho \rightarrow B_{\|\|_X}\left(0, \frac{\rho}{4}\right)$$

is a bijection.

Proof. First as $f(0) = 0 \in B_{\|\|_X}\left(0, \frac{\rho}{4}\right)$ we have that

$$0 \in B_{\|\|_X}\left(0, \frac{\rho}{4}\right) \bigcap f^{-1}\left(B_{\|\|_X}\left(0, \frac{\rho}{4}\right)\right) = U_\rho$$

For the function

$$f_{|U_\rho}: U_\rho \rightarrow B_{\|\|_X}\left(0, \frac{\rho}{4}\right)$$

we have:

surjectivity. Let $y \in B_{\|\|_X}\left(0, \frac{\rho}{4}\right)$ then by the previous conjecture [conjecture: 16.250] there exist a $x \in \overline{B_{\|\|_X}\left(0, \frac{\rho}{2}\right)}$ such that

$$f(x) = y \tag{16.365}$$

We prove now by contradiction that

$$\|x\| \neq \frac{\rho}{2} \tag{16.366}$$

As $y \in B_{\|\|_X}\left(0, \frac{\rho}{4}\right)$ we have that $\|f(x)\| < \frac{\rho}{4}$ so that $\varepsilon = \frac{\rho}{4} - \|y\| \in \mathbb{R}^+$. Given that f is continuous [because f is of class C^n hence of class C^0 by [theorem: 16.107]] there exist a $\delta' \in \mathbb{R}^+$ such that if $z \in U$ with $\|x - z\| < \delta'$ then we have $\|f(x) - f(z)\| < \varepsilon$. Define $\delta = \min\left(\delta', \frac{\rho}{2}\right) \in \mathbb{R}^+$ and take $x_1 = \left(1 + \frac{\delta}{\rho}\right) \cdot x$ then we have $\|x\| < \left(1 + \frac{\delta}{\rho}\right) \cdot \|x\| = \|x_1\|$ and $\|x_1\| = \left(1 + \frac{\delta}{\rho}\right) \cdot \|x\| = \left(1 + \frac{\delta}{\rho}\right) \cdot \frac{\rho}{2} = \frac{\rho}{2} + \frac{\delta}{2} < \frac{\rho}{2} + \frac{\rho}{4} < \frac{\rho}{2} + \frac{\rho}{2} = \rho$ so that

$$\frac{\rho}{2} = \|x\| < \|x_1\| < \rho \tag{16.367}$$

Further we have

$$\|x - x_1\| = \left\|x - \left(1 + \frac{\delta}{\rho}\right) \cdot x\right\| = \left\|\frac{\delta}{\rho} \cdot x\right\| = \frac{\delta}{\rho} \cdot \|x\| \underset{\|x\| = \frac{\rho}{2}}{=} \frac{\delta}{\rho} \cdot \frac{\rho}{2} < \delta < \delta'$$

so that

$$\|f(x) - f(x_1)\| < \varepsilon = \frac{\rho}{2} - \|y\|$$

Hence

$$\begin{array}{lll} \|f(x_1)\| & = & \|f(x_1) - f(x + f(x))\| \\ & \leqslant & \|f(x_1) - f(x)\| + \|f(x)\| \\ & < & \frac{\rho}{4} - \|y\| + \|f(x)\| \\ & \underset{\text{[eq: 16.365]}}{=} & \frac{\rho}{4} - \|y\| + \|y\| \\ & = & \frac{\rho}{4} \end{array}$$

proving that $f(x_1) \in B_{\|\|_X}\left(0, \frac{\rho}{4}\right) \subseteq \overline{B_{\|\|_X}\left(0, \frac{\rho}{4}\right)}$. Using the previous conjecture [conjecture: 16.250] there exist a $x_2 \in \overline{B_{\|\|_X}\left(0, \frac{\rho}{2}\right)}$ such that $f(x_1) = f(x_2)$. As $\|x_2\| \leqslant \frac{\rho}{2} < \rho$, $\|x_1\| < \rho$ and $f(x_1) = f(x_2)$ we have by [conjecture: 16.250] that $x_1 = x_2$. Then we have

$$\frac{\rho}{2} <_{\text{[eq: 16.367]}} \|x_1\| = \|x_2\| \leqslant \frac{\rho}{2}$$

leading to the contradiction $\frac{r}{2} < \frac{r}{2}$. Hence our assumption is wrong and we must have that

$$\|x\| \neq \frac{\rho}{2}$$

So as $x \in \overline{B_{\|\|_X}\left(0, \frac{\rho}{2}\right)}$ we have $x \in B_{\|\|_X}\left(0, \frac{\rho}{2}\right)$, further as $f(x) = y \in B_{\|\|_X}\left(0, \frac{\rho}{4}\right)$ it follows that $x \in f^{-1}\left(B_{\|\|_X}\left(0, \frac{\rho}{4}\right)\right)$ so that $x \in B_{\|\|_X}\left(0, \frac{\rho}{4}\right) \bigcap f^{-1}\left(B_{\|\|_X}\left(0, \frac{\rho}{4}\right)\right) = U_x$.

injectivity. Assume that there exist $x_1, x_2 \in U_\rho \subseteq B_{\|\|_X}\left(0, \frac{\rho}{4}\right)$ such that $f(x_1) = f(x_2) \in B_{\|\|_X}\left(0, \frac{\rho}{4}\right)$ so that by the previous conjecture [conjecture: 16.250] $x_1 = x_2$. □

Proceeding with the proof of [proposition: 16.249] we have as $f_{|U_\rho}: U_\rho \rightarrow B_{\|\|_X}\left(0, \frac{\rho}{4}\right)$ is a bijection that

$$g = (f_{|U_\rho})^{-1}: B_{\|\|_X}\left(0, \frac{\rho}{4}\right) \rightarrow U_\rho \text{ where } U_\rho = B_{\|\|_X}\left(0, \frac{\rho}{4}\right) \bigcap f^{-1}\left(B_{\|\|_X}\left(0, \frac{\rho}{4}\right)\right) \tag{16.368}$$

We prove now that g is continuous.

Conjecture 16.252. *g is continuous.*

Proof. Let $x \in U_\rho$ then $x = x = x - f(x) + f(x) \underset{\text{[eq: 16.352]}}{=} T(x) + f(x)$ so if $x_1, x_2 \in U_\rho$ then

$$\begin{aligned} \|x_1 - x_2\| &= \|T(x_1) + f(x_1) - T(x_2) - f(x_2)\| \\ &\leqslant \|T(x_1) - T(x_2)\| + \|f(x_1) - f(x_2)\| \end{aligned}$$

As [eq: 16.368] $U_\rho \subseteq B_{\|\|_X}(0, \frac{\rho}{4}) \subseteq \overline{B_{\|\|_X}(0, \frac{\rho}{4})}$ which is convex [see theorem: 16.157] and $\forall x \in B_{\|\|_X}(0, \frac{\rho}{4})$ we have $\|D_x T\| < \frac{1}{2}$ [see eq: 16.357] we can apply the main value theorem [theorem: 16.158] giving $\|T(x_1) - T(x_2)\| \leqslant \frac{1}{2} \cdot \|x_1 - x_2\|$ so that

$$\|x_1 - x_2\| \leqslant \frac{1}{2} \cdot \|x_1 - x_2\| + \|f(x_1) - f(x_2)\|$$

hence

$$\|x_1 - x_2\| \leqslant 2 \cdot \|f(x_1) - f(x_2)\|$$

To summarize we have proved that

$$\forall x_1, x_2 \in U_\rho \text{ we have } \|x_1 - x_2\| \leqslant 2 \cdot \|f(x_1) - f(x_3)\| \tag{16.369}$$

Let $y_1,\ y_2 \in B_{\|\|_X}(0, \frac{\rho}{4})$ then $g(y_1), g(y_2) \in U_\rho$ so that by [eq: 16.369] we have

$$\|g(y_1) - g(y_2)\| \leqslant 2 \cdot \|f(g(y_1)) - f(g(y_2))\| = 2 \cdot \|y_1 - y_2\| \tag{16.370}$$

Let $\varepsilon \in \mathbb{R}^+$ and $y \in B_{\|\|_X}(0, \frac{\rho}{4})$ then for all $y' \in B_{\|\|_X}(0, \frac{\rho}{4})$ with $\|y - y'\| < \frac{\varepsilon}{2}$ we have

$$\|g(y) - g(y')\| \leqslant 2 \cdot \|y - y'\| = 2 \cdot \frac{\varepsilon}{2} = \varepsilon$$

proving that g is continuous at y. As $y \in B_{\|\|_X}(0, \frac{\rho}{4})$ was chosen arbitrary we must conclude that g is continuous. □

Next we prove that g is differentiable on $B_{\|\|_X}(0, \frac{\rho}{4})$.

Conjecture 16.253. *g is differentiable on $B_{\|\|_X}(0, \frac{\rho}{4})$ and $\forall y \in B_{\|\|_X}(0, \frac{\rho}{4})$ we have $D_y g = (D_{g(y)} f)^{-1}$*

Proof. Let $y \in B_{\|\|_X}(0, \frac{\rho}{4})$ then by [eq: 16.368] $g(y) \in U_\rho \subseteq B_{\|\|_X}(0, \frac{\rho}{4}) \subseteq B_{\|\|_X}(0, \rho) \subseteq_{\text{[eq: 16.357]}} U$. Further using [eq: 16.358] we have that $D_{g(y)} f$ is a toplinear isomorphism so that

$$(D_{g(y)} f)^{-1} \text{ exists and } (D_{g(y)} f)^{-1} \in L(X, X) \tag{16.371}$$

Take $h \in \left(B_{\|\|_X}(0, \frac{\rho}{4})\right)_y$ so that $y + h \in B_{\|\|_X}(0, \frac{\rho}{4})$ then for k defined by

$$k = g(y + h) - g(y) \tag{16.372}$$

we have $k + g(y) = g(y + h) \in g\left(B_{\|\|_X}(0, \frac{\rho}{4})\right) \underset{\text{[eq: 16.368]}}{=} U_\rho \subseteq U_\rho \subseteq B_{\|\|_X}(0, \frac{\rho}{4}) \subseteq B_{\|\|_X}(0, \rho) \subseteq_{\text{[eq: 16.357]}} U$ so that

$$k = g(y + h) \in \left(B_{\|\|_X}\left(0, \frac{\rho}{4}\right)\right)_y$$

As f is Fréchet differentiable at $g(y)$ [for $g(y) \in U$] there exist by [theorem: 16.29] a ε-mapping $\zeta: U_{g(y)} \rightarrow X$ such that

$$f(g(y) + k) - f(g(y)) - D_{g(y)} f(k) = \|k\| \cdot \zeta(k) \tag{16.373}$$

As $f(g(y) + k) \underset{\text{[eq: 16.372]}}{=} f(g(y) + g(y + h) - g(y)) = f(g(y + h)) = y + h$ and $f(g(y)) = y$ it follows that $y + h - y - D_{g(y)} f(k) = \|k\| \cdot \zeta(k)$. Hence we have

$$h = D_{g(y)} f(k) + \|k\| \cdot \zeta(k) \tag{16.374}$$

Applying $(D_{g(y)} f)^{-1}$ to both sides of [eq: 16.374] and using linearity we have

$$\begin{aligned} (D_{g(y)} f)^{-1}(h) &= (D_{g(y)} f)^{-1}(D_{g(y)} f(k) + \|k\| \cdot \zeta(k)) \\ &= (D_{g(y)} f)^{-1}(D_{g(y)} f(k)) + \|k\| \cdot (D_{g(y)} f)^{-1}(\zeta(k)) \\ &= k + \|k\| \cdot (D_{g(y)} f)^{-1}(\zeta(k)) \\ &\underset{\text{[eq: 16.372]}}{=} g(y + h) - g(y) + \|g(y + h) - g(y)\| \cdot (D_{g(y)} f)^{-1}(\zeta\,(g(y + h) - g(y))) \end{aligned}$$

so that

$$g(y+h)-g(y)-(D_{g(y)}f)^{-1}(h)=-\|g(y+h)-g(y)\|\cdot(D_{g(y)}f)^{-1}(\zeta(g(y+h)-g(y))) \tag{16.375}$$

The above suggest that we use [theorem: 16.29] again, so define

$$\varepsilon\colon\left(B_{\|\|_X}\left(0,\frac{\rho}{4}\right)\right)_y\rightarrow X \text{ by } \begin{cases} 0 \text{ if } h=0 \\ -\frac{\|g(y+h)-g(y)\|}{\|h\|}\cdot(D_{g(y)}f)^{-1}(\zeta(g(y+h)-g(y))) \text{ if } h\neq 0\end{cases}$$

Using the fact that $g(y+0)-g(0)-(D_{g(y)}f)^{-1}(0)=0$ we have from the above definition and [eq: 16.375] that

$$g(y+h)-g(y)-(D_{g(y)}f)^{-1}(h)=\|h\|\cdot\varepsilon(h) \tag{16.376}$$

As $y, y+h\in B_{\|\|_X}\left(0,\frac{\rho}{4}\right)$ it follows from [eq: 16.370] that

$$\|g(y+h)-g(y)\|\leqslant 2\cdot\|y+h-y\|=2\cdot\|h\| \tag{16.377}$$

Let $\chi\in\mathbb{R}^+$ then by continuity of the ε-mapping ζ [see definition: 16.28] there exist a $\delta'_\chi\in\mathbb{R}^+$ such that for $h\in U_{g(y)}$ with $\|h\|=\|h-0\|<\delta'_\chi$ we have $\|\zeta(h)\|<\chi$. By continuity of g [see conjecture: 16.252] there exist a $\delta_\chi\in\mathbb{R}^+$ such that if $h\in B_{\|\|_X}\left(0,\frac{\rho}{4}\right)$ with $\|h\|<\delta$ we have $\|g(y+h)-g(y)\|<\delta'_\chi$. Further as $(g(y+h)-g(y))+g(y)=g(y+h)\in U_\rho\subseteq B_{\|\|_X}\left(0,\frac{\rho}{4}\right)\subseteq B_{\|\|_X}(0,\rho)\subseteq_{[\text{eq: } 16.357]} U$ it follows that $g(y+h)-g(y)\in U_{g(y)}$. So we have $\|\zeta(g(y+h)-g(y))\|<\zeta$. To summarize

$$\forall\chi\in\mathbb{R}^+ \text{ there exist a } \delta_\chi\in\mathbb{R}^+ \text{ such that } \|\zeta(g(+h)-g(y))\|<\chi \tag{16.378}$$

Take $\mu\in\mathbb{R}^+$ then given $\chi=\frac{\mu}{2\cdot(\|(D_{g(y)}f)^{-1}\|_{L(X,Y)}+1)}\in\mathbb{R}^+$ we have

$$\begin{array}{rcl}
\|\varepsilon(h)-\varepsilon(0)\| & = & \|\varepsilon(h)\| \\
& = & \begin{cases} \|0\| \text{ if } h=0 \\ \left\|-\frac{\|g(y+h)-g(y)\|}{\|h\|}\cdot(D_{g(y)}f)^{-1}(\zeta(g(y+h)-g(y)))\right\| \text{ if } h\neq 0\end{cases} \\
& = & \begin{cases} 0 \text{ if } h=0 \\ \frac{\|g(y+h)-g(y)\|}{\|h\|}\cdot\|(D_{g(y)}f)^{-1}(\zeta(g(y+h)-g(y)))\| \text{ if } h\neq 0\end{cases} \\
& \leqslant & \begin{cases} 0 \text{ if } h=0 \\ \frac{\|g(y+h)-g(y)\|}{\|h\|}\cdot\|(D_{g(y)}f)^{-1}\|_{L(X,X)}\|\zeta(g(y+h)-g(y))\| \text{ if } h\neq 0\end{cases} \\
& \leqslant_{[\text{eq: } 16.377]} & \begin{cases} 0 \text{ if } h=0 \\ \frac{2\cdot\|h\|}{\|h\|}\cdot\|(D_{g(y)}f)^{-1}\|_{L(X,X)}\cdot\|\zeta(g(y+h)-g(y))\| \text{ if } h\neq 0\end{cases} \\
& <_{[\text{eq: } 16.378]} & \begin{cases} 0 \text{ if } h=0 \\ \frac{2\cdot\|h\|}{\|h\|}\cdot\|(D_{g(y)}f)^{-1}\|_{L(X,X)}\cdot\chi \text{ if } h\neq 0\end{cases} \\
& = & \begin{cases} 0 \text{ if } h=0 \\ 2\cdot\|(D_{g(y)}f)^{-1}\|_{L(X,X)}\cdot\chi \text{ if } h\neq 0\end{cases} \\
& = & \begin{cases} 0 \text{ if } h=0 \\ 2\cdot\|(D_{g(y)}f)^{-1}\|_{L(X,X)}\cdot\frac{\mu}{2\cdot(\|(D_{g(y)}f)^{-1}\|_{L(X,Y)}+1)} \text{ if } h\neq 0\end{cases} \\
& < & \begin{cases} 0 \text{ if } h=0 \\ \mu \text{ if } h=0\end{cases} \\
& < & \mu
\end{array}$$

proving that ε is continuous at 0. Hence ε is a ε-mapping, as further by [theorem: 16.358] $(D_{g(y)}f)^{-1}\in L(X,X)$ it follows from [eq: 16.376] and [theorem: 16.29] that g is Fréchet differentiable at y with differential $D_y g=(D_{g(y)}f)^{-1}$. This concludes the proof [conjecture: 16.253] □

To complete the proof of [proposition: 16.249] we need to proof that g is of class C^n. Consider the inversion mapping:

$$\tau\colon \mathrm{GL}(X,X)\rightarrow L(X,X) \text{ defined by } \tau(L)=L^{-1} \text{ [see theorem: 16.239]}$$

As $f:U\rightarrow X$ is of class C^n we can define the function:

$$Df:U\rightarrow L(X,X) \text{ by } Df(x)=D_x f$$

Further by the previous conjecture [conjecture: 16.253] we can define the function

$$Dg: B_{\|\|_X}\left(0,\frac{\rho}{4}\right)\rightarrow L(X,X) \text{ by } Dg(y)=D_y g$$

If $y\in B_{\|\|_X}\left(0,\frac{\rho}{2}\right)$ then we have

$$(\tau\circ Df\circ g)(y)=\tau(Df(g(y)))=(Df(g(y)))^{-1}=(D_{g(y)}f)^{-1}=D_y g=(Dg)(y)$$

proving that

$$\tau\circ Df\circ g=Dg \tag{16.379}$$

Next we use mathematical induction to prove that g is of class C^n. So define

$$S=\{m\in\mathbb{N}|\text{If } m\leqslant n \text{ then } g \text{ is of class } C^m\}$$

then we have:

- $\boldsymbol{1\in S.}$ As τ is of class C^∞ [see theorem: 16.239] hence continuous, Df is of class C^r hence continuous and by [conjecture: 16.252] it follows from [eq: 16.379] that Dg is continuous. Hence g is of class C^1 which proves that $1\in S$.

- $\boldsymbol{m\in S\Rightarrow m+1\in S.}$ If $m+1\leqslant n$ then as f is of class n it follows from [theorem: 16.107] that f is of class $m+1$ so that by [theorem: 16.111] Df is of class C^m. Further as τ is of class C^m and g is of class C^m [because $m\in S$] it follows from the chain rule [theorem: 16.129] that $\tau\circ Df\circ g$ is of class C^m. Hence using [eq: 16.379] we conclude that Dg is of class C^m which by [theorem: 16.111] proves that$g$ is of class C^{m+1} so that $m+1\in S$.

Using induction it follows that $S=\mathbb{N}$ so that, as $n\in\mathbb{N}$ and $n<n$, g is of class C^n. As g is the inverse of $f_{|U_\rho}:U_\rho\rightarrow B_{\|\|_X}\left(0,\frac{\rho}{4}\right)$, $0\in U_\rho$ [see conjecture: 16.251], it follows that f is a local diffeomorphism of class C^n at 0. This concludes the proof of [proposition: 16.249] □

Next we must extend [proposition: 16.249] to a more general case.

Proposition 16.254. *Let $n\in\mathbb{N}$, $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$ be Banach spaces, U a open set in X, $0\in U$ and $f:U\rightarrow Y$ a function of class C^n such that $f(0)=0$ and $D_0 f\in L(X,Y)$ is a bijection then f is a local diffeomorphism of class C^n at 0.*

Proof. As $D_0 f\in L(X,X)$ is a bijection it follows from [theorem: 16.228] that

$$D_0 f \text{ is a toplinear isomorphism} \tag{16.380}$$

so that

$$(D_0 f)^{-1}: Y\rightarrow X$$

is well defined and is linear and continuous, hence by [theorem: 16.116] $(D_0 f)^{-1}$ is of class C^∞. So as f is of class C^n it follows from the chain rule [theorem: 16.129] that

$$h=(D_0 f)^{-1}\circ f:U\rightarrow X$$

is of class C^n. Then

$$\begin{aligned} h(0) &= ((D_0 f)^{-1}\circ f)(0)\\ &= (D_0 f)^{-1}(f(0))\\ &\underset{f(0)=0}{=} (D_0 f)^{-1}(0)\\ &\underset{\text{linearity}}{=} 0 \end{aligned} \tag{16.381}$$

Also

$$\begin{aligned} D_0 h &\underset{\text{[theorem: 16.36]}}{=} D_0((D_0 f)^{-1}\circ f)\\ &= D_{f(0)}(D_0 f)^{-1}\circ D_0 f\\ &\underset{\text{[theorem: 16.116]}}{=} (D_{f(0)} f)^{-1}\circ D_0 f\\ &\underset{f(0)=0}{=} (D_0 f)^{-1}\circ D_0 f\\ &= \mathrm{Id}_X \end{aligned} \tag{16.382}$$

So the conditions for [proposition: 16.249] are satisfied and we have that h is local diffeomorphism of class C^n at 0. Hence there exist open sets U_0, V_0' in X with $0 \in U_0$ and $0 = h(0) \in V_0'$ such that

$$h_{|U_0}: U_0 \rightarrow V_0' \text{ is a diffeomorphism of class } C^n \tag{16.383}$$

As $D_0 f$ is a toplinear isomorphism [see eq: 16.380] we have that

$$V_0 = (D_0 f)(V_0') \underset{[\text{theorem:: } 2.68]}{=} ((D_0 f)^{-1})^{-1}(V_0') \text{ is open in } Y \tag{16.384}$$

Further as $D_0 f \in L(X, Y)$ and is bijective it follows from [example: 16.242] it follows that $D_0 f$ is a diffeomorphism of class C^∞. Applying then [theorem: 16.243] proves that

$$(D_0 f)_{|V_0'}: V_0' \rightarrow D_0 f(V_0') = V_0 \text{ is a diffeomorphism of class } C^n \tag{16.385}$$

Next as $0 \in V_0'$ and $D_0 f$ is linear we have that $0 \in (D_0 f)_{|V_0'}(0) = D_0 f(0) \in D_0 f(V_0') = V_0$ so that

$$0 \in V_0$$

Take

$$g: (D_0 f)_{|V_0'} \circ h_{|U_0}: U_0 \rightarrow V_0$$

then as $h_{|U_0}$, $(D_0 f)_{|V_0'}$ are diffeomorphism of class C^n it follows from [theorem: 16.246] that

$$g \text{ is a diffeomorphism of class } C^n \tag{16.386}$$

Let now $x \in U_0$ then we have

$$\begin{aligned}
g(x) &= ((D_0 f)_{|V_0'} \circ h_{|U_0})(x) \\
&= (D_0 f)_{|V_0'}(h_{|U_0}(x)) \\
&= (D_0 f)(h_{|U_0}(x)) \\
&= (D_0 f)(h(x)) \\
&= (D_0 f)(((D_0 f)^{-1} \circ f)(x)) \\
&= (D_0 f)(D_0 f)^{-1}(f(x))) \\
&= f(x) \\
&= f_{|U_0}(x)
\end{aligned}$$

so that

$$g = f_{|U_0}$$

So that $f_{|U_0}: U_0 \rightarrow V_0$ is a diffeomorphism of class C^n which proves that f is a local diffeomorphism of class C^n at 0. □

Finally we extend the latest proposition to proof the inverse function theorem, to increase readability we repeat here the inverse function theorem as a conjecture, so as this conjecture is proved we have completed the proof of the inverse function theorem.

Conjecture 16.255. *Let $n \in \mathbb{N}$, $\langle X, \|\|_X \rangle$, $\langle Y, \|\|_Y \rangle$ be Banach spaces, U a open set in X, $x_0 \in U$, $f: U \rightarrow Y$ a function of class C^n such that $D_{x_0} f \in L_n(X, Y)$ is a bijection then f is a local diffeomorphism of class C^n at x_0. In other words there exist a open set U_{x_0} in X with $x_0 \in U_{x_0} \subseteq U$ and a open set $V_{f(x_0)}$ in Y with $f(x_0) \in V_{f(x_0)}$ such that*

$$f_{|U_{x_0}}: U_{x_0} \rightarrow V_{f(x_0)} \text{ is a diffeomorphism of class } C^n$$

Proof. Define $\mathcal{T}_1: X \rightarrow X$ by $\mathcal{T}_1(x) = x + x_0$ then by [example: 16.118] $\mathcal{T}_1$ is of class C^∞ with $D_0 \mathcal{T}_1 = \text{Id}_X$. Then

$$\mathcal{T}_1(0) = x \Rightarrow 0 \in W \text{ where } W = (\mathcal{T}_1)^{-1}(U) \tag{16.387}$$

As $\mathcal{T}_1$ is of class C^∞ hence continuous W is a open set, so applying [theorem: 16.97] we have

$$(\mathcal{T}_1)_{|W}: W \rightarrow X \text{ is of class } C^\infty \text{ with } D_0(\mathcal{T}_1)_{|W} = \text{Id}_X \tag{16.388}$$

Further

$$(\mathcal{T}_1)_{|W}(W) \underset{[\text{theorem: } 2.84]}{=} \mathcal{T}_1(W) = \mathcal{T}_1((\mathcal{T}_1)^{-1}(U)) \subseteq_{[\text{theorem: } 2.25]} U \tag{16.389}$$

Further define $\mathcal{T}_2: Y \rightarrow Y$ by $\mathcal{T}_2(x) = x + f(x_0)$ then by [example: 16.118] $\mathcal{T}_2$ is of class C^∞ with $D_0 \mathcal{T}_2 = \text{Id}_Y$. So we can construct the function

$$h = \mathcal{T}_2 \circ f \circ (\mathcal{T}_1)_{|W}: W \rightarrow Y$$

Using the chain rule [theorem: 16.129] it follows that h is of class C^∞ and

$$\begin{aligned}
D_0 h &= D_0(\mathcal{T}_2\circ f\circ(\mathcal{T}_1)_{|W})\\
&= D_0((\mathcal{T}_2\circ f)\circ(\mathcal{T}_1)_{|W})\\
&= D_{(\mathcal{T}_1)_{|W}(0)}(\mathcal{T}_2\circ f)\circ D_0(\mathcal{T}_1)_{|W}\\
&= D_{(\mathcal{T}_1)_{|W}(0)}(\mathcal{T}_2\circ f)\circ \mathrm{Id}_X\\
&= D_{(\mathcal{T}_1)_{|W}(0)}(\mathcal{T}_2\circ f)\\
&= D_{x_0}(\mathcal{T}_2\circ f)\\
&= D_{f(x_0)}\mathcal{T}_2\circ D_{x_0} f\\
&= \mathrm{Id}_Y\circ D_{x_0} f\\
&= D_{x_0} f
\end{aligned}$$

As $D_{x_0} f$ is a bijection it follows from the above that

$$D_0 h \text{ is a bijection}$$

Further we have

$$\begin{aligned}
h(0) &= (\mathcal{T}_2\circ f\circ(\mathcal{T}_1)_{|W})(0)\\
&= \mathcal{T}_2(f((\mathcal{T}_1)_{|W}(0)))\\
&= \mathcal{T}_2(f(x_0))\\
&= \mathcal{T}_2(f(x_0))\\
&= f(x_0)-f(x_0)\\
&= 0
\end{aligned}$$

So the conditions for the previous proposition [proposition: 16.254] are satisfied and we have that h is a local diffeomorphism of class C^n at 0. Hence there exist by [definition: 16.247] open sets U_0, V_0

$$0\in U_0\subseteq W,\ 0=h(0)\in V_0 \text{ and } h_{|U_0}: U_0\rightarrow V_0 \text{ is a diffeomorphism of class } C^n \tag{16.390}$$

Define now

$$\mathcal{T}_3: X\rightarrow X \text{ by } \mathcal{T}_3(x)=x-x_0$$

then we have that for $x\in X$

$$(\mathcal{T}_3\circ\mathcal{T}_1)(x)=\mathcal{T}_3(\mathcal{T}_1(x))=\mathcal{T}_3(x+x_0)=x+x_0-x_0=x=\mathrm{Id}_X(x)$$

$$(\mathcal{T}_1\circ\mathcal{T}_3)(x)=\mathcal{T}_1(\mathcal{T}_3(x))=\mathcal{T}_1(x-x_0)=x-x_0+x_0=x=\mathrm{Id}_X(x)$$

proving that $\mathcal{T}_3\circ\mathcal{T}_1=\mathrm{Id}_X=\mathcal{T}_3\circ\mathcal{T}_1$ so that

$$\mathcal{T}_3,\mathcal{T}_1 \text{ are bijections and } (\mathcal{T}_3)^{-1}=\mathcal{T}_1\wedge(\mathcal{T}_1)^{-1}=\mathcal{T}_3 \tag{16.391}$$

Let $U_{x_0}=(\mathcal{T}_3)^{-1}(U_0)$ then as $\mathcal{T}_3$ is a bijection we have that $\mathcal{T}_3(U_{x_0}) \underset{\text{[theorem: 2.55]}}{=} U_0$ hence

$$(\mathcal{T}_3)_{|U_{x_0}}: U_{x_0}\rightarrow U_0 \text{ is a bijection} \tag{16.392}$$

Further we have that

$$((\mathcal{T}_3)_{|U_{x_0}})^{-1} \underset{\text{[theorem: 2.90]}}{=} ((\mathcal{T}_3)^{-1})_{|\mathcal{T}_3(U_x)} \underset{\text{[eq: 16.391]}}{=} (\mathcal{T}_1)_{|U_0}: U_0\rightarrow U_{x_0} \tag{16.393}$$

We have also

$$U_{x_0}=(\mathcal{T}_3)^{-1}(U_0)\subseteq_{\text{[eq: 16.390]}}(\mathcal{T}_3)^{-1}(W)=((\mathcal{T}_1)^{-1})^{-1}(W) \underset{\text{[theorem: 2.68]}}{=} \mathcal{T}_1(W) \underset{\text{[eq: 16.389]}}{=} U$$

or

$$U_{x_0}\subseteq U \tag{16.394}$$

By [eq: 16.390] $0\in U_0$ so that $\mathcal{T}_3(x_0)=x_0-x_0=0\in U_0$ proving that

$$x_0\in(\mathcal{T}_3)^{-1}(U_0)=U_{x_0} \tag{16.395}$$

Using [example: 16.118] it follows that $\mathcal{T}_3,\mathcal{T}_1$ are of class C^∞. So applying [theorem: 16.97] we have that

$$(\mathcal{T}_3)_{|U_x}: U_{x_0}\rightarrow X,\ ((\mathcal{T}_3)_{|U_{x_0}})^{-1}: U_0\rightarrow X \text{ are of class } C^\infty \tag{16.396}$$

As $\mathcal{T}_3$ is of class C^∞ $\mathcal{T}_3$ is continuous so that

$$U_{x_0} = (\mathcal{T}_3)^{-1}(U_0) \text{ is a open set} \tag{16.397}$$

Hence by definition of a diffeomorphism [see definition: 16.240] it follows from [eqs: 16.392,16.396, 16.397] that

$$(\mathcal{T}_3)_{|U_{x_0}}: U_{x_0} \rightarrow U_0 \text{ is a diffeomorphism of class } C^n \tag{16.398}$$

Next define

$$\mathcal{T}_4: Y \rightarrow Y \text{ by } \mathcal{T}_4(y) = y + f(x_0) \text{ and } \mathcal{T}_5: Y \rightarrow Y \text{ by } \mathcal{T}_5(y) = y - f(x_0)$$

then we have that for $y \in Y$

$$(\mathcal{T}_4 \circ \mathcal{T}_5)(y) = \mathcal{T}_4(\mathcal{T}_5(y)) = \mathcal{T}_4(y - f(x_0)) = y - f(x_0) + f(x_0) = y = \mathrm{Id}_Y(y)$$

$$(\mathcal{T}_5 \circ \mathcal{T}_4)(y) = \mathcal{T}_5(\mathcal{T}_4(y)) = \mathcal{T}_5(y + f(x_0)) = y + f(x_0) - f(x_0) = y = \mathrm{Id}_Y(y)$$

proving that $\mathcal{T}_4 \circ \mathcal{T}_5 = \mathrm{Id}_Y = \mathcal{T}_5 \circ \mathcal{T}_4$ so that

$$\mathcal{T}_4, \mathcal{T}_5 \text{ are bijections and } (\mathcal{T}_4)^{-1} = \mathcal{T}_5 \wedge (\mathcal{T}_5)^{-1} = \mathcal{T}_4$$

Let $V_{f(x_0)} = \mathcal{T}_4(V_0)$ then we have

$$\mathcal{T}_4: V_0 \rightarrow V_{f(x_0)} \text{ is a bijection} \tag{16.399}$$

further we have

$$V_{f(x_0)} = \mathcal{T}_4(V_0) = ((\mathcal{T}_5)^{-1})(V_0) \underset{\text{[theorem: 2.68]}}{=} (\mathcal{T}_5)^{-1}(V_0) \tag{16.400}$$

and

$$((\mathcal{T}_4)_{|V_0})^{-1} \underset{\text{[theorem: 2.90]}}{=} ((\mathcal{T}_4)^{-1})_{|\mathcal{T}_4(V_0)} = ((\mathcal{T}_4)^{-1})_{|V_{f(x_0)}}: V_{f(x_0)} \rightarrow V_0 \tag{16.401}$$

Using [example: 16.118] it follows that $\mathcal{T}_4, \mathcal{T}_5$ are of class C^∞. So applying [theorem: 16.97] we have that

$$(\mathcal{T}_4)_{|V_0}: V_0 \rightarrow Y,\ ((\mathcal{T}_4)_{|V_0})^{-1}: V_{f(x_0)} \rightarrow Y \text{ are of class } C^\infty \tag{16.402}$$

As $\mathcal{T}_5$ is of class C^∞ hence continuous it follows that

$$V_{f(x_0)} \underset{\text{[eq: 16.400]}}{=} (\mathcal{T}_5)^{-1}(V_0) \text{ is a open set} \tag{16.403}$$

Hence by definition of a diffeomorphism [see definition: 16.240] it follows from [eq: 16.399, 16.402, 16.403] that

$$(\mathcal{T}_4)_{|V_0}: V_0 \rightarrow V_{f(x_0)} \text{ is a diffeomorphism of class } C^n \tag{16.404}$$

Define now

$$g = (\mathcal{T}_4)_{|V_0} \circ h_{|U_0} \circ (\mathcal{T}_3)_{|U_{x_0}}: U_{x_0} \rightarrow V_{f(x_0)}$$

then using [eqs 16.398, 16.390, 16.404] and [theorem: 16.246] it follows that

$$g: U_{x_0} \rightarrow V_{f(x_0)} \text{ is a diffeomorphism of class } C^n \tag{16.405}$$

Further we have for $x \in U_{x_0}$ that

$$\begin{aligned}
g(x) &= ((\mathcal{T}_4)_{|V_0} \circ h_{|U_0} \circ (\mathcal{T}_3)_{|U_{x_0}})(x) \\
&= (\mathcal{T}_4)_{|V_0}(h_{|U_0}((\mathcal{T}_3)_{|U_{x_0}}(x))) \\
&= \mathcal{T}_4(h(\mathcal{T}_3(x))) \\
&= \mathcal{T}_4((\mathcal{T}_2 \circ f \circ (\mathcal{T}_1)_{|W})(x - x_0)) \\
&= \mathcal{T}_4((\mathcal{T}_2 \circ f \circ \mathcal{T}_1)(x - x_0)) \\
&= \mathcal{T}_4(\mathcal{T}_2(f(\mathcal{T}_1(x - x_0)))) \\
&= \mathcal{T}_4(\mathcal{T}_2(f(x - x_0 + x_0))) \\
&= \mathcal{T}_4(\mathcal{T}_2(f(x))) \\
&= \mathcal{T}_4(f(x) + f(x_0)) \\
&= f(x) + f(x_0) - f(x_0) \\
&= f(x) \qquad (16.406) \\
&= f_{|U_{x_0}}(x)
\end{aligned}$$

which proves that $g = f_{|U_{x_0}}$. Hence by [eq: 16.405] $f_{|U_{x_0}}: U_{x_0} \rightarrow V_{x_0}$ is a diffeomorphism of class C^n. As $x_0 \in U_{x_0}$ [see eq: 16.395], $f(x_0) \underset{x_0 \in U_{x_0}}{=} g(x_0) \in V_{f(x_0)}$, $U_{x_0} \subseteq U$ it follows from the definition of a local diffeomorphism [see definition: 16.247] that f is a local diffeomorphism of class C^n at x_0

□

Finally as the above conjecture [conjecture: 16.255] is the same as the inverse theorem [theorem: 16.248] we have proved the inverse function theorem. □

Corollary 16.256. *Let $n, m \in \mathbb{N}$, $\langle \mathbb{K}^n, \|\|_n \rangle$ be the normed space equipped with the maximum norm $\|x\|_n = \max(\{|x_i| i \in \{1, \ldots, n\}\})$, U a open set in $\mathbb{K}^n$, $x_0 \in U$ and $f: \mathbb{K}^n \rightarrow \mathbb{K}^n$ a function of class C^m such that $\det\left(\left[\frac{\partial(f_1, \ldots, f_m)}{\partial(1, \ldots, n)}\right]_{x_0}\right) \neq 0$ then f is a local diffeomorphism of class C^n at x_0. In other words there exist a open set U_{x_0} in $\mathbb{K}^n$ with $x_0 \in U_{x_0} \subseteq U$ and a open set $V_{f(x_0)}$ in $\mathbb{K}^n$ with $f(x_0) \in V_{f(x_0)}$ such that*

$$f_{|U_{x_0}}: U_{x_0} \rightarrow V_{f(x_0)} \text{ is a diffeomorphism of class } C^n$$

Proof. Using the canonical basis $E = \{e_1, \ldots, e_n\}$ defined by $(e_i)_j = \delta_{i,j}$ we have by [theorem: 16.63] that

$$\mathcal{M}(D_{x_0} f; E, E) = \left[\frac{\partial(f_1, \ldots, f_n)}{\partial(1, \ldots, n)}\right]_{x_0}$$

As $\det\left(\left[\frac{\partial(f_1, \ldots, f_n)}{\partial(1, \ldots, n)}\right]_{x_0}\right) \neq 0$ it follows that $\det(\mathcal{M}(D_{x_0} f; E, E)) \neq 0$. So by [theorem: 11.340] $D_{x_0} f$ is a bijection. The rest of the proof follows by applying the inverse function theorem [theorem: 16.248]. □

The following consequence of the inverse function is actually the inverse function theorem as know in calculus.

Corollary 16.257. *Let $\langle \mathbb{K}, \|\rangle$ be the normed space of the real (complex) numbers, $U \subseteq \mathbb{K}$ a open set, $x_0 \in U$ and a function $f: U \rightarrow \mathbb{K}$ such that $\forall x \in U$ f has a derivative f'_x at x and the function*

$$f': U \rightarrow \mathbb{K} \text{ defined by } f'(x) = f'_x$$

is continuous. If $f'_{x_0} = f'(x_0) \neq 0$ then there exists open sets $U_{x_0}, V_{f(x_0)} \subseteq \mathbb{K}$ with $x_0 \in U_{x_0} \subseteq U$ and $f(x_0) \in V_{f(x_0)}$ such that:

1. *$f_{|U_{x_0}}: U_{x_0} \rightarrow V_{f(x_0)}$ is a bijection*
2. *$\forall y \in V_{f(x_0)}$ we have that $(f_{|U_{x_0}})^{-1}$ has a derivative $((f_{|U_{x_0}})^{-1})'_y$ at y such that*

$$((f_{|U_{x_0}})^{-1})': V_{f(x_0)} \rightarrow \mathbb{K} \text{ defined by } ((f_{|U_{x_0}})^{-1})'(y) = ((f_{|U_{x_0}})^{-1})'_y$$

   is continuous.
3. *$\forall x \in U_{x_0}$ we have that*

$$((f_{|U_{x_0}})^{-1})'_{f(x)} = \frac{1}{f'_x}$$

Proof. Using [theorem: 16.100] it follows that

$$f \text{ is of class } C^1$$

As $f'_{x_0} \neq 0$ we can define the function $g: \mathbb{K} \rightarrow \mathbb{K}$ by $g(t) = \frac{t}{f'_{x_0}}$ then we have $\forall t \in \mathbb{K}$

$$\begin{array}{rcl}
(D_{x_0} f \circ g)(t) & = & D_{x_0} f(g(t)) \\
& = & D_{x_0} f\left(\frac{t}{f'_{x_0}}\right) \\
& \underset{[\text{theorem: } 16.32]}{=} & \frac{t}{f'_{x_0}} \cdot f'_{x_0} \\
& = & t \\
& = & \mathrm{Id}_{\mathbb{K}}(t) \\
(g \circ D_{x_0} f)(t) & = & g(D_{x_0} f(t)) \\
& \underset{[\text{theorem: } 16.32]}{=} & g(t \cdot f'_{x_0}) \\
& = & \frac{t \cdot f'_{x_0}}{f'_{x_0}} \\
& = & t \\
& = & \mathrm{Id}_{\mathbb{K}}(t)
\end{array}$$

proving that $D_{x_0}f\circ g=\mathrm{Id}_{\mathbb{K}}=g\circ D_{x_0}f$. Hence

$$D_{x_0}f \text{ is a bijection and } (D_{x_0}f)^{-1}=g \tag{16.407}$$

So by the inverse function theorem [theorem: 16.248] there exists open sets U_{x_0}, $V_{f(x_0)}$ with $x_0\in U_{x_0}\subseteq U,\ f(x_0)\in V_{f(x_0)}$ such that

$$f_{|U_{x_0}}: U_{x_0}\rightarrow V_{f(x_0)} \text{ is a diffeomorphism of class } C^1$$

Hence we have that

$$(f_{|U_{x_0}})^{-1}: V_{f(x_0)}\rightarrow U_{x_0} \text{ is of class } C^1$$

so that by [theorem: 16.100] $\forall y\in V_{f(x_0)}$ we have that $(f_{|U_{x_0}})^{-1}$ has a derivative $((f_{|U_{x_0}})^{-1})'_y$ at y and the function

$$((f_{|U_{x_0}})^{-1})': V_{f(x_0)}\rightarrow\mathbb{K} \text{ with } ((f_{|U_{x_0}})^{-1})'(x)=((f_{|U_{x_0}})^{-1})'_x$$

is continuous. If $x\in U_{x_0}$ then $f(x)=f_{|U_{x_0}}(x)\in V_{f(x_0)}$ then as $f_{|U_{x_0}}\circ(f_{|U_{x_0}})^{-1}=\mathrm{Id}_{V_{f(x_0)}}$ we have by the chain rule that

$$\begin{aligned}
\mathrm{Id}_{V_{f(x_0)}} &\underset{[\text{theorem: } 16.22]}{=} D_{f(x)}(\mathrm{Id}_{V_{f(x_0)}})\\
&= D_{f(x)}(f_{|U_{x_0}}\circ(f_{|U_{x_0}})^{-1})\\
&\underset{[\text{theorem: } 16.36]}{=} D_{(f_{|U_{x_0}})^{-1}(f(x))}f_{|U_{x_0}}\circ D_{f(x)}((f_{|U_{x_0}})^{-1})\\
&= D_x f_{|U_{x_0}}\circ D_{f(x)}((f_{|U_{x_0}})^{-1})
\end{aligned}$$

Hence

$$\begin{aligned}
1 &= \mathrm{Id}_{V_{f(x_0)}}(1)\\
&= (D_x f_{|U_{x_0}}\circ D_{f(x)}((f_{|U_{x_0}})^{-1}))(1)\\
&= D_x f_{|U_{x_0}}(D_{f(x)}((f_{|U_{x_0}})^{-1})(1))\\
&\underset{[\text{theorem: } 16.32]}{=} D_x f_{|U_{x_0}}(1\cdot((f_{|U_{x_0}})^{-1})'_{f(x)})\\
&= ((f_{|U_{x_0}})^{-1})'_{f(x)}\cdot D_x f_{|U_{x_0}}(1)\\
&= ((f_{|U_{x_0}})^{-1})'_{f(x)}\cdot(f_{|U_{x_0}})'_x\\
&\underset{[\text{theorem: } 16.33]}{=} ((f_{|U_{x_0}})^{-1})'_{f(x)}\cdot f'_x
\end{aligned}$$

so that

$$((f_{|U_{x_0}})^{-1})'_{f(x)}=\frac{1}{f'_x}$$

□

Next we use the inverse function theorem to prove the implicit function theorem. First we need a little lemma.

Lemma 16.258. *Let X,Y,Z be vector spaces over the field $\mathbb{K}$. $L\in\mathrm{Hom}(X,Z)$ and $H\in\mathrm{Hom}(Y,Z)$ a linear isomorphism then*

$$T: X\cdot Y\rightarrow X\cdot Z \text{ defined by } T(x,y)=(x,L(x)+H(y))$$

is a linear isomorphism with inverse $S=T^{-1}$ defined by

$$S: X\cdot Z\rightarrow X\cdot Y \text{ by } S(x,z)=(x,-H^{-1}(L(x))+H^{-1}(z))$$

Proof. Let $(x_1,y_1),(x_2,y_2)\in X\cdot Y$ and $\alpha\in\mathbb{K}$ then

$$\begin{aligned}
T((x_1,y_1)+\alpha\cdot(x_2,y_2)) &= T(x_1+\alpha\cdot x_2, y_1+\alpha\cdot y_2)\\
&= (x_1+\alpha\cdot x_2, L(x_1+\alpha\cdot x_2)+H(y_1+\alpha\cdot y_2))\\
&= (x_1+\alpha\cdot x_2, L(x_1)+\alpha\cdot L(x_2)+H(y_1)+\alpha\cdot H(y_2))\\
&= (x_1, L(x_1)+H(y_1))+(\alpha\cdot x_2, \alpha\cdot L(x_2)+\alpha\cdot H(y_2))\\
&= (x_1, L(x_1)+H(y_1))+\alpha\cdot(x_2, L(x_2)+H(y_2))\\
&= T(x_1,y_1)+\alpha\cdot T(x_2,y_2)
\end{aligned}$$

proving that

$$T \text{ is a linear mapping}$$

Let $x\in X$, $y\in Y$ and $z\in Z$ then

$$
\begin{array}{rcl}
(T\circ S)(x,z) & = & T(S(x,z))\\
& = & T(x,-H^{-1}(L(x))+H^{-1}(z))\\
& = & (x,L(x)+H(-H^{-1}(L(x))+H^{-1}(z)))\\
& = & (x,L(x)-H(H^{-1}(L(x)))+H(H^{-1}(z)))\\
& = & (x,L(x)-L(x)+z)\\
& = & (x,z)\\
& = & \mathrm{Id}_{X\cdot Z}\\
(S\circ T)(x,y) & = & S(T(x,z))\\
& = & S(x,L(x)+H(z))\\
& = & (x,-H^{-1}(L(x))+H^{-1}(L(x)+H(z)))\\
& \underset{[\text{theorem: } 11.169]}{=} & (x,-H^{-1}(L(x))+H^{-1}(L(x))+H^{-1}(H(z)))\\
& = & (x,z)\\
& = & \mathrm{Id}_{X\cdot Y}
\end{array}
$$

proving that

$$S,T \text{ are bijections and } S=T^{-1}$$

□

Theorem 16.259. (Implicit Function Theorem (1)) *Let $n\in\mathbb{N}$, $\langle X,\|\|_X\rangle$, $\langle Y,\|\|_Y\rangle$ and $\langle Z,\|\|_Z\rangle$ be Banach spaces, U a open set in X, V a open set in Y and $f:U\cdot V\rightarrow Z$ be a function of class C^n. Using [theorem: 16.203] and [definition: 16.197] we have that*

$$D_1f:U\cdot V\rightarrow L(X,Z) \text{ defined by } D_1f(x,y)=D_{(x,y),1}f$$

$$D_2f:U\cdot V\rightarrow L(Y,Z) \text{ defined by } D_2f(x,y)=D_{(x,y),2}f$$

are well defined and continuous. If $\exists(x_0,y_0)\in U\cdot V$ such that $D_2f(x_0,y_0)$ is a bijection then we have:

1. *There exist a open set $W_0\subseteq U\cdot Z$ and a open set $V_0\subseteq V$ with $(x_0,f(x_0,y_0))\in W_0$ such that there exist a map $g:W_0\rightarrow V_0$ of class C^n such that $g(x_0,f(x_0,y_0))=y_0$ and*

$$\forall(x,z)\in W_0 \text{ we have } f(x,g(x,z))=z$$

2. *$\forall(x,z)\in W_0$ we have $D_2f(x,g(x,z))$ is a bijection so that $(D_2f(x,g(x,z)))^{-1}$ exist.*
3. *As g is of class C^n we have by [theorem: 16.203] and [definition: 16.197] that*

$$D_1g:W_0\rightarrow L(X,Y) \text{ defined by } D_1g(x)=D_{1,x}g \text{ is well defined and continuous}$$

$$D_2g:W_0\rightarrow L(Z,Y) \text{ defined by } D_2g(x)=D_{2,x}g \text{ is well defined and continuous}$$

   Then we have for the above functions that $\forall(x,z)\in W_0$ that:

   a. *$D_1g(x,z)=-(D_2f(x,g(x,z)))^{-1}\circ D_1f(x,g(x,z))$*
   b. *$D_2g(x,z)=(D_2f(x,g(x,z)))^{-1}$*

Proof. First, as we have to deal with different products hence different projections, we introduce the following notations for the different projections:

$$
\begin{array}{c}
\pi_1^{X\cdot Y}:X\cdot Y\rightarrow X \text{ defined by } \pi_1^{X\cdot Y}(x,y)=x\\
\pi_2^{X\cdot Y}:X\cdot Y\rightarrow Y \text{ defined by } \pi_2^{X\cdot Y}(x,y)=y\\
\pi_1^{X\cdot Z}:X\cdot Z\rightarrow X \text{ defined by } \pi_1^{X\cdot Z}(x,y)=x\\
\pi_2^{X\cdot Z}:X\cdot Z\rightarrow Z \text{ defined by } \pi_2^{X\cdot Z}(x,y)=y
\end{array}
$$

Note that by [theorems: 16.119, 16.97] we have that:

$$\pi_1^{X\cdot Y},\pi_2^{X\cdot Y},\pi_1^{X\cdot Z},\pi_2^{X\cdot Z} \text{ and their restrictions to open sets are of class } C^\infty \tag{16.408}$$

Define now the function

$$\varphi:U\cdot V\rightarrow X\cdot Z \text{ by } \varphi(x,y)\rightarrow(x,f(x,y)) \tag{16.409}$$

then $\forall(x,y)\in U\cdot V$ we have

$$\begin{array}{rcl}(\pi_1^{X\cdot Y}\circ\varphi)(x,y) & = & \pi_1^{X\cdot Y}(\varphi(x,y))\\ & = & \pi_1^{X\cdot Y}(x,fx,y)\\ & = & x\\ & = & \pi_1^{X\cdot Y}(x,y)\end{array}$$

proving, taking in account [eq: 16.408], that

$$\pi_1^{X\cdot Y}\circ\varphi=(\pi_1^{X\cdot Y})_{|U\cdot V}\text{ and }\pi_1^{X\cdot Y}\circ\varphi\text{ is of class }C^\infty \tag{16.410}$$

Further we have

$$\begin{array}{rcl}(\pi_2^{X\cdot Y}\circ\varphi)(x,y) & = & \pi_2^{X\cdot Y}(\varphi(x,y))\\ & = & \pi_2^{X\cdot Y}(x,f(x,y))\\ & = & f(x,y)\end{array}$$

proving that

$$\pi_2^{X\cdot Y}\circ\varphi=f\text{ so that }\pi_2^{X\cdot Y}\circ\varphi\text{ is of class }C^n \tag{16.411}$$

Using [theorem: 16.126] together with [eq: 16.410] and [eq: 16.411] proves that

$$\varphi\text{ is of class }C^n \tag{16.412}$$

So for $(x,y)\in U\cdot V$ we have that φ is Fréchet differentiable at (x,y) and we have

$$\begin{array}{rcl}D_{(x,y)}\varphi & \underset{\text{[theorem: 16.61]}}{=} & (D_{(x,y)}(\pi_1^{X\cdot Y}\circ\varphi),D_{(x,y)}(\pi_2^{X\cdot Y}\circ\varphi))_*\\ & \underset{\text{[eq: 16.410, 16.411]}}{=} & (D_{(x,y)}(\pi_1^{X\cdot Y})_{|U\cdot V},D_{(x,y)}f)_\star\\ & \underset{\text{[theorem: 16.24]}}{=} & (D_{(x,y)}\pi_1^{X\cdot Y},D_{(x,y)}f)_\star\\ & \underset{\text{[theorem: 16.119]}}{=} & (\pi_1^{X\cdot Y},D_{(x,y)}f)_\star\\ & \underset{\text{[theorem: 16.47]}}{=} & (\pi_1^{X\cdot Y},(D_{(x,y),1}f)\circ\pi_1^{X\cdot Y}+(D_{(x,y),2}f)\circ\pi_2^{X\cdot Y})_\star\\ & = & (\pi_1^{X\cdot Y},D_1f(x,y)\circ\pi_1^{X\cdot Y}+D_2f(x,y)\circ\pi_2^{X\cdot Y})_\star\end{array} \tag{16.413}$$

So if $h=(r,s)\in X\cdot Y$ we have that

$$\begin{array}{rcl}D_{(x,y)}\varphi(r,s) & \underset{\text{[eq: 16.413]}}{=} & (\pi_1^{X\cdot Y},D_1f(x,y)\circ\pi_1^{X\cdot Y}+D_2f(x,y)\circ\pi_2^{X\cdot Y})_\star(r,s)\\ & = & (\pi_1^{X\cdot Y}(r,s),(D_1f(x,y)\circ\pi_1^{X\cdot Y}+D_2f(x,y)\circ\pi_2^{X\cdot Y})(r,s))\\ & = & (r,(D_1f(x,y)\circ\pi_1^{X\cdot Y})(r,s)+(D_2f(x,y)\circ\pi_2^{X\cdot Y})(r,s))\\ & = & (r,D_1f(x,y)(\pi_1^{X\cdot Y}(r,s))+D_2f(x,y)(\pi_2^{X\cdot Y}(r,s)))\\ & = & (r,D_1f(x,y)(r)+D_2f(x,y)(s))\\ & = & (r,_1D_1f(x,y)(r)+D_2f(x,y)(s))\end{array}$$

Hence we have:

$$\forall(x,y)\in U\cdot V,\forall(r,s\in X\cdot Y\text{ we have }D_{(x,y)}\varphi(r,s)=(r,D_1f(x,y)(r)+D_2f(x,y)(s))) \tag{16.414}$$

Take $L=D_1f(x_0,y_0)\in L(X,Z)\subseteq\mathrm{Hom}(X,Y)$ and $H=D_2f(x_0,y_0)\in L(Y,Z)\subseteq\mathrm{Hom}(Y,Z)$ then by [theorem: 16.414] we have

$$D_{(x_0,y_0)}\varphi(r,s)=(r,L(r)+H(s))$$

Further as $H=D_2f(x_0,y_0)$ is supposed to be a bijection it follows from [lemma: 16.258] that

$$D_{(x_0,y_0)}\varphi\text{ is a bijection} \tag{16.415}$$

with $\forall(r,s)\in X\cdot Y$

$$(D_{(x_0,y_0)}\varphi)^{-1}(r,s)\;=\;(r,-H^{-1}(L(r))+H^{-1}(s))$$

so that

$$(D_{(x_0,y_0)}\varphi)^{-1}:X\cdot Z\rightarrow X\cdot Y$$

is defined by

$$(D_{(x_0,y_0)}\varphi)^{-1}(r,s)=(r,-(D_2f(x_0,y_0))^{-1}\circ D_1f(x_0,y_0)(r)+-(D_2f(x_0,y_0))^{-1}(s)) \tag{16.416}$$

Using the inverse function theorem [theorem: 16.248] it follows that there exist a open set W_1 in $X\cdot Y$ with $(x_0,y_0)\in W_1\subseteq U\cdot V$ and a open set W_2 in $X\cdot Z$ with $\varphi(x_0,y_0)\in W_2$ such that

$$\varphi_{|W_1}: W_1\rightarrow W_2 \text{ is a diffeomorphism of class } C^n \tag{16.417}$$

Using the definition of the product topology on $X\cdot Y$ [see: definition: 14.41] and [theorem: 14.47] there exists a open set U_0 in X and a open set in V_0 in Y such that $(x_0,y_0)\in U_0\cdot V_0\subseteq W_1$. Using [theorem: 16.243] we have, if we take $W_0=\varphi(U_0\cdot V_0)$, that

$$\varphi_{|U_0\cdot V_0}\underset{[\text{theorem: } 2.84]}{=}(\varphi_{|W_1})_{|U_0\cdot V}: U_0\cdot V_0\rightarrow W_0 \text{ is a diffeomorphism of class } C^n \tag{16.418}$$

So

$$\forall (x,z)\in W_0 \text{ we have } \quad x=\pi_1^{X\cdot Y}(\varphi_{|U_0\cdot V_0})^{-1}(x,y)\in\pi_1^{X\cdot Y}(U_0\cdot V_0)=U_0 \tag{16.419}$$

$$y=\pi_2^{X\cdot Y}(\varphi_{|U_0\cdot V_0})^{-1}(x,y)\in\pi_2^{X\cdot Y}(U_0,V_0)=V_0 \tag{16.420}$$

$$\forall (x,z)\in W_0 \text{ we have}$$

Further if $(x,y)\in U_0\cdot V_0$ then $(x,y)\in W_1\subseteq U\cdot V$ so that $x\in U$, as further $\varphi(x,y)=(x,f(x,y))\in X\cdot Z$ it follows that $\varphi(x,y)\in U\cdot Z$ so that

$$W_0=\varphi(U_0\cdot V_0)\subseteq U\cdot Z \tag{16.421}$$

Further we have that

$$(x_0,f(x_0,y_0))=\varphi(x_0,y_0)\in W_0 \tag{16.422}$$

If $y\in V_0$ then as $x_0\in U$ we have $(x_0,y)\in U_0\cdot V_0\subseteq W_1\subseteq U\cdot V$ so that $y\in V$ hence we have

$$V_0\subseteq V \tag{16.423}$$

Using [eq: 16.418] $(\varphi_{|U_0\cdot V})^{-1}: W_0\rightarrow U_0\cdot V_0$ exist and is of class C^n. Hence if $(x,z)\in W_0$ then

$$\begin{aligned}
(x,z) &= \varphi_{|U_0\cdot V_0}((\varphi_{|U_0\cdot V_0})^{-1}(x,z))\\
&= \varphi((\varphi_{|U_0\cdot V_0})^{-1}(x,z))\\
&\underset{[\text{eqs: } 16.419, 16.420]}{=} \varphi(\pi_1^{X\cdot Y}((\varphi_{|U_0\cdot V_0})^{-1}(x,z)),\pi_2^{X\cdot Y}((\varphi_{|U_0\cdot V_0})^{-1}(x,z)))\\
&= (\pi_1^{X\cdot Y}((\varphi_{|U_0\cdot V_0})^{-1}(x,z)),f(\pi_1^{X\cdot Y}((\varphi_{|U_0\cdot V_0})^{-1}(x,z)),\pi_2^{X\cdot Y}((\varphi_{|U_0\cdot V_0})^{-1}(x,z))))
\end{aligned}$$

so that

$$x = \pi_1^{X\cdot Y}((\varphi_{|U_0\cdot V_0})^{-1}(x,z)) \tag{16.424}$$

$$z = f(\pi_1^{X\cdot Y}((\varphi_{|U_0\cdot V_0})^{-1}(x,z)),\pi_2^{X\cdot Y}((\varphi_{|U_0\cdot V_0})^{-1}(x,z))) \tag{16.425}$$

Define

$$g: W_0\rightarrow V_0=\pi_2^{X\cdot Y}(U_0\cdot V_0) \text{ by } g=\pi_2^{X\cdot Y}\circ(\varphi_{|U_0\cdot V_0})^{-1} \tag{16.426}$$

Then we have given $(x,z)\in W_0$ that

$$\begin{aligned}
z &\underset{[\text{eq: } 16.425]}{=} f(\pi_1^{X\cdot Y}((\varphi_{|U_0\cdot V_0})^{-1}(x,z)),\pi_2^{X\cdot Y}((\varphi_{|U_0\cdot V_0})^{-1}(x,z)))\\
&= f(\pi_1^{X\cdot Y}((\varphi_{|U_0\cdot V_0})^{-1}(x,z)),g(x,z))\\
&\underset{[\text{eq: } 16.424]}{=} f(x,g(x,z))
\end{aligned}$$

proving that

$$\forall (x,z)\in W_0 \text{ that } f(x,g(x,z))=z \tag{16.427}$$

As $(x_0,y_0)\in U_0\cdot V_0$ $\varphi_{|U_0\cdot V_0}(x_0,y_0)=\varphi(x_0,y_0)=(x_0,f(x_0,y_0))$ it follows that

$$(x_0,y_0)=(\varphi_{|U_0\cdot V_0})^{-1}(x_0,f(x_0,y_0))$$

so that $y_0=\pi_2^{X\cdot Y}((\varphi_{|U_0\cdot V_0})^{-1}(x_0,f(x_0,y_0)))\underset{[\text{eqs: } 16.422, 16.426]}{=} g(x_0,f(x_0))$ proving that

$$g(x_0,f(x_0,y_0))=y_0 \tag{16.428}$$

As $\pi_2^{X\cdot Y}$ is of class C^∞ [see eq: 16.408] and by [eq: 16.418] $(\varphi_{|U_0\cdot V_0})^{-1}$ is of class C^n it follows from the chain rule [theorem: 16.129] that

$$g \text{ is of class } C^n \tag{16.429}$$

As $\varphi_{|U_0\cdot V_0}: U_0\cdot V_0 \rightarrow W_0$ is a diffeomorphism of class C^n [see eq: 16.418] it follows from [theorem: 16.245] that

$$\forall (x,y)\in U_0\cdot V_0 \text{ we have that } D_{(x,y)}\varphi \underset{\text{[theorem: 16.24]}}{=} D_{(x,y)}\varphi_{|U_0\cdot V_0} \text{ is bijective} \tag{16.430}$$

So part (1) of the theorem is proved by [eqs: 16.426, 16.427, 16.429].

Now for the next part. Take $r,s\in Y$ such that $D_2 f(x,y)(r) = D_2 f(x,y)(s)$ then we have

$$\begin{array}{rcl} D\varphi(x,y)(0,r) & \underset{\text{[eq: 16.414]}}{=} & (0, D_1 f(x,y)(0) + D_2 f(x,y)(r)) \\ & = & (0, D_1 f(x,y)(0) + D_2 f(x,y)(s)) \\ & \underset{\text{[eq: 16.414]}}{=} & D\varphi(x,y)(0,s) \end{array}$$

so that by [eq: 16.430] $(0,r) = (0,s)$ $r = s$ proving that

$$D_2 f(x,y) \text{ is injective} \tag{16.431}$$

If $z\in Z$ then as $D_{(x,y)}\varphi$ is a bijection [see eq: 16.430] there exist a $(r,s)\in X\cdot Z$ such that $D\varphi(r,s) = (0,z)$, hence [by [eq: 16.414] we have $(r, D_1 f(x,y)(r) + D_2 f(x,y)(s)) = (0,z)$ so that $r = 0 \wedge D_1 f(x,y)(r) + D_2 f(x,y)(s) = z$. So we have $z = D_{21} f(x,y)(0) + D_2 f(x,y)(s) = D_2 f(x,y)(s)$ which proves that $D_2 f(x,y)$ is surjective. Combining this with [eq: 16.431] proves that

$$\forall (x,y)\in U_0\cdot V_0 \text{ we have that } D_2 f(x,y) \text{ is bijective}$$

Take now $(x,z)\in W_0$ then by [eq: 16.419] $x\in U_0$ and by [eq: 16.426] $g(z,z)\in V_0$ so that by the above we have

$$\forall (x,z)\in W_0 \text{ we have } D_2 f(x, g(x,z)) \text{ is bijective so that } (D_2 f(x, g(x,z)))^{-1} \text{ exist} \tag{16.432}$$

Which proves part (2) of the theorem.

For the next part. Define

$$\psi: W_0 \rightarrow X\cdot Y \text{ by } \psi(x,z) = (x, g(x,z))$$

Then we have that for $(x,z)\in W_0$ that

$$\begin{array}{rcl} (\pi_1^{X\cdot Y}\circ\psi)(x,z) & = & \pi_1^{X\cdot Y}(\psi(x,z)) = \pi_1^{X\cdot Y}(x, g(x,z)) = x = \pi_1^{X\cdot Y}(x,z) \\ (\pi_2^{X\cdot Y}\circ\psi)(x,z) & = & \pi_2^{X\cdot Y}(\psi(x,z)) = \pi_2^{X\cdot Y}(x, g(x,z)) = g(x,z) \end{array}$$

proving that

$$\pi_1^{X\cdot Y}\circ\psi = \pi_1^{X\cdot Y} \text{ and } \pi_2^{X\cdot Y}\circ\psi = g$$

By [example: 16.119] $\pi_1^{X\cdot Y}$ is of class C^∞ and by [eq: 16.429] g is of class C^n, so using the above and [theorem: 16.126] it follows that

$$\psi \text{ is of class } C^n \text{ hence Fréchet differentiable on } W_0 \tag{16.433}$$

Define

$$f\circ\psi: W_0 \rightarrow Z$$

we have for $(x,z)\in W_0$ that

$$\begin{array}{rcl} (f\circ\psi)(x,z) & = & f(\psi(x,z)) \\ & = & f(x, g(x,z)) \\ & \underset{\text{[eq: 16.427]}}{=} & z \\ & = & \pi_2^{X\cdot Z}(x,z) \end{array}$$

proving that

$$f\circ\psi = \pi_2^{X\cdot Z} \tag{16.434}$$

Let $(x,z)\in W_0$. Then for $t\in X$ we have

$$((f\circ\psi)\circ I^{[1,(x,z)]})(t) = (f\circ\psi)(I^{[1,(x,z)]}(t)) \underset{\text{[definition: 16.39]}}{=} (f\circ\psi)(t,z) \underset{\text{[eq: 16.434]}}{=} z = C_z(t)$$

proving that

$$(f\circ\psi)\circ I^{[1,(x,z)]} = C_z$$

so that by [example: 16.26] $(f\circ\psi)\circ I^{[1,(x,z)]}$ is Fréchet differentiable at t which by [definition: 16.43] proves that $f\circ\psi$ has a 1-partial differential at t with

$$D_{(x,z),1}(f\circ\psi)\underset{\text{[definition: 16.43]}}{=}D_x((f\circ\psi)\circ I^{[1,(x,z)]})=D_xC_z\underset{\text{[example: 16.26]}}{=}0\in L(X,Z) \tag{16.435}$$

Also

$$\begin{aligned}((\pi_1^{X\cdot Z}\circ\psi)\circ I^{[1,(x,z)]})(t) &= \pi_1^{X\cdot Z}(\psi(I^{[1,(x,z)]}(t)))\\ &= \pi_1^{X\cdot Z}(\psi(t,z))\\ &= \pi_1^{X\cdot Z}(t,g(t,z))\\ &= t\\ &= \mathrm{Id}_X(t)\\ ((\pi_2^{X\cdot Z}\circ\psi)\circ I^{[1,(x,z)]})(t) &= \pi_2^{X\cdot Z}(\psi(I^{[1,(x,z)]}(t)))\\ &= \pi_2^{X\cdot Z}(\psi(t,z))\\ &= \pi_2^{X\cdot Z}(t,g(t,z))\\ &= g(t,z)\\ &= g(I^{[1,(x,z)]}(t))\\ &= (g\circ I^{[1,(x,z)]})(t)\end{aligned}$$

so that

$$(\pi_1^{X\cdot Z}\circ\psi)\circ I^{[1,(x,z)]}=I_X \text{ and } (\pi_2^{X\cdot Z}\circ\psi)\circ I^{[1,(x,z)]}=g\circ I^{[1,(x,z)]} \tag{16.436}$$

As g is of class C^n [see eq: 16.429] g is Fréchet differentiable on W_0 hence by [theorem: 16.203] and [definition: 16.197] that

$$D_1g: W_0\rightarrow L(X,Y) \text{ defined by } D_1g(x)=D_{1,x}g \text{ is well defined and continuous}$$

$$D_2g: W_0\rightarrow L(Z,Y) \text{ defined by } D_2g(x)=D_{2,x}g \text{ is well defined and continuous}$$

Given $h\in X$ we have

$$\begin{aligned} D_{(x,z),1}(f\circ\psi)(h) &\underset{\substack{\text{[eq: 16.433]}\\ \text{[theorem: 16.52]}}}{=} \\ (D_{\psi(x,z)}f\circ D_{(x,z),1}\psi)(h) &= \\ (D_{\psi(x,z)}f\circ(D_{(x,z)}(\psi\circ I^{[1,(x,z)]})))(h) &\underset{\text{[theorem: 16.61]}}{=} \\ (D_{\psi(x,z)}f\circ(D_{(x,z)}(\pi_1^{X\cdot Z}\circ(\psi\circ I^{[1,(x,z)]})),D_{(x,z)}(\pi_2^{X\cdot Z}\circ(\psi\circ I^{[1,(x,z)]})))_\star)(h) &\underset{\text{[eq: 16.436]}}{=} \\ (D_{\psi(x,z)}f\circ(D_{(x,z)}\mathrm{Id}_X,D_{(x,z)}(g\circ I^{[1,(x,z)]}))_*)(h) &\underset{\text{[example: 16.22]}}{=} \\ (D_{\psi(x,z)}f\circ(\mathrm{Id}_X,D_{(x,z)}(g\circ I^{[1,(x,z)]}))_*)(h) &= \\ (D_{\psi(x,z)}f\circ(\mathrm{Id}_X,D_{(x,z),1}g)_*)(h) &= \\ D_{\psi(x,z)}f((\mathrm{Id}_X,D_{(x,z),1}g)_*(h)) &= \\ D_{\psi(x,z)}f(h,D_{(x,z),1}g(h)) &\underset{\text{[theorem: 16.47]}}{=} \\ D_{\psi(x,z),1}f(h)+D_{\psi(x,z),2}f(D_{(x,z),1}g(h)) &= \\ D_{\psi(x,z),1}f(h)+(D_{\psi(x,z),2}f\circ D_{(x,z),1}g)(h) &= \\ (D_{\psi(x,z),1}f+D_{\psi(x,z).2}f\circ D_{(x,z),1}g)(h) &= \\ (D_1f(\psi(x,z))+D_2f(\psi(x,z))\circ D_1g(x,z))(h) &= \\ (D_1f(x,g(x,z))+D_2f(x,g(x,z))\circ D_1g(x,z))(h) & \end{aligned}$$

proving that

$$D_1f(x,g(x,z))+D_2f(x,g(x,z))\circ D_1g(x,z)=D_{(x,z),1}(f\circ\psi)\underset{\text{[eq: 16.435]}}{=}0$$

so that

$$D_2f(x,g(x,z))\circ D_1g(x,z)=-D_1f(x,g(x,z))$$

as $(D_2f(x,g(x,z))^{-1}$ exist [see eq: 16.432] and is linear by [theorem: 11.169] we have

$$\forall(x,z)\in W_0\ D_1g(x,z)=-(D_2f(x,g(x,z)))^{-1}\circ D_1f(x,g(x,z)) \tag{16.437}$$

which proves part (3.a) of the theorem.

For the remaining part (3.b) of the theorem note that for $(x,z)\in W_0$ we have for $t\in Z$ that

$$(\pi_2^{X\cdot Z}\circ I^{[2,(x,z)]})(t)=\pi_2^{X\cdot Z}(I^{[2,(x,z)]}(t))=\pi_2^{X\cdot Z}(x,t)=t=\mathrm{Id}_Z(t)$$

so that

$$\pi_2^{X\cdot Z_2}\circ I^{[2,(z,x)]}=\mathrm{Id}_Z \tag{16.438}$$

hence we have

$$\begin{aligned} D_{(x,z),2}(f\circ\psi) &\underset{[\text{eq: }16.434]}{=} D_{(x,z),2}(\pi_2^{X\cdot Z}) \\ &= D_{(x,z)}(\pi_2^{X\cdot Z}\circ I^{[2,(x,z)]}) \\ &\underset{[\text{eq: }16.438]}{=} D_{(x,z)}\mathrm{Id}_Z \\ &\underset{[\text{eq: }16.22]}{=} \mathrm{Id}_Z \end{aligned} \tag{16.439}$$

Further we have

$$\begin{aligned} (\pi_1^{X\cdot Z}\circ\psi\circ I^{[2,(x,z)]})(t) &= \pi_1^{X\cdot Z}(\psi(I^{[2,(x,z)]}(t))) \\ &= \pi_1^{X\cdot Z}(\psi(x,t)) \\ &= \pi_1^{X\cdot Z}(x,g(x,t)) \\ &= x \\ &= C_x(t) \\ (\pi_2^{X\cdot Z}\circ\psi\circ I^{[2,(x,z)]})(t) &= \pi_2^{X\cdot Z}(\psi(I^{[2,(x,z)]}(t))) \\ &= \pi_2^{X\cdot Z}(\psi(x,t)) \\ &= \pi_2^{X\cdot Z}(x,g(x,t)) \\ &= g(x,t) \\ &= g(I^{[2,(x,z)]}(t)) \\ &= (g\circ I^{[2,(x,z)]})(t) \end{aligned}$$

proving that

$$\pi_1^{X\cdot Z}\circ\psi\circ I^{[2,(x,z)]}=C_x\wedge\pi_2^{X\cdot Z}\circ\psi\circ I^{[2,(x,z)]}=g\circ I^{[2,(x,z)]} \tag{16.440}$$

We use the above to calculate $D_{(x,z),2}(f\circ\psi)$ in another way, let $h\in Z$ then

$$\begin{aligned} D_{(x,z),2}(f\circ\psi)(h) &\underset{\substack{[\text{eq: }16.433]\\ [\text{theorem: }16.52]}}{=} \\ (D_{\psi(x,z)}f\circ D_{(x,z),2}\psi)(h) &= \\ (D_{\psi(x,z)}f\circ D_{(x,z)}(\psi\circ I^{[2,(x,z)]}))(h) &\underset{[\text{theorem: }16.61]}{=} \\ (D_{\psi(x,z)}f\circ(D_{(x,z)}(\pi_1^{X\cdot Z}\circ(\psi\circ I^{[2,(x,z)]})),D_{(x,z)}(\pi_2^{X\cdot Z}(\psi\circ I^{[2,(x,z)]})))_*)(h) &\underset{[\text{eq: }16.440]}{=} \\ (D_{\psi(x,z)}f\circ(D_{(x,z)}C_x,D_{(x,z)}(g\circ I^{[2,(x,z)]}))_*)(h) &\underset{[\text{eq: }16.26]}{=} \\ (D_{\psi(x,z)}f\circ(0,D_{(x,z)}(g\circ I^{[2,(x,z)]}))_*)(h) &= \\ (D_{\psi(x,z)}f\circ(0,D_{(x,z),2}g)_*)(h) &= \\ D_{\psi(x,z)}f((0,D_{(x,z),2}g)_*(h)) &= \\ D_{\psi(x,z)}f(0(h),D_{(x,z),2}g(h)) &= \\ D_{\psi(x,z)}f(0,D_{(x,z),2}g(h)) &\underset{[\text{theorem: }16.47]}{=} \\ D_{\psi(x,z),1}f(0)+D_{\psi(x,z),2}f(D_{(x,z),2}g(h)) &= \\ D_{\psi(x,z),2}f(D_{(x,z),2}g(h)) &= \\ D_2f(\psi(x,z)(D_2g(x,z)(h))) &= \\ D_2f(x,g(x,z))(D_2g(x,z)(h)) &= \\ (D_2f(x,g(x,z))\circ D_2g(x,z))(h) & \end{aligned}$$

proving that

$$D_2 f(x, g(x, z)) \circ D_2 g(x, z) = D_{(x,z),2}(f \circ \psi) \underset{\text{[eq: 16.439]}}{=} \mathrm{Id}_Z$$

Hence we have, as $(D_2 f(x, g(x, z))^{-1}$ exist [see eq: 16.432]], that

$$\forall (x, z) \in W_0 \text{ we have } D_2 g(x, z) = (D_2 f(x, g(x, z))^{-1} \tag{16.441}$$

which proves part (3.b) of the theorem. □

A simpler version of the implicit function theorem is the following.

Theorem 16.260. *Let $n \in \mathbb{N}$, $\langle X, \|\|_X \rangle$, $\langle Y, \|\|_Y \rangle$ and $\langle Z, \|\|_Z \rangle$ be Banach spaces, U a open set in X, V a open set in Y and $f: U \cdot V \to Z$ of class C^n. Using [theorem: 16.203] and [definition: 16.197] we have that*

$$D_1 f: U \cdot V \to L(X, Z) \text{ defined by } D_1 f(x, y) = D_{(x,y),1} f$$

$$D_2 f: U \cdot V \to L(Y, Z) \text{ defined by } D_2 f(x, y) = D_{(x,y),2} f$$

are well defined and continuous. If $\exists (x_0, y_0) \in U \cdot V$ such that $D_2 f(x_0, y_0)$ is bijective then there exist a open set W with $x_0 \in W \subseteq U$ and a function $g: W \to Y$ of class C^n such that

$$g(x_0) = y_0 \text{ and } \forall x \in W \text{ we have } f(x, g(x)) = f(x_0, y_0)$$

$$\forall x \in W \ (D_2 f(x, g(x)))^{-1} \text{ exist and } Dg(x) = -(D_2 f(x, g(x)))^{-1} \circ D_1 f(x, g(x))$$

***Note* 16.261.** *As f is of class C^n we can use [theorem: 16.203] and [definition: 16.197] to define*

$$D_1 f: U \cdot V \to L(X, Z) \text{ by } D_1 f(x, y) = D_{(x,y),1} f$$

$$D_2 f: U \cdot V \to L(Y, Z) \text{ by } D_2 f(x, y) = D_{(x,y),2} f$$

and as g is of class C^n we can define

$$Dg: W \to L(X, Y) \text{ by } Dg(x) = D_x g$$

Proof. Using the implicit function theorem [theorem: 16.259] there exist a open set $W_0 \subseteq U \cdot Z$ and a open set $V_0 \subseteq V$ with $(x_0, y_0) \in W_0$ such that there exist a map $h: W_0 \to V_0$ of class C^n such that

$$(x_0, f(x_0, y_0)) \in W_0 \tag{16.442}$$

$$h(x_0, f(x_0, y_0)) = y_0 \tag{16.443}$$

and $\forall (x, z) \in W_0$ we have

$$f(x, h(x, z)) = z \tag{16.444}$$

$$D_1 h(x, z) = -(D_2 f(x, h(x, z)))^{-1} \circ D_1 f(x, h(x, z)) \tag{16.445}$$

$$D_2 h(x, z) = (D_2 f(x, h(x, z)))^{-1} \tag{16.446}$$

where

$$D_1 h: W_0 \to L(X, Y) \text{ is defined by } D_1 h(x, z) = D_{(x,z),1} h$$

$$D_2 h: W_0 \to L(Z, Y) \text{ is defined by } D_2 h(x, z) = D_{(x,z),2} h$$

Using the definition of the product topology on $X \cdot Y$ [see: definition: 14.41] and [theorem: 14.47] there exists a open set W in X and a open set O in Z such that

$$(x_0, f(x_0, y_0)) \in W \cdot O \subseteq W_0 \subseteq U \cdot Z$$

From the above it follows that $f(x_0, w_0) \in O$ so if $x \in W$ then by the above $(x, f(x_0, y_0)) \in W \cdot O$ which is the domain of h so the following function is well defined

$$g: W \to V \text{ where } g(x) = h(x, f(x_0, y_0)) \tag{16.447}$$

then we have

$$g(x_0) = h(x_0, f(x_0, y_0)) \underset{\text{[eq: 16.443]}}{=} y_0 \tag{16.448}$$

Further for $x \in W$ we have $(x, f(x_0, y_0)) \in W \cdot O \subseteq W_0$ so that

$$f(x, g(x)) = f(x, h(x, f(x_0, y_0))) \underset{\text{[eq: 16.444]}}{=} f(x_0, y_0)$$

hence

$$\forall x \in W \text{ we have } f(x, g(x)) = f(x_0, y_0) \tag{16.449}$$

Further we have for $t \in W$ that

$$\begin{aligned} (h \circ I^{[1,(x, f(x_0, y_0))]})(t) &= h(I^{[1,(x, f(x_0, y_0))]}(t)) \\ &= h(t, f(x_0, y_0)) \\ &\underset{\text{[eq: 16.447]}}{=} g(t) \end{aligned}$$

proving that

$$g = h \circ I^{[1,(x, f(x_0, y_0))]}$$

Using the chain rule [see theorem: 16.129] and the fact that h is of class C^n and $I^{[1,(x, f(x_0, y_0))]}$ is of class C^∞ [see example: 16.117] it follows that

$$g \text{ is of class } C^r \tag{16.450}$$

so that

$$Dg: W \to L(X, Y) \text{ defined by } Dg(x) = D_x g$$

is well defined. Further we have for $x \in W$ that

$$\begin{aligned} Dg(x) &= D_x g \\ &= D_x(h \circ I^{[1,(x, f(x_0, y_0))]}) \\ &\underset{\text{[definition: 16.43]}}{=} D_{(x, f(x_0, y_0)),1} h \\ &= D_1 h(x, f(x_0, y_0)) \\ &\underset{\text{[eq: 16.445]}}{=} -(D_2 f(x, h(x, f(x_0, y_0))))^{-1} \circ D_1 f(x, h(x, f(x_0, y_0))) \\ &\underset{\text{[theorem: 16.447]}}{=} -(D_2 f(x, g(x)))^{-1} \circ D_1 f(x, g(x)) \end{aligned} \tag{16.451}$$

The theorem is then proved by [eqs: 16.450, 16.448, 16.449 and 16.451]. □

Chapter 17

Fundamental theorem of algebra

The purpose of this chapter is to prove the fundamental theorem of algebra. This states that every higher order equation of the form

$$a_n \cdot x^n + a_{n-1} \cdot x^{n-1} + \cdots + a_0 = 0$$

This is mostly done using complex analysis. However in this case we will use algebraic operations to prove the theorem. First to define the concept of a higher order equation we must introduce the concept of polynomals.

17.1 Prerequisites

17.1.1 Polynomials

Definition 17.1. *If $n \in \mathbb{N}_0$ then a function $p: \mathbb{C} \rightarrow \mathbb{C}$ is a* ***polynomial of order n*** *if there exist a family $\{a_i\}_{i \in \{0, \ldots, n\}} \subseteq \mathbb{C}$ called the* ***coefficients of p*** *with $a_n \neq 0$ so that p is defined by*

$$p: \mathbb{C} \rightarrow \mathbb{C} \text{ where } p(z) = \sum_{i=0}^{n} a_i \cdot z^i$$

The collection of complex polynomials of order n is noted as $\mathcal{P}[n]$ so that

$$\mathcal{P}[n] = \{p \in \mathbb{C}^{\mathbb{C}} | p \text{ is a polynomial of order } n\}$$

The collection of all polynomials is noted as $\mathcal{P}$ so that

$$\mathcal{P} = \bigcup_{n \in \mathbb{N}_0} \mathcal{P}[n]$$

Note that in principle a polynomial can have different coefficients and be of different orders. We will see however that this is not the case and that every polynomial can be represented by a unique set of coefficients.

Lemma 17.2. *Let $\{a_i\}_{i \in \{0, \ldots, n\}} \subseteq \mathbb{C}$ such that $\forall x \in \mathbb{R} \subseteq \mathbb{C}$ we have $\sum_{i=0}^{n} a_i \cdot x^i = 0$ then $\forall i \in \{0, \ldots, n\}$ we have $a_i = 0$*

Proof. Assume that $\exists k \in \{0, \ldots, n\}$ such that $a_i \neq 0$. Define $\mathcal{N} = \{i \in \{0, \ldots, n\} | a_i \neq 0\}$ then $k \in \mathcal{N} \neq 0$ hence $m = \max(\mathcal{N})$ exist. So $\forall i \in \{m+1, \ldots, n\}$ we have $a_i = 0$. So that

$$\forall x \in \mathbb{R} \text{ we we have } 0 = p(x) = \sum_{i=0}^{n} a_i \cdot x^i = \sum_{i=0}^{m} a_i \cdot x^i + \sum_{i=m+1}^{n} a_i \cdot x^i = \sum_{i=0}^{m} a_i \cdot x^i$$

$\boldsymbol{m = 0.}$ Then we have $0 = p(1) = \sum_{i=0}^{0} a_0 \cdot 1^0 = a_0 \neq 0$ a contradiction.

$0<m$. Then as $a_m\neq 0$ we can take $x=\frac{\sum_{i=0}^{m-1}|a_i|}{|a_m|}+1$ then $1\leqslant x$ and we have by [theorem: 10.61] that $\forall j\in\{0,\ldots,m-1\}$ we have $x^j\leqslant x^{m-1}$. Hence

$$\begin{aligned}\left|\sum_{i=0}^{m-1} a_i\cdot x^i\right| &\leqslant \sum_{i=0}^{m-1} |a_i\cdot x^i|\\ &\underset{\text{[theorem: 10.83]}}{=} \sum_{i=0}^{m-1} |a_i|\cdot|x|^i\\ &\leqslant \sum_{i=0}^{m-1}|a_i|\cdot|x|^{m-1}\\ &= |x|^{m-1}\cdot\sum_{i=0}^{m-1}|a_i|\\ &= |x|^{m-1}\cdot|a_m|\cdot(x-1)\\ &< |x|^m\cdot|a_m|\\ &\underset{\text{[theorem: 10.83]}}{=} |-a_m\cdot x^m|\end{aligned}$$

Hence $\sum_{i=0}^{m-1}a_i\cdot x^i\neq -a_m x^m$ so that $\sum_{i=0}^m a_i\cdot x^i=\sum_{i=0}^{m-1}a_i\cdot x^i+a_m\cdot x^m\neq 0$ contradicting $0=p(x)=\sum_{i=0}^m a_i\cdot x^i$.

As in all possible cases we have a contradiction the assumption is wrong and thus $\forall i\in\{0,\ldots,n\}$ we have $a_i=0$. □

Note 17.3. The function $C_0:\mathbb{C}\rightarrow\mathbb{C}$ defined by $C_0(z)=0$ is not a polynomial.

Proof. Assume that C_0 is a polynomial then there exist a $\{a_i\}_{i\in\{0,\ldots,n\}}\subseteq\mathbb{C}$ with $a_n\neq 0$ such that $\forall z\in\mathbb{C}$ $\sum_{i=0}^n a_i\cdot z^i=C_0=0$. Then by [lemma: 17.2] we must have $a_n=0$ contradicting $a_n\neq 0$. □

Corollary 17.4. *Let $n\in\mathbb{N}_0$ and $\{a_i\}_{i\in\{1,\ldots,n\}}\subseteq\mathbb{C}$, $\{b_i\}_{i\in\{1,\ldots,n\}}\subseteq\mathbb{C}$ be finite families such that $\forall z\in\mathbb{C}$ we have*

$$\sum_{i=0}^n a_i\cdot z^i=\sum_{i=0}^n b_i\cdot z^i$$

then we have $\forall i\in\{0,\ldots,n\}$ $a_i=b_i$.

Proof. As $\forall z\in\mathbb{C}$ we have that $\sum_{i=0}^n a_i\cdot z^i=\sum_{i=0}^n b_i\cdot z^i$ it follows that

$$0=\sum_{i=0}^n a_i\cdot z^i-\sum_{i=0}^n b_i\cdot z^i=\sum_{i=0}^n(\alpha_i\cdot z^i-b_i\cdot z^i)=\sum_{i=0}^n(a_i-b_i)\cdot z^i$$

hence by [lemma: 17.2] it follows that $\forall\in\{0,\ldots,n\}$ $a_i=b_i$. □

Lemma 17.5. *Let $n,m\in\mathbb{N}_0$, $\{a_i\}_{i\in\{0,\ldots,n\}}\subseteq\mathbb{C}$ a family such that $a_n\neq 0$ if $n\neq 0$ and $\{b_i\}_{i\in\{0,\ldots,m\}}\subseteq\mathbb{C}$ a family such that $b_m\neq 0$ if $m\neq 0$. Then if $\forall z\in\mathbb{C}$ $\sum_{i=1}^n a_i\cdot z^i=\sum_{i=1}^m b_i\cdot z^i$ it follows that $n=m$ and $\forall i\in\{0,\ldots,n\}$ we have $a_i=b_i$ [in other words $\{a_i\}_{i\in\{0,\ldots,n\}}=\{b_i\}_{i\in\{0,\ldots,m\}}$.*

Proof. For $n,m\in\mathbb{N}_0$ we have the following different cases:

$n=m$. Then by [corollary: 17.4] we have $\forall i\in\{0,\ldots,n\}$ $a_i=b_i$ and trivially $n=m$.

$n<m$. Define then $\{a_i'\}_{i\in\{0,\ldots,m\}}\subseteq\mathbb{C}$ by $a_i'=\begin{cases}0 \text{ if } n<i\\ a_i \text{ if } i\in\{0,\ldots,n\}\end{cases}$ then we have $\forall z\in\mathbb{C}$

$$\begin{aligned}\sum_{i=0}^m a_i'\cdot z^i &= \sum_{i=0}^n a_i'\cdot z^i+\sum_{i=n+1}^m a_i'\cdot z^i\\ &= \sum_{i=0}^n a_i\cdot z^i+\sum_{i=n+1}^m 0\cdot z^i\\ &= \sum_{i=0}^n a_i\cdot z^i\\ &= \sum_{i=1}^m b_i\cdot z^i\end{aligned}$$

so that by [corollary: 17.4] we have that $0 = a'_m = b_m$. As $0 \leqslant n < m \Rightarrow m \neq 0$ we have $b_m \neq 0$ and we reach a contradiction. Hence this case will never occur.

$\boldsymbol{m < n}$**.** Define then $\{b'_i\}_{i \in \{0,\ldots,n\}} \subseteq \mathbb{C}$ by $b'_i = \left\{ \begin{smallmatrix} 0 \text{ if } m < i \\ b_i \text{ if } i \in \{0,\ldots,m\} \end{smallmatrix} \right.$ then we have $\forall z \in \mathbb{C}$

$$\begin{aligned} \sum_{i=0}^{n} b'_i \cdot z^i &= \sum_{i=0}^{m} b'_i \cdot z^i + \sum_{i=m+1}^{n} b'_i \cdot z^i \\ &= \sum_{i=0}^{m} b_i \cdot z^i + \sum_{i=m+1}^{n} 0 \cdot z^i \\ &= \sum_{i=0}^{m} b_i \cdot z^i \\ &= \sum_{i=1}^{a} a_i \cdot z^i \end{aligned}$$

so that by [corollary: 17.4] we have that $0 = b'_m = a_m$. As $0 \leqslant m < n \Rightarrow n \neq 0$ we have $a_n \neq 0$ and we reach a contradiction. Hence this case will never occur.

So the only valid cases is $n = m$ proving the lemma. □

The above ensures that the following definition is valid.

Definition 17.6. *Let* $\mathcal{C} = \{\{a_i\}_{i \in \{0,\ldots,n\}} \subseteq \mathbb{C} | a_n \neq 0 \text{ if } n \neq 0\}$ *then we define*

$$\mathrm{ord}_{\mathrm{coef}} \colon \mathcal{C} \to \mathbb{N}_0 \text{ by } \mathrm{ord}_{\mathrm{coef}}\left(\sum_{i=0}^{n} a_i \cdot z^i\right) = n$$

$$\mathrm{coef} \colon \mathcal{P} \to \mathcal{C} \text{ where } \mathrm{coef}(p) = \{a_i\}_{i \in \{1,\ldots,n\}} \in \mathcal{C} \text{ such that } \forall z \in \mathbb{C} \text{ we have } p(z) = \sum_{i=0}^{n} a_i \cdot z^i$$

and

$$\mathrm{ord} \colon \mathcal{P} \to \mathbb{N}_0 \text{ where } \mathrm{ord}(p) = \mathrm{ord}(\mathrm{coef}(p))$$

Actuall it turns out that the relation between a polynomial and its coefficients is a bijection

Theorem 17.7. *The function*

$$\mathrm{coef} \colon \mathcal{P} \to \mathcal{C} \text{ where } \mathrm{coef}(p) = \{a_i\}_{i \in \{1,\ldots,n\}} \in \mathcal{C} \text{ such that } \forall z \in \mathbb{C} \text{ we have } p(z) = \sum_{i=0}^{n} a_i \cdot z^i$$

is a bijection. In a sense we can identify a polynomial with its coefficients.

Proof.

injectivity. If $\mathrm{coef}(p) = \mathrm{coef}(q)$ then there exist a $\{a_i\}_{i \in \{0,\ldots,n\}} \subseteq \mathcal{C}$ such that $\forall z \in \mathbb{C}$

$$p(z) = \sum_{i=0}^{n} a_i \cdot z^i = q(z)$$

so that $p = q$.

surjectivity. If $\{a_i\}_{i \in \{1,\ldots,n\}} \in \mathcal{C}$ then $\{a_i\}_{i \in \{1,\ldots,n\}} \subseteq \mathbb{C}$ and $a_n \neq 0$ if $n \neq 0$. So if we define $p \colon \mathbb{C} \to \mathbb{C}$ by $p(z) = \sum_{i=0}^{n} a_i \cdot z^i$ then $\mathrm{coef}(p) = \{a_i\}_{i \in \{1,\ldots,n\}}$. □

Note 17.8. To simplify notation we use the following convention. Instead of saying that the polynomial p is defined by the coefficients $\{a_i\}_{i \in \{0,\ldots,n\}} \subseteq \mathbb{C}$ with $a_n \neq 0$ such that $\forall z \in \mathbb{C}$ $p(z) = \sum_{i=0}^{n} a_i \cdot z^i$ we simply say that $p(z) = \sum_{i=0}^{n} a_i \cdot z^i$.

Definition 17.9. *A polynomial* $p \in \mathcal{P}$ *is* ***non constant*** *if* $0 < \mathrm{ord}(p)$ *a polynomial* p *is* ***constant*** *if* $\mathrm{ord}(p) = 0$.

We will show now that the product of polynomials is again a polynomial. First we need a little lemma.

Lemma 17.10. *Let $m,n\in\mathbb{N}_0$ 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 that

$$\{(r,s)\in\{0,\ldots,n\}\times\{0,\ldots,m\}|r+s=i\}\bigcap\{(r,s)\in\{0,\ldots,n\}\times\{0,\ldots,m\}|r+s=j\}=\varnothing$$

Proof. Given $i\in\{0,\ldots,n+m\}$ define

$$I_i=\{(r,s)\in\{0,\ldots,n\}\times\{0,\ldots,m\}|r+s=i\}\subseteq\{0,\ldots,n\}\times\{0,\ldots,m\}$$

then we have trivially that

$$\bigcup_{i\in\{0,\ldots,n+m\}}I_i\subseteq\{0,\ldots,n\}\times\{0,\ldots,m\}\tag{17.1}$$

If $(r,s)\in\{0,\ldots,n\}\times\{0,\ldots,m\}$ then $0\leqslant r\leqslant n\wedge 0\leqslant s\leqslant m$ so that $r+s\leqslant n+m$ hence

$$(r,s)\in I_{r+s}\subseteq\bigcup_{i\in\{0,\ldots,n+m\}}I_i$$

Combining with [eq: 17.1] proves that

$$\{0,\ldots,n\}\times\{0,\ldots,m\}=\bigcup_{i\in\{0,\ldots,n+m\}}I_i$$

Let $i,j\in\{0,\ldots,n+m\}$ with $i\neq j$ then if $(r,s)\in I_i\bigcap I_j$ we have that $i=r+s=j$ contradiction $i\neq j$ so that

$$I_i\bigcap I_j=\varnothing$$

□

Theorem 17.11. *Let $p,q\in\mathcal{P}$ be two polynomials then*

$$p\cdot q:\mathbb{C}\rightarrow\mathbb{C}\text{ defined by }(p\cdot q)(z)=p(z)\cdot q(z)$$

is also a polynomial. If $p(z)=\sum_{i=0}^n a_i\cdot z^i$ and $q(z)=\sum_{i=0}^m b_i\cdot z^i$ then $p\cdot q$ is defined by

$$(p\cdot q)(z)=\sum_{i=0}^{n+m}\left(\sum_{(k,l)\in\{(r,s)\in\{0,\ldots,n\}\times\{0,\ldots,m\}|r+s=i\}}a_k\cdot b_l\right)\cdot z^i$$

from which it follows that

$$\operatorname{ord}(p\cdot q)=\operatorname{ord}(p)+\operatorname{ord}(q)$$

Proof. Let $z\in\mathbb{C}$ then

$$\begin{aligned}
(p\cdot q)(z) &= p(z)\cdot q(z)\\
&= \left(\sum_{k=0}^n a_k\cdot z^l\right)\cdot\left(\sum_{l=0}^m b_l\cdot z^i\right)\\
&= \left(\sum_{k\in\{0,\ldots,n\}}a_k\cdot z^k\right)\cdot\left(\sum_{l\in\{0,\ldots,m\}}b_l\cdot z^l\right)\\
&\underset{[\text{theorem: }11.46]}{=} \sum_{(k,l)\in\{0,\ldots,n\}\times\{0,\ldots,m\}}a_k\cdot b_l\cdot z^k\cdot z^l\\
&= \sum_{(k,l)\in\{0,\ldots,n\}\times\{0,\ldots,m\}}a_k\cdot b_l\cdot z^{k+l}\\
&\underset{[\text{lemma: }17.10,\text{ theorem: }11.43]}{=} \sum_{i=0}^{n+m}\left(\sum_{(k,l)\in\{(r,s)\in\{0,\ldots,n\}\times\{0,\ldots,m\}|r+s=i\}}a_k\cdot b_l\cdot z^{k+l}\right)\\
&= \sum_{i=0}^{n+m}\left(\sum_{(k,l)\in\{(r,s)\in\{0,\ldots,n\}\times\{0,\ldots,m\}|r+s=i\}}a_k\cdot b_l\cdot z^i\right)\\
&= \sum_{i=0}^{n+m}\left(\sum_{(k,l)\in\{(r,s)\in\{0,\ldots,n\}\times\{0,\ldots,m\}|r+s=i\}}a_k\cdot b_l\right)\cdot z^i
\end{aligned}$$

To finish the proof we must prove that $\sum_{(k,l)\in\{(r,s)\in\{0,\ldots,n\}\times\{0,\ldots,m\}|r+s=n+m\}} a_k\cdot b_l\neq 0$. Now if $(r,s)\in\{(r,s)\in\{0,\ldots,n\}\times\{0,\ldots,m\}|r+s=n+m\}$ then $r+s=n+m$. If $r\neq n$ then $r<n$ and as $s\leqslant m$ we have then $r+s<n+m$ contradicting $r+s=n+m$ so we must have that $r=n$ which as $r+s=n+m$ means that $s=m$. Hence $(k,l)\in\{(r,s)\in\{0,\ldots,n\}\times\{0,\ldots,m\}|r+s=n+m\}=\{(n,m)\}$ so that $\sum_{(k,l)\in\{(r,s)\in\{0,\ldots,n\}\times\{0,\ldots,m\}|r+s=n+m\}} a_k\cdot b_l=a_n\cdot b_m$, as $a_n\neq 0\neq b_m$ it follows that

$$\sum_{(k,l)\in\{(r,s)\in\{0,\ldots,n\}\times\{0,\ldots,m\}|r+s=n+m\}} a_k\cdot b_l\neq 0$$ □

We can now use mathematical induction to prove that a finite product of polynomials is again a polynomial.

Theorem 17.12. *Let $n\in\mathbb{N}$ and $\{p_i\}_{i\in\{1,\ldots,n\}}\subseteq\mathcal{P}$ a finite family of polynomials then if we define*

$$\prod_{i=1}^{n} p_i:\mathbb{C}\to\mathbb{C} \text{ by } \left(\prod_{i=1}^{n} p_i\right)(z)=\prod_{i=1}^{n} p_i(z)$$

then $\prod_{i=1}^{n} p_i$ is a polynomial with

$$\operatorname{ord}\left(\prod_{i=1}^{n} p_i\right)=\sum_{i=1}^{n}\operatorname{ord}(p_i)$$

Proof. We use mathematical induction to prove this, so define

$$S=\left\{n\in\mathbb{N}|\text{If }\{p_i\}_{i\in\{1,\ldots,n\}}\subseteq\mathcal{P}\text{ then }\prod_{i=1}^{n} p_i\in\mathcal{P}\text{ and }\operatorname{ord}\left(\prod_{i=1}^{n} p_i\right)=\prod_{i=1}^{n}\operatorname{ord}(p_i)\right\}$$

then we have:

$1\in S$. 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 $1\in\mathcal{S}$

$n\in S\Rightarrow n+1\in\mathcal{S}$. then $\forall z\in\mathbb{C}$ we have that

$$\left(\prod_{i=1}^{n+1} p_i\right)(x)=\prod_{i=1}^{n+1} p_i(z)=p_{n+1}(z)\cdot\left(\prod_{i=1}^{n} p_i\right)(z)$$

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 [lemma: 17.11] is a polynomial, further

$$\begin{aligned}\operatorname{ord}\left(\prod_{i=1}^{n+1} p_i\right) &= \operatorname{ord}\left(p_{n+1}\cdot\prod_{i=1}^{n} p_i\right)\\ &\underset{[\text{lemma: }17.12]}{=} \operatorname{ord}(p_{n+1})+\operatorname{ord}\left(\prod_{i=1}^{n} p_i\right)\\ &\underset{n\in\mathcal{S}}{=} \operatorname{ord}(p_{n+1})+\sum_{i=1}^{n}\operatorname{ord}(p_i)\\ &= \sum_{i=1}^{n+1}\operatorname{ord}(p_i)\end{aligned}$$

□

17.1.2 Divergent limits

In the proof of the fundamental theorem of algebra we need the concept not of convergent limits but the concept of divergent limits.

Definition 17.13. *Let $f:\mathbb{R}\to\mathbb{R}$ be a function then*

1. $\lim_{x\to\infty} f(x)=\infty$ *if $\forall M\in\mathbb{R}$ there exist a $N\in\mathbb{R}$ such that $\forall x\geqslant N$ we have $f(x)\geqslant M$*
2. $\lim_{x\to\infty} f(x)=-\infty$ *if $\forall M\in\mathbb{R}$ there exists a $N\in\mathbb{R}$ such that $\forall x\geqslant N$ we have $f(x)\leqslant M$*

Theorem 17.14. *Let $f:\mathbb{R}\to\mathbb{R}$ be a function and consider then for*

$$-f:\mathbb{R}\to\mathbb{R} \text{ defined by } (-f)(x)=-f(x)$$

we have:

1. *If* $\lim_{x\rightarrow\infty} f(x)=\infty$ *then* $\lim_{x\rightarrow\infty}(-f)(x)=-\infty$
2. *If* $\lim_{x\rightarrow\infty} f(x)=-\infty$ *then* $\lim_{x\rightarrow\infty}(-f)(x)=\infty$

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 $\lim_{x\rightarrow\infty}(-f)(x)=-\infty$
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 $\lim_{x\rightarrow\infty}(-f)(x)=\infty$ □

Lemma 17.15. *Let* $n\in\mathbb{N}$, $\{a_i\}_{i\in\{0,\ldots,n\}}\subseteq\mathbb{R}$ *then*

1. *If* $a_n>0$ *we have* $\lim_{x\rightarrow\infty}\left(\sum_{i=0}^n a_i\cdot x^i\right)=\infty$
2. *If* $a_n<0$ *we have* $\lim_{x\rightarrow\infty}\left(\sum_{i=0}^n a_i\cdot x^i\right)=-\infty$

Proof.

1. we prove this by induction so let

$$\mathcal{S}=\left\{n\in\mathbb{N}|\forall\{a_i\}_{i\in\{0,\ldots,n\}}\subseteq\mathbb{R}\text{ with }a_n>0\text{ then we have }\lim_{x\rightarrow\infty}\left(\sum_{i=0}^n a_i\cdot x^i\right)=\infty\right\}$$

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 $\lim_{x\rightarrow\infty}\sum_{i=0}^1 a_i\cdot x^i=\infty$. Hence $1\in S$.

$n\in\mathcal{S}\Rightarrow n+1\in\mathcal{S}$. Let $\{a_i\}_{i\in\{0,\ldots,n+1\}}\subseteq\mathbb{R}$ with $a_{n+1}>0$ then

$$\sum_{i=0}^{n+1} a_i\cdot x^i=\left(\sum_{i=1}^{n+1} a_i\cdot x^i\right)+a_0=x\cdot\left(\sum_{i=1}^{n+1} a_i\cdot x^{i-1}\right)+a_0=x\cdot\left(\sum_{i=0}^{n} a_{i+1}\cdot x^i\right)+a_0,$$

hence if we define $\{b_i\}_{i\in\{0,\ldots,n\}}$ by $b_i=a_{i+1}$, then we have

$$\sum_{i=1}^{n+1} a_i\cdot x^i=x\cdot\left(\sum_{i=0}^{n} b^i\cdot x^i\right)+a_0\wedge b_n=a_{n+1}>0 \tag{17.2}$$

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

$$0<\max(1,M-a_0)\leqslant\sum_{i=0}^n b^i\cdot x^i$$

Take then $N=\max(N',1)$ then we have if $N\leqslant x$ that $1\leqslant x$ so that

$$\max(1,M-a_0)\leqslant x\cdot\max(1,M-a_0)$$

and

$$x\cdot\max(1,M-a_0)\leqslant x\cdot\sum_{i=0}^n b^i\cdot x^i$$

so that

$$M-a_0\leqslant\max(1,M-a_0)\leqslant x\cdot\sum_{i=0}^n b^i\cdot x^i$$

so that

$$M\leqslant a_0+x\cdot\sum_{i=0}^n b^i\cdot x^i\underset{[\text{eq: }17.2]}{=}\sum_{i=1}^{n+1} a_i\cdot x^i$$

so that $\lim_{x\rightarrow\infty}\sum_{i=1}^{n+1} a_i\cdot x^i$ proveing that $n+1\in S$.

2. If $a_n<0$ then $0<(-a_n)$ and we have by (1) that $\lim_{x\rightarrow\infty}-\sum_{i=0}^n a_i\cdot x^i=\lim_{z\rightarrow\infty}\sum_{i=0}^n(-a_i)\cdot x^i=\infty$ and thus by [theorem: 17.14] that $\lim_{x\rightarrow\infty}\left(\sum_{i=0}^n a_i\cdot x^i\right)=\lim_{x\rightarrow\infty}\left(-\sum_{i=0}^n a_i\cdot x^i\right)=-\infty$. □

17.1.3 Propertues of $\mathbb{C}$ needed for the fundamental theorem

First we define a norm on $\mathbb{C}$ that does not need the square root.

Theorem 17.16. *The function* $\|\|: \mathbb{C} \to \mathbb{R}$ *defined by* $\|z\| = |\mathrm{Re}(z)| + |\mathrm{Img}(z)|$ *is a norm on the vector space* $\mathbb{C}$ *over* $\mathbb{R}$.

Proof. We have:

1. If $z \in \mathbb{C}$ then $\|z\| = |\mathrm{Re}(z)| + |\mathrm{Img}(z)| \geqslant 0$
2. If $\alpha \in \mathbb{R}$ then for z we have
$$\begin{array}{rcl} \|\alpha \cdot z\| & = & |\mathrm{Re}(\alpha \cdot z)| + |\mathrm{Img}(\alpha \cdot z)| \\ & \underset{[\text{theorem: } 10.78]}{=} & |\alpha \cdot \mathrm{Re}(z)| + |\alpha \cdot \mathrm{Img}(z)| \\ & = & |\alpha| \cdot |\mathrm{Re}(z)| + |\alpha| \cdot |\mathrm{Img}(z)| \\ & = & |\alpha| \cdot (|\mathrm{Re}(z)| + |\mathrm{Img}(z)|) \\ & = & |\alpha| \cdot \|z\| \end{array}$$
3. If $z_1, z_2 \in \mathbb{C}$ then we have
$$\begin{array}{rcl} \|z_1 + z_2\| & = & |\mathrm{Re}(z_1 + z_2)| + |\mathrm{Img}(z_1 + z_2)| \\ & \underset{[\text{theorem: } 10.78]}{=} & |\mathrm{Re}(z_1) + \mathrm{Re}(z_2)| + |\mathrm{Img}(z_1) + \mathrm{Img}(z_2)| \\ & \leqslant & |\mathrm{Re}(z_1)| + |\mathrm{Re}(z_2)| + |\mathrm{Img}(z_1)| + |\mathrm{Img}(z_2)| \\ & = & (|\mathrm{Re}(z_1)| + |\mathrm{Img}(z_1)|) + (|\mathrm{Re}(z_2)| + |\mathrm{Img}(z_2)|) \\ & = & \|z_1\| + \|z_2\| \end{array}$$
4. If $z \in \mathbb{C}$ satisfies $\|z\| = 0$ then $|\mathrm{Re}(z)| + |\mathrm{Img}(z)| = 0$ so that $|\mathrm{Re}(z)| = 0 = |\mathrm{Img}(z)|$ hence $\mathrm{Re}(z) = \mathrm{Img}(z) = 0$ so that $z = 0$. □

The above norm has the following properties.

Lemma 17.17. *The norm* $\|\|$ *has the following properties:*

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 by [theorem: 10.78] $\mathrm{Img}(z) = 0$ so that $\|z\| = |\mathrm{Re}(z)| + |\mathrm{Img}(z)| = |\mathrm{Re}(z)|$
2. Let $z = \mathrm{Re}(z) + \mathrm{Img}(z) \in \mathbb{C}$ then
$$\|z\| = |\mathrm{Re}(z)| + |\mathrm{Img}(z)| = |\mathrm{Re}(z)| + |-\mathrm{Img}(z)| = \|\mathrm{Re}(x) - i \cdot \mathrm{Img}(z)\| = \|\bar{z}\|$$
3. If $a, b, c, d \in \mathbb{R}$ then we have $0 \leqslant (|a| - |b|)^2 = |a|^2 + |b|^2 - 2 \cdot |a| \cdot |b|$ so that
$$2 \cdot |a| \cdot |b| < |a|^2 + |b|^2 \tag{17.3}$$
So $(|a| + |b|)^2 = |a|^2 + |b|^2 + 2 \cdot |a| \cdot |b| \leqslant_{[\text{eq: } 17.3]} |a|^2 + |b|^2 + |a|^2 + |b|^2 = 2 \cdot (|a|^2 + |b|^2)$ proving that
$$(|a| + |b|)^2 \leqslant 2 \cdot (|a|^2 + |b|^2) \tag{17.4}$$
Similar we have $0 \leqslant (|c| - |d|)^2 = |c|^2 + |d|^2 - 2 \cdot |c| \cdot |d|$ so that
$$2 \cdot |c| \cdot |d| < |c|^2 + |d|^2 \tag{17.5}$$

So $(|c|+|d|)^2=|c|^2+|d|^2+2\cdot|c|\cdot|d| \leqslant_{[\text{eq: } 17.5]} |c|^2+|d|^2+|c|^2+|d|^2=2\cdot(|c|^2+|d|^2)$ proving that

$$(|c|+|d|)^2 \leqslant 2\cdot(|c|^2+|d|^2) \tag{17.6}$$

Hence we have by [eqs: 17.4, 17.6]

$$\begin{array}{rcl} (|a|+|b|)^2\cdot(|c|+|d|)^2 & \leqslant_{[\text{eq: } 17.4]} & 2\cdot(|a|^2+|b|^2)\cdot(|c|+|d|)^2 \\ & \leqslant_{[\text{eq: } 17.6]} & 2\cdot(|a|^2+|b|^2)\cdot 2\cdot(|c|^2+|d|^2) \\ & = & 4\cdot(|a|^2+|b|^2)\cdot(|c|^2+|d|^2) \end{array}$$

proving that

$$(|a|+|b|)^2\cdot(|c|+|d|)^2 \leqslant 4\cdot(|a|^2+|b|^2)\cdot(|c|^2+|d|^2) \tag{17.7}$$

Now

$$\begin{array}{rcl} (a\cdot c-b\cdot d)^2+(a\cdot d+b\cdot c)^2 & = & a^2\cdot c^2+b^2\cdot d^2-2\cdot a\cdot c\cdot b\cdot d+a^2\cdot d^2+b^2\cdot c^2+4\cdot a\cdot d\cdot b\cdot c \\ & = & a^2\cdot c^2+b^2\cdot d^2+a^2\cdot d^2+b^2\cdot c^2 \\ & = & a^2\cdot(c^2+d^2)+b^2\cdot(d^2+c^2) \\ & = & (a^2+b^2)\cdot(c^2+d^2) \end{array}$$

substituting the above in [eq: 17.7] we have that

$$\begin{array}{rcl} (|a|+|b|)^2\cdot(|c|+|d|)^2 & \leqslant & 4\cdot(a\cdot c-b\cdot d)^2+(a\cdot d+b\cdot c)^2 \\ & = & 4\cdot|a\cdot c-b\cdot d|^2+|a\cdot d+b\cdot c|^2 \\ & \leqslant & 4\cdot|a\cdot c-b\cdot d|^2+|a\cdot d+b\cdot c|^2+8\cdot|a\cdot c-b\cdot d|\cdot|a\cdot d+b\cdot c| \\ & = & 4\cdot(|a\cdot c-b\cdot d|+|a\cdot d+b\cdot c|)^2 \end{array}$$

Using the fact that the square root is increasing on $\mathbb{R}^+$ [see theorem: 10.69] it follows that

$$(|a|+|b|)\cdot(|c|+|d|) \leqslant 2\cdot(|a\cdot c-b\cdot d|+|a\cdot d+b\cdot c|) \tag{17.8}$$

Let $z=a+i\cdot b\in\mathbb{C}$ and $w=c+i\cdot d\in\mathbb{C}$ then $z\cdot w=(a\cdot c-b\cdot d)+i\cdot(a\cdot d+b\cdot c)$ so that

$$\begin{array}{rcl} \|z\cdot w\| & = & |a\cdot c-b\cdot d|+|a\cdot d+b\cdot c| \\ & \geqslant_{[\text{eq: } 17.8]} & \frac{1}{2}\cdot(|a|+|b|)\cdot(|c|+|d|) \\ & = & \frac{1}{2}\cdot\|z\|\cdot\|w\| \end{array}$$

proving that

$$\frac{1}{2}\cdot\|z\|\cdot\|w\| \leqslant \|z\cdot w\| \tag{17.9}$$

Further we have

$$\begin{array}{rcl} \|z\cdot w\| & = & |a\cdot c-b\cdot d|+|a\cdot d+b\cdot c| \\ & \leqslant & |a|\cdot|c|+|b|\cdot|d|+|a|\cdot|d|+|b|\cdot|c| \\ & = & |a|\cdot(|c|+|d|)+|b|\cdot(|d|+|c|) \\ & = & (|a|+|b|)\cdot(|c|+|d|) \\ & = & \|z\|\cdot\|w\| \end{array}$$

proving that

$$\|z\cdot w\| \leqslant \|z\|\cdot\|w\|$$

(3) is proved by the above and [eq: 17.9].

4. If $z=a+i\cdot b$ and $w=c+i\cdot d$ so that $z\cdot w=(a\cdot c-b\cdot d)+i\cdot(a\cdot d+b\cdot c)$ then we have

$$\begin{array}{rcl} \|\mathrm{Re}(z\cdot w)\| & = & |\mathrm{Re}(z\cdot w)| \\ & = & |a\cdot c-b\cdot d| \\ & \leqslant & |a\cdot c-b\cdot d|+|a\cdot d+b\cdot c| \\ & = & \|z\cdot w\| \end{array}$$

5. We prove this by induction, so define

$$S = \{n \in \mathbb{N} | \|z^n\| \leqslant \|z\|^n\}$$

then we have:

$\boldsymbol{1 \in S}$. As $\|z^1\| = \|z\| = \|z\|^1$ it follows that $1 \in S$.

$\boldsymbol{n \in S \Rightarrow n+1 \in S}$. As $\|z^{n+1}\| = \|z \cdot z^n\| \leqslant_{(3)} \|z\| \cdot \|z^n\| \leqslant_{n \in \mathbb{N}} \|z\| \cdot \|z\|^n = \|z\|^{n+1}$ proving that $n+1 \in S$.

6. For $z = x + i \cdot y$ we for $n \in \mathbb{N}$ two cases:

$\boldsymbol{n=1}$. Then $\|\mathrm{Re}(z^1)\| = \|\mathrm{Re}(z)\| = |x| < |x| + |y| = \|z\| = \|z\|^1$ so that $\|\mathrm{Re}(z^1)\| \leqslant \|z\|^1$

$\boldsymbol{1 < n}$. Then $1 \leqslant n-1$ and we have

$$\|\mathrm{Re}(z^n)\| = \|\mathrm{Re}(z \cdot z^{n-1})\| \leqslant_{(4)} \|z\| \cdot \|z^{n-1}\| \leqslant_{(5)} \|z\| \cdot \|z\|^{n-1} = \|z\|^n$$

so that

$$\|\mathrm{Re}(z^n)\| \leqslant \|z\|^n$$

7. This is proved by induction, so define

$$S = \left\{n \in \mathbb{N} | \frac{\|z\|^n}{2^n} \leqslant \|z^n\|\right\}$$

then we have:

$\boldsymbol{1 \in S}$. As $\frac{1}{2} < 1$ we have $\frac{\|z^1\|}{2} \leqslant \|z^1\| = \|z\| = \|z\|^1$ so that $1 \in S$.

$\boldsymbol{n \in S \Rightarrow n+1 \in S}$. Then we have

$$\frac{\|z\| \cdot \|z^n\|}{2} \leqslant_{(3)} \|z \cdot z^n\| = \|z^{n+1}\| \tag{17.10}$$

As $n \in S$ we have $\frac{\|z\|^n}{2^n} \leqslant \|z^n\|$. So that

$$\frac{\|z\|^{n+1}}{2^{n+1}} = \frac{\|z\| \cdot \|z\|^n}{2 \cdot 2^n} \leqslant_{n \in S} \frac{\|z\|}{2} \cdot \|z\|^n \leqslant_{[\text{eq: } 17.10]} \|z^{n+1}\|$$

giving $n+1 \in S$. □

In the proof of the fundamental theorem of algebra we use the extreme value theorem [theorem: 14.235], so we need to look at continuous functions using the above norm. This is the subject of the following theorem.

Theorem 17.18. *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 $(\cdot)^n: \mathbb{C} \rightarrow \mathbb{C}$ defined by $(\cdot)^n(z) = z^n$ is continuous.*
3. *Given $n \in \mathbb{N}$, $\{\alpha_i\}_{i \in \{0,\ldots,n\}}$ then the function $p: \mathbb{C} \rightarrow \mathbb{C}$ defined by $p(z) = \sum_{i=0}^n \alpha_i \cdot z^i$ is continuous.*
4. *The function $\langle z\rangle: \mathbb{C} \rightarrow \mathbb{R}$ defined by $\langle z\rangle = \langle\rangle(z) = z \cdot \bar{z} = |z|^2$ is continuous*
5. *Given $n \in \mathbb{N}_0$, $\{a_i\}_{i \in \{0,\ldots,n\}}$ the function*

$$p \cdot \bar{p}: \mathbb{C} \rightarrow \mathbb{R} \text{ defined by } (p \cdot \bar{p})(z) = \left(\sum_{i=0}^n \alpha_i \cdot z^i\right) \cdot \overline{\left(\sum_{i=0}^n \alpha_i \cdot z^i\right)}$$

is continuous.

6. *The function $(\bar{\cdot}): \mathbb{C} \rightarrow \mathbb{C}$ defined by $(\bar{\cdot})(z) = \bar{z}$ 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 $p._z: \mathbb{R} \rightarrow \mathbb{C}$ defined by $p._z(\alpha) = p(\alpha \cdot z)$ is continuous.*
9. *The function $\mathrm{Re}: \mathbb{C} \rightarrow \mathbb{R}$ is continuous*
10. *If $f: \mathbb{C} \rightarrow \mathbb{C}$ and $g: \mathbb{C} \rightarrow \mathbb{C}$ are continuous functions in $\langle \mathbb{C}, \|\|\rangle$ then*

$$f \cdot g: \mathbb{C} \rightarrow \mathbb{C} \text{ define by } (f \cdot g) \text{ is continuous in } \langle \mathbb{C}, \|\|\rangle$$

Proof.

1. As $\forall z \in \mathbb{C}$ we have $(\alpha \cdot)(z) = \alpha \cdot z = \alpha \cdot \mathrm{Id}_{\mathbb{C}}(z)$ we have $(\alpha \cdot) = \alpha \cdot \mathrm{Id}_{\mathbb{C}}$. By [example: 14.132] $\mathrm{Id}_{\mathbb{C}}$ is continuous so that by [theorem: 14.144] $(\alpha \cdot)$ is continuous.

2. As $\forall z \in \mathbb{C}$ we have that $(\cdot)^n(z) = z^n = (\mathrm{Id}_{\mathbb{C}}(z))^n$ we have, as $\mathrm{Id}_{\mathbb{C}}$ is continuous, by [theorem: 14.147] that $(\cdot)^n$ is continuous.
3. Let $i \in \{1, \ldots, n\}$ then by (1) and (2) $(\alpha_i \cdot)$ and $(\cdot)^i$ are continuous. So, as the composition of continuous functions is continuous [see theorem: 14.138], it follows that $(\alpha_i \cdot) \circ (\cdot)^i$ is continous. Using now [theorem: 14.145] it follows that
$$\sum_{i=0}^{n} ((\alpha_i \cdot) \circ (\cdot)^i)$$
is continuous. Let $z \in \mathbb{C}$ then
$$\left(\sum_{i=0}^{n} ((\alpha_i \cdot) \circ (\cdot)^i)\right)(z) = \sum_{i=0}^{n} ((\alpha_i \cdot) \circ (\cdot)^i)(z) = \sum_{i=0}^{n} (\alpha_i \cdot)((\cdot)^i(z)) = \sum_{i=0}^{n} \alpha_i \cdot z^i = p(z)$$
it follows that $p = \sum_{i=0}^{n} ((\alpha_i \cdot) \circ (\cdot)^i)$. Hence p is continuous.
4. Let $z \in \mathbb{C}$ and $\varepsilon \in \mathbb{R}^+$. Take $\delta = \min\left(1, \frac{\varepsilon}{2 \cdot \|y\| + 1}\right) \in \mathbb{R}^+$ then if $y \in \mathbb{C}$ with $\|z - y\| < \delta$ then if we define $h = z - y$ we have $\|h\| < \delta$ and
$$\begin{array}{rcl} |\langle z \rangle - \langle y \rangle| & \underset{[\text{lemma: } 17.17]}{=} & \|\langle z \rangle - \langle y \rangle\| \\ & & \|\langle y + h \rangle - \langle y \rangle\| \\ & = & \|(y+h) \cdot \overline{(y+h)} - y \cdot \overline{y}\| \\ & \underset{[\text{theorem: } 10.83]}{=} & \|(y+h) \cdot (\overline{y} + \overline{h}) - y \cdot \overline{y}\| \\ & = & \|y \cdot \overline{y} + y \cdot \overline{h} + h \cdot \overline{y} + h \cdot \overline{h} - y \cdot \overline{y}\| \\ & = & \|y \cdot \overline{h} + h \cdot \overline{y} + h \cdot \overline{h}\| \\ & \leqslant & \|y \cdot \overline{h}\| + \|h \cdot \overline{y}\| + \|h \cdot \overline{h}\| \\ & \leqslant_{\underset{[\text{lemma: } 17.17(3)]}{=}} & \|y\| \cdot \|\overline{h}\| + \|h\| \cdot \|\overline{y}\| + \|h\| \cdot \|\overline{h}\| \\ & \underset{[\text{lemma: } 17.17(2)]}{=} & \|y\| \cdot \|h\| + \|h\| \cdot \|y\| + \|h\| \cdot \|h\| \\ & = & \|h\| \cdot (2 \cdot \|y\| + \|h\|) \\ & \leqslant_{\|h\| < \delta \leqslant 1} & \|h\| \cdot (2 \cdot \|y\| + 1) \\ & < & \frac{\varepsilon}{2 \cdot \|y\| + 1} \cdot (2 \cdot \|y\| + 1) \\ & = & \varepsilon \end{array}$$
proving continuity of $\langle \cdot \rangle$.
5. Let $z \in \mathbb{C}$ then
$$\begin{array}{rcl} (\langle \cdot \rangle \circ p)(z) & = & \langle \cdot \rangle(p(z)) \\ & = & \langle \cdot \rangle\left(\sum_{i=0}^{n} \alpha_i \cdot z^i\right) \\ & = & \left(\sum_{i=0}^{n} \alpha_i \cdot z^i\right) \cdot \overline{\left(\sum_{i=0}^{n} \alpha_i \cdot z^i\right)} \\ & = & (p \cdot p)(z) \end{array}$$
so that
$$(\langle \cdot \rangle \circ p) = p \cdot p$$
By (3),(4) p and $\langle \cdot \rangle$ are continous, hence using [theorem: 14.138] $p \cdot p$ is continuous.
6. Let $\varepsilon \in \mathbb{R}$, $z \in \mathbb{C}$ then if $z' \in \mathbb{C}$ with $\|z - z'\| < \varepsilon$ we have
$$\|\overline{z} - \overline{z'}\| \underset{[\text{theorem: } 10.81]}{=} \|\overline{z - z'}\| \underset{[\text{lemma: } 17.17]}{=} \|z - z'\| < \varepsilon$$
7. Let $z \in \mathbb{C}$, $\alpha \in \mathbb{R}$, given $\varepsilon > 0$ define $\delta = \frac{\varepsilon}{\|z\| + 1} \in \mathbb{R}^+$ 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.

8. Given $\alpha \in \mathbb{R}$ we have $(p_{\cdot z})(\alpha) = p(\alpha \cdot z) = p((\cdot z)(\alpha)) = (p \circ (\cdot z))(\alpha)$ so that $p_{\cdot z} = p \circ (\cdot z)$. 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 $\varepsilon \in \mathbb{R}^+$ we have for $\delta = \varepsilon$ that $\forall z' \in \mathbb{C}$ with $\|z - z'\| < \delta$ so 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.
10. Let $z \in \mathbb{C}$ and $\varepsilon \in \mathbb{R}^+$ then as f, g are continuous there exist a $\delta_1, \delta_2 \in \mathbb{R}^+$ such that
$$\forall z' \in \mathbb{C} \text{ with } \|z - z'\| < \delta_1 \text{ we have } \|f(z) - f(z')\| < \min\left(1, \frac{\varepsilon}{2 \cdot (1 + \|g(z)\|)}\right) \tag{17.11}$$
$$\forall z' \in \mathbb{C} \text{ with } \|z - z'\| < \delta_2 \text{ we have } \|g(z) - g(z')\| < \frac{\varepsilon}{2 \cdot (1 + \|f(z)\|)} \tag{17.12}$$
Defin $\delta = \min(\delta_1, \delta_2) \in \mathbb{R}^+$ then if $z' \in \mathbb{C}$ with $\|z'\| < \delta$ we have
$$\|f(z')\| \leqslant \|f(z) - (f(z) - f(z'))\| \leqslant \|f(z) - f(z')\| + \|f(z)\| < 1 + \|f(z\| \tag{17.13}$$
Now
$$\begin{array}{rc} \|(f \cdot g)(z) - (f \cdot g)(z')\| & = \\ \|f(z) \cdot g(z) - f(z') \cdot g(z')\| & = \\ \|f(z) \cdot g(z) - f(z') \cdot g(z) + f(z') \cdot g(z) - f(z') \cdot g(z')\| & \leqslant \\ \|f(z) \cdot g(z) - f(z') \cdot g(z)\| + \|f(z') \cdot g(z) - f(z') \cdot g(z')\| & = \\ \|(f(z) - f(z')) \cdot g(z)\| + \|f(z') \cdot (g(z) - g(z'))\| & \leqslant_{[\text{lemma: } 17.17(4)]} \\ \|f(z) - f(z')\| \cdot \|g(z)\| + \|f(z')\| \cdot \|g(z) - g(z')\| & <_{[\text{eq: } 17.13]} \\ \|f(z) - f(z')\| \cdot \|g(z)\| + (1 + \|f(z)\|) \cdot \|g(z) - g(z')\| & <_{[\text{eqs: } 17.11,\ 17.12} \\ \frac{\varepsilon}{2 \cdot (1 + \|g(z)\|)} \cdot \|g(z)\| + (1 + \|f(z)\|) \cdot \frac{\varepsilon}{2 \cdot (1 + \|f(z)\|)} & < \\ \frac{\varepsilon}{2} + \frac{\varepsilon}{2} & = \\ \varepsilon & \end{array}$$
proving that $f \cdot g$ is continuous at z. As $z \in \mathbb{C}$ was chosen arbitrary $f \cdot g$ is continuous.

□

The following very specific lemma will be used in the proof of the fundamental theorem of algebra.

Lemma 17.19. *Let $n \in \mathbb{N}_0$ then*

1. $(1 + i)^{4 \cdot n + 2} = 2 \cdot (-4)^n \cdot i$
2. $(1 - i)^{4 \cdot n + 2} = -2 \cdot (-4)^n \cdot i$

Proof. This is easely proved by induction.

1. Take $S = \{n \in \mathbb{N}_0 | (1 + i)^{4 \cdot n + 2} = 2 \cdot (-4)^n \cdot i\}$ then we have:

   0 ∈ S. $(1 + i)^{4 \cdot 0 + 2} = (1 + i)^2 = 1^2 + 2 \cdot i + i^2 = 1 + 2 \cdot i - 1 = 2 \cdot i = 2 \cdot (-4)^0 \cdot i$ so that $0 \in S$.

   n ∈ S ⇒ n + 1 ∈ S. We have
$$\begin{array}{rcl} (1 + i)^{4 \cdot (n+1) + 2} & = & (1 + i)^{(4 \cdot n + 2) + 4} \\ & = & (1 + i)^{(4 \cdot n + 2)} \cdot (1 + i)^4 \\ & \underset{n \in S}{=} & 2 \cdot (-4)^n \cdot i \cdot (1 + i)^4 \\ & = & 2 \cdot (-4)^n \cdot i \cdot (1 + i)^2 \cdot (1 + i)^2 \\ & = & 2 \cdot (-4)^n \cdot i \cdot (1 + 2 \cdot i + i^2) \cdot (1 + 2 \cdot i + i^2) \\ & = & 2 \cdot (-4)^n \cdot i \cdot (2 \cdot i) \cdot (2 \cdot i) \\ & = & 2 \cdot (-4)^n \cdot i \cdot (4 \cdot i^2) \\ & = & 2 \cdot (-4)^n \cdot i \cdot (-4) \\ & = & 2 \cdot (-4)^{n+1} \cdot i \end{array}$$

proving that $n+1\in S$.

2. Take $S=\{n\in\mathbb{N}_0|(1-i)^{4\cdot n+2}=-2\cdot(-4)^n\cdot i\}$ then we have:

   $\mathbf{0}\in\boldsymbol{S}$. $(1-i)^{4\cdot 0+2}=(1-i)^2=1^2-2\cdot i+i^2=1-2\cdot i-1=-2\cdot i=-2\cdot(-4)^0\cdot i$ so that $0\in S$.

   $\boldsymbol{n}\in\boldsymbol{S}\Rightarrow\boldsymbol{n}+\mathbf{1}\in\boldsymbol{S}$. We have

$$\begin{aligned}(1-i)^{4\cdot(n+1)+2} &= (1-i)^{(4\cdot n+2)+4}\\ &= (1-i)^{(4\cdot n+2)}\cdot(1-i)^4\\ &\underset{n\in S}{=} -2\cdot(-4)^n\cdot i\cdot(1-i)^4\\ &= -2\cdot(-4)^n\cdot i\cdot(1-i)^2\cdot(1-i)^2\\ &= -2\cdot(-4)^n\cdot i\cdot(1-2\cdot i+i^2)\cdot(1-2\cdot i+i^2)\\ &= -2\cdot(-4)^n\cdot i\cdot(-2\cdot i)\cdot(-2\cdot i)\\ &= -2\cdot(-4)^n\cdot i\cdot(4\cdot i^2)\\ &= -2\cdot(-4)^n\cdot(-4)\cdot i\\ &= -2\cdot(-4)^{n+1}\cdot i\end{aligned}$$

   proving that $n+1\in S$. □

We also use the binomial expansion of $(a+b)^n$ so we need the following definitions and theorems.

Definition 17.20. *Given $n\in\mathbb{N}_0$ then we define $n!\in\mathbb{N}_0$ by*

$$n!=\begin{cases}1 \text{ if } n=0\\ n\cdot(n-1) \text{ if } 0<n\end{cases}$$

Definition 17.21. *Given $n\in\mathbb{N}_0$ and $k\in\{0,\ldots,n\}$ then*

$$\binom{n}{k}=\frac{n!}{k!\cdot(n-k)!}$$

Theorem 17.22. *Let $n\in\mathbb{N}_0$ then we have*

1. $\binom{n}{0}=1$
2. $\binom{n}{n}=1$
3. *If $n\in\mathbb{N}$ and $0<k\leqslant n$ then* $\binom{n}{k}+\binom{n}{k-1}=\binom{n+1}{k}$

Proof.

1. $\binom{n}{0}=\frac{n!}{0!\cdot(n-0)!}=\frac{n!}{1\cdot n!}=1$
2. $\binom{n}{n}=\frac{n!}{n!\cdot(n-n)!}=\frac{n!}{n!\cdot 0!}=\frac{n!}{n!\cdot 1}=1$
3. We have

$$\begin{aligned}\binom{n}{k}+\binom{n}{m-1} &= \frac{n!}{k!\cdot(n-k)!}+\frac{n!}{(k-1)!\cdot(n-(k-1))!}\\ &= \frac{n!\cdot((n+1)-k)}{k!\cdot(n-k)!\cdot((n+1)-k)}+\frac{n!\cdot k}{k\cdot(k-1)!\cdot(n+1-k)}\\ &= \frac{n!\cdot(n+1-k)}{k!\cdot((n+1)-k)!}+\frac{n!\cdot k}{k!\cdot((n+1)-k)!}\\ &= \frac{n!\cdot(n+1)}{k!\cdot((n+1)-k)!}\\ &= \frac{(n+1)!}{k!\cdot((n+1)-k)!}\\ &= \binom{n+1}{k}\end{aligned}$$

□

We prove now the binomial formula.

Theorem 17.23. (Binomial Formula) *Let F be a field, $n\in\mathbb{N}$, $a,b\in F$ then*

$$(a+b)^n=\sum_{k=0}^{n}\binom{n}{k}\cdot a^k\cdot b^{n-k}$$

Proof. We prove this by induction, so defined

$$S=\left\{n\in\mathbb{N}|(a+b)^n=\sum_{k=0}^{n}\binom{n}{k}\cdot a^k\cdot b^{n-k}\right\}$$

then we have:

$1\in S$. As $\sum_{k=0}^{1}\binom{1}{k}\cdot a^k\cdot b^{n-k}=\binom{1}{0}\cdot a^0\cdot b^{1-0}+\binom{1}{1}\cdot a^1\cdot b^{1-1}\underset{[\text{theorem: } 17.22]}{=}a+b=(a+b)^1$

$n\in S\Rightarrow n+1\in S$. We have

$$\begin{aligned}
(a+b)^{n+1} &= \\
(a+b)^n\cdot(a+b) &\underset{n\in S}{=} \\
\left(\sum_{k=0}^{n}\binom{n}{k}\cdot a^k\cdot b^{n-k}\right)\cdot(a+b) &= \\
\sum_{k=0}^{n}\binom{n}{k}\cdot a\cdot a^k\cdot b^{n-k}+\sum_{k=0}^{n}\binom{n}{k}\cdot a^k\cdot b\cdot b^{n-k} &= \\
\sum_{k=0}^{n}\binom{n}{k}\cdot a^{k+1}\cdot b^{n-k}+\sum_{k=0}^{n}\binom{n}{k}\cdot a^k\cdot b^{n-k+1} &= \\
\sum_{k=1}^{n+1}\binom{n}{k-1}\cdot a^k\cdot b^{(n+1)-k}+\sum_{k=0}^{n}\binom{n}{k}\cdot a^k\cdot b^{n-k+1} &= \\
\binom{n}{n+1-1}\cdot a^{n+1}\cdot b^{(n+1)-(n+1)}+\sum_{k=1}^{n}\binom{n}{k-1}\cdot a^k\cdot b^{(n+1)-k}+\sum_{k=0}^{n}\binom{n}{k}\cdot a^k\cdot b^{n-k+1} &= \\
a^{n+1}+\sum_{k=1}^{n}\binom{n}{k-1}\cdot a^k\cdot b^{(n+1)-k}+\sum_{k=0}^{n}\binom{n}{k}\cdot a^k\cdot b^{n-k+1} &= \\
a^{n+1}+\sum_{k=1}^{n}\binom{n}{k-1}\cdot a^k\cdot b^{(n+1)-k}+\sum_{k=1}^{n}\binom{n}{k}\cdot a^k\cdot b^{n-k+1}+\binom{n}{0}\cdot a^0\cdot b^{n-0+1} &= \\
a^{n+1}+\sum_{k=1}^{n}\binom{n}{k-1}\cdot a^k\cdot b^{(n+1)-k}+\sum_{k=1}^{n}\binom{n}{k}\cdot a^k\cdot b^{n-k+1}+b^{n+1} &= \\
a^{n+1}+b^{n+1}+\sum_{k=1}^{n}\left(\binom{n}{k-1}+\binom{n}{k}\right)\cdot a^k\cdot b^{(n+1)-k} &= \\
a^{n+1}+b^{n+1}+\sum_{k=1}^{n}\binom{n+1}{k}\cdot a^k\cdot b^{(n+1)-k} &= \\
\binom{n+1}{n+1}\cdot a^{(n+1)}\cdot b^{(n+1)-(n+1)}+\binom{n+1}{0}\cdot a^n\cdot b^{(n+1)-0}+\sum_{k=1}^{n}\binom{n+1}{k}\cdot a^k\cdot b^{(n+1)-k} &= \\
\sum_{k=0}^{n+1}\binom{n+1}{k}\cdot a^k\cdot b^{(n+1)-k} &
\end{aligned}$$

proving that $n+1\in S$. □

Lemma 17.24. *$p\in\mathcal{P}$ be a polynomial [see definition: 17.9] and $z_0\in\mathbb{C}$ then*

$$p_{(+z_0)}:\mathbb{C}\rightarrow\mathbb{C} \text{ defined by } p_{(+z_0)}(z)=p(z_0+z)$$

is a polynomial with $\operatorname{ord}(p_{(+z_0)})=\operatorname{ord}(p)$.

Proof. We prove this by induction so let

$$S=\{n\in\mathbb{N}_0|\text{If } p\in\mathcal{P} \text{ with } \operatorname{ord}(p)=n \text{ then } p_{(+z_0)}\in\mathcal{P} \text{ with } \operatorname{ord}(p)=\operatorname{ord}(p_{(+z_0)})\}$$

then we have:

$\boldsymbol{0\in S}$. If $p\in\mathcal{P}$ with $\operatorname{ord}(p)=0$ then $p=\sum_{i=0}^{0} a_i\cdot z^i=a_0\cdot z^0=a_0$. If $z\in\mathbb{C}$ then $p_{(+z_0)}=p(z+z_0)=a_0$ so that $p_{(+z_0)}=p$. Hence $0\in S$

$\boldsymbol{n\in S\Rightarrow n+1\in S}$. Let $p\in\mathcal{P}$ with $\operatorname{ord}(p)=n+1$ then there exist a $\{a_i\}_{i\in\{0,\ldots,n+1\}}\subseteq\mathbb{C}$ such that $a_{n+1}\neq 0$. Define

$$\{b_i\}_{i\in\{0,\ldots,n\}} \text{ by } b_i=a_{i+1}$$

then $b_n=a_{n+1}\neq 0$ so that $q=\sum_{i=0}^{n} b_i\cdot z^i$ is a polynomial with $\operatorname{ord}(q)=n$. Further we have for $z\in\mathbb{C}$ that:

$$\begin{aligned} p(z) &= \sum_{i=0}^{n+1} a_i\cdot z^i \\ &= a_0+\sum_{i=1}^{n+1} a_i\cdot z^i \\ &= a_0+\sum_{i=1}^{n+1} a_i\cdot z\cdot z^{i-1} \\ &= a_0+z\cdot\sum_{i=1}^{n+1} a_i\cdot z^{i-1} \\ &= a_0+z\cdot\sum_{i=0}^{n} a_{i+1}\cdot z^i \\ &= a_0+z\cdot\sum_{i=0}^{n} b_i\cdot z^i \\ &= a_0+z\cdot q(z) \end{aligned} \tag{17.14}$$

As $n\in S$ we have that $q_{(+z_0)}$ is a polynomial if order n, hence there exist a $\{c_i\}_{i\in\{0,\ldots,n\}}\subseteq\mathbb{C}$ with $c_n\neq 0$ such that

$$\forall z\in\mathbb{C} \text{ we have } q(z_0+z)=q_{(+z_0)}=\sum_{i=0}^{n} c_i\cdot z^i \tag{17.15}$$

Further we have

$$\begin{aligned} p_{(+z_0)}(z) &= p(z_0+z) \\ &\underset{\text{[eq: 17.14]}}{=} a_0+(z_0+z)\cdot q(z_0+z) \\ &\underset{\text{[eq: 17.15]}}{=} a_0+(z_0+z)\cdot\sum_{i=0}^{n} c_i\cdot z^i \\ &= a_0+z_0\cdot\sum_{i=0}^{n} c_i\cdot z^i+z\cdot\sum_{i=0}^{n} c_i\cdot z^i \\ &= a_0+\sum_{i=0}^{n} c_i\cdot z_0\cdot z^i+\sum_{i=0}^{n} c_i\cdot z^{i+1} \\ &= a_0+\sum_{i=0}^{n} c_i\cdot z_0\cdot z^i+\sum_{i=1}^{n+1} c_{i-1}\cdot z^i \\ &= a_0+c_0\cdot z_0+\sum_{i=1}^{n} c_i\cdot z_0\cdot z^i+\sum_{i=1}^{n+1} c_{i-1}\cdot z^i \\ &= a_0+c_0\cdot z_0+\sum_{i=1}^{n} c_i\cdot z_0\cdot z^i+\sum_{i=1}^{n} c_{i-1}\cdot z^i+c_n\cdot z^{n+1} \\ &= (a_0+c_0\cdot z_0)+\sum_{i=1}^{n} (c_i\cdot z_0+c_{i-1})\cdot z^i+c_n\cdot z^{n+1} \\ &= \sum_{i=1}^{n+1} d_i\cdot z^i \end{aligned}$$

where

$$\{d_i\}_{i\in\{0,\ldots,n+1\}} \text{ is defined by } d_i = \begin{cases} a_0 + c_0\cdot z_0 \text{ if } i=0 \\ c_i\cdot z_0 + c_{i-1} \text{ if } i\in\{1,\ldots,n\} \\ c_n \text{ if } i=n+1 \end{cases}$$

As $d_{n+1}=c_n\neq 0$ it follows that $p_{(+z_0)}(z)$ is a polynomial with $\mathrm{ord}(p_{(z_0)})=n+1=\mathrm{ord}(p)$ which proves that $n+1\in S$. □

Lemma 17.25. *Let $p:\mathbb{C}\rightarrow\mathbb{C}$ be a polynomial of order $n\in\mathbb{N}$ then there exists a $1\leqslant m\leqslant n$ and a polynomial q of order $n-m$ such that $p(z)=p(0)+z^m\cdot q(z)$ and $q(0)\neq 0$.*

Proof. As p is a polynomial of order n there exist a $\{a_i\}_{i\in\{0,\ldots,n\}}\subseteq\mathbb{C}$ with $a_n\neq 0$ so that $\forall z\in\mathbb{C}$

$$p(z)=\sum_{i=0}^{n} a_i\cdot z^i$$

Then we have that $p(0)=\sum_{i=0}^{n} a_i\cdot 0^i = a_0+\sum_{i=1}^{n} a_i\cdot 0^i = a_0+\sum_{i=1}^{n} a_i\cdot 0 = a_0$ proving that

$$a_0 = p(0) \tag{17.16}$$

Define $K=\{i\in\{1,\ldots,n\}|a_i\neq 0\}$ then as $1\leqslant n$ and $a_n\neq 0$ $K\neq\varnothing$ so that $m=\min(K)$ exist and $\forall i\in\{0,\ldots,m-1\}$ we have $a_i=0$ and $a_m\neq 0$. Define now

$$\{c_i\}_{\in\{0,\ldots,n-m\}} \text{ by } c_i = a_{i+m} \text{ then } c_{n-m}=a_n\neq 0$$

and the polynomial q or order $n-m$ by

$$q(z)=\sum_{i=0}^{n-m} c_i\cdot z^i \tag{17.17}$$

then we have for $z\in\mathbb{C}$

$$\begin{aligned} p(z) &= \sum_{i=0}^{n} a_i\cdot z^i \\ &= \sum_{i\in\{0,\ldots,n\}} a_i\cdot z^i \\ &= a_0+\sum_{i\in\{1,\ldots,n\}} a_i\cdot z^i \\ &= a_0+\sum_{i\in\{1,\ldots,m-1\}} a_i\cdot z^i+\sum_{i\in\{m,\ldots,n\}} a_i\cdot z^i \\ &= a_0+\sum_{i\in\{1,\ldots,m-1\}} 0\cdot z^i+\sum_{i\in\{m,\ldots,n\}} a_i\cdot z^i \\ &= a_0+\sum_{i=m}^{n} a_i\cdot z^i \\ &= a_0+\sum_{i=m}^{n} a_i\cdot z^m\cdot z^{i-m} \\ &= a_0+z^m\cdot\sum_{i=m}^{n} a_i\cdot z^{i-m} \\ &= a_0+z^m\cdot\sum_{i=0}^{n-m} a_{i+m}\cdot z^i \\ &= a_0+z^m\cdot\sum_{i=0}^{n-m} c_i\cdot z^i \\ &\underset{[\text{eqs: } 17.16,\ 17.17]}{=} p(0)+z^m\cdot q(z) \end{aligned} \tag{17.18}$$

Finally we have $q(0) = \sum_{i=0}^{n-m} c_i \cdot 0^i = c_0 + \sum_{i=1}^{n-m} c_i \cdot 0^i = c_0 + \sum_{i=1}^{n-m} c_i \cdot 0 = c_0 = a_m \neq 0$ □

The following lemma shows how we can split up the sets $\{0, \ldots, 2 \cdot n\}$ and $\{0, \ldots, 2 \cdot n - 1\}$ in disjoint sets of even and odd numbers.

Lemma 17.26. *Let $n \in \mathbb{N}$ then we have the following **disjoint** unions*

1. $\{0, \ldots, 2 \cdot n\} = \{2 \cdot k | k \in \{0, \ldots, n\}\} \bigsqcup \{2 \cdot k + 1 | k \in \{0, \ldots, n-1\}\}$
2. $\{0, \ldots, 2 \cdot n - 1\} = \{2 \cdot k | k \in \{0, \ldots, n-1\}\} \bigsqcup \{2 \cdot k + 1 | k \in \{0, \ldots, n-1\}\}$

Proof.

1. We use induction in the proof, so let

$$S = \left\{n \in \mathbb{N} | \{0, \ldots, 2 \cdot n\} = \{2 \cdot k | k \in \{0, \ldots, n\}\} \bigcup \{2 \cdot k + 1 | k \in \{0, \ldots, n-1\}\}\right\}$$

then we have:

$1 \in S$. As $\{0, \ldots, 2 \cdot 1\} = \{0, 1, 2\} = \{0, 2\} \bigcup \{1\} = \{2 \cdot k | k \in \{0, 1\}\} \bigcup \{2 \cdot k + 1 | k \in \{0\}\}$ we have that $1 \in S$

$n \in S \Rightarrow n + 1 \in S$. We have

$$\begin{array}{rl} \{0, \ldots, 2 \cdot (n+1)\} & = \\ \{0, \ldots, 2 \cdot n\} \bigcup \{2 \cdot n + 1, 2 \cdot (n+1)\} & \underset{n \in S}{=} \\ \{2 \cdot k | k \in \{0, \ldots, n\}\} \bigcup \{2 \cdot k + 1 | k \in \{0, \ldots, n-1\}\} \bigcup \{2 \cdot n + 1, 2 \cdot (n+1)\} & = \\ \{2 \cdot k | k \in \{0, \ldots, n+1\}\} \bigcup \{2 \cdot k + 1 | k \in \{0, \ldots, n\}\} & = \\ \{2 \cdot k | k \in \{0, \ldots, n+1\}\} \bigcup \{2 \cdot k + 1 | k \in \{0, \ldots, (n+1) - 1\}\} & \end{array}$$

proving that $n + 1 \in S$.

Mathematical induction proves then that $\forall n \in \mathbb{N}$ we have

$$\{0, \ldots, 2 \cdot n\} = \{2 \cdot k | k \in \{0, \ldots, n\}\} \bigcup \{2 \cdot k + 1 | k \in \{0, \ldots, n-1\}\}$$

Further if $m \in \{2 \cdot k | k \in \{0, \ldots, n\}\} \bigcap \{2 \cdot k + 1 | k \in \{0, \ldots, n-1\}\}$ then there exist a $k \in \{0, \ldots, n\}$ and a $l \in \{0, \ldots. n-1\}$ such that $2 \cdot k = m = 2 \cdot l + 1$ so that by [theorem: 7.52] m is even and odd [= not even] leading to a contradiction. Hence we must have that

$$\{2 \cdot k | k \in \{0, \ldots, n\}\} \bigcap \{2 \cdot k + 1 | k \in \{0, \ldots, n-1\}\} = \varnothing$$

so that

$$\{0, \ldots, 2 \cdot n\} = \{2 \cdot k | k \in \{0, \ldots, n\}\} \bigsqcup \{2 \cdot k + 1 | k \in \{0, \ldots, n-1\}\}$$

2. We use induction in the proof, so let

$$S = \left\{n \in \mathbb{N} | \{0, \ldots, 2 \cdot n - 1\} = \{2 \cdot k | k \in \{0, \ldots, n-1\}\} \bigcup \{2 \cdot k + 1 | k \in \{0, \ldots, n-1\}\}\right\}$$

then we have:

$1 \in S$. As $\{0, \ldots, 2 \cdot 1 - 1\} = \{0, 1\} = \{0\} \bigcup \{1\} = \{2 \cdot k | k \in \{0\}\} \bigcup \{2 \cdot k + 1 | k \in \{0\}\}$ we have that $1 \in S$.

$n \in S \Rightarrow n + 1 \in S$. We have

$$\begin{array}{rl} \{0, \ldots, 2 \cdot (n+1) - 1\} & = \\ \{0, \ldots, 2 \cdot n + 1\} & = \\ \{0, \ldots, 2 \cdot n - 1\} \bigcup \{2 \cdot n, 2 \cdot n + 1\} & \underset{n \in S}{=} \\ \{2 \cdot k | k \in \{0, \ldots, n-1\}\} \bigcup \{2 \cdot k + 1 | k \in \{0, \ldots, n-1\}\} \bigcup \{2 \cdot n, 2 \cdot n + 1\} & = \\ \{2 \cdot k | k \in \{0, \ldots, n\}\} \bigcup \{2 \cdot k + 1 | k \in \{0, \ldots, n\}\} & = \\ \{2 \cdot k | k \in \{0, \ldots, (n+1) - 1\}\} \bigcup \{2 \cdot k + 1 | k \in \{0, \ldots, (n+1) - 1\}\} & \end{array}$$

so that $n+1\in S$.

Mathematical induction proves then that $\forall n\in\mathbb{N}$ we have

$$\{0,\ldots,2\cdot n-1\}=\{2\cdot k|k\in\{0,\ldots,n-1\}\}\bigcup\{2\cdot k+1|k\in\{0,\ldots,n-1\}\}$$

Further if $m\in\{2\cdot k|k\in\{0,\ldots,n-1\}\}\bigcap\{2\cdot k+1|k\in\{0,\ldots,n-1\}\}$ then there exist a $k\in\{0,\ldots,n-1\}$ and a $l\in\{0,\ldots.n-1\}$ such that $2\cdot k=m=2\cdot l+1$ so that by [theorem: 7.52] m is even and odd [= not even] leading to a contradiction. Hence we must have that

$$\{2\cdot k|k\in\{0,\ldots,n-1\}\}\bigcap\{2\cdot k+1|k\in\{0,\ldots,n-1\}\}=\varnothing$$

so that

$$\{0,\ldots,2\cdot n-1\}=\{2\cdot k|k\in\{0,\ldots,n-1\}\}\bigsqcup\{2\cdot k+1|k\in\{0,\ldots,n-1\}\}$$

□

Lemma 17.27. *Let $n\in\mathbb{N}$ then if we define*

$$A=\{(k,l)\in\{0,\ldots,n\}^2|k<l\}\text{ and for }m\in\{1,\ldots,2\cdot n-1\}\;B_m=\{(k,l)\in A|k+l=m\}\subseteq A$$

we have that

$$A=\bigsqcup_{m\in\{1,\ldots,2\cdot n-1\}}B_m$$

Proof. As $\forall m\in\{1,\ldots,2\cdot n-1\}$ $B_m\subseteq A$ it follows that

$$\bigcup_{m\in\{1,\ldots,2\cdot n-1\}}B_m\subseteq A \tag{17.19}$$

Further if $(k,l)\in A$ then $0\leqslant k<l\leqslant n$ so that $k+l<l+l\leqslant 2\cdot n$ hence $k+l\leqslant 2\cdot n-1$. Further as $0<l$ we have $0<k+l$ or $1\leqslant k+l$. Hence if we take $m=k+l$ we have $m\in\{1,\ldots,2\cdot n-1\}$ proving that $(k,l)\in B_m\subseteq\bigcup_{m\in\{1,\ldots,2\cdot n-1\}}B_m$ or that $A\subseteq\bigcup_{m\in\{1,\ldots,2\cdot n-1\}}B_m$. Combining this with [eq: 17.19] gives us

$$A=\bigcup_{m\in\{1,\ldots,2\cdot n-1\}}B_m$$

Let $m,m'\in\{1,\ldots,2\cdot n-1\}$ with $m\neq m'$ and assume that $(k,l)\in B_m\bigcap B_{m'}$ then $m=k+l=m'$ contradicting $m=m'$, so we must have that $B_m\bigcap B_{m'}=\varnothing$. Hence

$$A=\bigsqcup_{m\in\{1,\ldots,2\cdot n-1\}}B_m$$

□

Lemma 17.28. *Let $n\in\mathbb{N}$ then for $k=2\cdot n$ and $\zeta=\left(1+\frac{i}{k}\right)^2$ we have that $\mathrm{Re}(\zeta^k)<0<\mathrm{Img}(\zeta^k)$.*

Proof. We have

$$\begin{array}{rl}\zeta^k & \underset{\text{[theorem: 10.53}}{=}\\ \left(1+\dfrac{i}{k}\right)^{2\cdot k} & \underset{\text{[theorem: 17.23]}}{=}\\ \displaystyle\sum_{l=0}^{2\cdot k}\left(\binom{2\cdot k}{l}\cdot\left(\frac{i}{k}\right)^l\cdot 1^{2\cdot k-l}\right) & =\end{array}$$

$$\begin{aligned}
\sum_{l=0}^{2\cdot k}\left(\binom{2\cdot k}{l}\cdot\left(\frac{i}{k}\right)^{l}\right) &= \\
\sum_{l\in\{0,\ldots,2\cdot k\}}\left(\binom{2\cdot k}{l}\cdot\left(\frac{i}{k}\right)^{l}\right) &\underset{\text{[lemma: 17.26]}}{=} \\
\sum_{l\in\{2\cdot j|j\in\{0,\ldots,k\}\}\bigsqcup\{2\cdot j+1|j\in\{1,\ldots,k-1\}\}}\left(\binom{2\cdot k}{l}\cdot\left(\frac{i}{k}\right)^{l}\right) &\underset{\text{[theorem: 11.43]}}{=} \\
\sum_{l\in\{2\cdot j|j\in\{0,\ldots,k\}\}}\left(\binom{2\cdot k}{l}\cdot\left(\frac{i}{k}\right)^{l}\right)+\sum_{l\in\{2\cdot j+1|j\in\{1,\ldots,k-1\}\}}\left(\binom{2\cdot k}{l}\cdot\left(\frac{i}{k}\right)^{l}\right) &\underset{\text{[theorem: 11.36]}}{=} \\
\sum_{l\in\{0,\ldots,k\}}\left(\binom{2\cdot k}{2\cdot l}\cdot\left(\frac{i}{k}\right)^{2\cdot l}\right)+\sum_{l\in\{1,\ldots,k-1\}}\left(\binom{2\cdot k}{2\cdot l+1}\cdot\left(\frac{i}{k}\right)^{2\cdot l+1}\right) &
\end{aligned}$$

proving that

$$\zeta^{k}=A+B \tag{17.20}$$

where

$$A = \sum_{l\in\{0,\ldots,k\}}\left(\binom{2\cdot k}{2\cdot l}\cdot\left(\frac{i}{k}\right)^{2\cdot l}\right) \tag{17.21}$$

$$B = \sum_{l\in\{1,\ldots,k-1\}}\left(\binom{2\cdot k}{2\cdot l+1}\cdot\left(\frac{i}{k}\right)^{2\cdot l+1}\right) \tag{17.22}$$

Now

$$\begin{aligned}
A &\underset{\text{[eq: 17.21]}}{=} \\
\sum_{l\in\{0,\ldots,k\}}\left(\binom{2\cdot k}{2\cdot l}\cdot\left(\frac{i}{k}\right)^{2\cdot l}\right) &\underset{k=2\cdot n}{=} \\
\sum_{l\in\{0,\ldots,2\cdot n\}}\left(\binom{2\cdot k}{2\cdot l}\cdot\left(\frac{i}{k}\right)^{2\cdot l}\right) &\underset{\text{[lemma: 17.26]}}{=} \\
\sum_{l\in\{2\cdot j|j\in\{0,\ldots,n\}\}\bigsqcup\{2\cdot j+1|j\in\{0,\ldots,n-1\}\}}\left(\binom{2\cdot k}{2\cdot l}\cdot\left(\frac{i}{k}\right)^{2\cdot l}\right) &\underset{\text{[theorem: 11.43]}}{=} \\
\sum_{l\in\{2\cdot j|j\in\{0,\ldots,n\}\}}\left(\binom{2\cdot k}{2\cdot l}\cdot\left(\frac{i}{k}\right)^{2\cdot l}\right)+\sum_{l\in\{2\cdot j+1|j\in\{0,\ldots,n-1\}\}}\left(\binom{2\cdot k}{2\cdot l}\cdot\left(\frac{i}{k}\right)^{2\cdot l}\right) &\underset{\text{[theorem: 11.36]}}{=} \\
\sum_{l\in\{0,\ldots,n\}}\left(\binom{2\cdot k}{4\cdot l}\cdot\frac{i^{4\cdot l}}{k^{4\cdot l}}\right)+\sum_{l\in\{0,\ldots,n-1\}}\left(\binom{2\cdot k}{4\cdot l+2}\cdot\frac{i^{4\cdot l+2}}{k^{4\cdot l+2}}\right) &\underset{\text{[theorem: 10.79]}}{=} \\
\sum_{l\in\{0,\ldots,n\}}\left(\binom{2\cdot k}{4\cdot l}\cdot\frac{1}{k^{4\cdot l}}\right)-\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) &= \\
\binom{2\cdot k}{4\cdot 0}\cdot\frac{1}{k^{4\cdot 0}}+\binom{2\cdot k}{4\cdot 1}\cdot\frac{1}{k^{4\cdot 1}}+\sum_{l\in\{2,\ldots,n\}}\left(\binom{2\cdot k}{4\cdot l}\cdot\frac{1}{k^{4\cdot l}}\right)-\sum_{l\in\{2,\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}{4}\cdot\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)-\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) &= \\
1+\binom{2\cdot k}{4}\cdot\frac{1}{k^{4}}-\binom{2\cdot k}{4\cdot 0+2}\cdot\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)-\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}{4}\cdot\frac{1}{k^{4}}-\binom{2\cdot k}{2}\cdot\frac{1}{k^{2}}+\sum_{l\in\{2,\ldots,n\}}\left(\binom{2\cdot k}{4\cdot l}\cdot\frac{1}{k^{4\cdot l}}\right)-\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) &
\end{aligned}$$

So that

$$A=1+\binom{2\cdot k}{4}\cdot\frac{1}{k^{4}}-\binom{2\cdot k}{2}\cdot\frac{1}{k^{2}}+C-D \tag{17.23}$$

where

$$C = \sum_{l\in\{2,\ldots,n\}} \left(\binom{2\cdot k}{4\cdot l}\cdot\frac{1}{k^{4\cdot l}}\right) \tag{17.24}$$

$$D = \sum_{l\in\{1,\ldots,n-1\}} \left(\binom{2\cdot k}{4\cdot l+2}\cdot\frac{1}{k^{4\cdot l+2}}\right) \tag{17.25}$$

Now

$$\begin{aligned} C &= \sum_{l\in\{2,\ldots,n\}} \left(\binom{2\cdot k}{4\cdot l}\cdot\frac{1}{k^{4\cdot l}}\right) \\ &= \sum_{l\in\{1,\ldots,n-1\}} \left(\binom{2\cdot k}{4\cdot(l+1)}\cdot\frac{1}{k^{4\cdot(l+1)}}\right) \\ &= \sum_{l\in\{1,\ldots,n-1\}} \left(\binom{2\cdot k}{2\cdot(2\cdot l+1)+2}\cdot\frac{1}{k^{2\cdot(2\cdot l+1)+2}}\right) \\ &\underset{[\text{theorem: } 11.36]}{=} \sum_{l\in\{2\cdot j+1|j\in\{1,\ldots,n-1\}\}} \left(\binom{2\cdot k}{2\cdot l+2}\cdot\frac{1}{k^{2\cdot l+2}}\right) \end{aligned}$$

so that

$$C = \sum_{l\in\{2\cdot j+1|j\in\{1,\ldots,n-1\}\}} \left(\binom{2\cdot k}{2\cdot l+2}\cdot\frac{1}{k^{2\cdot l+2}}\right) \tag{17.26}$$

Further

$$\begin{aligned} D &= \sum_{l\in\{1,\ldots,n-1\}} \left(\binom{2\cdot k}{4\cdot l+2}\cdot\frac{1}{k^{4\cdot l+2}}\right) \\ &= \sum_{l\in\{1,\ldots,n-1\}} \left(\binom{2\cdot k}{2\cdot(2\cdot l+1)}\cdot\frac{1}{k^{2\cdot(2\cdot l+1)}}\right) \\ &\underset{[\text{theorem: } 11.36]}{=} \sum_{l\in\{2\cdot j+1|1,\ldots,n-1\}} \left(\binom{2\cdot k}{2\cdot l}\cdot\frac{1}{k^{2\cdot l}}\right) \end{aligned}$$

so that

$$D = \sum_{l\in\{2\cdot j+1|1,\ldots,n-1\}} \left(\binom{2\cdot k}{2\cdot l}\cdot\frac{1}{k^{2\cdot l}}\right) \tag{17.27}$$

Combining [eqs: 17.26, 17.27] gives us

$$\begin{aligned} C-D &= \sum_{l\in\{2\cdot j+1|j\in\{1,\ldots,n-1\}\}} \left(\binom{2\cdot k}{2\cdot l+2}\cdot\frac{1}{k^{2\cdot l+2}}\right) - \sum_{l\in\{2\cdot j+1|1,\ldots,n-1\}} \left(\binom{2\cdot k}{2\cdot l}\cdot\frac{1}{k^{2\cdot l}}\right) \\ &= \sum_{l\in\{2\cdot j+1|j\in\{1,\ldots,n-1\}\}} \left(\binom{2\cdot k}{2\cdot l+2}\cdot\frac{1}{k^{2\cdot l+2}} - \binom{2\cdot k}{2\cdot l}\cdot\frac{1}{k^{2\cdot l}}\right) \end{aligned}$$

substituting in [eq: 17.23] gives

$$A = 1+\binom{2\cdot k}{4}\cdot\frac{1}{k^4}-\binom{2\cdot k}{2}\cdot\frac{1}{k^2}+\sum_{l\in\{2\cdot j+1|j\in\{1,\ldots,n-1\}\}} \left(\binom{2\cdot k}{2\cdot l+2}\cdot\frac{1}{k^{2\cdot l+2}} - \binom{2\cdot k}{2\cdot l}\cdot\frac{1}{k^{2\cdot l}}\right) \tag{17.28}$$

and as A does not containg i anymore we have also

$$A\in\mathbb{R} \tag{17.29}$$

Let's calculate B:

$$\begin{aligned} B &\underset{[\text{eq: } 17.22]}{=} \\ \sum_{l\in\{1,\ldots,k-1\}} \left(\binom{2\cdot k}{2\cdot l+1}\cdot\left(\frac{i}{k}\right)^{2\cdot l+1}\right) &\underset{[\text{lemma: } 17.26]}{=} \\ \sum_{l\in\{2\cdot j|j\in\{1,\ldots,k-1\}\}\bigsqcup\{2\cdot j+1|j\in\{1,\ldots,k-1\}\}} \left(\binom{2\cdot k}{2\cdot l+1}\cdot\left(\frac{i}{k}\right)^{2\cdot l+1}\right) &\underset{[\text{theorem: } 11.43]}{=} \end{aligned}$$

$$
\begin{array}{rl}
\sum_{l\in\{2\cdot j|j\in\{1,\ldots,k-1\}\}}\left(\binom{2\cdot k}{2\cdot l+1}\cdot\left(\frac{i}{k}\right)^{2\cdot l+1}\right)+\sum_{l\in\{2\cdot j+1|j\in\{1,\ldots,k-1\}\}}\left(\binom{2\cdot k}{2\cdot l+1}\cdot\left(\frac{i}{k}\right)^{2\cdot l+1}\right) & \underset{[\text{theorem: } 11.36]}{=}\\
\sum_{l\in\{1,\ldots,k-1\}}\left(\binom{2\cdot k}{2\cdot(2\cdot l)+1}\cdot\left(\frac{i}{k}\right)^{2\cdot(2\cdot l)+1}\right)+\sum_{l\in\{2\cdot j+1|j\in\{1,\ldots,k-1\}\}}\left(\binom{2\cdot k}{2\cdot l+1}\cdot\left(\frac{i}{k}\right)^{2\cdot l+1}\right) & \underset{[\text{theorem: } 11.36]}{=}\\
\sum_{l\in\{1,\ldots,k-1\}}\left(\binom{2\cdot k}{2\cdot(2\cdot l)+1}\cdot\left(\frac{i}{k}\right)^{2\cdot(2\cdot l)+1}\right)+\sum_{l\in\{1,\ldots,k-1\}}\left(\binom{2\cdot k}{2\cdot(2\cdot l+1)+1}\cdot\left(\frac{i}{k}\right)^{2\cdot(2\cdot l+1)+1}\right) & =\\
\sum_{l\in\{1,\ldots,k-1\}}\left(\binom{2\cdot k}{4\cdot l+1}\cdot\left(\frac{i}{k}\right)^{4\cdot l+1}\right)+\sum_{l\in\{1,\ldots,k-1\}}\left(\binom{2\cdot k}{4\cdot l+3}\cdot\left(\frac{i}{k}\right)^{4\cdot l+3}\right) & =\\
\sum_{l\in\{1,\ldots,k-1\}}\left(\binom{2\cdot k}{4\cdot l+1}\cdot\frac{i^{4\cdot l+1}}{k^{4\cdot l+1}}+\sum_{l\in\{1,\ldots,k-1\}}\left(\binom{2\cdot k}{4\cdot l+3}\cdot\frac{i^{4\cdot l+3}}{k^{4\cdot l+3}}\right)\right. & \underset{[\text{theorem: } 10.79]}{=}\\
\sum_{l\in\{1,\ldots,k-1\}}\left(\binom{2\cdot k}{4\cdot l+1}\cdot\frac{i}{k^{4\cdot l+1}}-\sum_{l\in\{1,\ldots,k-1\}}\left(\binom{2\cdot k}{4\cdot l+3}\cdot\frac{i}{k^{4\cdot l+3}}\right)\right. & \underset{[\text{theorem: } 10.79]}{=}\\
i\cdot\left(\sum_{l\in\{1,\ldots,k-1\}}\left(\binom{2\cdot k}{4\cdot l+1}\cdot\frac{1}{k^{4\cdot l+1}}-\sum_{l\in\{1,\ldots,k-1\}}\left(\binom{2\cdot k}{4\cdot l+3}\cdot\frac{1}{k^{4\cdot l+3}}\right)\right)\right. & =\\
i\cdot\sum_{l\in\{1,\ldots,k-1\}}\left(\binom{2\cdot k}{4\cdot l+1}\cdot\frac{1}{k^{4\cdot l+1}}-\binom{2\cdot k}{4\cdot l+3}\cdot\frac{1}{k^{4\cdot l+3}}\right) & =\\
i\cdot\sum_{l\in\{1,\ldots,k-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) & \underset{[\text{theorem: } 11.36]}{=}\\
i\cdot\sum_{l\in\{2\cdot j+1|j\in\{1,\ldots,k-1\}\}}\left(\binom{2\cdot k}{2\cdot l-1}\cdot\frac{1}{k^{2\cdot l-1}}-\binom{2\cdot k}{2\cdot l+1}\cdot\frac{1}{k^{2\cdot l+1}}\right) &
\end{array}
$$

proving that

$$B=i\cdot B'$$

where

$$B'=\sum_{l\in\{2\cdot j+1|j\in\{1,\ldots,k-1\}\}}\left(\binom{2\cdot k}{2\cdot l-1}\cdot\frac{1}{k^{2\cdot l-1}}-\binom{2\cdot k}{2\cdot l+1}\cdot\frac{1}{k^{2\cdot l+1}}\right)\in\mathbb{R} \tag{17.30}$$

Hence as $\zeta^k=A+B=A+i\cdot B$ where $A,B\in\mathbb{R}$ [see eqs: 17.29, 17.30] it follows that

$$\operatorname{Re}(\zeta^k)=A \text{ and } \operatorname{Img}(\zeta^k)=B' \tag{17.31}$$

Now

$$
\begin{aligned}
\binom{2\cdot k}{2} &= \frac{(2\cdot k)!}{2!\cdot(2\cdot k-2)!}\\
&= \frac{(2\cdot k)\cdot(2\cdot k-1)\cdot(2\cdot k-2)!}{2\cdot(2\cdot k-2)!}\\
&= \frac{2\cdot k\cdot(2\cdot k-1)}{2}\\
&= 2\cdot k^2-k
\end{aligned} \tag{17.32}
$$

Now $2\leqslant 2\cdot n=k$ so that $4\leqslant 2\cdot k$ hence $0<2\cdot k-3, 2\cdot k-2$ which we use in the following

$$
\begin{aligned}
\binom{2\cdot k}{4} &= \frac{(2\cdot k)!}{4!\cdot(2\cdot k-4)!}\\
&= \frac{(2\cdot k)!}{4\cdot 3\cdot 2!\cdot(2\cdot k-4)!}\\
&\underset{0<2\cdot k-3,2\cdot k-2}{=} \frac{(2\cdot k)!\cdot(2\cdot k-2)\cdot(2\cdot k-3)}{4\cdot 3\cdot 2!\cdot(2\cdot k-4)!\cdot(2\cdot k-2)\cdot(2\cdot k-3)}
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{(2\cdot k)!\cdot(2\cdot k-2)\cdot(2\cdot k-3)}{4\cdot 3\cdot 2!\cdot(2\cdot k-2)\cdot(2\cdot k-3)\cdot(2\cdot k-4)!}\\
&= \frac{(2\cdot k)!\cdot(2\cdot k-2)\cdot(2\cdot k-3)}{4\cdot 3\cdot 2!\cdot(2\cdot k-2)!}\\
&= \frac{(2\cdot k)!\cdot(2\cdot k-2)\cdot(2\cdot k-3)}{2!\cdot(2\cdot k-2)!\cdot 4\cdot 3}\\
&= \binom{2\cdot k}{2}\cdot\frac{(2\cdot k-2)\cdot(2\cdot k-3)}{4\cdot 3}\\
&= \binom{2\cdot k}{2}\cdot\frac{4\cdot k^2-6\cdot k-4\cdot k+6}{12}\\
&= \binom{2\cdot k}{2}\cdot\frac{4\cdot k^2-10\cdot k+6}{12}\\
&= \binom{2\cdot k}{2}\cdot\frac{2\cdot k^2-5\cdot k+3}{6}
\end{aligned}
\tag{17.33}
$$

Hence

$$
\begin{aligned}
1+\binom{2\cdot k}{4}\cdot\frac{1}{k^4}-\binom{2\cdot k}{2}\cdot\frac{1}{k^2} &\underset{\text{[eq: 17.33]}}{=} 1+\binom{2\cdot k}{2}\cdot\frac{2\cdot k^2-5\cdot k+3}{6}\cdot\frac{1}{k^4}-\binom{2\cdot k}{2}\cdot\frac{1}{k^2}\\
&= 1+\binom{2\cdot k}{2}\cdot\left(\frac{2\cdot k^2-5\cdot k+3}{6}\cdot\frac{1}{k^4}-\frac{1}{k^2}\right)\\
&= 1+\binom{2\cdot k}{2}\cdot\frac{1}{k^2}\cdot\left(\frac{2\cdot k^2-5\cdot k+3}{6}\cdot\frac{1}{k^2}-1\right)\\
&= 1+\binom{2\cdot k}{2}\cdot\frac{1}{k^2}\cdot\left(\frac{2\cdot k^2}{6\cdot k^2}+\frac{-5\cdot k+3}{6}\cdot\frac{1}{k^2}-1\right)\\
&= 1+\binom{2\cdot k}{2}\cdot\frac{1}{k^2}\cdot\left(\frac{1}{3}+\frac{-5\cdot k+3}{6\cdot k^2}-1\right)\\
&= 1+\binom{2\cdot k}{2}\cdot\frac{1}{k^2}\cdot\left(-\frac{2}{3}+\frac{-5\cdot k+3}{6\cdot k^2}\right)\\
&= 1-\binom{2\cdot k}{2}\cdot\frac{1}{k^2}\cdot\left(\frac{2}{3}+\frac{5\cdot k-3}{6\cdot k^2}\right)\\
&\underset{\text{[eq: 17.32}}{=} 1-(2\cdot k^2-k)\cdot\frac{1}{k^2}\cdot\left(\frac{2}{3}+\frac{5\cdot k-3}{6\cdot k^2}\right)\\
&= 1-\left(2-\frac{1}{k}\right)\cdot\left(\frac{2}{3}+\frac{5\cdot k-3}{6\cdot k^2}\right)
\end{aligned}
$$

hence

$$1+\binom{2\cdot k}{4}\cdot\frac{1}{k^4}-\binom{2\cdot k}{2}\cdot\frac{1}{k^2}=1-\left(2-\frac{1}{k}\right)\cdot\left(\frac{2}{3}+\frac{5\cdot k-3}{6\cdot k^2}\right)\tag{17.34}$$

As $0<2\leqslant 2\cdot n\leqslant k$ we have $\frac{1}{k}<\frac{1}{2}$ or $-\frac{1}{2}<\frac{-1}{k}$ hence

$$\frac{3}{2}=2-\frac{1}{2}<2-\frac{1}{k}\tag{17.35}$$

Further as $2\leqslant k$ it follows that $10\leqslant 5\cdot k$ so that $0<7=10-3\leqslant 5\cdot k-3$ so that

$$0<5\cdot k-3\tag{17.36}$$

So

$$0<5\cdot k-3<\frac{2}{3}+\frac{5\cdot k-3}{6\cdot k^2}\tag{17.37}$$

further from [eq: 17.35] it follows that $-(5\cdot k-3)<0$ hence

$$-\frac{3}{2}\cdot\frac{1}{6\cdot k^2}\cdot(5\cdot k-3)<0\tag{17.38}$$

Multiplying both sides of [eq: 17.35] by $\frac{2}{3}+\frac{5\cdot k-3}{6\cdot k^2}$ we have as $0<_{[\text{eq: }17.37]}\frac{2}{3}+\frac{5\cdot k-3}{6\cdot k^2}$ gives

$$\frac{3}{2}\cdot\left(\frac{2}{3}+\frac{5\cdot k-3}{6\cdot k^2}\right)<\left(2-\frac{1}{k}\right)\cdot\left(\frac{2}{3}+\frac{5\cdot k-3}{6\cdot k^2}\right)$$

or

$$-\left(2-\frac{1}{k}\right)\cdot\left(\frac{2}{3}+\frac{5\cdot k-3}{6\cdot k^2}\right)<-\frac{3}{2}\cdot\left(\frac{2}{3}+\frac{5\cdot k-3}{6\cdot k^2}\right)$$

hence

$$\begin{aligned}1-\left(2-\frac{1}{k}\right)\cdot\left(\frac{2}{3}+\frac{5\cdot k-3}{6\cdot k^2}\right) &= 1-\frac{3}{2}\cdot\left(\frac{2}{3}+\frac{5\cdot k-3}{6\cdot k^2}\right)\\ &= 1-\frac{3}{2}\cdot\frac{2}{3}-\frac{3}{2}\cdot\frac{5\cdot k-3}{6\cdot k^2}\\ &= 1-1-\frac{3}{2}\cdot\frac{5\cdot k-3}{6\cdot k^2}\\ &= -\frac{3}{2}\cdot\frac{5\cdot k-3}{6\cdot k^2}\\ &<_{[\text{eq: }17.38]} 0\end{aligned}$$

This proves that $1-\left(2-\frac{1}{k}\right)\cdot\left(\frac{2}{3}+\frac{5\cdot k-3}{6\cdot k^2}\right)<0$ and by [eq: 17.34]

$$1+\binom{2\cdot k}{4}\cdot\frac{1}{k^4}-\binom{2\cdot k}{2}\cdot\frac{1}{k^2}<0 \tag{17.39}$$

Further for $l\in\{2\cdot j+1|j\in\{1,\ldots,n-1\}\}$ we have

$$\begin{aligned}&\binom{2\cdot k}{2\cdot l+2}\cdot\frac{1}{k^{2\cdot l+2}}-\binom{2\cdot k}{2\cdot l}\cdot\frac{1}{k^{2\cdot l}} =\\ &\frac{1}{k^{2\cdot l}}\cdot\left(\binom{2\cdot k}{2\cdot l+2}\cdot\frac{1}{k^2}-\binom{2\cdot k}{2\cdot l}\right) =\\ &\frac{1}{k^{2\cdot l}}\cdot\left(\frac{(2\cdot k)!}{(2\cdot l+2)!\cdot(2\cdot k-(2\cdot l+2))!}\cdot\frac{1}{k^2}-\frac{(2\cdot k)!}{(2\cdot l)!\cdot(2\cdot k-2\cdot l)!}\right) =\\ &\frac{(2\cdot k)!}{k^{2\cdot l}}\cdot\left(\frac{1}{(2\cdot l+2)!\cdot(2\cdot k-2\cdot l-2)!\cdot k^2}-\frac{1}{(2\cdot l)!\cdot(2\cdot k-2\cdot l)!}\right) =\\ &\frac{(2\cdot k)!}{k^{2\cdot l}}\cdot\left(\frac{1}{(2\cdot l+2)\cdot(2\cdot l+1)\cdot(2\cdot l)!\cdot(2\cdot k-2\cdot l-2)!\cdot k^2}-\frac{1}{(2\cdot l)!\cdot(2\cdot k-2\cdot l)!}\right) =\\ &\frac{(2\cdot k)!}{k^{2\cdot l}\cdot(2\cdot l)!}\cdot\left(\frac{1}{(2\cdot l+2)\cdot(2\cdot l+1)\cdot(2\cdot k-2\cdot l-2)!\cdot k^2}-\frac{1}{(2\cdot k-2\cdot l)!}\right) =\\ &\frac{(2\cdot k)!}{k^{2\cdot l}\cdot(2\cdot l)!}\cdot\left(\frac{1}{(2\cdot l+2)\cdot(2\cdot l+1)\cdot(2\cdot k-2\cdot l-2)!\cdot k^2}-\frac{1}{(2\cdot k-2\cdot l)\cdot(2\cdot k-2\cdot l-1)\cdot(2\cdot k-2\cdot l-2)!}\right) =\\ &\frac{(2\cdot k)!}{k^{2\cdot l}\cdot(2\cdot l)!\cdot(2\cdot k-2\cdot k-2)!}\cdot\left(\frac{1}{(2\cdot l+2)\cdot(2\cdot l+1)\cdot k^2}-\frac{1}{(2\cdot k-2\cdot l)\cdot(2\cdot k-2\cdot l-1)}\right) =\\ &\frac{(2\cdot k)!}{2\cdot k^{2\cdot l}\cdot(2\cdot l)!\cdot(2\cdot k-2\cdot k-2)!}\cdot\left(\frac{1}{(l+1)\cdot(2\cdot l+1)\cdot k^2}-\frac{1}{(k-l)\cdot(2\cdot k-2\cdot l-1)}\right) =\\ &\frac{(2\cdot k)!}{2\cdot k^{2\cdot l}\cdot(2\cdot l)!\cdot(2\cdot k-2\cdot k-2)!}\cdot\frac{(k-l)\cdot(2\cdot k-2\cdot l-1)-(l+1)\cdot(2\cdot l+1)\cdot k^2}{(l+1)\cdot(2\cdot l+1)\cdot k^2\cdot(k-l)\cdot(2\cdot k-2\cdot l-1)}\end{aligned}$$

proving that

$$\begin{aligned}&\binom{2\cdot k}{2\cdot l+2}\cdot\frac{1}{k^{2\cdot l+2}}-\binom{2\cdot k}{2\cdot l}\cdot\frac{1}{k^{2\cdot l}} =\\ &\frac{(2\cdot k)!}{2\cdot k^{2\cdot l}\cdot(2\cdot l)!\cdot(2\cdot k-2\cdot k-2)!}\cdot\frac{(k-l)\cdot(2\cdot k-2\cdot l-1)-(l+1)\cdot(2\cdot l+1)\cdot k^2}{(l+1)\cdot(2\cdot l+1)\cdot k^2\cdot(k-l)\cdot(2\cdot k-2\cdot l-1)}\end{aligned} \tag{17.40}$$

As $l\in\{2\cdot j+1|j\in\{1,\ldots,n-1\}\}$ we have that $3\leqslant l\leqslant 2\cdot(n-1)+1=2\cdot n-1=k-1<k$ so that

$$0<l<k \text{ and by [theorem: 10.50] } l^2<k^2 \tag{17.41}$$

hence

$$\begin{array}{rc}
-((k-l)\cdot(2\cdot k-2\cdot l-1)-(l+1)\cdot(2\cdot l+1)\cdot k^2) & = \\
(l+1)\cdot(2\cdot l+1)\cdot k^2-(k-l)\cdot(2\cdot k-2\cdot l-1) & = \\
2\cdot l^2\cdot k^2+l\cdot k^2+2\cdot l\cdot k^2+k^2-(2\cdot k^2-2\cdot k\cdot l-k-2\cdot l\cdot k+2\cdot l^2+1) & = \\
2\cdot l^2\cdot k^2+l\cdot k^2+2\cdot l\cdot k^2+k^2-2\cdot k^2+2\cdot k\cdot l+k+2\cdot l\cdot k-2\cdot l^2-1 & = \\
2\cdot l^2\cdot k^2+\underbrace{l\cdot k^2+2\cdot l\cdot k^2}_{1}+\underbrace{k^2-2\cdot k^2}_{2}+\underbrace{2\cdot k\cdot l}_{3}+k+\underbrace{2\cdot l\cdot k}_{3}-2\cdot l^2-1 & = \\
2\cdot l^2\cdot k^2+\underbrace{3\cdot l\cdot k^2}_{1}-\underbrace{k^2}_{2}+\underbrace{4\cdot k\cdot l}_{3}+k-2\cdot l^2-1 & = \\
2\cdot l^2\cdot k^2+3\cdot l\cdot k^2-k^2+4\cdot k\cdot l+k-2\cdot l^2-1 & >_{k>l} \\
2\cdot l^2\cdot k^2+3\cdot l\cdot k^2-k^2+\underbrace{4\cdot l^2}_{1}+k-\underbrace{2\cdot l^2}_{1}-1 & = \\
2\cdot l^2\cdot k^2+3\cdot l\cdot k^2-k^2+\underbrace{2\cdot l^2}_{1}+k-1 & = \\
2\cdot l^2\cdot k^2+3\cdot l\cdot k^2-k^2+2\cdot l^2+k-1 & >_{k^2>l^2} \\
2\cdot l^2\cdot k^2+3\cdot l\cdot k^2-\underbrace{l^2}_{2}+\underbrace{2\cdot l^2}_{2}+k-1 & = \\
2\cdot l^2\cdot k^2+3\cdot l\cdot k^2+\underbrace{l^2}_{2}+k-1 & = \\
2\cdot l^2\cdot k^2+3\cdot l\cdot k^2+l^2+k-1 & >_{2\leqslant 2\cdot n=k\Rightarrow k>1} \\
2\cdot l^2\cdot k^2+3\cdot l\cdot k^2+l^2+1-1 & = \\
2\cdot l^2\cdot k^2+3\cdot l\cdot k^2+l^2 & > \\
0 &
\end{array}$$

so that $-((k-l)\cdot(2\cdot k-2\cdot l-1)-(l+1)\cdot(2\cdot l+1)\cdot k^2)>0$ or

$$(k-l)\cdot(2\cdot k-2\cdot l-1)-(l+1)\cdot(2\cdot l+1)\cdot k^2<0$$

Using this [eq: 17.40] in together with the fact that $0<\frac{(2\cdot k)!}{2\cdot k^{2\cdot l}\cdot(2\cdot l)!\cdot(2\cdot k-2\cdot k-2)!}$ and $0<(l+1)\cdot(2\cdot l+1)\cdot k^2\cdot(k-l)\cdot(2\cdot k-2\cdot l-1)$ proves that

$$\binom{2\cdot k}{2\cdot l+2}\cdot\frac{1}{k^{2\cdot l+2}}-\binom{2\cdot k}{2\cdot l}\cdot\frac{1}{k^{2\cdot l}}<0 \tag{17.42}$$

Combining now [eqs: 17.28, 17.39, 17.42] proves that

$$\mathrm{Re}(\zeta^k)\underset{[\text{eq: }17.31]}{=}A<0 \tag{17.43}$$

Having estimated the value of real part of ζ we estimate now the imaginary part of ζ^k. Now for $l\in\{2\cdot j+1|j\in\{1,\ldots,k-1\}\}$ we have

$$\begin{array}{rc}
\binom{2\cdot k}{2\cdot l-1}\cdot\frac{1}{k^{2\cdot l-1}}-\binom{2\cdot k}{2\cdot l+1}\cdot\frac{1}{k^{2\cdot l+1}} & = \\
\frac{1}{k^{2\cdot l-1}}\cdot\left(\binom{2\cdot k}{2\cdot l-1}-\binom{2\cdot k}{2\cdot l+1}\cdot\frac{1}{k^2}\right) & = \\
\frac{1}{k^{2\cdot l-1}}\cdot\left(\frac{(2\cdot k)!}{(2\cdot l-1)!\cdot(2\cdot k-(2\cdot l-1))!}-\frac{(2\cdot k)!}{(2\cdot l+1)!\cdot(2\cdot k-(2\cdot l+1))!}\cdot\frac{1}{k^2}\right) & = \\
\frac{1}{k^{2\cdot l-1}}\cdot\left(\frac{(2\cdot k)!}{(2\cdot l-1)!\cdot(2\cdot k-2\cdot l+1)!}-\frac{(2\cdot k)!}{(2\cdot l+1)!\cdot(2\cdot k-2\cdot l-1)!}\cdot\frac{1}{k^2}\right) & = \\
\frac{(2\cdot k)!}{k^{2\cdot l-1}}\cdot\left(\frac{1}{(2\cdot l-1)!\cdot(2\cdot k-2\cdot l+1)!}-\frac{1}{(2\cdot l+1)!\cdot(2\cdot k-2\cdot l-1)!}\cdot\frac{1}{k^2}\right) & = \\
\frac{(2\cdot k)!}{k^{2\cdot l-1}}\cdot\left(\frac{1}{(2\cdot l-1)!\cdot(2\cdot k-2\cdot l+1)!}-\frac{1}{(2\cdot l+1)\cdot(2\cdot l)\cdot(2\cdot l-1)!\cdot(2\cdot k-2\cdot l-1)!}\cdot\frac{1}{k^2}\right) & = \\
\frac{(2\cdot k)!}{k^{2\cdot l-1}\cdot(2\cdot l-1)!}\cdot\left(\frac{1}{(2\cdot k-2\cdot l+1)!}-\frac{1}{(2\cdot l+1)\cdot(2\cdot l)\cdot(2\cdot k-2\cdot l-1)!}\cdot\frac{1}{k^2}\right) & =
\end{array}$$

$$\frac{(2\cdot k)!}{k^{2\cdot l-1}\cdot(2\cdot l-1)!}\cdot\left(\frac{1}{(2\cdot k-2\cdot l+1)\cdot(2\cdot k-2\cdot l)\cdot(2\cdot k-2\cdot l-1)!}-\frac{1}{(2\cdot l+1)\cdot(2\cdot l)\cdot(2\cdot k-2\cdot l-1)!}\cdot\frac{1}{k^2}\right) =$$

$$\frac{(2\cdot k)!}{k^{2\cdot l-1}\cdot(2\cdot l-1)!\cdot(2\cdot k-2\cdot l-1)!}\cdot\left(\frac{1}{(2\cdot k-2\cdot l+1)\cdot(2\cdot k-2\cdot l)}-\frac{1}{(2\cdot l+1)\cdot(2\cdot l)}\cdot\frac{1}{k^2}\right) =$$

$$\frac{(2\cdot k)!}{2\cdot k^{2\cdot l-1}\cdot(2\cdot l-1)!\cdot(2\cdot k-2\cdot l-1)!}\cdot\left(\frac{1}{(2\cdot k-2\cdot l+1)\cdot(k-l)}-\frac{1}{(2\cdot l+1)\cdot l}\cdot\frac{1}{k^2}\right) =$$

$$\frac{(2\cdot k)!}{2\cdot k^{2\cdot l-1}\cdot(2\cdot l-1)!\cdot(2\cdot k-2\cdot l-1)!}\cdot\left(\frac{1}{(2\cdot k-2\cdot l+1)\cdot(k-l)}-\frac{1}{(2\cdot l+1)\cdot l\cdot k^2}\right) =$$

$$\frac{(2\cdot k)!}{2\cdot k^{2\cdot l-1}\cdot(2\cdot l-1)!\cdot(2\cdot k-2\cdot l-1)!}\cdot\left(\frac{(2\cdot l+1)\cdot l\cdot k^2-(2\cdot k-2\cdot l+1)\cdot(k-l)}{(2\cdot k-2\cdot l+1)\cdot(k-l)\cdot(2\cdot l+1)\cdot l\cdot k^2}\right)$$

proving that

$$\binom{2\cdot k}{2\cdot l-1}\cdot\frac{1}{k^{2\cdot l-1}}-\binom{2\cdot k}{2\cdot l+1}\cdot\frac{1}{k^{2\cdot l+1}} = \frac{(2\cdot k)!}{2\cdot k^{2\cdot l-1}\cdot(2\cdot l-1)!\cdot(2\cdot k-2\cdot l-1)!}\cdot\left(\frac{(2\cdot l+1)\cdot l\cdot k^2-(2\cdot k-2\cdot l+1)\cdot(k-l)}{(2\cdot k-2\cdot l+1)\cdot(k-l)\cdot(2\cdot l+1)\cdot l\cdot k^2}\right) \tag{17.44}$$

We have for $l\in\{2\cdot j+1|j\in\{1,\ldots,k-1\}\}$ that $3=2\cdot 1+1\leqslant l<2\cdot(k-1)+1=2\cdot k-1<2\cdot k$ so that

$$1<l<2\cdot k \text{ and by [theorem: 10.50] that } 1<l^2$$

further as $2\leqslant 2\cdot n=k$ we have $1<k$ hence

$$0<k-1$$

$$\begin{array}{rl}
(2\cdot l+1)\cdot l\cdot k^2-(2\cdot k-2\cdot l+1)\cdot(k-l) & = \\
2\cdot l^2\cdot k^2+l\cdot k^2-2\cdot k^2+\underbrace{2\cdot k\cdot l+2\cdot k\cdot l}_{1}-2\cdot l^2-k+l & = \\
2\cdot l^2\cdot k^2+l\cdot k^2-2\cdot k^2+\underbrace{4\cdot k\cdot l}_{1}-2\cdot l^2-k+l & = \\
2\cdot l^2\cdot k^2+l\cdot k^2-2\cdot k^2+4\cdot k\cdot l-2\cdot l^2-k+l & >_{2\cdot k>l\Rightarrow 4\cdot k>2\cdot l} \\
2\cdot l^2\cdot k^2+l\cdot k^2-2\cdot k^2+2\cdot l^2-2\cdot l^2-k+l & = \\
2\cdot l^2\cdot k^2+l\cdot k^2-2\cdot k^2-k+l & = \\
2\cdot l^2\cdot k^2+l\cdot k^2-2\cdot k^2-k+l & >_{l^2>1} \\
2\cdot k^2+l\cdot k^2-2\cdot k^2-k+l & = \\
l\cdot k^2-k+l & >_{l>1} \\
k^2-k+l & = \\
k\cdot(k-1)+l & >_{0<k-1} \\
l & >_{l>0} \\
0 &
\end{array}$$

Using the above together with [eq: 17.44] and the fact that $0<\frac{(2\cdot k)!}{2\cdot k^{2\cdot l-1}\cdot(2\cdot l-1)!\cdot(2\cdot k-2\cdot l-1)!}$ and $0\leqslant(2\cdot k-2\cdot l+1)\cdot(k-l)\cdot(2\cdot l+1)\cdot l\cdot k^2$ results in

$$\binom{2\cdot k}{2\cdot l-1}\cdot\frac{1}{k^{2\cdot l-1}}-\binom{2\cdot k}{2\cdot l+1}\cdot\frac{1}{k^{2\cdot l+1}}>0$$

Substituting this in [eq: 17.30] it follows that

$$\mathrm{Img}(\zeta^k)\underset{\text{[eq: 17.31]}}{=}B'>0 \tag{17.45}$$

Summaring [eq: 17.43] gives finally

$$\mathrm{Re}(\zeta^k)<0<\mathrm{Img}(\zeta^k)$$

□

17.1.4 Proof of the fundamental theorem of algebra

After all this work we are ready to proof fundamental theorem of algebra.

Theorem 17.29. *Let $n\in\mathbb{N}$, $\{a_i\}_{i\in\{0,\ldots,n\}}\subseteq\mathbb{C}$ with $a_n\neq 0$ and $p:\mathbb{C}\rightarrow\mathbb{C}$ the polynomial defined by $p(z)=\sum_{i=0}a_0\cdot z^i$ then there exist a $z_0\in\mathbb{C}$ such that $p(z_0)$. In other words every non constant polynomial has a zero point.*

Proof. First define

$$p\cdot\overline{p}:\mathbb{C}\to\mathbb{C} \text{ by } (p\cdot\overline{p})(z)=p(z)\cdot\overline{p(z)}$$

then for $z\in\mathbb{C}$ we have

$$\begin{aligned}
&(p\cdot\overline{p})(z) =\\
&p(z)\cdot\overline{p(z)} =\\
&\left(\sum_{i=0}^{n} a_i\cdot z^i\right)\cdot\overline{\left(\sum_{i=0}^{n} a_i\cdot z^i\right)} \underset{[\text{theorem: }10.81]}{=}\\
&\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)} \underset{[\text{theorem: }10.81]}{=}\\
&\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) \underset{[\text{theorem: }11.43]}{=}\\
&\sum_{(i,j)\in\{0,\ldots,\}^2} a_i\cdot z^i\cdot\overline{a_j\cdot z^j} =\\
&\sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2|k=l\}} a_i\cdot z^i\cdot\overline{a_j\cdot z^j}+\sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2|k<l\}} a_i\cdot z^i\cdot\overline{a_j\cdot z^j}+\sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2|l<k\}} a_i\cdot z^i\cdot\overline{a_j\cdot z^j} =\\
&\sum_{i\in\{0,\ldots,n\}} a_i\cdot z^i\cdot\overline{a_i\cdot z^i}+\sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2|k<l\}} a_i\cdot z^i\cdot\overline{a_j\cdot z^j}+\sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2|l<k\}} a_i\cdot z^i\cdot\overline{a_j\cdot z^j} =\\
&\sum_{i\in\{0,\ldots,n\}} a_i\cdot z^i\cdot\overline{a_i\cdot z^i}+\sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2|k<l\}} a_i\cdot z^i\cdot\overline{a_j\cdot z^j}+\sum_{(j,i)\in\{(k,l)\in\{0,\ldots,n\}^2|k<l\}} a_i\cdot\overline{a_j}\cdot z^i\cdot(\overline{z})^j =\\
&\sum_{i\in\{0,\ldots,n\}} a_i\cdot z^i\cdot\overline{a_i\cdot z^i}+\sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2|k<l\}} a_i\cdot z^i\cdot\overline{a_j\cdot z^j}+\sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2|k<l\}} a_j\cdot z^j\cdot\overline{a_i\cdot z^i} =\\
&\sum_{i\in\{0,\ldots,n\}} a_i\cdot\overline{a_i}\cdot z^i\cdot(\overline{z})^i+\sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2|k<l\}} (a_i\cdot z^i\cdot\overline{a_j\cdot z^j}+a_j\cdot z^j\cdot\overline{a_i\cdot z^i}) \underset{[\text{theorem: }10.81]}{=}\\
&\sum_{i\in\{0,\ldots,n\}} a_i\cdot\overline{a_i}\cdot z^i\cdot(\overline{z})^i+\sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2|k<l\}} \left(a_i\cdot z^i\cdot\overline{a_j\cdot z^j}+\overline{a_i\cdot z^i\cdot\overline{a_j\cdot z^j}}\right) \underset{[\text{theorem: }10.81]}{=}\\
&\sum_{i\in\{0,\ldots,n\}} a_i\cdot\overline{a_i}\cdot z^i\cdot(\overline{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\overline{a_j\cdot z^j}) =\\
&\sum_{i\in\{0,\ldots,n\}} |a_i|^2\cdot|z|^{2\cdot i}+2\cdot\sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2|k<l\}} \mathrm{Re}(a_i\cdot z^i\cdot\overline{a_j\cdot z^j})
\end{aligned}$$

so we have that

$$(p\cdot\overline{p})(z)=\sum_{i\in\{0,\ldots,n\}} |a_i|^2\cdot|z|^{2\cdot i}+2\cdot\sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2|k<l\}} \mathrm{Re}(a_i\cdot z^i\cdot\overline{a_j\cdot z^j}) \tag{17.46}$$

Further we have

$$\begin{aligned}
&(p\cdot\overline{p})(z) =\\
&p(z)\cdot\overline{p(z)} \underset{0\leqslant p(z)\cdot\overline{p(z)}\in\mathbb{R}}{=}\\
&|p(z)\cdot\overline{p(z)}| \underset{[\text{lemma: }17.17]}{=}\\
&\|p(z)\cdot\overline{p(z)}\| \underset{[\text{eq: }17.46]}{=}\\
&\left\|\sum_{i\in\{0,\ldots,n\}} |a_i|^2\cdot|z|^{2\cdot i}+2\cdot\sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2|k<l\}} \mathrm{Re}(a_i\cdot z^i\cdot\overline{a_j\cdot z^j})\right\| \geqslant_{[\text{theorem: }14.87]}\\
&\left\|\sum_{i\in\{0,\ldots,n\}} |a_i|^2\cdot|z|^{2\cdot i}\right\|-\left\|2\cdot\sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2|k<l\}} \mathrm{Re}(a_i\cdot z^i\cdot\overline{a_j\cdot z^j})\right\| \underset{[\text{lemma: }17.17]}{=}\\
&\left|\sum_{i\in\{0,\ldots,n\}} |a_i|^2\cdot|z|^{2\cdot i}\right|-\left\|2\cdot\sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2|k<l\}} \mathrm{Re}(a_i\cdot z^i\cdot\overline{a_j\cdot z^j})\right\| =
\end{aligned}$$

$$\begin{array}{rl}
\sum_{i\in\{0,\ldots,n\}} |a_i|^2\cdot|z|^{2\cdot i} - 2\cdot\left\|\sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2|k<l\}} \mathrm{Re}(a_i\cdot z^i\cdot\overline{a_j\cdot z^j})\right\| & \geqslant \\
|a_n|^2\cdot|z|^{2\cdot n} - 2\cdot\left\|\sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2|k<l\}} \mathrm{Re}(a_i\cdot z^i\cdot\overline{a_j\cdot z^j})\right\| & \geqslant \\
|a_n|^2\cdot|z|^{2\cdot n} - 2\cdot\sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2|k<l\}} \|\mathrm{Re}(a_i\cdot z^i\cdot\overline{a_j\cdot z^j})\| & \underset{[\text{theorem: } 10.81]}{=} \\
|a_n|^2\cdot|z|^{2\cdot n} - 2\cdot\sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2|k<l\}} \|\mathrm{Re}(a_i\cdot\overline{a}_j\cdot z^i\cdot\overline{z^j})\| & \geqslant_{[\text{lemma: } 17.17(4)]} \\
|a_n|^2\cdot|z|^{2\cdot n} - 2\cdot\sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2|k<l\}} \|a_i\cdot\overline{a_j}\|\cdot\|z^i\cdot\overline{z^j}\| & \geqslant_{[\text{lemma: } 17.17(3)]} \\
|a_n|^2\cdot|z|^{2\cdot n} - 2\cdot\sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2|k<l\}} \|a_i\cdot\overline{a_j}\|\cdot\|z^i\|\cdot\|\overline{z^j}\| & \underset{[\text{lemma: } 17.17(2)]}{=} \\
|a_n|^2\cdot|z|^{2\cdot n} - 2\cdot\sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2|k<l\}} \|a_i\cdot\overline{a_j}\|\cdot\|z^i\|\cdot\|z^j\| & \geqslant_{[\text{lemma: } 17.17(5)]} \\
|a_n|^2\cdot|z|^{2\cdot n} - 2\cdot\sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2|k<l\}} \|a_i\cdot\overline{a_j}\|\cdot\|z\|^i\cdot\|z\|^j & = \\
|a_n|^2\cdot|z|^{2\cdot n} - 2\cdot\sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2|k<l\}} \|a_i\cdot\overline{a_j}\|\cdot\|z\|^{i+j} & \underset{0\leqslant|z|^{2\cdot n}\in\mathbb{R}\wedge[\text{lemma: } 17.17]}{=} \\
|a_n|^2\cdot\||z|^{2\cdot n}\| - 2\cdot\sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2|k<l\}} \|a_i\cdot\overline{a_j}\|\cdot\|z\|^{i+j} & = \\
|a_n|^2\cdot\|z^n\cdot\overline{z^n}\| - 2\cdot\sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2|k<l\}} \|a_i\cdot\overline{a_j}\|\cdot\|z\|^{i+j} & \geqslant_{[\text{lemma: } 17.17(3)]} \\
|a_n|^2\cdot\frac{\|z^n\|\cdot\|\overline{z^n}\|}{2} - 2\cdot\sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2|k<l\}} \|a_i\cdot\overline{a_j}\|\cdot\|z\|^{i+j} & \underset{[\text{lemma: } 17.17(2)]}{=} \\
|a_n|^2\cdot\frac{\|z^n\|\cdot\|z^n\|}{2} - 2\cdot\sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2|k<l\}} \|a_i\cdot\overline{a_j}\|\cdot\|z\|^{i+j} & = \\
|a_n|^2\cdot\frac{\|z^n\|^2}{2} - 2\cdot\sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2|k<l\}} \|a_i\cdot\overline{a_j}\|\cdot\|z\|^{i+j} & \geqslant_{[\text{lemma: } 17.17(7)]} \\
|a_n|^2\cdot\frac{\|z\|^{2\cdot n}}{2\cdot 2^n} - 2\cdot\sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2|k<l\}} \|a_i\cdot\overline{a_j}\|\cdot\|z\|^{i+j} & = \\
|a_n|^2\cdot\frac{\|z\|^{2\cdot n}}{2^{n+1}} - 2\cdot\sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2|k<l\}} \|a_i\cdot\overline{a_j}\|\cdot\|z\|^{i+j} &
\end{array}$$

so that

$$(p\cdot\overline{p})(z)\geqslant|a_n|^2\cdot\frac{\|z\|^{2\cdot n}}{2^{n+1}} - 2\cdot\sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2|k<l\}} \|a_i\cdot\overline{a_j}\|\cdot\|z\|^{i+j} \tag{17.47}$$

Using [lemma: 17.27] we have $A=\{(k,l)\in\{0,\ldots,n\}^2|k<l\}=\bigsqcup_{m\in\{1,\ldots,2\cdot n-1\}} B_m$ where $B_m=\{(k,l)\in A|k+l=m\}$. Using this in the above [eq: 17.47] gives

$$\begin{array}{rl}
\sum_{(i,j)\in\{(k,l)\in\{0,\ldots,n\}^2|k<l\}} \|a_i\cdot\overline{a_j}\|\cdot\|z\|^{i+j} & = \\
\sum_{(i,j)\in\bigsqcup_{m\in\{1,\ldots,2\cdot n-1\}} B_i} \|a_i\cdot\overline{a_j}\|\cdot\|z\|^{i+j} & \underset{[\text{theorem: } 11.44]}{=} \\
\sum_{m\in\{1,\ldots,2\cdot n-1\}}\left(\sum_{(i,j)\in\{(k,l)|k+l=m\}} \|a_i\cdot\overline{a_j}\|\cdot\|z\|^{i+j}\right) & =
\end{array}$$

$$\begin{array}{rl}\sum_{m\in\{1,\ldots,2\cdot n-1\}}\left(\sum_{(i,j)\in\{(k,l)|k+l=m\}}\|a_i\cdot\overline{a_j}\|\cdot\|z\|^m\right) & =\\ \sum_{m\in\{1,\ldots,2\cdot n-1\}}\left(\sum_{(i,j)\in\{(k,l)|k+l=m\}}\|a_i\cdot\overline{a_j}\|\right)\cdot\|z\|^m & =\\ \sum_{m\in\{1,\ldots,2\cdot n-1\}} b_m\cdot\|z\|^m & \end{array}$$

where $\forall m\in\{1,\ldots,2\cdot n-1\}$ $b_m=\sum_{(i,j)\in\{(k,l)|k+l=m\}}\|a_i\cdot\overline{a_j}\|$. Substituting this in [eq: 17.47] gives

$$\begin{array}{rcl}(p\cdot\overline{p})(z) & \geqslant & |a_n|^2\cdot\frac{\|z\|^{2\cdot n}}{2^{n+1}}-2\cdot\sum_{m\in\{1,\ldots,2\cdot n-1\}} b_m\cdot\|z\|^m\\ & = & |a_n|^2\cdot\frac{\|z\|^{2\cdot n}}{2^{n+1}}+\sum_{m\in\{1,\ldots,2\cdot n-1\}}(-2\cdot b_m)\cdot\|z\|^m\\ & = & 0\cdot\|z\|^0+|a_n|^2\cdot\frac{\|z\|^{2\cdot n}}{2^{n+1}}+\sum_{m\in\{1,\ldots,2\cdot n-1\}}(-2\cdot b_m)\cdot\|z\|^m\\ & = & \sum_{i=0}^{2\cdot n} c_i\cdot\|z\|^i\end{array}\tag{17.48}$$

where

$$\{c_i\}_{i\in\{0,\ldots,2\cdot n\}}\text{where } c_i=\begin{cases}0 \text{ if } i=0\\ -2\cdot b_i \text{ if } i\in\{1,\ldots,2\cdot n-1\}\\ \frac{1}{2^{n+1}}\cdot|a_n|^2 \text{ if } i=2\cdot n\end{cases}$$

then as $a_n\neq 0$ we have that $0<c_{2\cdot n}$. Hence using [lemma: 17.15] it follows that

$$\lim_{x\rightarrow\infty}\sum_{i=0}^{2\cdot n} c_i\cdot x^i=\infty$$

so there exist a $R'\in\mathbb{R}$ such that if $x\geqslant R'$ then $\sum_{i=0}^{2\cdot n} c_i\cdot x^i\geqslant(p\cdot\overline{p})(0)+1>(p\cdot\overline{p})(0)$. Take $R=\min(R',1)$ then we have using [eq: 17.48] that

$$0<R \text{ and } \forall z\in\mathbb{C} \text{ with } \|z\|\geqslant R \text{ we have } (p\cdot\overline{p})(z)\geqslant\sum_{i=0}^{2\cdot n} c_i\cdot\|z\|^i>(p\cdot\overline{p})(0)\tag{17.49}$$

Now as $\overline{B_{\|\|}(0,R)}$ is closed and bounded in $\langle\mathbb{C},\|\|\rangle$ [see theorems: 14.65, 14.78] it follows from [theorem: 14.277] that

$$\overline{B_{\|\|}(0,R)} \text{ is compact in } \langle\mathbb{C},\|\|\rangle\tag{17.50}$$

Now by [theorem: 17.18(5)] the function $p\cdot\overline{p}$ is continuous in $\langle\mathbb{C},\|\|\rangle$ so that using the above we can apply the extreme value theorem [theorem: 14.235]. Hence there exist a $z_0\in\overline{B_{\|\|}(0,R)}$ so that

$$\forall z\in\overline{B_{\|\|}(0,R)} \text{ we have } (p\cdot\overline{p})(z)\geqslant(p\cdot\overline{p})(z_0)\tag{17.51}$$

As $0\in\overline{B_{\|\|}(0,R)}$ it follows from the above that

$$(p\cdot\overline{p})(0)\geqslant(p\cdot\overline{p})(z_0)\tag{17.52}$$

If $z\in\mathbb{C}$ then we have either:

$\|z\|\leqslant R$. then by [eq: 17.51] $(p\cdot\overline{p})(z)\geqslant(p\cdot\overline{p})(z_0)$

$R<\|z\|$. then by [eq: 17.49] $(p\cdot\overline{p})(z)\geqslant\geqslant(p\cdot\overline{p})(0)\geqslant_{[\text{eq: } 17.52]}(p\cdot\overline{p})(z_0)$

so we have

$$\forall c\in\mathbb{C} \text{ that } (p\cdot\overline{p})(z)\geqslant(p\cdot\overline{p})(z_0)\tag{17.53}$$

Define now

$$r:\mathbb{C}\rightarrow\mathbb{C} \text{ by } r(z)=p(z+z_0) \text{ and } r\cdot\bar{r}:\mathbb{C}\rightarrow\mathbb{C} \text{ by } (r\cdot\bar{r})(z)=r(z)\cdot\overline{r(z)} \tag{17.54}$$

Then we have

$$r(0)=p(0+z_0)=p(z_0) \tag{17.55}$$

Further

$$\begin{array}{rcl} (r\cdot\bar{r})(z) & = & r(z)\cdot\overline{r(z)} \\ & = & p(z+z_0)\cdot\overline{p(z+z_0)} \\ & \geqslant_{[\text{eq: } 17.53]} & (p\cdot\bar{p})(z_0) \\ & = & p(z_0)\cdot\overline{p(z_0)} \\ & \underset{\text{eq: } 17.55}{=} & r(0)\cdot\overline{r(0)} \\ & = & (r\cdot\bar{r})(0) \end{array}$$

in other words

$$\forall z\in\mathbb{C} \text{ we have } (r\cdot\bar{r})(z)-(r\cdot\bar{r})(0)=r(z)\cdot\overline{r(z)}-r(0)\cdot\overline{r(0)}\geqslant 0 \tag{17.56}$$

By [lemma: 17.24] and the definition of r it follows that r is a polynomial with $\text{ord}(r)=\text{ord}(p)$. Further using [lemma: 17.25] there exist a $k\in\{1,\ldots,n\}$ a a polynomial of order $n-k$ such that

$$q(0)\neq 0 \text{ and } \forall z\in\mathbb{C} \text{ we have } r(z)=r(0)+z^k\cdot q(z) \tag{17.57}$$

Then we have

$$\begin{array}{rcl} 0 & \leqslant & r(z)\cdot\overline{r(z)}-r(0)\cdot\overline{r(0)} \\ & = & (r(0)+z^k\cdot q(z))\cdot(\overline{r(0)+z^k\cdot q(z)})-(r(0)+0^k\cdot q(0))\cdot(\overline{r(0)+0^k\cdot q(0)}) \\ & = & (r(0)+z^k\cdot q(z))\cdot(\overline{r(0)+z^k\cdot q(z)})-r(0)\cdot\overline{r(0)} \\ & \underset{[\text{theorem: } 10.81]}{=} & (r(0)+z^k\cdot q(z))\cdot(\overline{r(0)}+\overline{z^k\cdot q(z)})-r(0)\cdot\overline{r(0)} \\ & = & r(0)\cdot\overline{r(0)}+r(0)\cdot\overline{z^k\cdot q(z)}+z^k\cdot q(z)\cdot\overline{r(0)}+z^k\cdot q(z)\cdot\overline{z^k\cdot q(z)}-r(0)\cdot\overline{r(0)} \\ & = & r(0)\cdot\overline{z^k\cdot q(z)}+z^k\cdot q(z)\cdot\overline{r(0)}+z^k\cdot q(z)\cdot\overline{z^k\cdot q(z)} \\ & = & \overline{\overline{r(0)}\cdot z^k\cdot q(z)}+z^k\cdot q(z)\cdot\overline{r(0)}+z^k\cdot q(z)\cdot\overline{z^k\cdot q(z)} \\ & = & \overline{z^k\cdot q(z)\cdot\overline{r(0)}}+z^k\cdot q(z)\cdot\overline{r(0)}+z^k\cdot q(z)\cdot\overline{z^k\cdot q(z)} \\ & \underset{[\text{theorem: } 10.81]}{=} & 2\cdot\text{Re}(z^k\cdot q(z)\cdot\overline{r(0)})+z^k\cdot q(z)\cdot\overline{z^k\cdot q(z)} \\ & \underset{[\text{theoren: } 10.83]}{=} & 2\cdot\text{Re}(z^k\cdot q(z)\cdot\overline{r(0)})+|z^k\cdot q(z)|^2 \end{array}$$

hence

$$|z^k\cdot q(z)|^2+2\cdot\text{Re}(z^k\cdot q(z)\cdot\overline{r(0)})\geqslant 0 \tag{17.58}$$

Given $x\in\mathbb{C}$ and $\delta\in\mathbb{R}^+$ then taken $z=\delta\cdot x$ and substituting this in the above gives

$$\begin{array}{rcl} 0 & \leqslant & |(\delta\cdot x)^k\cdot q(\delta\cdot x)|^2+2\cdot\text{Re}((\delta\cdot x)^k\cdot q(\delta\cdot x)\cdot\overline{r(0)}) \\ & \underset{[\text{theorems: } 10.81, 10.78]}{=} & \delta^{2\cdot k}\cdot|x^k\cdot q(\delta\cdot x)|^2+2\cdot\delta^k\cdot\text{Re}(x^k\cdot q(\delta\cdot x)\cdot\overline{r(0)}) \end{array}$$

and by dividing both sides of the above by $\delta^k>0$ we have

$$0\leqslant\delta^k\cdot|x^k\cdot q(\delta\cdot x)|^2+2\cdot\text{Re}(x^k\cdot q(\delta\cdot x)\cdot\overline{r(0)}) \tag{17.59}$$

Given $x\in\mathbb{C}$ define

$$f_x:\mathbb{R}\rightarrow\mathbb{R} \text{ by } f_x(\delta)=\delta^k\cdot|x^k\cdot q(\delta\cdot x)|^2+2\cdot\text{Re}(x^k\cdot q(\delta\cdot x)\cdot\overline{r(0)}) \tag{17.60}$$

We prove now that f_x is continuous by showing that it is the sum of functions composed of the continuous functions in [theorem: 17.18]. So define

$$f_{x,1}:\mathbb{C}\rightarrow\mathbb{R} \text{ by } \langle\cdot\rangle\circ(x^k\cdot)\circ q\circ(x\cdot)$$

wich is continous as the composition the continuous mappings $\langle\cdot\rangle$, $(x^k\cdot)$, q and $(x\cdot)$ [see theorem: 17.18(1,3,4)]. Using the definition of $f_{x,1}$ we have:

$$f_{x,1}(\delta)=\langle\cdot\rangle(x^k\cdot)(q(x\cdot)(\delta))=\langle\cdot\rangle(x^k\cdot)(q(x\cdot\delta))=\langle\cdot\rangle(x^k\cdot q(x\cdot\delta))=|x^k\cdot q(\delta\cdot x)|^2 \tag{17.61}$$

Define

$$f_{x,2}\colon\mathbb{C}\rightarrow\mathbb{C} \text{ by } f_{x,2}=((\cdot)^k)\cdot f_{x,1} \text{ so that } f_{x,2}(\delta)=(\cdot)^k(\delta)\cdot f_{x,1}(\delta)\underset{[\text{eq: } 17.61]}{=}\delta^2\cdot|x^k\cdot q(\delta\cdot x)|^2 \tag{17.62}$$

which is continuous because it is the product of the continuous mappings $(\cdot)^k$ and $f_{x,1}$ [see theorem: 17.18(2,10). Define

$$f_{x,3}\colon\mathbb{C}\rightarrow\mathbb{R} \text{ by } f_{x,3}=\mathrm{Re}\circ(x^k\cdot)\circ(\overline{r(0)}\cdot)\circ q\circ(x\cdot)$$

which is continuous as the composition of continuous functions [see theorem: 17.18(1,3,4,9)]. Further using the definition of $f_{x,3}$ we have

$$\begin{aligned}
f_{x,3}(\delta) &= \mathrm{Re}((x^k\cdot)(\overline{r(0)}\cdot)(q((x\cdot)(\delta))))\\
&= \mathrm{Re}((x^k\cdot)(\overline{r(0)}\cdot)(q(x\cdot\delta)))\\
&= \mathrm{Re}((x^k\cdot)(\overline{r(0)}\cdot)(q(\delta\cdot x)))\\
&= \mathrm{Re}((x^k\cdot)(\overline{r(0)}\cdot q(\delta\cdot x)))\\
&= \mathrm{Re}((x^k\cdot)(q(\delta\cdot x)\cdot\overline{r(0)}))\\
&= \mathrm{Re}(x^k\cdot q(\delta\cdot x)\cdot\overline{r(0)})
\end{aligned} \tag{17.63}$$

Hence we have for $\delta\in\mathbb{R}$

$$\begin{aligned}
(f_{x,2}+f_{x,3})(\delta) &= (f_{x,2})_{|\mathbb{R}}(\delta)+(f_{x,3})_{|\mathbb{R}}(\delta)\\
&= f_{x,2}(\delta)+f_{x,3}\\
&\underset{[\text{eqs: } 17.62,17.63]}{=} \delta^2\cdot|x^k\cdot q(\delta\cdot x)|^2+\mathrm{Re}(x^k\cdot q(\delta\cdot x)\cdot\overline{r(0)})\\
&\underset{[\text{eq: } 17.60]}{=} f_x(\delta)
\end{aligned}$$

proving that

$$f_x=(f_{x,2})_{|\mathbb{R}}+(f_{x,3})_{|\mathbb{R}}$$

where by [theorem: 14.135] $(f_{x,2})_{|\mathbb{R}}$ and $(f_{x,3})_{|\mathbb{R}}$ are continuous. So f_x is the sum of two continuous function hence by [theorem: 14.145] it is continuous. To summarize

$$\forall x\in\mathbb{C}\ f_x\colon\mathbb{R}\rightarrow\mathbb{R} \text{ defined by } f(\delta)=\delta^2\cdot|x^k\cdot q(\delta\cdot x)|^2+\mathrm{Re}(x^k\cdot q(\delta\cdot x)\cdot\overline{r(0)}) \text{ is continuous} \tag{17.64}$$

Combining this result with [eq: 17.59] results in

$$\forall x\in\mathbb{C} \text{ we have } \forall\delta\in\mathbb{R}^+ \text{ that } 0\leqslant f_x(\delta) \tag{17.65}$$

What about $\delta=0$? Assume that $f_x(0)<0$ then $\varepsilon=-f_x(0)\in\mathbb{R}^+$ so by the continuoty of f_x at 0 there exist a $\zeta\in\mathbb{R}^+$ such that $\forall\delta\in\mathbb{R}$ with $\|f_x(0)-f_x(\zeta)\|<-f(0)$. So if we take $\delta=\frac{\zeta}{2}$ then we have

$$\left|f_x(0)-f_x\left(\frac{\zeta}{2}\right)\right|\underset{[\text{theorem: } 17.17(1)]}{=}\left\|f_x(0)-f_x\left(\frac{\zeta}{2}\right)\right\|<-f_x(0)$$

As $f_x(0)<_{\text{assumption}}0\leqslant f_x\left(\frac{\zeta}{2}\right)$ we have that $f_x\left(\frac{\zeta}{2}\right)-f_x(0)=\left|f_x(0)-f_x\left(\frac{\zeta}{2}\right)\right|<-f_x(0)$ so that $f_x\left(\frac{\zeta}{2}\right)<0$ contradiction $0\leqslant f_x\left(\frac{\zeta}{2}\right)$. Hence the assumption is wrong and we have for $x\in\mathbb{C}$ that

$$\begin{aligned}
0 &\leqslant f_x(0)\\
&= 0^2\cdot|x^k\cdot q(0\cdot x)|^2+\mathrm{Re}(x^k\cdot q(0\cdot x)\cdot\overline{r(0)})\\
&= \mathrm{Re}(x^k\cdot q(0)\cdot\overline{r(0)})
\end{aligned}$$

or as $(0)\cdot\overline{r(0)}\in\mathbb{C}$

$$0\leqslant\mathrm{Re}(x^k\cdot(a+i\cdot b)) \text{ where } q(0)\cdot\overline{r(0)}=a+i\cdot b\ \forall x\in\mathbb{C} \tag{17.66}$$

Now for $k \in \{1, \ldots, n\}$ we have the following exclusive cases:

k is odd. Then $k = 2 \cdot m + 1$ [see theorem: 7.51]. Substituting $x = 1$ and $x = -1$ in [eq: 17.66] gives

$$0 \leqslant \mathrm{Re}(1^k \cdot (a + i \cdot b)) \underset{[\text{theorem: } 10.55]}{=} \mathrm{Re}(a + i \cdot b) = a$$

$$0 \leqslant \mathrm{Re}((-1)^k \cdot (a + i \cdot b)) = \mathrm{Re}((-1)^{2 \cdot m + 1}) \underset{[\text{theorem: } 10.55]}{=} \mathrm{Re}(-a - i \cdot b) = -a$$

hence $0 \leqslant a \wedge a \leqslant 0$ so that

$$a = 0$$

Now for m we have either:

$m =$ even. Then $m = 2 \cdot l$ so that $k = 4 \cdot l + 1$. Substitute $x = i$ and $x = -i$ in [eq: 17.66] taking in account that $a = 0$ gives

$$0 \leqslant \mathrm{Re}(i^k \cdot i \cdot b) = \mathrm{Re}(i^{4 \cdot l + 1} \cdot i \cdot b) \underset{[\text{theorem: } 10.79]}{=} \mathrm{Re}(i \cdot i \cdot b) = \mathrm{Re}(-b) = -b$$

and

$$0 \leqslant \mathrm{Re}((-i)^k \cdot i \cdot b) = \mathrm{Re}((-1)^{2 \cdot m + 1} \cdot k^{4 \cdot l + 1} \cdot i \cdot b) \underset{[\text{theorem: } 10.55, 10.79]}{=} \mathrm{Re}(-i \cdot i \cdot b) = b$$

Hence $0 \leqslant b \wedge b \leqslant 0$ so that

$$b = 0$$

$m =$ odd. Then $m = 2 \cdot l + 1$ so that $k = 2 \cdot (2 \cdot l + 1) + 1 = 4 \cdot l + 3$ hence if we substitute $x = i$ and $x = -i$ in [eq: 17.66] taking in account that $a = 0$ we get

$$0 \leqslant \mathrm{Re}(i^k \cdot i \cdot b) = \mathrm{Re}(i^{4 \cdot l + 3} \cdot i \cdot b) \underset{[\text{theorem: } 10.79]}{=} \mathrm{Re}((-i) \wedge i \cdot b) = \mathrm{Re}(b) = b$$

and

$$0 \leqslant \mathrm{Re}((-i)^k \cdot i \cdot b) = \mathrm{Re}((-1)^{2 \cdot m + 1} \cdot i^{4 \cdot l + 3} \cdot i \cdot b) \underset{[\text{theorem: } 10.55, 10.79]}{=} \mathrm{Re}((-1) \cdot (-i) \cdot i \cdot b) = \mathrm{Re}(-b) = -b$$

Hence $0 \leqslant b \wedge b \leqslant 0$ so that

$$b = 0$$

So we have that $a = b = 0$ so that $q(0) \cdot \overline{r(0)} = 0 + i \cdot 0 = 0$, as $q(0) \neq 0$ [see eq: 17.57] we conclude that

$$\overline{r(0)} = 0$$

k is even. Then $k = 2 \cdot m$ [see theorem: 7.51] and for m we have either:

m is odd. Then $m = 2 \cdot l + 1$ so that $k = 4 \cdot l + 2$. Substituting $x = i$ in [eq: 17.66] gives

$$\begin{array}{rcl} 0 & \leqslant & \mathrm{Re}(i^k \cdot (a + i \cdot b)) \\ & = & \mathrm{Re}(i^{4 \cdot l + 2} \cdot (a + i \cdot b)) \\ & \underset{[\text{theorem: } 10.79]}{=} & \mathrm{Re}(-(a + i \cdot b)) \\ & = & \mathrm{Re}(-a - i \cdot b) \\ & = & -a \end{array}$$

Substituting $x = 1$ in [eq: 17.66] gives

$$0 \leqslant \mathrm{Re}(1^k \cdot (a + i \cdot b)) = \mathrm{Re}(a + i \cdot b) = a$$

Hence $0 \leqslant a \wedge a \leqslant 0$ proving that

$$a = 0$$

Substituting $x = (1 + i)$ in [eq: 17.66] gives

$$\begin{array}{rcl} 0 & \leqslant & \mathrm{Re}((1 + i)^k \cdot (a + i \cdot b)) \\ & = & \mathrm{Re}((1 + i)^{4 \cdot l + 2} \cdot i \cdot b) \\ & \underset{[\text{lemma: } 17.19]}{=} & \mathrm{Re}(2 \cdot (-4)^l \cdot i \cdot i \cdot b) \\ & = & \mathrm{Re}(-2 \cdot (-4)^l \cdot b) \\ & = & -2 \cdot (-4)^l \cdot b \end{array}$$

and substituting $x=(1-i)$ in [eq: 17.66] gives

$$\begin{array}{rcl} 0 & \leqslant & \mathrm{Re}((1-i)^k\cdot(a+i\cdot b)) \\ & = & \mathrm{Re}((1-i)^{4\cdot l+2}\cdot i\cdot b) \\ & \underset{\text{[lemma: 17.19]}}{=} & \mathrm{Re}(-2\cdot(-4)^l\cdot i\cdot i\cdot b) \\ & = & \mathrm{Re}(2\cdot(-4)^l\cdot b) \\ & = & 2\cdot(-4)^l\cdot b \end{array}$$

If now $(-4)^l\leqslant 0$ then we have $b\leqslant 0\wedge 0\leqslant b\Rightarrow b=0$ and if $0\leqslant(-4)^l$ then $0\leqslant b\wedge b\leqslant 0\Rightarrow b=0$. So we have $a=b=0$ proving that $q(0)\cdot\overline{r(0)}=0+i\cdot 0=0$, as $q(0)\neq 0$ [see eq: 17.57] we conclude that

$$\overline{r(0)}=0$$

***m* is even.** Then $m=2\cdot l$ so that $k=4\cdot l$. Substituting $x=1$ in [eq: 17.66] gives

$$0\leqslant\mathrm{Re}(1^k\cdot(a+i\cdot b))=\mathrm{Re}(a)=a$$

proving that

$$0\leqslant a \tag{17.67}$$

Take now $x=\left(1+\frac{i}{2\cdot l}\right)$ then $x^k=\left(1+\frac{i}{2\cdot l}\right)^{4\cdot l}=\left(\left(1+\frac{i}{m}\right)^2\right)^m$ so that using [lemma: 17.28] we have that

$$\mathrm{Re}(x^k)<0<\mathrm{Img}(x^k)$$

As $x^k\in\mathbb{C}$ we can write x^k as $x^k=c+i\cdot d$ and $c=\mathrm{Re}(x^k)$ and $d=\mathrm{Img}(x^k)$ so that

$$c<0<d \tag{17.68}$$

Subsituting $x^k=c+i\cdot d$ in [eq: 17.66] gives

$$\begin{array}{rcl} 0 & \leqslant & \mathrm{Re}(x^k\cdot(a+i\cdot b)) \\ & = & \mathrm{Re}((c+i\cdot d)\cdot(a+i\cdot b)) \\ & = & \mathrm{Re}(c\cdot a-b\cdot d+i\cdot(d\cdot a+c\cdot b)) \\ & = & c\cdot a-b\cdot d \end{array}$$

so that

$$0\leqslant c\cdot a-b\cdot d \tag{17.69}$$

Take $y=\overline{x}$ then $y^k=(\overline{x})^k\underset{\text{[theorem: 10.81]}}{=}\overline{x^k}$ so that $y^k=c-i\cdot d$, substituting this in [eq: 17.66] gives

$$\begin{array}{rcl} 0 & \leqslant & \mathrm{Re}(y^k\cdot(a+i\cdot b)) \\ & = & \mathrm{Re}((c-i\cdot d)\cdot(a+i\cdot b)) \\ & = & \mathrm{Re}(c\cdot a+b\cdot d+i\cdot(c\cdot b-d\cdot a)) \\ & = & c\cdot a+b\cdot d \end{array}$$

so that

$$0\leqslant c\cdot a+b\cdot d \tag{17.70}$$

Adding [eq: 17.69] to [eq: 17.70] gives $0\leqslant c\cdot a-b\cdot d+c\cdot a+b\cdot d=2\cdot c\cdot a$ so that $0\leqslant c\cdot a$. As by [eq: 17.68] $c<0$ it follows that $a\leqslant 0$ which combined with [eq: 17.67] proves that

$$a=0$$

Substituting this in [eqs: 17.69 and 17.70] yields $0\leqslant -b\cdot d\Rightarrow b\cdot d\leqslant 0$ and $0\leqslant b\cdot d$ so that $b\cdot d=0$, as by [eq: 17.68] $d\neq 0$ it follows that

$$b=0$$

Hence $q(0)\cdot\overline{r(0)}=a+i\cdot b=0$ and as $q(0)\neq 0$ [see eq: 17.57] it follows that

$$\overline{r(0)}=0$$

Hence in all cases we have $\overline{r(0)} = 0$ so that $|r(0)|^2 = r(0) \cdot \overline{r(0)} = 0$ from which it follows that $r(0)$. As by definition [see: 17.55] $r(0) = p(z_0)$ it follows that

$$p(z_0) = 0$$

and we proved that there exist a zero point of the polynomial p. □

Note 17.30. In the above proof we have used the extreme value theorem to find the existence of a zero point of p without actual providing a explicit expression to calculate the zero point.

Theorem 17.31. *Let $n \in \mathbb{N}$ and $\{a_i\}_{i \in \{0, \ldots, n\}} \subseteq \mathbb{R}$ with $a_n \neq 0$ then if we define*

$$p: \mathbb{C} \rightarrow \mathbb{C} \text{ is defined by } p(z) = \sum_{i=0}^{n} a_i \cdot z^i$$

[in other words p is a polynomial of order n with real coefficients] then for $z \in \mathbb{C}$ with $p(z) = 0$ we have that $p(\bar{z}) = 0$.

Proof. If $p(z) = 0$ then we have

$$\begin{array}{rcl} 0 & = & \bar{0} \\ & = & \overline{p(z)} \\ & = & \overline{\sum_{i=0}^{n} a_i \cdot z^i} \\ & \underset{\text{[theorem: 10.81]}}{=} & \sum_{i=0}^{n} \overline{a_i} \cdot (\bar{z})^i \\ & \underset{\text{[theorem: 10.81]} \wedge a_i \in \mathbb{R}}{=} & \sum_{i=0}^{n} a_i \cdot (\bar{z})^i \\ & = & p(\bar{z}) \end{array}$$

□

Theorem 17.32. *Let $n \in \mathbb{N}$ and $\{a_i\}_{i \in \{0, \ldots, n\}} \subseteq \mathbb{C}$ with $a_n \neq 0$ then if we define*

$$p: \mathbb{C} \rightarrow \mathbb{C} \text{ is defined by } p(z) = \sum_{i=0}^{n} a_i \cdot z^i$$

[in other words p is a polynomial of order n] and $z_0 \in \mathbb{C}$ such that $p(z_0) = 0$ then there exist a $m \in \{1, \ldots, n\}$ and a polynomial q with $\operatorname{ord}(q) = n - m$ *such that $\forall z \in \mathbb{C}$ we have $p(z) = (z - z_0)^m \cdot q(z)$ and $q(z_0) \neq 0$.*

Proof. Define

$$r: \mathbb{C} \rightarrow \mathbb{C} \text{ by } r(z) = p(z + z_0) \tag{17.71}$$

then by [lemma: 17.24]

$$r \text{ is a polynomial with } \operatorname{ord}(r) = \operatorname{ord}(p) \in \mathbb{N} \tag{17.72}$$

Using [lemma: 17.25] there exist a $m \in \{1, \ldots, n\}$ and a polynomial s such

$$s(0) \neq 0 \ \operatorname{ord}(s) = n - m \text{ and } \forall z \in \mathbb{C} \text{ we have } r(z) = r(0) + z^m \cdot s(z) \tag{17.73}$$

Then we have for $z \in \mathbb{C}$

$$\begin{array}{rcl} p(z) & = & p((z - z_0) + z_0) \\ & = & r(z - z_0) \\ & \underset{\text{[eq: 17.73]}}{=} & r(0) + (z - z_0)^m \cdot s(z - z_0) \\ & \underset{\text{[eq: 17.71}}{=} & p(0 + z_0) + (z - z_0)^m \cdot s(z - z_0) \\ & = & p(z_0) + (z - z_0)^m \cdot s(z - z_0) \\ & = & (z - z_0)^m \cdot s(z - z_0) \end{array} \tag{17.74}$$

Define

$$q\colon \mathbb{C} \to \mathbb{C} \text{ by } q(z) = s(z - z_0) = s(z + (-z_0)) \tag{17.75}$$

then by [lemma: 17.24] we have that

$$q \text{ is a polynomial and } \operatorname{ord}(q) = \operatorname{ord}(s) = n - m$$

Substituting this in [eq: 17.74] gives

$$p(z) = (z - z_0)^m \cdot q(z)$$

Further

$$q(z_0) = s(z_0 - z_0) = s(0) \neq 0$$

□

Theorem 17.33. *Let p be polynomial with $\operatorname{ord}(p) \in \mathbb{N}$ then there exists a $c \in \mathbb{C} \setminus \{0\}$ and a $\{z_i\}_{i \in \{1,\ldots,\operatorname{ord}(p)\}} \subseteq \mathbb{C}$ such that*

$$\forall z \in \mathbb{C} \text{ we have } p(z) = c \cdot \prod_{i=1}^{\operatorname{ord}(p)} (z - z_i)$$

Further more

$$\{z \in \mathbb{C} | p(z) = 0\} = \{z_i | i \in \{1, \ldots, \operatorname{ord}(p)\}\}$$

in other words the distinct number of zero points of a polynomial is less then the order of the polynomial.

Proof. We prove this by induction on the order of the polynomial. So let

$$S = \left\{ n \in \mathbb{N} | \text{If } p \text{ is a polynomial with } 1 \leqslant \operatorname{ord}(p) \leqslant n \text{ then there exists a } c \in \mathbb{C} \setminus \{0\} \text{ and } \{z_i\}_{i \in \operatorname{ord}(p)} \text{ such that } \forall z \in \mathbb{C}\ p(z) = c \cdot \prod_{i=1}^{\operatorname{ord}(p)} (z - z_i) \right\}$$

then we have:

$1 \in \mathbb{N}$. Then $\operatorname{ord}(p) = 1$ so $p\colon \mathbb{C} \to \mathbb{C}$ is defined by $p(z) = a_0 + a_1 \cdot z$ where $a_1 \neq 0$. Take $z_1 = -\frac{a_0}{a_1}$ and $c = a_1$ then $c \neq 0$ and $c \cdot \prod_{i=1}^{1} (z - z_i) = c \cdot (z - z_i) = a_1 \cdot \left(z + \frac{a_0}{a_1}\right) = a_1 \cdot z + a_0 = p(z)$. So we have that $1 \in S$

$n \in \mathbb{N} \to n + 1 \in \mathbb{N}$. Let p be a polymial with $1 \leqslant \operatorname{ord}(p) \leqslant n + 1$ then we have for $\operatorname{ord}(p)$ either:

$1 \leqslant \operatorname{ord}(p) \leqslant n$. As $n \in S$ there exist a $c \in \mathbb{C} \setminus \{0\}$ and $\{z_i\}_{i \in \{1,\ldots,\operatorname{ord}(p)\}}$ with $p(z) = c \cdot \prod_{i=1}^{\operatorname{ord}(p)} (z - z_i)$ so that in this case $n + 1 \in S$.

$1 \leqslant \operatorname{ord}(p) = n + 1$. Let p be a polynomial with $\operatorname{ord}(p) = n + 1 \in \mathbb{N}$. Using the fundamental theorem of algebra there exist a $y_0 \in \mathbb{C}$ such that $p(y_0) = 0$. Using the previous theorem [theorem: 17.32] there exist a $m \in \{1, \ldots, n + 1\}$ and a polynomial q of order $(n + 1) - m$ such that

$$p(z) = (z - y_0)^m \cdot q(z) \text{ and } q(y_0) \neq 0 \tag{17.76}$$

for m we have now either:

$m = n + 1$. Then $\operatorname{ord}(q) = n + 1 - m = 0$ so that there exist a $c \in \mathbb{C} \setminus \{0\}$ with

$$q\colon \mathbb{C} \to \mathbb{C} \text{ defined by } q(z) = c$$

Hence if we define $\{z_i\}_{i \in \{1,\ldots,n+1\}} \subseteq \mathbb{C}$ by $z_i = y_0$ then

$$\forall z \in \mathbb{C} \text{ we have } p(z) = c \cdot (z - y_0)^{n+1} = c \cdot \prod_{i=1}^{n+1} (z - y_0) = c \cdot \prod_{i=1}^{\operatorname{ord}(p)} (z - z_i)$$

proving that $n + 1 \in S$ in this case.

$m < n + 1$. Then $\operatorname{ord}(q) = (n + 1) - m > 0$ so that $1 \leqslant \operatorname{ord}(q) \leqslant n$, as $n \in S$ there exist a $c \in \mathbb{C} \setminus \{0\}$ and a $\{y_i\}_{i \in \{1,\ldots,\operatorname{ord}(q)\}} \subseteq \mathbb{C}$ such that $\forall z \in \mathbb{C}$ we have

$$q(z) = c \cdot \prod_{i=1}^{\operatorname{ord}(q)} (z - y_i) = c \cdot \prod_{i=1}^{(n+1)-m} (x - y_i) \tag{17.77}$$

Define now $\{z_i\}_{i\in\{1,\ldots,n+1\}}\subseteq\mathbb{C}$ by $z_i=\begin{cases} y_0 \text{ if } i\in\{1,\ldots,m\}\\ z_{i-m} \text{ if } i\in\{m+1,\ldots,n+1\}\end{cases}$ then we have

$$\begin{array}{rcl} p(z) & \underset{[\text{eq: } 17.76]}{=} & (z-y_0)^m\cdot q(z)\\ & = & \prod_{i=1}^{m}(z-y_0)\cdot q(z)\\ & = & \prod_{i=1}^{m}(z-z_i)\cdot q(z)\\ & \underset{[\text{eq: } 17.77]}{=} & c\cdot\left(\prod_{i=1}^{m}(z-z_i)\right)\cdot\left(\prod_{i=1}^{(n+1)-m}(x-y_i)\right)\\ & = & c\cdot\left(\prod_{i=1}^{m}(z-z_i)\right)\cdot\left(\prod_{i=m+1}^{n+1}(x-y_{i-m})\right)\\ & = & c\cdot\left(\prod_{i=1}^{m}(z-z_i)\right)\cdot\left(\prod_{i=m+1}^{n+1}(x-z_i)\right)\\ & = & c\cdot\prod_{i=1}^{n+1}(z-z_i)\\ & = & c\cdot\prod_{i=1}^{\operatorname{ord}(p)}(z-z_i)\end{array}$$

proving that in this case $n+1\in S$.

Mathethematical induction proves that $S=\mathbb{N}$. Hence if p is a polynomial with $\operatorname{ord}(p)\in\mathbb{N}$ then as $1\leqslant\operatorname{ord}(p)\in\mathbb{N}=S$ it follows that

$$\exists c\in\mathbb{C}\setminus\{0\} \text{ and } \{z_i\}_{i\in\{1,\ldots,\operatorname{ord}(p)\}}\subseteq\mathbb{C} \text{ such that } p(z)=c\cdot\prod_{i=1}^{\operatorname{ord}(p)}(z-z_i)$$

Further if $z\in\{z_i|i\in\{1,\ldots,\operatorname{ord}(p)\}\}$ then there exist a $k\in\{1,\ldots,\operatorname{ord}(p)\}$ such that $x=z_k$, Hence

$$\begin{array}{rcl} p(z) & = & p(z_k)\\ & = & c\cdot\prod_{i=1}^{\operatorname{ord}(p)}(z_k-z_i)\\ & = & c\cdot\prod_{i\in\{1,\ldots,\operatorname{ord}(p)\}}(z_k-z_i)\\ & = & c\cdot\prod_{i\in\{i_0\}}(z_k-z_i)\cdot\prod_{i\in\{1,\ldots,\operatorname{ord}(p)\}\setminus\{i_0\}}(z_k-z_i)\\ & = & c\cdot(z_k-z_k)\cdot\prod_{i\in\{1,\ldots,\operatorname{ord}(p)\}\setminus\{i_0\}}(z_k-z_i)\\ & = & c\cdot 0\cdot\prod_{i\in\{1,\ldots,\operatorname{ord}(p)\}\setminus\{i_0\}}(z_k-z_i)\\ & = & 0\end{array}$$

so that $z\in\{z\in\mathbb{C}|p(z)=0\}$. Hence

$$\{z_i|i\in\{1,\ldots,\operatorname{ord}(p)\}\}\subseteq\{z\in\mathbb{C}|p(z)=0\} \tag{17.78}$$

Further if $z\in\{z\in S|p(z)=0\}$ then $0=c\cdot\prod_{i=1}^{\operatorname{ord}(p)}(z-z_i)$, hence as $c\neq 0$, we have that

$$0=\prod_{i=1}^{\operatorname{ord}(p)}(z-z_i)$$

Assume that $\forall i\in\{1,\ldots,\mathrm{ord}(p)\}$ $z_i\neq z$ then $\forall i\in\{1,\ldots,\mathrm{ord}(p)\}$ $z-z_i\neq 0$, hence using [theorem: 11.50] $\prod_{i=1}^{\mathrm{ord}(p)}(z-z_i)\neq 0$ contradicting $0=\prod_{i=1}^{\mathrm{ord}(p)}(z-z_i)$. So there must exist a $k\in\{1,\ldots,\mathrm{ord}(p)\}$ such that $z=z_k$. So $\{z\in S|p(z)=0\}\subseteq\{z_i|i\in\{1,\ldots,\mathrm{ord}(p)\}\}$ which combined with [eq: 17.78] proves that

$$\{z_i|i\in\{1,\ldots,\mathrm{ord}(p)\}\}=\{z\in\mathbb{C}|p(z)=0\}$$

□

Lemma 17.34. *Let p be a polynomial such that $\forall x\in\mathbb{R}$ $p(x)\in\mathbb{R}$ then p has real coefficients.*

Proof. As p is a polynomial there exist a $\{a_i\}_{i\in n}\subseteq\mathbb{C}$ such that $\forall z\in\mathbb{C}$ $p(z)=\sum_{i=0}^{n}a_i\cdot z^i$. Let $x\in\mathbb{R}$ then as $p(x)\in\mathbb{R}$ we have

$$\begin{aligned}
0 &= \mathrm{Img}(p(x))\\
&= \mathrm{Img}\left(\sum_{i=0}^{n}a_i\cdot x^i\right)\\
&\underset{[\text{example: }11.164]}{=} \sum_{i=0}^{n}\mathrm{Img}(a_i\cdot x^i)\\
&\underset{[\text{theorem: }10.78]\wedge x^i\in\mathbb{R}}{=} \sum_{i=0}^{n}\mathrm{Img}(a_i)\cdot x^i
\end{aligned}$$

so that $\forall x\in\mathbb{R}$ we have $\sum_{i=0}^{n}\mathrm{Img}(a_i)\cdot x^i=0$. As $\mathrm{Img}(a_i)\in\mathbb{R}$ it follows from [theorem: 17.2] that $\forall i\in\{0,\ldots,n\}$ we have $\mathrm{Img}(a_i)=0$ so that $a_i\in\mathbb{R}$. □

Theorem 17.35. *Let p be a polynomial with $\mathrm{ord}(p)\in\mathbb{N}$ and **real coefficients** then there exists a $c\in\mathbb{R}\setminus\{0\}$, $m,M\in\mathbb{N}_0$, $\{\lambda_i\}_{i\in\{1,\ldots m\}}\subseteq\mathbb{R}$, $\{b_i\}_{i\in\{1,\ldots,M\}}\subseteq\mathbb{R}$ 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

$$p(z)=c\cdot\left(\prod_{i\in\{1,\ldots,m\}}(z-\lambda_i)\right)\cdot\left(\prod_{i\in\{1,\ldots,M\}}(z^2+b_i\cdot z+c_i)\right)$$

Note 17.36. *That if $m=0$ then $\{1,\ldots,m\}=\varnothing$ so that $\{\lambda_i\}_{i\in\varnothing}$ is the empty family [see example: 2.103] and $\prod_{i\in\varnothing}(z-\lambda_i)=1$ [see definition: 11.32]. Likewise if $M=0$ then $\{b_i\}_{i\in\{1,\ldots,M\}}$ and $\{c_i\}_{i\in\{1,\ldots,M\}}$ are empty families and $\prod_{i\in\{1,\ldots,M\}}(z^2+b_i\cdot z+c_i)=1$*

Proof. We prove this by induction so defined

$$S=\left\{n\in\mathbb{N}|\text{If }p\text{ is a polynomial with real coefficients and }1\leqslant\mathrm{ord}(p)\leqslant n\text{ then there exists a }c\in\mathbb{C},\ m,M\in\mathbb{N}_0,\ \{\lambda_i\}_{i\in\{1,\ldots,m\}}\subseteq\mathbb{R},\ \{b_i\}_{i\in\{1,\ldots,M\}}\subseteq\mathbb{R}\text{ and }\{c_i\}_{i\in\{1,\ldots,M\}}\subseteq\mathbb{R}\text{ such that }\forall z\in\mathbb{C}\text{ we have }p(z)=c\cdot\left(\prod_{i\in\{1,\ldots,M\}}(z-\lambda_i)\right)\cdot\left(\prod_{i\in\{1,\ldots,M\}}z^2+b_i\cdot z+c_i\right)\right\}$$

then we have:

- **$1\in S$.** If $1\leqslant\mathrm{ord}(p)\leqslant 1$ then $\mathrm{ord}(p)$ and there exist $a_0,a_1\in\mathbb{R}$ such that $a_1\neq 0$ and $p(z)=a_0+a_1\cdot z$. Take $m=1$, $M=0$, $\{\lambda_i\}_{i\{1,\ldots,m\}}$ defined by $\lambda_1=-\frac{a_0}{a_1}\in\mathbb{R}$, $\{b_i\}_{i\in\varnothing}=\{c_i\}_{i\in\varnothing}=\varnothing$ then we have

  $$c\cdot\left(\prod_{i\in\{1,\ldots,m\}}(z-\lambda_i)\right)\cdot\left(\prod_{i\in\{1,\ldots,0\}=\varnothing}z^2+b_i\cdot z+c_i\right)=a_1\cdot\left(z+\frac{a_0}{a_1}\right)=a_0+a_1\cdot z-p(z)$$

  Hence we must have that $1\in S$.

- **$n\in S\Rightarrow n+1\in S$.** Let p be a polynomial with real coefficients and $1\leqslant\mathrm{ord}(p)\leqslant n+1$. For $\mathrm{ord}(p)$ we have either:

  - **$\mathrm{ord}(p)<n+1$.** Then $\mathrm{ord}(p)\leqslant n$ so as $n\in S$ we have that then there exists a $c\in\mathbb{R}$, $m,M\in\mathbb{N}_0$, $\{\lambda_i\}_{i\in\{1,\ldots,m\}}\subseteq\mathbb{R}$, $\{b_i\}_{i\in\{1,\ldots,M\}}\subseteq\mathbb{R}$ and $\{c_i\}_{i\in\{1,\ldots,M\}}\subseteq\mathbb{R}$ such that $\forall z\in\mathbb{C}$ we have $p(z)=c\cdot\left(\prod_{i\in\{1,\ldots,M\}}(z-\lambda_i)\right)\cdot\left(\prod_{i\in\{1,\ldots,M\}}z^2+b_i\cdot z+c_i\right)$. This proves that $n+1\in S$ in this case.

  - **$\mathrm{ord}(p)=n+1$.** As $1\leqslant\mathrm{ord}(p)$ $\mathrm{ord}(p)\in\mathbb{N}$ so that by the previous theorem there exist $\{z_i\}_{i\in\{1,\ldots,\mathrm{ord}(p)\}}\subseteq\mathbb{C}$ and a $c'\in\mathbb{C}$ with $c'\neq 0$ such that

    $$\forall z\in\mathbb{C}\text{we have }p(z)=c'\cdot\prod_{i=1}^{\mathrm{ord}(p)}(z-z_i)=c'\cdot\prod_{i\in\{1,\ldots,\mathrm{ord}(p)\}}(z-z_i)\tag{17.79}$$

and

$$\{z_i|i\in\{1,\ldots,\mathrm{ord}(p)\}\}=\{z|p(z)=0\} \tag{17.80}$$

If $\{z_i\}_{i\in\{1,\ldots,\ldots,\mathrm{ord}(p)\}}\subseteq\mathbb{R}$ then by taking $m=\mathrm{ord}(p)$, $M=0$ and $\{b_i\}_{i\in\varnothing}=\{c_i\}_{i\in\varnothing}=\varnothing$ we have that

$$p(z)=c'\cdot\left(\prod_{i=1}^{\mathrm{ord}(p)}(z-z_i)\right)\cdot\left(\prod_{i\in\{1,\ldots,0\}=\varnothing}(z^2+b_i\cdot z+c_i)\right)$$

proving that $n+1\in S$ in this case. Hence we must only prove the case where there exist $k\in\{1,\ldots,n+1\}$ such that $z_k\notin\mathbb{R}$. Now as p has real coefficients it follows from [lemma: 17.31] that $p(\overline{z_k})=0$, hence by [eq: 17.80] there exist a $l\in\{1,\ldots,n+1\}$ such that $z_l=\overline{z_k}$. If $k=l$ then we would have that $z_k=\overline{z_k}$ so that by [theorem: 10.81] $z\in\mathbb{R}$ contradicting $z_k\notin\mathbb{R}$. Hence we must have that $k\neq l$ and $z_l=\overline{z_k}$. Next we have

$$\begin{aligned}
p(z) &= c'\cdot\prod_{i\in\{1,\ldots,\mathrm{ord}(p)\}}(z-z_i)\\
&= c'\cdot\left(\prod_{i\in\{1,\ldots,\mathrm{ord}(p)\}\setminus\{k,l\}}(z-z_i)\right)\cdot\left(\prod_{i\in\{k,l\}}(z-z_i)\right)\\
&= c'\cdot\left(\prod_{i\in\{1,\ldots,\mathrm{ord}(p)\}\setminus\{k,l\}}(z-z_i)\right)\cdot(z-z_k)\cdot(z-z_l)\\
&= c'\cdot\left(\prod_{i\in\{1,\ldots,\mathrm{ord}(p)\}\setminus\{k,l\}}(z-z_i)\right)\cdot(z-z_k)\cdot(z-\overline{z_k})\\
&= c'\cdot\left(\prod_{i\in\{1,\ldots,\mathrm{ord}(p)\}\setminus\{k,l\}}(z-z_i)\right)\cdot(z^2-z\cdot\overline{z_k}-z_k\cdot z+z_k\cdot\overline{z_k})\\
&= c'\cdot\left(\prod_{i\in\{1,\ldots,\mathrm{ord}(p)\}\setminus\{k,l\}}(z-z_i)\right)\cdot(z^2-(z_k+\overline{z_k})\cdot z+z_k\cdot\overline{z_k})\\
&= c'\cdot\left(\prod_{i\in\{1,\ldots,\mathrm{ord}(p)\}\setminus\{k,l\}}(z-z_i)\right)\cdot(z^2+d\cdot z+e)
\end{aligned}$$

where $d=-(z_k+\overline{z_k})=-\mathrm{Re}(z_k)\in\mathbb{R}$ and $e=z_k\cdot\overline{z_k}=|z_k|^2\in\mathbb{R}$. Hence we have

$$p(z)=c'\cdot\left(\prod_{i\in\{1,\ldots,\mathrm{ord}(p)\}\setminus\{k,l\}}(z-z_i)\right)\cdot(z^2+d\cdot z+e)\text{ where }d=-\mathrm{Re}(z_k)\in\mathbb{R}\wedge e=|z_k|^2\in\mathbb{R} \tag{17.81}$$

Now as $z_k\in\mathbb{C}\setminus\mathbb{R}$ we have that $0<\mathrm{img}(z_k)$ so that

$$d^2=4\cdot(\mathrm{Re}(z_k))^2<4\cdot(\mathrm{Re}(z_k))^2+4\cdot(\mathrm{Img}(z_k))^k=4\cdot|z^k|^2=4\cdot e$$

hence

$$d^2<4\cdot e \tag{17.82}$$

As $k\neq l$ and $k,l\in\{1,\ldots,\mathrm{ord}(p)\}$ we have by [theorem: 10.90] that $\mathrm{card}(\{1,\ldots,\mathrm{ord}(p)\}\setminus\{k,l\})=\mathrm{ord}(p)-2$. Hence there exist a bijection

$$\beta:\{1,\ldots,\mathrm{ord}(p)-2\}\to\{1,\ldots,\mathrm{ord}(p)\}\setminus\{k,l\}$$

so that by applying [theorem: 11.36] on [eq; 17.81] gives

$$p(z)=c'\cdot\left(\prod_{i\in\{1,\ldots,\mathrm{ord}(p)-2\}}(z-z_{\beta(i)})\right)\cdot(z^2+d\cdot z+e) \tag{17.83}$$

For $i\in\{1,\ldots,\mathrm{ord}(p)-2\}$ we have as $c'\neq 0$ that

$$p_i:\mathbb{C}\to\mathbb{C}\text{ defined by }p_{\beta(i)}(z)=z-z_{\beta(i)}=-z_{\beta(i)}+z\text{ is a polynomial with }\mathrm{ord}(p_{\beta(i)})=1$$

Define now

$$q:\mathbb{C}\to\mathbb{C} \text{ by } q(z)=c'\cdot\prod_{i\in\{1,\ldots,\mathrm{ord}(p)-2\}} p_i(z)=\prod_{i=1}^{\mathrm{ord}(p)-2} p_i(x)=\prod_{i=1}^{\mathrm{ord}(p)-2}(z-z_{\beta(i)})$$

so that by [eq: 17.83] we have

$$p(z)=q(z)\cdot(z^2+d\cdot z+e) \tag{17.84}$$

Using [theorem: 17.33] it follows then that

$$q \text{ is a polynomial with } \mathrm{ord}(q)=\mathrm{ord}(p)-2=n-1 \tag{17.85}$$

Let $x\in\mathbb{R}$ then by [eq: 17.82] $0<e-\frac{d^2}{4}<\left(x+\frac{d}{2}\right)^2+e-\frac{d^2}{4}$ and as

$$\left(x+\frac{d}{2}\right)^2+e-\frac{d^2}{4}=x^2+x\cdot d+\frac{d^2}{4}+e-\frac{d^2}{4}=x^2+d\cdot x+e$$

we have that $0<x^2+d\cdot x+e$ which as $p(x)\in\mathbb{R}$ proves that $q(x)=\frac{p(z)}{z^2+d\cdot x+e}\in\mathbb{R}$. Using [theorem: 17.34] it follows that q has real coefficients hence as $\mathrm{ord}(q)=n-1<n$ and $n\in S$ there exist $c\in\mathbb{C}$, $m,M'\in\mathbb{N}_0$, $\{\lambda_i\}_{i\in\{1,\ldots,m\}}\subseteq\mathbb{R}$, $\{b'_i\}_{i\in\{1,\ldots,M'\}}\subseteq\mathbb{R}$ and $\{c'_i\}_{i\in\{1,\ldots,M'\}}\subseteq\mathbb{R}$ such that $\forall i\in\{1,\ldots,M'\}$ ${b'}^2<\frac{c'}{4}$ and

$$\forall z\in\mathbb{C}\ q(z)=c\cdot\left(\prod_{i\in\{1,\ldots,m'\}}(z-\lambda'_i)\right)\cdot\left(\prod_{i\in\{1,\ldots,M'\}}(z^2+b'_i\cdot z+c'_i)\right) \tag{17.86}$$

Hence if we take M=M'+1, $\{c_i\}_{i\in M}\subseteq\mathbb{R}$ by $c_i=\begin{cases} d \text{ if } i=M=M'+1\\ b'_i \text{ if } i\in\{1,\ldots,M'\}\end{cases}$ and $\{c_i\}_{i\in\{1,\ldots,M\}}\subseteq\mathbb{R}$ by $b_i=\begin{cases} e \text{ if i=M=M'+1}\\ c'_i \text{ if } i\in\{1,\ldots,M\}\end{cases}$ then we have $\forall i\in\{1,\ldots,M\}$ that $b^2<\frac{c}{4}$

$$\begin{array}{rl} c\cdot\left(\prod_{i\in\{1,\ldots,m\}}(z-\lambda_i)\right)\cdot\left(\prod_{i\in\{1,\ldots,M\}}(z^2+b_i\cdot z+c_i)\right) & = \\ c\cdot\left(\prod_{i\in\{1,\ldots,m\}}(z-\lambda_i)\right)\cdot\left(\prod_{i\in\{1,\ldots,M'\}}(z^2+b_i\cdot z+c_i)\right)\cdot(z^2+b_M\cdot z+c_M) & = \\ c\cdot\left(\prod_{i\in\{1,\ldots,m\}}(z-\lambda_i)\right)\cdot\left(\prod_{i\in\{1,\ldots,M'\}}(z^2+b_i\cdot z+c_i)\right)\cdot(z^2+d\cdot z+e) & =_{[\text{eq: } 17.86]} \\ q(z)\cdot(z^2+d\cdot z+e) & =_{[\text{eq: } 17.84]} \\ p(z) & \end{array}$$

Hence we have that

$$p(z)=c\cdot\left(\prod_{i\in\{1,\ldots,m\}}(z-\lambda_i)\right)\cdot\left(\prod_{i\in\{1,\ldots,M\}}(z^2+b_i\cdot z+c_i)\right)$$

so that we have $n+1\in S$ in this case.

Using mathematical induction we have that $S=\mathbb{N}$. Hence if p is a polynomial of $\mathrm{ord}(q)$ then we can write $p(z)$ as

$$p(z)=c\cdot\left(\prod_{i\in\{1,\ldots,m\}}(z-\lambda_i)\right)\cdot\left(\prod_{i\in\{1,\ldots,M\}}(z^2+b_i\cdot z+c_i)\right) \tag{17.87}$$

where $c\in\mathbb{C}$, $m,M\in\mathbb{N}_0$, $\{\lambda_i\}_{i\in\{1,\ldots,m\}}\subseteq\mathbb{R}$, $\{b_i\}_{i\in\{1,\ldots,M\}}\subseteq\mathbb{R}$ and $\{c_i\}_{i\in\{1,\ldots,M\}}\subseteq\mathbb{R}$ and $\forall i\in\{1,\ldots,M\}$ we have $b_i^2<\frac{c}{4}$. Rest us to prove that $c\in\mathbb{R}$. Using [theorem: 17.33] we have that there exist a family $\{z_i\}_{i\in\{1,\ldots,\mathrm{ord}(p)\}}\subseteq\mathbb{C}$ such that $\{z_i|i\in\{1,\ldots,\mathrm{ord}(p)\}\}=\{z\in\mathbb{C}|p(z)=0\}$. Assume that $\{1,\ldots,\mathrm{ord}(p)+1\}\subseteq\{z\in\mathbb{C}|p(z)=0\}$ then we have

$$\begin{array}{rcl} \mathrm{ord}(p)+1 & \leqslant_{[\text{theorem: } 10.89]} & \mathrm{card}(\{z\in\mathbb{C}|p(z)=0\}) \\ & = & \mathrm{card}(\{z_i|i\in\{1,\ldots,\mathrm{ord}(p)\}\}) \\ & \leqslant_{[\text{theorem: } 10.94]} & \mathrm{card}(\{1,\ldots,\mathrm{ord}(p)\}) \\ & = & \mathrm{ord}(p) \end{array}$$

leading to the contradiction $\operatorname{ord}(p) < \operatorname{ord}(p) + 1 \leqslant \operatorname{ord}(p)$, hence $\{1, \ldots, \operatorname{ord}(p) + 1\} \subseteq \{z \in \mathbb{C} | p(z) = 0\}$, so that there exist a $x \in \{1, \ldots, \operatorname{ord}(p) + 1\} \subseteq \mathbb{R}$ with $x \notin \{z \in \mathbb{C} | p(z) = 0\}$ so that $p(x) \neq 0$. Hence by [eq: 17.87] $\left(\prod_{i \in \{1, \ldots, m\}} (z - \lambda_i)\right) \cdot \left(\prod_{i \in \{1, \ldots, M\}} (z^2 + b_i \cdot z + c_i)\right) \neq 0$ so that we have

$$c = \frac{p(x)}{\left(\prod_{i \in \{1, \ldots, m\}} (x - \lambda_i)\right) \cdot \left(\prod_{i \in \{1, \ldots, M\}} (x^2 + b_i \cdot x + c_i)\right)}$$

As $p(x) \in \mathbb{R}$ [because p has real coefficients], $\prod_{i \in \{1, \ldots, m\}} (x - \lambda_i) \in \mathbb{R}$ [because λ_i is real] and $\prod_{i \in \{1, \ldots, M\}} (x^2 + b_i \cdot x + c_i) \in \mathbb{R}$ [because $b_i, c_i \in \mathbb{R}$] it follows that $c \in \mathbb{R}$. $\square$

Index