Analysis

Marc Mertens

Email: marc.lisp@gmail.com

November 30, 2023

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:

$\wedge$	meaning	and
$\vee$	meaning	or
$\neg$	meaning	not
$\Rightarrow$	meaning	implies
$\Leftrightarrow$	meaning	is equivalent with
$\vdash, \vDash$	meaning	with, where
$\forall$	meaning	for all
$\exists$	meaning	there exists
$\exists!$	meaning	there exists a unique

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 define **sets** and **elements**, they are the same thing. 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:

$\boldsymbol{R \in R}$**.** Then R is a element and $R \notin R$ giving the contradiction $R \in R \wedge R \notin R$

$\boldsymbol{R \notin R}$**.** 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*

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

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 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. *$A \backslash B \subseteq A$*
6. *If C is a class such that $A \subseteq C$ and $B \subseteq C$ then $A \bigcup B \subseteq C$*
7. *If C is a class such that $A \subseteq C$ and D a class such that $B \subseteq D$ then $A \bigcup B \subseteq C \bigcup D$*
8. *If C is a class such that $C \subseteq A$ and $C \subseteq B$ then $C \subseteq A \bigcap B$*
9. *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. 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 $x \in C$
7. Using (1) $A \subseteq C \bigcup D$ and $B \subseteq C \bigcup D$, so using (6) we have $A \bigcup B \subseteq C \bigcup D$
8. If $x \in C$ then $x \in A$ and $x \in B$ so that $x \in A \bigcap B$
9. 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 are classes then we have*

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

Proof.

1.

   ⇒. If $x \in A \bigcup B \Rightarrow x \in A \underset{A \subseteq B}{\Rightarrow} x \in B$ and thus $A \bigcup B \subseteq B$. From the previous theorem [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 $A \subseteq B$. □

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

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

□

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

$y=y'$. 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:

$z=x$. 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 $x\in\{x',y'\}$

$z=y$. 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:

$z=x'$. 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 $x\in\{x,y\}$

$z=y$. 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.

$\Rightarrow$. If $(x,y)=(x',y')$ then by definition

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

By [lemma: 1.39] we have either:

$\{x\}=\{x'\}\wedge\{x,y\}=\{x',y'\}$. then, as $x\in\{x\}$, we have by definition $x=x'$, using [lemma: 1.39] again we have either:

$x=x'\wedge y=y'$. Then $x=x'\wedge y=y'$

$x = y' \wedge y = x'$. Then by [theorem: 1.8] and $x = x'$ we have $y' = x'$ so that by [theorem: 1.8] again $y = y'$. Hence we have $x = x' \wedge y = y'$

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

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

$\Leftarrow$. 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

   $$A\subseteq B\wedge C\subseteq B \tag{1.7}$$

   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

   $$B\subseteq A\wedge C\subseteq D \tag{1.8}$$

   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. $(A\times B)\backslash(C\times D)=((A\backslash C)\times B)\bigcup(A\times(B\backslash D))$
7. $(A\backslash B)\times C=(A\times C)\backslash(B\times C)$
8. $A\times(B\backslash C)=(A\times B)\backslash(A\times C)$

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

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

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

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

7. We have

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

8. We have

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

□

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 $\bigcup \mathcal{A} = \{x \mid \exists y \in \mathcal{A}$ 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 $\bigcap \mathcal{A} = \{x \mid \forall y \in \mathcal{A}$ 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.
$$\begin{array}{ccl} x \in \bigcup \{A\} & \Leftrightarrow & \exists y \in \{A\} \text{ with } x \in y \\ & \underset{y \in \{A\} \Leftrightarrow y = A}{\Leftrightarrow} & x \in A \end{array}$$

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

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

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

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

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

   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

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

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

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

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

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

giving finally

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

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

$$A \times B \text{ is a set} \qquad \square$$

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

$$B=D$$

⇐. Assume that $A=C\wedge B=D$ then

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

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.

⇒. 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 $A=D\wedge B=E$.

⇐. 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 $f\colon A\rightarrow B$ 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 $f\colon A\rightarrow B$ 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 $f\colon A\rightarrow B$ be a partial function then its domain noted as* dom(f) *and range noted as* 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 $f\colon A\rightarrow B$ 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 $f\colon A\rightarrow B$ a partial function then* dom(f) *and* range(f) *are sets*

Proof. Using [theorem: 2.11] we have that $\operatorname{dom}(f) \subseteq A$ and $\operatorname{range}(f) \subseteq B$, so applying the Axiom of Subsets [axiom: 1.54] proves that $\operatorname{dom}(f)$ and $\operatorname{range}(f)$ 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. $f(C) \subseteq \operatorname{range}(f)$
2. $f^{-1}(D) \subseteq \operatorname{dom}(f)$
3. $f(A) = \operatorname{range}(f)$
4. $f^{-1}(B) = \operatorname{dom}(f)$
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 $y \in \operatorname{range}(f)$. Hence

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

2. If $x \in f^{-1}(D)$ then there exists a $y \in D$ such that $(x, y) \in f$, which proves that $x \in \operatorname{dom}(f)$, hence

$$f^{-1}(D) \subseteq \operatorname{dom}(f)$$

3. If $y \in \operatorname{range}(f)$ 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 $\operatorname{range}(f) \subseteq f(A)$. From (1) we have $f(A) \subseteq \operatorname{range}(f)$, so using [theorem: 1.8]

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

4. If $x \in \operatorname{dom}(f)$ 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 $\operatorname{dom}(f) \subseteq f^{-1}(B)$. From (2) we have $f^{-1}(B) \subseteq \operatorname{dom}(f)$, so using [theorem: 1.8]

$$f^{-1}(B) = \operatorname{dom}(f)$$

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] range(f) and 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 $f: A \rightarrow B$ and $g: C \rightarrow D$ be partial functions then*
$$g \circ f: A \rightarrow D$$
is a partial function. We call $g \circ f: A \rightarrow D$ the ***composiiton*** *of $f: A \rightarrow B$ 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 $g \circ f: A \rightarrow D$ 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 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. dom$(g \circ f) =$ dom$(f) \bigcap f^{-1}($dom$(g))$
2. range$(g \circ f) = g($range$(f) \bigcap$ dom$(g))$
3. range$(g \circ f) \subseteq$ range(g)

Proof.

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

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

3. If $z \in \operatorname{range}(g \circ f)$ 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 $z \in \operatorname{range}(g)$. □

Theorem 2.23. *If $f: A \to B$ and $g: C \to D$ 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 $C \subseteq f^{-1}(f(C))$.*

Proof. 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 $C \subseteq f^{-1}(f(C))$. □

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

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

which proves that

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

Finally

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

□

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 $\mathrm{dom}(f) = A$ 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 $\text{dom}(f) = C$ 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 $\text{range}(f) \subseteq C$ we have that $f: A \rightarrow C$ is a function.*

Proof. If $(x, y) \in f$ then $y \in \text{range}(f)$ hence as $\text{range}(f) \subseteq C$ $y \in C$. 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 $\text{range}(f) = A$ (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

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

If we have defined a function $f\colon A\rightarrow B$ 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

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

or

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

or

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

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 $f\colon A\rightarrow B$ and $g\colon A\rightarrow B$ are equal if and only if*

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

Proof. Assume that $f\colon A\rightarrow B$ and $g\colon A\rightarrow B$ 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 $f\colon A\rightarrow B$ and $g\colon A\rightarrow B$ are equal. □

Using the new notation, composition of function is written as

Theorem 2.42. *If $f\colon A\rightarrow B$ and $g\colon C\rightarrow D$ 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 $f\colon A\rightarrow B$ 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)** $\varnothing\colon \varnothing \rightarrow B$

Proof. First $\varnothing \subseteq \varnothing \times B$ by [theorem: 1.18], if $x \in \operatorname{dom}(\varnothing)$ then $\exists y \in \varnothing$ such that $(x, y) \in \varnothing$ which is a contradiction, so by [theorem: 1.19] we have that $\operatorname{dom}(\varnothing) = \varnothing$. 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 $C_c\colon A \rightarrow B$ 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

$$C_c\colon A \rightarrow B \text{ is a partial function}$$

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

$$C_c\colon A \rightarrow B \, \text{is} \, a \, \text{function}$$ □

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

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 $\mathcal{X}_{A,B}\colon A \rightarrow \{0, 1\}$ is a function. □

Example 2.47. **(Identity Function)** Let A be a class then $\operatorname{Id}_A\colon A \rightarrow B$ is defined by

$$I_A = \{z | z = (x, x) \wedge x \in A\}$$

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

$$\operatorname{Id}_A\colon A \rightarrow A \text{ is a function}$$ □

Proposition 2.48. *Let $f\colon A \rightarrow B$ be a partial function then $f = f \circ \operatorname{Id}_A$ and $f = \operatorname{Id}_B \circ f$*

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 $(x, x) \in \operatorname{Id}_A$, as $(x, y) \in f$ we have $(x, y) \in \operatorname{Id}_A \circ f$. If $(x, y) \in f \circ \operatorname{Id}_A$ then $\exists x'$ such that $(x, x') \in \operatorname{Id}_A \wedge (x', y) \in f$. 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

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

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 $(y, y) \in \operatorname{Id}_B$, so $(x, y) \in \operatorname{Id}_B \circ f$. If $(x, y) \in \operatorname{Id}_B \circ f$ 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

$$f = \operatorname{Id}_B \circ f$$ □

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

Theorem 2.49. *If $f\colon A\rightarrow B$ 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 $f\colon A\rightarrow B$ 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 $f\colon A\rightarrow B$ 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 $f\colon A\rightarrow B$ 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 $f\colon A\rightarrow B$ is injective then $f\colon A\rightarrow C$ 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 $i_B\colon B\rightarrow A$ 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 $A \subseteq \operatorname{dom}(i_b)$ 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: 2.78] 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 $f\colon I\rightarrow\{B\}$ 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 $\mathrm{dom}(f)=\{0\}=I$. So we conclude that $f\colon\{0\}\rightarrow\{B\}$ is indeed a function. Further if $y\in\{B\}$ then $y=B$ and as $(0,B)\in f$ it follows that $y\in\mathrm{range}(f)$ or $\{B\}\subseteq\mathrm{range}(f)$, 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 $f\colon\{0\}\rightarrow\{B\}$ is a bijection. □

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

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

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

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

Theorem 2.68. *If $f\colon A\rightarrow B$ is bijective then*

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

Proof. First $f^{-1}\colon B\rightarrow A$ 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 $B=\mathrm{dom}(f^{-1})$ 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 2.69. *If $f\colon A\rightarrow B$ 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 2.70. *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: 2.68] 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 2.71. *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: 2.68] $f\circ f^{-1} = \mathrm{Id}_B$ and $f^{-1}\circ f = \mathrm{Id}_A$ which by [theorem: 2.70] proves that $f^{-1}: B\rightarrow A$ is a bijection. □

Proposition 2.72. *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

$$\begin{aligned}
f\circ g = \mathrm{Id}_B \quad & \Rightarrow \quad f^{-1}\circ(f\circ g) = f^{-1}\circ \mathrm{Id}_B \\
& \underset{\text{[proposition: 2.48}}{\Rightarrow} \quad f^{-1}\circ(f\circ g) = f^{-1} \\
& \underset{\text{[theorem: 2.21]}}{\Rightarrow} \quad (f^{-1}\circ f)\circ g = f^{-1} \\
& \underset{\text{[function: 2.68]}}{\Rightarrow} \quad \mathrm{Id}_B\circ g = f^{-1} \\
& \underset{\text{[proposition: 2.48}}{\Rightarrow} \quad g = f^{-1}
\end{aligned}$$

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 2.73. *We have*

1. If $f\colon A\rightarrow B$ *and* $g\colon C\rightarrow D$ *are injective functions with* $f(A)\subseteq C$ *then* $g\circ f\colon A\rightarrow D$ *is a injective function.*
2. *If* $f\colon A\rightarrow B$ *and* $g\colon C\rightarrow D$ *are injective functions with* $f(A)\subseteq C$ *then* $g\circ f\colon A\rightarrow g(f(A))$ *is a bijective function.*
3. *If* $f\colon A\rightarrow B$ *is a function and* $g\colon C\rightarrow D$ *a surjective function so that* $f(A)=C$ *then* $g\circ f\colon A\rightarrow D$ *is a surjective function.*
4. *If* $f\colon A\rightarrow B$ *is a injective function and* $g\colon C\rightarrow D$ *a bijective function so that* $f(A)=C$ *then* $g\circ f\colon A\rightarrow D$ *is a bijective function.*
5. *If* $f\colon A\rightarrow B$ *is a injective function and* $g\colon C\rightarrow D$ *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 $g\circ f\colon A\rightarrow D$ is injective so that by [theorem: 2.66] $g\circ f\colon A\rightarrow(g\circ f)(A)$ is a bijection. Further by [theorem: 2.23] $(g\circ f)(A)=g(f(A))$ so that $g\circ f\colon A\rightarrow g(f(A))$ 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 $g\circ f\colon A\rightarrow D$ is injective and surjective and thus by definition bijective.
5. By (3) $g\circ f$ is a bijection, so by [theorem: 2.68] we have that
$$\begin{aligned}
(g\circ f)^{-1}\circ(g\circ f)=\mathrm{Id}_A &\underset{\text{[associativity: 2.21]}}{\Rightarrow} ((g\circ f)^{-1}\circ g)\circ f=\mathrm{Id}_A\\
&\Rightarrow (((g\circ f)^{-1}\circ g)\circ f)\circ f^{-1}=\mathrm{Id}_A\circ f^{-1}\\
&\underset{\text{[proposition: 2.48]}}{\Rightarrow} (((g\circ f)^{-1}\circ g)\circ f)\circ f^{-1}=f^{-1}\\
&\underset{\text{[associativity: 2.21]}}{\Rightarrow} ((g\circ f)^{-1}\circ g)\circ(f\circ f^{-1})=f^{-1}\\
&\underset{\text{[theorem: 2.68]}}{\Rightarrow} ((g\circ f)^{-1}\circ g)\circ\mathrm{Id}_B=f^{-1}\\
&\underset{\text{[proposition: 2.48]}}{\Rightarrow} (g\circ f)^{-1}\circ g=f^{-1}\\
&\Rightarrow ((g\circ f)^{-1}\circ g)\circ g^{-1}=f^{-1}\circ g^{-1}\\
&\underset{\text{[associativity: 2.21]}}{\Rightarrow} (g\circ f^{-1})\circ(g\circ g^{-1})=f^{-1}\circ g^{-1}\\
&\underset{\text{[theorem: 2.68]}}{\Rightarrow} (g\circ f)^{-1}\circ\mathrm{Id}_A=f^{-1}\circ g^{-1}\\
&\underset{\text{[proposition: 2.48]}}{\Rightarrow} (g\circ f)^{-1}=f^{-1}\circ g^{-1}
\end{aligned}$$

□

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

Corollary 2.74. *We have*

1. If $f\colon A\rightarrow B$ *and* $g\colon B\rightarrow C$ *are injective functions then* $g\circ f\colon A\rightarrow C$ *is a injective function.*
2. *If* $f\colon A\rightarrow B$ *and* $g\colon B\rightarrow C$ *are surjective functions then* $g\circ f\colon A\rightarrow C$ *is a surjective function.*

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

Proof.

1. This follows from [theorem: 2.73 (1)] because $f(A) \subseteq B$.
2. This follows from [theorem: 2.73 (2)] because if f is surjective we have $f(A) = B$.
3. This follows from (1) and (2)
4. This follows from [theorem: 2.73 (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 2.75. *Let A be a class then there exists a bijection between A and $A^{\{0\}}$*

Proof. Given $x \in A$ define the function $f_x: \{0\} \rightarrow \{x\}$ 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

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

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

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

If $y \in A^{\{0\}}$ then $y: \{0\} \rightarrow A$ is a function, hence $0 \in \{0\} = \mathrm{dom}(y)$, 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 2.76. *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 $\gamma: \mathcal{P}(A) \rightarrow \{0, 1\}^A$ 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 $\mathcal{X}_{A,B}: A \rightarrow \{0, 1\}$ 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 $\mathrm{dom}(\gamma) = \mathcal{P}(A)$, hence

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

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

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

proving that $B = B'$. Hence

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

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:

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

$\boldsymbol{x \notin B}$**.** 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 $x \in A \backslash B$]

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 $x \in A \backslash B$ 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

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

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 $\mathbb{R} \backslash \{1\}$ produces a continuous function. This is the idea of the next definition

Definition 2.77. *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 2.78. *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]

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

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

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

3.

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

so

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

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

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

proving that

$$f\bigcup g\colon A\bigcup B\rightarrow C \text{ is a function} \qquad \square$$

Corollary 2.79. *Let $f\colon A\rightarrow B$ and $g\colon C\rightarrow D$ be functions such that $A\bigcap C=\varnothing$ then*

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

is a function.

Proof. Using [theorem: 2.33] we have that $f\colon A\rightarrow B\bigcup D$ and $g\colon C\rightarrow B\bigcup D$ are functions. Applying then the previous theorem [theorem: 2.78] proves that $f\bigcup g\colon A\bigcup C\rightarrow B\bigcup D$ is a function. □

Corollary 2.80. *Let $f\colon A\rightarrow B$ and $g\colon C\rightarrow D$ be bijections with $A\bigcap C=\varnothing$ and $B\bigcap D=\varnothing$ then*

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

is a bijection.

Proof. Using the previous theorem [theorem: 2.79] we have that $f\bigcup g\colon A\bigcup C\rightarrow B\bigcup D$ 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:

$y\in B$. 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 $x=x'$.

$y\in D$. 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 $f\bigcup g\colon A\bigcup C\rightarrow B\bigcup D$.

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

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

$y\in D$. 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 2.81. *Let $f\colon A\rightarrow B$ be a function a,b elements such that $a\notin A$ then*

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

is a function.

***Note* 2.82.** *A alternative definition of g is* $g(x)=\left\{\begin{smallmatrix} b \text{ if } x=a \\ f(x) \text{ if } x\in A\end{smallmatrix}\right.$

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: 2.79] so that

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

Theorem 2.83. *Let $f: A\rightarrow B$ be a partial function and $C\subseteq A$ a sub-class of A then we have:*

1. $\mathrm{dom}(f_{|C})=C\bigcap\mathrm{dom}(f)$
2. $\mathrm{range}(f_{|C})=f(C)$
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 $x\in\mathrm{dom}(f_{|C})$ 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 $x\in\mathrm{dom}(f)$, so that $x\in C\bigcap\mathrm{dom}(f)$. Hence

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

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

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

2. If $y\in\mathrm{range}(f_{|C})$ 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 $y\in\mathrm{range}(f_{|C})$. Hence using the Axiom of Extent [axiom: 1.5] we have

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

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

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

$$f_{|\operatorname{dom}(f)}=f \qquad \square$$

Theorem 2.85. *Let $f\colon A\rightarrow B$ 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:

$$(f^{-1})_{f(C)}=(f_{|C})^{-1} \qquad \square$$

Theorem 2.86. *Let $f\colon A\rightarrow B$ and $g\colon C\rightarrow D$ 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

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

Theorem 2.87. *Let $f\colon A\rightarrow B$ and $C\subseteq A$ a sub-class of A then $f_{|C}\colon C\rightarrow B$ is a function.*

Proof. Using [definition: 2.77] we have that $f_{|C}\colon C\rightarrow B$ is a partial function, as by [theorem: 2.83] $\operatorname{dom}(f_{|C})=C\bigcap\operatorname{dom}(f)\underset{f\text{ is a function}}{=}C\bigcap A\underset{C\subseteq A}{=}C$, it follows that $f_{|C}\colon C\rightarrow B$ is a function. □

The following theorem will be used for manifolds later

Theorem 2.88. *Let $f\colon A\rightarrow B$ and $g\colon C\rightarrow D$ be injections then we have*

1. *$f\colon A\rightarrow f(A)$ and $g\colon C\rightarrow f(C)$ are bijections*
2. *$\operatorname{dom}(f\circ g^{-1})=g(A\bigcap C)$*
3. *$f\circ g^{-1}\colon g(A\bigcap C)\rightarrow f(A\bigcap C)$ 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 $z\in\operatorname{dom}(f\circ g^{-1})$ 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

$$\operatorname{dom}(g\circ f^{-1})\subseteq g\left(A\bigcap C\right) \tag{2.27}$$

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 $A=\operatorname{dom}(f)$, 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 $z\in\operatorname{dom}(f\circ g^{-1})$. Hence $g(A\bigcap C)\subseteq\operatorname{dom}(g\circ f^{-1})$ which combined with [eq: 2.27].

$$\operatorname{dom}(g\circ f^{-1})=g\big(A\bigcap C\big)$$

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 $g\colon C\rightarrow B$ 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

$$\begin{aligned}(f\circ g^{-1}) &\underset{[\text{theorem: } 2.84]}{=} (f\circ g^{-1})_{\operatorname{dom}(f\circ g^{-1})}\\ &\underset{(1)}{=} (f\circ g^{-1})_{g(A\bigcap C)}\\ &\underset{[\text{theorem: } 2.86]}{=} f_{|g^{-1}(g(A\bigcap C))}\circ(g^{-1})_{g(A\bigcap C)}\\ &\underset{g\text{ is injective and [theorem: } 2.55]}{=} f_{|A\bigcap C}\circ(g^{-1})_{|g(A\bigcap C)}\\ &\underset{[\text{theorem: } 2.85]}{=} f_{|A\bigcap C}\circ(g_{|A\bigcap C})^{-1}\end{aligned}$$

□

2.2.5 Set operations and (Partial) Functions

Theorem 2.89. *Let $f\colon A\rightarrow B$ 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

$$f^{-1}(C\setminus D)\subseteq f^{-1}(C)\setminus f^{-1}(D)\tag{2.28}$$

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: 2.28] 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 2.90. *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
$$f\left(E \bigcup F\right) \subseteq f(E) \bigcup f(F) \tag{2.29}$$
If $y \in f(E) \bigcup f(F)$ then we have the following possibilities

   $y \in f(E)$. 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 $y \in f(E \bigcup F)$

   $y \in f(F)$. 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: 2.29] 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
$$f^{-1}\left(C \bigcup D\right) \subseteq f^{-1}(C) \bigcup f^{-1}(D) \tag{2.30}$$
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 2.30] 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
$$f\left(E \bigcap F\right) \subseteq f(E) \bigcap f(F) \tag{2.31}$$
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: 2.31] 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
$$f^{-1}\left(C \bigcap D\right) \subseteq f^{-1}(C) \bigcap f^{-1}(D) \tag{2.32}$$
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: 2.32] 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 2.91. *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 $x \in \operatorname{dom}(f)$. So $A \subseteq \operatorname{dom}(f)$ 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 2.92. *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 2.93. $f: \mathbb{R} \rightarrow \mathbb{R}$ 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 2.94. **(indexed set)** *Let $f\colon I\rightarrow A$ 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 2.95. **(unique indexed set)** *$A=\{f_i|i\in I\}$ is a **unique indexed set** if $f\colon I\rightarrow A$ 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 2.96. 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 2.97. *A Let I,B be classes then a family*

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

is actually a function

$$\boldsymbol{f\colon I\rightarrow B}$$

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 2.98. 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 $\boldsymbol{x\colon I\rightarrow B}$. However this is sometimes not feasible and in that case we state what the defining function of $\{x_i\}_{i\in I}$ is.

Example 2.99. The empty function $\varnothing\colon\varnothing\rightarrow V$ [see example: 2.44] defines the family $\{\varnothing_i\}_{i\in\varnothing}\subseteq V$. Further if $\{x_i\}_{i\in\varnothing}\subseteq V$ is a family 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 $x\colon\varnothing\rightarrow V$, 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 2.100. *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 $f\colon I\rightarrow\varnothing$ 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 2.101. Let A,B be classes then family $\{(\mathrm{Id}_A)_a\}_{a\in A}\subseteq B$ defined by the function Id_A: $A\rightarrow B$ is noted as $\{x\}_{x\in A}$.

We can now define the concept of a sub family

Definition 2.102. *Let $\{A_i\}_{i\in I}\subseteq B$ be a family of objects in B defined by the function $f: I\rightarrow B$ and $J\subseteq I$ then $\{A_i\}_{i\in J}\subseteq B$ is the family defined by the function $f_{|J}: J\rightarrow B$ [see: theorem: 2.87 for the proof that $f_J: I\rightarrow B$ is a function]*

Definition 2.103. *Let I, J, A, B be classes such that $I\bigcap J=\varnothing$ and*

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

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

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: 2.79 for the proof that $f\bigcup g: I\bigcup J\rightarrow A\bigcup B$ is indeed a function]

Definition 2.104. *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 2.105. *If you have a function $\boldsymbol{f: I\longrightarrow J}$ and a family $\boldsymbol{\{x_j\}_{j\in J}\subseteq A}$ [defined by the function $\boldsymbol{g: J\longrightarrow A}$] then*

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

is the family represented by the function

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

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

Definition 2.106. *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*

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

The motivation for this definition is the following theorem

Theorem 2.107. *If $\{x_i\}_{i\in I}\subseteq B$ is a family of objects in B with associated function f then*

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

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

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

Theorem 2.108. *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 $\mathrm{range}(x)=\{x_i|i\in I\}$. As $I$and B are sets, it follows from [theorem: 2.12] that $\mathrm{range}(x)$ 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 2.109. *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*

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

Definition 2.110. *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 2.111. *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 2.112. *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 $\bigcup_{i \in I} A_i = \bigcup \text{range}(A)$. Then we have

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

□

Corollary 2.113. *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 $x \in \bigcup_{i \in J} A_j$ then by [theorem: 2.112] 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: 2.112] and the definition of $\bigcup_{i \in I} A_{f(i)}$ we have $x \in \bigcup_{i \in I} A_{f(i)}$. Hence

$$\bigcup_{j \in J} A_j \subseteq \bigcup_{i \in I} A_{f(i)} \tag{2.33}$$

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] $(A \circ f)(i) \in \text{range}(A)$, means that there exists a $j \in J$ such that $A_j = (A \circ f)(i)$. Hence $x \in A_j$ proving by [theorem: 2.112] 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: 2.33] gives

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

□

Theorem 2.114. *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 range(A) is a set. Using the Axiom of Union [axiom: 1.61] $\bigcup$ range(A) is a set, so by definition $\bigcup_{i \in I} A_i$. is a set. □

Example 2.115. 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 $y \in \text{range}(A) = \text{range}(\varnothing)$ then x such that $(x, y) \in \varnothing$, a contradiction. Hence range$(A) = \varnothing$. So

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

Definition 2.116. *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*

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

Theorem 2.117. *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 2.118. *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 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 $\bigcap$ range(A) is a set, hence $\bigcap_{i \in I} A_i$ is a set. □

Example 2.119. 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 2.120. 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: 2.112] and [theorem: 2.117] about union and intersection of families without explicit referring to these theorems.

Theorem 2.121. *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 2.122. *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

$$\left(\bigcup_{i \in J} A_i\right) \bigcup \left(\bigcup_{i \in K} A_i\right) \subseteq \bigcup_{i \in I} A_i \tag{2.34}$$

   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: 2.34] 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

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

   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: 2.35] 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 2.123. *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 2.124. (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. $C\bigcup\left(\bigcap_{i\in I}A_i\right)=\bigcap_{i\in I}\left(C\bigcup A_i\right)$

3. $C\bigcap\left(\bigcap_{i\in I}A_i\right)=\bigcap_{i\in I}\left(C\bigcap A_i\right)$

4. $C\bigcup\left(\bigcup_{i\in I}A_i\right)=\bigcup_{i\in I}\left(C\bigcup A_i\right)$

Proof.

1. If $x\in C\bigcap\left(\bigcup_{i\in I}A_i\right)$ 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: 2.121] that $x\in\bigcup_{i\in I}A_i$. So

$$C\bigcap\left(\bigcup_{i\in I}A_i\right)\subseteq\bigcup_{i\in I}\left(C\bigcap A_i\right)\tag{2.36}$$

If $x\in\bigcup_{i\in}\left(C\bigcap A_i\right)$ 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 $x\in C\bigcap\left(\bigcup_{i\in I}A_i\right)$, proving that $\bigcup_{i\in I}\left(C\bigcap A_i\right)\subseteq C\bigcap\left(\bigcup_{i\in I}A_i\right)$. Combining this with [eq: 2.36] proves

$$C\bigcap\left(\bigcup_{i\in I}A_i\right)=\bigcup_{i\in I}\left(C\bigcap A_i\right)$$

2. If $x\in C\bigcup\left(\bigcap_{i\in I}A_i\right)$ then we have the following cases to consider:

   $\boldsymbol{x\in C}$**.** then $\forall i\in I$ we have $x\in C\bigcup A_i$ hence $x\in\bigcap_{i\in I}\left(C\bigcup A_i\right)$

   $\boldsymbol{x\in\bigcap_{i\in I}A_i}$**.** then $\forall i\in I$ we have $x\in A_i$ hence $x\in\bigcap_{i\in I}\left(C\bigcup A_i\right)$

which proves that

$$C\bigcup\left(\bigcap_{i\in I}A_i\right)\subseteq\bigcap_{i\in I}\left(C\bigcup A_i\right)\tag{2.37}$$

If $x\in\bigcap_{i\in I}\left(C\bigcup A_i\right)$ then we have two cases to consider:

   $\boldsymbol{x\in C}$**.** then $x\in C\bigcup\left(\bigcap_{i\in I}A_i\right)$

   $\boldsymbol{x\notin C}$**.** 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 $x\in C\bigcup\left(\bigcap_{i\in I}A_i\right)$

In all cases we have $x\in C\bigcup\left(\bigcap_{i\in I}A_i\right)$ proving that $\bigcap_{i\in I}\left(C\bigcup A_i\right)\subseteq C\bigcup\left(\bigcap_{i\in I}A_i\right)$, combining this with [eq: 2.37] gives

$$C\bigcup\left(\bigcap_{i\in I}A_i\right)\subseteq\bigcap_{i\in I}\left(C\bigcup A_i\right)$$

3. We have

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

Proving

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

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 2.125. *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: 2.123] 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

$$\left(\bigcup_{i\in I} A_i\right)\bigcup\left(\bigcup_{i\in I} B_i\right)\subseteq\bigcup_{i\in I}(A_i\bigcup B_i) \tag{2.38}$$

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 2.38 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: 2.123] 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 2.126. (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 $C\setminus(\bigcup_{i\in I} A_i)=\bigcap_{i\in I}(A_i\setminus C)$*
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

$$\begin{aligned}
C \setminus \left(\bigcup_{i\in I} A_i\right) &\underset{[\text{theorem: } 1.24]}{=} C \bigcap \left(\bigcup_{i\in I} A_i\right)^c \\
&\underset{(1)}{=} C \bigcap \left(\bigcap_{i\in I} (A_i)^c\right) \\
&\underset{[\text{theorem: } 2.124]}{=} \bigcap_{i\in I} \left(C \bigcap (A_i)^c\right) \\
&\underset{[\text{theorem: } 1.24]}{=} \bigcap_{i\in I} (C \setminus A_i)
\end{aligned}$$

4. We have

$$\begin{aligned}
C \setminus \left(\bigcap_{i\in I} A_i\right) &\underset{[\text{theorem: } 1.24]}{=} C \bigcap \left(\bigcap_{i\in I} A_i\right)^c \\
&\underset{(2)}{=} C \bigcap \left(\bigcup_{i\in I} (A_i)^c\right) \\
&\underset{[\text{theorem: } 2.124]}{=} \bigcup_{i\in I} \left(C \bigcap (A_i)^c\right) \\
&= \bigcup_{i\in I} (C \setminus A_i)
\end{aligned}$$

□

Theorem 2.127. *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.

$$\begin{aligned}\left(\bigcup_{i\in I} A_i\right)\setminus A &\underset{[\text{theorem: } 1.24]}{=} \left(\bigcup_{i\in I} A_i\right)\bigcap A^c\\ &\underset{[\text{theorem: } 1.30]}{=} A^c\bigcap\left(\bigcup_{i\in I} A_i\right)\\ &\underset{[\text{theorem: } 2.124]}{=} \bigcup_{i\in I}\left(A^c\bigcap A_i\right)\\ &\underset{[\text{theorem: } 1.30]}{=} \bigcup_{i\in I}\left(A_i\bigcap A^c\right)\\ &\underset{[\text{theorem: } 1.24]}{=} \bigcup_{i\in I}(A_i\setminus A)\end{aligned}$$

2.

$$\begin{aligned}\left(\bigcap_{i\in I} A_i\right)\setminus A &\underset{[\text{theorem: } 1.24]}{=} \left(\bigcap_{i\in I} A_i\right)\bigcap A^c\\ &\underset{[\text{theorem: } 1.30]}{=} A^c\bigcap\left(\bigcap_{i\in I} A_i\right)\\ &\underset{[\text{theorem: } 2.124]}{=} \bigcap_{i\in I}\left(A^c\bigcap A_i\right)\\ &\underset{[\text{theorem: } 1.30]}{=} \bigcap_{i\in I}\left(A_i\bigcap A^c\right)\\ &\underset{[\text{theorem: } 1.24]}{=} \bigcap_{i\in I}(A_i\setminus A)\end{aligned}$$

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.

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

□

Theorem 2.128. *Let $\{A_i\}_{i\in I}\subseteq B$ a family then*

1. *If $j\in I$ then $\left(\bigcup_{i\in I\backslash\{j\}}A_i\right)\bigcup A_j=\bigcup_{i\in I}A_i$*
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 $x\in\left(\bigcup_{i\in I\backslash\{j\}}A_i\right)\bigcup A_j$ then either $x\in A_j\subseteq\bigcup_{i\in I}A_i$ [see: 2.121], so that $x\in\bigcup_{i\in I}A_i$ or $x\in\bigcup_{i\in I\backslash\{j\}}A_i\Rightarrow\exists k\in I\backslash\{j\}$ 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 $i\in I\backslash\{j\}$ so that $x\in\bigcup_{i\in I\backslash\{j\}}A_i$ or $i=j$ giving $x\in A_j$, proving that $x\in\left(\bigcup_{i\in I\backslash\{j\}}A_i\right)\bigcup A_j$.
2. As $\{j\in I|A_j\neq\varnothing\}\subseteq I$ we have by [theorem: 2.122] that
$$\bigcup_{i\in\{j\in I|A_j\neq\varnothing\}}A_i\subseteq\bigcup_{i\in I}A_i \tag{2.39}$$
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: 2.39] 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 2.129. *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*

$$\bigcup_{i\in I}A_i=\bigcup_{(i,j)\in I\times J}B_{i,j}$$

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 2.130. *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

$$f\left(\bigcup_{i\in I} A_i\right) \subseteq \bigcup_{i\in I} f(A_i) \tag{2.40}$$

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: 2.40] 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

$$f^{-1}\left(\bigcup_{i\in I} B_i\right) \subseteq \bigcup_{i\in I} f^{-1}(A_i) \tag{2.41}$$

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: 2.41] 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
$$f^{-1}\left(\bigcap_{i\in I}B_i\right)\subseteq\bigcap_{i\in I}f^{-1}(B_i) \tag{2.42}$$
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 2.42 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 2.131. (Product of a family of sets) *Let $\{A_i\}_{i\in I}\subseteq B$ a family then the* ***product of $\{A_i\}_{i\in I}$*** *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 $\forall i\in I$ $f_i=f(i)\in A_i$.

The following shows that the product of a family of only one class is ‘almost‘ that class itself.

Example 2.132. Let $\{A_i\}_{i\in\{0\}}\subseteq\{B\}$ be the family in [example: 2.119] defined by $A:\{0\}\rightarrow\{B\}$ 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: 2.119] we have
$$B=\bigcup_{i\in\{0\}}A_i \tag{2.43}$$
hence
$$\left(\bigcup_{i\in\{0\}}A_i\right)^{\{0\}}=B^{\{0\}} \tag{2.44}$$
Let $f\in B^{\{0\}}\underset{\text{[eq: 2.44]}}{=}\left(\bigcup_{i\in\{0\}}A_i\right)$ 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 $\prod_{i\in\{0\}}A_i\subseteq\left(\bigcup_{i\in\{0\}}A_i\right)^{\{0\}}\underset{\text{[eq: 2.44]}}{=}B^{\{0\}}$ we have that
$$\prod_{i\in\{0\}}A_i=B^{\{0\}}$$

Now by [theorem: 2.75] 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 2.133. *Let* $\{A_i\}_{i\in\{0,1\}}\subseteq\{C,D\}$ *be the family in [example: 2.120] 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: 2.120]: we have that

$$\bigcup_{i\in\{0,1\}}A_i=C\bigcup D \tag{2.45}$$

so that

$$\left(\bigcup_{i\in\{0,1\}}A_i\right)^{\{0,1\}}=\left(C\bigcup D\right)^{\{0,1\}} \tag{2.46}$$

So

$$\prod_{i\in\{0,1\}}A_i=\left\{f|f\in\left(C\bigcup D\right)^{\{0,1\}}\text{ where }f(0)\in C\wedge f(1)\in D\right\} \tag{2.47}$$

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

$$f_{a,b}\subseteq\{0,1\}\times\left(C\bigcup D\right)\wedge f_{a,b}(0)\in C\wedge f_{a,b}(1)\in D \tag{2.48}$$

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: 2.48] this proves that

$$f_{a,b}{:}\,\{0,1\}\rightarrow C\bigcup D\text{ is a partial function} \tag{2.49}$$

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

$$f_{c,d}{:}\,\{0,1\}\rightarrow C\bigcup D\text{ is a function} \tag{2.50}$$

As by [eq: 2.48] $f_{c,d}(0)\in C\wedge f_{c,d}(1)\in D$ proving that

$$f_{c,d}\in\prod_{i\in\{0,1\}}A_i \tag{2.51}$$

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 $y=f_{c,d}\underset{\text{[eq: 2.51]}}{\Rightarrow}$, 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

$$\gamma\subseteq(C\times D)\times\left(\prod_{i\in\{0,1\}}A_i\right) \tag{2.52}$$

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:2.52] proves that

$$\gamma{:}\,C\times D\rightarrow\left(\prod_{i\in\{0,1\}}A_i\right)\text{ is a partial function} \tag{2.53}$$

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: 2.53] we have

$$\gamma : C \times D \rightarrow \left(\prod_{i \in \{0,1\}} A_i \right) \text{ is a function} \tag{2.54}$$

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

$$\gamma : C \times D \rightarrow \left(\prod_{i \in \{0,1\}} A_i \right) \text{ is a surjection}$$ □

Theorem 2.134. *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: 2.123] 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

$$\prod_{i \in I} A_i \subseteq \prod_{i \in I} B_i$$ □

Theorem 2.135. *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: 2.134]

$$\prod_{i \in I} \left(A_i \bigcap B_i \right) \subseteq \prod_{i \in I} A_i \text{ and } \prod_{i \in I} \left(A_i \bigcap B_i \right) \subseteq \prod_{i \in I} B_i$$

so that by [theorem: 1.25]

$$\prod_{i \in I} \left(A_i \bigcap B_i \right) \subseteq \left(\prod_{i \in I} A_i \right) \bigcap \left(\bigcup_{i \in I} B_i \right) \tag{2.55}$$

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

$$\begin{aligned} x : I \rightarrow \bigcup_{i \in I} A_i & \quad \text{is a function} \\ x : I \rightarrow \bigcup_{i \in I} B_i & \quad \text{is a function} \\ \forall i \in I & \quad \text{we have } x(i) \in A_i \bigcap B_i \end{aligned}$$

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

$$x\subseteq I\times\left(\bigcup_{i\in I}(A_i\bigcap B_i)\right)\text{ and }\forall i\in I\text{ we have }x(i)\in A_i\bigcap B_i \tag{2.56}$$

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 $\mathrm{dom}(x)=I$. Combining this with [eq: 2.56] 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 2.55 gives finally

$$\prod_{i\in I}(A_i\bigcap B_i)\subseteq\left(\prod_{i\in I}A_i\right)\bigcap\left(\bigcup_{i\in I}B_i\right) \qquad \square$$

The following theorem is a motivation for the notation A^B for the graphs of functions from B to A.

Theorem 2.136. *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* = ∅.** 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* ≠ ∅.** If $y\in\mathrm{range}(A)$ then $\exists x$ such that $(x,y)\in C_B=I\times\{B\}$, so that $y\in\{B\}$. Hence

$$\mathrm{range}(A)\subseteq\{B\} \tag{2.57}$$

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 $B\in\mathrm{range}(A)$. So if $y\in\{B\}$ then $y=B\in\mathrm{range}(A)$ proving that $\{B\}\subseteq\mathrm{range}(A)$ which combined with [eq: 2.57] gives

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

hence

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

so that

$$\left(\bigcup_{i\in I}A_i\right)^I=B^I \tag{2.58}$$

Now if $f\in B^I$ then $\forall i\in I$ we have $f(i)\in B=A(i)=A_i$ proving that

$$f\in\{f|f\in B^i\wedge\forall i\in I f(i)\in A_i\}\underset{[\text{eq: }2.58]}{=}\left\{f|f\in\left(\prod_{i\in I}A_i\right)^I\wedge\forall i\in I f(i)\in A_i\right\}=\prod_{i\in I}A_i$$

proving that

$$B^I\subseteq\prod_{i\in I}A_i \tag{2.59}$$

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: 2.59] proves that

$$B^I = \prod_{i\in I} A_i$$

□

Theorem 2.137. *Let I, J, B be classes, $f: I\rightarrow J$ a bijection and $\{A_j\}_{j\in J}$ then*

$$\beta: \prod_{j\in J} A_j \rightarrow \prod_{i\in I} A_{f(i)} \text{ where } \beta(x) = x\circ f$$

is a bijection.

Proof. First as $f: I\rightarrow J$ is a bijection, hence surjective, we have by [theorem: 2.113] that

$$\bigcup_{j\in J} A_j = \bigcup_{i\in I} A_{f(i)} \tag{2.60}$$

Let $x\in\prod_{j\in J} A_j$ then $x\in(\bigcup_{j\in J} A_j)^J$, which is equivalent with $x: J\rightarrow\bigcup_{j\in J} A_j$ is a function, and $\forall j\in J$ we have $x(j)\in A_j$. So $x\circ f: I\rightarrow\bigcup_{j\in J} A_j \underset{\text{[eq: 2.60]}}{=} \bigcup_{i\in I} A_{f(i)}$ 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

$$x\circ f\in\prod_{i\in I} A_{f(i)} \tag{2.61}$$

So

$$\beta: \prod_{j\in J} A_j \rightarrow \prod_{i\in I} A_{f(i)}$$

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 $y: I\rightarrow\bigcup_{i\in I} A_{f(i)} \underset{\text{[eq: 2.60]}}{=} \bigcup_{j\in J} A_j$ is a function and $\forall i\in I$ we have $y(i)\in A_{f(i)}$. As $f^{-1}: J\rightarrow I$ is a bijection we have that $y\circ f^{-1}: J\rightarrow\bigcup_{j\in J} A_j$ 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 2.138. *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: 2.102] 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 2.139. *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 $\bigcup_{j \in I} A_j \underset{\text{[theorem: 2.128]}}{=} A_i \bigcup \left(\bigcup_{j \in I \setminus \{i\}} A_j\right)$ we have by [theorem: 2.81] 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

$$y \in \left(\bigcup_{i \in I} A_i\right)^I \tag{2.62}$$

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

$$\forall j \in I \text{ we have } y_j \in A_j \tag{2.63}$$

From [eq: 2.62] and [eq: 2.63] it follows by

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

□

We introduce now the projection operator

Definition 2.140. *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

$$\pi_i \subseteq \left(\prod_{i \in I} A_i\right) \times A_i \tag{2.64}$$

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 $x \in \text{dom}(\pi_i)$ proving that $\prod_{j \in I} A_i \subseteq \text{dom}(\pi_i)$, 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.

$1 \Rightarrow 2$.

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 $A_i = A_j$.

$2 \Rightarrow 3$.

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 $A_i \bigcap A_j = \varnothing$

$3 \Rightarrow 1$.

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: 2.112], 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: 2.91] $R[] : A \rightarrow \mathcal{P}[A]$ 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

$$z\in R[x]\bigcap R[y]\Rightarrow zRx\wedge zRy\underset{\text{symmetry}}{\Rightarrow}xRz\wedge zRy\underset{\text{transitivity}}{\Rightarrow}xRy$$

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 xRx
   b. If xRy 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 yRx.
   c. If $xRy\wedge yRz$ 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 xRz.
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 yRx proving that $y\in R[x]$. So

$$A_i\subseteq R[x]$$

Take $y\in R[x]$ then yRx 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 yRx proving that $y\in R[x]$, hence

$$A_i\subseteq R[x]$$

Take $y\in R[x]$ then yRx 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 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

$$R/X \text{ is a set}$$

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 3.18. *$f: A \rightarrow B$ 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 3.19. *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 $x R_{|B} y \Rightarrow x \in B \wedge y \in B \wedge \mathrm{xRy} \Rightarrow \mathrm{yR}_{|B} x$

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 3.20. *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 3.21. (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** $\boldsymbol{R}$

Proof.

1. As for every $x \in A$ we have the unique $R[x] \in R/X$ it follows from [proposition: 2.91] that

$$f_R: A \rightarrow A/R \text{ is a function}$$

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

$$f_R: A \rightarrow A/R \text{ is surjective}$$

2. We have

$$\begin{array}{rcl} (x, y) \in R & \Leftrightarrow & x R y \\ & \underset{\text{[theorem: 3.11]}}{\Leftrightarrow} & R[x] = R[y] \\ & \Leftrightarrow & f_R(x) = f_R(y) \\ & \Leftrightarrow & (x, y) \in R_{f_R} \end{array}$$

□

We use the above to decompose every function as the composition of a surjection, bijection and injection.

Theorem 3.22. *Let A, B be sets and $f: A \rightarrow B$ a function and define the following functions:*

a) $s_f: A/R_f \rightarrow f(A)$ *where* $s_f = \{(R_f[x], f(x)) | x \in A\}$

b) $i_{f(A)}: f(A) \rightarrow B$ *where* $i_{f(A)} = \{(x, x) | x \in f(A)\}$ *[the inclusion function see [example: 2.53]*

c) $f_{R_f}: A \rightarrow A/R_f$ *where* $f_{R_f}(x) = R_f[x]$ *[theorem: 3.21]*

then

1. $s_f: A/R_f \rightarrow f(A)$ *is a bijection*

2. $i_{f(A)}: f(A) \rightarrow B$ *is a injective function*

3. $f_{R_f}: A \rightarrow A/R_f$ *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: 3.21] we have that

$$i_{f(A)}: f(A) \rightarrow B \text{ is a injective function}$$

and

$$f_{R_f}: A \rightarrow A/R_f \text{ is surjective function}$$

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: 3.20] we have $f(a)=f(a')$. As by [eq: 3.1] $y=f(a)\wedge y'=f(a')$ it follows that $y=y'$. So

$$s_f\colon A/R_f\rightarrow f(A) \text{ is a partial function}$$

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 $x\in\mathrm{dom}(s_f)$. So $A/R_f\subseteq\mathrm{dom}(f)$ which by [proposition: 2.26] proves that

$$s_f\colon A/R_f\rightarrow f(A) \text{ is a function}$$

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: 3.20] 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

$$s_f\colon A/R_f\rightarrow f(A) \text{ is injective} \tag{3.3}$$

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 $y\in\mathrm{range}(s_f)$. So $A/R_f\subseteq\mathrm{range}(s_f)$ which by [proposition: 2.51] proves that

$$s_f\colon A/R_f\rightarrow f(A) \text{ is surjective} \tag{3.4}$$

Combining [eq: 3.3] and [eq: 3.4] it follows that

$$s_f\colon A/R_f\rightarrow f(A) \text{ is a bijection}$$

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 3.23. *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 3.24. (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 3.25. *$\langle A, R\rangle$ 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 3.26. (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 3.27. (Partial ordered class) *$\langle A, R\rangle$ is a **partial ordered class** if A is a class and R is a order.*

Notation 3.28. *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 3.29. *If $\langle A, \leqslant\rangle$ is a pre-ordered or partial class and $x, y, z \in A$ then we define:*

$$\begin{array}{lll} x \leqslant y \leqslant z & \textit{is the same as} & x \leqslant y \wedge y \leqslant z \\ x \leqslant y < z & \textit{is the same as} & x \leqslant y \wedge y < z \\ x < y \leqslant z & \textit{is the same as} & x < y \wedge y \leqslant z \\ x < y < z & \textit{is the same as} & x < y \wedge y < z \end{array}$$

Definition 3.30. *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 3.31. *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

$$\begin{aligned}(x<y\vee x=y) &\Leftrightarrow ((x\leqslant y\wedge x\neq y)\vee x=y)\\ &\Leftrightarrow ((x\leqslant y\vee x=y)\wedge(x\neq y\vee x=y))\\ &\Leftrightarrow x\leqslant y\vee x=y\\ &\Leftrightarrow x\leqslant y\end{aligned}$$

□

Example 3.32. 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 3.33. *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 3.34. *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 3.35. 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 3.36. (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:

$\boldsymbol{x = u}$**.** Then $y \leqslant_B v$ and we have the following possibilities

$\boldsymbol{u = r}$**.** Then $v \leqslant_B s$ so that $y \leqslant_B s$ which as $x = r$ proves that

$$(x, y) \leqslant_{A \times B} (r, s)$$

$\boldsymbol{u \neq r}$**.** Then $u \leqslant_A r \underset{x=u}{\Rightarrow} x \leqslant_A r$ which as $x \neq r$ proves that

$$(x, y) \leqslant_{A \times B} (r, s)$$

$\boldsymbol{x \neq u}$**.** Then $x \leqslant_A u$ and we have the following possibilities

$\boldsymbol{u = r}$**.** 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 3.37. *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 3.38. *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 3.39. *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 3.40. (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 3.41. 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 3.42. *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: 3.34] 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 3.43. *Let $\langle A,\leqslant\rangle$ be a totally ordered class and $B\subseteq A$ then B is a chain [hence by [theorem: 3.42] $\langle B,\leqslant_{|B}\rangle$ 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 3.44. *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: 3.36]. 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 3.45. (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 3.46. *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 3.47. *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 3.48. *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 3.49. *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 3.50. *Let $\langle A, \leqslant_A \rangle$, $\langle B, \leqslant_B \rangle$ be two partial ordered classes, $D \subseteq B$ and*

$f: \langle A, \leqslant_A \rangle \rightarrow \langle D, (\leqslant_B)_{|D} \rangle$ be a order homeomorphism [see theorem: 3.34 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 $f(x) (\leqslant_B)_{|D} f(y) \underset{3.34}{\Rightarrow} f(x) \leqslant_B f(y)$ □

Theorem 3.51. *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: 3.34] 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: 3.34] 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: 3.34] 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: 2.73] 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 3.52. *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.

- $\Rightarrow$. 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

$$f(x) <_B f(y)$$

- $\Leftarrow$. 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 3.53. *If $\langle A, \leqslant_A \rangle$ and $\langle B, \leqslant_B \rangle$ are partially ordered classes and $f: A \rightarrow B$ a bijection then*

$f: \langle A, \leqslant_A \rangle \rightarrow \langle B, \leqslant_B \rangle$ is a order isomorphism $\Leftrightarrow$ $f: \langle A, \leqslant_A \rangle \rightarrow \langle B, \leqslant_B \rangle$ and $f^{-1}: \langle B, \leqslant_B \rangle \rightarrow \langle A, \leqslant_A \rangle$ are increasing functions

Proof. As $f: A \rightarrow B$ is a bijection we have by [theorems: 2.67, 2.71] that $f^{=1}: B \rightarrow A$ is a bijection.

- $\Rightarrow$. 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

$$f(f^{-1}(x)) = (f \circ f^{-1})(x) \underset{\text{[theorem: 2.68}}{=} x \leqslant_B y = (f \circ f^{-1})(y) = f(f^{-1}(y))$$

  which as f is a isomorphism proves that $f^{-1}(x) \leqslant_A f^{-1}(y)$, hence f^{-1} is increasing.

- $\Leftarrow$. 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 3.54. *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: 2.71] we have that $f^{-1}: B \rightarrow A$ is a bijection. By the previous theorem [theorem: 3.53] we have that f^{-1} is increasing. Further as by 2.72 $f = (f^{-1})^{-1}$ and by [theorem: 3.53] f is increasing it follows that $(f^{-1})^{-1}$ is increasing. Using then [theorem: 3.53] it follows that f^{-1} is a isomorphism.
3. This follows from [theorem: 3.51] □

Theorem 3.55. *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: 3.54] □

Theorem 3.56. *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: 3.37] 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 3.57. *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 3.58. *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 3.59. 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 3.60. The empty set $\varnothing$ can not have a maximal, minimal element, greatest element or least element.

Theorem 3.61. *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 $m = m'$. □

Theorem 3.62. *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 3.63. *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 3.64. *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. *$\upsilon(A) = \{x \in X | x \text{ is a upper bound of } A\}$ [the class of upper bound of A]*
6. *$\lambda(A) = \{x \in X | x \text{ is a lower bound of } A\}$ [the class of lower bounds of A]*

Example 3.65. 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 3.66. *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: 3.61], 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: 3.61], element. So $\inf(A)$ is the greatest lower bound [if it exist] and it is itself a lower bound.

Example 3.67. 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: 3.32] 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 3.68. *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: 3.38], 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 3.69. *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 3.70. *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 3.71. *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: 3.70] we have that $\upsilon(B) \subseteq \upsilon(A)$ so that by [lemma: 3.69]
$$\sup(A) = \min(\upsilon(A)) \leqslant \min(\upsilon(B)) = \sup(B)$$
2. Using [lemma: 3.70] we have that $\lambda(B) \subseteq \lambda(A)$ so that by [lemma: 3.69]
$$\inf(B) = \max(\lambda(B)) \leqslant \max(\lambda(A)) = \inf(A)$$
□

Theorem 3.72. *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 3.73. *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 3.74. (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 3.75. *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: 3.73] 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: 3.73] 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 3.76. *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: 3.53] 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 $x\underset{\text{[theorem: 2.69]}}{=}(f^{-1})(f(x))=(f^{-1})(y)\leqslant_A(f^{-1})(u)$, 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 $(f^{-1})(l)\leqslant_A(f^{-1})(y)=(f^{-1})(f(x))\underset{\text{[theorem: 2.69]}}{=}x$, 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 $y \underset{[\text{theorem: } 2.69}{=} f((f^{-1})(y)) = y \in f(\upsilon(A))$. 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 $y \underset{[\text{theorem: } 2.69}{=} f((f^{-1})(y)) = y \in f(\lambda(A))$. 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 3.77. *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: 3.76] 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: 3.76] $\sup(f(f^{-1}(A)))$ 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 3.78. *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 3.79. *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 3.80. *If $\langle X, \leqslant \rangle$ is a partial ordered class, $B \subseteq X$ then for $\langle B, \leqslant_{|B} \rangle$ [see theorem: 3.34] 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 3.81. *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: 3.38] 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 3.82. (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 3.83. *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: 3.81] 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: 3.81] 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 3.84. *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 3.85. *Let $\langle X, \leqslant \rangle$ be a well ordered class and $B \subseteq X$ then*

$$B \textit{ is a section} \Leftrightarrow B = X \textit{ or } B \textit{ is a initial segment of } X \textit{ [definition: 3.45]}$$

Proof.

$\Rightarrow$. 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: 3.45]. 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: 3.81] 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

$$S_{X,l} = B$$

$\Leftarrow$. 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 3.86. (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. □

Lemma 3.87. *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: 3.81] 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: 3.52] $f(f(c)) < f(c)$ 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 $x \leqslant f(x)$. □

Theorem 3.88. *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: 3.87] 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. □

Corollary 3.89. *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: 3.88] □

Theorem 3.90. *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: 2.73],[theorem: 3.51] 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: 3.56] 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: 3.88] this is impossible so the assumption is false, hence Y is not order isomorphic to a an initial segment of Y. □

Corollary 3.91. *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: 3.55] $Y \cong X$.

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: 3.90] 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: 3.90] is not allowed. □

Lemma 3.92. *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 $\mathcal{S}_{X,b}$ [using the order $\leqslant_{|S_{X,y}}$]*

Proof. First if $x \in S_{X,a}$ then $x < a \underset{\text{a<b}}{\Rightarrow} x < b$ 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: 3.85], $S_{X,a}$ is a initial segment of $S_{X,b}$. □

Theorem 3.93. *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: 3.55]

$$S_{Y,y} \cong S_{Y,y'} \tag{3.21}$$

Assume that $y \neq y'$ then, as $\langle Y, \leqslant_Y\rangle$ is well ordered we have by [theorem: 3.81] either:

- ***$y \leqslant y'$.*** then $y < y'$ so that by the previous lemma [lemma: 3.92] we have that $S_{Y,y}$ is a initial segment of $S_{Y,y'}$. Using [corollary: 3.89] we have then that $S_{Y,y'}$ is not order isomorphic with $S_{Y,y}$ contradicting [eq: 3.21].
- ***$y' < y$.*** then by the previous lemma [lemma: 3.92] we have that $S_{Y,y'}$ is a initial segment of $S_{Y,y}$. Using [corollary: 3.89] 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: 3.55] 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: 3.81] either:

- ***$x \leqslant x'$.*** then $x < x'$ so that by the previous lemma [lemma: 3.92] we have that $S_{X,x}$ is a initial segment of $S_{X,x'}$. Using [corollary: 3.89] we have then that $S_{X,x'}$ is not order isomorphic with $S_{X,x}$ contradicting [eq: 3.24].
- ***$x' \leqslant x$.*** then by the previous lemma [lemma: 3.92] we have that $S_{X,x'}$ is a initial segment of $S_{X,x}$. Using [corollary: 3.89] 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: 3.81] that $F(y) <_Y F(x)$. So using [theorem: 3.92] 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: 3.46]. 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: 3.90] 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: 3.56] 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: 3.46] that $S_{X,x} \subseteq S_{X,c}$. Hence by 2.87 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 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: 2.83] 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: 3.81] 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 $S_{X,c} \cong S_{Y,d} \underset{\text{[theorem: 3.55]}}{\Rightarrow} S_{Y,d} \cong S_{X,c}$. 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: 3.46] $S_{Y,y} \subseteq S_{Y,d}$. Hence by 2.87 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 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: 2.83] 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: 3.81] 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: 3.85] 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: 3.85] 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:

$\boldsymbol{X = C \wedge Y = D.}$ then by [eq: 3.43]

$$X \cong Y$$

Using theorem [theorem: 3.91] and the above we have that

$$X \text{ is not order isomorphic with a sub-class of Y}$$

$$Y \text{ is not order isomorphic with a sub-class of X}$$

$\boldsymbol{X = C \wedge Y \neq D.}$ 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: 3.90] we have that X is not order isomorphic to a subset of Y contradicting $X \cong D$ and $X \cong Y$. So

$$Y \text{ is not order isomorphic to } a \text{ initial segment of } X$$

$$X \text{ is not order isomorphic to } Y$$

$\boldsymbol{X \neq C \wedge Y = D.}$ then by [eq: 3.45] we have that C is a initial segment of X, which as by [eq: 3.43] $C \cong D \underset{\text{[theorem: 3.55]}}{\Rightarrow} Y = D \cong C$ 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: 3.90] we have that Y is not order isomorphic to a subset of X contradicting $Y \cong C$ and $Y \cong X$. So

$$X \text{ is not order isomorphic to } a \text{ initial segment of } Y$$

$$X \text{ is not order isomorphic to } Y$$

$\boldsymbol{X \neq C \wedge Y \neq D}$**.** 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 3.94. *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: 3.80], so using the previous [theorem: 3.93] 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: 3.90] we may not have that Y is order isomorphic with a sub-class of X. As by [theorem: 3.55] $Y \cong Y$ 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 3.95. *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 3.96. *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 3.97. (Choice Function) *Let A be a set then a* ***choice function for A*** *is a function $f: \mathcal{P}'(A) \rightarrow A$ 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 3.98. (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 3.99. *If $f: A \rightarrow B$ is a surjective function then there exists a injective function $g: B \rightarrow A$ such that $f \circ g = \mathrm{Id}_B$*

Proof. By the axiom of choice there exists a choice function

$$c:\mathcal{P}'(A)\rightarrow A \text{ such that } \forall A\in\mathcal{P}'(A) \text{ we have } c(A)\in A$$

If $f: A\rightarrow B$ 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

$$g: B\rightarrow Y \text{ by } g(y)=c(f^{-1}(\{y\}))$$

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

$$g: B\rightarrow Y \text{ is injective}$$ □

The important thing to remember in the above is that the axiom of choice ensures the existence of $g: B\rightarrow A$ but does not give a way to construct the function g itself.

We have the following equivalent statements of the axiom of choice

Theorem 3.100. *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 $\mathcal{A}$** 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 $f: I\rightarrow\bigcup_{i\in I}A_i$ 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: 3.98] to get a function

$$c:\mathcal{P}'(U)\rightarrow U \text{ such that } \forall A\in\mathcal{P}'(U) \text{ we have } c(A)\in A$$

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: 3.96] 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

$$C: \mathcal{P}'(A) \rightarrow A \text{ is a function} \tag{3.57}$$

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

$$C: \mathcal{P}'(A) \rightarrow A \text{ is a choice function}$$

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: 2.114] it follows that $\bigcup_{i\in I}A_i$ is a set. Using the Axiom of Choice [axiom: 3.98] 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: 2.121] 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 $\mathrm{range}(A)\subseteq\mathcal{P}'(\bigcup_{i\in I}A_i)$. 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: 3.96] 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). □

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 3.101. *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*

$$\pi_i:\prod_{j\in I}A_j\to A_i\textit{ defined by }\pi_j(x)=x(j)\textit{ [see definition: 2.140]}$$

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: 2.102] then $\forall j\in I\setminus\{i\}$ we have $A_j\neq\varnothing$. So we can use [theorem: 3.100 (3)] to find a function

$$f: I\setminus\{i\}\to\bigcup_{j\in I\setminus\{i\}}A_j\text{ such that }\forall j\in I\setminus\{i\}\text{ we have }f(j)\in A_j$$

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 $g: I\to\bigcup_{j\in I}A_j$ 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: 2.139] 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 3.102. *Let $\{A\}_{i\in I}\subseteq\mathcal{A}$ be a family of sets where $I,\mathcal{A}$ are sets then we have*

$$\prod_{i\in I}A_i\neq\varnothing\Leftrightarrow\forall i\in I\textit{ we have }A_i\neq\varnothing$$

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: 3.100 (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.

Corollary 3.103. *Let $\{A\}_{i \in I} \subseteq \mathcal{A}$ be a family of sets where $I, \mathcal{A}$ are sets then we have*

$$\prod_{i \in I} A_i = \varnothing \Leftrightarrow \exists i \in I \text{ such that } A_i = \varnothing$$

Proof. We proceed by contradiction to prove this

- **⇒.** Assume that $\forall i \in I$ we have that $A_i \neq \varnothing$ then by [theorem: 3.102] $\prod_{i \in I} A_i \neq \varnothing$ contradicting $\prod_{i \in I} A_i = \varnothing$. So the assumption is false or $\exists i \in I$ such that $A_i = \varnothing$.

- **⇐.** Assume that $\prod_{i \in I} A_i \neq 0$ then by [theorem: 3.102] we have $\forall i \in I$ that $A_i \neq \varnothing$ contradicting $\exists i \in I$ such that $A_i = 0$. Hence we must have $\prod_{i \in I} A_i = \varnothing$. □

The Axiom of Choice has also import consequences for partial ordered sets.

Theorem 3.104. *Let $\langle X, \leqslant \rangle$ be a partial ordered **set** such that:*

1. *X has a least element p*
2. *Every chain [see definition:3.40] of X has a supremum*

then there is a element $x \in X$ which has no immediate successor [see definition: 3.82]

Proof. We prove this by contradiction, so assume that $\forall x \in X$ there exist a immediate successor. Given $x \in X$ define $T_x = \{y | y \text{ is a immediate successor of } x\}$ then $T_x \neq \varnothing$ so that $T_x \in \mathcal{P}'(X)$. Using the Axiom of Choice [axiom: 3.98] 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: 2.91] to define the function

$$\mathrm{succ}: X \rightarrow X \text{ by } \mathrm{succ}(x) = c(T_x).$$

If $x \in X$ then $\mathrm{succ}(x) = c(T_x) \in T_x$ so that $\mathrm{succ}(x)$ is a immediate successor of x, to summarize

$$\mathrm{succ}: X \rightarrow X \text{ is a function such that } \forall x \in X \ \mathrm{succ}(x) \text{ is a immediate successor of } x \tag{3.60}$$

Before we can reach the contradiction we need to have some definitions and sub lemmas.

Definition 3.105. *$A \subseteq X$ is a **p-sequence** iff*

1. *$p \in A$*
2. *If $x \in A$ then $\mathrm{succ}(x) \in A$*
3. *If $C \subseteq A$ is a chain then $\sup(C) \in A$ [note that by hypothesis (2) $\sup(C)$ exist]*

Note 3.106. 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 3.107. *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: 3.107] we have that $\bigcap \{A \in \mathcal{P}(X) | A \text{ is a p-sequence}\}$ is a p-sequence and by definition $p \in \bigcap \{A \in \mathcal{P}(X) | A \text{ is a p-sequence}\}$. 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 3.108. *A element $x \in P$ is **select** if x is comparable with every element in P.*

Lemma 3.109. *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: 3.105 (2)] that $\operatorname{succ}(y) \in P$. Now as x is select we have that $x, \operatorname{succ}(y)$ are comparable, hence by [theorem: 3.38] 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 3.110. *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: 3.109] 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

$$A_x \text{ is a p-sequence} \qquad \square$$

Corollary 3.111. *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: 3.110] 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

$$P = A_x \qquad \square$$

Lemma 3.112. *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: 3.111] 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: 3.111] 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. □

Now for the last corollary in the proof.

Corollary 3.113. *P is a chain*

Proof. As by the previous lemma [lemma: 3.112] $\{x \in X | x \text{ is select}\}$ 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. □

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: 3.105 (3)] that $\sup(P) \in P$ and by [definition: 3.105 (2)] that $\operatorname{succ}(\sup(P)) \in P$ so that $\operatorname{succ}(\sup(P)) \leqslant \sup(P)$. As $\operatorname{succ}(\sup(P))$ is the immediate successor of $\sup(P)$ we have that $\sup(P) < \operatorname{succ}(\sup(P))$. Hence $\sup(P) < \sup(P)$ which is a contradiction. □

This was a long proof but it will be used in the following important theorem.

Definition 3.114. *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 3.115. (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: 3.32] 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: 3.41] 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:

$\boldsymbol{A \subseteq B}$**.** Then $x, y \in B$ which as B is a chain [using $\leqslant$] means that x, y are comparable [using the order $\leqslant$]

$\boldsymbol{B \subseteq A}$**.** 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: 3.67] 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: 3.104] 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. □

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 3.116. (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 3.117. *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: 3.114] 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: 3.57] 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 3.118. *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 3.78]*

Proof. Just like the proof of [theorem: 3.104] 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 $x \in B \wedge y \in B \backslash B \underset{\text{[theorem: 1.32]}}{=} \varnothing$ 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 3.119. *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 $x\in B'\backslash B\wedge b\in B$. 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 3.120. *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 $y\in B_{\mathcal{C}}\backslash B$ then $\exists(B',R')\in\mathcal{C}$ such that $y\in B'$ or as $y\in B_{\mathcal{C}}\backslash B$ that

   $$y\in B'\backslash B$$

   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 $y\in B_{\mathcal{C}}\backslash B$. So we have

   $$(B,R)\preccurlyeq(B',R')$$

   As $x\in B$ and $y\in B'\backslash B$ 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: 3.116] together with the above lemma [lemma: 3.120] 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

$$x\in X\backslash B_m\Rightarrow x\notin B_m.$$

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

$$\left\langle B_m\bigcup\{x\},R^*\right\rangle\text{ is well ordered}$$ □

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

$$X = B_m$$ □

As $\langle B_m, R_m \rangle$ is a well ordered the above proves that there exists a partial order R_m such that

$$\langle X, R_m \rangle = \langle B_m, R_m \rangle \text{ is well-ordered [by definition of } \mathcal{A}\ B_m \text{ is well ordered]}$$ □

We show now that Well Ordering implies the Axiom of Choice.

Theorem 3.121. *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 $c = \{(A, x) | A \in \mathcal{P}'(X) \wedge x \text{ is a least element of } A\}$. 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: 3.61] 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 $\mathcal{P}'(A) \subseteq \operatorname{dom}(c)$. 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 3.122. *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: 3.115]

2 ⇒ 3. This follows from [theorem: 3.117]

3 ⇒ 4. This follows from [theorem: 3.118]

4 ⇒ 1. This follows from [theorem: 3.121] □

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: 3.100] that the following statements are true.

Theorem 3.123.

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 change the definition of maximal element slightly.

Theorem 3.124. *Let $\langle X, \leqslant \rangle$ be a pre-ordered set [see definitions: 3.25, 3.24] 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: 3.33] 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: 3.57] 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 $x \leqslant m$. □

As a interesting application of the Axiom of Choice we prove that every function can be restricted to a injection or bijection.

Theorem 3.125. *Let X, Y be sets, $f: X \rightarrow Y$ a function then there exist a $Z \subseteq X$ such that:*

1. *$f_{|Z}: Z \rightarrow Y$ is a injection*
2. *$f_{|Z}(Z) = f(X)$*
3. *$f_{|Z}: Z \rightarrow f(X)$ 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: 3.98] there exist a function

$$c: \mathcal{P}'(X) \rightarrow X \text{ such that } \forall A \in \mathcal{P}'(X)\ (c)(A) \in A$$

Take

$$Z = c(\mathcal{A}) \subseteq X$$

and consider the restriction of f to Z

$$f_{|Z}: Z \rightarrow Y$$

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. □

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: 2.91], [notation: 2.92], [theorem: 2.112] and [theorem: 2.117] without explicit referring to them. This to avoid excessive notation and difference of notation between this text and standard mathematical practice. Another simplification of notation that we introduce is the following.

Notation 3.126. *If $f: A \times B \rightarrow C$ is a function then $f((x, y))$ is noted as $f(x, y)$*

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*

$$f: X \times X \rightarrow X$$

To avoid using excessive notation we use infix notation instead of the classic function call notation, so

$$f(x, y) \textit{ is noted as } x f y$$

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 $\odot: G \times G \rightarrow G$ 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: $X \rightarrow$ and $g: X \rightarrow X$ are functions, so that by [theorem: 2.28] $f \circ g: X \rightarrow X$ is a function, hence $f \circ g \in X^X$. So

$$\circ: X^X \times X^X \rightarrow X^X \text{ defined by } \circ(f, g) = f \circ g$$

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

$$y = y \odot e = y \circ (x \odot y') = (y \odot x) \odot y' = e \odot y' = y'$$

□

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

$$\begin{aligned} x &= x\odot e\\ &= x\odot(x^{-1}\odot(x^{-1})^{-1})\\ &= (x\odot x^{-1})\odot(x^{-1})^{-1}\\ &= e\circ(x^{-1})^{-1}\\ &= (x^{-1})^{-1}\end{aligned}$$

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: 2.73] $f \circ g$ 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: 2.68,2.71] 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: 2.71] that $f^{-1}: G \rightarrow F$ is a bijection. Take $x, y \in G$ then we have

$$\begin{array}{rcl} f^{-1}(x \oplus y) & = & f^{-1}(\mathrm{Id}_G(x) \oplus \mathrm{Id}_G(y)) \\ & \underset{\text{[theorem: 2.68}}{=} & f^{-1}((f \circ f^{-1})(x) \oplus (f \circ f^{-1})(y)) \\ & \underset{\text{[theorem: 2.42]}}{=} & f^{-1}(f(f^{-1}(x)) \oplus f(f^{-1}(y))) \\ & \underset{f \text{ is homeomorphism}}{=} & f^{-1}(f(f^{-1}(x) \odot f^{-1}(y))) \\ & \underset{\text{[theorem: 2.42]}}{=} & (f^{-1} \circ f)(f^{-1}(x) \odot f^{-1}(y)) \\ & \underset{\text{[theorem: 2.68}}{=} & \mathrm{Id}_F(f^{-1}(x) \odot f^{-1}(y)) \\ & = & f^{-1}(x) \odot f^{-1}(y) \end{array}$$

Further if e_F, e_G are the neutral elements of $\langle F, \oplus_F\rangle$, $\langle G, \oplus_G\rangle$ then

$$\begin{array}{rcl} e_F & = & \mathrm{Id}_F(e_f) \\ & = & (f^{-1} \circ f)(e_F) \\ & = & f^{-1}(f(e_F)) \\ & = & f^{-1}(e_G) \end{array}$$

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

$$\begin{array}{ccc} (g \circ f)(x \oplus_A y) & = & g(f(x \oplus_A y)) \\ & \underset{f \text{ is a homeomorphism}}{=} & g(f(x) \oplus_B f(y)) \\ & \underset{g \text{ is a homeomorphism}}{=} & g(f(x)) \oplus_C g(f(y)) \\ & = & (g \circ f)(x) \oplus_C (g \circ f)(y) \end{array}$$

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: 2.73] 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 $i \in I$

$$\begin{array}{ccc} (x \odot (y \odot z))(i) & = & x(i) \odot_i (y \odot z)(i) \\ & = & x(i) \odot_i (y(i) \odot_i z(i)) \\ & \underset{\langle A_i, \odot_i \rangle \text{ is a semi group}}{=} & (x(i) \odot y(i)) \odot z(i) \\ & = & (x \odot y)(i) \odot_i z(i) \\ & = & ((x \odot y) \odot z)(i) \end{array}$$

so that

$$x \odot (y \odot z) = (x \odot y) \odot z$$

neutral element. Let $x \in \prod_{i \in I} A_i$ then $\forall i \in I$

$$\begin{array}{ccc} (x \odot e)(i) & = & x(i) \odot_i e(i) \\ & = & x(i) \odot_i e_i \\ & \underset{\langle A_i, \odot_i \rangle \text{ is a semi group}}{=} & x(i) \\ (e \odot x)(i) & = & e(i) \odot_i x(i) \\ & = & e_i \odot_i x(i) \\ & \underset{\langle A_i, \odot_i \rangle \text{ is a semi group}}{=} & x(i) \end{array}$$

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

$$\begin{array}{ccc} (x \odot x^{-1})(i) & = & x(i) \odot_i (x^{-1})(i) \\ & = & x(i) \odot_i (x_i)^{-1} \\ & = & x_i \odot_i (x_i)^{-1} \\ & \underset{\langle A_i, \odot_i \rangle \text{ is a group}}{=} & e_i \\ & = & e(i) \\ (x^{-1} \odot x)(i) & = & (x^{-1})(i) \odot_i x(i) \\ & = & (x_i)^{-1} \odot_i x(i) \\ & = & (x_i)^{-1} \odot_i x_i \\ & = & e_i \\ & = & e(i) \end{array}$$

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 $\rhd: G \times X \to X$ where $\rhd(g, x) \underset{\text{notation}}{=} g \rhd x$ such that*
   a. *$\forall x \in X$ we have $e \rhd x = x$*
   b. *$\forall g, g' \in G$ and $\forall x \in X$ we have $(g \odot g') \rhd x = g \rhd (g' \rhd x)$*
2. *A **right group action** is a function $\lhd: X \times G \to X$ 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*

$$g_{\rhd}: X \to X \text{ by } g_{\rhd}(x) = g \rhd x$$

Definition 4.29. *Let $\langle G, \odot \rangle$ be a group, X, $\lhd$ a right group action and $g \in G$ then we define*

$$g_{\lhd}: X \to X \text{ by } g_{\lhd}(x) = x \lhd g$$

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 $x \in R \setminus \{0\}$ so that $\exists y \in R \setminus \{0\}$ 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*

$$f: \langle A, \oplus_A, \odot_A\rangle \rightarrow \langle B, \oplus_B, \odot_B\rangle \textit{ is a ring homeomorphism}$$

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*

$$f: \langle F, \oplus_F, \odot_F\rangle \rightarrow \langle H, (\oplus_G)_{|H \times H}, (\odot_G)_{|H \times H}\rangle \textit{ is a ring homeomorphism}$$

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

$$1_B = f(1_A) \qquad \square$$

Definition 4.46. *If* $\langle A, \oplus_A, \odot_A \rangle$ *and* $\langle B, \oplus_B, \odot_B \rangle$ *are rings then a function* $f: A \rightarrow B$ *is a ring isomorphism if it is a bijection and*

$$f: \langle A, \oplus_A, \odot_A \rangle \rightarrow \langle B, \oplus_B, \odot_B \rangle \text{ is a ring homeomorphism}$$

Theorem 4.47. *If* $\langle A, \oplus_A, \odot_A \rangle$ *and* $\langle B, \oplus_B, \odot_B \rangle$ *are rings and*

$$f: \langle A, \oplus_A, \odot_A \rangle \rightarrow \langle B, \oplus_B, \odot_B \rangle \text{ a ring homeomorphism [or ring isomorphism]}$$

then

$$f: \langle A, \oplus_A \rangle \rightarrow \langle B, \oplus_B \rangle \text{ is a group homeomorphism [or group isomorphism]}$$

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: 4.19, 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*

   $f: \langle A, \oplus_A, \odot_A \rangle \rightarrow \langle D, \oplus_B, \odot_B \rangle$ *and* $g: \langle B, \oplus_B, \odot_B \rangle \rightarrow \langle C, \oplus_C, \odot_C \rangle$ *are ring homeomorphisms*

   then

   $$g \circ f: \langle A, \oplus_A, \odot_A \rangle \rightarrow \langle C, \oplus_C, \odot_C \rangle \text{ is a ring homeomorphism}$$

   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*

   $f: \langle A, \oplus_A, \odot_A \rangle \rightarrow \langle D, \oplus_B, \odot_B \rangle$ *and* $g: \langle B, \oplus_B, \odot_B \rangle \rightarrow \langle C, \oplus_C, \odot_C \rangle$ *are ring isomorphisms*

   then

   $$g \circ f: \langle A, \oplus_A, \odot_A \rangle \rightarrow \langle g(f(A)), \oplus_C, \odot_C \rangle \text{ is a ring isomorphism}$$

   or as $g(f(A)) \underset{f: A \rightarrow D \text{ is a bijection}}{=} g(D)$

   $$g \circ f: \langle A, \oplus_A, \odot_A \rangle \rightarrow \langle g(f(A)), \oplus_C, \odot_C \rangle \text{ is a ring isomorphism}$$

Proof.

1. Let $x, y \in A$ then

$$\begin{aligned}
(g \circ f)(x \oplus_A y) &= g(f(x \oplus_A y)) \\
&\underset{f \text{ is a homeomorphism}}{=} g(f(x) \oplus_B f(y)) \\
&\underset{g \text{ is a homeomorphism}}{=} g(f(x)) \oplus_C g(f(y)) \\
&= (g \circ f)(x) \oplus_C (g \circ f)(y) \\
(g \circ f)(x \odot_A y) &= g(f(x \odot_A y)) \\
&\underset{f \text{ is a homeomorphism}}{=} g(f(x) \odot_B f(y)) \\
&\underset{g \text{ is a homeomorphism}}{=} g(f(x)) \odot_C g(f(y)) \\
&= (g \circ f)(x) \odot_C (g \circ f)(y) \\
(g \circ f)(1_A) &= g(f(1_A)) \\
&\underset{f \text{ is a homeomorphism}}{=} g(1_B) \\
&\underset{g \text{ is a homeomorphism}}{=} 1_C
\end{aligned}$$

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: 2.74] we have that $g \circ f : A \rightarrow g(f(A))$ is a bijection which combined with (1) proves that $g \circ f : \langle A, \oplus_A, \odot_A \rangle \rightarrow \langle g(f(A)), \oplus_C, \odot_C \rangle$ 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.

   $\Rightarrow$. This follows directly from the definition of a ordered ring.

   $\Leftarrow$. 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.

$$\begin{aligned} x \leqslant y \quad & \Leftrightarrow \quad x = y \vee x < y \\ & \underset{(1)}{\Leftrightarrow} \quad x = y \vee x \oplus z < y \oplus z \\ & \underset{x = y \Leftrightarrow x \oplus z = y \oplus z}{\Leftrightarrow} \quad x \oplus z = y \oplus z \vee x \oplus z < y \oplus z \\ & \Leftrightarrow \quad x \oplus z \leqslant y \oplus z \end{aligned}$$

3.

$$\begin{aligned} x < y &\underset{(1)}{\Leftrightarrow} x \oplus (-x) < y \oplus (-x) \\ &\Leftrightarrow 0 < y \oplus (-x) \end{aligned}$$

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:

$x = y$. Then $x \odot z = y \odot z$ contradicting $x \odot z < y \odot z$ so this case never occurs.

$y < x$. 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.

$x < y$. 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:

$x = y$. Then $x \odot z = y \odot z$ hence $x \odot z \leqslant y \odot z$

$x < y$. 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 $x \odot z \leqslant y \odot z$.

$\Leftarrow$. 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:

   $z = 0$. 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 $x \odot z \leqslant y \odot z$

   $0 < z$. 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. *$\odot: F\times F\rightarrow F$ 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

$$y=y\odot 1=y\circ(x\odot y')=(y\odot x)\odot y'=1\odot y'=y' \qquad \square$$

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

$$\begin{aligned}
1=x^{-1}\odot(x^{-1})^{-1} &\Rightarrow & x\odot 1=x\odot(x^{-1}\odot(x^{-1})^{-1})\\
&\underset{\text{neutral element}}{\Rightarrow} & x=x\odot(x^{-1}\odot(x^{-1})^{-1})\\
&\underset{\text{associativity}}{\Rightarrow} & x=(x\odot x^{-1})\odot(x^{-1})^{-1}\\
&\underset{\text{inverse element}}{\Rightarrow} & x=1\odot(x^{-1})^{-1}\\
&\underset{\text{neutral element}}{\Rightarrow} & x=(x^{-1})^{-1}
\end{aligned}$$

$\square$

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

- **⇒.** If $x=y$ then trivially $x^{-1}=y^{-1}$
- **⇐.** Then we have

$$\begin{aligned}
x^{-1}=y^{-1} &\Rightarrow x^{-1}\odot y=y^{-1}\odot y\\
&\underset{\text{inverse element}}{\Rightarrow} x^{-1}\odot y=1\\
&\Rightarrow x\odot(x^{-1}\odot y)=x\odot 1\\
&\underset{\text{neutral element}}{\Rightarrow} x\odot(x^{-1}\odot y)=x\\
&\underset{\text{associativity}}{\Rightarrow} (x\odot x^{-1})\odot y=x\\
&\underset{\text{inverse element}}{\Rightarrow} 1\odot y=x\\
&\underset{\text{neutral element}}{\Rightarrow} y=x
\end{aligned}$$

□

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
  - $\boldsymbol{x=0}$**.** Then clearly $0\vee y=0$ is true.
  - $\boldsymbol{x\neq 0}$**.** 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 $f: A \rightarrow B$ 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*

$$f:\langle A,\odot_A,\oplus_A\rangle\rightarrow\langle B,\odot_B,\oplus_B\rangle \textit{ is a field homeomorphism}$$

Note 4.65. If $\langle A,\odot_A,\oplus_A\rangle$ and $\langle B,\odot_B,\oplus_B\rangle$ are fields and

$$f:\langle A,\odot_A,\oplus_A\rangle\rightarrow\langle B,\odot_B,\oplus_B\rangle \text{ a field homeomorphism}$$

then

$$f:\langle A,\odot_A,\oplus_A\rangle\rightarrow\langle B,\odot_B,\oplus_B\rangle \text{ is a ring homeomorphism}$$

and

$$f:\langle A,\oplus_A\rangle\rightarrow\langle B,\oplus_B\rangle \text{ and } f:\langle A,\odot_A\rangle\rightarrow\langle B,\odot_B\rangle \text{ are group homeomorphisms}$$

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

$$f:\langle A,\odot_A,\oplus_A\rangle\rightarrow\langle B,\odot_B,\oplus_B\rangle \text{ is a ring homeomorphism}$$

Finally using [note: 4.43] we have that

$$f:\langle A,\oplus_A\rangle\rightarrow\langle B,\oplus_B\rangle \text{ and } f:\langle A,\odot_A\rangle\rightarrow\langle B,\odot_B\rangle \text{ are group homeomorphisms}$$ □

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*

$$f:\langle A,\oplus_A\rangle\rightarrow\langle B,\oplus_B\rangle \textit{ and } f:\langle A,\odot_A\rangle\rightarrow\langle B,\odot_B\rangle$$

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*

$$f:\langle A,\oplus_A,\odot_A\rangle\rightarrow\langle B,\oplus_B,\odot_B\rangle \textit{ a field homeomorphism [or field isomorphism]}$$

then

$$f:\langle A,\oplus_A,\odot_A\rangle\rightarrow\langle B,\oplus_B,\odot_B\rangle \textit{ a ring homeomorphism [or ring isomorphism]}$$

and

$$f:\langle A,\oplus_A\rangle\rightarrow\langle B,\oplus_B\rangle \textit{ is a group homeomorphism [or group isomorphism]}$$

and

$$f:\langle A,\odot_A\rangle\rightarrow\langle B,\odot_B\rangle \textit{ is a group homeomorphism [or group isomorphism]}$$

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*

$$f:\langle F,\oplus_F,\odot_F\rangle\rightarrow\langle H,(\oplus_G)_{|H\times H},(\odot_G)_{|H\times H}\rangle \textit{ is a field homeomorphism}$$

then

$$f:\langle F,\oplus_F,\odot_F\rangle\rightarrow\langle G,\oplus_G,\odot_G\rangle \textit{ is a field homeomorphism}$$

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: 2.71] we have that $f^{-1}: B \rightarrow A$ is a bijection. Further we have:

1. Take $x, y \in B$ then we have

$$\begin{aligned}
f^{-1}(x \oplus_B y) &= f^{-1}(\mathrm{Id}_B(x) \oplus_B \mathrm{Id}_B(y)) \\
&\underset{\text{[theorem: 2.68}}{=} f^{-1}((f \circ f^{-1})(x) \oplus_B (f \circ f^{-1})(y)) \\
&\underset{\text{[theorem: 2.42]}}{=} f^{-1}(f(f^{-1}(x)) \oplus_B f(f^{-1}(y))) \\
&\underset{f \text{ is homeomorphism}}{=} f^{-1}(f(f^{-1}(x) \oplus_A f^{-1}(y))) \\
&\underset{\text{[theorem: 2.42]}}{=} (f^{-1} \circ f)(f^{-1}(x) \oplus_A f^{-1}(y)) \\
&\underset{\text{[theorem: 2.68}}{=} \mathrm{Id}_A(f^{-1}(x) \oplus_A f^{-1}(y)) \\
&= f^{-1}(x) \oplus_A f^{-1}(y)
\end{aligned}$$

2. Take $x, y \in B$ then we have

$$\begin{aligned}
f^{-1}(x \odot_B y) &= f^{-1}(\mathrm{Id}_B(x) \odot_B \mathrm{Id}_B(y)) \\
&\underset{\text{[theorem: 2.68}}{=} f^{-1}((f \circ f^{-1})(x) \odot_B (f \circ f^{-1})(y)) \\
&\underset{\text{[theorem: 2.42]}}{=} f^{-1}(f(f^{-1}(x)) \odot_B f(f^{-1}(y))) \\
&\underset{f \text{ is homeomorphism}}{=} f^{-1}(f(f^{-1}(x) \odot_A f^{-1}(y))) \\
&\underset{\text{[theorem: 2.42]}}{=} (f^{-1} \circ f)(f^{-1}(x) \odot_A f^{-1}(y)) \\
&\underset{\text{[theorem: 2.68}}{=} \mathrm{Id}_A(f^{-1}(x) \odot_A f^{-1}(y)) \\
&= f^{-1}(x) \odot f^{-1}(y)
\end{aligned}$$

3. From $f(1_A) = 1_B$ it follows that

$$f^{-1}(1_B) = f^{-1}(f(1_B)) = (f^{-1} \circ f)(1_B) \underset{\text{[theorem: 2.68}}{=} \mathrm{Id}_A(1_B) = 1_B$$

□

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

$$\begin{array}{rcl} f(-x)\oplus_B f(x) & = & f((-x)\oplus_A x)\\ & = & f(0_A)\\ & \underset{(1)}{=} & 0_B\\ f(x)\oplus_B f(-x) & = & f(x\oplus_A(-x))\\ & = & f(0_A)\\ & \underset{(1)}{=} & 0_A \end{array}$$

so that

$$f(-x)=-f(x)$$

3. If $a\in A$ with $a\neq 0_A$ then

$$\begin{array}{rcl} f(a^{-1})\odot_B f(a) & = & f(a^{-1}\odot_A a)\\ & = & f(1_A)\\ & = & 1_B\\ f(a)\odot_B f(a^{-1}) & = & f(a\odot_A a^{-1})\\ & = & f(1_A)\\ & = & 1_B \end{array}$$

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*

   $f:\langle A,\oplus_A,\odot_A\rangle\rightarrow\langle D,\oplus_B,\odot_B\rangle$ and $g:\langle B,\oplus_B,\odot_B\rangle\rightarrow\langle C,\oplus_C,\odot_C\rangle$ are field homeomorphism then

   $$g\circ f:\langle A,\oplus_A,\odot_A\rangle\rightarrow\langle C,\oplus_C,\odot_C\rangle \text{ is a field homeomorphism}$$

   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*

   $f:\langle A,\oplus_A,\odot_A\rangle\rightarrow\langle B,\oplus_B,\odot_B\rangle$ and $g:\langle B,\oplus_B,\odot_B\rangle\rightarrow\langle C,\oplus_C,\odot_C\rangle$ are field isomorphisms

then

$$g\circ f\colon\langle A,\oplus_A,\odot_A\rangle\rightarrow\langle g(f(A)),\oplus_C,\odot_C\rangle \text{ is a field isomorphism}$$

or as $g(f(A))\underset{f:A\rightarrow D \text{ is a bijection}}{=}g(D)$

$$g\circ f\colon\langle A,\oplus_A,\odot_A\rangle\rightarrow\langle g(D),\oplus_C,\odot_C\rangle \text{ is a field isomorphism}$$

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: 2.74] we have that $g\circ f\colon A\rightarrow g(f(A))$ 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

$$\begin{aligned} x<y &\underset{(7)}{\Rightarrow} x\odot x^{-1}<y\odot x^{-1}\\ &\Rightarrow 1<y\odot x^{-1}\\ &\Rightarrow 1<x^{-1}\odot y\\ &\underset{(7)}{\Rightarrow} 1\odot y^{-1}<(x^{-1}\odot y)\odot y^{-1}\\ &\Rightarrow y^{-1}<x^{-1}\end{aligned}$$

⇐. 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.

⇒. If $0<x\leqslant y$ then by (15) we have that $0<x^{-1}$ and $0<y^{-1}$. Hence

$$\begin{aligned} x\leqslant y & \underset{(7)}{\Rightarrow} x\odot x^{-1}\leqslant y\odot x^{-1} \\ & \Rightarrow 1\leqslant y\odot x^{-1} \\ & \Rightarrow 1\leqslant x^{-1}\odot y \\ & \underset{(7)}{\Rightarrow} 1\odot y^{-1}\leqslant (x^{-1}\odot y)\odot y^{-1} \\ & \Rightarrow y^{-1}\leqslant x^{-1} \end{aligned}$$

⇐. 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 $1\leqslant x^{-1}$.

□

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*** $\mathbb{N}_0$ ***is defined*** *by*

$$\mathbb{N}_0 = \bigcap \{S|S \textit{ is a successor set}\}$$

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*

$$s: \mathbb{N}_0 \to \mathbb{N}_0 \textit{ where } s(n) = n \bigcup \{n\}$$

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. $\mathbb{N} = \mathbb{N}_0 \backslash \{0\}$

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*

$$s: \mathbb{N}_0 \rightarrow \mathbb{N}_0 \textit{ is injective}$$

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 $\bigcap \mathcal{G} \in \mathcal{G}$. □

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) ≠ 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

$$\lambda: \mathbb{N}_0 \rightarrow A \text{ is a function} \tag{5.8}$$

As $\lambda \in \mathcal{G}$ we have that $(0, a) \in \lambda \Rightarrow a = \lambda(0)$, further if $n \in \mathbb{N}_0 = \mathrm{dom}(\lambda)$ 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

$$\beta: \mathbb{N}_0 \rightarrow A \text{ such that } \beta(0) = a \text{ and } \forall n \in \mathbb{N}_0 \text{ we have } \lambda(s(n)) = f(\lambda(n))$$

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 $f: A \rightarrow A$ 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* $\lambda(s(n)) = f(\lambda(n))$
3. *If* $a \notin f(A)$ *and* $f: A \rightarrow A$ *is injective then* λ *is injective*

Proof. The first part is easy. Using recursion [theorem: 5.19] there exists a function

$$\lambda: \mathbb{N}_0 \rightarrow A$$

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 $f: A \rightarrow A$ a function, $a \in A$ and $\lambda: \mathbb{N}_0 \rightarrow A$ 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 $f: \{1,2,3\} \rightarrow f(\{1,2,3\})$ by $f(i) = \left\{ \begin{array}{l} 2 \text{ if } i = 1 \\ 3 \text{ if } i = 2 \\ 2 \text{ if } i = 1 \end{array} \right.$ (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 $\lambda: \mathbb{N}_0 - A$ is clearly not injective.

Example 5.23. Take $f:\{1,2,3\}\rightarrow\{1,2,3\}$ 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 $\lambda:\mathbb{N}_0\rightarrow\{1,2,3\}$ 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 $f:A\rightarrow A$ 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. $(f)^{s(n)}=f\circ(f)^n$

Proof. Let $a\in A$ and use the recursion [theorem: 5.19] to find a **unique** function

$$\lambda_a:\mathbb{N}_0\rightarrow A\text{ such that }\lambda_a(0)=a\text{ and }\forall n\in\mathbb{N}_0\ \lambda_a(s(n))=f(\lambda_a(n))$$

Define now

$$(f)^n:A\rightarrow A\text{ where }(f)^n(a)=\lambda_a(n)$$

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

$$\begin{array}{rcl}(f)^0 & = & \mathrm{Id}_A\\ (f)^1=(f)^{s(0)} & = & f\circ(f)^0=f\circ\mathrm{Id}_A=f\\ (f)^2=(f)^{s(1)} & = & f\circ(f)^1=f\circ f\\ (f)^3=(f)^{s(2)} & = & f\circ(f)^2=f\circ f\circ f\\ & \ldots & \\ (f)^n & = & \overbrace{f\circ\ldots\circ f}^{n\,\text{times}}\end{array}$$

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

$$\pi_1: \mathbb{N}_0\times A\rightarrow\mathbb{N}_0 \text{ where } \pi_1(n,x)=n$$

$$\pi_2: \mathbb{N}_0\times A\rightarrow A \text{ where } \pi_2(n,x)=x$$

Define now

$$\gamma: \mathbb{N}_0\times A\rightarrow\mathbb{N}_0\times A \text{ where } \gamma(x)=(s(\pi_1(x)), g(\pi_1(x),\pi_2(x))) \tag{5.10}$$

Using the iteration [theorem: 5.24] on the above functions gives $\forall n\in\mathbb{N}_0$ the existence of the function

$$(\gamma)^n: \mathbb{N}_0\times A\rightarrow\mathbb{N}_0\times A \text{ such that } (\gamma)^0=\mathrm{Id}_{\mathbb{N}_0\times A} \text{ and } \forall n\in\mathbb{N}_0 \text{ we have } (\gamma)^{s(n)}=\gamma\circ(\gamma)^n \tag{5.11}$$

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

$$\begin{array}{rcl} \pi_1((\gamma)^{s(n)}(0,a)) & \underset{\text{[eq: 5.11]}}{=} & \pi_1((\gamma\circ(\gamma)^n)(0,a)) \\ & = & \pi_1(\gamma((\gamma)^n(0,n))) \\ & \underset{\text{[eq: 5.10]}}{=} & \pi_1(\pi_1((\gamma)^n(0,n)), g(\pi_1((\gamma)^n(0,a)),\pi_2((\gamma)^n(0,a)))) \\ & \underset{n\in S\Rightarrow\pi_1((\gamma)^n(0,a))=n}{=} & \pi_1(n, g(n,\pi_2((\gamma)^n(0,a)))) \\ & = & n \end{array}$$

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

$$\lambda: \mathbb{N}_0 \rightarrow A \text{ by } \gamma(n) = \pi_2((\gamma)^n(0,a)) \tag{5.13}$$

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

$$\begin{array}{rcl}
\lambda(s(n)) & = & \pi_2((\gamma)^{s(n)}(0,a)) \\
 & \underset{[\text{eq: } 5.11]}{=} & \pi_2((\gamma \circ (\gamma)^n)(0,a)) \\
 & = & \pi_2(\gamma((\gamma)^n(0,a))) \\
 & \underset{[\text{eq: } 5.10}{=} & \pi_2(\pi_1((\gamma)^n(0,a)), g(\pi_1((\gamma)^n(0,a)), \pi_2((\gamma)^n(0,a)))) \\
 & = & g(\pi_1((\gamma)^n(0,a)), \pi_2((\gamma)^n(0,a))) \\
 & \underset{[\text{eq: } 5.12}{=} & g(n, \pi_2((\gamma)^n(0,a))) \\
 & \underset{[\text{eq: } 5.13}{=} & g(n, \lambda(n))
\end{array}$$

This proves the existence of the function we are searching for. Now for uniqueness assume that there is a

$$\beta: \mathbb{N}_0 \rightarrow A \text{ such that } \beta(0) = a \text{ and } \forall n \in \mathbb{N}_0 \text{ that } \beta(s(n)) = g(n, \beta(n))$$

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 $0\in S_n$

$\boldsymbol{k\in S_n\Rightarrow s(k)\in S_n}$**.** We have

$$\begin{array}{rcl} n+s(k) & \underset{\text{[lemma: 5.31}}{=} & s(n+k)\\ & \underset{k\in S_{n,m}}{=} & s(k+n)\\ & \underset{\text{[theorem: 5.30}}{=} & 1+(k+n)\\ & \underset{\text{[theorem: 5.32}}{=} & (1+k)+n\\ & \underset{\text{[theorem: 5.30}}{=} & s(k)+n \end{array}$$

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*

$$\alpha_n\colon\mathbb{N}_0\rightarrow\mathbb{N}_0 \text{ by } \alpha_n(m)=n+m$$

Then we define the multiplication operator as follows

$$\cdot\colon\mathbb{N}_0\times\mathbb{N}_0\rightarrow\mathbb{N}_0 \text{ by } n\cdot m=.(n,m)=(\alpha_n)^m(0)$$

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. □

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. $n\cdot s(m)=(\alpha_n)^{s(m)}(0)=(\alpha_n\circ(\alpha_n)^m)(0)=a_n((\alpha_n)^m(0))=\alpha_n(n\cdot m)=n+n\cdot m$. □

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:

$\boldsymbol{0\in S_{n,m}.}$ $(n+m)\cdot 0\underset{\text{[theorem: 5.36]}}{=}0\underset{\text{[theorem: 5.29]}}{=}0+0\underset{\text{[theorem: 5.36]}}{=}n\cdot 0+m\cdot 0$

$\boldsymbol{n\in S_{n,m}\Rightarrow s(n)\in S_{n,m}.}$ We have

$$\begin{array}{rcl} (n+m)\cdot s(k) & \underset{\text{[lemma: 5.38}}{=} & (n+m)\cdot k+(n+m) \\ & \underset{k\in S_{n,m}}{=} & (n\cdot k+m\cdot k)+(n+m) \\ & \underset{\text{[theorem: 5.32}}{=} & n\cdot k+(m\cdot k+(n+m)) \\ & \underset{\text{[theorem: 5.33]}}{=} & n\cdot k+(m\cdot k+(m+n)) \\ & \underset{\text{[theorem: 5.32}}{=} & n\cdot k+((m\cdot k+m)+n) \\ & \underset{\text{[theorem: 5.33]}}{=} & n\cdot k+(n+(m\cdot k+m)) \\ & \underset{\text{[theorem: 5.32}}{=} & (n\cdot k+n)+(m\cdot k+m) \\ & \underset{\text{[lemma: 5.38}}{=} & n\cdot s(k)+m\cdot s(k) \end{array}$$

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 $(n + m) \cdot k = n \cdot k + m \cdot k$. □

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 $n \cdot m = m \cdot n$. □

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 $S=\{k \in \mathbb{N}_0 | \forall n, m \in \mathbb{N}_0$ with $n+k=m+k$ 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: 3.81] 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: 3.81]. □

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

$$\begin{aligned}
n<m &\underset{k\in S}{\Leftrightarrow} n+k<m+k\\
&\underset{\text{[theorem: 5.50]}}{\Leftrightarrow} s(n+k)<s(m+k)\\
&\underset{\text{[theorem: 5.31}}{\Leftrightarrow} n+s(k)<m+s(k)
\end{aligned}$$

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 $k \in \mathbb{N}_0 \backslash \{0\}$ 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 $\exists! k \in \mathbb{N}_0 \backslash \{0\}$ 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:

$n=m$. Then $m\underset{\text{[theorem: 5.29]}}{=}n+0$ where $0\in\mathbb{N}_0$.

$n<m$. 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:

$k=0$. Then $m=n+0\underset{\text{[theorem: 5.29]}}{=}n$ so that $n\leqslant m$.

$0<k$. 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.

- **⇒.** This follows from the previous theorem [theorem: 5.67]
- **⇐.** 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.

- **⇒.** 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 $n - k \leqslant m - k$.
- **⇐.** 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:

$\sup(A) = 0$. 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 $\sup(A) = x \in A$.

$\sup(A) \neq 0$. 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: 3.68] 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* $n, m \in \mathbb{N}_0 \backslash \{0\}$ *then* $n \cdot m \in \mathbb{N}_0 \backslash \{0\}$.

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* $k \in \mathbb{N}_0 \backslash \{0\}$ *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 $l \in \mathbb{N}_0 \backslash \{0\}$ 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 $l, k \in \mathbb{N}_0 \backslash \{0\}$ 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:

   $k=0$. Then by [theorem: 5.36] we have $n\cdot k=0=m\cdot k$ so that $n\cdot k\leqslant m\cdot l$.

   $k\neq 0$. 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 $z\in\mathbb{N}_0\setminus\{0\}$ 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 $k\in\mathbb{N}_0\setminus\{0\}$ 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 $n\in\mathbb{N}_0\setminus\{0\}$ 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 $n\in\mathbb{N}_0\setminus\{0\}$ $n\neq 0$ 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

$$q' = \min \left(A_{n,m} \right)$$

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:

$\boldsymbol{0\in S}$***.*** As $0+n\underset{\text{[theorem: 5.29]}}{=}n\in X$ we have $0\in S$.

$\boldsymbol{i\in S\Rightarrow s(i)\in S}$***.*** 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*

$$\begin{aligned} \lambda(0) &= a \\ \forall i \in \{0, \ldots, n-1\} \textit{ we have } \lambda(i+1) &= g(i, \lambda(i)) \end{aligned}$$

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

$$\beta(0) = a \tag{5.18}$$

$$\forall i \in \mathbb{N}_0 \text{ we have } \beta(i+1) = g'(i, \beta(i)) \tag{5.19}$$

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

$$\begin{aligned} \lambda(i+1) &= \beta_{|\{0, \ldots, n\}}(i+1) \\ &\underset{i+1 \in \{0, \ldots, n\}}{=} \beta(i+1) \\ &\underset{\text{[theorem: 5.19]}}{=} g'(i, \beta(i)) \\ &\underset{i \in \{0, \ldots, n-1\}}{=} g'(i, \lambda(i)) \\ &\underset{i \in \{0, \ldots, n-1\}}{=} g(i, \lambda(i)) \end{aligned}$$

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 $\gamma\colon \{0,\ldots,n\} \rightarrow A$ 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:

$\boldsymbol{0 \in S}$. 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:

> **$\boldsymbol{i+1 \in \{0,\ldots,n\}}$.** 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 $i+1 \in S$
>
> **$\boldsymbol{i+1 \notin \{0,\ldots,n\}}$.** 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 $\lambda = \gamma$. □

In the above the function $\lambda\colon \mathbb{N}_0 \rightarrow A$ is specified by saying what $a \in A$ is and what the function $g\colon N_0 \times A \rightarrow A$ 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:*

$$f\colon \mathbb{N}_0 \rightarrow A$$

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 $g\colon \mathbb{N}_0 \times A \rightarrow A$ 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 $f\colon \mathbb{N}_0 \rightarrow A$ as follows*

$$f(n) = \left\{ \begin{array}{l} a \text{ if } n = 0 \\ G(n-1, f(n-1)) \text{ if } n \in \{1,\ldots\infty\} \end{array} \right.$$

Which is equivalent with the function defined by [theorem: 5.87] where $a \in A$ and $g\colon \mathbb{N}_0 \times A \rightarrow A$ 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*

$$\lambda\colon \{0,\ldots,n\} \rightarrow A$$

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: 5.88] where $a \in A$ and $g\colon \{0,\ldots,n-1\} \times A \rightarrow A$ is defined by $g(n,x) = G(n,x)$.

Example 5.92. (Faculty) $\text{fac}\colon \mathbb{N}_0 \rightarrow \mathbb{N}_0$ is defined by

$$\text{fac}(n) = \left\{ \begin{array}{l} 1 \text{ if n=0} \\ n \cdot \text{fac}(n-1) = ((n-1)+1) \cdot \text{fac}(n-1) \end{array} \right.$$

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

$$\begin{array}{rcl}
\mathrm{fac}(0) & = & 1\\
\mathrm{fac}(1) & = & g(0,\mathrm{fac}(0))=(0+1)\cdot\mathrm{fac}(0)=1\cdot\mathrm{fac}(0)=1\cdot 1=1\\
\mathrm{fac}(2) & = & g(1,\mathrm{fac}(1))=(1+1)\cdot\mathrm{fac}(1)=2\cdot\mathrm{fac}(1)=2\cdot 1=2\\
\mathrm{fac}(3) & = & g(2,\mathrm{fac}(2))=(2+1)\cdot\mathrm{fac}(2)=3\cdot\mathrm{fac}(2)=3\cdot 2=6\\
 & \ldots & \\
\mathrm{fac}(n) & = & g(n-1,\mathrm{fac}(n-1))=((n-1)+1)\cdot\mathrm{fac}(n-1)=n\cdot\mathrm{fac}(n-1)\\
 & \ldots &
\end{array}$$

or in other words without using g

$$\begin{array}{rcl}
\mathrm{fac}(0) & = & 1\\
\mathrm{fac}(1) & = & 1\cdot\mathrm{fac}(0)=1\cdot 1=1\\
\mathrm{fac}(2) & = & 2\cdot\mathrm{fac}(1)=2\cdot 1=2\\
\mathrm{fac}(3) & = & 3\cdot\mathrm{fac}(2)=3\cdot 2=6\\
 & \ldots & \\
\mathrm{fac}(n) & = & n\cdot\mathrm{fac}(n-1)\\
 & \ldots &
\end{array}$$

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 $f: A \to B$. 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. $\mathrm{Id}: A \to A$ is a bijection [see example: 2.64] proving that $A \approx A$
2. As $A \approx B$ there exist a bijection $f: A \to B$ but then by [theorem: 2.71] $f^{-1}|B \to A|$ is also a bijection, so that $B \approx A$.
3. If $A \approx B$ and $B \approx C$ then there exists bijections $f: A \to B$ and $g: B \to C$, using [theorem: 2.73] we have that $g \circ f: A \to C$ 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: 2.76] 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 $f: A \to B$*

Proof.

⇒. If $A \preccurlyeq B$ then there exist a set $C \subseteq B$ and a bijection $f: A \to C$, as a bijection is injective we have that $f: A \to C$ is injective and finally by [theorem: 2.52] $f: A \to B$ is a injection.

⇐. If $f: A \to C$ is a injection then by [theorem: 2.66] $f: A \to f(A)$ 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

$$f: A \rightarrow \mathcal{P}(A)$$

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:

$\boldsymbol{B = A}$. 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 $\mathcal{P}(A)$.

$\boldsymbol{B \subset A}$. 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 $g: B \rightarrow \mathcal{P}(A)$ exist, take the constant function $C_{\varnothing}: A \setminus B \rightarrow \mathcal{P}(A)$ where $C_{\varnothing}(x) = \varnothing$ and form then using [theorem: 2.78] the function

$$f = g \bigcup C_{\varnothing}: A \rightarrow \mathcal{P}(A)$$

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 $f: A \rightarrow B$ if and only there exist a surjection $g: B \rightarrow A$*

Proof.

⇒. Let $f: A \rightarrow B$ be a injection then by [theorem: 2.60] there exist a $g: B \rightarrow A$ 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 $g: B \rightarrow A$ be a surjection then by [theorem: 3.99] there exist a injective function $f: A \rightarrow B$. □

Corollary 6.10. *If A, B are sets then $A \preccurlyeq B$ if and only if there exist a surjection $f: B \rightarrow A$*

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 $f: A \rightarrow B$ and $g: C \rightarrow D$ then by [theorem: 2.80] there exists a bijection $f \bigcup g: A \bigcup C \rightarrow B \bigcup D$. 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 $f: A \rightarrow B$ and $g: C \rightarrow D$. Define

$$h: A \times C \rightarrow B \times D \text{ by } h(x, y) = (f(x), g(x))$$

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 $f\colon A\rightarrow B$ and $g\colon D\rightarrow C$. If $x\in A^C$ then $x\colon C\rightarrow A$ is a function, so $x\circ g\colon D\rightarrow A$ is a function, hence $f\circ(x\circ g)\colon D\rightarrow B$ is a function, proving that $f\circ(x\circ g)\in B^D$. Define now $h\colon A^C\rightarrow B^D$ by $h(x)=f\circ(x\circ g)$ then we have:

injectivity. If $x,y\in A^C$ satisfies $h(x)=h(y)$ then

$$\begin{array}{lll} f\circ(x\circ g)=f\circ(y\circ g) & \Rightarrow & f^{-1}\circ(f\circ(x\circ g))=f^{-1}\circ(f\circ(y\circ g)) \\ & \Rightarrow & (f^{-1}\circ f)\circ(x\circ g)=(f^{-1}\circ f)\circ(y\circ g) \\ & \Rightarrow & \mathrm{Id}_A\circ(x\circ g)=\mathrm{Id}_A\circ(y\circ g) \\ & \Rightarrow & x\circ g=y\circ g \\ & \Rightarrow & (x\circ g)\circ g^{-1}=(y\circ g)\circ g^{-1} \\ & \Rightarrow & x\circ(g\circ g^{-1})=y\circ(g\circ g^{-1}) \\ & \Rightarrow & x\circ\mathrm{Id}_C=y\circ\mathrm{Id}_C \\ & \Rightarrow x=y & \end{array}$$

surjectivity. If $y\in B^D$ then $y\colon D\rightarrow B$ is a function so that $y\circ g^{-1}\colon C\rightarrow B$ is a function, hence $f^{-1}\circ(y\circ g^{-1})\colon C\rightarrow A$ 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: 3.45] 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:

$0 \in S$. If $x \in S_0$ then $x < 0$ which by [theorem: 5.49] is false, so $S_0 = \varnothing = 0$ proving that $0 \in S$.

$n \in S \Rightarrow s(n) \in S$. 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:

$m = n$. Then by [theorem: 5.47] we have that $m < s(m) = s(n)$ so that $m \in S_{s(n)}$

$m \in S_n$. 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.

$\Rightarrow$. If $x \in S_n$ then $x < n$ which as $n \leqslant m$ proves that $x < m$ so that $x \in S_m$, hence $S_n \subseteq S_m$.

$\Leftarrow$. 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

$$\begin{array}{lcl} A \text{ is finite} & \Leftrightarrow & \exists n \in \mathbb{N}_0 \text{ such that } A \approx n \\ & \Leftrightarrow_{\sim\text{[theorem: 6.21]}} & \exists n \in \mathbb{N}_0 \text{ such that } A \approx \{1, \ldots, n\} \end{array}$$

□

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:

- **$\Rightarrow$.** 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

$$\{1, \ldots, n\} \approx A$$

- **$\Leftarrow$.** 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 Ais* ***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:

$\boldsymbol{i < n \wedge i' < n}$**.** Then $f(i) = g(i) = g(i') = f(i')$ which as f is injective proves that $i = i'$.

$\boldsymbol{i < n \wedge n \leqslant i'}$**.** 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.

$\boldsymbol{n \leqslant i \wedge i' < n}$**.** 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.

$\boldsymbol{n \leqslant i \wedge n \leqslant i'}$**.** 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

$\boldsymbol{i < n}$**.** Then $g(i) = f(i) = y$

$\boldsymbol{n < i}$**.** 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:

$\boldsymbol{0 \in S}$**.** 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: $\mathbb{N}_0 \rightarrow A$ 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: 3.123] there exists a order relation $\leqslant_A$ such that $\langle A,\leqslant_A\rangle$ is a well ordered set. Using [theorem: 3.93] 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 $f:A\rightarrow n$. As $B\subseteq A$ we have that $f_{|B}$: $B\rightarrow f(B)$ is a bijection [see theorems: 2.83, 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* ∈ *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* ∈ *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: 2.80] 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 ∈ *S*. If $n=0$ then $S_0=0=\varnothing$ so that $\bigcup_{i\in S_0}A_i=\bigcup_{i\in\varnothing}A_i\underset{\text{[example: 2.115}}{=}\varnothing$ which is finite, hence $0\in S$.

***n* ∈ *S* ⇒ *n* + 1 ∈ *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

$$\bigcup_{i\in S_{n+1}}A_i\underset{\text{[theorem: 2.128]}}{=}\left(\bigcup_{i\in S_{n+1}\setminus\{n\}}A_i\right)\bigcup A_n=\left(\bigcup_{i\in S_n}A_i\right)\bigcup A_n$$

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: 2.113] 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.

⇒. 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:

$$\begin{array}{rcl} g(x) & = & f(f^{-1}(x)+1)\\ & = & f(f^{-1}(f(f^{-1}(y)-1))+1)\\ & = & f((f^{-1}(y)-1)+1)\\ & = & f(f^{-1}(y))\\ & = & y \end{array}$$

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

$$\begin{array}{rcl} f(f^{-1}(x)+1) = f(f^{-1}(x')+1) & \underset{f \text{ is injective}}{\Rightarrow} & f^{-1}(x)+1 = f^{-1}(x')+1 \\ & \Rightarrow & f^{-1}(x) = f^{-1}(x') \\ & \underset{f^{-1} \text{ is injective}}{\Rightarrow} & x = x' \end{array}$$

$\boldsymbol{x \in B \wedge x' \notin B.}$ 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.

- **$\Rightarrow$.** If $A = \varnothing$ then by [example: 2.63] $\varnothing\colon \varnothing \rightarrow \oslash$ is a bijection, so as $0 = \varnothing$ we have $\#\varnothing = 0$.
- **$\Leftarrow$.** If $\#A = 0$ then as $0 = \varnothing$ there exists a bijection $f\colon \varnothing \rightarrow A$, 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 $f\colon B \rightarrow n = S_n$ and $g\colon A \rightarrow m = S_m$. 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

  $$h\colon A \times B \rightarrow S_{n \cdot m} \text{ where } h(x, y) = n \cdot g(x) + f(x)$$

  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 $f\colon A \rightarrow S_n$ and $g\colon B \rightarrow S_m$. 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

$$h\colon A \bigcup B \rightarrow S_{n+m} \text{ where } h(x) = \begin{cases} f(x) \text{ if } x \in A \\ n + g(x) \text{ if } x \in B \end{cases}$$

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 $h(x)=f(x)=y$

$\boldsymbol{n\leqslant y}$**.** 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:

$\boldsymbol{B=A}$**.** 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.

$\boldsymbol{B=\varnothing}$**.** 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 $\#B=0<\#A$.

$\boldsymbol{\varnothing\neq B\subset A}$**.** 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 $\#B=\#g(B)\leqslant\#(A)$. □

Theorem 6.44. *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 6.45. *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. □

Theorem 6.46. *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: 2.83] $f_{|n}$ 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 6.47. *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 6.48. *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: 6.47] 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 6.49. *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 $\forall i \in S_n$ we have $x_{\beta(i)} \leqslant x_{\beta(n)}$.*

Proof. We prove this by induction, so let

$$S = \{n \in \mathbb{N}_0 | \forall \{x_i\}_{i \in S_{n+1}} \subseteq X \text{ there exist a bijection } \beta: S_{n+1} \rightarrow S_{n+1} \text{ such that } \forall i \in S_n\ x_{\beta(i)} \leqslant x_{\beta(n)}\}$$

then we have:

$0 \in S$. If $\{x_i\}_{i \in S_1 = \{0\}} \subseteq X$ then for the bijection $\beta = \mathrm{Id}_{S_1}: S_1 \rightarrow S_1$ 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 $\alpha: S_{n+1} \rightarrow S_{n+1}$ 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

$$\beta: S_{(n+1)+1} \rightarrow S_{(n+1)+1} \text{ by } \beta(i) = \begin{cases} \alpha(i) \text{ if } i \in S_{n+1} \\ n+1 \text{ if } i = n+1 \end{cases}$$

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 $\beta: S_{(n+1)+1} \rightarrow S_{(n+1)+1}$ is a bijection. Let now $i \in S_{n+1}$ then we have the following possibilities:

$\boldsymbol{i = n}$. Then $x_{\beta(i)} = x_{\alpha(i)} = x_{\alpha(n)} \leqslant x_{n+1} = x_{\beta(n+1)}$.

$\boldsymbol{i \in S_n}$. 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 $n+1 \in S$.

$\boldsymbol{x_{n+1} < x_{\alpha(n)}}$. Define

$$\beta: S_{(n+1)+1} \rightarrow S_{(n+1)+1} \text{ by } \beta(i) = \begin{cases} \alpha(i) \text{ if } i \in S_n \\ n+1 \text{ if } i = n \\ \alpha(n) \text{ if } i = n+1 \end{cases}$$

then we have:

injectivity. Let $i, j \in S_{(n+1)+1}$ such that $\beta(i) = \beta(j)$ then we have the following possibilities:

$\boldsymbol{i \in S_n \wedge j \in S_n}$. Then $\alpha(i) = \beta(i) = \beta(j) = \alpha(j)$ which as β is a bijection gives $i = j$.

$\boldsymbol{i \in S_n \wedge j = n}$. 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.

$\boldsymbol{i \in S_n \wedge j = n+1}$. 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.

$\boldsymbol{i = n \wedge j \in S_n}$. 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.

$\boldsymbol{i = n \wedge j = n}$. Then $i = j$.

$\boldsymbol{i = n \wedge j = n+1}$. 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.

$\boldsymbol{i = n+1 \wedge j \in S_n}$. 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.

$\boldsymbol{i = n+1 \wedge j = n}$. 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.

$\boldsymbol{i = n+1 \wedge j = n+1}$. Then $i = j$.

surjectivity. Let $j \in S_{(n+1)+1}$ then we have the following possibilities to check:

$\boldsymbol{j = n+1}$. then $\beta(n) = j$

$\boldsymbol{j \in S_{n+1}}$. 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:

$\boldsymbol{i = n}$. Then $x_{\beta(i)} = x_{n+1} \leqslant x_{\alpha(n)} = x_{\beta(n+1)}$.

$\boldsymbol{i \in S_n}$. 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 6.50. *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: 6.49] 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 6.51. *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 $\beta: S_{n+1} \rightarrow S_{n+1}$ such that*

$$\forall i \in S_n \textit{ we have } x_{\beta(i)} \leqslant x_{\beta(i+1)}$$

Proof. We proof this by induction, so let

$S = \{n \in \mathbb{N}_0 | \forall \{x_i\}_{i \in n+1} \subseteq X$ there exist a bijection $\beta: S_{n+1} \rightarrow S_{n+1}$ 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: 6.49] 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.

$\boldsymbol{k=n+1\wedge l\in S_{n+1}}$**.** 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.

$\boldsymbol{k=n+1\wedge l=n+1}$**.** 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:

$\boldsymbol{l=n+1}$**.** Then $\beta(n+1)=\alpha(n+1)=k$

$\boldsymbol{l\in S_{n+1}}$**.** 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:

$\boldsymbol{i=n}$**.** 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

$$x_{\beta(i)}=x_{\alpha(\gamma(i))}\leqslant x_{\beta(n+1)}=x_{\beta(i+1)}$$

$\boldsymbol{i\in S_n}$**.** 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 6.52. *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 $\alpha\colon\{n,\ldots,m\}\rightarrow\{n,\ldots,m\}$ 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

$$\beta\colon\{n,\ldots,m\}\rightarrow S_{(m-n)+1}\text{ where }\beta(i)=i-n \tag{6.19}$$

and

$$\beta^{-1}\colon S_{(m-n)+1}\rightarrow\{n,\ldots,m\}\text{ where }\beta(i)=i+n \tag{6.20}$$

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: 6.51] a bijection

$$\gamma\colon S_{(m-n)+1}\rightarrow S_{(m-n)+1}\text{ such that }\forall i\in S_{m-n}\text{ we have }x_{\beta^{-1}(\gamma(i))}\leqslant x_{\beta^{-1}(\gamma(i+1))} \tag{6.21}$$

Define now the bijection

$$\alpha=\beta^{-1}\circ\gamma\circ\beta\colon\{n,\ldots,m\}\rightarrow\{n,\ldots,m\}$$

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 $\alpha\colon\{n,\ldots,m\}\rightarrow\{n,\ldots,m\}$ 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 6.53. *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:

$$x{:}\,S_{n+1}\to\bigcup_{i\in S_{n+1}}A_i \text{ by } x(i)=\begin{cases} y(i) \text{ if } i\in S_n\\ a \text{ if } i=n\end{cases}$$

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 6.54. *Let $n\in\mathbb{N}_0\backslash\{0\}$ 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

$$S=\left\{n\in\{1,\ldots,\infty\}|\text{If }\{A_i\}_{i\in S_n}\text{satisfies }\forall i\in S_n\ \ A_i\text{ is finite then }\prod_{i\in S_n}A_i\text{ is finite}\right\}$$

then we have:

$1\in S$. Using [example: 2.132] 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: 6.53] 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 6.55. *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 $n\in\mathbb{N}_0\backslash\{0\}$ such that $k\approx I$, so there exist a bijection $f{:}\,S_k\to I$. Using [theorem: 2.137] we have that there exists a bijection $\beta{:}\,\prod_{i\in I}A_i\to\prod_{i\in S_k}A_{f(i)}$ hence $\prod_{i\in I}A_i\approx\prod_{i\in S_k}A_{f(i)}$. By [theorem: 6.54] 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 6.56. *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: 3.94] 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 6.57. *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: 2.83] 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: 6.56] 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 6.58. $\mathbb{N}_0 \times \mathbb{N}_0 \approx \mathbb{N}_0$*, 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 6.59. *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 $\lambda\colon\mathbb{N}_0\rightarrow\mathbb{N}_0\times\mathbb{N}_0$ 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: 6.59] 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 $\lambda\colon\mathbb{N}_0\rightarrow\mathbb{N}_0\times\mathbb{N}_0$ is a bijection, proving that $\mathbb{N}_0\times\mathbb{N}_0$ is denumerable. □

Corollary 6.60. *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 6.61. *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: 6.57]

$$n \times \mathbb{N}_0 \text{ is either finite or denumerable}$$

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 6.62. *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: 6.61] 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 6.63. *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: 6.61] 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: 6.57] 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 6.64. *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 $\beta: \mathbb{N}_0 \rightarrow I$. 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

$$g: \mathbb{N}_0 \times \mathbb{N}_0 \rightarrow \bigcup_{i \in I} A_i \text{ by } g(n, m) = \alpha_{\beta(n)}(m)$$

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 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

$$g: \mathbb{N}_0 \times \mathbb{N}_0 \rightarrow \bigcup_{i \in I} A_i \text{ is surjective}$$

Now by [theorem: 6.58] 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: 6.10] 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: 6.57] 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 6.65. *Every subset of a denumerable set is countable*

Proof. This follows from [theorem: 6.57] and the definition of countable sets. □

Theorem 6.66. *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: 6.65] every subset of A is countable. □

Theorem 6.67. *Let A be a non empty set then the following are equivalent:*

1. *A is countable*

2. *There exists a surjection* $\beta: \mathbb{N}_0 \rightarrow A$
3. *There exists a injection* $\alpha: A \rightarrow \mathbb{N}_0$
4. *There exist a denumerable set* B *and a injection* $\alpha: A \rightarrow B$

Proof.

1 ⇒ 2. If A is countable then we have either:

***A* is finite.** Then $\exists n \in \mathbb{N}_0$ and a bijection $\alpha: n = S_n \rightarrow A$. As $A \neq \varnothing$ there exist a $a \in A$, this allows us to define the function

$$\beta: \mathbb{N}_0 \rightarrow A \text{ where } \beta(i) = \begin{cases} \alpha(i) \text{ if } i < n \\ a \text{ if } n \leqslant i \end{cases}$$

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 $\beta: \mathbb{N}_0 \rightarrow A$ is surjective.

***A* is denumerable.** Then $\mathbb{N}_0 \approx A$ so there exist a bijection, hence surjection, $\beta: \mathbb{N}_0 \rightarrow A$.

2 ⇒ 3. Given that there exists a surjection $\beta: \mathbb{N}_0 \rightarrow A$ and $A \neq \varnothing$ we have by [theorem: 6.9] the existence of a injection $\alpha: A \rightarrow \mathbb{N}_0$.

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: 6.56] 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 6.68. *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: 6.67]. 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: 6.58] 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: 6.67] it follows that $\bigcup_{i \in I} A_i$ is countable. □

Theorem 6.69. *If* $\{A_i\}_{i \in I}$ *is such that* $I \neq \varnothing \wedge I$ *is finite and* $\forall i \in I$ A_i *is countable then* $\bigcup_{i \in I} A_i$ *is countable. In other words the union of a finite family of countable sets is countable. If in addition* $\forall i \in I$ $A_i \neq \varnothing$ *and* $\forall i, j \in \mathbb{N}_0$ *with* $i \neq j$ $A_i \bigcap A_j = \varnothing$ *then* $\bigcup_{i \in I} A_i$ *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 countable there exist a surjection $\alpha_i: \mathbb{N}_0 \rightarrow A_i$ [see theorem: 6.67].] 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 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

$$g: n_0 \times \mathbb{N}_0 \rightarrow \bigcup_{i \in I} A_i \text{ is surjective}$$

Now by [theorem: 6.61] 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: 6.67] it follows that

$$\bigcup_{i \in I} A_i \text{ is countable}$$

Further if $\forall i \in I$ $A_i \neq \varnothing$ and $\forall i, j \in \mathbb{N}_0$ with $i \neq j$ $A_i \bigcap A_j = \varnothing$ then we can use a consequence of the axiom of choice [see theorem: 3.123] to find a function

$$\mathcal{C}: I \rightarrow \bigcup_{i \in I} A_i \text{ such that } \forall i \in I\ \mathcal{C}(i) \in A_i$$

If $\mathcal{C}(i) = \mathcal{C}(j)$ then $\mathcal{C}(i) \in A_i$ and $\mathcal{C}(i) = C(j) \in A_j$ so that $\mathcal{C}(i) \in A_i \bigcap A_j \Rightarrow A_i \bigcap A_j \neq \varnothing$. hence we must have $i = j$ [if $i \neq j$ then $A_i \bigcap A_j = \varnothing$]. So $\mathcal{C}: I \rightarrow \bigcup_{i \in I} A_i$ is a injection and $\mathcal{C}: I \rightarrow \mathcal{C}(I)$ is a bijection or $I \approx \mathcal{C}(I)$, as I is countable it follows from [theorem: 6.26] that $\mathcal{C}(I)$ is denumerable. As $\mathcal{C}(I) \subseteq \bigcup_{i \in I} A_i$ we have by [theorem: 6.29] that $\bigcup_{i \in I} A_i$ is not finite, so as $\bigcup_{i \in I} A_i$ is countable we have $\bigcup_{i \in I} A_i$ is denumerable. □

Theorem 6.70. *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: 6.62] $A \times B$ is denumerable hence countable.

***A* is denumerable and *B* is finite.** Then by [theorem: 6.62] $A \times B$ is denumerable hence countable.

***A* is denumerable and *B* is denumerable.** Then by [theorem: 6.60] $A \times B$ is denumerable hence countable. □

Lemma 6.71. *Let $n \in \mathbb{N}_0 \backslash \{0\}$ 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: 2.132] 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].

$\boldsymbol{n \in S \Rightarrow n+1 \in S}$. 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: 6.70] we have that $(\prod_{i \in S_n} A_i) \times A_n$ is countable. Finally by [theorem: 6.53] 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 6.72. *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: 2.137] there exists a bijection β: $\prod_{i \in I} A_i \rightarrow \prod_{i \in S_n} A_{f(i)}$ so that $\prod_{i \in S_n} A_{f(i)} \approx \prod_{i \in I} A_i$. Using the previous lemma [lemma: 6.71] $\prod_{i \in S_n} A_{f(i)}$ 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: 2.131] we can define the finite product of sets.

Definition 6.73. *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*

$$\prod_{i=1}^n A_i = \prod_{i \in \{1, \ldots, n\}} A_i \text{ [see definition: 2.131]}$$

In other words

$$\prod_{i=1}^n A_i \underset{\text{[defintion: 2.131]}}{=} \left\{ f | f \in \left(\bigcup_{i \in \{1, \ldots, n\}} A_i \right)^{\{1, \ldots, n\}} \text{ where } \forall i \in \{1, \ldots, n\} \text{ we have } f(i) \in A_i \right\}$$

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* 6.74.** *$x \in \prod_{i=1}^n A_i$ 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: 2.140]

Definition 6.75. *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 6.76. *If $n \in \mathbb{N}$ and $\{A_i\}_{i \in \{1, \ldots, n\}} \subseteq B$ a finite family of sets then $\forall i \in \{1, \ldots, n\}$ we have that π_i: $\prod_{i=1}^n A_i \rightarrow A_i$ is a surjection*

Proof. This was proved in [theorem: 3.101] □

Next we consider the special case where for $\forall i \in \{1, \ldots, n\}$ $A_i = A$ [see also theorem: 2.136]

Definition 6.77. *Let $n \in \mathbb{N}$, A a set then A^n is defined by*

$$A^n = \prod_{i=1}^{n} A_i \text{ where } \{A_i\}_{i \in \{1,\ldots,n\}} \subseteq \{A\} \text{ is defined by } C_A: \{1,\ldots,n\} \rightarrow \{A\}$$

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

$$A^n = \{f \mid f \in A^{\{1,\ldots,n\}} \text{ where } \forall i \in \{1,\ldots,n\} \text{ we have } f(i) \in A\} = A^{\{1,\ldots,n\}}$$

Using [notation: 6.74] 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 6.78. *Let $n \in \mathbb{N}$, A a set then $A^n = A^{\{1,\ldots,n\}} \underset{\text{def}}{=} \{f \mid f: \{1,\ldots,n\} \rightarrow A \text{ is a function}\}$*

Proof. This follows from [theorem: 2.136] □

Theorem 6.79. *If A is a set then*

$$\beta: A \rightarrow A^1 \text{ defined by } \beta(a) = \{(1,a)\}$$

is a bijection.

Proof. If $a \in A$ then $(1,a) \subseteq \{1\} \times A = \{1,\ldots,1\} \times A$ so that $\beta(a): \{1,\ldots,1\} \rightarrow A$ 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 $f: \{1\} \rightarrow A$ 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 6.80. *Let $\langle G, + \rangle$ be a semi-group withe neutral element 0, $n \in \mathbb{N}$ then if we define*

$$+^n: G^n \times G^n \rightarrow G^n \text{ by } x +^n y: \{1,\ldots,n\} \rightarrow G \text{ where } (x +^n y)(i) = x(i) + y(i)$$

$$0^n: \{1,\ldots,n\} \rightarrow G \text{ by } 0^n(i) = 0$$

or using [notation: 6.74]

$$\forall x, y \in G^n \text{ we have } x +^n y = (x_1 + y_1, \ldots, x_n + y_n) \text{ where } (x +^n y)_i = x_i + y_i$$

$$0^n \in G^n \text{ is } \left(\underbrace{0,\ldots,0}_{n}\right) \text{ so that } (0^n)_i = 0$$

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*

$$\sim x: \{1,\ldots,n\} \rightarrow G \text{ is defined by } (\sim x)(i) = -x(i)$$

or using [notation: 6.74]

$$\sim x = \sim(x_1, \ldots, x_n) = (-x_1, \ldots, -x_n) \text{ so that } (\sim x)_i = -x_i$$

Proof. This was actuall proved for a more general case in [theorem: 4.26]. □

Note 6.81. 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

$$\sim[(n+r,m+s)]=\sim[(n'+r',m'+s')] \qquad \square$$

The above theorem ensure that the following definition is well defined:

Definition 7.5. *The sum operator* $+:\mathbb{Z}\times\mathbb{Z}\rightarrow\mathbb{Z}$ *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 $\sim[(n,n)]=\sim[(0,0)]$. □

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

$$\begin{array}{rcl}(n+m)+k &=& (\sim[(n_1,m_1)]+\sim[(n_2,m_2)])+\sim[(n_3,m_3)]\\ &=& \sim[(n_1+n_2,m_1+m_2)]+\sim[(n_3,m_3)]\\ &=& [\sim((n_1+n_2)+n_3,(m_1+m_2)+m_3)]\\ &=& \sim[(n_1+(n_2+n_3),m_1+(m_2+m_3))]\\ &=& \sim[(n_1,m_1)]+\sim[(n_2+n_3,m_2+m_3)]\\ &=& \sim[(n_1,m_1)]+(\sim[(n_2,m_2)]+\sim[(n_3,m_3)])\\ &=& n+(m+k)\end{array}$$

Commutativity. If $n=\sim[(n_1,m_1)]$ and $m=\sim[(n_2,m_2)]$ then

$$\begin{array}{rcl}\sim[(n_1,m_1)]+\sim[(n_2,m_2)] &=& \sim[(n_1+n_2,m_1+m_2)]\\ &=& \sim[(n_2+n_1,m_2+m_1)]\\ &=& \sim[(n_2,m_2)]+\sim[(n_1,m_1)]\end{array}$$

Neutral element. If $k=\sim[(n,m)]\in\mathbb{Z}$ then

$$\begin{array}{rcl}0+k &\underset{\text{commutativity}}{=}& k+0\\ &=& \sim[(n,m)]+\sim[(0,0)]\\ &=& \sim[(n+0,m+0)]\\ &=& \sim[(n,m)]\\ &=& k\end{array}$$

Inverse element. If $k=\sim[(n,m)]$

$$\begin{array}{rcl}k+(-k) &\underset{\text{commutativity}}{=}& (-k)+k\\ &=& \sim[(m,n)]+\sim[(n,m)]\\ &=& \sim[(m+n,n+m)]\\ &=& \sim[(n+m,n+m)]\\ &\underset{\text{[theorem: 7.6]}}{=}& \sim[(0,0)]\end{array}$$

□

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

$$\begin{aligned}n\cdot r+m'\cdot r &= (n+m')\cdot r\\ &\underset{[\text{eq: }7.2]}{=} (m+n')\cdot r\\ &= m\cdot r+n'\cdot r\\ m\cdot s+n'\cdot s &= (m+n')\cdot s\\ &\underset{[\text{eq: }7.2]}{=} (n+m')\cdot s\\ &= n\cdot s+m'\cdot s\\ m'\cdot s+m'\cdot r' &= m'\cdot(s+r')\\ &\underset{[\text{eq: }7.2]}{=} m'\cdot(r+s')\\ &= m'\cdot r+m'\cdot s'\\ n'\cdot r+n'\cdot s' &= n'\cdot(r+s')\\ &\underset{[\text{eq: }7.2]}{=} n'\cdot(s+r')\\ &= n'\cdot s+n'\cdot r'\end{aligned}$$

so after summing (underlining common terms).

$$\begin{aligned}&n\cdot r+\underbrace{m'\cdot r}_{1}+m\cdot s+\underbrace{n'\cdot s}_{2}+\underbrace{m'\cdot s}_{3}+m'\cdot r'+\underbrace{n'\cdot r}_{4}+n'\cdot s' =\\ &m\cdot r+\underbrace{n'\cdot r}_{4}+n\cdot s+\underbrace{m'\cdot s}_{3}+\underbrace{m'\cdot r}_{1}+m'\cdot s'+\underbrace{n'\cdot s}_{2}+n'\cdot r'\end{aligned}$$

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

$$\begin{aligned} \sim[(n,m)] \cdot \sim[(r,s)] &= \sim[(n \cdot r + m \cdot s, m \cdot r + n \cdot s)] \\ &= \sim[(r \cdot n + s \cdot m, s \cdot n + r \cdot m)] \\ &= \sim[(r,s)] \cdot \sim[(n,m)] \end{aligned}$$

   Associativity. Let $\sim[(a,b)], \sim[(c,d)], \sim[(e,f)] \in \mathbb{Z}$ then

$$\begin{aligned} \sim[(a,b)] \cdot (\sim[(c,d)] \cdot \sim[(e,f)]) &= \\ \sim[(a,b)] \cdot (\sim[(c \cdot e + d \cdot f, d \cdot e + c \cdot f)]) &= \\ \sim[(a \cdot (c \cdot e + d \cdot f) + b \cdot (d \cdot e + c \cdot f), b \cdot (c \cdot e + d \cdot f) + a \cdot (d \cdot e + c \cdot f))] &= \\ \sim\left[\left(\underbrace{a \cdot (c \cdot e)}_{1} + \underbrace{a \cdot (d \cdot f)}_{2} + \underbrace{b \cdot (d \cdot e)}_{3} + \underbrace{b \cdot (c \cdot f)}_{4}, \underbrace{b \cdot (c \cdot e)}_{5} + \underbrace{b \cdot (d \cdot f)}_{6} + \underbrace{a \cdot (d \cdot e)}_{7} + \right.\right. & \\ \left.\left.\underbrace{a \cdot (c \cdot f)}_{8}\right)\right] &= \\ \sim\left[\left(\underbrace{(a \cdot c) \cdot e}_{1} + \underbrace{(b \cdot d) \cdot e}_{3} + \underbrace{(b \cdot c) \cdot f}_{4} + \underbrace{(a \cdot d) \cdot f}_{2}, \underbrace{(b \cdot c) \cdot e}_{5} + \underbrace{(a \cdot d) \cdot e}_{7} + \underbrace{(a \cdot c) \cdot f}_{8} + \right.\right. & \\ \left.\left.\underbrace{(b \cdot d) \cdot f}_{6}\right)\right] & \\ \sim[(a \cdot c + b \cdot d) \cdot e + (b \cdot c + a \cdot d) \cdot f, (b \cdot c + a \cdot d) \cdot e + (a \cdot c + b \cdot d) \cdot f] &= \\ \sim[(a \cdot c + b \cdot d, b \cdot c + a \cdot d)] \cdot \backsim[(e,f)] &= \\ (\sim[(a,b)] \cdot \sim[(c,d)]) \cdot \sim[(e,f)] & \end{aligned}$$

   Neutral element. If $n = \sim[(n,m)] \in \mathbb{Z}$ then we have

$$\begin{aligned} n \cdot 1 &\underset{\text{commutativity}}{=} 1 \cdot n \\ &= \sim[(1,0)] \cdot \sim[(n,m)] \\ &= \sim[(1 \cdot n + 0 \cdot m, 0 \cdot n + 1 \cdot m)] \\ &= \sim(n,m) \end{aligned}$$

3. Further we have:

Distributivity. If $\sim[(a,b)],\sim[(c,d)],\sim[(e,f)]\in\mathbb{Z}$ then

$$\begin{aligned}
\sim[(a,b)]\cdot(\sim[(c,d)]+\sim[(e,f)]) &=\\
\sim[(a,b)]\cdot\sim[(c+e,d+f)] &=\\
\sim[(a\cdot(c+e)+b\cdot(d+f),b\cdot(c+e)+a\cdot(d+f))] &=\\
\sim\left[\left(\underbrace{a\cdot c}_{1}+\underbrace{a\cdot e}_{2}+\underbrace{b\cdot d}_{3}+\underbrace{b\cdot f}_{4},\underbrace{b\cdot c}_{5}+\underbrace{b\cdot e}_{6}+\underbrace{a\cdot d}_{7}+\underbrace{a\cdot f}_{8}\right)\right] &=\\
\sim\left[\left(\underbrace{a\cdot c}_{1}+\underbrace{b\cdot d}_{3}+\underbrace{a\cdot e}_{2}+\underbrace{b\cdot f}_{4},\underbrace{b\cdot c}_{5}+\underbrace{a\cdot d}_{7}+\underbrace{b\cdot e}_{6}+\underbrace{a\cdot f}_{8}\right)\right] &=\\
\sim[(a\cdot c+b\cdot d,b\cdot c+a\cdot d)]+\sim[(a\cdot e+b\cdot f,b\cdot e+a\cdot f)] &=\\
\sim[(a,b)]\cdot\sim[(c,d)]+\sim[(a,b)]\cdot\sim[(e,f)] &=
\end{aligned}$$

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

$$\begin{aligned}
a\cdot c+(a+k)\cdot d=(a+k)\cdot c+a\cdot d &\quad \Rightarrow\\
\underbrace{a\cdot c}_{1}+\underbrace{a\cdot d}_{2}+k\cdot d=\underbrace{a\cdot c}_{1}+k\cdot c+\underbrace{a\cdot d}_{2} &\quad \Rightarrow\\
k\cdot d=k\cdot c &\quad \underset{k\neq 0\wedge[\text{theorem: }5.77]}{\Rightarrow}\\
d=c &
\end{aligned}$$

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

$$\begin{aligned}
(b+k)\cdot c+b\cdot d=b\cdot c+(b+k)\cdot d &\quad \Rightarrow\\
\underbrace{b\cdot c}_{1}+k\cdot c+\underbrace{b\cdot d}_{2}=\underbrace{b\cdot c}_{1}+\underbrace{b\cdot d}_{2}+k\cdot d &\quad \Rightarrow\\
k\cdot c=k\cdot d &\quad \underset{k\neq 0\wedge[\text{theorem: }5.77]}{\Rightarrow}\\
c=d &
\end{aligned}$$

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

$$x - y = x + (-y) = \sim[(n,0)] + \sim[(0,m)] = \sim[(n,m)] = z \qquad \square$$

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:

$\boldsymbol{n \leqslant m}$**.** 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 $z \in \mathbb{Z}_0^- \subseteq \mathbb{Z}_0^+ \bigcup \mathbb{Z}_0^-$.

$\boldsymbol{m < n}$**.** 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

$$\mathbb{Z}_0^+ \bigcap \mathbb{Z}_0^- = \{0\} \qquad \square$$

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

$$\begin{aligned}-(x+(-y)) &\underset{\text{[theorem: 4.8]}}{=} -x+(-(-y))\\ &\underset{\text{[theorem: 4.9]}}{=} -x+y\\ &= y+(-x)\end{aligned}$$

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

$$\begin{aligned}x\in\mathbb{Z}_0^+ &\underset{\text{[eq: 7.7]}}{\Leftrightarrow} x+(-0)\in\mathbb{Z}_0^+\\ &\Leftrightarrow 0\leqslant x\\ &\Leftrightarrow x\in\{x\in\mathbb{Z}|0\leqslant x\}\end{aligned}$$

proving

$$\mathbb{Z}_0^+=\{x\in\mathbb{Z}|0\leqslant x\}$$

Further

$$\begin{aligned} x\in\mathbb{Z}_0^- &\Leftrightarrow -x\in\mathbb{Z}_0^+\\ &\Leftrightarrow 0+(-x)\in\mathbb{Z}_0^+\\ &\Leftrightarrow x\leqslant 0\\ &\Leftrightarrow x\in\{x\in\mathbb{Z}|x\leqslant 0\}\end{aligned}$$

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. $$\begin{aligned} 0\leqslant x &\underset{[\text{theorem: } 7.22]}{\Leftrightarrow} x\in\mathbb{Z}_0^+\\ &\Leftrightarrow \exists k\in\mathbb{N}_0\,\text{such that}\,x=\sim[(k,0)]\\ &\Leftrightarrow \exists k\in\mathbb{N}_0\text{ such that } n+0=m+k\Leftrightarrow n=m+k\\ &\underset{[\text{theorem: } 5.62]}{\Leftrightarrow} m\leqslant n\end{aligned}$$

2. First

$$\begin{aligned} x\neq 0 &\Leftrightarrow \sim[(n,m)]\neq\sim[(0,0)]\\ &\Leftrightarrow n+0\neq m+0\\ &\Leftrightarrow n\neq m\end{aligned}$$

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

$$\begin{array}{rcl} n \cdot r + m \cdot s & = & (m + k) \cdot (s + l) + m \cdot s \\ & = & \underbrace{m \cdot s}_{1} + \underbrace{m \cdot l}_{2} + \underbrace{k \cdot s}_{3} + k \cdot l + \underbrace{m \cdot s}_{4} \\ m \cdot r + n \cdot s & = & m \cdot (s + l) + (m + k) \cdot s \\ & = & \underbrace{m \cdot s}_{1} + \underbrace{m \cdot l}_{2} + \underbrace{m \cdot s}_{4} + \underbrace{k \cdot s}_{3} \end{array}$$

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* $i_{\mathbb{N}_0 \rightarrow \mathbb{Z}}: \mathbb{N}_0 \rightarrow \mathbb{Z}_0^+$ *by* $i_{\mathbb{N}_0 \rightarrow \mathbb{Z}}(n) = \sim[(n,0)]$ *then*

$$i_{\mathbb{N}_0 \rightarrow \mathbb{Z}}: \langle \mathbb{N}_0, \leqslant \rangle \rightarrow \langle \mathbb{Z}_0^+, \leqslant \rangle \text{ is a order isomorphism}$$

Proof. Using [theorem: 7.17 (3)] it follows that

$$i_{\mathbb{N}_0 \rightarrow \mathbb{Z}}: \mathbb{N}_0 \rightarrow \mathbb{Z}_0^+ \text{ is a bijection}$$

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] $i_{\mathbb{N}_0 \rightarrow \mathbb{Z}}: \langle \mathbb{N}_0, \cdot \rangle \rightarrow \langle \mathbb{Z}_0^+, \cdot \rangle$ 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: 3.87] 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: 3.74].*

Proof. As by [theorem: 7.31] $\langle \mathbb{Z}_0^+, \leqslant \rangle$ is well-ordered it follows from [theorem: 3.81] 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: 3.54] 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: 3.76] $\sup((i_{\mathbb{N}_0\rightarrow\mathbb{Z}})^{-1}(A))$ 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

$$m = n \cdot q + r \quad \text{and} \quad 0 \leqslant r' < n$$

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

$$\begin{aligned} n \cdot q_1 + r_1 = n \cdot q_2 + r_2 &\Rightarrow n \cdot q_1 + (-(n \cdot q_2)) = r_2 + (-r_1) \\ &\Rightarrow n \cdot (q_1 + (-q_2)) = r_2 + (-r_1) && (7.13) \\ n \cdot q_1 + r_1 = n \cdot q_2 + r_2 &\Rightarrow n \cdot q_2 + (-(n \cdot q_1)) = r_1 + (-r_2) \\ &\Rightarrow n \cdot (q_2 + (-q_1)) = r_1 + (-r_2) && (7.14) \end{aligned}$$

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 $n \in \mathbb{Z} \setminus \{0\}$ 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 $n \in \mathbb{Z} \setminus \{0\}$ 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 $D_{n,m} = \{d \in \mathbb{Z}_0^+ | d \text{ is a common divisor of } n \text{ and } m\}$. 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

$$\begin{array}{lcl} z \text{ is odd} & \Leftrightarrow & z \text{ is not even} \\ & \underset{r=0 \Rightarrow z \text{ is evn}}{\Leftrightarrow} & z = 2 \cdot q + 1 \end{array}$$

□

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: 6.63] 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 $m\in\mathbb{Z}\backslash\{0\}$. 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. $\mathbb{Z}^*=\mathbb{Z}\backslash\{0\}$

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

$$n \cdot m' = m \cdot n' \wedge r \cdot s' = s \cdot r' \tag{8.1}$$

$$\begin{array}{rcl}
(n \cdot s + r \cdot m) \cdot (m' \cdot s') & = & (n \cdot s) \cdot (m' \cdot s') + (r \cdot m) \cdot (m' \cdot s') \\
& \underset{[\text{theorem: } 7.11]}{=} & (n \cdot m') \cdot (s \cdot s') + (r \cdot s') \cdot (m \cdot m') \\
& \underset{[\text{eq: } 8.1]}{=} & (m \cdot n') \cdot (s \cdot s') + (s \cdot r') \cdot (m \cdot m') \\
& \underset{[\text{theorem: } 7.11]}{=} & (n' \cdot s) \cdot (m \cdot s) + (r' \cdot m') \cdot (m \cdot s) \\
& \underset{[\text{theorem: } 7.11]}{=} & (n' \cdot s + r' \cdot m') \cdot (m \cdot s) \\
& \underset{[\text{theorem: } 7.11]}{=} & (m \cdot s) \cdot (n' \cdot s + r' \cdot m')
\end{array}$$

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 $+: \mathbb{Q} \times \mathbb{Q} \rightarrow \mathbb{Q}$ 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

$$\begin{array}{rcl}
\frac{a}{b} + \left(\frac{c}{d} + \frac{e}{f}\right) & = & \frac{a}{b} + \frac{c \cdot f + d \cdot e}{d \cdot f} \\
& = & \frac{a \cdot (d \cdot f) + b \cdot (c \cdot f + d \cdot e)}{b \cdot (d \cdot f)} \\
& \underset{[\text{theorem: } 7.11]}{=} & \frac{(a \cdot d) \cdot f + (c \cdot b) \cdot f + (b \cdot d) \cdot e}{(b \cdot d) \cdot f} \\
& \underset{[\text{theorem: } 7.11]}{=} & \frac{(a \cdot d + c \cdot b) \cdot f + (b \cdot d) \cdot e}{(b \cdot d) \cdot f} \\
& = & \frac{a \cdot d + c \cdot b}{b \cdot d} + \frac{e}{f} \\
& = & \left(\frac{a}{b} + \frac{c}{d}\right) + \frac{e}{f}
\end{array}$$

commutativity. Let $\frac{a}{b}, \frac{c}{d} \in \mathbb{Q}$ then

$$\begin{array}{rcl}
\frac{a}{b} + \frac{c}{d} & = & \frac{a \cdot d + b \cdot c}{b \cdot d} \\
& \underset{[\text{theorem: } 7.11]}{=} & \frac{c \cdot b + d \cdot a}{d \cdot b} \\
& = & \frac{c}{d} + \frac{a}{b}
\end{array}$$

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

$$\begin{array}{lcl}
\frac{a}{b}\cdot\frac{c}{d}+\frac{a}{b}\cdot\frac{e}{f} & = & \frac{a\cdot c}{b\cdot d}+\frac{a\cdot e}{b\cdot f}\\
 & & \frac{(a\cdot c)\cdot(b\cdot f)+(b\cdot d)\cdot(a\cdot e)}{(b\cdot d)\cdot(b\cdot f)}\\
 & \underset{\text{[theorem: 7.11]}}{=} & \frac{b\cdot(a\cdot(c\cdot f))+b\cdot(a\cdot(d\cdot e))}{b\cdot(b\cdot(d\cdot f))}\\
 & \underset{\text{[theorem: 7.11]}}{=} & \frac{b\cdot(a\cdot(c\cdot f)+a\cdot(d\cdot e))}{b\cdot(b\cdot(d\cdot f))}\\
 & \underset{b\neq 0\wedge\text{[theorem: 8.5]}}{=} & \frac{a\cdot(c\cdot f)+a\cdot(d\cdot e)}{b\cdot(d\cdot f)}\\
 & = & \frac{a}{b}\cdot\left(\frac{c\cdot f+d\cdot e}{d\cdot f}\right)\\
 & = & \frac{a}{b}\cdot\left(\frac{c}{d}+\frac{e}{f}\right)
\end{array}$$

commutativity. Let $\frac{a}{b},\frac{c}{d}\in\mathbb{Q}$ then

$$\begin{array}{lcl}
\frac{a}{b}\cdot\frac{c}{d} & = & \frac{a\cdot c}{b\cdot d}\\
 & \underset{\text{[theorem: 7.11]}}{=} & \frac{c\cdot a}{d\cdot b}\\
 & = & \frac{c}{d}\cdot\frac{a}{b}
\end{array}$$

neutral element. Let $\frac{a}{b}\in\mathbb{Q}$ then

$$\begin{array}{lcl}
\frac{1}{1}\cdot\frac{a}{b} & \underset{\text{commutativity}}{=} & \frac{a}{b}\cdot\frac{1}{1}\\
 & = & \frac{a\cdot 1}{b\cdot 1}\\
 & \underset{\text{[theorem: 7.11]}}{=} & \frac{a}{b}
\end{array}$$

associativity. Let $\frac{a}{b},\frac{c}{d},\frac{e}{f}\in\mathbb{Q}$ then

$$\begin{array}{lcl}
\frac{a}{b}\cdot\left(\frac{c}{d}\cdot\frac{e}{f}\right) & = & \frac{a}{b}\cdot\frac{c\cdot e}{d\cdot f}\\
 & = & \frac{a\cdot(c\cdot e)}{b\cdot(d\cdot f)}\\
 & \underset{\text{[theorem: 7.11]}}{=} & \frac{(a\cdot c)\cdot e}{(b\cdot d)\cdot f}\\
 & = & \frac{a\cdot c}{b\cdot d}\cdot\frac{e}{f}\\
 & = & \left(\frac{a}{b}\cdot\frac{c}{d}\right)\cdot\frac{e}{f}
\end{array}$$

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:*

$$\begin{aligned}\mathbb{Q}_0^+ &= \left\{\frac{a}{b}|(a,b)\in\mathbb{Z}\times\mathbb{Z}^*\wedge a\cdot b\in\mathbb{Z}_0^+\right\}\underset{\textit{[theorem: 7.22]}}{=}\left\{\frac{a}{b}|(a,b)\in\mathbb{Z}\times\mathbb{Z}^*\wedge 0\leqslant a\cdot b\right\}\\ \mathbb{Q}_0^- &= \left\{\frac{a}{b}|(a,b)\in\mathbb{Z}\times\mathbb{Z}^*\wedge a\cdot b\in\mathbb{Z}_0^-\right\}\underset{\textit{[theorem: 7.22]}}{=}\left\{\frac{a}{b}|(a,b)\in\mathbb{Z}\times\mathbb{Z}^*\wedge a\cdot b\leqslant 0\right\}\end{aligned}$$

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:

- $\boldsymbol{0 < a}$**.** 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 $c \cdot d \leqslant 0$.

- $\boldsymbol{a < 0}$**.** 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:

$\boldsymbol{r+(-q)\in\mathbb{Q}_0^+}$**.** Then $q\leqslant r$

$\boldsymbol{r+(-q)\in\mathbb{Q}_0^-}$**.** 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':

$\boldsymbol{0<n'\wedge 0<m'}$**.** Then $q=\frac{n'}{m'}$ so if we take $n=n'$ and $m=m'$ we have $0<n\wedge 0<m$ such that $q=\frac{n}{m}$.

$\boldsymbol{0<n'\wedge m'<0}$**.** 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.

$\boldsymbol{n'<0\wedge 0<m'}$**.** 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 $q = \frac{n}{m}$.

$\Leftarrow$. 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.

$\Rightarrow$. If $n = m$ then $\frac{n}{m} = \frac{n}{n} = \frac{1 \cdot n}{1 \cdot n} \underset{[\text{theorem: } 8.5]}{=} \frac{1}{1} = 1$.

$\Leftarrow$. 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

$$\begin{aligned}
n<m &\underset{[\text{theorem: } 7.27,\ 4.50]}{\Leftrightarrow} -m<-n\\
&\underset{(2.b)}{\Leftrightarrow} 1<\frac{-n}{-m}\\
&\underset{[\text{theorem: } 7.27,\ 4.40]}{\Leftrightarrow} 1<\frac{(-1)\cdot n}{(-1)\cdot m}\\
&\underset{[\text{theorem: } 8.5]}{=} 1<\frac{n}{m}
\end{aligned}$$

b. Then

$$\begin{aligned}
m<n &\underset{[\text{theorem: } 7.27,\ 4.50]}{\Leftrightarrow} -n<-m\\
&\underset{(2,a)}{\Leftrightarrow} \frac{-n}{-m}<1\\
&\underset{[\text{theorem: } 7.27,\ 4.40]}{\Leftrightarrow} \frac{(-1)\cdot n}{(-1)\cdot m}<1\\
&\underset{[\text{theorem: } 8.5]}{=} \frac{n}{m}<1
\end{aligned}$$

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

$$\begin{aligned}(q+s)+(-(r+s)) &\underset{[\text{theorem: } 4.8]}{=} (q+s)+((-r)+(-s))\\ &\underset{\text{commutativity}}{=} (q+s+((-s)+(-r))\\ &\underset{\text{associativity}}{=} q+(s+((-s)+(-r)))\\ &\underset{\text{associativity}}{=} q+((s+(-s))+(-r))\\ &= q+(0+(-r))\\ &= q+(-r)\end{aligned}$$

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

$$\begin{array}{rcl}
i_{\mathbb{Z}\rightarrow\mathbb{Q}}(x)\leqslant i_{\mathbb{Z}\rightarrow\mathbb{Q}}(y) & \Leftrightarrow & i_{\mathbb{Z}\rightarrow\mathbb{Q}}(y)+(-i_{\mathbb{Z}\rightarrow\mathbb{Q}}(x))\in\mathbb{Q}_0^+\\
 & \underset{[\text{eq: } 8.15]}{\Leftrightarrow} & \frac{y+(-x)}{1}\in\mathbb{Q}_0^+\\
 & \underset{\text{def } \mathbb{Q}_0^+}{\Leftrightarrow} & (y+(-x))\cdot 1\in\mathbb{Z}_0^+\\
 & \Leftrightarrow & y+(-x)\in\mathbb{Z}_0^+\\
 & \Leftrightarrow & x\leqslant y
\end{array}$$

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

$$i_{\mathbb{Z}\rightarrow\mathbb{Q}}:\langle\mathbb{Z},+,\cdot\rangle\rightarrow\langle\mathbb{Z}_{\mathbb{Q}},+,\cdot\rangle \text{ is a ring isomorphism}$$ □

Theorem 8.33. $\mathbb{Q}=\{x\cdot y^{-1}|x\in\mathbb{Z}_{\mathbb{Q}}\wedge y\in\mathbb{Z}_{\mathbb{Q}}\backslash\{0\}\}$

Proof. If $q\in\mathbb{Q}$ there exists $x'\in\mathbb{Z}$ and $y'\in\mathbb{Z}^*=\mathbb{Z}\backslash\{0\}$ 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 $y\in\mathbb{Z}_{\mathbb{Q}}\backslash\{0\}$. Further

$$q=\frac{x'}{y'}=\frac{x'}{1}\cdot\frac{1}{y'}=\frac{x'}{1}\cdot\left(\frac{y'}{1}\right)^{-1}=x\cdot y^{-1}\in\{x\cdot y^{-1}|x\in\mathbb{Z}_{\mathbb{Q}}\wedge y\in\mathbb{Z}_{\mathbb{Q}}\backslash\{0\}\}$$

proving

$$\mathbb{Q}\subseteq\{x\cdot y^{-1}|x\in\mathbb{Z}_{\mathbb{Q}}\wedge y\in\mathbb{Z}_{\mathbb{Q}}\backslash\{0\}\} \tag{8.16}$$

If $q\in\{x\cdot y^{-1}|x\in\mathbb{Z}_{\mathbb{Q}}\wedge y\in\mathbb{Z}_{\mathbb{Q}}\backslash\{0\}\}$ there exists a $x\in\mathbb{Z}_{\mathbb{Q}}$ and $y\in\mathbb{Z}_{\mathbb{Q}}\backslash\{0\}$ such that $q=x\cdot y^{-1}$. Using [theorem: 8.32] again there exists $x'\in\mathbb{Z}$ and $y'\in\mathbb{Z}\backslash\{0\}$ 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 $\{x\cdot y^{-1}|x\in\mathbb{Z}_{\mathbb{Q}}\wedge y\in\mathbb{Z}_{\mathbb{Q}}\backslash\{0\}\}\subseteq\mathbb{Q}$ which combined with [eq: 8.16] results in

$$\mathbb{Q}=\{x\cdot y^{-1}|x\in\mathbb{Z}_{\mathbb{Q}}\wedge y\in\mathbb{Z}_{\mathbb{Q}}\backslash\{0\}\}$$ □

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*

$$i_{\mathbb{N}_0\rightarrow\mathbb{Z}}:\mathbb{N}_0\rightarrow\mathbb{Z} \textit{is defined by } i_{\mathbb{N}_0\rightarrow\mathbb{Z}}(n)=\sim[n,0]$$
$$i_{\mathbb{Z}\rightarrow\mathbb{Q}}:\mathbb{Z}\rightarrow\mathbb{Q} \textit{ is defined by } i_{\mathbb{Z}\rightarrow\mathbb{Q}}(z)=\frac{z}{1}$$

Theorem 8.35. *We have that*

1. $\mathbb{N}_{0,\mathbb{Q}} = \left\{\frac{n}{1} | n \in \mathbb{Z}_0^+\right\}$
2. *If we define* $i_{\mathbb{N}_0 \rightarrow \mathbb{Q}}: \mathbb{N}_0 \rightarrow \mathbb{N}_{0,\mathbb{Q}}$ *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*
   $$i_{\mathbb{N}_0 \rightarrow \mathbb{Q}}: \langle \mathbb{N}_0, + \rangle \rightarrow \langle \mathbb{N}_{0,\mathbb{Q}}, + \rangle \text{ is a group isomorphism}$$
   b. $\langle \mathbb{N}_{0,\mathbb{Q}}, \cdot \rangle$ *is a sub semi-group of* $\langle \mathbb{Q}, \cdot \rangle$ *and*
   $$i_{\mathbb{N}_0 \rightarrow \mathbb{Q}}: \langle \mathbb{N}_0, \cdot \rangle \rightarrow \langle \mathbb{N}_{0,\mathbb{Q}}, \cdot \rangle \text{ is a group isomorphism}$$
   c.
   $$i_{\mathbb{N}_0 \rightarrow \mathbb{Q}}: \langle \mathbb{N}_0, \leqslant \rangle \rightarrow \langle \mathbb{N}_{0,\mathbb{Q}}, \leqslant \rangle \text{ is a group isomorphism}$$
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
$$\begin{array}{lcl} x \in \mathbb{N}_{0,\mathbb{Q}} & \Leftrightarrow & \exists n \subset \mathbb{N}_0 \text{ such that } x = (i_{\mathbb{Z} \rightarrow \mathbb{Q}} \circ i_{\mathbb{N}_0 \rightarrow \mathbb{Z}})(n) \\ & \Rightarrow & x = i_{\mathbb{Z} \rightarrow \mathbb{Q}}(i_{\mathbb{N}_0 \rightarrow \mathbb{Z}}(n)) \\ & \Rightarrow & x = \frac{\sim[(n,0)]}{1} \\ & \underset{\sim[(n,0)] \in \mathbb{Z}_0^+}{\Rightarrow} & x \in \left\{\frac{n}{1} | n \in \mathbb{Z}_0^+\right\} \end{array}$$
proving that
$$\mathbb{N}_{0,\mathbb{Q}} \subseteq \left\{\frac{n}{1} | n \in \mathbb{Z}_0^+\right\} \tag{8.17}$$
Further
$$\begin{array}{lcl} x \in \left\{\frac{n}{1} | n \in \mathbb{Z}_0^+\right\} & \Rightarrow & \exists n \in \mathbb{Z}_0^+ \text{ such that } x = \frac{n}{1} \\ & \underset{\text{definition of } \mathbb{Z}_0^+}{\Rightarrow} & \exists n' \in \mathbb{N}_0 \text{ such that } n = \sim[(n',0)] \\ & \Rightarrow & x = \frac{\sim[(n,0)]}{1} \\ & \Rightarrow & x = i_{\mathbb{Z} \rightarrow \mathbb{Q}}(i_{\mathbb{N}_0 \rightarrow \mathbb{Z}}(n')) \\ & \Rightarrow & (i_{\mathbb{Z} \rightarrow \mathbb{Q}} \circ i_{\mathbb{N}_0 \rightarrow \mathbb{Z}})(n') \\ & \Rightarrow & x \in \mathbb{N}_{0,\mathbb{Q}} \end{array}$$
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
   $$i_{\mathbb{N}_0 \rightarrow \mathbb{Z}}: \langle \mathbb{N}_0, + \rangle \rightarrow \langle \mathbb{Z}_0^+, + \rangle \text{ and } i_{\mathbb{Z} \rightarrow \mathbb{Q}}: \langle \mathbb{Z}, + \rangle \rightarrow \langle \mathbb{Z}_{\mathbb{Q}}, + \rangle \text{ are group isomorphisms}$$
   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 $i_{\mathbb{Z} \rightarrow \mathbb{Q}} \circ i_{\mathbb{N}_0 \rightarrow \mathbb{Z}}: \langle \mathbb{N}_0, + \rangle \rightarrow \langle (i_{\mathbb{Z} \rightarrow \mathbb{Q}} \circ i_{\mathbb{N}_0 \rightarrow \mathbb{Z}})(\mathbb{N}_0), + \rangle$ 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
   $$\mathbb{N}_{0,\mathbb{Q}} \text{ is a subgroup of } \langle \mathbb{Q}, + \rangle \text{ and } i_{\mathbb{N}_0 \rightarrow \mathbb{Q}}: \langle \mathbb{N}_0, + \rangle \rightarrow \langle \mathbb{N}_{0,\mathbb{Q}}, + \rangle \text{ is a group isomorphism.}$$
   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
   $$i_{\mathbb{N}_0 \rightarrow \mathbb{Z}}: \langle \mathbb{N}_0, \cdot \rangle \rightarrow \langle \mathbb{Z}_0^+, \cdot \rangle \text{ and } i_{\mathbb{Z} \rightarrow \mathbb{Q}}: \langle \mathbb{Z}, \cdot \rangle \rightarrow \langle \mathbb{Z}_{\mathbb{Q}}, \cdot \rangle \text{ are group isomorphisms}$$

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: 3.51] 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: 3.79] 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: 3.81] 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 $n'=n:\gcd(n,m)$ and $m'=m:\gcd(n,m)$ 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

$$\begin{aligned} \left(u + \frac{1}{n}\right) \cdot \left(u + \frac{1}{n}\right) < \frac{2}{1} &\Leftrightarrow u \cdot u + \frac{2}{1} \cdot u \cdot \frac{1}{n} + \frac{1}{n} \cdot \frac{1}{n} < \frac{2}{1} \\ &\Leftrightarrow \frac{2}{1} \cdot u \cdot \frac{1}{n} + \frac{1}{n} \cdot \frac{1}{n} < \frac{2}{1} - u \cdot u \\ &\Leftrightarrow \frac{2}{n} \cdot u + \frac{1}{n} \cdot \frac{1}{n} < \frac{2}{1} - u \cdot u \qquad (8.28) \end{aligned}$$

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

$$\begin{array}{rcl} \frac{2}{n} \cdot u + \frac{1}{n} \cdot \frac{1}{n} & \leqslant_{[\text{eq: } 8.29]} & \frac{2}{n} \cdot u + \frac{1}{n} \\ 1 & \leqslant_{[\text{eq: } 8.30]} & \frac{4}{n} + \frac{1}{n} \\ & = & \left(\frac{4}{1} + \frac{1}{1}\right) \cdot \frac{1}{n} \\ & = & \frac{5}{1} \cdot \frac{1}{n} \\ & = & \frac{5}{n} \\ & <_{[\text{eq: } 8.27]} & \frac{2}{1} - u \cdot u \end{array}$$

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

$$\begin{array}{rcl} \frac{2}{1} \cdot u < n' \cdot \left(u \cdot u + \frac{-2}{1}\right) & \Rightarrow & \frac{2}{1} \cdot u < \frac{n}{1} \cdot \left(u \cdot u + \frac{-2}{1}\right) \\ & \underset{0 < \frac{1}{n} \wedge [\text{theorems: } 8.29,\ 4.73]}{\Rightarrow} & \left(\frac{2}{1} \cdot u\right) \cdot \frac{1}{n} < \left(\frac{n}{1} \cdot \left(u \cdot u + \frac{-2}{1}\right)\right) \cdot \frac{1}{n} \\ & \Rightarrow & \frac{2}{n} \cdot u < u \cdot u + \frac{-2}{1} \\ & \Rightarrow & \frac{2}{1} < u \cdot u + \frac{-2}{n} \cdot u \end{array}$$

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: 3.68] 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] $i_{\mathbb{N}_0\rightarrow Q}:\mathbb{N}_0\rightarrow\mathbb{N}_{0,\mathbb{Q}}$ 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]

$$i_{\mathbb{Z}\rightarrow\mathbb{Q}}:\mathbb{Q}\rightarrow\mathbb{Z}_{\mathbb{Q}}$$

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 $f:\mathbb{Z}\times\mathbb{Z}\rightarrow\mathbb{Q}$ 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

$$f:\mathbb{Z}\times\mathbb{Z}\rightarrow\mathbb{Q}\,\text{is a surjection}$$

As $\mathbb{Z}$ is denumerable [see theorem: 7.53] we have by [theorem: 6.60] that $\mathbb{Z}\times\mathbb{Z}$ is denumerable, hence there exist a bijection $g:\mathbb{N}_0\rightarrow\mathbb{Z}\times\mathbb{Z}$, so $f\circ g:\mathbb{N}_0\rightarrow\mathbb{Q}$ is a surjection. By [theorem: 6.67] $\mathbb{Q}$ 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 $\mathbb{Q} \backslash \alpha \neq \varnothing$]*
3. $\forall q \in \alpha \wedge \forall r \in \mathbb{Q} \backslash \alpha$ *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 $\alpha \neq \varnothing \neq Q \backslash \alpha$ where every element in α is strict lower of elements in $\mathbb{Q} \backslash \alpha$ 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*

$$\mathbb{R} = \{\alpha \subseteq \mathbb{Q} | \alpha \textit{ is a Dedekind cut}\}$$

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 $r \in \mathbb{Q} \backslash \alpha$ 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* $\Leftrightarrow$ $q = \min(\mathbb{Q} \backslash \alpha)$ *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 $q \in \mathbb{Q} \backslash \alpha_q$ so that $\mathbb{Q} \backslash \alpha_q \neq \varnothing$.
3. If $r \in \alpha_q$ and $s \in \mathbb{Q} \backslash \alpha_q$ 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 $\alpha=\{r\in\mathbb{Q}|r<q\}=\alpha_q$

$\Leftarrow$. 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

$$\begin{array}{rcl} r-q & = & (q'+(k'\cdot k^{-1})\cdot(r'-q'))-(q'+((k'-1)\cdot k^{-1})\cdot(r'-q')) \\ & = & (q'+k'-(q'+(k'-1)))\cdot k^{-1}\cdot(r'-q') \\ & = & k^{-1}\cdot(r'-q') \\ & <\varepsilon & \text{[see eq: 9.1]} \end{array}$$

□

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 ℝ

9.1.2.1 Addition in 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 ℝ we must prove that $\alpha + \beta$ is a Dedekind cut, hence a element of 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 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 $\alpha_0\subseteq\alpha+-\alpha$- 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:

$0\in\alpha$. then $\alpha\in\mathbb{R}^+$ so that $\alpha\in\mathbb{R}^+\bigcup\mathbb{R}^-\bigcup\{0\}$

$0\notin\alpha$. then we have either:

- **$\min(\mathbb{Q}\backslash\alpha)$ does not exist.** As $0\notin\alpha$ we have $-0=0\in\mathbb{Q}\backslash\alpha$ and as $\min(\mathbb{Q}\backslash\alpha)$ does not exist there exist a $s\in\mathbb{Q}\backslash\alpha$ such that $s<0=-0$ so that $0\in-\alpha$. Hence $-\alpha\in\mathbb{R}^+$ proving that $\alpha\in\mathbb{R}^-\subseteq\mathbb{R}^+\bigcup\mathbb{R}^-\bigcup\{0\}$.
- **$\min(\mathbb{Q}\backslash\alpha)$ exist.** Then by [theorem: 9.4] $\alpha=\alpha_m$ where $m=\min(\mathbb{Q}\backslash\alpha)$
  - **$0=m$.** Then $\alpha=\alpha_0=0$ so that $\alpha\in\{0\}\subseteq\mathbb{R}^+\bigcup\mathbb{R}^-\bigcup\{0\}$.
  - **$0<m$.** Then $0\in\alpha_m=\alpha$ so that $\alpha\in\mathbb{R}^+\subseteq\mathbb{R}^+\bigcup\mathbb{R}^-\bigcup\{0\}$.
  - **$m<0$.** Then $0\notin\alpha_m=\alpha$ so that $-0=0\in\mathbb{Q}\backslash\alpha$ and as $m<0=-0$, $m\in\mathbb{Q}\backslash\alpha$ 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:

   $\boldsymbol{q \in \{r \in \mathbb{Q} | r \leqslant 0\}}$**.** 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 $q < r$.

   $\boldsymbol{q \notin \{r \in \mathbb{Q} | r \leqslant 0\}}$**.** 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 $r \in \mathbb{Q} \backslash \alpha \odot \beta$ 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 $t^{-1} \cdot s' \in \mathbb{Q} \backslash \alpha$ 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 $r \in \mathbb{Q} \backslash \alpha \odot \beta$, 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:

   $\boldsymbol{m \in \{r \in \mathbb{Q} | r \leqslant 0\}}$**.** 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.

   $\boldsymbol{m \notin \{r \in \mathbb{Q} | r \leqslant 0\}}$**.** 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:

- **$x \leqslant 0$.** As $\alpha \in \mathbb{R}^+$ we have that $0 \in \alpha$ and as $x \leqslant 0$ it follows from [theorem: 9.3] that $x \in \alpha$.
- **$0 < x$.** 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:

- **$x \leqslant 0$.** Then $x \in \{q \in \mathbb{Q} | q \leqslant 0\}$ so that $x \in \alpha_1 \odot \alpha$
- **$0 < x$.** 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

$$\begin{array}{lcl} q \in \alpha \odot \beta & \Leftrightarrow & q \in \{r \in \mathbb{Q} | r \leqslant 0\} \bigcup \{s \cdot t | (s, t) \in \alpha \times \beta \wedge 0 < s \wedge 0 < t\} \\ & \Leftrightarrow & q \leqslant 0 \vee \exists (s, t) \in \alpha \times \beta \text{ with } 0 < s \wedge 0 < t \text{ such that } q = s \cdot t \\ & \underset{\text{[theorem: 8.12}}{\Leftrightarrow} & q \leqslant 0 \vee \exists (s, t) \in \alpha \times \beta \text{ with } 0 < s \wedge 0 < t \text{ such that } q = t \cdot s \\ & \Leftrightarrow & q \leqslant 0 \vee \exists (t, s) \in \beta \times \alpha \text{ with } 0 < t \wedge 0 < s \text{ such that } q = s \cdot t \\ & \Leftrightarrow & q \in \beta \odot \alpha \end{array}$$

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

$x \leqslant 0$. 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: 9.30] 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 $x \underset{\text{[eq: 9.37]}}{=} s \cdot u + s \cdot v \in \alpha \odot \beta + s \odot \gamma$

$\boldsymbol{0<u\wedge 0<v}$**.** 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:

$\boldsymbol{x\leqslant 0}$**.** Then $x\in\mathbb{Q}_0^-$ so that $x\in\alpha\odot(\beta+\gamma)$

$\boldsymbol{0<x}$**.** Then we have to look at the following sub cases:

$\boldsymbol{r\leqslant 0\wedge t\leqslant 0}$**.** Then $x\underset{\text{[eq: 9.39]}}{=}r+t\leqslant 0$ a contradicting $0<x$, so this case does not occur.

$\boldsymbol{r\leqslant 0\wedge 0<t}$**.** 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

$$x\in\alpha\odot(\beta+\gamma)$$

$\boldsymbol{0<r\wedge t\leqslant 0}$**.** 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

$$x\in\alpha\odot(\beta+\gamma)$$

$\boldsymbol{0<r\wedge 0<t}$**.** 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:

$\boldsymbol{u=u'}$**.** 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

$$x\in\alpha\odot(\beta+\gamma)$$

$\boldsymbol{u<u'}$**.** 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

$$x\in\alpha\odot(\beta+\gamma)$$

$\boldsymbol{u'<u}$**.** 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

$$\alpha \odot (\beta + \gamma) = \alpha \odot \beta + \alpha \odot \gamma \qquad \square$$

Theorem 9.25. *Let $\alpha \in \mathbb{R}^+$ then* $\operatorname{inv}(\alpha)$ *defined by*

$$\begin{aligned} \operatorname{inv}(\alpha) &= \{r \in \mathbb{Q} | r \leqslant 0\} \bigcup \{s^{-1} | s \in \mathbb{Q} \backslash \alpha \textit{ with } 0 < s \textit{ and } \exists t \in \mathbb{Q} \backslash \alpha \vdash t < s\} \\ &\underset{\mathbb{Q}_0^- = \{r \in \mathbb{Q} | r \leqslant 0\}}{=} \mathbb{Q}_0^- \bigcup \{s^{-1} | s \in \mathbb{Q} \backslash \alpha \textit{ with } 0 < s \textit{ and } \exists t \in \mathbb{Q} \backslash \alpha \vdash t < s\} \end{aligned}$$

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

$$s^{-1} \in \{s^{-1} | s \in \mathbb{Q} \backslash \alpha \text{ with } 0 < s \text{ and } \exists t \in \mathbb{Q} \backslash \alpha \vDash t < s\}$$

so that $\exists t \in \mathbb{Q} \backslash \alpha$ such that $s^{-1} = t^{-1} \underset{\text{[theorem: 4.54]]}}{\Rightarrow} s = t$. So $s \in \mathbb{Q} \backslash \alpha$ contradicting $s \in \alpha$ hence $s^{-1} \notin \operatorname{inv}(\alpha)$ proving that

$$\operatorname{inv}(\alpha) \neq \mathbb{Q}$$

3. Let $q \in \operatorname{inv}(\alpha)$ and $r \in \mathbb{Q} \backslash \operatorname{inv}(\alpha)$. For q we have the following possibilities:

$q \leqslant 0$. Then as $r \in \mathbb{Q} \backslash \operatorname{inv}(\alpha)$ we have $r \notin \operatorname{inv}(\alpha)$ hence $r \notin \mathbb{Q}_0^-$ so that $0 < r$ giving

$$q < r$$

$0 < q$. Then $q \notin \mathbb{Q}_0^-$ hence, as $q \in \operatorname{inv}(\alpha)$, we have:

$$q \in \{s^{-1} | s \in \mathbb{Q} \backslash \alpha \text{ with } 0 < s \text{ and } \exists t \in \mathbb{Q} \backslash \alpha \vdash t < s\}$$

so there exist a $s \in \mathbb{Q} \backslash \alpha$ with $0 < s$ and $\exists t \in \mathbb{Q} \backslash \alpha \vdash t < s$ such that $q = s^{-1}$, as $q^{-1} = (s^{-1})^{-1} \underset{\text{[theorem: 4.53]}}{=} s$ we have that

$$q^{-1} \in \mathbb{Q} \backslash \alpha,\ 0 < q^{-1} \text{ and } \exists t \in \mathbb{Q} \backslash \alpha \vdash t < q^{-1} \tag{9.42}$$

Further as $r \in \mathbb{Q} \backslash \operatorname{inv}(\alpha)$ 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 $q^{-1} \in \mathbb{Q} \backslash \alpha$ [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 $r^{-1} \in \mathbb{Q} \backslash \alpha$ and we have to look at the following possibilities

$\forall t \in \mathbb{Q} \vDash r^{-1} \leqslant t$. Then as $q^{-1} \in \mathbb{Q} \backslash \alpha$ [see eq: 9.42] we have that $r^{-1} \leqslant q^{-1}$. If $r^{-1} = q^{-1}$ we have by [eq: 9.42] a $t \in \mathbb{Q} \backslash \alpha$ 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 $r^{-1} \in \mathbb{Q} \backslash \alpha$ and $0 < r^{-1}$ we have that $r = (r^{-1})^{-1} \in \{s^{-1} | s \in \mathbb{Q} \backslash \alpha$ with $0 < s$ and $\exists t \in \mathbb{Q} \backslash \alpha \vdash t < s\}$ so that $r \in \operatorname{inv}(\alpha)$ contradicting $r \in \mathbb{Q} \backslash \operatorname{inv}(\alpha)$ so this case does not occur.

Hence is all valid cases we have

$$q < r$$

4. Assume that $\operatorname{inv}(\alpha)$ has a greatest element m then we have

$$m \in \operatorname{inv}(\alpha) \text{ and } \forall s \in \operatorname{inv}(\alpha) \text{ we have } s \leqslant m \tag{9.44}$$

For m we have to look at the following possibilities:

$m \leqslant 0$. Using [definition: 9.1 (2)] $\varnothing \neq \mathbb{Q} \backslash \alpha$ so there exist a $r \in \mathbb{Q} \backslash \alpha$. 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 $r \in \mathbb{Q} \backslash \alpha$ we have by [definition: 9.1 (3)] that $r + 1 < r$ a contradiction, so we must have that $r + 1 \notin \alpha$ or $r + 1 \in \mathbb{Q} \backslash \alpha$. As further $r < r + 1$ and $0 < r + 1$ it follows that $(r+1)^{-1} \in \{s^{-1} | s \in \mathbb{Q} \backslash \alpha$ with $0 < s$ and $\exists t \in \mathbb{Q} \backslash \alpha \vdash t < s\}$ so that $(r+1)^{-1} \in \operatorname{inv}(\alpha)$ 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 $m \in \operatorname{inv}(\alpha)$, we have $m \in \{s^{-1} | s \in \mathbb{Q} \backslash \alpha$ with $0 < s$ and $\exists t \in \mathbb{Q} \backslash \alpha \vdash t < s\}$ so there exist a $s \in \mathbb{Q} \backslash \alpha$ with $0 < s$ and $\exists t \in \mathbb{Q} \backslash \alpha \vdash t < s$ such that $m = s^{-1}$, hence $m^{-1} = s$ so that:

$$m^{-1} \in \mathbb{Q} \backslash \alpha,\ 0 < m^{-1} \text{ and } \exists t \in \mathbb{Q} \backslash \alpha \text{ such that } t < m^{-1} \tag{9.46}$$

As $t \in \mathbb{Q} \backslash \alpha$ 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:

$x \leqslant 0$. Then as $0 < 1$ we have $x < 1$ hence $x \in \alpha_1$

$0 < x$. 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:

$t \leqslant 0$. Then from $0 < s$ we have that $x = s \cdot t \leqslant 0 < 1$ hence $x \in \alpha_1$.

$0 < t$. 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:

$x \leqslant 0$. Then $x \in \mathbb{Q}_0^-$ so that $x \in \alpha \odot \operatorname{inv}(\alpha)$.

$0 < x$. 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:

$s_3 < s_1$. 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

$$\begin{aligned}
s_1 \cdot (s_1 + \varepsilon)^{-1} &\underset{\text{[eq: 9.50]}}{=} s_1 \cdot (s_1 + s_1 \cdot (1 - x) \cdot x^{-1})^{-1} \\
&= s_1 \cdot (s_1 + s_1 \cdot x^{-1} - s_1 \cdot x^{-1} \cdot x)^{-1} \\
&= s_1 \cdot (s_1 + s_1 \cdot x^{-1} - s_1)^{-1} \\
&= s_1 \cdot (s_1 \cdot x^{-1})^{-1} \\
&\underset{\text{[theorem: 4.55]}}{=} s_1 \cdot (s_1^{-1} \cdot (x^{-1})^{-1}) \\
&= (x^{-1})^{-1} \\
&= x
\end{aligned}$$

proving using [eq: 9.54[that

$$x \in \alpha \odot \operatorname{inv}(\alpha)$$

$s_1 \leqslant s_3$. Then as $0 \in \alpha$ and $s_3 + \varepsilon \in \mathbb{Q}\backslash\alpha$ [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 $\text{inv}(\alpha)$ together with [eqs: 9.51, 9.55] that

$$(s_3+\varepsilon)^{-1}\in\text{inv}(\alpha) \tag{9.56}$$

Now by [eq: 9.49] $0<s_1\leqslant s_3\in\alpha$ and $0<(s_3+\varepsilon)^{-1}\in\text{inv}(\alpha)$ [see eq: 9.55, 9.51] so that

$$s_3\cdot(s_3+\varepsilon)^{-1}\in\alpha\odot\text{inv}(\alpha) \tag{9.57}$$

Now

$$\begin{aligned}
s_1\leqslant s_3 \quad & \underset{0<1-x}{\Rightarrow} && s_1\cdot(1-x)\leqslant s_3\cdot(1-x)\\
& \underset{0<x^{-1}}{\Rightarrow} && s_1\cdot(1-x)\cdot x^{-1}\leqslant s_3\cdot(1-x)\cdot x^{-1}\\
& \Rightarrow && s_3+s_1\cdot(1-x)\cdot x^{-1}\leqslant s_3+s_3\cdot(1-x)\cdot x^{-1}\\
& \underset{[\text{theorems: } 8.29,\ 4.73]}{\Rightarrow} && (s_3+s_1\cdot(1-x)\cdot x^{-1})^{-1}\geqslant(s_3+s_3\cdot(1-x)\cdot x^{-1})^{-1}\\
& \underset{[\text{eq: } 9.50]}{\Rightarrow} && (s_3+\varepsilon)^{-1}\geqslant(s_3+s_3\cdot(1-x)\cdot x^{-1})^{-1}\\
& \underset{0<s_1\leqslant s_3}{\Rightarrow} && s_3\cdot(s_3+\varepsilon)^{-1}\geqslant s_3\cdot(s_3+s_3\cdot(1-x)\cdot x^{-1})^{-1}
\end{aligned} \tag{9.58}$$

Further

$$\begin{aligned}
s_3\cdot(s_3+\varepsilon)^{-1} \quad & \geqslant_{[\text{eq: } 9.58]} && s_3\cdot(s_3+s_3\cdot(1-x)\cdot x^{-1})^{-1}\\
& = && s_3\cdot(s_3+s_3\cdot x^{-1}-s_3\cdot x\cdot x^{-1})^{-1}\\
& = && s_3\cdot(s_3+s_3\cdot x^{-1}-s_3)^{-1}\\
& = && s_3\cdot(s_3\cdot x^{-1})^{-1}\\
& \underset{[\text{theorem: } 4.55]}{=} && s_3\cdot s_3^{-1}\cdot(x^{-1})^{-1}\\
& = && x
\end{aligned}$$

which by [theorem: 9.3] and [eq: 9.57]proves that

$$x\in\alpha\odot\text{inv}(\alpha)$$

As in all cases $x\in\alpha\odot\text{in}(\alpha)$ it follows that $\alpha_1\subseteq\alpha\odot\text{inv}(\alpha)$ which combined with [eq: 9.48] proves finally that

$$\alpha_1=\alpha\odot\text{inv}(\alpha)$$ □

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*

$$\mathbb{R}\times\mathbb{R}=(\mathbb{R}^+\times\mathbb{R}^+)\bigcup(\mathbb{R}^+\times\mathbb{R}^-)\bigcup(\mathbb{R}^-\times\mathbb{R}^+)\bigcup(\mathbb{R}^-\times\mathbb{R}^-)\bigcup\big((\mathbb{R}\times\{0\})\bigcup(\{0\}\times\mathbb{R})\big)$$

where

$$\begin{gathered}
(\mathbb{R}^+\times\mathbb{R}^+)\bigcap(\mathbb{R}^+\times\mathbb{R}^-)=\varnothing\\
(\mathbb{R}^+\times\mathbb{R}^+)\bigcap(\mathbb{R}^-\times\mathbb{R}^+)=\varnothing\\
(\mathbb{R}^+\times\mathbb{R}^+)\bigcap(\mathbb{R}^-\times\mathbb{R}^-)=\varnothing\\
(\mathbb{R}^+\times\mathbb{R}^+)\bigcap\big((\mathbb{R}\times\{0\})\bigcup(\{0\}\times\mathbb{R})\big)=\varnothing\\
(\mathbb{R}^+\times\mathbb{R}^-)\bigcap(\mathbb{R}^-\times\mathbb{R}^+)=\varnothing\\
(\mathbb{R}^+\times\mathbb{R}^-)\bigcap(\mathbb{R}^-\times\mathbb{R}^-)=\varnothing\\
(\mathbb{R}^+\times\mathbb{R}^-)\bigcap\big((\mathbb{R}\times\{0\})\bigcup(\{0\}\times\mathbb{R})\big)=\varnothing\\
(\mathbb{R}^-\times\mathbb{R}^+)\bigcap(\mathbb{R}^-\times\mathbb{R}^-)=\varnothing\\
(\mathbb{R}^-\times\mathbb{R}^+)\bigcap\big((\mathbb{R}\times\{0\})\bigcup(\{0\}\times\mathbb{R})\big)=\varnothing\\
(\mathbb{R}^-\times\mathbb{R}^-)\bigcap\big((\mathbb{R}\times\{0\})\bigcup(\{0\}\times\mathbb{R})\big)=\varnothing
\end{gathered}$$

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\right.\\
&\qquad \left.(\{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\right.\\
&\qquad \left.\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\right.\\
&\qquad \left.(\{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\right.\\
&\qquad \left.\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\right.\right.\\
&\qquad \left.\left.(\{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\right.\\
&\qquad \left.\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}
$$

$$\begin{aligned}(\mathbb{R}^- \times \mathbb{R}^-)\bigcap\big((\mathbb{R}\times\{0\})\bigcup(\{0\}\times\mathbb{R})\big) &= \big((\mathbb{R}^-\times\mathbb{R}^-)\bigcap(\mathbb{R}\times\{0\})\big)\bigcup\big((\mathbb{R}^-\times\mathbb{R}^-)\bigcap \\ & \quad (\{0\}\times\mathbb{R})\big) \\ &= \big((\mathbb{R}^-\bigcap\mathbb{R})\times(\mathbb{R}^-\bigcap\{0\})\big)\bigcup\big((\mathbb{R}^-\bigcap\{0\})\times \\ & \quad (\mathbb{R}^-\bigcap\mathbb{R})\big) \\ &\underset{[\text{eq: } 9.60]}{=} \big((\mathbb{R}^-\bigcap\mathbb{R})\times\varnothing\big)\bigcup\big(\varnothing\times(\mathbb{R}^-\bigcap\mathbb{R})\big) \\ &= \varnothing\bigcup\varnothing \\ &= \varnothing \\ & \quad \square\end{aligned}$$

We use now [theorem: 9.20] and [theorem: 9.27] to define the multiplication operator on ℝ.

Definition 9.28. *The multiplication operator* $\cdot\colon\mathbb{R}\times\mathbb{R}\Rightarrow\mathbb{R}$ *is defined as*

$$\alpha\cdot\beta=\begin{cases}\alpha\odot\beta \textit{ if } (\alpha,b)\in\mathbb{R}^+\times\mathbb{R}^+\\ -((-\alpha)\odot\beta) \textit{ if } (\alpha,\beta)\in\mathbb{R}^-\times\mathbb{R}^+\\ -(\alpha\odot(-\beta)) \textit{ if } (\alpha,\beta)\in\mathbb{R}^+\times\mathbb{R}^-\\ (-\alpha)\odot(-\beta) \textit{ if } (\alpha,\beta)\in\mathbb{R}^-\times\mathbb{R}^-\\ 0 \textit{ if } (\alpha,\beta) \textit{ if } (\alpha,\beta)\in(\mathbb{R}\times\{0\})\bigcup(\{0\}\times\mathbb{R})\end{cases}$$

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]:

$(\alpha,\beta)\in\mathbb{R}^+\times\mathbb{R}^+$. Then

$$\begin{aligned}(-\alpha)\cdot\beta &\underset{-\alpha\in\mathbb{R}^-\wedge\beta\in\mathbb{R}^+}{=} -((-(-\alpha))\odot\beta) \\ &\underset{-(-\alpha)=\alpha}{=} -(\alpha\odot\beta) \\ &\underset{\alpha\in\mathbb{R}^+\wedge\beta\in\mathbb{R}^+}{=} -(\alpha\cdot\beta) \\ \alpha\cdot(-\beta) &\underset{\alpha\in\mathbb{R}^+\wedge-\beta\in\mathbb{R}^-}{=} -(\alpha\odot(-(-\beta))) \\ &= -(\alpha\odot\beta) \\ &\underset{\alpha\in\mathbb{R}^+\wedge\beta\in\mathbb{R}^+}{=} -(\alpha\cdot\beta)\end{aligned}$$

$(\alpha,\beta)\in\mathbb{R}^+\times\mathbb{R}^-$. Then

$$\begin{aligned}-(\alpha\cdot\beta) &\underset{\alpha\in\mathbb{R}^+\wedge\beta\in\mathbb{R}^-}{=} -(-(\alpha\odot(-\beta))) \\ &= \alpha\odot(-\beta) \\ &\underset{\alpha\in\mathbb{R}^+\wedge-\beta\in\mathbb{R}^-}{=} \alpha\cdot(-\beta) \qquad (9.63) \\ (-\alpha)\cdot\beta &\underset{-\alpha\in\mathbb{R}^-\wedge\beta\in\mathbb{R}^-}{=} (-(-\alpha))\odot(-\beta) \\ &= \alpha\odot(-\beta) \\ &\underset{\alpha\in\mathbb{R}^+\wedge-\beta\in\mathbb{R}^+}{=} \alpha\cdot(-\beta) \\ &\underset{[\text{eq:}9.63}{=} -(\alpha\cdot\beta)\end{aligned}$$

$(\alpha,\beta)\in\mathbb{R}^-\times\mathbb{R}^+$. Then

$$\begin{align*}
-(\alpha\cdot\beta) &\underset{\alpha\in\mathbb{R}^-\wedge\beta\in\mathbb{R}^+}{=} -(-((-\alpha)\odot\beta)) \\
&= (-\alpha)\odot\beta \\
&\underset{-\alpha\in\mathbb{R}^+\wedge\beta\in\mathbb{R}^+}{=} (-\alpha)\cdot\beta \tag{9.64}\\
\alpha\cdot(-\beta) &\underset{\alpha\in\mathbb{R}^-\wedge-\beta\in\mathbb{R}^-}{=} (-\alpha)\odot(-(-\beta)) \\
&= (-\alpha)\odot\beta \\
&\underset{-\alpha\in\mathbb{R}^+\wedge\beta\in\mathbb{R}^+}{=} (-\alpha)\cdot\beta \\
&\underset{[\text{eq: } 9.64]}{=} -(\alpha\cdot\beta)
\end{align*}$$

$(\alpha,\beta)\in\mathbb{R}^-\times\mathbb{R}^-$. Then

$$\begin{align*}
-(\alpha\cdot\beta) &\underset{\alpha\in\mathbb{R}^-\wedge\beta\in\mathbb{R}^-}{=} -((-\alpha)\odot(-\beta)) \tag{9.65}\\
(-\alpha)\cdot\beta &\underset{-\alpha\in\mathbb{R}^+\wedge\beta\in\mathbb{R}^-}{=} -((-\alpha)\odot(-\beta)) \\
&\underset{[\text{eq: } 9.65]}{=} -(\alpha\cdot\beta) \\
\alpha\cdot(-\beta) &\underset{\alpha\in\mathbb{R}^-\wedge-\beta\in\mathbb{R}^+}{=} -((-\alpha)\odot(-\beta)) \\
&\underset{[\text{eq: } 9.65]}{=} -(\alpha\cdot\beta)
\end{align*}$$

$(\alpha,\beta)\in(\{0\}\times\mathbb{R})\bigcup(\mathbb{R}\times\{0\})$. Then

$$\begin{align*}
-(\alpha\cdot\beta) &\underset{(\alpha,\beta)\in(\{0\}\times\mathbb{R})\bigcup(\mathbb{R}\times\{0\})}{=} 0 \\
&\underset{(-\alpha,\beta)\in(\{0\}\times\mathbb{R})\bigcup(\mathbb{R}\times\{0\})}{=} (-\alpha)\cdot\beta \\
-(\alpha\cdot\beta) &\underset{(\alpha,\beta)\in(\{0\}\times\mathbb{R})\bigcup(\mathbb{R}\times\{0\})}{=} 0 \\
&\underset{(\alpha,-\beta)\in(\{0\}\times\mathbb{R})\bigcup(\mathbb{R}\times\{0\})}{=} \alpha\cdot(-\beta)
\end{align*}$$

□

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

$$\begin{align*}
\alpha\cdot\beta+\alpha\cdot\beta &\underset{\alpha,\beta\in\mathbb{R}^+}{=} \alpha\odot\beta+\alpha\odot\beta \\
&\underset{\alpha,\beta\in\mathbb{R}^+\text{ and }[\text{theorem: } 9.24]}{=} \alpha\odot(\beta+\beta) \\
&= \alpha\odot((\beta+(-\gamma))+(\beta+\gamma)) \\
&\underset{[\text{eq: } 9.67+\text{theorem: } 9.24]}{=} (\alpha\odot(\beta+(-\gamma)))+\alpha\odot(\beta+\gamma) \\
&\underset{[\beta,-\gamma\in\mathbb{R}^++\text{theorem: } 9.24]}{=} \alpha\odot\beta+\alpha\odot(-\gamma)+\alpha\odot(\beta+\gamma) \\
&= \alpha\odot\beta+(-(-(\alpha\odot(-\gamma))))+\alpha\odot(\beta+\gamma) \\
&\underset{\alpha,\beta\in\mathbb{R}^+,\gamma\in\mathbb{R}^-,\beta+\gamma\in\mathbb{R}^+}{=} \alpha\cdot\beta+(-(\alpha\cdot\gamma))+\alpha\cdot(\beta+\gamma)
\end{align*}$$

so after adding $-(\alpha\cdot\beta)+\alpha\cdot\gamma$ to both sides gives

$$\begin{align*}
\alpha\cdot(\beta+\gamma) &= \alpha\cdot\beta+\alpha\cdot\beta+(-(\alpha\cdot\beta))+\alpha\cdot\gamma \\
&= \alpha\cdot\beta+\alpha\cdot\gamma
\end{align*}$$

□

For $\beta+\gamma\in\mathbb{R}$ we have three cases to consider:

$\boldsymbol{\beta+\gamma\in\mathbb{R}^+}$**.** Then $\alpha\cdot(\beta+\gamma)\underset{\text{[eq: 9.66]}}{=}\alpha\cdot\beta+\alpha\cdot\gamma$

$\boldsymbol{\beta+\gamma\in\mathbb{R}^-}$**.** 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

$$\begin{aligned}\alpha\cdot(\beta+\gamma) &= \alpha\cdot(-(-(\beta+\gamma)))\\ &\underset{\text{[theorem: 9.29]}}{=} -(\alpha\cdot(-(\beta+\gamma)))\\ &= -(\alpha\cdot((-\gamma)+(-\beta)))\\ &\underset{\text{[eq: 9.68]}}{=} -(\alpha\cdot(-\gamma)+\alpha\cdot(-\beta))\\ &= -(\alpha\cdot(-\beta)+\alpha\cdot(-\gamma))\\ &\underset{\text{[theorem: 4.8]}}{=} (-(\alpha\cdot(-\beta)))+(-(\alpha\cdot(-\gamma)))\\ &\underset{\text{[theorem: 9.29]}}{=} \alpha\cdot(-(-\beta))+\alpha\cdot(-(-\gamma))\\ &= \alpha\cdot\beta+\alpha\cdot\gamma\end{aligned}$$

$\boldsymbol{\beta+\gamma=0}$**.** Then $\gamma=-\beta$ and we have

$$\begin{aligned}\alpha\cdot(\beta+\gamma) &= \alpha\cdot 0\\ &= 0\\ &= \alpha\cdot\beta+(-(\alpha\cdot\beta))\\ &\underset{\text{[theorem: 9.29]}}{=} \alpha\cdot\beta+\alpha\cdot(-\beta)\\ &= \alpha\cdot\beta+\alpha\cdot\gamma\end{aligned}$$

□

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* $\alpha^{-1}=\begin{cases}\operatorname{inv}(\alpha) \text{ if } \alpha\in\mathbb{R}^+\\ -\operatorname{inv}(-\alpha) \text{ if } \alpha\in\mathbb{R}^-\end{cases}$

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. $+:\mathbb{R}\times\mathbb{R}\rightarrow\mathbb{R}$*is defined in [theorem: 9.15]*
2. $\cdot:\mathbb{R}\times\mathbb{R}\rightarrow\mathbb{R}$ *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 $\cdot:\mathbb{R}\times\mathbb{R}\rightarrow\mathbb{R}$ 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

$$\begin{aligned}\alpha\cdot\beta &\underset{\alpha,\beta\in\mathbb{R}^+}{=} \alpha\odot\beta\\ &\underset{[\text{theorem: }9.22]}{=} \beta\odot\alpha\\ &= \beta\cdot\alpha\end{aligned}$$

$(\alpha.\beta)\in\mathbb{R}^+\times\mathbb{R}^-$. Then

$$\begin{aligned}\alpha\cdot\beta &\underset{\alpha\in\mathbb{R}^+\wedge\beta\in\mathbb{R}^-}{=} -(\alpha\odot(-\beta))\\ &\underset{[\text{theorem: }9.22]}{=} -((-\beta)\odot\alpha)\\ &\underset{\alpha\in\mathbb{R}^+\wedge-\beta\in\mathbb{R}^+}{=} -((-\beta)\cdot\alpha)\\ &\underset{[\text{theorem: }9.29]}{=} -(-(\beta\cdot\alpha))\\ &= \beta\cdot\alpha\end{aligned}$$

$(\alpha,\beta)\in\mathbb{R}^-\times\mathbb{R}^+$. Then

$$\begin{aligned}\alpha\cdot\beta &\underset{\alpha\in\mathbb{R}^-\wedge\beta\in\mathbb{R}^+}{=} -((-\alpha)\odot\beta)\\ &\underset{[\text{theorem: }9.22]}{=} -(\beta\odot(-\alpha))\\ &\underset{-\alpha\in\mathbb{R}^+\wedge\beta\in\mathbb{R}^+}{=} -(\beta\cdot(-\alpha))\\ &\underset{[\text{theorem: }9.29]}{=} -(-(\beta\cdot\alpha))\\ &= \beta\cdot\alpha\end{aligned}$$

$(\alpha,\beta)\in\mathbb{R}^-\times\mathbb{R}^-$. Then

$$\begin{aligned}\alpha\cdot\beta &\underset{a,\beta\in\mathbb{R}^-}{=} (-\alpha)\odot(-\beta)\\ &\underset{[\text{theorem: }9.22]}{=} (-\beta)\odot(-\alpha)\\ &\underset{a,\beta\in\mathbb{R}^-}{=} \beta\cdot\alpha\end{aligned}$$

$(\alpha,\beta)\in(\{0\}\times\mathbb{R})\bigcup(\mathbb{R}\times\{0\})$. Then

$$\begin{aligned}\alpha\cdot\beta &= 0\\ &= \beta\cdot\alpha\end{aligned}$$

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

$$\begin{aligned}\alpha\cdot\alpha_1 &\underset{\text{commutativity}}{=} \alpha_1\cdot\alpha\\ &\underset{\alpha_1,\alpha\in\mathbb{R}^+}{=} \alpha_1\odot\alpha\\ &\underset{[\text{theorem: }9.21]}{=} \alpha\end{aligned}$$

$\alpha\in\mathbb{R}^-$. Then we have

$$\begin{aligned}\alpha\cdot\alpha_1 &\underset{\text{commutativity}}{=} \alpha_1\cdot\alpha\\ &\underset{\alpha_1\in\mathbb{R}^+\wedge\alpha\in\mathbb{R}^-}{=} -(\alpha_1\odot(-\alpha))\\ &\underset{[\text{theorem: }9.21]}{=} -(-\alpha)\\ &= \alpha\end{aligned}$$

$\boldsymbol{\alpha = 0}$**.** Then we have

$$\begin{array}{rcl}\alpha\cdot\alpha_1 & = & \alpha_1\cdot\alpha\\ & = & 0\\ & = & \alpha\end{array}$$

inverse element. Let $\alpha\in\mathbb{R}\setminus\{0\}$ then by [theorem: 9.17] we have to consider:

$\boldsymbol{\alpha\in\mathbb{R}^+}$**.** Then $\operatorname{inv}(\alpha)\in\mathbb{R}^+$ [see theorem: 9.25] and

$$\begin{array}{rcl}\alpha^{-1}\cdot\alpha & \underset{\text{commutativity}}{=} & \alpha\cdot\alpha^{-1}\\ & \underset{\alpha\in\mathbb{R}^+}{=} & \alpha\cdot\operatorname{inv}(\alpha)\\ & \underset{\operatorname{inv}(\alpha)\in\mathbb{R}^+\wedge\alpha\in\mathbb{R}^+}{=} & \alpha\odot\operatorname{inv}(\alpha)\\ & \underset{[\text{theorem: } 9.26]}{=} & \alpha_1\end{array}$$

$\boldsymbol{\alpha\in\mathbb{R}^-}$**.** Then $-\alpha\in\mathbb{R}^+$ and by [theorem: 9.17] $\operatorname{inv}(-\alpha)\in\mathbb{R}^+$, further

$$\begin{array}{rcl}\alpha^{-1}\cdot\alpha & \underset{\text{commutativity}}{=} & \alpha\cdot\alpha^{-1}\\ & \underset{\alpha\in\mathbb{R}^-}{=} & \alpha\cdot(-(\operatorname{inv}(-\alpha)))\\ & \underset{[\text{theorem: } 9.29]}{=} & -(\alpha\cdot\operatorname{inv}(-\alpha))\\ & \underset{[\text{theorem: } 9.29]}{=} & (-\alpha)\cdot\operatorname{inv}(-\alpha)\\ & \underset{[\text{theorem:: } 9.26]}{=} & \alpha_1\end{array}$$

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:

$\boldsymbol{\alpha\in\mathbb{R}^+\wedge\beta\in\mathbb{R}^+\wedge\gamma\in\mathbb{R}^+}$**.** Then we have

$$\begin{array}{rcl}\alpha\cdot(\beta\cdot\gamma) & = & \alpha\odot(\beta\odot\gamma)\\ & \underset{[\text{theorem: } 9.23]}{=} & (\alpha\odot\beta)\odot\gamma\\ & = & (\alpha\cdot\beta)\cdot\gamma\end{array}$$

$\boldsymbol{\alpha\in\mathbb{R}^+\wedge\beta\in\mathbb{R}^+\wedge\gamma\in\mathbb{R}^-}$**.** Then we have

$$\begin{array}{rcl}\alpha\cdot(\beta\cdot\gamma) & = & \alpha\cdot(-(-(\beta\cdot\gamma)))\\ & \underset{[\text{theorem: } 9.29]}{=} & \alpha\cdot(-(\beta\cdot(-\gamma)))\\ & \underset{[\text{theorem: } 9.29]}{=} & -(\alpha\cdot(\beta\cdot(-\gamma)))\\ & = & -(\alpha\odot(\beta\odot(-\gamma)))\\ & \underset{[\text{theorem: } 9.23]}{=} & -((\alpha\odot\beta)\odot(-\gamma))\\ & = & -((\alpha\cdot\beta)\cdot(-\gamma))\\ & \underset{[\text{theorem: } 9.29]}{=} & (\alpha\cdot\beta)\cdot(-(-\gamma))\\ & = & (\alpha\cdot\beta)\cdot\gamma\end{array}$$

$\boldsymbol{\alpha\in\mathbb{R}^+\wedge\beta\in\mathbb{R}^+\wedge\gamma=0}$**.** Then we have

$$\begin{array}{rcl}\alpha\cdot(\beta\cdot\gamma) & = & \alpha\cdot 0\\ & = & 0\\ & = & (\alpha\cdot\beta)\cdot 0\\ & = & (\alpha\cdot\beta)\cdot\gamma\end{array}$$

$\boldsymbol{\alpha\in\mathbb{R}^{+}\wedge\beta\in\mathbb{R}^{-}\wedge\gamma\in\mathbb{R}^{+}}$**.** Then we have

$$\begin{aligned}
\alpha\cdot(\beta\cdot\gamma) &= \alpha\cdot((-(-\beta))\cdot\gamma)\\
&\underset{[\text{theorem: }9.29]}{=} \alpha\cdot(-((-\beta)\cdot\gamma))\\
&\underset{[\text{theorem: }9.29]}{=} -(\alpha\cdot((-\beta)\cdot\gamma))\\
&= -(\alpha\odot((-\beta)\odot\gamma))\\
&\underset{[\text{theorem: }9.23]}{=} -((\alpha\odot(-\beta))\odot\gamma)\\
&= -((\alpha\cdot(-\beta))\cdot\gamma)\\
&\underset{[\text{theorem: }9.29]}{=} (-(\alpha\cdot(-\beta)))\cdot\gamma\\
&\underset{[\text{theorem: }9.29]}{=} (\alpha\cdot(-(-\beta)))\cdot\gamma\\
&= (\alpha\cdot\beta)\cdot\gamma
\end{aligned}$$

$\boldsymbol{\alpha\in\mathbb{R}^{+}\wedge\beta\in\mathbb{R}^{-}\wedge\gamma\in\mathbb{R}^{-}}$**.** Then we have

$$\begin{aligned}
\alpha\cdot(b\cdot\gamma) &= \alpha\cdot((-(-\beta))\cdot(-(-\gamma)))\\
&\underset{[\text{theorem: }9.29]}{=} \alpha\cdot(-((-\beta)\cdot(-(-\gamma))))\\
&\underset{[\text{theorem: }9.29]}{=} \alpha\cdot(-(-((-\beta)\cdot(-\gamma))))\\
&= \alpha\cdot((-\beta)\cdot(-\gamma))\\
&= \alpha\odot((-\beta)\odot(-\gamma))\\
&\underset{[\text{theorem: }9.23]}{=} (\alpha\odot(-\beta))\odot(-\gamma)\\
&= (\alpha\cdot(-\beta))\cdot(-\gamma)\\
&\underset{[\text{theorem: }9.29]}{=} -((\alpha\cdot(-\beta))\cdot\gamma)\\
&= -((-(\alpha\cdot\beta))\cdot\gamma)\\
&\underset{[\text{theorem: }9.29]}{=} -(-((a\cdot\beta)\cdot\gamma))\\
&= (\alpha\cdot\beta)\cdot\gamma
\end{aligned}$$

$\boldsymbol{\alpha\in\mathbb{R}^{+}\wedge\beta\in\mathbb{R}^{-}\wedge\gamma=0}$**.** 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

$$\begin{aligned}
\alpha \cdot (\beta \cdot \gamma) &= \alpha \cdot (0 \cdot \gamma)\\
&= \alpha \cdot 0\\
&= 0\\
&= 0 \cdot \gamma\\
&= (\alpha \cdot 0) \cdot \gamma\\
&= (\alpha \cdot \beta) \cdot \gamma
\end{aligned}$$

$\boldsymbol{\alpha \in \mathbb{R}^{-} \wedge \beta \in \mathbb{R}^{+} \wedge \gamma \in \mathbb{R}^{+}.}$ Then we have

$$\begin{aligned}
\alpha \cdot (\beta \cdot \gamma) &= (-(-\alpha)) \cdot (\beta \cdot \gamma)\\
&\underset{[\text{theorem: } 9.29]}{=} -((-\alpha) \cdot (\beta \cdot \gamma))\\
&= -((-\alpha) \odot (\beta \odot \gamma))\\
&\underset{[\text{theorem: } 9.23]}{=} -(((-\alpha) \odot \beta) \odot \gamma)\\
&= -(((-\alpha) \cdot \beta) \cdot \gamma)\\
&\underset{[\text{theorem: } 9.29]}{=} -((-(\alpha \cdot \beta)) \cdot \gamma)\\
&\underset{[\text{theorem: } 9.29]}{=} -(-((\alpha \cdot \beta) \cdot \gamma))\\
&= (\alpha \cdot \beta) \cdot \gamma
\end{aligned}$$

$\boldsymbol{\alpha \in \mathbb{R}^{-} \wedge \beta \in \mathbb{R}^{+} \wedge \gamma \in \mathbb{R}^{-}.}$ Then we have

$$\begin{aligned}
\alpha \cdot (\beta \cdot \gamma) &= (-(-\alpha)) \cdot (\beta \cdot (-(-\gamma)))\\
&\underset{[\text{theorem: } 9.29]}{=} -((-\alpha) \cdot (\beta \cdot (-(-\gamma))))\\
&\underset{[\text{theorem: } 9.29]}{=} -((-\alpha) \cdot (-(\beta \cdot (-\gamma))))\\
&\underset{[\text{theorem: } 9.29]}{=} -(-((-\alpha) \cdot (\beta \cdot (-\gamma))))\\
&= (-\alpha) \cdot (\beta \cdot (-\gamma))\\
&= (-\alpha) \odot (\beta \odot (-\gamma))\\
&\underset{[\text{theorem: } 9.23]}{=} ((-\alpha) \odot \beta) \odot (-\gamma)\\
&= ((-\alpha) \cdot \beta) \cdot (-\gamma)\\
&\underset{[\text{theorem: } 9.29]}{=} -(((-\alpha) \cdot \beta) \cdot \gamma)\\
&\underset{[\text{theorem: } 9.29]}{=} -((-(\alpha \cdot \beta)) \cdot \gamma)\\
&\underset{[\text{theorem: } 9.29]}{=} -(-((\alpha \cdot \beta) \cdot \gamma))\\
&= (\alpha \cdot \beta) \cdot \gamma
\end{aligned}$$

$\boldsymbol{\alpha \in \mathbb{R}^{-} \wedge \beta \in \mathbb{R}^{+} \wedge \gamma = 0.}$ Then we have

$$\begin{aligned}
\alpha \cdot (\beta \cdot \gamma) &= \alpha \cdot (\beta \cdot 0)\\
&= \alpha \cdot 0\\
&= 0\\
&= (\alpha \cdot \beta) \cdot 0\\
&= (\alpha \cdot \beta) \cdot \gamma
\end{aligned}$$

$\boldsymbol{\alpha \in \mathbb{R}^{-} \wedge \beta \in \mathbb{R}^{-} \wedge \gamma \in \mathbb{R}^{+}.}$ Then we have

$$\begin{aligned}
\alpha \cdot (\beta \cdot \gamma) &= (-(-\alpha)) \cdot ((-(-\beta)) \cdot \gamma)\\
&\underset{[\text{theorem: } 9.29]}{=} -((-\alpha) \cdot ((-(-\beta)) \cdot \gamma))\\
&\underset{[\text{theorem: } 9.29]}{=} -((-\alpha) \cdot (-((-\beta) \cdot \gamma)))
\end{aligned}$$

$$\begin{array}{rcl} & \underset{\text{[theorem: 9.29]}}{=} & -(-((-\alpha)\cdot((-\beta)\cdot\gamma))) \\ & = & (-\alpha)\cdot((-\beta)\cdot\gamma) \\ & = & (-\alpha)\odot((-\beta)\odot\gamma) \\ & \underset{\text{[theorem: 9.23]}}{=} & ((-\alpha)\odot(-\beta))\odot\gamma \\ & = & (\alpha\cdot\beta)\cdot\gamma \end{array}$$

$\boldsymbol{\alpha\in\mathbb{R}^-\wedge\beta\in\mathbb{R}^-\wedge\gamma\in\mathbb{R}^-}$**.** Then we have

$$\begin{array}{rcl} \alpha\cdot(\beta\cdot\gamma) & = & \alpha\cdot((-\beta)\odot(-\gamma)) \\ & = & -((-\alpha)\odot((-\beta)\odot(-\gamma))) \\ & \underset{\text{[theorem: 9.23]}}{=} & -(((-\alpha)\odot(-\beta))\odot(-\gamma)) \\ & = & -((\alpha\cdot\beta)\cdot(-\gamma)) \\ & \underset{\text{[theorem: 9.29]}}{=} & -(-((\alpha\cdot\beta)\cdot\gamma)) \\ & = & (\alpha\cdot\beta)\cdot\gamma \end{array}$$

$\boldsymbol{\alpha\in\mathbb{R}^-\wedge\beta\in\mathbb{R}^-\wedge\gamma=0}$**.** 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

$$\begin{array}{rcl} \alpha\cdot(\beta\cdot\gamma) & = & \alpha\cdot(0\cdot\gamma) \\ & = & \alpha\cdot 0 \\ & = & 0 \\ & = & 0\cdot\gamma \\ & = & (\alpha\cdot 0)\cdot\gamma \\ & = & (\alpha\cdot\beta)\cdot\gamma \end{array}$$

$\boldsymbol{\alpha = 0 \wedge \beta \in \mathbb{R}^+ \wedge \gamma \in \mathbb{R}^+}$**.** 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

$$\begin{aligned}\alpha\cdot(\beta\cdot\gamma) &= 0\cdot(\beta\cdot\gamma)\\ &= 0\\ &= 0\cdot\gamma\\ &= (0\cdot\beta)\cdot\gamma\\ &= (\alpha\cdot\beta)\cdot\gamma\end{aligned}$$

$\boldsymbol{\alpha = 0 \wedge \beta \in \mathbb{R}^- \wedge \gamma \in \mathbb{R}^+}$**.** Then we have

$$\begin{aligned}\alpha\cdot(\beta\cdot\gamma) &= 0\cdot(\beta\cdot\gamma)\\ &= 0\\ &= 0\cdot\gamma\\ &= (0\cdot\beta)\cdot\gamma\\ &= (\alpha\cdot\beta)\cdot\gamma\end{aligned}$$

$\boldsymbol{\alpha = 0 \wedge \beta \in \mathbb{R}^- \wedge \gamma \in \mathbb{R}^-}$**.** 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

$$\begin{aligned}\alpha\cdot(\beta\cdot\gamma) &= 0\cdot(\beta\cdot\gamma)\\ &= 0\\ &= 0\cdot\gamma\\ &= (0\cdot\beta)\cdot\gamma\\ &= (\alpha\cdot\beta)\cdot\gamma\end{aligned}$$

$\boldsymbol{\alpha = 0 \wedge \beta = 0 \wedge \gamma \in \mathbb{R}^+}$**.** Then we have

$$\begin{aligned}\alpha\cdot(\beta\cdot\gamma) &= 0\cdot(\beta\cdot\gamma)\\ &= 0\\ &= 0\cdot\gamma\\ &= (0\cdot\beta)\cdot\gamma\\ &= (\alpha\cdot\beta)\cdot\gamma\end{aligned}$$

$\boldsymbol{\alpha=0\wedge\beta=0\wedge\gamma\in\mathbb{R}^-}$**.** Then we have

$$\begin{aligned}\alpha\cdot(\beta\cdot\gamma) &= 0\cdot(\beta\cdot\gamma)\\ &= 0\\ &= 0\cdot\gamma\\ &= (0\cdot\beta)\cdot\gamma\\ &= (\alpha\cdot\beta)\cdot\gamma\end{aligned}$$

$\boldsymbol{\alpha=0\wedge\beta=0\wedge\gamma=0}$**.** Then we have

$$\begin{aligned}\alpha\cdot(\beta\cdot\gamma) &= 0\cdot(\beta\cdot\gamma)\\ &= 0\\ &= 0\cdot\gamma\\ &= (0\cdot\beta)\cdot\gamma\\ &= (\alpha\cdot\beta)\cdot\gamma\end{aligned}$$

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:

$\boldsymbol{\alpha\in\mathbb{R}^+\wedge\beta\in\mathbb{R}^+\wedge\gamma\in\mathbb{R}^+}$**.** Then

$$\begin{aligned}\alpha\cdot(\beta+\gamma) &= \alpha\odot(\beta+\gamma)\\ &\underset{\text{[theorem: 9.24]}}{=} \alpha\odot\beta+\alpha\odot\gamma\\ &= \alpha\cdot\beta+\alpha\cdot\gamma\end{aligned}$$

$\boldsymbol{\alpha\in\mathbb{R}^+\wedge\beta\in\mathbb{R}^+\wedge\gamma\in\mathbb{R}^-}$**.** Then

$$\alpha\cdot(b+\gamma)\underset{\text{[lemma: 9.30]}}{=}\alpha\cdot\beta+\alpha\cdot\gamma$$

$\boldsymbol{\alpha\in\mathbb{R}^+\wedge\beta\in\mathbb{R}^+\wedge\gamma=0}$**.** Then

$$\begin{aligned}\alpha\cdot(\beta+\gamma) &= \alpha\cdot(\beta+0)\\ &= \alpha\cdot\beta\\ &= \alpha\cdot\beta+0\\ &= \alpha\cdot\beta+\alpha\cdot 0\\ &= \alpha\cdot\beta+\alpha\cdot\gamma\end{aligned}$$

$\boldsymbol{\alpha\in\mathbb{R}^+\wedge\beta\in\mathbb{R}^-\wedge\gamma\in\mathbb{R}^+}$**.** Then

$$\begin{aligned}\alpha\cdot(\beta+\gamma) &= \alpha\cdot(\gamma+\beta)\\ &\underset{\text{[lemma: 9.30]}}{=} \alpha\cdot\gamma+\alpha\cdot\beta\\ &= \alpha\cdot\beta+\alpha\cdot\gamma\end{aligned}$$

$\boldsymbol{\alpha\in\mathbb{R}^+\wedge\beta\in\mathbb{R}^-\wedge\gamma\in\mathbb{R}^-}$**.** Then

$$\begin{aligned}\alpha\cdot(\beta+\gamma) &= \alpha\cdot(-(-(\beta+\gamma)))\\ &\underset{\text{[theorem: 9.29]}}{=} -(\alpha\cdot(-(\beta+\gamma)))\\ &\underset{\text{[theorem: 4.8]}}{=} -(\alpha\cdot((-\beta)+(-\gamma)))\\ &= -(\alpha\odot((-\beta)+(-\gamma)))\\ &\underset{\text{[theorem: 9.24]}}{=} -(\alpha\odot(-\beta)+\alpha\odot(-\gamma))\\ &= -(\alpha\cdot(-\beta)+\alpha\cdot(-\gamma))\\ &\underset{\text{[theorem: 4.8]}}{=} (-(\alpha\cdot(-\beta)))+(-(\alpha\cdot(-\gamma)))\\ &\underset{\text{[theorem: 9.29]}}{=} \alpha\cdot(-(-\beta))+\alpha\cdot(-(-\gamma))\\ &= \alpha\cdot\beta+\alpha\cdot\gamma\end{aligned}$$

$\boldsymbol{\alpha \in \mathbb{R}^+ \wedge \beta \in \mathbb{R}^- \wedge \gamma = 0.}$ Then

$$\begin{aligned}\alpha\cdot(\beta+\gamma) &= \alpha\cdot(\beta+0)\\ &= \alpha\cdot\beta\\ &= \alpha\cdot\beta+0\\ &= \alpha\cdot\beta+\alpha\cdot 0\\ &= \alpha\cdot\beta+\alpha\cdot\gamma\end{aligned}$$

$\boldsymbol{\alpha \in \mathbb{R}^+ \wedge \beta = 0 \wedge \gamma \in \mathbb{R}^+.}$ Then

$$\begin{aligned}\alpha\cdot(\beta+\gamma) &= \alpha\cdot(0+\gamma)\\ &= \alpha\cdot\gamma\\ &= 0+\alpha\cdot\gamma\\ &= \alpha\cdot\beta+\alpha\cdot\gamma\end{aligned}$$

$\boldsymbol{\alpha \in \mathbb{R}^+ \wedge \beta = 0 \wedge \gamma \in \mathbb{R}^-.}$ Then

$$\begin{aligned}\alpha\cdot(\beta+\gamma) &= \alpha\cdot(0+\gamma)\\ &= \alpha\cdot\gamma\\ &= 0+\alpha\cdot\gamma\\ &= \alpha\cdot\beta+\alpha\cdot\gamma\end{aligned}$$

$\boldsymbol{\alpha \in \mathbb{R}^+ \wedge \beta = 0 \wedge \gamma = 0.}$ Then

$$\begin{aligned}\alpha\cdot(\beta+\gamma) &= \alpha\cdot(0+\gamma)\\ &= \alpha\cdot\gamma\\ &= 0+\alpha\cdot\gamma\\ &= \alpha\cdot\beta+\alpha\cdot\gamma\end{aligned}$$

$\boldsymbol{\alpha \in \mathbb{R}^- \wedge \beta \in \mathbb{R}^+ \wedge \gamma \in \mathbb{R}^+.}$ Then

$$\begin{aligned}\alpha\cdot(\beta+\gamma) &= -((-\alpha)\odot(\beta+\gamma))\\ &\underset{[\text{theorem: } 9.24]}{=} -((-\alpha)\odot\beta+(-\alpha)\odot\gamma)\\ &\underset{[\text{theorem: } 4.8]}{=} (-((-\alpha)\odot\beta))+(-((-\alpha)\odot\gamma))\\ &= (-((-\alpha)\cdot\beta))+(-((-\alpha)\cdot\gamma))\\ &\underset{[\text{theorem: } 9.29]}{=} (-(-\alpha))\cdot\beta+(-(-\alpha))\cdot\gamma\\ &= \alpha\cdot\beta+\alpha\cdot\gamma\end{aligned}$$

$\boldsymbol{\alpha \in \mathbb{R}^- \wedge \beta \in \mathbb{R}^+ \wedge \gamma \in \mathbb{R}^-.}$ Then

$$\begin{aligned}\alpha\cdot(\beta+\gamma) &= (-(-\alpha))\cdot(\beta+\gamma)\\ &\underset{[\text{theorem: } 9.29]}{=} -((-\alpha)\cdot(\beta+\gamma))\\ &\underset{[\text{lemma: } 9.30]}{=} -((-\alpha)\cdot\beta+(-\alpha)\cdot\gamma)\\ &\underset{[\text{theorem: } 4.8]}{=} (-((-\alpha)\cdot\beta))+(-((-\alpha)\cdot\gamma))\\ &\underset{[\text{theorem: } 9.29]}{=} (-(-\alpha))\cdot\beta+(-(-\alpha))\cdot\gamma\\ &= \alpha\cdot\beta+\alpha\cdot\gamma\end{aligned}$$

$\boldsymbol{\alpha \in \mathbb{R}^- \wedge \beta \in \mathbb{R}^+ \wedge \gamma = 0.}$ Then

$$\begin{aligned}\alpha\cdot(\beta+\gamma) &= \alpha\cdot(\beta+0)\\ &= \alpha\cdot\beta\\ &= \alpha\cdot\beta+0\\ &= \alpha\cdot\beta+\alpha\cdot 0\\ &= \alpha\cdot\beta+\alpha\cdot\gamma\end{aligned}$$

$\boldsymbol{\alpha}\in\mathbb{R}^-\wedge\boldsymbol{\beta}\in\mathbb{R}^-\wedge\boldsymbol{\gamma}\in\mathbb{R}^+$**.** Then

$$\begin{aligned}
\alpha\cdot(\beta+\gamma) &= \alpha\cdot(\gamma+\beta)\\
&= (-(-\alpha))\cdot(\gamma+\beta)\\
&\underset{\text{[theorem: 9.29]}}{=} -((-\alpha)\cdot(\gamma+\beta))\\
&\underset{\text{[lemma: 9.30]}}{=} -((-\alpha)\cdot\gamma+(-\alpha)\cdot\beta)\\
&\underset{\text{[theorem: 4.8]}}{=} (-((-\alpha)\cdot\gamma))+(-((-\alpha)\cdot\beta))\\
&\underset{\text{[theorem: 9.29]}}{=} (-(-\alpha))\cdot\gamma+(-(-\alpha))\cdot\beta\\
&= \alpha\cdot\gamma+\alpha\cdot\beta\\
&= \alpha\cdot\beta+\alpha\cdot\gamma
\end{aligned}$$

$\boldsymbol{\alpha}\in\mathbb{R}^-\wedge\boldsymbol{\beta}\in\mathbb{R}^-\wedge\boldsymbol{\gamma}\in\mathbb{R}^-$**.** Then

$$\begin{aligned}
\alpha\cdot(\beta+\gamma) &= (-(-\alpha))\cdot(\beta+\gamma)\\
&\underset{\text{[theorem: 9.29]}}{=} -((-\alpha)\cdot(\beta+\gamma))\\
&\underset{\text{[theorem: 9.29]}}{=} (-\alpha)\cdot(-(\beta+\gamma))\\
&\underset{\text{[theorem: 4.8]}}{=} (-\alpha)\cdot((-\beta)+(-\gamma))\\
&= (-\alpha)\odot((-\beta)+(-\gamma))\\
&\underset{\text{[theorem: 9.24]}}{=} (-\alpha)\odot(-\beta)+(-\alpha)\odot(-\gamma)\\
&= (-\alpha)\cdot(-\beta)+(-\alpha)\cdot(-\gamma)\\
&\underset{\text{[theorem: 9.29]}}{=} (-(\alpha\cdot(-\beta)))+(-(\alpha\cdot(-\gamma)))\\
&\underset{\text{[theorem: 9.29]}}{=} (-(-(\alpha\cdot\beta)))+(-(-(\alpha\cdot\gamma)))\\
&= \alpha\cdot\beta+\alpha\cdot\gamma
\end{aligned}$$

$\boldsymbol{\alpha}\in\mathbb{R}^-\wedge\boldsymbol{\beta}\in\mathbb{R}^-\wedge\boldsymbol{\gamma}=\mathbf{0}$**.** Then

$$\begin{aligned}
\alpha\cdot(\beta+\gamma) &= \alpha\cdot(\beta+0)\\
&= \alpha\cdot\beta\\
&= \alpha\cdot\beta+0\\
&= \alpha\cdot\beta+\alpha\cdot 0\\
&= \alpha\cdot\beta+\alpha\cdot\gamma
\end{aligned}$$

$\boldsymbol{\alpha}\in\mathbb{R}^-\wedge\boldsymbol{\beta}=\mathbf{0}\wedge\boldsymbol{\gamma}\in\mathbb{R}^+$**.** Then

$$\begin{aligned}
\alpha\cdot(\beta+\gamma) &= \alpha\cdot(0+\gamma)\\
&= \alpha\cdot\gamma\\
&= 0+\alpha\cdot\gamma\\
&= \alpha\cdot 0+\alpha\cdot\gamma\\
&= \alpha\cdot\beta+\alpha\cdot\gamma
\end{aligned}$$

$\boldsymbol{\alpha}\in\mathbb{R}^-\wedge\boldsymbol{\beta}=\mathbf{0}\wedge\boldsymbol{\gamma}\in\mathbb{R}^-$**.** Then

$$\begin{aligned}
\alpha\cdot(\beta+\gamma) &= \alpha\cdot(0+\gamma)\\
&= \alpha\cdot\gamma\\
&= 0+\alpha\cdot\gamma\\
&= \alpha\cdot 0+\alpha\cdot\gamma\\
&= \alpha\cdot\beta+\alpha\cdot\gamma
\end{aligned}$$

$\boldsymbol{\alpha \in \mathbb{R}^- \wedge \beta = 0 \wedge \gamma = 0.}$ Then

$$\begin{aligned}\alpha\cdot(\beta+\gamma) &= \alpha\cdot(0+\gamma)\\ &= \alpha\cdot\gamma\\ &= 0+\alpha\cdot\gamma\\ &= \alpha\cdot 0+\alpha\cdot\gamma\\ &= \alpha\cdot\beta+\alpha\cdot\gamma\end{aligned}$$

$\boldsymbol{\alpha = 0 \wedge \beta \in \mathbb{R}^+ \wedge \gamma \in \mathbb{R}^+.}$ 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

$$\begin{aligned}\alpha\cdot(\beta+\gamma) &= 0\cdot(\beta+\gamma)\\ &= 0\\ &= 0+0\\ &= 0\cdot\beta+0\cdot\gamma\\ &= \alpha\cdot\beta+\alpha\cdot\gamma\end{aligned}$$

$\boldsymbol{\alpha = 0 \wedge \beta \in \mathbb{R}^- \wedge \gamma \in \mathbb{R}^+.}$ 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

$$\begin{aligned}\alpha\cdot(\beta+\gamma) &= 0\cdot(\beta+\gamma)\\ &= 0\\ &= 0+0\\ &= 0\cdot\beta+0\cdot\gamma\\ &= \alpha\cdot\beta+\alpha\cdot\gamma\end{aligned}$$

$\alpha = 0 \wedge \beta = 0 \wedge \gamma \in \mathbb{R}^+$. Then

$$\begin{array}{rcl}\alpha \cdot (\beta + \gamma) & = & 0 \cdot (\beta + \gamma)\\ & = & 0\\ & = & 0 + 0\\ & = & 0 \cdot \beta + 0 \cdot \gamma\\ & = & \alpha \cdot \beta + \alpha \cdot \gamma\end{array}$$

$\alpha = 0 \wedge \beta = 0 \wedge \gamma \in \mathbb{R}^-$. Then

$$\begin{array}{rcl}\alpha \cdot (\beta + \gamma) & = & 0 \cdot (\beta + \gamma)\\ & = & 0\\ & = & 0 + 0\\ & = & 0 \cdot \beta + 0 \cdot \gamma\\ & = & \alpha \cdot \beta + \alpha \cdot \gamma\end{array}$$

$\alpha = 0 \wedge \beta = 0 \wedge \gamma = 0$. 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* $\mathrm{inv}(\alpha_r) = \alpha_{r^{-1}}$

Proof. As $\alpha_r \in \mathbb{R}^+$ we have $0 \in \alpha_r$ hence $0 < r \Rightarrow 0 < r^{-1}$. Let $x \in \mathrm{inv}(\alpha_r)$ then we have for x the following to consider:

$x \leqslant 0$. Then $x \leqslant 0 < r^{-1}$ proving that $x \in \alpha_{r^{-1}}$.

$0 < x$. 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

$$\mathrm{inv}(\alpha_r) \subseteq \alpha_{r^{-1}} \tag{9.69}$$

Let $x\in\alpha_{r^{-1}}$ then for x we have either:

$x\leqslant 0$. Then $x\in\mathbb{Q}_0^-$ so that $x\in\operatorname{inv}(\alpha_r)$.

$0<x$. 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:

$x\leqslant 0$. Then as $0<r\wedge 0<s$ we have that $0<r\cdot s$ so that $x<r\cdot s$ proving that $x\in\alpha_{r\cdot s}$

$0<x$. 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:

$\boldsymbol{\alpha_r \in \mathbb{R}^+ \wedge \alpha_s \in \mathbb{R}^+}$**.** 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

$$\begin{aligned}\alpha_r \cdot \alpha_s &= -(\alpha_r \odot (-\alpha_s)) \\ &\underset{(2)}{=} -(\alpha_r \odot \alpha_{-s}) \\ &\underset{[\text{lemma: } 9.35]}{=} -\alpha_{r \cdot (-s)} \\ &\underset{[\text{theorem: } 4.40}{=} -\alpha_{-(r \cdot s)} \\ &\underset{(2)}{=} \alpha_{-(-(r \cdot s))} \\ &= \alpha_{r \cdot s}\end{aligned}$$

$\boldsymbol{\alpha_r \in \mathbb{R}^- \wedge \alpha_s \in \mathbb{R}^+}$**.** Then

$$\begin{aligned}\alpha_r \cdot \alpha_s &= -((-\alpha_r) \odot \alpha_s) \\ &= -(\alpha_{-r} \odot \alpha_s) \\ &\underset{[\text{lemma: } 9.35]}{=} -(\alpha_{(-r) \cdot s}) \\ &\underset{[\text{theorem: } 4.40}{=} -(\alpha_{-(r \cdot s)}) \\ &\underset{(2)}{=} \alpha_{-(-(r \cdot s))} \\ &= \alpha_{r \cdot s}\end{aligned}$$

$\boldsymbol{\alpha_r \in \mathbb{R}^- \wedge \alpha_s \in \mathbb{R}^-}$**.** Then

$$\begin{aligned}\alpha_r \cdot \alpha_s &= (-\alpha_r) \odot (-\alpha_s)\\ &\underset{(2)}{=} \alpha_{-r} \odot \alpha_{-s}\\ &\underset{[\text{lemma: } 9.35]}{=} \alpha_{(-r)\cdot(-s)}\\ &\underset{[\text{theorem: } 4.40]}{=} \alpha_{r\cdot s}\end{aligned}$$

$\boldsymbol{(\alpha_r, \alpha_s) \in (\{0\} \times \mathbb{R}) \bigcup (\mathbb{R} \times \{0\})}$**.** 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

$$\begin{aligned}\alpha_r \cdot \alpha_s &= 0 \cdot \alpha_s\\ &= 0\\ &= \alpha_0\\ &= \alpha_{0\cdot s}\\ &= \alpha_{r\cdot s}\end{aligned}$$

$\boldsymbol{(\alpha_r, \alpha_s) \in \mathbb{R} \times \{0\}}$**.** Then $\alpha_s = 0 = \alpha_0 \Rightarrow_{[\text{theorem: } 9.7]} s = 0$ and

$$\begin{aligned}\alpha_r \cdot \alpha_s &= \alpha_r \cdot 0\\ &= 0\\ &= \alpha_0\\ &= \alpha_{r\cdot 0}\\ &= \alpha_{r\cdot s}\end{aligned}$$

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 $(\alpha_r)^{-1} = \text{inv}(\alpha_r) \underset{[\text{lemma: } 9.34]}{=} \alpha_{r^{-1}}$

$\boldsymbol{\alpha_r \in \mathbb{R}^-}$**.** Then

$$\begin{aligned}(\alpha_r)^{-1} &= -\text{inv}(-\alpha_r)\\ &\underset{(2)}{=} -\text{inv}(\alpha_{-r})\\ &\underset{[\text{lemma: } 9.34]}{=} -\alpha_{(-r)^{-1}}\\ &\underset{[\text{theorems: } 8.30]}{=} -\alpha_{-(r^{-1})}\\ &\underset{(2)}{=} \alpha_{-(-(r^{-1}))}\\ &= \alpha_{r^{-1}}\end{aligned}$$

□

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

$$\begin{aligned} (y+z) + (-(x+z)) &\underset{\text{[theorem: 4.8]}}{=} (y+z) + ((-x) + (-z)) \\ &= (y+z) + ((-z) + (-x)) \\ &= y + (z + ((-z) + (-x))) \\ &= y + ((z + (-z)) + (-x)) \\ &= y + (0 + (-x)) \\ &= y + (-x) \end{aligned}$$

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

$\mathbf{0 < \boldsymbol{\alpha}}$**.** Then $\beta = 0 + \beta < \alpha + \beta = 0$ hence $\beta < 0$ contradicting $0 \leqslant \beta$, so this case does not occur.

$\mathbf{0 < \boldsymbol{\beta}}$**.** Then $\alpha = \alpha + 0 < \alpha + \beta = 0$ hence $\alpha < 0$ contradicting $0 \leqslant \alpha$, so this case does not occur.

$\boldsymbol{\alpha} = \boldsymbol{\beta} = \mathbf{0}$**.** 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.

⇒. 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

$$\alpha_r < \alpha_s$$

⇐. 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.

$$\begin{aligned} r \leqslant s &\Leftrightarrow r = s \vee r < s \\ &\Leftrightarrow \alpha_r = \alpha_s \vee r < s \\ &\underset{(1)}{\Leftrightarrow} \alpha_r = a_s \vee \alpha_r < \alpha_s \\ &\Leftrightarrow \alpha_r \leqslant \alpha_s \end{aligned}$$

□

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: 3.74] 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: 3.75] 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*

$$i_{\mathbb{Z} \rightarrow \mathbb{Q}}: \mathbb{Z} \rightarrow \mathbb{Z}_{\mathbb{Q}} \subseteq \mathbb{Q} \textit{ is defined by } i_{\mathbb{Z} \rightarrow \mathbb{Q}}(z) = \frac{z}{1} \textit{ [theorem: 8.32]}$$

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: 3.51] 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*

$$i_{\mathbb{N}_0}: \mathbb{N}_0 \rightarrow \mathbb{Z}_0^+ \subseteq \mathbb{Z} \text{ is defined by } i_{\mathbb{N}_0}(z) = \sim[(z, 0)] \text{ [theorem: 7.17] and}$$

$$i_{\mathbb{Z}\rightarrow\mathbb{R}}: \mathbb{Z} \rightarrow \mathbb{Z}_{\mathbb{R}} \text{ [theorem: 9.52]}$$

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: 3.51]

   $$i_{\mathbb{N}_0 \rightarrow \mathbb{R}} = i_{\mathbb{Z} \rightarrow \mathbb{R}} \circ i_{\mathbb{N}_0 \rightarrow \mathbb{Z}}: \langle \mathbb{N}_0, \leqslant \rangle \rightarrow \langle i_{\mathbb{Z} \rightarrow \mathbb{R}}(i_{\mathbb{N}_0 \rightarrow \mathbb{Z}}(\mathbb{N}_0)), \leqslant \rangle \text{ is a order isomorphism}$$

Using [eq: 9.84] it follows that

$$i_{\mathbb{N}_0 \rightarrow \mathbb{R}} = i_{\mathbb{Z} \rightarrow \mathbb{R}} \circ i_{\mathbb{N}_0 \rightarrow \mathbb{Z}} \colon \langle \mathbb{N}_0, \leqslant \rangle \rightarrow \langle \mathbb{N}_{0,\mathbb{R}}, \leqslant \rangle \text{ is a order isomorphism}$$

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) $i_{\mathbb{N}_0 \rightarrow \mathbb{R}} \colon \langle \mathbb{N}_0, \leqslant \rangle \rightarrow \langle \mathbb{N}_{0,\mathbb{R}}, \leqslant \rangle$ is a order isomorphism, hence by [theorem: 3.79] 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

$$i_{\mathbb{Z} \rightarrow \mathbb{Q}} \colon \langle \mathbb{Z}, +, \cdot \rangle \rightarrow \langle \mathbb{Z}_{\mathbb{Q}}, +, \cdot \rangle \text{ is a ring isomorphism} \tag{9.85}$$

$$i_{\mathbb{Q} \rightarrow \mathbb{R}} \colon \langle \mathbb{Q}, +, \cdot \rangle \rightarrow \langle \mathbb{Q}_{\mathbb{R}}, +, \cdot \rangle \text{ is a field isomorphism} \tag{9.86}$$

$$i_{\mathbb{Q} \rightarrow \mathbb{R}} \circ i_{\mathbb{Z} \rightarrow \mathbb{Q}} \colon \langle \mathbb{Z}, +, \cdot \rangle \rightarrow \langle \mathbb{Z}_{\mathbb{R}}, +, \cdot \rangle \text{ is a ring isomorphism} \tag{9.87}$$

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

$$\begin{aligned}(x,y)\cdot((x',y')\cdot(x'',y'')) &=\\ (x,y)\cdot(x'\cdot x''-y'\cdot y'',x'\cdot y''+y'\cdot x'') &=\\ (x\cdot(x'\cdot x''-y'\cdot y'')-y\cdot(x'\cdot y''+y'\cdot x''),x\cdot(x'\cdot y''+y'\cdot x''))+y\cdot(x'\cdot x''-y'\cdot y'') &=\\ (x\cdot(x'\cdot x'')-x\cdot(y'\cdot y'')-y\cdot(x'\cdot y'')-y\cdot(y'\cdot x''),x\cdot(x'\cdot y'')+x\cdot(y'\cdot x'')+y\cdot(x'\cdot x'')-\\ y\cdot(y'\cdot y'')) &=\end{aligned}$$

$$\begin{aligned}
&\Big(\underbrace{(x\cdot x')\cdot x''}_{1} - \underbrace{(x\cdot y')\cdot y''}_{2} - \underbrace{(y\cdot x')\cdot y''}_{2} - \underbrace{(y\cdot y')\cdot x''}_{1}, \quad \underbrace{(x\cdot x')\cdot y''}_{3} + \\
&\underbrace{(x\cdot y')\cdot x'' + (y\cdot x')\cdot x''}_{4} - \underbrace{(y\cdot y')\cdot y''}_{3}\Big) &=\\
&\Big(\underbrace{(x\cdot x' - y\cdot y')\cdot x''}_{1} - \underbrace{(x\cdot y' + y\cdot y')\cdot y''}_{2}, \underbrace{(x\cdot x' - y\cdot y')\cdot y''}_{3} + \underbrace{(x\cdot y' + y\cdot x')\cdot x''}_{4}\Big) &=\\
&(x\cdot x' - y\cdot y', x\cdot y' + y\cdot y')\cdot(x'', y'') &=\\
&((x,y)\cdot(x',y'))\cdot(x'',y'') &=
\end{aligned}$$

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

$$\begin{aligned}
(x,y)\cdot(1,0) &\underset{\text{commutativity}}{=} (1,0)\cdot(x,y)\\
&= (1\cdot x - 0\cdot y, 1\cdot y + 0\cdot x)\\
&= (x,y)
\end{aligned}$$

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

$$\begin{aligned}
(x,y)\cdot(x/(x^2+y^2), -y/(x^2+y^2)) &\underset{\text{commutativity}}{=}\\
(x/(x^2+y^2), -y/(x^2+y^2))\cdot(x,y) &=\\
(x^2/(x^2+y^2) + y^2/(x^2+y^2), x\cdot y/(x^2+y^2) + (-y)\cdot x/(x^2+y^2)) &=\\
((x^2+y^2)/(x^2+y^2), (x\cdot y - y\cdot x)/(x^2+y^2)) &=\\
(1,0) &=
\end{aligned}$$

distributivity. If $(x,y),(x',y'),(x'',y'')\in\mathbb{C}$ then

$$\begin{aligned}
(x,y)\cdot((x',y')+(x'',y'')) &= (x,y)\cdot(x'+x'', y'+y'')\\
&= (x\cdot(x'+x'') - y\cdot(y'+y''), x\cdot(y'+y'') + y\cdot(x'+x''))\\
&= (x\cdot x' + x\cdot x'' - y\cdot y' - y\cdot y'', x\cdot y' + x\cdot y'' + y\cdot x' + y\cdot x'')\\
&= ((x\cdot x' - y\cdot y') + (x\cdot x'' - y\cdot y''), (x\cdot y' + y\cdot x') + (x\cdot y'' + \\
&\quad y\cdot x''))\\
&= (x\cdot x' - y\cdot y', x\cdot y' + y\cdot x') + (x\cdot x'' - y\cdot y'', x\cdot y'' + y\cdot x'')\\
&= (x,y)\cdot(x',y') + (x,y)\cdot(x'',y'')
\end{aligned}$$

Finally as in $\mathbb{R}$ we have that $1\neq 0$ [see theorem: 9.32] we have that

$$0 = (0,0) \neq (1,0) = 1 \qquad \square$$

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 $\mathbb{R}_{\mathbb{C}}, \mathbb{Q}_{\mathbb{C}}, \mathbb{Z}_{\mathbb{C}}, \mathbb{N}_{\mathbb{C}}$ as follows*

$$\begin{aligned}\mathbb{R}_{\mathbb{C}} &= \{(x,0)|x\in\mathbb{R}\}\\ \mathbb{Q}_{\mathbb{C}} &= \{(x,0)|x\in\mathbb{Q}_{\mathbb{R}}\}\\ \mathbb{Z}_{\mathbb{C}} &= \{(x,0)|x\in\mathbb{Z}_{\mathbb{R}}\}\\ \mathbb{N}_{0,\mathbb{C}} &= \{(x,0)|x\in\mathbb{N}_{0,\mathbb{R}}\}\end{aligned}$$

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*

$$i_{\mathbb{R}\to\mathbb{C}}\colon \langle\mathbb{R},+,\cdot\rangle\to\langle\mathbb{R}_{\mathbb{C}},+,\cdot\rangle \text{ defined by } i_{\mathbb{R}\to\mathbb{C}}(r)=(r,0)$$

is a field isomorphism.

2. *$\mathbb{Q}_{\mathbb{C}}$ is a sub-field of $\langle\mathbb{C},+,\cdot\rangle$ and if we define $i_{\mathbb{Q}\to\mathbb{C}}=i_{\mathbb{R}\to\mathbb{C}}\circ i_{\mathbb{Q}\to\mathbb{R}}$ then*

$$i_{\mathbb{Q}\to\mathbb{C}}\colon \langle\mathbb{Q},+,\cdot\rangle\to\langle\mathbb{Q}_{\mathbb{C}},+,\cdot\rangle \text{ is a field isomorphism}$$

3. *$\mathbb{Z}_{\mathbb{C}}$ is a sub-ring of $\langle\mathbb{C},+,\cdot\rangle$ and if we define $i_{\mathbb{Z}\to\mathbb{C}}=i_{\mathbb{R}\to\mathbb{C}}\circ i_{\mathbb{Z}\to\mathbb{R}}$ then*

$$i_{\mathbb{Z}\to\mathbb{C}}\colon \langle\mathbb{Z},+,\cdot\rangle\to\langle\mathbb{Z}_{\mathbb{C}},+,\cdot\rangle \text{ is a field isomorphism}$$

4. *$\mathbb{N}_{0,\mathbb{C}}$ is a sub semi-group of $\langle\mathbb{C},+\rangle$ and if we define $i_{\mathbb{N}_0\to\mathbb{C}}=i_{\mathbb{R}\to\mathbb{C}}\circ i_{\mathbb{N}_0\to\mathbb{R}}$ then*

$$i_{\mathbb{N}_0\to\mathbb{C}}\colon \langle\mathbb{N}_0,+\rangle\to\langle\mathbb{N}_{0,\mathbb{C}},+\rangle \text{ is a group isomorphism}$$

5. *$\mathbb{N}_{0,\mathbb{C}}$ is a sub semi-group of $\langle\mathbb{C},\cdot\rangle$ and if we define $i_{\mathbb{N}_0\to\mathbb{C}}=i_{\mathbb{R}\to\mathbb{C}}\circ i_{\mathbb{N}_0\to\mathbb{R}}$ then*

$$i_{\mathbb{N}_0\to\mathbb{C}}\colon \langle\mathbb{N}_0,\cdot\rangle\to\langle\mathbb{N}_{0,\mathbb{C}},\cdot\rangle \text{ is a group isomorphism}$$

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 $x=(x',0)=i_{\mathbb{R}\to\mathbb{C}}(x')\in i_{\mathbb{R}\to\mathbb{C}}(\mathbb{R})$. Also if $x\in i_{\mathbb{R}\to\mathbb{C}}(\mathbb{R})$ then $\exists x'\in\mathbb{R}$ such that $x=i_{\mathbb{R}\to\mathbb{C}}(x')=(x',0)\in\mathbb{R}_{\mathbb{C}}$. So we have that

$$\mathbb{R}_{\mathbb{C}}=i_{\mathbb{R}\to\mathbb{C}}(\mathbb{R}) \tag{10.1}$$

Let $x,y\in\mathbb{R}$ then we have

$$i_{\mathbb{R}\to\mathbb{C}}(x+y)=(x+y,0)=(x,0)+(y,0)=i_{\mathbb{R}\to\mathbb{C}}(x)+i_{\mathbb{R}\to\mathbb{C}}(y)$$

and

$$i_{\mathbb{R}\to\mathbb{C}}(x\cdot y)=(x\cdot y,0)=(x\cdot y-0\cdot 0,x\cdot 0+0\cdot y)=(x,0)\cdot(y,0)=i_{\mathbb{R}\to\mathbb{C}}(x)\cdot i_{\mathbb{R}\to\mathbb{C}}(y)$$

$$i_{\mathbb{R}\to\mathbb{C}}(1)=(1,0)=1$$

proving that

$$i_{\mathbb{R}\to\mathbb{C}}\colon \langle\mathbb{R},+,\cdot\rangle\to\langle\mathbb{C},,+,\cdot\rangle \text{ is a field homeomorphism} \tag{10.2}$$

As $\langle\mathbb{R},+,\cdot\rangle$ is a field we can use [theorem: 4.70] together with the above proving that $i_{\mathbb{R}\to\mathbb{C}}(\mathbb{R})$ is a sub-field of $\langle\mathbb{C},+,\cdot\rangle$. Combining this with [eq: 10.1] gives

$$\mathbb{R}_{\mathbb{C}} \text{ is a sub-field of } \langle\mathbb{C}\dot{+},\cdot\rangle$$

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

   $\langle\mathbb{Q}_{\mathbb{R}},+,\cdot\rangle$ is a sub-field of $\langle\mathbb{R},+,\cdot\rangle$ and $i_{\mathbb{Q}\rightarrow\mathbb{R}}: \langle\mathbb{Q},+,\cdot\rangle \rightarrow \langle\mathbb{Q}_{\mathbb{R}},+,\cdot\rangle$ is a field isomorphism (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

   $\langle\mathbb{Z}_{\mathbb{R}},+,\cdot\rangle$ 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 field isomorphism (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

   $\langle\mathbb{N}_{0,\mathbb{R}},+\rangle$ 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 (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

$$i_{\mathbb{R}\rightarrow\mathbb{C}}\circ i_{\mathbb{N}_0\rightarrow\mathbb{R}}\colon\langle\mathbb{N}_0,+\rangle\rightarrow\langle i_{\mathbb{R}\rightarrow\mathbb{C}}(i_{\mathbb{C}\rightarrow\mathbb{R}}(\mathbb{N}_0)),+\rangle \text{ is a group isomorphism} \tag{10.12}$$

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

$$i_{\mathbb{N}_0\rightarrow\mathbb{C}}=i_{\mathbb{R}\rightarrow\mathbb{C}}\circ i_{\mathbb{N}_0\rightarrow\mathbb{R}}\colon\langle\mathbb{N}_0,+\rangle\rightarrow\langle\mathbb{N}_{0,\mathbb{C}},+\rangle \text{ is a group isomorphism}$$

5. By [theorem: 9.55] we have that

   $\langle\mathbb{N}_{0,\mathbb{R}},\cdot\rangle$ is a sub semi-group of $\langle\mathbb{R},\cdot\rangle$ and $i_{\mathbb{N}_0\rightarrow\mathbb{R}}\colon\langle\mathbb{N}_0,\cdot\rangle\rightarrow\langle\mathbb{N}_{0,\mathbb{R}},\cdot\rangle$ is a group isomorphism (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

   $$i_{\mathbb{R}\rightarrow\mathbb{C}}\circ i_{\mathbb{N}_0\rightarrow\mathbb{R}}\colon\langle\mathbb{N}_0,\cdot\rangle\rightarrow\langle i_{\mathbb{R}\rightarrow\mathbb{C}}(i_{\mathbb{C}\rightarrow\mathbb{R}}(\mathbb{N}_0)),\cdot\rangle \text{ is a group isomorphism} \tag{10.16}$$

   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

   $$i_{\mathbb{N}_0\rightarrow\mathbb{C}}=i_{\mathbb{R}\rightarrow\mathbb{C}}\circ i_{\mathbb{N}_0\rightarrow\mathbb{R}}\colon\langle\mathbb{N}_0,\cdot\rangle\rightarrow\langle\mathbb{N}_{0,\mathbb{C}},\cdot\rangle \text{ is a group isomorphism}$$

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 $x=(q,0)\underset{\text{[theorem: 10.5]}}{=}i_{\mathbb{R}\rightarrow\mathbb{C}}(q)$. 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

$$\begin{array}{rcl} x & = & i_{\mathbb{R}\rightarrow\mathbb{C}}(q)\\ & = & i_{\mathbb{R}\rightarrow\mathbb{C}}(n\cdot m^{-1})\\ & = & i_{\mathbb{R}\rightarrow\mathbb{C}}(n)\cdot i_{\mathbb{R}\rightarrow\mathbb{C}}(m^{-1})\\ & \underset{\text{[theorem: 4.70]}}{=} & i_{\mathbb{R}\rightarrow\mathbb{C}}(n)\cdot(i_{\mathbb{R}\rightarrow\mathbb{C}}(m))^{-1}\\ & \underset{\text{[theorem: 10.5]}}{=} & (n,0)\cdot(m,0)^{-1} \end{array}$$

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

$$\begin{array}{rcl}(n'\cdot m'^{-1},0) & \underset{\text{[theorem: 10.5]}}{=} & i_{\mathbb{R}\rightarrow\mathbb{C}}(n'\cdot m'^{-1})\\ & = & i_{\mathbb{R}\rightarrow\mathbb{C}}(n')\cdot i_{\mathbb{R}\rightarrow\mathbb{C}}(m'^{-1})\\ & \underset{\text{[theorem: 4.70]}}{=} & i_{\mathbb{R}\rightarrow\mathbb{C}}(n')\cdot(i_{\mathbb{R}\rightarrow\mathbb{C}}(m'))^{-1}\\ & = & n\cdot m^{-1}\\ & = & x\end{array}$$

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: 3.36] 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. *$i_{\mathbb{R}\rightarrow\mathbb{C}}$: $\langle\mathbb{R},\leqslant\rangle\rightarrow\langle\mathbb{R}_{\mathbb{C}},\leqslant\rangle$ 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] $i_{\mathbb{R}\rightarrow\mathbb{C}}: \mathbb{R} \rightarrow \mathbb{R}_{\mathbb{C}}$ is a bijection and second

$$\begin{array}{rcl} i_{\mathbb{R}\rightarrow\mathbb{R}_{\mathbb{C}}}(x) \leqslant i_{\mathbb{R}\rightarrow\mathbb{C}}(y) & \Leftrightarrow & (x,0) \leqslant (y,0) \\ & \underset{\text{definition}}{\Leftrightarrow} & x \leqslant y \end{array}$$

□

Theorem 10.9. *We have that*

1. $i_{\mathbb{R}\rightarrow\mathbb{C}}: \langle \mathbb{R}, \leqslant \rangle \rightarrow \langle \mathbb{R}_{\mathbb{C}}, \leqslant \rangle$ *is a order isomorphism*
2. $i_{\mathbb{Q}\rightarrow\mathbb{C}}: \langle \mathbb{Q}, \leqslant \rangle \rightarrow \langle \mathbb{Q}_{\mathbb{C}}, \leqslant \rangle$ *is a order isomorphism*
3. $i_{\mathbb{Z}\rightarrow\mathbb{C}}: \langle \mathbb{Z}, \leqslant \rangle \rightarrow \langle \mathbb{Z}_{\mathbb{C}}, \leqslant \rangle$ *is a order isomorphism*
4. $i_{\mathbb{N}_0\rightarrow\mathbb{C}}: \langle \mathbb{N}_0, \leqslant \rangle \rightarrow \langle \mathbb{N}_{0,\mathbb{C}}, \leqslant \rangle$ *is a order isomorphism*

Proof.

1. This is expressed by [theorem: 10.8].
2. Using (1) and [theorem: 9.49] we have that

$$i_{\mathbb{R}\rightarrow\mathbb{C}}: \langle \mathbb{R}, \leqslant \rangle \rightarrow \langle \mathbb{R}_{\mathbb{C}}, \leqslant \rangle \text{ and } i_{\mathbb{Q}\rightarrow\mathbb{R}}: \langle \mathbb{Q}, \leqslant \rangle \rightarrow \langle \mathbb{Q}_{\mathbb{R}}, \leqslant \rangle \text{ are order isomorphisms}$$

so using [theorem 3.51]

$$i_{\mathbb{Q}\rightarrow\mathbb{C}} = i_{\mathbb{R}\rightarrow\mathbb{C}} \circ i_{\mathbb{Q}\rightarrow\mathbb{R}}: \langle \mathbb{Q}, \leqslant \rangle \rightarrow \langle i_{\mathbb{R}\rightarrow\mathbb{C}}(i_{\mathbb{Q}\rightarrow\mathbb{R}}(\mathbb{Q})), \leqslant \rangle \text{ is a order isomorphism}$$

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

$$i_{\mathbb{Q}\rightarrow\mathbb{C}} = i_{\mathbb{R}\rightarrow\mathbb{C}} \circ i_{\mathbb{Q}\rightarrow\mathbb{R}}: \langle \mathbb{Q}, \leqslant \rangle \rightarrow \langle \mathbb{Q}_{\mathbb{C}}, \leqslant \rangle \text{ is a order isomorphism}$$

3. Using (1) and [theorem: 9.52] we have that

$$i_{\mathbb{R}\rightarrow\mathbb{C}}: \langle \mathbb{R}, \leqslant \rangle \rightarrow \langle \mathbb{R}_{\mathbb{C}}, \leqslant \rangle \text{ and } i_{\mathbb{Z}\rightarrow\mathbb{R}}: \langle \mathbb{Z}, \leqslant \rangle \rightarrow \langle \mathbb{Z}_{\mathbb{R}}, \leqslant \rangle \text{ are order isomorphisms}$$

so using [theorem 3.51]

$$i_{\mathbb{Z}\rightarrow\mathbb{C}} = i_{\mathbb{R}\rightarrow\mathbb{C}} \circ i_{\mathbb{Z}\rightarrow\mathbb{R}}: \langle \mathbb{Z}, \leqslant \rangle \rightarrow \langle i_{\mathbb{R}\rightarrow\mathbb{C}}(i_{\mathbb{Z}\rightarrow\mathbb{R}}(\mathbb{Z})), \leqslant \rangle \text{ is a order isomorphism}$$

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

$$i_{\mathbb{Z}\rightarrow\mathbb{C}} = i_{\mathbb{R}\rightarrow\mathbb{C}} \circ i_{\mathbb{Z}\rightarrow\mathbb{R}}: \langle \mathbb{Z}, \leqslant \rangle \rightarrow \langle \mathbb{Z}_{\mathbb{C}}, \leqslant \rangle \text{ is a order isomorphism}$$

4. Using (1) and [theorem: 9.55] we have that

$$i_{\mathbb{R}\rightarrow\mathbb{C}}: \langle \mathbb{R}, \leqslant \rangle \rightarrow \langle \mathbb{R}_{\mathbb{C}}, \leqslant \rangle \text{ and } i_{\mathbb{N}_0\rightarrow\mathbb{R}}: \langle \mathbb{N}_0, \leqslant \rangle \rightarrow \langle \mathbb{N}_{0,\mathbb{R}}, \leqslant \rangle \text{ are order isomorphisms}$$

so using [theorem 3.51]

$$i_{\mathbb{N}_0\rightarrow\mathbb{C}} = i_{\mathbb{R}\rightarrow\mathbb{C}} \circ i_{\mathbb{N}_0\rightarrow\mathbb{R}}: \langle \mathbb{N}_0, \leqslant \rangle \rightarrow \langle i_{\mathbb{R}\rightarrow\mathbb{C}}(i_{\mathbb{N}_0\rightarrow\mathbb{R}}(\mathbb{N}_0)), \leqslant \rangle \text{ is a order isomorphism}$$

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

$$i_{\mathbb{N}_0\rightarrow\mathbb{C}} = i_{\mathbb{R}\rightarrow\mathbb{C}} \circ i_{\mathbb{N}_0\rightarrow\mathbb{R}}: \langle \mathbb{N}_0, \leqslant \rangle \rightarrow \langle \mathbb{N}_{0,\mathbb{C}}, \leqslant \rangle \text{ is a order isomorphism}$$ □

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

$$i_{\mathbb{R} \rightarrow \mathbb{C}}\colon \langle \mathbb{R} \dot{+}, \cdot \rangle \rightarrow \langle \mathbb{R}, +, \cdot \rangle \text{ is a field isomorphism} \tag{10.21}$$

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 $x \cdot \leqslant y \cdot z$*
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: 3.74] in other words*

$\forall S \subseteq \mathbb{Z}_{\mathbb{C}}$ with $S \neq \varnothing$ such that $\exists \upsilon \in \mathbb{Z}_{\mathbb{C}}$ such that $\forall \alpha \in S$ we have $\alpha \leqslant \upsilon$ we have that $\sup(S)$ exist

Using [theorem: 3.75] this is equivalent with

$\forall S \subseteq \mathbb{Z}_{\mathbb{C}}$ with $S \neq \varnothing$ such that $\exists \lambda \in \mathbb{Z}_{\mathbb{C}}$ such that $\forall \alpha \in S$ we have $\lambda \leqslant \alpha$ we have that $\inf(S)$ 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: 3.77]

$$\langle \mathbb{Z}_{\mathbb{C}}, \leqslant \rangle \text{ is conditional complete}$$ □

Theorem 10.17. *$\langle \mathbb{Q}_{\mathbb{C}}, \leqslant \rangle$ is not conditional complete.*

Proof. Using [theorems: 10.9 and 3.54] 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: 3.77] $\langle\mathbb{Q}, \leqslant\rangle$ 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: 3.74] in other words*

$\forall S \subseteq \mathbb{R}_{\mathbb{C}}$ *with* $S \neq \varnothing$ *such that* $\exists \upsilon \in \mathbb{R}_{\mathbb{C}}$ *such that* $\forall \alpha \in S$ *we have* $\alpha \leqslant \upsilon$ *we have that* $\sup(S)$ *exist*

Using [theorem: 3.75] this is equivalent with

$\forall S \subseteq \mathbb{R}_{\mathbb{C}}$ *with* $S \neq \varnothing$ *such that* $\exists \lambda \in \mathbb{R}_{\mathbb{C}}$ *such that* $\forall \alpha \in S$ *we have* $\lambda \leqslant \alpha$ *we have that* $\inf(S)$ *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: 3.77]

$$\langle\mathbb{R}_{\mathbb{C}}, \leqslant\rangle \text{ is conditional complete}$$

□

Theorem 10.19. *$\langle\mathbb{N}_{0,\mathbb{C}}, \leqslant\rangle$ is well ordered [definition: 3.78]*

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: 3.79] 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 $\mathbb{R}_{\mathbb{C}} \backslash \mathbb{Q}_{\mathbb{C}} \neq \varnothing$. The set $\mathbb{R}_{\mathbb{C}} \backslash \mathbb{Q}$ 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 $s = \min(\{u \in \mathbb{R}_{\mathbb{C}} | u \text{ is a upper bound of } S\})$ 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 $\mathbb{R}_{\mathbb{C}} \backslash \mathbb{Q}_{\mathbb{C}} \neq \varnothing$. □

Theorem 10.21. *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* $\inf(S+T) = \inf(S) + \sup(T)$

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 $\langle\mathbb{R}, \leqslant\rangle$ is conditional complete [see theorem: 10.18] proves that

$$\sup(S+T) \text{ exist and } \sup(S+T) \leqslant \sup(S) + \sup(T) \tag{10.22}$$

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 $\langle \mathbb{R}, \leqslant \rangle$ is totally ordered we have by [theorem: 3.68] that there exists $s \in S$ and $t \in T$ such that $\sup(S) - \varepsilon/2 < s$ and $\sup(T) - \varepsilon/2 < t$. So

$$\begin{aligned} s+t &> \sup(S) - \varepsilon/2 + \sup(T) - \varepsilon/2 \\ &= \sup(S) + \sup(T) - (\varepsilon + \varepsilon)/2 \\ &= \sup(S) + \sup(T) - \varepsilon \\ &= \sup(S) + \sup(T) - \sup(S) - \sup(T) + \sup(S+T) \\ &= \sup(s+t) \end{aligned} \tag{10.23}$$

As $s+t \in S+T$ we have that $s+t \leqslant \sup(S+T)$ contradicting [eq: 10.23], so the assumption is wrong and we must have $\sup(S) + \sup(T) \leqslant \sup(S+T)$ which combined with [eq: 10.22] 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 $\langle \mathbb{R}, \leqslant \rangle$ is conditional complete [see theorem: 10.18] proves that

$$\inf(S+T) \text{ exist and } \inf(S) + \inf(T) \leqslant \inf(S+T) \tag{10.24}$$

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 $\langle \mathbb{R}, \leqslant \rangle$ is totally ordered we have by [theorem: 3.68] that there exists $s \in S$ and $t \in T$ such that $s < \inf(S) + \varepsilon/2$ and $t < \inf(T) + \varepsilon/2$. So

$$\begin{aligned} s+t &< \inf(S) + \varepsilon/2 + \inf(T) + \varepsilon/2 \\ &= \inf(S) + \inf(T) + \varepsilon \\ &= \inf(S) + \inf(T) + \inf(S+T) - \inf(S) - \inf(T) \\ &= \inf(S+T) \end{aligned} \tag{10.25}$$

As $s+t \in S+T$ we have that $\inf(S+T) \leqslant s+t$ contradicting [eq: 10.25], so the assumption is wrong and we must have $\inf(S+T) \leqslant \inf(S) + \inf(T)$ which combined with [eq: 10.24] proves that

$$\inf(S+T) = \inf(S) + \inf(T)$$

□

Corollary 10.22. *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 10.21] that $\sup(S + \{\alpha\})$ 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 10.21] 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*

$$\inf(S+\alpha) \text{ exist and } \inf(S+\alpha)=\inf(S)+\alpha$$ □

Theorem 10.23. *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

$$i_{\mathbb{Z}\rightarrow\mathbb{C}}: \langle\mathbb{Z},\leqslant\rangle\rightarrow\langle\mathbb{Z}_{\mathbb{C}},\leqslant\rangle \text{ is a order isomorphism} \tag{10.26}$$

$$i_{\mathbb{Z}\rightarrow\mathbb{Z}_{\mathbb{C}}}: \langle\mathbb{Z},+,\cdot\rangle\rightarrow\langle\mathbb{Z}_{\mathbb{C}},+,\cdot\rangle \text{ is a field isomorphism} \tag{10.27}$$

Let $x,y\in\mathbb{Z}_{\mathbb{C}}$ then by [eq: 10.26] there exists $x',y'\in\mathbb{Z}$ such that

$$x=i_{\mathbb{Z}\rightarrow\mathbb{C}}(x')\wedge y=i_{\mathbb{Z}\rightarrow\mathbb{C}}(y')\wedge x'<y' \tag{10.28}$$

1. Using [theorem: 7.28] we have that $x'+1\leqslant y'$ so that

   $$x+1 \underset{\text{[eq: 10.27]}}{=} i_{\mathbb{Z}\rightarrow\mathbb{C}}(x')+i_{\mathbb{Z}\rightarrow\mathbb{C}}(1) \underset{\text{[eq: 10.26]}}{=} i_{\mathbb{Z}\rightarrow\mathbb{C}}(x'+1)\leqslant_{\text{[eq: 10.26]}} i_{\mathbb{Z}}(y')=y$$

   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 10.24. (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

$y\leqslant 0$. Then for $n=1$ we have $y\leqslant 0<x=1\cdot x=n\cdot x$ proving that $y<n\cdot x$

$0<y$. 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: 3.68] 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

$$\sup(A)<n\cdot x+x=(n+1)\cdot x \tag{10.29}$$

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: 10.29]. So our assumption is wrong hence

$$\exists n\in\mathbb{N}_{0,\mathbb{R}} \text{ such that } y<n\cdot x$$ □

Corollary 10.25. *Let* $x\in\mathbb{R}_{\mathbb{C}}$ *then we have*

1. $\exists n\in\mathbb{N}_{0,\mathbb{C}}$ *such that* $x<n$
2. $\exists n\in\mathbb{Z}_{\mathbb{C}}$ *such that* $n\leqslant x<n+1$
3. $\exists n\in\mathbb{Z}_{\mathbb{C}}$ *such that* $n<x\leqslant n+1$
4. $\exists n\in\mathbb{Z}_{\mathbb{C}}$ *such that* $n-1\leqslant x<n$
5. $\exists n\in\mathbb{Z}_{\mathbb{C}}$ *such that* $n-1<x\leqslant n$
6. *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: 10.24] that there exist a $n\in\mathbb{N}_{0,\mathbb{C}}$ such that $x<n\cdot 1=n$
2. 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 $n\leqslant x<m=n+1$.
3. 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 $n<x\leqslant n+1$
4. 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 $n-1\leqslant x<n$.
5. 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 $n-1<x\leqslant n$.
6. Using the Archimedean Property [see theorem: 10.24] 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 10.26. *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: 10.25] 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: 10.25] 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: 10.25] 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 10.27. (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. *$\exists r\in\mathbb{R}_{\mathbb{C}}\backslash\mathbb{Q}_{\mathbb{C}}$ 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

   $0<x$. From $x<y$ we have that $0<y-x$ so that by [theorem: 10.14] $0<(y-x)^{-1}$. Using [corollary: 10.25] 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

   $$1+n\cdot x<n\cdot y \tag{10.30}$$

   and from $0<(y-x)^{-1}<n$ that

   $$0<n\underset{[\text{theorem: }10.14]}{\Rightarrow}0<n^{-1}=1/n \tag{10.31}$$

   Using [corollary: 10.25] there exist a $m\in\mathbb{N}_{0,\mathbb{R}}$ such that

   $$m-1\leqslant n\cdot x<m \tag{10.32}$$

   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

   $$x<q\in\mathbb{Q}_{\mathbb{C}} \tag{10.33}$$

   From [eq: 10.32] 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: 10.33] gives finally

   $$x<q<y\text{ where }q\in\mathbb{Q}_{\mathbb{C}}$$

   $x=0$. Then $0<y$ and by [corollary: 10.25] 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

   $$\forall x,y\in\mathbb{R}_{\mathbb{C}}\text{ with }0\leqslant x\wedge x<y\text{ there exist a }q\in\mathbb{Q}_{\mathbb{C}}\text{ such that }x<q<y \tag{10.34}$$

   So the only case left to prove is where $x<0$. Then for y we have either

   $0<y$. Then if we take $q=0\in\mathbb{N}_{0,\mathbb{C}}\subseteq\mathbb{Z}_{\mathbb{C}}\subseteq\mathbb{Q}_{\mathbb{C}}$ so that

   $$x<q<y\text{ with }q\in\mathbb{Q}_{\mathbb{C}}$$

   $y\leqslant 0$. 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: 10.34] 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 10.28. *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 10.29. *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

$$i_{\mathbb{N}_0 \rightarrow \mathbb{C}} \colon \langle \mathbb{N}_0, \leqslant \rangle \rightarrow \langle \mathbb{N}_{0,\mathbb{C}}, \leqslant \rangle \text{ is a order isomorphism} \tag{10.35}$$

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: 10.35] 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
$$i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(\{n, \ldots, m\}) \subseteq \{i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(n), \ldots, i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(m)\} \tag{10.36}$$
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: 10.35] $\exists x \in \mathbb{N}_0$ 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: 10.35] 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: 10.36] 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: 10.35] 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
$$i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(\{n, \ldots, \infty\}) \subseteq \{i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(n), \ldots, \infty\} \tag{10.37}$$
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: 10.35] $\exists x \in \mathbb{N}_0$ 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: 10.35] 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: 10.37] 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

$$\mathbb{N}_{0,\mathbb{C}} \underset{[\text{eq } 10.35]}{=} i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(\mathbb{N}_0) = i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(\{0, \ldots, \infty\}) \underset{(2)}{=} \{i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(0), \ldots, \infty\} = \{0, \ldots, \infty\} \qquad \square$$

Next we state Mathematical Induction for $\mathbb{N}_{0,\mathbb{C}}$ using Mathematical Induction using $\mathbb{N}_0$.

Theorem 10.30. *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, 3.54] we have that

$$i_{\mathbb{N}_0 \rightarrow \mathbb{C}} \colon \langle \mathbb{N}_0, + \rangle \rightarrow \langle \mathbb{N}_{0,\mathbb{C}}, + \rangle \text{ is a group isomorphisms and } i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(1) = 1 \tag{10.38}$$

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

$$i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(n + 1) \underset{[\text{eq: } 10.38]}{=} i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(n) + i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(1) \underset{[\text{eq: } 10.38]}{=} n' + 1 \in S$$

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

$$i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(\{k', \ldots, \infty\}) = i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(T) = i_{\mathbb{N}_0 \rightarrow \mathbb{C}}((i_{\mathbb{N}_0 \rightarrow \mathbb{C}})^{-1}(S)) \underset{[\text{eq: } 10.38] \text{ and } [\text{theorem: } 2.55]}{=} S \tag{10.39}$$

Next

$$\begin{aligned} i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(\{k', \ldots, \infty\}) &\underset{[\text{lemma: } 10.29]}{=} \{i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(k'), \ldots, \infty\} \\ &= \{k, \ldots, \infty\} \end{aligned}$$

which combined with [eq: 10.39] gives finally

$$S = \{k, \ldots, \infty\} \qquad \square$$

As a example of using recursion we have the following theorem.

Theorem 10.31. *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:

$\boldsymbol{n \in S}$**.** If $i \in \{n, \ldots, n\}$ then $i = n$ so that $x_i = x_n \leqslant x_n$ proving that $n \in S$.

$\boldsymbol{k \in S \Rightarrow k + 1 \in S}$**.** Let $i \in \{n, \ldots, k + 1\}$ then for i we have either:

$\boldsymbol{i = k + 1}$**.** Then $x_i = x_{k+1} \leqslant x_{k+1}$

$\boldsymbol{i \in \{n, \ldots, k\}}$**.** 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: 10.30] 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:

$\boldsymbol{n \in S}$**.** If $i \in \{n, \ldots, n\}$ then $i = n$ and $x_i = x_n \leqslant x_n$ proving that $n \in S$.

$\boldsymbol{k \in S \Rightarrow k+1 \in S}$**.** For $k+1$ we have either:

$\boldsymbol{k+1 \in \{m+1, \ldots, \infty\}}$**.** Then $k+1 \in S$

$\boldsymbol{k+1 \notin \{m+1, \ldots, \infty\}}$**.** Then $k < k+1 < m+1$ so that $k \notin \{m+1, \ldots, \infty\}$. For $i \in \{n, \ldots, k+1\}$ we have either:

$\boldsymbol{i = k+1}$**.** Then $x_i = x_{k+1} \leqslant x_{k+1}$

$\boldsymbol{i \in \{n, \ldots, k\}}$**.** 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 10.32. *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

$$i_{\mathbb{N}_0 \rightarrow \mathbb{C}}: \langle \mathbb{N}_0, + \rangle \rightarrow \langle \mathbb{N}_{0,\mathbb{C}}, + \rangle \text{ and } (i_{\mathbb{N}_0 \rightarrow \mathbb{C}})^{-1}: \langle \mathbb{N}_{0,\mathbb{C}}, + \rangle \rightarrow \langle \mathbb{N}_0, + \rangle \text{ are group isomorphisms} \quad (10.40)$$

and

$$i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(1) = 1 \quad (10.41)$$

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

$$(i_{\mathbb{N}_0 \rightarrow \mathbb{C}})^{-1}(n+1) = (i_{\mathbb{N}_0 \rightarrow \mathbb{C}})^{-1}(n) + (i_{\mathbb{N}_0 \rightarrow \mathbb{C}})^{-1}(1) \underset{\text{[eq: 10.41]}}{=} (i_{\mathbb{N}_0 \rightarrow \mathbb{R}})^{-1}(n) + 1 = n' + 1$$

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

$$\begin{aligned}(\gamma \circ i_{\mathbb{N}_0 \rightarrow \mathbb{R}})(n+1) &= \gamma(i_{\mathbb{N}_0 \rightarrow \mathbb{R}}(n+1))\\ &= \gamma(i_{\mathbb{N}_0 \rightarrow \mathbb{R}}(n) + i_{\mathbb{N}_0 \rightarrow \mathbb{R}}(1))\\ &= \gamma(i_{\mathbb{N}_0 \rightarrow \mathbb{R}}(n) + 1)\\ &= f(\gamma(i_{\mathbb{N}_0 \rightarrow \mathbb{R}}(n)))\\ &= f((\gamma \circ i_{\mathbb{N}_0 \rightarrow \mathbb{R}})(n))\end{aligned}$$

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 10.33. (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: 10.32] 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 10.34. *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

$$i_{\mathbb{N}_0 \rightarrow \mathbb{C}}: \langle \mathbb{N}_0, + \rangle \rightarrow \langle \mathbb{N}_{0,\mathbb{C}}, + \rangle \text{ and } (i_{\mathbb{N}_0 \rightarrow \mathbb{C}})^{-1}: \langle \mathbb{N}_{0,\mathbb{C}}, + \rangle \rightarrow \langle \mathbb{N}_0, + \rangle \text{ are group isomorphisms} \quad (10.42)$$

and

$$i_{\mathbb{N}_0 \rightarrow \mathbb{C}}(1) = 1 \quad (10.43)$$

Define now

$$h: \mathbb{N}_0 \times A \rightarrow A \text{ by } h(n, a) = g((i_{\mathbb{N}_0 \rightarrow \mathbb{C}})(n), a)$$

Using [theorem: 5.86] there exist a $\beta: \mathbb{N}_0 \rightarrow A$ 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

$$\lambda: \mathbb{N}_{0,\mathbb{C}} \rightarrow A \text{ by } \lambda = \beta \circ (i_{\mathbb{N}_0 \rightarrow \mathbb{C}})^{-1}$$

then we have:

1. $\lambda(0) = \beta((i_{\mathbb{N}_0 \rightarrow \mathbb{C}})^{-1}(0)) \underset{\text{[eq: 10.42]}}{=} \beta(0) = a$
2. If $n \in \mathbb{N}_{0,\mathbb{C}}$ then

$$\begin{array}{rcl}\lambda(n+1) & = & \beta((i_{\mathbb{N}_0 \rightarrow \mathbb{C}})^{-1}(n+1)) \\ & \underset{\text{[eq: 10.42]}}{=} & \beta((i_{\mathbb{N}_0 \rightarrow \mathbb{C}})^{-1}(n) + (i_{\mathbb{N}_0 \rightarrow \mathbb{C}})^{-1}(1)) \\ & \underset{\text{[eq: 10.43]}}{=} & \beta((i_{\mathbb{N}_0 \rightarrow \mathbb{C}})^{-1}(n) + 1) \\ & = & h((i_{\mathbb{N}_0 \rightarrow \mathbb{C}})^{-1}(n), \beta((i_{\mathbb{N}_0 \rightarrow \mathbb{C}})^{-1}(n))) \\ & = & h((i_{\mathbb{N}_0 \rightarrow \mathbb{C}})^{-1}(n), \lambda(n)) \\ & = & g(i_{\mathbb{N}_0 \rightarrow \mathbb{C}}((i_{\mathbb{N}_0 \rightarrow \mathbb{C}})^{-1}(n)), \lambda(n)) \\ & = & g(, \lambda(n))\end{array}$$

which proves existence. Now for uniqueness, assume that there is a

$$\gamma: \mathbb{N}_{0,\mathbb{C}} \rightarrow A \text{ such that } \beta(0) = a \text{ and } \forall n \in \mathbb{N}_{0,\mathbb{C}} \text{ that } \beta(n+1) = g(n, \beta(n))$$

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: 10.30] 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 10.35. *Let A be a set, $a \in A$ and $g: \mathbb{N}_{0,\mathbb{C}} \times A \rightarrow A$ then there exist a **unique** function*

$$\lambda: \mathbb{N}_{0,\mathbb{C}} \rightarrow A$$

such that:

1. $\lambda(0) = a$
2. $\forall n \in \{1, \ldots, \infty\}$ $\lambda(n) = g(n-1, \lambda(n-1))$

Proof. Using [theorem: 10.34] there exists a **unique** $\lambda: \mathbb{N}_{0,\mathbb{C}} \rightarrow A$ 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 10.36. *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*

$$\begin{array}{rcl}\lambda(0) & = & a \\ \forall i \in \{0, \ldots, n-1\} \text{ we have } \lambda(i+1) & = & g(i, \lambda(i))\end{array}$$

Proof. Define

$$g': \mathbb{N}_0 \times A \rightarrow A \text{ by } g'(i, x) = \left\{ \begin{array}{l} g(i, x) \text{ if } i \in \{0, \ldots, n-1\} \\ x \text{ if } i \in \{n, \ldots, \infty\} \end{array} \right.$$

then by [corollary: 5.87] there exists a $\beta\colon \mathbb{N}_0 \rightarrow A$ such that

$$\begin{aligned} \beta(0) &= a & (10.44)\\ \forall i \in \mathbb{N}_0 \text{ we have } \beta(i+1) &= g'(i, \beta(i)) & (10.45) \end{aligned}$$

Define now $\lambda\colon \{0, \ldots, n\} \rightarrow A$ by $\lambda = \beta_{|\{0, \ldots, n\}}$ then we have

$$\lambda(0) = \beta_{|\{0, \ldots, n\}}(0) \underset{0 \in \{0, \ldots, n\}}{=} \beta(0) \underset{[\text{theorem: } 10.44]}{=} a$$

and $\forall i \in \{0, \ldots, n-1\}$ we have

$$\begin{aligned} \lambda(i+1) &= \beta_{|\{0, \ldots, n\}}(i+1)\\ &\underset{i+1 \in \{0, \ldots, n\}}{=} \beta(i+1)\\ &\underset{[\text{theorem: } 10.45]}{=} g'(i, \beta(i))\\ &\underset{i \in \{0, \ldots, n-1\}}{=} g'(i, \lambda(i))\\ &\underset{i \in \{0, \ldots, n-1\}}{=} g(i, \lambda(i)) \end{aligned}$$

so that we found a function $\lambda\colon \{0, \ldots, n\} \rightarrow A$ 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 $\gamma\colon \{0, \ldots, n\} \rightarrow A$ 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:

$0 \in S$. 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:

$i + 1 \in \{0, \ldots, n\}$. $i + 1 \in \{0, \ldots, n\}$. 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 $i + 1 \in S$

$i + 1 \notin \{0, \ldots, n\}$. 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 10.37. *Let A be a set, $a \in A$ then we can define a function as follows:*

$$f\colon \mathbb{N}_{0,\mathbb{C}} \rightarrow A$$

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: 10.34] where $a \in A$ and $g\colon \mathbb{N}_{0,\mathbb{C}} \times A \rightarrow A$ is defined by $g(n, x) = G(n, x)$.

Another way to define a recursive function is based on [corollary: 10.35]

Definition 10.38. *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: 10.35] 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 10.39. *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: 10.36] 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 10.40. *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 10.41. 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 10.42. *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 10.43. *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*

$$\begin{array}{rcl}x^0 & = & 1\\ x^{n+1} & = & x\cdot x^n\end{array}$$

Note 10.44. 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 10.45. *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: 10.30]

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

   $\boldsymbol{0\in S_x}$**.** Using [corollary: 10.10] $0<1=x^0$ so that $0\in S_x$.

   $\boldsymbol{n\in S_x\Rightarrow n+1\in S_x}$**.** 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 10.46. *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:

$\boldsymbol{0\in S_{n,x}}$**.** We have $x^{n+0}=x^n=x^n\cdot 1=x^n\cdot x^0$ proving that $0\in S_{n,x}$.

$\boldsymbol{m\in S_{n,x}\Rightarrow m+1\in S_{n,x}}$**.** Then

$$\begin{aligned}
x^{n+(m+1)} &= x^{(n+m)+1}\\
&= x\cdot x^{(n+m)}\\
&= x^{n+m}\cdot x\\
&\underset{m\in S_{n,x}}{=} (x^n\cdot x^m)\cdot x\\
&= x^n\cdot(x^m\cdot x)\\
&= x^n\cdot(x\cdot x^m)\\
&= x^n\cdot x^{m+1}
\end{aligned}$$

proving that $m+1\in S_{n,x}$

So $\forall n\in\mathbb{N}_{0,\mathbb{C}}$ we have by mathematical induction [theorem: 10.30] 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 10.47. *Let $x\in\mathbb{C}\backslash\{0\}$ 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 $S=\{n\in\mathbb{N}_{0,\mathbb{C}}|(x^{-1})^n=(x^n)^{-1}\}$ 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

$$\begin{array}{lcl}(x^{-1})^{n+1} & = & (x^{-1})\cdot(x^{-1})^n\\ & \underset{n\in S}{=} & (x^{-1})\cdot(x^n)^{-1}\\ & \underset{\text{[theorem: 4.55]}}{=} & (x\cdot x^n)^{-1}\\ & = & (x^{n+1})^{-1}\end{array}$$

proving that $n+1\in S$. □

Theorem 10.48. *Let $n\in\mathbb{N}_{0,\mathbb{C}}$ then we have*

1. *If $n\neq 0$ then $0^n=0$ [note that by definition $0^0=1$]*
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. If $n\neq 0$ then $\exists m\in\mathbb{N}_{0,\mathbb{C}}$ such that $n=m+1$ 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:

   $0\in S$. $1^0=1$ by definition, proving that $0\in S$

   $n\in S\Rightarrow n+1\in S$. $1^{n+1}=1\cdot 1^n\underset{n\in S}{=}1\cdot 1=1$ 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:

   $0\in S$. $(-1)^0=1$ proving that $0\in S$.

   $n\in S\Rightarrow n+1\in S$. As $n\in S$ we have either:

   $(-1)^n=1$. Then $(-1)^{n+1}=(-1)\cdot(-1)^n=(-1)\cdot 1=-1$ so the $n+1\in S$

   $(-1)^n=-1$. 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. $(-1)^{2\cdot n}=(-1)^{(1+1)\cdot n}=(-1)^{n+n}\underset{\text{[theorem: 10.46]}}{=}(-1)^n\cdot(-1)^n\underset{\text{[theorem: 4.40] and (3)}}{=}1$
5. $(-1)^{2\cdot n+1}=(-1)\cdot(-1)^{2\cdot n}\underset{(4)}{=}(-1)\cdot 1=-1$ □

Theorem 10.49. *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. *$\forall\alpha,\beta\in\mathbb{C}$ 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:

   $1 \in S$. As $0 < \alpha < 1$ we have $0 < \alpha^1 < 1$ so that $1 \in S$

   $n \in S \Rightarrow n + 1 \in S$. 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:

      $1 \in S$. As $\alpha < \beta = \beta^1$ we have that $1 \in S$

      $n \in S \Rightarrow n + 1 \in S$. As $1 \leqslant \alpha < \beta \Rightarrow 1 < \beta$ we have by [theorem: 10.14] that

$$\alpha = 1 \cdot \alpha < \beta \cdot \alpha = \alpha \cdot \beta \tag{10.46}$$

      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: 10.46] 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:

      $1 \in S$. As $\alpha \leqslant \beta = \beta^1$ we have that $1 \in S$

      $n \in S \Rightarrow n + 1 \in S$. As $1 \leqslant \alpha \leqslant \beta \Rightarrow 1 \leqslant \beta$ we have by [theorem: 10.14] that

$$\alpha = 1 \cdot \alpha \leqslant \beta \cdot \alpha = \alpha \cdot \beta \tag{10.47}$$

      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: 10.47] 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 10.50. *$\forall n \in \mathbb{N}_{0,\mathbb{C}}$ 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

$0 \in S$. As $n = 0 < 1 = 2^0$ we have that $0 \in S$

$n \in S \Rightarrow n + 1 \in S$. For $n + 1$ we have the following cases to consider:

$n + 1 = 1$. Then $n + 1 = 1 < 2 = 2^1 = 2^{n+1}$ proving that in this case $n + 1 \in S$

$1 < n + 1$. Then by [theorem: 10.23] 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 10.51. *$\forall x \in \mathbb{R}_{\mathbb{C}}$ 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: 10.25] there exist a $n \in \mathbb{N}_{0,\mathbb{C}}$ such that $x < n$, using [theorem: 10.50] $n < 2^n$ so we have $x < 2^n$. □

Lemma 10.52. *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:

$0 \in S$. 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 $0 \in S$.

$n \in S \Rightarrow n + 1 \in S$. As $n \in S$ we have $n \cdot (x-1) \leqslant x^n - 1$ which as $0 < x$ prove that

$$x \cdot n \cdot (x-1) \leqslant x \cdot (x^n - 1) \tag{10.48}$$

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: 10.48] proves that

$$n \cdot (x-1) \leqslant x \cdot (x^n - 1) \tag{10.49}$$

Further

$$\begin{array}{rcl} x^{n+1} - 1 & = & x \cdot x^n - 1 \\ & = & x \cdot x^n - x + x - 1 \\ & = & x \cdot (x^n - 1) + (x - 1) \\ & \geqslant_{[\text{eq: } 10.49]} & n \cdot (x-1) + (x-1) \\ & = & (n+1) \cdot (x-1) \end{array}$$

proving that $n + 1 \in S$. □

Theorem 10.53. *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:

$m = 0$. Then $m = 0 < 1 = x^0$ proving the theorem in this case.

$m = 1$. Then $m = 1 < x = x^1$ proving the theorem in this case.

$1 < m$. Then $0 < m - 1$ and $0 < x - 1 \Rightarrow 0 < (x-1)^{-1}$ so that $(m-1)/(x-1)$ is defined. By [corollary: 10.25] 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: 10.52] $n \cdot (x-1) \leqslant (x^n - 1)$, hence $m - 1 \leqslant x^n - 1$ or $m < x^n$, proving the theorem in this case. □

Theorem 10.54. *If $x \in \mathbb{C}$ with $0 < x < 1$ and $n, m \in \mathbb{N}_{0,\mathbb{C}}$ 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:

$1 \in S_m$. As $0 < x$ we have by [theorem: 10.45] 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 $1 \in S_m$.

$n \in S_m \Rightarrow n + 1 \in S_m$. As $x < 1$ and by [theorem: 10.45] $0 < x^{m+n}$ 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: 10.30] 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. □

10.4 The square root in $\mathbb{R}_{\mathbb{C}}$

Theorem 10.55. *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

$$S_y\neq\varnothing \tag{10.50}$$

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

$$S_y \text{ is bounded above} \tag{10.51}$$

As $\langle\mathbb{R}_{\mathbb{C}},\leqslant\rangle$ is conditionally complete [see theorem: 10.18] we have thanks to [eqs: 10.50, 10.51] that

$$s_y=\sup(S_y) \text{ exist} \tag{10.52}$$

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

$$0<s_y \tag{10.53}$$

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

$$(s_y-\varepsilon)^2<s_y^2<(s_y+\varepsilon)^2 \tag{10.54}$$

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

$$y < (s_y + \varepsilon)^2 \tag{10.55}$$

As $\langle \mathbb{R}_{\mathbb{C}}, \leqslant \rangle$ is totally ordered and $s_y - \varepsilon < s_y$ we have by [theorem: 3.68] 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

$$(s_y - \varepsilon)^2 < y \tag{10.56}$$

Using [theorems: 10.14] on [eqs: 10.55, 10.56] we have that

$$-(s_y + \varepsilon)^2 < -y < -(s_y - \varepsilon)^2 \tag{10.57}$$

Adding [eq: 10.54] to [eq: 10.57] gives

$$(s_y - \varepsilon)^2 - (s_y + \varepsilon)^2 < s_y^2 - y < (s_y + \varepsilon)^2 - (s_y - \varepsilon)^2 \tag{10.58}$$

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: 10.58] gives

$$\forall \varepsilon \in \mathbb{R}_{\mathbb{C}} \text{ with } 0 < \varepsilon < s_y \text{ we have } -4 \cdot \varepsilon \cdot s_y^2 < s_y^2 - y < 4 \cdot \varepsilon \cdot s_y \tag{10.59}$$

or using [theorems: 10.14]

$$\forall \varepsilon \in \mathbb{R}_{\mathbb{C}} \text{ with } 0 < \varepsilon < s_y \text{ we have } -4 \cdot \varepsilon \cdot s_y^2 < y - s_y^2 < 4 \cdot \varepsilon \cdot s_y \tag{10.60}$$

Now as $0 < s_y$ we have by [theorem: 10.27] a $\varepsilon_0 \in \mathbb{R}_{\mathbb{C}}$ such that

$$0 < \varepsilon_0 < s_y \tag{10.61}$$

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: 10.60] 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: 10.59] 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 10.56. (Square Root) *Using the previous theorem [theorem: 10.55] 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 10.57. 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 10.58. $\sqrt{0}=0$ 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 10.59. $2\in\mathbb{R}_0^{+}$so $\sqrt{2}$ exist but $\sqrt{2}\notin\mathbb{Q}_{\mathbb{C}}$ so that $\sqrt{2}\in\mathbb{R}_{\mathbb{C}}\backslash\mathbb{Q}_{\mathbb{C}}$ or $\sqrt{2}$ is a irrational number.

Proof. Assume that $\sqrt{2}\in\mathbb{Q}_{\mathbb{C}}$ then as by [theorem: 10.5] $i_{\mathbb{Q}\rightarrow\mathbb{C}}\colon\langle\mathbb{Q},+,\cdot\rangle\rightarrow\langle\mathbb{Q}_{\mathbb{C}},+,\cdot\rangle$ 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 10.60. *$\sqrt{.}\colon\mathbb{R}_0^{+}\rightarrow\mathbb{R}_0^{+}$ 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 10.61. *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: 10.56] we have $x=\sqrt{x^2}=\sqrt{a}$

$x<0$. Then $0<-x$ and $(-x)^2=x^2=a$ so that by [definition: 10.56] 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 $-\sqrt{a}\leqslant x<\sqrt{a}\Rightarrow-\sqrt{a}\leqslant x\leqslant\sqrt{a}$.

$\Leftarrow$. 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.

$\Rightarrow$. 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 $-\sqrt{a} < x < \sqrt{a}$.

$\Leftarrow$. 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 10.62. *If $x \in \mathbb{R}_{\mathbb{C}}$ then $x \leqslant \sqrt{x^2}$*

Proof. Using [theorem: 10.14] we have that $0 \leqslant x^2 = x^2$, so by [theorem: 10.61] we have either:

$x = \sqrt{x^2}$. Then trivial $x \leqslant \sqrt{x^2}$

$x = -\sqrt{x^2}$. Then $0 \leqslant \sqrt{x^2} = -x$ so that by [theorem: 10.14] $x \leqslant -0 = 0$ which as $0 \leqslant \sqrt{x^2}$ proves that $x \leqslant \sqrt{x^2}$. □

Theorem 10.63. *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 10.64. *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 10.65. *$i \in \mathbb{C}$ is defined as $i = (0,1)$ so that as $0 \neq 1$ we have that $i \in \mathbb{C} \backslash \mathbb{R}_{\mathbb{C}}$.*

Theorem 10.66. *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

$$\begin{array}{rcl} x + i \cdot y &=& (x', 0) + (0, 1) \cdot (y', 0) \\ &=& (x', 0) + (0 \cdot y' - 1 \cdot 0, 0 \cdot 0 + 1 \cdot y') \\ &=& (x', 0) + (0, y') \\ &=& (x', y') \\ &=& z \end{array}$$

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

$$\begin{array}{rcl} (z', y') &=& z \\ &=& u + i \cdot v \\ &=& (u', 0) + (0, 1) \cdot (v', 0) \\ &=& (u', 0) + (0 \cdot v' - 1 \cdot 0, 0 \cdot 0 + 1 \cdot v') \\ &=& (u', 0) + (0, v') \\ &=& (u', v') \end{array}$$

so that $x' = u'$ and $y' = v'$ hence $x = (x', 0) = (u', 0) = u$ and $y = (y', 0) = (v', 0) = v$ proving uniqueness. □

Note 10.67. 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 10.68. (Real & Imaginary part) *The above theorem allows us to define the following functions*

$$\mathrm{Re}\colon \mathbb{C} \rightarrow \mathbb{R}_{\mathbb{C}} \textit{ is defined by } \mathrm{Re}(x + i \cdot y) = x$$

and

$$\mathrm{Img}\colon \mathbb{C} \rightarrow \mathbb{R}_{\mathbb{C}} \textit{ by } \mathrm{Img}(x + i \cdot y) = y$$

So $\forall z \in \mathbb{C}$ *we have that*

$$z = \mathrm{Re}(z) + i \cdot \mathrm{Img}(z)$$

Theorem 10.69. *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* $\mathrm{Img}(z_1 + z_2) = \mathrm{Img}(z_1) + \mathrm{Img}(z_2)$

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.

$$\begin{array}{rcl} \mathrm{Re}(z_1 + z_2) &=& \mathrm{Re}((\mathrm{Re}(z_1) + i \cdot \mathrm{Img}(z_1)) + (\mathrm{Re}(z_2) + i \cdot \mathrm{Img}(z_2))) \\ &=& \mathrm{Re}((\mathrm{Re}(z_1) + \mathrm{Re}(z_2)) + i \cdot (\mathrm{Img}(z_1) + \mathrm{Img}(z_2))) \\ &=& \mathrm{Re}(z_1) + \mathrm{Re}(z_2) \end{array}$$

4.

$$\begin{array}{rcl}\mathrm{Img}(z_1+z_2) & = & \mathrm{Img}((\mathrm{Re}(z_1)+i\cdot\mathrm{Img}(z_1))+(\mathrm{Re}(z_2)+i\cdot\mathrm{Img}(z_2)))\\ & = & \mathrm{Img}((\mathrm{Re}(z_1)+\mathrm{Re}(z_2))+i\cdot(\mathrm{Img}(z_1)+\mathrm{Img}(z_2)))\\ & = & \mathrm{Img}(z_1)+\mathrm{Img}(z_2)\end{array}$$

□

10.5.2 Norm on $\mathbb{C}$

Definition 10.70. (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 10.71. *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* $\overline{z_1+z_2}=\overline{z_{\backslash}}+\overline{z_2}$
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. $\forall z\in\mathbb{C}\backslash\{0\}$ *we have* $\overline{z^{-1}}=(\bar{z})^{-1}$
6. $\forall z_1\in\mathbb{C}$, $\forall z_2\in\mathbb{C}\backslash\{0\}$ *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* $\overline{-z}=-\bar{z}$

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.

$$\begin{array}{rcl}\overline{(\bar{z})} & = & \overline{\mathrm{Re}(z)-i\cdot\mathrm{Img}(z)}\\ & = & \overline{(\mathrm{Re}(z)+i\cdot(-\mathrm{Img}(z)))}\\ & = & \mathrm{Re}(z)-i\cdot(-\mathrm{Img}(z))\\ & = & \mathrm{Re}(z)+i\cdot\mathrm{Img}(z)\\ & = & z\end{array}$$

3.

$$\begin{array}{rcl}\overline{z_1+z_2} & = & \mathrm{Re}(z_1+z_2)-i\cdot\mathrm{Img}(z_1+z_2)\\ & \underset{[\text{theorem: }10.69]}{=} & \mathrm{Re}(z_1)+\mathrm{Re}(z_2)-i\cdot(\mathrm{Img}(z_1)+\mathrm{Img}(z_2))\\ & = & (\mathrm{Re}(z_1)-i\cdot\mathrm{Img}(z_1))+(\mathrm{Re}(z_2)-i\cdot\mathrm{Img}(z_2))\\ & = & \overline{z_1}+\overline{z_2}\end{array}$$

4.

$$\begin{array}{rcl}\overline{z_1\cdot z_2} & = & \overline{(\mathrm{Re}(z_1)+i\cdot\mathrm{Img}(z_1))\cdot(\mathrm{Re}(z_2)+i\cdot\mathrm{Img}(z_2))}\\ & = & \overline{\mathrm{Re}(z_1)\cdot\mathrm{Re}(z_2)-\mathrm{Img}(z_1)\cdot\mathrm{Img}(z_2)+i\cdot(\mathrm{Re}(z_1)\cdot\mathrm{Img}(z_2)+\mathrm{Img}(z_1)\cdot\mathrm{Re}(z_2))}\\ & = & \mathrm{Re}(z_1)\cdot\mathrm{Re}(z_2)-\mathrm{Img}(z_1)\cdot\mathrm{Img}(z_2)-i\cdot(\mathrm{Re}(z_1)\cdot\mathrm{Img}(z_2)+\mathrm{Img}(z_1)\cdot\mathrm{Re}(z_2))\\ & = & (\mathrm{Re}(z_1)-i\cdot\mathrm{Img}(z_1))\cdot(\mathrm{Re}(z_2)-i\cdot\mathrm{Img}(z_2))\\ & = & \overline{z_1}\cdot\overline{z_2}\end{array}$$

5. We have $\overline{z^{-1}}\cdot\overline{z}=\overline{z}\cdot\overline{z^{-1}}\underset{(4)}{=}\overline{z\cdot z^{-1}}=\overline{1}=\overline{1+i\cdot 0}=1-i\cdot 0=1$ proving that $(\overline{z})^{-1}=\overline{z^{-1}}$

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 $\overline{z}=\mathrm{Re}(z)-i\cdot\mathrm{Im}(z)\underset{\text{[theorem: 10.69]}}{=}\mathrm{Re}(z)-i\cdot 0=\mathrm{Re}(z)\underset{\text{[theorem: 10.69]}}{=}z$

   ⇐. If $z=\overline{z}$ 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: 10.69]

8.

$$\begin{aligned}z\cdot\overline{z} &= (\mathrm{Re}(z)+i\cdot\mathrm{Img}(z))\cdot(\mathrm{Re}(z)-i\cdot\mathrm{Img}(z))\\ &= \mathrm{Re}(z)\cdot\mathrm{Re}(z)-i\cdot\mathrm{Re}(z)\cdot\mathrm{Img}(z)+i\cdot\mathrm{Img}(z)\cdot\mathrm{Re}(z)+i\cdot(-i)\cdot\mathrm{Img}(z)\cdot\mathrm{Img}(z)\\ &= \mathrm{Re}(z)^2+\mathrm{Img}(z)^2\end{aligned}$$

   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 $z\cdot\overline{z}\in\mathbb{R}^+_{0,\mathbb{C}}$.

9. $z+\overline{z}=(\mathrm{Re}(z)+i\cdot\mathrm{Img}(z))+(\mathrm{Re}(z)-i\cdot\mathrm{Img}(z))=2\cdot\mathrm{Re}(z)$

10.

$$\begin{aligned}\overline{-z} &= \overline{-(\mathrm{Re}(z)+i\cdot\mathrm{Img}(z))}\\ &= \overline{-\mathrm{Re}(z)+i\cdot(-\mathrm{Img}(z))}\\ &= -\mathrm{Re}(z)-i\cdot(-\mathrm{Img}(z))\\ &= -(\mathrm{Re}(z)-i\cdot\mathrm{Img}(z))\\ &= -\overline{z}\end{aligned}$$

□

Using the above theorem we have that $\forall z\in\mathbb{C}$ $z\cdot\overline{z}\in\mathbb{R}^+_{0,\mathbb{C}}$ so that $\sqrt{z\cdot\overline{z}}\in\mathbb{R}^+_{0,\mathbb{C}}$ is well defined, hence the following definition makes sense.

Definition 10.72. (complex norm)

$$||:\mathbb{C}\rightarrow\mathbb{R}^+_{0,\mathbb{C}}\textit{ is defined by }|z|=\sqrt{z\cdot\overline{z}}\underset{\textit{[theorem: 10.71]}}{=}\sqrt{\mathrm{Re}(z)^2+\mathrm{Img}(z)^2}$$

Theorem 10.73. *$\forall z,z'\in\mathbb{C}$ we have*

1. $|z\cdot z'|=|z|\cdot|z'|$
2. $|\overline{z}|=|z|$
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*

$$|z|=\begin{cases} z\text{ if }0\leqslant z\\ -z\text{ if }z<0\end{cases}$$

Proof.

1.

$$\begin{aligned}
|z\cdot z'| &= \sqrt{(z\cdot z')\cdot\overline{(z\cdot z')}}\\
&\underset{[\text{theorem: }10.71]}{=} \sqrt{(z\cdot z')\cdot(\overline{z}\cdot\overline{z'})}\\
&= \sqrt{(z\cdot\overline{z})\cdot(z'\cdot\overline{z'})}\\
&\underset{[\text{theorem: }10.63]}{=} \sqrt{z\cdot\overline{z}}\cdot\sqrt{z'\cdot\overline{z'}}\\
&= |z|\cdot|z'|
\end{aligned}$$

2. $|\overline{z}|=\sqrt{\overline{z}\cdot\overline{(\overline{z})}}\underset{[\text{theorem: }10.71]}{=}\sqrt{\overline{z}\cdot z}=\sqrt{z\cdot\overline{z}}=|z|$

3. Using [corollary: 10.62] 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: 10.60] $\sqrt{.}$ 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

$$\begin{aligned}
|z+z'|^2 &\underset{(5)}{=} (z+z')\cdot\overline{(z+z')}\\
&\underset{[\text{theorem: }10.71]}{=} (z+z')\cdot(\overline{z}+\overline{z'})\\
&= z\cdot\overline{z}+z\cdot\overline{z'}+z'\cdot\overline{z}+z'\cdot\overline{z'}\\
&\underset{(4)}{=} |z|^2+|z'|^2+z\cdot\overline{z'}+z'\cdot\overline{z}\\
&\underset{[\text{theorem: }10.71]}{=} |z|^2+|z'|^2+z\cdot\overline{z'}+\overline{z\cdot\overline{(z')}}\\
&\underset{[\text{theorem: }10.71]}{=} |z|^2+|z'|^2+2\cdot\mathrm{Re}(z\cdot\overline{z'})\\
&\leqslant_{(3)} |z|^2+|z'|^2+2\cdot|z\cdot\overline{z'}|\\
&\underset{(1)}{=} |z|^2+|z'|^2+2\cdot|z|\cdot|\overline{z'}|\\
&= (|z|+|z'|)^2
\end{aligned}$$

which as $\sqrt{.}$ is increasing [see theorem: 10.60] and $0\leqslant|z+z'|,|z|+|z'|$ we have that

$$|z+z'|=\sqrt{|z+z'|^2}\leqslant\sqrt{(|z|+|z'|)^2}=|z|+|z'|$$

6. $|-z|=\sqrt{(-z)\cdot\overline{(-z)}}\underset{[\text{theorem: }10.71]}{=}\sqrt{(-z)\cdot(-\overline{z})}=\sqrt{z\cdot\overline{z}}=|z|$

7.

$$|-1|\underset{(6)}{=}|1|=\sqrt{1\cdot\overline{1}}\underset{1\in\mathbb{R}_{\mathbb{C}}\text{ and }[\text{theorem: }10.71]}{=}\sqrt{1\cdot1}=\sqrt{1}\underset{[\text{example: }10.58]}{=}1$$

and

$$|i|=\sqrt{i\cdot\overline{i}}\underset{[\text{theorem: }10.71]}{=}\sqrt{-(i\cdot i)}=\sqrt{1}=1$$

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 $\operatorname{Re}(x)=\operatorname{Img}(z)=0$ hence

$$|z|=\sqrt{\operatorname{Re}(z)^2+\operatorname{Img}(z)^2}=\sqrt{0^2+0^2}\underset{[\text{example: }10.58]}{=}0$$

⇐. If $|z|=0$ then $\sqrt{\operatorname{Re}(z)^2+\operatorname{Img}(z)^2}=0\underset{[\text{example: }10.58]}{=}\sqrt{0}$ so as $\sqrt{.}$ is injective it follows that $\operatorname{Re}(z)^2+\operatorname{Img}(z)^2=0$. Using [theorem: 10.14] we have that $0\leqslant\operatorname{Re}(z)^2,\operatorname{Img}(z)^2$ so that by [theorem: 10.14] again we have that

$$\operatorname{Re}(z)^2,\operatorname{Img}(z)^2\leqslant\operatorname{Re}(z)^2+\operatorname{Img}(z)^2$$

If now either $\operatorname{Re}(z)\neq 0$ or $\operatorname{Img}(z)\neq 0$ then by [theorem: 10.14] $0<\operatorname{Re}(z)$ or $0<\operatorname{Img}(z)$ which would give $0<\operatorname{Re}(z)^2+\operatorname{Img}(z)^2$ contradicting $\operatorname{Re}(z)^2+\operatorname{Img}(z)^2=0$, Hence we have that $\operatorname{Re}(z)=0=\operatorname{Img}(z)$ so that $z=\operatorname{Re}(z)+i\cdot\operatorname{Img}(z)=0$.

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 $\operatorname{Re}(z)\underset{[\text{theorem: }10.69]}{=}z\in\mathbb{R}^+_{0,\mathbb{C}}$ and $\operatorname{Img}(z)=0$ so that

$$z=\operatorname{Re}(z)=\sqrt{\operatorname{Re}(z)^2}=\sqrt{\operatorname{Re}(z)^2+\operatorname{Img}(z)^2}=|z|$$

proving that

$$\forall z\in\mathbb{R}^+_{0,\mathbb{C}}\text{ we have } z=|z| \tag{10.62}$$

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: 10.62] we have $z=|z|$

$z<0$. Then $0<-z$ so that by [eq: 10.62] we have $-z=|-z|\underset{(6)}{=}|z|$ proving $-z=|z|$. □

10.5.3 Finite sets

We define now a characterization of finite sets in terms of $\mathbb{N}_{0,\mathbb{C}}$.

Lemma 10.74. *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: 10.29] 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 10.75. *We have the following characterization of finite sets then*

$$I\text{ is finite}\quad\Leftrightarrow\quad\text{there exists a unique } k\in\mathbb{N}_{\mathbb{C},0}\text{ such that }\{1,\ldots,k\}\approx I$$

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: 10.74] 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 $m=(i_{\mathbb{N}_0\rightarrow\mathbb{C}})^{-1}(l)$ we have by [lemma: 10.74] that

$$\{1,\ldots,m\}\approx\{1,\ldots,i_{\mathbb{N}_0\rightarrow\mathbb{C}}(m)\}=\{1,\ldots,i_{\mathbb{N}_0\rightarrow\mathbb{C}}((i_{\mathbb{N}_0\rightarrow\mathbb{C}})^{-1}(l))\}=\{1,\ldots,l\}\approx I$$

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 $(i_{\mathbb{N}_0\rightarrow\mathbb{C}})^{-1}(k)=n=m=(i_{\mathbb{N}_0\rightarrow\mathbb{C}})^{-1}(l)$ giving as $(i_{\mathbb{N}_0\rightarrow\mathbb{C}})^{-1}$ is a bijection that $k=l$.

$\Leftarrow$. 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 $n=(i_{\mathbb{N}_0\rightarrow\mathbb{C}})^{-1}(k)$ we have by [lemma: 10.74] that

$$\{1,\ldots,n\}\approx\{1,\ldots,i_{\mathbb{N}_0\rightarrow\mathbb{C}}(n)\}=\{1,\ldots,i_{\mathbb{N}_0\rightarrow\mathbb{C}}((i_{\mathbb{N}_0\rightarrow\mathbb{C}})^{-1}(k))\}=\{1,\ldots,k\}\approx I$$

proving that $\{1,\ldots,n\}\approx I$, hence by [theorem: 6.22] I is finite. □

Corollary 10.76. *Let I be a non empty set then*

$$I \textit{ is finite} \Leftrightarrow \textit{there exists a } \textbf{\textit{unique}}\ k\in\mathbb{N}_{\mathbb{C}} \textit{ such that } \{0,\ldots,k-1\}\approx I$$

Proof. Let $k\in\mathbb{N}_{\mathbb{C}}\Rightarrow k-1\in\mathbb{N}_{\mathbb{C},0}$. Define

$$\tau:\{0,\ldots,k-1\}\rightarrow\{1,\ldots,k\} \text{ by } \tau(i)=i+1$$

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

$$\text{If } k\in\mathbb{N}_{\mathbb{C}} \text{ then } \{0,\ldots,k-1\}\approx\{1,\ldots,k\}$$

$\Rightarrow$. If I is finite there exists by [theorem: 10.75] a $k\in\mathbb{N}_{\mathbb{C},0}$ such that $\{1,\ldots,k\}\approx I$ hence there exists a bijection

$$\alpha:\{1,\ldots,k\}\rightarrow I.$$

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: 10.75] 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: 10.75] it follows then that I is finite. □

Definition 10.77. *Let I be a finite set then* $\text{card}(I)$ *is defined as follows*

$$\text{card}(I)=i_{\mathbb{N}_0\rightarrow\mathbb{C}}(\#(I))$$

where $\#(I)$ is defined in [definition: 6.38]

Theorem 10.78. *If I is a finite set then we have that*

$$\text{card}(I) \textit{ is the unique } k\in\mathbb{N}_{0,\mathbb{C}} \textit{ such that } \{1,\ldots,k\}\approx I$$

In particular if $I=\varnothing$ then $\text{card}(I)=0$ *[as $\{1,\ldots,0\}=\varnothing$]*

Proof. By definition we have that $\#(I)=n$ where $n\approx I$ then

$$\text{card}(I)=i_{\mathbb{N}_0\rightarrow\mathbb{C}}(\#(I))=i_{\mathbb{N}_0\rightarrow\mathbb{C}}(n)$$

As by [theorem: 6.21] we have that $\{1,\ldots,n\}\approx n$, further by [lemma: 10.74]

$$\{1,\ldots,n\}\approx\{1,\ldots,i_{\mathbb{N}_0\rightarrow\mathbb{C}}(n)\}=\{1,\ldots,\operatorname{card}(I)\}$$

so that $I\approx\{1,\ldots,\operatorname{card}(I)\}$. □

Theorem 10.79. *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 $B\subseteq A$ 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 $\operatorname{card}(B)=i_{\mathbb{N}_9\rightarrow\mathbb{C}}(\#(B))<i_{\mathbb{N}_0\rightarrow\mathbb{C}}(\#(A))=\operatorname{card}(B)$ □

Theorem 10.80. *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
$$\begin{array}{rcl}\operatorname{card}\left(A\bigcup B\right) & = & i_{\mathbb{N}_0\rightarrow\mathbb{C}}\left(\#\left(A\bigcup B\right)\right)\\ & = & i_{\mathbb{N}_0\rightarrow\mathbb{C}}(\#(A)+\#(B))\\ & = & i_{\mathbb{N}_0\rightarrow\mathbb{C}}(\#(A))+i_{\mathbb{N}_0\rightarrow\mathbb{C}}(\#(B))\\ & = & \operatorname{card}(A)+\operatorname{card}(B)\end{array}$$
□

Corollary 10.81. *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: 10.80] we have

$$\operatorname{card}\left(A\bigcup\{a\}\right)=\operatorname{card}(A)+\operatorname{card}(\{a\})=\operatorname{card}(A)+1$$ □

Corollary 10.82. *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: 10.81] 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 10.83. *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: 6.44]. □

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 10.84. *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 10.85. *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 10.86. *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 10.87. $\overline{\leqslant} \in \overline{\mathbb{R}} \times \overline{\mathbb{R}}$ *is defined as follows*

$$\begin{aligned}\{(-\infty, -\infty), (-\infty, \infty), (\infty, \infty)\} & \bigcup \{(x, \infty) | x \in \mathbb{R}_{\mathbb{C}}\} \\ & \bigcup \{(-\infty, x) | x \in \mathbb{R}_{C}\} \\ & \bigcup \{(x, y) \in \mathbb{R}_{\mathbb{C}} \times \mathbb{R}_{\mathbb{C}} | x \leqslant y\}\end{aligned}$$

Note 10.88. 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 10.89. $\langle \overline{\mathbb{R}}, \overline{\leqslant} \rangle$ *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 $x \overline{\leqslant} z$

$\boldsymbol{x} = -\infty \wedge \boldsymbol{y} = \infty \wedge \boldsymbol{z} = \infty$**.** then $x \overline{\leqslant} z$

$\boldsymbol{x} \in \mathbb{R}_{\mathbb{C}} \wedge \boldsymbol{y} = \infty \wedge \boldsymbol{z} = \infty$**.** then $x \overline{\leqslant} z$

$\boldsymbol{x} = \infty \wedge \boldsymbol{y} = -\infty \wedge \boldsymbol{z} = \infty$**.** then as $\neg(x \overline{\leqslant} y)$ this cases does not count

$\boldsymbol{x} = -\infty \wedge \boldsymbol{y} = -\infty \wedge \boldsymbol{z} = \infty$**.** then $x \overline{\leqslant} z$

$\boldsymbol{x} \in \mathbb{R}_{\mathbb{C}} \wedge \boldsymbol{y} = -\infty \wedge \boldsymbol{z} = \infty$**.** then as $\neg(x \overline{\leqslant} y)$ this case does not count

$\boldsymbol{x} = \infty \wedge \boldsymbol{y} \in \mathbb{R}_{\mathbb{C}} \wedge \boldsymbol{z} = \infty$**.** then as $\neg(x \overline{\leqslant} y)$ this case does not count

$\boldsymbol{x} = -\infty \wedge \boldsymbol{y} \in \mathbb{R}_{\mathbb{C}} \wedge \boldsymbol{z} = \infty$**.** then $x \overline{\leqslant} z$

$\boldsymbol{x} \in \mathbb{R}_{\mathbb{C}} \wedge \boldsymbol{y} \in \mathbb{R}_{\mathbb{C}} \wedge \boldsymbol{z} = \infty$**.** then $x \overline{\leqslant} z$

$\boldsymbol{x} = \infty \wedge \boldsymbol{y} = \infty \wedge \boldsymbol{z} = -\infty$**.** then as $\neg(y \overline{\leqslant} x)$ this case does not count

$\boldsymbol{x} = -\infty \wedge \boldsymbol{y} = \infty \wedge \boldsymbol{z} = -\infty$**.** then as $\neg(y \overline{\leqslant} x)$ this case does not count

$\boldsymbol{x} \in \mathbb{R}_{\mathbb{C}} \wedge \boldsymbol{y} = \infty \wedge \boldsymbol{z} = -\infty$**.** then as $\neg(y \overline{\leqslant} x)$ this case does not count

$\boldsymbol{x} = \infty \wedge \boldsymbol{y} = -\infty \wedge \boldsymbol{z} = -\infty$**.** then as $\neg(x \overline{\leqslant} y)$ this case does not count

$\boldsymbol{x} = -\infty \wedge \boldsymbol{y} = -\infty \wedge \boldsymbol{z} = -\infty$**.** then $x \overline{\leqslant} z$

$\boldsymbol{x} \in \mathbb{R}_{\mathbb{C}} \wedge \boldsymbol{y} = -\infty \wedge \boldsymbol{z} = -\infty$**.** then as $\neg(x \overline{\leqslant} y)$ this case does not count

$\boldsymbol{x} = \infty \wedge \boldsymbol{y} \in \mathbb{R}_{\mathbb{C}} \wedge \boldsymbol{z} = -\infty$**.** then as $\neg(y \overline{\leqslant} z)$ this case does not count

$\boldsymbol{x} = -\infty \wedge \boldsymbol{y} \in \mathbb{R}_{\mathbb{C}} \wedge \boldsymbol{z} = -\infty$**.** then as $\neg(y \overline{\leqslant} z)$ this case does not count

$\boldsymbol{x} \in \mathbb{R}_{\mathbb{C}} \wedge \boldsymbol{y} \in \mathbb{R}_{\mathbb{C}} \wedge \boldsymbol{z} = -\infty$**.** then as $\neg(y \overline{\leqslant} z)$ this case does not count

$\boldsymbol{x} = \infty \wedge \boldsymbol{y} = \infty \wedge \boldsymbol{z} \in \mathbb{R}_{\mathbb{C}}$**.** then as $\neg(y \overline{\leqslant} z)$ this case does not count

$\boldsymbol{x} = -\infty \wedge \boldsymbol{y} = \infty \wedge \boldsymbol{z} \in \mathbb{R}_{\mathbb{C}}$**.** then as $\neg(y \overline{\leqslant} z)$ this case does not count

$\boldsymbol{x} \in \mathbb{R}_{\mathbb{C}} \wedge \boldsymbol{y} = \infty \wedge \boldsymbol{z} \in \mathbb{R}_{\mathbb{C}}$**.** then as $\neg(\mathrm{y}\overline{\leqslant}\mathrm{z})$ this case does not count

$\boldsymbol{x} = \infty \wedge \boldsymbol{y} = -\infty \wedge \boldsymbol{z} \in \mathbb{R}_{\mathbb{C}}$**.** then as $\neg(x \overline{\leqslant} y)$ this case does not count

$\boldsymbol{x} - \infty \wedge \boldsymbol{y} = -\infty \wedge \boldsymbol{z} \in \mathbb{R}_{\mathbb{C}}$**.** then $x \overline{\leqslant} z$

$\boldsymbol{x} \in \mathbb{R}_{\mathbb{C}} \wedge \boldsymbol{y} = -\infty \wedge \boldsymbol{z} \in \mathbb{R}_{\mathbb{C}}$**.** then as $\neg(\mathrm{x}\overline{\leqslant}\mathrm{y})$ this case does not count

$\boldsymbol{x} = \infty \wedge \boldsymbol{y} \in \mathbb{R}_{\mathbb{C}} \wedge \boldsymbol{z} \in \mathbb{R}_{\mathbb{C}}$**.** then as $\neg(\mathrm{x}\overline{\leqslant}\mathrm{y})$ this case does not count

$\boldsymbol{x} = -\infty \wedge \boldsymbol{y} \in \mathbb{R}_{\mathbb{C}} \wedge \boldsymbol{z} \in \mathbb{R}_{\mathbb{C}}$**.** then $x \overline{\leqslant} z$

$\boldsymbol{x} \in \mathbb{R}_{\mathbb{C}} \wedge \boldsymbol{y} \in \mathbb{R}_{\mathbb{C}} \wedge \boldsymbol{z} \in \mathbb{R}_{\mathbb{C}}$**.** 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 $x, y \in \overline{\mathbb{R}}$:

$\boldsymbol{x} = \infty \wedge \boldsymbol{y} = \infty$**.** then $x \overline{\leqslant} y$

$\boldsymbol{x} = -\infty \wedge \boldsymbol{x} = \infty$**.** then $x \overline{\leqslant} y$

$\boldsymbol{x} \in \mathbb{R}_{\mathbb{C}} \wedge \boldsymbol{y} = \infty$**.** then $x \overline{\leqslant} y$

$\boldsymbol{x} = \infty \wedge \boldsymbol{y} = -\infty$**.** then $y \overline{\leqslant} x$

$\boldsymbol{x} = -\infty \wedge \boldsymbol{y} = -\infty$**.** then $x \overline{\leqslant} y$

$\boldsymbol{x} \in \mathbb{R}_{\mathbb{C}} \wedge \boldsymbol{y} = -\infty$**.** then $y \overline{\leqslant} x$

$\boldsymbol{x} = \infty \wedge \boldsymbol{y} \in \mathbb{R}_{\mathbb{C}}$**.** then $y \overline{\leqslant} x$

$\boldsymbol{x} = -\infty \wedge \boldsymbol{y} \in \mathbb{R}_{\mathbb{C}}$**.** then $x \overline{\leqslant} y$

$\boldsymbol{x} \in \mathbb{R}_{\mathbb{C}} \wedge \boldsymbol{y}_{\mathbb{C}} \in \mathbb{R}$**.** 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 10.90. *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	$\mathbb{N}$	$\mathbb{N}_{0,\mathbb{C}}\backslash\{0\}$
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	$\mathbb{R}_0^+$	$\{x\in\mathbb{R}\vert 0\leqslant x\}$
Positive real numbers	$\mathbb{R}^+$	$\{x\in\mathbb{R}\vert 0<x\}$
Non positve numbers	$\mathbb{R}_0^-$	$\{x\in\mathbb{R}\vert x\leqslant 0\}$
Negative real numbers	$\mathbb{R}^-$	$\{x\in\mathbb{R}\vert x<0\}$
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 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 $\{x_i\}_{i\in\{0,\ldots,n\}}\subseteq A$ z finite family of elements of A then*

$$\sum_{i=0}^{(.)} x_i\colon\{0,\ldots,n\}\rightarrow A$$

is recursively defined [see definition: 10.36] 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 group or semi-group. However sometimes we want to deal with finite products instead of finite sums. In that case we use the $\prod$ symbol. So all the theorems, lemmas, proposition in this text are valid if we replace $\sum$ by $\prod$. In a ring for example we have a addition operator and a multiplicative operator so we use $\sum$ for addition and $\prod$ and for the composition operator $\circ$ we typical use $\prod$.

Example 11.3. Let $\{2^i\}_{i\in\{0,\ldots,3\}}$ then $\sum_{i=0}^{3}2^i$

Proof.

$$\begin{array}{rcl}\displaystyle\sum_{i=0}^{3}2^i & = & \left(\displaystyle\sum_{i=0}^{2}2^i\right)+2^3\\ & = & \left(\left(\displaystyle\sum_{i=0}^{1}2^i\right)+2^2\right)+2^3\\ & = & \left(\left(\left(\displaystyle\sum_{i=0}^{0}2^i\right)+2^1\right)+2^2\right)+2^3\\ & = & ((2^0+2^1)+2^2)+2^3\\ & = & ((1+2)+4)+8\\ & = & 15\end{array}$$

□

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

$$S=\left\{n\in\mathbb{N}_0|\forall\{x_i\}_{i\in\{0,\ldots,n\}}\subseteq A\vdash\forall i\in\{0,\ldots,n\}\, x_i=0 \text{ we have } \sum_{i=0}^{n} x_i=0\right\}$$

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 $\sum_{i=}^{0} x_i=x_0=0$ 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:

$0\in S$. 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 $0\in S$.

***n*∈*S*⇒*n*+1∈*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, $n\in\mathbb{N}_0$ 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:

0∈*S*. If $\{x_i\}_{i\in\{0\}}\subseteq A$ then $\sum_{i=0}^{0}(-x_i)=-x_i=-\sum_{i=0}^{0}x_i$ proving that $0\in S$.

***n*∈*S*⇒*n*+1∈*S*.** If $\{x_i\}_{i\in\{0,\ldots,n+1\}}\subseteq A$ then

$$\begin{aligned}\sum_{i=0}^{n+1}(-x_i) &= \left(\sum_{i=0}^{n}(-x_i)\right)+(-x_{n+1})\\ &\underset{n\in S}{=} -\sum_{i=0}^{n}x_i+(-x_{n+1})\\ &= -\left(\left(\sum_{i=0}^{n}x_i\right)+x_{n+1}\right)\\ &= -\sum_{i=0}^{n+1}x_i\end{aligned}$$

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

$$S=\left\{n\in\mathbb{N}_0|\forall\{x_i\}_{i\in\{0,\ldots,n\}}\text{ we have }\sum_{i=0}^{n}(\alpha_i\cdot x_i)=\alpha\cdot\sum_{i=0}^{n}x_i\right\}$$

then we have:

0∈*S*. As for $\{x_i\}_{i\in\{0,\ldots,0\}}$ $\sum_{i=0}^{0}(\alpha\cdot x_i)=\alpha\cdot x_0=\alpha\cdot\sum_{i=0}^{0}x_i$ we have that $0\in S$

***n* ∈ *S* ⇒ *n* + 1 ∈ *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: 2.104] for $\{x_{i,k}\}_{k\in\{0,\ldots,n_2\}}$ and $\{x_{k,j}\}_{k\in\{0,\ldots,n_1\}}$ where $i\in\{0,\ldots,n_1\}\wedge j\in\{0,\ldots,n_2\}$ using 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:

0 ∈ 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 $0\in S_n$.

***m* ∈ S_n ⇒ *m* + 1 ∈ 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 $\{x_{i.j}\}_{(i,j)\in\{0,\ldots,n_1\}\times\{0,\ldots,n_2\}}$ then as $n_1\in\mathbb{N}_0$ we have $n_2\in\mathbb{N}_0=S_{n_1}$ 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 $\{x_{i+1}-x_i\}_{i\in\{0,\ldots,n-1\}}$ we have*

$$\sum_{i=0}^{n-1}(x_{i+1}-x_i)=x_n-x_0$$

Proof. We prove this by induction so let

$$S=\left\{n\in\{1,\ldots,\infty\}|\text{If }\{x_i\}_{i\in\{0,\ldots,n\}}\subseteq A\text{ then }\sum_{i=0}^{n-1}(x_{i+1}-x_i)=x_n-x_0\right\}$$

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 β: $\{0,\ldots,m-n\}\rightarrow\{n,\ldots m\}$ is defined by*

$$\beta(i)=n+i$$

[see definition: 2.105]

Proof. We must of course prove that β: $\{0,\ldots,m\}\rightarrow\{n,\ldots,m\}$ is indeed a bijection:

injectivity. If $\beta(i)=\beta(j)$ then $n+i=n+j\Rightarrow i=j$

surjectivity. If $j\in\{n,\ldots,m\}$ then $n\leqslant j\wedge j\leqslant m$ so that $0\leqslant j-n\wedge j-n\leqslant m-n$. Take $i=j-n$ then $i\in\{0,\ldots,m-n\}$ and $\beta(i)=i+n=(j-n)+n=j$ proving surjectivity. □

Up to now we have only defined the finite sum of 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.*

Using the above definition we can rephrase [theorems: 11.4, 11.5, 11.6, 11.9, 11.7, 11.10]

Theorem 11.14. *If $\langle A,+\rangle$ is a semi-group, $n,m\in\mathbb{N}_0$ with $n<m$ and $\{x_i\}_{i\in\{n,\ldots,m\}}\subseteq A$ then we have*

$$\sum_{i=n}^{k} x_i = \begin{cases} x_n \text{ if } k=n \\ (\sum_{i=n}^{k-1} x_i) + x^k \text{if } k\in\{n+1,\ldots,m\} \end{cases}$$

Proof. Let $k\in\{n,\ldots,m\}$ then we have either:

$\boldsymbol{k=n}$**.** Then $\sum_{i=n}^{k} x_i = \sum_{i=n}^{n} x_i = \sum_{i=0}^{n-n} x_{n+i} = \sum_{i=0}^{0} x_{n+i} = x_{n+0} = x_n$

$\boldsymbol{k\in\{n+1,\ldots,m\}}$**.** Then as $n<m$ we have $m-n\in\mathbb{N}$ so that by [theorem: 11.4] for $l\in\{1,\ldots,m-n\}$

$$\sum_{i=0}^{l} x_{n+i} = \left(\sum_{i=0}^{l-1} x_{n+i}\right) + x_{n+l} \tag{11.1}$$

$$\sum_{i=n}^{k} x_i = \sum_{i=0}^{k-n} x_{n+i} \underset{k-n\in\{1,\ldots,m-n\} \text{ and [eq: 11.1]}}{=} \left(\sum_{i=0}^{(k-n)-1} x_{n+i}\right) + x_{n+(k-n)} = \left(\sum_{i=n}^{k-1} x_i\right) + x_k$$ □

Theorem 11.15. *If $\langle A,+\rangle$ is a semi-group, $n,m\in\mathbb{N}_0$ with $n<m$ 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

$$\sum_{i=n}^{m} x_i \underset{\text{def}}{=} \sum_{i=0}^{m-m} x_{n+i} \underset{\text{[theorem: 11.5]}}{=} 0$$ □

Theorem 11.16. *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

$$\sum_{i=n}^{m} (x_i+y_i) = \sum_{i=0}^{m-n} (x_{n+i}+y_{n+i}) = \sum_{i=0}^{m-n} x_i + \sum_{i=0}^{m-n} y_i = \sum_{i=n}^{m} x_i + \sum y_{i_{i=n}^{m}}$$ □

Theorem 11.17. *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 11.18. *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: 2.104] for $\{x_{i,k}\}_{k\in\{n_2,\ldots,m_2\}}$ and $\{x_{k,j}\}_{k\in\{n_1,\ldots,m_1\}}$ where $i\in\{n_1,\ldots,m_1\}\wedge j\in\{n_2,\ldots,m_2\}$ using in $\sum_{j=n_2}^{m_2} x_{i,j}$ and $\sum_{i=n_1}^{m_1} x_{i,j}$.

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 11.19. *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 11.20. *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*

$$\sum_{i=n}^{m-n-1}(x_{i+1}-x_i) = x_m - x_n$$

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

$$\sum_{i=n}^{m}(x_{i+1}-x_i) = \sum_{i=0}^{n}(x_{(n+i)+1}-x_{n+i}) = x_m - x_n$$

□

Theorem 11.21. *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 $\sum_{i=0}^{1} x_i = (\sum_{i=0}^{0} x_i) + x_1 = x_0 + x_1 \underset{[\text{theorem: } 11.14]}{=} x_0 + \sum_{i=1}^{1} x_i$ proving that $1\in S$.

$\boldsymbol{n\in S\Rightarrow n+1\in S.}$ Let $\{x_i\}_{i\in\{0,\ldots,n+1\}}\subseteq A$ then we have

$$\begin{aligned}\sum_{i=0}^{n+1} x_i &= \left(\sum_{i=0}^{n} x_i\right)+x_{n+1}\\ &\underset{n\in S}{=} \left(x_0+\sum_{i=1}^{n} x_i\right)+x_{n+1}\\ &\underset{\text{associativity}}{=} x_0+\left(\left(\sum_{i=1}^{n} x_i\right)+x_{n+1}\right)\\ &\underset{[\text{theorem: 11.14}]}{=} x_0+\sum_{i=1}^{n+1} x_i\end{aligned}$$

proving that $n+1\in S$

□

11.1.2 Associativity

Theorem 11.22. *Let $\langle A,+\rangle$ be a semi-group $n,m\in\mathbb{N}_0$ with $n<m$ 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 let $k\in\{0,\ldots,n-1\}$ 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:

$\boldsymbol{1\in S.}$ 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 $1\in S$.

$\boldsymbol{n\in S\Rightarrow n+1\in S.}$ 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:

$\boldsymbol{k=n.}$ Then

$$\begin{aligned}\sum_{i=0}^{n+1} x_i &= \left(\sum_{i=0}^{n} x_i\right)+x_{n+1}\\ &= \sum_{i=0}^{n} x_i+\sum_{i=0}^{0} x_{(n+1)+i}\\ &= \sum_{i=0}^{n} x_i+\sum_{i=n+1}^{n+1} x\\ &\underset{k=n}{=} \sum_{i=0}^{k} x_i+\sum_{i=k+1}^{n+1} x_i\end{aligned}$$

proving that $n+1\in S$

$\boldsymbol{k \in \{0, \ldots, n-1\}}$. 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 $n < m \Rightarrow m - n \in \mathbb{N}$ we have for $\{x_i\}_{i \in \{n, \ldots, m\}}$ that for $k \in \{n, \ldots, m-1\} \Rightarrow k - n \in \{0, \ldots, m-n\}$.

$$\begin{aligned}
\sum_{i=n}^{m} x_i &= \sum_{i=0}^{m-n} x_{n+i}\\
&\underset{k-n \in \{0, \ldots, m-n\} \text{and [eq: 11.2]}}{=} \sum_{i=0}^{k-n} x_{n+i} + \sum_{i=(k-n)+1}^{m-n} x_{n+i}\\
&= \sum_{i=n}^{k} x_i + \sum_{i=k+1}^{m} x_i
\end{aligned}$$

proving the theorem. □

Theorem 11.23. **(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 $\{x_i\}_{i\in\{b_0,\ldots,e_1\}}$ we have then

$$\begin{aligned}\sum_{i=b_0}^{e_1} x_i &\underset{[\text{theorem: } 11.22]}{=} \sum_{j=b_0}^{e_0} x_j+\sum_{j=e_0+1}^{e_1} x_j\\ &= \sum_{j=b_0}^{e_0} x_j+\sum_{j=b_1}^{e_1} x_j\\ &= \sum_{i=0}^{0}\left(\sum_{j=b_i}^{e_i} x_j\right)+\sum_{j=b_1}^{e_1} x_j\\ &= \sum_{i=0}^{1}\left(\sum_{j=b_i}^{e_i} x_j\right)\end{aligned}$$

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 $\{x_i\}_{i\in\{0,\ldots,n+1\}}\subseteq A$. 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 $b_i\leqslant e_i<e_i+1\leqslant b_{i+1}$ proving by [theorem: 10.31] that $b_0\leqslant b_n\leqslant e_n$ so that $e_n\in\{b_0,\ldots,e_{n+1}-1\}$. Next

$$\begin{aligned}\sum_{j=b_0}^{e_{n+1}} x_j &\underset{e_n\in\{b_0,\ldots,e_{n+1}-1\}\ [\text{theorem: } 11.22]}{=} \sum_{j=b_0}^{e_n} x_j+\sum_{j\in e_n+1}^{e_{n+1}} x_j\\ &\underset{e_n+1=b_{n+1}}{=} \sum_{j=b_0}^{e_n} x_j+\sum_{j=b_{n+1}}^{e_{n+1}} x_j\\ &\underset{n\in S}{=} \sum_{i=0}^{n}\left(\sum_{j=b_i}^{e_i} x_j\right)+\sum_{j=b_{n+1}}^{e_{n+1}} x_j\\ &= \sum_{i=0}^{n+1}\left(\sum_{j=b_i}^{e_i} x_j\right)\end{aligned}$$

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 11.24. (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*

$$S_I=\{\sigma\in I^I|\sigma: I\rightarrow I \textit{ is a bijection}\}$$

Theorem 11.25. 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: 2.74] 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: 2.71] $\sigma^{-1}: I \rightarrow I$ is a bijection so that $\sigma^{-1} \in S_I$. Hence $\sigma \circ \sigma^{-1} \underset{[\text{theorem: } 2.68]}{=} \mathrm{Id}_I \underset{[\text{theorem: } 2.68]}{=} \sigma^{-1} \circ \sigma$ □

Theorem 11.26. *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. For $\sigma_{|I \setminus \{i\}}: I \setminus \{i\} \rightarrow I \setminus \{i\}$ we have:

injectivity. If $k, l \in I \setminus \{i\}$ is such that $\sigma_{|I \setminus \{i\}}(k) = \sigma_{|I \setminus \{i\}}(l)$ then $\sigma(k) = \sigma_{|I \setminus \{i\}}(k) = \sigma_{|I \setminus \{i\}}(l) = \sigma(l)$ which as σ is a bijection proves that $k = l$.

surjectivity. If $k \in I \setminus \{i\}$ then as σ is a bijection there exist $l \in I$ such that $\sigma(l) = k$. If $l = i$ then $k = \sigma(i) = i$ contradicting $k \in I \setminus \{i\}$. Hence $l \in I \setminus \{i\}$ so that $\sigma_{|I \setminus \{i\}}(l) = \sigma(l) = k$ proving surjectivity. □

We define now a special type of permutation a transposition

Theorem 11.27. (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. First note that if $i = j$ then $\left(i \underset{I}{\leftrightarrow} j\right) = i = j = \left(i \underset{I}{\leftrightarrow} j\right)$ o this is indeed a function. Next we prove that it is a bijection.

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* ∧ *l* = *i*.** Then trivially $k = i$

***k* = *i* ∧ *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* ∧ *l* ≠ *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* ∧ *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* ∧ *l* = *j*.** Then trivially $k = l$.

***k* = *j* ∧ *l* ≠ *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* ≠ *i*, *j* ∧ *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* ≠ *i*, *j* ∧ *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* ≠ *i*, *j* ∧ *l* ≠ *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 $\left(i \underset{I}{\leftrightarrow} j\right)(k) = \left(i \underset{I}{\leftrightarrow} j\right)(i) = j = l$

$\boldsymbol{l \neq i, j}$. 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 11.28. *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:

   $\boldsymbol{k = i}$. Then $\left(i \underset{I}{\leftrightarrow} j\right)(k) = j \underset{i=j}{=} i = k = \mathrm{Id}_I(k)$

   $\boldsymbol{k = j}$. Then $\left(i \underset{I}{\leftrightarrow} j\right)(k) = j \underset{i=j}{=} i = k = \mathrm{Id}_I(k)$

   $\boldsymbol{k \neq i, j}$. Then $\left(i \underset{I}{\leftrightarrow} j\right)(k) = j \underset{i=j}{=} i = k = \mathrm{Id}_I(k)$

2. If $k \in I$ then we have either:

   $\boldsymbol{k = i}$. Then $\left(i \underset{I}{\leftrightarrow} j\right)\left(\left(i \underset{I}{\longleftrightarrow} j\right)(k)\right) = \left(i \underset{I}{\leftrightarrow} j\right)(j) = i = k$

   $\boldsymbol{k = j}$. Then $\left(i \underset{I}{\leftrightarrow} j\right)\left(\left(i \underset{I}{\longleftrightarrow} j\right)(k)\right) = \left(i \underset{I}{\leftrightarrow} j\right)(i) = j = k$

   $\boldsymbol{k \neq i, j}$. Then $\left(i \underset{I}{\leftrightarrow} j\right)\left(\left(i \underset{I}{\longleftrightarrow} j\right)(k)\right) = k$

3. If $\in I$ then we have either:

   $\boldsymbol{k = i}$. Then $\left(i \underset{I}{\leftrightarrow} j\right)(k) = j = (j \leftrightarrow i)(k)$

   $\boldsymbol{k = j}$. Then $\left(i \underset{I}{\leftrightarrow} j\right)(k) = i = (j \leftrightarrow i)(k)$

   $\boldsymbol{k \neq i, j}$. $\left(i \underset{I}{\leftrightarrow} j\right)(k) = k = (j \leftrightarrow i)(k)$

4. If $k \in J$ then we have either:

   $\boldsymbol{k = i}$. Then $\left(i \underset{J}{\leftrightarrow} j\right)(k) = j = \left(i \underset{I}{\leftrightarrow} j\right)(k)$

   $\boldsymbol{k = j}$. Then $\left(i \underset{J}{\leftrightarrow} j\right)(k) = i = \left(i \underset{I}{\leftrightarrow} j\right)(k)$

   $\boldsymbol{k \neq i, j}$. Then $\left(i \underset{J}{\leftrightarrow} j\right)(k) = k = \left(i \underset{I}{\leftrightarrow} j\right)(k)$ □

Lemma 11.29. *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 $\left(\left(n \underset{\{0, \ldots n\}}{\leftrightarrow} k\right)(i)\right) = n+1$ 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 $\gamma_\sigma(i) \neq n+1 \Rightarrow \gamma_\sigma(i) \in \{0, \ldots, n\}$.

***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

$$\gamma_\sigma\colon\{0,\ldots,n\}\to\{0,\ldots,n\}$$

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:

***r* = *n* ∧ *s* = *n*.** Then trivially $r=s$.

***r* = *n* ∧ *s* ≠ *n*.** Then $\sigma(n+1)=\gamma_\sigma(r)=\gamma_\sigma(s)=\sigma\left(\left(n \underset{\{0,\ldots n\}}{\leftrightarrow} k\right)(n)\right)=\sigma(n)$ giving the contradiction $n+1=n$ so this case never occurs.

***r* ≠ *n* ∧ *s* = *n*.** Then $\sigma(n+1)=\gamma_\sigma(s)=\gamma_\sigma(r)=\sigma\left(\left(n \underset{\{0,\ldots n\}}{\leftrightarrow} k\right)(n)\right)=\sigma(n)$ giving the contradiction $n+1=n$ so this case never occurs.

***r* ≠ *n* ∧ *s* ≠ *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 l we have then either:

***s* = *n* + 1.** Then for $n\in\{0,\ldots,n\}$ we have $\gamma_\sigma(n)=\sigma(n+1)=\sigma(s)=r$ so that $\gamma_\sigma(n)=r$.

***s* = *n*.** If $k=n$ then $r=\sigma(s)=\sigma(n)=\sigma(k)=n+1\Rightarrow r=n+1$ contradicting $r\in\{0,\ldots,n\}$. So we must have that $k\neq n$ or $k\in\{0,\ldots,n-1\}$, hence

$$\gamma_\sigma(k)=\sigma\left(\left(n \underset{\{0,\ldots n\}}{\leftrightarrow} k\right)(k)\right)=\sigma(n)=\sigma(s)=r$$

proving that $\gamma_\sigma(k)=r$.

***s* ∈ {0, …, *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

$$\gamma_\sigma(s)=\sigma\left(\left(n \underset{\{0,\ldots n\}}{\leftrightarrow} k\right)(s)\right)=\sigma(s)=r$$

So in all cases we found a $l\in\{0,\ldots,n\}$ such that $\gamma_\sigma(l)=r$ proving surjectivity. □

Theorem 11.30. (Commutativity) *Let $\langle A,+\rangle$ be a Abelian semi-group, $n\in\mathbb{N}_0$, $\{x_i\}_{i\in\{0,\ldots,n\}}$ 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

$$S=\left\{n\in\mathbb{N}_0|\forall\{x_i\}_{i\in\{0,\ldots,n\}}\text{ and }\forall\sigma\in S_{\{0,\ldots,n\}}\text{ we have }\sum_{i=0}^n x_i=\sum_{i=0}^n x_{\sigma(i)}\right\}$$

then we have:

0 ∈ *S*. Let $\{x_i\}_{i\in\{0\}}\subseteq A$ and $\sigma\in S_{\{0\}}\Rightarrow\sigma=\mathrm{Id}_{\{0\}}$ then we have

$$\sum_{i=0}^0 x_i=x_0=x_{\mathrm{Id}_{\{0\}}(0)}=x_{\sigma(0)}=\sum_{i=0}^n x_i$$

proving that $0 \in S$.

***n* ∈ *S* ⇒ *n* + 1 ∈ *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 σ we have either:

***σ*(1) = 0.** Then as $\sigma: \{0,1\} \rightarrow \{0,1\}$ is a bijection we must have $\sigma(0) = 1$, so

$$\sum_{i=0}^{1} x_i = x_0 + x_1 = x_1 + x_0 = x_{\sigma(0)} + x_{\sigma(1)} = \sum_{i=0}^{1} x_{\sigma(i)}$$

***σ*(1) = 1.** Then as $\sigma: \{0,1\} \rightarrow \{0,1\}$ 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:

***σ*(*n* + 1) = *n* + 1.** Then by [theorem: 11.26] we have that $\sigma_{|\{0,\ldots,n\}} \in S_{\{0,\ldots,n\}}$ which as $n \in S$ proves that

$$\sum_{i=0}^{n} x_i \underset{n \in S}{=} \sum_{i=0}^{n} x_{\sigma(i)} \tag{11.3}$$

So

$$\begin{aligned} \sum_{i=0}^{n+1} x_i &= \left(\sum_{i=0}^{n} x_i\right) + x_{n+1} \\ &\underset{[\text{theorem: } 11.3]}{=} \left(\sum_{i=0}^{n} x_{\sigma(i)}\right) + x_{n+1} \\ &= \left(\sum_{i=0}^{n} x_{\sigma(i)}\right) + x_{\sigma(n+1)} \\ &= \sum_{i=0}^{n+1} x_{\sigma(i)} \end{aligned}$$

proving that $n+1 \in S$.

***σ*(*n* + 1) ∈ {0, . . . , *n*}.** Take then $k = \sigma^{-1}(n+1) \in \{0,\ldots,n\}$ then we have by [lemma: 11.29] 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:

$$\begin{aligned} \sum_{i=0}^{n} x_i &= \sum_{i=0}^{n} x_{\gamma_\sigma(i)} \\ &\underset{1<n+1=0\leqslant n-1}{=} \sum_{i=0}^{(n-1)+1} x_{\gamma_\sigma(i)} \\ &= \left(\sum_{i=0}^{n-1} x_{\gamma\sigma}(i)\right) + x_{\gamma_\sigma(n+1)} \\ &\underset{11.4}{=} \left(\sum_{i=0}^{n-1} x_{\sigma\left(\left(n \underset{\{0,\ldots,n\}}{\leftrightarrow} k\right)(i)\right)}\right) + x_{\sigma(n+1)} \end{aligned} \tag{11.5}$$

Further as $\left(n \underset{\{0,\ldots,n\}}{\leftrightarrow} k\right) \in S_{\{0,\ldots,n\}}$ we have

$$\begin{aligned}
\sum_{i=}^{n+1} x_{\sigma(i)} &= \\
\left(\sum_{i=0}^{n} x_{\sigma(i)}\right) + x_{\sigma(n+1)} &\underset{n\in S}{=} \\
\left(\sum_{i=0}^{n} x_{\sigma\left(\left(n \underset{\{0,\ldots,n\}}{\leftrightarrow}\right)(i)\right)}\right) + x_{\sigma(n+1)} &= \\
\left(\left(\sum_{i=0}^{n-1} x_{\sigma\left(\left(n \underset{\{0,\ldots,n\}}{\leftrightarrow}\right)(i)\right)}\right) + x_{\sigma\left(\left(n \underset{\{0,\ldots,n\}}{\leftrightarrow} k\right)(n)\right)}\right) + x_{\sigma(n+1)} &= \\
\left(\sum_{i=0}^{n-1} x_{\sigma\left(\left(n \underset{\{0,\ldots,n\}}{\leftrightarrow}\right)(i)\right)}\right) + \left(x_{\sigma\left(\left(n \underset{\{0,\ldots,n\}}{\leftrightarrow} k\right)(n)\right)} + x_{\sigma(n+1)}\right) &= \\
\left(\sum_{i=0}^{n-1} x_{\sigma\left(\left(n \underset{\{0,\ldots,n\}}{\leftrightarrow}\right)(i)\right)}\right) + \left(x_{\sigma(n+1)} + x_{\sigma\left(\left(n \underset{\{0,\ldots,n\}}{\leftrightarrow} k\right)(n)\right)}\right) &= \\
\left(\left(\sum_{i=0}^{n-1} x_{\sigma\left(\left(n \underset{\{0,\ldots,n\}}{\leftrightarrow}\right)(i)\right)}\right) + x_{\sigma(n+1)}\right) + x_{\sigma\left(\left(n \underset{\{0,\ldots,n\}}{\leftrightarrow} k\right)(n)\right)} &\underset{\text{[eq: 11.5]}}{=} \\
\left(\sum_{i=0}^{n} x_i\right) + x_{\sigma\left(\left(n \underset{\{0,\ldots,n\}}{\leftrightarrow} k\right)(n)\right)} &= \\
\left(\sum_{i=0}^{n} x_i\right) + x_{\sigma(k)} &\underset{k=\sigma^{-1}(n+1)}{=} \\
\left(\sum_{i=0}^{n} x_i\right) + x_{n+1} &= \\
\sum_{i=0}^{n+1} x_i &=
\end{aligned}$$

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 11.31. *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: 10.76] 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*

$$\sum_{i\in I} x_i = \sum_{i=0}^{n-1} x_{\beta(i)}$$

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

$$\sum_{i=0}^{n-1} x_{\beta(i)} \underset{\text{[theorem: 11.30]}}{=} \sum_{i=0}^{n-1} x_{\beta(\beta^{-1}(\gamma(i)))} = \sum_{i=0}^{n-1} x_{\gamma(i)}$$ □

Example 11.32. 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 11.33. *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

$$\sum_{i\in m-n} x_i = \sum_{i=0}^{m-n} x_{\sigma(i)} = \sum_{i=0}^{m-n} x_{n+i} \underset{\text{def}}{=} \sum_{i=n}^{m} x_i$$

□

Theorem 11.34. *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* = ∅.** Then as $\sigma: J\rightarrow I$ is a bijection we have $J = 0$ so that $\sum_{i\in I} x_i = 0 = \sum_{j\in J} x_{\sigma(j)}$

***I* ≠ ∅.** Then there exists a $n\in\mathbb{N}$ such that $\beta: \{0,\ldots,n-1\}\rightarrow I$ such that

$$\sum_{i\in I} x_i = \sum_{i=0}^{n-1} x_{\beta(i)}$$

So $\sigma^{-1}\circ\beta: \{0,\ldots,n-1\}\rightarrow J$ is a bijection so that

$$\sum_{j\in J} x_{\sigma(j)} = \sum_{i=0}^{n-1} x_{\sigma(\sigma^{-1}(\beta(i)))} = \sum_{i=0}^{n-1} x_{\beta(i)} = \sum_{i\in I} x_i$$ □

Using this more general definition of a finite sum we can rewrite [theorems: 11.15, 11.6, 11.8, 11.9 and 11.7].

Theorem 11.35. *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* = ∅.** Then $\sum_{i\in I} x_i = \varnothing$

***I* ≠ ∅.** 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 $\forall i\in\{0,\ldots,n-1\}$ $\beta(i)\in I$ 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 11.36. *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* = ∅.** 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* ≠ ∅.** Then there exists a $n\in\mathbb{N}$ and a bijection β: $\{0,\ldots,n-1\}\rightarrow I$ 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 11.37. *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 β: $\{0,\ldots,n-1\}\rightarrow I$ such that

$$\sum_{i\in I}(\alpha\cdot x_i)=\sum_{i=0}^{n-1}(\alpha\cdot x_{\beta(i)})\underset{[\text{theorem: }11.8]}{=}\alpha\cdot\sum_{i=0}^{n-1}x_{\beta(i)}=\alpha\cdot\sum_{i\in I}x_i$$ □

Theorem 11.38. *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 β: $\{0,\ldots,n-1\}\rightarrow I$ 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 11.39. *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

$$\sum_{i\in I}\left(\sum_{j\in J} x_{i,j}\right)=0 \underset{\text{[theorem: 11.35]}}{=} \sum_{j\in J} 0=\sum_{j\in J}\left(\sum_{i\in I} x_{i,j}\right)$$

***J* = ∅.** Then we have

$$\sum_{i\in I}\left(\sum_{j\in J} x_{i,j}\right)=\sum_{i\in I} 0 \underset{\text{[theorem: 11.35]}}{=} 0=\sum_{j\in J}\left(\sum_{i\in I} x_{i,j}\right)$$

***I* ≠ ∅ ∧ *J* ≠ ∅.** Then there exists $n, m\in\mathbb{N}$ and bijections

$$\alpha\colon\{0,\ldots,n-1\}\to I,\ \beta\colon\{0,\ldots,m-1\}\to J$$

such that

$$\forall j\in J \text{ we have } \sum_{i\in I} x_{i,j}=\sum_{i=0}^{n-1} x_{\alpha(i),j} \text{ and } \forall i\in I \text{ we have } \sum_{j\in J} x_{i,j}=\sum_{j=0}^{m-1} x_{i,\beta(j)} \tag{11.6}$$

so that

$$\begin{aligned}
\sum_{i\in I}\left(\sum_{j\in J} x_{i,j}\right) &\underset{\text{[eq: 11.6]}}{=} \sum_{i\in I}\left(\sum_{j=0}^{m-1} x_{i,\beta(j)}\right)\\
&= \sum_{i=0}^{n-1}\left(\sum_{j=0}^{m-1} x_{\alpha(i),\beta(j)}\right)\\
&\underset{\text{[theorem: 11.9]}}{=} \sum_{j=0}^{m-1}\left(\sum_{i=0}^{n-1} x_{\alpha(i),\beta(j)}\right)\\
&= \sum_{j\in J}\left(\sum_{i-0}^{n-1} x_{\alpha(i),j}\right)\\
&\underset{\text{[eq: 11.6]}}{=} \sum_{j\in J}\left(\sum_{i\in I} x_{i,j}\right)
\end{aligned}$$

□

We have the equivalent of [theorem: 11.22]

Lemma 11.40. *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* = ∅.** 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* = ∅.** 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* ≠ ∅.** Then there exists $n, m\in\mathbb{N}$ such $\alpha\colon\{0,\ldots,n-1\}\to N$ and $\beta\colon\{0,\ldots,m-1\}\to M$ are bijections and

$$\sum_{i\in N} x_i=\sum_{i=0}^{n-1} x_{\alpha(i)} \text{ and } \sum_{i\in M} x_i=\sum_{i=0}^{m-1} x_{\beta(i)} \tag{11.7}$$

Define now

$$\gamma:\{0,\ldots,n+m-1\}\rightarrow N\bigcup M \text{ by } \gamma(i)=\begin{cases}\alpha(i) \text{ where } i\in\{0,\ldots,n-1\}\\ \beta(i-n) \text{ where } i\in\{n,\ldots,n+m-1\}\end{cases}$$

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 $n\leqslant i\leqslant n+m-1$ so that $\gamma(j)=\beta(j-n)=\beta(i)=y$

So $\gamma:\{0,\ldots,n+m-1\}\rightarrow N\bigcup M=I$ is a surjection hence

$$\begin{aligned}\sum_{i\in I} x_i &= \sum_{i=0}^{n+m-1} x_{\gamma(i)}\\ &\underset{[\text{theorem: } 11.22]}{=} \sum_{i=0}^{n-1} x_{\gamma(i)}+\sum_{i=m}^{n+m-1} x_{\gamma(i)}\\ &= \sum_{i=0}^{n-1} x_{\alpha(i)}+\sum_{i=m}^{n+m-1} x_{\beta(i-n)}\\ &= \sum_{i=0}^{n-1} x_{\alpha(i)}+\sum_{i=0}^{m-1} x_{\beta((i+n)-n)}\\ &= \sum_{i=0}^{n-1} x_{\alpha(i)}+\sum_{i=0}^{m-1} x_{\beta(i)}\\ &\underset{[\text{theorem: } 11.7]}{=} \sum_{i\in N} x_i+\sum_{i\in M} x_i\end{aligned}$$

□

We have now the equivalence of [theorem: 11.23]

Theorem 11.41. *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, 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

$$J \bigcap I_{n+1} = \left(\bigcup_{i \in \{0, \ldots, n\}} I_i\right) \bigcap I_{n+1} = \bigcup_{i \in \{0, \ldots, n\}} \left(I_i \bigcap I_{n+1}\right) = \bigcup_{i \in \{0, \ldots, n\}} \varnothing = \varnothing$$

proving

$$\bigcup_{j \in \{0, \ldots, n+1\}} I_j = J \bigcup I_{n+1},\ J, I_{n+1} \text{ are finite and } J \bigcap I_{n+1} = \varnothing \tag{11.8}$$

So we have

$$\begin{aligned}
\sum_{i \in \bigcup_{i \in \{0, \ldots n\}} I_i} x_i &\underset{\text{[eq: 11.8] and [lemma: 11.40]}}{=} \sum_{i \in J} x_i + \sum_{i \in I_{n+1}} x_i \\
&= \sum_{i \in \bigcup_{j \in \{0, \ldots, n\}} I_j} x_i + \sum_{i \in I_{n+1}} x_i \\
&\underset{n \in S}{=} \sum_{i=0}^{n} \left(\sum_{j \in I_i} x_j\right) + \sum_{i \in I_{n+1}} x_i \\
&= \sum_{i=0}^{n+1} \left(\sum_{j \in I_i} x_i\right)
\end{aligned}$$

□

Corollary 11.42. *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 $I_i \bigcap I_j = \varnothing$ $\{x_i\}_{i \in \bigcup_{j \in I} I_j}$ 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 $\beta: \{0, \ldots, n-1\} \rightarrow I$. Using [theorem: 2.113] we have that

$$\bigcup_{j \in I} I_j = \bigcup_{j \in \{0, \ldots, n-1\}} I_{\beta(j)}$$

so that

$$\begin{aligned}
\sum_{i \in \bigcup_{j \in I} I_j} x_i &= \sum_{i \in \bigcup_{j \in \{0, \ldots, n-1\}} I_{\beta(j)}} x_i \\
&\underset{\text{[theorem: 11.41]}}{=} \sum_{i=0}^{n-1} \left(\sum_{j \in I_{\beta(i)}} x_j\right) \\
&= \sum_{i \in I} \left(\sum_{j \in I_i} x_j\right)
\end{aligned}$$

□

Corollary 11.43. *Let $\langle A,+\rangle$ be a Abelian semi-group, I,J finite sets and $\{x_{i,j}\}_{(i,j)\in I\times J}$ then*

$$\sum_{(i,j)\in I\times J} x_{i,j}=\sum_{i\in I}\left(\sum_{j\in J} x_{i,j}\right)\underset{\text{[theorem: 11.39]}}{=}\sum_{j\in J}\left(\sum_{i\in I} x_{i,j}\right)$$

Proof. Given $i\in I$ define $J_i=\{i\}\times J$ then s $I=\bigcup_{i\in I}\{i\}$ we have by [theorem: 2.127] 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

$$\forall i,j\in I \text{ with } i\neq j \text{ we have } \{i\}\times J\bigcap\{j\}\times J=\varnothing$$

So we can apply [theorem: 11.42] giving

$$\begin{aligned}\sum_{(i,j)\in I\times J} x_{i,j} &= \sum_{(i,j)\in\bigcup_{k\in I}J_k} x_{i,j}\\ &\underset{\text{[theorem: 11.42]}}{=} \sum_{k\in I}\left(\sum_{(i,j)\in J_k} x_{i,j}\right)\\ &\underset{(i,j)\in\{k\}\times J\Leftrightarrow j=k}{=} \sum_{k\in I}\left(\sum_{(k,j)\in J_k} x_{k,j}\right)\end{aligned}\tag{11.9}$$

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: 11.34]

$$\sum_{j\in J} x_{k,j}=\sum_{j\in J} x_{\beta_k(j)}\underset{\text{[theorem: 11.34]}}{=}\sum_{(k,j)\in J_k} x_{k,j}$$

which combined with [eq: 11.9] gives

$$\sum_{(i,j)\in I\times J} x_{i,j}=\sum_{k\in I}\left(\sum_{j\in J} x_{k,j}\right)$$

□

Theorem 11.44. *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: 6.80] 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 fro 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 $\boldsymbol{\in S \Rightarrow l+1 \in S}$**.** Let $\{x_j\}_{j\in\{0,\ldots,l+1\}}\subseteq A^n$ then we have

$$\begin{aligned}\left(\sum_{j=0}^{l+1} x_j\right)_i &= \left(\left(\sum_{j=0}^{l} x_j\right)+x_{l+1}\right)_i\\ &\underset{[\text{theorem: } 6.80]}{=} \left(\sum_{j=0}^{l} x_j\right)_i+(x_{l+1})_i\\ &\underset{l\in S}{=} \left(\sum_{j=0}^{l} (x_j)_i\right)+(x_{l+1})_j\\ &= \sum_{j=0}^{l+1} (x_j)_i\end{aligned}$$

proving that $l+1\in S$.

By mathematical induction it follows then that

$$\forall l\in\mathbb{N}_0 \text{ we have for } \{x_j\}_{j\in\{1,\ldots,l\}}\subseteq A^n \text{ and } i\in\{1,...,n\} \text{ that } \left(\sum_{j=0}^{l} x_j\right)_i=\sum_{j=0}^{l}(x_j)_i \tag{11.10}$$

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

$$\left(\sum_{j=k}^{l} x_j\right)_i=\left(\sum_{j=0}^{l-k} x_{j+k}\right)_i \underset{[\text{eq: } 11.10]}{=} \sum_{j=0}^{l-k}(x_{j+k})_i=\sum_{j=k}^{l}(x_j)_i$$

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:

   $\boldsymbol{I=\varnothing}$**.** Then

   $$\left(\sum_{j\in I} x_j\right)_i=\left(\sum_{j\in\varnothing} x_j\right)_i=0_i \underset{[\text{theorem: } 6.80]}{=} \sum_{j\in\varnothing}(x_j)_i$$

   $\boldsymbol{I\neq\varnothing}$**.** Then there exists a $l\in\mathbb{N}_0$ and bijection $\beta\colon\{0,\ldots,l\}\rightarrow I$ then

   $$\left(\sum_{j\in I} x_j\right)_i \underset{\text{def}}{=} \left(\sum_{j=0}^{l} x_{\beta(j)}\right)_i \underset{[\text{theorem: } 11.10]}{=} \sum_{j=0}^{l}(x_{\beta(j)})_i=\sum_{j\in I}(x_j)_i$$

□

Theorem 11.45. *For the ring $\langle\mathbb{R},\cdot,+\rangle$ we have*

1. *If $n\in\mathbb{N}_0$ and $\{\alpha_i\}_{i\in\{0,\ldots,n\}}\subseteq\mathbb{R}$ 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 $\{\alpha_i\}_{i\in\{n,\ldots,m\}}\subseteq\mathbb{R}$ 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 $\{\alpha_i\}_{i\in I}\subseteq\mathbb{R}$ defined by $\alpha_i=\alpha$ we have*
   a. $\sum_{i\in I}\alpha_i=\operatorname{card}(I)\cdot\alpha$
   b. $\prod_{i\in I}\alpha_i=\alpha^{\operatorname{card}(I)}$

Proof.

1.

   a. We prove this by induction, so let

   $$S=\left\{n\in\mathbb{N}_0|\text{If }\{\alpha_i\}_{i\in\{0,\ldots,n\}}\text{ is defined by }\alpha_i=\alpha\text{ then }\sum_{i=0}^{n}\alpha_i=(n+1)\cdot\alpha\right\}$$

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 $\{\alpha_i\}_{i\in\{1,\ldots,n+1\}}\subseteq\mathbb{R}$ be defined by $\alpha_i=\alpha$. Then

$$\begin{aligned}\sum_{i=0}^{n+1}\alpha_i &= \sum_{i=0}^{n}\alpha_i+\alpha_{n+1}\\ &= \sum_{i=0}^{n}\alpha_i+\alpha\\ &\underset{n\in S}{=} (n+1)\cdot\alpha+\alpha\\ &= ((n+1)+1)\cdot\alpha\end{aligned}$$

proving that $n+1\in S$

b. We prove this by induction, so let

$$S=\left\{n\in\mathbb{N}_0|\text{If }\{\alpha_i\}_{i\in\{0,\ldots,n\}}\text{ is defined by }\alpha_i=\alpha\text{ then }\prod_{i=0}^{n}\alpha_i=(n+1)\cdot\alpha\right\}$$

then we have:

$0\in S$. Then $\prod_{i=0}^{0}\alpha_i\underset{[\text{theorem: }11.32]}{=}\alpha_0=\alpha=\alpha^1=\alpha^{(0+1)}$ 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

$$\begin{aligned}\prod_{i=0}^{n+1}\alpha_i &= \left(\prod_{i=0}^{n}\alpha_i\right)\cdot\alpha_{n+1}\\ &= \left(\prod_{i=0}^{n}\alpha_i\right)\cdot\alpha\\ &\underset{n\in S}{=} \alpha^{n+1}\cdot\alpha\\ &= \alpha^{(n+1)+1}\end{aligned}$$

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 $\{0,\ldots,\mathrm{card}(I)-1\}\approx\{1,\ldots,\mathrm{card}(i)\}$ so that there exists a bijection $\beta:\{0,\ldots,\mathrm{card}(I)-1\}\to I$ so that

$$\sum_{i\in I}\alpha_i=\sum_{0}^{\mathrm{card}(I)-1}\alpha_{\beta(i)}\underset{(1.a)}{=}(\mathrm{card}(I)-1+1)\cdot\alpha=\mathrm{card}(I)\cdot\alpha$$

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 $\{0,\ldots,\mathrm{card}(I)-1\}\approx\{1,\ldots,\mathrm{card}(i)\}$ so that there exists a bijection $\beta:\{0,\ldots,\mathrm{card}(I)-1\}\to I$ so that

$$\prod_{i\in I}\alpha_i=\prod_{0}^{\mathrm{card}(I)-1}\alpha_{\beta(i)}\underset{(2.a)}{=}\alpha^{(\mathrm{card}(I)-1+1)}=\alpha^{\mathrm{card}(I)}$$ □

Theorem 11.46. *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 $\beta_j = \begin{cases} x & \text{if } j = i \\ \alpha_j & \text{if } j \in I \setminus \{i\} \end{cases}$.

3. *Let $n, m \in \mathbb{N}_0$ with $n \leqslant m$, $i \in \{n,...,m\}$, $x, y \in F$ 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 $\{\beta_j\}_{j \in \{0,\ldots,n\}}$ is defined by $\beta_j = \begin{cases} x & \text{if } j = i \\ \alpha_j & \text{if } j \in \{0, \ldots, n\} \setminus \{i\} \end{cases}$.

Proof.

1. We have

$$\begin{aligned}
\prod_{j \in I} \alpha_j &\underset{[\text{theorem: 11.41}]}{=} \left(\prod_{j \in I \setminus \{i\}} \alpha_j\right) \cdot \left(\prod_{j \in \{i\}} \alpha_j\right) \\
&\underset{[\text{theorem: 11.32}]}{=} \left(\prod_{j \in I \setminus \{i\}} \alpha_j\right) \cdot \alpha_i \\
&= \left(\prod_{j \in I \setminus \{i\}} \alpha_j\right) \cdot (x + y) \\
&= \left(\prod_{j \in I \setminus \{i\}} \alpha_j\right) \cdot x + \left(\prod_{j \in I \setminus \{i\}} \alpha_j\right) \cdot y \\
&= \left(\prod_{j \in I \setminus \{i\}} \beta_j\right) \cdot \beta_{n+1} + \left(\prod_{j \in I \setminus \{i\}} \gamma_j\right) \cdot \gamma_{n+1} \\
&\underset{[\text{theorem: 11.32}]}{=} \left(\prod_{j \in I \setminus \{i\}} \beta_j\right) \cdot \left(\prod_{j \in \{i\}} \beta_j\right) + \left(\prod_{j \in I \setminus \{i\}} \gamma_j\right) \cdot \left(\prod_{j \in \{i\}} \gamma_j\right) \\
&\underset{[\text{theorem: 11.41}]}{=} \prod_{j \in I} \beta_j + \prod_{j \in I} \gamma_j
\end{aligned}$$

2. We have

$$\begin{aligned}
\prod_{j\in I}\alpha_j &\underset{[\text{theorem: }11.41]}{=} \left(\prod_{j\in I\setminus\{i\}}\alpha_j\right)\cdot\left(\prod_{j\in\{i\}}\alpha_j\right)\\
&\underset{[\text{theorem: }11.32]}{=} \left(\prod_{j\in I\setminus\{i\}}\alpha_j\right)\cdot\alpha_i\\
&= \left(\prod_{j\in I\setminus\{i\}}\alpha_j\right)\cdot(\beta\cdot x)\\
&= \beta\cdot\left(\left(\prod_{j\in I\setminus\{i\}}\alpha_j\right)\cdot x\right)\\
&= \beta\cdot\left(\left(\prod_{j\in I\setminus\{i\}}\beta_j\right)\cdot\beta_i\right)\\
&\underset{[\text{theorem: }11.32]}{=} \beta\cdot\left(\left(\prod_{j\in I\setminus\{i\}}\beta_j\right)\cdot\left(\prod_{j\in\{i\}}\beta_j\right)\right)\\
&\underset{[\text{theorem: }11.41]}{=} \beta\cdot\prod_{j\in I}\beta_j
\end{aligned}$$

3. We have

$$\begin{aligned}
\prod_{i=n}^{m}\alpha_i &\underset{[\text{theorem: }11.33]}{=} \prod_{i\in\{n,\ldots,m\}}\alpha_i\\
&\underset{(1)}{=} \prod_{i\in\{n,\ldots,m\}}\beta_i+\prod_{i\in\{n,\ldots,m\}}\gamma_i\\
&\underset{[\text{theorem: }11.33]}{=} \prod_{i=n}^{m}\beta_i+\prod_{i=n}^{m}\gamma_i
\end{aligned}$$

4. We have

$$\begin{aligned}
\prod_{i=n}^{m}\alpha_i &\underset{[\text{theorem: }11.33]}{=} \prod_{i\in\{n,\ldots,m\}}\alpha_i\\
&\underset{(2)}{=} \beta\cdot\prod_{i\in\{n,\ldots,m\}}\beta_i\\
&\underset{[\text{theorem: }11.33]}{=} \beta\cdot\prod_{i=n}^{m}\beta_i
\end{aligned}$$

□

Theorem 11.47. *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 $\prod_{i=0}^{0} \alpha_i \underset{[\text{theorem: } 11.32]}{=} \alpha_0 \neq 0$ 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, $\alpha_{n+1} \neq 0$, we have, as a field is a integral domain [see: 4.58], 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:

$I = \varnothing$. Then $\prod_{i \in S} \alpha_i = \prod_{i \in \varnothing} \alpha_i = 1 \neq 0$.

$I \neq \varnothing$. Then there exists a bijection $\beta: \{0, \ldots, n-1\} \rightarrow I$ such that

$$\sum_{i \in I} \alpha_i = \sum_{i=0}^{n-1} \alpha_{\beta(i)} \neq 0 \text{ (by (1))}$$ □

11.2 Vector spaces

11.2.1 Definition

Definition 11.48. *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 i 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 11.49. 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. However in this book e use the convention of most books about mathematics and physics.

Theorem 11.50. *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 for $\alpha \in F \setminus \{0_f\}$ we have if $x \in V$ with $\alpha \odot x = 0_v$ that $x = 0_v$*
5. *If $x \in V \setminus \{0\}$ then we have if $\alpha \in F$ with $\alpha \cdot x = 0$ that $\alpha = 0$.*

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

$$0_v \underset{(3)}{=} \alpha^{-1} \odot 0_v \underset{(1)}{=} \alpha^{-1} \odot (\alpha \odot x) = (\alpha^{-1} \cdot \alpha) \odot x = 1 \odot x = x$$

5. Let $x \in V \setminus \{0\}$ and $\alpha \in F$ such that $\alpha \cdot x = 0$. Assume that $\alpha \neq 0$ then by (4) we have that $x = 0$ contradicting $x \in V \setminus \{0\}$, so we must have that $\alpha = 0$. □

Just as we have sub-groups, sub-rings and sub-fields we have sub-spaces of a vector space.

Definition 11.51. *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. *$\forall \alpha, \beta \in F$ and $x, y \in V$ we have that $(\alpha \odot x) \oplus (\beta \odot y) \in W$*

Theorem 11.52. *Let $\langle V, \oplus, \odot \rangle$ be a vector space over $\langle F, +, \cdot \rangle$ and $W \subseteq V$ a sub-space then $\langle V, \oplus_{|W \times W}, \odot_{|F \odot W} \rangle$ is a vector space over $\langle F, +, \cdot \rangle$.*

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

$$0_v \underset{[\text{theorem: } 11.50\ (1)]}{=} 0_f \odot w = (0_f + 0_f) \odot w = 0_f \odot w + 0 \odot w \in W$$

proving that $0_v \in W$. So $\forall x \in W$ 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 $-x \underset{[\text{theorem: } 11.50\ (2)]}{=} ((-1) \odot x) = ((-1) \odot x) \oplus 0_v \in W$ so that $-x \in W$. Hence $x \oplus_{|W \times W} (-x) \underset{\text{commutativity}}{=} (-x) \oplus_{|W \times W} x = (-x) \oplus x = 0_v$

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 $\alpha \odot_{|W \times W} (x \oplus_{|W \times W} y) = \alpha \odot (x \oplus y) = \alpha \odot x \oplus \alpha \odot y = \alpha \odot_{|W \times W} x \oplus_{|W \times W} \beta \odot_{|W \times W} y$
3. If $\alpha, \beta \in F$ and $x \in W$ then $(\alpha + \beta) \odot_{|W \times W} x = (\alpha + \beta) \odot x = \alpha \odot x \oplus \beta \odot x = \alpha \odot_{|W \times W} x \oplus_{|W \times W} \beta \odot_{|W \times W} x$
4. If $\alpha, \beta \in F$ and $x \in W$ then $(\alpha \cdot \beta) \odot_{|W \times W} x = (\alpha \cdot \beta) \odot x = \alpha \odot (\beta \odot x) = \alpha \odot_{|W \times W} (\beta \odot_{|W} x)$
5. If $x \in W$ then $1 \odot_{|F \times W} x = 1 \odot x = x$ □

Note 11.53. 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 11.54. 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. $\forall x, y, z \in \{e\}$ we have $x \oplus (y \oplus z) = e \oplus (e \oplus e) = e \oplus e = (e \oplus e) \oplus e = (x \oplus y) \oplus z$

neutral element. $\forall x \in \{0\}$ we have $x \oplus e = e \oplus e = e = e \oplus e = e \oplus x$ we have that e is the neutral element.

inverse element. $\forall x \in \{0\}$ we have $x \oplus e = e \oplus e = e$ so that e is the inverse element of x.

commutativity. We have trivial$\forall x, y \in \{0\}$ we have x@+y$= e \oplus e = y \oplus x$

For the remaining axioms we have

1. $\forall \alpha \in F$ and $x, y \in \{e\}$ we have
$$\alpha \odot (x + y) = \alpha \odot (e \oplus e) = \alpha \odot e = e = e \oplus e = \alpha \odot w \oplus \alpha \odot e = \alpha \odot x \oplus \alpha \odot y$$
2. $\forall \alpha, \beta \in F$ and $x \in \{e\}$ we have
$$(\alpha + \beta) \odot x = (\alpha + b) \odot e = e = e \oplus e = \alpha \odot e + \beta \odot e = \alpha \odot x + \beta \odot x$$
3. $\forall \alpha, \beta \in F$ and $x \in \{e\}$ we have $(\alpha \cdot \beta) \odot x = (\alpha \cdot \beta) \odot e = e = \alpha \odot e = \alpha \odot (b \odot e)$
4. $\forall x \in \{e\}$ we have $1 \odot x = 1 \odot e = e = x$ □

Every field is a vector spaces over itself.

Theorem 11.55. *If $\langle F, +, \cdot \rangle$ is a field then $\langle F, +, \cdot \rangle$ is a vector space over it self*

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 11.56. $\langle \mathbb{Q}, +, \cdot \rangle$ 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 11.57. $\langle \mathbb{C}, +, \cdot \rangle$ 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 11.58. *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 $\langle \mathrm{Rr}, +, \cdot \rangle$ then we say that $\langle \mathbb{V}, \oplus, \odot \rangle$ 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 11.59. (function space) *Let $\langle V, \oplus, \odot \rangle$ be a vector space over $\langle F, +, \cdot \rangle$, X a set then and define for $V^X = \{f | f: X \rightarrow V \text{ is a function}\}$ the following operations:*

$$\boxplus: V^X \times V^X \rightarrow V^X \textit{ defined by } f \boxplus g \text{ where } (f \boxplus g)(x) = f(x) \oplus g(y) \textit{ [pairwise addition]}$$

$$\boxdot: F \times V^X \rightarrow V^X \textit{ defined by } \alpha \boxdot f \textit{ where } (\alpha \boxdot f)(x) = \alpha \odot f(x)$$

then

$$\langle V^X, \boxplus, \boxdot \rangle \textit{ is a vector space over } \langle F, +, \cdot \rangle$$

where:

a) $C_o: X \rightarrow V$ is defined by $C_0(x) = e$ [the constant function [see example: 2.45] is the additive neutral element.

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 I have used different operator symbols for addition and multiplication. To simplify notation we use from now on always $+$ for addition and $\cdot$ for multiplication relaying on context to figure out which operator is associated with the symbols $+$ and $\cdot$. We also use 0 to denote the additive neutral element and $-x$ to note the inverse of x.

Referring to the power of a set [see definition: 6.77] we can

Theorem 11.60. *Let $\langle V,+,\cdot\rangle$ 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*

$$+\colon V^n\times V^n\rightarrow V^n \textit{ is defined by } (x_1,\ldots,x_n)+(y_1,\ldots,y_n)=(x_1+y_1,\ldots,x_n+y_n)$$

$$\cdot\colon F\times V^n\rightarrow V^n \textit{ is defined by } \alpha\cdot(x_1,\ldots,x_n)=(\alpha\cdot x_1,\ldots,\alpha\cdot x_n)$$

$$0=\underbrace{(0,\ldots,0)}_{n}\in V^n \textit{ is the additive neutral element in } V^n$$

$$\forall(x_1,\ldots,x_n)\in V^n \textit{ the additive negative is } (-x_1,\ldots,-x_n)$$

Proof. Note that using [theorem: 6.78] we have that

$$V^n=V^{\{1,\ldots,n\}}=\{f\,|\,f\colon\{1,\ldots,n\}\rightarrow V \text{ is a function}\}$$

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: 11.59] proves then that $\langle V^n,+,\cdot\rangle$ is a vector space. □

As a application of the above theorem we have that

Corollary 11.61. *Let $n\in\mathbb{N}$ and $\langle F,+,\cdot\rangle$ a field then $\langle F^n +,\cdot\rangle$ is a vector space over $\langle F,+,\cdot\rangle$*

Proof. This follows from [theorems: 11.55, 11.60] □

Using [example: 11.56, 11.57] and the above theorem [theorem: 11.60] we have then the following examples of vector spaces:

Example 11.62. 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 11.63. Note that in the proof of the previous theorem we use the fact that $V^n = \{f | f: \{1,...,n\} \rightarrow V$ 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: 6.73] and [theorem: 6.78] to understand the original definitions.

Definition 11.64. *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. $x + A = \{x + y | y \in A\} \underset{\text{commutativity}}{=} \{y + x | y \in A\} = A + x$
2. $A + B = \{x + y | x \in A \wedge y \in B\}$
3. $\alpha \cdot A = \{\alpha \cdot x | x \in A\}$
4. $G \cdot A = \{\gamma \cdot x | \gamma \in F \wedge x \in A\}$
5. $A - B = \{x - y | x \in A \wedge y \in B\}$

Theorem 11.65. *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. $$\begin{aligned} z \in (x + (y + A)) & \underset{(1)}{\Leftrightarrow} z - x \in y + A \\ & \underset{(1)}{\Leftrightarrow} (z - x) - y \in A \\ & \Leftrightarrow (z - (x + y)) \in A \\ & \Leftrightarrow z \in (x + y) + A \end{aligned}$$

□

Theorem 11.66. *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

$$\begin{aligned}((\alpha\cdot\beta)\cdot x)(i) &= (\alpha\cdot\beta)\cdot_i x(i)\\ &\underset{\langle V_i,+_i,\cdot_i\rangle\text{ is a vector space}}{=} \alpha\cdot_i(\beta\cdot_i x(i))\\ &= \alpha\cdot_i(\beta\cdot x)(i)\\ &= (\alpha\cdot(\beta\cdot x))(i)\\ &= (\alpha\cdot(\beta\cdot x))(i)\end{aligned}$$

   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 $1\cdot x=x$ □

11.3 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.

Theorem 11.67. *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 we have

$$\forall\{x_i\}_{i\in\{0,\ldots,n\}}V\text{ we have }\sum_{i=0}^{n}(\alpha\cdot x)=\alpha\cdot\sum_{i=0}^{n}x_i \tag{11.11}$$

Let now $n,m\in\mathbb{N}_0$ with $n\leqslant m$ then we have

$$\sum_{i=n}^{m}(\alpha\cdot x_i)=\sum_{i=0}^{m-n}(\alpha\cdot x_{n+i})\underset{[\text{eq: 11.11}]}{=}\alpha\cdot\sum_{i=0}^{m-n}x_{n+i}=\alpha\cdot\sum_{i=n}^{m}x_i$$

2. If I is finite and $\{x_i\}_{i\in I}$ with $\forall i\in I\ x_i=x$ then we have either:

$I=\varnothing$. Then

$$\sum_{i\in I}(\alpha\cdot x_i)=0=\alpha\cdot 0=\alpha\cdot\sum_{i\in I}x_i$$

$I\neq\varnothing$. Then there exist a $n\in\mathbb{N}$ and a bijection $\beta:\{0,\ldots,n-1\}$ such that

$$\sum_{i\in I}(\alpha\cdot x_i)=\sum_{i=0}^{n-1}(\alpha\cdot x_{\beta(i)})\underset{[\text{theorem: 11.11}]}{=}\alpha\cdot\sum_{i=0}^{n-1}x_{\beta(i)}=\alpha\cdot\sum_{i\in I}x_i$$ □

Theorem 11.68. *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:

$0\in S$. 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 $0\in S$.

$n\in S\Rightarrow n+1\in S$. 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

$$\forall n\in\mathbb{N}_0 \text{ and } \{\alpha_i\}_{i\in\{0,\ldots,n\}} \text{ we have } \sum_{i=0}^{n}(\alpha_i\cdot x)=\left(\sum_{i=0}^{n}\alpha_i\right)\cdot x \tag{11.12}$$

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

$$\sum_{i=n}^{m}(\alpha_i\cdot x)=\sum_{i=0}^{m-n}(\alpha_{n+i}\cdot x)\underset{[\text{eq: } 11.12]}{=}\left(\sum_{i=0}^{m-n}\alpha_i\right)\cdot x=\left(\sum_{i=n}^{m}\alpha_i\right)\cdot x$$

2. I is a finite set and $\{\alpha_i\}_{i\in I}\subseteq F$ then we have either:

$I=\varnothing$. Then

$$\sum_{i\in I}(\alpha_i\cdot x)=0=0\cdot x=\left(\sum_{i\in I}\alpha_i\right)$$

$I\neq\varnothing$. Then there exists a $n\in\mathbb{N}$ and a bijection $\beta:\{0,\ldots,n-1\}\rightarrow I$ such that

$$\sum_{i\in I}(\alpha_i\cdot x)=\sum_{i=0}^{n-1}(\alpha_{\beta(i)}\cdot x)\underset{[\text{theorem: } 11.12]}{=}\left(\sum_{i=0}^{n-1}\alpha_i\right)\cdot x=\left(\sum_{i\in I}\alpha_i\right)\cdot x$$

□

Theorem 11.69. *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

$$S=\left\{n\in\mathbb{N}_0|\forall\{x_i\}_{i\in\{0,\ldots,n\}} \text{ with } \forall i\in\{0,\ldots,n\}\, x_i=x \text{ we have } \sum_{i=0}^{n}x_i=(n+1)\cdot x\right\}$$

then we have:

$\mathbf{0}\in \boldsymbol{S}$. If $\{x_i\}_{i\in\{0\}}\subseteq V$ is such that $x_0=x$ then $\sum_{i=0}^{0} x_i=x=1\cdot x=(0+1)\cdot x$ proving that $0\in S$.

$\boldsymbol{n}\in \boldsymbol{S}\Rightarrow \boldsymbol{n}+\mathbf{1}\in \boldsymbol{S}$. 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

$$\forall\{x_i\}_{i\in\{0,\ldots,n\}} \text{ with } \forall i\in\{0,\ldots,n\}\, x_i=x \text{ we have } \sum_{i=0}^{n} x_i=(n+1)\cdot x \tag{11.13}$$

If now $n,m\in\mathbb{N}_0$ with $n\leqslant m$ then

$$\sum_{i=n}^{m} x_i=\sum_{i=0}^{m-n} x_{n+i}\underset{\text{[eq: 11.13]}}{=}(m-n)+1\cdot x$$

2. For I is finite and $\{x_i\}_{i\in I}$ with $\forall i\in I$ $x_i=x$ then we have either:

$\boldsymbol{I}=\varnothing$. Then

$$\sum_{i=I} x_i=0=0\cdot x=\operatorname{card}(I)\cdot x$$

$\boldsymbol{I}\neq\varnothing$. Then there exist a $n\in\mathbb{N}$ and a bijection $\beta\colon\{0,\ldots,n-1\}\rightarrow I\Rightarrow\{0,\ldots,n-1\}\approx I$ such that

$$\sum_{i\in I} x_i=\sum_{i=0}^{n-1} x_{\beta(i)}=(n-1+1)\cdot x=n\cdot x\underset{\text{[theorem: 10.78]}}{=}\operatorname{card}(I)\cdot x$$ □

Theorem 11.70. *Let $\langle V,+,\cdot\rangle$ be a vector space over $\langle F,+,\cdot\rangle$ and $\emptyset\neq W\subseteq V$ 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 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:

$\mathbf{0}\in \boldsymbol{S}$. 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 $0\in W$.

$\boldsymbol{n}\in \boldsymbol{S}\Rightarrow \boldsymbol{n}+\mathbf{1}\in \boldsymbol{S}$. Let $\{x_i\}_{i\in\{0,\ldots,n+1\}}$ then we have

$$\begin{aligned}\sum_{i=0}^{n+1} x_i &= \left(\sum_{i=0}^{n} x_i\right)+x_{n+1}\\ &= 1\cdot\left(\sum_{i=0}^{n} x_i\right)+1\cdot x_{n+1}\\ &\in W\,[\text{as } n\in S\Rightarrow\sum_{i=0}^{n} x_i\in S\wedge x_{n+1}\in S\end{aligned}$$

using induction we have then that

$$\forall n \in \mathbb{N}_0 \text{ and } \{x_i\}_{i \in \{0,\ldots,n\}} \subseteq W \text{ then } \sum_{i=0}^{n} x_i \in W \tag{11.14}$$

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

$$\sum_{i=n}^{m} x_i = \sum_{i=0}^{m-n} x_{n+i} \in_{[\text{eq: } 11.14]} W$$

2. Let I be finite then we have either:

   ***I* = ∅.** As $W \neq \varnothing$ we have by [theorem: 11.52] that $\langle W, +, \cdot\rangle$ is a vector space so that $0 \in W$. Hence $\sum_{i \in I} x_i = 0 \in W$.

   ***I* ≠ ∅.** As $I \neq$ there exists a $\in \mathbb{N}$ and a bijection $\beta: \{0, \ldots, n-1\} \rightarrow I$ such that

$$\sum_{i \in I} x_i = \sum_{i=0}^{n-1} x_{\beta(i)} \in W \text{ [see eq: 11.14]}$$

□

11.3.2 Linear (in)dependency

11.3.2.1 Span of a set

Definition 11.71. *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*

$$v = \sum_{i \in I} \alpha_i \cdot i$$

In the above the index set is not ordered, it will be usefully for induction arguments later, to use as index sets, sets of the form $\{1,\ldots,n\}$ [if $n=0$ then $\{1,\ldots,n\} = \varnothing$] and families $\{\alpha_i\}_{i \in \{1,\ldots,n\}} \subseteq V$ instead of a finite set of vectors. We have to be careful, for there could be $i, j \in \{1, \ldots, n\}$ with $i \neq j$ such that $v_i = v_j$ resulting in extra terms in the sum. To solve this we introduce that concept of ordered families and disjoint families.

Definition 11.72. (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**.*

Example 11.73. If $n = 0$ then $\{x_i\}_{i \in \{1,\ldots,n\}} \subseteq X$ is a ordered family where

$$\{x_i\}_{i \in \{1,\ldots,n\}} \subseteq X = \{\varnothing_i\}_{i \in \varnothing} \subseteq X$$

Proof. If $n = 0$ then $\{1, \ldots, n\} = \{1, \ldots, 0\} = \varnothing$ and we have by [example: 2.99] that

$$\{x_i\}_{i \in \{1,\ldots,n\}} \subseteq X = \{\varnothing_i\}_{i \in \varnothing} \subseteq X$$

□

Remember that a family $\{x_i\}_{i \in I} \subseteq X$ is actually the same as the function $x: I \rightarrow X$ [see definition: 2.97]. Then we have the concept of disjoint families where the defining function is a injection.

Definition 11.74. (Disjoint Family) *A family $\{x_i\}_{i \in I} \subseteq X$ is **disjoint** if $x: I \rightarrow X$ is a injection or equivalently, as $x_i \underset{\text{notation}}{=} x(i)$, $\forall i, j \in I$ with $x_i = x_j$ we have $i = j$.*

Theorem 11.75. *Let $n \in \mathbb{N}_0$, X be a set $\{x_i\}_{i \in \{1,\ldots,n\}} \subseteq X$ a ordered family then*

$$\{x_i\}_{i \in \{1,\ldots,n\}} \text{ is disjoint}$$
$$\Updownarrow$$
$$\operatorname{card}(\{x_i | i \in \{1, \ldots, n\}\}) = n$$

Proof.

- **$\Rightarrow$.** As the family $\{x_i\}_{i\in\{1,\ldots,n\}}\subseteq X$ is disjoint $x:\{1,...,n\}\rightarrow X$ is a injective function, hence $x:\{1,...,n\}\rightarrow x(\{1,...,n\})=\{x_{i|i\in\{1,\ldots,n\}}\}$ is a bijection so that $\{1,...,n\}\approx\{x_i|i\in\{1,...,n\}\}$. Hence

  $$\operatorname{card}(\{x_i|i\in\{1,\ldots,n\})=\operatorname{card}(\{1,\ldots,n\})=n$$

- **$\Leftarrow$.** As the family $\{x_i\}_{i\in\{1,\ldots,n\}}\subseteq X$ is represented by a function $x:\{1,\ldots,n\}\rightarrow X$. Using [theorem: 3.125] there exists a $J\subseteq\{1,\ldots,n\}$ such that

  $$x: J\rightarrow x(\{1,\ldots,n\})=\{x_i|i\in\{1,\ldots,n\}\}\text{ is a bijection}$$

  Hence $J\approx\{x_i|i\in\{1,\ldots,n\}\}$ so that

  $$\operatorname{card}(J)=\operatorname{card}(\{x_i|i\in\{1,\ldots,n\}\})=n$$

  Assume that $J\subset\{1,\ldots,n\}$ then by [theorem: 10.79] we have $\operatorname{card}(J)<\operatorname{card}(\{1,\ldots,n\})=n$ leading to the contradiction $n<n$. Hence we have that $J=\{1,...,n\}$ so that $x=x_{|J}$ proving that x is injective and thus that $\{x_i\}_{i\in\{1,\ldots,n\}}$ is disjoint. □

Theorem 11.76. *Let $\{v_i\}_{i\in I}\subseteq V$ be a disjoint family and $k\in I$ then*

$$\{v_i|i\in I\setminus\{k\}\}=\{v_i|i\in I\}\setminus\{v_k\}$$

Proof. If $v\in\{v_i|i\in I\setminus\{k\}\}$ then $\exists i\in I\setminus\{k\}$ such that $v=x_i$. Assume that $v=x_k$ then $x_i=x_k$ hence $i=k$ contradicting $i\in I\setminus\{k\}$ so $x\neq v_k$ which, as $i\in I$, proves that $x\in\{v_i|i\in I\}\setminus\{v_k\}$ or

$$\{v_i|i\in I\setminus\{k\}\}\subseteq\{v_i|i\in I\}\setminus\{v_k\}$$

On the other hand if $x\in\{v_i|i\in I\}\setminus\{v_k\}$ then $x\neq v_k$ and $\exists i\in I$ such that $x=v_i$, so we must have $i\neq k$, proving that $x=x_i\in\{v_i|i\in I\setminus\{k\}\}$. Hence

$$\{v_i|i\in I\}\setminus\{v_k\}\subseteq\{v_i|i\in I\setminus\{k\}\}$$ □

Theorem 11.77. *If $\{x_i\}_{i\in\{1,\ldots,n\}}\subseteq X$, $n\in\mathbb{N}_0$ is a ordered family then $\{x_i|i\in\{1,...,n\}\}$ is finite.*

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. □

Theorem 11.78. *Let X be a set, $n\in\mathbb{N}_0$, $\{x_i\}_{i\in\{1,\ldots,n\}}\subseteq X$ and $Y\subseteq\{x_i|i\in\{1,...,n\}\}$ then there exists a bijection $\beta:\{1,\ldots,\operatorname{card}(Y)\}\rightarrow J\subseteq\{1,\ldots,n\}$ such that $Y=\{x_{\beta(i)}|i\in\{1,\ldots,\operatorname{card}(Y)\}\}$ and $\{x_{\beta(i)}\}_{i\in\{1,\ldots,\operatorname{card}(Y)\}}\subseteq X$ is a distinct family.*

Proof. Let $x:\{1,\ldots,n\}\rightarrow X$ be the function defining the family $\{x_i\}_{i\in\{1,\ldots,n\}}\subseteq X$. Using a consequence of the Axiom of Choice [see theorem: 3.125] 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

$$x_{|J}\underset{[\text{theorem: }2.83]}{=}(x_{|I})_{|J}: J\rightarrow Y$$

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 $\mathrm{x}_{|I}$ it follows that $i=j$.

surjectivity. Let $y \in Y$ then, as $x_{|I}: I \rightarrow \{x_i | i \in \{1, \ldots, n\}\}$ 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 then $i \in (x_{|I})^{-1}(Y) = J$ 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: 6.42, 11.77] we have that J is finite with $\mathrm{card}(J) = \mathrm{card}(Y)$. Hence there exists a bijection

$$\beta: \{1, \ldots, \mathrm{card}(Y)\} \rightarrow J$$

Let $z \in \{x_{\beta(i)} | i \in \{1, \ldots, \mathrm{card}(Y)\}\}$ then there exist a $i \in \{1, \ldots, \mathrm{card}(Y)\}$ such that $z = x_{\beta(i)} = x(\beta(i)) \underset{\beta(i) \in J}{=} x_{|J}(\beta(i)) \in Y$ proving that

$$\{x_{\beta(i)} | i \in \{1, \ldots, \mathrm{card}(Y)\}\} \subseteq Y$$

Likewise if $z \in Y$ then, as $x_{|J}: J \rightarrow Y$ is a bijection, there exists a $j \in J$ such that $x_{|J}(j) = z$, further as $\beta: \{1, \ldots, \mathrm{card}(Y)\} \rightarrow J$ is a bijection there exist a $i \in \{1, \ldots, \mathrm{card}(Y)\}$ such that $j = \beta(i)$, hence $z = x_{|J}(\beta(i)) = x(\beta(i)) = x_{\beta(i)} \in \{x_{\beta(i)} | i \in \{1, \ldots, \mathrm{card}(Y)\}\}$. So we have also

$$Y \subseteq \{x_{\beta(i)} | i \in \{1, \ldots, \mathrm{card}(Y)\}\}$$

Hence

$$Y = \{x_{\beta(i)} | i \in \{1, \ldots, \mathrm{card}(Y)\}\}$$

Finally, let $i, j \in \{1,\ldots,\mathrm{card}(Y)\}$ such that $x_{\beta(i)} = x_{\beta(j)} \Rightarrow x(\beta(i) = x(\beta(j)$. As $i, j \in \{1,\ldots,\mathrm{card}(Y)\}$ we have $\beta(i), \beta(j) \in J \subseteq I$ so that $x_{|I}(\beta(i)) = x(\beta(i)) = x(\beta(j)) = x_{|I}(\beta(j))$. By injectivity of $x_{|I}$ it follows then that $\beta(i) = \beta(j)$ and by injectivity we have then $i = j$. Hence we have proved that

$$\{x_{\beta(i)}\}_{i \in \{1, \ldots, \mathrm{card}(Y)\}} \subseteq Y \text{ is a distinct family}$$ □

We have the following characterization of a finite set in terms of families.

Theorem 11.79. *Let X be a set then*

$$X \textit{ is finite}$$

$$\Downarrow$$

There exists a distinct ordered family $\{x_i\}_{i \in \{1, \ldots, n\}} \subseteq X$ such that $X = \{x_i | i \in \{1, \ldots, n\}\}$ where $n = \mathrm{card}(X)$

Proof. If X is finite we have by [theorems: 10.75,10.78] that there exists a $k \in \mathbb{N}_0$ and a bijection

$$x: \{1, \ldots, k\} \rightarrow X$$

so that $\{x_i\}_{i \in \{1, \ldots, k\}} \subseteq X$ is a ordered **disjoint** family. Further by surjectivity we have $\{x_i | i \in \{1, \ldots, k\}\} = x(\{1, \ldots, k\}) = X$. □

Theorem 11.80. *Let $\langle V, +, \cdot \rangle$ be a vector space over a field $\langle F, +, \cdot \rangle$, $W \subseteq V$ a finite set and $v \in V$ such that there exists a ordered family [not necessary distinct] $\{w_i\}_{i \in \{1, \ldots, n\}} \subseteq W$ and a $\{\alpha_i\}_{i \in \{1, \ldots, n\}} \subseteq F$ such that*

$$v = \sum_{i \in \{1, \ldots, n\}} \alpha_i \cdot w_i$$

then there exists a $\{\beta_w\}_{w \in \{w_i | i \in \{1, \ldots, n\}\}} \subseteq F$ such that

$$v = \sum_{w \in \{w_i | i \in \{1, \ldots, n\}\}} \beta_w \cdot w$$

Proof. By [theorem: 11.77] we have that

$$\{w_i|i\in\{1,\ldots,n\}\} \text{ is finite} \tag{11.15}$$

Let $w\in\{w_i|i\in\{1,\ldots,n\}\}$ and define $I_w=\{i\in\{1,\ldots,n\}|w_i=w\}$. Then if $j\in\{1,\ldots,n\}$ we have $w_j\in\{w_i|i\in\{1,\ldots,n\}\}$ so that for $w=w_j$ we have $j\in I_w$ proving that $\{1,\ldots,n\}\subseteq\bigcup_{w\in\{w_{i0|i\in\{1,\ldots,n\}}\}}I_w$. Further as $I_w\subseteq\{1,\ldots,n\}$ it follows that $\bigcup_{w\in\{w_{i0|i\in\{1,\ldots,n\}}\}}I_w\subseteq\{1,\ldots,n\}$. Hence we have

$$\{1,\ldots,n\}=\bigcup_{w\in\{w_i|i\in\{1,\ldots,n\}\}} I_w \tag{11.16}$$

Let $w,u\in\{w_i|i\in\{1,\ldots,n\}\}$ with $w\neq$vu. If $k\in I_w\bigcap I_u$ then $w=w_k=u$ contradicting $w\neq u$, hence we have

$$\forall w,u\in\{w_i|i\in\{1,\ldots,n\}\} \text{ with } w\neq u \text{ we have } I_w\bigcap I_u=\varnothing \tag{11.17}$$

Define now

$$\{\beta_w\}_{w\in\{w_i|i\in\{1,\ldots,n\}\}}\subseteq F \text{ by } \beta_w=\sum_{i\in I_w}\alpha_i \tag{11.18}$$

Then we have

$$\begin{aligned}
v &= \sum_{i\in\{1,\ldots,n\}} \alpha_i\cdot w_i\\
&\underset{\text{[theorem: 11.42] and [eqs: 11.15, 11.16, 11.17]}}{=} \sum_{w\in\{w_i|i\in\{1,\ldots,n\}\}}\left(\sum_{i\in I_w}\alpha_i\cdot w_i\right)\\
&\underset{i\in I_w\Rightarrow w_i=w}{=} \sum_{w\in\{w_i|i\in\{1,\ldots,n\}\}}\left(\sum_{i\in I_w}\alpha_i\cdot w\right)\\
&\underset{\text{[theorem: 11.68]}}{=} \sum_{w\in\{w_i|i\in\{1,\ldots,n\}\}}\left(\left(\sum_{i\in I_w}\alpha_i\right)\cdot w\right)\\
&\underset{\text{[theorem: 11.18]}}{=} \sum_{w\in\{w_i|i\in\{1,\ldots,n\}\}}\beta_w\cdot w
\end{aligned}$$

□

We have also a stronger opposite of the above theorem.

Theorem 11.81. *Let $\langle V,+,\cdot\rangle$ be a vector space over a field $\langle F,+,\cdot\rangle$, $W\subseteq V$ a finite set and $v\in V$ such that there exists a $\{\alpha_u\}_{u\in W}\subseteq F$ such that*

$$v=\sum_{u\in W}\alpha_u\cdot u$$

then there exists a ordered distinct family $\{w_i\}_{i\in\{1,\ldots,n\}}\subseteq W$ satisfying $W=\{w_i|i\in\{1,\ldots,n\}\}$ where $n=\operatorname{card}(W)$ and a $\{\beta_i\}_{i\in\{1,\ldots,n\}}\subseteq F$ such that

$$\{\beta_i|i\in\{1,\ldots,n\}\}=\{\alpha_w|w\in W\} \textit{ and } \sum_{i\in\{1,\ldots,n\}}\beta_i\cdot w_i=\sum_{w\in W}\alpha_w\cdot w$$

Proof. As W is finite we have by [theorems: 10.78] that for $n=\operatorname{card}(W)$ there exists a bijection

$$w:\{1,\ldots,n\}\rightarrow W$$

defining the ordered distinct family $\{w_i\}_{i\in\{1,\ldots,n\}}\subseteq W$ such that

$$W=w(\{1,\ldots,n\})=\{w_i|i\in\{1,\ldots,n\}\}$$

Define then

$$\{\beta_i\}_{i\in\{1,\ldots,n\}}\subseteq F \text{ by } \beta_i=\alpha_{w(i)}$$

If $x \in \{\beta_i | i \in \{1, \ldots, n\}\}$ then there exists a $i \in \{1, \ldots, n\}$ such that $x = \beta_i = \alpha_{w(i)}$ proving that $x \in \{\alpha_w | w \in W\}$. On the other hand if $x \in \{\alpha_w | w \in W\}$ then there exists a $w \in W$ such that $x = \alpha_w$ then we have by surjectivity a $i \in \{1, \ldots, n\}$ such that $w = w(i)$ hence $x = \alpha_{w(i)} = \beta_i \in \{\beta_i | i \in \{1, \ldots, n\}\}$. So we have

$$\{\beta_i | i \in \{1, \ldots, n\}\} = \{\alpha_w | w \in W\}$$

Further we have

$$\begin{aligned} \sum_{i \in \{1, \ldots, n\}} \beta_i \cdot w_i &= \sum_{i \in \{1, \ldots, n\}} \alpha_{w(i)} \cdot w(i) \\ &\underset{[\text{theorem: } 11.34]}{=} \sum_{w \in W} \alpha_w \cdot w \end{aligned}$$

□

The set of all linear combinations of a set is called the span of the set.

Definition 11.82. *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 have the following special case for the span of finite sets

Theorem 11.83. *Let $\langle V, +, \cdot \rangle$ be a vector space over a field $\langle F, +, \cdot \rangle$, and $\{v_i\}_{i \in \{1, \ldots, n\}} \subseteq V$ then*

$$\operatorname{span}(\{v_i | i \in \{1, \ldots, n\}\}) = \left\{ v \in V | \exists \{\alpha_i\}_{i \in \{1, \ldots, n\}} \subseteq F \text{ such that } v = \sum_{i \in \{1, \ldots, n\}} \alpha_i \cdot v_i \right\}$$

Proof. Let

$$S = \left\{ v \in V | \exists \{\alpha_i\}_{i \in \{1, \ldots, n\}} \subseteq F \text{ such that } v = \sum_{i \in \{1, \ldots, n\}} \alpha_i \cdot v_i \right\}$$

If $x \in S$ there exists a $\{\alpha_i\}_{i \in \{1, \ldots, n\}} \subseteq F$ such that $x = \sum_{i \in \{1, \ldots, n\}} \alpha_i \cdot v_i$. Using [theorem: 11.80] there exists $\{\beta_w\}_{w \in \{v_i | i \in \{1, \ldots, \}\}} \subseteq F$ such that

$$x = \sum_{w \in \{v_i | i \in \{1, \ldots, n\}\}} \beta_w \cdot w$$

As $\{v_i | i \in \{1, \ldots, n\}\}$ is finite [see theorem: 6.44] it follows from the above that

$$x \in \operatorname{span}(\{v_i | i \in \{1, \ldots, n\}\})$$

proving that

$$S \subset \operatorname{span}(\{v_i | i \in \{1, \ldots, n\}\}) \tag{11.19}$$

If $x \in \operatorname{span}(\{v_i | i \in \{1, \ldots, n\}\})$ then there exist a finite $J \subseteq \{v_i | i \in \{1, \ldots, n\}\}$ and a $\{\beta_w\}_{w \in J} \subseteq F$ such that

$$x = \sum_{w \in J} \beta_w \cdot w \tag{11.20}$$

Let $v: \{1, \ldots, n\} \to V$ be the function that defines $\{v_i\}_{i \in \{1, \ldots, n\}} \subseteq V$ then by [theorem: 3.125] there exists a $K \subseteq \{1, \ldots, n\}$ such that

$$v_{|K}: K \to v(\{1, \ldots, n\}) = \{v_i | i \in \{1, \ldots, n\}\} \text{ is a bijection}$$

take $L = (v_{|K})^{-1}(J)$ then

$$(v_{|k})_{|L}: L \to J \text{ is a bijection}$$

Define now

$$\{\alpha_i\}_{i\in\{1,\ldots,n\}}\subseteq F \text{ by } \alpha_i=\begin{cases}\beta_{v(i)} \text{ if } i\in L\\ 0 \text{ if } i\in\{1,\ldots,n\}\setminus L\end{cases}$$

then we have

$$\begin{aligned}\sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot v_i &\underset{[\text{theorem: } 11.41]}{=} \sum_{i\in\{1,\ldots,n\}\setminus L}\alpha_i\cdot v_i+\sum_{i\in L}\alpha_i\cdot v_i\\ &= \sum_{i\in\{1,\ldots,n\}\setminus L}0\cdot v_i+\sum_{i\in L}\beta_{v(i)}\cdot v(i)\\ &= \sum_{i\in L}\beta_{v(i)}\cdot v(i)\\ &\underset{L\subseteq K\subseteq\{1,\ldots,n\}}{=} \sum_{i\in L}\beta_{(v_{|K})_{|L}(i)}\cdot(v_{|K})_{|L}(i)\\ &\underset{[\text{theorem: } 11.34]}{=} \sum_{w\in J}\beta_w\cdot w\\ &\underset{[\text{eq: } 11.20]}{=} x\end{aligned}$$

proving that $x\in S$. Hence we have that $\operatorname{span}(\{v_i|i\in\{1,\ldots,n\}\})\subseteq S$ and combining this with [eq: 11.19] proves that

$$S=\operatorname{span}(\{v_i|i\in\{1,\ldots,n\}\})$$ □

Example 11.84. Let $\langle V,+,\cdot\rangle$ be a vector space over a field $\langle F,+,\cdot\rangle$ then $\operatorname{span}(\varnothing)=\{0\}$

Proof. Let $v\in\operatorname{span}(\varnothing)$ then there exists a finite $I\subseteq\varnothing\Rightarrow I=\varnothing$ and a $\{\alpha_w\}_{w\in\varnothing}\subseteq F$ such that

$$v=\sum_{w\in\varnothing}\alpha_w\cdot w\underset{[\text{definition: } 11.31]}{=}0$$

proving that

$$\operatorname{span}(\varnothing)=0$$ □

Example 11.85. Let $\langle V,+,\cdot\rangle$ be a vector space over a field $\langle F,+,\cdot\rangle$ then $\operatorname{span}(\{0\})=\{0\}$

Proof. Let $v\in\operatorname{span}(\{0\})$ 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:

$I=\varnothing$. Then $v=\sum_{u\in I}\alpha_u\cdot u=\sum_{u\in\varnothing}\alpha_u\cdot u\underset{[\text{definition: } 11.31]}{=}0$

$I=\{0\}$. Then $v=\sum_{u\in I}\alpha_u\cdot u=\sum_{u\in\{0\}}\alpha_u\cdot u\underset{[\text{theorem: } 11.35]}{=}\alpha_0\cdot 0=0$

proving that

$$\operatorname{span}(\{0\})=\{0\}$$

□

Theorem 11.86. *Let $\langle V,+,\cdot\rangle$ be a vector space over a field $\langle F,+,\cdot\rangle$ and $W\subseteq V$ then*

$$v\in\operatorname{span}(W)$$

$$\Updownarrow$$

$$\textit{There exists a finite } I\subseteq W \textit{ and a } \{\alpha_w\}_{w\in I}\subseteq F\setminus\{0\} \textit{ such that } v=\sum_{w\in I}\alpha_w\cdot w$$

Proof.

⇒. As $v\in\operatorname{span}(W)$ 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 $I=\{w\in J|\alpha_w\neq 0\}$ and define $\{\alpha_w\}_{w\in I}\subseteq F\setminus\{0\}$ by $\alpha_w=\beta_w$ then we have

$$\begin{aligned}
v &= \sum_{w\in J}\beta_w\cdot w\\
&\underset{[\text{theorem: }11.41]}{=} \sum_{w\in J\setminus I}\beta_w\cdot w+\sum_{w\in J}\beta_w\cdot w\\
&= \sum_{w\in J\setminus I}0\cdot w+\sum_{w\in J}\beta_w\cdot w\\
&\underset{[\text{theorem: }11.35]}{=} \sum_{w\in J}\beta_w\cdot w\\
&= \sum_{w\in J}\alpha_w\cdot w
\end{aligned}$$

⇐. As $F\setminus\{0\}\subseteq F$ this follows from the definition of $\operatorname{span}(W)$.

□

Theorem 11.87. *Let $\langle V,+,\cdot\rangle$ be a vector space over a field $\langle F,+,\cdot\rangle$ and $W\subseteq V$ then*

$$v\in\operatorname{span}(W)$$

$$\Downarrow$$

$$\exists\{w_i\}_{i\in\{1,\ldots,n\}}\subseteq W \text{ a distinct family and } \{\alpha_i\}_{i\in\{1,\ldots,n\}}\subseteq F\setminus\{0\} \text{ such that } v=\sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot w_i$$

and we have

$$\exists\{w_i\}_{i\in\{1,\ldots,n\}}\subseteq W \text{ and } \{\alpha_i\}_{i\in\{1,\ldots,n\}}\subseteq F \text{ such that } v=\sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot w_i$$

$$\Downarrow$$

$$v\in\operatorname{span}(W)$$

Proof.

⇒. As $v\in\operatorname{span}(W)$ then there exists a finite $I\subseteq W$ and a $\{\beta_w\}_{w\in I}\subseteq F$ such that $v=\sum_{w\in I}\alpha_w\cdot w$. Define $J=\{w\in I|\beta_w\neq 0\}\subseteq I$ then we have for $\{\beta_w\}_{w\in J}\subseteq F\setminus\{0\}$ that

$$\begin{aligned}
v &= \sum_{w\in I}\beta_w\cdot w\\
&\underset{[\text{theorem: }11.41]}{=} \sum_{w\in I\setminus J}\beta_w\cdot w+\sum_{w\in J}\beta_w\cdot w\\
&= \sum_{w\in I\setminus J}0\cdot w+\sum_{w\in J}\beta_w\cdot w\\
&\underset{[\text{theorem: }11.35]}{=} \sum_{w\in J}\beta_w\cdot w
\end{aligned}$$

Using [theorem: 11.81] there exists a ordered distinct family $\{w_i\}_{i\in\{1,\ldots,n\}}\subseteq J\subseteq W$ and a $\{\alpha_i\}_{i\in\{1,\ldots,n\}}\subseteq F$ such that $0\notin\{\beta_w|w\in J\}=\{\alpha_i|i\in\{1,\ldots,n\}\}$ so that $\{\alpha_i\}_{i\in\{1,\ldots,n\}}\subseteq F\setminus\{0\}$ and $\sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot w_i=\sum_{w\in I}\beta_w\cdot w=v$.

⇐. Assume that $\exists\{w_i\}_{i\in\{1,\ldots,n\}}\subseteq W$ a distinct family and $\{\beta_i\}_{i\in\{1,\ldots,n\}}\subseteq F$ such that $v=\sum_{i\in\{1,\ldots,n\}}\beta_i\cdot w_i$. Using [theorem: 11.77] $I=\{w_i|i\in\{1,...,n\}\}$ is finite, $I\subseteq W$ and using [theorem: 11.80] there exists a $\{\alpha_w\}_{w\in I}\subseteq F$ such that $v=\sum_{i\in\{1,\ldots,n\}}\beta_i\cdot w_i=\sum_{w\in I}\alpha_w\cdot w$ proving that $v\in\operatorname{span}(W)$ □

Theorem 11.88. *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

$$\sum_{u \in \{w\}} \alpha_u \cdot u \underset{[\text{theorem: } 11.32]}{=} \alpha_w \cdot w = 1 \cdot w = w$$

□

Theorem 11.89. *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 11.90. *Let $\langle V, +, \cdot \rangle$ be a vector space over a field $\langle F, +, \cdot \rangle$ and $W \subseteq V$ then*

$$\operatorname{span}(W) \text{ is a sub-space of } \langle V, +, \cdot \rangle$$

Proof. As $\varnothing \subseteq W$ we have $\{0\} \underset{[\text{example: } 11.84]}{=} \operatorname{span}(\varnothing) \subseteq_{[\text{theorem: } 11.89]} \operatorname{span}(W)$ proving that

$$0 \in W \text{ hence } W \neq \varnothing$$

Let $x, y \in \operatorname{span}(W)$ and $\alpha, \beta \in F$ 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

$$\{\gamma_w\}_{w \in K} \subseteq F \text{ by } \gamma_w = \begin{cases} \alpha \cdot \alpha_w \text{ if } w \in I \setminus J \\ \alpha \cdot \alpha_w + \beta \cdot \beta_w \text{ if } w \in I \bigcap J \\ \beta \cdot \beta_w \text{ if } w \in J \setminus I \end{cases}$$

then we have

$$\begin{aligned}
\sum_{w \in K} \gamma_w \cdot w &\underset{[\text{theorem: } 11.41]}{=} \\
\sum_{w \in I \setminus J} \gamma_w \cdot w + \sum_{w \in I \bigcap J} \gamma_w \cdot w + \sum_{w \in J \setminus I} \gamma_w \cdot w &= \\
\sum_{w \in I \setminus J} (\alpha \cdot \alpha_w) \cdot w + \sum_{w \in I \bigcap J} (\alpha \cdot \alpha_w + \beta \cdot \beta_w) \cdot w + \sum_{w \in J \setminus I} (\beta \cdot \beta_w) \cdot w &= \\
\sum_{w \in I \setminus J} \alpha \cdot (\alpha_w \cdot w) + \sum_{w \in I \bigcap J} (\alpha \cdot (\alpha_w \cdot w) + \beta \cdot (\beta_w \cdot w)) + \sum_{w \in J \setminus I} \beta \cdot (\beta_w \cdot w) &\underset{[\text{theorem: } 11.67]}{=} \\
\alpha \cdot \sum_{w \in I \setminus J} \alpha_w \cdot w + \sum_{w \in I \bigcap J} (\alpha \cdot (\alpha_w \cdot w) + \beta \cdot (\beta_w \cdot w)) + \beta \cdot \sum_{w \in J \setminus I} \beta_w \cdot w &\underset{[\text{theorem: } 11.36]}{=} \\
\alpha \cdot \sum_{w \in I \setminus J} \alpha_w \cdot w + \sum_{w \in I \bigcap J} \alpha \cdot (\alpha_w \cdot w) + \sum_{i \in I \bigcap J} \beta \cdot (\beta_w \cdot w) + \beta \cdot \sum_{w \in J \setminus I} \beta_w \cdot w &\underset{[\text{theorem: } 11.67]}{=} \\
\alpha \cdot \sum_{w \in I \setminus J} \alpha_w \cdot w + \alpha \cdot \sum_{w \in I \bigcap J} \alpha_w \cdot w + \beta \cdot \sum_{i \in I \bigcap J} \beta_w \cdot w) + \beta \cdot \sum_{w \in J \setminus I} \beta_w \cdot w &= \\
\alpha \cdot \left(\sum_{w \in I \setminus J} \alpha_w \cdot w + \sum_{w \in I \bigcap J} \alpha_w \cdot w \right) + \beta \cdot \left(\sum_{w \in J \setminus I} \beta_w \cdot w + \sum_{i \in I \bigcap J} \beta_w \cdot w) \right) &\underset{[\text{theorem: } 11.41]}{=} \\
\alpha \cdot \sum_{w \in I} \alpha_w \cdot w + \beta \cdot \sum_{w \in J} \beta_w \cdot w &= \\
\alpha \cdot x + \beta \cdot y &
\end{aligned}$$

proving that

$$\alpha \cdot x + \beta \cdot y \in \operatorname{span}(W)$$

Hence by [definition: 11.51] it follows that $\operatorname{span}(W)$ is a sub-space of $\langle V,+,\cdot\rangle$. □

Theorem 11.91. *Let $\langle V,+,\cdot\rangle$ be a vector space over a field $\langle F,+,\cdot\rangle$ and $W\subseteq V$ as 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: 11.70] 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: 11.88] proves

$$W=\operatorname{span}(W)$$ □

Corollary 11.92. *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: 11.90] $\operatorname{span}(W)$ is a sub-space of $\langle V,+,\cdot\rangle$, hence using [theorem: 11.92] we have that

$$\operatorname{span}(W)=\operatorname{span}(\operatorname{span}(w))$$ □

Theorem 11.93. *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:

***$w\in W$*.** Then $W=W\bigcup\{w\}$ so that $\operatorname{span}(W)=\operatorname{span}(W\bigcup\{w\})$

***$w\notin W$*.** 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

$$w=\sum_{u\in I}\alpha_u\cdot u \tag{11.21}$$

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

$$v=\sum_{u\in J}\beta_u\cdot u \tag{11.22}$$

For J we have either:

***$w\notin J$*.** Then $J\subseteq W$ so that $v\in\operatorname{span}(W)$

***$w\in J$*.** Take then the finite set $K=(I\bigcup J)\setminus\{u\}$then, as $I,J\subseteq W\Rightarrow I\bigcup J\subseteq W$, we have $K\subseteq W$. Further

$$\begin{aligned}K &= \left(I\bigcup J\right)\setminus\{w\}\\ &= \left((I\setminus J)\bigcup\left(I\bigcap J\right)\bigcup(J\setminus I)\right)|\{w\}\\ &= ((I\setminus J)\setminus\{w\})\bigcup\left(\left(I\bigcap J\right)\setminus\{w\}\right)\bigcup((J\setminus I)\setminus\{w\})\\ &\underset{w\notin W\Rightarrow w\notin I}{=} (I\setminus J)\bigcup\left(I\bigcap J\right)\bigcup((J\setminus I)\setminus\{w\})\end{aligned} \tag{11.23}$$

and we have

$$(I\setminus J)\bigcap\left(I\bigcap J\right)=\varnothing\wedge(I\setminus J)\bigcap((J\setminus I)\setminus\{w\})=\varnothing\wedge\left(I\bigcap J\right)\bigcap((J\setminus I)\setminus\{w\})=\varnothing \tag{11.24}$$

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{array}{rl}
\sum_{u\in K}\gamma_u\cdot u & \underset{[\text{theorem: } 11.41]}{=} \\
\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(u\cdot\alpha_u)+\beta_u\cdot u)+\sum_{u\in(J\setminus I)\setminus\{w\}}\beta_u\cdot u & \underset{[\text{theorem: } 11.67]}{=} \\
\beta_w\cdot\sum_{u\in I\setminus J}\alpha_u\cdot u+\sum_{u\in I\bigcap J}(\beta_w\cdot(u\cdot\alpha_u)+\beta_u\cdot u)+\sum_{u\in(J\setminus I)\setminus\{w\}}\beta_u\cdot u & \underset{[\text{theorem: } 11.36]}{=} \\
\beta_w\cdot\sum_{u\in I\setminus J}\alpha_u\cdot u+\sum_{u\in I\bigcap J}\beta_w\cdot(u\cdot\alpha_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: } 11.67]}{=} \\
\beta_w\cdot\sum_{u\in I\setminus J}\alpha_u\cdot u+\beta_w\cdot\sum_{u\in I\bigcap J}u\cdot\alpha_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}u\cdot\alpha_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: } 11.41]}{=} \\
\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:}11.21]}{=} \\
\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: } 11.41]}{=} \\
\beta_w\cdot w+\sum_{i\in J\setminus\{w\}}\beta_u\cdot u & \underset{[\text{theorem: } 11.32]}{=} \\
\sum_{i\in\{w\}}\beta_u\cdot u+\sum_{i\in J\setminus\{w\}}\beta_u\cdot u & \underset{[\text{theorem: } 11.41]}{=} \\
\sum_{i\in J}\beta_u\cdot u & \underset{[\text{eq: } 11.22]}{=} \\
v &
\end{array}$$

proving, as $K\subseteq W$, 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: 11.89] 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. □

11.3.2.2 Linear (in)dependent sets

Definition 11.94. (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 $\exists w\in I$ with $\alpha_w\neq 0$ such that $\sum_{u\in I}\alpha_u\cdot u=0$*

Example 11.95. 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 $\sum_{u\in\{0\}}\alpha_u\cdot u\underset{[\text{theorem: } 11.32]}{=}\alpha_0\cdot 0=0$ proving that W is linear dependent. □

Theorem 11.96. (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 11.97. *Let $\langle V,+,\cdot\rangle$ be a vector space over the field $\langle F,+,\cdot\rangle$ and $W\subseteq V$ then*

$$W \text{ is linear independent}$$
$$\Updownarrow$$
$$\text{for every finite } I\subseteq W \text{ we have } \forall\{\alpha_u\}_{u\in I}\subseteq F \text{ if } \sum_{u\in I}\alpha_u\cdot u=0 \text{ then } \forall u\in I\ \alpha_u=0$$

Proof.

⇒. As W is independent we have by definition that W is not dependent. Hence for every finite $I\subseteq W$ we have for every $\{\alpha_u\}_{u\in I}\subseteq F$ satisfying $\exists w\in I$ with $\alpha_w\neq 0$ that $\sum_{u\in I}\alpha_u\cdot u\neq 0$. So if we have a finite $I\subseteq W$ that satisfies $\sum_{u\in I}\alpha_u\cdot u=0$ we must have $\forall u\in I$ that $\alpha_u=0$ otherwise we reach the contradiction that $\sum_{u\in I}\alpha_u\cdot u\neq 0$.

⇐. 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 $\sum_{u\in I}\alpha_u\cdot u=0$. Using the hypothesis we have then that $\forall u\in I$ that $\alpha_u=0$ contradicting $\alpha_w\neq 0$. So we must have that W is linear independent. □

If W is finite then we have a simpler definition.

Theorem 11.98. *Let $\langle V,+,\cdot\rangle$ be a vector space over the field $\langle F,+,\cdot\rangle$ and $W\subseteq V$ a **finite** set then*

$$W \text{ is linear independent}$$
$$\Updownarrow$$
$$\forall\{\alpha_u\}_{u\in W}\subseteq F \text{ such that } \sum_{w\in W}\alpha_u\cdot u=0 \text{ we have } \forall u\in W \text{ we have } \alpha_u=0$$

Proof.

⇒. As W is finite and $W\subseteq W$ this follows from [theorem: 11.97].

⇐. 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

$$\begin{aligned}\sum_{u\in W}\beta_u\cdot u &\underset{[\text{theorem: } 11.41]}{=} \sum_{u\in W\setminus I}\beta_u\cdot u+\sum_{u\in I}\beta_u\cdot u\\ &= \sum_{u\in W\setminus I}0\cdot u+\sum_{u\in I}\alpha_u\cdot u\\ &\underset{[\text{theorem: } 11.35]}{=} \sum_{u\in I}\alpha_u\cdot u\\ &= 0\end{aligned}$$

By the hypothesis we have then 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. □

Example 11.99. 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 $\sum_{u\in\varnothing}\alpha_u\cdot u$ we have that $\alpha_u=0$ is satisfied vacuously. □

Example 11.100. 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: 11.32] $\alpha_v\cdot v=0$. 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: 11.98] that the finite set $\{v\}$ 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 11.101. *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 that $\exists w\in J$ with $\beta_w=0$ so that $\sum_{u\in J}\beta_u\cdot u=0$. So we have that

$$\begin{aligned}
0 &= \sum_{u\in J}\beta_u\cdot u\\
&\underset{\text{[theorem: 11.41]}}{=} \sum_{u\in J\setminus\{w\}}\beta_u\cdot u+\sum_{u\in\{w\}}\beta_u\cdot u\\
&\underset{\text{[theorem: 11.32]}}{=} \sum_{u\in J\setminus\{w\}}\beta_u\cdot u+\beta_w\cdot w
\end{aligned}$$

proving as $\beta_w\neq 0$ that

$$\begin{aligned}
w &= (\beta_w)^{-1}\cdot\left(-\sum_{u\in J\setminus\{w\}}\beta_u\cdot u\right)\\
&\underset{\text{[theorem: 11.38]}}{=} (\beta_w)^{-1}\cdot\sum_{u\in J\setminus\{w\}}-(\beta_u\cdot u)\\
&= (\beta_w)^{-1}\cdot\sum_{u\in J\setminus\{w\}}(-\beta_u)\cdot u\\
&\underset{\text{[theorem: 11.67]}}{=} \sum_{u\in J\setminus\{w\}}((\beta_w)^{-1}((-\beta_u)\cdot u))\\
&= \sum_{u\in J\setminus\{w\}}((\beta_w)^{-1}(-\beta_u))\cdot u
\end{aligned}$$

So if we define $I=J\setminus\{w\}$ and $\{\alpha_u\}_{u\in I}$ by $\beta_u=(\beta_w)^{-1}(-\beta_u)$ then $J\subseteq W\setminus\{w\}$ 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

$$\begin{aligned}
\sum_{u\in J}\beta_u\cdot u &\underset{\text{[theorem: 11.41]}}{=} \sum_{u\in I}\beta_u\cdot u+\sum_{u\in\{w\}}\beta_u\cdot u\\
&= \sum_{u\in I}\alpha_u\cdot u+\sum_{u\in\{w\}}\beta_u\cdot u\\
&\underset{\text{[theorem: 11.32]}}{=} \sum_{u\in I}\alpha_u\cdot u+\beta_w\cdot w\\
&= \sum_{u\in I}\alpha_u\cdot u+(-1)\cdot w\\
&= w+(-w)\\
&= 0
\end{aligned}$$

which as $\beta_w = -1 \neq 0$ proves that W is linear dependent. □

Related to a linear (in)dependent set is that of a linear (in)dependent family. In this text we need only families with a finite index. If we have to deal with infinite we use (in)dependent sets.

Definition 11.102. *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 dependent if there exist a $\{\alpha_i\}_{i\in\{1,\ldots,n\}} \subseteq F$ satisfying $\exists i \in \{1,\ldots, n\}$ such that*

$$\sum_{j\in\{1,\ldots,n\}\backslash\{i\}} \alpha_i \cdot x_i = 0$$

Definition 11.103. *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.*

We have a equivalent definition of a linear independent family.

Theorem 11.104. *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*

$$\{x_i\}_{i\in\{1,\ldots,n\}} \subseteq V \textit{ is linear independent}$$

$$\Updownarrow$$

$$\forall \{\alpha_i\}_{i\in\{1,\ldots,n\}} \subseteq F \textit{ such that } \sum_{i\in\{1,\ldots,n\}} \alpha_i \cdot x_i = 0 \textit{ we have } \forall i \in \{1,\ldots,n\}\ \alpha_i = 0$$

Proof.

⇒. As $\{\alpha_i\}_{i\in\{1,\ldots,n\}} \subseteq F$ is 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 have that $\sum_{i\in\{1,\ldots,n\}} \alpha_i \cdot x_i \neq 0$. 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 $\sum_{i\in\{1,\ldots,n\}} \alpha_i \cdot x_i = 0$. Using the hypothesis we have then that $\forall i \in \{1,\ldots,n\}$ that $\alpha_i = 0$ contradicting $\alpha_k \neq$. So we must have that $\{\alpha_i\}_{i\in\{1,\ldots,n\}} \subseteq F$ is linear independent. □

If $\{x_i\}_{i\in\{1,\ldots,n\}} \subseteq X$ is linear dependent family then it is not true that $\{x_i | i \in \{1,\ldots,n\}\}$ is a linear dependent set as the following example shows.

Example 11.105. 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 indpendent.

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: 11.100] linear independent. □

The opposite is however true as the following theorem shows.

Theorem 11.106. *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 family such 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: 10.83] 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 [theorem: 3.125] there exists a $J\subseteq\{1,\ldots,n\}$ so that

$$x_{|J}: J\rightarrow W \text{ is a bijection}$$

Define now

$$\{\beta_k\}_{k\in\{1,\ldots,n\}} \text{ by } \beta_k=\begin{cases} \alpha_{(x_{|J})(i)} \text{ if } k\in(x_{|J})^{-1}(I) \\ -0 \text{ if } k\in\{1,\ldots,n\}\setminus(x_{|J})^{-1}(I) \end{cases}$$

then

$$\begin{aligned} \sum_{i\in\{1,\ldots,n\}}\beta_i\cdot x_i &\underset{[\text{theorem: } 11.41]}{=} \sum_{i\in\{1,\ldots,n\}\setminus(x_{|J})^{-1}(I)}\beta_i\cdot x_i+\sum_{(x_{|J})^{-1}(I)}\beta_i\cdot x_i \\ &\underset{(x_{|J})^{-1}(I)\subseteq J}{=} \sum_{i\in\{1,\ldots,n\}\setminus(x_{|J})^{-1}(I)}0\cdot x_i+\sum_{i\in(x_{|J})^{-1}(I)}\alpha_{|x_{|J}(i)}\cdot x_{|J}(i) \\ &= \sum_{i\in\{1,\ldots,n\}\setminus(x_{|J})^{-1}(I)}0\cdot x_i+\sum_{i\in(x_{|J})^{-1}(I)}\alpha_{|x_{|J}(i)}\cdot x_{|J}(i) \\ &\underset{[\text{theorem: } 11.34]}{=} \sum_{i\in}\alpha_i\cdot i \\ &= 0 \end{aligned}$$

which as $\beta_{(x_{|J})^{-1}(l)}=\alpha_{x_{|}((x_{|J})^{-1}(l))}=\alpha_l\neq 0$ proves that

$$\{x_i\}_{i\in\{1,\ldots,n\}} \text{ is linear dependent}$$ □

Corollary 11.107. *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 linear independent family then $\{x_i|i\in\{1,\ldots,n\}\}$ is linear independent.*

Proof. Assume that $\{x_i|i\in\{1,\ldots,n\}\}$ is not linear independent, then $\{x_i|i\in\{1,\ldots,n\}\}$ must be linear dependent. Using the previous theorem [theorem: 11.106] it follows that $\{x_i\}_{i\in\{1,\ldots,n\}}\subseteq x$ is linear dependent, contradicting the linear independency of $\{x_i\}_{i\in\{1,\ldots,n\}}$. □

One difference between a set and a family is that a set can not contain duplicate elements while a family can have duplicate members (the same element with two different indexes), so the following theorem is unique for families.

Theorem 11.108. *Let $\langle V,+,\cdot\rangle$ be a vector space over a field $\langle F,+,\cdot\rangle$, $n\in\mathbb{N}$ and $\{x_k\}_{k\in\{1,\ldots,n\}}\subseteq X$ such that there exists $i,j\in\{1,\ldots,n\}$ with $i\neq j$ such that $x_i=x_j$ then $\{x_k\}_{k\in\{1,\ldots,n\}}$ is linear dependent.*

Proof. 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

$$\begin{aligned} \sum_{k\in\{1,\ldots,n\}}\alpha_k\cdot x_k &\underset{[\text{theorem: } 11.41]}{=} \sum_{k\in\{1,\ldots,n\}\setminus\{i,j\}}\alpha_k\cdot x_k+\sum_{k\in\{i\}}\alpha_k\cdot x_k+\sum_{k\in\{j\}}\alpha_k\cdot x_k \\ &\underset{[\text{theorem: } 11.32]}{=} \sum_{k\in\{1,\ldots,n\}\setminus\{i,j\}}\alpha_k\cdot x_k+\alpha_i\cdot x_i+\alpha_j\cdot x_j \\ &= \sum_{k\in\{1,\ldots,n\}\setminus\{i,j\}}0\cdot x_k+1\cdot x_i+(\alpha_j-1)\cdot x_j \\ &= x_i-x_j \\ &\underset{x_i=x_j}{=} 0 \end{aligned}$$

which as $\alpha_i=1\neq 0$ proves linear dependency. □

Corollary 11.109. *Let $\langle V,+,\cdot\rangle$ be a vector space over a field $\langle F,+,\cdot\rangle$, $n\in\mathbb{N}$ and $\{x_k\}_{k\in\{1,\ldots,n\}}\subseteq X$ is linear independent then*

$$\{x_i\}_{i\in\{1,\ldots,n\}}\subseteq X\text{ is distinct}$$

and

$$\operatorname{card}(\{x_i|i\in\{1,\ldots,n\}\})=n$$

Proof. If $\{x_i\}_{i\in\{1,\ldots,n\}}$ is not distinct then there exists $i,j\in\{1,\ldots,n\}$ such that $x_i=x_j$, hence by the previous theorem [theorem: 11.108] $\{x_i\}_{i\in\{1,\ldots,n\}}$ is linear dependent, contradicting linear independency of $\{x_i\}_{i\in\{1,\ldots,n\}}$. So we have that

$$\{x_i\}_{i\in\{1,\ldots,n\}}\subseteq X\ \text{ is distinct}$$

From the above it follows that $x:\{1,\ldots,n\}\rightarrow X$ is injective [x is the function defining the family] so that $x:\{1,\ldots,n\}\rightarrow x(\{1,\ldots,n\})=\{x_i|i\in\{1,\ldots,n\}\}$ is a bijection. Hence $\{1,\ldots,n\}\approx\{x_i|i\in\{1,\ldots,n\}\}$ proving that

$$\operatorname{card}(\{x_i|i\in\{1,\ldots,n\}\})$$

□

The above theorem shows why linear dependency of family $\{x_i\}_{i\in\{1,\ldots,n\}}$ does not imply linear dependency of $\{x_i|i\in\{1,\ldots,\}\}$, families can have duplicates while sets have none. To avoid this problem we have to work with distinct families.

Theorem 11.110. *Let $\langle V,+,\cdot\rangle$ be a vector space over a field $\langle F,+,\cdot\rangle$, $n\in\mathbb{N}$ and $\{x_k\}_{k\in\{1,\ldots,n\}}\subseteq V$ a ordered distinct family then we have:*

$$\{x_i\}_{i\in\{1,\ldots,n\}}\subseteq V\textit{ is linear dependent}$$
$$\Updownarrow$$
$$\{x_i|i\in\{1,\ldots,n\}\}\textit{ is linear dependent}$$

Proof.

⇒. As $\{x_i\}_{i\in\{1,\ldots,n\}}\subseteq V$ is distinct we have that the function $x:\{1,\ldots,n\}\rightarrow X$ is injective, hence

$$x:\{1,\ldots,n\}\rightarrow x(\{1,\ldots,n\})=\{x_i|i\in\{1,\ldots,n\}\}\text{ is a bijection}$$

so that $\{1,\ldots,n\}\approx\{x_i|i\in\{1,\ldots,n\}\}$ from which it follows that

$$I=\{x_i|i\in\{1,\ldots,n\}\}\text{ is finite}$$

As $\{x_i\}_{i\in\{1,\ldots,n\}}\subseteq V$ is linear dependent there exists a $\{\beta_i\}_{i\in\{1,\ldots,n\}}$ such that $\exists k\in\{1,\ldots,n\}$ with $\beta_k\neq 0$ and $\sum_{i\in\{1,\ldots,n\}}\beta_i\cdot x_i=0$. Define now

$$\{\alpha_w\}_{w\in I}\subseteq F\text{ by }\alpha_w=\beta_{x^{-1}(w)}$$

then we have

$$\begin{aligned}\sum_{w\in I}\alpha_w\cdot w &= \sum_{w\in I}\beta_{x^{-1}(w)}\cdot x^{-1}(w)\\ &\underset{x^{-1}:I\rightarrow\{1,\ldots,n\}\text{ is a bijection and [theorem: 11.34]}}{=} \sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot x_i\\ &= 0\end{aligned}$$

which as $\alpha_{x(k)}=\beta_{x^{-1}(x(k))}=\beta_k\neq 0$ and $I\subseteq I=\{x_i|i\in\{1,\ldots,n\}\}$ proves that

$$\{x_i|i\in\{1,\ldots,n\}\}\text{ is linear dependent}$$

⇐. This follows from [theorem: 11.106] □

Corollary 11.111. *Let $\langle V,+,\cdot\rangle$ be a vector space over a field $\langle F,+,\cdot\rangle$, $n\in\mathbb{N}$ and*

$\{x_k\}_{k\in\{1,\ldots,n\}}\subseteq V$ a ordered distinct family then we have:

$$\{x_i\}_{i\in\{1,\ldots,n\}}\subseteq V \textit{ is linear independent}$$
$$\Updownarrow$$
$$\{x_i|i\in\{1,\ldots,n\}\} \textit{ is linear independent}$$

Proof. This follows from [theorem: 11.110] by contra-position. □

We have a similar theorem as [theorem: 11.101].

Theorem 11.112. *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 $\{\beta_i\}_{i\in\{1,\ldots,n\}}\subseteq F$ and a $k\in\{1,\ldots,n\}$ such that $\beta_k\neq 0$ and $\sum_{i\in\{1,\ldots,n\}}\beta_i\cdot x_i=0$. So that

$$\begin{aligned}0 &= \sum_{i\in\{1,\ldots,n\}}\beta_i\cdot x_i\\ &\underset{[\text{theorem: }11.41]}{=} \sum_{i\in\{1,\ldots,n\}\setminus\{k\}}\beta_i\cdot x_i+\sum_{i\in\{k\}}\beta_i\cdot x_i\\ &\underset{[\text{theorem: }11.32]}{=} \sum_{i\in\{1,\ldots,n\}\setminus\{k\}}\beta_i\cdot x_i+\beta_k\cdot x_k\end{aligned}$$

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

$$\begin{aligned}\sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot x_i &\underset{[\text{theorem: }11.41]}{=} \sum_{i\in\{1,\ldots,n\}\setminus\{k\}}\alpha_i\cdot x_i+\sum_{i\in\{k\}}\alpha_i\cdot x_i\\ &\underset{[\text{theorem: }11.32]}{=} \sum_{i\in\{1,\ldots,n\}\setminus\{k\}}\alpha_i\cdot x_i+\alpha_k\cdot x_k\\ &= \sum_{i\in\{1,\ldots,n\}\setminus\{k\}}\beta_i\cdot x_i+(-1)\cdot x_k\\ &= x_k=x_k\\ &= 0\end{aligned}$$

proving, as $\alpha_k=-1\neq 0$, proves linear dependency. □

Example 11.113. Let $\langle V,+,\cdot\rangle$ be a vector space over a field $\langle F,+,\cdot\rangle$ and $\{x_1\}_{i\in\{1,2\}}$ a linear dependent family then there exists a $\alpha\in F$ such that $x_1=\alpha\cdot x_2$.

Proof. Using [theorem: 11.112] there exists a $k\in\{1,2\}$ and a $\{\alpha_i\}_{i\in\{1,2\}\setminus\{k\}}\subseteq F$ such that $x_k=\sum_{i\in\{1,2\}\setminus\{k\}}\alpha_i\cdot x_i$. For k we have then either:

$\boldsymbol{k=1.}$ Then $x_1=\sum_{i\in\{2\}}\alpha_i\cdot x_i=\alpha_2\cdot x_2$

$\boldsymbol{k=2.}$ Then $x_2=\sum_{i\in\{1\}}\alpha_i\cdot x_i=\alpha_1\cdot x_1$. If $x_1=0$ then $x_1=0\cdot x_2$ and if $x_1\neq 0$ we must have that $\alpha_1\neq 0$ so that $x_1=(\alpha_1)^{-1}\cdot x_2$. □

Theorem 11.114. *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 will be important later when we define the concept of a dimension.

Lemma 11.115. (Steinitz Lemma) *Let $\langle V,+,\cdot\rangle$ be a vector space over a field $\langle F,+,\cdot\rangle$. Suppose that $\{e_i\}_{i\in\{1,\ldots,n\}}\subseteq V$ is a ordered disjoint family such that $\{e_i|i\in\{1,\ldots,n\}\}$ is linear independent and $T\subseteq V$ satisfies* $\operatorname{span}(T)=V$. *Then there exists a $T_n\subseteq T$ with* $\operatorname{card}(T_n)=n$ *such that*

$$\operatorname{span}\big((T\setminus T_n)\bigcup\{e_i|i\in\{1,\ldots,n\}\}\big)=V$$

Proof. We prove this by induction so let

$S-\big\{k\in\mathbb{N}_0|\text{If }0\leqslant k\leqslant n\text{ then }\exists T'_k\subseteq T\text{ with }\operatorname{card}(T'_k)=k\text{ such that }\operatorname{span}\big((T\setminus T'_k)\bigcup\{e_i|i\in\{1,\ldots,$
$k\}\}\big)=V\big\}$

then we have:

$\boldsymbol{0\in S.}$ Take $T'_0=\varnothing\subseteq T$ then $\operatorname{card}(T'_0)=0$ and

$$V=\operatorname{span}(T)=\operatorname{span}\big((T\setminus\varnothing)\bigcup\varnothing\big)=\operatorname{span}\big((T\setminus T'_0)\bigcup\{e_i|i\in\{1,\ldots,0\}\}\big)$$

proving that $0\in S$.

$\boldsymbol{k\in S\Rightarrow k+1\in S.}$ If $0\leqslant k+1\leqslant n$ then $k<n$ and as $k\in S\subseteq\mathbb{N}_0$ we have $0\leqslant k$ so that $0\leqslant k<n$. Using the fact that $k\in S$ again we have that $\exists T'_k\subseteq T$ with $\operatorname{card}(T'_k)=k$ such that $\operatorname{span}((T\setminus T'_k)\bigcup\{e_i|i\in\{1,\ldots,k\}\})=V$. Define

$$T_k=(T\setminus T'_k)\bigcup\{e_i|i\in\{1,\ldots,k\}\}\tag{11.25}$$

then we have

$$\operatorname{span}(T_k)=V\tag{11.26}$$

As $e_{k+1}\in V=\operatorname{span}(T_k)$ there exists by [theorem: 11.86] a finite $I\subseteq T_k$ and a $\{\lambda_u\}_{u\in I}\subseteq F\setminus\{0\}$ such that

$$e_{k+1}=\sum_{u\in I}\lambda_u\cdot u\tag{11.27}$$

As $\{e_i|i\in\{1,\ldots,k+1\}\}\subseteq\{e_i|i\in\{1,\ldots,n\}\}$ a linear independent set we have by [theorem: 11.114] that

$$\{e_i|i\in\{1,\ldots,k+1\}\}\text{ is linear independent}\tag{11.28}$$

Assume now that $I\subseteq\{e_i|i\in\{1,\ldots,k\}\}\underset{\text{[theorem: 11.76]}}{=}\{e_i|i\in\{1,\ldots,k+1\}\}\setminus\{e_{k+1}\}$ then from [eq: 11.27] and [theorem: 11.101] it follows that $\{e_i|i\in\{1,\ldots,k+1\}\}$ is linear dependent contradicting [eq: 11.28]. Hence $I\nsubseteq\{e_i|i\in\{1,\ldots,k\}\}$, so

$$\exists t\in I\text{ with }t\notin\{e_i|i\in\{1,\ldots,k\}\}\tag{11.29}$$

As $t\in I\subseteq T_k\underset{\text{[eq: 11.25]}}{=}(T\setminus T'_k)\bigcup\{e_i|i\in\{1,\ldots,k\}\}$ we must have that $t\in T\setminus T'_k$ hence $t\notin T'_k$ and $t\in T$. Using [theorem: 10.81] it follows then that $\operatorname{card}(T'_k\bigcup\{t\})=\operatorname{card}(T'_k)+1=k+1$. To summarize we have

$$t\in T\text{ and }\operatorname{card}(T'_{k+1})=k+1\text{ where }T'_{k+1}=T_k\bigcup\{t\}\tag{11.30}$$

Further

$$\begin{aligned}
T_k\setminus\{t\} &\underset{\text{[eq: 11.25]}}{=} ((T\setminus T'_k)\bigcup\{e_i|i\in\{1,\ldots,k\}\})\setminus\{t\}\\
&= ((T\setminus T'_k)\setminus\{t\})\bigcup(\{e_i|i\in\{1,\ldots,k\}\}\setminus\{t\})\\
&\underset{\text{[eq: 11.29]}}{=} ((T\setminus T'_k)\setminus\{t\})\bigcup\{e_i|i\in\{1,\ldots,k\}\}\\
&\underset{\text{[theorem: 1.31]}}{=} (T\setminus(T'_k\bigcup\{t\}))\bigcup\{e_i|i\in\{1,\ldots,k\}\}
\end{aligned}\tag{11.31}$$

Let

$$T_{k+1}=(T\setminus T'_{k+1})\bigcup\{e_1|i\in\{1,\ldots,k+1\}\}\tag{11.32}$$

Then we have

$$\begin{aligned}
T_{k+1} &=\\
(T\setminus T'_{k+1})\bigcup\{e_1|i\in\{1,\ldots,k+1\}\} &=\\
(T\setminus T'_{k+1})\bigcup(\{e_1|i\in\{1,\ldots,k\}\}\bigcup\{e_{k+1}\}) &=\\
((T\setminus T'_{k+1})\bigcup\{e_1|i\in\{1,\ldots,k\}\})\bigcup\{e_{k+1}\} &\underset{\text{[eq: 11.30]}}{=}\\
((T\setminus(T'_k\bigcup\{t\}))\bigcup\{e_1|i\in\{1,\ldots,k\}\})\bigcup\{e_{k+1}\} &\underset{\text{[theorem: 11.31]}}{=}\\
(T_k\setminus\{t\})\bigcup\{e_{k+1}\}
\end{aligned}$$

proving that

$$T_{k+1}=(T_k\setminus\{t\})\bigcup\{e_{k+1}\}\tag{11.33}$$

Further

$$\begin{aligned}
e_{k+1} &\underset{\text{[eq: 11.27]}}{=} \sum_{u\in I}\lambda_u\cdot u\\
&\underset{\text{[theorem: 11.41]}}{=} \sum_{u\in I\setminus\{t\}}\lambda_u\cdot u+\sum_{u\in\{t\}}\lambda_u\cdot u\\
&\underset{\text{[theorem: 11.32]}}{=} \sum_{u\in I\setminus\{t\}}\lambda_u\cdot u+\lambda_t\cdot t
\end{aligned}\tag{11.34}$$

Now for e_{k+1} we have either:

$\boldsymbol{e_{k+1}\in I\setminus\{t\}}$**.** Then by [eq: 11.34] we have

$$\begin{aligned}
e_{k+1} &= \sum_{u\in I\setminus\{t\}}\lambda_u\cdot u+\lambda_t\cdot t\\
&\underset{\text{[eq: 11.27]}}{=} \sum_{u\in I\setminus\{t,e_{k+1}\}}\lambda_u\cdot u+\sum_{u\in\{e_{k+1}\}}\lambda_u\cdot u+\lambda_t\cdot t\\
&\underset{\text{[theorem: 11.32]}}{=} \sum_{u\in I\setminus\{t,e_{k+1}\}}\lambda_u\cdot u+\lambda_{e_{k+1}}\cdot e_{k+1}+\lambda_t\cdot t
\end{aligned}$$

as $\{\lambda_u\}_{u\in I}\subseteq F\setminus\{0\}\Rightarrow\lambda_t\neq 0$ we have

$$t=((\lambda_t)^{-1}\cdot(1-\lambda_{e_{k+1}}))\cdot e_{k+1}-(\lambda_t)^{-1}\cdot\sum_{u\in I\setminus\{t,e_{k+1}\}}\lambda_u\cdot u \tag{11.35}$$

Further for

$$J=(I\setminus\{t,e_{k+1}\})\bigcup\{e_{k+1}\}\subseteq(T_k\setminus\{t\})\bigcup\{e_{k+1}\}=T_{k+1} \tag{11.36}$$

Then we can define

$$\{\alpha_u\}_{u\in J}\text{ by }\alpha_u=\begin{cases}(\lambda_t)^{-1}\cdot(1-\lambda_{e_{k+1}})\text{ if }u=e_{k+1}\\(-(\lambda_t)^{-1}\cdot\lambda_u)\text{ if }u\in I\setminus\{t,e_{k+1}\}\end{cases}$$

and have

$$\begin{aligned}\sum_{u\in J}\alpha_u\cdot u &\underset{[\text{theorem: }11.41]}{=}\sum_{u\in I\setminus\{t,e_{k+1}\}}\alpha_u\cdot u+\sum_{u\in\{e_{k+1}\}}\alpha_u\cdot u\\&\underset{[\text{theorem: }11.32]}{=}\sum_{u\in I\setminus\{t,e_{k+1}\}}\alpha_u\cdot u+\alpha_{e_{k+1}}\cdot e_{k+1}\\&=\sum_{u\in I\setminus\{t,e_{k+1}\}}(-(\lambda_t)^{-1}\cdot\lambda_u)\cdot u+((\lambda_t)^{-1}\cdot(1-\lambda_{e_{k+1}}))\cdot e_{k+1}\\&\underset{[\text{theorem: }11.67]}{=}-(\lambda_t)^{-1}\cdot\sum_{u\in I\setminus\{t,e_{k+1}\}}\lambda_u\cdot u+((\lambda_t)^{-1}\cdot(1-\lambda_{e_{k+1}}))\cdot e_{k+1}\\&\underset{[\text{eq: }11.35]}{=}t\end{aligned}$$

proving by [eq: 11.36] that $t\in\operatorname{span}(T_{k+1})$

$\boldsymbol{e_{k+1}\notin I\setminus\{t\}}$**.** Then by [eq: 11.34] we have

$$t=(\lambda_t)^{-1}\cdot e_{k+1}-(\lambda_t)^{-1}\cdot\sum_{u\in I\setminus\{t\}}\lambda_u\cdot u \tag{11.37}$$

As $(I\setminus\{t\})\bigcap\{e_{k+1}\}=\varnothing$ we can define

$$J=(I\setminus\{t\})\bigcup\{e_{k+1}\}\subseteq(T_k\setminus\{t\})\bigcup\{e_{k+1}\}=T_{k+1}$$

$$\{\alpha_u\}_{u\in J}\subseteq F\text{ by }\alpha_i=\begin{cases}(\lambda_t)^{-1}\text{ if }e_{k+1}\\(-(\lambda_t)^{-1})\cdot\lambda_u\text{ if }u\in T_k\setminus\{t\}\end{cases}$$

so that

$$\begin{aligned}\sum_{u\in J}\alpha_u\cdot u &\underset{[\text{theorem: }11.41]}{=}\sum_{u\in J\setminus\{t\}}\alpha_u\cdot u+\sum_{u\in\{e_{k+1}\}}\alpha_u\cdot u\\&\underset{[\text{theorem: }11.32]}{=}\sum_{u\in J\setminus\{t\}}\alpha_u\cdot u+\alpha_{e_{k+1}}\cdot e_{k+1}\\&=\sum_{u\in J\setminus\{t\}}((-(\lambda_t)^{-1})\cdot\lambda_u)\cdot u+(\lambda_t)^{-1}\cdot e_{k+1}\\&=\sum_{u\in J\setminus\{t\}}(-(\lambda_t)^{-1})\cdot(\lambda_u\cdot u)+(\lambda_t)^{-1}\cdot e_{k+1}\\&\underset{[\text{theorem: }11.67]}{=}-(\lambda_t)^{-1}\cdot\sum_{u\in I\setminus\{t\}}\lambda_u\cdot u+(\lambda_t)^{-1}\cdot e_{k+1}\\&\underset{[\text{eq: }11.37]}{=}t\end{aligned}$$

proving that $t\in\operatorname{span}(T_{k+1})$.

So in all cases we have $t\in\operatorname{span}(T_{k+1})$ which by using [theorem: 11.93] proves that

$$\operatorname{span}\left(T_{k+1}\bigcup\{t\}\right)=\operatorname{span}(T_{k+1}) \tag{11.38}$$

AS

$$\begin{array}{rcl} T_{k+1}\bigcup\{t\} & \underset{\text{[eq: 11.33]}}{=} & ((T_k\setminus\{t\})\bigcup\{e_{k+1}\})\bigcup\{t\} \\ & = & ((T_k\setminus\{t\})\bigcup\{t\})\bigcup\{e_{k+1}\} \\ & = & T_k\bigcup\{e_{k+1}\} \\ & \supseteq & T_k \end{array}$$

we have by [theorem: 11.89]

$$V\underset{\text{[eq: 11.26]}}{=}\operatorname{span}(T_k)\leqslant\operatorname{span}\big(T_{k+1}\bigcup\{t\}\big)\underset{\text{[eq: 11.38]}}{=}\operatorname{span}(T_{k+1})\subseteq V$$

proving that that

$$\operatorname{span}\big((T\setminus T'_{k+1})\bigcup\{e_1|i\in\{1,\ldots,k+1\}\}\big)=\operatorname{span}(T_{k+1})=V$$

Finally by [eq: 11.30] we have that $\operatorname{card}(T'_{k+1})=k+1$, $t\in T$ and $T'_{k+1}=T'_k\bigcup\{t\}$ so that, as $T'_k\subseteq T$ and $t\in T$ we have $T'_{k+1}\subseteq T$, proving that $k+1\in S$ completing 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 $\operatorname{span}((T\setminus T_n)\bigcup\{e_1|i\in\{1,\ldots,n\}\})=V$. □

Corollary 11.116. *Let $\langle V,+,\cdot\rangle$ be a vector space over a field $\langle F,+,\cdot\rangle$. Suppose that $\{e_i\}_{i\in\{1,\ldots,n\}}\subseteq V$ is a finite disjoint family such that $\{e_i|i\in\{1,\ldots,n\}\}$ is linear independent 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: 11.79] that there exists a disjoint ordered family $\{t_i\}_{i\in\{1,\ldots,\operatorname{card}(T)\}}$ such that

$$T=\{t_i|i\in\{1,\ldots,\operatorname{card}(T)\}\}$$

Assume that $\operatorname{card}(T)<n$ so that $\operatorname{card}(T)+1\leqslant n$. As $\{e_i|i\in\{1,\ldots,\operatorname{card}(T)\}\}\subseteq\{e_i|i\in 1,\ldots,n\}$ a linear independent set we have by [theorem: 11.114] that

$$\{e_i|i\in\{1,\ldots,\operatorname{card}(T)\}\}\text{ is linear independent}$$

Using the Steinitz lemma [lemma: 11.115] there exists a $T_n\subseteq T$ with $\operatorname{card}(T_n)=\operatorname{card}(T)$ such that

$$\operatorname{span}\big((T\setminus T_n)\bigcup\{e_i|i\in\{1,\ldots,\operatorname{card}(T)\}\}\big)=V$$

As $T_n\subseteq T$ and $\operatorname{card}(T_n)=\operatorname{card}(T)$ we have by [theorem: 10.82] that $T_n=T$ or $T\setminus T_n=\varnothing$ so that we have

$$\operatorname{span}(\{e_i|i\in\{1,\ldots,\operatorname{card}(T)\}\})=V$$

Hence as $e_{\operatorname{card}(T)+1}\in V$ there exists a finite

$$I\subseteq\{e_i|i\in\{1,\ldots,\operatorname{card}(T)\}\}\underset{\text{[theorem: 11.76]}}{=}\{e_i|i\in\{1,\ldots,\operatorname{card}(T)+1\}\}\setminus\{e_{\operatorname{card}(T)+1}\}$$

and a $\{\alpha_u\}_{u\in I}$ such that

$$e_{\operatorname{card}(T)+1}=\sum_{u\in I}\alpha_u\cdot u$$

Using [theorem:11.101] it follows then that $\{e_i|i\in\{1,\ldots,\operatorname{card}(T)+1\}\}$ is linear dependent. By [theorem: 11.114] and the fact that $\{e_i|i\in\{1,\ldots,\operatorname{card}(T)+1\}\}\subseteq\{e_i|i\in\{1,\ldots,n\}\}$ we reach the conclusion that $\{e_i|i\in\{1,\ldots,n\}\}$ is linear dependent which contradicts the hypothesis that $\{e_i|i\in\{1,\ldots,n\}\}$ is linear independent. Hence we must have that $n\leqslant\operatorname{card}(T)$. □

Corollary 11.117. *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 $\operatorname{span}(T)=V$ then E is finite.*

Proof. Assume that E is infinite then by [theorem: 6.29] there exists a denumerable set $E'\subseteq E$. So there exists a bijection

$$e:\mathbb{N}_0\rightarrow E'\subseteq E$$

Define

$$\{e_i\}_{i\in\{1,\ldots,\mathrm{card}(T)+1\}}\subseteq V \text{ by } e_i=e(i)$$

Then $\forall i,j\in\{1,\ldots,\mathrm{card}(T)+1\}$ with $e_i=e_j$ we have $e(i)=e(j)$ so that $i=j$, hence

$$\{e_i\}_{i\in\{1,\ldots,\mathrm{card}(T)+1\}}\subseteq V \text{ is a finite disjoint family} \tag{11.39}$$

Further $\{e_i|i\in\{1,\ldots,\mathrm{card}(T)+1\}\}=e(\{1,\ldots,\mathrm{card}(T)+1\})\subseteq E'\subseteq E$ which as E is linear independent proves by [theorem: 11.114] that

$$\{e_i|i\in\{1,\ldots,\mathrm{card}(T)+1\}\} \text{ is linear independent} \tag{11.40}$$

Using [eqs: 11.39, 11.40] on [corollary: 11.116] proves that $\mathrm{card}(T)+1\leqslant\mathrm{card}(T)$ a contradiction, so E must be finite. □

11.3.3 Basis of a vector space

Definition 11.118. *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. $\mathrm{span}(B)=V$

Note 11.119. The basis that we define here is sometimes called a Hamel basis in books about vector spaces. In a Hamel basis every vector of the vector space is a finite linear combination of basis vectors. So we don't have to bother with infinite sums and convergence of these sums. Other kind of basis represent every vector as a infinite sum or a integral of the basis vectors.

Example 11.120. 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: 11.84] we have $\mathrm{span}(\varnothing)=\{0\}$ and by [example: 11.99] $\varnothing$ 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: 11.95] 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 11.121. *Let $\langle V,+,\cdot\rangle$ be a vector space over a field $\langle F,+,\cdot\rangle$ and $B\subseteq V$ then*

$$B \textit{ is a basis of } \langle V,+,\cdot\rangle$$

$$\Updownarrow$$

$$\forall v\in V \textit{ their exists a finite } I\subseteq B \textit{ and a } \textbf{unique } \{\alpha_u\}_{u\in I}\subseteq F \textit{ such that } v=\sum_{u\in I}\alpha_u\cdot e_u$$

Proof.

⇒. As $\mathrm{span}(V)=V$ and $v\in V$ there exists a finite $I\subseteq V$ 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

$$\begin{aligned}\sum_{u\in I}\gamma_u\cdot u &= \sum_{u\in I}(\alpha_u-\beta_u)\cdot u\\ &\underset{[\text{theorems: } 11.6,\ 11.38]}{=} \sum_{u\in I}\alpha_u\cdot u-\sum_{u\in I}\beta_u\cdot u\\ &= v-v\\ &= 0\end{aligned}$$

As B is linear independent we have $\forall u\in I$ that $\alpha_u-\beta_u=\gamma_u=0$ proving that

$$\{\alpha_u\}_{u\in I}\subseteq F=\{\beta_u\}_{u\in I}\subseteq F$$

⇐. 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 $\sum_{u\in I}\beta_u\cdot u=\sum_{u\in I}0\cdot u\underset{[\text{theorem: } 11.35]}{=}0$ then by uniqueness we have $\forall u\in I$ $\alpha_u=\beta_u=0$. So by definition

$$B \text{ is linear independent}$$ □

If B is a finite we have a simpler equivalence.

Theorem 11.122. *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*

$$B \text{ is a basis of } \langle V,+,\cdot\rangle$$

$$\Updownarrow$$

$$\forall v\in V \text{ their exists a } \textbf{unique } \{\alpha_u\}_{u\in B}\subseteq F \text{ such that } v=\sum_{u\in B}\alpha_u\cdot e_u$$

Proof.

⇒. If B is a basis we have by [theorem: 11.121] 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

$$\begin{aligned}\sum_{u\in B}\alpha_u\cdot u &\underset{[\text{theorem: } 11.41]}{=} \sum_{u\in B\setminus I}\alpha_u\cdot u+\sum_{u\in I}\alpha_u\cdot u\\ &= \sum_{u\in B\setminus I}0\cdot u+\sum_{u\in I}\beta_u\cdot u\\ &\underset{[\text{theorem: } 11.35]}{=} \sum_{u\in I}\beta_u\cdot u\\ &= v\end{aligned}$$

proving existence. For uniqueness 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

$$\begin{aligned}\sum_{u\in B}\gamma_u\cdot u &= \sum_{u\in B}(\alpha_u-\lambda_u)\cdot u\\ &\underset{[\text{theorems: } 11.6,\ 11.38]}{=} \sum_{u\in B}\alpha_u\cdot u-\sum_{u\in B}\lambda_u\cdot u\\ &= v-v\\ &= 0\end{aligned}$$

As B is linear independent we have $\forall u\in B$ that $\alpha_u-\lambda_u=\gamma_u=0$ proving that

$$\{\alpha_u\}_{u\in B}\subseteq F=\{\lambda_u\}_{u\in B}\subseteq F$$

⇐. 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 $\sum_{u\in I}\beta_u\cdot u=\sum_{u\in I}0\cdot u\underset{[\text{theorem: } 11.35]}{=}0$ then by uniqueness we have $\forall u\in I\ \alpha_u=\beta_u=0$. Using [theorem: 11.98] it follows then that

$$B \text{ is linear independent}$$

□

A other usefully alternative characterization of a basis is the following

Theorem 11.123. *Let $\langle V,+,\cdot\rangle$ be a vector space over a field $\langle F,+,\cdot\rangle$ and a distinct ordered family $\{e_i\}_{i\in\{1,\ldots,n\}}\subseteq V$ then for $B=\{e_i|i\in\{1,\ldots,n\}\}$ we have*

$$\operatorname{card}(B)=n$$

and

$$B \textit{ is a basis of } \langle V,+,\cdot\rangle$$

$$\Updownarrow$$

$$\forall v\in V \textit{ their exists a } \textbf{unique } \{\alpha_i\}_{i\in\{1,\ldots,n\}}\subseteq F \textit{ such that } v=\sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot e_i$$

Proof. Define

$$e:\{1,\ldots,n\}\rightarrow B \text{ by } e(i)=e_i$$

then we have:

injectivity. If $i,j\in\{1,\ldots,n\}$ with $e(i)=e(j)$ then $e_i=e(i)=e(j)=e_j$ so that by distinctness we have $i=j$.

surjectivity. If $u\in B=\{e_i|i\in\{1,\ldots,n\}\}$ there exists a $i\in\{1,\ldots,n\}$ so that $u=e_i=e(i)$.

Hence we have that

$$e:\{1,\ldots,n\}\rightarrow B \text{ and } e^{-1}:B\rightarrow\{1,\ldots,n\} \text{ are bijections} \tag{11.41}$$

so that

$$\operatorname{card}(B)=n$$

We are now ready to prove the equivalence.

⇒. Let $v\in V$. By [theorem: 11.122] there exists a **unique** $\{\alpha_u\}_{u\in B}\subseteq F$ 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

$$\begin{aligned}\sum_{i\in\{1,\ldots,n\}}\beta_i\cdot e_i &= \sum_{i\in\{1,\ldots,n\}}\alpha_{e(i)}\cdot e(i)\\ &\underset{[\text{theorem: } 11.34] \text{ and } [\text{eq: } 11.41]}{=} \sum_{u\in B}\alpha_u\cdot u\\ &= v\end{aligned}$$

proving existence. For uniqueness assume that $\{\gamma_i\}_{i\in\{1,\ldots,n\}}\subseteq F$ satisfies

$$\sum_{i\in\{1,\ldots,n\}}\gamma_i\cdot e_i$$

Define then

$$\{\lambda_u\}_{u\in B}\subseteq B \text{ by } \lambda_u=\gamma_{e^{-1}(u)}$$

then we have

$$\begin{aligned}\sum_{u\in B}\lambda_u\cdot u &= \sum_{u\in B}\gamma_{e^{-1}(u)}\cdot e(e^{-1}(u))\\ &\underset{[\text{theorem: } 11.34] \text{ and } [\text{eq: } 11.41]}{=} \sum_{i\in\{1,\ldots,n\}}\gamma_i\cdot e(i)\\ &= v\end{aligned}$$

so that by uniqueness we have $\forall u\in B$ that $\alpha_u=\lambda_u$. Now if $i\in\{1,\ldots,n\}$ then $\beta_i=\alpha_{e(i)}=\lambda_{e(i)}=\gamma_{e^{-1}(e(i))}=\gamma_i$. Hence we have $\{\beta_i\}_{i\in\{1,\ldots,n\}}\subseteq F=\{\gamma_i\}_{i\ni\{1,\ldots,n\}}\subseteq F$.

$\Leftarrow$. 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

$$\begin{aligned}\sum_{u\in B}\beta_u\cdot u &= \sum_{u\in B}\alpha_{e^{-1}(u)}\cdot e(e^{-1}(u))\\ &\underset{[\text{theorem: } 11.34]}{=} \sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot e(i)\\ &= \sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot e_i\\ &= v\end{aligned}$$

proving that $v\in\text{span}(B)$. Hence $V\subseteq\text{span}(B)\subseteq V$ so that

$$V=\text{span}(V)$$

For linear Independence, assume that $\{\gamma_u\}_{u\in B}\subseteq F$ is such that $\sum_{u\in B}\gamma_u\cdot u=0$. Define

$$\{\lambda_i\}_{i\in\{1,\ldots,n\}}\subseteq F \text{ by } \lambda_i=\gamma_{e(i)}$$

then we have

$$\begin{aligned}\sum_{i\in\{1,\ldots,n\}}\lambda_i\cdot e_i &= \sum_{i\in\{1,\ldots,n\}}\gamma_{e(i)}\cdot e(i)\\ &\underset{[\text{theorem: } 11.34]}{=} \sum_{u\in B}\gamma_u\cdot u\\ &= 0\end{aligned}$$

As for $\{\zeta_i\}_{i\in\{1,\ldots n\}}\subseteq F$ defined by $\zeta_i=0$ we have

$$\sum_{i\in\{1,\ldots,n\}}\zeta_i\cdot e_i=\sum_{i\in\{1,\ldots,n\}}0\cdot e_i\underset{[\text{theorem: } 11.35]}{=}0$$

we have by the uniqueness hypothesis that $\forall i \in \{1, \ldots, n\}$ $\lambda_i = \zeta_i = 0$. So $\forall u \in B$ we have $\gamma_u = \gamma_{e(e^{-1}(u))} = \lambda_{e^{-1}(u)} = 0$. Using [theorem: 11.98] it follows that

$$B \text{ is linear independent} \qquad \square$$

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 11.124. *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

$$\mathcal{A} = \{X \subseteq V | R \subseteq X \subseteq W \text{ and } X \text{ is linear independent}\} \tag{11.42}$$

then we have, as R is linear independent and $R \subseteq R \subseteq W$, we have that

$$R \in \mathcal{A} \tag{11.43}$$

Using [example: 3.32] we can order $\mathcal{A}$ by the inclusion operator, in other words

$$\langle \mathcal{A}, \subseteq \rangle \text{ is a partial ordered set} \tag{11.44}$$

Let $\mathcal{C} \subseteq \mathcal{A}$ be a non empty chain chain [see definition: 3.40]. If $X \in \mathcal{C}$ then $X \in \mathcal{A}$ so that $R \subseteq X \subseteq W$, hence $R \subseteq \bigcup_{X \in \mathcal{C}} X \subseteq W$ so that

$$R \subseteq B_{\mathcal{C}} \subseteq W \text{ and } \forall X \in \mathcal{C} \text{ we have } X \subseteq B_{\mathcal{C}} \text{ where } B_{\mathcal{C}} = \bigcup_{X \in \mathcal{C}} X \tag{11.45}$$

We use now mathematical induction on the size of finite sets to proof that

$$\forall A \subseteq B_{\mathcal{C}} \text{ such that } A \text{ is finite there } \exists X \in \mathcal{C} \text{ such that } A \subseteq X \tag{11.46}$$

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: 10.75] there exist a bijection

$$\beta: \{1, \ldots, 0\} = \varnothing \to A,$$

hence $A = \beta(\varnothing)$ proving that $A = \varnothing$. As $\mathcal{C} \neq \varnothing$ there exists a $X \in \mathcal{C}$ and trivially $A = \varnothing \subseteq X$. Hence $0 \in \mathcal{S}$.

$n \in \mathcal{S} \Rightarrow n + 1 \in \mathcal{S}$. Let $A \subseteq B_{\mathcal{C}}$ with $\{1, \ldots, n+1\} \approx A$ then there exist a bijection

$$\beta: \{1, \ldots, n+1\} \to A.$$

Take $A' = A \setminus \{\beta(n+1)\}$ and consider $\beta_{|\{1, \ldots, n\}}: \{1, \ldots, n\} \to A'$ 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 $i \in \{0, \ldots, n\}$, hence $y = \beta(i) = \beta_{|\{1, \ldots, n\}}(i)$.

proving that $\beta_{|\{1, \ldots, n\}}: \{1, \ldots, n\} \to A'$ 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 $A \backslash \{\beta(n+1)\} = A' \subseteq X'$ 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 if $A \subseteq B_{\mathcal{C}}$ such that A is finite we have by [theorem: 10.75] 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: 11.46] we have

$$\exists Y \in \mathcal{C} \text{ such that } I \subseteq Y \tag{11.47}$$

Further as $Y \in C \subseteq \mathcal{A}$ we have by [eq: 11.42] that

$$R \subseteq Y \subseteq W \text{ and } Y \text{ is linear independent} \tag{11.48}$$

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 $\forall u \in I$ α_u, which, as $I \subseteq B_{\mathcal{C}}$, I finite was chosen arbitrary, proves that

$$B_{\mathcal{C}} \text{ is linear independent} \tag{11.49}$$

Combining this with [eqs: 11.42, 11.45] 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: 11.43] $R \in \mathcal{A}$ 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

$$\text{Every chain } \mathcal{C} \text{ in } \mathcal{A} \text{ has a upper bound in } \mathcal{A}$$

So by Zorn's lemma [see theorem: 3.116] there exists a maximal element in $\mathcal{A}$, in other words

$$\exists B \in \mathcal{A} \text{ such that } \forall X \in \mathcal{A} \text{ we have } X \subseteq B \tag{11.50}$$

and as $B \in \mathcal{A}$ we have

$$R \subseteq B \subseteq W \text{ and } B \text{ is linear independent} \tag{11.51}$$

Let now $w \in W$ then we have either:

$w \in B$. As by [theorem: 11.88] $B \subseteq \text{span}(B)$ it follows that $w \in \text{span}(B)$.

$w \notin B$. 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: 11.50] it follows that $B \bigcup \{w\} \subseteq B$ so that $w \in B$ contradicting $w \notin B$. So we 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

$$w \in I \tag{11.52}$$

Further

$$\begin{aligned} 0 &= \sum_{u \in I} \alpha_u \cdot u \\ &\underset{\text{[theorem: 11.41]}}{=} \sum_{u \in I \backslash \{w\}} \alpha_u \cdot u + \sum_{u \in \{w\}} \alpha_u \cdot u \\ &\underset{\text{[theorem: 11.32]}}{=} \sum_{u \in I \backslash \{w\}} \alpha_u \cdot u + \alpha_w \cdot w \end{aligned} \tag{11.53}$$

Note as $I \subseteq B \bigcup \{w\}$ we have $I \backslash \{w\} \subseteq (B \bigcup \{w\}) \backslash \{w\} \subseteq B$ so that

$$I \backslash \{w\} \subseteq B \tag{11.54}$$

Assume $\alpha_w = 0$ then from the above we have $0 = \sum_{u \in I \setminus \{w\}} \alpha_u \cdot u$ which, as $I \setminus \{w\} \subseteq B$ [see eq: 11.54] 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: 11.53] proves that

$$w = (\alpha_w)^{-1} \cdot \left(- \sum_{u \in I \setminus \{w\}} \alpha_u \cdot u \right) \underset{\text{[theorems: 11.38,11.67]}}{=} \sum_{u \in I} ((\alpha_u)^{-1} \cdot (-\alpha_u)) \cdot u = \sum_{u \in I} \lambda_u \cdot u$$

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 in all cases we have $w \in \operatorname{span}(B)$ it follows that $W \subseteq \operatorname{span}(B)$, hence

$$V = \operatorname{span}(W) \subseteq_{\text{[theorem: 11.89]}} \operatorname{span}(\operatorname{span}(B)) \underset{\text{[theorem: 11.92]}}{=} \operatorname{span}(B) \subseteq V$$

proving that

$$\operatorname{span}(B) = V \tag{11.55}$$

By [eqs: 11.51, 11.55] we have that B is a basis of $\langle V, +, \cdot \rangle$ and $R \subseteq B \subseteq W$ proving the theorem. □

Corollary 11.125. *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 $\operatorname{span}(V) \underset{\text{[theorem: 11.91]}}{=} V$ we have by [theorem: 11.124] that there exist a basis B of $\langle V, +, \cdot \rangle$ with $R \subseteq B \subseteq V$. □

Corollary 11.126. *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:

$V = \{0\}$. By [example: 11.120] $\varnothing$ is a basis for V and trivially $\varnothing \subseteq W$.

$V \neq \{0\}$. As $V = \operatorname{span}(W)$ and $\operatorname{span}(\varnothing) \underset{\text{[example: 11.84]}}{=} \{0\}$ we must have $W \neq \varnothing$, also if $W = \{0\}$ then by [example: 11.85] $\operatorname{span}(W) = \{0\}$. Hence there exist a $w \in W$ such that $w \neq 0$. Using [example: 11.100] we have that $R = \{w\}$ is linear independent. As further $R = \{w\} \subseteq W$ we have by [theorem: 11.124] that there exist a basis B of $\langle V, +, \cdot \rangle$ such that $R \subseteq B \subseteq W$. □

Corollary 11.127. *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 $\operatorname{span}(V) \underset{\text{[theorem: 11.91]}}{=} V$ we have by [corollary: 11.126] that there exist a basis B of $\langle V, +, \cdot \rangle$. □

11.3.4 Dimension of a vector space

Lemma 11.128. *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

$$\operatorname{span}(B) = V \text{ and } B \text{ is linear independent and } B \text{ is finite}$$

Let A be another basis of $\langle V, +, \cdot \rangle$ then

$$\operatorname{span}(A) = V \text{ and } A \text{ is linear independent}$$

By [corollary: 11.117] it follows that A is finite. Further as A, B are finite there exists by [theorem: 11.79] finite distinct families $\{a_i\}_{i\in\{1,\ldots,\mathrm{card}(A)\}}\subseteq A\subseteq V$ and $\{b_i\}_{i\in\{1,\ldots,\mathrm{card}(B)\}}\subseteq A\subseteq V$ such that $A=\{a_i|i\in\{1,\ldots,\mathrm{card}(A)\}\}$ and $B=\{b_i|i\in\{1,\ldots,\mathrm{card}(B)\}\}$. Then as

$$V=\mathrm{span}(\{a_i|i\in\{1,\ldots,\mathrm{card}(A)\}\})\text{ and }\{b_i|i\in\{1,\ldots,\mathrm{card}(B)\}\}\text{ is linear independent}$$

we have by [corollary: 11.116] that

$$\mathrm{card}(B)\leqslant\mathrm{card}(A)$$

Likewise as

$$V=\mathrm{span}(\{b_i|i\in\{1,\ldots,\mathrm{card}(B)\}\})\text{ and }\{a_i|i\in\{1,\ldots,\mathrm{card}(A)\}\}\text{ is linear independent}$$

we have by [corollary: 11.116] that

$$\mathrm{card}(A)\leqslant\mathrm{card}(B)$$

So that

$$\mathrm{card}(A)=\mathrm{card}(B)$$ □

Corollary 11.129. *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: 11.128] B is finite a contradiction. So every basis of $\langle V,+,\cdot\rangle$ must be infinite. □

The above lemma and corollary ensures that the following definitions makes sense.

Definition 11.130. *Let $\langle V,+,\cdot\rangle$ be a vector space over a field $\langle F,+,\cdot\rangle$ then by [corollary: 11.127] there exists a basis B of $\langle V,+,\cdot\rangle$. If B is finite then we say that V is finite dimensional and if B is infinite [not finite] then we say that V is infinite dimensional. By [lemma: 11.128] and [corollary: 11.129] this definition is independent from the basis and V is either finite dimensional or infinite dimensional but not both.*

Definition 11.131. *Let $\langle V,+,\cdot\rangle$ be a vector space over a field $\langle F,+,\cdot\rangle$ then by [corollary: 11.127] there exists a basis B of $\langle V,+,\cdot\rangle$. The dimension of V noted by $\dim(V)$ is defined as follows:*

$$\dim(V)=\begin{cases}\infty\text{ if }B\text{ is infinite}\\ \mathrm{card}(B)\text{ if }B\text{ is finite}\end{cases}$$

By [lemma: 11.128] and [corollary: 11.129] the dimension is independent from the basis.

We show now that for a finite dimensional space every linear independent set that has as many elements as the the dimension is a basis.

Theorem 11.132. *Let $\langle V,+,\cdot\rangle$ be a finite dimensional vector space over a field $\langle F,+,\cdot\rangle$ with $\dim(V)=n$ where $n\in\mathbb{N}$ and $B\subseteq V$ a linear independent set with $\mathrm{card}(B)=n$ then B is a basis for $\langle V,+,\cdot\rangle$.*

Proof. As $\dim(V)=n$ there exist a basis B' with $\mathrm{card}(B')=n$. Using [theorem: 11.125] there exist a basis B'' such that $B\subseteq B''$. From [theorem: 11.128] we have that $\mathrm{card}(B'')=\mathrm{card}(B')=n$. So we have that $B\subseteq B''$ and $\mathrm{card}(B)=\mathrm{card}(B'')$, applying then [theorem: 10.82] proves that $B=B''$, hence B is a basis. □

Theorem 11.133. *Let $\langle V,+,\cdot\rangle$ be a finite dimensional vector space over a field $\langle F,+,\cdot\rangle$ with $\dim(V)=n$ where $n\in\mathbb{N}$ and $\{e_i\}_{i\in\{1,\ldots,n\}}\subseteq V$ a linear independent family then $\{e_i|i\in\{1,...,n\}\}$ is a basis of V*

Proof. Using [corollary: 11.109] we have that $\mathrm{card}(\{x_i|i\in\{1,\ldots,n\}\})=n$ and using [theorem: 11.107] it follows that $\{x_i|i\in\{1,\ldots,n\}\}$ is linear independent. Hence using the previous theorem [theorem: 11.132]] it follows that $\{e_i|i\in\{1,\ldots,n\}\}$ is a basis of $\langle V,+,\cdot\rangle$. □

Theorem 11.134. *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 $\operatorname{span}(\{v_i|i\in\{1,\ldots,n\}\})=V$ we have that*

$$\{v_i|i\in\{1,\ldots,n\}\} \text{ is a basis of } V$$

Proof. As $\operatorname{span}(\{v_i|i\in\{1,\ldots,n\}\})=V$ there exists by [theorem: 11.126] there exists a basis $B\subseteq\{v_i|i\in\{1,\ldots,n\}\}$. As $\dim(V)=n$ we must have that $\operatorname{card}(B)=n$, further by [theorem: 10.83] $\operatorname{card}(\{v_i|i\in\{1,\ldots,n\}\})\leqslant n$ and as $B\subseteq\{v_i|i\in\{1,\ldots,n\}\}$ we have by [theorem: 10.79] $n=\operatorname{card}(B)\leqslant\operatorname{card}(\{v_i|i\in\{1,\ldots,n\}\})$. So we have that $\operatorname{card}(\{v_i|i\in\{1,\ldots,n\}\})=\operatorname{card}(B)$. Finally from [theorem: 10.82] it follows that $B=\{v_i|i\in\{1,...,n\}\}$ hence $\{v_i|i\in\{1,...,n\}\}$ is a basis for V. □

Theorem 11.135. *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: 11.124] 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: 6.42] it follows that B_W is finite. Using [theorem: 10.79] we have that

$$\dim(W)=\operatorname{card}(B_W)\leqslant\operatorname{card}(B_V)=\dim(V).$$ □

Theorem 11.136. *Let $\langle V,+,\cdot\rangle$ be a finite dimensional vector space over $\langle F,+,\cdot\rangle$ with $\dim(V)=n$ then there exists a distinct ordered family $\{e_i\}_{i\in\{1,\ldots,n\}}\subseteq V$ such that $\{e_i|i\in\{1,\ldots,n\}\}$ is a basis for V.*

Proof. As $\dim(V)=n<\infty$ there exist by definition a finite basis $B\subseteq V$ with $\operatorname{card}(B)=n$. Hence there exists a bijection $e:\{1,\ldots,n\}\rightarrow B$ which defines a distinct ordered family $\{e_i\}_{i\in\{1,\ldots,n\}}$ such that $B=e(\{1,\ldots,n\})=\{e_i|i\in\{1,\ldots,n\}\}$. □

Theorem 11.137. *Let $\langle V,+,\cdot\rangle$ a vector space over a field $\langle F,+,\cdot\rangle$, $n\in\mathbb{N}$ and $\{e_i\}_{i\in\{1,\ldots,n\}}\subseteq V$ such that $\operatorname{span}(\{e_i|i\in\{1,\ldots,n\}\})=V$ then V is finite dimensional and $\dim(V)\leqslant n$*

Proof. Using [theorem: 10.83] we have that

$$\{e_i|i\in\{1,\ldots,n\}\} \text{ is finite and } \operatorname{card}(\{e_i|i\in\{1,\ldots,n\}\})\leqslant n$$

As $\sup(\{e_i|i\in\{1,...,n\}\})$ we have by [theorem: 11.126] that there exist a basis $B\subseteq\{e_i|i\in\{1,...,n\}\}$ of V. Using [theorem: 10.79] it follows that B is finite and $\operatorname{card}(B)\leqslant\operatorname{card}(\{e_i|i\in\{1,\ldots,n\}\})\leqslant n$. Hence

$$\dim(V)=\operatorname{card}(B)\leqslant n$$ □

Example 11.138. Let $\langle\{0\},,.\rangle$ be the trivial vector space over a field $\langle F,+,\cdot\rangle$ [see: 11.54] 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: 11.120] $\varnothing$ is a basis of $\langle\{0\},+,\cdot\rangle$ so that $\dim(\{0\})=\operatorname{card}(\varnothing)=0$. 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 $\operatorname{card}(B)=\dim(X)=0$. Hence $B=\varnothing$, so by [theorem: 11.84] $V=\operatorname{span}(B)=\operatorname{span}(\varnothing)=\{0\}$. □

Theorem 11.139. *Let $\langle F,+,\cdot\rangle$ a field and consider the vector space of $\langle F,+,\cdot\rangle$ over $\langle F,+,\cdot\rangle$ [see theorem: 11.55]. 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

$$0=\sum_{i\in\{1\}}\alpha_i\cdot i\underset{[\text{theorem: } 11.32]}{=}\alpha_1\cdot 1=\alpha_1$$

it follows that $\forall i\in\{1\}$ we have that $\alpha_i=0$ 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

$$v=v\cdot 1\underset{[\text{theorem: } 11.32]}{=}\sum_{i\in\{1\}}\alpha_i\cdot e_i$$

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 11.140. *We have [see: 11.56]*

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 11.141. *The vector space* $\langle\mathbb{C},+,\cdot\rangle$ *over* $\langle\mathbb{R},+,\cdot\rangle$ *[see example: 11.57] 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

$$\operatorname{span}(\{1,i\})=\mathbb{C}$$ □

Now for a basis of higher dimensional space we introduce the Kronecker delta.

Definition 11.142. (Kronecker delta) *Let* $n\in\mathbb{N}_0$ *and* $\langle F,+,\cdot\rangle$ *a field with additive neutral element* 0 *and multiplicative neutral element 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 11.143. *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

$$\begin{aligned}
\sum_{i \in \{1, \ldots, n\}} \delta_{i,j} \cdot \alpha_i &\underset{[\text{theorem: } 11.41]}{=} \sum_{i \in \{1, \ldots, n\} \backslash \{j\}} \delta_{i,j} \cdot \alpha_i + \sum_{i \in \{j\}} \delta_{i,j} \cdot \alpha_i \\
&\underset{[\text{theorem: } 11.32]}{=} \sum_{i \in \{1, \ldots, n\} \backslash \{j\}} \delta_{i,j} \cdot \alpha_i + \delta_{j,j} \cdot \alpha_j \\
&= \sum_{i \in \{1, \ldots, n\} \backslash \{j\}} (0 \cdot \alpha_i) + 1 \cdot \alpha_j \\
&\underset{[\text{theorem: } 11.35]}{=} \alpha_j
\end{aligned}$$

□

We have a similar theorem for vector spaces.

Theorem 11.144. *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

$$\begin{aligned}
\sum_{i \in \{1, \ldots, n\}} \delta_{i,j} \cdot x_i &\underset{[\text{theorem: } 11.41]}{=} \sum_{i \in \{1, \ldots, n\} \backslash \{j\}} \delta_{i,j} \cdot x_i + \sum_{i \in \{j\}} \delta_{i,j} \cdot x_i \\
&\underset{[\text{theorem: } 11.32]}{=} \sum_{i \in \{1, \ldots, n\} \backslash \{j\}} \delta_{i,j} \cdot x_i + \delta_{j,j} \cdot x_j \\
&= \sum_{i \in \{1, \ldots, n\} \backslash \{j\}} (0 \cdot x_i) + 1 \cdot x_j \\
&\underset{[\text{theorem: } 11.35]}{=} x_j
\end{aligned}$$

□

Theorem 11.145. *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: 11.61] define now $\{e_i\}_{i \in \{1, \ldots, n\}} \subseteq F^n$ by $e_i = ((e_i)_1, \ldots, (e_i)_n)$ where $(e_i)_j = \delta_{i,j}$ then*

$$\{e_i | i \in \{1, \ldots, n\}\} \textit{ is a basis for } \langle F^n, +, \cdot \rangle \text{ so that } \dim(F^n) = n$$

Proof. Let $i, j \in \{1, \ldots, n\}$ such that $e_i = e_j$. If $i \neq j$ then $1 = \delta_{i,i} = (e_i)_i = (e_j)_i = \delta_{j,i} = 0$ a contradiction so we must have that $i = j$. Hence

$$\{e_i\}_{i \in \{1, \ldots, n\}} \subseteq F^n \text{ is a ordered distinct family}$$

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 $\forall j \in \{1, \ldots, n\}$

$$\begin{aligned}
\left(\sum_{i \in \{1, \ldots, n\}} x_i \cdot e_i \right)_j &\underset{[\text{theorem: } 11.44]}{=} \sum_{i \in \{1, \ldots, n\}} (x_i \cdot e_i)_j \\
&= \sum_{i \in \{1, \ldots, n\}} x_i \cdot (e_i)_j \\
&= \sum_{i \in \{1, \ldots, n\}} x_i \cdot \delta_{i,j} \\
&\underset{[\text{theorem: } 11.143]}{=} x_j
\end{aligned}$$

proving that

$$\sum_{i\in\{1,\ldots,n\}} x_i\cdot e_i = x$$

Further if $\{y_i\}_{i\in\{1,\ldots,n\}}\subseteq F$ is such that $\sum_{i\in\{1,\ldots,n\}} y_i\cdot e_i$ then for every $j\in\{1,\ldots,n\}$ we have

$$\begin{aligned}
x_j &= \left(\sum_{i\in\{1,\ldots,n\}} y_i\cdot e_i\right)_j\\
&\underset{\text{[theorem: 11.44]}}{=} \sum_{i\in\{1,\ldots,n\}} (y_i\cdot e_i)_j\\
&= \sum_{i\in\{1,\ldots,n\}} u_i\cdot (e_i)_j\\
&= \sum_{i\in\{1,\ldots,n\}} y_i\cdot \delta_{i,j}\\
&\underset{\text{[theorem: 11.143]}}{=} y_j
\end{aligned}$$

proving that $\{x_i\}_{i\in\{1,\ldots,n\}}\subseteq F = \{y_i\}_{i\in\{1,\ldots,n\}}\subseteq F$. Using [theorem: 11.123] it follows that

$$\{e_i|i\in\{1,\ldots,n\}\} \text{ is a basis for } \langle F^n,+,\cdot\rangle \text{ and } \dim(F^n)=\operatorname{card}(\{e_i|i\in\{1,\ldots,n\}\})=n \qquad \square$$

11.3.5 Internal Direct Sum

Definition 11.146. *Let $\langle V,+,\cdot\rangle$ be a vector space over a field $\langle F,+,\cdot\rangle$, X,Y sub spaces of V then we say that V is the **internal direct sum** of X and Y, noted as*

$$V = X\oplus Y$$

if

$$V=\{x+y|x\in X\wedge y\in Y\}\textit{and } X\bigcap Y=\{0\}$$

Later in this book we introduce a more general concept of the **direct sum of vector spaces**, enabling us to build new vector spaces from existing vector spaces. The idea of a vector spaces that is the **internal direct sum of sub spaces** is that vectors in the vector space can be expressed as a **unique sum** of vectors in the sub-spaces. This is expressed in the following theorem.

Theorem 11.147. *Let $\langle V,+,\cdot\rangle$ be a vector space over a field $\langle F,+,\cdot\rangle$, X,Y sub-spaces of V then*

$$V=X\oplus Y$$
$$\Updownarrow$$
$$\forall v\in V \textit{ there exists } \boldsymbol{unique}\ (x,y)\in X\times Y \textit{ such that } v=x+y$$

Proof.

- **⇒.** Let $v\in V$ then, as $V=X\oplus Y$ we have $V=\{x+y|x\in X\wedge y\in Y\}$ there exists a $x\in X$ and a $y\in Y$ [so that $(x,y)\in X\times Y$] such that $v=x+y$, proving existence. Now for uniqueness assume that $\exists(x',y')\in X\times Y$ such that $v=x'+y'$, Then we have $x+y=x'+y'$ or $x-x'=y-y'$. As X is a subspace we have that $x-x'\in X$ hence $y-y'\in X$, likewise as Y is a sub-space we have $y-y'\in Y$ proving that $y-y'\in X\bigcap Y=\{0\}$. So $x-x'=y-y'=0$ proving that $x=x'$ and $y=y'$ or $(x,y)=(x',y')$.
- **⇐.** By the hypothesis we have $V=\{x+y|x\in X\wedge y\in Y\}$. Let $z\in X\bigcap Y\subseteq V$ then, as X,Y are sub spaces we have that $0\in X,Y$, we have $z=z+0=0+z$ where $(0,z),(z,0)\in X\times Y$. Using the uniqueness hypothesis we have $(0,z)=(z,0)$ proving that $z=0$. □

Example 11.148. 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 11.149. *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: 11.127] a basis $B_x\subseteq X$ for $\langle X,+,\cdot\rangle$. As $X\subseteq V\Rightarrow B_X\subseteq V$ and B_X is linear independent, there exists by [theorem: 11.125] a basis B for $\langle V,+,\cdot\rangle$ such that $B_X\subseteq V$. 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: 11.148] 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: 11.90] 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

$$v=\sum_{i\in I}\alpha_i\cdot i\underset{\text{[theorem: 11.41]}}{=}\sum_{i\in I\bigcap B_X}\alpha_i\cdot i+\sum_{i\in I\setminus B_X}\alpha_i\cdot i=x+y$$

where $x=\sum_{i\in I\setminus B_x}\alpha_i\cdot i\in\operatorname{span}(I\setminus B_X)$ and $y\in\operatorname{span}(I\bigcap B_X)$. As $I\setminus B_X\subseteq B\setminus B_x$ and $I\bigcap B_X\subseteq B_X$ we have by [theorem: 11.89] 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 $V\subseteq\{x+y|x\in X\wedge y\in Y\}\subseteq V$ hence

$$V=\{x+y|x\in X\wedge y\in Y\}$$

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 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

$$\sum_{i\in I}\alpha_i\cdot i=z=\sum_{i\in J}\beta_i\cdot i \tag{11.56}$$

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

$$\begin{aligned}\sum_{i\in I\bigcup J}\gamma_i\cdot i &\underset{\text{[theorem: 11.41]}}{=} \sum_{i\in I}\gamma_i\cdot i+\sum_{i\in J}\gamma_i\cdot i\\ &= \sum_{i\in I}\alpha_i\cdot i+\sum_{i\in J}(-\beta)\cdot i\\ &\underset{\text{[theorem: 11.38]}}{=} \sum_{i\in I}\alpha_i\cdot i-\sum_{i\in J}\beta_i\cdot i\\ &= 0\end{aligned}$$

As $I\bigcup J\subseteq B_X\bigcup(B\setminus B_X)=B$ a linear independent set we have that $\forall i\in I\bigcup J$ that $\gamma_i=0$, hence $\forall i\in I$ we have $\alpha_i=\gamma_i=0$, so $z=\sum_{i\in I}\alpha_i\cdot i=\sum_{i\in I}0\cdot i\underset{\text{[theorem: 11.35]}}{=}0$. So $X\bigcap Y\subseteq\{0\}$ proving as $\{0\}\subseteq X\bigcap Y$ that

$$V=X\oplus Y$$

□

Theorem 11.150. *Let $\langle V,+,\cdot\rangle$ be a vector space over a field $\langle F,+,\cdot\rangle$, X,Y a sub-spaces of $\langle V,+,\cdot\rangle$, $\langle Y,+,\cdot\rangle$ then there exists mappings*

$$\pi_X\colon V\to X\text{ and }\pi_Y\colon V\to Y$$

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. $(\pi_X)_{|X} = \operatorname{Id}_X$, $(\pi_X)_{|Y} = C_0$, $(\pi_Y)_{|X} = C_0$ *and* $(\pi_Y)_{|Y} = \operatorname{Id}_Y$

Proof. Let $v \in V = X \oplus Y$ then by [theorem: 11.147] there exists **unique** $v_X \in X$ and $v_Y \in Y$ such that $v = v_1 + v_2$. This allows us to define $\pi_X: V \rightarrow X$ and $\pi_Y: V \rightarrow Y$ by $\pi_X(v) = v_X$ and $\pi_Y(v)$ 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

$$\pi_X \in \operatorname{Hom}(V, X) \text{ and } \pi_Y \in \operatorname{Hom}(V, Y)$$

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

$$\forall v \in X\ v_X(v) = v = \operatorname{Id}_X(v),\ v_Y(v) = 0 = C_0(v) \text{ and } \forall v \in Y\ v_X(v) = 0 = C_0(v),\ v_Y(v) = v = \operatorname{Id}_Y(v)$$

proving that

$$(v_X)_{|X} = \operatorname{id}_X \wedge (v_X)_{|Y} = C_0 \wedge (v_Y)_{|Y} = \operatorname{Id}_Y \wedge (v_Y)_{|X} = C_0$$

□

Theorem 11.151. *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 $L \in \operatorname{Hom}(X, Z)$ then there exists a $K \in \operatorname{Hom}(V, Z)$ such that $K_{|X} = L$. In other words K is a extension of L.*

Proof. Using [theorem: 11.149] there exists a sub-space Y of $\langle V, +, \cdot\rangle$ such that

$$V = X \oplus Y$$

Using [theorem: 11.150] 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. $(\pi_X)_{|X} = \operatorname{Id}_X$, $(\pi_X)_{|Y} = C_0$, $(\pi_Y)_{|X} = C_0$ and $(\pi_Y)_{|Y} = \operatorname{Id}_Y$

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

$$K_{|X} = L$$ □

11.4 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 used later in Banach spaces, Hilbert spaces, differential analysis and so on.

Definition 11.152. (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 $L: X \rightarrow Y$ 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

$$\mathrm{Hom}\,(X, Y) = \{L \in Y^X | L: X \rightarrow Y \textit{ is a linear mapping}\}$$

*If $L \in \mathrm{Hom}(X, X)$ then $L: X \rightarrow Y$ is called a **linear transformation**.*

Theorem 11.153. *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 11.154. *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 $L: X \rightarrow Y$ that*

$$L: X \rightarrow Y \textit{ is a linear mapping}$$
$$\Updownarrow$$
$$\forall x, y \in X \textit{ and } \alpha \in F \textit{ we have } L(x + \alpha \cdot y) = L(x) + \alpha \cdot L(y)$$

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

$$L(\alpha \cdot x) = L(0 + \alpha \cdot x) = L(0) + \alpha \cdot L(x) \underset{[\text{theorem: } 11.153]}{=} 0 + \alpha \cdot L(x) = \alpha \cdot L(x)$$

proving by definition that $L: X \rightarrow Y$ is linear. □

Example 11.155. 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: 11.61] 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

$$\pi_i(\alpha \cdot x) = \alpha \cdot x_i = \alpha \cdot \pi_i(x)$$ □

Definition 11.156. (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 11.157. 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: 11.145, 11.141] then $\mathcal{C}: \mathbb{R}^2 \rightarrow \mathbb{C}$ defined by $C((x, y)) = x + i \cdot y$ 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: 10.66] $x = x'$ 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 11.158. *Let $\langle X, +, \cdot \rangle$ and $\langle Y, +, \cdot \rangle$ be vector spaces over the field $\langle F, +, \cdot \rangle$ and*

$$L: X \rightarrow Y \textit{ a linear isomorphism}$$

then

$$L^{-1}: Y \rightarrow X \textit{ is a linear isomorphism}$$

Proof. Using [theorem: 2.71] we have that

$$L^{-1}: Y \rightarrow X \text{ is a bijection}$$

Let $x, y \in Y$ and $\alpha \in F$ then we have

$$\begin{aligned}
L^{-1}(x) + L^{-1}(y) &= L^{-1}(L(L^{-1}(x) + L^{-1}(y))) \\
&= L^{-1}(L(L^{-1}(x)) + L(L^{-1}(y))) \\
&= L^{-1}(x + y)
\end{aligned}$$

Likewise

$$\begin{aligned}\alpha\cdot L^{-1}(x) &= L^{-1}(L(\alpha\cdot L^{-1}(x)))\\ &= L^{-1}(\alpha\cdot L(L^{-1}(x)))\\ &= L^{-1}(\alpha\cdot x)\end{aligned}$$

□

We look now at the composition of linear mappings.

Theorem 11.159. *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

$$L_2\circ L_1: X\rightarrow Z$$

$$\begin{aligned}(L_2\circ L_1)(x+y) &= L_2(L_1(x+y))\\ &= L_2(L_1(x)+L_1(y))\\ &= L_2(L_1(x))+L_2(L_1(y))\\ &= (L_2\circ L_1)(x)+(L_2\circ L_1)(y)\\ (L_2\circ L_1)(\alpha\cdot x) &= L_2(L_1(\alpha\cdot x))\\ &= L_2(\alpha\cdot L_1(x))\\ &= \alpha\cdot L_2(L_1(x))\\ &= \alpha\cdot(L_2\circ L_1)(x)\end{aligned}$$

□

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.

Theorem 11.160. *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: 11.159] we have that the following mapping is well defined

$$\circ:\mathrm{Hom}(X,X)\times\mathrm{Hom}(X,X)\rightarrow\mathrm{Hom}(X,X)\text{ where }\circ(L,K)=L\circ K$$

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 11.161. *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]

Corollary 11.162. *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 $L_1: X\rightarrow Y$ and $L_2: Y\rightarrow Z$ are linear isomorphism then $L_2\circ L_1: X\rightarrow Z$ is a linear isomorphism.*

Proof. This follows from [theorems: 2.73 and 11.159] □

In [theorem: 11.59] it is proved that $\langle Y^X,+,\cdot\rangle$ is a vector space where for $f,g\in Y^X$, $\alpha\in F$:

1. $f+g: X\rightarrow Y$ is defined by $(f+g)(x)=f(x)+g(x)$
2. $\alpha\cdot f: X\rightarrow Y$ is defined by $(\alpha\cdot f)(x)=\alpha\cdot f(x)$
3. $C_0=0: X\rightarrow Y$ is defined by $0(x)=0$
4. $(-f): X\rightarrow Y$ 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 11.163. *Let $\langle X,+,\cdot\rangle$ and $\langle Y,+,\cdot\rangle$ be vector spaces over the field $\langle F,+,\cdot\rangle$ then $\mathrm{Hom}(X,Y)$ is a sub-space of $\langle Y^X,+,\cdot\rangle$. Applying then [theorem: 11.52] proves that $\langle \mathrm{Hom}(X,Y),+,\cdot\rangle$ is vector space over the field $\langle F,+,\cdot\rangle$.*

Proof. Let $L_1,L_2\in\mathrm{Hom}(X,Y)$, $\alpha,\beta\in F$ then for $x,y\in X$ and $\gamma\in F$ we have

$$\begin{aligned}
(\alpha\cdot L_1+\beta\cdot L_2)(x+y) &= (\alpha\cdot L_1)(x+y)+(\beta\cdot L_2)(x+y)\\
&= \alpha\cdot L_1(x+y)+\beta\cdot L_2(x+y)\\
&= \alpha\cdot(L_1(x)+L_1(y))+\beta\cdot(L_2(x)+L_2(y))\\
&= \alpha\cdot L_1(x)+\alpha\cdot L_1(y)+\beta\cdot L_2(x)+\beta\cdot L_2(y)\\
&= \alpha\cdot L_1(x)+\beta\cdot L_2(x)+\alpha\cdot L_1(y)+\beta\cdot L_2(y)\\
&= (\alpha\cdot L_1)(x)+(\beta\cdot L_2)(y)+(\alpha\cdot L_1)(y)+(\beta\cdot L_2)(y)\\
&= (\alpha\cdot L_1+\beta\cdot L_2)(x)+(\alpha\cdot L_1+\beta\cdot L_2)(y)\\
(\alpha\cdot L_1+\beta\cdot L_2)(\gamma\cdot x) &= (\alpha\cdot L_1)(\gamma\cdot x)+(\beta\cdot L_2)(\gamma\cdot x)\\
&= \alpha\cdot L_1(\gamma\cdot x)+\beta\cdot L_2(\gamma\cdot x)\\
&= \alpha\cdot(\gamma\cdot L_1(x))+\beta\cdot(\gamma\cdot L_2(x))\\
&= \gamma\cdot(\alpha\cdot L_1(x))+\gamma\cdot(\beta\cdot L_2(x))\\
&= \gamma\cdot((\alpha\cdot L_1(x)))+\gamma\cdot((\beta\cdot L_2)(x))\\
&= \gamma\cdot((\alpha\cdot L_1)(x)+(\beta\cdot L_2)(x))\\
&= \gamma\cdot(\alpha\cdot L_1+\beta\cdot L_2)(x)\\
0(x+y) &= 0\\
&= 0+0\\
&= 0(x)+0(y)\\
0(\gamma\cdot x) &= 0\\
&= \gamma\cdot 0\\
&= \gamma\cdot 0(x)
\end{aligned}$$

proving that

$$\alpha\cdot L_1+\beta\cdot L_2\in\mathrm{Hom}(X,Y) \text{ and } 0\in\mathrm{Hom}(X,Y)\Rightarrow\mathrm{Home}(X,Y)$$

hence by definition we have that $\mathrm{Hom}(X,Y)$ is a sub-space of $\langle Y^X,+,\cdot\rangle$ □

Definition 11.164. (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 $\mathrm{Hom}(X,F)$ is called the **dual space** of X and noted as X^*.*

Theorem 11.165. *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 $L\in \operatorname{Hom}(X,F)=X^{\ast}$ such that $L(x)=1$*

Proof. As $x\neq 0$ we have by [theorem: 11.100] that $\{x\}$ is a linear independent, using [theorem: 11.125] there exists a basis B of $\langle V,+,\cdot\rangle$ such that $\{x\}\subseteq B$. Let $y\in X$ then by [theorem: 11.121] 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 $\alpha_w=\left\{\begin{smallmatrix}1 \text{ if } w=x\\ 0 \text{ if } w\in B\setminus\{x\}\end{smallmatrix}\right.$ that

$$\begin{aligned}
\sum_{w\in B}\alpha_w\cdot w &\underset{[\text{theorem: } 11.41]}{=} \sum_{w\in B\setminus\{x\}}\alpha_w\cdot w+\sum_{w\in\{x\}}\alpha_w\cdot w\\
&\underset{[\text{theorem: } 11.32]}{=} \sum_{w\in B\setminus\{x\}}\alpha_w\cdot w+\alpha_x\cdot x\\
&= \sum_{w\in B\setminus\{x\}}0\cdot w+1\cdot x\\
&\underset{[\text{theorem: } 11.35]}{=} 1\cdot x\\
&= x\\
&= \sum_{w\in B}\alpha_w^{(x)}\cdot w
\end{aligned}$$

so that by uniqueness we have $\{\alpha_w\}_{w\in B}\subseteq X=\{\alpha_w^{(x)}\}_{w\in B}\subseteq X$ then $L(x)=\alpha_x^{(x)}=\alpha_x=1$ or

$$L(x)=1 \tag{11.57}$$

Further we have if $y_1,y_2\in X$ then

$$\begin{aligned}
y_1+y_2 &= \sum_{w\in B}\alpha_w^{(y_1)}\cdot w+\sum_{w\in B}\alpha_w^{(y_2)}\cdot w\\
&\underset{[\text{theorem: } 11.36]}{=} \sum_{w\in B}\left(\alpha_w^{(y_1)}+\alpha_w^{(y_2)}\right)\cdot w\\
&= \sum_{w\in B}\alpha_w^{(y_1+y_2)}\cdot w
\end{aligned}$$

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

$$L(y_1+y_2)=L(y_1)+L(y_2) \tag{11.58}$$

Finally if $y\in X$ and $\alpha\in F$ we have

$$\begin{aligned}
\alpha\cdot w &= \alpha\cdot\sum_{w\in B}\alpha_w^{(y)}\cdot w\\
&\underset{[\text{theorem: } 11.67]}{=} \sum_{w\in B}\alpha\cdot\left(\alpha_w^{(y)}\cdot w\right)\\
&= \sum_{w\in B}\left(\alpha\cdot\alpha_w^{(y)}\right)\cdot w\\
&= \sum_{w\in B}\alpha_w^{(\alpha\cdot y)}\cdot w
\end{aligned}$$

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

$$L(\alpha\cdot y)=\alpha\cdot L(y) \tag{11.59}$$

The theorem is then proved by [eqs: 11.57, 11.58, 11.59]. □

Theorem 11.166. *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 $\forall L\in\operatorname{Hom}(X,Y)$ 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: 11.165] there exists a $L\in\mathrm{Hom}(X,Y)$ such that $L(x-y)=1$, so 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 11.167. *Let $\langle X,+,\cdot\rangle$ and $\langle Y,+,\cdot\rangle$ be vector spaces over a field $\langle F,+,\cdot\rangle$ and $L\in\mathrm{Hom}(X,Y)$ is 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

$$S=\left\{n\in\mathbb{N}_0|\text{If } \{\alpha_i\}_{i\in\{1,\ldots,n\}}\subseteq F \text{ and } \{x_i\}_{i\in\{1,\ldots,n\}}\subseteq X \text{ are families then } L\left(\sum_{i=0}^{n}\alpha_i\cdot x\right)=\sum_{i=0}^{n}\alpha_i\cdot L(x_i)\right\}$$

then we have:

$0\in S$. If $\{\alpha_i\}_{i\in\{0,\ldots,0\}}\subseteq F$ and $\{x_i\}_{i\in\{0,\ldots,n\}}\subseteq X$ are families then we have

$$\begin{aligned}L\left(\sum_{i=0}^{0}a_i\cdot x_i\right)&\underset{\text{def}}{=}L(a_0\cdot x_o)\\&=\alpha_0\cdot L(x_0)\\&\underset{\text{def}}{=}\sum_{i=0}^{0}\alpha_i\cdot L(x_i)\end{aligned}$$

$n\in S\Rightarrow n+1\in S$. 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

$$\begin{aligned}L\left(\sum_{i=0}^{n+1}\alpha_i\cdot x_i\right)&\underset{\text{def}}{=}L\left(\sum_{i=0}^{n}\alpha_i\cdot x_i+\alpha_{n+1}\cdot x_{n+1}\right)\\&=L\left(\sum_{i=0}^{n}\alpha_i\cdot x_i\right)+\alpha_{n+1}\cdot x_{n+1}\\&\underset{n\in S}{=}\sum_{i=0}^{n}\alpha_i\cdot x_i+\alpha_{n+1}\cdot x_{n+1}\\&\underset{\text{def}}{=}\sum_{i=0}^{n+1}\alpha_i\cdot x_i\end{aligned}$$

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

$$\begin{aligned}L\left(\sum_{i=n}^{m}\alpha_i\cdot x_i\right) &\underset{\text{def}}{=} L\left(\sum_{i=0}^{m-n}\alpha_{i+n}\cdot x_{i+n}\right)\\ &\underset{(1)}{=} \sum_{i=0}^{m-n}\alpha_{i+n}\cdot L(x_{i+n})\\ &\underset{\text{def}}{=} \sum_{i=n}^{m}\alpha_i\cdot L(x_i)\end{aligned}$$

3. If W is finite we have either:

$W=\varnothing$. Then $\sum_{w\in W}\alpha_w\cdot w=0$ so that

$$L\left(\sum_{w\in W}\alpha_w\cdot w\right)=L(0)\underset{[\text{theorem: }11.153]}{=}0=\sum_{w\in W}\alpha_w\cdot L(w)$$

$W\neq\varnothing$. Then there exist by definition a $n\in\mathbb{N}$ and a $\beta:\{0,\ldots,n-1\}\rightarrow W$ such that

$$\sum_{w\in W}\alpha_w\cdot w=\sum_{i=0}^{n-1}\alpha_{w(i)}\cdot w(i)$$

so that

$$\begin{aligned}L\left(\sum_{w\in W}\alpha_w\cdot w\right) &= L\left(\sum_{i=0}^{n-1}\alpha_{w(i)}\cdot w(i)\right)\\ &\underset{(1)}{=} \sum_{i=0}^{n-1}\alpha_{w(i)}\cdot L(w(i))\\ &\underset{[\text{theorem: }11.34]}{=} \sum_{w\in W}\alpha_w\cdot L(w)\end{aligned}$$

□

Example 11.168. Consider the vector space $\langle\mathbb{C},+,\cdot\rangle$ then we have that

1. $\mathrm{Img}:\mathbb{C}\rightarrow\mathbb{R}$ defined by $\mathrm{Img}(x+i\cdot y)=y$ is linear
2. $\mathrm{Re}:\mathbb{C}\rightarrow\mathbb{R}$ 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: 10.69].
2. This follows from [theorem: 10.69].
3. This follows from (1), (2) and [theorem: 11.167] □

Definition 11.169. (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 11.170. (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* $\mathrm{range}(L)$ *is defined by*

$$\mathrm{range}(L)=L(X)$$

Theorem 11.171. *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 $L(0) \underset{[\text{theorem: } 11.153]}{=} 0$ 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

$$L(x+y)=L(x)+L(y)=0+0=0$$

and

$$L(\alpha\cdot x)=\alpha\cdot L(x)=\alpha\cdot 0=0$$

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 $x,y\in\mathrm{range}(L)=L(X)$ 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

$$x+y\in L(X)=\mathrm{range}(L)$$

Likewise for $\alpha\cdot x'\in X$ we have

$$L(\alpha\cdot x')=\alpha\cdot L(x')=\alpha\cdot x$$

proving that

$$\alpha\cdot x\in L(X)=\mathrm{range}(L)$$

□

As $\mathrm{range}(L)$ is a vector space the following definition make sense.

Definition 11.172. (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 $L\in\mathrm{Hom}(X,Y)$ then the rank of L noted as $\mathrm{rank}(L)$ is defined by*

$$\mathrm{rank}(L)=\dim(\mathrm{range}(L))=\dim(L(X))$$

Theorem 11.173. *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*

$$L \text{ is injective} \Leftrightarrow \ker(L)=\{0\}$$

Proof.

⇒. As $L(0) \underset{[\text{theorem: } 11.153]}{=} 0$ we have $0\in\ker(L)$ so that $\{0\}\subseteq\ker(L)$. Let $x\in\ker(L)$ then $L(x)=0 \underset{[\text{theorem: } 11.153]}{=} L(0)$. 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 11.174. *Let $\langle X,+,\cdot\rangle$ and $\langle Y,+,\cdot\rangle$ be vector spaces over a field $\langle F,+,\cdot\rangle$, $W\subseteq X$ and $L\in\mathrm{Hom}(X,Y)$ then*

$$L(\mathrm{span}(W))\subseteq\mathrm{span}(L(W))$$

Further if L is injective [or equivalently $\ker(L) = \{0\}$*] then*

$$L(\mathrm{span}(W)) = \mathrm{span}(L(W))$$

Proof. For the first part. If $y \in L(\mathrm{span}(W))$ then $\exists x \in \mathrm{span}(W)$ such that $y = L(x)$. As $x \in \mathrm{span}(W)$ there exists by [theorem: 11.87] a family $\{w_i\}_{i \in \{1,\ldots,n\}} \subseteq W$ and $\{\alpha_i\}_{i \in 1,\ldots,n} \subseteq F$ such that

$$x = \sum_{i \in \{1,\ldots,n\}} \alpha_i \cdot w_i$$

Define

$$\{u_i\}_{i \in \{1,\ldots,n\}} \subseteq L(W) \text{ by } u_i = L(w_i)$$

then we have

$$\begin{aligned} L(x) &= L\left(\sum_{i \in \{1,\ldots,n\}} \alpha_i \cdot w_i\right) \\ &\underset{\text{[theorem: 11.167]}}{=} \sum_{i \in \{1,\ldots,n\}} \alpha_i \cdot L(w_i) \end{aligned}$$

By [theorem: 11.87] it follows that $y = L(x) \in \mathrm{span}(W)$ proving that

$$L(\mathrm{span}(W)) \subseteq \mathrm{span}(L(W)) \tag{11.60}$$

For the second part. Let $y \in \mathrm{span}(L(W))$ then there exists a finite $J \subseteq L(W)$ and a $\{\alpha_i\}_{i \in J} \subseteq F$ such that

$$y = \sum_{u \in J} \alpha_u \cdot u \tag{11.61}$$

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 $L(i) = L(j)$, as L is injective we have that $i = j \in W$ proving that

$$I \subseteq W \tag{11.62}$$

Define

$$\beta: I \rightarrow J \text{ by } \beta(i) = L(i)$$

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 $j = L(i) = \beta(i)$.

$$\beta: I \rightarrow J \text{ defined by } \beta(i) = L(i) \text{ is a bijection} \tag{11.63}$$

Hence as J is finite we have also that

$$I \text{ is finite} \tag{11.64}$$

Define

$$\{\gamma_i\}_{i \in I} \subseteq F \text{ by } \gamma_i = \alpha_{\beta(i)} \tag{11.65}$$

By [eqs: 11.62, 11.64] and the definition of a span we have that

$$\sum_{i \in I} \gamma_i \cdot i \in \mathrm{span}(W)$$

Further

$$\begin{aligned} L\left(\sum_{i \in I} \gamma_i \cdot i\right) &\underset{\text{[theorem: 11.167]}}{=} \sum_{i \in I} \gamma_i \cdot L(i) \\ &= \sum_{i \in I} \alpha_{\beta(i)} \cdot \beta(i) \\ &\underset{\text{[theorem: 11.34]}}{=} \sum_{i \in J} \alpha_i \cdot i \\ &= y \end{aligned}$$

proving that $y\in L(\text{span}(W))$. Hence $\text{span}(L(W))\subseteq L(\text{span}(W))$. Combining this with [eq: 11.60 proves that

$$L(\text{span}(W))=\text{span}(L(W)) \qquad \square$$

Theorem 11.175. *Let $\langle X,+,\cdot\rangle$ and $\langle Y,+,\cdot\rangle$ be vector spaces over a field $\langle F,+,\cdot\rangle$, $W\subseteq X$ and $L\in\text{Hom}(X,Y)$ a injective linear mapping then if W is linear independent then $L(W)$ is linear independent.*

Proof. Let $J\subseteq L(W)$ be a finite subset, $\{\alpha_i\}_{i\in J}\subseteq F$ such that

$$\sum_{i\in J}\alpha_i\cdot i=0 \tag{11.66}$$

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 $L(i)=L(j)$, as L is injective we have that $i=j\in W$ proving that

$$I\subseteq W \tag{11.67}$$

Define

$$\beta: I\rightarrow J \text{ by } \beta(i)=L(i)$$

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 $j=L(i)=\beta(i)$.

$$\beta: I\rightarrow J \text{ defined by } \beta(i)=L(i) \text{ is a bijection} \tag{11.68}$$

Hence as J is finite we have also that

$$I \text{ is finite} \tag{11.69}$$

Define now

$$\{\gamma_i\}_{i\in I}\subseteq F \text{ by } \gamma_i=\alpha_{\beta(i)}$$

Then we have

$$\begin{aligned} L\left(\sum_{i\in I}\gamma_i\cdot i\right) &\underset{[\text{theorem: } 11.167]}{=} \sum_{i\in I}\gamma_i\cdot L(i)\\ &= \sum_{i\in I}\alpha_{\beta(i)}\cdot\beta(i)\\ &\underset{[\text{theorem: } 11.34] \text{ and } [\text{eq: } 11.68]}{=} \sum_{i\in J}\alpha_i\cdot i\\ &\underset{[\text{theorem: } 11.66]}{=} 0\end{aligned}$$

So $\sum_{i\in I}\gamma_i\cdot i\in\ker(L)$, as L is injective we have by [theorem: 11.173] that $\ker(L)=\{0\}$, so that

$$\sum_{i\in I}\gamma_i\cdot i$$

As I is finite, $I\subseteq W$ and W is linear independent we have by [theorem: 11.97] that $\forall i\in I$ we have $\gamma_i=0$. 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: 11.97] proves that $L(W)$ is linear independent. □

Corollary 11.176. *Let $\langle X,+,\cdot\rangle$ and $\langle Y,+,\cdot\rangle$ be vector spaces over a field $\langle F,+,\cdot\rangle$, $B\subseteq V$ a basis of $\langle X,+,\cdot\rangle$ and $L\in\text{Hom}(X,Y)$ a injective linear mapping then $L(B)$ is a basis of* range(L)*.*

Proof. As $B\subseteq V$ we have

$$L(B)\subseteq L(X)=\text{range}(X)$$

As B is a basis of V we have that B is linear independent and $\text{span}(B)=X$. As L is linear independent we have by [theorem: 11.175] that

$$L(B)\,\text{is linear independent}$$

As B is a basis we have $\text{span}(B)=V$ so that

$$\begin{array}{rcl}\text{range}(L) & = & L(X)\\ & = & L(\text{span}(B))\\ & \underset{[\text{theorem: }11.174]}{=} & \text{span}(L(B))\end{array}$$

So $L(B)$ is a basis of $\text{range}(L)$. □

Corollary 11.177. *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 $L\in\text{Hom}(X,Y)$ a injective linear mapping then $L:X\rightarrow Y$ is a linear isomorphism*

Proof. As X is finite dimensional there exists a basis $B\subseteq X$ with $\text{card}(B)=\dim(X)$. Further as L is injective we have that $L_{|B}:B\rightarrow L(B)$ is a bijection so that $\text{card}(L(B))=\text{card}(B)=\dim(X)$. By [theorem: 11.175] we have that $L(B)\subseteq Y$ is linear independent. Using this, the fact that $\text{card}(L(B))=\dim(X)=\dim(Y)$ it follows from [theorem: 11.132] that $L(B)$ is a basis of Y. Hence

$$L(X)=L(\text{span}(B))\underset{[\text{theorem: }11.174]}{=}\text{span}(L(B))\underset{L(B)\text{ is a basis of Y}}{=}Y$$

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 11.178. *Let $\langle X,+,\cdot\rangle$ be a finite dimensional vector space over a filed $\langle F,+,\cdot\rangle$ and $L\in\text{Hom}(X,X)$ 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 11.179. *Let $\langle X,+,\cdot\rangle$ and $\langle Y,+,\cdot\rangle$ be vector spaces over a field $\langle F,+,\cdot\rangle$ and $L\in\text{Hom}(X,Y)$ 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

$$L:X\rightarrow Y\text{ is a bijection so that }\text{range}(L)=L(X)\tag{11.70}$$

Let B be a basis for X then

$$L_{|B}:B\rightarrow L(B)\text{ is a bijection or }B\approx L(B)\tag{11.71}$$

Further by the previous theorem [theorem: 11.176]

$$L(B)\text{ is a basis for }\langle Y,+,\cdot\rangle\tag{11.72}$$

So we have

1. If X is infinite dimensional then B is infinite so that by [theorem: 6.45] and [eq: 11.71] 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: 11.71] 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 $\{1,\ldots,n\}\approx B\approx L(B)$ such that

$$\dim(X)=n=\dim(Y)$$

□

The above theorem is a special case of a more general theorem.

Theorem 11.180. *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: 11.173] L is injective so that

$$L: X\rightarrow L(X)=\operatorname{range}(L)$$

is a isomorphism, hence by [theorem: 11.179] $\langle\operatorname{range}(L),+,\cdot\rangle$ is finite dimensional and $\dim(X)=\dim(\operatorname{range}(L))=\operatorname{rank}(L)$. Further $\dim(\ker(L))=\dim(\{0\})\underset{\text{]example: 11.138]}}{=}0$ 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 linear independent] of $\langle\ker(L),+,*\rangle$ and use [theorem: 11.125] 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: 10.80] that

$$\dim(X)=\operatorname{card}(B_x)=\operatorname{card}(B_x\setminus B_k)+\operatorname{card}(B_k)=\operatorname{card}(B_x\setminus B_k)+\dim(\ker(L)) \tag{11.73}$$

Consider the function

$$L_{|B_x\setminus B_k}: B_x\setminus B_k\rightarrow L(B'\setminus B)$$

then we have

injectivity. Let $x,y\in B'\setminus B_k$ 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: 11.122] 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

$$0=\sum_{u\in B_k}\alpha_u\cdot u+1\cdot y+(-1)\cdot x \tag{11.74}$$

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

$$B_x \text{ is the disjoint union of } ((B_x\setminus B_k)\setminus\{x,y\}),\{x\},\{y\} \text{ and } B_k \tag{11.75}$$

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

$$\begin{array}{rr}
\sum_{u\in B_x}\gamma_u\cdot u & \underset{\text{[theorem: 11.41]}}{=}\\
\sum_{u\in B_k}\gamma_u\cdot u+\sum_{u\in\{x\}}\gamma_u\cdot u+\sum_{u\in\{y\}}\gamma_u\cdot u+\sum_{u\in((B_x\setminus B_k)\setminus\{x,y\})}\gamma_u\cdot u & \underset{\text{[theorem: 11.32]}}{=}\\
\sum_{u\in B_k}\gamma_u\cdot u+\gamma_x\cdot x+\gamma_y\cdot y+\sum_{u\in((B_x\setminus B_k)\setminus\{x,y\})}\gamma_u\cdot u & =\\
\sum_{u\in B_k}\alpha_u\cdot u+1\cdot x+(-1)\cdot y+\sum_{u\in((B_x\setminus B_k)\setminus\{x,y\})}0\cdot u & \underset{\text{[theorem: 11.35]}}{=}\\
\sum_{u\in B_k}\alpha_u\cdot u+1\cdot x+(-1)\cdot y & \underset{\text{[eq: 11.74]}}{=}\\
0 &
\end{array}$$

As B_x is linear independent we have that $\forall u\in B_k\ \gamma_u=0$ contradicting the fact that $\alpha_x=1$. Hence we must have that $x=y$ proving injectivity.

surjectivity. This is trivial.

So we have proved that

$$L_{|B_x\setminus B_k}: B_x\setminus B_k\rightarrow L(B\setminus B_y) \text{ is a isomorphism} \tag{11.76}$$

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

$$\begin{array}{rcl}
y & = & L(x)\\
 & = & L\left(\sum_{u\in B_x}\alpha_u\cdot u\right)\\
 & \underset{\text{[theorem: 11.167]}}{=} & \sum_{u\in B_x}\alpha_u\cdot L(u)\\
 & \underset{\text{[theorem: 11.41]}}{=} & \sum_{u\in B_x\setminus B_k}\alpha_u\cdot L(u)+\sum_{u\in B_k}\alpha_u\cdot L(u)\\
 & = & \sum_{u\in B_x\setminus B_k}\alpha_u\cdot L(u)+\sum_{u\in B_k}\alpha_u\cdot 0\\
 & \underset{\text{[theorem: 11.35]}}{=} & \sum_{u\in B_x\setminus B_k}\alpha_u\cdot L(u)
\end{array} \tag{11.77}$$

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

$$\begin{array}{rcl}
\sum_{u\in L(B_x\setminus B_k)}\beta_u\cdot u & = & \sum_{u\in L(B_x\setminus B_k)}\alpha_{(L_{|B_x\setminus B_k})^{-1}(u)}\cdot(L_{|B_x\setminus B_k}((L_{|B_x\setminus B_k})^{-1}(u)))\\
 & \underset{\text{[theorem: 11.34]}}{=} & \sum_{u\in B_x\setminus B_k}\alpha_u\cdot L_{|B_x\setminus B_k}(u)\\
 & = & \sum_{u\in B_x\setminus B_k}\alpha_u\cdot L(u)\\
 & = & y
\end{array}$$

proving that $y\in\operatorname{span}(L(B_x\setminus B_k))$. Hence

$$L(X)\subseteq\operatorname{span}(L(B_x\setminus B_k)) \tag{11.78}$$

For the opposite inclusion:

$$\begin{array}{rcl} B_x\setminus B_k\subseteq X & \Rightarrow & L(B_x\setminus B_k)\subseteq L(X) \\ & \underset{\text{[theorem: 11.89]}}{\Rightarrow} & \operatorname{span}(L(B_x\setminus B_k))\subseteq\operatorname{span}(L(X)) \\ & \underset{\text{[theorems: 11.171,11.91]}}{\Rightarrow} & \operatorname{span}(L(B_x\setminus B_k))\subseteq L(x) \end{array}$$

which combined with [eq: 11.78] gives

$$L(X)=\operatorname{span}(L(B_x\setminus B_k)) \tag{11.79}$$

As $B_x\setminus B_k\subseteq B_x$ a linear independent set it follows from [theorem: 11.114] that $B_x\setminus B_k$ is linear independent. Using [eq: 11.76] we have that $L_{|B_x\setminus B_k}: B_x\setminus B_k\rightarrow L(B\setminus B_y)$ is a isomorphism, so by [theorem: 11.175]

$$L(B_x\setminus B_k)\text{ is linear independent}$$

The above together with [eq: 11.79] prove that

$$L(B_x\setminus B_k)\text{ is a basis of }L(X)$$

Using [eq: 11.76] 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: 11.73] gives finally

$$\dim(X)=\operatorname{rank}(L)+\dim(\ker(L))$$

So in all cases we have proved that

$$\dim(X)=\operatorname{rank}(L)+\dim(\ker(L))$$ □

11.5 Permutations

In [definition: 11.24] 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 11.181. *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: 11.24] $P_n=S_{\{1,\ldots,n\}}$ but evidently P_n is a shorter notation.

Example 11.182. Let $n\in\mathbb{N}$ then for

$$\rho_n:\{1,\ldots n\}\rightarrow\{1,\ldots,\}\text{ defined by }\rho_n(i)=n-i+1$$

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 $P_n\times\{1,\ldots,n+1\}\approx P_{n+1}$ and finally we will use mathematical induction to prove that $\forall n\in\mathbb{N}$ P_n is finite and $\operatorname{card}(P_n)=n!$. First we define the faculty $n!$.

Definition 11.183. *Let* $\operatorname{fac}\colon\mathbb{N}_0\rightarrow\mathbb{N}$ *is defined by*

$$\operatorname{fac}(n)=\begin{cases}1 \textit{ if } m=0\\ n\cdot\operatorname{fac}(n-1) \textit{ if } n\in\mathbb{N}\end{cases}$$

We will use the common convention and write $n!$ *for* $\operatorname{fac}(n)$*.*

Next we extend a permutation $\sigma\in P_n$ to a permutation $\sigma^{[i]}$ of P_{n+1}.

Lemma 11.184. *Let* $n\in\mathbb{N}$*,* $i\in\{1,\ldots,n+1\}$ *and* $\sigma\in P_n$ *then define*

$$\sigma^{[i]}\colon\{1,\ldots,n+1\}\rightarrow\{1,\ldots,n+1\}$$

by

1. *If* $i=n+1$ *then*

$$\sigma^{[n+1]}(j)=\begin{cases}n+1 \textit{ if } j=n+1\\ \sigma(j) \textit{ if } j\in\{1,\ldots,n\}\end{cases}$$

2. *If* $i\in\{1,\ldots,n\}$ *take* $k=\sigma^{-1}(i)$ *[hence* $\sigma(k)=i$*]*

$$\sigma^{[i]}(j)=\begin{cases}i \textit{ if } j{=}n{+}1\\ n+1 \textit{ if } j=k\\ \sigma(j) \textit{ if } j\in\{1,\ldots,n\}\setminus\{k\}\end{cases}$$

then we have that

$$\sigma^{[i]}\in P_{n+1} \textit{ and } \sigma^{[i]}(n+1)=i$$

Proof. For $i\in\{1,\ldots,n+1\}$ we have either:

$\boldsymbol{i=n+1}$**.** Then for $\sigma^{[n+1]}\colon\{1,\ldots,n+1\}\rightarrow\{1,\ldots,n+1\}$ 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 $\sigma^{[n+1]}(n+1)=n+1\notin\{1,...,n\}$]. So $\sigma(k)=\sigma^{[n+1]}(k)=\sigma^{[n+1]}(l)=\sigma(l)$ which as σ is a bijection proves that $k=l$.

$\boldsymbol{k=n+1}$**.** If $l\in\{1,\ldots,n\}$ then $\sigma^{[n+1]}(l)=\sigma(l)\in\{1,\ldots,n\}$ so that $\sigma^{[n+1]}(l)\neq n+1=\sigma^{[n+1]}(k)$ a 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:

$\boldsymbol{j=n+1}$**.** Then $n+1\in\{1,\ldots,n+1\}$ and $\sigma^{[n+1]}(n+1)=n+1=j$

$\boldsymbol{j\in\{1,\ldots,n\}}$**.** Then, as $\sigma\colon\{1,\ldots,n\}\rightarrow\{1,\ldots,n\}$ 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 $\sigma^{[n+1]}\colon \{1, \ldots, n+1\} \rightarrow \{1, \ldots, n+1\}$ is a bijection or

$$\sigma^{[n+1]} \in P_{n+1}$$

$\boldsymbol{i \in \{1, \ldots, n\}}$**.** Let

$$k = \sigma^{-1}[i] \Rightarrow k \in \{1, \ldots, n\} \wedge \sigma(k) = i \tag{11.80}$$

then we have for $\sigma^{[i]}\colon \{1, \ldots, n+1\} \rightarrow \{1, \ldots, n+1\}$ that

injectivity. If $\sigma^{[i]}(r) = \sigma^{[i]}(s)$ then we have to consider the following cases for r, s:

$\boldsymbol{r = n+1 \wedge s = n+1}$**.** Then $r = s$

$\boldsymbol{r = k \wedge s = n+1}$**.** 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.

$\boldsymbol{r \in \{1, \ldots, n\} \setminus \{k\} \wedge s = n+1}$**.** Then

$$\sigma(k) \underset{[\text{eq: } 11.80]}{=} i = \sigma^{[i]}(n+1) = \sigma^{[i]}(s) = \sigma^{[i]}[r] = \sigma[r]$$

so that, as σ is injective, $r = k$ contradicting $r \in \{1, \ldots, n\} \setminus \{k\}$. So this case does not occur.

$\boldsymbol{r = n+1 \wedge s = k}$**.** 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.

$\boldsymbol{r = k \wedge s = k}$**.** Then $r = s$

$\boldsymbol{r \in \{1, \ldots, n\} \setminus \{k\} \wedge s = k}$**.** 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.

$\boldsymbol{r = n+1 \wedge s \in \{1, \ldots, n\} \setminus \{k\}}$**.** 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.

$\boldsymbol{r = k \wedge s \in \{1, \ldots, n\} \setminus \{k\}}$**.** 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.

$\boldsymbol{r \in \{1, \ldots, n\} \setminus \{k\} \wedge s \in \{1, \ldots, n\} \setminus \{k\}}$**.** 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:

$\boldsymbol{j = n+1}$**.** Then $\sigma^{[i]}(k) = n+1 = j$

$\boldsymbol{j = i}$**.** Then $\sigma^{[i]}(n+1) = i = j$

$\boldsymbol{j \in \{1, \ldots, n\} \setminus \{i\}}$**.** Then as $\sigma\colon \{1, \ldots, n\} \rightarrow \{1, \ldots, n\}$ is a bijection there exists $l \in \{1, \ldots, n\}$ such that $\sigma(l) = j$. If $l = k$ then $j = \sigma(l) = \sigma(k) = i$ contradicting $j \in \{1, \ldots, n\} \setminus \{i\}$. So we must have that $l \in \{1, \ldots, n\} \setminus \{k\}$, hence $\sigma^{[i]}(l) = \sigma(l) = j$

proving surjectivity.

Hence $\sigma^{[i]}: \{1,\ldots,n+1\} \rightarrow \{1,\ldots,n+1\}$ is surjective or

$$\sigma^{[i]} \in P_{n+1}$$ □

Lemma 11.185. *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

$$k_1 \underset{[\text{theorem: } 11.184]}{=} \sigma_1^{[k_1]}(n+1) = \sigma_2^{[k_2]}(n+1) \underset{[\text{theorem: } 11.184]}{=} k_2$$

proving that

$$k_1 = k_2 \tag{11.81}$$

We have now to look at two cases for k_1, k_2 corresponding with the definitions of $\sigma_1^{[k_1]}$, $\sigma_2^{[k_2]}$

$\boldsymbol{k_1 = k_2 = n+1.}$ Then if $i \in \{1,\ldots,n\}$ we have

$$\sigma_1(i) \underset{[\text{theorem: } 11.184]}{=} \sigma_1^{[k_1]}(i) = \sigma_2^{[k_2]}(i) = \sigma_2(i)$$

proving that

$$\sigma_1 = \sigma_2$$

$\boldsymbol{k_1 = k_2 \neq n+1.}$ Let $l_1 = (\sigma_1)^{-1}(k_1)$ and $l_2 = (\sigma_2)^{-1}(k_2)$ then we have

$$\sigma_1^{[k_1]}(l_1) \underset{[\text{theorem: } 11.184]}{=} n+1 \underset{[\text{theorem: } 11.184]}{=} \sigma^{[k_2]}(l_2)$$

proving as $\sigma_1^{[k_1]}, \sigma_2^{[k_2]}$ are bijections that $l_1 = l_2$. Now if $i \in \{1,...,n\}$ then we have either:

$\boldsymbol{i = l_1 = l_2.}$ Then $\sigma_1(i) = \sigma_1(l_1) \underset{[\text{theorem: } 11.184]}{=} k_1 = k_2 \underset{[\text{theorem: } 11.184]}{=} \sigma_2(l_2) = \sigma_2(i)$

$\boldsymbol{i \neq l_1 = l_2.}$ Then $\sigma_1(i) \underset{[\text{theorem: } 11.184]}{=} \sigma_1^{[k_1]}(i) = \sigma_2^{[k_2]}(i) \underset{[\text{theorem: } 11.184]}{=} \sigma_2(i)$

proving that

$$\sigma_1 = \sigma_2$$

So in all cases we have $\sigma_1 = \sigma_2$ which together with [eq: 11.81] 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:

$\boldsymbol{\rho(n+1) = n+1.}$ Take $\sigma = \rho_{|\{1,\ldots,n\}}$ then by [theorem: 11.26] $\rho_{|\{1,\ldots,n\}} \in P_n$. Further $\forall i \in \{1,\ldots,n+1\}$ we have

$$\begin{aligned} \beta(n+1, \rho_{|\{1,\ldots,n\}})(i) &= (\rho_{|\{1,\ldots,n\}})^{[n+1]}(j) \\ &\underset{[\text{theorem: } 11.184]}{=} \begin{cases} n+1 \text{ if i=n+1} \\ \rho_{|\{1,\ldots,n+1\}}(i) \text{ if } i \in \{1,\ldots,n\} \end{cases} \\ &= \begin{cases} n+1 \text{ if i=n+1} \\ \rho(i) \text{ if } i \in \{1,\ldots,n\} \end{cases} \\ &= \rho(i) \end{aligned}$$

proving

$$\beta(n+1, \rho_{|\{1,\ldots,n\}}) = \rho$$

$\rho(n+1)\neq n+1$. Then $\rho(n+1)\in\{1,\ldots,n\}$. Define

$$k=\rho(n+1) \text{ and } l=\rho^{-1}(n+1) \text{ so that } \rho(l)=n+1 \tag{11.82}$$

Then we have $k\in\{1,\ldots,n\}$ and $l\in\{1,\ldots,n\}$ [for if $l=n+1$ we have $\rho(n+1)=\rho(l)\underset{[\text{eq: } 11.82]}{=}n+1$ 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 $\rho(i)=n+1\underset{[\text{eq: } 11.82]}{=}\rho(l)$ then, as ρ is injective, $i=l$ contradicting $i\in\{1,\ldots,n\}\setminus\{l\}$. To summarize we have

$$k,l\in\{1,\ldots,n\} \text{ and } \forall i\in\{1,\ldots,n\}\setminus\{l\} \quad \rho(i)\in\{1,\ldots,n\} \tag{11.83}$$

Define now

$$\sigma:\{1,\ldots,n\}\rightarrow\{1,\ldots,n\} \text{ by } \sigma(i)=\begin{cases} k \text{ if } i=l \\ \rho(i) \text{ if } i\in\{1,\ldots,n\}\setminus\{l\}\end{cases}$$

Then we have:

injectivity. Let $i,j\in\{1,\ldots,n\}$ such that $\sigma(i)=\sigma(j)$ then we have either:

$\sigma(i)=\sigma(j)=k$. If $i\neq l$ then $i\in\{1,\ldots,n\}\setminus\{l\}$ so that $\sigma(i)=\rho(i)$, hence $\rho(i)=k=\rho(n+1)$ which as ρ is injective proves that $i=n+1$ contradicting $i\in\{1,...,n\}$, hence we must have that $i=l$. Likewise if $j\neq l$ then $j\in\{1,...,n\}\setminus\{l\}$ so that $\sigma(j)=\rho(j)$, hence $\rho(j)=k=\rho(n+1)$ which as ρ is injective proves that $j=n+1$ contradicting $j\in\{1,...,n\}$. Hence we must have that $j=l$. So we have $i=l=j$ proving injectivity in this case.

$\sigma(i)=\sigma(j)\neq k$. 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:

$j=k$. Then $\sigma(l)=k=j$ where $l\in\{1,\ldots,n\}$ proving surjectivity in this case.

$j\neq k$. Then $j\in\{1,\ldots,n\}\setminus\{k\}\subseteq\{1,\ldots,n+1\}$. By surjectivity of $\rho:\{1,\ldots,n+1\}\rightarrow\{1,\ldots,n+1\}$ there exists a $i\in\{1,\ldots,n+1\}$ such that $\rho(i)=j$. Assume that $i=n+1$ then

$$j=\rho(i)=\rho(i)=\rho(n+1)\underset{[\text{eq: } 11.82]}{=}k$$

contradicting $j\neq k$. So we must have that $i\in\{1,\ldots,n\}$. Also if $i=l$ then $j=\rho(i)=\rho(l)\underset{[\text{eq: } 11.82]}{=}n+1$ 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 $\sigma:\{1,\ldots,n\}\rightarrow\{1,\ldots,n\}$ 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:

$i=n+1$. Then we have

$$\begin{aligned}\sigma^{[k]}(i) &= \sigma^{[k]}(n+1)\\ &\underset{[\text{theorem: } 11.184]}{=} k\\ &\underset{[\text{eq: } 11.82]}{=} \rho(n+1)\\ &= \rho(i)\end{aligned}$$

***i* = *l*.** Then as $\sigma(l) = k$ so that $l = \sigma^{-1}(k)$

$$\begin{array}{ccc} \sigma^{[k]}(i) & = & \sigma^{[k]}(l) \\ & \underset{\text{[theorem: 11.184] and } l = \sigma^{-1}(k)}{=} & n+1 \\ & \underset{\text{[eq: 11.82]}}{=} & \rho(l) \\ & = & \rho(i) \end{array}$$

***i* ∈ {1, …, *n*}\{*l*}.** As $\sigma(l) = k$ we have

$$i \neq l = \sigma^{-1}(k) \text{ and } i \neq n+1 \tag{11.84}$$

so that

$$\begin{array}{ccc} \sigma^{[k]}(i) & \underset{\text{[theorem: 11.184] and [eq: 11.84]}}{=} & \sigma(i) \\ & = & \rho(i) \end{array}$$

so in all cases we have $\rho(i) = \sigma^{[k]}(i) = \beta(k, \sigma)(i)$ for every $i \in \{1, \ldots, n\}$. 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 11.186. *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 ∈ *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* ∈ *S* ⇒ *n* + 1 ∈ *S*.** As $\{1,...,n+1\}$ is finite with $\operatorname{card}(\{1,...,n+1\}) \underset{\text{[theorem: 10.78]}}{=} n+1$ and, as $n \in S$, P_n is finite with $\operatorname{card}(P_n) = n!$, so we have by [theorem:10.80] 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: 11.185] we have then that

$$\{1, \ldots, n+1\} \times P_n \approx P_{n+1}$$

so that P_{n+1} is finite and $\operatorname{card}(P_{n+1}) = (n+1)!$, proving that $n+1 \in S$. □

Theorem 11.187. *Let $n \in \mathbb{N}$ then $\langle P_n, \circ \rangle$ is a group called the permutation group*

Proof. Using [definition: 11.24] we have that $P_n = S_{\{1, \ldots, n\}}$, the rest follows from [theorem: 11.25] □

Note 11.188. 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 11.189. *Let $n\in\mathbb{N}$. As $\langle P_n,\circ\rangle$ forms a group, giving a ordered family of $\{\sigma_i\}_{i\in\{1,\ldots,n\}}$ of permutations in P_n then $(\sigma_1\circ\cdots\circ\sigma_n)$ is defined by*

$$(\sigma_1\circ\cdots\circ\sigma_n)=\prod_{i=1}^{n}\sigma_i$$

(using $\circ$ for product).

Theorem 11.190. *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

$$\sigma_1\circ\cdots\circ\sigma_1=\prod_{i=1}^{1}\sigma_1\underset{\text{[theorem: 11.14]}}{=}\sigma_1$$

2. If $1<m$ then we have

$$\begin{aligned}\sigma_1\circ\cdots\circ\sigma_m &= \prod_{i=1}^{m}\sigma_i\\ &\underset{\text{[theorem: 11.14]}}{=} \left(\prod_{i=1}^{m-1}\sigma_i\right)\circ\sigma_m\\ &= (\sigma_1\circ\cdots\circ\sigma_{m-1})\circ\sigma_m\\ \sigma_1\circ\cdots\circ\sigma_m &= \prod_{i=1}^{m}\sigma_i\\ &\underset{\text{def}}{=} \prod_{i=0}^{m-1}\sigma_{i+1}\\ &\underset{\text{[theorem: 11.21}}{=} \sigma_{0+1}\circ\prod_{i=1}^{m-1}\sigma_{i+1}\\ &\underset{\text{def}}{=} \sigma_1\circ\prod_{i=2}^{m}\sigma_i\\ &= \sigma_1\circ(\sigma_2\circ\cdots\circ\sigma_m)\end{aligned}$$

□

11.5.1 Transpositions

Definition 11.191. (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 as [see: 11.27]*

$$\left(i\underset{n}{\leftrightarrow}j\right)=\left(i\underset{\{1,\ldots,n\}}{\leftrightarrow}j\right)$$

so that

$$\left(i\underset{n}{\leftrightarrow}j\right)(k)=\begin{cases}k \text{ if } k\in\{1,\ldots,n\}\setminus\{i,j\}\\ j \text{ if } k=i\\ i \text{ if } k=j\end{cases}$$

*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 11.192. *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. $\left(i \underset{n}{\leftrightarrow} j\right) \circ \left(i \underset{n}{\leftrightarrow} j\right) = \mathrm{Id}_{\{1, \ldots, n\}}$

Proof. This is proved in [theorem: 11.28] □

We can always extend a permutation on $\{1, \ldots, n\}$ to a permutation on $\{1, \ldots, n+1\}$ as the following theorem shows.

Theorem 11.193. *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

$$\sigma^{|\{1, \ldots, m\}}(i) = \left\{ \begin{array}{l} i \text{ if } i \in \{n+1, \ldots, m\} = \{1, \ldots, m\} \setminus \{1, \ldots, n\} \\ \sigma(i) \text{ if } i \in \{1, \ldots, n\} \end{array} \right. \in \{1, \ldots, m\}$$

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 $(\sigma_1 \circ \cdots \circ \sigma_n) = ((\sigma_1)^{|\{1, \ldots, m\}} \circ \cdots \circ (\sigma_n)^{|\{1, \ldots, m\}})$*

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:

   ***i*, *j* ∈ {1, …, *n*}.** Then we have $\sigma(i) = \sigma^{|\{1, \ldots, m\}}(i) = \sigma^{|\{1, \ldots, m\}}(j) = \sigma(j)$ proving, as σ is injective, that $i = j$.

   ***i* ∈ {1, …, *n*} ∧ *j* ∈ {*n* + 1, …, *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.

   ***i* ∈ {*n* + 1, …, *m*} ∧ *j* ∈ {1, …, *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.

   ***i*, *j* ∈ {*n* + 1, …, *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 $\sigma^{|\{1, \ldots m\}}: \{1, \ldots, m\} \Rightarrow \{1, \ldots, m\}$ is a bijection or $\sigma^{|\{1, \ldots, m\}} \in P_m$

2. Let $k \in \{1, \ldots, m\}$ then we have either:

   ***k* ∈ {1, …, *n*}.** Then as $i, j \in \{1, \ldots, n\}$ we have

$$\left(i \underset{n}{\leftrightarrow} j\right)^{|\{1, \ldots, m\}}(k) = \left(i \underset{n}{\leftrightarrow} j\right)(k) = \left\{ \begin{array}{l} i \text{ if } k = j \\ j \text{ if } k = i \\ k \text{ if } i \in \{1, \ldots, n\} \setminus \{i, j\} \end{array} \right. = \left(i \underset{m}{\leftrightarrow} j\right)(k)$$

$k \in \{n+1, \ldots, m\}$. Then as $i, j \in \{1, \ldots, n\}$ we have that $i, j \neq k$ so that

$$\left(i \underset{n}{\leftrightarrow} j\right)^{|\{1,\ldots,m\}}(k) = k = \left(i \underset{m}{\leftrightarrow} j\right)(k)$$

3. Let $i \in \{1, \ldots, m\}$ then we have either:

   $i \in \{1, \ldots, n\}$. Then

$$\begin{aligned}(\sigma \circ \rho)^{|\{1,\ldots,\}}(i) &= (\sigma \circ \rho)(i) \\ &= \sigma(\rho(i)) \\ &= \sigma^{|\{1,\ldots,m\}}(\rho^{|\{1,\ldots,m\}}(i)) \\ &= (\sigma^{|\{1,\ldots,m\}} \circ \rho^{|\{1,\ldots,m\}})(i)\end{aligned}$$

   $i \in \{n+1, \ldots, m\}$. Then

$$\begin{aligned}(\sigma \circ \rho)^{|\{1,\ldots,\}}(i) &= i \\ &= \sigma^{|\{1,\ldots,m\}}(i) \\ &= \sigma^{|\{1,\ldots,m\}}(\rho^{|\{1,\ldots,m\}}(i)) \\ &= (\sigma^{|\{1,\ldots,m\}} \circ \rho^{|\{1,\ldots,m\}})(i)\end{aligned}$$

   So $\forall i \in \{1, \ldots, m\}$ we have $(\sigma \circ \rho)^{|\{1,\ldots,\}}(i) = (\sigma^{|\{1,\ldots,m\}} \circ \rho^{|\{1,\ldots,m\}})(i)$ proving that

$$(\sigma \circ \rho)^{|\{1,\ldots,\}} = \sigma^{|\{1,\ldots,m\}} \circ \rho^{|\{1,\ldots,m\}}$$

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

$$\begin{aligned}(\sigma_1 \circ \cdots \circ \sigma_n)^{|\{1,\ldots,m\}} &\underset{[\text{theorem: } 11.190]}{=} (\sigma_1)^{|\{1,\ldots,m\}} \\ &\underset{[\text{theorem: } 11.190]}{=} ((\sigma_1)^{|\{1,\ldots,m\}}) \circ \cdots \circ (\sigma_k)^{|\{1,\ldots,m\}}\end{aligned}$$

   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

$$\begin{aligned}(\sigma_1 \circ \cdots \circ \sigma_{n+1})^{|\{1,\ldots,m\}} &\underset{[\text{theorem: } 11.190]}{=} \\ ((\sigma_1 \circ \cdots \circ \sigma_n) \circ \sigma_{n+1})^{|\{1,\ldots,m\}} &\underset{(3)}{=} \\ (\sigma_1 \circ \cdots \circ \sigma_n)^{|\{1,\ldots,m\}} \circ (\sigma_{n+1})^{|\{1,\ldots,m\}} &\underset{n \in S}{=} \\ ((\sigma_1)^{|\{1,\ldots,m\}} \circ \cdots \circ (\sigma_n)^{|\{1,\ldots,m\}}) \circ (\sigma_{n+1})^{|\{1,\ldots,m\}} &\underset{[\text{theorem: } 11.190]}{=} \\ ((\sigma_1)^{|\{1,\ldots,m\}} \circ \cdots \circ (\sigma_{n+1})^{|\{1,\ldots,m\}}) &\end{aligned}$$

   proving that $n+1 \in S$ □

Corollary 11.194. *Let $n \in \mathbb{N}$ and $\sigma \in P_{n+1}$ such that $\sigma(n+1) = n+1$ [so that by [theorem: 11.26] $\sigma_{|\{1,\ldots,n\}} \in P_n$] 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 $(\sigma_{|\{1,\ldots,n\}})^{|\{1,\ldots,n+1\}}(k) = \sigma_{|\{1,\ldots,n\}} = \sigma(n)$*

$k = n+1$. *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 11.195. *Let* $n \in \mathbb{N} \setminus \{1\}$ *and* $\sigma \in P_n$ *then there exists a family* $\left\{\left(i_k \underset{n}{\leftrightarrow} j_k\right)\right\}_{k \in \{1, \ldots, l\}} \subseteq P_n$ *with* $\forall k \in \{1, \ldots, l\}$ $i_k \neq j_k$ *such that*

$$\sigma = \left(i_1 \underset{n}{\leftrightarrow} j_1\right) \circ \cdots \circ \left(i_l \underset{n}{\leftrightarrow} j_l\right)$$

Proof. We use mathematical induction to prove this, so let

$S = \left\{n \in \{2, \ldots, \infty\} | \text{If } \sigma \in P_n \text{ then } \sigma = \left(i_1 \underset{n}{\leftrightarrow} j_1\right) \circ \cdots \circ \left(i_l \underset{n}{\leftrightarrow} j_l\right) \text{ where } \left\{\left(i_k \underset{n}{\leftrightarrow} j_k\right)\right\}_{k \in \{1, \ldots, l\}} \subseteq P_n \text{ satisfies } i_k \neq j_k \forall k \in \{1, \ldots, l\}\right\}$

then we have:

$\boldsymbol{2 \in S}$**.** The for $\sigma \in P_2$ we have for $\sigma(1)$ either:

$\boldsymbol{\sigma(1) = 1}$**.** Then we must have as σ is injective that $\sigma(1) = \sigma(2)$ hence $\sigma = \mathrm{Id}_{\{1, \ldots, 2\}}$. Define $\{(i_k \leftrightarrow j_k)\}_{k \in \{1, \ldots, 2\}}$ by $\left(i_1 \underset{2}{\leftrightarrow} j_1\right) = \left(i_2 \underset{2}{\leftrightarrow} j_2\right) = \left(1 \underset{2}{\leftrightarrow} 2\right)$, a family of **strict** transpositions. Then

$$\left(i_1 \underset{2}{\leftrightarrow} j_1\right) \circ \cdots \circ \left(i_2 \underset{2}{\leftrightarrow} j_2\right) = \left(1 \underset{2}{\leftrightarrow} 2\right) \circ \left(1 \underset{2}{\leftrightarrow} 2\right) \underset{[\text{theorem: } 11.192]}{=} \mathrm{Id}_{\{1, \ldots, 2\}} = \sigma$$

proving that in this cases $2 \in S$.

$\boldsymbol{\sigma(1) = 2}$**.** Then we must have as σ is injective that $\sigma(2) = \sigma(1)$ 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: 11.26] 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

$$\begin{aligned}
\sigma & \underset{[\text{theorem: } 11.194]}{=} (\sigma_{|\{1, \ldots, n\}})^{|\{1, \ldots, n+1\}} \\
& = \left(\left(i_1 \underset{n}{\leftrightarrow} j_1\right) \circ \cdots \circ \left(i_l \underset{n}{\leftrightarrow} j_l\right)\right)^{|\{1, \ldots, n+1\}} \\
& \underset{[\text{theorem: } 11.193]}{=} \left(\left(i_1 \underset{n}{\leftrightarrow} j_1\right)^{|\{1, \ldots, n+1\}} \circ \cdots \circ \left(i_l \underset{n}{\leftrightarrow} j_l\right)^{|\{1, \ldots, n+1\}}\right) \\
& \underset{[\text{theorem: } 11.193]}{=} \left(i_1 \underset{n+1}{\leftrightarrow} j_1\right) \circ \cdots \circ \left(i_l \underset{n+1}{\leftrightarrow} j_l\right)
\end{aligned}$$

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

$$\rho = \sigma \circ \left(i \underset{n+1}{\leftrightarrow} n+1\right) \text{ a strict transposition as } i \neq n+1 \tag{11.85}$$

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: 11.26] 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

$$\begin{array}{rcl}
\rho & \underset{[\text{theorem: } 11.194]}{=} & (\rho_{|\{1,\ldots,n\}})^{|\{1,\ldots,n+1\}} \\
 & = & \left((r_1 \underset{n}{\leftrightarrow} s_1)\circ\cdots\circ(r_l \underset{n}{\leftrightarrow} s_l)\right)^{|\{1,\ldots,n+1\}} \\
 & \underset{[\text{theorem: } 11.193]}{=} & \left((r_1 \underset{n}{\leftrightarrow} s_1)^{|\{1,\ldots,n+1\}}\circ\cdots\circ(r_l \underset{n}{\leftrightarrow} s_l)^{|\{1,\ldots,n+1\}}\right) \\
 & \underset{[\text{theorem: } 11.193]}{=} & (r_1 \underset{n+1}{\leftrightarrow} s_1)\circ\cdots\circ(r_l \underset{n+1}{\leftrightarrow} s_l)
\end{array} \tag{11.86}$$

Define now

$$\left\{\left(i_k \underset{n+1}{\leftrightarrow} j_k\right)\right\}_{k\in\{1,\ldots,l+1\}} \text{ by } (i_k \underset{n}{\leftrightarrow} j_k) = \begin{cases} \left(i \underset{n+1}{\leftrightarrow} n+1\right) \text{ if } k=l+1 \\ (r_k \underset{n+1}{\leftrightarrow} s_k) \text{ if } k\in\{1,\ldots,l\} \end{cases}$$

then we have

$$\begin{array}{rcl}
\sigma & \underset{[\text{theorem: } 11.192]}{=} & \sigma\circ\left(\left(i \underset{n+1}{\leftrightarrow} n+1\right)\circ\left(i \underset{n+1}{\leftrightarrow} n+1\right)\right) \\
 & \underset{\text{associativity}}{=} & \left(\sigma\circ\left(i \underset{n+1}{\leftrightarrow} n+1\right)\right)\circ\left(i \underset{n+1}{\leftrightarrow} n+1\right) \\
 & = & \rho\circ\left(i \underset{n+1}{\leftrightarrow} n+1\right) \\
 & \underset{[\text{theorem: } 11.86]}{=} & \left((r_1 \underset{n+1}{\leftrightarrow} s_1)\circ\cdots\circ(r_l \underset{n+1}{\leftrightarrow} s_l)\right)\circ\left(i \underset{n+1}{\leftrightarrow} n+1\right) \\
 & = & \left(\left(i_1 \underset{n+1}{\leftrightarrow} j_1\right)\circ\cdots\circ\left(i_l \underset{n+1}{\leftrightarrow} j_l\right)\right)\circ\left(i_{n+1} \underset{n+1}{\leftrightarrow} j_{n+1}\right) \\
 & \underset{[\text{theorem: } 11.190]}{=} & \left(i_1 \underset{n+1}{\leftrightarrow} j_1\right)\circ\cdots\circ\left(i_{l+1} \underset{n+1}{\leftrightarrow} j_{l+1}\right)
\end{array}$$

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. □

Permutations are typically used to per mutate the arguments of functions with several arguments. Remember that X^n is defined as $X^n \underset{[\text{definition: } 6.77]}{=} A^{\{1,\ldots,n\}}$

Definition 11.196. *Let X,Y be sets, $n\in\mathbb{N}$ and $f\colon X^n\rightarrow Y$ a function and $\sigma\in P_n$ then*

$$\sigma f\colon X^n\rightarrow Y$$

is defined by

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

or using the notation [notation: 6.74]

$$\sigma f(x)=f(x_{\sigma(1)},\ldots,x_{\sigma(n)})$$

Theorem 11.197. *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: 11.59] 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

$$\sigma(\alpha \cdot f) = \alpha \cdot \sigma f \qquad \square$$

Theorem 11.198. *Let X, Y be sets, $n \in \mathbb{N}$ and $f: X^n \rightarrow Y$ 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 $\mathrm{Id}_{\{1, \ldots, n\}} f = f$. □

Theorem 11.199. *Let X, Y be sets, $n \in \mathbb{N}$ and $f: X^n \rightarrow Y$ a function and $\sigma, \rho \in P_n$ then*

$$\sigma(\rho f) = (\rho \circ \sigma) f$$

Proof. $\forall x \in X^n$ we have

$$\sigma(\rho f)(x) = (\sigma f)(x \circ \rho) = f((x \circ \rho) \circ \sigma) = f(x \circ (\rho \circ \sigma)) = (\rho \circ \sigma) f(x)$$

so that $\sigma(\rho f) = (\rho \circ \sigma) f$. □

Theorem 11.200. *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 11.201. *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:

$S = \left\{l \in \mathbb{N} | \text{If } f: X^n \rightarrow Y \text{ satisfies } \forall i, j \in \{1, \ldots, n\} \left(i \underset{n}{\leftrightarrow} j\right) f = f \text{ and } \left\{\left(i_k \underset{n}{\leftrightarrow} j_k\right)\right\}_{k \in \{1, \ldots, l\}} \subseteq P_n \text{ then } \left(\left(i_1 \underset{n}{\leftrightarrow} j_1\right) \circ \cdots \circ \left(i_l \underset{n}{\leftrightarrow} j_l\right)\right) f = f\right\}$

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

$$\left(\left(i_1 \underset{n}{\leftrightarrow} j_1\right) \circ \cdots \circ \left(i_l \underset{n}{\leftrightarrow} j_l\right)\right) f \underset{\text{[theorem: 11.190]}}{=} \left(i_1 \underset{n}{\leftrightarrow} j_1\right) f \tag{11.87}$$

For i_1, j_1 we have either:

$i_1 = j_1$. Then

$$\left(i_1 \underset{n}{\leftrightarrow} j_1\right) f \underset{\text{[theorem: 11.190]}}{=} \mathrm{Id}_{\{1,, \ldots, n\}} f \underset{\text{[theorem: 11.198]}}{=} f$$

$i_1 \neq j_1$. Then by the hypothesis we have $\left(i_1 \underset{n}{\leftrightarrow} j_1\right) f = f$

combining this with [eq: 11.87] 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 $1\in S$.

$\boldsymbol{l\in S\Rightarrow l+1\in S}$**.** Let $f\colon X^n\to Y$ satisfies $\forall i,j\in\{1,\ldots,n\}$ $\left(i \underset{n}{\leftrightarrow} j\right)f=f$ and $\left\{\left(i_k \underset{n}{\leftrightarrow} j_k\right)\right\}_{k\in\{1,\ldots,l+1\}}\subseteq P_n$ then we have

$$\begin{aligned}
\left(\left(i_1 \underset{n}{\leftrightarrow} j_1\right)\circ\cdots\circ\left(i_{l+1} \underset{n}{\leftrightarrow} j_{l+1}\right)\right)f &\underset{[\text{eq: } 11.190]}{=} \\
\left(\left(\left(i_1 \underset{n}{\leftrightarrow} j_1\right)\circ\cdots\circ\left(i_l \underset{n}{\leftrightarrow} j_l\right)\right)\circ\left(i_{l+1} \underset{n}{\leftrightarrow} j_{l+1}\right)\right)f &\underset{[\text{theorem: } 11.199]}{=} \\
\left(i_{l+1} \underset{n}{\leftrightarrow} j_{l+1}\right)\left(\left(\left(i_1 \underset{n}{\leftrightarrow} j_1\right)\circ\cdots\circ\left(i_l \underset{n}{\leftrightarrow} j_l\right)\right)f\right) &\underset{n\in S}{=} \left(i_{l+1} \underset{n}{\leftrightarrow} j_{l+1}\right)f
\end{aligned} \tag{11.88}$$

For i_{l+1}, j_{l+1} we have either:

$\boldsymbol{i_{l+1}=j_{l+1}}$**.** Then

$$\left(i_{l+1} \underset{n}{\leftrightarrow} j_{l+1}\right)f \underset{[\text{theorem: } 11.190]}{=} \mathrm{Id}_{\{1,,\ldots,n\}}f \underset{[\text{theorem: } 11.198]}{=} f$$

$\boldsymbol{i_{l+1}\neq j_{l+1}}$**.** Then by the hypothesis we have $\left(i_{l+1} \underset{n}{\leftrightarrow} j_{l+1}\right)f=f$

combining this with [eq: 11.88] proves that $\left(\left(i_1 \underset{n}{\leftrightarrow} j_1\right)\circ\cdots\circ\left(i_{l+1} \underset{n}{\leftrightarrow} j_{l+1}\right)\right)f=f$. So $n+1\in S$ concluding the proof by induction. □

We can extend the above theorem to general permutation's.

Corollary 11.202. *Let X, Y be sets, $n\in\mathbb{N}$ and $f\colon X^n\to Y$ 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 $\sigma f=\mathrm{Id}_{\{1\}}f \underset{[\text{theorem: } 11.198]}{=} f$. To complete the proof we have to prove the case where $n\in\mathbb{N}\backslash\{1\}$. Using [theorem: 11.195] 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

$$\sigma f=\left(\left(i_1 \underset{n}{\leftrightarrow} j_1\right)\circ\cdots\circ\left(i_l \underset{n}{\leftrightarrow} j_l\right)\right)f \underset{[\text{theorem: } 11.201]}{=} f$$

proving the corollary. □

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 $\sum$. Further let $n\in\mathbb{N}_0$ then $\{(i,j)\in\{1,\ldots,n\}\times\{1,\ldots,n\}|i<j\}\subseteq\{1,\ldots,n\}\times\{1,\ldots,n\}$ a finite set [see [theorems: 10.75,10.80], 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 11.203. *Let $n\in\mathbb{N}$ then we define*

$$\Phi_n\colon\mathbb{Z}^n\to\mathbb{Z} \textit{ by } \Phi_n(x)=\prod_{(i,j)\in\{(i,j\in\{1,\ldots,n\}\cdot\{1,\ldots,n\})|i<j\}}(x_i-x_j)$$

Example 11.204.

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. $\Phi_2(x)=x_2-x_1$, as $\{(i,j)\in\{1,2\}\times\{1,2\}|i<j\}=\{(1,2)\}$ we have
$$\Phi_2(x)=\prod_{(i,j)\in\{(1,2)\}}(x_i-x_j)=x_1-x_2$$
3. $\Phi_3(x)=(x_1-x_2)\cdot(x_1-x_3)\cdot(x_2-x_3)$, as
$$\{(i,j)\in\{1,2,3\}\times\{1,2,3\}|i<j\}=\{(1,2),(1,3),(2,3)\}$$
we have
$$\Phi_3(x)=\prod_{(i,j)\in\{(1,2),(1,3),(2,3)\}}(x_i-x_j)=(x_1-x_2)\cdot(x_1-x_3)\cdot(x_2-x_3)$$
4. $\ldots$

Lemma 11.205. *Let $n \in \mathbb{N}$ then for $(x_1, \ldots, x_n) \in \mathbb{Z}^n$ defined by $x_i = i$ then $\Phi_n(x_1, \ldots, x_n) \neq 0$, 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: 11.47] it follows then that $\Phi_n(x_1, \ldots, x_n) = \prod_{(i,j) \in \{(i,j) \in \{1,\ldots,n\} \times \{1,\ldots,m\} | i < j\}} (x_i - x_j) \neq 0$. □

Proof. We prove this by induction so let

$$S = \{n \in \mathbb{N} | \Phi_n(1, \ldots, n) \neq 0\}$$

then we have:

$1 \in S$. Then $\Phi_1(1) = \prod_{(i,j) \in \{(i,j) \in \{1\} \times \{j\} | i < j\}} (i - j) = \prod_{(i,j) \in \varnothing} (i - j) = 1$

$n \in S \Rightarrow n + 1 \in S$. Then we have

□

Lemma 11.206. *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:

$i = k$. Then as $k < l$ and $i < j$ we have for j either:

$j < l$. Then $(i,j) \in J_0$ where $J_0 = \{(k,j) \in I | k < j < l\} \subseteq I$

$j = l$. Then $(i,j) \in J_1$ where $J_1 = \{(k,l)\} \subseteq I$

$l < j$. Then $(i,j) \in J_2$ where $J_2 = \{(k,j) \in I | l < j\} \subseteq I$

$i \neq k$. Then we have for j either:

$i = l$. Then as $i < j$ we have $(i,j) \in J_3 = \{(l,j) \in I | l < j\} \subseteq I$

$i \neq l$. Then we have for j either:

$j = k$. Then as $i < j \Rightarrow i < k$ we have $(i,j) \in J_4$ where $J_4 = \{(i,k) \in I | i < k\} \subseteq I$

$j = l$. Then as $i = k$ we have either:

$i < k$. Then $(i,j) \in J_5$ where $J_5 = \{(i,l) \in I | i < k\} \subseteq I$

$k < i$. Then $(i,j) \in J_6$ where $J_6 = \{(i,l) \in I | k < i < l\} \subseteq I$

$j \neq l, k$. 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

$(i,j) \in J_0 \bigcap J_1$. Then $i = k < j < l \wedge j = l$ a contradiction.

$(i,j) \in J_0 \bigcap J_2$. Then $i = k < j < l \wedge l < j$ a contradiction.

$(i,j) \in J_0 \bigcap J_3$. Then $i = k < j < l \wedge l < j$ a contradiction.

$(i,j) \in J_0 \bigcap J_4$. Then $i = k < j < l \wedge j = k$ a contradiction.

$(i,j) \in J_0 \bigcap J_5$. Then $i = k < j < l \wedge j = l$ a contradiction.

$(i,j) \in J_0 \bigcap J_6$. 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 $i=l\wedge l<j\wedge i\neq k$ 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 $i<k=j\wedge j=l$ contradicting $k<l$.

$(i,j)\in J_5\bigcap J_6$. 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

$$I=\bigcup_{i\in\{,\ldots,7\}} J_i \text{ and } \forall i,j\in I \text{ with } i\neq j \text{ we have } J_i\bigcap J_j=\varnothing$$

So that

$$\begin{aligned}\Phi_n(x_1,\ldots,x_n) &= \prod_{(i,j)\in I}(x_i-x_j)\\ &\underset{[\text{theorem: } 11.41]}{=} \prod_{m=0}^{7}\left(\prod_{(i,j)\in J_m}(x_i-x_j)\right)\\ &= \prod_{m=0}^{7} Q_m \end{aligned}\tag{11.89}$$

where

$$Q_m=\prod_{(i,j)\in J_m}(x_i-x_j)\tag{11.90}$$

and

$$\begin{aligned}\left(\left(k\underset{n}{\leftrightarrow}l\right)\Phi_n\right)(x_1,\ldots,x_n) &= \prod_{(i,j)\in I}\left(x_{(k\underset{n}{\leftrightarrow}l)(i)}-x_{(k\underset{n}{\leftrightarrow}l)(j)}\right)\\ &\underset{[\text{theorem: } 11.41]}{=} \prod_{m=0}^{7}\left(\prod_{(i,j)\in J_m}\left(x_{(k\underset{n}{\leftrightarrow}l)(i)}-x_{(k\underset{n}{\leftrightarrow}l)(j)}\right)\right)\\ &= \prod_{m=0}^{7} R_m\end{aligned}\tag{11.91}$$

where

$$R_m = \prod_{(i,j) \in J_m} \left(x_{(k \underset{n}{\leftrightarrow} l)(i)} - x_{(k \underset{n}{\leftrightarrow} l)(j)}\right) \tag{11.92}$$

Now we have:

***R*₀.**

$$\begin{aligned} R_0 &= \prod_{(i,j) \in J_0} \left(x_{(k \underset{n}{\leftrightarrow} l)(i)} - x_{(k \underset{n}{\leftrightarrow} l)(j)}\right) \\ &= \prod_{(i,j) \in \{(k,j) \in I | k < j < l\}} \left(x_{(k \underset{n}{\leftrightarrow} l)(i)} - x_{(k \underset{n}{\leftrightarrow} l)(j)}\right) \\ &= \prod_{(i,j) \in \{(k,j) \in I | k < j < l\}} (x_l - x_j) \\ &= \prod_{(i,j) \in J_0} (x_l - x_j) \end{aligned} \tag{11.93}$$

Define now

$$\beta: J_6 = \{(i,l) \in I | k < i < l\} \rightarrow J_0 = \{(k,j) \in I | k < j < l\} \text{ by } \beta(i,j) = (k,i)$$

then we have:

injectivity. Let $(i,j), (i',j') \in J_7$ 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 $(s,l) \in J_7$ which, as $\beta(s,l) = (k,s) = (r,s)$, proves surjectivity.

proving that

$$\beta: J_6 \rightarrow J_0 \text{ is a bijection}$$

Next

$$\begin{aligned} \prod_{(i,j) \in J_0} (x_l - x_j) &= \prod_{(i,j) \in J_0} (x_l - x_{(i,j)_2}) \\ &\underset{[\text{theorem: } 11.34]}{=} \prod_{(i,j) \in J_6} (x_l - x_{\beta(i,j)_2}) \\ &= \prod_{(i,j) \in J_6} (x_l - x_i) \\ &= \prod_{(i,j) \in J_6} ((-1) \cdot (x_i - x_l)) \\ &\underset{[\text{theorem: } 11.36]}{=} \left(\prod_{(i,j) \in J_6} (-1)\right) \cdot \prod_{(i,j) \in J_6} (x_i - x_l) \\ &\underset{[\text{theorem: } 11.45]}{=} (-1)^{\text{card}(J_6)} \cdot \prod_{(i,j) \in \{(i,l) \in I | k < i < l\}} (x_i - x_l) \\ &= (-1)^{\text{card}(J_6)} \cdot \prod_{(i,j) \in \{(i,l) \in I | k < i < l\}} (x_i - x_j) \\ &= (-1)^{\text{card}(j_6)} \cdot \prod_{(i,j) \in J_6} (x_i - x_l) \\ &= (-1)^{\text{card}(j_6)} \cdot Q_6 \end{aligned}$$

combining the above with [eq: 11.93] gives

$$R_0 = (-1)^{\text{card}(j_6)} \cdot Q_6 \tag{11.94}$$

***R*₁.** Then

$$\begin{aligned} R_1 &= \prod_{(i,j)\in J_1} \left(x_{(k\underset{n}{\leftrightarrow}l)(i)} - x_{(k\underset{n}{\leftrightarrow}l)(j)}\right) \\ &= \prod_{(i,j)\in\{(k,l)\}} \left(x_{(k\underset{n}{\leftrightarrow}l)(i)} - x_{(k\underset{n}{\leftrightarrow}l)(j)}\right) \\ &\underset{\text{[theorem: 11.32}}{=} x_{(k\underset{n}{\leftrightarrow}l)(k)} - x_{(k\underset{n}{\leftrightarrow}l)(l)} \\ &= x_l - x_k \\ &= (-1)\cdot(x_k - x_l) \\ &\underset{\text{[theorem: 11.32}}{=} (-1)\cdot \prod_{(i,j)\in\{(k,l)\}} (x_i - x_j) \\ &= (-1)\cdot \prod_{(i,j)\in J_1} (x_i - x_j) \\ &= (-1)\cdot Q_1 \end{aligned}$$

proving that

$$R_1 = (-1)\cdot Q_1 \tag{11.95}$$

***R*₂.** Then

$$\begin{aligned} R_3 &= \prod_{(i,j)\in J_2} \left(x_{(k\underset{n}{\leftrightarrow}l)(i)} - x_{(k\underset{n}{\leftrightarrow}l)(j)}\right) \\ &= \prod_{(i,j)\in\{(k,j)\in I|l<j\}} \left(x_{(k\underset{n}{\leftrightarrow}l)(i)} - x_{(k\underset{n}{\leftrightarrow}l)(j)}\right) \\ &= \prod_{(i,j)\in\{(k,j)\in I|l<j\}} (x_l - x_j) \\ &= \prod_{(i,j)\in J_2} (x_l - x_j) \end{aligned} \tag{11.96}$$

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

$$\begin{aligned} \prod_{(i,j)\in J_2} (x_l - x_j) &= \prod_{(i,j)\in J_2} (x_l - x_{(i,j)_2}) \\ &\underset{\text{[theorem: 11.34]}}{=} \prod_{(i,j)\in J_3} (x_l - x_{\beta(i,j)_2}) \\ &= \prod_{(i,j)\in J_3} (x_l - x_j) \\ &= \prod_{(i,j)\in\{(l,j)\in I|l<j\}} (x_l - x_j) \\ &= \prod_{(i,j)\in\{(l,j)\in I|l<j\}} (x_i - x_j) \\ &= \prod_{(i,j)\in J_3} (x_i - x_j) \\ &= Q_3 \end{aligned}$$

which combined with [eq: 11.96] proves that

$$R_2 = Q_3 \tag{11.97}$$

***R*₃.** We have

$$\begin{aligned}
R_3 &= \prod_{(i,j)\in J_3} \left(x_{(k \underset{n}{\leftrightarrow} l)(i)} - x_{(k \underset{n}{\leftrightarrow} l)(j)}\right)\\
&= \prod_{(i,j)\in\{(l,j)\in I|l<j\}} \left(x_{(k \underset{n}{\leftrightarrow} l)(i)} - x_{(k \underset{n}{\leftrightarrow} l)(j)}\right)\\
&= \prod_{(i,j)\in\{(k,j)\in I|l<j\}} (x_k - x_j)\\
&= \prod_{(i,j)\in J_3} (x_k - x_j) \qquad (11.98)
\end{aligned}$$

Define

$$\beta\colon J_2 = \{(k,j)\in I|l<j\} \rightarrow J_3 = \{(l,j)\in I|l<j\} \text{ by } \beta(i,j) = (l,j)$$

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 $l<s$, take (k,s) then $(k,s)\in J_3$ and $\beta(k,s)=(l,s)=(r,s)$ proving surjectivity.

hence we have

$$\beta\colon J_2 \rightarrow J_3 \text{ is a bijection}$$

Next

$$\begin{aligned}
\prod_{(i,j)\in J_3} (x_k - x_j) &= \prod_{(i,j)\in J_3} (x_k - x_{(i,j)_2})\\
&\underset{[\text{theorem: } 11.34]}{=} \prod_{(i,j)\in J_2} (x_k - x_{\beta(i,j)_2})\\
&= \prod_{(i,j)\in J_2} (x_k - x_{(l,j)_2})\\
&= \prod_{(i,j)\in J_2} (x_k - x_j)\\
&= \prod_{(i,j)\in\{(k,j)\in I|l<j\}} (x_k - x_j)\\
&= \prod_{(i,j)\in\{(k,j)\in I|l<j\}} (x_i - x_j)\\
&= \prod_{(i,j)\in J_2} (x_i - x_j)\\
&= Q_2
\end{aligned}$$

proving together with [eq: 11.98] that

$$R_3 = Q_2 \tag{11.99}$$

***R*₄.** We have

$$\begin{aligned}
R_4 &= \prod_{(i,j)\in J_4} \left(x_{(k \underset{n}{\leftrightarrow} l)(i)} - x_{(k \underset{n}{\leftrightarrow} l)(j)}\right)\\
&= \prod_{(i,j)\in\{(i,k)\in I|i<k\}} \left(x_{(k \underset{n}{\leftrightarrow} l)(i)} - x_{(k \underset{n}{\leftrightarrow} l)(j)}\right)\\
&= \prod_{(i,j)\in\{(i,k)\in I|i<k\}} (x_i - x_l)\\
&= \prod_{(i,j)\in J_4} (x_i - x_l) \qquad (11.100)
\end{aligned}$$

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 $(r,l) \in J_4$ and $\beta(r,l) = (r,k) = (r,s)$ proving surjectivity.

So

$$\begin{aligned}
\prod_{(i,j) \in J_4} (x_i - x_l) &= \prod_{(i,j) \in J_4} (x_{(i,j)_1} - x_l) \\
&\underset{\text{[theorem: 11.34]}}{=} \prod_{(i,j) \in J_5} (x_{\beta(i,j)_1} - x_l) \\
&= \prod_{(i,j) \in J_5} (x_{(i,k)_1} - x_l) \\
&= \prod_{(i,j) \in J_5} (x_i - x_l) \\
&= \prod_{(i,j) \in \{(i,l) \in I | i < k\}} (x_i - x_l) \\
&= \prod_{(i,j) \in \{(i,l) \in I | i < k\}} (x_i - x_j) \\
&= \prod_{(i,j) \in J_5} (x_i - x_j) \\
&= Q_5
\end{aligned}$$

combining this with [eq: 11.100] gives

$$R_4 = Q_5 \tag{11.101}$$

***R*₅.** We have

$$\begin{aligned}
R_5 &= \prod_{(i,j) \in J_5} \left(x_{(k \underset{n}{\leftrightarrow} l)(i)} - x_{(k \underset{n}{\leftrightarrow} l)(j)}\right) \\
&= \prod_{(i,j) \in \{(i,l) \in I | i < k\}} \left(x_{(k \underset{n}{\leftrightarrow} l)(i)} - x_{(k \underset{n}{\leftrightarrow} l)(j)}\right) \\
&= \prod_{(i,j) \in \{(i,l) \in I | i < k\}} (x_i - x_k) \\
&= \prod_{(i,j) \in J_5} (x_i - x_k) \qquad (11.102)
\end{aligned}$$

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

$$\begin{array}{lll}
\prod_{(i,j)\in J_5}(x_i-x_k) & = & \prod_{(i,j)\in J_5}(x_{(i,j)_1}-x_k)\\
 & \underset{[\text{theorem: } 11.34]}{=} & \prod_{(i,j)\in J_4}(x_{\beta(i,j)_1}-x_k)\\
 & = & \prod_{(i,j)\in J_4}(x_{(i,l)_1}-x_k)\\
 & = & \prod_{(i,j)\in J_4}(x_i-x_k)\\
 & = & \prod_{(i,j)\in\{(i,k)\in I|i<k\}}(x_i-x_k)\\
 & = & \prod_{(i,j)\in\{(i,k)\in I|i<k\}}(x_i-x_j)\\
 & = & \prod_{(i,j)\in J_4}(x_i-x_j)\\
 & = & Q_4
\end{array}$$

combining the above with [eq: 11.102] proves that

$$R_5=Q_4 \tag{11.103}$$

***R*₆.** We have

$$\begin{array}{lll}
R_6 & = & \prod_{(i,j)\in J_6}\left(x_{(k\underset{n}{\leftrightarrow}l)(i)}-x_{(k\underset{n}{\leftrightarrow}l)(j)}\right)\\
 & = & \prod_{(i,j)\in\{(i,l)\in I|k<i<l\}}\left(x_{(k\underset{n}{\leftrightarrow}l)(i)}-x_{(k\underset{n}{\leftrightarrow}l)(j)}\right)\\
 & = & \prod_{(i,j)\in\{(i,l)\in I|k<i<l\}}(x_i-x_k)\\
 & = & \prod_{(i,j)\in J_6}(x_i-x_k)
\end{array} \tag{11.104}$$

Define

$$\beta: J_0=\{(k,j)\in I|k<j<l\}\rightarrow J_6=\{(i,l)\in I|k<i<l\} \text{ by } \beta(i,j)=(j,l)$$

then we have:

injectivity. Let $(i,j),(i',j')\in J_0$ with $\beta(i,j)=\beta(i',j')$ then $i=k=i'$ and $(j,l)=(j',l)\Rightarrow j=j'$ then $(i,j)=(i',j')$.

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

$$\beta: J_0\rightarrow J_6 \text{ is a bijection so that } \operatorname{card}(J_0)=\operatorname{card}(J_6) \tag{11.105}$$

Further

$$\begin{aligned}
\prod_{(i,j)\in J_6}(x_i-x_k) &= \prod_{(i,j)\in J_6}(x_{(i,j)_1}-x_k)\\
&\underset{[\text{theorem: } 11.34]}{=} \prod_{(i,j)\in J_0}(x_{\beta(i,j)_1}-x_k)\\
&= \prod_{(i,j)\in J_0}(x_{(j,l)_1}-x_k)\\
&= \prod_{(i,j)\in J_0}(x_j-x_k)\\
&= \prod_{(i,j)\in J_0}((-1)\cdot(x_k-x_j))\\
&\underset{[\text{theorem: } 11.36]}{=} \left(\prod_{(i,j)\in J_0}(-1)\right)\cdot \prod_{(i,j)\in J_0}(x_k-x_j)\\
&\underset{[\text{theorem: } 11.45]}{=} (-1)^{\mathrm{card}(J_0)}\cdot \prod_{(i,j)\in J_0}(x_k-x_j)\\
&\underset{[\text{eq: } 11.105]}{=} (-1)^{\mathrm{card}(J_6)}\cdot \prod_{(i,j)\in J_0}(x_k-x_j)\\
&= (-1)^{\mathrm{card}(J_6)}\cdot \prod_{(i,j)\in\{(k,j)\in I|k<j<l\}}(x_k-x_j)\\
&= (-1)^{\mathrm{card}(J_6)}\cdot \prod_{(i,j)\in\{(k,j)\in I|k<j<l\}}(x_i-x_j)\\
&= (-1)^{\mathrm{card}(J_6)}\cdot \prod_{(i,j)\in J_0}(x_i-x_j)\\
&= (-1)^{\mathrm{card}(J_6)}\cdot Q_0
\end{aligned}$$

which together with [eq: 11.104] proves that

$$J_0=(-1)^{\mathrm{card}(J_6)}\cdot Q_0 \tag{11.106}$$

***R*₇.** We have

$$\begin{aligned}
J_7 &= \prod_{(i,j)\in J_7}\left(x_{(k\underset{n}{\leftrightarrow}l)(i)}-x_{(k\underset{n}{\leftrightarrow}l)(j)}\right)\\
&= \prod_{(i,j)\in\{(i,j)\in I|i\neq k,l\wedge j\neq k,l\}}\left(x_{(k\underset{n}{\leftrightarrow}l)(i)}-x_{(k\underset{n}{\leftrightarrow}l)(j)}\right)\\
&= \prod_{(i,j)\in\{(i,j)\in I|i\neq k,l\wedge j\neq k,l\}}(x_i-x_j)\\
&= \prod_{(i,j)\in J_7}(x_i-x_j)\\
&= Q_7
\end{aligned}$$

then

$$J_7=Q_7 \tag{11.107}$$

Finally we have

$$\begin{aligned}
\left(\left(k\underset{n}{\leftrightarrow}l\right)\Phi_n\right)(x_1,\ldots,x_n) &\underset{[\text{eq: } 11.91]}{=}\\
R_1\cdot R_2\cdot R_3\cdot R_4\cdot R_5\cdot R_6\cdot R_7 &\underset{[\text{eqs: } 11.94,11.95,11.97,11.99,11.101,11.103,11.106,11.107]}{=}\\
(-1)^{\mathrm{card}(J_6)}\cdot Q_6\cdot(-1)\cdot Q_1\cdot Q_3\cdot Q_2\cdot Q_5\cdot Q_4\cdot &\\
(-1)^{\mathrm{card}(J_6)}\cdot Q_0\cdot Q_7 &=\\
-(Q_1\cdot Q_2\cdot Q_3\cdot Q_4\cdot Q_5\cdot Q_6\cdot Q_7) &=\\
-\Phi_n(x_1,\ldots,x_n) &
\end{aligned}$$

proving that

$$\left(k \underset{n}{\leftrightarrow} l\right)\Phi_n = -\Phi_n$$ □

Corollary 11.207. *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

$$S = \left\{m \in \mathbb{N} | \left(\left(i_1 \underset{n}{\leftrightarrow} j_1\right) \circ \cdots \circ \left(i_m \underset{n}{\leftrightarrow} j_m\right)\right)\Phi_n = (-1)^m \cdot \Phi_n\right\}$$

then we have:

$1 \in S$. $\left(\left(i_1 \underset{n}{\leftrightarrow} j_1\right) \circ \cdots \circ \left(i_1 \underset{n}{\leftrightarrow} j_1\right)\right)\Phi_n = \left(i_1 \underset{n}{\leftrightarrow} j_1\right)\Phi_n \underset{[\text{theorem: } 11.206]}{=} (-1) \cdot \Phi_n = (-1)^1 \cdot \Phi_n$ proving that $1 \in S$.

$m \in S \Rightarrow m+1 \in S$. We have

$$\begin{aligned}
\left(\left(i_1 \underset{n}{\leftrightarrow} j_1\right) \circ \cdots \circ \left(i_{m+1} \underset{n}{\leftrightarrow} j_{m+1}\right)\right)\Phi_n & \underset{[\text{theorem: } 11.190]}{=} \\
\left(\left(\left(i_1 \underset{n}{\leftrightarrow} j_1\right) \circ \cdots \circ \left(i_m \underset{n}{\leftrightarrow} j_m\right)\right) \circ \left(i_{m+1} \underset{n}{\leftrightarrow} j_{m+1}\right)\right)\Phi_n & \underset{[\text{theorem: } 11.199]}{=} \\
\left(i_{m+1} \underset{n}{\leftrightarrow} j_{m+1}\right)\left(\left(\left(i_1 \underset{n}{\leftrightarrow} j_1\right) \circ \cdots \circ \left(i_m \underset{n}{\leftrightarrow} j_m\right)\right)\Phi_n\right) & \underset{n \in S}{=} \\
\left(i_{m+1} \underset{n}{\leftrightarrow} j_{m+1}\right)\left((-1)^m \cdot \Phi_n\right) & \underset{[\text{theorem: } 11.200]}{=} \\
(-1)^m \cdot \left(\left(i_{m+1} \underset{n}{\leftrightarrow} j_{m+1}\right)\Phi_n\right) & \underset{[\text{theorem: } 11.206]}{=} \\
(-1)^m \cdot (-1) \cdot \Phi_n & = \\
(-1)^{m+1} \cdot \Phi_n &
\end{aligned}$$

proving that $m+1 \in S$ □

Theorem 11.208. *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 $\sigma\Phi_1 = \sigma\mathrm{Id}_{\{1\}} \underset{[\text{theorem: } 11.198]}{=} \Phi_1$ so that if we take $\varepsilon_\sigma = 1$ that $\sigma\Phi_1 = \varepsilon_\sigma\Phi_1$.

$n \in \mathbb{N} \setminus \{1\}$. Using [theorem: 11.195] 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: 11.207] 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 δ_σ such that $\sigma\Phi_n = \delta_\sigma \cdot \Phi_n$ then $\varepsilon_\sigma \cdot \sigma = \delta_\sigma \cdot \sigma$. So $\varepsilon_\sigma \cdot \sigma(1,\ldots,n) = \delta_\sigma \cdot \sigma(1,\ldots,n)$, as by [theorem: 11.205] we have that $\sigma(1,\ldots,n) \neq 0$, so by multiplying both sides with $\sigma(1,\ldots,n)^{-1}$ we have $\varepsilon_\sigma = \delta_\sigma$. □

The above theorem ensures that the following definition makes sense.

Definition 11.209. *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* $\operatorname{sign}(\sigma) = 1$ *and* ***odd*** *if* $\operatorname{sign}(\sigma) = -1$

The concept of even or odd permutation follows from the following.

Remark 11.210. Let $n \in \mathbb{N} \setminus \{1\}$ and $\sigma \in P_n$ then by [theorem: 11.195] 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

$$\sigma\Phi_n = \left(\left(i_1 \underset{n}{\leftrightarrow} j_1\right) \circ \cdots \circ \left(i_m \underset{n}{\leftrightarrow} j_m\right)\right)\Phi_n \underset{\text{[theorem: 11.207]}}{=} (-1)^m \Phi_n$$

proving that

$$\operatorname{sign}(\sigma) = (-1)^m$$

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 event number of transpositions σ is even.

Theorem 11.211. *If* $n \in \mathbb{N}$ *then we have*

1. $\forall \sigma, \tau \in P_n$ *that* $\operatorname{sign}(\sigma \circ \tau) = \operatorname{sign}(\sigma) \cdot \operatorname{sign}(\tau)$
2. *If* $i, j \in \{1, \ldots, n\}$ *with* $i \neq j$ *then* $\operatorname{sign}\left(\left(i \underset{n}{\leftrightarrow} j\right)\right) = -1$
3. $\operatorname{sign}(\mathrm{Id}_{\{1,\ldots,n\}}) = 1$
4. $\forall \sigma \in P_n$ $\operatorname{sign}(\sigma) = \operatorname{sign}(\sigma^{-1})$

Proof.

1.
$$\begin{aligned} (\sigma \circ \tau)\Phi_n &\underset{\text{[theorem: 11.199]}}{=} \tau(\sigma\Phi_n) \\ &= \tau(\operatorname{sign}(\sigma) \cdot \Phi_n) \\ &\underset{\text{[theorem:11.200]}}{=} \operatorname{sign}(\sigma) \cdot (\tau\Phi_n) \\ &= \operatorname{sign}(\sigma) \cdot \operatorname{sign}(\tau) \cdot \Phi_n \end{aligned}$$

   proving that

$$\operatorname{sign}(\sigma \circ \tau) = \operatorname{sign}(\sigma) \cdot \operatorname{sign}(\tau)$$

2. As $\left(i \underset{n}{\leftrightarrow} j\right)\Phi \underset{\text{[theorem: 11.206]}}{=} (-1) \cdot \Phi$ we have

$$\operatorname{sign}\left(\left(i \underset{n}{\leftrightarrow} j\right)\right) = -1$$

3. As $\mathrm{Id}_{\{1,\ldots,n\}}\Phi_n \underset{\text{[theorem: 11.198]}}{=} \Phi_n$ we have

$$\operatorname{sign}(\mathrm{Id}_{\{1,\ldots,n\}}) = 1$$

4. We have

$$1 \underset{(2)}{=} \operatorname{sign}(\mathrm{Id}_{\{1,\ldots,n\}}) = \operatorname{sign}(\sigma \circ \sigma^{-1}) \underset{(1)}{=} \operatorname{sign}(\sigma) \cdot \operatorname{sign}(\sigma^{-1})$$

   so that $\operatorname{sign}(\sigma) = \operatorname{sign}(\sigma) \cdot 1 = (\operatorname{sign}(\sigma) \cdot \operatorname{sign}(\sigma)) \cdot \operatorname{sign}(\sigma^{-1}) \underset{\operatorname{sing}(\sigma) \in \{-1,1\}}{=} \operatorname{sign}(\sigma^{-1})$

□

The following permutation will be used in determinant functions later.

Definition 11.212. *Let* $n \in \mathbb{N}$ *and* $i, j \in \{1, \ldots, n\}$ *then we define*

$$\left(i \underset{n}{\rightsquigarrow} j\right) : \{1, \ldots, n\} \to \{1, \ldots, n\}$$

as follows:

1. *If* $i = j$ *then* $\left(i \underset{n}{\rightsquigarrow} j\right) := \mathrm{Id}_{\{1,\ldots,n\}}$

2. *If* $i < j$ *then for* $k \in \{1, \ldots, n\}$ *we have*

$$\left(i \underset{n}{\rightsquigarrow} j\right)(k) = \begin{cases} k \textit{ if } 1 \leqslant k < i \\ k + 1 \textit{ if } i \leqslant k < j \\ i \textit{ if } k = j \\ k \textit{ if } j < k \leqslant n \end{cases}$$

3. *If* $j < i$ *then for* $k \in \{1, \ldots, n\}$ *we have*

$$\left(i \underset{n}{\rightsquigarrow} j\right)(k) = \begin{cases} k \textit{ if } 1 \leqslant k < j \\ i \textit{ if } k = j \\ k - 1 \textit{ if } j < k \leqslant i \\ k \textit{ if } i < k \leqslant n \end{cases}$$

Example 11.213. To get a idea of the mapping $\left(i \underset{n}{\rightsquigarrow} j\right)$ we can look at some examples:

1. $n = 6$

$$\left(2 \underset{6}{\rightsquigarrow} 5\right)\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 4 \\ 5 \\ 2 \\ 6 \end{pmatrix}$$

2. $n = 6$

$$\left(5 \underset{6}{\rightsquigarrow} 2\right)\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{pmatrix} = \begin{pmatrix} 1 \\ 5 \\ 2 \\ 3 \\ 4 \\ 6 \end{pmatrix}$$

So we interpret $\left(i \underset{n}{\rightsquigarrow} j\right)$ as removing i from its position an inserting it after (before) j if $i < j$ $(j < i)$.

Next we prove that $\left(i \underset{n}{\rightsquigarrow} j\right)$ 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 11.214. *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* $\theta^{i>j}: \{1, \ldots, i - j\} \rightarrow \{j, \ldots, i - 1\}$ *by* $\theta^{i>j}(l) = j + (l - 1)$

Proof. Of course we must prove that the range is correct.

1. We have

$$\begin{aligned} l \in \{1, \ldots, j - i\} &\Leftrightarrow 1 \leqslant l \leqslant j - i \\ &\Leftrightarrow 0 \leqslant l - 1 \leqslant j - i - 1 \\ &\Leftrightarrow i - j + 1 \leqslant -(l - 1) \leqslant 0 \\ &\Leftrightarrow i + 1 \leqslant j - (l - 1) \leqslant j \\ &\Leftrightarrow i + 1 \leqslant \theta^{i<j}(l) \leqslant j \\ &\Leftrightarrow \theta^{i<j}(l) \in \{i + 1, \ldots, j\} \end{aligned}$$

2. We have

$$\begin{aligned} l\in\{1,\ldots,i-j\} &\Leftrightarrow 1\leqslant l\leqslant i-j\\ &\Leftrightarrow 0\leqslant l-1\leqslant i-j-1\\ &\Leftrightarrow j\leqslant j+(l-1)\leqslant i-1\\ &\Leftrightarrow j\leqslant \theta^{i>j}(l)\leqslant i-1\\ &\Leftrightarrow \theta^{i>j}(l)\in\{j,\ldots,i-1\}\\ &\quad\ \square \end{aligned}$$

Theorem 11.215. *If $n\in\mathbb{N}$ and $i,j\in\{1,\ldots,n\}$ then we have:*

1. *If $i<j$ then*
$$\big(i \underset{n}{\rightsquigarrow} j\big)=\big(i \underset{n}{\leftrightarrow} \theta^{i<j}(1)\circ\cdots\circ\big(i \underset{n}{\leftrightarrow} \theta^{i<j}(j-i)\big)\big)$$
2. *If $j<i$ then*
$$\big(i \underset{n}{\rightsquigarrow} j\big)=\big(i \underset{n}{\leftrightarrow} \theta^{i<j}(1)\circ\cdots\circ\big(i \underset{n}{\leftrightarrow} \theta^{i<j}(i-j)\big)$$
3. $\big(i \underset{n}{\rightsquigarrow} j\big)\in P_n$
4. $\operatorname{sign}\big(\big(i \underset{n}{\rightsquigarrow} j\big)\big)=(-1)^{|i-j|}$

Proof.

1. Let $n\in\mathbb{N}$. We proceed now by induction on $k=j-i$, so let

   $S_n=\big\{k\in\mathbb{N}|\text{If } i,j\in\{1,\ldots,n\}\text{with } i<j\wedge j-i=k \text{ then } \big(i \underset{n}{\rightsquigarrow} j\big)=\big(i \underset{n}{\leftrightarrow} \theta^{i<j}(1)\circ\cdots\circ \big(i \underset{n}{\leftrightarrow} \theta^{i<j}(j-i)\big)\big)=\big(i \underset{n}{\leftrightarrow} \theta^{i<j}(1)\circ\cdots\circ\big(i \underset{n}{\leftrightarrow} \theta^{i<j}(j-i)\big)\big)\big\}$

   then we have:

   $1\in S$. Let $i,j\in\{1,\ldots,n\}$ with $i<j\wedge j-i=1$. Then $\theta^{i<j}(1)=j-(1-1)=j=i+1$ and we have for $l\in\{1,\ldots,n\}$ either:

   $1\leqslant l<i$. Then

$$\begin{aligned} \big(i \underset{n}{\rightsquigarrow} j\big)(l) &= l\\ &\underset{l<i<i+1\wedge l<i<j}{=} \big(i \underset{n}{\leftrightarrow} i+1\big)(l)\\ &= \big(i \underset{n}{\leftrightarrow} \theta^{i<j}(l)\big)(l) \end{aligned}$$

   $i\leqslant l<j$. Then $i\leqslant l<i+1$ so that $l=i$ and

$$\begin{aligned} \big(i \underset{n}{\rightsquigarrow} j\big)(l) &= l+1\\ &\underset{l=i}{=} i+1\\ &= \big(i \underset{n}{\leftrightarrow} i+1\big)(i)\\ &\underset{l=i}{=} \big(i \underset{n}{\leftrightarrow} i+1\big)(l)\\ &= \big(i \underset{n}{\leftrightarrow} \theta^{i<j}(l)\big)(l) \end{aligned}$$

   $l=j$. Then $l=j=i+1$ and

$$\begin{aligned} \big(i \underset{n}{\rightsquigarrow} j\big)(l) &= i\\ &= \big(i \underset{n}{\leftrightarrow} i+1\big)(i+1)\\ &\underset{l=i+1}{=} \big(i \underset{n}{\leftrightarrow} i+1\big)(l)\\ &= \big(i \underset{n}{\leftrightarrow} \theta^{i<j}(l)\big)(l) \end{aligned}$$

$j<l\leqslant n$. Then

$$\begin{array}{rcl} \big(i \underset{n}{\rightsquigarrow} j\big)(l) & = & l \\ & \underset{i<j<l}{=} & \big(i \underset{n}{\leftrightarrow} i+1\big)(l) \\ & = & \big(i \underset{n}{\leftrightarrow} \theta^{i<j}(l)\big)(l) \end{array}$$

So we have in all cases $\big(i \underset{n}{\rightsquigarrow} j\big)(l)=\big(i \underset{n}{\leftrightarrow} \theta^{i<j}(l)\big)(l)$ so that

$$\big(i \underset{n}{\rightsquigarrow} j\big)=\big(i \underset{n}{\leftrightarrow} \theta^{i<j}(l)\big)=\big(i \underset{n}{\leftrightarrow} \theta^{i<j}(l)\big)\circ\cdots\circ\big(i \underset{n}{\leftrightarrow} \theta^{i<j}(l)\big)$$

proving that $1\in S_n$.

$k\in S\Rightarrow k+1\in S$. Let $i,j\in\{1,\ldots,n\}$ with $i<j$ and $j-i=k+1$. First we prove that

$$\big(i \underset{n}{\rightsquigarrow} j\big)=\big(i \underset{n}{\leftrightarrow} j\big)\circ\big(i \underset{n}{\rightsquigarrow} j-1\big) \tag{11.108}$$

Proof. For $l\in\{1,\ldots,n\}$ we have the following cases to consider:

$1\leqslant l<i$. Then

$$\begin{array}{rcl} \big(i \underset{n}{\leftrightarrow} j\big)\big(\big(i \underset{n}{\rightsquigarrow} j-1\big)(l)\big) & = & \big(i \underset{n}{\leftrightarrow} j\big)(l) \\ & \underset{l\neq i,j}{=} & l \\ & \underset{l<i}{=} & \big(i \underset{n}{\rightsquigarrow} j\big)(l) \end{array}$$

$i\leqslant l<j-1$. Then

$$\begin{array}{rcl} \big(i \underset{n}{\leftrightarrow} j\big)\big(\big(i \underset{n}{\rightsquigarrow} j-1\big)(l)\big) & = & \big(i \underset{n}{\leftrightarrow} j\big)(l+1) \\ & \underset{i<i+1\leqslant l+1<j}{=} & l+1 \\ & \underset{i\leqslant l+1<j}{=} & \big(i \underset{n}{\rightsquigarrow} j\big)(l) \end{array}$$

$l=j-1$. Then

$$\begin{array}{rcl} \big(i \underset{n}{\leftrightarrow} j\big)\big(\big(i \underset{n}{\rightsquigarrow} j-1\big)(l)\big) & = & \big(i \underset{n}{\leftrightarrow} j\big)(i) \\ & = & j \\ & \underset{i<j\Rightarrow i\leqslant j-1=l<j}{=} & \big(i \underset{n}{\rightsquigarrow} j\big)(j-1) \\ & \underset{l=j-1}{=} & \big(i \underset{n}{\rightsquigarrow} j\big)(l) \end{array}$$

$j-1<l\leqslant n$. Then $j\leqslant l$ so we have either:

$j=l$. Then

$$\begin{array}{rcl} \big(i \underset{n}{\leftrightarrow} j\big)\big(\big(i \underset{n}{\rightsquigarrow} j-1\big)(l)\big) & = & \big(i \underset{n}{\leftrightarrow} j\big)(l) \\ & \underset{l=j}{=} & i \\ & \underset{l=j}{=} & \big(i \underset{n}{\rightsquigarrow} j\big)(l) \end{array}$$

$j<l$. Then

$$\begin{array}{rcl} \big(i \underset{n}{\leftrightarrow} j\big)\big(\big(i \underset{n}{\rightsquigarrow} j-1\big)(l)\big) & \underset{j-1<j<l}{=} & \big(i \underset{n}{\leftrightarrow} j\big)(l) \\ & \underset{i<j<l\Rightarrow i,j\neq l}{=} & l \\ & \underset{j<l\leqslant n}{=} & \big(i \underset{n}{\rightsquigarrow} j\big)(l) \end{array}$$

So in all cases $\big(i \underset{n}{\leftrightarrow} j\big)\big(\big(i \underset{n}{\rightsquigarrow} j-1\big)(l)\big)=\big(i \underset{n}{\rightsquigarrow} j\big)(l)$ or $\big(i \underset{n}{\rightsquigarrow} j\big)=\big(i \underset{n}{\leftrightarrow} j\big)\circ\big(i \underset{n}{\rightsquigarrow} j-1\big)$ proving [eq: 11.108]. □

As $\theta^{i<j}(1)=j+(1-1)=j$ we have by [eq: 11.108] that

$$\big(i \underset{n}{\rightsquigarrow} j\big)=\big(i \underset{n}{\leftrightarrow} \theta^{i<j}(1)\big)\circ\big(i \underset{n}{\rightsquigarrow} j-1\big) \tag{11.109}$$

Further as $(j-1)-i=(k+1)-1=k$ we have as $k\in S_n$ that

$$\left(i\underset{n}{\rightsquigarrow}j-1\right)=\left(i\underset{n}{\leftrightarrow}\theta^{i<j-1}(1)\right)\circ\cdots\circ\left(i\underset{n}{\leftrightarrow}\theta^{i<j-1}(j-1-i)\right)$$

Let $l\in\{1,\ldots,j-i\}$ then $\theta^{i<j-1}(l)=j-1-(l-1)=j-((l+1)-1)=\theta^{i<j}(l+1)$ we have that

$$\left(i\underset{n}{\rightsquigarrow}j-1\right)=\left(i\underset{n}{\leftrightarrow}\theta^{i<j}(2)\right)\circ\cdots\circ\left(i\underset{n}{\leftrightarrow}\theta^{i<j-1}(j-i)\right)\tag{11.110}$$

Further

$$\begin{array}{rcl}\left(i\underset{n}{\rightsquigarrow}j\right) & \underset{[\text{eq: }11.109]}{=} & \left(i\underset{n}{\leftrightarrow}\theta^{i<j}(1)\right)\circ\left(i\underset{n}{\rightsquigarrow}j-1\right)\\ & = & \left(i\underset{n}{\leftrightarrow}\theta^{i<j}(1)\right)\circ\left(\left(i\underset{n}{\leftrightarrow}\theta^{i<j}(2)\right)\circ\cdots\circ\left(i\underset{n}{\leftrightarrow}\theta^{i<j-1}(j-i)\right)\right)\\ & = & \left(i\underset{n}{\leftrightarrow}\theta^{i<j}(1)\right)\circ\left(i\underset{n}{\leftrightarrow}\theta^{i<j}(2)\right)\circ\cdots\circ\left(i\underset{n}{\leftrightarrow}\theta^{i<j-1}(j-i)\right)\\ & = & \left(i\underset{n}{\leftrightarrow}\theta^{i<j}(1)\right)\circ\cdots\circ\left(i\underset{n}{\leftrightarrow}\theta^{i<j-1}(j-i)\right)\end{array}$$

proving that $k+1\in\mathcal{S}_n$.

2. Let $n\in\mathbb{N}$. We proceed now by induction on $k=j-i$, so let

   $S_n=\left\{k\in\mathbb{N}|\text{If } i,j\in\{1,\ldots,n\}\text{with } j<i\wedge i-j=k \text{ then } \left(i\underset{n}{\rightsquigarrow}j\right)=\left(i\underset{n}{\leftrightarrow}\theta^{i>j}(1)\circ\cdots\circ\left(i\underset{n}{\leftrightarrow}\theta^{i>j}(i-j)\right)\right)=\left(i\underset{n}{\leftrightarrow}\theta^{i>j}(1)\circ\cdots\circ\left(i\underset{n}{\leftrightarrow}\theta^{i>j}(i-j)\right)\right)\right\}$

   then we have:

   $\boldsymbol{1\in S}$. Let $i,j\in\{1,\ldots,n\}$ with $j<i$ and $i-j=1$ so that $j=i-1$. Then

   $$\theta^{i>j}(1)=j+(1-1)=j=i-1$$

   For $l\in\{1,\ldots,n\}$ we have either:

   $\boldsymbol{1\leqslant l<j}$. Then we have

   $$\begin{array}{rcl}\left(i\underset{n}{\rightsquigarrow}j\right) & = & l\\ & \underset{l<j<i\Rightarrow l\neq j,i-1}{=} & \left(i\underset{n}{\leftrightarrow}i-1\right)(l)\\ & = & \left(i\underset{n}{\leftrightarrow}\theta^{i>j}(1)\right)(l)\end{array}$$

   $\boldsymbol{l=j}$. Then we have $l=j=i-1$

   $$\begin{array}{rcl}\left(i\underset{n}{\rightsquigarrow}j\right)(l) & = & i\\ & = & \left(i\underset{n}{\leftrightarrow}i-1\right)(i-1)\\ & = & \left(i\underset{n}{\leftrightarrow}i-1\right)(l)\\ & = & \left(i\underset{n}{\leftrightarrow}\theta^{i>j}(1)\right)(l)\end{array}$$

   $\boldsymbol{j<l\leqslant i}$. Then we have $i-1<l\leqslant i$ so that $l=i$ and

   $$\begin{array}{rcl}\left(i\underset{n}{\rightsquigarrow}j\right)(l) & = & l-1\\ & \underset{l=i}{=} & i-1\\ & = & \left(i\underset{n}{\leftrightarrow}i-1\right)(i)\\ & \underset{l=i}{=} & \left(i\underset{n}{\leftrightarrow}i-1\right)(l)\\ & = & \left(i\underset{n}{\leftrightarrow}\theta^{i>j}(1)\right)(l)\end{array}$$

   $\boldsymbol{i<l\leqslant n}$. Then we have

   $$\begin{array}{rcl}\left(i\underset{n}{\rightsquigarrow}j\right)(l) & = & l\\ & \underset{j<i<l}{=} & \left(i\underset{n}{\leftrightarrow}i-1\right)(l)\\ & = & \left(i\underset{n}{\leftrightarrow}\theta^{i>j}(1)\right)(l)\end{array}$$

So in all cases we have $\big(i \underset{n}{\rightsquigarrow} j\big)(l) = \big(i \underset{n}{\leftrightarrow} \theta^{i>j}(1)\big)(l)$, hence

$$\big(i \underset{n}{\rightsquigarrow} j\big) = \big(i \underset{n}{\leftrightarrow} \theta^{i>j}(1)\big) = \big(i \underset{n}{\leftrightarrow} \theta^{i>j}(1)\big) \circ \cdots \circ \big(i \underset{n}{\leftrightarrow} \theta^{i>j}(1)\big)$$

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

$$\big(i \underset{n}{\rightsquigarrow} j\big) = \big(i \underset{n}{\leftrightarrow} j\big) \circ \big(i \underset{n}{\rightsquigarrow} j+1\big) \tag{11.111}$$

Proof. For $l \in \{1, \ldots, n\}$ we have either:

$\boldsymbol{1 \leqslant l < j.}$ Then

$$\begin{aligned}\big(i \underset{n}{\leftrightarrow} j\big)\big(\big(i \underset{n}{\rightsquigarrow} j+1\big)(l)\big) &\underset{l<j<j+1}{=} \big(i \underset{n}{\leftrightarrow} j\big)(l)\\ &\underset{l<j<i}{=} l\\ &\underset{l<j}{=} \big(i \underset{n}{\rightsquigarrow} j\big)(l)\end{aligned}$$

$\boldsymbol{l = j.}$ Then

$$\begin{aligned}\big(i \underset{n}{\leftrightarrow} j\big)\big(\big(i \underset{n}{\rightsquigarrow} j+1\big)(l)\big) &\underset{l=j<j+1}{=} \big(i \underset{n}{\leftrightarrow} j\big)(l)\\ &\underset{l=j}{=} \big(i \underset{n}{\leftrightarrow} j\big)(j)\\ &= i\\ &\underset{l=j}{=} \big(i \underset{n}{\rightsquigarrow} j\big)(l)\end{aligned}$$

$\boldsymbol{j < l \leqslant i.}$ Then we have $j + 1 \leqslant l$ so that either:

$\boldsymbol{l = j + 1.}$ Then

$$\begin{aligned}\big(i \underset{n}{\leftrightarrow} j\big)\big(\big(i \underset{n}{\rightsquigarrow} j+1\big)(l)\big) &= \big(i \underset{n}{\leftrightarrow} j\big)(i)\\ &= j\\ &\underset{l=j+1}{=} l - 1\\ &\underset{j<l\leqslant i}{=} \big(i \underset{n}{\rightsquigarrow} j\big)(l)\end{aligned}$$

$\boldsymbol{j + 1 < l.}$ Then

$$\begin{aligned}\big(i \underset{n}{\leftrightarrow} j\big)\big(\big(i \underset{n}{\rightsquigarrow} j+1\big)(l)\big) &= \big(i \underset{n}{\leftrightarrow} j\big)(l-1)\\ &\underset{j<j+1\leqslant l-1<l\leqslant i}{=} l - 1\\ &\underset{j<l\leqslant i}{=} \big(i \underset{n}{\rightsquigarrow} j\big)(l)\end{aligned}$$

$\boldsymbol{i < l \leqslant n.}$ Then

$$\begin{aligned}\big(i \underset{n}{\leftrightarrow} j\big)\big(\big(i \underset{n}{\rightsquigarrow} j+1\big)(l)\big) &\underset{i<l<l+1}{=} \big(i \underset{n}{\leftrightarrow} j\big)(l)\\ &\underset{j<i<l}{=} l\\ &\underset{i<l}{=} \big(i \underset{n}{\rightsquigarrow} j\big)(l)\end{aligned}$$

So in all cases we have $\big(i \underset{n}{\leftrightarrow} j\big)\big(\big(i \underset{n}{\rightsquigarrow} j+1\big)(l)\big) = \big(i \underset{n}{\rightsquigarrow} j\big)(l)$, hence $\big(i \underset{n}{\rightsquigarrow} j\big) = \big(i \underset{n}{\leftrightarrow} j\big) \circ \big(i \underset{n}{\rightsquigarrow} j+1\big)$, proving that [eq: 11.111]. □

Now $\theta^{i>j}(1) = j + (1 - 1) = j$ which combined with [eq: 11.111] gives

$$\big(i \underset{n}{\rightsquigarrow} j\big) = \big(i \underset{n}{\leftrightarrow} \theta^{i>j}(1)\big) \circ \big(i \underset{n}{\rightsquigarrow} j+1\big) \tag{11.112}$$

Further as $i - (j+1) = i - j - 1 = k + 1 - 1 = k \in \mathcal{S}_n$ we have that

$$\big(i \underset{n}{\rightsquigarrow} j+1\big) = \big(i \underset{n}{\leftrightarrow} \theta^{i>j+1}(1)\big) \circ \cdots \circ \big(i \underset{n}{\leftrightarrow} \theta^{i>j+1}(i-(j+1))\big)$$

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

$$\big(i \underset{n}{\rightsquigarrow} j+1\big) = \big(i \underset{n}{\leftrightarrow} \theta^{i>j}(2)\big) \circ \cdots \circ \big(i \underset{n}{\leftrightarrow} \theta^{i>j}(i-j)\big) \tag{11.113}$$

Next

$$\begin{array}{rcl} \big(i \underset{n}{\rightsquigarrow} j\big) & = & \big(i \underset{n}{\leftrightarrow} \theta^{i>j}(1)\big) \circ \big(i \underset{n}{\rightsquigarrow} j+1\big) \\ & \underset{\text{[eq: 11.113}}{=} & \big(i \underset{n}{\leftrightarrow} \theta^{i>j}(1)\big) \circ \Big(\big(i \underset{n}{\leftrightarrow} \theta^{i>j}(2)\big) \circ \cdots \circ \big(i \underset{n}{\leftrightarrow} \theta^{i>j}(i-j)\big)\Big) \\ & = & \big(i \underset{n}{\leftrightarrow} \theta^{i>j}(1)\big) \circ \big(i \underset{n}{\leftrightarrow} \theta^{i>j}(2)\big) \circ \cdots \circ \big(i \underset{n}{\leftrightarrow} \theta^{i>j}(i-j)\big) \\ & = & \big(i \underset{n}{\leftrightarrow} \theta^{i>j}(1)\big) \circ \cdots \circ \big(i \underset{n}{\leftrightarrow} \theta^{i>j}(i-j)\big) \end{array}$$

proving that $k+1 \in S_n$.

3. Let $i, j \in \{1, \ldots, n\}$ then we have either:

   $\boldsymbol{i=j}$**.** Then $\big(i \underset{n}{\rightsquigarrow} j\big) \underset{\text{def}}{=} \mathrm{Id}_{1,\ldots,n} \in P_n$

   $\boldsymbol{i<j}$**.** Then by (1) $\big(i \underset{n}{\rightsquigarrow} j\big) = \big(i \underset{n}{\leftrightarrow} \theta^{i<j}(1)\big) \circ \cdots \circ \big(i \underset{n}{\leftrightarrow} \theta^{i<j}(j-i)\big) \in P_n$

   $\boldsymbol{j<i}$**.** Then by (2) $\big(i \underset{n}{\rightsquigarrow} j\big) = \big(i \underset{n}{\leftrightarrow} \theta^{i>j}(1)\big) \circ \cdots \circ \big(i \underset{n}{\leftrightarrow} \theta^{i>j}(i-j)\big) \in P_n$

4. Let $i, j \in \{1, \ldots, n\}$ then we have either:

   $\boldsymbol{i=j}$**.** Then $|i-j| = 0$ and

   $$\mathrm{sign}\big(\big(i \underset{n}{\rightsquigarrow} j\big)\big) = \mathrm{sign}(\mathrm{Id}_{\{1,\ldots,\}}) \underset{\text{[theorem: 11.211]}}{=} 1 = (-1)^0 = (-1)^{|i-j|}$$

   $\boldsymbol{i<j}$**.** Then as $\theta^{i<j}(\{1,\ldots,j-i\}) \underset{\text{[definition: 11.214]}}{=} \{i+1,\ldots,j\}$ so that $\forall l \in \{1,\ldots,j-i\}$ we have that $i \neq \theta^{i<j}(l)$, so $\big(i \underset{n}{\leftrightarrow} \theta^{i<j}()\big)$ is a strict transposition, hence

   $$\mathrm{sign}\big(\big(i \underset{n}{\rightsquigarrow} j\big)\big) \underset{(1)}{=} \mathrm{sign}\big(\big(i \underset{n}{\leftrightarrow} \theta^{i<j}(1)\big) \circ \cdots \circ \big(i \underset{n}{\leftrightarrow} \theta^{i<j}(j-i)\big)\big) \underset{\text{[theorem: 11.210]}}{=} (-1)^{|i-j|}$$

   $\boldsymbol{j<i}$**.** Then as $\theta^{i<j}(\{1,\ldots,i-j\}) \underset{\text{[definition: 11.214]}}{=} \{j,\ldots,i-1\}$ so that $\forall l \in \{1,\ldots,i-j\}$ we have that $i \neq \theta^{i<j}(l)$, so $\big(i \underset{n}{\leftrightarrow} \theta^{i<j}()\big)$ is a strict transposition, hence

   $$\mathrm{sign}\big(\big(i \underset{n}{\rightsquigarrow} j\big)\big) \underset{(1)}{=} \mathrm{sign}\big(\big(i \underset{n}{\leftrightarrow} \theta^{i<j}(1)\big) \circ \cdots \circ \big(i \underset{n}{\leftrightarrow} \theta^{i<j}(j-i)\big)\big) \underset{\text{[theorem: 11.210]}}{=} (-1)^{|i-j|}$$

□

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.

Definition 11.216. *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*

$$(\ldots, x_{i-1}, a, x_{i+1}, \ldots) \text{ is equivalent with } x_i = a$$

using this notation we have

$$x = (x_1, \ldots, x_{i-1}, a, x_{i+1}, \ldots, x_n) \textit{ is equivalent with saying that } x_i = a$$

and if $i \neq j$ then

$$x = (x_1, \ldots, x_{i-1}, a, x_{i+1}, \ldots x_{j-1}, b, x_{j+1} \ldots, x_n) \textit{ is equivalent with } x_i = a \wedge x_j = b$$

...

Proposition 11.217. *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 $x\in\prod_{i\in\{1,\ldots,n\}}X_i$ if $x=(x_1,\ldots,x_{i-1},a,x_{i+1},\ldots,x_n)$ then*

$$x\circ\sigma \;=\; ((x\circ\sigma)_1,\ldots,(x\circ\sigma)_{\sigma^{-1}(i)-1},a,(x\circ\sigma)_{\sigma^{-1}(i)+1},\ldots,(x\circ\sigma)_n)$$

Proof. Let $i\in\{1,\ldots n\}$. As $x=(x_1,\ldots,x_{i-1},a,x_{i+1},\ldots,x_n)$ we have that $x(i)=x_i=a$. As

$$(x\circ\sigma)_{\sigma^{-1}(i)}=x(\sigma(\sigma^{-1}(i)))=x(i)=a$$

we have by definition that

$$x\circ\sigma=((x\circ\sigma)_1,\ldots,(x\circ\sigma)_{\sigma^{-1}(i)-1},a,(x\circ\sigma)_{\sigma^{-1}(i)+1},\ldots,(x\circ\sigma)_n)$$ □

Proof. For $i\in\{1,\ldots,n\}\setminus\{i\}$ we have $x_{\sigma(i)}=x(\sigma(i))=(x\circ\sigma)(i)=(x\circ\sigma)_i$ and $a=$ □

Definition 11.218. *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 $L\colon\prod_{j\in\{1,\ldots,n\}}X_j\rightarrow Y$ 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: 11.216] this can be written more clearly as

1. *$\forall u,v\in X_i$ and $\forall(x_1,\ldots,x_{i-1},u+v,x_{i+1},\ldots,x_n)\in\prod_{j\in\{1,\ldots,n\}}X_j$*

$$L(x_1,\ldots,x_{i-1},u+v,x_{i+1},\ldots,x_n)=L(x_1,\ldots,x_{i-1},u,x_{i+1},\ldots,x_n)+L(x_1,\ldots,x_{i-1},v,x_{i+1},\ldots,x_n)$$

2. *$\forall\alpha\in F$, $\forall u\in X_i$ and $\forall(x_1,\ldots,x_{i-1},\alpha\cdot u,x_{i+1},\ldots,x_n)\in\prod_{j\in\{1,\ldots,n\}}X$*

$$L(x_1,\ldots,x_{i-1},\alpha\cdot u,x_{i+1},\ldots,x_n)=\alpha\cdot L(x_1,\ldots,x_{i-1},u,x_{i+1},\ldots,x_n)$$

The set of graphs of multilinear mappings is noted as $\mathrm{Hom}(X_1,\ldots,X_n;Y)$, more specific:

$$\mathrm{Hom}(X_1,\ldots,X_n;Y)=\left\{L\in Y^{\prod_{i\in\{1,\ldots,n\}}X_i}|L\colon\prod_{i\in\{1,\ldots,n\}}X_i\rightarrow Y \text{ is multilinear}\right\}$$

Theorem 11.219. *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 $L\in\mathrm{Hom}(X_1,\ldots,X_n;Y)$ 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

$$x=(x_1,\ldots,x_{i-1},0,x_{i+1},\ldots,x_n)$$

so that

$$\begin{aligned}L(x) &= L(x_1,\ldots,x_{i-1},0,x_{i+1},\ldots,x_n)\\ &= L(x_1,\ldots,x_{i-1},0\cdot 0,x_{i+1},\ldots,x_n)\\ &= 0\cdot L(x_1,\ldots,x_{i-1},0,x_{i+1},\ldots,x_n)\\ &= 0\end{aligned}$$

□

Just a with linear functions we have a alternative and simpler condition for multilinearity.

Theorem 11.220. *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 $L:\prod_{i\in\{1,\ldots,n\}}X_i\rightarrow Y$ we have*

$$L:\prod_{i\in\{1,\ldots,n\}}X_i\rightarrow Y \text{ is a multilinear mapping}$$

$$\Updownarrow$$

$$\forall i\in\{1,\ldots,n\},\forall x,y\in X \text{ and } \alpha \text{ we have :}$$

$$L(x_1,\ldots,x_{i-1},u+\alpha\cdot v,x_{i+1},\ldots,x_n)=L(x_1,\ldots,x_{i-1},u,x_{i+1},\ldots,x_n)+\alpha\cdot L(x_1,\ldots,x_{i-1},v,x_{i+1},\ldots,x_n)$$

Proof. We have:

⇒. Let $i\in\{1,\ldots,n\}$, $x,y\in X_i$ and $\alpha\in F$ then we have

$$\begin{aligned}L(x_1,\ldots,x_{i-1,}u+\alpha\cdot v,x_{i+1},\ldots,x_n) &\underset{\text{multilinearity}}{=}\\ L(x_1,\ldots,x_{i-1,}u,x_{i+1},\ldots,x_n)+L(x_1,\ldots,x_{i-1,}\alpha\cdot v,x_{i+1},\ldots,x_n) &\underset{\text{multilinearity}}{=}\\ L(x_1,\ldots,x_{i-1,}u,x_{i+1},\ldots,x_n)+\alpha\cdot L(x_1,\ldots,x_{i-1,}v,x_{i+1},\ldots,x_n)\end{aligned}$$

⇐. Let $i\in\{1,\ldots,n\}$, $x,y\in X_i$ and $\alpha\in F$ then we have:

$$\begin{aligned}L(x_1,\ldots,x_{i-1,}u+v,x_{i+1},\ldots,x_n) &=\\ L(x_1,\ldots,x_{i-1,}u+1\cdot v,x_{i+1},\ldots,x_n) &=\\ L(x_1,\ldots,x_{i-1,}u,x_{i+1},\ldots,x_n)+1\cdot L(x_1,\ldots,x_{i-1,}v,x_{i+1},\ldots,x_n)\end{aligned}$$

and

$$\begin{aligned}L(x_1,\ldots,x_{i-1,}\alpha\cdot u,x_{i+1},\ldots,x_n) &=\\ L(x_1,\ldots,x_{i-1,}0+\alpha\cdot u,x_{i+1},\ldots,x_n) &=\\ L(x_1,\ldots,x_{i-1,}0,x_{i+1},\ldots,x_n)+\alpha\cdot L(x_1,\ldots,x_{i-1,}u,x_{i+1},\ldots,x_n) &\underset{[\text{theorem: } 11.219]}{=}\\ \alpha\cdot L(x_1,\ldots,x_{i-1,}u,x_{i+1},\ldots,x_n)\end{aligned}$$

□

A special case of the above is where $\{X_i\}_{i\in\{1,\ldots,n\}}$ satisfies $\forall i\in\{1,\ldots,n\}$ we have that $X_i=X$ so that $\prod_{i\in\{1,\ldots,n\}}X_i=X^n$. This leads to the definition of n-linear functions.

Definition 11.221. *Let $n\in\mathbb{N}$ and X,Y vector spaces over a field F then a multilinear mapping*

$$L:X^n\rightarrow Y$$

*is called a **n-linear mapping**. Further set of graphs of n-linear mappings is noted as* $\mathrm{Hom}(X^n;Y)$ *so*

$$\mathrm{Hom}(X^n;Y)=\mathrm{Hom}\left(\underbrace{X,\ldots,X}_{n};Y\right)$$

Example 11.222. Let $n\in\mathbb{N}$ and F a field then $L_\otimes:F^n\rightarrow F$ define by

$$L_\otimes(x)=\prod_{i\in\{1,\ldots,n\}}x_i$$

is multilinear or

$$L_{\otimes} \in \mathrm{Hom}(F^n; F)$$

Proof. This follows from [theorem: 11.46] □

To be able to use induction in proofs about multilinear mappings we have the following theorem.

Theorem 11.223. *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*

$$L_{\{\ldots a\}}\colon \prod_{j \in \{1,\ldots,n\}} X_j \rightarrow Y \text{ by } L_{\{\ldots a\}}(x_1,\ldots,x_n) = L(x_1,\ldots,x_n,a)$$

then

$$L_{\{\ldots a\}} \in \mathrm{Hom}(X_1,\ldots,X_{i-1},X_{i+1},\ldots,X_n)$$

Proof. Let $i \in \{1,...,n\}$, $\alpha \in F$, $u, v \in X_{n+1}$. If $(x_1,...,x_{i-1},x+y,x_{i+1},...,x_n) \in \prod_{j \in \{1,\ldots,n\}} X_j$ then

$$\begin{aligned} L_{\{\ldots a\}}(x_1,\ldots,x_{i-1},u+\alpha\cdot v,x_{i+1},\ldots,x_n) &= \\ L(x_1,\ldots,x_{i-1},u+\alpha\cdot v,x_{i+1},\ldots,x_n,a) &= \\ L(x_1,\ldots,x_{i-1},u,x_{i+1},\ldots,x_n,a) + \alpha\cdot L(x_1,\ldots,x_{i-1},v,x_{i+1},\ldots,x_n,a) &= \\ L_{\{\ldots a\}}(x_1,\ldots,x_{i-1},u,x_{i+1},\ldots,x_n) + \alpha\cdot L_{\{\ldots a\}}(x_1,\ldots,x_{i-1},v,x_{i+1},\ldots,x_n) & \end{aligned}$$

□

Definition 11.224. *Let $n \in \mathbb{N}$, $i \in \{1,\ldots,n\}$, $\{X_i\}_{j \in \{1,\ldots,n\}}$ a family of sets, Y a set, $(x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n) \in \prod_{j \in \{1,\ldots,n\}\setminus\{i\}} X_i$ and*

$$f\colon \prod_{j \in \{1,\ldots,n\}} X_i \rightarrow Y$$

then $f(x_1,\ldots,x_{i-1},*,x_{i+1},\ldots,x_n)\colon \prod_{i \in \{1,\ldots,n\}\setminus\{i\}} X_i \rightarrow Y$ is defined by

$$f(x_1,\ldots,x_{i-1},*,x_{i+1},\ldots,x_n)(x) = f(x_1,\ldots,x_{i-1},x,x_{i+1},\ldots,x_n)$$

Theorem 11.225. *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 $(x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n) \in \prod_{j \in \{1,\ldots,n\}\setminus\{i\}} X_i$ then*

$$L(x_1,\ldots,x_{i-1},*,x_{i+1},\ldots,x_n) \in \mathrm{Hom}(X_i, Y)$$

Proof. Let $\alpha \in F$, $u, v \in X_i$ then

$$\begin{aligned} L(x_1,\ldots,x_{i-1},*,x_{i+1},\ldots,x_n)(u+\alpha\cdot v) &= \\ L(x_1,\ldots,x_{i-1},u+\alpha\cdot v,x_{i+1},\ldots,x_n) &= \\ L(x_1,\ldots,x_{i-1},u,x_{i+1},\ldots,x_n) + \alpha\cdot L(x_1,\ldots,x_{i-1},v,x_{i+1},\ldots,x_n) &= \\ L(x_1,\ldots,x_{i-1},*,x_{i+1},\ldots,x_n)(u) + \alpha\cdot L(x_1,\ldots,x_{i-1},*,x_{i+1},\ldots,x_n)(v) & \end{aligned}$$

□

Example 11.226. 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

$$C_0(x_1,\ldots,u+v,\ldots,x_n) = 0 = 0 + 0 = C_0(x_1,\ldots,u,\ldots,x_n) + C_0(x_1,\ldots,v,\ldots,x_n)$$

and

$$C_0(x_1,\ldots,\alpha\cdot u,\ldots,x_n) = 0 = \alpha\cdot 0 = \alpha\cdot C_0(x_1,\ldots,u,\ldots,x_n)$$ □

Theorem 11.227. *Let $n \in \mathbb{N}$ and $\{X_i\}_{i \in \{1, \ldots, n\}}$ be a family of vector spaces over a field F and Y a vector space over the same field then $\operatorname{Hom}(X_1, \ldots, X_n; Y)$ is a sub-space of $\langle Y^{\prod_{i=1}^n X_i}, +, \cdot \rangle$. So by [theorem: 11.52] $\langle \operatorname{Hom}(X_1, \ldots, X_n; Y), +, \cdot \rangle$ is a vector space.*

Proof. First using [exercise: 11.226] we have that

$$C_0 \in \operatorname{Hom}(X_1, \ldots, X_n; Y) \Rightarrow \operatorname{Hom}(X_1, \ldots, X_n; Y) \neq \varnothing$$

Second if $\alpha, \beta \in F$ and $L_1, L_2 \in \operatorname{Hom}(x_1, \ldots, X_n; Y)$ and $i \in \{1, \ldots, n\}$ we have

$$\begin{array}{rl}
(\alpha \cdot L_1 + \beta \cdot L_2)(x_1, \ldots, x_{i-1}, x + y, x_{i+1}, \ldots, x_n) & = \\
\alpha \cdot L_1(x_1, \ldots, x_{i-1}, x + y, x_{i+1}, \ldots, x_n) + \beta \cdot L_2(x_1, \ldots, x_{i-1}, x + y, x_{i+1}, \ldots, x_n) & = \\
\underbrace{\alpha \cdot L_1(x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_n)}_{1} \quad + \quad \underbrace{\alpha \cdot L_1(x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n)}_{2} \quad + & \\
\underbrace{\beta \cdot L_2(x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_n)}_{1} + \underbrace{\beta \cdot L_2(x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n)}_{2} & = \\
\underbrace{(\alpha \cdot L_1 + \beta \cdot L_2)(x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_n)}_{1} + \underbrace{(\alpha \cdot L_1 + \beta \cdot L_2)(x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n)}_{2} &
\end{array}$$

and

$$\begin{array}{rl}
(\alpha \cdot L_1 + \beta \cdot L_2)(x_1, \ldots, x_{i-1}, \gamma \cdot x, x_{i+1}, \ldots, x_n) & = \\
\alpha \cdot L_1(x_1, \ldots, x_{i-1}, \gamma \cdot x, x_{i+1}, \ldots, x_n) + \beta \cdot L_2(x_1, \ldots, x_{i-1}, \gamma \cdot x, x_{i+1}, \ldots, x_n) & = \\
\alpha \cdot \gamma \cdot L_1(x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_n) + \beta \cdot \gamma \cdot L_2(x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_n) & = \\
\gamma \cdot (\alpha \cdot L_1(x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_n) + \beta \cdot L_2(x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_n)) & = \\
\gamma \cdot (\alpha \cdot L_1 + \beta \cdot L_2)(x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_n) &
\end{array}$$

proving that

$$\alpha \cdot L_1 + \beta \cdot L_2 \in \operatorname{Hom}(X_1, \ldots, X_n; Y)$$ □

Theorem 11.228. *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 $i \in \{1, \ldots, n\}$, $x \in \prod_{j \in \{1, \ldots, n\} \setminus \{i\}} X_i$, $\alpha \in F$ and $y, z \in X_i$ then

$$\begin{array}{rl}
(S \circ L)(x_1, \ldots, x_{i-1}, y + z, x_{i+1}, \ldots, x_n) & = \\
S(L(x_1, \ldots, x_{i-1}, y + z, x_{i+1}, \ldots, x_n)) & = \\
S(L(x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n) + L(x_1, \ldots, x_{i-1}, z, x_{i+1}, \ldots, x_n)) & = \\
(S \circ L)(x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n) + (S \circ L)(x_1, \ldots, x_{i-1}, z, x_{i+1}, \ldots, x_n) & =
\end{array}$$

and

$$\begin{array}{rcl}
(S \circ L)(x_1, \ldots, x_{i-1}, \alpha \cdot, x_{i+1}, \ldots, x_n) & = & S(L(x_1, \ldots, x_{i-1}, \alpha \cdot y, x_{i+1}, \ldots, x_n)) \\
& = & S(\alpha \cdot L(x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n)) \\
& = & \alpha \cdot S(L(x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n)) \\
& = & (S \circ L)(x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n)
\end{array}$$

□

Theorem 11.229. *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, $\{\alpha_i\}_{i \in \{1, \ldots, n\}} \subseteq F$ 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

$$S=\left\{k\in\mathbb{N}|\text{If } k\leqslant n \text{ and } L\in\mathrm{Hom}(X_1,\ldots,X_k;Y)\text{ then } L(\alpha_1\cdot x_1,\ldots,\alpha_k\cdot x_k)=\left(\prod_{i=1}^k\alpha_i\right)\cdot L(x_1,\ldots,x_k)\right\}$$

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 $L\in\mathrm{Hom}(X_1,\ldots,X_{k+1};Y)$ then by [theorem: 11.223] $L_{\{\ldots x_{k+1}\}}\in\mathrm{Hom}(X_1,\ldots,X_k;Y)$ so that as $k\in S$ we have

$$L_{\{\ldots x_k\}}(\alpha_1\cdot x_1,\ldots,\alpha_k\cdot x_k)=\left(\prod_{i=1}^k\alpha_i\right)\cdot L_{\{\ldots x_k\}}(x_1,\ldots,x_k) \tag{11.114}$$

Then we have

$$\begin{aligned}L(\alpha_1\cdot x_1,\ldots,\alpha_{k+1}\cdot x_{k+1}) &= \alpha_{k+1}\cdot L(\alpha_1\cdot x_1,\ldots,\alpha_{k-1}\cdot x_{k-1},x_{k+1})\\ &= \alpha_{k+1}\cdot L_{\{\ldots,x_{k+1}\}}(\alpha_1\cdot x_1,\ldots,a_k\cdot x_k)\\ &\underset{[\text{theorem: }11.114]}{=} \alpha_{k+1}\cdot\left(\prod_{i=1}^k\alpha_i\right)\cdot L_{\{\ldots x_k\}}(x_1,\ldots,x_k)\\ &= \left(\prod_{i=1}^{k+1}\alpha_i\right)\cdot L_{\{\ldots x_k\}}(x_1,\ldots,x_k)\\ &= \left(\prod_{i=1}^{k+1}\alpha_i\right)\cdot L(x_1,\ldots,x_{k+1})\end{aligned}$$

proving that $k+1\in S$.

Mathematical induction proves then that $S=\mathbb{N}$, so as $n\in\{1,\ldots,n\}$ 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 11.230. *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: 11.149] a sub-space Z_i of X_i such that $X_i=Y_i\oplus Z_i$. Using [theorem: 11.150] 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 $(x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)\in\prod_{j\in\{1,\ldots n\}\setminus\{i\}}X_i$ then we have:

$$\begin{aligned}&K(x_1,\ldots,x_{i-1},u+\alpha\cdot v,x_{i+1},\ldots,x_n) =\\ &L((\pi_{Y_1}(x_1),\ldots,\pi_{Y_{i-1}}(x_{i-1}),\pi_{Y_i}(u+\alpha\cdot v),\pi_{Y_{i+1}}(x_{i+1}),\ldots,\pi_{Y_n}(x_n))) =\\ &L((\pi_{Y_1}(x_1),\ldots,\pi_{Y_{i-1}}(x_{i-1}),\pi_{Y_i}(u)+\alpha\cdot\pi_{Y_i}(v),\pi_{Y_{i+1}}(x_{i+1}),\ldots,\pi_{Y_n}(x_n))) =\\ &L((\pi_{Y_1}(x_1),\ldots,\pi_{Y_{i-1}}(x_{i-1}),\pi_{Y_i}(u),\pi_{Y_{i+1}}(x_{i+1}),\ldots,\pi_{Y_n}(x_n)))+\alpha\cdot L((\pi_{Y_1}(x_1),\ldots,\pi_{Y_{i-1}}(x_{i-1}),\\ &\pi_{Y_i}(v),\pi_{Y_{i+1}}(x_{i+1}),\ldots,\pi_{Y_n}(x_n))) =\\ &K(x_1,\ldots,x_{i-1},\alpha,x_{i+1},\ldots,x_n)+\alpha\cdot K(x_1,\ldots,x_{i-1},v,x_{i+1},\ldots,x_n)\end{aligned}$$

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

$$\pi_i(x_i) = (\pi_i)_{|\mathrm{Y}_i} = \mathrm{Id}_{Y_i}(x_i) = x_i$$

hence

$$K(x_1, \ldots, x_n) = L(\pi_{\mathrm{Y}_1}(x_1), \ldots, \pi_{\mathrm{Yn}}(x_n)) = L(x_1, \ldots, x_n)$$

proving that

$$K_{|\prod_{i \in \{1, \ldots, n\}} Y_i} = L \qquad \square$$

Theorem 11.231. *Let $n \in \mathbb{N}$, X, Y vector spaces over the field F, $\{e_i\}_{i \in \{1, \ldots, \}} \subseteq X$ a distinct ordered family such that $\{e_i | i \in \{1, ..., n\}\}$ is a basis for X and $L \in \mathrm{Hom}(X^n; Y)$ then $\forall x \in X$ we have*

$$L(x_1, \ldots, x_n) = \sum_{k \in \{1, \ldots, n\}^n} \left(\prod_{j \in \{1, \ldots, n\}} \alpha_j^{k_j} \right) \cdot L(e_{k_1}, \ldots, e_{k_n})$$

where $\forall i \in \{1, ..., n\}$ $\{\alpha_j^i\}_{j \in \{1, \ldots, n\}}$ is the ***unique family*** *satisfying $x_i = \sum_{j \in \{1, \ldots, n\}} \alpha_i^j \cdot e_j$ [which exists by [theorem: 11.123]*

Proof. We prove this by induction on n, so let

$$S_n = \left\{ m \in \mathbb{N} | \text{If } m \leqslant n \text{ then } L(x_1, \ldots, x_n) = \sum_{k \in \{1, \ldots, n\}^m} \left(\prod_{j \in \{1, \ldots, m\}} \alpha_{k_j}^j \right) \cdot L(e_{k_1}, \ldots, e_{k_m}, x_{m+1}, \ldots, x_n) \right\}$$

where

$(e_{\rho_1}, \ldots, e_{\rho_m}, x_{m+1}, \ldots, x_n)$ 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

$\boldsymbol{1 \in S_n}$**.** If $m = 1$ then $\{1, \ldots, n\}^1$

and for $(x_1) \in X^1$

$$\begin{aligned} L(x_1, \ldots, x_n) &= L\left(\sum_{i \in \{1, \ldots, n\}} \alpha_1^i \cdot e_i, x_2, \ldots, x_n \right) \\ &\underset{[\text{theorem: } 11.225]}{=} \sum_{i \in \{1, \ldots, n\}} L(\alpha_1^i \cdot e_i, x_2, \ldots, x_n) \\ &\underset{[\text{theorem: } 11.225]}{=} \sum_{i \in \{1, \ldots, n\}} \alpha_1^i \cdot L(e_i, x_2, \ldots, x_n) \\ &= \sum_{i \in \{1, \ldots, n\}} \left(\prod_{j \in \{1\}} \alpha_j^i \right) \cdot L(e_i, x_2, \ldots, x_n) \end{aligned} \tag{11.115}$$

Define $\beta: \{1, \ldots, n\}^1 \rightarrow \{1, \ldots, n\}$ by $\beta(\rho) = \rho_1$

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

$$\beta: \{1, \ldots, n\}^1 \rightarrow \{1, \ldots, n\} \text{ is a bijection}$$

then we have

$$\begin{aligned}
\sum_{\rho\in\{1,\ldots,n\}^1}\left(\prod_{j\in\{1\}}\alpha_j^{\rho_j}\right)\cdot L(e_{\rho_j},x_2,\ldots,x_n) &= \\
\sum_{\rho\in\{1,\ldots,n\}^1}\left(\prod_{j\in\{1\}}\alpha_j^{\rho_1}\right)\cdot L(e_{\rho_1},x_2,\ldots,x_n) &= \\
\sum_{i\in\{1,\ldots,n\}^1}\left(\prod_{j\in\{1\}}\alpha_j^{i}\right)\cdot L(e_{\beta(\rho)},x_2,\ldots,x_n) &\underset{\text{[theorem: 11.34]}}{=} \\
\sum_{i\in\{1,\ldots,n\}}\left(\prod_{j\in\{1\}}\alpha_j^{i}\right)\cdot L(e_i,x_2,\ldots,x_n) &\underset{\text{[eq: 11.115}}{=} \\
L(x_1,\ldots,x_n) &
\end{aligned}$$

proving that $1\in S_n$.

$\boldsymbol{m\in S_n\Rightarrow m+1\in S_n}$**.** If $m+1\leqslant n$ then we have

$$\begin{aligned}
L(x_1,\ldots,x_n) &\underset{m\in S_n}{=} \\
\sum_{\rho\in\{1,\ldots,n\}^m}\left(\prod_{j\in\{1,\ldots,m\}}\alpha_j^{\rho_j}\right)\cdot L(e_{\rho_1},\ldots,e_{\rho_m},x_{m+1},\ldots,x_n) &= \\
\sum_{\rho\in\{1,\ldots,n\}^m}\left(\prod_{j\in\{1,\ldots,m\}}\alpha_j^{\rho_j}\right)\cdot L\left(e_{\rho_1},\ldots,\sum_{i\in\{1,\ldots,n\}}\alpha_{m+1}^i\cdot x_j,x_{m+1},\ldots,x_n\right) &= \\
\sum_{\rho\in\{1,\ldots,n\}^m}\left(\prod_{j\in\{1,\ldots,m\}}\alpha_j^{\rho_j}\right)\cdot\sum_{i\in\{1,\ldots,n\}}\alpha_{m+1}^i\cdot L(e_{\rho_1},\ldots,e_{\rho_m},e_i,,\ldots x_n) &= \\
\sum_{\rho\in\{1,\ldots,n\}^m}\left(\sum_{i\in\{1,\ldots,n\}}\left(\left(\prod_{j\in\{1,\ldots,m\}}\alpha_j^{\rho_j}\right)\cdot\alpha_{m+1}^i\right)\cdot L(e_{\rho_1},\ldots,e_{\rho_m},e_i,,\ldots x_n)\right) &=
\end{aligned}$$

so that

$$L(x_1,\ldots,x_n) = \sum_{\rho\in\{1,\ldots,n\}^m}\left(\sum_{i\in\{1,\ldots,n\}}A_{\rho,i}\right) \tag{11.116}$$

where

$$A_{\rho,i}=\left(\left(\prod_{j\in\{1,\ldots,m\}}\alpha_j^{\rho_j}\right)\cdot\alpha_{m+1}^i\right)\cdot L(e_{\rho_1},\ldots,e_{\rho_m},e_i,,\ldots x_n) \tag{11.117}$$

Given $\rho\in\{1,\ldots,n\}^m$ define

$$I_\rho=\{\rho\}\times\{1,\ldots,n\}\text{ and }\pi_2\colon I_\rho\rightarrow\{1,\ldots,n\}\text{ by }\pi_2(\rho,k)=k \tag{11.118}$$

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 $k\in\{1,...,n\}$ then for $(\rho,k)\in I_\rho$ we have $\pi_2(\rho,k)=k$ proving surjectivity.

hence:

$$\forall\rho\in\{1,\ldots,n\}\text{ we have }\pi_2\colon I_\rho\rightarrow\{1,\ldots,n\}\text{ is a bijection}$$

so that

$$\begin{aligned}
\sum_{k\in I_\rho}A_{k_1,k_2} &\underset{k\in I_\rho\Rightarrow k_1=\rho}{=} \sum_{k\in I_\rho}A_{\rho,\pi_2(k)} \\
&\underset{\text{[theorem: 11.34]}}{=} \sum_{i\in\{1,\ldots,n\}}A_{\rho,i}
\end{aligned}$$

which by substituting this in {eq: $m\in S_n\Rightarrow m+1\in S_n$] gives

$$L(x_1,\ldots,x_n)=\sum_{\rho\in\{1,\ldots,n\}^m}\left(\sum_{k\in I_\rho}A_{k_1,k_2}\right)\tag{11.119}$$

If $\rho,\sigma\in\{1,\ldots,n\}^m$ with $\rho\neq\sigma$ then if $k\in I_\rho\bigcap I_\sigma$ we have $k=(\rho,i),(\sigma,j)$ 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

$$\sum_{\rho\in I}A_{k_1,k_2}\underset{\text{[theorem: 11.42]}}{=}\sum_{\rho\in\{1,\ldots,n\}^m}\left(\sum_{k\in I_\rho}A_{k_1,k_2}\right)$$

which combined with [eq: 11.119] gives

$$L(x_1,\ldots,x_n)=\sum_{\rho\in I}A_{k_1,k_2}\tag{11.120}$$

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 we have $\forall i\in\{1,...,m+1\}$ that $k_{|\{1,\ldots,m\}}=l_{|\{1,\ldots,m\}}\Rightarrow\forall i\in\{1,...,m\}\ k(i)=l(i)$ 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 then $l\in\{1,\ldots,n\}^{m+1}$ 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 $\eta(l)=(l_{|\{1,\ldots,m\}},\lambda(m+1))=(\rho,k)=z$.

proving that

$$\eta:\{1,\ldots,n\}^{m+1}\rightarrow I\text{ is a bijection}$$

Using [theorem: 11.34] 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: 11.120] proves

$$L(x_1,\ldots,x_n)=\sum_{k\in\{1,\ldots,n\}^{m+1}}A_{\eta(k)_1,\eta(k)_2}\tag{11.121}$$

Now for $k\in\{1,\ldots,n\}^{m+1}$ we have:

$$\begin{aligned}
A_{\eta(k)_1,\eta(k)_2}&=\\
A_{k_{|\{1,\ldots,m\}},k_{m+1}}&=\\
\left(\left(\prod_{j\in\{1,\ldots,m\}}\alpha_j^{(k_{|\{1,\ldots,m\}})_j}\right)\cdot\alpha_{m+1}^{k_{m+1}}\right)\cdot L(e_{(k_{|\{1,\ldots,m\}})_1},\ldots,e_{(k_{|\{1,\ldots,m\}})_m},e_{k_{m+1}},,\ldots x_n)&=\\
\left(\left(\prod_{j\in\{1,\ldots,m\}}\alpha_j^{k_j}\right)\cdot\alpha_{m+1}^{k_{m+1}}\right)\cdot L(e_{k_1},\ldots,e_{k_m},e_{k_{m+1}},,\ldots x_n)&=\\
\left(\prod_{j\in\{1,\ldots,m+1\}}\alpha_j^{k_j}\right)\cdot L(e_{k_1},\ldots,e_{k_{m+1}},x_{m+2}\ldots,x_n)&
\end{aligned}$$

which combined with [eq: 11.121] 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 11.232. *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 11.233. *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 11.234. *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 11.235. *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

$$(-1) \cdot x + x = x + (-1) \cdot x = (-1) \cdot 1.x = (1 + (-1)) \cdot x = 0 \cdot x = 0$$

so that $-x = (-1) \cdot x$ □

Next we define the power in a field by recursion.

Definition 11.236. *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 11.237. *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 $n+1\in\mathcal{S}_m$ □

Proposition 11.238. *Let F be a field then for $n\in\mathbb{N}_0$ we have*

$$\begin{array}{rcl} 1^n & = & 1\\ (-1)^n & = & -1 \textit{ or } 1\\ (-1)^n & = & \left\{\begin{array}{l} -1 \textit{ if } n \textit{ is odd}\\ 1 \textit{ if } m \textit{ is even} \end{array}\right. \end{array}$$

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 $(-1)^{n+1}=(-1)\cdot(-1)^n=(-1)\cdot(-1) \underset{[\text{proposition: } 11.233]}{=} 1$

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 $(-1)^n=(-1)^{m+m} \underset{[\text{theorem: } 11.237]}{=} (-1)^m\cdot(-1)^m=1\cdot 1=1$

$(-1)^m=1$. Then

$$(-1)^n=(-1)^{m+m} \underset{[\text{theorem: } 11.237]}{=} (-1)^m\cdot(-1)^m=(-1)\cdot(-1) \underset{[\text{theorem: } 11.233]}{=} 1$$

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

$$0^n=0^{(n-1)+1}=0\cdot 0^{n-1}=0$$ □

All the above work is done so that we can define a mapping of $\{1,-1\}$ to F which we will use in the definition of skew symmetric multilinear mappings.

Definition 11.239. *Let $\langle F,+,\cdot\rangle$ be a field then we define*

$$\odot\colon\{1,-1\}\times F\rightarrow F \textit{ by } z\odot f=\left\{\begin{array}{l} f \textit{ if } z{=}1\\ -f \textit{ if } x=-1 \end{array}\right.$$

Theorem 11.240. *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 $(-1)^0\odot f=1\odot f\underset{\text{def}}{=}f=1\cdot f\underset{\text{[definition: 11.236}}{=}(-1)^0\cdot f$ 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

$$\begin{aligned}(-1)^{n+1}\odot f &= (-1)\odot f\\ &\underset{\text{def}}{=} -f\\ &\underset{\text{[theorem: 11.234]}}{=} (-1)\cdot f\\ &\underset{\text{[theorem: 11.238]}}{=} (-1)^{n+1}\cdot f\end{aligned}$$

proving that $n+1\in S$ in this case.

n is odd. Then $n+1$ is even

$$\begin{aligned}(-1)^{n+1}\odot f &= 1\odot f\\ &\underset{\text{def}}{=} f\\ &= 1\cdot f\\ &\underset{\text{[theorem: 11.238]}}{=} (-1)^{n+1}\cdot f\end{aligned}$$

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 $\operatorname{sign}(\sigma)\in\{-1,1\}\subseteq\mathbb{Z}$ so that if $f\in V$ $\operatorname{sign}(\sigma)\cdot f$ is actually $\operatorname{sign}(\alpha)\odot f$.

Definition 11.241. *Let X,Y be vector spaces over a field F, $m\in\mathbb{N}$ then a n-linear mapping $L\in\operatorname{Hom}(X^n;Y)$ [see definition: 11.221] is **symmetric** if $\forall\sigma\in P_n$ we have*

$$\sigma L=L$$

[see definition: 11.196].

More useful than symmetric n-linear mappings are skew-linear mappings.

Definition 11.242. *Let X,Y be vector spaces over a field F, $n\in\mathbb{N}$ then a n-linear mapping $L\in\operatorname{Hom}(X^n;Y)$ [see definition: 11.221] is **skew-symmetric** if $\forall\sigma\in P_n$ we have*

$$\sigma L=\operatorname{sign}(\sigma)\cdot L$$

[see definition: 11.196].

A trivial example of a mapping that is at the same time symmetric and skew-symmetric is the following.

Example 11.243. Let X,Y be vector space over a field F then if $L\in\operatorname{Hom}(X^1;Y)$ L is symmetric and skew-symmetric.

Proof. If $\sigma\in P_1$ then $\sigma=\operatorname{Id}_{\{1\}}$ so that $\operatorname{sign}(\sigma)=1$ and $\sigma L(x)=L(x\circ\sigma)=L$ proving that

$$\sigma L=L=1\cdot L=\operatorname{sign}(\sigma)\cdot L$$

□

If $L\in\operatorname{Hom}(X^n;Y)$ is symmetric then $\sigma L=L\in\operatorname{Hom}(X^n;Y)$ and if L is skew-symmetric we have $\sigma L=\operatorname{sign}(\sigma)\cdot L\in\operatorname{Hom}(X;Y)$. Actually this is a general result valid for every n-linear mapping.

Theorem 11.244. *Let X,Y be vector spaces over a field F, $n\in\mathbb{N}$ and $L\in\operatorname{Hom}(X^n;Y)$ a n-linear mapping then*

$$\forall\sigma\in P_n \text{ we have } \sigma L\in\operatorname{Hom}(X^n;Y)$$

Proof. Let $i\in\{1,\ldots,\}$, $\alpha\in F$ and $u,v\in X_i$. Then if

$$\begin{aligned}
z &= (x_1,\ldots,x_{i-1},u+\alpha\cdot v,x_{i+1},\ldots,x_n)\in X^n\\
r &= (x_1,\ldots,x_{i-1},u,x_{i+1},\ldots,x_n)\in X^n\\
s &= (x_1,\ldots,x_{i-1},v,x_{i+1},\ldots,x_n)\in X^n
\end{aligned}$$

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 $\forall j\in\{1,\ldots,n\}\setminus\{k\}$ we have:

$$\begin{aligned}
(z\circ\sigma)_j &= x(\sigma(j))=(r\circ\sigma)_j & (11.122)\\
(z\circ\sigma)_j &= x(\sigma(i))=(s\circ\sigma)_j & (11.123)
\end{aligned}$$

Then

$$\begin{aligned}
(\sigma L)(z) &=\\
L(z\circ\sigma) &\underset{[\text{theorem: }11.217]}{=}\\
L((z\circ\sigma)_1,\ldots,(z\circ\sigma)_{\sigma^{-1}(i)-1},u+\alpha\cdot v,(z\circ\sigma)_{\sigma^{-1}(i)+1},\ldots,(z\circ\sigma)_n) &=\\
L((z\circ\sigma)_1,\ldots,(z\circ\sigma)_{\sigma^{-1}(i)-1},u,(z\circ\sigma)_{\sigma^{-1}(i)+1},\ldots,(z\circ\sigma)_n)+\alpha\cdot L((z\circ\sigma)_1,\ldots,&\\
(z\circ\sigma)_{\sigma^{-1}(i)-1},v,(z\circ\sigma)_{\sigma^{-1}(i)+1},\ldots,(z\circ\sigma)_n) &\underset{[\text{eqs: }11.122,11.123]}{=}\\
L((r\circ\sigma)_1,\ldots,(r\circ\sigma)_{\sigma^{-1}(i)-1},u,(r\circ\sigma)_{\sigma^{-1}(i)+1},\ldots,(r\circ\sigma)_n)+\alpha\cdot L((s\circ\sigma)_1,\ldots,&\\
(s\circ\sigma)_{\sigma^{-1}(i)-1},v,(s\circ\sigma)_{\sigma^{-1}(i)+1},\ldots,(s\circ\sigma)_n) &\underset{[\text{theorem: }11.217]}{=}\\
L(r\circ\sigma)+\alpha\cdot L(s\circ\sigma)&
\end{aligned}$$

proving that

$$\begin{aligned}
(\sigma L)((x_1,\ldots,x_{i-1},u+\alpha\cdot v,x_{i+1},\ldots,x_n)) &=\\
(\sigma L)((x_1,\ldots,x_{i-1},u,x_{i+1},\ldots,x_n))+\alpha\cdot(\sigma L)((x_1,\ldots,x_{i-1},v,x_{i+1},\ldots,x_n))&
\end{aligned}$$

□

The next theorem shows we can construct a new skew-symmetric a n-linear mapping. First we need a little lemma:

Lemma 11.245. *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

$$\begin{aligned}
\rho\circ\sigma=\rho\circ\sigma' &\Rightarrow \rho^{-1}\circ(\rho\circ\sigma)=\rho^{-1}\circ(\rho\circ\sigma')\\
&\Rightarrow (\rho^{-1}\circ\rho)\circ\sigma=(\rho^{-1}\circ\rho)\circ\sigma'\\
&\Rightarrow \operatorname{Id}_{\{1,\ldots,n\}}\circ\sigma-\operatorname{Id}_{\{1,\ldots,n\}}\circ\sigma'\\
&\Rightarrow \sigma=\sigma'
\end{aligned}$$

proving injectivity.

surjectivity. Let $\sigma \in P_n$ then $\rho^{-1} \circ \sigma \in P_n$ [see theorem: 11.187] 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 11.246. *LetX, Y be vector spaces over a field F then given $n \in \mathbb{N}$ and a n-linear mapping $L \in \mathrm{Hom}(X^n; Y)$ then*

$$\sum_{\sigma \in P_n} \mathrm{sign}(\sigma) \cdot \sigma L$$

is a skew-symmetric mapping:

***Note* 11.247.** *The sum is well defined as P_n is finite, $\forall \sigma \in P_n\ \sigma L \in \mathrm{Hom}(X^n; Y) \Rightarrow \mathrm{sign}(\sigma) \cdot \sigma L \in \mathrm{Hom}(X^n; Y) \subseteq Y^{x^n}$ and Y^{X^n} is a vector space [see theorems: 11.186, 11.244 and 11.227].*

Proof. Using [theorem: 11.244] we have $\forall \sigma \in P_n$ that $\mathrm{sign}(\sigma) \cdot \sigma L \in \mathrm{Hom}(X^n; Y)$ so that by [theorem: 11.70]

$$\sum_{\sigma \in P_n} \mathrm{sign}(\sigma) \cdot \sigma L \in \mathrm{Hom}(X^n; Y) \text{ or } \sum_{\sigma \in P_n} \mathrm{sign}(\sigma) \cdot \sigma L \text{ is } n\text{-linear}$$

Next we have to prove that $\sum_{\sigma \in P_n} \mathrm{sign}(\sigma) \cdot \sigma L$ is skew-symmetric. So let $\rho \in P_n$ then we have for $x \in X^n$ that

$$\begin{aligned}
\left(\rho\left(\sum_{\sigma \in P_n} \mathrm{sign}(\sigma) \cdot \sigma L\right)\right)(x) &= \left(\sum_{\sigma \in P_n} \mathrm{sign}(\sigma) \cdot \sigma L\right)(x \circ \rho) \\
&= \sum_{\sigma \in P_n} (\mathrm{sign}(\sigma) \cdot \sigma L)(x \circ \rho) \\
&= \sum_{\sigma \in P_n} \mathrm{sign}(\sigma) \cdot (\sigma L)(x \circ \rho) \\
&= \sum_{\sigma \in P_n} \mathrm{sign}(\sigma) \cdot L((x \circ \rho) \circ \sigma) \\
&= \sum_{\sigma \in P_n} \mathrm{sign}(\sigma) \cdot L(x \circ (\rho \circ \sigma)) \\
&= \sum_{\sigma \in P_n} \mathrm{sign}(\sigma) \cdot ((\rho \circ \sigma)L)(x) \\
&\underset{\mathrm{sign}(\rho) \cdot \mathrm{sign}(\rho) = 1}{=} \sum_{\sigma \in P_n} (\mathrm{sign}(\rho) \cdot \mathrm{sign}(\rho)) \cdot \mathrm{sign}(\sigma) \cdot ((\rho \circ \sigma)L)(x) \\
&= \sum_{\sigma \in P_n} \mathrm{sign}(\rho) \cdot ((\mathrm{sign}(\rho) \cdot \mathrm{sign}(\sigma)) \cdot ((\rho \circ \sigma)L)(x)) \\
&= \mathrm{sign}(\rho) \cdot \sum_{\sigma \in P_n} (\mathrm{sign}(\rho) \cdot \mathrm{sign}(\sigma)) \cdot ((\rho \circ \sigma)L)(x) \\
&\underset{[\text{theorem: } 11.211]}{=} \mathrm{sign}(\rho) \cdot \sum_{\sigma \in P_n} \mathrm{sign}(\rho \circ \sigma) \cdot ((\rho \circ \sigma)L)(x) \\
&\underset{[\text{lemma: } 11.245]}{=} \mathrm{sign}(\rho) \cdot \sum_{\sigma \in P_n} \mathrm{sign}(T_\rho(\sigma)) \cdot (T_\rho(\sigma)L)(x) \\
&\underset{[\text{theorem: } 11.34]}{=} \mathrm{sign}(\rho) \cdot \sum_{\sigma \in P_n} \mathrm{sign}(\sigma) \cdot (\sigma L)(x)
\end{aligned}$$

proving that

$$\rho\left(\sum_{\sigma \in P_n} \mathrm{sign}(\sigma) \cdot \sigma L\right) = \mathrm{sign}(\rho) \cdot \sum_{\sigma \in P_n} \mathrm{sign}(\sigma) \cdot \sigma L$$

□

Theorem 11.248. *Let X, Y be vector spaces over a field F with characteristic zero [see definition: 10.42], $n \in \mathbb{N}$ and $L \in \mathrm{Hom}(X^n; Y)$ then the following are equivalent:*

1. *L is skew-symmetric*

2. *For every $x\in X^n$ such that $\exists i,j\in\{1,\ldots,n\}$ with $i\neq j$ and $x_i=x_j$ we have $L(x)=0$*

3. *For every $x\in X^n$ such that $\{x_i\}_{i\in\{1,\ldots,n\}}$ 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

$$\begin{aligned}
(x\circ(i\leftrightarrow j_n))(k) &= x\big((i\underset{n}{\leftrightarrow}j)(k)\big)\\
&= \begin{cases} x(i) \text{ if } k=j\\ x(j) \text{ if } k=i\\ x(k) \text{ if } k\in\{1,\ldots,n\}\setminus\{i,j\}\end{cases}\\
&\underset{x(i)=x(j)}{=} \begin{cases} x(k) \text{ if } k=j\\ x(k) \text{ if } k=i\\ x(k) \text{ if } k\in\{1,\ldots,n\}\setminus\{i,j\}\end{cases}\\
&= x(k)
\end{aligned}$$

proving that

$$x\circ\big(i\underset{n}{\leftrightarrow}j\big)=x$$

Then

$$\begin{aligned}
L(x) &= L\big(x\circ\big(i\underset{n}{\leftrightarrow}j\big)\big)\\
&\underset{L\text{ is skew-symmetric}}{=} \operatorname{sign}\big(\big(i\underset{n}{\leftrightarrow}j\big)\big)\cdot L(x)\\
&\underset{[\text{theorem: }11.211]}{=} (-1)\cdot L(x)\\
&= -L(x)
\end{aligned}$$

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 ⇒ 1. Assume that (2) hold then we prove the following

$$\forall x\in X^n \text{ we have } \forall i,j\in\{1,\ldots,n\} \text{ with } i\neq j \text{ that } L\big(x\circ\big(i\underset{n}{\leftrightarrow}j\big)\big)=-L(x) \tag{11.124}$$

Proof. Let $x\in X^n$, $i,j\in\{1,\ldots,n\}$ with $i\neq j$. Then

$$\begin{aligned}
0 &=\\
L(x_1,\ldots,x_{i-1},x_i+x_j,x_{i+1},\ldots,x_{j-1},x_i+x_j,x_{j+1},\ldots,x_n) &=\\
L(x_1,\ldots,x_{i-1},x_i+x_j,x_{i+1},\ldots,x_{j-1},x_i,x_{j+1},\ldots,x_n)+L(x_1,\ldots,x_{i-1},x_i+x_j,x_{i+1},\ldots,\\
x_{j-1},x_j,x_{j+1},\ldots,x_n) &=\\
L(x_1,\ldots,x_{i-1},x_i,x_{i+1},\ldots,x_{j-1},x_i,x_{j+1},\ldots,x_n)+L(x_1,\ldots,x_{i-1},x_j,x_{i+1},\ldots,x_{j-1},\\
x_i,x_{j+1},\ldots,x_n)+L(x_1,\ldots,x_{i-1},x_i,x_{i+1},\ldots,x_{j-1},x_j,x_{j+1},\ldots,x_n)+L(x_1,\ldots,x_{i-1},\\
x_j,x_{i+1},\ldots,x_{j-1},x_j,x_{j+1},\ldots,x_n) &=\\
0+L(x_1,\ldots,x_{i-1},x_j,x_{i+1},\ldots,x_{j-1},x_i,x_{j+1},\ldots,x_n)+L(x_1,\ldots,x_{i-1},x_i,x_{i+1},\ldots,x_{j-1},\\
x_j,x_{j+1},\ldots,x_n)+0 &=\\
L(x_1,\ldots,x_{i-1},x_j,x_{i+1},\ldots,x_{j-1},x_i,x_{j+1},\ldots,x_n)+L(x_1,\ldots,x_{i-1},x_i,x_{i+1},\ldots,x_{j-1},\\
x_j,x_{j+1},\ldots,x_n) &=\\
L\big(x\circ\big(i\underset{n}{\leftrightarrow}j\big)\big)+L(x)
\end{aligned}$$

proving that

$$L\big(x\circ\big(i\underset{n}{\leftrightarrow}j\big)\big)=-L(x)$$ □

Next we prove that

If $m \in \mathbb{N}$ $\{(i_k \leftrightarrow j_k)\}_{k \in \{,\ldots,m\}}$ is such that $\forall k \in \{1, \ldots, m\}$ $i_k \neq j_k$ then $\left(\left(i_1 \underset{n}{\leftrightarrow} j_1\right) \circ \cdots \circ \left(i_m \underset{n}{\leftrightarrow} j_m\right)\right) L = (-1)^m L$ (11.125)

Proof. We proceed by induction, so let

$S = \left\{m \in \mathbb{N} | \forall \{(i_k \leftrightarrow j_k)\}_{k \in \{,\ldots,m\}}\right.$ [family of **strict** transpositions] $\left.\left(\left(i_1 \underset{n}{\leftrightarrow} j_1\right) \circ \cdots \circ \left(i_m \underset{n}{\leftrightarrow} j_m\right)\right) L = (-1)^m L\right\}$ then we have:

$\boldsymbol{1 \in S}$.

$$\begin{array}{rcl} \left(\left(i_1 \underset{n}{\leftrightarrow} j_1\right) \circ \cdots \circ \left(i_1 \underset{n}{\leftrightarrow} j_1\right)\right) L & = & \left(i_1 \underset{n}{\leftrightarrow} j_1\right) L \\ & \underset{\text{[eq: 11.124]}}{=} & -L \\ & = & (-1)^1 \cdot L \end{array}$$

proving that $1 \in S$.

$\boldsymbol{m \in S \Rightarrow m + 1 \in S}$. We have

$$\begin{array}{rc} \left(\left(i_1 \underset{n}{\leftrightarrow} j_1\right) \circ \cdots \circ \left(i_{m+1} \underset{n}{\leftrightarrow} j_{m+1}\right)\right) L & = \\ \left(\left(\left(i_1 \underset{n}{\leftrightarrow} j_1\right) \circ \cdots \circ, \left(i_m \underset{n}{\leftrightarrow} j_m\right)\right) \circ (i_{m+1} \leftrightarrow)\right) L & \underset{\text{[theorem: 11.199]}}{=} \\ (i_{m+1} \leftrightarrow j_m)\left(\left(\left(i_1 \underset{n}{\leftrightarrow} j_1\right) \circ \cdots \circ, \left(i_m \underset{n}{\leftrightarrow} j_m\right)\right) L\right) & \underset{m \in S}{=} \\ (i_{m+1} \leftrightarrow j_m)((-1)^m L) & \underset{\text{[theorem: 11.197]}}{=} \\ (-1)^m \cdot ((i_{m+1} \leftrightarrow j_m) L) & \underset{\text{[eq: 11.124]}}{=} \\ (-1)^m \cdot (-L) & = \\ (-1)^m \cdot ((-1) \cdot L) & = \\ ((-1)^m \cdot (-1)) \cdot L & = \\ (-1)^{m+1} \cdot L & \end{array}$$

proving that $m + 1 \in S$. □

Let $\sigma \in P_n$ then by [theorem: 11.210] there exists a $\left\{\left(i_k \underset{n}{\leftrightarrow} j_k\right)\right\}_{k \in \{1, \ldots, m\}}$ of **strict** transpositions such that

$$\sigma = \left(i_1 \underset{n}{\leftrightarrow}\right) \circ \cdots \circ (i_m \leftrightarrow j_m) \text{ and } \operatorname{sign}(\sigma) = (-1)^m$$

So

$$\sigma L = \left(\left(i_1 \underset{n}{\leftrightarrow}\right) \circ \cdots \circ (i_m \leftrightarrow j_m)\right) L \underset{\text{[eq: 11.125]}}{=} (-1)^m \cdot L = \operatorname{sign}(\sigma) \cdot L$$

which proves that L is skew-symmetric.

$\boldsymbol{2 \Rightarrow 3}$. If $\{x_i\}_{i \in \{1, \ldots, n\}}$ is linear dependent then by [theorem: 11.112] there exists a $k \in \{1, \ldots, n\}$ and a $\{\alpha_i\}_{i \in \{1, \ldots, n\}} \subseteq F$ such that

$$x_k = \sum_{i \in \{1, \ldots, n\} \setminus \{k\}} \alpha_i \cdot x_i$$

Hence we have that

$$\begin{array}{rcl} L(x) & = & L(x_1, \ldots, x_{k-1}, x_k, x_{k+1}, \ldots, x_n) \\ & = & L\left(x_1, \ldots, x_{k-1}, \displaystyle\sum_{i \in \{1, \ldots, n\} \setminus \{k\}} \alpha_i \cdot x_i, x_{k+1}, \ldots, x_n\right) \\ & \underset{\text{[theorems: 11.225, 11.167]}}{=} & \displaystyle\sum_{i \in \{1, \ldots, n\} \setminus \{k\}} L(x_1, \ldots, x_{k-1}, x_i, x_{k+1}, \ldots, x_n) \end{array} \quad (11.126)$$

If $i\in\{1,\ldots,n\}\setminus\{k\}\Rightarrow i\neq k$ then we have

$$(x_1,\ldots,x_{k-1},x_i,x_{k+1},\ldots,x_n)_k=x_i=(x_1,\ldots,x_{k-1},x_i,x_{k+1},\ldots,x_n)_i$$

so that $L(x_1,\ldots,x_{k-1},x_i,x_{k+1},\ldots,x_n)=0$. Combining this with [eq: 11.126] gives finally

$$L(x)=0$$

$3\Rightarrow 2$. 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: 11.108] $\{x_i\}_{i\in\{1,\ldots,n\}}$ is linear dependent, hence by (3) $L(x)=0$ □

Corollary 11.249. *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: 11.138] $X=\{0\}$. Hence if $L\in\mathrm{Hom}(X^0;Y)$ we have for $x\in X^0$ that $x=(0)$ so that $L(x)=L(0)\underset{\text{[theorem: 11.219]}}{=}0$ proving the corollary.

$1<n$. Let $x\in X^n$ then for $\{x_i\}_{i\in\{1,\ldots,n\}}$ we have either:

$\{x_i\}_{i\in\{1,\ldots,n\}}$ is disjoint. As $\dim(X)<n$ there exist a basis $B\Rightarrow\mathrm{span}(B)=V$ with $\mathrm{card}(B)<n$. Assume that $\{x_i|i\in\{1,...,n\}\}$ is linear independent and disjoint we have by [theorem: 11.116] that $n\leqslant\mathrm{card}(B)$ leading to the contradiction $n<n$. Hence we must have that $\{x_i|i\in\{1,\ldots,n\}\}$ is linear dependent. By [theorem: 11.106] it follows that $\{x_i\}_{i\in\{1,\ldots,n\}}$ is linear dependent. Applying [theorem: 11.248 (4)] gives then

$$L(x)=0$$

$\{x_i\}_{i\in\{1,\ldots,n\}}$ is not disjoint. Hence by [definition: 11.74] there exists a $i,j\in\{1,\ldots,n\}$ with $i\neq j$ such that $x_i=x_j$. Applying [theorem: 11.248 (2)] gives then

$$L(x)=0$$

So in all cases we have $\forall x\in X$ that $L(x)=0$. □

Theorem 11.250. *Let $n\in\mathbb{N}$, X,Y vector spaces over the field F, $\{e_i\}_{i\in\{1,\ldots,\}}\subseteq X$ a distinct ordered family such that $\{e_i|i\in\{1,\ldots,n\}\}$ is a basis for X and $L\in\mathrm{Hom}(X^n;Y)$ such that L is skew-symmetric then $\forall x\in X$ we have:*

$$L(x_1,\ldots,x_n)=\sum_{\sigma\in P_n}\left(\prod_{j\in\{1,\ldots,m\}}\alpha_j^{\sigma(j)}\right)\cdot(\sigma L)(e_1,\ldots,e_n)$$

where $\forall i\in\{1,...,n\}$ $\{\alpha_j^i\}_{j\in\{1,\ldots,n\}}$ is the ***unique family*** *satisfying $x_i=\sum_{j\in\{1,\ldots,n\}}\alpha_i^j\cdot e_j$ [which exists by [theorem: 11.123]*

Proof. As $P_n=\{\sigma\in\{1,\ldots,n\}^{\{1,\ldots,n\}}|\sigma:\{1,\ldots,n\}\rightarrow\{1,\ldots,n\}$ is a bijection$\}$ we have $P_n\subseteq\{1,\ldots,n\}^{\{1,\ldots,n\}}\underset{\text{[definition: 6.78]}}{=}\{1,\ldots,n\}^n$. Let $\sigma\in\{1,\ldots,n\}^n\setminus P_n$ then $\sigma:\{1,\ldots,n\}\rightarrow\{1,\ldots,n\}$ is not a bijection. We have now two cases for $\sigma(\{1,\ldots,n\})$ to consider:

$\sigma(\{1,\ldots,n\})=\{1,\ldots,n\}$. Then σ can not be injective [otherwise $\sigma:\{1,...,n\}\rightarrow\{1,...,n\}$ is a bijection].

$\sigma(\{1,\ldots,n\})\neq\{1,\ldots,n\}$. Then as $\sigma(\{1,...,n\})\subseteq\{1,...,n\}$ we have $\sigma(\{1,...,n\})\subset\{1,...,n\}$ so that by [theorem: 10.79]

$$\mathrm{card}(\sigma(\{1,\ldots,n\}))<\mathrm{card}(\{1,\ldots,n\})=n$$

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 must be not injective.

As in all cases 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: 11.248 (2)] we have $L(e_{\sigma(1)},\ldots,e_{\sigma(n)}) = 0$. To summarize

$$\text{If } \sigma \in \{1,\ldots,n\} \setminus P_n \text{ then } L(e_{\sigma(1)},\ldots,e_{\sigma(n)}) = 0 \tag{11.127}$$

Now if $x \in X$ we have

$$\begin{aligned}
L(x_1,\ldots,x_n) &\underset{\text{[theorem: 11.231]}}{=} \sum_{\sigma \in \{1,\ldots,n\}^n} \left(\prod_{j \in \{1,\ldots,n\}} \alpha_j^{\sigma_j} \right) \cdot L(e_{\sigma_i},\ldots,e_{\sigma_n}) \\
&= \sum_{\sigma \in \{1,\ldots,n\}^n} \left(\prod_{j \in \{1,\ldots,n\}} \alpha_j^{\sigma(j)} \right) \cdot L(e_{\sigma(1)},\ldots,e_{\sigma(n)}) \\
&= \sum_{\sigma \in \{1,\ldots,n\}^n \setminus P_n} \left(\prod_{j \in \{1,\ldots,n\}} \alpha_j^{\sigma(j)} \right) \cdot L(e_{\sigma(1)},\ldots,e_{\sigma(n)}) + \\
&\quad \sum_{\sigma \in P_n} \left(\prod_{j \in \{1,\ldots,n\}} \alpha_j^{\sigma(j)} \right) \cdot L(e_{\sigma(1)},\ldots,e_{\sigma(n)}) \\
&= \sum_{\sigma \in \{1,\ldots,n\}^n \setminus P_n} \left(\prod_{j \in \{1,\ldots,n\}} \alpha_j^{\sigma(j)} \right) \cdot 0 + \sum_{k \in P_n} \left(\prod_{j \in \{1,\ldots,n\}} \alpha_j^{\sigma(j)} \right) \cdot \\
&\quad L(e_{\sigma(1)},\ldots,e_{\sigma(n)}) \\
&= \sum_{\sigma \in P_n} \left(\prod_{j \in \{1,\ldots,n\}} \alpha_j^{\sigma(j)} \right) \cdot L(e_{\sigma(1)},\ldots,e_{\sigma(n)}) \\
&= \sum_{\sigma \in P_n} \left(\prod_{j \in \{1,\ldots,n\}} \alpha_j^{\sigma(j)} \right) \cdot (\sigma L)(e_1,\ldots,e_n)
\end{aligned}$$

□

Corollary 11.251. *Let $n \in \mathbb{N}$, X, Y vector spaces over the field F, $\{e_i\}_{i \in \{1,\ldots,\}} \subseteq X$ a distinct ordered family such that $\{e_i | i \in \{1,\ldots,n\}\}$ is a basis for X and $L \in \operatorname{Hom}(X^n; Y)$ such that L is skew-symmetric then $\forall x \in X$ we have*

$$L(x_1,\ldots,x_n) = \left(\sum_{\rho \in P_n} \operatorname{sign}(\sigma) \cdot \prod_{j \in \{1,\ldots,n\}} \alpha^k \right) \cdot L(e_1,\ldots,e_n)$$

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: 11.123].

Proof. By [theorem: 11.250] we have

$$\begin{aligned}
L(x_1,\ldots,x_n) &= \sum_{\sigma \in P_n} \left(\prod_{j \in \{1,\ldots,m\}} \alpha_j^{\sigma(j)} \right) \cdot (\sigma L)(e_1,\ldots,e_n) \\
&\underset{L \text{ is skew-symmetric}}{=} \sum_{\sigma \in P_n} \left(\prod_{j \in \{1,\ldots,m\}} \alpha_j^{\sigma(j)} \right) \cdot (\operatorname{sign}(\sigma) \cdot L(e_1,\ldots,e_n)) \\
&= \left(\sum_{\sigma \in P_n} \left(\prod_{j \in \{1,\ldots,m\}} \alpha_j^{\sigma(j)} \right) \cdot \operatorname{sign}(\sigma) \right) \cdot L(e_1,\ldots,e_n)
\end{aligned}$$

□

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 11.252. *Let $n \in \mathbb{N}$, X, Y vector spaces over the field F, $\{e_i\}_{i \in \{1, \ldots,\}} \subseteq X$ a distinct ordered family such that $\{e_i | i \in \{1, \ldots, n\}\}$ is a basis for X then we have:*

1. *If $L_1, L_2 \in \mathrm{Hom}(X^n; Y)$ 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 $L \in \mathrm{Hom}(X^n; Y)$ is a skew-symmetric n-linear mapping such that $L(e_1, \ldots, e_n) = 0$ then*
$$L = C_0$$

Proof.

1. Using [theorem: 11.251] we have for $x \in X$
$$\begin{aligned} L_1(x) &= L_1(x_1, \ldots, x_n) \\ &= \left(\sum_{\sigma \in P_n} \left(\prod_{j \in \{1, \ldots, m\}} \alpha_j^{\sigma(j)} \right) \cdot \mathrm{sign}(\sigma) \right) \cdot L_1(e_1, \ldots, e_n) \\ &= \left(\sum_{\sigma \in P_n} \left(\prod_{j \in \{1, \ldots, m\}} \alpha_j^{\sigma(j)} \right) \cdot \mathrm{sign}(\sigma) \right) \cdot L_2(e_1, \ldots, e_n) \\ &= L_2(x_1, \ldots, x_n) \\ &= L_2(x) \end{aligned}$$
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: 11.123]. Proving that
$$L_1 = L_2$$
2. Using [theorem: 11.251] we have for $x \in X$ that
$$\begin{aligned} L(x) &= L(x_1, \ldots, x_n) \\ &= \left(\sum_{\sigma \in P_n} \left(\prod_{j \in \{1, \ldots, m\}} \alpha_j^{\sigma(j)} \right) \cdot \mathrm{sign}(\sigma) \right) \cdot L_1(e_1, \ldots, e_n) \\ &= \left(\sum_{\sigma \in P_n} \left(\prod_{j \in \{1, \ldots, m\}} \alpha_j^{\sigma(j)} \right) \cdot \mathrm{sign}(\sigma) \right) \cdot 0 \\ &= 0 \end{aligned}$$
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: 11.123]. So that
$$L = C_0$$ □

We are now ready to define determinant function.

Definition 11.253. *Let $n \in \mathbb{N}$, X a n-dimensional vector space over a field with characteristic zero then $\Delta \in \mathrm{Hom}(X^n; F)$ that is skew-symmetric is called a **determinant function** or **determinant mapping**.*

Example 11.254. 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

$$\begin{array}{rl} C_0(x_1,\ldots,x_{i-1},u+v,x_{i+1},\ldots,x_n) & = \\ 0 & = \\ 0+0 & = \\ C_0(x_1,\ldots,x_{i-1},u,x_{i+1},\ldots,x_n)+C_0(x_1,\ldots,x_{i-1},u,x_{i+1},\ldots,x_n) & = \end{array}$$

and

$$\begin{array}{rcl} C_0(x_1,\ldots,x_{i-1},\alpha\cdot u,x_{i+1},\ldots,x_n) & = & 0 \\ & = & \alpha\cdot 0 \\ & = & \alpha\cdot C_0(x_1,\ldots,x_{i-1},u,x_{i+1},\ldots,x_n) \end{array}$$

So

$$C_0\in\mathrm{Hom}(X^n;Y)$$

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 exists non trivial determinant function.

Theorem 11.255. *Let $n\in\mathbb{N}$, X a n-dimensional vector space over a field F of characteristic zero and $\{e_i\}_{i\in\{1,\ldots,\}}\subseteq X$ a distinct ordered family such that $\{e_i|i\in\{1,\ldots,n\}\}$ is 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 $f_i\colon X\rightarrow F$ by $f_i(x)=\alpha_i$ where $\{\alpha_i\}_{i\in\{1,\ldots,n\}}\subseteq F$ is the **unique** family such that $x=\sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot e_i$ that exist by [theorem: 11.123]. 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

$$\begin{array}{rcl} x+y & \underset{\text{def}\,f_i}{=} & \sum_{i\in\{1,\ldots,n\}}f_i(x)\cdot e_i+\sum_{i\in\{1,\ldots,n\}}f_i(y)\cdot e_i \\ & \underset{\text{[theorem: 11.36]}}{=} & \sum_{i\in\{1,\ldots,n\}}(f_i(x)\cdot e_i+f_i(y)\cdot e_i) \\ & = & \sum_{i\in\{1,\ldots,n\}}(f_i(x)+f_i(y))\cdot e_i \\ x+y & \underset{\text{def}\,f_i}{=} & \sum_{i\in\{1,\ldots,n\}}f_i(x+y)\cdot e_i \end{array}$$

so that by uniqueness we have

$$\forall i\in\{1,\ldots,n\}\text{ that }f_i(x+y)=f_i(x)+f_i(y) \tag{11.128}$$

Likewise

$$\begin{array}{rcl} \alpha\cdot x & \underset{\text{def}\,f_i}{=} & \alpha\cdot\sum_{i\in\{1,\ldots,n\}}f_i(x)\cdot e_i \\ & \underset{\text{[theorem: 11.67]}}{=} & \sum_{i\in\{1,\ldots,n\}}\alpha\cdot f_i(x)\cdot e_i \\ & = & \sum_{i\in\{1,\ldots,n\}}(\alpha\cdot f_i(x))\cdot e_i \\ \alpha\cdot x & \underset{\text{def}\,f_i}{=} & \sum_{i\in\{1,\ldots,n\}}f_i(\alpha\cdot x) \end{array}$$

so by uniqueness we have

$$\forall i \in \{1, \ldots, n\} \; f_i(\alpha \cdot x) = \alpha \cdot f_i(x) \tag{11.129}$$

Let $j \in \{1, \ldots, n\}$ then we have by [theorem: 11.143] that $e_j = \sum_{i \in \{1, \ldots, n\}} \delta_{i,j} \cdot e_j$ so that by uniqueness

$$\forall i \in \{1, \ldots, n\} \; f_i(e_j) = \delta_{i,j} \tag{11.130}$$

Define

$$\Psi \colon X^n \rightarrow F \text{ by } \Psi(x_1, \ldots, x_n) = \prod_{i \in \{1, \ldots, n\}} f_i(x_i)$$

then we have for $i \in \{1, \ldots, n\}$, $u, v \in X_i$ and $\alpha \in F$ that for

$$\begin{aligned} y &= (x_1, \ldots, x_{i-1}, u + v, x_{i+1}, \ldots, x_n) \\ z &= (x_1, \ldots, x_{i-1}, u, x_{i+1}, \ldots, x_n) \\ r &= (x_1, \ldots, x_{i-1}, v, x_{i+1}, \ldots, x_n) \\ s &= (x_1, \ldots, x_{i-1}, \alpha \cdot u, x_{i+1}, \ldots, x_n) \end{aligned}$$

$$\begin{aligned} \Psi(x_1, \ldots, x_{i-1}, u + v, x_{i+1}, \ldots, x_n) &= \prod_{j \in \{1, \ldots, n\}} y_j \\ &= \left(\prod_{j \in \{i\}} y_j \right) \cdot \prod_{j \in \{1, \ldots, n\} \setminus \{i\}} y_j \Bigg) \\ &= y_i \cdot \prod_{j \in \{1, \ldots, n\} \setminus \{i\}} y_j \\ &= (u + v) \cdot \prod_{j \in \{1, \ldots, n\} \setminus \{i\}} y_j \\ &= u \cdot \prod_{j \in \{1, \ldots, n\} \setminus \{i\}} y_j + v \cdot \prod_{j \in \{1, \ldots, n\} \setminus \{i\}} y_j \\ &= z_i \cdot \prod_{j \in \{1, \ldots, n\} \setminus \{i\}} z_j + r_i \cdot \prod_{j \in \{1, \ldots, n\} \setminus \{i\}} r_j \\ &= \prod_{j \in \{1, \ldots, n\}} z_j + \prod_{j \in \{1, \ldots, n\}} r_j \\ &= L(x_1, \ldots, x_{i-1}, u, x_{i+1}, \ldots, x_n) + L(x_1, \ldots, x_{i-1}, v, x_{i+1}, \ldots, x_n) \end{aligned}$$

and

$$\begin{aligned} \Psi(x_1, \ldots, x_{i-1}, \alpha \cdot u, x_{i+1}, \ldots, x_n) &= \prod_{j \in \{1, \ldots, n\}} s_j \\ &= \left(\prod_{j \in \{i\}} s_j \right) \cdot \prod_{j \in \{1, \ldots, n\} \setminus \{i\}} s_j \Bigg) \\ &= s_j \cdot \prod_{j \in \{1, \ldots, n\} \setminus \{i\}} s_j \\ &= (\alpha \cdot u) \cdot \prod_{j \in \{1, \ldots, n\} \setminus \{i\}} s_j \\ &= \alpha \cdot \left(u \cdot \prod_{j \in \{1, \ldots, n\} \setminus \{i\}} s_j \right) \\ &= \alpha \cdot \left(z_i \cdot \prod_{j \in \{1, \ldots, n\} \setminus \{i\}} z_j \right) \\ &= \alpha \cdot \prod_{j \in \{1, \ldots, n\}} z_j \\ &= \alpha \cdot L(x_1, \ldots, x_{i-1}, u, x_{i+1}, \ldots, x_n) \end{aligned}$$

So we have that

$$\Psi \in \mathrm{Hom}(X^n; F) \text{ or } \Psi \text{ is } n\text{-linear}$$

Using [theorem: 11.246] we create the skew symmetric n-linear mapping

$$\Delta: X^n \rightarrow F \text{ by } \Delta = \sum_{\sigma \in P_n} \mathrm{sign}(\sigma) \cdot \sigma\Psi \tag{11.131}$$

which is then by definition a determinant function. Let $\sigma \in P_n$ then we have either:

$\boldsymbol{\sigma = \mathrm{Id}_{\{1,\ldots,n\}}}$**.** Then

$$\sigma\Psi(e_1, \ldots, e_n) = \Psi(e_1, \ldots, e_n) = \prod_{i \in \{1,\ldots,n\}} f_i(e_i) \underset{\forall i \in I f_i(i) = 1}{=} 1$$

$\boldsymbol{\sigma \in P_n \setminus \{\mathrm{Id}_{\{1,\ldots,n\}}\}}$**.** Then there exists a $i \in \{1, \ldots, n\}$ such that $i \neq \mathrm{Id}_{\{\text{`},\ldots,n\}}(i) = \sigma(i)$ so that $f_i(\sigma(i)) = \delta_{i,\sigma(i)} = 0$ then

$$\begin{aligned} \sigma\Psi(e_1, \ldots, e_n) &= \Psi(e_{\sigma(1)}, \ldots, e_{\sigma(n)}) \\ &= \prod_{j \in \{1,\ldots,n\}} f_j(e_j) \\ &= \left(\prod_{j \in \{i\}} f_j(e_j)\right) \prod_{j \in \{1,\ldots,n\} \setminus \{i\}} f_j(e_j) \\ &= f_i(e_i) \cdot \prod_{j \in \{1,\ldots,n\} \setminus \{i\}} f_j(e_j) \\ &= 0 \cdot \prod_{j \in \{1,\ldots,n\} \setminus \{i\}} f_j(e_j) \\ &=0 \end{aligned}$$

So that

$$\sigma\Psi(e_1, \ldots, e_n) = \begin{cases} 0 \text{ if } \sigma \neq \mathrm{Id}_{\{1,\ldots,n\}} \\ 1 \text{ if } \sigma \in P_n \setminus \{\mathrm{Id}_{\{1,\ldots,n\}}\} \end{cases} \tag{11.132}$$

so that

$$\begin{aligned} \Delta(e_1, \ldots, e_n) &\underset{\text{[eq: 11.131]}}{=} \\ \sum_{\sigma \in P_n} \mathrm{sign}(\sigma) \cdot (\sigma\Psi)(e_1, \ldots, e_n) &\underset{\text{[theorem: 11.41]}}{=} \\ \sum_{\sigma \in P_n \setminus \{\mathrm{Id}_{\{1,\ldots,n\}}\}} \mathrm{sign}(\sigma) \cdot (\sigma\Psi)(e_1, \ldots, e_n) + (\mathrm{Id}_{\{1,\ldots,n\}}\Psi)(e_1, \ldots, e_n) &\underset{\text{[theorem: 11.32]}}{=} \\ \sum_{\sigma \in P_n \setminus \{\mathrm{Id}_{\{1,\ldots,n\}}\}} \mathrm{sign}(\sigma) \cdot 0 + 1 &\underset{\text{[eq: 11.132]}}{=} \\ 1 & \end{aligned}$$

proving that

$$\Delta(e_1, \ldots, e_n) = 1$$

□

Given a determinant function then we can generate a

Theorem 11.256. *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 $L \in \mathrm{Hom}(X^n; Y)$ there exist a* ***unique*** *$y \in Y$ such that*

$$\forall x \in X \text{ we have } L(x) = \Delta(x) \cdot y$$

Proof. Let $\{e_i\}_{i \in \{1,\ldots,n\}} \subseteq X$ a distinct family of vectors such that $\{e_i | i \in \{1, \ldots, n\}\}$ is a basis of X. As $\Delta \neq C_0$ we and by [theorem: 11.251]

$$\Delta(x_1, \ldots, x_n) = \left(\sum_{\rho \in P_n} \mathrm{sign}(\sigma) \cdot \prod_{j \in \{1,\ldots,n\}} \alpha^k\right) \cdot \Delta(e_1, \ldots, e_n)$$

we have that

$$\Delta(e_1,\ldots,e_n)\neq 0$$

This allows us to define

$$\{f_i\}_{i\in\{1,\ldots,n\}}\subseteq X \text{ by } f_i=\begin{cases}\Delta(e_1,\ldots,e_n)^{-1}\cdot e_1 \text{ if } i=1\\ e_i \text{ if } i\in\{2,\ldots,n\}\end{cases}$$

We prove now that

$$\{f_i\}_{i\in\{1,\ldots,n\}} \text{ is a basis for } X \tag{11.133}$$

Proof. Let $\{\alpha_i\}_{i\in\{1,\ldots,n\}}\subseteq F$ be a family such that $\sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot f_i=0$. Define

$$\{\beta_i\}_{i\in\{1,\ldots,n\}}\subseteq F \text{ by } \beta_i=\begin{cases}\Delta(e_1,\ldots,e_n)^{-1}\cdot\alpha_1 \text{ if } i=1\\ \alpha_i \text{ if } i\in\{2,\ldots,n\}\end{cases}$$

then we have

$$\begin{aligned}\sum_{i\in\{1,\ldots,n\}}\beta_i\cdot e_i &= \sum_{i\in\{2,\ldots,n\}}\beta_i\cdot e_i+\sum_{i\in\{1\}}\beta_i\cdot e_i\\ &= \sum_{i\in\{2,\ldots,n\}}\alpha_i\cdot f_i+\sum_{i\in\{1\}_i}(\Delta(e_1,\ldots,e_n)^{-1}\cdot\alpha_1)\cdot(\Delta(e_1,\ldots,e_n)\cdot f_1)\\ &= \sum_{i\in\{2,\ldots,n\}}\alpha_i\cdot f_i+\sum_{i\in\{1\}_i}\alpha_i\cdot f_i\\ &= \sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot f_i\\ &= 0\end{aligned}$$

As $\{e_i|i\in\{1,\ldots,n\}\}$ is a basis hence linear independent we have by [theorem: 11.111] that $\{e_i\}_{i\in\{1,\ldots,n\}}$ is linear independent, so $\forall i\in\{1,\ldots,n\}$ we have that $\beta_i=0$. Hence $\forall i\in\{1,\ldots,n\}$

$$\beta_i=\begin{cases}\Delta(e_1,\ldots,e_n)^{-1}\cdot\alpha_1 \text{ if } i=1\\ \alpha_i \text{ if } i\in\{2,\ldots,n\}\end{cases}=0$$

proving that $\{f_i\}_{i\in\{1,\ldots,n\}}$ is a linear independent family. Assume that $\{f_i|i\in\{1,\ldots,n\}\}$ is linear dependent then by [theorem: 11.106] it follows that $\{f_i\}_{i\in\{1,\ldots,n\}}$ is linear dependent contradicting the linear independence of $\{f_i\}_{i\in\{1,\ldots,n\}}$. Hence we must have that

$$\{f_i|i\in\{1,\ldots,n\}\} \text{ is linear independent} \tag{11.134}$$

Let $x\in X$ then as $\{e_i|i\in\{1,\ldots,n\}\}$ is a basis we have by [theorem: 11.123] that there exists a $\{\beta_i\}_{i\in\{1,\ldots,n\}}\subseteq F$ such that $\sum_{i\in\{1,\ldots,n\}}\beta_i\cdot e_i=x$. Define now

$$\{\alpha_i\}_{i\in\{1,\ldots,n\}} \text{ by } \alpha_i=\begin{cases}\Delta(e_1,\ldots,e_n)\cdot\beta_1 \text{ if } i=1\\ \beta_i \text{ if } i\in\{2,\ldots,n\}\end{cases}$$

then

$$\begin{aligned}\sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot f_i &= \sum_{i\in\{2,\ldots,n\}}\alpha_i\cdot f_i+\sum_{i\in\{1\}}\alpha_i\cdot f_i\\ &= \sum_{i\in\{2,\ldots,n\}}\beta_i\cdot e_i+\sum_{i\in\{1\}}(\Delta(e_1,\ldots,e_n)\cdot\beta_i)\cdot(\Delta(e_1,\ldots,e_n)^{-1}\cdot\alpha_i)\\ &= \sum_{i\in\{2,\ldots,n\}}\beta_i\cdot e_i+\sum_{i\in\{1\}}\alpha_i\cdot e_i\\ &= \sum_{i\in\{1,\ldots,n\}}\beta_i\cdot e_i\\ &= x\end{aligned}$$

proving by [theorem: 11.87] that $x\in\operatorname{span}(\{f_i\}_{i\in\{1,\ldots,n\}})$, hence $X\subseteq\operatorname{span}(\{f_i|i\in\{1,\ldots,n\}\})\subseteq X$. So

$$\operatorname{span}(\{f_i|i\in\{1,\ldots,n\}\})=X \tag{11.135}$$

From [eqs: 11.134,11.135] it follows that

$$\{f_i|i\in\{1,\ldots,n\}\} \text{ is a basis of } X \qquad\square$$

Now

$$\begin{array}{rcl}\Delta(f_1,\ldots,f_n) & = & \Delta(\Delta(e_1,\ldots,e_n)^{-1}\cdot e_1,e_2,\ldots,e_n)\\ & \underset{\Delta\in\operatorname{Hom}(X^n;Y)}{=} & \Delta(e_1,\ldots,e_n)^{-1}\cdot D(e_1,\ldots,e_n)\\ & = & 1\end{array} \tag{11.136}$$

Set

$$y=L(f_1,\ldots,f_n)$$

and define

$$K:X^n\rightarrow Y \text{ by } K(x_1,\ldots,x_n)=\Delta(x_1,\ldots,x_n)\cdot y$$

The for $i\in\{1,\ldots,n\}$, $u,v\in X$ we have that

$$\begin{array}{rl}K(x_1,\ldots,x_{i-1},u+v,x_{i+1},\ldots,x_n) & =\\ \Delta(x_1,\ldots,x_{i-1},u+v,x_{i+1},\ldots,x_n)\cdot y & =\\ (\Delta(x_1,\ldots,x_{i-1},u,x_{i+1},\ldots,x_n)+\Delta(x_1,\ldots,x_{i-1},v,x_{i+1},\ldots,x_n))\cdot y & =\\ \Delta(x_1,\ldots,x_{i-1},u,x_{i+1},\ldots,x_n)\cdot y+\Delta(x_1,\ldots,x_{i-1},v,x_{i+1},\ldots,x_n)\cdot y & =\\ K(x_1,\ldots,x_{i-1},u,x_{i+1},\ldots,x_n)+K(x_1,\ldots,x_{i-1},v,x_{i+1},\ldots,x_n) & \end{array}$$

and

$$\begin{array}{rcl}K(x_1,\ldots,x_{i-1},\alpha\cdot u,x_{i+1},\ldots,x_n) & = & \Delta(x_1,\ldots,x_{i-1},\alpha\cdot u,x_{i+1},\ldots,x_n)\cdot y\\ & = & (\alpha\cdot\Delta(x_1,\ldots,x_{i-1},u,x_{i+1},\ldots,x_n))\cdot y\\ & = & \alpha\cdot(\Delta(x_1,\ldots,x_{i-1},u,x_{i+1},\ldots,x_n)\cdot y)\\ & = & \alpha\cdot K(x_1,\ldots,x_{i-1},u,x_{i+1},\ldots,x_n)\end{array}$$

which proves that

$$K\in\operatorname{Hom}(X^n;Y)$$

Further if $\sigma\in P_n$ then if $x\in X^n$ we have

$$K(x\circ\sigma)=\Delta(x\circ\sigma)\cdot y=(\operatorname{sign}(\sigma)\cdot\Delta(x))\cdot y=\operatorname{sign}(\sigma)\cdot(\Delta(x)\cdot y)=\operatorname{sign}(\sigma)\cdot K(x)$$

proving that

$$K \text{ is a skew-symmetric linear function}$$

As $K(f_1,\ldots,f_n)=\Delta(f_1,\ldots,f_n)\cdot y\underset{\text{[eq: 11.136]}}{=}y=L(f_1,\ldots,f_n)$ we have by [theorem: 11.252] that $L=K$ or using the definition of K 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 $L(x)=\Delta(x)y$'. Then $\Delta(f_1,\ldots,f_n)\cdot y=\Delta(f_1,\ldots,f_n)\cdot y'\underset{\text{[eq: 11.136]}}{\Rightarrow}1\cdot y=1\cdot y'$ proving that $y=y'$. □

Corollary 11.257. *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:

$\boldsymbol{\Delta = C_0}$. Take $\alpha = 0$ then

$$\Delta' = C_0 = 0 \cdot \Delta$$

$\boldsymbol{\Delta \neq C_0}$. As Δ is a determinant function, hence a skew-symmetric n-linear mapping with range F, we have by the previous theorem [theorem: 11.256] that there exist a $\alpha \in F$ such that

$$\Delta' = \Delta \cdot \alpha = \alpha \cdot \Delta$$ □

Definition 11.258. *Let $n \in \mathbb{N}$, X a set, $y \in X$, $i \in \{1, \ldots, n\}$ and $x \in X^n$ then $(y, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$ is defined by*

$$(y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_k = \begin{cases} y \textit{ if } k = 1 \\ x_{k-1} \textit{ if } k \in \{2, \ldots, i\} \\ x_k \textit{ if } k \in \{i+1, \ldots, n\} \end{cases}$$

Lemma 11.259. *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:

$\boldsymbol{i = 1}$. Then by [definition: 11.212] we have that $\left(1 \underset{n}{\rightsquigarrow} i\right) = \left(1 \underset{n}{\rightsquigarrow} 1\right) = \mathrm{Id}_{\{1, \ldots, n\}}$, let $j \in \{1,\ldots,n\}$ then we have for $j \in \{1, \ldots, n\}$

$$\begin{aligned}
(x_i, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_j &= \begin{cases} x_i \text{ if } j = 1 \\ x_{i-1} \text{ if } j \in \{2, \ldots, i\} \\ x_k \text{ if } j \in \{i+1, \ldots, n\} \end{cases} \\
&\underset{i=1}{=} \begin{cases} x_1 \text{ if } j = 1 \\ x_{i-1} \text{ if } j \in \{2, \ldots, 1\} \\ x_k \text{ if } j \in \{2, \ldots, n\} \end{cases} \\
&= \begin{cases} x_1 \text{ if } j = 1 \\ x_{i-1} \text{ if } j \in \{2, \ldots, 1\} \\ x_k \text{ if } j \in \{2, \ldots, n\} \end{cases} \\
&= \begin{cases} x_1 \text{ if } j = 1 \\ x_k \text{ if } j \in \{2, \ldots, n\} \end{cases} \\
&= x_j \\
&= (x \circ \mathrm{Id}_{\{1, \ldots, n\}})_j \\
&= \left(x \circ \left(i \underset{n}{\rightsquigarrow} 1\right)\right)(j)
\end{aligned}$$

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 $j \in \{1, \ldots, n\}$ either:

$\boldsymbol{j = 1}$. Then

$$\begin{aligned}
(x_i, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_j &= (x_i, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)(1) \\
&= x_i \\
&= x(i) \\
&\underset{j=1 \wedge [\text{definition: } 11.212]}{=} x\left(\left(i \underset{n}{\rightsquigarrow} 1\right)(j)\right) \\
&= \left(x \circ \left(i \underset{n}{\rightsquigarrow} 1\right)\right)(j)
\end{aligned}$$

$j \in \{2, \ldots, i\}$. Then

$$\begin{aligned}(x_i, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_j &= x_{j-1}\\ &= x(j-1)\\ &\underset{1<j\leqslant i\wedge[\text{definition: } 11.212]}{=} x\big(\big(i \underset{n}{\rightsquigarrow} 1\big)(j)\big)\\ &= \big(x \circ \big(i \underset{n}{\rightsquigarrow} 1\big)\big)(j)\end{aligned}$$

$j \in \{i+1, \ldots, n\}$. Then

$$\begin{aligned}(x_i, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_j &= x_j\\ &= x(j)\\ &\underset{i<i+1<j\leqslant n\wedge[\text{definition: } 11.212]}{=} x\big(\big(i \underset{n}{\rightsquigarrow} 1\big)(j)\big)\\ &= \big(x \circ \big(i \underset{n}{\rightsquigarrow} 1\big)\big)(j)\end{aligned}$$

proving that

$$(x_i, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) = x \circ \big(i \underset{n}{\rightsquigarrow} 1\big) \qquad \square$$

Definition 11.260. *Let $n \in \mathbb{N}$, X a n-dimensional vector space over a field F with characteristic zero and $\Delta \in \operatorname{Hom}(X^n; F)$ 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: 11.258]

$$(y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_k = \begin{cases} y \text{ if } k = 1\\ x_{k-1} \text{ if } k \in \{2, \ldots, i\}\\ x_k \text{ if } k \in \{i+1, \ldots, n\}\end{cases}$$

To calculate $\underline{\Delta}(y, (x_1, \ldots, x_n))$ we need the following lemma.

Lemma 11.261. *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 we have for $z = (y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$ we have:*

1. *If $j + 1 \leqslant i$ then $(z_1, \ldots, z_j, x_i, z_{j+2}, \ldots, z_n) \circ \big(j + 1 \underset{n}{\rightsquigarrow} i\big) = (y, x_1, \ldots, x_{j-1}, x_j, \ldots, x_n)$*
2. *If $i + 1 \leqslant j$ then $(z_1, \ldots, z_{j-1}, x_i, z_{j+1}, \ldots, z_n) \circ \big(j \underset{n}{\rightsquigarrow} i + 1\big) = (y, x_1, \ldots, x_{j-1}, x_j, \ldots, x_n)$*

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 $a_k = (z_1, \ldots, z_j, x_i, z_{j+2}, \ldots, z_n)\big(\big(j + 1 \underset{n}{\rightsquigarrow} i\big)(x)\big)$ then using the definition of $\underset{n}{\rightsquigarrow}$ [see definition: 11.212] we must look at the following cases for $j + 1, i$

   $j + 1 = i$. Then by [definition: 11.212] we have that $\big(j + 1 \underset{n}{\rightsquigarrow} i\big) = \mathrm{Id}_{\{1, \ldots, n\}}$ so that

   $$a_k = (z_1, \ldots, z_j, x_i, z_{j+2}, \ldots, z_n)(k) \tag{11.137}$$

   now for $k \in \{1, \ldots, n\}$ we have to look at the following cases:

   $k = 1$. Then

   $$\begin{aligned}a_k &\underset{[\text{eq: } 11.137]}{=} (z_1, \ldots, z_j, x_i, z_{j+2}, \ldots, z_n)_k\\ &\underset{k=1<j+1}{=} z_k\\ &= (y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_k\\ &\underset{[\text{definition: } 11.258]\wedge k=1}{=} y\\ &\underset{[\text{definition: } 11.258]\cap k=1}{=} (y, x_1, \ldots, x_{j-1}, x_j, \ldots, x_n)_k\end{aligned}$$

$1 < k < j$. Then

$$\begin{aligned} a_k &\underset{\text{[eq: 11.137]}}{=} (z_1,\ldots,z_j,x_i,z_{j+2},\ldots,z_n)_k \\ &\underset{k<j<j+1}{=} z_k \\ &= (y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_k \\ &\underset{\text{[definition: 11.258]}\wedge 1<k<j<j+1=i}{=} x_{k-1} \\ &\underset{\text{[definition: 11.258]}\wedge 1<k<j}{=} (y,x_1,\ldots,x_{j-1},x_j,\ldots,x_n)_k \end{aligned}$$

$1 < k = j+1$. Then

$$\begin{aligned} a_k &\underset{\text{[eq: 11.137]}}{=} (z_1,\ldots,z_j,x_i,z_{j+2},\ldots,z_n)_k \\ &\underset{k=j+1}{=} x_i \\ &\underset{k=j+1=i}{=} x_k \\ &\underset{\text{[definition: 11.258]}\wedge 1<k=j+1=i}{=} (y,x_1,\ldots,x_{j-1},x_j,\ldots,x_n)_k \end{aligned}$$

$j+1 < k \leqslant n$. Then

$$\begin{aligned} a_k &\underset{\text{[eq: 11.137]}}{=} (z_1,\ldots,z_j,x_i,z_{j+2},\ldots,z_n)_k \\ &\underset{j+1<k}{=} z_k \\ &= (y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_k \\ &\underset{\text{[definition: 11.258]}\wedge i=j+1<k\Rightarrow i+1\leqslant k}{=} x_k \\ &\underset{\text{[definition: 11.258]}\wedge j+1<k}{=} (y,x_1,\ldots,x_{j-1},x_j,\ldots,x_n)_k \end{aligned}$$

$j+1 < i$. Then by [definition: 11.212] we have look at the following cases:

$1 = k$. Then

$$\begin{aligned} a_k &\underset{11.212}{=} (z_1,\ldots,z_j,x_i,z_{j+2},\ldots,z_n)_k \\ &\underset{k=1\leqslant j<j+1}{=} z_k \\ &\underset{k=1}{=} z_1 \\ &= y \\ &\underset{\text{[definition: 11.258\&k=1]}}{=} (y,x_1,\ldots,x_{j-1},x_j,\ldots,x_n)_k \end{aligned}$$

$1 < k < j+1$. Then

$$\begin{aligned} a_k &\underset{11.212}{=} (z_1,\ldots,z_j,x_i,z_{j+2},\ldots,z_n)_k \\ &\underset{k=1\leqslant j<j+1}{=} z_k \\ &= (y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_k \\ &\underset{\text{[definition: 11.258]}\wedge 1<k<j+1\leqslant i}{=} x_{k-1} \\ &\underset{\text{[definition: 11.258]}\wedge 1<k<j+1\Rightarrow 1<k\leqslant j}{=} (y,x_1,\ldots,x_{j-1},x_j,\ldots,x_n)_k \end{aligned}$$

$j+1 \leqslant k < i$. Then

$$\begin{aligned} a_k &\underset{11.212}{=} (z_1,\ldots,z_j,x_i,z_{j+2},\ldots,z_n)_{k+1} \\ &\underset{j+1\leqslant k<k+1}{=} z_{k+1} \\ &= (y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_{k+1} \\ &\underset{\text{[definition: 11.258]}\wedge 1<k+1\leqslant i}{=} x_{(k+1)-1} \\ &= x_k \\ &\underset{\text{[definition: 11.258]}\wedge 1<k<i}{=} (y,x_1,\ldots,x_{j-1},x_j,\ldots,x_n)_k \end{aligned}$$

$\boldsymbol{k=i.}$ Then

$$\begin{aligned}
a_k &\underset{11.212}{=} (z_1,\ldots,z_j,x_i,z_{j+2},\ldots,z_n)_{j+1}\\
&= x_i\\
&\underset{i=k}{=} x_k\\
&\underset{\text{[definition: 11.258]}j<j+1<i=k}{=} (y,x_1,\ldots,x_{j-1},x_j,\ldots,x_n)_k
\end{aligned}$$

$\boldsymbol{i<k\leqslant n.}$ Then

$$\begin{aligned}
a_k &\underset{11.212}{=} (z_1,\ldots,z_j,x_i,z_{j+2},\ldots,z_n)_k\\
&= z_k\\
&= (y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_k\\
&\underset{\text{[definition: 11.258]}\wedge i<k}{=} x_k\\
&\underset{\text{[definition: 11.258]}\wedge j+1<i<k}{=} (y,x_1,\ldots,x_{j-1},x_j,\ldots,x_n)_k
\end{aligned}$$

so in all cases we have $a_k=(y,x_1,\ldots,x_{j-1},x_j,\ldots,x_n)_k$ proving that

$$(z_1,\ldots,z_j,x_i,z_{j+2},\ldots,z_n)\circ\left(j+1\underset{n}{\rightsquigarrow}i\right)=(y,x_1,\ldots,x_{j-1},x_j,\ldots,x_n)$$

2. Let $k\in\{1,\ldots,n\}$ and take $b_k=(z_1,\ldots,z_{j-1},x_i,z_{j+1},\ldots,z_n)\left(\left(j\underset{n}{\rightsquigarrow}i+1\right)(k)\right)$ then using the definition of $\underset{n}{\rightsquigarrow}$ [see definition: 11.212] we must look at the following cases for $j,i+1$:

   $\boldsymbol{i+1=j.}$ Then by [definition: 11.212] we have that $\left(j\underset{n}{\rightsquigarrow}i+1\right)=\mathrm{Id}_{\{1,\ldots,n\}}$ so that

   $$b_k=(k) \tag{11.138}$$

   Now for $k\in\{1,\ldots,n\}$ we have to look at the following cases:

   $\boldsymbol{k=1.}$ Then

   $$\begin{aligned}
   b_k &\underset{\text{[eq: 11.138]}}{=} (z_1,\ldots,z_{j-1},x_i,z_{j+1},\ldots,z_n)_k\\
   &\underset{k=1<i+1=j}{=} z_1\\
   &= (y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_1\\
   &\underset{\text{[definition: 11.258]}\wedge k=1}{=} y\\
   &\underset{\text{[definition: 11.258]}\wedge k=1}{=} (y,x_1,\ldots,x_{j-1},x_j,\ldots,x_n)_k
   \end{aligned}$$

   $\boldsymbol{1<k<i.}$ Then

   $$\begin{aligned}
   b_k &\underset{\text{[eq: 11.138]}}{=} (z_1,\ldots,z_{j-1},x_i,z_{j+1},\ldots,z_n)_k\\
   &\underset{k<i<i+1=j}{=} z_k\\
   &= (y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_k\\
   &\underset{\text{[definition: 11.258]}\wedge 1<k<i}{=} x_{k-1}\\
   &\underset{\text{[definition: 11.258]}\wedge 1<k<i<i+1<j}{=} (y,x_1,\ldots,x_{j-1},x_j,\ldots,x_n)_k
   \end{aligned}$$

   $\boldsymbol{1<k=i.}$ Then

   $$\begin{aligned}
   b_k &\underset{\text{[eq: 11.138]}}{=} (z_1,\ldots,z_{j-1},x_i,z_{j+1},\ldots,z_n)_k\\
   &\underset{k=i<i+1=j}{=} z_k\\
   &= (y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_k\\
   &\underset{\text{[definition: 11.258]}\wedge 1<k=i}{=} x_{k-1}\\
   &\underset{\text{[definition: 11.258]}\wedge k=i<i+1=j}{=} (y,x_1,\ldots,x_{j-1},x_j,\ldots,x_n)_k
   \end{aligned}$$

$\boldsymbol{k = i+1.}$ Then

$$\begin{aligned}
b_k &\underset{\text{[eq: 11.138]}}{=} (z_1, \ldots, z_{j-1}, x_i, z_{j+1}, \ldots, z_n)_k \\
&\underset{k=i+1=j}{=} x_i \\
&\underset{k=i+1}{=} x_{k-1} \\
&\underset{\text{[definition: 11.258]} \wedge 1<k=i+1=j}{=} (y, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_k
\end{aligned}$$

$\boldsymbol{i+1 < k \leqslant n.}$ Then

$$\begin{aligned}
b_k &\underset{\text{[eq: 11.138]}}{=} (z_1, \ldots, z_{j-1}, x_i, z_{j+1}, \ldots, z_n)_k \\
&\underset{j=i+1<k}{=} z_k \\
&= (y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_k \\
&\underset{\text{[definition: 11.258]} \wedge i+1<k}{=}
\end{aligned}$$

$\boldsymbol{i+1 < j.}$ Then by [definition: 11.212] we have look at the following cases:

$\boldsymbol{k = 1.}$ Then

$$\begin{aligned}
b_k &\underset{11.212}{=} (z_1, \ldots, z_{j-1}, x_i, z_{j+1}, \ldots, z_n)_k \\
&\underset{k=1<i+1<j}{=} z_k \\
&= (y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_k \\
&\underset{\text{[definition: 11.258]} \wedge k=1}{=} y \\
&\underset{\text{[definition: 11.258]} \wedge k=1}{=} (y, x_1, \ldots, x_{j-1}, x_j, \ldots, x_n)_k
\end{aligned}$$

$\boldsymbol{1 < k < i+1.}$ Then

$$\begin{aligned}
b_k &\underset{11.212}{=} (z_1, \ldots, z_{j-1}, x_i, z_{j+1}, \ldots, z_n)_k \\
&\underset{k<i+1<j}{=} z_k \\
&= (y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_k \\
&\underset{\text{[definition: 11.258]} \wedge 1<k<i+1 \Rightarrow 1<k\leqslant i}{=} x_{k-1} \\
&\underset{\text{[definition: 11.258]} \wedge 1<k<i+1<j}{=} (y, x_1, \ldots, x_{j-1}, x_j, \ldots, x_n)_k
\end{aligned}$$

$\boldsymbol{1 < k = i+1.}$ Then

$$\begin{aligned}
b_k &\underset{11.212}{=} (z_1, \ldots, z_{j-1}, x_i, z_{j+1}, \ldots, z_n)_{i+1} \\
k = i+1 < j \quad & z_{i+1} \\
&= (y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_{i+1} \\
&\underset{\text{[definition: 11.258]} \wedge 1<i+1\leqslant i+1}{=} x_i \\
&= \underset{k=i+1}{=} x_{k-1} \\
&\underset{\text{[definition: 11.258]} \wedge 1<k=i+1<j}{=} (y, x_1, \ldots, x_{j-1}, x_j, \ldots, x_n)_k
\end{aligned}$$

$\boldsymbol{i+1 < k \leqslant j.}$ Then

$$\begin{aligned}
b_k &\underset{11.212}{=} (z_1, \ldots, z_{j-1}, x_i, z_{j+1}, \ldots, z_n)_{k-1} \\
&\underset{k\leqslant j \Rightarrow k-1<j}{=} z_{k-1} \\
&= (y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_{k-1} \\
&\underset{\text{[definition: 11.258]} \wedge i+1<k}{=} x_{k-1} \\
&\underset{\text{[definition: 11.258]} \wedge 1<i+1<k\leqslant j}{=} (y, x_1, \ldots, x_{j-1}, x_j, \ldots, x_n)_k
\end{aligned}$$

$\boldsymbol{j<k\leqslant n.}$ Then

$$\begin{array}{lcl} b_k & \underset{11.212}{=} & (z_1,\ldots,z_{j-1},x_i,z_{j+1},\ldots,z_n)(k) \\ & \underset{j<k}{=} & z_k \\ & = & (y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_k \\ & \underset{[\text{definition: } 11.258]\wedge i+1<j<k}{=} & x_k \\ & \underset{[\text{definition: } 11.258]\wedge j<k\Rightarrow j+1\leqslant k}{=} & (y,x_1,\ldots,x_{j-1},x_j,\ldots,x_n)_k \end{array}$$

So in all cases we have $b_k=(y,x_1,\ldots,x_{j-1},x_j,\ldots,x_n)_k$ proving that

$$(z_1,\ldots,z_{j-1},x_i,z_{j+1},\ldots,z_n)\circ\left(j\underset{n}{\rightsquigarrow}i+1\right)=(y,x_1,\ldots,x_{j-1},x_j,\ldots,x_n) \qquad \square$$

Theorem 11.262. *Let $n\in\mathbb{N}$, X a vector space over a field of characteristic zero with $\dim(X)=n$, $y\in X$, $(x_1,\ldots,x_n)\in X^n$ and $\Delta\in\mathrm{Hom}(X^n;F)$ 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

$$\sum_{i\in\{1,\ldots,n\}}(-1)^{i-1}\cdot\Delta(y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)\cdot x_i=\Delta(x_1,\ldots,x_n)\cdot y$$

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:

$\boldsymbol{\{x_i\}_{i\in\{1,\ldots,n\}}}$ **is linear dependent.** Then by [theorem: 11.248] we have that

$$\Delta(x_1,\ldots,x_n)=0 \tag{11.139}$$

So if we prove that $\underline{\Delta}(y,(x_1,\ldots,x_n))=0$ we are done for this case. Let $y\in X$ and $(x_1,\ldots,x_n)\in X^n$ and consider the following cases for $n\in\mathbb{N}$.

$\boldsymbol{n=1.}$ 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

$$\begin{array}{rcl} \underline{\Delta}(y,(x_1,\ldots,x_n)) & = & \displaystyle\sum_{i\in\{1,\ldots,1\}}(-1)^{i-1}\cdot\Delta(y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)\cdot x_i \\ & = & (-1)^0\cdot\Delta(y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)\cdot x_1 \\ & = & 0 \end{array}$$

$\boldsymbol{n=2.}$ As $\{x_i\}_{i\in\{1,2\}}$ is linear dependent it follows by [theorem: 11.113] that there exist a $\alpha\in F$ such that $x_1=\alpha\cdot$ Further

$$\begin{array}{rcl} \underline{\Delta}(y,(x_1,\ldots,x_n)) & = & \displaystyle\sum_{i\in\{1,\ldots,2\}}(-1)^{i-1}\cdot\Delta(y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)\cdot x_i \\ & = & (-1)^{1-1}\cdot\Delta(y,x_2)\cdot x_1+(-1)^{2-1}\cdot\Delta(y,x_1)\cdot x_2 \\ & = & \Delta(y,x_2)\cdot x_1-\Delta(y,x_1)\cdot x_2 \\ & = & \Delta(y,x_2)\cdot(\alpha\cdot x_1)-\Delta(y,\alpha\cdot x_2)\cdot x_1 \\ & = & \alpha\cdot(\Delta(y,x_2)\cdot x_1)-\alpha\cdot(\Delta(y\cdot x_2)\cdot x_1) \\ & = & 0 \end{array}$$

$\boldsymbol{2<n.}$ As $\{x_i\}_{i\in\{1,\ldots,n\}}$ is linear dependent there exists by [theorem: 11.112] a $k\in\{1,\ldots,n\}$ and a $\{\alpha_i\}_{i\in\{1,\ldots,n\}\setminus\{k\}}\subseteq F$ such that

$$x_k=\sum_{i\in\{1,\ldots,n\}\setminus\{k\}}\alpha_i\cdot x_i \tag{11.140}$$

So

$$\begin{aligned}\underline{\Delta}(y,(x_1,\ldots,x_n)) &= \sum_{i\in\{1,\ldots,n\}} ((-1)^{i-1}\cdot\Delta(y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n))\cdot x_i\\ &= A+B \qquad (11.141)\end{aligned}$$

where

$$\begin{aligned}A &= \sum_{i\in\{1,\ldots,n\}\setminus\{k\}} (-1)^{i-1}\cdot\Delta(y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)\cdot x_i \qquad (11.142)\\ B &= \sum_{i\in\{k\}} (-1)^{i-1}\cdot\Delta(y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)\cdot x_i \qquad (11.143)\end{aligned}$$

Now

$$\begin{aligned}B &= \sum_{i\in\{k\}} (-1)^{i-1}\cdot\Delta(y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)\cdot x_i\\ &= (-1)^{k-1}\cdot\Delta(y,x_1,\ldots,x_{k-1},x_{k+1},\ldots,x_n)\cdot x_k\\ &= (-1)\cdot((-1)^{k-1}\cdot\Delta(y,x_1,\ldots,x_{k-1},x_{k+1},\ldots,x_n)\cdot x_k) \qquad (11.144)\end{aligned}$$

For $i\in\{1,\ldots,n\}\setminus\{k\}$ we have the following cases:

***k* < *i*.** Let

$$z=(y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)$$

then

$$\begin{aligned}z_{k+1} &= (y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_{k+1}\\ &\underset{[\text{definition: } 11.258]\wedge 1<k+1\leqslant i}{=} x_{(k+1)-1}\\ &= x_k\\ &= \sum_{j\in\{1,\ldots,n\}\setminus\{k\}} \alpha_j\cdot x_j\end{aligned}$$

So that

$$\begin{aligned}\Delta(y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n) &=\\ \Delta(z) &=\\ \Delta\left(z_1,\ldots,z_k,\sum_{j\in\{1,\ldots,n\}\setminus\{k\}}\alpha_j\cdot x_j,z_{k+2},\ldots,x_n\right) &=\\ \sum_{j\in\{1,\ldots,n\}\setminus\{k\}}\alpha_j\cdot\Delta(z_1,\ldots,z_k,x_j,z_{k+2},\ldots,x_n) & \qquad (11.145)\end{aligned}$$

Now for $j\in\{1,\ldots,n\}\setminus\{k\}$ consider the following cases:

***j* < *i*.** Then we have

$$\begin{aligned}(z_1,\ldots,z_k,x_j,z_{k+2},\ldots,x_n)_{j+1} &\underset{j\neq k\Rightarrow j+1\neq k+1}{=}\\ z_{j+1} &=\\ (y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_{j+1} &\underset{[\text{definition: } 11.258]\wedge 1<j+1\leqslant i}{=}\\ x_j &=\\ (z_1,\ldots,z_k,x_j,z_{k+2},\ldots,x_n)_{k+1} &\end{aligned}$$

So we have $j\neq k\Rightarrow j+1\neq k+1$ and

$$(z_1,\ldots,z_k,x_j,z_{k+2},\ldots,x_n)_{k+1}=(y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_{j+1}$$

which by [theorem: 11.248] proves that

$$\Delta(z_1,\ldots,z_k,x_j,z_{k+2},\ldots,x_n)=0 \qquad (11.146)$$

***i* < *j*.** Then we have

$$\begin{aligned}(z_1,\ldots,z_k,x_j,z_{k+2},\ldots,x_n)_j &\underset{k<i\Rightarrow k+1\leqslant i<j\Rightarrow k+1\neq j}{=}\\ z_j &=\\ (y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_j &\underset{[\text{definition: } 11.258]\wedge i<j\Rightarrow i+1\leqslant j}{=}\\ x_j &=\\ (z_1,\ldots,z_k,x_j,z_{k+2},\ldots,x_n)_{k+1}\end{aligned}$$

So we have $k<i\Rightarrow k+1\leqslant i<j\Rightarrow k+1\neq j$ and

$$(z_1,\ldots,z_k,x_j,z_{k+2},\ldots,x_n)_j=(z_1,\ldots,z_k,x_j,z_{k+2},\ldots,x_n)_{k+1}$$

which by [theorem: 11.248] proves that

$$\Delta(z_1,\ldots,z_k,x_j,z_{k+2},\ldots,x_n)=0 \tag{11.147}$$

***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: 11.261] that

$$(z_1,\ldots,z_k,x_i,z_{k+2},\ldots,x_n)\circ\left(k+1\underset{n}{\rightsquigarrow}i\right)=(y,x_1,\ldots,x_{k-1},x_{k+1},\ldots,x_n)$$

Hence

$$\begin{aligned}\Delta(y,x_1,\ldots,x_{k-1},x_{k+1},\ldots,x_n) &=\\ \Delta\left((z_1,\ldots,z_k,x_i,z_{k+2},\ldots,x_n)\circ\left(k+1\underset{n}{\rightsquigarrow}i\right)\right) &=\\ \left(\left(k+1\underset{n}{\rightsquigarrow}i\right)\Delta\right)(z_1,\ldots,z_k,x_i,z_{k+2},\ldots,x_n) &\underset{[\text{theorem: } 11.215]}{=}\\ (-1)^{|k+1-i|}\cdot\Delta(z_1,\ldots,z_k,x_i,z_{k+2},\ldots,x_n) &\underset{k<i\Rightarrow k+1\leqslant i}{=}\\ (-1)^{i-k-1}\cdot\Delta(z_1,\ldots,z_k,x_i,z_{k+2},\ldots,x_n)\end{aligned}$$

which by multiplying both sides by $(-1)^{i-k-1}$ and the fact that $(-1)^{i-k-1}\cdot(-1)^{i-k-1}=1$ we have

$$\Delta(z_1,\ldots,z_k,x_i,z_{k+2},\ldots,x_n)=(-1)^{i-k-1}\cdot\Delta(y,x_1,\ldots,x_{k-1},x_{k+1},\ldots,x_n) \tag{11.148}$$

So that

$$\sum_{j\in\{1,\ldots,n\}\setminus\{k\}}\alpha_j\cdot\Delta(z_1,\ldots,z_k,x_j,z_{k+2},\ldots,x_n) = A_1+A_2+A_2 \tag{11.149}$$

where

$$\begin{aligned}A_1 &= \sum_{j\in\{1,\ldots,i-1\}\setminus\{k\}}\alpha_j\cdot\Delta(z_1,\ldots,z_k,x_j,z_{k+2},\ldots,x_n)\\ &\underset{[\text{eq: } 11.146]}{=} 0\\ A_2 &= \sum_{j\in\{i\}\setminus\{k\}}\alpha_j\cdot\Delta(z_1,\ldots,z_k,x_j,z_{k+2},\ldots,x_n)\\ &\underset{i\neq k\Rightarrow\{i\}\setminus\{k\}=\{i\}}{=} \alpha_i\cdot\Delta(z_1,\ldots,z_k,x_i,z_{k+2},\ldots,x_n)\\ &\underset{[\text{eq: } 11.148]}{=} (\alpha_i\cdot(-1)^{i-k-1}\cdot\Delta(y,x_1,\ldots,x_{k-1},x_{k+1},\ldots,x_n))\\ A_3 &= \sum_{j\in\{i+1,\ldots,n\}\setminus\{k\}}\alpha_j\cdot\Delta(z_1,\ldots,z_k,x_j,z_{k+2},\ldots,x_n)\\ &\underset{[\text{eq: } 11.147]}{=} 0\end{aligned}$$

proving by [eqs 11.145, 11.149] and the above that that

$$\begin{aligned}\Delta(y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n) &=\\ \alpha_i\cdot(-1)^{i-k-1}\cdot\Delta(y,x_1,\ldots,x_{k-1},x_{k+1},\ldots,x_n)\end{aligned}\tag{11.150}$$

***i* < *k*.** Let

$$z=(y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)$$

then

$$\begin{aligned}z_k &= (y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_k\\ &\underset{[\text{definition: } 11.258]\wedge i<k\Rightarrow i+1\leqslant k}{=} x_k\\ &= \sum_{j\in\{1,\ldots,n\}\setminus\{k\}}\alpha_j\cdot x_j\end{aligned}$$

So that

$$\begin{aligned}\Delta(y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n) &=\\ \Delta(z) &=\\ \Delta\left(z_1,\ldots,z_{k-1},\sum_{j\in\{1,\ldots,n\}\setminus\{k\}}\alpha_j\cdot x_j,z_{k+1},\ldots,x_n\right) &=\\ \sum_{j\in\{1,\ldots,n\}\setminus\{k\}}\alpha_j\cdot\Delta(z_1,\ldots,z_{k-1},x_j,z_{k+1},\ldots,x_n)\end{aligned}\tag{11.151}$$

Now for $j\in\{1,\ldots,n\}\setminus\{k\}$ we have the following cases:

***j* < *i*.** Then we have:

$$\begin{aligned}(z_1,\ldots,z_{k-1},x_j,z_{k+1},\ldots,x_n)_{j+1} &\underset{j<i\Rightarrow j+1\leqslant i<k}{=}\\ z_{j+1} &=\\ (y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_{j+1} &\underset{[\text{definition: } 11.258]\wedge j<i\Rightarrow 1<j+1\leqslant i}{=}\\ x_{(j+1)-1} &=\\ x_j &\underset{j\neq k}{=}\\ (z_1,\ldots,z_{k-1},x_j,z_{k+1},\ldots,x_n)_k\end{aligned}$$

So we have $j<i\Rightarrow j+1\leqslant i<k\Rightarrow j\neq k$ and

$$(z_1,\ldots,z_{k-1},x_j,z_{k+1},\ldots,x_n)_{j+1}=(z_1,\ldots,z_{k-1},x_j,z_{k+1},\ldots,x_n)_k$$

which by [theorem: 11.248] proves that

$$\Delta(z_1,\ldots,z_{k-1},x_j,z_{k+1},\ldots,x_n)=0\tag{11.152}$$

***i* < *j*.** Then we have:

$$\begin{aligned}(z_1,\ldots,z_{k-1},x_j,z_{k+1},\ldots,x_n)_j &\underset{j\neq k}{=}\\ z_j &=\\ (y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n) &\underset{[\text{definition: } 11.258]\wedge i<j\Rightarrow i+1\leqslant i}{=}\\ x_j &=\\ (z_1,\ldots,z_{k-1},x_j,z_{k+1},\ldots,x_n)_k\end{aligned}$$

So we have $j\neq k$ and

$$(z_1,\ldots,z_{k-1},x_j,z_{k+1},\ldots,x_n)_j=(z_1,\ldots,z_{k-1},x_j,z_{k+1},\ldots,x_n)_k$$

which by [theorem: 11.248] proves that

$$\Delta(z_1,\ldots,z_{k-1},x_j,z_{k+1},\ldots,x_n)=0\tag{11.153}$$

$i=j$. 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: 11.261] that

$$(z_1,\ldots,z_{k-1},x_i,z_{k+1},\ldots,x_n)\circ\left(k\underset{n}{\rightsquigarrow}i+1\right)=(y,x_1,\ldots,x_{k-1},x_{k+1},\ldots,x_n)$$

hence

$$\begin{array}{rc}
\Delta((y,x_1,\ldots,x_{k-1},x_{k+1},\ldots,x_n)) & = \\
\Delta\left((z_1,\ldots,z_{k-1},x_i,z_{k+1},\ldots,x_n)\circ\left(k\underset{n}{\rightsquigarrow}i+1\right)\right) & = \\
\left(\left(k\underset{n}{\rightsquigarrow}i+1\right)\Delta\right)(z_1,\ldots,z_{k-1},x_i,z_{k+1},\ldots,x_n) & = \\
\operatorname{sign}\left(\left(k\underset{n}{\rightsquigarrow}i+1\right)\right)\cdot\Delta(z_1,\ldots,z_{k-1},x_i,z_{k+1},\ldots,x_n) & \underset{\text{[theorem: 11.215]}}{=} \\
(-1)^{|k-i-1|}\cdot\Delta(z_1,\ldots,z_{k-1},x_i,z_{k+1},\ldots,x_n) & \underset{i<k\Rightarrow i+1<k}{=} \\
(-1)^{k-i-1}\cdot\Delta(z_1,\ldots,z_{k-1},x_i,z_{k+1},\ldots,x_n) & \underset{1=(-1)^{2\cdot i-2\cdot k}}{=} \\
(-1)^{k-i-1+2\cdot i-2\cdot k}\cdot\Delta(z_1,\ldots,z_{k-1},x_i,z_{k+1},\ldots,x_n) & = \\
(-1)^{i-k-1}\cdot\Delta(z_1,\ldots,z_{k-1},x_i,z_{k+1},\ldots,x_n) &
\end{array}$$

which by multiplying both sides by $(-1)^{i-k-1}$ and the fact that $(-1)^{i-k-1}\cdot(-1)^{i-k-1}=1$ gives

$$\Delta(z_1,\ldots,z_{k-1},x_j,z_{k+1},\ldots,x_n)=(-1)^{i-k-1}\cdot\Delta((y,x_1,\ldots,x_{k-1},x_{k+1},\ldots,x_n)) \tag{11.154}$$

So that

$$\sum_{j\in\{1,\ldots,n\}\setminus\{k\}}\alpha_j\cdot\Delta(z_1,\ldots,z_{k-1},x_j,z_{k+1},\ldots,x_n) = A_1+A_2+A_2 \tag{11.155}$$

Where

$$\begin{array}{lcl}
A_1 & = & \displaystyle\sum_{j\in\{1,\ldots,i-1\}\setminus\{k\}}\alpha_j\cdot\Delta(z_1,\ldots,z_{k-1},x_j,z_{k+1},\ldots,x_n) \\
 & \underset{\text{[eq: 11.152]}}{=} & 0 \\
A_2 & = & \displaystyle\sum_{j\in\{i\}\setminus\{k\}}\alpha_j\cdot\Delta(z_1,\ldots,z_{k-1},x_j,z_{k+1},\ldots,x_n) \\
 & \underset{i\neq k}{=} & \alpha_i\cdot\Delta(z_1,\ldots,z_{k-1},x_i,z_{k+1},\ldots,x_n) \\
 & \underset{\text{[eq: 11.154]}}{=} & \alpha_i\cdot(-1)^{i-k-1}\cdot\Delta(y,x_1,\ldots,x_{k-1},x_{k+1},\ldots,x_n) \\
A_3 & = & \displaystyle\sum_{j\in\{i+1,\ldots,n\}\setminus\{k\}}\alpha_j\cdot\Delta(z_1,\ldots,z_{k-1},x_j,z_{k+1},\ldots,x_n) \\
 & \underset{\text{[eq: 11.154]}}{=} & 0
\end{array}$$

proving by [eqs: 11.151, 11.155] and the above that

$$\Delta(y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)=\alpha_i\cdot(-1)^{i-k-1}\cdot\Delta(y,x_1,\ldots,x_{k-1},x_{k+1},\ldots,x_n) \tag{11.156}$$

So

$$\begin{array}{rc}
A & \underset{\text{[eq: 11.142]}}{=} \\
\displaystyle\sum_{i\in\{1,\ldots,n\}\setminus\{k\}}((-1)^{i-1}\cdot\Delta(y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n))\cdot x_i & = \\
B_1+B_2 &
\end{array}$$

where

$$\begin{aligned}
B_1 &= \sum_{i\in\{1,\ldots,k-1\}} ((-1)^{i-1}\cdot\Delta(y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n))\cdot x_i\\
&\underset{[\text{eq: }11.156]}{=} \sum_{i\in\{1,\ldots,k-1\}} ((-1)^{i-1}\cdot\alpha_i\cdot(-1)^{i-k-1}\cdot\Delta(y,x_1,\ldots,x_{k-1},x_{k+1},\ldots,\\
&\quad x_n))\cdot x_i\\
&= \sum_{i\in\{1,\ldots,k-1\}} ((-1)^{-k}\cdot\alpha_i\cdot\Delta(y,x_1,\ldots,x_{k-1},x_{k+1},\ldots,x_n))\cdot x_i\\
&= (-1)^{-k}\cdot\Delta(y,x_1,\ldots,x_{k-1},x_{k+1},\ldots,x_n)\cdot\sum_{i\in\{1,\ldots,k-1\}}\alpha_i\cdot x_i\\
&\underset{(-1)^{2\cdot k}=1}{=} (-1)^k\cdot\Delta(y,x_1,\ldots,x_{k-1},x_{k+1},\ldots,x_n)\cdot\sum_{i\in\{1,\ldots,k-1\}}\alpha_i\cdot x_i\\
B_2 &= \sum_{i\in\{1+1,\ldots,n\}} ((-1)^{i-1}\cdot\Delta(y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n))\cdot x_i\\
&\underset{[\text{eq: }11.148]}{=} \sum_{i\in\{1+1,\ldots,n\}} ((-1)^{i-1}\cdot\alpha_i\cdot(-1)^{i-k-1}\cdot\Delta(y,x_1,\ldots,x_{k-1},x_{k+1},\ldots,\\
&\quad x_n))\cdot x_i\\
&= \sum_{i\in\{1+1,\ldots,n\}} ((-1)^{-k}\cdot\alpha_i\cdot\Delta(y,x_1,\ldots,x_{k-1},x_{k+1},\ldots,x_n))\cdot x_i\\
&= (-1)^{-k}\cdot\Delta(y,x_1,\ldots,x_{k-1},x_{k+1},\ldots,x_n)\cdot\sum_{i\in\{k+1,\ldots,n\}}\alpha_i\cdot x_i\\
&\underset{(-1)^{2\cdot k}=1}{=} (-1)^k\cdot\Delta(y,x_1,\ldots,x_{k-1},x_{k+1},\ldots,x_n)\cdot\sum_{i\in\{k+1,\ldots,n\}}\alpha_i\cdot x_i
\end{aligned}$$

So that 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

$$\begin{aligned}
A &= (-1)^k\cdot\Delta(y,x_1,\ldots,x_{k-1},x_{k+1},\ldots,x_n)\\
&\underset{[\text{eq: }11.144]}{=} (-1)\cdot B
\end{aligned}$$

So that $A+B=0$ proving by [eq: 11.141] 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: 11.139] proves that

$$\underline{\Delta}(y,(x_1,\ldots,x_n))=0\cdot y\underset{[\text{eq: }11.139]}{=}\Delta(x_1,\ldots,x_n)\cdot y$$

$\{x_i\}_{i\in\{1,\ldots,n\}}$ is linear independent. Then by [corollary: 11.133] we have that

$$\{x_i|i\in\{1,\ldots,n\}\}\text{ is a basis of }X$$

So if $y\in X$ then there exists by [theorem: 11.123] a $\{\alpha_i\}_{i\in\{1,\ldots,n\}}\subseteq F$ such that

$$y=\sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot x_i \tag{11.157}$$

So that

$$\underline{\Delta}(y,(x_1,\ldots,x_n))\underset{\text{def}}{=}\sum_{i\in\{1,\ldots,n\}}(-1)^{i-1}\cdot\Delta(y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)\cdot x_i \tag{11.158}$$

Let $i\in\{1,\ldots,n\}$ then we have as Δ is n-linear that

$$\Delta(y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)=\sum_{j\in\{1,\ldots,n\}}\alpha_j\cdot\Delta(x_j,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n) \tag{11.159}$$

For i, j we have the following cases:

$\boldsymbol{i<j}$**.** Then

$$\begin{array}{lcl} (x_j,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_j & \underset{\text{[theorem: 11.258]}\wedge i<j\Rightarrow i+1\leqslant j}{=} & x_j \\ & \underset{\text{[theorem: 11.258]}}{=} & (x_j,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_1 \end{array}$$

which as $1\leqslant i<j\Rightarrow j\neq 1$ proves by [theorem: 11.248] that

$$\Delta(x_j,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)=0 \tag{11.160}$$

$\boldsymbol{j<i}$**.** Then

$$\begin{array}{lcl} (x_j,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_{j+1} & \underset{\text{[theorem: 11.258]}\wedge j<i\Rightarrow 1<j+1\leqslant i}{=} & x_{(j+1)-1} \\ & \underset{\text{[theorem: 11.258]}}{=} & x_1 \end{array}$$

which as $1<j+1$ proves by [theorem: 11.248] that

$$\Delta(x_j,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)=0 \tag{11.161}$$

$\boldsymbol{i=j}$**.** Let

$$A=(x_j,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)\circ\left(1\underset{n}{\rightsquigarrow}j\right) \tag{11.162}$$

then we have the following cases to consider:

$\boldsymbol{j=1}$**.** Then $\left(1\underset{n}{\rightsquigarrow}j\right)=\left(1\underset{n}{\rightsquigarrow}1\right)\underset{\text{[definition: 11.212]}}{=}\mathrm{Id}_{\{1,\ldots,n\}}$ so that $\forall k\in\{1,\ldots,n\}$

$$\begin{array}{lcl} A_k & = & (x_j,x_1,\ldots,x_{j-1},x_{j+1},\ldots x_n)_k \\ & \underset{\text{[theorem: 11.258]}}{=} & \begin{cases} x_j \text{ if } k=1 \\ x_{k-1} \text{ if } k\in\{2,\ldots,j\} \\ x_k \text{ if } k\in\{j+1,\ldots,n\} \end{cases} \\ & \underset{j=1\wedge\{2,\ldots,j\}=\varnothing}{=} & \begin{cases} x_1 \text{ if } k=1 \\ x_k \text{ if } k\in\{2,\ldots,n\} \end{cases} \\ & = & x_k \end{array}$$

so that

$$A=(x_1,\ldots,x_n)$$

$\boldsymbol{j\in\{2,\ldots,n\}}$**.** Then $1<j$ we have the following cases to consider (helped by [definition: 11.212]) for $k\in\{1,\ldots n\}$:

$\boldsymbol{1\leqslant k<j}$**.** Then

$$\begin{array}{lcl} A_k & = & (x_j,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_{(1\underset{n}{\rightsquigarrow}j)(k)} \\ & \underset{\text{[definition: 11.212]}}{=} & (x_j,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_{k+1} \\ & \underset{\text{[theorem: 11.258]}\wedge k<j\Rightarrow 1<k+1\leqslant j}{=} & x_{(k+1)-1} \\ & = & x_k \end{array}$$

$\boldsymbol{k=j}$**.** Then

$$\begin{array}{lcl} A_k & = & (x_j,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_{(1\underset{n}{\rightsquigarrow}j)(k)} \\ & \underset{\text{[definition: 11.212]}}{=} & (x_j,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_1 \\ & \underset{\text{[theorem: 11.258]}}{=} & x_j \\ & \underset{k=j}{=} & x_k \end{array}$$

$\boldsymbol{j<k\leqslant n}$**.** Then

$$\begin{array}{lcl} A_k & = & (x_j,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_{(1\underset{n}{\rightsquigarrow}j)(k)} \\ & \underset{\text{[definition: 11.212]}}{=} & (x_j,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_k \\ & \underset{\text{[theorem: 11.258]}\wedge j<k\Rightarrow j+1\leqslant k}{=} & x_k \end{array}$$

so that

$$A = (x_1, \ldots, x_n)$$

So in all cases $A = (x_1, \ldots, x_n)$ proving that

$$(x_j, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \circ \left(1 \underset{n}{\leadsto} j\right) = (x_1, \ldots, x_n) \tag{11.163}$$

So we have that

$$\begin{aligned}
\Delta(x_1, \ldots, x_n) &= \Delta\left((x_j, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \circ \left(1 \underset{n}{\leadsto} j\right)\right) \\
&= \left(\left(1 \underset{n}{\leadsto} j\right)\Delta\right)(x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \\
&= \operatorname{sign}\left(\left(1 \underset{n}{\leadsto} j\right)\right) \cdot \Delta(x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \\
&\underset{\text{[theorem: 11.215]}}{=} (-1)^{j-1} \cdot \Delta(x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \\
&\underset{i=j}{=} (-1)^{i-1} \cdot \Delta(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)
\end{aligned}$$

proving that

$$\Delta(x_j, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) = (-1)^{j-1} \cdot \Delta(x_1, \ldots, x_n) \tag{11.164}$$

Now

$$\sum_{j \in \{1, \ldots, n\}} \alpha_j \cdot \Delta(x_j, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) = C_1 + C_2 + C_3$$

where

$$\begin{aligned}
C_1 &= \sum_{j \in \{1, \ldots, i-1\}} \alpha_j \cdot \Delta(x_j, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \\
&\underset{\text{[eq: 11.161]}}{=} 0 \\
C_2 &= \sum_{j \in \{i\}} \alpha_j \cdot \Delta(x_j, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \\
&= \alpha_i \cdot \Delta(x_i, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \\
&\underset{\text{[eq: 11.165]}}{=} (-1)^{j-1} \cdot \alpha_i \cdot \Delta(x_1, \ldots, x_n) \\
C_3 &= \sum_{j \in \{i+1, \ldots, n\}} \alpha_j \cdot \Delta(x_j, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \\
&\underset{\text{[eq: 11.160]}}{=} 0
\end{aligned}$$

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: 11.159] proves

$$\Delta(y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \cdot x_i = \alpha_i \cdot (-1)^{i-1} \cdot \Delta(x_1, \ldots, x_n) \tag{11.165}$$

So

$$\begin{aligned}
\underline{\Delta}(y, (x_1, \ldots, x_n)) &\underset{\text{def}}{=} \sum_{i \in \{1, \ldots, n\}} (-1)^{i-1} \cdot \Delta(y, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \cdot x_i \\
&\underset{\text{[eq: 11.165]}}{=} \sum_{i \in \{1, \ldots, n\}} ((-1)^{i-1} \cdot \alpha_i \cdot (-1)^{i-1} \cdot \Delta(x_1, \ldots, x_n)) \cdot x_i \\
&= \Delta(x_1, \ldots, x_n) \cdot \sum_{i \in \{1, \ldots, n\}} \alpha_i \cdot x_i \\
&\underset{\text{[eq: 11.157}}{=} \Delta(x_1, \ldots, x_n) \cdot y
\end{aligned}$$

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

$$\underline{\Delta}(y,(x_1,\ldots,x_n))=\Delta(x_1,\ldots,x_n)\cdot y \qquad \square$$

We show no how given a determinant function and a linear mapping we can create a new determinant function.

Definition 11.263. *Let $n\in\mathbb{N}$, X a vector space over a field F such that* $\mathrm{dom}(X)=n$*,*
$\Delta\in\mathrm{Hom}(X^n;F)$ *a non trivial determinant function and* $L\in\mathrm{Hom}(X,X)$ *[a linear transformation] then we define*

$$\Delta_L\colon X^n\rightarrow F \text{ by } \Delta_L(x_1,\ldots,x_n)=\Delta(L(x_1),\ldots,L(x_1))$$

It turns out that Δ_L is also a determinant function.

Theorem 11.264. *Let $n\in\mathbb{N}$, X a vector space over a field F with characteristic zero such that* $\mathrm{dom}(X)=n$*, $\Delta\in\mathrm{Hom}(X^n;F)$ a non trivial determinant function and $L\in\mathrm{Hom}(X,X)$ 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:

$$\begin{aligned}
\Delta_L(x_1,\ldots,x_{i-1},u+v,x_{i+1},\ldots,x_n) &=\\
\Delta(L(x_1),\ldots,L(x_{i-1}),L(u+v),L(x_{i+1}),\ldots,L(x_n)) &=\\
\Delta(L(x_1),\ldots,L(x_{i-1}),L(u)+L(v),L(x_{i+1}),\ldots,L(x_n)) &=\\
\Delta(L(x_1),\ldots,L(x_{i-1}),L(u),L(x_{i+1}),\ldots,L(x_n))+\Delta(L(x_1),\ldots,L(x_{i-1}),L(v),L(x_{i+1}),\ldots,L(x_n)) &=\\
\Delta_L(x_1,\ldots,x_{i-1},u,x_{i+1},\ldots,x_n)+\Delta_L(x_1,\ldots,x_{i-1},v,x_{i+1},\ldots,x_n)&
\end{aligned}$$

and

$$\begin{aligned}
\Delta_L(x_1,\ldots,x_{i-1},\alpha\cdot u,x_{i+1},\ldots,x_n) &= L(L(x_1),\ldots,L(x_{i-1}),L(\alpha\cdot u),L(x_{i+1}),\ldots,L(x_n))\\
&= L(L(x_1),\ldots,L(x_{i-1}),\alpha\cdot L(u),L(x_{i+1}),\ldots,L(x_n))\\
&= \alpha\cdot L(L(x_1),\ldots,L(x_{i-1}),L(u),L(x_{i+1}),\ldots,L(x_n))\\
&= \alpha\cdot L(x_1,\ldots,x_{i-1},u,x_{i+1},\ldots,x_n)
\end{aligned}$$

proving that

$$\Delta_L\in\mathrm{Hom}(X^n,F) \text{ or } \Delta_L \text{ is } n\text{-linear}$$

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

$$\begin{aligned}
\Delta_L(x_1,\ldots,x_n) &= \Delta(L(x_1),\ldots,L(x_n))\\
&\underset{\text{[theorem: 11.248]}}{=} 0
\end{aligned}$$

which by [theorem: 11.248] proves that

$$\Delta_L \text{ is skew-symmetric} \qquad \square$$

Note 11.265. It is not true that Δ_L is a non trivial determinant function if D is non trivial determinant function. For example if $L=C_0\colon X\rightarrow X$ defined by $C_0(x)=0$ then if $(x_1,\ldots,x_n)\in X^n$ 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 11.266. *Let $n\in\mathbb{N}$, X a vector space over a field F such that $\dim(X)=n$ and $L\in\mathrm{Hom}(X,X)$ then there exist a $\alpha\in F$ such that for **every** non trivial determinant function $\Delta\in\mathrm{Hom}(X^n;F)$ we have*

$$\Delta_L=\alpha\cdot\Delta$$

Proof. Using [theorem: 11.255] there exist a non trivial determinant function Δ. 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: 11.264] Δ_L is also a determinant function we have by [theorem: 11.257] that there exists a $\alpha \in F$ such that

$$\Delta_L = \alpha \cdot \Delta \tag{11.166}$$

So in all cases there exists a $\alpha \in F$ such that $\Delta_L = \alpha \cdot \Delta$ proving existence. Assume that Δ' is another non trivial determent function then by [theorem: 11.257] there exist a $\lambda \in F$ such that

$$\Delta' = \lambda \cdot \Delta.$$

Let $(x_1, \ldots, x_n) \in X^n$ then

$$\begin{aligned} \Delta'_L(x_1, \ldots, x_n) &= \Delta'(L(x_1), \ldots, L(x_n)) \\ &= \lambda \cdot \Delta(L(x_1), \ldots, L(x_n)) \\ &= \lambda \cdot \Delta_L(x_1, \ldots, x_n) \end{aligned}$$

So

$$\begin{aligned} \Delta'_L &= \lambda \cdot \Delta_L \\ &\underset{[\text{eq: } 11.166]}{=} \lambda \cdot (\alpha \cdot \Delta) \\ &= \alpha \cdot (\lambda \cdot \Delta) \\ &= \alpha \cdot \Delta' \end{aligned}$$

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 11.267. *Let $n \in \mathbb{N}$, X a vector space over a field F with characteristic zero such that $\dim(X) = n$ and $L \in \operatorname{Hom}(X, X)$ then we define*

$\det: \operatorname{Hom}(X^n; Y) \rightarrow F$ *by* $\det(L) = \alpha$ *where for every determinant function Δ we have $\Delta_L = \alpha \cdot \Delta$*

In other words $\det(L)$ is the scalar such that for every non trivial determinant function Δ we have

$$\Delta_L = \det(L) \cdot \Delta$$

Theorem 11.268. *Let $n \in \mathbb{N}$, X a vector space over a field F with characteristic zero such that $\dim(X) = n$, $\{e_i\}_{i \in \{1, \ldots, n\}} \subseteq X$ defining the basis $\{e_i | i \in \{1,...,n\}\}$ for X and $L \in \operatorname{Hom}(X, X)$ then*

$$\det(L) = \Delta_L(e_1, \ldots, e_n) = \Delta(L(e_1), \ldots, L(e_n))$$

where $\Delta: X^n \rightarrow F$ is the determinant function such that $\Delta(e_1, \ldots, e_n)$ [see 11.255 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 11.269. 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

$$\lambda \cdot \operatorname{Id}_X: X \rightarrow X \text{ defined by } (\lambda \cdot \operatorname{Id}_X)(x) = \lambda \cdot x$$

then we have

$$\det(\lambda \cdot \operatorname{Id}_X) = \lambda^n$$

or in the case of $\lambda = 1$ that

$$\det(\mathrm{Id}_X) = 1$$

Proof. Let Δ be a non trivial determinant function then

$$\begin{array}{rcl} \Delta_{\lambda\cdot\mathrm{Id}_X}(x_1,\ldots,x_n) & = & \Delta(\lambda\cdot\mathrm{Id}_X(x_1),\ldots,\lambda\cdot\mathrm{Id}_X(x_n)) \\ & = & \Delta(\lambda\cdot x_1,\ldots,\lambda\cdot x_n) \\ & \underset{\text{[theorem: 11.229]}}{=} & \lambda^n\cdot D(x_1,\ldots,x_n) \end{array}$$

proving that $\Delta_{\lambda\cdot\mathrm{Id}_X} = \lambda^n\cdot\Delta$, hence

$$\det(\lambda\cdot\mathrm{Id}_x) = \lambda^n$$

□

Example 11.270. 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 11.271. *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. Let $\{e_i|i\in\{1,\ldots,n\}\}$ be a basis for X

⇒. As L is a linear isomorphism L is injective, it follows from [theorem: 11.175] that $L(\{e_i|i\in\{1,\ldots,n\}\}) = \{L(e_i)|i\in\{1,\ldots,n\}\}$ is linear independent. Using [theorem: 11.133] it follows then that

$$\{L(e_i)|i\in\{1,\ldots,n\}\} \text{ is a basis for } X$$

Using [theorem: 11.255] there exist then 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: 11.175] 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: 11.248] 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 $\{e_i|i\in\{1,\ldots,n\}\}$ is a basis we have that $\{\alpha_i\}_{i\in\{1,\ldots,n\}}\subseteq F$ 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: 11.173] proves that

$$L \text{ is injective}$$

As $\dim(X)$ is finite and $L: X\rightarrow X$ is injective it follows from [theorem: 11.177] 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 11.272. *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 $\{L_i\}_{i\in\{1,\ldots,m\}}\subseteq\operatorname{Hom}(X,X)$ 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

$$\begin{aligned}\det\left(\prod_{i=1}^{m+1} L_i\right) &= \det\left(\prod_{i=1}^{m} L_i\circ L_{m+1}\right)\\ &\underset{[\text{theorem: }11.271]}{=} \det\left(\prod_{i=1}^{m} L_i\right)\circ\det(L_{m+1})\\ &\underset{m\in S}{=} \prod_{i\in\{1,\ldots,m\}}\det(L_i)\cdot\det(L_{m+1})\\ &= \prod_{i\in\{1,\ldots,m+1\}}\det(L_i)\end{aligned}$$

proving that $m+1\in S$.

□

Definition 11.273. *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*

$$\overline{\Delta L}: X^n\rightarrow\operatorname{Hom}(X,X)\textit{ where }\overline{\Delta L}(x_1,\ldots,x_n): X\rightarrow X\textit{ is defined by}$$

$$\overline{\Delta L}(x_1,\ldots,x_n)(y)=\sum_{i\in\{1,\ldots,n\}}(-1)^{i-1}\cdot\Delta(y,L(x_1),\ldots,L(x_{i-1}),L(x_{i+1}),\ldots,L(x_n))\cdot x_i$$

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

$$\begin{aligned}&\overline{\Delta L}(x_1,\ldots,x_n)(u+v) =\\ &\sum_{i\in\{1,\ldots,n\}}(-1)^{i-1}\cdot\Delta(u+v,L(x_1),\ldots,L(x_{i-1}),L(x_{i+1}),\ldots,L(x_n))\cdot x_i =\\ &\sum_{i\in\{1,\ldots,n\}}(-1)^{i-1}\cdot(\Delta(u,L(x_1),\ldots,L(x_{i-1}),L(x_{i+1}),\ldots,L(x_n))+\Delta(v,L(x_1),\ldots,\\ &L(x_{i-1}),L(x_{i+1}),\ldots,L(x_n))) \underset{[\text{theorem: }11.36]}{=}\\ &\sum_{i\in\{1,\ldots,n\}}(-1)^{i-1}\cdot\Delta(u,L(x_1),\ldots,L(x_{i-1}),L(x_{i+1}),\ldots,L(x_n))\cdot x_i+\sum_{i\in\{1,\ldots,n\}}(-1)^{i}\cdot\\ &\Delta(v,L(x_1),\ldots,L(x_{i-1}),L(x_{i+1}),\ldots,L(x_n))\cdot x_i =\\ &\overline{\Delta L}(x_1,\ldots,x_n)(u)+\overline{\Delta L}(x_1,\ldots,x_n)(v)\end{aligned}$$

and

$$\begin{aligned}\overline{\Delta L}(x_1,\ldots,x_n)(\alpha\cdot u) &= \sum_{i\in\{1,\ldots,n\}}(-1)^{i-1}\cdot\Delta(\alpha\cdot u,L(x_1),\ldots,L(x_{i-1}),L(x_{i+1}),\ldots,L(x_n))\cdot x_i\\ &= \sum_{i\in\{1,\ldots,n\}}(-1)^{i-1}\cdot\alpha\cdot\Delta(u,L(x_1),\ldots,L(x_{i-1}),L(x_{i+1}),\ldots,L(x_n))\cdot x_i\\ &= \alpha\cdot\sum_{i\in\{1,\ldots,n\}}(-1)^{i}\cdot\Delta(u,L(x_1),\ldots,L(x_{i-1}),L(x_{i+1}),\ldots,L(x_n))\cdot x_i\\ &= \alpha\cdot\overline{\Delta L}(x_1,\ldots,x_n)(u)\end{aligned}$$

□

Lemma 11.274. *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

$$A=(t,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)\circ\left(i\underset{n}{\rightsquigarrow}j+1\right)\tag{11.167}$$

then we have as $j+1\leqslant i$ the following cases to consider for i:

$\boldsymbol{j+1=i}$**.** Then by [definition: 11.212] we have that $\left(i\underset{n}{\rightsquigarrow}j+1\right)=\mathrm{Id}_{\{1,\ldots,n\}}$ so that

$$A=(t,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)\tag{11.168}$$

Now for $k\in\{1,\ldots,n\}$ we have either (motivated by [definition: 11.258] the following possible cases

$\boldsymbol{k=1}$**.** Then

$$\begin{aligned}A_k &\underset{\text{[eq: 11.138]}}{=} (t,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_k\\ &\underset{\text{[definition: 11.258]}\wedge k=1}{=} t\\ &\underset{\text{[definition: 11.258]}\wedge k=1}{=} (t,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_k\end{aligned}$$

$\boldsymbol{2\leqslant k<j}$**.** Then

$$\begin{aligned}A_k &\underset{\text{[eq: 11.138]}}{=} (t,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_k\\ &\underset{\text{[definition: 11.258]}\wedge 2\leqslant k<j}{=} x_{k-1}\\ &\underset{\text{[definition: 11.258]}\wedge 2<k<j=i-1<i}{=} (t,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_k\end{aligned}$$

$\boldsymbol{k=j+1}$**.** Then

$$\begin{aligned}A_k &\underset{\text{[eq: 11.138]}}{=} (t,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_k\\ &\underset{\text{[definition: 11.258]}j+1=k}{=} x_k\\ &\underset{\text{[definition: 11.258]}i=j+1=k}{=} (t,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_k\end{aligned}$$

$\boldsymbol{j+1<k}$**.** Then

$$\begin{aligned}
A_k &\underset{\text{[eq: 11.138]}}{=} (t,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_k\\
&\underset{\text{[definition: 11.258]}j+1<k}{=} x_k\\
&\underset{\text{[definition: 11.258]}i=j+1<k\Rightarrow i+1\leqslant k}{=} (t,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_k
\end{aligned}$$

$\boldsymbol{j+1<i}$**.** Then we have motivated by [definition: 11.212] the following possible cases for k:

$\boldsymbol{k=1}$**.** Then

$$\begin{aligned}
A_k &= (t,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_{(i\underset{n}{\rightsquigarrow}j+1)(k)}\\
&\underset{\text{[definition: 11.212]}\wedge k=1<j+1}{=} (t,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_k\\
&\underset{\text{[definition: 11.258]}\wedge k=1}{=} t\\
&\underset{\text{[definition: 11.258]}\wedge k=1}{=} (t,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_k
\end{aligned}$$

$\boldsymbol{k\neq 1\wedge k<j+1}$**.** Then

$$\begin{aligned}
A_k &= (t,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_{(i\underset{n}{\rightsquigarrow}j+1)(k)}\\
&\underset{\text{[definition: 11.212]}\wedge k=1<j+1}{=} (t,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_k\\
&\underset{\text{[definition: 11.258]}\wedge 1<k<j+1\Rightarrow 1<k\leqslant j}{=} x_{k-1}\\
&\underset{\text{[definition: 11.258]}\wedge 1<k<j+1=i}{=} (t,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_k
\end{aligned}$$

$\boldsymbol{k\neq 1\wedge k=j+1}$**.** Then

$$\begin{aligned}
A_k &= (t,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_{(i\underset{n}{\rightsquigarrow}j+1)(k)}\\
&\underset{\text{[definition: 11.212]}\wedge k=j+1}{=} (t,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_i\\
&\underset{\text{[definition: 11.258]}\wedge j+1=k}{=} x_i\\
&\underset{x_i=x_j}{=} x_j\\
&= x_{(j+1)-1}\\
&\underset{\text{[definition: 11.258]}\wedge 1<j+1<i}{=} (t,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_{j+1}\\
&\underset{k=j+1}{=} (t,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_k
\end{aligned}$$

$\boldsymbol{k\neq 1\wedge j+1<k\leqslant i}$**.** Then

$$\begin{aligned}
A_k &= (t,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_{(i\underset{n}{\rightsquigarrow}j+1)(k)}\\
&\underset{\text{[definition: 11.212]}\wedge j+1<k\leqslant i}{=} (t,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_{k-1}\\
&\underset{\text{[definition: 11.258]}\wedge 1<k\leqslant i}{=} (t,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_k
\end{aligned}$$

$\boldsymbol{k\neq 1\wedge i<k}$**.** Then

$$\begin{aligned}
A_k &= (t,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_{(i\underset{n}{\rightsquigarrow}j+1)(k)}\\
&\underset{\text{[definition: 11.212]}\wedge i<k\leqslant n}{=} (t,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_k\\
&\underset{\text{[definition: 11.258]}\wedge j+1<i<k}{=} x_k\\
&\underset{\text{[definition: 11.258]}\wedge i<k\Rightarrow i+1\leqslant k}{=} (t,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_k
\end{aligned}$$

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

$$B = (t, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \circ \left(i+1 \underset{n}{\rightsquigarrow} j\right) \tag{11.169}$$

then we have as $i + 1 \leqslant j$ the following cases to consider for j:

$\boldsymbol{i+1=j.}$ Then by [definition: 11.212] we have that $\left(i \underset{n}{\rightsquigarrow} j+1\right) = \mathrm{Id}_{\{1,\ldots,n\}}$ so that

$$B = (t, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \tag{11.170}$$

Now for $k \in \{1,\ldots,n\}$ we have either (motivated by [definition: 11.258] following cases:

$\boldsymbol{k=1.}$ Then

$$\begin{aligned} B_k &\underset{\text{[eq: 11.170]}}{=} (t, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_k \\ &\underset{\text{[definition: 11.258]}\wedge k=1}{=} t \\ &\underset{\text{[definition: 11.258]}\wedge k=1}{=} (t, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_k \end{aligned}$$

$\boldsymbol{1 < k \leqslant i.}$ Then

$$\begin{aligned} B_k &\underset{\text{[eq: 11.170]}}{=} (t, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_k \\ &\underset{\text{[definition: 11.258]}\wedge 1<k\leqslant i<i+1=j}{=} x_{k-1} \\ &\underset{\text{[definition: 11.258]}\wedge 1<k\leqslant i}{=} (t, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_k \end{aligned}$$

$\boldsymbol{k=i+1.}$ Then

$$\begin{aligned} B_k &\underset{\text{[eq: 11.170]}}{=} (t, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_k \\ &\underset{\text{[definition: 11.258]}\wedge 1<k=i+1=j}{=} x_{k-1} \\ &\underset{\text{[definition: 11.258]}\wedge k=i+1}{=} (t, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_k \end{aligned}$$

$\boldsymbol{i+1 < k.}$ Then

$$\begin{aligned} B_k &\underset{\text{[eq: 11.170]}}{=} (t, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_k \\ &\underset{\text{[definition: 11.258]}\wedge j=i+1<k\Rightarrow j+1\leqslant k}{=} x_k \\ &\underset{\text{[definition: 11.258]}\wedge i+1<k}{=} (t, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_k \end{aligned}$$

$\boldsymbol{i+1 < j.}$ Then we have motivated by [definition: 11.212] the following possible cases for k:

$\boldsymbol{k=1.}$ Then

$$\begin{aligned} B_k &= (t, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_{(i+1 \underset{n}{\rightsquigarrow} j)(k)} \\ &\underset{\text{[theorem: 11.212]}\wedge 1=k<i+1}{=} (t, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_k \\ &\underset{\text{[definition: 11.258]}\wedge k=1}{=} t \\ &\underset{\text{[definition: 11.258]}\wedge k=1}{=} (t, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_k \end{aligned}$$

$\boldsymbol{k \neq 1 \wedge k < i+1.}$ Then

$$\begin{aligned} B_k &= (t, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_{(i+1 \underset{n}{\rightsquigarrow} j)(k)} \\ &\underset{\text{[definition: 11.212]}\wedge 1<k<i+1}{=} (t, x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)_k \\ &\underset{\text{[definition: 11.258]}\wedge 1<k<i+1=j}{=} x_{k-1} \\ &\underset{\text{[definition: 11.258]}\wedge 1<k<i+1\Rightarrow 1<k\leftarrow i}{=} (t, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)_k \end{aligned}$$

$\boldsymbol{k\neq 1\wedge i+1\leqslant k<j.}$ Then

$$\begin{aligned}
B_k &= (t,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_{(i+1 \underset{n}{\rightsquigarrow} j)(k)}\\
&\underset{[\text{definition: } 11.212]\wedge i+1\leqslant k<j}{=} (t,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_{k+1}\\
&\underset{[\text{definition: } 11.258]\wedge 1<k<j}{=} x_{(k+1)-1}\\
&= x_k\\
&\underset{[\text{definition: } 11.258]\wedge i+1\leqslant k}{=} (t,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_k
\end{aligned}$$

$\boldsymbol{k\neq 1\wedge k=j.}$ Then

$$\begin{aligned}
B_k &= (t,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_{(i+1 \underset{n}{\rightsquigarrow} j)(k)}\\
&\underset{[\text{definition: } 11.212]\wedge k=j}{=} (t,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_{i+1}\\
&\underset{[\text{definition: } 11.258]\wedge 1<i+1<j}{=} x_{(i+1)-1}\\
&= x_i\\
&\underset{x_i=x_j}{=} x_j\\
&\underset{j=k}{=} x_k\\
&\underset{[\text{definition: } 11.258]\wedge i<i+1<j=k}{=} (t,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_k
\end{aligned}$$

$\boldsymbol{k\neq 1\wedge j<k.}$ Then

$$\begin{aligned}
B_k &= (t,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_{(i+1 \underset{n}{\rightsquigarrow} j)(k)}\\
&\underset{[\text{definition: } 11.212]\wedge j<k}{=} (t,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)_k\\
&\underset{[\text{definition: } 11.258]\wedge j<k\rightarrow j+1\leqslant k}{=} x_k\\
&\underset{[\text{definition: } 11.258]\wedge i+1<j<k}{=} (t,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)_k
\end{aligned}$$

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

$$(t,x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)\circ\left(i+1 \underset{n}{\rightsquigarrow} j\right)=(t,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n) \qquad \square$$

The above lemma will be used to prove the following theorem that will be used to define the adjoint of a linear transformation.

Theorem 11.275. *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*

$$\overline{\Delta L}\in\operatorname{Hom}(X^n;\operatorname{Hom}(X,X)) \text{ and } \overline{\Delta L} \text{ is skew-symmetric}$$

Proof. Given $x\in X^n$, $t\in X$ then we have by definition

$$\begin{aligned}
\overline{\Delta L}(x)(t) &= \sum_{j\in\{1,\ldots,n\}}(-1)^{j-1}\cdot\Delta(t,L(x_1),\ldots,L(x_{j-1}),L(x_{j+1}),\ldots,L(x_n))\cdot x_j\\
&= \sum_{j\in\{1,\ldots,n\}}(-1)^{j-1}\cdot A_j(t,x)
\end{aligned}$$

where

$$A_j(t,x)=\Delta(t,L(x_1),\ldots,L(x_{j-1}),L(x_{j+1}),\ldots,L(x_n))\cdot x_j$$

Next we prove that $\overline{\Delta L}$ is n-linear. So let $(x_1,\ldots,x_n)\in X^n$, $u,v\in X$, $\alpha\in F$, $i\in\{1,\ldots,n\}$ and define

$$\begin{aligned} y &= (x_1,\ldots,x_{i-1},u+\alpha\cdot v,x_{i+1},\ldots,x_n) && (11.171)\\ r &= (x_1,\ldots,x_{i-1},u,x_{i+1},\ldots,x_n) && (11.172)\\ s &= (x_1,\ldots,x_{i-1},v,x_{i+1},\ldots,x_n) && (11.173)\end{aligned}$$

Then for $j\in\{1,\ldots,n\}$ we have either:

***j* = *i*.** Then for $k\in\{2,\ldots,n\}$ we have either:

***k* ⩽ *i*.** Then $k-1<i\Rightarrow k-1\neq i$ and

$$\begin{aligned}(t,L(y_1),\ldots,L(y_{i-1}),L(y_{i+1}),\ldots,L(y_n))_k &\underset{\text{[definition: 11.258]}}{=} L(y_{k-1})\\ &\underset{k-1\neq i\wedge\text{[eq: 11.172]}}{=} L(x_{k-1})\end{aligned}$$

***i* + 1 ⩽ *k* ⩽ *n*.** Then $i<k\Rightarrow i\neq k$ and

$$\begin{aligned}(t,L(y_1),\ldots,L(y_{i-1}),L(y_{i+1}),\ldots,L(y_n))_k &\underset{\text{[definition: 11.258]}}{=} L(y_k)\\ &\underset{k\neq i\wedge\text{[eq: 11.172]}}{=} L(x_k)\end{aligned}$$

proving

$$(t,L(y_1),\ldots,L(y_{i-1}),L(y_{i+1}),\ldots,L(y_n))=(t,L(x_1),\ldots,L(x_{i-1}),L(x_{i+1}),\ldots,L(x_n)) \quad (11.174)$$

hence

$$\begin{aligned} A_j(t,y) &\underset{i=j}{=}\\ A_i(t,y) &=\\ (t,L(y_1),\ldots,L(y_{i-1}),L(y_{i+1}),\ldots,L(y_n))\cdot y_i &\underset{\text{[eq: 11.174]}}{=}\\ (t,L(x_1),\ldots,L(x_{i-1}),L(x_{i+1}),\ldots,L(x_n))\cdot(u+\alpha\cdot v) &=\\ (t,L(x_1),\ldots,L(x_{i-1}),L(x_{i+1}),\ldots,L(x_n))\cdot u+\alpha\cdot(t,L(x_1),\ldots,L(x_{i-1}),L(x_{i+1}),\ldots,\\ L(x_n))\cdot v &=\\ A_i(t,r)+\alpha\cdot A_i(t,s) &\underset{i=j}{=}\\ A_j(t,r)+\alpha\cdot A_j(t,s)\end{aligned}$$

Hence we have

$$A_j(t,y)=A_j(t,r)+A_j(t,s)$$

***j* < *i*.** Then $j+1\leqslant i$ so that by [definition: 11.258]

$$(L(y_1),\ldots,L(y_{j-1}),L(y_{j+1}),\ldots,L(y_n))_i=L(y_i)\underset{\text{[eq: 11.171]}}{=}L(u+\alpha\cdot v)=L(u)+\alpha\cdot L(v)$$

so that

$$\begin{aligned} \Delta(t,y) &=\\ \Delta(t,L(y_1),\ldots,\Delta(y_{j-1}),L(u)+\alpha\cdot L(v),L(y_{j+1}),\ldots,L(y_n)) &=\\ \Delta(t,L(y_1),\ldots,\Delta(y_{j-1}),L(u),L(y_{j+1}),\ldots,L(y_n))+\alpha\cdot\Delta(t,L(y_1),\ldots,\\ \Delta(y_{j-1}),L(v),L(y_{j+1}),\ldots,L(y_n)) &\underset{\text{[eqs: 11.171,11.172,11.173]}}{=}\\ \Delta(t,L(r_1),\ldots,L(r_n))+\alpha\cdot\Delta(t,L(s_1),\ldots,L(s_n)) &=\\ \Delta_j(t,r)+\alpha\cdot\Delta(t,s)\end{aligned}$$

proving that

$$A_j(t,y)=A_j(t,r)+\alpha\cdot A_j(t,s)$$

***i* < *j*.** Then $1 < i + 1 \leqslant j$ so that by [definition: 11.258]

$$\begin{array}{rcl}(L(y_1), \ldots, L(y_{j-1}), L(y_{j+1}), \ldots, L(y_n))_{i+1} & = & L(y_{(i+1)-1})\\ & = & L(y_i)\\ & = & L(u + \alpha \cdot v)\\ & = & L(u) + \alpha \cdot L(v)\end{array}$$

so that

$$\begin{array}{rl}\Delta(t, y) & =\\ \Delta(t, L(y_1), \ldots, \Delta(y_{(j+1)-1}), L(u) + \alpha \cdot L(v), L(y_{(j+1)+1}), \ldots, L(y_n)) & =\\ \Delta(t, L(y_1), \ldots, \Delta(y_{(j+1)-1}), L(u), L(y_{(j+1)+1}), \ldots, L(y_n)) + \alpha \cdot \Delta(t, & \\ L(y_1), \ldots, \Delta(y_{(j+1)-1}), L(v), L(y_{(j+1)+1}), \ldots, L(y_n)) & \underset{[\text{eqs: } 11.171, 11.172, 11.173]}{=}\\ \Delta(t, L(r_1), \ldots, L(r_n)) + \alpha \cdot \Delta(t, L(s_1), \ldots, L(s_n)) & =\\ \Delta_j(t, r) + \alpha \cdot \Delta(t, s) & \end{array}$$

proving that

$$A_j(t, y) = A_j(t, r) + A_j(t, s)$$

So in all cases we have that

$$A_j(t, y) = A_j(t, r) + \alpha \cdot A_j(t, s) \tag{11.175}$$

Hence

$$\begin{array}{rl}\overline{\Delta L}(x_1, \ldots, x_{i-1}, u + \alpha \cdot v, x_{i+1}, \ldots, x_n)(t) & =\\ \overline{\Delta L}(y)(t) & =\\ \sum_{j \in \{1, \ldots, n\}} (-1)^j - 1 \cdot A_j(t, y) & \underset{[\text{eq: } 11.175]}{=}\\ \sum_{j \in \{1, \ldots, n\}} (-1)^{j-1} \cdot (A_j(t, r) + \alpha \cdot A_j(t, s)) & =\\ \sum_{j \in \{1, \ldots, n\}} (-1)^{j-1} \cdot A_j(t, r) + \alpha \cdot \sum_{j \in \{1, \ldots, n\}} (-1)^{j-1} \cdot A_j(t, s) & =\\ \overline{\Delta L}(r)(t) + \alpha \cdot \overline{\Delta L}(s)(t) & =\\ \overline{\Delta L}(x_1, \ldots, x_{i-1}, u, x_{i+1}, \ldots, x_n)(t) + \alpha \cdot \overline{\Delta L}(x_1, \ldots, x_{i-1}, v, x_{i+1}, \ldots, x_n)(t) & \end{array}$$

proving as $t \in X$ was chosen arbitrary that

$$\begin{array}{rl}\overline{\Delta L}(x_1, \ldots, x_{i-1}, u + \alpha \cdot v, x_{i+1}, \ldots, x_n) & =\\ \overline{\Delta L}(x_1, \ldots, x_{i-1}, u, x_{i+1}, \ldots, x_n) + \alpha \cdot \overline{\Delta L}(x_1, \ldots, x_{i-1}, v, x_{i+1}, \ldots, x_n) & \end{array}$$

hence

$$\overline{\Delta L} \in \mathrm{Hom}(X^n; \mathrm{Hom}(X, X))$$

Next we prove skew-symmetry. 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 for j the following two cases

***j* ≠ *k* ∧ *j* ≠ *l*.** Then we have that following sub cases:

***j* < *k* < *l*.** Then we have $j + 1 \leqslant k$ and $j + 1 \leqslant l$ hence

$$\begin{array}{rl}(t, L(x_1), \ldots, L(x_{j-1}), L(x_{j+1}), \ldots, L(x_n))_k & \underset{[\text{definition: } 11.258]}{=}\\ L(x_k) & \underset{x_k = x_l}{=}\\ L(x_l) & \underset{[\text{definition: } 11.258]}{=}\\ (t, L(x_1), \ldots, L(x_{j-1}), L(x_{j+1}), \ldots, L(x_n))_l & \end{array}$$

which as $k \neq l$ proves by [theorem: 11.248] 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

$$\begin{aligned}
(t, L(x_1), \ldots, L(x_{j-1}), L(x_{j+1}), \ldots, L(x_n))_{k+1} &\underset{\text{[definition: 11.258]}}{=} \\
L(x_{(k+1)-1}) &= \\
L(x_k) &\underset{x_k = x_l}{=} \\
L(x_l) &\underset{\text{[definition: 11.258]}}{=} \\
(t, L(x_1), \ldots, L(x_{j-1}), L(x_{j+1}), \ldots, L(x_n))_l &
\end{aligned}$$

which as $k \neq l \Rightarrow k + 1 \neq l + 1$ proves by [theorem: 11.248] that

$$\Delta(t, L(x_1), \ldots, L(x_{j-1}), L(x_{j+1}), \ldots, L(x_n)) = 0$$

***k* < *l* < *j*.** Then we have $1 < k + 1 \leqslant j$ and $1 < l + 1 \leqslant j$ so that

$$\begin{aligned}
(t, L(x_1), \ldots, L(x_{j-1}), L(x_{j+1}), \ldots, L(x_n))_{k+1} &\underset{\text{[definition: 11.258]}}{=} \\
L(x_{(k+1)-1}) &= \\
L(x_k) &\underset{x_k = x_l}{=} \\
L(x_l) &= \\
L(x_{(l+1)-1}) &\underset{\text{[definition: 11.258]}}{=} \\
(t, L(x_1), \ldots, L(x_{j-1}), L(x_{j+1}), \ldots, L(x_n))_{k+1} &
\end{aligned}$$

which as $k \neq l \Rightarrow k + 1 \neq l + 1$ proves by [theorem: 11.248] 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

$$\Delta(t, L(x_1), \ldots, L(x_{j-1}), L(x_{j+1}), \ldots, L(x_n)) = 0 \tag{11.176}$$

***j* = *k*.** For $j = k < l$ we have by [lemma: 11.274] and $L(x_k) = L()$ that

$$\begin{aligned}
(t, L(x_1), \ldots, L(x_{k-1}), L(x_{k+1}), \ldots, L(x_n)) \circ \left(k \underset{n}{\rightsquigarrow} l + 1\right) &= \\
(t, L(x_1), \ldots, L(x_{l-1}), L(x_{l+1}), \ldots, L(x_n))&
\end{aligned}$$

so that

$$\begin{aligned}
\Delta(t, L(x_1), \ldots, L(x_{l-1}), L(x_{l+1}), \ldots, L(x_n)) &= \\
\Delta\big((t, L(x_1), \ldots, L(x_{k-1}), L(x_{k+1}), \ldots, L(x_n)) \circ \left(k \underset{n}{\rightsquigarrow} l + 1\right)\big) &= \\
\left(\left(k \underset{n}{\rightsquigarrow} l + 1\right)\Delta\right)(t, L(x_1), \ldots, L(x_{k-1}), L(x_{k+1}), \ldots, L(x_n)) &= \\
\operatorname{sign}\left(\left(k \underset{n}{\rightsquigarrow} l + 1\right)\right) \cdot \Delta(t, L(x_1), \ldots, L(x_{k-1}), L(x_{k+1}), \ldots, L(x_n)) &\underset{\text{[theorem: 11.215]}}{=} \\
(-1)^{l+1-k} \cdot \Delta(t, L(x_1), \ldots, L(x_{k-1}), L(x_{k+1}), \ldots, L(x_n))&
\end{aligned}$$

So after multiplying both sides by $(-1)^{l+1-k}$ and the fact that $(-1)^{l+1-k} \cdot (-1)^{l+1-k} = 1$ it follows that:

$$\begin{aligned}
\Delta(t, L(x_1), \ldots, L(x_{k-1}), L(x_{k+1}), \ldots, L(x_n)) &= \\
(-1)^{l+1-k} \cdot \Delta(t, L(x_1), \ldots, L(x_{l-1}), L(x_{l+1}), \ldots, L(x_n))&
\end{aligned} \tag{11.177}$$

Now

$$\begin{array}{rcl} & \overline{\Delta L}(x_1,\ldots,x_n)(t) & = \\ \sum_{j\in\{1,\ldots,n\}} (-1)^j\cdot\Delta(t,L(x_1),\ldots,L(x_{j-1}),L(x_{j+1}),\ldots,L(x_n))\cdot x_j & & \underset{\text{[theorem: 11.41]}}{=} \\ & A+B+C & \end{array} \tag{11.178}$$

where

$$\begin{array}{rcl} A & = & \sum_{j\in\{1,\ldots,n\}\backslash\{k.l\}} (-1)^{j-1}\cdot\Delta(t,L(x_1),\ldots,L(x_{j-1}),L(x_{j+1}),\ldots,L(x_n))\cdot x_j \\ & \underset{\text{[theorem: 11.176]}}{=} & \sum_{j\in\{1,\ldots,n\}\backslash\{k.l\}} (-1)^{j-1}\cdot 0\cdot x_j \\ & = & 0 \\ B & = & \sum_{j\in\{k\}} (-1)^{j-1}\cdot\Delta(t,L(x_1),\ldots,L(x_{j-1}),L(x_{j+1}),\ldots,L(x_n))\cdot x_j \\ & & (-1)^{k-1}\cdot\Delta(t,L(x_1),\ldots,L(x_{k-1}),L(x_{k+1}),\ldots,L(x_n))\cdot x_k \\ & \underset{\text{[eq: 11.177]}}{=} & (-1)^{k-1}\cdot(-1)^{l+1-k}\cdot\Delta(t,L(x_1),\ldots,L(x_{l-1}),L(x_{l+1}),\ldots,L(x_n))\cdot x_k \\ & = & (-1)^l\cdot\Delta(t,L(x_1),\ldots,L(x_{l-1}),L(x_{l+1}),\ldots,L(x_n))\cdot x_k \\ & = & (-1)\cdot(-1)^{l-1}\cdot\Delta(t,L(x_1),\ldots,L(x_{l-1}),L(x_{l+1}),\ldots,L(x_n))\cdot x_k \\ & \underset{x_k=x_l}{=} & (-1)\cdot(-1)^{l-1}\cdot\Delta(t,L(x_1),\ldots,L(x_{l-1}),L(x_{l+1}),\ldots,L(x_n))\cdot x_l \\ C & = & \sum_{j\in\{l\}} (-1)^{j-1}\cdot\Delta(t,L(x_1),\ldots,L(x_{j-1}),L(x_{j+1}),\ldots,L(x_n))\cdot x_j \\ & = & (-1)^{l-1}\cdot\Delta(t,L(x_1),\ldots,L(x_{l-1}),L(x_{l+1}),\ldots,L(x_n))\cdot x_l \\ & = & -B \end{array}$$

so that using [eq: 11.178] $\overline{\Delta L}(x_1,...,x_n)(t)=0$, which, as $t\in X$ was chosen arbitrary, that $\overline{\Delta L}(x_1,..., x_n)=C_0$ [the neutral element in $\mathrm{Hom}(X,X)$] proving that

$$\overline{\Delta L} \text{ is skew-symmetric}$$

□

The above theorem allows us to define the adjoint of a linear map.

Definition 11.276. *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 there exists a function **adjoint***

$$\mathrm{adjoint}\colon \mathrm{Hom}(X,X)\rightarrow\mathrm{Hom}(X,X)$$

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 we must prove that for every L there exists only one $\mathrm{adjoint}(L)\in\mathrm{Hom}(X,X)$ such that $\overline{\Delta L}(x)=\Delta(x)\cdot\mathrm{adjoint}(L)$. First, as by the previous theorem [theorem: 11.275] $\overline{\Delta L}$ is a skew-symmetric n-linear map, we can use [theorem: 11.256] to get a **unique** $y\in\mathrm{Hom}(X,X)$ such that $\forall x\in X^n$ that

$$\overline{\Delta L}(x)=\Delta(x)\cdot y$$

This proves existence, next we have to prove that there exist only one. So let Δ' another non trivial determinant function and $y'\in\mathrm{Hom}(X,X)$ such that $\forall x\in X^n$ we have

$$\overline{\Delta' L}(x)=\Delta(x)\cdot y'$$

Using [theorem: 11.257] there exists a $\lambda\in F$ such that

$$\Delta'=\lambda\cdot\Delta$$

Then we have $\forall x\in X^n, t\in X$ that

$$\begin{array}{rcl}(\Delta'(x)\cdot y')(t) & = & \overline{\Delta' L}(x)(t)\\ & = & \sum_{j\in\{1,\ldots,n\}} (-1)^{j-1}\cdot\Delta'(t, L(x_1),\ldots,L(x_{j-1}),L(x_{j+1}),\ldots,L(x_n))\cdot x_j\\ & = & \sum_{j\in\{1,\ldots,n\}} (-1)^{j-1}\cdot\lambda\cdot\Delta(t, L(x_1),\ldots,L(x_{j-1}),L(x_{j+1}),\ldots,L(x_n))\cdot x_j\\ & = & \lambda\cdot\sum_{j\in\{1,\ldots,n\}} (-1)^{j-1}\cdot\Delta(t, L(x_1),\ldots,L(x_{j-1}),L(x_{j+1}),\ldots,L(x_n))\cdot x_j\\ & = & \lambda\cdot\overline{\Delta L}(x)(t)\\ & = & \lambda\cdot(\Delta(x)\cdot y)(t)\\ & = & (\Delta'(x)\cdot y)(t)\end{array}$$

which, as 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 11.277. If X is a one dimensional space with basis $\{e_1\}$ then if $L\in\mathrm{Hom}(X,X)$ then 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: 11.270] 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 11.278. *Let $n\in\mathbb{N}$, X a vector space over a field F of characteristic zero such that $\dim(X)=n$ and $L\in\operatorname{Hom}(X)$ then we have*

1. $\operatorname{adjoint}(L)\circ L=\det(L)\cdot\operatorname{Id}_X$
2. $L\circ\operatorname{adjoint}(L)=\det(L)\cdot\operatorname{Id}_X$

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: 11.255].

1. Let $x\in X$ then we have

$$\begin{array}{rc}
(\operatorname{adjoint}(L)\circ L)(x) & = \\
\operatorname{adjoint}(L)(L(x)) & = \\
1\cdot\operatorname{adjoint}(L)(L(x)) & = \\
\Delta(e_1,\ldots,e_n)\cdot\operatorname{adjoint}(L)(L(x)) & \underset{[\text{theorem: }11.276]}{=} \\
\overline{\Delta L}(e_1,\ldots,e_n)(L(x)) & \underset{[\text{definition: }11.273]}{=} \\
\sum_{i\in\{1,\ldots,n\}}(-1)^{i-1}\cdot\Delta(L(x),L(e_1),\ldots,L(e_{i-1}),L(e_{i+1}),\ldots,L(e_n))\cdot x_i & = \\
\sum_{i\in\{1,\ldots,n\}}(-1)^{i-1}\cdot\Delta_L(x,e_1,\ldots,e_{i-1},e_{i+1},\ldots,e_n)\cdot x_i & \underset{[\text{theorem, def: }11.264,11.260]}{=} \\
\underline{\Delta_L}(x,(e_1,\ldots,e_n)) & \underset{[\text{theorem: }11.262]}{=} \\
\Delta_L(e_1,\ldots,e_n)\cdot x & \underset{[\text{theorem: }11.267]}{=} \\
(\det(L)\cdot\Delta(e_1,\ldots,e_n))\cdot x & = \\
\det(L)\cdot x & = \\
\det(L)\cdot\operatorname{Id}_X(x) &
\end{array}$$

proving that

$$\operatorname{adjoint}(L)\circ L=\det(L)\cdot\operatorname{Id}_X$$

2. Let $x\in X$ then we have

$$\begin{array}{rc}
(L\circ\operatorname{adjoint}(L))(x) & = \\
L(\operatorname{adjoint}(L)(x)) & = \\
L(1\cdot\operatorname{adjoint}(L)(x)) & = \\
L(\Delta(e_1,\ldots,e_n)\cdot\operatorname{adjoint}(L)(x)) & \underset{[\text{theorem: }11.276]}{=} \\
L(\overline{\Delta L}(e_1,\ldots,e_n)(x)) & \underset{[\text{definition: }11.273]}{=} \\
L\left(\sum_{i\in\{1,\ldots,n\}}(-1)^{i-1}\cdot\Delta(x,L(e_1),\ldots,L(e_{i-1}),L(e_{i+1}),\ldots,L(e_n))\cdot e_i\right) & = \\
\sum_{i\in\{1,\ldots,n\}}(-1)^{i-1}\cdot\Delta(x,L(e_1),\ldots,L(e_{i-1}),L(e_{i+1}),\ldots,L(e_n))\cdot L(e_i) & \underset{[\text{def: }11.260]}{=} \\
\underline{\Delta}(x,(L(e_1),\ldots,L(e_n))) & \underset{[\text{theorem: }11.262]}{=} \\
\Delta(L(e_1),\ldots,L(e_n))\cdot x & = \\
\Delta_L(e_1,\ldots,e_n)\cdot x & \underset{[\text{theorem: }11.267]}{=} \\
(\det(L)\cdot\Delta(e_1,\ldots,e_n))\cdot x & = \\
\det(L)\cdot x & = \\
\det(L)\cdot\operatorname{Id}_X(x) &
\end{array}$$

proving that

$$L\circ\operatorname{adjoint}(L)=\det(L)\cdot\operatorname{Id}_X \qquad \square$$

The above lemma leads to a theorem that allows us to calculate the inverse of a linear mapping.

Theorem 11.279. *Let $n\in\mathbb{N}$, X a vector space over a field F of characteristic zero such that $\dim(X)=n$ and $L\in\operatorname{Hom}(X)$ then we have:*

1. *L is a isomorphism* $\Leftrightarrow \det(L)\neq 0$
2. *If L is a isomorphism then $L^{-1}=(\det(L))^{-1}\cdot\operatorname{adjoint}(L)$*

Proof.

1. This is already proved in [theorem: 11.271].
2. Using the previous lemma [lemma: 11.278] we have

$$\begin{aligned}L\circ\operatorname{adjoint}(L)=\det(L)\cdot\operatorname{Id}_X \quad&\Rightarrow\quad \operatorname{adjoint}(L)=L^{-1}\circ(\det(L)\cdot\operatorname{Id}_X)\\ &\Rightarrow\quad \operatorname{adjoint}(L)=\det(L)\cdot L^{-1}\circ\operatorname{Id}_X\\ &\Rightarrow\quad \operatorname{adjoint}(L)=\det(L)\cdot L^{-1}\\ &\underset{\det(L)\neq 0}{\Rightarrow}\quad (\det(L))^{-1}\cdot\operatorname{adjoint}(L)=L^{-1}\end{aligned}$$

$\square$

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 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 11.280. *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 $M\in\mathcal{M}_{n,m}(F)$ a 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

$$C=\left(\begin{array}{c} C_1\\ \vdots\\ C_n \end{array}\right)$$

In essence a $n\times m$ matrix can be set of a combination of n $1\times m$ row vectors or a combination 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 11.281. Let $n,m\in\mathbb{N}$, F be a field then $E\in\mathcal{M}_{n,m}(F)$ is defined by

$$\forall(i,j)\in\{1,\ldots,n\}\times\{1,\ldots,m\} \text{ we have } E_{i,j}=\delta_{i,j} \text{ [see definition: 11.142]}$$

this matrix is called the **identity matrix**. So

$$E=\left(\begin{array}{ccccc} 1 & 0 & \ldots & 0 & 1\\ 0 & 1 & \ddots & 0 & 0\\ \vdots & 0 & \ddots & 0 & 0\\ 0 & 0 & \ldots & 0 & 1 \end{array}\right)$$

Example 11.282. 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

$$0=\left(\begin{array}{ccc} 0 & \ldots & 0\\ \vdots & \ddots & \vdots\\ 0 & \ldots & 0 \end{array}\right)$$

We define now the different arithmetic operations on matrices.

Definition 11.283. *Let $n,m\in\mathbb{N}$ then we define*

1. **(sum)** $+:\mathcal{M}_{n,m}(F)\times\mathcal{M}_{n,m}(F)\rightarrow\mathcal{M}_{n,m}(F)$ *is defined by*

$$A+B\in\mathcal{M}_{n,m}(F) \textit{ where } \forall(i,j)\in\{1,\ldots,n\}\times\{1,\ldots,m\}\ (A+B)_{i,j}=A_{i,j}+B_{i,j}$$

2. **(scalar product)** $\cdot:F\times\mathcal{M}_{n,m}(F)\rightarrow\mathcal{M}_{n,m}(F)$ *is defined by*

$$\alpha\cdot A\in\mathcal{M}_{n,m}(F) \textit{ where } \forall(i,j)\in\{1,\ldots,n\}\times\{1,\ldots,m\}\ (\alpha\cdot A)_{i,j}=\alpha\cdot A_{i,j}$$

3. **(inner product)** *If $k\in\mathbb{N}$ then we define* $\cdot:\mathcal{M}_{n,m}(F)\times\mathcal{M}_{m,k}(F)\rightarrow\mathcal{M}_{n,k}(F)$ *by*

$$A\cdot B\in\mathcal{M}_{n,k}(F) \textit{ where } (A\cdot B)_{i,j}=\sum_{r\in\{1,\ldots,m\}}A_{i,r}\cdot B_{r,j}$$

Example 11.284. Then we have

$$\left(\begin{array}{ccc} 1 & 2 & 3\\ 4 & 4 & 2 \end{array}\right)+\left(\begin{array}{ccc} 0 & 2 & 6\\ 4 & 1 & 1 \end{array}\right)=\left(\begin{array}{ccc} 1 & 4 & 9\\ 8 & 5 & 3 \end{array}\right)$$

$$10\cdot\left(\begin{array}{ccc} 1 & 2 & 3\\ 4 & 4 & 2 \end{array}\right)=\left(\begin{array}{ccc} 10 & 20 & 30\\ 40 & 40 & 20 \end{array}\right)$$

$$\left(\begin{array}{ccc} 1 & 2 & 3\\ 4 & 4 & 2 \end{array}\right)\cdot\left(\begin{array}{cc} 1 & 2\\ 0 & 3\\ 2 & 0 \end{array}\right)=\left(\begin{array}{cc} 7 & 20\\ 8 & 24 \end{array}\right)$$

Theorem 11.285. *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. The neutral element is the null matrix 0 and for each $\mathrm{M} \in \mathcal{M}_{i,j}$ the additive inverse $-M$ is defined by $(-M)_{i,j} = -M_{i,j}$ $\forall (i,j) \in \{1, \ldots, n\} \times \{1, \ldots, m\}$.*

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

$$1\cdot A=A$$

□

Theorem 11.286. *Let $n\in\mathbb{N}$ and F a field then $\langle\mathcal{M}_{n,n}(F),\cdot\rangle$ is a semi-group with 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

$$\begin{aligned}((A\cdot B)\cdot C)_{i,j} &= \sum_{k\in\{1,\ldots,n\}}(A\cdot B)_{i,k}\cdot C_{k,j}\\ &= \sum_{k\in\{1,\ldots,n\}}\left(\sum_{l\in\{1,\ldots,n\}}A_{i,l}\cdot B_{l,k}\right)\cdot C_{k,j}\\ &\underset{[\text{theorem: } 11.67]}{=} \sum_{k\in\{1,\ldots,n\}}\left(\sum_{l\in\{1,\ldots,n\}}(A_{i,l}\cdot B_{l,k})\cdot C_{k,j}\right)\\ &\underset{[\text{theorem: } 11.39]}{=} \sum_{l\in\{1,\ldots,n\}}\left(\sum_{k\in\{1,\ldots,n\}}(A_{i,l}\cdot B_{l,k})\cdot C_{k,j}\right)\\ &= \sum_{l\in\{1,\ldots,n\}}\left(\sum_{k\in\{1,\ldots,n\}}A_{i,l}\cdot(B_{l,k})\cdot C_{k,j}\right)\\ &\underset{[\text{theorem: } 11.67]}{=} \sum_{l\in\{1,\ldots,n\}}A_{i,l}\cdot\left(\sum_{k\in\{1,\ldots,n\}}(B_{l,k})\cdot C_{k,j}\right)\\ &= \sum_{l\in\{1,\ldots,n\}}A_{i,l}\cdot(B\cdot C)_{l,j}\\ &= (A\cdot(B\cdot C))_{i,j}\end{aligned}$$

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

$$A\cdot E = A = E\cdot A$$ □

Note 11.287. That the inner product is not commutative as the following example shows:

$$\begin{aligned}
\begin{pmatrix}1&3\\2&4\end{pmatrix}\cdot\begin{pmatrix}4&1\\2&1\end{pmatrix} &= \begin{pmatrix}10&4\\16&6\end{pmatrix}\\
\begin{pmatrix}4&1\\2&1\end{pmatrix}\cdot\begin{pmatrix}1&3\\2&4\end{pmatrix} &= \begin{pmatrix}6&16\\4&10\end{pmatrix}
\end{aligned}$$

Also0 there exist non zero matrices that do not have a inverse, for example for

$$\begin{pmatrix}0&1\\0&0\end{pmatrix}\neq\begin{pmatrix}0&0\\0&0\end{pmatrix}$$

we have

$$\begin{pmatrix}a&b\\c&d\end{pmatrix}\cdot\begin{pmatrix}0&1\\0&0\end{pmatrix}=\begin{pmatrix}0&a\\0&c\end{pmatrix}\neq\begin{pmatrix}1&0\\0&1\end{pmatrix} \tag{11.179}$$

We define now the transpose of a matrix, which is essential the operation that interchange the column vector and row vectors of a matrix.

Definition 11.288. *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 and column rank of a matrix that eventually leads to the important concept of the rank from a matrix.

Definition 11.289. *Let F be a field, $n,m\in\mathbb{N}$ then we define:*

1. row: $\mathcal{M}_{n,m}(F)\times\{1,\ldots,n\}\to F^m$ *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. col: $\mathcal{M}_{n,m}(F)\times\{1,\ldots,m\}\to F^n$ *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* $\operatorname{cols}(M)$*, hence:*

$$\operatorname{cols}(M)=\{\operatorname{col}(M,i)|i\in\{1,\ldots,m\}\}\subseteq F^n$$

As $\langle F^n,+,\rangle$ is a vector space over F with $\dim(F^n)=n$, $\langle F^m,+,\cdot\rangle$ is a vector space over F with $\dim(F^m)=m$ *[see example: 11.145] and for $M\in\mathcal{M}_{n,m}(F)$* $\operatorname{span}(\operatorname{rows}(M))$ *and* $\operatorname{span}(\operatorname{cols}(M))$ *are sub spaces of F^m we have that* $\dim(\operatorname{span}(\operatorname{rows}(M)))$ *and* $\dim(\operatorname{span}(\operatorname{cols}(M)))$ *are defined with* $\dim(\operatorname{span}(\operatorname{rows}(M)))\leqslant m$ *and* $\dim(\operatorname{span}(\operatorname{cols}(M)))\leqslant n$ *[see theorems: 11.90, 11.135]. This allows us to define:*

1. *If $M\in\mathcal{M}_{n,m}(F)$ then the row rank of M noted as* $\operatorname{rrank}(M)$ *is defined by*

$$\operatorname{rrank}(M)=\dim(\operatorname{span}(\operatorname{rows}(M)))\leqslant m$$

2. *If $M\in\mathcal{M}_{n,m}(F)$ then the column rank of M noted as* $\operatorname{crank}(M)$ *is defined by*

$$\operatorname{crank}(M)=\dim(\operatorname{span}(\operatorname{cols}(M)))\leqslant n$$

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 11.290. *Let F a field, $n,m\subset\mathbb{N}$ and $M\in\mathcal{M}_{n,m}(F)$ then*

$$\operatorname{rrank}(M)=\operatorname{crank}(M)$$

Proof. Let

$$r=\operatorname{rrank}(M)=\dim(\operatorname{span}(\operatorname{rows}(M)))\leqslant m \text{ and } c=\operatorname{crank}(M)=\dim(\operatorname{span}(\operatorname{cols}(M)))\leqslant n$$

A $r=\dim(\operatorname{span}(\operatorname{rows}(M)))$ there exist by [theorem: 11.136] a distinct family $\{e_i\}_{i\in\{1,\ldots,r\}}\subseteq\operatorname{span}(\operatorname{rows}(M))\subseteq F^m$ such that $\{e_i|i\in\{1,\ldots,r\}\}$ is a basis of $\operatorname{span}(\operatorname{rows}(M))$. Let $i\in\{1,\ldots,n\}$ then there exist a $\{\lambda_{i,j}\}_{j\in\{1,\ldots,r\}}\subseteq F$ such that

$$\operatorname{row}(M,i)=\sum_{j\in\{1,\ldots,r\}}\lambda_{i,j}\cdot e_j \tag{11.180}$$

now let $k\in\{1,\ldots,m\}$ then

$$\begin{aligned}\operatorname{col}(M,k)_i &= M_{i,k}\\ &= \operatorname{row}(M,i)_k\\ &\underset{[\text{eq: } 11.180]}{=} \left(\sum_{j\in\{1,\ldots,r\}}\lambda_{i,j}\cdot e_j\right)_k\\ &\underset{[\text{theorem: } 11.44]}{=} \sum_{j\in\{1,\ldots,r\}}\lambda_{i,j}\cdot(e_j)_k\\ &= \sum_{j\in\{1,\ldots,r\}}(e_j)_k\cdot\lambda_{i,j}\end{aligned} \tag{11.181}$$

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: 11.181] 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: 11.87] 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: 11.89, 11.92]

$$\mathrm{span}(\mathrm{cols}(M)) \subseteq \mathrm{span}(\{f_i | i \in \{1, \ldots, r\}\})$$

As $\mathrm{span}(\{f_i | i \in \{1, \ldots, r\}\})$ is a vector space we have by [theorem: 11.137] that

$$\dim(\mathrm{span}(\{f_i | i \in \{1, \ldots, r\}\})) \leqslant r.$$

Using [theorem: 11.135] gives then:

$$\mathrm{rrank}(M) = \dim(\mathrm{span}(\mathrm{cols}(M))) \leqslant r = \mathrm{crank}(M) \tag{11.182}$$

Now we use the same reasoning to prove the opposite equation. As $c = \dim(\mathrm{span}(\mathrm{cols}(M)))$ there exist by [theorem: 11.136] a distinct family $\{g_i\}_{i \in \{1, \ldots, c\}} \subseteq \mathrm{span}(\mathrm{cols}(M)) \subseteq F^n$ such that $\{g_i | i \in \{1,\ldots, c\}\}$ is a basis of $\mathrm{span}(\mathrm{cols}(M))$. Let $i \in \{1,\ldots, m\}$ then there exists a $\{\alpha_{i,j}\}_{j \in \{1, \ldots, c\}} \subseteq F$ such that

$$\mathrm{col}(M, i) = \sum_{j \in \{1, \ldots, c\}} \alpha_{i,j} \cdot g_j \tag{11.183}$$

Now let $k \in \{1, \ldots, n\}$ then we have

$$\begin{aligned}(\mathrm{row}(M, k))_i &= M_{k,i}\\ &= (\mathrm{col}(M, i))_k\\ &\underset{[\text{eq: } 11.183]}{=} \left(\sum_{j \in \{1, \ldots, c\}} \alpha_{i,j} \cdot g_j\right)_k\\ &\underset{[\text{theorem: } 11.44]}{=} \sum_{j \in \{1, \ldots, c\}} \alpha_{i,j} \cdot (g_j)_k\\ &= \sum_{j \in \{1, \ldots, c\}} (g_j)_k \cdot \alpha_{i,j}\end{aligned} \tag{11.184}$$

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 [e: 11.184] we have

$$(\mathrm{row}(M, k))_i = \sum_{j \in \{1, \ldots, c\}} (g_j)_k \cdot (h_j)_i$$

so that

$$\mathrm{row}(M, k) = \sum_{j \in \{1, \ldots, c\}} (g_j)_k \cdot h_j$$

prove using [theorem: 11.87] that $\mathrm{row}(M, k) \in \mathrm{span}(\{h_i | i \in \{1, \ldots, c\}\})$. Hence

$$\mathrm{rows}(M) \subseteq \mathrm{span}(\{h_i | i \in \{1, \ldots, c\}\})$$

so that by that by [theorems: 11.89, 11.92]

$$\mathrm{span}(\mathrm{rows}(M)) \subseteq \mathrm{span}(\{h_i | i \in \{1, \ldots, c\}\})$$

As $\mathrm{span}(\{h_i | i \in \{1, \ldots, c\}\})$ is a vector space we have by [theorem: 11.137] that

$$\dim(\mathrm{span}(\{h_i | i \in \{1, \ldots, c\}\})) \leqslant c.$$

Using [theorem: 11.135] gives then:

$$\mathrm{rrank}(M) = \dim(\mathrm{span}(\mathrm{rows}(M))) \leqslant c = \mathrm{crank}(M)$$

combining the above with [eq: 11.182] proves

$$\mathrm{rrank}(M) = \mathrm{crank}(M)$$ □

The above theorem let us define the rank of a matrix.

Definition 11.291. *Let $n, m \in \mathbb{N}$, F a field and $M \in \mathcal{M}_{n,m}(F)$ then*

$$\operatorname{rank}(M) = \operatorname{rrank}(M) \underset{\text{[theorem: 11.290]}}{=} \operatorname{crank}(M)$$

Note: As by [definition: 11.289] $\operatorname{rrank}(M) \leqslant m$ *and* $\operatorname{crank}(M) \leqslant m$ *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 11.292. *Let $n, m \in \mathbb{N}$, X, Y finite dimensional vector spaces over a field F with basis $\{e_i | i \in \{1, \ldots, n\}\}, \{f_i | i \in \{1, \ldots, m\}\}$ defined by distinct families $E = \{e_i\}_{i \in \{1, \ldots, n\}} \subseteq X$, $F = \{f_i\}_{i \in \{1, \ldots, m\}} \subseteq Y$ then if $L_1, L_2 \in \operatorname{Hom}(X, Y)$ such that $\forall i \in \{1,\ldots,n\}$ $L_1(e_i) = L_2(e_i)$ 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 11.293. *Let $n, m \in \mathbb{N}$, X, Y finite dimensional vector spaces over a field F with basis $\{e_i | i \in \{1, \ldots, n\}\}, \{f_i | i \in \{1, \ldots, m\}\}$ defined by distinct families $E = \{e_i\}_{i \in \{1, \ldots, n\}} \subseteq X$, $F = \{f_i\}_{i \in \{1, \ldots, m\}} \subseteq Y$ 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: 11.123 ensures uniqueness and existence]. $\mathcal{M}(L; E, F)$ is called the matrix of L for the basis E and F.

Note 11.294. The matrix $M(L; E, F)$ depends clearly not only on L but also on E and F. If it is clear from the context which basis E and F is 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 11.295. *Let $n,m\in\mathbb{N}$, X,Y finite dimensional vector spaces over a field F with bases $\{e_i|i\in\{1,\ldots,n\}\},\{f_i|i\in\{1,\ldots,m\}\}$ defined by distinct families $E=\{e_i\}_{i\in\{1,\ldots,n\}}\subseteq X$, $F=\{f_i\}_{i\in\{1,\ldots,m\}}\subseteq Y$ and $L\in\operatorname{Hom}(X,Y)$ then we have:*

1. *If $x\in X$ then by [theorem: 11.123] 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

$$X=\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}\in\mathcal{M}_{n,1}(F)\text{ the column vector uniquely define by the components of }x$$

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

$$\begin{array}{lcl}
\sum_{j\in\{1,\ldots,m\}}L(x)_j\cdot f_j & = & L(x)\\
 & = & L\left(\sum_{i\in\{1,\ldots,n\}}x_i\cdot e_i\right)\\
 & \underset{[\text{theorem: }11.167]}{=} & \sum_{i\in\{1,\ldots,n\}}x_i\cdot L(e_i)\\
 & \underset{\text{def}}{=} & \sum_{i\in\{1,\ldots,n\}}x_i\cdot\left(\sum_{j\in\{1,\ldots,m\}}\mathcal{M}(L;E,F)_{j,i}\cdot f_j\right)\\
 & \underset{[\text{theorem: }11.67]}{=} & \sum_{i\in\{1,\ldots,n\}}\left(\sum_{j\in\{1,\ldots,m\}}x_i\cdot(\mathcal{M}(L;E,F)_{j,i}\cdot f_j)\right)\\
 & = & \sum_{i\in\{1,\ldots,n\}}\left(\sum_{j\in\{1,\ldots,m\}}(x_i\cdot\mathcal{M}(L;E,F)_{j,i})\cdot f_j\right)\\
 & \underset{[\text{theorem: }11.39]}{=} & \sum_{j\in\{1,\ldots,m\}}\left(\sum_{i\in\{1,\ldots,n\}}(x_i\cdot\mathcal{M}(L;E,F)_{j,i})\cdot f_j\right)\\
 & \underset{[\text{theorem: }11.67]}{=} & \sum_{j\in\{1,\ldots,m\}}\left(\sum_{i\in\{1,\ldots,n\}}x_i\cdot\mathcal{M}(L;E,F)_{j,i}\right)f_j\\
 & = & \sum_{j\in\{1,\ldots,m\}}\left(\sum_{i\in\{1,\ldots,n\}}\mathcal{M}(L;E,F)_{j,i}\cdot x_i\right)f_j
\end{array}$$

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 $M=\mathcal{M}(L;E,F)\in\mathcal{M}_{m,n}(F)$. If $y\in L(X)$ then there exists a $x\in X$ such that $y=L(x)$, then $x=\sum_{i\in\{1,\ldots,n\}}x_i\cdot e_i$ so that

$$y=L\left(\sum_{i\in\{1,\ldots,n\}}x_i\cdot e_i\right)\underset{[\text{theorem: }11.167]}{=}\sum_{i\in\{1,\ldots,n\}}x_i\cdot L(e_i)$$

proving by [theorem: 11.83] that $y\in\operatorname{span}(\{L(e_i)|i\in\{1,\ldots,n\}\})$, hence

$$L(X)\subseteq\operatorname{span}(\{L(e_i)|i\in\{1,\ldots,n\}\})\tag{11.185}$$

If $y\in\operatorname{span}(\{L(e_i)|i\in\{1,\ldots,n\}\})$ then by [theorem: 11.83] there exists a $\{\alpha_i\}_{i\in\{1,\ldots,n\}}\subseteq F$ such that

$$y=\sum_{i\in\{1,\ldots,n\}}\alpha_i\cdot L(e_i)\underset{[\text{theorem: }11.167]}{=}L\left(\sum_{i\in[1,\ldots,n]}e_i\right)\in L(X)$$

proving that $\operatorname{span}(\{L(e_i)|i\in\{1,\ldots,n\}\})\subseteq L(X)$. Combining this with [eq: 11.185] gives

$$L(X)=\operatorname{span}(\{L(e_i)|i\in\{1,\ldots,n\}\})\tag{11.186}$$

Using [theorem: 11.126] we have 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: 11.78] it follows that there exist a bijection $\beta:\{1,\ldots,\operatorname{rank}(L)\}\to J\subseteq\{1,\ldots,n\}$ such that

$\{L(e_{\beta(i)})\}_{i\in\{1,\ldots,\operatorname{rank}(L)\}}\subseteq L(X)$ is a distinct family and $B=\{L(e_{\beta(i)})|i\in\{1,\ldots,\operatorname{rank}(L)\}\}$ is a basis for $L(X)$ (11.187)

Define now

$$B_c=\{\operatorname{col}(M,\beta(i))|i\in\{1,\ldots,\operatorname{rank}(L)\}\}\subseteq\operatorname{cols}(M)$$

then by [theorem: 11.89] we have that

$$\operatorname{span}(B_c)\subseteq\operatorname{span}(\operatorname{cols}(M))\tag{11.188}$$

For the opposite inclusion let $c\in\operatorname{span}(\operatorname{cols}(M))=\operatorname{span}(\operatorname{col}(M,i)|i\in\{1,\ldots,n\})$ then by [theorem: 11.83] 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

$$\begin{aligned}c_k&=\left(\sum_{i\in\{1,\ldots,n\}}\lambda_i\cdot\operatorname{col}(M,i)\right)_k\\&\underset{[\text{theorem: }11.44]}{=}\sum_{i\in\{1,\ldots,n\}}\lambda_i\cdot(\operatorname{col}(M,i))_k\\&\underset{[\text{definition: }11.289]}{=}\sum_{i\in\{1,\ldots,n\}}\lambda_i\cdot M_{k,i}\end{aligned}\tag{11.189}$$

Consider now the sum $\sum_{i\in\{1,\ldots,n\}}\lambda_i\cdot L(e_i)$, then we have that

$$\sum_{i\in\{1,\ldots,n\}}\lambda_i\cdot L(e_i)\in\operatorname{span}(\{L(e_i)|i\in\{1,\ldots,n\}\})\underset{[\text{eq: }11.186]}{=}L(X)$$

By the above, [eq: 11.187] $\{L(e_{\beta(i)})|i\in\{1,\ldots,\operatorname{rank}(L)\}\}$ and [theorem: 11.123] there exist a $\{\gamma_i\}_{i\in\{1,\ldots,\operatorname{rank}(L)\}}\subseteq F$ such that

$$\begin{aligned}
\sum_{i\in\{1,\ldots,n\}}\lambda_i\cdot L(e_i) &= \sum_{j\in\{1,\ldots,\operatorname{rank}(L)\}}\gamma_j\cdot L(e_{\beta(j)})\\
&\underset{[\text{definition: }11.293]}{=} \sum_{j\in\{1,\ldots,\operatorname{rank}(L)\}}\gamma_j\cdot\left(\sum_{i\in\{1,\ldots,m\}}\mathcal{M}(L;E,F)_{i,\beta(j)}\cdot f_i\right)\\
&= \sum_{j\in\{1,\ldots,\operatorname{rank}(L)\}}\gamma_j\cdot\left(\sum_{i\in\{1,\ldots,m\}}M_{i,\beta(j)}\cdot f_i\right)\\
&\underset{[\text{theorem: }11.67]}{=} \sum_{j\in\{1,\ldots,\operatorname{rank}(L)\}}\left(\sum_{i\in\{1,\ldots,m\}}\gamma_j\cdot(M_{i,\beta(j)}\cdot f_i)\right)\\
&\underset{[\text{theorem: }11.39]}{=} \sum_{i\in\{1,\ldots,m\}}\left(\sum_{j\in\{1,\ldots,\operatorname{rank}(L)\}}\gamma_j\cdot(M_{i,\beta(j)}\cdot f_i)\right)\\
&= \sum_{i\in\{1,\ldots,m\}}\left(\sum_{j\in\{1,\ldots,\operatorname{rank}(L)\}}(\gamma_j\cdot M_{i,\beta(j)})\cdot f_i\right)\\
&\underset{[\text{theorem: }11.67]}{=} \sum_{i\in\{1,\ldots,m\}}\Bigg(\sum_{j\in\{1,\ldots,\operatorname{rank}(L)\}}\gamma_j\cdot M_{i,\beta(j)})\cdot f_i\\
\sum_{i\in\{1,\ldots,n\}}\lambda_i\cdot L(e_i) &\underset{[\text{definition: }11.293]}{=} \sum_{j\in\{1,\ldots,n\}}\lambda_j\cdot\left(\sum_{i\in\{1,\ldots,m\}}\mathcal{M}(L;E,F)_{i,j}\cdot f_i\right)\\
&= \sum_{j\in\{1,\ldots,n\}}\lambda_j\cdot\left(\sum_{i\in\{1,\ldots,m\}}M_{i,j}\cdot f_i\right)\\
&\underset{[\text{theorem: }11.67]}{=} \sum_{j\in\{1,\ldots,n\}}\left(\sum_{i\in\{1,\ldots,m\}}\lambda_j\cdot(M_{i,j}\cdot f_i)\right)\\
&\underset{[\text{theorem: }11.39]}{=} \sum_{i\in\{1,\ldots,m\}}\left(\sum_{j\in\{1,\ldots,n\}}\lambda_j\cdot(M_{i,j}\cdot f_i)\right)\\
&= \sum_{i\in\{1,\ldots,m\}}\left(\sum_{j\in\{1,\ldots,n\}}\lambda_j\cdot(M_{i,j}\cdot f_i)\right)\\
&= \sum_{i\in\{1,\ldots,m\}}\left(\sum_{j\in\{1,\ldots,n\}}(\lambda_j\cdot M_{i,j})\cdot f_i\right)\\
&\underset{[\text{theorem: }11.67]}{=} \sum_{i\in\{1,\ldots,m\}}\left(\sum_{j\in\{1,\ldots,n\}}\lambda_j\cdot M_{i,j}\right)\cdot f_i
\end{aligned}$$

Using the uniqueness of the expansion of a vector in a basis we have from the above that

$$\sum_{j\in\{1,\ldots,\operatorname{rank}(L)\}}\gamma_j\cdot M_{i,\beta(j)}=\sum_{j\in\{1,\ldots,n\}}\lambda_j\cdot M_{i,j}$$

and substituting this in [eq: 11.189] proves that $\forall k\in\{1,\ldots,m\}$

$$\begin{aligned}
c_k &= \sum_{j\in\{1,\ldots,\operatorname{rank}(L)\}}\gamma_j\cdot M_{k,\beta(j)}\\
&\underset{[\text{definition: }11.289]}{=} \sum_{j\in\{1,\ldots,\operatorname{rank}(L)\}}\gamma_j\cdot\operatorname{col}(M,\beta(j))_k\\
&\underset{[\text{theorem: }11.44]}{=} \left(\sum_{j\in\{1,\ldots,\operatorname{rank}(L)\}}\gamma_j\cdot\operatorname{col}(M,\beta(j))\right)_k
\end{aligned}$$

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 $\mathrm{span}(\mathrm{cols}(M))\subseteq\mathrm{span}(B_c)$ which together with [eq: 11.188] gives $\mathrm{span}(\mathrm{cols}(M))=\mathrm{span}(B_c)$ or using the definition of B_c.

$$\mathrm{span}\{\mathrm{col}(M,\beta(i))|i\in\{1,\ldots,\mathrm{rank}(L)\}\}=\mathrm{span}(\mathrm{cols}(M))\tag{11.190}$$

Next we prove that $\{\mathrm{col}(M,\beta(i))\}_{i\in\{1,\ldots,\mathrm{rank}(L)\}}$ is linear independent. Let $\{\alpha_i\}_{i\in\{1,\ldots,\mathrm{rank}(L)\}}\subseteq F$ such that

$$\sum_{i\in\{1,\ldots,\mathrm{rank}(L)\}}\alpha_i\cdot\mathrm{col}(M,\beta(i))=0$$

then $\forall k\in\{1,\ldots,m\}$ we have

$$\begin{aligned}
0 &= \left(\sum_{i\in\{1,\ldots,\mathrm{rank}(L)\}}\alpha_i\cdot\mathrm{col}(M,\beta(i))\right)_k\\
&\underset{[\text{theorem: }11.44]}{=} \sum_{i\in\{1,\ldots,\mathrm{rank}(L)\}}\alpha_i\cdot\mathrm{col}(M,\beta(i))_k\\
&\underset{[\text{definition: }11.289]}{=} \sum_{i\in\{1,\ldots,\mathrm{rank}(L)\}}\alpha_i\cdot M_{k,\beta(i)}
\end{aligned}\tag{11.191}$$

Now

$$\begin{aligned}
\sum_{i\in\{1,\ldots,\mathrm{rank}(L)\}}\alpha_i\cdot L(e_{\beta(i)}) &\underset{[\text{definition: }11.293]}{=}\\
\sum_{i\in\{1,\ldots,\mathrm{rank}(L)\}}\alpha_i\cdot\left(\sum_{j\in\{1,\ldots,m\}}\mathcal{M}(L;E,F)_{j,\beta(i)}\cdot f_j\right) &=\\
\sum_{i\in\{1,\ldots,\mathrm{rank}(L)\}}\alpha_i\cdot\left(\sum_{j\in\{1,\ldots,m\}}M_{j,\beta(i)}\cdot f_j\right) &\underset{[\text{theorem: }11.67]}{=}\\
\sum_{i\in\{1,\ldots,\mathrm{rank}(L)\}}\left(\sum_{j\in\{1,\ldots,m\}}\alpha_i\cdot(M_{j,\beta(i)}\cdot f_j)\right) &\underset{[\text{theorem: }11.39]}{=}\\
\sum_{j\in\{1,\ldots,m\}}\left(\sum_{i\in\{1,\ldots,\mathrm{rank}(L)\}}\alpha_i\cdot(M_{j,\beta(i)}\cdot f_j)\right) &=\\
\sum_{j\in\{1,\ldots,m\}}\left(\sum_{i\in\{1,\ldots,\mathrm{rank}(L)\}}(\alpha_i\cdot M_{j,\beta(i)})\cdot f_j\right) &\underset{[\text{theorem: }11.67]}{=}\\
\sum_{j\in\{1,\ldots,m\}}\left(\sum_{i\in\{1,\ldots,\mathrm{rank}(L)\}}\alpha_i\cdot M_{j,\beta(i)}\right)\cdot f_j &\underset{[\text{eq: }11.191]}{=}\\
\sum_{j\in\{1,\ldots,m\}}0\cdot f_j &=\\
0 &
\end{aligned}\tag{11.192}$$

Now by [eq: 11.187] $\{L(e_{\beta(i)})|i\in\{1,\ldots,\mathrm{rank}(L)\}\}$ is a basis, hence linear independent, and $\{L(e_{\beta(i)})\}_{i\in\{1,\ldots,\mathrm{rank}(L)\}}$ is distinct, so that by [theorem: 11.111] $\{L(e_{\beta(i)})\}_{i\in\{1,\ldots,\mathrm{rank}(L)\}}$ is linear independent. Hence applying [theorem: 11.104] on [eq: 11.192] results in

$$\forall i\in\{1,\ldots,m\}\text{ we have }\alpha_i=0$$

So $\{\mathrm{col}(M,\beta(i))\}_{i\in\{1,\ldots,\mathrm{rank}(L)\}}$ is linear independent or by [theorem: 11.107]

$$\{\mathrm{col}(M,\beta(i))|i\in\{1,\ldots,\mathrm{rank}(L)\}\}\text{ is linear independent}$$

Combining this with [eq: 11.190] proves that

$$\{\mathrm{col}(M,\beta(i))|i\in\{1,\ldots,\mathrm{rank}(L)\}\} \text{ is a basis for } \mathrm{span}(\mathrm{cols}(M))$$

Finally we have that

$$\mathrm{rank}(M)\underset{\text{def}}{=}\dim(\mathrm{cols}(M))=\mathrm{card}(\{\mathrm{col}(M,\beta(i))|i\in\{1,\ldots,\mathrm{rank}(L)\}\})\underset{[\text{theorem: } 11.109]}{=}\mathrm{rank}(L)$$

proving that

$$\mathrm{rank}(\mathcal{M}(L;E,F))=\mathrm{rank}(L)$$ □

Next we show that the mapping that associate a matrix with a linear mapping is 'linear', be aware that $\langle\mathcal{M}_{n,m}(L),+,\cdot\rangle$ is not a vector space so we can not talk about linear mappings between vector spaces.

Theorem 11.296. *Let $n,m\in\mathbb{N}$, X,Y finite dimensional vector spaces over a field F with basis $\{e_i|i\in\{1,\ldots,n\}\},\{f_i|i\in\{1,\ldots,m\}\}$ defined by distinct families $E=\{e_i\}_{i\in\{1,\ldots,n\}}\subseteq X$, $F=\{f_i\}_{i\in\{1,\ldots,m\}}\subseteq Y$ then*

$$\mathcal{M}(E,F)\colon \mathrm{Hom}(X,Y)\rightarrow\mathcal{M}_{m,n}(F) \text{ defined by } \mathcal{M}(E,F)(L)=M(L;E,F)$$

satisfies:

1. *$\forall\alpha\in F$ 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]*

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 in a sense $\mathrm{Hom}(X, Y)$ and $\mathcal{M}_{m,n}(F)$ are isomorphic (again $\mathcal{M}_{m,n}(F)$ is not a vector space).

Theorem 11.297. *Let $n, m \in \mathbb{N}$, X, Y finite dimensional vector spaces over a field F with basis $\{e_i | i \in \{1, \ldots, n\}\}, \{f_i | i \in \{1, \ldots, m\}\}$ defined by distinct families $E = \{e_i\}_{i \in \{1, \ldots, n\}} \subseteq X$, $F = \{f_i\}_{i \in \{1, \ldots, m\}} \subseteq Y$ then*

$$\mathcal{M}(E, F): \mathrm{Hom}(X, Y) \rightarrow \mathcal{M}_{m,n}(F)$$

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

$$\mathcal{M}(L_1; E, F) = \mathcal{M}(L_2; E, F) \tag{11.193}$$

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

$$\begin{array}{rcl} L_1(x) & = & L_1\left(\sum_{i \in \{1, \ldots, n\}} x_i \cdot e_i\right) \\ & \underset{[\text{theorem: } 11.167]}{=} & \sum_{i \in \{1, \ldots, n\}} x_i \cdot L_1(e_i) \\ & = & \sum_{i \in \{1, \ldots, n\}} x_i \cdot \left(\sum_{j \in \{1, \ldots, m\}} \mathcal{M}(L_1; E, F)_{j,i} \cdot f_j\right) \\ & \underset{[\text{eq: } 11.193]}{=} & \sum_{i \in \{1, \ldots, n\}} x_i \cdot \left(\sum_{j \in \{1, \ldots, m\}} \mathcal{M}(L_2; E, F)_{j,i} \cdot f_j\right) \\ & = & \sum_{i \in \{1, \ldots, n\}} x_i \cdot L_2(e_i) \\ & \underset{[\text{theorem: } 11.167]}{=} & L_2\left(\sum_{i \in \{1, \ldots, n\}} x_i \cdot e_i\right) \\ & = & L_2(x) \end{array}$$

proving that $L_1 = L_2$.

surjectivity. Let $M \in \mathcal{M}_{m,n}(F)$ then define $L: X \rightarrow Y$ by

$$L(x) = \sum_{i \in \{1, \ldots, m\}} \left(\sum_{j \in \{1, \ldots, n\}} M_{i,j} \cdot x_j\right) \cdot f_i$$

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

$$\begin{array}{rl} L(x + y) & = \\ \sum_{i \in \{1, \ldots, m\}} \left(\sum_{j \in \{1, \ldots, n\}} M_{i,j} \cdot (x_j + y_j)\right) \cdot f_i & = \\ \sum_{i \in \{1, \ldots, m\}} \left(\sum_{j \in \{1, \ldots, n\}} M_{i,j} \cdot x_j + \sum_{j \in \{1, \ldots, n\}} M_{i,j} \cdot y_j\right) & = \\ \sum_{i \in \{1, \ldots, m\}} \left(\sum_{j \in \{1, \ldots, n\}} M_{i,j} \cdot x_j\right) + \sum_{i \in \{1, \ldots, m\}} \left(\sum_{j \in \{1, \ldots, n\}} M_{i,j} \cdot y_j\right) & = \\ L(x) + L(y) & \end{array}$$

and

$$\begin{array}{rcl} L(\alpha \cdot x) & = & \sum_{i \in \{1, \ldots, m\}} \left(\sum_{j \in \{1, \ldots, n\}} M_{i,j} \cdot (\alpha \cdot x_j)\right) \cdot f_i \\ & = & \sum_{i \in \{1, \ldots, m\}} \left(\alpha \cdot \sum_{j \in \{1, \ldots, n\}} M_{i,j} \cdot x_j\right) \cdot f_i \\ & = & \alpha \cdot \sum_{i \in \{1, \ldots, m\}} \left(\sum_{j \in \{1, \ldots, n\}} M_{i,j} \cdot x_j\right) \cdot f_i \\ & = & \alpha \cdot L(x) \end{array}$$

proving that

$$L \in \mathrm{Hom}(X, Y)$$

Let $i \in \{1, \ldots, n\}$ then 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

$$\mathcal{M}(E, F)^{-1}(M) = L$$

□

Example 11.298. Let F be a field, $n, m \in \mathbb{N}$ and F a field and $E = \{e_i | i \in \{1, \ldots, n\}\} \subseteq F^n$, $F = \{f_i | i \in \{1, \ldots, m\}\} \subseteq F^m$ defined by $(e_i)_j = \delta^n_{i,j}$ and $\{f_i\}_j = \delta^m_{i,j}$ the canonical bases for the vector spaces F^n, F^m over F then for $M \in \mathcal{M}_{m,n}(F)$ we have for $x \in F^n$ 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 $x \in F^n$ $x_i = \sum_{j \in \{1, \ldots, n\}} \delta^n_{i,j} \cdot x_i = \sum_{i \in \{1, \ldots, n\}} x_i \cdot e_i$ so that for $k \in \{1, \ldots, m\}$

$$\begin{aligned}(\mathcal{M}(E, F)^{-1}(M)(x))_k &\underset{[\text{theorem: } 11.297]}{=} \left(\sum_{i \in \{1, \ldots, m\}} \left(\sum_{j \in \{1, \ldots, n\}} M_{i,j} \cdot x_j \right) \cdot f_i \right)_k\\ &\underset{[\text{theorem: } 11.44]}{=} \sum_{i \in \{1, \ldots, m\}} \left(\sum_{j \in \{1, \ldots, n\}} M_{i,j} \cdot x_j \right) \cdot (f_i)_k\\ &= \sum_{i \in \{1, \ldots, m\}} \left(\sum_{j \in \{1, \ldots, n\}} M_{i,j} \cdot x_j \right) \cdot \delta^m_{i,k}\\ &= \sum_{j \in \{1, \ldots, n\}} M_{k,j} \cdot x_j\end{aligned}$$

□

We have the following relation between the composition of linear mappings and the product of matrices.

Theorem 11.299. *Let $n,m,r\in\mathbb{N}$, X,Y,Z finite dimensional vector spaces over a field F with basis $\{e_i|i\in\{1,\ldots,n\}\},\{f_i|i\in\{1,\ldots,m\}\},\{g_i|i\in\{1,\ldots,r\}\}$ defined by distinct families $E=\{e_i\}_{i\in\{1,\ldots,n\}}\subseteq X$, $F=\{f_i\}_{i\in\{1,\ldots,m\}}\subseteq Y$, $G=\{g_i\}_{i\in\{1,\ldots,n\}}$ 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

$$\begin{aligned}
\sum_{k\in\{1,\ldots,r\}}\mathcal{M}(L_2\circ L_1;E,G)_{k,i}\cdot g_k &=\\
(L_2\circ L_1)(e_i) &=\\
L_2(L_1(e_i)) &=\\
L_2\left(\sum_{j\in\{1,\ldots,m\}}\mathcal{M}(L_1;F,G)_{j,i}\cdot f_j\right) &\underset{\text{[theorem: 11.167]}}{=}\\
\sum_{j\in\{1,\ldots,m\}}\mathcal{M}(L_1;F,G)_{j,i}\cdot L_2(f_j) &=\\
\sum_{j\in\{1,\ldots,m\}}\mathcal{M}(L_1;F,G)_{j,i}\cdot\left(\sum_{k\in\{1,\ldots,r\}}\mathcal{M}(L_2;F,G)_{k,j}\cdot g_k\right) &\underset{\text{[theorem: 11.67]}}{=}\\
\sum_{j\in\{1,\ldots,m\}}\left(\sum_{k\in\{1,\ldots,r\}}\mathcal{M}(L_1;F,G)_{j,i}\cdot(\mathcal{M}(L_2;F,G)_{k,j}\cdot g_k)\right) &\underset{\text{[theorem: 11.39]}}{=}\\
\sum_{k\in\{1,\ldots,r\}}\left(\sum_{j\in\{1,\ldots,m\}}(\mathcal{M}(L_2;F,G)_{k,j}\cdot\mathcal{M}(L_1;F,G)_{j,i})\cdot g_k\right) &\underset{\text{[theorem: 11.67]}}{=}\\
\sum_{k\in\{1,\ldots,r\}}\left(\left(\sum_{j\in\{1,\ldots,m\}}(\mathcal{M}(L_2;F,G)_{k,j}\cdot\mathcal{M}(L_1;F,G)_{j,i})\right)\cdot g_k\right) &=\\
\sum_{k\in\{1,\ldots,r\}}(\mathcal{M}(L_2;F.G)\cdot\mathcal{M}(L_1;E,F))_{k,i}\cdot g_k &
\end{aligned}$$

By the uniqueness of the expansion in a basis it follow then that $\forall i\in\{1,...,n\}$ and $\forall k\in\{1,...r\}$ 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 invertibility. 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 11.300. *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 11.301. *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 inverses of M then we have

$$N_1 = N_1 \cdot E = N_1 \cdot (M \cdot N_2) = (N_1 \cdot M) \cdot N_2 = E \cdot N_2 = N_2 \qquad \square$$

The above leads to the following definition

Definition 11.302. *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 11.303. *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 11.304. *Let $n \in \mathbb{N}$, X be a vector space over a field, $\{e_i | i \in \{1, \ldots, n\}\}$ defined by a distinct family $F = \{f_i\}_{i \in \{1, \ldots, n\}} \subseteq X$ and $L \in \operatorname{Hom}(X, X)$ then we have*

$$L \text{ has a inverse } L^{-1} \text{ [or equivalently by [theorem: 2.70] } L \text{ is a isomorphism]}$$
$$\Updownarrow$$
$$\mathcal{M}(L; F, F) \text{ is invertible}$$

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

$$\begin{aligned}
E &\underset{\text{[theorem: 11.296 (4)]}}{=} \mathcal{M}[\mathrm{Id}_X; F, F] \\
&= \mathcal{M}[L \circ L^{-1}; F, F] \\
&\underset{\text{[theorem: 11.299]}}{=} \mathcal{M}(L; F, F) \cdot \mathcal{M}(L^{-1}; F, F) \\
E &\underset{\text{[theorem: 11.296 (4)]}}{=} \mathcal{M}[\mathrm{Id}_X; F, F] \\
&= \mathcal{M}[L^{-1} \circ L; F, F] \\
&\underset{\text{[theorem: 11.299]}}{=} \mathcal{M}(L^{-1}; F, F) \cdot \mathcal{M}(L; F, F)
\end{aligned}$$

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 $K \underset{[\text{theorem: } 11.297]}{=} \mathcal{M}(F,F)^{-1}(\mathcal{M}(L;F,F)^{-1})$ then we have

$$\begin{aligned}
K \circ L & \underset{[\text{theorem: } 11.297]}{=} \\
\mathcal{M}(F,F)^{-1}(\mathcal{M}(F,F)(K \circ L)) & \underset{[\text{theorem:} 11.299]}{=} \\
\mathcal{M}(F,F)^{-1}(\mathcal{M}(F,F)(K) \cdot \mathcal{M}(F,F)(L)) & = \\
\mathcal{M}(F,F)^{-1}(\mathcal{M}(F,F)(\mathcal{M}(F,F)^{-1}(\mathcal{M}(L;F,F)^{-1})) \cdot \mathcal{M}(F,F)(L)) & = \\
\mathcal{M}(F,F)^{-1}(\mathcal{M}(L;F,F)^{-1} \cdot \mathcal{M}(F,F)(L)) & = \\
\mathcal{M}(F,F)^{-1}(\mathcal{M}(L;F,F)^{-1} \cdot \mathcal{M}(L;F,F)) & = \\
\mathcal{M}(F,F)^{-1}(E) & \underset{[\text{theorem: } 11.296]}{=} \\
\mathrm{Id}_X &
\end{aligned}$$

and

$$\begin{aligned}
L \circ K & \underset{[\text{theorem: } 11.297]}{=} \\
\mathcal{M}(F,F)^{-1}(\mathcal{M}(F,F)(L \circ K)) & \underset{[\text{theorem:} 11.299]}{=} \\
\mathcal{M}(F,F)^{-1}(\mathcal{M}(F,F)(L) \cdot \mathcal{M}(F,F)(K)) & = \\
\mathcal{M}(F,F)^{-1}(\mathcal{M}(F,F)(L) \cdot \mathcal{M}(F,F)(\mathcal{M}(F,F)^{-1}(\mathcal{M}(L;F,F)^{-1}))) & = \\
\mathcal{M}(F,F)^{-1}(\mathcal{M}(F,F)(L) \cdot \mathcal{M}(L;F,F)^{-1} \cdot) & = \\
\mathcal{M}(F,F)^{-1}(\mathcal{M}(L;F,F) \cdot \mathcal{M}(L;F,F)^{-1}) & = \\
\mathcal{M}(F,F)^{-1}(E) & \underset{[\text{theorem: } 11.296]}{=} \\
\mathrm{Id}_X &
\end{aligned}$$

which proves that $K \circ L = \mathrm{Id}_X = L \circ K$ so that

$$L \text{ is invertible with inverse } K = \mathcal{M}(F,F)^{-1}(\mathcal{M}(F,F)(L)) \qquad \square$$

We are now ready to define the determinant of a matrix.

Definition 11.305. (Determinant) *Let $n \in \mathbb{N}m$ F a field and $M \in \mathcal{M}_{n,n}(F)$ a square matrix then*

$$\det(M) = \sum_{\sigma \in P_n} \mathrm{sign}(\sigma) \cdot \left(\prod_{i \in \{1,\ldots,n\}} M_{i,\sigma(i)} \right)$$

Another notation for the determinant is $|M|$ so using this notation we have

$$|M| = \sum_{\sigma \in P_n} \mathrm{sign}(\sigma) \cdot \left(\prod_{i \in \{1,\ldots,n\}} M_{i,\sigma(i)} \right)$$

Example 11.306. 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

$$\det(M) = \sum_{\sigma \in \{\mathrm{Id}_{\{1\}}\}} \mathrm{sing}(\sigma) \cdot \left(\prod_{i \in \{1\}} M_{i,\sigma(i)} \right) = \mathrm{sign}(\mathrm{Id}_{\{1\}}) \left(\prod_{i \in \{1\}} M_{i,\mathrm{Id}_{\{1\}}(i)} \right) = M_{11}$$

Definition 11.307. (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 11.308. *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 $k\neq\sigma(k)$. Then

$$\begin{aligned}
\prod_{i\in\{1,\ldots,n\}} M_{i,\sigma(i)} &= \left(\prod_{i\in\{1,\ldots,n\}\setminus\{k\}} M_{i,j}\right)\cdot\left(\prod_{i\in\{k\}} M_{i\sigma(i)}\right)\\
&= \left(\prod_{i\in\{1,\ldots,n\}\setminus\{k\}} M_{i,j}\right)\cdot M_{k,\sigma(k)}\\
&= \left(\prod_{i\in\{1,\ldots,n\}\setminus\{k\}} M_{i,j}\right)\cdot M_{k,k}\cdot\delta_{k,\sigma(k)}\\
&= \left(\prod_{i\in\{1,\ldots,n\}\setminus\{k\}} M_{i,j}\right)\cdot M_{k,k}\cdot 0\\
&= 0
\end{aligned}$$

So that

$$\begin{aligned}
\det(M) &=\\
\sum_{\sigma\in P_n}\left(\mathrm{sign}(\sigma)\cdot\prod_{i\in\{1,\ldots,n\}} M_{i,\sigma(i)}\right)\\
\sum_{\sigma\in P_n\setminus\{\mathrm{Id}_{\{1,\ldots,n\}}\}}\left(\mathrm{sign}(\sigma)\cdot\prod_{i\in\{1,\ldots,n\}} M_{i,\sigma(i)}\right)+\sum_{\sigma\in\{\mathrm{Id}_{\{1,\ldots,n\}}\}}\left(\mathrm{sign}(\sigma)\cdot\prod_{i\in\{1,\ldots,n\}} M_{i,\sigma(i)}\right) &=\\
\sum_{\sigma\in P_n\setminus\{\mathrm{Id}_{\{1,\ldots,n\}}\}}\mathrm{sign}(\sigma)\cdot 0+\sum_{\sigma\in\{\mathrm{Id}_{\{1,\ldots,n\}}\}}\mathrm{sign}(\sigma)\cdot\left(\prod_{i\in\{1,\ldots,n\}} M_{i,\sigma(i)}\right) &=\\
\mathrm{sign}(\mathrm{Id}_{\{1,\ldots,n\}})\cdot\prod_{i\in\{1,\ldots,n\}} M_{i,\mathrm{Id}_{\{1,\ldots,n\}}(i)} &=\\
\prod_{i\in\{1,\ldots,n\}} M_{i,i} &=
\end{aligned}$$

□

The determinant is invariant under the transpose operation on a matrix.

Theorem 11.309. *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

$$(.)^{-1}: P_n\rightarrow P_n \text{ is a bijection}$$

First note that $\langle F,\cdot\rangle$ is a Abelian semi group

$$\prod_{i\in\{1,\ldots,n\}} M_{\sigma(i),i}=\prod_{i\in\{1,\ldots,n\}} M_{\sigma(i),\sigma(\sigma^{-1}(i))}\underset{[\text{theorem: }11.34]}{=}\prod_{i\in\{1,\ldots,n\}} M_{i,\sigma^{-1}(i)} \tag{11.194}$$

So that

$$\begin{aligned}
\det(M^T) &= \sum_{\sigma\in P_n}\operatorname{sign}(\sigma)\cdot\left(\prod_{i\in\{1,\ldots,n\}}(M^T)_{i,\sigma(i)}\right)\\
&= \sum_{\sigma\in P_n}\operatorname{sign}(\sigma)\cdot\left(\prod_{i\in\{1,\ldots,n\}}M_{\sigma(i),i}\right)\\
&\underset{\text{[eq: 11.194]}}{=} \sum_{\sigma\in P_n}\operatorname{sign}(\sigma)\cdot\left(\prod_{i\in\{1,\ldots,n\}}M_{i,\sigma^{-1}(i)}\right)\\
&\underset{\text{[theorem: 11.211]}}{=} \sum_{\sigma\in P_n}\operatorname{sign}(\sigma^{-1})\cdot\left(\prod_{i\in\{1,\ldots,n\}}M_{i,\sigma^{-1}(i)}\right)\\
&= \sum_{\sigma\in P_n}\operatorname{sign}((.)^{-1}(\sigma))\cdot\left(\prod_{i\in\{1,\ldots,n\}}M_{i,(.)^{-1}(\sigma)(i)}\right)\\
&\underset{\text{[eq: 11.194 and theorem: 11.34]}}{=} \sum_{\sigma\in P_n}\operatorname{sign}(\sigma)\cdot\left(\prod_{i\in\{1,\ldots,n\}}M_{i,\sigma(i)}\right)\\
&= \det(M)
\end{aligned}$$

□

Theorem 11.310. *Let $n\in\mathbb{N}$, X a vector space over a field F with characteristic zero with basis $E=\{e_i|i\in\{1,\ldots,n\}\}\subseteq X$ defined by the distinct family $\{e_i\}_{i\in\{1,\ldots,n\}}\subseteq X$ and $L\in\operatorname{Hom}(X,X)$ then we have:*

$$\det(L)=\det(\mathcal{M}(L;E,E))=\det(\mathcal{M}(E,E)(L))$$

Proof. Using [theorem: 11.255] there exists a determinant function Δ in F^n such that $\Delta(e_1,\ldots,e_n)=1$. Then we have

$$\Delta_L(e,\ldots,e_n)\underset{\text{[theorem: 11.267]}}{=}\det(L)\cdot\Delta(e_1,\ldots,e_n)=\det(L) \tag{11.195}$$

Let $M=\mathcal{M}(L,E,E)$ then we have

$$\begin{aligned}
\det(L) &\underset{\text{[eq: 11.195]}}{=} \Delta_L(e,\ldots,e_n)\\
&\underset{\text{def}}{=} \Delta(L(e_1),\ldots,L(e_n))\\
&= \Delta\left(\sum_{i\in\{1,\ldots,n\}}M_{i,1}\cdot e_1,\ldots,\sum_{i\in\{1,\ldots,n\}}M_{i,n}\cdot e_n\right)\\
&\underset{\text{[theorem: 11.250]}}{=} \sum_{\sigma\in P_n}\left(\prod_{i\in\{1,\ldots,n\}}M_{\sigma(i),i}\right)\cdot(\sigma\Delta)(e_1,\ldots,e_n)\\
&\underset{\Delta\text{ is skew-symmetry}}{=} \sum_{\sigma\in P_n}\left(\prod_{i\in\{1,\ldots,n\}}M_{\sigma(i),i}\right)\cdot(\operatorname{sign}(\sigma)\cdot\Delta(e_1,\ldots,e_n))\\
&= \sum_{\sigma\in P_n}\operatorname{sign}(\sigma)\cdot\left(\prod_{i\in\{1,\ldots,n\}}M_{\sigma(i),i}\right)\\
&= \det(M^T)\\
&\underset{\text{[theorem: 11.309]}}{=} \det(M)
\end{aligned}$$

□

We use now the above theorem to prove the following:

Theorem 11.311. *Let $n\in\mathbb{N}$, F is a field of characteristic zero then we have:*

1. *If $M_1,M_2\in\mathcal{M}_{n,n}(F)$ 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 F^n be the vector space over F with the canonical basis $F = \{e_i | i \in \{1, \ldots, n\}\}$ defined by $\{e_i\}_{i \in \{1, \ldots, n\}} \subseteq F^n$ where $\{e_i\}_k = \delta_{i,k}$ for $k \in \{1, \ldots, n\}$ [see: theorem: 11.145].

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:

$$\mathcal{M}(F, F)(L_1) = M_1 \text{ and } \mathcal{M}(F, F)(L_2) = M_2 \tag{11.196}$$

$$\begin{aligned}\mathcal{M}(F, F)(L_1 \circ L_2) &\underset{[\text{theorem: } 11.299]}{=} \mathcal{M}(F, F)(L_1) \cdot \mathcal{M}(F, F)(L_2)\\ &= \mathcal{M}(F, F)(\mathcal{M}[F, F]^{-1}(M_1)) \cdot \mathcal{M}(F, F)(\mathcal{M}[F, F]^{-1}(M_2))\\ &= M_1 \cdot M_2 \end{aligned}\tag{11.197}$$

hence:

$$\begin{aligned}\det(M_1 \cdot M_2) &\underset{[\text{eq: } 11.197]}{=} \det(\mathcal{M}(F, F)(L_1 \circ L_2))\\ &\underset{[\text{theorem: } 11.310]}{=} \det(L_1 \circ L_2)\\ &\underset{[\text{theorem: } 11.271]}{=} \det(L_1) \cdot \det(L_2)\\ &\underset{[\text{eq: } 11.196]}{=} \det(M_1) \cdot \det(M_2)\end{aligned}$$

2. As $\mathcal{M}(F, F)(\mathrm{Id}_{F^n}) \underset{[\text{theorem: } 11.296]}{=} E$ we have

$$\det(E) = \det(\mathcal{M}(F, F)(\mathrm{Id})) \underset{[\text{theorem: } 11.310]}{=} \det(\mathrm{Id}_{F^n}) \underset{[\text{theorem: } 11.271]}{=} 1$$

3. As $M \cdot M^{-1}$ we have $1 = \det(E) = \det(M \cdot M^{-1}) = \det(M) \cdot \det(M^{-1})$ hence $\det(M) \neq 0$ and $(\det(M))^{-1} = \det(M^{-1})$. □

Theorem 11.312. *Let $n \in \mathbb{N}$, F a field of characteristic zero, $M \in \mathcal{M}_{n,n}(F)$, $E = \{e_i | i \in \{1, \ldots, n\}\}$ the canonical basis for F^n [see: theorem: 11.145] 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

$$\begin{aligned}(L(e_i))_k &= (\mathcal{M}(E, E)^{-1}(M))_k\\ &\underset{[\text{example: } 11.298}{=} \sum_{j \in \{1, \ldots, n\}} M_{k,j} \cdot (e_i)_j\\ &= \sum_{j \in \{1, \ldots, n\}} M_{k,j} \cdot \delta_{j,i}\\ &= M_{k,i}\\ &= \mathrm{col}(M, i)_k\end{aligned}\tag{11.198}$$

So

$$\begin{aligned}\Delta(\mathrm{col}(M, 1), \ldots, \mathrm{col}(M, n)) &\underset{[\text{eq: } 11.198]}{=} \Delta(L(e_1), \ldots, L(e_n))\\ &= \Delta_L(e_1, \ldots, e_n)\\ &\underset{[\text{definition: } 11.267]}{=} \det(L) \cdot \Delta(e_1, \ldots, e_n)\\ &= \det(L) \cdot 1\\ &= \det(L)\\ &\underset{[\text{theorem: } 11.310]}{=} \det(\mathcal{M}(E, E)(L))\\ &= \det(M)\end{aligned}$$

proving that

$$\det(M) = \Delta(\operatorname{col}(M, 1), \ldots, \operatorname{col}(M, n)) \tag{11.199}$$

For the rows note that $\forall i, j \in \{1, \ldots, n\}$ we have

$$\operatorname{col}(M^T, i)_j = (M^T)_{j,i} = M_{i,j} = \operatorname{row}(M, i)_j$$

so that

$$\operatorname{col}(M^T, i) = \operatorname{row}(M, i) \tag{11.200}$$

and

$$\begin{aligned}\det(M) &\underset{\text{[theorem: 11.309]]}}{=} \det(M^T)\\ &\underset{\text{[eq: 11.199]}}{=} \Delta(\operatorname{col}(M, 1), \ldots, \operatorname{col}(M, n))\\ &\underset{\text{[eq: 11.200]}}{=} \Delta(\operatorname{row}(M, 1), \ldots, \operatorname{row}(M, n))\end{aligned}$$

giving finally

$$\Delta(\operatorname{row}(M, 1), \ldots, \operatorname{row}(M, n)) = \det(M) \qquad \square$$

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 11.313. *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 11.314. *Let $n \in \mathbb{N}$, F a field, $M \in \mathcal{M}_{n,n}(F)$ and $\sigma \in P_n$ then:*

1. $\det(M_\sigma) = \operatorname{sign}(\sigma) \cdot \det(M)$
2. $\det(M^\sigma) = \operatorname{sign}(\sigma) \cdot \det(M)$

Proof. Let $\{e_i | i \in \{1, \ldots, n\}\}$ the canonical basis for F^n [see: theorem: 11.145] then by [theorem: 11.255] 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 $\operatorname{row}(M_\sigma, i)_j = (M_\sigma)_{i,j} = M_{\sigma(i),j} = \operatorname{row}(M, \sigma(i))_j$ proving that

   $$\operatorname{row}(M_\sigma, i) = \operatorname{row}(M, \sigma(i)) \tag{11.201}$$

   so that

   $$\begin{aligned}\det(M_\sigma) &\underset{\text{[theorem: 11.312]}}{=} \Delta(\operatorname{row}(M_\sigma, 1), \ldots, \operatorname{row}(M_\sigma, n))\\ &\underset{\text{[eq: 11.201]}}{=} \Delta(\operatorname{row}(M, \sigma(1)), \ldots, \operatorname{row}(M, \sigma(i)))\\ &\underset{\text{skew}-\text{symmetry}}{=} \operatorname{sign}(\sigma) \cdot \Delta(\operatorname{row}(M, 1), \ldots, \operatorname{row}(M, n))\\ &\underset{\text{[theorem: 11.312]}}{=} \operatorname{sign}(\sigma) \cdot \det(M)\end{aligned}$$

2. $\forall i, j \in \{1, \ldots, n\}$ we have $\operatorname{col}(M^\sigma, i)_j = (M^\sigma)_{j,i} = M_{j,\sigma(i)} = \operatorname{col}(M, \sigma(i))_j$ proving that

   $$\operatorname{col}(M_\sigma, i) = \operatorname{col}(M, \sigma(i)) \tag{11.202}$$

so that

$$\begin{aligned}\det(M^{\sigma}) &\underset{[\text{theorem: } 11.312]}{=} \Delta(\mathrm{col}(M_{\sigma},1),\ldots,\mathrm{col}(M_{\sigma},n))\\ &\underset{[\text{eq: } 11.201]}{=} \Delta(\mathrm{col}(M,\sigma(1)),\ldots,\mathrm{col}(M,\sigma(i)))\\ &\underset{\text{skew}-\text{symmetry}}{=} \mathrm{sign}(\sigma)\cdot\Delta(\mathrm{col}(M,1),\ldots,\mathrm{col}(M,n))\\ &\underset{[\text{theorem: } 11.312]}{=} \mathrm{sign}(\sigma)\cdot\det(M)\end{aligned}$$

□

The above allows us to construct a test to determine if the determinant of a matrix is zero.

Corollary 11.315. *Let $n\in\mathbb{N}$, F a field with characteristic zero and $M\in\mathcal{M}_{n,n}(F)$ then we have*

$$\begin{gathered}\det(M)=0\\ \Updownarrow\\ \mathrm{rank}(M)<n\end{gathered}$$

Proof.

$\Rightarrow$**.** We prove this by contradiction, so assume that $n\leqslant\mathrm{rank}(M)$ then as by [definition: 11.291] $\mathrm{rank}(M)\leqslant n$ we must have that $\mathrm{rank}(M)=n\underset{[\text{definitions:} 11.289, 11.291]}{=}\dim(\mathrm{span}(\mathrm{cols}(M)))$. Using [theorem: 11.134] it follows that $\mathrm{cols}(M)=\{\mathrm{col}(M,i)|i\in\{1,\ldots,n\}\}$ is a basis for $\mathrm{span}(\mathrm{cols}(M))\subseteq F^n$, hence $\{\mathrm{col}(M,i)|i\in\{1,\ldots,n\}\}$ is linear independent. Using [theorem: 11.132] and the fact that $\mathrm{card}(\mathrm{col}(M,i)|i\in\{1,\ldots,n\})=\dim(\mathrm{span}(\mathrm{cols}(M)))=n$ it follows that

$$\{\mathrm{col}(M,i)|i\in\{1,\ldots,n\}\}\text{ is a basis for }F^n$$

By [theorem: 11.255] there exists a determinant function Δ in F^n such that

$$\Delta(\mathrm{cols}(M,1),\ldots,\mathrm{cols}(M,n))=1$$

Using [theorem: 11.312] we have $\det(M)=\Delta(\mathrm{cols}(M,1),\ldots,\mathrm{cols}(M,n))$ so that $\det(M)=1\neq 0$ contradicting the hypothesis $\det(M)=0$. Hence we must have that $\mathrm{rank}(M)<n$.

$\Leftarrow$**.** Let $\{e_i|i\in\{1,\ldots,n\}\}$ be the canonical basis for F^n, then by [theorems: 11.255,11.312] there exists a determinant function in F^n such that

$$\det(M)=\Delta(\mathrm{col}(M,1),\ldots,\mathrm{col}(M,n))$$

By the hypothesis we have $\dim(\mathrm{span}(\mathrm{cols}(M)))=\mathrm{rank}(M)<n$. We have now either:

$\{\mathrm{col}(M,i)\}_{i\in\{1,\ldots,n\}}$ **is distinct.** Then

$$\mathrm{card}(\mathrm{cols}(M))=\mathrm{card}(\{\mathrm{col}(M,i)|i\in\{1,\ldots,n\}\})\underset{[\text{theorem: } 11.109]}{=}n$$

If $\mathrm{cols}(M)$ is linear independent then $\mathrm{cols}(M)$ is a basis of $\mathrm{span}(\mathrm{cols}(M))$ so that

$$n=\dim(\mathrm{span}(\mathrm{cols}(M)))=\mathrm{rank}(M)<n$$

a contradiction. Hence we have that $\{\mathrm{col}(M,i)|i\in\{1,\ldots,n\}\}=\mathrm{cols}(M)$ is linear dependent. Then by [theorem: 11.248] $\Delta(\mathrm{col}(M,1),\ldots,\mathrm{col}(M,n))=0$ so that in this case

$$\det(M)=0$$

$\{\mathrm{col}(M,i)\}_{i\in\{1,\ldots,n\}}$ **is not distinct.** Then there exists $i,j\in\{1,\ldots,n\}$ with $i\neq j$ and $\mathrm{col}(M,i)=\mathrm{col}(M,j)$ so that by [theorem: 11.248] $\Delta(\mathrm{col}(M,1),\ldots,\mathrm{col}(M,n))=0$ and we have

$$\det(M)=0$$

□

We will now develop the necessary tools to calculate the inverse of a matrix (if the matrix is invertible).

Definition 11.316. *Let $n\in\mathbb{N}\backslash\{1\}$, X a set and $a\in X^{-1}$ then we define*

$$[+a]\colon X^{n-1}\rightarrow X^n \text{ by } [+a](x) \text{ where } \forall i\in\{1,\ldots,n\}\ ([+a](x))_i=\begin{cases} a \text{ if } i=1\\ x_{i-1} \text{ if } i\in\{2,\ldots,n\}\end{cases}$$

In other words $[+a](x_1,\ldots,x_{n-1})=(a,x_1,\ldots,x_n)$

Lemma 11.317. *Let $n\in\mathbb{N}\backslash\{1\}$, F a field and $0\in F$ the additive neutral element then*

$$[+0]\in\operatorname{Hom}(F^{n-1}\rightarrow F^n)$$

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

$$[+0](\alpha\cdot x)=\begin{cases}0 \text{ if } i=1\\ (\alpha\cdot x)_{i-1} \text{ if } i\in\{2,\ldots,n+1\}\end{cases}=\begin{cases}\alpha\cdot 0 \text{ if } i=1\\ \alpha\cdot x_{i-1} \text{ if } i\in\{2,\ldots,n+1\}\end{cases}=\alpha\cdot([+0](x))\quad\square$$

We use the above now to generate a new determinant function from a existing determinant function.

Theorem 11.318. *Let $n\in\mathbb{N}$, F a field with characteristic zero, $\{e_i|i\in\{1,\ldots,n\}\}$ the canonical basis for F^n [see: theorem: 11.145] then if*

$$\Delta\colon(F^n)^n\rightarrow F$$

is a determinant function then if we define

$$\Delta^-\colon(F^{n-1})^{n-1}\rightarrow F \text{ where } \Delta^-(x_1,\ldots,x_{n-1})=\Delta(e_1,[+0](x_1),\ldots,[+0](x_{n-1}))$$

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

$$(a_1,\ldots,a_{i-1},x+y,a_{i+1},\ldots,a_{n-1})\in(F^{n-1})^{n-1}$$

we have

$$\begin{aligned}\Delta^-(a_1,\ldots,a_{i-1},x+y,a_{i+1},\ldots,a_{n-1}) &=\\ \Delta(e_1,[+0](a_1),\ldots,[+0](a_{i-1}),[+0](x+y),[+0](a_{i+1}),\ldots,[+0](a_n)) &=\\ \Delta(e_1,[+0](a_1),\ldots,[+0](a_{i-1}),[+0](x)+[+0](y),[+0](a_{i+1}),\ldots,[+0](a_n)) &=\\ \Delta(e_1,[+0](a_1),\ldots,[+0](a_{i-1}),[+0](x),[+0](a_{i+1}),\ldots,[+0](a_n))+\Delta(e_1,[+0](a_1),\ldots,&\\ [+0](a_{i-1}),[+0](y),[+0](a_{i+1}),\ldots,[+0](a_n)) &=\\ \Delta^-(a_1,\ldots,a_{i-1},x,a_{i+1},\ldots,a_{n-1})+\Delta^-(a_1,\ldots,a_{i-1},y,a_{i+1},\ldots,a_{n-1})&\end{aligned}$$

and for

$$(a_1,\ldots,a_{i-1},\alpha\cdot x,a_{i+1},\ldots,a_{n-1})\in(F^{n-1})^{n-1}$$

we have

$$\begin{array}{r}\Delta^-(a_1,\ldots,a_{i-1},\alpha\cdot x,a_{i+1},\ldots,a_{n-1}) = \\ \Delta(e_{1,}[+0](a_1),\ldots,[+0](a_{i-1}),[+0](\alpha\cdot x),[+0](a_{i+1}),\ldots,[+0](a_n)) = \\ \alpha\cdot\Delta(e_1,[+0](a_1),\ldots,[+0](a_{i-1}),[+0](x),[+0](a_{i+1}),\ldots,[+0](a_n)) = \\ \alpha\cdot\Delta^-(a_1,\ldots,a_{i-1},x,a_{i+1},\ldots,a_{n-1})\end{array}$$

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

$$([+0](a_i))_k=\left\{\begin{array}{l}0\text{ if k=1}\\(a_i)_{k-1}\text{ if }k\in\{2,\ldots,n\}\end{array}\right.=\left\{\begin{array}{l}0\text{ if k=1}\\(a_j)_{k-1}\text{ if }k\in\{2,\ldots,n\}\end{array}\right.=([+0](a_j))_k$$

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: 11.248] 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

$$\begin{array}{rcl}([+0](e_i))_j & = & \left\{\begin{array}{l}0\text{ if }j=1\\(e_i)_{j-1}\text{ if }i\in\{2,\ldots,n\}\end{array}\right.\\ & = & \left\{\begin{array}{l}0\text{ if }j=1\\\delta_{i,j-1}\text{ if }i\in\{2,\ldots,n\}\end{array}\right.\\ & \underset{i+1=j\Rightarrow i=j-1}{=} & \delta_{i+1,j}\\ & = & (e_{i+1})_j\end{array}$$

proving that $[+0](e_i)=e_{i+1}$ so that

$$\Delta^-(e_1,\ldots,e_{n-1})=\Delta(e_1,[+0](e_1),\ldots,[+0](e_{n-1}))=\Delta(e_1,e_2,\ldots,e_n)=1 \qquad \square$$

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 11.319. *Let $n\in\mathbb{N}\setminus\{1\}=\{2,\ldots,\infty\}$, F a field and $i,j\in\{1,\ldots,n-1\}$ then we define:*

$$[i\boxplus j]:\mathcal{M}_{n,n}(F)\rightarrow\mathcal{M}_{n-1,n-1}(F)\text{ where }([i\boxplus j](M))_{k,l}=\left\{\begin{array}{l}M_{k,l}\text{ if }1\leqslant k<i\wedge 1\leqslant l<j\\M_{k+1,l}\text{ if }i\leqslant k\leqslant n-1\wedge 1\leqslant l<j\\M_{k,l+1}\text{ if }1\leqslant k<i\wedge j\leqslant l\leqslant n-1\\M_{k+1,l+1}\text{ if }i\leqslant k\leqslant n-1\wedge j\leqslant l\leqslant n-1\end{array}\right.$$

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 11.320.

$$[2\boxplus 3]\left(\begin{array}{cccc}1&2&3&4\\5&6&7&8\\10&20&30&40\\50&60&70&80\end{array}\right)=\left(\begin{array}{ccc}1&2&4\\10&20&40\\50&60&80\end{array}\right)$$

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 11.321. *Let $n \in \mathbb{N} \backslash \{1\}$, 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:

$$\begin{array}{rcl} 1 \leqslant k < i, 1 \leqslant l < k \Rightarrow ([i \boxplus j](M^T))_{k,l} & = & (M^T)_{k.l} \\ & = & M_{l,k} \\ & = & ([j \boxplus i]M)_{l,k} \\ & = & (([j \boxplus i]M)^T)_{l,k} \\ i \leqslant k \leqslant n-1, l < j \Rightarrow ([i \boxplus j](M^T))_{k,l} & = & (M^T)_{k+1,l} \\ & = & M_{l,k+1} \\ & = & ([j \boxplus i]M)_{l,k} \\ & = & (([j \boxplus i]M)^T)_{l,k} \\ k < i, j \leqslant l \leqslant n-1 \Rightarrow ([i \boxplus j](M^T))_{k,l} & = & (M^T)_{k,l+1} \\ & = & M_{l+1,k} \\ & = & ([j \boxplus i]M)_{l,k} \\ & = & (([j \boxplus i]M)^T)_{l,k} \\ i \leqslant k \leqslant n-1, j \leqslant l \leqslant n-1 \Rightarrow ([i \boxplus j](M^T))_{k,l} & = & (M^T)_{k+1,l+1} \\ & = & M_{l+1,k+1} \\ & = & ([j \boxplus i]M)_{l,k} \\ & = & (([j \boxplus i]M)^T)_{l,k} \end{array}$$

□

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 a 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 11.322. *Let $n \in \mathbb{N} \backslash \{1\}$, F a field and $M \in \mathcal{M}_{n,n}(F)$ then we define*

1. *Let $m \in \{2, \ldots, n\}$ 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. *Let $m \in \{1, \ldots, n-1\}$ 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 11.323. *Let $n \in \mathbb{N} \backslash \{1\}$, 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* $m_1\in\{1,...,n-1\}$ *and* $m_2\in\{1,...,m_2-m_1-1\}$ *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

$$\begin{aligned}([1\boxplus 1]M)_{i,j} &\underset{1\leqslant i,j\wedge[\text{definition: }11.319]}{=} M_{i+1,j+1}\\ &= ([>1]M)_{i,j}\end{aligned}$$

so we have $[1\boxplus 1]M=[>1]M$

2. $\forall i,j\in\{1,\ldots,n-1\}$ we have

$$\begin{aligned}([n\boxplus n]M)_{i,j} &\underset{1\leqslant i,j\leqslant n-1<n\wedge[\text{definition: }11.319]}{=} M_{i,j}\\ &\underset{1\leqslant i,j\leqslant n-1}{=} ([<n]M)_{i,j}\end{aligned}$$

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

$$[>(m_1+m_2)](M)\in\mathcal{M}_{n-(m_1+m_2),n-(m_1+m_2)}(F)\in\mathcal{M}_{n-(m_1+m_2),n-(m_1+m_2)}(F)$$

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

$$\begin{aligned}(([<m](M))^T)_{i,j} &= ([<m](M))_{j,i}\\ &= M_{j,i}\\ &= (M^T)_{j,i}\\ &= ([<m](M^T))_{i,j}\end{aligned}$$

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

$$(-1)^{i+j} \cdot \left|\begin{array}{cccccc} m_{1,1} & \ldots & m_{1,i-1} & m_{1,i+1} & \ldots & m_{1,n}\\ \vdots & & \vdots & \vdots & & \vdots\\ m_{j-1,1} & \ldots & m_{j-1,i-1} & m_{j-1,i+1} & \ldots & m_{j-1,n}\\ m_{j+1,1} & \ldots & m_{j+1,i-1} & m_{j+1}, i+1 & \ldots & m_{j+1,n}\\ \vdots & & \vdots & \vdots & & \vdots\\ m_{n,1} & \cdots & m_{n,i-1} & m_{n,i+1} & \cdots & m_{n,n} \end{array}\right|$$

The proof is done first for some simpler cases, first we start with proving 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 11.324. *Let $n \in \mathbb{N} \setminus \{1\}$ and $M \in M_{n,n}(F)$ such that $\forall i \in \{1, \ldots, n\}$*

$$\operatorname{row}(M,1)_i = \delta_{i,1} = \operatorname{col}(M,1)_i$$

then

$$\det(M) = \det([1 \boxplus 1](M))$$

Proof. Let $\{e_i | i \in \{1, \ldots, n\}\}$ be the canonical basis of F^n. As $\forall j \in \{1, \ldots, n\}$ we have $\operatorname{row}(M,1)_i = \delta_{i,1} = (e_1)_i$ it follows that

$$\operatorname{row}(M,1) = e_1$$

Further for $i \in \{2, \ldots, n\}$ we have

$$\operatorname{row}(M, i)_1 = M_{i,1} = \operatorname{col}(M, 1)_i = \delta_{1,i} \underset{i \neq 1}{=} 0 = ([+0](\operatorname{row}([1 \boxplus 1](M), i)))_1$$

and if $j \in \{2, \ldots, n-1\}$ we have

$$\operatorname{row}(M, i)_j = M_{i,j} = ([1 \boxplus 1](M))_{i-1,j-1} = ([+0](\operatorname{row}([1 \boxplus 1](M), i)))_j$$

proving that for $i \in \{2, \ldots, n\}$ that

$$\begin{aligned}
\det(M) &\underset{[\text{theorem: } 11.312]}{=} \Delta(\operatorname{row}(M, 1), \ldots, \operatorname{row}(M, n)) \\
&= \Delta(e_1, \operatorname{row}(M, 2), \ldots, \operatorname{row}(M, n)) \\
&= \Delta(e_1, [+0](\operatorname{row}([1 \boxplus 1])(M), 1), \ldots, [+0](\operatorname{row}([1 \boxplus 1](M), n-1))) \\
&= \Delta^{-}(\operatorname{row}([1 \boxplus 1](M), 1), \ldots, \operatorname{row}([1 \boxplus](M), n-1)) \qquad (11.203)
\end{aligned}$$

As Δ is a determinant function with $\Delta(e_1, \ldots, e_n)$ we have by [theorem: 11.318] that

$$\Delta^{-} \text{ is a matrix function and } \Delta^{-}(e_1, \ldots, e_{n-1}) = 1$$

which as $\{e_i | i \in \{1, \ldots, n-1\}\}$ is the canonical basis of F^{n-1} proves by [theorem: 11.312] that

$$\Delta^{-}(\operatorname{row}([1 \boxplus 1](M), 1), \ldots, \operatorname{row}([1 \boxplus](M), n-1)) = \det([1 \boxplus 1](M))$$

which combined with [eq: 11.203] proves that

$$\det(M) = \det([1 \boxplus 1](M))$$ □

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 11.325. *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* $\operatorname{row}(M, 1)_i = \delta_{i,j}$ *and* $\operatorname{col}(M, j)_i = \delta_{i,1}$ *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 $\operatorname{row}(M, 1)_i = \delta_{i,1} = \operatorname{col}(M, 1)_i$ so that the conditions of the previous lemma [lemma: 11.324] 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, so that M' is defined by [see definition: 11.212]

$$\forall k, l \in \{1, \ldots, n\} \; M'_{k,l} = M_{k,(j \underset{n}{\rightsquigarrow} 1)(l)}$$

then using [definition: 11.313] we have that

$$M' = M^{\left(j \underset{n}{\leadsto} 1\right)} \tag{11.204}$$

then we have $\forall k \in \{1, \ldots n\}$ we have

$$\begin{array}{lcl}
\operatorname{row}(M', 1)_k & = & M'_{1,k} \\
 & = & M_{1,\left(j \underset{n}{\leadsto} 1\right)(k)} \\
 & \underset{1<j\wedge[\text{definition: }11.212]}{=} & \begin{cases} M_{1,k} \text{ if } 1 \leqslant k < 1 \\ M_{1,j} \text{ if } k = 1 \\ M_{1,k-1} \text{ if } 1 < k \leqslant j \\ M_{1,k} \text{ if } j < k \leqslant n \end{cases} \\
 & = & \begin{cases} M_{1,j} \text{ if } k = 1 \\ M_{1,k-1} \text{ if } 1 < k \leqslant j \\ M_{1,k} \text{ if } j < k \leqslant n \end{cases} \\
 & = & \begin{cases} \operatorname{col}(M, j)_1 \text{ if } k = 1 \\ \operatorname{row}(M, 1)_{k-1} \text{ if } 1 < k \leqslant j \\ \operatorname{row}(M, 1)_k \text{ if } j < k \leqslant n \end{cases} \\
 & = & \begin{cases} \delta_{k,1} \text{ if } k = 1 \\ \delta_{k-1,j} \text{ if } 1 < k \leqslant j \\ \delta_{k,1} \text{ if } j < k \leqslant n \end{cases} \\
 & = & \begin{cases} 1 \text{ if } k = 1 \\ \delta_{k-1,j} \text{ if } 1 < k \leqslant j \\ \delta_{k,j} \text{ if } j < k \leqslant n \end{cases} \\
 & \underset{k\leqslant j\Rightarrow k-1<j\Rightarrow k-1\neq j}{=} & \begin{cases} 1 \text{ if } k = 1 \\ 0 \text{ if } 1 < k \leqslant j \\ \delta_{k,j} \text{ if } j < k \leqslant n \end{cases} \\
 & \underset{j<k\Rightarrow j\neq k}{=} & \begin{cases} 1 \text{ if } k = 1 \\ 0 \text{ if } 1 < k \leqslant j \\ 0 \text{ if } j < k \leqslant n \end{cases} \\
 & = & \delta_{k,1}
\end{array}$$

and we have also

$$\begin{array}{lcl}
\operatorname{col}(M', 1)_k & = & M'_{k,1} \\
 & = & M_{k,\left(j \underset{n}{\leadsto} 1\right)(1)} \\
 & \underset{1=1\wedge[\text{definition: }11.212]}{=} & M_{k,j} \\
 & = & \operatorname{col}(M, j)_k \\
 & = & \delta_{k,1}
\end{array}$$

Hence M' satisfies the conditions of the previous lemma [lemma: 11.324] giving us that

$$\det(M') = \det([1 \boxplus 1](M')) \tag{11.205}$$

Now if $k, l \in \{1, \ldots, n-1\}$ then $1 \leqslant k, l \leqslant n-1$ it follows from [definition: 11.319] that

$$([1 \boxplus 1](M'))_{k,l} = M'_{k+1,l+1} \tag{11.206}$$

Take $k\in\{1,\ldots,n-1\}$ then for $l\in\{1,\ldots,n-1\}$ we have the following possible cases:

$\boldsymbol{1\leqslant l<j.}$ Then

$$\begin{aligned}
M'_{k+1,l+1} &= M_{k+1,(j\underset{n}{\rightsquigarrow}1)(l+1)}\\
&\underset{l<j\Rightarrow 1<l+1\leqslant j\wedge[\text{definition: }11.212]}{=} M_{k+1,l}\\
&\underset{1\leqslant k\leqslant n-1,l<j\wedge[\text{definition: }11.319]}{=} ([1\boxplus j]M)_{k,l}
\end{aligned}$$

$\boldsymbol{j\leqslant l\leqslant n-1.}$ Then

$$\begin{aligned}
M'_{k+1,l+1} &= M_{k+1,(j\underset{n}{\rightsquigarrow}1)(l+1)}\\
&\underset{j\leqslant l\Rightarrow j<l+1\leqslant n\wedge[\text{definition: }11.212]}{=} M_{k+1,l+1}\\
&\underset{1\leqslant k\leqslant n-1\wedge j\leqslant l\leqslant n-1\wedge[\text{definition: }11.319]}{=} ([1\boxplus j](M))_{k,l}
\end{aligned}$$

So in all cases we gave $M'_{k+1,l+1}=([1\boxplus j](M))_{k,l}$ or combining this with [eq: 11.206] 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: 11.205] gives

$$\det(M')=\det([1\boxplus j](M)) \tag{11.207}$$

Further

$$\begin{aligned}
\det(M') &\underset{[\text{eq: }11.204]}{=} \det\big(M^{(j\underset{n}{\rightsquigarrow}1)}\big)\\
&\underset{[\text{theorem: }11.314]}{=} \operatorname{sign}\big(\big(j\underset{n}{\rightsquigarrow}1\big)\big)\cdot\det(M)\\
&\underset{[\text{theorem: }11.215]}{=} (-1)^{j-1}\cdot\det(M)
\end{aligned}$$

As $(-1)^{j+1}\cdot(-1)^{j-1}=(-1)^{2\cdot j}=1$, multiplying both sides of the above equation with $(-1)^{1+j}$ gives

$$\det(M)=(-1)^{j+1}\cdot\det(M')\underset{[\text{eq: }11.207]}{=}(-1)^{1+j}\cdot\det([1\boxplus j](M))$$

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

$$(-1)^{i+j}\cdot\begin{vmatrix}
m_{1,1} & \ldots & m_{1,i-1} & m_{1,i+1} & \ldots & m_{1,n}\\
\vdots & & \vdots & \vdots & & \vdots\\
m_{j-1,1} & \ldots & m_{j-1,i-1} & m_{j-1,i+1} & \ldots & m_{j-1,n}\\
m_{j+1,1} & \ldots & m_{j+1,i-1} & m_{j+1},i+1 & \ldots & m_{j+1,n}\\
\vdots & & \vdots & \vdots & & \vdots\\
m_{n,1} & \cdots & m_{n,i-1} & m_{n,i+1} & \cdots & m_{n,n}
\end{vmatrix}$$

Theorem 11.326. *Let $n\in\mathbb{N}\backslash\{1\}$, 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 $\forall k \in \{1,...,n\}$ we have $\mathrm{row}(M, 1)_k = \delta_{j,k}$ and $\mathrm{col}(M, j)_k = \delta_{i,k}$ so that the conditions for the previous lemma [lemma: 11.325] 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* ∈ {2, ..., *n*}.** We move not the i-the row to before the first row, so define M' by

  $$M' = M_{(i \underset{n}{\rightsquigarrow} 1)} \tag{11.208}$$

  or using [definition: 11.313] we have

  $$\forall k, l \in \{1, \ldots, n\} \text{ that } M'_{k,l} = M_{(i \underset{n}{\rightsquigarrow} 1)(k),l} \tag{11.209}$$

  Let $k \in \{1, \ldots, n\}$ then we have

  $$\begin{array}{lcl} \mathrm{row}(M', 1)_k & = & M'_{1,k} \\ & = & M_{(i \rightsquigarrow 1_n)(1),k} \\ & \underset{1=1 \wedge [\text{definition: } 11.212]}{=} & M_{i,k} \\ & = & \mathrm{row}(M, i)_k \\ & = & \delta_{j,k} \end{array}$$

  and

  $$\begin{array}{lcl} \mathrm{col}(M', j)_k & = & M'_{k,j} \\ & = & M_{(i \rightsquigarrow 1_n)(k),j} \\ & \underset{[\text{definition: } 11.212]}{=} & \begin{cases} M_{k,j} \text{ if } 1 \leqslant k < 1 \\ M_{i,j} \text{ if } k = 1 \\ M_{k-1,j} \text{ if } 1 \leqslant k \leqslant i \\ M_{k,j} \text{ if } i < k \leqslant \mathrm{nn} \end{cases} \\ & = & \begin{cases} M_{i,j} \text{ if } k = 1 \\ M_{k-1,j} \text{ if } 1 \leqslant k \leqslant i \\ M_{k,j} \text{ if } i < k \leqslant n \end{cases} \\ & = & \begin{cases} \mathrm{row}(M, i)_j \text{ if } k = 1 \\ \mathrm{col}(M, j)_{k-1} \text{ if } 1 \leqslant k \leqslant i \\ \mathrm{col}(M, j)_k \text{ if } i < k \leqslant n \end{cases} \\ & = & \begin{cases} \delta_{j,j} \text{ if } k = 1 \\ \delta_{i,k-1} \text{ if } 1 \leqslant k \leqslant i \\ \delta_{i,k} \text{ if } i < k \leqslant n \end{cases} \\ & = & \begin{cases} 1 \text{ if } k = 1 \\ \delta_{i,k-1} \text{ if } 1 \leqslant k \leqslant i \\ \delta_{i,k} \text{ if } i < k \leqslant n \end{cases} \\ & \underset{k \leqslant i \Rightarrow k-1 < i \Rightarrow k-1 \neq i}{=} & \begin{cases} 1 \text{ if } k = 1 \\ 0 \text{ if } 1 \leqslant k \leqslant i \\ \delta_{i,k} \text{ if } i < k \leqslant n \end{cases} \\ & \underset{i < k \Rightarrow i \neq k}{=} & \begin{cases} 1 \text{ if } k = 1 \\ 0 \text{ if } 1 \leqslant k \leqslant i \\ 0 \text{ if } i < k \leqslant n \end{cases} \\ & = & \delta_{1,k} \end{array}$$

So the conditions for the previous lemma [lemma: 11.325] are satisfied for M', hence

$$\det(M') = (-1)^{1+j} \cdot \det([1 \boxplus j](M')) \tag{11.210}$$

Now if $k, l \in \{1, \ldots, n-1\}$ we have the following possibilities for k, l:

$\boldsymbol{1 \leqslant k < i \wedge 1 \leqslant l < j.}$ Then

$$\begin{aligned}
([1 \boxplus j](M'))_{k,l} &\underset{1 \leqslant k \leqslant n-1 \wedge 1 \leqslant l < j \wedge [\text{definition: } 11.319]}{=} M'_{k+1,l} \\
&= M_{(i \underset{n}{\rightsquigarrow} 1)(k+1),l} \\
&\underset{1 \leqslant k < i \Rightarrow 1 < k+1 \leqslant i \wedge [\text{theorem: } 11.212]}{=} M_{k,l} \\
&\underset{1 \leqslant k \leqslant i \wedge 1 \leqslant l < j \wedge [\text{definition: } 11.319]}{=} ([i \boxplus j](M))_{k.l}
\end{aligned}$$

$\boldsymbol{i \leqslant k \leqslant n-1 \wedge 1 \leqslant l < j.}$ Then

$$\begin{aligned}
([1 \boxplus j](M'))_{k,l} &\underset{1 \leqslant k \leqslant n-1 \wedge 1 \leqslant l < j \wedge [\text{definition: } 11.319]}{=} M'_{k+1.l} \\
&= M_{(i \underset{n}{\rightsquigarrow} 1)(k+1),l} \\
&\underset{i \leqslant k < n-1 \Rightarrow i < k+1 \leqslant n \wedge [\text{theorem: } 11.212]}{=} M_{k+1,l} \\
&\underset{i \leqslant k \leqslant n-1 \wedge 1 \leqslant l < j \wedge [\text{definition: } 11.319]}{=} ([i \boxplus j](M))_{k,l}
\end{aligned}$$

$\boldsymbol{1 \leqslant k < i \wedge j \leqslant l < n-1.}$ Then

$$\begin{aligned}
([1 \boxplus j](M'))_{k,l} &\underset{1 \leqslant k \leqslant n-1 \wedge j \leqslant l < n-1 \wedge [\text{definition: } 11.319]}{=} M'_{k+1,l+1} \\
&= M_{(i \underset{n}{\rightsquigarrow} 1)(k+1),l+1} \\
&\underset{1 \leqslant k < i \Rightarrow 1 < k+1 \leqslant i \wedge [\text{theorem: } 11.212]}{=} M_{k,l+1} \\
&\underset{1 \leqslant k < i \wedge j \leqslant l < n-1 \wedge [\text{definition: } 11.319]}{=} ([i \boxplus j](M))_{k,l}
\end{aligned}$$

$\boldsymbol{i \leqslant k \leqslant n-1 \wedge j \leqslant l \leqslant n-1.}$ Then

$$\begin{aligned}
([1 \boxplus j](M'))_{k,l} &\underset{1 \leqslant k \leqslant n-1 \wedge j \leqslant l < n-1 \wedge [\text{definition: } 11.319]}{=} M'_{k+1,l+1} \\
&= M_{(i \underset{n}{\rightsquigarrow} 1)(k+1),l+1} \\
&\underset{i \leqslant k < n-1 \Rightarrow i < k+1 \leqslant n \wedge [\text{theorem: } 11.212]}{=} M_{k+1,l+1} \\
&\underset{i \leqslant k \leqslant n-1 \wedge j \leqslant l < n-1 \wedge [\text{definition: } 11.319]}{=} ([i \boxplus j](M))_{k,l}
\end{aligned}$$

So we have proved that $\forall k, l \in \{1, \ldots, n-1\}$ $([1 \boxplus j](M'))_{k.l} = ([i \boxplus j](M))$ giving

$$[1 \boxplus j](M') = [i \boxplus j](M)$$

Substituting the above in [eq: 11.210] gives then

$$\det(M') = (-1)^{1+j} \cdot \det([i \boxplus j](M)) \tag{11.211}$$

Now

$$\begin{aligned}
\det(M') &\underset{[\text{eq: } 11.208]}{=} \det\left(M_{(i \underset{n}{\rightsquigarrow} 1)}\right) \\
&\underset{[\text{theorem: } 11.314]}{=} \operatorname{sign}\left(\left(i \underset{n}{\rightsquigarrow} 1\right)\right) \cdot \det(M) \\
&\underset{[\text{theorem: } 11.215]}{=} (-1)^{i-1} \cdot \det(M)
\end{aligned}$$

Substituting the above in [eq: 11.211] 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 it 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 11.327. *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 11.328.

$$[2\oplus 3]\left(\begin{pmatrix}1 & 2 & 3 & 4\\ 5 & 6 & 7 & 8\\ 10 & 20 & 30 & 40\\ 11 & 12 & 13 & 14\end{pmatrix}\right)=\begin{pmatrix}1 & 0 & 3 & 4\\ 5 & 0 & 7 & 8\\ 10 & 1 & 30 & 40\\ 11 & 0 & 13 & 14\end{pmatrix}$$

Definition 11.329. (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 11.330. *Let $n\in\mathbb{N}$, X a finite dimensional vector space over a field with characteristic zero so that $\dim(X)=n$, $\{e_i\}_{i\in\{1,\ldots,n\}}\subseteq X$ a distinct family defining a basis $E=\{e_i|i\in\{1,\ldots,n\}\}$ 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

$$\mathcal{M}(\text{adjoint}(L);E,E)=\text{adjoint}(\mathcal{M}(L;E,E))$$

In other words the matrix of the adjoint of a linear transformation is the adjoint of the matrix of the linear transformation.

Proof. Using [theorem: 11.255] 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 $i\in\{1,\ldots,n\}$

$$\begin{aligned}
\sum_{i\in\{1,\ldots,n\}}\mathcal{M}(\text{adjoint}(L);E,E)_{i,j}\cdot e_i &\underset{\text{def}}{=}\\
\text{adjoint}(L)(e_j)\\
1\cdot\text{adjoint}(L)(e_j) &=\\
\Delta(e_1,\ldots,e_n)\cdot\text{adjoint}(L)(e_j) &\underset{[\text{definition: } 11.276]}{=}\\
\overline{\Delta L}(e_1,\ldots,e_n)(e_j) &\underset{[\text{definition: } 11.273]}{=}\\
\sum_{i\in\{1,\ldots,n\}}(-1)^{i-1}\cdot\Delta(e_j,L(e_1),\ldots,L(e_{i-1}),L(e_{i+1}),\ldots,L(e_n))\cdot e_i
\end{aligned}$$

From the uniqueness of expanding in a basis it follows that for

$\forall i,j\in\{1,\ldots,n\}$ we have $\mathcal{M}(\text{adjoint}(L);E,E)_{i,j}=(-1)^{i-1}\cdot\Delta(e_j,L(e_1),\ldots,L(e_{i-1}),L(e_{i+1}),\ldots,L(e_n))$ (11.212)

Given $i,j\in\{1,\ldots,n\}$ define

$$x^{(i,j)}\in X^n \text{ by } \forall k\in\{1,\ldots,n\}\ x^{(i,j)}_k=\begin{cases} e_j \text{ if } k=i\\ L(e_k) \text{ if } k\in\{1,\ldots,n\}\setminus\{i\}\end{cases} \tag{11.213}$$

Then we have for $i,j\in\{1,\ldots,n\}$ and $k\in\{1,\ldots,n\}$ that

$$\begin{aligned}
(e_j,L(e_1),\ldots,L(e_{i-1}),L(e_{i+1}),\ldots,L(e_n))_k &\underset{\text{[definition: 11.258]}}{=}\\
&\begin{cases} e_j \text{ if } k=1\\ L(e_{k-1}) \text{ if } k\in\{2,\ldots,i\}\\ L(e_k) \text{ if } k\in\{i+1,\ldots,n\}\end{cases}\\
&\begin{cases} x^{(i,j)}_i \text{ if } k=1\\ x^{(i,j)}_{k-1} \text{ if } k\in\{2,\ldots,j\}\\ x^{(i,j)}_k \text{ if } k\in\{j+1,\ldots,n\}\end{cases} =\\
\left(x^{(i,j)}_i,x^{(i,j)}_1,\ldots,x^{(i,j)}_{i-1},x^{(i,j)}_{i+1},\ldots,x^{(i,j)}_n\right) &\underset{\text{[theorem: 11.259]}}{=}\\
&\left(x^{(i,j)}_{(i\underset{n}{\rightsquigarrow}1)(1)},\ldots,x^{(i,j)}_{(i\underset{n}{\rightsquigarrow}1)(n)}\right)
\end{aligned}$$

So substituting the above in [eq: 11.212] gives $\forall i,j\in\{1,\ldots,n\}$ that

$$\begin{aligned}
\mathcal{M}(\text{adjoint}(L);E,E)_{i,j} &= (-1)^{i-1}\cdot\Delta\left(x^{(i,j)}_{(i\underset{n}{\rightsquigarrow}1)(1)},\ldots,x^{(i,j)}_{(i\underset{n}{\rightsquigarrow}1)(n)}\right)\\
&= (-1)^{i-1}\cdot\left(\left(i\underset{n}{\rightsquigarrow}1\right)\Delta\right)\left(x^{(i,j)}_1,\ldots,x^{(i,j)}_n\right)\\
&= (-1)^{i-1}\cdot\text{sign}\left(\left(i\underset{n}{\rightsquigarrow}1\right)\right)\cdot\Delta\left(x^{(i,j)}_1,\ldots,x^{(i,j)}_n\right)\\
&\underset{\text{[theorem: 11.215]}}{=} (-1)^{i-1}\cdot(-1)^{i-1}\cdot\Delta\left(x^{(i,j)}_1,\ldots,x^{(i,j)}_n\right)\\
&= \Delta\left(x^{(i,j)}_1,\ldots,x^{(i,j)}_n\right)
\end{aligned} \tag{11.214}$$

Let $i,j\in\{1,\ldots,n\}$ and define $M^{(i,j)}\in\mathcal{M}_{n,n}(F)$ by

$$M^{(i,j)}=[i\oplus j](\mathcal{M}(E,E)(L)) \tag{11.215}$$

so that by [definition: 11.327]

$$\forall r,s\in\{1,\ldots,n\} \text{ we have } (M^{(i,j)})_{r,s}=\begin{cases}(\mathcal{M}(E,E)(L))_{r,s} \text{ if } s\in\{1,\ldots,n\}\setminus\{i\}\\ \delta_{r,j} \text{ if } s=i\end{cases} \tag{11.216}$$

Let $L^{(i,j)}: X\rightarrow X$ be defined by $\mathcal{M}(E,E)^{-1}(M^{(i,j)})$ then we have by [theorem: 11.297] and the fact that $\forall i\in\{1,\ldots,n\}$ $e_i=\sum_{j\in\{1,\ldots,n\}}\delta_{i,j}\cdot e_j$ that

$$\begin{aligned}
L^{(i,j)}(e_r) &= \sum_{s\in\{1,\ldots,n\}}\left(\sum_{t\in\{1,\ldots,n\}}(M^{(i,j)})_{s,t}.\delta_{r,t}\right)\cdot e_s\\
&= \sum_{s\in\{1,\ldots,n\}}(M^{(i,j)})_{s,r}\cdot e_s
\end{aligned} \tag{11.217}$$

For $r \in \{1, \ldots, n\}$ we have either:

$\boldsymbol{r = i}$**.** Then

$$\begin{aligned}
L^{(i,j)}(e_r) &= L^{(i,j)}(e_i)\\
&\underset{[\text{eq: } 11.217]}{=} \sum_{s \in \{1, \ldots, n\}} (M^{(i,j)})_{s,i} \cdot e_s\\
&\underset{[\text{eq: } 11.216]}{=} \sum_{s \in \{1, \ldots, n\}} \delta_{j,s} \cdot e_s\\
&= e_j\\
&\underset{r=u \wedge [\text{eq: } 11.213]}{=} x_i^{(i,j)}\\
&= x_r^{(i,j)}
\end{aligned}$$

$\boldsymbol{r \neq i}$**.** Then

$$\begin{aligned}
L^{(i,j)}(e_r) &\underset{[\text{eq: } 11.217]}{=} \sum_{s \in \{1, \ldots, n\}} (M^{(i,j)})_{s,r} \cdot e_s\\
&\underset{[\text{eq: } 11.216]}{=} \sum_{s \in \{1, \ldots, n\}} (\mathcal{M}(E,E)(L))_{s,r} \cdot e_s\\
&= \sum_{s \in \{1, \ldots, n\}} (\mathcal{M}(L; E,E)(L))_{s,r} \cdot e_s\\
&= L(e_r)\\
&\underset{r \neq u \wedge [\text{eq: } 11.213]}{=} x_r^{i,j}
\end{aligned}$$

So we have proved that

$$\forall r \in \{1, \ldots, n\} \text{ we have that } L^{(i,j)}(e_r) = x_r^{(i,j)} \tag{11.218}$$

Substituting the above in [eq: 11.214] results in $\forall i, \in \{1, \ldots, n\}$

$$\begin{aligned}
\mathcal{M}(\text{adjoint}(L); E, E)_{i,j} &\underset{[\text{eq: } 11.214]}{=} \Delta(L^{(i,j)}(e_1), \ldots, L^{(i,j)}(e_n))\\
&= \Delta_{L^{(i,j)}}(e_1, \ldots, e_n)\\
&\underset{[\text{definition: } 11.267]}{=} \det(L^{(i,j)}) \cdot \Delta(e_1, \ldots, e_n)\\
&= \det(L^{(i,j)})\\
&\underset{[\text{theorem: } 11.310]}{=} \det(\mathcal{M}[E,E](L^{(i,j)}))\\
&= \det(\mathcal{M}[E,E](\mathcal{M}[E,E]^{-1}(M^{(i,j)})))\\
&= \det(M^{(i,j)})\\
&\underset{[\text{eq: } 11.215]}{=} \det([i \oplus j](\mathcal{M}[E,E](L)))\\
&\underset{[\text{definition: } 11.329]}{=} (\text{adjoint}(\mathcal{M}[E,E])(L))_{i,j}
\end{aligned}$$

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: 11.278] in terms of matrices instead of linear transformations.

Theorem 11.331. *Let $n \in \mathbb{N}$, X a finite dimensional vector space over a field with characteristic zero so that $\dim(X) = n$, $\{e_i\}_{i \in \{1, \ldots, n\}} \subseteq X$ a distinct family defining a basis $E = \{e_i | i \in \{1, \ldots, n\}\}$ 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. $\mathcal{M}(L;E,E)\cdot\operatorname{adjoint}(\mathcal{M}(L;E,E))=\det(\mathcal{M}(L;E,E))\cdot E$
2. $\operatorname{adjoint}(\mathcal{M}(L;E,E))\cdot\mathcal{M}(L;E,E)=\det(\mathcal{M}(L;E,E))\cdot E$

or if we write out the matrix product:

1. $\forall i,j\in\{1,\ldots,n\}$ *we have*
$$\sum_{k\in\{1,\ldots,n\}}\mathcal{M}(L;E,E)_{i,k}\cdot\operatorname{adjoint}(\mathcal{M}(L;E,E))_{k,j} = \det(\mathcal{M}(L;E,E))\cdot\delta_{i,j}$$
2. $\forall i,j\in\{1,\ldots,n\}$ *we have*
$$\sum_{k\in\{1,\ldots,n\}}\operatorname{adjoint}(\mathcal{M}(L;E,E))_{i,k}\cdot\mathcal{M}(L;E,E)_{k,j} = \det(\mathcal{M}(L;E,E))\cdot\delta_{i,j}$$

Proof.

1. We have
$$\begin{aligned}\det(\mathcal{M}(E,E)(L))\cdot E &\underset{[\text{theorem: }11.296]}{=} \det(\mathcal{M}(E,E)(L))\cdot\mathcal{M}(E,E)(\mathrm{Id}_X)\\ &\underset{[\text{theorem: }11.296]}{=} \mathcal{M}(E,E)(\det(\mathcal{M}(E,E)(L))\cdot\mathrm{Id}_X)\\ &\underset{[\text{theorem: }11.310]}{=} \mathcal{M}(E,E)(\det(L)\cdot\mathrm{Id}_X)\\ &\underset{[\text{theorem: }11.278]}{=} \mathcal{M}(E,E)(L\circ\operatorname{adjoint}(L))\\ &\underset{[\text{theorem: }11.299]}{=} \mathcal{M}(E,E)(L)\cdot\mathcal{M}(E,E)(\operatorname{adjoint}(L))\end{aligned}$$
2. We have
$$\begin{aligned}\det(\mathcal{M}(E,E)(L))\cdot E &\underset{[\text{theorem: }11.296]}{=} \det(\mathcal{M}(E,E)(L))\cdot\mathcal{M}(E,E)(\mathrm{Id}_X)\\ &\underset{[\text{theorem: }11.296]}{=} \mathcal{M}(E,E)(\det(\mathcal{M}(E,E)(L))\cdot\mathrm{Id}_X)\\ &\underset{[\text{theorem: }11.310]}{=} \mathcal{M}(E,E)(\det(L)\cdot\mathrm{Id}_X)\\ &\underset{[\text{theorem: }11.278]}{=} \mathcal{M}(E,E)(\operatorname{adjoint}(L)\circ L)\\ &\underset{[\text{theorem: }11.299]}{=} \mathcal{M}(E,E)(\operatorname{adjoint}(L))\cdot\mathcal{M}(E,E)(L)\end{aligned}$$

□

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 11.332. *Let* $n\in\mathbb{N}$*,* F *a field with characteristic zero and* $M\in\mathcal{M}_{n,n}(F)$ *then we have:*

1. $$M\cdot\operatorname{adjoint}(M)=\det(M)\cdot E,$$
hence $\forall i,j\in\{1,\ldots,n\}$
$$\sum_{k\in\{1,\ldots,n\}}M_{i,k}\cdot\operatorname{adjoint}(M)_{k,j}=\det(M)\cdot\delta_{i,j}$$
2. $$\operatorname{adjoint}(M)\cdot M=\det(M)\cdot E$$
hence $\forall i,j\in\{1,\ldots,n\}$
$$\sum_{k\in\{1,\ldots,n\}}\operatorname{adjoint}(M)_{i,k}\cdot M_{k,j}=\det(M)\cdot\delta_{i,j}$$
3. *If* $n>1$ *then* $\forall i,j\in\{1,\ldots,n\}$ *we have*
$$\operatorname{adjoint}(M)_{i,j}=(-1)^{i+j}\cdot\det([j\boxplus i](M))$$

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* $\operatorname{adjoint}(M)=(1)$

Proof.

1. Take $L=\mathcal{M}(E,E)^{-1}(M)$ so that $M=\mathcal{M}(E,E)(L)$ then we have that
$$\begin{aligned}M\cdot\operatorname{adjoint}(M) &= \mathcal{M}(E,E)(L)\cdot\operatorname{adjoint}(\mathcal{M}(E,E)(L))\\ &\underset{[\text{theorem: }11.331]}{=} \det(\mathcal{M}(E,E)(L))\cdot E\\ &= \det(M)\cdot E\end{aligned}$$

2. Take $L=\mathcal{M}(E,E)^{-1}(M)$ so that $M=\mathcal{M}(E,E)(L)$ then we have that
$$\begin{aligned}\operatorname{adjoint}(M)\cdot M &= \operatorname{adjoint}(\mathcal{M}(E,E)(L))\cdot\mathcal{M}(E,E)(L)\\ &\underset{[\text{theorem: }11.331]}{=} \det(\mathcal{M}(E,E)(L))\cdot E\\ &= \det(M)\cdot E\end{aligned}$$

3. First if $i,j\in\{1,\ldots,n\}$ then we have
$$\begin{aligned}\det([i\boxplus j](M)) &\underset{[\text{theorem: }11.309]}{=} \det(([i\oplus j](M))^T)\\ &= \det(([i\oplus j](M))^T)\cdot\delta_{i,i}\\ &\underset{(1)}{=} \sum_{k\in\{1,\ldots,n\}}(([i\oplus j](M))^T)_{i,k}\cdot\operatorname{adjoint}(([i\oplus j](M))^T)_{k,i}\\ &= \sum_{k\in\{1,\ldots,n\}}([i\oplus j](M))_{k,i}\cdot\operatorname{adjoint}(([i\oplus j](M))^T)_{k,i}\\ &\underset{[\text{definition: }11.327]}{=} \sum_{k\in\{1,\ldots,n\}}\delta_{k,j}\cdot\operatorname{adjoint}(([i\oplus j](M))^T)_{k,i}\\ &= \operatorname{adjoint}(([i\oplus j](M))^T)_{j,i}\\ &\underset{[\text{definition: }11.329]}{=} \det([j\oplus i](([i\oplus j](M))^T))\end{aligned}$$
giving
$$\det([i\boxplus j](M))=\det([j\oplus i](([i\oplus j](M))^T))\tag{11.219}$$
Then $\forall k\in\{1,\ldots,n\}$ we have
$$\begin{aligned}\operatorname{row}([j\oplus i](([i\oplus j](M))^T),i)_k &= ([j\oplus i](([i\oplus j](M))^T))_{i,k}\\ &\underset{[\text{definition: }11.327]}{=} \begin{cases}(([i\oplus j](M))^T)_{i,k}\text{ if }k\in\{1,\ldots,n\}\setminus\{j\}\\ \delta_{i,i}\text{ if }k=j\end{cases}\\ &= \begin{cases}([i\oplus j](M))_{k,i}\text{ if }k\in\{1,\ldots,n\}\setminus\{j\}\\ 1\text{ if }k=j\end{cases}\\ &\underset{[\text{definition: }11.327]}{=} \begin{cases}0\text{ if }k\in\{1,\ldots,n\}\setminus\{j\}\\ 1\text{ if }k=j\end{cases}\\ &= \delta_{k,j}\end{aligned}$$

giving

$$\operatorname{row}([j\oplus i](([i\oplus j](M))^T),i)_k=\delta_{k,j} \tag{11.220}$$

Further

$$\begin{aligned}
\operatorname{col}([j\oplus i](([i\oplus j](M))^T),j)_k &= ([j\oplus i](([i\oplus j](M))^T))_{k,j}\\
&\underset{[\text{definition: } 11.327]}{=} \begin{cases} (([i\oplus j](M))^T)_{k,j} \text{ if } j\in\{1,\ldots,n\}\setminus\{j\}\\ \delta_{k,i} \text{ if } j=j \end{cases}\\
&= \delta_{k,i}
\end{aligned}$$

giving

$$\operatorname{col}([j\oplus i](([i\oplus j](M))^T),j)_k=\delta_{k,i} \tag{11.221}$$

From [eqs: 11.220, 11.221] it follows that the conditions of [theorem: 11.326] 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: 11.219] results in

$$\det([i\boxplus j](M))=(-1)^{i+j}\cdot\det((i\boxplus j)([j\oplus i](([i\oplus j](M))^T))) \tag{11.222}$$

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:

$\boldsymbol{1\leqslant k<i\wedge 1\leqslant l<j.}$ Then

$$\begin{aligned}
((i\boxplus j)([j\oplus i](([i\oplus j](M))^T)))_{k.l} &\underset{[\text{definition: } 11.319]}{=}\\
([j\oplus i](([i\oplus j](M))^T))_{k,l} &\underset{l<j\Rightarrow l\neq j\wedge[\text{definition: } 11.327]}{=}\\
(([i\oplus j](M))^T)_{k,l} &=\\
([i\oplus j](M))_{l.k} &\underset{k<i\Rightarrow k\neq i\wedge[\text{definition: } 11.327]}{=}\\
M_{l,k} &=\\
(M^T)_{k,l} &\underset{[\text{definition: } 11.319]}{=}\\
([i\boxplus j](M^T))_{k,l}
\end{aligned}$$

$\boldsymbol{i\leqslant k\leqslant n-1\wedge 1\leqslant l<j.}$ Then

$$\begin{aligned}
((i\boxplus j)([j\oplus i](([i\oplus j](M))^T)))_{k.l} &\underset{[\text{definition: } 11.319]}{=}\\
([j\oplus i](([i\oplus j](M))^T))_{k+1,l} &\underset{l<j\Rightarrow l\neq j\wedge[\text{definition: } 11.327]}{=}\\
(([i\oplus j](M))^T)_{k+1,l} &=\\
([i\oplus j](M))_{l,k+1} &\underset{i\leqslant k\Rightarrow i\neq k+1\wedge[\text{definition: } 11.327]}{=}\\
M_{k+1,l} &=\\
(M^T)_{l,k+1} &\underset{[\text{definition: } 11.319]}{=}\\
([i\boxplus j](M^T))_{k,l}
\end{aligned}$$

$\boldsymbol{1\leqslant k<i\wedge j\leqslant l\leqslant n-1.}$ Then

$$\begin{aligned}
((i\boxplus j)([j\oplus i](([i\oplus j](M))^T)))_{k.l} &\underset{[\text{definition: } 11.319]}{=}\\
([j\oplus i](([i\oplus j](M))^T))_{k,l+1} &\underset{j\leqslant l\Rightarrow l+1\neq j\wedge[\text{definition: } 11.327]}{=}\\
(([i\oplus j](M))^T)_{k,l+1} &=\\
([i\oplus j](M))_{l+1,k} &\underset{k<i\Rightarrow k\neq i\wedge[\text{definition: } 11.327]}{=}\\
M_{l+1,k} &=\\
(M^T)_{k,l+1} &\underset{[\text{definition: } 11.319]}{=}\\
([i\boxplus j](M^T))_{k,l}
\end{aligned}$$

$\boldsymbol{i\leqslant k\leqslant n-1\wedge j\leqslant l\leqslant n-1.}$ Then

$$\begin{aligned}((i\boxplus j)([j\oplus i]((([i\oplus j](M))^T)))_{k.l} &\underset{[\text{definition: } 11.319]}{=}\\ ([j\oplus i]((([i\oplus j](M))^T))_{k+1,l+1} &\underset{j\leqslant l\Rightarrow j\neq l+1\wedge[\text{definition: } 11.327]}{=}\\ ((([i\oplus j](M))^T)_{k+1,l+1} &=\\ ([i\oplus j](M))_{l+1,k+1} &\underset{i\leqslant k\Rightarrow i\neq k+1\wedge[\text{definition: } 11.327]}{=}\\ M_{l+1,k+1} &=\\ (M^T)_{k+1,l+1} &\underset{[\text{definition: } 11.319]}{=}\\ ([i\boxplus j](M^T))_{k,l}\end{aligned}$$

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: 11.222] gives

$$\begin{aligned}\det([i\boxplus j](M)) &= (-1)^{i+j}\cdot\det([i\boxplus j](M^T))\\ &\underset{[\text{theorem: } 11.321]}{=} (-1)^{i+j}\cdot\det(([j\boxplus i](M))^T)\end{aligned}\tag{11.223}$$

Finally

$$\begin{aligned}\mathrm{adjoint}(M)_{i,j} &\underset{[\text{definition: } 11.329]}{=} \det([i\boxplus j](M))\\ &\underset{[\text{eq: } 11.223]}{=} (-1)^{i+j}\cdot\det(([j\boxplus i](M))^T)\\ &\underset{[\text{theorem: } 11.309]}{=} (-1)^{i+j}\cdot\det([j\boxplus i]M)\end{aligned}$$

4. Let $i\in\{1,\ldots,n\}$ then we have

$$\begin{aligned}\det(M) &= \det(M)\cdot\delta_{i,i}\\ &\underset{(2)}{=} \sum_{k\in\{1,\ldots,n\}}\mathrm{adjoint}(M)_{i,k}\cdot M_{k,i}\\ &\underset{(3)}{=} \sum_{k\in\{1,\ldots,n\}}(-1)^{i+k}\cdot\det([k\boxplus i](M))\cdot M_{k,i}\\ &= \sum_{k\in\{1,\ldots,n\}}(-1)^{i+k}\cdot M_{k,i}\cdot\det([k\boxplus i](M))\end{aligned}$$

5. Let $i\in\{1,\ldots,n\}$ then we have

$$\begin{aligned}\det(M) &= \det(M)\cdot\delta_{i,i}\\ &\underset{(1)}{=} \sum_{k\in\{1,\ldots,n\}}M_{i,k}\cdot\mathrm{adjoint}(M)_{k,i}\\ &\underset{(3)}{=} \sum_{k\in\{1,\ldots,n\}}M_{i,k}\cdot(-1)^{i+k}\cdot\det([i\boxplus k](M))\\ &= \sum_{k\in\{1,\ldots,n\}}(-1)^{i+k}\cdot M_{i,k}\cdot\det([i\boxplus k](M))\end{aligned}$$

6. This follows from [theorem: 11.332]

7. Using [theorem: 11.306] we have that for $M\in\mathcal{M}_{1,1}(F)$ that

$$\det(M)=M_{1,1}$$

Further

$$\mathrm{adjoint}(M)=\det([1\oplus 1](\ M_{1,1}\ ))=\det((\ 1\ ))=1 \qquad \square$$

We can use (4),(7) from the above to calculate the determinant of a matrix as show in the following example:

Example 11.333. Using [theorem: 11.332 (4)] we have

$$\begin{aligned} \begin{vmatrix} -2 & 2 & 3 \\ -1 & 1 & 3 \\ 2 & 0 & 1 \end{vmatrix} &= \\ (-1)^{1+1}\cdot M_{1,1}\cdot\det([1\boxplus 1]M)+(-1)^{1+2}\cdot M_{2,1}\cdot\det([2\boxplus 1]\cdot M)+(-1)^{1+3}\cdot M_{3,1}\cdot\det([3\boxplus 1]M) &= \\ -2\cdot\begin{vmatrix} 1 & 3 \\ 0 & 1 \end{vmatrix}+(-1)\cdot(-1)\cdot\begin{vmatrix} 2 & 3 \\ 0 & 1 \end{vmatrix}+2\cdot\begin{vmatrix} 2 & 3 \\ 1 & 3 \end{vmatrix} &= \\ -2\cdot(1\cdot|\,1\,|+(-1)\cdot 0\cdot|\,3\,|)+(2\cdot|\,1\,|+(-1)\cdot 0\cdot|\,3\,|)+2\cdot(2\cdot|\,3\,|+(-1)\cdot 1\cdot|\,3\,|) &= \\ -2+2+2\cdot 3 &= \\ 6 & \end{aligned}$$

To simplify calculation of the adjoint of a matrix we introduce the co-factor matrix

Definition 11.334. *Let $n\in\mathbb{N}\setminus\{1\}$, F a field with characteristic zero and $M\in\mathcal{M}_{n,n}(F)$ then the refactor matrix noted by* $\operatorname{cofactor}(M)$ *is defined by*

$$\operatorname{cofactor}(M)\in\mathcal{M}_{n,n}(M) \text{ where } \forall i,j\in\{1,\ldots,n\}\ \operatorname{cofactor}(M)_{i,j}=(-1)^{i+j}\cdot\det([i\boxplus j](M))$$

The use of the co-factor matrix is shown in the following theorem:

Theorem 11.335. *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

$$((\operatorname{cofactor}(M))^T)_{i,j}=\operatorname{cofactor}(M)_{j,i}=(-1)^{j+i}\cdot\det([j\boxplus i](M))\underset{\text{[theorem: 11.332(3}}{=}\operatorname{adjoint}(M)_{i,j}$$

proving that

$$\operatorname{adjoint}(M)=(\operatorname{cofactor}(M))^T$$ □

Example 11.336.

$$\begin{aligned} \operatorname{cofactor}\left(\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}\right) &= \begin{pmatrix} 1\cdot\begin{vmatrix} 5 & 6 \\ 8 & 9 \end{vmatrix} & -1\cdot\begin{vmatrix} 4 & 6 \\ 7 & 9 \end{vmatrix} & 1\cdot\begin{vmatrix} 4 & 5 \\ 7 & 8 \end{vmatrix} \\ -1\cdot\begin{vmatrix} 2 & 3 \\ 8 & 9 \end{vmatrix} & 1\cdot\begin{vmatrix} 1 & 3 \\ 7 & 9 \end{vmatrix} & -1\cdot\begin{vmatrix} 1 & 2 \\ 7 & 8 \end{vmatrix} \\ 1\cdot\begin{vmatrix} 2 & 3 \\ 5 & 6 \end{vmatrix} & -1\cdot\begin{vmatrix} 1 & 3 \\ 4 & 6 \end{vmatrix} & 1\cdot\begin{vmatrix} 1 & 2 \\ 4 & 5 \end{vmatrix} \end{pmatrix} \\ &= \begin{pmatrix} -3 & 6 & -3 \\ 6 & -12 & 6 \\ -3 & 6 & -3 \end{pmatrix} \end{aligned}$$

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 11.337. *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*
$$\operatorname{row}(M, i)_j = \delta_{i,j}$$
then
$$\det(M) = \det([<m](M))$$

Proof.

1. *We prove this by induction, so let*

$S_{n,M} = \{m \in \mathbb{N} | \text{If } 1 \leqslant m < n$ *and satisfies* $\forall i \in \{1, \ldots, m\}, \forall j \in \{1, \ldots, n\}$ $\operatorname{col}(M, i)_j = \delta_{i,j}$ *then* $\det(M) = \det([>m](M))\}$

then we have:

$1 \in S_{n,M}$. *Let $M \in \mathcal{M}_{n,n}(F)$ be such that $\forall j \in \{1,\ldots,n\}$ $\operatorname{col}(M,1)_j = \delta_{j,i}$ then we have*

$$\begin{aligned}
\det(M) &\underset{\text{[theorem: 11.332(4)]}}{=} \sum_{i \in \{1,\ldots,n\}} (-1)^{i+1} \cdot M_{i,1} \cdot \det([i \boxplus 1](M)) \\
&= \sum_{i \in \{1,\ldots,n\}} (-1)^{i+1} \cdot \delta_{i,1} \cdot \det([i \boxplus 1](M)) \\
&= (-1)^2 \cdot \det([1 \boxplus 1](M)) \\
&= \det([1 \boxplus 1](M)) \\
&\underset{\text{[theorem: 11.323]}}{=} \det([>1](M))
\end{aligned}$$

proving that $1 \in S_{n,M}$.

$m \in S_{n,M} \Rightarrow m+1 \in S_{n,M}$. *Let $M \in \mathcal{M}_{n,n}(F)$ be such that $\forall i \in \{1, \ldots, m+1\}$ and $\forall j \in \{1,\ldots,n\}$ we have $\operatorname{col}(M,i)_j = \delta_{i,j}$. As $m \in S+n, M$ and $\forall i \in \{1,\ldots,m\}, \forall j \in \{1,\ldots, n\}$ $\operatorname{col}(M, i)_j$ we have*

$$\det(M) = \det([>m](M)) \tag{11.224}$$

If $j \in \{1, \ldots, n-m\}$ then

$$\begin{aligned}
\operatorname{col}([>m](M), 1)_j &= ([>m](M))_{j,1} \\
&\underset{\text{[definition: 11.322]}}{=} M_{j+m,1+m} \\
&= \operatorname{col}(1+m, j+m) \\
&= \delta_{1+m,j+m} \\
&= \delta_{j,1}
\end{aligned} \tag{11.225}$$

Next

$$\begin{aligned}
\det(M) &\underset{\text{[eq: 11.224]}}{=} \\
\det([>m](M)) &\underset{\text{[theorem: 11.332(4)]}}{=} \\
\sum_{i\in\{1,\ldots,n\}} (-1)^{i+1}\cdot([>m](M))_{i,1}\cdot\det([i\boxplus 1]([>m])(M)) &= \\
\sum_{i\in\{1,\ldots,n\}} (-1)^{i+1}\cdot\mathrm{col}([>m](M),1)_j\cdot\det([i\boxplus 1]([>m])(M)) &\underset{\text{[eq: 11.225]}}{=} \\
\sum_{i\in\{1,\ldots,n\}} (-1)^{i+1}\cdot\delta_{j,1}\cdot\det([i\boxplus 1]([>m])(M)) &= \\
(-1)^2\cdot\det([1\boxplus 1]([>m])(M)) &\underset{\text{[theorem: 11.323]}}{=} \\
\det([>1]([>m](M))) &\underset{\text{[theorem: 11.323]}}{=} \\
\det([>m+1](M))
\end{aligned}$$

proving that $m+1\in S_{n,M}$.

Mathematical induction proves then that $S_{n,M}=\mathbb{N}$. *So if* $M\in\mathcal{M}_{n,n}(F)$ *and* $m\in\mathbb{N}$ *then* $m\in S_{n,m}$, *hence if* $1\leqslant m<n$ *and* $\forall i\in\{1,\ldots,m\},\forall j\in\{1,\ldots,n\}$ $\mathrm{col}(M,i)_j=\delta_{i,j}$ *then we have* $\det(M)=\det([>m](M))$.

2. *We prove this by induction, so let*

$S=\{n\in\{2,\ldots,n\}|$*If* $M\in\mathcal{M}_{n,n}(F)$ *satisfies* $\exists m\in\{2,\ldots,n\}$ *such that* $\forall i\in\{1,\ldots,m\}$, $\forall j\in\{1,\ldots,n\}$ *we have* $\mathrm{col}(M,i)_j=\delta_{i,j}$ *then* $\det(M)=\det([<m](M))\}$

then we have:

2 ∈ S. *Let* $M\in\mathcal{M}_{2,2}(F)$ *such that* $\exists m\in\{2,\ldots,2\}$ *such that* $\forall i\in\{m,\ldots,n\}$ *and* $\forall j\in\{1,\ldots,2\}$*we have* $\mathrm{col}(M,i)_j=\delta_{i,j}$ *then* $m=2$ *and we have*

$$\begin{aligned}
\det(M) &\underset{\text{[theorem: 11.332(4)]}}{=} \sum_{i\in\{1,\ldots,2\}} (-1)^{i+2}\cdot M_{i,2}\cdot\det([i\boxplus 2](M)) \\
&= \sum_{i\in\{1,\ldots,n\}} (-1)^{i+2}\cdot\mathrm{col}(M,2)_i\cdot\det([i\boxplus 2](M)) \\
&= \sum_{i\in\{1,\ldots,n\}} (-1)^{i+2}\cdot\delta_{2,i}\cdot\det([i\boxplus 2](M)) \\
&= (-1)^{2+2}\cdot\det([2\boxplus 2](M)) \\
&= \det([2\boxplus 2](M)) \\
&\underset{\text{[theorem: 11.323]}}{=} \det([<2](M))
\end{aligned}$$

proving that $2\in S$.

n ∈ S ⇒ n + 1 ∈ 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* $\mathrm{col}(M,i)_j=\delta_{i,j}$. *As*

$$[n+1\boxplus n+1](M)\underset{\text{[theorem: 11.323]}}{=}[<n+1](M) \tag{11.226}$$

it follows that $\forall i\in\{m,\ldots,n\}$ *and* $\forall j\in\{1,\ldots,n\}$ *we have*

$$\begin{aligned}
\mathrm{col}([n+1\boxplus n+1](M),i)_j &= ([n+1\boxplus n+1](M))_{j,i} \\
&\underset{i<n+1\wedge j<n+1\wedge\text{[definition: 11.319]}}{=} M_{j,i} \\
&= \mathrm{col}(M,i)_j \\
&= \delta_{i,j}
\end{aligned}$$

which as $[n+1 \boxplus n+1](M) \in \mathcal{M}_{n,n}(F)$ *and* $n \in S$ *proves that*

$$\det([n+1 \oplus n+1](M)) = \det([<m]([n+1 \oplus n+1](M)))$$

substituting [eq: 11.225] in the last term from the above gives

$$\det([n+1 \oplus n+1](M)) = \det([<m]([<n+1](M)))$$

As $m \in \{2, \ldots, n+1\}$ *it follows from [theorem: 11.323] that*

$$[<m]([<n+1](M)) = [<m](M)$$

so that

$$\det([n+1 \oplus n+1](M)) = \det([<m](M)) \tag{11.227}$$

Further we have

$$\begin{aligned}
\det(M) & \underset{\text{[theorem: 11.332(4)]}}{=} \sum_{i \in \{1, \ldots, n\}} (-1)^{i+(n+1)} \cdot M_{i,n+1} \cdot \det([i \boxplus n+1](M)) \\
& = \sum_{i \in \{1, \ldots, n\}} (-1)^{i+(n+1)} \cdot \mathrm{col}(M, n+1)_i \cdot \det([i \boxplus n+1](M)) \\
& = \sum_{i \in \{1, \ldots, n\}} (-1)^{i+(n+1)} \cdot \delta_{i,n+1} \cdot \det([i \boxplus n+1](M)) \\
& = (-1)^{(n+1)+(n+1)} \cdot \det([n+1 \boxplus n+1](M)) \\
& = \det([n+1 \boxplus n+1](M))
\end{aligned}$$

which combined wit [eq: 11.227] gives

$$\det(M) = \det([<m]M)$$

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*

$$\det(M^T) = \det([>m]M^T) \tag{11.228}$$

hence

$$\begin{aligned}
\det(M) & \underset{\text{[theorem: 11.309]}}{=} \det(M^T) \\
& \underset{\text{[eq: 11.228]}}{=} \det([>m]M^T) \\
& \underset{\text{[theorem: 11.323]}}{=} \det([>m](M^T)^T) \\
& = \det([<m](M))
\end{aligned}$$

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*

$$\det(M^T) = \det([<m]M^T) \tag{11.229}$$

hence

$$\begin{aligned}
\det(M) & \underset{\text{[theorem: 11.309]}}{=} \det(M^T) \\
& \underset{\text{[eq: 11.229]}}{=} \det([<m]M^T) \\
& \underset{\text{[theorem: 11.323]}}{=} \det([<m](M^T)^T) \\
& = \det([>m](M))
\end{aligned}$$

□

The next theorem shows sufficient and necessary conditions for a matrix to be invertible and how to calculate the inverse of the matrix.

Theorem 11.338. *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: 11.315] that

$$\mathrm{rank}(M) = m \Leftrightarrow \det(M) = 0$$

If $\det(M) \neq 0$ then $(\det(M))^{-1}$ exists and we have

$$\begin{array}{rcl}((\det(M))^{-1} \cdot \mathrm{adjoint}(M)) \cdot M & = & (\det(M))^{-1} \cdot (\mathrm{adjoint}(M) \cdot M)\\ & \underset{\text{[theorem: 11.332(2)]}}{=} & (\det(M))^{-1} \cdot (\det(M) \cdot E)\\ & = & E\\ M \cdot ((\det(M))^{-1} \cdot \mathrm{adjoint}(M)) & = & (\det(M))^{-1} \cdot (\mathrm{adjoint}(M) \cdot M)\\ & \underset{\text{[theorem: 11.332(1)]}}{=} & (\det(M))^{-1} \cdot (\det(M) \cdot E)\\ & = & E\end{array}$$

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

$$1 = \det(E) = \det(M \cdot M^{-1}) \underset{\text{[theorem: 11.311]}}{=} \det(M) \cdot \det(M^{-1})$$

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**. This allows us to extend statements about these elementary transformations to a general non singular operation once we know how these statements behaves under composition of transformations. We will use this to show how a Lebesgue measure transforms under a non singular transformation.

Definition 11.339. *Let $n \in \mathbb{N}$, X a finite dimensional vector space with $\dim(X) = n$ then a linear transformation $L \in \mathrm{Hom}(X, X)$ is nonsingular if $\det(L) \neq 0$. Using [theorem: 11.271] 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*

$$\mathrm{GL}(X) = \{L \in \mathrm{Hom}(X, X) | \det(L) \neq 0\}$$

We show now that $\mathrm{GL}(X)$ together with the composition forms a group

Theorem 11.340. *Let $n \in \mathbb{N}$, X a finite dimensional vector space with $\dim(X) = n$ then*

$\langle \mathrm{GL}(X), \circ\rangle$ is a group with neutral element Id_X and for every L L^{-1} as inverse.

Proof. First if $L_1, L_2 \in \mathrm{GL}(X)$ then $\det(L_1 \circ L_2) \underset{\text{[theorem: 11.271]}}{=} \det(L_1) \cdot \det(L_2)$ so that $L_1 \circ L_2 \in \mathrm{GL}(X)$ so that $\circ$ is a operator on $\mathrm{GL}(X)$. Further we have:

associativity. If $L_1, L_2, L_3 \in \mathrm{GL}(X)$ then

$$(L_1 \circ L_2) \circ L_3 \underset{\text{[function: 2.21]}}{=} L_1 \circ (L_2 \circ L_3)$$

neutral element. As $\det(\mathrm{Id}_X) \underset{\text{[theorem: 11.271]}}{=} 1 \neq 0$ we have that $\mathrm{Id}_X \in \mathrm{GL}(X)$, further by [theorem: 2.48] we have that $\forall L \in \mathrm{GL}(X)$ we have $\mathrm{Id}_X \circ L = L = L \circ \mathrm{Id}_X$

inverse element. If $L \in \mathrm{GL}(X)$ then $\det(L^{-1}) \underset{\text{[theorem: 11.271]}}{=} (\det(L))^{-1} \neq 0$ so that $L^{-1} \in \mathrm{GL}(X)$, further $L \circ L^{-1} \underset{\text{[theorem: 2.68]}}{=} \mathrm{Id}_X = \underset{\text{[theorem: 2.68]}}{=} L^{-1} \circ L$ □

In [definition: 11.160] we have already defined how to make the product of linear transformations in $\mathrm{Hom}(X, X)$. As $\mathrm{GL}(X) \subseteq \mathrm{Hom}(X, X)$ it follow that the finite product of nonsingular transformations is well defined. Furthermore the following theorem shows that the finite product of nonsingular transformations is itself nonsingular.

Theorem 11.341. *Let $n \in \mathbb{N}$, X a finite dimensional vector space with $\dim(X) = n$ and $\{L_i\}_{i \in \{1, \ldots, m\}} \subseteq \mathrm{GL}(X)$ then*

$$L_1 \circ \cdots \circ L_n \underset{\text{def}}{=} \prod_{i=1}^{m} L_i \in \mathrm{GL}(X)$$

Proof. This is easily prove by induction, so let

$$S = \left\{ m \in \mathbb{N} | \text{If } \{L_i\}_{i \in \{1, \ldots, m\}} \subseteq \mathrm{GL}(X) \text{ then } \prod_{i=1}^{m} L_i \in \mathrm{GL}(X) \right\}$$

then we have:

$1 \in S$. If $\{L_i\}_{i \in \{1\}} \subseteq \mathrm{GL}(X)$ we have $\prod_{i=1}^{1} L_i = L_1 \in \mathrm{GL}(X)$ proving that $1 \in S$

$m \in S \Rightarrow m + 1 \in S$. If $\{L_i\}_{i \in \{1, \ldots, m+1\}} \subseteq \mathrm{GL}(X)$ then as $m \in S$ we have that $\prod_{i=1}^{m} L_i \in \mathrm{Hom}(X, X)$ which, as $L_{m+1} \in \mathrm{Hom}(X)$, proves that $\prod_{i=1}^{m+1} L_i = (\prod_{i=1}^{m} L_i) \circ L_{m+1} \in \mathrm{Hom}(X, X)$, hence $m + 1 \in S$. □

We will now introduce a special class of nonsingular operators called elementary transformations.

Definition 11.342. *Let $n \in \mathbb{N}$, X a finite dimensional vector space with $\dim(X) = n$ over a field F and $\{e_i\}_{i \in \{1, \ldots, n\}} \subseteq X$ a distinct family such that $E = \{e_i | i \in \{1, \ldots, n\}\}$ is a basis for X then we define:*

1. *Let $i \in \{1, \ldots, n\}$ and $\alpha \in F$ define*

$$M^n_{[i,\alpha]} \in \mathcal{M}_{n,n}(F) \text{ by } \forall k, l \in \{1, \ldots, n\}\ (M^n_{[i,\alpha]})_{k,l} = \begin{cases} \delta_{k,l} \text{ if } k \in \{1, \ldots, n\} \setminus \{i\} \\ \alpha \cdot \delta_{k,i} \text{ if } k = i \end{cases}$$

and

$$\sigma^n_{[i,\alpha]} \in \mathrm{Hom}(X, X) \text{ by } \sigma^n_{[i,\alpha]} = \mathcal{M}(E, E)^{-1}(M^n_{[i,\alpha]})$$

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

$$\sigma^n_{[i,\alpha]}(e_i) = \alpha \cdot e_i$$

or more compact

$$\forall k \in \{1,\ldots,n\} \textit{ we have } \sigma^n_{[i,\alpha]}(e_k) = \begin{cases} e_k \textit{ if } k \in \{1,\ldots,n\} \backslash \{i\} \\ \alpha \cdot e_i \textit{ if } k = i \end{cases}$$

2. *Let $i, j \in \{1,\ldots,n\}$ with $i \neq j$ and define*

$$N^n_{[i,j]} \in \mathcal{M}_{n,n}(F) \textit{ by } \forall k, l \in \{1,\ldots,n\} \; (N^n_{[i,j]})_{k,l} = \begin{cases} \delta_{j,k} \textit{ if } l = i \\ \delta_{i,k} \textit{ if } l = j \\ \delta_{k,l} \textit{ if } l \in \{1,\ldots,n\} \backslash \{i,j\} \end{cases}$$

and

$$\tau^n_{[i,j]} \in \mathrm{Hom}(X,X) \textit{ by } \tau^n_{[i,j]} = \mathcal{M}(E,E)^{-1}(N^n_{[i,j]})$$

So that $\forall k \in \{1,\ldots,n\} \backslash \{i,j\}$ 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

$$\forall k \in \{1,\ldots,n\} \; \tau^n_{[i,j]}(e_k) = \begin{cases} e_i \textit{ if } k = j \\ e_j \textit{ if } k = i \\ e_k \textit{ if } k \in \{1,\ldots,n\} \backslash \{i,j\} \end{cases}$$

3. *Let $i, j \in \{1,\ldots,n\}$ with $i \neq j$, $\alpha \in F$ and define*

$$O^n_{[i,j,\alpha]} \in \mathcal{M}_{n,n}(F) \textit{ by } \forall k, l \in \{1,\ldots,n\} \; (O^n_{[i,j,\alpha]})_{k,l} = \begin{cases} \delta_{k.i} + \alpha \cdot \delta_{j,k} \textit{ if } l = i \\ \delta_{k,l} \textit{ if } l = i \end{cases}$$

and

$$\beta^n_{[i,j,\alpha]} \in \mathrm{Hom}(X,X) \textit{ by } \beta^n_{[a,i,j]} \mathcal{M}[E,E]^{-1}(O^n_{[i,j,\alpha]})$$

So that $\forall k \in \{1,\ldots,n\} \backslash \{i\}$ 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

$$\begin{aligned}\beta^n_{[i,j,\alpha]}(e_i) &= \sum_{l\in\{1,\ldots,n\}} (O^n_{[i,j,\alpha]})_{l,i}\cdot e_l\\ &= \sum_{l\in\{1,\ldots,n\}} (\delta_{l,i}+\alpha\cdot\delta_{l,j})\cdot e_l\\ &= \sum_{l\in\{1,\ldots,n\}} \delta_{l,i}\cdot e_l + \sum_{l\in\{1,\ldots,n\}} \alpha\cdot\delta_{l,j}\cdot e_l\\ &= e_i+\alpha\cdot e_j\end{aligned}$$

or in other words

$$\forall k\in\{1,\ldots,n\} \text{ we have } \beta^n_{[i,j,\alpha]}(e_i)=\begin{cases} e_k \text{ if } k\in\{1,\ldots,n\}\setminus\{i\}\\ e_i+\alpha\cdot e_j \text{ if } k=i\end{cases}$$

The above transformations have the following properties

Theorem 11.343. *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: 11.271]*

   $$\sigma^n_{[i,\alpha]} \text{ is a linear isomorphism} \Leftrightarrow \alpha\neq 0$$

2. *$\forall i,j\in\{1,\ldots,n\}$ with $i\neq j$ we have*

   $$\det(\tau^n_{[i,j]})=-1$$

   so that by [theorem: 11.271] $\tau^n_{[i,j]}$ 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: 11.271] $\beta^n_{[i,j,\alpha]}$ 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 $\{e_k\}_{k\in\{1,\ldots,n\}}\subseteq X$ be a disjoint family such that $\{e_k|k\in\{1,\ldots,n\}\}$ is a basis for X and let $\Delta: X^n\to F$ be the determinant function such that $\Delta(e_1,\ldots,e_n)=1$ [see theorem: 11.255] then we have:

1.

$$\begin{aligned}\det(\sigma^n_{[i,\alpha]}) &\underset{[\text{theorem: } 11.268]}{=} \Delta(\sigma^n_{[i,\alpha]}(e_1),\ldots,\sigma^n_{[i,\alpha]}(e_n))\\ &= \Delta(\sigma^n_{[i,\alpha]}(e_1),\ldots,\sigma^n_{[i,\alpha]}(e_{i-1}),\sigma^n_{[i,\alpha]}(e_1),\sigma^n_{[i,\alpha]}(e_{i+1})\ldots,\sigma^n_{[i,\alpha]}(e_n))\\ &\underset{[\text{definition: } 11.342]}{=} \Delta(e_1,\ldots,e_{i-1},\alpha\cdot e_i,e_{i+1},\ldots,e_n)\\ &= \alpha\cdot\Delta(e_1,\ldots,e_n)\\ &= \alpha\cdot 1\\ &= \alpha\end{aligned}$$

2. Let $k\in\{1,\ldots,n\}$ then

$$\tau^n_{[i,j]}(e_k) \underset{[\text{definition: } 11.342]}{=} \begin{cases} e_i \text{ if } k=j\\ e_j \text{ if } k=i\\ e_k \text{ if } k\in\{1,\ldots,n\}\setminus\{i,j\}\end{cases} = e_{\{i\leftrightarrow j_n\}(k)}$$

so that

$$\begin{aligned}\det(\tau^n_{[i,j]}) &\underset{[\text{theorem: } 11.268]}{=} \Delta(\tau^n_{[i,j]}(e_1),\ldots,\tau^n_{[i,j]}(e_n))\\ &= \Delta\big(e_{(i\underset{\text{nn}}{\leftrightarrow}j)(1)},\ldots,e_{\{i\underset{n}{\leftrightarrow}j\}(n)}\big)\\ &= \mathrm{sign}\big(i\underset{n}{\leftrightarrow}j\big)\cdot\Delta(e_1,\ldots,e_n)\\ &\underset{i\neq j\wedge[\text{theorem:}11.211]}{=} -1\cdot\Delta(e_1,\ldots,e_n)\\ &= -1\end{aligned}$$

3. We have

$$\begin{array}{rcl} \det(\beta^n_{[i,j,\alpha]}) & \underset{[\text{theorem: } 11.268]}{=} & \Delta \\ (\beta^n_{[i,j,\alpha]}(e_1),\ldots,\beta^n_{[i,j,\alpha]}(e_n)) & = & \\ \Delta(\beta^n_{[i,j,\alpha]}(e_1)\ldots,\beta^n_{[i,j,\alpha]}(e_{i-1}),\beta^n_{[i,j,\alpha]}(e_i),\beta^n_{[i,j,\alpha]}(e_{i+1}),\ldots,\beta^n_{[i,j,\alpha]}(e_n)) & = & \\ \Delta(e_1,\ldots,e_{i-1},e_i+\alpha\cdot e_j,e_{i-1},\ldots,e_n) & = & \\ \Delta(e_1,\ldots,e_{i-1},e_i,e_{i-1},\ldots,e_n)+\Delta(e_1,\ldots,e_{i-1},\alpha\cdot e_j,e_{i-1},\ldots,e_n) & = & \\ \Delta(e_1,\ldots,e_n)+\alpha\cdot\Delta(e_1,\ldots,e_{i-1},e_j,e_{i-1},\ldots,e_n) & \underset{[\text{theorem: } 11.248]}{=} & \\ 1+\alpha\cdot 0 & & \\ 1 & & \end{array}$$

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 11.344. *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*

$$L=T_1\circ\cdots\circ T_n\underset{\text{def}}{=}\prod_{i=1}^n T_1$$

Lemma 11.345. *Let X a finite dimensional vector space with $\dim(X) = n+1$ over a field F with characteristic zero with basis $\{e_i | i \in \{1, \ldots, n+1\}\}$ determined by the distinct family $\{e_i\}_{i \in \{1, \ldots, n+1\}} \subseteq X$, $m \in \{1, \ldots, n\}$ 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 $\forall k \in \{1, \ldots, n\}$ we have

$$\left(\prod_{i=1}^{m} \beta^{n+1}_{[n+1,i,\alpha_i]}\right)(e_k) = \left\{ \begin{array}{l} e_k \text{ if } k \in \{1, \ldots, n\} \\ \sum_{i \in \{1, \ldots, n\}} \alpha_i \cdot e_i + e_{n+1} \text{ if } k = n+1 \end{array} \right.$$

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:

   $1 \in S_{n,k}$. Let $\{\alpha_i\}_{i \in \{1, \ldots, 1\}} \subseteq F$ then
$$\left(\prod_{i=1}^{1} \beta^{n+1}_{[n+1,i,\alpha_i]}\right)(e_k) = \beta^{n+1}_{[n+1,1,\alpha_1]}(e_k) \underset{k \neq n+1[\text{definition: } 11.342]}{=} e_k$$
   proving that $1 \in S_{n,k}$.

   $m \in S_{n,k} \Rightarrow m+1 \in S_{n,k}$. Assume that $m+1 \leqslant n$ and let $\{\alpha_i\}_{i \in \{1, \ldots, m+1\}} \subseteq F$ then we have:
$$\begin{aligned} \left(\prod_{i=1}^{m+1} \beta^{n+1}_{[n+1,i,\alpha_i]}\right)(e_k) &= \left(\prod_{i=1}^{m} \beta^{n+1}_{[n+1,i,\alpha_i]} \circ \beta^{n+1}_{[n+1,m+1,\alpha_{m+1}]}\right)(e_k) \\ &= \left(\prod_{i=1}^{m} \beta^{n+1}_{[n+1,i,\alpha_i]}\right)(\beta^{n+1}_{[n+1,m+1,\alpha_{m+1}]}(e_k)) \\ &\underset{k \neq n+1[\text{theorem: } 11.342]}{=} \left(\prod_{i=1}^{m} \beta^{n+1}_{[n+1,i,\alpha_i]}\right)(e_k) \\ &\underset{m \in S_{n,k}}{=} e_k \end{aligned}$$
   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:

$1\in S_n$. If $\{\alpha_i\}_{i\in\{1\}}$ then

$$\begin{aligned}\left(\prod_{i=1}^{1}\beta^{n+1}_{[n+1,i,\alpha_i]}\right)(e_{n+1}) &= \beta^{n+1}_{[n+1,1,\alpha_1]}(e_{n+1})\\ &\underset{k=n+1[\text{definition: }11.342]}{=} e_{n+1}+\alpha_1\cdot e_1\\ &= e_{n+1}+\alpha_1\cdot e_1\end{aligned}$$

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

$$\begin{aligned}\left(\prod_{i=1}^{m+1}\beta^{n+1}_{[n+1,i,\alpha_i]}\right)(e_{n+1}) &=\\ \left(\prod_{i=1}^{m}\beta^{n+1}_{[n+1,i,\alpha_i]}\circ\beta^{n+1}_{[n+1,m+1,\alpha_i]}\right)(e_{n+1}) &=\\ \prod_{i=1}^{m}\beta^{n+1}_{[n+1,i,\alpha_i]}(\beta^{n+1}_{[n+1,m+1,\alpha_i]}(e_{n+1})) &=\\ \left(\prod_{i=1}^{m}\beta^{n+1}_{[n+1,i,\alpha_i]}\right)(e_{n+1}+\alpha_{m+1}\cdot e_{m+1}) &\underset{k=n+1[\text{definition: }11.342]}{=}\\ \left(\prod_{i=1}^{m}\beta^{n+1}_{[n+1,i,\alpha_i]}\right)(e_{n+1})+\alpha_{m+1}\cdot\left(\prod_{i=1}^{m}\beta^{n+1}_{[n+1,i,\alpha_i]}\right)(e_{m+1}) &\underset{m+1\leqslant n<n+1\text{ and }(1)}{=}\\ \left(\prod_{i=1}^{m}\beta^{n+1}_{[n+1,i,\alpha_i]}\right)(e_{n+1})+\alpha_{m+1}\cdot e_{m+1} &\underset{m\in S_n}{=}\\ \sum_{i=1}^{m}\alpha_i\cdot e_i+e_{n+1}+\alpha_{m+1}\cdot e_{m+1} &=\\ \sum_{i=1}^{m+1}\alpha_i\cdot e_i+e_{n+1}\end{aligned}$$

proving that $n+1\in S_n$. □

Lemma 11.346. *Let X a finite dimensional vector space with $\dim(X)=n+1$ over a field F with characteristic zero with basis $\{e_i|i\in\{1,\ldots,n+1\}\}$ determined by the distinct family $\{e_i\}_{i\in\{1,\ldots,n+1\}}\subseteq X$, $m\in\{1,\ldots,n\}$ 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

$$\left(\prod_{i=1}^{m}\beta^{n+1}_{[i,n+1,\alpha_i]}\right)(e_k)=\left\{\begin{array}{l}e_k+\alpha_k\cdot e_{n+1}\text{ if }k\in\{1,\ldots,m\}\\ e_k\text{ if }k\in\{m+1,\ldots,n+1\}\end{array}\right.$$

Proof.

1. We use induction on m for this proof, so let

   $S_n = \left\{ m \in \mathbb{N} | \text{If } m \leqslant n \text{ then } \forall \{\alpha_i\}_{i \in \{1, \ldots, m\}} \text{ and } \forall k \in \{m+1, \ldots, n+1\} \text{ we have } \left(\prod_{i=1}^{m} \beta^{n+1}_{[i,n+1,\alpha_i]} \right)(e_k) = e_k \right\}$

   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

   $$\left(\prod_{i=1}^{1} \beta^{n+1}_{[i,n+1,\alpha_i]} \right)(e_k) = \beta^{n+1}_{[1,n+1,\alpha_1]}(e_k) \underset{k \neq 1[\text{definition: } 11.342]}{=} e_k$$

   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 $k \in \{(m+1)+1, \ldots, n+1\} \Rightarrow k \neq m+1$ and $k \in \{m+1, \ldots, n+1\}$

   $$\begin{aligned} \left(\prod_{i=1}^{m+1} \beta^{n+1}_{[i,n+1,\alpha_i]} \right)(e_k) & = \\ \left(\prod_{i=1}^{m} \beta^{n+1}_{[i,n+1,\alpha_i]} \circ \beta^{n+1}_{[m+1,n+1,\alpha_{m+1}]} \right)(e_k) & \\ \left(\prod_{i=1}^{m} \beta^{n+1}_{[i,n+1,\alpha_i]} \right)(\beta^{n+1}_{[m+1,n+1,\alpha_{m+1}]}(e_k)) & \underset{k \neq 1m+1[\text{definition: } 11.342]}{=} \\ \left(\prod_{i=1}^{m} \beta^{n+1}_{[i,n+1,\alpha_i]} \right)(e_k) & \underset{m \in S_n \wedge k \in \{m+1, \ldots, n+1\}}{=} \\ \text{eZ}_k & \end{aligned}$$

   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

   $$\begin{aligned} \left(\prod_{i=1}^{1} \beta^{n+1}_{[i,n+1,\alpha_i]} \right)(e_1) & = (\beta^{n+1}_{[1,n+1,\alpha_1]})(e_1) \\ & \underset{k=1 \wedge [\text{definition: } 11.342]}{=} e_1 + \alpha_1 \cdot e_{n+1} \\ & = e_k + \alpha_k \cdot e_k \end{aligned}$$

   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

$$\begin{aligned}
\left(\prod_{i=1}^{m+1}\beta^{n+1}_{[i,n+1,\alpha_i]}\right)(e_k) &= \\
\left(\prod_{i=1}^{m+1}\beta^{n+1}_{[i,n+1,\alpha_i]}\right)(e_{m+1}) &= \\
\left(\prod_{i=1}^{m}\beta^{n+1}_{[i,n+1,\alpha_i]}\circ\beta^{n+1}_{[m+1,n+1,\alpha_{m+1}]}\right)(e_k) &= \\
\left(\prod_{i=1}^{m}\beta^{n+1}_{[i,n+1,\alpha_i]}\right)(\beta^{n+1}_{[m+1,n+1,\alpha_{m+1}]}(e_{m+1})) &\underset{[\text{definition: }11.342]}{=} \\
\left(\prod_{i=1}^{m}\beta^{n+1}_{[i,n+1,\alpha_i]}\right)(e_{m+1}+\alpha_{m+1}\cdot e_{n+1}) &= \\
\left(\prod_{i=1}^{m}\beta^{n+1}_{[i,n+1,\alpha_i]}\right)(e_{m+1})+\alpha_{m+1}\cdot \\
\left(\prod_{i=1}^{m}\beta^{n+1}_{[i,n+1,\alpha_i]}\right)(e_{n+1}) &\underset{(1)\wedge m+1,n+1\in\{m+1,\ldots n+1\}}{=} \\
e_{m+1}+\alpha_{m+1}\cdot e_{n+1} &\underset{m+1=k}{=} \\
e_k+\alpha_k\cdot e_{n+1}
\end{aligned}$$

$k\in\{1,\ldots,m\}$. Then

$$\begin{aligned}
\left(\prod_{i=1}^{m+1}\beta^{n+1}_{[i,n+1,\alpha_i]}\right)(e_k) &= \\
\left(\prod_{i=1}^{m}\beta^{n+1}_{[i,n+1,\alpha_i]}\circ\beta^{n+1}_{[m+1,n+1,\alpha_{m+1}]}\right)(e_k) &= \\
\left(\prod_{i=1}^{m}\beta^{n+1}_{[i,n+1,\alpha_i]}\right)(\beta^{n+1}_{[m+1,n+1,\alpha_{m+1}]}(e_k)) &\underset{k\neq m+1[\text{definition: }11.342]}{=} \\
\left(\prod_{i=1}^{m}\beta^{n+1}_{[i,n+1,\alpha_i]}\right)(e_k) &\underset{k\in\{1,\ldots,m\}\wedge m\in S_n}{=} \\
e_k+\alpha_k\cdot e_{n+1}
\end{aligned}$$

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 11.347. *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. This follows from [theorems: 11.14, 11.21]
2. If L_1, L_2 are composed of elementary transformations then there exists $k, l \in \mathbb{N}$ and families of elementary transformations $\{T_i\}_{i \in \{1, \ldots, k\}} \subseteq \operatorname{Elem}(X)$, $\{S_i\}_{i \in \{1, \ldots, l\}} \subseteq \operatorname{Elem}(X)$ such that
$$L_1 = \prod_{i=1}^{k} T_i \text{ and } L_2 = \prod_{i=1}^{l} S_i$$
Define then the family $\{E_i\}_{i \in \{1, \ldots, k+l\}} \subseteq \operatorname{GL}(X)$ by
$$\forall i \in \{1, \ldots, k+l\} \; E_i = \begin{cases} T_i \text{ if } i \in \{1, \ldots, k\} \\ S_{i-n} \text{ if } i \in \{l+1, \ldots, k+l\} \end{cases}$$
then we have
$$\begin{array}{rcl} L_1 \circ L_2 & = & \prod_{i=1}^{k} T_i \circ \prod_{i=1}^{l} S_i \\ & = & \prod_{i=1}^{k} T_i \circ \prod_{i=1}^{l} S_i \\ & \underset{\text{def}}{=} & \prod_{i=1}^{k} T_i \circ \prod_{i=k+1}^{k+l} S_{i-k} \\ & = & \prod_{i=1}^{k} E_i \circ \prod_{i=k+1}^{k+l} E_{i-k} \\ & \underset{[\text{theorem: } 11.22]}{=} & \prod_{i=1}^{k+l} E_i \end{array}$$
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 11.348. *Let $n \in \mathbb{N}$, X a finite dimensional vector space over a field F with basis $E = \{e_i | i \in \{1, ..., n+1\}\}$ where $\{e_i\}_{i \in \{1, \ldots, n+1\}} \subseteq X$ is a distinct family, $Y = \operatorname{span}(\{e_i | i \in \{1, ..., n\}\})$ a n-dimensional vector space with basis $F = \{e_i | i \in \{1, \ldots, n-1\}\}$ and $L \in \operatorname{Hom}(Y, Y)$ 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 $L^{[n]} \in \operatorname{Hom}(X, X)$ by
$$L^{[n+1]} = \mathcal{M}(E, E)^{-1}(M_{[n+1,L]})$$
then we have $\forall k \in \{1, \ldots, n\}$ that
$$\begin{array}{rcl} L^{[n+1]}(e_k) & = & \sum_{i \in \{1, \ldots, n+1\}} (M_{[n+1,L]})_{i,k} \cdot e_i \\ & = & \sum_{i \in \{1, \ldots, n\}} (M_{[n+1,L]})_{i,k} \cdot e_i + \sum_{i \in \{n+1\}} (M_{[n+1,L]})_{i,k} \cdot e_i \\ & = & \sum_{i \in \{1, \ldots, n\}} (M_{[n+1,L]})_{i,k} \cdot e_i + (M_{[n+1,L]})_{n+1,k} \cdot e_{n+1} \\ & = & \sum_{i \in \{1, \ldots, n\}} \mathcal{M}(L; F, F)_{i,k} \cdot e_i + \delta_{n+1,k} \cdot e_{n+1} \\ & = & L(e_k) + \delta_{n+1,k} \cdot e_{n+1} \\ & = & L(e_k) + 0 \\ & \underset{k \neq n}{=} & L(e_k) \end{array}$$

and

$$\begin{aligned} L^{[n+1]}(e_{n+1}) &= \sum_{i\in\{1,\ldots,n+1\}} (M_{[n+1,L]})_{i,n}\cdot e_i \\ &= \sum_{i\in\{1,\ldots,n+1\}} \delta_{i,n+1}\cdot e_i \\ &= e_{n+1} \end{aligned}$$

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 11.349. *Let $n\in\mathbb{N}$, X a finite dimensional vector space over a field F with basis $E=\{e_i|i\in\{1,\ldots,n+1\}\}$ where $\{e_i\}_{i\in\{1,\ldots,n+1\}}\subseteq X$ is a distinct family, $Y=\operatorname{span}(\{e_i|i\in\{1,\ldots,n\}\})$ a n-dimensional vector space with basis $F=\{e_i|i\in\{1,\ldots,n-1\}\}$ then we have:*

1. $(\mathrm{Id}_Y)^{[n+1]}=\mathrm{Id}_X$
2. *$\forall i\in\{1,\ldots,n\}$ and $\alpha\in F\setminus\{0\}$ we have $(\sigma^n_{[i,\alpha]})^{[n+1]}=\sigma^{n+1}_{[i,\alpha]}$*
3. *$\forall i,j\in\{1,\ldots,n\}$ with $i\neq j$ then $(\tau^n_{[i,j]})^{[n+1]}=\tau^{n+1}_{[i,j]}$*
4. *$\forall i,j\in\{1,\ldots,n\}$ with $i\neq j$ and $\alpha\in F\setminus\{0\}$ 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

$$\begin{aligned} (\mathrm{Id}_X)^{[n+1]}(e_k) &= \begin{cases} \mathrm{Id}_X(e_k) \text{ if } k\in\{1,\ldots,n\} \\ e_{n+1} \text{ if } k=n+1 \end{cases} \\ &= \begin{cases} e_k \text{ if } k\in\{1,\ldots,n\} \\ e_{n+1} \text{ if } k=n+1 \end{cases} \\ &= e_k \\ &= \mathrm{Id}_Y(e_k) \end{aligned}$$

so that by [theorem: 11.292] $(\mathrm{Id}_X)^{[n+1]}=\mathrm{Id}_Y$.

2. Let $k\in\{1,\ldots,n+1\}$ then

$$\begin{aligned} (\sigma^n_{[i,\alpha]})^{[n+1]}(e_k) &= \begin{cases} \sigma^n_{[i,\alpha]}(e_k) \text{ if } k\in\{1,\ldots,n\} \\ e_{n+1} \text{ if } k=n+1 \end{cases} \\ &= \begin{cases} \alpha\cdot e_i \text{ if } k\in\{1,\ldots,n\}\wedge k=i \\ e_k \text{ if } k\in\{1,\ldots,n\}\setminus\{i\} \\ e_{n+1} \text{ if } k=n+1 \end{cases} \\ &\underset{\{1,\ldots,n\}\setminus\{i\}\subseteq\{1,\ldots,n+1\}\setminus\{i\}}{=} \begin{cases} \sigma^{n+1}_{[i,\alpha]}(e_k) \text{ if } k\in\{1,\ldots,n\} \\ \sigma^{n+1}_{[i,\alpha]}(e_k) \text{ if } k\in\{1,\ldots,n\}\setminus\{i\} \\ e_{n+1} \text{ if } k=n+1 \end{cases} \\ &\underset{i\neq n+1}{=} \begin{cases} \sigma^{n+1}_{[i,\alpha]}(e_k) \text{ if } k\in\{1,\ldots,n\} \\ \sigma^{n+1}_{[i,\alpha]}(e_k) \text{ if } k\in\{1,\ldots,n\}\setminus\{i\} \\ \sigma^{n+1}_{[i,\alpha]}(e_k) \text{ if } k=n+1 \end{cases} \\ &= \sigma^{n+1}_{[i,\alpha]}(e_k) \end{aligned}$$

so that by [theorem: 11.292] $\sigma^n_{[i,\alpha]} = \sigma^{n+1}_{[i,\alpha]}$.

3. Let $k \in \{1, \ldots, n+1\}$ then

$$\begin{aligned}
(\tau^n_{[i,j]})^{[n+1]}(e_k) &= \begin{cases} \tau^n_{[i,j]}(e_k) \text{ if } k \in \{1, \ldots, n\} \\ e_{n+1} \text{ if } k = n+1 \end{cases} \\
&= \begin{cases} e_i \text{ if } k \in \{1, \ldots, n\} \wedge k = j \\ e_j \text{ if } k \in \{1, \ldots, n\} \wedge k = i \\ e_k \text{ if } k \in \{1, \ldots, n\} \setminus \{i, j\} \\ e_{n+1} \text{ if } k = n+1 \end{cases} \\
&\underset{\{1, \ldots, n\} \setminus \{i,j\} \subseteq \{1, \ldots, n+1\} \setminus \{i,j\}}{=} \begin{cases} \tau^{n+1}_{[i,j]}(e_k) \text{ if } k \in \{1, \ldots, n\} \wedge k = j \\ \tau^{n+1}_{[i,j]}(e_k) \text{ if } k \in \{1, \ldots, n\} \wedge k = i \\ \tau^{n+1}_{[i,j]}(e_k) \text{ if } k \in \{1, \ldots, n\} \setminus \{i, j\} \\ e_{n+1} \text{ if } k = n+1 \end{cases} \\
&\underset{i,j \neq n_{+1}}{=} \begin{cases} \tau^{n+1}_{[i,j]}(e_k) \text{ if } k \in \{1, \ldots, n\} \wedge k = j \\ \tau^{n+1}_{[i,j]}(e_k) \text{ if } k \in \{1, \ldots, n\} \wedge k = i \\ \tau^{n+1}_{[i,j]}(e_k) \text{ if } k \in \{1, \ldots, n\} \setminus \{i, j\} \\ \tau^{n+1}_{[i,j]}(e_k) \text{ if } k = n+1 \end{cases} \\
&= \tau^{n+1}_{[i,j]}(e_k)
\end{aligned}$$

proving by [theorem: 11.292] $\tau^n_{[i,j]} = \tau^{n+1}_{[i,j]}$.

4. Let $k \in \{1, \ldots, n+1\}$ then

$$\begin{aligned}
(\beta^n_{[i,j,\alpha]})^{[n+1]}(e_k) &= \begin{cases} \beta^n_{[i,j,\alpha]}(x_k) \text{ if } k \in \{1, \ldots n\} \\ e_{n+1} \text{ if } k = n+1 \end{cases} \\
&= \begin{cases} e_i + \alpha \cdot e_j \text{ if } k \in \{1, \ldots, n\} \wedge k = i \\ e_k \text{ if } k \in \{1, \ldots, n\} \setminus \{i\} \\ e_{n+1} \text{ if } k = n+1 \end{cases} \\
&= \begin{cases} \beta^{n+1}_{[i,j,\alpha]}(e_k) \text{ if } k \in \{1, \ldots, n\} \wedge k = i \\ e_k \text{ if } k \in \{1, \ldots, n\} \setminus \{i\} \\ e_{n+1} \text{ if } k = n+1 \end{cases} \\
&\underset{\{1, \ldots, n\} \setminus \{i\} \subseteq \{1, \ldots, n+1\} \setminus \{i\}}{=} \begin{cases} \beta^{n+1}_{[i,j,\alpha]}(e_k) \text{ if } k \in \{1, \ldots, n\} \wedge k = i \\ \beta^{n+1}_{[i,j,\alpha]}(e_k) \text{ if } k \in \{1, \ldots, n\} \setminus \{i\} \\ e_{n+1} \text{ if } k = n+1 \end{cases} \\
&\underset{i \neq n+1}{=} \begin{cases} \beta^{n+1}_{[i,j,\alpha]}(e_k) \text{ if } k \in \{1, \ldots, n\} \wedge k = i \\ \beta^{n+1}_{[i,j,\alpha]}(e_k) \text{ if } k \in \{1, \ldots, n\} \setminus \{i\} \\ \beta^{n+1}_{[i,j,\alpha]}(e_k) \text{ if } k = n+1 \end{cases} \\
&= \beta^{n+1}_{[i,j,\alpha]}(e_k)
\end{aligned}$$

proving by [theorem: 11.292] $\beta^n_{[i,j,\alpha]} = \beta^{n+1}_{[i,j,\alpha]}$. □

The above about extending a elementary transformation to a higher dimension also applies to nonsingular transformations that are composed of elementary transformation.

Lemma 11.350. *Let $n \in \mathbb{N}$, X a finite dimensional vector space over a field F with basis $E = \{e_i | i \in \{1, \ldots, n+1\}\}$ where $\{e_i\}_{i \in \{1, \ldots, n+1\}} \subseteq X$ is a distinct family, $Y = \operatorname{span}(\{e_i | i \in \{1, \ldots, n\}\})$ a n-dimensional vector space with basis $F = \{e_i | i \in \{1,\ldots, n-1\}\}$ 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:

***k* ∈ {1, …, *n*}.** Then

$$\begin{array}{rcl}
(K \circ L)^{[n+1]}(e_k) & = & (K \circ L)(e_k) \\
& = & K(L(e_k)) \\
& = & K\left(\sum_{i \in \{1, \ldots, n\}} \mathcal{M}(L; F, F)_{i,k} \cdot e_i\right) \\
& = & \sum_{i \in \{1, \ldots, n\}} \mathcal{M}(L; F, F)_{i,k} \cdot K(e_i) \\
& = & \sum_{i \in \{1, \ldots, n\}} \mathcal{M}(L; F, F)_{i,k} \cdot K^{[n+1]}(e_i) \\
& = & K^{[n+1]}\left(\sum_{i \in \{1, \ldots, n\}} \mathcal{M}(L; F, F)_{i,k} \cdot e_i\right) \\
& = & K^{[n+1]}\left(\sum_{i \in \{1, \ldots, n\}} \mathcal{M}(L; F, F)_{i,k} \cdot e_i + 0 \cdot e_{n+1}\right) \\
& \underset{k \neq n+1}{=} & K^{[n+1]}\left(\sum_{i \in \{1, \ldots, n\}} \mathcal{M}(L; F, F)_{i,k} \cdot e_i + \delta_{n+1,k} \cdot e_{n+1}\right) \\
& \underset{\text{[see: 11.348[}}{=} & K^{[n+1]}\left(\sum_{i \in \{1, \ldots, n\}} \left(M_{[n+1,L]_{i,k}}\right) \cdot e_i + (M_{[n+1,L]})_{n+1,k} \cdot e_{n+1}\right) \\
& = & K^{[n+1]}\left(\sum_{i \in \{1, \ldots, n+1\}} \left(M_{[n+1,L]_{i,k}}\right) \cdot e_i\right) \\
& = & K^{[n+1]}(L^{[n+1]}(e_k)) \\
& = & (K^{[n+1]} \circ L^{[n+1]})(e_k)
\end{array}$$

***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: 11.292] proves that

$$(K \circ L)^{[n+1]} = K^{[n+1]} \circ L^{[n+1]}$$

□

Lemma 11.351. *Let $n \in \mathbb{N}$, X a finite dimensional vector space over a field F with basis $E = \{e_i | i \in \{1, \ldots, n+1\}\}$ where $\{e_i\}_{i \in \{1, \ldots, n+1\}} \subseteq X$ is a distinct family, $Y = \operatorname{span}(\{e_i | i \in \{1, \ldots, n\}\})$ a n-dimensional vector space with basis $F = \{e_i | i \in \{1,\ldots, n-1\}\}$ then if $m \in \mathbb{N}, \{L_i\}_{i \in \{1, \ldots, m\}} \subseteq \mathrm{GL}(Y)$ we have that*

$$\left(\prod_{i=1}^{m} L_i\right)^{[n+1]} = \prod_{i=1}^{m} (L_i)^{[n+1]}$$

or in other words

$$(L_1 \circ, \ldots, \circ L_m)^{[n+1]} = ((L_1)^{[n+1]} \circ, \ldots, \circ (L_m)^{[n+1]})$$

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:

$\boldsymbol{1 \in S.}$ Let $\{L_i\}_{i \in \{1, \ldots, 1\}} \subseteq \mathrm{GL}(Y)$ then $(\prod_{i=1}^{1} L_i)^{[n+1]} = (L_1)^{[n+1]} = \prod_{i=1}^{1} (L_i)^{[n+1]}$

$\boldsymbol{m \in S \Rightarrow m+1 \in S.}$ Let $\{L_i\}_{i \in \{1, \ldots, n+1\}} \subseteq \mathrm{GL}(X)$ then we have

$$\begin{aligned} \left(\prod_{i=1}^{m+1} L_i\right)^{[n+1]} &= \left(\prod_{i=1}^{m} L_i \circ L_{m+1}\right)^{[n+1]} \\ &\underset{[\text{lemma: } 11.350]}{=} \left(\prod_{i=1}^{m} L_i\right)^{[n+1]} \circ (L_{m+1})^{[n+1]} \\ &\underset{n \in S}{=} \prod_{i=1}^{m} (L_i)^{[n+1]} \circ (L_{m+1})^{[n+1]} \\ &= \prod_{i=1}^{m+1} (L_i)^{[n+1]} \end{aligned}$$

proving that $m+1 \in S$. □

Corollary 11.352. *Let $n \in \mathbb{N}$, X a finite dimensional vector space over a field F with basis $E = \{e_i | i \in \{1,\ldots, n+1\}\}$ where $\{e_i\}_{i \in \{1, \ldots, n+1\}} \subseteq X$ is a distinct family, $Y = \operatorname{span}(\{e_i | i \in \{1,\ldots, n\}\})$ a n-dimensional vector space with basis $F = \{e_i | i \in \{1, \ldots, n-1\}\}$ then if $L \in \mathrm{GL}(Y)$ is composed of elementary transformations in Y then $L^{[n+1]}$ 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: 11.349] we have $\forall i \in \{1, \ldots, m\}$ that $(L_i)^{[n+1]} \in \mathrm{Elem}(X)$. So as

$$L^{[n+1]} \underset{[\text{theorem: } 11.351]}{=} \prod_{i=1}^{m} (T_i)^{[n+1]}$$

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 11.353. *Let $n\in\mathbb{N}$, X a finite dimensional vector space over a filed F with characteristic zero such that $\dim(X)=n$, $L\in\mathrm{GL}(X)$ 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 $L=\sigma^1_{[1,\alpha]}$. As L is nonsingular we have also that $0\neq\det(L)=\det(\sigma^1_{[1,\alpha]})\underset{\text{[theorem: 11.343]}}{=}\alpha$ so that $\sigma^1_{[1,\alpha]}$ is a elementary transformation. Hence $L=\prod_{i=1}^{1}\sigma^1_{[i,\alpha]}$ proving 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 $E=\{e_i|i\in\{1,\ldots,n+1\}\}$ a basis for X defined by the distinct family $\{e_i\}_{i\in\{1,\ldots,n\}}\subseteq X$. 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 we have

$$\forall i\in\{1,\ldots,n\} \text{ we have } L(e_i)=\sum_{k\in\{1,\ldots,n+1\}} M_{k,i}\cdot e_k \tag{11.230}$$

As $L\in\mathrm{GL}(X)$ we have that $\det(L)\neq 0$ hence

$$\begin{array}{lll} 0 & \neq & \det(L)\\ & \underset{\text{[theorem: 11.310]}}{=} & \det(M)\\ & \underset{\text{[theorem: 11.332]}}{=} & \displaystyle\sum_{i\in\{1,\ldots,n+1\}} (-1)^{i+(n+1)}\cdot M_{i,n+1}\cdot\det([i\boxplus n+1](M')) \end{array}$$

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

$$\exists i_0\in\{1,\ldots,n+1\} \text{ such that } \det([i_0\boxplus n+1](M))\neq 0 \tag{11.231}$$

For i_0 we have now two possibilities to consider:

$i_0=n+1$. 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

$$\exists T\in\mathrm{Elem}(X) \text{ such that } L=T\circ L_1 \text{ and } \det([n+1,n+1](\mathcal{M}(L_{1;E,E})))\neq 0 \tag{11.232}$$

$i_0\in\{1,\ldots,n\}$. Take $k,l\in\{1,\ldots,n\}$ and consider $([i_0\boxplus n+1](M))_{k,l}$ then for k we have either:

$1\leqslant k<i_0$. Then $([i_0\boxplus n+1](M))_{k,l}\underset{1\leqslant l<n+1\wedge\text{[definition: 11.319]}}{=}M_{k,l}$

$i_0\leqslant k\leqslant n$. Then $([i_0\boxplus n+1](M))_{k,l}\underset{1\leqslant l<n+1\wedge\text{[definition: 11.319]}}{=}M_{k+1,l}$

proving that

$$\forall k,l\in\{1,\ldots,n\} \text{ we have } ([i_0\boxplus n+1](M))_{k,l}=\begin{cases} M_{k,l} \text{ if } 1\leqslant k<i_0\\ M_{k+1,l} \text{ if } i_0\leqslant k\leqslant n\end{cases} \tag{11.233}$$

Use [definition: 11.313] to define $M'\in\mathcal{M}_{n,n}(F)$ giving

$M'=([i_0\boxplus n+1](M))_{(n\underset{\text{ni}}{\rightsquigarrow}i_0)}$ 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 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

$$\begin{aligned}
M'_{k,l} &= ([i_0 \boxplus n+1](M))_{(n \rightsquigarrow i_0)(k),l}\\
&\underset{\text{[definition: 11.212]}}{=} ([i_0 \boxplus n+1](M))_{k,l}\\
&\underset{\text{[eq: 11.233]}}{=} \begin{cases} M_{k,l} \text{ if } 1 \leqslant k < i_0\\ M_{k+1,l} \text{ if } i_0 \leqslant k \leqslant n \end{cases}\\
&\underset{i_0=n}{=} \begin{cases} M_{k,l} \text{ if } 1 \leqslant k < n\\ M_{k+1,l} \text{ if } k = i_0 \end{cases}\\
&\underset{i_0=n}{=} \begin{cases} M_{k,l} \text{ if } k \in \{1, \ldots, n\} \setminus \{i_0\}\\ M_{k+1,l} \text{ if } k = i_0 \end{cases}
\end{aligned}$$

$\boldsymbol{i_0 \in \{1, \ldots, n-1\}}$**.** Then we have

$$\begin{aligned}
M'_{k,l} &= ([i_0 \boxplus n+1](M))_{(n \rightsquigarrow i_0)(k),l}\\
&\underset{i_0<n \wedge \text{[definition: 11.212]}}{=} \begin{cases} ([i_0 \boxplus n+1](M))_{k,l} \text{ if } 1 \leqslant k < i_0\\ ([i_0 \boxplus n+1](M))_{n,l} \text{ if } k = i_0\\ ([i_0 \boxplus n+1](M))_{k-1,l} \text{ if } i_o < k \leqslant n\\ ([i_0 \boxplus n+1](M))_{k,l} \text{ if } n < k \leqslant n \end{cases}\\
&\underset{n<l\leqslant n \text{ is impossible}}{=} \begin{cases} ([i_0 \boxplus n+1](M))_{k,l} \text{ if } 1 \leqslant k < i_0\\ ([i_0 \boxplus n+1](M))_{n,l} \text{ if } k = i_0\\ ([i_0 \boxplus n+1](M))_{k-1,l} \text{ if } i_o < k \leqslant n \end{cases}\\
&\underset{k<i_0 \wedge l \leqslant n < n+1 \wedge \text{[definition: 11.319]}}{=} \begin{cases} M_{k,l} \text{ if } 1 \leqslant k < i_0\\ ([i_0 \boxplus n+1](M))_{n,l} \text{ if } k = i_0\\ ([i_0 \boxplus n+1](M))_{k-1,l} \text{ if } i_o < k \leqslant n \end{cases}\\
&\underset{i_0<n \wedge \text{[definition: 11.319]}}{=} \begin{cases} M_{k,l} \text{ if } 1 \leqslant k < i_0\\ M_{n+1,l} \text{ if } k = i_0\\ ([i_0 \boxplus n+1](M))_{k-1,l} \text{ if } i_o < k \leqslant n \end{cases}\\
&\underset{l \leqslant n < n+1 \wedge \text{[definition: 11.319]}}{=} \begin{cases} M_{k,l} \text{ if } 1 \leqslant k < i_0\\ M_{n+1,l} \text{ if } k = i_0\\ M_{(k-1)+1,l} \text{ if } i_o < k \leqslant n \end{cases}\\
&= \begin{cases} M_{k,l} \text{ if } k \in \{1, \ldots, n\} \setminus \{i_0\}\\ M_{k+1,l} \text{ if } k = i_0 \end{cases}
\end{aligned}$$

So in all cases we have

$$\forall k, l \in \{1, \ldots, n\} \text{ we have } M'_{k,l} = \begin{cases} M_{k,l} \text{ if } k \in \{1, \ldots, n\} \setminus \{i_0\}\\ M_{k+1,l} \text{ if } k = i_0 \end{cases} \tag{11.234}$$

Further

$$\begin{aligned}
\det(M') &= \det\big(([i_0 \boxplus n+1](M))_{(n \underset{\text{ni}}{\rightsquigarrow} i_0)}\big)\\
&\underset{\text{[theorem: 11.314]}}{=} \operatorname{sign}\big(\big(n \underset{\text{ni}}{\rightsquigarrow} i_0\big)\big) \cdot \det(([i_0 \boxplus n+1](M)))\\
&\neq 0
\end{aligned}$$

proving that

$$\det(M') \neq 0 \tag{11.235}$$

Define

$$L_1 = \tau^{n+1}_{[n+1,i_0]} \circ L \tag{11.236}$$

then we have $\forall i \in \{1, \ldots, \}$ that

$$\begin{aligned}
&\sum_{k\in\{1,\ldots,n+1\}} \mathcal{M}(L_1,E,E)_{k,i}\cdot e_k =\\
&L_1(e_i) =\\
&\tau^{n+1}_{[n+1,i_0]}(L(e_i))\\
&\tau^{n+1}_{[n+1,i_0]}\left(\sum_{k\in\{1,\ldots,n+1\}} M_{k,i}\cdot e_k\right) =\\
&\sum_{k\in\{1,\ldots,n+1\}} M_{k,i}\cdot \tau^{n+1}_{[n+1,i_0]}(e_k) =\\
&\sum_{k\in\{1,\ldots,n\}\setminus\{n+1,i_0\}} M_{k,i}\cdot \tau^{n+1}_{[n+1,i_0]}(e_k) + \sum_{k\in\{n+1\}} M_{k,i}\cdot \tau^{n+1}_{[n+1,i_0]}(e_k) + \sum_{k\in\{i_0\}} M_{k,i}\cdot \tau^{n+1}_{[n+1,i_0]}(e_k) =\\
&\sum_{k\in\{1,\ldots,n\}\setminus\{n+1,i_0\}} M_{k,i}\cdot \tau^{n+1}_{[n+1,i_0]}(e_k) + M_{n+1,i}\cdot \tau^{n+1}_{[n+1,i_0]}(e_{n+1}) + M_{i_0,i}\cdot \tau^{n+1}_{[n+1,i_0]}(e_{i_0}) =\\
&\sum_{k\in\{1,\ldots,n\}\setminus\{i_0\}} M_{k,i}\cdot \tau^{n+1}_{[n+1,i_0]}(e_k) + M_{n+1,i}\cdot e_{i_0} + M_{i_0,i}\cdot e_{n+1}
\end{aligned}$$

which as the expansion in a basis is unique proves that

$$\forall k,l \in \{1,\ldots,n+1\}\ (\mathcal{M}(L_1;E,E))_{k,l} = \begin{cases} M_{k,l} \text{ if } k \in \{1,\ldots,n\}\setminus\{i_0\}\\ M_{n+1,l} \text{ if } k = i_0\\ M_{i_0,l} \text{ if } k = n+1\end{cases} \tag{11.237}$$

Now $\forall k,l \in \{1,\ldots,n\}$ we have

$$\begin{aligned}
([n+1 \boxplus n+1](\mathcal{M}(L_1;E,E)))_{k.l} &\underset{k,l\leqslant n+1}{=} \mathcal{M}(L;E,E)_{k,l}\\
&\underset{\text{[eq: 11.236}}{=} \begin{cases} M_{k,l} \text{ if } k \in \{1,\ldots,n\}\setminus\{i_0\}\\ M_{n+1,l} \text{ if } k = i_0\\ M_{i_0,l} \text{ if } k = n+1\end{cases}\\
&\underset{k,l\leqslant n+1}{=} \begin{cases} M_{k,l} \text{ if } k \in \{1,\ldots,n\}\setminus\{i_0\}\\ M_{n+1,l} \text{ if } k = i_0\end{cases}\\
&\underset{\text{[eq: 11.234]}}{=} M'_{k,l}
\end{aligned}$$

so that $[n+1 \boxplus n+1](\mathcal{M}(L_1;E,E)) = M'$, hence by [eq: 11.235] that

$$\det([n+1 \boxplus n+1](\mathcal{M}(L_1;E,E))) \neq 0 \tag{11.238}$$

Next we have

$$\tau^{n+1}_{[n+1,i_0]} \circ L_1 \underset{\text{[eq: 11.236]}}{=} \tau^{n+1}_{[n+1,i_0]} \circ (\tau^{n+1}_{[n+1,i_0]} \circ L) \underset{\text{[theorem: 11.343]}}{=} \mathrm{Id}_X \circ L = L$$

So if we take $T = \tau^{n+1}_{[n+1,i_0]} \in \mathrm{Elem}(X)$ we have together with [eq: 11.238] that

$$\exists T \in \mathrm{Elem}(X) \text{ such that } L = T \circ L_1 \text{ and } \det([n+1,n+1](\mathcal{M}(L_1;E,E))) \neq 0 \tag{11.239}$$

So in all cases we have by [eqs: 11.232, 11.239] that

$$\exists T \in \mathrm{Elem}(X) \text{ such that } L = T \circ L_1 \text{ and } \det([n+1,n+1](\mathcal{M}(L_1;E,E))) \neq 0 \tag{11.240}$$

Let

$$N = \mathcal{M}(L_1;E,E) \tag{11.241}$$

Take now $Y = \operatorname{span}(\{e_i | i \in \{1, \ldots, n\}\})$ so that $F = \{e_i | i \in \{1, \ldots, n\}\}$ is a basis for Y. Define now $L_2 \in \operatorname{Hom}(Y, Y)$ by

$$L_2 = \mathcal{M}(F, F)^{-1}([n+1, n+1]N) \tag{11.242}$$

so that $\det(L_2) \underset{\text{[theorem: 11.310]}}{=} \det([n+1, n+1]N) = \det([n+1, n+1](\mathcal{M}(L_{1;E,E}))) \neq 0$ proving that $L_2 \in \mathrm{GL}(Y)$. Hence as $n \in S$ we have that

$$L_2 \text{ is composed of elementary transformations in } Y \tag{11.243}$$

Using the above and [corollary: 11.352] it follows that

$$L_3 = L_2^{[n+1]} \text{ is composed of elementary transformations in } X \tag{11.244}$$

Define now

$$L_4 = \prod_{i=1}^{n} \beta^{n+1}_{[n+1, i, N_{i,n+1}]} \tag{11.245}$$

then as L_4 and L_3 are composed of linear transformations it follows from [lemma: 11.347] that

$$L_5 = L_4 \circ L_3 \text{ is composed of linear transformations} \tag{11.246}$$

If $i \in \{1, \ldots, n\}$ then we have

$$\begin{aligned}
L_5(e_i) &= \left(\prod_{k=1}^{n} \beta^{n+1}_{[n+1,k,N_{k,n+1}]}\right)(L_3(e_i)) \\
&= \left(\prod_{k=1}^{n} \beta^{n+1}_{[n+1,k,N_{k,n+1}]}\right)(L_2^{[n+1]}(e_i)) \\
&\underset{\text{[definition: 11.348]}}{=} \left(\prod_{k=1}^{n} \beta^{n+1}_{[n+1,k,N_{k,n+1}]}\right)(L_2(e_i)) \\
&= \left(\prod_{k=1}^{n} \beta^{n+1}_{[n+1,k,N_{k,n+1}]}\right)\left(\sum_{j=\{1,\ldots,n\}} ([n+1, n+1]N)_{j,i} \cdot e_j\right) \\
&\underset{i,j \leqslant n+1 \wedge \text{[definition: 11.319]}}{=} \left(\prod_{k=1}^{n} \beta^{n+1}_{[n+1,k,N_{k,n+1}]}\right)\left(\sum_{j=\{1,\ldots,n\}} N_{j,i} \cdot e_j\right) \\
&= \sum_{j=\{1,\ldots,n\}} N_{j,i} \cdot \left(\prod_{k=1}^{n} \beta^{n+1}_{[n+1,k,N_{k,n+1}]}\right)(e_j) \\
&\underset{\text{[theorem: 11.345]}}{=} \sum_{j=\{1,\ldots,n\}} N_{j,i} \cdot e_j
\end{aligned}$$

and

$$\begin{aligned}
L_5(e_{n+1}) &= \left(\prod_{k=1}^{n} \beta^{n+1}_{[n+1,k,N_{k,n+1}]}\right)(L_3(e_i)) \\
&= \left(\prod_{k=1}^{n} \beta^{n+1}_{[n+1,k,N_{k,n+1}]}\right)(L_2^{[n+1]}(e_{n+1})) \\
&\underset{\text{[definition: 11.348]}}{=} \left(\prod_{k=1}^{n} \beta^{n+1}_{[n+1,k,N_{k,n+1}]}\right)(e_{n+1}) \\
&\underset{\text{[theorem: 11.345]}}{=} \sum_{j\in\{1,\ldots,n\}} N_{j,n+1} \cdot e_j + e_{n+1}
\end{aligned}$$

proving that

$$\forall i \in \{1, \ldots, n+1\}\ L_5(e_i) = \begin{cases} \sum_{j\in\{1,\ldots,n\}} N_{j,i} \cdot e_j \text{ if } i \in \{1, \ldots, n\} \\ \sum_{j\in\{1,\ldots,n\}} N_{j,n+1} \cdot e_j + e_{n+1} \text{ if } i = n+1 \end{cases} \tag{11.247}$$

Take $\alpha \in F$ [later we will choose a value for α] then as $\sigma^{n+1}_{[n1,\alpha]} \in \mathrm{Elem}(X)$ and L_5 is composed of elementary transformations [see eq: 11.246] we have by [lemma: 11.347] that

$$L_6 = \sigma^{n+1}_{[n+1,\alpha]} \circ L_5 \text{ is composed of elementary transformations in } X \tag{11.248}$$

For $i \in \{1, \ldots, n\}$ we have

$$\begin{aligned}
L_6(E_i) &= \sigma^{n+1}_{[n+1,\alpha]}(L_5(e_i)) \\
&\underset{[\text{eq: } 11.247]}{=} \sigma^{n+1}_{[n+1,\alpha]}\left(\sum_{j \in \{1,\ldots,n\}} N_{j,i} \cdot e_j\right) \\
&= \sum_{j \in \{1,\ldots,n\}} N_{j,i} \cdot \sigma^{n+1}_{[n+1,\alpha]}(e_j) \\
&\underset{[\text{definition: } 11.342]}{=} \sum_{j \in \{1,\ldots,n\}} N_{j,i} \cdot e_j
\end{aligned}$$

and

$$\begin{aligned}
L_6(e_{n+1}) &= \sigma^{n+1}_{[n+1,\alpha]}(L_5(e_i)) \\
&\underset{[\text{eq: } 11.247]}{=} \sigma^{n+1}_{[n+1,\alpha]}\left(\sum_{j \in \{1,\ldots,n\}} N_{j,n+1} \cdot e_j + e_{n+1}\right) \\
&= \sigma^{n+1}_{[n+1,\alpha]}\left(\sum_{j \in \{1,\ldots,n\}} N_{j,n+1} \cdot e_j\right) + \sigma^{n+1}_{[n+1,\alpha]}(e_{n+1}) \\
&\underset{[\text{definition: } 11.342]}{=} \sigma^{n+1}_{[n+1,\alpha]}\left(\sum_{j \in \{1,\ldots,n\}} N_{j,n+1} \cdot e_j\right) + \alpha \cdot e_{n+1} \\
&= \sum_{j \in \{1,\ldots,n\}} N_{j,n+1} \cdot \sigma^{n+1}_{[n+1,\alpha]}(e_j) + \alpha \cdot e_{n+1} \\
&\underset{[\text{definition: } 11.342]}{=} \sum_{j \in \{1,\ldots,n\}} N_{j,n+1} \cdot e_j + \alpha \cdot e_{n+1}
\end{aligned}$$

proving that

$$\forall i \in \{1,\ldots,n+1\} \text{ we have } L_6(e_i) = \begin{cases} \sum_{j \in \{1,\ldots,n\}} N_{j,i} \cdot e_j \text{ if } i \in \{1,\ldots,n\} \\ \sum_{j \in \{1,\ldots,n\}} N_{j,n+1} \cdot e_j + \alpha \cdot e_{n+1} \text{ if } n+1 \end{cases} \tag{11.249}$$

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

$$L_7 \text{ is composed of linear transformations.} \tag{11.250}$$

Further if $i \in \{1,\ldots,n\}$ we have

$$\begin{aligned}
L_7(e_i) &= \left(\prod_{i \in \{1,\ldots,n\}} \beta^{n+1}_{[i,n+1,B_i]}\right)(L_6(e_i)) \\
&\underset{[\text{eq: } 11.249]}{=} \left(\prod_{i \in \{1,\ldots,n\}} \beta^{n+1}_{[i,n+1,B_i]}\right)\left(\sum_{j \in \{1,\ldots,n\}} N_{j,i} \cdot e_j\right) \\
&= \sum_{j \in \{1,\ldots,n\}} N_{j,i} \cdot \left(\prod_{i \in \{1,\ldots,n\}} \beta^{n+1}_{[i,n+1,B_i]}\right)(e_j)
\end{aligned}$$

$$\begin{aligned}
&\underset{[\text{theorem: } 11.346]}{=} && \sum_{j\in\{1,\dots,n\}} N_{j,i}\cdot(e_j+B_j\cdot e_{n+1})\\
&= && \sum_{j\in\{1,\dots,n\}} N_{j,i}\cdot e_j+\sum_{j\in\{1,\dots,n\}}(N_{j,i}\cdot B_j)\cdot e_{n+1}\\
&= && \sum_{j\in\{1,\dots,n\}} N_{j,i}\cdot e_j+\left(\sum_{j\in\{1,\dots,n\}} N_{j,i}\cdot B_j\right)\cdot e_{n+1}
\end{aligned}$$

and

$$\begin{aligned}
L_7(e_{n+1}) &=\\
\left(\prod_{i\in\{1,\dots,n\}}\beta^{n+1}_{[i,n+1,B_i]}\right)(L_6(e_{n+1})) &\underset{[\text{eq: } 11.249]}{=}\\
\left(\prod_{i\in\{1,\dots,n\}}\beta^{n+1}_{[i,n+1,B_i]}\right)\left(\sum_{j\in\{1,\dots,n\}} N_{j,n+1}\cdot e_j+\alpha\cdot e_{n+1}\right) &=\\
\left(\prod_{i\in\{1,\dots,n\}}\beta^{n+1}_{[i,n+1,B_i]}\right)\left(\sum_{j\in\{1,\dots,n\}} N_{j,n+1}\cdot e_j\right)+\\
\left(\prod_{i\in\{1,\dots,n\}}\beta^{n+1}_{[i,n+1,B_i]}\right)(\alpha\cdot e_{n+1}) &=\\
\sum_{j\in\{1,\dots,n\}} N_{j,n+1}\cdot\left(\prod_{i\in\{1,\dots,n\}}\beta^{n+1}_{[i,n+1,B_i]}\right)(e_j)+\alpha\cdot\\
\left(\prod_{i\in\{1,\dots,n\}}\beta^{n+1}_{[i,n+1,B_i]}\right)(e_{n+1}) &\underset{[\text{theorem: } 11.346]}{=}\\
\sum_{j\in\{1,\dots,n\}} N_{j,n+1}\cdot(e_j+B_j\cdot e_{n+1})+\alpha\cdot e_{n+1} &=\\
\sum_{j\in\{1,\dots,n\}} N_{j,n+1}\cdot e_j+\left(\sum_{j\in\{1,\dots,n\}} N_{j,n+1}\cdot B_j\right)\cdot e_{n+1}+\alpha\cdot e_{n+1} &=\\
\sum_{j\in\{1,\dots,n\}} N_{j,n+1}\cdot e_j+\left(\alpha+\sum_{j\in\{1,\dots,n\}} N_{j,n+1}\cdot B_j\right)\cdot e_{n+1} &=
\end{aligned}$$

proving that $\forall i\in\{1,\dots,n\}$

$$L_7(e_i) = \begin{cases}\sum_{j\in\{1,\dots,n\}} N_{j,i}\cdot e_j+\left(\sum_{j\in\{1,\dots,n\}} N_{j,i}\cdot B_j\right)\cdot e_{n+1} \text{ if } i\in\{1,\dots,n\}\\ \sum_{j\in\{1,\dots,n\}} N_{j,n+1}\cdot e_j+\left(\alpha+\sum_{j\in\{1,\dots,n\}} N_{j,n+1}\cdot B_j\right)\cdot e_{n+1} \text{ if i=n+1}\end{cases} \tag{11.251}$$

As

$$\begin{aligned}
\det(([n+1\boxplus n+1](N))^T) &\underset{[\text{theorem: } 11.309]}{=} \det(([n+1\boxplus n+1](N)))\\
&\underset{[\text{eq: } 11.241]}{=} \det(([n+1\boxplus n+1](\mathcal{M}(L_1;E,E))))\\
&\neq_{[\text{eq:}11.240]} 0
\end{aligned}$$

$([n+1\boxplus n+1](N))^T$ 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

$$\begin{aligned}
\delta_{i,j} &= ((([n+1\boxplus n+1](N))^T)\cdot K)_{i,j}\\
&= \sum_{k\in\{1,\ldots,n\}} (([n+1\boxplus n+1](N))^T)_{i,k}\cdot K_{k,j}\\
&= \sum_{k\in\{1,\ldots,n\}} ([n+1\boxplus n+1](N))_{k,i}\cdot K_{k,j}\\
&\underset{[\text{theorem: } 11.319]}{=} \sum_{k\in\{1,\ldots,n\}} N_{k,i}\cdot K_{k,j}
\end{aligned}\tag{11.252}$$

Now we choose the content of $B\in\mathcal{M}_{n+1,1}(F)$ as follows

$$\forall i\in\{1,\ldots,n+1\}\ B_i=\begin{cases}\sum_{k\in\{1,\ldots,n\}} K_{i,k}\cdot N_{n+1,k} \text{ if } i\in\{1,\ldots,n\}\\ 1 \text{ if } i=n+1\end{cases}$$

then we have $\forall i\in\{1,\ldots,n\}$ that

$$\begin{aligned}
\sum_{j\in\{1,\ldots,n\}} N_{j,i}\cdot B_j &= \sum_{j\in\{1,\ldots,n\}} N_{j,i}\cdot\left(\sum_{k\in\{1,\ldots,n\}} K_{j,k}\cdot N_{n+1,k}\right)\\
&= \sum_{j\in\{1,\ldots,n\}}\left(\sum_{k\in\{1,\ldots,n\}} N_{j,i}\cdot K_{j,k}\cdot N_{n+1,k}\right)\\
&\underset{[\text{theorem: } 11.43]}{=} \sum_{k\in\{1,\ldots,n\}}\left(\sum_{j\in\{1,\ldots,n\}} N_{j,i}\cdot K_{j,k}\cdot N_{n+1,k}\right)\\
&= \sum_{k\in\{1,\ldots,n\}}\left(\sum_{j\in\{1,\ldots,n\}} N_{j,i}\cdot K_{j,k}\right)\cdot N_{n+1,k}\\
&\underset{[\text{eq: } 11.252}{=} \sum_{k\in\{1,\ldots,n\}} \delta_{i,k}\cdot N_{n+1,k}\\
&= N_{n+1,i}
\end{aligned}$$

Substituting the above result in [eq: 11.251] results in $\forall i\in\{1,\ldots,n+1\}$

$$L_7(e_i)=\begin{cases}\sum_{j\in\{1,\ldots,n\}} N_{j,i}\cdot e_j+N_{n+1}\cdot e_{n+1} \text{ if } i\in\{1,\ldots,n\}\\ \sum_{j\in\{1,\ldots,n\}} N_{j,n+1}\cdot e_j+\left(\alpha+\sum_{j\in\{1,\ldots,n\}} N_{j,n+1}\cdot B_j\right)\cdot e_{n+1} \text{ if i=n+1}\end{cases}$$

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

$$\begin{aligned}
L_7(e_i) &= \begin{cases}\sum_{j\in\{1,\ldots,n+1\}} N_{j,i}\cdot e_j \text{ if } i\in\{1,\ldots,n\}\\ \sum_{j\in\{1,\ldots,n\}} N_{j,n+1}\cdot e_j+N_{n+1,n+1}\cdot e_{n+1} \text{ if } i=n+1\end{cases}\\
&= \begin{cases}\sum_{j\in\{1,\ldots,n+1\}} N_{j,i}\cdot e_j \text{ if } i\in\{1,\ldots,n\}\\ \sum_{j\in\{1,\ldots,n+1\}} N_{j,n+1}\cdot e_j \text{ if } i=n+1\end{cases}\\
&= \sum_{j\in\{1,\ldots,n+1\}} N_{j,i}\cdot e_j\\
&\underset{[\text{eq: } 11.241]}{=} \sum_{j\in\{1,\ldots,n+1\}} \mathcal{M}(L_1;E,E)_{j,i}\cdot e_j\\
&= L_1(e_i)
\end{aligned}$$

So that $\forall i \in \{1, \ldots, n+1\}$ we have $L_7(e_i) = L_1(e_i)$ proving by [theorem: 11.292] that $L_7 = L_1$, hence L_1 is composed of elementary transformations [see eq: 11.250]. As $L \underset{\text{[eq: 11.240]}}{=} T \circ L_1$ and $T \in \mathrm{Elem}(X)$ we have by [lemma: 11.347] 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 11.354. *Let $n \in \mathbb{N}$, X a finite dimensional vector space over a filed F with characteristic zero such that $\dim(X) = n$, $L \in \mathrm{GL}(X)$ 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 $\forall i \in \{1, \ldots, m\}$ $\det(T_I) \neq 0$ [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: 11.353] 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

$$\det(L) = \det\left(\prod_{i=1}^{m} T_i\right) \underset{\text{[theorem: 11.272]}}{=} \prod_{i \in \{1, \ldots, m\}} \det(L_i) = \left(\prod_{i \in \{1, \ldots, m\} \setminus \{k\}} \det(L_i)\right) \cdot \det(L_k) = 0$$

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
Direct Sum

Chapter 13

Tensor Product

Index