

# Analysis



# Table of contents

<b>1 Classes</b>	7
1.1 Basic concepts about classes	7
1.2 Class Operations	9
1.3 Cartesian Products	14
1.4 Graphs	17
1.5 Sets	20
1.6 Pair and triples of classes	22
1.7 Families	24
<b>2 Partial functions, functions and relations</b>	35
2.1 Partial Functions and Functions	35
2.1.1 Partial functions	35
2.1.2 Functions	36
2.1.3 Injectivity, surjectivity and bijectivity	38
2.1.4 Operations on functions and partial functions	39
2.1.5 Images and preimages of functions and partial functions	46
2.1.6 Constructing families of sets	51
2.1.7 Product, Union and Intersection of a family of sets	52
2.1.8 Family of elements in a set	58
2.2 Relations	59
2.2.1 Relations	59
2.2.2 Equivalence relations	60
2.2.3 Equivalence relations and functions	62
2.3 Partially ordered classes	64
2.3.1 Order relation	64
2.3.2 Well-ordered classes	79
2.3.3 The Axiom of Choice	83
<b>3 Algebraic Constructs</b>	93
3.1 Groups	93
3.2 Rings	95
3.3 Fields	97
<b>4 The Natural Numbers</b>	99
4.1 Definition of the Natural Numbers	99
4.2 Recursion	101
4.3 Arithmetic of the natural numbers	107
4.4 Order relation on the natural numbers	109
4.4.1 Other forms of Mathematical induction	115
<b>5 Finite and Infinite Sets</b>	117
5.1 Introduction	117
5.2 Infinite and finite sets	119
5.3 Properties of denumerable sets	133
5.4 Some properties of countable sets	136
5.5 Sequences	139
5.6 Finite Cartesian product of sets	139

5.6.1 Finite product of sets and denumerability and countability . . . . .	147
5.6.2 Notation conventions . . . . .	148
<b>6 The Integer numbers . . . . .</b>	<b>153</b>
6.1 Definition and arithmetic's . . . . .	153
6.1.1 Power in $\mathbb{Z}$ . . . . .	156
6.2 Order relation of the Integers . . . . .	157
6.3 Denumerability of the integers . . . . .	165
<b>7 The rational numbers . . . . .</b>	<b>167</b>
7.1 Definition and arithmetic's . . . . .	167
7.2 Power in $\mathbb{Q}$ . . . . .	170
7.3 Order relation . . . . .	171
7.4 Denumerability of the rationals . . . . .	177
<b>8 The real numbers . . . . .</b>	<b>179</b>
8.1 Definition . . . . .	179
8.2 Arithmetic on $\mathbb{R}$ . . . . .	181
8.2.1 Addition in $\mathbb{R}$ . . . . .	181
8.2.2 Multiplication in $\mathbb{R}$ . . . . .	183
8.2.3 Power in $\mathbb{R}$ . . . . .	200
8.3 Order relation on $\mathbb{R}$ . . . . .	201
8.4 Square root . . . . .	213
<b>9 The complex numbers . . . . .</b>	<b>217</b>
9.1 Definition and arithmetic's . . . . .	217
9.2 Order relation on $\mathbb{C}$ . . . . .	220
9.3 Norm on $\mathbb{C}$ . . . . .	222
9.4 $\mathbb{R}$ and $\mathbb{R}_{\mathbb{C}}$ . . . . .	224
9.4.1 $\mathbb{N}_{\mathbb{C}}$ , $\mathbb{Z}_{\mathbb{C}}$ and $\mathbb{Q}_{\mathbb{C}}$ . . . . .	224
9.4.2 Properties of $\mathbb{R}_{\mathbb{C}}$ . . . . .	226
9.4.3 Square root on $\mathbb{R}_{\mathbb{C}}$ . . . . .	233
9.5 Summarize . . . . .	235
<b>10 Linear Algebra . . . . .</b>	<b>237</b>
10.1 Sums and products . . . . .	237
10.1.1 Sums and associativity . . . . .	239
10.1.2 Sums and commutativity . . . . .	241
10.2 Group of permutations . . . . .	254
10.3 Finite product of sets . . . . .	274
10.4 Vector Spaces . . . . .	275
10.4.1 Definition and examples . . . . .	275
10.4.2 Linear dependency . . . . .	279
10.5 Free vector space over a set . . . . .	290
10.6 Linear Mappings . . . . .	291
10.7 Multilinear Mappings . . . . .	306
10.8 Determinant Functions . . . . .	311
10.9 Matrices . . . . .	334
10.10 Expansion of linear transformation in elementair transformations . . . . .	358
10.11 Direct internal sum of sub spaces . . . . .	367
10.12 Tensor product of vector spaces . . . . .	371
<b>11 Topology . . . . .</b>	<b>389</b>
11.1 Topological spaces . . . . .	389
11.1.1 Closed sets . . . . .	390

11.1.2 Basis of a topological space . . . . .	392
11.1.3 Dense sets . . . . .	397
11.2 Metric Spaces . . . . .	398
11.3 Normed spaces . . . . .	402
11.4 Inner product spaces . . . . .	410
11.5 Continuity . . . . .	420
11.6 Linear maps and continuity . . . . .	430
11.7 Multilinear mappings and continuity . . . . .	434
11.8 Separation . . . . .	445
11.9 Topological Vector Spaces . . . . .	447
11.10 Compact Spaces . . . . .	447
11.10.1 Filter bases in Topological spaces . . . . .	453
11.10.2 Product of compact spaces . . . . .	460
11.11 Convergence . . . . .	462
11.11.1 Sequencess . . . . .	462
11.11.2 Complete spaces . . . . .	467
11.11.3 Examples of complete spaces . . . . .	471
11.11.4 Limit infinum and limit supremum in $\mathbb{R}$ . . . . .	475
11.11.5 Series . . . . .	479
11.11.6 Series of positive numbers . . . . .	481
11.11.7 Absolute convergent series . . . . .	487
11.11.8 Properties of complete spaces . . . . .	495
11.12 Integration in Banach space . . . . .	501
11.13 Connected Sets . . . . .	509
<b>12 Fundamental theorem of algebra and the Spectral theorems . . . . .</b>	<b>513</b>
12.1 Fundamental theorem of algebra . . . . .	513
12.1.1 Polynomials . . . . .	513
12.1.2 Divergent sequences . . . . .	516
12.1.3 Properties of $\mathbb{C}$ needed for the fundamental theorem . . . . .	516
12.1.4 Proof of the fundamental theorem of algebra . . . . .	530
12.2 Spectral theorem . . . . .	537
12.2.1 Polynomials and operators . . . . .	537
12.2.2 Eigenvectors and eigen values . . . . .	541
<b>14 Differentiability in normed vector spaces . . . . .</b>	<b>561</b>
14.1 Differentiability . . . . .	561
14.1.1 Differential . . . . .	561
14.1.2 Derivative of a function . . . . .	567
14.1.3 Properties of the differential . . . . .	567
14.1.4 Partial differentiation . . . . .	570
14.2 Higher order differentiability . . . . .	581
14.3 Differentiability on general sets . . . . .	598
14.4 Intermediate value and mean value theorems . . . . .	605
14.5 Symmetry of Higher Order Differentials . . . . .	621
14.6 Higher Order Partial Differentiation . . . . .	627
14.7 Inverse function theorem . . . . .	636
<b>17 The extended real numbers . . . . .</b>	<b>665</b>
17.1 Extented reals . . . . .	665
17.2 Topology on $\bar{\mathbb{R}}$ . . . . .	683
17.3 Sequences in $\bar{\mathbb{R}}$ . . . . .	688
17.3.1 Series in $\bar{\mathbb{R}}_+$ . . . . .	708
<b>18 The exponential function . . . . .</b>	<b>723</b>

18.1 Power series . . . . .	723
18.1.1 Definition and convergence domains . . . . .	723
18.1.2 Convergence criteria . . . . .	725
18.1.3 Differentiation of power series . . . . .	728
<b>18 Measure Theory . . . . .</b>	<b>751</b>
18.1 Prerequisites . . . . .	751
18.2 Algebras and $\sigma$ -algebras . . . . .	754
18.2.1 Borel algebra on $\mathbb{R}$ . . . . .	760
18.2.2 Borel algebra on $\mathbb{R}$ . . . . .	766
18.2.3 Borel algebra on $\mathbb{R}^n$ . . . . .	769
18.3 Measures . . . . .	779
18.3.1 Measure definition and properties . . . . .	779
18.3.2 Carathéodory theorem . . . . .	785
18.3.3 The Carathéodory theorem and the set of real numbers . . . . .	791
18.3.4 The Carathéodory theorem and $\mathbb{R}^n$ . . . . .	793
18.3.5 Properties of $\langle \mathbb{R}^n, \mathcal{L}^n, \lambda^n \rangle$ and $\langle \mathbb{R}^n, \mathcal{B}[\mathbb{R}^n], (\lambda^n) _{\mathcal{B}[\mathbb{R}^n]} \rangle$ . . . . .	802
18.3.6 Negligibility, Completeness and Regularity . . . . .	806
18.4 Measurable functions . . . . .	823
18.4.1 Dynkin Classes . . . . .	827
18.5 Measurable functions . . . . .	828
18.5.1 Measurable functions . . . . .	831
18.5.2 Example of measurable functions . . . . .	833
18.5.3 Properties of measurable functions . . . . .	837
18.5.4 Properties that hold everywhere . . . . .	840
18.6 The measure integral . . . . .	853
18.6.1 Integral of simple functions . . . . .	859
18.6.2 Integral of non negative functions . . . . .	859
18.6.3 Integral of integrable functions . . . . .	871
18.6.4 Properties of the integral . . . . .	877
18.6.5 Limit theorems for integrable functions . . . . .	877
18.6.6 Integral of complex functions . . . . .	877
<b>Index . . . . .</b>	<b>880</b>

# Chapter 1

## Classes

### 1.1 Basic concepts about classes

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

**Notation 1.1.** *We introduce the following rotational terms*

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

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

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

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

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

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

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

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

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

1.  $A = A$
2.  $A = B \Rightarrow B = A$

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

**Proof.**

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

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

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

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

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

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

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

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

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

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

**Definition 1.13.** *We define the empty class  $\emptyset$  by  $\emptyset = \{x | x \neq x\}$  which is the class with no elements for if  $x \in \emptyset$  then  $x \neq x$  which is a contradiction.*

**Note 1.14.** As  $\emptyset$  is an empty set we have that  $\forall x \in \emptyset$  every proposition about  $x$  is true.

**Theorem 1.15.** For every class  $A$  we have  $\emptyset \subseteq A$

**Proof.** If  $x \notin A \Rightarrow x \notin \emptyset$  is true as  $\emptyset$  has no elements so using contra-position [meaning  $x \in \emptyset \Rightarrow x \in A$  is equivalent with  $x \notin A \Rightarrow x \notin \emptyset$ ] it follows that  $\emptyset \subseteq A$ .  $\square$

**Theorem 1.16.** If  $A$  is a class such that  $x \in A$  yields a contradiction then  $A = \emptyset$

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

**Definition 1.17.** Two classes  $A, B$  are disjoint if  $A \cap B = \emptyset$  (they have no common elements)

**Definition 1.18.** If  $A, B$  are classes then  $A \setminus B = \{x | x \in A \wedge x \notin B\}$  is the class of elements in  $A$  that are not in  $B$ .

**Theorem 1.19.** Let  $A, B$  are classes then

1.  $A \setminus (A \setminus B) = B$
2.  $A \setminus \emptyset = A$

**Proof.**

1. Let  $x \in B$  then  $x \notin A \setminus B$  so that  $x \in A \setminus (A \setminus B)$ . Further if  $x \in A \setminus (A \setminus B)$  then  $x \in A \wedge x \notin A \setminus B$  so that

$$\begin{aligned} x \in A \wedge \neg(x \in A \wedge x \notin B) &= x \in A \wedge (x \notin A \vee x \in B) \\ &= (x \in A \wedge x \notin A) \vee (x \in A \wedge x \in B) \\ &\Rightarrow x \in B \end{aligned}$$

proving that  $x \in B$

2. If  $x \in A$  then as  $x \notin \emptyset$  we have that  $x \in A$  also if  $x \in A \setminus \emptyset$  then  $x \in A$  and  $x \notin \emptyset$  so that  $x \in A$   $\square$

## 1.2 Class Operations

**Theorem 1.20.** If  $A, B$  are classes then we have

1.  $A \subseteq A \cup B$  and  $B \subseteq A \cup B$
2.  $A \cap B \subseteq A$  and  $A \cap B \subseteq B$

**Proof.**

1.  $x \in A \Rightarrow x \in A \vee x \in B \Rightarrow x \in A \cup B$ ,  $x \in B \Rightarrow x \in A \vee x \in B \Rightarrow x \in A \cup B$
2.  $x \in A \cap B \Rightarrow x \in A \wedge x \in B \Rightarrow x \in A$  and thus  $A \cap B \subseteq A$ .  $x \in A \cap B \Rightarrow x \in B$  and thus  $A \cap B \subseteq B$   $\square$

**Theorem 1.21.** If  $A, B$  are classes then we have

1.  $A \subseteq B$  if and only if  $A \cup B = B$
2.  $A \subseteq B$  if and only if  $A \cap B = A$

**Proof.**

- 1.

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

- $\Leftarrow$ . If  $A \cup B = B$  then  $x \in A \Rightarrow x \in A \cup B$   $\underset{A \cup B = B}{\Rightarrow} x \in B$  and thus  $A \subseteq B$
- 2.
- $\Rightarrow$ . If  $x \in A \underset{A \subseteq B}{\Rightarrow} x \in B \Rightarrow x \in A \wedge x \in B \Rightarrow x \in A \cap B$  proving that  $A \subseteq A \cap B$ . From the previous theorem we have  $A \cap B \subseteq A$  so by 1.8 we have  $A \cap B = A$
- $\Leftarrow$ . If  $A \cap B = A$  we have  $x \in A \Rightarrow x \in A \cap B \Rightarrow (x \in A \wedge x \in B) \Rightarrow x \in B$  so  $A \subseteq B$ .  $\square$

**Theorem 1.22. (Absorption Laws)** *If  $A, B$  are classes then*

1.  $A \cup (A \cap B) = A$
2.  $A \cap (A \cup B) = A$

**Proof.**

1. By 1.20 we have  $A \cap B \subseteq A$  and thus by the previous theorem we have  $A \cup (A \cap B) = A$
2. By 1.20 we have  $A \subseteq A \cup B$  and thus by the previous theorem we have  $A \cap (A \cup B) = A$   $\square$

**Theorem 1.23. (DeMorgan's Law)** *For all classes  $A, B, C$  we have*

1.  $C \setminus (A \cap B) = (C \setminus A) \cup (C \setminus B)$
2.  $C \setminus (A \cup B) = (C \setminus A) \cap (C \setminus B)$

**Proof.**

1.

$$\begin{aligned} x \in C \setminus (A \cap B) &\Leftrightarrow x \in C \wedge x \notin (A \cap B) \\ &\Leftrightarrow x \in C \wedge \neg(x \in A \wedge x \in B) \\ &\Leftrightarrow x \in C \wedge (x \notin A \vee x \notin B) \\ &\Leftrightarrow (x \in C \wedge x \notin A) \vee (x \in C \wedge x \notin B) \\ &\Leftrightarrow x \in C \setminus A \vee x \in C \setminus B \\ &\Leftrightarrow x \in (C \setminus A) \cup (C \setminus B) \end{aligned}$$

2.

$$\begin{aligned} x \in C \setminus (A \cup B) &\Leftrightarrow x \in C \wedge x \notin (A \cup B) \\ &\Leftrightarrow x \in C \wedge \neg(x \in A \cup B) \\ &\Leftrightarrow x \in C \wedge \neg(x \in A \vee x \in B) \\ &\Leftrightarrow x \in C \wedge (x \notin A \wedge x \notin B) \\ &\Leftrightarrow (x \in C \wedge x \notin A) \wedge (x \in C \wedge x \notin B) \\ &\Leftrightarrow x \in (C \setminus A) \wedge x \in (C \setminus B) \\ &\Leftrightarrow x \in (C \setminus A) \cap (C \setminus B) \end{aligned}$$

$\square$

**Theorem 1.24.** *If  $A, B$  are classes then*

1.  $A \cup B = B \cup A$
2.  $A \cap B = B \cap A$

**Proof.**

1. We have

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

2. We have

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

**Theorem 1.25.** For any class  $A$  we have

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

**Proof.** This follows from  $A \subseteq A$  and 1.21.  $\square$

**Theorem 1.26. (associativity)** For any class  $A, B, C$  we have

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

**Proof.**

1. We have

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

2. We have

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

**Theorem 1.27.** For every class we have

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

**Proof.**

1. As  $\emptyset \subseteq A$  (see 1.15) it follows from 1.21 that  $A \bigcup \emptyset = A$
2. As  $\emptyset \subseteq A$  (see 1.15) it follows from 1.21 that  $A \bigcap \emptyset = \emptyset$

$\square$

**Theorem 1.28.** If  $A, B, C$  are classes then we have

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

5.  $A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C)$
6.  $A \cup B = A \cup (B \setminus A)$
7. If  $A \subseteq B$  then  $B = (B \setminus A) \cup A$
8.  $(A \cup B) \setminus C = (A \setminus C) \cup (B \setminus C)$
9.  $(A \cap B) \setminus C = (A \setminus C) \cap B = (B \setminus C) \cap A$
10. If  $A, B \subseteq C$  then  $A \setminus (C \setminus B) = A \cap B$
11. If  $A \subseteq B \subseteq C$  then  $C \setminus B \subseteq C \setminus A$

**Proof.**

1. We have

$$\begin{aligned}
 x \in A \cap (B \cup C) &\Leftrightarrow x \in A \wedge x \in B \cup C \\
 &\Leftrightarrow x \in A \wedge (x \in B \vee x \in C) \\
 &\Leftrightarrow (x \in A \wedge x \in B) \vee (x \in A \wedge x \in C) \\
 &\Leftrightarrow x \in A \cap B \vee x \in A \cap C \\
 &\Leftrightarrow x \in (A \cap B) \cup (A \cap C)
 \end{aligned}$$

2. We have

$$\begin{aligned}
 x \in A \cup (B \cap C) &\Leftrightarrow x \in A \vee x \in (B \cap C) \\
 &\Leftrightarrow x \in A \vee (x \in B \wedge x \in C) \\
 &\Leftrightarrow (x \in A \vee x \in B) \wedge (x \in A \vee x \in C) \\
 &\Leftrightarrow x \in A \cup B \wedge x \in A \cup C \\
 &\Leftrightarrow x \in (A \cup B) \cap (A \cup C)
 \end{aligned}$$

3. We have

$$\begin{aligned}
 x \in A \setminus (B \cup C) &\Leftrightarrow x \in A \wedge x \notin B \cup C \\
 &\Leftrightarrow x \in A \wedge \neg(x \in B \vee x \in C) \\
 &\Leftrightarrow x \in A \wedge (x \notin B \wedge x \notin C) \\
 &\Leftrightarrow (x \in A \wedge x \notin B) \wedge (x \in A \wedge x \notin C) \\
 &\Leftrightarrow x \in (A \setminus B) \wedge x \in (A \setminus C) \\
 &\Leftrightarrow x \in (A \setminus B) \cap (A \setminus C)
 \end{aligned}$$

4. We have

$$\begin{aligned}
 x \in A \setminus (B \cap C) &\Leftrightarrow x \in A \wedge x \notin (B \cap C) \\
 &\Leftrightarrow x \in A \wedge \neg(x \in B \wedge x \in C) \\
 &\Leftrightarrow x \in A \wedge (x \notin B \wedge x \notin C) \\
 &\Leftrightarrow (x \in A \wedge x \notin B) \wedge x \notin C \\
 &\Leftrightarrow x \in A \setminus B \wedge x \notin C \\
 &\Leftrightarrow x \in (A \setminus B) \setminus C
 \end{aligned}$$

5. We have

$$\begin{aligned}
 x \in A \setminus (B \cap C) &\Leftrightarrow x \in A \wedge x \notin B \cap C \\
 &\Leftrightarrow x \in A \wedge \neg(x \in B \wedge x \in C) \\
 &\Leftrightarrow x \in A \wedge (x \notin B \vee x \notin C) \\
 &\Leftrightarrow (x \in A \wedge x \notin B) \vee (x \in A \wedge x \notin C) \\
 &\Leftrightarrow x \in A \setminus B \vee x \in A \setminus C \\
 &\Leftrightarrow x \in (A \setminus B) \cup (A \setminus C)
 \end{aligned}$$

6.

$$\begin{aligned}
x \in A \bigcup (B \setminus A) &\Leftrightarrow x \in A \vee x \in B \setminus A \\
&\Leftrightarrow x \in A \vee (x \in B \wedge x \notin A) \\
&\Leftrightarrow (x \in A \vee x \in B) \wedge (x \in A \vee x \notin A) \\
&\Leftrightarrow x \in A \vee x \in B \\
&\Leftrightarrow x \in A \bigcup B
\end{aligned}$$

7. If  $A \subseteq B$  then by (6) we have  $A \bigcup (B \setminus A) \stackrel{(6)}{=} A \bigcup B \stackrel{1.21}{=} B$ 

8.

$$\begin{aligned}
x \in (A \bigcup B) \setminus C &\Leftrightarrow x \in A \bigcup B \wedge x \notin C \\
&\Leftrightarrow (x \in A \vee x \in B) \wedge x \notin C \\
&\Leftrightarrow (x \in A \wedge x \notin C) \vee (x \in B \wedge x \notin C) \\
&\Leftrightarrow x \in A \setminus C \vee x \in B \setminus C \\
&\Leftrightarrow x \in (A \setminus C) \bigcup (B \setminus C)
\end{aligned}$$

9.

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

10. Let  $A, B \subseteq C$  then

$$\begin{aligned}
x \in A \setminus (C \setminus B) &\Leftrightarrow x \in A \wedge x \notin (C \setminus B) \\
&\Leftrightarrow x \in A \wedge \neg(x \in C \wedge x \notin B) \\
&\Leftrightarrow x \in A \wedge (x \notin C \vee x \in B) \\
&\Leftrightarrow (x \in A \wedge x \notin C) \vee (x \in A \wedge x \in B) \\
&\stackrel{A \subseteq C}{\Leftrightarrow} x \in A \wedge x \in B \\
&\Leftrightarrow x \in A \bigcap B
\end{aligned}$$

11. Let  $A \subseteq B \subseteq C$  then if  $x \in C \setminus B$  we have that  $x \in C$  and  $x \notin B$ . If  $x \in A \subseteq B$  then  $x \in B$  contradicting  $x \notin B$  hence  $x \notin A$  proving that  $x \in C \setminus A$ .  $\square$ **Theorem 1.29.** Let  $A, B, C$  be classes such that  $A \bigcap B = \emptyset$ ,  $C = A \bigcup B$  then

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

**Proof.**

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

3.  $x \in A$  then as  $A \cap B = \emptyset$  we must have  $x \notin B$  so that  $x \in A \setminus B$  hence  $A \subseteq A \setminus B \subseteq A$  so that  $A \setminus B = A$   $\square$

### 1.3 Cartesian Products

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

**Definition 1.30.** *If  $a$  is an element then  $\{a\} = \{x|x=a\}$  is a class containing only one element. The class  $\{a\}$  is called a **singleton**.*

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

**Definition 1.31.** *If  $a, b$  are elements then  $\{a, b\} = \{x|x=a \vee x=b\}$  is called a **unordered pair**.*

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

**Axiom 1.32. (Axiom of Pairing)** *If  $a, b$  are elements then  $\{a, b\}$  is an element*

**Lemma 1.33.** *If  $a$  is an element then  $\{a, a\} = \{a\}$*

**Proof.**

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

**Theorem 1.34.** *If  $a$  is an element then  $\{a\}$  is an element*

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

**Theorem 1.35.** *If  $x, y, x', y'$  are elements then from  $\{x, y\} = \{x', y'\}$  we have  $(x=x' \wedge y=y') \vee (x=y' \wedge y=x')$*

**Proof.** Let's consider the following possible cases:

**$x = y$ .** so  $\{x, y\} = \{x\} = \{x', y'\}$  and thus as  $x' \in \{x', y'\} = \{x\}$  we must have  $x = x'$  and  $y' \in \{x', y'\} = \{x\} \Rightarrow y' = x = y$  proving that  $(x = x' \wedge y = y') \Rightarrow (x = x' \wedge y = y') \vee (x = y' \wedge y = x')$

**$x \neq y$ .** then we have from  $\{x, y\} = \{x', y'\}$  the following case to consider

**$x = x'$ .** then as  $y \in \{x', y'\}$  and  $x = x' \neq y$  we must have  $y = y'$  proving that  $x = x' \wedge y = y' \Rightarrow (x = x' \wedge y = y') \vee (x = y' \wedge y = x')$

**$x = y'$ .** then as  $y \in \{x', y'\}$  and  $y' = x \neq y$  we must have  $y = x'$  proving that  $x = y' \wedge y = x' \Rightarrow (x = x' \wedge y = y') \vee (x = y' \wedge y = x')$

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

We can now define an ordered pair of elements

**Definition 1.36.** *If  $a, b$  are elements then  $(a, b) = \{\{a\}, \{a, b\}\}$  (which is an element as  $\{a, b\}$  and  $\{a\}$  are elements)*

**Theorem 1.37.** *If  $(a, b) = (c, d)$  then  $a = c$  and  $b = d$*

**Proof.** If  $(a, b) = (c, d)$  then we have by definition  $\{\{a\}, \{a, b\}\} = \{\{c\}, \{c, d\}\}$ . Using 1.35 we have then the following possible cases

$\{a\} = \{c\} \wedge \{a, b\} = \{c, d\}$ . from this we conclude that  $a = c$  and using 1.35 again we have either

$a = c \wedge b = d$ . giving the required result

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

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

**Definition 1.38.** If  $A, B$  are classes then the class  $A \times B = \{z | z = (a, b) \text{ where } a \in A \wedge b \in B\}$  is noted as  $A \times B = \{(x, y) | x \in A \wedge y \in B\}$  and is named the **Cartesian product** of  $A$  and  $B$  (note that  $A \times B$  is a class as  $x \in A \wedge y \in B$  means that  $x, y$  are elements so that  $(x, y)$  is a element)

**Example 1.39.**  $\emptyset = \emptyset \times \emptyset$

**Proof.** If  $z \in \emptyset \times \emptyset$  then there exists  $x, y \in \emptyset$  such that  $z = (x, y)$  which is a contradiction so by 1.16 we have that  $\emptyset \times \emptyset = \emptyset$ .  $\square$

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

**Lemma 1.40.** If  $A, B$  are classes then we have  $A \times B = \emptyset \Leftrightarrow A = \emptyset \vee B = \emptyset$

**Proof.**

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

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

The following is a consequence of this definition

**Lemma 1.41.** If  $A, B, C, D$  are classes with  $A \subseteq B$  and  $C \subseteq D$  then  $A \times C \subseteq B \times D$

**Proof.**

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

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

For non empty sets we can have the opposite relations

**Lemma 1.42.** If  $A, B, C, D$  are classes with  $A \neq \emptyset \wedge B \neq \emptyset$  then if  $A \times B \subseteq C \times D$  we have  $A \subseteq C \wedge B \subseteq D$

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

Combining the Cartesian with the other class operations give

TODO

**Lemma 1.43.** For all classes  $A, B, C$  and  $D$  we have

$$1. A \times (B \cap C) = (A \times B) \cap (A \times C)$$

$$2. A \times (B \cup C) = (A \times B) \cup (A \times C)$$

3.  $(A \times B) \cap (C \times D) = (A \cap C) \times (B \cap D)$
4.  $(B \cap C) \times A = (B \times A) \cap (C \times A)$
5.  $(B \cup C) \times A = (B \times A) \cup (C \times A)$
6.  $(A \times B) \setminus (C \times D) = ((A \setminus C) \times B) \cup (A \times (B \setminus D))$
7.  $(A \setminus B) \times C = (A \times C) \setminus (B \times C)$
8.  $A \times (B \setminus C) = (A \times B) \setminus (A \times C)$

**Proof.**

1. We have

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

2. We have

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

3. We have

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

4. We have

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

5. We have

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

6. We have

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

7. We have

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

and

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

8. We have

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

and

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

□

**Notation 1.44.** From now on we note the class  $\{z \mid z = (x, y) \wedge P(x, y)\}$  as  $\{(x, y) \mid P(x, y)\}$ . This makes sense for if we want to prove that a class  $A$  of tuples is equal to a class  $B$  of tuples then we must prove that  $z \in A = \{(x, y) \mid P(x, y)\} \Leftrightarrow z \in B = \{(x, y) \mid Q(x, y)\}$  but this is equivalent to  $z = (x, y) \wedge P(x, y) \Leftrightarrow z = (x', y') \wedge Q(x', y')$  which because of the fact that  $(x, y) = z = (x', y')$  implies that  $x = x' \wedge y = y'$  is equivalent with  $P(x, y) = Q(x, y)$ .

## 1.4 Graphs

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

**Definition 1.45.** Let  $A, B$  be classes then a **graph** between  $A$  and  $B$  is a subclass of  $A \times B$

**Definition 1.46.** Let  $A, B$  be classes and  $R \subseteq A \times B$  a graph then the class  $R^{-1} = \{(x, y) \in B \times A \mid (y, x) \in R\}$  is the **inverse** of the graph  $R$

**Definition 1.47.** If  $A, B, C$  are classes and  $G \subseteq A \times B$  and  $H \subseteq B \times C$  are graphs then the class  $H \circ G = \{(x, y) \in A \times C \mid \exists z \in B \text{ such that } (x, z) \in G \wedge (z, y) \in H\}$  is the **composition** of  $G$  and  $H$

**Theorem 1.48.** If  $A, B, C, D$  are classes and  $J \subseteq A \times B$ ,  $H \subseteq B \times C$ ,  $G \subseteq C \times D$

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

**Proof.**

1. We have

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

2. We have

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

3. We have

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

□

**Definition 1.49.** If  $A, B$  are classes and  $G \subseteq A \times B$  is a graph then  $\text{dom}(G)$  is the class defined by  $\text{dom}(G) = \{x \in A \mid \exists y \in B \vdash (x, y) \in G\}$

**Definition 1.50.** If  $A, B$  are classes and  $G \subseteq A \times B$  is a graph then  $\text{range}(G)$  is the class defined by  $\text{range}(G) = \{y \in B \mid \exists x \vdash (x, y) \in G\}$

**Theorem 1.51.** If  $A, B, C$  are classes and  $G \subseteq B \times C$  and  $H \subseteq A \times B$  are graphs then

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

**Proof.**

- 1.

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

2.

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

3.

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

4.

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

□

**Theorem 1.52.** If  $A, B, C$  are classes,  $G \subseteq B \times C, H \subseteq A \times B$  are graphs with  $\text{range}(H) \subseteq \text{dom}(G)$  then  $\text{dom}(G \circ H) = \text{dom}(H)$

**Proof.** Using the previous theorem we have  $\text{dom}(G \circ H) \subseteq \text{dom}(H)$  so we must prove only the opposite inclusion. So

$$\begin{aligned}
x \in \text{dom}(H) &\Rightarrow \exists y \in B \vdash (x, y) \in H \\
&\Rightarrow y \in \text{range}(H) \\
&\stackrel{\text{range}(H) \subseteq \text{dom}(G)}{\Rightarrow} y \in \text{dom}(G) \\
&\Rightarrow \exists z \in C \vdash (y, z) \in G \\
&\Rightarrow (x, z) \in G \circ H \\
&\Rightarrow x \in \text{dom}(G \circ H)
\end{aligned}$$

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

**Theorem 1.53.**  $A, B$  are classes and  $G, H \subseteq A \times B$  are graphs then

1.  $\text{dom}(G \cup H) = \text{dom}(G) \cup \text{dom}(H)$
2.  $\text{range}(G \cup H) = \text{range}(G) \cup \text{range}(H)$

**Proof.**

1. If  $x \in \text{dom}(G \cup H)$  then there exists a  $y \in B$  such that  $(x, y) \in G \cup H$  then  $(x, y) \in G \Rightarrow x \in \text{dom}(G)$  or  $(x, y) \in H \Rightarrow x \in \text{dom}(H)$  proving that

$$\text{dom}(G \cup H) \subseteq \text{dom}(G) \cup \text{dom}(H).$$

On the other hand if  $x \in \text{dom}(G) \cup \text{dom}(H)$  then if  $x \in \text{dom}(G)$  we have that  $\exists y \in B$  such that  $(x, y) \in G \subseteq G \cup H \Rightarrow x \in \text{dom}(G \cup H)$  and if  $x \in \text{dom}(H)$  there exists a  $y \in B$  such that  $(x, y) \in H \subseteq G \cup H \Rightarrow x \in \text{dom}(G \cup H)$  proving that

$$\text{dom}(G) \cup \text{dom}(H) \subseteq \text{dom}(G \cup H).$$

2. If  $y \in \text{range}(G \cup H)$  then there exists a  $x \in A$  such that  $(x, y) \in G \cup H$  then  $(x, y) \in G \Rightarrow x \in \text{range}(G)$  or  $(x, y) \in H \Rightarrow x \in \text{range}(H)$  proving that

$$\text{range}(G \cup H) \subseteq \text{range}(G) \cup \text{range}(H).$$

On the other hand if  $y \in \text{range}(G) \cup \text{range}(H)$  then if  $y \in \text{range}(G)$  we have that  $\exists x \in A$  such that  $(x, y) \in G \subseteq G \cup H \Rightarrow y \in \text{range}(G \cup H)$  and if  $y \in \text{range}(H)$  there exists a  $x \in A$  such that  $(x, y) \in H \subseteq G \cup H \Rightarrow y \in \text{range}(G \cup H)$  proving that

$$\text{range}(G) \cup \text{range}(H) \subseteq \text{range}(G \cup H).$$

□

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

**Definition 1.54. (function graph)** A graph  $G$  is a function graph if  $\forall (x, y), (x, y') \in G$  we have  $y = y'$

**Theorem 1.55.** If  $F, G$  are function graphs then  $G \circ F$  is a function graph

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

## 1.5 Sets

A special type of classes are classes that them self are elements, we call these classes sets. To work with sets we have to introduce some new axioms.

**Definition 1.56.** A set is a class that is an element of a class. In formula form  $x$  is a set  $\Leftrightarrow \exists \mathcal{A}$  with  $x \in \mathcal{A}$

Rephrasing the Axiom of Pairing [see 1.32] we have then

**Theorem 1.57.** If  $x$  and  $y$  are sets then  $\{x, y\}$  is a set

**Axiom 1.58. (Axiom of Subsets)** Every subclass of a set is a set

**Theorem 1.59.** If  $A$  is a set and  $B$  a class then  $A \cap B$  is a set. In particular we have that the intersection of two sets is a set

**Proof.** By theorem 1.20 we have  $A \cap B \subseteq A$  and thus by the Axiom of Subsets (see 1.58) we have  $A \cap B$  is a set. □

**Axiom 1.60. (Axiom of Unions)** If  $\mathcal{A}$  is a set of sets then  $\bigcup_{A \in \mathcal{A}} A$  is a set where  $\bigcup_{A \in \mathcal{A}} A = \{x \mid \exists A \in \mathcal{A} \text{ such that } x \in A\}$

**Theorem 1.61.** If  $A, B$  are sets then  $A \cup B$  is a set

**Proof.** Take  $\mathcal{A} = \{A, B\}$  then  $\mathcal{A}$  is a set by 1.57 and thus a set of sets. Using the Axiom of Unions we have then that  $\bigcup_{C \in \mathcal{A}} C$  is a set. Next we prove that  $A \cup B = \bigcup_{C \in \mathcal{A}} C$

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

□

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

**Definition 1.62.** If  $A$  is a set then  $\mathcal{P}(A)$  is the class defined by  $\mathcal{P}(A) = \{B \mid B \subseteq A\}$

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

**Axiom 1.63. (Axiom of Power Sets)** *If  $A$  is a set then  $\mathcal{P}(A)$  is a set*

**Theorem 1.64.** *If  $A$  is a set and  $P$  a predicate about  $X$  then  $\{X | X \subseteq A \wedge P(X)\}$  is a set*

**Proof.** First as  $X \subseteq A \Rightarrow X$  is a set and thus an element (see 1.58) so we have that  $X \subseteq A \wedge P(X)$  is equivalent with  $(X \text{ is a element} \wedge X \subseteq A \wedge P(X))$  and thus  $B = \{X | X \subseteq A \wedge P(X)\}$  is a class. Now if  $X \in B \Rightarrow X \subseteq A \Rightarrow X \in \mathcal{P}(A) \Rightarrow B \subseteq \mathcal{P}(X)$  and thus by the axiom of subsets 1.58 we have that  $B$  is a set.  $\square$

**Definition 1.65.** *Let  $A$  be a set then  $\mathcal{P}'(A) = \mathcal{P}(A) \setminus \{\emptyset\}$  [this is a set as  $\mathcal{P}(A)$  is a set by 1.63 and 1.58]. It is essentially the set of all nonempty subsets of  $A$*

If  $A$  is a set then  $\{A\} = \{A, A\}$  is a set by 1.57 and thus by 1.61  $A \bigcup \{A\}$  is a set. This leads to the following definition

**Definition 1.66.** *If  $A$  is a set then the set  $s(A)$  is defined by  $s(A) = A \bigcup \{A\}$*

**Definition 1.67. (Successor Set)** *We say that a set  $A$  is a successor set iff*

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

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

**Axiom 1.68. (Axiom of Infinity)** *There exists a successor set*

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

**Theorem 1.69.**  $\emptyset$  is a set

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

**Example 1.70.** We start with the set  $0 = \emptyset$  it has no elements, the successor set  $1$  is the set  $\{\emptyset\} = \{0\}$  it has one element, and  $2$  is the set  $\{\emptyset, \{\emptyset\}\} = \{0, 1\}$  which has two elements [for if  $\emptyset = \{\emptyset\}$  we would have a contradiction from  $\{\emptyset\} \subseteq \emptyset \Rightarrow \emptyset \in \emptyset \underset{\emptyset = \{x | x \neq x\}}{\Rightarrow} \emptyset \neq \emptyset$ ]

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

**Lemma 1.71.** *If  $A, B$  are sets then  $A \times B \subseteq \mathcal{P}(\mathcal{P}(A \bigcup B))$*

**Proof.** We have

$$\begin{aligned}
 (x, y) \in A \times B &\Rightarrow (x, y) \underset{1.36}{=} \{\{x\}, \{x, y\}\} \wedge x \in A \wedge y \in B \\
 &\Rightarrow \{x\} \subseteq A \subseteq A \bigcup B \wedge \{x, y\} \subseteq A \bigcup B \\
 &\Rightarrow \{x\} \subseteq A \bigcup B \wedge \{x, y\} \subseteq A \bigcup B \\
 &\Rightarrow \{x\} \in \mathcal{P}(A \bigcup B) \wedge \{x, y\} \in \mathcal{P}(A \bigcup B) \\
 &\Rightarrow \{\{x\}, \{x, y\}\} \subseteq \mathcal{P}(A \bigcup B) \\
 &\Rightarrow \{\{x\}, \{x, y\}\} \in \mathcal{P}(\mathcal{P}(A \bigcup B)) \\
 &\Rightarrow (x, y) \in \mathcal{P}(\mathcal{P}(A \bigcup B))
 \end{aligned}$$

$\square$

**Theorem 1.72.** *If  $A, B$  are sets then  $A \times B$  is a set*

**Proof.** By 1.61 we have that  $A \cup B$  is a set, so that by 1.63  $\mathcal{P}(A \cup B)$  is a set and again by 1.63 we have that  $\mathcal{P}(\mathcal{P}(A \cup B))$  is a set. Now by the previous lemma we have that  $A \times B \subseteq \mathcal{P}(\mathcal{P}(A \cup B))$  and using 1.58 we have then that  $A \times B$  is a set.  $\square$

## 1.6 Pair and triples of classes

Although we have already defined the concept of a pair, we can not simply 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  $\langle A, B \rangle = \{A, \{B\}\}$  because this would mean that  $A, B$  are elements and not pure classes. So we need another way of forming pairs, triples and so on.

**Definition 1.73.** A pair  $\langle A, B \rangle$  of classes is the class defined by  $(A \times \{\emptyset\}) \cup \{B \times \{\{\emptyset\}\}\}$  /which is a class because  $\emptyset$  is a set [see 1.69] so  $\{\emptyset\}$  and  $\{\{\emptyset\}\}$  are sets [see 1.57] and thus  $\emptyset, \{\emptyset\}, \{\{\emptyset\}\}$  are classes which means that  $x \times \{\emptyset\}, \{y \times \{\{\emptyset\}\}\}$  are defined and thus their union is a class/.

**Theorem 1.74.** If  $\langle A, B \rangle$  and  $\langle A', B' \rangle$  are pairs then  $\langle A, B \rangle = \langle A', B' \rangle$  iff  $A = A' \wedge B = B'$

**Proof.**

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

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

so we have  $A = A'$ . Further

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

so we have that  $B = B'$

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

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

$\square$

**Definition 1.75.** A triple  $\langle A, B, C \rangle$  where  $A, B, C$  are classes is the class defined by  $(A \times \{\emptyset\}) \cup (B \times \{\{\emptyset\}\}) \cup (C \times \{\{\{\emptyset\}\}\})$  (which is a class because  $\emptyset$  is a set (see 1.69) so  $\{\emptyset\}, \{\{\emptyset\}\}$  and  $\{\{\{\emptyset\}\}\}$  are sets (see 1.57) thus  $\langle A, B, C \rangle$  is a class).

**Theorem 1.76.** If  $(A, B, C)$  and  $(A', B', C')$  are triples then  $\langle A, B, C \rangle = \langle A', B', C' \rangle$  iff  $A = A' \wedge B = B' \wedge C = C'$

**Proof.**

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

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

So we have  $A = A'$

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

So we have  $B = B'$

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

So we have  $C = C'$

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

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

□

## 1.7 Families

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

**Definition 1.77.** A family of classes is a pair  $\langle G, I \rangle$  where  $G$  is a graph with  $\text{dom}(G) \subseteq I$ . If  $\langle G, I \rangle$  is a family then we define  $\forall i \in I$  that  $G_i = \{x \mid (i, x) \in G\}$ . As the  $G_i$  are the most important objects of a family we note a family as  $\{G_i\}_{i \in I}$  instead of  $\langle G, I \rangle$ . The graph  $G$  is called the graph of the family  $\{G_i\}_{i \in I}$  and is noted as  $\text{graph}(\{G_i\}_{i \in I})$ . The class  $I$  of the family is called the index of the family. So we must have  $\text{dom}(\text{graph}(\{G_i\}_{i \in I})) \subseteq I$ .

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

**Example 1.78.** If  $\{G_i\}_{i \in I}$  is a family with  $I = \emptyset$  then  $\{G_i\}_{i \in \emptyset} = \langle G, \emptyset \rangle$  where  $G$  is a graph and  $\text{dom}(G) \subseteq \emptyset \Rightarrow \text{dom}(G) = \emptyset$ . If now  $(x, y) \in G$  then  $x \in \text{dom}(G) = \emptyset$  a contradiction so we must have that  $G = \emptyset$  hence  $\{G_i\}_{i \in I} = \langle \emptyset, \emptyset \rangle$ .

**Example 1.79.** If  $G = \{(1, a), (1, b), (1, c), (2, c), (2, d), (3, a)\}$  then we can write  $G$  as  $\{G_i\}_{i \in I}$  where  $I = \{1, 2, 3\}$  and  $G_1 = \{a, b, c\}$ ,  $G_2 = \{c, d\}$  and  $G_3 = \{a\}$

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

**Example 1.80.** If  $G = \{(1, a), (1, b), (1, c), (2, c), (2, d), (3, a)\}$  then we can write  $G$  as  $\{G_i\}_{i \in I}$  where  $I = \{1, 2, 3, 4\}$  and  $G_1 = \{a, b, c\}$ ,  $G_2 = \{c, d\}$ ,  $G_3 = \{a\}$  and  $G_4 = \emptyset$

However for a family of non empty sets we have the following theorem.

**Theorem 1.81.** Let  $\{G_i\}_{i \in I}$  be a family and  $J \subseteq I$  then if we take  $G_J = \{(i, x) \in G \mid i \in J\} \subseteq G$  then  $\langle G_J, J \rangle$  is a family with  $\forall j \in J (G_J)_j = G_j$ . So we note this family as  $\{G_j\}_{j \in J}$

**Proof.** First as  $G$  is a graph there exists classes  $A, B$  such that  $G \subseteq A \times B$  which as  $G_J \subseteq G$  proves that  $G_J \subseteq A \times B$ . Second if  $j \in \text{dom}(G_j)$  then  $\exists x \in B$  such that  $(j, x) \in G_J \Rightarrow j \in J$  proving that  $\text{dom}(G_J) \subseteq J$  which proves that

$$\langle G_J, J \rangle \text{ is a graph}$$

Let  $j \in J$  then if  $x \in (G_J)_j$  we have  $(j, x) \in G_J \subseteq G$  hence  $x \in G_j$  proving that  $(G_J)_j \subseteq G_j$ . Also if  $x \in G_j$  then  $(j, x) \in G$  and as  $j \in J$  we have that  $(j, x) \in G_J$  so that  $x \in (G_J)_j$  proving that  $G_j \subseteq (G_J)_j$ . So that

$$\forall j \in J \text{ we have } (G_J)_j = G_j$$

□

**Theorem 1.82.** If  $\{G_i\}_{i \in I}$  is a family of classes then we have  $\text{dom}(\text{graph}(\{G_i\}_{i \in I})) = I \Leftrightarrow \forall i \in I \models G_i \neq \emptyset$

**Proof.**

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

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

**Proposition 1.83.** Let  $\{G_i\}_{i \in I}$  be a family of classes with  $G = \text{graph}(\{G_i\}_{i \in I})$  and  $A$  a class then  $\langle F, I \rangle$  where  $F = \{(i, x) | (i, x) \in G \wedge x \in A\} \subseteq G$  is a family with  $\forall i \in F_i = G_i \cap A$ . Hence  $\langle F, I \rangle$  can be written as  $\{G_i \cap A\}_{i \in I}$ .

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

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

proving that  $F_i = G_i \cap A$   $\square$

**Proposition 1.84.** Let  $\{G_i\}_{i \in I}$  be a family of classes with  $G = \text{graph}(\{G_i\}_{i \in I})$  and  $A$  a class then  $\langle F, I \rangle$  where  $F = G \bigcup (I \times A)$  is a family with  $\forall i \in F_i = G_i \bigcup A$ . Hence  $\langle F, I \rangle$  can be written as  $\{G_i \bigcup A\}_{i \in I}$ .

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

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

proving that  $F_i = G_i \bigcup A$   $\square$

**Proposition 1.85.** Let  $\{G_i\}_{i \in I}$  be a family of classes with  $G = \text{graph}(\{G_i\}_{i \in I})$  and  $A$  a class then  $\langle F, I \rangle$  where  $F = \{(i, x) | (i, x) \in G \wedge x \notin A\} \subseteq G$  is a family with  $\forall i \in F_i = G_i \setminus A$ . Hence  $\langle F, I \rangle$  can be written as  $\{G_i \setminus A\}_{i \in I}$ .

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

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

proving that  $F_i = G_i \setminus A$   $\square$

**Proposition 1.86.** Let  $\{G_i\}_{i \in I}$  be a family of classes with  $G = \text{graph}(\{G_i\}_{i \in I})$  and  $A$  a class then  $\langle F, I \rangle$  where  $F = (I \times A) \setminus G \subseteq I \times A$  is a family with  $\forall i \in F_i = A \setminus G_i$ . Hence  $\langle F, I \rangle$  can be written as  $\{A \setminus G_i\}_{i \in I}$ .

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

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

proving that  $F_i = A \setminus G_i$   $\square$

**Lemma 1.87.** Let  $\{G_i\}_{i \in I}$  be a family of classes with  $G = \text{graph}(\{G_i\}_{i \in I})$  and  $A$  a class then  $\langle F, I \rangle$  where  $F = \{(i, (x, a)) | (i, x) \in G \wedge a \in A\}$  is a family with  $\forall i \in F_i = G_i \times A$ . Hence  $\langle F, I \rangle$  can be written as  $\{G_i \times A\}_{i \in I}$ .

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

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

proving that  $F_i = G_i \times A$  □

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

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

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

proving that  $F_i = A \times G_i$  □

**Lemma 1.89.** Let  $I, J$  be classes and  $\{A_{(i, j)}\}_{(i, j) \in I \times J}$  a family where  $A = \text{graph}(\{A_{(i, j)}\}_{(i, j) \in I \times J})$  then we have

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

**Proof.**

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

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

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

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

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

proving that  $F_i = A_{(i, j)}$  □

**Lemma 1.90.** Let  $\{G_i\}_{i \in I}$ ,  $\{F_j\}_{j \in J}$  be families with  $G = \text{graph}(\{G_i\}_{i \in I})$  and  $F = \text{graph}(\{F_j\}_{j \in J})$  then

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

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

**Proof.**

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

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

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

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

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

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

□

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

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

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

**Definition 1.91.** If  $\{G_i\}_{i \in I}$  is a indexed family of classes then  $\bigcup_{i \in I} G_i = \{x | \exists i \in I \vdash x \in G_i\}$  or written using the graph and index we have  $\bigcup_{i \in I} G_i = \{x | \exists i \vdash i \in I \wedge (i, x) \in \text{graph}(\{G_i\}_{i \in I})\}$

**Definition 1.92.** If  $\{G_i\}_{i \in I}$  is a non empty indexed family of classes then  $\bigcap_{i \in I} G_i = \{x | \forall i \in I \models x \in G_i\}$  or written using the graph and index we have  $\bigcap_{i \in I} G_i = \{x | \forall i \in I \models (i, x) \in \text{graph}(\{G_i\}_{i \in I})\}$

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

**Example 1.93.** If  $A$  and  $B$  are classes define  $I = \{0, 1\}$  where  $0, 1$  are different elements [we will see later that there exists two different elements] and  $G = (\{0\} \times A) \bigcup (\{1\} \times B)$  we have then that

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

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

1.

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

and

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

2.

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

3.

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

□

**Example 1.94.** If  $\{G_i\}_{i \in \emptyset} = \langle \emptyset, \emptyset \rangle$  (see 1.78) then we have  $\bigcup_{i \in \emptyset} G_i = \emptyset$

**Proof.** If  $x \in \bigcup_{i \in \emptyset} G_i$  then  $\exists i \in I$  such that  $x \in G_i$  which as  $I = \emptyset$  is impossible and thus  $\bigcup_{i \in \emptyset} G_i = \emptyset$  □

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

**Definition 1.95.** If  $\mathcal{A}$  is a class then  $\bigcap_{A \in \mathcal{A}} A = \{x \mid \forall A \in \mathcal{A} \models x \in A\}$  and  $\bigcup_{A \in \mathcal{A}} A = \{x \mid \exists A \in \mathcal{A} \models x \in A\}$

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

**Definition 1.96.** If  $\mathcal{A}$  is a class of classes then  $\{A\}_{A \in \mathcal{A}}$  is equal to  $\{A_i\}_{i \in \mathcal{A}}$  where  $\text{graph}(\{A_i\}_{i \in \mathcal{A}}) = \mathcal{I}^{\mathcal{A}} = \{(A, x) \mid A \in \mathcal{A} \wedge x \in A\}$

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

**Theorem 1.97.** If  $\mathcal{A}$  is a class then  $\bigcup_{A \in \mathcal{A}} (\mathcal{I}^{\mathcal{A}})_A = \bigcup_{A \in \mathcal{A}} A$  and  $\bigcap_{A \in \mathcal{A}} (\mathcal{I}^{\mathcal{A}})_A = \bigcap_{A \in \mathcal{A}} A$

**Proof.**

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

□

**Theorem 1.98.** If  $\{G_i\}_{i \in I}$  is a family of classes and  $B$  is a class then we have

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

**Proof.**

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

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

4.

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

□

**Theorem 1.99.** If  $\{G_i\}_{i \in I}$  is a family of classes then

1. If  $J \subseteq I$  then
  - a.  $\bigcup_{i \in J} G_i \subseteq \bigcup_{i \in I} G_i$
  - b.  $\bigcap_{i \in I} G_i \subseteq \bigcap_{i \in J} G_i$
2. If  $I = J \cup K$  then
  - a.  $(\bigcup_{i \in J} G_i) \cup (\bigcup_{i \in K} G_i)$
  - b.  $(\bigcap_{i \in J} G_i) \cap (\bigcap_{i \in K} G_i)$

**Proof.**

1.
  - a. If  $x \in \bigcup_{i \in J} G_i$  then  $\exists j \in J$  such that  $x \in G_j \subseteq \bigcup_{i \in I} G_i$  [see 1.98] proving that  $\bigcup_{i \in J} G_i \subseteq \bigcup_{i \in I} G_i$

- b. If  $x \in \bigcap_{i \in I} G_i$  then  $\forall i \in I$  we have  $x \in G_i$ . Now  $\forall i \in J$  we have  $i \in I$  so that  $x \in G_i$  giving  $x \in \bigcap_{i \in J} G_i$  proving that  $\bigcap_{i \in I} G_i \subseteq \bigcap_{i \in J} G_i$
- 2.

- a. As  $J, K \subseteq I$  we have using (1) that  $\bigcup_{i \in J} G_i \subseteq \bigcup_{i \in I} G_i \wedge \bigcup_{i \in K} G_i \subseteq \bigcup_{i \in I} G_i$ . Using then 1.98 we have

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

Further if  $x \in \bigcup_{i \in I} G_i$  then we have either  $i \in J$  so that  $x \in \bigcup_{i \in J} G_i$  or  $i \in K$  so that  $x \in \bigcup_{i \in K} G_i$  proving that  $x \in (\bigcup_{i \in J} G_i) \bigcup (\bigcup_{i \in K} G_i)$  or

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

- b. As  $J, K \subseteq I$  we have using (1) that  $\bigcap_{i \in I} G_i \subseteq \bigcap_{i \in J} A_i$  and  $\bigcap_{i \in I} G_i \subseteq \bigcap_{i \in K} A_i$  hence using 1.98 we have that

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

Futher if  $x \in (\bigcap_{i \in J} G_i) \bigcap (\bigcap_{i \in K} G_i)$  then  $x \in \bigcap_{i \in J} G_i$  and  $x \in \bigcap_{i \in K} G_i$ . So  $\forall i \in J$  we have  $x \in G_i$  and  $\forall i \in K$  we have  $x \in G_i$ . Hence if  $i \in I$  then either  $i \in J$  so that  $x \in G_i$  or  $i \in K$  so that  $x \in G_i$  proving that

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

□

**Theorem 1.100. (Generalized deMorgan's Laws)** Let  $\{G_i\}_{i \in I}$  be an indexed family of classes and  $A$  a class then

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

**Proof.**

1. This is proved by

$$\begin{aligned} x \in A \setminus \left( \bigcup_{i \in I} A_i \right) &\Leftrightarrow x \in A \wedge x \notin \bigcup_{i \in I} A_i \\ &\Leftrightarrow x \in A \wedge \neg(\exists i \in I \vdash x \in A_i) \\ &\Leftrightarrow x \in A \wedge \forall i \in I \models x \notin A_i \\ &\Leftrightarrow \forall i \in I \models x \in A \setminus A_i \\ &\Leftrightarrow x \in \bigcap_{i \in I} (A \setminus A_i) \end{aligned}$$

2. This is proved by

$$\begin{aligned} x \in A \setminus \left( \bigcap_{i \in I} A_i \right) &\Leftrightarrow x \in A \wedge x \notin \bigcap_{i \in I} A_i \\ &\Leftrightarrow x \in A \wedge \neg(\forall i \in I \models x \in A_i) \\ &\Leftrightarrow x \in A \wedge (\exists i \in I \models x \notin A_i) \\ &\Leftrightarrow \exists i \in I \models x \in A \setminus A_i \\ &\Leftrightarrow x \in \bigcup_{i \in I} (A \setminus A_i) \end{aligned}$$

□

**Theorem 1.101.** If  $\{A_i\}_{i \in I}$  and  $\{B_j\}_{j \in J}$  are indexed families of classes and a class  $C$  then

1.  $(\bigcup_{i \in I} A_i) \cap C = \bigcup_{i \in I} (A_i \cap C)$
2.  $(\bigcap_{i \in I} A_i) \cup C = \bigcap_{i \in I} (A_i \cup C)$
3.  $(\bigcup_{i \in I} A_i) \cup C = \bigcup_{i \in I} (A_i \cup C)$
4.  $(\bigcap_{i \in I} A_i) \cap C = \bigcap_{i \in I} (A_i \cap C)$

**Proof.**

1. This is proved by

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

2. This is proved by

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

- 3.

$$\begin{aligned} x \in \left( \bigcup_{i \in I} A_i \right) \cup C &\Leftrightarrow x \in \bigcup_{i \in I} A_i \vee x \in C \\ &\Leftrightarrow x \in C \vee x \in \bigcup_{i \in I} A_i \\ &\Leftrightarrow x \in C \vee (\exists i \in I \models x \in A_i) \\ &\Leftrightarrow \exists i \in I \models (x \in C \vee x \in A_i) \\ &\Leftrightarrow x \in \bigcup_{i \in I} (A_i \cup C) \end{aligned}$$

- 4.

$$\begin{aligned} x \in \left( \bigcap_{i \in I} A_i \right) \cap C &\Leftrightarrow x \in C \wedge x \in \bigcap_{i \in I} A_i \\ &\Leftrightarrow x \in C \wedge \forall i \in I \models x \in A_i \\ &\Leftrightarrow \forall i \in I \models (x \in A_i \wedge x \in C) \\ &\Leftrightarrow x \in \left( \bigcap_i (A_i \cap C) \right) \end{aligned}$$

□

**Theorem 1.102.** If  $A$  is a class and  $\{A_i\}_{i \in I}$  a family of classes then

1.  $A \setminus (\bigcap_{i \in I} A_i) = \bigcup_{i \in I} (A \setminus A_i)$
2.  $A \setminus (\bigcup_{i \in I} A_i) = \bigcap_{i \in I} (A \setminus A_i)$
3.  $(\bigcup_{i \in I} A_i) \setminus A = \bigcup_{i \in I} (A_i \setminus A)$

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

**Proof.**

1. We have

$$\begin{aligned}
 x \in A \setminus \left( \bigcap_{i \in I} A_i \right) &\Leftrightarrow x \in A \wedge \left( x \notin \bigcap_{i \in I} A_i \right) \\
 &\Leftrightarrow x \in A \wedge \neg(\forall i \in I \models x \in A_i) \\
 &\Leftrightarrow x \in A \wedge (\exists i \in I \vdash x \notin A_i) \\
 &\Leftrightarrow \exists i \in I \vdash x \in A \wedge x \notin A_i \\
 &\Leftrightarrow \exists i \in I \vdash x \in A \setminus A_i \\
 &\Leftrightarrow x \in \bigcup_{i \in I} (A \setminus A_i)
 \end{aligned}$$

2. We have

$$\begin{aligned}
 x \in A \setminus \left( \bigcup_{i \in I} A_i \right) &\Leftrightarrow x \in A \wedge \left( x \notin \bigcup_{i \in I} A_i \right) \\
 &\Leftrightarrow x \in A \wedge \neg(\exists i \in I \vdash x \in A_i) \\
 &\Leftrightarrow x \in A \wedge (\forall i \in I \models x \notin A_i) \\
 &\Leftrightarrow \forall i \in I \vdash x \in A \wedge x \notin A_i \\
 &\Leftrightarrow \forall i \in I \vdash x \in A \setminus A_i \\
 &\Leftrightarrow x \in \bigcap_{i \in I} (A \setminus A_i)
 \end{aligned}$$

- 3.

$$\begin{aligned}
 x \in \left( \bigcup_{i \in I} A_i \right) \setminus A &\Leftrightarrow x \notin A \wedge \exists i \in I \vdash x \in A_i \\
 &\Leftrightarrow \exists i \in I \vdash (x \notin A \wedge x \in A_i) \\
 &\Leftrightarrow \exists i \in I \vdash x \in A_i \setminus A \\
 &\Leftrightarrow x \in \bigcup_{i \in I} (A_i \setminus A)
 \end{aligned}$$

- 4.

$$\begin{aligned}
 x \in \left( \bigcap_{i \in I} A_i \right) \setminus A &\Leftrightarrow x \notin A \wedge \forall i \in I \models x \in A_i \\
 &\Leftrightarrow \forall i \in I \models (x \in A_i \wedge x \notin A) \\
 &\Leftrightarrow \forall i \in I \models x \in A_i \setminus A \\
 &\Leftrightarrow x \in \bigcap_{i \in I} (A_i \setminus A)
 \end{aligned}$$

5. We have

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

6. We have

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

7. We have

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

8. We have

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

□

**Theorem 1.103.** Let  $\{A_i\}_{i \in I}$  be a family of classes then  $\bigcup_{i \in I} A_i = \bigcup_{i \in \{j \in I \mid A_j \neq \emptyset\}} A_i$

**Proof.**

$$\begin{aligned}
 x \in \bigcup_{i \in I} A_i &\Rightarrow \exists i \in I \vdash x \in A_i \\
 &\stackrel{x \in A_i \Rightarrow A_i \neq \emptyset}{\Rightarrow} \exists i \in \{j \in I \mid A_j \neq \emptyset\} \vdash x \in A_i \\
 &\Rightarrow x \in \bigcup_{i \in \{j \in I \mid A_j \neq \emptyset\}} A_i \\
 x \in \bigcup_{i \in \{j \in I \mid A_j \neq \emptyset\}} A_i &\Rightarrow \exists i \in \{j \in I \mid A_j \neq \emptyset\} \vdash x \in A_i \\
 &\stackrel{\{j \in I \mid A_j \neq \emptyset\} \subseteq I}{\Rightarrow} \exists i \in I \vdash x \in A_i \\
 &\Rightarrow x \in \bigcup_{i \in I} A_i
 \end{aligned}$$

□



# Chapter 2

## Partial functions, functions and relations

### 2.1 Partial Functions and Functions

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

#### 2.1.1 Partial functions

**Definition 2.1.** A *partial function from A to B* is a triple  $\langle f, A, B \rangle$  where  $A, B$  are classes and  $f \subseteq A \times B$  is a function graph [hence  $\forall (x, y), (x, y') \in f$  we have  $y = y'$  [see 1.54]].

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

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

**Theorem 2.3.** If  $A, B$  are classes and  $f$  is a graph [hence there exists classes  $C, D$  such that  $f \subseteq C \times D$ ] then  $\langle f, A, B \rangle$  [or  $f: A \rightarrow B$ ] is a partial function if and only if

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

**Proof.**

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

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

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

**Definition 2.4.** Let  $f: A \rightarrow B$  be a partial function then if  $C \subseteq A$  and  $D \subseteq B$  we define  $f(C) = \{y \in B \mid \exists x \in C \vdash (x, y) \in f\}$  [called **the image of C by f**] and  $f^{-1}(D) = \{x \in A \mid \exists y \in D \vdash (x, y) \in f\}$  [called **the preimage of D by f**].

**Theorem 2.5.** If  $f: A \rightarrow B$  is a partial function and  $g: C \rightarrow D$  is a partial function then  $g \circ f: A \rightarrow D$  is a partial function with  $\text{range}(g \circ f) \subseteq \text{range}(g)$  and  $\text{dom}(g \circ f) \subseteq \text{dom}(f)$ .

**Proof.** We use the previous theorem [see 2.3] to prove this.

1. Using 1.51 we have that  $\text{range}(g \circ f) \subseteq \text{range}(g) \subseteq D$
2. Again using 1.51 we have that  $\text{dom}(g \circ f) \subseteq \text{dom}(f) \subseteq A$
3. As  $g, f$  are function graphs we have by 1.55 that  $g \circ f$  is a function graph.  $\square$

**Note 2.6.** If  $f: A \rightarrow B$  and  $g: C \rightarrow D$  are partial functions then

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

**Theorem 2.7.** If  $f: A \rightarrow B$  and  $g: C \rightarrow D$  are partial functions then  $\text{dom}(g \circ f) = \text{dom}(f) \cap f^{-1}(\text{dom}(g))$  and  $\text{range}(g \circ f) = g(\text{range}(f) \cap \text{dom}(g))$  for  $g \circ f: A \rightarrow D$

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

$$\text{dom}(g \circ f) \subseteq \text{dom}(f) \cap f^{-1}(\text{dom}(g)) \quad (2.1)$$

If  $x \in \text{dom}(f) \cap f^{-1}(\text{dom}(g))$  then  $x \in \text{dom}(f) \wedge x \in f^{-1}(\text{dom}(g))$  so there exists a  $y \in B$  such that  $(x, y) \in f$  and  $y' \in \text{dom}(g)$  with  $(x, y') \in f$ . As  $f$  is a function graph we have  $y = y'$  so that  $y \in \text{dom}(g)$ . As  $y \in \text{dom}(g)$  there exists a  $z$  such that  $(y, z) \in g$  hence  $(x, z) \in g \circ f$  or  $x \in \text{dom}(g \circ f)$ . So

$$\text{dom}(f) \cap f^{-1}(\text{dom}(g)) \subseteq \text{dom}(g \circ f) \quad (2.2)$$

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

$$\text{range}(g \circ f) \subseteq g(\text{range}(f) \cap \text{dom}(g)) \quad (2.3)$$

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

$$g(\text{range}(f) \cap \text{dom}(g)) \subseteq \text{range}(g \circ f) \quad (2.4)$$

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

### 2.1.2 Functions

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

**Definition 2.8. (function)** A partial function  $f: A \rightarrow B$  is a **function** if its graph satisfies  $\forall x \in A \models \exists y \in B \vdash (x, y) \in f$ . The class of all the graphs that satisfies  $\forall x \in A \models \exists y \in B \vdash (x, y) \in f$  is noted as  $B^A$ . So

$$B^A = \{f \subseteq A \times B \mid f \text{ is a function graph and } \forall x \in A \models \exists y \in B \vdash (x, y) \in f\}$$

**Proposition 2.9.** Let  $A, B, C$  are classes then  $B^A \subseteq C^A$

**Proof.** Let  $f \in B^A$  then  $f \subseteq A \times B$ , if  $(x, y) \in f \wedge (x', y) \in f$  then  $x = x'$  and  $\forall x \in A$  there exists a  $y \in B$  such that  $(x, y) \in f$ . Which as  $A \subseteq C$  gives  $f \subseteq A \times C$ , if  $(x, y) \in f \wedge (x', y) \in f$  then  $x = x'$  and  $\forall x \in A$  there exists a  $y \in C$  such that  $(x, y) \in f$ , proving that  $f \in C^A$ .  $\square$

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

**Proof.**

$\Rightarrow$ . If  $x \in \text{dom}(f) \Rightarrow \exists y \vdash (x, y) \in f \subseteq A \times B \Rightarrow x \in A \Rightarrow \text{dom}(f) \subseteq A$ . If  $x \in A \unders{f \text{ is a function}}{\Rightarrow} \exists y \in B \vdash (x, y) \in f \Rightarrow x \in \text{dom}(f) \Rightarrow A \subseteq \text{dom}(f)$ . So  $\text{dom}(f) = A$

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

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

**Example 2.11.**  $\emptyset: \emptyset \rightarrow \emptyset$  is a function [the null function].

**Proof.** Using 1.39 we have  $\emptyset = \emptyset \times \emptyset \Rightarrow \emptyset \subseteq \emptyset \times \emptyset$  the other requirements for a function are vacuously satisfied.  $\square$

**Notation 2.12.** If  $f: A \rightarrow B$  is a function then for every  $x \in A$  we have a unique  $y \in B$  such that  $(x, y) \in f$ . In many cases we have a mathematical expression like  $3 \cdot x + 1$  to calculate  $y$ . So we have  $\forall x \in \text{dom}(f)$  that  $(x, 3 \cdot x + 1) \in f$ . To express this we use the following equivalent notation to define a function based on a expression

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

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

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

or

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

or

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

In all of the above cases we actually means that  $\langle f, A, B \rangle$  is a function with  $f = \{(x, 3 \cdot x + 1) | x \in A\}$ . If we talk about a function  $f: A \rightarrow B$  and given  $x \in A$  we want to refer to the unique  $y$  such that  $(x, y) \in f$  then we use also the notation  $f(x)$  for  $y$ . Hence if  $z = f(x)$  then  $(x, z) \in f$ .

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

**Proof.**

$\Rightarrow$ . Assume that  $f = g$  then if  $x \in A = \text{dom}(f) = \text{dom}(g)$  then if  $y = f(x)$  and  $y' = g(x)$  we have that  $(x, y) \in f$  and  $(x, y') \in g$  which as  $f = g$  proves that  $(x, y), (x, y') \in g$  proving that  $y = y'$  hence  $f(x) = g(x)$ .

$\Leftarrow$ . Assume that  $\forall x \in A$  we have  $f(x) = g(x)$  then if  $(x, y) \in f$  then  $y = f(x) = g(x)$  so that  $(x, y) \in g$  and if  $(x, y) \in g$  we have that  $y = g(x) = f(x)$  so that  $(x, y) \in f$ . Hence  $f = g$ .  $\square$

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

**Proof.** First we show that  $\mathcal{X}_B$  is a function graph so let  $(x, y), (x, y') \in \mathcal{X}_B$  then we have for  $x$  either

$x \in B$ . then  $x \notin A \setminus B$  so that  $(x, y), (x, y') \in B \times \{1\} \Rightarrow y = 1 = y'$

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

Further if  $x \in A$  then either  $x \in B \Rightarrow (x, 1) \in \mathcal{X}_B$  or  $x \in A \setminus B$  then  $(x, 0) \in \mathcal{X}_B$  so that  $x \in \text{dom}(f)$  giving  $A \subseteq \text{dom}(f) \subseteq A$  proving that  $A = \text{dom}(f)$ .  $\square$

**Example 2.15.** Let  $X, Y$  be sets,  $y \in Y$  then  $C_{X,Y,y}: X \rightarrow Y$  where  $C_{X,Y,y} = \{(x, y) | x \in X\}$  is a function called a constant function.

**Note 2.16.** If  $X, Y$  are known from the context then we sometimes write  $C_y$  or  $C_{X,y}$  instead of  $C_{X,Y,y}$

**Proof.** First we prove that  $C_{X,Y,y}$  is a function graph so let  $(x, z), (x, z') \in C_{X,Y,y}$  then  $z = y = z'$ . Second if  $x \in X$  then  $(x, y \in C_{X,Y,y})$  proving that  $\text{dom}(C_{X,Y,y}) = X$ .  $\square$

### 2.1.3 Injectivity, surjectivity and bijectivity

**Definition 2.17. (Injectivity)** A partial function  $f: A \rightarrow B$  is injective iff  $(x, y) \in f \wedge (x', y) \in f \Rightarrow x = x'$ . Using the  $f(x)$  notation this is equivalent with  $f(x) = f(x') \Rightarrow x = x'$ .

**Theorem 2.18.** If  $f: A \rightarrow B$  and  $g: C \rightarrow D$  are injective partial functions then  $g \circ f: A \rightarrow D$  is a injective partial function.

**Proof.** By 2.5  $g \circ f$  is a partial function. Now if  $(x, z), (x', z) \in g \circ f$  then  $\exists y, y' \in B \vdash (x, y), (x', y) \in f \wedge (y, z), (y', z) \in g$   $\xrightarrow{g \text{ is injective}} y = y' \Rightarrow (x, y), (x', y) \in f \xrightarrow{f \text{ is injective}} x = x'$  proving that  $g \circ f$  is injective.  $\square$

**Definition 2.19.** A partial function  $f: A \rightarrow B$  is invertible if  $f^{-1}: B \rightarrow A$  is a partial function

**Lemma 2.20.** A partial function  $f: A \rightarrow B$  is invertible if and only if  $f$  is injective

**Proof.**

$\Rightarrow$ . Assume that  $f^{-1}: B \rightarrow A$  is a partial function then if  $(x, y), (x', y) \in f \Rightarrow (y, x), (y, x') \in f^{-1}$   $\xrightarrow{f^{-1} \text{ is a partial function}} x = x'$  so we have that  $f: A \rightarrow B$  is injective.

$\Leftarrow$ . Assume now that  $f$  is injective then we use 2.3 to prove that  $f^{-1}$  is a partial function.

1.  $x \in \text{range}(f^{-1}) \Rightarrow \exists y \vdash (y, x) \in f^{-1} \Rightarrow (x, y) \in f \Rightarrow x \in \text{dom}(f) \xrightarrow{f \text{ is a partial function}} x \in A$
2.  $x \in \text{dom}(f^{-1}) \Rightarrow \exists y \vdash (x, y) \in f^{-1} \Rightarrow (y, x) \in f \Rightarrow x \in \text{range}(f) \xrightarrow{f \text{ is a partial function}} x \in B$
3.  $(x, y), (x, y') \in f^{-1} \Rightarrow (y, x), (y', x) \in f \xrightarrow{f \text{ is injective}} y = y'$   $\square$

**Corollary 2.21.** If  $f: A \rightarrow B$  is a injective function then  $f^{-1}: f(A) \rightarrow A$  is a function

**Proof.** By the above lemma we have that  $f^{-1}: B \rightarrow A$  is a partial function, so  $f^{-1}$  is a function graph. Now

$$\begin{aligned} x \in \text{dom}(f^{-1}) &\Leftrightarrow \exists y \in A \vdash (x, y) \in f^{-1} \\ &\Leftrightarrow \exists y \in A \vdash (y, x) \in f \\ &\Leftrightarrow x \in f(A) \end{aligned}$$

proving that  $\text{dom}(f^{-1}) = f(A)$ . Further if  $(x, y) \in f^{-1} \subseteq B \times A$  then  $x \in \text{dom}(f^{-1}) = f(A)$  and  $y \in A$  hence  $f^{-1} \subseteq f(A) \times A$ . So  $f^{-1}: f(A) \rightarrow A$  is a function.  $\square$

**Definition 2.22. (Surjectivity)** A partial function  $f: A \rightarrow B$  is surjective iff  $\forall y \in B \models \exists x \in A \vdash (x, y) \in f$  /or  $\forall y \in B \models \exists x \in A \vdash y = f(x)$

**Note 2.23.** A partial function  $f: A \rightarrow B$  is surjective iff  $\text{range}(f) = B$

**Proof.**

$\Rightarrow$ . Assume that  $f$  is surjective. If  $y \in \text{range}(f) \Rightarrow \exists x \in A \vdash (x, y) \in f \subseteq A \times B \Rightarrow y \in B \Rightarrow \text{range}(f) \subseteq B$ . If  $y \in B \xrightarrow{f \text{ is surjective}} \exists x \in A \vdash (x, y) \in f \Rightarrow y \in \text{range}(f) \Rightarrow B \subseteq \text{range}(f)$ . So we conclude that  $\text{range}(f) = B$

$\Leftarrow$ . Assume that  $\text{range}(f) = B$ . Then if  $y \in B \Rightarrow y \in \text{range}(f) \Rightarrow \exists x \in A \vdash (x, y) \in f \Rightarrow \forall y \in B \vdash \exists x \in A \vdash (x, y) \in f \Rightarrow f$  is surjective.  $\square$

**Definition 2.24. (Bijection)** A function  $f: A \rightarrow B$  is bijective iff  $f$  is surjective and  $f$  is injective.

**Definition 2.25. (Bijective Classes)** If  $A$  and  $B$  are classes such that there exists a bijective function  $f: A \rightarrow B$  between  $A$  and  $B$  then we say that  $A$  and  $B$  are **bijective** or  $A$  and  $B$  are in a **one to one correspondence** and we note this as  $A \approx B$

**Example 2.26. (Identity Function)** If  $A$  is a class then  $\mathbb{1}_A: A \rightarrow A$  defined by  $\mathbb{1}_A = \{(x, x) | x \in A\}$  is a bijection.

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

**Example 2.27. (Inclusion Function)** If  $A$  is a class and  $B \subseteq A$  is a subclass then  $i_B: B \rightarrow A$  defined by  $i_B = \{(x, x) | x \in B\}$  is a injective function.

**Proof.** We have trivially that  $i_B \subseteq B \times B \subseteq B \times A$  and  $(x, y) \in i_B \wedge (x, y') \in i_B \Rightarrow y = x = y' \Rightarrow y = y'$  so  $i_B$  is a partial function. Next if  $x \in B$  then  $(x, x) \in i_B \Rightarrow \forall x \in B \vdash \exists y \in A \vdash (x, y) \in i_B \Rightarrow i_B$  is a function. Finally if  $(x, y) \in i_B \wedge (x', y) \in i_B \Rightarrow x = y = x' \Rightarrow x = x'$  so  $i_B$  is injective.  $\square$

**Example 2.28.**  $\emptyset: \emptyset \rightarrow \emptyset$  is a a bijection

**Proof.**

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

#### 2.1.4 Operations on functions and partial functions

**Theorem 2.29.** If  $f: A \rightarrow B$  is a partial function and  $C \subseteq A$  is a subclass of  $A$  then  $f|_C: C \rightarrow B$  with  $f|_C = \{(x, y) \in f | x \in C\}$  is a partial function with  $\text{dom}(f|_C) = \text{dom}(f) \cap C$ . Furthermore if  $f$  is a function then  $f|_C$  is a function

**Proof.** First as

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

For the domain note that

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

proving that

$$\text{dom}(f|_C) = \text{dom}(f) \cap C$$

Finally if  $f$  is a function then  $\text{dom}(f) = A$  so that  $\text{dom}(f|_C) = A \cap C \subseteq A$  proving that  $f|_C: C \rightarrow B$  is a function.  $\square$

**Theorem 2.30.** Let  $f: A \rightarrow B$  be a function and  $C \subseteq A$  and  $D \subseteq B$  then  $(f|_C)^{-1}(D) = C \cap f^{-1}(D)$

**Proof.** If  $x \in (f|_C)^{-1}(D)$  then [see 2.4]  $\exists y \in D$  such that  $(x, y) \in f|_C \subseteq C \times B$  so as  $f|_C \subseteq f$  we have  $x \in C \wedge (x, y) \in f$  proving that  $x \in f^{-1}(D) \cap C$ . If  $x \in C \cap f^{-1}(D)$  then  $x \in C$  and  $\exists y \in D$  such that  $(x, y) \in f$  so  $(x, y) \in f|_C$  proving that  $x \in (f|_C)^{-1}(D)$ .  $\square$

**Theorem 2.31.** If  $f: A \rightarrow B$  is a partial function,  $C, D \subseteq A$  then  $f|_{C \cap D} = (f|_C)|_D$ . Further if  $D \subseteq C$  then  $f|_D = (f|_C)|_D$

**Proof.** We have

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

proving that

$$f|_{C \cap D} = (f|_C)|_D \quad (2.5)$$

Further if  $D \subseteq C$  then  $C \cap D = D$  so that  $(f|_C)|_D = f|_{C \cap D} = f|_D$   $\square$

**Theorem 2.32.** If  $f: A \cup B \rightarrow C$  is a partial function then  $f = f|_A \cup f|_B$

**Proof.** We have

$$\begin{aligned} (x, y) \in f \subseteq (A \cup B) \times C &\Rightarrow (x \in A \vee x \in B) \wedge (x, y) \in f \\ &\Rightarrow (x \in A \wedge (x, y) \in f) \vee (x \in B \wedge (x, y) \in f) \\ &\Rightarrow (x, y) \in f|_A \vee (x, y) \in f|_B \\ &\Rightarrow (x, y) \in f|_A \cup f|_B \\ (x, y) \in f|_A \cup f|_B &\stackrel{f|_A, f|_B \subseteq f}{\Rightarrow} (x, y) \in f \bigcup f = f \end{aligned}$$

proving that

$$f = f|_A \cup f|_B \quad \square$$

**Theorem 2.33.** If  $f_1: B \rightarrow A$  and  $f_2: C \rightarrow A$  are partial functions with  $B \cap C = \emptyset$  then if  $f = f_1 \cup f_2$  the following hold

1.  $f: B \cup C \rightarrow A$  is a partial function and if  $f_1$  and  $f_2$  are functions then  $f: B \cup C \rightarrow A$  is a function
2.  $f_1 = f|_B$  and  $f_2 = f|_C$
3. If  $f_1, f_2$  are functions then if  $x \in B \Rightarrow f(x) = f_1(x)$  and if  $x \in C \Rightarrow f(x) = f_2(x)$

**Proof.** We begin the proof by proving

$$(x, y) \in f \wedge x \in B \Leftrightarrow (x, y) \in f_1 \quad (2.6)$$

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

Next we prove

$$(x, y) \in f \wedge x \in C \Leftrightarrow (x, y) \in f_2 \quad (2.7)$$

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

We proceed now as follows

1. If  $(x, y) \in f \subseteq B \times A \wedge (x, y') \in f \subseteq C \times A$  then  $x \in B \cup C$  and we have the following cases to consider

$$x \in B. \text{ then [see 2.6]} (x, y) \in f_1 \wedge (x, y') \in f_1 \xrightarrow{f_1 \text{ is a function}} y = y'$$

$$x \in C. \text{ then [see 2.7]} (x, y) \in f_2 \wedge (x, y') \in f_2 \xrightarrow{f_2 \text{ is a function}} y = y'$$

As we have also  $f = f_1 \cup f_2 \subseteq (B \times A) \cup (C \times A) \stackrel{1.43}{=} (B \cup C) \times A$  we have that

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

To prove that it is a function if  $f_1, f_2$  are functions, consider

$$\begin{aligned} x \in B \cup C &\Rightarrow x \in B \vee x \in C \\ &\xrightarrow{f_1, f_2 \text{ are functions}} (\exists y \in A \vdash (x, y) \in f_1) \vee (\exists y' \in A \vdash (x, y') \in f_2) \\ &\xrightarrow{f_1, f_2 \subseteq f_1 \cup f_2} (\exists y \in A \vdash (x, y) \in f_1 \cup f_2) \vee (\exists y' \in A \vdash (x, y') \in f_1 \cup f_2) \\ &\Rightarrow x \in \text{dom}(f_1 \cup f_2) \vee x \in \text{dom}(f_1 \cup f_2) \\ &\Rightarrow x \in \text{dom}(f_1 \cup f_2) \end{aligned}$$

and thus

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

2. We have

$$\begin{aligned} (x, y) \in f|_B &\Leftrightarrow (x, y) \in f \wedge x \in B \\ &\xrightarrow{2.6} (x, y) \in f_1 \\ (x, y) \in f|_C &\Leftrightarrow (x, y) \in f \wedge x \in C \\ &\xrightarrow{2.7} (x, y) \in f_2 \end{aligned}$$

3. Finally if  $f_1: B \rightarrow A, f_2: C \rightarrow A$  are functions then  $f: B \cup C \rightarrow A$  is a function [see 1] and for  $x \in B \cup C$  we have either

$$x \in B. \text{ Then as } \text{dom}(f) = B \cup C \text{ we have } \exists y \vdash (x, y) \in f \Rightarrow x \in B \wedge (x, y) \in f \xrightarrow{2.6} (x, y) \in f_1 \Rightarrow f(x) = y = f_1(x)$$

$$x \in C. \text{ Then as } \text{dom}(f) = B \cup C \text{ we have } \exists y \vdash (x, y) \in f \Rightarrow x \in C \wedge (x, y) \in f \xrightarrow{2.6} (x, y) \in f_2 \Rightarrow f(x) = y = f_2(x) \quad \square$$

**Notation 2.34.** If  $f_1: B \rightarrow A$  and  $f_2: C \rightarrow A$  are (partial) functions,  $B \cap C = \emptyset$  then  $f_1 \cup f_2: B \cup C \rightarrow A$  is noted as  $f: B \cup C \rightarrow A$  where  $f(x) = \begin{cases} f_1(x) & \text{if } x \in B \\ f_2(x) & \text{if } x \in C \end{cases}$ . This notation can easily be extended to a finite set of functions.

**Theorem 2.35.** If  $f: A \rightarrow B$  and  $g: C \rightarrow D$  are functions then  $g \circ f: A \cap f^{-1}(C) \rightarrow D$  is a function.

**Proof.** As  $f: A \rightarrow B$  and  $g: C \rightarrow D$  are functions we have by 2.5 that  $g \circ f: A \rightarrow D$  is a partial function. Using 2.7 we have that  $\text{dom}(g \circ f) = \text{dom}(f) \cap f^{-1}(\text{dom}(g)) \stackrel{\text{dom}(f) = A \wedge \text{dom}(g) = C}{=} A \cap f^{-1}(C)$  so that by 2.10 we have that  $g \circ f: A \cap f^{-1}(C) \rightarrow D$  is a function.  $\square$

**Theorem 2.36.** If  $f: A \rightarrow B$  and  $g: C \rightarrow D$  are functions with  $\text{range}(f) \subseteq C$  then  $g \circ f: A \rightarrow D$  is a function.

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

**Theorem 2.37.** *We have the following properties of the composition of partial functions*

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

**Proof.**

1.

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

2.

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

3.

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

$\square$

**Theorem 2.38.** *If  $f: A \rightarrow B$  and  $g: C \rightarrow D$  are functions with  $\forall x \in A \cap f^{-1}(C)$  we have  $(g \circ f)(x) = g(f(x))$*

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

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

$\square$

**Definition 2.39.** *A function  $f: A \rightarrow B$  is a **invertible** if  $f^{-1}: B \rightarrow A$  is a function.*

**Theorem 2.40.** *If  $f: A \rightarrow B$  is a function then  $f$  is invertible iff  $f$  is bijective*

**Proof.**

$\Rightarrow$ . Assume that  $f$  is invertible then it is invertible as a partial function and thus by 2.20 we have that  $f$  is injective. Now if  $x \in B \stackrel{f^{-1} \text{ is a function}}{\Rightarrow} \exists y \vdash (x, y) \in f^{-1} \Rightarrow (y, x) \in f$  proving that  $f$  is surjective. So we have that  $f$  is bijective.

$\Leftarrow$ . Assume that  $f$  is bijective then it is injective so that by 2.20  $f^{-1}$  is a partial function. Now if  $x \in B \stackrel{f \text{ is surjective}}{\Rightarrow} \exists y \vdash (y, x) \in f \Rightarrow (x, y) \in f^{-1} \Rightarrow x \in \text{dom}(f^{-1}) \Rightarrow B \subseteq \text{dom}(f^{-1}) \subseteq B$  so  $\text{dom}(f^{-1}) = B$  proving that  $f^{-1}$  is a function.  $\square$

**Theorem 2.41.** *If  $f: A \rightarrow B$  is a bijection then  $f^{-1}: B \rightarrow A$  is a bijection.*

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

**Theorem 2.42.** *If  $f: A \rightarrow B$  is bijective then*

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

**Proof.**

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

$$\begin{aligned}
 (x, y) \in f \circ f^{-1} &\Rightarrow \exists z \in B \vdash (x, z) \in f^{-1} \wedge (z, y) \in f \\
 &\Rightarrow \exists z \vdash (z, x) \in f \wedge (z, y) \in f \\
 &\xrightarrow{f \text{ is a function}} x = y \\
 &\Rightarrow (x, y) \in \mathbb{1}_B \\
 (x, y) \in \mathbb{1}_B &\Rightarrow x \in B \wedge x = y \\
 &\xrightarrow{f^{-1} \text{ is a function}} \exists z \vdash (x, z) \in f^{-1} \\
 &\Rightarrow \exists z \vdash (x, z) \in f^{-1} \wedge (z, x) \in f \\
 &\Rightarrow (x, x) \in f \circ f^{-1} \\
 &\xrightarrow{x=y} (x, y) \in f \circ f^{-1}
 \end{aligned}$$

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

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

$\square$

**Theorem 2.43.** *If  $f: A \rightarrow B$  and  $g: B \rightarrow A$  are two functions such that  $f \circ g = \mathbb{1}_B$  and  $g \circ f = \mathbb{1}_A$  then  $f$  is bijective (hence invertible) and  $g = f^{-1}$ .*

**Proof.** First we prove that  $f$  is bijective

1. (injectivity)

$$\begin{aligned}
 (x, y), (x', y) \in f \xrightarrow{y \in B \wedge g \text{ is a function}} \exists z \vdash (y, z) \in g \\
 &\Rightarrow (x, z), (x', z) \in g \circ f = \mathbb{1}_A \\
 &\Rightarrow x = z = x' \Rightarrow x = x'
 \end{aligned}$$

## 2. (surjectivity)

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

3. We have

$$\begin{aligned}
(x, y) \in g &\stackrel{g \subseteq B \times A}{\Rightarrow} y \in A \\
&\stackrel{\text{dom}(f) = A}{\Rightarrow} \exists z \vdash (y, z) \in f \\
&\Rightarrow (x, z) \in f \circ g = \mathbb{1}_B \\
&\Rightarrow x = z \\
&\Rightarrow (y, x) \in f \\
&\Rightarrow (x, y) \in f^{-1} \\
(x, y) \in f^{-1} &\Rightarrow (y, x) \in f \\
&\stackrel{f \subseteq A \times B}{\Rightarrow} x \in B \\
&\stackrel{\text{dom}(g) = B}{\Rightarrow} \exists z \vdash (x, z) \in g \\
&\Rightarrow (y, z) \in g \circ f = \mathbb{1}_A \\
&\Rightarrow y = z \\
&\Rightarrow (x, y) \in g
\end{aligned}$$

proving that  $g = f^{-1}$

□

We can summarize the two previous theorems in the following theorem

**Theorem 2.44.** *A function  $f: A \rightarrow B$  is invertible (or bijective) iff there exists a function  $g: B \rightarrow A$  such that  $f \circ g = \mathbb{1}_B$  and  $g \circ f = \mathbb{1}_A$ . If such a  $g$  exists then  $g = f^{-1}$ . Note that for  $g = f^{-1}$  we have  $g \circ f = \mathbb{1}_A \wedge f \circ g = \mathbb{1}_B$  so that  $(f^{-1})^{-1} = f$*

**Lemma 2.45.** *If  $f: A \rightarrow B$  is a function and  $\text{range}(f) \subseteq C$  then  $f: A \rightarrow C$  is a function*

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

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

□

**Theorem 2.46.** *If  $f: A \rightarrow B$  and  $g: C \rightarrow D$  are bijections and  $A \cap C = \emptyset = B \cap D$  then  $f \cup g: A \cup C \rightarrow B \cup D$  is a bijection.*

**Proof.** First from the two given functions we form the functions  $f: A \rightarrow B \cup D$  and  $g: C \rightarrow B \cup D$  [see 2.45]. Using 2.33 and  $A \cap C = \emptyset$  we have that  $f \cup g: A \cup C \rightarrow B \cup D$  is a function. We prove now that it is a bijection

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

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

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

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

**Theorem 2.47.** If  $f: A \rightarrow B$  is a function,  $A \neq \emptyset$  then  $f: A \rightarrow B$  is injective if and only if there exists a function  $g: B \rightarrow A$  such that  $g \circ f = \mathbb{1}_A$

**Proof.**

$\Rightarrow$ . Using 2.21 we have that  $f^{-1}: f(A) \rightarrow A$  is a function. Further we have that as  $A \neq \emptyset$  that there exists a  $a \in A$  which allows us to consider the function  $C_{B \setminus f(A), A, a}: B \setminus f(A) \rightarrow A$ . Then as  $B \cap (B \setminus f(A)) = \emptyset$  we can define the function  $g: B \rightarrow A$  where  $g = f^{-1} \bigcup C_{B \setminus f(A), A, a}$  [see 2.33]. Now we have if  $(x, y) \in g \circ f$  that  $\exists z \in B$  such that  $(x, z) \in f$  and  $(z, y) \in g$ . As  $(x, z) \in f$  we have that  $z \in f(A)$  so that by 2.33 (3) that  $(z, y) \in f^{-1} \Rightarrow (y, z) \in f \xrightarrow{(x, y \in f) \wedge f \text{ is injective}} x = y$  proving that  $(x, y) \in \mathbb{1}_A$ . Hence

$$g \circ f \subseteq \mathbb{1}_A$$

Further if  $(x, y) \in \mathbb{1}_A$  then  $x = y$  and, as  $x \in A = \text{dom}(f)$ , the existance of a  $z \in B$  such that  $(x, z) \in f$ . Hence  $(z, x) \in f^{-1} \subseteq g$  proving that  $(x, x) \in g \circ f \xrightarrow{x=y} (x, y) \in g \circ f$ . Hence

$$\mathbb{1}_A \subseteq g \circ f$$

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

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

$\square$

**Theorem 2.48.** A function  $f: A \rightarrow B$  is surjective if there exists a  $g: B \rightarrow A$  such that  $f \circ g = \mathbb{1}_B$ .

**Proof.**

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

proving surjectivity.  $\square$

**Theorem 2.49.** If  $f: A \rightarrow B$  and  $g: B \rightarrow C$  are partial functions with  $\text{range}(f) \subseteq \text{dom}(g)$  then we have for the function  $g \circ f: A \rightarrow C$

1. If  $f: A \rightarrow B$  and  $g: B \rightarrow C$  are injective then  $g \circ f: A \rightarrow C$  is injective.
2. If  $f: A \rightarrow B$  and  $g: B \rightarrow C$  are surjective then  $g \circ f: A \rightarrow C$  is surjective.
3. If  $f: A \rightarrow B$  and  $g: B \rightarrow C$  are bijective **functions** then  $g \circ f: A \rightarrow C$  is bijective. Or equivalent if  $f, g$  are invertible we have that  $g \circ f$  is invertible. Furthermore by 1.48 we have that  $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$ .

**Proof.**

1. Let  $(x, y), (x', y) \in g \circ f \Rightarrow \exists z, z' \vdash (x, z), (x', z') \in f, (z, y), (z', y) \in g \xrightarrow{g \text{ is injective}} z = z' \Rightarrow (x, z), (x', z) \in f \xrightarrow{f \text{ is injective}} x = x'$
2. If  $z \in C \xrightarrow{g \text{ is surjective}} \exists y \vdash (y, z) \in g \xrightarrow{g \subseteq B \times C} y \in B \xrightarrow{f \text{ is surjective}} \exists x \vdash (x, y) \in f \Rightarrow (x, z) \in g \circ f$  proving surjectivity.

3. This follows from (1) and (2)  $\square$

**Note 2.50.** Using the notation in 2.12 we have for  $f: A \rightarrow B$  that  $\forall x \in A (x, y) \in f \underset{\text{notation}}{=} y = f(x) \underset{\text{notation}}{=} (x, f(x)) \in f$ . Using this notation we have then that

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

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

this can easily extends to more then 2 partial functions.

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

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

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

## 2.1.5 Images and preimages of functions and partial functions

**Theorem 2.51.** Let  $f: A \rightarrow B, g: C \rightarrow D$  be partial functions and  $E \subseteq A$  then  $(g \circ f)|_E = (g|_{f(E)}) \circ (f|_E)$ . Further  $\text{dom}((g \circ f)|_E) = \text{dom}(f) \cap f^{-1}(\text{dom}(g)) \cap E$ . If  $f, g$  are functions [so that  $\text{dom}(f) = A$  and  $\text{dom}(g) = C$ ] then  $\text{dom}((g \circ f)|_E) = A \cap f^{-1}(C) \cap E \underset{f^{-1}(C) \subseteq}{=} f^{-1}(C) \cap E$  so that  $(g \circ f)|_E: f^{-1}(C) \cap E \rightarrow D$  is a function

**Proof.**

$$\begin{aligned} (x, y) \in (g \circ f)|_E &\Leftrightarrow (x, y) \in g \circ f \wedge x \in E \\ &\Leftrightarrow \exists z \in B \text{ such that } (x, z) \in f \wedge (z, y) \in g \wedge x \in E \\ &\Leftrightarrow \exists z \in B \text{ such that } z \in f(E) \wedge (x, z) \in f|_E \wedge (z, y) \in g \\ &\stackrel{2.29}{\Leftrightarrow} \exists z \in B \text{ such that } (x, z) \in f|_E \wedge (z, y) \in g|_{f(E)} \\ &\Leftrightarrow (z, y) \in g|_{f(E)} \circ f|_E \end{aligned}$$

proving that

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

For the domain note that  $\text{dom}((g \circ f)|_E) \underset{2.29}{=} \text{dom}(g \circ f) \cap E \underset{2.7}{=} \text{dom}(f) \cap f^{-1}(\text{dom}(g)) \cap E$ .  $\square$

**Theorem 2.52.** If  $f: A \rightarrow B$  is a injective function and  $C \subseteq A$  then  $f|_C: C \rightarrow f(C)$  is a bijection. Further we have that  $(f|_C)^{-1} = (f^{-1})|_{f(C)}$  so that the following bijections are equal  $(f|_C)^{-1}: f(C) \rightarrow C, (f^{-1})|_{f(C)}: f(C) \rightarrow C$

**Proof.** First we prove bijectivity

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

Next

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

which proves that

$$(f^{-1})|_{f(C)} = (f|_C)^{-1} \quad \square$$

**Theorem 2.53.** Let  $f: A \rightarrow B$  be a injective partial function [so that  $f^{-1}: B \rightarrow A$  is a partial function] then

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

**Note 2.54.** In the above theorem  $f^{-1}(C)$  is the preimage of  $C$  by  $f$  and  $(f^{-1})(C)$  is the image of  $C$  by  $f^{-1}$ .

**Proof.**

1. If  $y \in f^{-1}(C)$  then there exists a  $x \in C$  such that  $(y, x) \in f$  hence  $(x, y) \in f^{-1}$  whcih as  $x \in C$  proves that  $y \in (f^{-1})(C)$ . If  $y \in (f^{-1})(C)$  then  $\exists x \in C$  such that  $(x, y) \in f^{-1}$  hence  $(y, x) \in f$  which as  $x \in C$  means that  $y \in f^{-1}(C)$ .
2. If  $y \in (f^{-1})^{-1}(C)$  then  $\exists x \in C$  such that  $(y, x) \in f^{-1} \Rightarrow (x, y) \in f$  so that  $y \in f(C)$ . If  $y \in f(C)$  then there exists a  $x \in C$  such that  $(x, y) \in f \Rightarrow (y, x) \in f^{-1}$  hence  $y \in (f^{-1})^{-1}(C)$   $\square$

From now on we start using our alternative notation  $y = f(x) \Leftrightarrow (x, y) \in f \Leftrightarrow (x, f(x)) \in f$

**Theorem 2.55.** If  $f: A \rightarrow B$  is a bijection and  $C \subseteq A$  then  $f|_{(A \setminus C)}: A \setminus C \rightarrow B \setminus f(C)$  is a bijection

**Proof.**

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

**Theorem 2.56.** If  $f: A \rightarrow B$  is a function and  $C \subseteq B$  then  $f(f^{-1}(C)) \subseteq C$ . Further if  $f$  is surjective then  $f(f^{-1}(C)) = C$

**Proof.** If  $x \in f(f^{-1}(C)) \Rightarrow \exists y \in f^{-1}(C)$  so that  $x = f(y) \stackrel{y \in f^{-1}(C)}{\Rightarrow} f(y) \in C \Rightarrow x \in C$  proving that  $f(f^{-1}(C)) \subseteq C$ . Further if  $f$  is surjective then if  $y \in C \stackrel{f \text{ is surjective}}{\Rightarrow} \exists x \in A$  such that  $f(x) = y \in C \Rightarrow x \in f^{-1}(C) \Rightarrow y = f(x) \in f(f^{-1}(C))$   $\square$

**Theorem 2.57.** *If  $f: A \rightarrow B$  is a injective partial function and  $C \subseteq A$  then  $f^{-1}(f(C)) = C$*

**Proof.** If  $x \in f^{-1}(f(C)) \Rightarrow f(x) \in f(C) \Rightarrow \exists y \in C$  such that  $f(x) = f(y) \xrightarrow{f \text{ is injective}} x = y \Rightarrow x \in C$ . If  $x \in C \Rightarrow f(x) \in f(C) \Rightarrow x \in f^{-1}(f(C))$ .  $\square$

TODO

**Theorem 2.58.** *If  $f: A \rightarrow B$  is a function then we have*

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

**Proof.**

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

**Theorem 2.59.** *Let  $A, B, C$  be sets and  $f: A \rightarrow B$ ,  $g: B \rightarrow C$  functions then if  $D \subseteq C$  we have that  $(g \circ f)^{-1}(D) = f^{-1}(g^{-1}(D))$*

**Proof.**

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

$\square$

**Definition 2.60.** *If  $f: A \rightarrow B$  is a function and  $\{C_i\}_{i \in I}$  is a family so that  $\forall i \in I$  we have  $C_i \subseteq A$  [in other words  $\{C_i\}_{i \in I}$  is a family of sub-classes of  $A$ ]. Then we define  $\{f(C_i)\}_{i \in I}$  to be the family formed by  $\text{graph}(\{f(C_i)\}_{i \in I}) = \{(i, x) \mid i \in I \wedge x \in f(C_i)\}$ .*

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

**Theorem 2.61.** *If  $f: A \rightarrow B$  is a function and  $\{C_i\}_{i \in I}$  is a family of sub-classes of  $A$  and  $G = \langle \text{graph}(\{f(C_i)\}_{i \in I}), I \rangle$  then  $\forall i \in I$  we have that  $G_i = f(C_i)$*

**Proof.** If  $i \in I$  then if  $x \in G_i \Rightarrow (i, x) \in G \Rightarrow i \in I \wedge x \in f(C_i) \Rightarrow x \in f(C_i) \Rightarrow G_i \subseteq f(C_i)$ . If  $x \in f(C_i) \Rightarrow (i, x) \in G \Rightarrow x \in G_i \Rightarrow f(C_i) \subseteq G_i \Rightarrow f(C_i) = G_i$   $\square$

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

**Theorem 2.62.** If  $f: A \rightarrow B$  is a function, let  $\{C_i\}_{i \in I}$  be a non empty family [meaning  $I \neq \emptyset$ ] of sub-classes of  $A$  and let  $\{D_i\}_{i \in I}$  be a family of sub-classes of  $B$  then we have

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

**Proof.**

1.

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

2.

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

3.

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

proving

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

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

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

4.

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

□

**Theorem 2.63.** Let  $f: A \rightarrow B$  be injective then if  $X, Y \subseteq A$  we have  $f(X \setminus Y) = f(X) \setminus f(Y)$

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

$$f(X \setminus Y) \subseteq f(X) \setminus f(Y). \quad (2.9)$$

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

$$f(X) \setminus f(Y) \subseteq f(X \setminus Y). \quad (2.10)$$

Finally from (2.9) and (2.10) we have

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

**Theorem 2.64.** If  $f: A \rightarrow E$  and  $g: C \rightarrow E$  be injective functions [so that  $f: A \rightarrow f(A)$  and  $g: C \rightarrow g(C)$  are bijections (see 2.52)] then

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

**Proof.**

1. Using 2.7 we have that

$$\begin{aligned}
\text{dom}(f \circ g^{-1}) &= \text{dom}(g^{-1}) \bigcap (g^{-1})^{-1}(\text{dom}(f)) \\
&\stackrel{\text{dom}(f)=A}{=} \text{dom}(g^{-1}) \bigcap (g^{-1})^{-1}(A) \\
&\stackrel{(g^{-1}): g(C) \rightarrow C \text{ is bijective}}{=} g(C) \bigcap (g^{-1})^{-1}(A) \\
&\stackrel{2.53}{=} g(C) \bigcap g(A) \\
&\stackrel{2.62}{=} g(C \cap A)
\end{aligned}$$

2. We have

$$\begin{aligned}
f \circ g^{-1} &\stackrel{(1)}{=} (f \circ g^{-1})_{|g(A \cap C)} \\
&\stackrel{2.51}{=} f_{|g^{-1}(g(A \cap C))} \circ (g^{-1})_{g(A \cap C)} \\
&\stackrel{g \text{ is injective}}{=} f_{|A \cap C} \circ (g^{-1})_{|g(A \cap C)} \quad (2.11)
\end{aligned}$$

$$f_{|A \cap C} \circ (g_{|A \cap C})^{-1} \quad (2.12)$$

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

3. We have

$$\begin{aligned}
 (f \circ g^{-1})|_{g(D)} &\stackrel{(2)}{=} & (f|_{(A \cap C)} \circ (g^{-1})|_{g(A \cap C)})|_{g(D)} \\
 &\stackrel{2.51}{=} & (f|_{A \cap C})|_{(g^{-1})(g(D))} \circ ((g^{-1})|_{g(A \cap C)})|_{g(D)} \\
 &\stackrel{(g^{-1}):g(C) \rightarrow C \text{ is bijective}}{=} & (f|_{A \cap C})|_D \circ ((g^{-1})|_{g(A \cap C)})|_{g(D)} \\
 &\stackrel{2.31}{=} & f|_{(A \cap C) \cap D} \circ (g^{-1})|_{g(A \cap C) \cap g(D)} \\
 &\stackrel{D \subseteq A \cap C}{=} & f|_D \circ (g^{-1})|_{g(A \cap C) \cap D} \\
 &\stackrel{2.62}{=} & f|_D \circ (g^{-1})|_{g(D)} \\
 &\stackrel{D \subseteq A \cap C}{=} & f|_D \circ (g|_D)^{-1} \\
 &\stackrel{2.52}{=} & f|_D \circ (g|_D)^{-1}
 \end{aligned} \tag{2.13}$$

First note that by 2.52 and injectivity we have that  $g|_D: D \rightarrow g(D)$  is a bijection with  $(g|_D)^{-1}: g(D) \rightarrow D$  as its inverse bijection and that  $f|_D: D \rightarrow f(D)$  is a bijection. Using 2.49 we have then that the composition of these two functions  $(f|_D) \circ (g|_D)^{-1}: g(D) \rightarrow f(D)$  is a bijection where by (2.13)  $(f \circ g^{-1})|_{g(D)} = f|_D \circ (g|_D)^{-1}$   $\square$

**Theorem 2.65.** *If  $f: A \rightarrow B$  is a function and  $C \subseteq B$  then  ${}_C|f: f^{-1}(C) \rightarrow C$  defined by  ${}_C|f = f \cap (A \times C)$  is a function. Furthermore if  $f$  is injective, surjective or bijective then  ${}_C|f$  is also injective, surjective or bijective*

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

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

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

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

Finally if  $f$  is bijective then it is injective and surjective and we have just proved that then  ${}_C|f$  is injective and surjective and thus bijective.  $\square$

## 2.1.6 Constructing families of sets

**Theorem 2.66.** *If  $I, A$  are classes and  $f: I \rightarrow \mathcal{P}(A)$  a function. Define then the graph  $F = \{(i, x) \in I \times A | x \in f(i)\} \subseteq I \times A$  which has  $\text{dom}(F) \subseteq I$  and defines thus a family  $\{F_i\}_{i \in I}$  such that  $\forall i \in I$  that  $F_i = f(i)$ . So we can write  $\{F_i\}_{i \in I} = \{f(i)\}_{i \in I}$*

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

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

**Theorem 2.67.** *If  $\{F_i\}_{i \in I}$  is a family of classes then there is a graph  $F$  with  $\text{dom}(F) = I$ . Define then  $f \subseteq I \times \mathcal{P}(\bigcup_{i \in I} F_i)$  by  $f = \{(i, A) \in I \times \mathcal{P}(\bigcup_{i \in I} F_i) | A = F_i\}$  then  $f: I \rightarrow \mathcal{P}(\bigcup_{i \in I} F_i)$  is a function such that  $f(i) = F_i$ . In other words  $\{F_i\}_{i \in I} = \{f(i)\}_{i \in I} = \{f_i\}_{i \in I}$*

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

**Theorem 2.68.** *If  $\{F_i\}_{i \in I}$  is a family of classes defined and  $g: J \rightarrow I$  a surjection then by the above theorem we have the existence of a  $f: I \rightarrow \mathcal{P}(\bigcup_{i \in I} F_i)$  such that  $\{F_i\}_{i \in I} = \{f(i)\}_{i \in I}$ . We can form then  $f \circ g: J \rightarrow \mathcal{P}(\bigcup_{j \in J} F_i)$  which defines then  $\{(f \circ g)(j)\}_{j \in J} = \{f(g(j))\}_{j \in J}$  which is noted by  $\{F_{g(j)}\}_{j \in J}$  here  $F_{g(j)} \equiv f(g(j)) = (f \circ g)(j)$ . We have then that  $\bigcup_{i \in I} F_i = \bigcup_{j \in J} F_{g(j)}$*

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

### 2.1.7 Product, Union and Intersection of a family of sets

**Axiom 2.69. (Axiom of Replacement)** *If  $A$  is a set and  $f: A \rightarrow B$  is a surjective function then  $B$  is a set.*

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

**Lemma 2.71.** *If  $\{A_i\}_{i \in I}$  is a family of sets then  $\{A_i \mid i \in I\} \stackrel{\text{definition}}{=} \{x \mid \exists i \in I \vdash x = A_i\}$  (which as  $A_i$  is a set and thus a element is a valid definition of a class) then*

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

**Proof.**

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

$\square$

**Theorem 2.72.** *If  $\{A_i\}_{i \in I}$  is a family of sets with  $I$  a set then  $\bigcup_{i \in I} A_i$  is a set.*

**Proof.** Define  $\varphi: I \rightarrow \{A_i \mid i \in I\}$  by  $\varphi = \{(i, A_i) \mid i \in I\}$  then  $\varphi$  is a surjective function and thus by the axiom of replacement 2.69 we have that  $\{A_i \mid i \in I\}$  is a set. Using then 1.60 we have that  $\bigcup_{i \in I} A_i = \bigcup_{A \in \{A_i \mid i \in I\}} A$  is a set.  $\square$

**Definition 2.73.** *If  $A$  and  $B$  are sets then the class  $B^A = \{f \mid f \subseteq A \times B \wedge f: A \rightarrow B \text{ is a function}\}$ , as  $A \times B$  is a set (see 1.72) and thus by the axiom of subsets (see 1.58) we have that  $f$  is a set so  $B^A$  is indeed a valid class by the axiom of class construction (see 1.9)*

**Note 2.74.** Using the above definition we see that  $B^A$  is the set of function graphs of functions between  $A$  and  $B$  or in other words  $f \in B^A$  is equivalent with  $f: A \rightarrow B$  is function. In the future we will often not make a distinction between a function and its function graph. So we will use terms like the function  $f \in B^A$  when we actually mean the function graph  $f \in B^A$ . We will also say that a function  $f: A \rightarrow B$  has a certain property  $P$  when we actually mean that the function graph has a certain property  $P$ . Hence if we define a class (set if  $B^A$  is a set)  $\mathcal{A} = \{f \in B^A \mid f \text{ satisfies } P\}$  then we call this the class (set) of functions satisfying  $P$ .

**Theorem 2.75.** *Let  $A, B, C$  be classes such that  $B \subseteq C$  then  $B^A \subseteq C^A$*

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

**Theorem 2.76.** *If  $A$  and  $B$  are sets then  $B^A$  is a set*

**Proof.**

If  $f \in B^A$  then  $f \subseteq A \times B$  and thus  $f \in \mathcal{P}(A \times B)$  so that we have  $B^A \subseteq \mathcal{P}(A \times B)$ . Now by 1.72 we have that  $A \times B$  is a set, using the axiom of power sets (see 1.63) we have that  $\mathcal{P}(A \times B)$  is a set and finally by the axiom of subsets (see 1.58) we have that  $B^A$  is a set.  $\square$

**Theorem 2.77.** *If  $A$  is a set then there exists a bijection between  $2^A$  and  $\mathcal{P}(A)$*

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

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

so  $\gamma$  is a function, next we have to prove that it is bijective

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

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

and thus  $D = E$

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

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

and thus  $f = C_B = \gamma(B)$   $\square$

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

**Definition 2.78. (pre-tuple)** A pre-tuple  $\langle f, I \rangle$  is a pair of a class and a function graph [see 1.54 [so  $\forall (x, y), (x, y') \in f$  we have  $y = y'$ ] and  $\text{dom}(f) = I$ . Note that because of this definition we have that  $\forall i \in I$  there exists a unique  $y$  such that  $(i, y) \in f$  (just like with functions) we use also the same notation to specify this unique value as  $f(i)$ , essential saying  $y = f(i)$  is the same as saying  $(i, y) \in f$ .]

**Example 2.79.**  $\langle \emptyset, \emptyset \rangle$  is a pretuple for  $\forall (x, y), (x, y') \in \emptyset$  we have vacuously that  $y = y'$  and  $\text{dom}(\emptyset) = \emptyset$

**Theorem 2.80.** Let  $\langle f, I \rangle$  be a pretuple and let  $J \subseteq I$  then if we define  $f|_J = \{(i, x) \in f | i \in J\}$  then  $\langle f|_J, J \rangle$  is a pretuple. Note that if  $(i, x) \in f|_J \Rightarrow (i, x) \in f \Rightarrow f|_J(i) = f(i)$ .

**Proof.** If  $(i, x), (i, y) \in f|_J \Rightarrow (i, x), (i, y) \in f \Rightarrow x = y$  so  $f|_J$  is a function graph. Also if  $i \in \text{dom}(f|_J) \Rightarrow \exists y$  such that  $(i, y) \in f|_J \Rightarrow i \in J \Rightarrow \text{dom}(f|_J) \subseteq J$ . If  $i \in J \Rightarrow i \in I = \text{dom}(f) \Rightarrow \exists y$  such that  $(i, y) \in f \stackrel{i \in J}{\Rightarrow} (i, y) \in f|_J$ .  $\square$

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

**Definition 2.81.** Let  $\{A_i\}_{i \in I}$  be a family of sets where  $I$  is also a set (see 1.77) then we define  $\prod_{i \in I} A_i = \{f \mid f \subseteq I \times (\bigcup_{i \in I} A_i) \wedge \langle f, I \rangle \text{ is a pretuple} \wedge \forall i \in I \models f(i) \in A_i\}$ . Note that  $\prod_{i \in I} A_i$  is well defined as a class as by 2.72  $\bigcup_{i \in I} A_i$  is a set so that  $I \times (\bigcup_{i \in I} A_i)$  is a set which by the axiom of subsets (see 1.58) gives that  $f$  is a set. The elements of  $\prod_{i \in \{1, \dots, n\}} A_i$  are called tuples.

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

**Theorem 2.82.** Let  $\{A_i\}_{i \in I}$  be a family of sets where  $I$  is also a set (see 1.77) then  $\prod_{i \in I} A_i$  is a set

**Proof.**

If  $f \in \prod_{i \in I} A_i$  then  $f \subseteq I \times (\bigcup_{i \in I} A_i)$  so that  $f \in \mathcal{P}(I \times (\bigcup_{i \in I} A_i))$ , from this it follows that  $\prod_{i \in I} A_i \subseteq \mathcal{P}(A \times B)$ . Now  $I$  is a set, using 2.72 we have that  $\bigcup_{i \in I} A_i$  is a set and by 1.72 we have that  $I \times (A_i)$  is a set, next using the axiom of power sets (see 1.63) we have that  $\mathcal{P}(I \times (\bigcup_{i \in I} A_i))$  is a set and finally by the axiom of subsets (see 1.58) we have that  $\prod_{i \in I} A_i$  is a set.  $\square$

**Example 2.83.** If  $\{A_i\}_{i \in \emptyset}$  is the empty family of sets then  $\bigcup_{i \in \emptyset} A_i = \emptyset$  [if  $x \in \bigcup_{i \in \emptyset} A_i$  there exists a  $i \in I$  such that  $x \in A_i$  which as  $I = \emptyset$  is impossible], so in this cases  $\prod_{i \in \emptyset} A_i = \{\emptyset\}$  (it only contains the empty graph and  $\langle \emptyset, \emptyset \rangle$  is a pretuple and of course  $\forall i \in I \models \emptyset(i) \in A_i$  is satisfied vacuously).

**Theorem 2.84.** Let  $\{A_i\}_{i \in I}$  be a family of sets where  $I$  is also a set then if  $f: I \rightarrow B$  is a function ( $f$  is the graph of our function) such that  $\forall i \in I$  we have  $f(i) \in A_i$  then  $f \in \prod_{i \in I} A_i$

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

**Theorem 2.85.** Let  $\{A_i\}_{i \in I}$  be a family of set,  $\beta: I \rightarrow J$  a bijection then  $\lambda: \prod_{i \in I} A_i \rightarrow \prod_{i \in J} A_{\beta^{-1}(i)}$  defined by  $\lambda(x) = x \circ \beta^{-1}$  [which is well defined as  $\beta$  is a bijection] is a bijection

**Proof.** First as  $x, \beta^{-1}$  are function graphs and  $\text{range}(\beta^{-1}) = I$ ,  $\text{dom}(x) = I$  we have by 1.52 that  $\text{dom}(x \circ \beta^{-1}) = I$  so  $\langle f, I \rangle$  is a pretuple. As  $\forall i \in J$  we have that  $x(\beta^{-1}(i)) \in A_{\beta^{-1}(i)}$  we have proved that  $x \circ \beta^{-1} \in \prod_{i \in J} A_{\beta^{-1}(i)}$ . Further we have

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

**surjectivity.** let  $y \in \prod_{i \in J} A_{\beta^{-1}(i)}$  and define  $x = y \circ \beta$  then as  $y, \beta^{-1}$  are function graphs and  $\text{range}(\beta) = I$ ,  $\text{dom}(y) = J$  we have by 1.52 that  $\text{dom}(y \circ \beta) = I$ . As  $\forall i \in I$  we have  $y(\beta(i)) \in A_{\beta^{-1}(\beta(i))} = A_i$  we have that  $y \circ \beta \in \prod_{i \in I} A_i$ . Further we have  $\lambda(x) = (y \circ \beta) \circ \beta^{-1} = y$  which proves that  $\lambda$  is surjective.

So  $\lambda$  is a bijection.  $\square$

**Theorem 2.86.** Let  $\{A_i\}_{i \in I}$  be a family of sets where  $\forall i \in I$  we have  $A_i = A$  then  $\prod_{i \in I} A_i = A^I$

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

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

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

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

So using the definition of a partial function (see 2.3) and a function (see 2.10 we have that  $\langle f, I, A \rangle$  is a function which means that by definition of  $A^I$  (see 2.73) we have  $f \in A^I$ . This proves that  $A \prod_{i \in I} A_i \subseteq A^I$ . Next if  $f \in A^I$  then  $f \subseteq I \times A = I \times \bigcup_{i \in I} A_i$  and  $\langle f, I, A \rangle$  is a function, as also  $\forall i \in I \models f(i) \in A$  [as  $\text{range}(f) \subseteq A$ ] we have by 2.84 that  $f \in \prod_{i \in I} A_i$  proving that  $A^I \subseteq \prod_{i \in I} A_i$ . So finally we conclude that  $A^I = \prod_{i \in I} A_i$ .  $\square$

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

**Theorem 2.87.** *Let  $I$  be a set,  $\{A_i\}_{i \in I}$  be a family of sets then we have  $f \in \prod_{i \in I} A_i \Leftrightarrow \langle f, I \rangle$  is a pretuple  $\wedge \forall i \in I \models f(i) \in A_i$*

**Proof.**

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

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

The following notation will be used a lot

**Notation 2.88.** *Let  $I$  be a set,  $\{A_i\}_{i \in I}$  be a family of sets and  $x \in \prod_{i \in I} A_i$  then given  $i \in I$  we note  $x(i)$  by  $x_i$  so  $x_i = x(i)$  and  $x_i = y \Leftrightarrow (i, y) \in x$ . Using the previous theorem we have then  $x \in \prod_{i \in I} A_i \Leftrightarrow \langle x, I \rangle$  is a pretuple and  $\forall i \in I$  we have  $x_i \in A_i$ .*

*If  $n \in \mathbb{N}_0$  (see 4.7) then if  $I = \{1, \dots, n\} / \{1, \dots, n\} = \{x \in \mathbb{N}_0 \mid 1 \leq x \wedge x \leq n\}$  we use the notation  $x = (x_1, \dots, x_n)$  to mean that  $\langle x, I \rangle$  is a pretuple and  $\forall i \in \{1, \dots, n\}$  we have  $x(i) = x_i$ , using this we have then the equivalence  $(x_1, \dots, x_n) \in \prod_{i \in \{1, \dots, n\}} A_i \Leftrightarrow \forall i \in \{1, \dots, n\} \models x_i \in A_i$ , corresponding with the typical mathematical convention for finite products of sets.*

**Theorem 2.89.** *If  $I$  is a set and  $\{B_i\}_{i \in I}$ ,  $\{A_i\}_{i \in I}$  families of sets such that  $\forall i \in I \models B_i \subseteq A_i$  then we have  $\prod_{i \in I} B_i \subseteq \prod_{i \in I} A_i$ .*

**Proof.** If  $x \in \prod_{i \in I} B_i$  then using 2.87 we have  $\langle x, I \rangle$  is a pretuple and  $\forall i \in I$  we have  $x(i) \in A_i \subseteq B_i$  so that  $\langle x, I \rangle$  is a pretuple and  $\forall i \in I$  we have  $x(i) \in A_i$  proving by 2.87 that  $x \in \prod_{i \in I} B_i$ .  $\square$

**Lemma 2.90.** *Let  $I$  be a set and  $\{A_i\}_{i \in I}$  a family of sets then  $\prod_{i \in I} A_i = \emptyset$  if and only if  $\exists i \in I$  with  $A_i = \emptyset$*

**Proof.**

$\Rightarrow$ . Let  $\prod_{i \in I} A_i = \emptyset$ , assume now that  $\forall i \in I$  we have  $A_i \neq \emptyset$  then using the Axiom of Choice (see 2.210) there exists a function  $f: I \rightarrow \bigcup_{i \in I} A_i$  such that  $\forall i \in I$  we have  $f(i) \in A_i$ , so  $\langle f, I \rangle$  is a pretuple and as  $\forall i \in I$  we have  $f(i) \in A_i$  we conclude by 2.87 that  $f \in \prod_{i \in I} A_i \Rightarrow f \in \emptyset$  a contradiction. So we must have that  $\exists i \in I$  such that  $A_i = \emptyset$ .

$\Leftarrow$ . Let  $\exists i \in I$  such that  $A_i = \emptyset$  then if  $f \in \prod_{i \in I} A_i$  we must have that  $f(i) \in A_i = \emptyset$  a contradiction. So we must have that  $\prod_{i \in I} A_i = \emptyset$   $\square$

**Definition 2.91.** *Let  $\{A_i\}_{i \in I}$  be a family of sets where  $I$  is a set then  $\forall i \in I$  we define  $\pi_i: \prod_{i \in I} A_i \rightarrow A_i$  by  $\pi_i = \{(x, x(i)) \mid x \in \prod_{i \in I} A_i\}$ . We have then that  $\pi_i$  is a function called the projection function. Given  $x \in \prod_{i \in I} A_i$  we have  $y = \pi_i(x) \Leftrightarrow (x, y) \in \pi_i \xrightarrow{x \in \prod_{i \in I} A_i} y = x(i)$  so that we have in a shorter notation that  $\pi_i(x) = x(i) = x_i$ .*

**Proof.** We still have to prove that  $\pi_i$  is a function, for this we use 2.3

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

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

**Theorem 2.92.** Let  $I$  be a set and  $\{A_i\}_{i \in I}$ ,  $\{B_i\}_{i \in I}$  two families of sets such that  $\forall x \in \prod_{i \in I} A_i$  we have that  $\forall i \in I$  we have that  $\pi_i(x) \in B_i$  then  $\prod_{i \in I} A_i \subseteq \prod_{i \in I} B_i$

**Proof.** Let  $x \in \prod_{i \in I} A_i$  then  $x$  is a pre-tuple, further as  $\forall i \in I$  we have that  $x(i) = \pi_i(x) \in B_i$  so that by 2.84 we have that  $x \in \prod_{i \in I} B_i$   $\square$

**Theorem 2.93.** Let  $\{A_i\}_{i \in I}$  be a family of **non empty** sets where  $I$  is a set then  $\forall i \in I$   $\pi_i: \prod_{i \in I} A_i \rightarrow A_i$  is a surjection.

**Proof.** Let  $i \in I$  and  $x \in A_i$  then as  $\forall j \in I \setminus \{i\}$  we have  $A_j \neq \emptyset$  we can use the Axiom of Choice (see 2.210) to find a function  $f: I \setminus \{i\} \rightarrow \bigcup_{j \in I \setminus \{i\}} A_j$  such that  $\forall j \in I \setminus \{i\}$  we have  $f(j) \in A_j$ . Using 2.33 we have then that for  $g = f \cup \{(i, x)\}$  we have that  $g: I \rightarrow \bigcup_{i \in I} A_i$  is a function so  $\langle g, I \rangle$  is a pretuple and as  $\forall j \in I$  we have  $g(j) \in A_j$  we conclude by 2.87 that  $g \in \prod_{j \in \{1, \dots, n\}} A_j$ . Finally as  $g(i) = x$  we have that  $(g, x) \in \pi_i$  or  $\pi_i(g) = x$  hence  $\pi_i$  is surjective.  $\square$

**Note 2.94.** That the condition that the family of sets is empty is necessary for if  $A_1 = \emptyset$  and  $A_2 = \{1\}$  then  $\prod_{i \in \{1, 2\}} A_i = \emptyset$  and for  $1 \in A_2$  we can not find a  $x \in \prod_{i \in \{1, 2\}} A_i$  such that  $\pi_2(x) = 1$

**Theorem 2.95.** Let  $I$  be a set and  $\{A_i\}_{i \in I}$ ,  $\{B_i\}_{i \in I}$  families of sets with  $\forall i \in I$  we have  $A_i \neq \emptyset$  (which by 2.90 is equivalent with  $\prod_{i \in I} A_i \neq \emptyset$ ) then  $\prod_{i \in I} A_i \subseteq \prod_{i \in I} B_i \Leftrightarrow \forall i \in I$  we have  $A_i \subseteq B_i$

**Proof.**

$\Rightarrow$ . Assume that  $\exists i \in I$  with  $A_i \not\subseteq B_i$  then there exists a  $x \in A_i$  with  $x \notin B_i$ . Now by the Axiom of Choice (see 2.210) there exists a function  $f': I \rightarrow \bigcup_{i \in I} A_i$  such that  $\forall i \in I \models f'(i) \in A_i$ , define then  $f: I \rightarrow \bigcup_{i \in I} A_i$  by  $f(j) = \begin{cases} f'(j) & \text{if } j \in I \setminus \{i\} \\ x & \text{if } j = i \end{cases}$  which is a function (see 2.33) (so  $\langle f, I \rangle$  is a pretuple) and as  $\forall j \in I$  we have  $f(j) \in A_j$  we have by 2.87 that  $f \in \prod_{i \in I} A_i$  but as  $f(i) = x \notin B_i$  we have that  $\prod_{j \in I} A_j \not\subseteq \prod_{j \in I} B_j$  a contradiction. So  $\forall i \in I$  we must have  $A_i \subseteq B_i$

$\Leftarrow$ . This follows from 2.89  $\square$

**Note 2.96.** The condition that  $\forall i \in I$  we have  $B_i \neq \emptyset$  is essential in the above theorem. For example if  $B_1 = \emptyset$ ,  $B_2 = \{1, 2\}$  and  $A_1 = \{1\}$ ,  $A_2 = \{2\}$  then clearly  $B_1 \subseteq A_1$  but  $B_2 \not\subseteq A_2$ , still we have  $\prod_{i \in \{1, 2\}} B_i = \emptyset \subseteq \prod_{i \in \{1, 2\}} A_i$ .

**Corollary 2.97.** Let  $I$  be a set then and  $\{A_i\}_{i \in I}$ ,  $\{B_i\}_{i \in I}$  be families of sets such that  $\forall i \in I$  we have  $A_i \neq \emptyset$  then  $\prod_{i \in I} A_i = \prod_{i \in I} B_i$  if and only if  $\forall i \in I$  we have  $A_i = B_i$

**Proof.** First as  $\forall i \in I$  we have  $A_i \neq \emptyset$  so that by 2.90  $\prod_{i \in I} A_i \neq \emptyset$  and thus by 2.90 we have that  $\forall i \in I$  we have  $B_i \neq \emptyset$ .

$\Rightarrow$ . If  $\prod_{i \in I} A_i = \prod_{i \in I} B_i$  then  $\prod_{i \in I} A_i \subseteq \prod_{i \in I} B_i$  and  $\prod_{i \in I} B_i \subseteq \prod_{i \in I} A_i \stackrel{2.95}{\Rightarrow} \forall i \in I$  we have  $A_i \subseteq B_i \wedge B_i \subseteq A_i \Rightarrow A_i = B_i$

$\Leftarrow$ . If  $\forall i \in I$  we have  $A_i = B_i \Rightarrow A_i \subseteq B_i \wedge B_i \subseteq A_i \stackrel{2.89}{\Rightarrow} \prod_{i \in I} A_i \subseteq \prod_{i \in I} B_i \wedge \prod_{i \in I} B_i \subseteq \prod_{i \in I} A_i$  proving that  $\prod_{i \in I} A_i = \prod_{i \in I} B_i$   $\square$

**Note 2.98.** The condition that  $\forall i \in I$  we have  $B_i \neq \emptyset$  is essential in the above theorem. For example if  $B_1 = \emptyset$ ,  $B_2 = \{1, 2\}$  and  $A_1 = \{1\}$ ,  $A_2 = \emptyset$  then clearly  $A_1 \neq B_1 \wedge A_2 \neq B_2$  and still we have  $\prod_{i \in \{1, 2\}} B_i = \emptyset = \prod_{i \in \{1, 2\}} A_i$ .

**Theorem 2.99.** If  $I$  is a set and  $\{A_i\}_{i \in I}$ ,  $\{B_i\}_{i \in I}$  families of sets then  $(\prod_{i \in I} A_i) \cap (\prod_{i \in I} B_i) = \prod_{i \in I} (A_i \cap B_i)$ .

**Proof.**

1. If  $x \in (\prod_{i \in I} A_i) \cap (\prod_{i \in I} B_i)$  then using 2.87  $\langle x, I \rangle$  is a pretuple and  $\forall i \in I$  we have  $x(i) \in A_i \wedge x(i) \in B_i \Rightarrow x(i) \in A_i \cap B_i$  so using 2.87 we have  $x \in \prod_{i \in I} (A_i \cap B_i)$ , proving that

$$\left( \prod_{i \in I} A_i \right) \cap \left( \prod_{i \in I} B_i \right) \subseteq \prod_{i \in I} (A_i \cap B_i) \quad (2.14)$$

2. If  $x \in \prod_{i \in I} (A_i \cap B_i)$  then  $\langle x, I \rangle$  is a pretuple and  $\forall i \in I$  we have  $x(i) \in A_i \cap B_i \Rightarrow x(i) \in A_i \wedge x(i) \in B_i$  proving by 2.87 that  $x \in \prod_{i \in I} A_i \wedge x \in \prod_{i \in I} B_i \Rightarrow x \in (\prod_{i \in I} A_i) \cap (\prod_{i \in I} B_i)$  proving that

$$\prod_{i \in I} (A_i \cap B_i) \subseteq \left( \prod_{i \in I} A_i \right) \cap \left( \prod_{i \in I} B_i \right) \quad (2.15)$$

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

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

**Theorem 2.100.** Let  $I, J$  be sets and  $\{\{A_{i,j}\}_{j \in J}\}_{i \in I}$  be a family of families of sets then  $\prod_{i \in I} (\prod_{j \in J} A_{i,j}) = \prod_{j \in J} (\prod_{i \in I} A_{i,j})$

**Proof.** First if  $x \in \prod_{i \in I} (\prod_{j \in J} A_{i,j})$  then  $\forall i \in I$  we have  $x \in \prod_{j \in J} A_{i,j}$  2.87  $\Rightarrow \langle x, J \rangle$  is a pretuple and  $\forall j \in J$  we have  $x(j) \in A_{i,j} \Rightarrow \langle x, J \rangle$  is a pretuple and  $\forall i \in I, \forall j \in J \models x(j) \in A_{i,j} \Rightarrow \forall j \in J \models x(j) \in \prod_{i \in I} A_{i,j}$  2.87  $\Rightarrow x \in \prod_{j \in J} (\prod_{i \in I} A_{i,j})$  proving that

$$\prod_{i \in I} \left( \prod_{j \in J} A_{i,j} \right) \subseteq \prod_{j \in J} \left( \prod_{i \in I} A_{i,j} \right) \quad (2.16)$$

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

$$\prod_{j \in J} \left( \prod_{i \in I} A_{i,j} \right) \subseteq \prod_{i \in I} \left( \prod_{j \in J} A_{i,j} \right) \quad (2.17)$$

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

**Theorem 2.101.** If  $I$  is a set and  $\{A_i\}_{i \in I}$ ,  $\{B_i\}_{i \in I}$  families of sets then  $(\prod_{i \in I} A_i) \cup (\prod_{i \in I} B_i) \subseteq \prod_{i \in I} (A_i \cup B_i)$

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

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

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

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

**Example 2.102.** If  $f \in \prod_{i \in \{0,1\}} (A_i \cup B_i)$  is such that  $f(0) \in A_0$  and  $f(1) \in B_1$  then if  $A_0 \cap B_1 = \emptyset$  we can not have  $f \in \prod_{i \in \{0,1\}} A_i$  (because  $f(1) \notin A_1$ ) and we can not have  $f \in \prod_{i \in \{0,1\}} B_i$  because  $f(0) \in A$

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

**Theorem 2.103.** Let  $I$  be a set  $\{A_i\}_{i \in I}$  a family of sets then if  $K, L \subseteq I$  such that  $I = K \cup L$  and  $\emptyset = K \cap L$  then  $f|_K \in \prod_{i \in K} A_i$ ,  $f|_L \in \prod_{i \in L} A_i$  and  $\beta: \prod_{i \in I} A_i \rightarrow (\prod_{i \in K} A_i) \times (\prod_{i \in L} A_i)$  defined by  $\beta(f) = (f|_K, f|_L)$  is a bijection.

**Proof.** First note that for  $f \in \prod_{i \in I} A_i$  we have that  $\langle I, f \rangle$  is a pretuple and  $\forall i \in I f(i) \in A_i$ . Using 2.80 it follows that  $\langle K, f|_K \rangle$  and  $\langle L, f|_L \rangle$  are pre-tuples and as we have trivially that  $\forall i \in K f|_K(i) = f(i) \in A_i$  and  $\forall i \in L f|_L(i) = f(i) \in A_i$  we conclude that  $f|_K \in \prod_{i \in K} A_i \wedge f|_L \in \prod_{i \in L} A_i$ . Hence  $\beta$  is indeed a function between  $\prod_{i \in I} A_i$  and  $(\prod_{i \in K} A_i) \times (\prod_{i \in L} A_i)$ . To prove that  $\beta$  is a bijection consider:

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

$i \in K$ . then  $f(i) = f|_K(i) = g|_K(i) = g(i)$

$i \in L$ . then  $f(i) = f|_L(i) = g|_L(i) = g(i)$

proving that  $f = g$

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

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

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

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

## 2.1.8 Family of elements in a set

**Definition 2.104.** If  $I, A$  are sets and  $x \in A^I$  then we note  $x$  as  $\{x_i\}_{i \in I}$  where  $x_i = x(i)$  and call  $\{x_i\}_{i \in I}$  a family of elements of  $A$  (or family in  $A$  in short). We say that the family is non empty if  $I \neq \emptyset$ .

**Example 2.105.** If  $A$  is a set then  $1_A: A \rightarrow A \in A^A$  defines a family of elements in  $A$  which we note by  $\{x\}_{x \in A}$  (so  $\{x\}_{x \in A} = \{(1_A)_i\}_{i \in A}$ ) and call the family obtained by self indexing  $A$ .

**Definition 2.106.** If  $\{x_i\}_{i \in I}$  is a family of elements of  $A$  (so  $\{x_i\}_{i \in I} \in A^I$ ) and  $J \subseteq I$  then  $\{x_i\}_{i \in J}$  is defined by  $(\{x_i\}_{i \in I})|_J \in B^J$ .

**Notation 2.107.** If  $A$  is a set and  $x = \{x_i\}_{i \in I}$  is a family of elements in  $A$ ,  $J \subseteq I$  and  $\sigma: J \rightarrow J$  a bijection then as  $x$  is a function  $x: I \rightarrow A$  we can consider the function  $x \circ \sigma: J \rightarrow A$  which represent again a family of elements in  $A$  which will be noted as  $\{x_{\sigma(i)}\}_{i \in J}$ . Here  $\forall i \in J$  we have  $x_{\sigma(i)} \equiv (x \circ \sigma)(i) = x(\sigma(i)) \equiv x_{\sigma(i)} = (x \circ \sigma)_i$

**Definition 2.108.** If  $I, A$  are sets and  $\{x_i\}_{i \in I}$  a family of elements in  $A$  then  $\{x_i: i \in I\}$  is defined to be equal to  $\{x_i\}_{i \in I}(I) = \{y \in A \mid \exists i \in I \vdash y = \{x_i\}_{i \in I}(i)\} \subseteq A$  (so that  $\{x_i: i \in I\}$  must be a set).

**Example 2.109.** If  $A$  is a set then  $\{x\}_{x \in A}(I) = A$  and this is consistent with the notation  $A = \{x | x \in A\}$

**Theorem 2.110.** If  $A, I$  are sets and  $\sigma: I \rightarrow I$  a bijection (this is called a permutation) then if  $\{x_i\}_{i \in I}$  is a family of elements in  $A$  then for the family  $\{x_{\sigma_i}\}_{i \in I} (= \{x_i\}_{i \in I} \circ \sigma)$  we have  $\{x_i | i \in I\} = \{x_{\sigma_i} | i \in I\}$

**Proof.**

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

□

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

1. If  $A$  is a set then  $\{A_i\}_{i \in I}$  is a family in  $\mathcal{P}(A)$  so  $\{A_i\}_{i \in I}$  is a function from  $I$  to  $\mathcal{P}(A)$
2. If  $A$  is a set then  $\{A_i\}_{i \in I}$  is a family of sets (see 1.77) such that  $\forall i \in I \models A_i \subseteq A$  and thus a graph with  $\text{domain}(\text{graph}(\{A_i\}_{i \in I})) \subseteq I$

However 2.66 and 2.67 shows that the two definitions (1),(2) are essentially the same if it's about subset's of a certain set.

**Definition 2.111.** Let  $I, A$  be sets,  $B \subseteq A$  and  $\{x_i\}_{i \in I}$  a family of elements in  $A$  then we say that  $\{x_i\}_{i \in I} \subseteq B$  if  $\forall i \in I$  we have  $\{x_i | i \in I\} \subseteq B$  (or  $\{x_i\}_{i \in I}(I) \subseteq B$ )

## 2.2 Relations

### 2.2.1 Relations

**Definition 2.112.** Let  $A$  be a class then a relation in  $A$  is any subclass of  $A \times A$

**Definition 2.113.** If  $A$  is a class then a relation  $R$  in  $A$  is

1. (reflexive) iff

$$\forall x \in A \vdash (x, x) \in R$$

2. (symmetric) iff

$$(x, y) \in R \Rightarrow (y, x) \in R$$

3. (anti-symmetric) iff

$$(x, y) \in R \wedge (y, x) \in R \Rightarrow x = y$$

4. (transitive) iff

$$(x, y) \in R \wedge (y, z) \in R \Rightarrow (x, z) \in R$$

**Notation 2.114.** Let  $A$  be a class and  $R$  is a relation in  $A$  then we note  $(x, y) \in R$  as  $xRy$

**Definition 2.115.** If  $A$  is a class then  $I_A = \{(x, x) | x \in A\}$  is the diagonal graph.

### 2.2.2 Equivalence relations

**Definition 2.116.** If  $A$  is a class then a relation in  $A$  is a equivalence relation if it is reflexive, symmetric and transitive.

**Definition 2.117.** If  $A$  is a set then a partition of  $A$  is a family of nonempty subsets of  $A$   $\{A_i\}_{i \in I}$  with the following properties

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

**Theorem 2.118.** If  $A$  is a set and  $\{A_i\}_{i \in I}$  is a family of non empty subsets of  $A$  then  $\{A_i\}_{i \in I}$  is a partition of  $A$  iff

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

**Proof.**

$\Rightarrow$

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

$\Leftarrow$

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

**Definition 2.119.** If  $A$  is a set and  $R$  a equivalence relation in  $A$  then given a  $x \in A$  the equivalence class  $R[x]$  of  $x$  modulo  $R$  is

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

(note that because  $R[x] \subseteq A$  and  $A$  is a set we have by the axiom of subset 1.58 that  $R[x]$  is a set)

**Theorem 2.120.** If  $R$  is a equivalence relation in a set  $A$  then

$$xRy \Leftrightarrow R[x] = R[y]$$

**Proof.**

$\Rightarrow$

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

$\Leftarrow$

If  $R[x] = R[y]$  then as  $xRx \Rightarrow x \in R[x] = R[y] \Rightarrow xRy$   $\square$

**Theorem 2.121.** Let  $A$  be a set and  $R$  a equivalence relation in  $A$  then  $\{R[x]\}_{x \in A}$  is a partition of  $A$ .  $\{R[x]\}_{x \in A}$  is the partition corresponding with  $R$

**Proof.** First for every  $x \in A$  we have  $xRx$  so that  $x \in R[x]$  and thus  $\{R[x]\}_{x \in A}$  is a family of non empty subsets of  $A$ . We use 2.118 to prove that  $\{R[x]\}_{x \in A}$  is a partition

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

The above theorem has also a opposite

**Theorem 2.122.** *If  $\{A_i\}_{i \in I}$  is a partition of  $A$  then define  $R = \{(x, y) \mid \exists i \in I \vdash x \in A_i \wedge y \in A_i\}$  then  $R$  is a equivalence relation,  $\forall i \in I \vdash \exists x \in A \vdash A_i = R[x]$  and  $\forall x \in A \vdash \exists i \in I \vdash R[x] = A_i$ . We call  $R$  to be the equivalence relation corresponding with the partition  $\{A_i\}_{i \in I}$*

**Proof.** First we prove that  $R$  is a equivalence relation

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

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

$$\begin{aligned} z \in R[x] &\Leftrightarrow zRx \\ &\Leftrightarrow \exists j \in I \vdash z \in A_j \wedge x \in A_j \\ &\stackrel{j=i \text{ as } x \in A_i \cap A_j}{\Leftrightarrow} z \in A_i \wedge x \in A_i \\ &\stackrel{x \in A_i}{\Leftrightarrow} z \in A_i \end{aligned}$$

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

**Definition 2.123.** *If  $A$  is set and  $R$  is a equivalence relation on  $A$  then the set  $A/R$  is defined to be  $\{R[x] \mid x \in A\}$  (note that as  $R[x] \subseteq A$  we have that  $A/R \subseteq \mathcal{P}(A)$  which because of axiom of subsets 1.58 and 1.63 that  $A/R$  is a set)*

**Theorem 2.124.** *If  $f: A \rightarrow B$  is a function and  $R$  is a equivalence relation on  $B$  then  $f\langle R \rangle = \{(x, y) \mid f(x)Rf(y)\}$  is a equivalence relation.  $f\langle R \rangle$  is called the pre-image of  $R$  by  $f$*

**Proof.**

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

**Theorem 2.125.** *If  $A$  is a class with a equivalence relation  $R$  then if  $B \subseteq A$  then  $R|_B = \{(x, y) \mid x \in B \wedge y \in B \wedge (x, y) \in R\}$  is a equivalence relation.*

**Proof.**

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

**Definition 2.126.** *Let  $R, R'$  be equivalence relations on  $A$  then  $R$  is a refinement of  $R'$  if  $R \subseteq R'$  (we express this also by saying that  $R$  is finer then  $R'$  and that  $R'$  is coarser then  $R$ )*

**Theorem 2.127.** *If  $R, R'$  are equivalence relations on  $A$  with  $R \subseteq R'$  then we have  $z \in R'[x] \Rightarrow R[z] \subseteq R'[x]$*

**Proof.** If  $z \in R'[x]$  then if  $y \in R[z] \Rightarrow yRz \Rightarrow yR'z \xrightarrow{z \in R'[x]} yR'z \wedge zR'x \Rightarrow yR'x \Rightarrow y \in R'[x]$   $\square$

**Corollary 2.128.** *If  $R, R'$  are equivalence relations on  $A$  with  $R \subseteq R'$  then  $\forall x \in A \vdash R[x] \subseteq R'[x]$*

**Proof.** If  $x \in A$  then as  $x \in R'[x] \xrightarrow{\text{previous theorem}} R[x] \subseteq R'[x]$   $\square$

**Corollary 2.129.** If  $P, R$  are equivalence relations on a set  $A$  with  $P \subseteq R$ . Then the quotient of  $\frac{R}{P}$  is a relation in  $A/P$  defined as follows

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

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

**Proof.** This is proved by

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

### 2.2.3 Equivalence relations and functions

**Theorem 2.130.** If  $f: A \rightarrow B$  is a function then  $R_f \subseteq A \times A$  defined by  $R_f = \{(x, y) \in A \times A | f(x) = f(y)\}$  is a equivalence relation in  $A$ , it is called the equivalence relation determined by  $f$

**Proof.**

1. **(reflexivity)** If  $x \in A \Rightarrow f(x) = f(x) \Rightarrow (x, x) \in R_f$
2. **(symmetry)** If  $x R_f y \Rightarrow f(x) = f(y) \Rightarrow f(y) = f(x) \Rightarrow y R_f x$
3. **(transitive)** If  $x R_f y$  and  $y R_f z$  then  $f(x) = f(y) \wedge f(y) = f(z) \Rightarrow f(x) = f(z) \Rightarrow x R_f z$   $\square$

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

**Theorem 2.131.** If  $R$  is a equivalence relation in a set  $A$  then  $f_R: A \rightarrow A/R$  defined by  $f_R = \{(x, R[x]) | x \in A\}$  is a function and we have then  $R_{f_R} = R$ . Furthermore we have that  $f_R$  is surjective. We call  $f_R$  the canonical function of  $R$

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

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

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

$$\begin{aligned} (x, y) \in R &\stackrel{2.120}{\Leftrightarrow} R[x] = R[y] \\ &\Leftrightarrow f_R(x) = f_R(y) \\ &\Leftrightarrow (x, y) \in R_{f_R} \end{aligned}$$

proving that  $R = R_{f_R}$ .

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

**Theorem 2.132.** If  $f: A \rightarrow B$  is a function where  $A, B$  are sets, define then  $s_f: A/R_f \rightarrow f(A)$  with  $s_f = \{(R_f[x], f(x)) | x \in A\}$  then we have  $f = e_{f(A)} \circ s_f \circ f_{R_f}$  where

1.  $f_{R_f}: A \rightarrow R_f$  is defined in 2.131 and is proved there to be surjective

2.  $e_{f(A)}: f(A) \rightarrow B$  is the inclusion function (see 2.27) which is injective (see 2.27)
3.  $s_f$  is bijective

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

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

1.  $\text{range}(s_f) \subseteq f(A)$ . If  $y \in \text{range}(s_f) \Rightarrow \exists x \vdash (x, y) \in s_f \Rightarrow \exists z \in A \vdash x = R_f[z], y = f(z) \Rightarrow y \in f(A) \Rightarrow \text{range}(s_f) \subseteq f(A)$
2.  $\text{dom}(s_f) = A / R_f$ . If  $x \in \text{dom}(s_f) \Rightarrow \exists y \vdash (x, y) \in s_f \Rightarrow \exists z \in A \vdash x = R_f[z] \wedge y = f(z) \Rightarrow x \in A / R_f \Rightarrow \text{dom}(s_f) \subseteq A / R_f$ . Now if  $x \in A / R_f \Rightarrow \exists z \in A \vdash x = R_f[z] \Rightarrow (R_f[z], f(z)) \in s_f \Rightarrow (x, f(z)) \in s_f \Rightarrow x \in \text{dom}(s_f) \Rightarrow A / R_f \subseteq \text{dom}(s_f)$  so we conclude that  $\text{dom}(s_f) = A / R_f$
3. If  $(x, y), (x, y') \in s_f \Rightarrow y = y'$ . This is the most important one to prove, so let  $(x, y), (x, y') \in s_f$  then there exists a  $z, z' \in A \vdash x = R_f[z] \wedge y = f(z) \wedge x = R_f[z'] \wedge y' = f(z')$  now from  $R_f[z] = x = R_f[z'] \Rightarrow R_f[z] = R_f[z']$  we have by 2.120 that  $z R_f z' \Rightarrow f(z) = f(z') \Rightarrow y = y'$

Next we must prove that  $s_f$  is a bijection

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

Finally we must prove that  $f = e_{f(A)} \circ s_f \circ f_{R_f}$  now if  $x \in A$  then we have by 2.38 that

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

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

□

So we conclude that  $A / R_f \approx f(A)$ . Further if  $f$  is surjective then  $f(A) = \text{range}(f) = B$  and then  $e_{f(A)}$  is  $i_{f(A)}$  which is a bijection so that by 2.49 we have that  $e_{f(A)} \circ s_f: A / R_f \rightarrow B$  is a bijection so  $A / R_f \approx B$ .

**Theorem 2.133.** Let  $f: A \rightarrow B$  be a function between sets  $A$  and  $B$ . Let  $R \subseteq R_f$  be a equivalence relation in  $A$  which is finer then  $R_f$  then we can define  $f / R: A / R \rightarrow B$  by  $f / R = \{(R[x], f(x)) | x \in A\}$ . Then we have that  $f$  is a function (called the **quotient of  $f$  by  $R$** ). Furthermore we have that  $\frac{R_f}{R} = R_{f/R}$

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

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

Now for the last part

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

□

## 2.3 Partially ordered classes

### 2.3.1 Order relation

**Definition 2.134. (Pre-order)** If  $A$  is a class then a relation  $R$  in  $A$  is a pre-order if it is reflexive and transitive or in other words

1.  $\forall x \in A \vdash xRx$  (reflexivity)
2.  $\forall x, y \in A \vdash xRy \wedge yRz \Rightarrow xRz$  (transitivity)

**Definition 2.135.** A class  $A$  together with a pre-order  $R$  on the class is called a pre-ordered class and noted by  $\langle A, R \rangle$  (if  $A$  is a set then  $\langle A, R \rangle$  is called a pre-ordered set)

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

**Definition 2.136. (Order relation)** If  $A$  is a class then a relation  $R$  in  $A$  is a order relation if it is reflexive, anti-symmetric and transitive or in other words

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

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

**Example 2.137.** If  $\mathcal{A}$  is a class of classes then  $\langle \mathcal{A}, \subseteq \rangle$  is a partially ordered class

**Proof.**

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

□

**Example 2.138.** Let  $A$  be a class,  $\langle B, \leq \rangle$  a partially ordered class then  $\langle B^A, \leq \rangle$  where  $\leq = \{(f, g) \in B^A \mid \forall x \in A \text{ we have } f(x) \leq g(x)\}$  forms a partially ordered class. Here  $B^A$  is a the class of all functions from  $A$  to  $B$ . Note that we say  $f \leq g$  if  $\forall x \in A$  we have  $f(x) \leq g(x)$

**Proof.**

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

□

**Definition 2.139.** If we have a class  $A$  together with a order relation  $R$  in  $A$  then the pair  $\langle A, R \rangle$  (see 1.73) is called a partially ordered class. If  $A$  is a set then  $\langle A, R \rangle$  is called a partially ordered set.

**Notation 2.140.** If  $\langle A, R \rangle$  is a partially ordered class (set) then  $R$  is noted as  $\leq$  so  $x \leq y$  is the same as  $(x, y) \in R$ . In words we say that  $x$  is smaller then  $y$ . Also  $y \geq x$  ( $y$  is greater then  $x$ ) is another way of saying  $x \leq y$ . Finally  $x < y$  ( $x$  is strictly less then  $y$ ) is a abbreviation for  $x \leq y \wedge x \neq y$  and  $y > x$  ( $y$  is strictly greater then  $x$ ) means  $x < y \wedge x \neq y$ .

**Theorem 2.141.** If  $\langle A, \leq \rangle$  is a partially ordered set then

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

**Proof.**

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

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

□

**Theorem 2.142.** If  $\langle A, \leq \rangle$  is a pre-ordered class then

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

**Proof.**

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

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

□

**Theorem 2.143.** If  $\langle A, \leq \rangle$  is a partial ordered (pre-ordered) class and  $B \subseteq A$  then we can define the relation  $\leq_{|B} = \bigcap (B \times B) \subseteq B \times B$  we have then that  $\langle B, \leq_{|B} \rangle$  is a partial ordered (pre-ordered) class. In general we use the simpler notation  $\langle B, \leq \rangle$  for the order relation (pre-order) on  $B$ .

**Proof.** The proof is quite trivial

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

□

**Definition 2.144.** If  $\langle A, \leq_A \rangle$  and  $\langle B, \leq_B \rangle$  are partially ordered sets then the lexical order  $\leq_{A,B} \subseteq (A \times B) \times (A \times B)$  is defined by  $\leq_{A,B} = \{((x, y), (x', y')) \in (A \times B) \times (A \times B) | ((x \neq x') \wedge (x \leq_A x')) \vee ((x = x') \wedge (y \leq_B y'))\}$

**Theorem 2.145.** If  $\langle A, \leq_A \rangle$  and  $\langle B, \leq_B \rangle$  are partially ordered sets then  $\langle A \times B, \leq_{A,B} \rangle$  is a partially ordered set

**Proof.**

1. If  $(x, y) \in A \times B$  then  $x = x \wedge y \leq_B y \Rightarrow (x, y) \leq_{A,B} (x, y)$
2. If  $(x, y) \leq_{A,B} (x', y')$  and  $(x', y') \leq_{A,B} (x, y)$  then if  $x \neq x'$  we must have  $x \leq_A x'$  and  $x' \leq_A x$  giving  $x = x'$  a contradiction. So we must have  $x = x'$  but then we have  $y \leq_B y'$  and  $y' \leq_B y$  and thus  $y = y'$
3. If  $(x, y) \leq_{A,B} (x', y')$  and  $(x', y') \leq_{A,B} (x'', y'')$  then consider the following possible cases for  $x, x'$ 
  - a.  $x = x' \Rightarrow y \leq_B y'$  now we have the following possibilities for  $x'$  and  $x''$ 
    - i.  $x' = x'' \Rightarrow y' \leq_B y'' \Rightarrow x = x'' \wedge y \leq_B y'' \Rightarrow (x, y) \leq_{A,B} (x'', y'')$

- ii.  $x' \neq x'' \Rightarrow x' \leq_A x'' \xrightarrow{x=x' \Rightarrow x \leq_A x'} x \neq x'' \wedge x \leq_A x'' \Rightarrow (x, y) \leq_{A,B} (x'', y'')$
- b.  $x \neq x' \Rightarrow x \leq_A x'$  now we have the following possibilities for  $x'$  and  $x''$ 
  - i.  $x' = x'' \Rightarrow x' \leq_A x'' \wedge x \neq x'' \Rightarrow x \neq x'' \wedge x \leq_A x'' \Rightarrow (x, y) \leq_{A,B} (x'', y'')$
  - ii.  $x' \neq x'' \Rightarrow x' \leq_A x'' \Rightarrow x \leq_A x''$  now we have the following cases
    - A.  $x = x'' \Rightarrow x' \leq_A x \Rightarrow x = x'$  a contradiction so this case does not occur
    - B.  $x \neq x'' \Rightarrow (x, y) \leq_{A,B} (x'', y'')$   $\square$

**Definition 2.146.** If  $\langle A, \leq \rangle$  is a partially ordered class then  $x, y \in A$  are comparable if  $x \leq y$  or  $y \leq x$

**Theorem 2.147.** If  $\langle A, \leq \rangle$  is a partially ordered class and  $x, y \in A$  are comparable then we have either  $x \leq y$  or  $y < x$

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

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

**Definition 2.148.** A partially ordered (pre-ordered) class  $\langle A, \leq \rangle$  is called a **fully ordered class**, **fully pre-ordered class**, **totally ordered class** (**totally pre-ordered class**) or a **linear ordered class** (**linear pre-ordered class**) if  $\forall x, y \in A \models x \leq y \vee y \leq x$ . A subclass  $B \subseteq A$  is called a **fully ordered (pre-ordered) subclass of A** or a **linear ordered (pre-ordered) subclass** or a **chain** of  $A$  if  $\forall x, y \in B \models x \leq y \vee y \leq x$ . So a **fully ordered (sub)class** is a (sub)class where every element is comparable with every other element. Also in a **fully ordered class** we have for every  $x, y$  either  $x \leq y$  or  $y < x$  (see previous theorem).

**Theorem 2.149.** Let  $\langle A, \leq \rangle$  be a totally ordered class,  $B \subseteq A$  then  $\langle B, \leq|_B \rangle$  is totally ordered (see 2.143)

**Proof.** Using 2.143 we have that  $\langle B, \leq|_B \rangle$  is partially ordered. If now  $x, y \in B \subseteq A$  then by the total ordering of  $\leq$  we have that  $(x, y) \in \leq$  or  $(y, x) \in \leq$  which as  $(x, y), (y, x) \in B \times B$  is equivalent with  $(x, y) \in \leq \cap B \times B$  or  $(y, x) \in \leq \cap B \times B$ .  $\square$

**Theorem 2.150.** Let  $\langle A, \leq_A \rangle$  and  $\langle B, \leq_B \rangle$  be fully ordered classes then  $\langle A \times B, \leq_{A,B} \rangle$  is a fully ordered class (see 2.145)

**Proof.** First  $\langle A \times B, \leq_{A,B} \rangle$  is a partially ordered class by 2.145. If now  $(x, y), (x', y') \in A \times B$  then we have for  $x, x'$  either

$x = x'$ . As  $\langle B, \leq_B \rangle$  is fully ordered we have either

$y \leq y'$ . then  $(x, y) \leq (x', y')$

$y' < y$ . then  $(x', y') \leq (x, y)$

$x \neq x'$ . As  $\langle A, \leq_A \rangle$  is fully ordered we have either

$x \leq x'$ . then  $(x, y) \leq (x', y')$

$x' \leq x$ . then  $(x', y') \leq (x, y)$   $\square$

**Definition 2.151.** If  $\langle A, \leq \rangle$  is a partially ordered class and  $a \in A$  then the **initial segment of A determined by a**  $S_a$  is defined by

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

**Theorem 2.152.** If  $\langle A, \leq \rangle$  and  $P$  is a initial segment of  $A$  and  $Q$  a initial segment of  $P$  (using the induced order relation on  $P$ ) then  $Q$  is a initial segment of  $A$

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

$$\begin{aligned}
 x \in S_b &= \{x \in A \mid x < b\} &\Leftrightarrow & x \in A \wedge x < b \\
 &\xrightarrow{x < b < a \Rightarrow x < a} && x \in A \wedge x < a \wedge x < b \\
 &\Rightarrow && x \in P \wedge x < b \\
 &\Rightarrow && x \in P \wedge x <_{|P} b \\
 &\Rightarrow && x \in Q \\
 x \in Q &\Rightarrow && x \in P \wedge x <_{|P} b \\
 &\Rightarrow && x \in P \wedge x < b \\
 &\xrightarrow{P \subseteq A} && x \in A \wedge x < b \\
 &\Rightarrow && x \in S_b \\
 &\square
 \end{aligned}$$

**Definition 2.153.** If  $\langle A, \leq \rangle$  is a partially ordered class then a **cut** of  $A$  is a pair  $\langle C, D \rangle$  such that

1.  $C \neq \emptyset \wedge D \neq \emptyset$
2.  $C \cap D = \emptyset$
3.  $x \in C \wedge y \leq x \Rightarrow y \in C$
4.  $x \in D \wedge y \geq x \Rightarrow y \in D$

**Definition 2.154.** If  $\langle A, \leq_A \rangle$  and  $\langle B, \leq_B \rangle$  are partially ordered classes then a function  $f: A \rightarrow B$  is **increasing** if  $\forall x, y \in A$  we have

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

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

**Theorem 2.155.** If  $\langle A, \leq_A \rangle$ ,  $\langle B, \leq_B \rangle$  and  $\langle C, \leq_C \rangle$  are partially ordered classes then  $f: A \rightarrow B$  and  $g: B \rightarrow C$  are increasing functions then  $g \circ f$  is a increasing function. Further if  $f$  and  $g$  is strictly increasing then  $g \circ f$  is strictly increasing.

**Proof.** If  $x \leq y \Rightarrow f(x) \leq f(y) \Rightarrow g(f(x)) \leq g(f(y)) \Rightarrow (g \circ f)(x) \leq (g \circ f)(y)$ . If  $g$  is strictly increasing then  $x < y \Rightarrow f(x) < f(y) \Rightarrow g(f(x)) < g(f(y)) \Rightarrow (g \circ f)(x) < (g \circ f)(y)$   $\square$

**Definition 2.156.** If  $\langle A, \leq_A \rangle$  and  $\langle B, \leq_B \rangle$  are partially ordered classes then a function  $f: A \rightarrow B$  is **strictly increasing** if  $\forall x, y \in A$  we have

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

**Definition 2.157.** If  $\langle A, \leq_A \rangle$  and  $\langle B, \leq_B \rangle$  are partially ordered classes then a function  $f: A \rightarrow B$  is a **isomorphism** iff  $f$  is bijective and  $\forall x, y \in A$  we have

$$x \leq_A y \Leftrightarrow f(x) \leq_B f(y)$$

**Notation 2.158.** If  $\langle A, \leq_A \rangle$  and  $\langle B, \leq_B \rangle$  are partially ordered class and there exists a isomorphism  $f: A \rightarrow B$  between  $A$  and  $B$  then we say that  $A$  and  $B$  are isomorphic noted by  $A \cong B$

**Theorem 2.159.** If  $\langle A, \leq_A \rangle$  and  $\langle B, \leq_B \rangle$  are partially ordered classes and  $f: A \rightarrow B$  a **isomorphism** then

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

**Proof.**

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

**Theorem 2.160.** If  $\langle A, \leqslant_A \rangle$  and  $\langle B, \leqslant_B \rangle$  are partially ordered classes then a bijective function  $f: A \rightarrow B$  is a isomorphism iff  $f: A \rightarrow B$  and  $f^{-1}: B \rightarrow A$  are increasing functions

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

1. If  $f$  is a isomorphism then  $x \leqslant_A y \Rightarrow f(x) \leqslant_B f(y) \Rightarrow f$  is a increasing function. If  $x, y \in B$  with  $x \leqslant_A y$  then  $f(f^{-1}(x)) = x \leqslant_A y = f(f^{-1}(y)) \xrightarrow{f \text{ is a isomorphism}} f^{-1}(x) \leqslant_B f^{-1}(y)$  so  $f^{-1}$  is a increasing function.
2. Suppose  $f, f^{-1}$  are increasing functions then if  $x \leqslant_A y \xrightarrow{f \text{ is increasing}} f(x) \leqslant_B f(y)$ . Suppose that  $f(x) \leqslant_B f(y) \xrightarrow{f^{-1} \text{ is increasing}} f^{-1}(f(x)) \leqslant_A f^{-1}(f(y)) \Rightarrow x \leqslant y$   $\square$

**Theorem 2.161.** If  $\langle A, \leqslant_A \rangle$ ,  $\langle B, \leqslant_B \rangle$  and  $\langle C, \leqslant_C \rangle$  are partially ordered classes then

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

**Proof.**

1. By 2.26 we have that  $1_A: A \rightarrow A$  is a bijection then as  $x = 1_A(x)$  and  $y = 1_A(y)$  so that  $x \leqslant y \Leftrightarrow 1_A(x) \leqslant 1_A(y)$ .
2. If  $f: A \rightarrow B$  is a isomorphism then by 2.41 we have that  $f^{-1}: B \rightarrow A$  is a bijection. By the previous theorem we have that  $f^{-1}$  is increasing. So assume  $f^{-1}(x) \leqslant f^{-1}(y) \xrightarrow{f^{-1} \text{ is increasing}} f(f^{-1}(x)) \leqslant f(f^{-1}(y)) \Rightarrow x \leqslant y$ .
3. Using 2.49 we have that  $g \circ f$  is a bijection and that  $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$ . Now using 2.155 and the fact that  $f, g, f^{-1}, g^{-1}$  are increasing we have that  $g \circ f$  and  $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$  are increasing. Using 2.160 we have then that  $g \circ f$  is a isomorphism.  $\square$

**Theorem 2.162.** 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.  $\square$

**Theorem 2.163.** If  $\langle A, \leqslant_A \rangle$  is a fully ordered class and  $\langle B, \leqslant_B \rangle$  is a partially ordered class then a bijective and increasing function  $f: A \rightarrow B$  is a isomorphism

**Proof.** Suppose that  $f(x) \leqslant_B f(y)$  then since  $A$  is fully ordered we have that  $x, y$  are comparable therefore by 2.147 we have the following exclusive cases

1.  $x \leqslant_A y$  in this case or theorem is proved
2.  $y <_A x$  in this case we would have  $f(y) \leqslant_B f(x) \Rightarrow f(y) = f(x) \xrightarrow{f \text{ is injective}} x = y$  a contradiction. So this case does not occurs.  $\square$

**Definition 2.164.** If  $\langle X, \leqslant \rangle$  is a pre-ordered class and  $A \leqslant X$  then

1.  $m$  is a maximal element of  $A$  iff  $m \in A$  and if  $x \in A$  with  $m \leqslant x$  then  $x \leqslant m$

2.  $m$  is a minimal element of  $A$  iff  $m \in A$  and if  $x \in A$  with  $x \leq m$  then  $m \leq x$

**Definition 2.165.** If  $\langle X, \leq \rangle$  is a partially ordered class and  $A \subseteq X$  then

1.  $m$  is a maximal element of  $A$  iff  $m \in A$  and  $\forall x \in A$  if  $x \geq m \Rightarrow x = m$
2.  $m$  is a minimal element of  $A$  iff  $m \in A$  and  $\forall x \in A$  if  $x \leq m \Rightarrow x = m$
3.  $m$  is the greatest element of  $A$  iff  $m \in A$  and  $\forall x \in A \vdash x \leq m$
4.  $m$  is the least element of  $A$  iff  $m \in A$  and  $\forall x \in A \vdash m \leq x$

note that (3) and (4) suggests that there is only one greatest and least element, this is indeed true as is proved in the next theorem.

**Note 2.166.** There is a difference between the definition of a maximal element and a greatest element. If  $m$  is a greatest element of  $A$  then  $m$  is comparable with every element in  $A$ . This is not needed if  $m$  is a maximal element of  $A$ . The same goes for minimal element and least element.

**Theorem 2.167.** If  $\langle X, \leq \rangle$  is a partially ordered class and  $A \subseteq X$  then if  $A$  has a greatest element then this is unique. If  $A$  has a least element then it is unique. The unique greatest element of  $A$  if it exists is called the maximum of  $A$  and noted by  $\max(A)$ . The unique least element if it exists is called the minimum of  $A$  and noted by  $\min(A)$ .

Note: The existence of a maximum is not the same as the existence of a maximal element, the existence of a minimum is not the same as the existence of a minimal element.

**Proof.**

1. If  $m, m'$  are greatest elements then  $m, m' \in A \Rightarrow m \leq m' \wedge m' \leq m \Rightarrow m = m'$
2. If  $m, m'$  are least elements then  $m, m' \in A \Rightarrow m \leq m' \wedge m' \leq m \Rightarrow m = m'$

□

**Theorem 2.168.** Let  $\langle X, \leq \rangle$  be a partially ordered class and let  $B, C \subseteq X$  sub-classes such that  $\max(B), \max(C)$  exists [or  $\min(B), \min(C)$  exists] then if  $\forall x \in B$  there exists a  $y \in C$  with  $x \leq y$  [or  $\forall x \in B$  there exists a  $y \in C$  such that  $y \leq x$ ] then  $\max(B) \leq \max(C)$  [or  $\min(C) \leq \min(B)$ ]

**Proof.** If  $\max(B), \max(C)$  exists then  $\max(B) \in B$  and thus there exist a  $c \in C$  such that  $\max(B) \leq c \leq \max(C) \Rightarrow \max(B) \leq \max(C)$ . If  $\min(B), \min(C)$  exists then as  $\min(B) \in B$  there exists a  $c \in C$  such that  $\min(C) \leq c \leq \min(B) = \min(C) \leq \min(B)$ . □

**Definition 2.169.** If  $\langle A, \leq \rangle$  is a pre-ordered class and  $B \subseteq A$  then

1.  $u \in A$  is a upper bound of  $B$  iff  $\forall x \in B \vdash x \leq u$
2.  $B$  is bounded above if it has a upper bound
3.  $l \in A$  is a lower bound of  $B$  iff  $\forall x \in B \vdash l \leq x$
4.  $B$  is bounded below if it has a lower bound

**Definition 2.170.** If  $\langle A, \leq \rangle$  is a partially ordered class and  $B \subseteq A$  a subclass of  $A$  then

1.  $u \in A$  is a upper bound of  $B$  iff  $\forall x \in B \vdash x \leq u$
2.  $B$  is bounded above if it has a upper bound
3.  $l \in A$  is a lower bound of  $B$  iff  $\forall x \in B \vdash l \leq x$
4.  $B$  is bounded below if it has a lower bound
5.  $v(B) = \{u \in A \mid u \text{ is a upper bound of } B\}$  (the class of all the upper bounds of  $B$ )
6.  $\lambda(B) = \{l \in A \mid l \text{ is a lower bound of } B\}$  (the class of all the lower bounds of  $B$ )
7. If  $v(B)$  has a least element then this is called the **least upper bound** or **supremum** of  $B$  which is noted as  $\sup(B)$  (it is **unique** by the previous theorem).

8. If  $\lambda(B)$  has a greatest element then this is called the **greatest lower bound** or **infimum** of  $B$  which is noted as  $\inf(B)$  (it is **unique** by the previous theorem).

**Theorem 2.171.** If  $\langle A, \leq \rangle$  is a partially ordered class and  $B \subseteq A$  then we have

1. If  $\inf(B)$  exists then  $\forall a \in A$  with  $\inf(B) < a$  there exists a  $b \in B$  such that  $\inf(B) \leq b < a$
2. If  $\sup(B)$  exists then  $\forall a \in A$  with  $a < \sup(B)$  there exists a  $b \in B$  such that  $a < b \leq \sup(B)$

**Proof.**

1. We proceed by contradiction so assume that  $\exists a \in A$  with  $\inf(B) < a$  so that  $\forall b \in B$  we have that  $\neg(\inf(B) \leq b < a) = \neg(\inf(B) \leq b \wedge b < a) = b < \inf(B) \vee a \leq b$   $\Rightarrow$   $a \leq b$  inf(B) is a lower bound of B  $\Rightarrow \inf(B) \leq b$   $b \Rightarrow a$  is a lower bound of  $B$  and thus by definition of  $\inf(B)$  we have  $a \leq \inf(B)$  contradicting  $\inf(B) < a$ . So we must have that  $\forall a \in A$  with  $\inf(B) < a$  there exist a  $b \in B$  such that  $\inf(B) \leq b < a$ .
2. We proceed by contradiction so assume that  $\exists a \in A$  with  $a < \sup(B)$  so that  $\forall b \in B$  we have that  $\neg(a < b \leq \sup(B)) = \neg(a < b \wedge b \leq \sup(B)) = b < a \vee \sup(B) < b$   $\Rightarrow$   $b < a$  is a upper bound of  $B$   $\Rightarrow b \leq \sup(B)$   $b \leq a \Rightarrow a$  is a upper bound of  $B$  and thus by definition of  $\sup(B)$  we have  $\sup(B) \leq a$  contradicting  $a < \sup(B)$ . So we must have that  $\forall a \in A$  with  $a < \sup(B)$  there exist a  $b \in B$  such that  $a < b \leq \sup(B)$ .  $\square$

**Theorem 2.172.** Let  $\mathcal{A}$  be a class of classes partially ordered by inclusion  $(\langle \mathcal{A}, \subseteq \rangle)$  then if  $\mathcal{B} \subseteq \mathcal{A}$  then

1. If  $\bigcap_{B \in \mathcal{B}} B \in \mathcal{A} \Rightarrow \inf(\mathcal{B}) = \bigcap_{B \in \mathcal{B}} B$
2. If  $\bigcup_{B \in \mathcal{B}} B \in \mathcal{A} \Rightarrow \sup(\mathcal{B}) = \bigcup_{B \in \mathcal{B}} B$

**Proof.**

1. If  $B \in \mathcal{B} \Rightarrow \bigcap_{B \in \mathcal{B}} B \subseteq B \Rightarrow \bigcap_{B \in \mathcal{B}} B \in \lambda(\mathcal{B})$  now if  $C \in \lambda(\mathcal{B})$  then  $\forall B \in \mathcal{B}$  we have  $C \subseteq B \Rightarrow C \subseteq \bigcap_{B \in \mathcal{B}} B$  so  $\bigcap_{B \in \mathcal{B}} B$  is the greatest element of  $\lambda(\mathcal{B})$
2. If  $B \in \mathcal{B} \Rightarrow B \subseteq \bigcup_{B \in \mathcal{B}} B \Rightarrow \bigcup_{B \in \mathcal{B}} B \in v(\mathcal{B})$  now if  $C \in v(\mathcal{B})$  then  $\forall B \in \mathcal{B}$  we have  $B \subseteq C \Rightarrow \bigcup_{B \in \mathcal{B}} B \subseteq C$  so  $\bigcup_{B \in \mathcal{B}} B$  is the least element of  $v(\mathcal{B})$   $\square$

**Theorem 2.173.** If  $\langle A, \leq \rangle$  is a partially ordered class then if  $B, C \subseteq A$  with  $B \subseteq C$  we have

1. If  $B$  has a greatest element  $b$  and  $C$  has a greatest element  $c$  then  $b \leq c$
2. If  $B$  has a least element  $b$  and  $C$  has a least element  $c$  then  $c \leq b$

**Proof.**

1.  $b \in B \Rightarrow b \in C \xrightarrow{\forall x \in C \vdash x \leq c} b \leq c$
2.  $b \in B \Rightarrow b \in C \xrightarrow{\forall x \in C \vdash c \leq x} c \leq b$   $\square$

**Theorem 2.174.** If  $\langle A, \leq \rangle$  is a partially ordered class then if  $B, C \subseteq A$  with  $B \subseteq C$  we have

1.  $v(C) \subseteq v(B)$
2.  $\lambda(C) \subseteq \lambda(B)$

**Proof.**

1. If  $x \in v(C)$  then  $\forall b \in B \vdash b \in C \Rightarrow b \leq x \Rightarrow x \in v(B)$
2. If  $x \in \lambda(C)$  then  $\forall b \in B \vdash b \in C \Rightarrow x \leq b \Rightarrow x \in \lambda(B)$   $\square$

**Theorem 2.175.** If  $\langle A, \leq \rangle$  is a partially ordered class then if  $B, C \subseteq A$  with  $B \subseteq C$

1. If the supremum exists for  $B, C$  then  $\sup(B) \leq \sup(C)$

2. If the infinum exists for  $B, C$  then  $\inf(C) \leq \inf(B)$

**Proof.**

1. By 2.174 we have  $v(C) \subseteq v(B)$  so using 2.173 we have  $\sup(B) = \text{least element of } v(B) \leq \text{least element of } v(C) = \sup(C)$
2. By 2.174 we have  $\lambda(C) \subseteq \lambda(B)$  so using 2.173 we have  $\inf(C) = \text{greatest element of } \lambda(C) \leq \text{greatest element of } B = \inf(B)$   $\square$

**Theorem 2.176.** If  $\langle A, \leq \rangle$  is a partially ordered class,  $B, C \subseteq A$  are subsets of  $A$  such that  $\forall b \in B$  there exists a  $c \in C$  such that  $b \leq c$  then  $\sup(B) \leq \sup(C)$  (assuming that the supremum of  $B, C$  exists). On the other hand if  $\forall b \in B$  there exists a  $c \in C$  such that  $c \leq b$  then  $\inf(C) \leq \inf(B)$  (assuming that the infinum of  $B, C$  exists).

**Proof.** If  $b \in B$  then there exists a  $c \in C$  such that  $b \leq c \leq \sup(C)$  ( $\sup(C)$  is a upper bound of  $C$ ) so that  $\sup(C)$  is a upper bound of  $B$  and thus we have  $\sup(B) \leq \sup(C)$ .

On the other hand if  $b \in B$  then there exists a  $c \in C$  such that  $\inf(C) \leq c \leq b$  so that  $\inf(C)$  is a lower bound of  $B$  so that  $\inf(C) \leq \inf(B)$   $\square$

**Theorem 2.177.** If  $\langle A, \leq \rangle$  is a partially ordered class with  $B \subseteq A$  then  $B \subseteq v(\lambda(B))$  and  $B \subseteq \lambda(v(B))$

**Proof.**

1. If  $x \in B \Rightarrow \forall y \in \lambda(B) \vdash y \leq x \Rightarrow x \in v(\lambda(B))$
2. If  $x \in B \Rightarrow \forall y \in v(B) \vdash x \leq y \Rightarrow x \in \lambda(v(B))$   $\square$

**Theorem 2.178.** If  $\langle A, \leq \rangle$  is a partially ordered class with  $B \subseteq A$  a subclass then

1. If  $\lambda(B)$  has a supremum then  $B$  has a infinum and  $\sup(\lambda(B)) = \inf(B)$
2. If  $v(B)$  has a infinum then  $B$  has a supremum and  $\inf(v(B)) = \sup(B)$

**Proof.**

1. Let  $s = \sup(\lambda(B))$  then if  $b \in B$  we have  $\forall y \in \lambda(B)$  that  $y \leq b$  hence  $b \in v(\lambda(B))$  thus ( $s$  is least upper bound of  $\lambda(B)$ )  $s \leq b$ . As  $b$  has been chosen arbitrary we have that  $s \in \lambda(B)$ . Now if  $d \in \lambda(B)$  then  $d \leq s$  [because  $s = \sup(\lambda(B)) \Rightarrow s$  is the least element of  $v(\lambda(B)) \Rightarrow s \in v(\lambda(B))$ ] so  $s$  is the greatest element of  $\lambda(B)$  hence  $s = \inf(B)$ .
2. Let  $s = \inf(v(B))$  then if  $b \in B$  we have  $\forall y \in v(B)$  that  $b \leq y$  hence  $b \in \lambda(v(B))$  thus ( $s$  is greatest lower bound of  $v(B)$ )  $b \leq s$ . As  $b$  has been chosen arbitrary we have  $s \in v(B)$ . Now if  $d \in v(B)$  then  $s \leq d$  [because  $s = \inf(v(B)) \Rightarrow s$  is the greatest element of  $\lambda(v(B)) \Rightarrow s \in \lambda(v(B))$ ] so  $s$  is the least element of  $v(B)$  hence  $s = \sup(B)$   $\square$

**Definition 2.179.** If  $\langle A, \leq \rangle$  is a partially ordered class. If every nonempty subclass of  $A$  that is bounded above has a supremum then  $A$  is said to be **conditional complete**.

Next theorem shows that conditional completeness can be defined also with lower bounds and infinum

**Theorem 2.180.** If  $\langle A, \leq \rangle$  is a partially ordered class then the following are equivalent

1. Every nonempty subclass of  $A$  that is bounded above has a supremum
2. Every nonempty subclass of  $A$  that is bounded below has a infinum

**Proof.**

$$1 \Rightarrow 2$$

Let  $B \subseteq A$  which is bounded below  $\Rightarrow \lambda(B) \neq \emptyset$ . Now as  $B \neq \emptyset$  there exists a  $b \in B$  and by definition of  $\lambda(B)$  we have  $\forall y \in \lambda(B)$  that  $y \leq b$ . So  $\lambda(B)$  is nonempty and bounded above. From the hypothesis we have then that  $\sup(\lambda(B))$  exists but then by 2.178 we have that  $B$  has a infinum.

$$2 \Rightarrow 1$$

Let  $B \subseteq A$  which is bounded above  $\Rightarrow v(B) \neq \emptyset$ . Now as  $B \neq \emptyset$  there exists a  $b \in B$  and by definition of  $v(B)$  we have  $\forall y \in v(B)$  that  $b \leq y$ . So  $v(B)$  is nonempty and bounded below. From the hypothesis we have then that  $\inf(v(B))$  exists and then by 2.178 we have that  $B$  has a supremum.  $\square$

We show now that a order preserving isomorphism conserves the supremum and infimum

**Lemma 2.181.** *Let  $\langle A, \leq_A \rangle$ ,  $\langle B, \leq_B \rangle$  be partially ordered classes and  $f: A \rightarrow B$  a order preserving isomorphism (see isomorphism) then*

1. *If  $b$  is a upper bound of  $f(S)$  then  $(f^{-1})(b)$  is a upper bound of  $f^{-1}(S)$*
2. *If  $l$  is a lower bound of  $f(S)$  then  $(f^{-1})(l)$  is a lower bound of  $f^{-1}(S)$*
3. *If  $S \subseteq A$  so that  $\sup(S)$  exists then  $\sup(f(S))$  exists and is equal to  $f(\sup(S))$*
4. *If  $S \subseteq A$  so that  $\inf(S)$  exists then  $\inf(f(S))$  exists and is equal to  $f(\inf(S))$*

**Proof.** As  $f$  is a isomorphism we have using 2.160 that  $f$  and  $f^{-1}$  are increasing functions.

1. Let  $b$  be a upper bound of  $f(S)$  then if  $a \in f^{-1}(S)$  we have  $f(a) \in S$  so that  $f(a) \leq_B b$  and thus  $a = f^{-1}(f(a)) \leq_A f^{-1}(b)$  proving that  $f^{-1}(b)$  is a upper bound
2. Let  $l$  be a lower bound of  $f(S)$  then if  $a \in f^{-1}(S)$  we have  $f(a) \in S$  so that  $l \leq_B f(a)$  and thus  $f^{-1}(l) \leq_A f^{-1}(f(a)) = a$  proving that  $f^{-1}(l)$  is a lower bound.
3. If  $S \subseteq A$  so that  $\sup(S)$  exist then if  $a \in f(S)$  there exists  $a' \in S$  such that  $a = f(a')$ , as  $\sup(S)$  is a upper bound of  $S$  we have that  $a' \leq_A \sup(S)$  so that  $a = f(a') \leq_B f(\sup(S))$  proving that

$$f(\sup(S)) \text{ is a upper bound of } f(S) \quad (2.18)$$

If now  $b$  is a upper bound of  $f(S)$  then using (1)  $(f^{-1})(b)$  is a upper bound of  $f^{-1}(f(S))$   $f$  is a bijection  $\equiv S$  hence  $\sup(S) \leq_A (f^{-1})(b)$  so that  $f(\sup(S)) \leq_B f((f^{-1})(b)) = b$  which proves together with 2.18 that  $f(\sup(S)) = \sup(f(S))$

4. If  $S \subseteq A$  so that  $\inf(S)$  exists then if  $a \in f(S)$  there exists a  $a' \in S$  such that  $a = f(a')$ , as  $\inf(S)$  is a lower bound of  $S$  we have that  $\inf(S) \leq_A a'$  so that  $f(\inf(S)) \leq_B f(a') = a$  proving that

$$f(\inf(S)) \text{ is a lower bound of } f(S) \quad (2.19)$$

If now  $l$  is a lower bound of  $f(S)$  then using (2)  $(f^{-1})(b)$  is a lower bound of  $f^{-1}(f(S))$   $f$  is a bijection  $\equiv S$  hence  $(f^{-1})(b) \leq_A \inf(S)$  so that  $b = f((f^{-1})(b)) \leq_B f(\inf(S)) = b$  which proves together with 2.18 that  $f(\inf(S)) = \inf(f(S))$   $\square$

**Corollary 2.182.** *Let  $\langle A, \leq_A \rangle$  be a conditionally complete partial ordered set,  $\langle B, \leq_B \rangle$  a partially ordered set and  $f: A \rightarrow B$  a isomorphism (see isomorphism) then  $\langle B, \leq_B \rangle$  is conditionally complete.*

**Proof.** Let  $\emptyset \neq S \subseteq B$  be a non empty set such that there exists a upper bound  $b$  then using the previous theorem (see 2.181) we have that  $f^{-1}(b)$  is a upper bound of  $f^{-1}(S)$ . As  $S \neq \emptyset$  there exists a  $a \in S$  and as  $f$  is a bijection there exists a  $a' \in A$  such that  $f(a') = a$  hence  $a' \in f^{-1}(S)$ . From the conditionally completeness of  $\langle A, \leq_A \rangle$  we have that  $\sup(f^{-1}(S))$  exists. Using 2.181 we have that  $S = f(f^{-1}(S))$  has a supremum.  $\square$

**Definition 2.183.** *Let Let  $\langle A, \leq \rangle$  be a pre-ordered set then if  $a, b \in A$  we define*

1.  $[a, b] = \{x \in A | a \leq x \wedge x \leq b\}$
2.  $]a, b] = \{x \in A | a < x \wedge x \leq b\}$
3.  $[a, b[ = \{x \in A | a \leq x \wedge x < b\}$

4.  $]a, b[ = \{x \in A \mid a < x \wedge x < b\}$
5.  $]-\infty, b] = \{x \in A \mid x \leq b\}$
6.  $]-\infty, b[ = \{x \in A \mid x < b\}$
7.  $[a, \infty[ = \{x \in A \mid a \leq x\}$
8.  $]a, \infty[ = \{x \in A \mid a < x\}$
9.  $]-\infty, \infty[ = A$

**Note 2.184.** If  $b < a$  then  $[a, b] = [a, b[ = ]a, b] = ]a, b[ = \emptyset$  and if  $b \leq a$  then  $]a, b] = [a, b[ = ]a, b[ = \emptyset$

**Proof.** If  $b < a$  then we have

1. If  $z \in [a, b]$  then  $b < a \wedge a \leq z \wedge z \leq b$  so that  $b < z \wedge z \leq b$  a contradiction.
2. If  $z \in ]a, b]$  then  $b < a \wedge a < z \wedge z \leq b$  so that  $b < z \wedge z \leq b$  a contradiction.
3. If  $z \in [a, b[$  then  $b < a \wedge a \leq z \wedge z < b$  so that  $b < z \wedge z < b$  a contradiction.
4. If  $z \in ]a, b]$  then  $b < a \wedge a < z \wedge z \leq b$  so that  $b < z \wedge z < b$  a contradiction.

Finally if  $b \leq a$  then we have

1. If  $z \in [a, b[$  then  $b \leq a \wedge a \leq z \wedge z < b$  so that  $b \leq z \wedge z < b$  a contradiction.
2. If  $z \in ]a, b]$  then  $b \leq a \wedge a < z \wedge z \leq b$  so that  $b < z \wedge z \leq b$  a contradiction.
3. If  $z \in ]a, b[$  then  $b \leq a \wedge a < z \wedge z < b$  so that  $b < z \wedge z < b$  a contradiction.

□

**Definition 2.185.** Let  $\langle A, \leq \rangle$  be a partial ordered class that is conditional complete then  $I \subseteq A$  is a **generalized interval** if  $\forall x, y \in I$  we have  $[x, y] \in I$ .

We have the following equivalent conditions for a generalized interval.

**Theorem 2.186.** Let  $\langle A, \leq \rangle$  be a partial ordered class that is conditional complete and  $I \subseteq A$  then the following are equivalent

1.  $I$  is a generalized interval
2.  $\forall x, y \in I$  we have that  $]x, y[ \subseteq I$

**Proof.**

**1 ⇒ 2.** Let  $x, y \in I$  then if  $z \in ]x, y[$  we have  $x < z \wedge z < y \Rightarrow x \leq z \wedge z \leq y \Rightarrow z \in [x, y] \subseteq I$  proving that  $]x, y[ \subseteq I$

**2 ⇒ 1.** Let  $x, y \in I$  then if  $z \in [x, y]$  we have either

**$x = y$ .** Then  $x \leq z \wedge z \leq x \Rightarrow z = x \in I$

**$x < y$ .** Then we have either

**$z = x$ .** Then  $z \in I$

**$z = y$ .** Then  $z \in I$

**$z \neq x \wedge z \neq y$ .** Then  $x < z \wedge z < y$  proving that  $z \in ]x, y[ \subseteq I$

So in all cases we have that  $z \in I$  proving that  $[x, y] \subseteq I$

□

**Theorem 2.187.** Let  $\langle A, \leq \rangle$  be a fully ordered class that is conditional complete and  $\emptyset \neq I \subseteq A$  a generalized interval then we have either

1. If  $I$  is bounded below and above (so that by conditional completeness  $\inf(I)$  and  $\sup(I)$  exists) we have:
  - a. If  $\inf(I) \in I \wedge \sup(I) \in I$  then  $I = [\inf(I), \sup(I)]$
  - b. If  $\inf(I) \notin I \wedge \sup(I) \in I$  then  $I = ]\inf(I), \sup(I)]$

- c. If  $\inf(I) \in I \wedge \sup(I) \notin I$  then  $I = [\inf(I), \sup(I)[$
- d. If  $\inf(I) \notin I \wedge \sup(I) \notin I$  then  $I = ]\inf(I), \sup(I)[$
- 2. If  $I$  is bounded below and not bounded above (so that by conditional completeness  $\inf(I)$  exists) we have
  - a. If  $\inf(I) \in I$  then  $I = [\inf(I), \infty[$
  - b. If  $\inf(I) \notin I$  then  $I = ]\inf(I), \infty[$
- 3. If  $I$  is bounded above and not bounded below (so that by conditional completeness  $\sup(I)$  exists) we have
  - a. If  $\sup(I) \in I$  then  $I = ]-\infty, \sup(I)]$
  - b. If  $\sup(I) \notin I$  then  $I = ]-\infty, \sup(I)[$
- 4. If  $I$  is not bounded above and not bounded below we have  $I = A$

In other words a generalized interval  $I$  has either one of the following forms

- $[\inf(I), \sup(I)]$
- $]\inf(I), \sup(I)]$
- $[\inf(I), \sup(I)[$
- $]\inf(I), \sup(I)[$
- $]\inf(I), \infty]$
- $]\inf(I), \infty[$
- $]-\infty, \sup(I)]$
- $]-\infty, \sup(I)[$
- $[\inf(I), \infty[$
- $]\inf(I), \infty[$
- $A$

### Proof.

1.
  - a. Let  $z \in I$  then using the definition of a supremum and infimum and the fact that  $\inf(I) \in I \wedge \sup(I) \in I$  we have  $\inf(I) \leq z \wedge z \leq \sup(I)$  so that

$$I \subseteq [\inf(I), \sup(I)] \quad (2.20)$$

As  $I$  is a generalized interval we have further that  $[\inf(I), \sup(I)] \subseteq I$  which using 2.20 proves

$$I = [\inf(I), \sup(I)]$$

1.
  - b. Let  $z \in I$  then using the definition of a supremum and infimum and the fact that  $\inf(I) \notin I \wedge \sup(I) \in I$  we have that  $\inf(I) < z \wedge z \leq \sup(I)$  so that

$$I \subseteq ]\inf(I), \sup(I)] \quad (2.21)$$

If  $z \in ]\inf(I), \sup(I)]$  we have  $\inf(I) < z \wedge z \leq \sup(I)$ . Using 2.171 there exists a  $x \in I$  such that  $\inf(I) \leq x < z$  or  $z \in [x, \sup(I)] \subseteq I$  [as  $I$  is a generalized interval and  $x, \sup(I) \in I$ ]. This proves that  $]\inf(I), \sup(I)] \subseteq I$ . Combining this with 2.21 proves

$$I = ]\inf(I), \sup(I)]$$

1.
  - c. Let  $z \in I$  then using the definition of a supremum and infimum and the fact that  $\inf(I) \in I \wedge \sup(I) \notin I$  we have that  $\inf(I) \leq z \wedge z < \sup(I)$  so that

$$I \subseteq [\inf(I), \sup(I)[ \quad (2.22)$$

If  $z \in [\inf(I), \sup(I)]$  we have  $\inf(I) \leq z \wedge z < \sup(I)$ . Using 2.171 there exists a  $y \in I$  such that  $z < y \leq \sup(I)$  or  $z \in [\inf(I), y] \subseteq I$  [as  $I$  is a generalized interval and  $\inf(I), y \in I$ ]. This proves that  $[\inf(I), \sup(I)] \subseteq I$ . Combining this with 2.22 proves

$$I = [\inf(I), \sup(I)]$$

- d. Let  $z \in I$  then using the definition of a supremum and infimum and the fact that  $\inf(I) \notin I \wedge \sup(I) \notin I$  we have that  $\inf(I) < z \wedge z < \sup(I)$  so that

$$I \subseteq ]\inf(I), \sup(I)[ \quad (2.23)$$

If  $z \in ]\inf(I), \sup(I)[$  we have  $\inf(I) < z \wedge z < \sup(I)$ . Using 2.171 there exists  $x, y \in I$  such that  $\inf(I) \leq x < z \wedge z < y \leq \sup(I)$  or  $z \in ]x, y[ \subseteq I$  [as  $I$  is a generalized interval and  $x, y \in I$ ]. This proves that  $]inf(I), sup(I)[ \subseteq I$ . Combining this with 2.23 gives

$$I = ]\inf(I), \sup(I)[$$

2.

- a. Let  $z \in I$  then using the definition of the infimum and the fact that  $\inf(I) \in I$  we have that  $\inf(I) \leq z$  proving that

$$I \subseteq [\inf(I), \infty[ \quad (2.24)$$

If  $z \in [\inf(I), \infty[$  then  $\inf(I) \leq z$ . As  $I$  is not bounded above there exists a  $x \in I$  such that  $\neg(x \leq z) \xrightarrow{A \text{ is fully ordered}} z < x$ . Hence  $z \in [\inf(I), x] \subseteq I$  [as  $I$  is a generalized interval and  $\inf(I), x \in I$ ]. So  $[\inf(I), \infty[ \subseteq I$  and combining this with 2.24 proves

$$I = [\inf(I), \infty[$$

- b. Let  $z \in I$  then using the definition of the infimum and the fact that  $\inf(I) \notin I$  we have that  $\inf(I) < z$  proving that

$$I \subseteq ]\inf(I), \infty[ \quad (2.25)$$

If  $z \in ]\inf(I), \infty[$  then  $\inf(I) < z$ . Using 2.171 there exists a  $x \in I$  such that  $\inf(I) \leq x < z$ . As  $I$  is not bounded above there exists a  $y \in I$  such that  $\neg(y \leq z) \xrightarrow{A \text{ is fully ordered}} z < y$ . Hence  $z \in ]x, y[ \subseteq I$  [as  $I$  is a generalized interval and  $x, y \in I$ ]. So  $]inf(I), \infty[ \subseteq I$  and combining this with 2.25 gives

$$I = ]\inf(I), \infty[$$

3.

- a. Let  $z \in I$  then using the definition of the supremum and the fact that  $\sup(I) \in I$  we have  $z \leq \sup(I)$  proving that

$$I \subseteq ]-\infty, \sup(I)] \quad (2.26)$$

If  $z \in ]-\infty, \sup(I)]$  then  $z \leq \sup(I)$ . As  $I$  is not bounded below there exists a  $x \in I$  such that  $\neg(z \leq x) \xrightarrow{A \text{ is fully ordered}} x < z$ . Hence  $z \in [x, \sup(I)] \subseteq I$  [as  $I$  is a generalized interval and  $x, \sup(I) \in I$ ]. So  $]-\infty, \sup(I)] \subseteq I$  which combined with 2.26 gives

$$I = ]-\infty, \sup(I)]$$

- b. Let  $z \in I$  then using the definition of the supremum and the fact that  $\sup(I) \notin I$  we have that  $z < \sup(I)$  proving that

$$I \subseteq ]-\infty, \sup(I)[ \quad (2.27)$$

If  $z \in ]-\infty, \sup(I)[$  then  $z < \sup(I)$ . Using 2.171 there exists a  $y \in I$  such that  $z < y \leq \sup(I)$ . As  $I$  is not bounded below there exists a  $x \in I$  such that  $\neg(z \leq x) \xrightarrow{A \text{ is fully ordered}} x < z$ . Hence  $z \in ]x, y[ \subseteq I$  [as  $I$  is a generalized interval and  $x, y \in I$ ]. So  $]-\infty, \sup(I)[ \subseteq I$  which combined with 2.27 proves

$$I = ]-\infty, \sup(I)[$$

4. If  $z \in A$  then as  $I$  is not bounded below and not bounded above there exists  $x, y \in I$  such that  $x < z \wedge z < y$  giving  $z \in ]x, y[ \subseteq I$  [as  $I$  is a generalized interval and  $x, y \in I$ ]. This proves that  $A \subseteq I$  and as trivially we have  $I \subseteq A$  it follows that

$$I = A$$

□

We have also the opposite of the above theorem

**Theorem 2.188.** *Let  $\langle A, \leq \rangle$  be a fully ordered class that is conditional complete then we have given  $a, b \in A$  that*

1.  $[a, b]$  is a generalized set
2.  $[a, b[$  is a generalized set
3.  $]a, b]$  is a generalized set
4.  $]a, b[$  is a generalized set
5.  $[a, \infty[$  is a generalized set
6.  $]a, \infty[$  is a generalized set
7.  $]-\infty, a]$  is a generalized set
8.  $]-\infty, a[$  is a generalized set

**Proof.**

1. If  $x, y \in [a, b]$  and  $z \in [x, y]$  then we have  $x \leq z \xrightarrow{a \leq x} a \leq z$  and  $z \leq y \xrightarrow{y \leq b} z \leq b$  proving that  $z \in [a, b]$ . Hence  $[x, y] \subseteq [a, b]$  from which it follows that

$$[a, b] \text{ is a generalized interval}$$

2. If  $x, y \in [a, b[$  and  $z \in [x, y]$  then we have  $x \leq z \xrightarrow{a \leq x} a \leq z$  and  $z \leq y \xrightarrow{y < b} z < b$  proving that  $z \in [a, b[$ . Hence  $[x, y] \subseteq [a, b[$  from which it follows that

$$[a, b[ \text{ is a generalized interval}$$

3. If  $x, y \in ]a, b]$  and  $z \in [x, y]$  then we have  $x \leq z \xrightarrow{a < x} a < z$  and  $z \leq y \xrightarrow{y \leq b} z \leq b$  proving that  $z \in ]a, b]$ . Hence  $[x, y] \subseteq ]a, b]$  from which it follows that

$$]a, b] \text{ is a generalized interval}$$

4. If  $x, y \in ]a, b[$  and  $z \in [x, y]$  then we have  $x \leq z \xrightarrow{a < x} a < z$  and  $z \leq y \xrightarrow{y < b} z < b$  proving that  $z \in ]a, b[$ . Hence  $[x, y] \subseteq ]a, b[$  from which it follows that

$$]a, b[ \text{ is a generalized interval}$$

5. If  $x, y \in [a, \infty[$  and  $z \in [x, y]$  we have  $x \leq z \xrightarrow{a \leq x} a \leq z$  proving that  $z \in [a, \infty[$  from which it follows that

$$[a, \infty[ \text{ is a generalized interval}$$

6. If  $x, y \in ]a, \infty[$  and  $z \in [x, y]$  we have  $x \leq z \xrightarrow{a < x} a < z$  proving that  $z \in ]a, \infty[$  from which it follows that

$$]a, \infty[ \text{ is a generalized interval}$$

7. If  $x, y \in ]-\infty, a]$  and  $z \in [x, y]$  we have  $z \leq y \xrightarrow{y \leq a} z \leq a$  proving that  $z \in ]-\infty, a]$  from which it follows that

$$]-\infty, a] \text{ is a generalized interval}$$

8. If  $x, y \in ]-\infty, a[$  and  $z \in [x, y]$  we have  $z \leq y \xrightarrow{y < a} z < a$  proving that  $z \in ]-\infty, a[$  from which it follows that

$$]-\infty, a[ \text{ is a generalized interval}$$

□

**Theorem 2.189.** Let  $\langle A, \leq \rangle$  be a totally ordered class then

1.  $\forall a, b \in A$  we have if  $]-\infty, a] = ]-\infty, b]$  then  $a = b$
2.  $\forall a, b \in A$  we have if  $]-\infty, a[ = ]-\infty, b[$  then  $a = b$
3.  $\forall a, b \in A$  we have if  $[a, \infty[ = [b, \infty[$  then  $a = b$
4.  $\forall a, b \in A$  we have if  $]a, \infty[ = ]b, \infty[$  then  $a = b$
5.  $\forall a, b, c, d$  with  $[a, b] \neq \emptyset$  [so that by 2.184  $a \leq b \wedge c \leq d$ ] we have if  $[a, b] = [c, d]$  then  $a = c \wedge b = d$
6.  $\forall a, b, c, d$  with  $]a, b] \neq \emptyset$  [so that by 2.184  $a < b \wedge c < d$ ] we have if  $]a, b] = ]c, d]$  then  $a = c \wedge b = d$
7.  $\forall a, b, c, d$  with  $[a, b[ \neq \emptyset$  [so that by 2.184  $a < b \wedge c < d$ ] we have if  $[a, b[ = [c, d]$  then  $a = c \wedge b = d$
8.  $\forall a, b, c, d$  with  $]a, b[ \neq \emptyset$  [so that by 2.184  $a < b \wedge c < d$ ] we have if  $]a, b[ = ]c, d[$  then  $a = c \wedge b = d$

**Proof.**

1. As  $a \in ]-\infty, a] = ]-\infty, b]$  we have  $a \leq b$ , further from  $b \in ]-\infty, b] = ]-\infty, a]$  we have  $b \leq a$ , hence  $a = b$ .

2. As  $A$  is totally ordered we have for  $a, b$  either

**$a = b$ .** then (2) is proved

**$a < b$ .** then  $a \in ]-\infty, b[ = ]-\infty, a[$  giving the contradiction  $a < a$

**$b < a$ .** then  $b \in ]-\infty, a[ = ]-\infty, b[$  giving the contradiction  $b < b$

so the only valid conclusion is that  $a = b$ .

3. As  $a \in [a, \infty[ = [b, \infty[$  we have  $b \leq a$ , further from  $b \in [b, \infty[ = [a, \infty[$  we have  $a \leq b$ , hence  $a = b$

4. As  $A$  is totally ordered we have for  $a, b$  either

**$a = b$ .** then (4) is proved

**$a < b$ .** then  $b \in ]a, \infty[ = ]b, \infty[$  giving the contradiction  $b < b$

**$b < a$ .** then  $a \in ]b, \infty[ = ]a, \infty[$  giving the contradiction  $a < a$

so the only valid conclusion is that  $a = b$

5. As  $a, b \in [a, b] = [c, d]$  we have that  $c \leq a \wedge b \leq d$ , further from  $c, d \in [c, d] = [a, b]$  it follows that  $a \leq c \wedge d \leq b$ , hence  $a = c \wedge b = d$

6. As  $a < b$  we have  $b \in ]a, b] = ]c, d] \Rightarrow c < b \leq d \Rightarrow c < d$  giving

$$a < b \wedge c < d \Rightarrow d \in ]c, d] \quad (2.28)$$

Further as  $b \in ]a, b] = ]c, d] \Rightarrow b \leq d$  and  $d \in ]c, d] = ]a, b] \Rightarrow d \leq b$  giving

$$b = d \quad (2.29)$$

For  $a, b, c, d$  we have either

**$a = c \wedge b = d$ .** then (6) is proved

**$a = c \wedge b < d$ .** contradicting  $b = d$  [see 2.29].

**$a = c \wedge d < b$ .** contradicting  $b = d$  [see 2.29].

**$a < c \wedge b = d$ .** then  $a < c < d = b \Rightarrow c \in ]a, b] = ]c, d] \Rightarrow c < c$  a contradiction.

**$a < c \wedge b < d$ .** contradicting  $b = d$  [see 2.29].

**$a < c \wedge d < b$ .** contradicting  $b = d$  [see 2.29].

**$c < a \wedge b = d$ .** then  $c < a < b = d \Rightarrow a \in ]c, d] = ]a, b] \Rightarrow a < a$  a contradiction.

**$c < a \wedge b < d$ .** contradicting  $b = d$  [see 2.29].

**$c < a \wedge d < b$ .** contradicting  $b = d$  [see 2.29].

so the only valid case is that  $a = b$  and  $c = d$ .

7. As  $a < b$  we have  $a \in [a, b] = [c, d] \Rightarrow c \leq a < d$  giving

$$a < b \wedge c < d \Rightarrow c \in [c, d] \quad (2.30)$$

Further as  $a \in [a, b] = [c, d] \Rightarrow c \leq a$  and  $c \in [c, d] = [a, b] \Rightarrow a \leq c$  giving

$$a = c \quad (2.31)$$

For  $a, b, c, d$  we have either

$a = c \wedge b = d$ . then (7) is proved.

$a = c \wedge b < d$ . then  $c = a < b < d \Rightarrow b \in [c, d] = [a, b] \Rightarrow b < b$  a contradiction.

$a = c \wedge d < b$ . then  $a = c < d < b \Rightarrow d \in [a, b] = [c, d] \Rightarrow d < d$  a contradiction.

$a < c \wedge b = d$ . contradicting  $a = c$  [see 2.31]

$a < c \wedge b < d$ . contradicting  $a = c$  [see 2.31]

$a < c \wedge d < b$ . contradicting  $a = c$  [see 2.31]

$c < a \wedge b = d$ . contradicting  $a = c$  [see 2.31]

$c < a \wedge b < d$ . contradicting  $a = c$  [see 2.31]

$c < a \wedge d < b$ . contradicting  $a = c$  [see 2.31]

so the only valid case is that  $a = b$  and  $c = d$ .

8. As  $\emptyset \neq ]a, b] = [c, d]$  there exists a  $x \in A$  such that  $x \in ]a, b] = [c, d]$  giving

$$a < x < b \wedge c < x < d \quad (2.32)$$

For  $a, b, c, d$  we have either

$a = c \wedge b = d$ . then (8) is proved

$a = c \wedge b < d$ . then  $c = a < x < b < d \Rightarrow b \in ]c, d] = ]a, b]$  a contradiction.

$a = c \wedge d < b$ . then  $a = c < x < d < b \Rightarrow d \in ]a, b] = ]c, d]$  a contradiction.

$a < c \wedge b = d$ . then  $a < c < x < d = b \Rightarrow c \in ]a, b] = ]c, d]$  a contradiction.

$a < c \wedge b < d$ . then  $c < x < b < d \Rightarrow b \in ]c, d] = ]a, b]$  a contradiction.

$a < c \wedge d < b$ . then  $a < x < d < b \Rightarrow d \in ]a, b] = ]c, d]$  a contradiction.

$c < a \wedge b = d$ . then  $c < a < x < b = d \Rightarrow a \in ]c, d] = ]a, b]$  a contradiction.

$c < a \wedge b < d$ . then  $c < a < x < b < d \Rightarrow a \in ]c, d] = ]a, b]$  a contradiction.

$c < a \wedge d < b$ . then  $c < a < x < d < b \Rightarrow d \in ]c, d] = ]a, b]$  a contradiction.

so the only valid case is that  $a = b$  and  $c = d$ .  $\square$

### 2.3.2 Well-ordered classes

**Definition 2.190.** A partially ordered class  $\langle A, \leq \rangle$  is **well-ordered** if every nonempty subclass of  $A$  has a least element.

**Theorem 2.191.** If  $\langle A, \leq \rangle$  is a partially ordered class and  $B \subseteq A$  then we have the following for  $\langle B, \leq \rangle$  (see 2.143)

1. If  $\langle A, \leq \rangle$  is fully ordered then  $\langle B, \leq \rangle$  is fully ordered
2. If  $\langle A, \leq \rangle$  is well-ordered then  $\langle B, \leq \rangle$  is well-ordered

**Proof.**

1. If  $x, y \in B \Rightarrow x, y \in A \Rightarrow x \leq y \vee y \leq x \Rightarrow x \leq_{|B} y \vee y \leq_{|B} x$
2. If  $C \subseteq B$  is a nonempty set then  $C \subseteq A$  then there exists a least element  $c$  of  $C$ . So  $c \in C$  and  $\forall x \in C \vdash c \leq x \Rightarrow c \in C$  and  $\forall x \in C \vdash c \leq_{|B} x$   $\square$

**Theorem 2.192.** A well-ordered class  $\langle A, \leq \rangle$  is fully ordered

**Proof.** If  $x, y \in A \Rightarrow \{x, y\}$  is a nonempty subclass and thus it has a least element. We have then the following cases

1.  $x$  is the least element then  $x \leq y$
2.  $y$  is the least element then  $y \leq x$

□

**Theorem 2.193.** *A well-ordered class  $\langle A, \leq \rangle$  is conditional complete.*

**Proof.** If  $B$  is a non empty subclass that is bounded above then  $v(B) \neq \emptyset$  and thus  $v(B)$  has a least element which is by definition  $\sup(B)$  □

**Definition 2.194.** *Let  $\langle A, \leq \rangle$  be a partially ordered class and  $x, y \in B$  then  $y$  is a **immediate successor** of  $x$  if  $x < y$  and there is no element  $a \in A$  with  $x < a < y$*

**Theorem 2.195.** *Let  $\langle A, \leq \rangle$  be a well-ordered class then every element in  $A$  (with the exception of the greatest element) has a immediate successor*

**Proof.** If  $x \in A$  and  $x$  is not the greatest element then  $B = \{y \in A \mid y > x\}$  is nonempty and thus by well-ordering there exists a least element  $b \in B$ . Then  $x < b$  and if  $x < a < b$  we would have  $a \in B$  and  $a < b$  now  $b$  is the least element and thus  $b \leq a \Rightarrow a < a \Rightarrow a \neq a$  a contradiction, □

**Definition 2.196.** *Let  $\langle A, \leq \rangle$  be a partially ordered class and  $B \subseteq A$  a subclass then  $B$  is a **section** of  $A$  iff*

$$\forall x \in A \vdash \text{if } y \in B \wedge x \leq y \Rightarrow x \in B$$

**Theorem 2.197.** *Let  $\langle A, \leq \rangle$  be a well-ordered class then  $B \subseteq A$  is a section of  $A$  iff  $A = B$  or  $B$  is a initial segment of  $A$*

**Proof.**

⇒

If  $B$  is a section of  $A$  then if  $B = A$  we are done, so consider  $B \neq A$  then  $A \setminus B \neq \emptyset$ . Because  $A$  is well-ordered  $A \setminus B$  has a least element  $l$  we prove then that  $B = S_l$

$$\begin{aligned} x \in S_l &\Rightarrow x \in A \wedge x < l \\ &\Rightarrow x \in B [\text{if } x \notin B \Rightarrow x \in A \setminus B \Rightarrow l \text{ is not the least element}] \\ x \in B &\Rightarrow \text{if } l \leq x \underset{B \text{ is a section}}{\Rightarrow} l \in B \text{ contradicting } l \in A \setminus B \Rightarrow x < l \Rightarrow x \in S_l \end{aligned}$$

⇐

If  $A = B$  or  $B$  is a initial segment of  $A$  then either

1.  $A = B$  then if  $\forall x \in A \vdash y \in B \wedge x \leq y$  we have  $x \in A$
2.  $\exists b \in A \vdash B = S_b = \{x \in A \mid x < b\}$  so if  $x \in A$  and there exists a  $y \in B$  with  $x \leq y \underset{y < b}{\Rightarrow} x < b \Rightarrow x \in B$  □

**Theorem 2.198. (Transfinite Induction)** *Let  $\langle A, \leq \rangle$  be a well-ordered class and let  $P(x)$  be a statement which is either false or true for a element  $x \in A$  which satisfies*

*If  $P(y)$  is true for every  $y < x$  then  $P(x)$  is true*

*then  $P(x)$  is true for every  $x \in A$*

**Proof.** Suppose  $P(x)$  is not true for every  $x \in A$  then  $B = \{y \in A \mid P(y) \text{ is false}\}$  is not empty. As  $A$  is well-ordered there exists a least element  $b$  of  $B$ . Now if  $x < b$  then if  $x \in B \Rightarrow x < b \leq x \Rightarrow x < x$  a contradiction. So if  $x < b$  then  $x \notin B$  and thus  $P(x)$  is true. By the hypothesis we have then that  $P(b)$  is true which means  $b \notin B$  contradicting the fact that  $b$  is the least element of  $B$ . So we must have that  $P(x)$  is true for every  $x \in A$ . □

**Theorem 2.199.** *If  $\langle A, \leq \rangle$  is a well-ordered class,  $B \subseteq A$  and  $f: A \rightarrow B$  a isomorphism then  $\forall x \in A \vdash x \leq f(x)$*

**Proof.** We prove this by contradiction so assume that there exists a  $x \in A$  such that  $f(x) < x$ . Then we have that  $C = \{x \in A \mid f(x) < x\} \neq \emptyset$  and by well-ordering it has a least element  $c$ . As  $c \in C$  we have  $f(c) < c$  and so using the fact that  $f$  is a isomorphism we have by 2.159 that  $f(f(c)) < f(c)$  so  $f(c) \in C$  but as  $c$  is the least element we have  $f(c) < c \leq f(c) \Rightarrow f(c) < f(c)$  a contradiction. So we must conclude that  $C = \emptyset$  proving our theorem.  $\square$

**Theorem 2.200.** Let  $\langle A, \leq \rangle$  be a well-ordered class then there does not exists a isomorphism from  $A$  to a subclass of an initial segment of  $A$ .

**Proof.**

We prove this again by contradiction. So assume that there exists a  $a \in A$ , a  $B \subseteq S_a$  and a isomorphism  $f: A \rightarrow B$ . Then by the previous theorem we have  $a \leq f(a) \Rightarrow \neg(f(a) < a)$  and thus  $f(a) \notin S_a \Rightarrow f(a) \notin B$  but this contradicts the fact that  $\text{range}(f) = B$ . Hence the theorem must be true.  $\square$

As a initial segment is a subclass of itself so we have the following corollary

**Corollary 2.201.** Let  $\langle A, \leq \rangle$  be a well-ordered class then there does not exists a isomorphism from  $A$  to a initial segment of  $A$ .

**Theorem 2.202.** If  $\langle A, \leq_A \rangle, \langle B, \leq_B \rangle$  be well-ordered classes then if  $A$  is isomorphic with an initial segment of  $B$  we have that  $B$  is not isomorphic with any subclass of  $A$

**Proof.** Let  $b \in B$  and let  $f: A \rightarrow S_b$  be a isomorphism from  $A$  to a initial segment of  $B$ . We proceed now by contradiction, so assume that there exists a  $C \subseteq A$  and a isomorphism  $g: B \rightarrow C$  then  $g: B \rightarrow A$  is a injective increasing function hence  $f \circ g: B \rightarrow S_b$  is a injective increasing function (see 2.49 and 2.155) and thus  $f \circ g: B \rightarrow (f \circ g)(B) \subseteq S_b$  is a injective increasing function that by 2.163, 2.192 is a isomorphism. So we have a isomorphism of  $B$  to a subclass of one of its initial segments which by 2.200 is impossible. Proving the theorem by contradiction.  $\square$

**Lemma 2.203.** Let  $\langle A, \leq \rangle$  be a well-ordered and  $a, b \in A$  with  $a < b$  then  $S_a$  is a initial segment of  $S_b$  (using the induced order relation on  $S_b$ ).

**Proof.** First

$$\begin{aligned} x \in S_a &\Rightarrow x < a \\ &\stackrel{a < b}{\Rightarrow} x < b \\ &\Rightarrow x \in S_b \end{aligned}$$

so we have  $S_a \subseteq S_b$ . Now if  $x \in S_b$  and there is a  $y \in S_a$  such that  $x \leq y \Rightarrow x \leq y < a \Rightarrow x < a \Rightarrow x \in S_a$ . So  $S_a$  is a section of  $S_b$  and as  $a \notin S_a \wedge a \in S_b$  we have  $S_a \neq S_b$  so that using 2.197  $S_a$  is a initial segment of  $S_b$ .  $\square$

**Theorem 2.204.** Let  $\langle A, \leq_A \rangle, \langle B, \leq_B \rangle$  be well-ordered classes then exactly one of the following cases holds

1.  $A$  is isomorphic with  $B$
2.  $A$  is isomorphic with an initial segment of  $B$
3.  $B$  is isomorphic with an initial segment of  $A$

**Proof.** Let's define  $C = \{x \in A \mid \exists y \in B \vdash S_x \cong S_y\}$  we have then the following

$$x \in C \Rightarrow \exists ! y \vdash S_x \cong S_y \text{ (there is only on } y \text{ such that } S_x \cong S_y\text{)}$$

Suppose that given  $x \in C$  there exists  $y, y'$  with  $y \neq y'$  such that  $S_x \cong S_y$  and  $S_x \cong S_{y'}$  giving  $S_y \cong S_{y'}$  and  $S_{y'} \cong S_y$  (see 2.162) then we have as  $B$  is well-ordered and thus fully ordered the following cases

1.  $y \leq_B y' \Rightarrow y <_B y'$  so by the previous lemma we have that  $S_y$  is a initial segment of  $S_{y'}$  and then from  $S_{y'} \cong S_y$  we have that  $S_{y'}$  is isomorphic to a initial segment of itself which is forbidden by 2.201.

2.  $y' \leq_B y \Rightarrow y' <_B y$  so by the previous lemma we have that  $S_{y'}$  is a initial segment of  $S_y$  and then from  $S_y \cong S_{y'}$  we have that  $S_y$  is isomorphic to a initial segment of itself which is forbidden by 2.201.

So we must conclude that  $y = y'$  proving our assessment.

Using the above we can define a function  $F: C \rightarrow B$  where  $F = \{(x, y) | S_x \approx S_y\}$  so if  $x \in C$  then  $S_x \cong S_{F(x)}$ . We prove now that  $F: C \rightarrow D$  where  $D = \text{range}(F) \subseteq B$  is a isomorphism.

1. **(injective)** If  $F(x) = F(x') = d$  then  $S_x \cong S_d$  and  $S_{x'} \cong S_d \Rightarrow S_x \cong S_{x'}$  now by fully ordering we have if  $x \neq x'$

- a.  $x \leq_A x' \Rightarrow x <_A x'$  so  $S_x$  is a initial segment of  $S_{x'}$  which by 2.201 conflicts with  $S_x \cong S_{x'}$
- b.  $x' \leq_A x \Rightarrow x' <_A x$  so  $S_{x'}$  is a initial segment of  $S_x$  which by 2.201 conflicts with  $S_{x'} \cong S_x$  proving that  $x \neq x'$  leads to a contradiction, so we must have  $x = x'$ .

2. **(surjective)** This is trivial as  $D = \text{range}(F)$

3. **(increasing)** Suppose  $x \leq_A y$  then  $S_x \cong S_{F(x)}$  and  $S_y \cong S_{F(y)}$ . Proceed now by contradiction so assume  $\neg(F(x) \leq_B F(y))$  or  $F(y) <_B F(x)$  then  $S_{F(y)}$  is a initial segment of  $S_{F(x)}$  and from  $x \leq_A y$  we have  $S_x \subseteq S_y$  which proves that

- a.  $S_y$  is isomorphic with  $S_{F(y)}$  a initial segment of  $S_{F(x)}$
- b.  $S_{F(x)}$  is isomorphic with  $S_x$  a subclass of  $S_y$

which by 2.202 is forbidden. So we are left with  $F(x) \leq_B F(y)$  and thus  $F$  is increasing.

From 1,2 and 3 and 2.163 we have that  $F: C \rightarrow D$  is a isomorphism.

Next we show that **C is a section of A**. So if  $x \in A$  and there is a  $c \in C$  such that  $x \leq_A c$  then  $S_c \cong S_{F(c)}$  and thus there exists a isomorphism  $g: S_c \rightarrow S_{F(c)}$  we prove now that

$$g|_{S_x}: S_x \rightarrow S_{g(x)} \quad \text{is a isomorphism}$$

First  $x \leq_A c \Rightarrow S_x \subseteq S_c$  and if  $y \in S_x \Rightarrow y <_A x \xrightarrow{g \text{ is a isomorphism}} g(y) <_B g(x) \Rightarrow g(y) \in S_{g(x)}$  and thus  $g|_{S_x}: S_x \rightarrow S_{g(x)}$  is a function. We prove now that it is a isomorphism

- 1. **(injective)** If  $g|_{S_x}(y) = g|_{S_x}(y') \Rightarrow g(y) = g(y') \xrightarrow{g \text{ is a isomorphism}} y = y'$
- 2. **(surjective)** If  $y \in S_{g(x)} \Rightarrow y <_B g(x) \in S_{F(c)} \Rightarrow g(x) <_B F(c) \Rightarrow y <_B F(c) \xrightarrow{g \text{ is surjective}} \exists r \in S_c$  such that  $g(r) = y$ . Now if  $x \leq_A r$  then  $g(x) \leq_B g(r) = y$  contradicting  $y <_B g(x)$  so we must have  $r <_A x \Rightarrow r \in S_x$  and  $y = g(r) = g|_{S_x}(r) \Rightarrow \text{range}(g|_{S_x}) = S_{g(x)}$
- 3. **(increasing)** If  $s, r \in S_x$  and  $s \leq_A r \Rightarrow g|_{S_x}(s) = g(s) \leq_B g(r) = g|_{S_x}(r)$

So by using 2.163 we have that  $g|_{S_x}$  is a isomorphism. Thus  $S_x \cong S_{g(x)}$  and thus  $x \in C$  proving that **C is a section of A**.

Next we prove that **D is a section of B**, so if  $y \in B$  is such that there is a  $d \in D$  with  $y \leq_B d$ . Then as  $d \in D = \text{range}(F)$  there exists a  $c \in C \vdash F(c) = d$  so that  $S_c \cong S_d \Rightarrow S_d \approx S_c$ . So there exists a isomorphism  $g: S_d \rightarrow S_c$ . Now from  $y \leq_B d$  we have that  $S_y \subseteq S_d$  and if  $x \in S_y \Rightarrow x <_B y \xrightarrow{g \text{ is a isomorphism}} g(x) <_A g(y) \Rightarrow g(x) \in S_{g(y)}$  so  $g|_{S_y}: S_y \rightarrow S_{g(y)}$  is a function, we prove now that

$$g|_{S_y}: S_y \rightarrow S_{g(y)} \quad \text{is a isomorphism}$$

- 1. **(injective)** If  $g|_{S_y}(x) = g|_{S_y}(x') \Rightarrow g(x) = g(x') \xrightarrow{g \text{ is injective}} x = x'$
- 2. **(surjective)** If  $z \in S_{g(y)} \Rightarrow z <_A g(y) \in S_c \Rightarrow g(y) <_A c \Rightarrow z <_A c \Rightarrow z \in S_c$  so there exists a  $x \in S_d$  such that  $g(x) = z$ . Now if  $y \leq_B x \Rightarrow g(y) \leq_A g(x) = z \Rightarrow g(y) \leq_A z \Rightarrow z \notin S_{g(y)}$  contradicting  $z \in S_{g(y)}$  so we must have  $x <_B y \Rightarrow x \in S_y$  and  $g|_{S_y}(x) = g(x) = z \Rightarrow \text{range}(g|_{S_y}) = S_{g(y)}$
- 3. **(increasing)** If  $s, r \in S_y$  and  $s \leq_B r \Rightarrow g|_{S_y}(s) = g(s) \leq_A g(r) = g|_{S_y}(r)$

So using 2.163 we have that  $g|_{S_y}: S_y \rightarrow S_{g(y)}$  is a isomorphism. Thus  $S_y \cong S_{g(y)}$  and thus as  $g(y) \in S_c \subseteq A$  we have  $g(y) \in C$  and  $F(g(y)) = y \Rightarrow y \in D$  proving that **D is a section of B**.

We have thus proved that  $F: C \rightarrow D$  is a isomorphism where  $C$  is a section of  $A$  and  $D$  is a section of  $B$ . Using 2.197 we have then the following cases for the isomorphism  $F: C \rightarrow D$

1.  $C = A \wedge D = B$  then  $A$  is isomorphic with  $B$
2.  $C = A \wedge D$  is a initial segment of  $B$  then  $A$  is isomorphic with a initial segment of  $B$
3.  $C$  is a initial segment of  $A \wedge D = B$  then  $B$  is isomorphic with a initial segment of  $B$
4.  $C$  is a initial segment of  $A$  and  $D$  is a initial segment of  $B$ . In this case there exists a  $a \in A$  and a  $b \in B$  so that  $C = S_a$  and  $D = S_b$  and  $S_a \cong S_b$  but then  $a \in C \Rightarrow a \in S_a \Rightarrow a < a$  a contradiction. So this case does not occur.

We conclude that or theorem is proved if we prove that 1,2 and 3 can not occur at the same time but this follows from 2.202.  $\square$

**Corollary 2.205.** *Let  $\langle A, \leq_A \rangle$  be a well-ordered class then every subclass of  $A$  is isomorphic with  $A$  or an initial segment of  $A$*

**Proof.** If  $B \subseteq A$  then  $\langle B, \leq \rangle$  is a well-ordered class and using the previous theorem we have the following exclusive cases

1.  $B$  is isomorphic with  $A$
2.  $B$  is isomorphic with a initial segment of  $A$  then as  $B \cong B$  we use 2.202 to reach a contradiction. So this can not occur.
3.  $B$  is isomorphic with a initial segment of  $A$   $\square$

### 2.3.3 The Axiom of Choice

**Definition 2.206.** *Let  $A$  be a set then a **choice** function for  $A$  is a function  $c: \mathcal{P}'(A) \rightarrow A$  such that  $\forall B \in \mathcal{P}'(A) \vdash c(B) \in B$  (see 1.65)*

We can then state the axiom of choice as follows

**Axiom 2.207. (Axiom of Choice)** *Every set has a choice function.*

Using the Axiom of Choice we can prove the opposite from 2.48

**Theorem 2.208.** *If  $f: A \rightarrow B$  is a surjective function then there exists a function  $g: B \rightarrow A$  such that  $f \circ g = 1_B$*

**Proof.** If  $f: A \rightarrow B$  is surjective. Then  $\forall y \in B$  we have that  $f^{-1}(\{y\})$  is a non empty subset of  $A \Rightarrow f^{-1}(\{y\}) \in \mathcal{P}'(A)$ . By the axiom of choice there exists a choice function  $c: \mathcal{P}'(A) \rightarrow A$ . Define now  $g: B \rightarrow A$  such that  $g(y) = c(f^{-1}(y))$ . Then if  $y \in B$  we have  $g(y) = c(f^{-1}(y)) \in f^{-1}(\{y\}) \Rightarrow f(g(y)) = y$  and thus  $f \circ g = 1_B$ . we prove now that  $g$  is injective. If  $g(y) = g(y')$  then we have  $f(g(y)) = f(g(y')) \Rightarrow 1_B(y) = 1_B(y') \Rightarrow y = y'$ .  $\square$

Using the above theorem and 2.48 we have

**Theorem 2.209.** *A function  $f: A \rightarrow B$  is surjective if and only if there exists a function  $g: B \rightarrow A$  such that  $f \circ g = 1_B$*

We have the following equivalent definitions of the axiom of choice

**Theorem 2.210.** *The following are equivalent*

1. *The Axiom of Choice*
2. *Let  $\mathcal{A}$  be a set of sets such that  $A \in \mathcal{A} \Rightarrow A \neq \emptyset$  and if  $A, B \in \mathcal{A}$  with  $A \neq B$  then  $A \cap B = \emptyset$  (in other words  $\mathcal{A}$  is a set of nonempty mutually disjoint sets) then there exists a set  $C$  (called the choice set for  $\mathcal{A}$ ) such that  $\forall A \in \mathcal{A} \vdash \exists! a \in A$  so that  $a \in C$  and if  $a \in C$  then  $\exists A \in \mathcal{A}$  such that  $a \in A$  (in other words a choice set consists of exactly one element of every element of  $\mathcal{A}$ )*

3. If  $\{A_i\}_{i \in I}$  is a family of nonempty sets then  $\exists c: I \rightarrow \bigcup_{i \in I} A_i$  such that  $\forall i \in I \vdash c(i) \in A_i$  (in other words  $\prod_{i \in I} A_i \neq \emptyset$ )

**Proof.**

1  $\Rightarrow$  2

So let  $\mathcal{A}$  be a set of nonempty mutually disjoint sets and let  $\mathbb{A} = \bigcup_{X \in \mathcal{A}} X$  then we have  $\mathcal{A} \subseteq \mathcal{P}'(\mathbb{A})$ , now by the axiom of choice there is a function  $c: \mathcal{P}'(\mathcal{A}) \rightarrow \mathbb{A}$ . Take then  $C = c(\mathcal{A})$  then if  $A \in \mathcal{A} \subseteq \mathcal{P}'(\mathbb{A}) \Rightarrow c(A) \in A$  and also  $c(A) \in C$  so  $C$  consist of a element of all sets in  $\mathcal{A}$ . If now  $A \in \mathcal{A}$  and  $X \in \mathcal{A} \vdash c(X) \in A \xrightarrow{c(X) \in X} X \cap A \neq \emptyset \Rightarrow X = A$  (as we have mutually disjoint sets).

So if there exists a  $A \in \mathcal{A}$  with  $a, a' \in A$  and  $a, a' \in C = c(\mathcal{A})$  then  $\exists X, X' \in \mathcal{A} \vdash c(X) = a \in A$ ,  $c(X') = a' \in A \Rightarrow X = A = X' \xrightarrow{X = X' = A \wedge (X, a) \in c, (X', a') \in c} (A, a), (A, a') \in c \xrightarrow{c \text{ is a function}} a = a'$  proving that  $C$  is indeed the choice set for  $\mathcal{A}$

2  $\Rightarrow$  1

Let  $A$  be a set and let  $B \subseteq A$  be a nonempty subset of  $A$ . Form then  $P_B = \{(B, x) | x \in B\}$ , Then we have that  $P_B$  is nonempty as  $B$  is nonempty. Also if  $z \in P_B \cap P_{B'} \Rightarrow z = (B, x) \wedge x \in B \wedge z = (B', x') \wedge x' \in B' \Rightarrow B = B' \Rightarrow P_B = P_{B'}$ . So if  $P_B \neq P_{B'}$  then  $P_B \cap P_{B'} = \emptyset$ . From this it follows that  $\mathcal{A} = \{P_B | B \subseteq A \wedge B \neq \emptyset\}$  is set of nonempty mutually disjoint sets (note that  $P_B \subseteq \mathcal{P}(A) \times A$  and thus  $\mathcal{A} \subseteq \mathcal{P}(\mathcal{P}(A) \times A)$  is indeed a set). From the hypothesis we have then the existence of a choice set  $C$ . If  $B \in \mathcal{P}'(A)$  then  $P_B \in \mathcal{A} \Rightarrow \exists (B, b) \in P_B \vdash (B, b) \in C$ . If  $c \in C$  then  $\exists B \in \mathcal{P}'(A) \vdash c \in P_B \Rightarrow \exists b \in B \subseteq A \vdash (B, b) = c \Rightarrow C \subseteq \mathcal{P}'(A) \times A$ . Also if  $(B, b) \in C \Rightarrow \exists P_B \vdash (B, b) \in P_B \Rightarrow B = B' \wedge b \in B' \Rightarrow b \in B$ . So if  $(B, b), (B, b') \in C \Rightarrow b, b' \in B \Rightarrow (B, b), (B, b') \in P_B \xrightarrow{C \text{ is a choice set}} (B, b) = (B, b') \Rightarrow b = b'$ . So we have proved that  $C: \mathcal{P}'(A) \rightarrow A$  is a function such that  $\forall B \in \mathcal{P}'(A) \vdash C(B) \in B$ . So  $C$  is the choice function we would search for.

1  $\Rightarrow$  3

If  $\{A_i\}_{i \in I}$  is a family of nonempty sets, form then  $A = \bigcup_{i \in I} A_i$  by the axiom of choice there exists a  $c: \mathcal{P}'(A) \rightarrow A$  such that  $\forall B \in \mathcal{P}'(A)$  we have  $c(B) \in B$ . Define then the function  $x: I \rightarrow \mathcal{A}$  with  $x = \{(i, y) | i \in I \wedge y = c(A_i)\} \subseteq I \times A$  [then  $\text{dom}(x) = I$  and  $(i, y), (i, y') \in x \Rightarrow y = c(A_i) = y' \Rightarrow y = y'$  so  $x$  is a function and  $\forall i \in I \vdash x(i) = c(A_i) \in A_i$ ] and thus  $x \in \prod_{i \in I} A_i$

3  $\Rightarrow$  1

If  $A$  is a set define then the graph  $\mathcal{B} = \{(B, b) | B \in \mathcal{P}'(A) \wedge b \in B\}$ . We have then  $\text{dom}(\mathcal{B}) \subseteq \mathcal{P}'(A)$  and if  $B \in \mathcal{P}'(A)$  there exists a  $b \in B$  and thus  $(B, b) \in \mathcal{B}$  proving that  $\text{dom}(\mathcal{B}) = \mathcal{P}'(A)$ . This defines then the family  $\{\mathcal{B}_B\}_{B \in \mathcal{P}'(A)}$ . Now given  $B \in \mathcal{P}'(A)$  we have

$$\begin{aligned} x \in B &\Leftrightarrow B \in \mathcal{P}'(A) \wedge x \in B \\ &\Leftrightarrow (B, x) \in \mathcal{B} \\ &\Leftrightarrow x \in \{y | (B, y) \in \mathcal{B}\} \\ &\Leftrightarrow x \in \mathcal{B}_B \end{aligned}$$

proving that  $\mathcal{B}_B = B$ . So  $\bigcup_{B \in \mathcal{P}'(A)} \mathcal{B}_B \subseteq A$  and as  $x \in A \Rightarrow \{x\} \in \mathcal{P}'(A) \Rightarrow x \in \{x\} = \mathcal{B}_{\{x\}} \Rightarrow x \in \bigcup_{B \in \mathcal{P}'(A)} \mathcal{B}_B$  so  $A = \bigcup_{B \in \mathcal{P}'(A)} \mathcal{B}_B$  and by the hypothesis there exists a  $c: \mathcal{P}'(A) \rightarrow A$  with  $\forall B \in \mathcal{P}'(A) \vdash c(B) \in B$  and  $c$  is the sought for choice function.  $\square$

**Corollary 2.211.** Let  $A$  and  $B$  be sets and assume that  $\forall x \in A$  there exists a  $y \in B$  such that  $P(x, y)$  is true then there exists a function  $f: A \rightarrow B$  such that  $\forall x \in A$  we have  $P(f(x))$  is true

**Proof.** Define the graph  $G = \{(x, y) \in A \times B | P(x, y)\}$  then this defines (see 1.77) a family  $\{F_a\}_{a \in A}$  with  $F_a = \{y \in |P(a, y)| \neq \emptyset$  so using the equivalences of the Axiom of Choice (see 2.210) there exists a function  $f: A \rightarrow \bigcup_{a \in A} F_a$  with  $\forall a \in A$  that  $f(a) \in F_a \Rightarrow P(a, f(a))$  is true.  $\square$

As a direct application of the Axiom of Choice we prove the following

**Theorem 2.212.** Let  $\langle A, \leq \rangle$  be a partially ordered set such that

1.  $A$  has a least element  $p$
2. Every chain of  $A$  has a supremum

then there is a element  $x \in A$  which has no immediate successor

**Proof.** We will prove this by contradiction. So assume that  $\forall x \in A$  there exists a immediate successor, we will then derive a contradiction from this.

First  $\forall x \in A$  define  $T_x = \{y: y \text{ is a immediate successor of } x\}$  then  $T_x \neq \emptyset \in \mathcal{P}'(A)$ . By the axiom of choice there exists a choice function from  $c: \mathcal{P}'(A) \rightarrow A$  such that  $\forall A \in \mathcal{P}'(A) \models c(A) \in A$ . This lets us define a function  $f: A \rightarrow A$  where  $f = \{(x, y): x \in A \wedge y = c(T_x)\} \subseteq A \times A$  [if  $x \in A \Rightarrow T_x \neq \emptyset \Rightarrow (x, c(T_x)) \in f \Rightarrow \text{dom}(f) = A$  and if  $(x, y), (x, y') \in f \Rightarrow y = c(T_x) = y' \Rightarrow y = y'$  thus  $f: A \rightarrow A$  is indeed a function] and if  $x \in A \Rightarrow f(x) \in T_x$  so  $f(x)$  is a immediate successor of  $x$ . So we have defined a function

$$f: A \rightarrow A \quad \forall x \in A \models f(x) \text{ is a immediate successor of } x$$

Given the least element  $p$  of  $A$  we define the concept of a **p-sequence** as follows

**Definition 2.213.** A subset  $B \subseteq A$  is called a **p-sequence** iff

1.  $p \in B$
2. If  $x \in B \Rightarrow f(x) \in B$
3. If  $C$  is a chain of  $B$  (and thus of  $A$ ) then  $\sup(C) \in B$

Note that  $A$  is trivially a p-sequence. We have now the following lemma

**Lemma 2.214.** Every intersection of p-sequences is a p-sequence

**Proof.** If  $\mathcal{A}$  is a collection of p-sequences then

1.  $\forall A \in \mathcal{A} \models p \in A \Rightarrow p \in \bigcap_{A \in \mathcal{A}} A$
2. If  $x \in \bigcap_{A \in \mathcal{A}} A \Rightarrow \forall A \in \mathcal{A} \models x \in A \Rightarrow \forall A \in \mathcal{A} \models f(x) \in A \Rightarrow f(x) \in \bigcap_{A \in \mathcal{A}} A$
3. If  $C$  is a chain in  $\bigcap_{A \in \mathcal{A}} A$  then every element in  $C$  is comparable and  $C \subseteq \bigcap_{A \in \mathcal{A}} A \Rightarrow \forall A \in \mathcal{A} \models C \subseteq A \Rightarrow \forall A \in \mathcal{A} \models C$  is a chain in  $A \Rightarrow \forall A \in \mathcal{A} \models \sup(C) \in A \Rightarrow \sup(C) \in \bigcap_{A \in \mathcal{A}} A$

So we have that  $\bigcap_{A \in \mathcal{A}} A$  is a p-sequence □

Take now  $\mathcal{A} = \{B \in \mathcal{P}(A): B \text{ is a } p\text{-sequence}\}$  which is nonempty as it contains  $A$  and define the intersection of all p-sequences

**Theorem 2.215.** Let  $\{A_i\}_{i \in I}$  be a family of sets and  $b: J \rightarrow I$  a bijection then

$$\begin{aligned} \bigcap_{i \in I} A_i &= \bigcap_{j \in J} A_{b(j)} \\ \bigcup_{i \in I} A_i &= \bigcup_{j \in J} A_{b(j)} \end{aligned}$$

**Proof.** First we prove that □

$$P = \bigcap_{B \in \mathcal{A}} B$$

then  $P$  is a p-sequence and (as  $p \in P$  we have  $P \neq \emptyset$ ) and if  $B$  is a p-sequence then  $P \subseteq B$ . Next we define the concept of **select** elements in  $P$ .

**Definition 2.216.** A element  $x \in P$  is called **select** if it is comparable with every element  $y$  of  $P$

We have then the following lemma

**Lemma 2.217.** Suppose  $x$  is **select** then if  $y \in P$  with  $y < x$  we have  $f(y) \leq x$

**Proof.** If  $y \in P$  then as  $P$  is a p-sequence we have by (2) that  $f(y) \in P$ . Now  $x$  is select so we have either  $f(y) \leq x$  or  $x < f(y)$ . If  $x < f(y)$  then from  $y < x$  we have  $y < x < f(y)$  contradicting that  $f(y)$  is the immediate successor of  $y$ , so we must have  $f(y) \leq x$   $\square$

Next we prove the following lemma

**Lemma 2.218.** Suppose  $x$  is select then  $B_x = \{y \in P \mid y \leq x \text{ or } f(x) \leq y\}$  is a p-sequence.

**Proof.**

1. Since  $p$  is the least element of  $A$  we have  $p \leq x \Rightarrow p \in B_x$
  2. Suppose  $y \in B_x$  then  $y \leq x$  or  $y \geq f(x)$  lets break this in the following three cases
    - a.  $y < x$ . Then by the previous lemma we have  $f(y) \leq x \Rightarrow f(y) \in B_x$
    - b.  $y = x$ . Then  $f(y) = f(x)$  thus  $f(x) \leq f(y) \Rightarrow f(y) \in B_x$
    - c.  $y \geq f(x)$  Then as  $f(y)$  is the immediate successor of  $y$  we have  $y < f(y) \Rightarrow f(x) < f(y) \Rightarrow f(x) \leq f(y) \Rightarrow f(y) \in B_x$
  3. If  $C$  is a chain in  $B_x$  (and thus in  $A$ ) let  $s = \sup(C)$  then we have the following excluding cases
    - a.  $\exists y \in C \vdash f(x) \leq y \nRightarrow f(x) \leq s \Rightarrow s \in B_x$
    - b.  $\forall y \in C \models \neg(f(x) \leq y)$ . Now  $\forall y \in C$  as  $y \in C \subseteq B_x$  we have  $y \leq x$  or  $f(x) \leq y$ , as the last contradicts  $\neg(f(x) \leq y)$ , we have always  $y \leq x$  and thus  $x$  is a upper bound of  $C$ . As  $s$  is the least upper bound we must have  $s \leq x \Rightarrow s \in B_x$
- So in all cases we have  $f(y) \in B_x$
- So in all valid cases  $s \in B_x$   $\square$

We can easily derive the following corollary.

**Corollary 2.219.** If  $x$  is select then  $\forall y \in P$  we have  $y \leq x$  or  $f(x) \leq y$

**Proof.** As  $P \subseteq B_x$  ( $P$  is the intersection of all the p-sequences) and by definition of  $B_x$  we have  $B_x \subseteq P \Rightarrow P = B_x$   $\square$

We can use this to prove the following lemma

**Lemma 2.220.** The set of all select elements  $\{x \in P \mid x \text{ is select}\}$  is a p-sequence

**Proof.**

1.  $p$  is select because  $\forall x \in P \subseteq A$  we have  $p \leq x$  so it is comparable with every element of  $p$
2. If  $x$  is select then, by the previous corollary,  $\forall y \in P$  we have either
  - a.  $y \leq x$ . Then as  $x < f(x) \Rightarrow y < f(x) \Rightarrow y \leq f(x)$
  - b.  $f(x) \leq y$

So  $f(x)$  is comparable with every element of  $P$  and  $f(x)$  is thus select.

3. Let  $C$  be a chain of select elements. Then as  $C$  is a chain in  $A$  we have that  $s = \sup(C)$  exists. Then  $\forall y \in P$  we have the following possibilities for  $C$ 
  - a.  $\exists x \in C \vdash y \leq x$  then as  $x \leq s$  we have  $y \leq s$
  - b.  $\forall x \in C \models \neg(y \leq x)$  then as  $y \in P$  we have  $\forall x \in C$  either  $y \leq x$  (contradicted by  $\neg(y \leq x)$ ) or  $f(x) \leq y$  so  $\forall x \in C$  we have  $f(x) \leq y \nRightarrow x < f(x) \leq y$  and thus  $y$  is a upper bound of  $C$ . Ans as  $s$  is the least upper bound of  $S$  we have  $s \leq y$

So  $s$  is comparable with every  $y \in P$  and thus  $s$  is select.  $\square$

**Corollary 2.221.**  $P$  is fully-ordered

**Proof.**  $S = \{x \in P \mid x \text{ is select}\} \subseteq P$  and by the previous lemma we have that  $S$  is a p-sequence thus  $P \subseteq S$  so  $P = S = \{x \in P \mid x \text{ is select}\}$ . Or every element in  $P$  is select and thus comparable with every other element so  $P$  is fully-ordered.  $\square$

Now for the final proof, we show that this corollary causes a contradiction. As  $P$  is by the corollary a chain we have that  $s = \sup(P)$  exists by the hypothesis. But as  $P$  is a p-sequence we find by (2) that  $f(s) \in P$  hence  $f(s) \leq s$  but as  $f(s)$  is the successor of  $s$  we have  $s < f(s) \Rightarrow s < s$  a contradiction. So we have derived a contradiction and the theorem must be true.  $\square$

We show now that the following theorem follows from the Axiom of Choice (via 2.212)

**Theorem 2.222. (Hausdorff's Maximal Principle)** *Every partially ordered set  $\langle A, \leq \rangle$  has a maximal chain  $C$  (see 2.165 if  $D$  is a chain and if  $C \subseteq D \Rightarrow C = D$ )*

**Proof.** Define  $\mathcal{C} = \{B \in \mathcal{P}(A) \mid B \text{ is a chain}\}$  be the set [is  $\subseteq \mathcal{P}(A)$  and applying 1.63 and 1.58 gives  $\mathcal{C}$  is a set] of all chains in  $A$  and order  $\mathcal{C}$  by inclusion (see 2.137) so  $\langle \mathcal{C}, \subseteq \rangle$  is a partially ordered set. Notice that  $\mathcal{C}$  has a least element namely  $\emptyset$ . Let now  $\mathcal{D} \subseteq \mathcal{C}$  be a chain in  $\mathcal{C}$  and define

$$K = \bigcup_{C \in \mathcal{D}} C$$

then we prove that  $K$  is a chain in  $A$  and thus  $K \in \mathcal{C}$ . Indeed if  $x, y \in K$  then  $\exists C_1, C_2 \in \mathcal{D}$  such that  $x \in C_1$  and  $y \in C_2$ . Now as  $\mathcal{D}$  is a chain in  $\mathcal{C}$  we have either

1.  $C_1 \subseteq C_2 \Rightarrow x, y \in C_2$   $C_2$  is a chain in  $A$   $x$  and  $y$  are comparable
2.  $C_2 \subseteq C_1 \Rightarrow x, y \in C_1$   $C_1$  is a chain in  $A$   $x$  and  $y$  are comparable

**Theorem 2.223.** *Let  $\{A_i\}_{i \in I}$  be a family of sets and  $b: J \rightarrow I$  a bijection then*

$$\begin{aligned} \bigcap_{i \in I} A_i &= \bigcap_{j \in J} A_{b(j)} \\ \bigcup_{i \in I} A_i &= \bigcup_{j \in J} A_{b(j)} \end{aligned}$$

**Proof.** First we prove that

$\square$

proving that  $K$  is a chain. Using 2.172 we have that  $K = \sup(\mathcal{D})$  exists. So we can apply 2.212 (which is derived from the Axiom of Choice) to find an element  $C \in \mathcal{C}$  which has no immediate successor.  $C$  is (by definition of  $C$ ) a chain. We prove now by contradiction that  $C$  is maximal. If there exists a chain  $D$  in  $A$  with  $C \subseteq D$  and  $D \neq C$  then  $\exists d \in D \setminus C$ . Then  $C \cup \{d\}$  is a chain for if  $x, y \in C \cup \{d\} \subseteq D$  we have as  $D$  is a chain that  $x, y$  are comparable. Now if there is a set  $X$  with  $C \subset X \subset C \cup \{d\}$  then  $X = X \cup (X \setminus C) \subseteq C \cup \{d\} \Rightarrow (X \setminus C) \subseteq C \cup \{d\}$  and if  $x \in (X \setminus C) \Rightarrow x \notin C \wedge x \in C \cup \{d\} \Rightarrow x = d$  so  $(X \setminus C) = \{d\}$  so  $X = C \cup \{d\}$  contradicting  $X \subset C \cup \{d\}$ . So for every chain  $D$  in  $A$  with  $C \subseteq D$  we must have  $C = D$  proving maximality of  $C$ .  $\square$

**Definition 2.224.** *A partially ordered set  $\langle A, \leq \rangle$  is said to be **inductive** when every chain in  $A$  has an upper bound.*

**Lemma 2.225. (Zorn's Lemma)** *Every inductive set has at least one maximal element. In other words, if  $\langle A, \leq \rangle$  is a partially ordered set such that every chain in  $A$  has an upper bound, then  $A$  has a maximal element.*

We prove now that from the Hausdorff's maximal theorem follows directly Zorn's lemma, without using either 2.210 or 2.212, so it does not depend indirectly on the Axiom of Choice, but directly on Hausdorff's maximal theorem.

**Theorem 2.226.** *Hausdorff's maximal theorem implies Zorn's lemma*

**Proof.** Let  $\langle A, \leq \rangle$  be a partially ordered set such that every chain in  $A$  has an upper bound. By the Hausdorff's Maximal Principle (see 2.222) there exists a maximal chain  $C$  in  $A$ , by the hypothesis there exists an upper bound  $u$  for  $C$ . We prove now by contradiction that  $u$  is a maximal element of  $A$ . So assume that  $u$  is not a maximal element of  $A$  then there exists a  $x \in A \vdash u < x$ . If now  $x \in C \xrightarrow{u \text{ is an upper bound}} x \leq u < x \Rightarrow x < x$  a contradiction so we must have that  $x \notin C$ . Now if  $r, s \in C \cup \{x\}$  then we have the following cases

1.  $r = s = x \xrightarrow{\text{reflexivity}} r \leq s$
2.  $r = x, s \in C \Rightarrow s \leq u < x = r \Rightarrow s \leq r$
3.  $r \in C, s = x \Rightarrow r \leq u < x = s \Rightarrow r \leq s$
4.  $r, s \in C \xrightarrow{C \text{ is a chain}} s \leq r \text{ or } r \leq s$

proving that  $C \cup \{x\}$  is a chain with (because  $x \notin C$ )  $C \subset C \cup \{x\}$  contradicting that  $C$  is a maximal chain. So we must have that  $u$  is a maximal element of  $A$ .  $\square$

We state now the Well-Ordering Theorem

**Theorem 2.227. (Well-Ordering Theorem)** *Given a set  $A$  then there exists a order relation  $\leq$  on  $A$  such that  $\langle A, \leq \rangle$  forms a well-ordered set.*

We prove then that from the Zorn's lemma follows directly the Well-Ordering Theorem without using either 2.210, 2.222 or 2.212 so it depends directly on Zorn's Lemma.

**Theorem 2.228.** *Zorn's Lemma implies the Well-Ordering Theorem*

**Proof.** Let  $A$  be an arbitrary set and let  $B \subseteq A$  a subset. A relation  $G$  on  $B$  is a subclass of the set  $B \times B$  (see 1.72) and we have by 1.58 that  $G$  is a set. So the pair  $(B, G)$  is an element (see 1.36) we define then the class  $\mathcal{A} = \{(B, G) | B \in \mathcal{P}(A) \wedge G \text{ is a order relation on } B \text{ so that } \langle B, G \rangle \text{ is well-ordered}\}$ . We define now  $\prec \subseteq \mathcal{A} \times \mathcal{A}$  by

$$(B, G) \prec (B', G') \Leftrightarrow B \subseteq B' \text{ and } G \subseteq G' \text{ and } [x \in B \wedge y \in B' \setminus B \Rightarrow (x, y) \in G'] \quad (2.33)$$

**Lemma 2.229.**  $\langle \mathcal{A}, \prec \rangle$  forms a order relation

**Proof.**

1. **(reflectivity)** If  $(B, G) \in \mathcal{A}$  then
  - a.  $B \subseteq B$
  - b.  $G \subseteq G$
  - c.  $x \in B \wedge y \in B \setminus B = \emptyset \Rightarrow (x, y) \in G$  is satisfied because  $y \in \emptyset$  can not occur so  $x \in B \wedge y \in B \setminus B$  is always false.

proving  $(B, G) \prec (B, G)$
2. **(anti-symmetry)** If  $(B, G) \prec (B', G') \wedge (B', G') \prec (B, G) \Rightarrow B \subseteq B' \subseteq B \wedge G \subseteq G' \subseteq G \Rightarrow B = B' \wedge G = G' \Rightarrow (B, G) = (B', G')$  proving anti-symmetry.
3. **(transitivity)** If  $(B, G) \prec (B', G') \wedge (B', G') \prec (B'', G'')$  then we have
  - a.  $B \subseteq B' \wedge B' \subseteq B'' \Rightarrow B \subseteq B''$
  - b.  $G \subseteq G' \wedge G' \subseteq G'' \Rightarrow G \subseteq G''$
  - c. If  $x \in B \wedge y \in B'' \setminus B$  we have then for  $y$  the following cases
    - i.  $y \in B' \Rightarrow y \in B' \setminus B \Rightarrow (x, y) \in G' \xrightarrow{G' \subseteq G''} (x, y) \in G''$

ii.  $y \notin B' \Rightarrow y \in B'' \setminus B' \Rightarrow (x, y) \in G''$

So in all cases we have  $(x, y) \in G''$

proving  $(B, G) \prec (B'', G'')$  □

We now have the following lemma.

**Lemma 2.230.** *If  $\mathcal{C} \subseteq \mathcal{A}$  is a chain in  $\langle \mathcal{A}, \prec \rangle$  then if*

$$\begin{aligned} B_{\mathcal{C}} &= \bigcup_{(B, G) \in \mathcal{C}} B \\ G_{\mathcal{C}} &= \bigcup_{(B, G) \in \mathcal{C}} G \end{aligned}$$

we have  $(B_{\mathcal{C}}, G_{\mathcal{C}}) \in \mathcal{A}$

**Proof.**

1.  $\forall (B, G) \in \mathcal{C}$  we have  $B \subseteq A$  so if  $x \in B_{\mathcal{C}}$  then  $\exists (B, G) \in \mathcal{C} \vdash x \in B \subseteq A \Rightarrow x \in A$  and thus  $B_{\mathcal{C}} \subseteq A$
2. We must prove now that  $G_{\mathcal{C}}$  is a order relation on  $B_{\mathcal{C}}$ 
  - a. **(reflectivity)** If  $x \in B_{\mathcal{C}} \Rightarrow \exists (B, G) \in \mathcal{C} \vdash x \in B \Rightarrow (x, x) \in G \Rightarrow (x, x) \in G_{\mathcal{C}}$
  - b. **(anti-symmetry)** If  $(x, y) \in G_{\mathcal{C}} \wedge (y, x) \in G_{\mathcal{C}}$  then  $\exists (B, G), (B', G') \in \mathcal{C} \vdash (x, y) \in G \wedge (y, x) \in G'$  now because  $\mathcal{C}$  is a chain we have either
    - i.  $(B, G) \prec (B', G')$  then  $G \subseteq G' \Rightarrow (x, y), (y, x) \in G' \Rightarrow x = y$
    - ii.  $(B', G') \prec (B, G)$  then  $G' \subseteq G \Rightarrow (x, y), (y, x) \in G \Rightarrow x = y$
  - c. **(transitivity)** If  $(x, y), (y, z) \in G_{\mathcal{C}}$  then  $\exists (B, G), (B', G') \in \mathcal{C} \vdash (x, y) \in G \wedge (y, z) \in G'$ . Now because  $\mathcal{C}$  is a chain we have either
    - i.  $(B, G) \prec (B', G') \Rightarrow G \subseteq G' \Rightarrow (x, y), (y, z) \in G' \Rightarrow (x, z) \in G' \Rightarrow (x, z) \in G_{\mathcal{C}}$
    - ii.  $(B', G') \prec (B, G) \Rightarrow G' \subseteq G \Rightarrow (x, y), (y, z) \in G \Rightarrow (x, z) \in G \Rightarrow (x, z) \in G_{\mathcal{C}}$
3. Now we must prove well-ordering. Suppose that  $D \subseteq B_{\mathcal{C}}$  and  $D \neq \emptyset$  then there exists a  $x \in D \Rightarrow x \in B_{\mathcal{C}} \Rightarrow \exists (B, G) \in \mathcal{C} \vdash x \in B \Rightarrow D \cap B \neq \emptyset$ . Now  $D \cap B \subseteq B$  hence by well-ordering of  $\langle B, G \rangle$  we have that there exists a least element  $b \in B$  in  $B \Rightarrow \forall y \in B \models (b, y) \in G \Rightarrow \forall y \in B \cap D \models (b, y) \in G$ . We prove now that  $b$  is the least element of  $D$ .

**Proof.** If  $x \in D \Rightarrow \exists (B', G')$  with  $x \in B'$  then we have the following cases

- a.  $x \in B \Rightarrow x \in B \cap D \Rightarrow (b, x) \in G \Rightarrow (b, x) \in G_{\mathcal{C}}$
- b.  $x \notin B \Rightarrow x \in B' \setminus B \wedge b \in B$  now we have two cases
  - i.  $(B, G) \prec (B', G') \underset{b \in B \wedge x \in B' \setminus B}{\Rightarrow} (b, x) \in B' \Rightarrow (b, x) \in G_{\mathcal{C}}$
  - ii.  $(B', G') \prec (B, G) \Rightarrow B' \subseteq B \Rightarrow x \in B$  contradicting  $x \notin B$  so this case does not apply.

So in all cases that are valid we have  $(b, x) \in G_{\mathcal{C}}$  proving that  $b$  is the least element of  $D$  and thus proving that  $\langle B_{\mathcal{C}}, G_{\mathcal{C}} \rangle$  is a well-ordered set. □ □

Next we have the finally lemma before the Well-Ordering theorem

**Lemma 2.231.** *If  $\mathcal{C}$  is a chain in  $\langle \mathcal{A}, \prec \rangle$  then  $(B_{\mathcal{C}}, G_{\mathcal{C}})$  is a upper bound of  $\mathcal{C}$*

**Proof.** If  $(B, G) \in \mathcal{C}$  then clearly  $B \subseteq B_{\mathcal{C}}$  and  $G \subseteq G_{\mathcal{C}}$  so to prove that  $(B, G) \prec (B_{\mathcal{C}}, G_{\mathcal{C}})$  we must prove that if  $x \in B$  and  $y \in B_{\mathcal{C}} \setminus B$  that  $(x, y) \in G_{\mathcal{C}}$ . So assume that  $x \in B$  and  $y \in B_{\mathcal{C}} \setminus B$  then  $\exists (B', G') \vdash y \in B'$ . We have then as  $y \notin B$  that  $B' \not\subseteq B$  hence as  $\mathcal{C}$  is a chain and we can't have  $(B', G') \prec (B, G)$  we must have  $(B, G) \prec (B', G')$  and thus as  $B \subseteq B'$  we have  $y \in B' \setminus B \Rightarrow (x, y) \in G' \Rightarrow (x, y) \in B_{\mathcal{C}}$  proving that  $(B, G) \prec (B_{\mathcal{C}}, G_{\mathcal{C}})$ . □

Finally we can prove our well-ordering theorem. Using Zorn's lemma (2.225) we have a maximal element  $(B_m, G_m) \in \mathcal{A}$ . We prove now that  $B_m = A$  proving that  $(A, G_m) \in \mathcal{A}$  and that  $\langle A, G_m \rangle$  is well-ordered.

**Proof.** Assume that  $B_m \neq A$  then as  $B_m \subseteq A$  we have  $\exists x \in A \setminus B_m$  define then  $G^* = G_m \cup \{(b, x) \mid b \in B_m\} \cup \{(x, x)\}$  (a disjoint union), we prove then that  $\langle B_m \cup \{x\}, G^* \rangle$  is a well-ordered set

**Proof.** Note that if  $(x, r) \in G^*$  then we have the following possibilities

1.  $(x, r) \in G_m$  which is impossible because  $G_m \subseteq B_m \times B_m$  and  $x \notin B_m$
2.  $(x, r) \in \{(b, x) \mid b \in B_m\}$  which is impossible as  $x \notin B_m$
3.  $(x, r) \in \{(x, x)\} \Rightarrow r = x$

So from  $(x, r) \in G^*$  it follows that  $(x, r) = (x, x) \in \{(x, x)\}$  we now prove that  $G^*$  is a well-ordering.

1. **(reflexive)** If  $y \in B_m \cup \{x\}$  then we have the following cases

- a.  $y \in B_m \Rightarrow (y, y) \in G_m \Rightarrow (y, y) \in G^*$
- b.  $y \in \{x\} \Rightarrow (y, y) = (x, x) \in G^*$

2. **(anti-symmetry)** If  $(r, s), (s, r) \in G^*$  we have either

- a.  $(r, s) \in G_m \underset{G_m \subseteq B_m \times B_m}{\Rightarrow} r, s \in B_m \underset{r, s \neq x \wedge (s, r) \in G^*}{\Rightarrow} (s, r) \in G_m \Rightarrow r = s$
- b.  $(r, s) \in \{(b, x) \mid b \in B_m\} \Rightarrow s = x \underset{(x, r) \in G^*}{\Rightarrow} r = x \Rightarrow r = s$
- c.  $(r, s) \in \{(x, x)\} \Rightarrow r = x = s$

3. **(transitive)** If  $(r, s), (s, t) \in G^*$  then we have the following cases

- a.  $(r, s) \in G_m$  we have then the following sub cases
  - i.  $(s, t) \in G_m \Rightarrow (r, t) \in G_m \Rightarrow (r, t) \in G^*$
  - ii.  $(s, t) = (b, x) \Rightarrow t = x \Rightarrow (r, t) = (r, x) \in G^* \Rightarrow (r, t) \in G^*$
  - iii.  $(s, t) = (x, x) \Rightarrow t = x \Rightarrow (r, t) = (r, x) \in G^* \Rightarrow (r, t) \in G^*$
- b.  $(r, s) \in \{(b, x) \mid b \in B_m\} \Rightarrow s = x \Rightarrow (x, t) \in G^* \Rightarrow t = x \Rightarrow (r, t) \in \{(b, x) \mid b \in B_m\} \Rightarrow (r, t) \in G^*$
- c.  $(r, s) \in \{(x, x)\} \Rightarrow r = s = x \Rightarrow (x, t) \in G^* \Rightarrow t = x \Rightarrow (r, t) = (x, x) \in G^*$

so  $G^*$  is indeed a order relation. Now if  $C \subseteq B_m \cup \{x\}$  is a nonempty set then we have the following possibilities

1.  $C \cap B_m \neq \emptyset \Rightarrow \emptyset \neq C \cap B_m \subseteq B_m \underset{G_m \text{ is well-ordered}}{\Rightarrow} \exists l \in C \cap B_m \vdash \forall y \in C \cap B_m \models (l, y) \in G_m$ .

Now if  $y \in C$  then we have the following possibilities

- a.  $y \in B_m \Rightarrow y \in C \cap B_m \Rightarrow (l, y) \in G_m \Rightarrow (l, y) \in G^*$
- b.  $y \notin B_m \Rightarrow y = x \Rightarrow (l, y) \in \{(b, x) \mid b \in B_m\} \Rightarrow (l, y) \in G^*$

So we conclude that  $C$  has a least element.

2.  $C \cap B_m = \emptyset \Rightarrow C = \{x\} \Rightarrow x$  is the least element of  $C$ .

So we have that  $(B_m \cup \{x\}, G^*) \in \mathcal{A}$

Next we have  $(B_m, G_m) \prec (B_m \cup \{x\}, G^*)$  as

1.  $B_m \subseteq B_m \cup \{x\}$
2.  $G_m \subseteq G^*$
3. If  $r \in B_m$  and  $s \in (B_m \cup \{x\}) \setminus B_m \Rightarrow s = x \Rightarrow (r, s) \in \{(b, x) \mid b \in B_m\} \Rightarrow (r, s) \in G^*$

But as  $x \notin B_m$  we have  $B_m \neq B_m \cup \{x\}$  and thus  $(B_m, G_m) \neq (B_m \cup \{x\}, G^*)$  so we have a contradiction as by maximality of  $(B_m, G_m)$  in  $\langle \mathcal{A}, \prec \rangle$  we would have  $(B_m, G_m) = (B_m \cup \{x\}, G^*)$ . So we must conclude that  $B_m = A$   $\square$

So finally  $(A, G_m) \in \mathcal{A} \Rightarrow \langle A, G_m \rangle$  is a well-ordered set.  $\square$

$\square$

Using the above we have proved that

Axiom Of Choice  $\Rightarrow$  Hausdorff's Maximal Principle  $\Rightarrow$  Zorn's Lemma  $\Rightarrow$  Well-Ordering Theorem.

We prove now that the Axiom of Choice follows from the Well-Ordering Theorem.

**Theorem 2.232.** *The Well-Ordering Theorem implies the Axiom of Choice*

**Proof.** If  $A$  is any set then choose a well-ordering  $\langle A, \leq \rangle$  of  $A$ . Define then  $c: \mathcal{P}'(A) \rightarrow A$  by  $c = \{(B, b) | B \in \mathcal{P}'(A) \wedge b \text{ is least element of } B\} \subseteq \mathcal{P}'(A) \times A$ . We have then that if  $(B, b), (B, b') \in c$   $\xrightarrow{\text{least element is unique}} b = b'$  proving that  $c$  is a partial function. And if  $B \in \mathcal{P}'(A) \Rightarrow \emptyset \neq B \subseteq A \xrightarrow{\text{well-ordering}} \exists b \vdash b \text{ is least element in } B \Rightarrow (B, b) \in c$  proving that  $\text{dom}(c) = \mathcal{P}'(A)$  and thus that  $c: \mathcal{P}'(A) \rightarrow A$  is a function. Finally if  $B \in \mathcal{P}'(A)$  we have  $c(B) \in B$  as the least element of a set is in the set itself.  $\square$

So we have proved the following equivalences.

**Theorem 2.233.** *The following statements are equivalent*

1. *Axiom of Choice*
2. *Hausdorff's Maximal Principle*
3. *Zorn's Lemma*
4. *Every set can be well-ordered*

For the rest of this document we assume that the Axiom of Choice is valid and thus also that the Hausdorff's Maximal Principle, Zorn's Lemma and 'Every set can be well-ordered' holds true.

As an example of applying the Axiom of Choice let's prove the following theorem.

**Theorem 2.234.** *Assume  $X, Y$  sets and  $f: X \rightarrow Y$  a function then there exists a  $Z \subseteq X$  such that  $f|_Z: X \rightarrow Y$  is injective and  $f|_Z(X) = f(X)$  (which ensures that  $f|_Z: Z \rightarrow f(X)$  is a bijection).*

**Proof.** First define  $\mathcal{A} = \{f^{-1}(\{y\}) | y \in f(X)\}$ . If  $A \in \mathcal{A}$  there exists a  $y \in f(X)$  such that  $A = f^{-1}(\{y\}) \Rightarrow A \subseteq X$ , as  $y \in f(X)$  there exists a  $x \in X$  such that  $y = f(x) \Rightarrow x \in f^{-1}(\{y\}) = A \Rightarrow A \neq \emptyset \Rightarrow A \in \mathcal{P}'(X)$  so we conclude that  $\mathcal{A} \subseteq \mathcal{P}'(X)$ . By the Axiom of Choice (see 2.207) there exists a choice function  $c: \mathcal{P}'(X) \rightarrow X$  such that if  $A \in \mathcal{P}'(X)$  then  $c(A) \in A$ . Take now  $Z = c(\mathcal{A})$  and consider  $f|_Z: Z \rightarrow Y$  and prove that is injective. Let  $z_1, z_2 \in Z$  is such that  $f|_Z(z_1) = f|_Z(z_2) \Rightarrow f(z_1) = f(z_2)$ . As  $z_1, z_2 \in Z = c(\mathcal{A})$  there exists  $A_1, A_2 \in \mathcal{A}$  such that  $z_1 = c(A_1) \in A_1, z_2 = c(A_2) \in A_2$  and thus as  $A_1, A_2 \in \mathcal{A}$  there exists  $y_1, y_2 \in f(X)$  such that  $A_1 = f^{-1}(\{y_1\}), A_2 = f^{-1}(\{y_2\})$  so  $z_1 \in f^{-1}(\{y_1\}), z_2 \in f^{-1}(\{y_2\}) \Rightarrow f(z_1) = y_1, f(z_2) = y_2 \xrightarrow{f(z_1) = f(z_2)} y_1 = y_2 \Rightarrow A_1 = f^{-1}(\{y_1\}) = f^{-1}(\{y_2\}) = A_2 \Rightarrow z_1 = c(A_1) = c(A_2) = z_2 \Rightarrow z_1 = z_2$  proving that  $f|_Z: Z \rightarrow Y$  is injective.

If now  $y \in f(X)$  then  $f^{-1}(\{y\}) \in \mathcal{A} \Rightarrow x = c(f^{-1}(\{y\})) \in c(\mathcal{A}) = Z \Rightarrow x \in Z$  and as also  $x = c(f^{-1}(\{y\})) \in f^{-1}(\{y\})$  we have  $f(x) = y$  so we have found a  $x \in Z$  such that  $y = f(x) \Rightarrow f(X) \subseteq f(Z)$ . From  $Z \subseteq X$  we have  $f(Z) \subseteq f(X)$  and we conclude thus that  $f(X) = f(Z)$  as must be proved.  $\square$

Finally we prove that Zorn's lemma is also valid if we replace partial ordered sets by pre-ordered sets.

**Theorem 2.235.** *Let  $\langle A, \leq \rangle$  be a pre-ordered set (see 2.135) such that every chain  $C \subseteq A$  has an upper bound (see 2.169) then  $A$  has a maximal element (see 2.164)*

**Proof.** Use 2.142 to construct the partial ordered set  $\langle A/\sim, \leq \rangle$  where  $x \sim y$  iff  $x \leq y \wedge y \leq x$  and where  $\sim[x] < \sim[y]$  iff  $x < y$ . Let now  $C \subseteq A/\sim$  be a chain in  $A/\sim$  (in the pre-order sense). Form then  $C' = \bigcup_{S \in C} S \subseteq A$ . If  $x, y \in C'$  then there exists  $\sim[x'], \sim[y'] \in C$  such that  $x \in \sim[x'], y \in \sim[y']$  or  $x \leq x' \wedge x' \leq y$  and  $y \leq y' \wedge y' \leq y$ . As  $C$  is a chain we have either:

1.  $\sim[x'] \leq \sim[y'] \xrightarrow{2.142} x' \leq y' \Rightarrow x \leq y$
2.  $\sim[y'] \leq \sim[x'] \xrightarrow{2.142} y' \leq x' \Rightarrow y \leq x$

proving that  $C'$  is a chain in  $A$ . By the hypothesis there exists a upper bound  $u$  of  $C'$  in  $A$ , in other words  $\exists u \in U$  such that  $\forall a \in A$  we have  $a \leq u$ . Consider now  $\sim[u]$  and take  $\sim[z] \in C$  then as  $z \in \sim[z]$  we have  $z \in C'$  and thus  $z \leq u$  proving that  $\sim[z] \leq \sim[u]$ . This proves that every chain in  $A/\sim$  has a upper bound and thus by Zorn's lemma there exists a maximal element  $\sim[m] \in A$  of  $A$ . If now  $x \in A$  with  $m \leq x$  then  $\sim[m] \leq \sim[x]$  so by maximality of  $\sim[m]$  we have then that  $\sim[m] = \sim[x]$  giving  $x \leq m$  and thus by 2.164 that  $m$  is a maximal element in the pre-order of  $A$ .  $\square$

# Chapter 3

## Algebraic Constructs

### 3.1 Groups

**Definition 3.1.** A semi-group is a pair  $\langle G, \odot \rangle$  where  $G$  is a set and  $\odot: G \times G \rightarrow G$  a function such that (here  $\odot(x, y)$  is noted as  $x \odot y$ )

1. (Neutral Element)  $\exists e \in G \vdash \forall x \in G \vdash x \odot e = e \odot x$
2. (Associativity)  $\forall x, y, z \in G$  we have  $x \odot (y \odot z) = (x \odot y) \odot z$

**Lemma 3.2.** If  $\langle G, \odot \rangle$  is a semi-group then  $G \neq \emptyset$

**Proof.** This follows from (1) in the definition. □

**Theorem 3.3.** If  $\langle G, \odot \rangle$  is a semi-group then there exists only one neutral element.

**Proof.** If  $e, e'$  are neutral elements then  $e = e' \odot e = e'$  □

**Definition 3.4.** A group is a semi-group  $\langle G, \odot \rangle$  with the extra condition

1. (Inverse Element)  $\forall x \in G \vdash \exists y \in G \vdash x \odot y = e = y \odot x$

**Theorem 3.5.** If  $\langle G, \odot \rangle$  is a group then  $\forall x \in G \vdash \exists! y \in G \Rightarrow x \odot y = e = y \odot x$  this unique element is noted as  $x^{-1}$  □

**Proof.** Take  $x \in G$  and suppose that  $y, y'$  is such that  $x \odot y = e = y \odot x$  and  $x \odot y' = e = y' \odot x$  then  $y = y \odot e = y \odot (x \odot y') = (y \odot x) \odot y' = e \odot y' = y'$  □

**Theorem 3.6.** If  $\langle G, \odot \rangle$  is a group then  $\forall x \in G$  we have  $(x^{-1})^{-1} = x$

**Proof.** If  $x \in G$  then  $x \odot x^{-1} = e = x^{-1} \odot x$  and  $(x^{-1})^{-1} \odot x^{-1} = e = x^{-1} \odot (x^{-1})^{-1}$ , using the previous theorem we have then  $x = (x^{-1})^{-1}$  □

**Definition 3.7.** A group or semi-group  $\langle G, \odot \rangle$  is **abelian** or **commutative**. If the following is satisfied

1. (Commutativity)  $\forall x, y \in G \vdash x \odot y = y \odot x$

**Definition 3.8.** If  $\langle G, \odot \rangle$  is a semi-group then  $F \subseteq G$  is a sub-semi-group of  $G$  if

1.  $\forall x, y \in F \vdash x \odot y \in F$
2.  $e \in F$

**Theorem 3.9.** If  $\langle G, \odot \rangle$  is a semi-group then if  $F \subseteq G$  is a sub-semi-group then  $\langle F, \odot|_{F \times F} \rangle$  is a semi-group. For simplicity we note this as  $\langle F, \odot \rangle$

**Proof.**

1. If  $x, y \in F \Rightarrow x \odot y \in F \Rightarrow x \odot|_{F \times F} y \in F \Rightarrow \odot|_{F \times F}: F \times F \rightarrow F$  is a function
2.  $e \in F \Rightarrow x \odot|_{F \times F} e = x \odot e = x = e \odot x = e \odot|_{F \times F} x$
3. If  $x, y, z \in F \Rightarrow x \odot|_{F \times F} (y \odot|_{F \times F} z) = x \odot (y \odot z) = (x \odot y) \odot z = (x \odot|_{X \times Y} y) \odot|_{F \times F} z$  □

**Definition 3.10.** If  $\langle G, \odot \rangle$  is a group then  $F \subseteq G$  is a sub-group of  $G$  if

1.  $\forall x, y \in F \models x \odot y \in F$
2.  $\forall x \in F \models x^{-1} \in F$
3.  $e \in F$  where  $e$  is the unit in  $G$ .

**Theorem 3.11.** If  $\langle G, \odot \rangle$  is a group then if  $F \subseteq G$  is a sub-group we have that  $\langle F, \odot|_{F \times F} \rangle$  is a group noted as  $\langle F, \odot \rangle$  for simplicity.

**Proof.** As  $e \in F$  we have using 3.9 that  $\langle F, \odot \rangle$  is a semi-group. Also if  $x \in F \Rightarrow x^{-1} \in F = x \odot|_{F \times F} x^{-1} = x \odot x^{-1} = e = x^{-1} \odot x = x^{-1} \odot|_{F \times F} x$   $\square$

**Definition 3.12.** If  $\langle F, \odot \rangle$  and  $\langle G, \oplus \rangle$  are semi-groups then a **group homeomorphism** is a function  $f: F \rightarrow G$  such that  $\forall x, y \in F$  we have  $f(x \odot y) = f(x) \oplus f(y)$

**Theorem 3.13.** If  $\langle F, \odot \rangle$  and  $\langle G, \oplus \rangle$  are groups with units  $e_F, e_G$  and  $f: F \rightarrow G$  a group homeomorphism then we have

1.  $f(e_F) = f(e_G)$
2.  $\forall a \in F$  we have  $f(a^{-1}) = f(a)^{-1}$

**Proof.**

1.  $f(e_F) = f(e_F \cdot e_F) = f(e_F) \cdot f(e_F)$  multiply both sides by  $f(e_F)^{-1}$   $e_G = f(e_F)$
2. If  $a \in F \Rightarrow f(a^{-1} \cdot a) = f(e_F) \stackrel{(1)}{=} e_G \stackrel{(2)}{=} f(e_F) = f(a \cdot a^{-1}) \Rightarrow f(a^{-1}) \cdot f(a) = e_G = f(a) \cdot f(a^{-1}) \Rightarrow f(a^{-1}) = f(a)^{-1}$   $\square$

**Definition 3.14.** If  $\langle F, \odot \rangle$  and  $\langle G, \oplus \rangle$  are semi-groups then a **group isomorphism** is a function  $f: F \rightarrow G$  that is a group homeomorphism and a bijection.

**Theorem 3.15.** If  $\{\langle F_i, \odot_i \rangle\}_{i \in I}$  is a family of (abelian) groups (or semi-groups) then we have:

1.  $\forall x, y \in \prod_{i \in I} F_i$  we have:
  - a.  $\forall i \in I$  that  $x(i) \odot_i y(i) \in F_i$
  - b.  $x \odot y: I \rightarrow \bigcup_{i \in I} F_i$  defined by  $i \rightarrow x(i) \odot_i y(i)$  is a element of  $\prod_{i \in I} F_i$
2.  $\odot: \prod_{i \in I} F_i \times \prod_{i \in I} F_i \rightarrow \prod_{i \in I} F_i$  defined by  $(x, y) \rightarrow x \odot y$  is a function [here we use  $x \odot y$  to actually note the function  $\langle x \odot y, I, \bigcup_{i \in I} F_i \rangle$ ]
3.  $\langle \prod_{i \in I} F_i, \odot \rangle$  is a (abelian) group (or semi group) with neutral element  $0: I \rightarrow \bigcup_{i \in I} F_i$  defined by  $i \rightarrow 0_i \in F_i$  (where  $0_i$  is the neutral element in  $F_i$ ) and if  $x \in \prod_{i \in I} F_i$  then  $-x: I \rightarrow \bigcup_{i \in I} F_i$  defined by  $i \rightarrow -x(i) \in F_i$  (where  $-x(i)$  is the inverse element of  $x(i)$ )

**Proof.**

1. If  $x, y \in \prod_{i \in I} F_i$  then we have
  - a. As  $x(i), y(i) \in F_i$  by the definition of  $\prod_{i \in I} F_i$  we have as  $\odot: F_i \odot_i F_i \rightarrow F_i$  is a function we have that  $x(i) \odot_i y(i) \in F_i$
  - b. As  $\forall i \in I$  we have by  $x(i) \odot_i y(i) \in F_i$  we have by definition of  $\prod_{i \in I} F_i$  that  $x \odot y \in \prod_{i \in I} F_i$
2. This follows from the definition of  $\odot$  and the fact that  $x \odot y \in F_i$
3. We have if  $\{\langle F_i, \odot_i \rangle\}_{i \in I}$  is a family of semi-groups then we have
  - **(associativity)**  $\forall i \in I$  we have  $((x \odot y) \odot z)(i) = (x \odot y)(i) \odot_i z(i) = (x(i) \odot_i y(i)) \odot_i z(i) \stackrel{\langle F_i, \odot_i \rangle \text{ is a semi-group}}{=} x(i) \odot_i (y(i) \odot_i z(i)) = x(i) \odot_i (y \odot z)(i) = (x \odot (y \odot z))(i)$  resulting in  $(x \odot y) \odot z = x \odot (y \odot z)$

- **(neutral element)**  $\forall i \in I$  we have  $(x \odot 0)(i) = x(i) \odot_i 0(i) = x(i) \odot_i 0_i = x(i) = 0_i \odot_i x(i) = (0 \odot x)$  proving that  $0 \odot x = x = x \odot 0$

If  $\{\langle F_i, \odot_i \rangle\}_{i \in I}$  is also a group then we have

- **(inverse element)**  $\forall i \in I$  we have  $(x \odot (-x))(i) = x(i) \odot_i ((-x)(i)) = x(i) \odot_i (-x(i)) = 0_i = 0(i) = 0_i = (-x(i)) \odot x(i) = (-x)(i) \odot_i x(i) = ((-x) \odot x)(i)$  giving  $x \odot (-x) = 0 = (-x) \odot x$

and in case of a abelian family we have

- **(commutativity)**  $\forall i \in I$  we have  $(x \odot y)(i) = x(i) \odot_i y(i) = y(i) \odot_i x(i) = (y \odot x)(i)$  giving  $x \odot y = y \odot x$   $\square$

**Definition 3.16.** Let  $\langle G, \odot \rangle$  be a group with neutral element  $0_G$  and let  $X$  be a set then we have the following definitions:

1. A **left group action** is a function  $\triangleright: G \times X \rightarrow X$  where  $\triangleright(g, x) =_{\text{noted}} g \triangleright x$  such that
  - $\forall x \in X$  we have  $0_G \triangleright x = x$
  - $\forall g, g' \in G$  and  $\forall x \in X$  we have  $(g \odot g') \triangleright x = g \triangleright (g' \triangleright x)$
2. A **right group action** is a function  $\triangleleft: X \times G \rightarrow X$  where  $\triangleleft(x, g) =_{\text{noted}} x \triangleleft g$  such that
  - $\forall x \in X$  we have  $x \triangleleft 0_G = x$
  - $\forall g, g' \in G$  and  $\forall x \in X$  we have  $x \triangleleft (g \odot g') = (x \triangleleft g) \triangleleft g'$

**Note 3.17.** If  $\langle G, \odot \rangle$  is a group and  $X$  as set with a left group action  $\triangleright$  then you think that we can always define a right group action  $\triangleleft$  as follows:  $\triangleleft: X \times G \rightarrow X$  is defined by  $(x, g) \rightarrow x \triangleleft g \equiv g \triangleright g$ , however this will in general not give you a right group action as we have not in general that 2 (b) is not fulfilled for  $x \triangleleft (g \odot g') =_{\text{definition}} (g \odot g') \triangleright x = g \triangleright (g' \triangleright x) =_{\text{definition}} (g' \triangleright x) \triangleleft g =_{\text{definition}} (x \triangleleft g') \triangleleft g$  which in general is not the same as  $(x \triangleleft g) \triangleleft g'$ .

**Definition 3.18.** Let  $\langle G, \odot \rangle$  be a group,  $X$  a set then if  $\triangleright$  ( $\triangleleft$ ) is a left (right) group action then we define if  $g \in G$

1.  $g_{\triangleright}: X \rightarrow X$  by  $x \rightarrow g_{\triangleright}(x) = g \triangleright x$
2.  $g_{\triangleleft}: X \rightarrow X$  by  $x \rightarrow g_{\triangleleft}(x) = x \triangleleft g$

**Definition 3.19.** Let  $\langle G, \odot \rangle$  be a group with neutral element  $0_G$  and let  $X$  be a set then we have the following definitions for a left (right) group action  $\triangleright$  ( $\triangleleft$ ):

1.  $\triangleright$  (or  $\triangleleft$ ) is **faithful** iff  $g_{\triangleright} = 1_X$  (or  $g_{\triangleleft} = 1_X$ ) if and only if  $g = 0_G$  where  $1_X: X \rightarrow X$  is the identity mapping. Equivalently this means that  $\{g \in G \mid \forall x \in X \mid g \triangleright x = x\} = \{0_G\}$  (or  $\{g \in G \mid \forall x \in X \mid x \triangleleft g = x\} = \{0_G\}$ )
2.  $\triangleright$  (or  $\triangleleft$ ) is **transitive** iff  $\forall x_1, x_2 \in X$  there exists a  $g \in G$  such that  $g \triangleright x_1 = x_2$  (or  $x_1 \triangleleft g = x_2$ )
3.  $\triangleright$  (or  $\triangleleft$ ) is **free** iff  $\forall x \in X$  we have  $\{g \in G \mid g \triangleright x = x\} = \{0_G\}$  (or  $\{g \in G \mid x \triangleleft g = x\} = \{0_G\}$ )

## 3.2 Rings

**Definition 3.20. (Ring)** A triple  $\langle R, \oplus, \odot \rangle$  is a ring iff

1.  $R$  is a set
2.  $\oplus: R \times R \rightarrow R$  is a function
3.  $\odot: R \times R \rightarrow R$  is a function
4.  $\langle R, \oplus \rangle$  is a abelian group
  - (Associative)**  $\forall a, b, c \in R \models a \oplus (b \oplus c) = (a \oplus b) \oplus c$
  - (Neutral element)**  $\exists 0 \in R \models \forall a \in R \models a \oplus 0 = a = 0 \oplus a$

- c. **(Inverse)**  $\forall a \in R \text{ there } \exists -a \in R \models a \oplus (-a) = 0 = (-a) \oplus a$
- d. **(Commutative)**  $\forall a, b \in R \models a \oplus b = b \oplus a$
- 5. **(Distributive)**  $\forall a, b, c \in R \models a \odot (b \oplus c) = (a \odot b) \oplus (a \odot c)$
- 6. **(Neutral element)**  $\exists 1 \in R \models \forall a \in R \models a \odot 1 = a = 1 \odot a$
- 7. **(Commutative)**  $\forall a, b \in R \models a \odot b = b \odot a$
- 8. **(Associative)**  $\forall a, b, c \in R \models a \odot (b \odot c) = (a \odot b) \odot c$

Note that in a ring we have also that  $\langle R, \odot \rangle$  is a abelian semi-group.

**Definition 3.21.** If  $\langle R, \oplus, \odot \rangle$  is a ring then a **zero divisor** is a  $x \in R \setminus \{0\}$  so that there exists a  $b \in R \setminus \{0\}$  such that  $a \cdot b = 0$ .

**Definition 3.22.** A ring  $\langle R, \oplus, \odot \rangle$  is a **integral domain** if it does not contain a **zero divisor**

**Definition 3.23. (Sub-Ring)** If  $\langle R, \oplus, \odot \rangle$  is a ring then a subset  $S \subseteq R$  is a sub ring iff

- 1.  $\forall a, b \in S$  we have  $a \oplus b \in S$
- 2.  $\forall a, b \in S$  we have  $a \odot b \in S$
- 3.  $\forall a \in S$  we have  $-a \in S$
- 4.  $1 \in S$  where 1 is the unit of multiplication  $\odot$
- 5.  $0 \in S$  where 0 is the unit of multiplication  $\oplus$

**Theorem 3.24.** If  $\langle R, \oplus, \odot \rangle$  is a ring and  $S \subseteq R$  a sub-ring then  $\langle S, \oplus|_{S \times S}, \odot|_{S \times S} \rangle$  is ring. For simplicity we note this ring as  $\langle S, \oplus, \odot \rangle$ .

**Proof.**

- 1. As  $S \subseteq R$  then by 1.58 S is a set.
- 2. As  $\oplus: R \times R \rightarrow R$  is a function we have by 2.29 that  $\oplus|_{S \times S}: S \times S \rightarrow S$  is a function.
- 3. As  $\odot: R \times R \rightarrow R$  is a function we have by 2.29 that  $\odot|_{S \times S}: S \times S \rightarrow S$  is a function.
- 4.  $\langle S, \oplus|_{S \times S} \rangle$  is a abelian group
  - a. **(Associative)**  $\forall a, b, c \in S$  we have  $(a \oplus|_{S \times S} b) \oplus|_{S \times S} c = (a \oplus b) \oplus c = a \oplus (b \oplus c) = a \oplus|_{S \times S} (b \oplus|_{S \times S} c)$
  - b. **(Neutral element)** As  $0 \in S$  we have  $\forall a \in S \Rightarrow a \oplus|_{S \times S} 0 = a \oplus 0 = a = 0 \oplus a = 0 \oplus|_{S \times S} a$
  - c. **(Inverse element)** If  $a \in S$  then there exists  $a - a \in S$  with  $a \oplus (-a) = 0 = (-a) \oplus a \Rightarrow a \oplus|_{S \times S} (-a) = a \oplus (-a) = 0 = (-a) \oplus a = (-a) \oplus|_{S \times S} a$
  - d. **(Commutativity)** If  $a, b \in S$  then  $a \oplus|_{S \times S} b = a \oplus b = b \oplus a = b \oplus|_{S \times S} a$
- 5. **(Distributivity)**  $\forall a, b, c \in S$  we have  $a \odot|_{S \times S} (b \oplus|_{S \times S} c) = a \odot (b \oplus c) = (a \odot b) \oplus (a \odot c) = (a \odot|_{S \times S} b) \oplus|_{S \times S} (a \odot|_{S \times S} c)$
- 6. **(Neutral Element)** As  $1 \in S$  then we have if  $a \in S$  that  $1 \odot|_{S \times S} a = 1 \odot a = a \odot 1 = a \odot|_{S \times S} 1$
- 7. **(Commutativity)** If  $a, b \in S$  then  $a \odot|_{S \times S} b = a \odot b = b \odot a = b \odot|_{S \times S} a$
- 8. **(Associative)** If  $a, b, c \in S$  then  $a \odot|_{S \times S} (b \odot|_{S \times S} c) = a \odot (b \odot c) = (a \odot b) \odot c = (a \odot|_{S \times S} b) \odot|_{S \times S} c$   $\square$

**Theorem 3.25.** If  $\langle R, \oplus, \odot \rangle$  is a ring with neutral element 0 in  $\langle R, + \rangle$  then  $\forall x \in R$  we have  $x \odot 0 = 0 \odot x = 0$

**Proof.** If  $x \in R$  then  $0 = (0 \odot x) \oplus (- (0 \odot x)) \stackrel{0 \oplus 0 = 0}{=} ((0 \oplus 0) \odot x) \oplus (- (0 \odot x)) = [(0 \odot x) \oplus (0 \odot x)] \oplus (- (0 \odot x)) = 0 \odot x \oplus [(0 \odot x) + (- (0 \odot x))] = 0 \odot x + 0 = 0 \odot x$   $\square$

**Definition 3.26.** If  $\langle A, \oplus_A, \odot_A \rangle$  and  $\langle B, \oplus_B, \odot_B \rangle$  are rings then a function  $f: A \rightarrow B$  is a ring homeomorphism iff

1.  $\forall a, b \in A$  we have  $f(a \oplus_A b) = f(a) \oplus_B f(b)$
  2.  $\forall a, b \in A$  we have  $f(a \odot_A b) = f(a) \odot_B f(b)$
  3.  $f(1_A) = 1_B$  where  $1_A$  is the multiplicative inverse in  $A$  and  $1_B$  is the multiplicative inverse in  $B$ .

Note that a ring homeomorphism  $f: A \rightarrow B$  for the rings  $\langle A, \oplus_A, \odot_A \rangle, \langle B, \oplus_B, \odot_B \rangle$  is automatically a group homeomorphism for the groups  $\langle A, \oplus_A \rangle, \langle B, \oplus_B \rangle$ . Using 3.13 we have then the following theorem

**Theorem 3.27.** If  $\langle A, \oplus_A, \odot_A \rangle$  and  $\langle B, \oplus_B, \odot_B \rangle$  are rings with additive units  $0_A, 0_B$  and  $f: A \rightarrow B$  a ring homeomorphism then we have

1.  $f(0_A) = 0_B$
  2.  $\forall a \in A$  we have  $f(-a) = -f(a)$

**Definition 3.28.** If  $\langle A, \oplus_A, \odot_A \rangle$  and  $\langle B, \oplus_B, \odot_B \rangle$  are rings then a function  $f: A \rightarrow B$  is a ring homeomorphism if it is a ring homeomorphism and a bijection.

### 3.3 Fields

**Definition 3.29. (Field)** A ring  $\langle F, \oplus, \odot \rangle$  is a field if

1.  $F$  is a set
  2.  $\oplus: F \times F \rightarrow F$  is a function
  3.  $\odot: F \times F \rightarrow F$  is a function
  4.  $\langle F, \oplus, \odot \rangle$  is a Ring
    - a.  $\langle F, \oplus \rangle$  is an abelian group
      - i. (Associative)  $\forall a, b, c \in F \models a \oplus (b \oplus c) = (a \oplus b) \oplus c$
      - ii. (Neutral element)  $\exists 0 \in F \vdash \forall a \in F \models a \oplus 0 = a = 0 \oplus a$
      - iii. (Inverse)  $\forall a \in F \text{ there } \exists -a \in F \models a \oplus (-a) = 0 = (-a) \oplus a$
      - iv. (Commutative)  $\forall a, b \in F \vdash a \oplus b = b \oplus a$
    - b. (Distributive)  $\forall a, b, c \in F \models a \odot (b \oplus c) = (a \odot b) \oplus (a \odot c)$
    - c. (Neutral element)  $\exists 1 \in F \vdash \forall a \in F \models a \odot 1 = a = 1 \odot a$
    - d. (Commutative)  $\forall a, b \in F \models a \odot b = b \odot a$
    - e. (Associative)  $\forall a, b, c \in F \models a \odot (b \odot c) = (a \odot b) \odot c$
  5. (Inverse for non zero element) If  $a \in F \setminus \{0\}$  then  $\exists b \in F \vdash b \odot a = 1$

So a field is a ring that has also for every non-zero element a multiplicative inverse. We prove now that the inverse for non zero element is unique.

**Theorem 3.30.** If  $\langle F, \oplus, \odot \rangle$  is a field then  $\forall a \in F$  with  $a \neq 0$  there exists a **unique** inverse element, we note this inverse element by  $a^{-1}$ .

**Proof.** Let  $a \in F$  with  $a \neq 0$  then by definition there exists a  $b \in F$  such that  $b \odot a = 1$ , assume that there exists also a  $b' \in F$  such that  $b' \odot a = 1$ . Then we have that  $(b \odot a) \odot b' = b' \Rightarrow b \odot (a \odot b') = b'$  by associativity  $\Rightarrow b \odot (b' \odot a) = b' \Rightarrow b \odot 1 = b' \Rightarrow 1 \odot b = b' \Rightarrow b = b'$  by commutativity  $\square$

**Definition 3.31.** If  $\langle F, \oplus, \odot \rangle$  is a field then a subset  $S \subseteq F$  is a sub-field iff the following is full filled

1.  $\forall a, b \in S \models a \oplus b \in S$
2.  $\forall a, b \in S \models a \odot b \in S$
3.  $\forall a \in S \models -a \in S$
4.  $\forall a \in S \setminus \{0\} \models a^{-1} \in S$
5.  $0 \in S$  (where 0 is the additive neutral element of  $F$ )
6.  $1 \in S$  (where 1 is the multiplicative neutral element of  $F$ )

**Theorem 3.32.** If  $\langle F, \oplus, \odot \rangle$  is a field and  $S \subseteq F$  a sub field then  $\langle S, \oplus|_S, \odot|_S \rangle$  (for simplicity of notation we note this as  $\langle S, \oplus, \odot \rangle$ )

**Proof.** Using 3.24 we have that  $\langle S, \oplus|_S, \odot|_S \rangle$  is a sub ring so we only have to prove that every non-zero element in  $S$  has a multiplicative inverse. If  $x \in S \setminus \{0\} \subseteq F \setminus \{0\} \Rightarrow x^{-1}$  exists, then by the definition of a sub-field  $x^{-1} \in S$  and we have  $x \odot|_S x^{-1} = x \odot x^{-1} = 1 = x^{-1} \odot x = x^{-1} \odot|_S x$ .  $\square$

**Definition 3.33.** If  $\langle F_1, \odot_1, \oplus \rangle$  and  $\langle F_2, \odot_2, \oplus_2 \rangle$  are fields with multiplicative units  $1_1, 1_2$  then a function  $f: F_1 \rightarrow F_2$  is a field homeomorphism iff

1.  $\forall a, b \in F_1$  we have  $f(a \odot_1 b) = f(a) \odot_2 f(b)$
2.  $\forall a, b \in F_1$  we have  $f(a \oplus_1 b) = f(a) \oplus_2 f(b)$
3.  $f(1_1) = 1_2$

If  $f$  is also a bijection then we call  $f$  a field isomorphism.

Note that a field homeomorphism  $f: A \rightarrow B$  for the fields  $\langle A, \oplus_A, \odot_A \rangle, \langle B, \oplus_B, \odot_B \rangle$  is automatically a group homeomorphism for the groups  $\langle A, \oplus_A \rangle, \langle B, \oplus_B \rangle$ . Using 3.13 we have then the following theorems

**Theorem 3.34.** If  $\langle F_1, \odot_1, \oplus \rangle$  and  $\langle F_2, \odot_2, \oplus_2 \rangle$  are fields with multiplicative units  $1_1, 1_2$  and  $f: F_1 \rightarrow F_2$  is a field isomorphism then  $f^{-1}: F_2 \rightarrow F_1$  is a field isomorphism

**Proof.**

1. If  $a, b \in F_2$  then  $f((f^{-1})(a) \odot_1 (f^{-1})(b)) = f((f^{-1})(a)) \odot_2 f((f^{-1})(b)) = f((f^{-1})(a)) \odot_2 f((f^{-1})(b)) = a \odot_2 b$  hence  $(f^{-1})(a \odot_2 b) = (f^{-1})(f((f^{-1})(a) \odot_1 (f^{-1})(b))) = (f^{-1})(a) \odot_1 (f^{-1})(b)$
2. If  $a, b \in F_1$  then  $f((f^{-1})(a) \oplus_1 (f^{-1})(b)) = f((f^{-1})(a)) \oplus_2 f((f^{-1})(b)) = a \oplus_2 b$  hence  $(f^{-1})(a \oplus_2 b) = (f^{-1})(f((f^{-1})(a) \oplus_1 (f^{-1})(b))) = (f^{-1})(a) \oplus_1 (f^{-1})(b)$
3. From  $f(1_1) = 1_2$  it follows that  $(f^{-1})(1_2) = 1_1$   $\square$

**Theorem 3.35.** If  $\langle A, \oplus_A, \odot_A \rangle$  and  $\langle B, \oplus_B, \odot_B \rangle$  are fields with additive units  $0_A, 0_B$  and multiplicative units  $1_A, 1_B$  and  $f: A \rightarrow B$  a field homeomorphism then we have

1.  $f(0_A) = 0_B$
2.  $\forall a \in A$  we have  $f(-a) = -f(a)$
3.  $\forall a \in A$  with  $a \neq 0_A$  we have  $f(a^{-1}) = (f(a))^{-1}$

**Proof.** As field homeomorphism  $f: A \rightarrow B$  for the fields  $\langle A, \oplus_A, \odot_A \rangle, \langle B, \oplus_B, \odot_B \rangle$  is automatically a group homeomorphism for the groups  $\langle A, \oplus_A \rangle, \langle B, \oplus_B \rangle$  we have by 3.13 that (1) and (2) are valid. As for (3), if  $a \in A$  with  $a \neq 0_A$  then there exists a  $a^{-1}$  such that  $a^{-1} \cdot a = 1_A$  hence  $1_B = f(1_A) = f(a^{-1} \cdot a) = f(a^{-1}) \cdot f(a)$  proving that  $f(a^{-1}) = (f(a))^{-1}$ .  $\square$

# Chapter 4

## The Natural Numbers

### 4.1 Definition of the Natural Numbers

Recall the definition of successor (see 1.66), successor set (see 1.67) and the Axiom of Infinity (1.68).

**Definition 4.1.** *If  $A$  is a set then the successor of  $A$  is defined to be the set  $s(A) = A \cup \{A\}$*

**Definition 4.2.** *A set  $A$  is a successor set if*

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

**Axiom 4.3. (Axiom of Infinity)** *There exists a successor set*

We can now define the set of natural numbers  $\mathbb{N}$  as follows

**Definition 4.4. (Natural Numbers)** *Let  $\mathbb{S} = \{S : S \text{ is a successor set}\}$  then  $\mathbb{N} = \bigcap_{S \in \mathbb{S}} S$  is the set of natural numbers. In other words the set of Natural Numbers is the intersection of all successor sets. Note that there exists a successor set  $S$  and  $\mathbb{N} \subseteq S$  so we have by 1.58 that  $\mathbb{N}$  is a set.*

As for all successor sets  $S$  we have  $\emptyset \in S, s(\emptyset) \in S, s(s(\emptyset)) \in S, \dots$  so thus  $\emptyset, s(\emptyset), s(s(\emptyset)), \dots \in \mathbb{N}$

**Definition 4.5.** *The successor function  $s : \mathbb{N} \rightarrow \mathbb{N}$  is the function defined by  $n \rightarrow s(n)$*

**Definition 4.6.** *We define the numbers*

1.  $0 = \emptyset \in \mathbb{N}$
2.  $1 = s(\emptyset) = \emptyset \cup \{\emptyset\} = \{\emptyset\} = \{0\} \in \mathbb{N}$
3.  $2 = s(1) = 1 \cup \{1\} = \{\emptyset\} \cup \{1\} = \{\emptyset, 1\} = \{\emptyset, \{\emptyset\}\} \in \mathbb{N}$
4.  $3 = s(2) = 2 \cup \{2\} = \{\emptyset, 1\} \cup \{2\} = \{\emptyset, 1, 2\} = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$
5.  $\dots$

**Definition 4.7.** *We use the following notations for subsets of  $\mathbb{N}$*

1.  $\mathbb{N}_0 = \mathbb{N} \setminus \{0\}$
2.  $\mathbb{N}_1 = \mathbb{N} \setminus \{0, 1\}$
3.  $\mathbb{N}_2 = \mathbb{N} \setminus \{0, 1, 2\}$
4.  $\dots$

**Theorem 4.8.** *If  $n \in \mathbb{N} \Rightarrow s(n) \neq 0$*

**Proof.** By definition  $s(n) = n \cup \{n\} \Rightarrow n \in s(n) \Rightarrow s(n) \neq \emptyset = 0$   $\square$

**Theorem 4.9.** *If  $n \in \mathbb{N} \Rightarrow s(n) \in \mathbb{N}$*

**Proof.** If  $n \in \mathbb{N}$  then if  $S$  is a successor set we have that  $n \in S \Rightarrow s(n) \in S \Rightarrow s(n) \in \bigcap_{S \in \mathbb{S}} S = \mathbb{N}$   $\square$

**Theorem 4.10. (Mathematical Induction)** *If  $X \subseteq \mathbb{N}$  has the following properties*

1.  $0 \in X$
2.  $n \in X \Rightarrow s(n) \in X$

*then we have  $X = \mathbb{N}$*

**Proof.** By (1) and (2) we have that  $X$  is a successor set and thus by 4.4 we have  $\mathbb{N} \subseteq X$  so that from  $X \subseteq \mathbb{N}$  we have  $X = \mathbb{N}$   $\square$

**Theorem 4.11.** *Let  $m, n \in \mathbb{N}$  be natural numbers then if  $m \in s(n)$  then  $m \in n \vee m = n$*

**Proof.** As  $s(n) = n \bigcup \{n\}$  then from  $m \in s(n)$  we have either  $m \in n \vee m \in \{n\}$  so we have either  $m \in n \vee m = n$   $\square$

**Definition 4.12.** *A set  $A$  is **transitive** if  $\forall x \in A \vdash x \subseteq A$*

**Theorem 4.13.**  *$\forall n \in \mathbb{N}$  we have that  $n$  is transitive*

**Proof.** We prove this by mathematical induction so let  $S = \{n \in \mathbb{N} \mid n \text{ is transitive}\} \subseteq \mathbb{N}$ . We have then

1.  $0 = \emptyset$  so  $\forall x \in \emptyset \vdash x \subseteq \emptyset$  is satisfied vacuously and thus  $0 \in S$
2. If  $n \in S$  then  $\forall m \in s(n)$  we have by the previous theorem (4.11)
  - a.  $m \in n \xrightarrow{n \text{ is transitive}} m \subseteq n \subseteq n \bigcup \{n\} = s(n) \Rightarrow m \subseteq s(n)$
  - b.  $m = n \Rightarrow m \subseteq n \bigcup \{n\} = s(n) \Rightarrow m \subseteq s(n)$

Using mathematical induction we have then that  $S = \mathbb{N}$  so if  $n \in \mathbb{N} \Rightarrow n \in S \Rightarrow n$  is transitive.  $\square$

**Theorem 4.14.** *If  $n \in \mathbb{N} \Rightarrow n \neq s(n)$*

**Proof.** We prove this by mathematical induction, so let  $S = \{n \in \mathbb{N} \mid n \neq s(n)\}$  then we have

1.  $0 \neq \emptyset \bigcup \{\emptyset\} = s(0)$  [as  $\emptyset \notin \emptyset = 0$ ]  $\Rightarrow 0 \in S$
2. If  $n \in S$  then  $n \neq s(n)$  we prove now by contradiction that  $s(n) \neq s(s(n))$ . So assume that  $s(n) = s(s(n))$  now as  $s(s(n)) = s(n) \bigcup \{s(n)\}$  we have  $s(n) \in s(s(n)) = s(n) \Rightarrow s(n) \in s(n) = n \bigcup \{n\} \xrightarrow{n \neq s(n)} s(n) \in n \xrightarrow{4.13} s(n) \subseteq n \Rightarrow n \bigcup \{n\} \subseteq n \subseteq n \bigcup \{n\} \Rightarrow n = n \bigcup \{n\} = s(n)$  contradicting  $n \neq s(n)$ . So we must have  $s(n) \neq s(s(n)) \Rightarrow s(n) \in S$

Using 4.10 we have then that  $S = \mathbb{N} \Rightarrow$  if  $n \in \mathbb{N} \Rightarrow n \in S \Rightarrow n \neq s(n)$   $\square$

**Theorem 4.15.** *If  $n, m \in \mathbb{N}$  is such that  $s(n) = s(m)$  then  $n = m$*

**Proof.** If  $s(n) = s(m)$  then as  $n \in n \bigcup \{n\} = s(n) \Rightarrow n \in s(m)$  then using 4.11 we have either

1.  $n \in m \xrightarrow{m \text{ is transitive}} n \subseteq m$ . Now  $m \in s(m) = s(n) \Rightarrow m \in s(n)$  and by 4.11 we have then
  - a.  $m \in n$  now we have by transitivity of  $n$  that  $m \subseteq n$  and thus  $n = m$
  - b.  $m = n$  in this case the theorem is proved.
2.  $n = m$  in this case the theorem is proved.  $\square$

We have proved now that  $\mathbb{N}$  full fills the **Peano** axiom's

**Theorem 4.16. (Peano Axioms)**  *$\mathbb{N}$  satisfies the following so-called Peano Axioms*

1.  $0 \in \mathbb{N}$  (see 4.6)
2. If  $n \in \mathbb{N} \Rightarrow s(n) \in \mathbb{N}$  (see 4.9)
3.  $\forall n \in \mathbb{N}$  we have  $s(n) \neq 0$  (see 4.8)

4. If  $X \subseteq \mathbb{N}$  is such that
  - a.  $0 \in X$
  - b.  $n \in X \Rightarrow s(n) \in X$
 then  $X = \mathbb{N}$  (see 4.10)
5. If  $n, m \in \mathbb{N}$  is such that  $s(n) = s(m)$  then  $n = m$  (see 4.15)

**Theorem 4.17.** If  $n \in \mathbb{N} \wedge n \neq 0$  then  $\exists! m \in \mathbb{N}$  such that  $n = s(m)$

**Proof.** Define  $A = \{n \in \mathbb{N} | (n = 0) \vee (\exists! m \in \mathbb{N} \vdash n = s(m))\}$  then we have

1.  $0 \in A$
  2. If  $n \in A \Rightarrow n \in \mathbb{N}$  and  $s(n) = s(n)$  further if  $m \in \mathbb{N}$  such that  $s(n) = s(m) \Rightarrow n = m \Rightarrow \exists! r \in \mathbb{N} [\text{just take } r = n] \vdash s(n) = s(r)$  and thus  $s(n) \in a$
- So if  $n \in \mathbb{N} \wedge n \neq 0 \Rightarrow n \in A \wedge n \neq 0 \Rightarrow \exists! m \in \mathbb{N} \vdash n = s(m)$   $\square$

## 4.2 Recursion

**Theorem 4.18. (Recursion)** If  $A$  is a set,  $a \in A$  and  $f: A \rightarrow A$  is a function then there exists a unique function  $\lambda: \mathbb{N} \rightarrow A$  such that

1.  $\lambda(0) = a$
2.  $\forall n \in \mathbb{N} \vdash \lambda(s(n)) = f(\lambda(n))$

**Proof.** Let  $\mathcal{G} = \{G | G \subseteq \mathbb{N} \times A \text{ such that } (0, a) \in G \text{ and } \forall n \in \mathbb{N} \vdash (n, x) \in G \Rightarrow (s(n), f(x)) \in G\}$ . We have then that  $\mathcal{G} \neq \emptyset$  as for  $\mathbb{N} \times A$  we have  $(0, a) \in \mathbb{N} \times A$  and if  $(n, x) \in \mathbb{N} \times A$  then as  $s(n) \in \mathbb{N} \wedge f(x) \in A \Rightarrow (s(n), f(x)) \in \mathbb{N} \times A \Rightarrow \mathbb{N} \times A \in \mathcal{G}$ . We prove now that

$$\lambda = \bigcap_{G \in \mathcal{G}} G \in \mathcal{G}$$

**Proof.**

1.  $\lambda \subseteq \mathbb{N} \times A$  as  $\forall G \in \mathcal{G}$  we have  $G \subseteq \mathbb{N} \times A$
2.  $\forall G \in \mathcal{G}$  we have  $(0, a) \in G \Rightarrow (0, a) \in \lambda$
3. If  $(n, x) \in \lambda \Rightarrow \forall G \in \mathcal{G}$  we have  $(n, x) \in G \Rightarrow (s(n), f(x)) \in G \Rightarrow (s(n), f(x)) \in \lambda$

So we have that  $\lambda \in \mathcal{G}$   $\square$

We shown now that  $\text{dom}(\lambda) = \mathbb{N}$ . We have

1.  $\text{dom}(\lambda) \subseteq \mathbb{N}$
2.  $0 \in \text{dom}(\lambda)$  as  $(0, a) \in \lambda$
3. If  $n \in \text{dom}(\lambda) \Rightarrow \exists x \in A \vdash (n, x) \in \lambda \Rightarrow (s(n), f(x)) \in \lambda \Rightarrow s(n) \in \text{dom}(\lambda)$

Using mathematical induction (see 4.10) we have then that  $\text{dom}(\lambda) = \mathbb{N}$ . We prove now using mathematical induction that  $\lambda$  is a partial function (from  $\text{dom}(\lambda) = \mathbb{N}$  we have then that  $\lambda$  is a function). So let  $S = \{n \in \mathbb{N} | \exists! x \vdash (n, x) \in \lambda\} \subseteq \mathbb{N}$  then we have

1.  $0 \in S$ . We prove this by contradiction, so let  $(0, a), (0, x) \in \lambda$  with  $a \neq x \Rightarrow (0, a) \neq (0, x)$  Take then  $\lambda' = \lambda \setminus (0, x) \subset \lambda$ . We have then
  - a.  $(0, a) \in \lambda \setminus (0, x) \Rightarrow (0, a) \in \lambda'$
  - b. If  $(n, y) \in \lambda' \setminus \lambda \vdash (n, y) \in \lambda \Rightarrow (s(n), f(y)) \in \lambda \xrightarrow{4.8 \Rightarrow s(n) \neq 0} (s(n), f(y)) \in \lambda'$

this proves that  $\lambda' \in \mathcal{G}$  but then we have  $\lambda \subseteq \lambda' \Rightarrow \lambda \subset \lambda'$  a contradiction. So we must have that  $0 \in S$

2. If  $n \in S \Rightarrow s(n) \in S$ . Again we prove this by contradiction. So assume that  $n \in S \Rightarrow \exists! x \vdash (n, x) \in \lambda \Rightarrow (s(n), f(x)) \in \lambda$ , assume now that  $\exists y \neq f(x) \vdash (s(n), y) \in \lambda$ . Then we can form  $\lambda' = \lambda \setminus (s(n), y)$  where because  $y \neq f(x)$  we have  $\lambda' \subset \lambda$ . For  $\lambda'$  we have then that

$$a. (0, a) \in \lambda \xrightarrow{0 \neq s(n) \Rightarrow (0, a) \neq (s(n), y)} (0, a) \in \lambda'$$

b. If  $(m, z) \in \lambda' \Rightarrow (m, z) \in \lambda \Rightarrow (s(m), f(z)) \in \lambda$  we have then two sub cases

$$i. s(m) = s(n) \text{ then by 4.15 we have that } m = n \Rightarrow (m, z) = (n, z) \xrightarrow{n \in S \Rightarrow (n, z), (n, x) \in \lambda \Rightarrow z = x} z = x \Rightarrow (s(m), f(z)) = (s(n), f(x)) \neq (s(n), y) \Rightarrow (s(m), f(z)) \in \lambda'$$

$$ii. s(m) \neq s(n) \text{ then } (s(m), f(z)) \neq (s(n), y) \Rightarrow (s(m), f(z)) \in \lambda'$$

or in all cases  $(s(m), f(z)) \in \lambda'$

This proves that  $\lambda' \in \mathcal{G}$  but then we have  $\lambda \subseteq \lambda' \xrightarrow{\lambda' \subset \lambda} \lambda \subset \lambda$  a contradiction. So we must conclude that  $\forall y \vdash (s(n), y) \in \lambda$  that  $y = f(x)$  and thus that  $\exists! y (= f(x)) \vdash (s(n), y) \in \lambda$  proving that  $s(n) \in S$

Using mathematical induction we have then that  $S = \mathbb{N}$  and thus that  $\lambda$  is a partial function and because we have already proved that  $\text{dom}(\lambda) = \mathbb{N}$  we have that  $\lambda: \mathbb{N} \rightarrow A$  is a function. By the fact that  $\lambda \in \mathcal{G}$  we have also that

$$1. (0, a) \in \lambda$$

$$2. \text{ if } (n, x) \in \lambda \Rightarrow (s(n), f(x)) \in \lambda$$

proving that  $\lambda$  is the function we search for. Now to prove that  $\lambda$  is unique suppose that there exists a  $\lambda'$  satisfying (1) and (2) above, define  $T = \{n | \lambda(n) = \lambda'(n)\}$ , then we have

$$1. \text{ From } \lambda(0) = a = \lambda'(0) \text{ that } 0 \in T$$

$$2. \text{ If } n \in T \Rightarrow \lambda(n) = \lambda'(n) = d \Rightarrow (n, d) \in \lambda \wedge (n, d) \in \lambda' \xrightarrow{(2) \text{ f is a function}} (s(n), f(d)) \in \lambda \wedge (s(n), f(d)) \in \lambda' \Rightarrow \lambda(s(n)) = f(d) = \lambda'(s(n)) \Rightarrow s(n) \in T$$

By using mathematical induction again we have then that  $T = \mathbb{N}$  so if  $n \in \mathbb{N} \Rightarrow n \in T \Rightarrow \lambda(n) = \lambda'(n) \Rightarrow \lambda = \lambda'$   $\square$

As a corollary of the above theorem we have the following

**Corollary 4.19.** *If  $A$  is a set,  $a \in A$  and  $f: A \rightarrow A$  is a injective function then there exists a unique function  $\lambda: \mathbb{N} \rightarrow A$  such that*

$$1. \lambda(0) = a$$

$$2. \forall n \in \mathbb{N} \models \lambda(s(n)) = f(\lambda(n))$$

$$3. \text{ If } a \notin f(A) \text{ then } \lambda \text{ is injective}$$

**Proof.** By recursion (see 4.18) we have that there exists a unique function  $\lambda: \mathbb{N} \rightarrow A$  such that (1) and (2) are satisfied. Now if  $a \notin f(A)$  then we must show that if  $\lambda(m) = \lambda(n)$  then  $m = n$ . We do this by induction so let  $n \in \mathbb{N}$  and define  $S = \{m \in \mathbb{N} | \text{if } n \in \mathbb{N} \text{ is such that } \lambda(m) = \lambda(n) \Rightarrow m = n\}$  then we have

$$1. \text{ Then } 0 \in S \text{ as we have the following if } \lambda(n) = \lambda(0)$$

$$a. n = 0 \text{ then } 0 \in S$$

$$b. n \neq 0 \text{ then by 4.17 there exists a } k \in \mathbb{N} \text{ such that } n = s(k) \text{ but then } \lambda(n) = \lambda(0) = a \text{ implies } \lambda(s(k)) = a \Rightarrow f(\lambda(k)) = a \text{ which is false as } a \notin f(A) \text{ and thus } \lambda(n) = \lambda(0) \Rightarrow n = 0 \text{ is true or } 0 \in S \text{ again.}$$

$$2. \text{ If } m \in S \text{ then if } k \in \mathbb{N} \vdash \lambda(m) = \lambda(k) \Rightarrow m = k. \text{ Suppose now that there exists a } n \in \mathbb{N} \vdash \lambda(s(m)) = \lambda(n) \text{ then we have the following possibilities}$$

$$a. n = 0 \text{ then } a = \lambda(0) = \lambda(s(m)) = f(\lambda(m)) \text{ is not allowed because } a \notin f(A) \text{ so } \lambda(s(m)) = \lambda(n) \text{ is false and thus } \lambda(s(m)) = \lambda(n) \Rightarrow s(m) = n \text{ is true and thus } s(m) \in S_n$$

- b.  $n \neq 0$  then by 4.17 we have that  $\exists k \in \mathbb{N} \vdash n = s(k)$  so from  $\lambda(s(m)) = \lambda(n)$  we have  $\lambda(s(m)) = \lambda(s(k)) \Rightarrow f(\lambda(m)) = f(\lambda(k)) \xrightarrow{f \text{ is injective}} \lambda(m) = \lambda(k) \xrightarrow{m \in S} m = k \Rightarrow s(m) = s(k) \Rightarrow s(m) = n \Rightarrow s(m) \in S$

Using 4.10 we have then that  $S = \mathbb{N}$  and thus if there exists a  $n \in \mathbb{N}$  and a  $m \in \mathbb{N}$  with  $s(m) = s(n) \Rightarrow n \in S \Rightarrow m = n$  proving injectivity.  $\square$

**Theorem 4.20.** Suppose  $A$  is a set  $a \in A$  and  $f: A \rightarrow A$  a function, then there exists a **unique** function  $\varphi: \mathbb{N} \rightarrow A$  such that  $\varphi(0) = a$  and  $\varphi \circ s = f \circ \varphi$ .

**Proof.** Define the function  $\sigma: \mathbb{N} \times A \rightarrow A$  by  $(n, x) \rightarrow (s(n), f(x))$ . Given  $a \in A$  we define a subset  $R \subseteq \mathbb{N} \times A$  to be a-closed if the following is true

1.  $(0, a) \in R$
2.  $\sigma(R) \subseteq R$

We have that  $\mathbb{N} \times A$  is a-closed for

1.  $(0, a) \in \mathbb{N} \times A$
2. If  $(m, y) \in \sigma(\mathbb{N} \times A) \Rightarrow \exists (n, x) \in \mathbb{N} \times A \vdash (m, y) = (s(n), f(x)) \in \mathbb{N} \times A \Rightarrow \sigma(\mathbb{N} \times A) \subseteq \mathbb{N} \times A$ .

If  $\mathcal{R}$  is a set of a-closed subsets of  $\mathbb{N} \times A$  then we have

1.  $\forall R \in \mathcal{R}$  we have  $(0, a) \in R \Rightarrow (0, a) \in \bigcap_{R \in \mathcal{R}} R$
2. if  $y \in \sigma(\bigcap_{R \in \mathcal{R}} R) \Rightarrow \exists x \in \bigcap_{R \in \mathcal{R}} R \vdash y = \sigma(x) \Rightarrow \exists x \vdash (\forall R \in \mathcal{R} \vdash x \in R) \wedge y = \sigma(x) \Rightarrow \forall R \in \mathcal{R} \vdash y \in \sigma(R) \Rightarrow y \in \bigcap_{R \in \mathcal{R}} \sigma(R) \Rightarrow \sigma(\bigcap_{R \in \mathcal{R}} R) \subseteq \bigcap_{R \in \mathcal{R}} \sigma(R) \xrightarrow{\sigma(R) \subseteq R} \sigma(\bigcap_{R \in \mathcal{R}} R) \subseteq \bigcap_{R \in \mathcal{R}} R$

proving that  $\bigcap_{R \in \mathcal{R}} R$  is also a-closed. Define now the set  $\mathcal{R}_a = \{R \in \mathcal{P}(\mathbb{N} \times A) \mid R \text{ is a-closed}\}$  set (because  $\subseteq \mathcal{P}(\mathbb{N} \times A)$  which is set because  $\mathbb{N} \times A$  is a set and 1.63) and  $R_a = \bigcap_{R \in \mathcal{R}_a} R$  then we have by the above that  $R_a$  is a-closed and is further the smallest a-closed set. We have then

- a)  $(0, a) \in \{(0, a)\} \cup \sigma(R_a)$
- b)  $\sigma(\{(0, a)\} \cup \sigma(R_a)) \subseteq \sigma(\{(0, a)\}) \cup \sigma(\sigma(R_a)) \subseteq_{\sigma(0, a) \in R_a} R_a \cup \sigma(\sigma(R_a)) \subseteq_{\sigma(R_a) \subseteq R_a} R_a \cup R_a = R_a$

proving that  $\{(0, a)\} \cup R_a$  is a-closed and then from the minimality of  $R_a$  we have

$$R_a \subseteq \{(0, a)\} \cup \sigma(R_a)$$

As  $\{(0, a)\} \subseteq R_a$  and  $\sigma(R_a) \subseteq R_a$  we have

$$\{(0, a)\} \cup \sigma(R_a) \subseteq R_a$$

and thus we finally have

$$R_a = \{(0, a)\} \cup \sigma(R_a)$$

We say now that  $n \in \mathbb{N}$  is a-paired-up iff

- a) There exists  $y \in A$  such that  $(n, y) \in R_a$
- b) If also  $(n, y') \in R_a \Rightarrow y = y'$

Let's define now the set of a-paired-up numbers  $P_a = \{n \in \mathbb{N} \mid n \text{ is a-paired-up}\} \subseteq \mathbb{N}$  then we proof that

1.  $0 \in P_a$

**Proof.** First  $(0, a) \in R_a$ . Second if  $(0, a') \in R_a$  with  $a \neq a' \Rightarrow (0, a) \neq (0, a') \xrightarrow{(0, a') \in R_a = \{(0, a)\} \cup \sigma(R_a)} (0, a') \in \sigma(R_a)$  so there exists a  $(n, a'') \in R_a$  with  $(0, a') = \sigma(n, a'') = (s(n), f(a''))$  giving  $0 = s(n)$  a contradiction. So we must conclude that  $a = a'$ .  $\square$

2. If  $n \in P_a \Rightarrow s(n) \in P_a$

**Proof.** As  $n \in P_a$  then  $n$  is a-paired-up so there exists a **unique**  $y \in A$  such that  $(n, y) \in R_a$   $\xrightarrow[R_a \text{ is a-closed}]{} \sigma(n, y) \in R_a \Rightarrow (s(n), f(y)) \in R_a$ . Suppose now that  $(s(n), y') \in R_a = \{(0, a)\} \cup \sigma(R_a) \xrightarrow[s(n) \neq 0]{} (s(n), y') \in \sigma(R_a) \Rightarrow \exists (m, c) \in R_a \vdash (s(n), y') = \sigma(m, c) = (s(m), f(c)) \Rightarrow s(n) = s(m) \wedge y' = f(c) \xrightarrow[4.15]{} n = m \wedge y' = f(c) \Rightarrow (n, c) \in R_a \wedge y' = f(c) \xrightarrow[(n, y) \in R_a]{} (n, c), (n, y) \in R_a \wedge y' = f(c) \xrightarrow[n \text{ is a-paired-up}]{} c = y \wedge y' = f(c) \Rightarrow y' = f(y)$ . So  $s(n)$  is a-paired-up and thus  $s(n) \in P_a$   $\square$

Using mathematical induction (see 4.10) we have then that  $P_a = \mathbb{N}$ . Define now  $\varphi: \mathbb{N} \rightarrow A$  with  $\varphi = R_a$  then we have that if  $n \in \mathbb{N} \Rightarrow n \in P_a$  so there exists a  $y \in A \vdash (n, a) \in R_a = \varphi$  and thus  $\text{dom}(\varphi) = \mathbb{N}$ . Second if  $(n, y), (n, y') \in \varphi$  then as  $n \in \mathbb{N} \Rightarrow n \in P_a$  we have that  $n$  is a-paired-up and thus  $y = y'$ . So we conclude that  $\varphi: \mathbb{N} \rightarrow A$  is indeed a function. Now as  $R_a$  is a-closed we have  $(0, a) \in R_a = \varphi \Rightarrow \varphi(0) = a$  and thus

$$\varphi(0) = a$$

Second if  $n \in \mathbb{N}$  then  $n \in P_a$  so there exists a  $y \in A$  such that  $(n, y) \in R_a = \varphi \Rightarrow y = \varphi(n)$  then as  $R_a$  is a-closed we have  $(s(n), f(y)) \in R_a = \varphi$  and thus  $(\varphi \circ s)(n) = \varphi(s(n)) = f(y) = f(\varphi(n)) = (f \circ \varphi)(n)$  proving that

$$\varphi \circ s = f \circ \varphi$$

So we have found that the function  $\varphi: \mathbb{N} \rightarrow A$  satisfies the requirements of the theorem. Let's now prove that it is unique. Assume that there exists also a  $\varphi': \mathbb{N} \rightarrow A$  satisfying

$$\begin{aligned} \varphi'(0) &= a \\ \varphi' \circ s &= f \circ \varphi \end{aligned}$$

Take then  $E = \{n \in \mathbb{N} \mid \varphi(n) = \varphi'(n)\}$  then we have

1.  $\varphi(0) = a = \varphi'(0) \Rightarrow 0 \in E$
2. If  $n \in E$  then for  $s(n)$  we have  $\varphi(s(n)) = (\varphi \circ s)(n) = (f \circ \varphi)(n) = f(\varphi(n)) \xrightarrow[n \in E \Rightarrow \varphi(n) = \varphi'(n)]{} f(\varphi'(n)) = (f \circ \varphi')(n) = (\varphi' \circ s)(n) = \varphi'(s(n))$  and thus  $s(n) \in E$

Using mathematical induction (see 4.10) we have then that  $E = \mathbb{N}$  so  $\forall n \in \mathbb{N}$  we have  $n \in E \Rightarrow \varphi(n) = \varphi'(n)$  and thus  $\varphi = \varphi'$   $\square$

This theorem allows us to iterate the application of a function in the following way: If  $a \in A$  and  $f: A \rightarrow A$  is a function then there exists a function  $\varphi: \mathbb{N} \rightarrow A$  such that

$$\begin{aligned} \varphi(0) &= a \\ \varphi(1) &= \varphi(s(0)) = f(\varphi(0)) = f(a) \\ \varphi(2) &= \varphi(s(1)) = f(\varphi(1)) = f(f(a)) \\ \varphi(3) &= \varphi(s(2)) = f(\varphi(2)) = f(f(f(a))) \\ &\dots \\ \varphi(n) &= \overbrace{f(f(\dots(f(a))))}^{n \text{ times}} \end{aligned}$$

As application of the above theorem we have the following iteration theorem

**Theorem 4.21. (Iteration)** *Let  $A$  be a set and  $f: A \rightarrow A$  be a function. Then  $\forall n \in \mathbb{N}$  there exists a function  $(f)^n: A \rightarrow A$  such that*

1.  $(f)^0 = 1_A$
2.  $(f)^{s(n)} = f \circ (f)^n$

**Proof.** If  $A = \emptyset$  then  $\emptyset: \emptyset \rightarrow \emptyset$  then as  $1_{\emptyset}: \emptyset \rightarrow \emptyset$  is also the empty function  $\emptyset$  then  $(\emptyset)^n = \emptyset$  satisfies (1), (2). If  $A \neq \emptyset$  then let  $n \in \mathbb{N}$  and  $a \in A$  and use the previous theorem 4.20 to find a function  $\varphi_a: \mathbb{N} \rightarrow A$  such that  $\varphi_a(0) = a$  and  $\varphi_a \circ s = f \circ \varphi_a$ . Define now  $(f)^n: A \rightarrow A$  by  $a \rightarrow \varphi_a(n)$  then we have

1.  $\forall a \in A$  that  $(f)^0(a) = \varphi_a(0) = a \Rightarrow (f)^0 = i_A$
2.  $\forall a \in A$  then  $(f)^{s(n)}(a) = \varphi_a(s(n)) = f(\varphi_a(n)) = f((f)^n(a)) = (f \circ (f)^n)(a) \Rightarrow (f)^{s(n)} = f \circ (f)^n$   $\square$

As a illustration of this iteration let  $f: A \rightarrow A$  then we have

$$\begin{aligned}
 (f)^0 &= 1_A \\
 (f)^1 &= (f)^{s(0)} = f \circ (f)^0 = f \circ i_A = f \\
 (f)^2 &= (f)^{s(1)} = f \circ (f)^1 = f \circ f \\
 (f)^3 &= (f)^{s(2)} = f \circ (f)^2 = f \circ f \circ f \\
 &\dots \\
 (f)^n &= \overbrace{f \circ \dots \circ f}^{\text{n times}}
 \end{aligned}$$

**Example 4.22.** Let  $\langle A, \oplus \rangle$  be a group and  $a \in A$  define then  $\oplus_a: A \rightarrow A$  by  $x \rightarrow \oplus_a(x) = x \oplus a$  we define then given  $n \in \mathbb{N}$   $a \langle \oplus \rangle n = (\oplus_a)^n(\nu)$  where  $\nu$  is the neutral element in the group . We have then that

$$\begin{aligned}
 a \langle \oplus \rangle 0 &= (\oplus_a)^0(\nu) = i_A(\nu) = \nu \\
 a \langle \oplus \rangle s(n) &= ((\oplus_a)^n(\nu)) \oplus a \\
 &= (a \langle \oplus \rangle n) \oplus a
 \end{aligned}$$

Sometimes we consider a group to be additive or multiplicative this is either noted as  $\langle A, + \rangle$  with neutral element 0 or  $\langle A, \cdot \rangle$  with neutral element 1 then we note  $a \langle + \rangle n$  as  $a \cdot n$  as and  $a \langle \cdot \rangle n$  as  $a^n$  we have then

1. Additive group  $\langle A, + \rangle$  with neutral element 0

$$\begin{aligned}
 a \cdot 0 &= 0 \\
 a \cdot s(n) &= (a \cdot n) + a
 \end{aligned}$$

so we have

$$\begin{aligned}
 a \cdot 0 &= 0 \\
 a \cdot 1 &= a \cdot s(0) = (a \cdot 0) + a = 0 + a = a \\
 a \cdot 2 &= a \cdot s(1) = (a \cdot 1) + a = a + a \\
 a \cdot 3 &= a \cdot s(2) = (a \cdot 2) + a = (a + a) + a \\
 &\dots
 \end{aligned}$$

2. Multiplicative group  $\langle A, \cdot \rangle$  with neutral element 1

$$\begin{aligned}
 a^0 &= 1 \\
 a^{s(n)} &= (a^n) \cdot a
 \end{aligned}$$

so we have

$$\begin{aligned}
 a^0 &= 1 \\
 a^1 &= a^{s(0)} = (a^0) \cdot a = 1 \cdot a = a \\
 a^2 &= a^{s(1)} = (a^1) \cdot a = a \cdot a \\
 a^3 &= a^{s(2)} = (a \cdot a) \cdot a \\
 &\dots
 \end{aligned}$$

**Definition 4.23.** If  $\langle F, +, \cdot \rangle$  is a field with multiplicative unit  $u$  and additive neutral element  $e$ , let  $n \in \mathbb{N}_0$  define take then  $f: F \rightarrow F$  defined by  $x \rightarrow x + u$  then  $n \cdot u = (f)^n(e)$ .

**Example 4.24.** If  $\langle F, +, \cdot \rangle$  is a field with multiplicative unit  $u$  then we have

$$\begin{aligned} 0 \cdot u &= (f)^0(e) = 1_F(e) = e \\ 1 \cdot u &= (f)^1(e) = f^{(s(0))}(e) = f(f^0(e)) = f(e) = e + u = u \\ 2 \cdot u &= (f)^2(e) = f^{s(1)}(e) = f(f^1(e)) = f(u) = u + u \\ 3 \cdot u &= (f)^3(e) = f^{s(2)}(e) = f(f^2(e)) = f(u + u) = u + u + u \\ &\dots \end{aligned}$$

**Definition 4.25.** If  $\langle F, +, \cdot \rangle$  is a field with multiplicative unit  $u$  and neutral element  $e$  then  $F$  is of characterization zero if  $\forall n \in \mathbb{N}_0$  we have  $n \cdot u \neq e$

**Theorem 4.26. (Recursion on  $\mathbb{N}$  - Step Form)** Let  $A$  be a set,  $a \in A$  and  $g: \mathbb{N} \times A \rightarrow A$  a function. Then there exists a unique function  $f: \mathbb{N} \rightarrow A$  satisfying

$$\begin{aligned} f(0) &= a \\ \forall n \in \mathbb{N} \text{ we have } f(s(n)) &= g(n, f(n)) \end{aligned}$$

**Proof.** First define the projection functions  $\pi_1$  and  $\pi_2$  as follows

$$\begin{aligned} \pi_1: \mathbb{N} \times A &\rightarrow \mathbb{N} \text{ by } (n, x) \rightarrow \pi_1(n, x) = n \\ \pi_2: \mathbb{N} \times A &\rightarrow \mathbb{N} \text{ by } (n, x) \rightarrow \pi_2(n, x) = x \end{aligned}$$

we have then  $\forall (n, x) \in \mathbb{N} \times A$  that  $(n, x) = (\pi_1(n, x), \pi_2(n, x))$ . Next define  $\gamma$  by

$$\gamma: \mathbb{N} \times A \rightarrow \mathbb{N} \times A \text{ by } (n, x) \rightarrow \gamma(n, x) = (s(n), g(n, x))$$

By using 4.21 we have  $\forall n \in \mathbb{N}$  the existence of a function  $(\gamma)^n: \mathbb{N} \times A \rightarrow \mathbb{N} \times A$  such that

$$\begin{aligned} (\gamma)^0 &= 1_{\mathbb{N} \times A} \\ (\gamma)^{s(n)} &= \gamma \circ (\gamma)^n \end{aligned}$$

We prove now by induction that  $\forall n \in \mathbb{N}$  we have  $\pi_1((\gamma)^n(0, a)) = n$ . So define  $B = \{n \in \mathbb{N} \mid \pi_1((\gamma)^n(0, a)) = n\}$  then we have

1.  $\pi_1((\gamma)^0(0, a)) = \pi_1(i_{\mathbb{N} \times A}(0, a)) = \pi_1(0, a) = 0$  and thus we have that  $0 \in B$ .
2. If  $n \in B$  then  $\pi_1((\gamma)^{s(n)}(0, a)) = \pi_1(\gamma((\gamma)^n(0, a))) = \pi_1(\gamma(\pi_1((\gamma)^n(0, a)), \pi_2((\gamma)^n(0, a)))) \stackrel{n \in B}{=} \pi_1(\gamma(n, \pi_2((\gamma)^n(0, a)))) = \pi_1(s(n), g(n, \pi_2((\gamma)^n(0, a)))) = s(n)$  and thus  $s(n) \in B$

Using induction (see 4.10) we have then that  $B = \mathbb{N}$  and thus proved our statement.

Define now  $f: \mathbb{N} \rightarrow A$  by  $n \rightarrow f(n) = \pi_2((\gamma)^n(0, a))$  then we have

1.  $f(0) = \pi_2((\gamma)^0(0, a)) = \pi_2(i_{\mathbb{N} \times A}(0, a)) = \pi_2(0, a) = a$
2. If  $n \in \mathbb{N}$  then  $f(s(n)) = \pi_2((\gamma)^{s(n)}(0, a)) = \pi_2(\gamma((\gamma)^n(0, a))) = \pi_2(\gamma(\pi_1((\gamma)^n(0, a)), \pi_2((\gamma)^n(0, a)))) = \pi_2(\gamma(n, \pi_2((\gamma)^n(0, a)))) = \pi_2(s(n), g(n, \pi_2((\gamma)^n(0, a)))) = g(n, \pi_2((\gamma)^n(0, a))) = g(n, f(n))$

proving that  $f$  is the function we search for. We are thus left with proving uniqueness. So assume that there exists a  $f': \mathbb{N} \rightarrow A$  such that

$$\begin{aligned} f'(0) &= a \\ f'(s(n)) &= g(n, f'(n)) \end{aligned}$$

and define then  $C = \{n \in \mathbb{N} \mid f(n) = f'(n)\}$  then we have

1.  $f(0) = a = f'(0)$  and thus  $0 \in C$
2. If  $n \in \mathbb{N}$  then  $f(s(n)) = g(n, f(n)) = g(n, f'(n)) = f'(s(n))$  and thus  $s(n) \in C$

Using mathematical induction (see 4.10) we have that  $C = \mathbb{N}$  and thus  $f = f'$   $\square$

### 4.3 Arithmetic of the natural numbers

**Definition 4.27. (Addition)** Let  $m, n \in \mathbb{N}$  then we define  $+: \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$  by  $(n, m) \rightarrow n+m = (s)^m(n)$  (here  $s$  is the successor function)

**Example 4.28.** If  $n \in \mathbb{N}$  then we have

$$\begin{aligned} n+0 &= (s)^0(n) = i_{\mathbb{N}}(n) = n \\ n+1 &= (s)^1(n) = (s)^{s(0)}(n) = (s \circ (s)^0)(n) = s(i_{\mathbb{N}}(n)) = s(n) \\ n+2 &= (s)^2(n) = (s)^{s(1)}(n) = (s \circ (s)^1)(n) = s(s(n)) \\ &\dots \end{aligned}$$

Using the definition of  $+$  we immediately have

**Theorem 4.29.** If  $n \in \mathbb{N}$  then  $n+0 = 0+n = n$

**Proof.**

1.  $n+0 = (s)^0(n) = 1_{\mathbb{N}}(n) = n$
2. We prove this by mathematical induction, define  $S = \{n \in \mathbb{N} | 0+n = n\} \subseteq \mathbb{N}$  then we have
  - a.  $0+0 \stackrel{(1)}{=} 0 \Rightarrow 0 \in S$
  - b. If  $n \in S \Rightarrow 0+n = n$  now for  $s(n)$  we have  $0+s(n) = (s)^{s(n)}(0) = (s \circ s^{(n)})(0) = s(s^{(n)}(0)) = s(0+n) = s(n) \Rightarrow s(n) \in S$

Using mathematical induction 4.10 we have  $S = \mathbb{N}$  and thus if  $n \in \mathbb{N} \Rightarrow n \in S \Rightarrow 0+n = n$   $\square$

**Theorem 4.30.** If  $n \in \mathbb{N}$  then  $s(n) = n+1 = 1+n$

**Proof.**

1.  $n+1 = (s)^1(n) = (s)^{s(0)}(n) = (s \circ (s)^0)(n) = (s \circ i_{\mathbb{N}})(n) = s(n)$
2.  $1+n = s(n)$  is proved by induction, so define  $S = \{n \in \mathbb{N} | 1+n = s(n)\}$  then we have
  - a.  $1+0 \stackrel{4.29}{=} 1 = s(0) \Rightarrow 0 \in S$
  - b. If  $n \in S \Rightarrow 1+n = s(n)$ . Now  $1+s(n) = (s)^{s(n)}(1) = (s \circ (s)^n)(1) = s((s)^n(1)) = s(1+n) = s(s(n))$  and thus  $s(n) \in S$

from 4.10 we have then that  $S = \mathbb{N} \Rightarrow n \in \mathbb{N} \Rightarrow n \in S \Rightarrow 1+n = s(n)$   $\square$

**Lemma 4.31.** If  $n, m \in \mathbb{N}$  then  $n+s(m) = s(n+m)$

**Proof.**  $n+s(m) = (s)^{s(m)}(n) = (s \circ (s)^m)(n) = s((s)^m(n)) = s(n+m)$   $\square$

**Theorem 4.32. (Associativity)** If  $n, m, k \in \mathbb{N}$  then  $(n+m)+k = n+(m+k)$

**Proof.** The proof is by mathematical induction given  $n, m \in \mathbb{N}$  define  $S_{n,m} = \{k \in \mathbb{N} | (n+m)+k = n+(m+k)\}$  then we have

1.  $(n+m)+0 \stackrel{4.29}{=} n+m \stackrel{4.29}{=} n+(m+0) \Rightarrow 0 \in S_{n,m}$
2. If  $k \in S_{n,m} \Rightarrow (n+m)+k = n+(m+k)$ . We have then  $(n+m)+s(k) \stackrel{4.31}{=} s((n+m)+k) = s(n+(m+k)) \stackrel{4.31}{=} n+s(m+k) \stackrel{4.31}{=} n+(m+s(k))$  and thus  $s(k) \in S_{n,m}$

So if  $n, m, k \in \mathbb{N}$  then  $k \in S_{n,m} \Rightarrow (n+m)+k = n+(m+k)$   $\square$

**Theorem 4.33. (Commutativity)** If  $n, m \in \mathbb{N}$  then  $n+m = m+n$

**Proof.** Again we prove this by induction. So let  $n \in \mathbb{N}$  define then  $S_n = \{m \in \mathbb{N} | n+m = m+n\}$  then we have

1.  $n+0 = n = 0+n$  (see 4.29) and thus  $0 \in S_n$

2. If  $m \in S_n \Rightarrow n+m = m+n$ . We have then  $n+s(m) \stackrel{4.31}{=} s(n+m) = s(m+n) \stackrel{4.30}{=} 1+(m+n) \stackrel{4.32}{=} (1+n)+m \stackrel{4.30}{=} s(n)+m$  so we have that  $s(m) \in S_n$

Using 4.10 we have then that  $S_n = \mathbb{N}$ . So if  $n, m \in \mathbb{N} \Rightarrow m \in S_n \Rightarrow n+m = m+n$   $\square$

**Corollary 4.34.**  $\langle \mathbb{N}, + \rangle$  forms a abelian semi-group

**Proof.** This follows from 4.32, 4.29 and 4.33  $\square$

**Definition 4.35. (Multiplication)** Given  $n \in \mathbb{N}$  define  $\alpha_n: \mathbb{N} \rightarrow \mathbb{N}$  defined by  $m \rightarrow \alpha_n(m) = n+m$ . Then we define  $\cdot: \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$  by  $(n, m) \rightarrow n \cdot m = (\alpha_n)^m(0)$

**Example 4.36.** We have the following examples to see how multiplication works by repeating summation

$$\begin{aligned} 2 \cdot 0 &= (\alpha_2)^0(0) = i_{\mathbb{N}}(0) = 0 \\ 2 \cdot 1 &= (\alpha_2)^1(0) = (\alpha_2)^{s(0)}(0) = (\alpha_2 \circ (\alpha_2)^0)(0) = \alpha_2(0) = 2+0 = 2 \\ 2 \cdot 2 &= (\alpha_2)^2(0) = (\alpha_2)^{s(1)}(0) = (\alpha_2((\alpha_2)^1(0))) = \alpha_2(2) = 2+2 = 4 \end{aligned}$$

**Theorem 4.37. (Absorbing Element)** If  $n \in \mathbb{N}$  then  $n \cdot 0 = 0 = 0 \cdot n$

**Proof.**

1.  $n \cdot 0 = (\alpha_n)^0(0) = 1_{\mathbb{N}}(0) = 0$
2. We prove by induction that  $0 \cdot n = 0$  so define  $S = \{n \in \mathbb{N} \mid 0 \cdot n = 0\} \subseteq \mathbb{N}$  then we have
  - a.  $0 \cdot 0 \stackrel{(1)}{=} 0 \Rightarrow 0 \in S$
  - b. If  $n \in S$  then  $0 \cdot n = 0$  Now  $0 \cdot s(n) = (\alpha_0)^{s(n)}(0) = (\alpha_0 \circ (\alpha_0)^n)(0) = \alpha_0((\alpha_0)^n(0)) = \alpha_0(0 \cdot n) \stackrel{4.29}{=} \alpha_0(0) = 0+0 = 0$  and thus  $s(n) \in S$

proving by 4.10 that  $S = \mathbb{N}$   $\square$

**Theorem 4.38. (Neutral Element)** If  $n \in \mathbb{N}$  then  $n \cdot 1 = n = 1 \cdot n$

**Proof.**

1.  $n \cdot 1 = (\alpha_n)^1(0) = (\alpha_n)^{s(0)}(0) = \alpha_n((\alpha_n)^0(0)) = \alpha_n(0) = n+0 \stackrel{4.29}{=} n$
2. We prove by induction that  $1 \cdot n = n$  so define  $S = \{n \in \mathbb{N} \mid 1 \cdot n = n\}$  then we have
  - a.  $1 \cdot 0 \stackrel{4.37}{=} 0 \Rightarrow 0 \in S$
  - b. If  $n \in S \Rightarrow 1 \cdot n = n$  Now  $1 \cdot s(n) = (\alpha_1)^{s(n)}(0) = \alpha_1((\alpha_1)^n(0)) = \alpha_1(1 \cdot n) = \alpha_1(n) = 1+n \stackrel{4.30}{=} s(n)$  so  $s(n) \in S$

Using 4.10 we have then  $S = \mathbb{N}$   $\square$

**Theorem 4.39.** If  $n, m \in \mathbb{N}$  then  $n \cdot s(m) = n+n \cdot m \stackrel{4.33}{=} n \cdot m+n$

**Proof.**  $n \cdot s(m) = (\alpha_n)^{s(m)}(0) = \alpha_n((\alpha_n)^m(0)) = \alpha_n(n \cdot m) = n+n \cdot m$   $\square$

**Theorem 4.40. (Distributivity)**  $\forall n, m, k \in \mathbb{N}$  we have  $(n+m) \cdot k = n \cdot k + m \cdot k$

**Proof.** If  $n, m \in \mathbb{N}$  define  $S_{n,m} = \{k \in \mathbb{N} \mid (n+m) \cdot k = n \cdot k + m \cdot k\}$  then we have

1.  $(n+m) \cdot 0 \stackrel{4.37}{=} 0 \stackrel{4.29}{=} 0+0 \stackrel{4.37}{=} n \cdot 0 + m \cdot 0 \Rightarrow 0 \in S_{n,m}$
2. Assume that  $k \in S_{n,m}$  then  $(n+m) \cdot k = n \cdot k + m \cdot k$ . Now  $(n+m) \cdot s(k) \stackrel{4.39}{=} (n+m) \cdot k + (n+m) \cdot (n \cdot k + m \cdot k) \stackrel{4.32}{=} (n \cdot k + n) + (m \cdot k + m) \stackrel{4.39}{=} n \cdot s(k) + m \cdot s(k) \Rightarrow s(k) \in S_{n,m}$

If now  $n, m, k \in \mathbb{N} \Rightarrow k \in S_{n,m} \Rightarrow (n+m) \cdot k = n \cdot k + m \cdot k$   $\square$

**Theorem 4.41. (Commutativity)** *If  $n, m \in \mathbb{N}$  then  $n \cdot m = m \cdot n$*

**Proof.** We prove this by induction, let  $S_n = \{m \in \mathbb{N} \mid n \cdot m = m \cdot n\}$  then we have

1. Using 4.37 we have  $n \cdot 0 = 0 = 0 \cdot n \Rightarrow 0 \in S_n$
2. If  $m \in S_n \Rightarrow n \cdot m = m \cdot n$ . Now  $n \cdot s(m) = n + n \cdot m \underset{m \in S_n}{=} n + m \cdot n \underset{4.38}{=} 1 \cdot n + m \cdot n \underset{4.40}{=} (1+m) \cdot n \underset{4.30}{=} s(m) \cdot n \Rightarrow s(m) \in S_n$

This proves by 4.10 that  $S_n = \mathbb{N}$  and thus if  $n, m \in \mathbb{N} \Rightarrow m \in S_n \Rightarrow n \cdot m = m \cdot n$   $\square$

**Theorem 4.42. (Associativity)** *If  $n, m, k \in \mathbb{N}$  then  $(n \cdot m) \cdot k = n \cdot (m \cdot k)$*

**Proof.** Given  $n, m \in \mathbb{N}$  define  $S_{n,m} = \{k \in \mathbb{N} \mid (n \cdot m) \cdot k = n \cdot (m \cdot k)\}$  then we have

1.  $(n \cdot m) \cdot 0 \underset{4.37}{=} 0 \underset{4.37}{=} n \cdot 0 = n \cdot (m \cdot 0) \Rightarrow 0 \in S_{n,m}$
2. If  $k \in S_{n,m}$  then  $(n \cdot m) \cdot k = n \cdot (m \cdot k)$ . Now  $(n \cdot m) \cdot s(k) \underset{4.39}{=} (n \cdot m) \cdot k + n \cdot m = n \cdot (m \cdot k) + n \cdot m \underset{4.41}{=} (m \cdot k) \cdot n + m \cdot n \underset{4.40}{=} (m \cdot k + m) \cdot n \underset{4.39}{=} (m \cdot s(k)) \cdot n \underset{4.41}{=} n \cdot (m \cdot s(k)) \Rightarrow s(k) \in S_{n,m}$

Using mathematical induction 4.10 we have  $S_{n,m} = \mathbb{N}$  so if  $n, m, k \in \mathbb{N} \Rightarrow k \in S_{n,m} \Rightarrow (n \cdot m) \cdot k = n \cdot (m \cdot k)$   $\square$

**Corollary 4.43.**  $\langle \mathbb{N}, \cdot \rangle$  is a abelian semi-group

**Proof.** This follows from 4.41, 4.42 and 4.38.  $\square$

**Theorem 4.44.** *If  $n, m, k \in \mathbb{N}$  is such that  $n+k = m+k \Rightarrow n = m$ , (and thus we have  $n = m \Leftrightarrow n+k = m+k$ )*

**Proof.** We prove this by induction so let  $S_{n,m} = \{k \in \mathbb{N} \mid n+k = m+k \Rightarrow n = m\}$  then we have

1.  $n+0 = m+0 \underset{4.29}{\Rightarrow} n = m \Rightarrow 0 \in S_{n,m}$
2. If  $k \in S_{n,m} \Rightarrow n+k = m+k \Rightarrow n = m$ . For  $s(k)$  if we have  $n+s(k) = m+s(k) \underset{4.31}{\Rightarrow} s(n+k) = s(m+k) \underset{4.15}{\Rightarrow} n+k = m+k \Rightarrow n = m \Rightarrow s(k) \in S_{n,m}$

So if  $n, m, k \in \mathbb{N} \Rightarrow k \in S_{n,m} \Rightarrow n+k = m+k \Rightarrow n = m$   $\square$

## 4.4 Order relation on the natural numbers

**Definition 4.45.** Define the relation  $\leq \in \mathbb{N} \times \mathbb{N}$  by  $\leq = \{(n, m) \in \mathbb{N} \times \mathbb{N} \mid n \in m \vee n = m\}$  so we say that  $n \leq m$  iff  $n \in m \vee n = m$ . If  $n \leq m$  and  $n \neq m$  then we say that  $n < m$ .

**Theorem 4.46.**  $\langle \mathbb{N}, \leq \rangle$  is a partially ordered set

**Proof.**

1. **(reflectivity)** If  $m \in \mathbb{N} \Rightarrow m = m \Rightarrow m \leq m$
2. **(anti-symmetry)** If  $n, m \in \mathbb{N}$  and  $n \leq m \wedge m \leq n \Rightarrow (n \in m \vee n = m) \wedge (m \in n \vee n = m) \Rightarrow (n \in m \wedge m \in n) \vee (n = m)$  so we have the following cases
  - a.  $n = m \Rightarrow$  anti-symmetry is proved
  - b.  $n \in m \wedge m \in n$  then by 4.13  $n \subseteq m \wedge m \subseteq n \Rightarrow n = m$  proving again anti-symmetry
3. **(transitivity)** If  $n \leq m \wedge m \leq k$  then we have the following possibilities
  - a.  $n \in m \wedge m \in k$  then by 4.13  $n \subseteq m \wedge m \subseteq k \Rightarrow n \subseteq k \Rightarrow n \leq k$
  - b.  $n \in m \wedge m = k$  then by 4.13  $n \subseteq m \wedge m = k \Rightarrow n \subseteq k \Rightarrow n \leq k$
  - c.  $n = m \wedge m \in k$  then by 4.13  $n = m \wedge m \subseteq k \Rightarrow n \subseteq k \Rightarrow n \leq k$

d.  $n = m \wedge m = k$  then  $n = k \Rightarrow n \leq k$

So in all cases we have  $n \leq k$  proving transitivity.  $\square$

**Theorem 4.47.** *If  $n \in \mathbb{N} \Rightarrow 0 \leq n$*

**Proof.** We prove this by mathematical induction, so let  $S = \{n \in \mathbb{N} \mid 0 \leq n\} \subseteq \mathbb{N}$  then we have

1.  $0 = 0 \Rightarrow 0 \leq 0 \Rightarrow 0 \in S$
2. Suppose  $n \in S$  then  $0 \leq n$  now by definition  $s(n) = n \cup \{n\} \Rightarrow n \in s(n) \Rightarrow n \leq s(n) \stackrel{4.46 \wedge 0 \leq n}{=} 0 \leq s(n) \Rightarrow s(n) \in S$

Using 4.10 we have then  $S = \mathbb{N}$  proving our theorem.  $\square$

**Theorem 4.48.** *If  $n \in \mathbb{N} \Rightarrow n < s(n)$*

**Proof.** From  $n \in n \cup \{n\} = s(n)$  we have  $n \leq s(n) \stackrel{4.14}{\Rightarrow} n < s(n)$   $\square$

**Theorem 4.49.** *If  $n \in \mathbb{N}$  then  $k \in n \Leftrightarrow k < n$*

**Proof.**

1. **( $k \in n \Rightarrow k < n$ )** We prove this by induction so let  $S = \{n \in \mathbb{N} \mid \text{if } k \in n \Rightarrow k < n\}$  then
  - a. If  $n = 0$  then  $n = \emptyset \Rightarrow k \in n$  is impossible and thus ' $k \in 0 \Rightarrow k < n$ ' is always true giving that  $0 \in S$
  - b. If  $n \in S$  then if  $k \in s(n) = n \cup \{n\}$  we have
    - i. **( $k \in n$ )** then as  $n \in S$  we have  $k < n$  so  $k \neq n$ . From  $k \in n \subseteq n \cup \{n\} = s(n)$  we have  $k \leq s(n)$ . Now if  $k = s(n) = n \cup \{n\}$  then as  $n \in \{n\} \Rightarrow n \in n \cup \{n\} \Rightarrow n \in k \Rightarrow n \leq k \stackrel{k < n \Rightarrow k \leq n \text{ and 4.46}}{\Rightarrow} n = k$  contradicting  $k \neq n$  so that we have  $k \neq s(n)$  and thus  $k < s(n)$
    - ii. **( $k = n$ )** then we have  $k = n < s(n) \Rightarrow k < s(n)$   
from (i) and (ii) it follows then that  $s(n) \in S$

Using mathematical induction we have then  $S = \mathbb{N}$  proving that if  $n \in \mathbb{N} = S$  then if  $k \in n \Rightarrow k < n$

2. **( $k < n \Rightarrow k \in n$ )** If  $k < n$  then  $k \leq n$  and  $k \neq n$  so  $k \in n \vee k = n$  and  $k \neq n$  giving  $k \in n$   $\square$

**Theorem 4.50.** *If  $k, n, m \in \mathbb{N}$  then we have*

1.  $k \leq n$  and  $n < m$  then  $k < m$
2.  $k < n$  and  $n \leq m$  then  $k < m$

**Proof.** We have

1. If  $k \leq n$  and  $n < m$  then we have as  $n < m \Rightarrow n \leq m$  that  $k \leq m$ . Now for  $k$  we have the following possibilities:
  - a. **( $k < n$ )** then using the previous theorem we have  $k \in n$  and  $n \in m$ . If now  $k = m$  then we have  $n \in k$ . From  $n \in k$  and  $k \in n$  it follows then that  $k \leq n$  and  $n \leq k$  giving  $k = n$  contradicting  $k < n$ . So we must have  $k \neq m$  and thus  $k < m$ .
  - b. **( $k = n$ )** if now  $k = m$  then  $m = n$  contradicting  $n < m$  so we must have  $k \neq m$  and thus we have  $k < m$ .
2. If  $k < n$  and  $n \leq m$  then we have  $k \leq m$ . Now for  $n$  we have the following possibilities
  - a. **( $n < m$ )** Using the previous theorem we have  $n \in m$ . If now  $k = m$  then  $n \in k \Rightarrow n \leq k$  this with  $k < n \Rightarrow k \leq n$  means  $k = n$  contradicting  $k < n$ . So we have  $k \neq m$  and thus  $k < m$ .
  - b. **( $n = m$ )** then from  $k < n \stackrel{\text{previous theorem}}{\Rightarrow} k \in n$  we have  $k \in m \stackrel{\text{previous theorem}}{\Rightarrow} k < m$   $\square$

**Theorem 4.51.** For  $n, m \in \mathbb{N}$  we have  $n < m \Rightarrow s(n) \leq m$

**Proof.** We prove this by mathematical induction, so let  $S_n = \{m \in \mathbb{N} \mid n < m \Rightarrow s(n) \leq m\}$  we have then

1. As  $n < 0$  means that  $n \in 0 = \emptyset$  which is false so  $n < 0 \Rightarrow s(n) \leq 0$  is true or  $0 \in S_n$
2. If  $m \in S_n$  then  $n < m \Rightarrow s(n) \leq m$ , for  $s(n) < m$  we have then  $s(n) \in m$  we have now to prove that  $s(m) \in S_n$  or

$$n < s(m) \Rightarrow s(n) \leq s(m)$$

Now as  $n < s(m) \Rightarrow n \in s(m)$  by 4.11 then we have either

$$\text{a. } n \in m \Rightarrow n < m \Rightarrow s(n) \leq m \xrightarrow{4.46 \text{ and } m \in s(m) \Rightarrow m \cup \{m\} \Rightarrow m \leq s(m)} s(n) \leq s(m)$$

$$\text{b. } n = m \Rightarrow s(n) = s(m) \Rightarrow s(n) \leq s(m)$$

we conclude thus that  $s(m) \in S_n$

Using 4.10 we have then that  $S_n = \mathbb{N}$ . So if  $n, m \in \mathbb{N} \Rightarrow m \in S_n \Rightarrow n < m \Rightarrow s(n) \leq m$

□

**Theorem 4.52.**  $\langle \mathbb{N}, \leq \rangle$  is well-ordered

**Proof.** We prove this by contradiction, so let  $\emptyset \neq A \subseteq \mathbb{N}$  be nonempty without a least element. Define then  $S_A = \{n \in \mathbb{N} \mid \forall m \in A \text{ we have } n \leq m\}$  then as  $A$  has no least element we must have  $S_A \cap A = \emptyset$  [Otherwise  $x \in S_A \cap A$  would mean that  $x$  is a least element of  $A$ ].

We prove now by induction that  $S_A = \mathbb{N}$

1. By 4.47 we have  $\forall m \in A \subseteq \mathbb{N}$  that  $0 \leq m \Rightarrow 0 \in S_A$
2. If  $n \in S_A$  then  $\forall m \in A$  we have  $n \leq m$ . Now if  $n \in A$  then  $n$  is a least element of  $A$  (which we assumed to be not the case) so we must have  $\forall m \in A \vdash n < m$ , using 4.51 we have then that  $\forall m \in A s(n) \leq m$  and thus that  $s(n) \in S_A$ .

By 4.10 we have thus that  $S_A = \mathbb{N}$  but this means that  $S_A \cap A = \mathbb{N} \cap A = A \neq \emptyset$  contradicting that  $S_A \cap A = \emptyset$ . So every nonempty subset of  $\mathbb{N}$  must have a least element.

□

**Theorem 4.53.**  $\langle \mathbb{N}, \leq \rangle$  is fully-ordered

**Proof.** This follows from the well-orderings of  $\langle N, \leq \rangle$  and 2.192.

□

**Theorem 4.54.**  $\langle \mathbb{N}, \leq \rangle$  is conditional complete.

**Proof.** This follows from the fact that  $\langle \mathbb{N}, \leq \rangle$  is well-ordered and 2.193

□

**Theorem 4.55.**  $\forall n, m \in \mathbb{N} \text{ then } n < m \Leftrightarrow s(n) < s(m)$

**Proof.**

1. ( $n < m$ ) From 4.51 we have then  $s(n) \leq m$  and as  $m < s(m)$  we have thus  $s(n) < s(m)$
2. ( $s(n) < s(m)$ ) Assume that  $m \leq n$  then we have either  $m = n$  which leads to the contradiction  $s(m) = s(n)$ , or  $m < n \xrightarrow{(1)} s(m) < s(n) \Rightarrow s(m) < s(m)$  again a contradiction. So we must conclude that  $n < m$ .

□

**Theorem 4.56.** If  $n, m, k \in \mathbb{N}$  then  $n < m \Leftrightarrow n+k < m+k$ . Notice that from  $n = m \Leftrightarrow n+k = m+k$  we have also the equivalence  $n \leq m \Leftrightarrow n+k \leq m+k$

**Proof.** We prove this by induction on  $k$  so let  $B = \{k \in \mathbb{N} \mid \text{if } n, m \in \mathbb{N} \text{ then } n \leq m \Leftrightarrow n+k \leq m+k\}$  then we have

1. If  $k = 0$  then clearly we have for  $n, m \in \mathbb{N}$  that  $n \leq m \Leftrightarrow n+0 \leq m+0 \Leftrightarrow n+k \leq m+k \Rightarrow 0 \in B$

2. If  $k \in B$  then we have  $n \leq m \Leftrightarrow n+k \Leftrightarrow m+k \stackrel{\text{previous theorem}}{\Leftrightarrow} s(n+k) = s(m+k)$   $\Leftrightarrow_{s(n+k)=n+s(k), s(m+k)=m+s(k)} n+s(k) = m+s(k)$  so we have  $s(k) \in B$

using mathematical induction we have  $B = \mathbb{N}$  proving our theorem.  $\square$

**Theorem 4.57.** *If  $n, m \in \mathbb{N}$  with  $n < s(m)$  then  $n \leq m$*

**Proof.** If we would assume the opposite then  $m < n$  and using (4.51) we have then  $s(m) \leq n < s(m)$  a contradiction. So we must have  $n \leq m$   $\square$

**Theorem 4.58.** *If  $n, m \in \mathbb{N}$  and  $n < m$  then there exists a  $k \in \mathbb{N}, k \neq 0$  such that  $n+k = m$*

**Proof.** Let  $S_n = \{m \in \mathbb{N} \mid n < m \Rightarrow \exists k \in \mathbb{N} \mid k \neq 0 \vdash n+k = m\}$  then we have

1.  $n < 0$  is false as from 4.47  $0 \leq n \Rightarrow 0 < 0 \Rightarrow 0 \neq 0$  a contradiction so we have that  $n < 0 \Rightarrow \exists k \in \mathbb{N}, k \neq 0 \vdash n+k = m$  is true, so  $0 \in S_n$
2. If  $m \in S_n$  suppose now that  $n < s(m)$  then as  $\mathbb{N}$  is fully ordered (see 4.53) we have the following cases

a.  $m \leq n$  here we have the following cases

- i.  $m < n \stackrel{4.51}{\Rightarrow} s(m) \leq n \Rightarrow s(m) < s(m)$  a contradiction so this case does not apply.
- ii.  $m = n \Rightarrow s(n) = s(m)$  contradiction  $n < s(m)$  so this case does not apply.

b.  $n \leq m$  here we have the following possibilities

- i.  $n = m$  then  $n+1 \stackrel{4.30}{=} s(n) = s(m) \stackrel{1 \neq 0}{\neq} s(m) \in S_n$
- ii.  $n < m \stackrel{m \in S_n}{\Rightarrow} \exists k \in \mathbb{N}, k \neq 0 \vdash n+k = m \Rightarrow n+s(k) \stackrel{4.31}{=} s(n+k) = s(m)$ . As we have  $k \in s(k) = k \cup \{k\} \Rightarrow k \leq s(k)$  and  $0 < k \Rightarrow 0 < s(k) \Rightarrow 0 \neq s(k)$  we conclude that  $s(m) \in S_n$

Using mathematical induction 4.10 we have then that  $S_n = \mathbb{N}$ . So if  $n, m \in \mathbb{N} \Rightarrow m \in S_n \Rightarrow n < m \Rightarrow \exists k \in \mathbb{N}, k \neq 0 \vdash n+k = m$ .  $\square$

**Lemma 4.59.** *If  $n \in \mathbb{N}, k \in \mathbb{N}_0 \Rightarrow n < n+k$*

**Proof.** We prove this by mathematical induction, so let  $k \in \mathbb{N}_0$  and  $S_k = \{n \in \mathbb{N} \mid n < n+k\}$  then we have

1.  $0+k \stackrel{4.29}{=} k$  and as  $k \neq 0$  we have  $0 < k = 0+k \Rightarrow 0 \in S_k$
2. If  $n \in S_k$  then by 4.48 we have  $n+k < s(n+k) \stackrel{4.31}{=} s(n)+k$  also as  $n \in S_k$  we have  $n < n+k \Rightarrow s(n) \leq n+k \Rightarrow s(n) < s(n)+k \Rightarrow s(n) \in S_k$

By 4.10 we have  $S_k = \mathbb{N}$ . So if  $k \in \mathbb{N}_0$  and  $n \in \mathbb{N} \Rightarrow n \in S_k$  and thus  $n < n+k$   $\square$

**Theorem 4.60.** *If  $n, k \in \mathbb{N}$  then  $n+k = 0 \Rightarrow n = k = 0$*

**Proof.** Suppose that  $k \neq 0$  then as  $0 \leq n \stackrel{4.59}{\Rightarrow} 0 \leq n < n+k = 0 \Rightarrow 0 < 0 \Rightarrow 0 \neq 0$  and thus we reach a contradiction. So  $k = 0$  but then  $n = n+0 = n+k = 0$  and thus also  $n = 0$   $\square$

**Theorem 4.61.** *If  $n, m \in \mathbb{N}$  then  $n < m \Leftrightarrow \exists! k \in \mathbb{N}_0 \vdash n+k = m$ .*

**Proof.**

$\Rightarrow$

We prove this by mathematical induction, so let  $S_n = \{m \in \mathbb{N} \mid \text{if } n < m \Rightarrow \exists k \in \mathbb{N}_0 \vdash n+k = m\}$  then we have

1. We have by 4.47 that  $0 \leq n$  so if  $n < 0 \Rightarrow 0 < 0$  a contradiction. So we can never have  $n < 0 \Rightarrow 0 \in S_n$

2. If  $m \in S_n$  suppose then that  $n < s(m)$  then we have the following possible cases for  $m$

a.  $n = m \Rightarrow s(m) = s(n) \stackrel{4.30}{=} n+1 \stackrel{1 \neq 0}{\Rightarrow} s(m) \in S_n$

b.  $n < m \stackrel{m \in S_n}{\Rightarrow} \exists k \in \mathbb{N}_0 \vdash n+k = m \stackrel{4.31}{\Rightarrow} n+s(k) = s(n+k) = s(m)$  and as from  $0 \leq k \stackrel{k \neq 0}{\Rightarrow} 0 < k$  and by 4.48  $k < s(k) \Rightarrow 0 < s(k) \Rightarrow s(k) \neq 0$  and as we just have proved that  $n+s(k) = s(m)$  we have  $s(m) \in S_n$

c.  $m < n \stackrel{4.51}{\Rightarrow} s(m) \leq n \stackrel{n < s(m)}{\Rightarrow} s(m) < s(m)$  is a contradiction so this case does not apply

In all the cases that applies we had  $s(m) \in S_n$

Using 4.10 we have then  $S_n = \mathbb{N}$  so if  $m, n \in \mathbb{N} \Rightarrow m \in S_n \Rightarrow$  if  $n < m \Rightarrow \exists k \in \mathbb{N}_0 \vdash n+k = m$ . Finally to prove uniqueness assume that  $n+k = m = n+k'$  then by 4.44 we have  $k = k'$ .

$\Leftarrow$

Suppose  $\exists k \in \mathbb{N}_0 \vdash n+k = m$  then for  $n, m$  we have by 4.53 that either  $n = m$ ,  $m < n$  or  $n < m$ . Now if  $n = m$  then as  $k \in \mathbb{N}_0$  we have by 4.59 that  $n < n+k = m = n$  a contradiction. Also if  $m < n$  then again from 4.59 we have  $n < n+k = m \Rightarrow n < n$  again a contradiction. So the only possibility left is  $n < m$ .

□

**Corollary 4.62.** *If  $n, m \in \mathbb{N}$  then  $n \leq m \Leftrightarrow \exists! k \in \mathbb{N} \vdash n+k = m$ .*

**Proof.**

$\Rightarrow$  if  $n \leq m$  then we have either

1.  $n < m \stackrel{\text{previous theorem}}{\Rightarrow} \exists! k \in \mathbb{N}_0 \vdash n+k = m \Rightarrow \exists k \in \mathbb{N} \vdash n+k = m$

2.  $n = m \Rightarrow m = n+0 \Rightarrow \exists k \in \mathbb{N} (k = 0) \vdash n+k = m$

Finally to prove uniqueness assume that  $n+k = m = n+k'$  then by 4.44 we have  $k = k'$ .

$\Leftarrow$  If  $\exists k \in \mathbb{N} \vdash n+k = m$  then we have for  $k$  the following possibilities

1.  $k = 0 \Rightarrow n = m \Rightarrow n \leq m$

2.  $k \neq 0 \Rightarrow k \in \mathbb{N}_0$  and using the previous theorem we have then  $n < m \Rightarrow n \leq m$

□

**Definition 4.63.** *If  $n, m \in \mathbb{N}$ ,  $n \leq m$  then the unique  $k$  that by the previous theorem exists so that  $n+k = m$  is noted by  $m-n$  so  $n+(m-n) = m$ .*

**Note 4.64.** *If  $n \in \mathbb{N} \Rightarrow n-n=0$  (this is trivial because of uniqueness and the fact that  $n+0=n$ )*

**Theorem 4.65.** *If  $n, m \in \mathbb{N}$  and  $i \in \mathbb{N}$  such that  $n \leq i \leq m$  then  $0 \leq i-n \leq m-n$*

**Proof.** First if  $m-n < i-n \Rightarrow m = (m-n)+n < (i-n)+n = i \Rightarrow m < i \leq m \Rightarrow m < m$  a contradiction so we have that  $i-n \leq m-n$ . By applying the above on  $n \leq i$  we find that  $0 = n-n \leq i-n \Rightarrow 0 \leq i-n$

□

**Theorem 4.66.** *If  $n \in \mathbb{N}$  then there does not exist a  $k \in \mathbb{N}$  such that  $n < k < s(n)$*

**Proof.** We prove this by mathematical induction, so take  $S = \{n \in \mathbb{N} \mid \text{there does not exist } k \in \mathbb{N} \text{ such that } n < k < s(n)\}$ . We have then

1. If  $0 < k < 1 \stackrel{4.51}{\Rightarrow} s(0) = 1 \leq s(k) \leq 1 \Rightarrow s(k) = 1 = s(0) \stackrel{4.15}{\Rightarrow} k = 0$  contradicting  $0 < k$ . So we conclude that  $0 \in S$

2. Let  $n \in S$ . Assume now that  $\exists k \in \mathbb{N}$  such that  $s(n) < k < s(s(n))$  then from 4.61 we have  $\exists l \in \mathbb{N}_0$  such that  $k = s(n)+l = s(n+l)$ . Take then  $k' = n+l$  then as  $l \in \mathbb{N}_0$  we have by 4.61 again that  $n < k'$ . From  $s(k') = k < s(s(n)) \Rightarrow s(k') < s(s(n)) \stackrel{4.55}{\Rightarrow} k' < s(n)$  so we have  $n < k' < s(n)$  contradicting  $n \in S$ . So the assumption turns out to be wrong and thus  $s(n) \in \mathbb{N}$

So by mathematical induction we have  $S = \mathbb{N}$  proving our theorem.

□

**Theorem 4.67.** If  $\emptyset \neq A \subseteq \mathbb{N}$  is a set such that  $\sup(A)$  exists then  $\sup(A) \in A$

**Proof.** Suppose that  $A \neq \emptyset$  and that  $\sup(A)$  exists. We have now the following cases for  $\sup(A)$

1. ( $\sup(A) = 0$ ) As  $A \neq \emptyset$  there exists a  $x \in A$  and as a sup is an upper bound we have  $x \leq 0$  and from 4.47 we have  $0 \leq x \Rightarrow x = 0 \Rightarrow \sup(A) \in A$
2. ( $\sup(A) \neq 0$ ) Using 4.17 we have  $\exists n \in \mathbb{N}$  such that  $\sup(A) = s(n)$ . Then as  $n < s(n) = \sup(A)$  there must exist a  $x \in A$  with  $n < x \leq \sup(A)$  (otherwise  $n$  would be the lowest upper bound). If  $x \neq \sup(A)$  then we have  $n < x < s(n)$  which is forbidden by 4.66 so we must have  $x = \sup(A) \Rightarrow \sup(A) \in A$ .  $\square$

**Theorem 4.68.** If  $n, m, r, s \in \mathbb{N}$  then if  $n < m \wedge r < s \Rightarrow n+r < m+s$

**Proof.** If  $n < m$  and  $r < s$  then there exists  $k, l \in \mathbb{N}_0$  such that  $m = n+k$  and  $s = r+l$  so  $m+s = (n+k)+(r+l) = (n+r)+(k+l)$ . Now as  $0 < k$  we have by 4.59 that  $0 < k+l \Rightarrow k+l \in \mathbb{N}_0$ . Using 4.61 we have then that  $n+r < m+s$ .  $\square$

**Lemma 4.69.** If  $n, m \in \mathbb{N}_0$  then  $n \cdot m \in \mathbb{N}_0$

**Proof.** We prove this by mathematical induction. If  $n \in \mathbb{N}_0$  then let  $S_n = \{m \in \mathbb{N} \mid m \neq 0 \Rightarrow n \cdot m \neq 0\}$  then we have

1.  $0 \in S$  [ $0 \neq 0$  is false so  $[0 \neq 0 \Rightarrow n \cdot m \neq 0]$  is true]
2. If  $m \in S_n$  then  $n \cdot s(m) = n \cdot m + n$  as  $m \in S_n$  we have  $n \cdot m \in S_n \Rightarrow 0 < n \cdot m \xrightarrow{4.59 \text{ and } n \in S_n} 0 < n \cdot m < n \cdot m + n \Rightarrow 0 < n \cdot m + n = n \cdot s(m) \Rightarrow n \cdot s(m) \neq 0$

Using 4.10 we have then that  $S_n = \mathbb{N}$  So if  $n, m \in \mathbb{N}_0 \Rightarrow 0 \neq m \in S_n \Rightarrow n \cdot m \neq 0$   $\square$

**Theorem 4.70.** If  $n, m, k \in \mathbb{N}$  then if  $k \neq 0$  and  $n < m \Rightarrow k \cdot n < k \cdot m$

**Proof.** As  $n < m \xrightarrow{4.61} \exists l \in \mathbb{N}_0$  such that  $n+l = m \Rightarrow k \cdot (n+l) = k \cdot m \Leftrightarrow k \cdot n + k \cdot l = k \cdot m$ . As we have  $0 \leq n$  and  $n < m \Rightarrow 0 < m \Rightarrow m \in \mathbb{N}_0 \xrightarrow{4.69} k \cdot l \in \mathbb{N}_0 \xrightarrow{4.61} k \cdot n < k \cdot m$   $\square$

**Corollary 4.71.** If  $n, m, k \in \mathbb{N}$  then if  $n \leq m \Rightarrow n \cdot k \leq m \cdot k$

**Proof.** If  $n \leq m$  then we have the following cases

1.  $n = m \Rightarrow n \cdot k = m \cdot k \Rightarrow n \cdot k \leq m \cdot k$
2.  $n < m$  then we have for  $k$  the following cases
  - a.  $k = 0 \xrightarrow{4.37} n \cdot k = 0 = m \cdot k \Rightarrow n \cdot k \leq m \cdot k$
  - b.  $k \neq 0 \xrightarrow{\text{previous theorem}} n \cdot k < m \cdot k \Rightarrow n \cdot k \leq m \cdot k$

$\square$

**Theorem 4.72.** If  $n, m \in \mathbb{N}$ ,  $k \in \mathbb{N}_0$  and  $n \cdot k = m \cdot k \Rightarrow n = m$

**Proof.** We have by 4.53 that  $n < m, m < n$  or  $n = m$ . Now if  $n < m \xrightarrow{4.70} n \cdot k < m \cdot k$  a contradiction, also if  $m < n \Rightarrow m \cdot k < n \cdot k$  a contradiction. So the only thing left is that  $n = m$ .  $\square$

**Theorem 4.73. (Archimedean property of  $\mathbb{N}$ )** If  $x, y \in \mathbb{N}$  and  $x > 0$  then there exists a  $z \in \mathbb{N}$  such that  $z \cdot x > y$

**Proof.** We have the following possibilities for  $y$

1.  $y \leq x$  then as  $1 < 2 = s(1)$  (see 4.48) we have by 4.70 that  $y \leq x = x \cdot 1 < 2 \cdot x$  so taking  $z = 2$  we have  $z \cdot z > y$
2.  $x < y$  then by 4.61 there exists a  $k \in \mathbb{N}_0$  such that  $y = x+k$  now as  $0 < x$  we have by 4.51 that  $1 = s(0) \leq x \Rightarrow 1 \leq x$ , also we have that  $x+k < s(x+k)$ . This gives using 4.71  $s(x+k) \cdot 1 \leq s(x+k) \cdot x \Rightarrow y = x+k < s(x+k) \leq s(x+k) \cdot x \Rightarrow y < s(x+k) \cdot y$ , so taking  $z = s(x+k)$  we have  $z \cdot x > y$   $\square$

**Theorem 4.74. (Division Algorithm)** *If  $m, n \in \mathbb{N}$  and  $n > 0$  then there exists a unique  $r \in \mathbb{N}$  with  $0 \leq r < n$  and a unique  $q \in \mathbb{N}$  such that  $m = n \cdot q + r$*

**Proof.** First we prove the existence of  $r, q \in \mathbb{N}$ . For  $m$  we have the following cases to consider

1. ( $m = 0$ ) In this case we can take  $r = 0, q = 0$  and then  $0 = r < n$  and  $n \cdot q + r = n \cdot 0 + 0 = 0 = m$
2. ( $0 < m$ ) We have then the following cases for  $n$ 
  - a. ( $n = 1$ ) Then we can take  $q = m, r = 0$  and then  $0 = r < n$  and  $n \cdot q + r = 1 \cdot m + 0 = m$
  - b. ( $1 < n$ ) Then from 4.70 we have  $m < n \cdot m$  so  $m \in A_{n,m} = \{x \in \mathbb{N} | m < n \cdot x \wedge x \leq m\} \Rightarrow A_{n,m} \neq \emptyset$ . Using well-ordering (see 4.52) there exists a  $q' = \min(A_{n,m})$ . We have  $q' \neq 0$  [because if  $q' = 0$  then from  $q' \in A_{n,m}$  we have  $0 < m < n \cdot 0 = 0 \Rightarrow 0 < 0$  a contradiction] and thus there exists a  $q \in \mathbb{N}$  such that  $q' = s(q) = q + 1$ . From  $q < q' \leq m$  we have by the definition of the minimum that  $q \notin A_{n,m}$  and thus  $n \cdot q \leq m$  so we can define  $r = m - n \cdot q$  where  $r \in \mathbb{N}$  and thus  $m = n \cdot q + r$ . If now  $n \leq r \Rightarrow n + n \cdot q \leq r + n \cdot q = m \Rightarrow n \cdot (q + 1) \leq m \Rightarrow n \cdot q' \leq m$  and from  $q' \in A_{n,m}$  we have  $m < n \cdot q'$  so we reach the contradiction  $n \cdot q' \leq m < n \cdot q' \Rightarrow n \cdot q < n \cdot q'$  and we conclude that  $0 \leq r < n$ .

Now to prove uniqueness assume that  $m = n \cdot q + r = n \cdot q'' + r''$  then we have if  $q \neq q''$  that either the following is true

1. ( $q < q''$ ) then  $q + 1 \leq q'' \Rightarrow n \cdot (q + 1) \leq n \cdot q'' \Rightarrow n \cdot q + n \leq n \cdot q'' \Rightarrow n \cdot q + n + r + r'' \leq n \cdot q'' + r + r'' \Rightarrow m + n + r'' \leq m + r \Rightarrow n + r'' \leq r \Rightarrow n \leq r$  contradicting  $r < n$
2. ( $q'' < q$ ) then  $q'' + 1 < q \Rightarrow n \cdot (q'' + 1) \leq n \cdot q \Rightarrow n \cdot q'' + n \leq n \cdot q \Rightarrow n \cdot q'' + n + r + r'' \leq n \cdot q + r + r'' \Rightarrow m + n + r \leq m + r'' \Rightarrow n + r \leq r'' \Rightarrow n \leq r''$  contradicting  $r'' < n$

as neither of the above cases can be true we must have  $q = q''$  but then  $n \cdot q = n \cdot q'' \xrightarrow{n \cdot q + r = n \cdot q'' + r''} r = r''$  proving uniqueness.  $\square$

#### 4.4.1 Other forms of Mathematical induction

Once we have discussed the order relation on  $\mathbb{N}$  we can write down some other forms of mathematical induction.

**Definition 4.75.** Given  $n \in \mathbb{N}$  we define  $\{n, \dots\}$  to be equal to  $\{i \in \mathbb{N} | n \leq i\}$

**Definition 4.76.** Given  $n, m \in \mathbb{N}$  with  $n \leq m$  we define  $\{n, \dots, m\} = \{i \in \mathbb{N} | n \leq i \wedge i \leq m\}$ , note that if  $m < n$  then  $\{n, \dots, m\} = \emptyset$

We use this in the following theorem

**Theorem 4.77.** *If  $n \in \mathbb{N}$  and  $X \subseteq \{n, \dots\}$  is such that*

1.  $n \in X$
2. if  $i \in X \Rightarrow i + 1 \in X$

*then  $X = \{n, \dots\}$*

**Proof.** Take  $B = \{i \in \mathbb{N} | i + n \in X\}$  then we have

1. As  $n \in X$  we have  $0 + n \in X \Rightarrow 0 \in B$
2. If  $i \in B$  then  $i + n \in X \Rightarrow (i + 1) + n = (i + n) + 1 \in X \Rightarrow i + 1 \in B$

by mathematical induction (see 4.10) we have then  $B = \mathbb{N}$ . If now  $i \in \{n, \dots\}$  then  $n \leq i \Rightarrow i - n \in \mathbb{N} = B \Rightarrow (i - n) + n \in X \Rightarrow i \in X \Rightarrow \{n, \dots\} \subseteq X \subseteq \{n, \dots\} \Rightarrow X = \{n, \dots\}$   $\square$

**Theorem 4.78.** *If  $n \in \mathbb{N}$  and  $P(i)$  is a predicate defined  $\forall i \in \{n, \dots\}$  depending on  $i$  is such that*

1.  $P(n)$  is true

2. If  $n \leq i$  and  $P(i)$  is true then  $P(i+1)$  is true  
then  $\forall i \in \{n, \dots\}$  we have that  $P(i)$  is true

**Proof.** Define  $X = \{i \in \{n, \dots\} \mid P(i) \text{ is true}\}$  then we have

1.  $n \in X$
2. if  $i \in X \Rightarrow i+1 \in X$

thus  $X = \{n, \dots\}$  and if  $i \in \{n, \dots\} \Rightarrow i \in X \Rightarrow P(i)$  is true  $\square$

The last form is the form of induction used in most mathematical text. In this text however we use mostly the form 4.77 as I think this forces you to express the induction hypotheses in a much clearer form. As a example of induction and recursion lets define the composition of  $n$  functions.

**Definition 4.79.** Let  $n \in \mathbb{N}_0$ ,  $\{X_i\}_{i \in \{1, \dots, n+1\}}$  a family of sets,  $\{f_i\}_{i \in \{1, \dots, n\}}$  with  $\forall i \in \{1, \dots, n\}$  we have  $f_i \subseteq X_i \times X_{i+1}$  then  $f_n \circ \dots \circ f_1 \subseteq X_1 \times X_{n+1}$  is defined by

1.  $f_1 \circ \dots \circ f_1 = f_1$
2.  $f_{n+1} \circ \dots \circ f_1 = f_{n+1} \circ (f_n \circ \dots \circ f_1)$

**Theorem 4.80.** Let  $n \in \mathbb{N}_0$ ,  $\{X_i\}_{i \in \{1, \dots, n+1\}}$  a family of sets,  $\{f_i\}_{i \in \{1, \dots, n\}}$  a family of functions  $f_i: X_i \rightarrow X_{i+1}$  then  $f_n \circ \dots \circ f_1: X_1 \rightarrow X_{n+1}$  is a function

**Proof.** We use mathematical induction to prove this so let  $S = \{n \in \mathbb{N}_0 \mid \text{if } \{f_i\}_{i \in \{1, \dots, n\}} \text{ is a family of functions } f_i: X_i \rightarrow X_{i+1} \text{ then } f_n \circ \dots \circ f_1: X_1 \rightarrow X_{n+1} \text{ is a function}\}$ , then we have

**1  $\in S$ .** If  $\{f_i\}_{i \in \{1, \dots, 1\}}$  is a family of functions  $f_i: X_i \rightarrow X_{i+1}$  then  $f_1: X_1 \rightarrow X_2$  is a function so that  $f_1 \circ \dots \circ f_1 = f_1: X_1 \rightarrow X_2$  is a function proving that  $1 \in S$

**$n \in S \Rightarrow n+1 \in S$ .** If  $\{f_i\}_{i \in \{1, \dots, n+1\}}$  is a family of functions  $f_i: X_i \rightarrow X_{i+1}$  then as  $n \in S$  we have  $f_n \circ \dots \circ f_1: X_1 \rightarrow X_{n+1}$  is a function, then as  $f_{n+1}: X_{n+1} \rightarrow X_{n+2} = X_{(n+1)+1}$  is a function by the hypothesis we have that  $f_{n+1} \circ \dots \circ f_1 = f_{n+1} \circ (f_n \circ \dots \circ f_1): X_1 \rightarrow X_{(n+1)+1}$  is a function proving that  $n+1 \in S$

Using mathematical induction we have that  $S = \mathbb{N}_0$  proving the theorem.  $\square$

We use now the associativity of the composition of functions to prove a more general result

**Theorem 4.81.** Let  $n \in \mathbb{N}_0 \setminus \{1\}$ ,  $m \in \mathbb{N}$   $\{X_i\}_{i \in \{1, \dots, n+1\}}$  a family of sets,  $\{f_i\}_{i \in \{1, \dots, n+m\}}$  a family of functions  $f_i: X_i \rightarrow X_{i+1}$  then we have that  $(f_{n+m} \circ \dots \circ f_n) \circ (f_{n-1} \circ \dots \circ f_1) = (f_{n+m} \circ \dots \circ f_1)$

**Proof.** We prove this by induction on  $m$  so let  $S_n = \{m \in \mathbb{N} \mid (f_{n+m} \circ \dots \circ f_n) \circ (f_{n-1} \circ \dots \circ f_1) = (f_{n+m} \circ \dots \circ f_1)\}$  then we have

**$0 \in S_n$ .** In this case we have  $(f_{n+0} \circ \dots \circ f_n) \circ (f_{n-1} \circ \dots \circ f_1) = (f_n \circ \dots \circ f_n) \circ (f_{n-1} \circ \dots \circ f_1) = f_n \circ (f_{n-1} \circ \dots \circ f_1) = (f_n \circ \dots \circ f_1)$  proving that  $0 \in S$

**$m \in S_n \Rightarrow m+1 \in S_n$ .** Then for  $(f_{n+(m+1)} \dots f_n) \circ (f_{n-1} \circ \dots \circ f_1) = ((f_{n+m} \circ (f_{n+m} \circ \dots \circ f_n)) \circ (f_{n-1} \circ \dots \circ f_1)) \stackrel{2.37}{=} f_{n+(m+1)} \circ ((f_{n+m} \circ \dots \circ f_n) \circ (f_{n-1} \circ \dots \circ f_1)) \stackrel{m \in S_n}{=} f_{n+(m+1)} \circ (f_{n+m} \circ \dots \circ f_1) = (f_{n+(m+1)} \circ \dots \circ f_1)$  proving that  $m+1 \in S$

Mathematical induction proves then the theorem.  $\square$

# Chapter 5

## Finite and Infinite Sets

### 5.1 Introduction

**Definition 5.1.** Two classes  $A$  and  $B$  are called equipotent if there exists a bijection  $f: A \rightarrow B$ , we note this as  $A \approx B$

**Definition 5.2.** Given two classes  $A$  and  $B$  then  $A \preccurlyeq B$  if there is a  $C \subseteq B$  such that  $A \approx C$

**Definition 5.3.** Given two classes  $A$  and  $B$  then  $A \prec B$  if  $A \preccurlyeq B$  and  $\neg(A \approx B)$

**Theorem 5.4.** Given two classes  $A$  and  $B$  then  $A \preccurlyeq B$  if and only if there exists a injection  $f: A \rightarrow C$  where  $C \subseteq B$

**Proof.**

$$\begin{aligned}
 A \preccurlyeq B &\Leftrightarrow \text{there exists a } C \subseteq B \text{ and a bijection } f: A \rightarrow C \\
 &\Rightarrow f: A \rightarrow B \text{ is a injection} \\
 f: A \rightarrow B \text{ is a injection} &\Rightarrow f: A \rightarrow f(A) \text{ is a bijection} \\
 &\stackrel{f(A) \subseteq B}{\Rightarrow} A \preccurlyeq B
 \end{aligned}$$

□

**Theorem 5.5.** If  $A$  is a set then there exists no surjective function between  $A$  and  $\mathcal{P}(A)$

**Proof.** We prove this by contradiction so assume that there is a surjective function  $f: A \rightarrow \mathcal{P}(A)$  define then  $B = \{x \in A \mid x \notin f(x)\}$ . As  $B \subseteq A$  we have that  $B \in \mathcal{P}(A)$  and by surjectivity of  $f$  there exists a  $y \in A$  such that  $B = f(y)$ . Now if  $y \in B$  then  $y \notin f(y) = B \Rightarrow y \notin B$  a contradiction. If  $y \notin B$  then as  $B = f(y)$  we have  $y \notin f(y)$  and thus  $y \in B$  again a contradiction. So all the cases  $y \in B$  or  $y \notin B$  gives a contradiction and thus a surjective  $f: A \rightarrow \mathcal{P}(A)$  can not exist. □

**Corollary 5.6.** If  $A$  is a set then no subset of  $A$  can be equipotent with  $\mathcal{P}(A)$  (or with  $2^A$ )

**Proof.** First we prove that no subset of  $A$  can be equipotent with  $\mathcal{P}(A)$ . If  $B \subseteq A$  then we have the following possible cases:

1. ( $B = A$ ) then by 5.5 we can not have a bijection (which is surjective) between  $A$  and  $\mathcal{P}(A)$ , so  $B$  can not be equipotent with  $\mathcal{P}(A)$
2. ( $B \subset A$ ) in this case  $A \setminus B$  is non-empty and  $B \cap (A \setminus B) = \emptyset$ . Assume now that  $B$  is equipotent with  $\mathcal{P}(A)$  then a bijection  $f': B \rightarrow \mathcal{P}(A)$  exists. Form then the function (see 2.34)  $f: A \rightarrow \mathcal{P}(A)$  defined by  $x \rightarrow f(x) = \begin{cases} \emptyset & \text{if } x \in A \setminus B \\ f'(x) & \text{if } x \in B \end{cases}$ . If now  $y \in \mathcal{P}(A)$  there exists by bijectivity of  $f'$  a  $x \in B$  such that  $f'(x) = y \Rightarrow f(x) = y \Rightarrow f$  is surjective which is impossible by 5.5. So we must have that  $B$  is not equipotent with  $\mathcal{P}(A)$

Next if  $B \subseteq A$  and  $B$  is equipotent with  $2^A$  then there exists a bijection  $f: B \rightarrow 2^A$ . By 2.77  $2^A$  and  $\mathcal{P}(A)$  are bijective so there exists a bijection  $g: 2^A \rightarrow \mathcal{P}(A)$ , but this means that  $g \circ f: B \rightarrow \mathcal{P}(A)$  is bijective which we have just proved is impossible. So  $B$  and  $\mathcal{P}(A)$  cannot be equipotent. □

**Corollary 5.7.** If  $A$  is a set then  $A$  cannot be equipotent with any  $B \supseteq \mathcal{P}(A)$  or with any  $B \supseteq 2^A$

**Proof.** Suppose that  $A$  is equipotent with  $B \supseteq \mathcal{P}(A)$  then there exists a bijection  $f: A \rightarrow B$ , using 2.65 we have that  $B|f: f^{-1}(\mathcal{P}(A)) \rightarrow \mathcal{P}(A)$  is a bijection which by the previous corollary is impossible.

Suppose that  $A$  is equipotent with  $B \supseteq 2^A$  then there exists a bijection  $f: A \rightarrow B$ , using 2.65 we have that  $B|f: f^{-1}(2^A) \rightarrow 2^A$  is a bijection which by the previous corollary is impossible.  $\square$

**Theorem 5.8.** *If  $A, B$  are classes then there exists a injection  $f: A \rightarrow B$  if and only if there exists a surjection  $g: B \rightarrow A$*

**Proof.**

If there exists a injection  $f: A \rightarrow B$  then by 2.47 there exists a function  $g: B \rightarrow A$  such that  $g \circ f = i_A$ . Then if  $y \in A$  we have  $y = i_A(y) = g(f(y)) \Rightarrow g$  is surjective.

If there exists a surjection  $g: B \rightarrow A$  then by 2.48 there exists a  $f: A \rightarrow B$  such that  $g \circ f = i_A$ . Given  $x, x' \in A$  if  $f(x) = f(x') \Rightarrow g(f(x)) = g(f(x')) \Rightarrow (g \circ f)(x) = (g \circ f)(x') = i_A(x) = i_A(x') \Rightarrow x = x'$  and this we have that  $f$  is injective.  $\square$

Using the above theorem and 5.4 we have the following corollary

**Corollary 5.9.** *If  $A, B$  are classes then  $A \preceq B$  if and only if there exists a surjection  $f: B \rightarrow A$*

**Theorem 5.10.** *If  $A, B, C, D$  are classes where  $A \cap C = \emptyset$  and  $B \cap D$  are sets and  $f: A \rightarrow B$  and  $g: C \rightarrow D$  are bijections then  $f \cup g: A \cup C \rightarrow B \cup D$  is a bijection*

**Proof.** As  $f: A \rightarrow B$  and  $g: C \rightarrow D$  are functions then clearly  $f: A \rightarrow B \cup D$  and  $g: C \rightarrow B \cup D$  are also functions and thus by 2.34  $f \cup g: A \cup C \rightarrow B \cup D$  is a function. Now as  $f, g$  are bijective we have the existence of the bijections  $f^{-1}: B \rightarrow A$ ,  $g^{-1}: D \rightarrow C$  meaning that  $f^{-1}: B \rightarrow A \cup C$  and  $g^{-1}: D \rightarrow A \cup C$  are functions and by 2.34 we have  $f^{-1} \cup g^{-1}: B \cup D \rightarrow A \cup C$  is also a function. Now  $\forall x \in B \cup D$  we have either

1. ( $x \in B$ ) then  $(f^{-1} \cup g^{-1})(x) = f^{-1}(x) \in A \Rightarrow (f \cup g)((f^{-1} \cup g^{-1})(x)) = (f \cup g)(f^{-1}(x)) = f(f^{-1}(x)) = i_B(x) = x = i_{B \cup D}(x)$
2. ( $x \in D$ ) then  $(f^{-1} \cup g^{-1})(x) = g^{-1}(x) \in C \Rightarrow (f \cup g)((f^{-1} \cup g^{-1})(x)) = (f \cup g)(g^{-1}(x)) = g(g^{-1}(x)) = i_D(x) = x = i_{B \cup D}(x)$

So we have  $(f \cup g) \circ (f^{-1} \cup g^{-1}) = i_{B \cup D}$ . Also  $\forall y \in A \cup C$  we have either

1. ( $y \in A$ ) then  $(f \cup g)(y) = f(y) \in B$  so that  $(f^{-1} \cup g^{-1})((f \cup g)(y)) = (f^{-1} \cup g^{-1})(f(y)) = f^{-1}(f(y)) = i_A(y) = y = i_{A \cup C}(y)$
2. ( $y \in C$ ) then  $(f \cup g)(y) = g(y) \in D$  so that  $(f^{-1} \cup g^{-1})((f \cup g)(y)) = (f^{-1} \cup g^{-1})(g(y)) = g^{-1}(g(y)) = i_D(y) = y = i_{A \cup C}(y)$

so we have then  $(f^{-1} \cup g^{-1}) \circ (f \cup g) = i_{A \cup C}$ . This proves that  $f \cup g: A \cup C \rightarrow B \cup D$  is a bijective function with inverse function  $f^{-1} \cup g^{-1}: B \cup D \rightarrow A \cup C$   $\square$

**Corollary 5.11.** *If  $A, B, C, D$  are classes with  $A \cap C = \emptyset$ ,  $B \cap D = \emptyset$  and  $A \approx B$  and  $C \approx D$  then  $(A \cup C) \approx (B \cup D)$*

**Proof.** This is ease as  $A \approx B$ ,  $C \approx D$  implies that there exists bijections  $f: A \rightarrow B$  and  $g: C \rightarrow D$ . Using the previous theorem we haven the bijection  $f \cup g: A \cup C \rightarrow B \cup D$  and thus that  $(A \cup C) \approx (B \cup D)$   $\square$

**Theorem 5.12.** *If  $A, B, C, D$  are classes and  $A \approx C$  and  $B \approx D$  then  $(A \times B) \approx (C \times D)$*

**Proof.** As  $A \approx C$  and  $B \approx D$  we have the existence of the bijections  $f: A \rightarrow C$  and  $g: B \rightarrow D$ . Define  $h: A \times B \rightarrow C \times D$  by  $(x, y) \rightarrow (f(x), g(y))$  then we have

1. If  $(x, y), (x', y') \in A \times B$  such that  $h((x, y)) = h((x', y'))$  then  $(f(x), g(y)) = (f(x'), g(y')) \Rightarrow f(x) = f(x') \wedge g(y) = g(y') \Rightarrow x = x' \wedge y = y' \Rightarrow (x, y) = (x', y')$  proving injectivity.

2. If  $(u, v) \in C \times D$  then  $u \in C \wedge v \in D \xrightarrow{f, g \text{ are surjective}} \exists x \in A, y \in B \vdash u = f(x) \wedge v = g(y) \Rightarrow (u, v) = (f(x), g(y)) = h((x, y))$

This proves that  $h: A \times B \rightarrow C \times D$  is bijective and thus  $(A \times B) \approx (C \times D)$   $\square$

**Theorem 5.13.** *If  $A, B, C, D$  are sets such that  $A \approx B$  and  $C \approx D$  then  $A^C \approx B^D$*

**Proof.** As  $A \approx B$  and  $C \approx D$  there exists bijections  $f: A \rightarrow B$  and  $g: D \rightarrow C$ . Define now  $h: A^C \rightarrow B^D$  by  $s \in A^C \rightarrow h(s) = f \circ s \circ g \in B^D$  [as  $s: C \rightarrow A$  we have that  $f \circ s \circ g: D \rightarrow B \Rightarrow f \circ s \circ g \in B^D$ ]. We have now

1. If  $s, t \in A^C$  is such that  $h(s) = h(t)$  then  $f \circ s \circ g = f \circ t \circ g \Rightarrow f^{-1} \circ (f \circ s \circ g) = f^{-1} \circ (f \circ t \circ g) \Rightarrow s \circ g = t \circ g \Rightarrow (s \circ g) \circ g^{-1} = (t \circ g) \circ g^{-1} = s = t$  proving injectivity.
2. If  $u \in B^D$  then  $u: D \rightarrow B$  and thus  $f^{-1} \circ u \circ g^{-1}: C \rightarrow A \Rightarrow f^{-1} \circ u \circ g^{-1} \in A^C$  and  $h(f^{-1} \circ u \circ g^{-1}) = f \circ (f^{-1} \circ u \circ g^{-1}) \circ g = u$  proving surjectivity.  $\square$

As we have that  $\mathcal{P}(A) \approx 2^A$  and  $\mathcal{P}(B) \approx 2^B$  and clearly  $2 \approx 2$  we have the following corollary from the above theorem

**Theorem 5.14.** *If  $A, B$  are sets such that  $A \approx B$  then  $\mathcal{P}(A) \approx \mathcal{P}(B)$  and  $2^A \approx 2^B$*

## 5.2 Infinite and finite sets

First every natural number is in reality a set this is expressed in the following definition and theorem.

**Definition 5.15.** *If  $n \in \mathbb{N}$  then  $S_n = \{m \in \mathbb{N} \mid m < n\}$  (this is a initial segment of  $\mathbb{N}$  (see 2.151))*

**Theorem 5.16.** *If  $n \in \mathbb{N}$  then  $n = S_n$*

**Proof.** Define  $A = \{n \in \mathbb{N} \mid n = S_n\}$  then we have

1. If  $x \in S_0 \Rightarrow 0 \leq x < 0 \Rightarrow 0 < 0$  a contradiction so we have  $S_0 = \emptyset = 0 \Rightarrow 0 \in A$
2. If  $n \in A$  then  $n = S_n$  if we take then  $s(n) = s(S_n) = S_n \cup \{S_n\} = S_n \cup \{n\}$ . Now if  $m \in s(n) \Rightarrow m \in S_n$  or  $m = n \Rightarrow m < n$  or  $m = n < s(n) \Rightarrow m \in S_{s(n)}$ . If  $m \in S_{s(n)}$  then  $m < s(n)$  we have then either
  - a.  $(m < n) \Rightarrow m \in S_n \Rightarrow m \in s(n)$
  - b.  $(n < m) \Rightarrow s(n) \leq m < s(n) \Rightarrow s(n) < s(n)$  a contradiction, so this case can not occur.
  - c.  $(n = m) \Rightarrow m \in \{n\} \Rightarrow m \in S_n \cup \{n\} \Rightarrow m \in s(n)$

So we have proved that  $s(n) = S_{s(n)}$  and thus that  $s(n) \in A$

By the principle of induction (see 4.10) we have that  $A = \mathbb{N}$  and thus  $\forall n \in \mathbb{N}$  that  $n = S_n$ .  $\square$

**Theorem 5.17. (Recursion on  $S_{n+1}$  - Step Form)** *Let  $A$  be a set,  $a \in A$  and  $g: S_n \times A \rightarrow A$  a function. Then there exists a unique function  $f: S_{n+1} \rightarrow A$  satisfying*

$$\begin{aligned} f(0) &= a \\ \forall i \in S_n \text{ we have } f(i+1) &= g(i, f(i)) \end{aligned}$$

**Proof.** Define  $g': \mathbb{N} \times A \rightarrow A$  by  $(i, x) \rightarrow g'(i, x) = \begin{cases} g(i, x) & \text{if } i < n \\ x & \text{if } n \leq i \end{cases}$  then by 4.26 there exists a function  $f': \mathbb{N} \rightarrow A$  such that

$$\begin{aligned} f'(0) &= a \\ \forall i \in \mathbb{N} \text{ we have } f'(s(i)) &= f'(i+1) = g'(i, f'(i)) \end{aligned}$$

Define now  $f: S_{n+1} \rightarrow A$  by  $f = f'_{|S_{n+1}}$  then we have

1.  $f(0) \underset{0 \in S_{n+1} \text{ (as } 0 < n+1)}{=} f'(0) = a$
2.  $\forall i \in S_n$  we have  $i+1 \in S_{n+1}$  and thus  $f(i+1) = f'(i+1) = g'(i, f'(i)) \underset{i < n}{=} g(i, f'(i)) \underset{i \in S_n \subseteq S_{n+1}}{=} g(i, f(i))$

so we proved that  $f$  is the sought for function. Now to prove that it is unique let  $h: S_{n+1} \rightarrow A$  be such that :

1.  $h(0) = a$
2.  $\forall i \in S_n$  we have  $h(i+1) = g(i, h(i))$

Take now  $B = \{i \in \mathbb{N} \mid i \notin S_{n+1} \vee (i \in S_{n+1} \wedge h(i) = f(i))\}$  then we have

1.  $h(0) = a = f(0)$  and thus  $0 \in B$
2. If  $i \in B$  then we have the following cases

- a.  $(s(i) \notin S_{n+1}) \Rightarrow s(i) \in B$
- b.  $(s(i) \in S_{n+1})$  then  $i+1 < n+1 \Rightarrow i < n \Rightarrow i \in S_n$  and  $h(s(i)) = h(i+1) = g(i, h(i)) \underset{i \in B \wedge i \in S_n \subseteq S_{n+1} \Rightarrow h(i) = f(i)}{=} g(i, f(i)) = f(i+1) = f(s(i))$  so  $s(i) \in B$

Using mathematical induction (see 4.10) we have that  $B = \mathbb{N}$ . Now if  $i \in S_{n+1} \Rightarrow i \in \mathbb{N} = B \underset{i \in S_{n+1}}{\Rightarrow} h(i) = f(i)$  and thus  $h = f$   $\square$

We can create a second form of the above theorem as follows

**Theorem 5.18. (Recursion on  $\mathbb{N}$  - Step Form)** Let  $A$  be a set  $a \in A$  and  $g: \mathbb{N} \times A \rightarrow A$  a function then there exists a function  $f: \mathbb{N} \rightarrow A$  such that

$$\begin{aligned} f(0) &= a \\ \forall i \in \mathbb{N} \text{ we have } f(i+1) &= g(i, f(i)) \end{aligned}$$

**Proof.** This is trivial by using 4.26 and the fact that  $s(n) = n + 1$   $\square$

**Corollary 5.19.** Let  $A$  be a set,  $n \in \mathbb{N}$ ,  $a \in A$  and  $g: \{n, \dots\} \times A \rightarrow A$  a function then there exists a function  $f: \{n, \dots\} \rightarrow A$  such that

$$\begin{aligned} f(n) &= a \\ \forall i \in \{n, \dots\} \text{ we have } f(i+1) &= g(i, f(i)) \end{aligned}$$

**Proof.** Take  $g': \mathbb{N} \times A \rightarrow A$  by  $(i, a) \rightarrow g(i+n, a)$  then by the previous theorem there exists a  $f': \mathbb{N} \rightarrow A$  such that

$$\begin{aligned} f'(0) &= a \\ \forall i \in \mathbb{N} \text{ we have } f'(i+1) &= g'(i, f'(i)) \end{aligned}$$

Take now  $f: \{n, \dots\} \rightarrow A$  by  $i \rightarrow f'(i-n)$  then we have

$$\begin{aligned} f(n) &= f'(n-n) = f'(0) = a \\ \forall i \in \{n, \dots\} \text{ we have } i-n &\in \mathbb{N} \text{ and } f(i+1) = f'((i-n)+1) = g'((i-n), f'(i-n)) = g(i, f(i)) \end{aligned}$$

$\square$

**Example 5.20.** Define  $g: \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$  by  $g(i, n) = (i+1) \cdot n$  and take  $a = 1$  then using the above theorem there exists a function  $f: \mathbb{N} \rightarrow \mathbb{N}$  such that

$$\begin{aligned} f(0) &= 1 \\ \forall i \in \mathbb{N} \text{ we have } f(i+1) &= (i+1) \cdot f(i) \end{aligned}$$

We note  $f(i)$  as  $i!$  called the faculty and as some values we have

$$\begin{aligned} 0! &= 1 \\ 1! &= (1+0)! \\ &\quad (1+0) \cdot 0! = 1 \\ 2! &= (1+1)! \\ &= 2 \cdot 1! = 2 \\ 3! &= 3 \cdot 2! = 6 \\ &\dots \end{aligned}$$

The step form of recursion leads to the following corollary using finite sequences and sequences.

**Corollary 5.21.** *Let  $A$  be a set,  $a \in A$ ,  $\mathcal{M} \subseteq \{\{x_i\}_{i \in \{0, \dots, n\}} \mid n \in \mathbb{N} \wedge x_i \in A\}$  [the set of finite families of elements in  $A$  (or the set of function from  $S_n \rightarrow A$ ,  $n \in \mathbb{N}_0$ ] and  $\rho: \mathcal{M} \rightarrow A$  a function then there exists a function  $f: \mathbb{N} \rightarrow A$  such that*

1.  $f(0) = a$
2.  $\forall n \in \mathbb{N}$  we have  $f(s(n)) = \rho(\{f(i)\}_{i \in \{0, \dots, n\}})$

**Proof.** Let  $\{c_i\}_{i \in \{0, \dots, 0\}}$  be defined by  $c_0 = a$  and define now  $g: \mathcal{M} \rightarrow \mathcal{M}$  by  $\{x_i\}_{i \in \{0, \dots, n\}} \mapsto g(\{x_i\}_{i \in \{0, \dots, n\}})$  where  $\forall i \in \{0, \dots, n+1\}$  we have  $g(\{x_i\}_{i \in \{0, \dots, n\}}) = \begin{cases} x_i & \text{if } 0 \leq i \leq n \\ \rho(\{x_i\}_{i \in \{0, \dots, n\}}) & \text{if } i = n+1 \end{cases}$  using recursion (see 4.18) we have then the existence of a function  $h: \mathbb{N} \rightarrow \mathcal{M}$  where

1.  $h(0) = \{c_i\}_{i \in \{0, \dots, 0\}}$
2.  $\forall n \in \mathbb{N}$  we have  $h(n+1) = g(h(n))$

Define then  $S = \{n \in \mathbb{N} \mid \text{dom}(h(n)) = \{0, \dots, n\}\}$  then we have

1. If  $n = 0$  then  $\text{dom}(h(0)) = \text{dom}(\{c_i\}_{i \in \{0, \dots, 0\}}) = \{0, \dots, 0\}$
2. If  $n \in S$  then  $\text{dom}(h(n+1)) = \text{dom}(g(h(n))) \stackrel{\text{definition of } h}{=} \{0, \dots, \text{dom}(h(n)) + 1\} \stackrel{n \in S}{=} \{0, \dots, n+1\}$

Using mathematical induction we have thus that  $S = \mathbb{N} \Rightarrow \forall n \in \mathbb{N}$  that  $\text{dom}(h(n)) = \{0, \dots, n\}$ . Define now  $F = \{n \in \mathbb{N} \mid \text{if } i \leq n \wedge j \leq i \text{ then } h(n)_i = h(i)_j\}$  then we have

1. If  $n = 0$  then from  $i \leq n \wedge j \leq i$  we have  $i = j = 0$  then  $h(n)_j = h(0)_0 = h(i)_j$  or  $0 \in F$
2. If  $n \in F$  then we have that  $\text{dom}(h(n+1)) = \{0, \dots, n+1\}$  and if  $i \leq n+1 \wedge j \leq i$  we have the following cases
  - a.  $(i = n+1)$  then  $h(n+1)_j = h(i)_j$
  - b.  $(i < n+1)$  then  $i \leq n$  and if  $j \leq i$  we have  $h(n+1)_j = g(h(n))_j \stackrel{j < n+1 \text{ and } \text{definition of } g}{=} h(n)_j \stackrel{n \in F}{=} h(i)_j$

So in both cases we have  $h(n+1)_j = h(i)_j$  proving that  $n+1 \in F$

Using mathematical induction we have then that  $F = \mathbb{N}$  or  $\forall n \in \mathbb{N}$  we have then if  $i \leq n$  and  $j \leq i$  then  $h(n)_j = h(i)_j$ . Define now  $f: \mathbb{N} \rightarrow A$  by  $f(n) = h(n)_n$  then we have

1.  $f(0) = h(0)_0 = c_0 = a$
2.  $f(n+1) = h(n+1)_{n+1} = g(h(n))_{n+1} = \rho(h(n))$ , where if  $k \in \text{dom}(h(n)) = \{0, \dots, n\}$  we have that  $h(n)_k \stackrel{k \leq n, k \leq n+1}{=} h(k)_k = f(k)$  so that  $h(n) = \{f(i)\}_{i \in \{0, \dots, n\}}$  giving that  $f(n+1) = \rho(\{f_i\}_{i \in \{1, \dots, n\}})$   $\square$

**Lemma 5.22.** *Let  $A$  be a set  $\mathcal{M} = \{\{x_i\}_{i \in \{0, \dots, n\}} \mid n \in \mathbb{N} \wedge \forall i \in \{0, \dots, n\} \text{ we have } x_i \in A\}$ ,  $m \in \mathbb{N}$  and  $\mathcal{M}_m = \{\{x_i\}_{i \in \{m, \dots, n\}} \mid n \in \{m, \dots\} \wedge \forall i \in \{m, \dots, n\} \text{ we have } x_i \in A\}$ . Define then  $T_m: \mathcal{M} \rightarrow \mathcal{M}_m$  by  $\{x_i\}_{i \in \{0, \dots, n\}} \rightarrow T_m(\{x_i\}_{i \in \{0, \dots, m\}}) = \{(T_m(\{x_i\}_{i \in \{0, \dots, n\}}))_j\}_{j \in \{m, \dots, n+m\}}$  where for  $j \in \{m, \dots, n+m\}$  we have  $(T_m(\{x_i\}_{i \in \{0, \dots, n\}}))_j = x_{j-m}$  then we have*

1.  $T_m$  is a bijection

2. The inverse of  $T_m$  is  $T_{-m}$ :  $\mathcal{M}_m \rightarrow \mathcal{M}$  defined by  $\{x_i\}_{i \in \{m, \dots, n\}} \rightarrow \{(T_{-m}(\{x_i\}_{i \in \{m, \dots, n\}}))_j\}_{j \in \{0, \dots, n-m\}}$  where if  $j \in \{0, \dots, n-m\}$  then  $(T_{-m}(\{x_i\}_{i \in \{m, \dots, n\}}))_j = x_{j+m}$

As a shorthand we note  $\{T_m(\{x_i\}_{i \in \{0, \dots, n\}})\}_{j \in \{m, \dots, n+m\}}$  notation  $\{x_{i-m}\}_{i \in \{m, \dots, n+m\}}$  and  $\{T_m(\{x_i\}_{i \in \{m, \dots, n\}})\}_{j \in \{0, \dots, n-m\}}$  notation  $\{x_{i+m}\}_{i \in \{0, \dots, n-m\}}$ .

**Proof.**

- (injectivity)** If  $\{x_i\}_{i \in \{0, \dots, n_1\}}, \{y_i\}_{i \in \{0, \dots, n_2\}} \in \mathcal{M}$  then if  $T_m(\{x_i\}_{i \in \{0, \dots, n_1\}}) = T_m(\{y_i\}_{i \in \{0, \dots, n_2\}})$  we must have that  $n_1 + m = n_2 + m \Rightarrow n_1 = n_2 = n$  (consider families as functions which must be equal so their domain must be equal) and  $\forall i \in \{0, \dots, n\}$  we have  $i + m \in \{m, \dots, n+m\}$  so that  $x_i = x_{(i+m)-m} = (T_m(\{x_i\}_{i \in \{0, \dots, n\}}))_{i+m} = (T_m(\{y_i\}_{i \in \{0, \dots, n\}}))_{i+m} = y_{(i+m)-m} = y_i$  proving that  $\{x_i\}_{i \in \{0, \dots, n_1\}} = \{y_i\}_{i \in \{0, \dots, n_2\}}$
- (surjectivity)** Let  $\{x_i\}_{i \in \{m, \dots, n\}} \in \mathcal{M}_m$  define then  $\{y_i\}_{i \in \{0, \dots, n-m\}}$  by  $\forall i \in \{0, \dots, n-m\}$  we have  $y_i = x_{i+m}$  then  $\forall i \in \{m, \dots, n\}$  we have  $(T_m(\{y_i\}_{i \in \{0, \dots, n-m\}}))_i = y_{i-m} = x_{(i+m)-m} = x_i$  proving that  $T_m(\{y_i\}_{i \in \{0, \dots, n-m\}}) = \{x_i\}_{i \in \{m, \dots, n\}}$

Finally to prove that  $T_{-m}$  is the inverse of  $T_m$ , take  $\{x_i\}_{i \in \{0, \dots, n\}}$  then  $T_{-m}(T_m(\{x_i\}_{i \in \{0, \dots, n\}})) = T_{-m}(\{(T_m(x_i)_{i \in \{0, \dots, n\}})_j\}_{j \in \{m, \dots, n+m\}}) = \{(T_{-m}(\{(T_m(x_i)_{i \in \{0, \dots, n\}})_j\}_{j \in \{m, \dots, n+m\}}))_k\}_{k \in \{0, \dots, (m+n)-m\}} = \{(T_{-m}(\{x_{j-m}\}_{j \in \{m, \dots, n+m\}}))_k\}_{k \in \{0, \dots, n\}} = \{x_{(k-m)+m}\}_{k \in \{0, \dots, n\}} = \{x_k\}_{k \in \{0, \dots, n\}}$   $\square$

**Corollary 5.23.** Let  $A$  be a set,  $a \in A, m \in \mathbb{N}$ ,  $\mathcal{M}_m = \{\{x_i\}_{i \in \{m, \dots, n\}} | n \in \{m, \dots, n\} \wedge x_i \in A\}$  /the set of finite families of elements in  $A$  and  $\rho: \mathcal{M}_m \rightarrow A$  a function then there exists a function  $f: \{m, \dots\} \rightarrow A$  such that

- $f(m) = a$
- $\forall n \in \{m, \dots\}$  we have  $f(n+1) = f(s(n)) = \rho(\{f(i)\}_{i \in \{m, \dots, n\}})$

**Proof.** If we define  $\mathcal{M} = \{\{x_i\}_{i \in \{0, \dots, n\}} | n \in \mathbb{N} \wedge \forall i \in \{0, \dots, n\}$  we have  $x_i \in A\}$  then using the previous lemma (see 5.22) we have the bijections  $T_m: \mathcal{M} \rightarrow \mathcal{M}_m$  and  $T_{-m}: \mathcal{M}_m \rightarrow \mathcal{M}$ , define now  $\rho': \mathcal{M} \rightarrow A$  by  $\rho' = \rho \circ T_m$  then  $\rho': \mathcal{M} \rightarrow A$  and by 5.21 there exists a  $f': \mathbb{N} \rightarrow A$  such that

- $f'(0) = a$
- $f'(n+1) = \rho'(\{f(i)\}_{i \in \{0, \dots, n\}})$

Define now  $f: \{m, \dots\} \rightarrow A$  by  $f(i) = f'(i-m)$  so that if  $j \in \mathbb{N}$  then  $f'(j) = f(j+m)$  then we have

- $f(m) = f'(m-m) = f'(0) = a$
- $\forall n \in \{m, \dots\}$  we have  $f(n+1) = f'((n+1)-m) = f'((n-m)+1) = \rho'(\{f'(i)_{i \in \{0, \dots, n-m\}}\}) = \rho(T_m(\{f'(i)_{i \in \{0, \dots, n-m\}}\})) = \rho(T_m(\{f(i+m)\}_{i \in \{0, \dots, n-m\}})) = \rho(T_m(T_{-m}(\{f(i)\}_{i \in \{m, \dots, n\}}))) = \rho(\{f(i)\}_{i \in \{m, \dots, n\}})$   $\square$

**Corollary 5.24.** Let  $A$  be a set,  $a \in A, m \in \mathbb{N}$ ,  $\mathcal{M} = \{\{x_i\}_{i \in \{m, \dots, n\}} | n \in \{m, \dots\} \wedge \forall i \in \{m, \dots, n\}$  we have  $x_i \in A\}$  /the set of finite families of elements in  $A$  and  $\rho: \mathcal{M} \rightarrow A$  a function so that  $\forall \{x_i\}_{i \in \{m, \dots, n\}} \in \mathcal{M}$  we have that  $P(\{x_i\}_{i \in \{m, \dots, n\}}, \rho(\{x_i\}_{i \in \{m, \dots, n\}}))$  is true then there exists a function  $f: \{m, \dots\} \rightarrow A$  such that

- $f(m) = a$
- $\forall n \in \mathbb{N}$  we have  $f(s(n)) = \rho(\{f(i)\}_{i \in \{m, \dots, n\}})$
- $\forall n \in \mathbb{N}$  we have that  $P(\{f(i)\}_{i \in \{m, \dots, n\}}, f(n+1))$  is true

**Proof.** First we use the previous corollary to prove that there exists a  $f: \{m, \dots\} \rightarrow A$  such that

- $f(m) = a$
- $\forall n \in \{m, \dots\}$  we have  $f(s(n)) = \rho(\{f(i)\}_{i \in \{m, \dots, n\}})$

We proceed now by induction so let  $S = \{n \in \{m, \dots\} \mid P(\{f(i)\}_{i \in \{m, \dots, n\}}, f(n+1)) \text{ is true}\}$  then we have

1. if  $n = m$  then as  $P(\{f(i)\}_{i \in \{m, \dots, m\}}, f(m+1)) = P(\{f(i)\}_{i \in \{m, \dots, m\}}, \rho(\{f(i)\}_{i \in \{m, \dots, m\}}))$  is true by the hypothesis so that  $m \in S$
2. If  $n \in S$  then by assumption we have that  $P(\{f(i)\}_{i \in \{m, \dots, n\}}, \rho(\{f(i)\}_{i \in \{m, \dots, n\}})) = P(\{f(i)\}_{i \in \{m, \dots, n\}}, f(n+1))$  so we have that  $n+1 \in S$

Which by mathematical induction proves the theorem.  $\square$

Another variant of recursion and families is the following

**Theorem 5.25.** *Let  $A$  be a set and  $\mathcal{M} = \{\{x_i\}_{i \in \{0, \dots, n\}} \mid \forall i \in \{0, \dots, n\} \text{ we have } x_i \in A\}$ ,  $\mathcal{N} \subseteq \mathcal{M}$  and  $\rho: \mathcal{N} \rightarrow A$  such that for  $\rho(\{x_i\}_{i \in \{0, \dots, n\}})$  we have  $\rho(\{x_i\}_{i \in \{0, \dots, n\}}) \in A$  and  $\{x'_i\}_{i \in \{0, \dots, n+1\}} \in \mathcal{N}$  (where  $x'_i = \begin{cases} x_i & \text{if } i \in \{0, \dots, n\} \\ \rho(\{x_i\}_{i \in \{0, \dots, n\}}) & \text{if } i = n+1 \end{cases}$ ). Then for  $a \in A$  with  $\{a\}_{i \in \{0, \dots, 0\}} \in \mathcal{N}$  there exists a function  $f: \mathbb{N} \rightarrow A$  such that*

1.  $f(0) = a$
2.  $\forall n \in \mathbb{N} \text{ we have } f(n+1) = \rho(\{f(i)\}_{i \in \{0, \dots, n\}})$
3.  $\forall n \in \mathbb{N} \text{ we have } \{f(i)\}_{i \in \{0, \dots, n\}} \in \mathcal{N}$

**Proof.** Take  $\{a\}_{i \in \{0, \dots, 0\}} \in \mathcal{N}$  (by the hypothesis) and define  $g: \mathcal{N} \rightarrow \mathcal{N}$  by  $\{x_i\}_{i \in \{0, \dots, n\}} \mapsto g(\{x_i\}_{i \in \{0, \dots, n\}})$  where  $\forall i \in \{0, \dots, n+1\}$  we have  $g(\{x_i\}_{i \in \{0, \dots, n\}}) = \begin{cases} x_i & \text{if } 0 \leq i \leq n \\ \rho(\{x_i\}_{i \in \{0, \dots, n\}}) & \text{if } i = n+1 \end{cases} \in \mathcal{N}$  (by the hypothesis), using recursion (see 4.18) we have then the existence of a function  $h: \mathbb{N} \rightarrow \mathcal{N}$  where

1.  $h(0) = \{a\}_{i \in \{0, \dots, 0\}}$
2.  $\forall n \in \mathbb{N} \text{ we have } h(n+1) = g(h(n))$

Define then  $S = \{n \in \mathbb{N} \mid \text{dom}(h(n)) = \{0, \dots, n\}\}$  then we have

1. If  $n = 0$  then  $\text{dom}(h(0)) = \text{dom}(\{a\}_{i \in \{0, \dots, 0\}}) = \{0, \dots, 0\}$
2. If  $n \in S$  then  $\text{dom}(h(n+1)) = \text{dom}(g(h(n))) \stackrel{\text{definition of } h}{=} \{0, \dots, \text{dom}(h(n))+1\} \stackrel{n \in S}{=} \{0, \dots, n+1\}$

Using mathematical induction we have thus that  $S = \mathbb{N} \Rightarrow \forall n \in \mathbb{N} \text{ that } \text{dom}(h(n)) = \{0, \dots, n\}$ . Define now  $F = \{n \in \mathbb{N} \mid \text{if } i \leq n \wedge j \leq i \text{ then } h(n)_i = h(i)_j\}$  then we have

1. If  $n = 0$  then from  $i \leq n \wedge j \leq i$  we have  $i = j = 0$  then  $h(n)_j = h(0)_0 = h(i)_j$  or  $0 \in F$
2. If  $n \in F$  then we have that  $\text{dom}(h(n+1)) = \{0, \dots, n+1\}$  and if  $i \leq n+1 \wedge j \leq i$  we have the following cases
  - a.  $(i = n+1)$  then  $h(n+1)_j = h(i)_j$
  - b.  $(i < n+1)$  then  $i \leq n$  and if  $j \leq i$  we have  $h(n+1)_j = g(h(n))_j \stackrel{j < n+1 \text{ and definition of } g}{=} h(n)_j \stackrel{n \in F}{=} h(i)_j$

So in both cases we have  $h(n+1)_j = h(i)_j$  proving that  $n+1 \in F$

Using mathematical induction we have then that  $F = \mathbb{N}$  or  $\forall n \in \mathbb{N}$  we have then if  $i \leq n$  and  $j \leq i$  then  $h(n)_j = h(i)_j$ . Define now  $f: \mathbb{N} \rightarrow A$  by  $f(n) = h(n)_n$  then we have

1.  $f(0) = h(0)_0 = c_0 = a$
2.  $f(n+1) = h(n+1)_{n+1} = g(h(n))_{n+1} = \rho(h(n))$ , where if  $k \in \text{dom}(h(n)) = \{0, \dots, n\}$  we have that  $h(n)_k \stackrel{k \leq n, k \leq n+1}{=} h(k)_k = f(k)$  so that  $h(n) = \{f(i)\}_{i \in \{0, \dots, n\}}$  giving that  $f(n+1) = \rho(\{f_i\}_{i \in \{1, \dots, n\}})$

Let  $\mathcal{W} = \{n \in \mathbb{N} \mid \{f(i)\}_{i \in \{0, \dots, n\}} \in \mathcal{N}\}$  then we have:

1.  $\{f(i)\}_{i \in \{0, \dots, 0\}} = \{f(0)\}_{i \in \{0, \dots, 0\}} = \{a\}_{i \in \{0, \dots, 0\}} \in \mathcal{N}$
2. If  $n \in \mathcal{W}$  then  $\{f(i)\}_{i \in \{0, \dots, n+1\}} = \rho(\{f(i)\}_{i \in \{0, \dots, n+1\}}) \in \mathcal{N}$  as  $\{f(i)\}_{i \in \{0, \dots, n\}} \in \mathcal{N}$  [ $n \in \mathcal{W}$ ] and the hypothesis about  $\rho$ .

Applying mathematical induction on  $\mathcal{W}$  gives then  $\mathcal{W} = \mathbb{N}$  and thus (3) in our theorem.  $\square$

**Corollary 5.26.** Let  $A$  be a set and  $\mathcal{M}_m = \{\{x_i\}_{i \in \{m, \dots, n\}} \mid n \in \{m, \dots\} \wedge \forall i \in \{0, \dots, n\} \text{ we have } x_i \in A\}$ ,  $\mathcal{N}_m \subseteq \mathcal{M}_m$  and  $\rho: \mathcal{N}_m \rightarrow A$  such that for  $\rho(\{x_i\}_{i \in \{m, \dots, n\}})$  we have  $\rho(\{x_i\}_{i \in \{m, \dots, n\}}) \in A$  and  $\{x'_i\}_{i \in \{m, \dots, n+1\}} \in \mathcal{N}_m$  (where  $x'_i = \begin{cases} x_i & \text{if } i \in \{m, \dots, n\} \\ \rho(\{x_i\}_{i \in \{m, \dots, n\}}) & \text{if } i = n+1 \end{cases}$ ). Then for  $a \in A$  with  $\{a\}_{i \in \{m, \dots, n\}} \in \mathcal{N}_m$  there exists a function  $f: \mathbb{N} \rightarrow A$  such that

1.  $f(m) = a$
2.  $\forall n \in \{m, \dots\}$  we have  $f(n+1) = \rho(\{f(i)\}_{i \in \{m, \dots, n\}})$
3.  $\forall n \in \{m, \dots\}$  we have  $\{f(i)\}_{i \in \{m, \dots, n\}} \in \mathcal{N}$

**Proof.** If we define  $\mathcal{M} = \{\{x_i\}_{i \in \{0, \dots, n\}} \mid n \in \mathbb{N} \wedge \forall i \in \{0, \dots, n\} \text{ we have } x_i \in A\}$  then using 5.22 we have the bijections  $T_m: \mathcal{M} \rightarrow \mathcal{M}_m$  and  $T_{-m}: \mathcal{M}_m \rightarrow \mathcal{M}$ . Define then  $\mathcal{N} = T_{-m}(\mathcal{N}_m)$  so that we have the bijections  $T_{-m|N_m}: \mathcal{N}_m \rightarrow \mathcal{N}$  and  $T_{m|N}: \mathcal{N} \rightarrow \mathcal{N}_m$ . Define then  $\rho': \mathcal{N} \rightarrow A$  as  $\rho' = \rho \circ T_{m|N}$  then if  $\{x_i\}_{i \in \{0, \dots, n\}} \in \mathcal{N}$  we have  $\rho'(\{x_i\}_{i \in \{0, \dots, n\}}) = \rho(T_m(\{x_i\}_{i \in \{0, \dots, n\}})) = \rho(\{x_{i-m}\}_{i \in \{m, \dots, n+m\}})$ . If we then create  $\{x'_i\}_{i \in \{0, \dots, n+1\}} \in \mathcal{N}$  by  $x'_i = \begin{cases} x_i & \text{if } i \in \{0, \dots, n\} \\ \rho'(\{x_i\}_{i \in \{0, \dots, n\}}) & \text{if } i = n+1 \end{cases} = \begin{cases} x_i & \text{if } i \in \{0, \dots, n\} \\ \rho(T_m(\{x_i\}_{i \in \{0, \dots, n\}})) & \text{if } i = n+1 \end{cases}$  then we have for  $T_{m|N}(\{x'_i\}_{i \in \{0, \dots, n+1\}})$  that if  $k \in \{m, \dots, n+m+1\}$   $(T_{m|N}(\{x'_i\}_{i \in \{0, \dots, n+1\}}))_k = x'_{j-m} = \begin{cases} T_m(\{x_i\}_{i \in \{0, \dots, n\}})_k & \text{if } k \in \{m, \dots, n+m\} \\ \rho(T_m(\{x_i\}_{i \in \{0, \dots, n\}})) & \text{if } k = n+m+1 \end{cases}$  so by the properties of  $\rho$  we have that  $T_m(\{x'_i\}_{i \in \{0, \dots, n+1\}}) \in \mathcal{N}_m$  so that  $\{x'_n\}_{i \in \{0, \dots, n+1\}} \in \mathcal{N}$ . Also using the fact that  $\{a\}_{i \in \{m, \dots, n\}} \in \mathcal{N}_m$  we have that  $\{a\}_{i \in \{0, \dots, 0\}} = T_{-m}(\{a\}_{i \in \{0, \dots, 0\}}) \in \mathcal{N}$ . Using the above theorem (see 5.25) there exists then a  $f': \mathbb{N} \rightarrow A$  such that

1.  $f'(0) = a$
2.  $\forall n \in \mathbb{N}$  we have  $f'(n+1) = \rho'(\{f'(i)\}_{i \in \{0, \dots, n\}})$
3.  $\forall n \in \mathbb{N}$  we have  $\{f'(i)\}_{i \in \{0, \dots, n\}} \in \mathcal{N}$

Define then  $f: \{m, \dots\} \rightarrow A$  by  $i \rightarrow f(i) = f'(i-m)$  or if  $i \in \mathbb{N}$  then  $f'(i) = f(i+m)$  then we have

1.  $f(m) = f'(m-m) = f'(0) = a$
2.  $\forall n \in \{m, \dots\}$  we have  $f(n+1) = f'((n-m)+1) = \rho'(\{f'(i)\}_{i \in \{0, \dots, (n-m)\}}) = \rho'(\{f(i+m)\}_{i \in \{0, \dots, (n-m)\}}) = \rho'(T_{-m}(\{f_i\}_{i \in \{m, \dots, n\}})) = \rho(T_m(T_{-m}(\{f_i\}_{i \in \{m, \dots, n\}}))) = \rho(\{f(i)\}_{i \in \{m, \dots, n\}})$
3.  $\forall n \in \{m, \dots\}$  we have  $\{f'(i)\}_{i \in \{0, \dots, n-m\}} \in \mathcal{N}$  so that  $T_m(\{f'(i)\}_{i \in \{0, \dots, n-m\}}) \in T_m(T_{-m}(\mathcal{N}_m)) = \mathcal{N}_m$ , as  $T_m(\{f'(i)\}_{i \in \{0, \dots, n-m\}}) = \{f'(i-m)\}_{i \in \{m, \dots, n\}} = \{f(i)\}_{i \in \{m, \dots, n\}}$  this means that  $\{f(i)\}_{i \in \{m, \dots, n\}} \in \mathcal{N}$   $\square$

**Definition 5.27.** A set  $A$  is called finite if  $\exists n \in \mathbb{N}$  such that  $n \approx A$ . A set that is not finite is called a infinite set.

**Note 5.28.**  $\emptyset: \emptyset \rightarrow 0 = \emptyset$  is a bijection as all the conditions of a bijection are vacuously true, so the empty set is finite.

**Notation 5.29.** If  $A \neq \emptyset$  is a finite set then there exists a bijection  $a: S_{n+1} \rightarrow A$  and we note then  $A = \{a_0, \dots, a_n\}$  where  $a_i = a(i)$  (note that sometimes if  $n \in \mathbb{N}_0$  we write  $A = \{b_1, \dots, b_n\}$  where  $b_i = a(i+1)$  and  $a: S_n \rightarrow A$  is a bijection)

**Definition 5.30.** A set  $A$  is called **denumerable** or **infinite countable** if  $A \approx \mathbb{N}$

**Definition 5.31.** A set  $A$  is called **countable** if it is either **finite** or **infinite countable**

**Lemma 5.32.** If  $A$  is a **denumerable** set and  $x \in A$  then  $A \setminus \{x\}$  is a **denumerable** set.

**Proof.** If  $A$  is countable then  $A \approx \mathbb{N}$  and thus there exists a bijection  $f: \mathbb{N} \rightarrow A$ . By surjectivity of  $f$  and  $x \in A$  there exists a  $n \in \mathbb{N}$  such that  $f(n) = x$ . Define now  $g: \mathbb{N} \rightarrow A \setminus \{x\}$  by  $g(m) = \begin{cases} f(m) & \text{if } m < n \\ f(m+1) & \text{if } n \leq m \end{cases}$ . We prove now that  $g$  is bijective:

1. **(injectivity)** If  $g(m) = g(m')$  then we have the following possible cases
  - a.  $(m < n \wedge m' < n)$  then we have  $f(m) = g(m) = g(m') = f(m') \Rightarrow f(m) = f(m')$   $\underset{f \text{ is injective}}{\Rightarrow} m = m'$

- b.  $(m < n \wedge n \leq m')$  then we have  $f(m) = g(m) = g(m') = f(m'+1) \Rightarrow f(m) = f(m'+1) \Rightarrow m = m'+1$ . Then as  $n \leq m' \Rightarrow n < n+1 \leq m'+1 = m$  we have from  $m < n$  that  $m < m$  a contradiction, so this case does not apply.
  - c.  $(n \leq m \wedge m' < n)$  then we have  $f(m+1) = g(m) = g(m') = f(m') \Rightarrow f(m+1) = f(m') \Rightarrow m+1 = m' \underset{n \leq m}{\nRightarrow} n+1 \leq m+1 = m' \underset{m' < n}{\nRightarrow} n+1 < n < n+1 \Rightarrow n+1 < n+1$  again a contradiction, so we can ignore this case as it never occurs.
  - d.  $(n \leq m \wedge n \leq m')$  then  $f(m+1) = g(m) = g(m') = f(m'+1) \Rightarrow f(m+1) = f(m'+1) \Rightarrow m+1 = m'+1 \Rightarrow m = m'$
2. **(surjectivity)** If  $y \in A \setminus \{x\}$  then there exists by surjectivity of  $g$  a  $i \in \mathbb{N}$  such that  $f(i) = y$ . We can not have that  $i = n$  as then  $y = f(i) = x \notin A \setminus \{x\}$ . So the only cases for  $i$  are
- a.  $(i < n)$  then  $g(i) = f(i) = y \Rightarrow g(i) = y$
  - b.  $(n < i)$  then  $n+1 \leq i \Rightarrow n \leq i-1 \Rightarrow g(i-1) = f((i-1)+1) = f(i) = y \Rightarrow g(i-1) = y$
- proving surjectivity.  $\square$

**Lemma 5.33.** *If  $n \in \mathbb{N}$  then  $n$  has no denumerable subset*

**Proof.** We prove this by induction on  $n$ , so let  $S = \{n \in \mathbb{N} \mid n \text{ does not contain a denumerable subset}\}$ .

1. First as  $0 = \emptyset$  we have if  $A \subseteq 0 = \emptyset \Rightarrow \emptyset \subseteq A \subseteq \emptyset \Rightarrow A = \emptyset$ . If now  $\mathbb{N} \approx A$  then there exists a bijection  $f: \mathbb{N} \rightarrow A$  and thus  $f(0) \in A \Rightarrow A \neq \emptyset$  a contradiction. So  $0$  does not contain a denumerable subset and we have  $0 \in S$ .
2. Assume now that  $n \in S$ . We proceed now by contradiction, so assume that there exists a  $A \subseteq n+1 = s(n) = n \cup \{n\}$  such that  $A$  is denumerable. If  $n \notin A \Rightarrow A \subseteq n$  which is impossible as  $n \in S \Rightarrow n \in A$ . Now if  $x \in A \setminus \{n\} \subseteq A \subseteq n \cup \{n\}$  then  $x \in n \cup \{n\}$  and from  $x \neq n$  we have that  $x \in n$  so we have  $A \setminus \{n\} \subseteq n$  and as by the previous lemma we would have that  $A \setminus \{n\}$  is denumerable, we reach the contradiction that  $n \notin S$ . So we can not have a countable subset of  $n+1$  and thus  $n+1 \in S$

Using 4.10 we conclude that  $S = \mathbb{N}$  proving our theorem.  $\square$

**Theorem 5.34.** *If  $A$  is an infinite set if and only  $A$  has a denumerable subset. So any denumerable set must be infinite.*

**Proof.** First we prove that if  $A$  is infinite then  $A$  has a denumerable subset. Using the 'well ordering theorem' which is derived from the axiom of choice (see 2.227) there exists a order  $\leq_A$  on  $A$  such that  $\langle A, \leq_A \rangle$  is a well ordered set. Using 2.204 and the fact that  $\langle \mathbb{N}, \leq \rangle$  is also well ordered (see 4.52) we have exactly one of the following:

1.  $(\langle \mathbb{N}, \leq \rangle \text{ is isomorphic with } \langle A, \leq_A \rangle)$  but this would mean that  $A \approx \mathbb{N}$  and thus  $A$  has a denumerable subset.
2.  $(\langle \mathbb{N}, \leq \rangle \text{ is isomorphic with an initial segment of } \langle A, \leq_A \rangle)$  but this would again mean that  $A$  has a denumerable subset.
3.  $(\langle A, \leq_A \rangle \text{ is isomorphic with an initial segment of } \langle \mathbb{N}, \leq \rangle)$  But then there exists a  $n \in \mathbb{N}$  so that  $A \approx S_n = n$  and thus  $A$  is finite which contradicts the fact that  $A$  is infinite, so this case does not apply.

Second assume that  $A$  has a denumerable subset  $B$ . If  $A$  is finite then there exists a  $n \in \mathbb{N}$  such that  $A \approx n$  and thus there exists a bijection  $f: A \rightarrow n$ . Then  $f(B) \subseteq n$  and as  $f|_B: B \rightarrow f(B)$  is a bijection and as  $\mathbb{N} \approx B$  we have the existence of a bijection  $g: \mathbb{N} \rightarrow B$  so  $f|_B \circ g: \mathbb{N} \rightarrow f(B)$  is a bijection and thus  $f(B)$  is a denumerable subset of  $n$  which is impossible by the previous lemma. we must thus conclude that  $A$  is not finite and thus is infinite.  $\square$

**Corollary 5.35.** *Every set with an infinite subset is infinite.*

**Proof.** If  $A$  is a set such that there exists an infinite set  $B$  with  $B \subseteq A$  then as  $B$  is infinite we have by 5.34 the existence of a denumerable set  $C \subseteq B$  but then  $C \subseteq A$  and thus  $A$  has a denumerable subset and by 5.34 we have that  $A$  is infinite.  $\square$

**Corollary 5.36.** *Every subset of a finite set is finite.*

**Proof.** If a finite set would contain a infinite subset then by the previous theorem the finite set would be infinite.  $\square$

**Theorem 5.37.** *If  $A, B$  are finite sets then  $A \cup B$  is finite*

**Proof.** As  $A$  is finite we have by 5.36 that  $A \setminus B$  is finite, this together with the finiteness of  $B$  give the existence of two bijections  $f: A \setminus B \rightarrow n = S_n$  and  $g': B \rightarrow m = S_m$  where  $n, m \in \mathbb{N}$ . Define now  $g: B \rightarrow \{i \in \mathbb{N} \mid n \leq i \wedge i < n+m\}$  by  $g(x) = g'(x) + n$  then this function is bijective

1. **(injectivity)** If  $g(x) = g(x')$  then  $g'(x) + n = g'(x') + n \xrightarrow{4.44} g'(x) = g'(x') \Rightarrow x = x'$
2. **(surjectivity)** If  $i \in \{i \in \mathbb{N} \mid n \leq i \wedge i < n+m\}$  then  $n \leq i < n+m$  and thus using 4.62 there exists a  $k \in \mathbb{N}$  such that  $n+k = i$ . We must have  $k < m$  [if  $m \leq k \Rightarrow n+m \leq n+k = i < n+m \Rightarrow$  we get the contradiction  $n+m \neq n+m$ ] and thus  $k \in S_m = m$ . So using the fact that  $g'$  is bijective and thus surjective we have  $\exists x \in B$  such that  $g'(x) = k \Rightarrow g(x) = g'(x) + n = k + n = i$ .

Further if  $i \in n \cap \{i \in \mathbb{N} \mid n \leq i \wedge i < n+m\} \xrightarrow{n=S_n} i < n \wedge n \leq i \Rightarrow i \neq i$  a contradiction, so we have  $n \cap \{i \in \mathbb{N} \mid n \leq i \wedge i < n+m\} = \emptyset$ . So we can use 2.46 to construct the bijection  $f \cup g: A \cup B = (A \setminus B) \cup B \rightarrow n \cup \{i \in \mathbb{N} \mid n \leq i \wedge i < n+m\} = S_{n+m} = n+m$  and thus we have proved that  $A \cup B \approx n+m$ .  $\square$

**Lemma 5.38.** *If  $\{A_i\}_{i \in S_n}$  is a finite family of finite sets then  $\bigcup_{i \in S_n} A_i$  is finite.*

**Proof.** We prove this by induction on  $n$ . So define  $S = \{n \in \mathbb{N} \mid \text{if } \{A_i\}_{i \in S_n} \text{ is a family of finite set then } \bigcup_{i \in S_n} A_i \text{ is finite}\}$  then we have

1. If  $n = 0$  take then  $\{A_i\}_{i \in S_0}$  then there exists a graph  $A$  with  $\text{dom}(A) = \emptyset$  so if  $x \in \bigcup_{i \in S_0} A_i$  then there exists a  $i \in S_0$  such that  $x \in A_i$  which is a contradiction as  $S_0 = \emptyset$  so we have that  $\bigcup_{i \in S_0} A_i = \emptyset$  and thus finite. So we have  $0 \in S$
2. Assume that  $n \in S$  and take  $n+1$  then we have for the family of finite sets  $\{A_i\}_{i \in S_{n+1}}$  that there exists a graph  $A$  with  $\text{dom}(A) = S_{n+1}$ . If we take the family  $A' = \{(x, y) \in A \mid x \in S_n\}$  [so  $\text{dom}(A') = S_n$  and  $A'_i = \{(x, y) \in A' \mid x = i\}$ ] then  $\bigcup_{i \in S_{n+1}} A_i = A_n \cup (\bigcup_{i \in S_n} A'_i)$

**Proof.**

- a. If  $x \in \bigcup_{i \in S_{n+1}} A_i$  then  $\exists i \in \text{dom}(A) = S_{n+1} \Rightarrow i < n+1$  such that  $x \in A_i$  then we have the following cases

- i.  $(i = n)$  then  $x \in A_n \Rightarrow x \in A_n \cup (\bigcup_{i \in S_n} A'_i)$
- ii.  $(i \neq n)$  then  $i < n$  [if  $n \leq i \xrightarrow{i \neq n} n < i < n+1 \Rightarrow n+1 \leq i < n+1 \Rightarrow n+1 \neq n+1$  a contradiction] and thus  $i \in S_n$ . As  $x \in A_i$  we have  $(i, x) \in A \xrightarrow{i \in S_n} (i, x) \in A' \Rightarrow x \in A'_i \Rightarrow x \in \bigcup_{i \in S_n} A'_i \Rightarrow x \in A_n \cup (\bigcup_{i \in S_n} A'_i)$

in all cases we have  $x \in A_n \cup (\bigcup_{i \in S_n} A'_i)$

- b. If  $x \in A_n \cup (\bigcup_{i \in S_n} A'_i)$  then we have the following cases

- i.  $(x \in A_n) \Rightarrow x \in \bigcup_{i \in S_{n+1}} A_i$
- ii.  $(x \notin A_n) \Rightarrow x \in \bigcup_{i \in S_n} A'_i \Rightarrow \exists i \in S_n \vdash x \in A'_i \Rightarrow (i, x) \in A' \subseteq A \Rightarrow x \in A_i \Rightarrow x \in \bigcup_{i \in S_{n+1}} A_i$

in all cases  $x \in \bigcup_{i \in S_{n+1}} A_i$   $\square$

As  $n \in S$  we have  $\bigcup_{i \in S_n} A'_i$  is finite and as by assumption  $A_n$  is finite we have by 5.37 that  $\bigcup_{i \in S_{n+1}} A_i = A_n \cup (\bigcup_{i \in S_n} A'_i)$  is finite and thus  $n+1 \in S$ .

By induction we have then  $S = \mathbb{N}$  proving our theorem.  $\square$

**Theorem 5.39.** *If  $\{A_i\}_{i \in I}$  is a family of finite sets with  $I$  finite then  $\bigcup_{i \in I} A_i$  is finite*

**Proof.** As  $I$  is finite we have the existence of a bijection  $f: n = S_n \rightarrow I$  ( $n \in \mathbb{N}$ ) and thus using 2.68 we have that  $\bigcup_{i \in I} A_i = \bigcup_{j \in S_n} A_{f(j)}$  which is finite by the previous lemma.  $\square$

**Theorem 5.40.** *A set  $A$  is infinite if and only if  $\exists B \subset A$  such that  $B \approx A$  (if  $A$  is equipotent with a proper subset of itself)*

**Proof.**

First, suppose  $A$  is infinite then by 5.34 there exists a denumerable subset  $B \subseteq A$ . So there exists a bijection  $b: \mathbb{N} \rightarrow B$  define then the function  $f: A \rightarrow A$  by  $f(x) = \begin{cases} x & \text{if } x \in A \setminus B \\ b(b^{-1}(x)+1) & \forall x \in B \end{cases}$  then we have that

1. ( $f(A) = A \setminus \{b(0)\}$ )

**Proof.** If  $y \in f(A)$  then there exists a  $x \in A$  such that  $y = f(x)$  and then either

- a. ( $x \in A \setminus B$ ) and thus  $f(x) = x \in A \setminus B \Rightarrow f(x) = x \neq b(0) \in B \Rightarrow f(x) \in A \setminus \{b(0)\}$
- b. ( $x \in B$ ) and thus as  $b^{-1}(x) \in \mathbb{N} \Rightarrow b^{-1}(x) \geq 0 \Rightarrow b^{-1}(x)+1 > 0$ , if we would have  $b(b^{-1}(x)+1) = b(0)$  then as  $b$  is a bijection we have  $0 = b^{-1}(x)+1 > 0$  giving the contradiction  $0 > 0$  so we must have that  $f(x) = b(b^{-1}(x)+1) \neq b(0) \Rightarrow f(x) \in A \setminus \{b(0)\}$  giving  $f(A) \subseteq A \setminus \{b(0)\}$ .

Also if  $y \in A \setminus \{b(0)\}$  then we have either

- a. ( $y \notin B$ )  $\Rightarrow y \in (A \setminus \{b(0)\}) \setminus B$  then  $f(y) = y \Rightarrow y \in f(A)$
- b. ( $y \in B$ ) then as  $b$  is a bijection  $y = b(b^{-1}(y))$  and as  $y \neq b(0)$  we can not have  $b^{-1}(y) = 0$  [otherwise  $y = b(b^{-1}(y)) = b(0)$ ], so  $0 < b^{-1}(y) \Rightarrow 1 \leq b^{-1}(y) \Rightarrow \exists m \in \mathbb{N} \vdash m+1 = b^{-1}(y)$  and thus  $b(m+1) = b(b^{-1}(y)) = y$ . Take now  $x = b(m) \in B \subseteq A$  then  $f(x) = b(b^{-1}(x)+1) = b(m+1) = y \Rightarrow y \in f(A)$ .

giving that  $A \setminus \{b(0)\} \subseteq f(A)$  and thus together with the already proven  $f(A) \subseteq A \setminus \{b(0)\}$  that  $f(A) = A \setminus \{b(0)\}$   $\square$

2. ( $f$  is injective)

**Proof.** If  $f(x) = f(x')$  then we have the following possible cases to consider

- a. ( $x, y \in A \setminus B$ ) then  $x = f(x) = f(x') = x' \Rightarrow x = x'$
- b. ( $x \in A \setminus B \wedge x' \in B$ ) then  $f(x') = f(x) = x \in A \setminus B \Rightarrow f(x') \notin B$  contradicting the fact that  $f(x') = b(b^{-1}(x')+1) \in B$ , so this case does not apply.
- c. ( $x \in B \wedge x' \in A \setminus B$ ) then  $f(x) = f(x') = x' \in A \setminus B \Rightarrow f(x) \notin B$  contradicting the fact that  $f(x) = b(b^{-1}(x)+1) \in B$  so this case does not apply.
- d. ( $x, x' \in B$ ) then we have  $b(b^{-1}(x)+1) = f(x) = f(x') = b(b^{-1}(x')+1) \stackrel{\text{b is injective}}{\Rightarrow} b^{-1}(x)+1 = b^{-1}(x')+1 \Rightarrow b^{-1}(x) = b^{-1}(x') \Rightarrow x = x'$

so in all the cases that do not lead to a contradiction we have  $x = x'$  proving injectivity.  $\square$

From (1) and (2) we conclude that  $f: A \rightarrow A \setminus \{b(0)\}$  is a bijection and thus that  $A$  is bijective with a proper subset of itself.

Second assume that there exists a proper subset  $B \subset A$  and a bijection  $f: A \rightarrow B$ , giving a injective function  $f: A \rightarrow A$  with  $f(A) = B$ . As  $B \subset A$  there exists a  $c \in A \setminus B$  and thus  $c \notin f(A)$ . Using 4.19 there exists a injective function  $\lambda: \mathbb{N} \rightarrow A$  such that

1.  $\lambda(0) = c$

$$2. \forall n \in \mathbb{N} \models \lambda(s(n)) = f(\lambda(n))$$

So we found a bijective function  $\lambda: \mathbb{N} \rightarrow \lambda(\mathbb{N}) \subseteq A$  giving a denumerable set  $\lambda(\mathbb{N}) \subseteq A$  which by 5.34 means that  $A$  is infinite.  $\square$

**Theorem 5.41.** *If  $n, m \in \mathbb{N}$  and  $n \approx m$  then  $n = m$*

**Proof.** Assume  $n \approx m$  then

1. If  $n < m \xrightarrow{n=S_n, m=S_m} n \subset m$  and thus  $m$  is bijective to a proper subset of itself, which means by the previous theorem that  $m$  is infinite contradicting the finiteness of  $m$  (as it is bijective with itself)
2. If  $m < n \xrightarrow{n=S_n, m=S_m} m \subset n$  and thus  $n$  is bijective to a proper subset of itself, which means by the previous theorem that  $n$  is infinite contradicting the finiteness of  $n$  (as it is bijective with itself)

as  $\mathbb{N}$  is totally ordered we must conclude that  $n = m$ .  $\square$

The previous theorem leads to the following observation: If  $A$  is a finite set then there exists a  $n \in \mathbb{N}$  such that  $n \approx A$ , if there was also a  $n' \in \mathbb{N}$  such that  $n' \approx A$  and thus  $n \approx n'$  and this leads to the following definition of the number of elements in a set.

**Definition 5.42.** *If  $A$  is a finite set then  $\exists! n \in \mathbb{N} \models n \approx A$  this number is called the number of elements in  $A$  and is noted as  $\#(A)$ .*

We have already proved that a finite union of a finite sets is finite (see 5.39) but we can go further in case of a family of pairwise disjoint sets and calculate the number of elements in the union.

**Theorem 5.43.** *Let  $\{A_i\}_{i \in \{1, \dots, n\}}$  be such that  $\forall i \in \{1, \dots, n\}$  we have that  $A_i$  is finite,  $\#(A_i) = n_i$  and  $\forall i, j \in \mathbb{N}_0$  with  $i \neq j$  we have  $A_i \cap A_j = \emptyset$  then  $\#(\bigcup_{i \in \{1, \dots, n\}} A_i) = \sum_{i=1}^n \#(A_i)$*

**Proof.** We prove this by induction on  $n$  so let  $\mathcal{S} = \{n \in \mathbb{N}_0 \mid \text{if } \{A_i\}_{i \in \{1, \dots, n\}} \text{ is a family of pairwise disjoint finite sets then } \#(\bigcup_{i \in \{1, \dots, n\}} A_i) = \sum_{i=1}^n \#(A_i)\}$  then we have

**1  $\in \mathcal{S}$ .** As  $\#(\bigcup_{i \in \{1, \dots, 1\}} A_i) = \#(A_1) = \sum_{i=1}^1 \#(A_i)$  we have  $1 \in \mathcal{S}$

**$n \in \mathcal{S} \Rightarrow n + 1 \in \mathcal{S}$ .** If  $\{A_i\}_{i \in \{1, \dots, n+1\}}$  is a pairwise disjoint family of sets then we have either

$$A_{n+1} = \emptyset. \text{ then } \#(A_{n+1}) = 0 \text{ and } \#(\bigcup_{i \in \{1, \dots, n+1\}} A_i) = \#(\bigcup_{i \in \{1, \dots, n\}} A_i) = \sum_{i=1}^n \#(A_i) \text{ proving } n+1 \in \mathcal{S}$$

**$A_{n+1} \neq \emptyset$ .** Define then  $N = \sum_{i=1}^n \#(A_i)$  then we have as  $n \in \mathcal{S}$  the existence of a bijection  $\sigma_1: \{1, \dots, N\} \rightarrow \bigcup_{i=1}^n A_i$  and as  $A_{n+1}$  is finite there exists a bijection  $\sigma_2: \{1, \dots, \#(A_{n+1})\} \rightarrow A_{n+1}$ . Define  $\sigma: \{1, \dots, N + \#(A_{n+1})\} \rightarrow \bigcup_{i \in \{1, \dots, n+1\}} A_i$  by  $i \rightarrow \sigma(i) = \begin{cases} \sigma_1(i) & \text{if } i \in \{1, \dots, N\} \\ \sigma_2(i - N) & \text{if } i \in \{N+1, \dots, N + \#(A_{n+1})\} \end{cases}$ , then we have

**$\sigma$  is injective.** Take  $i, j \in \{1, \dots, N + \#(A_{n+1})\}$  such that  $\sigma(i) = \sigma(j)$  then we have either

$$i, j \in \{1, \dots, N\}. \text{ then } \sigma_1(i) = \sigma(i) = \sigma(j) = \sigma_1(j) \text{ proving } i = j$$

$$i, j \in \{N+1, \dots, N + \#(A_{n+1})\}. \text{ then } \sigma_2(i - N) = \sigma(i) = \sigma(j) = \sigma_2(j - N) \text{ proving } i - N = j - N \Rightarrow i = j$$

$$i \in \{1, \dots, N\} \wedge j \in \{N+1, \dots, N + \#(A_{n+1})\}. \text{ then } \sigma(i) = \sigma_1(i) \in \bigcup_{i \in \{1, \dots, n\}} A_i \text{ and } \sigma(i) = \sigma(j) = \sigma_2(j - N) \in A_{n+1} \text{ proving that } \sigma(i) \in (\bigcup_{i \in \{1, \dots, n\}} A_i) \cap A_{n+1} = \emptyset \text{ a contradiction so this case does not occur.}$$

$j \in \{1, \dots, N\} \wedge i \in \{N+1, \dots, N+\#(A_{n+1})\}$ . then  $\sigma(j) = \sigma_j(j) \in \bigcup_{i \in \{1, \dots, n\}} A_i$  and  $\sigma(j) = \sigma(i) = \sigma_2(i - N) \in A_{n+1}$  proving that  $\sigma(j) \in (\bigcup_{i \in \{1, \dots, n\}} A_i) \cap A_{n+1} = \emptyset$  a contradiction so this case does not occur.

$\sigma$  is surjective. If  $y \in \bigcup_{i \in \{1, \dots, n+1\}} A_i$  then we have either

$y \in A_{n+1}$ . then as  $\sigma_2$  is a bijection there exists a  $i \in \{1, \dots, \#(A_{n+1})\}$  such that  $\sigma_2(j) = y$ . Take  $j = i + N \in \{N+1, \dots, N+\#(A_{n+1})\}$  then  $\sigma(j) = \sigma_2(j - N) = \sigma_2(i) = y$

$y \in \bigcup_{i \in \{1, \dots, n\}} A_i$ . then as  $\sigma_1$  is a bijection there exists a  $i \in \{1, \dots, \sum_{i=1}^N \#(A_i)\}$  such that  $\sigma_1(i) = y \Rightarrow \sigma(i) = \sigma_1(i) = y$

As  $\sigma$  is a bijection we have that  $\#(\bigcup_{i \in \{1, \dots, n+1\}} A_i) = N + \#(A_{n+1}) = \sum_{i=1}^N \#(A_i) + \#(A_{n+1}) = \sum_{i=1}^{N+1} \#(A_i)$  proving  $N+1 \in \mathcal{S}$   $\square$

**Theorem 5.44.** If  $A, B$  are finite sets with  $n = \#(A)$  and  $m = \#(B)$  then  $A \times B$  is finite and has size  $n \cdot m$

**Proof.** If  $A = \emptyset$  (or  $B = \emptyset$ ) then  $\#(A) = 0$  (or  $\#(B) = 0$ ) and then  $A \times B = \emptyset$  and  $\#(A \times B) = 0 = \#(A) \cdot \#(B)$ . So assume that  $A, B \neq \emptyset$  then there exists bijections  $a: B \rightarrow S_n$ ,  $b: B \rightarrow S_m$  where  $n, m \in \mathbb{N}_0$ . Define now  $f: A \times B \rightarrow S_{m \cdot n}$  by  $(x, y) \rightarrow a(x) \cdot m + b(y)$  (as  $\forall x \in A, \forall y \in B$  we have  $0 \leq a(x) < n$  and  $0 \leq b(y) < m$  we have indeed that  $\forall (x, y) \in A \times B$  that  $0 \leq f(x, y) < n \cdot m + m$ ). We prove now that  $f$  is a bijection:

1. **(injectivity)** If  $f(x, y) = f(x', y')$  then  $a(x) \cdot m + b(y) = a(x') \cdot m + b(y')$  and using the fact that  $b(y), b(y') < m$  and 4.74 we have that  $a(x) = a(x')$  and  $b(y) = b(y')$   $\xrightarrow[a, b \text{ are bijections}]{} x = x'$  and  $y = y' \Rightarrow (x, y) = (x', y')$  proving injectivity.
2. **(surjectivity)** If  $z \in S_{m \cdot n}$  then  $0 \leq z < m \cdot n$ , using the Division Algorithm (see 6.48) there exists a  $q, r$  so that  $z = q \cdot m + r$  and  $0 \leq r < m \Rightarrow r \in S_m$ , also  $0 \leq q < n$  [if  $n \leq q \Rightarrow n \cdot m \leq q \cdot m \Rightarrow n \cdot m + r \leq q \cdot m + r = z \Rightarrow n \cdot m + r \leq z \Rightarrow n \cdot m \leq z < n \cdot m$  a contradiction] and thus  $q \in S_n$ . Now as  $a, b$  are bijections there exists  $x \in A$  and  $y \in B$  such that  $a(x) = q$  and  $b(y) = r$  and thus  $f(x, y) = a(x) \cdot m + b(y) = q \cdot m + r = z$ .  $\square$

**Theorem 5.45.** If  $A$  is a finite set and  $B \subseteq A$  then  $B, A \setminus B$  are finite and  $\#(B) \leq \#(A)$  and  $\#(A \setminus B) = \#(A) - \#(B)$ . Note that from the last equation it follows that  $\#(A \setminus B) + \#(B) = \#(A)$  and thus that  $\#(A \setminus B) \leq \#(B)$

**Proof.** As  $A$  is finite with  $n = \#(A)$  then there exists a bijection  $a: n = S_n \rightarrow A$  we have now to consider the following cases:

1. **( $B = A$ )** then obviously  $B$  is finite and  $A \setminus B = \emptyset$  and thus  $A \setminus B$  is finite and  $\#(A \setminus B) = \#(A) - \#(B) = \#(A) - \#(A) = 0$
2. **( $B = \emptyset$ )** then obviously  $B$  is finite and  $A \setminus B = A$  and thus  $A \setminus B$  is finite and  $\#(A \setminus B) = \#(A) = \#(A) - 0 = \#(A) - \#(B)$
3. **( $\emptyset \neq B \subset A$ )** Let now  $X = \{n \in \{2, \dots, \}\}$  If  $A$  is such that  $\#(A) = n$  and  $\emptyset \neq B \subset A$  then there exists a bijection  $f: S_n \rightarrow A$  and  $a, m \in S_n$  such that  $f|_{S_m}: S_m \rightarrow B$  is a bijection} then we have
  - a. If  $n = 2$  then if  $\#(A) = 2$  then  $2 = \{0, 1\} = S_2$  and there exists a bijection  $f': \{0, 1\} \rightarrow A$ , now as  $\emptyset \neq B \subset A$  there exists a  $a \in A \setminus B$  and as  $f'$  is a bijection then there exists a  $i \in \{0, 1\}$  such that  $f'(i) = a$  we have now two cases:
    - i. **( $i = 0$ )** then  $B = \{f'(1)\}$  so take the bijection  $f: \{0, 1\} \rightarrow A$  with  $\begin{cases} f(0) = f'(1) \\ f(1) = f'(0) \end{cases}$  then  $f|_{S_1}: \{0\} \rightarrow \{f'(1)\} = B$  is a bijection
    - ii. **( $i = 1$ )** then  $B = \{f'(0)\}$  and take then  $f = f'$  so that  $f|_{S_1} = f|_{S_1}: \{0\} \rightarrow \{f'(0)\} = B$  is a bijection

we can thus conclude that  $2 \in X$

- b. Assume now that  $n \in X$ . If now  $\#(A) = n+1$  then there exists a bijection  $f': S_{n+1} \rightarrow A$ , now as  $\emptyset \neq B \subset A$  there exists a  $a \in A \setminus B$  and thus  $\emptyset \neq B \subseteq A \setminus \{a\}$ . As  $f'$  is a bijection there exists a  $i \in S_{n+1} = \{0, \dots, n\}$  such that  $f'(i) = a$ . We have now two cases to consider:

- i. ( $i = n$ ) take then  $f = f'$
- ii. ( $i \neq n$ ) take then  $a' = f'(n)$  then  $a \neq a'$  [otherwise  $i = n$ ]. Then  $f'_{|S_{n+1} \setminus \{i, n\}}: S_{n+1} \setminus \{i, n+1\} \rightarrow A \setminus \{a, a'\}$  is a bijection (see 2.55). Define further the bijection  $f'': \{i, n\} \rightarrow \{a, a'\}$  by  $\begin{cases} f''(i) = a' \\ f''(n) = a \end{cases}$  so that we have a bijection (see 2.46)  $f = f'_{|S_{n+1} \setminus \{i, n\}} \cup f'': S_{n+1} \rightarrow A$  with  $(n) = f''(n) = a$ .

So we have found a bijection  $f: S_{n+1} \rightarrow A$  such that  $f(n) = a$ . We have now the following cases to consider for  $A \setminus \{a\}$  (remember  $B \subseteq A \setminus \{a\}$ ):

- i. ( $\emptyset \neq B = A \setminus \{\alpha\}$ ) then we take  $m = n$  and  $f_{|S_n}: S_n \rightarrow A \setminus \{a\} = B$  is a bijection (as  $f(\{n\}) = \{a\}$  and thus  $n+1 \in X$ )
- ii. ( $\emptyset \neq B \subset A \setminus \{a\}$ ) then as  $n = \#(A \setminus \{a\})$  and  $n \in X$  there exists a bijection  $f''': S_n \rightarrow A \setminus \{a\}$  and a  $m \in S_n$  such that  $f'''_{|S_m}: S_m \rightarrow B$  is a bijection, take then the bijection  $f^{\text{iv}}: \{n\} \rightarrow \{a\}$  by  $f^{\text{iv}}(n) = a$  and construct then the bijection  $f^v = f''' \cup f^{\text{iv}}: S_{n+1} \rightarrow A$ . We have then that  $f^v_{|S_m} = f'''_{|S_m}$  is a bijection from  $S_m$  to  $B$  proving that  $n+1 \in X$

Using mathematical induction (see 4.77) we have that  $X = \{2, \dots\}$ , now as  $\emptyset \neq B \subset A$  we must have  $2 \leq \#(A) = n \Rightarrow n = \#(A) \in \{2, \dots\} = X$  so there exists a bijection  $f: S_n \rightarrow A$  and a  $m \in S_n$  such that  $f_{|S_m}: S_m \rightarrow B$  is a bijection, this proves that  $B$  is finite and  $\#(B) = m$ . Now if  $i \in \{m, \dots, n-1\}$  we have  $f(i) \notin B$  [if  $f(i) \in B$  then as  $f_{|S_m}$  is a bijection there exists a  $j < m$  such that  $f(j) = f_{|S_m}(j) = f(i) \Rightarrow i = j < m$  contradicting  $i \in \{m, \dots, n-1\}$ ] and thus  $f(\{m, \dots, n-1\}) \subseteq A \setminus B$ . If  $y \in A \setminus B$  then as  $f$  is a bijection there exists a  $i \in S_n$  such that  $y = f(i)$  if now  $i < m \Rightarrow i \in S_m \Rightarrow f(i) \in B$  a contradiction so we must have  $i \in \{m, \dots, n-1\} \Rightarrow A \setminus B \subseteq f(\{m, \dots, n-1\})$ . So  $f(\{m, \dots, n-1\}) = A \setminus B$  and thus by 2.55 we have that  $f_{|\{m, \dots, n-1\}}: \{m, \dots, n-1\} \rightarrow A \setminus B$  is a bijection. Define now  $g: S_{n-m} \rightarrow \{m, \dots, n-1\}$  (note that from  $m \in S_n \Rightarrow m < n \Rightarrow 0 < n-m \in \mathbb{N}$ ) by  $i \rightarrow g(i) = i+m$  then  $g$  is a bijection:

- a. (**injectivity**) if  $g(i) = g(j) \Rightarrow i+m = j+m \Rightarrow i = j$
- b. (**surjectivity**) If  $j \in \{m, \dots, n-1\} \Rightarrow m \leq j < n-1 \Rightarrow 0 \leq j-m < n-m-1 \Rightarrow j-m \in S_{n-m}$  and  $g(j-m) = j$

So we have then a bijection  $f_{|\{m, \dots, n-1\}} \circ g: S_{n-m} \rightarrow A \setminus B$  proving that  $A \setminus B$  is finite and  $\#(A \setminus B) = n-m = \#(A) - \#(B)$   $\square$

**Theorem 5.46.** *If  $A$  is a finite set then if  $B \subset A$  we have that  $\#(B) < \#(A)$*

**Proof.** Using the previous theorem we have that  $B$  is finite and  $\#(B) \leq \#(A)$ , if now  $\#(A) = \#(B) = n$  then  $A \approx n \approx B$  and thus  $A \approx B$  which by 5.40 would give that  $A$  is infinite contradicting the finiteness of  $A$ . So we must have  $\#(B) < \#(A)$   $\square$

**Theorem 5.47.** *If  $A$  is a set,  $n \in \mathbb{N}$  and  $f: S_n \rightarrow A$  a surjection then  $A$  is finite and  $\#(A) \leq n$ .*

**Proof.** If  $n = 0$  then  $S_0 = \emptyset$  and if  $f: \emptyset \rightarrow A$  is a surjection then if  $x \in A$  there exists a  $i \in \emptyset$  such that  $x = f(i)$  which would mean that  $\emptyset \neq \emptyset$  a contradiction, so in this cases we must have  $A = \emptyset \Rightarrow A$  is finite and  $\#(A) = 0 \leq 0$  proving the theorem for the case  $n = 0$ . We prove the case  $n \in \mathbb{N}_0$  by mathematical induction so let  $X = \{n \in \{1, \dots\} \mid$  if  $A$  is a set such that there exists a surjection  $f: S_n \rightarrow A$  then  $A$  is finite and  $\#(A) = n\}$  we have then that

1. If  $n = 1$  then if  $A$  is such that there exists a surjection  $f: S_1 = \{0\} \rightarrow A$  then  $A = \{f(0)\}$  which trivially means that  $A$  is finite and that  $\#(A) = 1 \leq 1 \Rightarrow 1 \in X$

2. If now  $n \in X$  then if  $A$  is such that there exists a surjection  $f: S_{n+1} \rightarrow A$ . Then as  $f(n) \in A$  we can not have  $A = \emptyset$  and we are left with the following two cases
- $(\forall a \in A \models a = f(n)) \Rightarrow A = \{f(n)\}$  and thus  $A$  is finite with  $\#(A) = 1 \leq n+1$ . This gives  $n+1 \in X$
  - $(\exists a \in A \models a \neq f(n))$  Define then  $g: S_n \rightarrow A \setminus \{f(n)\}$  by

$$\begin{aligned} g(i) &= f(i) & \text{if } i \in S_n \setminus f^{-1}(\{f(n)\}) & [\Rightarrow g(i) \in A \setminus \{f(n)\}] \\ &= a & \text{if } i \in S_n \cap f^{-1}(\{f(n)\}) & [\Rightarrow g(i) \in A \setminus \{f(n)\}] \end{aligned}$$

If now  $y \in A \setminus \{f(n)\}$  then as  $f$  is a surjection there exists a  $i \in S_{n+1}$  such that  $f(i) = y$ . We can not have  $i = n$  [As then  $f(i) = f(n) \Rightarrow y \in \{f(n)\}$ ], also we can not have  $i \in f^{-1}(\{f(n)\})$  [as then  $f(i) \in \{f(n)\} \Rightarrow y \in \{f(n)\}$  so  $i \in S_n \setminus f^{-1}(\{f(n)\})$  and thus  $g(i) = f(i) = y$ ]. This proves that  $g$  is a surjection. As  $n \in X$  we have that  $A \setminus \{f(n)\}$  is finite and  $m = \#(A \setminus \{f(n)\}) \leq n$ . So there exists a bijection  $h: S_m \rightarrow A \setminus \{f(n)\}$ , combine this with the bijection  $h': \{m\} \rightarrow \{f(n)\}$  using 2.46 to form the bijection  $h \cup h': S_{m+1} \rightarrow A$ . So  $A$  is finite and  $\#(A) = m+1 \leq n+1$  and this proves that  $n+1 \in X$

Using mathematical induction (4.77) we have that  $X = \{1, \dots\} = \mathbb{N}_0$  proving the theorem for  $n \in \mathbb{N}_0$   $\square$

**Corollary 5.48.** *If  $A, B$  are sets where  $A$  is finite and  $f: A \rightarrow B$  is a surjection then  $B$  is finite and  $\#(B) \leq \#(A)$*

**Proof.** If  $A$  is finite and  $\#(A) = n$  then there exists a bijection  $b: n \rightarrow A$  so  $f \circ g: S_n \rightarrow B$  is a surjection and by the previous theorem we have then  $B$  is finite and  $\#(B) \leq n = \#(A)$   $\square$

**Theorem 5.49.** *Let  $A, B$  be sets,  $A$  infinite and  $f: A \rightarrow B$  a injection then  $B$  is infinite*

**Proof.** Assume that  $B$  is finite then  $f(A) \subseteq B$  is finite and there is a bijection  $f: \{1, \dots, n\} \rightarrow f(A)$  so that as  $f: A \rightarrow f(A)$  is a bijection we find that  $f^{-1} \circ h: \{1, \dots, n\} \rightarrow A$  is a bijection making  $A$  finite a contradiction.  $\square$

**Theorem 5.50.** *Let  $\langle A, \leq \rangle$  be a fully ordered non empty finite set then  $\max(A)$  (or  $\min(A)$ ) exists (see 2.167)*

**Proof.** We prove this by induction on  $\#(A) = n$ . So let  $S = \{n \in \mathbb{N}_0 = \{1, \dots\} \mid \text{if } \#(A) = n \text{ then } \max(A) \text{ exists}\}$  then we have :

1. If  $n = 1$  then  $A = \{a\}$  with  $\max(A)$
2. If  $n \in S$  let then  $\#(A) = n+1$  so there exists a bijection  $b: \{1, \dots, n+1\} \rightarrow A$ , take then  $A \setminus \{b(n+1)\}$  which is bijective to  $\{1, \dots, n\}$  so that  $\#(A \setminus \{b(n+1)\}) = n$  and thus  $m' = \max(A \setminus \{b(n+1)\})$  exist. We have now two cases to consider:
  - a.  $(m' \leq b(n+1))$  then  $\max(A) = b(n+1)$
  - b.  $(b(n+1) \leq m')$  then  $\max(A) = m'$

The proof for the minimum is similar.  $\square$

**Theorem 5.51.** *If  $A$  is finite and  $f: \mathbb{N}_0 \rightarrow A$  is a function then  $\exists a \in A$  such that  $\{m \in \mathbb{N}_0 \mid f(x) = a\} = f^{-1}(\{a\})$  is infinite.*

**Proof.** Assume that the theorem is not valid then  $\forall a \in A$  we have that  $f^{-1}(\{a\})$  is finite so that  $\bigcup_{a \in A} f^{-1}(\{a\})$ , being a finite union of finite sets, is finite. Now from  $\mathbb{N}_0 = f^{-1}(A) = f^{-1}(\bigcup_{a \in A} \{a\}) = \bigcup_{a \in A} f^{-1}(\{a\})$  we would then conclude that  $\mathbb{N}_0$  is finite which is a contradiction.  $\square$

**Corollary 5.52.** *If  $A$  is finite and  $f: \mathbb{N}_0 \rightarrow A$  is a function then  $\exists a \in A$  so that  $\forall n \in \mathbb{N}_0$  there exists a  $m \in \{i \in \mathbb{N}_0 \mid i \geq n\}$  so that  $f(m) = a$*

**Proof.** By the previous theorem  $\exists a \in A$  such that  $f^{-1}(\{a\}) = \{m \in \mathbb{N}_0 | f(m) = a\}$  is infinite. We proceed non by contradiction. So assume that  $\exists n \in \mathbb{N}_0$  such that  $\forall m \geq n$  we have  $f(m) \neq a$ . If then  $m \in f^{-1}(\{m\}) \Rightarrow f(m) = a \Rightarrow m < n \Rightarrow f^{-1}(\{a\}) \subseteq S_n$  meaning that  $f^{-1}(\{a\})$  is finite contradicting the fact that  $f^{-1}(\{a\})$  is infinite.  $\square$

**Theorem 5.53.** Let  $\langle X, \leq \rangle$  be a fully ordered set and let  $A \subseteq X$  be a non empty finite set. Then there exists a bijection  $i: \{1, \dots, \#(A)\} \rightarrow A$  such that  $\forall k \in \{1, \dots, \#(A) - 1\}$  we have  $i(k) < i(k + 1)$ . Second for such a bijection we must have  $\forall k, l \in \{1, \dots, \#(A)\}$  with  $k < l$  that  $i(k) < i(l)$ . Third the above bijection is unique.

**Proof.** If  $\#(A) = 1$  we take the bijection  $i = 1_{\{1\}}$  then for  $k \in \{1, \dots, 0\} = \emptyset$  we have that  $i(l) < i(k)$  is satisfied vacuously, also  $\forall k, l \in \{1, \dots, 1\}$  with  $k < l$  we have that  $i(k) < i(l)$  is satisfied vacuously. So we only have to prove the theorem for  $\#(A) > 1$ . We prove the rest by induction so let  $\mathcal{S} = \{n \in \{2, \dots\} | \text{if } \#(A) = n \text{ then there exists a bijection } i: \{1, \dots, n\} \rightarrow A \text{ with } \forall k \in \{1, \dots, n-1\} \text{ we have } i(k) < i(k+1)\}$ . We have then:

1. If  $n = 2$  then there exists a bijection  $b: \{1, \dots, 2\} = \{1, 2\} \rightarrow A$  so  $A = \{b(1), b(2)\}$ . If  $b(1) < b(2)$  we have found our bijection. If  $b(2) < b(1)$  construct then  $i: \{1, 2\} \rightarrow A$  by  $i(1) = b(2)$  and  $i(2) = b(1)$  which is obviously a bijection with  $i(1) < i(2)$ . This proves that  $2 \in \mathcal{S}$ .
2. Assume that  $n \in \mathcal{S}$  and that we have a finite  $A$  with  $\#(A) = n + 1$ . As  $A$  is finite a maximum exists (see 5.50) take then  $A \setminus \{\max(A)\}$ . Using 5.45 we have that  $\#(A \setminus \{\max(A)\}) = \#(A) - \#(\{\max(A)\}) = (n + 1) - 1 = n$ . As  $n \in \mathcal{S}$  there exists a bijection  $i': \{1, \dots, n\} \rightarrow A \setminus \{\max(A)\}$  such that  $\forall k \in \{1, \dots, n-1\}$  we have  $i'(k) < i'(k+1)$ . Define then the bijection  $i: \{1, \dots, n+1\} \rightarrow A$  by  $k \rightarrow i(k) = \begin{cases} i'(k) & \text{if } k \in \{1, \dots, n\} \\ \max(A) & \text{if } k = n+1 \end{cases}$  (see 2.46) then if  $k \in \{1, \dots, (n+1)-1\} = \{1, \dots, n\}$  we have either  $k \in \{1, \dots, n-1\}$  and then  $i(k) = i'(k) < i'(k+1) = i(k+1)$ , if  $k = n$  then  $i(k) \leq \max(A) = i(k+1)$ . As  $i(k) = i'(k) \in A \setminus \{\max(A)\}$  we must have  $i(k) \neq \max(A) = i(k+1)$  so that we have  $i(k) < i(k+1)$ . This proves that  $n+1 \in \mathcal{S}$

Using induction we have  $\mathcal{S} = \{2, \dots\}$  proving the existence. Next we prove by induction that if  $i: \{1, \dots, \#(A)\} \rightarrow A$  is a bijection such that  $\forall k \in \{1, \dots, \#(A) - 1\}$  we have  $i(k) < i(k + 1)$  then we have  $\forall k, l \in \{1, \dots, \#(A)\}$  that  $i(k) < i(l)$ . So let  $k \in \{1, \dots, \#(A)\}$  and take  $\mathcal{P}_k = \{m \in \mathbb{N}_0 | \text{if } k + m \leq \#(A) \text{ then } i(k) < i(k + m)\}$  then we have

1. If  $m = 1$  then if  $k + 1 < \#(A)$  we have  $k \leq \#(A) - 1$  so that  $i(k) < i(k + 1) = i(k + m)$  proving that  $1 \in \mathcal{P}_k$
2. Let  $m \in \mathcal{P}_k$  then if  $k + (m + 1) \leq \#(A)$  we have  $(k + m) \leq \#(A) - 1$  so that  $i(k + m) < i((k + m) + 1) = i(k + (m + 1))$  and as  $m \in \mathcal{P}_k$  and  $k + m \leq \#(A)$  we have  $i(k) < i(k + m)$  so that  $i(k) < i(k + (m + 1))$  proving that  $m + 1 \in \mathcal{P}_k$

Using induction we have that  $\mathcal{P}_k = \mathbb{N}_0$  so if  $k, l \in \{1, \dots, \#(A)\}$  and  $k < l$  then  $m = l - k \in \mathbb{N}_0 = \mathcal{P}_k$  and as  $k + m = l \leq \#(A)$  we have that  $i(k) < i(k + m) = i(l)$ .

Finally to prove uniqueness suppose that  $n = \#(A)$  and  $i_1: \{1, \dots, n\} \rightarrow A$ ,  $i_2: \{1, \dots, n\} \rightarrow A$  are bijections such that if  $k \in \{1, \dots, n-1\}$  then  $i_1(k) < i_1(k) \wedge i_2(k) < i_2(k+1)$ . Take now  $\mathcal{Q} = \{k \in \mathbb{N}_0 | \text{if } k \leq n \text{ then } \forall l \in \{1, \dots, k\} \text{ we have } i_1(l) = i_2(l)\}$  then we have:

1. If  $k = 1$  then if  $i_1(1) \neq i_2(1)$  we have either
  - $i_1(1) < i_2(1)$  then as  $i_1(1) \in A$  there exists a  $l \in \{1, \dots, n\}$  such that  $i_2(l) = i_1(1)$ . If now  $l > 1$  we must have  $i_2(1) < i_2(l) = i_1(1) < i_2(1)$  a contradiction, so we must have  $l = 1$  but this means that  $i_2(1) = i_1(1)$  contradicting  $i_1(1) < i_2(1)$ . So this case leads to a contradiction.
  - $i_2(1) < i_1(1)$  then as  $i_2(1) \in A$  there exists a  $l \in \{1, \dots, n\}$  such that  $i_1(l) = i_2(1)$ . If now  $l > 1$  we must have  $i_1(1) < i_1(l) = i_2(1) < i_1(1)$  a contradiction, so we must have  $l = 1$  but this means that  $i_1(1) = i_2(1)$  contradicting  $i_2(1) < i_1(1)$ . So this case leads to a contradiction.

we must thus conclude that  $i_1(1) = i_2(1)$  proving that  $1 \in \mathcal{Q}$

2. Let  $k \in \mathcal{Q}$  then if  $k+1 \leq n$  we prove by contradiction that  $i_1(k+1) = i_2(k+1)$ . So assume that  $i_1(k+1) \neq i_2(k+1)$  then we have either

- a.  $i_1(k+1) < i_2(k+1)$  then as  $i_1(k+1) \in A$  there exists a  $l \in \{1, \dots, n\}$  such that  $i_2(l) = i_1(k+1)$ . If now  $l = k+1$  we would have the contradiction  $i_1(k+1) = i_2(l) = i_2(k+1) > i_1(k+1)$ , if  $l > k+1$  then  $i_1(k+1) = i_2(l) > i_2(k+1) > i_1(k+1)$  again a contradiction, if  $l < k+1$  then  $l \leq k$  so that  $i_1(l) = i_2(l) = i_1(k+1) > i_1(l)$  again a contradiction. So all the cases leads to a contradiction.
- b.  $i_2(k+1) < i_1(k+1)$  then as  $i_2(k+1) \in A$  there exists a  $l \in \{1, \dots, n\}$  such that  $i_1(l) = i_2(k+1)$ . If now  $l = k+1$  we would have the contradiction  $i_2(k+1) = i_1(l) = i_1(k+1) > i_2(k+1)$ , if  $l > k+1$  then  $i_2(k+1) = i_1(l) > i_1(k+1) > i_2(k+1)$  again a contradiction, if  $l < k+1$  then  $l \leq k$  so that  $i_2(l) = i_1(l) = i_2(k+1) > i_2(l)$  again a contradiction. So all the cases leads to a contradiction.

so we must have  $i_1(k+1) = i_2(k+1)$  or  $k+1 \in \mathcal{Q}$

By induction we have thus that  $\mathcal{Q} = \mathbb{N}_0$ . So if  $k \in \{1, \dots, n\}$  then  $k \in \mathcal{Q}$  and  $k \leq n$  so that  $i_1(k) = i_2(k)$  proving that  $i_1 = i_2$ .  $\square$

### 5.3 Properties of denumerable sets

**Lemma 5.54.** *Every subset of  $\mathbb{N}$  is finite or denumerable.*

**Proof.** As  $\langle \mathbb{N}, \leq \rangle$  is well-ordered (see 4.52) we can use 2.205. So if  $A \subseteq \mathbb{N}$  we have either that  $A$  is isomorphic with  $\mathbb{N}$  (and thus denumerable as a isomorphism is a bijection) or  $A$  is isomorphic with a segment  $S_n = n$  and thus finite.  $\square$

**Theorem 5.55.** *Every subset of a denumerable set is finite or denumerable.*

**Proof.** Let  $A$  be a denumerable set and take  $B \subseteq A$ . From the denumerability of  $A$  there exists a bijective function  $f: \mathbb{N} \rightarrow A$ . Then using the previous theorem we have that  $f^{-1}(B) \subseteq \mathbb{N}$  is either denumerable or finite, so we have the following cases:

- 1. ( $f^{-1}(B)$  is enumerable) Then there exists a bijection  $g: \mathbb{N} \rightarrow f^{-1}(B)$ , together with the bijection  $f|_{f^{-1}(B)}: f^{-1}(B) \rightarrow B$  this means that  $B$  is bijective with  $\mathbb{N}$  and is thus denumerable.
- 2. ( $f^{-1}(B)$  is finite) Then there exists a  $n \in \mathbb{N}$  such that there exists a bijection  $g: n \rightarrow f^{-1}(B)$ , together with the bijection  $f|_{f^{-1}(B)}: f^{-1}(B) \rightarrow B$  this means that  $B$  is bijective with  $n$  and is thus finite.

$\square$

**Theorem 5.56.** *If  $0 < m$  then there exists a number noted by  $m-1$  such that*

1.  $(m-1)+1 = m$
2.  $m-1 = m'-1 \Leftrightarrow m = m'$
3.  $(m+1)-1 = m$

**Note 5.57.** As  $0 < m \Rightarrow 1 \leq m$  we have that this definition is equivalent with the definition 4.63.

**Proof.**

1. If  $0 < m$  then  $1 \leq m$  and thus by 4.62 there exists a  $k \in \mathbb{N}$  such that  $1+k = k+1 = m$  so we can take  $m-1$  to be  $k$ .
2. If  $m-1 = m'-1 \Rightarrow m = (m-1)+1 = (m'-1)+1 = m' \Rightarrow m = m'$ . If  $m = m'$  then  $(m-1)+1 = m = m' = (m'-1)+1 \Rightarrow s(m-1) = s(m'-1) \xrightarrow{4.15} m-1 = m'-1$
3. We have  $((m+1)-1)+1 = m+1 \Rightarrow s((m+1)-1) = s(m) \xrightarrow{4.15} (m+1)-1 = m$   $\square$

**Theorem 5.58.**  $\mathbb{N} \times \mathbb{N} \approx \mathbb{N}$

**Proof.** First define the function  $f: \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N} \times \mathbb{N}$  (using the previous theorem to define  $m-1$ )

$$f(k, m) = \begin{cases} (0, k+1) & \text{if } m = 0 \\ (k+1, m-1) & \text{if } m \neq 0 (\Rightarrow 0 < m) \end{cases}$$

then if  $f(k, m) = f(k', m')$  we have the following cases for  $m, m'$

1.  $(m, m' = 0)$  then  $(0, k+1) = f(k, m) = f(k', m') = (0, k'+1) \Rightarrow k+1 = k'+1 \Rightarrow k = k' \xrightarrow[m=0=m']{} (k, m) = (k', m')$
2.  $(m > 0, m' = 0)$  then  $(k+1, m-1) = f(k, m) = f(k', m') = (0, k'+1) \Rightarrow k+1 = 0 \xrightarrow[0 \leq k \Rightarrow 0 < k+1]{} 0 < 0$  a contradiction, this case can not occur
3.  $(m = 0, m' > 0)$  then  $(0, k+1) = f(k, m) = f(k', m') = (k'+1, m'-1) \Rightarrow 0 = k'+1 \xrightarrow[0 \leq k' \Rightarrow 0 < k'+1]{} 0 > 0$  a contradiction, this case can not occur
4.  $(m > 0, m' > 0)$  then  $(k+1, m-1) = f(k, m) = f(k', m') = (k'+1, m'-1) \Rightarrow k+1 = k'+1 \wedge m-1 = m'-1 \Rightarrow k = k' \wedge m = m' \Rightarrow (k, m) = (k', m')$

this proves that  $f: \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N} \times \mathbb{N}$  is a injective function. If now  $f(k, m) = (0, 0)$  then we have the following cases to consider for  $m$

1.  $(m = 0)$  then  $(0, 0) = f(k, m) = (0, k+1) \Rightarrow 0 = k+1 \xrightarrow[0 \leq k \Rightarrow 0 < k+1]{} 0 < 0$  a contradiction.
2.  $(m \neq 0)$  then  $(0, 0) = f(k, m) = (k+1, m-1) \Rightarrow 0 = k+1 \xrightarrow[0 \leq k \Rightarrow 0 < k+1]{} 0 < 0$  a contradiction.

so we have  $f(k, m) \neq (0, 0)$  or  $(0, 0) \notin f(\mathbb{N} \times \mathbb{N})$ .

Using 4.19 there exists a injective function  $\lambda: \mathbb{N} \rightarrow \mathbb{N} \times \mathbb{N}$  such that

1.  $\lambda(0) = (0, 0)$
2.  $\forall n \in \mathbb{N} \models \lambda(n+1) = f(\lambda(n))$

We prove now the following **property (a)** of  $\lambda$ , assume that there exists a  $n, m \in \mathbb{N}$  such that  $\lambda(n) = (0, m)$  then  $\forall k \in \mathbb{N}$  with  $\exists l \in \mathbb{N}$  such that  $k+l = m$  we have  $\lambda(n+k) = (k, l)$ . The proof is by induction on  $k$  so let  $S_{n,m} = \{k \in \mathbb{N} \mid \exists l \in \mathbb{N} \text{ such that } k+l = m \text{ then } \lambda(n+k) = (k, l)\}$  then:

1. If  $k=0$  then if  $l=m$  we have  $k+l=m$  and  $\lambda(n+k) = \lambda(n) = (0, m) = (k, l)$  so we have  $0 \in S_{n,m}$
2. Assume  $k \in S_{n,m}$  then consider  $k+1$  if there exists a  $l \in \mathbb{N}$  such that  $(k+1)+l = m$  then  $k+(l+1) = m$  and as  $k \in S_{n,m}$  we have  $\lambda(n+k) = (k, l+1)$ , now as  $0 \leq l \Rightarrow 0 < l+1$  we have  $\lambda(n+(k+1)) = \lambda((n+k)+1) = f(\lambda(n+k)) = f(k, l+1) = (k+1, l)$  so  $k+1 \in S_{n,m}$

Thus if  $k+l = m \Rightarrow k \in \mathbb{N} = S_{n,m} \Rightarrow \lambda(n+k) = (k, l)$

We shown now by induction that  $\lambda$  is surjective, so let  $S = \{n \in \mathbb{N} \mid \forall (k, m) \in \mathbb{N} \times \mathbb{N} \vdash k+m = n \models \exists n' \in \mathbb{N} \vdash \lambda(n') = (k, m)\}$ .

1. If  $n=0$  take  $k=m \in \mathbb{N}$  is such that  $k+m=0$  then we must have  $k=0=m$  [if  $m > 0$  then  $0=k+m > 0 \Rightarrow 0 \neq 0$  and if  $k > 0$  again  $0=k+m > 0 \Rightarrow 0 \neq 0$ ] so that  $\lambda(0) = (0, 0) = (k, m)$  and thus  $0 \in S$ .
2. Assume now that  $n \in S$  and consider  $n+1$ . If  $(k, m) \in \mathbb{N} \times \mathbb{N}$  is such that  $k+m = n+1$ . we can have then the following cases for  $k$ 
  - a.  $(k = 0)$  then  $k = k+m = n+1$  and thus  $(k, m) = (0, m) = (0, n+1) = f(n, 0)$ . As  $n+0 = n \in S$  there exists a  $n'' \in \mathbb{N}$  such that  $(n, 0) = \lambda(n'')$  and thus  $\lambda(n''+1) = f(\lambda(n'')) = f(n, 0) = (k, m)$  so we have found a  $n' = n''+1$  such that  $\lambda(n') = (k, m)$
  - b.  $(k \neq 0)$  then  $0 < k \Rightarrow k-1$  exists and as  $0 \leq m \Rightarrow 0 < m+1$  we have  $f(k-1, m+1) = ((k-1)+1, (m+1)-1) = (k, m)$ . Now  $((k+m)-1)+0 = (n+1)-1 = n \in S$  so that  $\exists n'' \in \mathbb{N}$  such that  $\lambda(n'') = ((k+m)-1, 0)$ . So  $\lambda(n''+1) = f((k+m)-1, 0) = (0, k+m)$ . Using **property a** of  $\lambda$  we have then that  $\lambda((n''+1)+k) = (k, m)$  so we have found a  $n' = (n''+1)+k$  such that  $\lambda(n') = (k, m)$

If thus  $(k, m) \in \mathbb{N} \times \mathbb{N}$  then  $n = k+m \in \mathbb{N} = S \Rightarrow \exists n' \in \mathbb{N} \vdash \lambda(n') = (k, m)$  proving that  $\lambda: \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$  is surjective and as we have by construction of  $\lambda$  it is injective we have that  $\lambda: \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$  is a bijection and thus  $\mathbb{N} \approx \mathbb{N} \times \mathbb{N}$   $\square$

**Corollary 5.59.** *If  $A, B$  are denumerable then  $A \times B$  is denumerable*

**Proof.** If  $A, B$  are denumerable then  $\mathbb{N} \approx A$  and  $\mathbb{N} \approx B$  we have by 5.12 that  $A \times B \approx \mathbb{N} \times \mathbb{N} \approx \mathbb{N}$  so  $A \times B \approx \mathbb{N}$   $\square$

**Theorem 5.60.** *If  $\{A_i\}_{i \in B}$  is a denumerable family of denumerable sets (so  $B$  is denumerable and  $\forall i \in B$  we have that  $A_i$  is a denumerable set) then  $\bigcup_{i \in B} A_i$  is denumerable.*

**Proof.** As  $\{A_i\}_{i \in B}$  is a denumerable family of denumerable sets there exists a graph  $A$  with  $\text{dom}(G) = B$  where  $\mathbb{N} \approx B$ , so there is a bijection  $f: \mathbb{N} \rightarrow B$ , and  $\forall i \in B$  we have  $A_i \approx \mathbb{N} \Rightarrow \forall i \in B$  there exists a bijection  $f_i: \mathbb{N} \rightarrow A_i$ . So define  $\forall i \in B$  the non empty set (see 2.76)  $\mathcal{A}_i = \{f \in A_i^{\mathbb{N}}: f \text{ is a bijection}\}$ , then  $\mathcal{A} = \bigcup_{i \in B} \mathcal{A}_i$  is a set and  $\forall i \in B$  we have  $\mathcal{A}_i \subseteq \mathcal{A}$ . Using the axiom of choice (see 2.207) we have then the existence of a choice function  $c: \mathcal{P}'(\mathcal{A}) \rightarrow \mathcal{A}$  with  $c(\mathcal{A}_i) \in \mathcal{A}_i \Rightarrow c(\mathcal{A}_i)$  is a bijection from  $\mathbb{N} \rightarrow A_i$ . Define now  $F: \mathbb{N} \times \mathbb{N} \rightarrow \bigcup_{i \in B} A_i$  by  $F(n, m) = c(\mathcal{A}_{f(n)})(m)$  then  $F$  is surjective. For if  $y \in \bigcup_{i \in B} A_i$  then there exists a  $i \in B$  such that  $y \in A_i$  and as  $f$  is bijective there exists a  $n \in \mathbb{N}$  such that  $f(n) = i$ , also as  $c(\mathcal{A}_{f(n)}): \mathbb{N} \rightarrow A_{f(n)} = A_i$  there exists a  $m \in \mathbb{N}$  such that  $c(\mathcal{A}_{f(n)})(m) = y$  and thus  $F(n, m) = c(\mathcal{A}_{f(n)})(m)$ . As  $\mathbb{N} \times \mathbb{N} \approx \mathbb{N}$  there exists a bijection  $\varphi: \mathbb{N} \rightarrow \mathbb{N} \times \mathbb{N}$  and thus  $F \circ \varphi: \mathbb{N} \rightarrow \bigcup_{i \in B} A_i$  is a surjective function. Using 5.9 we have then that  $\bigcup_{i \in B} A_i \approx E \subseteq \mathbb{N}$ . Using 5.55 we have that  $E$  is finite or denumerable hence  $\bigcup_{i \in B} A_i$  is either finite or denumerable. As  $A_{f(1)} \subseteq \bigcup_{i \in B} A_i$  and  $A_{f(1)}$  is denumerable  $\xrightarrow{5.34} A_{f(1)}$  is infinite  $\xrightarrow{5.35} \bigcup_{i \in B} A_i$  is not finite so  $\bigcup_{i \in B} A_i$  is a denumerable set.  $\square$

**Corollary 5.61.** *The union of two denumerable sets is denumerable.*

**Proof.** If  $A, B$  are denumerable sets then there exists bijections  $f: \mathbb{N} \rightarrow A$  and  $g: \mathbb{N} \rightarrow B$ . Define now  $F: \{1, 2\} \times \mathbb{N} \rightarrow A \bigcup B$  by  $F(1, n) = f(n)$  and  $F(2, n) = g(n)$  then  $F$  is a surjection as if  $x \in A \bigcup B$  we have either  $x \in A \xrightarrow{f \text{ is bijective}} \exists n \in \mathbb{N} \vdash x = f(n) = F(1, n)$  or  $x \in B \xrightarrow{g \text{ is bijective}} \exists n \in \mathbb{N} \vdash x = g(n) = F(2, n)$ . Now  $\{1, 2\} \times \mathbb{N} \subseteq \mathbb{N} \times \mathbb{N}$  then by 5.55 and the fact that  $\mathbb{N} \approx \mathbb{N} \times \mathbb{N}$  we have that  $\{1, 2\} \times \mathbb{N}$  is either finite or denumerable, as  $\mathbb{N} \approx \{1\} \times \mathbb{N} \subseteq \{1, 2\} \times \mathbb{N}$  we have that  $\{1, 2\} \times \mathbb{N}$  is infinite and thus is denumerable. So there exists a bijection  $\varphi: \mathbb{N} \rightarrow \{1, 2\} \times \mathbb{N}$  and thus a surjection  $F \circ \varphi: \mathbb{N} \rightarrow A \bigcup B$  which means by 5.9 that  $A \bigcup B \approx E \subseteq \mathbb{N}$ . Using 5.55 we have that  $E$  is finite or denumerable hence  $A \bigcup B$  is either finite or denumerable, as  $A \subseteq A \bigcup B$  is a infinite subset we must have that  $A \bigcup B$  is not finite and thus that  $A \bigcup B$  is denumerable.  $\square$

Using mathematical induction we can extend this to a finite union of denumerable sets

**Corollary 5.62.** *Let  $n \in \mathbb{N}_0$ ,  $\{A_i\}_{i \in \{1, \dots, n\}}$  a finite family of denumerable sets then  $\bigcup_{i \in \{1, \dots, n\}} A_i$  is denumerable.*

**Proof.** We prove this by induction so let  $\mathcal{S} = \{n \in \mathbb{N}_0 \mid \forall \{A_i\}_{i \in \{1, \dots, n\}} \forall i \in \{1, \dots, n\} A_i \text{ is denumerable we have that } \bigcup_{i \in \{1, \dots, n\}} A_i \text{ is denumerable}\}$  then we have

$1 \in \mathcal{S}$ . If  $\{A_i\}_{i \in \{1, \dots, 1\}}$  is a family of denumerable sets then  $\bigcup_{i \in \{1, \dots, 1\}} A_i = A_1$  is denumerable proving that  $1 \in \mathcal{S}$

$n \in \mathcal{S} \Rightarrow n + 1 \in \mathcal{S}$ . Let  $\{A_i\}_{i \in \{1, \dots, n+1\}}$  be a family of denumerable sets then as  $n \in \mathcal{S}$  we have that  $\bigcup_{i \in \{1, \dots, n\}} A_i$  is denumerable. So  $\bigcup_{i \in \{1, \dots, n+1\}} A_i = (\bigcup_{i \in \{1, \dots, n\}} A_i) \bigcup A_{n+1}$  is denumerable by 5.61, proving that  $n + 1 \in \mathcal{S}$ .  $\square$

As also the finite union of finite sets is finite (see 5.39) we have the following corollary of the above corollary.

**Corollary 5.63.** *The finite union of countable sets is countable.*

**Corollary 5.64.** *Let  $A$  be a non empty finite set,  $B$  a denumerable set then  $A \times B$  is denumerable*

**Proof.** First as  $A$  is finite and non empty there exists a bijection  $\sigma: \{1, \dots, n\} \rightarrow A$  where  $n \in \mathbb{N}_0$ . So we have if  $x = (x_1, x_2) \in A \times B$  then  $x_1 \in A \wedge x_2 \in B \xrightarrow{\exists i \in \{1, \dots, n\} \vdash \sigma(i) = x_1} x_1 = \sigma(i) \in \{\sigma(i)\} \wedge x_2 \in B$ , proving that  $x = (x_1, x_2) \in \bigcup_{i \in \{1, \dots, n\}} (\{\sigma(i)\} \times B)$ . If  $x = (x_1, x_2) \in \bigcup_{i \in \{1, \dots, n\}} (\{\sigma(i)\} \times B) \Rightarrow \exists i \in \{1, \dots, n\}$  so that  $x_1 = \sigma(i) \subseteq A \wedge x_2 \in B \Rightarrow x = (x_1, x_2) \in A \times B$ . So we have

$$\bigcup_{i \in \{1, \dots, n\}} (\{\sigma(i)\} \times B) = A \times B \quad (5.1)$$

As  $B$  is denumerable there exists a bijection  $\tau: \mathbb{N}_0 \rightarrow B$ . Then  $\forall i \in \{1, \dots, n\}$  we have that  $\rho_i: \mathbb{N}_0 \rightarrow \{\sigma(i)\} \times B$  defined by  $\tau(j) = (\sigma(i), \tau(j))$  is bijection for we have

**injectivity.** If  $\rho_i(k) = \rho_i(l)$  then  $(\sigma(i), \tau(k)) = (\sigma(i), \tau(l)) \Rightarrow \tau(k) = \tau(l) \xrightarrow{\tau \text{ is bijection}} k = l$

**surjectivity.** If  $(x, y) \in \{\sigma(i)\} \times B$  then as  $\tau$  is a bijection there exists a  $k \in \mathbb{N}_0$  such that  $\tau(k) \in B$  so that  $\rho_i(k) = (\sigma(i), \tau(k)) \in \{\sigma(i)\} \times B$ .

The above proves that  $\forall i \in \{1, \dots, n\}$  we have that  $\{\sigma(i)\} \times B$  is denumerable, the theorem follows then from 5.62 and 5.1.  $\square$

## 5.4 Some properties of countable sets

Using the definition of **countability** and 5.55 we have the following theorem

**Theorem 5.65.** *Every subset of a denumerable set is countable.*

As the subsets of finite sets are finite we have that the following is true.

**Corollary 5.66.** *Every subset of a countable set is countable.*

**Theorem 5.67.** *If  $X$  is a non-empty set then the following are equivalent*

1.  $X$  is countable
2. There exists a surjective function  $f: \mathbb{N} \rightarrow X$
3. There exists a injective function  $f: X \rightarrow \mathbb{N}$
4. There exists a denumerable set  $Y$  and a injection  $f: X \rightarrow Y$

**Proof.**

**1  $\Rightarrow$  2.** If  $X$  is countable then we have the following possibilities

1. ( **$X$  is finite**) Then  $\exists n \in \mathbb{N}$  such that  $n \approx X$  so there exists a bijection  $f': n = S_n \rightarrow X$ , also as  $X \neq \emptyset$  there exists a  $x \in X$ . Define now  $f: \mathbb{N} \rightarrow X$  by

$$f(n) = \begin{cases} f'(n) & \text{if } i < n \text{ (or } i \in S_n\text{)} \\ x & \text{if } n \leq i \end{cases}$$

which is surjective because  $f'$  is surjective.

2. ( **$X$  is infinite**) Then  $\mathbb{N} \approx X \Rightarrow$  there exists a bijection (thus surjection) between  $\mathbb{N}$  and  $X$

**2  $\Rightarrow$  3.** If a surjective function  $f: \mathbb{N} \rightarrow X$  exists then by 2.208 there exists a injective function  $g: X \rightarrow \mathbb{N}$

**3  $\Rightarrow$  1.** If  $X$  is finite then it is of course countable. So assume that  $X$  is infinite. Then from the existence of the injective function  $f: X \rightarrow \mathbb{N}$  we have that  $f: X \rightarrow f(X)$  is a bijection and thus  $f(X)$  is infinite. From 5.54 we have then that  $f(X)$  is denumerable  $\Rightarrow X \approx f(X) \approx \mathbb{N} \Rightarrow X \approx \mathbb{N}$  and thus  $X$  is countable.

**3  $\Rightarrow$  4.** As  $\mathbb{N}$  is denumerable (4) follows automatically

**4  $\Rightarrow$  1.** As  $Y$  is denumerable there exists a bijection  $g: Y \rightarrow \mathbb{N}$  so that  $g \circ f: X \rightarrow \mathbb{N}$  is injective [see 2.18] which by 3  $\Rightarrow$  1 proves that  $X$  is countable.  $\square$

**Corollary 5.68.** *Let  $X$  be a non-empty set and  $Y$  a denumerable set then if  $f: X \rightarrow Y$  is an injection it follows that  $Y$  is countable*

**Proof.**

$\square$

**Lemma 5.69.** *If  $n \in \mathbb{N}$  then  $n \times \mathbb{N}$  and  $\mathbb{N} \times n$  are countable. If  $n \neq 0$  then  $n \times \mathbb{N}$  and  $\mathbb{N} \times n$  are denumerable.*

**Proof.** First  $f: n \times \mathbb{N} \rightarrow \mathbb{N} \times n$  define by  $(i, j) \Rightarrow f(i, j) = (j, i)$  is bijective

1. **(injectivity)** If  $f(i, j) = f(i', j') \Rightarrow (j, i) = (j', i') \Rightarrow i = i' \wedge j = j' \Rightarrow (i, j) = (i', j')$
2. **(surjectivity)** If  $(i, j) \in \mathbb{N} \times n \Rightarrow i \in \mathbb{N} \wedge j \in n \Rightarrow (j, i) \in n \times \mathbb{N}$  and  $f(j, i) = (i, j)$

so we have  $n \times \mathbb{N} \approx \mathbb{N} \times n$ . We must thus only prove that  $n \times \mathbb{N}$  is denumerable. First  $n \times \mathbb{N} = S_n \times \mathbb{N} \subseteq \mathbb{N} \times \mathbb{N}$  so we have by 5.55 that  $n \times \mathbb{N}$  is either finite or denumerable and is thus countable. If  $n \neq 0$  then  $0 < n = S_n \Rightarrow 0 \in n \Rightarrow \{0\} \subseteq \mathbb{N} \Rightarrow \{0\} \times \mathbb{N} \subseteq n \times \mathbb{N}$ . And as it is trivial to prove that  $f: \mathbb{N} \rightarrow \{0\} \times \mathbb{N}: i \rightarrow (0, i)$  is a bijection we have that  $\{0\} \times \mathbb{N}$  is denumerable and thus by 5.34 it is infinite, so by 5.35  $n \times \mathbb{N}$  is not finite and must thus be denumerable.  $\square$

**Theorem 5.70.** *If  $A, B$  are countable sets then  $A \times B$  is countable*

**Proof.** If  $A, B$  are countable sets then we have the following cases to consider

1. **( $A$  finite  $\wedge$   $B$  finite)** then  $A \times B$  is finite by 5.44 and thus countable.
2. **( $A$  finite  $\wedge$   $B$  denumerable)** Then  $\exists n \in \mathbb{N}$  such that  $A \approx n$  and  $B \approx \mathbb{N} \Rightarrow A \times B \approx n \times \mathbb{N}$  which by the above lemma is countable so we have that  $A \times B$  is countable.
3. **( $A$  denumerable  $\wedge$   $B$  finite)** Then  $\exists n \in \mathbb{N}$  such that  $A \approx \mathbb{N}$  and  $B \approx n \Rightarrow A \times B \approx \mathbb{N} \times n$  which by the above lemma is countable; so we have that  $A \times B$  is countable.
4. **( $A$  denumerable  $\wedge$   $B$  denumerable)** Then by 5.59 we have that  $A \times B$  is denumerable and thus countable.  $\square$

**Lemma 5.71.** *If  $\{A_i\}_{i \in S_n}$  is a family of countable sets then  $\bigcup_{i \in S_n} A_i$  is countable.*

**Proof.** We prove this by induction on  $n$ . So define  $S = \{n \in \mathbb{N} \mid \text{if } \{A_i\}_{i \in S_n} \text{ is a family of countable sets then } \bigcup_{i \in S_n} A_i \text{ is countable}\}$  then we have

1. If  $n = 0$  take then  $\{A_i\}_{i \in S_0}$  then there exists a graph  $A$  with  $\text{dom}(A) = \emptyset$  so if  $x \in \bigcup_{i \in S_0} A_i$  then there exists a  $i \in S_0$  such that  $x \in A_i$  which is a contradiction as  $S_0 = \emptyset$  so we have that  $\bigcup_{i \in S_0} A_i = \emptyset$  and thus finite and countable. So we have  $0 \in S$
2. Assume that  $n \in S$  and take  $n+1$  then we have for the family of countable sets  $\{A_i\}_{i \in S_{n+1}}$  that there exists a graph  $A$  with  $\text{dom}(A) = S_{n+1}$ . If we take  $A' = \{(x, y) \in A \mid x \in S_n\}$  (so  $\text{dom}(A') = S_n$ ) then  $\bigcup_{i \in S_{n+1}} A_i = A_n \bigcup (\bigcup_{i \in S_n} A'_i)$

**Proof.**

- a. If  $x \in \bigcup_{i \in S_{n+1}} A_i$  then  $\exists i \in \text{dom}(A) = S_{n+1} \Rightarrow i < n+1$  such that  $x \in A_i$  then we have the following cases
    - i. **( $i = n$ )** then  $x \in A_n \Rightarrow x \in A_n \bigcup (\bigcup_{i \in S_n} A'_i)$
    - ii. **( $i \neq n$ )** then  $i < n$  [if  $n \leq i \Rightarrow n < i < n+1 \Rightarrow n+1 \leq i < n+1 \Rightarrow n+1 \neq n+1$  a contradiction] and thus  $i \in S_n$ . As  $x \in A_i$  we have  $(i, x) \in A \Rightarrow (i, x) \in A' \Rightarrow x \in \bigcup_{i \in S_n} A'_i \Rightarrow x \in A_n \bigcup (\bigcup_{i \in S_n} A'_i)$
- in all cases we have  $x \in A_n \bigcup (\bigcup_{i \in S_n} A'_i)$

b. If  $x \in A_n \cup (\bigcup_{i \in S_n} A'_i)$  then we have the following cases

$$\text{i. } (x \in A_n) \Rightarrow x \in \bigcup_{i \in S_{n+1}} A_i$$

$$\text{ii. } (x \notin A_n) \Rightarrow x \in \bigcup_{i \in S_n} A'_i \Rightarrow \exists i \in S_n \vdash x \in A'_i \Rightarrow (i, x) \in A' \subseteq A \Rightarrow x \in A_i \Rightarrow x \in \bigcup_{i \in S_{n+1}} A_i$$

in all cases  $x \in \bigcup_{i \in S_{n+1}} A_i$  □

As  $n \in S$  we have  $\bigcup_{i \in S_n} A'_i$  is countable and as by assumption  $A_n$  is finite we have by 5.63 that  $\bigcup_{i \in S_{n+1}} A_i = A_n \cup (\bigcup_{i \in S_n} A'_i)$  is finite and thus  $n+1 \in S$ .

By induction we have then  $S = \mathbb{N}$  proving our theorem. □

**Lemma 5.72.** *If  $\{A_i\}_{i \in I}$  is a **finite** family ( $I$  is finite) of **countable** sets then  $\bigcup_{i \in I} A_i$  is **countable**.*

**Proof.** As  $I$  is finite we have the existence of a bijection  $f: n = S_n \rightarrow I$  ( $n \in \mathbb{N}$ ) and thus using 2.68 we have that  $\bigcup_{i \in I} A_i = \bigcup_{j \in S_n} A_{f(j)}$  which is countable by the previous lemma. □

**Theorem 5.73.** *If  $\{A_i\}_{i \in B}$  is a **denumerable** family of **countable** sets (so  $B$  is denumerable and  $\forall i \in B$  we have that  $A_i$  is a **countable** set) then  $\bigcup_{i \in B} A_i$  is **countable**. Further if we have in addition that  $\forall i \in B$   $A_i \neq \emptyset$  and  $\forall i, j \in B$  with  $i \neq j$  that  $A_i \cap A_j = \emptyset$  then  $\bigcup_{i \in B} A_i$  is denumerable.*

**Proof.** As  $\{A_i\}_{i \in B}$  is a denumerable family of countable sets, there exists a graph  $A$  with  $\text{dom}(G) = B$  where  $\mathbb{N} \approx B$ , so there is a bijection  $f: \mathbb{N} \rightarrow B$ , and  $\forall i \in B$  we have that  $A_i$  is countable. So  $\forall i \in B$  there exists a bijection  $g: N \rightarrow A_i$  where  $N$  is either  $\mathbb{N}$  or  $n \in \mathbb{N}$ . So if we define  $\forall i \in B$  the non empty set (see 2.76)  $\mathcal{A}_i = \{f \in A_i^{\mathbb{N}} \mid f \text{ is a bijection}\} \cup (\bigcup_{n \in \mathbb{N}} \{f \in A_i^n \mid f \text{ is a bijection}\})$ , then  $\mathcal{A} = \bigcup_{i \in B} \mathcal{A}_i$  is a set and  $\forall i \in B$  we have  $\mathcal{A}_i \subseteq \mathcal{A}$ . Using the axiom of choice (see 2.207) we have then the existence of a choice function  $c: \mathcal{P}'(\mathcal{A}) \rightarrow \mathcal{A}$  so that  $c(\mathcal{A}_i) \in \mathcal{A}_i \Rightarrow c(\mathcal{A}_i)$  is a bijection from  $\mathbb{N} \rightarrow A_i$  or a bijection from  $n \rightarrow A_i$ . Define now  $F: \mathbb{N} \times \mathbb{N} \rightarrow \bigcup_{i \in B} A_i$  by  $F(n, m) = c(\mathcal{A}_{f(n)})(m)$ .

We can then prove that  $F$  is surjective. For if  $y \in \bigcup_{i \in B} A_i \Rightarrow \exists i \in B \vdash y \in A_i \xrightarrow{f: \mathbb{N} \rightarrow B \text{ is bijective}} \exists n \in \mathbb{N} \vdash f(n) = i \Rightarrow y \in A_{f(n)}$ . We have now the following cases for  $c(A_i) = c(A_{f(n)})$ :

1.  $(\exists m \in \mathbb{N} \vdash c(A_{f(n)}) : m \rightarrow A_{f(n)} \text{ is a bijection}) \Rightarrow \exists j \in m \vdash c(A_{f(n)})(j) = y \Rightarrow F(n, j) = c(A_{f(n)})(j) = y$
2.  $(c(A_{f(n)}) : \mathbb{N} \rightarrow A_{f(n)} \text{ is a bijection}) \Rightarrow \exists j \in \mathbb{N} \vdash c(A_{f(n)})(j) = y \Rightarrow F(n, j) = c(A_{f(n)})(j) = y$

proving that  $F$  is surjective.

As  $\mathbb{N} \times \mathbb{N} \approx \mathbb{N}$  there exists a bijection  $\varphi: \mathbb{N} \rightarrow \mathbb{N} \times \mathbb{N}$  and thus  $F \circ \varphi: \mathbb{N} \rightarrow \bigcup_{i \in B} A_i$  is a surjective function. Using 5.9 we have then that  $\bigcup_{i \in B} A_i \approx E \subseteq \mathbb{N}$ . Using 5.55 we have that  $E$  is finite or denumerable hence  $\bigcup_{i \in B} A_i$  is either finite or denumerable and thus countable.

Finally if in addition  $\forall i \in B A_i = \emptyset$  and  $\forall i, j \in B \vdash i \neq j$  we have  $A_i \cap A_j$ . Using the axiom of choice (see 2.207) there exists a function  $\tau: B \rightarrow \bigcup_{i \in B} A_i$  such that  $\forall i \in B$  we have  $\tau(i) \in A_i$ . If  $\tau(i) = \tau(j)$  then  $\tau(i) \in A_i \wedge \tau(i) = \tau(j) \in A_j \Rightarrow \tau(i) \in A_i \cap A_j$  from which it follows that  $i = j$  [if  $i \neq j$  then  $\tau(i) \in A_i \cap A_j = \emptyset$  a contradiction], hence  $\tau: B \rightarrow \bigcup_{i \in B} A_i$  is a injection. From this it follows that  $\tau: B \rightarrow \tau(B) \subseteq \bigcup_{i \in B} A_i$  is a bijection, hence  $\tau \circ f: \mathbb{N} \rightarrow \tau(B)$  is a bijection which means that  $\tau(B)$  is denumerable and as  $\tau(B) \subseteq \bigcup_{i \in B} A_i$  we have by 5.34 that  $\bigcup_{i \in B} A_i$  is infinite and as it is countable it must be finite. □

Using the previous lemma and the above theorem we have then the following theorem.

**Theorem 5.74.** *If  $\{A_i\}_{i \in B}$  is a **countable** family of **countable** sets then we have that  $\bigcup_{i \in B} A_i$  is **countable**.*

**Proof.** If  $\{A_i\}_{i \in B}$  is a countable family of countable sets then we have either:

1. ( **$B$  is finite**) then by 5.72 we have that  $\bigcup_{i \in B} A_i$  is countable.

2. ( $B$  is denumerable) then by 5.74 we have that  $\bigcup_{i \in B} A_i$  is countable.  $\square$

## 5.5 Sequences

**Definition 5.75.** If  $X$  is a set then a family  $\{x_i\}_{i \in \mathbb{N}_0}$  of elements in  $X$  is called a sequence.

**Proposition 5.76.** Let  $\langle X, \leq \rangle$  be a partially ordered set and  $\{x_i\}_{i \in \mathbb{N}_0} \subseteq X$  a sequence of elements in  $X$  such that  $\forall i \in \mathbb{N}_0$  we have  $x_i \leq x_{i+1}$  then  $\forall i, j \in \mathbb{N}_0$  with  $i < j$  we have  $x_i \leq x_j$ .

**Proof.** We prove this by induction so given  $i \in \mathbb{N}_0$  take  $\mathcal{S}_i = \{n \in \mathbb{N}_0 \mid x_i \leq x_{i+n}\}$  then we have

$1 \in \mathcal{S}_i$ . by assumption we have  $x_i \leq x_{i+1}$  proving that  $1 \in \mathcal{S}_i$

$n \in \mathcal{S}_i \Rightarrow n+1 \in \mathcal{S}_i$ . as  $n \in \mathcal{S}_i$  we have  $x_i \leq x_{i+n}$  and by the hypothesis we have  $x_{i+n} \leq x_{(i+n)+1} \leq x_i + (n+1)$  so by transitivity we have  $x_i \leq x_{i+(n+1)}$  proving that  $n+1 \in \mathcal{S}_i$

Mathematical induction proves that  $\mathbb{N}_0 = \mathcal{S}_i$ . So if  $i, j \in \mathbb{N}_0$  with  $i < j$  then  $j - i \in \mathbb{N}_0 = \mathcal{S}_i$  and  $x_i \leq x_{i+(j-i)} = x_j$  proving the proposition.  $\square$

**Proposition 5.77.** Let  $\langle X, \leq \rangle$  be a partially ordered set and  $\{x_i\}_{i \in \mathbb{N}_0} \subseteq X$  a sequence of elements in  $X$  such that  $\forall i \in \mathbb{N}_0$  we have  $x_{i+1} \leq x_i$  then  $\forall i, j \in \mathbb{N}_0$  with  $i < j$  we have  $x_j \leq x_i$ .

**Proof.** We prove this by induction so given  $i \in \mathbb{N}_0$  take  $\mathcal{S}_i = \{n \in \mathbb{N}_0 \mid x_{i+n} \leq x_i\}$  then we have

$1 \in \mathcal{S}_i$ . by assumption we have  $x_{i+1} \leq x_i$  proving that  $1 \in \mathcal{S}_i$

$n \in \mathcal{S}_i \Rightarrow n+1 \in \mathcal{S}_i$ . as  $n \in \mathcal{S}_i$  we have  $x_{i+n} \leq x_i$  and by the hypothesis we have  $x_{i+(n+1)} = x_{(i+n)+1} \leq x_i$  so by transitivity we have  $x_{i+(n+1)} \leq x_i$  proving that  $n+1 \in \mathcal{S}_i$

Mathematical induction proves that  $\mathbb{N}_0 = \mathcal{S}_i$ . So if  $i, j \in \mathbb{N}_0$  with  $i < j$  then  $j - i \in \mathbb{N}_0 = \mathcal{S}_i$  and  $x_j = x_{i+(j-i)} \leq x_i$  proving the proposition.  $\square$

**Corollary 5.78.** Let  $\mathcal{A}$  be a class of classes and  $\{A_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{A}$  such that  $\forall i \in \mathbb{N}_0$  we have  $A_i \subseteq A_{i+1}$  [or  $A_{i+1} \subseteq A_i$ ] then  $\forall i, j \in \mathbb{N}_0$  with  $i < j$  we have  $A_i \subseteq A_j$  [or  $A_j \subseteq A_i$ ]

**Proof.** This follows because  $\langle \mathcal{A}, \subseteq \rangle$  is a partially ordered class (see 2.137) so that we can apply 5.76 and 5.77.  $\square$

## 5.6 Finite Cartesian product of sets

First some notations about the generalized product of sets in the finite case (see 2.81 and 2.88)

**Notation 5.79.** Let  $n, m \in \mathbb{N}_0$ ,  $\{A_i\}_{i \in \{n, \dots, m\}}$  be a finite family of sets then we note  $x \in \prod_{i \in \{n, \dots, m\}} A_i$  by  $x = (x_n, \dots, x_m)$  meaning  $x = \{(n, x_n), \dots, (m, x_m)\} \subseteq \{n, \dots, m\} \times (\bigcup_{i \in \{n, \dots, m\}} A_i)$  such that  $\langle x, \{n, \dots, m\} \rangle$  is a pretuple and  $\forall i \in \{n, \dots, m\}$  we have  $x_i \in A_i$  (see the definition of  $\prod_{i \in \{n, \dots, m\}} A_i$  at 2.81)

Using the above notation we can rephrase theorem 2.103 as follows

**Theorem 5.80.** Let  $n, k \in \mathbb{N}_0$  such that  $1 \leq k < n$  and  $\{A_i\}_{i \in \{n, \dots, m\}}$  a finite family then given  $x \in \prod_{i \in \{1, \dots, n\}} A_i$  we have that  $x|_{\{1, \dots, k\}} \in \prod_{i \in \{1, \dots, k\}} A_i$ ,  $x|_{\{k+1, \dots, n\}} \in \prod_{i \in \{k+1, \dots, n\}} A_i$  and  $\beta: \prod_{i \in \{1, \dots, n\}} A_i \rightarrow (\prod_{i \in \{1, \dots, k\}} A_i) \times (\prod_{i \in \{k+1, \dots, n\}} A_i)$  defined by  $\beta(x) = (x|_{\{1, \dots, k\}}, x|_{\{k+1, \dots, n\}})$  is a bijection. Or using the above notation we reformulate that as, if  $(x_1, \dots, x_n) \in \prod_{i \in \{1, \dots, n\}} A_i$  then  $(x_1, \dots, x_k) \in \prod_{i \in \{1, \dots, k\}} A_i$ ,  $(x_{k+1}, \dots, x_n) \in \prod_{i \in \{k+1, \dots, n\}} A_i$  and  $\beta((x_1, \dots, x_n)) = ((x_1, \dots, x_k), (x_{k+1}, \dots, x_n))$ .

The next theorem shows how to extend a tuple

**Theorem 5.81.** Let  $n \in \mathbb{N}_0$ ,  $\{A_i\}_{i \in \{1, \dots, n+1\}}$  be a finite family of sets and consider  $\{A_i\}_{i \in \{1, \dots, n\}}$  the finite subfamily formed by restriction to  $\{1, \dots, n\}$  then if  $x \in \prod_{i \in \{1, \dots, n\}} A_i$ ,  $y \in A_{n+1}$  define  $x+y = x \bigcup \{(n+1, y)\}$  then  $x+y \in \prod_{i \in \{1, \dots, n+1\}} A_i$  and  $(x+y)|_{\{1, \dots, n\}} = x$ .

**Proof.** As  $x \subseteq \{1, \dots, n\} \times (\bigcup_{i \in \{1, \dots, n\}} A_i)$  we have then if  $(i, z) \in x$  either

$$(i, z) \in x. \text{ so that } (i, z) \in \{1, \dots, n\} \times (\bigcup_{i \in \{1, \dots, n\}} A_i) \subseteq_{1.41} \{1, \dots, n+1\} \times (\bigcup_{i \in \{1, \dots, n\}} A_{n+1})$$

$$(i, z) \in \{(n+1, y)\}. \text{ so that } i = n+1 \wedge z = y \underset{y \in A_{n+1}}{\Rightarrow} i \in \{1, \dots, n+1\} \wedge y \in \bigcup_{i \in \{1, \dots, n+1\}} A_i \Rightarrow (i, z) \in \{1, \dots, n+1\} \times (\bigcup_{i \in \{1, \dots, n+1\}} A_i)$$

so we have in all cases  $(i, y) \in \{1, \dots, n+1\} \times (\bigcup_{i \in \{1, \dots, n+1\}} A_i)$  proving that

$$x+y \subseteq \{1, \dots, n+1\} \times \left( \bigcup_{i \in \{1, \dots, n+1\}} A_i \right). \quad (5.2)$$

Also if  $(i, z), (i, z') \in x+y$  we have the following cases to consider then

$$i \in \{1, \dots, n\}. \text{ then } (i, z), (i, z') \in x \underset{\langle x, \{1, \dots, n\} \rangle \text{ is a pretuple}}{\Rightarrow} z = z'$$

$$i = n+1. \text{ then } z = y = z'$$

proving that  $x+y$  is a function graph. If  $i \in \{1, \dots, n+1\}$  then if  $i = n+1$  we have  $(n+1, y) \in x+y$  and if  $i \in \{1, \dots, n\}$  there exists as  $\text{dom}(x) = \{1, \dots, n\}$  a  $z$  such that  $(i, z) \in x \subseteq x+y$  proving that  $\text{dom}(x+y) = \{1, \dots, n+1\}$ . So we have proved that

$$\langle x+y, \{1, \dots, n+1\} \rangle \text{ is a pretuple} \quad (5.3)$$

If  $i \in \{1, \dots, n+1\}$  then either  $i \in \{1, \dots, n\} \Rightarrow (i, x+y(i)) \in x \Rightarrow x+y(i) = x(i) \in A_i$  or  $i = n+1$  then  $x+y(i) = x+y(n+1) = y \in A_{n+1}$  so that

$$\forall i \in \{1, \dots, n+1\} \text{ we have } x+y(i) \in A_i \quad (5.4)$$

Using 5.2, 5.3 and 5.4 we have that  $x+y \in \prod_{i \in \{1, \dots, n+1\}} A_i$  as also  $(x+y)|_{\{1, \dots, n\}} = \{(i, z) \in x+y \mid i \in \{1, \dots, n\}\} = x$  we have proved the theorem.  $\square$

Next we define the finite Cartesian product of sets and show its relation with the normal product of a family of sets (see 2.81).

**Definition 5.82. (finite Cartesian product of sets)** Let  $n \in \mathbb{N}_0$  and  $\{A_i\}_{i \in \{1, \dots, n\}}$  a finite family of sets then we define  $\bigotimes_{i \in \{1, \dots, n\}} A_i$  recursively as follows:

$$n = 1. \bigotimes_{i \in \{1, \dots, 1\}} A_i = A_1$$

$$n > 1. \bigotimes_{i \in \{1, \dots, n\}} A_i = (\bigotimes_{i \in \{1, \dots, n-1\}} A_i) \times A_n \text{ (see 1.38)}$$

**Example 5.83.** Let  $\{A_i\}_{i \in \{1, \dots, 3\}}$  be a finite family of sets then  $\bigotimes_{i \in \{1, \dots, 3\}} A_i = (\bigotimes_{i \in \{1, \dots, 2\}} A_i) \times A_3 = ((\bigotimes_{i \in \{1, \dots, 1\}} A_i) \times A_2) \times A_3 = (A_1 \times A_2) \times A_3$  and  $x \in \bigotimes_{i \in \{1, \dots, 3\}} A_i \Leftrightarrow \exists x_1 \in A_1, x_2 \in A_2, x_3 \in A_3$  so that  $x = ((x_1, x_2), x_3)$

**Theorem 5.84.** Let  $\{A_i\}_{i \in \{1, \dots, n\}}, \{B_i\}_{i \in \{1, \dots, n\}}$  be finite families of sets such that  $\forall i \in \{1, \dots, n\}$  we have  $B_i \subseteq A_i$  then  $\bigotimes_{i \in \{1, \dots, n\}} B_i \subseteq \bigotimes_{i \in \{1, \dots, n\}} A_i$

**Proof.** We prove this by induction, so let  $\mathcal{S} = \{n \in \mathbb{N}_0 \mid \text{If } \{B_i\}_{i \in \{1, \dots, n\}}, \{A_i\}_{i \in \{1, \dots, n\}} \text{ are finite families of sets such that } \forall i \in \{1, \dots, n\} B_i \subseteq A_i \text{ then } \bigotimes_{i \in \{1, \dots, n\}} B_i \subseteq \bigotimes_{i \in \{1, \dots, n\}} A_i\}$ . Then we have

$$1 \in \mathcal{S}. \text{ If } x \in \bigotimes_{i \in \{1, \dots, 1\}} B_i = B_1 \subseteq A_1 = \bigotimes_{i \in \{1, \dots, 1\}} A_i \text{ proving that } \bigotimes_{i \in \{1, \dots, 1\}} B_i \subseteq \bigotimes_{i \in \{1, \dots, 1\}} A_i, \text{ hence } 1 \in \mathcal{S}$$

**$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$ .** Let  $x \in \prod_{i \in \{1, \dots, n+1\}} B_i$  then as  $n+1 > 1$  then  $\exists y \in \bigotimes_{i \in \{1, \dots, n\}} B_i$  and  $\exists z \in B_{n+1}$  such that  $x = (y, z)$ , As  $n \in \mathcal{S}$  we have that  $y \in \bigotimes_{i \in \{1, \dots, n\}} A_i$ , further  $z \in B_{n+1} \subseteq A_{n+1}$ . Applying 1.41 gives  $x = (y, z) \in (\bigotimes_{i \in \{1, \dots, n\}} A_i) \times A_{n+1} = \bigotimes_{i \in \{1, \dots, n+1\}} A_i$  proving that  $\bigotimes_{i \in \{1, \dots, n+1\}} B_i \subseteq \bigotimes_{i \in \{1, \dots, n+1\}} A_i$ . Hence  $n+1 \in \mathcal{S}$ .  $\square$

Next we show that with every  $x \in \prod_{i \in \{1, \dots, n\}} A_i$  we can construct a  $[x] \in \bigotimes_{i \in \{1, \dots, n\}} A_i$

**Definition 5.85. (product tuple)** Let  $n \in \mathbb{N}_0$ ,  $A = \{A_i\}_{i \in \{1, \dots, n\}}$  be a finite family of sets,  $x \in \prod_{i \in \{1, \dots, n\}} A_i$  then we define  $[x]_n \in \bigotimes_{i \in \{1, \dots, n\}} A_i$  recursively by

**$n = 1$ .**  $[x]_1 = \pi_1(x)$

**$n > 1$ .**  $[x]_n = ([x]_{\{1, \dots, n-1\}}]_{n-1}, \pi_n(x))$

where  $\pi_i$  is the projection function defined in 2.91.

**Example 5.86.** For  $1 \in \mathbb{N}_0$  we have that for  $(x) \in \prod_{i \in \{1, \dots, 1\}} A_i$  that  $[x]_1 = \pi_1(x) = x$

**Example 5.87.** Let  $x = (x_1, x_2, x_3) \in \prod_{i \in \{1, \dots, 3\}} A_i$  then

$$\begin{aligned} [x]_3 &= ([x]_{\{1, \dots, 2\}}]_2, \pi_3(x)) \\ &= (([x]_{\{1\}}]_1, \pi_2(x)), \pi_3(x)) \\ &= ((\pi_1(x), \pi_2(x)), \pi_3(x)) \in \bigotimes_{i \in \{1, \dots, 3\}} A_i \end{aligned}$$

**Lemma 5.88.** Let  $n \in \mathbb{N}_0$ ,  $\{A_i\}_{i \in \mathbb{N}_0}$  be a family of sets,  $x \in \prod_{i \in \{1, \dots, n\}} A_i$  and  $\{B_i\}_{i \in \{1, \dots, n\}}$  a family of sets such that  $\forall i \in \{1, \dots, n\}$  we have  $B_i \subseteq A_i$  then  $x \in \prod_{i \in \{1, \dots, n\}} B_i \Leftrightarrow [x]_n \in \bigotimes_{i \in \{1, \dots, n\}} B_i$

**Proof.**

$\Rightarrow$ . We prove this by induction, let  $\mathcal{S} = \{n \in \mathbb{N}_0 \mid \forall \{A_i\}_{i \in \{1, \dots, n\}}, x \in \prod_{i \in \{1, \dots, n\}} A_i, \{B_i\}_{i \in \{1, \dots, n\}}$  such that  $\forall i \in \{1, \dots, n\} B_i$  we have if  $x \in \prod_{i \in \{1, \dots, n\}} B_i$  that  $[x]_n \in \bigotimes_{i \in \{1, \dots, n\}} B_i\}$  then we have

**$1 \in \mathcal{S}$ .** As  $x \in \prod_{i \in \{1, \dots, 1\}} B_i \Rightarrow \pi_1(x) = x(1) \in B_1$  so that  $[x]_1 = \pi_1(x) \in B_1 = \bigotimes_{i \in \{1, \dots, 1\}} B_i$  proving that  $1 \in \mathcal{S}$ .

**$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$ .** Assume for  $x \in \prod_{i \in \{1, \dots, n+1\}} A_i$  that  $x \in \prod_{i \in \{1, \dots, n+1\}} B_i$  then using 5.80 we have that  $x|_{\{1, \dots, n\}} \in \prod_{i \in \{1, \dots, n\}} A_i$  and  $x|_{\{1, \dots, n\}} \in \prod_{i \in \{1, \dots, n\}} B_i$ . Using the induction hypothesis  $n \in \mathcal{S}$  we have that  $[x|_{\{1, \dots, n\}}]_n \in \bigotimes_{i \in \{1, \dots, n\}} B_i$  and as also  $\pi_{n+1}(x) = x(n+1) \in B_{n+1}$  it follows that  $([x|_{\{1, \dots, n\}}]_n, \pi_{n+1}(x)) \in (\bigotimes_{i \in \{1, \dots, n\}} B_i) \times B_{n+1} = \bigotimes_{i \in \{1, \dots, n+1\}} B_i$ . Hence  $x \in \bigotimes_{i \in \{1, \dots, n+1\}} B_i$  proving that  $n+1 \in \mathcal{S}$

$\Leftarrow$ . We prove this by induction, let  $\mathcal{S} = \{n \in \mathbb{N}_0 \mid \forall \{A_i\}_{i \in \{1, \dots, n\}}, x \in \prod_{i \in \{1, \dots, n\}} A_i, \{B_i\}_{i \in \{1, \dots, n\}}$  such that  $\forall i \in \{1, \dots, n\} B_i \subseteq A_i$  we have if  $[x]_n \in \bigotimes_{i \in \{1, \dots, n\}} B_i$  then  $x \in \prod_{i \in \{1, \dots, n\}} B_i\}$  then

**$1 \in \mathcal{S}$ .** If  $[x]_1 \in \bigotimes_{i \in \{1, \dots, 1\}} B_i$  then  $x(1) = \pi_1(x) \in B_1$  so that by 2.87 we have that  $x \in \prod_{i \in \{1, \dots, 1\}} B_i$ , hence  $1 \in \mathcal{S}$

**$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$ .** Assume for  $x \in \prod_{i \in \{1, \dots, n+1\}} A_i$  that  $[x]_{n+1} \in \bigotimes_{i \in \{1, \dots, n+1\}} B_i = (\bigotimes_{i \in \{1, \dots, n\}} B_i) \times B_{n+1}$  then  $[x|_{\{1, \dots, n\}}]_n \in \bigotimes_{i \in \{1, \dots, n\}} B_i$  and  $\pi_{n+1}(x) \in B_{n+1} \Rightarrow x(n+1) \in B_{n+1}$ , As  $n \in \mathcal{S}$  it follows that  $x|_{\{1, \dots, n\}} \in \prod_{i \in \{1, \dots, n\}} B_i$ . So  $\forall i \in \{1, \dots, n+1\}$  we have  $x(i) \in B_i$  from which it follows using 2.87 that  $x \in \prod_{i \in \{1, \dots, n+1\}} B_i$ . This proves that  $n+1 \in \mathcal{S}$ .  $\square$

We prove now that  $[.]_n$  can be used to define a bijection between  $\prod_{i \in \{1, \dots, n\}} A_i$  and  $\bigotimes_{i \in \{1, \dots, n\}} A_i$ .

**Theorem 5.89.** Let  $n \in \mathbb{N}_0$ ,  $\{A_i\}_{i \in \{1, \dots, n\}}$  be a family of sets, then for the function  $\mathcal{P}_n: \prod_{i \in \{1, \dots, n\}} A_i \rightarrow \bigotimes_{i \in \{1, \dots, n\}} A_i$  defined by  $\mathcal{P}_n(x) = [x]_n$  we have that

1.  $\mathcal{P}_n$  is a bijection
2.  $\forall \{B_i\}_{i \in \{1, \dots, n\}}$  with  $\forall i \in \{1, \dots, n\} B_i \subseteq A_i$  we have for  $\mathcal{P}_n: \prod_{i \in \{1, \dots, n\}} A_i \rightarrow \bigotimes_{i \in \{1, \dots, n\}} A_i$  that  $\mathcal{P}_n(\prod_{i \in \{1, \dots, n\}} B_i) = \bigotimes_{i \in \{1, \dots, n\}} B_i$  (note that this is sensible because of 5.84) for as  $\mathcal{P}_n$  is a bijection  $\prod_{i \in \{1, \dots, n\}} B_i = (\mathcal{P}_n^{-1})(\bigotimes_{i \in \{1, \dots, n\}} B_i)$

**Proof.**

1. To prove that  $\mathcal{P}_n$  is a bijection we prove that  $\mathcal{P}_n$  is a injection and surjection by induction

**injectivity.** let  $\mathcal{S} = \{n \in \mathbb{N}_0 \mid \text{if } \{A_i\}_{i \in \{1, \dots, n\}} \text{ is a family of sets then } \mathcal{P}_n: \prod_{i \in \{1, \dots, n\}} A_i \rightarrow \bigotimes_{i \in \{1, \dots, n\}} A_i \text{ is injectivity}\}$  then we have

$1 \in \mathcal{S}$ . then if  $\mathcal{P}_1(x) = \mathcal{P}_1(y) \Rightarrow [x]_1 = [y]_1 \Rightarrow \pi_1(x) = \pi_1(y)$  so that  $x(1) = y(1)$  proving that the functions  $x, y: \{1\} \rightarrow A_1$  are the same. Hence  $1 \in \mathcal{S}$

$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$ . Assume that  $x, y \in \prod_{i \in \{1, \dots, n+1\}} A_i$  such that  $\mathcal{P}_{n+1}(x) = \mathcal{P}_{n+1}(y)$  then

$$\begin{aligned} \mathcal{P}_{n+1}(x) = \mathcal{P}_{n+1}(y) &\Rightarrow [x]_{n+1} = [y]_{n+1} \\ &\Rightarrow ([x|_{\{1, \dots, n\}}]_n, \pi_{n+1}(x)) = ([y|_{\{1, \dots, n\}}]_n, \pi_{n+1}(y)) \\ &\Rightarrow [x|_{\{1, \dots, n\}}]_n = [y|_{\{1, \dots, n\}}]_n \wedge \pi_{n+1}(x) = \pi_{n+1}(y) \\ &\Rightarrow [x|_{\{1, \dots, n\}}]_n = [y|_{\{1, \dots, n\}}]_n \wedge x(n+1) = y(n+1) \\ &\stackrel{n \in \mathcal{S}}{\Rightarrow} x|_{\{1, \dots, n\}} = y|_{\{1, \dots, n\}} \wedge x(n+1) = y(n+1) \\ &\Rightarrow \forall i \in \{1, \dots, n\} \models x(i) = y(i) \wedge x(n+1) = y(n+1) \\ &\Rightarrow x = y \end{aligned}$$

proving that  $\mathcal{P}_{n+1}$  is injective and thus that  $n+1 \in \mathcal{S}$ .

**surjectivity.** let  $\mathcal{S} = \{n \in \mathbb{N}_0 \mid \text{if } \{A_i\}_{i \in \{1, \dots, n\}} \text{ is a family of sets then } \mathcal{P}_n: \prod_{i \in \{1, \dots, n\}} A_i \rightarrow \bigotimes_{i \in \{1, \dots, n\}} A_i \text{ is surjective}\}$  then we have

$1 \in \mathcal{S}$ . then  $\bigotimes_{i \in \{1, \dots, 1\}} A_i = A_1$  and if  $x \in A_1$  define  $\hat{x}: \{1\} \rightarrow A_1$  by  $\hat{x}(1) = x \in A_1$  then  $\hat{x} \in \prod_{i \in \{1, \dots, 1\}} A_i$  (see 2.84) and  $\mathcal{P}_1(\hat{x}) = [\hat{x}]_1 = \pi_1(\hat{x}) = x$  proving that  $\mathcal{P}_1$  is surjective.

$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$ . let  $x \in \bigotimes_{i \in \{1, \dots, n+1\}} A_i$  then as  $n+1 > 1$  we have that  $x = (y, z)$  where  $y \in \bigotimes_{i \in \{1, \dots, n\}} A_i$  and  $z \in A_{n+1}$ . As  $n \in \mathcal{S}$  there exists a  $\hat{y} \in \prod_{i \in \{1, \dots, n\}} A_i$  such that  $\mathcal{P}_n(\hat{y}) = y$ . Define now  $\hat{x} \in \prod_{i \in \{1, \dots, n+1\}} A_i$  as  $\hat{y} + z$  (see 5.81). Then

$$\begin{aligned} \mathcal{P}_{n+1}(\hat{x}) &= (\mathcal{P}_n(\hat{x}|_{\{1, \dots, n\}}), \pi_{n+1}(\hat{x})) \\ &\stackrel{5.81}{=} (\mathcal{P}_n(\hat{y}), z) \\ &= (y, z) \end{aligned}$$

proving that  $\mathcal{P}_{n+1}$  is surjective so that  $n+1 \in \mathcal{S}$

2. If  $x \in \mathcal{P}_n(\prod_{i \in \{1, \dots, n\}} B_i)$  then  $\exists \hat{x} \in \prod_{i \in \{1, \dots, n\}} B_i \subseteq \prod_{i \in \{1, \dots, n\}} A_i$  such  $x = [\hat{x}]_n$ , using 5.88 we have that  $[\bar{x}]_n \in \bigotimes_{i \in \{1, \dots, n\}} B_i$  proving that  $\mathcal{P}_n(\prod_{i \in \{1, \dots, n\}} B_i) \subseteq \bigotimes_{i \in \{1, \dots, n\}} B_i$ . Finally if  $x \in \bigotimes_{i \in \{1, \dots, n\}} B_i$  we have as  $\mathcal{P}_n$  is a bijection that there exists a  $\hat{x} \in \prod_{i \in \{1, \dots, n\}} B_i$  such that  $x = [\hat{x}]_n$ . Applying then 5.88 we have  $\hat{x} \in \prod_{i \in \{1, \dots, n\}} B_i$  so that  $x \in \mathcal{P}_n(\prod_{i \in \{1, \dots, n\}} B_i)$  proving that  $\bigotimes_{i \in \{1, \dots, n\}} B_i \subseteq \mathcal{P}_n(\prod_{i \in \{1, \dots, n\}} B_i)$ .  $\square$

**Lemma 5.90.** Let  $n \in \mathbb{N}_0$ ,  $\{A_i\}_{i \in \{1, \dots, n\}}$  a finite family of sets then  $\bigotimes_{i \in \{1, \dots, n\}} A_i = \emptyset \Leftrightarrow \exists i \in \{1, \dots, n\} \models A_i = \emptyset$

**Proof.** We prove this by induction, so let  $\mathcal{S} = \{n \in \mathbb{N}_0 \mid \forall \{A_i\}_{i \in \{1, \dots, n\}} \text{ we have } \bigotimes_{i \in \{1, \dots, n\}} A_i = \emptyset \Leftrightarrow \exists i \in \{1, \dots, n\} \models A_i = \emptyset\}$  then

**1  $\in \mathcal{S}$ .** As  $\bigotimes_{i \in \{1, \dots, 1\}} A_i = A_1$  we have  $\bigotimes_{i \in \{1, \dots, 1\}} A_i = \emptyset \Leftrightarrow A_1 = \emptyset$  proving that  $1 \in \mathcal{S}$

**$n \in \mathcal{S} \Rightarrow n + 1 \in \mathcal{S}$ .** then we have

1. If  $\bigotimes_{i \in \{1, \dots, n+1\}} A_i = \emptyset$  we have

$$\begin{aligned} \bigotimes_{i \in \{1, \dots, n+1\}} A_i = \emptyset &\stackrel{n+1 > 1}{\Leftrightarrow} \left( \bigotimes_{i \in \{1, \dots, n\}} A_i \right) \times A_{n+1} = \emptyset \\ &\stackrel{1.40}{\Leftrightarrow} \left( \bigotimes_{i \in \{1, \dots, n\}} A_i \right) = \emptyset \vee A_{n+1} = \emptyset \\ &\stackrel{n \in \mathcal{S}}{\Leftrightarrow} (\exists i \in \{1, \dots, n\} \models A_i = \emptyset) \vee A_{n+1} = \emptyset \\ &\Leftrightarrow \exists i \in \{1, \dots, n+1\} \models A_i = \emptyset \end{aligned}$$

□

Using the above bijection we can define the projection function.

**Definition 5.91.** Let  $n \in \mathbb{N}_0$ ,  $\{A_i\}_{i \in \{1, \dots, n\}}$  a finite family of sets,  $i \in \{1, \dots, n\}$  then we define  $\pi_i^\times$ :  $\bigotimes_{j \in \{1, \dots, n\}} A_j \rightarrow A_i$  by  $\pi_i^\times = \pi_i \circ \mathcal{P}_n^{-1}$

**Example 5.92.** For the special case  $\{A_i\}_{i \in \{1, \dots, 1\}}$  we have that  $\pi_1^\times = \mathbb{1}_{A_1}$

**Proof.** Let  $x \in A_1 \stackrel{\text{def}}{=} \bigotimes_{i \in \{1, \dots, 1\}} A_i$  and define  $x' \in \prod_{i \in \{1, \dots, 1\}} A_i$  by  $x' = \{(1, x)\}$  then  $\mathcal{P}_1(x') = [x'] = \pi_1(x') = x$  so that  $\mathcal{P}_1^{-1}(x) = x'$  hence  $\pi_1^\times(x) = \pi_1(\mathcal{P}_1^{-1}(x)) = \pi_1(x') = x$  □

**Theorem 5.93.** Let  $n \in \mathbb{N}_0$ ,  $\{A_i\}_{i \in \{1, \dots, n\}}$  a finite family of sets then we have

1.  $\forall x, y \in \bigotimes_{i \in \{1, \dots, n\}} A_i$  we have  $x = y \Leftrightarrow \forall i \in \{1, \dots, n\}$  we have  $\pi_i^+(x) = \pi_i^+(y)$
2. If  $\forall i \in \{1, \dots, n\} A_i \neq \emptyset$  then  $\forall i \in \{1, \dots, n\}$  we have that  $\pi_i^\times: \bigotimes_{i \in \{1, \dots, n\}} A_i \rightarrow A_i$  is a surjection
3. Let  $\{B_i\}_{i \in \{1, \dots, n\}}$  another finite family of sets then if  $\forall x \in \bigotimes_{i \in \{1, \dots, n\}} A_i$  we have that  $\forall i \in \{1, \dots, n\} \pi_i^\times(x) \in B_i$  then  $\bigotimes_{i \in \{1, \dots, n\}} A_i \subseteq \bigotimes_{i \in \{1, \dots, n\}} B_i$

**Proof.**

1. If  $x, y \in \bigotimes_{i \in \{1, \dots, n\}} A_i$  then

$$\begin{aligned} x = y &\stackrel{\mathcal{P}_n \text{ is a bijection}}{\Leftrightarrow} \mathcal{P}_n^{-1}(x) = \mathcal{P}_n^{-1}(y) \\ &\Leftrightarrow \forall i \in \{1, \dots, n\} \text{ we have } (\mathcal{P}_n^{-1}(x))(i) = (\mathcal{P}_n^{-1}(y))(i) \\ &\Leftrightarrow \forall i \in \{1, \dots, n\} \text{ we have } \pi_i(\mathcal{P}_n^{-1}(x)) = \pi_i(\mathcal{P}_n^{-1}(y)) \\ &\Leftrightarrow \forall i \in \{1, \dots, n\} \text{ we have } \pi_i^\times(x) = \pi_i^\times(y) \end{aligned}$$

2. As  $\forall i \in \{1, \dots, n\} A_i \neq \emptyset$  we have by 2.93 that  $\forall i \in \{1, \dots, n\} \pi_i$  is a surjection. Take  $i \in \{1, \dots, n\}$  and  $x \in A_i$  then there exists a  $\hat{x} \in \prod_{i \in \{1, \dots, n\}} A_i$  such that

$$\begin{aligned} x &= \pi_1(\hat{x}) \\ &\stackrel{\text{bijective}}{=} \pi_i((\mathcal{P}_n^{-1})(\mathcal{P}_n(\hat{x}))) \\ &= \pi_i^\times(\mathcal{P}_n(\hat{x})) \end{aligned}$$

which as  $\mathcal{P}_n(x) \in \bigotimes_{i \in \{1, \dots, n\}} A_i$  proves that  $\pi_i^\times$  is a surjection.

3. We have to consider two cases

$\exists i \in \{1, \dots, n\} \vdash A_i = \emptyset$ . then using 5.90 we have that  $\bigotimes_{i \in \{1, \dots, n\}} A_i = \emptyset$  so that

$$\bigotimes_{i \in \{1, \dots, n\}} A_i \subseteq \bigotimes_{i \in \{1, \dots, n\}} B_i.$$

$\forall i \in \{1, \dots, n\} \models A_i \neq \emptyset$ . Take  $i \in \{1, \dots, n\}$  then by (2)  $\pi_i^\times$  is a surjection, so if  $x \in A_i$  there exists a  $\hat{x} \in \bigotimes_{i \in \{1, \dots, n\}} A_i$  such that  $x = \pi_i^\times(\hat{x})$ . As by the hypothesis we have that  $\pi_i^\times(\hat{x}) \in B_i$  it follows that  $x \in B_i$ . Hence  $\forall i \in \{1, \dots, n\}$  we have  $A_i \subseteq B_i$  and using 5.84 it follows that

$$\bigotimes_{i \in \{1, \dots, n\}} A_i \subseteq \bigotimes_{i \in \{1, \dots, n\}} B_i. \quad \square$$

In the next theorem we examine the product of product of sets.

**Theorem 5.94.** *Let  $n, k \in \mathbb{N}_0$  such that  $1 \leq k < n$  and  $\{A_i\}_{i \in \{1, \dots, n\}}$  a finite family of sets then there exists a bijection  $\beta: \bigotimes_{i \in \{1, \dots, n\}} A_i \rightarrow (\bigotimes_{i \in \{1, \dots, k\}} A_i) \times (\bigotimes_{i \in \{1, \dots, n-k\}} A_{k+i})$  such that given  $x \in \bigotimes_{i \in \{1, \dots, n\}} A_i$  we have*

$$\forall i \in \{1, \dots, k\} \pi_{n,i}^\times(x) = \pi_{k,i}(\pi_1(\beta(x)))$$

and

$$\forall i \in \{k+1, \dots, n\} \pi_{n,i}^\times(x) = \pi_{n-k,i-k}(\pi_2(\beta(x)))$$

where  $\pi_{n,i}^\times$  is the projection function between  $\bigotimes_{i \in \{1, \dots, n\}} A_i$  and  $A_i$ ,  $\pi_{k,i}^\times$  is the projection function between  $\bigotimes_{i \in \{1, \dots, k\}} A_i$  and  $A_i$ ,  $\pi_{n-k,i}^\times$  is the projection function between  $\bigotimes_{i \in \{1, \dots, n-k\}} A_{k+i}$  and  $A_{k+i}$ ,  $\pi_1$  is the projection function between  $(\bigotimes_{i \in \{1, \dots, k\}} A_i) \times (\bigotimes_{i \in \{1, \dots, n-k\}} A_{k+i})$  and  $\bigotimes_{i \in \{1, \dots, k\}} A_i$  and  $\pi_2$  is the projection function between  $(\bigotimes_{i \in \{1, \dots, k\}} A_i) \times (\bigotimes_{i \in \{1, \dots, n-k\}} A_{k+i})$  and  $\bigotimes_{i \in \{1, \dots, n-k\}} A_{k+i}$ .

**Proof.** We use the following bijections and functions in the constructing of the bijection  $\beta$ :

$$\alpha: \prod_{i \in \{1, \dots, n\}} A_i \rightarrow \left( \prod_{i \in \{1, \dots, k\}} A_i \right) \times \left( \prod_{i \in \{k+1, \dots, n\}} A_i \right) \text{ where } \alpha(x) = (x|_{\{1, \dots, k\}}, x|_{\{k+1, \dots, n\}})$$

which is a bijection because 5.80. As  $\rho: \{k+1, \dots, n\} \rightarrow \{1, \dots, n-k\}$  defined by  $\rho(i) = i - k$  is a bijection with inverse  $\rho^{-1}: \{1, \dots, n-k\} \rightarrow \{k+1, \dots, n\}$  defined by  $(\rho^{-1})(i) = i + k$ . Further we have by 2.85 that

$$\lambda: \prod_{i \in \{k+1, \dots, n\}} A_i \rightarrow \prod_{i \in \{1, \dots, n-k\}} A_{i+k} \text{ defined by } \lambda(x) = x \circ \rho^{-1}$$

is a bijection. Finally we need the following

$$\mathcal{P}_n: \prod_{i \in \{1, \dots, n\}} A_i \rightarrow \bigotimes_{i \in \{1, \dots, n\}} A_i \text{ (see 5.89)}$$

$$\mathcal{P}_k: \prod_{i \in \{1, \dots, k\}} A_i \rightarrow \bigotimes_{i \in \{1, \dots, k\}} A_i \text{ (see 5.89)}$$

$$\mathcal{P}_{n-k}: \prod_{i \in \{1, \dots, n-k\}} A_i \rightarrow \bigotimes_{i \in \{1, \dots, n-k\}} A_{i+k}$$

We are ready now to define

$$\beta: \bigotimes_{i \in \{1, \dots, n\}} A_i \rightarrow \left( \bigotimes_{i \in \{1, \dots, k\}} A_i \right) \times \left( \bigotimes_{i \in \{1, \dots, n-k\}} A_{k+i} \right)$$

by

$$\beta(x) = (\mathcal{P}_k(\pi_1(\alpha(\mathcal{P}_n^{-1}(x)))), \mathcal{P}_{n-k}(\lambda(\pi_2(\alpha(\mathcal{P}_n^{-1}(x))))))$$

Then we have for  $x \in \bigotimes_{i \in \{1, \dots, n\}} A_i$  that for  $i \in \{1, \dots, k\}$

$$\begin{aligned} \pi_{k,i}^{\times}(\pi_1(\beta(x))) &= \pi_{k,i}^{\times}(\mathcal{P}_k(\pi_1(\alpha(\mathcal{P}_n^{-1}(x))))) \\ &= \pi_{k,i}(\mathcal{P}_k^{-1}(\mathcal{P}_k(\pi_1(\alpha(\mathcal{P}_n^{-1}(x)))))) \\ &= \pi_{k,i}(\pi_1(\alpha(\mathcal{P}_n^{-1}(x)))) \\ &= \pi_{k,i}(\mathcal{P}_n^{-1}(x)_{\{1, \dots, k\}}) \\ &= \mathcal{P}_n^{-1}(x)_{\{1, \dots, k\}}(i) \\ &= \mathcal{P}_n^{-1}(x)(i) \\ &= \pi_{n,i}(\mathcal{P}_n^{-1}(x)) \\ &= \pi_{n,i}^{\times}(x) \end{aligned} \quad (5.5)$$

$$= \pi_{n,i}^{\times}(x) \quad (5.6)$$

and for  $i \in \{k+1, \dots, n-k\}$

$$\begin{aligned} \pi_{n-k,i-k}^{\times}(\pi_2(\beta(x))) &= \pi_{n-k,i-k}^{\times}(\mathcal{P}_{n-k}(\lambda(\pi_2(\alpha(\mathcal{P}_n^{-1}(x)))))) \\ &= \pi_{n-k,i-k}(\mathcal{P}_{n-k}^{-1}(\mathcal{P}_{n-k}(\lambda(\pi_2(\alpha(\mathcal{P}_n^{-1}(x))))))) \\ &= \pi_{n-k,i-k}(\lambda(\pi_2(\alpha(\mathcal{P}_n^{-1}(x))))) \\ &= \pi_{n-k,i-k}(\lambda(\mathcal{P}_n^{-1}(x)_{\{k+1, \dots, n\}})) \\ &= \lambda(\mathcal{P}_n^{-1}(x)_{\{k+1, \dots, n\}})(i-k) \\ &= \mathcal{P}_n^{-1}(x)_{\{k+1, \dots, n\}}(i) \\ &= \mathcal{P}_n^{-1}(x)(i) \\ &= \pi_{n,i}(\mathcal{P}_n^{-1}(x)) \\ &= \pi_{n,i}^{\times}(x) \end{aligned} \quad (5.7)$$

$$= \pi_{n,i}^{\times}(x) \quad (5.8)$$

The above proves part of the theorem, for the rest we have to prove that  $\beta$  is a bijection:

**injectivity.** Let  $x, y \in \bigotimes_{i \in \{1, \dots, n\}} A_i$  such that  $\beta(x) = \beta(y)$  then we have  $\forall i \in \{1, \dots, n\}$

$$\begin{aligned} \pi_{k,i}^{\times}(\pi_1(\beta(x))) &= \pi_{k,i}^{\times}(\pi_1(\beta(y))) \xrightarrow[5.5, 5.7]{} \mathcal{P}_n^{-1}(x)(i) = \mathcal{P}_n^{-1}(y)(i) \\ &\Rightarrow \mathcal{P}_n^{-1}(x) = \mathcal{P}_n^{-1}(y) \\ &\xrightarrow[\mathcal{P}_n \text{ is bijective}]{} x = y \end{aligned}$$

proving injectivity.

**surjectivity.** Let  $(x, y) \in (\bigotimes_{i \in \{1, \dots, k\}} A_i) \times (\bigotimes_{i \in \{k+1, \dots, n\}} A_i)$  take then

$$z = \mathcal{P}_n(\alpha^{-1}((\mathcal{P}_k^{-1}(x)), \lambda^{-1}(\mathcal{P}_{n-k}^{-1}(y))))$$

then we have

$$\begin{aligned} \mathcal{P}_k(\pi_1(\alpha(\mathcal{P}_n^{-1}(z)))) &= \mathcal{P}_k(\pi_1(\alpha(\mathcal{P}_n^{-1}(\mathcal{P}_n(\alpha^{-1}((\mathcal{P}_k^{-1}(x)), \lambda^{-1}(\mathcal{P}_{n-k}^{-1}(y)))))))) \\ &= \mathcal{P}_k(\pi_1(\alpha(\alpha^{-1}((\mathcal{P}_k^{-1}(x)), \lambda^{-1}(\mathcal{P}_{n-k}^{-1}(y)))))) \\ &= \mathcal{P}_k(\pi_1((\mathcal{P}_k^{-1}(x)), \lambda^{-1}(\mathcal{P}_{n-k}^{-1}(y)))) \\ &= \mathcal{P}_k(\mathcal{P}_k^{-1}(x)) \\ &= x \end{aligned}$$

and

$$\begin{aligned} \mathcal{P}_{n-k}(\lambda(\pi_2(\alpha(\mathcal{P}_n^{-1}(z))))) &= \mathcal{P}_{n-k}(\lambda(\pi_2(\alpha(\mathcal{P}_n^{-1}(\mathcal{P}_n(\alpha^{-1}((\mathcal{P}_k^{-1}(x)), \lambda^{-1}(\mathcal{P}_{n-k}^{-1}(y)))))))))) \\ &= \mathcal{P}_{n-k}(\lambda(\pi_2(\alpha(\alpha^{-1}((\mathcal{P}_k^{-1}(x)), \lambda^{-1}(\mathcal{P}_{n-k}^{-1}(y)))))))) \\ &= \mathcal{P}_{n-k}(\lambda(\pi_2((\mathcal{P}_k^{-1}(x)), \lambda^{-1}(\mathcal{P}_{n-k}^{-1}(y)))))) \\ &= \mathcal{P}_{n-k}(\lambda(\lambda^{-1}(\mathcal{P}_{n-k}^{-1}(y)))) \\ &= \mathcal{P}_{n-k}(\mathcal{P}_{n-k}^{-1}(y)) \\ &= y \end{aligned}$$

so that using the above

$$\begin{aligned} \beta(z) &= (\mathcal{P}_k(\pi_1(\alpha(\mathcal{P}_n^{-1}(z)))), \mathcal{P}_{n-k}(\lambda(\pi_2(\alpha(\mathcal{P}_n^{-1}(z)))))) \\ &= (x, y) \end{aligned}$$

proving that  $\beta$  is surjective.  $\square$

Next we prove that  $\bigotimes_{i \in \{1, \dots, n\}} A_i$  has the same properties as  $\prod_{i \in \{1, \dots, n\}} A_i$  (more specific 2.90, 2.95, 2.97, 2.99, 2.100, 2.101)

**Theorem 5.95.** Let  $n \in \mathbb{N}_0$ ,  $\{A_i\}_{i \in \{1, \dots, n\}}$ ,  $\{B_i\}_{i \in \{1, \dots, n\}}$  a finite family of sets with  $\forall i \in \{1, \dots, n\}$   $A_i \neq \emptyset$  (which by the previous theorem means that  $\bigotimes_{i \in \{1, \dots, n\}} A_i \neq \emptyset$ ) then  $\bigotimes_{i \in \{1, \dots, n\}} A_i \subseteq \bigotimes_{i \in \{1, \dots, n\}} B_i \Leftrightarrow A_i \subseteq B_i$

**Proof.** We prove this by induction, so let  $\mathcal{S} = \{n \in \mathbb{N}_0 \mid \{A_i\}_{i \in \{1, \dots, n\}}, \{B_i\}_{i \in \{1, \dots, n\}} \text{ with } \exists i \in \{1, \dots, n\} \models A_i \neq \emptyset \text{ then } \bigotimes_{i \in \{1, \dots, n\}} A_i \subseteq \bigotimes_{i \in \{1, \dots, n\}} B_i \Leftrightarrow A_i \subseteq B_i\}$  then we have

**1  $\in \mathcal{S}$ .** As  $\bigotimes_{i \in \{1, \dots, 1\}} A_i = A_1$  and  $\bigotimes_{i \in \{1, \dots, 1\}} B_i = B_1$  we have  $\bigotimes_{i \in \{1, \dots, 1\}} A_i \subseteq \bigotimes_{i \in \{1, \dots, 1\}} B_i$  if and only  $A_1 \subseteq B_1$  proving that  $1 \in \mathcal{S}$

**$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$ .** Take  $\bigotimes_{i \in \{1, \dots, n+1\}} A_i, \bigotimes_{i \in \{1, \dots, n+1\}} B_i$  then

$\Rightarrow$ . If  $\forall i \in \{1, \dots, n+1\}$  we have that  $A_i \neq \emptyset$  we have using the theorem 5.90 that

$$\bigotimes_{i \in \{1, \dots, n\}} A_i \neq \emptyset \wedge A_{n+1} \neq \emptyset \quad (5.9)$$

$$\begin{aligned} \bigotimes_{i \in \{1, \dots, n+1\}} A_i \subseteq \bigotimes_{i \in \{1, \dots, n+1\}} B_i &\stackrel{1 \leq n+1}{\Rightarrow} \left( \bigotimes_{i \in \{1, \dots, n\}} A_i \right) \times A_{n+1} \subseteq \\ &\quad \left( \bigotimes_{i \in \{1, \dots, n\}} B_i \right) \times B_{n+1} \\ &\stackrel{5.9 \text{ and } 1.42}{\Rightarrow} \bigotimes_{i \in \{1, \dots, n\}} A_i \subseteq \bigotimes_{i \in \{1, \dots, n\}} B_i \wedge A_{n+1} \subseteq \\ &\quad B_{n+1} \\ &\stackrel{n \in \mathcal{S}}{\Rightarrow} (\forall i \in \{1, \dots, n\} \models A_i \subseteq B_i) \wedge A_{n+1} \subseteq B_{n+1} \\ &\Rightarrow \forall i \in \{1, \dots, n+1\} \models A_i \subseteq B_i \end{aligned}$$

$\Leftarrow$ . Using 5.84 we have if  $\forall i \in \{1, \dots, n+1\} \models A_i \subseteq B_i$  then  $\bigotimes_{i \in \{1, \dots, n+1\}} A_i \subseteq \bigotimes_{i \in \{1, \dots, n+1\}} B_i$ .

So we have proved that  $\bigotimes_{i \in \{1, \dots, n+1\}} A_i \subseteq \bigotimes_{i \in \{1, \dots, n+1\}} B_i \Leftrightarrow \forall i \in \{1, \dots, n+1\} \models A_i \subseteq B_i$  from which we conclude that  $n+1 \in \mathcal{S}$   $\square$

**Corollary 5.96.** Let  $n \in \mathbb{N}_0$ ,  $\{A_i\}_{i \in \{1, \dots, n\}}$ ,  $\{B_i\}_{i \in \{1, \dots, n\}}$  finite families of sets with  $\forall i \in \{1, \dots, n\}$   $A_i \neq \emptyset$  then  $\bigotimes_{i \in \{1, \dots, n\}} A_i = \bigotimes_{i \in \{1, \dots, n\}} B_i \Leftrightarrow \forall i \in \{1, \dots, n\}$  we have  $A_i = B_i$

**Proof.**

$\Rightarrow$ . As  $\forall i \in \{1, \dots, n\} A_i \neq \emptyset$  we have by 5.90 that  $\emptyset \neq \bigotimes_{i \in \{1, \dots, n\}} A_i = \bigotimes_{i \in \{1, \dots, n\}} B_i$  or applying 5.90 again we have  $\forall i \in \{1, \dots, n\} B_i \neq \emptyset$ . Then we have

$$\begin{aligned} \bigotimes_{i \in \{1, \dots, n\}} A_i = \bigotimes_{i \in \{1, \dots, n\}} B_i &\Rightarrow \bigotimes_{i \in \{1, \dots, n\}} A_i \subseteq \bigotimes_{i \in \{1, \dots, n\}} B_i \\ &\stackrel{5.95}{\Rightarrow} \forall i \in \{1, \dots, n\} \models A_i \subseteq B_i \end{aligned}$$

and similar

$$\begin{aligned} \bigotimes_{i \in \{1, \dots, n\}} A_i = \bigotimes_{i \in \{1, \dots, n\}} B_i &\Rightarrow \bigotimes_{i \in \{1, \dots, n\}} B_i \subseteq \bigotimes_{i \in \{1, \dots, n\}} A_i \\ &\stackrel{5.95}{\Rightarrow} \forall i \in \{1, \dots, n\} \models B_i \subseteq A_i \end{aligned}$$

so we have  $\forall i \in \{1, \dots, n\}$  that  $A_i = B_i$

$\Leftarrow$ . This is trivial  $\square$

**Theorem 5.97.** Let  $n \in \mathbb{N}_0$ ,  $\{A_i\}_{i \in \{1, \dots, n\}}$ ,  $\{B_i\}_{i \in \{1, \dots, n\}}$  finite families of sets then  $(\bigotimes_{i \in \{1, \dots, n\}} A_i) \cap (\bigotimes_{i \in \{1, \dots, n\}} B_i) = \bigotimes_{i \in \{1, \dots, n\}} (A_i \cap B_i)$

**Proof.** First as  $\forall i \in \{1, \dots, n\}$  we have that  $A_i \cap B_i \subseteq A_i, B_i$  we have using 5.84 that  $\bigotimes_{i \in \{1, \dots, n\}} (A_i \cap B_i) \subseteq \bigotimes_{i \in \{1, \dots, n\}} A_i, \bigotimes_{i \in \{1, \dots, n\}} B_i$  proving that

$$\bigotimes_{i \in \{1, \dots, n\}} (A_i \cap B_i) \subseteq \left( \bigotimes_{i \in \{1, \dots, n\}} A_i \right) \cap \left( \bigotimes_{i \in \{1, \dots, n\}} B_i \right) \quad (5.10)$$

For the opposite inclusion let  $x \in (\bigotimes_{i \in \{1, \dots, n\}} A_i) \cap (\bigotimes_{i \in \{1, \dots, n\}} B_i)$  then  $x \in \bigotimes_{i \in \{1, \dots, n\}} A_i \wedge x \in \bigotimes_{i \in \{1, \dots, n\}} B_i$ . Using 5.93 (3) we have that  $\forall i \in \{1, \dots, n\} \pi_i^\times(x) \in A_i$  and  $\pi_i^\times(x) \in B_i$  so that  $\pi_i^\times(x) \in A_i \cap B_i$ . Finally using 5.93 (3) again we have  $(\bigotimes_{i \in \{1, \dots, n\}} A_i) \cap (\bigotimes_{i \in \{1, \dots, n\}} B_i) \subseteq \bigotimes_{i \in \{1, \dots, n\}} (A_i \cap B_i)$  which together with 5.10 gives

$$\left( \bigotimes_{i \in \{1, \dots, n\}} A_i \right) \cap \left( \bigotimes_{i \in \{1, \dots, n\}} B_i \right) = \bigotimes_{i \in \{1, \dots, n\}} (A_i \cap B_i)$$

□

Actually we can generalize the above to a arbitrary collection of family of sets with almost the same proof.

**Theorem 5.98.** Let  $n \in \mathbb{N}_0$ ,  $I$  be a set, and  $\{\{A_{i,j}\}_{j \in \{1, \dots, n\}}\}_{i \in I}$  a family of finite families of sets then  $\bigcap_{i \in I} (\bigotimes_{j \in \{1, \dots, n\}} A_{i,j}) = \bigotimes_{j \in \{1, \dots, n\}} (\bigcap_{i \in I} A_{i,j})$

**Proof.** First  $\forall i \in I$  we have  $\forall j \in \{1, \dots, n\}$  that  $\bigcap_{k \in I} A_{k,j} \subseteq A_{i,j}$  so using 5.84 it follows that  $\bigotimes_{j \in \{1, \dots, n\}} (\bigcap_{k \in I} A_{k,j}) \subseteq \bigotimes_{k \in \{1, \dots, n\}} A_{i,k}$  proving that

$$\bigotimes_{j \in \{1, \dots, n\}} \left( \bigcap_{i \in I} A_{i,j} \right) \subseteq \bigcap_{i \in I} \left( \bigotimes_{j \in \{1, \dots, n\}} A_{i,j} \right) \quad (5.11)$$

For the opposite inclusion let  $x \in \bigcap_{i \in I} (\bigotimes_{j \in \{1, \dots, n\}} A_{i,j})$  then  $\forall i \in I x \in \bigotimes_{j \in \{1, \dots, n\}} A_{i,j}$ , using 5.93 (3) we have that  $\forall j \in \{1, \dots, n\} \pi_j^\times(x) \in A_{i,j}$ , so that  $\pi_j^\times(x) \in \bigcap_{i \in I} A_{i,j}$ . Finally using 5.93 (3) again we have  $\bigcap_{i \in I} (\bigotimes_{j \in \{1, \dots, n\}} A_{i,j}) \subseteq \bigotimes_{j \in \{1, \dots, n\}} (\bigcap_{i \in I} A_{i,j})$  which together with 5.11 gives

$$\bigcap_{i \in I} \left( \bigotimes_{j \in \{1, \dots, n\}} A_{i,j} \right) = \bigotimes_{j \in \{1, \dots, n\}} \left( \bigcap_{i \in I} A_{i,j} \right)$$

□

**Theorem 5.99.** Let  $n \in \mathbb{N}_0$ ,  $\{A_i\}_{i \in \{1, \dots, n\}}$ ,  $\{B_i\}_{i \in \{1, \dots, n\}}$  be finite family of sets then  $(\bigotimes_{i \in \{1, \dots, n\}} A_i) \cup (\bigotimes_{i \in \{1, \dots, n\}} B_i) \subseteq \bigotimes_{i \in \{1, \dots, n\}} (A_i \cup B_i)$

**Proof.** As  $\forall i \in \mathbb{N}_0$  we have  $A_i \subseteq A_i \cup B_i$  and  $B_i \subseteq A_i \cup B_i$  we have by 5.84 that  $\bigotimes_{i \in \{1, \dots, n\}} A_i \subseteq \bigotimes_{i \in \{1, \dots, n\}} (A_i \cup B_i)$  and  $\bigotimes_{i \in \{1, \dots, n\}} B_i \subseteq \bigotimes_{i \in \{1, \dots, n\}} (A_i \cup B_i)$  so that  $(\bigotimes_{i \in \{1, \dots, n\}} A_i) \cup (\bigotimes_{i \in \{1, \dots, n\}} B_i) \subseteq \bigotimes_{i \in \{1, \dots, n\}} (A_i \cup B_i)$

□

**Note 5.100.** The opposite inclusion is not true for example if  $A_1 = \{1\}$ ,  $A_2 = \{2\}$ ,  $B_1 = \{3\}$  and  $B_2 = \{4\}$  then  $(A_1 \times A_2) \cup (B_1 \times B_2) = \{(1, 2), (3, 4)\}$  but  $(A_1 \cup B_1) \times (A_2 \cup B_2) = \{(1, 3), (1, 4), (2, 3), (2, 4)\}$

### 5.6.1 Finite product of sets and denumerability and countability

**Theorem 5.101.** Let  $\{A_i\}_{i \in \{1, \dots, n\}}$  be a finite family of sets then we have

1. If  $\forall i \in \{1, \dots, n\} \models A_i$  is finite then  $\bigotimes_{i \in \{1, \dots, n\}} A_i$  is finite and  $\#(\bigotimes_{i \in \{1, \dots, n\}} A_i) = \prod_{i=1}^n \#(A_i)$

2. If  $\forall i \in \{1, \dots, n\} \models A_i$  is denumerable then  $\bigotimes_{i \in \{1, \dots, n\}} A_i$  is denumerable

**Proof.** We prove this by induction

1. Define  $\mathcal{S} = \{n \in \mathbb{N}_0 \mid \text{If } \{A_i\}_{\{1, \dots, n\}} \text{ is a finite family of finite sets then } \bigotimes_{i \in \{1, \dots, n\}} A_i \text{ is finite and } \#(\bigotimes_{i \in \{1, \dots, n\}} A_i) = \prod_{i=1}^n \#(A_i)\}$  then we have

**1  $\in \mathcal{S}$ .** If  $\{A_i\}_{i \in \{1, \dots, 1\}}$  is finite family of finite sets then  $\bigotimes_{i \in \{1, \dots, 1\}} A_i = A_1$  is finite and  $\prod_{i=1}^1 \#(A_i) = \#(A_1) = \#(\bigotimes_{i \in \{1, \dots, 1\}} A_i)$  proving that  $1 \in \mathcal{S}$

**$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$ .** If  $\{A_i\}_{i \in \{1, \dots, n+1\}}$  is a finite family of finite sets then as  $n \in \mathcal{S}$  we have that  $\bigotimes_{i \in \{1, \dots, n\}} A_i$  is finite and  $\#(\bigotimes_{i \in \{1, \dots, n\}} A_i) = \prod_{i=1}^n \#(A_i)$ .

Now  $\bigotimes_{i \in \{1, \dots, n+1\}} A_i = (\bigotimes_{i \in \{1, \dots, n\}} A_i) \otimes A_{n+1}$  which is by 5.44 finite and  $\#(\bigotimes_{i \in \{1, \dots, n+1\}} A_i) = \#((\bigotimes_{i \in \{1, \dots, n\}} A_i) \otimes A_{n+1}) = \#(\bigotimes_{i \in \{1, \dots, n\}} A_i) \cdot \#(A_{n+1}) = (\prod_{i=1}^n \#(A_i)) \cdot \#(A_{n+1}) = \prod_{i=1}^{n+1} \#(A_i)$  showing that  $n+1 \in \mathcal{S}$

2. Define  $\mathcal{S} = \{n \in \mathbb{N}_0 \mid \text{If } \{A_i\}_{\{1, \dots, n\}}$  is a finite family of denumerable sets then  $\bigotimes_{i \in \{1, \dots, n\}} A_i$  is denumerable} then we have

**1  $\in \mathcal{S}$ .** If  $\{A_i\}_{i \in \{1, \dots, 1\}}$  is finite family of denumerable sets then  $\bigotimes_{i \in \{1, \dots, 1\}} A_i = A_1$  is denumerable proving that  $1 \in \mathcal{S}$

**$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$ .** If  $\{A_i\}_{i \in \{1, \dots, n+1\}}$  is a finite family of denumerable sets then as  $n \in \mathcal{S}$  we have that  $\bigotimes_{i \in \{1, \dots, n\}} A_i$  is denumerable. Now  $\bigotimes_{i \in \{1, \dots, n+1\}} A_i = (\bigotimes_{i \in \{1, \dots, n\}} A_i) \otimes A_{n+1}$  which is by 5.59 denumerable showing that  $n+1 \in \mathcal{S}$   $\square$

**Theorem 5.102.** Let  $\{A_i\}_{i \in \{1, \dots, n\}}$  be a finite family of sets then we have

1. If  $\forall i \in \{1, \dots, n\} \models A_i$  is finite then  $\prod_{i \in \{1, \dots, n\}} A_i$  is finite and  $\#(\prod_{i \in \{1, \dots, n\}} A_i) = \prod_{i=1}^n \#(A_i)$
2. If  $\forall i \in \{1, \dots, n\} \models A_i$  is denumerable then  $\prod_{i \in \{1, \dots, n\}} A_i$  is denumerable

**Proof.**

1. Let  $\{A_i\}_{i \in \{1, \dots, n\}}$  be a finite family of finite sets then using the previous lemma we have that  $\bigotimes_{i \in \{1, \dots, n\}} A_i$  is finite with  $\#(\bigotimes_{i \in \{1, \dots, n\}} A_i) = \prod_{i=1}^n \#(A_i)$ . As by 5.89  $\bigotimes_{i \in \{1, \dots, n\}} A_i$  is bijective to  $\prod_{i \in \{1, \dots, n\}} A_i$  we have that  $\prod_{i \in \{1, \dots, n\}} A_i$  is finite and  $\#(\prod_{i \in \{1, \dots, n\}} A_i) = \prod_{i=1}^n \#(A_i)$ .
2. Let  $\{A_i\}_{i \in \{1, \dots, n\}}$  be a finite family of denumerable sets then using the previous lemma we have that  $\bigotimes_{i \in \{1, \dots, n\}} A_i$  is denumerable. As by 5.89  $\bigotimes_{i \in \{1, \dots, n\}} A_i$  is bijective to  $\prod_{i \in \{1, \dots, n\}} A_i$  we have that  $\prod_{i \in \{1, \dots, n\}} A_i$  is denumerable.  $\square$

### 5.6.2 Notation conventions

In the above we have proved that  $\prod_{i \in I} A_i$  and  $\bigotimes_{i \in \{1, \dots, n\}} A_i$  have similar properties and also that if  $I = \{1, \dots, n\}$  we can set up a bijection between  $\prod_{i \in \{1, \dots, n\}} A_i$  and  $\bigotimes_{i \in \{1, \dots, n\}} A_i$ .  $\prod_{i \in I} A_i$  is of course the most general definition because it also work for non finite families of sets. However there is not a easy way to use a induction argument if  $I = \{1, \dots, n\}$  as  $\prod_{i \in \{1, \dots, n\}} A_i$  is not defined using recursion, this is the big benefit of using  $\bigotimes_{i \in \{1, \dots, n\}} A_i$ . Of course we can use the bijection  $\mathcal{P}_n: \prod_{i \in \{1, \dots, n\}} A_i \rightarrow \bigotimes_{i \in \{1, \dots, n\}} A_i$  to prove a theorem in  $\bigotimes_{i \in \{1, \dots, n\}} A_i$  and then the inverse bijection to state the theorem for  $\prod_{i \in \{1, \dots, n\}} A_i$  but this a little cumbersome, so in most cases we work in the finite case only with  $\bigotimes_{i \in \{1, \dots, n\}} A_i$ . To avoid double notation and also to be consistent with existing literature we introduce the following notation conventions.

**Notation 5.103. (Conventions for product of sets)**

1. Unless mentioned otherwise we use the symbol  $\prod_{i \in \{1, \dots, n\}} A_i$  to mean  $\bigotimes_{i \in \{1, \dots, n\}} A_i$  and if  $I$  is not finite  $\prod_{i \in I} A_i$  is equal to the generalized product.
2. For  $\prod_{i \in \{1, \dots, n\}} A_i$  [which is using (1)  $\bigotimes_{i \in \{1, \dots, n\}} A_i$ ] we use  $\pi_i$  to mean  $\pi_i^\times$
3. For  $\prod_{i \in I} A_i$  [the general not finite case]  $\pi_i$  is the general projection function (see 2.91)
4. If  $x \in \prod_{i \in \{1, \dots, n\}} A_i$  [which is using (1)  $\bigotimes_{i \in \{1, \dots, n\}} A_i$ ] then as by 5.93 (1)  $x$  is uniquely defined by the values of  $\pi_i^\times(x) \stackrel{(2)}{=} \pi_i(x)$  we note  $x_i$  instead of  $\pi_i(x) \stackrel{(3)}{=} \pi_i^\times(x)$ . Further we use the notation  $x = (x_1, \dots, x_n)$  to mean that  $\forall i \in \{1, \dots, n\} \pi_i(x) = x_i$ .
5. Using the above notation we have that  $\prod_{i \in \{1, \dots, 1\}} A_i = A_1$  and if  $\forall x \in \prod_{i \in \{1, \dots, 1\}} A_i = A_1$  we have  $x_1 = \pi_1^\times(x) \stackrel{5.92}{=} x$
6. If  $x \in \prod_{i \in I} A_i$  [the general not finite case] then we use the notation referenced in 2.88 where  $x_i$  means  $x(i)$
7. Let  $I, A$  be sets then  $A^I = \prod_{i \in I} C_i$  where  $\{C_i\}_{i \in I}$  is defined by  $C_i = A$  (this is actually a theorem see 2.86)
8. If  $n \in \mathbb{N}_0$  then and  $A$  a set then  $A^n = \prod_{i \in \{1, \dots, n\}} C_i \stackrel{(1)}{=} \bigotimes_{i \in \{1, \dots, n\}} C_i$  where  $\{C_i\}_{i \in \{1, \dots, n\}}$  is define by  $C_i = A$

Using the above theorem we can restate the basic theorems about product of sets

1. Let  $I$  be a set and  $\{A_i\}_{i \in I}$  a family of **non empty** sets then  $\forall i \in I \pi_i: \prod_{i \in I} A_i \rightarrow A_i$  is a surjection (see 2.93)
2. Let  $n \in \mathbb{N}_0$ ,  $\{A_i\}_{i \in \{1, \dots, n\}}$  a finite family of non empty sets then  $\forall i \in \{1, \dots, n\} \pi_i: \prod_{i \in \{1, \dots, n\}} A_i \rightarrow A_i$  is a surjection (see 5.93 (2))
3. Let  $n \in \mathbb{N}_0$ ,  $\{A_i\}_{i \in \{1, \dots, n+1\}}$  a finite family of sets then  $\prod_{i \in \{1, \dots, n+1\}} A_i = (\prod_{i \in \{1, \dots, n\}} A_i) \times A_{n+1}$  and  $\forall (x_1, \dots, x_n) \in \prod_{i \in \{1, \dots, n\}} A_i$  and  $x \in A_{n+1}$  we have that  $((x_1, \dots, x_n), x) = (x_1, \dots, x_{n+1})$  (see 5.82)
4. Let  $n, k \in \mathbb{N}_0$  with  $1 \leq k < n$  and  $\{A_i\}_{i \in \{1, \dots, n\}}$  then there exists a bijection  $\beta: \prod_{i \in \{1, \dots, n\}} A_i \rightarrow (\prod_{i \in \{1, \dots, k\}} A_i) \times (\prod_{i \in \{1, \dots, n-k\}} A_{i+k})$  such that  $\beta((x_1, \dots, x_n)) = ((x_1, \dots, x_k), (x_{k+1}, \dots, x_n))$
5. Let  $I$  be a set and  $\{A_i\}_{i \in I}$ ,  $\{B_i\}_{i \in I}$  two families of sets such that  $\forall x \in \prod_{i \in I} A_i$  we have that  $\forall i \in I x_i \in B_i$  then  $\prod_{i \in I} A_i \subseteq \prod_{i \in I} B_i$  (see 2.92)
6. Let  $n \in \mathbb{N}_0$  and  $\{A_i\}_{i \in \{1, \dots, n\}}$ ,  $\{B_i\}_{i \in \{1, \dots, n\}}$  two finite families of sets such that  $\forall x \in \prod_{i \in \{1, \dots, n\}} A_i$  we have that  $\forall i \in \{1, \dots, n\} x_i \in B_i$  then  $\prod_{i \in \{1, \dots, n\}} A_i \subseteq \prod_{i \in \{1, \dots, n\}} B_i$  (see 5.93 (3))
7. Let  $I$  be a set and  $\{A_i\}_{i \in I}$  a family of sets then  $\prod_{i \in I} A_i = \emptyset \Leftrightarrow \exists i \in I$  such that  $A_i = \emptyset$  (see 2.90)
8. Let  $n \in \mathbb{N}_0$  and  $\{A_i\}_{i \in I}$  a finite family of sets then  $\prod_{i \in \{1, \dots, n\}} A_i = \emptyset \Leftrightarrow \exists i \in \{1, \dots, n\}$  such that  $A_i = \emptyset$  (see 5.90)
9. Let  $I$  be a set,  $\{A_i\}_{i \in I}$ ,  $\{B_i\}_{i \in I}$  two families of sets then we have
  - a. If  $\forall i \in I$  we have  $A_i \subseteq B_i$  then  $\prod_{i \in I} A_i \subseteq \prod_{i \in I} B_i$  (see 2.89)
  - b. If  $\forall i \in I$  we have  $A_i \neq \emptyset$  then  $\prod_{i \in I} A_i \subseteq \prod_{i \in I} B_i \Leftrightarrow \forall i \in I \models A_i \subseteq B_i$  (see 2.95)
10. Let  $n \in \mathbb{N}_0$ ,  $\{A_i\}_{i \in \{1, \dots, n\}}$ ,  $\{B_i\}_{i \in \{1, \dots, n\}}$  two finite families of sets then we have
  - a. If  $\forall i \in \{1, \dots, n\}$  we have  $A_i \subseteq B_i$  then  $\prod_{i \in \{1, \dots, n\}} A_i \subseteq \prod_{i \in \{1, \dots, n\}} B_i$  (see 5.84)
  - b. If  $\forall i \in \{1, \dots, n\}$  we have  $A_i \neq \emptyset$  then  $\prod_{i \in \{1, \dots, n\}} A_i \subseteq \prod_{i \in \{1, \dots, n\}} B_i \Leftrightarrow \forall i \in \{1, \dots, n\} \models A_i \subseteq B_i$  (see 5.95)
11. Let  $I, J$  be sets and  $\{\{A_{i,j}\}_{j \in J}\}_{i \in I}$  a family of families of sets then  $\bigcap_{i \in I} (\prod_{j \in J} A_{i,j}) = \prod_{j \in J} (\bigcap_{i \in I} A_{i,j})$  (see 2.100)

12. Let  $I$  be a set,  $n \in \mathbb{N}_0$  and  $\{\{A_{i,j}\}_{j \in \{1, \dots, n\}}\}_{i \in I}$  a family of finite families of sets then  $\bigcap_{i \in I} (\prod_{j \in \{1, \dots, n\}} A_{i,j}) = \prod_{j \in \{1, \dots, n\}} (\bigcap_{i \in I} A_{i,j})$  (see 5.98)
13. If  $I$  is a set and  $\{A_i\}_{i \in I}$ ,  $\{B_i\}_{i \in I}$  families of sets then  $(\prod_{i \in I} A_i) \cup (\prod_{i \in I} B_i) \subseteq \prod_{i \in I} (A_i \cup B_i)$  (see 2.101)
14. If  $n \in \mathbb{N}_0$  and  $\{A_i\}_{i \in \{1, \dots, n\}}$ ,  $\{B_i\}_{i \in \{1, \dots, n\}}$  finite families of sets then  $(\prod_{i \in \{1, \dots, n\}} A_i) \cup (\prod_{i \in \{1, \dots, n\}} B_i) \subseteq \prod_{i \in \{1, \dots, n\}} (A_i \cup B_i)$  (see 5.99)

We have the following corollary for (8) and (9) above

**Corollary 5.104.** *Equality for a product of sets*

1. If  $I$  is a set and  $\{A_i\}_{i \in I}$ ,  $\{B_i\}_{i \in I}$  be two family of sets such that  $\forall i \in I A_i \neq \emptyset$  then  $\prod_{i \in I} A_i = \prod_{i \in I} B_i \Leftrightarrow \forall i \in I$  we have  $A_i = B_i$
2. If  $n \in \mathbb{N}_0$  and  $\{A_i\}_{i \in \{1, \dots, n\}}$ ,  $\{B_i\}_{i \in \{1, \dots, n\}}$  be two family of sets such that  $\forall i \in \{1, \dots, n\} A_i \neq \emptyset$  then  $\prod_{i \in \{1, \dots, n\}} A_i = \prod_{i \in \{1, \dots, n\}} B_i \Leftrightarrow \forall i \in I$  we have  $A_i = B_i$

**Proof.**

1.

$\Rightarrow$ . As  $\forall i \in I$  we have  $A_i \neq \emptyset$  we have by 5.103 (6) that  $\prod_{i \in I} A_i \neq \emptyset$  hence as  $\prod_{i \in I} A_i = \prod_{i \in I} B_i$  we have by 5.103 (6) that  $\forall i \in I B_i \neq \emptyset$ . From  $\prod_{i \in I} A_i = \prod_{i \in I} B_i$  we have  $\prod_{i \in I} A_i \subseteq \prod_{i \in I} B_i$   $\stackrel{5.103 (8)}{\Rightarrow} \forall i \in I A_i \subseteq B_i$  and  $\prod_{i \in I} B_i \subseteq \prod_{i \in I} A_i$   $\stackrel{5.103 (8)}{\Rightarrow} \forall i \in I B_i \subseteq A_i$ . So we conclude that  $\forall i \in I A_i = B_i$ .

$\Leftarrow$ . This is trivial

2.

$\Rightarrow$ . As  $\forall i \in \{1, \dots, n\}$  we have  $A_i \neq \emptyset$  we have by 5.103 (6) that  $\prod_{i \in \{1, \dots, n\}} A_i \neq \emptyset$  hence as  $\prod_{i \in \{1, \dots, n\}} A_i = \prod_{i \in \{1, \dots, n\}} B_i$  we have by 5.103 (6) that  $\forall i \in \{1, \dots, n\} B_i \neq \emptyset$ . From  $\prod_{i \in \{1, \dots, n\}} A_i = \prod_{i \in \{1, \dots, n\}} B_i$  we have  $\prod_{i \in \{1, \dots, n\}} A_i \subseteq \prod_{i \in \{1, \dots, n\}} B_i$   $\stackrel{5.103 (8)}{\Rightarrow} \forall i \in \{1, \dots, n\} A_i \subseteq B_i$  and  $\prod_{i \in \{1, \dots, n\}} B_i \subseteq \prod_{i \in \{1, \dots, n\}} A_i$   $\stackrel{5.103 (8)}{\Rightarrow} \forall i \in \{1, \dots, n\} B_i \subseteq A_i$ . So we conclude that  $\forall i \in \{1, \dots, n\} A_i = B_i$ .

$\Leftarrow$ . This is trivial □

**Theorem 5.105.** Let  $n, m \in \mathbb{N}_0$  and  $\{A_i\}_{i \in \{1, \dots, m\}}$  a family of **non empty** sets such that

$$\forall i \in \{1, \dots, m\} \text{ there exists a } \{B_{i,j}\}_{j \in \{1, \dots, n\}} \text{ such that } A_i = \prod_{j \in \{1, \dots, n\}} B_{i,j}$$

$$\exists \{C_i\}_{i \in \{1, \dots, n\}} \text{ such that } \bigcup_{i \in \{1, \dots, m\}} A_i = \prod_{i \in \{1, \dots, n\}} C_i$$

then we have

1.  $\bigcup_{i \in \{1, \dots, m\}} A_i = \prod_{j \in \{1, \dots, n\}} (\bigcup_{i \in \{1, \dots, m\}} B_{i,j})$
2.  $\forall i \in \{1, \dots, n\} C_i = \bigcup_{j \in \{1, \dots, m\}} B_{j,i}$

**Proof.** First

$$\begin{aligned} x = (x_1, \dots, x_n) \in \bigcup_{i \in \{1, \dots, m\}} A_i &\Rightarrow \exists i \in \{1, \dots, m\} \vdash (x_1, \dots, x_n) \in A_i \\ &\Rightarrow \forall j \in \{1, \dots, n\} \vdash x_j \in B_{i,j} \subseteq \bigcup_{k \in \{1, \dots, m\}} B_{k,j} \\ &\Rightarrow (x_1, \dots, x_n) \in \prod_{j \in \{1, \dots, n\}} \left( \bigcup_{k \in \{1, \dots, m\}} B_{k,j} \right) \end{aligned}$$

proving

$$\bigcup_{i \in \{1, \dots, m\}} A_i \subseteq \prod_{j \in \{1, \dots, n\}} \left( \bigcup_{k \in \{1, \dots, m\}} B_{k,j} \right) \quad (5.12)$$

Second as  $\forall i \in \{1, \dots, m\} \ \emptyset \neq A_i$  it follows from 5.103 (7) that

$$\forall j \in \{1, \dots, n\} \models B_{i,j} \neq \emptyset \quad (5.13)$$

Given  $i \in \{1, \dots, m\}$  take  $j \in \{1, \dots, n\}$  and  $x_j \in B_{i,j}$  then by the above there exists  $\forall k \in \{1, \dots, n\} \setminus \{j\}$  a  $x_k \in B_{i,k}$  so that  $(x_1, \dots, x_k) \in \prod_{l \in \{1, \dots, n\}} B_{l,j} = A_i \subseteq \bigcup_{i \in \{1, \dots, m\}} A_i = \prod_{l \in \{1, \dots, n\}} C_l$  proving that  $x_j \in C_j$  hence

$$\forall j \in \{1, \dots, n\} \models x_j \in B_{i,j} \Rightarrow x_j \in C_j \quad (5.14)$$

Next

$$\begin{aligned} x = (x_1, \dots, x_n) \in \prod_{j \in \{1, \dots, n\}} \left( \bigcup_{k \in \{1, \dots, m\}} B_{k,j} \right) &\Rightarrow \forall j \in \{1, \dots, n\} \models x_j \in \bigcup_{k \in \{1, \dots, m\}} B_{k,j} \\ &\Rightarrow \exists k \in \{1, \dots, n\} \models x_j \in B_{k,j} \\ &\stackrel{5.14}{\Rightarrow} x_j \in C_j \\ &\Rightarrow (x_1, \dots, x_n) \in \prod_{j \in \{1, \dots, n\}} C_j = \bigcup_{i \in \{1, \dots, m\}} A_i \end{aligned}$$

proving that  $\prod_{j \in \{1, \dots, n\}} (\bigcup_{k \in \{1, \dots, m\}} B_{k,j}) \subseteq \bigcup_{i \in \{1, \dots, m\}} A_i$  and using 5.12 this gives

$$\prod_{j \in \{1, \dots, n\}} \left( \bigcup_{k \in \{1, \dots, m\}} B_{k,j} \right) = \bigcup_{i \in \{1, \dots, m\}} A_i$$

proving (1) of the lemma. Finally (2) follows from 5.104 □



# Chapter 6

## The Integer numbers

### 6.1 Definition and arithmetic's

The problem with the natural numbers is that we have no inverse number for every natural number, meaning that  $\mathbb{N}$  can not be a group. In other words given  $n \in \mathbb{N}$  with  $n \neq 0$  there does not exist a  $n'$  such that  $n + n' = 0$ . This is solved by introducing the whole numbers.

**Definition 6.1.** We define the relation  $\sim$  on  $\mathbb{N} \times \mathbb{N}$  by  $\sim = \{( (n, m), (n', m') ) \in (\mathbb{N} \times \mathbb{N}) \times (\mathbb{N} \times \mathbb{N}) \mid n + m' = m + n' \}$

**Theorem 6.2.**  $\sim$  is a equivalence relation on  $\mathbb{N}$

**Proof.**

1. **(reflectivity)** If  $(n, m) \in \mathbb{N} \times \mathbb{N}$  then  $n + m \stackrel{4.33}{=} m + n \Rightarrow (n, m) \sim (n, m)$
2. **(Symmetry)** If  $(n, m) \sim (n', m') \Rightarrow n + m' = m + n' \stackrel{4.33}{\Rightarrow} m' + n = n' + m \Rightarrow n' + m = m' + n \Rightarrow (n', m') \sim (n, m)$
3. **(Transitivity)** If  $(n, m) \sim (n', m')$  and  $(n', m') \sim (n'', m'') \Rightarrow n + m' = m + n'$  and  $n' + m'' = m' + n'' \Rightarrow (n + m') + (n' + m'') = (m + n') + (m' + n'') \stackrel{4.33, 4.32}{\Rightarrow} (n + m'') + (m' + n') = (m + n'') + (m' + n') \stackrel{4.44}{\Rightarrow} n + m'' = m + n'' \Rightarrow (n, m) \sim (n'', m'')$   $\square$

**Definition 6.3.** We define the set of integer numbers  $\mathbb{Z}$  by  $\mathbb{Z} = \mathbb{N} / \sim$  (see 2.123)

So we have that  $\mathbb{Z} = \{ \sim[(n, m)] \mid (n, m) \in \mathbb{N} \times \mathbb{N} \}$

**Theorem 6.4.** If  $\sim[(n, m)] \in \mathbb{Z}$  then if  $k \in \mathbb{N}$  we have  $\sim[(n, m)] = \sim[(n+k, m+k)]$

**Proof.**  $n + (m+k) = (n+m) + k = (m+n) + k = m + (n+k) \Rightarrow (n, m) \sim (n+k, m+k) \Rightarrow \sim[(n, m)] = \sim[(n+k, m+k)]$   $\square$

**Theorem 6.5.**  $+: \mathbb{Z} \times \mathbb{Z} \rightarrow \mathbb{Z}$  defined by  $(\sim[(n, m)], \sim[r, s]) \rightarrow \sim[(n, m)] + \sim[(r, s)] = \sim[(n+r, m+s)]$  is a function. Note that we have here two additions + one defined in  $\mathbb{N}$  the other in  $\mathbb{Z}$ .

**Proof.** We must prove that if  $\sim[(n, m)] = \sim[(n', m')]$  and  $\sim[(r, s)] = \sim[(r', s')]$  then  $\sim[(n+r, m+s)] = \sim[(n'+r', m'+s')]$ . Now from  $\sim[(n, m)] = \sim[(n', m')]$  and  $\sim[(r, s)] = \sim[(r', s')]$  we have by 2.120 that  $(n, m) \sim (n', m')$  and  $(r, s) \sim (r', s') \Rightarrow n + m' = m + n'$  and  $r + s' = s + r'$ . So  $(n+r) + (m'+s') = (n+m') + (r+s') = (m+n') + (s+r') = (m+s) + (n'+r') \Rightarrow (n+r, m+s) \sim (n'+r', m'+s') \Rightarrow \sim[(n+r, m+s)] = \sim[(n'+r', m'+s')]$   $\square$

**Theorem 6.6. ( $(\mathbb{Z}, +)$  is a group)** We have the following properties for  $+: \mathbb{Z} \times \mathbb{Z} \rightarrow \mathbb{Z}$

1. **(Associativity)**  $\forall n, m, k \in \mathbb{Z}$  we have  $(x+y)+z = x+(y+z)$
2. **(Neutral element)**  $\forall n \in \mathbb{Z}$  there exists a  $0 \in \mathbb{Z}$  with  $n+0 = 0+n = n$  (here  $0 = \sim[(1, 1)]$ , note that 0 here is not the same as the 0 in  $\mathbb{N}$ , although the  $1 \in \mathbb{N}$ )
3. **(Commutativity)**  $\forall n, m \in \mathbb{Z}$  then  $n+m = m+n$

4. (**Inverse Element**)  $\forall n \in \mathbb{Z}$  there exists a **unique** element  $-n$  such that  $(-n) + n = 0 = n + (-n)$ . If  $n = \sim[(n_1, n_2)] \Rightarrow -n = \sim[(n_2, n_1)]$  (note as inverse elements are always unique we don't have to prove that  $-n$  is defined independent of its representation).

**Proof.**

1. If  $n = \sim[(n_1, n_2)]$ ,  $m = \sim[(m_1, m_2)]$  and  $k = \sim[(k_1, k_2)]$  then  $(n+m)+k = \sim[(n_1+m_1, n_2+m_2)] + k = \sim[((n_1+m_1)+k_1, (n_2+m_2)+k_2)] \stackrel{4.32}{=} \sim[(n_1+(m_1+k_1), n_2+(m_2+k_2))] = \sim[(n_1, n_2)] + \sim[(m_1+k_1, m_2+k_2)] = n + ([m_1, m_2] + \sim[(k_1, k_2)]) = n + (m+k)$
2. If  $n = \sim[(n_1, n_2)]$  then  $\sim[(1, 1)] + \sim[(n_1, n_2)] = \sim[(1 + n_1, 1 + n_2)]$  now  $(1+n_1)+n_2 \stackrel{4.32, 4.33}{=} (1+n_2)+n_1 \Rightarrow (1+n_1, 1+n_2) \sim (n_1, n_2) \Rightarrow \sim[(1, 1)] + \sim[(n_1, n_2)] = \sim[(n_1, n_2)]$ . Also  $\sim[(n_1, n_2)] + \sim[(1, 1)] = \sim[(n_1+1, n_2+1)] \stackrel{4.33}{=} \sim[(1+n_1, 1+n_2)] \stackrel{\text{already proved}}{=} \sim[(n_1, n_2)]$
3. If  $n = \sim[(n_1, n_2)]$  and  $m = \sim[(m_1, m_2)]$  then  $n+m = \sim[(n_1+m_1, n_2+m_2)] \stackrel{4.33}{=} \sim[(m_1+n_1, m_2+n_2)] = m+n$
4. If  $n = \sim[(n_1, n_2)]$  define then  $-n = \sim[(n_2, n_1)]$  then  $n+(-n) = \sim[(n_1+n_2, n_2+n_1)] \stackrel{4.33}{=} \sim[(n_1+n_2, n_1+n_2)]$  now  $(n_1+n_2)+1 = (n_1+n_2)+1 \Rightarrow (n_1+n_2, n_1+n_2) = \sim[(1, 1)] \Rightarrow n + (-n) = 0$  and by (3) we have  $(-n)+n=0$   $\square$

**Theorem 6.7.** If  $n \in \mathbb{N}$  then  $0 = \sim[(n, n)]$

**Proof.**  $0 = \sim[(1, 1)]$  and  $n+1 = n+1 \Rightarrow (1, 1) \sim (n, n) \Rightarrow \sim[(n, n)] = \sim[(1, 1)]$   $\square$

**Theorem 6.8.**  $\because \mathbb{Z} \times \mathbb{Z} \rightarrow \mathbb{Z}$  defined by  $(\sim[(m, n)], \sim[(k, r)]) \rightarrow \sim[(m, n)] \cdot \sim[(k, r)] = \sim[(m \cdot k + n \cdot r, n \cdot k + m \cdot r)]$  is a function. Note that we have two multiplications  $\cdot$  here one defined in  $\mathbb{N}$  the other in  $\mathbb{Z}$ .

**Proof.** If  $(m, n) \sim (m', n')$  and  $(k, r) \sim (k', r')$  then we have  $m+n' = n+m' \wedge k+r' = r+k'$  so we have then

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

proving that our definition is independent of the representation and that  $\cdot$  is a function.  $\square$

**Theorem 6.9. ( $\langle \mathbb{Z}, \cdot \rangle$  is a abelian semi-group)** If  $n, m, k \in \mathbb{Z}$  then we have

1. **(Commutativity)**  $n \cdot m = m \cdot n$
2. **(Associativity)**  $n \cdot (m \cdot k) = (n \cdot m) \cdot k$
3. **(Neutral Element)** There exist a  $1 = \sim[(2, 1)]$  such that  $n \cdot 1 = 1 \cdot n = n$  (note that we have two 1 here a  $1 \in \mathbb{Z}$  and a  $1 \in \mathbb{N}$ )
4. **(Distributivity)**  $n \cdot (m+k) = n \cdot m + n \cdot k$
5. **(There does not exists a zero divisor)**  $n \cdot m = 0 \Rightarrow n = 0 \vee m = 0$
6.  $(-1) \cdot (-1) = 1$

Note that the symbol 1 is here used as the unit of  $\mathbb{N}$  and  $\mathbb{Z}$ , of course these units are different objects, but context will always tell if  $1 \in \mathbb{N}$  or  $1 = \sim[(2, 1)] \in \mathbb{Z}$ .

**Proof.** Let  $n = \sim[(n_1, n_2)]$ ,  $m = \sim[(m_1, m_2)]$  and  $k = \sim[(k_1, k_2)]$  then we have

1.  $n \cdot m = \sim[(n_1, n_2)] \cdot \sim[(m_1, m_2)] = \sim[(n_1 \cdot m_1 + n_2 \cdot m_2, n_2 \cdot m_1 + n_1 \cdot m_2)] \stackrel{4.41}{=} \sim[(m_1 \cdot n_1 + m_2 \cdot n_2, m_1 \cdot n_2 + m_2 \cdot n_1)] =$
2.  $n \cdot (m \cdot k) = n \cdot \sim[(m_1 \cdot k_1 + m_2 \cdot k_2, m_2 \cdot k_1 + m_1 \cdot k_2)] = \sim[(n_1 \cdot (m_1 \cdot k_1 + m_2 \cdot k_2) + n_2 \cdot (m_2 \cdot k_1 + m_1 \cdot k_2), n_2 \cdot (m_1 \cdot k_1 + m_2 \cdot k_2) + n_1 \cdot (m_2 \cdot k_1 + m_1 \cdot k_2))] = \sim[((n_1 \cdot m_1) \cdot k_1 + (n_1 \cdot m_2) \cdot k_2 + (n_2 \cdot m_2) \cdot k_1 + (n_2 \cdot m_1) \cdot k_2, (n_2 \cdot m_1) \cdot k_1 + (n_2 \cdot m_2) \cdot k_2 + (n_1 \cdot m_1) \cdot k_2) = \sim[((n_1 \cdot m_1 + n_2 \cdot m_2) \cdot k_1 + (n_1 \cdot m_2 + n_2 \cdot m_1) \cdot k_2, (n_2 \cdot m_1 + n_1 \cdot m_2) \cdot k_1 + (n_2 \cdot m_2 + n_1 \cdot m_1) \cdot k_2)] = \sim[(n_1 \cdot m_1 + n_2 \cdot m_2, n_1 \cdot m_2 + n_2 \cdot m_1)] \cdot \sim[(k_1, k_2)] = (\sim[(n_1, n_2)] \cdot \sim[(m_1, m_2)]) \cdot \sim[(k_1, k_2)] = (n \cdot m) \cdot k$
3.  $1 \cdot n = \sim[2, 1] \cdot \sim[(n_1, n_2)] = \sim[(2 \cdot n_1 + 1 \cdot n_2, 1 \cdot n_1 + 2 \cdot n_2)] = \sim[((1+1) \cdot n_1 + n_2, n_1 + (1+1) \cdot n_2)] = \sim[(n_1 + n_1 + n_2, n_1 + n_2 + n_2)] = \sim[(n_1, n_2)] = n$  as  $n_1 + n_1 + n_2 = n_1 + n_2 + n_2 + n_1$ . The fact that  $n \cdot 1 = n$  follows from (1).
4.  $n \cdot (m+k) = n \cdot \sim[(m_1+k_1, m_2+k_2)] = \sim[(n_1 \cdot (m_1+k_1) + n_2 \cdot (m_2+k_2), n_2 \cdot (m_1+k_1) + n_1 \cdot (m_2+k_2))] = \sim[(n_1 \cdot m_1 + n_1 \cdot k_1 + n_2 \cdot m_2 + n_2 \cdot k_2, n_2 \cdot m_1 + n_2 \cdot k_1 + n_1 \cdot m_2 + n_1 \cdot k_2)] = \sim[((n_1 \cdot m_1 + n_2 \cdot m_2 + n_1 \cdot k_1 + n_2 \cdot k_2, n_1 \cdot m_2 + n_2 \cdot m_1 + n_2 \cdot k_1 + n_1 \cdot k_2))] = \sim[(n_1 \cdot m_1 + n_2 \cdot m_2, n_1 \cdot m_2 + n_2 \cdot m_1)] + \sim[(n_1 \cdot k_1 + n_2 \cdot k_2, n_2 \cdot k_1 + n_1 \cdot k_2)] = \sim[(n_1, n_2)] \cdot \sim[(m_1, m_2)] + \sim[(n_1, n_2)] \cdot \sim[(k_1, k_2)] = n \cdot m + n \cdot k$
5. If  $n \cdot m = 0$  then  $\sim[(n_1 \cdot m_1 + n_2 \cdot m_2, n_2 \cdot m_1 + n_1 \cdot m_2)] = \sim[(1, 1)] \Rightarrow$

$$\begin{aligned} n_1 \cdot m_1 + n_2 \cdot m_2 + 1 &= n_2 \cdot m_1 + n_1 \cdot m_2 + 1 \\ &\stackrel{4.44}{=} \\ n_1 \cdot m_1 + n_2 \cdot m_2 &= n_2 \cdot m_1 + n_1 \cdot m_2 \end{aligned}$$

Suppose now that  $\sim[(n_1, n_2)] \neq \sim[(1, 1)] \Rightarrow n_1 + 1 \neq n_2 + 1$  then if  $n_1 = n_2 \Rightarrow n_1 + 1 = n_2 + 1$  a contradiction so we have  $n_1 \neq n_2$ . So the following cases are possible

- a.  $n_1 < n_2 \stackrel{4.61}{\Rightarrow} \exists l \in \mathbb{N}_0 \Rightarrow n_1 + l = n_2 \Rightarrow n_1 \cdot m_1 + (n_1 + l) \cdot m_2 = (n_1 + l) \cdot m_1 + n_1 \cdot m_2 \Rightarrow l \cdot m_2 + (n_1 \cdot m_1 + n_1 \cdot m_2) = l \cdot m_1 + (n_1 \cdot m_1 + n_1 \cdot m_2) \stackrel{4.44}{\Rightarrow} l \cdot m_1 = l \cdot m_2 \stackrel{4.72}{\Rightarrow} m_1 = m_2 \Rightarrow m_1 + 1 = m_2 + 1 \Rightarrow \sim[(m_1, m_2)] = \sim[(1, 1)] = 0$
- b.  $n_2 < n_1 \stackrel{4.61}{\Rightarrow} \exists l \in \mathbb{N}_0 \Rightarrow n_2 + l = n_1 \Rightarrow (n_2 + l) \cdot m_1 + n_2 \cdot m_2 = n_2 \cdot m_1 + (n_2 + l) \cdot m_2 \Rightarrow l \cdot m_1 + (n_2 \cdot m_1 + n_2 \cdot m_2) = l \cdot m_2 + (n_2 \cdot m_1 + n_2 \cdot m_2) \stackrel{4.44}{\Rightarrow} l \cdot m_2 = l \cdot m_1 \stackrel{4.72}{\Rightarrow} m_2 = m_1 \Rightarrow m_2 + 1 = m_1 + 1 \Rightarrow \sim[(m_1, m_2)] = \sim[(1, 1)] = 0$
6.  $(-1) \cdot (-1) = \sim[(1, 2)] \cdot \sim[(1, 2)] = \sim[(1 \cdot 1 + 2 \cdot 2, 2 \cdot 1 + 1 \cdot 2)] = \sim[(5, 4)] = \sim[(2, 1)] = 1 \quad \square$

From 6.6 and 6.9 we have that

**Theorem 6.10.**  $\langle \mathbb{Z}, +, \cdot \rangle$  is a ring even more  $\langle \mathbb{Z}, +, \cdot \rangle$  forms a integral domain

**Notation 6.11.** Whenever we write  $a - b$  we actually mean  $a + (-b)$  (where  $-b$  is the inverse of  $b$  in the group  $\langle \mathbb{Z}, + \rangle$ ).

**Theorem 6.12. (Absorbing element)** If  $n \in \mathbb{Z}$  then  $0 \cdot n = 0 = n \cdot 0$

**Proof.** If  $n = \sim[(n_1, n_2)]$  then we have

$$\begin{aligned}
 0 \cdot n &= \sim[(1, 1)] \cdot \sim[(n_1, n_2)] \\
 &= \sim[(1 \cdot n_1 + 1 \cdot n_2, 1 \cdot n_1 + 1 \cdot n_2)] \\
 &= \sim[(n_1 + n_2, n_1 + n_2)] \\
 &\stackrel{6.4}{=} \sim[(n_1, n_1)] = \sim[(0 + n_1, 0 + n_1)] \\
 &\stackrel{6.4}{=} \sim[(0, 0)] \\
 &\stackrel{6.4}{=} \sim[(0 + 1, 0 + 1)] \\
 &\stackrel{6.4}{=} \sim[(1, 1)] \\
 &= 0
 \end{aligned}$$

The remaining is proved by commutativity of  $\cdot$ . □

**Lemma 6.13.** If  $n \in \mathbb{Z}$  then  $-n = (-1) \cdot n$

**Proof.** First  $1 = \sim[(2, 1)] \Rightarrow -1 = \sim[(1, 2)] \stackrel{6.4}{=} \sim[(0, 1)]$ . Now if  $n \in \mathbb{Z} \Rightarrow n = \sim[(n_1, n_2)]$  then we have

$$\begin{aligned}
 (-1) \cdot n &= \sim[(0 \cdot n_1 + 1 \cdot n_2, 0 \cdot n_2 + 1 \cdot n_1)] \\
 &= \sim[(n_2, n_1)] \\
 &= -n
 \end{aligned}$$

□

**Theorem 6.14.** If  $n, m \in \mathbb{Z}$  then  $-(n \cdot m) = (-n) \cdot m = n \cdot (-m)$

**Proof.**  $-(n \cdot m) = (-1) \cdot (n \cdot m) = ((-1) \cdot n) \cdot m \stackrel{6.13}{=} (-n) \cdot m$ , also  $-(n \cdot m) = -(m \cdot n) = (-m) \cdot n = n \cdot (-m)$  □

**Theorem 6.15.** If  $n, k, r \in \mathbb{Z}$  and  $r \neq 0$  then we have  $n \cdot r = k \cdot r \Rightarrow n = k$

**Proof.** From  $n \cdot r = k \cdot r$  we get  $n \cdot r + (-k \cdot r) = k \cdot r + (-k \cdot r) = 0 \Rightarrow n \cdot r + (-k) \cdot r = 0 \stackrel{6.9}{\Rightarrow} (n + (-k)) \cdot r = 0 \stackrel{6.9}{\Rightarrow} n + (-k) = 0 \Rightarrow (n + (-k)) + k = k \Rightarrow n + ((-k) + k) = k \Rightarrow n + 0 = k \Rightarrow n = k$  □

### 6.1.1 Power in $\mathbb{Z}$

**Definition 6.16.** As  $\langle \mathbb{Z}, \cdot \rangle$  is a abelian semi-group we have by 4.22 that given a  $a \in \mathbb{Z}$  and  $n \in \mathbb{N}$  that there exists a  $a^n$  such that

$$\begin{aligned}
 a^0 &= 1 \\
 a^{n+1} &= a^n \cdot a \stackrel{\text{abelian}}{=} a \cdot a^n
 \end{aligned}$$

**Theorem 6.17.** If  $n, n' \in \mathbb{N}$  and  $a \in \mathbb{Z}$  then  $a^{n'+n} = a^{n'} \cdot a^n$

**Proof.** We prove this by induction on  $n$ . So let  $X = \{n \in \mathbb{N} | a^{n'+n} = a^{n'} \cdot a^n\}$  then we have

1. If  $n = 0$  then  $a^{n'+n} = a^{n'+0} = a^{n'} = a^{n'} \cdot 1 = a^{n'} \cdot a^0 \Rightarrow 0 \in X$
2. If  $n \in X$  then  $a^{n'+(n+1)} = a^{(n'+n)+1} = a^{(n'+n)} \cdot a \stackrel{n \in X}{=} (a^{n'} \cdot a^n) \cdot a = a^{n'} \cdot (a^n \cdot a) = a^{n'} \cdot a^{n+1}$  and thus  $n+1 \in X$

Using mathematical induction (see 4.10) we have  $X = \mathbb{N}$  proving the theorem □

**Theorem 6.18.** If  $n \in \mathbb{N}$  and  $a, b \in \mathbb{Z}$  then we have  $(a \cdot b)^n = a^n \cdot b^n$

**Proof.** We prove this by induction so take  $\mathcal{S} = \{n \in \mathbb{N} | (a \cdot b)^n = a^n \cdot b^n\}$  then we have

**$n = 0$ .** then  $(a \cdot b)^0 = 1 = 1 \cdot 1 = a^0 \cdot b^0$

$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$ . then  $(a \cdot b)^{n+1} = (a \cdot b)^n \cdot (a \cdot b) = (a^n \cdot b^n) \cdot (a^n \cdot a) \cdot (b^n \cdot b) = a^{n+1} \cdot b^{n+1}$   $\square$

**Theorem 6.19.** *In  $\mathbb{Z}$  we have*

$$\begin{aligned} 0^n &= 0 \text{ (if } n \neq 0) \\ 1^n &= 1 \\ (-1)^n &= -1 \text{ or } 1 \\ (-1)^{2 \cdot n} &= 1 \\ (-1)^{2 \cdot n+1} &= -1 \end{aligned}$$

**Proof.**

1. If  $n \neq 0 \Rightarrow \exists m \in \mathbb{N} \vdash n = m+1$  then  $0^n = 0^{(m+1)} = 0^m \cdot 0 = 0$
2.  $1^n = 1$  is proved by induction on  $n$ , let  $X = \{n \in \mathbb{N} \mid 1^n = 1\}$  then
  - a.  $1^0 = 1 \Rightarrow 0 \in X$
  - b. If  $n \in X \Rightarrow 1^{n+1} = 1^n \cdot 1 \underset{n \in X}{=} 1 \cdot 1 = 1 \Rightarrow n+1 \in X$   
so  $X = \mathbb{N}$
3.  $(-1)^n = \pm 1$  is proved by induction on  $n$ , let  $X = \{n \in \mathbb{N} \mid (-1)^n = -1 \text{ or } 1\}$  then
  - a.  $(-1)^0 = 1 \Rightarrow 0 \in X$
  - b. If  $n \in X$  then  $(-1)^{n+1} = (-1)^n \cdot (-1) \underset{n \in X}{=} (-1) \cdot (-1) \vee 1 \cdot (-1) = 1 \vee -1 \Rightarrow n+1 \in X$   
so  $X = \mathbb{N}$
4.  $(-1)^{2 \cdot n} = (-1)^{(1+1) \cdot n} = (-1)^{n+n} = (-1)^n \cdot (-1)^n \underset{(3)}{=} (-1) \cdot (-1) \text{ or } 1 \cdot 1 = 1$
5.  $(-1)^{2 \cdot n+1} = (-1)^{2 \cdot n} \cdot (-1) \underset{(4)}{=} 1 \cdot (-1) = -1$   $\square$

## 6.2 Order relation of the Integers

**Definition 6.20.**  $\mathbb{N}_{\mathbb{Z}} = \{\sim[(s(n), 1)] \mid n \in \mathbb{N}\} \subseteq \mathbb{Z}$ ,  $\mathbb{N}_{\mathbb{Z}}$  is called the collection of positive whole numbers.

**Note 6.21.**  $0 = \sim[(1, 1)] \in \mathbb{N}_{\mathbb{Z}}$ ,  $1 = \sim[(2, 1)] \in \mathbb{N}_{\mathbb{Z}}$

**Theorem 6.22.** *We have the following*

1.  $\langle \mathbb{N}_{\mathbb{Z}}, + \rangle$  is a sub-semi-group of  $\langle \mathbb{Z}, + \rangle$
2.  $\langle \mathbb{N}_{\mathbb{Z}}, \cdot \rangle$  is a sub-semi-group of  $\langle \mathbb{Z}, \cdot \rangle$
3.  $i_{\mathbb{N}}: \mathbb{N} \rightarrow \mathbb{N}_{\mathbb{Z}}$  defined by  $n \rightarrow \sim[(s(n), 1)]$  forms a
  - a. group isomorphism between  $\langle \mathbb{N}, + \rangle \rightarrow \langle \mathbb{N}_{\mathbb{Z}}, + \rangle$
  - b. group isomorphism between  $\langle \mathbb{N}, \cdot \rangle \rightarrow \langle \mathbb{N}_{\mathbb{Z}}, \cdot \rangle$
4. For every  $z \in \mathbb{Z}$  we have  $\exists x, y \in \mathbb{N}_z$  such that  $z = x + (-y)$

As  $0, 1 \in \mathbb{N}_{\mathbb{Z}}$  we have by 3.9 that  $\langle \mathbb{N}_{\mathbb{Z}}, + \rangle$  and  $\langle \mathbb{N}_{\mathbb{Z}}, \cdot \rangle$  are semi-groups

**Proof.**

1. We have the following

- a.  $n, m \in \mathbb{N}_{\mathbb{Z}}$  then  $\exists n', m' \in \mathbb{N}$  such that  $n = \sim[(s(n'), 1)]$ ,  $m = \sim[(s(m'), 1)]$  then  $n+m = \sim[(s(n') + s(m'), 1+1)] = \sim[((n'+m'+1)+1, 1+1)] \underset{6.4}{=} \sim[(n'+m'+1, 1)] = \sim[(s(n'+m'), 1)] \in \mathbb{N}_{\mathbb{Z}}$

$$\text{b. } 0 = \sim[(1, 1)] = \sim[(s(0), 1)] \in \mathbb{N}_{\mathbb{Z}}$$

using 3.8 we have our proof.

2. We have the following

$$\text{a. } n, m \in \mathbb{N}_{\mathbb{Z}} \text{ then } \exists n', m' \in \mathbb{N} \text{ such that } n = \sim[(s(n'), 1)], m = \sim[(s(m'), 1)]. \text{ we have then that } n \cdot m = \sim[(s(n') \cdot s(m') + 1 \cdot 1, 1 \cdot s(m') + 1 \cdot s(m'))] = \sim[s(n') \cdot s(m') + 1, s(m') + s(n')] \sim [(n'+1)(m'+1) + 1, n'+1+m'+1] = \sim[(n' \cdot m' + n' + m' + 1, n' + m' + 1)] = \sim[(n' \cdot m' + 1, 1)] \in \mathbb{N}_{\mathbb{Z}}$$

$$\text{b. } 1 = \sim[(2, 1)] = \sim[(s(1), 1)] \in \mathbb{N}_{\mathbb{Z}}$$

so using 3.8 we have our proof.

3. First we prove that the function  $i_{\mathbb{N}}: \mathbb{N} \rightarrow \mathbb{N}_{\mathbb{Z}}$  is a bijection

$$\text{a. (Injectivity) if } i_{\mathbb{N}}(n) = i_{\mathbb{N}}(m) \Rightarrow \sim[(s(n), 1)] = \sim[s(m), 1] \Rightarrow (s(n), 1) \sim (s(m), 1) \Rightarrow s(m) + 1 = s(n) + 1 \xrightarrow{4.44} s(m) = s(n) \xrightarrow{4.15} m = n$$

$$\text{b. (Surjectivity) If } n \in \mathbb{N}_{\mathbb{Z}} \text{ then there exists a } n' \in \mathbb{N} \vdash n = \sim[(s(n'), 1)] = i_{\mathbb{N}}(n')$$

Next we prove that  $i_{\mathbb{N}}(n+m) = i_{\mathbb{N}}(n) + i_{\mathbb{N}}(m)$  and  $i_{\mathbb{N}}(n \cdot m) = i_{\mathbb{N}}(n) \cdot i_{\mathbb{N}}(m)$

$$\text{a. } i_{\mathbb{N}}(n+m) = \sim[(s(n+m), 1)] \xrightarrow{6.4} \sim[(s(n+m)+1, 1+1)] = \sim[((n+m)+1)+1, 1+1] = \sim[((n+1)+(m+1), 1+1)] = \sim[(n+1, 1)] + \sim[(m+1, 1)] = \sim[(s(n), 10)] + \sim[(s(m), 1)] = i_{\mathbb{N}}(n) + i_{\mathbb{N}}(m)$$

$$\text{b. } i_{\mathbb{N}}(n \cdot m) = \sim[(s(n \cdot m), 1)] = \sim[(n \cdot m + 1, 1)] \xrightarrow{6.4} \sim[(n \cdot m, 0)] = \sim[(n, 0)] \cdot \sim[(m, 0)] \xrightarrow{6.4} \sim[(n+1, 1)] + \sim[(m+1, 1)] = \sim[(s(n), 1)] + \sim[(s(m), 1)] = i_{\mathbb{N}}(n) \cdot i_{\mathbb{N}}(m)$$

$$\text{4. If } z \in \mathbb{Z} \text{ then } z = \sim[(n, m)] = \sim[(n, 0)] + \sim[(0, m)] = \sim[(n, 0)] + (-(\sim[(m, 0)])) \xrightarrow{6.4} \sim[(s(n), 1)] + (-(\sim[(s(m), 1)])) = x + (-y) \text{ where } x = \sim[(s(n), 1)] \in \mathbb{N}_{\mathbb{Z}} \text{ and } y = \sim[(s(m), 1)] \in \mathbb{N}_{\mathbb{Z}} \quad \square$$

**Definition 6.23.**  $-\mathbb{N}_{\mathbb{Z}} = \{-n \mid n \in \mathbb{N}_{\mathbb{Z}}\}$ ,  $-\mathbb{N}_{\mathbb{Z}}$  is the collection of negative numbers.

**Lemma 6.24.**  $\mathbb{Z} = \mathbb{N}_{\mathbb{Z}} \cup (-\mathbb{N}_{\mathbb{Z}})$ , so the whole numbers is the union of the positive and negative numbers.

**Proof.** First as  $\mathbb{N}_{\mathbb{Z}} \subseteq \mathbb{Z}$  and thus  $-\mathbb{N}_{\mathbb{Z}} \subseteq \mathbb{Z}$  we have  $\mathbb{N}_{\mathbb{Z}} \cup (-\mathbb{N}_{\mathbb{Z}}) \subseteq \mathbb{Z}$ . Next assume that  $z \in \mathbb{Z}$  then by 6.22 we have there exists a  $x, y \in \mathbb{N}_{\mathbb{Z}}$  such that  $z = x + (-y)$ . Using 6.22 again there exists  $x', y' \in \mathbb{N}$  with  $x = i_{\mathbb{N}}(x')$  and  $y = i_{\mathbb{N}}(y')$  where  $i_{\mathbb{N}}$  is the isomorphism defined by  $i_{\mathbb{N}}: \mathbb{N} \rightarrow \mathbb{N}_{\mathbb{Z}}$   $n \rightarrow f(n) = \sim[(s(n), 1)]$ . We have then the following cases to consider ( $\mathbb{N}, \leq$  is fully ordered)

1.  $x' = y'$  then  $x = i_{\mathbb{N}}(x') = i_{\mathbb{N}}(y') = y \Rightarrow z = x + (-x) = 0 = \sim[(1, 1)] = \sim[(s(0), 1)] \in \mathbb{N}_{\mathbb{Z}} \Rightarrow z \in \mathbb{N}_{\mathbb{Z}} \cup (-\mathbb{N}_{\mathbb{Z}})$
2.  $x' < y'$  then by 4.61 there exists a  $k' \in \mathbb{N}_0$  such that  $y' = x' + k'$ . Take then  $k = i_{\mathbb{N}}(k') \in \mathbb{N}_{\mathbb{Z}}$ , by the linearity of  $i_{\mathbb{N}}$  we have then  $y = i_{\mathbb{N}}(y') = i_{\mathbb{N}}(x' + k') = i_{\mathbb{N}}(x') + i_{\mathbb{N}}(k') = x + k$  and thus  $z = x + (-y) = x + (-1) \cdot y = x + (-1) \cdot (x + k) = x + ((-x) + (-k)) = -k \in -\mathbb{N}_{\mathbb{Z}} \Rightarrow z \in -\mathbb{N}_{\mathbb{Z}} \Rightarrow z \in \mathbb{N}_{\mathbb{Z}} \cup (-\mathbb{N}_{\mathbb{Z}})$
3.  $y' < x'$  then by 4.61 there exists a  $k' \in \mathbb{N}_0$  such that  $x' = y' + k'$ , if  $k = i_{\mathbb{N}}(k') \in \mathbb{N}_{\mathbb{Z}}$  we have then  $x = i_{\mathbb{N}}(x') = i_{\mathbb{N}}(y') + i_{\mathbb{N}}(k') = y + k$  and thus  $z = x + (-y) = y + k + (-y) = k \in \mathbb{N}_{\mathbb{Z}} \Rightarrow z \in \mathbb{N}_{\mathbb{Z}} \Rightarrow z \in \mathbb{N}_{\mathbb{Z}} \cup (-\mathbb{N}_{\mathbb{Z}})$   $\square$

We can now define a order relation on  $\mathbb{Z}$ .

**Definition 6.25. (Order relation in  $\mathbb{Z}$ )** We define  $\leq \subseteq \mathbb{Z} \times \mathbb{Z}$  by  $\leq = \{(n, m) \mid m + (-n) \in \mathbb{N}_{\mathbb{Z}}\}$  in other words  $n \leq m$  iff  $m + (-n) \in \mathbb{N}_{\mathbb{Z}}$

**Theorem 6.26.**  $\langle \mathbb{Z}, \leq \rangle$  is a partially ordered set that is fully-ordered

**Proof.**

1. **(reflectivity)** If  $n \in \mathbb{Z}$  then  $n + (-n) = 0 = \sim[(1, 1)] = \sim[(s(0), 1)] \in \mathbb{N}_{\mathbb{Z}}$

2. (**anti-symmetry**) If  $n, m \in \mathbb{Z}$ ,  $n \leq m$  and  $m \leq n$  then  $m + (-n) \in \mathbb{N}_{\mathbb{Z}}$  and  $n + (-m) \in \mathbb{N}_{\mathbb{Z}}$ .  
As  $n, m \in \mathbb{Z}$  we have  $n = \sim[(n_1, n_2)]$  and  $m = \sim[(m_1, m_2)]$  then

$$\begin{aligned} n + (-m) &= \sim[(n_1 + m_2, n_2 + m_1)] = \sim[(s(k), 1)] \text{ where } k \in \mathbb{N} \\ m + (-n) &= \sim[(m_1 + n_2, m_2 + n_1)] = \sim[(s(l), 1)] \text{ where } l \in \mathbb{N} \end{aligned}$$

we have then

$$\begin{aligned} n_1 + m_2 + 1 &= n_2 + m_1 + s(k) \\ &= n_2 + m_1 + k + 1 \\ &\Leftrightarrow \\ n_1 + m_2 &= n_2 + m_1 + k \\ m_1 + n_2 + 1 &= m_2 + n_1 + s(l) \\ &= m_2 + n_1 + l + 1 \\ &\Leftrightarrow \\ m_1 + n_2 &= m_2 + n_1 + l \end{aligned}$$

summing the two equalities we get

$$\begin{aligned} n_1 + m_2 + m_1 + n_2 &= n_1 + m_2 + m_1 + n_2 + k + l \\ &\stackrel{4.44}{\Leftrightarrow} \\ 0 &= k + l \end{aligned}$$

Using 4.60 we have then that  $k = l = 0$  but then we have from  $n_1 + m_2 = n_2 + m_1 + k$  that  $n_1 + m_2 = n_2 + m_1$  or  $(n_1, n_2) \sim (m_1, m_2) \Rightarrow \sim[(n_1, n_2)] = \sim[(m_1, m_2)] \Rightarrow n = m$

3. (**transitivity**) If  $n, m, k \in \mathbb{Z}$  then if  $n \leq m$  and  $m \leq k$  we have  $m + (-n) \in \mathbb{N}_{\mathbb{Z}}$  and  $k + (-m) \in \mathbb{N}_{\mathbb{Z}}$  then as we have that  $\mathbb{N}_{\mathbb{Z}}$  is a sub-semi-group we have that  $\mathbb{N}_{\mathbb{Z}} \ni (k + (-m)) + (m + (-n)) = k + ((-m) + m) + (-n) = k + 0 + (-n) = k + (-n) \Rightarrow k \leq n$
4. (**fully ordering**) If  $n, m \in \mathbb{Z}$  then by 6.24 we have for  $n + (-m) \in \mathbb{Z}$  that either

- $n + (-m) \in \mathbb{N}_{\mathbb{Z}} \Rightarrow m \leq n$
- $n + (-m) \in -\mathbb{N}_{\mathbb{Z}} \Rightarrow -(n + (-m)) \in \mathbb{N}_{\mathbb{Z}} \Rightarrow m + (-n) \in \mathbb{N}_{\mathbb{Z}} \Rightarrow n \leq m$  proving fully ordering.  $\square$

**Theorem 6.27.** If  $n, m \in \mathbb{Z}$  with  $n \leq m$  (or  $n < m$ ) then  $-m \leq -n$  (or  $-m < -n$ )

**Proof.**

- If  $n \leq m$  then  $m + (-n) \in \mathbb{N}_{\mathbb{Z}} \Rightarrow (-n) + (-(-m)) = (-n) + m = m + (-n) \in \mathbb{N}_{\mathbb{Z}} \Rightarrow -m \leq -n$
- If  $n < m \Rightarrow n \leq m \wedge n \neq m \Rightarrow -m \leq -n \wedge -n \neq -m \Rightarrow -m < -n$   $\square$

**Theorem 6.28.** If  $n \in \mathbb{Z}$  then we have  $n \in \mathbb{N}_{\mathbb{Z}} \Leftrightarrow 0 \leq n$ . Note that this is the motivation to call  $\mathbb{N}_{\mathbb{Z}}$  the set of positive numbers.

**Proof.**

$$\begin{aligned} 0 \leq z &\Leftrightarrow n + (-0) \in \mathbb{N}_{\mathbb{Z}} \\ &\stackrel{6.13}{\Leftrightarrow} n + (-1) \cdot 0 \in \mathbb{N}_{\mathbb{Z}} \\ &\stackrel{4.37}{\Leftrightarrow} n + 0 \in \mathbb{N}_{\mathbb{Z}} \\ &\Leftrightarrow n \in \mathbb{N}_{\mathbb{Z}} \end{aligned}$$

$\square$

**Theorem 6.29.** If  $n \in \mathbb{Z}$  so  $n = \sim[(n_1, n_2)]$  then  $0 < n$  iff  $n_2 < n_1$

**Proof.**

$\Rightarrow$

Assume that  $0 < n$  then  $n \neq 0$  and  $0 \leq n$  so  $0 \neq 0$  and  $n \in \mathbb{N}_Z \Rightarrow \exists m \in \mathbb{N} \vdash n = \sim[(n_1, n_2)] \sim[(s(m), 1)] \Rightarrow n_1 + 1 = n_2 + s(m) = n_2 + (m + 1) = (n_2 + m) + 1 \Rightarrow n_1 = n_2 + m$ . Now if  $m = 0 \Rightarrow n_1 = n_2 \Rightarrow n = \sim[(n_1, n_1)] = 0 \neq n$  a contradiction so we have  $m \in \mathbb{N}_0$  but then by 4.61 we have  $n_2 < n_1$ .

$\Leftarrow$  If  $n_2 < n_1 \xrightarrow[4.61]{6.7} \exists m \in \mathbb{N}_0 \vdash n_1 = n_2 + m \Rightarrow n_1 + 1 = n_2 + s(m) \Rightarrow \sim[(n_1, n_2)] \sim[(s(m), 1)] \Rightarrow n \in \mathbb{N}_Z \Rightarrow 0 \leq n$ , now if  $n = 0 = \sim[(1, 1)]$  then  $s(m) + 1 = 1 + 1 \Rightarrow m + (1 + 1) = 0 + (1 + 1) \Rightarrow m = 0$  a contradiction so  $n \neq 0$  and thus  $0 < n$   $\square$

**Corollary 6.30.** *If  $n \in \mathbb{Z}$  so  $n = \sim[(n_1, n_2)]$  then  $0 \leq n$  iff  $n_2 \leq n_1$*

**Proof.**

$\Rightarrow$

If  $0 \leq n$  then we have either

1.  $n = 0 \xrightarrow[6.7]{} n_1 = n_2 \Rightarrow n_2 \leq n_1$
2.  $0 < n \xrightarrow{\text{previous theorem}} n_2 < n_1 \Rightarrow n_2 \leq n_1$

$\Leftarrow$

If  $n_2 \leq n_1$  then we have either

1.  $n_2 = n_1 \Rightarrow n = \sim[(n_1, n_1)] = 0 \Rightarrow 0 \leq n$
2.  $n_2 < n_1 \Rightarrow 0 < n \Rightarrow 0 \leq n$

$\square$

**Theorem 6.31.**  $i_{\mathbb{N}}: \mathbb{N} \rightarrow \mathbb{N}_Z$  defined by  $n \rightarrow i_{\mathbb{N}}(n) = \sim[(s(n), 1)]$  is order preserving, so if  $n, m \in \mathbb{N}$  with  $n \leq m$  then  $i_{\mathbb{N}}(n) \leq i_{\mathbb{N}}(m)$

**Proof.** If  $n \leq m$  then we have the following cases

1.  $n = m \Rightarrow i_{\mathbb{N}}(n) = i_{\mathbb{N}}(m) \Rightarrow \mathbb{N}_Z \ni 0 = i_{\mathbb{N}}(m) + (-i_{\mathbb{N}}(n)) \Rightarrow i_{\mathbb{N}}(n) \leq i_{\mathbb{N}}(m)$
2.  $n < m$  then  $i_{\mathbb{N}}(m) + (-i_{\mathbb{N}}(n)) = \sim[(s(m), 1)] + (-(\sim[s(n), 1])) = \sim[(s(m), 1)] + \sim[(1, s(n))] = \sim[(s(m) + 1, s(n) + 1)] = \sim[(s(m), s(n))] = \sim[(m + 1, n + 1)] = \sim[(m, n)]$ . By 6.29 we have from  $n < m$  that  $0 < \sim[(m, n)] = i_{\mathbb{N}}(m) + (-i_{\mathbb{N}}(n)) \Rightarrow 0 \leq i_{\mathbb{N}}(m) + (-i_{\mathbb{N}}(n)) \xrightarrow{6.28} i_{\mathbb{N}}(m) + (-i_{\mathbb{N}}(n)) \in \mathbb{N}_Z \Rightarrow i_{\mathbb{N}}(n) \leq i_{\mathbb{N}}(m)$   $\square$

**Theorem 6.32.** *If  $n, m, k \in \mathbb{Z}$  then  $n < m \Rightarrow n + k < m + k$*

**Proof.** From  $n < m \Rightarrow n \neq m$  and  $n \leq m$  we have  $n + (-m) \in \mathbb{N}_Z \Rightarrow n + 0 + (-m) \in \mathbb{N}_Z \Rightarrow n + k + (-k) + (-m) \in \mathbb{N}_Z \Rightarrow (n + k) + (-m + k) \in \mathbb{N}_Z \Rightarrow n + k \leq m + k$ . If now  $n + k = m + k \Rightarrow n + k + (-k) = m + k + (-k) \Rightarrow n + 0 = m + 0 \Rightarrow n = m$  a contradiction with  $n < m$  so we have  $n + k < m + k$ .  $\square$

**Corollary 6.33.** *If  $n, m, k \in \mathbb{Z}$  then  $n \leq m \Rightarrow n + k \leq m + k$*

**Proof.** If  $n \leq m$  then we have either

1.  $n = m \Rightarrow n + k = m + k \Rightarrow n + k \leq m + k$
2.  $n < m \xrightarrow{\text{previous theorem}} n + k < m + k \Rightarrow n + k \leq m + k$   $\square$

**Theorem 6.34.** *If  $n, m \in \mathbb{Z}$  then  $n < m \Leftrightarrow 0 < m + (-n)$*

**Proof.**

$\Rightarrow$

If  $n < m \Rightarrow n \leq m$  and  $n \neq m$  so  $n + (-m) \in \mathbb{N}_Z \xrightarrow{6.28} 0 \leq n + (-m)$ . Now if  $0 = n + (-m) \Rightarrow 0 + m = n + (-m) + m \Rightarrow m = n + 0 = n$  contradicting  $n < m$ .

$\Leftarrow$

If  $0 < m + (-n)$  then by the previous corollary we have  $0 + n < m + (-n) + n \Rightarrow n < m + 0 \Rightarrow n < m$   $\square$

**Corollary 6.35.** *If  $n, m \in \mathbb{Z}$  then  $n \leq m \Leftrightarrow 0 \leq m + (-n)$*

**Proof.**

$$\begin{aligned} n \leq m &\Leftrightarrow n + (-m) \in \mathbb{N}_{\mathbb{Z}} \\ &\Leftrightarrow 0 \leq n + (-m) \\ &\quad \square \end{aligned}$$

**Theorem 6.36.** If  $n, m \in \mathbb{Z}$  with  $0 < n$  and  $0 < m \Rightarrow 0 < n \cdot m$

**Proof.** If  $n = \sim[(n_1, n_2)], m = \sim[(m_1, m_2)]$  then from  $0 < n$  and  $0 < m$  we have by 6.29  $n_2 < n_1$  and  $m_2 < m_1 \Rightarrow \exists k, l \in \mathbb{N}_0 \vdash n_1 = n_2 + k$  and  $m_1 = m_2 + l$  this gives

$$\begin{aligned} n_1 \cdot m_1 + n_2 \cdot m_2 &= (n_2 + k) \cdot (m_2 + l) + n_2 \cdot m_2 \\ &= n_2 \cdot m_2 + n_2 \cdot m_2 + n_2 \cdot l + m_2 \cdot k + k \cdot l \\ n_2 \cdot m_1 + n_1 \cdot m_2 &= n_2 \cdot (m_2 + l) + (n_2 + k) \cdot m_2 \\ &= n_2 \cdot m_2 + n_2 \cdot m_2 + n_2 \cdot l + m_2 \cdot k \end{aligned}$$

So using 4.61 and  $(0 < k \text{ then by 4.70 and } 0 < l \text{ we have } 0 < k \cdot l)$  we have that  $n_2 \cdot m_1 + n_1 \cdot m_2 < n_1 \cdot m_1 + n_2 \cdot m_2$ . Now  $n \cdot m = \sim[(n_1 \cdot m_1 + n_2 \cdot m_2, n_2 \cdot m_1 + n_1 \cdot m_2)]$  and thus we have  $0 < n \cdot m \quad \square$

**Theorem 6.37.** If  $n, m \in \mathbb{Z}$  and  $k \in \mathbb{Z} \vdash k > 0$  then  $n < m \Rightarrow n \cdot k < m \cdot k$

**Proof.** From  $n < m$  we have by 4.56 that  $0 < m + (-n)$  then by the above theorem we have  $0 < (m + (-n)) \cdot k = m \cdot k + (-1) \cdot n \cdot k = m \cdot k + (-1)(n \cdot k) = m \cdot k + (-(n \cdot k)) \Rightarrow n \cdot k < m \cdot k \quad \square$

**Corollary 6.38.** If  $n, m \in \mathbb{Z}$  and  $k \in \mathbb{Z} \vdash k \geq 0$  then  $n \leq m \Rightarrow n \cdot k \leq m \cdot k$

**Proof.** For  $k$  we have the following possibilities

1.  $k = 0 \Rightarrow n \cdot k = 0 = m \cdot k \Rightarrow n \cdot k \leq m \cdot k$
2.  $0 < k$  then for  $n \leq m$  we have either
  - a.  $n = m \Rightarrow n \cdot k = m \cdot k \Rightarrow n \cdot k \leq m \cdot k$
  - b.  $n < m \Rightarrow n \cdot k < m \cdot k = n \cdot k \leq m \cdot k$   $\square$

**Theorem 6.39.** If  $n \in \mathbb{Z}$  then  $0 \leq n \cdot n$

**Proof.** Let  $n = \sim[(n_1, n_2)]$  then  $n \cdot n = \sim[(n_1 \cdot n_1 + n_2 \cdot n_2, n_2 \cdot n_1 + n_1 \cdot n_2)]$ . From the totality of order in  $\mathbb{N}$  we have either

1.  $n_1 = n_2$  then  $n = \sim[(n_1, n_1)] = 0 \Rightarrow n \cdot n = 0 \cdot 0 = 0 \Rightarrow 0 \leq n \cdot n$
2.  $n_1 < n_2$  then we have  $\exists k \in \mathbb{N}_0$  with  $n_2 = n_1 + k$

$$\begin{aligned} n_1 \cdot n_1 + n_2 \cdot n_2 &= n_1 \cdot n_1 + (n_1 + k) \cdot (n_1 + k) \\ &= (n_1 \cdot n_1 + n_1 \cdot n_1 + n_1 \cdot k + n_1 \cdot k) + k \cdot k \\ n_2 \cdot n_1 + n_1 \cdot n_2 &= (n_1 + k) \cdot n_1 + n_1 \cdot (n_1 + k) \\ &= (n_1 n_1 + n_1 \cdot n_1 + n_1 \cdot k + n_1 \cdot k) \end{aligned}$$

So as  $0 < k \cdot k$  we have  $n_2 \cdot n_1 + n_1 \cdot n_2 < n_1 \cdot n_1 + n_2 \cdot n_2 \Rightarrow 0 < n \cdot n \Rightarrow 0 \leq n \cdot n$

3.  $n_2 < n_1$  then we have  $\exists k \in \mathbb{N}_0$  with  $n_1 = n_2 + k$

$$\begin{aligned} n_1 \cdot n_1 + n_2 \cdot n_2 &= (n_2 + k) \cdot (n_2 + k) + n_2 \cdot n_2 \\ &= (n_2 \cdot n_2 + n_2 \cdot n_2 + n_2 \cdot k + n_2 \cdot k) + k \cdot k \\ n_2 \cdot n_1 + n_1 \cdot n_2 &= n_2 \cdot (n_2 + k) + (n_2 + k) \cdot n_2 \\ &= (n_2 n_2 + n_2 \cdot n_2 + n_2 \cdot k + n_2 \cdot k) \end{aligned}$$

So as  $0 < k \cdot k$  we have  $n_2 \cdot n_1 + n_1 \cdot n_2 < n_1 \cdot n_1 + n_2 \cdot n_2 \Rightarrow 0 < n \cdot n \Rightarrow 0 \leq n \cdot n$

$\square$

**Theorem 6.40. (Archimedean property of  $\mathbb{Z}$ )** If  $m, n \in \mathbb{Z}$  with  $m > 0$  then there exists a  $k \in \mathbb{N}_{\mathbb{Z}}$  with  $n < k \cdot m$

**Proof.** We have the following cases for  $n$

1.  $n \leq 0$  then if we take  $k = 1 \in \mathbb{N}_{\mathbb{Z}}$  then  $n \leq 0 < m = 1 \cdot m = k \cdot m \Rightarrow n < k \cdot m$
2. If  $n > 0 \Rightarrow n \in \mathbb{N}_{\mathbb{Z}}$  as also  $m > 0$  we have  $m \in \mathbb{N}_{\mathbb{Z}}$  and there exists a  $n', m' \in \mathbb{N}$  such that  $n = i_{\mathbb{Z}}(n'), m = i_{\mathbb{Z}}(m')$ . Assume now that  $m' \leq 0 \Rightarrow i_{\mathbb{Z}}(m') \leq i_{\mathbb{Z}}(0) = \sim[(s(0), 1)] = \sim[(1, 1)] = 0$  a contradiction so we must have that  $m' > 0$ . Then by 4.73 there exists a  $k' \in \mathbb{N}$  such that  $n' < k' \cdot m' \Rightarrow n = i_{\mathbb{Z}}(n') < i_{\mathbb{Z}}(k' \cdot m') = i_{\mathbb{Z}}(k') \cdot i_{\mathbb{Z}}(m') = k \cdot m$  where  $k = i_{\mathbb{Z}}(k') \in \mathbb{N}_{\mathbb{Z}}$   $\square$

**Theorem 6.41.**  $\langle \mathbb{N}_{\mathbb{Z}}, \leq \rangle$  is well-ordered.

**Proof.** Let  $A \subseteq \mathbb{N}_{\mathbb{Z}}$  be a non empty subset of  $\mathbb{N}_{\mathbb{A}}$  then as  $i_{\mathbb{N}}: \mathbb{N} \rightarrow \mathbb{N}_{\mathbb{Z}}$  is a bijection we have that  $i_{\mathbb{N}}^{-1}(A)$  is non empty and thus by the well-ordering of the natural numbers (see 4.52) there exists a  $m' = \min(i_{\mathbb{N}}^{-1}(A)) \in i_{\mathbb{N}}^{-1}(A)$ . Take now  $m = i_{\mathbb{N}}(m')$  then we have

1.  $m' \in i_{\mathbb{N}}^{-1}(A) \Rightarrow m = i_{\mathbb{N}}(m') \in A$
2. If  $a \in A$   $\underset{i_{\mathbb{N}} \text{ is bijective}}{\Rightarrow} \exists a' = i_{\mathbb{N}}^{-1}(a) \in i_{\mathbb{N}}^{-1}(A)$   $\underset{m' \text{ is least element of } A}{\Rightarrow} m' \leq a' \underset{i_{\mathbb{N}} \text{ is order preserving (see 6.31)}}{\Rightarrow} m = i_{\mathbb{N}}(m') \leq i_{\mathbb{N}}(a') = a \Rightarrow m = \min(A)$   $\square$

**Corollary 6.42.**  $\langle \mathbb{N}_{\mathbb{Z}}, \leq \rangle$  is conditionally complete (any non empty subset in  $\mathbb{N}_{\mathbb{Z}}$  bounded above has a supremum (lowest upper bound). Furthermore if  $\emptyset \neq A \subseteq \mathbb{N}$  then if  $\sup(A)$  exists we have  $\sup(A) \in A$

**Proof.** The fact that  $\langle \mathbb{N}_{\mathbb{Z}}, \leq \rangle$  is conditionally complete follows from the fact that  $\langle \mathbb{N}_{\mathbb{Z}}, \leq \rangle$  is well-ordered (see 6.41) and 2.193. Now if  $\emptyset \neq A \subseteq \mathbb{N}_{\mathbb{Z}}$  then there exists a  $a \in A \Rightarrow i_{\mathbb{N}}^{-1}(a) \in i_{\mathbb{N}}^{-1}(A) \Rightarrow \emptyset \neq i_{\mathbb{N}}^{-1}(A)$ . Also if  $M = \sup(A)$  take then  $m = i_{\mathbb{N}}^{-1}(M)$ . If now  $a \in A \Rightarrow a \leq m$  [For if  $m < a \underset{i_{\mathbb{N}} \text{ is order preserving and injective}}{\Rightarrow} M = i_{\mathbb{N}}(m) < i_{\mathbb{N}}(a) \in A$  contradicting the fact that  $M$  is a upper bound of  $A$ ]. Furthermore if  $m'$  is another upper bound of  $i_{\mathbb{N}}^{-1}(A)$  then if  $a \in A$  we have  $i_{\mathbb{N}}^{-1}(a) \in i_{\mathbb{N}}^{-1}(A) \Rightarrow i_{\mathbb{N}}^{-1}(a) \leq m' \underset{i_{\mathbb{N}} \text{ is order preserving}}{\Rightarrow} a = i_{\mathbb{N}}(i_{\mathbb{N}}^{-1}(a)) \leq i_{\mathbb{N}}(m') \Rightarrow i_{\mathbb{N}}(m')$  is a upper bound of  $A$  and as  $M$  is the least upper bound we have  $M \leq i_{\mathbb{N}}(m')$  and then  $i_{\mathbb{N}}^{-1}(M) \leq m'$  [otherwise is  $m' < i_{\mathbb{N}}^{-1}(M) \underset{i_{\mathbb{N}} \text{ is order preserving}}{\Rightarrow} i_{\mathbb{N}}(m') < i_{\mathbb{N}}(i_{\mathbb{N}}^{-1}(M)) = M \leq i_{\mathbb{N}}(m') \Rightarrow i_{\mathbb{N}}(m') < i_{\mathbb{N}}(m')$  a contradiction] so  $m = i_{\mathbb{N}}^{-1}(M)$  is the lowest upper bound of  $i_{\mathbb{N}}^{-1}(A)$  or in other words  $m = \sup(i_{\mathbb{N}}^{-1}(A))$ , using 4.67 we have then that  $i_{\mathbb{N}}^{-1}(M) = m \in i_{\mathbb{N}}^{-1}(A) \Rightarrow M = i_{\mathbb{N}}(i_{\mathbb{N}}^{-1}(M)) \in A \Rightarrow \sup(A) \in M$   $\square$

**Lemma 6.43.**  $\forall n, m \in \mathbb{N}_{\mathbb{Z}}$  then  $n < m \Leftrightarrow \exists k \in \mathbb{N}_{\mathbb{Z}} \setminus \{0\}$  such that  $m = n + k$

**Proof.**

( $\Rightarrow$ )

Take  $n' = i_{\mathbb{N}}^{-1}(n)$  and  $m' = i_{\mathbb{N}}^{-1}(m)$  then  $n' < m'$  [otherwise if  $m' \leq n'$  we have by the order preserving properties of  $i_{\mathbb{N}}$  that  $m = i_{\mathbb{N}}(i_{\mathbb{N}}^{-1}(m)) \leq i_{\mathbb{N}}(i_{\mathbb{N}}^{-1}(n)) = n \Rightarrow m \leq n < m \Rightarrow m < m$  a contradiction]. Using 4.61 there exists a  $k' \in \mathbb{N} \setminus \{0\}$  such that  $m' = n' + k'$  then using the fact that  $i_{\mathbb{N}}$  is a isomorphism we have  $m = i_{\mathbb{N}}(i_{\mathbb{N}}^{-1}(m)) = i_{\mathbb{N}}(m') = i_{\mathbb{N}}(n' + k') = i_{\mathbb{N}}(n') + i_{\mathbb{N}}(k') = i_{\mathbb{N}}(i_{\mathbb{N}}^{-1}(n)) + k = n + k$  where  $k = i_{\mathbb{N}}(k') \in \mathbb{N}_{\mathbb{Z}}$  and  $k \neq 0$  [otherwise if  $k = 0 \Rightarrow i_{\mathbb{N}}(0) = 0 = k = i_{\mathbb{N}}(k') \underset{i_{\mathbb{N}} \text{ is a bijection}}{\Rightarrow} k' = 0$  contradicting  $k' \in \mathbb{N} \setminus \{0\}$ ]

( $\Leftarrow$ )

If  $m = n + k$  where  $k \in \mathbb{N}_{\mathbb{Z}} \setminus \{0\} \Rightarrow 0 < k = (n + k) + (-n) = m + (-n) \Rightarrow 0 < m + (-n) \Rightarrow n = 0 + n < m + (-n) + n = m \Rightarrow n < m$   $\square$

**Definition 6.44.** Given  $z \in \mathbb{Z}$  then  $|z| \in \mathbb{N}_{\mathbb{Z}}$  is defined by  $|z| = \begin{cases} z & \text{if } 0 \leq z \\ -z & \text{if } z < 0 \end{cases}$ . So if  $z \geq 0 \Rightarrow z = |z|$

**Note 6.45.** Given  $z, z' \in \mathbb{Z}$  then its easy to prove that  $|z \cdot z'| = |z| \cdot |z'|$

**Proof.** We have the following cases for  $z$  and  $z'$  to consider

1.  $(0 \leq z \wedge 0 \leq z')$  then  $0 \leq z \cdot z'$  and  $|z| \cdot |z'| = z \cdot z' = |z \cdot z'|$

2.  $(0 \leq z \wedge z' < 0)$  then  $z \cdot z' \leq 0$  and  $z \cdot z' = -|z \cdot z'| \Rightarrow -(z \cdot z') = |z \cdot z'|$ , also  $-(z \cdot z') = z \cdot (-z') = |z| \cdot |z'| \Rightarrow |z| \cdot |z'| = |z \cdot z'|$
3.  $(z < 0 \wedge 0 \leq z')$  then  $z \cdot z' \leq 0$  and  $z \cdot z' = -|z \cdot z'| \Rightarrow -(z \cdot z') = |z \cdot z'|$ , also  $-(z \cdot z') = (-z) \cdot z' = |z| \cdot |z'| \Rightarrow |z| \cdot |z'| = |z \cdot z'|$
4.  $(z < 0 \wedge z' < 0)$  then  $0 \leq -z \wedge 0 \leq -z' \Rightarrow 0 \leq (-z) \cdot (-z') = z \cdot z' \Rightarrow |z \cdot z'| = z \cdot z' = (-|z|) \cdot (-|z'|) = |z| \cdot |z'|$   $\square$

**Lemma 6.46.** *If  $m \in \mathbb{Z}$  is such that  $0 < m \Rightarrow 1 \leq m$*

**Proof.** If  $m \in \mathbb{Z}$  is such that  $0 < m$  then using 6.28 we have  $m \in \mathbb{N}_{\mathbb{Z}}$ . We can not have  $i_{\mathbb{N}}^{-1}(m) = 0$  [for then  $m = i_{\mathbb{N}}(i_{\mathbb{N}}^{-1}(m)) = i_{\mathbb{N}}(0) = \sim[(s(0), 1)] = \sim[(1, 1)] = 0$ ] and thus  $i_{\mathbb{N}}^{-1}(m) \neq 0$ . Using 4.47 we have then  $0 < i_{\mathbb{N}}^{-1}(m)$  which by 4.51 means that  $1 = s(0) \leq i_{\mathbb{N}}^{-1}(m) \Rightarrow 1 \leq i_{\mathbb{N}}^{-1}(m)$ . From the order preserving of  $i_{\mathbb{N}}$  we have then  $i_{\mathbb{N}}(1) \leq i_{\mathbb{N}}(i_{\mathbb{N}}^{-1}(m)) = m \Rightarrow \sim[(s(1), 1)] \leq m \Rightarrow 1 = \sim[(2, 1)] \leq m \Rightarrow 1 \leq m$   $\square$

**Theorem 6.47.** *If  $n, m \in \mathbb{Z}$ ,  $m \neq 0$  and  $n|m$  then  $n \leq |m|$*

**Proof.** From  $n|m$  there exists a  $q$  such that  $n \cdot q = m$  and as  $m \neq 0$  we must have  $q \neq 0$  (otherwise  $m = n \cdot 0 = 0$ ). Then we have the following cases to consider

1.  $(0 < m \wedge n \leq 0)$  in this case we have trivially  $n < m = |m| \Rightarrow n \leq |m|$
2.  $(0 < m \wedge 0 < n)$  we must have  $0 < q$  [if  $q \leq 0 \Rightarrow q < 0 \Rightarrow m = n \cdot q < n \cdot 0 = 0 \Rightarrow m < 0 < m \Rightarrow m < m$  a contradiction] so using the previous lemma we have then  $1 \leq q \Rightarrow n = n \cdot 1 \leq n \cdot q = m = |m| \Rightarrow n \leq |m|$
3.  $(m < 0 \wedge n \leq 0)$  here  $0 < -m = |m| \Rightarrow n < |m| \Rightarrow n \leq |m|$
4.  $(m < 0 \wedge 0 < n)$  then  $q < 0$  [otherwise if  $0 \leq q \Rightarrow 0 < q \Rightarrow 0 < n \cdot q = m \Rightarrow 0 < 0$  a contradiction] and we have  $0 < -q \Rightarrow 1 \leq -q \Rightarrow n = n \cdot 1 \leq n \cdot (-q) = -(n \cdot q) = -m = |m| \Rightarrow n \leq |m|$   $\square$

**Theorem 6.48. (Division Algorithm)** *If  $m, n \in \mathbb{Z}$  and  $n > 0$  then there exists a unique  $r \in \mathbb{N}$  with  $0 \leq r < n$  and a unique  $q \in \mathbb{Z}$  such that  $m = n \cdot q + r$*

**Proof.** Let  $A_{m,n} = \{m + n \cdot q \mid q \in \mathbb{Z} \wedge 0 \leq m + n \cdot q\} \subseteq \mathbb{N}_{\mathbb{Z}}$ . Using the Archimedean property of  $\mathbb{Z}$  (see 6.40) there exists a  $k \in \mathbb{N}_{\mathbb{Z}}$  such that  $k \cdot n > -m \Rightarrow m + n \cdot k > 0 \Rightarrow m + n \cdot k \in A_{m,n} \Rightarrow A_{m,n} \neq \emptyset$ . By the fact that  $(\mathbb{N}_{\mathbb{Z}}, \leq)$  is well-ordered we have then that  $r' = \min(A_{m,n})$  exists. From  $r' \in A_{m,n}$  we have then also  $0 \leq r'$  and a  $q' \in \mathbb{N}$  such that  $r' = m + n \cdot q'$ . For the relation between  $r'$  and  $n$  we have then the following possibilities

1.  $(r' = n)$  in this case we have  $m + n \cdot q' = r' = n \Rightarrow m = n \cdot (1 - q')$ , thus taking  $q = (1 - q')$  and  $r = 0 \Rightarrow 0 \leq r < n$  we have  $m = n \cdot q + r$
2.  $(r' < n)$  then  $r' = m + n \cdot q' \Rightarrow m = n \cdot (-q') + r'$  so if  $r = r' \Rightarrow 0 \leq r < n$  and  $q = -q'$  then  $m = n \cdot q + r$
3.  $(n < r')$  using 6.43 there exists a  $k \in \mathbb{N}_{\mathbb{Z}} \setminus \{0\} \Rightarrow 0 < k$  such that  $r' = n + k \Rightarrow n + k = m + n \cdot q' \Rightarrow 0 < k = m + n \cdot (q' - 1) \Rightarrow k \in A_{mn}$ . From  $r' = n + k \Rightarrow k < r' \Rightarrow r' \neq \min(A_{m,n})$  a contradiction. So this case does not apply.

Now to prove uniqueness assume that there exists another  $q'', r'' \in \mathbb{Z}$  with  $0 \leq r'' < n$  and  $m = n \cdot q'' + r''$ . We have then that  $n \cdot q'' + r'' = n \cdot q + r \Rightarrow n \cdot (q - q'') = r'' - r$ . We prove now by contradiction that  $r = r''$ , so assume that  $r \neq r''$  then we have either

1.  $(r < r'')$  then  $0 < r'' - r = n \cdot (q - q'')$  which as  $n > 0$  means that  $0 < q - q''$  [otherwise from  $q - q'' \leq 0$  we would have  $n \cdot (q - q'') \leq 0$  contradicting  $0 < n \cdot (q - q'')$ ] so using 6.46 we have  $1 \leq q - q''$ . Also from  $0 \leq r, r'' < n$  we have  $r'' - r < n - r \leq n \Rightarrow n \cdot (q - q'') < n \Rightarrow q - q'' < 1$  [if  $1 \leq q - q'' \Rightarrow n = n \cdot 1 \leq n \cdot (q - q'') < n \Rightarrow n < n$  a contradiction]. So we finally reach the conclusion that  $1 \leq q - q'' < 1 \Rightarrow 1 < 1$  a contradiction.

2. ( $r'' < r$ ) then  $0 < r - r'' = n \cdot (q'' - q)$  which as  $n > 0$  means that  $0 < q'' - q$  [otherwise from  $q'' - q \leq 0$  we would have  $n \cdot (q'' - q) \leq 0$  contradicting  $0 < n \cdot (q'' - q)$ ] so using 6.46 we have  $1 \leq q'' - q$ . Also from  $0 \leq r, r'' < n$  we have  $r - r'' < n - r'' \leq n \Rightarrow n \cdot (q'' - q) < n \Rightarrow q'' - q < 1$  [if  $1 \leq q'' - q \Rightarrow n = n \cdot 1 \leq n \cdot (q'' - q) < n \Rightarrow n < n$  a contradiction]. So we finally reach the conclusion that  $1 \leq q'' - q < 1 \Rightarrow 1 < 1$  a contradiction.  $\square$

**Definition 6.49.** Given  $n, m \in \mathbb{Z}$  then we say that  $n$  divides  $m$ , noted by  $n|m$  if there exists a  $q \in \mathbb{Z}$  such that  $q \cdot n = m$

**Note 6.50.** If  $n|m$  then we have also  $(-n)|m$  and this proves that then also  $(|d|)|m$

**Proof.** If  $n|m$  then there exists a  $q \in \mathbb{Z}$  such that  $m = n \cdot q = (-n) \cdot (-q)$  where  $-q \in \mathbb{Z}$  so that  $(-n)|m$   $\square$

**Definition 6.51.** If  $n, m \in \mathbb{Z}$  and  $n|m$  then there exists a unique (see 6.48)  $q \in \mathbb{Z}$  such that  $n \cdot q = m$ . This number is called the quotient of  $m$  and  $n$  and noted by  $m/n$ . So if  $n|m$  then there

**Definition 6.52.** Given  $n, m \in \mathbb{Z}$  then  $d$  is a common divisor for  $n$  and  $m$  if  $d|n$  and  $d|m$ . The greatest common divisor of  $m$  and  $n$  noted by  $\gcd(n, m) = \sup(D_{n,m}) \in \mathbb{N}_Z \setminus \{0\}$  if it exists, where  $D_{n,m} = \{d \in \mathbb{N}_Z \setminus \{0\} \mid d|n \wedge d|m\}$ . Note that  $1 \in D_{n,m} \Rightarrow D_{n,m} \neq \emptyset$ , note also that from 6.42 we have that if  $\gcd(n, m) = \sup(D_{n,m})$  exists then  $\gcd(n, m) \in D_{n,m} \Rightarrow \gcd(n, m)|n$  and  $\gcd(n, m)|m$

**Theorem 6.53.** Given  $n, m \in \mathbb{Z}$  then  $\gcd(n, m)$  exists  $\Leftrightarrow n \neq 0$  or  $m \neq 0$

**Proof.** Define  $D_{n,m} = \{d \in \mathbb{Z} \mid d|n \wedge d|m\}$  so that  $\gcd(n, m) = \sup(D_{n,m})$ . The proof is delivered in two parts

1. ( $\Rightarrow$ ) So assume that  $\gcd(n, m)$  exists and assume that  $n = 0 = m$ . Then if  $q \in \mathbb{Z}$  we have  $q \cdot 0 = 0 = n = m \Rightarrow q|n \wedge q|m \Rightarrow \mathbb{Z} = D_{n,m}$ . Now from  $0 < 1 \underset{\text{add } \gcd(n, m)}{\Rightarrow} \gcd(n, m) < \gcd(n, m) + 1 \in D_{n,m}$  contradicting the fact that  $\gcd(n, m)$  is a upper bound of  $D_{n,m}$ .
2. ( $\Leftarrow$ ) Suppose that  $n \neq 0$  or  $m \neq 0$  then we have the following cases
  - a. ( $n \neq 0$ ) from 6.47 we have then if  $d \in D_{n,m}$  then  $d|n \Rightarrow d < |n|$  so  $D_{n,m}$  has a upper bound and by conditional completeness (see 6.42) we have that  $\sup(D_{n,m})$  exists.
  - b. ( $m \neq 0$ ) from 6.47 we have then if  $d \in D_{n,m}$  then  $d|m \Rightarrow d < |m|$  so  $D_{n,m}$  has a upper bound and by conditional completeness (see 6.42) we have that  $\sup(D_{n,m})$  exists.  $\square$

**Theorem 6.54.** If  $n, m \in \mathbb{Z}$  with  $m \neq 0$  then by the above theorem  $\gcd(n, m)$  exists and from the fact that  $\gcd(n, m)|m$ ,  $\gcd(n, m)|n$  we can consider  $n/\gcd(n, m)$  and  $m/\gcd(n, m)$ . We have now that if  $d|(n/\gcd(n, m))$  and  $d|(m/\gcd(n, m))$  then  $d=1$  or  $d=-1$  furthermore  $\gcd(n/\gcd(n, m), m/\gcd(n, m))=1$

**Proof.** Take  $n' = n/\gcd(n, m)$  and  $m' = m/\gcd(n, m)$  then we have  $n' \cdot \gcd(n, m) = n$  and  $m' \cdot \gcd(n, m) = m$ . If  $d|n'$  then there exists a  $n''$  such that  $n'' \cdot d = n' \Rightarrow n'' \cdot d \cdot \gcd(n, m) = n' \cdot \gcd(n, m) = n$  or  $(d \cdot \gcd(n, m))|n \underset{0 < \gcd(n, m)}{\Rightarrow} (|d| \cdot \gcd(n, m))|n$ . Also if  $d|m'$  then there exists a  $m''$  such that  $m'' \cdot d = m' \Rightarrow m'' \cdot d \cdot \gcd(n, m) = m' \cdot \gcd(n, m) = m$  or  $(d \cdot \gcd(n, m))|m \underset{0 < \gcd(n, m)}{\Rightarrow} (|d| \cdot \gcd(n, m))|m$ . Now  $0 \leq |d|$  and as  $d \neq 0$  [otherwise  $m = m'' \cdot d \cdot \gcd(n, m) = 0$  contradicting  $m \neq 0$ ] we have  $0 < |d|$  and as  $0 < \gcd(n, m)$  we have  $0 < |d| \cdot \gcd(n, m)$  and thus  $|d| \cdot \gcd(n, m) \in D_{n,m} \Rightarrow |d| \cdot \gcd(n, m) \leq \gcd(n, m)$  and thus  $|d| \leq 1$  [otherwise  $\gcd(n, m) < |d| \cdot \gcd(n, m)$ ]. As we have also by 6.46 and  $0 < |d|$  that  $1 \leq |d|$  we must conclude that  $|d|=1$  and thus  $d=1$  or  $d=-1$ . As  $D_{n',m'} = \{d \in \mathbb{N}_Z \setminus \{0\} \mid d|n' \wedge d|m'\} = \{1\}$  we have  $\gcd(n', m')=1$ .  $\square$

**Definition 6.55.** Given  $z \in \mathbb{Z}$  then we say  $z$  is even if  $2|z$ . If  $z$  is not even then we call  $Z$  odd.

**Theorem 6.56.** Given  $z \in \mathbb{Z}$  then we have

1.  $z$  is even  $\Leftrightarrow \exists m \in \mathbb{Z} \vdash z = 2 \cdot m$

2.  $z$  is odd  $\Leftrightarrow \exists m \in \mathbb{Z} \vdash z = 2 \cdot m + 1$
3.  $z$  can not be even and odd

**Proof.**

1. This follows directly from the definition of  $2|z$
2. Using the division algorithm (see 6.48) there exists unique  $n, m \in \mathbb{Z}$  with  $z = 2 \cdot n + m$  where  $0 \leq m < 2$ . If now  $1 < m \Rightarrow 0 < m - 1 \stackrel{6.46}{\Rightarrow} 1 \leq m - 1 \Rightarrow 2 \leq m < 2 \Rightarrow 2 < 2$  a contradiction so we have  $0 \leq m \leq 1$ . We have then
  - a. ( **$z$  is odd**) then  $z$  is not even so we must have  $0 < m \Rightarrow 1 \leq m \leq 1 \Rightarrow m = 1 \Rightarrow z = 2 \cdot m + 1$
  - b. ( **$z = 2 \cdot m + 1$** ) from the uniqueness of  $n, m$  we can not have  $z = 2 \cdot m' + 0$  because we would have then the contradiction  $1 = 0$
3. If  $z$  is even and odd then  $\exists m, m' \in \mathbb{Z}$  such that  $2 \cdot m = z = 2 \cdot m' + 1$  hence  $2 \cdot (m - m') = 1$  we have now the following 3 cases to consider

**$m = m'$ .** then  $2 \cdot 0 = 1 \Rightarrow 0 = 1$  a contradiction

**$m < m'$ .** then  $m + 1 \leq m' \Rightarrow m - m' \leq -1 \Rightarrow 2 \cdot (m - m') \leq -2 \Rightarrow 1 \leq -2$  a contradiction

**$m' < m$ .** then  $m' + 1 \leq m \Rightarrow 1 \leq m - m' \Rightarrow 2 \leq 2 \cdot (m - m') = 1 \Rightarrow 2 \leq 1$  a contradiction

so in all cases we have a contradiction.  $\square$

**Theorem 6.57.** *Given  $m \in \mathbb{Z}$  such that  $z \cdot z$  is even then we have that  $z$  is even*

**Proof.** We prove this by contradiction, so assume that  $z$  is odd then by the above theorem we have  $z = 2 \cdot n + 1$  and thus  $z \cdot z = (2 \cdot n + 1) \cdot (2 \cdot n + 1) = 2 \cdot 2 \cdot n \cdot n + 2 \cdot n \cdot 1 + 1 \cdot 2 \cdot n + 1 \cdot 1 = 2 \cdot (2 \cdot n) + 2 \cdot n + 2 \cdot n + 1 = 2 \cdot (2 \cdot n + n + n) + 1$  meaning that  $z \cdot z$  is odd contradicting the fact that  $z \cdot z$  is even.  $\square$

### 6.3 Denumerability of the integers

We are going to prove now that the set of integers is **denumerable**.

**Lemma 6.58.**  $\mathbb{N}_{\mathbb{Z}}$  is **denumerable**

**Proof.** This follows directly from 6.22 (3) where it is proved that  $i_{\mathbb{Z}}: \mathbb{N} \rightarrow \mathbb{N}_{\mathbb{Z}}$  is a bijection.  $\square$

**Theorem 6.59.**  $\mathbb{Z}$  is **denumerable**.

**Proof.** Using 6.24 we have that  $\mathbb{Z} = \mathbb{N}_{\mathbb{Z}} \cup (-\mathbb{N}_{\mathbb{Z}})$  where  $\mathbb{N}_{\mathbb{Z}} = \{-n \mid n \in \mathbb{N}_{\mathbb{Z}}\}$ . We can now construct the trivial bijection  $f: \mathbb{N}_{\mathbb{Z}} \rightarrow (-\mathbb{N}_{\mathbb{Z}})$  defined by  $n \rightarrow -n$ . So we have  $\mathbb{N}_{\mathbb{Z}} \approx (-\mathbb{N}_{\mathbb{Z}})$  and by the previous lemma that  $\mathbb{N}_{\mathbb{Z}} \approx \mathbb{N}$  and thus  $(-\mathbb{N}_{\mathbb{Z}}) \approx \mathbb{N}$ . This means that  $\mathbb{Z}$  is the union of two denumerable sets and thus by 5.61 that  $\mathbb{Z}$  is denumerable.  $\square$



# Chapter 7

## The rational numbers

### 7.1 Definition and arithmetic's

Just like we defined integer numbers using a equivalence relation on  $\mathbb{N}$  we define the rational numbers using a equivalence relation in  $\mathbb{Z}$

**Definition 7.1.**  $\mathbb{Z}_0 = \mathbb{Z} \setminus \{0\}$  (here  $0 = \sim[(1, 1)]$ )

**Definition 7.2.**  $\simeq \subseteq (\mathbb{Z} \times \mathbb{Z}_0) \times (\mathbb{Z} \times \mathbb{Z}_0)$  is the relation defined on  $\mathbb{Z} \times \mathbb{Z}_0$  by  $\simeq = \{( (n, m), (r, k) ) \in \mathbb{Z} \times \mathbb{Z}_0 \mid n \cdot k = m \cdot r \}$ . In other words if  $(n, m), (r, k) \in \mathbb{Z} \times \mathbb{Z}_0$  then  $(n, m) \simeq (r, k)$  iff  $n \cdot k = m \cdot r$

**Theorem 7.3.**  $\simeq$  is a equivalence relation on  $\mathbb{Z} \times \mathbb{Z}_0$

**Proof.**

1. **(reflexive)** If  $(n, m) \in \mathbb{Z} \times \mathbb{Z}_0$  then  $n \cdot m = m \cdot n \xrightarrow{6.9} (n, m) \simeq (n, m)$
2. **(symmetry)** If  $(n, m), (r, k) \in \mathbb{Z} \times \mathbb{Z}_0$  and  $(n, m) \simeq (r, k)$  then  $n \cdot k = m \cdot r \xrightarrow{6.9} r \cdot m = k \cdot n \Rightarrow (r, k) \simeq (n, m)$
3. **(transitive)** If  $(i, j), (n, m), (r, k) \in \mathbb{Z} \times \mathbb{Z}_0$  and  $(i, j) \simeq (n, m)$  and  $(n, m) \simeq (r, k)$  then we have  $i \cdot m = j \cdot n$  and  $n \cdot k = m \cdot r$  then  $(i \cdot m) \cdot k = (j \cdot n) \cdot k = j \cdot (n \cdot k) = j \cdot (m \cdot r) \Rightarrow (i \cdot k) \cdot m = (j \cdot r) \cdot m \xrightarrow{m \neq 0, 6.15} i \cdot k = j \cdot r \Rightarrow (i, j) \simeq (k, r)$   $\square$

**Definition 7.4.** We define the set of rationals  $\mathbb{Q}$  by  $\mathbb{Q} / \simeq$ , we note  $\simeq[(n, k)] \in \mathbb{Q} / \simeq$  as  $\frac{n}{k}$ ,  $n$  is called the denominator,  $k$  is the nominator. We have then that  $\frac{n}{k} = \frac{n'}{k'} \text{ iff } (n, k) \sim (n', k') \Leftrightarrow n \cdot k' = k \cdot n'$

**Theorem 7.5.**  $+: \mathbb{Q} \times \mathbb{Q} \rightarrow \mathbb{Q}$  defined by  $(\frac{n}{m}, \frac{r}{k}) \rightarrow \frac{n}{m} + \frac{r}{k} = \frac{n \cdot k + m \cdot r}{m \cdot k}$  is a function

**Proof.** First note that as  $m, k \in \mathbb{Z}_0$  then  $m \cdot k \in \mathbb{Z}_0$  [if  $m \cdot k = 0 \xrightarrow{6.9} m = n = 0$  contradicting  $m, k \in \mathbb{Z}_0$ ] we have thus that  $\frac{n \cdot k + m \cdot r}{m \cdot k} \in \mathbb{Q}$

Second, suppose that  $\frac{n}{m} = \frac{n'}{m'}$  and  $\frac{r}{k} = \frac{r'}{k'}$  then we have  $n \cdot m' = m \cdot n'$  and  $r \cdot k' = k \cdot r'$ . Now

$$\begin{aligned}
 (n \cdot k + m \cdot r) \cdot (m' \cdot k') &= (n \cdot k) \cdot (m' \cdot k') + (m \cdot r) \cdot (m' \cdot k') \\
 &= (n \cdot m') \cdot (k \cdot k') + (r \cdot k') \cdot (m \cdot m') \\
 &= (m \cdot n') \cdot (k \cdot k') + (k \cdot r') \cdot (m \cdot m') \\
 &= (n' \cdot k') \cdot (m \cdot k) + (m' \cdot r') \cdot (m \cdot k) \\
 &= (m \cdot k) \cdot (n' \cdot k' + m' \cdot r') \\
 &\Rightarrow \\
 \frac{n \cdot k + m \cdot r}{m \cdot k} &= \frac{n' \cdot k' + m' \cdot r'}{m' \cdot k'}
 \end{aligned}$$

meaning that the function is well defined.  $\square$

**Theorem 7.6.** If  $k \in \mathbb{Z}_0$  and  $\frac{a}{b} \in \mathbb{Q}$  then  $\frac{a}{b} = \frac{a \cdot k}{b \cdot k}$

**Proof.**  $a \cdot (b \cdot k) = (b \cdot a) \cdot k = b \cdot (a \cdot k) \Rightarrow \frac{a}{b} = \frac{a \cdot k}{b \cdot k}$  □

**Theorem 7.7.**  $\langle \mathbb{Q}, + \rangle$  forms a **abelian group** with neutral element  $0 = \frac{0}{1}$ . Note that the same symbol 0 is used for neutral elements in  $\mathbb{Z}$  and  $\mathbb{Q}$ , context tell's us always what we mean by 0.

**Proof.**

1. **(associative)** If  $\frac{a}{b}, \frac{c}{d}, \frac{e}{f} \in \mathbb{Q}$  then we have

$$\begin{aligned} \frac{a}{b} + \left( \frac{c}{d} + \frac{e}{f} \right) &= \frac{a}{b} + \frac{c \cdot f + d \cdot e}{d \cdot f} \\ &= \frac{a \cdot (d \cdot f) + b \cdot (c \cdot f + d \cdot e)}{b \cdot (d \cdot f)} \\ &= \frac{(a \cdot d) \cdot f + (b \cdot c) \cdot f + (b \cdot d) \cdot e}{(b \cdot d) \cdot f} \\ &= \frac{(a \cdot d + b \cdot c) \cdot f + (b \cdot d) \cdot e}{(b \cdot d) \cdot f} \\ &= \frac{a \cdot d + b \cdot c}{b \cdot d} + \frac{e}{f} \\ &= \left( \frac{a}{b} + \frac{c}{d} \right) + \frac{e}{f} \end{aligned}$$

2. **(neutral element)** We define the neutral element in  $\mathbb{Q}$  to be  $0 = \frac{0}{1}$  (note that one  $0 \in \mathbb{Z}$  and the other is in  $\mathbb{Q}$  context will always tell us which is which). If  $\frac{a}{b} \in \mathbb{Q}$  then we have

$$\begin{aligned} \frac{a}{b} + 0 &= \frac{a}{b} + \frac{0}{1} \\ &= \frac{a \cdot 1 + b \cdot 0}{b \cdot 1} \\ &= \frac{a + 0}{b} \\ &= \frac{a}{b} \\ 0 + \frac{a}{b} &= \frac{0}{1} + \frac{a}{b} \\ &= \frac{0 \cdot b + 1 \cdot a}{1 \cdot b} \\ &= \frac{a}{b} \end{aligned}$$

3. **(inverse)** If  $\frac{a}{b} \in \mathbb{Q}$  then we prove that the inverse is  $\frac{-a}{b}$ .

**Proof.** We have

$$\begin{aligned} \frac{a}{b} + \frac{-a}{b} &= \frac{a \cdot b + b \cdot (-a)}{b \cdot b} \\ &= \frac{b \cdot (a + (-a))}{b \cdot b} \\ &= \frac{b \cdot 0}{b \cdot b} \\ &\stackrel{7.6}{=} \frac{0}{b} \\ &= \frac{b \cdot 0}{b \cdot 1} \\ &= \frac{0}{1} = 0 \\ \frac{-a}{b} + \frac{a}{b} &= \frac{(-a) \cdot b + b \cdot a}{b \cdot b} \end{aligned}$$

$$\begin{aligned}
&= \frac{b \cdot (-a+a)}{b \cdot b} \\
&= \frac{b \cdot 0}{b \cdot b} \\
&= 0
\end{aligned}$$

□

4. **(commutative)** If  $\frac{a}{b}, \frac{c}{d} \in \mathbb{Q}$  then

$$\begin{aligned}
\frac{a}{b} + \frac{c}{d} &= \frac{a \cdot d + b \cdot c}{b \cdot d} \\
&= \frac{c \cdot b + d \cdot a}{d \cdot b} \\
&= \frac{c}{d} + \frac{a}{b}
\end{aligned}$$

□

**Notation 7.8.** Whenever we write  $a - b$  we mean  $a + (-b)$

**Theorem 7.9.**  $\cdot: \mathbb{Q} \times \mathbb{Q} \rightarrow \mathbb{Q}$  defined by  $(\frac{n}{m}, \frac{r}{k}) \mapsto \frac{n}{m} \cdot \frac{r}{k} = \frac{n \cdot r}{m \cdot k}$  is a function.

**Proof.** First as  $m, k \in \mathbb{Z}_0$  we have that  $m \cdot k \in \mathbb{Z}_0$  [if  $m \cdot k = 0 \Rightarrow m = k = 0$  by 6.9] so we have that  $\frac{n \cdot r}{m \cdot k} \in \mathbb{Q}$ . Secondly if  $\frac{n}{m} = \frac{n'}{m'}$  and  $\frac{r}{k} = \frac{r'}{k'}$  then we have  $n \cdot m' = m \cdot n'$  and  $r \cdot k' = k \cdot r'$ . We have then

$$\begin{aligned}
(n \cdot r) \cdot (m' \cdot k') &= (n \cdot m') \cdot (r \cdot k') \\
&= (m \cdot n') \cdot (k \cdot r') \\
&= (m \cdot k) \cdot (n' \cdot r') \\
&\Rightarrow \\
\frac{n \cdot r}{m \cdot k} &= \frac{n' \cdot r'}{m' \cdot k'}
\end{aligned}$$

proving that the function is well-defined. □

**Theorem 7.10.**  $\langle \mathbb{Q}, +, \cdot \rangle$  forms a **field**, the neutral element is  $1 = \frac{1}{1}$  and if  $\frac{a}{b} \in \mathbb{Q} \setminus \{0\}$  the inverse for multiplication is  $\frac{b}{a}$  (so  $(\frac{a}{b})^{-1} = \frac{b}{a}$ ). Note that 1 stands here for the unit in  $\mathbb{Z}$  and  $\mathbb{Q}$ , context tell's us which is which.

**Proof.** By 7.7 we have that  $\langle \mathbb{Q}, + \rangle$  is a abelian group, we must thus only prove the additional properties (see 3.29) of the multiplication in a field.

1. **(Distributive)** If  $\frac{a}{b}, \frac{c}{d}, \frac{e}{f} \in \mathbb{Q}$  then we have

$$\begin{aligned}
\frac{a}{b} \cdot \left( \frac{c}{d} + \frac{e}{f} \right) &= \frac{a}{b} \cdot \left( \frac{c \cdot f + e \cdot d}{d \cdot f} \right) \\
&= \frac{a \cdot (c \cdot f + e \cdot d)}{b \cdot (d \cdot f)} \\
&\stackrel{7.6}{=} \frac{b \cdot (a \cdot (c \cdot f + e \cdot d))}{b \cdot (b \cdot (d \cdot f))} \\
&= \frac{(a \cdot c) \cdot (b \cdot f) + (a \cdot e) \cdot (b \cdot d)}{(b \cdot d) \cdot (b \cdot f)} \\
&= \frac{a \cdot c}{b \cdot d} + \frac{a \cdot e}{b \cdot f} \\
&= \left( \frac{a}{b} \cdot \frac{c}{d} \right) + \left( \frac{a}{b} \cdot \frac{e}{f} \right)
\end{aligned}$$

2. (**Neutral**) The neutral element  $1 = \frac{1}{1}$  (note that we have two 1's here context will always tell if  $1 \in \mathbb{Z}$  or  $1 \in \mathbb{Q}$ ), we have then if  $\frac{a}{b} \in \mathbb{Q}$  that

$$\begin{aligned} 1 \cdot \frac{a}{b} &= \frac{1}{1} \cdot \frac{a}{b} \\ &= \frac{1 \cdot a}{1 \cdot b} \\ &= \frac{a}{b} \\ &= \frac{a \cdot 1}{b \cdot 1} \\ &= \frac{a}{b} \cdot \frac{1}{1} \\ &= \frac{a}{b} \cdot 1 \end{aligned}$$

3. (**Commutative**) If  $\frac{a}{b}, \frac{c}{d} \in \mathbb{Q}$  then we have

$$\begin{aligned} \frac{a}{b} \cdot \frac{c}{d} &= \frac{a \cdot c}{b \cdot d} \\ &= \frac{c \cdot a}{d \cdot b} \\ &= \frac{c}{d} \cdot \frac{a}{b} \end{aligned}$$

4. (**Associative**) If  $\frac{a}{b}, \frac{c}{d}, \frac{e}{f} \in \mathbb{Q}$  then we have

$$\begin{aligned} \frac{a}{b} \cdot \left( \frac{c}{d} \cdot \frac{e}{f} \right) &= \frac{a}{b} \cdot \frac{c \cdot e}{d \cdot f} \\ &= \frac{a \cdot (c \cdot e)}{b \cdot (d \cdot f)} \\ &= \frac{(a \cdot c) \cdot e}{(b \cdot d) \cdot f} \\ &= \frac{a \cdot c}{b \cdot d} \cdot \frac{e}{f} \\ &= \left( \frac{a}{b} \cdot \frac{c}{d} \right) \cdot \frac{e}{f} \end{aligned}$$

5. (**Inverse for non zero element**) If  $\frac{a}{b} \in \mathbb{Q} \setminus \{0\}$  then  $\frac{a}{b} \neq \frac{0}{1}$ . Now if  $a = 0$  then  $a \cdot 1 = 0 = b \cdot 0 \Rightarrow \frac{a}{b} = \frac{0}{1}$  a contradiction, so we must have  $a \neq 0$  and thus  $\frac{b}{a} \in \mathbb{Q}$ . we prove now that  $\left(\frac{a}{b}\right)^{-1} = \frac{b}{a}$

$$\begin{aligned} \frac{b}{a} \cdot \frac{a}{b} &= \frac{b \cdot a}{a \cdot b} \\ &= \frac{a \cdot b}{a \cdot b} \\ &= \frac{(a \cdot b) \cdot 1}{(a \cdot b) \cdot 1} \\ &\stackrel{7.6}{=} \frac{1}{1} \\ &= 1 \end{aligned}$$

□

## 7.2 Power in $\mathbb{Q}$

**Definition 7.11.** As  $\langle \mathbb{Q}, \cdot \rangle$  is a abelian semi-group we have by 4.22 that given  $a \in \mathbb{Q}$  and  $n \in \mathbb{N}$  that there exists a  $a^n$  such that

$$\begin{aligned} a^0 &= 1 \\ a^{n+1} &= a^n \cdot a \underset{\text{abelian}}{=} a \cdot a^n \end{aligned}$$

**Theorem 7.12.** If  $n, n' \in \mathbb{N}$  and  $a \in \mathbb{Q}$  then  $a^{n'+n} = a^{n'} \cdot a^n$

**Proof.** We prove this by induction on  $n$ . So let  $X = \{n \in \mathbb{N} | a^{n'+n} = a^{n'} \cdot a^n\}$  then we have

1. If  $n = 0$  then  $a^{n'+n} = a^{n'+0} = a^{n'} = a^{n'} \cdot 1 = a^{n'} \cdot a^0 \Rightarrow 0 \in X$
2. If  $n \in X$  then  $a^{n'+(n+1)} = a^{(n'+n)+1} = a^{(n'+n)} \cdot a \underset{n \in X}{=} (a^{n'} \cdot a^n) \cdot a = a^{n'} \cdot (a^n \cdot a) = a^{n'} \cdot a^{n+1}$  and thus  $n+1 \in X$

Using mathematical induction (see 4.10) we have  $X = \mathbb{N}$  proving the theorem  $\square$

**Theorem 7.13.** If  $n \in \mathbb{N}$  and  $a, b \in \mathbb{Q}$  then we have  $(a \cdot b)^n = a^n \cdot b^n$

**Proof.** We prove this by induction so take  $\mathcal{S} = \{n \in \mathbb{N} | (a \cdot b)^n = a^n \cdot b^n\}$  then we have

$$n = 0. \text{ then } (a \cdot b)^0 = 1 = 1 \cdot 1 = a^0 \cdot b^0$$

$$n \in \mathcal{S} \Rightarrow n + 1 \in \mathcal{S}. \text{ then } (a \cdot b)^{n+1} = (a \cdot b)^n \cdot (a \cdot b) = (a^n \cdot b^n) \cdot (a^n \cdot a) \cdot (b^n \cdot b) = a^{n+1} \cdot b^{n+1} \quad \square$$

**Theorem 7.14.** In  $\mathbb{Q}$  we have

$$\begin{aligned} 0^n &= 0 \text{ (if } n \neq 0) \\ 1^n &= 1 \\ (-1)^n &= -1 \text{ or } 1 \\ (-1)^{2 \cdot n} &= 1 \\ (-1)^{2 \cdot n+1} &= -1 \end{aligned}$$

**Proof.**

1. If  $n \neq 0 \Rightarrow \exists m \in \mathbb{N} \vdash n = m+1$  then  $0^n = 0^{(m+1)} = 0^m \cdot 0 = 0$
2.  $1^n = 1$  is proved by induction on  $n$ , let  $X = \{n \in \mathbb{N} | 1^n = 1\}$  then
  - a.  $1^0 = 1 \Rightarrow 0 \in X$
  - b. If  $n \in X \Rightarrow 1^{n+1} = 1^n \cdot 1 \underset{n \in X}{=} 1 \cdot 1 = 1 \Rightarrow n+1 \in X$   
so  $X = \mathbb{N}$
3.  $(-1)^n = \pm 1$  is proved by induction on  $n$ , let  $X = \{n \in \mathbb{N} | (-1)^n = -1 \text{ or } 1\}$  then
  - a.  $(-1)^0 = 1 \Rightarrow 0 \in X$
  - b. If  $n \in X$  then  $(-1)^{n+1} = (-1)^n \cdot (-1) \underset{n \in X}{=} (-1) \cdot (-1) \vee 1 \cdot (-1) = 1 \vee -1 \Rightarrow n+1 \in X$   
so  $X = \mathbb{N}$
4.  $(-1)^{2 \cdot n} = (-1)^{(1+1) \cdot n} = (-1)^{n+n} = (-1)^n \cdot (-1)^n \underset{(3)}{=} (-1) \cdot (-1) \text{ or } 1 \cdot 1 = 1$
5.  $(-1)^{2 \cdot n+1} = (-1)^{2 \cdot n} \cdot (-1) \underset{(4)}{=} 1 \cdot (-1) = -1 \quad \square$

## 7.3 Order relation

**Theorem 7.15.**  $\text{sign}: \mathbb{Q} \rightarrow \{1, -1\} \subseteq \mathbb{N}$  defined by

$$\text{sign}\left(\frac{a}{b}\right) = \begin{cases} 1 & \text{if } 0 \leq a \cdot b \\ -1 & \text{if } \neg(0 \leq a \cdot b) \end{cases}$$

is a well defined function.

**Proof.** If  $\frac{a}{b} = \frac{a'}{b'}$  then we must prove that  $\text{sign}\left(\frac{a}{b}\right) = \text{sign}\left(\frac{a'}{b'}\right)$ . As  $\frac{a}{b} = \frac{a'}{b'}$  then  $a \cdot b' = b \cdot a'$ . We have now the following cases

1.  $\text{sign}\left(\frac{a}{b}\right) = 1 \Rightarrow 0 \leq a \cdot b$  then by 6.39 we have  $0 \leq b' \cdot b'$  giving by 6.38 that  $0 \cdot (b' \cdot b') \leq (a \cdot b) \cdot (b' \cdot b') \Rightarrow 0 \leq (a \cdot b') \cdot (b \cdot b') \Rightarrow 0 \leq (b \cdot a') \cdot (b \cdot b') \Rightarrow 0 \leq (a' \cdot b') \cdot (b \cdot b)$ . We proceed now by contradiction that  $0 \leq a' \cdot b'$ , so assume that  $a' \cdot b' < 0 \xrightarrow{6.39, 6.37} (a' \cdot b') \cdot (b \cdot b) < 0$  contradicting  $0 \leq (a' \cdot b') \cdot (b \cdot b)$  so  $0 \leq a' \cdot b' \Rightarrow \text{sign}\left(\frac{a'}{b'}\right) = 1$
2.  $\text{sign}\left(\frac{a}{b}\right) = -1 \Rightarrow \neg(0 \leq a \cdot b) \xrightarrow{2.148, 6.26} a \cdot b < 0 \xrightarrow{6.39, 6.37} (a \cdot b) \cdot (b' \cdot b') < 0 \Rightarrow (a \cdot b') \cdot (b \cdot b') < 0 \Rightarrow (a' \cdot b) \cdot (b \cdot b') < 0 \Rightarrow (a' \cdot b') \cdot (b \cdot b) < 0$ . Suppose now that  $0 \leq (a' \cdot b')$  then by 6.38, 6.39 we have  $0 \leq (a' \cdot b') \cdot (b \cdot b) < 0 \Rightarrow 0 < 0$  a contradiction, so we must have  $(a' \cdot b') < 0$  or  $\text{sign}\left(\frac{a'}{b'}\right) = -1$   $\square$

**Lemma 7.16.** If  $\text{sign}(q) = -1$  then  $\text{sign}(-q) = 1$

**Proof.** If  $q = \frac{a}{b}$  then from  $\text{sign}(q) = 1$  we have  $a \cdot b < 0$  and thus by 6.27 we have  $-0 < -(a \cdot b) \Rightarrow 0 < (-a) \cdot b \Rightarrow 0 \leq (-a) \cdot b \Rightarrow \text{sign}\left(\frac{-a}{b}\right) = 1 \Rightarrow \text{sign}(-q) = 1$   $\square$

**Definition 7.17. (Order Relation  $\mathbb{Q}$ )**  $\leq \subseteq \mathbb{Q} \times \mathbb{Q}$  is defined by  $\leq = \{(q, r) \in \mathbb{Q} \mid \text{sign}(r + (-q)) = 1\}$  (note that this definition is independent of the representation of a rational number as sign is a well-defined function). So we have  $q \leq r \Leftrightarrow \text{sign}(r + (-q)) = 1$

**Lemma 7.18.** If  $q \in \mathbb{Q}$  with  $0 \leq q$  and  $q \leq 0$  then  $q = 0$

**Proof.** Let  $q = \frac{a}{b}$  then  $-q = \frac{-a}{b}$ . From  $0 \leq q \wedge q \leq 0$  we have  $1 = \text{sign}(q + (-0)) = \text{sign}(q)$  and  $1 = \text{sign}(0 + (-q)) = \text{sign}(-q)$  and thus  $\text{sign}\left(\frac{a}{b}\right) = 1 \wedge \text{sign}\left(\frac{-a}{b}\right) = 1$ . This gives  $0 \leq a \cdot b \wedge 0 \leq (-a) \cdot b \xrightarrow{6.27} 0 \leq a \cdot b \wedge (-((-a) \cdot b)) \leq -0 \Rightarrow 0 \leq a \cdot b \leq 0 \Rightarrow a \cdot b = 0 \xrightarrow{b \neq 0 \wedge 6.9} a = 0 \Rightarrow q = \frac{0}{b} = \frac{0}{1} = 0$   $\square$

**Lemma 7.19.** If  $q, r, s \in \mathbb{Q}$  with  $q \leq r \Rightarrow q + s \leq r + s$

**Proof.**  $q \leq r \Rightarrow 1 = \text{sign}(r + (-q)) = \text{sign}(r + (s + (-s)) + (-q)) = \text{sign}((r + s) + (-(-q + s))) \Rightarrow q + s \leq r + s$   $\square$

**Lemma 7.20.** If  $q, r \in \mathbb{Q}$  with  $0 \leq q$  and  $0 \leq r$  then  $0 \leq q + r$

**Proof.** Let  $q = \frac{a}{b}$  and  $r = \frac{a'}{b'}$ . From  $0 \leq q$  we have  $1 = \text{sign}(q + (-0)) = \text{sign}(q)$  and from  $0 \leq r$  we have  $1 = \text{sign}(r + (-0)) = \text{sign}(r)$ . This gives us  $0 \leq a \cdot b$  and  $0 \leq a' \cdot b'$ . Now  $\frac{a}{b} + \frac{a'}{b'} = \frac{a \cdot b' + b \cdot a'}{b \cdot b'}$  and  $\square$

**Theorem 7.21.**  $\langle \mathbb{Q}, \leq \rangle$  forms a partially ordered set that is fully-ordered

**Proof.**

1. **(Reflexivity)** If  $q \in \mathbb{Q} \Rightarrow q - q = 0 = \frac{0}{1} \Rightarrow \text{sign}(q - q) = \text{sign}\left(\frac{0}{1}\right) \xrightarrow[0 \leq 0 = 0 \cdot 1]{0 \leq 0 = 0 \cdot 1} 1 \Rightarrow q \leq q$
2. **(Anti-symmetry)** If  $q \leq r$  and  $r \leq q$  then we have using 7.19 that  $q + (-q) \leq r + (-q)$  and  $r + (-r) \leq q + (-r) \Rightarrow 0 \leq r + (-q)$  and  $0 \leq q + (-r) = -(r + (-q)) \Rightarrow 0 \leq r + (-q) \leq 0 \xrightarrow{7.18} r + (-q) = 0 \Rightarrow r = q$
3. **(Transitivity)** If  $q \leq r$  and  $r \leq s$  then using 7.19 we have  $0 = q + (-q) \leq r + (-q)$  and  $0 = r + (-r) \leq s + (-r)$ , using 7.20 we have then that  $0 \leq (r + (-q)) + (s + (-r)) = s + (-q)$  which using 7.19 again gives  $q = 0 + q \leq (s + (-q)) + q \Rightarrow q \leq s$  proving transitivity.
4. **(Fully-ordered)** Let  $q, r \in \mathbb{Q}$  then for  $q + (-r)$  we have either
  - a.  $\text{sign}(q + (-r)) = 1 \Rightarrow r \leq q$
  - b.  $\text{sign}(q + (-r)) = -1$  then by 7.16 we have  $\text{sign}(-(q + (-r))) = 1 \Rightarrow \text{sign}(r + (-q)) = 1 \Rightarrow q \leq r$   $\square$

**Theorem 7.22.** If  $q, r \in \mathbb{Q}$  with  $q \leq r$  then we have  $-r \leq -q$

**Proof.** If  $q \leq r$  then  $\text{sign}(r-q) = 1$ . Now  $(-q) - (-r) = (-q) + (-(-r)) = (-q) + r = r + (-q) = r - q$  and this  $\text{sign}((-q) - (-r)) = \text{sign}(r-q) = 1 \Rightarrow -r \leq -q$   $\square$

**Theorem 7.23.** If  $q \in \mathbb{Q} \Rightarrow q < q+1$  and  $q-1 < q$

**Proof.** First in  $\mathbb{N}$  we have  $0 < 1$  so that using the injectivity and order preserving off  $i_{\mathbb{N}}$  we have  $0 = \sim(s(0), 1) = i_{\mathbb{N}}(0) < i_{\mathbb{N}}(1) = \sim(s(1), 1) = 1$  so that  $0 < 1$  in  $\mathbb{Z}$ . In the same way we have that  $0 = \frac{0}{1} = i_{\mathbb{Z}}(0) < i_{\mathbb{Z}}(1) = \frac{1}{1} = 1$  so that  $0 < 1$  in  $\mathbb{Q}$ . Now we have  $0 < 1 \Rightarrow 0 \leq 1 \xrightarrow{7.19} 0+q \leq 1+q \Rightarrow q \leq q+1$  if now  $q = q+1$  then  $0 = q+(-q) = q+(-q)+1 = 1$  contradicting  $0 < 1$  so we have that  $q < q+1$ . Finally  $q < q+1 \Rightarrow q \leq q+1 \Rightarrow q+(-1) \leq q+1+(-1) \Rightarrow q-1 \leq q$ , if now  $q-1 = q$  then  $-1 = 0 \Rightarrow 0 = 1$  a contradiction, so we have  $q-1 < q$ .  $\square$

**Definition 7.24.**  $\mathbb{Z}_{\mathbb{Q}} = \left\{ \frac{a}{1} \mid a \in \mathbb{Z} \right\}$

**Theorem 7.25.**  $\langle \mathbb{Z}_{\mathbb{Q}}, +, \cdot \rangle$  is a sub-ring of  $\langle \mathbb{Q}, +, \cdot \rangle$  and  $i_{\mathbb{Z}}: \mathbb{Z} \rightarrow \mathbb{Z}_{\mathbb{Q}}$  defined by  $n \rightarrow \frac{n}{1}$  is a ring-isomorphism that is order preserving

**Proof.** First we prove that  $\langle \mathbb{Z}_{\mathbb{Q}}, +, \cdot \rangle$  is a sub ring of  $\mathbb{Q}$ . If  $m, n \in \mathbb{Z}_{\mathbb{Q}}$  then there exists a  $m', n' \in \mathbb{Z}$  then  $m = \frac{m'}{1}, n = \frac{n'}{1}$  we have then:

1.  $m+n = \frac{m'}{1} + \frac{n'}{1} = \frac{m' \cdot 1 + n' \cdot 1}{1 \cdot 1} = \frac{m'+n'}{1} \in \mathbb{Z}_{\mathbb{Q}}$
2.  $-\frac{n}{1} = \frac{-n}{1} \in \mathbb{Z}_{\mathbb{Q}}$
3.  $m \cdot n = \frac{m' \cdot n'}{1 \cdot 1} = \frac{m' \cdot n'}{1} \in \mathbb{Z}_{\mathbb{Q}}$
4.  $1 = \frac{1}{1} \in \mathbb{Z}_{\mathbb{Q}}$
5.  $0 = \frac{0}{1} \in \mathbb{Z}_{\mathbb{Q}}$

proving that  $\langle \mathbb{Z}_{\mathbb{Q}}, +, \cdot \rangle$  is a sub-ring of  $\langle \mathbb{Q}, +, \cdot \rangle$ . Next prove that  $i_{\mathbb{Z}}$  is a ring isomorphism

1. If  $m, n \in \mathbb{Z}$  then  $i_{\mathbb{Z}}(m+n) = \frac{m+n}{1} = \frac{m \cdot 1 + n \cdot 1}{1 \cdot 1} = \frac{m}{1} + \frac{n}{1} = i_{\mathbb{Z}}(m) + i_{\mathbb{Z}}(n)$
2. If  $m, n \in \mathbb{Z}$  then  $i_{\mathbb{Z}}(m \cdot n) = \frac{m \cdot n}{1} = \frac{m \cdot n}{1 \cdot 1} = \frac{m}{1} \cdot \frac{n}{1} = i_{\mathbb{Z}}(m) \cdot i_{\mathbb{Z}}(n)$
3.  $i_{\mathbb{Z}}(1) = \frac{1}{1} = 1$  (note that 1 stands for the unit in  $\mathbb{Z}$  and  $\mathbb{Q}$ )

Secondly we have to prove that  $i_{\mathbb{Z}}$  is a bijection.

1. **(Injectivity)** If  $i_{\mathbb{Z}}(m) = i_{\mathbb{Z}}(n) \Rightarrow \frac{m}{1} = \frac{n}{1} \Rightarrow m \cdot 1 = n \cdot 1 \Rightarrow m = n$
2. **(Surjectivity)** If  $a \in \mathbb{Z}$  then  $\exists a' \in \mathbb{Q}$  with  $a = \frac{a'}{1} = i_{\mathbb{Z}}(a')$

Finally to prove that  $i_{\mathbb{Z}}$  is order preserving, assume that  $m, n \in \mathbb{Z}$  with  $n \leq m$  then we have  $i_{\mathbb{Z}}(m) + (-i_{\mathbb{Z}}(n)) = \frac{m}{1} + \frac{-n}{1} = \frac{m \cdot 1 + (-n) \cdot 1}{1 \cdot 1} = \frac{m-n}{1}$ , now from  $n \leq m$  we have  $0 \leq m-n$  so that  $(m-n) \cdot 1 = m-n \geq 0 \Rightarrow \text{sign}(i_{\mathbb{Z}}(m) + (-i_{\mathbb{Z}}(n))) = \text{sign}(\frac{m-n}{1}) = 1 \Rightarrow i_{\mathbb{Z}}(n) \leq i_{\mathbb{Z}}(m)$ .  $\square$

**Lemma 7.26.** If  $q \in \mathbb{Q}$  with  $q > 0$  then there exists  $a, b \in \mathbb{N}_{\mathbb{Z}} \setminus \{0\}$  such that  $q = \frac{a}{b}$

**Proof.** If  $q \in \mathbb{Q}$  then there exist a  $a', b' \in \mathbb{Z}, b' \neq 0$  such that  $q = \frac{a'}{b'}$  and as  $q > 0$   $a' \cdot b' \geq 0$ . We have the following possibilities (as  $a' = 0$  would mean that  $q = 0$  a contradiction).

1.  $a' < 0 \wedge b' > 0$  then using 6.37  $a' \cdot b' < 0 \cdot b' = 0$  contradicting  $a' \cdot b' \geq 0$ , so this case does not apply.
2.  $a' > 0 \wedge b' < 0$  then using 6.37  $a' \cdot b' < 0 \cdot a' = 0$  contradicting  $a' \cdot b' \geq 0$ , so this case does not apply.
3.  $a' < 0 \wedge b' < 0$  then using 6.27 we have  $a = -a' > 0, b = -b' > 0$  and  $\frac{a}{b} = \frac{(-1) \cdot a'}{(-1) \cdot b'} = \frac{a'}{b'} = q$
4.  $a' > 0 \wedge b' > 0$  then we take  $a = a', b = b'$  and have  $q = \frac{a}{b}$   $\square$

**Lemma 7.27.** If  $r \in \mathbb{Q}$  with  $0 < r$  (or  $0 \leq r$ ) then for  $s \in \mathbb{Q}, s > 0$  we have  $0 < s \cdot r$  (or  $0 \leq s \cdot r$ )

**Proof.** Using the previous lemma (see 7.26) we have that  $s = \frac{a}{b}$  where  $a, b \in \mathbb{N}_{\mathbb{Z}} \setminus \{0\} \Rightarrow a, b > 0$ . Then

1. **( $0 < r$ )** Again by the previous lemma we have  $r = \frac{c}{d}$  where  $c, d > 0$ . Now  $s \cdot r = \frac{a}{b} \cdot \frac{c}{d} = \frac{a \cdot c}{b \cdot d}$ . Using 6.36 we have  $0 < a \cdot c, 0 < b \cdot d$  and using 6.36 again we have  $0 < (a \cdot c) \cdot (b \cdot d)$  so that  $\text{sign}(s \cdot r - 0) = \text{sign}(s \cdot r) = 1 \Rightarrow 0 \leq s \cdot r$ . If  $s \cdot r = 0 \Rightarrow \frac{a \cdot c}{b \cdot d} = \frac{0}{1} \Rightarrow a \cdot c = 0$  contradicting  $0 < a \cdot c \Rightarrow s \cdot r \neq 0 \Rightarrow 0 < s \cdot r$
2. **( $0 \leq r$ )** The we have the following cases
  - a.  $r = 0 \Rightarrow r \cdot s = 0 \Rightarrow 0 \leq r \cdot s$
  - b.  $0 < r \Rightarrow 0 < s \cdot r = < 0 \leq r \cdot s$

□

**Theorem 7.28.** If  $r, q \in \mathbb{Q}$  with  $r < q$  (or  $r \leq q$ ) and  $s \in \mathbb{Q}, 0 < s$  then  $s \cdot r < s \cdot q$  (or  $s \cdot r \leq s \cdot q$ )

**Proof.**

1. **( $r < q$ )** In this case we have  $0 < (q - r)$  and by the previous lemma we have then  $0 < s \cdot (q - r) = s \cdot q - s \cdot r \Rightarrow s \cdot r < s \cdot q$
2. **( $r \leq q$ )** In this case  $0 \leq (q - r)$  and by the previous lemma we have then  $0 \leq s \cdot (q - r) = s \cdot q - s \cdot r \Rightarrow s \cdot r \leq s \cdot q$

□

**Definition 7.29.**  $\mathbb{N}_{\mathbb{Q}} = \left\{ \frac{a}{1} \mid a \in \mathbb{N}_{\mathbb{Z}} \right\}$

**Theorem 7.30.**  $\langle \mathbb{N}_{\mathbb{Q}}, + \rangle$  is a sub-semi-group of  $\langle \mathbb{Q}, + \rangle$ , furthermore if  $a, b \in \mathbb{N}_{\mathbb{Q}}$  and  $a \leq b$  then  $b - a \in \mathbb{N}_{\mathbb{Q}}$

**Proof.** Let's first prove that  $\langle \mathbb{N}_{\mathbb{Q}}, + \rangle$  is a sub-semi-group of  $\langle \mathbb{Q}, + \rangle$ .

1. If  $a, b \in \mathbb{N}_{\mathbb{Q}}$  then there exists  $a', b' \in \mathbb{N}_{\mathbb{Z}}$  such that  $a = \frac{a'}{1}, b = \frac{b'}{1}$  and then  $a + b = \frac{a'}{1} + \frac{b'}{1} = \frac{\frac{a'}{1} \cdot 1 + \frac{b'}{1} \cdot 1}{1 \cdot 1} = \frac{a' + b'}{1}$ . As  $\mathbb{N}_{\mathbb{Z}}$  is a semi-group (being a non empty semi-sub-group of  $\mathbb{Z}$ , we have  $a' + b' \in \mathbb{N}_{\mathbb{Z}} \Rightarrow a + b \in \mathbb{N}_{\mathbb{Q}}$ )
2.  $1 = \frac{1}{1} \in \mathbb{N}_{\mathbb{Q}}$

Next if  $a \leq b \Rightarrow \text{sign}(b - a) = a$  then  $b - a = b + (-a) = \frac{b'}{1} + \frac{-a'}{1} = \frac{b' - a'}{1}$  and as  $\text{sign}(b - a) = 1$  we have then  $(b' - a') \cdot 1 \geq 0 \Rightarrow b' - a' \geq 0 \Rightarrow b' - a' \in \mathbb{N}_{\mathbb{Z}} \Rightarrow b - a \in \mathbb{N}_{\mathbb{Q}}$

□

**Theorem 7.31.**  $\langle \mathbb{N}_{\mathbb{Q}}, \leq \rangle$  is well-ordered

**Proof.** Let  $A \subset \mathbb{N}_{\mathbb{Q}}$  be a non-empty subset of  $\mathbb{Q}$  then  $i_{\mathbb{Z}}^{-1}(A)$  is non-empty as  $i_{\mathbb{Z}}: \mathbb{Z} \rightarrow \mathbb{Z}_{\mathbb{Q}}$  is a bijection. Also if  $a \in i_{\mathbb{Z}}^{-1}(A)$  then  $\frac{a}{1} = i_{\mathbb{Z}}(a) \in A \subseteq \mathbb{N}_{\mathbb{Q}} \Rightarrow \frac{a}{1} = \frac{a'}{1}$  where  $a' \in \mathbb{N}_{\mathbb{Z}} \Rightarrow a \cdot 1 = a' \cdot 1 \Rightarrow a = a' \Rightarrow a \in \mathbb{N}_{\mathbb{Z}}$  so we have that  $i_{\mathbb{Z}}^{-1}(A) \subseteq \mathbb{N}_{\mathbb{Z}}$ . By well-ordering of  $\mathbb{N}_{\mathbb{Z}}$  (see 6.41) there exists a  $m' = \min(i_{\mathbb{Z}}^{-1}(A))$ . Take now  $m = i_{\mathbb{Z}}(m')$  then we have

1.  $m' \in i_{\mathbb{Z}}^{-1}(A) \Rightarrow m = i_{\mathbb{Z}}(m') \in A$
2.  $\forall a \in A$  we have by bijectivity of  $i_{\mathbb{Z}}$   $a' = i_{\mathbb{Z}}^{-1}(a) \in i_{\mathbb{Z}}^{-1}(A) \Rightarrow m' \leq a' \Rightarrow m = i_{\mathbb{Z}}(m') \leq i_{\mathbb{Z}}(a') = a \Rightarrow m = \min(A)$

□

**Theorem 7.32.** If  $a \in \mathbb{N}_{\mathbb{Q}}$  then  $0 \leq a$  and if  $a \neq 0$  then  $1 \leq a$

**Proof.** If  $a \in \mathbb{N}_{\mathbb{Q}}$  then there exists a  $a' \in \mathbb{N}_{\mathbb{Z}}$  (so  $0 \leq a'$  (see 6.28)) with  $a = \frac{a'}{1}$ , but then  $0 \leq a' = a' \cdot 1 \Rightarrow \text{sign}(a - 0) = \text{sign}(a) = 1 \Rightarrow 0 \leq a$ . If  $a \neq 0$  then  $a' \neq 0$  and as  $i_{\mathbb{N}}: \mathbb{N} \rightarrow \mathbb{N}_{\mathbb{Z}}$  is a bijection, there exists a  $a'' \in \mathbb{N}$  with  $a' = i_{\mathbb{N}}(a'')$ , as  $a' \neq 0$  we must have that  $a'' \neq 0$  (otherwise  $i_{\mathbb{N}}(0) = \sim[(s(0), 1)] = \sim[(1, 1)] = 0$ ) and thus  $0 < a''$  (by 4.47). Using 4.51 we have then that  $1 = s(0) \leq a''$  and using the order preserving of  $i_{\mathbb{N}}$  we have then  $1 = \sim[(s(1), 1)] = i_{\mathbb{N}}(1) \leq i_{\mathbb{N}}(a'') = a' \Rightarrow 1 \leq a'$ . As  $a - 1 = \frac{a'}{1} - \frac{1}{1} = \frac{a' - 1}{1}$  and  $(a' - 1) \cdot 1 = a' + (-1) \geq 0$  (as  $a' \geq 1$ ) we have  $\text{sign}(a - 1) = 1$  or  $1 \leq a$ . □

**Theorem 7.33. (Archimedean Property of  $\mathbb{Q}$ )** If  $x, y \in \mathbb{Q}$  with  $x > 0$  then there exists a  $z \in \mathbb{N}_{\mathbb{Q}}$  such that  $z \cdot x > y$

**Proof.** For  $y \in \mathbb{Q}$  we have the following possibilities

1.  $y \leq 0$  take then  $z = 1 \in \mathbb{Q}$  then as the unit 1 in  $\mathbb{Z}$  is in  $\mathbb{N}_{\mathbb{Z}}$  we have that  $1 = \frac{1}{1} \in \mathbb{N}_{\mathbb{Q}}$  and  $z \cdot x = 1 \cdot x = x > 0 \geq y \Rightarrow z \cdot x > y$ .
2.  $0 < y$ . As also  $x > 0$  we have by the previous lemma (see 7.26) that there exists  $p, q, r, s \in \mathbb{N}_{\mathbb{Z}} \setminus \{0\}$  such that  $x = \frac{p}{q}$ ,  $y = \frac{r}{s}$  as  $p > 0, s > 0 \Rightarrow p \cdot s > 0$  then by 6.40 there exists a  $z' \in \mathbb{N}_{\mathbb{Z}}$  such that  $z' \cdot p \cdot s > q \cdot r \Rightarrow z' \cdot p \cdot s > q \cdot r$ . Define  $z = \frac{z'}{1} \in \mathbb{N}_{\mathbb{Q}}$  then  $z \cdot x - y = \frac{z' \cdot p}{1 \cdot q} - \frac{r}{s} = \frac{z' \cdot p \cdot s + (-q \cdot r)}{p \cdot q}$ , as  $p > 0, q > 0 \Rightarrow p \cdot q > 0 \Rightarrow (p \cdot q) \cdot (z' \cdot p \cdot s + (-q \cdot r)) > 0 \Rightarrow \text{sign}(z \cdot x - y) = 1 \Rightarrow z \cdot x > y$ .  $\square$

**Lemma 7.34.** If  $s \in \mathbb{Q}$  then  $0 < s$  iff  $0 < s^{-1}$

**Proof.** If  $0 < s$  then by 7.26 we have  $s = \frac{a}{b}$  where  $a, b > 0$  as  $s^{-1} = \frac{b}{a}$  we have  $b \cdot a > 0 \Rightarrow \text{sign}(s^{-1} - 0) = \text{sign}(s^{-1}) = 1 \Rightarrow 0 < s^{-1}$ . If  $0 < s^{-1}$  then if  $s \leq 0$  we have by 7.27  $s \cdot s^{-1} \leq s^{-1} \cdot 0 \Rightarrow 1 \leq 0$  a contradiction, so we must have  $0 < s$ .  $\square$

**Lemma 7.35.** If  $r, q \in \mathbb{Q}$  then we have

1.  $0 < r < 1 \Rightarrow 1 < r^{-1}$
2.  $1 < r \Rightarrow r^{-1} < 1$
3.  $1 < r$  and  $0 < s \Rightarrow s < r \cdot s$
4.  $r < 1$  and  $0 < s \Rightarrow s \cdot r < s$
5.  $0 < r < s \Rightarrow s^{-1} < r^{-1}$
6.  $0 < r^{-1} < s^{-1} \Rightarrow s < r$
7. If  $r \neq 0$  then  $-(r^{-1}) = (-r)^{-1}$

**Proof.**

1. If  $0 < r < 1 \xrightarrow{7.34} 0 < r^{-1} \xrightarrow{7.28} r \cdot r^{-1} < 1 \cdot r^{-1} \Rightarrow 1 < r^{-1}$
2. If  $1 < r \xrightarrow{0 < 1} 0 < r \xrightarrow{7.34} 0 < r^{-1} \xrightarrow{7.28} 1 \cdot r^{-1} < r \cdot r^{-1} \Rightarrow r^{-1} < 1$
3. If  $0 < s$  and  $1 < r$  then by 7.28 we have  $1 \cdot s < r \cdot s = r \cdot s \Rightarrow s < r \cdot s$
4. If  $0 < s$  and  $r < 1$  then by 7.28 we have  $r \cdot s < 1 \cdot s \Rightarrow s \cdot r < s$
5. If  $0 < r < s$  then by the previous theorem we have  $0 < r^{-1}, s^{-1}$  and given  $r < s \xrightarrow{7.28} r \cdot r^{-1} < s \cdot r^{-1} \xrightarrow{(3)} 1 < s \cdot r^{-1} \xrightarrow{7.28} s^{-1} < s^{-1}(s \cdot r^{-1}) = r^{-1} \Rightarrow s^{-1} < r^{-1}$
6. If  $0 < r^{-1} < s^{-1}$  then using the previous lemma we have  $0 < r, 0 < s$  so using  $r^{-1} < s^{-1} \xrightarrow{7.28} r^{-1} \cdot r < s^{-1} \cdot r \Rightarrow 1 < s^{-1} \cdot r \xrightarrow{(3)} 1 \cdot s < s \cdot (s^{-1} \cdot r) = s < r$
7. If  $r \neq 0$  then  $-r \neq 0$  so  $r^{-1}$  and  $(-r)^{-1}$  exists. Now if  $r \neq 0$  then  $r = \frac{a}{b}$ ,  $a, b \neq 0$  and  $-(r^{-1}) = -\left(\frac{b}{a}\right) = \left(\frac{-b}{a}\right) = \left(\frac{a}{-b}\right)^{-1} = \left(\frac{-1}{-1} \cdot \frac{a}{-b}\right)^{-1} = \left(\frac{-a}{b}\right)^{-1} = (-r)^{-1}$   $\square$

**Theorem 7.36. ( $\mathbb{Q}$  is dense)** If  $x, y \in \mathbb{Q}$  with  $x < y$  then  $\exists z \in \mathbb{Q} \vdash x < z < y$

**Proof.** Take  $z = \frac{1}{2} \cdot (x+y) \in \mathbb{Q}$ , we have then  $x < y \xrightarrow{6.32} x+x < x+y$  and  $x+y < y+y$  so  $\frac{2}{1} \cdot x = (1+1) \cdot x = x+x < x+y$  and  $x+y < \frac{2}{1} \cdot y$ . Using the fact that  $0 < \frac{1}{2}$  and 6.37 we have then  $x = \frac{1}{2} \cdot \frac{2}{1} \cdot x < \frac{1}{2} \cdot (x+y)$  and  $\frac{1}{2} \cdot (x+y) < \frac{1}{2} \cdot \frac{2}{1} \cdot y = y$  and thus  $x < \frac{1}{2} \cdot (x+y) < y$   $\square$

The following theorem shows that  $\langle \mathbb{Q}, \leq \rangle$  is not well-ordered and this is the reason that we extend the rational numbers to the set of real numbers.

**Lemma 7.37.**  $\forall r \in \mathbb{Q}$  we have  $r \cdot r \neq 2 = \frac{2}{1}$

**Proof.** We prove this by contradiction so assume that  $\exists r' \in \mathbb{Q}$  such that  $r' \cdot r' = 2$ . Of course  $r' \neq 0$  [otherwise  $0 = r' \cdot r' = 2 \Rightarrow 0 = 2 \neq 0$  a contradiction]. If now  $r' < 0$  take then  $r = -r' \Rightarrow r \cdot r = (-r') \cdot (-r') = r' \cdot r' = 2$  otherwise take  $r = r'$ . So we have that  $\exists r \in \mathbb{Q}$  with  $0 < r$  and  $r \cdot r = 2$ . Then there exists a  $m, n \in \mathbb{Z}$  such that  $r = \frac{n}{m}$  and  $m \neq 0$ . Take now  $n' = n/\gcd(n, m)$  and  $m' = m/\gcd(n, m)$  then using 6.54 we have  $d|n'$  and  $d|m' \Rightarrow d = 1$  or  $d = -1$ . As  $n = n' \cdot \gcd(n, m)$  and  $m = m' \cdot \gcd(n, m)$  we have  $r = \frac{n' \cdot \gcd(n, m)}{m' \cdot \gcd(n, m)} = \frac{n'}{m'}$ . Now if  $r \cdot r = 2 \Rightarrow \frac{n' \cdot n'}{m' \cdot m'} = \frac{2}{1} \Rightarrow n' \cdot n' = 2 \cdot m' \cdot m'$  then we have that  $n' \cdot n'$  is even and using 6.57 we have that  $n'$  is even and thus there exists a  $k \in \mathbb{Z}$  such that  $n' = 2 \cdot k$ . From this it follows that  $2 \cdot m' \cdot m' = n' \cdot n' = 2 \cdot k \cdot 2 \cdot k \Rightarrow m' \cdot m' = 2 \cdot (k \cdot k)$  proving that  $m' \cdot m'$  is even and by 6.57 that  $m'$  is even and thus the existence of a  $l \in \mathbb{Z}$  such that  $m' = 2 \cdot l$ . We have thus that  $2|m'$  and  $2|n'$  but then we must have either  $2 = -1$  or  $2 = 1$  both of which are impossible so we reach a contradiction.  $\square$

**Theorem 7.38.**  $\langle \mathbb{Q}, \leq \rangle$  is not conditional complete, so there exists a non-empty set which is bounded above that does not have a least upper bound.

**Proof.** Consider the set  $A = \{r \in \mathbb{Q} \mid r > 0 \wedge r \cdot r < \frac{2}{1}\}$ , then as  $0 < \frac{4}{3}$  and  $\frac{2}{1} - \frac{4}{3} = \frac{18-16}{9} = \frac{2}{9} > 0 \Rightarrow \frac{4}{3} \cdot \frac{4}{3} < 2$  we have  $\frac{4}{3} \in A$  so that  $A \neq \emptyset$ . Also if  $x \in A$  then  $x \leq \frac{2}{1}$  [if  $\frac{2}{1} < x$   $x \in A \Rightarrow 0 < x \Rightarrow \frac{2}{1} \cdot x < x \cdot x$  and  $\frac{2}{1} \cdot \frac{2}{1} < \frac{2}{1} \cdot x \Rightarrow \frac{4}{1} < x \cdot x \Rightarrow \frac{2}{1} < x \cdot x$  contradicting  $x \in A \Rightarrow x \cdot x < \frac{2}{1}$ ] so  $\frac{2}{1}$  is a upper bound of  $A$ . We prove now by contradiction that  $u = \sup(A)$  does not exists. So assume that  $\sup(A)$  exists. As  $\frac{4}{2} \in A$  and  $\frac{4}{3} - 1 = \frac{4-3}{3} = \frac{1}{3} > 0 \Rightarrow 1 < \frac{4}{3} \leq u \Rightarrow 0 < 1 < u$  we can have now only the following cases for  $u \cdot u$  ( $\langle \mathbb{Q}, \leq \rangle$  is fully-ordered)

1.  $(u \cdot u = \frac{2}{1})$  Using the previous lemma we have shown that this is impossible.
2.  $(u \cdot u < \frac{2}{1})$  Then  $u \in A$ . Now given a  $n \in \mathbb{N}_{\mathbb{Z}} \setminus \{0\}$  we have

$$\begin{aligned} \left(u + \frac{1}{n}\right) \left(u + \frac{1}{n}\right) &= u \cdot u + u \cdot \frac{1}{n} + u \cdot \frac{1}{n} + \frac{1}{n} \cdot \frac{1}{n} \\ &= u \cdot u + 2 \cdot u \cdot \frac{1}{n} + \frac{1}{n} \cdot \frac{1}{n} \end{aligned}$$

and thus

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

Since  $\frac{2}{1}$  is a upper bound of  $A$  we have  $u \leq \frac{2}{1}$ . Now if  $n \in \mathbb{N}_{\mathbb{Z}} \setminus \{0\}$  then  $0 < n$  and using 6.46 we have then  $1 \leq n \Rightarrow 0 \leq n-1$  and as  $\frac{1}{n} - \frac{1}{n} \cdot \frac{1}{n} = \frac{1}{n} - \frac{1}{n \cdot n} = \frac{n-1}{n \cdot n} \geq 0 \Rightarrow \frac{1}{n} \cdot \frac{1}{n} \leq \frac{1}{n}$  and so we have

$$\begin{aligned} 2 \cdot u \cdot \frac{1}{n} + \frac{1}{n} \cdot \frac{1}{n} &\leq \frac{2}{1} \cdot \frac{2}{1} \cdot \frac{1}{n} + \frac{1}{n} \\ &\leq \frac{5}{n} \end{aligned}$$

Since  $0 < 5$  and by assumption  $0 < 2 - u \cdot u$  we have by the Archimedean property (see 7.33) the existence of a  $n'_0 \in \mathbb{N}_{\mathbb{Q}}$  (thus  $n'_0 = \frac{n_0}{1}$ ,  $n_0 \in \mathbb{N}_{\mathbb{Z}}$ ) such that

$$5 < n'_0 \cdot (2 - u \cdot u) = \frac{n_0}{1} \cdot (2 - u \cdot u)$$

as  $n_0 \neq 0$  (otherwise  $5 = 0$ ) we have then by multiplying both sides of the above by  $\frac{1}{n_0}$  that

$$\frac{5}{n_0} < 2 - u \cdot u$$

and this gives that  $2 \cdot u \cdot \frac{1}{n_0} + \frac{1}{n_0} \cdot \frac{1}{n_0} \leq \frac{5}{n_0} < \frac{2}{1} - u \cdot u$  and thus that  $\left(u + \frac{1}{n_0}\right) \cdot \left(u + \frac{1}{n_0}\right) < \frac{2}{1}$  which as  $0 < u < u + \frac{1}{n_0}$  means that  $u + \frac{1}{n_0} \in A$  contradicting then fact that  $u$  is a upper-bound of  $A$ . Hence this case is impossible.

3.  $(\frac{2}{1} < u \cdot u)$  If  $n \in \mathbb{N}_Z \setminus \{0\}$  then  $u - \frac{1}{n}$  can not be a upper bound of  $A$  as  $u - \frac{1}{n} < u$  (because  $0 < \frac{1}{n}$ ) and  $u$  is the least upper bound. So there exists a  $r \in A$  such that  $u - \frac{1}{n} < r$ . Now as  $0 < n$  we have by 6.46  $1 \leq n \Rightarrow 0 \leq n - 1$  and thus  $1 - \frac{1}{n} = \frac{n-1}{n} \geq 0 \Rightarrow \frac{1}{n} \leq 1 < u \Rightarrow 0 < u - \frac{1}{n}$  and thus from  $u - \frac{1}{n} < r$  we have  $(u - \frac{1}{n}) \cdot (u - \frac{1}{n}) < (u - \frac{1}{n}) \cdot r$  and  $(u - \frac{1}{n}) \cdot r < r \cdot r \Rightarrow (u - \frac{1}{n}) \cdot (u - \frac{1}{n}) < r \cdot r < \frac{2}{1} \Rightarrow (u - \frac{1}{n}) \cdot (u - \frac{1}{n}) < \frac{2}{1} \Rightarrow u \cdot u - \frac{2}{1} \cdot u \cdot \frac{1}{n} + \frac{1}{n} \cdot \frac{1}{n} < \frac{2}{1}$  and thus we have  $u \cdot u - \frac{2}{1} \cdot u \cdot \frac{1}{n} < u \cdot u - \frac{2}{1} \cdot u \cdot \frac{1}{n} + \frac{1}{n} \cdot \frac{1}{n} < \frac{2}{1}$ . Now  $0 < \frac{2}{1} \cdot u$  and  $0 < u \cdot u - \frac{2}{1}$  so using the Archimedean property there exists a  $n_0 \in \mathbb{N}_Z$  such that  $\frac{2}{1} \cdot u < n_0 \cdot (u \cdot u - \frac{2}{1})$  and as  $n_0 \neq 0$  [otherwise  $2 \cdot u < 0$ ] we have  $n \in \mathbb{N}_Z \setminus \{0\}$  but also by multiplying the last expression by  $\frac{1}{n_0}$  that  $\frac{2}{n_0} \cdot u < u \cdot u - \frac{2}{1} \Rightarrow \frac{2}{1} < u \cdot u - \frac{2}{n_0} \cdot u = u \cdot u - \frac{2}{1} \cdot u \cdot \frac{1}{n_0} < 2 \Rightarrow 2 < 2$  a contradiction. So this case is also impossible.

As (1),(2) and (3) can never occur we have finally reach a contradiction. So  $\sup(A)$  does not exists.  $\square$

## 7.4 Denumerability of the rationals

**Lemma 7.39.** *The sets  $\mathbb{N}_Q$  and  $\mathbb{Z}_Q$  are denumerable*

**Proof.** The proof that  $\mathbb{Z}_Q$  is denumerable is simple, using 7.25 we have that  $\mathbb{Z}_Q \approx \mathbb{Z}$  and as we know by 6.59 that  $\mathbb{Z} \approx \mathbb{N}$  we have that  $\mathbb{Z}_Q \approx \mathbb{N}$ , so we have that  $\mathbb{Z}_Q$  is enumerable. To prove that  $\mathbb{N}_Q$  is denumerable first note that by definition (see 7.29) we have  $\mathbb{N}_Q = \left\{ \frac{a}{1} \mid a \in \mathbb{N}_Z \right\}$ . Define now  $f: \mathbb{N}_Z \rightarrow \mathbb{N}_Q$ , we prove then that  $f$  is bijective:

1. **(injective)** If  $f(a) = f(a') \Rightarrow \frac{a}{1} = \frac{a'}{1} \Rightarrow a \cdot 1 = a' \cdot 1 \Rightarrow a = a'$
2. **(surjective)** If  $y \in \mathbb{N}_Q \Rightarrow \exists a \in \mathbb{N}_Z$  such that  $y = \frac{a}{1} = f(a)$

So we have that  $\mathbb{N}_Z \approx \mathbb{N}_Q$ , as by 6.58 we have  $\mathbb{N}_Z \approx \mathbb{N}$  we must have  $\mathbb{N}_Q \approx \mathbb{N}$  and thus that  $\mathbb{N}_Q$  is denumerable.  $\square$

**Theorem 7.40.** *The set  $\mathbb{Q}$  is denumerable*

**Proof.** Define the mapping  $f: \mathbb{Z} \times \mathbb{Z} \rightarrow \mathbb{Q}$  by

$$f(x, y) = \begin{cases} \frac{x}{y} & \text{if } (x, y) \in \mathbb{Z} \times \mathbb{Z}_0 = \mathbb{Z} \times (\mathbb{Z} \setminus \{0\}) \\ 0 & \text{if } (x, y) \in \mathbb{Z} \times \{0\} \end{cases}$$

then we prove that  $f$  is a surjection. If  $y \in \mathbb{Q}$  then there exists a  $a \in \mathbb{Z}, b \in \mathbb{Z}_0$  such that  $y = \frac{a}{b} = f(a, b)$ . Now by 6.59 we have that  $\mathbb{Z}$  is denumerable and thus by 5.59 we have that  $\mathbb{Z} \times \mathbb{Z}$  is denumerable so there exists a bijection  $g: \mathbb{N} \rightarrow \mathbb{Z} \times \mathbb{Z}$  and thus a surjection  $f \circ g: \mathbb{N} \rightarrow \mathbb{Q}$ . Using 5.67 we have then that  $\mathbb{Q}$  is countable and thus finite or denumerable. As  $\mathbb{Q}$  contains the denumerable and thus infinite subset we have by 5.36 that  $\mathbb{Q}$  is infinite and thus denumerable.  $\square$



# Chapter 8

## The real numbers

### 8.1 Definition

**Definition 8.1. (Dedekind's Cut)** A subset  $\alpha \subseteq \mathbb{Q}$  is a Dedekind's cut if the following properties are true

1.  $\alpha \neq \emptyset$
2.  $\alpha \neq \mathbb{Q}$
3.  $\forall r \in \alpha \wedge \forall s \in \mathbb{Q} \setminus \alpha \text{ we have } r < s$
4.  $\alpha$  does not have a greatest element

**Definition 8.2. ( $\mathbb{R}$ )** We define the set of real numbers to be the set of Dedekind's cuts.  $\mathbb{R} = \{\alpha \subseteq \mathbb{Q} \mid \alpha \text{ is a Dedekind's cut}\}$

**Lemma 8.3.**  $\forall \alpha \in \mathbb{R}, \forall r \in \alpha \text{ and } \forall s \in \mathbb{Q} \vdash s \leq r \text{ we have } s \in \alpha$

**Proof.** We prove this by contradiction, so assume that there exists a  $\alpha \in \mathbb{R}$  and a  $s \in \mathbb{Q}$  with  $s \leq r$  and  $s \notin \alpha \Rightarrow s \in \mathbb{Q} \setminus \alpha \xrightarrow[8.1, 3]{} r < s \leq r \Rightarrow r < r$  a contradiction.  $\square$

**Theorem 8.4. (Rational cuts)** If  $r \in \mathbb{Q}$  then  $\alpha_r = \{x \in \mathbb{Q} \mid x < r\}$  is a cut, called a rational cut. Furthermore we have

1.  $\alpha_r = \alpha_s \text{ iff } r = s$
2.  $\alpha$  is a rational cut iff  $r = \min(\mathbb{Q} \setminus \alpha)$  exists and in that case  $\alpha = \alpha_r$

**Proof.**

First we prove that  $\alpha_r$  is a cut.

1. If  $r \in \mathbb{Q}$  then there exists  $a, b \in \mathbb{Z}$  with  $b \neq 0$  such that  $r = \frac{a}{b}$ , now if  $b < 0$  then we can always take  $a' = -a, b' = -a$  such that  $\frac{a'}{b'} = \frac{-a}{-b} = \frac{(-1) \cdot a}{(-1) \cdot b} = \frac{a}{b} = q$ . So we can always assume that  $b > 0$ . Taken now  $q = \frac{a+(-1)}{b}$  then  $r - q = \frac{a}{b} + \frac{(a+(-1))}{b} = \frac{-1}{b}$  and as  $0 < b$  we have  $-b < 0 \Rightarrow (-1) \cdot b < 0 \Rightarrow q < r$  ( $q = r$  is impossible as  $b \cdot (a+(-1)) \cdot b = b \cdot a + (-b) \neq b \cdot a$ ). This proves that  $\alpha_r \neq \emptyset$ .
2. As we don't have  $r < r$  we have that  $r \notin \alpha_r \Rightarrow \alpha_r \neq \mathbb{Q}$
3. If  $w \in \alpha_r$  then  $w < r$  and if  $u \in \mathbb{Q} \setminus \alpha_r$  then  $\neg(u < r) \Rightarrow r \leq u$  and thus  $w < u$
4. If  $m = \max(\alpha_r)$  is the greatest element of  $\alpha_r$  then as  $m \in \alpha_r$  we have  $m < r$  then by the dense theorem of rational numbers (see 7.36) there exists a  $q \in \mathbb{Q}$  such that  $m < q < r \Rightarrow q \in \alpha_r$  contradiction the definition of a greatest element.

Next we prove that  $\alpha_r = \alpha_s \Leftrightarrow r = s$

1.  $\Rightarrow$  If  $r \neq s$  then we have either
  - a.  $r < s \Rightarrow r \in \alpha_s \setminus \alpha_r \Rightarrow \alpha_r \neq \alpha_s$  a contradiction
  - b.  $s < r \Rightarrow s \in \alpha_r \setminus \alpha_s \Rightarrow \alpha_s \neq \alpha_r$  a contradiction

So we conclude that  $r = s$

2.  $\Leftarrow$  If  $r = s$  then clearly  $\alpha_r = \alpha_s$

Finally we prove the equivalence of ' $\alpha$  is a rational cut iff the least element  $r$  of  $\mathbb{Q} \setminus \alpha$  exists and in that case  $\alpha = \alpha_r$ '.

1. If  $\alpha$  is a rational cut then  $\exists r \in \mathbb{Q} \vdash \alpha = \alpha_r$ . Then as  $r \in \alpha_r = \alpha$  we have that  $r \in \mathbb{Q} \setminus \alpha$ . If  $s \in \mathbb{Q} \setminus \alpha$  then  $s \notin \alpha = \alpha_r \Rightarrow \neg(s < r)$  which by the fact that  $\langle \mathbb{Q}, \leq \rangle$  is fully-ordered means that  $r \leq s$  and thus that  $r$  is the least element of  $\mathbb{Q} \setminus \alpha$  or  $r = \min(\mathbb{Q} \setminus \alpha)$ .
  2. Assume that  $\alpha$  is a cut and that  $r = \min(\mathbb{Q} \setminus \alpha)$  exists then we claim that  $\alpha = \alpha_r$ .
    - a. If  $s \in \alpha$  then as  $r \in \mathbb{Q} \setminus \alpha$  we have by the definition of a cut (see 8.1-3) that  $s < r \Rightarrow s \in \alpha_r$ . So we have  $\alpha \subseteq \alpha_r$
    - b. If  $s \in \alpha_r \subseteq \mathbb{Q}$  then we have  $s < r$  and thus  $s \notin \mathbb{Q} \setminus \alpha$  (otherwise  $r$  would not be the least element). From this it follows that  $s \in \alpha \Rightarrow \alpha_r \subseteq \alpha$ .
- (a) and (b) gives finally  $\alpha = \alpha_r$   $\square$

A consequence of the above and the fact that  $\mathbb{Q}$  is not empty proves the following corollary.

**Corollary 8.5.**  $\mathbb{R} \neq \emptyset$  The set of real numbers is not empty.

**Definition 8.6.** We define the set  $\mathbb{Q}_{\mathbb{R}}$  as follows  $\mathbb{Q}_{\mathbb{R}} = \{\alpha_r \mid r \in \mathbb{Q}\}$ , so  $\mathbb{Q}_r$  is the set of all the rational cuts.

**Theorem 8.7.** If  $\alpha \in \mathbb{R}$  then  $\forall \varepsilon \in \mathbb{Q} \setminus \{0\}$  there  $\exists r \in \alpha$  and  $\exists s \notin \alpha$  such that  $s - r < \varepsilon$

**Proof.** Take  $\alpha \in \mathbb{R}$  and  $\varepsilon \in \mathbb{Q}$  by the definition of a cut (see 8.1 1,2) there exists a  $r' \in \alpha$  and a  $s' \notin \alpha$ . Using the definition again (see 8.1 3) we have  $r' < s'$ . Using the Archimedean property of  $\mathbb{Q}$  (see 7.33) there exists a  $k \in \mathbb{N}_{\mathbb{Q}}$  such that  $s' - r' < k \cdot \varepsilon$ . Now as  $r' < s' \Rightarrow 0 < s' - r'$  we can't have that  $k = 0 \Rightarrow 0 < k$  so  $n^{-1}$  exists, also if  $k^{-1} \leq 0 \Rightarrow k \cdot k^{-1} \leq 0 \Rightarrow 1 \leq 0 < 1$  a contradiction, so we must conclude that  $0 < k^{-1}$ . From  $s' - r' < k \cdot \varepsilon$  it follows then that  $k^{-1} \cdot (s' - r') < k^{-1} \cdot k \cdot \varepsilon \Rightarrow k^{-1} \cdot (s' - r') < \varepsilon$ . Define now  $A = \{n \in \mathbb{N}_{\mathbb{Q}} \mid r' + (n \cdot k^{-1}) \cdot (s' - r') \notin \alpha\} \subseteq \mathbb{N}_{\mathbb{Q}}$ . As we have  $r' + (k \cdot k^{-1}) \cdot (s' - r') = r' + (s' - r') = s' \notin \alpha$  we have that  $k \in A \Rightarrow A$  is non empty. Using the fact that  $\mathbb{N}_{\mathbb{Q}}$  is well-ordered (see 7.31) there exists a  $k' = \min(A)$ . Now if  $k' = 0$  then we would have that  $r' + 0 \cdot (s' - r') = r' \notin \alpha$  contradicting that  $r' \in \alpha$  so we can't have  $k' = 0$ , using 7.32 we must then have  $k' \geq 1$ . Using 7.30 we have then that  $k' - 1 \in \mathbb{N}_{\mathbb{Q}}$ . By the definition of a minimum and  $k' - 1 < k'$  (see 7.23) we have also  $k' - 1 \in A \Rightarrow r' + (k' - 1) \cdot k^{-1} \cdot (s' - r') \in \alpha$ . If we now define

$$\begin{aligned} r &= r' + ((k' - 1) \cdot k^{-1}) \cdot (s' - r') \\ s &= r' + (k' \cdot k^{-1}) \cdot (s' - r') \end{aligned}$$

then we have that  $r \in \alpha$  and  $s \notin \alpha$  (as  $k' \in A$ ). Now  $s - r = r' + (k' \cdot k^{-1}) \cdot (s' - r') - (r' + ((k' - 1) \cdot k^{-1}) \cdot (s' - r')) = (s' - r') \cdot k^{-1} < \varepsilon$ .  $\square$

**Theorem 8.8. (Negative cut)** If  $\alpha \in \mathbb{R}$  then  $-\alpha = \{-s \mid s \notin \alpha \wedge s \neq \min(\mathbb{Q} \setminus \alpha)\}$  is a Dedekind's cut called the negative cut of  $\alpha$ . Note that  $s \neq \min(\mathbb{Q} \setminus \alpha)$  does not mean that  $\min(\mathbb{Q} \setminus \alpha)$  must exists, but that  $\alpha$  may not be the minimum element of  $\mathbb{Q} \setminus \alpha$  (in other words if the minimum exists it may not be  $\alpha$ ).

**Proof.**

1. Since  $\alpha \neq \mathbb{Q}$  there exists a  $\exists s \in \mathbb{Q} \vdash s \notin \alpha$  then  $s + 1 \notin \alpha$  [if  $s + 1 \in \alpha$  then as  $s < s + 1$  (see 7.23) and  $s \in \mathbb{Q} \setminus \alpha$  we have by 8.1 that  $s + 1 < s < s + 1 \Rightarrow s + 1 < s + 1$  a contradiction]. As  $s, s + 1 \in \mathbb{Q} \setminus \alpha$ ,  $s < s + 1$  we can't have  $s + 1 = \min(\mathbb{Q} \setminus \alpha) \Rightarrow s + 1 \in -\alpha \Rightarrow -\alpha \neq \emptyset$
2. Since  $\alpha \neq \emptyset$  there exists a  $r \in \alpha \subseteq \mathbb{Q}$  then  $-r \notin \alpha \Rightarrow \alpha \neq \mathbb{Q}$
3. Let now  $r \in -\alpha$  and  $s \in \mathbb{Q} \setminus (-\alpha)$  suppose now that  $s \leq r \stackrel{7.22}{\Rightarrow} -r \leq -s$ . From  $r \in -\alpha$  we have  $-r \notin \alpha$  as  $s \in \mathbb{Q} \setminus (-\alpha)$  we have  $s \notin -\alpha$  and thus  $\neg(-s \in \alpha \wedge -s \neq \min(\mathbb{Q} \setminus \alpha)) \Rightarrow -s \notin \alpha \vee -s = \min(\mathbb{Q} \setminus \alpha)$  so we have the following cases to consider
  - a.  $-s \in \alpha$ . Now  $-r \leq -s \Rightarrow \neg(-s < -r)$  together with  $-s \in \alpha$  and  $-r \in \mathbb{Q} \setminus \alpha$  contradicts the definition of  $\alpha$  as a cut (see 8.1 , 3). So this cases is impossible.
  - b.  $-s = \min(\mathbb{Q} \setminus \alpha)$  As  $-s \leq -r$  we must have  $-s = -r \Rightarrow s = r \Rightarrow s \in (-\alpha) \cap (\mathbb{Q} \setminus (-\alpha)) = \emptyset$  a contradiction. So this case is also impossible.

As (a) and (b) are impossible we must conclude that  $r < s$

4. We prove by contradiction that  $\max(-\alpha)$  does not exist. So assume that  $\max(-\alpha)$  exists, then as  $\max(-\alpha) \in -\alpha$  we must have that  $-\max(-\alpha) \notin \alpha \Rightarrow -\max(-\alpha) \in \mathbb{Q} \setminus \alpha$  and  $-\max(-\alpha) \neq \min(\mathbb{Q} \setminus \alpha)$ . We consider now the two only cases for  $\min(\mathbb{Q} \setminus \alpha)$
- $\min(\mathbb{Q} \setminus \alpha)$  does not exist. If now  $\forall s \in \mathbb{Q} \setminus \alpha$  we have  $-\max(-\alpha) \leq s$  we would have as  $-\max(-\alpha) \in \mathbb{Q} \setminus \alpha$  that  $-\max(-\alpha) = \min(\mathbb{Q} \setminus \alpha)$  contradicting the fact that we assumed that  $\min(\mathbb{Q} \setminus \alpha)$  does not exist. So we must have that  $\exists s \in \mathbb{Q} \setminus \alpha$  so that  $s < -\max(-\alpha) \Rightarrow \max(-\alpha) < -s$ , as  $s \in \mathbb{Q} \setminus \alpha$  and  $\min(\mathbb{Q} \setminus \alpha)$  does not exist so  $-s \neq \min(\mathbb{Q} \setminus \alpha)$  we must have that  $-s \in -\alpha \Rightarrow -s \leq \max(-\alpha)$  but this contradicts  $\max(-\alpha) < -s$ . So we reach a final contradiction.
  - $\min(\mathbb{Q} \setminus \alpha)$  exists. As  $-\max(-\alpha) \in \mathbb{Q} \setminus \alpha$  and  $-\max(-\alpha) \neq \min(\mathbb{Q} \setminus \alpha)$  we have  $\min(\mathbb{Q} \setminus \alpha) < \max(-\alpha)$ . Using the density theorem (see 7.36) there exists a  $s \in \mathbb{Q}$  such that  $\min(\mathbb{Q} \setminus \alpha) < s < -\max(-\alpha)$ . As  $\min(\mathbb{Q} \setminus \alpha) \in \mathbb{Q} \setminus \alpha$  we would have by the fact that  $\alpha$  is a cut (see 8.1 3) that we can not have  $s \in \alpha$  [as this would mean  $[\min(\mathbb{Q} \setminus \alpha) \in \mathbb{Q} \setminus \alpha] s < \min(\mathbb{Q} \setminus \alpha) < s$  a contradiction]. So  $s \in \mathbb{Q} \setminus \alpha$  and  $\min(\mathbb{Q} \setminus \alpha) \neq s \Rightarrow -s \in -\alpha$ , from  $s < -\max(-\alpha) \Rightarrow -s \leq \max(-\alpha) < -s \Rightarrow -s < -s$  which is a contradiction. 7.22

As all the possible cases give a contradiction we must conclude that  $\max(-\alpha)$  does not exist.

(1),(2),(3) and (4) proves that indeed  $-\alpha$  is a cut. □

For rational cuts there is an easy way to construct its negative cut as is expressed in the following theorem.

**Theorem 8.9.** *If  $r \in \mathbb{Q}$  then  $-\alpha_r = \alpha_{-r}$*

**Proof.**

$$\begin{aligned}
 x \in -\alpha_r &\Leftrightarrow -x \notin \alpha_r \wedge -x \neq \min(\mathbb{Q} \setminus \alpha_r) \\
 &\stackrel{8.4 \Rightarrow r = \min(\mathbb{Q} \setminus \alpha_r)}{\Leftrightarrow} -x \notin \alpha_r \wedge -x \neq r \\
 &\Leftrightarrow r \leq -x \wedge -x \neq r \\
 &\Leftrightarrow -x > r \\
 &\Leftrightarrow x < -r \\
 &\Leftrightarrow x \in \alpha_{-r}
 \end{aligned}$$

□

## 8.2 Arithmetic on $\mathbb{R}$

### 8.2.1 Addition in $\mathbb{R}$

**Definition 8.10.** *If  $\alpha, \beta \in \mathbb{R}$  then we define  $\alpha + \beta = \{r+s \mid r \in \alpha, s \in \beta\}$*

**Lemma 8.11.**  *$\forall \alpha \in \mathbb{R}$  and  $\varepsilon \in \mathbb{Q} \setminus \{0\}$  there exists a  $r \in \alpha$  such that  $r + \varepsilon \notin \alpha$*

**Proof.** Using 8.7 we have there exists a  $r \in \alpha$  and a  $s \notin \alpha$  such that  $s - r < \varepsilon \Rightarrow s < r + \varepsilon$ . Assume now that  $r + \varepsilon \in \alpha$  then as  $s \notin \alpha \Rightarrow s \in \mathbb{Q} \setminus \alpha$  we have by 8.1, 3 that  $r + \varepsilon < s$  contradicting  $s < r + \varepsilon$  so we must have that  $r + \varepsilon \notin \alpha$ . □

**Theorem 8.12.**  *$\forall \alpha, \beta \in \mathbb{R}$  then  $\alpha + \beta \in \mathbb{R}$*

**Proof.** Given  $\alpha, \beta \in \mathbb{R}$  we must prove that  $\alpha + \beta$  is a Dedekind's cut.

- ( $\alpha + \beta \neq \emptyset$ )** Since  $\alpha \neq \emptyset \wedge \beta \neq \emptyset \Rightarrow \exists r \in \alpha \wedge \exists s \in \beta \Rightarrow r + s \in \alpha + \beta \Rightarrow \alpha + \beta \neq \emptyset$
- ( $\alpha + \beta \neq \mathbb{Q}$ )** Given  $\varepsilon = \frac{1}{2} \in \mathbb{Q}$  there exists by 8.11 a  $r' \in \alpha$  and a  $s' \in \beta$  such that  $r' + \frac{1}{2} \notin \alpha$  and  $s' + \frac{1}{2} \in \beta$ . Now  $\forall r \in \alpha$  and  $\forall s \in \beta$  we have by 8.1, 3 that  $r < r' + \frac{1}{2}$  and  $s < s' + \frac{1}{2} \Rightarrow r + s < r' + s' + \frac{1}{2} + \frac{1}{2} = r' + s' + 1 \Rightarrow r + s < r' + s' + 1$ . Now if  $r' + s' + 1 \in \alpha + \beta$  then there exists a  $r \in \alpha$  and a  $s \in \beta$  such that  $r' + s' + 1 = r + s$  and by the above we would have the contradiction  $r' + s' + 1 = r + s < r' + s' + 1$  so we conclude that  $r' + s' + 1 \notin \alpha + \beta \Rightarrow \alpha + \beta \neq \mathbb{Q}$

3. ( $\forall w \in \alpha + \beta \wedge \forall v \in \mathbb{Q} \setminus (\alpha + \beta)$  we have  $w < v$ ) Let  $w \in \alpha + \beta$  then there exists a  $r \in \alpha$  and a  $s \in \beta$  such that  $w = r + s$ . Assume now that there exists a  $v \in \mathbb{Q} \setminus (\alpha + \beta)$  such that  $v \leq w$ . We have then  $v \leq r + s \Rightarrow v - s \leq r$ , using 8.3 we have then  $v - s \in \alpha$  and thus  $v = (v - s) + s \in \alpha + \beta$  contradicting  $v \in \mathbb{Q} \setminus (\alpha + \beta)$  and thus our assumption. So we must have  $\forall v \in \mathbb{Q} \setminus (\alpha + \beta)$  that  $w < v$ .
4. ( $\alpha + \beta$  does not have a maximum) Suppose that  $m = \max(\alpha + \beta)$  then as  $m \in \alpha + \beta$  there exists a  $r \in \alpha$  and a  $\beta \in \beta$  such that  $m = r + s$ . As  $\max(\alpha)$  and  $\max(\beta)$  does not exists there exists a  $r' \in \alpha$  such that  $r < r' \Rightarrow r + s < r' + s$  which as  $r' + s \in \alpha + \beta$  would give  $r' + s \leq m = r + s < r' + s \Rightarrow r' + s < r' + s$  a contradiction. The only conclusion left is that  $\alpha + \beta$  does not have a maximum

□

**Theorem 8.13.**  $\langle \mathbb{R}, + \rangle$  forms a abelian group, the identity element is  $\alpha_0$  which we will note as usually as 0 (so  $0 = \alpha_0$  where the 0 in the lower index is the identity in  $\langle \mathbb{Q}, + \rangle$ ). If  $\alpha \in \mathbb{R}$  then its inverse is  $-\alpha$ .

**Proof.** We rely here heavily on the fact that  $\langle \mathbb{Q}, + \rangle$  forms a abelian group (see 7.7)

1. (**associative**) If  $\alpha, \beta, \gamma \in \mathbb{R}$  then we have

$$\begin{aligned} z \in (\alpha + \beta) + \gamma &\Leftrightarrow z = r + s \wedge r \in (\alpha + \beta) \wedge s \in \gamma \\ &\Leftrightarrow z = (u + t) + s \wedge u \in \alpha \wedge t \in \beta \wedge s \in \gamma \\ &\stackrel{7.7}{\Leftrightarrow} z = u + (t + s) \wedge u \in \alpha \wedge t \in \beta \wedge s \in \gamma \\ &\Leftrightarrow z = u + v \wedge u \in \alpha \wedge v \in \beta + \gamma \\ &\Leftrightarrow z \in \alpha + (\beta + \gamma) \end{aligned}$$

So we have  $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$

2. (**commutative**) If  $\alpha, \beta \in \mathbb{R}$  then we have

$$\begin{aligned} z \in \alpha + \beta &\Leftrightarrow z = r + s \wedge r \in \alpha \wedge s \in \beta \\ &\stackrel{7.7}{\Leftrightarrow} z = s + r \wedge r \in \alpha \wedge s \in \beta \\ &\Leftrightarrow z \in \beta + \alpha \end{aligned}$$

giving  $\alpha + \beta = \beta + \alpha$

3. (**neutral element**) Let  $\alpha \in \mathbb{R}$ . If  $z \in \alpha + \alpha_0$  then  $z = r + s$  where  $r \in \alpha$  and  $s < 0 \Rightarrow r + s < r \Rightarrow r + s \leq r \Rightarrow (r + s) \in \alpha \Rightarrow z \in \alpha \Rightarrow (\alpha + \alpha_0) \subseteq \alpha$ . If  $z \in \alpha$  then as  $\alpha$  has no maximum (8.14), there exists a  $z' \in \alpha$  with  $z < z'$ , then we have  $z - z' < 0 \Rightarrow z - z' \in \alpha_0 \Rightarrow z = z' + (z - z') \in \alpha + \alpha_0 \Rightarrow \alpha \subseteq (\alpha + \alpha_0)$ . This gives then  $\alpha = \alpha + \alpha_0$  and by commutativity we have also  $\alpha = \alpha_0 + \alpha$ .

4. (**inverse**) Let  $\alpha \in \mathbb{R}$  take then  $-\alpha$  and consider the following possible cases for  $\min(\mathbb{Q} \setminus \alpha)$

- a. ( **$\min(\mathbb{Q} \setminus \alpha)$  does not exists**) If  $x \in \alpha_0$  then  $x < 0$  and thus  $0 < -x$  using 8.11 there exists a  $r \in \alpha$  such that  $r - x = r + (-x) \notin \alpha$ , as  $\min(\mathbb{Q} \setminus \alpha)$  does not exists, we have  $r - x \neq \min(\mathbb{Q} \setminus \alpha)$ . Thus by definition of  $-\alpha$  we have  $-(r - x) \in -\alpha \Rightarrow x - r \in -\alpha \Rightarrow x = r + (x - r) \in \alpha + (-\alpha) \Rightarrow \alpha_0 \subseteq \alpha + (-\alpha)$ . Also if  $x \in \alpha + (-\alpha)$  then  $x = r + s$  where  $r \in \alpha$  and  $s \in -\alpha \Rightarrow -s \notin \alpha$ , using 8.1 3 we have then that  $r < -s \Rightarrow r + s < 0 \Rightarrow x = r + s \in \alpha_0 \Rightarrow \alpha + (-\alpha) \subseteq \alpha_0$ . Combining our two results gives  $\alpha_0 = \alpha + (-\alpha)$ .

- b. ( **$\min(\mathbb{Q} \setminus \alpha)$  exists**) If  $r = \min(\mathbb{Q} \setminus \alpha)$  exist then by 8.4 we have that  $\alpha = \alpha_r$ , by 8.9 we have  $-\alpha = -\alpha_r = \alpha_{-r}$ . So we have  $z \in \alpha + (-\alpha) = \alpha_r + \alpha_{-r}$  then there exists a  $s < r$  and a  $t < -r$  such that  $z = s + t$ , this gives  $s + t < r + t$  and  $t + r < r + (-r) = 0 \Rightarrow s + t < 0 \Rightarrow z < 0 \Rightarrow z \in \alpha_0 \Rightarrow \alpha + (-\alpha) \subseteq \alpha_0$ . If  $z \in \alpha_0 \Rightarrow z < 0 \Rightarrow 0 < -z$  and as  $0 < \frac{1}{2}$  we have by 7.28 that  $0 < (-z) \cdot \frac{1}{2} = -(z \cdot \frac{1}{2}) \Rightarrow (z \cdot \frac{1}{2}) < 0$ . From this it follows that  $r + (z \cdot \frac{1}{2}) < r \Rightarrow (r + (z \cdot \frac{1}{2})) \in \alpha_r$  and  $-r + (z \cdot \frac{1}{2}) < -r \Rightarrow (-r + (z \cdot \frac{1}{2})) \in \alpha_{-r}$ . Finally  $z = (\frac{1}{2} \cdot z) + (\frac{1}{2} \cdot z) = (r + (z \cdot \frac{1}{2})) + (-r + (z \cdot \frac{1}{2})) \in \alpha_r + \alpha_{-r} = \alpha + (-\alpha) \Rightarrow \alpha_0 \subseteq \alpha + (-\alpha)$ . Combining our two results gives  $\alpha_0 = \alpha + (-\alpha)$

Using commutativity we have the also  $\alpha_0 = (-\alpha) + \alpha$  □

Next we define the set of positive and negative numbers

### 8.2.2 Multiplication in $\mathbb{R}$

**Definition 8.14.** The set of positive real numbers  $\mathbb{R}_+$  is defined by  $\mathbb{R}_+ = \{\alpha \in \mathbb{R} \mid 0 < \alpha\} \subseteq \mathbb{R}$ , the set of negative real numbers  $\mathbb{R}_-$  is defined by  $\mathbb{R}_- = \{\alpha \mid -\alpha \in \mathbb{R}_+\}$

**Theorem 8.15.**  $\mathbb{R} = \mathbb{R}_+ \cup \mathbb{R}_- \cup \{0\}$  and  $\mathbb{R}_+ \cap \{0\} = \emptyset = \mathbb{R}_- \cap \{0\}$  and  $\mathbb{R}_+ \cap \mathbb{R}_- = \emptyset$ . In other words if  $\alpha \in \mathbb{R}$  then we have either only one of  $\alpha \in \mathbb{R}_+$  or  $\alpha \in \mathbb{R}_-$  or  $\alpha = 0$ .

**Proof.** First we prove that  $\mathbb{R} = \mathbb{R}_+ \cup \mathbb{R}_- \cup \{0\}$ , as  $\mathbb{R}_+, \mathbb{R}_-, \{0\} \subseteq \mathbb{R}$  we have  $\mathbb{R}_+ \cup \mathbb{R}_- \cup \{0\} \subseteq \mathbb{R}$ . Now if  $\alpha \in \mathbb{R}$  then we have the following possible cases

1.  $(0 \in \alpha) \Leftrightarrow \alpha \in \mathbb{R}_+$
2.  $(0 \notin \alpha)$  here we have again two possible cases
  - a.  $(0 = \min(\mathbb{Q} \setminus \alpha)) \stackrel{8.4}{\Leftrightarrow} \alpha = \alpha_0 = 0 \Leftrightarrow \alpha \in \{0\}$
  - b.  $(0 \neq \min(\mathbb{Q} \setminus \alpha)) \stackrel{0 \in \alpha \text{ and definition of } -\alpha \text{ (see 8.8)}}{\Leftrightarrow} 0 = -0 \in -\alpha \Leftrightarrow -\alpha \in \mathbb{R}_+ \Leftrightarrow \alpha \in \mathbb{R}_-$

from (1),(2.a),(2.b) we conclude that  $\mathbb{R} \subseteq \mathbb{R}_+ \cup \mathbb{R}_- \cup \{0\}$  and thus we have  $\mathbb{R} = \mathbb{R}_+ \cup \mathbb{R}_- \cup \{0\}$ .

Now if  $0 \in \mathbb{R}_+$  then  $0 = \alpha_0 \in \mathbb{R}_+ \Rightarrow 0 \in \alpha_0 \Rightarrow 0 < 0$  a contradiction, so we have  $\{0\} \cap \mathbb{R}_+ = \emptyset$ .

If  $0 \in \mathbb{R}_-$  then  $-0 \in \mathbb{R}_+ \Rightarrow -0 \stackrel{\text{neutral element}}{=} \stackrel{\Rightarrow}{-0+0} \stackrel{\text{inverse element}}{=} 0 \in \mathbb{R}_+$  which we have already prove to be impossible.

Finally if  $\alpha \in \mathbb{R}_+ \cap \mathbb{R}_-$  then  $\alpha \in \mathbb{R}_+ \wedge \alpha \in \mathbb{R}_- \Rightarrow \alpha \in \mathbb{R}_+ \wedge -\alpha \in \mathbb{R}_+ \Rightarrow 0 \in \alpha \wedge 0 \in -\alpha \Rightarrow 0 \in \alpha \wedge -0 \in -\alpha$  which is a contradiction by the definition of the negative (see 8.8). □

Let's now proceed to definition of multiplication in  $\mathbb{R}$ , first we define multiplication on the set of positive reals

**Definition 8.16.**  $\forall \alpha, \beta \in \mathbb{R}_+$  then  $\alpha \odot \beta = \{r \in \mathbb{Q} \mid r \leq 0\} \cup \{s \cdot t \mid s \in \alpha \wedge t \in \beta, s > 0 \wedge t > 0\}$

We have then the following theorem

**Theorem 8.17.**  $\forall \alpha, \beta \in \mathbb{R}_+$  we have that  $\alpha \odot \beta \in \mathbb{R}_+$

**Proof.** Let's first prove that  $\alpha \odot \beta \in \mathbb{R}$

1.  $(\alpha \odot \beta \neq \emptyset)$  As  $0 \in \{r \in \mathbb{Q} \mid r \leq 0\} \Rightarrow 0 \in \alpha \odot \beta \Rightarrow \alpha \odot \beta \neq \emptyset$
2.  $(\mathbb{Q} \setminus (\alpha \odot \beta) \neq \emptyset)$  As  $\alpha, \beta \in \mathbb{R}_+$  we have  $0 \in \alpha$  and  $0 \in \beta$ , from the fact that  $\alpha, \beta$  do not have a maximum we must then have the existence of a  $s_1 \in \alpha$  and a  $t_1 \in \beta$  such that  $0 < s_1, 0 < t_1$ . As  $1 = \frac{1}{1} \in \mathbb{Q} \setminus \{0\}$  there exists by 8.11 a  $s_2 \in \alpha$  and a  $t_2 \in \beta$  such that  $s_2 + 1 \notin \alpha$  and  $t_2 + 1 \in \beta$ . Take  $s = \max(s_1, s_2), t = \max(t_1, t_2)$  then  $s \in \alpha, t \in \beta$  [if  $s \notin \alpha \stackrel{8.1.3}{\Rightarrow} s_1 < s, s_2 < s \Rightarrow s \notin \{s_1, s_2\}$  a contradiction, if  $t \notin \beta \stackrel{8.1.3}{\Rightarrow} t_1 < t, t_2 < t \Rightarrow t \notin \{t_1, t_2\}$  a contradiction] and  $0 < s, 0 < t$  and  $s_2 + 1 \leq s + 1, t_2 + 1 \leq t + 1$  [as  $s_2 \leq \max(s_1, s_2) = s, t_2 \leq \max(t_1, t_2) = t$ ]. We have then also  $s + 1 \notin \alpha, t + 1 \notin \beta$  [if  $s + 1 \in \alpha$  then by 8.1.3 and  $s_2 + 1 \notin \alpha$  we have the contradiction  $s + 1 < s_2 + 1 \leq s + 1 \Rightarrow s + 1 < s + 1$ , if  $t + 1 \in \beta$  then by 8.1.3 and  $t_2 + 1 \notin \beta$  we have the contradiction  $t + 1 < t_2 + 1 \leq t + 1 \Rightarrow t + 1 < t + 1$ ]. We claim now that  $s \cdot t + s + t + 1 \notin \alpha \odot \beta$  by contradiction, so let  $s \cdot t + s + t \in \alpha \odot \beta$ . First  $0 < s, 0 < t \stackrel{7.28}{\Rightarrow} 0 < s \cdot t$  and thus  $0 < s \cdot t + s \Rightarrow 0 < s \cdot t + s + t \Rightarrow 0 < s \cdot t + s + t + 1 \Rightarrow s \cdot t + s + t + 1 \notin \{r \in \mathbb{Q} \mid r \leq 0\}$ , from this it follows that then we must have  $s \cdot t + s + t + 1 \in \{s \cdot t \mid s \in \alpha \wedge t \in \beta \wedge 0 < s \wedge 0 < t\}$  so there exists a  $s' \in \alpha, 0 < s'$  and a  $t' \in \beta, 0 < t'$  such that  $s \cdot t + s + t + 1 = s' \cdot t'$ . Using 8.1.3 and  $s + 1 \notin \alpha, t + 1 \notin \beta$  we have that  $s' < s + 1, t' < t + 1 \stackrel{0 < s', 0 < t' < t + 1}{\Rightarrow} s' \cdot t' < (s + 1) \cdot t', (s + 1) \cdot t = t' \cdot (s + 1) < (s + 1) \cdot (t + 1) \Rightarrow s' \cdot t' < (s + 1) \cdot (t + 1) = s \cdot t + s + t + 1$  and this gives then the contradiction  $s \cdot t + s + t + 1 < s \cdot t + s + t + 1$ . So we conclude that  $s \cdot t + s + t + 1 \notin \alpha \odot \beta \Rightarrow s \cdot t + s + t + 1 \in \mathbb{Q} \setminus (\alpha \odot \beta) \Rightarrow \mathbb{Q} \setminus (\alpha \odot \beta) \neq \emptyset$

3. If  $r \in \alpha \odot \beta$  and  $s \in \mathbb{Q} \setminus (\alpha \odot \beta)$  then we have either
- $(r \in \{s \in \mathbb{Q} | s \leq 0\}) \Rightarrow r \leq 0$  now for  $s \in \mathbb{Q} \setminus (\alpha \odot \beta)$  we have  $s \notin \{t \in \mathbb{Q} | t \leq 0\} \Rightarrow 0 < s \Rightarrow r < s$
  - $(r \in \{s \in \mathbb{Q} | s \leq 0\}) \Rightarrow r \in \{s \cdot t | s \in \alpha \wedge t \in \beta \wedge 0 < s \wedge 0 < t\}$  and  $0 \leq r$  so  $r = s' \cdot t'$  where  $s' \in \alpha \wedge t' \in \beta \wedge s' > 0 \wedge t' > 0$  so we have by 7.28 that  $0 < s' \cdot t' = r$ . Suppose now that  $s \in \mathbb{Q} \setminus (\alpha \odot \beta)$  and that  $s \leq r$ , we will derive a contradiction from this. As  $s \notin \alpha \odot \beta$  we have  $s \neq r$  and  $s \notin \{s \in \mathbb{Q} | s \leq 0\} \Rightarrow 0 < s$  and thus  $0 < s < r$  using 7.10 we have that  $s^{-1}$  exists and by 7.34  $0 < s^{-1}$ . Take now  $t = s^{-1} \cdot r \Rightarrow s \cdot t = r$  then we have  $0 < t$ , using 7.28 and  $s < r$  this gives  $s^{-1} s < s^{-1} \cdot r = t \Rightarrow 1 < t$ . As  $r = s \cdot t = s' \cdot t'$  and  $0 < 1 < t$  we have  $s = t^{-1} \cdot (s' \cdot t')$ . From  $1 < t$  we have by 7.35 that  $t^{-1} < 1$  and by 7.35 and  $0 < s'$  we get  $t^{-1} \cdot s' < s'$  and thus  $t^{-1} \cdot s' \in \alpha$  [if  $t^{-1} \cdot s' \in \alpha \Rightarrow s' \leq t^{-1} \cdot s' < s' \Rightarrow s' < s'$  a contradiction]. From  $t^{-1} \cdot s' \in \alpha$  and  $t' \in \beta$  we have  $s = (t^{-1} \cdot s') \cdot t' \in \alpha \odot \beta$  contradiction  $s \in \mathbb{Q} \setminus (\alpha \odot \beta)$  so we must have  $r < s$ .
4. We prove now by contradiction that  $\max(\alpha \odot \beta)$  does not exist. So assume that  $m = \max(\alpha \odot \beta)$  exists then as  $m \in \alpha \odot \beta$  we have to consider the following two cases
- $(m \in \{s \in \mathbb{Q} | s \leq 0\}) \Rightarrow m \leq 0$ , now  $\alpha, \beta \in \mathbb{R}_+ \Rightarrow 0 \in \alpha, 0 \in \beta$  and as  $\alpha, \beta$  don't have a maximum there exists a  $s' \in \alpha, 0 < s'$  and a  $t' \in \beta, 0 < t'$  so  $0 < s' \cdot t' \in \alpha \odot \beta$  and thus  $m < s' \cdot t'$  contradicting the fact that  $m = \max(\alpha \odot \beta)$ , so we reach a contradiction.
  - $(m \in \{s \in \mathbb{Q} | s \leq 0\}) \Rightarrow 0 < m$  and  $m \in \{s \cdot t | s \in \alpha \wedge t \in \beta \wedge 0 < s \wedge 0 < t\} \Rightarrow \exists s \in \alpha, 0 < s, \exists t \in \beta, 0 < t$  with  $m = s \cdot t$ . As  $\max(\alpha)$  and  $\max(\beta)$  does not exist we have that  $\exists s' \in \alpha \vdash 0 < s < s', \exists t' \in \beta \vdash 0 < t < t' \Rightarrow m = s \cdot t < s' \cdot t' \in \alpha \odot \beta$  contradicting the maximality of  $m$ .  $\square$

Using the above definition we can define multiplication on  $\mathbb{R}$ , first we prove the following lemma.

**Lemma 8.18.**  $\mathbb{R} \times \mathbb{R}$  is the disjoint union of  $\mathbb{R}_+ \times \mathbb{R}_+$ ,  $\mathbb{R}_- \times \mathbb{R}_-$ ,  $\mathbb{R}_+ \times \mathbb{R}_-$ ,  $\mathbb{R}_- \times \mathbb{R}_+$ ,  $(\{0\} \times \mathbb{R}) \cup (\mathbb{R} \times \{0\})$

**Proof.** First as  $\mathbb{R}_+, \mathbb{R}_-, \{0\} \subseteq \mathbb{R}$  and  $(\{0\} \times \mathbb{R}) \cup (\mathbb{R} \times \{0\}) \subseteq \mathbb{R} \times \mathbb{R}$  we have  $(\mathbb{R}_+ \times \mathbb{R}_+) \cup (\mathbb{R}_- \times \mathbb{R}_-) \cup (\mathbb{R}_+ \times \mathbb{R}_-) \cup (\mathbb{R}_- \times \mathbb{R}_+) \cup ((\{0\} \times \mathbb{R}) \cup (\mathbb{R} \times \{0\})) \subseteq \mathbb{R} \times \mathbb{R}$ . Second if  $(x, y) \in \mathbb{R} \times \mathbb{R}$  then using 8.15 we have  $(x \in \mathbb{R}_+ \vee x \in \mathbb{R}_- \vee x = 0) \wedge (y \in \mathbb{R}_+ \vee y \in \mathbb{R}_- \vee y = 0) \Rightarrow (x \in \mathbb{R}_+ \wedge y \in \mathbb{R}_+) \vee (x \in \mathbb{R}_+ \wedge y \in \mathbb{R}_-) \vee (x \in \mathbb{R}_- \wedge y \in \mathbb{R}_+) \vee (x \in \mathbb{R}_- \wedge y \in \mathbb{R}_-) \vee (x \in 0 \wedge y \in \mathbb{R}_+) \vee (x \in 0 \wedge y \in \mathbb{R}_-) \vee (x \in 0 \vee y = 0) \Rightarrow (x, y) \in (\mathbb{R}_+ \times \mathbb{R}_+) \cup (\mathbb{R}_- \times \mathbb{R}_-) \cup (\mathbb{R}_+ \times \mathbb{R}_-) \cup (\mathbb{R}_- \times \mathbb{R}_+) \cup ((\{0\} \times \mathbb{R}) \cup (\mathbb{R} \times \{0\})) \Rightarrow \mathbb{R} \times \mathbb{R} \subseteq (\mathbb{R}_+ \times \mathbb{R}_+) \cup (\mathbb{R}_- \times \mathbb{R}_-) \cup (\mathbb{R}_+ \times \mathbb{R}_-) \cup (\mathbb{R}_- \times \mathbb{R}_+) \cup ((\{0\} \times \mathbb{R}) \cup (\mathbb{R} \times \{0\}))$ . This proves that  $(\mathbb{R}_+ \times \mathbb{R}_+) \cup (\mathbb{R}_- \times \mathbb{R}_-) \cup (\mathbb{R}_+ \times \mathbb{R}_-) \cup (\mathbb{R}_- \times \mathbb{R}_+) \cup ((\{0\} \times \mathbb{R}) \cup (\mathbb{R} \times \{0\})) = \mathbb{R} \times \mathbb{R}$ . Next to prove that the union is a disjoint union. From the fact that  $\mathbb{R}$  is the disjoint union of  $\mathbb{R}_+$ ,  $\mathbb{R}_-$  and  $\{0\}$  we have the following

- If  $(x, y) \in (\{0\} \times \mathbb{R}) \cup (\mathbb{R} \times \{0\})$  then  $x = 0 \vee y = 0$  so  $(x, y) \notin \mathbb{R}_+ \times \mathbb{R}_+$ ,  $\mathbb{R}_+ \times \mathbb{R}_-$ ,  $\mathbb{R}_- \times \mathbb{R}_+$ ,  $\mathbb{R}_- \times \mathbb{R}_-$  and thus  $((\{0\} \times \mathbb{R}) \cup (\mathbb{R} \times \{0\})) \cap (\mathbb{R}_+ \times \mathbb{R}_+) = \emptyset$ ,  $((\{0\} \times \mathbb{R}) \cup (\mathbb{R} \times \{0\})) \cap (\mathbb{R}_+ \times \mathbb{R}_-) = \emptyset$ ,  $((\{0\} \times \mathbb{R}) \cup (\mathbb{R} \times \{0\})) \cap (\mathbb{R}_- \times \mathbb{R}_+) = \emptyset$ ,  $((\{0\} \times \mathbb{R}) \cup (\mathbb{R} \times \{0\})) \cap (\mathbb{R}_- \times \mathbb{R}_-) = \emptyset$
- If  $(x, y) \in \mathbb{R}_+ \times \mathbb{R}_+$  then  $x \neq 0, x \notin \mathbb{R}_-, y \neq 0, y \notin \mathbb{R}_-$  so  $(x, y) \notin (\{0\} \times \mathbb{R}) \cup (\mathbb{R} \times \{0\})$ ,  $\mathbb{R}_+ \times \mathbb{R}_-$ ,  $\mathbb{R}_- \times \mathbb{R}_+$ ,  $\mathbb{R}_- \times \mathbb{R}_-$  and thus  $(\mathbb{R}_+ \times \mathbb{R}_+) \cap (\mathbb{R}_+ \times \mathbb{R}_-) = \emptyset$ ,  $(\mathbb{R}_+ \times \mathbb{R}_+) \cap (\mathbb{R}_- \times \mathbb{R}_+) = \emptyset$ ,  $(\mathbb{R}_+ \times \mathbb{R}_+) \cap (\mathbb{R}_- \times \mathbb{R}_-) = \emptyset$ ,  $(\mathbb{R}_+ \times \mathbb{R}_+) \cap ((\{0\} \times \mathbb{R}) \cup (\mathbb{R} \times \{0\})) = \emptyset$
- If  $(x, y) \in \mathbb{R}_+ \times \mathbb{R}_-$  then  $x \neq 0, x \notin \mathbb{R}_-, y \neq 0, y \notin \mathbb{R}_+$  so  $(x, y) \notin (\{0\} \times \mathbb{R}) \cup (\mathbb{R} \times \{0\})$ ,  $\mathbb{R}_+ \times \mathbb{R}_+$ ,  $\mathbb{R}_- \times \mathbb{R}_+$ ,  $\mathbb{R}_- \times \mathbb{R}_-$  and thus  $(\mathbb{R}_+ \times \mathbb{R}_-) \cap (\mathbb{R}_+ \times \mathbb{R}_+) = \emptyset$ ,  $(\mathbb{R}_+ \times \mathbb{R}_-) \cap (\mathbb{R}_- \times \mathbb{R}_+) = \emptyset$ ,  $(\mathbb{R}_+ \times \mathbb{R}_-) \cap (\mathbb{R}_- \times \mathbb{R}_-) = \emptyset$ ,  $(\mathbb{R}_+ \times \mathbb{R}_+) \cap ((\{0\} \times \mathbb{R}) \cup (\mathbb{R} \times \{0\})) = \emptyset$
- If  $(x, y) \in \mathbb{R}_- \times \mathbb{R}_+$  then  $x \neq 0, x \notin \mathbb{R}_+, y \neq 0, y \notin \mathbb{R}_-$  so  $(x, y) \notin (\{0\} \times \mathbb{R}) \cup (\mathbb{R} \times \{0\})$ ,  $\mathbb{R}_+ \times \mathbb{R}_+$ ,  $\mathbb{R}_- \times \mathbb{R}_-$ ,  $\mathbb{R}_+ \times \mathbb{R}_-$  and thus  $(\mathbb{R}_- \times \mathbb{R}_+) \cap (\mathbb{R}_+ \times \mathbb{R}_-) = \emptyset$ ,  $(\mathbb{R}_- \times \mathbb{R}_+) \cap (\mathbb{R}_+ \times \mathbb{R}_+) = \emptyset$ ,  $(\mathbb{R}_- \times \mathbb{R}_+) \cap (\mathbb{R}_- \times \mathbb{R}_-) = \emptyset$ ,  $(\mathbb{R}_+ \times \mathbb{R}_-) \cap ((\{0\} \times \mathbb{R}) \cup (\mathbb{R} \times \{0\})) = \emptyset$

5. If  $(x, y) \in \mathbb{R}_- \times \mathbb{R}_-$  then  $x \neq 0, x \notin \mathbb{R}_+, y \neq 0, y \notin \mathbb{R}_+$  so  $(x, y) \notin (\{0\} \times \mathbb{R}) \cup (\mathbb{R} \times \{0\}), \mathbb{R}_+ \times \mathbb{R}_+, \mathbb{R}_+ \times \mathbb{R}_-, \mathbb{R}_- \times \mathbb{R}_+$  and thus  $(\mathbb{R}_- \times \mathbb{R}_-) \cap (\mathbb{R}_+ \times \mathbb{R}_+) = \emptyset, (\mathbb{R}_- \times \mathbb{R}_-) \cap (\mathbb{R}_+ \times \mathbb{R}_-) = \emptyset, (\mathbb{R}_- \times \mathbb{R}_-) \cap (\mathbb{R}_- \times \mathbb{R}_+) = \emptyset, (\mathbb{R}_- \times \mathbb{R}_-) \cap ((\{0\} \times \mathbb{R}) \cup (\mathbb{R} \times \{0\}))) = \emptyset$   $\square$

Using the previous lemma and 2.34 we can define  $\cdot: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$

**Definition 8.19.** We define  $\cdot: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$   $(\alpha, \beta) \rightarrow \alpha \cdot \beta$  as follows

$$\alpha \cdot \beta = \begin{cases} \alpha \odot \beta & \text{if } (\alpha, \beta) \in \mathbb{R}_+ \times \mathbb{R}_+ \\ -((- \alpha) \odot \beta) & \text{if } (\alpha, \beta) \in \mathbb{R}_- \times \mathbb{R}_+ \\ -(\alpha \times (-\beta)) & \text{if } (\alpha, \beta) \in \mathbb{R}_+ \times \mathbb{R}_- \\ (-\alpha) \odot (-\beta) & \text{if } (\alpha, \beta) \in \mathbb{R}_- \times \mathbb{R}_+ \\ 0 & \text{if } (\alpha, \beta) \in (\{0\} \times \mathbb{R}) \cup (\mathbb{R} \times \{0\}) \end{cases}$$

If we want to prove something about multiplication we have to consider each time 5 different cases, luckily the following lemma will allow us to bring down the different cases.

**Lemma 8.20.**  $\forall (\alpha, \beta) \in \mathbb{R} \times \mathbb{R}$  we have  $-(\alpha \cdot \beta) = (-\alpha) \cdot \beta = \alpha \cdot (-\beta)$

**Proof.** We have to consider the following 5 cases

1.  $((\alpha, \beta) \in \mathbb{R}_+ \times \mathbb{R}_+)$  We have then

a.  $-\alpha \in \mathbb{R}_-$  and (see 3.6)  $\alpha = -(-\alpha)$  then we have

$$\begin{aligned} -(\alpha \cdot \beta) &= -(\alpha \odot \beta) \\ &= -(-(-\alpha) \odot \beta) \\ &= -(-(-\alpha \cdot \beta)) \\ &= -(-\alpha \cdot \beta) \end{aligned}$$

b.  $-\beta \in \mathbb{R}_-$  and (see 3.6)  $\beta = -(-\beta)$  then we have

$$\begin{aligned} -(\alpha \cdot \beta) &= -(\alpha \odot \beta) \\ &= -(\alpha \odot (-(-\beta))) \\ &= -(-(\alpha \cdot (-\beta))) \\ &= (\alpha \cdot (-\beta)) \end{aligned}$$

2.  $((\alpha, \beta) \in \mathbb{R}_+ \times \mathbb{R}_-)$  We have then

a.  $-\alpha \in \mathbb{R}_-$  and  $\alpha = -(-\alpha)$  then we have

$$\begin{aligned} -(\alpha \cdot \beta) &= -(-(\alpha \odot (-\beta))) \\ &= -(-(-\alpha) \odot (-\beta)) \\ &= ((-\alpha) \cdot \beta) \end{aligned}$$

b.  $-\beta \in \mathbb{R}_+$  and  $\beta = -(-\beta)$  then we have

$$\begin{aligned} -(\alpha \cdot \beta) &= -(-(\alpha \odot (-\beta))) \\ &= (\alpha \odot (-\beta)) \\ &= (\alpha \cdot (-\beta)) \end{aligned}$$

3.  $((\alpha, \beta) \in \mathbb{R}_- \times \mathbb{R}_+)$  We have then

a.  $-\alpha \in \mathbb{R}_+$  and  $\alpha = -(-\alpha)$  then we have

$$\begin{aligned} -(\alpha \cdot \beta) &= -(-(-\alpha \odot \beta)) \\ &= (-\alpha \cdot \beta) \end{aligned}$$

b.  $-\beta \in \mathbb{R}_-$  and  $\beta = -(-\beta)$  then we have

$$\begin{aligned} -(\alpha \cdot \beta) &= -(-(-\alpha \odot \beta)) \\ &= (-\alpha \odot \beta) \\ &= (-\alpha \odot (-(-\beta))) \\ &= (\alpha \cdot (-\beta)) \end{aligned}$$

4.  $((\alpha, \beta) \in \mathbb{R}_- \times \mathbb{R}_-)$  We have then

a.  $-\alpha \in \mathbb{R}_+$  and  $\alpha = -(-\alpha)$  then we have

$$\begin{aligned} -(\alpha \cdot \beta) &= -(-\alpha \odot (-\beta)) \\ &= -(-(-\alpha \cdot \beta)) \\ &= (-\alpha \cdot \beta) \end{aligned}$$

b.  $-\beta \in \mathbb{R}_+$  and  $\beta = -(-\beta)$  then we have

$$\begin{aligned} -(\alpha \cdot \beta) &= -(-\alpha \odot -\beta) \\ &= -(-(\alpha \cdot (-\beta))) \\ &= (\alpha \cdot (-\beta)) \end{aligned}$$

5.  $((\alpha, \beta) \in (\{0\} \times \mathbb{R}) \cup (\mathbb{R} \times \{0\}))$  then we have the following (non exclusive cases)

a.  $\alpha = 0$  We have then

$$\begin{aligned} -(\alpha \cdot \beta) &= -(0 \cdot \beta) \\ &= 0 \\ &= (0 \cdot \beta) \\ &= (-0 \cdot \beta) \\ &= (-\alpha \cdot \beta) \\ -(\alpha \cdot \beta) &= -(0 \cdot \beta) \\ &= 0 \\ &= (0 \cdot (-\beta)) \\ &= (\alpha \cdot (-\beta)) \end{aligned}$$

b.  $\beta = 0$  We have then

$$\begin{aligned} -(\alpha \cdot \beta) &= -(\alpha \cdot 0) \\ &= 0 \\ &= (-\alpha, 0) \\ &= (-\alpha, \beta) \\ -(\alpha \cdot \beta) &= -(\alpha \cdot 0) \\ &= 0 \\ &= (\alpha \cdot (-0)) \\ &= (\alpha \cdot (-\beta)) \end{aligned}$$

□

**Theorem 8.21.**  $\forall \alpha \in \mathbb{R}_+$  we have that  $\text{rep}(\alpha) = \{r \in \mathbb{Q} | r \leq 0\} \cup \{s^{-1} | s \in \mathbb{Q} \setminus \alpha \wedge s > 0 \wedge s \neq \min(\mathbb{Q} \setminus \alpha)\}$  is a cut. This is called the reciprocal cut of  $\alpha$ .

**Proof.** We have

1. **( $\text{rep}(\alpha) \neq \emptyset$ )** As  $0 \in \{r \in \mathbb{Q} | r \leq 0\}$  we have  $0 \in \text{rep}(\alpha) \Rightarrow \text{rep}(\alpha) \neq \emptyset$
2. **( $\text{rep}(\alpha) \neq \mathbb{Q}$ )** As  $0 \in \alpha$  then from the fact that there does not exist a maximum for  $\alpha$  there exists a  $s \in \alpha$  with  $s > 0$ . Using 7.34 we have  $0 < s^{-1} \Rightarrow s^{-1} \notin \{r \in \mathbb{Q} | r \leq 0\}$  and as  $s \in \alpha$  we have  $s^{-1} \notin \{s^{-1} | s \in \mathbb{Q} \setminus \alpha \wedge s > 0 \wedge s \neq \min(\mathbb{Q} \setminus \alpha)\}$  so  $s^{-1} \in \mathbb{Q} \setminus \text{rep}(\alpha)$

3. ( $\forall r \in \text{rep}(\alpha) \wedge \forall s \in \mathbb{Q} \setminus \text{rep}(\alpha)$  we have  $r < s$ ) Let  $r \in \text{rep}(\alpha)$  and  $s \in \mathbb{Q} \setminus \text{rep}(\alpha) \Rightarrow s \notin \text{rep}(\alpha)$ . We have for  $r$  the following exclusive cases

- a. ( $r \leq 0$ ) then as  $s \notin \text{rep}(\alpha)$  we must have  $s \in \{r \in \mathbb{Q} | r \leq 0\} \Rightarrow 0 < s \Rightarrow r < s$
- b. ( $0 < r$ ) then  $0 < r^{-1}, r^{-1} \in \mathbb{Q} \setminus \alpha \Rightarrow r^{-1} \notin \alpha$  and  $r^{-1} \neq \min(\mathbb{Q} \setminus \alpha)$ . As  $s \in \mathbb{Q} \setminus \text{rep}(\alpha)$  we have  $0 < s \Rightarrow 0 < s^{-1}$ . Consider now the following possible cases for  $s^{-1}$  then

$$\text{i. } (s^{-1} = \min(\mathbb{Q} \setminus \alpha)) \text{ then as } r^{-1} \notin \alpha \Rightarrow r^{-1} \notin \mathbb{Q} \setminus \alpha \xrightarrow{s^{-1} = \min(\mathbb{Q} \setminus \alpha)} s^{-1} \leq r^{-1} \xrightarrow{r^{-1} \neq \min(\mathbb{Q} \setminus \alpha)} s^{-1} < r^{-1} \xrightarrow{7.35} r < s$$

ii. ( $s^{-1} \neq \min(\mathbb{Q} \setminus \alpha)$ ) As  $s \notin \text{rep}(\alpha)$  we have  $0 < s$  [ $s \notin \{r \in \mathbb{Q} | r \leq 0\}$ ] and thus by 7.35 we have  $0 < s^{-1}$ . We must then have  $s^{-1} \in \alpha$  [otherwise  $s^{-1} \in \{s^{-1} | s \in \mathbb{Q} \setminus \alpha \wedge s > 0 \wedge s \neq \min(\mathbb{Q} \setminus \alpha)\} \subseteq \text{rep}(\alpha)$ ] and using 8.1, 3 we have then  $s^{-1} < r^{-1} \xrightarrow{7.35} r < s$

4. ( $\text{rep}(\alpha)$  does not have a greatest element) We prove this by contradiction. So assume that  $m = \max(\text{rep}(\alpha)) \in \text{rep}(\alpha)$  exists, then we derive a contradiction from this. We can now have for  $m \in \text{rep}(\alpha)$  the following possibilities

- a. ( $m \leq 0$ ) Using 8.1, 2 there exists a  $s \in \mathbb{Q} \setminus \alpha \Rightarrow s \notin \alpha$ , as  $\alpha \in \mathbb{R}_+ \Rightarrow 0 \in \alpha \xrightarrow{8.1, 3} 0 < s \Rightarrow 0 < s < s+1$ . From  $0 < s < s+1$  we have  $s+1 \notin \alpha$  [for otherwise as  $s \notin \alpha$  we have by 8.1, 3 that  $s+1 < s$  contradicting  $s < s+1$ ], as  $s < s+1$  we have  $s+1 \neq \min(\mathbb{Q} \setminus \alpha)$  and from  $0 < s+1$  we have  $0 < (s+1)^{-1}$  and thus we have  $(s+1)^{-1} \in \text{rep}(\alpha)$ . Then  $0 \leq m < (s+1)^{-1} \leq m \Rightarrow m < m$  and we reach a contradiction.
- b. ( $0 < m$ ) then as  $m \in \text{rep}(\alpha)$  we must have  $0 < m^{-1}, m^{-1} \notin \alpha$  and  $m^{-1} \neq \min(\mathbb{Q} \setminus \alpha)$ . From the last inequality we have the existence of a  $t \in \mathbb{Q} \setminus \alpha$  such that  $t < m^{-1}$ . Using the fact that  $\mathbb{Q}$  is dense (see 7.36) there exists a  $s \in \mathbb{Q}$  such that  $t < s < m^{-1}$  then we must have  $s \notin \alpha$  [for otherwise by 8.1, 3 I would have  $s < t$  contradicting  $t < s$ ]. As  $\alpha \in \mathbb{R}_+ \Rightarrow 0 \in \alpha \Rightarrow 0 < t < s < m^{-1}$  then as  $t \notin \alpha$  we must have  $s \notin \alpha$  [otherwise we must have  $s < t$  contradicting  $t < s$ ]. Since  $t < s$  and  $t, s \in \mathbb{Q} \setminus \alpha$  we have  $s \neq \min(\mathbb{Q} \setminus \alpha)$ . We must then from  $0 < s \Rightarrow 0 < s^{-1}, s \neq \min(\mathbb{Q} \setminus \alpha), s \in \mathbb{Q} \setminus \alpha$  conclude that  $s^{-1} \in \text{rep}(\alpha)$ . From  $0 < s < m^{-1}$  we have  $m = (m^{-1})^{-1} < s^{-1} \Rightarrow m < s^{-1}$  contradicting the maximality of  $m$ .  $\square$

**Definition 8.22.**  $\mathbb{R}_0 = \mathbb{R} \setminus \{0\}$ . Note that by 8.15 we have  $\mathbb{R}_0 = \mathbb{R}_+ \cup \mathbb{R}_-$

We are now ready to define the multiplicative inverse of non zero real numbers.

**Definition 8.23.** If  $r \in \mathbb{R}_0$  then we define  $r^{-1}$  as follows  $r^{-1} = \begin{cases} \text{rep}(r) & \text{if } r \in \mathbb{R}_+ \\ -\text{rep}(-r) & \text{if } r \in \mathbb{R}_- \end{cases}$

Actual the reciprocal cut of a cut takes a rather simple form in the case of positive rational cuts.

**Lemma 8.24.** If  $r \in \mathbb{Q}$  with  $\alpha_r \in \mathbb{R}_0$  then  $r^{-1} = \alpha_{r^{-1}}$  (note that there are two different inverses here one in  $\mathbb{Q}$  and one in  $\mathbb{R}_0$ )

**Proof.** First we take the case of  $a_r \in \mathbb{R}_+$  then we have as  $0 \in \alpha_r \Rightarrow 0 < r \Rightarrow 0 < r^{-1}$ . For  $x \in (\alpha_r)^{-1}$  the following possible cases

1. ( $x \leq 0$ ) then  $x \leq 0 < r^{-1} \Rightarrow x \in \alpha_{r^{-1}}$
2. ( $0 < x$ ) then we must have  $0 < x^{-1}, x^{-1} \in \mathbb{Q} \setminus \alpha_r, x^{-1} \neq \min(\mathbb{Q} \setminus \alpha_r) \xrightarrow{8.4} r \Rightarrow r < x^{-1} \xrightarrow{7.35} (x^{-1})^{-1} < r^{-1} \Rightarrow x < r^{-1} \Rightarrow x \in \alpha_{r^{-1}}$

so we have  $(\alpha_r)^{-1} \subseteq \alpha_{r^{-1}}$

If  $x \in \alpha_{r^{-1}}$  then we have the following possible cases for  $x$

1. ( $x \leq 0$ ) Then  $x \in (\alpha_r)^{-1}$  trivially.
2. ( $0 < x$ ) Then as  $x \in \alpha_{r^{-1}} \Rightarrow x < r^{-1} \xrightarrow{7.35} r < x^{-1} \Rightarrow r = \min(\mathbb{Q} \setminus \alpha_r) \neq x^{-1}, 0 < x^{-1}$  and  $x^{-1} \in \alpha_r \Rightarrow x^{-1} \in \mathbb{Q} \setminus \alpha_r \Rightarrow x \in (\alpha_r)^{-1}$

so we have  $\alpha_{r^{-1}} \in (\alpha_r)^{-1}$ . And we conclude thus that  $\alpha_{r^{-1}} = (\alpha_r)^{-1}$ .

$$\text{Second if } \alpha_r \in \mathbb{R}_- \text{ then } (\alpha_r)^{-1} = -(\text{rep}(-\alpha)) \underset{-\alpha \in \mathbb{R}_+ \text{ and previous}}{=} -(\alpha_{(-r)^{-1}}) \underset{7.35}{=} -(\alpha_{-(r^{-1})}) \underset{8.9}{=} \alpha_{(-(r^{-1}))} = \alpha_{r^{-1}} \quad \square$$

**Theorem 8.25. ( $\mathbb{R}$  forms a field)**  $\langle \mathbb{R}, +, \cdot \rangle$  forms a field. The neutral element (identity) from multiplication is  $\alpha_1$ , to simplify our notations in the future we note  $\alpha_1$  as 1 (again 1 means different things in different fields). The multiplicative inverse of  $\alpha$  is  $\alpha^{-1}$

**Proof.** First we have already proved that  $\langle \mathbb{R}, + \rangle$  is a abelian group (see 8.13) we prove then the following for multiplication. In the proof we make heavily use of the lemma 8.20.

1. **(Commutativity)** we have to consider the following five cases

a.  $((\alpha, \beta) \in \mathbb{R}_+ \times \mathbb{R}_+)$  Here we have

$$\begin{aligned} \alpha \cdot \beta &= \alpha \odot \beta \\ &= \{r \in \mathbb{Q} | r \leq 0\} \bigcup \{s \cdot t | s \in \alpha, t \in \beta, s > 0, t > 0\} \\ &\stackrel{7.10}{=} \{r \in \mathbb{Q} | r \leq 0\} \bigcup \{t \cdot s | s \in \alpha, t \in \beta, s > 0, t > 0\} \\ &= \beta \odot \alpha \\ &= \beta \cdot \alpha \end{aligned}$$

b.  $((\alpha, \beta) \in \mathbb{R}_+ \times \mathbb{R}_-)$  Here we have

$$\begin{aligned} \alpha \cdot \beta &= -(\alpha \cdot (-\beta)) \\ &\stackrel{1.a}{=} -(\beta \cdot (-\alpha)) \\ &= \beta \cdot \alpha \end{aligned}$$

c.  $((\alpha, \beta) \in \mathbb{R}_- \times \mathbb{R}_+)$  Here we have

$$\begin{aligned} \alpha \cdot \beta &= -(-\alpha \cdot \beta) \\ &\stackrel{1.a}{=} -(\beta \cdot (-\alpha)) \\ &= \beta \cdot \alpha \end{aligned}$$

d.  $((\alpha, \beta) \in \mathbb{R}_- \times \mathbb{R}_-)$  Here we have

$$\begin{aligned} \alpha \cdot \beta &= -(-(-\alpha \cdot (-\beta))) \\ &= ((-\alpha) \cdot (-\beta)) \\ &\stackrel{1.a}{=} ((-\beta) \cdot (-\alpha)) \\ &= \beta \cdot \alpha \end{aligned}$$

e.  $((\alpha, \beta) \in ((\{0\} \times \mathbb{R}) \bigcup (\mathbb{R} \times \{0\})))$  Here we have

$$\begin{aligned} \alpha \cdot \beta &= 0 \\ &= \beta \cdot \alpha \end{aligned}$$

2. **(Neutral element)**

First note that  $0 < 1$  so  $0 \in \alpha_1$  we have the to consider the following cases

a.  $(\alpha \in \mathbb{R}_+)$  If  $x \in \alpha \odot \alpha_1$  then we have either

- i.  $(x \leq 0)$  As  $\alpha \in \mathbb{R}_+ \Rightarrow 0 \in \alpha$  then using 8.3 we have  $x \in \alpha$
- ii.  $(0 < x)$  So there exists a  $s \in \alpha, t \in \alpha_1 \Rightarrow t < 1$  with  $0 < s, t$  and  $x = s \cdot t$ , from  $t < 1$  we have  $x = s \cdot t < s \Rightarrow x \leq s$ , as  $s \in \alpha$  we have by 8.3 that  $x \in \alpha$ .

So we conclude that  $\alpha \odot \alpha_1 \subseteq \alpha$ . If  $x \in \alpha$  then we have the following possible cases

- i.  $(x \leq 0)$  then  $x \in \alpha \odot \alpha_1$
- ii.  $(0 < x)$  then as  $\max(\alpha)$  does not exists there exists a  $s \in \alpha$  such that  $0 < x < s$  then  $0 < x \cdot s^{-1} < 1 \Rightarrow 0 < x \cdot s^{-1} \in \alpha_1$  and from  $x = s \cdot (x \cdot s^{-1})$  we have  $x \in \alpha \odot \alpha_1$

this gives  $\alpha \subseteq \alpha \odot \alpha_1$ . Or to summarize we have  $\alpha = \alpha \odot \alpha_1$ . Now we have

$$\begin{aligned}\alpha &= \alpha \odot \alpha_1 \\ &= \alpha \cdot \alpha_1 \\ &\stackrel{\text{commutativity}}{=} \alpha_1 \cdot \alpha\end{aligned}$$

b. ( $\alpha \in \mathbb{R}_-$ ) Here we have

$$\begin{aligned}\alpha \cdot \alpha_1 &= -((-\alpha) \cdot \alpha_1) \\ &\stackrel{2.a}{=} -(-\alpha) \\ &= \alpha \\ &\stackrel{\text{commutativity}}{=} \alpha_1 \cdot \alpha\end{aligned}$$

c. ( $\alpha = 0$ ) Here we have

$$\begin{aligned}\alpha \cdot \alpha_1 &= 0 \\ &= \alpha_1 \cdot \alpha\end{aligned}$$

3. (Inverse for non zero element) As  $\alpha \in \mathbb{R}_0$  we have the following cases to consider

a. ( $\alpha \in \mathbb{R}_+$ ) Take then  $x \in \alpha \odot \alpha^{-1} = \alpha \odot \text{rep}(\alpha)$  then we have either

- i. ( $x \leq 0$ ) then as  $0 < 1$  we have  $x < 1 \Rightarrow x \in \alpha_1$
- ii. ( $0 < x$ ) then  $\exists s \in \alpha \vdash 0 < s$  and  $\exists t \in \alpha^{-1} \vdash t < 1$  such that  $s \cdot t = x$ . Now for  $t$  we have the following cases

- A. ( $t \leq 0$ ) then from  $0 < s$  we have  $x = s \cdot t \leq 0 < 1 \Rightarrow x < 1 \Rightarrow x \in \alpha_1$
- B. ( $0 < t$ ) then  $t^{-1} \in \mathbb{Q} \setminus \alpha$ ,  $0 < t^{-1}$  and  $t^{-1} \neq \min(\mathbb{Q} \setminus \alpha)$ . As  $s \in \alpha$  we have from  $t^{-1} \in \mathbb{Q} \setminus \alpha$  and 8.1, 3 that  $s < t^{-1} \Rightarrow x = s \cdot t < 1 \Rightarrow x \in \alpha_1$

Take now  $x \in \alpha_1 \Rightarrow x < 1$  then we have either

- i. ( $x \leq 0$ )  $\Rightarrow x \in \alpha \odot \text{rep}(\alpha)$
- ii. ( $0 < x$ ) then we have  $x^{-1} \in \mathbb{Q} \setminus \alpha$ ,  $0 < x^{-1}$  and  $x^{-1} \neq \min(\mathbb{Q} \setminus \alpha)$ . From  $0 < x^{-1}$  we have by 7.34 that  $0 < x \Rightarrow 0 < x < 1 \Rightarrow 0 < 1 - x$ . Since  $\alpha \in \mathbb{R}_+ \Rightarrow 0 \in \alpha$  and as  $\max(\alpha)$  does not exists there exists a  $s_1 \in \alpha$  such that  $0 < s_1$ . From  $0 < x < 1$  we have  $1^{-1} = 1 < x^{-1}$  and thus  $0 < 1 - x < x^{-1}(x - 1) \Rightarrow 0 < s_1 \cdot (x - 1) \cdot x^{-1}$ . If we note  $\varepsilon = s_1 \cdot (x - 1) \cdot x^{-1}$  and  $0 < \varepsilon$ . Using 8.11 there exists a  $s_2 \in \alpha$  such that  $\varepsilon + s_2 \notin \alpha$ . Then we can find a  $s_3 \in \alpha$  such that  $\varepsilon + s_2 \leq \varepsilon + s_3 \in \mathbb{Q} \setminus \alpha$ ,  $\varepsilon + s_3 \neq \min(\mathbb{Q} \setminus \alpha)$  [If  $\varepsilon + s_2 \neq \min(\mathbb{Q} \setminus \alpha)$  we take  $s_3 = s_2$ , else if  $\varepsilon + s_2 = \min(\mathbb{Q} \setminus \alpha)$  then as  $\max(\alpha)$  does not exists there exists a  $s_3 \in \alpha$  with  $s_2 < s_3 \Rightarrow \varepsilon + s_2 < \varepsilon + s_3$ , now  $\varepsilon + s_3 \in \mathbb{Q} \setminus \alpha$  [otherwise  $\varepsilon + s_3 \in \alpha \Rightarrow \varepsilon + s_3 < \varepsilon + s_2$  contradicting  $\varepsilon + s_2 < \varepsilon + s_3$ ], we have then  $\varepsilon + s_3 \neq \min(\mathbb{Q} \setminus \alpha)$ ]. Take now  $s = \max(s_1, s_3)$ . Then as  $0 \in \alpha$  we have as  $\varepsilon + s_2 \in \mathbb{Q} \setminus \alpha$  by 8.1 that  $0 < \varepsilon + s_2 \leq \varepsilon + s$ ,  $\varepsilon + s \in \mathbb{Q} \setminus \alpha$  [if  $\varepsilon + s \in \alpha$   $\varepsilon + s_2 \in \mathbb{Q} \setminus \alpha$ , 8.1  $\varepsilon + s < \varepsilon + s_2$  contradicting  $\varepsilon + s_2 < \varepsilon + s$ ] and  $\varepsilon + s \neq \min(\mathbb{Q} \setminus \alpha)$  hence  $0 < (\varepsilon + s)^{-1} \in \text{rep}(\alpha)$ . As  $s_1, s_3 \in \alpha \Rightarrow s = \max(s_1, s_3) \Rightarrow s \in \alpha$  and from  $0 < s_1 \leq \max(x_1, s_2) \Rightarrow 0 < s$ . The last two results gives us  $s \cdot (\varepsilon + s)^{-1} \in \alpha \odot \text{rep}(\alpha)$ . Now from  $s_1 \leq s$  we have by  $0 < x^{-1}, (1-x)$  that  $\varepsilon = s_1 \cdot (1-x) \cdot x^{-1} \leq s \cdot (1-x) \cdot x^{-1} \Rightarrow 0 < \varepsilon + s \leq s + s \cdot (1-x) \cdot x^{-1} = s \cdot (1 + (1-x) \cdot x^{-1}) \Rightarrow s^{-1} \cdot (1 + (1-x) \cdot x^{-1}) = (s \cdot (1 + (1-x) \cdot x^{-1}))^{-1} \leq (\varepsilon + s)^{-1} \Rightarrow (1 + (1-x) \cdot x^{-1})^{-1} \leq s \cdot (\varepsilon + s)^{-1}$ . Now  $1 + (1-x) \cdot x^{-1} = 1 + x^{-1} - x \cdot x^{-1} = x^{-1} \Rightarrow x \leq s \cdot (\varepsilon + s)^{-1}$ , as  $s \cdot (\varepsilon + s)^{-1} \in \alpha \odot \text{rep}(\alpha)$  we have by 8.3 that  $x \in \alpha \odot \text{rep}(\alpha)$

So we have proved that  $\alpha \odot \text{rep}(\alpha) = \alpha_1$  and thus

$$\begin{aligned}\alpha \cdot \alpha^{-1} &= \alpha \odot \alpha^{-1} \\ &= \alpha \odot \text{rep}(\alpha) \\ &= \alpha_1 \\ &\stackrel{\text{commutativity}}{=} \alpha^{-1} \cdot \alpha\end{aligned}$$

b. ( $\alpha \in \mathbb{R}_-$ ) We have then

$$\begin{aligned}
 \alpha \cdot \alpha^{-1} &= \alpha \cdot (-(\text{rep}(-\alpha))) \\
 &= -(\alpha \cdot (-\alpha)^{-1}) \\
 &= -(-((- \alpha) \cdot (-\alpha)^{-1})) \\
 &= ((-\alpha) \cdot (-\alpha)^{-1}) \\
 &\stackrel{3.b}{=} a_1 \\
 &\stackrel{\text{commutativity}}{=} \alpha^{-1} \cdot \alpha
 \end{aligned}$$

4. (**Distributive**) We have the following cases to consider

a.  $\alpha \in \mathbb{R}_+, \beta \in \mathbb{R}_+, \gamma \in \mathbb{R}_+$  We have then

$$x \in \alpha \cdot (\beta + \gamma) \Rightarrow x \in \alpha \odot (\beta + \gamma)$$

we have then two possible cases

- i. ( $x \leq 0$ ) then  $x \in \alpha \odot \beta$  also  $0 \in \alpha \odot \gamma \Rightarrow x = x + 0 \in (\alpha \odot \beta) + (\alpha \odot \gamma)$
- ii. ( $0 < x$ ) then  $x = s \cdot t$  where  $s \in \alpha \wedge 0 < s$  and  $t \in \beta + \gamma \wedge 0 < t \Rightarrow \exists u \in \beta \wedge v \in \gamma$  with  $t = u + v$ . Using 7.10 we have that  $x = s \cdot t = s \cdot (u + v) = s \cdot u + s \cdot v$ . We have now the following possibilities

- A.  $u \leq 0 \wedge v \leq 0$  this case does not apply because it gives rise to the contradiction  $0 < t = u + v \leq 0 \Rightarrow 0 < 0$ .
- B.  $u \leq 0 \wedge 0 < v \stackrel{0 < s}{\Rightarrow} s \cdot u \leq 0 \Rightarrow s \cdot u \in \alpha \odot \beta \wedge s \cdot t \in \alpha \odot \gamma \Rightarrow x = s \cdot u + s \cdot v \in (\alpha \odot \beta) + (\alpha \odot \gamma) \Rightarrow x \in (\alpha \odot \beta) + (\alpha \odot \gamma)$
- C.  $0 < u \wedge v \leq 0 \Rightarrow s \cdot v \leq 0 \Rightarrow s \cdot u \in \alpha \odot \beta \wedge s \cdot t \in \alpha \odot \gamma \Rightarrow x = s \cdot u + s \cdot v \in (\alpha \odot \beta) + (\alpha \odot \gamma) \Rightarrow x \in (\alpha \odot \beta) + (\alpha \odot \gamma)$
- D.  $0 < u \wedge 0 < v \Rightarrow s \cdot u \in \alpha \odot \beta \wedge s \cdot t \in \alpha \odot \gamma \Rightarrow x = s \cdot u + s \cdot v \in (\alpha \odot \beta) + (\alpha \odot \gamma) \Rightarrow x \in (\alpha \odot \beta) + (\alpha \odot \gamma)$

So we have  $\alpha \odot (\beta + \gamma) \subseteq \alpha \odot \beta + \alpha \odot \gamma$

Consider now

$$x \in \alpha \odot \beta + \alpha \odot \gamma \text{ then } x = r + t \text{ with } r \in \alpha \odot \beta \wedge t \in \alpha \odot \gamma$$

we have then the following possibilities

- i. ( $x \leq 0$ )  $\Rightarrow x \in \alpha \odot (\beta + \gamma)$
- ii. ( $0 < x$ ) we have now the following possible sub-cases
  - A. ( $r \leq 0 \wedge t \leq 0$ ) then we would have the contradiction  $x \leq 0$  so this case does not applies.
  - B. ( $r \leq 0 \wedge 0 < t$ ) then  $\exists u \in \alpha, \exists v \in \gamma$  such that  $t = u \cdot v \wedge 0 < u \wedge 0 < v \Rightarrow t = u \cdot v = u \cdot (0 + v) \in \alpha \odot (\beta + \gamma)$  since  $r \leq 0 \Rightarrow x = r + t \leq 0 + t \Rightarrow x \leq t$  we have by  $t \in \alpha \odot (\beta + \gamma)$  and 8.3 that  $x \in \alpha \odot (\beta + \gamma)$
  - C. ( $0 < r \wedge t \leq 0$ ) then  $\exists u \in \alpha, \exists v \in \gamma$  such that  $r = u \cdot v \Rightarrow r = u \cdot v + u \cdot 0 = u \cdot (v + 0) \in \alpha \odot (\beta + \gamma)$  since  $t \leq 0 \Rightarrow x = r + t \leq r \Rightarrow x \leq r$  and from 8.3 and  $r \in \alpha \odot (\beta + \gamma)$  we would have  $x \in \alpha \odot (\beta + \gamma)$
  - D. ( $0 < r \wedge 0 < t$ ) then there exists  $u \in \alpha, v \in \beta, u' \in \alpha, v' \in \gamma$  with  $r = u \cdot v \wedge t = u' \cdot v' \wedge 0 < u \wedge 0 < v \wedge 0 < u' \wedge 0 < v'$ . We have now to consider the following possibilities
    1. ( $u = u'$ ) then  $x = u \cdot v + u' \cdot v' = u \cdot v + u \cdot v' = u \cdot (v + v') \in \alpha \odot (\beta + \gamma) \Rightarrow x \in \alpha \odot (\beta + \gamma)$
    2. ( $u < u'$ ) then as  $0 < u' \wedge 0 < v + v'$  we have  $u' \cdot (v + v') \in \alpha \odot (\beta + \gamma)$  and from  $u < u' \Rightarrow u \cdot v < u' \cdot v' \Rightarrow x = u \cdot v + u' \cdot v' \leq u' \cdot v + u' \cdot v' = u' \cdot (v + v') \in \alpha \odot (\beta + \gamma) \stackrel{8.3}{\Rightarrow} x \in \alpha \odot (\beta + \gamma)$

3.  $(u' < u)$  then as  $0 < u \wedge 0 < v + v'$  we have  $u \cdot (v + v') \in \alpha \odot (\beta + \gamma)$ .  
 From  $u' < u$  we have  $u' \cdot v' < u \cdot v' \Rightarrow x = u \cdot v + u' \cdot v' < u \cdot v + u \cdot v' = u \cdot (v + v') \in \alpha \odot (\beta + \gamma) \stackrel{8.3}{\Rightarrow} x \in \alpha \odot (\beta + \gamma)$

So we have  $\alpha \odot \beta + \alpha \odot \gamma \subseteq \alpha \odot (\beta + \gamma) \Rightarrow \alpha \odot (\beta + \gamma) = \alpha \odot \beta + \alpha \odot \gamma$ . So finally we have

$$\begin{aligned} \alpha \cdot (\beta + \gamma) &= \alpha \odot (\beta + \gamma) \\ &= \alpha \odot \beta + \alpha \odot \gamma \\ &= \alpha \cdot \beta + \alpha \cdot \gamma \end{aligned}$$

b.  $\alpha \in \mathbb{R}_+, \beta \in \mathbb{R}_+, \gamma \in \mathbb{R}_-$  We have now two sub-cases

- i.  $(\beta + \gamma \in \mathbb{R}_+)$  First as  $\gamma \in \mathbb{R}_-$  we have  $-\gamma \in \mathbb{R}_+$  and thus we have using (4.a) that

$$\begin{aligned} \alpha \cdot (\beta + (-\gamma)) &= \alpha \cdot \beta + \alpha \cdot (-\gamma) \\ &= \alpha \cdot \beta + (-(\alpha \cdot \gamma)) \end{aligned}$$

now note that a  $\beta, -\gamma \in \mathbb{R}_+$  we have  $0 \in \beta, 0 \in -\gamma \Rightarrow 0 = 0 + 0 \in \beta + (-\gamma) \Rightarrow \beta + (-\gamma) \in \mathbb{R}_+$ , next as we have that

$$\beta + \beta = (\beta + (-\gamma)) + (\beta + \gamma)$$

we have then by multiplication both sides with  $\alpha$

$$\begin{aligned} \alpha \cdot (\beta + \beta) &= \alpha \cdot ((\beta + (-\gamma)) + (\beta + \gamma)) \\ &\stackrel{\beta \in \mathbb{R}_+, \beta + \gamma \in \mathbb{R}_+, \beta + \gamma \in \mathbb{R}_+ \text{ and 4.a.}}{\Rightarrow} \\ \alpha \cdot \beta + \alpha \cdot \beta &= \alpha \cdot (\beta + (-\gamma)) + \alpha \cdot (\beta + \gamma) \\ &\stackrel{\text{previous remark}}{=} \alpha \cdot \beta + (-(\alpha \cdot \gamma)) + \alpha \cdot (\beta + \gamma) \\ &\Rightarrow \\ \alpha \cdot \beta + \alpha \cdot \gamma &= \alpha \cdot (\beta + \gamma) \end{aligned}$$

- ii.  $(\beta + \gamma \in \mathbb{R}_-)$  Then we have  $-(\beta + \gamma) = (-\gamma) + (-\beta) \in \mathbb{R}_+, -\gamma \in \mathbb{R}_+, -\beta \in \mathbb{R}_-$  and thus

$$\begin{aligned} \alpha \cdot (\beta + \gamma) &= -(\alpha \cdot (-(\beta + \gamma))) \\ &= -(\alpha \cdot ((-\beta) + (-\gamma))) \\ &= -(\alpha \cdot ((-\gamma) + (-\beta))) \\ &\stackrel{4.a.i}{=} -(\alpha \cdot (-\gamma) + \alpha \cdot (-\beta)) \\ &= -(-(\alpha \cdot \gamma) + (-(\alpha \cdot \beta))) \\ &= \alpha \cdot \gamma + \alpha \cdot \beta \\ &= \alpha \cdot \beta + \alpha \cdot \gamma \end{aligned}$$

c.  $\alpha \in \mathbb{R}_+, \beta \in \mathbb{R}_+, \gamma = 0$  Here we have

$$\begin{aligned} \alpha \cdot (\beta + \gamma) &= \alpha \cdot (\beta + 0) \\ &= \alpha \cdot \beta \\ &= \alpha \cdot \beta + 0 \\ &= \alpha \cdot \beta + \alpha \cdot 0 \\ &= \alpha \cdot \beta + \alpha \cdot \gamma \end{aligned}$$

d.  $\alpha \in \mathbb{R}_+, \beta \in \mathbb{R}_-, \gamma \in \mathbb{R}_+$  Here we have

$$\begin{aligned} \alpha \cdot (\beta + \gamma) &= \alpha \cdot (\gamma + \beta) \\ &\stackrel{4.b}{=} \alpha \cdot \gamma + \alpha \cdot \beta \\ &= \alpha \cdot \beta + \alpha \cdot \gamma \end{aligned}$$

e.  $\alpha \in \mathbb{R}_+, \beta \in \mathbb{R}_-, \gamma \in \mathbb{R}_-$  then we have  $-\beta, -\gamma \in \mathbb{R}_+$  and then we have

$$\begin{aligned}\alpha \cdot (\beta + \gamma) &= -(\alpha \cdot (-(\beta + \gamma))) \\ &= -(\alpha \cdot ((-\beta) + (-\gamma))) \\ &\stackrel{4.a}{=} -(\alpha \cdot (-\beta) + \alpha \cdot (-\gamma)) \\ &= -(-(\alpha \cdot \beta) + (-(\alpha \cdot \gamma))) \\ &= \alpha \cdot \beta + \alpha \cdot \gamma\end{aligned}$$

f.  $\alpha \in \mathbb{R}_+, \beta \in \mathbb{R}_-, \gamma = 0$  We have then

$$\begin{aligned}\alpha \cdot (\beta + \gamma) &= \alpha \cdot \beta \\ &= \alpha \cdot \beta + 0 \\ &= \alpha \cdot \beta + \alpha \cdot \gamma\end{aligned}$$

g.  $\alpha \in \mathbb{R}_+, \beta = 0, \gamma \in \mathbb{R}_+$  then we have

$$\begin{aligned}\alpha \cdot (\beta + \gamma) &= \alpha \cdot \gamma \\ &= 0 + \alpha \cdot \gamma \\ &= \alpha \cdot \beta + \alpha \cdot \gamma\end{aligned}$$

h.  $\alpha \in \mathbb{R}_+, \beta = 0, \gamma \in \mathbb{R}_-$  then we have

$$\begin{aligned}\alpha \cdot (\beta + \gamma) &= \alpha \cdot \gamma \\ &= 0 + \alpha \cdot \gamma \\ &= \alpha \cdot \beta + \alpha \cdot \gamma\end{aligned}$$

i.  $\alpha \in \mathbb{R}_+, \beta = 0, \gamma = 0$  then we have

$$\begin{aligned}\alpha \cdot (\beta + \gamma) &= \alpha \cdot \gamma \\ &= 0 + \alpha \cdot \gamma \\ &= \alpha \cdot \beta + \alpha \cdot \gamma\end{aligned}$$

j.  $\alpha \in \mathbb{R}_-, \beta \in \mathbb{R}_+, \gamma \in \mathbb{R}_+$  then we have  $-\alpha \in \mathbb{R}_+$  and thus

$$\begin{aligned}\alpha \cdot (\beta + \gamma) &= -((-\alpha) \cdot (\beta + \gamma)) \\ &\stackrel{4.a}{=} -((-\alpha) \cdot \beta + (-\alpha) \cdot \gamma) \\ &= -(-(\alpha \cdot \beta) + (-(\alpha \cdot \gamma))) \\ &= \alpha \cdot \beta + \alpha \cdot \gamma\end{aligned}$$

k.  $\alpha \in \mathbb{R}_-, \beta \in \mathbb{R}_+, \gamma \in \mathbb{R}_-$  then we have  $-\alpha \in \mathbb{R}_+$  and thus

$$\begin{aligned}\alpha \cdot (\beta + \gamma) &= -((-\alpha) \cdot (\beta + \gamma)) \\ &\stackrel{4.b}{=} -((-\alpha) \cdot \beta + (-\alpha) \cdot \gamma) \\ &= -(-(\alpha \cdot \beta) + (-(\alpha \cdot \gamma))) \\ &= \alpha \cdot \beta + \alpha \cdot \gamma\end{aligned}$$

l.  $\alpha \in \mathbb{R}_-, \beta \in \mathbb{R}_+, \gamma = 0$  then we have  $-\alpha \in \mathbb{R}_+$  and thus

$$\begin{aligned}\alpha \cdot (\beta + \gamma) &= -((-\alpha) \cdot (\beta + \gamma)) \\ &\stackrel{4.c}{=} -((-\alpha) \cdot \beta + (-\alpha) \cdot \gamma) \\ &= -(-(\alpha \cdot \beta) + (-(\alpha \cdot \gamma))) \\ &= \alpha \cdot \beta + \alpha \cdot \gamma\end{aligned}$$

m.  $\alpha \in \mathbb{R}_-, \beta \in \mathbb{R}_-, \gamma \in \mathbb{R}_+$  then we have  $-\alpha \in \mathbb{R}_+$  and thus

$$\begin{aligned}\alpha \cdot (\beta + \gamma) &= -((-\alpha) \cdot (\beta + \gamma)) \\ &\stackrel{4.d}{=} -((-\alpha) \cdot \beta + (-\alpha) \cdot \gamma) \\ &= -(-(\alpha \cdot \beta) + (-(\alpha \cdot \gamma))) \\ &= \alpha \cdot \beta + \alpha \cdot \gamma\end{aligned}$$

n.  $\alpha \in \mathbb{R}_-, \beta \in \mathbb{R}_-, \gamma \in \mathbb{R}_-$  then we have  $-\alpha \in \mathbb{R}_+$  and thus

$$\begin{aligned}\alpha \cdot (\beta + \gamma) &= -((- \alpha) \cdot (\beta + \gamma)) \\ &\stackrel{4.e}{=} -((- \alpha) \cdot \beta + (- \alpha) \cdot \gamma) \\ &= -(-(\alpha \cdot \beta) + (-(\alpha \cdot \gamma))) \\ &= \alpha \cdot \beta + \alpha \cdot \gamma\end{aligned}$$

o.  $\alpha \in \mathbb{R}_-, \beta \in \mathbb{R}_-, \gamma = 0$  then we have  $-\alpha \in \mathbb{R}_+$  and thus

$$\begin{aligned}\alpha \cdot (\beta + \gamma) &= -((- \alpha) \cdot (\beta + \gamma)) \\ &\stackrel{4.f}{=} -((- \alpha) \cdot \beta + (- \alpha) \cdot \gamma) \\ &= -(-(\alpha \cdot \beta) + (-(\alpha \cdot \gamma))) \\ &= \alpha \cdot \beta + \alpha \cdot \gamma\end{aligned}$$

p.  $\alpha \in \mathbb{R}_-, \beta = 0, \gamma \in \mathbb{R}_+$  then we have  $-\alpha \in \mathbb{R}_+$  and thus

$$\begin{aligned}\alpha \cdot (\beta + \gamma) &= -((- \alpha) \cdot (\beta + \gamma)) \\ &\stackrel{4.g}{=} -((- \alpha) \cdot \beta + (- \alpha) \cdot \gamma) \\ &= -(-(\alpha \cdot \beta) + (-(\alpha \cdot \gamma))) \\ &= \alpha \cdot \beta + \alpha \cdot \gamma\end{aligned}$$

q.  $\alpha \in \mathbb{R}_-, \beta = 0, \gamma \in \mathbb{R}_-$  then we have  $-\alpha \in \mathbb{R}_+$  and thus

$$\begin{aligned}\alpha \cdot (\beta + \gamma) &= -((- \alpha) \cdot (\beta + \gamma)) \\ &\stackrel{4.h}{=} -((- \alpha) \cdot \beta + (- \alpha) \cdot \gamma) \\ &= -(-(\alpha \cdot \beta) + (-(\alpha \cdot \gamma))) \\ &= \alpha \cdot \beta + \alpha \cdot \gamma\end{aligned}$$

r.  $\alpha \in \mathbb{R}_-, \beta = 0, \gamma = 0$  then we have  $-\alpha \in \mathbb{R}_+$  and thus

$$\begin{aligned}\alpha \cdot (\beta + \gamma) &= -((- \alpha) \cdot (\beta + \gamma)) \\ &\stackrel{4.i}{=} -((- \alpha) \cdot \beta + (- \alpha) \cdot \gamma) \\ &= -(-(\alpha \cdot \beta) + (-(\alpha \cdot \gamma))) \\ &= \alpha \cdot \beta + \alpha \cdot \gamma\end{aligned}$$

s.  $\alpha \in 0, \beta \in \mathbb{R}_+, \gamma \in \mathbb{R}_+$  then we have

$$\begin{aligned}\alpha \cdot (\beta + \gamma) &= 0 \\ &= 0+0 \\ &= \alpha \cdot \beta + \alpha \cdot \gamma\end{aligned}$$

t.  $\alpha \in 0, \beta \in \mathbb{R}_+, \gamma \in \mathbb{R}_-$  then we have

$$\begin{aligned}\alpha \cdot (\beta + \gamma) &= 0 \\ &= 0+0 \\ &= \alpha \cdot \beta + \alpha \cdot \gamma\end{aligned}$$

u.  $\alpha \in 0, \beta \in \mathbb{R}_+, \gamma = 0$  then we have

$$\begin{aligned}\alpha \cdot (\beta + \gamma) &= 0 \\ &= 0+0 \\ &= \alpha \cdot \beta + \alpha \cdot \gamma\end{aligned}$$

v.  $\alpha \in 0, \beta \in \mathbb{R}_-, \gamma \in \mathbb{R}_+$  then we have

$$\begin{aligned}\alpha \cdot (\beta + \gamma) &= 0 \\ &= 0+0 \\ &= \alpha \cdot \beta + \alpha \cdot \gamma\end{aligned}$$

w.  $\alpha \in 0, \beta \in \mathbb{R}_-, \gamma \in \mathbb{R}_-$  then we have

$$\begin{aligned}\alpha \cdot (\beta + \gamma) &= 0 \\ &= 0+0 \\ &= \alpha \cdot \beta + \alpha \cdot \gamma\end{aligned}$$

x.  $\alpha \in 0, \beta \in \mathbb{R}_-, \gamma = 0$  then we have

$$\begin{aligned}\alpha \cdot (\beta + \gamma) &= 0 \\ &= 0+0 \\ &= \alpha \cdot \beta + \alpha \cdot \gamma\end{aligned}$$

y.  $\alpha \in 0, \beta = 0, \gamma \in \mathbb{R}_+$  then we have

$$\begin{aligned}\alpha \cdot (\beta + \gamma) &= 0 \\ &= 0+0 \\ &= \alpha \cdot \beta + \alpha \cdot \gamma\end{aligned}$$

z.  $\alpha \in 0, \beta = 0, \gamma \in \mathbb{R}_-$  then we have

$$\begin{aligned}\alpha \cdot (\beta + \gamma) &= 0 \\ &= 0+0 \\ &= \alpha \cdot \beta + \alpha \cdot \gamma\end{aligned}$$

aa.  $\alpha \in 0, \beta = 0, \gamma = 0$  then we have

$$\begin{aligned}\alpha \cdot (\beta + \gamma) &= 0 \\ &= 0+0 \\ &= \alpha \cdot \beta + \alpha \cdot \gamma\end{aligned}$$

## 5. (Associative)

a.  $\alpha \in \mathbb{R}_+, \beta \in \mathbb{R}_+, \gamma \in \mathbb{R}_+$  First take  $x \in \alpha \cdot (\beta \cdot \gamma)$  then we have the following cases to consider

i. ( $x \leq 0$ ) then  $x \in \{r \in \mathbb{Q} | r \leq 0\} \Rightarrow x \in (\alpha \cdot \beta) \cdot \gamma$

ii. ( $0 < x$ ) then there exists a  $s \in \alpha, t \in (\beta \cdot \gamma)$  with  $x = s \cdot t$  and  $0 < s, t$ , as  $0 < t \in \beta \cdot \gamma$  we must have that there exists a  $u \in \beta, v \in \gamma$  such that  $t = u \cdot v$  and  $0 < u, v$ . So we have  $x = s \cdot t = s \cdot (u \cdot v) \stackrel{7.10}{=} (s \cdot u) \cdot v$ . But  $s \cdot u \in \alpha \cdot \beta \Rightarrow x = (s \cdot u) \cdot v \in (\alpha \cdot \beta) \cdot \gamma \Rightarrow x \in (\alpha \cdot \beta) \cdot \gamma$

it follows that  $\alpha \cdot (\beta \cdot \gamma) \subseteq (\alpha \cdot \beta) \cdot \gamma$ . Now if  $x \in (\alpha \cdot \beta) \cdot \gamma \stackrel{\text{commutativity}}{=} \gamma \cdot (\alpha \cdot \beta) \stackrel{\text{commutativity}}{=} \gamma \cdot (\beta \cdot \alpha) \subseteq$  we have just proved this  $(\gamma \cdot \beta) \cdot \alpha \stackrel{\text{commutativity twice}}{=} \alpha \cdot (\beta \cdot \gamma) \Rightarrow x \in \alpha \cdot (\beta \cdot \gamma)$  proving that  $(\alpha \cdot \beta) \cdot \gamma \subseteq \alpha \cdot (\beta \cdot \gamma)$  and thus finally  $\alpha \cdot (\beta \cdot \gamma) = (\alpha \cdot \beta) \cdot \gamma$

b.  $\alpha \in \mathbb{R}_+, \beta \in \mathbb{R}_+, \gamma \in \mathbb{R}_-$  in this case we have  $-\gamma \in \mathbb{R}_+$

$$\begin{aligned}\alpha \cdot (\beta \cdot \gamma) &= \alpha \cdot (-(\beta \cdot (-\gamma))) \\ &= -(\alpha \cdot (\beta \cdot (-\gamma))) \\ &\stackrel{5.a}{=} -((\alpha \cdot \beta) \cdot (-\gamma)) \\ &= (\alpha \cdot \beta) \cdot \gamma\end{aligned}$$

c.  $\alpha \in \mathbb{R}_+, \beta \in \mathbb{R}_+, \gamma = 0$  in this case we have

$$\begin{aligned}\alpha \cdot (\beta \cdot \gamma) &= \alpha \cdot (\beta \cdot 0) \\ &= \alpha \cdot 0 \\ &= 0 \\ &= (\alpha \cdot \beta) \cdot 0 \\ &= (\alpha \cdot \beta) \cdot \gamma\end{aligned}$$

d.  $\alpha \in \mathbb{R}_+, \beta \in \mathbb{R}_-, \gamma \in \mathbb{R}_+$  we have then  $-\beta \in \mathbb{R}_+$  and

$$\begin{aligned}\alpha \cdot (\beta \cdot \gamma) &= \alpha \cdot (-((- \beta) \cdot \gamma)) \\ &= -(\alpha \cdot ((-\beta) \cdot \gamma)) \\ &\stackrel{5.a}{=} -((\alpha \cdot (-\beta)) \cdot \gamma) \\ &= -((-(\alpha \cdot \beta)) \cdot \gamma) \\ &= (\alpha \cdot \beta) \cdot \gamma\end{aligned}$$

e.  $\alpha \in \mathbb{R}_+, \beta \in \mathbb{R}_-, \gamma \in \mathbb{R}_-$  we have then  $-\beta, -\gamma \in \mathbb{R}_+$

$$\begin{aligned}\alpha \cdot (\beta \cdot \gamma) &= \alpha \cdot (-((- \beta) \cdot \gamma)) \\ &= \alpha \cdot ((-\beta) \cdot (-\gamma)) \\ &= (\alpha \cdot (-\beta)) \cdot (-\gamma) \\ &= -((\alpha \cdot (-\beta)) \cdot \gamma) \\ &= -((-(\alpha \cdot \beta)) \cdot \gamma) \\ &= (\alpha \cdot \beta) \cdot \gamma\end{aligned}$$

f.  $\alpha \in \mathbb{R}_+, \beta \in \mathbb{R}_-, \gamma = 0$  here we have

$$\begin{aligned}\alpha \cdot (\beta \cdot \gamma) &= \alpha \cdot (\beta \cdot 0) \\ &= \alpha \cdot 0 \\ &= 0 \\ &= (\alpha \cdot \beta) \cdot 0 \\ &= (\alpha \cdot \beta) \cdot \gamma\end{aligned}$$

g.  $\alpha \in \mathbb{R}_+, \beta = 0, \gamma \in \mathbb{R}_+$  we have now

$$\begin{aligned}\alpha \cdot (\beta \cdot \gamma) &= \alpha \cdot (0 \cdot \gamma) \\ &= \alpha \cdot 0 \\ &= 0 \\ &= 0 \cdot \gamma \\ &= (\alpha \cdot 0) \cdot \gamma \\ &= (\alpha \cdot \beta) \cdot \gamma\end{aligned}$$

h.  $\alpha \in \mathbb{R}_+, \beta = 0, \gamma \in \mathbb{R}_-$  we have now

$$\begin{aligned}\alpha \cdot (\beta \cdot \gamma) &= \alpha \cdot (0 \cdot \gamma) \\ &= \alpha \cdot 0 \\ &= 0 \\ &= 0 \cdot \gamma \\ &= (\alpha \cdot 0) \cdot \gamma \\ &= (\alpha \cdot \beta) \cdot \gamma\end{aligned}$$

i.  $\alpha \in \mathbb{R}_+, \beta = 0, \gamma = 0$  we have now

$$\begin{aligned}\alpha \cdot (\beta \cdot \gamma) &= \alpha \cdot (0 \cdot \gamma) \\ &= \alpha \cdot 0 \\ &= 0 \\ &= 0 \cdot \gamma \\ &= (\alpha \cdot 0) \cdot \gamma \\ &= (\alpha \cdot \beta) \cdot \gamma\end{aligned}$$

j.  $\alpha \in \mathbb{R}_-, \beta \in \mathbb{R}_+, \gamma \in \mathbb{R}_+$  here  $-\alpha \in \mathbb{R}_+$  and

$$\begin{aligned}\alpha \cdot (\beta \cdot \gamma) &= -((- \alpha) \cdot (\beta \cdot \gamma)) \\ &\stackrel{5.a}{=} -((( - \alpha) \cdot \beta) \cdot \gamma) \\ &= -((-(\alpha \cdot \beta)) \cdot \gamma) \\ &= (\alpha \cdot \beta) \cdot \gamma\end{aligned}$$

k.  $\alpha \in \mathbb{R}_-, \beta \in \mathbb{R}_+, \gamma \in \mathbb{R}_-$  here  $-\alpha \in \mathbb{R}_+$  and

$$\begin{aligned}\alpha \cdot (\beta \cdot \gamma) &= -((- \alpha) \cdot (\beta \cdot \gamma)) \\ &\stackrel{5.b}{=} -((( - \alpha) \cdot \beta) \cdot \gamma) \\ &= -((-(\alpha \cdot \beta)) \cdot \gamma) \\ &= (\alpha \cdot \beta) \cdot \gamma\end{aligned}$$

l.  $\alpha \in \mathbb{R}_-, \beta \in \mathbb{R}_+, \gamma = 0$  here  $-\alpha \in \mathbb{R}_+$  and

$$\begin{aligned}\alpha \cdot (\beta \cdot \gamma) &= -((- \alpha) \cdot (\beta \cdot \gamma)) \\ &\stackrel{5.c}{=} -((( - \alpha) \cdot \beta) \cdot \gamma) \\ &= -((-(\alpha \cdot \beta)) \cdot \gamma) \\ &= (\alpha \cdot \beta) \cdot \gamma\end{aligned}$$

m.  $\alpha \in \mathbb{R}_-, \beta \in \mathbb{R}_-, \gamma \in \mathbb{R}_+$  here  $-\alpha \in \mathbb{R}_+$  and

$$\begin{aligned}\alpha \cdot (\beta \cdot \gamma) &= -((- \alpha) \cdot (\beta \cdot \gamma)) \\ &\stackrel{5.d}{=} -((( - \alpha) \cdot \beta) \cdot \gamma) \\ &= -((-(\alpha \cdot \beta)) \cdot \gamma) \\ &= (\alpha \cdot \beta) \cdot \gamma\end{aligned}$$

n.  $\alpha \in \mathbb{R}_-, \beta \in \mathbb{R}_-, \gamma \in \mathbb{R}_-$  here  $-\alpha \in \mathbb{R}_+$  and

$$\begin{aligned}\alpha \cdot (\beta \cdot \gamma) &= -((- \alpha) \cdot (\beta \cdot \gamma)) \\ &\stackrel{5.e}{=} -((( - \alpha) \cdot \beta) \cdot \gamma) \\ &= -((-(\alpha \cdot \beta)) \cdot \gamma) \\ &= (\alpha \cdot \beta) \cdot \gamma\end{aligned}$$

o.  $\alpha \in \mathbb{R}_-, \beta \in \mathbb{R}_-, \gamma = 0$  here  $-\alpha \in \mathbb{R}_+$  and

$$\begin{aligned}\alpha \cdot (\beta \cdot \gamma) &= -((- \alpha) \cdot (\beta \cdot \gamma)) \\ &\stackrel{5.f}{=} -((( - \alpha) \cdot \beta) \cdot \gamma) \\ &= -((-(\alpha \cdot \beta)) \cdot \gamma) \\ &= (\alpha \cdot \beta) \cdot \gamma\end{aligned}$$

p.  $\alpha \in \mathbb{R}_-, \beta = 0, \gamma \in \mathbb{R}_+$  here  $-\alpha \in \mathbb{R}_+$  and

$$\begin{aligned}\alpha \cdot (\beta \cdot \gamma) &= -((- \alpha) \cdot (\beta \cdot \gamma)) \\ &\stackrel{5.g}{=} -((( - \alpha) \cdot \beta) \cdot \gamma) \\ &= -((-(\alpha \cdot \beta)) \cdot \gamma) \\ &= (\alpha \cdot \beta) \cdot \gamma\end{aligned}$$

q.  $\alpha \in \mathbb{R}_-, \beta = 0, \gamma \in \mathbb{R}_-$  here  $-\alpha \in \mathbb{R}_+$  and

$$\begin{aligned}\alpha \cdot (\beta \cdot \gamma) &= -((- \alpha) \cdot (\beta \cdot \gamma)) \\ &\stackrel{5.h}{=} -((( - \alpha) \cdot \beta) \cdot \gamma) \\ &= -((-(\alpha \cdot \beta)) \cdot \gamma) \\ &= (\alpha \cdot \beta) \cdot \gamma\end{aligned}$$

r.  $\alpha \in \mathbb{R}_-, \beta = 0, \gamma = 0$  here  $-\alpha \in \mathbb{R}_+$  and

$$\begin{aligned}\alpha \cdot (\beta \cdot \gamma) &= -((-\alpha) \cdot (\beta \cdot \gamma)) \\ &\stackrel{5.i}{=} -(((-\alpha) \cdot \beta) \cdot \gamma) \\ &= -((-(\alpha \cdot \beta)) \cdot \gamma) \\ &= (\alpha \cdot \beta) \cdot \gamma\end{aligned}$$

s.  $\alpha \in 0, \beta \in \mathbb{R}_+, \gamma \in \mathbb{R}_+$  in this case we have

$$\begin{aligned}\alpha \cdot (\beta \cdot \gamma) &= 0 \cdot (\beta \cdot \gamma) \\ &= 0 \\ &= 0 \cdot \gamma \\ &= (0 \cdot \beta) \cdot \gamma \\ &= (\alpha \cdot \beta) \cdot \gamma\end{aligned}$$

t.  $\alpha \in 0, \beta \in \mathbb{R}_+, \gamma \in \mathbb{R}_-$  in this case we have

$$\begin{aligned}\alpha \cdot (\beta \cdot \gamma) &= 0 \cdot (\beta \cdot \gamma) \\ &= 0 \\ &= 0 \cdot \gamma \\ &= (0 \cdot \beta) \cdot \gamma \\ &= (\alpha \cdot \beta) \cdot \gamma\end{aligned}$$

u.  $\alpha \in 0, \beta \in \mathbb{R}_+, \gamma = 0$  in this case we have

$$\begin{aligned}\alpha \cdot (\beta \cdot \gamma) &= 0 \cdot (\beta \cdot \gamma) \\ &= 0 \\ &= 0 \cdot \gamma \\ &= (0 \cdot \beta) \cdot \gamma \\ &= (\alpha \cdot \beta) \cdot \gamma\end{aligned}$$

v.  $\alpha \in 0, \beta \in \mathbb{R}_-, \gamma \in \mathbb{R}_+$  in this case we have

$$\begin{aligned}\alpha \cdot (\beta \cdot \gamma) &= 0 \cdot (\beta \cdot \gamma) \\ &= 0 \\ &= 0 \cdot \gamma \\ &= (0 \cdot \beta) \cdot \gamma \\ &= (\alpha \cdot \beta) \cdot \gamma\end{aligned}$$

w.  $\alpha \in 0, \beta \in \mathbb{R}_-, \gamma \in \mathbb{R}_-$  in this case we have

$$\begin{aligned}\alpha \cdot (\beta \cdot \gamma) &= 0 \cdot (\beta \cdot \gamma) \\ &= 0 \\ &= 0 \cdot \gamma \\ &= (0 \cdot \beta) \cdot \gamma \\ &= (\alpha \cdot \beta) \cdot \gamma\end{aligned}$$

x.  $\alpha \in 0, \beta \in \mathbb{R}_-, \gamma = 0$  in this case we have

$$\begin{aligned}\alpha \cdot (\beta \cdot \gamma) &= 0 \cdot (\beta \cdot \gamma) \\ &= 0 \\ &= 0 \cdot \gamma \\ &= (0 \cdot \beta) \cdot \gamma \\ &= (\alpha \cdot \beta) \cdot \gamma\end{aligned}$$

y.  $\alpha \in 0, \beta = 0, \gamma \in \mathbb{R}_+$  in this case we have

$$\begin{aligned}\alpha \cdot (\beta \cdot \gamma) &= 0 \cdot (\beta \cdot \gamma) \\ &= 0 \\ &= 0 \cdot \gamma \\ &= (0 \cdot \beta) \cdot \gamma \\ &= (\alpha \cdot \beta) \cdot \gamma\end{aligned}$$

z.  $\alpha \in 0, \beta = 0, \gamma \in \mathbb{R}_-$  in this case we have

$$\begin{aligned}\alpha \cdot (\beta \cdot \gamma) &= 0 \cdot (\beta \cdot \gamma) \\ &= 0 \\ &= 0 \cdot \gamma \\ &= (0 \cdot \beta) \cdot \gamma \\ &= (\alpha \cdot \beta) \cdot \gamma\end{aligned}$$

aa.  $\alpha \in 0, \beta = 0, \gamma = 0$  in this case we have

$$\begin{aligned}\alpha \cdot (\beta \cdot \gamma) &= 0 \cdot (\beta \cdot \gamma) \\ &= 0 \\ &= 0 \cdot \gamma \\ &= (0 \cdot \beta) \cdot \gamma \\ &= (\alpha \cdot \beta) \cdot \gamma\end{aligned}$$

□

**Notation 8.26.** If  $x, y \in \mathbb{R}, x \neq 0$  then we note

1.  $x^{-1} \stackrel{\text{notation}}{=} \frac{1}{x}$
2.  $y \cdot x^{-1} = \frac{y}{x}$

**Theorem 8.27.**  $\forall r, s \in \mathbb{Q}$  we have

1.  $\alpha_r + \alpha_s = \alpha_{r+s}$
2.  $\alpha_r \cdot \alpha_s = \alpha_{r \cdot s}$
3. if  $\alpha_r \neq 0$  then  $(\alpha_r)^{-1} = \alpha_{r^{-1}}$

**Proof.**

1. If  $x \in \alpha_r + \alpha_s$  then there exists  $u \in \alpha_r, v \in \alpha_s$  with  $x = u+v$ . From  $u \in \alpha_r, v \in \alpha_s$  we have  $u < r, v < s \Rightarrow u+v < r+v, r+v < r+s \Rightarrow u+v < r+s \Rightarrow x = u+v \in \alpha_{r+s} \Rightarrow \alpha_r + \alpha_s \subseteq \alpha_{r+s}$ . If  $x \in \alpha_{r+s} \Rightarrow x < r+s \Rightarrow x - r < s$ . Using the density theorem of  $\mathbb{Q}$  (see 7.36) there exists a  $z \in \mathbb{Q}$  with  $x - r < z < s$ . Then  $z \in \alpha_s$  and  $z - (x - r) > 0 \Rightarrow -z < 0 \Rightarrow r + (-z) < r \Rightarrow r - z \in \alpha_r$ . As  $\alpha_r + \alpha_s \ni r - z + z = r - (z - (x - r)) + z = r - z + x - r + z = x \Rightarrow x \in \alpha_r + \alpha_s \Rightarrow \alpha_{r+s} \subseteq \alpha_r + \alpha_s$ . So we conclude that  $\alpha_r + \alpha_s = \alpha_{r+s}$ .

2. We have the following possible cases for  $\alpha_r, \alpha_s$  (see 8.15)

- a.  $(\alpha_r \in \mathbb{R}_+, \alpha_s \in \mathbb{R}_+)$  First as  $0 \in \alpha_r, 0 \in \alpha_s \Rightarrow 0 < r \wedge 0 < s$ . Take  $x \in \alpha_r \odot \alpha_s$  then we have the following possibilities:

- i.  $(x \leq 0)$  as  $0 < r, 0 < s$  we have  $0 < r \cdot s \Rightarrow x \leq 0 < r \cdot s \Rightarrow x \in \alpha_{r \cdot s}$

- ii.  $(0 < x)$  in this case we have  $x = u \cdot v$  where  $u \in \alpha_r, v \in \alpha_s, 0 < u, 0 < v$ . As  $0 < u < r, 0 < v < s \Rightarrow u \cdot v < r \cdot v, r \cdot v < r \cdot s \Rightarrow u \cdot v < r \cdot s \Rightarrow x < r \cdot s \Rightarrow x \in \alpha_{r \cdot s}$

we conclude thus that  $\alpha_r \odot \alpha_s \subseteq \alpha_{r \cdot s}$ . Take now  $x \in \alpha_{r \cdot s} \Rightarrow x < r \cdot s$  we have then the following cases for  $x$

- i.  $(x \leq 0)$  then trivially  $x \in \alpha_r \odot \alpha_s$

ii. **(0 <  $x$ )** then  $0 < x$ . As  $0 < r$  we have by the density theorem (see 7.36) that there exists a  $\varepsilon_1 \in \mathbb{Q}$  such that  $0 < \varepsilon_1 < r$ . From  $x < r \cdot s$  we have  $0 < r \cdot s - x \underset{0 < s \Rightarrow 0 < s^{-1}}{\Rightarrow} 0 < (r \cdot s - x) \cdot s^{-1}$  and by the density theorem again we have that there exists a  $\varepsilon_2 \in \mathbb{Q}$  such that  $0 < \varepsilon_2 < (r \cdot s - x) \cdot s^{-1}$ . Take now  $\varepsilon = \min(\varepsilon_1, \varepsilon_2)$  then  $0 < \varepsilon \leq \varepsilon_1 < r$  and  $0 < \varepsilon \leq \varepsilon_2 < (r \cdot s - x) \cdot s^{-1} \Rightarrow -((r \cdot s - x) \cdot s^{-1}) < -\varepsilon \Rightarrow r - ((r \cdot s - x) \cdot s^{-1}) < r - \varepsilon \Rightarrow x \cdot s^{-1} < r - \varepsilon \underset{0 < x, 0 < s^{-1} \Rightarrow 0 < x \cdot s^{-1}}{\Rightarrow} 0 < x \cdot s^{-1} < r - \varepsilon \underset{7.35}{\Rightarrow} 0 < (r - \varepsilon)^{-1} < x^{-1} \cdot s \underset{0 < x}{\Rightarrow} 0 < x \cdot (r - \varepsilon)^{-1} < s \Rightarrow 0 < x \cdot (r - \varepsilon)^{-1} \in \alpha_s$ . As  $\varepsilon < 0 \Rightarrow 0 < -\varepsilon$  and  $\varepsilon < r$  we have  $0 < r - \varepsilon = r + (-\varepsilon) < r \Rightarrow 0 < r - \varepsilon \in \alpha_r$  and thus  $(r - \varepsilon)(x \cdot (r - \varepsilon)^{-1}) \in \alpha_r \odot \alpha_s \Rightarrow x \in \alpha_r \odot \alpha_s$ . We conclude thus that  $\alpha_{r \cdot s} \subseteq \alpha_r \odot \alpha_s$ .

So we finally reach the conclusion that  $\alpha_{r \cdot s} = \alpha_r \odot \alpha_s$ . As  $\alpha_r, \alpha_s \in \mathbb{R}_+$  we have  $\alpha_r \cdot \alpha_s = \alpha_r \odot \alpha_s = \alpha_{r \cdot s} \Rightarrow \alpha_r \cdot \alpha_s = \alpha_{r \cdot s}$

- b. **( $\alpha_r \in \mathbb{R}_+, \alpha_s \in \mathbb{R}_-$ )** Then  $\alpha_{-s} = -\alpha_s \in \mathbb{R}_+$  and  $\alpha_r \cdot \alpha_s = \alpha_r \cdot (-(-\alpha_s)) = -(\alpha_r \cdot (-\alpha_s)) = -(\alpha_r \cdot \alpha_{-s}) \underset{8.9}{=} -\alpha_{r \cdot (-s)} = -\alpha_{-(r \cdot s)} \underset{8.9}{=} -(-(\alpha_{r \cdot s})) = \alpha_{r \cdot s}$
- c. **( $\alpha_r \in \mathbb{R}_+, \alpha_s = 0$ )**  $\Rightarrow \alpha_s = \alpha_0 \underset{8.4}{\Rightarrow} s = 0$ . Then  $\alpha_r \cdot \alpha_s = \alpha_r \cdot 0 = 0 = \alpha_0 = \alpha_{r \cdot 0} = \alpha_{r \cdot s}$
- d. **( $\alpha_r \in \mathbb{R}_-, \alpha_s \in \mathbb{R}_+$ )** Then  $\alpha_r \cdot \alpha_s = \alpha_s \cdot \alpha_r \underset{(b)}{=} \alpha_{s \cdot r} = \alpha_{r \cdot s}$
- e. **( $\alpha_r \in \mathbb{R}_-, \alpha_s \in \mathbb{R}_-$ )** Then  $\alpha_{-r} = -\alpha_r \in \mathbb{R}_+, \alpha_{-s} = -\alpha_s \in \mathbb{R}_+$  and  $\alpha_r \cdot \alpha_s = -(-(-\alpha_r)) \cdot -(-\alpha_s) = (-\alpha_r) \cdot (-\alpha_s) = \alpha_{-r} \cdot \alpha_s \underset{(a)}{=} \alpha_{(-r) \cdot (-s)} = \alpha_{r \cdot s}$
- f. **( $\alpha_r \in \mathbb{R}_-, \alpha_s = 0$ )**  $\Rightarrow \alpha_s = \alpha_0 \underset{8.4}{\Rightarrow} s = 0$ . Then  $\alpha_r \cdot \alpha_s = \alpha_r \cdot 0 = 0 = \alpha_0 = \alpha_{r \cdot 0} = \alpha_{r \cdot s}$
- g. **( $\alpha_r = 0, \alpha_s \in \mathbb{R}_+$ )**  $\Rightarrow \alpha_r = \alpha_0 \Rightarrow r = 0$ . Then  $\alpha_r \cdot \alpha_s = 0 \cdot \alpha_s = 0 = \alpha_0 = \alpha_{0 \cdot s} = \alpha_{r \cdot s}$
- h. **( $\alpha_r = 0, \alpha_s \in \mathbb{R}_-$ )**  $\Rightarrow \alpha_r = \alpha_0 \Rightarrow r = 0$ . Then  $\alpha_r \cdot \alpha_s = 0 \cdot \alpha_s = 0 = \alpha_0 = \alpha_{0 \cdot s} = \alpha_{r \cdot s}$
- i. **( $\alpha_r = 0, \alpha_s = 0$ )**  $\Rightarrow \alpha_r = \alpha_0 \Rightarrow r = 0$ . Then  $\alpha_r \cdot \alpha_s = 0 \cdot \alpha_s = 0 = \alpha_0 = \alpha_{0 \cdot s} = \alpha_{r \cdot s}$

3. This follows from 8.24.  $\square$

**Theorem 8.28.**  $\langle \mathbb{Q}_{\mathbb{R}}, +, \cdot \rangle$  forms a sub-field of  $\langle \mathbb{R}, +, \cdot \rangle$ , further the function  $i_{\mathbb{Q}}: \mathbb{Q} \rightarrow \mathbb{Q}_{\mathbb{R}}$  defined by  $r \rightarrow \alpha_r$  is a field isomorphism. So we can consider  $\mathbb{Q}_{\mathbb{R}}$  as the set of rational numbers embedded in the set of reals.

**Proof.** First we prove that  $\langle \mathbb{Q}_{\mathbb{R}}, +, \cdot \rangle$  is a sub field

1. If  $x, y \in \mathbb{Q}_{\mathbb{R}} \Rightarrow \exists r, s \in \mathbb{Q} \vdash x = \alpha_r$  and  $y = \alpha_s$  and thus  $x + y = \alpha_r + \alpha_s \underset{\text{previous theorem}}{=} \alpha_{r+s} \in \mathbb{Q}_{\mathbb{R}}$
2. If  $x, y \in \mathbb{Q}_{\mathbb{R}} \Rightarrow \exists r, s \in \mathbb{Q} \vdash x = \alpha_r$  and  $y = \alpha_s$  and thus  $x \cdot y = \alpha_r \cdot \alpha_s \underset{\text{previous theorem}}{=} \alpha_{r \cdot s} \in \mathbb{Q}_{\mathbb{R}}$
3. If  $x \in \mathbb{Q}_{\mathbb{R}} \setminus \{0\}$  then  $\exists r \in \mathbb{Q}$  with  $x = \alpha_r \neq 0 = \alpha_0 \Rightarrow r \in \mathbb{Q} \setminus \{0\} \Rightarrow r^{-1}$  exists and if we take  $x' = \alpha_{r^{-1}}$  then we have  $x' \cdot x \underset{\text{commutativity}}{=} x \cdot x' = \alpha_r \cdot \alpha_{r^{-1}} = \alpha_{r \cdot r^{-1}} = \alpha_1 = 1$ , so  $x^{-1} = x' = \alpha_{r^{-1}} \in \mathbb{Q}_{\mathbb{R}}$
4.  $0 = \alpha_0 \in \mathbb{Q}_{\mathbb{R}}$
5.  $1 = \alpha_1 \in \mathbb{Q}_{\mathbb{R}}$

Next we prove that  $i_{\mathbb{Q}}$  is a bijection

1. **(injectivity)** If  $r, s \in \mathbb{Q}$  is such that  $i_{\mathbb{Q}}(r) = i_{\mathbb{Q}}(s) = \alpha_r = \alpha_s \underset{8.4}{\Rightarrow} r = s$
2. **(surjectivity)** If  $x \in \mathbb{Q}_{\mathbb{R}} \Rightarrow \exists r \in \mathbb{Q} \vdash x = \alpha_r = i_{\mathbb{Q}}(r)$ .

Finally we prove the homeomorphism properties

1. If  $r, s \in \mathbb{Q}_{\mathbb{R}} \Rightarrow i_{\mathbb{Q}}(r) + i_{\mathbb{Q}}(s) = \alpha_r + \alpha_s \underset{\text{previous theorem}}{=} \alpha_{r+s} = i_{\mathbb{Q}}(r+s)$
2. If  $r, s \in \mathbb{Q}_{\mathbb{R}} \Rightarrow i_{\mathbb{Q}}(r \cdot s) = \alpha_{r \cdot s} \underset{\text{previous theorem}}{=} \alpha_r \cdot \alpha_s = i_{\mathbb{Q}}(r) \cdot i_{\mathbb{Q}}(s)$
3.  $i_{\mathbb{Q}}(1) = \alpha_1 = 1$   $\square$

**Corollary 8.29.** *The set  $\mathbb{Q}_{\mathbb{R}}$  is denumerable*

**Proof.** As  $\mathbb{Q}$  is denumerable (see 7.40) there exists a bijection  $\beta: \mathbb{N} \rightarrow \mathbb{Q}$ , as by the previous theorem we have that  $i_{\mathbb{Q}}: \mathbb{Q} \rightarrow \mathbb{Q}_{\mathbb{R}}$  is a bijection we have that  $i_{\mathbb{Q}} \circ \beta: \mathbb{N} \rightarrow \mathbb{Q}_{\mathbb{R}}$  is a bijection proving that  $\mathbb{Q}_{\mathbb{R}}$  is denumerable.  $\square$

### 8.2.3 Power in $\mathbb{R}$

**Definition 8.30.** *As  $\langle \mathbb{R}, \cdot \rangle$  is a abelian semi-group we have by 4.22 that given  $a \in \mathbb{R}$  and  $n \in \mathbb{N}$  that there exists a  $a^n$  such that*

$$\begin{aligned} a^0 &= 1 \\ a^{n+1} &= a^n \cdot a \underset{\text{abelian}}{=} a \cdot a^n \end{aligned}$$

**Theorem 8.31.** *If  $n, n' \in \mathbb{N}$  and  $a \in \mathbb{R}$  then  $a^{n'+n} = a^{n'} \cdot a^n$*

**Proof.** We prove this by induction on  $n$ . So let  $X = \{n \in \mathbb{N} | a^{n'+n} = a^{n'} \cdot a^n\}$  then we have

1. If  $n=0$  then  $a^{n'+n} = a^{n'+0} = a^{n'} = a^{n'} \cdot 1 = a^{n'} \cdot a^0 \Rightarrow 0 \in X$
2. If  $n \in X$  then  $a^{n'+(n+1)} = a^{(n'+n)+1} = a^{(n'+n)} \cdot a \underset{n \in X}{=} (a^{n'} \cdot a^n) \cdot a = a^{n'} \cdot (a^n \cdot a) = a^{n'} \cdot a^{n+1}$  and thus  $n+1 \in X$

Using mathematical induction (see 4.10) we have  $X = \mathbb{N}$  proving the theorem  $\square$

**Theorem 8.32.** *If  $n \in \mathbb{N}$  and  $a, b \in \mathbb{R}$  then we have  $(a \cdot b)^n = a^n \cdot b^n$*

**Proof.** We prove this by induction so take  $\mathcal{S} = \{n \in \mathbb{N} | (a \cdot b)^n = a^n \cdot b^n\}$  then we have

$n=0$ . then  $(a \cdot b)^0 = 1 = 1 \cdot 1 = a^0 \cdot b^0$

$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$ . then  $(a \cdot b)^{n+1} = (a \cdot b)^n \cdot (a \cdot b) = (a^n \cdot b^n) \cdot (a \cdot b) = a^{n+1} \cdot b^{n+1} \square$

**Theorem 8.33.** *In  $\mathbb{R}$  we have*

$$\begin{aligned} 0^n &= 0 \text{ (if } n \neq 0\text{)} \\ 1^n &= 1 \\ (-1)^n &= -1 \text{ or } 1 \\ (-1)^{2 \cdot n} &= 1 \\ (-1)^{2 \cdot n+1} &= -1 \end{aligned}$$

**Proof.**

1. If  $n \neq 0 \Rightarrow \exists m \in \mathbb{N} \vdash n = m+1$  then  $0^n = 0^{(m+1)} = 0^m \cdot 0 = 0$
2.  $1^n = 1$  is proved by induction on  $n$ , let  $X = \{n \in \mathbb{N} | 1^n = 1\}$  then
  - a.  $1^0 = 1 \Rightarrow 0 \in X$
  - b. If  $n \in X \Rightarrow 1^{n+1} = 1^n \cdot 1 \underset{n \in X}{=} 1 \cdot 1 = 1 \Rightarrow n+1 \in X$   
so  $X = \mathbb{N}$
3.  $(-1)^n = \pm 1$  is proved by induction on  $n$ , let  $X = \{n \in \mathbb{N} | (-1)^n = -1 \text{ or } 1\}$  then
  - a.  $(-1)^0 = 1 \Rightarrow 0 \in X$
  - b. If  $n \in X$  then  $(-1)^{n+1} = (-1)^n \cdot (-1) \underset{n \in X}{=} (-1) \cdot (-1) \vee 1 \cdot (-1) = 1 \vee -1 \Rightarrow n+1 \in X$   
so  $X = \mathbb{N}$
4.  $(-1)^{2 \cdot n} = (-1)^{(1+1) \cdot n} = (-1)^{n+n} = (-1)^n \cdot (-1)^n \underset{(3)}{=} (-1) \cdot (-1) \text{ or } 1 \cdot 1 = 1$
5.  $(-1)^{2 \cdot n+1} = (-1)^{2 \cdot n} \cdot (-1) \underset{(4)}{=} 1 \cdot (-1) = -1 \quad \square$

### 8.3 Order relation on $\mathbb{R}$

**Theorem 8.34.** If  $\alpha, \beta \in \mathbb{R}_+$  then  $\alpha + \beta \in \mathbb{R}_+$  and  $\alpha \cdot \beta \in \mathbb{R}_+$

**Proof.** If  $\alpha, \beta \in \mathbb{R}$  then  $0 \in \alpha \wedge 0 \in \beta$  then  $0 = 0 + 0 \in \alpha + \beta \Rightarrow \alpha + \beta \in \mathbb{R}_+$ . Also  $0 \in \{r \in \mathbb{Q} | r \leq 0\} \subseteq \alpha \odot \beta \underset{a, b \in \mathbb{R}_+}{=} \alpha \cdot \beta \Rightarrow \alpha \cdot \beta \in \mathbb{R}_+$   $\square$

We define now relations  $< \subseteq \mathbb{R} \times \mathbb{R}$  and  $\leq \subseteq \mathbb{R} \times \mathbb{R}$  as follows

**Definition 8.35.**  $< = \{(a, \beta) \in \mathbb{R} \times \mathbb{R} | \beta + (-\alpha) \in \mathbb{R}_+\}$ , so  $\alpha < \beta \Leftrightarrow (a, \beta) \in < \Leftrightarrow \beta + (-\alpha) = \beta - \alpha \in \mathbb{R}_+$ .

**Definition 8.36.**  $\leq = \{(\alpha, \beta) \in \mathbb{R} \times \mathbb{R} | \alpha = \beta \vee \beta + (-\alpha) \in \mathbb{R}_+\}$ , so  $\alpha \leq \beta \Leftrightarrow (\alpha, \beta) \in \leq \Leftrightarrow \alpha = \beta \vee \beta + (-\alpha) \in \mathbb{R}_+ \Leftrightarrow \alpha = \beta \vee \alpha < \beta$

**Theorem 8.37.**  $\forall \alpha, \beta \in \mathbb{R}$  we have

1.  $\alpha < \beta$  iff  $\alpha \subset \beta$
2.  $\alpha \leq \beta$  iff  $\alpha \subseteq \beta$

**Proof.** Be care full in this prove  $<$  can mean a relation in  $\mathbb{R}$  or in  $\mathbb{Q}$ , context tells us which is which.

1.
  - a. **( $\alpha < \beta \Rightarrow \alpha \subset \beta$ )** If  $\alpha < \beta \Rightarrow \beta + (-\alpha) \in \mathbb{R}_+ \Rightarrow 0 \in \beta + (-\alpha)$ . If  $r \in \alpha \underset{8.8}{\Rightarrow} -r \notin -\alpha \Rightarrow -r \in \mathbb{Q} \setminus (-\alpha)$ , now as  $0 \in \beta + (-\alpha)$  there exists a  $s \in \beta, t \in -\alpha$  such that  $0 = s + t \Rightarrow s = -t$ . By 8.1,  $t \in -\alpha$  and  $-r \in \mathbb{Q} \setminus (-\alpha)$  we have  $t < (-r) \underset{7.22}{\Rightarrow} r < (-t) = s \in \beta \underset{8.3}{\Rightarrow} r \in \beta$  and thus we have  $\alpha \subseteq \beta$ . If now  $\alpha = \beta \Rightarrow \beta + (-\alpha) = 0 = \alpha_0 \Rightarrow 0 \in \beta + (-\alpha) = \alpha_0 \Rightarrow 0 < 0$  a contradiction, so we must have  $\alpha \neq \beta \Rightarrow \alpha \subset \beta$ .
  - b. **( $\alpha \subset \beta \Rightarrow \alpha < \beta$ )** If  $\alpha \subset \beta \Rightarrow \exists r \in \beta$  with  $r \notin \alpha \Rightarrow r \in \mathbb{Q} \setminus \alpha$  as  $\max(\beta)$  does not exists there exists a  $r' \in \beta$  with  $r < r'$ . If  $r' \in \alpha$  then using 8.1 we have  $r' < r$  contradicting  $r < r'$  and thus we must have  $r' \notin \alpha \Rightarrow r' \in \mathbb{Q} \setminus \alpha$ . Hence  $r, r' \in \mathbb{Q} \setminus \alpha$  and  $r < r' \Rightarrow r' \neq \min(\mathbb{Q} \setminus \alpha) \Rightarrow -r' \in -\alpha$ . We have then as  $r' \in \beta, -r' \in -\alpha$  and  $0 = r' + (-r') \in \beta + (-\alpha) \Rightarrow \beta + (-\alpha) \in \mathbb{R}_+ \Rightarrow \alpha < \beta$
2.
  - a. **( $\alpha \leq \beta \Rightarrow \alpha \subseteq \beta$ )** As  $\alpha \leq \beta$  we have the following two cases
    - i. **( $\alpha < \beta$ )** Using 1.a we have then  $\alpha \subset \beta \Rightarrow \alpha \subseteq \beta$
    - ii. **( $\alpha = \beta$ )**  $\Rightarrow \alpha \subseteq \beta$
  - b. **( $\alpha \subseteq \beta \Rightarrow \alpha \leq \beta$ )** As  $\alpha \subseteq \beta$  we have the following two exclusive cases to consider
    - i. **( $\alpha = \beta$ )** then  $\alpha = \beta \Rightarrow \alpha \leq \beta$
    - ii. **( $\alpha \subset \beta$ )** then using 1.b we have  $\alpha < \beta \Rightarrow \alpha \leq \beta$   $\square$

**Theorem 8.38.**  $\langle \mathbb{R}, \leq \rangle$  forms a partially ordered set that is fully-ordered

**Proof.**

1. **(Reflexivity)** If  $\alpha \in \mathbb{R} \Rightarrow \alpha \subseteq \alpha \underset{8.37}{\Rightarrow} \alpha \leq \alpha$
2. **(Anti-symmetry)** If  $\alpha, \beta \in \mathbb{R}$  with  $\alpha \leq \beta$  and  $\beta \leq \alpha$  then by 8.37 we have  $\alpha \subseteq \beta$  and  $\beta \subseteq \alpha$  so  $\alpha = \beta$
3. **(Transitivity)** If  $\alpha, \beta, \gamma \in \mathbb{R}$  with  $\alpha \leq \beta$  and  $\beta \leq \gamma$  then by 8.37 we have  $\alpha \subseteq \beta$  and  $\beta \subseteq \gamma \Rightarrow \alpha \subseteq \gamma$  and by 8.37 again we have  $\alpha \leq \gamma$
4. **(Fully-ordered)** If  $\alpha, \beta \in \mathbb{R}$  then by 8.15 we have for  $\alpha + (-\beta)$  either
  - a. **( $\alpha + (-\beta) = 0$ )** From this we have  $\alpha = \beta \Rightarrow \alpha \subseteq \beta \underset{8.37}{\Rightarrow} \alpha \leq \beta$

- b.  $(\alpha + (-\beta)) \in \mathbb{R}_+$  From this we have by definition  $\beta \leq \alpha$
- c.  $(\alpha + (-\beta)) \in \mathbb{R}_-$  In this case we have  $-(\alpha + (-\beta)) \in \mathbb{R}_+ \Rightarrow (-\alpha) + \beta \in \mathbb{R}_+ \Rightarrow \beta + (-\alpha) \in \mathbb{R}_+$  and thus  $\alpha \leq \beta$   $\square$

**Theorem 8.39.** *We have the following for the set of reals*

1.  $\alpha \in \mathbb{R}_+$  iff  $0 < \alpha$
2.  $\alpha \in \mathbb{R}_-$  iff  $\alpha < 0$
3. If  $\alpha, \beta \in \mathbb{R}$  then we have the following exclusive possibilities
  - a.  $\alpha < \beta$
  - b.  $\beta < \alpha$
  - c.  $\alpha = \beta$
4. If  $\alpha, \beta \in \mathbb{R}$  with  $\alpha < \beta \Rightarrow -\beta < -\alpha$
5. If  $\alpha, \beta, \gamma \in \mathbb{R}$  with  $\alpha < \beta \Rightarrow \alpha + \gamma < \beta + \gamma$
6. If  $\alpha, \beta \in \mathbb{R}$  and  $\alpha < \beta$  with  $\gamma \in \mathbb{R}_+$  then  $\alpha \cdot \gamma < \beta \cdot \gamma$  [as we have  $\alpha = \beta \Rightarrow \alpha \cdot \lambda = \beta \cdot \lambda$  we have also that  $\alpha \leq \beta$  implies  $\alpha \cdot \lambda \leq \beta \cdot \lambda$ ]
7. If  $\alpha, \beta \in \mathbb{R}$  and  $\alpha < \beta$  with  $\gamma \in \mathbb{R}_-$  then  $\beta \cdot \gamma < \alpha \cdot \gamma$  [as we have  $\alpha = \beta \Rightarrow \alpha \cdot \lambda = \beta \cdot \lambda$  we have also that  $\alpha \leq \beta$  implies  $\beta \cdot \lambda \leq \alpha \cdot \lambda$ ]
8. If  $0 < x \Rightarrow 0 < x^{-1}$
9. If  $0 < x < y \Rightarrow y^{-1} < x^{-1}$
10. If  $x \in \mathbb{R} \Rightarrow 0 \leq x^2$  and if  $x \neq 0$  then  $0 < x^2$
11. If  $x, y \in \mathbb{R}_+ \cup \{0\}$  are such that  $x < y$  then  $x^2 < y^2$  (so  $\mathbb{R}_+ : \{x \in \mathbb{R} | 0 \leq x\} \rightarrow \{x \in \mathbb{R} | 0 \leq x\}$  is a strictly increasing function)
12. If  $\alpha \in \mathbb{R}_+$  and  $n \in \mathbb{N}$  then  $\alpha^n \in \mathbb{R}_+$
13. If  $\alpha \in \mathbb{R}_+$  so that  $0 < \alpha < 1$  then if  $n \in \mathbb{N}$  we have  $0 < \alpha^n < \alpha$
14. If  $\alpha \in \mathbb{R}$  such that  $1 \leq \alpha$ 
  - a. if  $\alpha < \beta$  then  $\alpha < \beta^n$
  - b. if  $\alpha \leq \beta$  then  $\alpha \leq \beta^n$

**Proof.**

1. If  $\alpha \in \mathbb{R}_+$  then  $\alpha + (-0) = \alpha \in \mathbb{R}_+ \Rightarrow 0 < \alpha$ . On the other hand if  $0 < \alpha$  then  $\mathbb{R}_+ \ni \alpha + (-0) = \alpha \Rightarrow \alpha \in \mathbb{R}_+$ .
2. If  $\alpha \in \mathbb{R}_-$  then  $-\alpha \in \mathbb{R}_+ \Rightarrow 0 + (-\alpha) \in \mathbb{R}_+ \Rightarrow \alpha < 0$ . If  $\alpha < 0 \Rightarrow 0 + (-\alpha) \in \mathbb{R}_+ \Rightarrow -\alpha \in \mathbb{R}_+ \Rightarrow \alpha \in \mathbb{R}_-$
3. If  $\alpha, \beta \in \mathbb{R}$  then we have by 8.15 the following exclusive possible cases for  $\alpha + (-\beta)$ 
  - a.  $\alpha + (-\beta) \in \mathbb{R}_+ \Leftrightarrow \beta < \alpha$
  - b.  $\alpha + (-\beta) \in \mathbb{R}_- \Leftrightarrow -(\alpha + (-\beta)) \in \mathbb{R}_+ \Leftrightarrow \beta + (-\alpha) \in \mathbb{R}_+ \Leftrightarrow \alpha < \beta$
  - c.  $\alpha + (-\beta) = 0 \Leftrightarrow \alpha = \beta$
4. If  $\alpha, \beta \in \mathbb{R}$  with  $\alpha < \beta$  then  $\beta + (-\alpha) \in \mathbb{R}_+ \Rightarrow (-\alpha) + (-(-\beta)) \in \mathbb{R}_+ \Rightarrow -\beta < -\alpha$
5. If  $\alpha, \beta, \gamma \in \mathbb{R}$  with  $\alpha < \beta$  then  $\beta + (-\alpha) \in \mathbb{R}_+ \Rightarrow \beta + \gamma + (-\gamma) + (-\alpha) \in \mathbb{R}_+ \Rightarrow (\beta + \gamma) + (-(\alpha + \gamma)) \in \mathbb{R}_+ \Rightarrow \alpha + \gamma < \beta + \gamma$
6. If  $\alpha, \beta \in \mathbb{R}$  with  $\alpha < \beta$  and  $\gamma \in \mathbb{R}_+$  then  $\beta + (-\alpha) \in \mathbb{R}_+$  then using 8.34 we have  $(\beta + (-\alpha)) \cdot \gamma \in \mathbb{R}_+ \Rightarrow \beta \cdot \gamma + (-\alpha) \cdot \gamma \in \mathbb{R}_+ \Rightarrow (\beta \cdot \gamma) + (-(\alpha \cdot \gamma)) \in \mathbb{R}_+ \Rightarrow \alpha \cdot \gamma < \beta \cdot \gamma$
7. If  $\alpha, \beta \in \mathbb{R}$  with  $\alpha < \beta$  and  $\gamma \in \mathbb{R}_-$  then  $\beta + (-\alpha) \in \mathbb{R}_+$  and  $-\gamma \in \mathbb{R}_+$  then using 8.34 we have  $\mathbb{R}_+ \ni (\beta + (-\alpha)) \cdot (-\gamma) = -((\beta + (-\alpha)) \cdot \gamma) = -(\beta \cdot \gamma + (-(\alpha \cdot \gamma))) = (\alpha \cdot \gamma) + (-(\beta \cdot \gamma)) \Rightarrow \beta \cdot \gamma < \alpha \cdot \gamma$

8. Let  $0 < x$  consider then the following possible cases for  $x^{-1}$  are
- $(x^{-1} = 0)$  then  $1 = x \cdot x^{-1} = 0$  contradicting  $1 \neq 0$
  - $(x^{-1} < 0)$  then using (6)  $1 = x \cdot x^{-1} < 0 \cdot x = 0 \Rightarrow 1 < 0$  contradicting  $0 < 1$
  - $(0 < x^{-1})$  this is the only case left and proves thus (8)
9. If  $0 < x < y \xrightarrow{(8)} 0 < x^{-1}, y^{-1}$  and using (6) we have  $1 = x \cdot x^{-1} < y \cdot x^{-1} \Rightarrow 1 < y \cdot x^{-1} \Rightarrow y^{-1} = 1 \cdot y^{-1} < (y \cdot x^{-1}) \cdot y^{-1} = x^{-1} \Rightarrow y^{-1} < x^{-1}$
10. If  $x \in \mathbb{R}$  then we have
- $(x = 0) \Rightarrow x^2 = x \cdot x = 0 \cdot 0 = 0 \Rightarrow 0 \leq x^2$
  - $(x < 0) \Rightarrow -0 < -x \Rightarrow 0 < -x \Rightarrow 0 < (-x) \cdot (-x) = (-1)^2(x \cdot x) = x^2 \Rightarrow 0 < x^2$
  - $(0 < x) \Rightarrow 0 \cdot x < x \cdot x \Rightarrow 0 < x^2$
11. If  $x, y \in \mathbb{R} \cup \{0\}$  is such that  $x < y$  then we have to consider the following cases for  $x$
- $(x = 0)$  then  $x^2 = 0 \cdot 0 < y \Rightarrow x^2 = 0 \underset{(6)}{=} 0 \cdot y < y \cdot y = y^2 \Rightarrow x^2 < y^2$
  - $(0 < x)$  then from  $0 < x < y$  we have  $x, y \in \mathbb{R}_+$  and using (6) we get  $x \cdot x < y \cdot x$  and  $x \cdot y < y \cdot y$  giving  $x \cdot x < y \cdot y \Rightarrow x^2 < y^2$
12. If  $\alpha \in \mathbb{R}_+$  then  $0 < \alpha$  we proceed now by induction on  $n$  to prove that  $\forall n \in \mathbb{N}$  we have  $0 < \alpha^n$ .  
So let  $X = \{n \in \mathbb{N} \mid \text{if } 0 < \alpha \text{ then } 0 < \alpha^n\}$  then we have:
- $\alpha^0 = 1 > 0 \Rightarrow 0 \in X$
  - If  $n \in X$  then if  $0 < \alpha$  we have  $0 < \alpha^n$ . Now  $\alpha^{n+1} = \alpha^n \cdot \alpha \underset{0 < \alpha^n, \alpha \in \mathbb{R}_+ \text{ and (6)}}{\Rightarrow} 0 < \alpha^n \cdot \alpha = \alpha^{n+1} \Rightarrow n+1 \in X$

Using mathematical induction (see 4.10) we have  $X = \mathbb{N}$  proving our assertion.

13. Let  $\alpha \in \mathbb{R}_+$  with  $0 < \alpha < 1$  then if  $n \in \mathbb{N}$  we have by (12) already  $0 < \alpha^n$ . Now to prove  $\alpha^n < \alpha$  take  $S = \{n \in \mathbb{N} \mid \text{if } \alpha < 1 \text{ then } \alpha^n < \alpha\}$  then we have

- if  $n = 1$  then  $\alpha^1 = \alpha < 1$  so that  $1 \in S$
- if  $n \in S$  then  $\alpha^{n+1} = \alpha^n \cdot \alpha < \alpha$  [as  $n \in S \Rightarrow \alpha^n < 1 \underset{0 < \alpha \text{ and (6)}}{\Rightarrow} \alpha^n \cdot \alpha < 1 \cdot \alpha = \alpha$ ]

14.

- a. We prove this by induction so let  $\mathcal{S} = \{n \in \mathbb{N} \mid \alpha < \beta^n\}$  then we have

$1 \in \mathcal{S}$ . this follows from  $\alpha < \beta = \beta^1$

$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$ . as  $n \in \mathcal{S}$  we have  $\alpha < \beta^n$  we have by (6) that  $\alpha \cdot \beta < \beta \cdot \beta^n = \beta^{n+1}$  and  $1 \leq \alpha \Rightarrow \beta \leq \alpha \cdot \beta \underset{\alpha \leq \beta}{\Rightarrow} \alpha < \alpha \cdot \beta < \beta^{n+1}$  proving that  $n+1 \in \mathcal{S}$

- b. We prove this by induction so let  $\mathcal{S} = \{n \in \mathbb{N} \mid \alpha \leq \beta^n\}$  then we have

$1 \in \mathcal{S}$ . this follows from  $\alpha \leq \beta = \beta^1$

$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$ . as  $n \in \mathcal{S}$  we have  $\alpha \leq \beta^n$  we have by (6) that  $\alpha \cdot \beta \leq \beta \cdot \beta^n = \beta^{n+1}$  and  $1 \leq \alpha \Rightarrow \beta \leq \alpha \cdot \beta \underset{\alpha \leq \beta}{\Rightarrow} \alpha \leq \alpha \cdot \beta \leq \beta^{n+1}$  proving that  $n+1 \in \mathcal{S}$

□

**Theorem 8.40.** If  $\alpha, \beta \in \mathbb{R}$  and  $0 \leq \alpha, \beta$  then from  $\alpha + \beta = 0$  we have  $\alpha = \beta = 0$

**Proof.** We have then to consider the following cases

- $(0 < \alpha) \Rightarrow \beta = 0 + \beta < \alpha + \beta = 0 \Rightarrow \beta < 0$  a contradiction
- $(0 < \beta) \Rightarrow \alpha = 0 + \alpha < \alpha + \beta = 0 \Rightarrow \alpha < 0$  a contradiction
- $(0 = \alpha, \beta)$  this is what we have to prove

□

**Lemma 8.41.** If  $r, s \in \mathbb{Q}$  with  $r < s$  then  $\alpha_r < \alpha_s$  (and thus we have from  $r \leq s \Rightarrow r = s$  or  $r < s \Rightarrow \alpha_r = \alpha_s$  or  $\alpha_r < \alpha_s \Rightarrow \alpha_r \leq \alpha_s$ )

**Proof.** If  $r, s \in \mathbb{Q}$  with  $r < s$  then if  $u \in \alpha_r \Rightarrow u < r \Rightarrow u < s \Rightarrow u \in \alpha_s \Rightarrow \alpha_r \subseteq \alpha_s$ . Also  $r < s \Rightarrow r \in \alpha_s$  and  $r \notin \alpha_r \Rightarrow \alpha_r \neq \alpha_s \Rightarrow \alpha_r < \alpha_s$   $\square$

**Theorem 8.42.** *The field isomorphism  $i_{\mathbb{Q}}: \mathbb{Q} \rightarrow \mathbb{Q}_{\mathbb{R}}$  defined by  $q \rightarrow \alpha_q$  is order preserving.*

**Proof.** This is trivial by using the previous lemma. If  $r, s \in \mathbb{Q}$  with  $r \leq s \Rightarrow \alpha_r \leq \alpha_s \Rightarrow i_{\mathbb{Q}}(r) \leq i_{\mathbb{Q}}(s)$   $\square$

**Theorem 8.43. (R is conditional complete)**  *$\langle \mathbb{R}, \leq \rangle$  is conditional complete (see 2.179). Using the definition of conditional completeness this means that  $\forall S \subseteq \mathbb{R} \vdash S \neq \emptyset$  for which there exists a  $b \in \mathbb{R}$  such that  $\forall s \in S$  we have  $s \leq b$  (existence of a upper bound) we have the existence of  $\sup(S)$  (a lowest upper bound). In other words: any nonempty set in  $\mathbb{R}$  with a upper bound has a lowest upper bound. Using 2.180 we see that  $\forall S \subseteq \mathbb{R} \vdash S \neq \emptyset$  for which there exists a  $b \in \mathbb{R}$  such that  $\forall s \in S$  we have  $b \leq s$  (existence of a lower bound) there exists a  $\inf(S)$  (in other words every non empty set in  $\mathbb{R}$  with a lower bound has a greatest lower bound).*

**Proof.** Let  $S \subseteq \mathbb{R}$  with  $S \neq \emptyset$  and the existence of a  $\beta \in \mathbb{R}$  such that  $\forall \alpha \in S$  we have  $\alpha \leq \beta$ . Define then  $\gamma = \{r \in \mathbb{Q} \mid \exists \alpha \in S \vdash r \in \alpha\}$  (in another notation  $\gamma = \bigcup_{\alpha \in S} \alpha$ ). We prove now that  $\gamma$  is a Dedekind's cut and thus that  $\gamma \in \mathbb{R}$

1. ( $\gamma \neq \emptyset$ ) As  $S \neq \emptyset$  there exists a  $\alpha \in S \subseteq \mathbb{R}$  and by 8.1, 1 there exists a  $r \in \alpha \subseteq \mathbb{Q} \Rightarrow r \in \gamma \Rightarrow \gamma \neq \emptyset$
2. ( $\gamma \neq \mathbb{Q}$ ) If  $r \in \gamma \Rightarrow \exists \alpha \in S$  such that  $r \in \alpha$  and as  $\alpha \leq \beta \Rightarrow \alpha \subseteq \beta \Rightarrow r \in \beta$  and thus we have  $\gamma \subseteq \beta$  and as by 8.1, 2 we have  $\beta \neq \mathbb{Q}$  we have also  $\gamma \neq \mathbb{Q}$
3. ( $\forall r \in \gamma \wedge \forall s \in \mathbb{Q} \setminus \gamma$  we have  $r < s$ ) So let  $r \in \gamma$  and  $s \in \mathbb{Q} \setminus \gamma$ . From  $r \in \gamma$  we have that  $\exists \alpha \in S \vdash r \in \alpha$ . From  $s \in \mathbb{Q} \setminus \gamma$  we have that  $\forall \tau \in S$  we have that  $s \notin \tau$  so in particular we have  $s \notin \alpha \Rightarrow s \in \mathbb{Q} \setminus \alpha$ . Using 8.1, 3 for  $\alpha$  we have thus  $r < s$ .
4. (**max ( $\gamma$ ) does not exists**) We prove this by contradiction, so assume  $m = \max(\gamma)$  exists. Then as  $m \in \gamma$  there exists a  $\alpha \in S$  such that  $m \in \alpha$  and as  $\max(\alpha)$  does not exists there exist a  $m' \in \alpha$  with  $m < m'$ , and from  $m' \in \alpha \in S$  we derive  $m' \in \gamma \Rightarrow m' < m < m' \Rightarrow m' < m'$  a contradiction. So  $\max(\gamma)$  does not exists.

Next we prove that  $\gamma$  is a upper bound for  $S$ . If  $\alpha \in S$  then  $\forall r \in \alpha$  we have  $r \in \{s \in \mathbb{Q} \mid \exists \alpha \in S \vdash s \in \alpha\} = \gamma \Rightarrow \alpha \subseteq \gamma \Rightarrow \alpha \leq \gamma$ . Finally we prove that it is the least upper bound. So if  $\gamma' \in v(S) = \{u \in \mathbb{R} \mid u \text{ is a upper bound of } S\}$  then if  $r \in \gamma \Rightarrow \exists \alpha \in S \vdash r \in \alpha$ , as  $\alpha \leq \gamma' \Rightarrow \alpha \subseteq \gamma'$  we have  $r \in \gamma'$  and thus we have  $\gamma \subseteq \gamma' \Rightarrow \gamma \leq \gamma'$  and thus is a least element of  $v(S)$  and thus we have  $\gamma = \sup(S)$   $\square$

**Theorem 8.44.** *Let  $S \subseteq \mathbb{R}$  then we have for  $-S = \{-s, s \in S\}$  that*

1. *If  $\sup(S)$  exists then  $\inf(-S)$  exists and  $\inf(-S) = -\sup(S)$*
2. *If  $\inf(S)$  exists then  $\sup(-S)$  exists and  $\sup(-S) = -\inf(S)$*

**Proof.**

1. As  $\sup(S)$  exists we have  $\forall s \in -S$  that  $-s \in S$  so that  $-s \leq \sup(S)$  or  $-\sup(S) \leq s$  so  $-S$  is bounded below by  $-\sup(S)$  proving by the conditional completeness of  $\mathbb{R}$  [see 8.43] that

$$\inf(-S) \text{ exists and } -\sup(S) \leq \inf(-S). \quad (8.1)$$

Assume now that  $-\sup(S) < \inf(-S)$  then  $-\inf(-S) < \sup(S)$  so there exist a  $s \in S$  such that  $-\inf(-S) < s \leq \sup(S)$  so that  $-s < \inf(-S)$  which as  $-s \in -S$  gives  $-s < \inf(-S) \leq -s$  a contradicton. So we must have that  $\inf(-S) \leq -\sup(S)$  which together with (8.1) proves that

$$\inf(-S) = -\sup(S)$$

2. As  $\inf(S)$  exists we have  $\forall s \in -S$  that  $-s \in S$  so that  $\inf(S) \leq -s$  or  $s \leq -\inf(S)$  so  $-S$  is bounded above by  $-\inf(S)$  proving by the conditional completeness of  $\mathbb{R}$  [see 8.43] that

$$\sup(-S) \text{ exists and } \sup(-S) \leq -\inf(S) \quad (8.2)$$

Assume now that  $\sup(-S) < -\inf(S)$  then  $\inf(S) < -\sup(-S)$  so there exist a  $s \in S$  such that  $\inf(S) \leq s < -\sup(-S)$  so that  $\sup(-S) < -s$  which as  $-s \in -S$  gives  $-s \leq \sup(-S) < -s$  a contradiction. So we must have that  $-\inf(S) \leq \sup(-S)$  which together with (8.2) proves that

$$\sup(-S) = -\inf(S)$$

□

**Theorem 8.45.** *If  $S \subseteq \mathbb{R}$  has a supremum then if  $\alpha \in \mathbb{R}$  with  $\alpha \geq 0$  then  $\alpha \cdot S = \{\alpha \cdot s \mid s \in S\}$  has a supremum where  $\sup(\{\alpha \cdot s \mid s \in S\}) = \alpha \cdot \sup(S)$*

**Proof.** If  $\alpha = 0$  then  $\alpha \cdot S = \{0\}$  with  $\sup(\alpha \cdot S) = \sup(\{0\}) = 0$  so we only have to prove the case that  $\alpha > 0$ . By the definition of  $\sup(S)$  we have that  $\forall s \in S$  that  $s \leq \sup(S)$  so if  $y \in \alpha \cdot S$  then  $y = \alpha \cdot s \leq \alpha \cdot \sup(S)$  [as  $\alpha > 0$ ] proving that  $\alpha \cdot \sup(S)$  is an upper bound for  $\alpha \cdot S$ . So by the above theorem  $\alpha \cdot S$  has a  $\sup(\alpha \cdot S)$  with by definition  $\sup(\alpha \cdot S) \leq \alpha \cdot \sup(S)$ . If now  $\sup(\alpha \cdot S) < \alpha \cdot \sup(S)$  then  $\frac{1}{\alpha} \cdot \sup(\alpha \cdot S) < \sup(S)$  and there exists a  $s \in S$  with  $\frac{1}{\alpha} \cdot \sup(\alpha \cdot S) < s \leq \sup(S)$  giving  $\sup(\alpha \cdot S) < \alpha \cdot s \leq \alpha \cdot \sup(S)$  which as  $\alpha \cdot s \in \alpha \cdot S$  contradicts the fact that  $\sup(\alpha \cdot S)$  is the supremum of  $\alpha \cdot S$  so we must have  $\sup(\alpha \cdot S) = \alpha \cdot \sup(S)$ . □

**Theorem 8.46.** *If  $S, T \subseteq \mathbb{R}$  have a supremum then  $S + T = \{s + t \mid s \in S \wedge t \in T\}$  has a supremum with  $\sup(S + T) = \sup(S) + \sup(T)$*

**Proof.** Let  $r \in S + T$  then  $\exists s \in S, \exists t \in T$  such that  $r = s + t$ . Using the fact that the supremum is an upper bound we have that  $r = s + t \leq \sup(S) + \sup(T)$ . Using the definition of the supremum we have then that

$$\sup(S + T) \leq \sup(S) + \sup(T) \quad (8.3)$$

Assume now that  $\sup(S + T) < \sup(S) + \sup(T)$  then if we take  $\varepsilon = \sup(S) + \sup(T) - \sup(S + T)$  we have that  $0 < \varepsilon$ . So  $\sup(S) - \frac{\varepsilon}{2} < \sup(S)$  and  $\sup(T) - \frac{\varepsilon}{2} < \sup(T)$  and by the definition of the supremum there exists a  $s \in S$  and a  $t \in T$  such that  $\sup(S) - \frac{\varepsilon}{2} < s \leq \sup(S) + \frac{\varepsilon}{2}$  and  $\sup(T) - \frac{\varepsilon}{2} < t \leq \sup(T)$ . So  $s + t > \sup(S) - \frac{\varepsilon}{2} + \sup(T) - \frac{\varepsilon}{2} = \sup(S) + \sup(T) - \varepsilon = \sup(S) + \sup(T) - \sup(S) - \sup(T) + \sup(S + T) = \sup(S + T)$  or  $s + t > \sup(S + T)$  which conflict the fact that the supremum is an upper bound. Hence  $\sup(S) + \sup(T) \leq \sup(S + T)$  which together with (8.3) proves that

$$\sup(S + T) = \sup(S) + \sup(T) \quad \square$$

**Corollary 8.47.** *If  $S \subseteq \mathbb{R}$  has a supremum and  $x \in \mathbb{R}$  then  $S + x = \{s + x \mid s \in S\}$  has a supremum with  $\sup(S + x) = \sup(S) + x$*

**Proof.** Define  $T = \{x\}$  then  $\sup(T) = x$  exists and  $S + T = \{s + t \mid s \in S \wedge t \in T\} = \{s = x \mid s \in S\} = S + x$ . Hence using 8.46 we have that  $\sup(S + x)$  exists and  $\sup(S + x) = \sup(S) + \sup(T) = \sup(S) + x$ . □

**Theorem 8.48.** *If  $S, T \subseteq \mathbb{R}$  have a infimum then  $S + T = \{s + t \mid s \in S \wedge t \in T\}$  has a infimum with  $\inf(S + T) = \inf(S) + \inf(T)$*

**Proof.** Let  $r \in S + T$  then  $\exists s \in S, \exists t \in T$  such that  $r = s + t$ . Using the fact that the infimum is a lower bound we have that  $\inf(S) + \inf(T) \leq s + t = r$ . Using the definition of the infimum we have then that

$$\inf(S) + \inf(T) \leq \inf(S + T) \quad (8.4)$$

Assume now that  $\inf(S) + \inf(T) < \inf(S + T)$  then if we take  $\varepsilon = \inf(S + T) - \inf(S) - \inf(T)$  we have that  $0 < \varepsilon$ . So  $\inf(S) < \inf(S) + \frac{\varepsilon}{2}$  and  $\inf(T) < \inf(T) + \frac{\varepsilon}{2}$  and by the definition of the infimum there exists a  $s \in S$  and a  $t \in T$  such that  $\inf(S) \leq s < \inf(S) + \frac{\varepsilon}{2}$  and  $\inf(T) \leq t < \inf(T) + \frac{\varepsilon}{2}$ . So  $s + t < \inf(S) + \inf(T) + \varepsilon = \inf(S + t) + \inf(T) + \inf(S + T) - \inf(S) - \inf(T) = \inf(S + T)$  or  $s + t < \inf(S + T)$  which conflict the fact that the infimum is a lower bound. Hence  $\inf(S + T) \leq \inf(S) + \inf(T)$  which together with (8.4) proves that

$$\inf(S + T) = \inf(S) + \inf(T) \quad \square$$

**Corollary 8.49.** If  $S \subseteq \mathbb{R}$  has a infimum and  $x \in \mathbb{R}$  then  $S + x = \{s + x \mid s \in S\}$  has a infimum with  $\inf(S + x) = \inf(S) + x$

**Proof.** Define  $T = \{x\}$  then  $\inf(T) = x$  exists and  $S + T = \{s + t \mid s \in S \wedge t \in T\} = \{s = x \mid s \in S\} = S + x$ . Hence using 8.48 we have that  $\inf(S + x)$  exists and  $\inf(S + x) = \inf(S) + \inf(T) = \inf(S) + x \quad \square$

**Theorem 8.50.**  $\langle \mathbb{Q}_{\mathbb{R}}, \leq \rangle$  is not conditional complete

**Proof.** Using the fact that  $\langle \mathbb{Q}, \leq \rangle$  is not conditional complete (see 7.38) we have that there exists a nonempty set  $A' \subseteq \mathbb{Q}$  which is bounded above by a  $u' \in \mathbb{Q}$  such that  $\sup(A')$  does not exists. Define now  $A = i_{\mathbb{Q}}(A')$  and  $u = i_{\mathbb{Q}}(u')$  then we have the following:

1. As  $A'$  is not empty there exists a  $x \in A'$  then  $i_{\mathbb{Q}}(x) \in A \Rightarrow A$  is not empty.
2. If  $a \in A \Rightarrow \exists a' \in A'$  such that  $a = i_{\mathbb{Q}}(a')$ , as  $u$  is a upper bound of  $A'$  we have that  $a' \leq u' \Rightarrow a = i_{\mathbb{Q}}(a') \leq i_{\mathbb{Q}}(u') = u \Rightarrow A$  is bounded above by  $u$ .

So  $A$  is a non-empty set that is bounded above. Assume now that  $s = \sup(A)$  exists. Take now  $s' = i_{\mathbb{Q}}^{-1}(s)$  then we have:

1. If  $a' \in A'$  then if  $s' < a'$  we have by the order preserving of  $i_{\mathbb{Q}}$  and the fact that it is injective that  $s = i_{\mathbb{Q}}(s') < i_{\mathbb{Q}}(a') \in A$  so  $s$  is not a upper bound of  $A$  contradicting the fact that  $s$  is the lowest upper bound. So we must have that  $a' \leq s'$  and as  $a'$  was chosen arbitrary  $s'$  is a upper bound of  $A'$
2. Let  $b'$  be another upper bound of  $A'$  and  $b = i_{\mathbb{Q}}(b')$ . If now  $a \in A$  then  $a' = i_{\mathbb{Q}}^{-1}(a) \in A'$  [as  $i_{\mathbb{Q}}(a') = a \in A$ ] and thus  $a' \leq b'$  using then the order preserving of  $i_{\mathbb{Q}}$  we have then  $a = i_{\mathbb{Q}}(a') \leq i_{\mathbb{Q}}(b') = b \Rightarrow b$  is a upper bound of  $A$  and as  $s$  is the least upper bound we have  $s \leq b$ . If now  $b' < s'$  then we have using order preserving and injectivity of  $i_{\mathbb{Q}}$  that  $b = i_{\mathbb{Q}}(b') < i_{\mathbb{Q}}(s') = s \Rightarrow b < s$  contradicting the fact that  $s$  is the lowest upper bound of  $A$ . So we must have that  $s' \leq b'$

From the above we conclude that  $s' = \sup(A')$  is a lower upper bound of  $A'$  contradicting the fact that we have chosen  $A'$  so that no upper bound exists. So we conclude that  $A$  has no lowest upper bound and as  $A$  is a non empty set that has a upper bound we conclude that  $\langle \mathbb{Q}_{\mathbb{R}}, \leq \rangle$  is not conditional complete.  $\square$

**Corollary 8.51.**  $\mathbb{Q}_{\mathbb{R}} \subset \mathbb{R}$  so there exists a  $r \in \mathbb{R}$  that is not in  $\mathbb{Q}$ . In other words we have  $\mathbb{R} \setminus \mathbb{Q}_{\mathbb{R}} \neq \emptyset$ .  $\mathbb{R} \setminus \mathbb{Q}_{\mathbb{R}}$  is called the set of irrational numbers.

**Proof.** As  $\langle \mathbb{Q}_{\mathbb{R}}, \leq \rangle$  is not conditional complete (see previous theorem) there exists a non-empty set  $A \subseteq \mathbb{Q}_{\mathbb{R}}$  with a upper bound  $u$  so that  $\sup(A)$  does not exists in  $\mathbb{Q}_{\mathbb{R}}$ . Because  $\mathbb{Q}_{\mathbb{R}} \subseteq \mathbb{R}$  we have  $\emptyset \neq A \subseteq \mathbb{R}$  and that  $A$  has the upper bound  $u \in \mathbb{Q}_{\mathbb{R}} \subseteq \mathbb{R}$ . As  $\mathbb{R}$  is conditional complete (see 8.43) there exists a lowest upper bound  $s = \sup(A)$ , now if  $s \in \mathbb{Q}_{\mathbb{R}}$  it would be a upper bound of  $A$  and if  $b \in \mathbb{Q}_{\mathbb{R}}$  of  $A$  is another upper bound of  $A$  it is also a upper bound of  $A$  in  $\mathbb{R}$  and thus  $s \leq b$  so  $s$  would be the supremum of  $A$  in  $\mathbb{Q}_{\mathbb{R}}$  contradicting the fact that  $\sup(A)$  does not exists in  $\mathbb{Q}_{\mathbb{R}}$ . So  $s \notin \mathbb{Q}_{\mathbb{R}}$ .  $\square$

**Definition 8.52.**  $\mathbb{Z}_{\mathbb{R}} = \{\alpha_r \mid r \in \mathbb{Z}_{\mathbb{Q}}\} \subseteq \mathbb{Q}_{\mathbb{R}} \subseteq \mathbb{R}$ . Note that  $0 = \alpha_0, 1 = \alpha_1$  are elements of  $\mathbb{Z}_{\mathbb{R}}$ .

**Theorem 8.53.**  $\langle \mathbb{Z}_{\mathbb{R}}, +, \cdot \rangle$  forms a sub ring of  $\langle \mathbb{R}, +, \cdot \rangle$ . Also  $i_{\mathbb{Q}_{\mathbb{Z}}} : \mathbb{Z} \rightarrow \mathbb{Z}_{\mathbb{R}}$  defined by  $x \rightarrow \alpha_{\frac{x}{1}}$  is a ring isomorphism that is also order preserving. So we can consider  $\mathbb{Z}_{\mathbb{R}}$  as the set of integer numbers embedded in the set of reals.

**Proof.** Let  $x, y \in \mathbb{Z}_{\mathbb{R}}$  then we have the existence of a  $u, v \in \mathbb{Z}_{\mathbb{Q}} \subseteq \mathbb{Q}$  so that  $x = \alpha_u$  and  $y = \alpha_v$ . Further using the fact that  $\mathbb{Z}_{\mathbb{Q}}$  is a sub-ring of  $\mathbb{Q}$  (see 7.25) we have that  $u+v \in \mathbb{Z}_{\mathbb{Q}}$  and  $u \cdot v \in \mathbb{Z}_{\mathbb{Q}}$  and  $-u \in \mathbb{Z}_{\mathbb{Q}}$ . This gives then

1.  $x+y = \alpha_u + \alpha_v = \alpha_{u+v} \in \mathbb{Z}_{\mathbb{R}}$  8.27

2.  $x \cdot y = \alpha_u \cdot \alpha_v \stackrel{8.27}{=} \alpha_u \cdot \alpha_v = \alpha_{u \cdot v} \in \mathbb{Z}_{\mathbb{R}}$
3.  $-x = -(\alpha_u) \stackrel{8.9}{=} \alpha_{-u} \in \mathbb{Z}_{\mathbb{R}}$  (as  $-u \in \mathbb{Z}_{\mathbb{Q}}$ )
4.  $1 \in \mathbb{Z}_{\mathbb{Q}}$  (see definition)
5.  $0 \in \mathbb{Z}_{\mathbb{Q}}$  (see definition)

this proves that  $\langle \mathbb{Z}_{\mathbb{R}}, +, \cdot \rangle$  is a sub-ring of  $\langle \mathbb{R}, +, \cdot \rangle$ . Now to prove that  $i_{\mathbb{Q}_{\mathbb{Z}}}$  is a order preserving ring isomorphism we first prove that it is a bijection:

1. **(injective)** If  $x, y \in \mathbb{Z}$  are such that  $\alpha_{\frac{x}{1}} = \alpha_{\frac{y}{1}}$  then by 8.4 we have  $\frac{x}{1} = \frac{y}{1} \Rightarrow x \cdot 1 = y \cdot 1 \Rightarrow x = y$  proving injectivity.
2. **(surjective)** If  $z \in \mathbb{Z}_{\mathbb{R}}$  then there exists a  $r \in \mathbb{Z}_{\mathbb{Q}}$  such that  $z = \alpha_r$ . As  $r \in \mathbb{Z}_{\mathbb{Q}}$  there exists a  $u \in \mathbb{Z}$  such that  $r = \frac{u}{1}$ , but then  $z = \alpha_{\frac{u}{1}} = i_{\mathbb{Q}_{\mathbb{Z}}}(u)$  proving surjectivity.

Finally to prove that  $i_{\mathbb{Q}_{\mathbb{Z}}}$  is a order preserving ring isomorphism, note that

1.  $\forall x, y \in \mathbb{Z}$  we have that  $\frac{x+y}{1} = \frac{x}{1} + \frac{y}{1}$  so  $i_{\mathbb{Q}_{\mathbb{Z}}}(x+y) = \alpha_{\frac{x+y}{1}} = \alpha_{\frac{x}{1} + \frac{y}{1}} \stackrel{8.27}{=} \alpha_{\frac{x}{1}} + \alpha_{\frac{y}{1}} = i_{\mathbb{Q}_{\mathbb{Z}}}(x) + i_{\mathbb{Q}_{\mathbb{Z}}}(y)$
2.  $\forall x, y \in \mathbb{Z}$  we have that  $\frac{x \cdot y}{1} = \frac{x}{1} \cdot \frac{y}{1}$  so  $i_{\mathbb{Q}_{\mathbb{Z}}}(x \cdot y) = \alpha_{\frac{x \cdot y}{1}} = \alpha_{\frac{x}{1} \cdot \frac{y}{1}} \stackrel{8.27}{=} \alpha_{\frac{x}{1}} \cdot \alpha_{\frac{y}{1}} = i_{\mathbb{Q}_{\mathbb{Z}}}(x) \cdot i_{\mathbb{Q}_{\mathbb{Z}}}(y)$
3.  $\forall x, y \in \mathbb{Z}$  with  $x \leq y$  we have by 7.25 that  $\frac{x}{1} = i_{\mathbb{Z}}(x) \leq i_{\mathbb{Z}}(y) = \frac{y}{1} \Rightarrow \frac{x}{1} \leq \frac{y}{1}$ . Using 8.41 we have then  $\alpha_{\frac{x}{1}} \leq \alpha_{\frac{y}{1}}$  and thus  $i_{\mathbb{Q}_{\mathbb{Z}}}(x) \leq i_{\mathbb{Q}_{\mathbb{Z}}}(y)$ .

□

**Corollary 8.54.**  $\mathbb{Z}_{\mathbb{R}}$  is denumerable

**Proof.** As  $\mathbb{Z}$  is denumerable (see 6.59) there exists a bijection  $\beta: \mathbb{N} \rightarrow \mathbb{Z}$  and as  $i_{\mathbb{Q}_{\mathbb{Z}}}: \mathbb{Z} \rightarrow \mathbb{Z}_{\mathbb{R}}$  is a bijection by the previous theorem we have the bijection  $i_{\mathbb{Q}_{\mathbb{Z}}} \circ \beta: \mathbb{N} \rightarrow \mathbb{Z}_{\mathbb{R}}$  proving that  $\mathbb{Z}_{\mathbb{R}}$  is denumerable. □

**Definition 8.55.**  $\mathbb{N}_{\mathbb{R}} = \{\alpha_r \mid r \in \mathbb{N}_{\mathbb{Q}}\} \subseteq \mathbb{Z}_{\mathbb{R}} \subseteq \mathbb{Q}_{\mathbb{R}} \subseteq \mathbb{R}$  (for definition of  $\mathbb{N}_{\mathbb{Q}}$  see 7.29). Note that  $0 = \alpha_0, 1 = \alpha_1$  are elements of  $\mathbb{Z}_{\mathbb{R}}$

**Theorem 8.56.**  $\langle \mathbb{N}_{\mathbb{R}}, + \rangle$  forms a sub-semi-group of  $\langle \mathbb{R}, + \rangle$ ,  $\langle \mathbb{N}_{\mathbb{R}}, \cdot \rangle$  forms a sub-semi-group of  $\langle \mathbb{R}, \cdot \rangle$ . Also  $i_{\mathbb{Q}_{\mathbb{N}}}: \mathbb{N} \rightarrow \mathbb{N}_{\mathbb{R}}$  defined by  $n \rightarrow \alpha_r$  where  $r = \frac{i_{\mathbb{N}}(n)}{1}$  is a semi-group isomorphism that is also order preserving (for the addition  $+$  and multiplication  $\cdot$ ). So we can consider  $\mathbb{N}_{\mathbb{R}}$  as the set of natural number embedded in the set of reals. Also  $0 = i_{\mathbb{Q}_{\mathbb{N}}}(0)$  and  $\forall n \in \mathbb{N}_{\mathbb{R}}$  we have  $0 \leq n$

**Proof.** Let  $x, y \in \mathbb{N}_{\mathbb{R}}$  then there exists  $r, s \in \mathbb{N}_{\mathbb{Q}}$  such that  $x = \alpha_r$  and  $y = \alpha_s$ . From  $r, s \in \mathbb{N}_{\mathbb{Q}}$  we have the existence of  $u, v \in \mathbb{N}_{\mathbb{Z}}$  such that  $r = \frac{u}{1}$  and  $s = \frac{v}{1}$ . Using the above gives then

1.  $x+y = \alpha_r + \alpha_s \stackrel{8.27}{=} \alpha_{r+s}$ . Also  $r+s = \frac{u}{1} + \frac{v}{1} = \frac{u \cdot 1 + v \cdot 1}{1 \cdot 1} = \frac{u+v}{1}$  where  $u+v \in \mathbb{N}_{\mathbb{Z}}$  (see 6.22) and this gives that  $r+s \in \mathbb{N}_{\mathbb{Q}} \Rightarrow x+y = \alpha_{r+s} \in \mathbb{N}_{\mathbb{R}}$
2.  $0 \in \mathbb{N}_{\mathbb{R}}$  (see note in the definition)

we therefore conclude that  $\langle \mathbb{N}_{\mathbb{R}}, + \rangle$  is indeed a sub-semi-group of  $\langle \mathbb{R}, + \rangle$ .

Next we prove that  $\langle \mathbb{N}_{\mathbb{R}}, \cdot \rangle$  is a sub-semi-group of  $\langle \mathbb{R}, \cdot \rangle$

1.  $x \cdot y = \alpha_r \cdot \alpha_s \stackrel{8.27}{=} \alpha_{r \cdot s}$ . Also  $r \cdot s = \frac{u}{1} \cdot \frac{v}{1} = \frac{u \cdot v}{1}$  where  $u \cdot v \in \mathbb{N}_{\mathbb{Z}}$  (see 6.22) and thus  $r \cdot s \in \mathbb{N}_{\mathbb{Q}} \Rightarrow x \cdot y = \alpha_{r \cdot s} \in \mathbb{N}_{\mathbb{R}}$
2.  $1 \in \mathbb{N}_{\mathbb{R}}$  (see note in the definition)

Next we prove that  $i_{\mathbb{Q}_{\mathbb{N}}}$  is a order preserving semi-group isomorphism. First we prove that  $i_{\mathbb{Q}_{\mathbb{N}}}$  is a bijection:

1. **(Injective)** If  $n, m$  are such that  $i_{\mathbb{Q}_{\mathbb{N}}}(n) = i_{\mathbb{Q}_{\mathbb{N}}}(m)$  then if  $r = \frac{i_{\mathbb{N}}(n)}{1}, s = \frac{i_{\mathbb{N}}(m)}{1}$  we have that  $\alpha_r = \alpha_s \stackrel{8.4}{\Rightarrow} r = s \Rightarrow i_{\mathbb{N}}(n) \cdot 1 = i_{\mathbb{N}}(m) \cdot 1 \Rightarrow i_{\mathbb{N}}(n) = i_{\mathbb{N}}(m) \stackrel{i_{\mathbb{N}} \text{ is a bijection}}{\Rightarrow} n = m$

2. (**Surjective**) If  $x \in \mathbb{N}_{\mathbb{R}}$  then there exists a  $r \in \mathbb{N}_{\mathbb{Q}}$  such that  $x = \alpha_r$ , as  $r \in \mathbb{N}_{\mathbb{Q}}$  there exists a  $u \in \mathbb{N}_{\mathbb{Z}}$  such that  $r = \frac{u}{1}$ , as  $u \in \mathbb{N}_{\mathbb{Z}}$  there exists by 6.22 a  $n \in \mathbb{N}$  such that  $u = i_{\mathbb{N}}(n) \Rightarrow r = \frac{i_{\mathbb{N}}(n)}{1} \Rightarrow x = \alpha_r = i_{\mathbb{Q}_{\mathbb{N}}}(n)$

Next we prove that it is a semi-group isomorphism

1. If  $n, m \in \mathbb{N}$  then  $n+m \in \mathbb{N}$  and by 6.22 we have  $i_{\mathbb{N}}(n+m) = i_{\mathbb{N}}(n) + i_{\mathbb{N}}(m)$  and thus  $\frac{i_{\mathbb{N}}(n+m)}{1} = \frac{i_{\mathbb{N}}(n) + i_{\mathbb{N}}(m)}{1} = \frac{i_{\mathbb{N}}(n)}{1} + \frac{i_{\mathbb{N}}(m)}{1}$ . So we have that  $i_{\mathbb{Q}_{\mathbb{N}}}(n+m) = \alpha_{\frac{i_{\mathbb{N}}(n+m)}{1}} = \alpha_{\frac{i_{\mathbb{N}}(n)}{1} + \frac{i_{\mathbb{N}}(m)}{1}} \stackrel{8.27}{=} \alpha_{\frac{i_{\mathbb{N}}(n)}{1}} + \alpha_{\frac{i_{\mathbb{N}}(m)}{1}} = i_{\mathbb{Q}_{\mathbb{N}}}(n) + i_{\mathbb{Q}_{\mathbb{N}}}(m)$
2. If  $n, m \in \mathbb{N}$  then  $n \cdot m \in \mathbb{N}$  and by 6.22 we have  $i_{\mathbb{N}}(n \cdot m) = i_{\mathbb{N}}(n) \cdot i_{\mathbb{N}}(m)$  and thus  $\frac{i_{\mathbb{N}}(n \cdot m)}{1} = \frac{i_{\mathbb{N}}(n) \cdot i_{\mathbb{N}}(m)}{1} = \frac{i_{\mathbb{N}}(n)}{1} \cdot \frac{i_{\mathbb{N}}(m)}{1}$ . So we have that  $i_{\mathbb{Q}_{\mathbb{N}}}(n \cdot m) = \alpha_{\frac{i_{\mathbb{N}}(n \cdot m)}{1}} = \alpha_{\frac{i_{\mathbb{N}}(n)}{1} \cdot \frac{i_{\mathbb{N}}(m)}{1}} \stackrel{8.27}{=} \alpha_{\frac{i_{\mathbb{N}}(n)}{1}} \cdot \alpha_{\frac{i_{\mathbb{N}}(m)}{1}} = i_{\mathbb{Q}_{\mathbb{N}}}(n) \cdot i_{\mathbb{Q}_{\mathbb{N}}}(m)$

Next to prove order preserving:

1. If  $n, m \in \mathbb{N} \vdash n \leq m$  then using 6.31 we have  $i_{\mathbb{N}}(n) \leq i_{\mathbb{N}}(m)$  and then using 7.25 we have  $\frac{i_{\mathbb{N}}(n)}{1} = i_{\mathbb{Z}}(i_{\mathbb{N}}(n)) \leq i_{\mathbb{Z}}(i_{\mathbb{N}}(m)) = \frac{i_{\mathbb{N}}(m)}{1}$ . So finally using 8.41 we have  $i_{\mathbb{Q}_{\mathbb{N}}}(n) = \alpha_{\frac{i_{\mathbb{N}}(n)}{1}} \leq \alpha_{\frac{i_{\mathbb{N}}(m)}{1}} = i_{\mathbb{Q}_{\mathbb{N}}}(m)$  proving that  $i_{\mathbb{Q}_{\mathbb{N}}}$  is order preserving.

Next if  $0 \in \mathbb{N}$  then  $i_{\mathbb{N}}(0) = \sim[(s(0), 1)] = \sim[(1, 1)] = 0 \in \mathbb{Z} \Rightarrow \frac{i_{\mathbb{N}}(0)}{1} = \frac{0}{1} = 0 \in \mathbb{Q} \Rightarrow i_{\mathbb{Q}_{\mathbb{N}}}(0) = \alpha_0 = 0 \in \mathbb{R}$ .

Also if  $1 \in \mathbb{N} \Rightarrow i_{\mathbb{N}}(1) = \sim[(s(1), 1)] = \sim[(2, 1)] = 1 \in \mathbb{Z} \Rightarrow \frac{i_{\mathbb{N}}(1)}{1} = \frac{1}{1} = 1 \in \mathbb{Q} \Rightarrow i_{\mathbb{Q}_{\mathbb{N}}}(1) = \alpha_1 = 1 \in \mathbb{R}$ .

Finally to prove that  $\forall n \in \mathbb{N}_{\mathbb{R}}$  we have  $0 \leq n$ . If  $n \in \mathbb{N}_{\mathbb{R}}$  then by the above there exists a  $m \in \mathbb{N}$  such that  $i_{\mathbb{Q}_{\mathbb{N}}}(m) = n$ . Using 4.47 we have  $0 \leq m$  and using the order preserving properties of  $i_{\mathbb{Q}_{\mathbb{N}}}$  we have  $0 = i_{\mathbb{Q}_{\mathbb{N}}}(0) \leq i_{\mathbb{Q}_{\mathbb{N}}}(m) = n \Rightarrow 0 \leq n$   $\square$

As  $u = i_{\mathbb{Q}_{\mathbb{N}}}$  is a bijection.

**Corollary 8.57.**  $\mathbb{N}_{\mathbb{R}}$  is denumerable

**Theorem 8.58.** We have the following properties concerning  $\mathbb{N}_{\mathbb{R}}$  and  $\mathbb{Z}_{\mathbb{R}}$  (here  $-\mathbb{N}_{\mathbb{R}} = \{n \vdash -n \in \mathbb{N}_{\mathbb{R}}\} \subseteq \mathbb{Z}_{\mathbb{R}}$ )

1.  $\mathbb{Z}_{\mathbb{R}} = \mathbb{N}_{\mathbb{R}} \cup (-\mathbb{N}_{\mathbb{R}})$  and  $\mathbb{N}_{\mathbb{R}} \cap (-\mathbb{N}_{\mathbb{R}}) = \{0\}$
2. If  $n \in \mathbb{Z}_{\mathbb{R}}$  and  $0 \leq n \Rightarrow n \in \mathbb{N}_{\mathbb{R}}$  (so  $\mathbb{N}_{\mathbb{R}}$  is indeed the set of positive integers)
3.  $\langle \mathbb{N}_{\mathbb{R}}, + \rangle$  is a semi-group of  $\langle \mathbb{Z}_{\mathbb{R}}, + \rangle$
4.  $\langle \mathbb{N}_{\mathbb{R}}, \cdot \rangle$  is a sub-semi-group of  $\langle \mathbb{Z}_{\mathbb{R}}, \cdot \rangle$

**Proof.**

1. As  $\mathbb{N}_{\mathbb{R}}, -\mathbb{N}_{\mathbb{R}} \subseteq \mathbb{Z}_{\mathbb{R}}$  we have  $\mathbb{N}_{\mathbb{R}} \cup (-\mathbb{N}_{\mathbb{R}}) \subseteq \mathbb{Z}_{\mathbb{R}}$ . If  $x \in \mathbb{Z}_{\mathbb{R}}$  then  $x = \frac{n}{1}$  where  $n \in \mathbb{Z}$  then using 6.24 we have either

- a.  $n \in \mathbb{N}_{\mathbb{Z}} \Rightarrow x \in \mathbb{N}_{\mathbb{R}}$
- b.  $n \in -\mathbb{N}_{\mathbb{Z}} \Rightarrow -n \in \mathbb{N}_{\mathbb{Z}} \Rightarrow -x = \frac{-n}{1} \in \mathbb{N}_{\mathbb{R}} \Rightarrow x \in -\mathbb{N}_{\mathbb{R}}$

and thus  $x \in \mathbb{N}_{\mathbb{Z}} \cup (-\mathbb{N}_{\mathbb{Z}}) \Rightarrow \mathbb{Z}_{\mathbb{R}} \subseteq \mathbb{N}_{\mathbb{R}} \cup (-\mathbb{N}_{\mathbb{R}}) \Rightarrow \mathbb{Z}_{\mathbb{R}} = \mathbb{N}_{\mathbb{R}} \cup (-\mathbb{N}_{\mathbb{R}})$ . Finally if  $x \in \mathbb{N}_{\mathbb{R}} \cap (-\mathbb{N}_{\mathbb{R}})$  then  $0 \leq x$  (as  $x \in \mathbb{N}_{\mathbb{R}}$ ) also  $-x \in \mathbb{N}_{\mathbb{R}} \Rightarrow -x = \frac{n}{1}$  where  $n \in \mathbb{N}_{\mathbb{Z}} \stackrel{6.28}{\Rightarrow} 0 \leq n \stackrel{0 \leq 1 \Rightarrow 0 \leq n}{\Rightarrow} \text{sign}(-x) = 1 \Rightarrow 0 \leq -x \Rightarrow x \leq -0 = 0 \Rightarrow 0 \leq x \leq 0 \Rightarrow x = 0 \Rightarrow \mathbb{N}_{\mathbb{R}} \cap (-\mathbb{N}_{\mathbb{R}}) = \{0\}$

2. If  $n \in \mathbb{Z}_{\mathbb{R}}$  and  $0 \leq n$  then  $\exists q \in \mathbb{Z}_{\mathbb{Q}}$  such that  $n = \alpha_q$  and thus  $\exists z \in \mathbb{Z}$  such that  $q = \frac{z}{1}$ . As  $0 = \alpha_0 \leq \alpha_q$  we must have  $0 \leq q$  [if  $q < 0$  then by 8.41 we have  $\alpha_q < \alpha_0$  contradicting  $\alpha_0 \leq \alpha_q$ ]. As  $0 \leq q \Rightarrow \text{sign}(q) = 1 \Rightarrow 0 \leq z \cdot 1 = z \stackrel{6.28}{\Rightarrow} z \in \mathbb{N}_{\mathbb{Z}} \Rightarrow q = \frac{z}{1} \in \mathbb{N}_{\mathbb{Q}} \Rightarrow n = \alpha_q \in \mathbb{N}_{\mathbb{R}} \Rightarrow n \in \mathbb{N}_{\mathbb{R}}$

3. We can use the same prove as in 8.56

4. We can use the same prove as in 8.56

□

**Theorem 8.59.**  $\langle \mathbb{N}_{\mathbb{R}}, \leq \rangle$  is well-ordered

**Proof.** If  $B \subseteq \mathbb{N}_{\mathbb{R}}$  is a nonempty subset of  $\mathbb{N}_{\mathbb{R}}$  then  $\exists b \in B \xrightarrow{i_{\mathbb{Q}_{\mathbb{N}}} \text{ is surjective}} \exists n \in \mathbb{N} \vdash b = i_{\mathbb{Q}_{\mathbb{N}}}(n) \Rightarrow i_{\mathbb{Q}_{\mathbb{N}}}^{-1}(B)$  is a nonempty subset of  $\mathbb{N}$ . As  $\langle \mathbb{N}, \leq \rangle$  is well-ordered (see 4.52) there exist a least element  $m = \min(i_{\mathbb{Q}_{\mathbb{N}}}^{-1}(B))$  so  $m \in i_{\mathbb{Q}_{\mathbb{N}}}^{-1}(B)$  and  $\forall n \in i_{\mathbb{Q}_{\mathbb{N}}}^{-1}(B)$  we have  $m \leq n$ . Now  $M = i_{\mathbb{Q}_{\mathbb{N}}}(m) \in B$  (as  $m \in i_{\mathbb{Q}_{\mathbb{N}}}^{-1}(B)$ ) and if  $N \in B \xrightarrow{i_{\mathbb{Q}_{\mathbb{N}}} \text{ is surjective}} \exists n \in \mathbb{N} \vdash i_{\mathbb{Q}_{\mathbb{N}}}(n) = N \in B \Rightarrow n \in i_{\mathbb{Q}_{\mathbb{N}}}^{-1}(B) \Rightarrow m \leq n \xrightarrow{i_{\mathbb{Q}_{\mathbb{N}}} \text{ is order preserving}} M = i_{\mathbb{Q}_{\mathbb{N}}}(m) \leq i_{\mathbb{Q}_{\mathbb{N}}}(n) = N \Rightarrow M \leq N$  or  $M = \min(B)$  and  $B$  has a least element. □

**Lemma 8.60.** If  $n \in \mathbb{N}_{\mathbb{R}}$  and  $0 < n \Rightarrow 1 \leq n$ .

**Proof.** As  $n \in \mathbb{N}_{\mathbb{R}}$  there exists by surjectivity of  $i_{\mathbb{Q}_{\mathbb{N}}}$  a  $n' \in \mathbb{N}$  such that  $i_{\mathbb{Q}_{\mathbb{N}}}(n') = n$ . Using 4.47 we have  $0 \leq n'$ , if  $n' = 0 \Rightarrow n = i_{\mathbb{Q}_{\mathbb{N}}}(n') = i_{\mathbb{Q}_{\mathbb{N}}}(0) = 0$  contradicting  $0 < n$ , this leaves us to conclude that  $0 < n'$ . Using 4.51 we have then  $1 = s(n) \leq n'$  and using the order preserving properties of  $i_{\mathbb{Q}_{\mathbb{N}}}$  and the fact that  $i_{\mathbb{Q}_{\mathbb{N}}}(1) = 1$  (see 8.56 that  $1 = i_{\mathbb{Q}_{\mathbb{N}}}(1) \leq i_{\mathbb{Q}_{\mathbb{N}}}(n') = n \Rightarrow 1 \leq n$ ) □

**Theorem 8.61. (Archimedean Property)** If  $x, y \in \mathbb{R}$  with  $0 < x$  then  $\exists n \in \mathbb{N}_{\mathbb{R}}$  such that  $y < n \cdot x$

**Proof.** We have the following cases to consider for  $y$

1. ( $y \leq 0$ ) then take  $1 \in \mathbb{N}_{\mathbb{R}}$  then we have  $y \leq 0 < x \Rightarrow y < x = 1 \cdot x$
2. ( $0 < y$ ) we prove the theorem here by contradiction, so assume that  $\forall n \in \mathbb{N}_{\mathbb{R}}$  we have  $n \cdot x \leq y$ . Define  $A = \{n \cdot x \mid n \in \mathbb{N}_{\mathbb{R}}\}$  then  $\forall t \in A$  we have  $t \leq y$  and thus  $y$  is an upper bound of  $A$ . From conditional completeness (see 9.43)  $\sup(A)$  exists. As  $x > 0$  we have using 9.41 that  $-x < 0 \Rightarrow \sup(A) - x < \sup(A)$  as  $\sup(A)$  is the least upper bound of  $A$   $\sup(A) - x$  is not an upper bound of  $A$  so there exists a  $n \in \mathbb{N}_{\mathbb{R}}$  such that  $\sup(A) - x < n \cdot x \Rightarrow \sup(A) < n \cdot x + x = (n+1) \cdot x$ , now as  $n, 1 \in \mathbb{N}_{\mathbb{R}}$  we have using 8.56 that  $n+1 \in \mathbb{N}_{\mathbb{R}} \Rightarrow (n+1) \cdot x \in A$  and from  $\sup(A) < (n+1) \cdot x \in A$  we have that  $\sup(A)$  is not an upper bound of  $A$  and reach thus a contradiction and the initial hypothesis is false. So we must have that  $\exists n \in \mathbb{N}_{\mathbb{R}}$  such that  $y < n \cdot x$ . □

The following theorem shows some of the consequences of the Archimedean property of  $\mathbb{R}$

**Theorem 8.62.** Let  $x \in \mathbb{R}$  then the following holds:

1.  $\exists n \in \mathbb{N}_{\mathbb{R}} \vdash x < n$
2.  $\exists n \in \mathbb{Z}_{\mathbb{R}} \vdash n \leq x < n+1$
3. If  $0 < x$  then  $\exists n \in \mathbb{N}_{\mathbb{R}} \setminus \{0\}$  such that  $\frac{1}{n} < x$
4.  $\exists n \in \mathbb{Z}_{\mathbb{R}}$  such that  $n-1 \leq x < n$
5.  $\exists n \in \mathbb{Z}_{\mathbb{R}}$  such that  $n < x \leq n+1$
6.  $\exists n \in \mathbb{Z}_{\mathbb{R}}$  such that  $n-1 < x \leq n$

**Proof.**

1. As  $0 < 1 \in \mathbb{R}$  we have by the Archimedean property (see 8.61) that  $\exists n \in \mathbb{N}_{\mathbb{R}}$  such that  $x < 1 \cdot n = n$
2. We have for  $y$  two possibilities

**$0 \leq x$ .** then by (1) we have  $A = \{n \in \mathbb{N} \mid y < n\} \neq \emptyset$  and by the well ordering of  $\mathbb{N}_{\mathbb{R}}$  there exists a least element  $m \in A \Rightarrow x < n$  then as  $m-1 < m$  we have  $\neg(x < m-1) \Rightarrow m-1 \leq x$ . So if we take  $n = m-1 \in \mathbb{N}_{\mathbb{R}} \subseteq \mathbb{Z}_{\mathbb{R}}$  we have  $n \leq x < n+1$

**$x < 0$ .** then we have  $0 < -x$  then we have that  $A = \{n \in \mathbb{N} \mid -y \leq n\} \neq \emptyset$  and there exists a least element  $m \in A \Rightarrow -x \leq m$  and as  $m - 1 < m$  we have  $m - 1 < -x \Rightarrow -m \leq x < -m + 1$ , take  $n = -m \in \mathbb{Z}_{\mathbb{R}}$  then  $n \leq x < n + 1$

3. Using the Archimedean property on 1 (see 8.61) we have  $\exists n \in \mathbb{N}_{\mathbb{R}}$  such that  $1 < x \cdot n$ , as  $n = 0$  would give  $1 = 0$  a contradiction we have that  $0 < n$  and using 8.39 we have then  $0 < \frac{1}{n} \Rightarrow \frac{1}{n} < x$
4. Let  $x \in \mathbb{R}$  then by (2) there exists a  $n \in \mathbb{Z}_{\mathbb{R}}$  such that  $m \leq x \leq m + 1$ , take then  $n = m + 1 \in \mathbb{Z}_{\mathbb{R}}$  then  $n - 1 = m \leq x < m + 1 = n$
5. As  $x \in \mathbb{R}$  there exists by (2) a  $m \in \mathbb{Z}_{\mathbb{R}}$  such that  $m \leq x < m + 1$ . We have then two cases to consider:

**$m = x$ .** take then  $n = x - 1 \in \mathbb{Z}_{\mathbb{R}}$  giving  $n = x - 1 < x = n + 1$

**$m < x$ .** take then  $n = m \in \mathbb{Z}_{\mathbb{R}}$  giving  $n = m < x < m + 1 = n + 1$

6. As  $x \in \mathbb{R}$  then by (4) there exists a  $m \in \mathbb{Z}_{\mathbb{R}}$  such that  $m - 1 \leq x < m$ . We have then two cases to consider

**$m - 1 = x$ .** take then  $n = x - 1 = m - 2 \in \mathbb{Z}_{\mathbb{R}}$  giving  $n = x - 1 < x = n + 1$

**$m - 1 < x$ .** take then  $n = m \in \mathbb{Z}_{\mathbb{R}}$  then  $n - 1 = m - 1 < x < m = n$  □

**Corollary 8.63.** *Let  $x, y \in \mathbb{R}$  then we have*

1. If  $\forall \varepsilon \in \mathbb{R}_+$  we have  $x \leq y + \varepsilon$  then  $x \leq y$
2. If  $\forall \varepsilon \in \mathbb{R}_+$  we have  $x < y + \varepsilon$  then  $x \leq y$
3. If  $\forall n \in \mathbb{N}_{\mathbb{R}} \setminus \{0\}$  we have  $x \leq y + \frac{1}{n}$  then  $x \leq y$
4. If  $\forall n \in \mathbb{N}_{\mathbb{R}} \setminus \{0\}$  we have  $x < y + \frac{1}{n}$  then  $x \leq y$
5. If  $\forall \varepsilon \in \mathbb{R}_+$ ,  $a \in \mathbb{R} \vdash 0 \leq a$  we have  $x \leq y + \varepsilon \cdot a$  then  $x \leq y$
6. If  $\forall \varepsilon \in \mathbb{R}_+$ ,  $a \in \mathbb{R} \vdash 0 \leq a$  we have  $x < y + \varepsilon \cdot a$  then  $x \leq y$
7. If  $\forall n \in \mathbb{N}_{\mathbb{R}} \setminus \{0\}$ ,  $a \in \mathbb{R} \vdash 0 \leq a$  we have  $x \leq y + \frac{1}{n} \cdot a$  then  $x \leq y$
8. If  $\forall n \in \mathbb{N}_{\mathbb{R}} \setminus \{0\}$ ,  $a \in \mathbb{R} \vdash 0 \leq a$  we have  $x < y + \frac{1}{n} \cdot a$  then  $x \leq y$

**Proof.**

1. Assume that  $y < x \Rightarrow 0 < x - y$  then by the above theorem (3) there exists a  $n \in \mathbb{N}_{\mathbb{R}} \setminus \{0\}$  such that for  $\varepsilon = \frac{1}{n}$  we have  $\varepsilon < x - y \Rightarrow y + \varepsilon < x \leq y + \varepsilon \Rightarrow \varepsilon < \varepsilon$  a contradiction so we must have  $x \leq y$
2.  $x < y + \varepsilon \Rightarrow x \leq y + \varepsilon \xrightarrow{(1)} x \leq y$
3. If  $\varepsilon \in \mathbb{N}_{\mathbb{R}} \setminus \{0\}$  then by the previous theorem there exists a  $0 < \frac{1}{n} < \varepsilon \Rightarrow x \leq y + \frac{1}{n} < y + \varepsilon \xrightarrow{(2)} x \leq y$
4.  $x < y + \frac{1}{n} \Rightarrow x \leq y + \frac{1}{n} \xrightarrow{(4)} x \leq y$
5. Given  $\varepsilon \in \mathbb{R}_+$  take  $\varepsilon' = \frac{\varepsilon}{a+1} > 0$  then from  $x \leq y + \varepsilon' \cdot a = y + \frac{a}{a+1} \cdot \varepsilon < y + \varepsilon \xrightarrow{(2)} x \leq y$
6.  $x < y + \varepsilon \cdot a \Rightarrow x \leq y + \varepsilon \cdot a \xrightarrow{(5)} x \leq y$
7. Given  $n \in \mathbb{N}_{\mathbb{R}} \setminus \{0\}$  then as  $0 < \frac{1}{n \cdot (a+1)}$  there exists by the above theorem a  $m \in \mathbb{N}_{\mathbb{R}} \setminus \{0\}$  such that  $\frac{1}{m} < \frac{1}{n \cdot (a+1)} \Rightarrow \frac{1}{m} \cdot a < \frac{1}{n} \cdot \frac{a}{a+1} < \frac{1}{n}$  so that  $x \leq y + \frac{1}{m} \cdot a < y + \frac{1}{n} \xrightarrow{(3)} x \leq y$
8.  $x < y + \frac{1}{n} \cdot a \Rightarrow x \leq y + \frac{1}{n} \cdot a \xrightarrow{(7)} x \leq y$  □

**Theorem 8.64. (Density Theorem)** *If  $x, y \in \mathbb{R}$  such that  $x < y$  then  $\exists r \in \mathbb{Q}_{\mathbb{R}}$  and  $\exists i \in \mathbb{R} \setminus \mathbb{Q}_{\mathbb{R}}$  such that  $x < r < y$  and  $x < i < y$*

**Proof.** First we prove that  $\exists r \in \mathbb{Q}$  such that  $x < r < y$ , we have then the following possible cases for  $x$

1. **(0 < x)** then from  $x < y$  we have  $0 < y - x \Rightarrow (y - x)^{-1}$  is defined and by 9.41 we have  $0 < (y - x)^{-1}$ , then using the Archimedean property (see 9.55) and  $0 < 1$  we have  $\exists n \in \mathbb{N}_{\mathbb{R}}$  such that  $0 < (y - x)^{-1} < n \cdot 1 = n \underset{0 < (y - x)^{-1}}{\Rightarrow} 1 < n \cdot (y - x) \Rightarrow 1 < n \cdot y - n \cdot x \Rightarrow n \cdot x + 1 < n \cdot y$  and  $0 < n$ . As  $0 < x$  and  $n \in \mathbb{N}_{\mathbb{R}} \Rightarrow 0 \leq n \Rightarrow 0 \leq n \cdot x \underset{9.56}{\Rightarrow} \exists m \in \mathbb{N}_{\mathbb{R}} \setminus \{0\}$  such that  $m - 1 \leq n \cdot x < m \Rightarrow m \leq n \cdot x + 1 < n \cdot y$ . Using the fact that  $n, m \in \mathbb{N}_{\mathbb{R}} \subseteq \mathbb{Z}_{\mathbb{R}} \subseteq \mathbb{Q}_{\mathbb{R}}$  so as  $\mathbb{Q}_{\mathbb{R}}$  is a sub-field in  $\mathbb{R}$  (see 8.28) and  $0 < n$  we have  $r = m \cdot n^{-1} \in \mathbb{Q}_{\mathbb{R}}$ , from  $0 < n$  we have  $0 < n^{-1}$  and thus  $r = m \cdot n^{-1} < n \cdot y \cdot n^{-1} = y \Rightarrow r < y$ . Also from  $n \cdot x < m$  we have  $x = n^{-1} \cdot n \cdot x < n^{-1} \cdot m = r \Rightarrow x < r$ . So we have  $x < r < y$ .
2. **(0 = x)** then  $0 < y$  and by 9.56 there exists a  $n \in \mathbb{N}_{\mathbb{R}} \setminus \{0\} \subseteq \mathbb{Q}_{\mathbb{R}} \Rightarrow 0 < n \underset{9.41}{\Rightarrow} 0 < n^{-1}$  such that for  $n^{-1} \in \mathbb{Q}_{\mathbb{R}}$  (as  $\mathbb{Q}_{\mathbb{R}}$  is a sub-field see 8.28) we have  $0 < n^{-1} < x$ . So taking  $r = n^{-1}$  we have  $0 = x < r < y$
3. **(x < 0)** We have now the following possibilities for  $y$ 
  - a. **(0 < y)** then if we take  $r = 0 \in \mathbb{Q}_{\mathbb{R}} \Rightarrow x < r < y$
  - b. **(y ≤ 0)** then  $x < y \leq 0 \Rightarrow 0 \leq -y < -x$  and we can use (1) or (2) to find a  $r' \in \mathbb{Q}_{\mathbb{R}}$  such that  $-y < r' < -x \Rightarrow x < -r' < y$  so using  $r = -r' \in \mathbb{Q}_{\mathbb{R}}$  we have  $x < r < y$

Next we prove the existence of a  $i \in \mathbb{R} \setminus \mathbb{Q}_{\mathbb{R}}$  such that  $x < i < y$ . As  $\mathbb{R} \setminus \mathbb{Q}_{\mathbb{R}} \neq \emptyset$  (see 8.51) there exists a  $z \in \mathbb{R} \setminus \mathbb{Q}_{\mathbb{R}}$  then from  $x < y$  we have  $x + z < y + z$  and by (1),(2) and (3) we have that there exists a  $r \in \mathbb{Q}_{\mathbb{R}}$  such that  $x + z < r < y + z \Rightarrow x < r - z < y$ . So if  $i = r - z$  then we have  $x < i < y$ . Now if  $i \in \mathbb{Q}_{\mathbb{R}}$  then as  $\mathbb{Q}_{\mathbb{R}}$  is a sub-field (see 8.28) and  $r \in \mathbb{Q}_{\mathbb{R}}$  we have  $z = -i + r \in \mathbb{Q}_{\mathbb{R}} \in \mathbb{R} \setminus \mathbb{Q}_{\mathbb{R}}$ .  $\square$

**Definition 8.65.** If  $x \in \mathbb{R}$  then  $|x|$  is defined by  $|x| = \begin{cases} x & \text{if } 0 \leq x \\ -x & \text{if } x < 0 \end{cases}$  and is called the absolute value

**Theorem 8.66.** The absolute value  $||$  satisfies the following

1.  $\forall x \in \mathbb{R}$  we have  $|x| \geq 0$
2.  $\forall x, \alpha \in \mathbb{R}$  we have  $|\alpha \cdot x| = |\alpha| \cdot |x|$
3.  $\forall x, y \in \mathbb{R}$  we have  $|x + y| \leq |x| + |y|$
4.  $\forall x \in \mathbb{R}$  we have  $|x| = 0 \Leftrightarrow x = 0$

**Proof.**

1. This is trivial based on the definition.
2. Given  $x, \alpha \in \mathbb{R}$  we have the following cases:
  - a. **(0 ≤ x, α)** then  $0 \leq x \cdot \alpha$  and  $|\alpha \cdot x| = \alpha \cdot x = |\alpha| \cdot |x|$
  - b. **(0 ≤ x, α < 0)** then  $\alpha \cdot x < 0$  and  $|\alpha \cdot x| = -(\alpha \cdot x) = (-\alpha) \cdot x = |\alpha| \cdot |x|$
  - c. **(x < 0, 0 ≤ α)** then  $\alpha \cdot x < 0$  and  $|\alpha \cdot x| = -(\alpha \cdot x) = \alpha \cdot (-x) = |\alpha| \cdot |x|$
  - d. **(x < 0, α < 0)** then  $0 < -(\alpha) \cdot (-x) = \alpha \cdot x$  and  $|\alpha \cdot x| = \alpha \cdot x = (-\alpha) \cdot (-x) = |\alpha| \cdot |x|$
3. Given  $x, y \in \mathbb{R}$  we have
  - a. **(0 ≤ x, y)** then  $0 \leq x + y$  and  $|x + y| = x + y = |x| + |y| \Rightarrow |x + y| \leq |x| + |y|$
  - b. **(0 ≤ x, y < 0)** then we have the following cases
    - i. **(0 ≤ x + y)** then  $|x + y| = x + y \underset{y < 0 \Rightarrow y < 0 < -y = |y|}{\Rightarrow} x + y < x + |y| = |x| + |y| \Rightarrow |x + y| \leq |x| + |y|$
    - ii. **(x + y < 0)** then  $|x + y| = -x - y \underset{0 \leq x \Rightarrow -x \leq 0 \leq x}{\Rightarrow} -x - y \leq x - y \underset{-y = |y|}{=} x + |y| = |x| + |y| \Rightarrow |x + y| \leq |x| + |y|$
  - c. **(x < 0, 0 ≤ y)** then if we take  $x' = y$  and  $y' = x$  we have  $|x + y| = |y + x| = |x' + y'| \leq |x'| + |y'|$  (case (b)) and thus  $|x + y| \leq |x| + |y|$

- d.  $(x < 0, y < 0)$  then  $x + y < 0 \Rightarrow |x + y| = -(x + y) = (-x) + (-y) = |x| + |y| \Rightarrow |x + y| \leq |x| + |y|$
4. If  $x = 0 \Rightarrow |x| = x = 0$ , if  $|x| = 0$  then if  $x < 0$  we have  $0 < -x = |x| = 0$  a contradiction so we must have  $0 \leq x$  but then  $x = |x| = 0 \Rightarrow x = 0$   $\square$

We prove now some useful theorems needed for the calculations of limits, first some definitions to simplify our notation.

**Definition 8.67.** Let  $n \in \mathbb{N}$  and  $x \in \mathbb{R}$  then  $n \cdot x = \begin{cases} 0 & \text{if } n=0 \\ x + (n-1) \cdot x & \text{if } n > 1 \end{cases}$ . Hence  $0 \cdot x = 0$  (last 0 is the neutral element in  $\mathbb{R}$ ),  $1 \cdot x = x + 0 \cdot x = x + 0 = x$ ,  $2 \cdot x = x + 1 \cdot x = x + x$ , ....

The above operation is consistent with the multiplication in  $\mathbb{R}$  as the following theorem proves.

**Theorem 8.68.** Let  $n \in \mathbb{N}$  then  $n \cdot x = i_{\mathbb{Q}_N}(n) \cdot x$  hence if 1 is the multiplicative unit in  $\mathbb{R}$  we have  $n \cdot 1 = i_{\mathbb{Q}_N}(n)$

**Proof.** By 8.56 we have that  $i_{\mathbb{Q}_N}(0) = 0 \in \mathbb{R}$  and  $i_{\mathbb{Q}_N}(1) = 1$ . We prove now the theorem by induction so let  $S = \{n \in \mathbb{N} \mid n \cdot x = i_{\mathbb{Q}_N}(n) \cdot x\}$  then we have

**0 ∈ S.** As  $0 \cdot x = 0 = i_{\mathbb{Q}_N}(0) \cdot x$

**n ∈ S = < n + 1 ∈ S.** we have that  $(n+1) \cdot x = x + ((n+1)-1) \cdot x = x + i_{\mathbb{Q}_N}(n) \cdot x = i_{\mathbb{Q}_N}(1) \cdot x + i_{\mathbb{Q}_N}(n) \cdot x = i_{\mathbb{Q}_N}(n+1) \cdot x$   $\square$

Motivated by the above theorem and the fact that we already use the same symbols for the neutral element and multiplicative unit in  $\mathbb{R}$  and  $\mathbb{N}$  we introduce the following notation.

**Notation 8.69.** If  $n \in \mathbb{N}$  then we note  $i_{\mathbb{Q}_N}(n)$  also by  $n$ . So if in a operation (sum, multiplication or comparison) we mix elements of  $\mathbb{N}$  and elements of  $\mathbb{R}$  we implicitly assume that we have applied  $i_{\mathbb{Q}_N}$  to the elements of  $\mathbb{N}$ .

**Theorem 8.70.**  $\forall n \in \mathbb{N}$  we have that  $n < 2^n$  [and thus by 8.62 we have also  $\forall a \in \mathbb{R} \exists n \in \mathbb{N}$  such that  $a < n < 2^n$ ]

**Proof.** This is proved by induction so let  $S = \{n \in \mathbb{N} \mid n < 2^n\}$ . Then we have that

1. If  $n = 0$  then  $0 < 1 = 2^0$  so that  $0 \in S$
2. If  $n \in S$  then  $i_{\mathbb{Q}_N}(n) < 2^n$  then consider the following cases for  $n+1$  (as  $0 < n+1$ )
  - a.  $(n+1 = 1)$  then  $n+1 = 1 < 2 = 2^1 = 2^{n+1} \Rightarrow n+1 \in S$
  - b.  $(1 < n+1)$  then  $2 \leq n+1 \Rightarrow 1 \leq n$  so that  $n+1 \leq n+n = 2 \cdot n < 2 \cdot 2^n = 2^{n+1}$  proving that  $n+1 \in S$   $\square$

**Theorem 8.71.** Let  $x \in \mathbb{R}$  with  $x > 1$  then  $x^n - 1 \geq n \cdot (x - 1) \forall n \in \mathbb{N}$

**Proof.** We prove this by induction, so let  $S = \{n \in \mathbb{N} \mid x^n - 1 \geq n \cdot (x - 1)\}$  then we have

1. if  $n = 0$  then  $x^n - 1 = x^0 - 1 = 1 - 1 = 0 \geq 0 = 0 \cdot (x - 1) = n \cdot (x - 1) \Rightarrow 0 \in S$
2. If  $n \in S$  then  $x^{n+1} - 1 = x \cdot x^{n+1} - 1 = x \cdot (x^n - 1) + (x - 1) \geq x \cdot (n \cdot (x - 1)) + (x - 1) = (x - 1) \cdot (x \cdot n + 1) >_{x > 1} n \cdot x \geq n \cdot (x - 1) \cdot (n + 1) \Rightarrow n + 1 \in S$   $\square$

**Theorem 8.72.** If  $N \in \mathbb{N}_{\mathbb{R}} \setminus \{0\}$  and  $x \in \mathbb{R}$  with  $x > 1$  then there exists a  $n \in \mathbb{N}_{\mathbb{R}} \setminus \{0\}$  such that  $x^n > N$

**Proof.** Consider the following cases for  $N \in \mathbb{N}_{\mathbb{R}} \setminus \{0\}$

1. ( $N = 1$ ) then  $N = 1 < x = x^1$  proving the theorem
2. ( $N > 1$ ) take  $\delta = \frac{N-1}{x-1} > 0$  then by the Archimedean property (8.62) there exists a  $n \in \mathbb{N}_{\mathbb{R}}$  such that  $\delta < n$  as by definition of  $\delta$  we have  $N - 1 = \delta \cdot (x - 1) \leq n \cdot (x - 1) < x^n - 1$  (using the previous theorem) so that  $N < x^n$   $\square$

**Theorem 8.73.** If  $x \in \mathbb{R}$  with  $0 < x < 1$  and  $n, m \in \mathbb{N}_{\mathbb{R}}$  with  $n < m$  then  $x^m < x^n$

**Proof.** We prove this by induction, so let  $m \in \mathbb{N}_{\mathbb{R}}$  and take  $S_m = \{n \in \mathbb{N} \mid x^{m+n} < x^m\}$  then we have

1. if  $n = 1$  then  $x^{m+n} = x^{n+1} = x \cdot (x^n) <_{x < 1 \Rightarrow x \cdot x^n < x^n} x^n \Rightarrow 1 \in S_m$
2. if  $n \in S_m$  then  $x^{m+(n+1)} = x \cdot x^{m+n} <_{x < 1} x^{m+n} <_{n \in S_m} x^m$

Using mathematical induction we have then that  $S_m = \mathbb{N}$ . If now  $n, m \in \mathbb{N}_{\mathbb{R}}$  with  $n < m$  then  $k = m - n > 0$  or  $k \in \mathbb{N} = S_n$  so that  $x^m = x^{n+k} < x^n$   $\square$

**Theorem 8.74.** If  $\varepsilon \in \mathbb{R}_+$  and  $x \in \mathbb{R}$  such that  $0 < x < 1$  then  $\exists N \in \mathbb{N}_{\mathbb{R}} \setminus \{0\}$  such that  $0 < x^n < \varepsilon$  if  $n \geq N$

**Proof.** As  $\varepsilon > 0$  we have that  $\frac{1}{\varepsilon}$  is defined and by the Archimedean property (see 8.62) there exists a  $n \in \mathbb{N}_{\mathbb{R}}$  such that  $0 < \frac{1}{\varepsilon} < n$ , as  $0 < x < 1$  we have that  $1 < \frac{1}{x}$  so that by 8.72 there exists a  $N \in \mathbb{N}_{\mathbb{R}} \setminus \{0\}$  such that  $n < x^N$  and thus we have that  $0 < \frac{1}{\varepsilon} < \left(\frac{1}{x}\right)^N = \frac{1}{x^N}$   $\xrightarrow{x, \varepsilon > 0} 0 < x^N < \varepsilon$ . If now we have  $n \geq N$  then by the previous theorem we have  $0 < x^n \leq x^N < \varepsilon$   $\square$

## 8.4 Square root

**Theorem 8.75.** The function  ${}^2: \{x \in \mathbb{R} \mid 0 \leq x\} \rightarrow \{x \in \mathbb{R} \mid 0 \leq x\}$  defined by  $x \mapsto x^2$  (where  $0 \leq x^2$  by 9.41 so this is indeed a function) is a bijection.

**Proof.**

1. **(injectivity)** If  $x, y \geq 0$  is such that  $x^2 = y^2$  then either
  - a. **( $x = 0$ )** then  $0 = x^2 = y^2 \Rightarrow y^2 = 0 \xrightarrow{9.41} y = 0 \Rightarrow x = y$
  - b. **( $y = 0$ )** then  $0 = y^2 = x^2 \Rightarrow x = 0 \Rightarrow x = y$
  - c. **( $0 < x, y$ )** then if we assume that  $x \neq y$  then we have the following cases
    - i. **( $x < y$ )** then  $x^2 < y \cdot x \xrightarrow{x < y \Rightarrow y \cdot x < y^2} x^2 < y^2$  contradicting  $x^2 = y^2$
    - ii. **( $y < x$ )** then similar we have  $y^2 < x^2$  contradicting  $x^2 = y^2$

In all the cases we have a contradiction so we must have  $x = y$ .

2. **(surjectivity)** Given  $y \in \{x \in \mathbb{R} \mid 0 \leq x\}$  then  $0 \leq y$  and we have the following cases :
  - a. **( $y = 0$ )** then  $y = 0 = 0 \cdot 0 = 0^2 \Rightarrow 0^2 = y$
  - b. **( $y = 1$ )** then  $1^2 = 1 \cdot 1 = 1 = y \Rightarrow 1^2 = y$
  - c. **( $0 < y \neq 1$ )** take then  $S_y = \{t \in \mathbb{R} \mid 0 \leq t \wedge t^2 \leq y\}$ . Then as  $0^2 = 0 < y$  we have  $0 \in S_y$  and thus  $0 \notin S_y$ . Now as  $y \neq 1$  we have the following possibilities for  $y$ 
    - i. **( $y < 1$ )** Take now  $t \in S_y$  and suppose  $1 < t$ . Then from  $1 < t$  and  $0 < 1 < t \Rightarrow 0 < t$  and 9.41 we have  $t \leq t^2$ . From  $y < 1$  and  $1 < t$  we have then  $y < t$  and using  $t \leq t^2$  we have then  $y < t^2$  which contradicts  $t \in S_y \Rightarrow t^2 \leq y$ . So we must have  $t \leq 1$  and thus we have that 1 is a upper bound of  $S_t$ .
    - ii. **( $1 < y$ )** Take now  $t \in S_y$  and suppose  $y < t$ . Then from  $1 < y$  and  $y < t$  we have  $1 < t$ . Using  $1 < t \Rightarrow 0 < t$  and 9.41 on  $1 < t$  we have  $1 \cdot t < t \cdot t \Rightarrow t < t^2 \xrightarrow{y < t} y < t^2$  contradicting  $t \in S_y \Rightarrow t^2 \leq y$ . So we must have  $t \leq y$  and thus we have that  $y$  is a upper bound of  $S_y$ .

Using (i) and (ii) above we see that  $S_y$  has a upper bound and is not empty. From the conditional completeness on the reals (see 9.43) we have that there exists a lowest upper bound  $s_y = \sup(S_y)$ . For  $S_y$  we consider now again the two possible cases for  $y \neq 1$

- i. **( $1 < y$ )** then  $1^2 = 1 < y \Rightarrow 1 \in S_y$  and thus  $S_y$  contains a  $t$  with  $0 < t$

ii. ( $y < 1$ ) then from  $0 < y$  and 9.41 we have  $y \cdot y < 1 \cdot y \Rightarrow y^2 < y \Rightarrow y \in S_y$  and thus  $S_y$  contains a  $t$  with  $0 < t$ .

From the above (i) and (ii) we have, as  $\forall t \in S_y$  that  $t \leq s_y = \sup(S_y)$ , that

$$0 < s_y. \quad (8.5)$$

Now from  $0 < s_y$  there exists by 9.56 a  $\varepsilon_0 \in \mathbb{R}$  such that  $0 < \varepsilon_0 < s_y$ . Now for all  $\varepsilon \in \mathbb{R}$  with  $0 < \varepsilon < s_y$  we have  $0 < s_y - \varepsilon < s_y < s_y + \varepsilon$ , then using 9.41 we have  $(s_y - \varepsilon)^2 < s_y \cdot (s_y - \varepsilon) \wedge (s_y - \varepsilon) \cdot s_y < s_y^2$  and  $s_y^2 < s_y \cdot (s_y + \varepsilon) \wedge s_y \cdot (s_y + \varepsilon) < (s_y + \varepsilon)^2$  giving finally  $(s_y - \varepsilon)^2 < s_y^2 < (s_y + \varepsilon)^2$ . As  $s_y$  is an upper bound of  $S_y$  and  $s_y < s_y + \varepsilon$  we must have  $s_y + \varepsilon \notin S_y$  and as  $0 < s_y < s_y + \varepsilon$  we must have that  $y < (s_y + \varepsilon)^2$ . Using the fact that  $s_y$  is the lowest bound on  $S_y$  and the fact that  $s_y - \varepsilon < s_y$  there exists a  $f \in S_y$  such that  $s_y - \varepsilon < f \leq s_y$ . Using 9.41 we have then that  $f \cdot (s_y - \varepsilon) < f^2 \wedge (s_y - \varepsilon)^2 < f \cdot (s_y - \varepsilon) \Rightarrow (s_y - \varepsilon)^2 < f^2$  and as  $f \in S_y$  we have  $f^2 \leq y$  giving us  $(s_y - \varepsilon)^2 < f^2 \leq y < (s_y + \varepsilon)^2$ . To summarize we have proved that if  $0 < \varepsilon_0 < s_y$  that

$$(s_y - \varepsilon)^2 < s_y^2 < (s_y + \varepsilon)^2 \quad (8.6)$$

$$(s_y - \varepsilon)^2 < y < (s_y + \varepsilon)^2 \quad (8.7)$$

Using 9.41 on 8.7 gives

$$-(s_y + \varepsilon)^2 < -y < -(s_y - \varepsilon)^2 \quad (8.8)$$

and adding 8.6 and 8.8 gives

$$(s_y - \varepsilon)^2 - (s_y + \varepsilon)^2 < s_y^2 - y < (s_y + \varepsilon)^2 - (s_y - \varepsilon)^2 \quad (8.9)$$

As  $(s_y + \varepsilon)^2 - (s_y - \varepsilon)^2 = s_y^2 + 2 \cdot s_y \cdot \varepsilon + \varepsilon^2 - s_y^2 + 2 \cdot s_y \cdot \varepsilon - \varepsilon^2 = 4 \cdot \varepsilon \cdot s_y$  and using 8.9 we have then

$$-4 \cdot \varepsilon \cdot s_y < s_y^2 - y < 4 \cdot \varepsilon \cdot s_y \text{ or } -4 \cdot \varepsilon \cdot s_y < y - s_y^2 < 4 \cdot \varepsilon \cdot s_y \quad (8.10)$$

For  $s_y^2 - y$  we have now the following possible cases to consider

i. ( $s_y^2 - y < 0$ ) Take then  $\delta = y - s_y^2$  then  $0 < \delta$ . Take now  $\varepsilon = \min\left(\frac{\delta}{4 \cdot s_y}, \varepsilon_0\right)$  (which is well defined as  $0 < s_y$  see 8.5) then we have  $0 < \varepsilon \leq \varepsilon_0 < s_y$  so we have by 8.10 that  $\delta = y - s_y^2 < 4 \cdot \varepsilon \cdot s_y$ . As by definition of  $\varepsilon$  we have  $\varepsilon \leq \frac{\delta}{4 \cdot s_y} \Rightarrow 4 \cdot \varepsilon \cdot s_y \leq \delta < 4 \cdot \varepsilon \cdot s_y$  and we reach a contradiction.

ii. ( $0 < s_y^2 - y$ ) Take then  $\delta = s_y^2 - y$  then we have (see 8.10)  $0 < \delta$ . Take now  $\varepsilon = \min\left(\frac{\delta}{4 \cdot s_y}, \varepsilon_0\right)$  then we have  $0 < \varepsilon \leq \varepsilon_0 < s_y$ , so we have by 8.10 that  $\delta = s_y^2 - y < 4 \cdot \varepsilon \cdot s_y$ . As by definition of  $\varepsilon$  we have  $\varepsilon \leq \frac{\delta}{4 \cdot s_y} \Rightarrow 4 \cdot \varepsilon \cdot s_y \leq \delta < 4 \cdot \varepsilon \cdot s_y$  and we reach a contradiction.

iii. ( $s_y^2 - y = 0$ ) Then  $s_y^2 = y$

So (iii) is the only possible case [as (i) and (ii) leads to a contradiction] we must have  $s_y^2 = y$ .

So in cases (a), (b) and (c) we have found a  $t$  such that  $t^2 = y$  proving surjectivity.  $\square$

**Definition 8.76.** Using the above theorem we have that the function  ${}^2: \{x \in \mathbb{R} | 0 \leq x\} \rightarrow \{x \in \mathbb{R} | 0 \leq x\}$  has a inverse mapping. This mapping is called the square root function and is noted by  $\sqrt{\cdot}$ . So  $\sqrt{\cdot}: \{x \in \mathbb{R} | 0 \leq x\} \rightarrow \{x \in \mathbb{R} | 0 \leq x\}$  is mapping such that  $\sqrt{\cdot} \circ {}^2 = i_{\{x \in \mathbb{R} | 0 \leq x\}} = {}^2 \circ \sqrt{\cdot}$  or  $\forall x \in \{x \in \mathbb{R} | 0 \leq x\}$  we have  $(\sqrt{x})^2 = x$  and  $\sqrt{x^2} = x$

**Note 8.77.** The requirement that  $0 \leq x$  to have  $\sqrt{x^2} = x = \sqrt{x^2}$  is really necessary as  ${}^2: \mathbb{R} \rightarrow \{x \in \mathbb{R} | 0 \leq x\}$  is definitely not a injection and thus a bijection as  $(-1)^2 = 1 = 1^2$

**Theorem 8.78.**  $\sqrt{\cdot}: \{x \in \mathbb{R} | 0 \leq x\} \rightarrow \{x \in \mathbb{R} | 0 \leq x\}$  is a strictly increasing function

**Proof.** If  $x, y \in \{x \in \mathbb{R} | 0 \leq x\}$  is such that  $x < y$ . Assume then that  $\sqrt{y} \leq \sqrt{x}$  then by 8.39 we have  $(\sqrt{y})^2 \leq (\sqrt{x})^2 \Rightarrow y \leq x$  contradicting  $x < y$  so we must have  $\sqrt{x} < \sqrt{y}$   $\square$

**Theorem 8.79.** If  $x \in \mathbb{R}$  and  $0 \leq a \in \mathbb{R}$  then we have

1.  $x^2 = a$  if and only if  $x = \sqrt{a}$  or  $x = -\sqrt{a}$
2.  $x^2 \leq a$  if and only if  $-\sqrt{a} \leq x \leq \sqrt{a}$
3.  $x^2 < a$  if and only if  $-\sqrt{a} < x < \sqrt{a}$

**Proof.**

1.

$\Rightarrow$ . For  $x$  we have either

$0 \leq x$ . then using 8.76 we have that  $x = (\sqrt{\cdot}) \circ (x^2)(x) = (\sqrt{\cdot})(x^2) = (\sqrt{\cdot})(a) = \sqrt{(a)}$ .

$x < 0$ . then  $0 < -x$  so that using the case  $(0 < x)$  we have that  $-x = \sqrt{a}$  or  $x = -\sqrt{a}$

$\Leftarrow$ . If  $x = \sqrt{a}$  then  $x^2 = (\sqrt{a})^2 = a$ , if  $x = -\sqrt{a}$  then  $x^2 = (-\sqrt{a})^2 = (\sqrt{a})^2 = a$ .

2.

$\Rightarrow$ . For  $x$  we have either

$0 \leq x$ . Then as  $0 \leq \sqrt{a}$  we have that  $-\sqrt{a} \leq 0 \leq x$ . Further assume that  $\sqrt{a} < x$  then, as  $(x^2)$  is strictly increasing [see 8.39], we have that  $a = (\sqrt{a})^2 < x^2$  contradicting  $x^2 \leq a$ , hence we must have that  $x \leq \sqrt{a}$ . So we have  $-\sqrt{a} \leq x \leq \sqrt{a}$

$x < 0$ . then  $0 < -x$  and we can apply the case  $(0 \leq x)$  to find that  $-\sqrt{a} \leq -x \leq \sqrt{a}$ , multiplying this by  $(-1)$  gives  $-\sqrt{a} \leq x \leq \sqrt{a}$

$\Leftarrow$ . Consider the following cases for  $x$

$0 \leq x$ . then from  $x \leq \sqrt{a}$  we have, as  $(x^2)$  is increasing [see 9.41], that  $x^2 \leq (\sqrt{a})^2 = a$ .

$x < 0$ . then from  $-\sqrt{a} \leq x$  we have  $0 < -x < \sqrt{a}$ . As  $(x^2)$  is increasing [see 8.39] it follows that  $x^2 = (-x)^2 < (\sqrt{a})^2 = a$

3.

$\Rightarrow$ . Using (2) we have  $-\sqrt{a} \leq x \leq \sqrt{a}$ . If  $x = \sqrt{a}$  then  $x^2 = (\sqrt{a})^2 = a$  contradicting  $x^2 < a$ , if  $x = -\sqrt{a}$  then  $(-x) = \sqrt{a}$  so that  $x^2 = (-\sqrt{a})^2 = (\sqrt{a})^2 = a$  contradicting  $x^2 < a$ . So we must have that  $-\sqrt{a} < x < \sqrt{a}$ .

$\Leftarrow$ . Using (2) we have that  $x^2 \leq a$ . If  $x^2 = a$  then by (1)  $x$  would be either  $\sqrt{a}$  or  $-\sqrt{a}$ , where both contradicts  $-\sqrt{a} < x < \sqrt{a}$  so we must have that  $x^2 < a$ .

$\square$

**Theorem 8.80.** If  $x, y \in \{x \in \mathbb{R} | 0 \leq x\}$  then  $\sqrt{x \cdot y} = \sqrt{x} \cdot \sqrt{y}$

**Proof.** As  $(\sqrt{x \cdot y})^2 = x \cdot y = (\sqrt{x})^2 \cdot (\sqrt{y})^2 = (\sqrt{x} \cdot \sqrt{y})^2$  we have by the fact that  $(x^2)$  is a bijection and thus injective so we have  $\sqrt{x \cdot y} = \sqrt{x} \cdot \sqrt{y}$   $\square$

**Theorem 8.81.** Given  $x, y \in \{x \in \mathbb{R} | 0 \leq x\}$  then we have  $\sqrt{x+y} \leq \sqrt{x} + \sqrt{y}$

**Proof.** We prove this by contradiction, so assume that  $\sqrt{x} + \sqrt{y} < \sqrt{x+y}$   $\Rightarrow -\sqrt{x+y} < -(\sqrt{x} + \sqrt{y})$  and by multiplying by  $\sqrt{x} + \sqrt{y} \geq 0$  we have  $-\sqrt{x+y} \cdot (\sqrt{x} + \sqrt{y}) < -(\sqrt{x} + \sqrt{y}) \cdot (\sqrt{x} + \sqrt{y}) = -(x+y+2\sqrt{x+y})$ . Now we have by 8.39 that  $0 \leq (\sqrt{x+y} - (\sqrt{x} + \sqrt{y}))^2 = (\sqrt{x+y})^2 + (\sqrt{x} + \sqrt{y})^2 - 2 \cdot \sqrt{x+y} \cdot (\sqrt{x} + \sqrt{y}) = x+y + (\sqrt{x})^2 + (\sqrt{y})^2 + 2\sqrt{x} \cdot \sqrt{y} - 2 \cdot \sqrt{x+y} \cdot (\sqrt{x} + \sqrt{y}) = 2 \cdot x + 2 \cdot y + 2\sqrt{x} \cdot \sqrt{y} - 2 \cdot \sqrt{x+y} \cdot (\sqrt{x} + \sqrt{y})$   $\Rightarrow$  multiply by  $0 < \frac{1}{2}$   $0 \leq x+y + \sqrt{x} \cdot \sqrt{y} - \sqrt{x+y} \cdot (\sqrt{x} + \sqrt{y}) < x+y + \sqrt{x} \cdot \sqrt{y} - (\sqrt{x} + \sqrt{y}) \cdot (\sqrt{x} + \sqrt{y}) = x+y + \sqrt{x} \cdot \sqrt{y} - (x+y+2\sqrt{x+y}) = -\sqrt{x+y} \leq 0 \Rightarrow 0 < 0$  a contradiction. So we must have  $\sqrt{x+y} \leq \sqrt{x} + \sqrt{y}$   $\square$



# Chapter 9

## The complex numbers

### 9.1 Definition and arithmetic's

**Definition 9.1.** The space  $\langle \mathbb{C}, +, \cdot \rangle$  of complex numbers is defined by: (note that  $+, \cdot$  has different meanings in  $\mathbb{C}$  and  $\mathbb{R}$ )

1.  $\mathbb{C} = \mathbb{R} \times \mathbb{R}$
2.  $+ : \mathbb{C} \times \mathbb{C} \rightarrow \mathbb{C}$  by  $((x, y), (x', y')) \rightarrow (x, y) + (x', y') \stackrel{\text{defined}}{=} (x + x', y + y')$
3.  $\cdot : \mathbb{C} \times \mathbb{C} \rightarrow \mathbb{C}$  by  $((x, y), (x', y')) \rightarrow (x, y) \cdot (x', y') \stackrel{\text{defined}}{=} (x \cdot x' - y \cdot y', x \cdot y' + x' \cdot y)$

**Theorem 9.2.**  $\langle \mathbb{C}, +, \cdot \rangle$  forms a field. The unit element is  $(1, 0)$  noted also as  $1$ , the neutral element is  $(0, 0)$  noted as  $0$ ,  $-(x, y) = (-x, -y)$  and if  $(x, y) \neq (0, 0)$  then  $(x, y)^{-1} = \left( \frac{x^2 \cdot y}{x^3 \cdot y + x \cdot y^3}, \frac{-y^2 \cdot x}{x^3 \cdot y + x \cdot y^3} \right)$  if  $x \neq 0 \wedge y \neq 0$ ,  $(x, y)^{-1} = (x^{-1}, 0)$  if  $x \neq 0, y = 0$  and  $(x, y)^{-1} = (0, -y^{-1})$  if  $x = 0, y \neq 0$

**Proof.** First we prove that  $\langle \mathbb{C}, + \rangle$  forms a abelian group using the fact that  $\langle \mathbb{R}, +, \cdot \rangle$  is a field.

1. **(associativity)**  $\forall (x, y), (x', y'), (x'', y'') \in \mathbb{C}$  we have  $(x, y) + ((x', y') + (x'', y'')) = (x, y) + (x' + x'', y' + y'') = (x + (x' + x''), y + (y' + y'')) = ((x + x') + x'', (y + y') + y'') = (x + x', y + y') + (x'', y'') = ((x, y) + (x', y')) + (x'', y'')$
2. **(neutral element)**  $\forall (x, y) \in \mathbb{C}$  we have  $(x, y) + (0, 0) = (x + 0, y + 0) = (x, y) = (x + 0, y + 0) = (x, y) + (0, 0)$
3. **(inverse element)**  $\forall (x, y) \in \mathbb{C}$  we have  $(x, y) + (-x, -y) = (x + (-x), y + (-y)) = (0, 0) = (-x + x, -y + y) = (-x, -y) + (x, y)$ . So  $(-x, -y)$  is the additive inverse of  $(x, y)$ .
4. **(commutativity)**  $\forall (x, y), (x', y') \in \mathbb{C}$  we have  $(x, y) + (x', y') = (x + x', y + y') = (x' + x, y' + y) = (x', y') + (x, y)$

Next we prove the rest of the axioms of a field

1. **(distributive)**  $\forall (x, y), (x', y'), (x'', y'')$  we have  $(x, y) \cdot ((x', y') + (x'', y'')) = (x, y) \cdot (x' + x'', y' + y'') = (x \cdot (x' + x'') - y \cdot (y' + y''), x \cdot (y' + y'') + y \cdot (x' + x'')) = (x \cdot x' + x \cdot x'' - y \cdot y' - y \cdot y'', x \cdot y' + x \cdot y'' + y \cdot x' + y \cdot x'') = (x \cdot x' - y \cdot y', x \cdot y' + y \cdot x') + (x \cdot x'' - y \cdot y'', x \cdot y'' + y \cdot x'') = (x, y) \cdot (x', y') + (x, y) \cdot (x'', y'')$
2. **(neutral element)**  $\forall (x, y) \in \mathbb{C}$  we have  $(1, 0) \cdot (x, y) = (1 \cdot x - 0 \cdot y, 1 \cdot y + 0 \cdot x) = (x, y)$  and  $(x, y) \cdot (1, 0) = (x \cdot 1 + y \cdot 0, x \cdot 0 + y \cdot 1) = (x, y)$
3. **(commutative)**  $\forall (x, y), (x', y') \in \mathbb{C}$  we have  $(x, y) \cdot (x', y') = (x \cdot x' - y \cdot y', x \cdot y' + y \cdot x') = (x' \cdot x - y' \cdot y, y' \cdot x + x' \cdot y) = (x' \cdot x - y' \cdot y, x' \cdot y + y' \cdot x) = (x', y') \cdot (x, y)$
4. **(associative)**  $\forall (x, y), (x', y'), (x'', y'')$  then we have  $(x, y) \cdot ((x', y') \cdot (x'', y'')) = (x, y) \cdot (x' \cdot x'' - y' \cdot y'', x' \cdot y'' + y' \cdot x'') = (x \cdot (x' \cdot x'' - y' \cdot y'') - y \cdot (x' \cdot y'' + y' \cdot x''), x \cdot (x' \cdot y'' + y' \cdot x'') + y \cdot (x' \cdot x'' - y' \cdot y'')) = (x \cdot x' \cdot x'' - x \cdot y' \cdot y'' - y \cdot x' \cdot y'' - y \cdot y' \cdot x'', x \cdot x' \cdot y'' + x \cdot y' \cdot x'' + y \cdot x' \cdot x'' - y \cdot y' \cdot y'') = ((x \cdot x' - y \cdot y') \cdot x'' - (x \cdot y' + y \cdot x') \cdot y'', (x \cdot x' - y \cdot y') \cdot y'' + (x \cdot y' + y \cdot x') \cdot x'') = (x \cdot x' - y \cdot y', x \cdot y' + y \cdot x') \cdot (x'', y'') = ((x, y) \cdot (x', y')) \cdot (x'', y'')$

5. (**multiplicative inverse**) Given  $(x, y) \in \mathbb{C}$  with  $(x, y) \neq (0, 0)$ , then we have 3 cases to consider

**$x \neq 0 \wedge y \neq 0$ .** Now  $x^3 \cdot y + x \cdot y^3 = x \cdot (x^2 \cdot y + y^3) = x \cdot y \cdot (x^2 + y^2)$  now using 9.41 we have  $0 < x^2, y^2 \Rightarrow 0 < (x^2 + y^2) \underset{x, y \neq 0}{\Rightarrow} x \cdot y \cdot (x^2 + y^2) \Rightarrow x^3 \cdot y + x \cdot y^3 \neq 0$ . So  $\left( \frac{x^2 \cdot y}{x^3 \cdot y + x \cdot y^3}, \frac{-y^2 \cdot x}{x^3 \cdot y + x \cdot y^3} \right)$  is defined and  $(x, y) \cdot \left( \frac{x^2 \cdot y}{x^3 \cdot y + x \cdot y^3}, \frac{-y^2 \cdot x}{x^3 \cdot y + x \cdot y^3} \right) = \left( \frac{x \cdot x^2 \cdot y + y \cdot y^2 \cdot x}{x^3 \cdot y + x \cdot y^3}, \frac{x \cdot -y^2 \cdot x + y \cdot x^2 \cdot y}{x^3 \cdot y + x \cdot y^3} \right) = \left( \frac{x^3 \cdot y + x \cdot y^3}{x^3 \cdot y + x \cdot y^3}, \frac{-x^2 \cdot y^2 + x^2 \cdot y^2}{x^3 \cdot y + x \cdot y^3} \right) = (1, 0) \underset{(3)}{=} \left( \frac{x^2 \cdot y}{x^3 \cdot y + x \cdot y^3}, \frac{-y^2 \cdot x}{x^3 \cdot y + x \cdot y^3} \right)$ .

**$x \neq 0 \wedge y = 0$ .** then  $(x, y) \cdot (x^{-1}, 0) = (x, 0) \cdot (x^{-1}, 0) = (x \cdot x^{-1} - 0 \cdot 0, x \cdot 0 + 0 \cdot x^{-1}) = (x \cdot x^{-1}, 0) = 1$

**$x = 0 \wedge y \neq 0$ .** then  $(x, y) \cdot (0, y^{-1}) = (0, y) \cdot (0, y^{-1}) = (0 \cdot 0 - y \cdot (-y^{-1}), 0 \cdot y^{-1} + y \cdot 0) = (1, 0)$   $\square$

**Notation 9.3.** If  $x, y \in \mathbb{C}$  with  $y \neq 0$  then we note  $x \cdot y^{-1}$  as  $\frac{x}{y}$  hence as  $x^{-1} = 1 \cdot x^{-1} = \frac{1}{y}$  we have also that  $y^{-1} = \frac{1}{y}$

**Definition 9.4.** Given  $(x, y) \in \mathbb{C}$  we define  $\overline{(x, y)} = (x, -y)$

**Lemma 9.5.** If  $z, z' \in \mathbb{C}$  then  $\overline{z + z'} = \bar{z} + \bar{z}'$ ,  $\overline{z \cdot z'} = \bar{z} \cdot \bar{z}'$  and  $\bar{\bar{z}} = z$

**Proof.** If  $z = (x, y), z' = (x', y')$  then

1.  $\overline{(x, y) + (x', y')} = \overline{(x + x', y + y')} = (x + x', -y - y') = (x, -y) + (x', -y') = \overline{(x, y)} + \overline{(x', y')}$
2.  $\overline{(x, y) \cdot (x', y')} = \overline{(x \cdot x' - y \cdot y', x \cdot y' + y \cdot x')} = (x \cdot x' - y \cdot y', -x \cdot y' - y \cdot x') = (x \cdot x' - (-y) \cdot (-y'), x \cdot (-y') + (-y) \cdot x') = (x, -y) \cdot (x', -y') = \overline{(x, y)} \cdot \overline{(x', y')}$
3.  $\bar{\bar{z}} = \overline{\overline{(x, y)}} = \overline{(x, -y)} = (x, y) = z$   $\square$

**Definition 9.6.** As  $\langle \mathbb{C}, \cdot \rangle$  is a abelian semi-group we have by 4.22 that given a  $a \in \mathbb{C}$  and  $n \in \mathbb{N}$  that there exists a  $a^n$  such that

$$\begin{aligned} a^0 &= 1 \\ a^{n+1} &= a^n \cdot a \underset{\text{abelian}}{=} a \cdot a^n \end{aligned}$$

**Theorem 9.7.** If  $n, n' \in \mathbb{N}$  and  $a \in \mathbb{C}$  then  $a^{n'+n} = a^{n'} \cdot a^n$

**Proof.** We prove this by induction on  $n$ . So let  $X = \{n \in \mathbb{N} | a^{n'+n} = a^{n'} \cdot a^n\}$  then we have

1. If  $n = 0$  then  $a^{n'+n} = a^{n'+0} = a^{n'} = a^{n'} \cdot 1 = a^{n'} \cdot a^0 \Rightarrow 0 \in X$
2. If  $n \in X$  then  $a^{n'+(n+1)} = a^{(n'+n)+1} = a^{(n'+n)} \cdot a \underset{n \in X}{=} (a^{n'} \cdot a^n) \cdot a = a^{n'} \cdot (a^n \cdot a) = a^{n'} \cdot a^{n+1}$  and thus  $n+1 \in X$

Using mathematical induction (see 4.10) we have  $X = \mathbb{N}$  proving the theorem  $\square$

**Theorem 9.8.** If  $n \in \mathbb{N}$  and  $a, b \in \mathbb{C}$  then we have  $(a \cdot b)^n = a^n \cdot b^n$

**Proof.** We prove this by induction so take  $\mathcal{S} = \{n \in \mathbb{N} | (a \cdot b)^n = a^n \cdot b^n\}$  then we have

**$n = 0$ .** then  $(a \cdot b)^0 = 1 = 1 \cdot 1 = a^0 \cdot b^0$

**$n \in \mathcal{S} \Rightarrow n + 1 \in \mathcal{S}$ .** then  $(a \cdot b)^{n+1} = (a \cdot b)^n \cdot (a \cdot b) = (a^n \cdot b^n) \cdot (a^n \cdot a) \cdot (b^n \cdot b) = a^{n+1} \cdot b^{n+1}$   $\square$

**Corollary 9.9.** If  $z \in \mathbb{C}$  then  $\forall n \in \mathbb{N}$  we have that  $\overline{z^n} = \bar{z}^n$

**Proof.** This is easily proved by induction so let  $\mathcal{S} = \{n \in \mathbb{N} | \overline{z^n} = \bar{z}^n\}$  then we have

**$1 \in \mathcal{S}$ .** from  $\bar{z}^1 = \bar{z} = \bar{z}^1$  it follows that  $1 \in \mathcal{S}$

$n \in \mathcal{S} \Rightarrow n + 1 \in \mathcal{S}$ . we have

$$\begin{aligned}\overline{z^{n+1}} &= \overline{z^n \cdot z} \\ &\stackrel{9.5}{=} \overline{z^n} \cdot \bar{z} \\ &\stackrel{n \in \mathcal{S}}{=} \bar{z}^n \cdot \bar{z} \\ &= \bar{z}^n\end{aligned}$$

proving that  $n + 1 \in \mathcal{S}$

□

**Definition 9.10.**  $\mathbb{R}_{\mathbb{C}} = \{(x, y) \in \mathbb{C} | y = 0\} \subseteq \mathbb{C}$

**Theorem 9.11.** If  $\alpha = (\alpha', 0) \in \mathbb{R}_{\mathbb{C}}$  and  $z = (x, y) \in \mathbb{C}$  then  $\alpha \cdot z = (\alpha' \cdot x, \alpha' \cdot y)$

**Proof.**  $\alpha \cdot z = (\alpha', 0) \cdot (x, y) = (\alpha' \cdot x - 0 \cdot y, \alpha' \cdot y + 0 \cdot x) = (\alpha' \cdot x, \alpha' \cdot y)$

□

**Theorem 9.12.** If  $x \in \mathbb{C}$  so that  $x = \bar{x}$  then  $x \in \mathbb{R}_{\mathbb{C}}$

**Proof.** Let  $x = (a, b)$  then if  $x = \bar{x}$  we have  $(a, b) = (a, -b) \Rightarrow b = -b \Rightarrow 2 \cdot b = 0 \Rightarrow b = 0 \Rightarrow x = (a, 0) \in \mathbb{R}_{\mathbb{C}}$

□

**Theorem 9.13.**  $\langle \mathbb{R}_{\mathbb{C}}, +, \cdot \rangle$  forms a sub field (see 3.31)

**Proof.** We have to prove the following (using the definition of a sub field)

$\forall x, y \in \mathbb{R}_{\mathbb{C}} \models x + y \in \mathbb{R}_{\mathbb{C}}$ . as  $x, y \in \mathbb{R}_{\mathbb{C}}$  there exists  $x', y' \in \mathbb{R}$  so that  $x = (x', 0) \wedge y = (y', 0)$  so that  $x + y = (x', 0) + (y', 0) = (x' + y', 0 + 0) = (x + x', 0) \in \mathbb{R}$

$\forall x, y \in \mathbb{R}_{\mathbb{C}} \models x \cdot y \in \mathbb{R}_{\mathbb{C}}$ . as  $x, y \in \mathbb{R}_{\mathbb{C}}$  there exists  $x', y' \in \mathbb{R}$  so that  $x = (x', 0) \wedge y = (y', 0)$  hence we have  $x \cdot y = (x' \cdot y' - 0 \cdot 0, x' \cdot 0 + y' \cdot 0) = (x' \cdot y', 0) \in \mathbb{R}_{\mathbb{C}}$

$\forall x \in \mathbb{R}_{\mathbb{C}} \models -x \in \mathbb{R}_{\mathbb{C}}$ . as  $x \in \mathbb{R}_{\mathbb{C}}$  there exists a  $x' \in \mathbb{R}$  such that  $x = (x', 0)$  hence  $-x = (-x', 0) = (-x', 0) \in \mathbb{R}_{\mathbb{C}}$

$0 \in \mathbb{R}_{\mathbb{C}}$ . this is trivial as  $0 = (0, 0) \in \mathbb{R}_{\mathbb{C}}$  (see 9.2)

$1 \in \mathbb{R}_{\mathbb{C}}$ . this is trivial as  $1 = (1, 0) \in \mathbb{R}_{\mathbb{C}}$  (see 9.2)

□

The following theorem combined with the above that we can identify  $\mathbb{R}$  with the sub field  $\mathbb{R}_{\mathbb{C}}$  of  $\mathbb{C}$

**Theorem 9.14.**  $i_{\mathbb{R}}: \mathbb{R} \rightarrow \mathbb{R}_{\mathbb{C}}$  defined by  $i_{\mathbb{R}}(x) = (x, 0)$  is a injection with  $i_{\mathbb{R}}(\mathbb{R}) = \mathbb{R}_{\mathbb{C}}$  such that  $i_{\mathbb{R}}: \mathbb{R} \rightarrow \mathbb{R}_{\mathbb{C}}$  is a field isomorphism (see 3.33)

**Proof.** First  $i_{\mathbb{R}}(x) = i_{\mathbb{R}}(y)$  then  $(x, 0) = (y, 0) \Rightarrow x = y$ . Second if  $x \in \mathbb{R}_{\mathbb{C}}$  then  $\exists x' \in \mathbb{R}$  such that  $x = (x', 0) = i_{\mathbb{R}}(x')$  proving that  $i_{\mathbb{R}}: \mathbb{R} \rightarrow \mathbb{R}_{\mathbb{C}}$  is a bijection. To prove that  $i_{\mathbb{R}}$  is a field isomorphism we must prove that it conserve the addition and multiplication operations.

1. If  $x, y \in \mathbb{R}$  then  $i_{\mathbb{R}}(x) + i_{\mathbb{R}}(y) = (x, 0) + (y, 0) = (x + y, 0 + 0) = (x + y, 0) = i_{\mathbb{R}}(x + y)$
2. If  $x, y \in \mathbb{R}$  then  $i_{\mathbb{R}}(x) \cdot i_{\mathbb{R}}(y) = (x, 0) \cdot (y, 0) = (x \cdot y - 0 \cdot 0, x \cdot 0 + y \cdot 0) = (x \cdot y, 0) = i_{\mathbb{R}}(x \cdot y)$
3.  $i_{\mathbb{R}}(0) = (0, 0)$  the neutral element in  $\mathbb{R}_{\mathbb{C}}$  and  $\mathbb{C}$
4.  $i_{\mathbb{R}}(1) = (1, 0)$  the neutral multiplicative element in  $\mathbb{R}_{\mathbb{C}}$  and  $\mathbb{C}$

□

**Definition 9.15.**  $i \in \mathbb{C}$  is defined by  $i = (0, 1)$  and  $1 \in \mathbb{R}_{\mathbb{C}}$  is defined by  $1 = (1, 0)$

**Note 9.16.** As  $\bar{i} = \overline{(0, 1)} = (1, 0) = 1$

**Definition 9.17.** If  $z = (x, y) \in \mathbb{C}$  we have  $\operatorname{Re}(z) = (x, 0) \in \mathbb{R}_{\mathbb{C}}$  and  $\operatorname{Img}(z) = (y, 0) \in \mathbb{R}_{\mathbb{C}}$

**Lemma 9.18.** Let  $z \in \mathbb{C}$  then  $z = \operatorname{Re}(z) + i \cdot \operatorname{Img}(z)$  further if  $z = x + i \cdot y$  where  $x, y \in \mathbb{R}_{\mathbb{C}}$  then  $x = \operatorname{Re}(z) \wedge y = \operatorname{Img}(z)$

**Proof.** Take  $z \in \mathbb{C}$  then we have  $z = (z, y)$  further  $i \cdot \operatorname{Img}(z) = (0, 1) \cdot (y, 0) = (0 \cdot y - 1 \cdot 0, 0 \cdot 0 + 1 \cdot y) = (0, y)$  hence  $\operatorname{Re}(z) + i \cdot \operatorname{Img}(z) = (x, 0) + (0, y) = (x, y) = z$ . Finally if  $z = x + i \cdot y$  where  $x, y \in \mathbb{R}_{\mathbb{C}}$  then  $\exists x', y' \in \mathbb{R}$  such that  $x = (x', 0) \wedge y = (y', 0)$  so  $z = (x', 0) + i \cdot (y', 0) = (x', 0) + (0, 1) \cdot (y', 0) = (x', 0) + (0 \cdot y' - 1 \cdot 0, 0 \cdot 0 + 1 \cdot y') = (x', 0) + (0, y') = (x', y')$  hence  $\operatorname{Re}(z) = (x', 0) = x$  and  $\operatorname{Img}(z) = (y', 0) = y$   $\square$

**Theorem 9.19.** Let  $z \in \mathbb{C}$  then  $2 \cdot \operatorname{Re}(z) = z + \bar{z}$  and  $2 \cdot \operatorname{Img}(z) = z - \bar{z}$

**Proof.** Let  $z = (x, y)$  then  $z + \bar{z} = (x, y) + (x, -y) = (2 \cdot x, 0) = 2 \cdot (x, 0) = 2 \cdot \operatorname{Re}(z)$ , further  $z - \bar{z} = (x, y) - (x, -y) = (0, 2 \cdot y) = 2 \cdot (0, y) = 2 \cdot \operatorname{Img}(z)$   $\square$

The above motivates the following notation of complex numbers

**Notation 9.20.** Let  $z \in \mathbb{C}$  then we can write  $z$  in a unique way as  $z = x + i \cdot y$  where  $x = \operatorname{Re}(z)$ ,  $y = \operatorname{Img}(z) \in \mathbb{R}_{\mathbb{C}}$

**Theorem 9.21.**  $i^2 = -1$

**Proof.**  $i^2 = (0, 1) \cdot (0, 1) = (0 \cdot 0 - 1 \cdot 1, 0 \cdot 1 + 1 \cdot 0) = (-1, 0) = -(1, 0) = -1$  (the unit in  $\mathbb{C}$ )  $\square$

**Corollary 9.22.** If  $n \in \mathbb{N}$  then  $i^{4 \cdot n} = 1$ ,  $i^{4 \cdot n+1} = i$ ,  $i^{4 \cdot n+2} = -1$ ,  $i^{4 \cdot n+3} = -i$

**Proof.** We prove the first part by induction so let  $\mathcal{S} = \{k \in \mathbb{N} \mid i^{4 \cdot k} = 1\}$  then we have

**$n = 0$ .** as  $n = 0$  we have  $i^{4 \cdot 0} = i^0 = 1$  and  $i^{4 \cdot 0+1} = i^1 = i$

**$n \in \mathcal{S} \Rightarrow n + 1$ .** we have

$$\begin{aligned} i^{4 \cdot (n+1)} &= i^{4 \cdot n+4} \\ &= i^{4 \cdot n} \cdot i \cdot i \cdot i \cdot i \\ &= i^{4 \cdot n} \cdot (-1) \cdot (-1) \\ &= i^{4 \cdot n} \\ &\stackrel{n \in \mathcal{S}}{=} 1 \end{aligned}$$

proving that  $n + 1$ .

Mathematical induction proves then that  $i^{4 \cdot n} = 1$ . Next  $i^{4 \cdot n+1} = i^{4 \cdot n} \cdot i = i$ ,  $i^{4 \cdot n+2} = i^{4 \cdot n+1} \cdot i = i \cdot i = -1$  and finally  $i^{4 \cdot n+3} = i^{4 \cdot n+2} \cdot i = (-1) \cdot i = -i$   $\square$

## 9.2 Order relation on $\mathbb{C}$

**Notation 9.23.** To avoid any confusion we use  $\leqslant$  for the order relation on  $\mathbb{R}$ .  $\leqslant_{\mathbb{C}}$  is the order relation on  $\mathbb{C}$  and  $\leqslant_{\mathbb{R}}$  will be the order relation on  $\mathbb{R}_{\mathbb{C}}$

We use lexical ordering (see 2.145) to define a partial order on  $\mathbb{C}$  that is fully ordered.

**Definition 9.24.** Define  $\leqslant_{\mathbb{C}} \in \mathbb{C} \times \mathbb{C}$  by  $(x, y) \leqslant_{\mathbb{C}} (x', y')$  iff either

1.  $x < x'$
2.  $x = x'$  and  $y \leqslant y'$

Using 2.145 together with 8.38 we have that  $\leqslant$  is a partial order relation on  $\mathbb{C}$  that is fully ordered

As  $\mathbb{R}_{\mathbb{C}} \subseteq \mathbb{C}$  we can use the above and 2.143 that

**Theorem 9.25.** *If  $x, y \in \mathbb{R}$  with  $x \leq y \Leftrightarrow i_{\mathbb{R}}(x) \leq_{\mathbb{C}} i_{\mathbb{R}}(y)$  and if  $x < y \Leftrightarrow i_{\mathbb{R}}(x) <_{\mathbb{C}} i_{\mathbb{R}}(y)$*

**Proof.** We have the following cases to consider for  $x \leq y$

$x < y$ . then  $i_{\mathbb{R}}(x) = (x, 0) \leq_{\mathbb{C}} (y, 0) = i_{\mathbb{R}}(y)$  and as  $(x, 0) \neq (y, 0)$  [if  $(x, 0) = (y, 0) \Rightarrow x = y$ ] we have  $i_{\mathbb{R}}(x) < i_{\mathbb{R}}(y)$

$x = y$ . then as  $0 \leq 0$  we have  $(x, 0) = i_{\mathbb{R}}(x) \leq_{\mathbb{C}} i_{\mathbb{R}}(y) = (y, 0)$

For the opposite implication if  $i_{\mathbb{R}}(x) \leq_{\mathbb{C}} i_{\mathbb{R}}(y)$  then we have either

$i_{\mathbb{R}}(x) < i_{\mathbb{R}}(y)$ . assume that  $x \not\leq y$  then by 9.41 we have  $y \leq x$  hence by the first part of the theorem we have  $i_{\mathbb{R}}(y) \leq_{\mathbb{C}} i_{\mathbb{R}}(x) <_{\mathbb{C}} i_{\mathbb{R}}(y)$  giving the contradiction  $i_{\mathbb{R}}(y) <_{\mathbb{C}} i_{\mathbb{R}}(y)$ . Hence we must have  $x < y$

$i_{\mathbb{R}}(x) = i_{\mathbb{R}}(y)$ . as  $i_{\mathbb{R}}$  is a bijection we conclude then that  $x = y$  □

**Definition 9.26.**  $\leq_{\mathbb{R}} \subseteq \mathbb{R}_{\mathbb{C}} \times \mathbb{R}_{\mathbb{C}}$  is defined by  $\{(x, 0), (y, 0) | x, y \in \mathbb{R} \wedge x \leq y\}$

**Theorem 9.27.**  $\leq_{\mathbb{R}} = \leq_{\mathbb{C}} \cap (\mathbb{R}_{\mathbb{C}} \times \mathbb{R}_{\mathbb{C}})$  hence using 2.149 we have that  $\langle \mathbb{R}_{\mathbb{C}}, \leq_{\mathbb{R}} \rangle$  is a fully ordered set.

**Proof.** If  $(z, z') \in \mathbb{R}_{\mathbb{C}}$  there exists  $x, x' \in \mathbb{R}$  such that  $z = (x, 0) \wedge z' = (x', 0)$ , hence if  $(z, z') \in \leq_{\mathbb{R}}$  we have  $x \leq x'$  which proves that  $(z, z') \in \leq_{\mathbb{C}}$ , also if  $(z, z') \in \leq_{\mathbb{C}}$  we must have  $x \leq x'$  proving that  $(z, z') \in \leq_{\mathbb{R}}$ . So we conclude that (see 2.143)

$$\leq_{\mathbb{R}} = \leq_{\mathbb{C}} \cap \mathbb{R}_{\mathbb{C}} \times \mathbb{R}_{\mathbb{C}}$$
□

**Corollary 9.28.**  $i_{\mathbb{R}}: \mathbb{R} \rightarrow \mathbb{R}_{\mathbb{C}}$  is a order preserving field isomorphism (see isomorphism and 3.33)

**Proof.** This follows from 9.14, 9.27 and 9.25 □

**Theorem 9.29.** *If  $x, y \in \mathbb{R}_{\mathbb{C}}$  with  $0 \leq_{\mathbb{R}} x, y$  and  $x + y = 0$  then  $x = y = 0$*

**Proof.** First there exists a  $x', y' \in \mathbb{R}$  such that  $x = i_{\mathbb{R}}(x') = (x', 0)$ ,  $y = i_{\mathbb{R}}(y') = (y, 0)$  and as  $0 = (0, 0) = x + y = (x', 0) + (y', 0) = (x' + y', 0)$  we have  $x' + y' = 0$ . Also from  $0 \leq_{\mathbb{R}} x$  we have  $0 \leq x'$  [if  $x' < 0$  then  $0 = i_{\mathbb{R}}(x') <_{\mathbb{R}} i_{\mathbb{R}}(0) = x \Rightarrow x <_{\mathbb{R}} 0 \leq_{\mathbb{R}} x \Rightarrow x <_{\mathbb{R}} x$  a contradiction] and similar we have  $0 \leq y'$ . So using 9.42 we have then  $x' = 0 = y'$  but this means that  $x = y = (0, 0) = 0$  □

**Definition 9.30.**  $\mathbb{R}_{\mathbb{C}+} = \{\alpha \in \mathbb{R}_{\mathbb{C}} | \alpha >_{\mathbb{R}} 0\}$  and  $\mathbb{R}_{\mathbb{C}-} = \{\alpha \in \mathbb{R}_{\mathbb{C}} | \alpha <_{\mathbb{R}} 0\}$

Based on the properties (see 9.410) of the reals and 9.13 we have the following theorem

**Theorem 9.31.** *We have the following for the set of reals embedded in the complex*

1. *If  $\alpha, \beta \in \mathbb{R}_{\mathbb{C}}$  then we have the following exclusive possibilities*

a.  $\alpha <_{\mathbb{R}} \beta$

b.  $\beta <_{\mathbb{R}} \alpha$

c.  $\alpha = \beta$

2. *If  $\alpha, \beta \in \mathbb{R}_{\mathbb{C}}$  with  $\alpha <_{\mathbb{R}} \beta \Rightarrow -\beta <_{\mathbb{R}} -\alpha$*

3. *If  $\alpha, \beta, \gamma \in \mathbb{R}_{\mathbb{C}}$  with  $\alpha <_{\mathbb{R}} \beta \Rightarrow \alpha + \gamma <_{\mathbb{R}} \beta + \gamma$*

4. *If  $\alpha, \beta \in \mathbb{R}_{\mathbb{C}}$  and  $\alpha <_{\mathbb{R}} \beta$  with  $\gamma \in \mathbb{R}_{\mathbb{C}+}$  then  $\alpha \cdot \gamma <_{\mathbb{R}} \beta \cdot \gamma$*

5. *If  $\alpha, \beta \in \mathbb{R}_{\mathbb{C}}$  and  $\alpha <_{\mathbb{R}} \beta$  with  $\gamma \in \mathbb{R}_{\mathbb{C}-}$  then  $\beta \cdot \gamma <_{\mathbb{R}} \alpha \cdot \gamma$*

6. *If  $0 <_{\mathbb{R}} \alpha \Rightarrow 0 <_{\mathbb{R}} \alpha^{-1}$*

7. If  $0 <_{\mathbb{R}} \alpha <_{\mathbb{R}} \beta \Rightarrow \beta^{-1} <_{\mathbb{R}} \alpha^{-1}$
8. If  $\alpha \in \mathbb{R}_{\mathbb{C}} \Rightarrow 0 \leq_{\mathbb{R}} \alpha^2$  and if  $\alpha \neq 0$  then  $0 <_{\mathbb{R}} \alpha^2$
9. If  $\alpha, \beta \in \mathbb{R}_{\mathbb{C}+} \cup \{0\}$  are such that  $\alpha <_{\mathbb{R}} \beta$  then  $\alpha^2 <_{\mathbb{R}} \beta^2$  (so  $\alpha^2: \{x \in \mathbb{R}_{\mathbb{C}} | 0 \leq_{\mathbb{R}} x\} \rightarrow \{x \in \mathbb{R}_{\mathbb{C}} | 0 \leq_{\mathbb{R}} x\}$  is a strictly increasing function)
10. If  $\alpha = (\alpha', 0) \in \mathbb{R}_{\mathbb{C}+}$  and  $n \in \mathbb{N}$  then  $\alpha^n = (\alpha'^n, 0) \Rightarrow \alpha^n \in \mathbb{R}_{\mathbb{C}+}$

**Proof.** First note that if  $\alpha, \beta \in \mathbb{R}_{\mathbb{C}}$  then there exists a  $\alpha', \beta'$  such that  $\alpha = (\alpha', 0), \beta = (\beta', 0)$  and  $\alpha \leq_{\mathbb{R}} \beta$  if and only if  $\alpha' \leq \beta'$  also  $\alpha = \beta$  if and only if  $\alpha' = \beta'$  so that we have also  $\alpha <_{\mathbb{R}} \beta$  if and only if  $\alpha' < \beta'$

1. If  $\alpha = (\alpha', 0), \beta = (\beta', 0)$  then as for  $\alpha', \beta'$  we have the following possibilities

- a.  $\alpha' < \beta'$
- b.  $\beta' = \alpha'$
- c.  $\alpha' = \beta'$

we have the following possibilities for  $\alpha, \beta$

- a.  $\alpha <_{\mathbb{R}} \beta$
- b.  $\beta <_{\mathbb{R}} \alpha$
- c.  $\alpha = \beta$

2. If  $(\alpha', 0) = \alpha <_{\mathbb{R}} \beta = (\beta', 0) \Rightarrow \alpha' < \beta' \Rightarrow -\beta' < -\alpha' \Rightarrow -\alpha <_{\mathbb{R}} -\beta$
3. If  $(\alpha', 0) = \alpha <_{\mathbb{R}} \beta = (\beta', 0) \Rightarrow \alpha' < \beta'$  and  $\gamma = (\gamma', 0)$  then  $\alpha' + \gamma' < \beta' + \gamma' \Rightarrow \alpha + \gamma = (\alpha' + \gamma', 0) <_{\mathbb{R}} (\beta' + \gamma', 0) = \beta + \gamma$
4. If  $\alpha, \beta \in \mathbb{R}_{\mathbb{C}}$  and  $\gamma \in \mathbb{R}_{\mathbb{C}+}$  then  $\alpha = (\alpha', 0), \beta = (\beta', 0), \gamma = (\gamma', 0)$  where  $\alpha, \beta \in \mathbb{R}$  and  $\gamma \in \mathbb{R}_+$  then from  $\alpha <_{\mathbb{R}} \beta$  it follows that  $\alpha' < \beta' \xrightarrow{\gamma' \in \mathbb{R}_+} \alpha' \cdot \gamma' < \beta' \cdot \gamma' \Rightarrow \alpha \cdot \gamma = (\alpha' \cdot \gamma', 0) <_{\mathbb{R}} (\beta' \cdot \gamma', 0) = \beta \cdot \gamma$
5. If  $\alpha, \beta \in \mathbb{R}_{\mathbb{C}}$  and  $\gamma \in \mathbb{R}_{\mathbb{C}-}$  then  $\alpha = (\alpha', 0), \beta = (\beta', 0), \gamma = (\gamma', 0)$  where  $\alpha, \beta \in \mathbb{R}$  and  $\gamma \in \mathbb{R}_-$  then from  $\alpha <_{\mathbb{R}} \beta$  it follows that  $\alpha' < \beta' \xrightarrow{\gamma' \in \mathbb{R}_-} \beta' \cdot \gamma' < \alpha' \cdot \gamma' \Rightarrow \beta \cdot \gamma = (\beta' \cdot \gamma', 0) <_{\mathbb{R}} (\alpha' \cdot \gamma', 0) = \alpha \cdot \gamma$
6. If  $0 = (0, 0) <_{\mathbb{R}} \alpha = (\alpha', 0) \Rightarrow 0 < \alpha' \Rightarrow 0 < \alpha'^{-1} \Rightarrow 0 = (0, 0) <_{\mathbb{R}} (\alpha'^{-1}, 0) = \alpha^{-1}$
7. If  $0 <_{\mathbb{R}} (\alpha', 0) = \alpha <_{\mathbb{R}} \beta = (\beta', 0) \Rightarrow 0 < \alpha' < \beta' = > \beta'^{-1} < \alpha'^{-1} \Rightarrow \beta^{-1} = (\beta'^{-1}, 0) <_{\mathbb{R}} (\alpha'^{-1}, 0) < \alpha^{-1}$
8. If  $\alpha = (\alpha', 0) \in \mathbb{R}_{\mathbb{C}}$  then  $0 \leq_{\mathbb{R}} (<_{\mathbb{R}}) \alpha'^2$  ( $<$  if  $\alpha' \neq 0$ ) so we have  $0 \leq_{\mathbb{R}} (<_{\mathbb{R}}) (\alpha'^2, 0) = \alpha^2$  ( $<_{\mathbb{R}}$  if  $\alpha = (\alpha', 0) \neq (0, 0) = 0$ )
9. If  $\alpha = (\alpha', 0), \beta = (\beta', 0) \in \mathbb{R}_{\mathbb{C}+} \cup \{0\}$  are such that  $\alpha <_{\mathbb{R}} \beta$  then  $\alpha', \beta' \in \mathbb{R}_+ \cup \{0\}$  so  $\alpha'^2 < \beta'^2 \Rightarrow \alpha^2 = (\alpha'^2, 0) <_{\mathbb{R}} (\beta'^2, 0) = \beta^2$
10. We prove this by induction, so let  $S = \{n \in \mathbb{N} | \text{if } (\alpha', 0) = \alpha \in \mathbb{R}_{\mathbb{C}+} \text{ then } \alpha^n = (\alpha'^n, 0)\}$  then we have
  - a. if  $n = 0$  then by definition we have  $\alpha^0 = (\alpha', 0) = (\alpha'^0, 0) \Rightarrow 0 \in S$
  - b. If  $n \in S$  then  $\alpha^{n+1} = \alpha^n \cdot a = (\alpha'^n, 0) \cdot (\alpha', 0) = (\alpha'^{n+1}, 0) \Rightarrow n+1 \in S$

□

### 9.3 Norm on $\mathbb{C}$

**Definition 9.32.** Given  $(x, y) \in \mathbb{C}$  we define  $|(x, y)| = \left( \sqrt{x^2 + y^2}, 0 \right) \in \mathbb{R}_{\mathbb{C}}$  (which is defined as  $0 \leq x^2 + y^2 \in \mathbb{R}$ ) and this is called the **complex norm**

**Theorem 9.33.**  $\forall z, z' \in \mathbb{C}$  we have

1.  $0 \leqslant_{\mathbb{R}} |z|$
2.  $|z \cdot z'| = |z| \cdot |z'|$
3.  $|\bar{z}| = |z|$
4.  $\operatorname{Re}(z) \leqslant_{\mathbb{R}} |z|$  and  $\operatorname{Img}(z) \leqslant_{\mathbb{R}} |z|$
5.  $z \cdot \bar{z} = |z|^2$
6. If  $x, y \in \mathbb{R}_{\mathbb{C}}$  with  $0 \leqslant_{\mathbb{R}} x, y$  then from  $x^2 \leqslant_{\mathbb{R}} y^2$  we have  $x \leqslant_{\mathbb{R}} y$
7.  $|z + z'| \leqslant_{\mathbb{R}} |z| + |z'|$
8.  $|1| = |-1| = |i| = 1$
9.  $|z| \leqslant |\operatorname{Re}(z)| + |\operatorname{Img}(z)|$
10.  $|z| = 0 \Leftrightarrow z = 0$
11. If  $z \in \mathbb{R}_{\mathbb{C}}$  then  $|z| = |\operatorname{Re}(z)|$
12. If  $z \neq 0$  then  $\left| \frac{1}{z} \right| = \frac{1}{|z|}$

**Proof.**

1. Let  $z = (x, y)$  then as  $\sqrt{\cdot}: \{x \in \mathbb{R} | 0 \leqslant x\} \rightarrow \{x \in \mathbb{R} | 0 \leqslant x\}$  we have that  $0 \leqslant_{\mathbb{R}} \sqrt{x^2 + y^2}$  given then two cases to consider:

$$\mathbf{0} < \sqrt{x^2 + y^2}. \text{ then } 0 <_{\mathbb{R}} \left( \sqrt{x^2 + y^2}, 0 \right) = |z| \text{ proving } 0 \leqslant_{\mathbb{R}} |z|$$

$$\mathbf{0} = \sqrt{x^2 + y^2}. \text{ then } 0 = (0, 0) = \left( \sqrt{x^2 + y^2}, 0 \right) = |z| \text{ proving } 0 \leqslant_{\mathbb{R}} |z|$$

2. If  $z = (x, y) \in \mathbb{C}$ ,  $z' = (x', y') \in \mathbb{C}$  then  $z \cdot z' = (x \cdot x' - y \cdot y', x \cdot y' + y \cdot x')$  and  $|z \cdot z'| = \left( \sqrt{(x \cdot x' - y \cdot y')^2 + (x \cdot y' + y \cdot x')^2}, 0 \right) = \left( \sqrt{x^2 \cdot x'^2 + y^2 \cdot y'^2 - 2 \cdot x \cdot x' \cdot y \cdot y' + x^2 \cdot y'^2 + y^2 \cdot x'^2 + 2 \cdot x \cdot x' \cdot y \cdot y'}, 0 \right) = \left( \sqrt{x^2 \cdot x'^2 + y^2 \cdot y'^2 + x^2 \cdot y'^2 + y^2 \cdot x'^2}, 0 \right) = \left( \sqrt{x^2 \cdot (x'^2 + y'^2) + y^2 \cdot (x'^2 + y'^2)}, 0 \right) = \left( \sqrt{(x^2 + y^2) \cdot (x'^2 + y'^2)}, 0 \right) = |z| \cdot |z'|$

3. If  $z = (x, y)$  then  $|\bar{z}| = |(x, -y)| = \left( \sqrt{x^2 + (-y)^2}, 0 \right) = \left( \sqrt{x^2 + y^2}, 0 \right) = |z|$

4. If  $z = (x, y)$  then  $\operatorname{Re}(z) = (x, 0)$  and  $\operatorname{Img}(z) = (0, y)$  Now as  $0 \leqslant y^2$  we have that  $x^2 \leqslant x^2 + y^2$   $\sqrt{\cdot}$  is increasing (see 9.72)  $\sqrt{|x|^2} \leqslant \sqrt{x^2 + y^2}$  hence as  $x \leqslant |x| = \sqrt{|x|^2}$  we have  $x \leqslant \sqrt{x^2 + y^2}$  so that  $\operatorname{Re}(z) = (x, 0) \leqslant_{\mathbb{R}} \left( \sqrt{x^2 + y^2}, 0 \right) = |z|$ . Further as  $0 \leqslant x^2$  we have that  $y^2 \leqslant x^2 + y^2$   $\sqrt{\cdot}$  is increasing (see 9.72)  $\sqrt{|y|^2} \leqslant \sqrt{x^2 + y^2}$  hence as  $y \leqslant |y| = \sqrt{|y|^2}$  we have  $y \leqslant \sqrt{x^2 + y^2}$  so that  $\operatorname{Img}(z) = (0, y) \leqslant_{\mathbb{R}} \left( \sqrt{x^2 + y^2}, 0 \right) = |z|$ .

5. If  $z = (x, y)$  then  $z \cdot \bar{z} = (x, y) \times (x, -y) = (x \cdot x + y \cdot y, -x \cdot y + y \cdot x) = (x^2 + y^2, 0) \underset{0 \in \mathbb{R}^2}{=} \left( \left( \sqrt{x^2 + y^2} \right)^2, 0 \right) = \left( \sqrt{x^2 + y^2}, 0 \right) \cdot \left( \sqrt{x^2 + y^2}, 0 \right) = |z|^2$

6. Let  $z_1, z_2 \in \mathbb{R}_{\mathbb{C}}$  then there exists  $x', y' \in \mathbb{R}$  such that  $x = (x', 0) \wedge y = (y', 0)$  hence  $x^2 = (x' \cdot x' - 0, 0, x', 0 + 0 \cdot x') = (x'^2, 0)$  and  $y^2 = (y', 0) \cdot (y', 0) = (y' \cdot y' - 0 \cdot 0, y' \cdot 0 + 0 \cdot y') = (y'^2, 0)$  now as  $0 \leqslant x, y$  we have that  $0 \leqslant x', y'$  and as  $\sqrt{\cdot}$  is a increasing function on 9.72 we have that  $x' \leqslant y$  hence  $x = (x', 0) = i_{\mathbb{R}}(x') \leqslant_{\mathbb{R}_{9.28}} i_{\mathbb{R}}(y') = (y', 0) = y$

7. Given  $x, y \in \mathbb{C}$  then we have by (5)

$$\begin{aligned}
 |x+y|^2 &\stackrel{(5)}{=} (x+y) \cdot \overline{(x+y)} \\
 &\stackrel{9.5}{=} (x+y) \cdot (\bar{x}+\bar{y}) \\
 &= x \cdot \bar{x} + x \cdot \bar{y} + y \cdot \bar{x} + y \cdot \bar{y} \\
 &\stackrel{(5)}{=} |x|^2 + |y|^2 + (x \cdot \bar{y} + \bar{x} \cdot \bar{y}) \\
 &\stackrel{9.19}{=} |x|^2 + |y|^2 + 2 \cdot \operatorname{Re}(x \cdot \bar{y}) \\
 &\leq_{\mathbb{R}(4)} |x|^2 + |y|^2 + 2 \cdot |x \cdot \bar{y}| \\
 &\stackrel{(2)}{=} |x|^2 + |y|^2 + 2 \cdot |x| \cdot |y| \\
 &= (|x| + |y|)^2
 \end{aligned}$$

hence using (6) we have  $|x+y| \leq_{\mathbb{R}} |x| + |y|$

8.

$$\begin{aligned}
 |1| &= |(1, 0)| = \sqrt{1^2 + 0^2} = 1 \\
 |-1| &= |(-1, 0)| = \sqrt{(-1)^2 + 0^2} = 1 \\
 |i| &= |(0, 1)| = \sqrt{0^2 + 1^2} = 1
 \end{aligned}$$

9. As  $|z| = |\operatorname{Re}(z) + i \cdot \operatorname{Img}(z)|$  we have using (7) that

$$\begin{aligned}
 |z| &\leq |\operatorname{Re}(z)| + |i \cdot \operatorname{Img}(z)| \\
 &\stackrel{(2)}{=} |\operatorname{Re}(z)| + |i| \cdot |\operatorname{Img}(z)| \\
 &\stackrel{(8)}{=} |\operatorname{Re}(z)| + |\operatorname{Img}(z)|
 \end{aligned}$$

10. If  $z = (x, y)$  then if  $z = 0 = (0, 0)$  we have  $|z| = (\sqrt{0^2 + 0^2}, 0) = (\sqrt{0}, 0) = 0$ . On the other side if  $|z| = 0$  then  $(\sqrt{x^2 + y^2}, 0) = 0 = (0, 0) \Rightarrow (\sqrt{x^2 + y^2}) = 0 = 0^2 \stackrel{0 \leq x^2 + y^2}{\Rightarrow} x^2 + y^2 = 0$ . If now  $x \neq 0 \Rightarrow 0 < |x| \Rightarrow 0 < |x|^2 = x^2 \stackrel{0 \leq y^2 \text{ and } 9.41}{\Rightarrow} 0 < x^2 + y^2$  a contradiction. So we must have that  $x = 0$  and thus  $0 = x^2 + y^2 = 0^2 + y^2 = y^2$ , if  $y \neq 0 \Rightarrow 0 < |y| \Rightarrow 0 < |y|^2 = y^2 = 0$  a contradiction so we have  $y = 0$  and thus  $z = (0, 0) = 0$

11. If  $z \in \mathbb{R}_{\mathbb{C}}$  then  $\exists z' \in \mathbb{R}$  such that  $z = (z', 0)$  hence  $\operatorname{Re}(z) = z$  so that  $|\operatorname{Re}(z)| = |z|$

12. As  $1 = z \cdot \frac{1}{z}$  we have  $1 \stackrel{(8)}{=} |1| = |z \cdot \frac{1}{z}| \stackrel{(2)}{=} |z| \cdot \left| \frac{1}{z} \right| \Rightarrow \left| \frac{1}{z} \right| = \frac{1}{|z|}$ .

□

## 9.4 $\mathbb{R}$ and $\mathbb{R}_{\mathbb{C}}$

### 9.4.1 $\mathbb{N}_{\mathbb{C}}$ , $\mathbb{Z}_{\mathbb{C}}$ and $\mathbb{Q}_{\mathbb{C}}$

**Definition 9.34.** Define  $\mathbb{N}_{\mathbb{C}} \subseteq \mathbb{Z}_{\mathbb{C}} \subseteq \mathbb{Q}_{\mathbb{C}} \subseteq \mathbb{R}_{\mathbb{C}} \subseteq \mathbb{C}$  by  $\mathbb{N}_{\mathbb{C}} = \{(x, 0) \in \mathbb{R}_{\mathbb{C}} \mid x \in \mathbb{N}_{\mathbb{R}}\}$ ,  $\mathbb{Z}_{\mathbb{C}} = \{(x, 0) \mid x \in \mathbb{Z}_{\mathbb{R}}\}$ ,  $\mathbb{Q}_{\mathbb{C}} = \{(x, 0) \in \mathbb{R}_{\mathbb{C}} \mid x \in \mathbb{Q}_{\mathbb{R}}\}$

**Theorem 9.35.**  $\langle \mathbb{N}_{\mathbb{C}}, + \rangle$ ,  $\langle \mathbb{N}_{\mathbb{C}}, \cdot \rangle$  forms a semi-groups and if we define  $i_{\mathbb{R}_{\mathbb{N}}} : \mathbb{N} \rightarrow \mathbb{N}_{\mathbb{C}}$  by  $i_{\mathbb{R}_{\mathbb{N}}} = i_{\mathbb{R}_{|\mathbb{N}_{\mathbb{R}}}} \circ i_{\mathbb{Q}_{\mathbb{N}}}$  (see 8.56) then  $i_{\mathbb{R}_{\mathbb{N}}}$  is a order preserving semi-group isomorphism (for  $+$  and  $\cdot$ )

**Proof.** First we prove that  $\langle \mathbb{N}_{\mathbb{C}}, + \rangle$  and  $\langle \mathbb{N}_{\mathbb{C}}, \cdot \rangle$  are semi-groups. If  $x, y \in \mathbb{N}_{\mathbb{C}}$  then we have  $x = (x', 0)$ ,  $y = (y', 0)$  and  $x', y' \in \mathbb{N}_{\mathbb{R}}$  so that

1.  $x + y = (x', 0) + (y', 0) = (x' + y', 0) \in \mathbb{N}_{\mathbb{C}}$  as  $x' + y' \in \mathbb{N}_{\mathbb{R}}$  (for  $\mathbb{N}_{\mathbb{R}}$  is a semi-group)
2.  $0 = (0, 0) \in \mathbb{N}_{\mathbb{C}}$

3.  $x \cdot y = (x' \cdot y' - 0 \cdot 0, x' \cdot 0 + 0 \cdot y') = (x' \cdot y', 0) \in \mathbb{N}_{\mathbb{C}}$  (as  $x' \cdot y' \in \mathbb{N}_{\mathbb{C}}$  which is a semi-group)
4.  $1 = (1, 0) \in \mathbb{N}_{\mathbb{C}}$

To prove that  $i_{\mathbb{R}_{\mathbb{N}}}$  is a order preserving semi-group isomorphism, note that by 8.56  $i_{\mathbb{Q}_{\mathbb{N}}}$  is a order preserving semi-group isomorphism and that by 9.14, 9.25  $i_{\mathbb{R}}$  is a order preserving isomorphism and thus  $i_{\mathbb{R}_{\mathbb{N}}}$  is a order preserving semi-group isomorphism.  $\square$

**Theorem 9.36.**  $\langle \mathbb{Z}_{\mathbb{C}}, +, \cdot \rangle$  forms a sub ring of  $\langle \mathbb{C}, +, \cdot \rangle$  [as a field  $\langle \mathbb{C}, +, \cdot \rangle$  is also a ring] with additive neutral element  $0 = (0, 0)$  and multiplicative unit  $(1, 0)$ . Further there exists a order preserving ring isomorphism between  $\mathbb{Z}$  and  $\mathbb{Z}_{\mathbb{C}}$ . So we can consider  $\mathbb{Z}_{\mathbb{C}}$  as a embedding of  $\mathbb{Z}$  in  $\mathbb{C}$ .

**Proof.** If  $x, y \in \mathbb{Z}_{\mathbb{C}}$  then there exists a  $x', y' \in \mathbb{Z}_{\mathbb{R}}$  such that  $x = (x', 0) \wedge y = (y', 0)$ . We have then that

1.  $x + y = (x', 0) + (y', 0) = (x' + y', 0) \in \mathbb{Z}_{\mathbb{C}}$  as  $x' + y' \in \mathbb{Z}_{\mathbb{R}}$  (see 8.53)
2.  $x \cdot y = (x', 0) \cdot (y', 0) = (x' \cdot y' - 0 \cdot 0, x' \cdot 0 + y' \cdot 0) = (x' \cdot y', 0) \in \mathbb{Z}_{\mathbb{C}}$  as  $x' \cdot y' \in \mathbb{Z}_{\mathbb{R}}$  (see 8.53)
3.  $-x = -(x', 0) \stackrel{9.2}{=} (-x', -0) = (-x', 0) \in \mathbb{Z}_{\mathbb{C}}$  as  $-x' \in \mathbb{Z}_{\mathbb{R}}$  (see 8.53)
4.  $1 = (1, 0) \in \mathbb{Z}_{\mathbb{C}}$
5.  $0 = (0, 0) \in \mathbb{Z}_{\mathbb{C}}$

Next  $i_{\mathbb{R}_{\mathbb{Z}_{\mathbb{C}}}}: \mathbb{Z}_{\mathbb{R}} \rightarrow \mathbb{Z}_{\mathbb{C}}$  is a order preserving field (and thus ring) isomorphism (see 9.14, 9.25) hence as by 8.53  $i_{\mathbb{Q}_{\mathbb{Z}}}: \mathbb{Z} \rightarrow \mathbb{Z}_{\mathbb{R}}$  is a order preserving ring isomorphism we have that  $i_{\mathbb{R}_{\mathbb{Z}_{\mathbb{C}}}} \circ i_{\mathbb{Q}_{\mathbb{Z}}}: \mathbb{Z} \rightarrow \mathbb{Z}_{\mathbb{C}}$  is a order preserving ring isomorphism.  $\square$

**Theorem 9.37.**  $\langle \mathbb{Q}_{\mathbb{C}}, +, \cdot \rangle$  forms a sub field of  $\langle \mathbb{C}, +, \cdot \rangle$  with additive neutral element  $0 = (0, 0)$  and multiplicative unit  $(1, 0)$ . Further there exists a order preserving field isomorphism between  $\mathbb{Q}$  and  $\mathbb{Q}_{\mathbb{C}}$ . So we can consider  $\mathbb{Q}_{\mathbb{C}}$  as a embedding of  $\mathbb{Q}$  in  $\mathbb{C}$ .

**Proof.** If  $x, y \in \mathbb{Q}_{\mathbb{C}}$  then there exists a  $x', y' \in \mathbb{Q}_{\mathbb{R}}$  such that  $x = (x', 0) \wedge y = (y', 0)$ . We have then that

1.  $x + y = (x', 0) + (y', 0) = (x' + y', 0) \in \mathbb{Q}_{\mathbb{C}}$  as  $x' + y' \in \mathbb{Q}_{\mathbb{R}}$  (see 8.28)
2.  $x \cdot y = (x', 0) \cdot (y', 0) = (x' \cdot y' - 0 \cdot 0, x' \cdot 0 + y' \cdot 0) = (x' \cdot y', 0) \in \mathbb{Z}_{\mathbb{C}}$  as  $x' \cdot y' \in \mathbb{Z}_{\mathbb{R}}$  (see 8.28)
3.  $-x = -(x', 0) \stackrel{9.2}{=} (-x', -0) = (-x', 0) \in \mathbb{Q}_{\mathbb{C}}$  as  $-x' \in \mathbb{Q}_{\mathbb{R}}$  (see 8.28)
4.  $1 = (1, 0) \in \mathbb{Q}_{\mathbb{C}}$
5.  $0 = (0, 0) \in \mathbb{Q}_{\mathbb{C}}$

Next  $i_{\mathbb{R}_{\mathbb{Q}_{\mathbb{C}}}}: \mathbb{Q}_{\mathbb{R}} \rightarrow \mathbb{Q}_{\mathbb{C}}$  is a order preserving field isomorphism (see 9.14, 9.25) hence as by 8.28  $i_{\mathbb{Q}}: \mathbb{Q} \rightarrow \mathbb{Q}_{\mathbb{R}}$  is a order preserving ring isomorphism we have that  $i_{\mathbb{R}_{\mathbb{Q}_{\mathbb{C}}}} \circ i_{\mathbb{Q}}: \mathbb{Q} \rightarrow \mathbb{Q}_{\mathbb{C}}$  is a order preserving ring isomorphism.  $\square$

**Theorem 9.38.**  $\mathbb{N}_{\mathbb{C}}, \mathbb{Z}_{\mathbb{C}}$  and  $\mathbb{Q}_{\mathbb{C}}$  are denumerable.

**Proof.** This follows from the fact that  $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$  are denumerable and isomorphic (and thus bijective) with  $\mathbb{N}_{\mathbb{C}}, \mathbb{Z}_{\mathbb{C}}$  and  $\mathbb{Q}_{\mathbb{C}}$ .  $\square$

Next we prove that  $\mathbb{C}$  is isomorphic to  $\mathbb{R}_{\mathbb{C}}^2$

**Theorem 9.39.** The map  $\mathcal{C}: \mathbb{R}_{\mathbb{C}} \times \mathbb{R}_{\mathbb{C}} \rightarrow \mathbb{C}$  defined by  $\mathcal{C}((x, y)) = x + i \cdot y$  is a bijection.

**Proof.** First we prove that  $\mathcal{C}$  is a bijection

**injectivity.** Assume that  $\mathcal{C}((x_1, y_1)) = \mathcal{C}((x_2, y_2))$  then there exists  $x'_1, y'_1, x'_2, y'_2 \in \mathbb{R}$  such that  $x_1 = (x'_1, 0), x_2 = (x'_2, 0), y_1 = (y'_1, 0)$  and  $y_2 = (y'_2, 0)$  hence  $(x'_1, y'_1) = (x', 0) + (0, y'_1) = (x', 0) + (0, 1) \cdot (y'_1, 0) = x_1 + i \cdot y_1 = \mathcal{C}((x_1, y_1)) = \mathcal{C}((x_2, y_2)) = x_2 + i \cdot y_2 = (x'_2, 0) + (0, 1) \cdot (y'_2, 0) = (x'_2, 0) + (0, y'_2) = (x'_2, y'_2)$  proving that  $x'_1 = x'_2 \wedge y'_1 = y'_2 \Rightarrow x_1 = x_2 \wedge y_1 = y_2 \Rightarrow (x_1, y_1) = (x_2, y_2)$  proving injectivity.

**surjectivity.** Take  $z \in \mathbb{C}$  then there exists  $x', y' \in \mathbb{R}$  such that  $z = (x', y') = (x', 0) + (0, y') = (x', 0) + (0, 1) \cdot (y', 0) = x + i \cdot y$  where  $x = (x', 0), y = (y', 0) \in \mathbb{R}_{\mathbb{C}}$  hence  $z = \mathcal{C}((x, y))$  proving that  $\mathcal{C}$  is bijective.  $\square$

### 9.4.2 Properties of $\mathbb{R}_{\mathbb{C}}$

We show now that  $\mathbb{R}$  and  $\mathbb{R}_{\mathbb{C}}$  has the same properties

**Definition 9.40.**  $\mathbb{R}_{\mathbb{C}+} = \{z \in \mathbb{R}_{\mathbb{C}} \mid 0 <_{\mathbb{R}} z\}$  and  $\mathbb{R}_{\mathbb{C}-} = \{z \in \mathbb{R}_{\mathbb{C}} \mid z <_{\mathbb{R}} 0\}$

**Theorem 9.41.** *We have the following for the set of the reals*

1.  $\alpha \in \mathbb{R}_{\mathbb{C}+}$  iff  $0 <_{\mathbb{R}} \alpha$
2.  $\alpha \in \mathbb{R}_{\mathbb{C}-}$  iff  $\alpha <_{\mathbb{R}} 0$
3. If  $\alpha, \beta \in \mathbb{R}_{\mathbb{C}}$  then we have the following exclusive possibilities
  - a.  $\alpha <_{\mathbb{R}} \beta$
  - b.  $\beta <_{\mathbb{R}} \alpha$
  - c.  $\alpha = \beta$
4. If  $\alpha, \beta \in \mathbb{R}_{\mathbb{C}}$  with  $\alpha <_{\mathbb{R}} \beta \Rightarrow -\beta <_{\mathbb{R}} -\alpha$
5. If  $\alpha, \beta, \gamma \in \mathbb{R}_{\mathbb{C}}$  with  $\alpha <_{\mathbb{R}} \beta \Rightarrow \alpha + \gamma <_{\mathbb{R}} \beta + \gamma$
6. If  $\alpha, \beta \in \mathbb{R}_{\mathbb{C}}$  and  $\alpha <_{\mathbb{R}} \beta$  with  $\gamma \in \mathbb{R}_{\mathbb{C}+}$  then  $\alpha \cdot \gamma <_{\mathbb{R}} \beta \cdot \gamma$  [as we have  $\alpha = \beta \Rightarrow \alpha \cdot \lambda = \beta \cdot \lambda$  we have also that  $\alpha \leq \beta$  implies  $\alpha \cdot \lambda \leq \beta \cdot \lambda$ ]
7. If  $\alpha, \beta, \gamma, \delta \in \mathbb{R}_{\mathbb{C}+}$  with  $\alpha < \beta \wedge \gamma < \delta$  then  $\alpha \cdot \gamma < \beta \cdot \delta$
8. If  $\alpha, \beta \in \mathbb{R}_{\mathbb{C}}$  and  $\alpha <_{\mathbb{R}} \beta$  with  $\gamma \in \mathbb{R}_{\mathbb{C}-}$  then  $\beta \cdot \gamma <_{\mathbb{R}} \alpha \cdot \gamma$  [as we have  $\alpha = \beta \Rightarrow \alpha \cdot \lambda = \beta \cdot \lambda$  we have also that  $\alpha \leq \beta$  implies  $\beta \cdot \lambda \leq \alpha \cdot \lambda$ ]
9. If  $0 <_{\mathbb{R}} x \Rightarrow 0 <_{\mathbb{R}} x^{-1}$
10. If  $0 <_{\mathbb{R}} x <_{\mathbb{R}} y \Rightarrow y^{-1} <_{\mathbb{R}} x^{-1}$
11. If  $x \in \mathbb{R}_{\mathbb{C}} \Rightarrow 0 \leq_{\mathbb{R}} x^2$  and if  $x \neq 0$  then  $0 <_{\mathbb{R}} x^2$
12. If  $x, y \in \mathbb{R}_{\mathbb{C}+} \cup \{0\}$  are such that  $x <_{\mathbb{R}} y$  then  $x^2 <_{\mathbb{R}} y^2$  (so  ${}^2: \{x \in \mathbb{R}_{\mathbb{C}} \mid 0 \leq_{\mathbb{R}} x\} \rightarrow \{x \in \mathbb{R}_{\mathbb{R}} \mid 0 \leq x\}$  is a strictly increasing function)
13. If  $\alpha \in \mathbb{R}_{\mathbb{C}+}$  and  $n \in \mathbb{N}$  then  $\alpha^n \in \mathbb{R}_{\mathbb{C}+}$
14. If  $\alpha \in \mathbb{R}_{\mathbb{C}+}$  so that  $0 <_{\mathbb{R}} \alpha <_{\mathbb{R}} 1$  then if  $n \in \mathbb{N}$  we have  $0 <_{\mathbb{R}} \alpha^n <_{\mathbb{R}} \alpha$
15. If  $\alpha \in \mathbb{R}_{\mathbb{C}}$  such that  $1 \leq_{\mathbb{R}} \alpha$ 
  - a. if  $\alpha <_{\mathbb{R}} \beta$  then  $\alpha <_{\mathbb{R}} \beta^n$
  - b. if  $\alpha \leq_{\mathbb{R}} \beta$  then  $\alpha \leq_{\mathbb{R}} \beta^n$
  - c. if  $n, m \in \mathbb{N}$  with  $n \leq m$  we have  $\alpha^n \leq \alpha^m$

#### Proof.

1. This follows trivially from the definition of  $\mathbb{R}_{\mathbb{C}+}$
2. This follows trivially from the definition of  $\mathbb{R}_{\mathbb{C}+}$
3. If  $\alpha, \beta \in \mathbb{R}_{\mathbb{C}}$  then by 9.14 there exists  $\alpha', \beta' \in \mathbb{R}$  so that  $\alpha = i_{\mathbb{R}}(\alpha')$  and  $\beta = i_{\mathbb{R}}(\beta')$  using 8.39 (3) we have then for  $\alpha', \beta' \in \mathbb{R}$  the following excluding cases
  - $\alpha' < \beta'$ . then using 9.25 we have  $\alpha = i_{\mathbb{R}}(\alpha') <_{\mathbb{R}} i_{\mathbb{R}}(\beta') = \beta$
  - $\beta' < \alpha'$ . then using 9.25 we have  $\alpha = i_{\mathbb{R}}(\beta') <_{\mathbb{R}} i_{\mathbb{R}}(\alpha') = \beta$
  - $\alpha' = \beta'$ . then using 9.14 we have  $\alpha = i_{\mathbb{R}}(\alpha') = i_{\mathbb{R}}(\beta') = \beta$

4. If  $\alpha, \beta \in \mathbb{R}_{\mathbb{C}}$  with  $\alpha <_{\mathbb{R}} \beta$  then by 9.14 there exists  $\alpha', \beta' \in \mathbb{R}$  so that  $\alpha = i_{\mathbb{R}}(\alpha')$ ,  $\beta = i_{\mathbb{R}}(\beta')$ , and using 9.25 we have  $\alpha' < \beta'$ . Using 8.39 we have  $-\beta' < -\alpha' \xrightarrow[9.25]{3.35} i_{\mathbb{R}}(-\beta') <_{\mathbb{R}} i_{\mathbb{R}}(-\alpha') \xrightarrow[3.35]{} -i_{\mathbb{R}}(\beta') <_{\mathbb{R}} -i_{\mathbb{R}}(\alpha') \Rightarrow -\beta <_{\mathbb{R}} \alpha$
5. If  $\alpha, \beta, \gamma \in \mathbb{R}_{\mathbb{C}}$  then by 9.14 there exists  $\alpha', \beta', \gamma' \in \mathbb{R}$  so that  $\alpha = i_{\mathbb{R}}(\alpha')$ ,  $\beta = i_{\mathbb{R}}(\beta')$ ,  $\gamma = i_{\mathbb{R}}(\gamma')$ . As  $\alpha <_{\mathbb{R}} \beta$  we have  $i_{\mathbb{R}}(\alpha') <_{\mathbb{R}} i_{\mathbb{R}}(\beta')$  and thus by 9.25  $\alpha' < \beta'$ . Using then 8.39 (5) we have that  $\alpha' + \gamma' < \beta' + \gamma'$  so that using the fact that  $i_{\mathbb{R}}$  is a field isomorphism we have  $\alpha + \gamma = i_{\mathbb{R}}(\alpha') + i_{\mathbb{R}}(\beta') = i_{\mathbb{R}}(\alpha' + \gamma') <_{\mathbb{R}} i_{\mathbb{R}}(\beta' + \gamma') = i_{\mathbb{R}}(\beta') + i_{\mathbb{R}}(\gamma') = \beta + \gamma$
6. If  $\alpha, \beta \in \mathbb{R}_{\mathbb{C}}$  with  $\alpha <_{\mathbb{R}} \beta$ ,  $\lambda \in \mathbb{R}_{\mathbb{C}+}$  then by 9.14 there exists  $\alpha', \beta', \lambda' \in \mathbb{R}$  so that  $\alpha = i_{\mathbb{R}}(\alpha')$ ,  $\beta = i_{\mathbb{R}}(\beta')$ ,  $\lambda = i_{\mathbb{R}}(\lambda')$ . As  $\alpha <_{\mathbb{R}} \beta$  we have  $i_{\mathbb{R}}(\alpha') <_{\mathbb{R}} i_{\mathbb{R}}(\beta')$  and thus by 9.25  $\alpha' < \beta'$ , further as  $0 = i_{\mathbb{R}}(0)$  we have also  $0 < \lambda'$ . Using 8.39 (6) we have then  $\alpha' \cdot \lambda' < \beta' \cdot \lambda'$  and hence  $\alpha \cdot \lambda = i_{\mathbb{R}}(\alpha') \cdot i_{\mathbb{R}}(\lambda') = i_{\mathbb{R}}(\alpha' \cdot \lambda') <_{\mathbb{R}} i_{\mathbb{R}}(\beta' \cdot \lambda') = i_{\mathbb{R}}(\beta') \cdot i_{\mathbb{R}}(\lambda') = \beta \cdot \lambda$ .
7. If  $\alpha, \beta, \gamma, \delta \in \mathbb{R}_{\mathbb{C}+}$  with  $\alpha < \beta \wedge \gamma < \delta$  then by (6) we have  $\alpha \cdot \gamma < \beta \cdot \gamma$  and  $\beta \cdot \gamma < \beta \cdot \delta$  so that  $\alpha \cdot \gamma < \beta \cdot \delta$
8. If  $\alpha, \beta \in \mathbb{R}_{\mathbb{C}}$  with  $\alpha <_{\mathbb{R}} \beta$ ,  $\lambda \in \mathbb{R}_{\mathbb{C}-}$  then by 9.14 there exists  $\alpha', \beta', \lambda' \in \mathbb{R}$  so that  $\alpha = i_{\mathbb{R}}(\alpha')$ ,  $\beta = i_{\mathbb{R}}(\beta')$ ,  $\lambda = i_{\mathbb{R}}(\lambda')$ . As  $\alpha <_{\mathbb{R}} \beta$  we have  $i_{\mathbb{R}}(\alpha') <_{\mathbb{R}} i_{\mathbb{R}}(\beta')$  and thus by 9.25  $\alpha' < \beta'$ , further as  $0 = i_{\mathbb{R}}(0)$  we have also  $\lambda' < 0$ . Using 8.39 (7) we have then  $\beta' \cdot \lambda' < \alpha' \cdot \lambda'$  and hence  $\beta \cdot \lambda = i_{\mathbb{R}}(\beta') \cdot i_{\mathbb{R}}(\lambda') = i_{\mathbb{R}}(\beta' \cdot \lambda') <_{\mathbb{R}} i_{\mathbb{R}}(\alpha' \cdot \lambda') = i_{\mathbb{R}}(\alpha') \cdot i_{\mathbb{R}}(\lambda') = \alpha \cdot \lambda$ .
9. If  $x \in \mathbb{R}_{\mathbb{C}}$  with  $0 <_{\mathbb{R}} x$  then there exists a  $x' \in \mathbb{R}$  such that  $i_{\mathbb{R}}(x') = x$ , using 9.25 and  $0 = i_{\mathbb{R}}(0)$  we have  $0 < x'$ . Using 8.39 (8) we have then that  $0 < x'^{-1} \Rightarrow 0 = i_{\mathbb{R}}(0) <_{\mathbb{R}} i_{\mathbb{R}}(x'^{-1}) \xrightarrow[3.35]{} (i_{\mathbb{R}}(x'))^{-1} = x^{-1}$
10. If  $x, y \in \mathbb{R}_{\mathbb{C}}$  with  $0 <_{\mathbb{R}} x <_{\mathbb{R}} y$  then using 9.14 and 9.25 there exists  $x', y' \in \mathbb{R}$  such that  $0 < x' < y'$  hence using 8.39 we have that  $y'^{-1} < x'^{-1}$  and thus  $y^{-1} = (i_{\mathbb{R}}(y'))^{-1} \xrightarrow[3.35]{} i_{\mathbb{R}}(y'^{-1}) <_{\mathbb{R}} i_{\mathbb{R}}(x'^{-1}) \xrightarrow[3.35]{} (i_{\mathbb{R}}(x'))^{-1} = x^{-1}$
11. If  $x \in \mathbb{R}$  then using 9.14 there exists a  $x' \in \mathbb{R}$  such that  $x = i_{\mathbb{R}}(x')$ . Using 8.39 (10) we have  $0 \leq x'^2$  and thus  $0 = i_{\mathbb{R}}(0) \leq i_{\mathbb{R}}(x'^2) = i_{\mathbb{R}}(x') \cdot i_{\mathbb{R}}(x') = x \cdot x = x^2$ . Further if  $x \neq 0$  then as  $i_{\mathbb{R}}$  is a bijection we have that  $x' \neq 0$  and thus by 8.39 (10) we have  $0 < x'^2$  so that  $0 = i_{\mathbb{R}}(0) <_{\mathbb{R}} i_{\mathbb{R}}(x'^2) = i_{\mathbb{R}}(x') \cdot i_{\mathbb{R}}(x') = x \cdot x = x^2$
12. Let  $x, y \in \mathbb{R}_{\mathbb{C}+} \cup \{0\}$  are such that  $x <_{\mathbb{R}} y$  we have using 9.14 and 9.25 that  $\exists x', y' \in \mathbb{R}$  such that  $x = i_{\mathbb{R}}(x')$ ,  $y = i_{\mathbb{R}}(y')$  and  $x' < y'$ . Using 8.39 we have then  $x'^2 < y'^2$  so that  $x^2 = i_{\mathbb{R}}(x') \cdot i_{\mathbb{R}}(x') = i_{\mathbb{R}}(x' \cdot x') = i_{\mathbb{R}}(x'^2) <_{\mathbb{R}} i_{\mathbb{R}}(y'^2) = i_{\mathbb{R}}(y' \cdot y') = i_{\mathbb{R}}(y') \cdot i_{\mathbb{R}}(y') = y^2$
13. If  $\alpha \in \mathbb{R}_{\mathbb{C}+}$  then  $0 <_{\mathbb{R}} \alpha$  we proceed now by induction on  $n$  to prove that  $\forall n \in \mathbb{N}$  we have  $0 < \alpha^n$ . So let  $X = \{n \in \mathbb{N} \mid \text{if } 0 <_{\mathbb{R}} \alpha \text{ then } 0 <_{\mathbb{R}} \alpha^n\}$  then we have:
  - a.  $\alpha^0 = 1 >_{\mathbb{R}} 0 \Rightarrow 0 \in X$
  - b. If  $n \in X$  then if  $0 <_{\mathbb{R}} \alpha$  we have  $0 <_{\mathbb{R}} \alpha^n$ . Now  $\alpha^{n+1} = \alpha^n \cdot \alpha \xrightarrow[0 < \alpha^n, \alpha \in \mathbb{R}_+ \text{ and (6)}]{} 0 <_{\mathbb{R}} \alpha^n \cdot \alpha = \alpha^{n+1} \Rightarrow n+1 \in X$

Using mathematical induction (see 4.10) we have  $X = \mathbb{N}$  proving our assertion.

14. Let  $\alpha \in \mathbb{R}_{\mathbb{C}+}$  with  $0 <_{\mathbb{R}} \alpha <_{\mathbb{R}} 1$  then if  $n \in \mathbb{N}$  we have by (12) already  $0 <_{\mathbb{R}} \alpha^n$ . Now to prove  $\alpha^n <_{\mathbb{R}} \alpha$  take  $S = \{n \in \mathbb{N} \mid \text{if } \alpha <_{\mathbb{R}} 1 \text{ then } \alpha^n <_{\mathbb{R}} \alpha\}$  then we have
  - a. if  $n = 1$  then  $\alpha^1 = \alpha <_{\mathbb{R}} 1$  so that  $1 \in S$
  - b. if  $n \in S$  then  $\alpha^{n+1} = \alpha^n \cdot \alpha <_{\mathbb{R}} \alpha$  [as  $n \in S \Rightarrow \alpha^n <_{\mathbb{R}} 1 \xrightarrow[0 < \alpha \text{ and (6)}]{} \alpha^n \cdot \alpha <_{\mathbb{R}} 1 \cdot \alpha = \alpha$ ]
15.
  - a. We prove this by induction so let  $\mathcal{S} = \{n \in \mathbb{N} \mid \alpha <_{\mathbb{R}} \beta^n\}$  then we have
    - 1  $\in \mathcal{S}$ .** this follows from  $\alpha <_{\mathbb{R}} \beta = \beta^1$
    - $n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$ .** as  $n \in \mathcal{S}$  we have  $\alpha <_{\mathbb{R}} \beta^n$  we have by (6) that  $\alpha \cdot \beta <_{\mathbb{R}} \beta \cdot \beta^n = \beta^{n+1}$  and  $1 \leq_{\mathbb{R}} \alpha \Rightarrow \beta \leq_{\mathbb{R}} \alpha \cdot \beta \xrightarrow[\alpha < \beta]{} \alpha <_{\mathbb{R}} \alpha \cdot \beta <_{\mathbb{R}} \beta^{n+1}$  proving that  $n+1 \in \mathcal{S}$

b. We prove this by induction so let  $\mathcal{S} = \{n \in \mathbb{N} \mid \alpha \leq_{\mathbb{R}} \beta^n\}$  then we have

**1  $\in \mathcal{S}$ .** this follows from  $\alpha \leq_{\mathbb{R}} \beta = \beta^1$

**$n \in \mathcal{S} \Rightarrow n + 1 \in \mathcal{S}$ .** as  $n \in \mathcal{S}$  we have  $\alpha \leq_{\mathbb{R}} \beta^n$  we have by (6) that  $\alpha \cdot \beta \leq_{\mathbb{R}} \beta \cdot \beta^n = \beta^{n+1}$  and  $1 \leq_{\mathbb{R}} \alpha \Rightarrow \beta \leq_{\mathbb{R}} \alpha \cdot \beta \Rightarrow \alpha \leq_{\mathbb{R}} \alpha \cdot \beta \leq_{\mathbb{R}} \beta^{n+1}$  proving that  $n + 1 \in \mathcal{S}$

c. We prove this by induction so let  $n \in \mathbb{N}$  and  $\mathcal{S}_n = \{k \in \mathbb{N} \mid \alpha^n \leq_{\mathbb{R}} \alpha^{n+k}\}$  then we have

**$k = 0$ .** then  $\alpha^n = \alpha^{n+0}$  proving that  $0 \in \mathcal{S}_n$

**$k \in \mathcal{S}_n \Rightarrow k + 1 \in \mathcal{S}_n$ .** then  $\alpha^n \leq_{\mathbb{R}} \alpha^{n+k} \Rightarrow \alpha^{n+1} \leq_{\mathbb{R}} \alpha^{n+k+1}$ , as  $1 \leq_{\mathbb{R}} \alpha$  we have  $\alpha^n \leq_{\mathbb{R}} \alpha^{n+1}$  we have  $\alpha^n \leq_{\mathbb{R}} \alpha^{n+k+1}$  proving that  $k + 1 \in \mathcal{S}_n$

Using mathematical induction we have then  $\mathcal{S}_n = \mathbb{N}$ , so if  $n, m \in \mathbb{N}$  with  $n \leq m$  we have that  $m - n \in \mathbb{N}$  and thus  $\alpha^n \leq_{\mathbb{R}} \alpha^{n+(m-n)} = \alpha^m$   $\square$

**Theorem 9.42.** If  $\alpha, \beta \in \mathbb{R}_{\mathbb{C}}$  with  $0 \leq \alpha, \beta$  then from  $\alpha + \beta = 0$  we have  $\alpha = \beta = 0$

**Proof.** We have to consider the following cases

**$0 < \alpha$ .** then  $\beta = 0 + \beta < \alpha + \beta = 0$  giving the contradiction  $\beta < 0$

**$0 < \beta$ .** then  $\alpha = 0 + \alpha < \beta + \alpha = 0$  giving the contradiction  $\alpha < 0$

**$\alpha = 0 = \beta$ .** this is what we have to prove  $\square$

**Theorem 9.43. ( $\mathbb{R}_{\mathbb{C}}$  is conditional complete)**  $\langle \mathbb{R}_{\mathbb{C}}, \leq_{\mathbb{R}} \rangle$  is conditional complete (see 2.179) Using the definition of conditional completeness this means that  $\forall S \subseteq \mathbb{R} \vdash S \neq \emptyset$  for which there exists  $a, b \in \mathbb{R}$  such that  $\forall s \in S$  we have  $s \leq b$  (existence of a upper bound) we have the existence of  $\sup(S)$  (a lowest upper bound). In other words: any nonempty set in  $\mathbb{R}$  with a upper bound has a lowest upper bound. Using 2.180 we see that  $\forall S \subseteq \mathbb{R} \vdash S \neq \emptyset$  for which there exists  $a, b \in \mathbb{R}$  such that  $\forall s \in S$  we have  $b \leq s$  (existence of a lower bound) there exists a  $\inf(S)$  (in other words every non empty set in  $\mathbb{R}$  with a lower bound has greatest lower bound).

**Proof.** As by 9.28  $i_{\mathbb{R}}: \mathbb{R} \rightarrow \mathbb{R}_{\mathbb{C}}$  is a order preserving isomorphism,  $\langle R, \leq \rangle$  is conditionally complete we have by 2.182 that  $\langle \mathbb{R}_{\mathbb{C}}, \leq_{\mathbb{R}} \rangle$  is conditionally complete.  $\square$

**Theorem 9.44.** Let  $S \subseteq \mathbb{R}_{\mathbb{C}}$  then we have for  $-S = \{-s, s \in S\}$  that

1. If  $\sup(S)$  exists then  $\inf(-S)$  exists and  $\inf(-S) = -\sup(S)$
2. If  $\inf(S)$  exists then  $\sup(-S)$  exists and  $\sup(-S) = -\inf(S)$

**Proof.**

1. As  $\sup(S)$  exists we have  $\forall s \in -S$  that  $-s \in S$  so that  $-s \leq \sup(S)$  or  $-\sup(S) \leq s$  so  $-S$  is bounded below by  $-\sup(S)$  proving by the conditional completeness of  $\mathbb{R}$  [see 8.43] that

$$\inf(-S) \text{ exists and } -\sup(S) \leq \inf(-S). \quad (9.1)$$

Assume now that  $-\sup(S) < \inf(-S)$  then  $-\inf(-S) < \sup(S)$  so there exist a  $s \in S$  such that  $-\inf(-S) < s \leq \sup(S)$  so that  $-s < \inf(-S)$  which as  $-s \in -S$  gives  $-s < \inf(-S) \leq -s$  a contradiction. So we must have that  $\inf(-S) \leq -\sup(S)$  which together with (9.1) proves that

$$\inf(-S) = -\sup(S)$$

2. As  $\inf(S)$  exists we have  $\forall s \in -S$  that  $-s \in S$  so that  $\inf(S) \leq -s$  or  $s \leq -\inf(S)$  so  $-S$  is bounded above by  $-\inf(S)$  proving by the conditional completeness of  $\mathbb{R}$  [see 8.43] that

$$\sup(-S) \text{ exists and } \sup(-S) \leq -\inf(S) \quad (9.2)$$

Assume now that  $\sup(-S) < -\inf(S)$  then  $\inf(S) < -\sup(-S)$  so there exist a  $s \in S$  such that  $\inf(S) \leq s < -\sup(-S)$  so that  $\sup(-S) < -s$  which as  $-s \in -S$  gives  $-s \leq \sup(-S) < -s$  a contradiction. So we must have that  $-\inf(S) \leq \sup(-S)$  which together with (9.2) proves that

$$\sup(-S) = -\inf(S) \quad \square$$

**Theorem 9.45.** If  $S \subseteq \mathbb{R}_{\mathbb{C}}$  has a supremum and  $\alpha \in \mathbb{R}_{\mathbb{C}}$  with  $\alpha \geq 0$  then  $\alpha \cdot S = \{\alpha \cdot s \mid s \in S\}$  has a supremum equal to  $\alpha \cdot \sup(S)$

**Proof.** If  $\alpha = 0$  then  $\alpha \cdot S = \{0\}$  with  $\sup(\alpha \cdot S) = \sup(\{0\}) = 0 = 0 \cdot \sup(S)$  so we only have to prove the case that  $\alpha > 0$ . By the definition of  $\sup(S)$  we have that  $\forall s \in S$  that  $s \leq_{\mathbb{R}} \sup(S)$  so if  $y \in \alpha \cdot S$  then  $y = \alpha \cdot s \leq_{\mathbb{R}} \alpha \cdot \sup(S)$  [as  $\alpha >_{\mathbb{R}} 0$ ] proving that  $\alpha \cdot \sup(S)$  is an upper bound for  $\alpha \cdot S$ . By the above theorem  $\alpha \cdot S$  has a  $\sup(\alpha \cdot S)$  with by definition  $\sup(\alpha \cdot S) \leq_{\mathbb{R}} \alpha \cdot \sup(S)$ . If now  $\sup(\alpha \cdot S) <_{\mathbb{R}} \alpha \cdot \sup(S)$  then  $\frac{1}{\alpha} \cdot \sup(\alpha \cdot S) <_{\mathbb{R}} \sup(S)$  and there exists a  $s \in S$  with  $\frac{1}{\alpha} \cdot \sup(\alpha \cdot S) <_{\mathbb{R}} s \leq_{\mathbb{R}} \sup(S)$  giving  $\sup(\alpha \cdot S) <_{\mathbb{R}} \alpha \cdot s \leq_{\mathbb{R}} \alpha \cdot \sup(S)$  which as  $\alpha \cdot s \in \alpha \cdot S$  contradicts the fact that  $\sup(\alpha \cdot S)$  is the supremum of  $\alpha \cdot S$  so we must have  $\sup(\alpha \cdot S) = \alpha \cdot \sup(S)$ .  $\square$

**Theorem 9.46.** If  $S, T \subseteq \mathbb{R}_{\mathbb{C}}$  have a supremum then  $S + T = \{s + t \mid s \in S \wedge t \in T\}$  has a supremum with  $\sup(S + T) = \sup(S) + \sup(T)$

**Proof.** Let  $r \in S + T$  then  $\exists s \in S, \exists t \in T$  such that  $r = s + t$ . Using the fact that the supremum is an upper bound we have that  $r = s + t \leq \sup(S) + \sup(T)$ . Using the definition of the supremum we have then that

$$\sup(S + T) \leq \sup(S) + \sup(T) \quad (9.3)$$

Assume now that  $\sup(S + T) < \sup(S) + \sup(T)$  then if we take  $\varepsilon = \sup(S) + \sup(T) - \sup(S + T)$  we have that  $0 < \varepsilon$ . So  $\sup(S) - \frac{\varepsilon}{2} < \sup(S)$  and  $\sup(T) - \frac{\varepsilon}{2} < \sup(T)$  and by the definition of the supremum there exists a  $s \in S$  and a  $t \in T$  such that  $\sup(S) - \frac{\varepsilon}{2} < s \leq \sup(S) + \frac{\varepsilon}{2}$  and  $\sup(T) - \frac{\varepsilon}{2} < t \leq \sup(T)$ . So  $s + t > \sup(S) - \frac{\varepsilon}{2} + \sup(T) - \frac{\varepsilon}{2} = \sup(S) + \sup(T) - \varepsilon = \sup(S) + \sup(T) - \sup(S) - \sup(T) + \sup(S + T) = \sup(S + T)$  or  $s + t > \sup(S + T)$  which conflict the fact that the supremum is an upper bound. Hence  $\sup(S) + \sup(T) \leq \sup(S + T)$  which together with (9.3) proves that

$$\sup(S + T) = \sup(S) + \sup(T) \quad \square$$

**Corollary 9.47.** If  $S \subseteq \mathbb{R}_{\mathbb{C}}$  has a supremum and  $x \in \mathbb{R}_{\mathbb{C}}$  then  $S + x = \{s + x \mid s \in S\}$  has a supremum with  $\sup(S + x) = \sup(S) + x$

**Proof.** Define  $T = \{x\}$  then  $\sup(T) = x$  exists and  $S + T = \{s + t \mid s \in S \wedge t \in T\} = \{s + x \mid s \in S\} = S + x$ . Hence using 9.46 we have that  $\sup(S + x)$  exists and  $\sup(S + x) = \sup(S) + \sup(T) = \sup(S) + x$ .  $\square$

**Theorem 9.48.** If  $S, T \subseteq \mathbb{R}_{\mathbb{C}}$  have a infimum then  $S + T = \{s + t \mid s \in S \wedge t \in T\}$  has a infimum with  $\inf(S + T) = \inf(S) + \inf(T)$

**Proof.** Let  $r \in S + T$  then  $\exists s \in S, \exists t \in T$  such that  $r = s + t$ . Using the fact that the infimum is a lower bound we have that  $\inf(S) + \inf(T) \leq s + t = r$ . Using the definition of the infimum we have then that

$$\inf(S) + \inf(T) \leq \inf(S + T) \quad (9.4)$$

Assume now that  $\inf(S) + \inf(T) < \inf(S + T)$  then if we take  $\varepsilon = \inf(S + T) - \inf(S) - \inf(T)$  we have that  $0 < \varepsilon$ . So  $\inf(S) < \inf(S) + \frac{\varepsilon}{2}$  and  $\inf(T) < \inf(T) + \frac{\varepsilon}{2}$  and by the definition of the infimum there exists a  $s \in S$  and a  $t \in T$  such that  $\inf(S) \leq s < \inf(S) + \frac{\varepsilon}{2}$  and  $\inf(T) \leq t < \inf(T) + \frac{\varepsilon}{2}$ . So  $s + t < \inf(S) + \inf(T) + \varepsilon = \inf(S + t) + \inf(T) + \inf(S + T) - \inf(S) - \inf(T) = \inf(S + T)$  or  $s + t < \inf(S + T)$  which conflict the fact that the infimum is a lower bound. Hence  $\inf(S + T) \leq \inf(S) + \inf(T)$  which together with (9.4) proves that

$$\inf(S + T) = \inf(S) + \inf(T) \quad \square$$

**Corollary 9.49.** If  $S \subseteq \mathbb{R}_{\mathbb{C}}$  has a infimum and  $x \in \mathbb{R}_{\mathbb{C}}$  then  $S + x = \{s + x \mid s \in S\}$  has a infimum with  $\inf(S + x) = \inf(S) + x$

**Proof.** Define  $T = \{x\}$  then  $\inf(T) = x$  exists and  $S + T = \{s + t \mid s \in S \wedge t \in T\} = \{s + x \mid s \in S\} = S + x$ . Hence using 9.48 we have that  $\inf(S + x)$  exists and  $\inf(S + x) = \inf(S) + \inf(T) = \inf(S) + x$ .  $\square$

**Theorem 9.50.**  $\langle \mathbb{Q}_{\mathbb{C}}, \leq \rangle$  is not conditionally complete

**Proof.** We prove this by contradiction, so assume that  $\mathbb{Q}_{\mathbb{C}}$  is conditionally complete then as  $\mathbb{Q}$  is order isomorphic with  $\mathbb{Q}$  by 9.37 we have by 2.182 that  $\mathbb{Q}$  is conditionally complete contradiction the fact that  $\mathbb{Q}$  is not conditionally complete by 7.38.  $\square$

**Corollary 9.51.**  $\mathbb{Q}_{\mathbb{C}} \subset \mathbb{R}_{\mathbb{C}}$  [there exists a  $r \in \mathbb{R}_{\mathbb{C}}$  such that  $r \notin \mathbb{Q}_{\mathbb{C}}$ ]. In other words we have that  $\mathbb{R}_{\mathbb{C}} \setminus \mathbb{Q}_{\mathbb{C}} \neq \emptyset$ .  $\mathbb{R}_{\mathbb{C}} \setminus \mathbb{Q}_{\mathbb{C}}$  is called the set of irrational numbers.

**Proof.** As  $\langle \mathbb{Q}_{\mathbb{C}}, \leq \rangle$  is not conditional complete (see previous theorem) there exists a non-empty set  $A \subseteq \mathbb{Q}_{\mathbb{C}}$  with an upper bound  $u$  so that  $\sup(A)$  does not exist in  $\mathbb{Q}_{\mathbb{C}}$ . Because  $\mathbb{Q}_{\mathbb{C}} \subseteq \mathbb{R}_{\mathbb{C}}$  we have  $\emptyset \neq A \subseteq \mathbb{R}_{\mathbb{C}}$  and that  $A$  has the upper bound  $u \in \mathbb{Q}_{\mathbb{C}} \subseteq \mathbb{R}$ . As  $\mathbb{R}_{\mathbb{C}}$  is conditionally complete (see 9.43) there exists a lowest upper bound  $s = \sup(A)$ , now if  $s \in \mathbb{Q}_{\mathbb{R}}$  it would be an upper bound of  $A$  and if  $b \in \mathbb{Q}_{\mathbb{R}}$  of  $A$  is another upper bound of  $A$  it is also an upper bound of  $A$  in  $\mathbb{R}_{\mathbb{C}}$  and thus  $s \leq b$  so  $s$  would be the supremum of  $A$  in  $\mathbb{Q}_{\mathbb{R}}$  contradicting the fact that  $\sup(A)$  does not exist in  $\mathbb{Q}_{\mathbb{R}}$ . So  $s \notin \mathbb{Q}_{\mathbb{R}}$ .  $\square$

**Theorem 9.52.** We have the following properties concerning  $\mathbb{N}_{\mathbb{C}}$  and  $\mathbb{Z}_{\mathbb{C}}$  (here  $-\mathbb{N}_{\mathbb{C}} = \{n \vdash -n \in \mathbb{N}_{\mathbb{R}}\} \subseteq \mathbb{Z}_{\mathbb{R}}$ )

1.  $\mathbb{Z}_{\mathbb{C}} = \mathbb{N}_{\mathbb{C}} \cup (-\mathbb{N}_{\mathbb{C}})$  and  $\mathbb{N}_{\mathbb{C}} \cap (-\mathbb{N}_{\mathbb{C}}) = \{0\}$
2. If  $n \in \mathbb{Z}_{\mathbb{C}}$  and  $0 \leq n \Rightarrow n \in \mathbb{N}_{\mathbb{C}}$  (so  $\mathbb{N}_{\mathbb{C}}$  is indeed the set of positive integers)
3.  $\langle \mathbb{N}_{\mathbb{C}}, + \rangle$  is a semi-group of  $\langle \mathbb{Z}_{\mathbb{C}}, + \rangle$
4.  $\langle \mathbb{N}_{\mathbb{C}}, \cdot \rangle$  is a sub-semi-group of  $\langle \mathbb{Z}_{\mathbb{C}}, \cdot \rangle$

**Proof.**

1. As  $\mathbb{N}_{\mathbb{C}}, -\mathbb{N}_{\mathbb{C}} \subseteq \mathbb{Z}_{\mathbb{C}}$  we have  $\mathbb{N}_{\mathbb{C}} \cup (-\mathbb{N}_{\mathbb{C}}) \subseteq \mathbb{Z}_{\mathbb{C}}$ . If  $x \in \mathbb{Z}_{\mathbb{C}}$  then  $\exists x' \in \mathbb{Z}_{\mathbb{R}}$  such that  $x = (x', 0)$  as  $\mathbb{Z}_{\mathbb{R}} = \mathbb{N}_{\mathbb{R}} \cup (-\mathbb{N}_{\mathbb{R}})$  we have that either  $x' \in \mathbb{N}_{\mathbb{R}}$  or  $-x' \in \mathbb{N}_{\mathbb{R}}$  hence  $(x', 0) \in \mathbb{N}_{\mathbb{C}}$  or  $(-x', 0) \in \mathbb{N}_{\mathbb{C}}$  proving that  $x \in \mathbb{N}_{\mathbb{C}} \cup (-\mathbb{N}_{\mathbb{C}})$
2. If  $n = (n', 0) \in \mathbb{Z}_{\mathbb{C}}$  and  $0 \leq_{\mathbb{R}} n = (n', 0)$  then  $0 \leq n' = n' \in \mathbb{N}_{\mathbb{R}} \Rightarrow n = (n', 0) \in \mathbb{N}_{\mathbb{C}}$
3. If  $x, y \in \mathbb{N}_{\mathbb{C}}$  then we have  $x = (x', 0), y = (y', 0)$  and  $x', y' \in \mathbb{N}_{\mathbb{R}}$  so that
  - a.  $x + y = (x', 0) + (y', 0) = (x' + y', 0) \in \mathbb{N}_{\mathbb{C}}$  as  $x' + y' \in \mathbb{N}_{\mathbb{R}}$  (for  $\mathbb{N}_{\mathbb{R}}$  is a semi-group)
  - b.  $0 = (0, 0) \in \mathbb{N}_{\mathbb{C}}$
4. If  $x, y \in \mathbb{N}_{\mathbb{C}}$  then we have  $x = (x', 0), y = (y', 0)$  and  $x', y' \in \mathbb{N}_{\mathbb{R}}$  so that
  - a.  $x \cdot y = (x' \cdot y' - 0 \cdot 0, x' \cdot 0 + 0 \cdot y') = (x' \cdot y', 0) \in \mathbb{N}_{\mathbb{C}}$  (as  $x' \cdot y' \in \mathbb{N}_{\mathbb{C}}$  which is a semi-group)
  - b.  $1 = (1, 0) \in \mathbb{N}_{\mathbb{C}}$

**Lemma 9.53.** If  $n \in \mathbb{N}_{\mathbb{C}}$  and  $0 < n \Rightarrow 1 \leq_{\mathbb{R}} n$ .

**Proof.** As  $n = (n', 0)$  with  $n' \in \mathbb{N}_{\mathbb{R}}$  we have from  $0 \leq_{\mathbb{R}} n$  that  $0 < n'$  hence using 8.60 we have  $1 \leq n'$  proving that  $1 = (1, 0) \leq (n', 0) = n$ .  $\square$

**Corollary 9.54.** If  $n, m \in \mathbb{Z}_{\mathbb{C}}$  with  $n < m$  then  $n + 1 \leq m$

**Proof.** If  $n < m$  then  $0 < m - n \in \mathbb{Z}_{\mathbb{C}}$  [using 9.52 (2)]. Using the previous lemma 9.53 we have then that  $1 \leq m - n \Rightarrow n + 1 \leq m$ .  $\square$

**Theorem 9.55. (Archimedean Property)** If  $x \in \mathbb{R}_{\mathbb{C}}$  with  $0 <_{\mathbb{R}} x$  then  $\exists n \in \mathbb{N}_{\mathbb{R}}$  such that  $y <_{\mathbb{R}} n \cdot x$

**Proof.** As  $x, y \in \mathbb{R}_{\mathbb{C}}$  there exists  $x', y' \in \mathbb{R}$  such that  $x = (x', 0), y = (y', 0)$ , using the Archimedean property of  $\mathbb{R}$  (see 8.61) there exists a  $n' \in \mathbb{N}_{\mathbb{R}}$  such that  $y' < n' \cdot x'$  hence  $y = (y', 0) <_{\mathbb{R}} (n' \cdot x', 0) = (n', 0) \cdot (x', 0) = n \cdot x$  where  $n = (n', 0) \in \mathbb{N}_{\mathbb{C}}$ .  $\square$

**Theorem 9.56.** Let  $x \in \mathbb{R}_{\mathbb{C}}$  then the following holds

1.  $\exists n \in \mathbb{N}_{\mathbb{C}} \vdash x <_{\mathbb{R}} n$
2.  $\exists n \in \mathbb{Z}_{\mathbb{C}} \vdash n \leqslant_{\mathbb{R}} x <_{\mathbb{R}} n + 1$
3. If  $0 <_{\mathbb{R}} x$  then  $\exists n \in \mathbb{N}_{\mathbb{R}} \setminus \{0\}$  such that  $\frac{1}{n} <_{\mathbb{R}} x$
4. For  $y \in \mathbb{R}_{\mathbb{C}}$  with  $x < y$  we have that  $n \in \mathbb{N}_{\mathbb{R}} \setminus \{0\}$  such that  $x < x + \frac{1}{n} < y$
5. If  $0 \leqslant_{\mathbb{R}} x$  then  $\exists n \in \mathbb{N}_{\mathbb{R}} \setminus \{0\}$  such that  $n - 1 \leqslant_{\mathbb{R}} x <_{\mathbb{R}} n$
6.  $\exists n \in \mathbb{Z}_{\mathbb{C}} \vdash n <_{\mathbb{R}} x \leqslant_{\mathbb{R}} n + 1$
7.  $\exists n \in \mathbb{Z}_{\mathbb{C}} \vdash n - 1 <_{\mathbb{R}} x \leqslant_{\mathbb{R}} n$

**Proof.** As  $x \in \mathbb{R}_{\mathbb{C}}$  we have that  $x = (x', 0)$  where  $x' \in \mathbb{R}$

1. As  $0 <_{\mathbb{R}} 1 \in \mathbb{R}_{\mathbb{C}}$  we have by the Archimedean property (see 9.55) that  $\exists n \in \mathbb{N}_{\mathbb{C}}$  such that  $x <_{\mathbb{R}} 1 \cdot n = n$
2. There exists by 8.62 (2) a  $n' \in \mathbb{Z}_{\mathbb{R}}$  such that  $n' \leqslant x' < n' + 1$  hence we have  $(n', 0) \leqslant_{\mathbb{R}} (x', 0) <_{\mathbb{R}} (n' + 1, 0) = (n', 0) + (1, 0)$  or if we take  $n = (n', 0) \in \mathbb{Z}_{\mathbb{C}}$  that  $n \leqslant_{\mathbb{R}} x <_{\mathbb{R}} n + 1$
3. As  $0 <_{\mathbb{R}} x = (x', 0)$  we must have  $0 < x'$  and thus using 8.62 (3) there exists a  $n' \in \mathbb{N}_{\mathbb{R}} \setminus \{0\}$  such that  $\frac{1}{n'} < x'$  hence  $(\frac{1}{n'}, 0) <_{\mathbb{R}} (x', 0) = x$  if we take  $n = (n', 0) \in \mathbb{N}_{\mathbb{C}} \setminus \{0\}$  then using 9.2 we have that  $\frac{1}{n} = n^{-1} = (n'^{-1}, 0) = (\frac{1}{n'}, 0)$  proving that  $\frac{1}{n} <_{\mathbb{R}} x$
4. As  $x < y$  we have  $0 < y - x$  hence using (3) there exists a  $n \in \mathbb{N}_{\mathbb{R}} \setminus \{0\}$  such that  $\frac{1}{n} < y - x \Rightarrow x + \frac{1}{n} < n$
5. As  $0 \leqslant_{\mathbb{R}} x$  we have  $0 \leqslant x'$  we have by 8.62 that there exists  $n' \in \mathbb{N}_{\mathbb{R}} \setminus \{0\}$  so that  $n' - a \leqslant x' < n'$  hence we have  $(n' - 1, 0) \leqslant (x', 0) \leqslant (n', 0)$  so  $n - 1 \leqslant_{\mathbb{R}} x <_{\mathbb{R}} n$  if we define  $n = (n', 0) \in \mathbb{N}_{\mathbb{C}}$
6. There exists by 8.62R  $\subseteq Q$  then  $R \cap Q = R \neq \emptyset$ . Combining this with 19.66 gives (5) a  $n' \in \mathbb{Z}_{\mathbb{R}}$  such that  $n' < x' \leqslant n' + 1$  hence we have  $(n', 0) <_{\mathbb{R}} (x', 0) \leqslant_{\mathbb{R}} (n' + 1, 0) = (n', 0) + (1, 0)$  or if we take  $n = (n', 0) \in \mathbb{Z}_{\mathbb{C}}$  that  $n <_{\mathbb{R}} x \leqslant_{\mathbb{R}} n + 1$
7. There exists by 8.62 (6) a  $n' \in \mathbb{Z}_{\mathbb{R}}$  such that  $n' - 1 < x' \leqslant n'$  hence we have  $(n', 0) - (1, 0) = (n' - 1, 0) <_{\mathbb{R}} (x', 0) < \leqslant_{\mathbb{R}} (n', 0) = (n', 0)$  or if we take  $n = (n', 0) \in \mathbb{Z}_{\mathbb{C}}$  that  $n - 1 <_{\mathbb{R}} x \leqslant_{\mathbb{R}} n + 1$   $\square$

**Corollary 9.57.** Let  $x, y \in \mathbb{R}_{\mathbb{C}}$  then we have

1. If there exists a  $c \in \mathbb{R}_{\mathbb{C}}$  with such that  $\forall n \in \mathbb{N}_{\mathbb{C}} \setminus \{0\}$  with  $\frac{1}{n} \leqslant c$  we have  $x \leqslant_{\mathbb{R}} y + \frac{1}{n}$  then  $x \leqslant_{\mathbb{R}} y$
2. If  $\forall n \in \mathbb{N}_{\mathbb{C}} \setminus \{0\}$  we have  $x \leqslant y + \frac{1}{n}$  then  $x \leqslant y$
3. If there exists a  $c \in \mathbb{R}_{\mathbb{C}}$  with such that  $\forall \varepsilon \in \mathbb{R}_{\mathbb{C}}$  with  $0 < \varepsilon \leqslant c$  we have  $x \leqslant_{\mathbb{R}} y + \varepsilon$  then  $x \leqslant_{\mathbb{R}} y$
4. If  $\forall \varepsilon \in \mathbb{R}_{\mathbb{C}}$  with  $0 < \varepsilon$  we have  $x \leqslant_{\mathbb{R}} y + \varepsilon$  then  $x \leqslant_{\mathbb{R}} y$
5. If there exists a  $a \in \mathbb{R}_{\mathbb{C}}$  with  $0 \leqslant_{\mathbb{R}} a$  such that  $\forall \varepsilon \in \mathbb{R}_{\mathbb{C}}$  with  $0 \leqslant_{\mathbb{R}} \varepsilon$  we have  $x \leqslant_{\mathbb{R}} y + \varepsilon \cdot a$  then  $x \leqslant_{\mathbb{R}} y$
6. If  $\forall \varepsilon \in \mathbb{R}_{\mathbb{C}}$  with  $0 <_{\mathbb{R}} \varepsilon$  we have  $x \leqslant (1 + \varepsilon) \cdot y$  then  $x \leqslant_{\mathbb{R}} y$

**Proof.**

1. Assume that  $y <_{\mathbb{R}} x$  then  $0 <_{\mathbb{R}} (x - y)$  then as  $0 <_{\mathbb{R}} c$  we have  $0 <_{\mathbb{R}} \min(x - y, c)$ . Using the above theorem [see 9.56 (3)] there exists a  $n \in \mathbb{N}_{\mathbb{C}} \setminus \{0\}$  such that  $0 <_{\mathbb{R}} \frac{1}{n} <_{\mathbb{R}} c$  and  $0 <_{\mathbb{R}} \frac{1}{n} <_{\mathbb{R}} x - y$ . Hence  $y + \frac{1}{n} <_{\mathbb{R}} x \leqslant_{\mathbb{R}} y + \frac{1}{n}$  leading to the contradiction  $y + \frac{1}{n} <_{\mathbb{R}} y + \frac{1}{n}$ , so we must have  $x \leqslant_{\mathbb{R}} y$ .
2. As  $\forall n \in \mathbb{N}_{\mathbb{C}} \setminus \{0\}$  we have  $x \leqslant_{\mathbb{R}} y + \frac{1}{n}$  we have certainly  $\forall n \in \mathbb{N}_{\mathbb{C}} \setminus \{0\}$  with  $\frac{1}{n} \leqslant_{\mathbb{R}} c$  that  $x \leqslant_{\mathbb{R}} y + \frac{1}{n}$ . Using (1) it follows that  $x \leqslant_{\mathbb{R}} y$ .

3. As  $\forall n \in \mathbb{N}_{\mathbb{C}} \setminus \{0\}$  with  $\frac{1}{n} \leq_{\mathbb{R}} c$  we have  $\frac{1}{n} \in \mathbb{R}_{\mathbb{C}}$  with  $0 < \frac{1}{n} \leq_{\mathbb{R}} c$  it follows from (1) that  $x \leq_{\mathbb{R}} y$
4. As  $\forall n \in \mathbb{N}_{\mathbb{C}} \setminus \{0\}$  we have  $\frac{1}{n} \in \mathbb{R}_{\mathbb{C}}$  with  $0 <_{\mathbb{R}} \frac{1}{n}$  it follows from (1) that  $x \leq_{\mathbb{R}} y$
5. For  $a$  we have two cases to consider:

**$a = 0$ .** Then  $x \leq_{\mathbb{R}} y + \varepsilon \cdot a = y$  proving that  $x \leq_{\mathbb{R}} y$

**$a \neq 0$ .** Then  $0 <_{\mathbb{R}} a$ , given  $\varepsilon \in \mathbb{R}_{\mathbb{C}}$  with  $\varepsilon >_{\mathbb{R}} 0$  construct  $\varepsilon' = \frac{\varepsilon}{a} >_{\mathbb{R}} 0$  then as by the hypothese  $x \leq_{\mathbb{R}} y + \varepsilon' \cdot a$  we have  $x \leq_{\mathbb{R}} y + \varepsilon$ , applying then (4) proves that  $x \leq_{\mathbb{R}} y$ .

6. For  $y$  we have to consider the following possibilities

**$y \leq 0$ .** Then as  $1 \in \mathbb{R}_{\mathbb{C}}$  and  $0 < 1$  we have  $x \leq (1+1) \cdot y$  we have  $x - y \leq y \leq 0$  so that  $x \leq y$

**$0 < y$ .** Then given a  $\varepsilon \in \mathbb{R}_{\mathbb{C}}$  with  $0 < \varepsilon$  we have  $x \leq (1+\varepsilon) \cdot y \Rightarrow \frac{x}{y} < 1 + \varepsilon$  which by (4) means that  $\frac{x}{y} \leq 1$  so that  $x \leq y$   $\square$

**Theorem 9.58. (Density Theorem)** *If  $x, y \in \mathbb{R}_{\mathbb{C}}$  such that  $x < y$  then  $\exists q \in \mathbb{Q}_{\mathbb{C}}$  and  $\exists r \in \mathbb{R}_{\mathbb{C}} \setminus \mathbb{Q}_{\mathbb{C}}$  such that  $x <_{\mathbb{R}} q <_{\mathbb{R}} y$  and  $x <_{\mathbb{R}} r <_{\mathbb{R}} y$*

**Proof.** First as  $x, y \in \mathbb{R}_{\mathbb{C}}$  with  $x < y$  there exists  $x', y' \in \mathbb{R}$  such that  $x = (x', 0), y = (y', 0)$  and  $x' < y'$ . Using then the density of  $\mathbb{Q}$  in  $\mathbb{R}$  (see 8.64) there exists a  $q' \in \mathbb{Q}_{\mathbb{C}}$  and  $r' \in \mathbb{R} \setminus \mathbb{Q}_{\mathbb{C}}$  such that  $x' < q' < y'$  and  $x' < r' < y'$  and thus  $(x', 0) <_{\mathbb{R}} (q', 0) <_{\mathbb{R}} (y', 0)$  and  $(x', 0) <_{\mathbb{R}} (r', 0) <_{\mathbb{R}} (y', 0)$ . Hence if we define  $q = (q', 0) \in \mathbb{R}_{\mathbb{C}}$  and  $r = (r', 0) \in \mathbb{R}_{\mathbb{C}} \setminus \mathbb{Q}_{\mathbb{C}}$  then we have  $x <_{\mathbb{R}} q <_{\mathbb{R}} y$  and  $x <_{\mathbb{R}} r <_{\mathbb{R}} y$   $\square$

To simplify notation as we mix elements of  $\mathbb{N}$  and  $\mathbb{C}$  we introduce the following definition and notation.

**Definition 9.59.** Let  $n \in \mathbb{N}$  and  $x \in \mathbb{R}_{\mathbb{C}}$  then  $n \cdot x = \begin{cases} 0 & \text{if } n=0 \\ x + (n-1) \cdot x & \text{if } n > 1 \end{cases}$ . Hence  $0 \cdot x = 0$  (last 0 is the neutral element in  $\mathbb{R}_{\mathbb{C}}$ ),  $1 \cdot x = x + 0 \cdot x = x + 0 = x$ ,  $2 \cdot x = x + 1 \cdot x = x + x$ , ...

The above operation is consistent with the multiplication in  $\mathbb{R}_{\mathbb{C}}$  as the following theorem proves.

**Theorem 9.60.** Let  $n \in \mathbb{N}$  then  $n \cdot x = i_{\mathbb{R}_{\mathbb{N}}}(n) \cdot x$  hence if 1 is the multiplicative unit in  $\mathbb{R}_{\mathbb{C}}$  we have  $n \cdot 1 = i_{\mathbb{R}_{\mathbb{N}}}(n)$

**Proof.** By 9.35 we have that  $i_{\mathbb{R}_{\mathbb{N}}}(0) = 0 \in \mathbb{R}_{\mathbb{C}}$  and  $i_{\mathbb{R}_{\mathbb{N}}}(1) = 1$ . We prove now the theorem by induction so let  $\mathcal{S} = \{n \in \mathbb{N} | n \cdot x = i_{\mathbb{R}_{\mathbb{N}}}(n) \cdot x\}$  then we have

**$0 \in \mathcal{S}$ .** As  $0 \cdot x = 0 = i_{\mathbb{R}_{\mathbb{N}}}(0) \cdot x$

**$n \in \mathcal{S} = < n + 1 \in \mathcal{S}$ .** we have that  $(n+1) \cdot x = x + ((n+1)-1) \cdot x = x + i_{\mathbb{R}_{\mathbb{N}}}(n) \cdot x = i_{\mathbb{R}_{\mathbb{N}}}(1) \cdot x + i_{\mathbb{R}_{\mathbb{N}}}(n) \cdot x = i_{\mathbb{R}_{\mathbb{N}}}(n+1) \cdot x = x$   $\square$

Motivated by the above theorem and the fact that we already use the same symbols for the neutral element and multiplicative unit in  $\mathbb{R}$  and  $\mathbb{N}$  we introduce the following notation.

**Notation 9.61.** If  $n \in \mathbb{N}$  then we note  $i_{\mathbb{R}_{\mathbb{N}}}(n)$  also by  $n$ . So if in a operation (sum, multiplication or comparison) we mix elements of  $\mathbb{N}$  and elements of  $\mathbb{R}$  we implicitly assume that we have applied  $i_{\mathbb{R}_{\mathbb{N}}}$  to the elements of  $\mathbb{N}$ .

**Theorem 9.62.**  $\forall n \in \mathbb{N}_{\mathbb{C}}$  we have that  $n <_{\mathbb{R}} 2^n$  [and thus by 9.56 we have also  $\forall a \in \mathbb{R}_{\mathbb{C}} \exists n \in \mathbb{N}_{\mathbb{C}}$  such that  $a <_{\mathbb{R}} n_{\mathbb{R}} < 2^n$ ] (where  $2 = (2, 0) \in \mathbb{N}_{\mathbb{C}}$ )

**Proof.** This is proved by induction so let  $S = \{n \in \mathbb{N} | n <_{\mathbb{R}} 2^n\}$  then we have that

1. If  $n = 0$  then  $0 < 1 = 2^0 \Rightarrow 0 = 0 <_{\mathbb{R}} 1 = 2^0$  so that  $0 \in S$
2. If  $n \in S$  then  $n <_{\mathbb{R}} 2^n$  then consider the following cases for  $n+1$  (as  $0 <_{\mathbb{R}} n+1$ )
  - a.  **$(n+1 = 1)$**  then  $n+1 = 1 <_{\mathbb{R}} 2 = 2^1 = 2^{n+1} \Rightarrow n+1 \in S$
  - b.  **$(1 <_{\mathbb{R}} n+1)$**  then  $2 \leq_{\mathbb{R}} n+1 \Rightarrow 1 \leq_{\mathbb{R}} n$  so that  $n+1 \leq_{\mathbb{R}} n+n = 2 \cdot n <_{\mathbb{R}} 2 \cdot 2^n = 2^{n+1}$  proving that  $n+1 \in S$   $\square$

**Theorem 9.63.** Let  $x \in \mathbb{R}_{\mathbb{C}}$  with  $x >_{\mathbb{R}} 1$  then  $x^n - 1 \geq_{\mathbb{R}} n \cdot (x - 1) \forall n \in \mathbb{N}$

**Proof.** We prove this by induction, so let  $S = \{n \in \mathbb{N} | x^n - 1 \geq n \cdot (x - 1)\}$  then we have

1. if  $n = 0$  then  $x^n - 1 = x^0 - 1 = 1 - 1 = 0 \geq_{\mathbb{R}} 0 = 0 \cdot (x - 1) = n \cdot (x - 1) \Rightarrow 0 \in S$
2. If  $n \in S$  then  $x^{n+1} - 1 = x \cdot x^{n+1} - 1 = x \cdot (x^n - 1) + (x - 1) \geq_{\mathbb{R}} x \cdot (n \cdot (x - 1)) + (x - 1) = (x - 1) \cdot (x \cdot n + 1) \geq_{\mathbb{R}_{x >_{\mathbb{R}} 1 \Rightarrow n \cdot x \geq_{\mathbb{R}} n}} (x - 1) \cdot (n + 1) \Rightarrow n + 1 \in S$   $\square$

**Theorem 9.64.** If  $N \in \mathbb{N}_{\mathbb{C}} \setminus \{0\}$  and  $x \in \mathbb{R}_{\mathbb{C}}$  with  $x >_{\mathbb{R}} 1$  then there exists a  $n \in \mathbb{N}_{\mathbb{C}} \setminus \{0\}$  such that  $x^n >_{\mathbb{R}} N$

**Proof.** Consider the following cases for  $N \in \mathbb{N}_{\mathbb{C}} \setminus \{0\}$

1. ( $N = 1$ ) then  $N = 1 <_{\mathbb{R}} x = x^1$  proving the theorem
2. ( $N > 1$ ) take  $\delta = \frac{N-1}{x-1} >_{\mathbb{R}} 0$  then by the Archimedean property (9.56) there exists a  $n \in \mathbb{N}_{\mathbb{C}}$  such that  $\delta <_{\mathbb{R}} n$  as by definition of  $\delta$  we have  $N - 1 = \delta \cdot (x - 1) \leq_{\mathbb{R}} n \cdot (x - 1) <_{\mathbb{R}} x^n - 1$  (using the previous theorem) so that  $N <_{\mathbb{R}} x^n$   $\square$

**Corollary 9.65.** Let  $x, y \in \mathbb{R}_{\mathbb{C}}$  with  $x >_{\mathbb{R}} 1$  then there exists a  $n \in \mathbb{N}_{\mathbb{C}} \setminus \{0\}$  such that  $x^n >_{\mathbb{R}} y$

**Proof.** Using 9.56 there exists a  $N \in \mathbb{N}_{\mathbb{C}} \setminus \{0\}$  such that  $y <_{\mathbb{R}} N$ , using the previous theorem there exists then a  $n \in \mathbb{N}_{\mathbb{C}} \setminus \{0\}$  such that  $N <_{\mathbb{R}} x^n$  hence  $y <_{\mathbb{R}} x^n$   $\square$

**Theorem 9.66.** If  $x \in \mathbb{R}_{\mathbb{C}}$  with  $0 <_{\mathbb{R}} x <_{\mathbb{R}} 1$  and  $n, m \in \mathbb{N}_{\mathbb{C}}$  with  $n <_{\mathbb{R}} m$  then  $x^m <_{\mathbb{R}} x^n$

**Proof.** We prove this by induction, so let  $m \in \mathbb{N}_{\mathbb{C}}$  and take  $S_m = \{n \in \mathbb{N} | x^{m+n} <_{\mathbb{R}} x^m\}$  then we have

1. if  $n = 1$  then  $x^{m+n} = x^{n+1} = x \cdot (x^n) <_{\mathbb{R}_{x <_{\mathbb{R}} 1 \Rightarrow x \cdot x^n <_{\mathbb{R}} x^n}} x^n \Rightarrow 1 \in S_m$
2. if  $n \in S_m$  then  $x^{m+(n+1)} = x \cdot x^{m+n} <_{\mathbb{R}_{x <_{\mathbb{R}} 1}} x^{m+n} <_{\mathbb{R}_{n \in S_m}} x^m$

Using mathematical induction we have then that  $S_m = \mathbb{N}$ . If now  $n, m \in \mathbb{N}_{\mathbb{R}}$  with  $n <_{\mathbb{R}} m$  then  $k = m - n >_{\mathbb{R}} 0$  or  $k \in \mathbb{N} = S_n$  so that  $x^m = x^{n+k} <_{\mathbb{R}} x^n$   $\square$

**Theorem 9.67.** If  $\varepsilon \in \mathbb{R}_{\mathbb{C}+}$  and  $x \in \mathbb{R}_{\mathbb{C}}$  such that  $0 <_{\mathbb{R}} x <_{\mathbb{R}} 1$  then  $\exists N \in \mathbb{N}_{\mathbb{C}} \setminus \{0\}$  such that  $0 <_{\mathbb{R}} x^n <_{\mathbb{R}} \varepsilon$  if  $n \geq N$

**Proof.** As  $\varepsilon >_{\mathbb{R}} 0$  we have that  $\frac{1}{\varepsilon}$  is defined and by the Archimedean property (see 9.56) there exists a  $n \in \mathbb{N}_{\mathbb{C}}$  such that  $0 <_{\mathbb{R}} \frac{1}{\varepsilon} <_{\mathbb{R}} n$ , as  $0 <_{\mathbb{R}} x <_{\mathbb{R}} 1$  we have that  $1 <_{\mathbb{R}} \frac{1}{x}$  so that by 9.64 there exists a  $N \in \mathbb{N}_{\mathbb{C}} \setminus \{0\}$  such that  $n <_{\mathbb{R}} x^N$  and thus we have that  $0 <_{\mathbb{R}} \frac{1}{\varepsilon} <_{\mathbb{R}} \left(\frac{1}{x}\right)^N = \frac{1}{x^N} \Rightarrow 0 <_{\mathbb{R}} x^N <_{\mathbb{R}} \varepsilon$ . If now we have  $n \geq N$  then by the previous theorem we have  $0 <_{\mathbb{R}} x^n \leq_{\mathbb{R}} x^N <_{\mathbb{R}} \varepsilon$   $\square$

**Theorem 9.68.** Let  $x, y \in \mathbb{R}_{\mathbb{C}+}$  such that  $y < x$  then  $\forall n \in \mathbb{N}$  we have  $y^n < x^n$

**Proof.** We prove this by induction, so take  $\mathcal{S} = \{n \in \mathbb{N} | 0 < y^n < x^n\}$  then we have

**1  $\in \mathcal{S}$ .** then  $0 < y^1 = y < x = x^1$  proving that  $1 \in \mathcal{S}$

**$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$ .** then as  $n \in \mathcal{S}$  we have  $0 < y^n < x^n$  then as  $0 < y$  we have that  $0 < y^n \cdot y < x^n \cdot y \Rightarrow 0 < y^{n+1} < x^n \cdot y$ , further as  $y < x$  and  $0 < y^n < x^n$  we have  $y \cdot x^n < x^n \cdot x = x^{n+1}$ . Combining these two inequalities gives  $0 < y^{n+1} < x^{n+1}$  so that  $n+1 \in \mathcal{S}$ .  $\square$

**Corollary 9.69.** Let  $x \in \mathbb{R}_{\mathbb{C}+}$  then  $\forall n \in \mathbb{N}$  we have  $0 < x^n$

**Proof.** If  $n = 0$  then  $0 < 1 = x^0$ , if  $n \in \mathbb{N}$  then using the previous theorem we have  $0 = 0^n < x^0$   $\square$

### 9.4.3 Square root on $\mathbb{R}_{\mathbb{C}}$

**Definition 9.70.** Let  $x = (x', 0) \in \mathbb{R}_{\mathbb{C}}$  with  $0 \leq_{\mathbb{R}} x$  then we define  $\sqrt{x} = (\sqrt{x'}, 0)$  (which is well defined as from  $x \geq 0$  it follows that  $x' \geq 0$ )

**Theorem 9.71.** Let  $\circ^2: \{x \in \mathbb{R}_C \mid 0 \leq x\} \rightarrow \{x \in \mathbb{R}_C \mid 0 \leq x\}$  be the map defined by  $x \mapsto x^2$  then  $\sqrt{\cdot}: \{x \in \mathbb{R}_C \mid 0 \leq x\} \rightarrow \{x \in \mathbb{R}_C \mid 0 \leq x\}$  defined by  $x \mapsto \sqrt{x}$  is the inverse of  $\circ^2$ . Hence  $\circ^2$  and  $\sqrt{\cdot}$  are bijections.

**Proof.** Let  $z = (z', 0) \in \{x \in \mathbb{R}_C \mid 0 \leq x\}$  then  $(\sqrt{\cdot} \circ \circ^2)(z) = \sqrt{(z^2)} = \sqrt{(z', 0) \cdot (z', 0)} = \sqrt{((z')^2, 0)} = (\sqrt{(z')^2}, 0) = (z', 0) = z$ . Further we have  $(\circ^2 \circ \sqrt{\cdot})(z) = (\sqrt{z})^2 = (\sqrt{z'}, 0)^2 = (\sqrt{z'}, 0) \cdot (\sqrt{z'}, 0) = ((\sqrt{z'})^2, 0) = (z', 0) = z$   $\square$

The restriction to positive real numbers in the above is needed because for example  $(-1)^2 = 1 = 1^2$  and  $i^2 = -1 = (-i)^2$ .

**Theorem 9.72.**  $\sqrt{\cdot}: \{x \in \mathbb{R}_C \mid 0 \leq x\} \rightarrow \{x \in \mathbb{R}_C \mid 0 \leq x\}$  is a strictly increasing function

**Proof.** If  $x, y \in \{x \in \mathbb{R}_C \mid 0 \leq x\}$  is such that  $x < y$ . Assume then that  $\sqrt{y} \leq \sqrt{x}$  then by 9.41 we have  $(\sqrt{y})^2 \leq (\sqrt{x})^2 \Rightarrow y \leq x$  contradicting  $x < y$  so we must have  $\sqrt{x} < \sqrt{y}$   $\square$

**Theorem 9.73.** If  $x \in \mathbb{R}_C$  and  $0 \leq a \in \mathbb{R}_C$  then we have

1.  $x^2 = a$  if and only if  $x = \sqrt{a}$  or  $x = -\sqrt{a}$
2.  $x^2 \leq a$  if and only if  $-\sqrt{a} \leq x \leq \sqrt{a}$
3.  $x^2 < a$  if and only if  $-\sqrt{a} < x < \sqrt{a}$

**Proof.**

1.

$\Rightarrow$ . For  $x$  we have either

$0 \leq x$ . then using 9.71 we have that  $x = (\sqrt{\cdot}) \circ (\circ^2)(x) = (\sqrt{\cdot})(x^2) = (\sqrt{\cdot})(a) = \sqrt{(a)}$ .

$x < 0$ . then  $0 < -x$  so that using the case  $(0 < x)$  we have that  $-x = \sqrt{a}$  or  $x = -\sqrt{a}$

$\Leftarrow$ . If  $x = \sqrt{a}$  then  $x^2 = (\sqrt{a})^2 = a$ , if  $x = -\sqrt{a}$  then  $x^2 = (-\sqrt{a})^2 = (\sqrt{a})^2 = a$ .

2.

$\Rightarrow$ . For  $x$  we have either

$0 \leq x$ . Then as  $0 \leq \sqrt{a}$  we have that  $-\sqrt{a} \leq 0 \leq x$ . Further assume that  $\sqrt{a} < x$  then, as  $(\circ^2)$  is strictly increasing [see 9.41], we have that  $a = (\sqrt{a})^2 < x^2$  contradicting  $x^2 \leq a$ , hence we must have that  $x \leq \sqrt{a}$ . So we have  $-\sqrt{a} \leq x \leq \sqrt{a}$

$x < 0$ . then  $0 < -x$  and we can apply the case  $(0 \leq x)$  to find that  $-\sqrt{a} \leq -x \leq \sqrt{a}$ , multiplying this by  $(-1)$  gives  $-\sqrt{a} \leq x \leq \sqrt{a}$

$\Leftarrow$ . Consider the following cases for  $x$

$0 \leq x$ . then from  $x \leq \sqrt{a}$  we have, as  $(\circ^2)$  is increasing [see 9.41], that  $x^2 \leq (\sqrt{a})^2 = a$ .

$x < 0$ . then from  $-\sqrt{a} \leq x$  we have  $0 < -x < \sqrt{a}$ . As  $(\circ^2)$  is increasing [see 9.41] it follows that  $x^2 = (-x)^2 < (\sqrt{a})^2 = a$

3.

$\Rightarrow$ . Using (2) we have  $-\sqrt{a} \leq x \leq \sqrt{a}$ . If  $x = \sqrt{a}$  then  $x^2 = (\sqrt{a})^2 = a$  contradicting  $x^2 < a$ , if  $x = -\sqrt{a}$  then  $(-x) = \sqrt{a}$  so that  $x^2 = (-x)^2 = (\sqrt{a})^2 = a$  contradicting  $x^2 < a$ . So we must have that  $-\sqrt{a} < x < \sqrt{a}$ .

$\Leftarrow$ . Using (2) we have that  $x^2 \leq a$ . If  $x^2 = a$  then by (1)  $x$  would be either  $\sqrt{a}$  or  $-\sqrt{a}$ , where both contradicts  $-\sqrt{a} < x < \sqrt{a}$  so we must have that  $x^2 < a$ .  $\square$

**Theorem 9.74.** If  $x, y \in \{x \in \mathbb{R}_{\mathbb{C}} \mid 0 \leq x\}$  then  $\sqrt{x \cdot y} = \sqrt{x} \cdot \sqrt{y}$

**Proof.** As  $(\sqrt{x \cdot y})^2 = x \cdot y = (\sqrt{x})^2 \cdot (\sqrt{y})^2 = (\sqrt{x} \cdot \sqrt{y})^2$  we have by the fact that  $\{x \in \mathbb{R}_{\mathbb{C}} \mid 0 \leq x\} \rightarrow \{x \in \mathbb{R} \mid 0 \leq x\}$  is a bijection and thus injective so we have  $\sqrt{x \cdot y} = \sqrt{x} \cdot \sqrt{y}$   $\square$

**Theorem 9.75.** Given  $x, y \in \{x \in \mathbb{R}_{\mathbb{C}} \mid 0 \leq x\}$  then we have  $\sqrt{x+y} \leq \sqrt{x} + \sqrt{y}$

**Proof.** We prove this by contradiction, so assume that  $\sqrt{x} + \sqrt{y} < \sqrt{x+y}$   $\xrightarrow{9.41} -\sqrt{x+y} < -(\sqrt{x} + \sqrt{y})$  and by multiplying by  $\sqrt{x} + \sqrt{y} \geq 0$  we have  $-\sqrt{x+y} \cdot (\sqrt{x} + \sqrt{y}) < -(\sqrt{x} + \sqrt{y}) \cdot (\sqrt{x} + \sqrt{y}) = -(x+y+2 \cdot \sqrt{x} \cdot \sqrt{y})$ . Now we have by 9.41 that  $0 \leq (\sqrt{x+y} - (\sqrt{x} + \sqrt{y}))^2 = (\sqrt{x+y})^2 + (\sqrt{x} + \sqrt{y})^2 - 2 \cdot \sqrt{x+y} \cdot (\sqrt{x} + \sqrt{y}) = x+y + (\sqrt{x})^2 + (\sqrt{y})^2 + 2 \cdot \sqrt{x} \cdot \sqrt{y} - 2 \cdot \sqrt{x+y} \cdot (\sqrt{x} + \sqrt{y}) = 2 \cdot x + 2 \cdot y + 2 \cdot \sqrt{x} \cdot \sqrt{y} - 2 \cdot \sqrt{x+y} \cdot (\sqrt{x} + \sqrt{y}) \xrightarrow{\text{multiply by } \frac{1}{2}} 0 \leq x+y + \sqrt{x} \cdot \sqrt{y} - \sqrt{x+y} \cdot (\sqrt{x} + \sqrt{y}) < x+y + \sqrt{x} \cdot \sqrt{y} - (x+y+2 \cdot \sqrt{x} \cdot \sqrt{y}) = -\sqrt{x} \cdot \sqrt{y} \leq 0 \Rightarrow 0 < 0$  a contradiction. So we must have  $\sqrt{x+y} \leq \sqrt{x} + \sqrt{y}$   $\square$

We have now the biggest reason to introduce the complex numbers

**Theorem 9.76.** If  $z \in \mathbb{R}_{\mathbb{C}}$  then there exists a  $x \in \mathbb{C}$  such that  $z = x^2$

**Proof.** We have two cases to consider for  $z$

$0 \leq z$ . take then  $x = \sqrt{z}$  hence  $x^2 = (\sqrt{z})^2 = z$

$z < 0$ . then  $0 \leq (-1) \cdot z$  take then  $x = i \cdot \sqrt{(-1) \cdot z}$  then  $x^2 = i^2 \cdot (\sqrt{(-1) \cdot z})^2 = (-1) \cdot (-1) \cdot z = z$   $\square$

## 9.5 Summarize

To summarize what we have done in the above:

1.  $\langle \mathbb{C}, +, \cdot \rangle$  forms a field
2.  $\langle \mathbb{C}, \leq_{\mathbb{C}} \rangle$  is a total ordered set
3. We have embedded  $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$  and  $\mathbb{R}$  in  $\mathbb{C}$  as the sets  $\mathbb{N}_{\mathbb{C}}, \mathbb{Z}_{\mathbb{C}}, \mathbb{Q}_{\mathbb{C}}$  and  $\mathbb{R}_{\mathbb{C}}$  such that
  - a.  $\mathbb{N}_{\mathbb{C}} \subseteq \mathbb{Z}_{\mathbb{C}} \subseteq \mathbb{Q}_{\mathbb{C}} \subseteq \mathbb{C}$
  - b.  $\langle \mathbb{N}_{\mathbb{C}}, + \rangle$  is a sub semi group of  $\langle \mathbb{C}, + \rangle$
  - c.  $\langle \mathbb{N}_{\mathbb{C}}, \cdot \rangle$  is a sub semi group of  $\langle \mathbb{C}, \cdot \rangle$
  - d.  $\langle \mathbb{Z}_{\mathbb{C}}, +, \cdot \rangle$  is a sub ring of  $\langle \mathbb{C}, +, \cdot \rangle$
  - e.  $\langle \mathbb{R}_{\mathbb{C}}, +, \cdot \rangle$  is a sub field of  $\langle \mathbb{C}, +, \cdot \rangle$
  - f.  $\langle \mathbb{R}_{\mathbb{C}}, \leq_R \rangle$  is a conditional complete total ordered set
  - g.  $\mathbb{R}_{\mathbb{C}}$  has the Archimedean property
  - h.  $\mathbb{Q}_{\mathbb{C}}$  is dense in  $\mathbb{R}_{\mathbb{C}}$
  - i.  $\forall z \in \mathbb{R}_{\mathbb{C}}$  there exists a  $x \in \mathbb{R}$  such that  $x^2 = z$  (later we will extend this by proving the fundamental theorem of Algebra)
  - j.  $\mathbb{R}_{\mathbb{C}}^2$  is bijection with  $\mathbb{C}$
4. We have defined a norm  $\| \cdot \|$  on  $\mathbb{C}$  such that
  - a.  $0 \leq |z| \in \mathbb{R}_{\mathbb{C}}$
  - b.  $|z \cdot z'| = |z| \cdot |z'|$
  - c.  $\operatorname{Re}(z) \leq |z|$
  - d.  $z \cdot \bar{z} = |z|^2$

- e.  $|z + y| \leq |x| + |y|$
- f.  $|z| = 0 \Leftrightarrow z = 0$
- g. If  $z \in \mathbb{R}_{\mathbb{C}}$  then  $|z| = |\operatorname{Re}(z)|$

For the rest of this book we don't have to talk anymore about the original  $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$  (with the exception of  $\mathbb{N}$  that we need for induction). So to simplify notation we will from now on use  $\mathbb{N}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$  when we actually means  $\mathbb{N}_{\mathbb{C}}, \mathbb{Q}_{\mathbb{C}}, \mathbb{R}_{\mathbb{C}}, \mathbb{C}$  and  $\leq$  when we use  $\leq_{\mathbb{R}}$  so that  $\mathbb{N} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}$ .

# Chapter 10

## Linear Algebra

### 10.1 Sums and products

**Definition 10.1.** If  $n \in \mathbb{N}$  then  $\{0, \dots, n\} = S_{n+1} = \{i \in \mathbb{N} \mid i < n+1\} = n+1$  (see 5.16)

**Definition 10.2.** If  $n, m \in \mathbb{N}$  with  $n \leq m$  then  $\{n, \dots, m\} = \{x \in \mathbb{N} \mid n \leq x \leq m\}$

**Definition 10.3. (Finite sum)** Let  $\langle A, + \rangle$  be a semi-group, and  $x = \{x_i\}_{i \in \{0, \dots, n\}}$  a family of elements of  $A$  (see 2.104) then  $\forall j \in \{0, \dots, n\} \sum_{i=0}^j x_i$  is defined as follows via recursion (see 5.17) as follows. Define  $g: S_n \times A \rightarrow A$  by  $(i, a) \rightarrow a + x_{i+1}$  (note if  $n = 0$  then  $g$  is the empty function) then we can use 5.17 to define a  $f: S_{n+1} = \{0, \dots, n\} \rightarrow A$  such that

$$\begin{aligned} f(0) &= x_0 \\ \forall j \in S_n \text{ we have } f(j+1) &= g(j, f(j)) \end{aligned}$$

we define then  $\sum_{i=0}^j x_i = f(j)$ .

**Note 10.4.** The index  $i$  used in  $\sum_{i=0}^j x_i$  and  $x = \{x_i\}_{i \in \{0, \dots, n\}}$  is actually part of a bad (but traditional notation), we could easily use  $k$  instead of  $i$ . Probably a better notation would be  $\sum_{\{0, \dots, n\}} x$  (so no index is used). However the use of indexes allows use to use formula's based on the index, like in  $x = \{10^i\}_{i \in \{0, \dots, n\}}$  and  $\sum_{i=0}^n 10^i$  so we will stick with this notation.

**Note 10.5.** If  $\langle A, \cdot \rangle$  is a semi-group, where the operator  $\cdot$  is interpreted as a product and not a sum, then we note  $\sum_{i=n}^m x_i$  as  $\prod_{i=n}^m x_i$

**Theorem 10.6.** If  $\langle A, + \rangle$  is a semi-group and  $\{x_i\}_{i \in \{0, \dots, n\}}$  then we have

1.  $\sum_{i=0}^0 x_i = x_0$
2.  $\forall j \in \{1, \dots, n\}$  we have  $\sum_{i=0}^j x_i = (\sum_{i=0}^{j-1} x_i) + x_j$
3.  $\forall j \in \{0, \dots, n-1\}$  we have  $\sum_{i=0}^{j+1} x_i = (\sum_{i=0}^j x_i) + x_{j+1}$

**Proof.**

1.  $\sum_{i=0}^0 x_i = f(0) = x_0$
2. If  $j \in \{1, \dots, n\}$  then  $j = (j-1)+1$  and  $\sum_{i=0}^j x_i = f((j-1)+1) = g(j-1, f(j-1)) = f(j-1) + x_j = \sum_{i=0}^{j-1} x_i + x_j$
3. If  $j \in S_n$  we have  $\sum_{i=0}^{j+1} x_i = f(j+1) = g(j, f(j)) = f(j) + x_{j+1} = (\sum_{i=0}^j x_i) + x_{j+1}$   $\square$

**Example 10.7.** If  $\langle A, + \rangle$  is a semi-group (additive with neutral element 0) and  $\{x_i\}_{i \in \{0, \dots, n\}}$  is such that  $\forall i \in \{0, \dots, n\}$  we have  $x_i = a \in A$  then we have  $\sum_{i=0}^n x_i = a \cdot (n+1)$  (see 4.22)

**Proof.** If  $n \in \mathbb{N}$  then we have  $X = \{n \in \mathbb{N} \mid \text{if } \{x_i\}_{i \in \{0, \dots, n\}} \text{ is such that } \forall i \in \{0, \dots, n\} \text{ we have } x_i = a \text{ and thus } \sum_{i=0}^n x_i = a \cdot (n+1)\}$  then :

1. Then if  $n = 0$  we have  $\sum_{i=0}^n x_i = \sum_{i=0}^0 x_i = x_0 = a = a \cdot 1 \Rightarrow 0 \in X$

2. If  $n \in X$  then if  $\{x_i\}_{i \in \{0, \dots, n+1\}}$  is such that  $\forall i \in \{0, \dots, n+1\}$  we have  $x_i = a$  then we have  $\sum_{i=0}^{n+1} x_i = (\sum_{i=0}^n x_i) + x_{n+1} \stackrel{n \in X}{=} (a \cdot n) + a = a \cdot (n+1) \Rightarrow n+1 \in X$

Using mathematical induction we have then  $X = \mathbb{N}$  proving the theorem.  $\square$

**Example 10.8.** If  $\langle A, \cdot \rangle$  is a semi-group (multiplicative with neutral element 1) and  $\{x_i\}_{i \in \{0, \dots, n\}}$  is such that  $\forall i \in \{0, \dots, n\}$  we have  $x_i = a \in A$  then we have  $\prod_{i=0}^n x_i = a^{n+1}$  (see 4.22)

**Proof.** If  $n \in \mathbb{N}$  then we have  $X = \{n \in \mathbb{N} \mid \text{if } \{x_i\}_{i \in \{0, \dots, n\}} \text{ is such that } \forall i \in \{0, \dots, n\} \text{ we have } x_i = a \text{ and thus } \prod_{i=0}^n x_i = a^{n+1}\}$  then :

1. Then if  $n = 0$  we have  $\prod_{i=0}^n x_i = \prod_{i=0}^0 x_i = x_0 = a = a \cdot 1 \Rightarrow 0 \in X$
2. If  $n \in X$  then if  $\{x_i\}_{i \in \{0, \dots, n+1\}}$  is such that  $\forall i \in \{0, \dots, n+1\}$  we have  $x_i = a$  then we have  $\prod_{i=0}^{n+1} = (\prod_{i=0}^n x_i) \cdot x_{n+1} \stackrel{n \in X}{=} (a^n) \cdot a = a^{n+1} \Rightarrow n+1 \in X$

Using mathematical induction we have then  $X = \mathbb{N}$  proving the theorem.  $\square$

**Example 10.9.** If  $x = \{i\}_{i \in \{0, \dots, 3\}}$  then

$$\begin{aligned} \sum_{i=0}^3 i &= \left( \sum_{i=0}^2 i \right) + 3 \\ &= \left( \left( \sum_{i=0}^1 i \right) + 2 \right) + 3 \\ &= \left( \left( \left( \sum_{i=0}^0 i \right) + 1 \right) + 2 \right) + 3 \\ &= ((0+1)+2)+3 = 6 \end{aligned}$$

**Lemma 10.10.** If  $\{\langle A_i, +_i \rangle\}_{i \in I}$  is a family of abelian groups then by 3.15 we have that  $\langle \prod_{i \in I} A_i, + \rangle$  is a abelian group (here  $(x+y)(i) \stackrel{\text{notation}}{=} (x+y)_i \stackrel{\text{definition}}{=} x_i +_i y_i$  and if  $\{x_j\}_{j \in \{0, \dots, n\}}$  is a family in  $\prod_{i \in I} A_i$  then  $\forall i \in I$  we have  $(\sum_{i=0}^n x_j)(i) \stackrel{\text{notation}}{=} (\sum_{i=0}^n x_j)_i = \sum_{i=0}^n (x_j)(i) \stackrel{\text{notation}}{=} \sum_{i=0}^n (x_j)_i$  (last sum is based on  $+_i$ )

**Proof.** This is easily proved by induction so let  $S = \{n \in \mathbb{N} \mid \text{if } \{x_j\}_{j \in \{0, \dots, n\}} \text{ is a family in } \prod_{i \in I} A_i \text{ then } (\sum_{i=0}^n x_j)(i) = \sum_{i=0}^n (x_j)(i)\}$  then we have:

1.  $(\sum_{j=0}^0 x_j)(i) = (x_0)(i) = \sum_{j=0}^0 (x_0)(i) \Rightarrow 0 \in S$
2. If  $n \in S$  take then a family  $\{x_i\}_{i \in \{0, \dots, n+1\}}$  then we have  $(\sum_{j=0}^{n+1} x_j)(i) = ((\sum_{j=0}^n x_j) + x_{n+1})(i) = (\sum_{j=0}^n x_j)(i) +_i x_{n+1}(i) \stackrel{n \in S}{=} (\sum_{j=0}^n x_j(i)) +_i x_{n+1}(i) = \sum_{j=0}^{n+1} x_j(i) \Rightarrow n+1 \in S$

So by mathematical induction we have  $S = \mathbb{N}$  completing our proof.  $\square$

**Lemma 10.11.** If  $\langle A, + \rangle$  is a abelian semi-group and  $\{x_i\}_{i \in \{0, \dots, n\}}$ ,  $\{y_i\}_{i \in \{0, \dots, n\}}$  are families of elements in  $A$  then  $\sum_{i=0}^n (x_i + y_i) = (\sum_{i=0}^n x_i) + (\sum_{i=0}^n y_i)$

**Proof.** We prove this by induction on  $n$  so let  $X = \{n \in \mathbb{N}_0 \mid \text{if } \{x_i\}_{i \in \{0, \dots, n\}}, \{y_i\}_{i \in \{0, \dots, n\}} \text{ then } \sum_{i=0}^n (x_i + y_i) = (\sum_{i=0}^n x_i) + (\sum_{i=0}^n y_i)\}$  then we have:

1. If  $n = 0$  then  $\sum_{i=0}^0 (x_i + y_i) = x_0 + y_0 = (\sum_{i=0}^0 x_i) + (\sum_{i=0}^0 y_i) \Rightarrow 0 \in X$
2. Assume that  $n \in X$  then if  $\{x_i\}_{i \in \{0, \dots, n+1\}}$ ,  $\{y_i\}_{i \in \{0, \dots, n+1\}}$  we have  $\sum_{i=0}^{n+1} (x_i + y_i) = (\sum_{i=0}^n (x_i + y_i)) + (x_{n+1} + y_{n+1}) \stackrel{n \in X}{=} ((\sum_{i=0}^n x_i) + (\sum_{i=0}^n y_i)) + (x_{n+1} + y_{n+1}) \stackrel{\text{commutativity, associativity}}{=} ((\sum_{i=0}^n x_i) + x_{n+1}) + ((\sum_{i=0}^n y_i) + y_{n+1}) = (\sum_{i=0}^{n+1} x_i) + (\sum_{i=0}^{n+1} y_i)$  proving that  $n+1 \in X$

and using mathematical induction we have  $X = \mathbb{N}_0$  proving the theorem.  $\square$

**Theorem 10.12.** Let  $\langle A, + \rangle$  be a abelian semi-group and  $n, m \in \mathbb{N}$  and  $\{x_{i,j}\}_{(i,j) \in \{0, \dots, n\} \times \{0, \dots, m\}}$  then we have  $\sum_{i=0}^n (\sum_{j=0}^m x_{i,j}) = \sum_{j=0}^m (\sum_{i=0}^n x_{i,j})$

**Proof.** We prove this by induction so let  $S_n = \{m \in \mathbb{N} \mid \text{if } \{x_{i,j}\}_{(i,j) \in \{0, \dots, n\} \times \{0, \dots, m\}} \text{ is a family in } A \text{ then } \sum_{i=0}^n (\sum_{j=0}^m x_{i,j}) = \sum_{j=0}^m (\sum_{i=0}^n x_{i,j})\}$  then we have :

1. If  $m = 0$  then  $\sum_{i=0}^n (\sum_{j=0}^0 x_{i,j}) = \sum_{i=0}^n x_{i,0} = \sum_{j=0}^0 (\sum_{i=0}^n x_{i,j}) \Rightarrow 0 \in S_n$
2. Assume that  $m \in S_n$  and take  $\{x_{i,j}\}_{(i,j) \in \{0, \dots, n\} \times \{0, \dots, m+1\}}$  then  $\sum_{i=0}^n (\sum_{j=0}^{m+1} x_{i,j}) = \sum_{i=0}^n ((\sum_{j=0}^m x_{i,j}) + x_{i,m+1}) \stackrel{\text{previous theorem}}{=} \sum_{i=0}^n (\sum_{j=0}^m x_{i,j}) + \sum_{i=0}^n x_{i,m+1} \stackrel{m \in S_n}{=} \sum_{j=0}^m (\sum_{i=0}^n x_{i,j}) + \sum_{i=0}^n x_{i,m+1} = \sum_{j=0}^{m+1} (\sum_{i=0}^n x_{i,j}) \Rightarrow m+1 \in S_n$

using mathematical induction we have that  $S_n = \mathbb{N}$  proving the theorem.  $\square$

### 10.1.1 Sums and associativity

**Definition 10.13.** If  $\langle A, + \rangle$  is a semi-group,  $n, m \in \mathbb{N}$  with  $n \leq m$  and  $\{x_i\}_{i \in \{n, \dots, m\}}$  a family in  $A$  then we define  $\sum_{i=n}^m x_i$  to be equal to  $\sum_{i=0}^{m-n} x_{i+n}$ .

**Theorem 10.14.** If  $\langle A, + \rangle$  is a semi-group  $n, m, k \in \mathbb{N}$  with  $n \leq m$  and  $\{x_i\}_{i \in \{n+k, \dots, m+k\}}$  then  $\sum_{i=n}^m x_{i+k} = \sum_{i=n+k}^{m+k} x_i$

**Proof.**  $\sum_{i=n}^m x_{i+k} \stackrel{\text{definition}}{=} \sum_{i=0}^{m-n} x_{(i+n)+k} = \sum_{i=0}^{m-n} x_{i+(n+k)} \stackrel{\text{definition}}{=} \sum_{i=n+k}^{(m-n)+n+k} x_i = \sum_{i=n+k}^{m+n} x_i$   $\square$

**Lemma 10.15.** If  $\langle A, + \rangle$  is a semi-group  $n, m \in \mathbb{N}$  with  $n \leq m$  and  $\{x_i\}_{i \in \{n, \dots, m+1\}}$  then

1.  $\sum_{i=n}^n x_i = x_n$
2.  $\sum_{i=n}^{m+1} x_i = \sum_{i=n}^m x_i + x_{m+1}$

**Proof.**

1.  $\sum_{i=n}^n x_i = \sum_{i=0}^0 x_{i+n} = x_{0+n} = x_n$
2.  $\sum_{i=n}^{m+1} x_i = \sum_{i=0}^{(m-n)+1} x_{i+n} = \sum_{i=0}^{(m-n)} x_{i+n} + x_{(m-n)+1+n} = \sum_{i=n}^m x_i + x_{m+1}$   $\square$

**Example 10.16.** Let  $\langle A, + \rangle$  be a semi group,  $n, m \in \mathbb{N}$  with  $n < m$  and  $\{x_i\}_{i \in \{n, \dots, m\}}$  be such that  $\forall i \in \{n, \dots, m\}$  we have  $x_i = 0$  then  $\sum_{i=n}^m x_i = 0$

**Proof.** We prove this by induction, so let  $S = \{i \in \mathbb{N} \mid \text{If } \{x_i\}_{i \in \{n, \dots, n+i\}} \text{ is such that } \forall i \in \{n, \dots, n+i\} \text{ we have } x_i = 0 \text{ then } \sum_{i=n}^{n+i} x_i = 0\}$  then we have

$0 \in S$ . then  $\sum_{i=n}^{n+0} x_i = x_n = 0$  proving that  $0 \in S$

$i \in S \Rightarrow i \in S$ . take then  $\sum_{i=n}^{n+(i+1)} x_i = (\sum_{i=n}^{n+i} x_i) + x_{n+(i+1)} \stackrel{i \in S}{=} 0 + 0 = 0$  proving that  $i+1 \in S$

Using mathematical induction we have then that  $S = \mathbb{N}$  so that  $m - n \in S$  and thus  $0 = \sum_{i=n}^{(m-n)+n} x_i = \sum_{i=n}^m x_i$ .  $\square$

**Theorem 10.17.** Let  $\langle A, + \rangle$  be a semi-group,  $n \in \mathbb{N}_0$ ,  $\{x_i\}_{i \in \{0, \dots, n\}}$  a family of element of  $A$  and  $m \in \{0, \dots, n-1\}$  then  $\sum_{i=0}^n x_i = (\sum_{i=0}^m x_i) + (\sum_{i=m+1}^n x_i)$

**Proof.** We prove this by induction on  $n$  so let  $S = \{n \in \mathbb{N}_0 \mid \text{if } \{x_i\}_{i \in \{0, \dots, n\}} \subseteq A \text{ and } m \in \{0, \dots, n-1\} \text{ then } \sum_{i=0}^n x_i = (\sum_{i=0}^m x_i) + (\sum_{i=m+1}^n x_i)\}$  then we have

$1 \in S$ . Let  $\{x_i\}_{i \in \{0, \dots, 1\}} \subseteq A$  then if  $m \in \{0, \dots, 1-1\} \Rightarrow m = 0$  and  $\sum_{i=0}^1 x_i = (\sum_{i=0}^0 x_i) + x_1 \stackrel{10.15}{=} (\sum_{i=0}^0 x_i) + (\sum_{i=1}^1 x_i) = (\sum_{i=0}^1 x_i) + (\sum_{i=m+1}^1 x_i)$  proving that  $1 \in S$

$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$ . Let  $\{x_i\}_{i \in \{1, \dots, n+1\}} \subseteq A$  and take  $m \in \{1, \dots, (n+1)-1\} = \{0, \dots, n\}$  then we have either

$\mathbf{m = n}$ . Then  $\sum_{i=0}^{n+1} x_i \stackrel{10.6}{=} (\sum_{i=0}^n x_i) + x_{n+1} \stackrel{10.15}{=} (\sum_{i=0}^n x_i) + (\sum_{i=n+1}^{n+1} x_i) = (\sum_{i=0}^m x_i) + (\sum_{i=m+1}^{n+1} x_i)$  proving that in this case  $n+1 \in \mathcal{S}$

$\mathbf{m \in \{0, \dots, n-1\}}$ . Then as  $n \in \mathcal{S}$  we have for the subsequence  $\{x_i\}_{i \in \{0, \dots, n\}}$  that

$$\sum_{i=0}^n x_i = \left( \sum_{i=0}^m x_i \right) + \left( \sum_{i=m+1}^n x_i \right) \quad (10.1)$$

so that

$$\begin{aligned} \sum_{i=0}^{n+1} x_i &\stackrel{10.6}{=} \left( \sum_{i=0}^n x_i \right) + x_{n+1} \\ &\stackrel{10.1}{=} \left( \left( \sum_{i=0}^m x_i \right) + \left( \sum_{i=m+1}^n x_i \right) \right) + x_{n+1} \\ &\stackrel{\text{associativity}}{=} \left( \sum_{i=0}^m x_i \right) + \left( \left( \sum_{i=m+1}^n x_i \right) + x_{n+1} \right) \\ &\stackrel{10.15}{=} \left( \sum_{i=0}^m x_i \right) + \left( \sum_{i=m+1}^{n+1} x_i \right) \end{aligned}$$

proving that in this case  $n+1 \in \mathcal{S}$

Mathematical induction proves then the theorem.  $\square$

**Theorem 10.18. (Generalized Associativity)** Let  $\langle A, + \rangle$  be a semi-group. Take  $\{(b_j, e_j)\}_{j \in \{0, \dots, m\}} \subseteq \mathbb{N} \times \mathbb{N}$  a sequence of natural number pairs such that  $b_0 = 0$ ,  $\forall j \in \{0, \dots, m\}$  we have  $b_j \leq e_j$  and  $\forall j \in \{0, \dots, m-1\}$  we have  $e_j + 1 = b_{j+1}$ . Then we have for a family  $\{x_i\}_{i \in \{0, \dots, e_m\}} \subseteq A$  that  $\sum_{i=0}^{e_m} x_i = \sum_{k=0}^m (\sum_{l=b_k}^{e_k} x_l)$

**Proof.** We prove this by induction on  $m$  so take  $\mathcal{S} = \{m \in \mathbb{N} \mid \text{let } \{(b_j, e_j)\}_{j \in \{0, \dots, m\}} \subseteq \mathbb{N} \times \mathbb{N} \text{ be such that } b_0 = 0, \forall j \in \{0, \dots, m\} \text{ we have } b_j \leq e_j \text{ and } \forall j \in \{0, \dots, m-1\} \text{ we have } e_j + 1 = b_{j+1}\}$ . If  $\{x_i\}_{i \in \{0, \dots, e_m\}} \subseteq A$  then  $\sum_{i=0}^{e_m} x_i = \sum_{k=0}^m (\sum_{l=b_k}^{e_k} x_l)$ . We have then

$\mathbf{0 \in \mathcal{S}}$ . Then  $b_0 = 0, e_m = e_0$  so if  $\{x_i\}_{i \in \{0, \dots, e_0\}} \subseteq A$  then  $\sum_{i=0}^{e_m} x_i = \sum_{k=0}^{e_0} x_i \stackrel{10.6}{=} \sum_{k=0}^0 (\sum_{l=0}^{e_0} x_l) = \sum_{k=0}^0 (\sum_{l=b_0}^{e_0} x_l)$  proving that  $0 \in \mathcal{S}$ .

$\mathbf{m \in \mathcal{S} \Rightarrow m+1 \in \mathcal{S}}$ . Take  $\{(b_j, e_j)\}_{j \in \{0, \dots, m+1\}} \subseteq \mathbb{N} \times \mathbb{N}$  with  $b_0 = 0, \forall j \in \{0, \dots, m+1\}$  we have  $b_j \leq e_j$  and  $\forall j \in \{0, \dots, (m+1)-1\} = \{0, \dots, m\}$  we have  $e_j + 1 = b_{j+1}$  and take  $\{x_i\}_{i \in \{0, \dots, e_{m+1}\}} \subseteq A$ . Then we have still  $b_0 = 0, \forall j \in \{0, \dots, m\} b_j \leq e_j$  and  $\forall j \in \{0, \dots, m-1\}$  we have  $e_j + 1 = b_{j+1}$ , so as  $m \in \mathcal{S}$  we have for the subsequence  $\{x_i\}_{i \in \{0, \dots, e_m\}}$  that

$$\sum_{i=0}^{e_m} x_i = \sum_{k=0}^m \left( \sum_{l=b_k}^{e_k} x_l \right) \quad (10.2)$$

Next as  $0 \leq b_0 \leq e_0 \leq b_m \leq e_m < e_m + 1 = b_{m+1} \leq e_{m+1}$  we have  $0 \leq e_m \leq e_{m+1} - 1$

$$\begin{aligned} \sum_{i=0}^{e_{m+1}} x_i &\stackrel{10.17 \wedge e_m \in \{0, \dots, e_{m+1}-1\}}{=} \left( \sum_{i=0}^{e_m} x_i \right) + \left( \sum_{i=e_m+1}^{e_{m+1}} x_i \right) \\ &= \left( \sum_{i=0}^{e_m} x_i \right) + \left( \sum_{i=b_{m+1}}^{e_{m+1}} x_i \right) \\ &\stackrel{10.2}{=} \sum_{k=0}^m \left( \sum_{l=b_k}^{e_k} x_l \right) + \left( \sum_{l=b_{m+1}}^{e_{m+1}} x_l \right) \\ &\stackrel{10.6}{=} \sum_{k=0}^{m+1} \left( \sum_{l=b_k}^{e_k} x_l \right) \end{aligned}$$

proving that  $m+1 \in \mathcal{S}$

Mathematical induction completes the theorem.  $\square$

### 10.1.2 Sums and commutativity

**Definition 10.19.** If  $I$  is a set and  $\sigma: I \rightarrow I$  a bijective function then  $\sigma$  is called a **permutation**. The set of all the permutations on  $I$  is noted as  $S_I = \{f \in I^I \mid f \text{ is a permutation}\}$

**Theorem 10.20.** Let  $I$  be a set and  $S_I$  its permutation set then  $\langle S_I, \circ \rangle$  is a group

**Proof.** First  $\circ: S_I \times S_I \rightarrow S_I$  is a function as if  $(f, g) \in S_I \times S_I$  we have that  $f \circ g \in S_I$  (see 2.36)

1. if  $f, g, h \in S_I$  then  $(f \circ g) \circ h = f \circ (g \circ h)$  (see 2.37)
2. if  $f \in S_I$  then  $1_A \in S_I$  and by 2.37 we have  $1_I \circ f = f = f \circ 1_I$
3. if  $f \in S_I$  then as  $f$  is a bijection and we have thus that  $f \circ f^{-1} = 1_A = f^{-1} \circ f$   $\square$

**Theorem 10.21.** If  $\sigma: I \rightarrow I$  is a permutation such that  $\sigma(i) = i$  then  $\sigma|_{I \setminus \{i\}}: I \setminus \{i\} \rightarrow I \setminus \{i\}$  is a permutation

**Proof.**

1. **(injective)** If  $\sigma|_{I \setminus \{i\}}(j) = \sigma|_{I \setminus \{i\}}(j') \Rightarrow \sigma(j) = \sigma(j')$   $\underset{\sigma \text{ is injective}}{\Rightarrow} j = j'$
2. **(surjective)** If  $l \in I \setminus \{i\}$   $\underset{\sigma \text{ is surjective}}{\Rightarrow} \exists k \in I \vdash \sigma(k) = l$  if  $k = i$  then  $l = \sigma(i) = i$  contradicting  $l \in I \setminus \{i\}$  so  $k \in I \setminus \{i\}$  and  $\sigma|_{I \setminus \{i\}}(k) = \sigma(k) = l$   $\square$

**Definition 10.22.** Let  $n, m \in \mathbb{N}$  with  $n \leq m$  and  $i, j \in \{n, \dots, m\}$  then  $(i \leftrightarrow_{n, m} j): \{n, \dots, m\} \rightarrow \{n, \dots, m\}$  is defined by  $(i \leftrightarrow_{n, m} j)(k) = \begin{cases} k & \text{if } k \neq i, j \\ j & \text{if } k = i \\ i & \text{if } k = j \end{cases}$ . If  $n = 1$  then we note  $(i \leftrightarrow_{1, m} j)$  as  $(i \leftrightarrow_m j)$ .

**Theorem 10.23.** If  $n, m \in \mathbb{N}$  with  $n \leq m$  and  $i, j \in \{n, \dots, m\}$  then  $(i \leftrightarrow_{n, m} j): \{n, \dots, m\} \rightarrow \{n, \dots, m\}$  is a **permutation** called a **transposition**. Note that trivially  $(i \leftrightarrow_{n, m} j) = (j \leftrightarrow_{n, m} i)$  and that  $(i \leftrightarrow_{n, m} i) = \mathbb{1}_{\{n, \dots, m\}}$  [the identity function].

**Proof.** We have to prove injectivity and surjectivity

**injectivity.** If  $(i \leftrightarrow_n j)(k) = (i \leftrightarrow_n j)(k')$  then we have the following cases for  $k, k'$ :

$k = i \wedge k' = i$ . Then we have evidently that  $k = k'$

$k = i \wedge k' = j$ . Then  $k' = j = (i \leftrightarrow_{n, m} j)(k) = (i \leftrightarrow_{n, m} j)(k') = i = k$

$k = i \wedge k' \neq i, j$ . Then  $j = (i \leftrightarrow_{n, m} j)(k) = (i \leftrightarrow_{n, m} j)(k') = k' \neq i, j$  a contradiction so this case does not count.

$k = j \wedge k' = i$ . Then  $k' = i = (i \leftrightarrow_{n, m} j)(k) = (i \leftrightarrow_{n, m} j)(k') = j = k$  and thus  $k = k'$

$k = j \wedge k' = j$ . Then we have evidently that  $k = k'$

$k = j \wedge k' \neq i, j$ . Then  $i = (i \leftrightarrow_{n, m} j)(k) = (i \leftrightarrow_{n, m} j)(k') = k' \neq i, j$  a contradiction so this case does not count.

$k \neq i, j \wedge k' = i$ . Then  $k = (i \leftrightarrow_{n, m} j)(k) = (i \leftrightarrow_{n, m} j)(k') = k'$  and thus  $k = k'$ .

$k \neq i, j \wedge k' = j$ . Then  $k = (i \leftrightarrow_{n, m} j)(k) = (i \leftrightarrow_{n, m} j)(k') = i$  a contradiction so that this case does not count.

$k \neq i, j \wedge k' \neq i, j$ . Then  $k = (i \leftrightarrow_{n, m} j)(k) = (i \leftrightarrow_{n, m} j)(k') = k'$  and thus  $k = k'$ .

in all the cases that apply we have  $k = k'$  proving injectivity.

**surjectivity.** If  $l \in \{n, \dots, m\}$  then we have the following cases for  $l$

$l = i$ . then  $l = (i \leftrightarrow_{n, m} j)(j)$

$$\begin{aligned} l = j. & \text{ then } l = (i \leftrightarrow_{n,m} j)(i) \\ l \neq i, j. & \text{ then } l = (i \leftrightarrow_{n,m} j)(l) \end{aligned}$$

□

**Lemma 10.24.** If  $n, m \in \mathbb{N}$  with  $n \leq m$  then if  $(i, j) \in \{n, \dots, m\}$  we have that  $(i \leftrightarrow_{n,m+1} j)|_{\{n, \dots, m\}} = (i \leftrightarrow_{n,m} j)$

**Proof.** If  $k \in \{n, \dots, m\}$  then we have either

1. ( $k = i$ ) then  $(i \leftrightarrow_{n,m+1} j)(k) = (i \leftrightarrow_{n,m+1} j)(i) = j = (i \leftrightarrow_{n,m} j)(i) = (i \leftrightarrow_{n,m} j)(k)$
2. ( $k = j$ ) then  $(i \leftrightarrow_{n,m+1} j)(k) = (i \leftrightarrow_{n,m+1} j)(j) = i = (i \leftrightarrow_{n,m} j)(j) = (i \leftrightarrow_{n,m} j)(k)$
3. ( $k \neq i, j$ ) then  $(i \leftrightarrow_{n,m+1} j)(k) = k = (i \leftrightarrow_{n,m} j)(k)$

$$\text{so } (i \leftrightarrow_{n,m+1} j)|_{\{0, \dots, n\}} = (i \leftrightarrow_{n,m} j)$$

□

**Lemma 10.25.** If  $n, m \in \mathbb{N}$  with  $n \leq m$  and  $i, j \in \{n, \dots, m\}$  then  $(i \leftrightarrow_{n,m} j) \circ (i \leftrightarrow_{n,m} j) = \mathbb{1}_{\{n, \dots, m\}}$

**Proof.** We have the following cases for  $k \in \{n, \dots, m\}$

1. ( $k = i$ ) then  $[(i \leftrightarrow_{n,m} j) \circ (i \leftrightarrow_{n,m} j)](k) = (i \leftrightarrow_{n,m} j)((i \leftrightarrow_{n,m} j)(k)) = (i \leftrightarrow_{n,m} j)(j) = i = k$
2. ( $k = j$ ) then  $[(i \leftrightarrow_{n,m} j) \circ (i \leftrightarrow_{n,m} j)](k) = (i \leftrightarrow_{n,m} j)((i \leftrightarrow_{n,m} j)(k)) = (i \leftrightarrow_{n,m} j)(i) = j = k$
3. ( $k \neq i, j$ ) then  $[(i \leftrightarrow_{n,m} j) \circ (i \leftrightarrow_{n,m} j)](k) = (i \leftrightarrow_{n,m} j)((i \leftrightarrow_{n,m} j)(k)) = (i \leftrightarrow_{n,m} j)(k) = k$

proving the lemma. □

The use of commutativity and associativity of abelian group gives the following

**Theorem 10.26.** If  $\langle A, + \rangle$  is a **abelian semi-group**,  $n \in \mathbb{N}$ ,  $\{x_i\}_{i \in \{0, \dots, n\}}$  a family of elements in  $A$ ,  $\sigma: \{0, \dots, n\} \rightarrow \{0, \dots, n\}$  a permutation then  $\sum_{k=0}^n x_i = \sum_{k=0}^n x_{\sigma(k)}$

**Proof.** We prove this by mathematical induction so let  $\mathcal{S} = \{n \in \mathbb{N} \mid \text{if } \sigma \in \mathcal{S}_{\{0, \dots, n\}} \text{ and } \{x_i\}_{i \in \{0, \dots, n\}} \subseteq A \text{ then } \sum_{i=0}^n x_{\sigma(i)} = \sum_{i=0}^n x_i\}$  then we have

**0 ∈ S.** In this cases  $\sigma = \mathbb{1}_{\{0, \dots, 0\}}$  so that  $\sum_{i=0}^0 x_{\sigma(i)} = \sum_{i=0}^0 x_i$  proving that  $0 \in \mathcal{S}$ .

**n ∈ S ⇒ n + 1 ∈ S.** Then we have for  $n + 1$  either

**n + 1 = 1.** Then if  $\{x_i\}_{i \in \{0, 1\}} \subseteq A$  and  $\sigma$  is either  $(0 \leftrightarrow_{0,1} 0) = \mathbb{1}_{\{0,1\}}$  or  $(0 \leftrightarrow_{0,1} j)$  hence  $\sum_{i=0}^1 x_{\sigma(i)} = x_{\sigma(0)} + x_{\sigma(1)} \stackrel{\text{commutativity}}{=} x_0 + x_1 = \sum_{i=0}^1 x_i$  proving that  $n + 1 \in \mathcal{S}$

**1 < n + 1.** Then if  $\{x_i\}_{i \in \{0, \dots, n+1\}} \subseteq A$  we have for  $\sigma(n+1)$  either :

**σ(n + 1) = n + 1.** then by 10.21 we have that  $\sigma' = \sigma|_{\{0, \dots, n\}} \in \mathcal{S}_{\{0, \dots, n\}}$ . Hence

$$\begin{aligned} \sum_{i=0}^{n+1} x_{\sigma(i)} & \stackrel{10.6}{=} \left( \sum_{i=0}^n x_{\sigma(i)} \right) + x_{\sigma(n+1)} \\ & = \left( \sum_{i=0}^n x_{\sigma(i)} \right) + x_{n+1} \\ & = \left( \sum_{i=0}^n x_{\sigma'(i)} \right) + x_{n+1} \\ & \stackrel{\sigma' \in \mathcal{S}_{\{0, \dots, n\}} \wedge n \in \mathcal{S}}{=} \left( \sum_{i=0}^n x_i \right) + x_{n+1} \\ & \stackrel{10.6}{=} \sum_{i=0}^{n+1} x_i \end{aligned}$$

proving that  $n + 1 \in \mathcal{S}$

$\sigma(n+1) \neq n+1$ . Then  $k = \sigma^{-1}(n+1) \neq n+1$  so that  $k \in \{0, \dots, n\}$  then we have as  $(n \leftrightarrow_n k) \in \mathcal{S}_{\{0, \dots, n\}}$

$$\begin{aligned}
 \sum_{i=0}^{n+1} x_{\sigma(i)} &\stackrel{10.6}{=} \left( \sum_{i=0}^n x_{\sigma(i)} \right) + x_{\sigma(n+1)} \\
 &\stackrel{n \in \mathcal{S}}{=} \left( \sum_{i=0}^n x_{\sigma((n \leftrightarrow_{0,n} k)(i))} \right) + x_{\sigma(n+1)} \\
 &\stackrel{10.6}{=} \left( \left( \sum_{i=0}^{n-1} x_{\sigma((n \leftrightarrow_n k)(i))} \right) + x_{\sigma((n \leftrightarrow_n k)(n))} \right) + x_{\sigma(n+1)} \\
 &= \left( \left( \sum_{i=0}^{n-1} x_{\sigma((n \leftrightarrow_{0,n} k)(i))} \right) + x_{\sigma(k)} \right) + x_{\sigma(n+1)} \\
 &= \left( \left( \sum_{i=0}^{n-1} x_{\sigma((n \leftrightarrow_{0,n} k)(i))} \right) + x_{n+1} \right) + x_{\sigma(n+1)} \\
 &\stackrel{\text{associativity}}{=} \left( \sum_{i=0}^{n-1} x_{\sigma((n \leftrightarrow_{0,n} k)(i))} \right) + (x_{n+1} + x_{\sigma(n+1)}) \\
 &\stackrel{\text{commutativity}}{=} \left( \sum_{i=0}^{n-1} x_{\sigma((n \leftrightarrow_{0,n} k)(i))} \right) + (x_{\sigma(n+1)} + x_{n+1}) \\
 &\stackrel{\text{associativity}}{=} \left( \left( \sum_{i=0}^{n-1} x_{\sigma((n \leftrightarrow_{0,n} k)(i))} \right) + x_{\sigma(n+1)} \right) + x_{n+1} \quad (10.3)
 \end{aligned}$$

As  $\sigma(n+1) \neq n+1$  we have  $\sigma(n+1) \in \{0, \dots, n\}$ . Further if  $\exists i \in \{1, \dots, n-1\}$  such that  $\sigma((n \leftrightarrow_{0,n} k)(i)) = n+1 = \sigma(k)$  then  $(n \leftrightarrow_{0,n} k)(i) = k = (n \rightarrow_{0,n} k)(n)$  so that  $i = n$  a contradiction, hence  $\forall i \in \{0, \dots, n-1\}$  we have  $\sigma((n \leftrightarrow_{0,n} k))(i) \in \{0, \dots, n-1\}$ . If we define then  $\forall i \in \{0, \dots, n\} \in \gamma(i) = \begin{cases} \sigma((n \leftrightarrow_{0,n} k)(i)) & \text{if } i \in \{0, \dots, n-1\} \\ \sigma(n+1) & \text{if } i = n \end{cases}$  then  $\gamma(i) \in \{0, \dots, n\}$  hence  $\gamma: \{0, \dots, n\} \rightarrow \{0, \dots, n\}$  is a function. We prove now that  $\gamma$  is a bijection

**injectivity.** Assume that  $\gamma(i) = \gamma(i')$  then we have either

$i = n = i'$ . then  $i = i'$

$i = n \wedge i' \neq n$ . then  $\sigma(n+1) = \gamma(i) = \gamma(i') = \sigma((n \leftrightarrow_{0,n} k)(i'))$  so that  $n+1 = (n \leftrightarrow_{0,n} k)(i') \in \{0, \dots, n\}$  a contradiction. So this case does not apply.

$i \neq n \wedge i' = n$ . then  $\sigma(n+1) = \gamma(i') = \gamma(i) = \sigma((n \leftrightarrow_n k)(i))$  so that  $n+1 = (n \leftrightarrow_{0,n} k)(i) \in \{0, \dots, n\}$  a contradiction. So this case does not apply.

$i \neq n \wedge i' \neq n$ . then  $\sigma((n \leftrightarrow_{0,n} k)(i)) = \gamma(i) = \gamma(i') = \sigma((n \leftrightarrow_{0,n} k)(i))$  so that  $(n \leftrightarrow_n k)(i) = (n \leftrightarrow_n k)(i')$  proving that  $i = i'$

hence in all valid cases we have  $i = i'$  proving injectivity.

**surjectivity.** Let  $l \in \{0, \dots, n\}$  then by the surjectivity of  $\sigma$  there exists a  $m \in \{0, \dots, n+1\}$  such that  $\sigma(m) = l$ . We have then the following cases:

$m = n+1$ . then  $\gamma(n) = \sigma(n+1) = l$  giving  $\gamma(n) = l$

$m = n$ . then  $k \neq n$  [otherwise  $n+1 = \sigma(k) = \sigma(n) = l$  leading to the contradiction  $n+1 = l \in \{0, \dots, n\}$ . Hence  $k \in \{0, \dots, n-1\}$  so that  $\gamma(k) = \sigma((n \leftrightarrow_{0,n} k)(k)) = \sigma(n) = \sigma(m) = l$  giving  $\gamma(k) = l$ ]

$m \in \{0, \dots, n-1\}$ . then  $m \neq k$  [otherwise  $n+1 = \sigma(k) = \sigma(m) = l \in \{0, \dots, n\}$  a contradiction]. Hence  $\gamma(m) = \sigma((n \leftrightarrow_{0,n} k)(m)) \stackrel{m \neq n, k}{=} \sigma(m) = l$  proving that  $\gamma(m) = l$

So in all cases we can find a  $i \in \{0, \dots, n\}$  such that  $\gamma(i) = l$  proving surjectivity.

The above proves that  $\gamma \in \mathcal{S}_{\{0, \dots, n\}}$  giving

$$\begin{aligned} \left( \left( \sum_{i=0}^{n-1} x_{\sigma((n \leftrightarrow_{0,n} k)(i))} \right) + x_{\sigma(n+1)} \right) &= \left( \left( \sum_{i=0}^{n-1} x_{\gamma(i)} \right) + x_{\gamma(n)} \right) \\ &\stackrel{10.6}{=} \sum_{i=0}^n x_{\gamma(i)} \\ &\stackrel{n \in \mathcal{S}}{=} \sum_{i=0}^n x_i \end{aligned}$$

so that

$$\begin{aligned} \left( \left( \sum_{i=0}^{n-1} x_{\sigma((n \leftrightarrow_{0,n} k)(i))} \right) + x_{\sigma(n+1)} \right) + x_{n+1} &= \left( \sum_{i=0}^n x_i \right) + x_{n+1} \\ &\stackrel{10.6}{=} \sum_{i=0}^{n+1} x_i \end{aligned}$$

which using 10.3 proves that  $\sum_{i=0}^{n+1} x_{\sigma(i)} = \sum_{i=0}^{n+1} x_i$ . Hence  $n+1 \in \mathcal{S}$ .

So in all case we have that  $n+1 \in \mathcal{S}$  completing the proof by mathematical induction.  $\square$

**Theorem 10.27.** Let  $\langle A, + \rangle$  be a abelian group then if  $n, m \in \mathbb{N}$  with  $n \leq m$  then if  $\{x_i\}_{i \in \{n, \dots, m+1\}}$  is a family in  $A$  we have  $\sum_{i=n}^m (x_{i+1} - x_i) = x_{m+1} - x_n$

**Proof.** We prove this by induction so let  $\mathcal{S}_n = \{k \in \mathbb{N} \mid \text{if } \{x_i\}_{i \in \{n, \dots, (n+k)+1\}} \text{ is a family in } A \text{ then } \sum_{i=n}^{(n+k)} (x_{i+1} - x_i) = x_{(n+k)+1} - x_n\}$  then we have:

1. If  $k=0$  then for  $\{x_i\}_{i \in \{n, \dots, (n+0)+1\}} = \{x_i\}_{i \in \{n, \dots, n+1\}}$  we have  $\sum_{i=n}^n (x_{i+1} - x_i) = x_{n+1} - x_n$  proving that  $0 \in \mathcal{S}_n$
2. If  $k \in \mathcal{S}_n$  then we have if  $\{x_i\}_{i \in \{n, \dots, (n+(k+1))+1\}}$  that  $\sum_{i=n}^{(n+(k+1))} (x_{i+1} - x_i) = \sum_{i=n}^{(n+k)+1} (x_{i+1} - x_i) = \sum_{i=n}^{n+k} (x_{i+1} - x_i) + (x_{((n+k)+1)+1} - x_{((n+k)+1)}) \stackrel{k \in \mathcal{S}_n}{=} (x_{(n+k)+1} - x_n) + (x_{(n+(k+1))+1} - x_{(n+k)+1}) = (x_{(n+(k+1))+1} - x_n)$  proving that  $k+1 \in \mathcal{S}_n$

Using mathematical induction we have then that  $\mathcal{S}_n = \mathbb{N}$  so if  $n, m \in \mathbb{N}$  with  $n \leq m$  then  $k = m - n \in \mathbb{N} = \mathcal{S}_n$  so that  $\sum_{i=n}^m (x_{i+1} - x_i) = \sum_{i=n}^{(n+k)} (x_{i+1} - x_i) = x_{(n+k)+1} - x_n = x_{m+1} - x_n$   $\square$

**Theorem 10.28.** Let  $\langle A, + \rangle$  be a abelian group then if  $n, m \in \mathbb{N}$  with  $n \leq m$  then if  $\{x_i\}_{i \in \{n, \dots, m+1\}}$  is a family in  $A$  we have  $\sum_{i=n}^m (x_i - x_{i+1}) = x_n - x_{m+1}$

**Proof.** We prove this by induction so let  $\mathcal{S}_n = \{k \in \mathbb{N} \mid \text{if } \{x_i\}_{i \in \{n, \dots, (n+k)+1\}} \text{ is a family in } A \text{ then } \sum_{i=n}^{(n+k)} (x_i - x_{i+1}) = x_n - x_{(n+k)+1}\}$  then we have:

1. If  $k=0$  then for  $\{x_i\}_{i \in \{n, \dots, (n+0)+1\}} = \{x_i\}_{i \in \{n, \dots, n+1\}}$  we have  $\sum_{i=n}^n (x_i - x_{i+1}) = x_n - x_{n+1}$  proving that  $0 \in \mathcal{S}_n$
2. If  $k \in \mathcal{S}_n$  then we have if  $\{x_i\}_{i \in \{n, \dots, (n+(k+1))+1\}}$  that  $\sum_{i=n}^{(n+(k+1))} (x_i - x_{i+1}) = \sum_{i=n}^{(n+k)+1} (x_i - x_{i+1}) = \sum_{i=n}^{n+k} (x_i - x_{i+1}) + (x_{((n+k)+1)+1} - x_{((n+k)+1)}) \stackrel{k \in \mathcal{S}_n}{=} (x_n - x_{(n+k)+1}) + (x_{n+(k+1)} - x_{(n+k)+1}) = (x_n - x_{(n+(k+1))+1})$  proving that  $k+1 \in \mathcal{S}_n$

Using mathematical induction we have then that  $\mathcal{S}_n = \mathbb{N}$  so if  $n, m \in \mathbb{N}$  with  $n \leq m$  then  $k = m - n \in \mathbb{N} = \mathcal{S}_n$  so that  $\sum_{i=n}^m (x_i - x_{i+1}) = \sum_{i=n}^{(n+k)} (x_i - x_{i+1}) = x_n - x_{(n+k)+1} = x_n - x_{m+1}$   $\square$

**Theorem 10.29.** Let  $\langle F, +, \cdot \rangle$  be a field,  $\alpha \in F$  and  $\{x_i\}_{i \in \{0, \dots, n\}}$  a family in  $F$  then  $\sum_{i=0}^n (\alpha \cdot x_i) = \alpha \cdot \sum_{i=0}^n x_i$

**Proof.** This is proved by induction so let  $S = \{n \in \mathbb{N} \mid \text{if } \{x_i\}_{i \in \{0, \dots, n\}} \text{ is a family in } F \text{ and } \alpha \in F \text{ then } \sum_{i=0}^n \alpha \cdot x_i = \alpha \cdot \sum_{i=0}^n x_i\}$  then we have

1. If  $n = 0$  then  $\sum_{i=0}^0 \alpha \cdot x_i = \alpha \cdot x_0 = \alpha \cdot \sum_{i=1}^0 x_i$  proving that  $0 \in S$
2. If  $n \in S$  then we have  $\sum_{i=0}^{n+1} \alpha \cdot x_i = \sum_{i=0}^n \alpha \cdot x_i + \alpha \cdot x_{n+1} \stackrel{n \in S}{=} \alpha \cdot \sum_{i=0}^n x_i + \alpha \cdot x_{n+1} = \alpha \cdot (\sum_{i=0}^n x_i + x_{n+1}) = \alpha \cdot \sum_{i=0}^{n+1} x_i$  proving that  $n+1 \in S$

Mathematical induction completes then the theorem.  $\square$

**Theorem 10.30.** Let  $\mathbb{R}$  be the real field,  $n \in \mathbb{N}$  and  $\{c\}_{i \in \{0, \dots, n\}}$  be constant family in  $\mathbb{R}$  then we have that  $\sum_{i=0}^n c = (n+1) \cdot c$

**Proof.** Then we have

1. If  $n = 0$  then  $\sum_{i=0}^0 c = c = 1 \cdot n = (0+1) \cdot n$  proving that  $0 \in S$
2. If  $n \in S$  then  $\sum_{i=0}^{n+1} c = \sum_{i=0}^n c + c \stackrel{n \in S}{=} (n+1) \cdot c + c = ((n+1)+1) \cdot c$  proving that  $n+1 \in S$   $\square$

**Theorem 10.31.** Let  $\mathbb{R}$  (or  $\mathbb{R}_{\mathbb{C}}$ ) be the real field (embedded real field in the complex numbers,  $\{x_i\}_{i \in \{0, \dots, n\}}$  a family in  $\mathbb{R}$  (or  $\mathbb{R}_{\mathbb{C}}$ ) such that  $\forall i \in \{0, \dots, n\}$  we have  $0 \leq x_i$  then  $0 \leq (\leq) \sum_{i=0}^n x_i$  (where we have  $<$  if  $\exists i \in \{0, \dots, n\}$  such that  $0 < x_i$ )

**Proof.** We prove this by induction so let  $S = \{n \in \mathbb{N} \mid \text{if } \{x_i\}_{i \in \{0, \dots, n\}} \text{ is a family of non negative numbers then } 0 \leq (\leq) \sum_{i=0}^n x_i \text{ where we have } < \text{ if at least one of the } x'_i \text{ is } > 0\}$ , we have then

1. if  $n = 0$  then  $0 \leq x_0 = \sum_{i=0}^0 x_i$  and if  $\exists i \in (0, \dots, 0)$  such that  $0 < x_i = x_0$  then we have  $0 < x_0 = \sum_{i=0}^0 x_i$  proving that  $0 \in S$
2. If  $n \in S$  then we have  $\sum_{i=0}^{n+1} x_i = \sum_{i=0}^n x_i + x_{n+1} \geq \sum_{i=0}^n x_i + 0 \geq_{n \in S} 0$ . If now  $\exists i \in \{0, \dots, n+1\}$  such that  $x_i > 0$  then we have two cases to consider
  - a. ( $i \in \{0, \dots, n\}$ ) then as  $n \in S$  we have  $0 < \sum_{i=0}^n x_i$  so that  $\sum_{i=0}^{n+1} x_i \geq \sum_{i=0}^n x_i > 0$
  - b. ( $i = n+1$ ) then we have  $\sum_{i=0}^{n+1} x_i = \sum_{i=0}^n x_i + x_{n+1} > \sum_{i=0}^n x_i + 0 \geq_{n \in S} 0$

This proves that  $n+1 \in S$

By mathematical induction we have then that  $S = \mathbb{N}$  proving the theorem  $\square$

**Corollary 10.32.** Let  $\mathbb{R}$  (or  $\mathbb{R}_{\mathbb{C}}$ ) be the real field (embedded real field in the complex numbers,  $\{x_i\}_{i \in \{0, \dots, n\}}$ ,  $\{y_i\}_{i \in \{0, \dots, n\}}$  a family in  $\mathbb{R}$  (or  $\mathbb{R}_{\mathbb{C}}$ ) such that  $\forall i \in \{0, \dots, n\}$  we have  $x_i \leq y_i$  then  $\sum_{i=0}^n x_i \leq (\leq) \sum_{i=0}^n y_i$  (where we have  $<$  if  $\exists i \in \{0, \dots, n\}$  such that  $x_i < y_i$ )

**Proof.** Take the family  $\{y_i + (-x_i)\}_{i \in \{0, \dots, n\}}$  then  $\forall i \in \{0, \dots, n\}$  we have  $x_i + (-y_i) \geq 0$  so by 10.31 we have

$$\sum_{i=0}^n (y_i + (-x_i)) \geq (0) > \text{if we have one } y_i > x_i \quad (10.4)$$

Now  $\sum_{i=0}^n (y_i + (-x_i)) \stackrel{10.11}{=} \sum_{i=0}^n y_i + \sum_{i=0}^n (-x_i) \stackrel{10.29}{=} \sum_{i=0}^n y_i - \sum_{i=0}^n x_i$  which together with 10.4 we have  $\sum_{i=0}^n y_i \geq (0) \sum_{i=0}^n x_i > \text{if we have one } y_i > x_i$   $\square$

**Definition 10.33. (Support)** If  $\langle A, + \rangle$  is a semi-group and  $\{x_i\}_{i \in I}$  a family of elements in  $A$ . Then  $\text{support}(\{x_i\}_{i \in I}) = \{i \in I \mid x_i \neq 0\}$  where 0 is the neutral element

**Definition 10.34. (Finite support)** If  $\langle A, + \rangle$  is a semi-group and  $\{x_i\}_{i \in I}$  a family of elements in  $A$  has finite support if  $\text{support}(\{x_i\}_{i \in I})$  is finite.

**Example 10.35.** If  $\langle A, + \rangle$  is a semi-group and  $\{x_i\}_{i \in I}$  is a family of elements in  $A$  where  $I$  is finite then as  $\text{support}(\{x_i\}_{i \in I}) \subseteq I$  we have by 5.45 that  $\text{support}(\{x_i\}_{i \in I})$  is finite and thus  $\{x_i\}_{i \in I}$  has finite support.

**Definition 10.36. (generalized sum)** Let  $\langle A, + \rangle$  be a abelian semi-group and  $\{x_i\}_{i \in I}$  a family of elements in  $A$  with finite support. Then  $\sum_{i \in I} x_i$  is defined to be

1. If  $\text{support}(\{x_i\}_{i \in I}) = \emptyset$  then  $\sum_{i \in I} x_i = 0$  (neutral element of the semi-group)
2. If  $\text{support}(\{x_i\}_{i \in I}) \neq \emptyset$  then  $\sum_{i \in I} x_i = \sum_{i=0}^{n-1} x_{\beta(i)}$  where  $\beta: \{0, \dots, n-1\} = S_n \rightarrow \text{support}(\{x_i\}_{i \in I})$  is a bijection (which must exists because of finiteness of the support)

We must of course prove that (2) of the definition is independent of the choosen bijection

**Proof.** As  $\text{support}(\{x_i\}_{i \in I})$  is finite there exists a unique (see 5.42)  $n \in \mathbb{N}$  such that  $n = S_n \approx \text{support}(\{x_i\}_{i \in I})$ , which as  $\text{support}(\{x_i\}_{i \in I}) \neq \emptyset$  means that  $n > 0$ . If  $\beta': \{0, \dots, n-1\} \rightarrow I$  is another bijection then we have that  $\tau = \beta'^{-1} \circ \beta: \{0, \dots, n-1\} \rightarrow \{0, \dots, n-1\}$  is also a bijection and  $\sum_{i=0}^{n-1} x_{\beta'(i)} = \sum_{i=0}^{n-1} x_{\beta'((\beta'^{-1} \circ \beta)(i))} = \sum_{i=0}^{n-1} x_{\beta'(\beta'^{-1}(\beta(i)))} = \sum_{i=0}^{n-1} x_{\beta(i)}$   $\square$

**Example 10.37.** Let  $\langle A, + \rangle$  be a abelian semi-group and  $\{x_i\}_{i \in I} \subseteq A$  such that  $\forall i \in I$  we have  $x_i = 0$  then  $\sum_{i \in I} x_i = 0$

**Proof.** By the hypothese we have that  $\sup(\{x_i\}_{i \in I}) = \emptyset$  so that by definition  $\sum_{i \in I} x_i = 0$   $\square$

We have for a multiplicative group a equivalent definition, only we note then the neutral element with 1 and use  $\prod_{i \in I} x_i$  instead of the summation symbol all the following theorems are valid for a multiplicative group if we do this substitution. So we use then the following definition:

**Definition 10.38. (generalized product)** Let  $\langle A, + \rangle$  be a multiplicative abelian semi-group and  $\{x_i\}_{i \in I}$  a family of elements in  $A$  with finite support  $[\text{support}(A) = \{x \in A \mid x \neq 1\}]$  Then  $\prod_{i \in I} x_i$  is defined to be

1. If  $\text{support}(\{x_i\}_{i \in I}) = \emptyset$  then  $\prod_{i \in I} x_i = 1$  (neutral element of the semi-group)
2. If  $\text{support}(\{x_i\}_{i \in I}) \neq \emptyset$  then  $\prod_{i \in I} x_i = \prod_{i=0}^{n-1} x_{\beta(i)}$  where  $\beta: \{0, \dots, n-1\} = S_n \rightarrow \text{support}(\{x_i\}_{i \in I})$  is a bijection (which must exists because of finiteness of the support)

**Lemma 10.39.** Let  $\langle X, + \rangle$  be a abelian semi-group,  $I$  a finite set with  $n = \#(I)$  and  $\{x_i\}_{i \in I} \subseteq X$  a finite family of elements in  $X$  then

1. If  $I = \emptyset$  then  $\sum_{i \in I} x_i = 0$
2. If  $I \neq \emptyset$  then  $\sum_{i \in I} x_i = \sum_{i=0}^n x_{\beta(i)}$  where  $n = \#(I) - 1$  and  $\beta: \{0, \dots, n\} \rightarrow I$  is a bijection.
3. For every bijection  $\alpha: \{0, \dots, \#(I) - 1\} \rightarrow I$  we have that  $\sum_{i \in I} x_i = \sum_{i=0}^n x_{\alpha(i)}$

**Proof.**

1. If  $I = \emptyset$  then  $\sup(\{x_i\}_{i \in I}) = \emptyset$  so that  $\sum_{i \in I} x_i = 0$  by definition.
2. Let  $S = \text{support}(\{x_i\}_{i \in I}) \subseteq I$  then  $S$  is finite (so that  $\sum_{i \in I} x_i$  is defined) and also  $T = I \setminus S = \{i \in I \mid x_i = 0\}$  is finite. We have then the following cases for  $S$

$S = \emptyset$ . then  $T = I \setminus S = I \neq \emptyset$  and thus there exists a bijection  $\beta: \{0, \dots, n\} \rightarrow I$  where  $n = \#(I) - 1$  then  $\sum_{i=0}^n x_i = \sum_{i=0}^n 0 = 0$   $\stackrel{\text{def}}{=} \sum_{i \in I} x_i$

$S = I$ . then as  $I \neq \emptyset$  there exists a bijection  $\beta: \{0, \dots, n\} \rightarrow I = S$  where  $n = \#(I) - 1 = \#(S) - 1$  and  $\sum_{i \in I} x_i = \sum_{i=0}^n x_{\beta(i)}$

$\emptyset \neq S \subset I$ . then  $S, T \neq \emptyset$  then there exists bijections  $\beta_1: \{0, \dots, n_1\} \rightarrow S$  and  $\beta_2: \{0, \dots, n_2\} \rightarrow T$  where  $n_1 = \#(S) - 1, n_2 = \#(T) - 1$ . Define then  $\beta: \{0, \dots, n_1 + n_2 + 1\} \rightarrow I$  by  $\beta(i) = \begin{cases} \beta_1(i) & \text{if } i \in \{0, \dots, n_1\} \\ \beta_2(i - n_1 - 1) & \text{if } i \in \{n_1 + 1, \dots, n_1 + n_2 + 1\} \end{cases}$  then we prove that  $\beta$  is a bijection

**injectivity.** Let  $i, j \in \{0, \dots, n_1 + n_2\}$  with  $\beta(i) = \beta(j)$ . For  $i$  we must then consider the following cases:

$i \in \{0, \dots, n_1\}$ . then we can not have that  $j \in \{n_1 + 1, \dots, n_1 + n_2\}$  for then  $\beta(i) = \beta_1(i) \in S$  and  $\beta(j) = \beta_2(j - n_1 - 1) \in T$  which as  $\beta(i) = \beta(j)$  would mean that  $\emptyset \neq S \cap T = \emptyset$  a contradiction. So  $j \in \{0, \dots, n_1\}$  and then  $\beta_1(i) = \beta(i) = \beta(j) = \beta_1(j) \stackrel{\beta_1 \text{ is a bijection}}{\Rightarrow} i = j$

$i \in \{n_1 + 1, \dots, n_1 + n_2 + 1\}$ . then we can not have that  $j \in \{0, \dots, n_1\}$  for then  $\beta(i) = \beta_2(i - n_1 - 1) \in S$  and  $\beta(j) = \beta_1(j) \in T$  which as  $\beta(i) = \beta(j)$  would mean that  $\emptyset \neq S \cap T = \emptyset$  a contradiction. So  $j \in \{n_1 + 1, \dots, n_1 + n_2\}$  and then  $\beta_2(i) = \beta(i) = \beta(j) = \beta_2(j) \xrightarrow{\beta_2 \text{ is a bijection}} i = j$

**surjectivity.** If  $x \in I$  then either  $x \in S$  and then  $\exists i \in \{0, \dots, n_1\} \subseteq \{0, \dots, n_1 + n_2\}$  such that  $x = \beta_1(i) = \beta(i)$  or  $x \in T$  and then there exists a  $i \in \{0, \dots, n_2\}$  such that  $\beta_2(i) = x$  hence for  $i + n_1 + 1 \in \{n_1, \dots, n_1 + n_2 + 1\}$  we have  $\beta(i + n_1 + 1) = \beta_2(i) = x$

Finally  $\sum_{i=0}^{n_1+n_2+1} x_{\beta(i)} \stackrel{10.17}{=} \sum_{i=0}^{n_1} x_{\beta(i)} + \sum_{i=n_1+1}^{n_1+n_2+1} x_{\beta(i)} = \sum_{i=0}^{n_1} x_{\beta_1(i)} + \sum_{i=0}^{n_2} x_{\beta(i+n_1+1)} = \sum_{i=0}^{n_1} x_{\beta_1(i)} + \sum_{i=0}^{n_2} x_{\beta_2(i)} = \sum_{i \in I} x_i + 0$  hence if we take  $n = n_1 + n_2 + 1$  we have a bijection  $\beta: \{0, \dots, n\} \rightarrow I$  such that  $\sum_{i=0}^n x_{\beta(i)} = \sum_{i \in I} x_i$

3. If  $\alpha: \{0, \dots, n\} \rightarrow I$  (where  $n = \#(I) - 1$ ) is another bijection then  $\beta^{-1} \circ \alpha: \{0, \dots, \#(I) - 1\}$  is a bijection and we have  $\sum_{i \in I} x_i = \sum_{i=0}^n x_{\beta(i)} \stackrel{10.26}{=} \sum_{i=0}^n x_{\beta((\beta^{-1} \circ \alpha)(i))} = \sum_{i=0}^n x_{\alpha(i)}$   $\square$

**Lemma 10.40.** Let  $\langle A, + \rangle$  be a abelian semi-group and  $\{x_i\}_{i \in \{n, \dots, m\}}$  a family in  $A$  where  $n \leq m$ ,  $n, m \in \mathbb{N}$  then  $\sum_{i \in \{n, \dots, m\}} x_i = \sum_{i=n}^m x_i$

**Proof.** Let  $\beta: \{0, \dots, m - n\} \rightarrow \{n, \dots, m\}$  be the bijection defined by  $\beta(i) = i + n$  then using 10.39 we have that  $\sum_{i \in \{n, \dots, m\}} x_i = \sum_{i=0}^{m-n} x_{\beta(i)} = \sum_{i=0}^{m-n} x_{i+n} \stackrel{10.13}{=} \sum_{i=n}^m x_i$   $\square$

**Theorem 10.41.** Let  $\langle A, + \rangle$  be a abelian semi-group and  $\{x_i\}_{i \in I}$  a family in  $A$  with finite support then if  $\text{support}(\{x_i\}_{i \in I}) \subseteq J \subseteq I$  we have  $\sum_{i \in I} x_i = \sum_{i \in J} x_i$

**Proof.** If  $k \in \text{support}(\{x_i\}_{i \in I}) \Rightarrow x_k \neq 0 \xrightarrow{\text{support}(\{x_i\}_{i \in I} \subseteq J)} x_k \neq 0 \wedge k \in J \Rightarrow k \in \text{support}(\{x_i\}_{i \in J})$ , also if  $k \in \text{support}(\{x_i\}_{i \in J}) \Rightarrow k \in J \wedge x_k \neq 0 \xrightarrow{J \subseteq I} k \in \text{support}(\{x_i\}_{i \in I})$ , this proves that  $\text{support}(\{x_i\}_{i \in I}) = \text{support}(\{x_i\}_{i \in J})$ . So if  $b: \{0, \dots, n - 1\} = S_n \rightarrow \text{support}(\{x_i\}_{i \in I}) = \text{support}(\{x_i\}_{i \in J})$  is a bijection then  $\sum_{i \in I} x_i = \sum_{i=0}^{n-1} x_{b_i} = \sum_{i \in J} x_i$ .  $\square$

**Note 10.42.** If  $\langle A, \cdot \rangle$  is a abelian semi-group, where the operator is interpreted as a product, and  $\{x_i\}_{i \in I}$  has finite support then  $\sum_{i \in I} x_i$  is written as  $\prod_{i \in I} x_i$ . In general we use for a product 1 as the unit so if  $\text{support}(\{x_i\}_{i \in I}) = \emptyset$  then  $\prod_{i \in I} x_i = 1$

**Lemma 10.43.** Let  $\langle A, + \rangle$  be a abelian semi-group,  $\{x_i\}_{i \in I}$  a family of elements in  $A$  with finite support and  $f: J \rightarrow I$  a bijection then  $\{x_{f(i)}\}_{i \in J}$  has finite support given by  $\text{support}(\{x_{f(i)}\}_{i \in J}) = f^{-1}(\text{support}(\{x_i\}_{i \in I}))$  and  $\sum_{i \in J} x_{f(i)} = \sum_{i \in I} x_i = \sum_{i \in f(J)} x_i$

**Proof.** First we prove the assertion about the support

$$\begin{aligned} j \in \text{support}(\{x_{f_i}\}_{i \in J}) &\Leftrightarrow x_{f_j} \neq 0 \\ &\Leftrightarrow x(f(j)) \neq 0 \\ &\Leftrightarrow f(j) \in \text{support}(\{x_i\}_{i \in I}) \\ &\Leftrightarrow j \in f^{-1}(\text{support}(\{x_i\}_{i \in I})) \end{aligned}$$

As  $f: J \rightarrow I$  is a bijection we have that  $f|_{f^{-1}(\text{support}(\{x_i\}_{i \in I}))}: f^{-1}(\text{support}(\{x_i\}_{i \in I})) = \text{support}(\{x_{f_i}\}_{i \in J}) \rightarrow \text{support}(\{x_i\}_{i \in I})$  is a bijection and thus  $\text{support}(\{x_{f_i}\}_{i \in J}) \approx \text{support}(\{x_i\}_{i \in I})$  so  $\text{support}(\{x_{f_i}\}_{i \in J})$  is also finite and there exists a  $n \in \mathbb{N}$  such that  $S_n \approx \text{support}(\{x_{f_i}\}_{i \in J})$  and  $S_n \approx \text{support}(\{x_i\}_{i \in I})$ . So if  $h: S_n \rightarrow \text{support}(\{x_{f_i}\}_{i \in J})$  is a bijection then  $\sum_{j \in J} x_{f_j} = \sum_{j=0}^{n-1} x_{f_{h_j}}$ . But also  $g = f|_{f^{-1}(\text{support}(\{x_i\}_{i \in I}))} \circ h: S_n \rightarrow \text{support}(\{x_i\}_{i \in I})$  is a bijection so that  $\sum_{i \in I} x_i = \sum_{i=0}^{n-1} x_{g_i} = \sum_{i=0}^{n-1} g_{f_{h_i}} = \sum_{j \in J} x_{f_j}$   $\square$

**Lemma 10.44.** If  $\langle A, + \rangle$  is a abelian semi-group  $\{x_i\}_{i \in I}$  a family of elements with finite support and  $J, K$  are sets such that  $J \cap K = \emptyset$  and  $I = J \cup K$  then  $\{x_i\}_{i \in J}$  and  $\{x_i\}_{i \in K}$  have finite support and  $\sum_{i \in I} x_i = (\sum_{i \in J} x_i) + (\sum_{i \in K} x_i)$

**Proof.** First as  $J \cup K = I \Rightarrow J, K \subseteq I$  now if  $i \in \text{support}(\{x_i\}_{i \in J}) \Rightarrow i \in J \wedge x_i \neq 0 \Rightarrow i \in I \wedge x_i \neq 0 \Rightarrow i \in \text{support}(\{x_i\}_{i \in I}) \Rightarrow \text{support}(\{x_i\}_{i \in J}) \subseteq \text{support}(\{x_i\}_{i \in I})$ , similar we have  $\text{support}(\{x_i\}_{i \in K}) \subseteq \text{support}(\{x_i\}_{i \in I})$ . Using 5.45 we have then that  $\text{support}(\{x_i\}_{i \in J})$  and  $\text{support}(\{x_i\}_{i \in K})$  are finite and thus that  $\{x_i\}_{i \in J}$  and  $\{x_i\}_{i \in K}$  have finite support. Also we have  $\text{support}(\{x_i\}_{i \in J}) \cup \text{support}(\{x_i\}_{i \in K}) \subseteq \text{support}(\{x_i\}_{i \in I})$ . If  $i \in \text{support}(\{x_i\}_{i \in I}) \Rightarrow i \in I \wedge x_i \neq 0 \underset{I=J \cup K}{\Rightarrow} (i \in J \vee i \in K) \wedge x_i \neq 0 \Rightarrow (i \in J \wedge x_i \neq 0) \vee (i \in K \wedge x_i \neq 0) \Rightarrow i \in \text{support}(\{x_i\}_{i \in J}) \cup \text{support}(\{x_i\}_{i \in K}) \Rightarrow \text{support}(\{x_i\}_{i \in I}) \subseteq \text{support}(\{x_i\}_{i \in J}) \cup \text{support}(\{x_i\}_{i \in K})$  and thus we have  $\text{support}(\{x_i\}_{i \in J}) \cup \text{support}(\{x_i\}_{i \in K}) = \text{support}(\{x_i\}_{i \in I})$ . Also if  $i \in \text{support}(\{x_i\}_{i \in J}) \cap \text{support}(\{x_i\}_{i \in K}) \Rightarrow (i \in J \wedge x_i \neq 0) \wedge (i \in K \wedge x_i \neq 0) \Rightarrow i \in I \wedge i \in J \Rightarrow i \in I \cap J = \emptyset$  which is a contradiction, so we must have  $\text{support}(\{x_i\}_{i \in J}) \cap \text{support}(\{x_i\}_{i \in K}) = \emptyset$  We have then the following cases to consider:

1. (**support**( $\{x_i\}_{i \in J}$ ), **support**( $\{x_i\}_{i \in K}$ ) =  $\emptyset$ ) in this case we have  $\text{support}(\{x_i\}_{i \in I}) = \emptyset$  and thus  $\sum_{i \in I} x_i = 0 = 0 + 0 = (\sum_{i \in J} x_i) + (\sum_{i \in K} x_i)$
2. (**support**( $\{x_i\}_{i \in J}$ ) =  $\emptyset$ , **support**( $\{x_i\}_{i \in K}$ )  $\neq \emptyset$ ) we must then have  $\text{support}(\{x_i\}_{i \in K}) = \text{support}(\{x_i\}_{i \in I})$  and if  $b: \{0, \dots, n-1\} \rightarrow \text{support}(\{x_i\}_{i \in I}) = \text{support}(\{x_i\}_{i \in K})$  is a bijection then we have  $\sum_{i \in I} x_i = \sum_{i=0}^{n-1} x_{b_i} = \sum_{i \in K} x_i = 0 + (\sum_{i \in K} x_i) = (\sum_{i \in J} x_i) + (\sum_{i \in K} x_i)$
3. (**support**( $\{x_i\}_{i \in J}$ )  $\neq \emptyset$ , **support**( $\{x_i\}_{i \in K}$ ) =  $\emptyset$ ) we must then have  $\text{support}(\{x_i\}_{i \in J}) = \text{support}(\{x_i\}_{i \in I})$  and if  $b: \{0, \dots, n-1\} \rightarrow \text{support}(\{x_i\}_{i \in I}) = \text{support}(\{x_i\}_{i \in J})$  is a bijection then we have  $\sum_{i \in I} x_i = \sum_{i=0}^{n-1} x_{b_i} = \sum_{i \in J} x_i = (\sum_{i \in J} x_i) + 0 = (\sum_{i \in J} x_i) + (\sum_{i \in K} x_i)$
4. (**support**( $\{x_i\}_{i \in J}$ ), **support**( $\{x_i\}_{i \in K}$ )  $\neq \emptyset$ ) then as the supports are finite there exists bijections  $f': \{0, \dots, n_1-1\} \rightarrow \text{support}(\{x_i\}_{i \in J})$ ,  $g: \{0, \dots, n_2-1\} \rightarrow \text{support}(\{x_i\}_{i \in K})$ . Define now  $f: \{n_2, \dots, (n_1+n_2)-1\} \rightarrow \text{support}(\{x_i\}_{i \in J})$  by  $i \rightarrow f(i) = f'(i-n_2)$  [which is well defined as if  $i \in \{n_2, \dots, (n_1+n_2)-1\} \Rightarrow n_2 \leq i \leq n_2 + (n_1-1) \Rightarrow 0 \leq i - n_2 \leq n_1-1$ ] then  $f$  is a bijection.
  - a. (**injectivity**) if  $f(i) = f(j) \Rightarrow f'(i-n_2) = f'(j-n_2) \underset{f' \text{ is bijective}}{\Rightarrow} i - n_2 = j - n_2 \Rightarrow i = j$
  - b. (**surjectivity**) if  $j \in \text{support}(\{x_i\}_{i \in J})$  then as  $f'$  is a bijection there exists a  $i \in \{0, \dots, n_1-1\}$  such that  $f'(i) = j$ . From  $0 \leq i \leq n_1-1$  we have  $n_2 \leq i + n_2 \leq (n_1+n_2)-1$  and  $f(i+n_2) = f'((i+n_2)-n_2) = f'(i) = j$

Using  $\{0, \dots, n_2-1\} \cap \{n_2, \dots, (n_1+n_2)-1\} = \emptyset$ ,  $\text{support}(\{x_i\}_{i \in J}) \cap \text{support}(\{x_i\}_{i \in K}) = \emptyset$  and 2.46 we have that  $g \cup f: \{0, \dots, n_2-1\} \cup \{n_2, \dots, (n_1+n_2)-1\} = \{0, \dots, (n_1+n_2)-1\} \rightarrow \text{support}(\{x_i\}_{i \in J}) \cup \text{support}(\{x_i\}_{i \in K}) = \text{support}(\{x_i\}_{i \in I})$  is a bijection. Now  $\sum_{i \in I} x_i = \sum_{i=0}^{(n_1+n_2)-1} x_{(g \cup f)_i} \underset{10.17}{=} (\sum_{i=0}^{n_2-1} x_{(g \cup f)_i}) + (\sum_{i=n_2}^{(n_1+n_2)-1} x_{(g \cup f)_i}) = (\sum_{i=0}^{n_2-1} x_{(g \cup f)(i)}) + (\sum_{i=n_2}^{(n_1+n_2)-1} x_{(g \cup f)(i)}) = (\sum_{i=0}^{n_2-1} x_{g(i)}) + (\sum_{i=n_2}^{(n_1+n_2)-1} x_{f(i)}) = (\sum_{i=0}^{n_2-1} x_{g_i}) + (\sum_{i=n_2}^{(n_1+n_2)-1} x_{f'(i-n_2)}) = (\sum_{i \in J} x_j) + (\sum_{i=0}^{n_1-1} x_{f'((i-n_2)+n_2)}) = (\sum_{i \in J} x_j) + (\sum_{i=0}^{n_1-1} x_{f'_i}) = (\sum_{i \in J} x_i) + (\sum_{i \in K} x_i)$   $\square$

**Theorem 10.45.** Let  $\langle A, + \rangle$  be a abelian semi-group,  $\{x_i\}_{i \in I}$  a family of elements in  $A$  with finite support and  $\{I_i\}_{i \in \{0, \dots, n\}}$ ,  $n \in \mathbb{N}_0$  a family of sets in  $\mathcal{P}(I)$  such that  $\bigcup_{i \in \{0, \dots, n\}} I_i = I$  and  $\forall k, l \in \{0, \dots, n\}$  we have  $I_k \cap I_l = \emptyset$  then we have that  $\sum_{i \in I} x_i = \sum_{i=0}^n (\sum_{j \in I_i} x_j)$

**Proof.** We prove this by mathematical induction on  $n$  (see 4.77) so let  $X = \{n \in \{1, \dots\} = \mathbb{N}_0 \mid \{x_i\}_{i \in I}$  is a family of elements in  $A$  with finite support and  $\{I_i\}_{i \in \{1, \dots, n\}}$  is a family of subsets of  $I$  such that  $\bigcup_{i \in \{0, \dots, n\}} I_i = I$  and  $\forall k, l \in \{0, \dots, n\}$  we have  $I_k \cap I_l = \emptyset \Rightarrow \sum_{i \in I} x_i = \sum_{i=0}^n (\sum_{j \in I_i} x_j)$  then we have:

1. If  $n = 1$  then  $I = I_0 \cup I_1$  and  $I_0 \cap I_1 = \emptyset$  and using the previous lemma we have  $\sum_{i \in I} x_i = (\sum_{i \in I_0} x_i) + (\sum_{i \in I_1} x_i) = \sum_{i=0}^1 (\sum_{j \in I_i} x_j)$  so  $1 \in X$

2. Assume that  $n \in X$  then if  $\{x_i\}_{i \in I}$  is a family of elements with finite support is such that  $I = \bigcup_{i \in \{0, \dots, n+1\}} I_i$  then  $I = (\bigcup_{i \in \{0, \dots, n\}} I_i) \cup I_{n+1}$  (see 1.101) and we have then

$$\begin{aligned} \sum_{i \in I} x_i &\stackrel{\text{previous lemma}}{=} \sum_{i \in \bigcup_{i \in \{0, \dots, n\}} I_i} x_i + \sum_{i \in I_{n+1}} x_i \\ &\stackrel{n \in X}{=} \left( \sum_{i=0}^n \left( \sum_{j \in I_i} x_j \right) \right) + \sum_{i \in I_{n+1}} x_i \\ &= \sum_{i=0}^{n+1} \left( \sum_{j \in I_i} x_j \right) \end{aligned}$$

and thus  $n+1 \in X$

Using 4.77 we have that  $X = \mathbb{N}_0$  □

**Theorem 10.46.** Let  $\langle A, + \rangle$  be a abelian semi-group,  $\{x_i\}_{i \in I}$  a family of elements in  $A$  with finite support and  $\{I_j\}_{j \in J}$  a family such that  $\forall i, j \in J$  with  $j \neq i$  we have  $I_i \cap I_j = \emptyset$  and  $\bigcup_{j \in J} I_j = I$  then we have  $\sum_{i \in I} x_i = \sum_{j \in J} (\sum_{i \in I_j} x_i)$

**Proof.** Define

$$J_0 = \{j \in J \mid I_j \cap \text{support}(\{x_i\}_{i \in I}) \neq \emptyset\} \subseteq J \quad (10.5)$$

then trivially  $\bigcup_{j \in J_0} I_j \cap \text{support}(\{x_i\}_{i \in I}) \subseteq \text{support}(\{x_i\}_{i \in I})$ , also if  $i \in \text{support}(\{x_i\}_{i \in I})$   $I = \bigcup_{j \in J} I_j \Rightarrow \exists j \in J$  with  $i \in I_j \cap \text{support}(\{x_i\}_{i \in I}) \Rightarrow I_j \cap \text{support}(\{x_i\}_{i \in I}) \neq \emptyset \Rightarrow j \in J_0 \Rightarrow i \in \bigcup_{j \in J_0} I_j \cap \text{support}(\{x_i\}_{i \in I})$  giving that

$$\bigcup_{j \in J_0} I_j \cap \text{support}(\{x_i\}_{i \in I}) = \text{support}(\{x_i\}_{i \in I}) \quad (10.6)$$

Now as  $\forall j \in J_0$  we have by definition that  $I_j \cap \text{support}(\{x_i\}_{i \in I}) \neq \emptyset$  we can use the axiom of choice (see 2.210) to find a function  $c: J_0 \rightarrow \bigcup_{j \in J_0} I_j \cap \text{support}(\{x_i\}_{i \in I}) = \text{support}(\{x_i\}_{i \in I})$  such that  $\forall j \in J_0$  we have  $c(j) \in I_j \cap \text{support}(\{x_i\}_{i \in I})$ . Now if  $j_1, j_2 \in J_0$  is such that  $c(j_1) = c(j_2)$  then  $c(j_1), c(j_2) \in I_{j_1} \cap \text{support}(\{x_i\}_{i \in I}) \cap I_{j_2} \cap \text{support}(\{x_i\}_{i \in I}) \Rightarrow c(j_1), c(j_2) \in I_{j_1} \cap I_{j_2} \stackrel{I_j \text{ are pairwise disjoint}}{=} c(j_1) = c(j_2)$  proving that  $c$  is a injection. So  $J_0 \approx c(J_0) \subseteq \text{support}(\{x_i\}_{i \in I})$  which is finite and thus by 5.36 we have that

$$J_0 \text{ must be finite.} \quad (10.7)$$

As  $J_0$  is finite there exists a  $b: \{0, \dots, n-1\} \rightarrow J_0$  which is a bijection. Also as  $I_j \cap \text{support}(\{x_i\}_{i \in I}) \subseteq \text{support}(\{x_i\}_{i \in I})$  which is finite we have that  $S_i = \sum_{i \in I_j \cap \text{support}(\{x_i\}_{i \in I})} x_i$  is well defined and we have also that  $\sum_{j \in J_0} S_i$  is well defined. We have then  $\sum_{j \in J_0} S_i = \sum_{j \in J_0} (\sum_{i \in I_j \cap \text{support}(\{x_i\}_{i \in I})} x_i) \stackrel{\text{definition}}{=} \sum_{j=0}^{n-1} (\sum_{i \in I_{b(j)} \cap \text{support}(\{x_i\}_{i \in I})} x_i) = \sum_{i \in \text{support}(\{x_i\}_{i \in I})} x_i$  (using 10.45, 10.5 and the fact that  $I_{b(j)} \cap \text{support}(\{x_i\}_{i \in I})$  are disjoints (as  $b$  is a bijection). So we have proved that

$$\sum_{j \in J_0} \left( \sum_{i \in I_j \cap \text{support}(\{x_i\}_{i \in I})} x_i \right) = \sum_{i \in \text{support}(\{x_i\}_{i \in I})} x_i \quad (10.8)$$

Now if  $k \in \text{support}(\{\sum_{j \in I_i \cap \text{support}(\{x_i\}_{i \in I})} x_j\}_{i \in J_0})$  then  $\sum_{j \in I_k \cap \text{support}(\{x_i\}_{i \in I})} x_i \neq 0 \Rightarrow \text{support}(\{x_i\}_{i \in I_k \cap \text{support}(\{x_i\}_{i \in I})}) \neq \emptyset \Rightarrow \exists l \in I_k \cap \text{support}(\{x_i\}_{i \in I}) \mid x_l \neq 0 \Rightarrow I_k \cap \text{support}(\{x_i\}_{i \in I}) \neq \emptyset \Rightarrow k \in J_0$  so we have

$$\text{support} \left( \left\{ \sum_{j \in I_i \cap \text{support}(\{x_i\}_{i \in I})} x_j \right\}_{i \in J_0} \right) \subseteq J_0 \subseteq I \quad (10.9)$$

so using 10.41 and 10.8,10.9 twice we have

$$\sum_{j \in I} \left( \sum_{j \in I_i \cap \text{support}(\{x_i\}_{i \in I})} x_i \right) = \sum_{i \in I} x_i \quad (10.10)$$

also if  $l \in \text{support}(\{x_i\}_{i \in I_k}) \Rightarrow l \in I_k \wedge x_l \neq 0 \Rightarrow l \in I_k \cap \text{support}(\{x_i\}_{i \in I}) \Rightarrow l \in I_k$  giving

$$\text{support}(\{x_i\}_{i \in I_k}) \subseteq I_k \cap \text{support}(\{x_i\}_{i \in I}) \subseteq I_k \quad (10.11)$$

so that we can use 10.10,10.11 and 10.41 to find finally

$$\sum_{j \in J} \left( \sum_{i \in I_j} x_i \right) = \sum_{i \in I} x_i \quad \square$$

**Corollary 10.47.** Let  $\langle A, + \rangle$  be a abelian semi-group,  $I, J$  sets and  $\{x_{i,j}\}_{(i,j) \in I \times J}$  a family of elements in  $A$  with finite support then  $\sum_{(i,j) \in I \times J} x_{i,j} = \sum_{i \in I} (\sum_{j \in J} x_{i,j}) = \sum_{j \in J} (\sum_{i \in I} x_{i,j})$

**Proof.**

1. Define  $\forall i \in I \ J_i = \{i\} \times J$  then we have clearly  $I \times J = \bigcup_{i \in I} J_i$  and if  $i, j \in I$  with  $i \neq j$  we have if  $(n, m) \in J_i \cap J_j \Rightarrow n = i, m = j \Rightarrow i = j$  contradicting  $i \neq j$  so we have  $J_i \cap J_j = \emptyset$ . Using then the previous theorem 10.46 we have that  $\sum_{(i,j) \in I \times J} x_{i,j} = \sum_{i \in I} (\sum_{(k,l) \in J_i} x_{k,l})$ . Define now the bijection  $b_i: J \rightarrow J_i$  by  $j \rightarrow b_i(j) = (i, j)$  then we have by 10.43 that  $\sum_{(k,l) \in J_i} x_{k,l} = \sum_{l \in J} x_{b_i(k)} = \sum_{l \in J} x_{i,l}$  giving that  $\sum_{(i,j) \in I \times J} x_{i,j} = \sum_{i \in I} (\sum_{j \in J} x_{i,j})$ .
2. Define  $\forall j \in J \ I_j = I \times \{j\}$  then we have clearly  $I \times J = \bigcup_{j \in J} I_j$  and if  $i, j \in J$  with  $i \neq j$  we have if  $(n, m) \in I_i \cap I_j \Rightarrow n = i, m = j \Rightarrow i = j$  contradicting  $i \neq j$  so we have  $I_i \cap I_j = \emptyset$ . Using then the previous theorem 10.46 we have that  $\sum_{(i,j) \in I \times J} x_{i,j} = \sum_{j \in J} (\sum_{(k,l) \in I_j} x_{k,l})$ . Define now the bijection  $h_i: I \rightarrow I_i$  by  $j \rightarrow h_i(j) = (j, i)$  then we have by 10.43 that  $\sum_{(k,l) \in I_j} x_{k,l} = \sum_{l \in I} x_{h_i(l)} = \sum_{l \in I} x_{l,i}$  giving that  $\sum_{(i,j) \in I \times J} x_{i,j} = \sum_{j \in J} (\sum_{i \in I} x_{i,j})$ .  $\square$

**Theorem 10.48.** Let  $\langle A, + \rangle$  be a abelian semi-group,  $n \in \mathbb{N}_0 \setminus \{1\}$ ,  $\{N_i\}_{i \in \{1, \dots, n\}}$  such that  $N_i$  is finite and non empty and  $f: \prod_{i \in \{1, \dots, n\}} A_i \rightarrow A$  then  $\sum_{\gamma \in \prod_{i \in \{1, \dots, n\}} A_i} f(\gamma) = \sum_{\gamma \in \prod_{i \in \{1, \dots, n\}} A_i} f(\gamma_1, \dots, \gamma_n) = \sum_{\gamma \in \prod_{i \in \{1, \dots, n-1\}} A_i} (\sum_{i \in A_n} f(\lambda_1, \dots, \gamma_n, i))$

**Proof.** Define  $\{x_{(i,j)}\}_{(i,j) \in (\prod_{i \in \{1, \dots, n-1\}} A_i) \times A_n}$  by  $x_{(i,j)} = f(i_1, \dots, i_{n-1}, j)$  then we have

$$\begin{aligned} \sum_{(i,j) \in (\prod_{i \in \{1, \dots, n-1\}} A_i) \times A_n} x_{(i,j)} &\stackrel{10.47}{=} \sum_{i \in \prod_{i \in \{1, \dots, n-1\}} A_i} \left( \sum_{j \in A_n} x_{(i,j)} \right) \\ &= \sum_{i \in \prod_{i \in \{1, \dots, n-1\}} A_i} \left( \sum_{j \in A_n} f(i_1, \dots, i_{n-1}, j) \right) \end{aligned} \quad (10.12)$$

Further

$$\begin{aligned} \sum_{(i,j) \in (\prod_{i \in \{1, \dots, n-1\}} A_i) \times A_n} x_{(i,j)} &= \sum_{(i,j) \in (\prod_{i \in \{1, \dots, n-1\}} A_i) \times A_n} x_{((i_1, \dots, i_{n-1}), j)} \\ &\stackrel{5.103}{=} \sum_{(i_1, \dots, i_n) \in \prod_{i \in \{1, \dots, n\}} A_i} x_{(i_1, \dots, i_n)} \\ &= \sum_{(i_1, \dots, i_n) \in \prod_{i \in \{1, \dots, n\}} A_i} f(i_1, \dots, i_n) \end{aligned}$$

Using this with 10.12 proves the theorem.  $\square$

**Lemma 10.49.** If  $\langle A, + \rangle$  is a semi-group and  $\{x_i\}_{i \in \{0, \dots, n\}}$  is a set of elements in  $A$  such that  $\forall i \in \{1, \dots, 0\}$  we have that  $x_i = 0$  (the neutral element of the semi-group) then  $\sum_{i=0}^n x_i = 0$

**Proof.** This is proved by induction on  $n$ . So let  $X = \{n \in \mathbb{N} \mid \text{if } \{x_i\}_{i \in \{0, \dots, n\}} \text{ is a family of elements in } A \text{ that are all 0 then } \sum_{i=0}^n x_i\}$  then we have

1. If  $n = 0$  then  $\{x_i\}_{i \in \{0, \dots, 0\}}$  has  $x_0 = 0$  then  $\sum_{i=0}^0 x_i = x_0 = 0$  so  $0 \in X$
2. If  $n \in X$  then if  $\{x_i\}_{i \in \{0, \dots, n+1\}}$  of 0's we have  $\sum_{i=0}^{n+1} x_i = (\sum_{i=0}^n x_i) + x_{n+1}$   $\stackrel{x_{n+1}=0}{=} \sum_{i=0}^n x_i \stackrel{x_i \in X}{=} 0$

So using mathematical induction we have  $X = \mathbb{N}$  proving the theorem.  $\square$

**Theorem 10.50.** If  $\langle A, + \rangle$  is a abelian semi-group and  $\{x_i\}_{i \in I}$  is a family of elements in  $A$  with  $I$  is finite and non-empty (and thus have finite support) then if  $h: S_n \rightarrow I$  is a bijection (which must exists as  $I$  is finite) we have that  $\sum_{i \in I} x_i = \sum_{i=0}^{n-1} x_{h(i)}$

**Proof.** Let  $K = \text{support}(\{x_i\}_{i \in I})$  then if  $K = I$  we have if  $h: S_n \rightarrow I = \text{support}(\{x_i\}_{i \in I})$  is a bijection that by definition of  $\sum_{i \in I} x_i = \sum_{i=0}^{n-1} x_{h(i)}$ . So assume that  $K \subset I$  then if  $m = \#(K)$  we have by 5.45 that  $\#(I \setminus K) = n - m$ . So there exists bijections  $f_1: \{0, \dots, m-1\} \rightarrow K$  and  $f_2: \{0, \dots, (n-m)-1\} \rightarrow I \setminus K$ , define then the (easily proved) bijection  $f_3: \{m, \dots, n-1\} \rightarrow I \setminus K: i \rightarrow f_2(i-m)$ . This allows us to create a bijection (see 2.46)  $f_1 \cup f_3: \{0, \dots, n-1\} \rightarrow I$ . Then  $\sum_{i=0}^{n-1} x_{(f_1 \cup f_3)(i)} \stackrel{10.17}{=} \sum_{i=0}^{m-1} x_{(f_1 \cup f_3)(i)} + \sum_{i=m}^{n-1} x_{(f_1 \cup f_3)(i)} = \sum_{i=0}^{n-1} x_{(f_1)(i)} + \sum_{i=m}^{n-1} x_{(f_3)(i)} \stackrel{\text{K is support}(\{x_i\}_{i \in I})}{=} \sum_{i \in I} x(i) + \sum_{i=m}^{n-1} x_{(f_3)(i)} = \sum_{i \in I} x_i + \sum_{i=0}^{(n-m)-1} x_{(f_3)(i+m)} \stackrel{\text{previous lemma and } x_{(f_3)(i+m)} = x_{(f_2(i))} = 0}{=} \sum_{i \in I} x_i$  or summarized  $\sum_{i=0}^{n-1} x_{(f_1 \cup f_3)(i)} = \sum_{i \in I} x_i$ . Now if  $h: S_n = \{0, \dots, n-1\} \rightarrow I$  is a bijection then  $\sigma = (f_1 \cup f_3)^{-1} \circ h: S_n \rightarrow S_n$  is a bijection (or permutation) so using 10.26 we have that  $\sum_{i \in I} x_i = \sum_{i=0}^{n-1} x_{(f_1 \cup f_3)(i)} = \sum_{i=0}^{n-1} x_{(f_1 \cup f_3)((f_1 \cup f_3)^{-1}(h(i)))} = \sum_{i=0}^{n-1} x(h(i)) = \sum_{i=0}^{n-1} x_{h(i)}$  proving the theorem.  $\square$

A variant of the above theorem is expressed in the following lemma

**Lemma 10.51.** If  $\langle A, + \rangle$  is a abelian semi-group and  $\{x_i\}_{i \in I}$  a family of elements in  $A$  with  $\text{support}(\{x_i\}_{i \in I})$  is finite and  $\text{support}(\{x_i\}_{i \in I}) \subseteq J \subseteq I$  where  $J$  is a finite set then if  $h: S_n \rightarrow J$  is a bijection we have that  $\sum_{i=0}^{n-1} x_{h(i)} = \sum_{i \in I} x_i$

**Proof.** As  $\text{support}(\{x_i\}_{i \in I}) = \text{support}(\{x_i\}_{i \in J})$  we have if  $g: S_n \rightarrow \text{support}(\{x_i\}_{i \in I}) = \text{support}(\{x_i\}_{i \in J})$  is bijection that  $\sum_{i=0}^{n-1} x_i \stackrel{10.50}{=} \sum_{i \in J} x_i = \sum_{i \in I} x_i$   $\square$

**Theorem 10.52.** If  $\langle A, + \rangle$  is a abelian semi-group and  $\{x_i\}_{i \in I}$  and  $\{y_i\}_{i \in I}$  are family of sets in  $A$  with finite support then  $\{x_i + y_i\}_{i \in I}$  is a family with finite support and  $\sum_{i \in I} (x_i + y_i) = (\sum_{i \in I} x_i) + (\sum_{i \in I} y_i)$

**Proof.** If  $i \in \text{support}(\{x_i + y_i\}_{i \in I})$  then  $x_i + y_i \neq 0$ , now if  $x_i = 0 = y_i \Rightarrow x_i + y_i = 0$  a contradiction so we have either  $x_i \neq 0$  or  $y_i \neq 0$  and thus  $i \in \text{support}(\{x_i\}_{i \in I}) \cup \text{support}(\{y_i\}_{i \in I})$  and thus  $\text{support}(\{x_i + y_i\}_{i \in I}) \subseteq \text{support}(\{x_i\}_{i \in I}) \cup \text{support}(\{y_i\}_{i \in I})$ . As  $\text{support}(\{x_i\}_{i \in I})$ ,  $\text{support}(\{y_i\}_{i \in I})$  are finite we have  $\text{support}(\{x_i\}_{i \in I}) \cup \text{support}(\{y_i\}_{i \in I})$  is finite (see 5.37) and thus by 5.36 we have that  $\text{support}(\{x_i + y_i\}_{i \in I})$  is finite. If  $h: S_n \rightarrow \text{support}(\{x_i\}_{i \in I}) \cup \text{support}(\{y_i\}_{i \in I})$  is a bijection then as  $\text{support}(\{x_i + y_i\}_{i \in I}) \subseteq \text{support}(\{x_i\}_{i \in I}) \cup \text{support}(\{y_i\}_{i \in I}) \subseteq I$  and  $\text{support}(\{x_i\}_{i \in I})$ ,  $\text{support}(\{y_i\}_{i \in I}) \subseteq \text{support}(\{x_i\}_{i \in I}) \cup \text{support}(\{y_i\}_{i \in I})$  we have  $\sum_{i \in I} (x_i + y_i) \stackrel{\text{previous lemma}}{=} \sum_{i=0}^{n-1} (x_{h_i} + y_{h_i}) \stackrel{10.11}{=} (\sum_{i=0}^{n-1} x_{h_i}) + (\sum_{i=0}^{n-1} y_{h_i}) \stackrel{\text{previous lemma}}{=} \sum_{i \in I} x_i + \sum_{i \in I} y_i$   $\square$

**Corollary 10.53.** If  $\langle A, + \rangle$  is a abelian additive semi-group with unit 0 and  $\{x_i\}_{i \in I}$  is a finite family of elements in  $A$  then given  $a \in A$  we have for  $\{x_i + a\}_{i \in I}$  that  $\sum_{i \in I} (x_i + a) = (\sum_{i \in I} x_i) + (a \cdot \#(I))$

**Proof.** Define  $\{a_i\}_{i \in I}$  by  $\forall i \in I$  we have  $a_i = a$  then  $\{x_i + a\}_{i \in I} = \{x_i + a_i\}_{i \in I}$ , also let  $h: S_n \rightarrow I$  be a bijection where  $n = \#(I)$  then  $\sum_{i \in I} (x_i + a) = \sum_{i \in I} (x_i + a_i) = \sum_{i \in I} x_i + \sum_{i \in I} a_i = \sum_{i \in I} x_i + \sum_{i=0}^{n-1} a_{h_i} \stackrel{a_{h_i}=a \text{ and } 10.7}{=} \sum_{i \in I} x_i + a \cdot n = \sum_{i \in I} x_i + a \cdot \#(I)$   $\square$

**Corollary 10.54.** If  $\langle A, \cdot \rangle$  is a abelian multiplicative semi-group with unit 1 and  $\{x_i\}_{i \in I}$  is a finite family of elements in  $A$  then given  $a \in A$  we have for  $\{x_i \cdot a\}_{i \in I}$  that  $\prod_{i \in I} (x_i \cdot a) = (\prod_{i \in I} x_i) \times (a^{\#(I)})$

**Proof.** Define  $\{a_i\}_{i \in I}$  by  $\forall i \in I$  we have  $a_i = a$  then  $\{x_i \cdot a\}_{i \in I} = \{x_i \cdot a_i\}_{i \in I}$ , also let  $h: S_n \rightarrow I$  be a bijection where  $n = \#(I)$  then  $\prod_{i \in I} (x_i \cdot a) = \prod_{i \in I} (x_i \cdot a_i) = \prod_{i \in I} x_i \cdot \prod_{i \in I} a_i = \prod_{i \in I} x_i \cdot \prod_{i=0}^{n-1} a_{h_i} = a$  and 10.8  $\prod_{i \in I} x_i \cdot a^n = \sum_{i \in I} x_i \cdot a^n$   $\square$

If we have a ring then we have the following interesting properties for

**Theorem 10.55.** If  $\langle R, +, \cdot \rangle$  is a commutative ring with neutral element and unit element 1 then if  $\{x_i\}_{i \in I}$  is a family of elements in  $A$  with finite support then

1. If  $\exists i \in I$  with  $x_i = 0$  then  $\prod_{i \in I} x_i = 0$
2. If  $\langle R, +, \cdot \rangle$  is a integral domain (such that if  $x, y \in R \setminus \{0\}$  then  $x \cdot y \neq 0$ ) and  $\forall i \in I$  we have that  $x_i \neq 0$  then  $\text{support}(\{x_i^{-1}\}_{i \in I}) = \text{support}(\{x_i\}_{i \in I})$  is finite (so that  $\prod_{i \in I} x_i^{-1}$  is defined),  $\prod_{i \in I} x_i \neq 0$  and  $(\prod_{i \in I} x_i)^{-1} = \prod_{i \in I} x^{-1}$

**Proof.**

1.  $\exists i \in I$  such that  $x_i = 0$  then as  $I$  is the disjoint union of  $\{i\}$  and  $I \setminus \{i\}$  we have that  $\prod_{i \in I} x_i = (\prod_{i \in I \setminus \{i\}} x_i) \cdot (\prod_{i \in \{i\}} x_i) = (\prod_{i \in I} x_i) \cdot x_i = (\prod_{i \in I} x_i) \cdot 0 = 0$
2. First if  $i \in \text{support}(\{x_i\}_{i \in I})$  then  $x_i \neq 1$ , if now  $x_i^{-1} = 1$  then  $1 = x_1 \cdot x_i^{-1} = x_i \cdot 1 = x_i \neq 1$  yielding contradiction so that  $x_i^{-1} \neq 1 \Rightarrow i \in \text{support}(\{x_i^{-1}\}_{i \in I})$ . If  $i \in \text{support}(\{x_i^{-1}\}_{i \in I})$  then  $x_i^{-1} \neq 1$ , if now  $x_i = 1$  then  $1 = x_i^{-1} \cdot x_i = x_i^{-1} \cdot 1 = x_i^{-1} \neq 1$ , a contradiction so that  $x_i \neq 1$  and thus  $i \in \text{support}(\{x_i\}_{i \in I})$ . We prove now by induction that if  $\{x_i\}_{i \in \{0, \dots, n\}}$  is such that  $\forall i \in \{1, \dots, n\}$  we have  $x_i \neq 0$  then  $\prod_{i=0}^n x_i^{-1} = (\prod_{i=0}^n x_i)^{-1}$  so let  $S = \{n \in \mathbb{N}_0 \mid \{x_i\}_{i \in \{0, \dots, n\}} \text{ is such that } \forall i \in \{0, \dots, n\} \text{ we have } x_i \neq 0 \text{ then } \prod_{i=0}^n x_i \neq 0 \text{ and } \prod_{i=0}^n x_i^{-1} = (\prod_{i=0}^n x_i)^{-1}\}$  then we have
  - a. If  $n = 0$  then  $0 \neq x_0 = \prod_{i=0}^0 x_i^{-1} = x_0^{-1} = (\prod_{i=0}^0 x_i)^{-1}$  so that  $0 \in S$
  - b. If  $n \in S$  then  $\prod_{i=0}^{n+1} x_i = (\prod_{i=0}^n x_i) \cdot x_{n+1} \neq 0$  (as  $n \in S$  and  $x_{n+1} \neq 0$ ), also  $\prod_{i=0}^{n+1} x_i^{-1} = (\prod_{i=0}^n x_i^{-1}) \cdot x_{n+1}^{-1} \stackrel{n \in S}{=} (\prod_{i=0}^n x_i)^{-1} \cdot x_{n+1}^{-1} = ((\prod_{i=0}^n x_i) \cdot x_{n+1})^{-1} = (\prod_{i=0}^{n+1} x_i)^{-1}$  so that  $n+1 \in S$ .

Now if  $\text{support}(\{x_i^{-1}\}_{i \in I}) = \text{support}(\{x_i\}_{i \in I}) = \emptyset$  then  $\prod_{i \in I} x_i^{-1} = 1 = 1^{-1} = (\prod_{i \in I} x_i)^{-1}$ , if  $\text{support}(\{x_i\}_{i \in I}) = \text{support}(\{x_i^{-1}\}_{i \in I}) \neq \emptyset$  then there exists a bijection  $b: \{0, \dots, n\} \rightarrow \text{support}(\{x_i\}_{i \in I}) = \text{support}(\{x_i^{-1}\}_{i \in I})$  such that  $\prod_{i \in I} x_i^{-1} = \prod_{i=0}^n x_{b(i)}^{-1} = (\prod_{i=0}^n x_{b(i)})^{-1} = (\prod_{i \in I} x_i)^{-1}$   $\square$

**Lemma 10.56.** If  $\{\langle A_i, +_i \rangle\}_{i \in I}$  is a family of abelian semi-groups then by 3.15 we have that  $\langle \prod_{i \in I} A_i, + \rangle$  is a abelian semi-group. If now  $\{x_j\}_{j \in J}$  is a family of elements in  $\prod_{i \in I} A_i$  with finite support then  $\forall i \in I$  we have  $(\sum_{j \in J} x_j)(i) \stackrel{\text{notation}}{=} (\sum_{j \in J} x_j)_i = \sum_{j \in J} (x_j)_i \stackrel{\text{notation}}{=} \sum_{j \in J} (x_j)(i)$

**Proof.** As  $\{x_j\}_{j \in J}$  has finite support then we have the following cases to consider:

1. ( $\text{support}(\{x_j\}_{j \in J}) = \emptyset$ ) Then if  $i \in I$  and  $j \in J \Rightarrow x_j = 0 \Rightarrow (x_j)(i) = 0 \Rightarrow \text{support}(\{(x_j)(i)\}) = \emptyset$  and we have  $(\sum_{j \in J} x_j)(i) = 0(i) = 0 = \sum_{j \in J} (x_j)(i)$
2. ( $\text{support}(\{x_j\}_{j \in J}) \neq \emptyset$ ) then  $\forall i \in I$  if  $j \in \text{support}(\{x_j(i)\}) \Rightarrow x_j(i) \neq 0 \Rightarrow x_j \neq 0 \Rightarrow j \in \text{support}(\{x_j\}_{j \in J})$  giving  $\text{support}(\{x_j(i)\}_{j \in J}) \subseteq \text{support}(\{x_j\}_{j \in J})$  proving that  $\text{support}(\{x_j(i)\}_{j \in J})$  is finite so the sums are defined and we have a bijection  $b: \{0, \dots, n-1\} \rightarrow \text{support}(\{x_j\}_{j \in J})$  then by 10.51 we have  $(\sum_{j \in J} x_j)(i) = (\sum_{j=0}^{n-1} x_{b_j})(i) \stackrel{10.10}{=} \sum_{j=0}^{n-1} (x_{b_j})(i) \stackrel{10.51}{=} \sum_{j \in J} x_j(i)$   $\square$

**Theorem 10.57.** Let  $\mathbb{R}$  (or  $\mathbb{R}_{\mathbb{C}}$ ) be the real field (or the real field embedded in  $\mathbb{C}$ ) and  $\{x_i\}_{i \in I}$  be such that  $\forall i \in \sup(\{x_i\}_{i \in I})$  we have  $x_i = c$  then we have that  $\sum_{i \in I} x_i = (\#(\sup(\{x_i\}_{i \in I}))) \cdot c$

**Proof.** If  $b: \{0, \dots, n-1\} \rightarrow \sup(\{x_i\}_{i \in I})$  is a bijection where  $n = \#(\sup(\{x_i\}_{i \in I}))$  then we have that  $\sum_{i \in I} x_i = \sum_{i=0}^{n-1} x_{b(i)} = \sum_{i=0}^{n-1} c \stackrel{10.30}{=} ((n-1)+1) \cdot c = n \cdot c = \#(\sup(\{x_i\}_{i \in I}))$   $\square$

**Theorem 10.58.** Let  $\langle F, +, \cdot \rangle$  be a field  $\alpha \in F$  and  $\{x_i\}_{i \in I}$  a family with finite support then  $\sum_{i \in I} \alpha \cdot x_i = \alpha \cdot \sum_{i \in I} x_i$

**Proof.** Let  $b: \{0, \dots, n-1\} \rightarrow \text{support}(\{x_i\}_{i \in I})$  then we have  $\sum_{i \in I} \alpha \cdot x_i = \sum_{i=0}^n \alpha \cdot x_{b(i)} = \alpha \cdot \sum_{i=0}^n x_{b(i)} = \alpha \cdot \sum_{i \in I} x_i$   $\square$

**Corollary 10.59.** Let  $\langle A, +, \cdot \rangle$  be a field,  $I, J$  sets and  $\{x_i\}_{i \in I}$ ,  $\{y_j\}_{j \in J}$  families of elements in  $I$ ,  $J$  with finite support then  $\sum_{(i,j) \in I \times J} (x_i \cdot y_j) = (\sum_{i \in I} x_i) \cdot (\sum_{j \in J} y_j)$

**Proof.** As for a field  $\langle A, + \rangle$  is a Abelian group we have by 10.47 that  $\sum_{(i,j) \in I \times J} (x_i \cdot y_j) = \sum_{i \in I} (\sum_{j \in J} (x_i \cdot y_j)) \stackrel{10.58}{=} \sum_{i \in I} x_i \cdot (\sum_{j \in J} y_j) \stackrel{10.58}{=} (\sum_{i \in I} x_i) \cdot (\sum_{j \in J} y_j)$   $\square$

**Theorem 10.60.** Let  $\langle R, +, \cdot \rangle$  be a ring,  $n \in \mathbb{N}_0$  and  $\{\{a_{i,j}\}_{j \in N_i}\}_{i \in \{1, \dots, n\}}$  such that  $\forall i \in \{1, \dots, n\}$   $N_i$  is finite and non empty then  $\prod_{i \in \{1, \dots, n\}} (\sum_{j \in N_i} a_{i,j}) = \sum_{\gamma \in \prod_{i \in \{1, \dots, n\}} N_i} (\prod_{i \in \{1, \dots, n\}} a_{i, \gamma_i})$

**Proof.** We prove this by induction on  $n$  so let  $\mathcal{S} = \{n \in \mathbb{N}_0 \mid \forall \{\{a_{i,j}\}_{j \in N_i}\}_{i \in \{1, \dots, n\}}, \text{ such that } N_i \text{ is finite non empty, we have } \prod_{i \in \{1, \dots, n\}} (\sum_{j \in N_i} a_{i,j}) = \sum_{\gamma \in \prod_{i \in \{1, \dots, n\}} N_i} (\prod_{i \in \{1, \dots, n\}} a_{i, \gamma_i})\}$  then

**1  $\in \mathcal{S}$ .** If  $\{\{a_{i,j}\}_{j \in N_i}\}_{i \in \{1, \dots, 1\}}$   $N_1$  finite non empty then

$$\begin{aligned} \prod_{i \in \{1, \dots, 1\}} \left( \sum_{j \in N_i} a_{i,j} \right) &= \sum_{j \in N_1} a_{1,j} \\ &\stackrel{5.103}{=} \sum_{\gamma \in \prod_{i \in \{1, \dots, 1\}} N_i} a_{1, \gamma_1} \\ &= \prod_{i \in \{1, \dots, 1\}} \left( \sum_{\gamma \in \prod_{i \in \{1, \dots, 1\}} N_i} a_{i, \gamma_i} \right) \end{aligned}$$

proving that  $1 \in \mathcal{S}$

**$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$ .** Let  $\{\{a_{i,j}\}_{j \in N_i}\}_{i \in \{1, \dots, n+1\}}$  be a finite family such that  $N_i$  is finite and non empty then

$$\begin{aligned} \prod_{i \in \{1, \dots, n+1\}} \left( \sum_{j \in A_i} a_{i,j} \right) &= \left( \prod_{i \in \{1, \dots, n\}} \left( \sum_{j \in A_i} a_{i,j} \right) \right) \cdot \left( \sum_{j \in A_{n+1}} a_{n+1,j} \right) \\ &\stackrel{n \in \mathcal{S}}{=} \left( \sum_{\gamma \in \prod_{i \in \{1, \dots, n\}} N_i} \left( \prod_{i \in \{1, \dots, n\}} a_{i, \gamma_i} \right) \right) \cdot \left( \sum_{j \in A_{n+1}} a_{n+1,j} \right) \\ &\stackrel{10.58}{=} \sum_{\gamma \in \prod_{i \in \{1, \dots, n\}} N_i} \left( \left( \prod_{i \in \{1, \dots, n\}} a_{i, \gamma_i} \right) \cdot \sum_{j \in A_{n+1}} a_{n+1,j} \right) \\ &\stackrel{10.58}{=} \sum_{\gamma \in \prod_{i \in \{1, \dots, n\}} N_i} \left( \sum_{j \in A_{n+1}} \left( \left( \prod_{i \in \{1, \dots, n\}} a_{i, \gamma_i} \right) \cdot a_{n+1,j} \right) \right) \\ &\stackrel{10.48}{=} \sum_{\gamma \in \prod_{i \in \{1, \dots, n+1\}} A_i} \left( \left( \prod_{i \in \{1, \dots, n\}} a_{i, \gamma_i} \right) \cdot a_{n+1, \gamma_{n+1}} \right) \\ &= \sum_{\gamma \in \prod_{i \in \{1, \dots, n+1\}} A_i} \left( \prod_{i \in \{1, \dots, n+1\}} a_{i, \gamma_i} \right) \end{aligned}$$

proving that  $n+1 \in \mathcal{S}$  □

**Theorem 10.61.** Let  $\mathbb{R}$  (or  $\mathbb{R}_{\mathbb{C}}$ ) be the real field (embedded real field in the complex numbers,  $\{x_i\}_{i \in I}$  a family in  $\mathbb{R}$  (or  $\mathbb{R}_{\mathbb{C}}$ ) with finite support such that  $\forall i \in I$  we have  $0 \leq x_i$  then  $0 \leq \langle \sum_{i \in I} x_i \rangle$  (where we have  $<$  if  $\text{support}(\{x_i\}_{i \in I}) \neq 0$ ) and  $\sum_{i \in I} x_i = 0$  if  $\text{support}(\{x_i\}_{i \in I}) = 0$

**Proof.** We have the following cases to consider:

1. ( $\text{support}(\{x_i\}_{i \in I}) = \emptyset$ ) then  $\sum_{i \in I} x_i = 0$  by definition
2. ( $\text{support}(\{x_i\}_{i \in I}) \neq \emptyset$ ) then let  $b: \{0, \dots, n-1\} \rightarrow \text{support}(\{x_i\}_{i \in I})$  then we have that  $n > 0 \Rightarrow x_{b(0)} > 0$  so that we have  $\sum_{i \in I} x_i = \sum_{i=0}^{n-1} x_{b(i)} > 0$  using 10.31 □

**Theorem 10.62.** Let  $\mathbb{R}$  (or  $\mathbb{R}_{\mathbb{C}}$ ) be the real field (embedded real field in the complex numbers,  $\{x_i\}_{i \in I}$ ,  $\{y_i\}_{i \in I}$  families in  $\mathbb{R}$  (or  $\mathbb{R}_{\mathbb{C}}$ ) such that  $\forall i \in I$  we have  $x_i \leq y_i$  then we have  $\sum_{i \in I} x_i \leq \langle \sum_{i \in I} y_i \rangle$  (where we have  $<$  if there is a  $x_i < y_i$ )

**Proof.** Consider the family  $\{y_i + (-x_i)\}_{i \in I}$  then we have  $\forall i \in I$  that  $y_i + (-x_i) \geq 0$  and we have thus by 10.61 that

$$\sum_{i \in I} (y_i + (-x_i)) \geq 0 (> 0) \text{ (where we have } > \text{ if there exists } a y_i > x_i) \quad (10.13)$$

Now  $\sum_{i \in I} (y_i + (-x_i)) \stackrel{10.52}{=} \sum_{i \in I} y_i + \sum_{i \in I} (-x_i) \stackrel{10.58}{=} \sum_{i \in I} y_i + (-\sum_{i \in I} x_i) = \sum_{i \in I} y_i - \sum_{i \in I} x_i$  which together with 10.13 gives that  $\sum_{i \in I} y_i \geq (\geq) \sum_{i \in I} x_i$  where we have  $>$  if there exists a  $y_i > x_i$  □

**Theorem 10.63.** Let  $\delta \in \mathbb{R}_+$ ,  $\{x_i\}_{i \in I}$  is a family in  $\mathbb{R}_+$  with finite support( $\{x_i\}_{i \in I}\}) =  $\{i \in I \mid x_i \neq 1\} \neq \emptyset$  such that  $\forall i \in I$  we have  $0 < x_i < \delta$  then  $0 < \prod_{i \in I} x_i < \delta^{\#(\text{support}(\{x_i\}_{i \in I}))} < \delta < 1$$

**Proof.** First we prove by induction that  $\prod_{i=0}^n x_i < \delta^{n+1}$  so let  $S = \{n \in \mathbb{N} \mid \text{if } \{x_i\}_{i \in \{0, \dots, n\}} \text{ is such that } \forall i \in \{0, \dots, n\} 0 < x_i < \delta \text{ then } 0 < \prod_{i=0}^n x_i < \delta^{n+1}\}$  then we have

1. If  $n = 0$  then  $0 < x_0 = \prod_{i=0}^0 x_i = x_0 < \delta = \delta^{0+1}$  so that  $0 \in S$
2. Let  $n \in S$  and  $\{x_i\}_{i \in \{0, \dots, n+1\}}$  such that  $\forall i \in \{0, \dots, n+1\}$  we have  $0 < x_i < \delta$  then we have as  $0 < x_{n+1} < \delta < 1$  and  $n \in S$  we have  $0 < (\prod_{i=0}^n x_i) \cdot x_{n+1} = \prod_{i=0}^{n+1} x_i = (\prod_{i=0}^n x_i) \cdot x_{n+1} < \delta^n \cdot x_{n+1} < \delta^n \cdot \delta < \delta^{n+1}$  so that  $n+1 \in S$

Second if  $\emptyset \neq \text{support}(\{x_i\}_{i \in I})$  is finite then there exists a bijection  $b: \{0, \dots, \#(\text{support}(\{x_i\}_{i \in I})) - 1\} \rightarrow \text{support}(\{x_i\}_{i \in I})$  and  $0 < \prod_{i=0}^{\#(\text{support}(\{x_i\}_{i \in I})) - 1} x_{b(i)} = \prod_{i \in I} x_i = \prod_{i=0}^{\#(\text{support}(\{x_i\}_{i \in I})) - 1} x_{b(i)} < \delta^{\#(\text{support}(\{x_i\}_{i \in I}))}$ . The rest follows from 9.41 (13). □

## 10.2 Group of permutations

**Definition 10.64.** Given  $n \in \mathbb{N}_0$  we define  $P_n = \{\sigma \mid \sigma: \{1, \dots, n\} \rightarrow \{1, \dots, n\} \text{ is a permutation}\}$  /  $P_n$  is the collection of all permutations on  $\{1, \dots, n\}$  / or equivalently the collection of bijections from  $\{1, \dots, n\}$  to  $\{1, \dots, n\}$ . So  $P_n = S_{\{1, \dots, n\}}$  / see 10.19

**Example 10.65.** Let  $n \in \mathbb{N}_0$  then  $\iota_n: \{1, \dots, n\} \rightarrow \{1, \dots, n\}$  defined by  $j \rightarrow \iota_n(j) = n - j + 1$  is a permutation (further it is trivial that if  $m \in \{1, \dots, n\}$  then  $(\iota_n)_{\mid \{1, \dots, m\}} = \iota_m$ ). Here  $\iota_n(n) = n - n + 1 = 1$ ,  $\iota_n(n-1) = 2, \dots, \iota_n(1) = n - 1 + 1 = n$ .

**Proof.**

1. (**injectivity**) If  $\iota_n(j) = \iota_n(k) \Rightarrow n - j = b - k \Rightarrow -j = -k \Rightarrow j = k$
2. (**surjectivity**) If  $k \in \{1, \dots, n\}$  take then  $j = n - k + 1 \in \{1, \dots, n\}$  so that  $\iota_n(n - k + 1) = n - (n - k + 1) + 1 = k$  □

**Theorem 10.66.**  $\langle P_n, \circ \rangle$  forms a group ( $\circ$  is composition of functions) called the permutation group

**Proof.**

1. Given  $\sigma, \tau \in P_n$  we have that  $\sigma \circ \tau$  is again a element of  $\sigma$  (composition of bijections is again a bijection), so  $\circ$  defines a function  $P_n \times P_n \rightarrow P_n$
2. **(Associativity)**  $\forall \sigma, \tau, \kappa \in P_n$  and  $\forall i \in \{1, \dots, n\}$  we have

$$\begin{aligned} ((\sigma \circ \tau) \circ \kappa)(i) &= (\sigma \circ \tau)(\kappa(i)) \\ &= \sigma(\tau(\kappa(i))) \\ &= \sigma((\tau \circ \kappa)(i)) \\ &= (\sigma \circ (\tau \circ \kappa))(i) \end{aligned}$$

and thus we have  $((\sigma \circ \tau) \circ \kappa) = \sigma \circ (\tau \circ \kappa)$

3. **(neutral element)**  $\forall \sigma \in P_n$  we have  $\sigma \circ 1_{\{1, \dots, n\}} = \sigma = 1_{\{1, \dots, n\}} \circ \sigma$  so  $1_{\{1, \dots, n\}}$  the identity mapping is the neutral element in  $P_n$
4. **(inverse)**  $\forall \sigma \in P_n$  we have as  $\sigma$  is a bijection the existence of a inverse  $\sigma^{-1}$  so that  $\sigma \circ \sigma^{-1} = i_{\{1, \dots, n\}} = \sigma^{-1} \circ \sigma$   $\square$

Note that  $\langle P_n, \circ \rangle$  is not commutative.

**Definition 10.67.** As  $\langle P_n, \circ \rangle$  forms a group, giving a finite family of  $\{\sigma_i\}_{i \in \{1, \dots, n\}}$  of permutations in  $P_n$  we have  $\sum_{i=1}^n \sigma_i$  defined ( $\circ$  being the sum operator) we use however the notation  $(\sigma_1 \circ \dots \circ \sigma_n)$  instead of  $\sum_{i=1}^n \sigma_i$

**Note 10.68.** As  $(\sigma_1 \circ \dots \circ \sigma_n) = \sum_{i=1}^n \sigma_i = \sum_{i=0}^{n-1} \sigma_{i+1}$  we have

$$\begin{aligned} (\sigma_1 \circ \dots \circ \sigma_1) &= \sum_{i=0}^0 \sigma_{i+1} = \sigma_1 \\ \text{if } n > 1 \text{ then } n &= (n-1)+1 \text{ and } (\sigma_1 \circ \dots \circ \sigma_n) = \sum_{i=0}^{(n-1)+1} \sigma_{i+1} = \left( \sum_{i=0}^{n-1} \sigma_{i+1} \right) \circ \sigma_{(n-1)+1} \\ &= (\sigma_1 \circ \dots \circ \sigma_{n-1}) \circ \sigma_n \\ \text{if } n > 1 \text{ then } (n-1)+1 &= \sum_{i=0}^{(n-1)+1} \sigma_{i+1} \\ &\stackrel{10.17, m=0}{=} \left( \sum_{i=0}^0 \sigma_{i+1} \right) \circ \left( \sum_{i=1}^{(n-1)+1} \sigma_{i+1} \right) \\ &= \sigma_1 \circ (\sigma_2 \circ \dots \circ \sigma_n) \end{aligned}$$

Summarizing we have that

$$(\sigma_1 \circ \dots \circ \sigma_1) = \sigma_1$$

and if  $n > 1$  then

$$\begin{aligned} (\sigma_1 \circ \dots \circ \sigma_n) &= (\sigma_1 \circ \dots \circ \sigma_{n-1}) \circ \sigma_n \\ &= \sigma_1 \circ (\sigma_2 \circ \dots \circ \sigma_n) \end{aligned}$$

We define now a special permutation called a transposition.

**Definition 10.69.** If  $n \in \mathbb{N}_0$  and  $i, j \in \{1, \dots, n\}$  then  $(i \leftrightarrow_n j) \in P_n$  is defined by  $(i \leftrightarrow_n j) = (i \leftrightarrow_{1, n} j)$  [see 10.23] hence  $(i \leftrightarrow_n j)(k) = \begin{cases} k & \text{if } k \in \{1, \dots, n\} \setminus \{i, j\} \\ j & \text{if } k = i \\ i & \text{if } k = j \end{cases}$ . We call  $(i \leftrightarrow_n j)$  a trasposition.

A transposition  $(i \leftrightarrow_n j)$  is **strict** if  $i \neq j$

**Note 10.70.** It is trivially to prove the following statements

1. If  $i = j$  then  $(i \leftrightarrow_n j) = \mathbb{1}_{\{1, \dots, n\}}$
2.  $(i \leftrightarrow_n j) = (j \leftrightarrow_n i)$
3.  $(i \leftrightarrow_n j) \circ (i \leftrightarrow_n j) = \mathbb{1}_{\{1, \dots, n\}}$

**Lemma 10.71.** Let  $n \in \mathbb{N}_0 \setminus \{1\}$  with  $X, Y$  sets and  $f: X^n \rightarrow Y$  a function so that for all  $i, j \in \{1, \dots, n\}$  we have that  $f(x_{(i \leftrightarrow_n j)(1)}, \dots, x_{(i \leftrightarrow_n j)(n)}) = f(x_1, \dots, x_n)$  then  $\forall \sigma \in P_n$  we have  $f(x_{\sigma(1)}, \dots, x_{\sigma(n)}) = f(x_1, \dots, x_n)$

**Proof.** We prove this by induction so let  $\mathcal{S} = \{n \in \{2, \dots, \infty\} \mid \text{Let } (x_1, \dots, x_n) \in X^n \text{ and } f: X^n \rightarrow Y \text{ be such that } \forall i, j \in \{1, \dots, n\} f(x_{(i \leftrightarrow_n j)(1)}, \dots, x_{(i \leftrightarrow_n j)(n)}) = f(x_1, \dots, x_n) \text{ then } \forall \sigma \in P_n f(x_{\sigma(1)}, \dots, x_{\sigma(n)}) = f(x_1, \dots, x_n)\}$  then we have:

**2  $\in \mathcal{S}$ .** Let  $(x_1, x_2) \in X^2$  and  $f: X^2 \rightarrow Y$  be such that  $f(x_1, x_2) = f(x_2, x_1)$ . Given  $\sigma \in P_n$  we have either:

- $\sigma(1) = 1$ .** then as  $\sigma$  is injective we must have  $\sigma(2) = 2$  hence  $f(x_{\sigma(1)}, x_{\sigma(2)}) = f(x_1, x_2)$
- $\sigma(1) = 2$ .** then as  $\sigma$  is injective we must have  $\sigma(2) = 1$  hence  $f(x_{\sigma(1)}, x_{\sigma(2)}) = f(x_2, x_1) = f(x_1, x_2)$

hence in all cases we have  $f(x_{\sigma(1)}, x_{\sigma(2)}) = f(x_1, x_2)$  proving that  $2 \in \mathcal{S}$

**$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$ .** Let  $(x_1, \dots, x_{n+1}) \in X^{n+1}$  and  $f: X^{n+1} \rightarrow Y$  be such that  $\forall i, j \in \{1, \dots, n+1\}$  we have  $f(x_{(i \leftrightarrow_{n+1} j)(1)}, \dots, x_{(i \leftrightarrow_{n+1} j)(n+1)}) = f(x_1, \dots, x_{n+1})$ . Define given  $x \in X$   $g_x: X^n \rightarrow Y$  by  $g(y_1, \dots, y_n) = f(y_1, \dots, y_n, x)$  then given  $(i, j) \in \{1, \dots, n\}$  we have

$$\begin{aligned} g_{x_{n+1}}(x_{(i \leftrightarrow_n j)(1)}, \dots, x_{(i \leftrightarrow_n j)(n)}) &= f(x_{(i \leftrightarrow_n j)(1)}, \dots, x_{(i \leftrightarrow_n j)(n)}, x_{n+1}) \\ &\stackrel{10.24}{=} f(x_{(i \leftrightarrow_{n+1} j)_{\{1, \dots, n\}}(1)}, \dots, x_{(i \leftrightarrow_{n+1} j)_{\{1, \dots, n\}}(n)}, \\ &\quad x_{n+1}) \\ &\stackrel{i, j \in \{1, \dots, n\}}{=} f(x_{(i \leftrightarrow_{n+1} j)_{\{1, \dots, n\}}(1)}, \dots, x_{(i \leftrightarrow_{n+1} j)_{\{1, \dots, n\}}(n+1)}) \\ &= f(x_1, \dots, x_n, x_{n+1}) \\ &= g_{x_{n+1}}(x_1, \dots, x_n) \end{aligned}$$

using then the fact that  $n \in \mathcal{S}$  we have that

$$\forall \sigma \in P_n \text{ we have } g(x_{\sigma(1)}, \dots, x_{\sigma(n)}) = g(x_1, \dots, x_n) \quad (10.14)$$

Let  $\sigma \in P_{n+1}$  take then  $l = \sigma^{-1}(n+1)$  so that  $n+1 = \sigma(l)$ . Define then  $\rho \in P_{n+1}$  by  $\rho = \sigma \circ (l \leftrightarrow_{n+1} n+1)$  then we have that  $\rho(n+1) = \sigma((l \leftrightarrow_{n+1} n+1)(n+1)) = \rho(l) = n+1$ . So if we define  $\rho' = \rho|_{\{n, \dots, n\}}$  then by 10.21 we have that  $\rho' \in P_n$ .

$$\begin{aligned} f(x_{\sigma(1)}, \dots, x_{\sigma(n+1)}) &= f(x_{\sigma((l \leftrightarrow_{n+1} n+1)(1))}, \dots, x_{\sigma((l \leftrightarrow_{n+1} n+1)(n))}, x_{\sigma((l \leftrightarrow_{n+1} n+1)(n+1))}) \\ &= f(x_{\rho(1)}, \dots, x_{\rho(n+1)}) \\ &= f(x_{\rho(1)}, \dots, x_{\rho(n)}, x_{n+1}) \\ &= f(x_{\rho|_{\{1, \dots, n\}}(1)}, \dots, x_{\rho|_{\{1, \dots, n\}}(n)}, x_{n+1}) \\ &= f(x_{\rho'(1)}, \dots, x_{\rho'(n)}, x_{n+1}) \\ &= g_{x_{n+1}}(x_{\rho'(1)}, \dots, x_{\rho'(n+1)}) \\ &\stackrel{10.14}{=} g_{x_{n+1}}(x_1, \dots, x_n) \\ &= f(x_1, \dots, x_{n+1}) \end{aligned}$$

proving that  $n+1 \in P_{n+1}$ .

The theorem is then proved by mathematical induction.  $\square$

**Lemma 10.72.** Let  $n \in \mathbb{N}_0 \setminus \{1\}$ ,  $X, Y$  sets and  $f: X^n \rightarrow Y$  such that  $\forall i \in \{1, \dots, n-1\}$  we have that  $f(x_{(i \leftrightarrow_n i+1)(1)}, \dots, x_{(i \leftrightarrow_n i+1)(n)}) = f(x_1, \dots, x_n)$  then  $\forall i, j \in \{1, \dots, n\}$  we have that  $f(x_{(i \leftrightarrow_n j)(1)}, \dots, x_{(i \leftrightarrow_n j)(n)})$

**Proof.** We prove this by induction on  $n$ , so let  $\mathcal{S} = \{n \in \mathbb{N}_0 \setminus \{1\} \mid \text{Let } (x_1, \dots, x_n) \in X^n \text{ and } f: X^n \rightarrow Y \text{ be such that } \forall i \in \{1, \dots, n-1\} \text{ } f(x_{(i \leftrightarrow_n i+1)(1)}, \dots, x_{(i \leftrightarrow_n i+1)(n)}) = f(x_1, \dots, x_n) \text{ then } \forall i, j \in \{1, \dots, n\} \text{ we have } f(x_{(i \leftrightarrow_n j)(1)}, \dots, x_{(i \leftrightarrow_n j)(n)}) = f(x_1, \dots, x_n)\}$  then we have

**2  $\in \mathcal{S}$ .** Let  $(x_1, x_2) \in X^2$  and  $f: X^2 \rightarrow Y$  be such that  $\forall i \in \{1, \dots, 2-1\} = \{1\}$  we have that  $f(x_{(i \leftrightarrow_2 i+1)(1)}, f(x_{(i \leftrightarrow_2 i+1)(2)})) = f(x_1, x_2) \Rightarrow f(x_2, x_1) = f(x_1, x_2)$ . Let  $i, j \in \{1, 2\}$  then as  $(i \leftrightarrow_2 j) = (j \leftrightarrow_2 i)$  we may always assume that  $i \leq j$ . For  $i, j$  we have then either

**$i = 1 \wedge j = 1$ .** then  $(i \leftrightarrow_2 j) = \mathbb{1}_{\{1, 2\}}$  hence  $f(x_{(i \leftrightarrow_2 j)(1)}, x_{(i \leftrightarrow_2 j)(2)}) = f(x_1, x_2)$

**$i = 1 \wedge j = 2$ .** then  $(i \leftrightarrow_2 j) = (1 \leftrightarrow_2 2)$  hence  $f(x_{(i \leftrightarrow_2 j)(1)}, x_{(i \leftrightarrow_2 j)(2)}) = f(x_2, x_1) = f(x_1, x_2)$

**$i = 2 \wedge j = 2$ .** then  $(i \leftrightarrow_2 j) = \mathbb{1}_{\{1, 2\}}$  hence  $f(x_{(i \leftrightarrow_2 j)(1)}, x_{(i \leftrightarrow_2 j)(2)}) = f(x_1, x_2)$

**$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$ .** Let  $(x_1, \dots, x_{n+1})$  and  $f: X^{n+1} \rightarrow Y$  be such that  $\forall i \in \{1, \dots, (n+1)-1\} = \{1, \dots, n\}$   $f(x_{(i \leftrightarrow_{n+1} i+1)(1)}, \dots, x_{(i \leftrightarrow_{n+1} i+1)(n+1)}) = f(x_1, \dots, x_{n+1})$ . Define given  $x \in X$   $g_x: X^n \rightarrow Y$  by  $g_x(y_1, \dots, y_n) = f(y_1, \dots, y_n, x)$  then we have  $\forall i \in \{1, \dots, n-1\}$  that  $(i \leftrightarrow_n j) \stackrel{10.24}{=} (i \leftrightarrow_{n+1} j)|_{\{1, \dots, n\}}$  giving

$$\begin{aligned}
 g_{x_{n+1}}(x_{(i \leftrightarrow_{n+1} i+1)(1)}, \dots, x_{(i \leftrightarrow_{n+1} i+1)(n)}) &= f(x_{(i \leftrightarrow_{n+1} i+1)(1)}, \dots, x_{(i \leftrightarrow_{n+1} i+1)(n)}, x_{n+1}) \\
 &\stackrel{10.24}{=} f(x_{(i \leftrightarrow_{n+1} i+1)|_{\{1, \dots, n\}}(1)}, \dots, \\
 &\quad x_{(i \leftrightarrow_{n+1} i+1)|_{\{1, \dots, n\}}(n)}, x_{n+1}) \\
 &\stackrel{i, i+1 \leq n < n+1}{=} f(x_{(i \leftrightarrow_{n+1} i+1)|_{\{1, \dots, n\}}(1)}, \dots, \\
 &\quad x_{(i \leftrightarrow_{n+1} i+1)|_{\{1, \dots, n\}}(n+1)}) \\
 &= f(x_1, \dots, x_n, x_{n+1}) \\
 &= g_{x_{n+1}}(x_1, \dots, x_n)
 \end{aligned} \tag{10.15}$$

Let  $i, j \in \{1, \dots, n+1\}$  then as  $(i \leftrightarrow_{n+1} j) = (j \leftrightarrow_{n+1} i)$  we may always assume that  $i \leq j$ . For  $i, j$  we have then the following cases:

**$i = j$ .** then  $(i \leftrightarrow_{n+1} j) = \mathbb{1}_{\{1, \dots, n+1\}}$  so that  $f(x_{(i \leftrightarrow_{n+1} j)(1)}, \dots, x_{(i \leftrightarrow_{n+1} j)(n+1)}) = f(x_1, \dots, x_{n+1})$

**$i = n \wedge j = n+1$ .** then  $(i \leftrightarrow_{n+1} j) = (n \leftrightarrow_{n+1} n+1)$  so that  $f(x_{(i \leftrightarrow_{n+1} j)(1)}, \dots, x_{(i \leftrightarrow_{n+1} j)(n+1)}) = f(x_1, \dots, x_{n+1})$

**$i \neq n \wedge j = n+1$ .** then  $i < n < n+1$  so that by 10.24 we have

$$(i \leftrightarrow_n n) = (i \leftrightarrow_{n+1} n)|_{\{1, \dots, n\}} \tag{10.16}$$

Define  $\rho = (n \leftrightarrow_{n+1} n+1) \circ (i \leftrightarrow_{n+1} n) \circ (n \leftrightarrow_{n+1} n+1)$  then we have for  $k \in \{1, \dots, n+1\}$  that

**$k = n+1$ .** then

$$\begin{aligned}
 \rho(k) &= (n \leftrightarrow_{n+1} n+1)((i \leftrightarrow_{n+1} n)((n \leftrightarrow_{n+1} n+1)(k))) \\
 &= (n \leftrightarrow_{n+1} n+1)((i \leftrightarrow_{n+1} n)(n)) \\
 &= (n \leftrightarrow_{n+1} n+1)(i) \\
 &\stackrel{i \neq n < n+1}{=} i \\
 &\stackrel{k=n+1=j}{=} (i \leftrightarrow_{n+1} j)(k)
 \end{aligned} \tag{10.17}$$

**$k = n$ .** then

$$\begin{aligned}
 \rho(k) &= (n \leftrightarrow_{n+1} n+1)((i \leftrightarrow_{n+1} n)((n \leftrightarrow_{n+1} n+1)(k))) \\
 &= (n \leftrightarrow_{n+1} n+1)((i \leftrightarrow_{n+1} n)(n+1)) \\
 &\stackrel{i \neq n < n+1}{=} (n \leftrightarrow_{n+1} n+1)(n+1) \\
 &= n \\
 &\stackrel{i \neq n = k = n < n+1 = j}{=} (i \leftrightarrow_{n+1} j)(k)
 \end{aligned}$$

$k = i$ . then

$$\begin{aligned}
 \rho(k) &= (n \leftrightarrow_{n+1} n + 1)((i \leftrightarrow_{n+1} n)((n \leftrightarrow_{n+1} n + 1)(k))) \\
 &\stackrel{k=i < n < n+1}{=} (n \leftrightarrow_{n+1} n + 1)((i \leftrightarrow_{n+1} n)(i)) \\
 &= (n \leftrightarrow_{n+1} n + 1)(n) \\
 &= n + 1 \\
 &\stackrel{j=n+1 \wedge i=k}{=} (i \leftrightarrow_{n+1} j)(k)
 \end{aligned}$$

$k \neq i, n, n+1$ . then

$$\begin{aligned}
 \rho(k) &= (n \leftrightarrow_{n+1} n + 1)((i \leftrightarrow_{n+1} n)((n \leftrightarrow_{n+1} n + 1)(k))) \\
 &\stackrel{k \neq n, n+1}{=} (n \leftrightarrow_{n+1} n + 1)((i \leftrightarrow_{n+1} n)(k)) \\
 &\stackrel{k \neq i, n}{=} (n \leftrightarrow_{n+1} n + 1)(k) \\
 &\stackrel{k \neq n, n+1}{=} k \\
 &\stackrel{k \neq i, j=k+1 \neq k}{=} (i \leftrightarrow_{n+1} j)(k)
 \end{aligned}$$

which proves that

$$\sigma \circ (n \leftrightarrow_{n+1} n + 1) = \rho = (i \leftrightarrow_{n+1} j) \text{ where } \sigma = (n \leftrightarrow_{n+1} n + 1) \circ (i \leftrightarrow_{n+1} n) \quad (10.18)$$

Further

$$\begin{aligned}
 \sigma(n+1) &= (n \leftrightarrow_{n+1} n + 1)((i \leftrightarrow_{n+1} n)(n+1)) \\
 &\stackrel{i < n < n+1}{=} (n \leftrightarrow_{n+1} n + 1)(n+1) \\
 &= n \\
 &= (n \leftrightarrow_{n+1} n + 1)(n+1)
 \end{aligned} \quad (10.19)$$

Hence we have that

$$\begin{aligned}
 &f(x_{(i \leftrightarrow_{n+1} j)(1)}, \dots, x_{(i \leftrightarrow_{n+1} j)(n)}, x_{(i \leftrightarrow_{n+1} j)(n+1)}) \stackrel{10.18}{=} \\
 &f(x_{\sigma((n \leftrightarrow_{n+1} n + 1)(1))}, \dots, x_{\sigma((n \leftrightarrow_{n+1} n + 1)(n+1))}) = \\
 &\quad f(x_{\sigma(1)}, \dots, x_{\sigma(n+1)}) = \\
 &\quad g_{x_{\sigma(n+1)}}(x_{\sigma(1)}, \dots, x_{\sigma(n)}) = \\
 &g_{x_{\sigma(n+1)}}(x_{(n \leftrightarrow_{n+1} n + 1)(i \leftrightarrow_{n+1} j)(1)}, \dots, x_{(n \leftrightarrow_{n+1} n + 1)(i \leftrightarrow_{n+1} j)(n)}) \stackrel{10.16}{=} \\
 &g_{x_{\sigma(n+1)}}(x_{(n \leftrightarrow_{n+1} n + 1)(i \leftrightarrow_{n+1} n)(1)}, \dots, x_{(n \leftrightarrow_{n+1} n + 1)(i \leftrightarrow_{n+1} n)(n)}) \stackrel{n \in \mathcal{S} \text{ and } 10.15}{=} \\
 &\quad g_{x_{\sigma(n+1)}}(x_{(n \leftrightarrow_{n+1} n + 1)(1)}, \dots, x_{(n \leftrightarrow_{n+1} n + 1)(n)}) = \\
 &\quad f(x_{(n \leftrightarrow_{n+1} n + 1)(1)}, \dots, x_{(n \leftrightarrow_{n+1} n + 1)(n)}, x_{\sigma(n+1)}) \stackrel{10.19}{=} \\
 &\quad f(x_{(n \leftrightarrow_{n+1} n + 1)(1)}, \dots, x_{(n \leftrightarrow_{n+1} n + 1)(n+1)}) = \\
 &\quad f(x_1, \dots, x_{n+1})
 \end{aligned}$$

proving that  $f(x_{(i \leftrightarrow_{n+1} j)(1)}, \dots, x_{(i \leftrightarrow_{n+1} j)(n+1)}) = f(x_1, \dots, x_{n+1})$

$i, j \in \{1, \dots, n\}$ . then we have

$$\begin{aligned}
 &f(x_{(i \leftrightarrow_{n+1} j)(1)}, \dots, x_{(i \leftrightarrow_{n+1} j)(n+1)}) = \\
 &g_{x_{(i \leftrightarrow_{n+1} j)(n+1)}}(x_{(i \leftrightarrow_{n+1} j)(1)}, \dots, x_{(i \leftrightarrow_{n+1} j)(n)}) \stackrel{10.24}{=} \\
 &\quad g_{x_{(i \leftrightarrow_{n+1} j)(n+1)}}(x_{(i \leftrightarrow_{n+1} j)(1)}, \dots, x_{(i \leftrightarrow_{n+1} j)(n)}) \stackrel{10.15 \text{ and } n \in \mathcal{S}}{=} \\
 &\quad g_{x_{(i \leftrightarrow_{n+1} j)(n+1)}}(x_1, \dots, x_n) = \\
 &\quad f(x_1, \dots, x_n, x_{(i \leftrightarrow_{n+1} j)(n+1)}) \stackrel{i, j \neq n+1}{=} \\
 &\quad f(x_1, \dots, x_{n+1}) =
 \end{aligned}$$

proving that  $f(x_{(i \leftrightarrow_{n+1} j)(1)}, \dots, x_{(i \leftrightarrow_{n+1} j)(n+1)}) = f(x_1, \dots, x_{n+1})$

So in all cases we have that  $f(x_{(i \leftrightarrow_{n+1} j)(1)}, \dots, x_{(i \leftrightarrow_{n+1} j)(n+1)}) = f(x_1, \dots, x_{n+1})$  proving that  $n+1 \in \mathcal{S}$ .

Mathematical induction proves then the lemma.  $\square$

The above two lemmas can be summarized as follows:

**Theorem 10.73.** *Let  $n \in \mathbb{N}_0 \setminus \{1\}$ ,  $X, Y$  sets and  $f: X^n \rightarrow Y$  such that  $\forall i \in \{1, \dots, n-1\}$  we have that  $f(x_{(i \leftrightarrow_n i+1)(1)}, \dots, x_{(i \leftrightarrow_n i+1)(n)}) = f(x_1, \dots, x_n)$  then  $\forall \sigma \in P_n$  we have that  $f(x_{\sigma(1)}, \dots, x_{\sigma(n)}) = f(x_1, \dots, x_n)$*

**Proof.** Using the previous lemma [see 10.72] it follows that  $\forall i, j \in \{1, \dots, n\}$  we have  $f(x_{(i \leftrightarrow_{n+1} j)(1)}, \dots, x_{(i \leftrightarrow_{n+1} j)(n)}) = f(x_1, \dots, x_n)$ . Hence applying 10.71 it follows that  $\forall \sigma \in P_n$  we have that  $f(x_{\sigma(1)}, \dots, x_{\sigma(n)}) = f(x_1, \dots, x_n)$   $\square$

Next we will prove that  $P_n$  is finite and show how to compute the number of elements in  $P_n$ . First we will show how given a  $\sigma \in P_n$  we construct a new  $\sigma^{[i]} \in P_{n+1}$  such that  $\sigma^{[i]}(n+1) = i$

**Theorem 10.74.** *Given  $n \in \mathbb{N}_0$ ,  $\sigma \in P_n$  and  $i \in \{1, \dots, n+1\}$  define then  $\sigma^{[i]}: \{1, \dots, n+1\} \rightarrow \{1, \dots, n+1\}$  as follows:*

1. *If  $i = n+1$  then  $\forall j \in \{1, \dots, n+1\}$  we have  $\sigma^{[i]}(j) = \begin{cases} \sigma(j) & \text{if } j \in \{1, \dots, n\} \\ n+1 & \text{if } j = n+1 \end{cases}$*
2. *If  $i \in \{1, \dots, n\}$  take then  $k = \sigma^{-1}(i)$  (so that  $\sigma(k) = i$ ) then  $\forall j \in \{1, \dots, n+1\}$  we have  $\sigma^{[i]}(j) = \begin{cases} \sigma(j) & \text{if } j \in \{1, \dots, n\} \setminus \{k\} \\ i & \text{if } j = n+1 \\ n+1 & \text{if } j = k \end{cases}$*

*we have then that  $\sigma^{[i]}$  is a bijection and thus  $\sigma^{[i]} \in P_{n+1}$ . Further from the definition of  $\sigma^{[i]}$  it is immediate clear that  $\sigma^{[i]}(n+1) = i$ .*

**Proof.** We have the following cases to consider

**$i = n+1$ .** If we define the bijection  $\sigma': \{n+1\} \rightarrow \{n+1\}$  by  $\sigma'(n+1) = n+1$  then as  $\{1, \dots, n\} \cap \{n+1\} = \emptyset$  and  $\sigma^{[i]} = \sigma \cup \sigma'$  we have by 2.46 that  $\sigma^{[i]}$  is a bijection.

**$i \in \{1, \dots, n\}$ .** If  $k = \sigma^{-1}(i) \in \{1, \dots, n\}$  we have using 2.55 that  $\sigma|_{\{1, \dots, n\} \setminus \{k\}}: \{1, \dots, n\} \setminus \{k\} \rightarrow \{1, \dots, n\} \setminus \{i\}$  is a bijection. Define  $\sigma': \{k, n+1\} \rightarrow \{i, n+1\}$  by  $\sigma'(j) = \begin{cases} n+1 & \text{if } j = k \\ i & \text{if } j = n+1 \end{cases}$  then because  $n+1 \neq k$  [as  $k \in \{1, \dots, n\}$ ] it follows that  $\sigma'$  is a bijection. We have then trivially that  $\sigma^{[i]} = \sigma|_{\{1, \dots, n\} \setminus \{k\}} \cup \sigma'$ . Using the fact that  $\{1, \dots, n\} \setminus \{k\} \cap \{n+1, k\} = \emptyset$  and  $\{1, \dots, n\} \setminus \{i\} \cap \{i, n+1\} = \emptyset$  we can use 2.46 to conclude that  $\sigma^{[i]}$  is a bijection.  $\square$

We prove now that  $\{1, \dots, n+1\} \times P_n \approx P_{n+1}$  which will allow us to prove by induction that  $P_n$  is finite and have size  $n!$ .

**Theorem 10.75.** *Let  $n \in \mathbb{N}_0$  and  $\Delta: \{1, \dots, n+1\} \times P_n \rightarrow P_{n+1}$  defined by  $\Delta(i, \sigma) = \sigma^{[i]}$  [the previous theorem guarantees that  $\sigma^{[i]} \in P_{n+1}$ ].  $\Delta$  is then proven to be a bijection.*

**Proof.**

1. **(injective)** Assume that  $\Delta(k_1, \sigma_1) = \Delta(k_2, \sigma_2)$  then we have  $\sigma_1^{[k_1]} = \sigma_2^{[k_2]}$ . So  $k_1 = \sigma_1^{[k_1]}(n+1) = \sigma_2^{[k_2]}(n+1) = k_2 \Rightarrow k_1 = k_2$ . We can now consider two cases
  - a.  **$(k_1 = k_2 = n+1)$**  If  $j \in \{1, \dots, n\}$  then  $\sigma_1(j) = \sigma_1^{[n+1]}(j) = \sigma_1^{[k]}(j) = \sigma_2^{[k]}(j) = \sigma_2^{[n+1]}(j) = \sigma_2(j) \Rightarrow \sigma_1(j) = \sigma_2(j) \Rightarrow \sigma_1 = \sigma_2$
  - b.  **$(k_1 = k_2 \neq n+1)$**  take now  $l_1 = \sigma_1^{-1}(k_1), l_2 = \sigma_2^{-1}(k_2)$  then  $\sigma_2^{[k_2]}(l_1) = \sigma_1^{[k_1]}(l_1) = n+1 = \sigma_2^{[k_2]}(l_2)$   $\sigma_2^{[k_2]}$  is a bijection  $\Rightarrow l_1 = l_2$ . Now if  $j \in \{1, \dots, n\}$  then we have the following cases
    - i.  **$(j = l_1 = l_2)$**  then  $\sigma_1(j) = \sigma_1(l_1) = k_1 = k_2 = \sigma_2(l_2) = \sigma_2(j)$
    - ii.  **$(j \neq l_1 = l_2)$**  then  $\sigma_1(j) = \sigma_1^{[k_1]}(j) = \sigma_2^{[k_2]}(j) = \sigma_2(j)$

so we conclude that  $\sigma_1 = \sigma_2$  again.

from (a) and (b) it follows then  $\sigma_1 = \sigma_2$  proving injectivity.

2. (**surjective**) Let  $\sigma \in P_{n+1}$  then we have two cases

- a. ( $\sigma(n+1) = n+1$ ) then clearly  $\sigma = (\sigma|_{\{1, \dots, n\}})^{[n+1]} = \Delta(n+1, \sigma|_{\{1, \dots, n\}})$
- b. ( $\sigma(n+1) \in \{1, \dots, n\}$ ) let now  $k = \sigma(n+1) \in \{1, \dots, n\}$  and  $l = \sigma^{-1}(n+1)$  then  $l \in \{1, \dots, n\}$  [if  $l = n+1 \Rightarrow \sigma(l) \in \{1, \dots, n\} \Rightarrow \sigma(l) \neq n+1$  a contradiction so  $l \in \{1, \dots, n\}$ ] and if  $j \in \{1, \dots, n\} \setminus \{l\}$  we have  $\sigma(j) \in \{1, \dots, n\}$  [if  $\sigma(j) = n+1 = \sigma(l) \xrightarrow{\sigma \text{ is bijective}} l = j$  a contradicting  $j \in \{1, \dots, n\} \setminus \{l\}$ ]. Define now  $\sigma': \{1, \dots, n\} \rightarrow \{1, \dots, n\}$  by

$$\begin{aligned}\sigma'(j) &= \sigma(j) \text{ if } j \in \{1, \dots, n\} \setminus \{l\} \\ &= k \text{ if } j = l\end{aligned}$$

We prove now that  $\sigma'$  is a bijection and thus  $\sigma' \in P_n$ .

### Proof.

i. (**surjective**) Let  $j \in \{1, \dots, n\}$  then we must consider the following cases

- A. ( $j = k$ ) then  $\sigma'(l) = k = j \Rightarrow \sigma'(l) = j$
- B. ( $j \in \{1, \dots, n\} \setminus \{k\}$ ) by surjectivity of  $\sigma$  there exists a  $i \in \{1, \dots, n+1\}$  such that  $\sigma(i) = j$ . We can not have  $i = n+1$  as then  $j = \sigma(i) = \sigma(n+1) = k$  contradicting  $j \in \{1, \dots, n\} \setminus \{k\}$ . Also we cannot have  $i = l$  as then  $j = \sigma(i) = \sigma(l) = n+1$  contradicting  $j \in \{1, \dots, n\} \setminus \{k\}$ . So we must have  $i \in \{1, \dots, n\} \setminus \{l\}$  and thus  $\sigma'(i) = \sigma(i) = j \Rightarrow \sigma'(i) = j$

this proves surjectivity.

ii. (**injective**) Let  $i, j \in \{1, \dots, n\}$  be such that  $\sigma'(i) = \sigma'(j)$  then we must consider the following cases

- A. ( $\sigma'(i) = \sigma'(j) = k$ ) then  $i = l$  [if  $i \neq l$  we have  $i \in \{1, \dots, n\} \setminus \{l\} \Rightarrow \sigma(n+1) = k = \sigma'(i) = \sigma(i) \xrightarrow{\sigma \text{ is a bijection}} i = n+1$  contradicting  $i \in \{1, \dots, n\} \setminus \{l\}$ ], also  $j = l$  [if  $j \neq l$  we have  $j \in \{1, \dots, n\} \setminus \{l\} \Rightarrow \sigma(n+1) = k = \sigma'(j) = \sigma(j) \xrightarrow{\sigma \text{ is a bijection}} j = n+1$  contradicting  $j \in \{1, \dots, n\} \setminus \{l\}$ ] so we conclude that  $i = l = j \Rightarrow i = j$ .
- B. ( $\sigma'(i) = \sigma'(j) \neq k$ ) then  $i, j \neq l$  [otherwise we have  $\sigma'(i) = \sigma'(l) = k = \sigma'(l) = \sigma'(j)$  by the definition of  $\sigma'$ . So  $i, j \in \{1, \dots, n\} \setminus \{l\}$  and thus  $\sigma(i) = \sigma'(i) = \sigma'(j) = \sigma(j) \xrightarrow{\sigma \text{ is bijective}} i = j$ ]

proving injectivity.  $\square$

Consider now  $\sigma'^{[k]} \in P_{n+1}$  and take  $j \in \{1, \dots, n+1\}$  then we have the following cases to consider

i. ( $j = n+1$ ) then

$$\begin{aligned}\sigma'^{[k]}(j) &= \sigma'^{[k]}(n+1) \\ &= k \\ &= \sigma(n+1) \\ &= \sigma(j)\end{aligned}$$

ii. ( $j = l$ ) then as  $\sigma'(l) = k \Rightarrow l = \sigma'^{-1}(k)$

$$\begin{aligned}\sigma'^{[k]}(j) &= \sigma'^{[k]}(l) \\ &= n+1 \\ &= \sigma(l) \\ &= \sigma(j)\end{aligned}$$

iii. ( $j \in \{1, \dots, n\} \setminus \{l\}$ ) then as  $j \neq l = \sigma'^{-1}(k), j \neq n+1$  we have

$$\begin{aligned}\sigma'^{[k]}(j) &= \sigma'(j) \\ &= \sigma(j)\end{aligned}$$

this proves that  $\sigma'^{[k]} = \sigma$  and thus  $\sigma = \Delta(k, \sigma')$  proving surjectivity.  $\square$

We are now finally ready to prove that  $P_n$  is a finite set and count its number of elements.

**Theorem 10.76.** *Let  $n \in \mathbb{N}$  then  $P_n$  is finite and  $\#(P_n) = n!$*

**Proof.** First if  $n = 0$  then  $P_n = P_0$  and  $\emptyset: \emptyset \rightarrow \emptyset$  is the only bijection in  $P_n$  and then  $\#(P_n) = 1 = 0!$ . Assume now that  $n \in \mathbb{N}_0$  then we prove by induction that  $P_n$  is finite with  $\#(P_n) = n!$ . So let  $B = \{n \in \{1, \dots\} | P_n \text{ is finite and } \#(P_n) = (n)!\}$ , we have then:

1. If  $n = 1$  then  $P_n = P_1$  and the only bijection between  $S_1 = \{0\}$  and  $S_1 = \{0\}$  is  $1_{\{0\}}$  so  $P_1$  is finite and  $\#(P_n) = \#(P_1) = 1 = 1!$  so  $0 \in B$
2. Assume that  $n \in B$  then  $P_n$  is finite and  $\#(P_n) = (n)!$ . Then by the previous theorem there exists a bijection  $\Delta: \{1, \dots, n+1\} \times P_n \rightarrow P_{n+1}$ . As  $\{1, \dots, n+1\}$  is finite with  $\#(\{1, \dots, n+1\}) = n+1$  and  $P_n$  is finite and  $\#(P_n) = n$  we have by 5.44 that  $\{1, \dots, n+1\} \times P_n$  is finite and  $\#(\{1, \dots, n+1\} \times P_n) = (n+1) \cdot (n)! = (n+1)!$ . So there exists a bijection  $b: (n+1)! \rightarrow \{1, \dots, n+1\} \times P_n$  and thus a bijection  $\Delta \circ b: (n+1)! \rightarrow P_{n+1}$ . So  $P_{n+1}$  is finite and  $\#(P_{n+1}) = (n+1)!$  or  $n+1 \in B$

By mathematical induction (see 4.77) we have then  $B = \{1, \dots\} = \mathbb{N}_0$   $\square$

**Lemma 10.77.** *If  $m, n \in \mathbb{N}_0$  and  $\{\sigma_i\}_{i \in \{1, \dots, m\}}$  a family of permutations in  $P_n$  then we have that  $(\sigma_1 \circ \dots \circ \sigma_m)^{[n+1]} = (\sigma_1^{[n+1]} \circ \dots \circ \sigma_m^{[n+1]})$*

**Proof.** First we prove that if  $\sigma, \tau \in P_n$  that  $(\sigma \circ \tau)^{[n+1]} = \sigma^{[n+1]} \circ \tau^{[n+1]}$ , so if  $i \in \{1, \dots, n+1\}$  then we have the following cases:

1. ( $i = n+1$ )

$$\begin{aligned} (\sigma \circ \tau)^{[n+1]}(i) &= n+1 \\ &= \sigma^{[n+1]}(n+1) \\ &= \sigma^{[n+1]}(\tau^{[n+1]}(n+1)) \\ &= (\sigma^{[n+1]} \circ \tau^{[n+1]})(n+1) \end{aligned}$$

2. ( $i \in \{1, \dots, n\}$ )

$$\begin{aligned} (\sigma \circ \tau)^{[n+1]}(i) &= (\sigma \circ \tau)(i) \\ &= \sigma(\tau(i)) \\ &\stackrel{\tau(i) \in \{1, \dots, n\}}{=} \sigma^{[n+1]}(\tau(i)) \\ &\stackrel{i \in \{1, \dots, n\}}{=} \sigma^{[n+1]}(\tau^{[n+1]}(i)) \\ &= (\sigma^{[n+1]} \circ \tau^{[n+1]})(i) \end{aligned}$$

proving that indeed  $(\sigma \circ \tau)^{[n+1]} = \sigma^{[n+1]} \circ \tau^{[n+1]}$

We prove this by mathematical induction (see 4.77) so let  $B = \{m \in \{1, \dots\} | (\sigma_1 \circ \dots \circ \sigma_n)^{[n+1]} = (\sigma_1^{[n+1]} \circ \dots \circ \sigma_n^{[n+1]})\}$  then we have

1.  $(\sigma_1 \circ \dots \circ \sigma_1)^{[n+1]} = \sigma_1^{[n+1]} = (\sigma_1^{[n+1]} \circ \dots \circ \sigma_1^{[n+1]}) \Rightarrow 1 \in B$
2. if  $m \in B$  then we have  $(\sigma_1 \circ \dots \circ \sigma_{m+1})^{[n+1]} = ((\sigma_1 \circ \dots \circ \sigma_m) \circ \sigma_{m+1})^{[n+1]} = (\sigma_1 \circ \dots \circ \sigma_m)^{[n+1]} \circ \sigma_{m+1}^{[n+1]} \stackrel{m \in B}{=} (\sigma_1^{[n+1]} \circ \dots \circ \sigma_m^{[n+1]}) \circ \sigma_{m+1}^{[n+1]} = (\sigma_1^{[n+1]} \circ \dots \circ \sigma_m^{[n+1]})$

so  $B = \{1, \dots\}$  proving our theorem.  $\square$

**Lemma 10.78.** *If  $n \in \mathbb{N}_0$  and  $i, j \in \{1, \dots, n\}$  then  $(i \leftrightarrow_{n+1} j) = ((i \leftrightarrow_n j))^{[n+1]}$ . Note that if  $(i \leftrightarrow_n j)$  is a strict  $(i \leftrightarrow_n j)$  transposition ( $i \neq j$ ) then trivially  $((i \leftrightarrow_n j))^{[n+1]}$  is a strict transposition.*

**Proof.** If  $k \in \{1, \dots, n+1\}$  then we have the following cases to consider

1. ( $k = n+1$ )

$$\begin{aligned} (i \leftrightarrow_{n+1} j)(k) &= (i \leftrightarrow_{n+1} j)(n+1) \\ &\stackrel{i, j \in \{1, \dots, n\} \Rightarrow i, j \neq n+1}{=} n+1 \\ &= ((i \leftrightarrow_n j))^{[n+1]}(n+1) \\ &= ((i \leftrightarrow_n j))^{[n+1]}(k) \end{aligned}$$

2. ( $k = i$ )

$$\begin{aligned} (i \leftrightarrow_{n+1} j)(k) &= j \\ &= (i \leftrightarrow_n j)(i) \\ &\stackrel{i \in \{1, \dots, n\} \Rightarrow i \neq n+1}{=} ((i \leftrightarrow_n j))^{[n+1]}(i) \\ &= ((i \leftrightarrow_n j))^{[n+1]}(k) \end{aligned}$$

3. ( $k = j$ )

$$\begin{aligned} (i \leftrightarrow_{n+1} j)(k) &= i \\ &= (i \leftrightarrow_n j)(j) \\ &\stackrel{j \in \{1, \dots, n\} \Rightarrow j \neq n+1}{=} ((i \leftrightarrow_n j))^{[n+1]}(j) \\ &= ((i \leftrightarrow_n j))^{[n+1]}(k) \end{aligned}$$

4. ( $k \neq i, j, n+1$ )

$$\begin{aligned} (i \leftrightarrow_{n+1} j)(k) &\stackrel{k \neq i, j}{=} k \\ &\stackrel{k \neq i, j, n+1}{=} (i \leftrightarrow_n j)(k) \\ &\stackrel{k \neq n+1}{=} ((i \leftrightarrow_n j))^{[n+1]}(k) \end{aligned}$$

So having considered all the cases we conclude that  $(i \leftrightarrow_n j) = ((i \leftrightarrow_n j))^{[n+1]}$   $\square$

By using mathematical induction we can extend the above lemma

**Lemma 10.79.** Given  $n, m \in \mathbb{N}_0$ ,  $0 < m$  and  $\{(k_i \leftrightarrow_n l_i)\}_{i \in \{1, \dots, m\}}$  is a family of transpositions then  $((k_1 \leftrightarrow_{n+1} l_1) \circ \dots \circ (k_m \leftrightarrow_{n+1} l_m)) = (((k_1 \leftrightarrow_n l_1) \circ \dots \circ (k_m \leftrightarrow_n l_m)))^{[n+1]}$

**Proof.**  $((((k_1 \leftrightarrow_n l_1) \circ \dots \circ (k_m \leftrightarrow_n l_m)))^{[n+1]}) \stackrel{\text{previous lemma}}{=} ((k_1 \leftrightarrow_n l_1)^{[n+1]} \circ \dots \circ (k_m \leftrightarrow_n l_m)^{[n+1]})$   $\square$

**Theorem 10.80.** Given  $n \in \mathbb{N}_0$ ,  $n > 1$  then if  $\sigma \in P_n$  there exists a  $\{(k_i \leftrightarrow_n l_i)\}_{i \in \{1, \dots, m\}}$ ,  $\forall i \in \{1, \dots, m\} \models k_i \neq l_i$  and  $k_i, l_i \in \{1, \dots, n\}$  (a family of strict transpositions) such that  $\sigma = ((k_i \leftrightarrow_n l_i) \circ \dots \circ (k_m \leftrightarrow_n l_m))$

**Proof.** We prove this by induction on  $n$  so let  $B = \{n \in \{2, \dots\} \mid \forall \sigma \in P_n \text{ there exists a } \{(k_i \leftrightarrow_n l_i)\}_{i \in \{1, \dots, m\}} \text{ of strict transpositions so that } ((k_1 \leftrightarrow_n l_1) \circ \dots \circ (k_m \leftrightarrow_n l_m))\}$ . We have then :

1. If  $n = 2$  then we have the following possibilities for  $\sigma$

a.  $\sigma$  is defined by

$$\begin{aligned} \sigma(1) &= 1 \\ \sigma(2) &= 2 \end{aligned}$$

so  $\sigma = i_{\{1, 2\}}$  and we can write  $\sigma = i_{\{1, 2\}} = (1 \leftrightarrow_2 2) \circ (1 \leftrightarrow_2 2)$  so if  $\{(k_i, l_i)\}_{i \in \{1, \dots, 2\}} = \{(1 \leftrightarrow_2 2), (1 \leftrightarrow_2 2)\}$  then  $\sigma = ((k_1 \leftrightarrow_2 l_1) \dots (k_2 \leftrightarrow_2 l_2))$

b.  $\sigma$  is defined by

$$\begin{aligned}\sigma(1) &= 2 \\ \sigma(2) &= 1\end{aligned}$$

and then clearly  $\sigma = (1 \leftrightarrow_2 2)$  so if  $\{(k_i, l_i)\}_{i \in \{1, \dots, 2\}} = \{(1 \leftrightarrow_2 2)\}$  then  $\sigma = ((k_1 \leftrightarrow_2 l_1) \dots (k_2 \leftrightarrow_2 l_2))$

proving that  $2 \in B$

2. Assume that  $n \in B$  then given  $\sigma \in P_{n+1}$  we can consider the following cases

a. **( $\sigma(n+1) = n+1$ )** then we have  $\sigma = (\sigma|_{\{1, \dots, n\}})^{[n+1]}$  (see definition of  $\sigma^{[i]}$ ) and as  $\sigma|_{\{1, \dots, n\}} \in P_n$  we have as  $n \in B$  that there exists a family of strict transpositions in  $P_n \{(k_i \leftrightarrow_n l_i)\}_{i \in \{1, \dots, m\}}$  so that  $\sigma|_{\{1, \dots, n\}} = ((k_1 \leftrightarrow_n l_1) \circ \dots \circ (k_m \leftrightarrow_n l_m))$  and thus  $\sigma = (((k_1 \leftrightarrow_n l_1) \circ \dots \circ (k_m \leftrightarrow_n l_m)))^{[n+1]} =_{\text{previous lemma}} ((k_1 \leftrightarrow_{n+1} l_1) \circ \dots \circ (k_m \leftrightarrow_{n+1} l_m))$ . So we have found a family of strict transpositions  $\{(k_i \leftrightarrow_{n+1} l_i)\}_{i \in \{1, \dots, m\}}$  in  $P_{n+1}$  so that  $\sigma = ((k_1 \leftrightarrow_{n+1} l_1) \circ \dots \circ (k_m \leftrightarrow_{n+1} l_m))$  and thus  $n+1 \in B$

b. **( $\sigma(n+1) \neq n+1$ )** Take then  $i = \sigma(n+1) \Rightarrow i \neq n+1$  then we have  $\sigma = i_{\{1, \dots, n+1\}} \circ \sigma = ((i \leftrightarrow_{n+1} n+1) \circ (i \leftrightarrow_{n+1} n+1)) \circ \sigma = (i \leftrightarrow_{n+1} n+1) \circ ((i \leftrightarrow_{n+1} n+1) \circ \sigma) = (i \leftrightarrow_{n+1} n+1) \circ \sigma'$  where  $\sigma' = (i \leftrightarrow_{n+1} n+1) \circ \sigma$ . We have then that  $\sigma'(n+1) = (i \leftrightarrow_{n+1} n+1)(\sigma(n+1)) = (i \leftrightarrow_{n+1} n+1)(i) = n+1$ . Just like in (a) we have then that  $\sigma' = (\sigma'|_{\{1, \dots, n\}})^{[n+1]}$  and as  $\sigma'|_{\{1, \dots, n\}} \in P_n$  we have as  $n \in B$  that there exists a family of strict transpositions  $\{(k_i \leftrightarrow_n l_i)\}_{i \in \{1, \dots, m\}}$  such that  $\sigma'|_{\{1, \dots, n\}} = ((k_1 \leftrightarrow_n l_1) \circ \dots \circ (k_m \leftrightarrow_n l_m))$  and thus  $\sigma' = (\sigma'|_{\{1, \dots, n\}})^{[n+1]} = (((k_1 \leftrightarrow_n l_1) \circ \dots \circ (k_m \leftrightarrow_n l_m)))^{[n+1]} = ((k_1 \leftrightarrow_{n+1} l_1) \circ \dots \circ (k_m \leftrightarrow_{n+1} l_m))$ . Define now  $f = \{(r_i \leftrightarrow_{n+1} s_i)\}_{i \in \{1, \dots, m+1\}}$  by  $\forall o \in \{1, \dots, m+1\}$

$$\begin{aligned}f(o) &= (i \leftrightarrow_{n+1} n+1) \text{ if } o=1 \text{ (a strict transposition)} \\ &= (k_{o-1} \leftrightarrow_{n+1} l_{o-1}) \text{ if } o \in \{2, \dots, m+1\}\end{aligned}$$

then  $((r_1 \leftrightarrow_{n+1} s_1) \circ \dots \circ (r_{m+1} \leftrightarrow_{n+1} s_{m+1})) = (r_1 \leftrightarrow_{n+1} s_1) \circ ((r_2 \leftrightarrow_{n+1} s_2) \circ \dots \circ (r_{m+1} \leftrightarrow_{n+1} s_{m+1})) = (i \leftrightarrow_{n+1} n+1) \circ ((k_1 \leftrightarrow_{n+1} l_1) \circ \dots \circ (k_m \leftrightarrow_{n+1} l_m)) = (i \leftrightarrow_{n+1} n+1) \circ \sigma' = \sigma$  proving that  $n+1 \in B$

Using mathematical induction (see 4.77) we have  $B = \{2, \dots\}$  and thus if  $n \in \mathbb{N}_0, n > 1$  we have  $n \in B$  and thus if  $\sigma \in P_n$  then we can write it as a combination of strict transpositions.  $\square$

**Theorem 10.81.** Given  $n \in \mathbb{N}_0$ ,  $\{x_i\}_{i \in \{1, \dots, n\}}$  a family of elements in  $X$  and  $\emptyset \neq Y \subseteq \{x_i | i \in \{1, \dots, n\}\}$  then there exists a  $m \in \{1, \dots, n\}$  and a permutation  $\sigma \in P_n$  such that  $\forall i \in \{1, \dots, m\}$  we have  $x_{\sigma_i} \in Y$  and  $\forall i \in \{m+1, \dots, n\}$  we have  $x_{\sigma_i} \notin Y$  [and thus  $x_{\sigma_i} \in X \setminus Y$ ]

**Proof.** We prove this by induction so let  $A = \{n \in \{1, \dots\} | \text{if } \{x_i\}_{i \in \{1, \dots, n\}}$  is a family of elements of a set  $X$  and  $\emptyset \neq Y \subseteq \{x_i | i \in \{1, \dots, n\}\}$  then there exists a  $m \in \{1, \dots, n\}$  and a permutation  $\sigma \in P_n$  such that  $\forall i \in \{1, \dots, m\}$  we have  $x_{\sigma_i} \in Y$  and  $\forall i \in \{m+1, \dots, n\}$  we have  $x_{\sigma_i} \notin Y\}$  we have then :

1. If  $n = 1$  then  $\{x_i\}_{i \in \{1, \dots, 1\}}$  is a function from  $\{1\} \rightarrow X$  so  $\{x_i | i \in \{1, \dots, 1\}\} = \{x_1\}$  and if  $\emptyset \neq Y \subseteq \{x_1\}$  we have  $Y = \{x_1\}$  and we can that  $m = 1$  and  $\sigma = 1_{\{1\}}$  so that if  $i \in \{1, \dots, 1\} \Rightarrow i = 1 \Rightarrow x_i = x_1 \in \{x_1\} = Y$  and if  $i \in \{2, \dots, 1\} = \emptyset$  we have that  $x_{\sigma_i} \notin Y$  is full filled vacuously. So  $1 \in A$
2. Assume that  $n \in A$  then if  $\{x_i\}_{i \in \{1, \dots, n+1\}}$  is a family of elements of  $X$  and  $\emptyset \neq Y \subseteq \{x_i | i \in \{1, \dots, n+1\}\}$ . There are two cases to consider:
  - a. **( $Y = \{x_i | i \in \{1, \dots, n+1\}\}$ )** In this case we can take  $m = n+1$  and  $\sigma = 1_{\{1, \dots, n+1\}}$  and then  $\forall i \in \{1, \dots, m\} = \{1, \dots, n+1\}$  we have  $x_{\sigma_i} = x_i \in \{x_i | i \in \{1, \dots, n+1\}\} = Y$  and if  $i \in \{m+1, \dots, n+1\} = \emptyset$  then  $x_{\sigma_i} \notin Y$  is full filled vacuously. So we conclude that  $n+1 \in A$

b. ( $Y \subset \{x_i | i \in \{1, \dots, n+1\}\}$ ) then there exists a  $y \in X$  such that  $y \notin Y$ . As  $y \in X$  there exists a  $j \in \{1, \dots, n+1\}$  such that  $x_j = y$ . Take then  $\sigma' \in P_{n+1}$  to be  $\sigma' = (j \leftrightarrow_{n+1} n+1)$ . We have then for  $\{x_{\sigma'_i}\}_{i \in \{1, \dots, n+1\}}$  that  $x_{\sigma'_{n+1}} = x_{(j \leftrightarrow_{n+1} n+1)(n+1)} = x_j = y \notin Y$ . This means that for  $\{x_{\sigma'_i}\}_{i \in \{1, \dots, n\}}$  the restriction (as a function) of  $\{x_{\sigma'_i}\}_{i \in \{1, \dots, n+1\}}$  to  $\{1, \dots, n\}$  we have that if  $z \in Y$  then from  $Y \subseteq \{x_i | i \in \{1, \dots, n+1\}\}$  that there exists a  $i \in \{1, \dots, n+1\}$  such that  $z = x_i$ , as  $\sigma'$  is a bijection there exists a  $k \in \{1, \dots, n+1\}$  such that  $i = \sigma'(k)$ , now if  $k = n+1$  then  $i = \sigma'(k) = \sigma'(n+1) = j \Rightarrow z = x_i = x_j = y \notin Y$  a contradiction, so  $k \in \{1, \dots, n\}$  and thus  $z \in \{x_{\sigma'_i} | i \in \{1, \dots, n\}\}$  proving that  $\emptyset \neq Y \subseteq \{x_{\sigma'_i} | i \in \{1, \dots, n\}\}$ . Using the fact that  $n \in A$  there exists a  $m \in \{1, \dots, n\}$  and a  $\sigma'' \in P_n$  [ $\sigma'': \{1, \dots, m\} \rightarrow \{1, \dots, n\}$  is a bijection] such that  $\forall i \in \{1, \dots, m\}$  we have  $x_{\sigma''_i} \in Y$  and  $\forall i \in \{m+1, \dots, n\}$  we have  $x_{\sigma''_i} \notin Y$ . Using the bijection  $i_{n+1}: \{n+1\} \rightarrow \{n+1\}$  we construct (see 2.46) the bijection  $\sigma''' = \sigma'' \cup i_{\{n+1\}}: \{1, \dots, n+1\} \rightarrow \{1, \dots, n+1\}$  so  $\sigma''' \in P_{n+1}$ . Take then  $\sigma = \sigma' \circ \sigma'''$  then  $\sigma \in P_{n+1}$  and if  $i \in \{1, \dots, m\}$  we have  $\sigma(i) = \sigma'(\sigma'''(i)) \underset{\{1, \dots, m\} \subseteq \{1, \dots, n\}}{=} \sigma'(\sigma''(i))$  and thus  $x_{\sigma_i} = x_{\sigma''_i} \in Y$ . Also if  $i \in \{m+1, \dots, n+1\}$  then we have either

- i. ( $i = n+1$ ) then  $\sigma(i) = \sigma'(\sigma'''(n+1)) = \sigma'(n+1) = j \Rightarrow x_{\sigma_i} = y \notin Y$
- ii. ( $i \neq n+1$ ) then  $i \in \{m+1, \dots, n\}$  and  $\sigma(i) = \sigma'(\sigma'''(i)) \underset{\{m+1, \dots, n\} \subseteq \{1, \dots, n\}}{=} \sigma'(\sigma''(i))$  and thus  $x_{\sigma_i} = x_{\sigma''_i} \notin Y$

Having found our  $m \in \{1, \dots, n\} \subseteq \{1, \dots, n+1\}$  and  $\sigma \in P_{n+1}$  we can say that  $n+1 \in A$  Using mathematical induction (see 4.77) we have then that  $A = \{1, \dots\} = \mathbb{N}_0$  proving the theorem.  $\square$

**Definition 10.82.** If  $X$  is a set,  $n \in \mathbb{N}_0$  define then the family  $\{X_i\}_{i \in \{1, \dots, n\}}$  by  $\forall i \in \{1, \dots, n\}$  we have  $X_i = X$  then  $X^n = \prod_{i \in \{1, \dots, n\}} X_i$ . Using the definition of a 2.81 we have then  $X^n = \{f | f: \{1, \dots, n\} \rightarrow X \text{ is a function}\}$ . If  $x \in X^n$  then  $\forall i \in \{1, \dots, n\}$  we note  $x(i)$  as  $x_i$  this is the reason why another notation for the elements of  $X^n$  is  $(x_1, \dots, x_n)$  (especially if we have a formula to calculate  $x(x) = x_i$ ). If  $x \in X^n$  and  $\sigma \in P_n$  then  $x \circ \sigma: \{1, \dots, n\} \rightarrow X$  and thus  $x \circ \sigma \in X^n$  we note this element as  $(x_{\sigma(1)}, \dots, x_{\sigma(n)})$  or  $(x_{\sigma_1}, \dots, x_{\sigma_n})$

**Example 10.83.** If  $n = 1$  then  $X^1 = \{f: f: \{1\} \rightarrow x\}$  we can now find a trivial bijection between  $X$  and  $X^1$  given  $x \in X$  from  $\{1\}(x): \{1\} \rightarrow X$  by  $\{1\}(x)(1) = x$

1. (**injectivity**) If  $\{1\}(x) = \{1\}(y) \Rightarrow x = \{1\}(x)(1) = \{1\}(y) = y \Rightarrow x = y$
2. (**surjectivity**) If  $f \in X^1$  take then  $x = f(1)$  then  $\{1\}(x)(1) = x = f(1) \Rightarrow \{1\}(x) = f$

**Definition 10.84.** Let  $X, Y$  be sets,  $n \in \mathbb{N}_0$ ,  $f: X^n \rightarrow Y$  a function and  $\sigma \in P_n$  then  $\sigma f: X^n \rightarrow Y$  is defined by  $\forall x \in X^n$  we have  $(\sigma f)(x) = f(x \circ \sigma)$  or in other notation  $(\sigma f)(x_1, \dots, x_n) = f(x_{\sigma_1}, \dots, x_{\sigma_n})$

**Theorem 10.85.** Let  $X, Y$  be sets,  $\sigma, \tau \in P_n$  and  $f: X^n \rightarrow Y$  then  $\tau(\sigma f) = (\tau \circ \sigma) f$

**Proof.** If  $x \in X^n$  then

$$\begin{aligned} (\tau(\sigma f))(x) &= (\sigma f)(x \circ \tau) \\ &= f((x \circ \tau) \circ \sigma) \\ &= f(x \circ (\tau \circ \sigma)) \\ &= ((\tau \circ \sigma) f)(x) \end{aligned}$$

and thus we conclude that  $\tau(\sigma f) = (\tau \circ \sigma) f$   $\square$

**Definition 10.86.** Given  $n \in \mathbb{N}_0$  then  $\Phi_n: \mathbb{Z}^n \rightarrow \mathbb{Z}$  is defined by  $\Phi_n(x) = \prod_{(i, j) \in \{(i, j) \in \{1, \dots, n\} \times \{1, \dots, n\} | i < j\}} (x_i - x_j)$

**Example 10.87.**

1.  $\Phi_1(2) = 1$  as  $\{(i, j) \in \{1\} \times \{1\} | i < j\} = \emptyset$  (here 1 is neutral element in  $\langle \mathbb{Z}, \cdot \rangle$ )
2.  $\Phi_2(1, 2) = \prod_{(i, j) \in \{(1, 2)\}} (x_i - x_j) = (1-2)$
3.  $\Phi_3(1, 2, 3) = (1-2) \cdot (1-3) \cdot (2-3) = 2$

**Theorem 10.88.** Given  $n \in \mathbb{N}_0$ ,  $1 < n$  then if  $k, l \in \{1, \dots, n\}$ ,  $k \neq l$  and  $\sigma = (k \leftrightarrow_n l)$  is a strict transposition then  $\sigma \Phi_n = (-1) \cdot \Phi_n$

**Proof.** First define  $I = \{(i, j) \in \{1, \dots, n\} \times \{1, \dots, n\} \mid i < j\}$  which is finite, then if  $k \neq l$  we have the following possibilities:

1. ( $k < l$ ) If  $(i, j) \in I \Rightarrow i < j$  then we have the following excluding possibilities

- a. ( $i = k, j = l$ )  $\Rightarrow (i, j) \in \{(k, l)\} = I_1 \subseteq I$
- b. ( $i = k, j \neq l$ ) then we have the following sub-cases

i. ( $j < l$ )  $\underset{k=i < j}{\Rightarrow} k < j < l \Rightarrow (i, j) \in \{(k, j) \in I \mid k < j < l\} = I_2 \subseteq I$

ii. ( $l < j$ )  $\Rightarrow (i, j) \in \{(k, j) \in I \mid l < j\} = I_3 \subseteq I$

- c. ( $i \neq k$ ) with the following sub-cases

i. ( $i = l$ )  $\underset{i < j}{\Rightarrow} (i, j) \in \{(l, j) \in I \mid l < j\} = I_4 \subseteq I$

ii. ( $i \neq l$ ) with the following sub-cases

A. ( $j = k$ )  $\underset{i < j}{\Rightarrow} i < k \Rightarrow (i, j) \in \{(i, k) \in I \mid i < k\} = I_5 \subseteq I$

B. ( $j = l$ ) then we have the following sub-cases

1. ( $i < k$ )  $\Rightarrow (i, j) \in \{(i, l) \in I \mid i < k\} = I_6 \subseteq I$

2. ( $k < i$ )  $\underset{i < j = l}{\Rightarrow} (i, j) \in \{(i, l) \in I \mid k < i < l\} = I_7 \subseteq I$

C. ( $j \neq l, k$ )  $\Rightarrow (i, j) \in \{(i, j) \in I \mid i \neq k, l \wedge j \neq k, l\} = I_8 \subseteq I$

proving that  $I \subseteq \bigcup_{k \in \{1, \dots, 8\}} I_k \subseteq I \Rightarrow I = \bigcup_{k \in \{1, \dots, 8\}} I_k$ . Now for the intersections of the form  $I_i \cap I_j, i \neq j$  we have the following cases to consider:

- a.  $((i, j) \in I_1 \cap I_2)$  then  $i = k \wedge j = l \wedge j < l \Rightarrow$  a contradiction
- b.  $((i, j) \in I_1 \cap I_3)$  then  $i = k \wedge j = l \wedge l < j \Rightarrow$  a contradiction
- c.  $((i, j) \in I_1 \cap I_4)$  then  $i = k \wedge j = l \wedge l < j \Rightarrow$  a contradiction
- d.  $((i, j) \in I_1 \cap I_5)$  then  $i = k \wedge j = l \wedge i < k \Rightarrow$  a contradiction
- e.  $((i, j) \in I_1 \cap I_6)$  then  $i = k \wedge j = l \wedge i < k \Rightarrow$  a contradiction
- f.  $((i, j) \in I_1 \cap I_7)$  then  $i = k \wedge j = l \wedge k < i < l \Rightarrow$  a contradiction
- g.  $((i, j) \in I_1 \cap I_8)$  then  $i = k \wedge j = l \wedge i \neq k \Rightarrow$  a contradiction
- h.  $((i, j) \in I_2 \cap I_3)$  then  $i = k \wedge k < j < l \wedge l < j \Rightarrow$  a contradiction
- i.  $((i, j) \in I_2 \cap I_4)$  then  $i = k \wedge k < j < l \wedge i = l \wedge l < j \Rightarrow$  a contradiction
- j.  $((i, j) \in I_2 \cap I_5)$  then  $i = k \wedge k < j < l \wedge k = j \Rightarrow$  a contradiction
- k.  $((i, j) \in I_2 \cap I_6)$  then  $i = k \wedge k < j < l \wedge j = l \wedge i < k \Rightarrow$  a contradiction
- l.  $((i, j) \in I_2 \cap I_7)$  then  $i = k \wedge k < j < l \wedge j = l \wedge k < i < l \Rightarrow$  a contradiction
- m.  $((i, j) \in I_2 \cap I_8)$  then  $i = k \wedge k < j < l \wedge i, j \neq k, l \Rightarrow$  a contradiction
- n.  $((i, j) \in I_3 \cap I_4)$  then  $i = k \wedge l < j \wedge i = l \wedge l < j \Rightarrow$  a contradiction
- o.  $((i, j) \in I_3 \cap I_5)$  then  $i = k \wedge l < j \wedge i < k \wedge j = k \Rightarrow$  a contradiction
- p.  $((i, j) \in I_3 \cap I_6)$  then  $i = k \wedge l < j \wedge j = l \wedge i < k \Rightarrow$  a contradiction
- q.  $((i, j) \in I_3 \cap I_7)$  then  $i = k \wedge l < j \wedge j = l \wedge k < i < l \Rightarrow$  a contradiction
- r.  $((i, j) \in I_3 \cap I_8)$  then  $i = k \wedge l < j \wedge i, j \neq k, l \Rightarrow$  a contradiction
- s.  $((i, j) \in I_4 \cap I_5)$  then  $i = l \wedge l < j \wedge j = k \wedge i < k \Rightarrow l < k$  contradicting  $k < l$
- t.  $((i, j) \in I_4 \cap I_6)$  then  $i = l \wedge l < j \wedge j = l \wedge i < k \Rightarrow l < l$  a contradiction
- u.  $((i, j) \in I_4 \cap I_7)$  then  $i = l \wedge l < j \wedge j = l \wedge k < i < l \Rightarrow l < l$  a contradiction
- v.  $((i, j) \in I_4 \cap I_8)$  then  $i = l \wedge l < j \wedge i, j \neq k, l \Rightarrow$  s contradiction
- w.  $((i, j) \in I_5 \cap I_6)$  then  $j = k \wedge i < k \wedge j = l \wedge i < k \Rightarrow k = l$  contradicting  $k < l$
- x.  $((i, j) \in I_5 \cap I_7)$  then  $j = k \wedge i < k \wedge j = l \wedge k < i < l \Rightarrow i < i$  a contradiction

- y.  $((i, j) \in I_5 \cap I_8)$  then  $j = k \wedge i < k \wedge i, j \neq k, l \Rightarrow$  a contradiction
- z.  $((i, j) \in I_6 \cap I_7)$  then  $j = l \wedge i < k \wedge j = l \wedge k < i < l \Rightarrow i < i$  a contradiction
- aa.  $((i, j) \in I_6 \cap I_8)$  then  $j = l \wedge i < k \wedge i, j \neq k, l \Rightarrow$  a contradiction
- ab.  $((i, j) \in I_7 \cap I_8)$  then  $j = l \wedge k < i < l \wedge i, j \neq k, l \Rightarrow$  a contradiction

So we have (taking commutativity of  $\cap$  in account) that  $\forall i, j \in \{1, \dots, n\}$  with  $i \neq j$  we have  $I_i \cap I_j = \emptyset$ . Using 10.45 we have that

$$\begin{aligned}\Phi(x_1, \dots, x_n) &= \left( \prod_{k=1}^8 \left( \prod_{(i,j) \in I_k} (x_i - x_j) \right) \right) \\ &= Q_1 \cdot Q_2 \cdot Q_3 \cdot Q_4 \cdot Q_5 \cdot Q_6 \cdot Q_7 \\ \text{here } Q_m &= \prod_{(i,j) \in I_m} (x_i - x_j)\end{aligned}$$

and

$$\begin{aligned}\sigma \Phi(x_1, \dots, x_n) &= \Phi(x_{\sigma_1}, \dots, x_{\sigma_n}) \\ &= \prod_{k=1}^8 \left( \prod_{(i,j) \in I_k} (x_{\sigma_i} - x_{\sigma_j}) \right) \\ &= P_1 \cdot P_2 \cdot P_3 \cdot P_4 \cdot P \\ \text{here } P_m &= \prod_{(i,j) \in I_m} (x_{\sigma_i} - x_{\sigma_j})\end{aligned}$$

Now we have the following 8 products to calculate

- a.  $P_1 = \prod_{(i,j) \in I_1} (x_{\sigma_i} - x_{\sigma_j}) = (x_{\sigma_k} - x_{\sigma_l}) = (x_l - x_k) = -(x_k - x_l) = -\prod_{(i,j) \in I_1} (x_i - x_j) = -Q_1$
- b.  $P_2 = \prod_{(i,j) \in I_2} (x_{\sigma_i} - x_{\sigma_j}) = \prod_{(i,j) \in \{(k,j) \in I \mid k < j < l\}} (x_{\sigma_i} - x_{\sigma_j}) = \prod_{(i,j) \in \{(k,j) \in I \mid k < j < l\}} (x_{\sigma_k} - x_{\sigma_l}) = \prod_{(i,j) \in \{(k,j) \in I \mid k < j < l\}} (x_l - x_j) = \prod_{(i,j) \in I_2} A_{(i,j)}$  where  $A_{(i,j)} = x_l - x_j$ . Now we construct the following function  $b: I_7 = \{(i, l) \in I \mid k < i < l\} \rightarrow I_2 = \{(k, j) \in I \mid k < j < l\}$  by  $(i, j) \rightarrow b(i, j) = (k, i)$  this is a bijection as
  - i. **(injectivity)** If  $b(i, j) = b(i', j')$   $\underset{(i,j) \in I_7}{\Rightarrow} j = l = j' \wedge (k, i) = (k, i') \Rightarrow i = i'$  and thus  $(i, j) = (i', j')$ .
  - ii. **(surjectivity)** If  $(i, j) \in I_2 \Rightarrow k = i \wedge k < j < l \Rightarrow (i, j) = (k, j)$  and  $(j, l) \in I_7$  and  $b(j, l) = (k, j) = (i, j)$

Using 10.43 (applied to products) we get  $\prod_{(i,j) \in I_2} A_{(i,j)} = \prod_{(i,j) \in I_7} A_{b(i,j)} = \prod_{(i,j) \in I_7} A_{(k,i)} = \prod_{(i,j) \in I_7} (x_l - x_i) = \prod_{(i,j) \in I_7} (-1) \cdot (x_i - x_l) \underset{10.54}{=} \left( \prod_{(i,j) \in I_7} (x_i - x_l) \right) (-1)^{\#(I_7)} \underset{(i,j) \in I_7 \Rightarrow j = l}{\Rightarrow} \left( \prod_{(i,j) \in I_7} (x_i - x_j) \right) \cdot (-1)^{\#(I_7)} = Q_7 \cdot (-1)^{\#(I_7)}$  so we have proved that

$$P_2 = (-1)^{\#(I_7)} \cdot Q_7$$

- c.  $P_3 = \prod_{(i,j) \in I_3} (x_{\sigma_i} - x_{\sigma_j}) \underset{(i,j) \in I_3 \Rightarrow i = k \text{ and } j \neq l}{=} \prod_{(i,j) \in I_3} (x_l - x_j) = \prod_{(i,j) \in I_3} A_{(i,j)}$  where here  $A_{i,j} = x_l - x_j$ . Define now  $b: I_4 = \{(l, j) \in I \mid l < j\} \rightarrow I_3 = \{(k, j) \in I \mid l < j\}$  defined by  $(i, j) \rightarrow (k, j)$  then  $b$  is bijective:

- i. **(injectivity)** If  $b(i, j) = b(i', j')$   $\underset{(i,j), (i',j') \in I_4}{\Rightarrow} i = l = i' \wedge (k, j) = (k, j') \Rightarrow i = i' \wedge j = j' \Rightarrow (i, j) = (i', j')$
- ii. **(surjectivity)** If  $(i, j) \in I_3 \Rightarrow i = k \wedge l < j \Rightarrow (i, j) = (k, j) \wedge (l, j) \in I_4$  and  $b(l, j) = (k, j) = (i, j)$

Using 10.43 we have  $\prod_{(i,j) \in I_3} A_{(i,j)} = \prod_{(i,j) \in I_4} A_{b(i,j)} = \prod_{(i,j) \in I_4} A_{(k,j)} = \prod_{(i,j) \in I_4} (x_l - x_j) \underset{(i,j) \in I_4 \Rightarrow i = l}{\Rightarrow} \prod_{(i,j) \in I_4} (x_i - x_j) = Q_4$  so we have proved that

$$P_3 = Q_4$$

- d.  $P_4 = \prod_{(i,j) \in I_4} (x_{\sigma_i} - x_{\sigma_j})$   $\underset{(i,j) \in I_4 \Rightarrow i=l \wedge k < l < j}{=} \prod_{(i,j) \in I_4} (x_k - x_j) = \prod_{(i,j) \in I_4} A_{(i,j)}$   
 where  $A_{(i,j)} = x_k - x_j$ . Define now  $b: I_3 = \{(k, j) \in I \mid l < j\} \rightarrow I_4 = \{(l, j) \in I \mid l < j\}$  by  $(i, j) \rightarrow b(i, j) = (l, j)$  then  $b$  is bijective:

- i. **(injectivity)** If  $b(i, j) = b(i', j')$   $\underset{(i,j),(i',j') \in I_3}{\Rightarrow} i = k = i' \wedge (l, j) = (l, j') \Rightarrow (i, j) = (i', j')$
- ii. **(surjectivity)** If  $(i, j) \in I_4 \Rightarrow i = l \wedge l < j \Rightarrow (i, j) = (l, j) \wedge (k, j) \in I_3$  and  $b(k, j) = (l, j) = (i, j)$

Using 10.43 we have then that  $\prod_{(i,j) \in I_4} A_{(i,j)} = \prod_{(i,j) \in I_3} A_{b(i,j)} = \prod_{(i,j) \in I_3} A_{(l,j)} = \prod_{(i,j) \in I_3} (x_k - x_j)$   $\underset{(i,j) \in I_3 \Rightarrow i=k}{=} \prod_{(i,j) \in I_3} (x_i - x_j) = Q_3$ , so we have proved that

$$P_4 = Q_3$$

- e.  $P_5 = \prod_{(i,j) \in I_5} (x_{\sigma_i} - x_{\sigma_j})$   $\underset{(i,j) \in I_5 \Rightarrow j=k \wedge i < k < l}{=} \prod_{(i,j) \in I_5} (x_i - x_l) = \prod_{(i,j) \in I_5} A_{(i,j)}$  where  $A_{(i,j)} = x_i - x_l$ . Define now  $b: I_6 = \{(i, l) \in I \mid i < k\} \rightarrow I_5 = \{(i, k) \in I \mid i < k\}$  by  $(i, j) = b(i, j) = (i, k)$  then  $b$  is a bijection :

- i. **(injectivity)** If  $b(i, j) = b(i', j')$   $\underset{(i,j),(i',j') \in I_6}{\Rightarrow} j = l = j' \wedge (i, k) = (i', k) \Rightarrow (i, j) = (i', j')$
- ii. **(surjectivity)** If  $(i, j) \in I_5 \Rightarrow (i, j) = (i, k) \wedge i < k \Rightarrow (i, l) \in I_6$  and  $b(i, l) = (i, k) = (i, j)$

Using 10.43 we have  $\prod_{(i,j) \in I_5} A_{(i,j)} = \prod_{(i,j) \in I_6} A_{b(i,j)} = \prod_{(i,j) \in I_6} A_{(i,k)} = \prod_{(i,j) \in I_6} (x_i - x_l)$   $\underset{(i,j) \in I_6 \Rightarrow j=l}{=} \prod_{(i,j) \in I_6} (x_i - x_j) = Q_6$ , so we have proved that

$$P_5 = Q_6$$

- f.  $P_6 = \prod_{(i,j) \in I_6} (x_{\sigma_i} - x_{\sigma_j})$   $\underset{(i,j) \in I_6 \Rightarrow j=l \wedge i < k < l}{=} \prod_{(i,j) \in I_6} (x_i - x_k) = \prod_{(i,j) \in I_6} A_{(i,j)}$  where  $A_{(i,j)} = x_i - x_k$ . Define now  $b: I_5 = \{(i, k) \in I \mid i < k\} \rightarrow I_6 = \{(i, l) \mid i < k\}$  by  $(i, j) \rightarrow b(i, j) = (i, l)$  then  $b$  is a bijection :

- i. **(injectivity)** If  $b(i, j) = b(i', j')$   $\underset{(i,j),(i',j') \in I_5}{\Rightarrow} j = k = j' \wedge (i, l) = (i', l) \Rightarrow (i, j) = (i', j')$
- ii. **(surjectivity)** If  $(i, j) \in I_6 \Rightarrow (i, j) = (i, l) \cap i < k \Rightarrow (i, k) \in I_5$  and  $b(i, k) = (i, l) = (i, j)$

Using 10.43 we have  $\prod_{(i,j) \in I_6} A_{(i,j)} = \prod_{(i,j) \in I_5} A_{b(i,j)} = \prod_{(i,j) \in I_5} A_{(i,l)} = \prod_{(i,j) \in I_5} (x_i - x_k)$   $\underset{(i,j) \in I_5 \Rightarrow j=k}{=} \prod_{(i,j) \in I_5} (x_i - x_j) = Q_5$ , so we have proved that

$$P_6 = Q_5$$

- g.  $P_7 = \prod_{(i,j) \in I_7} (x_{\sigma_i} - x_{\sigma_j})$   $\underset{(i,j) \in I_7 \Rightarrow j=l \wedge k < i < l}{=} \prod_{(i,j) \in I_7} (x_i - x_k) = \prod_{(i,j) \in I_7} A_{(i,j)}$  where  $A_{(i,j)} = x_i - x_k$ . Define now  $b: I_2 = \{(k, j) \in I \mid k < j < l\} \rightarrow I_7 = \{(i, l) \in I \mid k < i < l\}$  by  $(i, j) \rightarrow (j, l)$  then  $b$  is a bijection:

- i. **(injectivity)** If  $b(i, j) = b(i', j')$   $\underset{(i,j),(i',j') \in I_7}{\Rightarrow} i = k = i' \wedge (j, l) = (j', l) \Rightarrow (i, j) = (i', j')$
- ii. **(surjectivity)** If  $(i, j) \in I_7 \Rightarrow (i, j) = (i, l) \wedge k < i < l \Rightarrow (k, i) \in I_2$  and  $b(k, i) = (i, l) = (i, j)$

Using 10.43 we have that  $\prod_{(i,j) \in I_7} A_{(i,j)} = \prod_{(i,j) \in I_2} A_{b(i,j)} = \prod_{(i,j) \in I_2} A_{(j,l)} = \prod_{(i,j) \in I_2} (x_j - x_k)$   $\underset{(i,j) \in I_2 \Rightarrow i=k}{=} \prod_{(i,j) \in I_2} (x_j - x_i) = \prod_{(i,j) \in I_2} ((-1) \cdot (x_i - x_j)) = (\prod_{(i,j) \in I_2} (x_i - x_j)) \cdot (-1)^{\#(I_2)}$   $\underset{I_2 \approx I_7 \Rightarrow \#(I_2) = \#(I_7)}{=} Q_2 \cdot (-1)^{\#(I_7)}$ , this proves that

$$P_7 = (-1)^{\#(I_7)} \cdot Q_2$$

- h.  $P_8 = \prod_{(i,j) \in I_8} (x_{\sigma_i} - x_{\sigma_j})$   $\underset{(i,j) \in I_8 \Rightarrow i,j \neq k,l}{=} \prod_{(i,j) \in I_8} (x_i - x_j) = Q_8$

So using the above easy but elaborate calculations we have that

$$\begin{aligned}
 \sigma\Phi(x_1, \dots, x_n) &= P_1 \cdot P_2 \cdot P_3 \cdot P_4 \cdot P_5 \cdot P_6 \cdot P_7 \cdot P_8 \\
 &= (-1) \cdot Q_1 \cdot (-1)^{\#(I_7)} \cdot Q_7 \cdot Q_4 \cdot Q_3 \cdot Q_6 \cdot Q_5 \cdot (-1)^{\#(I_7)} \cdot Q_2 \cdot Q_8 \\
 &= (-1) \cdot Q_1 \cdot Q_2 \cdot Q_3 \cdot Q_4 \cdot Q_5 \cdot Q_6 \cdot Q_7 \cdot Q_8 \\
 &= -\Phi(x_1, \dots, x_n)
 \end{aligned}$$

where we used associativity, commutativity and 6.19.

2. ( $l < k$ ) Take then  $l' = k$  and  $k' = l$ , then we have  $k' < l'$  and if  $(k \leftrightarrow_n l) = (l' \leftrightarrow_n k') = (k' \leftrightarrow_n l')$  then we have

$$\begin{aligned}
 (k \leftrightarrow_n l)\Phi(x_1, \dots, x_n) &= (k' \leftrightarrow_n l')\Phi(x_1, \dots, x_n) \\
 &\stackrel{\text{(use (1))}}{=} -\Phi(x_1, \dots, x_n) \\
 &\square
 \end{aligned}$$

**Lemma 10.89.** *If  $\{x_i\}_{i \in I}$  is a finite family of elements in  $\mathbb{Z}$  where  $I$  is a finite set and  $\forall i \in I$  we have  $x_i \neq 0$  then  $\prod_{i \in I} x_i \neq 0$*

**Proof.** First we prove by induction on  $n \in \mathbb{N}_0$  that if  $\{x_i\}_{i \in \{0, \dots, n-1\}}$  is such that  $x_i \neq 0$  then  $\prod_{i=0}^{n-1} x_i \neq 0$ . So let  $X = \{n \in \{1, \dots\} \mid \text{If } \{x_i\}_{i \in \{0, \dots, n-1\}} \text{ is such that } x_i \neq 0 \Rightarrow \prod_{i=0}^{n-1} x_i \neq 0\}$  then

1.  $\prod_{i=0}^{1-1} x_i = x_0 \neq 0 \Rightarrow 1 \in X$
2. If  $n \in X \Rightarrow \prod_{i=0}^{(n+1)-1} = \prod_{i=0}^{(n-1)+1} x_i = (\prod_{i=0}^{n-1} x_i) \cdot x_n \neq 0$  because  $n \in X$  and  $(\mathbb{Z}, +, \cdot)$  is a integral domain. So  $n+1 \in X$

Using mathematical induction (see 4.77) we have  $X = \{1, \dots\} = \mathbb{N}_0$ .

Now if  $I$  is finite and  $n = \#(I)$  and  $h: S_n \rightarrow I$  a bijection then by 10.50 we have  $\prod_{i \in I} x_i = \prod_{i=0}^{n-1} x_{h_i} \neq 0$  [previous proof and the fact that  $x_{h_i} \neq 0$ ]  $\square$

**Theorem 10.90.** *If  $n \in \mathbb{N}_0$  then  $\forall \sigma \in P_n$  there exists a unique  $\varepsilon_\sigma \in \{-1, 1\}$  such that  $\sigma\Phi = \varepsilon_\sigma\Phi$  (where  $\Phi$  is as defined in 10.86). Further if  $\sigma = ((i_1 \leftrightarrow j_1) \circ \dots \circ (i_m \leftrightarrow j_m))$  then  $\varepsilon_\sigma = (-1)^m$  (which by 6.19 means that  $\varepsilon_\sigma = -1$  if  $\sigma$  is a odd number of strict transpositions and 1 if  $\sigma$  is a even number of strict transpositions).*

**Proof.**

We do this proof in two stages

1. (**Existence**) If  $n = 1$  then  $\sigma = i_{\{1\}}$  and  $\sigma\Phi(x) = \Phi(x \circ \sigma) = \Phi(x) \Rightarrow \sigma\Phi = \Phi$ . If  $\sigma \in P_n, n \in \{2, \dots\}$  then by 10.80  $\sigma = ((i_1 \leftrightarrow_n j_1) \circ \dots \circ (i_m \leftrightarrow_n j_m))$ . We prove now by induction on  $n$  that  $\sigma\Phi = (-1)^m\Phi$ . So let  $X = \{m \in \{1, \dots\} \mid \text{If } \sigma = ((i_1 \leftrightarrow_n j_1) \circ \dots \circ (i_m \leftrightarrow_n j_m)) \text{ then } \sigma\Phi = (-1)^m\Phi\}$  then we have

- a. If  $n = 1 \Rightarrow \sigma = (i_1 \leftrightarrow_n j_1)$  and using 10.88 we have  $\sigma\Phi = -\Phi = (-1)^1\Phi$  so  $1 \in X$
- b. If  $m \in X$  then if  $\sigma = ((i_1 \leftrightarrow_n i_1) \circ \dots \circ (i_{m+1} \leftrightarrow_n j_{m+1})) = \sigma' \circ (i_{m+1} \leftrightarrow_n j_{m+1})$  where  $\sigma' = ((i_1 \leftrightarrow_n i_1) \circ \dots \circ (i_m \leftrightarrow_n j_m))$  and using 10.85 we have  $(\sigma' \circ (i_{m+1} \leftrightarrow_n j_{m+1}))\Phi = \sigma'((i_{m+1} \leftrightarrow_n j_{m+1})\Phi) = \sigma'(-\Phi) = -\sigma'\Phi \stackrel{n \in X}{=} -(-1)^m\Phi = (-1)^{m+1}\Phi$  proving that  $m+1 \in X$

Using induction (see 4.10) we have  $X = \{1, \dots\}$  proving the existence.

2. (**Uniqueness**) Assume that there exists  $\varepsilon_\sigma, \varepsilon_{\sigma'} \in \{-1, 1\}$  such that  $\sigma\Phi = \varepsilon_\sigma\Phi$  and  $\sigma\Phi = \varepsilon'_{\sigma'}\Phi$  then  $\varepsilon_\sigma\Phi = \varepsilon'_{\sigma'}\Phi$ . Now if  $(x_1, \dots, x_n) = (1, \dots, n)$  then if  $i < j$  we have  $x_i - x_j = i - j \neq 0$  so that  $\Phi(1, \dots, n) = \prod_{(i, j) \in \{(i, j) \in \{1, \dots, n\} \times \{1, \dots, n\} \mid i < j\}} (x_i - x_j) \neq 0$  (see above lemma) and then from  $\varepsilon_\sigma\Phi(1, \dots, n) = \varepsilon_{\sigma'}\Phi(1, \dots, n)$  we have by multiplying by  $\Phi(1, \dots, n)^{-1}$  that  $\varepsilon_\sigma = \varepsilon_{\sigma'} \quad \square$

**Definition 10.91.** *Given  $n \in \mathbb{N}_0$  and  $\sigma \in P_n$  the unique  $\varepsilon_\sigma \in \{-1, 1\}$  associated with  $\sigma$  (see the previous theorem) is called the **sign** of  $\sigma$  and noted  $\text{sign}(\sigma)$ . This defines a mapping  $\text{sign}: P_n \rightarrow \{-1, 1\}$  by  $\sigma \rightarrow \text{sign}(\sigma)$  where using the previous theorem  $\text{sign}(\sigma) = -1$  if we can write  $\sigma$  as a odd number of strict transpositions and  $\text{sign}(\sigma)$  if we can write  $\sigma$  as a even number of transpositions (this even works if  $n = 1$  and  $\sigma = 1_{\{0\}}$  for then  $\text{sign}(\sigma) = 1 = (-1)^0$ )*

**Lemma 10.92.**  $\langle \{-1, 1\}, \cdot \rangle$  forms a abelian group

1. ( $\because \{-1, 1\} \rightarrow \{-1, 1\}$  is a mapping)
  - a.  $-1 \cdot -1 = 1 \in \{-1, 1\}$
  - b.  $-1 \cdot 1 = -1 \in \{-1, 1\}$
  - c.  $1 \cdot -1 = -1 \in \{-1, 1\}$
  - d.  $1 \cdot 1 = 1 \in \{-1, 1\}$
2. (associativity) This follows from the fact that  $\cdot$  is associative in  $\mathbb{Z}$
3. (neutral element) 1 is the neutral element as  $1 \cdot -1 = -1 \cdot 1 = 1$  and  $1 \cdot 1 = 1$
4. (inverse element) 1 has as inverse element 1 and  $-1$  has as inverse element  $-1$
5. (commutativity)  $1 \cdot -1 = -1 = -1 \cdot 1$

**Theorem 10.93.** If  $n \in \mathbb{N}_0$  then  $\text{sign}: P_n \rightarrow \{-1, 1\}$  is a homeomorphism (see 3.13) from the group  $\langle P_n, \circ \rangle$  to the multiplicative group  $\langle \{-1, 1\}, \cdot \rangle$  or in other words:

1.  $\forall \sigma, \tau \in P_n$  we have  $\text{sign}(\sigma \circ \tau) = \text{sign}(\sigma) \cdot \text{sign}(\tau)$
2.  $\text{sign}(i_{\{1, \dots, n\}}) = 1$

**Proof.**

1.  $(\sigma \circ \tau)\Phi = \sigma(\tau\Phi) = \sigma(\text{sign}(\tau)\Phi) = \text{sign}(\tau) \cdot \sigma\Phi = \text{sign}(\tau) \cdot (\text{sign}(\sigma) \cdot \Phi) = (\text{sign}(\tau) \cdot \text{sign}(\sigma)) \cdot \Phi \Rightarrow \text{sign}(\sigma \circ \tau) = \text{sign}(\sigma) \cdot \text{sign}(\tau)$
2.  $i_{\{1, \dots, n\}}\Phi = \Phi = 1 \cdot \Phi \Rightarrow \text{sign}(i_{\{1, \dots, n\}}) = 1$  □

**Corollary 10.94.** Given  $n \in \mathbb{N}_0$  then if  $\sigma \in P_n \Rightarrow \text{sign}(\sigma) = \text{sign}(\sigma^{-1})$

**Proof.** As  $1 = \text{sign}(i_{\{1, \dots, n\}}) = \text{sign}(\sigma \circ \sigma^{-1}) \stackrel{\text{previous theorem}}{=} \text{sign}(\sigma) = \text{sign}(\sigma^{-1}) \Rightarrow \text{sign}(\sigma) = \text{sign}(\sigma^{-1})$  □

**Definition 10.95.** If  $n \in \mathbb{N}_0$  and  $i, j \in \{1, \dots, n\}$  then we define  $(i \rightsquigarrow_n j)$  as follows

1. ( $i = j$ ) then  $(i \rightsquigarrow_n j) = i_{\{1, \dots, n\}}$
2. ( $i < j$ ) then  $(i \rightsquigarrow_n j)$  is defined by

$$\begin{aligned} (i \rightsquigarrow_n j)(k) &= k \text{ if } k < i \\ &= k+1 \text{ if } i \leq k < j \\ &= i \text{ if } k = j \\ &= k \text{ if } j < k \end{aligned}$$

3. ( $j < i$ ) then  $(i \rightsquigarrow_n j)$  is defined by

$$\begin{aligned} (i \rightsquigarrow_n j)(k) &= k \text{ if } k < j \\ &= i \text{ if } k = j \\ &= k-1 \text{ if } j < k \leq i \\ &= k \text{ if } i < k \end{aligned}$$

**Example 10.96.**

1.  $n = 6$

$$(2 \rightsquigarrow_6 5) \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 4 \\ 5 \\ 2 \\ 6 \end{pmatrix}$$

2.  $n=6$

$$(5 \rightsquigarrow_6 2) \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{pmatrix} = \begin{pmatrix} 1 \\ 5 \\ 2 \\ 3 \\ 4 \\ 6 \end{pmatrix}$$

So we interpret  $(i \rightsquigarrow_n j)$  as removing  $i$  from its position and inserting it after (before)  $j$  if  $i < j$  ( $j < i$ ).

**Example 10.97.** In case of  $n=6$  we have

$$\begin{aligned}
 (2 \leftrightarrow_6 5) \circ (2 \leftrightarrow_6 4) \circ (2 \leftrightarrow_6 3) \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{pmatrix} &= (2 \leftrightarrow_6 5) \circ (2 \leftrightarrow_6 4) \begin{pmatrix} 1 \\ 3 \\ 2 \\ 4 \\ 5 \\ 6 \end{pmatrix} \\
 &= (2 \leftrightarrow_6 5) \begin{pmatrix} 1 \\ 3 \\ 4 \\ 2 \\ 5 \\ 6 \end{pmatrix} \\
 &= \begin{pmatrix} 1 \\ 3 \\ 4 \\ 5 \\ 2 \\ 6 \end{pmatrix} \\
 &= (2 \rightsquigarrow_6 5) \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{pmatrix} \\
 (5 \leftrightarrow_6 2) \circ (5 \leftrightarrow_6 3) \circ (5 \leftrightarrow_6 4) \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{pmatrix} &= (5 \leftrightarrow_6 2) \circ (5 \leftrightarrow_6 3) \begin{pmatrix} 1 \\ 2 \\ 3 \\ 5 \\ 4 \\ 6 \end{pmatrix} \\
 &= (5 \leftrightarrow_6 2) \begin{pmatrix} 1 \\ 2 \\ 5 \\ 3 \\ 4 \\ 6 \end{pmatrix}
 \end{aligned}$$

$$\begin{aligned}
&= \begin{pmatrix} 1 \\ 5 \\ 2 \\ 3 \\ 4 \\ 6 \end{pmatrix} \\
&= (5 \rightsquigarrow_6 2) \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{pmatrix}
\end{aligned}$$

**Theorem 10.98.** If  $n \in \mathbb{N}_0$  then if  $i, j \in \{1, \dots, n\}$  then  $\text{sign}(i \rightsquigarrow_n j) = (-1)^{|i-j|}$ . Further if  $i \neq j$  then  $(i \rightsquigarrow_n j) = ((i \leftrightarrow_n j_1) \circ \dots \circ (i \leftrightarrow_n j_{|i-j|}))$  with  $\forall k \in \{1, \dots, |i-j|\}$  where

$$\begin{aligned}
j_k &= j - (k-1) \text{ if } i < j \\
&= j + (k-1) \text{ if } j < i
\end{aligned}$$

or in other words

$$\begin{aligned}
(i \rightsquigarrow_n j) &= ((i \leftrightarrow_n j) \circ (i \leftrightarrow_n j-1) \circ \dots \circ (i \leftrightarrow_n i+1)) \text{ if } i < j \\
&= ((i \leftrightarrow_n j) \circ (i \leftrightarrow_n j+1) \circ \dots \circ (i \leftrightarrow_n i-1)) \text{ if } j < i
\end{aligned}$$

**Proof.** The proof is very simple but a little bit complex if you do it formally. Given  $i, j \in \{1, \dots, n\}$  we have the following cases:

1. ( $i = j$ ) then  $|i-j|=0$  and  $(i \rightsquigarrow_n j) = i_{\{1, \dots, n\}}$  then  $\text{sign}(i \rightsquigarrow_n j) = 1 = (-1)^0 = (-1)^{|i-j|}$
2. ( $i < j$ ) then we take  $k = j - i$  and prove the theorem by induction on  $k$  so take  $X = \{k \in \{1, \dots\} \mid \text{if } i - j = k \text{ then } \text{sign}(i \rightsquigarrow_n j) = (-1)^k \text{ and } (i \rightsquigarrow_n j) = ((i \leftrightarrow_n j_1) \circ \dots \circ (i \leftrightarrow_n j_k)) \text{ where } j_k = j - (k-1)\}$  then we have :

a. ( $k = 1$ ) then  $j = i + 1$  now if  $l \in \{1, \dots, n\}$  we have :

i. ( $l < i$ ) then we have

$$\begin{aligned}
(i \rightsquigarrow_n i+1)(l) &= l \\
&\stackrel{l \neq i, i+1}{=} (i \leftrightarrow_n i+1)(l)
\end{aligned}$$

ii. ( $i \leq l < i+1$ ) then  $l = i$  and

$$\begin{aligned}
(i \rightsquigarrow_n i+1)(l) &= l+1 \\
&= i+1 \\
&= (i \leftrightarrow_n i+1)(i) \\
&\stackrel{l=i}{=} (i \leftrightarrow_n i+1)(l)
\end{aligned}$$

iii. ( $l = i+1$ ) then

$$\begin{aligned}
(i \rightsquigarrow_n i+1)(l) &= i \\
&\stackrel{l=i+1}{=} (i \leftrightarrow_n i+1)(l)
\end{aligned}$$

iv. ( $i+1 < l$ ) then

$$\begin{aligned}
(i \rightsquigarrow_n i+1)(l) &= l \\
&\stackrel{l \neq i, i+1}{=} (i \leftrightarrow_n i+1)(l)
\end{aligned}$$

So  $(i \rightsquigarrow_n i+1) = (i \leftrightarrow_n i+1)$  and  $\text{sign}(i \rightsquigarrow_n i+1) = (-1)^1$  and thus  $1 \in X$

- b. ( $k \in X$ ) we must now prove that  $k+1 \in X$ . From  $j - i = k + 1$  we have that  $j = i + k + 1$  and  $j - 1 = i + k > i$ . First we prove that  $(i \rightsquigarrow_n j) = (i \leftrightarrow_n j) \circ (i \rightsquigarrow_n j - 1)$  and this is seen for if  $l \in \{1, \dots, n\}$  then we have the following cases:

- i. ( $l < i$ ) then

$$(i \leftrightarrow_n j) \circ (i \rightsquigarrow_n j - 1)(l) = \begin{cases} (i \leftrightarrow_n j)(l) & l \neq i, j \\ (i \rightsquigarrow_n j)(l) & l < i \end{cases}$$

- ii. ( $i \leq l < j - 1$ ) then

$$(i \leftrightarrow_n j) \circ (i \rightsquigarrow_n j - 1)(l) = \begin{cases} (i \leftrightarrow_n j)(l+1) & i < i+1 \leq l+1 < j \\ (i \rightsquigarrow_n j)(l) & i \leq l+1 < j \end{cases}$$

- iii. ( $l = j - 1$ ) then

$$(i \leftrightarrow_n j) \circ (i \rightsquigarrow_n j - 1)(l) = \begin{cases} (i \leftrightarrow_n j)(i) & i < j \Rightarrow i \leq j-1 = l < j \\ (i \rightsquigarrow_n j)(j-1) & l = j-1 \end{cases}$$

- iv. ( $j - 1 < l$ ) then we have the following cases

- A. ( $j = l$ ) then

$$(i \leftrightarrow_n j) \circ (i \rightsquigarrow_n j - 1)(l) = \begin{cases} (i \leftrightarrow_n j)(l) & j-1 \leq l = j \\ (i \leftrightarrow_n j)(j) & l = j \\ i & l = j-1 \\ (i \rightsquigarrow_n j)(l) & l = j \end{cases}$$

- B. ( $j < l$ ) then

$$(i \leftrightarrow_n j) \circ (i \rightsquigarrow_n j - 1)(l) = \begin{cases} (i \leftrightarrow_n j)(l) & j-1 \leq j < l \\ l & i < j < l \Rightarrow i, j \neq l \\ (i \rightsquigarrow_n j)(l) & j < l \end{cases}$$

proving indeed that  $(i \leftrightarrow_n j) \circ (i \rightsquigarrow_n j - 1) = (i \rightsquigarrow_n j)$ .

Now we have

$$\begin{aligned} (i \rightsquigarrow_n j) &= (i \leftrightarrow_n j) \circ (i \rightsquigarrow_n j - 1) \\ &\stackrel{(j-1)-i=i+(k+1)-1-i=k \wedge k \in X}{=} (i \leftrightarrow_n j) \circ [(i \leftrightarrow_n j - 1) \circ (i \leftrightarrow_n (j-1) - 1) \circ \dots \circ (i \leftrightarrow_n i + 1)] \\ &= ((i \leftrightarrow_n j) \circ (i \rightsquigarrow_n j - 1) \circ \dots \circ (i \leftrightarrow_n i + 1)) \end{aligned}$$

Also  $\text{sign}(i \rightsquigarrow_n j) = \text{sign}(i \leftrightarrow_n j) \cdot \text{sign}(i \rightsquigarrow_n j - 1) = (-1) \cdot (-1)^{j-1-i} = (-1)^{j-i}$ , proving that  $k+1 \in X$  and thus  $X = \{1, \dots\} = \mathbb{N}_0$  proving our assertion.

3. ( $j < i$ ) then we take  $k = i - j$  and prove the theorem by induction on  $k$  so take  $X = \{k \in \{1, \dots\} \mid \text{if } i - j = k \text{ then } \text{sign}(i \rightsquigarrow_n j) = (-1)^k\}$  and  $(i \rightsquigarrow_n j) = ((i \leftrightarrow_n j_1) \circ \dots \circ (i \leftrightarrow_n j_k))$  where  $j_k = j + (k-1)$  then we have :

- a. ( $k = 1$ ) so  $j = i - 1$  then we have the following cases for  $l \in \{1, \dots, n\}$

- i. ( $l < j$ ) then

$$(i \rightsquigarrow_n i - 1)(l) = \begin{cases} l & l < j = i-1 \\ (i \leftrightarrow_n i - 1)(l) & l < j \leq i-1 < i \end{cases}$$

ii.  $(l = j)$  then  $j = i - 1$

$$(i \rightsquigarrow_n i - 1)(l) = i$$

$$\underset{l=j=i-1}{\equiv} (i \leftrightarrow_n i - 1)(l)$$

iii.  $(j < l \leq i)$  then  $i - 1 = j < l \leq i \Rightarrow l = i$  and thus

$$(i \rightsquigarrow_n i - 1)(l) \underset{i-1 \leq l \leq i}{\equiv} l - 1$$

$$\underset{l=i}{\equiv} i - 1$$

$$= (i \leftrightarrow_n i - 1)(i)$$

$$\underset{i=l}{\equiv} (i \leftrightarrow_n i - 1)(l)$$

iv.  $(i < l)$  then

$$(i \rightsquigarrow_n i - 1)(l) \underset{i, i-1 < l}{\equiv} l$$

$$\underset{i, i-1 \neq l}{\equiv} (i \leftrightarrow_n i - 1)(l)$$

So  $(i \rightsquigarrow_n i - 1) = (i \leftrightarrow_n i - 1)$  and  $\text{sign}(i \rightsquigarrow_n i - 1) = \text{sign}(i \leftrightarrow_n i - 1) = (-1) = (-1)^1 = (-1)^{|i-j|}$  proving that  $1 \in X$

b.  $(k \in X)$  we have then to prove that  $k + 1 \in X$ . From  $i - j = k + 1$  it follows that  $j = i - k - 1$  and  $j + 1 < i - k < i$ . First we prove that  $(i \rightsquigarrow_n j) \circ (i \rightsquigarrow_n j + 1) = (i \rightsquigarrow_n j)$ . So let  $l \in \{1, \dots, n\}$  then we have the following cases

i.  $(l < j)$  then

$$(i \leftrightarrow_n j) \circ (i \rightsquigarrow_n j + 1)(l) \underset{l < j < j+1}{\equiv} (i \leftrightarrow_n j)(l)$$

$$\underset{l < j < i}{\equiv} l$$

$$\underset{l < j}{\equiv} (i \rightsquigarrow_n j)(l)$$

ii.  $(l = j)$  then

$$(i \leftrightarrow_n j) \circ (i \rightsquigarrow_n j + 1)(l) \underset{l=j < j+1}{\equiv} (i \leftrightarrow_n j)(l)$$

$$\underset{l=j}{\equiv} i$$

$$\underset{l=j}{\equiv} (i \rightsquigarrow_n j)(l)$$

iii.  $(j < l \leq i)$  Here we have the following sub-cases :

A.  $(l = j + 1)$  then

$$(i \leftrightarrow_n j) \circ (i \rightsquigarrow_n j + 1)(l) \underset{l=j+1}{\equiv} (i \leftrightarrow_n j)(i)$$

$$= j$$

$$\underset{l=j+1}{\equiv} l - 1$$

$$\underset{j < l \leq i}{\equiv} (i \rightsquigarrow_n j)(l)$$

B.  $(j+1 < l \leq i)$  then

$$(i \leftrightarrow_n j) \circ (i \rightsquigarrow_n j + 1)(l) = (i \leftrightarrow_n j)(l - 1)$$

$$\underset{j < j+1 \leq l-1 < l \leq i}{\equiv} l - 1$$

$$\underset{j < l \leq i}{\equiv} (i \rightsquigarrow_n j)(l)$$

iv.  $(i < l)$  then

$$(i \leftrightarrow_n j) \circ (i \rightsquigarrow_n j + 1)(l) = (i \leftrightarrow_n j)(l)$$

$$\underset{j < i < l}{\equiv} l$$

$$\underset{i < l}{\equiv} (i \rightsquigarrow_n j)(l)$$

So we have that

$$\begin{aligned}
 (i \rightsquigarrow_n j) &= (i \leftrightarrow_n j) \circ (i \rightsquigarrow_n j + 1) \\
 \text{induction and } i - (j+1) = k &\quad (i \leftrightarrow_n j) \circ [(i \leftrightarrow_n j + 1) \circ (i \leftrightarrow_n j + 2) \circ \dots \circ (i \leftrightarrow_n i - 1)] \\
 &= ((i \leftrightarrow_n j) \circ (i \leftrightarrow_n j + 1) \circ \dots \circ (i \leftrightarrow_n i - 1))
 \end{aligned}$$

also  $\text{sign}(i \rightsquigarrow_n j) = \text{sign}(i \leftrightarrow_n j) \cdot \text{sign}(i \rightsquigarrow_n j) \underset{k \in X}{=} (-1) \cdot (-1)^k = (-1)^{k+1}$  proving that  $k+1 \in X$  and by mathematical induction that  $X = \{1, \dots\} = \mathbb{N}_0$   $\square$

**Theorem 10.99.** *Given  $\tau \in P_n$  define then  $T_\tau: P_n \rightarrow P_n$  by  $\sigma \mapsto T_\tau(\sigma) = \tau \circ \sigma$  is a bijection*

**Proof.**

1. **(surjectivity)** If  $\sigma \in P_n$  take then  $\sigma' = \tau^{-1} \circ \sigma \in P_n$  then  $T_\tau(\sigma') = \tau \circ (\tau^{-1} \circ \sigma) = (\tau \circ \tau^{-1}) \circ \sigma = 1_{\{1, \dots, n\}} \circ \sigma = \sigma$  so we have found a  $\sigma'$  such that  $T_\tau(\sigma') = \sigma$
2. **(injectivity)** If  $T_\tau(\sigma) = T_\tau(\sigma') \Rightarrow \tau \circ \sigma = \tau \circ \sigma' \Rightarrow \tau^{-1} \circ (\tau \circ \sigma) = \tau^{-1} \circ (\tau \circ \sigma') = (\tau^{-1} \circ \tau) \circ \sigma = (\tau^{-1} \circ \tau) \circ \sigma' = 1_{\{1, \dots, n\}} \circ \sigma = 1_{\{1, \dots, n\}} \circ \sigma' = \sigma = \sigma'$   $\square$

### 10.3 Finite product of sets

**Definition 10.100.** *Let  $n \in \mathbb{N}_0$ ,  $A$  a set then we define  $A^n$  as  $A^n = \prod_{i \in \{1, \dots, n\}} A_i$  where  $\{A_i\}_{i \in \{1, \dots, n\}}$  where  $\forall i \in \{1, \dots, n\}$  we have  $A_i = A$ .*

**Theorem 10.101.** *Let  $n \in \mathbb{N}_0$  and  $A$  a set then we have that  $A^n = A^{\{1, \dots, n\}}$*

**Proof.** See 2.86  $\square$

We can of course define the power of a set by cartesian products

**Definition 10.102.** *Let  $n \in \mathbb{N}_0$ ,  $A$  a set then  $A^{(\times)n}$  is defined recursively by:*

- $n = 1$ .** then  $A^{(\times)1} = A$
- $n > 1$ .** then  $A^{(\times)n} = (A^{(\times)(n-1)}) \times A$

**Theorem 10.103.** *Let  $n \in \mathbb{N}_0$ ,  $A$  a set then  $A^{(\times)n} = \bigotimes_{i \in \{1, \dots, n\}} A_i$  where  $\{A_i\}_{i \in \{1, \dots, n\}}$  is such that  $A_i = A$*

**Proof.** This is trivially proved by recursion so let  $S_A = \{n \in \mathbb{N}_0 | A^{(\times)n} = \bigotimes_{i \in \{1, \dots, n\}} A_i \text{ where } \{A_i\}_{i \in \{1, \dots, n\}} \text{ is such that } A_i = A\}$  then we have

$1 \in S$ .  $A \stackrel{\text{definition}}{=} A^{(\times)1} \stackrel{\text{definition}}{=} \bigotimes_{i \in \{1, \dots, 1\}} A_i$  where  $\{A_i\}_{i \in \{1, \dots, 1\}}$  is such that  $A_1 = A$  proving that  $1 \in S$ .

$n \in S \Rightarrow n+1 \in S$ . Let  $\{A_i\}_{i \in \{1, \dots, n+1\}}$  such that  $A_i = A$  then we have  $A^{(\times)(n+1)} = (A^{(\times)n}) \times A \underset{n \in S}{=} (\bigotimes_{i \in \{1, \dots, n\}} A_i) \times A = (\bigotimes_{i \in \{1, \dots, n\}} A_i) \times A_{n+1} = \bigotimes_{i \in \{1, \dots, n+1\}} A_i$  proving  $n+1 \in S$   $\square$

Let's now compare the two definitions of the finite power of a set.

**Theorem 10.104.** *Let  $n \in \mathbb{N}_0$  and  $A$  as set then there exists a bijection  $\mathcal{P}: A^{(\times)n} \rightarrow A^n$  so that*

1. If  $B \subseteq A$  then  $\mathcal{P}(B^{(\times)n}) = B^n$
2.  $\forall [x_1, \dots, x_n] \in B^{(\times)n}$  we have  $\mathcal{P}([x_1, \dots, x_n]) = (x_1, \dots, x_n)$

**Proof.** This follows from the previous theorem and 5.89  $\square$

**Theorem 10.105.** Let  $n \in \mathbb{N}_0$  and  $\{A_i\}_{i \in \{1, \dots, n\}}$  be such that  $\forall i \in \{1, \dots, n\}$  we have  $A_i$  is finite then  $\bigotimes_{i \in \{1, \dots, n\}} A_i$  and  $\prod_{i \in \{1, \dots, n\}} A_i$  are finite. Furthermore we have that  $\#(\bigotimes_{i \in \{1, \dots, n\}} A_i) = \#(\prod_{i \in \{1, \dots, n\}} A_i) = \prod_{i=1}^n \#(A_i)$

**Proof.** We prove this by mathematical induction so let  $S = \{n \in \mathbb{N}_0 \mid \{A_i\}_{i \in \{1, \dots, n\}} \text{ a finite family of finite sets then } \bigotimes_{i \in \{1, \dots, n\}} A_i \text{ is finite with } \#(\bigotimes_{i \in \{1, \dots, n\}} A_i) = \prod_{i \in \{1, \dots, n\}} \#(A_i)\}$  then we have

$1 \in S$ . If  $\{A_i\}_{i \in \{1, \dots, n\}}$  is a finite family of finite sets then  $\bigotimes_{i \in \{1, \dots, 1\}} A_i = A_i$  which is finite and  $\#(\bigotimes_{i \in \{1, \dots, 1\}} A_i) = \#(A_1) = \prod_{i=1}^1 \#(A_i)$  so that  $1 \in S$

$n \in S \Rightarrow n+1 \in S$ . Let  $\{A_i\}_{i \in \{1, \dots, n+1\}}$  then  $\bigotimes_{i \in \{1, \dots, n+1\}} A_i = (\bigotimes_{i \in \{1, \dots, n\}} A_i) \times A_{n+1}$  which, as  $\bigotimes_{i \in \{1, \dots, n\}} A_i$  is finite ( $n \in S$ ) and  $A_{n+1}$  is finite, is finite by 5.44.

Also  $\#(\bigotimes_{i \in \{1, \dots, n+1\}} A_i) = \#((\bigotimes_{i \in \{1, \dots, n\}} A_i) \times A_{n+1}) \stackrel{5.44}{=} \#(\bigotimes_{i \in \{1, \dots, n\}} A_i) \cdot \#(A_{n+1}) \stackrel{n \in S}{=} (\prod_{i=1}^n \#(A_i)) \cdot \#(A_{n+1}) = \prod_{i=1}^{n+1} \#(A_i)$  proving that  $n+1 \in S$

As  $\bigotimes_{i \in \{1, \dots, n\}} A_i$  and  $\prod_{i \in \{1, \dots, n\}} A_i$  are bijective we have proved the rest of our theorem.  $\square$

A easy consequence of the above theorem is the following theorem.

**Theorem 10.106.** If  $n \in \mathbb{N}_0$  and  $X$  is a finite set then  $X^n$  and  $X^{(\times)n}$  is finite and  $\#(X^{(\times)n}) = \#(X^n) = \#(X)^n$

**Proof.** See the previous theorem.  $\square$

## 10.4 Vector Spaces

### 10.4.1 Definition and examples

**Definition 10.107.** 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  $\odot: F \times V \rightarrow V$  (here we note  $\odot(x, y)$  as  $x \odot y$  and call this scalar multiplication of a scalar and a vector) such that

1.  $\forall \alpha \in F, \forall x, y \in V$  we have  $\alpha \odot (x \oplus y) = (\alpha \odot x) \oplus (\alpha \odot y)$
2.  $\forall \alpha, \beta \in F, \forall x \in V$  we have  $(\alpha + \beta) \odot x = (\alpha \odot x) \oplus (\beta \odot x)$
3.  $\forall \alpha, \beta \in F, \forall x \in V$  we have  $(\alpha \cdot \beta) \odot x = \alpha \odot (\beta \odot x)$
4.  $\forall x \in V$  we have  $1 \odot x = x$  (1 is the neutral element in  $\langle F, \cdot \rangle$ )

Elements of  $V$  are called **vectors**.

**Theorem 10.108.** Let  $\langle V, \oplus, \odot \rangle$  be a vector space over a field  $\langle F, +, \cdot \rangle$  then we have

1.  $\forall x \in V$  we have  $0 = 0 \odot x$  ( $0$  is the neutral element in  $\langle V, \oplus \rangle$  and  $0$  is the neutral element in  $\langle F, + \rangle$ )
2.  $\forall x \in V$  we have  $(-1) \odot x = -x$  where  $-1$  is the inverse element of  $1$  in  $\langle F, + \rangle$  and  $-x$  is the inverse element of  $x$  in  $\langle V, \oplus \rangle$
3.  $\forall \alpha \in F$  we have  $\alpha \odot 0 = 0$

**Proof.**

1.  $0 = 0 \odot x \oplus (-0 \odot x) \stackrel{0=0+0}{=} (0+0) \odot x \oplus (-0 \odot x) = (0 \odot x + 0 \odot x) \oplus (-0 \odot x) = 0 \odot x \oplus (0 \odot x \oplus (-0 \odot x)) = 0 \odot x \oplus 0 = 0 \odot x$

2.  $x \oplus ((-1) \odot x) = (1 \odot x) \oplus ((-1) \odot x) = (1 + (-1)) \odot x = 0 \odot x \stackrel{(1)}{=} O \Rightarrow (-1) \odot x = -x$
3.  $\alpha \odot O \stackrel{(1)}{=} \alpha \odot (0 \odot O) \stackrel{3.25}{=} (\alpha \cdot 0) \odot O = 0 \odot O = O$

□

**Theorem 10.109.** Let  $\langle V, \oplus, \odot \rangle$  be a vector space over a field  $\langle F, +, \cdot \rangle$  then if  $\alpha \in F \setminus \{0\}$  ( $0$  is the neutral element in  $\langle F, + \rangle$ ) we have  $\forall x \in V \alpha \odot x = O \Rightarrow x = O$  ( $O$  the neutral element in  $\langle V, \oplus \rangle$ )

**Proof.** As  $\alpha \neq 0$  we have that  $\alpha^{-1}$  exists.  $O = \alpha^{-1} \odot O \stackrel{\alpha \odot x = O}{=} \alpha^{-1} \odot (\alpha \odot x) = (\alpha^{-1} \cdot \alpha) \odot x = 1 \odot x = x$  □

**Definition 10.110.** If  $\langle V, \oplus, \odot \rangle$  is a vector space over a field  $\langle F, +, \cdot \rangle$  and  $\emptyset \neq W \subseteq V$  is a non-empty set such that  $\forall \alpha, \beta \in F$  and  $\forall x, y \in W$  we have that  $\alpha \odot x \oplus \beta \odot y \in W$  then  $W$  is called a subspace of  $V$

**Theorem 10.111.** If  $\langle V, \oplus, \odot \rangle$  is a vector space over a field  $\langle F, +, \cdot \rangle$  and  $\emptyset \neq W \subseteq V$  is a subspace of  $V$  then  $\langle W, \oplus|_{W \times W}, \odot|_{F \times W} \rangle$  is a vector space over a field  $\langle F, +, \cdot \rangle$ . Note that in general we write the subspace as  $\langle W, \oplus, \odot \rangle$  and we assume it is now that we use here the restrictions to  $W$  of the vector operations.

**Proof.** First we must prove that  $\langle W, \oplus|_{W \times W} \rangle$  is a abelian group

1.  $\oplus|_{W \times W}: W \times W \rightarrow W$  is a function, for this to be true we must have  $\forall x, y \in W$  that  $x \oplus|_{W \times W} y = x \oplus y \in W$ . Now  $x \oplus y = (1 \odot x) \oplus (1 \odot y) \in W$  as  $W$  is a subspace.
2. **(associativity)** If  $x, y, z \in W$  then  $x \oplus|_{W \times W} (y \oplus|_{W \times W} z) = x \oplus (y \oplus z) = (x \oplus y) \oplus z = (x \odot|_{W \times W} y) \oplus|_{W \times W} z$
3. **(neutral element)** If  $O$  is the neutral element of  $\langle V, \oplus \rangle$  then as  $\emptyset \neq W$  there exists a  $w \in W$  but then  $O = O \oplus O \stackrel{10.108}{=} 0 \odot x \oplus 0 \odot x \in W$  as  $W$  is a subspace. Now if  $x \in W$  then  $x \oplus|_{W \times W} O = x \oplus O = x = O \oplus x = O \oplus|_{W \times W} x$
4. **(inverse element)** If  $x \in W$  then  $-x \stackrel{10.108}{=} (-1) \odot x \oplus O \in W$  (see (3) and  $W$  is a subspace) and  $x \oplus|_{W \times W} (-x) = x \oplus (-x) = O = (-x) \oplus x = (-x) \oplus|_{W \times W} x$
5. **(commutativity)** If  $x, y \in W$  then  $x \oplus|_{W \times W} y = x \oplus y = y \oplus x = y \oplus|_{W \times W} x$

Next we prove the remaining axioms of a vector space

1.  $\odot|_{F \times W}: F \times W \rightarrow W$  is a function, for this to be true we must have that  $\forall \alpha \in F, \forall x \in W$  that  $\alpha \odot|_{F \times W} x = \alpha \odot x \in W$ . But  $\alpha \odot x = \alpha \odot x + O = \alpha \odot x + 0 \odot x \in W$
2. If  $\alpha \in F$  and  $x, y \in W$  then  $\alpha \odot|_{W \times W} (x \oplus|_{W \times W} y) = \alpha \odot (x \oplus y) = \alpha \odot x \oplus \alpha \odot y = \alpha \odot|_{W \times W} x \oplus|_{W \times W} \beta \odot|_{W \times W} y$
3. If  $\alpha, \beta \in F$  and  $x \in W$  then  $(\alpha + \beta) \odot|_{W \times W} x = (\alpha + \beta) \odot x = \alpha \odot x \oplus \beta \odot x = \alpha \odot|_{W \times W} x \oplus|_{W \times W} \beta \odot|_{W \times W} x$
4. If  $\alpha, \beta \in F$  and  $x \in W$  then  $(\alpha \cdot \beta) \odot|_{W \times W} x = (\alpha \cdot \beta) \odot x = \alpha \odot (\beta \odot x) = \alpha \odot|_{W \times W} (\beta \odot|_{W \times W} x)$
5. If  $x \in W$  then  $1 \odot|_{F \times W} x = 1 \odot x = x$  □

Let's now look at some examples of vector spaces.

First every field is a vector space over itself.

**Theorem 10.112.** If  $\langle F, +, \cdot \rangle$  is a field then  $\langle F, +, \cdot \rangle$  is a vector space over the field  $\langle F, +, \cdot \rangle$

**Proof.** As  $F$  is a field we have that (see 3.29)  $\langle F, + \rangle$  is a abelian group and further from the axioms of a field we have :

1.  $\forall \alpha \in F, \forall x, y \in F$  we have  $\alpha \cdot (x + y) = \alpha \cdot x + \alpha \cdot y$
2.  $\forall \alpha, \beta \in F, \forall x \in F$  we have  $(\alpha + \beta) \cdot x = \alpha \cdot x + \beta \cdot x$

3.  $\forall \alpha, \beta \in F, \forall x \in F$  we have  $(\alpha \cdot \beta) \cdot x = \alpha \cdot (\beta \cdot x)$
4.  $\forall x \in F$  we have  $1 \cdot x = x$

□

As  $\langle \mathbb{R}, +, \cdot \rangle$  is a field (see 8.25) we have thus

**Corollary 10.113.**  $\langle \mathbb{R}, +, \cdot \rangle$  is a vector space over itself.  $\langle \mathbb{C}, +, \cdot \rangle$  is a vector space over itself

We can also make  $\langle \mathbb{C}, +, \cdot \rangle$  a vector space over  $\langle \mathbb{R}, +, \cdot \rangle$

**Theorem 10.114.**  $\langle \mathbb{C}, +, \cdot \rangle$  is a vector space over  $\langle \mathbb{R}, +, \cdot \rangle$  where the scalar multiplication  $\cdot: \mathbb{R} \times \mathbb{C} \rightarrow \mathbb{C}$  is defined to be  $(\alpha, (x, y)) \rightarrow (\alpha \cdot x, \alpha \cdot y)$

**Proof.** As  $\langle \mathbb{C}, +, \cdot \rangle$  is a field we have that  $\langle \mathbb{C}, + \rangle$  is a abelian group. We have then

1.  $\forall \alpha \in \mathbb{R}, \forall (x, y), (x', y') \in \mathbb{C}$  we have  $\alpha \cdot ((x, y) + (x', y')) = \alpha \cdot (x + x', y + y') = (\alpha \cdot (x + x'), \alpha \cdot (y + y')) = (\alpha \cdot x + \alpha \cdot x', \alpha \cdot y + \alpha \cdot y') = (\alpha \cdot x, \alpha \cdot y) + (\alpha \cdot x', \alpha \cdot y') = \alpha \cdot (x, y) + \alpha \cdot (x', y')$
2.  $\forall \alpha, \beta \in \mathbb{R}, \forall (x, y) \in \mathbb{C}$  we have  $(\alpha + \beta) \cdot (x, y) = ((\alpha + \beta) \cdot x, (\alpha + \beta) \cdot y) = (\alpha \cdot x + \beta \cdot x, \alpha \cdot y + \beta \cdot y) = (\alpha \cdot x, \alpha \cdot y) + (\beta \cdot x, \beta \cdot y) = \alpha \cdot (x, y) + \beta \cdot (x, y)$
3.  $\forall \alpha, \beta \in \mathbb{R}, \forall (x, y) \in \mathbb{C}$  we have  $(\alpha \cdot \beta) \cdot (x, y) = ((\alpha \cdot \beta) \cdot x, (\alpha \cdot \beta) \cdot y) = (\alpha \cdot (\beta \cdot x), \alpha \cdot (\beta \cdot y)) = \alpha \cdot (\beta \cdot x, \beta \cdot y) = \alpha \cdot (\beta \cdot (x, y))$
4.  $\forall (x, y) \in \mathbb{C}$  we have  $1 \cdot (x, y) = (1 \cdot x, 1 \cdot y) = (x, y)$

□

Let's see now how we can create a vector space based on a existing vector space.

**Theorem 10.115.** Let  $\langle V, \oplus, \odot \rangle$  be a vector space over a field  $\langle F, +, \cdot \rangle$ ,  $X$  a set then  $V^X = \{f | f: X \rightarrow V \text{ is a function}\}$  (the set of functions from  $X$  to  $V$  (see 2.73)). Define then  $\boxplus: V^X \times V^X \rightarrow V^X$  by  $(f, g) \rightarrow f + g$  where  $(f + g)(x) = f(x) + g(x)$  and  $\boxdot: F \times V^X \rightarrow V^X$  by  $(\alpha, f) \rightarrow \alpha \boxdot f$  where  $(\alpha \boxdot f)(x) = \alpha \odot f(x)$ . Then  $\langle V^X, \boxplus, \boxdot \rangle$  forms a vector space over  $\langle F, +, \cdot \rangle$  we note this vector space as  $M(X, V)$ . The neutral element in  $M(X, V)$  is  $C_{X,0}$  (see 2.15)

**Proof.** First we prove that  $\langle V^X, \boxplus \rangle$  forms a abelian group with units  $C_{X,0}: X \rightarrow V$  where  $C_{X,0}(x) = 0$  for every  $x$  in  $X$ . So let  $f, g, h \in V^X$  then we have:

1. **(associativity)**  $\forall x \in X$  we have  $f \boxplus (g \boxplus h)(x) = f(x) \oplus (g \boxplus h)(x) = f(x) \oplus (g(x) \oplus h(x)) \stackrel{\langle V, \oplus, \odot \rangle \text{ is a vectorspace}}{=} (f(x) \oplus g(x)) \oplus h(x) = (f \boxplus g)(x) \oplus h(x) = ((f \boxplus g) \boxplus)(x) \Rightarrow f \boxplus (g \boxplus h) = (f \boxplus g) \boxplus h$
2. **(commutativity)**  $\forall x \in X$  we have  $(f \boxplus g)(x) = f(x) \oplus g(x) \stackrel{\langle V, \oplus, \odot \rangle \text{ is a vectorspace}}{=} g(x) \oplus f(x) = (g \boxplus f)(x) \Rightarrow f \boxplus g = g \boxplus f$
3. **(neutral element)**  $\forall x \in X$  we have  $(f \boxplus 0)(x) = f(x) \oplus C_{X,0}(x) = f(x) \oplus 0 = f(x) \Rightarrow f \boxplus C_{X,0} = f$  and  $C_{X,0} \boxplus f = f$  follows then from commutativity.
4. **(inverse element)** Define  $-f: X \rightarrow V$  by  $x \rightarrow -f(x)$  then we have  $\forall x \in X$  that  $(f \boxplus (-f))(x) = f(x) \oplus (-f(x)) \stackrel{\langle V, \oplus, \odot \rangle \text{ is a vectorspace}}{=} 0 \Rightarrow f \boxplus (-f) = C_{X,0}$  and  $(-f) \boxplus f = C_{X,0}$  follows from commutativity.

For the remaining vector space axioms note that:

1. If  $\alpha \in F$  and  $f, g \in V^X$  then  $\forall x \in X$  we have  $(\alpha \boxdot (f \boxplus g))(x) = \alpha \odot (f \boxplus g)(x) = \alpha \odot (f(x) \oplus g(x)) \stackrel{\langle V, \oplus, \odot \rangle \text{ is a vectorspace}}{=} \alpha \odot f(x) \oplus \alpha \odot g(x) = (\alpha \boxdot f)(x) \oplus (\alpha \boxdot g)(x) = ((\alpha \boxdot f) \boxplus (\alpha \boxdot g))(x) \Rightarrow \alpha \boxdot (f \boxplus g) = \alpha \boxdot f \boxplus \alpha \boxdot g$
2. If  $\alpha, \beta \in F$  and  $g \in V^X$  then  $\forall x \in X$  we have  $((\alpha + \beta) \boxdot f)(x) = (\alpha + \beta) \odot f(x) \stackrel{\langle V, \oplus, \odot \rangle \text{ is a vectorspace}}{=} \alpha \odot f(x) + \beta \odot f(x) = (\alpha \boxdot f)(x) \oplus (\beta \boxdot f)(x) = ((\alpha \boxdot f) \boxplus (\beta \boxdot f))(x) \Rightarrow (\alpha + \beta) \boxdot f = (\alpha \boxdot f) \boxplus (\beta \boxdot f)$

3. If  $\alpha, \beta \in F$  and  $f \in V^X$  then  $\forall x \in X$  we have  $((\alpha \cdot \beta) \boxdot f)(x) = (\alpha \cdot \beta) \odot f(x) \underset{\langle V, \oplus, \odot \rangle \text{ is a vectorspace}}{=} \alpha \cdot (\beta \cdot f(x)) = \alpha \cdot ((\beta \boxdot f)(x)) = (\alpha \boxdot (\beta \boxdot f))(x) \Rightarrow (\alpha \cdot \beta) \boxdot f = \alpha \boxdot (\beta \boxdot f)$
4. If  $1 \in F$  and  $f \in V^x$  then  $\forall x \in V$  we have  $(1 \boxdot f)(x) = 1 \odot f(x) \underset{\langle V, \oplus, \odot \rangle \text{ is a vectorspace}}{=} f(x) \Rightarrow 1 \boxdot f = f$   $\square$

**Note 10.116.** To avoid excessive use of different use of different operator notations we use from now on only the symbols  $+, \cdot$  instead of  $\oplus, \boxdot, \odot$  and  $\boxdot$  and assume that care full inspection of the context where a symbol is used is enough to find to which space (field ...) a operator belongs. This is similar to our use of 0, 1.

If  $\langle V, +, \cdot \rangle$  is a vector space over  $\langle F, +, \cdot \rangle$  and if  $n \in \mathbb{N}_0$  then by 10.101 we have that  $V^n = V^{\{1, \dots, n\}}$  and by applying the above theorem we have automatically that the following corollary is true.

**Corollary 10.117.** If  $\langle V, +, \cdot \rangle$  is a vector space over  $\langle F, +, \cdot \rangle$  and  $n \in \mathbb{N}_0$  then  $\langle V^n, +, \cdot \rangle$  is a vector space over  $\langle F, +, \cdot \rangle$  where  $+: V^n \times V^n \rightarrow V^n$  is defined by  $(x, y) \rightarrow x + y$  where  $\forall i \in \{1, \dots, n\}$  we have  $(x + y)_i = (x + y)(i) = x(i) + y(i) = x_i + y_i$  and  $\cdot: F \times V^n \rightarrow V^n$  is defined by  $(\alpha, x) \rightarrow \alpha \cdot x$  where  $\forall i \in \{1, \dots, n\}$  we have  $(\alpha \cdot x)_i = (\alpha \cdot x)(i) = \alpha \cdot x(i) = \alpha \cdot x_i$

This with the fact that a field forms a vector space over itself gives us the following corollaries (where we have assumed the above meaning of  $+, \cdot$  in  $F^n$ )

**Corollary 10.118.** If  $\langle F, +, \cdot \rangle$  is a field and  $n \in \mathbb{N}_0$  then  $\langle F^n, +, \cdot \rangle$  forms a vector space over the field  $\langle F, +, \cdot \rangle$

**Corollary 10.119.** If  $n \in \mathbb{N}_0$  then

1.  $\langle \mathbb{R}^n, +, \cdot \rangle$  forms a vector space over  $\langle \mathbb{R}, +, \cdot \rangle$
2.  $\langle \mathbb{C}^n, +, \cdot \rangle$  forms a vector space over  $\langle \mathbb{C}, +, \cdot \rangle$
3.  $\langle \mathbb{C}^n, +, \cdot \rangle$  forms a vector space over  $\langle \mathbb{R}, +, \cdot \rangle$

**Theorem 10.120.** 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.  $+: \prod_{i \in I} V_i \times \prod_{i \in I} V_i \rightarrow \prod_{i \in I} V_i$  is defined by  $(x, y) \rightarrow x + y$  where  $x + y: I \rightarrow \bigcup_{i \in I} V_i$  is defined by  $(x + y)(i) = x(i) +_i y(i) = x_i +_i y_i$  (see 3.15)
2.  $\cdot: F \times \prod_{i \in I} V_i \rightarrow \prod_{i \in I} V_i$  is defined by  $(\alpha, x) \rightarrow \alpha \cdot x$  where  $\alpha \cdot x: I \rightarrow \bigcup_{i \in I} V_i$  is defined by  $(\alpha \cdot x)(i) = \alpha \cdot_i x(i) = \alpha \cdot_i x_i$

then we have that  $\langle \prod_{i \in I} V_i, +, \cdot \rangle$  is a vector space over  $\langle F, +, \cdot \rangle$

**Proof.** From 3.15 it follows that  $\langle \prod_{i \in I} V_i, + \rangle$  is a abelian group. Next if  $\alpha \in F$  and  $x \in V_i$  we have by the fact that  $\langle V_i, +_i, \cdot_i \rangle$  is a group that  $\alpha \cdot_i x(i) \in V_i$  that  $\alpha \cdot x: I \rightarrow \bigcup_{i \in I} V_i$  is a element of  $\prod_{i \in I} V_i$  and thus that  $\cdot: F \times \prod_{i \in I} V_i \rightarrow \prod_{i \in I} V_i$  is indeed a function. Now that we have proved that (1) and (2) are well defined we prove the rest of the vector space axioms.

1. If  $\alpha \in F$  and  $x, y \in \prod_{i \in I} V_i$  then  $\forall i \in I$  we have  $(\alpha \cdot (x + y))(i) = \alpha \cdot_i (x + y)(i) = \alpha \cdot_i (x(i) +_i y(i)) \underset{\langle V_i, +_i, \cdot_i \rangle \text{ is a vectorspace}}{=} \alpha \cdot_i x(i) +_i \alpha \cdot_i y(i) = (\alpha \cdot x)(i) +_i (\alpha \cdot y)(i) = (\alpha \cdot x + \alpha \cdot y)(i)$  and thus  $\alpha \cdot (x + y) = \alpha \cdot x + \alpha \cdot y$
2. If  $\alpha, \beta \in F$  and  $x \in \prod_{i \in I} V_i$  then  $\forall i \in I$  we have  $((\alpha + \beta) \cdot x)(i) = (\alpha + \beta) \cdot_i x(i) \underset{\langle V_i, +_i, \cdot_i \rangle}{=} \alpha \cdot_i x(i) +_i \beta \cdot_i x(i) = (\alpha \cdot x)(i) +_i (\beta \cdot x)(i) = (\alpha \cdot x + \beta \cdot x)(i)$  and thus  $(\alpha + \beta) \cdot x = \alpha \cdot x + \beta \cdot x$
3. If  $\alpha, \beta \in F$  and  $x \in \prod_{i \in I} V_i$  then  $\forall i \in I$  we have  $((\alpha \cdot \beta) \cdot x)(i) = (\alpha \cdot \beta) \cdot_i x(i) \underset{\langle V_i, +_i, \cdot_i \rangle \text{ is a vectorspace}}{=} \alpha \cdot_i (\beta \cdot x(i)) = \alpha \cdot_i (\beta \cdot x)(i) = (\alpha \cdot (\beta \cdot x))(i)$  and thus  $(\alpha \cdot \beta) \cdot x = \alpha \cdot (\beta \cdot x)$

4. If 1 is the unit in  $F$  and  $x \in \prod_{i \in I} V_i$  then  $\forall i \in I$  we have  $(1 \cdot x)(i) = 1 \cdot x(i) \underset{\langle V_i, +_i, \cdot_i \rangle \text{ is a vector space}}{=} x(i)$  and thus  $1 \cdot x = x$   $\square$

**Note 10.121.** To avoid having to use excessive notation we do not use anymore the notation:  $\langle V, +, \cdot \rangle$  is a vector space over a field  $\langle F, +, \cdot \rangle$

Instead we just say that  $V$  is a vector space over the field  $F$  and assume that  $+, \cdot$  are the additive and multiplicative operators in  $V, F$  (using context to find out which is which). Also we assume that the multiplicative units in  $V$  is 1 and the neutral element in  $V$  and  $F$  is 0 (again context will tell us which is which).

**Definition 10.122.** If  $V$  is a vector space over a field  $F$  and  $A, B \subseteq V, G \subseteq F, x \in V$  and  $\alpha \in F$  then we have the following definitions

1.  $x + A = \{x + y \mid y \in A\} = \{y + x \mid y \in A\} = A + x$
2.  $A + B = \{x + y \mid x \in A \wedge y \in B\}$
3.  $\alpha \cdot A = \{\alpha \cdot x \mid x \in A\}$
4.  $G \cdot A = \{\alpha \cdot x \mid \alpha \in G \wedge x \in A\}$
5.  $A - B = \{x - y \mid x \in A \wedge y \in B\}$

**Theorem 10.123.** If  $V$  is a vector space over a field  $F$  then we have

1.  $\forall A \subseteq V, x \in V$  we have  $y \in x + A \Leftrightarrow y - x \in A$
2.  $\forall A, B \subseteq V, x \in V$  we have  $x + (A \cup B) = (x + A) \cup (x + B)$
3.  $\forall A, B \subseteq V, x \in V$  we have  $x + (A \cap B) = (x + A) \cap (x + B)$
4.  $\forall A \subseteq V, x, y \in V$  we have  $x + (y + A) = (x + y) + A$

**Proof.**

1. If  $y \in x + A$  there exists a  $z \in A$  such that  $y = x + z \Rightarrow y - x = z \in A$ . If  $y - x \in A \Rightarrow y = x + (y - x) \in x + A$ .
2.  $y \in x + (A \cup B) \Leftrightarrow y - x \in A \cup B \Leftrightarrow (y - x \in A) \vee (y - x \in B) \Leftrightarrow (y \in x + A) \vee (y \in x + B) \Leftrightarrow y \in (x + A) \cup (x + B)$
3.  $y \in x + (A \cap B) \Leftrightarrow y - x \in A \cap B \Leftrightarrow (y - x \in A) \wedge (y - x \in B) \Leftrightarrow (y \in x + A) \wedge (y \in x + B) \Leftrightarrow y \in (x + A) \cap (x + B)$
4.  $z \in (x + (y + A)) \Leftrightarrow z - x \in y + A \Leftrightarrow (z - x) - y \in A \Leftrightarrow (z - (x + y)) \in A \Leftrightarrow z \in (x + y) + A \quad \square$

## 10.4.2 Linear dependency

**Lemma 10.124.** If  $V$  is a vector space over a field  $F, \alpha \in F, n \in \mathbb{N}$  and  $\{x_i\}_{i \in \{0, \dots, n\}}$  a family in  $V$  then  $\sum_{i=0}^n \alpha \cdot x_i = \alpha \cdot \sum_{i=0}^n x_i$

**Proof.** We prove this by mathematical induction so let  $X = \{n \in \mathbb{N} \mid \text{if } \{x_i\}_{i \in \{0, \dots, n\}}$  is a family of elements in  $V$  then  $\sum_{i=0}^n \alpha \cdot x_i = \alpha \cdot \sum_{i=0}^n x_i\}$  we have then:

1. For  $n = 0$  we have  $\sum_{i=0}^0 \alpha \cdot x_i = \alpha \cdot x_0 = \alpha \cdot \sum_{i=0}^0 x_i$  so  $0 \in X$
2. If  $n \in X$  and if  $\{x_i\}_{i \in \{0, \dots, n+1\}}$  is a family of elements in  $V$  then  $\sum_{i=0}^{n+1} \alpha \cdot x_i = (\sum_{i=0}^n \alpha \cdot x_i) + (\alpha \cdot x_{n+1}) \underset{n \in X}{=} \alpha \cdot (\sum_{i=0}^n \alpha \cdot x_i) + \alpha \cdot x_{n+1} = \alpha \cdot (\sum_{i=0}^n x_i + x_{n+1}) = \alpha \cdot \sum_{i=1}^{n+1} x_i$  so  $n+1 \in X$

Using mathematical induction we have then  $X = \mathbb{N}$  proving our theorem.  $\square$

**Lemma 10.125.** *If  $V$  is a vector space over a field  $F$ ,  $\{x_i\}_{i \in I}$  a family of sets with finite support and  $\alpha \in F$  then  $\{\alpha \cdot x_i\}_{i \in I}$  is a family of sets with finite support (so  $\sum_{i \in I} \alpha \cdot x_i$  is defined) and  $\sum_{i \in I} \alpha \cdot x_i = \alpha \cdot \sum_{i \in I} x_i$*

**Proof.** If  $i \in \text{support}(\{\alpha \cdot x_i\}_{i \in I})$  then  $\alpha \cdot x_i \neq 0$ , now if  $x_i = 0$  then by 10.108 we have  $\alpha \cdot x_i = 0$  contradicting  $\alpha \cdot x_i \neq 0$  so we must have  $x_i \neq 0$  and thus  $i \in \text{support}(\{x_i\}_{i \in I})$  or in other words  $\text{support}(\{\alpha \cdot x_i\}_{i \in I}) \subseteq \text{support}(\{x_i\}_{i \in I})$ . So as  $\text{support}(\{\alpha \cdot x_i\}_{i \in I})$  is finite we have by 5.36 that  $\text{support}(\{\alpha \cdot x_i\}_{i \in I})$  is finite.

Using the definition of the generalized sum (see 10.36 and 10.51) we have  $\sum_{i \in I} \alpha \cdot x_i = \sum_{i=0}^{n-1} \alpha \cdot x_{b(i)}$  where  $b: S_n \rightarrow \text{support}(\{x_i\}_{i \in I})$  is a bijection. By the previous lemma we have then that  $\sum_{i=0}^{n-1} \alpha \cdot x_{b(i)} = \alpha \cdot \sum_{i=0}^{n-1} x_{b(i)} = \alpha \cdot \sum_{i \in I} x_i$  and thus  $\sum_{i \in I} \alpha \cdot x_i = \alpha \cdot \sum_{i \in I} x_i$ .  $\square$

**Lemma 10.126.** *If  $V$  is a vector space over a field  $F$ ,  $v \in V$  and  $\{\alpha_i\}_{i \in \{0, \dots, n\}}$  a family in  $F$  then  $\sum_{i=0}^n \alpha_i \cdot v = (\sum_{i=0}^n \alpha_i) \cdot v$*

**Proof.** We prove this by induction so let  $S = \{n \in \mathbb{N} \mid \text{if } \{\alpha_i\}_{i \in \{0, \dots, n\}} \text{ is a family in } F \text{ and } v \in V \text{ then } \sum_{i=0}^n \alpha_i \cdot v = (\sum_{i=0}^n \alpha_i) \cdot v\}$ . So we have:

1. If  $n = 0$  then  $\sum_{i=0}^0 \alpha_i \cdot v = \alpha_0 \cdot v = (\sum_{i=0}^0 \alpha_i) \cdot v$
2. If  $n \in S$  then if  $\{\alpha_i\}_{i \in \{0, \dots, n+1\}}$  is a family in  $F$  and  $v \in V$  then  $\sum_{i=0}^{n+1} \alpha_i \cdot v = (\sum_{i=0}^n \alpha_i \cdot v) + \alpha_{n+1} \cdot v = (\sum_{i=0}^n \alpha_i) \cdot v + \alpha_{n+1} \cdot v = (\sum_{i=0}^n \alpha_i + \alpha_{n+1}) \cdot v = (\sum_{i=0}^{n+1} \alpha_i) \cdot v$  and thus  $n+1 \in S$

Using mathematical induction we have then  $S = \mathbb{N}$  proving our theorem.  $\square$

**Lemma 10.127.** *If  $V$  is a vector space over a field  $F$ ,  $v \in V$  and  $\{\alpha_i\}_{i \in I}$  a family in  $F$  with finite support then  $\{\alpha_i \cdot v\}_{i \in I}$  has finite support and  $\sum_{i \in I} \alpha_i \cdot v = (\sum_{i \in I} \alpha_i) \cdot v$*

**Proof.** If  $i \in \text{support}(\{\alpha_i \cdot v\}_{i \in I}) \Rightarrow \alpha_i \cdot v \neq 0 \Rightarrow \alpha_i \neq 0 \Rightarrow i \in \text{support}(\{\alpha_i\}_{i \in I}) \Rightarrow \text{support}(\{\alpha_i \cdot v\}_{i \in I}) \subseteq \text{support}(\{\alpha_i\}_{i \in I})$  and thus  $\text{support}(\{\alpha_i \cdot v\}_{i \in I})$  is finite. If  $b: S_n \rightarrow \text{support}(\{\alpha_i\}_{i \in I})$  then using the definition of the generalized sum (see 10.36 and 10.51) we have  $\sum_{i \in I} \alpha_i \cdot v = \sum_{i=0}^{n-1} \alpha_i \cdot v = (\sum_{i=0}^{n-1} \alpha_i) \cdot v = (\sum_{i \in I} \alpha_i) \cdot v$ .  $\square$

**Theorem 10.128.** *If  $V$  is a vector space over a field  $F$ ,  $\{v_i\}_{i \in I}$  a family in  $V$  and  $\{\alpha_i\}_{i \in I}$  a family in  $F$  with finite support then  $\{\alpha_i \cdot v_i\}_{i \in I}$  is a family in  $V$  with finite support and  $\text{support}(\{\alpha_i \cdot v_i\}_{i \in I}) \subseteq \text{support}(\{\alpha_i\}_{i \in I})$ .*

**Proof.** If  $i \in \text{support}(\{\alpha_i \cdot v_i\}_{i \in I}) \Rightarrow \alpha_i \cdot v_i \neq 0 \Rightarrow \alpha_i \neq 0$  [otherwise  $\alpha_i \cdot v_i = 0$ ]  $\Rightarrow i \in \text{support}(\{\alpha_i\}_{i \in I}) \Rightarrow \text{support}(\{\alpha_i \cdot v_i\}_{i \in I}) \subseteq \text{support}(\{\alpha_i\}_{i \in I})$ . The theorem follows then from 5.36 and the fact that  $\text{support}(\{\alpha_i\}_{i \in I})$  is finite.  $\square$

The above theorem ensures that the next definition make sense.

**Definition 10.129.** *Let  $V$  be a vector space over a field  $F$  then if  $\{v_i\}_{i \in I}$  is a family in  $V$  and  $v \in V$  then  $v$  is a **linear combination of  $\{v_i\}_{i \in I}$**  if there exists a family  $\{\alpha_i\}_{i \in I}$  with finite support such that  $v = \sum_{i \in I} \alpha_i \cdot v_i$  (sum is defined as  $\{\alpha_i \cdot v_i\}_{i \in I}$  has finite support by the previous theorem). If  $W \subseteq V$  then a **linear combination of  $W$**  is defined to be the **linear combination** of the family  $\{w\}_{w \in W}$  (see 2.111).*

**Definition 10.130.** *Let  $V$  be a vector space over a field  $F$ ,  $W \subseteq V$  in  $V$  then  $\mathcal{S}(W)$  the **linear span of  $W$**  is defined by  $\mathcal{S}(W) = \{v \in V \mid \exists \{\alpha_w\}_{w \in W} \text{ with finite support such that } v = \sum_{w \in W} \alpha_w \cdot w\}$ . If  $\{v_i\}_{i \in I}$  is a family of elements in  $V$  then  $\mathcal{S}(\{v_i\}_{i \in I})$  is defined by  $\mathcal{S}(\{v_i\}_{i \in I}) \stackrel{\text{definition}}{=} \mathcal{S}(\{v_i \mid i \in I\})$  (see 2.108).*

**Definition 10.131.** *Let  $V$  be a vector space over a field  $F$ ,  $A \subseteq V$  then if  $W \subseteq V$  we say that  $W$  spans  $A$  if  $\mathcal{S}(W) = A$*

**Theorem 10.132.** Let  $V$  be a vector space over a field  $F$ ,  $W \subseteq V$  then  $\mathcal{S}(W)$  is a subspace of  $V$ . If  $W \neq \emptyset$  then  $\mathcal{S}(W)$  is a vector space. (and thus by 10.111  $\mathcal{S}(W)$  is a sub space).

**Proof.** Let  $\alpha, \beta \in F$  and  $x, y \in \mathcal{S}(W)$  then there exists  $\{\alpha_w\}_{w \in W}, \{\beta_w\}_{w \in W}$  such that  $x = \sum_{w \in W} \alpha_w \cdot w, y = \sum_{w \in W} \beta_w \cdot w$ . So  $\alpha \cdot x + \beta \cdot y = \alpha \cdot \sum_{w \in W} \alpha_w \cdot w + \beta \cdot \sum_{w \in W} \beta_w \cdot w \stackrel{10.125}{=} \sum_{w \in W} (\alpha \cdot \alpha_w + \beta \cdot \beta_w) \cdot w \in \mathcal{S}(W)$ .

If  $W \neq \emptyset$  then there exists a  $w \in W$  and then  $0 = 0 \cdot w \in \mathcal{S}(W)$ , using 10.111 we have then that  $\mathcal{S}(W)$  is a vector space.  $\square$

**Theorem 10.133.** Let  $V$  be a vector space over a field  $F$  and  $W \subseteq V$  a subspace of  $V$  then

1. If  $\{w_i\}_{i \in \{0, \dots, n\}}$  is a family in  $W$  then  $\sum_{i=0}^n w_i \in W$
2. If  $\{w_i\}_{i \in I}$  is a family in  $W$  with finite support then  $\sum_{i \in I} w_i \in W$

**Proof.**

1. We proceed by induction, let  $S = \{n \in \mathbb{N} \mid \text{If } \{w_i\}_{i \in \{0, \dots, n\}} \text{ is a family in } W \text{ then } \sum_{i=0}^n w_i \in W\}$ , then:
  - a.  $\sum_0^0 w_i = w_0 \in W \Rightarrow 0 \in S$
  - b. If  $n \in S$  take then  $\{w_i\}_{i \in \{0, \dots, n+1\}}$  in  $W$  then  $\sum_{i=0}^{n+1} w_i = (\sum_{i=0}^n w_i) + w_{n+1} \in W$  [As  $\sum_{i=0}^n w_i \in W$  (because  $n \in S$ ),  $w_{n+1} \in W$  and  $W$  is a subspace] so  $n+1 \in S$

Using mathematical induction we have then  $S = \mathbb{N}$  proving (1).

2. If  $\{w_i\}_{i \in I}$  is a family with finite support then by 10.36 there exists a bijection  $b: \{0, \dots, n-1\} \rightarrow \text{support}(\{w_i\}_{i \in I})$  such that  $\sum_{i \in I} w_i = \sum_{i=0}^{n-1} w_{b(i)} \in W$  (see (1))  $\square$

It is now easy to write a equivalent definition of  $\mathcal{S}(W)$  using finite families only

**Theorem 10.134.** Let  $V$  be a vector space over a field  $F$  and  $W \subseteq V$  then  $\mathcal{S}(W) = \{x \in V \mid \exists \{\alpha_i\}_{i \in \{0, \dots, n-1\}} \text{ a finite family in } F \text{ and } \exists \{v_i\}_{i \in \{0, \dots, n-1\}} \text{ a finite family in } V \text{ such that } x = \sum_{i \in \{0, \dots, n-1\}} \alpha_i \cdot v_i\}$  (note we use  $n-1$  to indicate that  $\{0, \dots, n-1\}$  can be the empty set (so that the sum is 0)).

**Proof.** Let  $\mathcal{V} = \{x \in V \mid \exists \{\alpha_i\}_{i \in \{1, \dots, n\}} \text{ a finite family in } F \text{ and } \exists \{v_i\}_{i \in \{1, \dots, n\}} \text{ a finite family in } V \text{ such that } x = \sum_{i \in \{1, \dots, n\}} \alpha_i \cdot v_i\}$ . If now  $x \in \mathcal{V}$  then there exists a  $\{\alpha_i\}_{i \in \{0, \dots, n\}}$  and  $\{v_i\}_{i \in \{0, \dots, n\}}$  such that  $x = \sum_{i \in \{0, \dots, n\}} \alpha_i \cdot v_i = \sum_{i=0}^n \alpha_i \cdot v_i$  which as  $\forall i \in \{1, \dots, n\}$  we have  $\alpha_i \cdot v_i \in \mathcal{S}(W)$  [as  $\mathcal{S}(W)$  is a vector space and  $W \subseteq \mathcal{S}(W)$ ] means by the above theorem that  $x \in \mathcal{S}(W)$  so we have proved that

$$\mathcal{V} \subseteq \mathcal{S}(W) \quad (10.20)$$

If  $x \in \mathcal{S}(W)$  then there exists a family  $\{\alpha_w\}_{w \in W}$  with finite support such that  $x = \sum_{w \in W} \alpha_w \cdot w$ . As  $\text{support}(\{\alpha_w \cdot w\}_{w \in W})$  is finite there exists a bijection  $b: \{0, \dots, n-1\} \rightarrow \text{support}(\{\alpha_w \cdot w\}_{w \in W})$ . If we define then the finite families  $\{\alpha_{b(i)}\}_{i \in \{0, \dots, n-1\}}, \{b(i)\}_{i \in \{0, \dots, n-1\}}$  then we have that  $x = \sum_{w \in W} \alpha_w \cdot w \stackrel{10.36}{=} \sum_{i=0}^{n-1} \alpha_{b(i)} \cdot b(i) = \sum_{i \in \{0, \dots, n-1\}} \alpha_{b(i)} \cdot b(i) \in \mathcal{V}$  proving that

$$\mathcal{S}(W) \subseteq \mathcal{V} \quad (10.21)$$

Using 10.20 and 10.21 we have  $\mathcal{V} = \mathcal{S}(W)$  proving our theorem.  $\square$

**Theorem 10.135.** Let  $V$  be a vector space over a field  $F$  and  $W \subseteq V$  then we have  $W \subseteq \mathcal{S}(W)$

**Proof.** If  $w_0 \in W$  define then  $\{\alpha_w\}_{w \in W}$  a family in  $F$  by  $\alpha_w = \begin{cases} 1 & \text{if } w = w_0 \\ 0 & \text{if } w \in W \setminus \{w_0\} \end{cases}$ , this family clearly has finite support and  $\sum_{w \in W} \alpha_w \cdot w = \sum_{w \in W \setminus \{w_0\}} \alpha_w \cdot w + \sum_{w \in \{w_0\}} \alpha_w \cdot w = \sum_{w \in W \setminus \{w_0\}} 0 \cdot w + 1 \cdot w_0 = w_0$  so  $w_0 \in \mathcal{S}(W)$   $\square$

**Theorem 10.136.** *Let  $V$  be a vector space over the field  $F$  and  $W \subseteq V$  a subspace of  $V$  then  $\mathcal{S}(W) = W$*

**Proof.** By 10.135 we have that  $W \subseteq \mathcal{S}(W)$ . Now if  $v \in \mathcal{S}(W)$  then  $\exists \{\alpha_w\}_{w \in W}$  in  $F$  with finite support such that  $v = \sum_{w \in W} \alpha_w \cdot w$ . Using 10.133 and the fact that  $\alpha_w \cdot w \in W$  we have  $\sum_{w \in W} \alpha_w \cdot w \in W \Rightarrow v \in W \Rightarrow \mathcal{S}(W) \subseteq W$  and thus  $W = \mathcal{S}(W)$ .  $\square$

**Corollary 10.137.** *If  $V$  is a vector space over a field  $F$  and  $W \subseteq V$  then  $\mathcal{S}(W) = \mathcal{S}(\mathcal{S}(W))$*

**Proof.** If  $W = \emptyset$  then  $\mathcal{S}(W) = \emptyset$  and thus  $\mathcal{S}(\mathcal{S}(W)) = \emptyset = \mathcal{S}(W)$ . If  $W \neq \emptyset$  then use the above theorem and the fact that  $\mathcal{S}(W)$  is a subspace (see 10.132) is  $W \neq \emptyset$ .  $\square$

**Theorem 10.138.** *Let  $V$  be a vector space over a field  $F$ ,  $W_1, W_2 \subseteq V$  with  $W_1 \subseteq W_2$  then  $\mathcal{S}(W_1) \subseteq \mathcal{S}(W_2)$*

**Proof.** We have the following cases to consider:

1.  $(W_1 = W_2)$  then trivially  $\mathcal{S}(W_1) = \mathcal{S}(W_2) \Rightarrow \mathcal{S}(W_1) \subseteq \mathcal{S}(W_2)$
2.  $(W_1 \subset W_2)$  if  $v \in \mathcal{S}(W_1)$  then there exists a  $\{\alpha_w\}_{w \in W_1}$  in  $F$  with finite support such that  $v = \sum_{w \in W_1} \alpha_w \cdot w$ . Define then  $\{\beta_w\}_{w \in W_2}$  by  $\beta_w = \begin{cases} \alpha_w & \text{if } w \in W_1 \\ 0 & \text{if } w \in W_2 \setminus W_1 \end{cases}$  then  $\text{support}(\{\alpha_w\}_{w \in W_1}) = \text{support}(\{\beta_w\}_{w \in W_2})$  and thus  $\{\beta_w\}_{w \in W_2}$  has finite support, further  $W_2 \ni \sum_{w \in W_2} \beta_w \cdot w = \sum_{w \in W_2 \setminus W_1} \beta_w \cdot w + \sum_{w \in W_1} \beta_w \cdot w = \sum_{w \in W_2 \setminus W_1} 0 \cdot w + \sum_{w \in W_1} \alpha_w \cdot w = v$  so  $\mathcal{S}(W_1) \subseteq \mathcal{S}(W_2)$   $\square$

**Definition 10.139.** *Let  $V$  be a vector space over a field  $F$  and  $\{v_i\}_{i \in I}$  a family of elements in  $V$  then  $\{v_i\}_{i \in I}$  is **linear independent** if  $\forall \{\alpha_i\}_{i \in I}$  in  $F$  with finite support such that  $\sum_{i \in I} \alpha_i \cdot v_i = 0$  we have that  $\forall i \in I \alpha_i = 0$  (in other words  $\text{support}(\{\alpha_i\}_{i \in I}) = \emptyset$ ). If  $\{v_i\}_{i \in I}$  is not linear independent it is **linear dependent**.*

Note that we can also define linear dependency and linear dependency for a set based on the fact that we can give as set  $W$  define the family  $\{w\}_{w \in W}$  by self indexing (see 2.105)

**Definition 10.140.** *If  $V$  is a vector space over a field  $F$  and  $W \subseteq V$  then  $W$  is linear independent (or linear dependent) if  $\{w\}_{w \in W}$  is linear independent (or linear dependent)*

The following theorem gives a relation between linear independence of a family and its associated set.

**Theorem 10.141.** *If  $V$  is a vector space over a field  $F$  and  $\{v_i\}_{i \in I}$  a family in  $V$ . If  $b: J \rightarrow I$  is a bijection then  $\{v_i\}_{i \in I}$  is linear dependent (or linear independent) then  $\{v_{b_j}\}_{j \in J}$  is linear dependent (or linear independent).*

**Proof.** First we observe that if  $\{\alpha_i\}_{i \in I}$  is a family in  $F$  with finite support we can define a family  $\{\alpha'_j\}_{j \in J}$  by  $\alpha'_j = \alpha_{b(j)}$  and thus  $\sum_{j \in J} \alpha'_j \cdot v_{b_j} = \sum_{j \in J} \alpha_{b(j)} \cdot v_{b(j)} \stackrel{10.43}{=} \sum_{i \in I} \alpha_i \cdot v_i$ . Also if  $\{\beta_j\}_{j \in J}$  is a family in  $F$  with finite support we can define a family  $\{\beta''_i\}_{i \in I}$  by  $\beta''_i = \beta_{b^{-1}(i)}$  so that  $\beta''_{b(i)} = \beta_{b^{-1}(b(i))} = \beta_i$  we have then that  $\sum_{j \in J} \beta_j \cdot v_{b_j} = \sum_{j \in J} \beta''_{b(i)} \cdot v_{b(i)} = \sum_{i \in I} \beta''_i \cdot v_i$

1. If  $\{v_i\}_{i \in I}$  is linear dependent then there exists a  $\{\alpha_i\}_{i \in I}$  such that  $\exists i_0 \in I \vdash \alpha_{i_0} \neq 0$  and  $0 = \sum_{i \in I} \alpha_i \cdot v_i = \sum_{j \in J} \alpha'_j \cdot v_{b_j}$  where for  $j_0 = b(i_0)$  we have  $\alpha'_{j_0} = \alpha_{b(i_0)} \neq 0$  proving that  $\{v_{b_j}\}_{j \in J}$  is linear dependent.
2. If now  $\{v_i\}_{i \in I}$  is linear independent and  $\{\beta_j\}_{j \in J}$  a family in  $F$  such that  $0 = \sum_{j \in J} \beta_j \cdot v_{b_j} = \sum_{i \in I} \beta''_i \cdot v_i$  and by linear Independence we have then that  $\forall i \in I$  we have  $\beta''_i = 0$ . Then  $\forall j \in J$  we have  $\beta_j = \beta''_{b^{-1}(i)} = 0$  proving linear Independence of  $\{v_{b_j}\}_{j \in J}$   $\square$

**Theorem 10.142.** Let  $V$  be a vector space over a field  $F$ ,  $\{v_i\}_{i \in I}$  a family in  $V$  with  $\exists i, j \in I, i \neq j$  with  $v_i = v_j$  then  $\{v_i\}_{i \in I}$  is linear dependent.

**Proof.** If  $i, j \in I$  with  $i \neq j$  and  $v_i = v_j$  define then  $\{\alpha_i\}_{i \in I}$  by  $\alpha_k = \begin{cases} 1 & \text{if } k = i \\ -1 & \text{if } k = j \\ 0 & \text{if } k \in I \setminus \{i, j\} \end{cases}$  then  $\{\alpha_i\}_{i \in I}$  has finite support and is not all zero. We have then  $\sum_{k \in I} \alpha_k \cdot v_k = \sum_{k \in I \setminus \{i, j\}} \alpha_k \cdot v_k + \sum_{k \in \{i, j\}} \alpha_k \cdot v_k + \sum_{k \in \{j\}} \alpha_k \cdot v_k = \sum_{k \in I \setminus \{i, j\}} 0 \cdot v + \alpha_i \cdot v_i + \alpha_j \cdot v_j = 1 \cdot v_i + (-1) \cdot v_j = v_i + (-v_j) \underset{v_i = v_j}{=} 0$ .  $\square$

**Corollary 10.143.** Let  $V$  be a vector space over a field  $F$  and  $\{v_i\}_{i \in I}$  a linear independent family then if  $v_i = v_j$  we have that  $i = j$  (so  $\{v_i\}_{i \in I}: I \rightarrow \{v_i | i \in I\}$  forms a bijection).

**Proof.** If  $v_i = v_j$  with  $i \neq j$  then  $\{v_i\}_{i \in I}$  is linear dependent contradicting the linear Independence.  $\square$

**Theorem 10.144.** If  $V$  is a vector space over over a field  $F$  and  $\{v_i\}_{i \in I}$  is a linear independent family in  $V$  then  $\{v_i | i \in I\}$  is also linear independent.

**Proof.** Define  $V = \{v_i | i \in I\}$  then using the above corollary we have that  $v: I \rightarrow V$  defined by  $i \rightarrow v_i$  is a bijection. Assume now that  $\{\alpha_w\}_{w \in V}$  is a family in  $F$  with finite support such that  $\sum_{w \in V} \alpha_w \cdot w = 0$ . Define then  $\{\beta_i\}_{i \in I}$  by  $\beta_i = \alpha_{v(i)}$ , then  $\text{support}(\{\beta_i\}_{i \in I}) = v^{-1}(\text{support}(\{\alpha_w\}_{w \in V}))$  is finite, and we have  $\sum_{i \in I} \beta_i \cdot v_i = \sum_{i \in I} \alpha_{v(i)} \cdot v(i) \underset{10.43}{=} \sum_{w \in V} \alpha_w \cdot w = 0$  so by linear Independence of  $\{v_i\}_{i \in I}$  we have that  $\forall i \in I$  we have  $\forall i \in I$  that  $\beta_i = 0$ . Finally  $\forall w \in V$  we have  $v^{-1}(w) \in I$  and  $0 = \beta_{v^{-1}(w)} = \alpha_{v(v^{-1}(w))} = \alpha_w \Rightarrow \forall w \in V$  we have  $\alpha_w = 0$  proving that  $V = \{v_i | i \in I\}$  is linear independent.  $\square$

Note that the opposite is not true, for example if  $0 \neq v \in V$  and we define  $\{v_i\}_{i \in \{1, 2\}}$  by  $v_1 = v = v_2$  then  $\{v_i | i \in \{1, 2\}\} = \{v\}$  is a linear independent but  $\{v_i\}_{i \in \{1, 2\}}$  is not as  $v_1 = v_2$ . However if there are no duplicates in a family then the opposite is true.

**Theorem 10.145.** If  $V$  is a vector space over a field  $F$  and  $\{v_i\}_{i \in I}$  a family such that  $\forall i, j \in I$  with  $v_i = v_j$  we have  $i = j$ . Then we have

1.  $\{v_i\}_{i \in I}$  is linear independent if and only if  $\{v_i | i \in I\}$  is linear independent
2.  $\{v_i\}_{i \in I}$  is linear dependent if and only if  $\{v_i | i \in I\}$  is linear dependent

**Proof.** Define  $V = \{v_i | i \in I\}$  then  $v: I \rightarrow V$  defined by  $i \rightarrow v(i) = v_i$  (the actual definition of a family) is a bijection.

1. Assume that  $\{v_i\}_{i \in I}$  is linear independent then using 10.144 we have that  $\{v_i | i \in I\}$  is linear independent.
2. Assume that  $\{v_i | i \in I\}$  is linear independent. Then if  $\{\alpha_i\}_{i \in I}$  is a family in  $F$  with finite support such that  $0 = \sum_{i \in I} \alpha_i \cdot v_i$  define then  $\{\beta_w\}_{w \in V}$  by  $\beta_w = \alpha_{v^{-1}(w)}$  then  $\text{support}(\{\beta_w\}_{w \in V}) = v(\text{support}(\{\alpha_i\}_{i \in I}))$  is finite,  $\beta_{v(i)} = \alpha_{v^{-1}(v(i))} = \alpha_i$  and  $0 = \sum_{i \in I} \alpha_i \cdot v_i = \sum_{i \in I} \beta_{v(i)} \cdot v(i) \underset{10.43}{=} \sum_{w \in V} \beta_w \cdot w$  and by linear independence of  $V = \{v_i | i \in I\}$  we have  $\forall w \in V$  that  $\beta_w = 0$ . If now  $i \in I$  then  $v(i) \in V \Rightarrow \alpha_i = \beta_{v(i)} = 0$  so we have  $\forall i \in I$  that  $\alpha_i$  proving that  $\{v_i\}_{i \in I}$  is linear independent.
3. Assume that  $\{v_i\}_{i \in I}$  is linear dependent then if  $\{v_i | i \in I\}$  would be linear independent we would have by (2) the contradiction that  $\{v_i\}_{i \in I}$  would be linear independent. So we must have that  $\{v_i | i \in I\}$  is linear dependent.
4. Assume that  $\{v_i | i \in I\}$  is linear dependent then if  $\{v_i\}_{i \in I}$  would be linear independent we would have by (1) the contradiction that  $\{v_i | i \in I\}$  would be linear independent. So we must have that  $\{v_i\}_{i \in I}$  is linear dependent.  $\square$

**Theorem 10.146.** Let  $V$  be a vector space over a field  $F$  then  $\{v_i\}_{i \in I}$  in  $V$  is linear dependent if and only if  $\exists i \in I$  and a  $\{\alpha_k\}_{k \in I \setminus \{i\}}$  with finite support such that  $v_i = \sum_{k \in I \setminus \{i\}} \alpha_k \cdot v_k$

**Proof.**

If  $\{v_i\}_{i \in I}$  is linear dependent then there exists a family  $\{\alpha_i\}_{i \in I}$  in  $F$  with finite support such that  $\sum_{i \in I} \alpha_i \cdot v_i = 0$  and  $\exists i \in I$  such that  $\alpha_i \neq 0$ . We have then that  $0 = \sum_{j \in I \setminus \{i\}} \alpha_i \cdot v_i + \sum_{j \in \{i\}} \alpha_j \cdot v_j = \sum_{j \in I \setminus \{i\}} \alpha_j \cdot v_j + \alpha_i \cdot v_i$  and thus as  $\alpha_i \neq 0$  we have

$$v_i = (-\alpha_i^{-1}) \cdot \sum_{j \in I \setminus \{i\}} \alpha_j \cdot v_j \quad (10.22)$$

If we take now  $\{\beta_j\}_{j \in I \setminus \{i\}}$  by  $\beta_j = -\alpha_i^{-1} \cdot \alpha_j$  which has trivially finite support (as  $\{\alpha_i\}_{i \in I}$  has finite support) then we have  $\sum_{j \in I \setminus \{i\}} \beta_j \cdot v_i = \sum_{j \in I \setminus \{i\}} (-\alpha_i^{-1} \cdot \alpha_j) \cdot v_j = \sum_{j \in I \setminus \{i\}} (-\alpha_i^{-1}) \cdot (\alpha_j \cdot v_j) = (-\alpha_i^{-1}) \cdot \sum_{j \in I \setminus \{i\}} \alpha_j \cdot v_j \stackrel{10.22}{=} v_i \Rightarrow \sum_{j \in I \setminus \{i\}} \alpha_j \cdot v_j = v_i$ .

If there exists a  $i \in I$  and a  $\{\alpha_k\}_{k \in I \setminus \{i\}}$  with finite support such that  $v_i = \sum_{k \in I \setminus \{i\}} \alpha_k \cdot v_k$  then

$$-v_i + \sum_{k \in I \setminus \{i\}} \alpha_k \cdot v_k = 0 \quad (10.23)$$

take now  $\{\beta_k\}_{k \in I}$  by  $\beta_k = \begin{cases} \alpha_k & \text{if } k \in I \setminus \{i\} \\ -1 & \text{if } k = i \end{cases}$  then we have  $\sum_{k \in I} \beta_k \cdot v_k = \sum_{k \in I \setminus \{i\}} \beta_k \cdot v_k + \sum_{k \in \{i\}} \beta_k \cdot v_k = \sum_{k \in I \setminus \{i\}} \alpha_k \cdot v_k + (-v_i) \stackrel{10.23}{=} 0$  so we have  $\sum_{k \in I} \beta_k \cdot v_k = 0$  and  $\beta_i = -1 \neq 0$  proving that  $\{v_i\}_{i \in I}$  is linear dependent.  $\square$

**Theorem 10.147.** Let  $V$  be a vector space over a field  $F$ ,  $\{v_i\}_{i \in I}$  a family in  $V$  with  $0 \in \{v_i | i \in I\}$  then  $\{v_i\}_{i \in I}$  is linear dependent.

**Proof.** As  $0 \in \{v_i | i \in I\}$  there exists a  $i \in I$  such that  $0 = v_i$ , if we take now  $\{\alpha_k\}_{k \in I \setminus \{i\}}$  defined by  $\alpha_k = 0$  if  $k \in I \setminus \{i\}$ . Then  $\sum_{k \in I \setminus \{i\}} \alpha_k \cdot v_k = \sum_{k \in I \setminus \{i\}} 0 \cdot v_k = 0 = v_i$ . Using the above theorem we have then that  $\{v_i\}_{i \in I}$  is linear dependent.  $\square$

**Definition 10.148.** If  $S, I, J$  are sets and  $\{x_i\}_{i \in I}$ ,  $\{y_j\}_{j \in J}$  are families in  $S$  then we say that  $\{x_i\}_{i \in I} \preceq \{y_j\}_{j \in J}$  if  $I \subseteq J$  and  $\forall i \in I$  we have  $x_i = y_i$ . We say that the  $\{y_j\}_{j \in J}$  is a **extension** of  $\{x_i\}_{i \in I}$ . Note that if  $X, Y \in S$  with  $X \subseteq Y$  then trivially we have  $\{x\}_{x \in X} \preceq \{y\}_{y \in Y}$

**Theorem 10.149.** Let  $V$  be a vector space over a field  $F$ ,  $\{v_i\}_{i \in I}$  a linear dependent family in  $V$  and  $\{w_i\}_{i \in J}$  a family in  $V$  such that  $\{v_i\}_{i \in I} \preceq \{w_i\}_{i \in J}$ . In particular this means that if  $W_1, W_2 \subseteq V$  is such that  $W_1$  is linear dependent and  $W_1 \subseteq W_2$  then  $W_2$  is linear dependent.

**Proof.** As  $\{v_i\}_{i \in I}$  is linear dependent there exists a family  $\{\alpha_i\}_{i \in I}$  in  $F$  with finite support not all zero such that  $\sum_{i \in I} \alpha_i \cdot v_i$ . Then as  $I \subseteq J$  we can take then  $\{\beta_i\}_{i \in J}$  defined by  $\beta_i = \begin{cases} \alpha_i & \text{if } i \in I \\ 0 & \text{if } i \in J \setminus I \end{cases}$  which has trivially finite support and is not all zero. We have then that  $\sum_{i \in J} \beta_i \cdot w_i = \sum_{i \in I} \beta_i \cdot w_i + \sum_{i \in J \setminus I} \beta_i \cdot w_i = \sum_{i \in J \setminus I} 0 \cdot w_i + \sum_{i \in I} \alpha_i \cdot w_i \stackrel{\forall i \in I \models v_i = w_i}{=} \sum_{i \in I} \alpha_i \cdot v_i = 0$  proving that  $\{w_i\}_{i \in J}$  is linear dependent.  $\square$

**Definition 10.150.** Let  $V$  be a vector space over a field  $F$  then a family  $\{v_i\}_{i \in I}$  in  $V$  is a basis if

1.  $\mathcal{S}(\{v_i | i \in I\}) = V$
2.  $\{v_i\}_{i \in I}$  is linear independent

**Theorem 10.151.** Let  $V$  be a vector space over a field  $F$  then if  $\{v_i\}_{i \in I}$  is a basis of  $V$  and  $b: J \rightarrow I$  is a bijection we have that  $\{v_{b_i}\}_{i \in I}$  is a basis of  $V$ .

**Proof.** Using 10.141 we have that  $\{v_{b_i}\}_{i \in I}$  is linear independent. If  $v \in \{v_{b(i)} | i \in J\}$  then  $\exists i \in J$  such that  $v = v_{b_i} \in \{v_i | i \in I\} \Rightarrow \{v_{b_i} | i \in J\} \subseteq \{v_i | i \in I\}$  also if  $v \in \{v_i | i \in I\}$  then  $\exists i \in I$  such that  $v = v_i = v_{b(b^{-1}(i))} \in \{v_{b_j} | j \in J\} \Rightarrow \{v_i | i \in I\} \subseteq \{v_{b_j} | j \in J\}$  giving that  $\{v_i | i \in I\} = \{v_{b_j} | j \in J\}$  and thus  $V = \mathcal{S}(\{v_i | i \in I\}) = \mathcal{S}(\{v_{b_j} | j \in J\})$ .  $\square$

**Theorem 10.152.** If  $V$  is a non trivial vector space (so  $V \neq \{0\}$ ) with a basis  $\{v_i\}_{i \in I}$  then  $I \neq \emptyset$

**Proof.** If  $V$  is non trivial then  $\exists v \in V$  such that  $v \neq 0$ . Then there must exists a  $\{\alpha_i\}_{i \in I}$  with finite support such that  $v = \sum_{i \in I} \alpha_i \cdot v_i$ , if now  $\text{support}(\{\alpha_i\}_{i \in I}) = \emptyset$  then we would have  $0 = \sum_{i \in I} \alpha_i \cdot v_i = v \neq 0$  a contradiction. So there must exists a  $i \in I$  such that  $\alpha_i \neq 0 \Rightarrow I \neq \emptyset$ .  $\square$

**Theorem 10.153.** Let  $V$  be a vector space over a field  $F$  and  $\{v_i\}_{i \in I}$  a family in  $V$  then  $\{v_i\}_{i \in I}$  is a basis of  $V$  if and only if  $\forall v \in V$  there exists a unique family  $\{\alpha_i\}_{i \in I}$  in  $F$  with finite support such that  $v = \sum_{i \in I} \alpha_i \cdot v_i$ .

**Proof.**

If  $\{v_i\}_{i \in I}$  is a basis then  $\{v_i\}_{i \in I}$  is linear independent and  $\mathcal{S}(\{v_i | i \in I\}) = V$  so if  $v \in V$  then there exists a  $\{\alpha_i\}_{i \in I}$  with finite support such that  $v = \sum_{i \in I} \alpha_i \cdot v_i$ . Now to prove uniqueness let's assume that there exists another  $\{\beta_i\}_{i \in I}$  with finite support so that  $v = \sum_{i \in I} \beta_i \cdot v_i$  then we have  $0 = v + (-1) \cdot v = \sum_{i \in I} \alpha_i \cdot v_i + (-1) \cdot \sum_{i \in I} \beta_i \cdot v_i = \sum_{i \in I} \alpha_i \cdot v_i + \sum_{i \in I} (-\beta_i) \cdot v_i = \sum_{i \in I} (\alpha_i - \beta_i) \cdot v_i$ .  $\{v_i\}_{i \in I}$  is independent  $\Rightarrow \forall i \in I$  we have  $\alpha_i - \beta_i = 0 \Rightarrow \alpha_i = \beta_i$  and thus we find that  $\{\alpha_i\}_{i \in I} = \{\beta_i\}_{i \in I}$ .

Assume now that  $\{v_i\}_{i \in I}$  is a family in  $V$  such that  $\forall v \in V$  there exists a unique  $\{\alpha_i\}_{i \in I}$  with finite support so that  $v = \sum_{i \in I} \alpha_i \cdot v_i$ . Then clearly  $\mathcal{S}(\{v_i\}_{i \in I}) = V$ . If now  $\{\alpha_i\}_{i \in I}$  is a family in  $F$  with finite support so that  $0 = \sum_{i \in I} \alpha_i \cdot v_i$  if we define now  $\{\beta_i\}_{i \in I}$  where  $\forall i \in I$  we have  $\beta_i$  then we have also  $0 = \sum_{i \in I} \beta_i \cdot v_i$ , from uniqueness it follows that  $\forall i \in I$  we have  $\alpha_i = \beta_i = 0$  and thus we have proved linear Independence of  $\{v_i\}_{i \in I}$ .  $\square$

**Notation 10.154.** Let  $V$  be a vector space over a field  $F$  with basis  $\{e_i\}_{i \in I}$  then given  $v \in V$  there exists a unique  $\{\alpha_i\}_{i \in I}$  in  $F$  such that  $v = \sum_{i \in I} \alpha_i \cdot e_i$  we note then  $\forall i \in I$  then unique element  $\alpha_i \in F$  as  $v_i$  and call this the  $i$ -the coordinate of  $v$  in the basis  $\{e_i\}_{i \in I}$ . So using this notation we have  $v = \sum_{i \in \{1, \dots, n\}} v_i \cdot e_i$

**Definition 10.155.** If  $S, I$  are sets then if  $\{x_i\}_{i \in I}$  is a family of elements in  $S$  and  $T \subseteq S$  then we say that  $\{x_i\}_{i \in I} \subseteq T$  if  $\{x_i | i \in I\} \subseteq T$

The next theorem says that every linear independent family that is part of a subset that spans the vector space can be extended to a basis of the vector space.

**Theorem 10.156.** Let  $V$  be a non trivial vector space over a field  $F$ ,  $W \subseteq V$  is such that  $\mathcal{S}(W) = V$ ,  $R$  a linear independent set (see 10.140) such that  $R \subseteq W$  then there exist a set  $B$  such that  $\{b\}_{b \in B}$  forms a basis of  $V$  and  $R \subseteq B \subseteq W$

**Proof.** Using the notation of the theorem define the following set of sets

$$\mathcal{A}(R, W) = \{X \subseteq V | R \subseteq X \subseteq W \text{ and } X \text{ is linear independent}\} \quad (10.24)$$

then using 2.137 we have that

$$\langle \mathcal{A}(R, W), \subseteq \rangle \text{ is a partial ordered set} \quad (10.25)$$

Assume now that  $\mathcal{C} \subseteq \mathcal{A}(R, W)$  is a non-empty chain (see 2.148) then if we define  $B_{\mathcal{C}} = \bigcup_{X \in \mathcal{C}} X$  we have

$$\forall X \in \mathcal{C} \text{ we have that } X \subseteq B_{\mathcal{C}} = \bigcup_{X \in \mathcal{C}} X \quad (10.26)$$

Now if  $X \in \mathcal{C} \subseteq \mathcal{A}(R, W)$  then by 10.24 we have  $R \subseteq X \subseteq W$  and thus using 10.26 we have

$$R \subseteq B_{\mathcal{C}} \subseteq W \quad (10.27)$$

We prove now the following assertion by mathematical induction

$$\text{If } A \subseteq B_{\mathcal{C}} \text{ is a finite subset of } B_{\mathcal{C}} \text{ then there exists a } C \in \mathcal{C} \text{ with } A \subseteq C \quad (10.28)$$

**Proof.** Let  $S_{\mathcal{C}} = \{n \in \mathbb{N} | \text{if } A \subseteq B_{\mathcal{C}} \text{ is such that } S_n \approx A \text{ then there exists a } C \in \mathcal{C} \text{ such that } A \subseteq C\}$  then we have:

1. If  $n = 0$  then if  $A \subseteq B_{\mathcal{C}}$  is such that  $\emptyset = S_0 \approx A \Rightarrow A = \emptyset$  as  $\mathcal{C}$  is not empty there exists a  $C \in \mathcal{C}$  and then  $A = \emptyset \subseteq C \Rightarrow 0 \in S_{\mathcal{C}}$

2. Assume that  $n \in S_{\mathcal{C}}$  if now  $A \subseteq B_{\mathcal{C}}$  is such that  $S_{n+1} \approx A$  then there exists a bijection  $b: \{0, \dots, n\} = S_{n+1} \rightarrow A$ , form then the bijection  $b|_{S_n}: S_n = \{0, \dots, n-1\} \rightarrow A \setminus \{b(n)\}$ , so as then  $S_n \approx A \setminus \{b(n)\} \subseteq B_{\mathcal{C}}$  we have as  $n \in S_{\mathcal{C}}$  that there exists a  $C' \in \mathcal{C}$  such that  $A \setminus \{b(n)\} \subseteq C'$ . Also as  $b(n) \in A \subseteq B_{\mathcal{C}} = \bigcup_{X \in \mathcal{C}} X$  there exists a  $C'' \in \mathcal{C}$  such that  $b(n) \in C''$ . Now as  $\mathcal{C}$  is a chain we have the following possibilities

- a.  $(C' \subseteq C'')$  then  $A \setminus \{b(n)\} \subseteq C' \subseteq C''$  and  $b(n) \in C'' \Rightarrow A = A \setminus \{b(n)\} \cup \{b(n)\} \subseteq C''$  so if we take  $C = C''$  we have  $A \subseteq C \in S_{\mathcal{C}}$
- b.  $(C'' \subseteq C')$  then  $A \setminus \{b(n)\} \subseteq C'$  and  $b(n) \in C'' \subseteq C'$  so  $A = A \setminus \{b(n)\} \cup \{b(n)\} \subseteq C'$ , take then  $C = C'$  and then we have  $A \subseteq C \in S_{\mathcal{C}}$

so in both cases we have found a  $C \in \mathcal{C}$  such that  $A \subseteq C$  proving that  $n+1 \in S_{\mathcal{C}}$ .

Using mathematical induction we have then that  $S_{\mathcal{C}} = \mathbb{N}$  proving that 10.28 is indeed true  $\square$

Now if  $\{\alpha_v\}_{v \in B_{\mathcal{C}}}$  is a family in  $F$  with finite support such that  $0 = \sum_{v \in B_{\mathcal{C}}} \alpha_v \cdot v$  then  $\text{support}(\{\alpha_v\}_{v \in B_{\mathcal{C}}}) \subseteq B_{\mathcal{C}}$  is finite and we can use 10.28 to prove that:

$$\exists Y \in \mathcal{C} \text{ such that } \text{support}(\{\alpha_v\}_{v \in B_{\mathcal{C}}}) \subseteq Y \quad (10.29)$$

As  $Y \in \mathcal{C} \subseteq \mathcal{A}(R, W)$  we have by 10.24 that  $Y$  is linear independent. Define now  $\{\beta_v\}_{v \in Y}$  by  $\beta_v = \begin{cases} \alpha_v & \text{if } v \in B_{\mathcal{C}} \\ 0 & \text{if } v \in Y \setminus B_{\mathcal{C}} \end{cases}$  then we have trivially that  $\text{support}(\{\alpha_v\}_{v \in B_{\mathcal{C}}}) = \text{support}(\{\beta_v\}_{v \in Y})$  so that  $\{\beta_v\}_{v \in Y}$  has finite support. Also  $\sum_{v \in Y} \beta_v \cdot v = \sum_{v \in Y \setminus B_{\mathcal{C}}} \beta_v \cdot v + \sum_{v \in B_{\mathcal{C}}} \beta_v \cdot v = \sum_{v \in Y \setminus B_{\mathcal{C}}} 0 \cdot v + \sum_{v \in B_{\mathcal{C}}} \alpha_v \cdot v = 0$  so by linear independence of  $Y$  that  $\forall v \in Y$  we have  $\beta_v = 0$  and as  $B_{\mathcal{C}} \subseteq Y$  we have  $\forall v \in B_{\mathcal{C}}$  we have  $\alpha_v = \beta_v = 0$  proving that

$$B_{\mathcal{C}} \text{ is linear independent} \quad (10.30)$$

Using 10.30, 10.27 and the definition of  $\mathcal{A}(R, W)$  (see 10.24) we have

$$B_{\mathcal{C}} \in \mathcal{A}(R, W) \quad (10.31)$$

Using 10.31 together with 10.26 let's us finally conclude that

$$\text{If } \mathcal{C} \text{ is a non empty chain in } \mathcal{A}(R, W) \text{ then it is bounded above by } B_{\mathcal{C}} \quad (10.32)$$

If  $\mathcal{C} = \emptyset$  is the empty chain then as  $R \in \mathcal{A}(R, W)$  we have vacuously that  $\forall C \in \mathcal{C}$  we have  $C \subseteq R$  where  $R \in \mathcal{A}(R, W)$  so  $R$  is a upper bound of  $\mathcal{C}$ . This together with 10.32 gives us

$$\text{Every chain in } \mathcal{A}(R, W) \text{ has a upper bound in } \mathcal{A}(R, W) \quad (10.33)$$

Using Zorn's lemma (see 2.225) we have that

$$\exists M \in \mathcal{A}(R, W) \text{ such that } \forall X \in \mathcal{A}(R, W) \text{ we have } X \subseteq M \quad (10.34)$$

and by 10.24 we have also

$$R \subseteq M \subseteq W \text{ and } M \text{ is linear independent} \quad (10.35)$$

Let now  $w \in W$  then we have the following cases to consider:

1.  $(w \in M)$  then as  $M \subseteq S(M)$  (see 10.135) we have  $w \in S(M)$ .
2.  $(w \notin M)$  then  $\{w\} \cup M$  must be linear dependent [for if it is linear dependent then as  $R \subseteq M \subseteq M \cup \{w\} \subseteq W$  we would have  $M \cup \{w\}$  contradicting the maximality of  $M$ ]. Hence there exists a  $\{\alpha_v\}_{v \in M \cup \{w\}}$  in  $F$  with finite support such that  $\sum_{v \in M \cup \{w\}} \alpha_v \cdot v = 0$  and there exists a  $v_0 \in M \cup \{w\}$  such that  $\alpha_{v_0} \neq 0$ . Now  $\text{support}(\{\alpha_v\}_{v \in M \cup \{w\}}) = \{v \in M \cup \{w\} \mid \alpha_v \neq 0\} = \{v \in M \mid \alpha_v \neq 0\} \cup \{w \in \{w\} \mid \alpha_w \neq 0\} = \text{support}(\{\alpha_v\}_{v \in M}) \cup \{w \in \{w\} \mid \alpha_w \neq 0\}$ . If now  $\alpha_w = 0$  then  $\text{support}(\{\alpha_v\}_{v \in M \cup \{w\}}) = \text{support}(\{\alpha_v\}_{v \in M})$  and as  $0 = \sum_{v \in M \cup \{w\}} \alpha_v \cdot w = \sum_{v \in M} \alpha_v \cdot v + \sum_{v \in \{w\}} \alpha_v \cdot v = \sum_{v \in M} \alpha_v \cdot v + 0 \cdot w = \sum_{v \in M} \alpha_v \cdot v$  and from independence from  $M$  we would then have that  $\alpha_{v_0} = 0$  a contradiction. So we must conclude that  $\alpha_w \neq 0$  and then from  $0 = \sum_{v \in M \cup \{w\}} \alpha_v \cdot v = \sum_{v \in M} \alpha_v \cdot v + \sum_{v \in \{w\}} \alpha_v \cdot v = \sum_{v \in M} \alpha_v \cdot v + \alpha_w \cdot w \Rightarrow w = (-\alpha_w^{-1}) \cdot \sum_{v \in M} \alpha_v \cdot v = \sum_{v \in M} (-\alpha_w^{-1} \cdot \alpha_v) \cdot v$  proving again that  $w \in S(M)$

From the above we conclude that  $W \subseteq \mathcal{S}(M)$ . Using 10.149 we have  $V = \mathcal{S}(W) \subseteq \mathcal{S}(\mathcal{S}(M)) \stackrel{10.137}{=} \mathcal{S}(M) \subseteq V \Rightarrow \mathcal{S}(M) = V$  which together with 10.35 means that  $\{m\}_{m \in M}$  is basis for  $V$  and  $R \subseteq M \subseteq W$ .  $\square$

**Corollary 10.157.** *Let  $V$  be a non trivial vector space over a field  $F$  and let  $W \subseteq V$  be such that  $\mathcal{S}(W) = V$  then there exist a set  $B$  with  $B \subseteq W$  such that  $\{b\}_{b \in B}$  forms a basis of  $V$ . As  $\mathcal{S}(V) = V$  we see directly that every non trivial vector space has a basis.*

**Proof.** As  $V \neq \{0\}$  there exists a  $v_0 \in V$  such that  $v_0 \neq 0$ , as  $V = \mathcal{S}(W)$  there exist a  $\{\alpha_v\}_{v \in W}$  a family in  $F$  with finite support so that  $v_0 = \sum_{w \in W} \alpha_w \cdot w$ . Now if  $W = \{0\}$  or  $W = \emptyset$  then  $\text{support}(\{\alpha_w \cdot w\}_{w \in W}) = \emptyset$  and thus  $v_0 = \sum_{w \in W} \alpha_w \cdot w = 0$  a contradiction so there must exists a  $w \in W$  with  $w \neq 0$ . If now there exists a  $\{\alpha_v\}_{v \in \{w\}}$  such that  $0 = \sum_{v \in \{w\}} \alpha_v \cdot v = \alpha_w \cdot w \Rightarrow \alpha_w = 0 \Rightarrow \forall v \in \{w\}$  we have  $\alpha_w = 0$  proving that  $\{w\} \subseteq W$  is a linear independent set. Using the previous theorem we have then to find a linear independent set  $B$  with  $\{w\} \subseteq B \subseteq W$  such that  $\{b\}_{b \in B}$  is a basis of  $V$ .  $\square$

**Corollary 10.158.** *Let  $V$  be a non trivial vector space over a field  $F$  and let  $\{v_i\}_{i \in I}$  be a family in  $V$  such that  $\mathcal{S}(\{v_i | i \in I\}) = V$  then there exists a  $J \subseteq I$  such that  $\{v_i\}_{i \in J}$  is a basis of  $V$  [equivalent there exists a  $\{w_i\}_{i \in J}$  such that  $\{w_i\}_{i \in J}$  is a basis and  $\{w_i\}_{i \in J} \preccurlyeq \{v_i\}_{i \in I}$ ].*

**Proof.** First as  $v = \{v_i\}_{i \in I}$  is a family of elements in  $V$  we have by definition of a family of elements (see 2.104) that it is a function  $v: I \rightarrow V$  with  $i \rightarrow v(i) = v_i$ . We can use 2.234 to find a subset  $J' \subseteq I$  so that  $v|_{J'}: J' \rightarrow v(I) = \{v_i | i \in I\}$  is a bijection. Now as  $\mathcal{S}(\{v_i | i \in I\}) = V$  we can use 10.157 to find a set  $B$  with  $B \subseteq \{v_i | i \in I\}$  such that  $\{b\}_{b \in B}$  forms a basis for  $V$ . Take now  $J = (v|_{J'})^{-1}(B) \subseteq J'$  then  $(v|_{J'})|_J = v|_J: J \rightarrow B$  forms a bijection such that  $\forall j \in J$  we have  $v|_J(j) = v(j) = v_j$  and defines the family  $\{v_i\}_{i \in J} \preccurlyeq \{v_i\}_{i \in I}$ . We prove now that  $\{v_i\}_{i \in J}$  forms a basis. Now if  $\{\alpha_i\}_{i \in J}$  is a family in  $F$  with finite support such that  $0 = \sum_{i \in J} \alpha_i \cdot v_i$  define then  $\{\beta_b\}_{b \in B}$  such that  $\beta_b = \alpha_{(v|_J)^{-1}(b)}$  then  $\beta_{v_i} = \beta_{v|_J(i)} = \alpha_{(v|_J)^{-1}(v|_J(i))} = \alpha_i$  and thus  $0 = \sum_{i \in J} \beta_{v_i} \cdot v_i \stackrel{v|_J: J \rightarrow B \text{ is bijective and 10.43}}{=} \sum_{b \in B} \beta_b \cdot b$  and as  $\{b\}_{b \in B}$  is linear independent we have  $\forall b \in B$  that  $\beta_b = 0$ , now if  $i \in J$  then  $\alpha_i = \beta_{v_i} = 0$  so  $\forall i \in J$  we have  $\alpha_i = 0$  proving that  $\{v_i\}_{i \in J}$  is linear independent. As also  $\{v_i | i \in J\} = v|_J(J) = B$  we have  $\mathcal{S}(\{v_i | i \in J\}) = \mathcal{S}(B) = V$  and all conditions are fulfilled to make  $\{v_i\}_{i \in J}$  a basis of  $V$ .  $\square$

**Corollary 10.159.** *Let  $V$  be a non trivial vector space over a field  $F$  then if  $W \subseteq V$  is a linear independent subset of  $V$  then there exists a set  $B$  with  $W \subseteq B \subseteq F$  such that  $\{b\}_{b \in B}$  is a basis of  $V$ .*

**Proof.** As  $\mathcal{S}(V) = V$  and  $W \subseteq V$  we can use 10.156 to find a set  $W \subseteq B \subseteq F$  such that  $\{b\}_{b \in B}$  is a basis for  $V$ .  $\square$

The following is a useful theorem if we have a finite dimensional basis to express it as a family based on the indexing set  $\{1, \dots, n\}$

**Theorem 10.160.** *If  $V$  is a vector space over a field  $F$  and  $B \subseteq V$  a set such that  $\{b\}_{b \in B}$  is a basis of  $V$  and  $\#(B) = n \in \mathbb{N}_0$  (so  $B$  is finite) then there exists a basis  $\{v_i\}_{i \in \{1, \dots, n\}}$  in  $V$*

**Proof.** If  $\#(B) = n > 0$  then there exists a bijection  $v: \{1, \dots, n\} \rightarrow B$  which defines the family  $\{v_i\}_{i \in \{1, \dots, n\}}$ . If now  $\{\alpha_i\}_{i \in \{1, \dots, n\}}$  is a family in  $F$  with finite support such that  $\sum_{i \in \{1, \dots, n\}} \alpha_i \cdot v_i = 0$ , define then  $\{\beta_b\}_{b \in B}$  by  $\beta_b = \alpha_{v^{-1}(b)}$  then we have  $\beta_{v_i} = \beta_{v(i)} = \alpha_{v^{-1}(v(i))} = \alpha_i$  and thus  $0 = \sum_{i \in \{1, \dots, n\}} \beta_{v_i} \cdot v_i \stackrel{\text{10.43 and } v \text{ is a bijection}}{=} \sum_{b \in B} \beta_b \cdot b$ . By linear independence of  $\{b\}_{b \in B}$  we have  $\forall b \in B$  that  $\beta_b = 0$ . If now  $i \in \{1, \dots, n\}$  then  $\alpha_i = \beta_{v(i)} = 0$  so we have proved that  $\{v_i\}_{i \in \{1, \dots, n\}}$  is linear independent. Finally  $\mathcal{S}(\{v_i | i \in \{1, \dots, n\}\}) = \mathcal{S}(B) = V$  so  $\{v_i\}_{i \in \{1, \dots, n\}}$  spans  $V$ .  $\square$

**Definition 10.161. (Kronicker Delta)** *Let  $n \in \mathbb{N}_0$  and  $F$  a field with unit element 1 and neutral element 0 then we define  $\delta_n: \{1, \dots, n\} \times \{1, \dots, n\} \rightarrow F$  by  $(i, j) \rightarrow \delta_n(i, j) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$ . If  $n$  is understood to be known then we write just  $\delta_{i,j}$  instead of  $\delta_n(i, j)$ .*

**Theorem 10.162.** *If  $n \in \mathbb{N}_0$ ,  $F$  is field and  $\{\alpha_i\}_{i \in \{1, \dots, n\}}$  a family in  $F$  then  $\sum_{i=1}^n \delta_{i,j} \cdot \alpha_i = \alpha_j$*

**Proof.**  $\sum_{i=1}^n \delta_{i,j} \cdot \alpha_i = \sum_{i \in \{1, \dots, n\}} \delta_{i,j} \cdot \alpha_i = \sum_{i \in \{1, \dots, n\} \setminus \{j\}} \delta_{i,j} \cdot \alpha_i + \sum_{i \in \{j\}} \delta_{i,j} \cdot \alpha_i = \sum_{i \in \{1, \dots, n\} \setminus \{j\}} 0 \cdot \alpha_i + \delta_{j,j} \cdot \alpha_j = 0 + 1 \cdot \alpha_j = \alpha_j$   $\square$

**Definition 10.163.** If  $F$  is a field,  $n \in \mathbb{N}_0$  and  $F^n$  the vector space defined over  $F$  (see 10.118 for the definition) define then the family  $\{\mathcal{E}_i\}_{i \in \{1, \dots, n\}}$  in  $F^n$  by  $(\mathcal{E}_i)_j = \mathcal{E}_i(j) = \delta_{i,j} \forall i, j \in \{1, \dots, n\}$

**Theorem 10.164.** If  $F$  is a field  $n \in \mathbb{N}_0$  and  $F^n$  the vector space over  $F$  defined in 10.118 then  $\{\mathcal{E}_i\}_{i \in \{1, \dots, n\}}$  is a basis of  $F^n$

**Proof.** If  $x = (x_1, \dots, x_n) \in F^n$  then  $x = \{x_i\}_{i \in \{1, \dots, n\}}$  is a family in  $F$  and  $\forall i \in \{1, \dots, n\}$  we have  $(\sum_{j \in \{1, \dots, n\}} x_j \cdot \mathcal{E}_j)_i = (\sum_{j=1}^n x_j \cdot \mathcal{E}_j)_i \stackrel{10.10}{=} \sum_{j=1}^n (x_j \cdot \mathcal{E}_j)_i = \sum_{j=1}^n x_j \cdot (\mathcal{E}_j)_i = \sum_{j=1}^n x_j \cdot \delta_{j,i} \stackrel{10.162}{=} x_i$  so we have  $x = \sum_{j \in \{1, \dots, n\}} x_j \cdot \mathcal{E}_j$  and thus we have  $\mathcal{S}(\{\mathcal{E}_i | i \in \{1, \dots, n\}\}) = F^n$ .

Now to prove linear independence assume that there exists a  $\{\alpha_j\}_{j \in \{1, \dots, n\}}$  such that  $0 = (0, \dots, 0) = \sum_{j \in \{1, \dots, n\}} \alpha_j \cdot \mathcal{E}_j$  then we have  $0 = (\sum_{j \in \{1, \dots, n\}} \alpha_j \cdot \mathcal{E}_j)_i = (\sum_{j=1}^n \alpha_j \cdot \mathcal{E}_j)_i \stackrel{10.10}{=} \sum_{j=1}^n (\alpha_j \cdot \mathcal{E}_j)_i = \sum_{j=1}^n \alpha_j \cdot (\mathcal{E}_j)_i = \sum_{j=1}^n \alpha_j \cdot \delta_{j,i} = \alpha_i$  proving that  $\forall i \in \{1, \dots, n\}$  we have  $\alpha_i = 0$   $\square$

Applying the above to the fields  $\mathbb{R}$  and  $\mathbb{C}$  we have the following corollaries:

**Corollary 10.165.**  $\{\mathcal{E}_i\}_{i \in \{1, \dots, n\}}$  is a basis of the vector space  $\mathbb{R}^n$  over the field  $\mathbb{R}$

**Corollary 10.166.**  $\{\mathcal{E}_i\}_{i \in \{1, \dots, n\}}$  is a basis of the vector space  $\mathbb{C}^n$  over the field  $\mathbb{C}$

We proceede now to construct the basis for a product of finite vector spaces based of these vector spaces. First we need a little lemma.

**Lemma 10.167.** Let  $n \in \mathbb{N}_0$ ,  $\{n_i\}_{i \in \{1, \dots, n\}} \subseteq \mathbb{N}_0$  then if we define  $\{N_i\}_{i \in \{1, \dots, n\}}$  by  $N_i = \{(i, j) | j \in \{1, \dots, n_i\}\}$  we have

1. The function  $\beta: \bigcup_{i \in \{1, \dots, n\}} N_i \rightarrow \{1, \dots, \sum_{i=1}^n n_i\}$  by  $\beta(i, j) = (\sum_{k=1}^{i-1} n_k) + j$  is a bijection
2. The inverse bijection  $\tau = \beta^{-1}: \{1, \dots, \sum_{i=1}^n n_i\} \rightarrow \bigcup_{i \in \{1, \dots, n\}} N$  satisfies  $\forall k \in \{1, \dots, \sum_{i=1}^n n_i\}$  that  $\tau(k)_1 \in \{1, \dots, n\}$ ,  $\tau(k) \in \{1, \dots, n_{\tau(k)_1}\}$  and  $\forall l \in \{1, \dots, n\}$  we have if  $i \in \tau^{-1}(N_l)$  that  $\tau(i)_1 = l$
3.  $\forall k \in \{1, \dots, n\}$  we have that the function  $\gamma_k: \tau^{-1}(N_k) \rightarrow \{1, \dots, n_k\}$  defined by  $\gamma_k(i) = \tau(i)_2$  is a bijection.

**Proof.**

1. If  $(i, j) \in \bigcup_{k \in \{1, \dots, n\}} N_k$  then  $\exists k \in \{1, \dots, n\}$  such that  $(i, j) \in N_k$  hence  $(i, j) \in \{(k, l) | l \in \{1, \dots, n_k\}\}$  proving that  $i = k$  and  $j \in \{1, \dots, n_k\} = \{1, \dots, n_i\}$ . This gives

$$\forall (i, j) \in \bigcup_{k \in \{1, \dots, n\}} N_k \Rightarrow j \in \{1, \dots, n_i\} \text{ and } i \in \{1, \dots, n\} \quad (10.36)$$

**injectivity.** Let  $(i, j), (r, s) \in \bigcup_{k \in \{1, \dots, n\}} N_k$  be such that  $\beta(i, j) = \beta(r, s)$  then we have  $(\sum_{k=1}^{i-1} n_k) + j = (\sum_{k=1}^{r-1} n_k) + s$ . Assume that  $i \neq r$  then we may assume without loosing generality that  $i < r$  (otherwise exchange  $i$  with  $r$ ). Using 10.36 we have that  $j \in \{1, \dots, n_i\}$  hence  $(\sum_{k=1}^{i-1} n_k) + j \leq (\sum_{k=1}^{i-1} n_k) + n_i = \sum_{k=1}^i n_k \leq i < r \Rightarrow i \leq r-1 \sum_{k=1}^{r-1} n_k < (\sum_{k=1}^{r-1} n_k) + s$  proving that  $(\sum_{k=1}^{i-1} n_k) + j < (\sum_{k=1}^{r-1} n_k) + s$  in contradiction with  $(\sum_{k=1}^{i-1} n_k) + j = (\sum_{k=1}^{r-1} n_k) + s$ . So we must have that  $i = r$  and thus  $(\sum_{k=1}^{i-1} n_k) + j = (\sum_{k=1}^{i-1} n_k) + s \Rightarrow j = s$ , hence  $(i, j) = (r, s)$  proving injectivity.

**surjectiviy.** If  $i \in \{1, \dots, \sum_{k=1}^n n_k\}$  then as  $\sum_{i=1}^{1-1} n_k = \sum_{i=1}^0 n_k = 0 < i$  we have that  $\mathcal{S} = \{k \in \{1, \dots, n\} | \sum_{l=1}^{k-1} n_l < i\} \neq \emptyset$  hence  $i_m = \max(\mathcal{S})$  exists. We have now two possible cases for  $i_m$ :

**$i_m = n$ .** hence  $\sum_{l=1}^{n-1} n_l < i \leq \sum_{l=1}^n n_l$  then  $0 < i - \sum_{l=1}^{n-1} n_l \leq \sum_{l=1}^n n_l - \sum_{l=1}^{n-1} n_l = n_n$  so that if we take  $s = i - \sum_{l=1}^{n-1} n_l$  we have that  $s \in \{1, \dots, n_n\}$  and  $i = \sum_{l=1}^{n-1} n_l + s$  proving that  $(i, s) \in N_n$  and  $\beta(s) = i$

$i_m < n$ . hence  $i_m + 1 \notin \mathcal{S}$  so that  $\sum_{l=1}^{i_m-1} n_l < i \leq \sum_{l=1}^{(i_m+1)-1} n_l = \sum_{l=1}^{i_m} n_l$  then  $0 < i - \sum_{l=1}^{i_m-1} n_l \leq \sum_{l=1}^{i_m} n_l - \sum_{l=1}^{i_m-1} n_l = n_{i_m}$  so that if we take  $s = i - \sum_{l=1}^{i_m-1} n_l$  we have that  $s \in \{1, \dots, n_{i_m}\}$  and  $i = \sum_{l=1}^{i_m-1} n_l + s$  proving that  $(i, s) \in N_{i_m}$  and  $\beta(s) = i$

2. If  $k \in \{1, \dots, \sum_{i=1}^n N_i\}$  then  $\tau(k) = (\tau(k)_1, \tau(k)_2) \in \bigcup_{k \in \{1, \dots, n\}} N_k$  hence using 10.36 we have that  $\tau(k)_1 \in \{1, \dots, n\}$  and  $\tau(k)_2 \in \{1, \dots, n_{\tau(k)_1}\}$ . If now  $l \in \{1, \dots, n\}$  and  $i \in \tau^{-1}(N_l)$  then  $(\tau(i)_1, \tau(i)_2) \in N_l = \{(l, i) | i \in \{1, \dots, n_l\}\}$  so that  $\tau(i)_1 = l$ .
3. Let  $k \in \{1, \dots, n\}$  then we have

**injectivity.** If  $i, j \in \tau^{-1}(N_k)$  and  $\gamma_k(i) = \gamma_k(j)$  then  $\tau_k(i)_2 = \tau_k(j)_1$  and as by (2) we have that  $\tau(i)_1 = k = \tau(j)_1$  we conclude that  $\tau(i) = \tau(j)$  hence  $i = j$ .

**surjectivity.** Let  $l \in \{1, \dots, n_k\}$  then  $(k, l) \in N_k$  so as  $\tau$  is a bijection there exists a  $i \in \{1, \dots, \sum_{j=1}^n n_j\}$  such that  $\tau(i) = (k, l) \in N_k$ . By definition we have then that  $i \in \tau^{-1}(N_k)$  and  $\gamma_k(i) = \tau(i)_2 = l \in \{1, \dots, n_k\}$   $\square$

**Theorem 10.168.** Let  $n \in \mathbb{N}_0$ ,  $\{V_i\}_{i \in \{1, \dots, n\}}$  be a finite family of non trivial vector spaces over a field  $F$  with  $\forall i \in \{1, \dots, n\}$   $\{e_{i,j}\}_{j \in \{1, \dots, n_i\}}$  a basis for  $V_i$ . Given  $k \in \{1, \dots, \sum_{i=1}^n n_i\}$  define  $f_k = ((f_k)_1, \dots, (f_k)_n) \in \prod_{i \in \{1, \dots, n\}} V_i$  by  $(f_k)_i = e_{\tau(k)_1, \tau(k)_2}$  if  $\tau(k)_1 = i$  and  $(f_k)_i = 0$  if  $\tau(k)_1 \in \{1, \dots, n\} \setminus \{i\}$  (here  $\tau: \{1, \dots, \sum_{i=1}^n n_i\} \rightarrow \bigcup_{i \in \{1, \dots, n\}} N_i$  is defined in the previous lemma, which also ensures that  $\tau(k)_1 \in \{1, \dots, n\}$  and  $\tau(k)_2 \in \{1, \dots, n_{\tau(k)_1}\}$  as it should). Then  $\{f_k\}_{k \in \{1, \dots, \sum_{i=1}^n n_i\}}$  is a basis of  $\prod_{i \in \{1, \dots, n\}} V_i$  (which is a vector space by 10.120)

**Proof.** Define  $N = \sum_{i=1}^n n_i$ . Then if  $k \in \{1, \dots, n\}$  and  $l \in \tau^{-1}(N_k)$  then  $\tau(l) \in N_k$  and thus  $\tau(l)_1 = k$  and  $\tau(l)_2 \in \{1, \dots, n_k\}$  so that  $(f_l)_k = e_{k, \tau(l)_2} \in \{e_{k,j} | j \in \{1, \dots, n_k\}\}$ . If  $l \in \{1, \dots, N\} \setminus \tau^{-1}(N_k)$  then  $\tau(l)_1 \neq k$  [if  $\tau(l)_1 = k$  then by the previous lemma we have that  $\tau(l)_2 \in \{1, \dots, n_k\}$  hence  $\tau(l) \in N_k \Rightarrow l \in \tau^{-1}(k)$  a contradiction], so  $(f_l)_k = 0$ . To summarize we have proved that

$$\forall k \in \{1, \dots, n\} \models \text{if } l \in \tau^{-1}(N_k) \text{ then } (f_l)_k = e_{k, \tau(l)_2} \in \{e_{k,j} | j \in \{1, \dots, n_k\}\} \text{ and if } l \in \{1, \dots, N\} \setminus \tau^{-1}(N_k) \text{ then } (f_l)_k = 0 \quad (10.37)$$

Let now  $\{\alpha_i\}_{i \in \{1, \dots, N\}} \subseteq F$  then we have given  $k \in \{1, \dots, n\}$  that

$$\begin{aligned} \left( \sum_{i=1}^N \alpha_i \cdot f_i \right)_k &= \sum_{i=1}^N \alpha_i \cdot (f_i)_k \\ &= \sum_{i \in \{1, \dots, N\}} \alpha_i \cdot (f_i)_k \\ &= \sum_{i \in \{1, \dots, N\} \setminus \tau^{-1}(N_k)} \alpha_i \cdot (f_i)_k + \sum_{i \in \tau^{-1}(N_k)} \alpha_i \cdot (f_i)_k \\ &\stackrel{10.37}{=} 0 + \sum_{i \in \tau^{-1}(N_k)} \alpha_i \cdot e_{k, \tau(i)_2} \\ &= \sum_{i \in \tau^{-1}(N_k)} \alpha_i \cdot e_{k, \tau(i)_2} \end{aligned}$$

proving that

$$\forall k \in \{1, \dots, n\} \models \left( \sum_{i=1}^N \alpha_i \cdot f_i \right)_k = \sum_{i \in \tau^{-1}(N_k)} \alpha_i \cdot e_{k, \tau(i)_2} \quad (10.38)$$

If now  $\{\alpha_i\}_{i \in \{1, \dots, N\}} \subseteq F$  is such that  $0 = \sum_{i=1}^N \alpha_i \cdot f_i$  we have by the above that  $\forall k \in \{1, \dots, n\}$  that  $0 = \sum_{i \in \tau^{-1}(N_k)} \alpha_i \cdot e_{k, \tau(i)_2}$ . As  $\{e_{k,j}\}_{j \in \{1, \dots, n_k\}}$  is a basis we have using 10.37 that  $\forall i \in \tau^{-1}(N_k)$   $\alpha_i = 0$ . If  $m \in \{1, \dots, N\}$  we have that  $\tau(m) \in \bigcup_{k \in \{1, \dots, n\}} N_k$  so that there exists a  $k \in \{1, \dots, n\}$  such that  $\tau(m) \in N_k \Rightarrow m \in \tau^{-1}(N_k)$  hence  $\alpha_m = 0$ . So we proved that  $\forall m \in \{1, \dots, N\}$  we have  $\alpha_m = 0$  and thus that

$$\{f_i\}_{i \in \{1, \dots, N\}} \text{ is linear independent} \quad (10.39)$$

To prove completeness let  $x = (x_1, \dots, x_n) \in \prod_{i \in \{1, \dots, n\}} V_i$  then we have  $\forall k \in \{1, \dots, n\}$  that there exists a  $\{\alpha_{k,i}\}_{i \in \{1, \dots, n_k\}}$  such that  $x_k = \sum_{i=1}^{n_k} \alpha_{k,i} \cdot e_{k,i}$ . Define now  $\{\beta_k\}_{k \in \{1, \dots, N\}}$  by  $\beta_k = \alpha_{\tau(k)_1, \tau(k)_2}$ . We have then  $\forall k \in \{1, \dots, n\}$  that if  $i \in \tau^{-1}(N_k)$  that  $(\tau(i)_1, \tau(i)_2) = \tau(i) \in N_k \Rightarrow \tau(i)_1 = k \wedge \tau(i)_2 \in \{1, \dots, n_k\}$  so that  $\beta_i = \alpha_{\tau(i)_1, \tau(i)_2} = \alpha_{k, \tau(i)_2}$  giving

$$\forall k \in \{1, \dots, n\} \text{ we have if } i \in \tau^{-1}(N_k) \text{ that } \beta_k = \alpha_{k, \tau(i)_2} \quad (10.40)$$

hence we have  $\forall k \in \{1, \dots, n\}$  that

$$\begin{aligned} \left( \sum_{i=1}^N \beta_i \cdot f_i \right)_k &\stackrel{10.38}{=} \sum_{i \in \tau^{-1}(N_k)} \beta_i \cdot e_{k, \tau(i)_2} \\ &\stackrel{10.40}{=} \sum_{i \in \tau^{-1}(N_k)} \alpha_{k, \tau(i)_2} \cdot e_{k, \tau(i)_2} \end{aligned}$$

If we use now the bijection  $\gamma_k: \tau^{-1}(N_k) \rightarrow \{1, \dots, n_k\}$  defined by  $\gamma_k(i) = \tau(i)_2$  (see the previous lemma 10.167 (3)) then we have that  $(\sum_{i=1}^N \beta_i \cdot f_i)_k = \sum_{i \in \tau^{-1}(N_k)} \alpha_{k, \gamma_k(i)} \cdot e_{k, \gamma_k(i)} \stackrel{10.43}{=} \sum_{i \in \{1, \dots, n_k\}} \alpha_{k, i} \cdot e_{k, i} = x_k$ . So we have proved that  $\forall k \in \{1, \dots, n\}$  we have that  $(\sum_{i=1}^N \beta_i \cdot f_i)_k = x_k$  where  $x = (x_1, \dots, x_n) \in \prod_{i \in \{1, \dots, n\}} V_i$  hence  $x = \sum_{i=1}^N \beta_i \cdot f_i$  proving that  $\{f_i\}_{i \in \{1, \dots, N\}}$  is covering  $\prod_{i \in \{1, \dots, n\}} V_i$ .  $\square$

## 10.5 Free vector space over a set

In this section we show that given a non empty set we can construct a vector space that has this set as a basis.

**Definition 10.169.** Given a non empty set  $X$  and a field  $\langle F, +, \cdot \rangle$  define then  $\mathcal{F}(X, F) = \{f: X \rightarrow \mathbb{K} \mid f^{-1}(F \setminus \{0\}) \text{ is finite}\}$ . In other words  $\mathcal{F}(X, F)$  is the set of functions from  $X$  to  $F$  that are nonzero only for a finite number of points in  $X$ .

**Definition 10.170.** Given a set  $X$  and a field  $F$  then we define

1. If  $f, g \in \mathcal{F}(X, F)$  then  $f + g: X \rightarrow F$  is defined by  $x \mapsto f(x) + g(x)$
2. If  $\alpha \in F$  and  $f \in \mathcal{F}(X, F)$  then  $\alpha \cdot f: X \rightarrow F$  is defined by  $x \mapsto \alpha \cdot f(x)$
3.  $\forall a \in X$  we define  $\delta_a \in \mathcal{F}(X, F)$  by  $\delta_a(x) = \begin{cases} 1 & \text{if } x = a \\ 0 & \text{if } x \in X \setminus \{a\} \end{cases}$

**Theorem 10.171.** Given a non empty set  $X$  and a field  $\langle F, +, \cdot \rangle$  then  $\forall f, g \in \mathcal{F}(X, F)$  and  $\alpha \in F$  we have that  $\langle \mathcal{F}(X, F), +, \cdot \rangle$  is a vector space over  $F$  (see previous definition for  $+, \cdot$ ) with basis  $\{\delta_a\}_{a \in X}$  and the zero element  $0: X \rightarrow F$  defined by  $x \mapsto 0$ . This vector space is called the **free vector space over a set**. The set  $X$  is embedded in  $\mathcal{F}(X, F)$  by the injection  $\delta: X \rightarrow \mathcal{F}(X, F)$  defined by  $a \mapsto \delta_a$  if  $\langle F, +, \cdot \rangle$  has characterization 0.

**Proof.** If  $f, g \in \mathcal{F}(X, F)$  then if  $x \in (f + g)^{-1}(F \setminus \{0\})$  we have  $(f + g)(x) \neq 0 \Rightarrow f(x) + g(x) \neq 0 \Rightarrow f(x) \neq 0 \vee g(x) \neq 0 \Rightarrow x \in f^{-1}(F \setminus \{0\}) \cup g^{-1}(F \setminus \{0\}) \Rightarrow (f + g)^{-1}(F \setminus \{0\}) \subseteq f^{-1}(F \setminus \{0\}) \cup g^{-1}(F \setminus \{0\})$  which means as  $f^{-1}(F \setminus \{0\})$ ,  $g^{-1}(F \setminus \{0\})$  are finite that  $(f + g)^{-1}(F \setminus \{0\})$  is finite and thus that  $f + g \in \mathcal{F}(X, F)$  proving that  $+: \mathcal{F}(X, F) \times \mathcal{F}(X, F) \rightarrow \mathcal{F}(X, F)$  is well defined.

Also if  $\alpha \in F$  and  $f \in \mathcal{F}(X, F)$  then if  $x \in (\alpha \cdot f)^{-1}(F \setminus \{0\})$  we have  $\alpha \cdot f(x) \neq 0 \Rightarrow f(x) \neq 0 \Rightarrow x \in f^{-1}(F \setminus \{0\}) \Rightarrow (\alpha \cdot f)^{-1}(F \setminus \{0\}) \subseteq f^{-1}(F \setminus \{0\})$  which as  $f^{-1}(F \setminus \{0\})$  is finite proves that  $\alpha \cdot f \in \mathcal{F}(X, F)$  and thus that  $\cdot: F \times \mathcal{F}(X, F) \rightarrow \mathcal{F}(X, F)$  is well defined. Now to prove that it is a vector space:

1.  $\langle \mathcal{F}(X, F), + \rangle$  is a Abelian group
  - a. **(Associativity)** If  $f, g, h \in \mathcal{F}(X, F)$  then  $\forall x \in X$  we have  $(f + (g + h))(x) = f(x) + (g + h)(x) = f(x) + (g(x) + h(x)) = (f(x) + g(x)) + h(x) = (f + g)(x) + h(x) = ((f + g) + h)(x)$  proving that  $f + (g + h) = (f + g) + h$
  - b. **(Commutativity)** If  $f, g \in \mathcal{F}(X, F)$  then  $\forall x \in X$  we have  $(f + g)(x) = f(x) + g(x) = g(x) + f(x) = (g + f)(x)$  proving that  $f + g = g + f$

- c. **(Neutral Element)** If  $f \in \mathcal{F}(X, F)$  then  $(f + 0)(x) = f(x) + 0(x) = f(x)$  giving  $f + 0 = f$   $\underset{\text{commutativity}}{=} 0 + f$
- d. **(Inverse Element)** If  $f \in \mathcal{F}(X, F)$  define then  $-f: X \rightarrow F$  by  $x \mapsto -f(x)$  which as  $x \in (-f)^{-1}(F \setminus \{0\})$  then  $(-f)(x) \neq 0 \Rightarrow -f(x) \neq 0 \Rightarrow f(x) \neq 0 \Rightarrow x \in f^{-1}(F \setminus \{0\}) \Rightarrow (-f)^{-1}(F \setminus \{0\}) \subseteq f^{-1}(F \setminus \{0\})$  which is finite proving that  $-f \in \mathcal{F}(X, F)$ . Now  $\forall x \in X$  we have  $(f + (-f))(x) = f(x) + (-f(x)) = 0 = 0(x)$  proving that  $f + (-f) = 0$   $\underset{\text{commutativity}}{=} (-f) + f = 0$

## 2. Vector spaces axioms

- a. If  $\alpha \in F$  and  $f, g \in \mathcal{F}(X, F)$  then  $\forall x \in X$  we have  $(\alpha \cdot (f + g))(x) = \alpha \cdot (f + g)(x) = \alpha \cdot (f(x) + g(x)) = \alpha \cdot f(x) + \alpha \cdot g(x) = (\alpha \cdot f)(x) + (\alpha \cdot g)(x) = ((\alpha \cdot f) + (\alpha \cdot g))(x)$  proving that  $\alpha \cdot (f + g) = \alpha \cdot f + \alpha \cdot g$
- b. If  $\alpha, \beta \in F$  and  $f \in \mathcal{F}(X, F)$  then  $\forall x \in X$  we have  $((\alpha + \beta) \cdot f)(x) = (\alpha + \beta) \cdot f(x) = \alpha \cdot f(x) + \beta \cdot f(x) = ((\alpha \cdot f)(x) + (\beta \cdot f)(x)) = ((\alpha \cdot f) + (\beta \cdot f))(x)$  so that  $(\alpha + \beta) \cdot f = \alpha \cdot f + \beta \cdot f$
- c. If  $\alpha, \beta \in F$  and  $f \in \mathcal{F}(X, F)$  then  $\forall x \in X$  we have  $((\alpha \cdot \beta) \cdot f)(x) = (\alpha \cdot \beta) \cdot f(x) = \alpha \cdot (\beta \cdot f(x)) = \alpha \cdot ((\beta \cdot f)(x)) = (\alpha \cdot (\beta \cdot f))(x)$  giving that  $(\alpha \cdot \beta) \cdot f = \alpha \cdot (\beta \cdot f)$
- d. If  $f \in \mathcal{F}(X, F)$  then  $\forall x \in X$  we have  $(1 \cdot f)(x) = 1 \cdot f(x) = f(x)$  proving that  $1 \cdot f = f$

Next we have to prove that  $\{\delta_a\}_{a \in X}$  is a basis of  $\mathcal{F}(X, F)$

1. If  $f \in \mathcal{F}(X, F)$  then  $f^{-1}(F \setminus \{0\})$  is finite and we have two cases to consider

- a.  $(f^{-1}(F \setminus \{0\}) = \emptyset)$  then  $f = 0$  and as  $X$  is non empty there exists a  $a \in X$  and we have  $f = 0 \cdot \delta_a = 0$
- b.  $(f^{-1}(F \setminus \{0\}) \neq \emptyset)$  then there exists a bijection  $b: \{1, \dots, n\} \rightarrow f^{-1}(F \setminus \{0\})$  defining a family  $\{b_i\}_{i \in \{1, \dots, n\}}$  such that  $\forall y \in f^{-1}(F \setminus \{0\})$  we have that  $\exists i \in \{1, \dots, n\}$  such that  $b_i = y$  and  $\forall i, j \in \{1, \dots, n\}$  we have  $i \neq j \Rightarrow b_i \neq b_j$  Take now  $\sum_{i \in \{1, \dots, n\}} f(b_i) \cdot \delta_{b_i}$  then we have for  $x \in X$
- $(f(x) = 0)$  then  $x \notin f^{-1}(F \setminus \{0\})$  so  $\forall i \in \{1, \dots, n\}$  we have  $b_i \neq x$  and thus  $\delta_{b_i}(x) = 0$  giving that  $(\sum_{i \in \{1, \dots, n\}} f(b_i) \cdot \delta_{b_i})(x) = \sum_{i \in \{1, \dots, n\}} f(b_i) \cdot \delta_{b_i}(x) = \sum_{i \in \{1, \dots, n\}} f(b_i) \cdot 0 = 0 = f(x)$ .
  - $(f(x) \neq 0)$  then  $x \in f^{-1}(F \setminus \{0\})$  so there exists a unique  $i \in \{1, \dots, n\}$  such that  $b_i = x$  and thus for  $j \in \{1, \dots, n\}$  we have  $\delta_j(x) = \delta_j(b_i) = \delta_{i,j}$  so that  $(\sum_{j \in \{1, \dots, n\}} f(b_j))(x) = \sum_{j \in \{1, \dots, n\}} f(b_j) \cdot \delta_{b_j}(x) = \sum_{j \in \{1, \dots, n\}} f(b_j) \cdot \delta_{i,j} = f(b_i) = f(x)$
- proving that  $\sum_{i \in \{1, \dots, n\}} f(b_i) \cdot \delta_{b_i} = f$

So  $\mathcal{S}(\{\delta_a\}_{a \in X}) = \mathcal{F}(X, F)$

2. To prove linear independence, let  $\{\alpha_a\}_{a \in X}$  be a family in  $F$  with finite support such that  $\sum_{a \in X} \alpha_a \cdot \delta_a = 0$  then  $\forall x \in X$  we have  $0 = \sum_{a \in X} \alpha_a \cdot \delta_a(x) \underset{\delta_a(x) = 0 \text{ if } x \neq a}{=} \alpha_x \cdot \delta_x(x) = \alpha_x \Rightarrow \alpha_x = 0 \Rightarrow \forall x \in X$  we have  $\alpha_x = 0$ .

Finally we have to prove that  $\delta: X \rightarrow \mathcal{F}(X, F)$  defined by  $a \mapsto \delta_a$  is a injection, so assume that  $\delta_{a_1} = \delta_{a_2}$  then  $\forall x \in X$  we have  $\delta_{a_1}(x) = \delta_{a_2}(x)$  so if  $a_1 \neq a_2$  then  $1 = \delta_{a_1}(a_1) = \delta_{a_2}(a_1) = 0$  which is a contradiction if  $F$  has characterization 0 so we must have that  $a_1 = a_2$ .  $\square$

## 10.6 Linear Mappings

**Definition 10.172.** If  $X$  and  $Y$  are vector spaces over a field  $F$  then a function  $L: X \rightarrow Y$  is a **linear mapping** if  $\forall \alpha, \beta \in F$  and  $\forall x, y \in X$  we have  $L(\alpha \cdot x + \beta \cdot y) = \alpha \cdot L(x) + \beta \cdot L(y)$

**Example 10.173.** Let  $n \in \mathbb{N}_0$ ,  $\{X_i\}_{i \in \{1, \dots, n\}}$  a family of vector spaces over a field  $F$  then given  $i \in \{1, \dots, n\}$  we have that  $\pi_i: \prod_{j \in \{1, \dots, n\}} X_j \rightarrow X_i$  is linear

**Proof.** This follows trivially of the definition of  $\pi_i$  □

**Definition 10.174.** If  $X$  and  $Y$  are vector spaces over a field  $F$  then a linear mapping  $L: X \rightarrow Y$  is regular if  $L$  is injective. and the vector space  $\prod_{i \in \{1, \dots, n\}} X_i$  (see 10.120)

**Lemma 10.175.** Let  $X$  and  $Y$  be vector spaces over a field  $F$  and  $L: X \rightarrow Y$  is a linear mapping then  $L(0) = 0$

**Proof.**  $L(0) = L(0 \cdot 0) = 0 \cdot L(0) = 0$  □

**Theorem 10.176.** Let  $X$  and  $Y$  be vector spaces over a field  $F$  and  $L: X \rightarrow Y$  a linear mapping then

1. If  $n \in \mathbb{N}_0$  and  $\{\alpha_i\}_{i \in \{1, \dots, n\}}$ ,  $\{x_i\}_{i \in \{1, \dots, n\}}$  are families in  $F$  and  $X$  then  $L(\sum_{i=1}^n \alpha_i \cdot x_i) = \sum_{i=1}^n \alpha_i \cdot L(x_i)$
2. If  $\{\alpha_i\}_{i \in I}$  is a family in  $F$  with finite support and  $\{x_i\}_{i \in I}$  a family in  $X$  then  $L(\sum_{i \in I} \alpha_i \cdot x_i) = \sum_{i \in I} \alpha_i \cdot L(x_i)$
3. If  $A \subseteq X$  and  $\{\alpha_a\}_{a \in A}$  then  $L(\sum_{a \in A} \alpha_a \cdot a) = \sum_{a \in A} \alpha_a \cdot L(a)$

**Proof.**

1. This is proved by induction, let  $S = \{n \in \{1, \dots, n\} \mid \{\alpha_i\}_{i \in \{1, \dots, n\}}$  is a family in  $F$  and  $\{x_i\}_{i \in \{1, \dots, n\}}$  is a family in  $X$  then  $L(\sum_{i=1}^n \alpha_i \cdot x_i) = \sum_{i=1}^n \alpha_i \cdot L(x_i)\}$ , we have then

- a. If  $n = 1$  then  $L(\sum_{i=1}^1 \alpha_i \cdot x_i) = L(\alpha_1 \cdot x_1) = \alpha_1 \cdot L(x_1) = \sum_{i=1}^1 \alpha_i \cdot L(x_i)$  and thus  $1 \in S$
- b. If  $n \in S$  then if  $\{\alpha_i\}_{i \in \{1, \dots, n+1\}}$  is a field in  $F$  and  $\{x_i\}_{i \in \{1, \dots, n+1\}}$  is a field in  $X$  then we have  $L(\sum_{i=1}^{n+1} \alpha_i \cdot x_i) = L(\sum_{i=1}^n \alpha_i \cdot x_i + \alpha_{n+1} \cdot x_{n+1}) = L(\sum_{i=1}^n \alpha_i \cdot x_i) + \alpha_{n+1} \cdot L(x_{n+1}) \stackrel{n \in S}{=} \sum_{i=1}^n \alpha_i \cdot L(x_i) + \alpha_{n+1} \cdot L(x_{n+1}) = \sum_{i=1}^{n+1} \alpha_i \cdot L(x_i)$  and thus  $n+1 \in S$

by mathematical induction we have then  $S = \{1, \dots, n\} = \mathbb{N}_0$  proving (1).

2. As  $\text{support}(\{\alpha_i \cdot x_i\}_{i \in I}) \subseteq \text{support}(\{\alpha_i\}_{i \in I})$  (see 10.128). Now if  $i \in \text{support}(\{\alpha_i \cdot L(x_i)\}_{i \in I}) \Rightarrow \alpha_i \cdot L(x_i) \neq 0 \Rightarrow \alpha_i \neq 0 \Rightarrow i \in \text{support}(\{\alpha_i\}_{i \in I})$ , so  $\text{support}(\{\alpha_i \cdot x_i\}_{i \in I}), \text{support}(\{\alpha_i \cdot L(x_i)\}_{i \in I}) \subseteq \text{support}(\{\alpha_i\}_{i \in I})$  proving that  $\text{support}(\{\alpha_i \cdot x_i\}_{i \in I}), \text{support}(\{\alpha_i \cdot L(x_i)\}_{i \in I})$  is finite and thus that both sums in (2) are well defined. If now  $b: \{0, \dots, n\} \rightarrow \text{support}(\{\alpha_i\}_{i \in I})$  then we have  $L(\sum_{i \in I} \alpha_i \cdot x_i) = L(\sum_{i=0}^n \alpha_{b_i} \cdot x_{b_i}) = L(\sum_{i=1}^{n+1} \alpha_{b_{i-1}} \cdot x_{b_{i-1}}) \stackrel{(1)}{=} \sum_{i=1}^{n+1} \alpha_{b_{i-1}} \cdot L(x_{b_{i-1}}) = \sum_{i=0}^n \alpha_{b_i} \cdot L(x_{b_i}) = \sum_{i \in I} \alpha_i \cdot L(x_i)$

3. This follows from applying (2) to the families  $\{\alpha_a\}_{a \in A}$  and  $\{a\}_{a \in A}$  □

**Example 10.177.** If  $\{x_i\}_{i \in I}$  is a family with finite support in  $\mathbb{C}$  then as  $\text{Im}: \mathbb{C} \rightarrow \mathbb{R}$  defined by  $x \rightarrow \text{Im}(x)$  and  $\text{Re}: \mathbb{C} \rightarrow \mathbb{R}$  defined by  $x \rightarrow \text{Re}(x)$  are trivially linear we have that  $\text{Im}(\sum_{i \in I} x_i) = \sum_{i \in I} \text{Im}(x_i)$  and  $\text{Re}(\sum_{i \in I} x_i) = \sum_{i \in I} \text{Re}(x)$

**Theorem 10.178.** Let  $X$  be a vector space over a field  $F$  and  $x \in X$  with  $x \neq 0$  then there exists a linear mapping  $L: X \rightarrow F$  such that  $L(x) = 1$

**Proof.** If  $x \in X$  with  $x \neq 0$  then we have that  $\{x\}$  is linearly independent (see 10.109) and thus by 10.159 (and as  $x \neq 0$  we have that  $X$  is not trivial) there exists a set  $B$  with  $x \in B$  such that  $\{b\}_{b \in B}$  is a basis of  $X$ . Define now  $L: X \rightarrow F$  as follows, if  $y \in X$  then there exists a unique  $\{\alpha_b^{(y)}\}_{b \in B}$  with finite support such that  $y = \sum_{b \in B} \alpha_b^{(y)} \cdot b$  we define then that  $L(y) = \sum_{b \in B} \alpha_b^{(y)}$ . We prove now that  $L$  satisfies the requirement of the theorem.

1. ( $L(x) = 1$ ) As  $x \in B$  we can write  $x = 1 \cdot x$  so that  $\{\alpha_b^{(x)}\}_{b \in B}$  is defined by  $\alpha_x^{(x)} = 1$  and if  $b \neq x$  then  $\alpha_b^{(x)} = 0$  so that  $L(x) = \sum_{b \in B} \alpha_b^{(x)} = \alpha_x^{(x)} = 1$

2. ( **$L$  is linear**) If  $y_1, y_2 \in X$  and  $\alpha, \beta \in F$  then if  $y_1 = \sum_{b \in B} \alpha_b^{(y_1)} \cdot b$ ,  $y_2 = \sum_{b \in B} \alpha_b^{(y_2)} \cdot b$  we have  $\alpha \cdot y_1 + \beta \cdot y_2 = \alpha \cdot (\sum_{b \in B} \alpha_b^{(y_1)}) + \beta \cdot (\sum_{b \in B} \alpha_b^{(y_2)}) = \sum_{b \in B} (\alpha \cdot \alpha_b^{(y_1)} + \beta \cdot \alpha_b^{(y_2)}) \cdot b$  so that  $L(\alpha \cdot y_1 + \beta \cdot y_2) = L(\sum_{b \in B} (\alpha \cdot \alpha_b^{(y_1)} + \beta \cdot \alpha_b^{(y_2)}) \cdot b) = \sum_{b \in B} (\alpha \cdot \alpha_b^{(y_1)} + \beta \cdot \alpha_b^{(y_2)}) = \alpha \cdot \sum_{b \in B} \alpha_b^{(y_1)} + \beta \cdot \sum_{b \in B} \alpha_b^{(y_2)} = \alpha \cdot L(y_1) + \beta \cdot L(y_2)$   $\square$

We can use the above theorem to prove the following handy theorem

**Theorem 10.179.** *If  $X$  is a vector space over a field  $F$  with characterization zero,  $x, y \in X$  and if for every linear mapping  $L: X \rightarrow F$  we have  $L(x) = L(y)$  then we must have  $x = y$*

**Proof.** Assume that  $x \neq y \Rightarrow x - y \neq 0$  then using the previous theorem we can find a linear mapping  $L$  such that  $L(x - y) = 1$  but then we have  $1 = L(x - y) = L(x) - L(y) = 0$  giving the contradiction  $1 = 0$ .  $\square$

**Theorem 10.180.** *If  $X$  is a vector space over a field  $F$  and  $\{e_i\}_{i \in \{1, \dots, n\}}$  a linear independent set then there exists a linear mapping  $L: X \rightarrow F$  such that  $L(e_1) = 1$  and  $L(e_i) = 0, i \in \{2, \dots, k\}$ .*

**Proof.** As  $\{e_i\}_{i \in \{1, \dots, n\}}$  is linear independent it can be extended by 10.159 to a basis  $\{f_i\}_{i \in \{1, \dots, k\}}$  of  $X$  (so that  $\forall i \in \{1, \dots, n\}$  we have  $e_i = f_i$ ). Given now  $y \in X$  where  $y = \sum_{i \in \{1, \dots, k\}} y_i \cdot f_i$  is the unique expansion of  $y$  in the basis  $\{f_i\}_{i \in \{1, \dots, k\}}$ . Define then  $L(y) = y_1$ , this gives us the function (because the expansion is unique) that we are searching for.

1. ( $y = e_1$ ) then  $y = e_1 = f_1 = 1 \cdot f_1$  or  $y_1 = 1$  so that  $L(y) = 1$
2. ( $y = e_i$  where  $i \in \{2, \dots, n\}$ ) then  $y = e_i = 1 \cdot e_i$  or  $y_1 = 0$  so that  $L(e_i) = 0$
3. ( **$L$  is linear**) If  $x = \sum_{i \in \{1, \dots, k\}} x_i, y = \sum_{i \in \{1, \dots, k\}} y_i \in X$  and  $\alpha, \beta \in F$  then  $\alpha \cdot x + \beta \cdot y = \alpha \cdot \sum_{i \in \{1, \dots, k\}} x_i + \beta \cdot \sum_{i \in \{1, \dots, k\}} y_i = \sum_{i \in \{1, \dots, k\}} (\alpha \cdot x_i + \beta \cdot y_i)$  so that  $L(\alpha \cdot x + \beta \cdot y) = \alpha \cdot x_1 + \beta \cdot y_1 = \alpha \cdot L(x) + \beta \cdot L(y)$   $\square$

**Definition 10.181.** *If  $X, Y$  are vector spaces over a field  $F$  then a function  $L: X \rightarrow Y$  is a **linear isomorphism** iff*

1.  $L$  is a bijection
2.  $L$  is linear

**Example 10.182.**  $\mathcal{C}: \langle \mathbb{R}^2, +, \cdot \rangle \rightarrow \langle \mathbb{C}, +, \cdot \rangle$  defined by  $\mathcal{C}((x, y)) = x + i \cdot y$  is a isomorphism (here we consider  $\langle \mathbb{C}, +, \cdot \rangle$  a real vector space (see 10.114))

**Proof.** Using 9.39 we have that  $\mathcal{C}$  is a bijection so we have only to prove linearity. So let  $\alpha, \beta \in \mathbb{R}$  and  $(x_1, y_1), (x_2, y_2) \in \mathbb{R}^2$  then  $\mathcal{C}(\alpha \cdot (x_1, y_1) + \beta \cdot (x_2, y_2)) = \mathcal{C}((\alpha \cdot x_1 + \beta \cdot x_2, \alpha \cdot y_1 + \beta \cdot y_2)) = \alpha \cdot x_1 + \beta \cdot y_1 + i \cdot (\alpha \cdot y_1 + \beta \cdot y_2) = \alpha \cdot (x_1 + i \cdot y_1) + \beta \cdot (x_2 + i \cdot y_2) = \alpha \cdot \mathcal{C}((x_1, y_1)) + \beta \cdot \mathcal{C}((x_2, y_2))$   $\square$

**Theorem 10.183.** *If  $X, Y$  are vector spaces over a field  $F$  and  $L: X \rightarrow Y$  a linear isomorphism then  $L^{-1}: Y \rightarrow X$  is defined and is also a linear isomorphism.*

**Proof.** Using 2.41 we have that  $L^{-1}: Y \rightarrow X$  is defined and is also a bijection. If  $\alpha, \beta \in F$  and  $x, y \in X$  then there exists a  $x' = L^{-1}(x), y' = L^{-1}(y)$  such that  $L(x') = x, L(y') = y$ . Now  $L^{-1}(\alpha \cdot x + \beta \cdot y) = L^{-1}(\alpha \cdot L(x') + \beta \cdot L(y')) = L^{-1}(L(\alpha \cdot x' + \beta \cdot y')) = \alpha \cdot x' + \beta \cdot y' = \alpha \cdot L^{-1}(x) + \beta \cdot L^{-1}(y)$  proving the linearity of  $L^{-1}$ .  $\square$

**Theorem 10.184.** *If  $X, Y, Z$  are vector spaces over a field  $F$  and  $L_1: X \rightarrow Y, L_2: Y \rightarrow Z$  linear mappings (linear isomorphism) then  $L_2 \circ L_1: X \rightarrow Z$  is a linear mapping (linear isomorphism).*

**Proof.** First if  $\alpha, \beta \in F$  and  $x, y \in X$  then  $(L_2 \circ L_1)(\alpha \cdot x + \beta \cdot y) = L_2(L_1(\alpha \cdot x + \beta \cdot y)) = L_2(\alpha \cdot L_1(x) + \beta \cdot L_1(y)) = \alpha \cdot L_2(L_1(x)) + \beta \cdot L_2(L_1(y)) = \alpha \cdot (L_2 \circ L_1)(x) + \beta \cdot (L_2 \circ L_1)(y)$ .

Second if  $L_1, L_2$  are linear isomorphism they are linear and thus  $L_2 \circ L_1$  is linear (see above) (and a bijection (see 2.49)).  $\square$

**Theorem 10.185.** Let  $\langle X, +, \cdot \rangle$  be a vector space over a field  $\langle F, +, \cdot \rangle$ ,  $Y$  a set and  $L: X \rightarrow Y$  a bijection. Define then  $+_L: Y \times Y \rightarrow Y$  by  $(x, y) \rightarrow x +_L y = L(L^{-1}(x) + L^{-1}(y))$  and  $\cdot_L: F \times Y \rightarrow Y$  by  $(\alpha, x) \rightarrow \alpha \cdot_L x = L(\alpha \cdot L^{-1}(x))$ . Then  $\langle Y, +_L, \cdot_L \rangle$  forms a vector space over  $\langle F, +, \cdot \rangle$  with zero element  $L(0)$  and  $L: X \rightarrow Y$  is then a isomorphism. Further if  $\{e_i\}_{i \in I}$  is a basis in  $X$  then  $\{L(e_i)\}_{i \in I}$  is a basis in  $\langle Y, +_L, \cdot_L \rangle$

**Proof.** First we prove that  $\langle Y, +_L \rangle$  is a abelian group

1. **(associativity)**  $x +_L (y +_L z) = x +_L (L(L^{-1}(x) + L^{-1}(y)) + L^{-1}(z)) = L(L^{-1}(x) + L^{-1}(L(L^{-1}(y) + L^{-1}(z)))) = L(L^{-1}(x) + (L^{-1}(y) + L^{-1}(z))) = L((L^{-1}(x) + L^{-1}(y)) + L^{-1}(z)) = L(L^{-1}(L(L^{-1}(x) + L^{-1}(y))) + L^{-1}(z)) = L(L^{-1}(x +_L y) + L^{-1}(z)) = ((x +_L y) +_L z)$
2. **(commutativity)**  $x +_L y = L(L^{-1}(x) + L^{-1}(y)) = L(L^{-1}(y) + L^{-1}(x)) = y +_L x$
3. **(neutral element)**  $x +_L L(0) = L(L^{-1}(x) + L^{-1}(L(0))) = L(L^{-1}(x) + 0) = L(L^{-1}(x)) = x$   
so  $x +_L L(0) = x \stackrel{\text{commutativity}}{=} L(0) +_L x$

Next we prove the remaining requirements of a vector space

1.  $\alpha \cdot_L (x +_L y) = \alpha \cdot_L (L(L^{-1}(x) + L^{-1}(y))) = L(\alpha \cdot L^{-1}(L(L^{-1}(x) + L^{-1}(y)))) = L(\alpha \cdot (L^{-1}(x) + L^{-1}(y))) = L(\alpha \cdot L^{-1}(x) + \alpha \cdot L^{-1}(y)) = L(L^{-1}(L(\alpha \cdot L^{-1}(x))) + L^{-1}(L(\alpha \cdot L^{-1}(y)))) = L(\alpha \cdot L^{-1}(x)) +_L L(\alpha \cdot L^{-1}(y)) = \alpha \cdot_L x +_L \alpha \cdot_L y$
2.  $(\alpha + \beta) \cdot_L x = L((\alpha + \beta) \cdot L^{-1}(x)) = L(\alpha \cdot L^{-1}(x) + \beta \cdot L^{-1}(x)) = L(L^{-1}(L(\alpha \cdot L^{-1}(x))) + L^{-1}(L(\beta \cdot L^{-1}(x)))) = L(\alpha \cdot L^{-1}(x)) +_L L(\beta \cdot L^{-1}(x)) = \alpha \cdot_L x + \beta \cdot_L x$
3.  $(\alpha \cdot \beta) \cdot_L x = L((\alpha \cdot \beta) \cdot L^{-1}(x)) = L(\alpha \cdot (\beta \cdot L^{-1}(x))) = L(\alpha \cdot L^{-1}(L(\beta \cdot L^{-1}(x)))) = \alpha \cdot_L L(\beta \cdot L^{-1}(x)) = \alpha \cdot_L (\beta \cdot_L x)$

Next we prove that  $L$  is linear (and as it is bijective) thus a isomorphism

$$L(\alpha \cdot x + \beta \cdot y) = L(L^{-1}(L(\alpha \cdot x)) + L^{-1}(L(\beta \cdot y))) = L(\alpha \cdot x) +_L L(\beta \cdot y) = L(\alpha \cdot L^{-1}(L(x))) + L(\beta \cdot L^{-1}(L(y))) = \alpha \cdot_L L(x) + \beta \cdot_L L(y)$$

Finally if  $\{e_i\}_{i \in I}$  is a basis for  $\langle X, +, \cdot \rangle$  then we have for  $\{L(e_i)\}_{i \in I}$

1. If  $\{\alpha_i\}_{i \in I}$  is a family with finite support such that  $0 = L(0) = \sum_{i \in I} \alpha_i \cdot L(e_i) \stackrel{L \text{ is linear}}{=} L(\sum_{i \in I} \alpha_i \cdot e_i) \stackrel{L \text{ is bijective}}{\Rightarrow} 0 = \sum_{i \in I} \alpha_i \cdot e_i \stackrel{\{e_i\}_{i \in I} \text{ is linear independent}}{\Rightarrow} \forall i \in I \text{ we have } \alpha_i = 0$  proving that

$$\{L(e_i)\}_{i \in I} \text{ is independent} \quad (10.41)$$

2. If  $y \in Y$  then there exists a  $x \in X$  such that  $y = L(x)$ , as  $\{e_i\}_{i \in I}$  is a basis there exists a  $\{\alpha_v\}_{v \in \{e_i|i \in I\}}$  with finite support such that  $x = \sum_{v \in \{e_i|i \in I\}} \alpha_v \cdot v$  and thus  $y = L(x) = L(\sum_{v \in \{e_i|i \in I\}} \alpha_v \cdot v) = \sum_{v \in \{e_i|i \in I\}} \alpha_v \cdot L(v)$  so

$$y = \sum_{v \in \{e_i|i \in I\}} \alpha_v \cdot L(v) \quad (10.42)$$

Now as  $L: X \rightarrow Y$  is a bijection we have that  $L|_{\{e_i|i \in I\}}: \{e_i|i \in I\} \rightarrow L(\{e_i|i \in I\}) = \{L(e_i)|i \in I\}$  is a bijection, now define  $\{\beta_v\}_{v \in \{L(e_i)|i \in I\}}$  by  $\beta_v = \alpha_{L^{-1}(v)}$  then we have  $\beta_{L(v)} = \alpha_{L^{-1}(L(v))} = \alpha_v$  and thus using 10.42 we have  $y = \sum_{v \in \{e_i|i \in I\}} \beta_{L(v)} \cdot L(v) = \sum_{v \in \{e_i|i \in I\}} \beta_{L|_{\{e_i|i \in I\}}(v)} \cdot L|_{\{e_i|i \in I\}}(v) \stackrel{10.43}{=} \sum_{v \in \{L(e_i)|i \in I\}} \beta_v \cdot v$  and thus  $y \in \mathcal{S}(\{L(e_i)|i \in I\})$  and this proves that:

$$Y = \mathcal{S}(\{L(e_i)|i \in I\}) \quad (10.43)$$

Using 10.41 and 10.43 we have then that  $\{L(e_i)\}_{i \in I}$  is linear independent.  $\square$

**Definition 10.186.** Let  $X, Y$  be vector spaces over a field  $F$  then  $\text{Hom}(X, Y) = \{L \in M(X, Y) | L \text{ is a linear mapping}\}$  (for a definition of  $M(X, Y)$  see 10.115). [In other words  $\text{Hom}(X, Y)$  is the set of all linear mappings from  $X$  to  $Y$ ]

**Theorem 10.187.** Let  $X, Y$  be a vector space over a field  $F$  then  $\text{Hom}(X, Y)$  is a non empty subspace of  $M(X, Y)$  (see 10.115 for definition of  $M(X, Y)$ ) and thus by 10.111  $\text{Hom}(X, Y)$  is a vector space.

**Proof.** First if we define  $0: X \rightarrow Y$  by  $x \rightarrow 0(x) = 0$  then as  $0(\alpha \cdot x + \beta \cdot y) = 0 = 0 + 0 = \alpha \cdot 0(x) + \beta \cdot 0(y)$  so  $0 \in \text{Hom}(X, Y) \Rightarrow \emptyset \neq \text{Hom}(X, Y)$ .

Finally if  $\alpha, \beta \in F$ ,  $L_1, L_2 \in \text{Hom}(X, Y)$  then if  $\alpha', \beta' \in F$  and  $x, y \in X$  we have  $(\alpha \cdot L_1 + \beta \cdot L_2)(\alpha' \cdot x + \beta' \cdot y) = \alpha \cdot L_1(\alpha' \cdot x + \beta' \cdot y) + \beta \cdot L_2(\alpha' \cdot x + \beta' \cdot y) = \alpha \cdot (\alpha' \cdot L_1(x)) + \alpha \cdot (\beta' \cdot L_1(y)) + \beta \cdot (\alpha' \cdot L_2(x)) + \beta \cdot (\beta' \cdot L_2(y)) = \alpha' \cdot (\alpha \cdot L_1(x) + \beta \cdot L_2(x)) + \beta' \cdot (\alpha \cdot L_1(y) + \beta \cdot L_2(y)) = \alpha' \cdot (\alpha \cdot L_1 + \beta \cdot L_2)(x) + \beta' \cdot (\alpha \cdot L_1 + \beta \cdot L_2)(y)$  proving that  $\alpha \cdot L_1 + \beta \cdot L_2$  is linear and thus that  $\alpha \cdot L_1 + \beta \cdot L_2 \in \text{Hom}(X, Y)$ .  $\square$

**Definition 10.188.** Let  $V$  be a vector space over a field  $F$  (which is itself a vector space by 10.112) then we note the vector space  $\text{Hom}(V, F)$  as  $V^*$  and call it the **dual space** of  $V$

**Theorem 10.189.** If  $V$  is a finite vector space over the field  $F$  with a basis  $\{e_i\}_{i \in \{1, \dots, m\}}$  (so  $\dim(V) = m$ ). Then  $V^*$  is finite dimensional with basis  $\{e_i^*\}_{i \in \{1, \dots, m\}}$  where  $\forall i \in \{1, \dots, m\}$  we define  $e_i^*: V \rightarrow F$  by  $e_i^*(\sum_{i \in \{1, \dots, m\}} \lambda_i \cdot e_i) = \lambda_i$

**Proof.** First we have to prove that  $e_i^*$  is linear so if we take  $x = \sum_{i \in \{1, \dots, m\}} \lambda_i \cdot e_i$ ,  $y = \sum_{i \in \{1, \dots, m\}} \lambda'_i \cdot e_i$  and  $\alpha, \beta \in F$  then  $\alpha \cdot x + \beta \cdot y = \alpha \cdot (\sum_{i \in \{1, \dots, m\}} \lambda_i \cdot e_i) + \beta (\sum_{i \in \{1, \dots, m\}} \lambda'_i \cdot e_i) = \sum_{i \in \{1, \dots, m\}} (\alpha \cdot \lambda_i + \beta \cdot \lambda'_i)$  so that  $e_i^*(\alpha \cdot x + \beta \cdot y) = \alpha \cdot \lambda_i + \beta \cdot \lambda'_i = \alpha \cdot e_i^*(x) + \beta \cdot e_i^*(y)$  proving linearity.

Second we prove that  $\mathcal{S}(\{e_i^*\}_{i \in \{1, \dots, m\}}) = V^*$ , if  $f \in V^*$  then  $f: V \rightarrow F$  is a linear mapping. Define now  $\{\alpha_i\}_{i \in \{1, \dots, m\}}$  by  $\alpha_i = f(e_i)$  then  $\forall x \in V$  we have  $x = \sum_{i \in \{1, \dots, m\}} \lambda_i \cdot e_i \Rightarrow f(x) = \sum_{i \in \{1, \dots, m\}} \lambda_i \cdot f(e_i) = \sum_{i \in \{1, \dots, m\}} \alpha_i \cdot \lambda_i = \sum_{i \in \{1, \dots, m\}} \alpha_i \cdot e_i^*(x) = (\sum_{i \in \{1, \dots, m\}} \alpha_i \cdot e_i^*)(x)$  proving that  $f = \sum_{i \in \{1, \dots, m\}} \alpha_i \cdot e_i^*$  and thus that  $V^* \subseteq \mathcal{S}(\{e_i^*\}_{i \in \{1, \dots, m\}})$  which together with the trivial  $\mathcal{S}(\{e_i^*\}_{i \in \{1, \dots, m\}}) \subseteq V^*$  we have  $V^* = \mathcal{S}(\{e_i^*\}_{i \in \{1, \dots, m\}})$ .

Finally to prove that  $\{e_i\}_{i \in \{1, \dots, m\}}$  is linear independent assume that  $\sum_{i \in \{1, \dots, m\}} \alpha_i \cdot e_i^* = 0$  (the mapping  $V \rightarrow F$  with  $0(x) = 0$ ) then  $\forall i \in \{1, \dots, m\}$  we have that  $0 = 0(e_i) = (\sum_{j \in \{1, \dots, m\}} \alpha_j \cdot e_j^*(e_i)) = \sum_{j \in \{1, \dots, m\}} \alpha_j \cdot \delta_{i,j} = \alpha_i \Rightarrow \alpha_i = 0$  proving linear independence.  $\square$

**Definition 10.190.** If  $X$  is a vector space over a field  $F$  then  $\text{GL}(X) = \{L \in \text{Hom}(X, X) \mid L \text{ is a bijection}\} \subseteq \text{Hom}(X, X)$ . So  $\text{GL}(X)$  is the set of all the linear isomorphism's from  $X \rightarrow X$ .

**Theorem 10.191.** If  $X$  is a vector space over a field  $F$  then  $(\text{GL}(X), \circ)$  is a group with neutral element  $1_X$  (here  $\circ$  is the function composition operator). This group is called the **general linear group**. Furthermore if we defined  $\triangleright: \text{GL}(X) \times X \rightarrow X$  by  $(L, x) \rightarrow \triangleright(L, x) = L \triangleright x$  defined then  $\triangleright$  is a **faithful left action** (see 3.19, 3.16)

**Proof.** First we must prove that  $\circ: \text{GL}(X) \times \text{GL}(X) \rightarrow \text{GL}(X)$  is indeed a function, essential we must prove that if  $L_1, L_2 \in \text{GL}(X)$  then that  $L_1 \circ L_2 \in \text{GL}(X)$  this follows directly from 10.184. Now for the rest of the group axioms:

1. **(associativity)** This follows from the associativity of the composition of functions (see 1.48)
2. **(neutral element)** If  $L \in \text{GL}(X)$  then trivially we have  $1_X \in \text{GL}(X)$  and  $L \circ 1_X = L = 1_X \circ L$
3. **(inverse element)** If  $L \in \text{GL}(X)$  then using 10.183 we have that  $L^{-1} \in \text{GL}(X)$  and trivially we have  $L \circ L^{-1} = 1_X = L^{-1} \circ L$

To prove that  $\triangleright$  is a left action note that:

1.  $\forall x \in X$  we have  $1_X \triangleright x = 1_X(x) = x \Rightarrow 1_X \triangleright x = x$
2.  $\forall L, L' \in \text{GL}(X)$  and  $\forall x \in X$  we have  $(L \circ L') \triangleright x = (L \circ L')(x) = L(L'(x)) = L \triangleright (L'(x)) = L \triangleright (L' \triangleright x)$

To prove that the left action is faithful, if  $L \in \text{GL}(X)$  take then  $L_{\triangleright}: X \rightarrow X$  defined by  $x \rightarrow L_{\triangleright}(x) = L \triangleright x = L(x)$ , we have then the following:

1. If  $L_{\triangleright} = 1_X$  then  $\forall x \in X$  we have  $L(x) = L_{\triangleright}(x) = 1_X(x) = x \Rightarrow L = 1_X$  the neutral element in  $\text{GL}(X)$
2. If  $L = 1_X$  then  $\forall x \in X$  we have  $L_{\triangleright}(x) = L(x) = 1_X(x) = x \Rightarrow L_{\triangleright} = 1_X$   $\square$

**Theorem 10.192.** Let  $X$  be a vector space over a field  $F$  and let  $\emptyset \neq Y \subseteq X$  be a subspace of  $X$  then  $\sim_Y$  defined by  $\sim_Y = \{(x, y) \in X \times X \mid x - y \in Y\}$  is a equivalence relation.

**Proof.** First as  $\emptyset \neq Y$  we have that  $Y$  is a vector space over the field  $F$  and thus if  $0$  is the neutral element in  $X$  then  $0 \in Y$ . Next we prove that  $\sim_X$  is a equivalence relation

1. **(reflexive)** If  $x \in X$  then  $x - x = 0 \in Y \Rightarrow x \sim_Y x$
2. **(symmetric)** If  $x, y \in X$  and  $x \sim_Y y$  then  $x - y \in Y \Rightarrow (-1) \cdot (x - y) \in Y \Rightarrow y - x \in Y \Rightarrow y \sim_Y x$
3. **(transitive)** If  $x, y, z \in X$  and  $x \sim_Y y, y \sim_Y z$  then we have  $x - y \in Y$  and  $y - z \in Y$  and thus  $x - z = (x - y) + (y - z) \in Y \Rightarrow x \sim_Y z$   $\square$

**Theorem 10.193.** Let  $X$  be a vector space over a field  $F$  and  $\emptyset \neq Y \subseteq X$  a subspace of  $X$  then if we define  $X/Y = X/\sim_Y = \{\sim_Y[x] \mid x \in X\}$  and

1.  $+: X/Y \times X/Y \rightarrow X/Y$  defined by  $(\sim_Y[x], \sim_Y[y]) \rightarrow \sim_Y[x] + \sim_Y[y] = \sim_Y[x + y]$
2.  $\cdot: F \times X/Y \rightarrow X/Y$  defined by  $(\alpha, \sim_Y[x]) \rightarrow \alpha \cdot \sim_Y[x] = \sim_Y[\alpha \cdot x]$

then these are well defined functions and  $\langle X/Y, +, \cdot \rangle$  forms a vector space over  $F$  with neutral element  $Y = \sim_Y[0]$ , this vector space is called the **factor space**. Also the canonical surjective function (see 2.131)  $\pi_Y: X \rightarrow X/Y$  defined by  $x \rightarrow \pi_Y(x) = \sim_Y[x]$  is a linear mapping.

**Proof.** First we prove that the operations are well defined. If  $\sim_Y[x] = \sim_Y[x']$  and  $\sim_Y[y] = \sim_Y[y']$  then we have  $x - x' \in Y$  and  $y - y' \in Y$  and we have:

1.  $(x + y) - (x' + y') = (x - x') + (y - y') \in Y \Rightarrow (x + y) \sim_Y (x' + y')$  so we have  $\sim_Y[x + y] = \sim_Y[x' + y']$  proving that  $+: X/Y \times X/Y \rightarrow X/Y$  is well defined.
2.  $\alpha \cdot x - \alpha \cdot x' = \alpha \cdot (x - x') \in Y \Rightarrow \alpha \cdot x \sim_Y \alpha \cdot x'$  so we have  $\sim_Y[\alpha \cdot x] = \sim_Y[\alpha \cdot x']$  proving that  $\cdot: F \times X/Y \rightarrow X/Y$  is well defined.

Next we prove that  $\langle X/Y, +, \cdot \rangle$  is a vector space over the field  $F$ .

1. **(Abelian group axioms)** Let  $\sim_Y[x], \sim_Y[y], \sim_Y[z] \in X/Y$ 
  - a. **(associativity)**  $\sim_Y[x] + (\sim_Y[y] + \sim_Y[z]) = \sim_Y[x] + \sim_Y[y + z] = \sim_Y[x + (y + z)] = \sim_Y[(x + y) + z] = \sim_Y[x + y] + \sim_Y[z] = (\sim_Y[x] + \sim_Y[y]) + \sim_Y[z]$
  - b. **(commutativity)**  $\sim_Y[x] + \sim_Y[y] = \sim_Y[x + y] = \sim_Y[y + x] = \sim_Y[y] + \sim_Y[x]$
  - c. **(neutral element)** As  $Y \neq \emptyset$  we have that  $Y$  is a subspace and thus  $0 \in Y$ , if  $y \in Y \Rightarrow y = y - 0 \in Y \Rightarrow y \sim_Y 0 \Rightarrow y \in \sim_Y[0] \Rightarrow Y \subseteq \sim_Y[0]$ , if  $y \in \sim_Y[0]$  then  $y = y - 0 \in Y \Rightarrow \sim_Y[0] \subseteq Y$  so we have

$$\sim_Y[0] = Y \tag{10.44}$$

Then  $\sim_Y[x] + Y = \sim_Y[x] + \sim_Y[0] = \sim_Y[x + 0] = \sim_Y[x] \Rightarrow x = \sim_Y[x] + Y \underset{\text{commutativity}}{=} Y + \sim_Y[x]$

- d. **(inverse element)**  $\sim_Y[x] + \sim_Y[-x] = \sim_Y[x - x] = \sim_Y[0] = Y \Rightarrow Y = \sim_Y[x] + \sim_Y[-x] \underset{\text{commutativity}}{=} \sim_Y[-x] + \sim_Y[x]$

## 2. (Vector space axioms)

- a. Let  $\alpha \in F$  and  $\sim_Y[x], \sim_Y[y] \in X/Y$  then  $\alpha \cdot (\sim_Y[x] + \sim_Y[y]) = \alpha \cdot \sim_Y[x + y] = \sim_Y[\alpha \cdot (x + y)] = \sim_Y[\alpha \cdot x + \alpha \cdot y] = \sim_Y[\alpha \cdot x] + \sim_Y[\alpha \cdot y] = \alpha \cdot \sim_Y[x] + \alpha \cdot \sim_Y[y]$
- b. Let  $\alpha, \beta \in F$  and  $\sim_Y[x] \in X/Y$  then  $(\alpha + \beta) \cdot \sim_Y[x] = \sim_Y[(\alpha + \beta) \cdot x] = \sim_Y[\alpha \cdot x + \beta \cdot x] = \sim_Y[\alpha \cdot x] + \sim_Y[\beta \cdot x] = \alpha \cdot \sim_Y[x] + \beta \cdot \sim_Y[x]$
- c. Let  $\alpha, \beta \in F$  and  $\sim_Y[x] \in X/Y$  then  $(\alpha \cdot \beta) \cdot \sim_Y[x] = \sim_Y[(\alpha \cdot \beta) \cdot x] = \sim_Y[\alpha \cdot (\beta \cdot x)] = \alpha \cdot \sim_Y[\beta \cdot x] = \alpha \cdot (\beta \cdot \sim_Y[x])$
- d. If  $\sim_Y[x] \in X/Y$  then we have  $1 \cdot \sim_Y[x] = \sim_Y[1 \cdot x] = \sim_Y[x]$

Finally if  $\alpha, \beta \in F$  and  $x, y \in X$  then  $\pi_Y(\alpha \cdot x + \beta \cdot y) = \sim_Y[\alpha \cdot x + \beta \cdot y] = \sim_Y[\alpha \cdot x] + \sim_Y[\beta \cdot y] = \alpha \cdot \sim_Y[x] + \beta \cdot \sim_Y[y] = \alpha \cdot \pi_Y(x) + \beta \cdot \pi_Y(y)$  proving the linearity of  $\pi_Y$ .  $\square$

**Definition 10.194. (Direct Internal Sum)** If  $V$  is a vector space over a field  $F$  then  $V = X \oplus Y$  iff  $X, Y$  are non empty sub spaces of  $V$ ,  $X \cap Y = \{0\}$  and  $V = \{x + y \mid (x, y) \in X \times Y\}$

**Theorem 10.195.** If  $V$  is a vector space over a field  $F$  then if  $V = X \oplus Y$  then  $\forall a \in V$  we have that there exists **unique**  $x \in X, y \in Y$  such that  $a = x + y$

**Proof.** If  $a \in V = X \oplus Y$  then there exists  $x \in X, y \in Y$  such that  $a = x + y$ . Assume now that also  $a = x' + y'$  (where  $x' \in X, y' \in Y$ ) then  $0 = a - a = (x - x') + (y - y') \Rightarrow x - x' = y' - y \Rightarrow X \ni x - x' = y' - y \in Y \Rightarrow (x - x'), (y' - y) \in X \cap Y = \{0\} \Rightarrow x = x', y = y'$   $\square$

**Theorem 10.196.** If  $V$  is a vector space over a field  $F$  and  $\emptyset \neq X \subseteq V$  is a sub space of  $V$  then there exist a sub space  $\emptyset \neq Y \subseteq V$  such that  $V = X \oplus Y$

**Proof.** As  $X \neq \emptyset$   $X$  is a vector space and thus  $0 \in X$  we have then the following cases to consider for  $X$

1. ( $X = V$ ) take then  $Y = \{0\}$  then  $X \cap Y = V \cap \{0\} = \{0\}$  and if  $v \in V \Rightarrow v = v + 0$  where  $(v, 0) \in V \times \{0\} = X \times Y \Rightarrow V = X \oplus Y$
2. ( $X = \{0\}$ ) take then  $Y = V$  then  $X \cap Y = \{0\} \cap Y = \{0\}$  and if  $v \in V \Rightarrow v = 0 + v$  where  $(0, v) \in \{0\} \times V = X \times Y \Rightarrow V = X \oplus Y$
3. ( $X \neq \{0\}, V$ ) Then as  $X \neq \{0\}$  we have that  $X$  is not a trivial vector space. Using 10.157 we have the existence of a  $B_X \subseteq X$  such that  $\{b\}_{b \in B_X}$  is a basis for  $X$ . Using the linear independence of  $B_X$  and 10.159 we find a  $B$  with  $B_X \subseteq B$  and  $\{b\}_{b \in B}$  is a basis for  $V$ . If now  $B_X = B$  then  $X = \mathcal{S}(B_X) = \mathcal{S}(B) = V$  contradicting  $X \neq V$  so we must have  $B_X \subset B$  and thus have  $\emptyset \neq B \setminus B_X$ . Using 10.132 we have then that  $\mathcal{S}(B \setminus B_X)$  is a vector space over  $F$  (and also a subspace of  $V$ ), take then  $Y = \mathcal{S}(B \setminus B_X)$ . If now  $v \in V$  then as  $\{b\}_{b \in B}$  is a basis of  $V$  there exists a  $\{\alpha_b\}_{b \in B}$  with finite support such that  $v = \sum_{b \in B} \alpha_b \cdot b = \sum_{b \in B_X} \alpha_b \cdot b + \sum_{b \in B \setminus B_X} \alpha_b \cdot b = x + y$  where  $x = \sum_{b \in B_X} \alpha_b \cdot b \in \mathcal{S}(B_X) = X$  and  $y = \sum_{b \in B \setminus B_X} \alpha_b \cdot b \in \mathcal{S}(B \setminus B_X) = Y$  also if  $x \in X \subseteq V, y \in Y \subseteq V$  we have  $x + y \in V$ . So we have

$$v \in V \Leftrightarrow \exists (x, y) \in X \times Y \text{ such that } v = x + y \quad (10.45)$$

If now  $z \in X \cap Y$  then there exists families  $\{\beta_b\}_{b \in B_X}, \{\tau_b\}_{b \in B \setminus B_X}$  in  $F$  with finite support such that  $\sum_{b \in B_X} \beta_b \cdot b = z = \sum_{b \in B \setminus B_X} \tau_b \cdot b \Rightarrow \sum_{b \in B_X} \beta_b \cdot b - \sum_{b \in B \setminus B_X} \tau_b \cdot b = 0$ . Define now  $\{\gamma_b\}_{b \in B}$  by  $\gamma_b = \begin{cases} \beta_b & \text{if } b \in B_X \\ -\tau_b & \text{if } b \in B \setminus B_X \end{cases}$  which, because  $\text{support}(\{\gamma_b\}_{b \in B}) \subseteq \text{support}(\{\beta_b\}_{b \in B_X}) \cup \text{support}(\{\tau_b\}_{b \in B \setminus B_X})$ , has finite support. The  $\sum_{b \in B} \gamma_b \cdot b = \sum_{b \in B_X} \gamma_b \cdot b + \sum_{b \in B \setminus B_X} \gamma_b \cdot b = \sum_{b \in B_X} \beta_b \cdot b + \sum_{b \in B \setminus B_X} (-1) \cdot \tau_b \cdot b = \sum_{b \in B_X} \beta_b \cdot b - \sum_{b \in B \setminus B_X} \tau_b \cdot b = 0$ . Using the independence of  $B$  we have then  $\forall b \in B$  we have  $\gamma_b = 0$ , so  $\forall b \in B_X \subseteq B$  we have  $\alpha_b = \beta_b = 0$  and thus  $z = \sum_{b \in B_X} \beta_b \cdot b = \sum_{b \in B_X} 0 \cdot b = 0$  proving that

$$X \cap Y = \{0\} \quad (10.46)$$

From 10.45 and 10.46 we have then  $V = X \oplus Y$   $\square$

**Theorem 10.197.** If  $V$  is a non trivial vector space over a field  $F$  with a basis  $\{b_i\}_{i \in \{1, \dots, n\}}$  then if  $\{a_i\}_{i \in \{1, \dots, n+1\}}$  is a finite family in  $V$  we have that  $\{a_i\}_{i \in \{1, \dots, n+1\}}$  is linear dependent.

**Proof.** We prove this by induction so let  $S = \{n \in \{1, \dots\} \mid \text{If } V \text{ is a vector space with a finite basis } \{b_i\}_{i \in \{1, \dots, m\}} \text{ then if } \{a_i\}_{i \in \{1, \dots, n+1\}} \text{ is a finite family in } V \text{ then } \{a_i\}_{i \in \{1, \dots, n\}} \text{ is linear dependent}\}$  we have then:

1. ( $n = 1$ ) If  $V$  is a vector space with a basis  $\{v_i\}_{i \in \{1, \dots, m\}}$  where  $m \leq 1$  then as  $V$  is not trivial we have by 10.152 that  $\{1, \dots, m\} \neq \emptyset \Rightarrow m = 1$  and thus  $\{v_i\}_{i \in \{1\}}$  is a basis of  $V$ . If now  $A = \{a_i\}_{i \in \{1, \dots, 2\}} \Rightarrow A = \{a_1, a_2\}$  then there exists a  $\{\alpha_i\}_{i \in \{1\}}, \{\beta_i\}_{i \in \{1\}}$  such that  $a_1 = \sum_{i \in \{1\}} \alpha_i \cdot v_i = \alpha_1 \cdot v_1$  and  $\beta_1 = \sum_{i \in \{1\}} \beta_i \cdot v_i = \beta_1 \cdot v_1$ . We have then the following possibilities:

- a.  $(\alpha_1 = 0 \vee \beta_1 = 0) \Rightarrow a_1 = 0 \vee a_2 = 0 \Rightarrow A$  is linear dependent

- b.  $(\alpha_1, \alpha_2 \neq 0)$  define then  $\{\gamma_i\}_{i \in \{1,2\}}$  such that  $\gamma_{a_1} = 1$  and  $\gamma_{a_2} = -\alpha_1 \cdot \beta_1^{-1}$  then  $\sum_{i \in \{1,2\}} \gamma_i \cdot a_i = 1 \cdot a_1 - \alpha_1 \cdot \beta_1^{-1} = \alpha_1 \cdot v_1 - \alpha_1 \cdot \beta_1^{-1}(\beta_1 \cdot v_1) = \alpha_1 \cdot v_1 - \alpha_1 \cdot v_1 = 0$  which as  $\gamma_{a_1} = 1 \neq 0$  proves that  $A$  is linear dependent.

So in all cases we have that  $A$  is linear dependent and thus we have that  $1 \in S$ .

2.  $(n \in S)$  If  $V$  is a vector space with a basis  $\{b_i\}_{i \in \{1, \dots, m\}}$  where  $m \leq n+1$ . As  $V$  is non trivial we have by 10.152 that  $\{1, \dots, m\} \neq \emptyset \Rightarrow 1 \leq m \leq n+1$ . Let now  $A = \{a_i\}_{i \in \{1, \dots, n+2\}}$  a family in  $V$ . We have now the following cases:

- a.  $(m \leq n)$  then as  $n \in S$ ,  $m \leq n$  and we have that  $\{a_i\}_{i \in \{1, \dots, n+1\}}$  is linear dependent. Finally using 10.149 and  $\{a_i\}_{i \in \{1, \dots, n+1\}} \preceq \{a_i\}_{i \in \{1, \dots, n+2\}}$  we have that  $A = \{a_i\}_{i \in \{1, \dots, n+2\}}$  is linear dependent.

- b.  $(n < m)$  then we have  $n < m \leq n+1 \Rightarrow n+1 \leq m \leq n+1 \Rightarrow m = n+1$ . We have now the following possibilities for  $a_1, a_{n+2}$

- i.  $(a_1 = a_{n+2})$  then we have that  $A$  is linear dependent

- ii.  $(a_1 \neq a_{n+2} \wedge (a_{n+2} = 0 \vee a_1 = 0))$  then  $A$  is linear dependent.

- iii.  $(a_1 \neq a_{n+2} \wedge (a(n+2), a(1) \neq 0))$ . Take then

$$V_1 = V / \mathcal{S}(\{a_{n+2}\}) \quad (10.47)$$

and the surjection  $\pi_{\mathcal{S}(\{a_{n+2}\})}: V \rightarrow V_1$ . Define now  $\{\bar{b}_i\}_{i \in \{1, \dots, n+1\}}$  where  $\forall i \in \{1, \dots, n+1\}$  we have  $\bar{b}_i = \pi_{\mathcal{S}(\{a_{n+2}\})}(b_i)$ . Now if  $y \in V_1$   $\pi_{\mathcal{S}(\{a_{n+2}\})}$  is a surjection  $\exists x \in V$  such that  $\pi_{\mathcal{S}(\{a_{n+2}\})}(x) = y$ . As  $\{b_i\}_{i \in \{1, \dots, n+1\}}$  is a basis of  $V$  there exists a  $\{\alpha_i\}_{i \in \{1, \dots, n+1\}}$  with finite support such that  $x = \sum_{i \in \{1, \dots, n+1\}} \alpha_i \cdot b_i$  and thus  $y = \pi_{\mathcal{S}(\{a_{n+2}\})}(\sum_{i \in \{1, \dots, n+1\}} \alpha_i \cdot b_i) \underset{\pi_{\mathcal{S}(\{a_{n+2}\})} \text{ is linear}}{=} \sum_{i \in \{1, \dots, n+1\}} \alpha_i \cdot \pi_{\mathcal{S}(\{a_{n+2}\})}(b_i) = \sum_{i \in \{1, \dots, n+1\}} \alpha_i \cdot \bar{b}_i$  proving that

$$\mathcal{S}(\{\bar{b}_i \mid i \in \{1, \dots, n+1\}\}) = V_1 \quad (10.48)$$

For  $a_1$  we have now the following cases to consider:

- A.  $(\sim_{\mathcal{S}(\{a_{n+2}\})}[a_1] = 0)$  as then neutral element in  $V_1$  is given by  $0 = \sim_{\mathcal{S}(\{a_{n+2}\})}[0]$  we have by 2.120 that  $a_1 \sim_{\mathcal{S}(\{a_{n+2}\})} 0 \Rightarrow a_1 = a_1 - 0 \in \mathcal{S}(\{a_{n+2}\}) \Rightarrow \exists \gamma \in F$  such that  $a_1 = \gamma \cdot a_{n+2}$  and using 10.146 we have then that  $A$  is linear dependent.

- B.  $(\sim_{\mathcal{S}(\{a_{n+2}\})}[a_1] \neq 0)$  from this it follows that  $V_1$  is a non trivial vector space. Using 10.48 together with 10.158 there exists a

$$J \subseteq \{1, \dots, n+1\} \text{ such that } \{\bar{b}_i\}_{i \in J} \text{ is a basis of } V_1 \quad (10.49)$$

As  $0 \neq a_{n+2} \in V$  there exists a  $\{\lambda_i\}_{i \in \{1, \dots, n+1\}}$  such that  $a_{n+2} = \sum_{i \in \{1, \dots, n+1\}} \lambda_i \cdot b_i$  where not all  $\lambda_i = 0$  [as  $a_{n+2} \neq 0$ ]. We have then as  $a_{n+2} \in \mathcal{S}(\{a_{n+2}\})$  that  $0 = \sim_{\mathcal{S}(\{a_{n+2}\})}[a_{n+2}] = \pi_{\mathcal{S}(\{a_{n+2}\})}(a_{n+2}) = \pi_{\mathcal{S}(\{a_{n+2}\})}(\sum_{i \in \{1, \dots, n+1\}} \lambda_i \cdot b_i) = \sum_{i \in \{1, \dots, n+1\}} \lambda_i \cdot \pi_{\mathcal{S}(\{a_{n+2}\})}(b_i) = \sum_{i \in \{1, \dots, n+1\}} \lambda_i \cdot \bar{b}_i$  proving that  $\{\bar{b}_i\}_{i \in \{1, \dots, n+1\}}$  is linear dependent and thus that

$$J \subset \{1, \dots, n+1\} \underset{5.46}{\Rightarrow} \#(J) < n+1 \Rightarrow \#(J) \leq n \quad (10.50)$$

As  $\#(J) = m \leq n$  there exists a bijection  $\sigma: \{1, \dots, m\} \rightarrow J$  such that (see 10.151).

$$\{\bar{b}_{\sigma_i}\}_{i \in \{1, \dots, m\}}$$
 is a basis of  $V_1$  where  $m \leq n \quad (10.51)$

If we define  $\bar{A} = \{\bar{a}_i\}_{i \in \{1, \dots, n+1\}}$  a family in  $V_1$  where  $\bar{a}_i = \pi_{\mathcal{S}(\{a_{n+2}\})}(a_i)$  then using 10.51 and  $n \in S$  we have that  $\bar{A} = \{\bar{a}_i\}_{i \in \{1, \dots, n+1\}}$  is linear dependent. So there exists a  $\{\varepsilon_i\}_{i \in \{1, \dots, n+2\}}$  with not every  $\varepsilon_i = 0$  such that  $0 = \sum_{i \in \{1, \dots, n+1\}} \varepsilon_i \cdot \bar{a}_i = \sum_{i \in \{1, \dots, n+1\}} \varepsilon_i \cdot \pi_{\mathcal{S}(\{a_{n+2}\})}(a_i) = \pi_{\mathcal{S}(\{a_{n+1}\})}(a_i) = \pi_{\mathcal{S}(\{a_{n+2}\})}(\sum_{i \in \{1, \dots, n+1\}} \varepsilon_i \cdot a_i)$  and this gives that  $\sum_{i \in \{1, \dots, n+1\}} \varepsilon_i \cdot a_i \in \mathcal{S}(\{a_{n+2}\})$ . So there exists a  $\mu \in F$  such that  $\sum_{i \in \{1, \dots, n+1\}} \varepsilon_i \cdot a_i = \mu \cdot a_{n+2} \Rightarrow \sum_{i \in \{1, \dots, n+1\}} \varepsilon_i \cdot a_i - \mu \cdot a_{n+2} = 0$ . if we define now  $\{\kappa_i\}_{i \in \{1, \dots, n+2\}}$  by  $\kappa_i = \begin{cases} \varepsilon_i & \text{if } i \in \{1, \dots, n+1\} \\ -\mu & \text{if } i = n+2 \end{cases}$  which is not all zeroes [as the  $\varepsilon_i$ 's are]. We have then  $\sum_{i \in \{1, \dots, n+2\}} \kappa_i \cdot a_i = \sum_{i \in \{1, \dots, n+1\}} \kappa_i \cdot a_i + \sum_{i \in \{n+2\}} \kappa_i \cdot a_i = \sum_{i \in \{1, \dots, n+1\}} \varepsilon_i \cdot a_i - \mu \cdot a_{n+2} = 0$  proving that  $A$  is linear dependent.

As  $A$  is linear dependent in all possible cases we have then  $n+1 \in S$ .

Using mathematical induction we have  $S = \{1, \dots\} = \mathbb{N}_0$  which proves our theorem.  $\square$

**Corollary 10.198.** *Let  $V$  be a non trivial vector space over a field  $F$  with a basis  $\{b_i\}_{i \in \{1, \dots, n\}}$  then every independent set  $A$  is finite and  $\#(A) \leq n$*

**Proof.** If  $A$  is infinite and independent then by 5.34 there exists a  $C \subseteq A$  with  $C \approx \mathbb{N}$  so there exists a bijection  $c: \mathbb{N} \rightarrow C$  and then  $\{c_i\}_{i \in \{1, \dots, n+1\}}$  is linear dependent by the previous axiom. Using 10.144 we have that  $\{c_i | i \in \{1, \dots, n+1\}\}$  is linear dependent. As  $\{c_i | i \in \{1, \dots, n+1\}\} \subseteq C \subseteq A$  this proves that  $A$  is linear dependent (see 10.149) contradicting the linear independence of  $A$ . So we must have that  $A$  is finite.

If now  $A$  is finite with  $m = \#(A) > n$  then there exists a bijection  $a: \{1, \dots, m\} \rightarrow A$ , as  $n < m$  we have  $n+1 \leq m$  and thus by the previous theorem we have that  $\{a_i\}_{i \in \{1, \dots, n+1\}}$  is linear dependent. Using 10.144 we have then that  $\{a_i | i \in \{1, \dots, n+1\}\}$  is linear dependent and thus by 10.149 we have as  $\{a_i | i \in \{1, \dots, n+1\}\} \subseteq A$  that  $A$  is linear dependent contradicting its linear independence. So we must have that  $\#(A) \leq n$ .  $\square$

**Definition 10.199.** *Let  $V$  be a vector space over a field  $F$  with a basis  $\{b_i\}_{i \in I}$  then a basis is finite if  $\{b_i | i \in I\}$  is finite. We define  $\#(\{b_i\}_{i \in I}) = \#(\{b_i | i \in I\})$  and call this number the number of elements in  $\{b_i\}_{i \in I}$ . Note that as for a basis  $\{b_i\}_{i \in I}$  we have that  $I$  and  $\{b_i | i \in I\}$  are bijective (see 10.143) we have that  $\{b_i | i \in I\}$  is finite if and only if  $I$  is finite and  $\#(\{b_i\}_{i \in I}) = \#(I)$ .*

**Theorem 10.200.** *Let  $V$  be a non trivial vector space over a field  $F$  with a finite basis  $\{b_i\}_{i \in I}$  such that  $\#\{b_i | i \in I\} = n$  then every basis  $\{c_j\}_{j \in J}$  of  $V$  is finite and  $\#(\{c_j\}_{j \in J}) \leq n$ .*

**Proof.** If  $\{b_i\}_{i \in I}$  is a finite basis with  $n$  elements then  $\#(I) = n$  so there exists a bijection  $h: \{1, \dots, n\} \rightarrow I$  and using 10.151 we have that  $\{b_{h(i)}\}_{i \in \{1, \dots, n\}}$  is a basis for  $V$ . Then if  $\{c_j\}_{j \in J}$  is another basis (and thus by 10.144  $\{c_j | j \in J\}$  is linear independent) we must using the previous corollary have that  $\{c_j | j \in J\}$  is finite and  $\#(c_j | j \in J) \leq n$ .  $\square$

**Corollary 10.201.** *Let  $V$  be a non trivial vector space over a field  $F$  with a finite basis  $\{b_i\}_{i \in I}$  then every other basis  $\{c_j\}_{j \in J}$  is finite and  $\#(\{b_i\}_{i \in I}) = \#(\{c_j\}_{j \in J})$*

**Proof.** If  $\{b_i\}_{i \in I}$  is a finite basis with  $n$  elements then by the previous theorem we have that  $\{c_j\}_{j \in J}$  is also finite and  $m = \#(\{c_j\}_{j \in J}) \leq n$ . Using the previous theorem again we must have as  $\{c_j\}_{j \in J}$  is finite that  $n = \#(\{b_i\}_{i \in I}) \leq m$  giving that finally we have  $n = m$ .  $\square$

The above corollary assures us then that the following definition make sense

**Definition 10.202.** *Let  $V$  be a vector space over a field  $F$  then  $\dim(V)$  is defined as follows*

1. If  $V = \{0\}$  ( $V$  is trivial) then  $\dim(V) = 0$

2. If  $V$  is non trivial and has a finite basis  $\{b_i\}_{i \in I}$  then  $\dim(V) = n$  (by the previous corollary this number is unique)
3. If  $V$  is non trivial and has a basis which is not finite (by the previous corollary we can then have no finite basis) then  $\dim(V) = \infty$ . Note that  $\dim(V) = \infty$  means  $V$  is infinite and non trivial, so we assume that  $\infty$  is different from every  $n \in \mathbb{N}$ .

$V$  is finite dimension if it has a finite basis or if it is trivial.

**Theorem 10.203.** Let  $V$  be a vector space over  $F$  then  $\dim(V) = 0$  is equivalent with  $V = \{0\}$

**Proof.** If  $V = \{0\}$  then by definition we have  $\dim(V) = 0$ . If  $\dim(V) = 0$  then by definition we can not have that  $V$  is infinite dimensional so  $V$  is finite dimensional. If now  $x \in V$  with  $x \neq 0$  then  $\{x\}$  is linear independent [if  $\lambda \cdot x = 0$  with  $\lambda \neq 0$  then  $0 = \lambda^{-1} \cdot (\lambda \cdot x) = (\lambda^{-1} \cdot \lambda) \cdot x = 1 \cdot x = x$  contradicting  $x \neq 0$ ]. By 10.159 there exists a  $B$  (which must be finite dimensional) such that  $x \in B$  and  $\{b\}_{b \in B}$  is a basis of  $V$ . Now as  $x \in B \Rightarrow \#(\{b\}_{b \in B}) = \#(B) > 0 \Rightarrow \dim(V) > 0$  contradicting  $\dim(V) = 0$ . So we must have that  $x = 0$  and this means that  $V = \{0\}$ .  $\square$

**Theorem 10.204.** Let  $n \in \mathbb{N}_0$  be  $\{V_i\}_{i \in \{1, \dots, n\}}$  be a finite family of finite dimensional vector spaces with  $\forall i \in \{1, \dots, n\} \dim(V_i) = n_i$  then  $\prod_{i=1}^n V_i$  (which is a vector space see 10.120) is a finite dimensional vector space with  $\dim(\prod_{i=1}^n V_i) = \sum_{i=1}^n n_i$

**Proof.** This follows from the construction of a basis for  $\prod_{i=1}^n V_i$  (see 10.168)  $\square$

**Theorem 10.205.** Let  $V$  be a vector space over  $F$  with  $n = \dim(V) \in \mathbb{N}_0$  (hence by the above theorem a non trivial vector space) then if  $\{e_i\}_{i \in \{1, \dots, m\}}$  with  $m \in \{1, \dots, n\}$  is a linear independent set in  $V$  then there exists a basis  $\{f_i\}_{i \in \{1, \dots, n\}}$  of  $V$  such that  $\forall i \in \{1, \dots, m\}$  we have  $e_i = f_i$

**Proof.** If  $m = n$  then the proof is trivial so we may assume that  $m < n$ . Take  $R = \{e_i \mid i \in \{1, \dots, m\}\} \subseteq V$  then using 10.145 we have that  $R$  is linear independent, using 10.156 together with  $\mathcal{S}(V) = V$  we have that there exists a set  $B$  with  $R \subseteq B$  such that  $\{b\}_{b \in B}$  is a basis of  $V$ . As  $\dim(V) = n$  we have that  $\#(B) = n$  hence using 5.45 we have that  $\#(K) = \#(B \setminus R) = \#(B) - \#(R) = n - m > 0$ . So there exists a  $\beta: \{1, \dots, m - n\} \rightarrow B \setminus R$  hence if we define  $f: \{1, \dots, n\} \rightarrow B$  by  $f_i = f(i) = \begin{cases} e_i & \text{if } i \in \{1, \dots, n - m\} \\ \beta((i - (n - m)) \text{ if } i \in \{(n - m) + 1, \dots, n\} \end{cases}$  we have  $f(\{1, \dots, n\}) \subseteq R \cup (B \setminus R) = B$  so that  $\mathcal{S}(\{f_i\}_{i \in \{1, \dots, n\}}) = \mathcal{S}(\{f_i \mid i \in \{1, \dots, n\}\}) = \mathcal{S}(f(\{1, \dots, n\})) = \mathcal{S}(B) = V$  and as  $B$  is independent we have by 10.145 that  $\{f_i\}_{i \in \{1, \dots, n\}}$  is linear independent. This proves that  $\{f_i\}_{i \in \{1, \dots, n\}}$  is a basis and that  $\forall i \in \{1, \dots, m\}$  we have  $e_i = f_i$ .  $\square$

The above theorem and definition suggest the following definition of equality of dimensions of a vector space (taking in account the  $\infty$ ).

**Definition 10.206.** If  $V$  is a vector space over a field  $F$  and  $W$  a vector space over a field  $G$  then  $\dim(V) = \dim(W)$  if and only if

1.  $V$  and  $W$  are both finite dimensional and  $\dim(V) = \dim(W)$
2.  $V$  and  $W$  are both infinite dimensional

As a consequence of this definition we have the following

**Remark 10.207.** If  $V$  and  $W$  are vector spaces and  $\dim(V) = \dim(W)$  then we have

1. If  $V$  is trivial then  $W$  is trivial
2. If  $V$  is non trivial and finite dimensional then  $W$  is non trivial and finite dimensional with the same number of elements in its basis as  $V$
3. If  $V$  is infinite dimensional then  $W$  is infinite dimensional

**Theorem 10.208.** If  $V$  is a finite dimensional vector space over a field  $F$  with  $\dim(V) = n \in \mathbb{N}_0$  then there exists a basis  $\{e_i\}_{i \in \{1, \dots, n\}}$  of  $V$

**Proof.** As  $\dim(V) \neq 0$   $V$  is not trivial and by definition there exists a basis  $\{f_i\}_{i \in I}$  with  $\#(I) = n$  and as  $\#(I) = n$  there exists a bijection  $b: \{1, \dots, n\} \rightarrow I$  then by 10.151 we have that  $\{f_{b(i)}\}_{i \in \{1, \dots, n\}}$  is a basis for  $V$ . So if we take  $\{e_i\}_{i \in \{1, \dots, n\}}$  where  $e_i = f_{b(i)}$  we have proved our theorem.  $\square$

**Theorem 10.209.** Let  $n \in \mathbb{N}_0$ ,  $V$  a  $n$ -dimensional vector space over a field  $F$  and  $\{e_i\}_{i \in \{1, \dots, n\}}$  a family of independent vectors then  $\{e_i\}_{i \in \{1, \dots, n\}}$  is a basis of  $V$

**Proof.** Using 10.144 we have that  $E = \{e_i | i \in \{1, \dots, n\}\}$  is a linear independent set. Then by 10.159 we have the existence of a set  $B$  with  $E \subseteq B$  and  $B$  is a basis of  $V$  so that  $\mathcal{S}(B) = V$ . Take now  $e: \{1, \dots, n\} \rightarrow E$  by  $i \rightarrow e_i$  then we have by 10.143 that  $e$  is a bijection and thus that  $\#(E) = n$ . Now if  $E \subset B$  then by 5.46 we have that  $n = \#(E) < \#(B) \stackrel{\dim(V)=0}{=} n$  giving the contradiction  $n < n$ , so we must have  $E = B$  and thus  $\mathcal{S}(E) = \mathcal{S}(B) = V$  proving together with the linear independence of  $\{e_i\}_{i \in \{1, \dots, n\}}$  that  $\{e_i\}_{i \in \{1, \dots, n\}}$  is a basis of  $V$ .  $\square$

**Theorem 10.210.** Let  $V$  be a finite dimensional vector space over  $F$  and  $\emptyset \neq W \subseteq V$  a sub space of  $V$  then we have that  $W$  is finite dimensional and  $\dim(W) \leq \dim(V)$

**Proof.** As  $W \neq 0$  we have that  $W$  is indeed a vector space. Now we have the following cases for  $\dim(W)$ :

1. ( $\dim(W) = 0$ ) then we have that  $V = \{0\}$  and thus we must have that  $W = \{0\}$  giving  $\dim(W) = 0$  so that  $W$  is finite and  $\dim(W) \leq \dim(V)$
2. ( $\dim(W) \neq 0$ ) we have now the following cases for  $W$ 
  - a. ( $W = \{0\}$ ) then  $\dim(W) = 0 \leq \dim(V)$  giving  $W$  is finite and  $\dim(W) \leq \dim(V)$
  - b. ( $W \neq \{0\}$ ) if now  $\{b_i\}_{i \in I}$  is a basis for  $W$  then by 10.159 then there exists a  $B$  such that  $\{b_i | i \in I\} \subseteq B$  and  $\{b\}_{b \in B}$  is a basis for  $V$ . From the finite dimensionality of  $V$  it follows that  $\{b_i | i \in I\}$  is finite so that by definition  $\{b_i\}_{i \in I}$  is finite. Also using 5.45 we have that  $\dim(W) = \#(\{b_i | i \in I\}) \leq \#(B) = \dim(V)$  so we have  $W$  is finite and  $\dim(W) \leq \dim(V)$   $\square$

**Theorem 10.211.** Let  $V$  be a vector space over  $F$  and let  $A = \mathcal{S}(\{a_1, \dots, a_n\})$ ,  $n \in \mathbb{N}_0$  then  $A$  is finite dimensional and  $\dim(A) \leq n$

**Proof.** We have to consider the following cases:

1. ( $A = \mathcal{S}(\{0\})$ ) as  $\mathcal{S}(\{0\}) = \{0\}$  we have  $\dim(A) = 0 \leq n$
2. ( $A = \mathcal{S}(\{a_1, \dots, a_n\})$  with not every  $a_i = 0$ ) then  $\mathcal{S}(\{a_1, \dots, a_n\})$  is non trivial and as  $\mathcal{S}(\{a_1, \dots, a_n\}) = A$  there exists by 10.157 a set  $B \subseteq \{a_1, \dots, a_n\}$  such that  $\{b\}_{b \in B}$  forms a basis of  $A$ . Using 5.45 we have that  $B$  is finite and  $\dim(A) = \#(B) \leq n$   $\square$

**Definition 10.212.** Let  $V$  and  $W$  be vector spaces over the field  $F$  and  $L: V \rightarrow W$  be a linear mapping then  $\ker(L) = \{x \in V | L(x) = 0\} = L^{-1}(\{0\})$ . We call  $\ker(L)$  the kernel of  $L$ . Note that as  $L(0) = 0$  we have  $0 \in \ker(L)$

**Theorem 10.213.** Let  $V$  and  $W$  be vector spaces over the field  $F$  and let  $L: V \rightarrow W$  be a linear mapping then  $L$  is regular (see 10.174) if and only if  $\ker(L) = \{0\}$

**Proof.** if  $\ker(L) = \{0\}$  then if  $x, y \in W$  such that  $L(x) = L(y) \Rightarrow 0 = L(x) - L(y) = L(x - y) \Rightarrow x - y \in \ker(L) = \{0\} \Rightarrow x - y = 0 \Rightarrow x = y$ , proving the regularity of  $L$ . If  $L$  is injective and  $x \in \ker(L)$  then  $0 = L(x) \stackrel{L(0)=0}{\Rightarrow} L(0) = L(x) \Rightarrow x = 0$  if  $L$  is injective so that  $\ker(L) = \{0\}$   $\square$

**Theorem 10.214.** Let  $V, W$  be vector spaces over a field  $F$  with  $\dim(V) = n \in \mathbb{N}_0$  (so  $V$  is finite dimensional and non trivial),  $L: V \rightarrow W$  a linear mapping and  $\{b_i\}_{i \in I}$  a basis for  $V$  then we have that  $L$  is regular if and only if  $\{L(b_i)\}_{i \in I}$  is linear independent. Furthermore if  $\dim(W) = n$  then we have that  $L$  is regular if and only if  $L$  is a isomorphism.

**Proof.**

1. If  $L$  is regular (=injective) then if  $\{\alpha_i\}_{i \in I}$  is a family with finite support such that  $0 = \sum_{i \in I} \alpha_i \cdot L(b_i) = \sum_{i \in I} L(\alpha_i \cdot b_i)$   $\stackrel{L \text{ is linear and 10.176}}{=} L(\sum_{i \in I} \alpha_i \cdot b_i) \Rightarrow \sum_{i \in I} \alpha_i \cdot b_i \in \ker(L)$   $\stackrel{L \text{ is regular and previous theorem}}{=} \{0\} \Rightarrow \sum_{i \in I} \alpha_i \cdot b_i \stackrel{\{\alpha_i\}_{i \in I} \text{ is a basis}}{\Rightarrow} \forall i \in I \text{ we have } \alpha_i = 0$  proving that  $\{L(b_i)\}_{i \in I}$  is linear independent.
2. If  $\{L(b_i)\}_{i \in I}$  is linear independent then if  $x \in \ker(L)$  we have  $L(x) = 0$ , as  $x \in V$  and  $\{b_i\}_{i \in I}$  is a basis for  $V$  there exists a  $\{\alpha_i\}_{i \in I}$  with finite support such that  $x = \sum_{i \in I} \alpha_i \cdot b_i \Rightarrow 0 = L(x) = L(\sum_{i \in I} \alpha_i \cdot b_i) = \sum_{i \in I} \alpha_i \cdot L(b_i) \stackrel{\{L(b_i)\}_{i \in I} \text{ is linear independent}}{\Rightarrow} \forall i \in I \text{ we have } \alpha_i = 0 \Rightarrow x = \sum_{i \in I} \alpha_i \cdot b_i = \sum_{i \in I} 0 \cdot b_i = 0$  so that  $\ker(L) = \{0\}$  proving using the above theorem that  $L$  is regular.
3. If  $L$  is injective (=regular) then  $L: \{b_i | i \in I\} \rightarrow \{L(b_i) | i \in I\} = L(\{b_i | i \in I\})$  is a bijection and thus  $\dim(V) = n = \#(\{b_i | i \in I\}) = \#(\{L(b_i) | i \in I\})$ . By (1) we have that  $\{L(b_i)\}_{i \in I}$  is linear independent and using 10.159 there exists a  $B'$  such that  $\{L(b_i) | i \in I\} \subseteq B'$  and  $\{b\}_{b \in B'}$  is a basis for  $W$  and thus  $\#(B') = \dim(W) = n$ . If now  $\{L(b_i) | i \in I\} \subset B'$  then by 5.46 we have  $n = \#(\{L(b_i) | i \in I\}) < \#(B') = n$  a contradiction. So we have  $\{L(b_i) | i \in I\} = B'$  giving that  $\mathcal{S}(\{L(b_i) | i \in I\}) = \mathcal{S}(B') = W$  and this together with the linear independence of  $\{L(b_i)\}_{i \in I}$  proves that  $\{L(b_i)_{i \in I}\}$  is a basis of  $W$ . So if  $y \in W$  then there exists a family  $\{\alpha_i\}_{i \in I}$  with finite support such that  $y = \sum_{i \in I} \alpha_i \cdot L(b_i) = L(\sum_{i \in I} \alpha_i \cdot b_i) \in L(V)$  proving that  $W \subseteq L(V)$  and thus that  $L$  is surjective. Together with injectivity and linearity this means that  $L$  is a isomorphism.
4. If  $L$  is a isomorphism then evidently it is also injective.  $\square$

**Theorem 10.215.** Let  $V$  and  $W$  be vector spaces over a field  $F$  and  $L: V \rightarrow W$  a linear mapping then  $\ker(L)$  is a subspace of  $V$ . As  $0 \in \ker(L)$  this also implies that  $\ker(L)$  is a vector space.

**Proof.** If  $x, y \in \ker(L)$  and  $\alpha, \beta \in F$  then  $L(\alpha \cdot x + \beta \cdot y) = \alpha \cdot L(x) + \beta \cdot L(y) = \alpha \cdot 0 + \beta \cdot 0 = 0$  so  $\alpha \cdot x + \beta \cdot y \in \ker(L)$   $\square$

**Theorem 10.216.** Let  $V$  and  $W$  be vector spaces over a field  $F$  and  $L: V \rightarrow W$  a linear mapping then  $L(V)$  is a subspace of  $W$ . As  $0 = L(0) \in L(V)$  this also implies that  $L(V)$  is a vector space.

**Proof.** If  $x, y \in L(V)$  and  $\alpha, \beta \in F$  then  $\exists x', y' \in V$  such that  $x = L(x'), y = L(y')$  then  $L(V) \ni L(\alpha \cdot x' + \beta \cdot y') = \alpha \cdot L(x') + \beta \cdot L(y') = \alpha \cdot x + \beta \cdot y$   $\square$

**Definition 10.217.** Let  $V, W$  be vector spaces over a field  $F$  then  $\text{rank}(L) = \dim(L(V))$

**Theorem 10.218.** Let  $V, W$  be vector spaces over a field  $F$  and  $L: V \rightarrow W$  a linear mapping then if  $V$  is finite dimensional we have that  $L(V)$  is finite dimensional and  $\dim(L(V)) = \text{rank}(L) \leq \dim(V)$ . Furthermore if  $W$  is finite dimensional then  $\text{rank}(L) \leq \dim(W)$

**Proof.** First we consider the trivial case where  $\dim(V) = 0$ , then  $V = \{0\}$  and thus  $L(V) = \{0\} \Rightarrow \text{rank}(L) = \dim(L(V)) = 0 \leq \dim(V) = 0$ , also if  $W$  is finite dimensional then  $\text{rank}(L) = 0 \leq \dim(W)$ . Second take the non trivial case and assume that  $\{b_i\}_{i \in I}$  is a basis of  $V$  with  $\#(\{b_i | i \in I\}) = n$  then if  $y \in L(V)$  there exists a  $x \in V$  such that  $y = L(x)$  and a  $\{\alpha_i\}_{i \in I}$  with finite support such that  $x = \sum_{i \in I} \alpha_i \cdot b_i$  so that  $y = L(x) = L(\sum_{i \in I} \alpha_i \cdot b_i) = \sum_{i \in I} \alpha_i \cdot L(b_i)$  proving that

$$y \in L(V) \text{ then } \exists \{\alpha_i\}_{i \in I} \text{ with finite support such that } y = \sum_{i \in I} \alpha_i \cdot L(b_i) \quad (10.52)$$

Now given the fact that  $L: \{b_i | i \in I\} \rightarrow L(\{b_i | i \in I\}) = \{L(b_i) | i \in I\}$  is a surjection and the finiteness of  $\{b_i | i \in I\}$  we have by 5.48 that

$$L(\{b_i | i \in I\}) \text{ is finite and } \#(\{L(b_i) | i \in I\}) \leq \#(\{b_i | i \in I\}) = \dim(V) \quad (10.53)$$

So let  $m = \#(\{L(b_i) | i \in I\})$  then  $m \leq \dim(V)$  and there exists a bijection  $h: \{0, \dots, m-1\} \rightarrow \{L(b_i) | i \in I\}$ . Define then the family  $\{I_i\}_{i \in \{0, \dots, m-1\}}$  where  $I_i = \{j \in I | L(b_j) = h(i)\}$ . If now  $i, j \in \{0, \dots, m-1\}$  with  $i \neq j$  then if  $k \in I_i \cap I_j$  we have  $L(b_k) = h(i), h(j) \Rightarrow h(i) = h(j) \Rightarrow i = j$  a contradiction so we must have that  $I_i \cap I_j = \emptyset$ . Also as  $I_i \subseteq I$  we have  $\bigcup_{i \in \{0, \dots, m-1\}} I_i \subseteq I$ . If  $k \in I$  then  $L(b_k) \in \{L(b_i) | i \in I\} \Rightarrow h^{-1}(L(b_k)) \in \{0, \dots, m-1\}$  and  $h(h^{-1}(L(b_k))) = L(b_k) \Rightarrow k \in I_{h^{-1}(L(b_k))} \in \bigcup_{i \in \{0, \dots, m-1\}} I_i$  proving that  $I \subseteq \bigcup_{i \in \{0, \dots, m-1\}} I_i$  and finally that  $I = \bigcup_{i \in \{0, \dots, m-1\}} I_i$  where  $I_i$  are disjoint sets. If now  $y \in L(V)$  then by 10.52 there exists a  $\{\alpha_i\}_{i \in I}$  such that  $y = \sum_{i \in I} \alpha_i \cdot L(b_i)$  10.45  $= \sum_{i=0}^{m-1} (\sum_{k \in I_i} \alpha_k \cdot L(b_k)) = \sum_{i=0}^{m-1} (\sum_{k \in I_i} \alpha_k \cdot h(i)) = \sum_{i=0}^{m-1} (\sum_{k \in I_i} \alpha_k) \cdot h(i) = \sum_{i=0}^{m-1} \beta_{h(i)} \cdot h(i) = \sum_{v \in \{L(b_i) | i \in I\}} \beta_v \cdot v$  if we define  $\{\beta_v\}_{v \in \{L(b_i) | i \in I\}}$  by  $\beta_v = \left( \sum_{k \in I_{h^{-1}(v)}} \alpha_k \right)$  so that  $\beta_{h(i)} = \left( \sum_{k \in I_{h^{-1}(h(i))}} \alpha_k \right) = \sum_{k \in I_i} \alpha_k$ . This proves that  $y \in \mathcal{S}(\{L(b_i) | i \in I\})$  and thus

$$L(V) \subseteq \mathcal{S}(\{L(b_i) | i \in I\}) \quad (10.54)$$

As  $L(\{b_i | i \in I\}) \subseteq V$  we have by 10.138 that  $\mathcal{S}(\{b_i | i \in I\}) \subseteq \mathcal{S}(L(V))$   $L(V) \text{ is a vectorspace} = L(V)$  and this gives with 10.54 that

$$L(V) = \mathcal{S}(\{L(b_i) | i \in I\}) \quad (10.55)$$

Using 10.157 there exists a  $E \subseteq \{L(b_i) | i \in I\}$  such that  $\{e\}_{e \in E}$  is a basis of  $L(V)$  using 5.45 and 10.53 we have that  $E$  is finite and  $\dim(L(V)) = \#(E) \leq \#(\{L(b_i) | i \in I\}) \leq \dim(V)$ . This proves that  $L(V)$  is finite dimensional and that

$$\dim(L(V)) = \text{rank}(L) \leq \dim(V) \quad (10.56)$$

Lastly if  $W$  is finite dimensional and as  $L(V)$  is a subspace of  $W$  we have by 10.210 that

$$\text{rank}(L) = \dim(L(V)) \leq \dim(W)$$

□

**Theorem 10.219.** *Let  $V, W$  be vector spaces over a field  $F$ ,  $V$  is finite dimensional and  $L: V \rightarrow W$  a linear mapping then  $L$  is surjective if and only if  $W$  is finite dimensional and  $\text{rank}(L) = \dim(W)$*

**Proof.** If  $L$  is surjective then  $L(V) = W$  and by using the previous theorem we have that  $L(V)$  and thus  $W$  is finite dimensional and  $\text{rank}(L) = \dim(L(V)) = \dim(W)$ .

If  $W$  is finite dimensional and  $\text{rank}(L) = \dim(W)$  then  $\dim(L(V)) = \dim(W)$ , we must then consider the following two cases:

1. ( $\dim(W) = 0$ ) then  $W = \{0\}$  and we must have  $L(V) = \{0\}$  so  $L: \{0\} \rightarrow \{0\}$  is as  $L(0) = 0$  surjective.
2. ( $\dim(W) \neq 0$ ) then  $0 \neq \text{rank}(L) = \dim(L(V))$  so there exists a basis  $\{b_i\}_{i \in I}$  of  $L(V)$  such that  $\#(\{b_i | i \in I\}) = \dim(W)$ . Using 10.159 there exists a  $B$  such that  $\{b_i | i \in I\} \subseteq B$  and  $\{b\}_{b \in B}$  is a basis of  $W$ . If now  $\{b_i | i \in I\} \subset B$  then by 5.46 we would have the contradiction  $\dim(W) = \#(\{b_i | i \in I\}) < \#(B) = \dim(W)$  so we must have that  $\{b_i | i \in I\} = B$ . So  $L(V) = \mathcal{S}(\{b_i | i \in I\}) = \mathcal{S}(B) = W$  proving surjectivity. □

**Theorem 10.220.** *Let  $V, W$  be vector spaces over a field  $F$ ,  $V$  is finite dimensional and  $L: V \rightarrow W$  a linear mapping, then  $\dim(V) = \dim(\ker(L)) + \text{rank}(L)$*

**Proof.** We divide the proof in two cases:

1. ( $\ker(L) = \{0\}$ ) then by 10.213 we have that  $L$  is injective, also we have that  $\dim(\ker(L)) = 0$ . Let now  $\{b_i\}_{i \in I}$  be a basis of  $V$  then we have  $\dim(V) = \#(\{b_i | i \in I\})$ . Using 10.214 we have then by regularity of the basis that  $\{L(b_i)\}_{i \in I}$  is linear independent. If  $y \in L(V)$  then there exists a  $x \in V$  such that  $y = L(x)$  and a  $\{\alpha_i\}_{i \in I}$  with finite support such that  $x = \sum_{i \in I} \alpha_i \cdot b_i$  and thus we have  $y = L(\sum_{i \in I} \alpha_i \cdot b_i) = \sum_{i \in I} \alpha_i \cdot L(b_i)$  proving that  $L(V) \subseteq \mathcal{S}(\{L(b_i) | i \in I\}) \subseteq L(V) \Rightarrow L(V) = \mathcal{S}(\{L(b_i) | i \in I\})$ . This proves that  $\{L(b_i)\}_{i \in I}$  is a basis for  $L(V)$  and thus  $\text{rank}(L) = \dim(L(V)) = \#(\{b_i | i \in I\}) = \dim(V) = \dim(W)$ . As  $\dim(\ker(L)) = 0$  we have then finally  $\dim(\ker(L)) + \text{rank}(L) = \dim(V)$ .

2. ( $\ker(L) \neq \{0\}$ ) Let  $\{b_i\}_{i \in I}$  be a basis for  $\ker(L)$ , define then  $B = \{b_i | i \in I\}$ , and using 10.159 extend it to a  $B'$  such that  $B \subseteq B'$  and  $\{w\}_{w \in B'}$  is a basis for  $V$ . As  $V$  is finite dimensional we have that  $B'$  and thus also  $B$  and  $B' \setminus B$  is finite, furthermore we have (see 5.45) that

$$\#(B' \setminus B) + \#(B) = \#(B') \quad (10.57)$$

We prove now that

$$L|_{B' \setminus B}: B' \setminus B \rightarrow L(B' \setminus B) \text{ is a bijection} \quad (10.58)$$

**Proof.**

- a. (**injectivity**) If  $x, y \in B' \setminus B$  is such that  $L|_{B' \setminus B}(x) = L|_{B' \setminus B}(y) \Rightarrow L(x) = L(y) \Rightarrow L(x) - L(y) = 0 \Rightarrow L(x - y) = 0 \Rightarrow x - y \in \ker(L)$ . As  $B$  is a basis for  $\ker(L)$  (see 10.144 and  $\{b_i\}_{i \in I}$  is a basis of  $\ker(L)$ ). There exists a family  $\{\alpha_w\}_{w \in B}$  such that  $x - y = \sum_{w \in B} \alpha_w \cdot w$  and thus

$$0 = \sum_{w \in B} \alpha_w \cdot w + 1 \cdot y + (-1) \cdot x \quad (10.59)$$

Assume now that  $x \neq 0$  and define using the fact that  $x, y \notin B$  the family  $\{\lambda_w\}_{w \in B'}$  by

$$\lambda_w = \begin{cases} \alpha_w & \text{if } w \in B \\ 1 & \text{if } w \in \{y\} \\ -1 & \text{if } w \in \{x\} \\ 0 & \text{if } w \in B' \setminus (B \cup \{x, y\}) \end{cases} \quad \text{then we have } \sum_{w \in B'} \lambda_w \cdot w = \sum_{w \in B} \lambda_w \cdot w + \sum_{w \in \{y\}} \lambda_w \cdot w.$$

$$w + \sum_{w \in \{x\}} \lambda_w \cdot w + \sum_{w \in B' \setminus (B \cup \{x, y\})} \lambda_w \cdot w = \sum_{w \in B} \alpha_w \cdot w + 1 \cdot y + (-1) \cdot x + 0 \stackrel{10.59}{=} 0.$$

Using the fact that  $B'$  is a basis we must have that  $\forall w \in W \models \lambda_w = 0 \Rightarrow \forall w \in B$  we have  $\alpha_w = \lambda_w = 0$  proving that  $x - y = 0 \Rightarrow x = y$  contradicting the assumption that  $x \neq y$ . So we must have that  $x = y$  proving that  $L|_{B' \setminus B}$  is bijective.

- b. (**surjectivity**) This is trivial □

Now if  $y \in L(V)$  there exists a  $x \in V$  and  $\{\alpha_w\}_{w \in B'}$  such that  $y = L(x)$  and  $x = \sum_{w \in B'} \alpha_w \cdot w$  so that  $y = L(\sum_{w \in B'} \alpha_w \cdot w) = \sum_{w \in B'} \alpha_w \cdot L(w) = \sum_{w \in B' \setminus B} \alpha_w \cdot L(w) + \sum_{w \in B} \alpha_w \cdot L(w) = \sum_{w \in B' \setminus B} \alpha_w \cdot w + \sum_{w \in B} \alpha_w \cdot 0 = \sum_{w \in B' \setminus B} \alpha_w \cdot L(w) = \sum_{w \in B' \setminus B} \alpha'_{L(w)} \cdot L(w) = \sum_{w \in L(B' \setminus B)} \alpha'_w \cdot w$  [where  $\{\alpha'_w\}_{w \in L(B' \setminus B)}$  is defined by  $\alpha'_w = \alpha_{L^{-1}_{|B' \setminus B}(w)}$  so that  $\alpha'_{L(w)} = \alpha_{L^{-1}_{|B' \setminus B}(L(w))} = \alpha_w$ ]. This proves that

$$L(V) \subseteq \mathcal{S}(L(B' \setminus B)) \quad (10.60)$$

As we have also  $B' \setminus B \subseteq V \Rightarrow L(B' \setminus B) \subseteq L(V) \Rightarrow \mathcal{S}(L(B' \setminus B)) \subseteq \mathcal{S}(L(V))$   $\underset{L(V) \text{ is a vectorspace}}{=} L(V)$  and this together with 10.60 gives

$$L(V) = \mathcal{S}(L(B' \setminus B)) \quad (10.61)$$

Now if  $\{\lambda_w\}_{w \in L(B' \setminus B)}$  is a family in  $F$  with finite support such that  $0 = \sum_{w \in L(B' \setminus B)} \lambda_w \cdot w$   $\underset{L|_{B' \setminus B}: B' \setminus B \rightarrow L(B' \setminus B) \text{ is a bijection}}{=}$   $\sum_{w \in B' \setminus B} \lambda_{L|_{B' \setminus B}(w)} \cdot L|_{B' \setminus B}(w) = \sum_{w \in B' \setminus B} \lambda_{L(w)} \cdot L(w) = L(\sum_{w \in B' \setminus B} \lambda_{L(w)} \cdot w)$  so  $\sum_{w \in B' \setminus B} \lambda_{L(w)} \cdot w \in \ker(L)$ , so there exists a  $\{\beta_w\}_{w \in B}$  such that  $\sum_{w \in B' \setminus B} \lambda_{L(w)} \cdot w = \sum_{w \in B} \beta_w \cdot w \Rightarrow \sum_{w \in B' \setminus B} \lambda_{L(w)} \cdot w + \sum_{w \in B} (-\beta_w) \cdot w = 0$ . Define now  $\{\gamma_w\}_{w \in B'}$  by  $\gamma_w = \begin{cases} \lambda_{L(w)} & \text{if } w \in B' \setminus B \\ -\beta_w & \text{if } w \in B \end{cases}$  then  $\sum_{w \in B'} \gamma_w \cdot w = \sum_{w \in B' \setminus B} \gamma_w \cdot w + \sum_{w \in B} \gamma_w \cdot w = \sum_{w \in B' \setminus B} \lambda_{L(w)} \cdot w + \sum_{w \in B} (-\beta_w) \cdot w = 0$  and by the independence of  $B'$  we must have  $\forall w \in B'$  that  $\gamma_w = 0 \Rightarrow \forall w \in B' \setminus B$  we have  $\lambda_{L(w)} = 0$ . If now  $v \in L(B' \setminus B)$  then  $\exists v' \in B' \setminus B$  such that  $v = L(v')$  and  $\lambda_v = \lambda_{L(v')} = 0$  proving that

$$L(B' \setminus B) \text{ is linear independent} \quad (10.62)$$

Using 10.61 and 10.62 we have that  $L(B' \setminus B)$  is a basis for  $L(V)$  and thus we have  $\dim(L(V)) = \#(L(B' \setminus B)) \underset{L|_{B' \setminus B}: B' \setminus B \rightarrow L(B' \setminus B) \text{ is a bijection}}{=} \#(B' \setminus B)$  and thus  $\dim(L(V)) + \dim(\ker(L)) = \#(B' \setminus B) + \#(B) \underset{10.57}{=} \#(B') = \dim(V)$  so we have finally

$$\dim(L(V)) + \dim(\ker(L)) = \dim(V) \quad \square$$

or

$$\text{rank}(L) + \dim(\ker(L)) = \dim(V) \quad (10.63)$$

**Corollary 10.221.** *Let  $V, W$  be finite dimensional vector spaces over a field  $F$  and  $L: V \rightarrow W$  a regular linear mapping then  $\dim(V) \leq \dim(W)$*

**Proof.**  $\dim(V) = \dim(\ker(L)) + \text{rank}(L) \underset{\ker(L) = \{0\}}{=} 0 + \text{rank}(L) = \text{rank}(L) \leq \dim(W)$   $\square$

**Corollary 10.222.** *Let  $V, W$  be vector spaces over a field  $F$ ,  $V$  is finite dimensional and  $L \in \text{Hom}(V, W)$  then  $L$  is injective if and only if  $\text{rank}(L) = \dim(V)$ . Furthermore  $L$  is a isomorphism if and only if  $\dim(W) = \dim(V)$  and  $\text{rank}(L) = \dim(V)$*

**Proof.**

1. Assume that  $L$  is injective then  $\ker(L) = 0 \Rightarrow \dim(V) = \text{rank}(L)$
2. Assume that  $\text{rank}(L) = \dim(V)$  then by 10.220 we have that  $\ker(L) = 0$  which by 10.203 means that  $\ker(L) = \{0\}$  proving that  $L$  is regular.
3. Assume that  $L$  is a isomorphism then  $L$  is surjective and by 10.219 we have that  $W$  is finite dimensional and  $\text{rank}(L) = \dim(W)$ . From injectivity and (1) we have that  $\dim(V) = \text{rank}(L)$  and thus  $\dim(W) = \dim(V)$ .
4. Assume that  $\dim(W) = \dim(V)$  and  $\text{rank}(L) = \dim(V)$ . Form  $\dim(V) = \dim(W)$  and the fact that  $V$  is finite dimensional we must have that  $W$  is finite dimensional. From  $\text{rank}(L) = \dim(V) = \dim(W)$  we have by 10.219 that  $L$  is surjective. By (2) we have also that  $L$  is injective and thus  $L$  is bijective and thus a isomorphism.  $\square$

**Theorem 10.223.** *Let  $U, V$  and  $W$  be vector spaces over a field  $F$  where  $V$  is finite dimensional then we have:*

1. *If  $L: U \rightarrow V$  is a linear mapping and  $H: V \rightarrow W$  is a regular linear mapping then  $\text{rank}(L) = \text{rank}(H \circ L)$*
2. *If  $L: U \rightarrow V$  is a isomorphism and  $H: V \rightarrow W$  is a linear mapping then  $\text{rank}(H) = \text{rank}(H \circ L)$*

**Proof.**

1. If  $H: V \rightarrow W$  is injective and linear this means that  $H|_{L(U)}: L(U) \rightarrow H(L(U))$  is a isomorphism (by 10.216  $L(U)$  is a vector space). Using 10.222 and the fact that  $L(U)$  is finite dimensional (see 10.210 and the finite dimensionality of  $V$ ) we have then that  $\dim(L(U)) = \dim(H(L(U)))$  so that  $\text{rank}(H \circ L) = \dim(H(L(U))) = \dim(L(U)) = \text{rank}(L)$
2. If  $x \in \ker(H \circ L) \Rightarrow H(L(x)) = 0 \Rightarrow L(x) \in \ker(H) \Rightarrow x \in L^{-1}(\ker(H))$  then

$$\ker(H \circ L) \subseteq L^{-1}(\ker(H)) \quad (10.64)$$

Also if  $x \in L^{-1}(\ker(H)) \Rightarrow L(x) \in \ker(H) \Rightarrow H(L(x)) = 0 \Rightarrow x \in \ker(H \circ L)$  this gives together with 10.64 that

$$\ker(H \circ L) = L^{-1}(\ker(H)) \quad (10.65)$$

From the fact that  $L$  is a bijection we have that  $L|_{L^{-1}(\ker(H))}: L^{-1}(\ker(H)) = \ker(H \circ L) \rightarrow L(L^{-1}(\ker(H))) = \ker(H)$  is a bijection and as  $\ker(H \circ L)$  and  $\ker(H)$  are vector spaces we have that  $L|_{\ker(H \circ L)} = L|_{L^{-1}(\ker(H))}: \ker(H \circ L) \rightarrow \ker(H)$  is a isomorphism. Using 10.222 we have then that

$$\dim(\ker(H \circ L)) = \dim(\ker(H)) \quad (10.66)$$

Now as by 10.183  $L^{-1}: V \rightarrow U$  is a isomorphism and using 10.222 and the finite dimensionality of  $V$  we have

$$\dim(V) = \dim(U) \text{ (so } U \text{ is finite)} \quad (10.67)$$

Now as  $V$  is finite dimensional we can use 10.220 to get

$$\dim(\ker(H)) + \text{rank}(H) = \dim(V) \quad (10.68)$$

From the finiteness of  $U$  (see 10.68) and the linearity (see 10.184) of  $H \circ L: U \rightarrow W$  and again 10.220 we have

$$\dim(\ker(H \circ L)) + \text{rank}(H \circ L) = \dim(U) \quad (10.69)$$

So using 10.69, 10.66 we have  $\dim(\ker(H)) + \text{rank}(H \circ L) = \dim(U) = \dim(V)$  giving that

$$\text{rank}(H \circ L) = \text{rank}(H) \quad \square$$

**Theorem 10.224.** *Given  $n, m \in \mathbb{N}$ ,  $F$  a field and  $F^n$  a vector spaces over  $F$  (defined in 10.118),  $Y$  a vector space over  $F$  and  $\sigma \in P_n$  (see 10.64 then if  $L: F^n \rightarrow Y$  is a linear mapping then we have the following:*

1.  $\sigma L$  is a linear mapping (see 10.84)
2. If  $L$  is a isomorphism then  $\sigma L$  is a isomorphism

**Proof.**

1. **(linearity)** If  $x = (x_1, \dots, x_n) \in F^n$ ,  $y = (y_1, \dots, y_n) \in F^n$  [so actually  $x: \{1, \dots, n\} \rightarrow F$ ,  $y: \{1, \dots, n\} \rightarrow F$ ] and  $\alpha, \beta \in F$  then  $\forall i \in \{1, \dots, n\}$  we have  $((\alpha \cdot x + \beta \cdot y) \circ \sigma)(i) = (\alpha \cdot x + \beta \cdot y)(\sigma(i)) = \alpha \cdot x(\sigma(i)) + \beta \cdot y(\sigma(i)) = \alpha \cdot (x \circ \sigma)(i) + \beta \cdot (y \circ \sigma)(i) = (\alpha \cdot (x \circ \sigma) + \beta \cdot (y \circ \sigma))(i)$  so that  $(\alpha \cdot x + \beta \cdot y) \circ \sigma = \alpha \cdot (x \circ \sigma) + \beta \cdot (y \circ \sigma)$ . So  $\sigma L(\alpha \cdot x + \beta \cdot y) = L((\alpha \cdot x + \beta \cdot y) \circ \sigma) = L(\alpha \cdot (x \circ \sigma) + \beta \cdot (y \circ \sigma)) = \alpha \cdot L(x \circ \sigma) + \beta \cdot L(y \circ \sigma) = \alpha \cdot \sigma L(x) + \beta \cdot \sigma L(y)$
2. **(bijectivity)**
  - a. **(injectivity)** Let  $x, y \in F^n$  is such that  $\sigma L(x) = \sigma L(y) \Rightarrow L(x \circ \sigma) = L(y \circ \sigma)$   $\xrightarrow[L \text{ is isomorphism}]{} x \circ \sigma = y \circ \sigma \Rightarrow (x \circ \sigma) \circ \sigma^{-1} = (y \circ \sigma) \circ \sigma^{-1} \Rightarrow x \circ (\sigma \circ \sigma^{-1}) = y \circ (\sigma \circ \sigma^{-1}) \Rightarrow x = y$
  - b. **(surjectivity)** Let  $y \in Y$  then by surjectivity of  $L$  there exists a  $x: \{1, \dots, n\} \rightarrow F$  such that  $L(x) = y$ . Taken then  $x \circ \sigma^{-1}: \{1, \dots, n\} \rightarrow F$  so that  $\sigma L(x \circ \sigma^{-1}) = L((x \circ \sigma^{-1}) \circ \sigma) = L(x \circ (\sigma^{-1} \circ \sigma)) = L(x) = y$   $\square$

**Theorem 10.225. (Universal property)** *Let  $X$  be a vector space over the field  $F$ ,  $\emptyset \neq Y \subseteq X$  a sub vector space of  $X$  and  $Z$  another vector space over the field  $F$  and  $\varphi: X \rightarrow Z$  a linear mapping such that  $Y \subseteq \ker(\varphi)$  then there exists a unique linear mapping  $h: X/Y \rightarrow Z$  such that  $\varphi = h \circ \pi_Y$*

**Proof.** Define  $h: X/Y \rightarrow Z$  by  $h(\sim_Y[x]) \rightarrow \varphi(x)$ , this is a well defined function for if  $\sim_Y[x] = \sim_Y[x']$  then  $x - x' \in Y \subseteq \ker(\varphi) \Rightarrow \varphi(x - x') = 0 \xrightarrow[\varphi \text{ is linear}]{} \varphi(x) - \varphi(x') = 0 \Rightarrow \varphi(x) = \varphi(x')$  so that  $h(\sim_Y[x]) = \varphi(x) = \varphi(x') = h(\sim_Y[x'])$ . Also  $h$  is linear as  $h(\alpha \cdot \sim_Y[x] + \beta \cdot \sim_Y[y]) = h(\sim_Y[\alpha \cdot x + \beta \cdot y]) = \varphi(\alpha \cdot x + \beta \cdot y) \xrightarrow[\varphi \text{ is linear}]{} \alpha \cdot \varphi(x) + \beta \cdot \varphi(y) = \alpha \cdot h(\sim_Y[x]) + \beta \cdot h(\sim_Y[y])$ . Next if  $x \in X$  then  $\pi_Y[x] = \sim_Y[x]$  so that  $(h \circ \pi_Y)(x) = h(\pi_Y(x)) = h(\sim_Y[x]) = \varphi(x)$  so that  $h \circ \pi_Y = \varphi$ . Finally if  $h': X/Y \rightarrow Z$  is such that  $h' \circ \pi_Y = \varphi$  then  $\forall \sim_Y[x] \in X/Y$  we have  $h(\sim_Y[x]) = \varphi(x) = h'(\pi_Y(x)) = h'(\sim_Y[x])$  so that  $h = h'$  proving uniqueness.  $\square$

## 10.7 Multilinear Mappings

**Definition 10.226.** *Let  $n \in \mathbb{N}_0$  and  $i \in \{1, \dots, n\}$  then we define  $\{1, \dots, i-1, i+1, \dots, n\}$  to be  $\{1, \dots, n\} \setminus \{i\}$ . If  $i, j \in \{1, \dots, n\}$  then we define  $\{1, \dots, i-1, i+1, \dots, j-1, j+1, \dots, n\}$  by  $\{1, \dots, n\} \setminus \{i, j\}$ .*

**Definition 10.227.** *If  $n \in \mathbb{N}_0$  and  $\{X_i\}_{i \in \{1, \dots, n\}}$  a family of sets and  $x = (x_1, \dots, x_n) \in \prod_{i \in \{1, \dots, n\}} X_i$  is such that  $x_i = a$  for  $i \in \{1, \dots, n\}$  then we note this as  $(x_1, \dots, x_{i-1}, a, x_{i+1}, \dots, x_n)$ . In the special cases that  $i = 1$  or  $i = n$  we sometimes use  $(a, x_2, \dots, x_n)$  or  $(x_1, \dots, x_{n-1}, a)$  instead of  $(x_1, \dots, x_{i-1}, a, x_{i+1}, \dots, x_n)$ .*

**Definition 10.228.** If  $n \in \mathbb{N}_0$  an  $\{X_i\}_{i \in \{1, \dots, n\}}$  a family of sets then if  $i \in \{1, \dots, n\}$ ,  $a \in X_i$  and  $x \in \prod_{i \in \{1, \dots, n\}} X_i$  then we define  $x_{i \rightarrow a}$  by  $(x_{i \rightarrow a})_j = \begin{cases} x_i & \text{if } j \in \{1, \dots, n\} \setminus \{i\} \\ a & \text{if } j = i \end{cases}$  (or using the above definition we have  $(x_1, \dots, x_n)_{i \rightarrow a} = (x_1, \dots, x_{i-1}, a, x_{i+1}, \dots, x_n)$ )

**Theorem 10.229.** If  $n \in \mathbb{N}_0$  an  $\{X_i\}_{i \in \{1, \dots, n\}}$  a family of sets then if  $i \in \{1, \dots, n\}$ ,  $a \in X_i$  and  $x \in \prod_{i \in \{1, \dots, n\}} X_i$  then if  $i, j \in \{1, \dots, n\}$  and  $a \in X_i, b \in X_j$  we have  $(x_{i \rightarrow a})_{j \rightarrow b} = (x_{j \rightarrow b})_{i \rightarrow a}$

**Proof.** If  $k \in \{1, \dots, n\}$  then we have  $[(x_{i \rightarrow a})_{j \rightarrow b}]_k = \begin{cases} (x_{i \rightarrow a})_k & \text{if } k \in \{1, \dots, n\} \setminus \{j\} \\ b & \text{if } k = j \end{cases} = \begin{cases} x_k & \text{if } k \in \{1, \dots, n\} \setminus \{j, i\} \\ a & \text{if } k = i \\ b & \text{if } k = j \end{cases} = \begin{cases} (x_{j \rightarrow b})_k & \text{if } k \in \{1, \dots, n\} \setminus \{i\} \\ a & \text{if } k = i \end{cases} = [(x_{j \rightarrow b})_{i \rightarrow a}]_k$  so we have  $(x_{i \rightarrow a})_{j \rightarrow b} = (x_{j \rightarrow b})_{i \rightarrow a}$   $\square$

**Definition 10.230.** Let  $n \in \mathbb{N}_0$  and  $\{X_i\}_{i \in \{1, \dots, n\}}$  a family of vector spaces over a field  $F$ ,  $\prod_{i \in \{1, \dots, n\}} X_i$  the product vector space over the field  $F$  and  $Y$  a vector space over the field  $F$  then a mapping  $L: \prod_{i \in \{1, \dots, n\}} X_i \rightarrow Y$  is **multilinear** if  $\forall i \in \{1, \dots, n\}$  and  $\forall x \in \prod_{i \in \{1, \dots, n\}} X_i$  with  $x_i = \alpha \cdot y + \beta \cdot z$  where  $y, z \in X_i$  and  $\alpha, \beta \in F$  we have  $L(x) = \alpha \cdot L(x_{i \rightarrow y}) + \beta \cdot L(x_{i \rightarrow z})$ . Or using the definitions above we say that  $L: \prod_{i \in \{1, \dots, n\}} X_i \rightarrow Y$  is multilinear if  $\forall i \in \{1, \dots, n\}$  and  $\forall (x_1, \dots, x_{i-1}, \alpha \cdot x + \beta \cdot y, x_{i+1}, \dots, x_n) \in \prod_{i \in \{1, \dots, n\}} X_i$  where  $x, y \in X_i$  and  $\alpha, \beta \in F$  we have  $L(x_1, \dots, x_{i-1}, \alpha \cdot x + \beta \cdot y, x_{i+1}, \dots, x_n) = \alpha \cdot L(x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_n) + \beta \cdot L(x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_n)$

**Example 10.231.** 0:  $\prod_{i \in \{1, \dots, n\}} X_i \rightarrow Y$  is multilinear

**Proof.** If  $(x_1, \dots, x_{i-1}, \alpha \cdot y + \beta \cdot z, x_{i+1}, \dots, x_n)$  then  $0(x_1, \dots, x_{i-1}, \alpha \cdot y + \beta \cdot z, x_{i+1}, \dots, x_n) = 0 = 0 + 0 = \alpha \cdot 0 + \beta \cdot 0 = \alpha \cdot 0(x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_n) + \beta \cdot L(x_1, \dots, x_{i-1}, z, x_{i+1}, \dots, x_n)$   $\square$

**Example 10.232.** Let  $n \in \mathbb{N}_0$  and  $F$  a field then  $\odot: F^n \rightarrow F$  defined by  $x = (x_1, \dots, x_n) \rightarrow \odot(x_1, \dots, x_n) = \prod_{i \in \{1, \dots, n\}} x_i$  is a multilinear mapping.

**Proof.** We prove this by induction on  $n$  so let  $S = \{n \in \mathbb{N}_0 | \text{If } (x_1, \dots, x_i, \alpha \cdot x + \beta \cdot y, x_{i+1}, \dots, x_n) \in F^n \text{ then } \prod_{j \in \{1, \dots, n\}} (x_1, \dots, x_{i-1}, \alpha \cdot x + \beta \cdot y, x_{i+1}, \dots, x_n)_j = \alpha \cdot \prod_{j \in \{1, \dots, n\}} (x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_n)_{ij} + \beta \cdot \prod_{j \in \{1, \dots, n\}} (x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_n)_j \text{ for all } i \in \{1, \dots, n\}\}$  then we have :

1. If  $n = 1$  then  $\prod_{i \in \{1, \dots, 1\}} (x_1, \dots, x_{i-1}, \alpha \cdot x + \beta \cdot y, x_{i+1}, \dots, x_n) = \alpha \cdot x + \beta \cdot y = \alpha \cdot \prod_{i \in \{1, \dots, 1\}} (x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_n)_i + \beta \cdot \prod_{i \in \{1, \dots, 1\}} (x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_n)_i$  so we have  $1 \in S$
2. Let  $n \in S$  then for  $n+1$  we have for  $i \in \{1, \dots, n+1\}$  the following cases

a.  $(i \in \{1, \dots, n\})$  So that we have  $\prod_{j \in \{1, \dots, n+1\}} (x_1, \dots, x_{i-1}, \alpha \cdot x + \beta \cdot y, x_{i+1}, \dots, x_{n+1})_j = (\prod_{j \in \{1, \dots, n\}} (x_1, \dots, x_{i-1}, \alpha \cdot x + \beta \cdot y, x_{i+1}, \dots, x_{n+1})_j) \cdot x_{n+1} = (\prod_{j \in \{1, \dots, n\}} (x_1, \dots, x_{i-1}, \alpha \cdot x + \beta \cdot y, x_{i+1}, \dots, x_n)_j) \cdot x_{n+1} = (\alpha \cdot \prod_{j \in \{1, \dots, n\}} (x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_n)_j) \cdot x_{n+1} + \beta \cdot (\prod_{j \in \{1, \dots, n\}} (x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_n)_j) \cdot x_{n+1} = \alpha \cdot (\prod_{j \in \{1, \dots, n\}} (x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_n)_j) \cdot x_{n+1} + \beta \cdot (\prod_{j \in \{1, \dots, n\}} (x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_n)_j) \cdot x_{n+1} = \alpha \cdot \prod_{j \in \{1, \dots, n+1\}} (x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_{n+1})_j + \beta \cdot \prod_{j \in \{1, \dots, n+1\}} (x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_{n+1})_j$  so we have  $n+1 \in S$

b.  $(i = n+1)$  Then we have that  $\prod_{j \in \{1, \dots, n+1\}} (x_1, \dots, x_{i-1}, \alpha \cdot x + \beta \cdot y, x_{i+1}, \dots, x_{n+1})_j = (\prod_{j \in \{1, \dots, n+1\}} (x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_{n+1})_j) \cdot (\alpha \cdot x + \beta \cdot y) = \alpha \cdot (\prod_{j \in \{1, \dots, n+1\}} (x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_{n+1})_j) \cdot x + \beta \cdot (\prod_{j \in \{1, \dots, n+1\}} (x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_{n+1})_j) \cdot y = \alpha \cdot \prod_{j \in \{1, \dots, n+1\}} (x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_{n+1})_j + \beta \cdot \prod_{j \in \{1, \dots, n+1\}} (x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_{n+1})_j$  so that  $n+1 \in S$

Using mathematical induction we have that  $S = \mathbb{N}_0$  proving our theorem.  $\square$

**Example 10.233.** Let  $n \in \mathbb{N}_0$  and  $\{X_i\}_{i \in \{1, \dots, n\}}$  a family of vector spaces over a field  $F$  and  $\{L_i\}_{i \in \{1, \dots, n\}}$  a family in  $\bigcup_{i \in \{1, \dots, n\}} \text{Hom}(X_i, F)$  such that  $\forall i \in \{1, \dots, n\}$  we have  $L_i \in \text{Hom}(X_i, F)$  then if we define  $\prod_{i \in \{1, \dots, n\}} L_i: \prod_{i \in \{1, \dots, n\}} X_i \rightarrow F$  by  $(x_1, \dots, x_n) \rightarrow (\prod_{i \in \{1, \dots, n\}} L_i)(x_1, \dots, x_n) = \prod_{i \in \{1, \dots, n\}} L_i(x_i)$  is multilinear

**Proof.** Let  $(x_1, \dots, x_{i-1}, \alpha \cdot x + \beta \cdot y, x_{i+1}, \dots, x_n) \in \prod_{i \in \{1, \dots, n\}} X_i$  then  $(\prod_{j \in \{1, \dots, n\}} L_j)(x_1, \dots, x_{i-1}, \alpha \cdot x + \beta \cdot y, x_i, \dots, x_n) = \prod_{j \in \{1, \dots, n\}} (L(x_1), \dots, L(x_{i-1}), L(\alpha \cdot x + \beta \cdot y), L(x_{i+1}), \dots, L(x_n))_j = \prod_{j \in \{1, \dots, n\}} (L(x_1), \dots, L(x_{i-1}), \alpha \cdot L(x) + \beta \cdot L(y), L(x_{i+1}), \dots, L(x_n))_j \stackrel{\text{previous theorem}}{=} \alpha \cdot \prod_{j \in \{1, \dots, n\}} (L(x_1), \dots, L(x_{i-1}), L(x), L(x_{i+1}), \dots, L(x_n))_j + \beta \cdot \prod_{j \in \{1, \dots, n\}} (L(x_1), \dots, L(x_{i-1}), L(y), L(x_{i+1}), \dots, L(x_n))_j = \alpha \cdot ((\prod_{i \in \{1, \dots, n\}} L_i)(x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_n)) + \beta \cdot ((\prod_{i \in \{1, \dots, n\}} L_i)(x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_n))$   $\square$

**Definition 10.234.** Let  $n \in \mathbb{N}_0$  and  $\{X_i\}_{i \in \{1, \dots, n\}}$  a family of vector spaces over a field  $F$ ,  $Y$  a vector space over a field  $F$  then  $\text{Hom}(X_1, \dots, X_n; Y) = \left\{ L \in Y^{\prod_{i \in \{1, \dots, n\}} X_i} \mid L \text{ is multilinear} \right\}$  /So  $\text{Hom}(X_1, \dots, X_n; Y)$  is the set of multilinear mappings  $L: \prod_{i \in \{1, \dots, n\}} X_i \rightarrow Y$ .

**Theorem 10.235.** Let  $n \in \mathbb{N}_0$  and  $\{X_i\}_{i \in \{1, \dots, n\}}$  a family of vector spaces over a field  $F$ ,  $Y$  a vector space over a field  $F$  then  $\text{Hom}(X_1, \dots, X_n; Y)$  is a sub space of  $Y^{\prod_{i \in \{1, \dots, n\}} X_i}$ . Further as  $0 \in \text{Hom}(X_1, \dots, X_n, Y)$  it follows that  $\text{Hom}(X_1, \dots, X_n; Y)$  is a vector space.

**Proof.** If  $\alpha, \beta \in F$  and  $L_1, L_2 \in \text{Hom}(X_1, \dots, X_n; Y)$  then we have to prove that  $\alpha \cdot L_1 + \beta \cdot L_2 \in \text{Hom}(X_1, \dots, X_n, Y)$ . So if  $\alpha', \beta' \in F$ ,  $i \in \{1, \dots, n\}$ ,  $y, z \in X_i$  and  $(x_1, \dots, x_{i-1}, \alpha' \cdot y + \beta' \cdot z, x_{i+1}, \dots, x_n)$  then  $(\alpha \cdot L_1 + \beta \cdot L_2)(x_1, \dots, x_{i-1}, \alpha' \cdot y + \beta' \cdot z, x_{i+1}, \dots, x_n) = \alpha \cdot L_1(x_1, \dots, x_{i-1}, \alpha' \cdot y + \beta' \cdot z, x_{i+1}, \dots, x_n) + \beta \cdot L_2(x_1, \dots, x_{i-1}, \alpha' \cdot y + \beta' \cdot z, x_{i+1}, \dots, x_n) \stackrel{L_1, L_2 \text{ is multilinear}}{=} \alpha \cdot (\alpha' \cdot L_1(x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_n) + \beta' \cdot L_1(x_1, \dots, x_{i-1}, z, x_{i+1}, \dots, x_n)) + \beta \cdot (\alpha' \cdot L_2(x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_n) + \beta' \cdot L_2(x_1, \dots, x_{i-1}, z, x_{i+1}, \dots, x_n)) = \alpha' \cdot (\alpha \cdot L_1(x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_n) + \beta \cdot L_2(x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_n)) + \beta' \cdot (\alpha \cdot L_1(x_1, \dots, x_{i-1}, z, x_{i+1}, \dots, x_n) + \beta \cdot L_2(x_1, \dots, x_{i-1}, z, x_{i+1}, \dots, x_n)) = \alpha' \cdot (\alpha \cdot L_1 + \beta \cdot L_2)(x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_n) + \beta' \cdot (\alpha \cdot L_1 + \beta \cdot L_2)(x_1, \dots, x_{i-1}, z, x_{i+1}, \dots, x_n)$  proving that  $\alpha \cdot L_1 + \beta \cdot L_2$  is multilinear and thus element of  $\text{Hom}(X_1, \dots, X_n, Y)$ .  $\square$

**Theorem 10.236.** Let  $n \in \mathbb{N}_0$  and  $\{X_i\}_{i \in \{1, \dots, n\}}$  a family of vector spaces over a field  $F$ ,  $Y, Z$  vector spaces over a field  $F$  then if  $K \in \text{Hom}(X_1, \dots, X_n; Y)$  is a multilinear mapping and  $L \in \text{Hom}(Y, Z)$  a linear mapping then  $L \circ K$  is a multilinear mapping (so  $L \circ K \in \text{Hom}(X_1, \dots, X_n; Z)$ )

**Proof.** Let  $\alpha, \beta \in F$ ,  $i \in \{1, \dots, n\}$   $(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n) \in \prod_{j \in \{1, \dots, i-1, i+1, \dots, n\}} X_j$ ,  $x, y \in X_i$  then we have  $(L \circ K)(x_1, \dots, x_{i-1}, \alpha \cdot x + \beta \cdot y, x_{i+1}, \dots, x_n) = L(K(x_1, \dots, x_{i-1}, \alpha \cdot x + \beta \cdot y, x_{i+1}, \dots, x_n)) \stackrel{K \text{ is multilinear}}{=} L(\alpha \cdot K(x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_n) + \beta \cdot K(x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_n)) \stackrel{L \text{ is linear}}{=} \alpha \cdot L(K(x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_n)) + \beta \cdot L(K(x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_n)) = \alpha \cdot (L \circ K)(x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_n) + \beta \cdot (L \circ K)(x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_n)$ .  $\square$

**Definition 10.237.** Let  $n \in \mathbb{N}_0$  and  $\{X_i\}_{i \in \{1, \dots, n+1\}}$  a family of vector spaces over a field  $F$ , then if  $x \in \prod_{i \in \{1, \dots, n\}} X_i$  and  $a \in X_{n+1}$  we define  $x \dots a \in \prod_{i \in \{1, \dots, n+1\}} X_i$  by  $(x \dots a)_k = \begin{cases} x_k & \text{if } k \in \{1, \dots, n\} \\ a & \text{if } k = n+1 \end{cases}$  /in other words we have that  $x \dots a = (x_1, \dots, x_n, a)$

**Theorem 10.238.** Let  $n \in \mathbb{N}_0$  and  $\{X_i\}_{i \in \{1, \dots, n+1\}}$  a family of vector spaces over a field  $F$ , then if  $x \in \prod_{i \in \{1, \dots, n\}} X_i$  and  $a \in X_{n+1}$  then if  $i \in \{1, \dots, n\}$  and  $b \in X_i$  then we have that  $(x \dots a)_{j \rightarrow b} = (x_{j \rightarrow b}) \dots a$

**Proof.** Let  $k \in \{1, \dots, n+1\}$  then  $[(x \dots a)_{j \rightarrow b}]_k = \begin{cases} (x \dots a)_k & \text{if } k \in \{1, \dots, n+1\} \setminus \{j\} \\ b & \text{if } k = j \end{cases} = \begin{cases} x_k & \text{if } k \in \{1, \dots, n\} \setminus \{j\} \\ a & \text{if } k = n+1 \\ b & \text{if } k = j \end{cases} = \begin{cases} (x_{j \rightarrow b})_k & \text{if } k \in \{1, \dots, n\} \\ a & \text{if } k = n+1 \end{cases} = [(x_{j \rightarrow b}) \dots a]_k$  so that  $(x \dots a)_{j \rightarrow b} = (x_{j \rightarrow b}) \dots a$ .  $\square$

**Definition 10.239.** Let  $n \in \mathbb{N}_0$  and  $\{X_i\}_{i \in \{1, \dots, n+1\}}$  a family of vector spaces over a field  $F$ ,  $Y$  a vector space over  $F$  then if  $L: \prod_{i \in \{1, \dots, n+1\}} X_i \rightarrow Y$  is a mapping and  $a \in X_{n+1}$  then we define  $L_{(\dots a)}: \prod_{i \in \{1, \dots, n\}} X_i \rightarrow Y$  by  $x \rightarrow L_{(\dots a)}(x \dots a)$  /in other words  $L_{(\dots a)}(x_1, \dots, x_n) = L(x_1, \dots, x_n, a)$

**Theorem 10.240.** Let  $n \in \mathbb{N}_0$  and  $\{X_i\}_{i \in \{1, \dots, n+1\}}$  a family of vector spaces over a field  $F$ ,  $Y$  a vector space over  $F$  then if  $L: \prod_{i \in \{1, \dots, n+1\}} X_i \rightarrow Y$  is a multilinear mapping and  $a \in X_{n+1}$  then  $L_{(\dots a)}: \prod_{i \in \{1, \dots, n\}} X_i \rightarrow Y$  is multilinear.

**Proof.** Take  $i \in \{1, \dots, n\}$ ,  $\alpha, \beta \in F$ ,  $y, z \in X_i$  and  $x \in \prod_{i \in \{1, \dots, n\}} X_i$  such that  $x_i = \alpha \cdot y + \beta \cdot z$  then for  $x_{\dots a} \in \prod_{i \in \{1, \dots, n+1\}}$  we have that  $(x_{\dots a})_{i \neq n+1} = x_i = \alpha \cdot y + \beta \cdot z$  so that we have that  $L_{(\dots a)}(x) = L(x_{\dots a})$   $\stackrel{L \text{ is multilinear}}{=} \alpha \cdot L((x_{\dots a})_{i \rightarrow y}) + \beta \cdot L((x_{\dots a})_{i \rightarrow z})$   $\stackrel{10.238}{=} \alpha \cdot L((x_{i \rightarrow y})_{\dots a}) + \beta \cdot L((x_{i \rightarrow z})_{\dots a}) = \alpha \cdot L_{(\dots a)}(x_{i \rightarrow y}) + \beta \cdot L_{(\dots a)}(x_{i \rightarrow z})$   $\square$

If we use the notation that is introduced in 10.227 then the proof becomes almost trivial as seen in the following version of the above proof. From now on we try to use the easy notation on our proofs.

**Proof.** Take  $i \in \{1, \dots, n\}$ ,  $\alpha, \beta \in F$ ,  $y, z \in X_i$  and  $x = (x_1, \dots, x_{i-1}, \alpha \cdot y + \beta \cdot z, x_{i+1}, \dots, x_n) \in \prod_{i \in \{1, \dots, n\}} X_i$  then  $L_{(\dots a)}(x_1, \dots, x_{i-1}, \alpha \cdot y + \beta \cdot z, x_{i+1}, \dots, x_n, a) = L(x_1, \dots, x_{i-1}, \alpha \cdot y + \beta \cdot z, x_{i+1}, \dots, x_n, a)$   $\stackrel{L \text{ is multilinear}}{=} \alpha \cdot L(x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_n, a) + \beta \cdot L(x_1, \dots, x_{i-1}, z, x_{i+1}, \dots, x_n, a) = \alpha \cdot L_{(\dots a)}(x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_n) + \beta \cdot L(x_1, \dots, x_{i-1}, z, x_{i+1}, \dots, x_n)$   $\square$

The above theorem can be used in the induction step of proofs like the following theorem will show.

**Theorem 10.241.** Let  $n \in \mathbb{N}_0$ ,  $\{X_i\}_{i \in \{1, \dots, n\}}$  a family of vector spaces over a field  $F$ ,  $Y$  a vector space over  $F$  and  $\alpha = (\alpha_1, \dots, \alpha_n) \in F^n$ ,  $x = (x_1, \dots, x_n) \in \prod_{i \in \{1, \dots, n\}} X_i$ ,  $L: \prod_{i \in \{1, \dots, n\}} X_i \rightarrow Y$  a multilinear function then if we define  $\alpha \cdot x \in \prod_{i \in \{1, \dots, n\}} X_i$  by  $\forall i \in \{1, \dots, n\}$   $(\alpha \cdot x)_i = \alpha_i \cdot x_i$  or  $\alpha \cdot z = (\alpha_1 \cdot x_1, \dots, \alpha_n \cdot x_n)$  we have that  $L(\alpha \cdot x) = (\prod_{i \in \{1, \dots, n\}} \alpha_i) \cdot L(x)$

**Proof.** We prove this by induction on  $n$  so let  $S = \{n \in \{1, \dots, n\} \mid \text{If } L: \prod_{i \in \{1, \dots, n\}} X_i \rightarrow Y \text{ is multilinear and } \alpha \in F^n, x \in \prod_{i \in \{1, \dots, n\}} X_i \text{ then } L(\alpha \cdot x) = (\prod_{i \in \{1, \dots, n\}} \alpha_i) \cdot L(x)\}$  then we have:

1. If  $n = 1$  then  $\alpha = (\alpha_1)$ ,  $x = (x_1)$  and  $L(\alpha \cdot x) = L(\alpha_1 \cdot x_1) = \alpha_1 \cdot L(x_1) = (\prod_{i \in \{1\}} \alpha_i) \cdot L(x_1) = (\prod_{i \in \{1\}} \alpha_i) \cdot L(x)$  so  $1 \in S$
2. If  $n \in S$  take then  $\alpha = (\alpha_1, \dots, \alpha_{n+1}) \in F^{n+1}$ ,  $x \in \prod_{i \in \{1, \dots, n+1\}} X_i$ ,  $L: \prod_{i \in \{1, \dots, n+1\}} X_i \rightarrow Y$  a multilinear mapping then as trivially  $\alpha \cdot x = (\alpha \cdot x|_{\{1, \dots, n\}})_{(\dots \alpha_{n+1} \cdot x_{n+1})}$  we have  $L(\alpha \cdot x) = L((\alpha \cdot x|_{\{1, \dots, n\}})_{(\dots \alpha_{n+1} \cdot x_{n+1})}) = L_{(\dots \alpha_{n+1} \cdot x_{n+1})}(\alpha \cdot x|_{\{1, \dots, n\}})$   $\stackrel{n \in S \text{ and 10.240}}{=} (\prod_{i \in \{1, \dots, n\}} \alpha_i) \cdot L_{(\dots \alpha_{n+1} \cdot x_{n+1})}(x|_{n+1}) = (\prod_{i \in \{1, \dots, n\}} \alpha_i) \cdot L(x_1, \dots, x_n, \alpha_{n+1} \cdot x_{n+1})$   $\stackrel{L \text{ is multilinear}}{=} (\prod_{i \in \{1, \dots, n\}} \alpha_i) \cdot (\alpha_{n+1} \cdot L(x_1, \dots, x_{n+1})) = ((\prod_{i \in \{1, \dots, n\}} \alpha_i) \cdot \alpha_{n+1}) \cdot L(x) = (\prod_{i \in \{1, \dots, n+1\}} \alpha_i) \cdot L(x)$  proving that  $n+1 \in S$

Using mathematical induction we have then that  $S = \{1, \dots\} = \mathbb{N}_0$  proving the theorem.  $\square$

**Theorem 10.242.** Let  $n \in \mathbb{N}_0$ ,  $\{X_i\}_{i \in \{1, \dots, n\}}$  a family of vector spaces over a field  $F$ ,  $Y$  a vector-space over  $F$  and  $L: \prod_{i \in \{1, \dots, n\}} X_i \rightarrow Y$  is a multi-linear function then if  $x \in \prod_{i \in \{1, \dots, n\}} X_i$  is such that  $\exists i \in \{1, \dots, n\}$  such that  $x_i = 0$  then  $L(x) = 0$

**Proof.**  $L(x_1, \dots, x_{i-1}, 0, x_i, \dots, x_n) = L(x_1, \dots, x_{i-1}, 0 \cdot 0, x_{i+1}, \dots, x_n) = 0 \cdot L(x_1, \dots, x_{i-1}, 0, x_{i+1}, \dots, x_n) = 0$   $\square$

**Theorem 10.243.** Let  $n \in \mathbb{N}_0$ ,  $\{X_i\}_{i \in \{1, \dots, n\}}$  a family of vector spaces over a field  $F$ ,  $i \in \{1, \dots, n\}$ ,  $a \in X_i$  and  $\sigma \in P_n$  then  $(x_1, \dots, x_{i-1}, a, x_{i+1}, \dots, x_n) \circ \sigma = ((x \circ \sigma)_1, \dots, (x \circ \sigma)_{\sigma^{-1}(i)-1}, a, (x \circ \sigma)_{\sigma^{-1}(i)+1}, \dots, (x \circ \sigma)_{n+1})$  (in other words if  $x_i = a$  then  $(x \circ \sigma)_{\sigma^{-1}(i)} = a$ )

**Proof.** If  $x = (x_1, \dots, x_{i-1}, a, x_{i+1}, \dots, x_n)$  then this is by definition equivalent to  $x_i = a$ , then we have  $(x \circ \sigma)_{\sigma^{-1}(i)} = x(\sigma(\sigma^{-1}(i))) = x(i) = x_i = a$ .  $\square$

**Theorem 10.244.** Let  $Y$  be a vector space over a field  $F$ ,  $n \in \mathbb{N}_0$ ,  $\{X_i\}_{i \in \{1, \dots, n\}}$  a family of vector spaces over  $F$  and then

1. If  $m \in \mathbb{N}_0$  and  $\{L_i\}_{i \in \{0, \dots, m\}}$  is a family in  $\text{Hom}(X_1, \dots, X_n; Y)$  then if we define  $\sum_{i=0}^m L_i: \prod_{i \in \{1, \dots, n\}} X_i \rightarrow Y$  by  $x \rightarrow \sum_{i=0}^m L_i(x)$  we have that  $\sum_{i=0}^m L_i$  is a multilinear function.
2. If  $\{L_i\}_{i \in I}$  is a finite family (meaning  $I$  is finite) in  $\text{Hom}(X_1, \dots, X_n; Y)$  then if we define  $\sum_{i \in I} L_i: \prod_{i \in \{1, \dots, n\}} X_i \rightarrow Y$  by  $x \rightarrow \sum_{i \in I} L_i(x)$  (which is defined as  $I$  is finite) we have that  $\sum_{i \in I} L_i$  is multilinear.

**Proof.**

1. Let  $\alpha, \beta \in F$ ,  $i \in \{1, \dots, n\}$ ,  $y, z \in X_i$  and  $x = (x_1, \dots, x_{i-1}, \alpha \cdot y + \beta \cdot z, x_{i+1}, \dots, x_n) \in \prod_{j \in \{1, \dots, n\}} X_i$  then  $(\sum_{j=0}^m L_j)(x) = \sum_{j=0}^m L_j(x) \underset{L_i \text{ are multilinear}}{=} \sum_{j=0}^m (\alpha \cdot L_j(x_{i \rightarrow y}) + \beta \cdot L_j(x_{i \rightarrow z})) = \sum_{j=0}^m \alpha \cdot L_j(x_{i \rightarrow y}) + \sum_{j=0}^m \beta \cdot L_j(x_{i \rightarrow z}) = \alpha \cdot \sum_{j=0}^m L_j(x_{i \rightarrow y}) + \beta \cdot \sum_{j=0}^m L_j(x_{i \rightarrow z}) = \alpha \cdot (\sum_{j=0}^m L_j)(x_{i \rightarrow y}) + \beta \cdot (\sum_{j=0}^m L_j)(x_{i \rightarrow z})$
2. Let  $\alpha, \beta \in F$ ,  $i \in \{1, \dots, n\}$ ,  $y, z \in X_i$  and  $x = (x_1, \dots, x_{i-1}, \alpha \cdot y + \beta \cdot z, x_{i+1}, \dots, x_n) \in \prod_{j \in \{1, \dots, n\}} X_i$  then  $(\sum_{j \in I} L_j)(x) = \sum_{j \in I} L_j(x) \underset{L_i \text{ are multilinear}}{=} \sum_{j \in I} (\alpha \cdot L_j(x_{i \rightarrow y}) + \beta \cdot L_j(x_{i \rightarrow z})) = \sum_{j \in I} \alpha \cdot L_j(x_{i \rightarrow y}) + \sum_{j \in I} \beta \cdot L_j(x_{i \rightarrow z}) = \alpha \cdot \sum_{j \in I} L_j(x_{i \rightarrow y}) + \beta \cdot \sum_{j \in I} L_j(x_{i \rightarrow z}) = \alpha \cdot (\sum_{j \in I} L_j)(x_{i \rightarrow y}) + \beta \cdot (\sum_{j \in I} L_j)(x_{i \rightarrow z}) \quad \square$

**Theorem 10.245.** Let  $n \in \mathbb{N}_0$ ,  $\{X_i\}_{i \in \{1, \dots, n\}}$  a family of vector spaces over a field  $F$ ,  $Y$  a vector space over a field  $F$  then if  $L \in \text{Hom}(X_1, \dots, X_n; Y)$  and  $i \in \{1, \dots, n\}$ :

1. If  $m \in \mathbb{N}_0$  and  $\{\alpha_j\}_{j \in \{1, \dots, m\}}$  is a family in  $F$  and  $\{y_j\}_{j \in \{1, \dots, m\}}$  a family in  $X_i$  then if  $x = (x_1, \dots, x_{i-1}, \sum_{j=1}^m \alpha_j \cdot x_j, x_{i+1}, \dots, x_n) \in \prod_{i \in \{1, \dots, n\}} X_i$  we have  $L(x_1, \dots, x_{i-1}, \sum_{j=1}^m \alpha_j \cdot x_j, x_{i+1}, \dots, x_n) = \sum_{j=1}^m \alpha_j L(x_1, \dots, x_{i-1}, y_j, x_{i+1}, \dots, x_n)$
2. If  $\{\alpha_j\}_{j \in I}$  is a family in  $F$  with finite support and  $\{y_j\}_{j \in I}$  a family in  $X_i$  then if  $x = (x_1, \dots, x_{i-1}, \sum_{j \in I} \alpha_j \cdot x_j, x_{i+1}, \dots, x_n) \in \prod_{i \in \{1, \dots, n\}} X_i$  we have  $L(x_1, \dots, x_{i-1}, \sum_{j \in I} \alpha_j \cdot x_j, x_{i+1}, \dots, x_n) = \sum_{j \in I} \alpha_j L(x_1, \dots, x_{i-1}, y_j, x_{i+1}, \dots, x_n)$

**Proof.**

1. We prove this by induction on  $m$  so let  $S_i = \{m \in \{1, \dots\} \mid \text{if } \{\alpha_j\}_{j \in \{1, \dots, m\}}, \{y_j\}_{j \in \{1, \dots, m\}} \text{ are families in } F, X_i \text{ then if } (x_1, \dots, x_{i-1}, \sum_{j=1}^m \alpha_j \cdot y_j, x_{i+1}, \dots, x_n) \text{ we have } L(x_1, \dots, x_{i-1}, \sum_{j=1}^m \alpha_j \cdot y_j, x_{i+1}, \dots, x_n) = \sum_{j=1}^m \alpha_j \cdot L(x_1, \dots, x_{i-1}, y_j, x_{i+1}, \dots, x_n)\}$  then we have:
  - a. If  $m=1$  then if  $\{\alpha_j\}_{j \in \{1\}}, \{y_j\}_{j \in \{1\}}$  are families in  $F, X_i$  then  $L(x_1, \dots, x_{i-1}, \sum_{j=1}^1 \alpha_j \cdot x_j, x_{i+1}, \dots, x_n) = L(x_1, \dots, x_{i-1}, \alpha_1 \cdot x_1, x_{i+1}, \dots, x_n) = \alpha_1 \cdot L(x_1, \dots, x_{i-1}, y_1, x_{i+1}, \dots, x_n) = \sum_{j=1}^1 \alpha_j \cdot L(x_1, \dots, x_{i-1}, y_j, x_{i+1}, \dots, x_n)$  proving that  $1 \in S_i$ .
  - b. If  $m \in S_i$  then if  $\{\alpha_j\}_{j \in \{1, \dots, m+1\}}, \{y_j\}_{j \in \{1, \dots, m+1\}}$  are families in  $F, X_i$  then  $L(x_1, \dots, x_{i-1}, \sum_{j=1}^{m+1} \alpha_j \cdot y_j, x_{i+1}, \dots, x_n) = L(x_1, \dots, x_{i-1}, \sum_{j=1}^m \alpha_j \cdot y_j + \alpha_{m+1} \cdot y_{m+1}, x_{i+1}, \dots, x_n) = L(x_1, \dots, x_{i-1}, \sum_{j=1}^m \alpha_j \cdot y_j, x_{i+1}, \dots, x_n) + \alpha_{m+1} \cdot L(x_1, \dots, x_{i-1}, y_{m+1}, x_{i+1}, \dots, x_n) \underset{m \in S_i}{=} \sum_{j=1}^m \alpha_j \cdot L(x_1, \dots, x_{i-1}, y_j, x_{i+1}, \dots, x_n) + \alpha_{m+1} \cdot L(x_1, \dots, x_{i-1}, y_{m+1}, x_{i+1}, \dots, x_n) = \sum_{j=1}^{m+1} \alpha_j \cdot L(x_1, \dots, x_{i-1}, y_j, x_{i+1}, \dots, x_n)$  proving that  $m+1 \in S_i$ .

Using mathematical induction we have that  $S_i = \{1, \dots\} = \mathbb{N}_0$

2. Let  $\{\alpha_j\}_{j \in I}$  be a family in  $F$  with finite support,  $\{y_j\}_{j \in I}$  a family in  $X_i$ . We have  $\text{support}(\{\alpha_j \cdot x_j\}_{j \in I}) \subseteq \text{support}(\{\alpha_j\}_{j \in I})$  (see 10.128). Now if  $k \in \text{support}(\{\alpha_j \cdot L(x_1, \dots, x_{i-1}, y_j, x_{i+1}, \dots, x_n)\}_{j \in I}) \Rightarrow \alpha_k \cdot L(x_1, \dots, x_{i-1}, y_k, x_{i+1}, \dots, x_n) \neq 0 \Rightarrow \alpha_k \neq 0 \Rightarrow k \in \text{support}(\{\alpha_j\}_{j \in I})$ , so  $\text{support}(\{\alpha_j \cdot x_j\}_{j \in I}), \text{support}(\{\alpha_j \cdot L(x_1, \dots, x_{i-1}, y_j, x_{i+1}, \dots, x_n)\}_{j \in I}) \subseteq \text{support}(\{\alpha_j\}_{j \in I})$  proving that  $\text{support}(\{\alpha_j \cdot x_j\}_{j \in I}), \text{support}(\{\alpha_j \cdot L(x_1, \dots, x_{i-1}, y_j, x_{i+1}, \dots, x_n)\}_{j \in I})$  is finite and thus that both sums in (2) are well defined. If now  $b: \{0, \dots, m\} \rightarrow \text{support}(\{\alpha_i\}_{i \in I})$  then we have  $L(x_1, \dots, x_{i-1}, \sum_{j \in I} \alpha_j \cdot y_j, x_{i+1}, \dots, x_n) = L(x_1, \dots, x_{i-1}, \sum_{j=0}^m \alpha_{b_j} \cdot y_{b_j}, x_{i+1}, \dots, x_n) = L(x_1, \dots, x_{i-1}, \sum_{j=1}^{m+1} \alpha_{b_{j-1}} \cdot y_{b_{j-1}}, x_{i+1}, \dots, x_n) \underset{(1)}{=} \sum_{j=1}^{m+1} \alpha_{b_{j-1}} \cdot L(x_1, \dots, x_{i-1}, y_{b_{j-1}}, x_{i+1}, \dots, x_n) = \sum_{j=0}^m \alpha_{b_j} \cdot L(x_1, \dots, x_{i-1}, y_{b_j}, x_{i+1}, \dots, x_n) = \sum_{j \in I} \alpha_j \cdot L(x_1, \dots, x_{i-1}, y_j, x_{i+1}, \dots, x_n) \quad \square$

The following theorem proves that given a family of sub spaces of a family of vectorspaces we can always extend multilinear mappings on the sub spaces to multilinear mappings on the vector space.

**Theorem 10.246.** *Let  $n \in \mathbb{N}_0$  and  $\{X_i\}_{i \in \{1, \dots, n\}}$  a family of vector spaces over a field  $F$ ,  $Y$  a vector space over  $F$  and let  $\{Y_i\}_{i \in \{1, \dots, n\}}$  a family of non empty sub spaces [i.e.  $\forall i \in \{1, \dots, n\}$  we have that  $Y_i$  is a non empty subspace of  $X_i$ ]. If  $L \in \text{Hom}(Y_1, \dots, Y_n; Y)$  then we can find a  $K \in \text{Hom}(X_1, \dots, X_n; Y)$  such that  $K|_{\prod_{i \in \{1, \dots, n\}} Y_i} = L$*

**Proof.** Using 10.196 we have  $\forall i \in \{1, \dots, n\}$  that there exists a subspace  $Z_i$  such that  $X_i = Y_i \oplus Z_i$  and thus  $\forall x \in X_i \exists! y \in Y_i \wedge \exists! z \in Z_i$  such that  $x = y + z$ , this defines two functions  $\tau_i: X_i \rightarrow Y_i$  and  $\rho_i: X_i \rightarrow Z_i$  such that  $x = \tau_i(x) + \rho_i(x)$ . Note that if  $\alpha, \beta \in F$  and  $x, y \in X_i$  then we have  $x = \tau_i(x) + \rho_i(x), y = \tau_i(y) + \rho_i(y)$  then  $\alpha \cdot x + \beta \cdot y = \alpha \cdot (\tau_i(x) + \rho_i(x)) + \beta \cdot (\tau_i(y) + \rho_i(y)) = (\alpha \cdot \tau_i(x) + \beta \cdot \tau_i(y)) + (\alpha \cdot \rho_i(x) + \beta \cdot \rho_i(y)) \xrightarrow{\text{extension is unique in direct sum}} \tau_i(\alpha \cdot x + \beta \cdot y) = \alpha \cdot \tau_i(x) + \beta \cdot \tau_i(y) \wedge \rho_i(\alpha \cdot x + \beta \cdot y) = \alpha \cdot \rho_i(x) + \beta \cdot \rho_i(y)$  proving that  $\tau_i, \rho_i$  are linear. Given now  $L$  define  $K: \prod_{i \in \{1, \dots, n\}} X_i \rightarrow Y$  by  $(x_1, \dots, x_n) \rightarrow K(x_1, \dots, x_n) = L(\tau_1(x_1), \dots, \tau_n(x_n))$ . We have now  $K(x_1, \dots, x_{i-1}, \alpha \cdot x + \beta \cdot y, x_{i+1}, \dots, x_n) = L(\tau_1(x_1), \dots, \tau_{i-1}(x_{i-1}), \tau_i(\alpha \cdot x + \beta \cdot y), \tau_{i+1}(x_{i+1}), \dots, \tau_n(x_n)) \xrightarrow{\tau_i \text{ is linear}} L(\tau_1(x_1), \dots, \tau_{i-1}(x_{i-1}), \alpha \cdot \tau_i(x) + \beta \cdot \tau_i(y), \tau_{i+1}(x_{i+1}), \dots, \tau_n(x_n)) \xrightarrow{L \text{ is multilinear}} \alpha \cdot L(\tau_1(x_1), \dots, \tau_{i-1}(x_{i-1}), \tau_i(x), \tau_{i+1}(x_{i+1}), \dots, \tau_n(x_n)) + \beta \cdot L(\tau_1(x_1), \dots, \tau_{i-1}(x_{i-1}), \tau_i(y), \tau_{i+1}(x_{i+1}), \dots, \tau_n(x_n)) = \alpha \cdot K(x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_n) + \beta \cdot K(x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_n)$  proving that  $K$  is multilinear or  $K \in \text{Hom}(X_1, \dots, X_n; Y)$ .

If now  $(x_1, \dots, x_n) \in \prod_{i \in \{1, \dots, n\}} Y_i \subseteq \prod_{i \in \{1, \dots, n\}} X_i$  then  $\forall i \in \{1, \dots, n\}$  we have  $x_i = x_i + 0 \xrightarrow{x_i \in Y_i, 0 \in Z_i} \tau_i(x_i) = x_i$  and thus  $K(x_1, \dots, x_n) = L(\tau_1(x_1), \dots, \tau_n(x_n)) = L(x_1, \dots, x_n)$  proving that  $K|_{\prod_{i \in \{1, \dots, n\}} Y_i} = L$   $\square$

## 10.8 Determinant Functions

First we have to introduce the concept of  $-1$  which we need in addition to  $1$  for the definition of skew-symmetric functions.

**Definition 10.247.** *If  $F$  is a field and  $1$  is the multiplicative unit then  $-1$  is the (sum) inverse of  $1$  so  $-1$  is such that  $(-1) + 1 = 0 = 1 + (-1)$*

We have then the following theorem for a vector space

**Theorem 10.248.** *If  $X$  is a vector field over  $F$  then we have that  $\forall x \in X$  that  $-x = (-1) \cdot x$*

**Proof.** We have  $(-1) \cdot x + x = (-1) \cdot x + 1 \cdot x = ((-1) + 1) \cdot x = 0 \cdot x = 0 = 0 \cdot x = (1 + (-1)) \cdot x = 1 \cdot x + (-1) \cdot x = x + (-1) \cdot x$   $\square$

Next we have a generalization of power's in  $\mathbb{Z}$

**Definition 10.249.** *If  $F$  is a field then as  $\langle F, \cdot \rangle$  is a abelian semi-group we have by 4.22 that given  $a \in F$  and  $n \in \mathbb{N}$  that there exists a  $a^n$  such that*

$$\begin{aligned} a^0 &= 1 \\ a^{n+1} &= a^n \cdot a \xrightarrow{\text{abelian}} a \cdot a^n \end{aligned}$$

**Theorem 10.250.** *If  $n, n' \in \mathbb{N}$  and  $a \in F$  then  $a^{n'+n} = a^{n'} \cdot a^n$*

**Proof.** We prove this by induction on  $n$ . So let  $X = \{n \in \mathbb{N} | a^{n'+n} = a^{n'} \cdot a^n\}$  then we have

1. If  $n = 0$  then  $a^{n'+n} = a^{n'+0} = a^{n'} = a^{n'} \cdot 1 = a^{n'} \cdot a^0 \Rightarrow 0 \in X$
2. If  $n \in X$  then  $a^{n'+(n+1)} = a^{(n'+n)+1} = a^{(n'+n)} \cdot a \xrightarrow{n \in X} (a^{n'} \cdot a^n) \cdot a = a^{n'} \cdot (a^n \cdot a) = a^{n'} \cdot a^{n+1}$  and thus  $n+1 \in X$

Using mathematical induction (see 4.10) we have  $X = \mathbb{N}$  proving the theorem  $\square$

**Theorem 10.251.** *In  $F$  we have*

$$\begin{aligned} 0^n &= 0 \text{ (if } n \neq 0) \\ 1^n &= 1 \\ (-1)^n &= -1 \text{ or } 1 \\ (-1)^{2 \cdot n} &= 1 \\ (-1)^{2 \cdot n+1} &= -1 \end{aligned}$$

**Proof.**

1. If  $n \neq 0 \Rightarrow \exists m \in \mathbb{N} \vdash n = m+1$  then  $0^n = 0^{(m+1)} = 0^m \cdot 0 = 0$
2.  $1^n = 1$  is proved by induction on  $n$ , let  $X = \{n \in \mathbb{N} \mid 1^n = 1\}$  then
  - a.  $1^0 = 1 \Rightarrow 0 \in X$
  - b. If  $n \in X \Rightarrow 1^{n+1} = 1^n \cdot 1 \underset{n \in X}{=} 1 \cdot 1 = 1 \Rightarrow n+1 \in X$   
so  $X = \mathbb{N}$
3.  $(-1)^n = \pm 1$  is proved by induction on  $n$ , let  $X = \{n \in \mathbb{N} \mid (-1)^n = -1 \text{ or } 1\}$  then
  - a.  $(-1)^0 = 1 \Rightarrow 0 \in X$
  - b. If  $n \in X$  then  $(-1)^{n+1} = (-1)^n \cdot (-1) \underset{n \in X}{=} (-1) \cdot (-1) \vee 1 \cdot (-1) = 1 \vee -1 \Rightarrow n+1 \in X$   
so  $X = \mathbb{N}$
4.  $(-1)^{2 \cdot n} = (-1)^{(1+1) \cdot n} = (-1)^{n+n} = (-1)^n \cdot (-1)^n \underset{(3)}{=} (-1) \cdot (-1) \text{ or } 1 \cdot 1 = 1$
5.  $(-1)^{2 \cdot n+1} = (-1)^{2 \cdot n} \cdot (-1) \underset{(4)}{=} 1 \cdot (-1) = -1$   $\square$

We can now make identification between  $(1, -1) \in \mathbb{Z}$  and  $(1, -1) \in F$  such that  $(-1)^n$  is the same in  $\mathbb{Z}$  and  $F$  as is proved in the next theorem.

**Theorem 10.252.** *If  $(1, -1) \in \mathbb{Z}$  and  $(1', -1') \in F$  [we use ' to show that we are talking about  $(1, -1) \in F$ ] then  $\forall n \in \mathbb{N}$  we have that  $(-1)^n = (-1')^n$  if we equate 1 with  $1'$  and  $-1$  with  $-1'$*

**Proof.** This is easily proved by induction on  $n \in \mathbb{N}$  so let  $S = \{n \in \mathbb{N} \mid (-1)^n = (-1')^n\}$  then we have:

1. If  $n = 0$  then  $(-1)^0 = 1 = 1' = (-1')^0$  and thus  $n \in S$
2. If  $n \in S$  then  $(-1)^n = (-1')^n$  and we have to consider the following cases
  - a.  $((-1)^n = 1 = 1' = (-1')^n)$  so  $(-1)^{n+1} = (-1) \cdot (-1)^n = (-1) \cdot 1 = -1 = -1' = (-1')$ .  
 $1' = (-1') \cdot (-1)^n = (-1')^n$  and thus  $n+1 \in S$
  - b.  $((-1)^n = -1 = -1' = (-1')^n)$   $(-1)^{n+1} = (-1) \cdot (-1)^n = (-1) \cdot (-1) = 1 = 1' = (-1') \cdot (-1') = (-1') \cdot (-1)^n = (-1')^n$  so  $n+1 \in S$

So in all cases we have  $n+1 \in S$

By mathematical induction we have then  $S = \mathbb{N}$  proving our theorem.  $\square$

For the following definition remember by 10.100 that if  $n \in \mathbb{N}_0$ ,  $X$  is a set and  $\{X_i\}_{i \in \{1, \dots, n\}}$  is a family of sets such that  $X_i = X$  then  $X^n = \prod_{i \in \{1, \dots, n\}} X_i$  (which is by 10.101 equal to  $X^{\{1, \dots, n\}}$ ).

**Definition 10.253.** *Let  $n \in \mathbb{N}_0$  and  $X, Y$  vector spaces over a field  $F$  then we say that  $L: X^n \rightarrow Y$  is a  $n$ -linear mapping if  $L$  is a multilinear mapping. We note  $\text{Hom}(X^n; Y) = \{L \in Y^{X^n} \mid L \text{ is multilinear}\}$  for the set of  $n$ -linear mappings.*

**Note 10.254.** Using 10.235 we have that  $\text{Hom}(X^n, Y)$  is a vector space and a subspace of  $Y^{X^n}$

**Definition 10.255.** Let  $X, Y$  be vector spaces over the field  $F$ ,  $n \in \mathbb{N}_0$  then a  $n$ -linear mapping  $L \in \text{Hom}(X^n; Y)$  is **skew symmetric** if  $\forall \sigma \in P_n$  we have  $\sigma L = \text{sign}(\sigma) \cdot L$

**Example 10.256.** If  $X, Y$  are vector spaces over the field  $F$  and  $n = 1$  we have that if  $L \in \text{Hom}(X^1; Y)$  then  $L$  is skew symmetric

**Proof.** If  $\sigma \in P_1$  then  $\sigma = 1_{\{1\}}$  with  $\text{sign}(\sigma) = 1$  and  $\forall x \in X^1$  we have  $\sigma L(x) = L(x \circ \sigma) = L(x) = \text{sign}(\sigma) \cdot L(x)$  proving that  $\sigma L = \text{sign}(\sigma) \cdot L$   $\square$

**Lemma 10.257.** Let  $X, Y$  be vector spaces over a field  $F$ ,  $n \in \mathbb{N}_0$  and  $L \in \text{Hom}(X^n; Y)$  a  $n$ -linear mapping then  $\sigma L$  is a multilinear mapping.

**Proof.** Let  $i \in \{1, \dots, n\}$ ,  $\alpha, \beta \in F$   $y, z \in X_i$  and  $x = (x_1, \dots, x_{i-1}, \alpha \cdot y + \beta \cdot z, x_{i+1}, \dots, x_n) \in \prod_{i \in \{1, \dots, n\}} X_i$  then  $\sigma L(x) = L((x_1, \dots, x_{i-1}, \alpha \cdot y + \beta \cdot z, x_{i+1}, \dots, x_n) \circ \sigma) \stackrel{10.243}{=} L(x_{\sigma(1)}, \dots, x_{\sigma(\sigma^{-1}(i)-1)}, \alpha \cdot y + \beta \cdot z, x_{\sigma(\sigma^{-1}(i)+1)}, \dots, x_{\sigma(n)}) = \alpha \cdot L(x_{\sigma(1)}, \dots, x_{\sigma(\sigma^{-1}(i)-1)}, y, x_{\sigma(\sigma^{-1}(i)+1)}, \dots, x_{\sigma(n)}) + \beta \cdot L(x_{\sigma(1)}, \dots, x_{\sigma(\sigma^{-1}(i)-1)}, z, x_{\sigma(\sigma^{-1}(i)+1)}, \dots, x_{\sigma(n)}) = \alpha \cdot L((x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_n) \circ \sigma) + \beta \cdot L((x_1, \dots, x_{i-1}, z, x_{i+1}, \dots, x_n) \circ \sigma) = \alpha \cdot \sigma L(x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_n) + \beta \cdot \sigma L(x_1, \dots, x_{i-1}, z, x_{i+1}, \dots, x_n)$  proving the multilinearity of  $\sigma L$ .  $\square$

**Theorem 10.258.** Let  $X, Y$  be vector spaces over a field  $F$  then given  $n \in \mathbb{N}_0$  and a  $n$ -linear mapping  $L \in \text{Hom}(X^n; Y)$  then  $\sum_{\sigma \in P_n} \text{sign}(\sigma) \cdot \sigma L: X^n \rightarrow Y$  defined by  $x \rightarrow \sum_{\sigma \in P_n} \text{sign}(\sigma) \cdot \sigma L$  is skew symmetric (the sum is well defined for we have proved that  $P_n$  is a finite set (see 10.76)) and  $\sum_{\sigma \in P_n} \text{sign}(\sigma) \cdot \sigma L \in \text{Hom}(X^n, Y)$ .

**Proof.** First as  $\text{Hom}(X^n; Y)$  is a vector space and the above lemma we have that  $\text{sign}(\sigma) \cdot \sigma L$  is  $n$ -linear, then using 10.244 we have that  $\sum_{\sigma \in P_n} \text{sign}(\sigma) \cdot \sigma L$  is  $n$ -linear and thus  $\sum_{\sigma \in P_n} \text{sign}(\sigma) \cdot \sigma L \in \text{Hom}(X^n; Y)$ . Next we have to prove that  $\sum_{\sigma \in P_n} \text{sign}(\sigma) \cdot \sigma L$  is skew symmetric. Take  $\tau \in P_n$  then  $\forall x \in X^n$  we have

$$\begin{aligned}
 \tau \left( \sum_{\sigma \in P_n} \text{sign}(\sigma) \cdot \sigma L \right) (x) &= \left( \sum_{\sigma \in P_n} \text{sign}(\sigma) \cdot \sigma L \right) (x \circ \tau) \\
 &= \sum_{\sigma \in P_n} \text{sign}(\sigma) \cdot \sigma L(x \circ \tau) \\
 &= \sum_{\sigma \in P_n} \text{sign}(\sigma) \cdot L((x \circ \tau) \circ \sigma) \\
 &= \sum_{\sigma \in P_n} \text{sign}(\sigma) \cdot L(x \circ (\tau \circ \sigma)) \\
 &= \sum_{\sigma \in P_n} \text{sign}(\sigma) \cdot (\tau \circ \sigma) L(x) \\
 &\stackrel{\text{sign}(\tau) \cdot \text{sign}(\tau) = 1}{=} \sum_{\sigma \in P_n} \text{sign}(\tau) \cdot \text{sign}(\tau) \cdot \text{sign}(\sigma) \cdot (\tau \circ \sigma) L(x) \\
 &\stackrel{10.93}{=} \text{sign}(\tau) \cdot \sum_{\sigma \in P_n} \text{sign}(\tau \circ \sigma) \cdot (\tau \circ \sigma) L(x) \\
 &\stackrel{10.99}{=} \text{sign}(\tau) \cdot \sum_{\sigma \in P_n} \text{sign}(T_\tau(\sigma)) \cdot T_\tau(\sigma) L(x) \\
 &\stackrel{10.99 \text{ and } 10.43}{=} \text{sign}(\tau) \cdot \sum_{\sigma \in T_\tau(P_n)} \text{sign}(\sigma) \cdot \sigma L(x) \\
 &\stackrel{10.99}{=} \text{sign}(\tau) \cdot \sum_{\sigma \in P_n} \text{sign}(\sigma) \cdot \sigma L(x) \\
 &= \text{sign}(\tau) \cdot \left( \sum_{\sigma \in P_n} \text{sign}(\sigma) \cdot L \right) (x)
 \end{aligned}$$

proving that  $\tau(\sum_{\sigma \in P_n} \text{sign}(\sigma) \cdot \sigma L) = \text{sign}(\tau) \cdot \sum_{\sigma \in P_n} \text{sign}(\sigma) \cdot \sigma L$  and thus skew symmetry.  $\square$

**Theorem 10.259.** Let  $X, Y$  be vector spaces over a field  $F$  that has characteristics of zero (see 4.25) then given  $n \in \mathbb{N}_0$  and a  $n$ -linear mapping  $L \in \text{Hom}(X^n; Y)$  the the following are equivalent:

1.  $L$  is skew symmetric
2. If  $x = (x_1, \dots, x_n) \in X^n$  is such that  $\exists i, j \in \{1, \dots, n\}$  with  $i \neq j$  and  $x_i = x_j$  then  $L(x) = 0$
3. If  $x = \{x_1, \dots, x_n\} \in X^n$  is such that  $\{x_i\}_{i \in \{1, \dots, n\}}$  is linear independent then  $L(x) = 0$

**Proof.**

1. **(1  $\Rightarrow$  2)** Let  $x \in X^n$  with  $i, j \in \{1, \dots, n\}$ ,  $i \neq j$  and  $x_i = x_j$  then we have  $(x \circ (i \leftrightarrow_n j))_k = \begin{cases} x_j & \text{if } k = i \\ x_i & \text{if } k = j \\ x_k & \text{if } k \in \{1, \dots, n\} \setminus \{i, j\} \end{cases}$   $x_i = x_j \Rightarrow \begin{cases} x_i & \text{if } k = i \\ x_j & \text{if } k = j \\ x_k & \text{if } k \in \{1, \dots, n\} \setminus \{i, j\} \end{cases} = x_k$  proving that  $x \circ (i \leftrightarrow_n j) = x$ . Now as  $L$  is skew symmetric we have that  $(i \leftrightarrow_n j)L = \text{sign}(i \leftrightarrow_n j)L$  so that we have ( $u$  is unit in  $F$ )  $-L(x) = \text{sign}(i \leftrightarrow_n j) \cdot L(x) = (i \leftrightarrow_n j)L(x) = L(x \circ (i \leftrightarrow_n j)) = L(x) \Rightarrow L(x) + L(x) = 0 \Rightarrow 0 = u \cdot L(x) + u \cdot L(x) = (1+1) \cdot L(x) = (2 \cdot u) \cdot L(x) \stackrel{2 \cdot u \neq 0 \text{ and 10.109}}{=} L(x) = 0$
2. **(2  $\Rightarrow$  1)** Let  $x = (x_1, \dots, x_n) \in X^n$ ,  $i, j \in \{1, \dots, n\}$  such that  $i \neq j$  define then  $x'$  by  $x'_k = \{x_1, \dots, x_{i-1}, x_i + x_j, x_{i+1}, \dots, x_{j-1}, x_i + x_j, x_{j+1}, \dots, x_n\}$  then as  $x'_i = x_i + x_j = x'_j$  we have that

$$L(x') = 0 \quad (10.70)$$

By using multilinearity we have  $L(x') = L(x'_{i \rightarrow x_i}) + L(x'_{i \rightarrow x_j})$  giving together with 10.70 that

$$L(x'_{i \rightarrow x_i}) + L(x'_{i \rightarrow x_j}) = 0 \quad (10.71)$$

Using multilinearity of  $L$  again we have as  $(x'_{i \rightarrow x_i})_j \stackrel{i \neq j}{=} x'_j = x_i + x_j$  that  $L(x'_{i \rightarrow x_i}) = L((x'_{i \rightarrow x_i})_{j \rightarrow x_i}) + L((x'_{i \rightarrow x_i})_{j \rightarrow x_j}) \stackrel{((x'_{i \rightarrow x_i})_{j \rightarrow x_j})_i = x_i = ((x'_{i \rightarrow x_i})_{j \rightarrow x_i})_j}{=} 0 + L((x'_{i \rightarrow x_i})_{j \rightarrow x_j}) = L(x)$  so we have

$$L(x'_{i \rightarrow x_i}) = L(x) \quad (10.72)$$

As  $(x'_{i \rightarrow x_j})_j = x_i + x_j$  we have by multilinearity that  $L(x'_{i \rightarrow x_j}) = L((x'_{i \rightarrow x_j})_{j \rightarrow x_i}) + L((x'_{i \rightarrow x_j})_{j \rightarrow x_j}) \stackrel{((x'_{i \rightarrow x_j})_{j \rightarrow x_j})_i = x_j = ((x'_{i \rightarrow x_j})_{j \rightarrow x_j})_j}{=} L((x'_{i \rightarrow x_j})_{j \rightarrow x_i}) + 0 = L((x'_{i \rightarrow x_j})_{j \rightarrow x_i})$  giving

$$L(x'_{i \rightarrow x_j}) = L((x'_{i \rightarrow x_j})_{j \rightarrow x_i}) \quad (10.73)$$

For every  $k \in \{1, \dots, n\}$  we have  $((x'_{i \rightarrow x_j})_{j \rightarrow x_j})_k = \begin{cases} (x'_{i \rightarrow x_j})_k & \text{if } k \in \{1, \dots, n\} \setminus \{j\} \\ x_i & \text{if } k = j \end{cases} = \begin{cases} x_k & \text{if } k \in \{1, \dots, n\} \setminus \{i, j\} \\ x_i & \text{if } k = i \\ x_j & \text{if } k = j \end{cases} = (x \circ (i \leftrightarrow_n j))_k$  giving using 10.73 that

$$L(x'_{i \rightarrow x_j}) = L(x \circ (i \leftrightarrow_n j)) \quad (10.74)$$

Using 10.71, 10.72 and 10.74 we have then  $L(x) + L(x \circ (i \leftrightarrow_n j)) \Rightarrow -L(x) = L(x \circ (i \leftrightarrow_n j))$  or

$$\text{If } i, j \in \{1, \dots, n\}, i \neq j \text{ then } (-1) \cdot L = (i \leftrightarrow_n j)L \quad (10.75)$$

We prove now by induction on  $m$  that if  $\{(k_i \leftrightarrow l_i)\}_{i \in \{1, \dots, m\}}$  is a family where  $k_i, l_i \in \{1, \dots, n\}$  and  $k_i \neq l_i$  that

$$[(k_1 \leftrightarrow_n l_1) \circ \dots \circ (k_m \leftrightarrow_n l_m)]L = (-1)^m \cdot L \quad (10.76)$$

**Proof.** So let  $S = \{m \in \{1, \dots, \} \mid \{(k_i \leftrightarrow_n l_i)\}_{i \in \{1, \dots, \}} \text{ is a family where } k_i, l_i \in \{1, \dots, n\} \text{ and } k_i \neq l_i\}$  then  $[(k_1 \leftrightarrow_n l_1) \circ \dots \circ (k_m \leftrightarrow_n l_m)]L = (-1)^m \cdot L$  then we have:

- a. By 10.75 we have that  $1 \in S$
- b. If  $n \in S$  and if  $\{(k_i \leftrightarrow l_i)\}_{i \in \{1, \dots, m+1\}}$  is a family where  $k_i, l_i \in \{1, \dots, n\}$  and  $k_i \neq l_i$  then  $[(k_1 \leftrightarrow_n l_1) \circ \dots \circ (k_{m+1} \leftrightarrow_n l_{m+1})]L = [[(k_1 \leftrightarrow_n l_1) \circ \dots \circ (k_m \leftrightarrow_n l_m)] \circ (k_{m+1} \leftrightarrow_n l_{m+1})] \stackrel{10.85}{=} [(k_1 \leftrightarrow_n l_1) \circ \dots \circ (k_m \leftrightarrow_n l_m)]((k_{m+1} \leftrightarrow_n l_{m+1})L) = (-1) \cdot [(k_1 \leftrightarrow_n l_1) \circ \dots \circ (k_m \leftrightarrow_n l_m)]L \stackrel{m \in S}{=} (-1) \cdot ((-1)^m \cdot L) \stackrel{(-1)^m = +/-1}{=} (-1)^{m+1} \cdot L$  proving that  $m+1 \in S$ .

Using induction we have then that  $S = \{1, \dots\} = \mathbb{N}_0$  and thus 10.76.  $\square$

If now  $\sigma \in P_n$  then by 10.80 there exists a  $\{(k_i \leftrightarrow_n l_i)\}_{i \in \{1, \dots, m\}}$  such that  $k_i, l_i \in \{1, \dots, n\}$  and  $k_i \neq l_i$  and  $\sigma = (k_1 \leftrightarrow_n l_1) \circ \dots \circ (k_m \leftrightarrow_n l_m)$  where by 10.91 we have that  $(-1)^m = \text{sign}(\sigma)$ , using 10.76 we finally arrive at

$$\sigma L = \text{sign}(\sigma) \cdot L$$

proving that  $L$  is skew symmetric.

3. **(2  $\Rightarrow$  3)** If  $x = (x_1, \dots, x_n) \in X^n$  is such that  $\{x_i\}_{i \in \{1, \dots, m\}}$  is linear dependent then using 10.146 there exists a  $i \in \{1, \dots, n\}$ ,  $\{\alpha_j\}_{j \in \{1, \dots, m\} \setminus \{i\}}$  such that  $x_i = \sum_{j \in \{1, \dots, m\} \setminus \{i\}} \alpha_j \cdot x_j$  then as  $x = x_{i \rightarrow x_i}$  we have  $L(x) = L(x_{i \rightarrow x_i})$   $\underset{L \text{ is multilinear and 10.245}}{=} \sum_{j \in \{1, \dots, m\} \setminus \{i\}} \alpha_j \cdot L(x_1, \dots, x_{i-1}, x_j, x_{i+1}, \dots, x_n)$   $\underset{j \neq i \text{ and } (x_1, \dots, x_{i-1}, x_j, x_{i+1}, \dots, x_n)_j = x_j = (x_1, \dots, x_{i-1}, x_j, \dots, x_n)_i}{=} \sum_{j \in \{1, \dots, m\} \setminus \{i\}} \alpha_j \cdot 0 = 0$
4. **(3  $\Rightarrow$  2)** If  $x = (x_1, \dots, x_n) \in X^n$  is such that  $\exists i, j \in \{1, \dots, n\}$  where  $i \neq j$  and  $x_i = x_j$  then  $\{x_i\}_{i \in \{1, \dots, n\}}$  is linear dependent so we have  $L(x) = 0$   $\square$

**Corollary 10.260.** Let  $X, Y$  be vector spaces over a field  $F$  that has characteristics of zero then given  $n \in \mathbb{N}_0, n > 1$ ,  $L \in \text{Hom}(X^n; Y)$  a skew symmetric  $n$ -linear mapping. If now  $X$  is finite dimensional with  $\dim(X) < n$  then  $L = 0$  (the zero mapping)

**Proof.** If  $x = (x_1, \dots, x_n) \in X^n$  then we have the following possibilities

1.  $(\exists i, j \in \{1, \dots, n\} \text{ with } i \neq j \text{ such that } x_i = x_j)$  then using the previous theorem we have as  $L$  is skew symmetric we have  $L(x)$
2.  $(\neg(\exists i, j \in \{1, \dots, n\} \text{ with } i \neq j \text{ such that } x_i = x_j))$  then we have  $\forall i, j \in \{1, \dots, n\}$  we have  $x_i \neq x_j$  so  $x: \{1, \dots, n\} \rightarrow \{x_i | i \in \{1, \dots, n\}\}$  is a bijection and thus  $\#(x_i | i \in \{1, \dots, n\}) = n > \dim(X)$  so by 10.198 we have that  $\{x_i | i \in \{1, \dots, n\}\}$  is linear dependent and using 10.145 we have that  $\{x_i\}_{i \in \{1, \dots, n\}}$  is linear dependent which by the above theorem means that  $L(x) = 0$

So in all cases fro  $x \in X^n$  we have  $L(x) = 0$  and thus  $L = 0$   $\square$

**Theorem 10.261.** Given  $n \in \mathbb{N}_0$   $X, Y$  vector spaces over a field  $F$  of characterization 0,  $\dim(X) = n$  then if  $\{e_i\}_{i \in \{1, \dots, n\}}$  is a basis of  $X$  we have for  $x = (x_1, \dots, x_n) \in X^n$  that  $\forall i \in \{1, \dots, n\}$  there exists a  $\{\xi_i^j\}_{j \in \{1, \dots, n\}}$  such that  $x_i = \sum_{j=1}^n \xi_i^j \cdot e_j$ . Now if  $L \in \text{Hom}(X^n; Y)$  we have

$$L(x_1, \dots, x_n) = \sum_{i=(i_1, \dots, i_n) \in \{1, \dots, n\}^m} \left( \prod_{j \in \{1, \dots, n\}} \xi_j^{i_j} \right) \cdot L(e_{i_1}, \dots, e_{i_n})$$

**Proof.** We proof this by induction on  $n$  so let  $n \in \mathbb{N}_0$  and  $S_n = \{m \in \{1, \dots\} | \text{if } m \leq n \text{ then } L(x_1, \dots, x_n) = \sum_{(i_1, \dots, i_m) \in \{1, \dots, n\}^m} (\prod_{j \in \{1, \dots, m\}} \xi_j^{i_j}) \cdot L(e_{i_1}, \dots, e_{i_m}, x_{m+1}, \dots, x_n)\}$  where  $(e_{i_1}, \dots, e_{i_m}, x_{m+1}, \dots, x_n)_k = \begin{cases} e_{i_k} & \text{if } k \in \{1, \dots, m\} \\ x_k & \text{if } k \in \{m+1, \dots, n\} \end{cases}$  we have then :

1. If  $m = 1$  then

$$\begin{aligned} L(x_1, \dots, x_n) &= L\left(\sum_{i \in \{1, \dots, n\}} \xi_1^i \cdot e_i, x_2, \dots, x_n\right) \\ &= \sum_{i \in \{1, \dots, n\}} \xi_1^i \cdot L(e_i, x_2, \dots, x_n) \\ &= \sum_{i \in \{1, \dots, n\}} \left( \prod_{j \in \{1\}} \xi_j^n \right) \cdot L(e_i, x_2, \dots, x_n) \end{aligned} \tag{10.77}$$

Now  $b: \{1, \dots, n\}^1 \rightarrow \{1, \dots, n\}$  defined by  $i \in \{1, \dots, n\}^1 \rightarrow b(i) = i_1$  is injective and bijective

- a. **(injectivity)** if  $b(i) = b(j) \Rightarrow \forall k \in \{1\}$  we have  $i_k = i_1 = b(i) = b(j) = j_1 = j_k$  so  $i = j$

b. **(surjectivity)** If  $k \in \{1, \dots, n\}$  then if we define  $i \in \{1, \dots, n\}^1$  by  $i_1 = k$  we have  $b(i) = i_1 = k$

Using 10.43 and the above equation we have

$$\begin{aligned} L(x_1, \dots, x_n) &= \sum_{i=(i_1) \in \{1, \dots, n\}^1} \left( \prod_{j \in \{1\}} \xi_j^{b(i_1)} \right) \cdot L(e_{b(i_1)}, x_1, \dots, x_n) \\ &\stackrel{b(i)=i_1}{=} \sum_{i=(i_1) \in \{1, \dots, n\}^1} \left( \prod_{j \in \{1\}} \xi_j^{i_1} \right) \cdot L(e_{i_1}, x_1, \dots, x_n) \end{aligned}$$

proving that  $1 \in S_n$ .

2. If  $m \in S$  then if  $m+1 \leq n$  we have

$$\begin{aligned} L(x_1, \dots, x_n) &\stackrel{m \in S_n}{=} \sum_{i=(1, \dots, n)^m} \left( \prod_{j \in \{1, \dots, m\}} \xi_j^{i_j} \right) \cdot L(e_{i_1}, \dots, e_{i_m}, x_{m+1}, \dots, x_n) \\ &= \sum_{i=(1, \dots, n)^m} \left( \prod_{j \in \{1, \dots, m\}} \xi_j^{i_j} \right) \cdot L\left(e_{i_1}, \dots, e_{i_m}, \sum_{k \in \{1, \dots, n\}} \xi_{m+1}^k e_k, \dots, x_n\right) \\ &= \sum_{i \in \{1, \dots, n\}^m} \left( \sum_{k \in \{1, \dots, n\}} \left( \left( \prod_{j \in \{1, \dots, m\}} \xi_j^{i_j} \right) \cdot \xi_{m+1}^k \right) \cdot L(e_{i_1}, \dots, e_{i_m}, e_k, \dots, x_n) \right) \\ &= \sum_{i \in \{1, \dots, n\}^m} \left( \sum_{k \in \{1, \dots, m\}} A_{i,k} \right) \end{aligned}$$

so we have

$$L(x_1, \dots, x_n) = \sum_{i \in \{1, \dots, n\}^m} \left( \sum_{k \in \{1, \dots, m\}} A_{i,k} \right) \quad (10.78)$$

where

$$A_{i,k} = \left( \left( \prod_{j \in \{1, \dots, m\}} \xi_j^{i_j} \right) \cdot \xi_{m+1}^k \right) \cdot L(e_{i_1}, \dots, e_{i_m}, e_k, \dots, x_n) \quad (10.79)$$

Given  $i \in \{1, \dots, n\}^m$  define now  $I_i = \{i\} \times \{1, \dots, n\}$  and  $b_i: \{i\} \times \{1, \dots, n\} \rightarrow \{1, \dots, n\}$  by  $(i, k) \mapsto k$  which is trivially a bijection. Using 10.43 and 10.78 we have then  $\sum_{k \in \{1, \dots, m\}} A_{i,k} = \sum_{k \in I_i} A_{i,b_i(k)}$   $\stackrel{k=(i,l) \Rightarrow k_1=i, b_i(k)=k_2}{=} \sum_{k \in I_i} A_{k_1, k_2}$

$$L(x_1, \dots, x_n) = \sum_{i \in \{1, \dots, n\}^m} \left( \sum_{k \in I_i} A_{k_1, k_2} \right) \quad (10.80)$$

If now  $i, j \in \{1, \dots, n\}^m$  and  $i \neq j$  then if  $m \in I_i \cap I_j$  we have that there exists a  $k, l \in \{1, \dots, n\}$  such that  $m = (i, k)$  and  $m = (j, l)$  but then  $(i, k) = (j, l) \Rightarrow i = j$  contradicting  $i \neq j$  so we have  $I_i \cap I_j$ . If we take now

$$I = \bigcup_{i \in \{1, \dots, n\}^m} I_i \quad (10.81)$$

which is a finite union (see 10.106) of disjoint set, we can use the 10.46 to get

$$\sum_{i \in I} A_{i_1, i_2} = \sum_{i \in \{1, \dots, n\}^m} \left( \sum_{k \in I_i} A_{k_1, k_2} \right) \quad (10.82)$$

Define now  $h: \{1, \dots, n\}^{m+1} \rightarrow \bigcup_{i \in \{1, \dots, n\}^m} I_i$  by  $i \mapsto h(i) = (i|_{\{1, \dots, m\}}, i_{m+1})$  then we have:

a. **(injectivity)** If  $h(i) = h(j)$  then  $(i|_{\{1, \dots, m\}}, i_{m+1}) = (j|_{\{1, \dots, m\}}, j_{m+1}) \Rightarrow i|_{\{1, \dots, m\}} = j|_{\{1, \dots, m\}}$  and  $i_{m+1} = j_{m+1} \Rightarrow i = j$ .

- b. **(surjectivity)** If  $z \in \bigcup_{i \in \{1, \dots, n\}^m} I_i$  then there exists a  $i \in \{1, \dots, n\}^m$  such that  $z \in I_i$  and thus there exists a  $k \in \{1, \dots, n\}$  such that  $z = (i, k)$ . Define then  $l \in \{1, \dots, n\}^{m+1}$  by  $\forall r \in \{1, \dots, n\}$ ,  $l_r = \begin{cases} i_r & \text{if } r \in \{1, \dots, m\} \\ k & \text{if } r = m+1 \end{cases}$  then  $h(l) = (l_{\{1, \dots, m\}}, l_{m+1}) = (i, k)$

which proves that  $h$  is a bijection. Using 10.43 and 10.82 we have then that

$$\sum_{i \in \{1, \dots, n\}^{m+1}} A_{h(i)_1, h(i)_2} = \sum_{i \in I} A_{i_1, i_2} \quad (10.83)$$

Now  $A_{h(i)_1, h(i)_2} = A_{i_{\{1, \dots, m\}}, i_{m+1}} = \left( \left( \prod_{j \in \{1, \dots, m\}} \xi_j^{(i_{\{1, \dots, m\}})_j} \right) \cdot \xi_{m+1}^{i_{m+1}} \right) \cdot L(e_{i_1}, \dots, e_{i_m}, e_{i_{m+1}}, \dots, e_n) = \left( \prod_{j \in \{1, \dots, m+1\}} \xi_j^{i_j} \right) \cdot L(e_{i_1}, \dots, e_{i_{m+1}}, \dots, e_n)$  so that by using 10.79, 10.82 and the above gives finally that

$$L(x_1, \dots, x_n) = \sum_{i \in \{1, \dots, n\}^{m+1}} \left( \prod_{j \in \{1, \dots, m+1\}} \xi_j^{i_j} \right) \cdot L(e_{i_1}, \dots, e_{i_{m+1}}, \dots, x_n)$$

proving that  $m+1 \in S_n$ .

Using mathematical induction we have then  $S_n = \{1, \dots, n\} = \mathbb{N}_0$  and thus if  $n \in \mathbb{N}_0 = S_n$  and  $n \leq n$  we have  $L(x_1, \dots, x_n) = \sum_{i \in \{1, \dots, n\}^m} \left( \prod_{j \in \{1, \dots, n\}} \xi_j^{i_j} \right) \cdot L(e_{i_1}, \dots, e_{i_n})$   $\square$

**Theorem 10.262.** Given  $n \in \mathbb{N}_0$   $X, Y$  vector spaces over a field  $F$  of characterization 0,  $\dim(X) = n$  then if  $\{e_i\}_{i \in \{1, \dots, n\}}$  is a basis of  $X$  we have for  $x = (x_1, \dots, x_n) \in X^n$  that  $\forall i \in \{1, \dots, n\}$  there exists a  $\{\xi_i^j\}_{j \in \{1, \dots, n\}}$  such that  $x_i = \sum_{j=1}^n \xi_i^j \cdot e_j$ . Now if  $L \in \text{Hom}(X^n, Y)$  is a skew symmetric  $n$ -linear mapping then

$$L(x_1, \dots, x_n) = \sum_{\sigma \in P_n} \left( \prod_{j \in \{1, \dots, m\}} \xi_j^{\sigma(j)} \right) \cdot \sigma L(e_1, \dots, e_n)$$

**Proof.** Now as  $P_n$  is the set of all bijections between  $\{1, \dots, n\}$  and  $\{1, \dots, n\}$  we have that  $P_n \subseteq \{1, \dots, n\}^n$ . Now if  $i \in \{1, \dots, n\}^n \setminus P_n$  then  $i$  is not a bijection and thus not a injection [if  $i$  would be injective then  $i: \{1, \dots, n\} \rightarrow i(\{1, \dots, n\})$  would be a bijection, and if  $i(\{1, \dots, n\}) = \{1, \dots, n\}$  we would have that  $i$  is a bijection so we must have that  $i(\{1, \dots, n\}) \subset \{1, \dots, n\}$  giving  $n = \#\{1, \dots, n\} = \#\{1, \dots, n\} \approx \#\{1, \dots, n\} = \#\{i(\{1, \dots, n\})\} < \#\{1, \dots, n\} = n$  reaching the contradiction  $n < n$ , so there exists a  $k, l \in \{1, \dots, n\}$  such that  $i_k = i_l \Rightarrow e_{i_k} = e_{i_l}$  and thus  $L(e_{i_1}, \dots, e_{i_n}) = 0$ . Using the previous theorem we have then

$$\begin{aligned} L(x_1, \dots, x_n) &= \sum_{\sigma \in \{1, \dots, n\}^n} \left( \prod_{j \in \{1, \dots, n\}} \xi_j^{\sigma(j)} \right) \cdot L(e_{\sigma_1}, \dots, e_{\sigma_n}) \\ &= \sum_{\sigma \in \{1, \dots, n\}^n \setminus P_n} \left( \prod_{j \in \{1, \dots, n\}} \xi_j^{\sigma(j)} \right) \cdot L(e_{\sigma_1}, \dots, e_{\sigma_n}) + \sum_{\sigma \in P_n} \left( \prod_{j \in \{1, \dots, n\}} \xi_j^{\sigma(j)} \right) \cdot L(e_{\sigma_1}, \dots, e_{\sigma_n}) \\ &= \sum_{\sigma \in \{1, \dots, n\}^n \setminus P_n} \left( \prod_{j \in \{1, \dots, n\}} \xi_j^{\sigma(j)} \right) \cdot 0 + \sum_{\sigma \in P_n} \left( \prod_{j \in \{1, \dots, n\}} \xi_j^{\sigma(j)} \right) \cdot L(e_{\sigma_{(1)}}, \dots, e_{\sigma(n)}) \\ &= \sum_{\sigma \in P_n} \left( \prod_{j \in \{1, \dots, n\}} \xi_j^{\sigma(j)} \right) \cdot L(e_{\sigma_{(1)}}, \dots, e_{\sigma(n)}) \\ &= \sum_{\sigma \in P_n} \left( \prod_{j \in \{1, \dots, n\}} \xi_j^{\sigma(j)} \right) \cdot \sigma L(e_1, \dots, e_n) \end{aligned}$$

$\square$

**Corollary 10.263.** Given  $n \in \mathbb{N}_0$   $X, Y$  vector spaces over a field  $F$  of characterization 0,  $\dim(X) = n$  then if  $\{e_i\}_{i \in \{1, \dots, n\}}$  is a basis of  $X$  we have for  $x = (x_1, \dots, x_n) \in X^n$  that  $\forall i \in \{1, \dots, n\}$  there exists a  $\{\xi_i^j\}_{j \in \{1, \dots, n\}}$  such that  $x_i = \sum_{j=1}^n \xi_i^j \cdot e_j$ . Now if  $L \in \text{Hom}(X^n; Y)$  is a skew symmetric  $n$ -linear mapping then

$$L(x_1, \dots, x_n) = \left( \sum_{\sigma \in P_n} \text{sign}(\sigma) \cdot \prod_{j \in \{1, \dots, n\}} \xi_j^{\sigma(j)} \right) \cdot L(e_1, \dots, e_n)$$

Proving that  $L$  depends only on the values of  $L$  at  $(e_1, \dots, e_n)$  [in other words if  $L_1, L_2$  are two skew symmetric  $n$ -linear mappings such that  $L_1(e_1, \dots, e_n) = L_2(e_1, \dots, e_n)$  then  $L_1 = L_2$ ]. This also proves that if  $L$  is a  $n$ -linear skew symmetric mapping such that  $L(e_1, \dots, e_n) = 0$  then  $L = 0$ .

**Proof.** As  $L$  is a  $n$ -linear skew symmetric mapping we have using the previous theorem that

$$\begin{aligned} L(x_1, \dots, x_n) &= \sum_{\sigma \in P_n} \left( \prod_{j \in \{1, \dots, n\}} \xi_j^{\sigma(j)} \right) \cdot \sigma L(e_1, \dots, e_n) \\ &\stackrel{L \text{ is skewsymmetric}}{=} \sum_{\sigma \in P_n} \left( \prod_{j \in \{1, \dots, n\}} \xi_j^{\sigma(j)} \right) \cdot (\text{sign}(\sigma) \cdot L(e_1, \dots, e_n)) \\ &= \left( \sum_{\sigma \in P_n} \text{sign}(\sigma) \cdot \prod_{j \in \{1, \dots, n\}} \xi_j^{\sigma(j)} \right) \cdot L(e_1, \dots, e_n) \end{aligned}$$

□

**Definition 10.264.** Given  $n \in \mathbb{N}_0$  let  $X$  be a  $n$ -dimensional vector space over a field  $F$  then a determinant function  $\Delta$  in  $X$  is a skew symmetric  $n$ -linear mapping  $\Delta: X^n \rightarrow F$

**Theorem 10.265.** Given  $n \in \mathbb{N}_0$  let  $X$  be a  $n$ -dimensional vector space over a field  $F$  of characterization zero then if  $\{e_i\}_{i \in \{1, \dots, n\}}$  is a basis for  $X$  there exists a determinant function  $\Delta$  such that  $\Delta(e_1, \dots, e_n) = 1$  (the unit in  $F$ ). This proves that there exist a non trivial determinant function (meaning a determinant function different from the zero mapping).

**Proof.** Let  $\{e_i\}_{i \in \{1, \dots, n\}}$  be a basis in  $X$  then  $\forall i \in \{1, \dots, n\}$  define  $f_i: X \rightarrow F$  by  $x = \sum_{i \in \{1, \dots, n\}} \alpha_i \cdot x_i \rightarrow f_i(x) = \alpha_i$  (which is well defined as for every  $x \in X$  there exist a unique expansion of  $x$  in the basis  $e_i$  and is linear [as for if  $\alpha, \beta \in F$  and  $x, y \in X$  with  $x = \sum_{i \in \{1, \dots, n\}} \alpha_i \cdot e_i$ ,  $y = \sum_{i \in \{1, \dots, n\}} \beta_i \cdot e_i$  then  $f_i(\alpha \cdot x + \beta \cdot y) = f_i(\alpha \cdot \sum_{j \in \{1, \dots, n\}} \alpha_j \cdot e_j + \beta \cdot \sum_{j \in \{1, \dots, n\}} \beta_j \cdot e_j) = f_i(\sum_{j \in \{1, \dots, n\}} (\alpha \cdot \alpha_j + \beta \cdot \beta_j) \cdot e_j) = \alpha \cdot \alpha_i + \beta \cdot \beta_i = \alpha \cdot f_i(x) + \beta \cdot f_i(y)$ ]. As  $\forall j \in \{1, \dots, n\}$  we have  $e_j = \sum_{i \in \{1, \dots, n\}} \delta_{i,j} \cdot e_i$  we have that  $f_i(e_j) = \delta_{i,j}$ . Given  $x = (x_1, \dots, x_n) \in X^n$  define then

$$\Phi: X^n \rightarrow F \text{ by } (x_1, \dots, x_n) \rightarrow \Phi(x_1, \dots, x_n) = \prod_{j \in \{1, \dots, n\}} f_j(x_j)$$

If  $i \in \{1, \dots, n\}$ ,  $\alpha, \beta \in F$  and  $x, y \in X$  then we have

$$\begin{aligned} \Phi(x_1, \dots, x_{i-1}, \alpha \cdot x + \beta \cdot y, x_{i+1}, \dots, x_n) &= f_i(\alpha \cdot x + \beta \cdot y) \cdot \prod_{j \in \{1, \dots, n\} \setminus \{i\}} f_j(x_j) \\ &= (\alpha \cdot f_i(x) + \beta \cdot f_i(y)) \cdot \prod_{j \in \{1, \dots, n\} \setminus \{i\}} f_j(x_j) \\ &= \alpha \cdot \left( f_i(x) \cdot \prod_{j \in \{1, \dots, n\} \setminus \{i\}} f_j(x_j) \right) + \beta \cdot \left( f_i(y) \cdot \prod_{j \in \{1, \dots, n\} \setminus \{i\}} f_j(x_j) \right) \\ &= \alpha \cdot \Phi(x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_n) + \beta \cdot \Phi(x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_n) \end{aligned}$$

proving that  $\Phi$  is  $n$ -linear. Using 10.258 we create the skew symmetric  $n$ -linear mapping

$$\Delta: X^n \rightarrow F \text{ by } \Delta = \sum_{\sigma \in P_n} \text{sign}(\sigma) \cdot \sigma \Phi$$

that is then by definition a determinant function. Now as  $f_i(e_j) = \delta_{i,j}$  we have if  $\sigma \in P_n \setminus 1_{\{1, \dots, n\}}$  that there exists a  $i \in \{1, \dots, n\}$  such that  $\sigma(i) \neq i$  and thus  $f_i(e_{\sigma(i)}) = 0$  and thus  $\sigma\Phi(e_1, \dots, e_n) = \prod_{j \in \{1, \dots, n\}} f_j(e_{\sigma(j)}) = f_i(e_i) \cdot \prod_{j \in \{1, \dots, n\} \setminus \{i\}} f_j(e_j) = 0 \cdot \prod_{j \in \{1, \dots, n\} \setminus \{i\}} f_j(e_j) = 0$ , if  $\sigma = 1_{\{1, \dots, n\}}$  then  $\sigma\Phi(e_1, \dots, e_n) = \prod_{j \in \{1, \dots, n\}} f_j(e_j) = \prod_{j \in \{1, \dots, n\}} 1 = 1$  this gives then

$$\sigma\Phi(e_1, \dots, e_n) = \begin{cases} 0 & \text{if } \sigma \neq 1_{\{1, \dots, n\}} \\ 1 & \text{if } \sigma = 1_{\{1, \dots, n\}} \end{cases}$$

So we have finally  $\Delta(e_1, \dots, e_n) = \sum_{\sigma \in P_n} \text{sign}(\sigma) \cdot \sigma\Phi(e_1, \dots, e_n) = \sum_{\sigma \in \{1_{\{1, \dots, n\}}\}} \text{sign}(\sigma) \cdot \sigma\Phi(e_1, \dots, e_n) + \sum_{\sigma \in P_n \setminus \{1_{\{1, \dots, n\}}\}} \text{sign}(\sigma) \cdot \sigma\Phi(e_1, \dots, e_n) = 1 + 9 = 10$   $\square$

**Theorem 10.266.** *Given  $n \in \mathbb{N}_0$  let  $X$  be a  $n$ -dimensional vector space over a field  $F$  of characterization zero,  $Y$  a vector space over  $F$  and  $\Delta$  a non zero determinant function. Then for every skew symmetric  $n$ -linear function  $L: X^n \rightarrow Y$  there exists a unique  $y \in Y$  such that  $\forall (x_1, \dots, x_n) \in X^n$  we have that  $L(x_1, \dots, x_n) = \Delta(x_1, \dots, x_n) \cdot y$*

**Proof.** Let  $\{e_i\}_{i \in \{1, \dots, n\}}$  a basis of  $X$  then using 10.263 and the fact that  $\Delta \neq 0$  we must have that  $\Delta(e_1, \dots, e_n) \neq 0$ . Define now  $\{e'_i\}_{i \in \{1, \dots, n\}}$  by  $e'_i = \begin{cases} \Delta(e_1, \dots, e_n)^{-1} \cdot e_1 & \text{if } i = 1 \\ e_i & \text{if } i \in \{2, \dots, n\} \end{cases}$ . We prove now that  $\{e'_i\}_{i \in \{1, \dots, n\}}$  is still a basis for  $X$ .

**Proof.** If  $\{\alpha_i\}_{i \in \{1, \dots, n\}}$  is a family in  $F$  such that  $\sum_{i \in \{1, \dots, n\}} \alpha_i \cdot e'_i = 0$  then if we define  $\{\alpha'_i\}_{i \in \{1, \dots, n\}}$  by  $\alpha'_i = \begin{cases} \Delta(e_1, \dots, e_n)^{-1} \cdot \alpha_1 & \text{if } i = 1 \\ \alpha_i & \text{if } i \in \{2, \dots, n\} \end{cases}$  we have  $\alpha'_i \cdot e_i = \begin{cases} \alpha_1 \cdot \Delta(e_1, \dots, e_n)^{-1} \cdot e_1 = \alpha_1 \cdot e'_1 & \text{if } i = 1 \\ \alpha_i \cdot e'_i & \text{if } i \in \{2, \dots, n\} \end{cases} = \alpha_i \cdot e'_i$  and  $\sum_{i \in \{1, \dots, n\}} \alpha'_i \cdot e_i = \sum_{i \in \{1, \dots, n\}} \alpha_i \cdot e'_i = 0$ . So  $\forall i \in \{1, \dots, n\}$  we have  $\alpha'_i = 0 \Rightarrow 0 = \begin{cases} \Delta(e_1, \dots, e_n)^{-1} \cdot \alpha_1 & \text{if } i = 1 \\ \alpha_i & \text{if } i \in \{2, \dots, n\} \end{cases} \Rightarrow \alpha_i = 0$  proving linear Independence. If  $x \in X$  then there exists a  $\{\beta_i\}_{i \in \{1, \dots, n\}}$  such that  $\sum_{i \in \{1, \dots, n\}} \beta_i \cdot e_i = x$ , so if we take  $\{\beta'_i\}_{i \in \{1, \dots, n\}}$  where  $\beta'_i = \begin{cases} \Delta(e_1, \dots, e_n) \cdot \beta & \text{if } i = 1 \\ \beta_i & \text{if } i \in \{2, \dots, n\} \end{cases}$  then we have  $\beta'_i \cdot e'_i = \begin{cases} \Delta(e_1, \dots, e_n) \cdot (\Delta(e_1, \dots, e_n)^{-1} \cdot e_1) & \text{if } i = 1 \\ \beta_i \cdot e'_i & \text{if } i \in \{2, \dots, n\} \end{cases} = \beta_i \cdot e_i$ . So  $\sum_{i \in \{1, \dots, n\}} \beta'_i \cdot e'_i = \sum_{i \in \{1, \dots, n\}} \beta_i \cdot e_i = x$   $\square$

Now  $\Delta(e'_1, \dots, e'_n) = \Delta(\Delta(e_1, \dots, e_n)^{-1} \cdot e_1, \dots, e_n) = \Delta(e_1, \dots, e_n)^{-1} \cdot \Delta(e_1, \dots, e_n) = 1$ . Set  $y = L(e'_1, \dots, e'_n)$  and define  $K: X^n \rightarrow Y$  by  $(x_1, \dots, x_n) \rightarrow K(x_1, \dots, x_n) = \Delta(x_1, \dots, x_n) \cdot y$ , then because  $\Delta$  is a skew symmetric  $n$ -linear mapping it is trivial to show that  $K$  is also skew symmetric and  $n$ -linear mapping. Now  $K(e'_1, \dots, e'_n) = \Delta(e'_1, \dots, e'_n) \cdot y = 1 \cdot y = y = L(e'_1, \dots, e'_n)$  and using 10.263 we have then  $K = L$  proving that  $L(x_1, \dots, x_n) = \Delta(x_1, \dots, x_n) \cdot y$ . Now to prove that  $y$  is unique assume that also  $L(x_1, \dots, x_n) = \Delta(x_1, \dots, x_n) \cdot y'$  then  $\Delta(e'_1, \dots, e'_n) \cdot y = \Delta(e'_1, \dots, e'_n) \cdot y' \Rightarrow 1 \cdot y = 1 \cdot y' \Rightarrow y = y'$ .  $\square$

**Corollary 10.267.** *Given  $n \in \mathbb{N}_0$  let  $X$  be a  $n$ -dimensional vector space over a field  $F$  of characterization zero. Let  $\Delta$  be a non zero determinant function then every other determinant function is a scalar multiple of  $\Delta$*

**Proof.** If  $\Delta': X^n \rightarrow F$  is a another determinant function then by the above theorem there exists a  $y \in F$  such that  $\Delta' = \Delta \cdot y$   $\square$

**Definition 10.268.** *Let  $n \in \mathbb{N}_0$ ,  $X$  a set then given  $i \in \{1, \dots, n\}$  and  $y \in X$  and  $x = (x_1, \dots, x_n) \in X^n$  we define  $(y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n) \in X^n$  to be defined by  $\forall k \in \{1, \dots, n\}$  we have  $(y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_k = \begin{cases} y & \text{if } k = 1 \\ x_{k-1} & \text{if } k \in \{2, \dots, i\} \\ x_k & \text{if } k \in \{i+1, \dots, n\} \end{cases}$*

**Example 10.269.** if  $n = 1$  and  $x = (x_1) \in X^1$  and  $y \in X$  then by the above definition we have that  $(y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n) = (y)$

**Lemma 10.270.** *Let  $n \in \mathbb{N}_0$ ,  $X$  a vector space over a field  $F$  then given  $i \in \{1, \dots, n\}$ ,  $j \in \{1, \dots, n\}$  and  $y \in X$  and  $x = (x_1, \dots, x_n) \in X^n$  then we have*

1.  $(j+1 \leq i)$  then  $((y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_{j+1 \rightarrow x_i} \circ (j+1 \rightsquigarrow_n i)) = (y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)$

2.  $(i+1 \leq j)$  then  $((y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_{j \rightarrow x_i} \circ (j \rightsquigarrow_n i+1)) = (y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)$

**Proof.**

1.  $(j+1 < i)$  Let  $k \in \{1, \dots, n\}$  then we have (using the definition of  $\rightsquigarrow$ ) the following cases to consider for  $A_k = ((y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_{j+1 \rightarrow x_i} \circ (j+1 \rightsquigarrow_n i))_k$ :

a.  $(k=1)$  then

$$\begin{aligned} A_k &\stackrel{k=1 < j+1}{=} ((y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_{j+1 \rightarrow x_i})_k \\ &\stackrel{k=1 \neq j+1}{=} (y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_k \\ &\stackrel{k=1}{=} y \\ &\stackrel{k=1}{=} (y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_k \end{aligned}$$

b.  $(k \neq 1 \wedge k < j+1)$  then

$$\begin{aligned} A_k &= ((y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_{j+1 \rightarrow x_i})_k \\ &\stackrel{k \neq j+1}{=} (y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_k \\ &\stackrel{k \neq 1 \wedge k < j+1 \leq i}{=} x_{k-1} \\ &\stackrel{k \neq 1 \wedge k < j+1 \Rightarrow k \leq j}{=} (y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_k \end{aligned}$$

c.  $(k \neq 1 \wedge j+1 \leq k < i)$  then

$$\begin{aligned} A_k &= ((y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_{j+1 \rightarrow x_i})_{k+1} \\ &\stackrel{j+1 \leq k < k+1}{=} (y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_{k+1} \\ &\stackrel{k \neq 1 \wedge k < i+1 \Rightarrow k+1 \leq i}{=} (y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_k \end{aligned}$$

d.  $(k \neq 1 \wedge k = i)$  then

$$\begin{aligned} A_k &= ((y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_{j+1 \rightarrow x_i})_{j+1} \\ &= x_i \\ &\stackrel{j < i = k}{=} (y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_k \end{aligned}$$

e.  $(k \neq 1 \wedge i < k)$  then

$$\begin{aligned} A_k &= ((y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_{j+1 \rightarrow x_i})_k \\ &\stackrel{j+1 < i < k}{=} (y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_k \\ &= x_k \\ &\stackrel{j+1 < i < k}{=} (y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_k \end{aligned}$$

So  $\forall k \in \{1, \dots, n\}$  we have  $((y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_{j+1 \rightarrow x_i} \circ (j+1 \rightsquigarrow_n i))_k = (y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_k$  or  $(y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_{j+1 \rightarrow x_i} \circ (j+1 \rightsquigarrow_n i) = (y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)$ .

2.  $(j+1 = i)$  Then by the definition of  $(j+1 \rightsquigarrow_n i)$  we have  $(j+1 \rightsquigarrow_n i) = 1_{\{1, \dots, n\}}$  and thus  $(y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_{j+1 \rightarrow x_i} \circ (j+1 \rightsquigarrow_n i) = (y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_{j+1 \rightarrow x_i}$  and we have then  $\forall k \in \{1, \dots, n\}$  to consider for  $A_k = ((y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_{j+1 \rightarrow x_i} \circ (j+1 \rightsquigarrow_n i))_k = ((y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_{j+1 \rightarrow x_i})_k$  the following cases:

a.  $(k=1)$  then

$$\begin{aligned} A_k &\stackrel{k=1 < j+1}{=} (y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_k \\ &\stackrel{k=1}{=} y \\ &\stackrel{k=1}{=} (y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_k \end{aligned}$$

b.  $(k \neq 1 \wedge k < j)$  then

$$\begin{aligned} A_k &\stackrel{k < j \leq j+1}{=} (y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_k \\ &\stackrel{k < j \leq j+1 = i}{=} x_{k-1} \\ &\stackrel{k \neq 1 \wedge k < j}{=} (y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_k \end{aligned}$$

c.  $(k \neq 1 \wedge k = j)$  then

$$\begin{aligned} A_k &\stackrel{k = j \neq j+1}{=} (y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_k \\ &\stackrel{k = j < j+1 = i}{=} x_{k-1} \\ &\stackrel{k \neq 1 \wedge k = j}{=} (y, x_1, \dots, x_{j-1}, x_j, \dots, x_n)_k \end{aligned}$$

d.  $(k \neq 1 \wedge k = j+1)$  then

$$\begin{aligned} A_k &\stackrel{k = j+1}{=} x_i \\ &\stackrel{k = j+1 = i}{=} x_k \\ &\stackrel{j < j+1 = k}{=} (y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_k \end{aligned}$$

e.  $(k \neq 1 \wedge j+1 < k)$  then

$$\begin{aligned} A_k &\stackrel{j+1 < k}{=} (y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_k \\ &\stackrel{i = j+1 < k \Rightarrow i+1 \leq k}{=} x_k \\ &\stackrel{j < j+1 < k}{=} (y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_k \end{aligned}$$

So  $\forall k \in \{1, \dots, n\}$  we have  $((y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_{j+1 \rightarrow x_i} \circ (j+1 \rightsquigarrow_n i))_k = (y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_k$  or  $(y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_{j+1 \rightarrow x_i} \circ (j+1 \rightsquigarrow_n i) = (y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_k$ .

3.  $(i+1 < j)$  Let  $k \in \{1, \dots, n\}$  then using the definition of  $\rightsquigarrow$  we have to consider the following cases for  $A_k = ((y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_{j \rightarrow x_i} \circ (j \rightsquigarrow_n i+1))_k$

a.  $(k=1)$  then

$$\begin{aligned} A_k &\stackrel{k=1 \leq i+1}{=} ((y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_{j \rightarrow x_i})_k \\ &\stackrel{k=1 < i+1 < j}{=} (y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_k \\ &\stackrel{k=1}{=} y \\ &\stackrel{k=1}{=} (y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_k \end{aligned}$$

b.  $(k \neq 1 \wedge k < i+1)$  then

$$\begin{aligned} A_k &= ((y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_{j \rightarrow x_i})_k \\ &\stackrel{k < i+1 < j}{=} (y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_k \\ &\stackrel{k \neq 1 \wedge k < i+1 \Rightarrow k \leq i}{=} x_{k-1} \\ &\stackrel{k < i+1 < j}{=} (y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_k \end{aligned}$$

c.  $(k \neq 1 \wedge k = i+1)$  then

$$\begin{aligned} A_k &= ((y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_{j \rightarrow x_i})_j \\ &= x_i \\ &\stackrel{k \neq 1 \wedge k = i+1 < j}{=} (y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_{k=i+1} \end{aligned}$$

d.  $(k \neq 1 \wedge i+1 < k \leq j)$  then

$$\begin{aligned} A_k &= ((y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_{j \rightarrow x_i})_{k-1} \\ &\stackrel{k \leq j \Rightarrow k-1 < j}{=} (y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_{k-1} \\ &\stackrel{i+1 < k \Rightarrow i < k-1}{=} x_{k-1} \\ &\stackrel{k \neq 1 \wedge k \leq j}{=} (y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_k \end{aligned}$$

e.  $(k \neq 1 \wedge j < k)$  then

$$\begin{aligned}
 A_k &= ((y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_{j \rightarrow x_i})_k \\
 &\stackrel{j < k}{=} (y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_k \\
 &\stackrel{i+1 < j < k \Rightarrow i < k}{=} x_k \\
 &\stackrel{k \neq 1 \wedge j < k}{=} (y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_k
 \end{aligned}$$

So  $\forall k \in \{1, \dots, n\}$  we have  $((y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_{j \rightarrow x_i} \circ (j \rightsquigarrow_n i + 1))_k = (y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_k$  and thus  $(y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_{j \rightarrow x_i} \circ (j \rightsquigarrow_n i + 1) = (y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)$

4.  $(i+1 = j)$  Then by the definition of  $(j \rightsquigarrow_n i + 1)$  we have that  $(j \rightsquigarrow_n i + 1) = 1_{\{1, \dots, n\}}$  and thus we have that  $(y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_{j \rightarrow x_i} \circ (j \rightsquigarrow_n i + 1) = (y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_{j \rightarrow x_i}$  then  $\forall k \in \{1, \dots, n\}$  we have for  $A_k = ((y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_{j \rightarrow x_i} \circ (j \rightsquigarrow_n i + 1))_k = ((y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_{j \rightarrow x_i})_k$  the following cases to consider

a.  $(k = 1)$  then

$$\begin{aligned}
 A_k &\stackrel{k=1 \neq i+1=j}{=} (y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_k \\
 &\stackrel{k=1}{=} y \\
 &\stackrel{k=1}{=} (y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_k
 \end{aligned}$$

b.  $(k \neq 1 \wedge k < i)$  then

$$\begin{aligned}
 A_k &\stackrel{k < i < i+1=j}{=} (y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_k \\
 &\stackrel{k \neq 1 \wedge k < i}{=} x_{k-1} \\
 &\stackrel{k \neq 1 \wedge k < i < i+1=j}{=} (y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_k
 \end{aligned}$$

c.  $(k \neq 1 \wedge k = i)$  then

$$\begin{aligned}
 A_k &\stackrel{k=i \neq i+1=j}{=} (y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_k \\
 &\stackrel{k=i \leq i}{=} x_{k-1} \\
 &\stackrel{k=i < i+1=j}{=} (y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_k
 \end{aligned}$$

d.  $(k \neq 1 \wedge k = i + 1)$  then

$$\begin{aligned}
 A_k &\stackrel{k=i+1=j}{=} x_i \\
 &\stackrel{k=i+1}{=} x_{k-1} \\
 &\stackrel{k=i+1=j \leq j}{=} (y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_k
 \end{aligned}$$

e.  $(k \neq 1 \wedge i + 1 < k)$  then

$$\begin{aligned}
 A_k &\stackrel{j=i+1 < k}{=} (y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_k \\
 &\stackrel{i < i+1 < k}{=} x_k \\
 &\stackrel{j=i+1 < k}{=} (y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_k
 \end{aligned}$$

So  $\forall k \in \{1, \dots, n\}$  we have  $((y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_{j \rightarrow x_i} \circ (j \rightsquigarrow_n i + 1))_k = (y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_k$  and thus  $(y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_{j \rightarrow x_i} \circ (j \rightsquigarrow_n i + 1) = (y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)$

□

**Definition 10.271.** Let  $n \in \mathbb{N}_0$  then if  $X$  is a vector space over a field  $F$  and  $\Delta: X^n \rightarrow F$  a determinant function in  $X$  then we define  $\hat{\Delta}: X \times X^n \rightarrow X$  by  $(y, (x_1, \dots, x_n)) \rightarrow \hat{\Delta}(y, (x_1, \dots, x_n)) = \sum_{j \in \{1, \dots, n\}} (-1)^{j-1} \Delta(y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n) \cdot x_j$

**Theorem 10.272.** Let  $n \in \mathbb{N}_0$  then if  $X$  is a  $n$ -dimensional vector space over a field  $F$  of characterization zero and  $\Delta: X^n \rightarrow F$  a determinant function in  $X$  if  $y \in X$  and  $(x_1, \dots, x_n) \in X^n$  then  $\hat{\Delta}(y, (x_1, \dots, x_n)) = \Delta(x_1, \dots, x_n) \cdot y$

**Proof.** Then we have to consider the following cases for  $(x_1, \dots, x_n) \in X^n$ :

1. ( $\{x_i\}_{i \in \{1, \dots, n\}}$  are linear dependent) then by 10.259 we have that  $\Delta(x_1, \dots, x_n) = 0$  we must prove that  $\hat{\Delta}(y, (x_1, \dots, x_n)) = 0$ . Consider the following cases for  $n$ :

- ( $n = 1$ ) In this case let  $\{e_1\}$  be the basis of  $X$  then if  $y \in X$  and  $x = (x_1) \in X^1$  there exists a  $\alpha, \beta \in F$  such that  $y = \alpha \cdot e_1$ ,  $x_1 = \beta \cdot e_1$  and thus  $\hat{\Delta}(y, (x_1)) \stackrel{10.269}{=} (-1)^0 \cdot \Delta(y) \cdot x_1 = \Delta(\alpha \cdot e_1) \cdot (\beta \cdot e_1) = \Delta(\beta \cdot e_1) \cdot (\alpha \cdot e_1) = \Delta(x_1) \cdot y$
- ( $n = 2$ ) In this case  $\hat{\Delta}(y, (x_1, x_2)) = (-1)^0 \cdot \Delta(y, x_2) \cdot x_1 + (-1)^1 \cdot \Delta(y, x_1) \cdot x_2 = \Delta(y, x_2) \cdot x_1 - \Delta(y, x_1) \cdot x_2$ . Now as  $\{x_i\}_{i \in \{1, \dots, 2\}}$  is linear dependent there exists by 10.146 a  $\alpha \neq 0$  such that  $x_1 = \alpha \cdot x_2$  and thus  $\hat{\Delta}(y, (x_1, x_2)) = \Delta(y, x_2) \cdot (\alpha \cdot x_2) - \Delta(y, \alpha \cdot x_2) \cdot x_2 = \alpha \cdot (\Delta(y, x_2) \cdot x_2) - \alpha \cdot (\Delta(y, x_2) \cdot x_2) = 0$
- ( $n > 2$ ) In this case as  $\{x_i\}_{i \in \{1, \dots, n\}}$  are linear independent there exists by 10.146 a  $i_0 \in \{1, \dots, n\}$  and  $\{\alpha_i\}_{i \in \{1, \dots, n\} \setminus \{i_0\}}$  such that

$$x_{i_0} = \sum_{i \in \{1, \dots, n\} \setminus \{i_0\}} \alpha_i \cdot x_i \quad (10.84)$$

Now  $\hat{\Delta}(y, (x_1, \dots, x_n)) = \sum_{j \in \{i_0\}} (-1)^j \cdot \Delta(y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n) \cdot x_j + \sum_{j \in \{1, \dots, n\} \setminus \{i_0\}} (-1)^j \cdot \Delta(y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n) \cdot x_j = (-1)^{i_0} \cdot \Delta(y, x_1, \dots, x_{i_0-1}, x_{i_0+1}, \dots, x_n) + \sum_{j \in \{1, \dots, n\} \setminus \{i_0\}} (-1)^j \cdot \Delta(y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n) \cdot x_j$  giving

$$\hat{\Delta}(y, (x_1, \dots, x_n)) = A + B \quad (10.85)$$

where

$$A = (-1)^{i_0} \cdot \Delta(y, x_1, \dots, x_{i_0-1}, x_{i_0+1}, \dots, x_n) \cdot x_{i_0}$$

$$B = \sum_{j \in \{1, \dots, n\} \setminus \{i_0\}} (-1)^j \cdot \Delta(y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n) \cdot x_j$$

Using 10.84 we have then that

$$A = (-1)^{i_0} \cdot \Delta(y, x_1, \dots, x_{i_0-1}, x_{i_0+1}, \dots, x_n) \cdot \sum_{i \in \{1, \dots, n\} \setminus \{i_0\}} \alpha_i \cdot x_i \quad (10.86)$$

Now for  $j \in \{1, \dots, n\} \setminus \{i_0\}$  we have two cases to consider:

- i. ( $i_0 < j$ ) then using 10.84 and the multilinearity of  $\Delta$  we have

$$\Delta(y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n) \cdot x_j = \sum_{k \in \{1, \dots, n\} \setminus \{i_0\}} \alpha_k \cdot \Delta((y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_{i_0+1 \rightarrow x_k}) \cdot x_j \quad (10.87)$$

now for  $k \in \{1, \dots, n\} \setminus \{i_0\}$  consider the following cases

- A. ( $k < j$ ) then  $((y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_{i_0+1 \rightarrow x_k})_{k+1} \stackrel{10.259}{=} (y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_{k+1} \stackrel{10.259}{=} x_k$  and  $((y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_{i_0+1 \rightarrow x_k})_{i_0+1} = x_k$  this together with  $k \neq i_0 \Rightarrow k+1 \neq i_0+1$  means by 10.259 that

$$\Delta((y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_{i_0+1 \rightarrow x_k}) = 0$$

- B. ( $j < k$ ) then  $((y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_{i_0+1 \rightarrow x_k})_k \stackrel{10.259}{=} (y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_{j < k} = x_k$  and  $((y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_{i_0+1 \rightarrow x_k})_{i_0+1} = x_k$  this together with  $i_0 < j < k \Rightarrow i_0+1 \neq k$  and 10.259 that

$$\Delta((y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_{i_0+1 \rightarrow x_k}) = 0$$

Using (1) and (2) we conclude that  $\Delta((y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_{i_0+1 \rightarrow x_k}) = \delta_{j,k} \cdot \Delta((y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_{i_0+1 \rightarrow x_k})$  and thus using 10.87 we have  $\Delta(y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n) \cdot x_j = \sum_{k \in \{1, \dots, n\} \setminus \{i_0\}} \delta_{j,k} \cdot (\alpha_k \cdot (\Delta((y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_{i_0+1 \rightarrow x_k}) \cdot x_j)) = \alpha_j \cdot \Delta((y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_{i_0+1 \rightarrow x_j}) \cdot x_j$  giving :

$$\Delta(y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n) \cdot x_j = \alpha_j \cdot (\Delta((y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_{i_0+1 \rightarrow x_j}) \cdot x_j) \quad (10.88)$$

Now  $(i_0 + 1 \rightsquigarrow j) \Delta((y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_{i_0+1 \rightarrow x_j}) = \Delta((y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_{i_0+1 \rightarrow x_j})_{i_0 < j \Rightarrow i_0+1 \leq j}$  and 10.270 using 10.98 and the fact that  $\Delta$  is skew symmetric and  $i_0 + 1 < j$  we have then that  $(-1)^{j-i_0-1} \Delta((y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_{i_0+1 \rightarrow x_i}) = (i_0 + 1 \rightsquigarrow j) \Delta((y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_{i_0+1 \rightarrow x_j}) = \Delta((y, x_1, \dots, x_{i_0-1}, x_{i_0+1}, \dots, x_n))$  and by multiplying both sides by  $(-1)^{j-i_0-1}$  and the fact that  $1 = (-1)^{j-i_0-1} \cdot (-1)^{j-i_0-1}$  gives using 10.88 that

$$\Delta(y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n) \cdot x_j = (-1)^{j-i_0-1} \cdot (\alpha_j \cdot (\Delta(y, x_1, \dots, x_{i_0-1}, x_{i_0+1}, \dots, x_n) \cdot x_j)) \quad (10.89)$$

ii.  $(j < i_0)$  Then using 10.84 and the multilinearity of  $\Delta$  we have that

$$\Delta(y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n) \cdot x_j = \sum_{k \in \{1, \dots, n\} \setminus \{i_0\}} \alpha_k \cdot \Delta((y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_{i_0 \rightarrow x_k}) \cdot x_j \quad (10.90)$$

now for  $k \in \{1, \dots, n\} \setminus \{i_0\}$  consider the following cases:

A.  $(k < j)$  then  $((y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_{i_0 \rightarrow x_k})_{k+1 \leq j < i_0 \Rightarrow k+1 \leq k < i_0} = (y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_{k+1} = x_k$  and  $((y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_{i_0 \rightarrow x_k})_{i_0} = x_k$ , this together with  $k+1 \leq j < i_0$  and 10.259 gives

$$\Delta((y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_{i_0 \rightarrow k}) = 0$$

B.  $(j < k)$  then  $((y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_{i_0 \rightarrow x_k})_{k \in \{1, \dots, n\} \setminus \{i_0\} \Rightarrow k \neq i_0} = (y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_k = x_k$  and  $((y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_{i_0 \rightarrow x_k})_{i_0} = x_k$ , this together with  $k \neq i_0$  and 10.259 gives

$$\Delta((y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_{i_0 \rightarrow x_k}) = 0$$

Using (1) and (2) we find that  $\Delta((y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_{i_0 \rightarrow x_k}) = \delta_{i,j} \cdot \Delta((y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_{i_0 \rightarrow x_k})$  and thus using 10.90  $\Delta(y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n) \cdot x_j = \sum_{k \in \{1, \dots, n\} \setminus \{i_0\}} \delta_{j,k} \cdot (\alpha_k \cdot (\Delta((y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_{i_0 \rightarrow x_k}) \cdot x_j)) = \alpha_j \cdot \Delta((y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_{i_0 \rightarrow x_j}) \cdot x_j$  giving that

$$\Delta(y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n) \cdot x_j = \alpha_j \cdot \Delta((y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_{i_0 \rightarrow x_j}) \cdot x_j \quad (10.91)$$

Now  $(i_0 \rightsquigarrow j+1) \Delta((y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_{i_0 \rightarrow x_j}) = \Delta((y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_{i_0 \rightarrow x_j} \circ (i_0 \rightsquigarrow j+1))_{j < i_0 \Rightarrow j+1 \leq i_0 \text{ and 10.270}} = \Delta(y, x_1, \dots, x_{i_0-1}, x_{i_0+1}, \dots, x_n)$  so using the skew symmetry of  $\Delta$ , 10.98 and  $j < i_0 \Rightarrow j \leq i_0 + 1$  we have that  $(-1)^{i_0+1-j} \cdot \Delta((y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_{i_0 \rightarrow x_j}) = \Delta(y, x_1, \dots, x_{i_0-1}, x_{i_0+1}, \dots, x_n)$  and dividing both sides by  $(-1)^{i_0+1-j}$  we get  $\Delta((y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_{i_0 \rightarrow x_j}) = (-1)^{j-i_0-1} \Delta(y, x_1, \dots, x_{i_0-1}, x_{i_0+1}, \dots, x_n)$  and thus

$$\Delta(y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n) \cdot x_j = (-1)^{j-i_0-1} \cdot (\alpha_j \cdot (\Delta(y, x_1, \dots, x_{i_0-1}, x_{i_0+1}, \dots, x_n) \cdot x_j)) \quad (10.92)$$

Using 10.85, 10.89 and 10.92 we have  $B = \sum_{j \in \{1, \dots, n\} \setminus \{i_0\}} (-1)^j \cdot \Delta(y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n) \cdot x_j = \sum_{j \in \{1, \dots, i_0-1\}} (-1)^j \cdot \Delta(y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n) \cdot x_j + \sum_{j \in \{i_0+1, \dots, n\}} (-1)^j \cdot \Delta(y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n) \cdot x_j = \sum_{j \in \{1, \dots, i_0-1\}} (-1)^j((-1)^{j-i_0-1} \cdot (\alpha_j \cdot (\Delta(y, x_1, \dots, x_{i_0-1}, x_{i_0+1}, \dots, x_n) \cdot x_j))) + \sum_{j \in \{i_0+1, \dots, n\}} (-1)^j((-1)^{j-i_0-1} \cdot (\alpha_j \cdot (\Delta(y, x_1, \dots, x_{i_0-1}, x_{i_0+1}, \dots, x_n) \cdot x_j))) = \sum_{j \in \{1, \dots, n\} \setminus \{i_0\}} (-1)^j((-1)^{j-i_0-1} \cdot (\alpha_j \cdot (\Delta(y, x_1, \dots, x_{i_0-1}, x_{i_0+1}, \dots, x_n) \cdot x_j))) = \Delta(y, x_1, \dots, x_{i_0-1}, x_{i_0+1}, \dots, x_n) \cdot \sum_{j \in \{1, \dots, n\} \setminus \{i_0\}} (-1)^{2j-i_0-1} \cdot (\alpha_j \cdot x_j) \stackrel{(-1)^{2j-i_0-1} = (-1)^{2j} \cdot (-1)^{-i_0} \cdot (-1)^{-1} = -(-1)^{i_0}}{=} -(-1)^{i_0} \cdot \Delta(y, x_1, \dots, x_{i_0-1}, x_{i_0+1}, \dots, x_n) \cdot \sum_{j \in \{1, \dots, n\} \setminus \{i_0\}} \alpha_j \cdot x_j \stackrel{10.86}{=} -A$  so we have finally that  $A + B = A + (-A) = 0$  and thus by 10.85 that  $\hat{\Delta}(y, (x_1, \dots, x_n)) = 0$  as should be

2. ( $\{x_i\}_{i \in \{1, \dots, n\}}$  are linear independent) and thus as  $\dim(X) = n$  we have by 10.159 and 10.201 that  $\{x_i\}_{i \in \{1, \dots, n\}}$  forms a basis of  $X$ . So if  $y \in X$  there exists a unique  $\{\alpha_i\}_{i \in \{1, \dots, n\}}$  such that

$$y = \sum_{i \in \{1, \dots, n\}} \alpha_i \cdot x_i \quad (10.93)$$

Then we have

$$\hat{\Delta}(y, (x_1, \dots, x_n)) = \sum_{j \in \{1, \dots, n\}} (-1)^{j-1} \cdot \Delta(y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n) \cdot x_j \quad (10.94)$$

and by multilinearity of  $\Delta$  we have then that

$$\Delta(y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n) \cdot x_j = \sum_{i \in \{1, \dots, n\}} \alpha_i \cdot \Delta(x_i, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n) \cdot x_j \quad (10.95)$$

Now for  $\Delta(x_i, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)$  we have if :

- $(i < j)$  then  $(x_i, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_1 = x_i = (x_i, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_{i+1}$  which as  $1 \neq i+1$  would mean by 10.259 that  $\Delta(x_i, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n) = 0$ .
- $(j < i)$  then  $(x_i, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_1 = x_i = (x_i, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_i$  which as  $1 \leq j < i$  means by 10.259 that  $\Delta(x_i, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n) = 0$

So taking (a) and (b) in account we have  $\Delta(x_i, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n) = \delta_{i,j} \cdot \Delta(x_i, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)$  and thus by 10.95 we have that  $\Delta(y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n) \cdot x_j = \sum_{i \in \{1, \dots, n\}} \delta_{i,j} (\alpha_i \cdot \Delta(x_i, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n) \cdot x_j) = \alpha_j \cdot (\Delta(x_j, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n) \cdot x_j)$  giving:

$$\Delta(y, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n) \cdot x_j = \alpha_j \cdot (\Delta(x_j, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n) \cdot x_j) \quad (10.96)$$

Now if  $A = (x_j, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n) \circ (1 \rightsquigarrow_n j)$  we have the following cases for  $j \in \{1, \dots, n\}$

- $(j = 1)$  then  $(1 \rightsquigarrow_n j) = 1_{\{1, \dots, n\}}$  so  $A = (x_j, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)$  and  $\forall k \in \{1, \dots, n\}$  we have  $(x_j, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_k = \begin{cases} x_j = x_1 \text{ if } k = 1 \\ x_{k-1} \text{ if } k < j = 1 \text{ which is impossible} \\ x_k \text{ if } k \in \{2, \dots, n\} \end{cases} = x_k$  so  $A = (x_1, \dots, x_n)$ .
- $(j \in \{2, \dots, n\})$  then  $1 < j$  and by definition of  $(1 \rightsquigarrow_n j)$  for  $k \in \{1, \dots, n\}$  we have the following cases to consider for  $A_k = ((x_j, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n) \circ (1 \rightsquigarrow_n j))_k$ :

- $(k < 1)$  this can not happen
- $(1 \leq k < j)$  then  $A_k = (x_j, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_{k+1} \stackrel{2 \leq k+1 \leq j}{=} x_{(k+1)-1} = x_k$
- $(k = j)$  then  $A_k = (x_j, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_1 = x_j \stackrel{k=j}{=} x_k$
- $(j < k)$  then  $A_k = (x_j, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_k \stackrel{j < k}{=} x_k$

so in all cases we have  $A_k = x_k$  so  $A = (x_1, \dots, x_n)$

As in (a) en (b) we have  $A = (x_1, \dots, x_n)$  so we have  $(x_j, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n) \circ (1 \rightsquigarrow_n j) = (x_1, \dots, x_n)$  and thus  $(1 \rightsquigarrow_n j) \Delta(x_j, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n) = \Delta((x_j, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n) \circ (1 \rightsquigarrow_n j)) = \Delta(x_1, \dots, x_n)$ . Using the fact that  $\Delta$  is skew symmetric and 10.98 we have then that  $(-1)^{j-1} \cdot \Delta(x_j, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n) = \Delta(x_1, \dots, x_n)$ . Using the last equation, 10.94 and 10.96 we have then  $\hat{\Delta}(y, (x_1, \dots, x_n)) = \sum_{j \in \{1, \dots, n\}} \alpha_j \cdot (\Delta(x_1, \dots, x_n) \cdot x_j) = \Delta(x_1, \dots, x_n) \cdot \sum_{j \in \{1, \dots, n\}} \alpha_j \cdot x_j \stackrel{10.93}{=} \Delta(x_1, \dots, x_n) \cdot y$ .

So in all case (1) and (2) we have  $\hat{\Delta}(y, (x_1, \dots, x_n)) = \Delta(x_1, \dots, x_n) \cdot y$   $\square$

**Definition 10.273.** Let  $n \in \mathbb{N}_0$ ,  $X$  be a  $n$ -dimensional vector space over a field  $F$ ,  $\Delta$  a non trivial determinant function  $\Delta: X^n \rightarrow F$  and  $L: X \rightarrow X$  a linear mapping ( $L \in \text{Hom}(X, X)$ ) (we call this sometimes a linear transformation). Then  $\Delta_L: X^n \rightarrow F$  is defined by  $(x_1, \dots, x_n) \rightarrow \Delta_L(x_1, \dots, x_n) = \Delta(L(x_1), \dots, L(x_n))$

**Theorem 10.274.** Let  $n \in \mathbb{N}_0$ ,  $X$  be a  $n$ -dimensional vector spaces over a field  $F$  of characterization zero,  $\Delta$  a non trivial determinant function  $\Delta: X^n \rightarrow F$  in  $X$  and  $L \in \text{Hom}(X, X)$  a linear transformation then  $\Delta_L$  is a determinant function.

**Proof.**

1. **(Multilinearity)**

$$\begin{aligned}
 \Delta_L(x_1, \dots, x_{i-1}, \alpha \cdot x + \beta \cdot y, x_{i+1}, \dots, x_n) &= \Delta(L(x_1), \dots, L(x_{i-1}), L(\alpha \cdot x + \beta \cdot y), \\
 &\quad L(x_{i+1}), \dots, L(x_n)) \\
 &\stackrel{L \text{ is linear}}{=} \Delta(L(x_1), \dots, L(x_{i-1}), \alpha \cdot L(x) + \beta \cdot L(y), \\
 &\quad L(x_{i+1}), \dots, L(x_n)) \\
 &\stackrel{\Delta \text{ is multilinear}}{=} \alpha \cdot \Delta(L(x_1), \dots, L(x_{i-1}), L(x), \\
 &\quad L(x_{i+1}), \dots, L(x_n)) + \beta \cdot \Delta(L(x_1), \dots, \\
 &\quad L(x_{i-1}), L(y), L(x_{i+1}), \dots, L(x_n)) \\
 &= \alpha \cdot \Delta_L(x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_n) + \beta \cdot \\
 &\quad \Delta_L(x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_n)
 \end{aligned}$$

2. **(Skew Symmetry)** Assume that we have  $(x_1, \dots, x_n) \in X^n$  such that  $\exists i, j \in \{1, \dots, n\}$  with  $i \neq j$  and  $x_i = x_j$  then  $L(x_i) = L(x_j)$  then  $\Delta_L(x_1, \dots, x_n) = \Delta(L(x_1), \dots, L(x_n)) \stackrel{\Delta \text{ is skew symmetric and 10.259}}{=} 0$  so by 10.259 we have that  $\Delta_L$  is skew symmetric.  $\square$

**Theorem 10.275.** Let  $n \in \mathbb{N}_0$ ,  $X$  be a  $n$ -dimensional vector spaces over a field  $F$  of characterization zero,  $\Delta$  a non trivial determinant function  $\Delta: X^n \rightarrow F$  in  $X$  and  $L \in \text{Hom}(X, X)$  a linear transformation then there exists a unique scalar  $\alpha \in F$  such that  $\Delta_L = \alpha \cdot \Delta$  (unique meaning that  $\alpha$  is only determined by  $L$  not by  $\Delta$ ) or if  $\Delta': X^n \rightarrow F$  is another non trivial determinant function then  $\Delta'_L = \alpha \cdot \Delta'$ .

**Proof.**

- (existence)** Using 10.267 and the fact that by the above theorem  $\Delta_L$  is determinant function we have the existence of a  $\alpha \in F$  such that  $\Delta_L = \alpha \cdot \Delta$ .
- (uniqueness)** If  $\Delta'$  is another non trivial determinant function then by 10.267 there exists a  $\alpha'$  such that  $\Delta' = \lambda \cdot \Delta$ . So if  $(x_1, \dots, x_n) \in X^n$  then  $\Delta'_L(x_1, \dots, x_n) = \Delta'(L(x_1), \dots, L(x_n)) = \lambda \cdot \Delta(L(x_1), \dots, L(x_n)) = \lambda \cdot \Delta_L(x_1, \dots, x_n) \Rightarrow \Delta'_L = \lambda \cdot \Delta_L = \lambda \cdot (\alpha \cdot \Delta) = \alpha \cdot (\lambda \cdot \Delta) = \alpha \cdot \Delta' \Rightarrow \Delta'_L = \alpha \cdot \Delta'$  if now also  $\Delta'_L = \alpha' \cdot \Delta'$  then by 10.266 we must have  $\alpha = \alpha'$ .  $\square$

Now by 10.265 there always exists a non trivial determinant function so that the following definition make sense (together with the above theorem)

**Definition 10.276.** Let  $n \in \mathbb{N}_0$ ,  $X$  be a  $n$ -dimensional vector spaces over a field  $F$  of characterization zero,  $L: X \rightarrow X$  be a linear transformation [ $L \in \text{Hom}(X, X)$ ]. Then  $\det(L)$  is defined so that  $\Delta_L = \det(L) \cdot \Delta$  where  $\Delta$  is any non trivial determinant function (by 10.265 such a determinant function exists and by the previous theorem  $\det(X)$  is independent of the determinant function). So  $\det: \text{Hom}(X, X) \rightarrow F$  defined by  $L \rightarrow \det(L)$  is a well defined function.

**Example 10.277.** Let  $X$  be a  $n$ -dimensional vector space over a field  $F$  of characterization zero then if  $\lambda \in F$  we have that  $\lambda \cdot 1_X: X \rightarrow X$  is trivially a linear transformation. Now if  $\Delta$  is a trivial determinant function we have that  $\Delta_{\lambda \cdot 1_X}(x_1, \dots, x_n) = \Delta(\lambda \cdot 1_X(x_1), \dots, \lambda \cdot 1_X(x_n)) = \Delta(\lambda \cdot x_1, \dots, \lambda \cdot x_n) = \lambda^n \cdot \Delta(x_1, \dots, x_n)$  so we have that  $\det(\lambda \cdot 1_X) = \lambda^n$ . Or if we take  $\lambda = 1$  (the unit in  $F$ ) that  $\det(1_X) = 1$ .

**Example 10.278.** If  $X$  is a 1 dimensional space then if  $\{e_1\}$  is the basis of  $X$  then if  $x_1 = \alpha \cdot e_1 \in X$  we have for  $L \in \text{Hom}(X, X)$  that  $L(x_1) = \alpha \cdot L(e_1)$ , as  $L(e_1) \in X$  there exists a  $\beta$  such that  $L(e_1) = \beta \cdot e_1$ . So we have  $\Delta_L(x_1) = \Delta(L(x_1)) = \Delta(\alpha \cdot L(e_1)) = \Delta(\alpha \cdot \beta \cdot e_1) = \beta \cdot \Delta(\alpha \cdot e_1) = \beta \cdot \Delta(x)$  or we have that  $\det(L) = \beta$  (where  $L(e_1) = \beta \cdot e_1$ ).

**Example 10.279.** If  $X$  is a  $n$ -dimensional space with basis  $\{e_i\}_{i \in \{1, \dots, n\}}$  over a field  $F$  of characterization zero,  $\alpha \in F$ ,  $i \in \{1, \dots, n\}$  and  $L: X \rightarrow X$  defined by  $L(e_k) = \begin{cases} e_k & \text{if } k \neq i \\ \alpha \cdot e_k & \text{if } k = i \end{cases}$  then  $\det(L) = \alpha$

**Proof.**  $\Delta_L(e_1, \dots, e_n) = \Delta(L(e_1), \dots, L(e_{i-1}), L(e_i), L(e_{i+1}), \dots, L(e_n)) = \Delta(e_1, \dots, e_{i-1}, \alpha \cdot e_i, e_{i+1}, \dots, e_n) = \alpha \cdot \Delta(e_1, \dots, e_n)$  proving that  $\det(L) = \alpha$   $\square$

**Example 10.280.** If  $X$  is a  $n$ -dimensional space with basis  $\{e_i\}_{i \in \{1, \dots, n\}}$  over a field  $F$  of characterization zero,  $i, j \in \{1, \dots, n\}$  with  $i \neq j$  and  $L: X \rightarrow X$  defined by  $L(e_k) = \begin{cases} e_k & \text{if } k \neq i, j \\ e_i & \text{if } k = j \\ e_j & \text{if } k = i \end{cases} \Rightarrow L(e_k) = e_{(i \leftrightarrow_n j)(k)}$  then  $\det(L) = -1$

**Proof.**  $\Delta_L(e_1, \dots, e_n) = \Delta(L(e_1), \dots, L(e_n)) = \Delta(e_{(i \leftrightarrow_n j)}, \dots, e_{(i \leftrightarrow_n j)(n)}) = \text{sign}((i \leftrightarrow_n j)) \cdot \Delta(e_1, \dots, e_n) = (-1) \cdot \Delta(e_1, \dots, e_n)$  proving that  $\det(L) = -1$   $\square$

**Example 10.281.** If  $X$  is a  $n$ -dimensional space with basis  $\{e_i\}_{i \in \{1, \dots, n\}}$  over a field of characterization zero, let  $i, j \in \{1, \dots, n\} \vdash i \neq j$  and  $\alpha \neq 0$  define then  $L$  by  $L(e_k) = \begin{cases} e_k & \text{if } k \neq i \\ e_i + \alpha \cdot e_j & \text{if } k = i \end{cases}$  then  $\det(L) = 1$

**Proof.**

$$\begin{aligned} \Delta_L(e_1, \dots, e_n) &= \Delta(L(e_1), \dots, L(e_{i-1}), L(e_i), L(e_{i+1}), \dots, L(e_n)) \\ &= \Delta(e_1, \dots, e_{i-1}, e_i + \alpha \cdot e_j, e_{i+1}, \dots, e_n) \\ &= \Delta(e_1, \dots, e_{i-1}, e_i, e_{i+1}, \dots, e_n) + \alpha \cdot \Delta(e_1, \dots, e_{i-1}, e_j, e_{i+1}, \dots, e_n) \\ &\stackrel{10.259}{=} \Delta(e_1, \dots, e_n) + \alpha \cdot 0 \\ &= \Delta(e_1, \dots, e_n) \end{aligned}$$

proving that  $\det(L) = 1$   $\square$

**Theorem 10.282.** Let  $n \in \mathbb{N}_0$ ,  $X$  be a  $n$ -dimensional vector spaces over a field  $F$  of characterization zero then we have

1.  $\det(1_X) = 1$  (see for the proof the previous example)
2. A linear transformation  $L: X \rightarrow X$  is regular (meaning injective and thus a isomorphism by 10.214) if and only if  $\det(L) \neq 0$
3. If  $L_1, L_2: X \rightarrow X$  are linear transformations then  $\det(L_1 \circ L_2) = \det(L_1) \cdot \det(L_2)$ . From this if follows that if  $L: X \rightarrow X$  is a bijective linear transformation (a linear isomorphism) then  $L \circ L^{-1} = 1_X \Rightarrow \det(L) \cdot \det(L^{-1}) = \det(1_X) = 1 \Rightarrow \det(L^{-1}) = \det(L)^{-1}$

**Proof.**

1. This was proved in the previous example.
2. First we prove that regularity implies that  $\det(L) \neq 0$ . If  $\{e_i\}_{i \in \{1, \dots, n\}}$  is a basis of  $X$  then as  $L$  is a regular linear transformation we can use 10.214 and 10.209 to prove that  $\{L(e_i)\}_{i \in \{1, \dots, n\}}$  is a basis of  $X$ . Then using the determinant function defined in 10.265 we have that

$$\begin{aligned} 1 &= \Delta(L(e_1), \dots, L(e_n)) \\ &= \Delta_L(e_1, \dots, e_n) \\ &= \det(L) \cdot \Delta(e_1, \dots, e_n) \end{aligned}$$

If now  $\det(L) = 0$  then we have  $1 = 0$  which is a contradiction as  $F$  is field of characterization zero, so

$$\det(L) \neq 0$$

Next we prove that if  $\det(L) \neq 0$  then  $L$  is regular. Using 10.265 find a  $\Delta'$  such that

$$\Delta'(e_1, \dots, e_n) = 1$$

then

$$\begin{aligned} \Delta'(L(e_1), \dots, L(e_n)) &= \det(L) \cdot \Delta'(e_1, \dots, e_n) \\ &= \det(L) \\ &\neq 0 \end{aligned}$$

So we must have that  $\{L(e_i)\}_{i \in \{1, \dots, n\}}$  is linear independent [if not we have by 10.259 that  $\Delta'(L(e_1), \dots, L(e_n)) = 0$ ]. Using 10.214 we have then that  $L$  is regular.

3. If  $(x_1, \dots, x_n) \in X^n$  and  $\Delta$  any non trivial determinant function (which must exists by 10.265) then we have

$$\begin{aligned} \Delta_{L_1 \circ L_2}(x_1, \dots, x_n) &= \Delta((L_1 \circ L_2)(x_1), \dots, (L_1 \circ L_2)(x_n)) \\ &= \Delta(L_1(L_2(x_1)), \dots, L_1(L_2(x_n))) \\ &= \Delta_{L_1}(\Delta(L_2(x_1)), \dots, L_2(x_n)) \\ &= \det(L_1) \cdot \Delta(L_2(x_1), \dots, L_2(x_n)) \\ &= \det(L_1) \cdot \Delta_{L_2}(x_1, \dots, x_n) \\ &= (\det(L_1) \cdot \det(L_2)) \cdot \Delta(x_1, \dots, x_n) \end{aligned}$$

and thus by definition of  $\det(L_1 \circ L_2)$  we have that  $\det(L_1 \circ L_2) = \det(L_1) \cdot \det(L_2)$ .  $\square$

**Definition 10.283.** Let  $n \in \mathbb{N}_0$ ,  $X$  be a  $n$ -dimensional vector space over a field  $F$  of characterization zero,  $\Delta$  a non trivial determinant function in  $X$ ,  $L \in \text{Hom}(X, X)$  then we define  $\check{\Delta}_L: X^n \rightarrow \text{Hom}(X, X)$  by  $(x_1, \dots, x_n) \rightarrow \check{\Delta}_L(x_1, \dots, x_n)$  where  $\check{\Delta}_L(x_1, \dots, x_n): X \rightarrow X$  is defined by  $x \rightarrow \check{\Delta}_L(x_1, \dots, x_n)(x) = \sum_{i \in \{1, \dots, n\}} (-1)^{i-1} \cdot \Delta(x, L(x_1), \dots, L(x_{i-1}), L(x_{i+1}), \dots, L(x_n)) \cdot x_i$

**Proof.** Of course we must prove that  $\check{\Delta}_L(x_1, \dots, x_n)$  is a linear mapping. So let  $\alpha, \beta \in F$ ,  $x, y \in X$  and  $(x_1, \dots, x_n) \in X^n$  then we have

$$\begin{aligned} \check{\Delta}_L(x_1, \dots, x_n)(\alpha \cdot x + \beta \cdot y) &= \sum_{i \in \{1, \dots, n\}} (-1)^{i-1} \cdot \Delta(\alpha \cdot x + \beta \cdot y, L(x_1), \dots, L(x_{i-1}), \\ &\quad L(x_{i+1}), \dots, L(x_n)) \cdot x_i \\ &\stackrel{L \text{ is multilinear}}{=} \sum_{i \in \{1, \dots, n\}} (-1)^{i-1} \cdot (\alpha \cdot \Delta(x, L(x_1), \dots, L(x_{i-1}), L(x_{i+1}), \dots, \\ &\quad L(x_n)) + \beta \cdot \Delta(y, L(x_1), \dots, L(x_{i-1}), L(x_{i+1}), \dots, L(x_n))) \cdot x_i \\ &= \alpha \cdot \sum_{i \in \{1, \dots, n\}} (-1)^{i-1} \cdot \Delta(x, L(x_1), \dots, L(x_{i-1}), L(x_{i+1}), \dots, \\ &\quad L(x_n)) \cdot x_i + \beta \cdot \sum_{i \in \{1, \dots, n\}} (-1)^{i-1} \cdot \Delta(y, L(x_1), \dots, L(x_n)) \cdot x_i \\ &= \alpha \cdot \check{\Delta}_L(x_1, \dots, x_n)(x) + \beta \cdot \check{\Delta}_L(x_1, \dots, x_n)(y) \end{aligned}$$

$\square$

**Example 10.284.**

1. If  $n = 1$  then  $\check{\Delta}_L(x_1)(x) = (-1)^{1-1} \cdot \Delta(x) \cdot x_1 = \Delta(x) \cdot x_1$

2. If  $n = 2$  then  $\check{\Delta}_L(x_1, x_2) = (-1)^{1-1} \cdot \Delta(x, L(x_2)) \cdot x_1 + (-1)^{2-1} \cdot \Delta(x, L(x_1)) \cdot x_2 = \Delta(x, L(x_2)) \cdot x_1 - \Delta(x, L(x_1)) \cdot x_2$

**Lemma 10.285.** Let  $n \in \mathbb{N}_0$ ,  $X$  as set and  $x = (x_1, \dots, x_n) \in X^n$ ,  $i, j \in \{1, \dots, n\}$  with  $i \neq j$  and  $x_i = x_j$  and  $t \in X$  then we have the following

1. ( $j < i$ ) then  $(t, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n) \circ (i \rightsquigarrow_n j + 1) = (t, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$
2. ( $i < j$ ) then  $(t, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n) \circ (i + 1 \rightsquigarrow_n j) = (t, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$

**Proof.**

1. ( $j < i$ ) then  $j + 1 \leq i$  let then  $a = (t, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n) \circ (i \rightsquigarrow_n j + 1)$  then we have the following to consider:

- a. ( $j + 1 = i$ ) then  $a = (t, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)$  and we have for  $k \in \{1, \dots, n\}$  that
  - i. ( $k = 1$ ) then  $a_k = t = (t, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_k$
  - ii. ( $2 \leq k \leq j$ ) then  $a_k = x_{k-1} \underset{k \leq j < i}{\equiv} (t, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_k$
  - iii. ( $k = j + 1$ ) then  $a_k = x_k = x_i = x_j = x_{k-1} \underset{j+1=i}{\equiv} (t, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_k$
  - iv. ( $j + 1 < k$ ) then  $a_k = x_k \underset{i=j+1 < k}{\equiv} (t, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_k$

this proves that  $(t, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n) \circ (i \rightsquigarrow_n j + 1) = (t, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$ .

- b. ( $j + 1 < i$ ) then we have by the definition of  $(i \rightsquigarrow_n j + 1)$  the following cases to consider for  $k \in \{1, \dots, n\}$  in  $a = (t, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n) \circ (i \rightsquigarrow_n j + 1)$ :

- i. ( $k = 1$ ) then as  $k = 1 < j + 1$  we have

$$\begin{aligned} a_k &= (t, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_k \\ &\underset{k=1}{\equiv} t \\ &\underset{k=1}{\equiv} (t, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_k \end{aligned}$$

- ii. ( $k \neq 1 \wedge k < j + 1$ ) then

$$\begin{aligned} a_k &= (t, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_k \\ &\underset{k < j+1 < i}{\equiv} (t, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_k \end{aligned}$$

- iii. ( $k \neq 1 \wedge k = j + 1$ ) then

$$\begin{aligned} a_k &= (t, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_i \\ &\underset{j+1 < i}{\equiv} x_i \\ &\underset{x_i=x_j}{\equiv} x_j \\ &\underset{j+1 < i}{\equiv} (t, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_{j+1} \\ &\underset{k=j+1}{\equiv} (t, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_k \end{aligned}$$

- iv. ( $k \neq 1 \wedge j + 1 < k \leq i$ ) then

$$\begin{aligned} a_k &= (t, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_{k-1} \\ &\underset{k \leq i}{\equiv} (t, x+1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_k \end{aligned}$$

- v. ( $k \neq 1 \wedge i < k$ ) then

$$\begin{aligned} a_k &= (t, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_k \\ &\underset{j < j+1 < i < k}{\equiv} x_k \\ &\underset{i < k}{\equiv} (t, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_k \end{aligned}$$

this proves that  $(t, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n) \circ (i \rightsquigarrow_n j + 1) = (t, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$ .

2. ( $i < j$ ) then  $i+1 \leq j$  and we have by the definition of  $(i+1 \rightsquigarrow_n j)$  the following cases to consider for  $a = (t, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n) \circ (i+1 \rightsquigarrow_n j)$

a. ( $i+1 = j$ ) then we have that  $a = (t, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)$  and we have for  $k \in \{1, \dots, n\}$  that :

i. ( $k = 1$ ) then  $a_k = t = (t, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_k$

ii. ( $k \leq i$ ) then as  $k \leq i < j$  we have  $a_k = x_{k-1} \stackrel{k \leq i}{=} (t, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_k$

iii. ( $k = i+1 = j$ ) then as  $k = j \leq j$  we have  $a_k = x_{k-1} \stackrel{k=i+1 \Rightarrow k-1=i}{=} x_i \stackrel{x_i=x_j}{=} x_j = x_k \stackrel{i < i+1=k}{=} (t, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_k$

iv. ( $j < k$ ) then we have  $a_k = x_k \stackrel{i < j < k}{=} (t, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_k$

this proves that  $(t, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n) \circ (i+1 \rightsquigarrow_n j) = (t, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$

b. ( $i+1 < j$ ) then we have by the definition of  $(i+1 \rightsquigarrow_n j)$  the following cases to consider for  $k \in \{1, \dots, n\}$  in  $a = (t, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n) \circ (i \rightsquigarrow_n j+1)$ :

i. ( $k = 1$ ) then

$$\begin{aligned} a_k &\stackrel{1 < i+1}{=} (t, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_k \\ &\stackrel{k=1}{=} t \\ &\stackrel{k=1}{=} (t, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_k \end{aligned}$$

ii. ( $k \neq 1 \wedge k < i+1$ ) then

$$\begin{aligned} a_k &= (t, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_k \\ &\stackrel{k < i+1 < j}{=} x_{k-1} \\ &\stackrel{k < i+1 \Rightarrow k \leq i}{=} (t, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_k \end{aligned}$$

iii. ( $k \neq 1 \wedge i+1 \leq k < j$ ) then

$$\begin{aligned} a_k &= (t, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_{k+1} \\ &\stackrel{k < j \Rightarrow k+1 \leq j}{=} x_k \\ &\stackrel{i < i+1 \leq k}{=} (t, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_k \end{aligned}$$

iv. ( $k \neq 1 \wedge k = j$ ) then

$$\begin{aligned} a_k &= (t, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_{i+1} \\ &\stackrel{i+1 < j}{=} x_i \\ &\stackrel{x_i=x_j}{=} x_j \\ &\stackrel{k=j}{=} x_k \\ &\stackrel{i+1 < j = k \Rightarrow i < k}{=} (t, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_k \end{aligned}$$

v. ( $k \neq 1 \wedge j < k$ ) then

$$\begin{aligned} a_k &= (t, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)_k \\ &\stackrel{j < k}{=} x_k \\ &\stackrel{i < i+1 < j < k}{=} (t, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)_k \end{aligned}$$

this proves that  $(t, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n) \circ (i+1 \rightsquigarrow_n j) = (t, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$

□

**Theorem 10.286.** Let  $n \in \mathbb{N}_0$ ,  $X$  be a  $n$ -dimensional vector space over a field  $F$  of characterization zero,  $\Delta$  a non trivial determinant function in  $X$ ,  $L \in \text{Hom}(X, X)$  then  $\check{\Delta}_L: X^n \rightarrow \text{Hom}(X, X)$  is  $n$ -linear and skew symmetric.

**Proof.** Let  $i \in \{1, \dots, n\}$ ,  $\alpha, \beta \in F$  and  $x, y \in X$ ,  $(x_1, \dots, x_n) \in X^n$  such that  $x_i = \alpha \cdot x + \beta \cdot y$  then we have for  $t \in X$ , define then  $\forall j \in \{1, \dots, n\}$

$$A_j(x_1, \dots, x_n) = \Delta(t, L(x_1), \dots, L(x_{j-1}), L(x_{j+1}), \dots, L(x_n)) \cdot x_j \quad (10.97)$$

then we have by definition that

$$\check{\Delta}_L(x_1, \dots, x_n)(t) = \sum_{j \in \{1, \dots, n\}} (-1)^{j-1} \cdot A_j(x_1, \dots, x_n) \quad (10.98)$$

We have now the following cases for  $j$  to consider

1. ( $j = i$ ) then

$$\begin{aligned} A_j(x_1, \dots, x_n) &= \Delta(t, L(x_1), \dots, L(x_{i-1}), L(x_{i+1}), \dots, L(x_n)) \cdot (\alpha \cdot x + \beta \cdot y) \\ &= \alpha \cdot \Delta(t, L(x_1), \dots, L(x_{i-1}), L(x_{i+1}), \dots, L(x_n)) \cdot x + \beta \cdot \Delta(t, L(x_1), \dots, L(x_{i-1}), \\ &\quad L(x_{i+1}), \dots, L(x_n)) \cdot x \\ &= \alpha \cdot A_j((x_1, \dots, x_n)_{i \rightarrow x}) + \beta \cdot A_j((x_1, \dots, x_n)_{i \rightarrow y}) \end{aligned}$$

2. ( $j < i$ ) then  $(t, L(x_1), \dots, L(x_{j-1}), L(x_{j+1}), \dots, L(x_n))_i = L(x_i) = L(\alpha \cdot x + \beta \cdot y) = \alpha \cdot L(x) + \beta \cdot L(y)$  and thus by multi linearity of  $\Delta$  we have

$$\begin{aligned} A_j(x_1, \dots, x_n) &= \Delta(t, L(x_1), \dots, L(x_{j-1}), L(x_j), \dots, L(x_n)) \\ &= \alpha \cdot \Delta((t, L(x_1), \dots, L(x_{j-1}), L(x_j), \dots, L(x_n))_{i \rightarrow L(x)}) + \beta \cdot \Delta((t, L(x_1), \dots, \\ &\quad L(x_{j-1}), L(x_j), \dots, L(x_n))_{i \rightarrow L(y)}) \\ &= \alpha \cdot A_j((x_1, \dots, x_n)_{i \rightarrow x}) + \beta \cdot A_j((x_1, \dots, x_n)_{i \rightarrow y}) \end{aligned}$$

3. ( $i < j$ ) then  $(t, L(x_1), \dots, L(x_{j-1}), L(x_{j+1}), \dots, L(x_n))_{i+1} = L(x_i) = L(\alpha \cdot x + \beta \cdot y) = \alpha \cdot L(x) + \beta \cdot L(y)$  and thus by multi linearity of  $\Delta$  we have

$$\begin{aligned} A_j(x_1, \dots, x_n) &= \Delta(t, L(x_1), \dots, L(x_{j-1}), L(x_{j+1}), \dots, L(x_n)) \\ &= \alpha \cdot \Delta((t, L(x_1), \dots, L(x_{j-1}), L(x_j), \dots, L(x_n))_{i \rightarrow L(x)}) + \beta \cdot \Delta((t, L(x_1), \dots, \\ &\quad L(x_{j-1}), L(x_j), \dots, L(x_n))_{i \rightarrow L(y)}) \\ &= \alpha \cdot A_j((x_1, \dots, x_n)_{i \rightarrow x}) + \beta \cdot A_j((x_1, \dots, x_n)_{i \rightarrow y}) \end{aligned}$$

so taken (1),(2) and (3) in account we have by 10.98 that

$$\begin{aligned} \check{\Delta}_L(x_1, \dots, x_n)(t) &= \alpha \cdot \sum_{j \in \{1, \dots, n\}} (-1)^{j-1} \cdot A_j((x_1, \dots, x_n)_{i \rightarrow x}) + \beta \cdot \sum_{j \in \{1, \dots, n\}} (-1)^{j-1} \cdot \\ &\quad A_j((x_1, \dots, x_n)_{i \rightarrow y}) \\ &= \alpha \cdot \check{\Delta}_L((x_1, \dots, x_n)_{i \rightarrow x}) + \beta \cdot \check{\Delta}_L((x_1, \dots, x_n)_{i \rightarrow y}) \end{aligned}$$

proving that  $\check{\Delta}_L$  is indeed multi linear.

To prove that  $\check{\Delta}_L$  is skew symmetric let  $x = (x_1, \dots, x_n) \in X^n$ ,  $t \in X$  such that there exists a  $i_0, j_0 \in \{1, \dots, n\}$  with  $i_0 \neq j_0$  and  $x_{i_0} = x_{j_0}$  then we have the following cases to consider for  $\Delta(t, L(x_1), \dots, L(x_{j-1}), L(x_{j+1}), \dots, L(x_n))$

1. ( $i_0 < j_0$ ) look then at the following cases for  $j \in \{1, \dots, n\}$

a. ( $j \neq i_0, j_0$ ) then we have the following sub cases to consider

i. ( $j < i_0 < j_0$ ) then  $(t, L(x_1), \dots, L(x_{j-1}), L(x_{j+1}), \dots, L(x_n))_{i_0} = L(x_{i_0})_{x_{i_0}=x_{j_0}} L(x_{j_0}) = (t, L(x_1), \dots, L(x_{j-1}), L(x_{j+1}), \dots, L(x_n))_{j_0}$  so that using the skew symmetry of  $\Delta$  we have  $\Delta(t, L(x_1), \dots, L(x_{j-1}), L(x_{j+1}), \dots, L(x_n)) = 0$

ii. ( $i_0 < j < j_0$ ) then  $(t, L(x_1), \dots, L(x_{j-1}), L(x_{j+1}), \dots, L(x_n))_{i_0+1} = L(x_{i_0})_{x_{i_0}=x_{j_0}} L(x_{j_0}) = (t, L(x_1), \dots, L(x_{j-1}), L(x_{j+1}), \dots, L(x_n))_{j_0}$  so that using the skew symmetry of  $\Delta$  we have  $\Delta(t, L(x_1), \dots, L(x_{j-1}), L(x_{j+1}), \dots, L(x_n)) = 0$

- iii.  $(i_0 < j_0 < j)$  then  $(t, L(x_1), \dots, L(x_{j-1}), L(x_{j+1}), \dots, L(x_n))_{i_0+1} = L(x_{i_0}) \underset{x_{i_0}=x_{j_0}}{=} L(x_{j_0}) = (t, L(x_1), \dots, L(x_{j-1}), L(x_{j+1}), \dots, L(x_n))_{j_0+1}$  so that using the skew symmetry of  $\Delta$  we have  $\Delta(t, L(x_1), \dots, L(x_{j-1}), L(x_{j+1}), \dots, L(x_n)) = 0$

so we have

$$\text{if } j \neq i_0, j_0 \Rightarrow \Delta(t, L(x_1), \dots, L(x_{j-1}), L(x_{j+1}), \dots, L(x_n)) = 0 \quad (10.99)$$

- b.  $(j = i_0)$  for  $j = i_0 < j_0$  so using the above lemma together with  $L(x_{i_0}) = L(x_{j_0})$  we have that  $(t, L(x_1), \dots, L(x_{i_0-1}), L(x_{i_0+1}), \dots, L(x_n)) \circ (j_0 \rightsquigarrow_n i_0 + 1) = (t, L(x_1), \dots, L(x_{j_0-1}), L(x_{j_0+1}), \dots, L(x_n))$ , using the skew symmetry of  $\Delta$  and 10.98 we have that  $(-1)^{j_0-i_0-1} \Delta(t, L(x_1), \dots, L(x_{i_0-1}), L(x_{i_0+1}), \dots, L(x_n)) = (j_0 \rightsquigarrow_n i_0 + 1) \Delta(t, L(x_1), \dots, L(x_{i_0-1}), L(x_{i_0+1}), \dots, L(x_n)) \circ (j_0 \rightsquigarrow_n i_0 + 1) = \Delta(t, L(x_1), \dots, L(x_{j_0-1}), L(x_{j_0+1}), \dots, L(x_n))$  giving thus

$$(-1)^{j_0-i_0-1} \Delta(t, L(x_1), \dots, L(x_{i_0-1}), L(x_{i_0+1}), \dots, L(x_n)) = \Delta(t, L(x_1), \dots, L(x_{j_0-1}), L(x_{j_0+1}), \dots, L(x_n)) \quad (10.100)$$

Now

$$\begin{aligned} \check{\Delta}_L(x_1, \dots, x_n)(t) &= \sum_{j \in \{1, \dots, n\}} (-1)^{j-1} \cdot \Delta(t, L(x_1), \dots, L(x_{j-1}), L(x_{j+1}), \dots, L(x_n)) \cdot x_j \\ &= \sum_{j \in \{1, \dots, n\} \setminus \{i_0, j_0\}} (-1)^{j-1} \cdot \Delta(t, L(x_1), \dots, L(x_{j-1}), L(x_{j+1}), \dots, L(x_n)) \cdot x_j + \sum_{j \in \{i_0\}} (-1)^{j-1} \cdot \Delta(t, L(x_1), \dots, L(x_{j-1}), L(x_{j+1}), \dots, L(x_n)) \cdot x_j + \sum_{j \in \{j_0\}} (-1)^{j-1} \cdot \Delta(t, L(x_1), \dots, L(x_{j-1}), L(x_{j+1}), \dots, L(x_n)) \cdot x_j \\ &\stackrel{10.99}{=} \sum_{j \in \{1, \dots, n\} \setminus \{i_0, j_0\}} (-1)^{j-1} \cdot 0 \cdot x_j + (-1)^{i_0-1} \cdot \Delta(t, L(x_1), \dots, L(x_{i_0-1}), L(x_{i_0+1}), \dots, L(x_n)) \cdot x_{i_0} + (-1)^{j_0-1} \cdot \Delta(t, L(x_1), \dots, L(x_{j_0-1}), L(x_{j_0+1}), \dots, L(x_n)) \cdot x_{j_0} \\ &\stackrel{10.100}{=} (-1)^{i_0-1} \cdot (-1) \Delta(t, L(x_1), \dots, L(x_{j_0-1}), \dots, L(x_n)) \cdot x_{j_0} + (-1)^{j_0-1} \cdot \Delta(t, L(x_1), \dots, L(x_{j_0-1}), L(x_{j_0+1}), \dots, L(x_n)) \cdot x_{j_0} \\ &\stackrel{x_{i_0}=x_{j_0}}{=} (-1)^{j_0-1} \cdot (-1) \Delta(t, L(x_1), \dots, L(x_{j_0-1}), \dots, L(x_n)) \cdot x_{j_0} + (-1)^{j_0-1} \cdot \Delta(t, L(x_1), \dots, L(x_{j_0-1}), L(x_{j_0+1}), \dots, L(x_n)) \cdot x_{j_0} \\ &= 0 \end{aligned}$$

2.  $(j_0 < i_0)$  take then  $i'_0 = j_0$  and  $j'_0 = i_0$  then  $i'_0 < j'_0$  and by using (1) we have that  $\check{\Delta}_L(x_1, \dots, x_n) = 0$

So in (1) and (2) we have proved that  $\check{\Delta}_L(x_1, \dots, x_n) = 0$  which means by 10.259 that  $\check{\Delta}_L$  is skew symmetric.  $\square$

Now using the fact that by the above theorem  $\check{\Delta}_L: X^n \rightarrow \text{Hom}(X, X)$  is n-linear and skew symmetric together with 10.266 there exists a unique  $y \in \text{Hom}(X, X)$  such that  $\check{\Delta}_L(x_1, \dots, x_n) = \Delta(x_1, \dots, x_n) \cdot y$ . We call this unique  $y$  the adjoint of  $L$  and note it by  $\text{adjoint}(L)$ . This leads directly to the following definition.

**Definition 10.287.** Let  $n \in \mathbb{N}_0$ ,  $X$  be a  $n$ -dimensional vector spaces over a field  $F$  with characteristics zero,  $L \in \text{Hom}(X, X)$  a linear transformation. Then there exists a unique (it does only dependent on  $L$ )  $\text{adjoint}(L) \in \text{Hom}(X, X)$  such that for every non trivial determinant function  $\Delta$  we have  $\check{\Delta}_L(x_1, \dots, x_n) = \Delta(x_1, \dots, x_n) \cdot \text{adjoint}(L)$

**Proof.** To prove uniqueness let  $\Delta'$  be another non trivial determinant function and  $\text{adjoint}'(L)$  such that  $\tilde{\Delta}'_L(x_1, \dots, x_n) = \Delta'(x_1, \dots, x_n) \cdot \text{adjoint}'(L)$ . Now using 10.267 there exists a  $\lambda \in F$  such that  $\Delta' = \lambda \cdot \Delta$ . Then we have

$$\begin{aligned}
 \Delta'(x_1, \dots, x_n) \cdot \text{adjoint}'(L)(t) &= \tilde{\Delta}'_L(x_1, \dots, x_n)(t) \\
 &= \sum_{j \in \{1, \dots, n\}} (-1)^{j-1} \cdot \Delta'(t, L(x_1), \dots, L(x_{j-1}) \cdot L(x_{j+1}), \dots, L(x_n)) \cdot x_j \\
 &= \sum_{\substack{j \in \{1, \dots, n\} \\ x_j}} (-1)^{j-1} \cdot \lambda \cdot \Delta(t, L(x_1), \dots, L(x_{j-1}) \cdot L(x_{j+1}), \dots, L(x_n)) \cdot \\
 &\quad x_j \\
 &= \lambda \cdot \sum_{\substack{j \in \{1, \dots, n\} \\ x_j}} (-1)^{j-1} \cdot \Delta(t, L(x_1), \dots, L(x_{j-1}) \cdot L(x_{j+1}), \dots, L(x_n)) \cdot \\
 &\quad x_j \\
 &= \lambda \cdot \tilde{\Delta}_L(x_1, \dots, x_n)(t) \\
 &= \lambda \cdot (\Delta(x_1, \dots, x_n) \cdot \text{adjoint}(L)(t)) \\
 &= \Delta'(x_1, \dots, x_1) \cdot \text{adjoint}(L)(t)
 \end{aligned}$$

As  $\Delta'$  is non trivial there exists a  $(x_1, \dots, x_n)$  such that  $\Delta'(x_1, \dots, x_n) \neq 0$  then by dividing by  $\Delta'(x_1, \dots, x_n)$  we get  $\text{adjoint}'(L) = \text{adjoint}(L)$ .  $\square$

**Example 10.288.** If  $n=1$  then if  $\{e_1\}$  is a basis of  $X$  then if  $x \in X$  we have  $x = \alpha \cdot e_1$  and  $L(x) = \alpha \cdot L(e_1)$ , from 10.284 we have that  $\tilde{\Delta}_L(x_1)(x) = \Delta(x) \cdot x_1$  and if  $x_1 = \beta \cdot e_1$  we have  $\tilde{\Delta}_L(x_1)(x) = \alpha \cdot \beta \cdot \Delta(e_1) \cdot e_1 = \Delta(\beta \cdot e_1) \cdot (\alpha \cdot e_1) = \Delta(x_1) \cdot x$  so that  $\text{adjoint}(L)(x) = x$  or we have that  $\text{adjoint}(L) = i_X$ . Note also that if  $L(e_1) = \gamma \cdot e_1$  we have by 10.278 that  $\det(L) = \gamma$  so that we have

1.  $\text{adjoint}(L(x)) = L(x) = L(\alpha \cdot e_1) = \alpha \cdot L(e_1) = \alpha \cdot (\gamma \cdot e_1) = \gamma \cdot (\alpha \cdot e_1) = \gamma \cdot x = \det(L) \cdot x = \det(L) \cdot 1_X(x)$  or  $\text{adjoint}(L) \circ L = \det(L) \cdot 1_X$
2.  $L(\text{adjoint}(x)) = L(x) \underset{\text{see (1)}}{=} \det(L) \cdot 1_X(x)$  or  $L \circ \text{adjoint}(L) = \det(L) \cdot 1_X$

Actually we can generalize this result to any dimension as the following theorem shows

**Theorem 10.289.** Let  $n \in \mathbb{N}_0$ ,  $X$  be a  $n$ -dimensional vector space over a field of characteristic zero,  $L \in \text{Hom}(X, X)$  then we have

1.  $\text{adjoint}(L) \circ L = \det(L) \cdot 1_X$
2.  $L \circ \text{adjoint}(L) = \det(L) \cdot 1_X$

**Proof.**

1. Let  $(e_1, \dots, e_n)$  a basis for  $X$  and let  $\Delta$  the non zero determinant function defined in 10.265 such that  $\Delta(e_1, \dots, e_n) = 1$  then if  $x \in X$  we have

$$\begin{aligned}
 (\text{adjoint}(L) \circ L)(x) &= \text{adjoint}(L)(L(x)) \\
 &= \Delta(e_1, \dots, e_n) \cdot \text{adjoint}(L)(L(x)) \\
 &= \tilde{\Delta}_L(e_1, \dots, e_n)(L(x)) \\
 &= \sum_{i \in \{1, \dots, n\}} (-1)^{i-1} \cdot \Delta(L(x), L(e_1), \dots, L(e_{i-1}), L(e_{i+1}), \dots, L(e_n)) \cdot \\
 &\quad L(e_i) \\
 &= \sum_{i \in \{1, \dots, n\}} (-1)^{i-1} \cdot \Delta_L(x, e_1, \dots, e_{i-1}, e_{i+1}, \dots, e_n) \\
 &= \hat{\Delta}_L(x, (e_1, \dots, e_n)) \\
 &\stackrel{10.272}{=} \Delta_L(e_1, \dots, e_n) \cdot x \\
 &\stackrel{10.276}{=} (\det(L) \cdot \Delta(e_1, \dots, e_n)) \cdot x \\
 &= \det(L) \cdot x \\
 &= (\det(L) \cdot 1_X)(x)
 \end{aligned}$$

2. Let  $(e_1, \dots, e_n)$  is a basis of  $X$  and choose  $\Delta$  such that  $\Delta(e_1, \dots, e_n) = 1$  then we have

$$\begin{aligned}
 (L \circ \text{adjoint}(L))(t) &= L(\text{adjoint}(L)(t)) \\
 &= L(\Delta(e_1, \dots, e_n) \cdot \text{adjoint}(L)(t)) \\
 &= L(\hat{\Delta}_L(e_1, \dots, e_n)(t)) \\
 &= L\left(\sum_{j \in \{1, \dots, n\}} (-1)^{j-1} \cdot \Delta(t, L(e_1), \dots, L(e_{j-1}), L(e_{j+1}), \dots, \right. \\
 &\quad \left. L(e_n)) \cdot e_j\right) \\
 &\stackrel{L \text{ is linear}}{=} \sum_{j \in \{1, \dots, n\}} (-1)^{j-1} \cdot \Delta(t, L(e_1), \dots, L(e_{j-1}), L(e_{j+1}), \dots, L(e_n)) \cdot \\
 &\quad L(e_j) \\
 &= \hat{\Delta}_L(t, (L(e_1), \dots, L(e_n))) \\
 &\stackrel{10.272}{=} \Delta(L(e_1), \dots, L(e_n)) \cdot t \\
 &= \Delta_L(e_1, \dots, e_n) \cdot t \\
 &= (\det(L) \cdot \Delta(e_1, \dots, e_n)) \cdot t \\
 &= \det(L) \cdot t \\
 &= (\det(L) \cdot 1_X)(t)
 \end{aligned}$$

□

**Corollary 10.290.** Let  $n \in \mathbb{N}_0$ ,  $X$  be a  $n$ -dimensional vector space over a field  $F$  of characterization zero,  $L \in \text{Hom}(X, X)$  then  $L$  is a isomorphism if and only if  $\det(L) \neq 0$  and if  $L$  is a isomorphism then  $L^{-1} = (\det(L))^{-1} \cdot \text{adjoint}(L)$

**Proof.** Using 10.282 and 10.214 we have that  $L$  is a isomorphism if and only if  $\det(L) \neq 0$  proving the first part of the corollary. For the second part if  $L$  is a isomorphism then  $\det(L) \neq 0$  and using the above theorem we have

$$\begin{aligned}
 \frac{\text{adjoint}(L)}{\det(L)} \circ L &= 1_X \\
 L \circ \frac{\text{adjoint}(L)}{\det(L)} &= 1_X
 \end{aligned}$$

proving that  $L^{-1} = \frac{\text{adjoint}(L)}{\det(L)} = (\det(L))^{-1} \cdot \text{adjoint}(L)$

□

## 10.9 Matrices

**Definition 10.291.** Let  $F$  be a field,  $n, m \in \mathbb{N}_0$  then a  $n \times m$  matrix in  $F$  is a mapping  $M: \{1, \dots, n\} \times \{1, \dots, m\} \rightarrow F$  (so a  $n \times m$  matrix is a element of  $F^{\{1, \dots, n\} \times \{1, \dots, m\}}$ ). We note  $M_{i,j}$  for  $M(i, j)$  and call  $\{1, \dots, n\}$  the row indexes of  $M$  and  $\{1, \dots, m\}$  the column indexes of  $M$ . A  $1 \times m$  matrix is called is called a row matrix and a  $n \times 1$  matrix is called a column matrix.

**Notation 10.292.** A  $n \times m$  matrix  $M$  is also noted as

$$\begin{pmatrix} M_{1,1} & \dots & M_{1,m} \\ \dots & \dots & \dots \\ M_{n,1} & \dots & M_{n,m} \end{pmatrix}$$

a row matrix  $R$  is noted as

$$(R_{1,1} \ \dots \ R_{1,m})$$

and a column matrix  $C$  is noted as

$$\begin{pmatrix} C_{1,1} \\ \vdots \\ C_{n,1} \end{pmatrix}$$

**Example 10.293.** Given  $n \in \mathbb{N}_0$  then the  $n \times n$  matrix  $E$  defined by  $E: \{1, \dots, n\} \times \{1, \dots, n\}$  defined by  $(i, j) \rightarrow E(i, j) = E_{i,j} = \delta_{i,j}$  or in matrix notation

$$\begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix}$$

**Definition 10.294.** Let  $F$  be a field then we define the sum, scalar product, product as follows

1. If  $A, B$  are  $n \times m$  matrices then  $A + B: \{1, \dots, n\} \times \{1, \dots, m\} \rightarrow F$  is defined by  $(i, j) \rightarrow (A + B)(i, j) = A(i, j) + B(j)$  or in our shorter notation  $(A + B)_{i,j} = A_{i,j} + B_{i,j}$
2. If  $A$  is a  $n \times m$  matrix and  $\alpha \in F$  then  $\alpha \cdot A: \{1, \dots, n\} \times \{1, \dots, m\} \rightarrow F$  is defined by  $(i, j) \rightarrow (\alpha \cdot A)(i, j) = \alpha \cdot A(i, j)$  or in our shorter notation  $(\alpha \cdot A)_{i,j} = \alpha \cdot A_{i,j}$
3. If  $A$  is a  $n \times m$  matrix and  $B$  is a  $m \times k$  matrix then  $A \cdot B: \{1, \dots, n\} \times \{1, \dots, k\} \rightarrow F$  is defined by  $(i, j) \rightarrow (A \cdot B)(i, j) = \sum_{l=1}^m A(i, l) \cdot B(l, j)$  or in our shorter notation  $(A \cdot B)_{i,j} = \sum_{l=1}^m A_{i,l} \cdot B_{l,j}$

**Definition 10.295.** Let  $F$  be a field,  $n, m \in \mathbb{N}_0$  then  $\mathcal{M}(n \times m, F) = F^{\{1, \dots, n\} \times \{1, \dots, m\}}$  (the set of  $n \times m$  matrices)

**Theorem 10.296.** Let  $F$  be a field,  $n, m \in \mathbb{N}_0$  then  $\langle \mathcal{M}(n \times m, F), +, \cdot \rangle$  forms a vector space over  $F$ . The zero element is the matrix  $0$  where  $\forall i, j \in \{1, \dots, n\} \times \{1, \dots, m\}$  we have  $0_{i,j} = 0$  which as usually noted by  $0$ .

**Proof.** First we prove the group axioms ;

1. **(associativity)** If  $A, B, C \in \mathcal{M}(n \times m, F)$  then  $((A + B) + C)_{i,j} = (A + B)_{i,j} + C_{i,j} = ((A_{i,j} + B_{i,j}) + C_{i,j}) = A_{i,j} + (B_{i,j} + C_{i,j}) = A_{i,j} + (B + C)_{i,j} = (A + (B + C))_{i,j} \Rightarrow ((A + B) + C) = (A + (B + C))$
2. **(commutativity)** If  $A, B \in \mathcal{M}(n \times m, F)$  then  $(A + B)_{i,j} = A_{i,j} + B_{i,j} = B_{i,j} + A_{i,j} = (B + A)_{i,j} \Rightarrow A + B = B + A$
3. **(neutral element)**  $(A + 0)_{i,j} + A_{i,j} + 0_{i,j} = A_{i,j} + 0 = A_{i,j} \Rightarrow A + 0 = A$  using commutativity we have then  $0 + A = a$
4. **(inverse element)** If  $A \in \mathcal{M}(n \times m, F)$  define then  $-A$  by  $\forall (i, j) \in \{1, \dots, n\} \times \{1, \dots, m\}$  we have  $(-A)_{i,j} = -A_{i,j}$  so that  $(A + (-A))_{i,j} = A_{i,j} + (-A)_{i,j} = A_{i,j} + (-A_{i,j}) = 0 \Rightarrow A + (-A) = 0$  and using commutativity we have then also  $(-A) + A = 0$

For the rest of the axioms note:

1. If  $A, B \in \mathcal{M}(n \times m, F)$  and  $\alpha \in F$  then  $(\alpha \cdot (A + B))_{i,j} = \alpha \cdot (A + B)_{i,j} = \alpha \cdot (A_{i,j} + B_{i,j}) = \alpha \cdot A_{i,j} + \alpha \cdot B_{i,j} = (\alpha \cdot A)_{i,j} + (\alpha \cdot B)_{i,j} = (\alpha \cdot A + \alpha \cdot B)_{i,j} \Rightarrow \alpha \cdot (A + B) = \alpha \cdot A + \alpha \cdot B$
2. If  $\alpha, \beta \in F$  and  $A \in \mathcal{M}(n \times m, F)$  then  $((\alpha + \beta) \cdot A)_{i,j} = (\alpha + \beta) \cdot A_{i,j} = \alpha \cdot A_{i,j} + \beta \cdot A_{i,j} = (\alpha \cdot A)_{i,j} + (\beta \cdot A)_{i,j} = (\alpha \cdot A + \beta \cdot A)_{i,j} \Rightarrow (\alpha + \beta) \cdot A = \alpha \cdot A + \beta \cdot A$
3. If  $\alpha, \beta \in F$  and  $A \in \mathcal{M}(n \times m, F)$  then  $((\alpha \cdot \beta) \cdot A)_{i,j} = (\alpha \cdot \beta) \cdot A_{i,j} = \alpha \cdot (\beta \cdot A_{i,j}) = \alpha \cdot (\beta \cdot A)_{i,j} = (\alpha \cdot (\beta \cdot A))_{i,j} \Rightarrow (\alpha \cdot \beta) \cdot A = \alpha \cdot (\beta \cdot A)$

4. If 1 is the neutral element in  $F$ ,  $A \in \mathcal{M}(n \times m, F)$  then  $(1 \cdot A)_{i,j} = 1 \cdot A_{i,j} = A_{i,j} \Rightarrow (1 \cdot A) = A$   $\square$

**Theorem 10.297.** Given  $n \in \mathbb{N}_0$  and a field  $F$  then  $\langle \mathcal{M}(n \times n, F), \cdot \rangle$  is a semi-group with multiplicative neutral element  $E$

**Proof.**

**associativity.** Given  $A, B, C \in \mathcal{M}(n \times n, F)$  then for  $i, j \in \{1, \dots, n\}$  we have

$$\begin{aligned}
 ((A \cdot B) \cdot C)_{i,j} &= \sum_{k=1}^n (A \cdot B)_{i,k} \cdot C_{k,j} \\
 &= \sum_{k=1}^n \left( \sum_{l=1}^n A_{i,l} \cdot B_{l,k} \right) \cdot C_{k,j} \\
 &= \sum_{k=1}^n \left( \sum_{l=1}^n ((A_{i,l} \cdot B_{l,k}) \cdot C_{k,j}) \right) \\
 &\stackrel{10.47}{=} \sum_{l=1}^n \left( \sum_{k=1}^n (A_{i,l} \cdot B_{l,k}) \cdot C_{k,j} \right) \\
 &= \sum_{l=1}^n \left( \sum_{k=1}^n A_{i,l} \cdot (B_{l,k} \cdot C_{k,j}) \right) \\
 &= \sum_{l=1}^n A_{i,l} \cdot \left( \sum_{k=1}^n (B_{l,k} \cdot C_{k,j}) \right) \\
 &= \sum_{l=1}^n A_{i,l} \cdot (B \cdot C)_{l,j} \\
 &= ((A \cdot B) \cdot C)_{i,j}
 \end{aligned}$$

proving that  $(A \cdot B) \cdot C = A \cdot (B \cdot C)$

**neutral element.** Let  $A \in \mathcal{M}(n \times n, F)$  and  $i, j \in \{1, \dots, n\}$  then

$$\begin{aligned}
 (A \cdot E)_{i,j} &= \sum_{k=1}^n A_{i,k} \cdot E_{k,j} \\
 &= \sum_{k=1}^n A_{i,k} \cdot \delta_{k,j} \\
 &= A_{i,j}
 \end{aligned}$$

proving that  $A \cdot E = A$  and

$$\begin{aligned}
 (E \cdot A)_{i,j} &= \sum_{k=1}^n E_{i,k} \cdot A_{k,j} \\
 &= \sum_{k=1}^n \delta_{i,k} \cdot A_{k,j} \\
 &= A_{i,j}
 \end{aligned}$$

proving that  $E \cdot A = A$   $\square$

**Definition 10.298.** Let  $F$  be a field,  $n, m \in \mathbb{N}_0$  and  $M \in \mathcal{M}(n \times m, F)$  then  $M^T \in \mathcal{M}(m \times n, F)$  is defined by  $(M^T)_{i,j} = M_{j,i}$

**Theorem 10.299.** Let  $F$  be a field  $n, m \in \mathbb{N}_0$  and  $M, N \in \mathcal{M}(n \times n, F)$  then  $(M \cdot N)^T = N^T \cdot M^T$

**Proof.** Let  $i, j \in \{1, \dots, n\}$  then

$$\begin{aligned}
 ((M \cdot N)^T)_{i,j} &= (M \cdot N)_{j,i} \\
 &= \sum_{k=1}^n M_{j,k} \cdot N_{k,i} \\
 &= \sum_{k=1}^n (M^T)_{k,j} \cdot (N^T)_{i,k} \\
 &= (N^T \cdot M^T)_{i,j}
 \end{aligned}$$

□

**Definition 10.300.** Let  $F$  be a field,  $n, m \in \mathbb{N}_0$  and  $M \in \mathcal{M}(n \times m)$  then we define the following:

1.  $\forall i \in \{1, \dots, n\}$   $\text{row}(M, i) = (M_{i,1}, \dots, M_{i,m}) \in F^m$  (so  $\forall j \in \{1, \dots, m\}$  we have  $\text{row}(M, i)_j = M_{i,j}$ )
2.  $\forall i \in \{1, \dots, m\}$   $\text{col}(M, i) = (M_{1,i}, \dots, M_{n,i}) \in F^n$  (so  $\forall j \in \{1, \dots, n\}$  we have  $\text{col}(M, i)_j = M_{j,i}$ )
3.  $\text{rows}(M) = \{\text{row}(M, i) \mid i \in \{1, \dots, n\}\} \subseteq F^m$
4.  $\text{cols}(M) = \{\text{col}(M, i) \mid i \in \{1, \dots, m\}\} \subseteq F^n$
5. The row rank of  $M$  is the  $\dim(\mathcal{S}(\text{rows}(M)))$  [which as  $\dim(F^m) = m$  (see 10.164) and 10.210) must be finite and lower than  $m$ ]
6. The column rank of  $M$  is the  $\dim(\mathcal{S}(\text{cols}(M)))$  [which as  $\dim(F^n) = n$  (see 10.164) and 10.210) must be finite and lower than  $n$ ]

**Theorem 10.301.** Let  $F$  be a field,  $n, m \in \mathbb{N}_0$  and  $M \in \mathcal{M}(n \times m, F)$  then the column rank and the row rank of  $M$  are equal. The column (or row) rank of the matrix is called the rank of the matrix, noted by  $\text{rank}(M)$ . Note that by the previous definition we must have that  $\text{rank}(M) \leq \min(n, m)$ .

**Proof.** Lets assume that the column rank of  $M$  is  $c$  and the row rank of  $M$  is  $r$  then we have

1. As the column rank of  $M$  is  $c$  we have that  $\dim(\mathcal{S}(\text{cols}(M))) = r$  and there exists thus a basis  $\{e_i\}_{i \in \{1, \dots, c\}}$  of  $\mathcal{S}(\text{cols}(M))$ . This means that  $\forall i \in \{1, \dots, m\}$  there exists a  $\{\lambda_{i,j}\}_{j \in \{1, \dots, c\}}$  such that  $\text{col}(M, i) = \sum_{j=1}^c \lambda_{i,j} \cdot e_j$ . Now  $\forall k \in \{1, \dots, n\}$  and  $\forall i \in \{1, \dots, m\}$  we have  $\text{row}(M, k)_i = M_{k,i} = \text{col}(M, i)_k = (\sum_{j=1}^c \lambda_{i,j} \cdot e_j)_k \stackrel{10.10}{=} \sum_{j=1}^c \lambda_{i,j} \cdot (e_j)_k$ . If we define now  $\forall j \in \{1, \dots, c\}$   $f_j = (\lambda_{1,j}, \dots, \lambda_{m,j}) \in F^m$  so that  $\forall i \in \{1, \dots, m\} \models (f_j)_i = \lambda_{i,j}$  then we have that  $\text{row}(M, k)_i = \sum_{j=1}^c (f_j)_i \cdot (e_j)_k = \sum_{j=1}^c (e_j)_k \cdot (f_j)_i = (\sum_{j=1}^c (e_j)_k \cdot f_j)_i \Rightarrow \text{row}(M, k) = \sum_{j=1}^c (e_j)_k \cdot f_j$  proving that  $\forall k \in \{1, \dots, m\} \models \text{row}(M, k) \in \mathcal{S}(\{f_j \mid j \in \{1, \dots, c\}\}) \Rightarrow \text{rows}(M) \subseteq \mathcal{S}(\{f_j \mid j \in \{1, \dots, c\}\})$ . From this and 10.138 we have that  $\mathcal{S}(\text{rows}(M)) \subseteq \mathcal{S}(\mathcal{S}(\{f_j \mid j \in \{1, \dots, c\}\})) \stackrel{10.137}{=} \mathcal{S}(\{f_j \mid j \in \{1, \dots, c\}\})$ . Using 10.158 we have that  $\dim(\mathcal{S}(\{f_j \mid j \in \{1, \dots, c\}\})) \leq c$  and as  $\mathcal{S}(\text{rows}(M)) \subseteq \mathcal{S}(\{f_j \mid j \in \{1, \dots, c\}\})$  we have by 10.210 that

$$r = \dim(\mathcal{S}(\text{rows}(M))) \leq c. \quad (10.101)$$

2. As the row rank of  $M$  is  $r$  we have that  $\dim(\mathcal{S}(\text{rows}(M))) = r$  and there exists thus a basis  $\{e_i\}_{i \in \{1, \dots, r\}}$  of  $\mathcal{S}(\text{rows}(M))$ . This means that  $\forall i \in \{1, \dots, n\}$  there exists a  $\{\lambda_{i,j}\}_{j \in \{1, \dots, r\}}$  such that  $\text{row}(M, i) = \sum_{j=1}^r \lambda_{i,j} \cdot e_j$ . Now  $\forall k \in \{1, \dots, m\}$  and  $\forall i \in \{1, \dots, n\}$  we have that  $\text{col}(M, k)_i = M_{i,k} = \text{row}(M, i)_k = (\sum_{j=1}^r \lambda_{i,j} \cdot e_j)_k = \sum_{j=1}^r \lambda_{i,j} (e_j)_k$ . If we define now  $\forall j \in \{1, \dots, r\}$   $f_j = (\lambda_{1,j}, \dots, \lambda_{n,j}) \in F^n$  then  $\text{col}(M, k)_i = \sum_{j=1}^r (f_j)_i \cdot (e_j)_k = \sum_{j=1}^r (e_j)_k \cdot (f_j)_i = (\sum_{j=1}^r (e_j)_k \cdot f_j)_i \Rightarrow \text{col}(M, k) = \sum_{j=1}^r (e_j)_k \cdot f_j$  proving that  $\forall k \in \{1, \dots, m\} \models \text{col}(M, k) \in \mathcal{S}(\{f_j \mid j \in \{1, \dots, r\}\}) \Rightarrow \text{cols}(M) \subseteq \mathcal{S}(\{f_j \mid j \in \{1, \dots, r\}\})$ . From this and 10.138 we have that  $\mathcal{S}(\text{cols}(M)) \subseteq \mathcal{S}(\mathcal{S}(\{f_j \mid j \in \{1, \dots, r\}\})) \stackrel{10.137}{=} \mathcal{S}(\{f_j \mid j \in \{1, \dots, r\}\})$ . Using 10.158 we have that  $\dim(\mathcal{S}(\{f_j \mid j \in \{1, \dots, r\}\})) \leq r$  and as  $\mathcal{S}(\text{cols}(M)) \subseteq \mathcal{S}(\{f_j \mid j \in \{1, \dots, r\}\})$  we have by 10.210 that

$$c = \dim(\mathcal{S}(\text{cols}(M))) \leq r \quad (10.102)$$

Using 10.101 and 10.102 we have finally  $c=r$  proving the theorem.  $\square$

**Definition 10.302.** Let  $X, Y$  be finite dimensional vector spaces over a field  $F$  with  $\dim(X) = n$ ,  $\dim(Y) = m$  and let  $L: X \rightarrow Y$  be a linear mapping ( $L \in \text{Hom}(X, Y)$ ). Then if  $\{e_i\}_{i \in \{1, \dots, n\}}$  is a basis for  $X$ ,  $\{f_i\}_{i \in \{1, \dots, m\}}$  is a basis of  $Y$  we have  $\forall i \in \{1, \dots, n\}$  the existence of a unique family  $\{L_{i,j}\}_{j \in \{1, \dots, m\}}$  such that  $L(e_i) = \sum_{j \in \{1, \dots, m\}} L_{i,j} \cdot f_j$ . This defines a matrix  $\mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{1, \dots, m\}}) \in M(m \times n, F)$  by  $\mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{1, \dots, m\}})_{i,j} = L_{i,j} = (L(e_j))_i$  so that we have  $L(e_i) = \sum_{j \in \{1, \dots, m\}} \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{1, \dots, m\}})_{i,j} \cdot f_j$ . If  $\{e_i\}_{i \in \{1, \dots, n\}}$  and  $\{f_i\}_{i \in \{1, \dots, m\}}$  is assumed to be known we use the shorter notation  $\mathcal{M}(L)$ .

**Definition 10.303.** Let  $X$  be a finite dimensional space with base  $\{e_i\}_{i \in \{1, \dots, n\}}$ ,  $L \in \text{Hom}(X) = \text{Hom}(X, X)$  then  $\mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}) = \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}})$

**Example 10.304.** Let  $X$  be a finite dimensional vector space over a field  $F$  with  $\dim(X) = n$  and  $\{e_i\}_{i \in \{1, \dots, n\}}$  a basis for  $X$  then we have the following equivalences for  $L \in \text{Hom}(X)$  :

1.  $L = 1_X$
2.  $\mathcal{M}(1_X, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}}) = E$

**Proof.**

**1  $\Rightarrow$  2.**

$$\begin{aligned}
 \mathcal{M}(1_X, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}})_{i,j} &= 1_X(e_j)_i \\
 &= (e_j)_i \\
 &= \delta_{i,j} \\
 &= \delta_{i,j} \\
 &= E_{i,j}
 \end{aligned}$$

**2  $\Rightarrow$  1.** Let  $x \in X \Rightarrow x = \sum_{i=1}^n x_i \cdot e_i$  then

$$\begin{aligned}
 L(x) &= L\left(\sum_{i=1}^n x_i \cdot e_i\right) = \sum_{i=1}^n x_i \cdot L(e_i) \\
 &= \sum_{i=1}^n x_i \cdot \left(\sum_{j=1}^m E_{i,j} \cdot e_j\right) \\
 &= \sum_{i=1}^n x_i \cdot \left(\sum_{j=1}^m \delta_{i,j} \cdot e_j\right) \\
 &= \sum_{i=1}^n x_i \cdot e_i \\
 &= x = 1_X(x)
 \end{aligned}$$

proving that  $L = 1_X$   $\square$

**Theorem 10.305.** Let  $X, Y$  be finite dimensional vector spaces over a field  $F$  with  $\dim(X) = n$ ,  $\dim(Y) = m$ ,  $\{e_i\}_{i \in \{1, \dots, n\}}$ ,  $\{f_i\}_{i \in \{1, \dots, m\}}$  basis for  $X, Y$  and let  $L \in \text{Hom}(X, Y)$  then we have

1. If  $x = \sum_{i \in \{1, \dots, n\}} x_i \cdot e_i$  and  $L(x) = \sum_{i \in \{1, \dots, n\}} L(x)_i \cdot f_i$  (see 10.154) then  $L(x)_i = \sum_{i \in \{1, \dots, n\}} \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{1, \dots, m\}})_{i,i} \cdot x_i$
2.  $\text{rank}(L) = \text{rank}(\mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{1, \dots, m\}}))$

**Proof.**

1. If  $x = \sum_{i \in \{1, \dots, n\}} x_i \cdot e_i$  then  $L(x) = L(\sum_{i \in \{1, \dots, n\}} x_i \cdot e_i) = \sum_{i \in \{1, \dots, n\}} x_i \cdot L(e_i) = \sum_{i \in \{1, \dots, n\}} L(e_i) \cdot x_i = \sum_{i \in \{1, \dots, n\}} (\sum_{j \in \{1, \dots, m\}} \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{1, \dots, m\}})_{j,i} \cdot f_j) \cdot x_i = \sum_{i \in \{1, \dots, n\}} (\sum_{j \in \{1, \dots, m\}} (\mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{1, \dots, m\}})_{j,i} \cdot f_j) \cdot x_i) \stackrel{10.47}{=} \sum_{j \in \{1, \dots, m\}} (\sum_{i \in \{1, \dots, n\}} (\mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{1, \dots, m\}})_{j,i} \cdot f_j) \cdot x_i) = \sum_{j \in \{1, \dots, m\}} (\sum_{i \in \{1, \dots, n\}} \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{1, \dots, m\}})_{j,i} \cdot x_i) \cdot f_j \stackrel{L(x) = \sum_{i \in \{1, \dots, m\}} L(x)_j \cdot f_j \text{ and expansion in a base is unique}}{=} L(x)_j = \sum_{i \in \{1, \dots, n\}} \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{1, \dots, m\}})_{j,i} \cdot x_i$

2. Let now  $M = \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{1, \dots, m\}})$ . To prove that  $\text{rank}(L) = \text{rank}(M)$  we proceed as follows:

- If  $y \in L(X)$  then there exists a  $x = \sum_{i \in \{1, \dots, n\}} x_i \cdot e_i \in X$  such that  $y = L(x)$  giving that  $y = L(\sum_{i \in \{1, \dots, n\}} x_i \cdot e_i) = \sum_{i \in \{1, \dots, n\}} x_i \cdot L(e_i) \in \mathcal{S}(\{L(e_i)\}_{i \in \{1, \dots, n\}})$  proving that  $L(X) \subseteq \mathcal{S}(\{L(e_i)\}_{i \in \{1, \dots, n\}})$ . If  $y \in \mathcal{S}(\{L(e_i)\}_{i \in \{1, \dots, n\}})$  then there exists a  $\{\alpha_i\}_{i \in I}$  such that  $y = \sum_{i \in \{1, \dots, n\}} \alpha_i \cdot L(e_i) = L(\sum_{i \in \{1, \dots, n\}} \alpha_i \cdot e_i) \in L(X)$  proving that  $\mathcal{S}(\{L(e_i)\}_{i \in \{1, \dots, n\}}) \subseteq L(X)$ . So we have

$$\mathcal{S}(\{L(e_i)\}_{i \in \{1, \dots, n\}}) = L(X) \quad (10.103)$$

Using 10.103 and 10.158 there exists a  $J \subseteq \{1, \dots, n\}$  such that  $\{L(e_i)\}_{i \in J}$  forms a basis of  $L(X)$ . Now as  $\text{rank}(L) = \dim(L(X))$  we must have that  $\#(J) = \text{rank}(L)$  giving the following:

$$\begin{aligned} \exists \{i_j\}_{j \in \{1, \dots, \text{rank}(L)\}} & \quad \text{in} & \{1, \dots, \\ n\} \text{ such that } \{L(e_{i_j})\}_{j \in \{1, \dots, \text{rank}(L)\}} & \text{is a basis of } L(X) \end{aligned} \quad (10.104)$$

- Define now  $B_c = \{\text{col}(M, i_j) | j \in \{1, \dots, \text{rank}(L)\}\} \subseteq \text{cols}(M)$  so that by 10.138 we have  $\mathcal{S}(B_c) \subseteq \mathcal{S}(\text{cols}(M))$ . If  $c \in \mathcal{S}(\text{cols}(M)) \Rightarrow \exists \{\lambda_i\}_{i \in \{1, \dots, n\}}$  such that  $c = \sum_{i \in \{1, \dots, n\}} \lambda_i \cdot \text{col}(M, i)$ , so  $\forall k \in \{1, \dots, m\}$  we have  $c_k = (\sum_{i \in \{1, \dots, n\}} \lambda_i \cdot \text{col}(M, i))_k \stackrel{10.10}{=} \sum_{i \in \{1, \dots, n\}} \lambda_i \cdot \text{col}(M, i)_k = \sum_{i \in \{1, \dots, n\}} \lambda_i \cdot M_{k,i}$ . Take now  $y = \sum_{i \in \{1, \dots, n\}} \lambda_i \cdot L(e_i) \in L(X)$  (see 10.103) then as  $\{L(e_{i_j})\}_{j \in \{1, \dots, \text{rank}(L)\}}$  is a basis of  $L(X)$  (see 10.104) we have that  $\exists \{\beta_i\}_{i \in \{1, \dots, \text{rank}(k)\}}$  such that  $y = \sum_{l \in \{1, \dots, \text{rank}(k)\}} \beta_l \cdot L(e_{i_l}) = \sum_{l \in \{1, \dots, \text{rank}(L)\}} \beta_l \cdot (\sum_{j \in \{1, \dots, m\}} M_{j,i_l} \cdot f_j) = \sum_{j \in \{1, \dots, m\}} (\sum_{l \in \{1, \dots, \text{rank}(L)\}} \beta_l \cdot M_{j,i_l}) \cdot f_j$ . So  $\sum_{j \in \{1, \dots, m\}} (\sum_{l \in \{1, \dots, \text{rank}(L)\}} \beta_l \cdot M_{j,i_l}) \cdot f_j = y \stackrel{\text{definition of } y}{=} \sum_{i \in \{1, \dots, n\}} \lambda_i \cdot L(e_i) = \sum_{i \in \{1, \dots, n\}} \lambda_i \cdot (\sum_{j \in \{1, \dots, m\}} M_{j,i} \cdot f_j) = \sum_{j \in \{1, \dots, m\}} (\sum_{i \in \{1, \dots, n\}} \lambda_i \cdot M_{j,i}) \cdot f_j$  and given the uniqueness of the expansion in the base of  $\{f_j\}_{j \in \{1, \dots, m\}}$  we get that  $\sum_{l \in \{1, \dots, \text{rank}(L)\}} \beta_l \cdot M_{j,i_l} = \sum_{i \in \{1, \dots, n\}} \lambda_i \cdot M_{j,i} = c_i \Rightarrow c_i = \sum_{l \in \{1, \dots, \text{rank}(L)\}} \beta_l \cdot \text{col}(M, i_l) \Rightarrow c = \sum_{l \in \{1, \dots, \text{rank}(L)\}} \beta_l \cdot \text{col}(M, i_l) \in \mathcal{S}(B_c)$ . This proves the following

$$\mathcal{S}(\{\text{col}(M, i_j) | j \in \{1, \dots, \text{rank}(L)\}\}) = \mathcal{S}(\text{cols}(M)) \quad (10.105)$$

- Let now  $\{\lambda_i\}_{i \in \{1, \dots, \text{rank}(L)\}}$  be such that  $0 = \sum_{j \in \{1, \dots, \text{rank}(L)\}} \lambda_j \cdot \text{col}(M, i_j)$  we have that  $\forall k \in \{1, \dots, m\}$  that  $0 = 0_j = \sum_{j \in \{1, \dots, \text{rank}(L)\}} \lambda_j \cdot \text{col}(M, i_j)_k = \sum_{j \in \{1, \dots, \text{rank}(L)\}} \lambda_j \cdot M_{k,i_j}$ . Take now  $y = \sum_{l \in \{1, \dots, \text{rank}(L)\}} \lambda_l \cdot L(e_{i_l}) \in L(X)$  (by 10.103) then  $y = \sum_{l \in \{1, \dots, \text{rank}(L)\}} \lambda_l \cdot (\sum_{j \in \{1, \dots, m\}} M_{j,i_l} \cdot f_j) = \sum_{j \in \{1, \dots, m\}} (\sum_{l \in \{1, \dots, \text{rank}(L)\}} \lambda_l \cdot M_{j,i_l}) \cdot f_j = \sum_{j \in \{1, \dots, m\}} 0 \cdot f_j = 0$  giving that  $0 = y = \sum_{l \in \{1, \dots, \text{rank}(L)\}} \lambda_l \cdot L(e_{i_l}) \stackrel{\{L(e_{i_l})\}_{l \in \{1, \dots, \text{rank}(L)\}} \text{ is a basis}}{\Rightarrow} \forall l \in \{1, \dots, \text{rank}(L)\}$  we have  $\lambda_l = 0$ . This proves that

$$\{\text{col}(M, i_j)\}_{j \in \{1, \dots, \text{rank}(L)\}}$$
 is linear independent  $\quad (10.106)$

Using 10.105 and 10.106 we conclude that  $\{\text{col}(M, i_j)\}_{j \in \{1, \dots, \text{rank}(L)\}}$  forms a basis of  $\mathcal{S}(\text{cols}(M))$  so that  $\text{rank}(M) = \text{rank}(L)$   $\square$

**Theorem 10.306.** *Let  $X, Y$  be finite dimensionl vector spaces over a field  $F$  with dimensions  $n, m \in \mathbb{N}_0$ ,  $\{e_i\}_{i \in \{1, \dots, n\}}$ ,  $\{f_i\}_{i \in \{1, \dots, m\}}$  basis of  $X, Y$ ,  $\alpha \in F$  and  $L \in \text{Hom}(X, Y)$  then we have  $\mathcal{M}(\alpha \cdot L, \{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{1, \dots, m\}}) = \alpha \cdot \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{1, \dots, m\}})$*

**Proof.**  $\mathcal{M}(\alpha \cdot L, \{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{1, \dots, m\}})_{i,j} = (\alpha \cdot L(e_j))_i = \alpha \cdot (L(e_j)_i) = \alpha \cdot \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{1, \dots, m\}})_{i,j}$   $\square$

**Theorem 10.307.** *Let  $X, Y$  and  $Z$  be vector spaces over a field  $F$  with dimensions  $n, m$  and  $r$ . Let  $\{e_i\}_{i \in \{1, \dots, n\}}$ ,  $\{f_i\}_{i \in \{1, \dots, m\}}$  and  $\{g_i\}_{i \in \{1, \dots, r\}}$  be the basis of  $X, Y$  and  $Z$ ,  $L_1 \in \text{Hom}(X, Y)$  and  $L_2 \in \text{Hom}(Y, Z)$  then  $\mathcal{M}(L_2 \circ L_1, \{e_i\}_{i \in \{1, \dots, n\}}, \{g_i\}_{i \in \{1, \dots, r\}}) = \mathcal{M}(L_2, \{f_i\}_{i \in \{1, \dots, m\}}, \{g_i\}_{i \in \{1, \dots, r\}}) \cdot \mathcal{M}(L_1, \{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{1, \dots, m\}})$*

**Proof.**  $\forall i \in \{1, \dots, n\}$  we have

$$\begin{aligned}
 (L_2 \circ L_1)(e_i) &= L_2(L_1(e_i)) \\
 &= L_2 \left( \sum_{k \in \{1, \dots, m\}} \mathcal{M}((L_1, \{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{0, \dots, m\}})_{k,i} \cdot f_k) \right) \\
 &= \sum_{k \in \{1, \dots, m\}} \mathcal{M}((L_1, \{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{0, \dots, m\}})_{k,i} \cdot L_2(f_k)) \\
 &= \sum_{k \in \{1, \dots, m\}} \mathcal{M}(L_1, \{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{0, \dots, m\}})_{k,i} \cdot \left( \sum_{l \in \{1, \dots, r\}} \mathcal{M}(L_2, \{f_i\}_{i \in \{1, \dots, m\}}, \{g_i\}_{i \in \{1, \dots, r\}})_{l,k} \cdot g_l \right) \\
 &= \sum_{l \in \{1, \dots, r\}} \left( \sum_{k \in \{1, \dots, m\}} \mathcal{M}(L_1, \{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{0, \dots, m\}})_{k,i} \cdot \mathcal{M}(L_2, \{f_i\}_{i \in \{1, \dots, m\}}, \{g_i\}_{i \in \{1, \dots, r\}})_{l,k} \cdot g_l \right) \\
 &= \sum_{l \in \{1, \dots, r\}} \left( \sum_{k \in \{1, \dots, m\}} \mathcal{M}(L_2, \{f_i\}_{i \in \{1, \dots, m\}}, \{g_i\}_{i \in \{1, \dots, r\}})_{l,k} \cdot \mathcal{M}(L_1, \{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{0, \dots, m\}})_{k,i} \right) \cdot g_l \\
 &= \sum_{l \in \{1, \dots, r\}} (\mathcal{M}(L_2, \{f_i\}_{i \in \{1, \dots, m\}}, \{g_i\}_{i \in \{1, \dots, r\}}) \cdot \mathcal{M}(L_1, \{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{0, \dots, m\}}))_{l,i} \cdot g_l
 \end{aligned}$$

By the definition of  $\mathcal{M}$  we have further  $\sum_{l \in \{1, \dots, r\}} \mathcal{M}(L_2 \circ L_1, \{e_i\}_{i \in \{1, \dots, n\}}, \{g_i\}_{i \in \{1, \dots, r\}})_{l,i} \cdot g_l = (L_2 \circ L_1)(e_i) = \sum_{l \in \{1, \dots, r\}} (\mathcal{M}(L_2, \{f_i\}_{i \in \{1, \dots, m\}}, \{g_i\}_{i \in \{1, \dots, r\}}) \cdot \mathcal{M}(L_1, \{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{0, \dots, m\}}))_{l,i} \cdot g_l$  and as the expansion in a basis is unique we have finally

$$\mathcal{M}(L_2 \circ L_1, \{e_i\}_{i \in \{1, \dots, n\}}, \{g_i\}_{i \in \{1, \dots, r\}})_{l,i} = (\mathcal{M}(L_2, \{f_i\}_{i \in \{1, \dots, m\}}, \{g_i\}_{i \in \{1, \dots, r\}}) \cdot \mathcal{M}(L_1, \{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{0, \dots, m\}}))_{l,i}$$

proving the theorem.  $\square$

**Example 10.308.** Let  $n \in \mathbb{N}_0$ ,  $\sigma \in P_n$ ,  $F$  a field and  $\{\varepsilon_i\}_{i \in \{1, \dots, n\}}$  the basis for  $F^n$  then we have for  $\sigma 1_{F^n} \in \text{Hom}(F^n, F^n)$  :

1.  $\mathcal{M}(\sigma 1_{F^n}, \{\varepsilon_i\}_{i \in \{1, \dots, n\}}, \{\varepsilon_i\}_{i \in \{1, \dots, n\}})_{i,j} = \delta_{i,\sigma(j)}$
2. If  $L \in \text{Hom}(F^n, F^m)$  then  $\sigma L = L \circ \sigma 1_{F^n} \in \text{Hom}(F^n, F^m)$

$$3. \mathcal{M}(\sigma L, \{\varepsilon_i\}_{i \in \{1, \dots, n\}}, \{\varepsilon_i\}_{i \in \{1, \dots, m\}}, \{\varepsilon_i\}_{i \in \{1, \dots, m\}})_{i,j} = \mathcal{M}(L \circ \sigma 1_{F^n}, \{\varepsilon_i\}_{i \in \{1, \dots, n\}}, \{\varepsilon_i\}_{i \in \{1, \dots, m\}}, \{\varepsilon_i\}_{i \in \{1, \dots, m\}})_{i,j} = \mathcal{M}(\sigma L, \{\varepsilon_i\}_{i \in \{1, \dots, n\}}, \{\varepsilon_i\}_{i \in \{1, \dots, m\}}, \{\varepsilon_i\}_{i \in \{1, \dots, m\}})_{i, \sigma(j)}$$

**Proof.** First by 10.224 we have that  $\sigma i_F$  and  $\sigma L$  are linear. Next we have

1.  $(\sigma 1_{F^n}(\varepsilon_i))_j \stackrel{\text{think of } \varepsilon_i \text{ as a function of } \{1, \dots, n\} \rightarrow F}{=} (1_F(\varepsilon_i \circ \sigma))_j = (\varepsilon_i \circ \sigma)(j) = \varepsilon_i(\sigma(j)) = (\varepsilon_i)_{\sigma(j)} = \delta_{i, \sigma(j)}$ . Using 10.164 we have that  $\sigma 1_{F^n}(\varepsilon_i) = \sum_{j \in \{1, \dots, n\}} (\sigma 1_{F^n}(\varepsilon_i))_j \cdot \varepsilon_j$  giving that  $\mathcal{M}(\sigma 1_F, \{\varepsilon_i\}_{i \in \{1, \dots, n\}}, \{\varepsilon_i\}_{i \in \{1, \dots, n\}})_{i,j} = \delta_{i, \sigma(j)}$
2. If  $x \in F^n$  (and thus  $x: \{1, \dots, n\} \rightarrow F$ ) then we have  $\sigma L(x) = L(x \circ \sigma) = L(1_{F^n}(x \circ \sigma)) = L(\sigma 1_{F^n}(x)) = (L \circ \sigma 1_{F^n})(x)$  so  $\sigma L = L \circ \sigma 1_{F^n}$
3. Using 10.307 we have that  $\mathcal{M}(L \circ \sigma 1_{F^n}, \{\varepsilon_i\}_{i \in \{1, \dots, n\}}, \{\varepsilon_i\}_{i \in \{1, \dots, m\}})_{k,l} = (\mathcal{M}(L, \{\varepsilon_i\}_{i \in \{1, \dots, n\}}, \{\varepsilon_i\}_{i \in \{1, \dots, m\}}) \cdot \mathcal{M}(1_{F^n}, \{\varepsilon_i\}_{i \in \{1, \dots, n\}}, \{\varepsilon_i\}_{i \in \{1, \dots, n\}}))_{k,l} = \sum_{j \in \{1, \dots, n\}} \mathcal{M}(L, \{\varepsilon_i\}_{i \in \{1, \dots, n\}}, \{\varepsilon_i\}_{i \in \{1, \dots, m\}})_{k,j} \cdot \mathcal{M}(1_{F^n}, \{\varepsilon_i\}_{i \in \{1, \dots, n\}}, \{\varepsilon_i\}_{i \in \{1, \dots, n\}})_{j,l} = \sum_{j \in \{1, \dots, n\}} \mathcal{M}(L, \{\varepsilon_i\}_{i \in \{1, \dots, n\}}, \{\varepsilon_i\}_{i \in \{1, \dots, m\}})_{k,j} \cdot \delta_{j, \sigma(l)} = \mathcal{M}(L, \{\varepsilon_i\}_{i \in \{1, \dots, n\}}, \{\varepsilon_i\}_{i \in \{1, \dots, m\}})_{k, \sigma(l)}$   $\square$

**Theorem 10.309.** Let  $X, Y$  be vector spaces over a field  $F$  with dimensions  $n, m$  and basis  $\{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{1, \dots, m\}}$ . Then  $\mathcal{M}(\{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{1, \dots, m\}}): \text{Hom}(X, Y) \rightarrow M(m \times n, F)$  defined by  $L \mapsto \mathcal{M}(\{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{1, \dots, m\}})(L) = \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{1, \dots, m\}})$  is a bijection. Furthermore if  $M \in M(m \times n, F)$  then  $L = \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{1, \dots, m\}})^{-1}(M)$  is defined by  $x = \sum_{i \in \{1, \dots, n\}} x_i \cdot e_i \mapsto L(x) = \sum_{i \in \{1, \dots, m\}} (\sum_{j \in \{1, \dots, n\}} M_{i,j} \cdot x_j) \cdot f_i$ . In other words we can define a  $L \in \text{Hom}(X, Y)$  by specifying how  $\forall i \in \{1, \dots, n\} L(e_i)$  is written in terms of the basis  $\{f_i\}_{i \in \{1, \dots, m\}}$  as this defines  $\mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{1, \dots, m\}})$ .

**Proof.**

1. **(injectivity)** Let  $L_1, L_2 \in \text{Hom}(X, Y)$  with  $\mathcal{M}(\{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{1, \dots, m\}})(L_1) = \mathcal{M}(\{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{1, \dots, m\}})(L_2)$  then we have that  $\mathcal{M}(L_1, \{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{1, \dots, m\}}) = \mathcal{M}(L_2, \{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{1, \dots, m\}})$ . If now  $x = \sum_{i \in \{1, \dots, n\}} x_i \cdot e_i$  then  $L_1(x) = \sum_{j \in \{1, \dots, m\}} (\sum_{i \in \{1, \dots, n\}} \mathcal{M}(L_1, \{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{1, \dots, m\}})_{i,j} \cdot x_j) \cdot f_j = \sum_{j \in \{1, \dots, m\}} (\sum_{i \in \{1, \dots, n\}} \mathcal{M}(L_2, \{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{1, \dots, m\}})_{i,j} \cdot x_j) \cdot f_j = L_2(x)$  proving that  $L_1 = L_2$
2. **(surjectivity)** Let  $M \in M(m \times n, F)$  and define  $L: X \rightarrow Y$  by  $x = \sum_{i \in \{1, \dots, n\}} x_i \mapsto L(x) = \sum_{i \in \{1, \dots, m\}} (\sum_{j \in \{1, \dots, n\}} M_{i,j} \cdot x_j) \cdot f_i$  then we have:
  - a. **( $L \in \text{Hom}(X, Y)$ )** For if  $\alpha, \beta \in F$ ,  $x = \sum_{i \in \{1, \dots, n\}} x_i \cdot e_i \in X$ ,  $y = \sum_{i \in \{1, \dots, n\}} y_i \cdot e_i \in X$  then  $L(\alpha \cdot x + \beta \cdot y) = \sum_{i \in \{1, \dots, m\}} (\sum_{j \in \{1, \dots, n\}} M_{i,j} \cdot (\alpha \cdot x + \beta \cdot y)_j) \cdot f_i = \sum_{i \in \{1, \dots, m\}} (\sum_{j \in \{1, \dots, n\}} M_{i,j} \cdot (\alpha \cdot x_j + \beta \cdot y_j)) \cdot f_i = \sum_{i \in \{1, \dots, m\}} (\alpha \cdot \sum_{j \in \{1, \dots, n\}} M_{i,j} \cdot x_j + \beta \cdot \sum_{j \in \{1, \dots, n\}} M_{i,j} \cdot y_j) \cdot f_i = \alpha \cdot \sum_{i \in \{1, \dots, m\}} (\sum_{j \in \{1, \dots, n\}} M_{i,j} \cdot x_j) + \beta \cdot \sum_{i \in \{1, \dots, m\}} (\sum_{j \in \{1, \dots, n\}} M_{i,j} \cdot y_j) = \alpha \cdot L(x) + \beta \cdot L(y)$
  - b. **( $\mathcal{M}(\{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{1, \dots, m\}})(L) = M$ )** For by definition  $(\mathcal{M}(\{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{1, \dots, m\}})(L))_{j,i} = (\mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{1, \dots, m\}})(L))_{j,i} = L(e_i)_{j,i} = \sum_{k \in \{1, \dots, n\}} M_{j,k} \cdot (e_i)_k \stackrel{e_i = \sum_{j \in \{1, \dots, m\}} \delta_{i,j} \cdot e_j}{=} \sum_{k \in \{1, \dots, n\}} M_{j,k} \cdot \delta_{i,k} = M_{j,i}$  proving (b).  $\square$

**Definition 10.310.** Let  $n \in \mathbb{N}_0$ ,  $F$  a field then  $M \in M(n \times n, F)$  then a  $N \in M(n \times n, F)$  is called a **inverse** of  $M$  if  $M \cdot N = E = N \cdot M$

The next theorem proves that the inverse is of a matrix is unique

**Theorem 10.311.** Let  $n \in \mathbb{N}_0$ ,  $F$  a field and assume that  $M \in M(n \times n, F)$  has two inverses  $N, N'$  then  $N = N'$

**Proof.** Let  $N, N'$  be two inverses of  $M$  then  $N = N \cdot E = N \cdot (M \cdot N') = (N \cdot M) \cdot N' = E \cdot N' = N'$   $\square$

**Definition 10.312.** Let  $n \in \mathbb{N}_0$ ,  $F$  a field and  $M \in \mathcal{M}(n \times n, F)$  then if  $M$  has a inverse we note the unique inverse of  $M$  by  $M^{-1}$ .

**Theorem 10.313.** Let  $n \in \mathbb{N}_0$ ,  $F$  a field and  $M \in \mathcal{M}(n \times n, F)$  then if  $M$  has a inverse  $M^{-1}$  then  $M^{-1}$  has also a inverse and  $(M^{-1})^{-1} = M$

**Proof.** This is trivial as  $(M^{-1}) \cdot M = E = M \cdot (M^{-1}) \cdot M$  so that  $M$  is a inverse of  $M^{-1}$  hence  $(M^{-1})^{-1} = M$ .  $\square$

**Theorem 10.314.** If  $X, Y$  are finite  $n$ -dimensional vector spaces over a field  $\mathcal{F}$  with basis  $\{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{1, \dots, n\}}$  and  $L \in \text{Hom}(X, Y)$  then we have the following equivalences

1.  $L$  has a inverse  $L^{-1}$  (or equivalent  $L$  is a isomorphism)
2.  $\mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{1, \dots, n\}})$  has a inverse

Further if  $L^{-1}$  exists then  $\mathcal{M}(L^{-1}, \{f_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}}) = \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{1, \dots, n\}})^{-1}$

**Proof.**

**1  $\Rightarrow$  2.** We have

$$\begin{aligned} E &= \mathcal{M}(1_Y, \{f_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{1, \dots, n\}}) \\ &= \mathcal{M}(L \circ L^{-1}, \{f_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{1, \dots, n\}}) \\ &\stackrel{10.307}{=} \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{1, \dots, n\}}) \cdot \mathcal{M}(L^{-1}, \{f_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}}) \end{aligned}$$

and

$$\begin{aligned} E &= \mathcal{M}(1_X, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}}) \\ &= \mathcal{M}(L^{-1} \circ L, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}}) \\ &\stackrel{10.307}{=} \mathcal{M}(L^{-1}, \{f_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}}) \cdot \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{1, \dots, n\}}) \end{aligned}$$

proving that  $\mathcal{M}(L^{-1}, \{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{1, \dots, n\}}) = \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{1, \dots, n\}})^{-1}$  and thus that  $\mathcal{M}(L^{-1}, \{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{1, \dots, n\}})$  has a inverse.

**2  $\Rightarrow$  1.** Define  $K \in \text{Hom}(Y, X)$  by  $K(f_i) = \sum_{j=1}^n \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{1, \dots, n\}})_{i,j}^{-1} \cdot f_j$  so that  $\mathcal{M}(K, \{f_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}}) = \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{1, \dots, n\}})^{-1}$ . Hence we have that

$$\begin{aligned} \mathcal{M}(K \circ L, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}}) &\stackrel{10.307}{=} \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{1, \dots, n\}})^{-1} \cdot \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{1, \dots, n\}}) \\ &= E \end{aligned}$$

proving (see 10.304) that  $K \circ L = 1_X$ . Further we have

$$\begin{aligned} \mathcal{M}(L \circ K, \{f_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{1, \dots, n\}}) &\stackrel{10.307}{=} \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{1, \dots, n\}}) \cdot \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{1, \dots, n\}})^{-1} \\ &= E \end{aligned}$$

proving (see 10.304) that  $L \circ K = 1_X$ .  $\square$

**Theorem 10.315.** Let  $F$  be a field,  $m, n \in \mathbb{N}_0$ ,  $\{\varepsilon_i\}_{i \in \{1, \dots, n\}}, \{\varepsilon_i\}_{i \in \{1, \dots, m\}}$  the canonical basis of  $F^n$ ,  $F^m$  and  $\mathcal{M}(\{\varepsilon_i\}_{i \in \{1, \dots, n\}}, \{\varepsilon_i\}_{i \in \{1, \dots, m\}}): \text{Hom}(F^n, F^m) \rightarrow M(m \times n, F)$  (defined above) then if  $x \in F^n$  we have for  $M \in M(m \times n, F)$  that  $(\mathcal{M}(\{\varepsilon_i\}_{i \in \{1, \dots, n\}}, \{\varepsilon_i\}_{i \in \{1, \dots, m\}})^{-1}(M)(x))_j = \sum_{i \in \{1, \dots, n\}} M_{j,i} \cdot x_i$ . Or if we identify  $F^n$  with  $M(n \times 1, F)$  and  $F^m$  with  $M(m \times 1, F)$  that  $\mathcal{M}(\{\varepsilon_i\}_{i \in \{1, \dots, n\}}, \{\varepsilon_i\}_{i \in \{1, \dots, m\}})^{-1}(M)(x) = M \cdot x$

**Proof.** If  $x \in F^n$  (so  $x: \{1, \dots, n\} \rightarrow F$ ) then (see 10.164)  $x = \sum_{i \in \{1, \dots, n\}} x_i \cdot \varepsilon_i$ , and by definition (see 10.309) we have  $\mathcal{M}(\{\varepsilon_i\}_{i \in \{1, \dots, n\}}, \{\varepsilon_i\}_{i \in \{1, \dots, m\}})^{-1}(M)(x) = \sum_{i \in \{1, \dots, m\}} (\sum_{j \in \{1, \dots, n\}} M_{i,j} x_j) \cdot \varepsilon_i$  so that  $(\mathcal{M}(\{\varepsilon_i\}_{i \in \{1, \dots, n\}}, \{\varepsilon_i\}_{i \in \{1, \dots, m\}})^{-1}(M)(x))_k = (\sum_{i \in \{1, \dots, m\}} (\sum_{j \in \{1, \dots, n\}} M_{i,j} x_j) \cdot \varepsilon_i)_k = \sum_{i \in \{1, \dots, m\}} (\sum_{j \in \{1, \dots, n\}} M_{i,j} x_j) \cdot (\varepsilon_i)_k = \sum_{i \in \{1, \dots, m\}} (\sum_{j \in \{1, \dots, n\}} M_{i,j} x_j) \cdot \delta_{i,k} = \sum_{j \in \{1, \dots, n\}} M_{k,j} \cdot x_j$  proving the theorem.  $\square$

**Definition 10.316.** If  $n \in \mathbb{N}_0$ ,  $F$  a field and  $M \in M(n \times n, F)$  then  $\det(M) = \sum_{\sigma \in P_n} \text{sign}(\sigma) \cdot (\prod_{i \in \{1, \dots, n\}} M_{\sigma(i), i})$ . Another notation that we use is  $|M| = \det(M)$

**Example 10.317.** If  $n = 1$  then  $M = (M_{1,1})$  and  $P_n = \{1_{\{1\}}\}$  and  $|M| = \sum_{\sigma \in \{1_{\{1\}}\}} (\prod_{i \in \{1\}} M_{i, \sigma(i)}) = M_{1,1}$

**Example 10.318.** If  $n \in \mathbb{N}_0$ ,  $F$  a field and  $M \in M(n \times n, F)$  a diagonal matrix meaning that  $\forall i, j \in \{1, \dots, n\}$  we have  $M_{i,j} = \delta_{i,j} \cdot M_{i,i}$  then  $\det(M) = \prod_{i=1}^n M_{i,i}$

**Proof.** First if  $\sigma \in P_n$  is such that  $\sigma \neq 1_{\{1, \dots, n\}}$  then  $\exists i_0 \in \{1, \dots, n\}$  such that  $\sigma(i_0) \neq i_0 + 0$  hence  $\prod_{i \in \{1, \dots, n\}} M_{\sigma(i), i} = (\prod_{i \in \{1, \dots, n\} \setminus \{i_0\}} M_{\sigma(i_0), i_0 + 0}) \cdot M_{\sigma(i_0), i_0 + 0} = 0$ . So  $\det(M) = \sum_{\sigma \in P_n} (\prod_{i \in \{1, \dots, n\}} M_{\sigma(i), i}) = \left( \sum_{\sigma \in P_n \setminus \{1_{\{1, \dots, n\}}\}} (\prod_{i \in \{1, \dots, n\}} M_{\sigma(i), i}) \right) + \prod_{i \in \{1, \dots, n\}} M_{1_{\{1, \dots, n\}}(i), i} = 0 + \prod_{i \in \{1, \dots, n\}} M_{i,i} = \prod_{i \in \{1, \dots, n\}} M_{i,i}$   $\square$

**Theorem 10.319.** Let  $F$  be a field,  $n \in \mathbb{N}_0$  and  $M \in M(n \times n, F)$  then  $\det(M) = \sum_{\sigma \in P_n} \text{sign}(\sigma) \cdot (\prod_{i \in \{1, \dots, n\}} M_{i, \sigma(i)})$

**Proof.** First note that from the commutativity of multiplication we have

$$\prod_{i \in \{1, \dots, n\}} M_{\sigma(i), i} \stackrel{10.26 \text{ and } \sigma \text{ is a bijection}}{=} \prod_{i \in \{1, \dots, n\}} M_{\sigma(\sigma^{-1}(i)), \sigma^{-1}(i)} = \prod_{i \in \{1, \dots, n\}} M_{i, \sigma^{-1}(i)}$$

from which it follows that

$$\begin{aligned} \det(M) &= \sum_{\sigma \in P_n} \text{sign}(\sigma) \cdot \left( \prod_{i \in \{1, \dots, n\}} M_{i, \sigma^{-1}(i)} \right) \\ &\stackrel{10.94}{=} \sum_{\sigma \in P_n} \text{sign}(\sigma^{-1}) \cdot \left( \prod_{i \in \{1, \dots, n\}} M_{i, \sigma^{-1}(i)} \right) \\ &\stackrel{-1 \text{ is a bijection from } P_n \rightarrow P_n \text{ and 10.43}}{=} \sum_{\tau \in P_n} \text{sign}(\tau) \cdot \left( \prod_{i \in \{1, \dots, n\}} M_{i, \tau(i)} \right) \end{aligned}$$

$\square$

**Corollary 10.320.** Let  $F$  be a field,  $n \in \mathbb{N}_0$  and  $M \in M(n \times n, F)$  then  $\det(M) = \det(M^T)$

**Proof.**  $\det(M) = \sum_{\sigma \in P_n} \text{sign}(\sigma) \cdot (\prod_{i \in \{1, \dots, n\}} M_{\sigma(i), i}) \stackrel{\text{previous theorem}}{=} \sum_{\sigma \in P_n} \text{sign}(\sigma) \cdot (\prod_{i \in \{1, \dots, n\}} M_{i, \sigma(i)}) = \sum_{\sigma \in P_n} \text{sign}(\sigma) \cdot (M^T)_{\sigma(i), i} = \det(M^T)$   $\square$

**Theorem 10.321.** If  $X$  is a  $n$ -dimensional space over a field  $F$  of characterization 0 with basis  $\{e_i\}_{i \in \{1, \dots, n\}}$  then if  $L \in \text{Hom}(X, X)$  we have  $\det(L) = \det(\mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}}))$

**Proof.** Let  $\Delta$  be a determinant function such that  $\Delta(e_1, \dots, e_n) = 1$  (see 10.265) then we have  $\Delta_L(e_1, \dots, e_n) \stackrel{\text{definition } \Delta_L}{=} \Delta(L(e_1), \dots, L(e_n)) \stackrel{10.276}{=} \det(L) \cdot \Delta(e_1, \dots, e_n) = \det(L)$  so we have

$$\Delta_L(e_1, \dots, e_n) = \det(L) \tag{10.107}$$

As a determinant function is a skew symmetric linear mapping from  $X^n \rightarrow F$  we have (using  $M = \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}})$ )

$$\begin{aligned}
 \Delta_L(e_1, \dots, e_n) &= \Delta(L(e_1), \dots, L(e_n)) \\
 &= \Delta\left(\sum_{i \in \{1, \dots, n\}} M_{i,1} \cdot e_i, \dots, \sum_{i \in \{1, \dots, n\}} M_{i,n} \cdot e_i\right) \\
 &\stackrel{10.262}{=} \sum_{\sigma \in P_n} \left( \left( \prod_{i \in \{1, \dots, n\}} M_{\sigma(i),i} \right) \cdot \sigma \Delta(e_1, \dots, e_n) \right) \\
 &\stackrel{\Delta \text{ is skewsymmetric}}{=} \left( \sum_{\sigma \in P_n} \text{sign}(\sigma) \cdot \prod_{i \in \{1, \dots, n\}} M_{\sigma(i),i} \right) \cdot \Delta(e_1, \dots, e_n) \\
 &= \det(M) \cdot \Delta(e_1, \dots, e_n) \\
 &= \det(M)
 \end{aligned}$$

Using 10.107 we have finally  $\det(L) = \det(M)$  □

**Theorem 10.322.** *If  $F$  is a field of characterization zero,  $n \in \mathbb{N}_0$  and  $M_1, M_2 \in M(n \times n, F)$  then*

$$\begin{aligned}
 \det(M_1 \cdot M_2) &= \det(M_1) \cdot \det(M_2) \\
 \det(E) &= 1
 \end{aligned}$$

**Proof.** Let  $F^n$  be the canonical vector space over  $F$  with the canonical basis  $\{\varepsilon_i\}_{i \in \{1, \dots, n\}}$  and let  $\mathcal{M} = \mathcal{M}(\{\varepsilon_i\}_{i \in \{1, \dots, n\}}, \{\varepsilon_i\}_{i \in \{1, \dots, n\}}): \text{Hom}(F^n, F^n) \rightarrow M(n \times n, F)$  be the bijection from 10.309. Take then  $L_1 = \mathcal{M}^{-1}(M_1)$ ,  $L_2 = \mathcal{M}^{-1}(M_2)$  then we have as  $\mathcal{M}(L_1 \circ L_2) \stackrel{10.307}{=} \mathcal{M}(L_1) \cdot \mathcal{M}(L_2) = \mathcal{M}(\mathcal{M}^{-1}(M_1)) \cdot \mathcal{M}(\mathcal{M}^{-1}(M_2)) = M_1 \cdot M_2$

$$\begin{aligned}
 \det(M_1 \cdot M_2) &= \det(\mathcal{M}(L_1 \circ L_2)) \\
 &\stackrel{10.321}{=} \det(L_1 \circ L_2) \\
 &\stackrel{10.282}{=} \det(L_1) \cdot \det(L_2) \\
 &\stackrel{10.321}{=} \det(M_1) \cdot \det(M_2)
 \end{aligned}$$

For the last part  $\mathcal{M}(1_{F^n})_{i,j} = (1_{F^n}(\varepsilon_j))_i = (\varepsilon_j)_i = \delta_{i,j} = E_{i,j}$  so that we have

$$\begin{aligned}
 \det(E) &= \det(\mathcal{M}(1_{F^n})) \\
 &\stackrel{10.321}{=} \det(1_{F^n}) \\
 &\stackrel{10.277}{=} 1
 \end{aligned}$$
□

**Theorem 10.323.** *Let  $n \in \mathbb{N}_0$ ,  $F$  a field of characterization zero,  $M \in M(n \times n, F)$  and  $\{\varepsilon_i\}_{i \in \{1, \dots, n\}}$  be the canonical basis of  $F^n$  then if  $\Delta$  is a determinant function in  $F^n$  such that  $\Delta(\varepsilon_1, \dots, \varepsilon_n) = 1$  then  $\det(M) = \Delta(\text{col}(M, 1), \dots, \text{col}(M, n)) = \Delta(\text{row}(M, 1), \dots, \text{row}(M, n))$*

**Proof.** Take  $\mathcal{M} = \mathcal{M}(\{\varepsilon_i\}_{i \in \{1, \dots, n\}}, \{\varepsilon_i\}_{i \in \{1, \dots, n\}}): \text{Hom}(F^n, F^n) \rightarrow M(n \times n, F)$  be the bijection from 10.309. If  $M \in M(n \times n, F)$  define then  $L = \mathcal{M}^{-1}(M)$  then we have by 10.315 that  $(L(\varepsilon_i))_j = \sum_{k \in \{1, \dots, n\}} M_{j,k} \cdot (\varepsilon_i)_k = \sum_{k \in \{1, \dots, n\}} M_{j,k} \cdot \delta_{i,k} = M_{j,i} = \text{col}(M, i)_j$  so that

$$L(\varepsilon_i) = \text{col}(M, i) \tag{10.108}$$

So we have

$$\begin{aligned}
 \Delta(\text{col}(M, 1), \dots, \text{col}(M, n)) &= \Delta(L(\varepsilon_1), \dots, L(\varepsilon_n)) \\
 &= \Delta_L(\varepsilon_1, \dots, \varepsilon_n) \\
 &= \det(L) \cdot \Delta(\varepsilon_1, \dots, \varepsilon_n) \\
 &= \det(L) \\
 &= \det(\mathcal{M}(L)) \\
 &= \det(\mathcal{M}(\mathcal{M}^{-1}(L))) \\
 &= \det(M)
 \end{aligned}$$

This proves that  $\det(M) = \Delta(\text{col}(M, 1), \dots, \text{col}(M, n))$ . For the rest note that  $\text{col}(M^T, i)_j = (M^T)_{j,i} = M_{i,j} = \text{row}(M, i)_j$  giving

$$\begin{aligned}\Delta(\text{row}(M, 1), \dots, \text{row}(M, n)) &= \Delta(\text{col}(M^T, 1), \dots, \text{col}(M^T, n)) \\ &= \det(M^T) \\ &\stackrel{10.320}{=} \det(M) \\ &\square\end{aligned}$$

**Definition 10.324.** If  $F$  is a field of characterization zero,  $M \in M(n \times n, F)$  and  $\sigma: \{1, \dots, n\} \rightarrow \{1, \dots, n\}$  then we define

1.  $M_\sigma \in M(n \times n, F)$  by  $\forall k, l \in \{1, \dots, n\}$   $(M_\sigma)_{k,l} = M_{\sigma(k),l}$
2.  $M^\sigma \in M(n \times n, F)$  by  $\forall k, l \in \{1, \dots, n\}$   $(M^\sigma)_{k,l} = M_{k,\sigma(l)}$

Essential the original matrices with rows or columns permuted.

**Corollary 10.325.** If  $F$  is a field of characterization zero,  $n \in \mathbb{N}_0$  and  $M \in M(n \times n, F)$  then if  $\sigma: \{1, \dots, n\} \rightarrow \{1, \dots, n\}$  is a permutation then we have  $\det(M_\sigma) = \text{sign}(\sigma) \cdot \det(M)$  and  $\det(M^\sigma) = \text{sign}(\sigma) \cdot \det(M)$

**Proof.**  $\forall i \in \{1, \dots, n\}$  we have  $\text{row}(M_\sigma, i)_j = (M_\sigma)_{i,j} = M_{\sigma(i),j} = \text{row}(M, \sigma(i))$ , then  $\det(M_\sigma) = \Delta(\text{row}(M, \sigma(1)), \dots, \text{row}(M, \sigma(n))) = \text{sign}(\sigma) \cdot \Delta(\text{row}(M, 1), \dots, \text{row}(M, n)) = \text{sign}(\sigma) \cdot \det(M)$ .

$\forall i \in \{1, \dots, n\}$  we have  $\text{col}(M^\sigma, i)_j = (M^\sigma)_{j,i} = M_{j,\sigma(i)} = \text{col}(M, \sigma(i))$ , then  $\det(M^\sigma) = \Delta(\text{col}(M^\sigma, 1), \dots, \text{col}(M^\sigma, n)) = \text{sign}(\sigma) \cdot \Delta(\text{col}(M, 1), \dots, \text{col}(M, n)) = \text{sign}(\sigma) \cdot \det(M)$ .  $\square$

**Corollary 10.326.** Let  $n \in \mathbb{N}_0$ ,  $F$  a field of characterization zero then if  $M \in M(n \times n, F)$  we have

$$\det(M) = 0 \Leftrightarrow \text{rank}(M) < n$$

**Proof.** This follows from  $\det(M) = \Delta(\text{row}(M, 1), \dots, \text{row}(M, n)) = \Delta(\text{col}(M, 1), \dots, \text{col}(M, n))$  and 10.259 (skew symmetric mappings an linear Independence) and the definition of the rank of a matrix (see 10.301).  $\square$

**Definition 10.327.** Let  $n \in \mathbb{N}_0$ ,  $X$  a set,  $a \in X$  and  $x \in X^n$  define  $[+a]x \in X^{n+1}$  by  $\forall i \in \{1, \dots, n+1\}$   $([+a]x)_i = \begin{cases} a & \text{if } i = 1 \\ x_{i-1} & \text{if } i > 1 \end{cases}$ . If  $x = (x_1, \dots, x_n)$  then we sometimes write  $[+a]x$  as  $(a, x_1, \dots, x_n)$

**Lemma 10.328.** Let  $n \in \mathbb{N}_0$ ,  $F$  a field and  $0 \in F$  the neutral element in  $F$ , then  $[+0]: F^n \rightarrow F^{n+1}$  defined by  $x \rightarrow [+0]x$  is a linear mapping.

**Proof.** Let  $\alpha, \beta \in F$  and  $x, y \in F^n$  then for  $i \in \{1, \dots, n\}$  we have  $([+0](\alpha \cdot x + \beta \cdot y))_i = \begin{cases} 0 & \text{if } i = 1 \\ (\alpha \cdot x + \beta \cdot y)_{i-1} & \text{if } i > 1 \end{cases} = \begin{cases} \alpha \cdot 0 + \beta \cdot 0 & \text{if } i = 1 \\ \alpha \cdot x_{i-1} + \beta \cdot y_{i-1} & \text{if } i > 1 \end{cases} = \alpha \cdot [+0]x + \beta \cdot [+0]y$ .  $\square$

**Lemma 10.329.** Let  $n \in \mathbb{N}_0 \setminus \{1\}$ ,  $F$  a field of characterization zero,  $\{\varepsilon_i\}_{i \in \{1, \dots, n\}}$  the canonical basis of  $F^n$  then if  $\Delta: (F^n)^n \rightarrow F$  is a determinant function then  $\Delta_{-1}: (F^{n-1})^{n-1} \rightarrow F$  defined by  $(a_1, \dots, a_{n-1}) \rightarrow \Delta_{-1}(a_1, \dots, a_{n-1}) = \Delta(\varepsilon_1, [+0]a_1, \dots, [+0]a_{n-1})$  is a determinant function. Further if  $\Delta(\varepsilon_1, \dots, \varepsilon_n) = 1$  then  $\Delta_{-1}(\varepsilon_1, \dots, \varepsilon_{n-1}) = 1$

**Proof.** We prove first that  $\Delta_{-1}$  is a determinant function

1. **(multilinearity)** If  $\alpha, \beta \in F$ ,  $x, y \in F^{n-1}$ ,  $i \in \{1, \dots, n-1\}$  and  $a = (a_1, \dots, a_{i-1}, \alpha \cdot x + \beta \cdot y, a_{i+1}, \dots, a_{n-1}) \in (F^{n-1})^{n-1}$  then  $\Delta_{-1}(a_1, \dots, a_{i-1}, \alpha \cdot x + \beta \cdot y, a_{i+1}, \dots, a_{n-1}) = \Delta(\varepsilon_1, [+0]a_1, \dots, [+0]a_{i-1}, [+0]a_i, \dots, [+0]a_{n-1})$   $\stackrel{\text{previous lemma}}{=} \Delta(\varepsilon_1, [+0]a_1, \dots, [+0]a_{i-1}, \alpha \cdot [+0]x + \beta \cdot [+0]y, [+0]a_{i+1}, \dots, [+0]a_{n-1})$   $\stackrel{\Delta \text{ is multilinear}}{=} \alpha \cdot \Delta(\varepsilon_1, [+0]a_1, \dots, [+0]a_{i-1}, [+0]x, [+0]a_{i+1}, \dots, [+0]a_{n-1}) + \beta \cdot (\varepsilon_1, [+0]a_1, \dots, [+0]a_{i-1}, [+0]y, [+0]a_{i+1}, \dots, [+0]a_{n-1}) = \alpha \cdot \Delta_{-1}(a_1, \dots, a_{i-1}, x, a_{i+1}, \dots, a_{n-1}) + \beta \cdot \Delta_{-1}(a_1, \dots, a_{i-1}, y, a_{i+1}, \dots, a_{n-1})$
2. **(skew symmetry)** If  $\{a_1, \dots, a_{n-1}\}$  is such that there exists a  $i \neq j$  with  $a_i = a_j \Rightarrow [+0]a_i = [+0]a_j$  so that  $\Delta_{-1}(a_1, \dots, a_{n-1}) = \Delta(\varepsilon_1, [+0]a_1, \dots, [+0]a_{n-1}) \stackrel{10.259}{=} 0$  which by 10.259 again means that  $\Delta_{-1}$  is skew symmetric.

Next if  $\Delta(\varepsilon_1, \dots, \varepsilon_n) = 1$  and as for  $i, j \in \{1, \dots, n-1\}$  we have  $([+0]\varepsilon_i)_j = \begin{cases} 0 & \text{if } j=1 \\ (\varepsilon_i)_{j-1} & \text{if } 1 < j \end{cases} = \delta_{i+1,j} = (\varepsilon_{i+1})_j$  so that  $[+0]\varepsilon_i = \varepsilon_{i+1}$  and thus  $\Delta_{-1}(\varepsilon_1, \dots, \varepsilon_{n-1}) = \Delta(\varepsilon_1, [+0]\varepsilon_1, \dots, [+0]\varepsilon_{n-1}) = \Delta(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n) = 1$   $\square$

**Definition 10.330.** Let  $F$  be a field,  $n \in \mathbb{N}_0 \setminus \{1\}$  then we define for  $i, j \in \{1, \dots, n\}$   $[i \boxplus j] : M(n \times n, F) \rightarrow M(n-1 \times n-1, F)$  by  $M \rightarrow [i \boxplus j]M$  where

$$\begin{aligned} ([i \boxplus j]M)_{k,l} &= M_{k,l} \text{ if } k < i \wedge l < j \\ &= M_{k+1,l} \text{ if } i \leq k \leq n-1 \wedge l < j \\ &= M_{k,l+1} \text{ if } k < i \wedge j \leq l \leq n-1 \\ &= M_{k+1,l+1} \text{ if } i \leq k \leq n-1 \wedge j \leq l \leq n-1 \end{aligned}$$

**Lemma 10.331.** Let  $F$  be a field,  $n \in \mathbb{N}_0 \setminus \{1\}$  then for  $i, j \in \{1, \dots, n\}$  and  $M \in M(n \times n, F)$  we have  $[i \boxplus j]M^T = ([j \boxplus i]M)^T$

**Proof.** We have for  $k, l \in \{1, \dots, n-1\}$  the following cases to consider

$$\begin{aligned} k < i, l < k \Rightarrow ([i \boxplus j]M^T)_{k,l} &= M_{k,l}^T \\ &= M_{l,k} \\ &= ([j \boxplus i]M)_{l,k} \\ &= ([j \boxplus i]M)_{k,l}^T \\ i \leq k \leq n-1, l < j \Rightarrow ([i \boxplus j]M^T)_{k,l} &= M_{k+1,l}^T \\ &= M_{l,k+1} \\ &= ([j \boxplus i]M)_{l,k} \\ &= ([j \boxplus i]M)_{k,l}^T \\ k < i, j \leq l \leq n-1 \Rightarrow ([i \boxplus j]M^T)_{k,l} &= M_{k,l+1}^T \\ &= M_{l+1,k} \\ &= ([j \boxplus i]M)_{l,k} \\ &= ([j \boxplus i]M)_{k,l}^T \\ i \leq k \leq n-1, j \leq l \leq n-1 \Rightarrow ([i \boxplus j]M^T)_{k,l} &= M_{k+1,l+1}^T \\ &= M_{l+1,k+1} \\ &= ([j \boxplus i]M)_{l,k} \\ &= ([j \boxplus i]M)_{k,l}^T \end{aligned}$$

$\square$

**Lemma 10.332.** Let  $n \in \mathbb{N}_0 \setminus \{1\}$  and let  $M \in M(n \times n, F)$  such that  $\forall i \in \{1, \dots, n\}$  we have  $\text{row}(M, 1)_i = \delta_{i,1} = \text{col}(M, 1)_i$  then  $\det(M) = \det([1 \boxplus 1M])$

**Proof.** Consider now the following cases for  $i \in \{1, \dots, n\}$  and  $\text{row}(M, i)$

1. ( $i = 1$ ) then  $\forall j \in \{1, \dots, n\} \models \text{row}(M, i)_j = \delta_{i,1} = (\varepsilon_1)_i$
2. ( $i \neq 1$ ) then if  $j = 1$  we have  $\text{row}(M, i)_j = \text{row}(M, i)_1 = \text{col}(M, 1)_i = \delta_{i,1} \underset{i \neq 1}{=} 0$  and if  $j \in \{2, \dots, n\}$  we have  $\text{row}(M, i)_j = M_{i,j} = ([1 \boxplus 1]M)_{i-1,j-1}$  so that we have that  $\text{row}(M, i) = [+0]\text{row}([1 \boxplus 1]M, i-1)$ . Let now  $\Delta$  be a determinant function in  $F^n$  such that  $\Delta(\varepsilon_1, \dots, \varepsilon_n)$  then we have by 10.323 that :

$$\begin{aligned} \det(M) &= \Delta(\text{row}(M, 1), \dots, \text{row}(M, n)) \\ &= \Delta(\varepsilon_1, \text{row}(M, 2), \dots, \text{row}(M, n)) \\ &= \Delta(\varepsilon_1, \text{row}([1 \boxplus 1]M, 1), \dots, \text{row}([1 \boxplus 1]M, n-1)) \\ &\stackrel{10.329}{=} \Delta(\text{row}([1 \boxplus 1]M), \dots, \text{row}([1 \boxplus 1]M, n-1)) \\ &\stackrel{10.329 \text{ and } 10.323}{=} \det([1 \boxplus 1M]) \end{aligned}$$

$\square$

**Lemma 10.333.** Let  $n \in \mathbb{N}_0 \setminus \{1\}$  and let  $M \in M(n \times n, F)$  such that  $\exists j \in \{1, \dots, n\}$  such that  $\forall i \in \{1, \dots, n\}$  we have  $\text{row}(M, 1)_i = \delta_{j,i}$  and  $\text{col}(M, j)_i = \delta_{i,1}$  then  $\det(M) = (-1)^{1+j} \cdot \det([1 \boxplus j]M)$

**Proof.** We have the following cases to consider for  $j \in \{1, \dots, m\}$  :

1. ( $j = 1$ ) then  $\forall i \in \{1, \dots, n\}$  we have  $\text{col}(M, j)_i = \delta_{i,1} = \text{row}(M, 1)_i$  so we can use the previous lemma and find that  $\det(M) = \det([1 \boxplus 1]M) = (-1)^{1+1} \det([1 \boxplus 1]M) = (-1)^{1+1} \cdot \det([1 \boxplus 1]M)$
2. ( $j \in \{2, \dots, n\}$ ) define then  $M'$  by  $\forall k, l \in \{1, \dots, n\}$   $M'_{k,l} = \text{col}(M', l)_k = \text{col}(M, j \rightsquigarrow_n 1(l))_k = M_{j \rightsquigarrow_n 1(l)}$  then we have  $\forall k \in \{1, \dots, n\}$  that

$$\begin{aligned}
 \text{row}(M', 1)_k &= \text{col}(M', k)_1 \\
 &= \text{col}(M, j \rightsquigarrow_n 1(k))_1 \\
 &= \begin{cases} \text{col}(M, j)_1 \text{ if } k = 1 \\ \text{col}(M, k-1)_1 \text{ if } 1 < k \leq j \\ \text{col}(M, k)_1 \text{ if } j < k \end{cases} \\
 &= \begin{cases} \delta_{1,1} \text{ if } k = 1 \\ \text{row}(M, 1)_{k-1} \text{ if } 1 < k \leq j \\ \text{row}(M, 1)_k \text{ if } j < k \end{cases} \\
 &= \begin{cases} 1 \text{ if } k = 1 \\ \delta_{j,k-1} \text{ if } 1 < k \leq j \Rightarrow k-1 \neq j \\ \delta_{j,k} \text{ if } j < k \end{cases} \\
 &= \begin{cases} 1 \text{ if } k = 1 \\ 0 \text{ if } 1 < k \leq j \\ 0 \text{ if } j < k \end{cases} \\
 &= \delta_{1,k}
 \end{aligned}$$

and we have also

$$\begin{aligned}
 \text{col}(M', 1)_k &= \text{col}(M, j \rightsquigarrow_n 1(1))_k \\
 &= \text{col}(M, j)_k \\
 &= \delta_{1,k}
 \end{aligned}$$

So we have  $\forall k \in \{1, \dots, n\}$  that  $\text{row}(M', 1)_k = \delta_{1,k} = \text{col}(M', 1)_k$  so by the previous lemma (see 10.332) we have

$$\det(M') = \det([1 \boxplus 1]M') \quad (10.109)$$

Now for  $k, l \in \{1, \dots, n-1\}$  we have as  $1 \leq k, l \leq n-1$  that  $([1 \boxplus 1]M')_{k,l} = M'_{k+1,l+1}$  and we must look at the following possible cases for  $l$  where  $1 \leq k \leq n-1$

a. ( $1 \leq k \leq n-1 \wedge l < j$ ) then

$$\begin{aligned}
 M'_{k+1,l+1} &= \text{col}(M', l+1)_{k+1} \\
 &= \text{col}(M', j \rightsquigarrow_n 1(l+1))_{k+1} \\
 &\stackrel{l < j \Rightarrow 1 < l+1 \leq j}{=} \text{col}(M, l)_{k+1} \\
 &= M_{k+1,l} \\
 &\stackrel{1 \leq k < k+1, l < j}{=} ([1 \boxplus j]M)_{k,l}
 \end{aligned}$$

b.  $(1 \leq k \leq n-1 \wedge j \leq l \leq n-1)$  then

$$\begin{aligned}
 M'_{k+1, l+1} &= \text{col}(M', l+1)_{k+1} \\
 &= \text{col}(M, j \rightsquigarrow_n 1(l+1))_{k+1} \\
 &\stackrel{j \leq l \Rightarrow j < l+1}{=} \text{col}(M, l+1)_{k+1} \\
 &= M_{k+1, l+1} \\
 &\stackrel{1 \leq k < k+1, j < l+1}{=} ([1 \boxplus j]M)_{k, l}
 \end{aligned}$$

The above means that  $[1 \boxplus 1]M' = [1 \boxplus j]M$  and given 10.109 we have then

$$\det(M') = \det([1 \boxplus j]M) \quad (10.110)$$

Now take a determinant function  $\Delta$  in  $F^n$  such that  $\Delta(\varepsilon_1, \dots, \varepsilon_n) = 1$  then by 10.323 we have

$$\begin{aligned}
 \det(M') &= \Delta(\text{col}(M', 1), \dots, \text{col}(M', n)) \\
 &= \Delta(\text{col}(M, j \rightsquigarrow_n 1(1)), \dots, \text{col}(M, j \rightsquigarrow_n 1(n))) \\
 &\stackrel{10.98}{=} (-1)^{j-1} \Delta(\text{col}(M, 1), \dots, \text{col}(M, n)) \\
 &\stackrel{10.323}{=} (-1)^{j-1} \det(M)
 \end{aligned}$$

as  $\frac{1}{(-1)^{j-1}} = (-1)^{j-1} = (-1)^2 \cdot (-1)^{j-1} = (-1)^{j+1}$  we have  $(-1)^{j+1} \cdot \det(M') = \det(M)$  and using 10.110 this gives finally  $\det(M) = (-1)^{j+1} \cdot \det([1 \boxplus j]M)$   $\square$

**Theorem 10.334.** Let  $n \in \mathbb{N}_0 \setminus \{1\}$  and for  $M \in M(n \times n, F)$  there exists a  $i, j \in \{1, \dots, n\}$  such that  $\text{row}(M, i)_k = \delta_{j,k}$  and  $\text{col}(M, j)_k = \delta_{i,k}$  then  $\det(M) = (-1)^{i+j} \det([i \boxplus j]M)$

**Proof.** We have the following cases to consider for  $i \in \{1, \dots, n\}$

- ( $i = 1$ )** then  $\forall k \in \{1, \dots, n\}$  we have  $\text{row}(M, 1)_k = \delta_{j,k}$  and  $\text{col}(M, j)_k = \delta_{1,k}$  so using the previous lemma 10.333 we have that  $\det(M) = (-1)^{j+1} \det([1 \boxplus j]M) \stackrel{i=1}{=} (-1)^{i+j} \det([i \boxplus j]M)$
- ( $i \in \{2, \dots, n\}$ )** define then  $M'$  by  $\forall k, l \in \{1, \dots, n\}$   $M'_{k,l} = \text{row}(M', k)_l = \text{row}(M, i \rightsquigarrow_n 1(k))_l = M_{i \rightsquigarrow_n 1(k), l}$  then we have  $\forall k \in \{1, \dots, n\}$  that

$$\begin{aligned}
 \text{row}(M', 1)_k &= \text{row}(M, i \rightsquigarrow 1(1))_k \\
 &= \text{row}(M, i)_k \\
 &= \delta_{j,k}
 \end{aligned}$$

and

$$\begin{aligned}
 \text{col}(M', j)_k &= \text{row}(M', k)_j \\
 &= \text{row}(M, i \rightsquigarrow_n 1(k))_j \\
 &= \begin{cases} \text{row}(M, i)_j \text{ if } k = 1 \\ \text{row}(M, k-1)_j \text{ if } 1 < k \leq i \\ \text{row}(M, k)_j \text{ if } i < k \end{cases} \\
 &= \begin{cases} \delta_{j,j} \text{ if } k = 1 \\ \text{col}(M, j)_{k-1} \text{ if } 1 < k \leq i \\ \text{col}(M, j)_k \text{ if } i < k \end{cases} \\
 &= \begin{cases} 1 \text{ if } k = 1 \\ \delta_{i,k-1} \text{ if } 1 < k \leq i \Rightarrow k-1 \neq i \\ \delta_{i,k} \text{ if } i < k \end{cases} \\
 &= \begin{cases} 1 \text{ if } k = 1 \\ 0 \text{ if } 1 < k \leq i \\ 0 \text{ if } i < k \end{cases} \\
 &= \delta_{1,k}
 \end{aligned}$$

By using 10.333 we have then that

$$\det(M') = (-1)^{j+1} \det([1 \boxplus j]M') \quad (10.111)$$

Now if  $k, l \in \{1, \dots, n-1\}$  then we have for  $([1 \boxplus j]M')_{k,l}$  the following cases to consider for  $k, l$  :

a.  $(k < i \wedge l < j)$  then

$$\begin{aligned} ([1 \boxplus j]M')_{k,l} &\underset{1 \leq k \wedge l < j}{=} M'_{k+1,l} \\ &= \text{row}(M', k+1)_l \\ &= \text{row}(M, i \rightsquigarrow_n 1(k+1))_l \\ &\underset{1 < k+1 \leq i}{=} \text{row}(M, k)_l \\ &= M_{k,l} \\ &\underset{k < i \wedge l < j}{=} ([i \boxplus j]M)_{k,l} \end{aligned}$$

b.  $(k < i \wedge j \leq l \leq n-1)$  then

$$\begin{aligned} ([1 \boxplus j]M')_{k,l} &\underset{1 \leq k \wedge j \leq l \leq n-1}{=} M'_{k+1,l+1} \\ &= \text{row}(M', k+1)_{l+1} \\ &= \text{row}(M, i \rightsquigarrow_n 1(k+1))_{l+1} \\ &\underset{1 < k+1 \leq i}{=} \text{row}(M, k)_{l+1} \\ &= M_{k,l+1} \\ &\underset{k < i \wedge j \leq l \leq n-1}{=} ([i \boxplus j]M)_{k,l} \end{aligned}$$

c.  $(i \leq k \leq n-1 \wedge l < j)$  then

$$\begin{aligned} ([1 \boxplus j]M')_{k,l} &\underset{1 \leq k \wedge l < j}{=} M'_{k+1,l} \\ &= \text{row}(M', k+1)_l \\ &= \text{row}(M, i \rightsquigarrow_n 1(k+1))_l \\ &\underset{i < k+1}{=} \text{row}(M, k+1)_l \\ &= M_{k+1,l} \\ &\underset{i \leq k \leq n-1 \wedge l < j}{=} ([i \boxplus j]M)_{k,l} \end{aligned}$$

d.  $(i \leq k \leq n-1 \wedge j \leq l \leq n-1)$  then

$$\begin{aligned} ([1 \boxplus j]M')_{k,l} &\underset{1 \leq k \wedge j \leq l \leq n-1}{=} M'_{k+1,l+1} \\ &= \text{row}(M', k+1)_{l+1} \\ &= \text{row}(M, i \rightsquigarrow_n 1(k+1))_{l+1} \\ &\underset{i < k+1}{=} \text{row}(M, k+1)_{l+1} \\ &= M_{k+1,l+1} \\ &\underset{i \leq k \leq n-1 \wedge j \leq l \leq n-1}{=} ([i \boxplus j]M)_{k,l} \end{aligned}$$

So we have that  $[1 \boxplus j]M' = [i \boxplus j]M$  and thus by 10.111 we have

$$\det(M') = (-1)^{j+1} \det([i \boxplus j]M) \quad (10.112)$$

Now if  $\Delta$  is a determinant function in  $F^n$  such that  $\Delta(\varepsilon_1, \dots, \varepsilon_n) = 1$  we have by 10.323 that

$$\begin{aligned} \det(M') &= \Delta(\text{row}(M', 1), \dots, \text{row}(M', n)) \\ &= \Delta(\text{row}(M, i \rightsquigarrow_n 1(1)), \dots, \text{row}(M, i \rightsquigarrow_n 1(n))) \\ &\underset{\text{skewsymmetry and 10.98}}{=} (-1)^{i-1} \cdot \Delta(\text{row}(M, 1), \dots, \text{row}(M, n)) \\ &\underset{10.323}{=} (-1)^{i-1} \cdot \det(M) \end{aligned}$$

Using the fact that  $\frac{1}{(-1)^{-1}} = (-1)^{i-1} = (-1)^2 \cdot (-1)^{i-1} = (-1)^{i+1}$  we have  $\det(M) = (-1)^{i+1} \det(M')$  and this together with 10.112 gives finally  $\det(M) = (-1)^{i+1} (-1)^{j+1} \det([i \boxplus j]M)$  or  $\det(M) = (-1)^{i+j} \det([i \boxplus j]M)$  what he have to prove.  $\square$

**Definition 10.335.** Let  $F$  be a field,  $n \in \mathbb{N}_0 \setminus \{1\}$  and  $M \in M(n \times n, F)$  then we define  $[< m]M \in M(m-1 \times m-1, F)$  and  $[> m]M \in M(n-m \times n-m, F)$  by

$$\begin{aligned} 1 < m \leq n &\Rightarrow ([< m]M)_{k,l} = M_{k,l} \text{ where } k, l \in \{1, \dots, m-1\} \\ 1 \leq m < n &\Rightarrow ([> m]M)_{k,l} = M_{k+m, l+m} \text{ where } k, l \in \{1, \dots, n-m\} \end{aligned}$$

**Lemma 10.336.** Let  $F$  be a field,  $n \in \mathbb{N}_0 \setminus \{1\}$  then we have for  $M \in M(n \times n, F)$  then the following holds:

1.  $[1 \boxplus 1]M = [> 1]M$
2.  $[n \boxplus n]M = [< n]M$
3. If  $1 < m_1 \leq n$  and  $1 < m_2 \leq m_1 - 1$  then  $[< m_2]([< m_1]M) = [< m_2]M$
4. If  $1 \leq m_1 < n$  and  $1 \leq m_2 < n - m_1$  then  $[> m_2]([> m_1]M) = [> m_1 + m_2]M$
5. If  $1 < m \leq n$  then  $([< m]M)^T = [< m]M^T$
6. If  $1 \leq m < n$  then  $([> m]M)^T = [> m]M^T$

**Proof.**

1.  $\forall i, j \in \{1, \dots, n-1\}$  we have

$$\begin{aligned} ([1 \boxplus 1]M)_{i,j} &\stackrel{1 \leq i, j}{=} M_{i+1, j+1} \\ &= ([> 1]M)_{i,j} \end{aligned}$$

so we have  $[1 \boxplus 1]M = [> 1]M$

2.  $\forall i, j \in \{1, \dots, n-1\}$  we have

$$\begin{aligned} ([n \boxplus n]M)_{i,j} &\stackrel{i, j \leq n-1 < n}{=} M_{i,j} \\ &\stackrel{1 \leq i, j \leq n-1}{=} ([< n]M)_{i,j} \end{aligned}$$

so we have  $[n \boxplus n]M = [< n]M$

3. If  $1 \leq i, j \leq m_2 - 1$  then we have

$$\begin{aligned} ([< m_2]([< m_1]M))_{i,j} &\stackrel{1 \leq i, j \leq m_2-1}{=} ([< m_1]M)_{i,j} \\ &\stackrel{1 \leq i, j \leq m_2-1 < m_2 \leq m_1-1}{=} M_{i,j} \\ &\stackrel{1 \leq i, j \leq m_2-1}{=} ([< m_2]M)_{i,j} \end{aligned}$$

so we have  $[< m_2]([< m_1]M) = [< m_2]M$

4. First note that  $[> m_1]M \in M(n-m_1, n-m_1, F) \Rightarrow [> m_2]([> m_1]M) \in M(n-m_1-m_2 \times n-m_1-m_2, F) = M(n-(m_1+m_2) \times n-(m_1+m_2), F) \ni [> m_1+m_2]M$ , so the dimensions are ok. Next if  $1 \leq i, j \leq n-(m_1+m_2)$  then we have

$$\begin{aligned} ([> m_2]([> m_1]M))_{i,j} &= ([> m_1]M)_{i+m_2, j+m_2} \\ &= M_{i+m_2+m_1, j+m_2+m_1} \\ &= M_{i+(m_1+m_2), j+(m_1+m_2)} \\ &= ([> m_1+m_2]M)_{i,j} \end{aligned}$$

so  $[> m_2]([> m_1]M) = [> m_1+m_2]M$

5. For  $1 \leq i, j \leq m-1$  we have

$$\begin{aligned} ([< m]M)_{i,j}^T &= ([< m]M)_{j,i} \\ &= M_{j,i} \\ &= M_{i,j}^T \\ &= ([< m]M^T)_{i,j} \end{aligned}$$

$$\text{so } ([<m]M)^T = [<m]M^T$$

6. For  $1 \leq i, j \leq n - m$  we have

$$\begin{aligned} ([>m]M)_{i,j}^T &= ([>m]M)_{j,i} \\ &= M_{j+m, i+m} \\ &= M_{i+m, j+m}^T \\ &= ([>m]M^T)_{i,j} \end{aligned}$$

$$\text{so } ([>m]M)^T = [>m]M^T \quad \square$$

**Definition 10.337.** Given  $n, m \in \mathbb{N}_0$ ,  $F$  a field and  $M \in M(n \times m, F)$  then if  $i \in \{1, \dots, n\}$  and  $j \in \{1, \dots, m\}$  we define  $[i \oplus j]M \in M(n \times m, F)$  by  $\forall k \in \{1, \dots, n\}, \forall l \in \{1, \dots, m\}$  we have

$$([i \oplus j]M)_{k,l} = \begin{cases} M_{k,l} & \text{if } l \neq i \\ \delta_{k,j} & \text{if } l = i \end{cases}. \text{ Essentially we replace the } i\text{-the column by the column } \begin{pmatrix} \delta_{1,j} \\ \dots \\ \delta_{n,j} \end{pmatrix}$$

**Definition 10.338.** Given  $n \in \mathbb{N}_0$ ,  $F$  a field and  $M \in M(n \times n, F)$  then  $\text{adjoint}(M) \in M(n \times n, F)$  is defined by  $\text{adjoint}(M)_{i,j} = \det([i \oplus j]M)$ .  $\text{adjoint}(M)$  is called the adjoint of the matrix  $M$  and  $\text{adjoint}(M)_{i,j} = \det([i \oplus j]M)$  is the  $i, j$ -the co-factor of  $M$ .

**Example 10.339.** If  $n = 1$  then  $\text{adjoint}(M) = (\text{adjoint}(M)_{1,1}) = (\det([1 \oplus 1]M)) = (\det(1)) = (1)$

**Theorem 10.340.** If  $X$  is a  $n$ -dimensional vector space ( $n \in \mathbb{N}_0$ ) over a field  $F$  of a characterization zero,  $L \in \text{Hom}(X, X)$  a linear transformation and  $\{e_i\}_{i \in \{1, \dots, n\}}$  a basis of  $X$  then  $\mathcal{M}(\text{adjoint}(L), \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}}) = \text{adjoint}(\mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}}))$ . Essentially the matrix of the adjoint of a linear transformation is the adjoint of the matrix of a linear transformation.

**Proof.** Let  $\Delta$  be a determinant function in  $X$  such that  $\Delta(e_1, \dots, e_n) = 1$  then we have

$$\begin{aligned} \sum_{k \in \{1, \dots, n\}} \mathcal{M}(\text{adjoint}(L), \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}})_{k,j} \cdot e_k &\stackrel{\text{definition of } \mathcal{M}}{=} \text{adjoint}(L)(e_j) \\ &\stackrel{\Delta(e_1, \dots, e_n) = 1}{=} \Delta(e_1, \dots, e_n) \cdot \text{adjoint}(L)(e_j) \\ &\stackrel{10.287}{=} \check{\Delta}_L(e_1, \dots, e_n)(e_j) \\ &= \sum_{k \in \{1, \dots, n\}} (-1)^{k-1} \cdot \Delta(e_j, L(e_1), \dots, \\ &\quad L(e_{k-1}), L(e_{k+1}), \dots, L(e_n)) \cdot e_k \end{aligned}$$

From the uniqueness of a expansion in a basis we have then that

$$\mathcal{M}(\text{adjoint}(L), \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}})_{k,j} = (-1)^{k-1} \Delta(e_j, L(e_1), \dots, L(e_{k-1}), L(e_{k+1}), \dots, L(e_n)) \quad (10.113)$$

Define now  $\{b_i^{k,j}\}_{i \in \{1, \dots, n\}}$  by  $b_i^{j,k} = \begin{cases} L(e_i) & \text{if } i \neq k \\ e_j & \text{if } i = k \end{cases}$  then we have that

$$\begin{aligned} \Delta(e_j, L(e_1), \dots, L(e_{k-1}), L(e_{k+1}), \dots, L(e_n)) &= \Delta(b_k^{k,j}, b_1^{k,j}, \dots, b_{k-1}^{k,j}, b_{k+1}^{k,j}, \dots, b_n^{k,j}) \\ &= \Delta(b_{1 \rightsquigarrow n k(1)}^{k,j}, b_{1 \rightsquigarrow n k(2)}^{k,j}, \dots, b_{1 \rightsquigarrow n k(k)}^{k,j}, b_{1 \rightsquigarrow n k(k+1)}^{k,j}, \dots, \\ &\quad b_{1 \rightsquigarrow n (n)}^{k,j}) \\ &= (1 \rightsquigarrow n k) \Delta(b_1^{k,j}, \dots, b_n^{k,j}) \\ &= (-1)^{k-1} \Delta(b_1^{k,j}, \dots, b_n^{k,j}) \end{aligned}$$

Using the above and 10.113 we have

$$\mathcal{M}(\text{adjoint}(L), \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}})_{k,j} = \Delta(b_1^{k,j}, \dots, b_n^{k,j}) \quad (10.114)$$

Define now

Define now  $\forall i, j \in \{1, \dots, n\}$  the matrix  $M^{i,j} \in M(n \times n, F)$  by  $\forall r, s \in \{1, \dots, n\}$  we have

$$(M^{i,j})_{r,s} = \begin{cases} L(e_s)_r \text{ if } s \neq i \\ \delta_{j,r} \text{ if } s = i \end{cases} \quad (10.115)$$

so that  $(M^{k,j})_{r,s} = \begin{cases} (b_s^{k,j})_r \text{ if } s \neq k \\ (e_j)_r \text{ if } s = k \end{cases} = \begin{cases} (b_s^{k,j})_r \text{ if } s \neq k \\ (b_s^{k,j})_r \text{ if } s = k \end{cases} = (b_s^{k,j})_r$ . If we take now the linear mapping  $L^{i,j}: X \rightarrow X$  by  $L^{i,j} = \mathcal{M}(\{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}})^{-1}(M^{i,j})$  then by 10.309 and the fact that  $e_t = \sum_{l \in \{1, \dots, n\}} \delta_{l,t} \cdot e_l$  we have that

$$\begin{aligned} L^{i,j}(e_t) &= \sum_{r \in \{1, \dots, n\}} \left( \sum_{s \in \{1, \dots, n\}} (M^{i,j})_{r,s} \cdot \delta_{t,s} \right) \cdot e_r \\ &= \sum_{r \in \{1, \dots, n\}} (M^{i,j})_{r,t} \cdot e_r \end{aligned}$$

If now  $i = k$  then

$$\begin{aligned} L^{k,j}(e_i) &= \sum_{r \in \{1, \dots, n\}} (M^{i,j})_{r,i} \cdot e_r \\ &= \sum_{r \in \{1, \dots, n\}} (b_i^{k,j})_r \cdot e_r \\ &= b_i^{k,j} \end{aligned}$$

This together with 10.114 gives

$$\begin{aligned} \mathcal{M}(\text{adjoint}(L), \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}})_{k,j} &= \Delta(L^{k,j}(e_1), \dots, L^{k,j}(e_n)) \\ &= \det(L_{k,j}) \cdot \Delta(e_1, \dots, e_n) \\ &= \det(L_{k,j}) \\ &\stackrel{10.321}{=} \det(M^{k,j}) \end{aligned}$$

so we have

$$\mathcal{M}(\text{adjoint}(L), \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}})_{k,j} = \det(M^{k,j}) \quad (10.116)$$

Now for  $r, s \in \{1, \dots, n\}$  we have the following cases to consider for  $(M^{i,j})_{r,s}$

1. ( $s \neq i$ ) then

$$\begin{aligned} (M^{i,j})_{r,s} &= L(e_s)_r \\ &= \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}})_{r,s} \\ &= ([i \oplus j] \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}})) \end{aligned}$$

2. ( $s = k$ ) then

$$\begin{aligned} (M^{i,j})_{r,s} &= \delta_{j,r} \\ &= ([i \oplus j] \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}}))_{r,s} \end{aligned}$$

So that  $M^{i,j} = [i \oplus j] \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}})$ . Using this and 10.116 gives us  $\mathcal{M}(\text{adjoint}(L), \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}})_{k,j} = \det([k \oplus j] \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}}))$  definition of adjoint of a matrix  $\stackrel{\text{adjoint}}{=} \text{adjoint}(\mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}}))_{k,j}$  or finally

$$\mathcal{M}(\text{adjoint}(L), \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}}) = \text{adjoint}(\mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}})) \quad \square$$

**Theorem 10.341.** *If  $X$  is a  $n$ -dimensional vector space ( $n \in \mathbb{N}_0$ ) over a field  $F$  of characterization 0,  $L \in \text{Hom}(X, X)$  and  $\{e_i\}_{i \in \{1, \dots, n\}}$  a basis of  $X$  then we have the following*

1.  $\mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}}) \cdot \text{adjoint}(\mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}})) = \det(\mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}})) \cdot E$
2.  $\text{adjoint}(\mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}})) \cdot \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}}) = \det(\mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}})) \cdot E$

or if we write out the matrix products directly we have

1.  $\sum_{k \in \{1, \dots, n\}} \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}})_{i,k} \cdot \text{adjoint}(\mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}}))_{k,j} = \det(\mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}})) \cdot \delta_{i,j}$
2.  $\sum_{k \in \{1, \dots, n\}} \text{adjoint}(\mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}}))_{i,k} \cdot \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}})_{k,j} = \det(\mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}})) \cdot \delta_{i,j}$

**Proof.**

1. Using 10.289 we have that  $L \circ \text{adjoint}(L) = \det(L) \cdot 1_X$  so that  $\mathcal{M}(L \circ \text{adjoint}(L), \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}}) = \mathcal{M}(\det(L) \cdot 1_X, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}}) \xrightarrow{10.307} \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}}) \cdot \mathcal{M}(\text{adjoint}(L), \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}}) = \mathcal{M}(\det(L) \cdot 1_X, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}}) \xrightarrow{10.306 \text{ and } 10.304} \det(L) \cdot E \xrightarrow{10.321 \text{ and } 10.340} \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}}) \cdot \text{adjoint}(\mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}})) = \det(\mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}})) \cdot E$
2. Using 10.289 we have that  $\text{adjoint}(L) \circ L = \det(L) \cdot 1_X$  so that  $\mathcal{M}(\text{adjoint}(L) \circ L, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}}) = \mathcal{M}(\det(L) \cdot 1_X, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}}) \xrightarrow{10.307} \mathcal{M}(\text{adjoint}(L), \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}}) \cdot \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}}) = \mathcal{M}(\det(L) \cdot 1_X, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}}) \xrightarrow{10.306 \text{ and } 10.304} \det(L) \cdot E \xrightarrow{10.321 \text{ and } 10.340} \text{adjoint}(\mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}})) \cdot \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}}) = \det(\mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}})) \cdot E \quad \square$

The above theorem gives the reformulation of 10.289 in matrix form, the next theorem works just for matrices alone (which is made possible by the bijection between  $M(m \times n, F)$  and  $\text{Hom}(F^n, F^m)$  (see 10.309 and 10.315).

**Theorem 10.342.** *If  $F$  is a field of characteristic zero then if  $n \in \mathbb{N}_0$  we have for  $M \in M(n \times n, F)$  that*

1.  $M \cdot \text{adjoint}(M) = \det(M) \cdot E$  (or  $\sum_{k \in \{1, \dots, n\}} M_{i,k} \cdot \text{adjoint}(M)_{k,j} = \det(M) \cdot \delta_{i,j}$ )
2.  $\text{adjoint}(M) \cdot M = \det(M) \cdot E$  (or  $\sum_{k \in \{1, \dots, n\}} \text{adjoint}(M)_{i,k} \cdot M_{k,j} = \det(M) \cdot \delta_{i,j}$ )
3. If  $n > 1$  then  $\text{adjoint}(M)_{i,j} = (-1)^{i+j} \cdot \det([j \boxplus i]M)$
4. If  $n > 1$  then  $\det(M) = \sum_{i \in \{1, \dots, n\}} (-1)^{i+j} M_{i,j} \cdot \det([i \boxplus j]M)$  where  $j \in \{1, \dots, n\}$  (this is the expansion of the determinant with respect to the  $j$ -the column).
5. If  $n > 1$  then  $\det(M) = \sum_{j \in \{1, \dots, n\}} (-1)^{i+j} M_{i,j} \cdot \det([i \boxplus j]M)$  where  $i \in \{1, \dots, n\}$  (this is the expansion of the determinant with respect to the  $i$ -the row)
6.  $\det(E) = 1$
7. If  $M \in M(1 \times 1, F)$  then  $\text{adjoint}(M) = (1)$  and  $\det(M) = M_{1,1}$

**Proof.**

1. Let  $\mathcal{M}(\{\varepsilon_i\}_{i \in \{1, \dots, n\}}, \{\varepsilon_i\}_{i \in \{1, \dots, n\}}): \text{Hom}(F^n, F^n) \rightarrow M(n \times n, F)$  be the bijection defined in 10.309. If  $M \in M(n \times n, F)$  take then  $L = \mathcal{M}(\{\varepsilon_i\}_{i \in \{1, \dots, n\}}, \{\varepsilon_i\}_{i \in \{1, \dots, n\}})^{-1}(M)$  then we have by 10.341 that  $\mathcal{M}(L, \{\varepsilon_i\}_{i \in \{1, \dots, n\}}, \{\varepsilon_i\}_{i \in \{1, \dots, n\}}) \cdot \text{adjoint}(\mathcal{M}(L, \{\varepsilon_i\}_{i \in \{1, \dots, n\}}, \{\varepsilon_i\}_{i \in \{1, \dots, n\}})) = \det(\mathcal{M}(L, \{\varepsilon_i\}_{i \in \{1, \dots, n\}}, \{\varepsilon_i\}_{i \in \{1, \dots, n\}})) \cdot E$  which as  $M = \mathcal{M}(\{\varepsilon_i\}_{i \in \{1, \dots, n\}}, \{\varepsilon_i\}_{i \in \{1, \dots, n\}})(L) = \mathcal{M}(L, \{\varepsilon_i\}_{i \in \{1, \dots, n\}}, \{\varepsilon_i\}_{i \in \{1, \dots, n\}})$  means that  $M \cdot \text{adjoint}(M) = \det(M) \cdot E$ .
2. Let  $\mathcal{M}(\{\varepsilon_i\}_{i \in \{1, \dots, n\}}, \{\varepsilon_i\}_{i \in \{1, \dots, n\}}): \text{Hom}(F^n, F^n) \rightarrow M(n \times n, F)$  be the bijection defined in 10.309. If  $M \in M(n \times n, F)$  take then  $L = \mathcal{M}(\{\varepsilon_i\}_{i \in \{1, \dots, n\}}, \{\varepsilon_i\}_{i \in \{1, \dots, n\}})^{-1}(M)$  then we have by 10.341 that  $\text{adjoint}(\mathcal{M}(L, \{\varepsilon_i\}_{i \in \{1, \dots, n\}}, \{\varepsilon_i\}_{i \in \{1, \dots, n\}})) \cdot \mathcal{M}(L, \{\varepsilon_i\}_{i \in \{1, \dots, n\}}, \{\varepsilon_i\}_{i \in \{1, \dots, n\}}) = \det(\mathcal{M}(L, \{\varepsilon_i\}_{i \in \{1, \dots, n\}}, \{\varepsilon_i\}_{i \in \{1, \dots, n\}})) \cdot E$  which as  $M = \mathcal{M}(\{\varepsilon_i\}_{i \in \{1, \dots, n\}}, \{\varepsilon_i\}_{i \in \{1, \dots, n\}})(L) = \mathcal{M}(L, \{\varepsilon_i\}_{i \in \{1, \dots, n\}}, \{\varepsilon_i\}_{i \in \{1, \dots, n\}})$  means that  $\text{adjoint}(M) \cdot M = \det(M) \cdot E$ .

3. Note first that if  $r, s \in \{1, \dots, n\}$  then we have

$$\begin{aligned}
 \det([r \oplus s]M) &= \det(([r \oplus s]M)^T) \\
 &= \det(([r \oplus s]M)^T) \cdot \delta_{r,r} \\
 &\stackrel{\text{see (1)}}{=} \sum_{k \in \{1, \dots, n\}} ([r \oplus s]M)_{r,k}^T \cdot \text{adjoint}(([r \oplus s]M)^T)_{k,r} \\
 &= \sum_{k \in \{1, \dots, n\}} ([r \oplus s]M)_{k,r} \cdot \text{adjoint}(([r \oplus s]M)^T)_{k,r} \\
 &\stackrel{([r \oplus s]M)_{k,r} = \delta_{k,s}}{=} \sum_{k \in \{1, \dots, n\}} \delta_{k,s} \cdot \text{adjoint}(([r \oplus s]M)^T)_{k,r} \\
 &= \text{adjoint}(([r \oplus s]M)^T)_{s,r} \\
 &\stackrel{\text{definition of adjoint}}{=} \det([s \oplus r]([r \oplus s]M)^T)
 \end{aligned}$$

this gives

$$\det([r \oplus s]M) = \det([s \oplus r]([r \oplus s]M)^T) \quad (10.117)$$

If now  $j \in \{1, \dots, n\}$  then we have

$$\begin{aligned}
 \text{row}([s \oplus r]([r \oplus s]M)^T, r)_j &= ([s \oplus r]([r \oplus s]M)^T)_{r,j} \\
 &= \begin{cases} \delta_{r,r} = 1 \text{ if } j = s \\ ([r \oplus s]M)_{r,j}^T \text{ if } j \neq s \end{cases} \\
 &= \begin{cases} 1 \text{ if } j = s \\ ([r \oplus s]M)_{j,r} \text{ if } j \neq s \end{cases} \\
 &= \begin{cases} 1 \text{ if } j = s \\ \delta_{j,s} \text{ if } j \neq s \end{cases} \\
 &= \begin{cases} 1 \text{ if } j = s \\ 0 \text{ if } j \neq s \end{cases} \\
 &= \delta_{j,s}
 \end{aligned}$$

giving

$$\text{row}([s \oplus r]([r \oplus s]M)^T, r)_j = \delta_{j,s} \quad (10.118)$$

Also we have that

$$\begin{aligned}
 \text{col}([s \oplus r]([r \oplus s]M)^T, s)_j &= ([s \oplus r]([r \oplus s]M)^T)_{j,s} \\
 &= \delta_{j,r}
 \end{aligned}$$

giving

$$\text{col}([s \oplus r]([r \oplus s]M)^T, s)_j = \delta_{j,r} \quad (10.119)$$

Using 10.118, 10.119 and 10.334 we have then

$$\det([s \oplus r]([r \oplus s]M)^T) = (-1)^{r+s} \det([r \boxplus s]([s \oplus r]([r \oplus s]M)^T)) \quad (10.120)$$

Consider now the following cases for  $i, j \in \{1, \dots, n-1\}$  in  $([r \boxplus s]([s \oplus r]([r \oplus s]M)^T))_{i,j}$ :

a.  $(i < r \wedge j < s)$  then

$$\begin{aligned}
 ([r \boxplus s]([s \oplus r]([r \oplus s]M)^T))_{i,j} &= (([s \oplus r]([r \oplus s]M)^T))_{i,j} \\
 &\stackrel{j < s \Rightarrow j \neq s}{=} ([r \oplus s]M)_{i,j}^T \\
 &= ([r \oplus s]M)_{j,i} \\
 &\stackrel{j < s \Rightarrow j \neq s}{=} M_{j,i} \\
 &= M_{i,j}^T \\
 &\stackrel{i < r \wedge j < s}{=} ([r \boxplus s]M^T)_{i,j}
 \end{aligned}$$

b.  $(r \leq i \leq n-1 \wedge j < s)$  then

$$\begin{aligned}
 ([r \boxplus s]([s \oplus r]([r \oplus s]M)^T))_{i,j} &= (([s \oplus r]([r \oplus s]M)^T))_{i+1,j} \\
 &\stackrel{j < s \Rightarrow j \neq s}{=} ([r \oplus s]M)_{i+1,j}^T \\
 &= ([r \boxplus s]M)_{j,i+1} \\
 &\stackrel{r \leq i \Rightarrow r \neq i+1}{=} M_{j,i+1} \\
 &= M_{i+1,j}^T \\
 &\stackrel{r \leq i < n-1, j < s}{=} ([r \boxplus s]M^T)_{i,j}
 \end{aligned}$$

c.  $(i < r \wedge s \leq j \leq n-1)$  then

$$\begin{aligned}
 ([r \boxplus s]([s \oplus r]([r \oplus s]M)^T))_{i,j} &= (([s \oplus r]([r \oplus s]M)^T))_{i,j+1} \\
 &\stackrel{s \leq j \Rightarrow s \neq j+1}{=} ([r \oplus s]M)_{i,j+1}^T \\
 &= ([r \boxplus s]M)_{j+1,i} \\
 &\stackrel{i < r \Rightarrow i \neq r}{=} M_{j+1,i} \\
 &= M_{i,j+1}^T \\
 &\stackrel{i < r \wedge s \leq j \leq n-1}{=} ([r \boxplus s]M^T)_{i,j}
 \end{aligned}$$

d.  $(r \leq i \leq n-1, s \leq j \leq n-1)$  then

$$\begin{aligned}
 ([r \boxplus s]([s \oplus r]([r \oplus s]M)^T))_{i,j} &= (([s \oplus r]([r \oplus s]M)^T))_{i+1,j+1} \\
 &\stackrel{s \leq j \Rightarrow s \neq j+1}{=} ([r \oplus s]M)_{i+1,j+1}^T \\
 &= ([r \oplus s]M)_{j+1,i+1} \\
 &\stackrel{r \leq i \Rightarrow r \neq i+1}{=} M_{j+1,i+1} \\
 &= M_{i+1,j+1}^T \\
 &= ([r \boxplus s]M^T)_{i,j}
 \end{aligned}$$

Which means that  $([r \boxplus s]([s \oplus r]([r \oplus s]M)^T)) = ([r \boxplus s]M^T) \stackrel{10.331}{=} ([s \oplus r]M)^T$ , This together with  $\text{adjoint}(M)_{r,s} = \det([r \oplus s]M^T)$ , 10.117 and 10.120 gives that

$$\text{adjoint}(M)_{r,s} = (-1)^{r+s} \det(([s \oplus r]M)^T) \stackrel{10.320}{=} (-1)^{r+s} \det([s \oplus r]M)$$

4. If  $i \in \{1, \dots, n\}$  then we have

$$\begin{aligned}
 \det(M) &= \det(M) \cdot \delta_{i,i} \\
 &\stackrel{\text{see (2)}}{=} \sum_{k \in \{1, \dots, n\}} \text{adjoint}(M)_{i,k} \cdot M_{k,i} \\
 &\stackrel{\text{see (3)}}{=} \sum_{k \in \{1, \dots, n\}} (-1)^{i+k} \det([k \boxplus i]M) \cdot M_{k,i} \\
 &= \sum_{k \in \{1, \dots, n\}} (-1)^{i+k} M_{k,i} \cdot \det([k \boxplus i]M)
 \end{aligned}$$

5. If  $i \in \{1, \dots, n\}$  then we have

$$\begin{aligned}
 \det(M) &= \det(M) \cdot \delta_{i,i} \\
 &\stackrel{\text{see (1)}}{=} \sum_{k \in \{1, \dots, n\}} M_{i,k} \cdot \text{adjoint}(M)_{k,i} \\
 &\stackrel{\text{see (3)}}{=} \sum_{k \in \{1, \dots, n\}} (-1)^{i+k} M_{i,k} \cdot \det([i \boxplus k]M)
 \end{aligned}$$

6. This follows from 10.322.

7. This follows from 10.339 and 10.317 □

Using the above theorem we can calculate the determinant and adjoint of every matrix.

**Example 10.343.** Using (4) from the above theorem and  $i = 1$  then

$$\begin{aligned} \begin{vmatrix} -2 & 2 & 3 \\ -1 & 1 & 3 \\ 2 & 0 & 1 \end{vmatrix} &= (-1)^{1+1}M_{1,1}\det([1 \boxplus 1]M) + (-1)^{1+2}M_{2,1}\det([2 \boxplus 1]M) + (-1)^{1+3}M_{3,1}\det([3 \boxplus 1]M) \\ &= -2 \cdot \begin{vmatrix} 1 & 3 \\ 0 & 1 \end{vmatrix} + (-1) \cdot (-1) \cdot \begin{vmatrix} 2 & 3 \\ 0 & 1 \end{vmatrix} + 2 \cdot \begin{vmatrix} 2 & 3 \\ 1 & 3 \end{vmatrix} \\ &= -2 \cdot (1 \cdot |1| + (-1) \cdot 0 \cdot |3|) + (2 \cdot |1| + (-1) \cdot 0 \cdot |3|) + 2 \cdot (2 \cdot |3| + (-1) \cdot 1 \cdot |3|) \\ &= -2 + 2 + 2 \cdot 3 = 6 \end{aligned}$$

**Example 10.344.** Using (3) from the above we have

$$\begin{aligned} \text{adjoint}\left(\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}\right) &= \begin{pmatrix} 1 \cdot \begin{vmatrix} 5 & 6 \\ 8 & 9 \end{vmatrix} & -1 \cdot \begin{vmatrix} 2 & 3 \\ 8 & 9 \end{vmatrix} & 1 \cdot \begin{vmatrix} 2 & 3 \\ 5 & 6 \end{vmatrix} \\ -1 \cdot \begin{vmatrix} 4 & 6 \\ 7 & 9 \end{vmatrix} & 1 \cdot \begin{vmatrix} 1 & 3 \\ 7 & 9 \end{vmatrix} & -1 \cdot \begin{vmatrix} 1 & 3 \\ 4 & 6 \end{vmatrix} \\ 1 \cdot \begin{vmatrix} 4 & 5 \\ 7 & 8 \end{vmatrix} & -1 \cdot \begin{vmatrix} 1 & 2 \\ 7 & 8 \end{vmatrix} & 1 \cdot \begin{vmatrix} 1 & 2 \\ 4 & 5 \end{vmatrix} \end{pmatrix} \\ &= \begin{pmatrix} -3 & 6 & -3 \\ 6 & -12 & 6 \\ -3 & 6 & -3 \end{pmatrix} \end{aligned}$$

The following corollary allows us to calculate in a much faster way the determinant for matrices where some of the leading (trailing) rows (columns) are unit rows (columns).

**Corollary 10.345.** If  $F$  is a field of characterization zero then if  $n \in \mathbb{N}_0$  we have the following for  $M \in M(n \times n, F)$

1. If  $\exists m \in \mathbb{N}_0 \vdash 1 \leq m < n$  such that  $\forall i \vdash i \leq m$  we have  $\text{col}(M, i)_j = \delta_{i,j}$  then  $\det(M) = \det([>m]M)$
2. If  $\exists m \in \mathbb{N}_0 \vdash 1 < m \leq n$  such that  $\forall i \vdash m \leq i \leq n$  we have  $\text{col}(M, i)_j = \delta_{i,j}$  then  $\det(M) = \det([<m]M)$
3. If  $\exists m \in \mathbb{N}_0 \vdash 1 \leq m < n$  such that  $\forall i \vdash i \leq m$  we have  $\text{row}(M, i)_j = \delta_{i,j}$  then  $\det(M) = \det([>m]M)$
4. If  $\exists m \in \mathbb{N}_0 \vdash 1 < m \leq n$  such that  $\forall i \vdash m \leq i \leq n$  we have  $\text{row}(M, i)_j = \delta_{i,j}$  then  $\det(M) = \det([<m]M)$

**Proof.**

1. We use induction on  $m$  to prove this, so let  $E_n = \{m \in \mathbb{N}_0 \mid \text{if } 1 \leq m < n \text{ and } \forall i \in \mathbb{N}_0 \vdash i \leq m \text{ we have } \text{col}(M, i)_j = \delta_{i,j} \text{ then } \det(M) = \det([>m]M)\}$  then we have:

- a. If  $m = 1$  then by the previous theorem we have

$$\begin{aligned} \det(M) &= \sum_{i \in \{1, \dots, n\}} (-1)^{i+1}M_{i,1} \cdot \det([i \boxplus 1]M) \\ &= \sum_{i \in \{1, \dots, n\}} (-1)^{i+1}\text{col}(M, 1)_i \cdot \det([i \boxplus 1]M) \\ &= \sum_{i \in \{1, \dots, n\}} (-1)^{i+1} \cdot \delta_{1,i} \det([i \boxplus 1]M) \\ &= (-1)^{1+1} \det([1 \boxplus 1]M) = \det([1 \boxplus 1]M) \\ &= \det([>1]M) \end{aligned}$$

so we have  $1 \in E_n$

- b. If  $m \in E_n$  then if  $m + 1$  is such that  $1 \leq m + 1 < n$  and  $\forall i \vdash i \leq m + 1$  we have  $\text{col}(M, i)_j = \delta_{i,j}$ . We have then as  $m \in E_n \subseteq \mathbb{N}_0 \Rightarrow 1 \leq m < m + 1 < n$  and  $\forall i \vdash i \leq m$  we have  $\text{col}(M, i) = \delta_{i,j}$  so as  $m \in E_n$  we have

$$\det(M) = \det([>m]M). \quad (10.121)$$

Now if  $j \in \{1, \dots, n-m\}$  then  $\text{col}([\geq m]M, 1)_j = ([\geq m]M)_{j,1} = M_{j+m, 1+m} = \text{col}(M, 1+m)_{j+m} = \delta_{m+1, j+m}$ . Using the previous theorem we have then

$$\begin{aligned}
 \det([\geq m]M) &= \sum_{i \in \{1, \dots, n-m\}} (-1)^{i+1} ([\geq m]M)_{i,1} \cdot \det([i \boxplus 1]([\geq m]M)) \\
 &= \sum_{i \in \{1, \dots, n-m\}} (-1)^{i+1} \text{col}([\geq m]M, 1)_i \cdot \det([i \boxplus 1]([\geq m]M)) \\
 &= \sum_{i \in \{1, \dots, n-m\}} (-1)^{i+1} \delta_{m+1, i+m} \cdot \det([i \boxplus 1]([\geq m]M)) \\
 &\stackrel{m+1=i+m \Rightarrow i=1}{=} (-1)^2 \cdot \det([1 \boxplus 1]([\geq m]M)) \\
 &\stackrel{10.336}{=} \det([\geq 1]([\geq m]M)) \\
 &\stackrel{10.336}{=} \det([\geq m+1]M)
 \end{aligned}$$

Using the above with 10.121 gives  $\det(M) = \det([\geq m+1]M)$  and thus  $m+1 \in E_n$ .

Using mathematical induction we have then  $E_n = \mathbb{N}_0$  proving (1).

2. We use induction on  $n$  to prove this, so let  $E = \{n \in \{2, \dots\} \mid \text{if } 1 < m \leq n \text{ and } \forall i \in \mathbb{N}_0 \vdash m \leq i \leq n \text{ and for } M \in M(n \times n, F) \text{ we have } \text{col}(M, i)_j = \delta_{i,j} \text{ then } \det(M) = \det([\leq m]M)\}$  then we have:

- a. If  $n = 2$  then  $m = 2$  and by the previous theorem we have

$$\begin{aligned}
 \det(M) &= \sum_{i \in \{1, \dots, 2\}} (-i)^{i+2} \cdot M_{i,2} \cdot \det([i \boxplus 2]M) \\
 &= \sum_{i \in \{1, \dots, 2\}} (-i)^{i+2} \cdot \text{col}(M, 2)_i \cdot \det([i \boxplus 2]M) \\
 &= \sum_{i \in \{1, \dots, 2\}} (-i)^{i+2} \cdot \delta_{i,2} \cdot \det([i \boxplus 2]M) \\
 &= (-1)^{2+2} \cdot \det([2 \boxplus 2]M) \\
 &= \det([2 \boxplus 2]M) \\
 &\stackrel{10.336}{=} \det([\leq 2]M) \stackrel{m=2}{=} \det([\leq m]M)
 \end{aligned}$$

so we have  $2 \in E$

- b. Let  $n \in E$  then if  $1 < m \leq n+1$  is such that  $\forall i \in \mathbb{N}_0 \vdash m \leq i \leq n+1$  we have  $\text{col}(M, i)_j = \delta_{i,j}$  then we have for  $[n+1 \boxplus n+1]M \stackrel{10.336}{=} [\leq m+1]M$  that  $\forall i \in \mathbb{N}_0 \vdash m \leq i \leq n+1$

$$\begin{aligned}
 \text{col}([n+1 \boxplus n+1]M, i)_j &\stackrel{i \leq n < n+1}{=} M_{j,i} \\
 &= \text{col}(M, j)_i \\
 &\stackrel{m \leq i \leq n+1 \Rightarrow m \leq i \leq n+1}{=} \delta_{i,j}
 \end{aligned}$$

so as  $n \in E$  we have  $\det([\leq m+1]M) \stackrel{10.336}{=} \det([\leq m]([\leq m+1]M)) \stackrel{m \leq ((m+1)-1) \wedge 10.336}{=} \det([\leq m])$  so that  $n+1 \in E$ .

Using mathematical induction we have then  $E = \{2, \dots\} = \mathbb{N}_0 \setminus \{1\}$  proving (2)

3. As  $\text{row}(M, i)_j = M_{i,j} = M_{j,i}^T = \text{col}(M^T, i)_j$  we have  $\forall i \in \mathbb{N}_0 \vdash i \leq m$  that  $\text{col}(M^T, i)_j = \delta_{i,j}$  and thus by (1) we have  $\det([\geq m]M^T) \stackrel{10.336}{=} \det([\geq m]M)^T \stackrel{10.320}{=} \det([\geq m]M)$

4. As  $\text{row}(M, i)_j = M_{i,j} = M_{j,i}^T = \text{col}(M^T, i)_j$  we have  $\forall i \in \mathbb{N}_0 \vdash m \leq i \leq n$  that  $\text{col}(M^T, i)_j = \delta_{i,j}$  and thus by (1) we have  $\det([\leq m]M^T) \stackrel{10.336}{=} \det([\leq m]M)^T \stackrel{10.320}{=} \det([\leq m]M)$   $\square$

**Theorem 10.346.** *If  $F$  is a field of characterization zero then for  $n \in \mathbb{N}_0$  and  $M \in M(n \times n, F)$  we have*

$$\begin{aligned}
 \det(M) &\neq 0 \\
 \Leftrightarrow & \\
 \exists M^{-1} \in M(n \times n, F) \text{ with } M^{-1} \cdot M &= E = M \cdot M^{-1} \text{ where } M^{-1} = \frac{1}{\det(M)} \cdot \text{adjoint}(M)
 \end{aligned}$$

**Proof.** Using 10.342 we have  $M \cdot \text{adjoint}(M) = \text{adjoint}(M) \cdot M = \det(M) \cdot E$  and thus we have:

1. ( $\det(M) \neq 0$ ) then  $\det(M)^{-1} = \frac{1}{\det(M)}$  exists and we have  $M \cdot \left( \frac{1}{\det(M)} \cdot \text{adjoint}(M) \right) = \left( \frac{1}{\det(M)} \cdot \text{adjoint}(M) \right) \cdot M = \frac{1}{\det(M)} \cdot \det(M) \cdot E = E$ . So  $M^{-1}$  exists and  $M^{-1} = \frac{1}{\det(M)} \cdot \text{adjoint}(M)$
2. (**Assume  $M^{-1}$  exists**) then  $M \cdot M^{-1} = E$  and  $1 = \det(E) = \det(M \cdot M^{-1}) = \det(M) \cdot \det(M^{-1}) \Rightarrow \det(M) \neq 0$   $\square$

## 10.10 Expansion of linear transformation in elementair transformations

**Definition 10.347.** Let  $n \in \mathbb{N}_0$ ,  $X$  be a  $n$ -dimensional vector space over a field  $F$  of characterization zero with a basis  $\{e_i\}_{i \in \{1, \dots, n\}}$  then a elementair transformation is a **elementair transformation** on  $X$  if it is of one of the following forms

1.  $\forall i \in \{1, \dots, n\}$  define  $\sigma_n(i, \alpha): X \rightarrow X$  is defined by  $\forall k \in \{1, \dots, n\}$  we have  $\sigma_n(i, \alpha)(e_k) = \begin{cases} e_k & \text{if } k \neq i \\ \alpha \cdot e_i & \text{if } k = i \end{cases}$  where  $i \in \{1, \dots, n\}$  and  $\alpha \in F$
2.  $\forall i, j \in \{1, \dots, n\}$  define  $\tau_n(i, j): X \rightarrow X$  is defined by  $\forall k \in \{1, \dots, n\}$  we have  $\tau_n(i, j)(e_k) = \begin{cases} e_k & \text{if } k \neq i, j \\ e_j & \text{if } k = i \\ e_i & \text{if } k = j \end{cases}$  where  $i, j \in \{1, \dots, n\}$
3.  $\forall i, j \in \{1, \dots, n\} \mid i \neq j$  define  $\beta_n(i, j, \alpha): X \rightarrow X$  is defined by  $\forall k \in \{1, \dots, n\}$  we have  $\beta_n(i, j, \alpha)(e_k) = \begin{cases} e_k & \text{if } k \neq i \\ e_i + \alpha \cdot e_j & \text{if } k = i \end{cases}$  where  $i, j \in \{1, \dots, n\}$  and  $\alpha \in F$

**Note 10.348.** Let  $n \in \mathbb{N}_0$ ,  $X$  a  $n$ -dimensional vector space over a field  $F$  of characterization zero with basis  $\{e_i\}_{i \in \{1, \dots, n\}}$  then we have

1.  $\forall i \in \{1, \dots, n\}, \forall \alpha \in F$  we have  $\det(\sigma_n(i, \alpha)) = \alpha$ , so  $\sigma_n(i, \alpha)$  is a isomorphism if and only if  $\alpha \neq 0$
2.  $\forall i, j \in \{1, \dots, n\}$  we have  $\det(\tau_n(i, j)) = -1$  if  $i \neq j$  and  $\det(\tau_n(i, j)) = 1$  if  $i = j$ , so  $\tau_n$  is a isomorphism
3.  $\forall i, j \in \{1, \dots, n\}, \forall \alpha \in F$  we have  $\det(\beta_n(i, j, \alpha)) = 1$
4. If  $i = j$  then  $\tau_n(i, j) = 1_X$
5.  $\tau_n(i, j) \circ \tau_n(i, j) = 1_X$  proving that  $\tau_n(i, j) = \tau_n(i, j)^{-1}$

**Proof.**

1. If  $i, j \in \{1, \dots, n\}$  with  $i = j$  then we have if  $k \in \{1, \dots, n\}$  that  $\tau_n(i, j)(e_k) = \begin{cases} e_k & \text{if } k \in \{1, \dots, n\} \setminus \{i\} \\ e_j & \text{if } k = i \\ e_i & \text{if } k = i \end{cases} \stackrel{i=j \Rightarrow e_i=e_j}{=} \begin{cases} e_k & \text{if } k \in \{1, \dots, n\} \setminus \{i, j\} \\ e_i & \text{if } k = i \\ e_j & \text{if } k = j \end{cases} = e_k = 1_X(e_k)$  proving that  $\tau_n(i, j) = 1_X$
2. Let  $i, j \in \{1, \dots, n\}$  then we have for  $k \in \{1, \dots, n\}$  either  
 $\mathbf{k} \in \{1, \dots, n\} \setminus \{i, j\}$ . then  $(\tau_n(i, j) \circ \tau_n(i, j))(e_k) = \tau_n(i, j)(\tau_n(i, j)(e_k)) = \tau_n(i, j)(e_k) = e_k = 1_X(e_k)$   
 $\mathbf{k} = i$ .  $(\tau_n(i, j) \circ \tau_n(i, j))(e_k) = \tau_n(i, j)(\tau_n(i, j)(e_k)) = \tau_n(i, j)(e_j) = e_i = e_k = 1_X(e_k)$   
 $\mathbf{k} = j$ .  $(\tau_n(i, j) \circ \tau_n(i, j))(e_k) = \tau_n(i, j)(\tau_n(i, j)(e_k)) = \tau_n(i, j)(e_i) = e_j = e_k = 1_X(e_k)$   
which proves that  $\tau_n(i, j) \circ \tau_n(i, j) = 1_X$   $\square$

**Theorem 10.349.** Let  $n \in \mathbb{N}_0$ ,  $X$  be a  $n$ -dimensional vectorspace over a field  $F$  of characterization zero with basis  $\{e_i\}_{i \in \{1, \dots, n\}}$  then we have

1.  $\forall i \in \{1, \dots, n\}, \forall \alpha \in F$  we have  $\det(\sigma_n(i, \alpha)) = \alpha$ , so  $\sigma_n(i, \alpha)$  is a isomorphism if and only if  $\alpha \neq 0$

2.  $\forall i, j \in \{1, \dots, n\}$  we have  $\det(\tau_n(i, j)) = -1$  if  $i \neq j$  and  $\det(\tau_n(i, j)) = 1$  if  $i = j$ , so  $\tau_n$  is a isomorphism
3.  $\forall i, j \in \{1, \dots, n\}, \forall \alpha \in F$  we have  $\det(\beta_n(i, j, \alpha)) = 1$

**Proof.** See the following examples 10.279. 10.280 and 10.281, together with the fact that if  $i = j$  we have that  $\tau_n(i, j) = 1_X$ .  $\square$

The following theorem show how to extend a elementair function.

**Definition 10.350.** Let  $n \in \mathbb{N}_0 \setminus \{1\}$ ,  $X$  a  $n$ -dimensional vectorspace over a field  $F$  of characterization zero with basis  $\{e_i\}_{i \in \{1, \dots, n\}}$ ,  $Y = S(\{e_i | i \in \{1, \dots, n-1\}\})$  (see 10.130) a  $(n-1)$ -dimensional subvector space of  $X$  then if  $L \in \text{Hom}(Y, Y)$  we define  $L_{\langle n \rangle} \in \text{Hom}(X, X)$  by  $L_{\langle n \rangle}(e_k) = \begin{cases} L(e_k) & \text{if } k \in \{1, \dots, n-1\} \\ e_n & \text{if } k = n \end{cases}$

**Theorem 10.351.** Let  $n \in \mathbb{N}_0 \setminus \{1\}$ ,  $X$  a  $n$ -dimensional vectorspace over a field  $F$  of characterization zero with basis  $\{e_i\}_{i \in \{1, \dots, n\}}$ ,  $Y = S(\{e_i | i \in \{1, \dots, n-1\}\})$  (see 10.130) a  $(n-1)$ -dimensional subvector space of  $X$  then if  $L: Y \rightarrow Y$  is elementair transformation on  $X$  then  $L_{\langle n \rangle}: X \rightarrow X$  defined by  $L_{\langle n \rangle}(e_k) = \begin{cases} L(e_k) & \text{if } k \in \{1, \dots, n-1\} \\ e_n & \text{if } k = n \end{cases}$  is a elementair trasnformation on  $Y$ .

**Proof.** As  $L$  is a elementair transformation it is of one of the following forms

1.  $L = \sigma_{n-1}(i, \alpha)$ ,  $i \in \{1, \dots, n-1\}, \alpha \in F$  then if  $k \in \{1, \dots, n\}$  we have  $L_{\langle n \rangle}(e_k) = \begin{cases} L(e_k) & \text{if } k \in \{1, \dots, n-1\} \\ e_n & \text{if } k = n \end{cases} = \begin{cases} \sigma_{n-1}(i, j)(e_k) & \text{if } k \in \{1, \dots, n-1\} \\ e_k & \text{if } k = n \end{cases} = \begin{cases} e_k & \text{if } k \in \{1, \dots, n-1\} \setminus \{i\} \\ \alpha \cdot e_i & \text{if } k = i \\ e_k & \text{if } k = n \end{cases} = \begin{cases} e_k & \text{if } k \in \{1, \dots, n\} \setminus \{i\} \\ \alpha \cdot e_i & \text{if } k = i \end{cases} = \sigma_n(i, \alpha)(e_k)$ , which proves that  $L_{\langle n \rangle} = \sigma_n(i, \alpha)$
2.  $L = \tau_{n-1}(i, j)$ ,  $i, j \in \{1, \dots, n-1\}$  then if  $k \in \{1, \dots, n\}$  we have  $L_{\langle n \rangle}(e_k) = \begin{cases} L(e_k) & \text{if } k \in \{1, \dots, n-1\} \\ e_n & \text{if } k = n \end{cases} = \begin{cases} \tau_{n-1}(i, j)(e_k) & \text{if } k \in \{1, \dots, n-1\} \\ e_k & \text{if } k = n \end{cases} = \begin{cases} e_k & \text{if } k \in \{1, \dots, n-1\} \setminus \{i, j\} \\ e_i & \text{if } k = j \\ e_j & \text{if } k = i \\ e_k & \text{if } k = n \end{cases} = \tau_n(i, j)(e_k)$ , which proves that  $L_{\langle n \rangle} = \tau_n(i, j)$
3.  $L = \beta_{n-1}(i, j, \alpha)$ ,  $i, j \in \{1, \dots, n-1\}, \alpha \in F$  then if  $k \in \{1, \dots, n\}$  we have  $L_{\langle n \rangle}(e_k) = \begin{cases} L(e_k) & \text{if } k \in \{1, \dots, n-1\} \\ e_n & \text{if } k = n \end{cases} = \begin{cases} \beta_n(i, j) & \text{if } k \in \{1, \dots, n-1\} \\ e_k & \text{if } k = n \end{cases} = \begin{cases} e_k & \text{if } k \in \{1, \dots, n-1\} \setminus \{i\} \\ e_i + \alpha \cdot e_j & \text{if } k = i \\ e_k & \text{if } k = n \end{cases} = \begin{cases} e_k & \text{if } k \in \{1, \dots, n\} \setminus \{i\} \\ e_i + \alpha \cdot e_j & \text{if } k = i \end{cases} = \beta_n(i, j, \alpha)(e_k)$  which proves that  $L_{\langle n \rangle} = \beta_n(i, j, \alpha)$   $\square$

**Corollary 10.352.** Let  $n \in \mathbb{N}_0 \setminus \{1\}$ ,  $X$  a  $n$ -dimensional vector space over a field  $F$  of characterization zero with basis  $\{e_i\}_{i \in \{1, \dots, n\}}$ ,  $Y = S(\{e_i | i \in \{1, \dots, n\}\})$  (see 10.130) a  $(n-1)$ -dimensional subvector space of  $X$  then if  $L = L_m \circ \dots \circ L_1$  where  $\{L_i\}_{i \in \{1, \dots, n\}}$  a familie of elementair transformations on  $Y$  then we have  $L_{\langle n \rangle} = (L_m)_{\langle n \rangle} \circ \dots \circ (L_1)_{\langle n \rangle}$  also a composition of elementair transformations (by the above theorem).

**Proof.** This is proved by induction on  $m$ . so let  $S = \{m \in \mathbb{N}_0 | \text{If } L = L_m \circ \dots \circ L_1 \text{ is the composition of } m \text{ elementair transformations on } Y \text{ then } L_{\langle n \rangle} = ((L_m)_{\langle n \rangle} \circ \dots \circ (L_1)_{\langle n \rangle})\}$  then we have

**1  $\in S$ .** If  $L = L_1 \circ \dots \circ L_1 = L_1$  then  $L_{\langle n \rangle} = (L_1)_{\langle n \rangle} = (L_1)_{\langle n \rangle} \circ \dots \circ (L_1)_{\langle n \rangle}$

**$m \in S \Rightarrow m+1 \in S$ .** First if  $L = L_2 \circ L_1$  then  $\forall i \in \{1, \dots, n\}$  we have  $L_{\langle n \rangle}(e_i) = \begin{cases} L(e_i) & \text{if } i \in \{1, \dots, n-1\} \\ e_n & \text{if } i = n \end{cases} = \begin{cases} L_2(L_1(e_i)) & \text{if } i \in \{1, \dots, n-1\} \\ e_n & \text{if } i = n \end{cases} = \begin{cases} (L_2)_{\langle n \rangle}((L_1)_{\langle n \rangle}(e_i)) & \text{if } i \in \{1, \dots, n-1\} \\ (L_2)_{\langle n \rangle}((L_1)_{\langle n \rangle}(e_i)) & \text{if } i = n \end{cases} = (L_2)_{\langle n \rangle}((L_1)_{\langle n \rangle}(e_i))$  proving that

$$(L_2 \circ L_1)_{\langle n \rangle} = (L_2)_{\langle n \rangle} \circ (L_1)_{\langle n \rangle} \quad (10.122)$$

Let now  $L = L_{m+1} \circ \dots \circ L_1$  then  $L_{\langle n \rangle} = (L_{m+1} \circ \dots \circ L_1)_{\langle n \rangle} = (L_{m+1} \circ (L_m \circ \dots \circ L_1))_{\langle n \rangle} \stackrel{10.122}{=} (L_{m+1})_{\langle n \rangle} \circ (L_m \circ \dots \circ L_1)_{\langle n \rangle} \stackrel{m \in S}{=} (L_{m+1})_{\langle n \rangle} \circ ((L_m)_{\langle n \rangle} \circ \dots \circ (L_1)_{\langle n \rangle}) = ((L_{m+1})_{\langle n \rangle} \circ \dots \circ (L_1)_{\langle n \rangle})$  proving that  $m+1 \in S$   $\square$

Using the composition of  $n$  function on linear transformations we have the following theorem

**Lemma 10.353.** Let  $n, m \in \mathbb{N}_0$ ,  $X$  a  $n$ -dimensional vector space over a field of characterization zero,  $\{L_i\}_{i \in \{1, \dots, n\}}$  a family of linear transformations  $L_i: X \rightarrow X$  then we have

1.  $L_m \circ \dots \circ L_1: X \rightarrow X$  is a linear transformation
2.  $\det(L_m \circ \dots \circ L_1) = \prod_{i \in \{1, \dots, n\}} \det(L_i)$
3. If  $L$  is regular then  $\forall i \in \{1, \dots, m\}$  we have that  $L_i$  is regular (thus a isomorphism because of 10.214).

**Proof.** We use induction to prove this theorem

1. Let  $S = \{m \in \mathbb{N}_0 \mid \text{If } \{L_i\}_{i \in \{1, \dots, m\}} \text{ is a family of linear transformations on } X \text{ then } L_m \circ \dots \circ L_1 \text{ is a linear transformation on } X\}$  then we have

**1  $\in S$ .** If  $\{L_i\}_{i \in \{1, \dots, 1\}}$  is a family of linear transformations then  $L_1 \circ \dots \circ L_1 = L_1: X \rightarrow X$  is a linear transformation proving that  $1 \in S$

**$m \in S \Rightarrow m + 1 \in S$ .** If  $\{L_i\}_{i \in \{1, \dots, m+1\}}$  is a family of linear transformations on  $X$  then as  $m \in S$  we have that  $L_m \circ \dots \circ L_1: X \rightarrow X$  is a linear transformation on  $X$ . As  $L_{m+1}: X \rightarrow X$  is also a linear transformation on  $X$  we have by 10.184 that  $L_{m+1} \circ \dots \circ L_1 = L_{m+1} \circ (L_m \circ \dots \circ L_1): X \rightarrow X$  is a linear transformation on  $X$ , proving that  $m+1 \in S$ .

2. Let  $S = \{m \in \mathbb{N}_0 \mid \text{If } \{L_i\}_{i \in \{1, \dots, m\}} \text{ is a family of linear transformations on } X \text{ then } \det(L_m \circ \dots \circ L_1) = \prod_{i \in \{1, \dots, m\}} \det(L_i)\}$  then we have

**1  $\in S$ .** If  $\{L_i\}_{i \in \{1, \dots, 1\}}$  is a family of linear transformations then  $\det(L_1 \circ \dots \circ L_1) = \det(L_1) = \prod_{i \in \{1, \dots, 1\}} \det(L_i)$  proving that  $1 \in S$

**$m \in S \Rightarrow m + 1 \in S$ .** If  $\{L_i\}_{i \in \{1, \dots, m+1\}}$  is a family of linear transformations on  $X$  then  $\det(L_{m+1} \circ \dots \circ L_1) = \det(L_{m+1} \circ (L_m \circ \dots \circ L_1)) \stackrel{10.282}{=} \det(L_{m+1}) \cdot \det(L_m \circ \dots \circ L_1) \stackrel{m \in S}{=} \det(L_{m+1}) \cdot \prod_{i \in \{1, \dots, m\}} \det(L_i) = \prod_{i \in \{1, \dots, m+1\}} \det(L_i)$

3. Assume that  $\{L_i\}_{i \in \{1, \dots, m\}}$  is a family of linear transformations then by 10.282 we have that  $0 \neq \det(L_m \circ \dots \circ L_1) = \prod_{i \in \{1, \dots, m\}} \det(L_i)$  so that  $\forall i \in \{1, \dots, m\}$  we must have that  $\det(L_i) \neq 0$  and thus by 10.282 again we have that  $L_i$  is a regular linear transformation.  $\square$

Next we need the following two lemmas

**Lemma 10.354.** Let  $n \in \mathbb{N}_0$ ,  $X$  a  $(n+1)$ -dimensional vector space over a field  $F$  of characterization zero with basis  $\{e_i\}_{i \in \{1, \dots, n+1\}}$  then if  $\{\alpha_i\}_{i \in \{1, \dots, n\}}$  is a family of values in  $F$  then  $\forall k \in \{1, \dots, n\}$  we have  $(\beta_{n+1}(n+1, n, \alpha_n) \circ \dots \circ \beta_{n+1}(n+1, 1, \alpha_1))(e_k) = e_k$

**Proof.** Let  $k \in \{1, \dots, n\}$  take then  $S_k = \{m \in \mathbb{N}_0 \mid \text{If } m \leq n \text{ then } (\beta_{n+1}(n+1, m, \alpha_m) \circ \dots \circ \beta_{n+1}(n+1, 1, \alpha_1))(e_k) = e_k\}$  then we have

**1  $\in S$ .**  $(\beta_{n+1}(n+1, 1, \alpha_1) \circ \dots \circ \beta_{n+1}(n+1, 1, \alpha_1))(e_k) = \beta_{n+1}(n+1, 1, \alpha_1)(k) \stackrel{k \neq n+1}{=} e_k$  proving that  $1 \in S_k$

**$m \in S_k$ .** Take  $m+1$  and assume that  $m+1 \leq n$  then we have  $(\beta_{n+1}(n+1, m+1, \alpha_{m+1}) \circ \dots \circ \beta_{n+1}(n+1, 1, \alpha_1))(x_k) = (\beta_{n+1}(n+1, m+1, \alpha_{m+1}) \circ (\beta_{n+1}(n+1, m, \alpha_m) \circ \dots \circ \beta_{n+1}(n+1, 1, \alpha_1)))(x_k) = \beta_{n+1}(n+1, m+1, \alpha_{m+1})((\beta_{n+1}(n+1, m, \alpha_m) \circ \dots \circ \beta_{n+1}(n+1, 1, \alpha_1))(x_k)) \stackrel{m \in S_k, m \leq n}{=} \beta_{n+1}(n+1, m+1, \alpha_{m+1})(x_k) \stackrel{k \neq n+1}{=} x_k$  proving that  $m+1 \in S_k$

By mathematical induction we have then that  $S_k = \mathbb{N}_0$ , as  $n \in \mathbb{N}_0 \Rightarrow n \in S_k$  and  $n \leq n$ , we have  $\forall k \in \{1, \dots, n\}$  that  $(\beta_{n+1}(n+1, n, \alpha_n) \circ \dots \circ \beta_{n+1}(n+1, 1, \alpha_1))(e_k) = e_k$   $\square$

**Lemma 10.355.** Let  $n \in \mathbb{N}_0$ ,  $X$  a  $(n+1)$ -dimensional vector space over a field  $F$  of characterization zero with basis  $\{e_i\}_{i \in \{1, \dots, n+1\}}$  and  $\{\alpha_i\}_{i \in \{1, \dots, n\}}$  a family of values in  $F$  then  $(\beta_{n+1}(n+1, n, \alpha_n) \circ \dots \circ \beta_{n+1}(n+1, 1, \alpha_1))(e_{n+1}) = (\sum_{k=1}^n \alpha_k \cdot e_k) + e_{n+1}$

**Proof.** We prove this by induction so let  $S = \{m \in \mathbb{N}_0 \mid \text{If } m \leq n \text{ then } (\beta_{n+1}(n+1, m, \alpha_m) \circ \dots \circ \beta_{n+1}(n+1, 1, \alpha_1))(e_{n+1}) = (\sum_{k=1}^m \alpha_k \cdot e_k) + e_{n+1}\}$  then we have

**1  $\in S$ .**  $(\beta_{n+1}(n+1, 1, \alpha_1) \circ \dots \circ \beta_{n+1}(n+1, 1, \alpha_1))(e_{n+1}) = \beta_{n+1}(n+1, 1, \alpha_1)(e_{n+1}) = \alpha_1 \cdot e_1 + e_{n+1} = (\sum_{i=1}^1 \alpha_k \cdot e_k) + e_{n+1}$  proving that  $1 \in S$

$m \in S \Rightarrow m + 1 \in S$ . Take  $m + 1$  and assume that  $m + 1 \leq n$  then  $(\beta_{n+1}(n+1, m+1, \alpha_{m+1}) \circ \dots \circ \beta_{n+1}(n+1, 1, \alpha_1))(e_{n+1}) = (\beta_{n+1}(n+1, m+1, \alpha_{m+1}) \circ (\beta_{n+1}(n+1, m, \alpha_m) \circ \dots \circ \beta_{n+1}(n+1, 1, \alpha_1)))(e_{n+1}) = \beta_{n+1}(n+1, m+1, \alpha_{m+1})((\beta_{n+1}(n+1, m, \alpha_m) \circ \dots \circ \beta_{n+1}(n+1, 1, \alpha_1))(e_{n+1}))$   $\underset{m \leq n \wedge m \in S}{=} \beta_{n+1}(n+1, m+1, \alpha_{m+1})((\sum_{k=1}^m \alpha_k \cdot e_k) + e_{n+1}) = (\sum_{k=1}^m \alpha_k \cdot e_k) + e_{n+1} = \beta_{n+1}(n+1, m+1, \alpha_{m+1})(e_k) + \beta_{n+1}(n+1, m+1, \alpha_{m+1})(e_{n+1}) = (\sum_{k=1}^m \alpha_k \cdot e_k) + \alpha_{m+1} \cdot e_{m+1} + e_{n+1} = (\sum_{k=1}^{m+1} \alpha_k \cdot e_k) + e_{n+1}$  proving that  $m + 1 \in S$

By mathematical induction we have then that  $S = \mathbb{N}_0$ , so as  $n \in \mathbb{N}_0 \Rightarrow n \in S$  and  $n \leq n$  we have  $(\beta_{n+1}(n+1, n, \alpha_n) \circ \dots \circ \beta_{n+1}(n+1, 1, \alpha_1)) = (\sum_{k=1}^n \alpha_k \cdot e_k) + e_{n+1}$   $\square$

**Lemma 10.356.** Let  $n \in \mathbb{N}_0$ ,  $X$  a  $(n+1)$ -dimensional vector space over a field  $F$  of characterization zero with basis  $\{e_i\}_{i \in \{1, \dots, n+1\}}$  and  $\{a_i\}_{i \in \{1, \dots, n\}}$  a family in  $F$  then if  $m \in \mathbb{N}_0$  with  $m \leq n$  is such that  $i \in \{m+1, \dots, n+1\}$  we have that  $(\beta_{n+1}(m, n+1, \alpha_m) \circ \dots \circ \beta_{n+1}(1, n+1, \alpha_1))(e_i) = e_i$

**Proof.** This is proved by induction so let  $S = \{m \in \mathbb{N}_0 \mid \text{If } m \leq n \text{ then if } i \in \{m+1, \dots, n+1\} \text{ we have } (\beta_{n+1}(m, n+1, \alpha_m) \circ \dots \circ \beta_{n+1}(1, n+1))(e_i) = e_i\}$  then we have

$1 \in S$ . If  $i \in \{2, \dots, n+1\}$  then we have  $i \neq 1$  so that  $(\beta_{n+1}(1, n+1, \alpha_1) \circ \dots \circ \beta_{n+1}(1, n+1, \alpha_1))(e_i) = \beta_{n+1}(1, n+1, \alpha_1)(e_i) = e_i$  proving that  $1 \in S$

$m \in S \Rightarrow m + 1 \in S$ . If  $m + 1 \leq n$  then if  $i \in \{(m+1)+1, \dots, n+1\}$  we have  $i \in \{(m+1), \dots, n+1\}$  and  $(\beta_{n+1}(m+1, n+1, \alpha_{m+1}) \circ \dots \circ \beta_{n+1}(1, n+1, \alpha_1))(e_1) = \beta_{n+1}(m+1, n+1, \alpha_{m+1})((\beta_{n+1}(m, n+1, \alpha_m) \circ \dots \circ \beta_{n+1}(1, n+1, \alpha_1))(e_i))$   $\underset{m \in S \wedge i \in \{(m+1), \dots, n+1\}}{=} \beta_{n+1}(m+1, n+1, \alpha_{m+1})(e_i) = e_i$  proving that  $m + 1 \in S$

By mathematical induction we have then that  $S = \mathbb{N}_0$  so if  $m \in \mathbb{N}_0 = S$  with  $m \leq n$  we have if  $i \in \{m+1, \dots, n+1\}$  that  $(\beta_{n+1}(m, n+1, \alpha_m) \circ \dots \circ \beta_{n+1}(1, n+1, \alpha_1))(e_i) = e_i$ .  $\square$

**Lemma 10.357.** Let  $n \in \mathbb{N}_0$ ,  $X$  a  $(n+1)$ -dimensional vector space over a field  $F$  of characterization zero with basis  $\{e_i\}_{i \in \{1, \dots, n+1\}}$  and  $\{a_i\}_{i \in \{1, \dots, n\}}$  a matrix in  $F$  then given  $i \in \{1, \dots, n\}$  we have  $(\beta_{n+1}(n, n+1, \alpha_n) \circ \dots \circ \beta_{n+1}(1, n+1, \alpha_1))(e_i) = e_i + \alpha_i \cdot e_{n+1}$

**Proof.** We prove this by induction so let  $S_i = \{m \in \mathbb{N}_0 \mid \text{If } m \leq n \text{ then if } i \in \{1, \dots, m\} \text{ we have } (\beta_{n+1}(m, n+1, \alpha_m) \circ \dots \circ \beta_{n+1}(1, n+1, \alpha_1))(e_i) = e_i + \alpha_i \cdot e_{n+1}\}$  then we have

$1 \in S_i$ . For if  $m = 1$  then if  $i \in \{1, \dots, 1\} \Rightarrow i = 1$  we have  $(\beta_{n+1}(1, n+1, \alpha_1) \circ \dots \circ \beta_{n+1}(1, n+1, \alpha_1))(e_i) = \beta_{n+1}(1, n+1, \alpha_1)(e_1) = e_1 + \alpha_1 \cdot e_{n+1}$  proving that  $1 \in S_i$

$m \in S \Rightarrow m + 1 \in S$ . Assume that  $m + 1 \leq n$  and that  $i \in \{1, \dots, m+1\}$  then we have either

$i \in \{1, \dots, m\}$ . then  $(\beta_{n+1}(m+1, n+1, \alpha_{m+1}) \circ \dots \circ \beta_{n+1}(1, n+1, \alpha_1))(e_i) = \beta_{n+1}(m+1, n+1, \alpha_{m+1})((\beta_{n+1}(m, n+1, \alpha_m) \circ \dots \circ \beta_{n+1}(1, n+1, \alpha_1))(e_i))$   $\underset{m \in S \wedge i \in \{1, \dots, m\} \wedge m \leq n}{=} \beta_{n+1}(m+1, n+1, \alpha_{m+1})(e_i + \alpha_i \cdot e_{n+1}) = \beta_{n+1}(m+1, n+1, \alpha_{m+1})(e_i) + \alpha_i \cdot \beta_{n+1}(m+1, n+1, \alpha_{m+1})(e_{n+1})$   $\underset{i, n+1 \neq m+1}{=} e_i + \alpha_i \cdot e_{n+1}$

$i = m + 1$ . then  $(\beta_{n+1}(m+1, n+1, \alpha_{m+1}) \circ \dots \circ \beta_{n+1}(1, n+1, \alpha_1))(e_i) = \beta_{n+1}(m+1, n+1, \alpha_{m+1})((\beta_{n+1}(m, n+1, \alpha_m) \circ \dots \circ \beta_{n+1}(1, n+1, \alpha_1))(e_i))$   $\underset{\text{Lemma 10.356}}{=} \beta_{n+1}(m+1, n+1, \alpha_{m+1})(e_i) = e_{m+1} + \alpha_{m+1} \cdot e_{n+1} = e_i + \alpha_i \cdot e_{n+1}$

This proves that  $m + 1 \in S$

Using mathematical induction we have  $\mathbb{N}_0 = S$  so as  $n \in \mathbb{N}_0 = S$  and  $n \leq n$  we have  $(\beta_{n+1}(n, n+1, \alpha_n) \circ \dots \circ \beta_{n+1}(1, n+1, \alpha_1))(e_i) = e_i + \alpha_i \cdot e_{n+1}$   $\square$

**Theorem 10.358.** Let  $n \in \mathbb{N}_0$ ,  $X$  a  $n$ -dimensional space over a field  $F$  of characterization zero,  $L \in \text{Hom}(X, X)$  a regular linear transformation (hence a isomorphism) then there exists a family  $\{L_i\}_{i \in \{1, \dots, n\}}$  of elementair regular transformations so that  $L = L_n \circ \dots \circ L_1$

**Proof.** We prove this by induction on  $n$  so let  $S = \{n \in \mathbb{N}_0 \mid \text{If } X \text{ is a } n\text{-dimensional vector space over } F, L \in \text{Hom}(X, X) \text{ a regular transformation then } \exists \{L_i\}_{i \in \{1, \dots, n\}} \text{ of regular elementair transformations so that } L = L_n \circ \dots \circ L_1\}$ , we have then

$1 \in S$ . If  $X$  is one dimensional it has a base  $\{e_i\}_{i \in \{1, \dots, 1\}}$  then if  $L \in \text{Hom}(X, X)$  we have

$L(e_1) = \alpha \cdot e_1$  so  $L = \sigma_1(1, \alpha)$  a elementair function, which must be regular as  $L$  is regular, from  $L = \sigma_1(1, \alpha) = \sigma_1(1, \alpha) \circ \dots \circ \sigma_1(1, \alpha)$  we have then that  $1 \in S$ .

$n \in S \Rightarrow n+1 \in S$ . Take now  $X$  a  $(n+1)$ -dimensional space and  $L \in \text{Hom}(X, X)$  a regular transformation. Define now  $M' = \{M'_{i,j}\}_{i,j \in \{1, \dots, n+1\} \times \{1, \dots, n+1\}} = \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n+1\}}, \{e_i\}_{i \in \{1, \dots, n+1\}})$  the associated matrix of  $L$  so that  $L$  is defined by

$$\forall i \in \{1, \dots, n\} \models L(e_i) = \sum_{k=1}^{n+1} M'_{k,i} e_k \quad (10.123)$$

As  $L$  is regular we have that  $0 \neq \det(L) \stackrel{10.321}{=} \det(M') \stackrel{10.342, j=n+1}{=} \sum_{i \in \{1, \dots, n+1\}} (-1)^{i+(n+1)} M'_{i,j} \cdot \det([i \boxplus (n+1)] M')$  so that  $\exists i_0 \in \{1, \dots, n+1\}$  so that

$$\det(i_0 \boxplus (n+1) M') \neq 0 \quad (10.124)$$

We have now for  $i_0$  the following possibilities

$i_0 = n+1$ . So  $\det([n+1 \boxplus n+1] M) \neq 0$ , take then  $T = 1_{\{1, \dots, n+1\}} = \tau_{n+1}(n+1, n+1)$  a elementair transformation and  $L_1 = L$  then we have  $L = T \circ L$  and

There exists a elementair transformation  $T$  such that  $L = T \circ L_1, M = \mathcal{M}(L_1, \{e_i\}_{i \in \{1, \dots, n+1\}}, \{e_i\}_{i \in \{1, \dots, n+1\}}), \det([n+1 \boxplus n+1] M) \neq 0$  (10.125)

$i_0 \leq n$ . Consider now  $([i_0 \boxplus (n+1)] M')_{k,l}$  then we have the following cases (based on the definition) for  $k, l \in \{1, \dots, n\}$

$k < i_0 \wedge l < n+1$ . then  $([i_0 \boxplus (n+1)] M')_{k,l} = M'_{k,l}$

$i_0 \leq k \leq n \wedge l < n+1$ . then  $([i_0 \boxplus (n+1)] M')_{k,l} = M'_{k+1,l}$

$k < i_0 \wedge n+1 \leq l \leq n$ . does not occur as  $n < n+1$

$i_0 \leq k \leq n \wedge n+1 \leq l \leq n$ . does not occur as  $n < n+1$

proving that

$$\forall k, l \in \{1, \dots, n\} \text{ we have } ([i_0 \boxplus (n+1)] M')_{k,l} = \begin{cases} M'_{k,l} & \text{if } 1 \leq k < i_0 \\ M'_{k+1,l} & \text{if } i_0 \leq k \leq n \end{cases} \quad (10.126)$$

Take now  $(n \rightsquigarrow n i_0): \{1, \dots, n\} \rightarrow \{1, \dots, n\}$  (see 10.95 for the definition) then we have for  $([i_0 \boxplus (n+1)] M')_{(n \rightsquigarrow n i_0)}$  the following cases

$i_0 = n$ . then  $\forall k, l \in \{1, \dots, n\}$  we have  $(([i_0 \boxplus (n+1)] M')_{(n \rightsquigarrow n n)})_{k,l} = ([i \boxplus (n+1)] M')_{k,l} = \begin{cases} M'_{k,l} & \text{if } 1 \leq k < n \\ M'_{n+1,l} & \text{if } k = n \end{cases} = \begin{cases} M'_{k,l} & \text{if } k \in \{1, \dots, n\} \setminus \{i_0\} \\ M'_{n+1,l} & \text{if } k = i_0 \end{cases}$

$i_0 < n$ . then  $\forall k, l \in \{1, \dots, n\}$  we have

$$\begin{aligned} (([i_0 \boxplus (n+1)] M')_{(n \rightsquigarrow n i_0)})_{k,l} &= ([i_0 \boxplus (n+1)] M')_{(n \rightsquigarrow n i_0)(k),l} \\ &= \begin{cases} ([i_0 \boxplus (n+1)] M')_{k,l} & \text{if } k < i_0 \\ ([i_0 \boxplus (n+1)] M')_{n,l} & \text{if } k = i_0 \\ ([i_0 \boxplus (n+1)] M')_{k-1,l} & \text{if } i_0 < k \leq n \\ ([i_0 \boxplus (n+1)] M')_{k,l} & \text{if } n < k \text{ (impos.)} \end{cases} \end{aligned}$$

$$\stackrel{10.126}{=} \begin{cases} M'_{k,l} & \text{if } k < i_0 \\ M'_{n+1,l} & \text{if } k = i_0 \\ M'_{(k-1)+1,l} & \text{if } i_0 < k \leq n \end{cases}$$

$$= \begin{cases} M'_{k,l} & \text{if } k < i_0 \\ M'_{n+1,l} & \text{if } k = i_0 \\ M'_{k,l} & \text{if } i_0 < k \leq n \end{cases}$$

$$= \begin{cases} M'_{k,l} & \text{if } k \in \{1, \dots, n\} \setminus \{i_0\} \\ M'_{n+1,l} & \text{if } k = i_0 \end{cases}$$

So we have that in all cases that  $\forall k, l \in \{1, \dots, n\}$  we have  $(([i_0 \boxplus (n+1)]M_i)_{(n \rightsquigarrow_n i_0)})_{k,l} = \begin{cases} M'_{k,l} & \text{if } k \neq i_0 \\ M'_{n+1,l} & \text{if } k = i_0 \end{cases}$  so we have

$$\text{If } M'' \text{ is defined by } \forall k, l \in \{1, \dots, n\} \models M''_{k,l} = \begin{cases} M'_{k,l} & \text{if } k \neq i_0 \\ M'_{n+1,l} & \text{if } k = i_0 \end{cases} \text{ then } M'' = ([i_0 \boxplus (n+1)]M')_{(n \rightsquigarrow_n i_0)} \quad (10.127)$$

Now we have that  $\det(M'') = \det(([i_0 \boxplus (n+1)]M')_{(n \rightsquigarrow_n i_0)}) \underset{10.325}{=} \text{sign}(n \rightsquigarrow_n i_0) \cdot \det([i_0 \boxplus (n+1)]M') \neq 0$  (using 10.124) proving that

$$\det(M'') \neq 0 \quad (10.128)$$

Define now  $L_1 = \tau_{n+1}(n+1, i_0) \circ L$  and let  $M = \mathcal{M}(L_1, \{e_i\}_{i \in \{1, \dots, n+1\}}, \{e_i\}_{i \in \{1, \dots, n+1\}})$  then we have that  $\forall i \in \{1, \dots, n\}$

$$\begin{aligned} L_1(e_i) &= \tau_{n+1}(n+1, i_0)(L(e_i)) \\ &= \tau_{n+1}(n+1, i_0) \left( \sum_{k \in \{1, \dots, n\}} M'_{k,i} \cdot e_k \right) \\ &= \sum_{k \in \{1, \dots, n\}} M'_{k,i} \cdot \tau_{n+1}(n+1, i_0)(e_k) \\ &= \sum_{k \in \{1, \dots, n\} \setminus \{n+1, i_0\}} M'_{k,i} \cdot \tau_{n+1}(n+1, i_0)(e_k) + M'_{i_0,i} \cdot \tau_{n+1}(n+1, i_0)(e_{i_0}) + M'_{n+1,i} \cdot \tau_{n+1}(n+1, i_0)(e_{n+1}) \\ &= \sum_{k \in \{1, \dots, n\} \setminus \{n+1, i_0\}} M'_{k,i} \cdot e_k + M'_{i_0,i} \cdot e_{n+1} + M'_{n+1,i} \cdot e_{i_0} \\ &= \sum_{k \in \{1, \dots, n\} \setminus \{n+1, i_0\}} M_{k,i} \cdot e_k + M_{i_0,i} \cdot e_{n+1} + M_{n+1,i} \cdot e_{i_0} \end{aligned}$$

proving by the uniqueness of the expansion in a basis that  $M$  is defined by  $\forall k, l \in \{1, \dots, n+1\}$  we have  $M_{k,l} = \begin{cases} M_{k,l} & \text{if } k \in \{1, \dots, n\} \setminus \{n+1, i_0\} \\ M_{n+1,l} & \text{if } k = i_0 \\ M_{i_0,l} & \text{if } k = n+1 \end{cases}$ . Now  $\forall k, l \in \{1, \dots, n\}$  we have

$$\begin{aligned} ([n+1 \boxplus n+1]M)_{k,l} &\underset{k, l \leq n+1}{=} M_{k,l} \\ &\underset{i_0 \leq n \leq n+1}{=} \begin{cases} M_{k,l} & \text{if } k \in \{1, \dots, n\} \setminus \{i_0\} \\ M_{n+1} & \text{if } k = i_0 \end{cases} \\ &= M'' \end{aligned}$$

Using 10.128 we have then that  $\det([n+1 \boxplus n+1]M) = \det(M'') \neq 0$ . As  $L_1 = \tau_{n+1}(n+1, i_0) \circ L \underset{10.348}{\Rightarrow} L = \tau_{n+1}(n+1, i_0) \circ L_1$ . Take  $T = \tau_{n+1}(n+1, i_0)$  a elementair transformation then we have

There exist a elementair transformation  $T$  such that  $L = T \circ L_1$ ,  $M = \mathcal{M}(L_1, \{e_i\}_{i \in \{1, \dots, n+1\}}, \{e_i\}_{i \in \{1, \dots, n+1\}})$ ,  $\det([n+1, n+1]M) \neq 0$  (10.129)

Using 10.125 and 10.129 we have then that

$$\exists T \text{ a elementair transformation such that } L = T \circ L_1 \quad (10.130)$$

and

$$\text{If } M = \mathcal{M}(L_1, \{e_1\}_{i \in \{1, \dots, n+1\}}, \{e_i\}_{i \in \{1, \dots, n+1\}}) \text{ then } \det([n+1 \boxplus n+1]M) \neq 0 \quad (10.131)$$

Take now  $Y = \mathcal{S}(\{e_1, \dots, e_n\})$  (see 10.130) a  $n$ -dimensional subspace of  $X$  take then

$$L'_2 \in \text{Hom}(Y, Y) \text{ defined by } \forall i \in \{1, \dots, n\} \models L'_2(e_i) = \sum_{k \in \{1, \dots, n\}} M_{k,i} \cdot e_k \quad (10.132)$$

then we have  $\det(L'_2) = \det([n+1 \boxplus n+1]M) \neq 0$  so as  $n \in S$  there exists a family of elementair transformations  $\{T'_i\}_{i \in \{1, \dots, m\}}$  such that  $L'_2 = T'_m \circ \dots \circ T'_1$ . If we take then  $L_2 = (L_2)_{\langle n+1 \rangle}$  and  $\forall i \in \{1, \dots, m\} T_i = (T'_i)_{\langle n+1 \rangle}$  then using 10.352 we have  $L_2 = T_m \circ \dots \circ T_1$  giving

If  $L_2$  defined by  $\forall i \in \{1, \dots, n+1\} \models L_2(e_i) = \begin{cases} \sum_{k \in \{1, \dots, n\}} M_{k,i} \cdot e_k \\ L(e_{n+1}) = e_{n+1} \end{cases}$  then there exists a family  $\{T_i\}_{i \in \{1, \dots, m\}}$  of elementair functions in  $X$  such that  $L_2 = T_m \circ \dots \circ T_1$  (10.133)

Define now  $L_3$  by

$$L_3 = (\beta_{n+1}(n+1, n, M_{n,n+1}) \circ \dots \circ \beta_{n+1}(n+1, 1, M_{1,n+1})) \circ L_2 \quad (10.134)$$

then we have for  $i \in \{1, \dots, n\}$  that

$$\begin{aligned} L_3(e_i) &= (\beta_{n+1}(n+1, n, M_{n,n+1}) \circ \dots \circ \beta_{n+1}(n+1, 1, M_{1,n+1}))(L_2(e_i)) \\ &= (\beta_{n+1}(n+1, n, M_{n,n+1}) \circ \dots \circ \beta_{n+1}(n+1, 1, M_{1,n+1})) \left( \sum_{k \in \{1, \dots, n\}} M_{k,i} \cdot e_k \right) \\ &= \sum_{k \in \{1, \dots, n\}} M_{k,i} \cdot (\beta_{n+1}(n+1, n, M_{n,n+1}) \circ \dots \circ \beta_{n+1}(n+1, 1, M_{1,n+1}))(e_k) \\ &\stackrel{10.354}{=} \sum_{k \in \{1, \dots, n\}} M_{k,i} \cdot e_i \end{aligned}$$

and for  $n+1$  we have

$$\begin{aligned} L_3(e_{n+1}) &= (\beta_{n+1}(n+1, n, M_{n,n+1}) \circ \dots \circ \beta_{n+1}(n+1, 1, M_{1,n+1}))(e_{n+1}) \\ &\stackrel{10.355}{=} \sum_{k \in \{1, \dots, n\}} M_{k,n+1} \cdot e_k + e_{n+1} \end{aligned}$$

which proves that

$$\forall i \in \{1, \dots, n+1\} \text{ we have } L_3(e_i) = \begin{cases} \sum_{k \in \{1, \dots, n\}} M_{k,i} \cdot e_k \text{ if } i \in \{1, \dots, n\} \\ \left( \sum_{k \in \{1, \dots, n\}} M_{k,n+1} \cdot e_k \right) + e_{n+1} \end{cases} \quad (10.135)$$

Next take  $\alpha \in \mathbb{R}$  define then

$$L_4 = \sigma_{n+1}(n+1, \alpha) \circ L_3 \quad (10.136)$$

Then we have  $\forall i \in \{1, \dots, n\}$  that

$$\begin{aligned} L_4(e_i) &= \sigma_{n+1}(n+1, \alpha)(L_3(e_i)) \\ &\stackrel{10.135}{=} \sigma_{n+1}(n+1, \alpha) \left( \sum_{k \in \{1, \dots, n\}} M_{k,i} \cdot e_k \right) \\ &= \sum_{k \in \{1, \dots, n\}} M_{k,i} \cdot \sigma_{n+1}(n+1, \alpha)(e_k) \\ &= \sum_{k \in \{1, \dots, n\}} M_{k,i} \cdot e_k \end{aligned}$$

and for  $n+1$  we have

$$\begin{aligned} L_4(e_{n+1}) &= \sigma_{n+1}(n+1, \alpha)(L_3(e_{n+1})) \\ &\stackrel{10.135}{=} \sigma_{n+1}(n+1, \alpha) \left( \left( \sum_{k \in \{1, \dots, n\}} M_{k,n+1} \cdot e_k \right) + e_{n+1} \right) \\ &= \left( \sum_{k \in \{1, \dots, n\}} M_{k,n+1} \cdot \sigma_{n+1}(n+1, \alpha)(e_k) \right) + \sigma_{n+1}(n+1, \alpha)(e_{n+1}) \\ &= \sum_{k \in \{1, \dots, n\}} M_{k,n+1} \cdot e_k + \alpha \cdot e_{n+1} \end{aligned}$$

which proves that

$$\forall i \in \{1, \dots, n+1\} \models L_4(e_i) = \begin{cases} \sum_{k \in \{1, \dots, n\}} M_{k,i} \cdot e_k & \text{if } i \in \{1, \dots, n\} \\ (\sum_{k \in \{1, \dots, n\}} M_{k,n+1} \cdot e_k) + \alpha \cdot e_{n+1} & \text{if } i = n+1 \end{cases} \quad (10.137)$$

Next given  $B = M(n \times 1, F)$  (we will define  $B$  later) and take

$$L_5 = (\beta_{n+1}(n, n+1, B_n) \circ \dots \circ \beta_{n+1}(1, n+1, B_1)) \circ L_4 \quad (10.138)$$

then we have  $\forall i \in \{1, \dots, n\}$  that

$$\begin{aligned} L_5(e_i) &= (\beta_{n+1}(n, n+1, B_n) \circ \dots \circ \beta_{n+1}(1, n+1, B_1))(L_4(e_i)) \\ &\stackrel{10.137}{=} (\beta_{n+1}(n, n+1, B_n) \circ \dots \circ \beta_{n+1}(1, n+1, B_1)) \left( \sum_{k \in \{1, \dots, n\}} M_{k,i} \cdot e_k \right) \\ &= \sum_{k \in \{1, \dots, n\}} M_{k,i} \cdot (\beta_{n+1}(n, n+1, B_n) \circ \dots \circ \beta_{n+1}(1, n+1, B_1))(e_k) \\ &\stackrel{10.357}{=} \sum_{k \in \{1, \dots, n\}} M_{k,i} \cdot (e_k + B_k \cdot e_{n+1}) \\ &= \left( \sum_{k \in \{1, \dots, n\}} M_{k,i} \cdot e_k \right) + \left( \sum_{k \in \{1, \dots, n\}} M_{k,i} \cdot B_k \right) \cdot e_{n+1} \end{aligned}$$

and for  $n+1$

$$\begin{aligned} L_5(e_{n+1}) &= (\beta_{n+1}(n, n+1, B_n) \circ \dots \circ \beta_{n+1}(1, n+1, B_1))(L_4(e_{n+1})) \\ &\stackrel{10.137}{=} (\beta_{n+1}(n, n+1, B_n) \circ \dots \circ \beta_{n+1}(1, n+1, B_1)) \left( \left( \sum_{k \in \{1, \dots, n\}} M_{k,n+1} \cdot e_k \right) + \alpha \cdot e_{n+1} \right) \\ &= \left( \sum_{k \in \{1, \dots, n\}} M_{k,n+1} \cdot (\beta_{n+1}(n, n+1, B_n) \circ \dots \circ \beta_{n+1}(1, n+1, B_1))(\beta_{n+1}(n, n+1, B_n) \circ \dots \circ \beta_{n+1}(1, n+1, B_1))(e_k) \right) + \alpha \cdot (\beta_{n+1}(n, n+1, B_n) \circ \dots \circ \beta_{n+1}(1, n+1, B_1))(e_{n+1}) \\ &\stackrel{10.357}{=} \left( \sum_{k \in \{1, \dots, n\}} M_{k,n+1} \cdot (e_k + B_k \cdot e_{n+1}) \right) + \alpha \cdot (\beta_{n+1}(n, n+1, B_n) \circ \dots \circ \beta_{n+1}(1, n+1, B_1))(e_{n+1}) \\ &\stackrel{10.356}{=} \left( \sum_{k \in \{1, \dots, n\}} M_{k,n+1} \cdot (e_k + B_k \cdot e_{n+1}) \right) + \alpha \cdot e_{n+1} \\ &= \left( \sum_{k \in \{1, \dots, n\}} M_{k,n+1} \cdot e_k \right) + \left( \alpha + \sum_{k \in \{1, \dots, n\}} M_{k,n+1} \cdot B_k \right) \cdot e_{n+1} \end{aligned}$$

proving that  $\forall i \in \{1, \dots, n+1\}$  we have

$$\begin{aligned} L_5(e_i) &= \\ &\begin{cases} \left( \sum_{k \in \{1, \dots, n\}} M_{k,i} \cdot e_k \right) + \left( \sum_{k \in \{1, \dots, n\}} M_{k,i} \cdot B_k \right) \cdot e_{n+1} & \text{if } i \in \{1, \dots, n\} \\ \left( \sum_{k \in \{1, \dots, n\}} M_{k,n+1} \cdot e_k \right) + \left( \alpha + \sum_{k \in \{1, \dots, n\}} M_{k,n+1} \cdot B_k \right) \cdot e_{n+1} & \text{if } i = n+1 \end{cases} \quad (10.139) \end{aligned}$$

Now  $\sum_{k \in \{1, \dots, n\}} M_{k,i} \cdot B_k = \sum_{k \in \{1, \dots, n\}} M_{i,k}^T \cdot B_k$  is the matrix product  $([n+1 \boxplus n+1]M)^T \cdot B$ , we want this product to be  $A \in M(n \times 1, F)$  where  $\forall i \in \{1, \dots, n\} A_i = M_{n+1,i}$  so that we have the correct factor for  $e_{n+1}$  in 10.139, so we must have

$$([n+1 \boxplus n+1]M)^T \cdot B = A \quad (10.140)$$

Now as  $\det(([n+1 \boxplus n+1]M)^T) \underset{10.320}{=} \det([n+1 \boxplus n+1]M) \neq 0$  [by using 10.131] there exists a inverse of  $([n+1 \boxplus n+1]M)^T$ , take then  $N = (([n+1 \boxplus n+1]M)^T)^{-1}$  so that

$$\forall i, j \in \{1, \dots, n\} \text{ we have } \delta_{i,j} = \sum_{k \in \{1, \dots, n\}} (M^T)_{i,k} \cdot N_{k,j} = \sum_{k \in \{1, \dots, n\}} M_{k,i} \cdot N_{k,j} \quad (10.141)$$

Multiplying then by  $N$  in 10.140 gives  $N \cdot A = N \cdot ([n+1 \boxplus n+1]M)^T \cdot B = (([n+1 \boxplus n+1]M)^T)^{-1} \cdot ([n+1 \boxplus n+1]M)^T \cdot B = B \Rightarrow N \cdot A = B$ , so we take  $B$  to be defined by

$$\forall i \in \{1, \dots, n\} \models B_i = \sum_{l \in \{1, \dots, n\}} N_{i,l} \cdot M_{n+1,l} \quad (10.142)$$

Then we have  $\forall i \in \{1, \dots, n\}$  that

$$\begin{aligned} \sum_{k \in \{1, \dots, n\}} M_{k,i} \cdot B_k &= \sum_{k \in \{1, \dots, n\}} M_{k,i} \cdot \left( \sum_{l \in \{1, \dots, n\}} N_{k,l} \cdot M_{n+1,l} \right) \\ &= \sum_{k \in \{1, \dots, n\}} \left( \sum_{l \in \{1, \dots, n\}} (M_{k,i} \cdot N_{k,l}) \cdot M_{n+1,l} \right) \\ &\underset{10.47}{=} \sum_{l \in \{1, \dots, n\}} \left( \sum_{k \in \{1, \dots, n\}} M_{k,i} \cdot N_{k,l} \right) \cdot M_{n+1,l} \\ &\underset{10.141}{=} \sum_{l \in \{1, \dots, n\}} \delta_{i,l} \cdot M_{n+1,l} \\ &= M_{n+1,i} \end{aligned}$$

Using the above in 10.139 gives that

$$\forall i \in \{1, \dots, n\} \models L_5(e_i) = \sum_{k \in \{1, \dots, n+1\}} M_{k,i} \cdot e_k \quad (10.143)$$

Take now  $\alpha = M_{n+1,n+1} - \sum_{k \in \{1, \dots, n\}} M_{k,n+1} \cdot B_k$  then we have using 10.139 that

$$L_5(e_{n+1}) = \sum_{k \in \{1, \dots, n+1\}} M_{k,n+1} \cdot e_k \quad (10.144)$$

So we have that  $\mathcal{M}(L_5, \{e_i\}_{i \in \{1, \dots, n+1\}}, \{e_i\}_{i \in \{1, \dots, n+1\}}) = \mathcal{M}(L_1, \{e_i\}_{i \in \{1, \dots, n+1\}}, \{e_i\}_{i \in \{1, \dots, n+1\}}) \underset{10.309}{\Rightarrow} L_5 = L_1$  which proves by 10.130, 10.133, 10.134, 10.134, 10.136 and 10.138 that  $L = T \circ (\beta_{n+1}(n, n+1, B_n) \circ \dots \circ \beta_{n+1}(1, n+1, B_1)) \circ \sigma_{n+1}(n+1, \alpha) \circ (\beta_{n+1}(n+1, n, M_{n,n+1}) \circ \dots \circ \beta_{n+1}(n+1, 1, M_{1,n+1})) \circ (T_m \circ \dots \circ T_1)$ . Define now  $\forall i \in \{1, \dots, 2+m+2 \cdot n\}$  the elementair transformation  $E_i$  by

$$E_i = \begin{cases} T_i \text{ if } i \in \{1, \dots, m\} \\ \beta_{n+1}(n+1, i-m, M_{i-m,n+1}) \text{ if } i \in \{m+1, \dots, m+n\} \\ \sigma_{n+1}(n+1, \alpha) \text{ if } i = m+n+1 \\ \beta_{n+1}(i-(m+n+1), n+1, B_{(i-(m+n+1))}) \text{ if } i \in \{m+n+2, \dots, m+2 \cdot n+1\} \\ T \text{ if } i = m+2 \cdot n+2 \end{cases}$$

then we have  $L = E_{m+2 \cdot n+2} \circ (E_{m+2 \cdot n+1}, \dots, E_{m+n+2}) \circ E_{m+n+1} \circ (E_{m+n} \circ \dots \circ E_{m+1}) \circ (E_m \circ \dots \circ E_1)$ , repeately applying the general associativity of the composition of functions (see 4.81) gives then finally

$$L = E_{m+2 \cdot n+2} \circ \dots \circ E_1$$

Using 10.353 we have as  $L$  is regular that then also every  $E_i$  is a regular transformation, so that we have proved that  $L$  is the composition of regular elementair transformations and thus that  $n+1 \in S$ .

Using mathematical induction we have then that  $S = \mathbb{N}_0$  which proves the theorem.  $\square$

## 10.11 Direct internal sum of sub spaces

We extend now the concept of a direct sum introduced in 10.194 to a direct sum of a arbitrary finite number of sub spaces. First we consider sum's of vector spaces.

**Definition 10.359.** If  $V$  is a vector space over a field  $F$  and  $\{V_i\}_{i \in I}$  is a **finite** family of non empty sub spaces then  $\sum_{i \in I} V_i = \{v \in V \mid \exists \{v_i\}_{i \in I} \text{ with } \forall i \in I \text{ we have } v_i \in V_i \text{ such that } v = \sum_{i \in I} v_i\}$ .

**Example 10.360.** If  $V$  is a vector space over a field  $F$  and  $\emptyset = \{V_i\}_{i \in \emptyset}$  is the empty family then  $\sum_{i \in \emptyset} V_i = \{0\}$

**Proof.** If  $x \in \sum_{i \in \emptyset} V_i$  then there exists a  $\{v_i\}_{i \in \emptyset} = \emptyset$ , a family where  $\forall i \in \emptyset \models v_i \in V_i$  is satisfied vacuously where  $x = \sum_{i \in \emptyset} v_i$ . As  $\text{support}(\{v_i\}_{i \in \emptyset}) = \emptyset$  we have  $\sum_{i \in \emptyset} v_i = 0$  giving that  $x = \{0\}$   $\square$

**Example 10.361.** If  $V$  is a vector space then if  $\{V_k\}_{k \in \{k\}}$  is a finite family of sub spaces then  $\sum_{k \in \{k\}} V_k = V_k$

**Proof.** If  $v \in \sum_{k \in \{k\}} V_k$  then there exists a  $\{v_i\}_{i \in \{k\}}$  such that  $v = \sum_{i \in \{k\}} v_i = v_k \in V_k$ . If  $v \in V_k$  then if we take  $\{v_i\}_{i \in \{k\}}$  where  $v_k = v$  then we have  $v = v_k = \sum_{i \in \{k\}} v_i \in \sum_{i \in \{k\}} V_i$   $\square$

As the sum of sub spaces is based on the generalized sum a lot of its properties transfer to the sum of sub spaces. This is expressed in the next theorems.

**Theorem 10.362.** If  $V$  is vector space over a field  $F$  and  $\{V_i\}_{i \in I}$  is a **finite** family of non empty sub spaces then if  $h: J \rightarrow I$  is a bijection then we have that  $\sum_{j \in J} V_{h(j)} = \sum_{i \in I} V_i$

**Proof.** If  $v \in \sum_{j \in J} V_{h(j)}$  then  $v = \sum_{j \in J} v_j$  where  $\forall j \in J$  we have  $v_j \in V_{h(j)}$  and if we define  $\{w_i\}_{i \in I}$  by  $w_i = v_{h^{-1}(i)}$  then  $w_{h(j)} = v_{h^{-1}(h(j))} = v_j \in V_{h(j)}$  and  $v = \sum_{j \in J} w_{h(j)} \stackrel{10.43}{=} \sum_{i \in I} w_i \in \sum_{i \in I} V_i$  proving that  $\sum_{j \in J} V_{h(j)} \subseteq \sum_{i \in I} V_i$ .

If  $v \in \sum_{i \in I} V_i$  then  $v = \sum_{i \in I} v_i$  where  $\forall i \in I$  we have  $v_i \in V_i$ . Define then  $\{w_j\}_{j \in J}$  by  $\forall j \in J$   $w_j = v_{h(j)} \in V_{h(j)}$  then  $w_{h^{-1}(i)} = v_{h(h^{-1}(i))} = v_i$  which as  $h^{-1}$  is a bijection we can use 10.43 again to prove that  $v = \sum_{i \in I} w_{h^{-1}(i)} = \sum_{j \in J} w_j \in \sum_{j \in J} V_{h(j)}$  proving that  $\sum_{i \in I} V_i \subseteq \sum_{j \in J} V_{h(j)}$ .  $\square$

**Theorem 10.363.** If  $V$  is vector space over a field  $F$  and  $\{V_i\}_{i \in I}$  is a **finite** family of non empty sub spaces then if  $i \in I$  we have that  $\sum_{i \in I} V_i = (\sum_{j \in I \setminus \{i\}} V_j) + V_i$

**Proof.** If  $v \in \sum_{j \in I} V_j$  then  $v = \sum_{j \in I} v_j$  where  $\forall j \in I$  we have  $v_j \in V_j$ . From this it follows that  $v = \sum_{j \in I} v_j = \sum_{j \in I \setminus \{i\}} v_j + \sum_{j \in \{i\}} v_j = \sum_{j \in I \setminus \{i\}} v_j + v_i \in (\sum_{j \in I \setminus \{i\}} V_j) + V_i$  proving that  $\sum_{i \in I} V_i \subseteq (\sum_{j \in I \setminus \{i\}} V_j) + V_i$ .

If  $v \in (\sum_{j \in I \setminus \{i\}} V_j) + V_i$  then there exists a  $\{v_j\}_{j \in I \setminus \{i\}}$  and a  $w \in V_i$  such that  $v = \sum_{j \in I \setminus \{i\}} v_j + w$  define now  $\{w_j\}_{j \in I}$  by  $w_j = \begin{cases} v_j & \text{if } j \in I \setminus \{i\} \\ w & \text{if } j = i \end{cases}$  then  $v = \sum_{j \in I \setminus \{i\}} v_j + w = \sum_{j \in I \setminus \{i\}} w_j + w_i = \sum_{j \in I \setminus \{i\}} w_j + \sum_{j \in \{i\}} w_j = \sum_{j \in I} w_j \in \sum_{j \in I} V_j$  proving that  $(\sum_{j \in I \setminus \{i\}} V_j) + V_i \subseteq \sum_{j \in I} V_j$ .  $\square$

We can even generalize the above theorem

**Theorem 10.364.** *If  $V$  is a vector space over a field  $F$  and  $\{V_i\}_{i \in I}$  is a **finite** family of non empty sub spaces then if  $I = I_1 \cup I_2$  where  $I_1 \cap I_2$  then we have  $\sum_{i \in I} V_i = (\sum_{i \in I_1} V_i) + (\sum_{i \in I_2} V_i)$*

**Proof.** If  $v \in \sum_{j \in I} V_j$  then  $v = \sum_{j \in I} v_j$  where  $\forall j \in I$  we have  $v_j \in V_j$ . From this it follows that  $v = \sum_{j \in I} v_j = \sum_{j \in I_1} v_j + \sum_{j \in I_2} v_j \in \sum_{j \in I_1} V_j + \sum_{j \in I_2} V_j$  proving that  $\sum_{j \in I} V_j \subseteq \sum_{j \in I_1} V_j + \sum_{j \in I_2} V_j$ .

If  $v \in \sum_{j \in I_1} V_j + \sum_{j \in I_2} V_j$  then  $v = u + w$  where  $u = \sum_{j \in I_1} u_j$  and  $w = \sum_{j \in I_2} w_j$ . Define now (using  $I_1 \cap I_2$ )  $\{v_i\}_{i \in I}$  by  $v_i = \begin{cases} u_i & \text{if } i \in I_1 \\ w_i & \text{if } i \in I_2 \end{cases}$  then  $\sum_{j \in I} V_j \ni \sum_{j \in I} v_j = \sum_{j \in I_1} v_j + \sum_{j \in I_2} v_j = \sum_{j \in I_1} u_j + \sum_{j \in I_2} w_j = v$  proving that  $\sum_{j \in I_1} V_j + \sum_{j \in I_2} V_j \subseteq \sum_{j \in I} V_j$   $\square$

**Theorem 10.365.** *If  $V$  is a vector space over a field  $F$  and  $\{V_i\}_{i \in I}$  is a **finite** family of non empty sub spaces then  $\sum_{i \in I} V_i$  is a non empty subspace of  $V$  and thus a vector space.*

**Proof.** First as  $0 = \sum_{i \in I} 0$  we have that  $0 \in \sum_{i \in I} V_i$  and thus  $\sum_{i \in I} V_i \neq \emptyset$ . Second if  $\alpha, \beta \in F, x, y \in \sum_{i \in I} V_i$  then there exists  $\{x_i\}_{i \in I}, \{y_i\}_{i \in I}$  such that  $x = \sum_{i \in I} x_i, y = \sum_{i \in I} y_i$  and thus we have that

$$\begin{aligned} \alpha \cdot x + \beta \cdot y &= \alpha \cdot \left( \sum_{i \in I} x_i \right) + \beta \cdot \left( \sum_{i \in I} y_i \right) \\ &= \left( \sum_{i \in I} (\alpha \cdot x_i) \right) + \left( \sum_{i \in I} (\beta \cdot y_i) \right) \\ &= \sum_{i \in I} (\alpha \cdot x_i + \beta \cdot y_i) \\ &\in \sum_{i \in I} V_i \text{ (as } \alpha x_i + \beta y_i \in V_i) \end{aligned}$$

$\square$

**Definition 10.366.** *If  $V$  is a vector space over a field  $F$  then  $V = \sum_{i \in I}^\oplus V_i$  if  $\{V_i\}_{i \in I}$  is a **finite** family of non empty sub spaces and  $\forall v \in V$  there exists a **unique** family  $\{v_i\}_{i \in I}$  in  $V$  such that  $\forall i \in I$  we have  $v_i \in V_i$  and  $v = \sum_{i \in I} v_i$  (note that the sum exists because  $I$  is finite).*

*If  $I = \{1, \dots, n\}$  then we sometimes write  $V = V_1 \oplus \dots \oplus V_n$  instead of  $V = \sum_{i \in I}^\oplus V_i$  and in the particular case of  $I = \{1, \dots, 2\}$  as  $V = V_1 \oplus V_2$*

**Example 10.367.** If  $V$  is a vector space over a field  $V$  then  $\{V_i\}_{i \in \{1\}}$  where  $V_1 = V$  is a finite family of non empty sub spaces such that  $\forall v \in V$  the unique  $\{v_i\}_{i \in \{1\}}$  where  $v = v_1$  has  $v = v_1 = \sum_{i \in \{1\}} v_i$  so that  $V = \sum_{i \in \{1\}}^\oplus V_i$

**Theorem 10.368.** *If  $V$  is a vector space over a field  $F$  and  $\{V_i\}_{i \in I}$  is a finite family of non empty sub spaces of  $V$  then  $V = \sum_{i \in I}^\oplus V_i$  if and only if  $V = \sum_{i \in I} V_i$  and  $\forall i \in I$  we have that  $V_i \cap (\sum_{j \in I \setminus \{i\}} V_j) = \{0\}$  where  $0$  is the neutral element of  $V$*

**Proof.**

1.  $(\Rightarrow)$  As  $V = \sum_{i \in I}^\oplus V_i$  we have by definition that  $V \subseteq \sum_{i \in I} V_i \subseteq V \Rightarrow V = \sum_{i \in I} V_i$ . Let  $i \in I$  and take  $v \in V_i \cap (\sum_{j \in I \setminus \{i\}} V_j)$  then  $v \in V_i$  and there exists a  $\{v_j\}_{j \in I \setminus \{i\}}$  such that  $\forall j \in I \setminus \{i\}$  we have  $v_j \in V_j$  and  $v = \sum_{j \in I \setminus \{i\}} v_j$ . Define now  $\{w_j\}_{j \in I}$  where  $w_j = \begin{cases} v_j & \text{if } j \in I \setminus \{i\} \\ -v & \text{if } j = i \end{cases}$  then we have that  $\sum_{j \in I} w_j = (\sum_{j \in I \setminus \{i\}} w_j) + (\sum_{j \in \{i\}} w_j) = (\sum_{j \in I \setminus \{i\}} v_j) + (-v) = v + (-v) = 0$  then as  $0 = \sum_{j \in I} 0$  we have by the uniqueness of the expansion of  $0$  that  $\forall j \in I$  we have that  $w_j = 0$  and thus that  $v = -w_i = -0 = 0$ .

2. ( $\Leftarrow$ ) As  $V = \sum_{i \in I} V_i$  we have if  $v \in V$  that there exists a  $\{v_i\}_{i \in I}$  with  $\forall i \in I \models v_i \in V_i$  and  $v = \sum_{i \in I} v_i$ . Suppose now that  $\{w_i\}_{i \in I}$  is another family with  $\forall i \in I \models w_i \in V_i$  and  $v = \sum_{i \in I} w_i$  which is different from  $\{v_i\}_{i \in I}$  so that  $\exists i_0 \in I$  such that  $v_{i_0} \neq w_{i_0}$ . then we have that  $0 = v + (-v) = (\sum_{i \in I} v_i) + (-\sum_{i \in I} w_i) = \sum_{i \in I} (v_i - w_i) = (\sum_{i \in I \setminus \{i_0\}} (v_i - w_i)) + (\sum_{i \in \{i_0\}} (v_i - w_i)) = (\sum_{i \in I \setminus \{i_0\}} (v_i - w_i)) + (v_{i_0} - w_{i_0}) \Rightarrow w_{i_0} - v_{i_0} = \sum_{i \in I \setminus \{i_0\}} (v_i - w_i) \in \sum_{i \in I \setminus \{i_0\}} (v_i - w_i) \Rightarrow w_{i_0} - v_{i_0} \in V_{i_0} \cap (\sum_{i \in I \setminus \{i_0\}} V_i)$  which by the assumption means that  $w_{i_0} - v_{i_0} = 0$  contradicting  $w_{i_0} \neq v_{i_0}$ . This proves then by contradiction that  $\forall i \in I$  we have  $v_i = w_i$  proving uniqueness of the expansion.  $\square$

**Theorem 10.369.** If  $V$  is a vector space over a field,  $\{V_i\}_{i \in I}$  a family of sub spaces such that  $V = \sum_{i \in I}^{\oplus} V_{h(i)}$  where  $h: J \rightarrow I$  is a bijection then  $\sum_{i \in I}^{\oplus} V_{h(i)} = \sum_{j \in J}^{\oplus} V_j$

**Proof.** By the definition of direct sum we must have  $V = \sum_{i \in I} V_{h(i)}$  and then by 10.362 we have  $V = \sum_{j \in J} V_j$ . If now  $j \in J$  and  $V_j \cap \sum_{i \in J \setminus \{j\}} V_i$ . Define  $i_j = h^{-1}(j) \Rightarrow h(i_j) = j$  and take then  $f = h|_{I \setminus \{i_j\}}: I \setminus \{i_j\} \rightarrow J \setminus \{j\}$  then  $f$  is a bijection and we have  $\sum_{i \in I \setminus \{i_j\}} V_{h(i)} = \sum_{j \in J \setminus \{j\}} V_{f(j)}$   $\stackrel{10.362}{=} \sum_{j \in J \setminus \{j\}} V_j$  so we have that  $V_j \cap \sum_{i \in J \setminus \{j\}} V_i = V_{h(i_j)} \cap \sum_{i \in I \setminus \{i_j\}} V_{h(i)} = \sum_{i \in I \setminus \{i_j\}} V_{h(i)} = \sum_{i \in I \setminus \{i_j\}} \{0\}$   $\square$

**Theorem 10.370.** If  $V$  is a vector space over  $F$  and  $\{V_i\}_{i \in I}$  is a finite family of non empty sub spaces of  $V$  then (as by 10.365  $\sum_{j \in I \setminus \{i\}} V_j$  is a vector space over  $F$  (with  $\{V_j\}_{j \in I \setminus \{i\}}$  as non empty subsets))  $\sum_{j \in I \setminus \{i\}} V_i = \sum_{j \in I \setminus \{i\}}^{\oplus} V_j$  and  $V = (\sum_{j \in I \setminus \{i\}}^{\oplus} V_j) \oplus V_i$

**Proof.**

1. ( $\sum_{j \in I \setminus \{i\}} V_i = \sum_{j \in I \setminus \{i\}}^{\oplus} V_i$ ) As we trivially have that  $\sum_{j \in I \setminus \{i\}} V_i = \sum_{j \in I \setminus \{i\}} V_i$  we must prove that  $\forall k \in I \setminus \{i\}$  we have that  $V_k \cap (\sum_{j \in (I \setminus \{i\}) \setminus \{k\}} V_j) = \{0\}$ . Now if  $x \in V_k \cap (\sum_{j \in (I \setminus \{i\}) \setminus \{k\}} V_j)$  then  $x \in V_k$  and there exists a  $\{x_j\}_{j \in (I \setminus \{i\}) \setminus \{k\}}$  such that  $\forall j \in (I \setminus \{i\}) \setminus \{k\}$  we have  $x_j \in V_j$  and  $x = \sum_{j \in (I \setminus \{i\}) \setminus \{k\}} x_j$ , if we define then  $\{y_j\}_{j \in I \setminus \{k\}}$  by  $y_j = \begin{cases} x_j & \text{if } j \neq i \\ 0 & \text{if } j = i \end{cases} \in V_j$  then  $x = \sum_{j \in (I \setminus \{i\}) \setminus \{k\}} x_j + 0 = \sum_{j \in (I \setminus \{i\}) \setminus \{k\}} y_j + y_i = \sum_{j \in I \setminus \{k\}} y_j \in \sum_{j \in I \setminus \{k\}} V_j$  giving that  $x \in V_k \cap (\sum_{j \in I \setminus \{k\}} V_j) = \{0\}$  so that  $x = 0$ .
2. ( $(V = (\sum_{j \in I \setminus \{i\}}^{\oplus} V_j) \oplus V_i)$  Using 10.363 we have that  $(\sum_{k \in I \setminus \{i\}} V_j) + V_i$  so we must prove that if  $(\sum_{j \in I \setminus \{i\}} V_j) \cap V_i = \{0\}$  which is true because  $V = \sum_{i \in I}^{\oplus} V_i$  by 10.368.  $\square$

**Lemma 10.371.** If  $V$  is a vector space over a field  $F$ ,  $V_1, V_2 \subseteq V$  non empty finite dimensional sub spaces of  $V$  (which may be trivial) then  $\dim(V_1 \oplus V_2) = \dim(V_1) + \dim(V_2)$

**Proof.** We have the following cases for  $V_1, V_2$  to consider:

**$V_1, V_2$  are trivial.** If  $V_1, V_2$  are trivial then  $V_1 = \{0\} = V_2$  hence  $V_1 \oplus V_2 = \{0\}$  so that  $\dim(V \oplus V_2) = 0 = 0 + 0 + \dim(V_1) + \dim(V_2)$

**$V_1$  is trivial,  $V_2$  is not trivial.** Then  $V_1 = \{0\}$  and  $V_2$  has a basis  $\{e_i\}_{i \in \{1, \dots, n\}}$ . If  $x \in V_1 \oplus V_2$  there exists a  $x_1 \in V_1, x_2 \in V_2$  such that  $x = x_1 + x_2 \stackrel{V_1 = \{0\}}{=} 0 + x_2$  proving that  $x \in V_2$  hence there exists a  $\{\alpha_i\}_{i \in \{1, \dots, n\}}$  such that  $x = \sum_{i=1}^n \alpha_i \cdot e_i$ . As  $\{e_i\}_{i \in \{1, \dots, n\}}$  are by definition linear independent we have that  $\{e_i\}_{i \in \{1, \dots, n\}}$  is a basis of  $V_1 \oplus V_2$ . So  $\dim(V_1 \oplus V_2) = n = \dim(V_2) = 0 + \dim(V_2) = \dim(V_1) + \dim(V_2)$ .

**$V_1$  is not trivial,  $V_2$  is trivial.** This is the same as the above if we interchange the roles of  $V_1$  and  $V_2$

**$V_1, V_2$  are not trivial.** Define then  $\varphi: \prod_{i \in \{1, \dots, 2\}} V_i \rightarrow V_1 \oplus V_2$  by  $\varphi(v_1, v_2) = v_1 + v_2$  then  $\varphi$  is a isomorphism:

**injectivity.** If  $\varphi(v_1, v_2) = \varphi(u_1, u_2) \Rightarrow v_1 + v_2 = u_1 + u_2 \underset{\text{uniqueness of sum}}{\Rightarrow} (v_1, v_2) = (u_1, u_2)$

**surjectivity.** If  $v \in V_1 \oplus V_2$  then there exists a  $v_1 \in V_1, v_2 \in V_2$  such that  $v = v_1 + v_2 = \varphi(v_1, v_2)$

**linearity.** If  $(v_1, v_2), (u_1, u_2) \in \prod_{i \in \{1, \dots, 2\}} V_i$  and  $\alpha, \beta \in F$  then  $\varphi(\alpha \cdot (v_1, v_2) + \beta \cdot (u_1, u_2)) = \varphi(\alpha \cdot v_1 + \beta \cdot u_1, \alpha \cdot v_2 + \beta \cdot u_2) = \alpha \cdot v_1 + \beta \cdot u_1 + \alpha \cdot v_2 + \beta \cdot u_2 = \alpha \cdot (v_1 + u_1) + \beta \cdot (v_2 + u_2) = \alpha \cdot \varphi(v_1, u_1) + \beta \cdot \varphi(v_2, u_2)$

Hence  $\dim(V_1 \oplus V_2) \underset{10.222}{=} \dim(\prod_{i \in \{1, \dots, 2\}} V_i) \underset{10.204}{=} \sum_{i=1}^2 \dim(V_i) = \dim(V_1) + \dim(V_2)$   $\square$

**Theorem 10.372.** If  $V$  is a vector spaces over a field  $F$ ,  $n \in \mathbb{N}_0 \setminus \{1\}$ , and  $\{V_i\}_{i \in \{1, \dots, n\}}$  a finite family of non empty finite dimensional such that  $V = V_1 \oplus \dots \oplus V_n$  then  $\dim(V) = \sum_{i=1}^n \dim(V_i)$

**Proof.** We prove this by induction on  $n$  so let  $\mathcal{S} = \{n \in \{2, \dots, \infty\} \mid \text{if } V \text{ is a vector space } \{V_i\}_{i \in \{1, \dots, n\}} \text{ finite dimensional non empty subspace of } V \text{ such that } V = V_1 \oplus \dots \oplus V_n \text{ then } \dim(V) = \sum_{i=1}^n \dim(V_i)\}$  then we have

**2  $\in \mathcal{S}$ .** This follows from the previous lemma (see 10.371)

**$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$ .** If  $\{V_i\}_{i \in \{1, \dots, n+1\}}$  is a family of non empty finite subspaces of  $V$  then we have by 10.370 that  $V = (V_1 \oplus \dots \oplus V_{n+1}) = (V_1 \oplus \dots \oplus V_n) \oplus V_{n+1}$  so that  $\dim(V) = \dim((V_1 \oplus \dots \oplus V_n) \oplus V_{n+1}) \underset{10.371}{=} \dim(V_1 \oplus \dots \oplus V_n) + \dim(V_{n+1}) \underset{n \in \mathcal{S}}{=} (\sum_{i=1}^n \dim(V_i)) + \dim(V_{n+1}) = \sum_{i=1}^{n+1} \dim(V_i)$  proving that  $n+1 \in \mathcal{S}$

Induction finishes then the proof.  $\square$

This theorem can be generalized also

**Theorem 10.373.** If  $V$  is a vector space over a field  $F$  and  $\{V_i\}_{i \in I}$  is a finite family of non empty sub spaces of  $V$  and  $I = I_1 \cup I_2, I_1 \cap I_2 = \emptyset$  (by 10.365 we have then that  $\sum_{i \in I_1} V_i, \sum_{i \in I_2} V_i$  are vector spaces over  $F$  with sub spaces  $\{V_i\}_{i \in I_1}, \{V_i\}_{i \in I_2}$ ). Then  $\sum_{i \in I_1} V_i = \sum_{i \in I_1}^{\oplus} V_i, \sum_{i \in I_2} V_i = \sum_{i \in I_2}^{\oplus} V_i$  (more general if  $J \subseteq I$  then  $\sum_{j \in J}^{\oplus} V_j$  exist) and  $V = (\sum_{i \in I_1}^{\oplus} V_i) \oplus (\sum_{i \in I_2}^{\oplus} V_i)$

**Proof.**

1. Let  $J \subseteq I$  and prove that  $\sum_{j \in J} V_j = \sum_{j \in J}^{\oplus} V_j$  then as we trivially have  $\sum_{j \in J} V_j = \sum_{j \in J}^{\oplus} V_j$  we must prove that  $\forall i \in J$  we have  $V_i \cap (\sum_{j \in J \setminus \{i\}} V_j) = \{0\}$ . Now if  $x \in V_i \cap (\sum_{j \in J \setminus \{i\}} V_j)$  then  $x \in V_i$  and  $x = \sum_{j \in J \setminus \{i\}} x_j$  where for the family  $\{x_j\}_{j \in J \setminus \{i\}}$  we have  $x_j \in X_j$ . Define now  $\{y_j\}_{j \in J \setminus \{i\}}$  by  $y_j = \begin{cases} x_j & \text{if } j \in J \setminus \{i\} \\ 0 & \text{if } j \in (I \setminus \{i\}) \setminus (J \setminus \{i\}) \end{cases}$  then  $x = \sum_{j \in J \setminus \{i\}} x_j + 0 = \sum_{j \in J \setminus \{i\}} x_j + \sum_{j \in (I \setminus \{i\}) \setminus (J \setminus \{i\})} 0 = \sum_{j \in J \setminus \{i\}} y_j + \sum_{j \in (I \setminus \{i\}) \setminus (J \setminus \{i\})} y_j = \sum_{j \in I \setminus \{i\}} y_j \in \sum_{j \in I \setminus \{i\}} V_j$  giving that  $x \in V_i \cap (\sum_{j \in I \setminus \{i\}} V_j) = \{0\}$  giving that  $x = 0$  or  $V_i \cap (\sum_{j \in I \setminus \{i\}} V_j) = \{0\}$ . Applying the above for  $J = I_1, I_2$  proves the first assertion of the theorem.

2. Using 10.364 we have that  $V = \sum_{j \in I}^{\oplus} V_j \underset{\text{definition of direct sum}}{=} \sum_{j \in I} V_j = \sum_{j \in I_1} V_j + \sum_{j \in I_2} V_j$ . So we must only prove that  $(\sum_{j \in I_1} V_j) \cap (\sum_{j \in I_2} V_j) = \{0\}$  and do this using proof by contradiction. If  $x \in (\sum_{j \in I_1} V_j) \cap (\sum_{j \in I_2} V_j)$  and  $x \neq 0$  then  $0 \neq x = \sum_{i \in I_1} x_i = \sum_{i \in I_2} y_i$  where  $\forall i \in I_1 \models x_i \in V_i, \forall i \in I_2 \models y_i \in V_i$ . As  $x \neq 0$  there exists a  $i \in I_1$  such that  $V_i \ni x_i \neq 0$  define now  $\{v_j\}_{j \in I \setminus \{i\}}$  by  $v_j = \begin{cases} -x_j & \text{if } j \in I_1 \setminus \{i\} \\ y_j & \text{if } j \in I_2 \end{cases}$  then  $\sum_{j \in I \setminus \{i\}} V_j \ni \sum_{j \in I \setminus \{i\}} v_j = \sum_{j \in I_1 \setminus \{i\}} v_j + \sum_{j \in I_2} v_j = \sum_{j \in I_1 \setminus \{i\}} (-x_j) + \sum_{j \in I_2} y_j = -\sum_{j \in I_1 \setminus \{i\}} x_j + x = -(\sum_{j \in I_1 \setminus \{i\}} x_j + x_i) + x_i + x = -x + x_i + x = x_i$  proving that  $0 \neq x_i \in V_i \cap (\sum_{j \in I \setminus \{i\}} V_j) = \{0\}$  which is a contradiction.  $\square$

**Theorem 10.374.** *If  $V$  is a vector space over a field  $F$  and  $\{V_i\}_{i \in I}$  is a finite family of non empty non trivial sub spaces of  $V$  and  $I = \bigcup_{i \in \{1, \dots, m\}} I_i$  where  $I_i$ 's are pairwise distinct then we have  $\sum_{i \in \{1, \dots, m\}} (\sum_{j \in I_i}^\oplus V_j) = \sum_{i \in \{1, \dots, m\}} (\sum_{j \in I_i}^\oplus V_j) = \sum_{i \in I}^\otimes V_i$  (if  $\sum_{i \in I}^\otimes V_i$  exist)*

**Proof.** The proof is by induction, let  $S = \{m \in \mathbb{N}_0 \mid \text{If } I = \bigcup_{i < \{1, \dots, m\}} I_i \text{ (pairwise disjoint union) and } \sum_{i \in I}^\oplus V_i \text{ exist then } \sum_{i \in I}^\oplus V_i = \sum_{i \in \{1, \dots, m\}} (\sum_{j \in I_i}^\oplus V_j)\}$  then we have

1. ( $m = 1$ ) Here  $I = I_1$  and  $\sum_{i \in \{1, \dots, 1\}} (\sum_{j \in I_i}^\oplus V_j) = \sum_{j \in I_1}^\oplus V_j = \sum_{j \in I}^\oplus V_j \Rightarrow 1 \in S$
2. ( $m \in S$ ) If  $I = \bigcup_{j \in \{1, \dots, m+1\}} I_j$  (pairwise disjoint union) and  $\sum_{j \in I}^\oplus V_j$  exist take then  $J = \bigcup_{i \in \{1, \dots, m\}} I_i$  then by the previous theorem  $\sum_{j \in J}^\oplus V_j$  exists and as  $I = J \bigcup I_{m+1}$  and trivially  $J \cap I_{m+1} = \emptyset$  we have

$$\begin{aligned} \sum_{i \in I}^\oplus V_i &= \left( \sum_{j \in J}^\oplus V_j \right) \oplus \left( \sum_{i \in I_{m+1}}^\oplus V_i \right) \\ &\stackrel{m \in S}{=} \left( \sum_{i \in \{1, \dots, m\}} \left( \sum_{j \in I_i}^\oplus V_i \right) \right) \oplus \left( \sum_{j \in I_{m+1}}^\oplus V_i \right) \\ &\stackrel{10.370}{=} \sum_{i \in \{1, \dots, m+1\}} \left( \sum_{j \in I_i}^\oplus V_i \right) \end{aligned}$$

proving that  $m+1 \in S$

By induction we have then that  $S = \mathbb{N}_0$  which proves the theorem. □

## 10.12 Tensor product of vector spaces

**Definition 10.375. (Tensor Product)** *If  $\{V_i\}_{i \in \{1, \dots, n\}}$  is a finite non empty family of vector spaces over a field  $F$  then a **tensor product** of  $\{V_i\}_{i \in \{1, \dots, n\}}$  is a pair  $\langle P, \nu \rangle$  where*

1.  $P$  is a vector space over the field  $F$
2.  $\nu \in \text{Hom}(V_1, \dots, V_n; P)$
3.  $\mathcal{S}(\nu(\prod_{i \in \{1, \dots, n\}} V_i)) = P$
4. If  $\varphi \in \text{Hom}(V_1, \dots, V_n; U)$  then there exists a  $h \in \text{Hom}(P, U)$  such that  $\varphi = h \circ \nu$

We prove now that the tensor product is uniquely determined to within a canonical isomorphism.

**Theorem 10.376.** *If  $\{V_i\}_{i \in \{1, \dots, n\}}$  is a finite non empty family of vector spaces over a field  $F$  and  $\langle P, \nu \rangle$  and  $\langle Q, \mu \rangle$  be tensor products of  $\{V_i\}_{i \in \{1, \dots, n\}}$  then there exists a **unique** bijective linear function (= linear isomorphism)  $k: P \rightarrow Q$  such that  $k \circ \nu = \mu$*

**Proof.** As  $\mu \in \text{Hom}(V_1, \dots, V_n; Q)$  we have by the fact that  $\langle P, \nu \rangle$  is a tensor product there exists a  $k \in \text{Hom}(P, Q)$  such that

$$k \circ \nu = \mu. \tag{10.145}$$

Also as  $\nu \in \text{Hom}(V_1, \dots, V_n; P)$  we have by the fact that  $\langle Q, \mu \rangle$  is a tensor product there exist a  $h \in \text{Hom}(Q, P)$  such that

$$h \circ \mu = \nu \tag{10.146}$$

Using 10.145 and 10.146 we have that

$$k \circ h \circ \mu = \mu \quad (10.147)$$

$$h \circ k \circ \nu = \nu \quad (10.148)$$

If now  $x \in P$  then as  $\mathcal{S}(\nu(\prod_{i \in \{1, \dots, n\}} V_i)) = P$  there exists a  $\{\alpha_i\}_{i \in \{1, \dots, k\}}$  in  $F$  and a  $\{y_i\}_{i \in \{1, \dots, k\}}$  in  $\nu(\prod_{i \in \{1, \dots, k\}} V_i)$  such that  $x = \sum_{i \in \{1, \dots, k\}} \alpha_i \cdot y_i$ , now  $\forall i \in \{1, \dots, k\}$  there exists a unique  $x_i$  such that  $\nu(x_i) = y_i$ , which defines a family  $\{x_i\}_{i \in \{1, \dots, k\}}$  in  $\prod_{i \in \{1, \dots, n\}} V_i$  such that  $x = \sum_{i \in \{1, \dots, k\}} \alpha_i \cdot \nu(x_i)$  so that  $(h \circ k)(x) = (h \circ k)(\sum_{i \in \{1, \dots, k\}} \alpha_i \cdot \nu(x_i)) \stackrel{h \circ k \text{ is linear}}{=} \sum_{i \in \{1, \dots, k\}} \alpha_i \cdot (h \circ k)(\nu(x_i)) = \sum_{i \in \{1, \dots, k\}} \alpha_i \cdot (h \circ k \circ \nu)(x_i) \stackrel{10.148}{=} \sum_{i \in \{1, \dots, k\}} \alpha_i \cdot \nu(x_i) = x$  proving that

$$h \circ k = 1_P \quad (10.149)$$

If now  $x \in Q$  then as  $\mathcal{S}(\mu(\prod_{i \in \{1, \dots, n\}} V_i)) = Q$  there exists a  $\{\alpha_i\}_{i \in \{1, \dots, k\}}$  in  $F$  and a  $\{y_i\}_{i \in \{1, \dots, k\}}$  in  $\mu(\prod_{i \in \{1, \dots, k\}} V_i)$  such that  $x = \sum_{i \in \{1, \dots, k\}} \alpha_i \cdot y_i$ , now  $\forall i \in \{1, \dots, k\}$  there exists a unique  $x_i$  such that  $\mu(x_i) = y_i$ , which defines a family  $\{x_i\}_{i \in \{1, \dots, k\}}$  in  $\prod_{i \in \{1, \dots, n\}} V_i$  such that  $x = \sum_{i \in \{1, \dots, k\}} \alpha_i \cdot \mu(x_i)$  so that  $(k \circ h)(x) = (k \circ h)(\sum_{i \in \{1, \dots, k\}} \alpha_i \cdot \mu(x_i)) \stackrel{k \circ h \text{ is linear}}{=} \sum_{i \in \{1, \dots, k\}} \alpha_i \cdot (k \circ h)(\mu(x_i)) = \sum_{i \in \{1, \dots, k\}} \alpha_i \cdot (k \circ h \circ \mu)(x_i) \stackrel{10.148}{=} \sum_{i \in \{1, \dots, k\}} \alpha_i \cdot \mu(x_i) = x$  proving that

$$k \circ h = 1_Q \quad (10.150)$$

This proves that  $k$  and  $h$  are bijections (and as they are linear) that  $k, h$  are linear isomorphism's.

To prove uniqueness, suppose that there is another bijection  $k' \in \text{Hom}(P, Q)$  such that  $k' \circ \nu = \mu$  then  $k' \circ \nu = k \circ \nu$ , again as  $x \in P = \mathcal{S}(\nu(\prod_{i \in \{1, \dots, n\}} V_i))$  there exists a  $\{\alpha_i\}_{i \in \{1, \dots, n\}}$  in  $F$  and a  $\{x_i\}_{i \in \{1, \dots, n\}}$  in  $\prod_{i \in \{1, \dots, n\}} V_i$  such that  $x = \sum_{i \in \{1, \dots, n\}} \alpha_i \cdot \nu(x_i)$  so that  $k'(x) = k'(\sum_{i \in \{1, \dots, n\}} \alpha_i \cdot \nu(x_i)) = \sum_{i \in \{1, \dots, n\}} \alpha_i \cdot k'(\nu(x_i)) \stackrel{k' \circ \nu = k \circ \nu}{=} \sum_{i \in \{1, \dots, n\}} \alpha_i \cdot k(\nu(x_i)) = k(\sum_{i \in \{1, \dots, n\}} \alpha_i \cdot \nu(x_i)) = k(x)$  so that we have  $k' = k$   $\square$

Because of the previous theorem it makes sense to introduce a special notation for the tensor product.

**Notation 10.377.** If  $\langle P, \nu \rangle$  is a tensor product of  $\{V_i\}_{i \in \{1, \dots, n\}}$  then we note  $P$  as  $V_1 \otimes \dots \otimes V_n$  or  $\otimes_{i \in \{1, \dots, n\}} V_i$  and  $\nu$  as  $\otimes$ . Also if  $(x_1, \dots, x_n) \in \prod_{i \in \{1, \dots, n\}} V_i$  then  $\nu(x)$  is noted as  $x_1 \otimes \dots \otimes x_n$ .

**Example 10.378.** If  $\{V_i\}_{i \in \{1\}}$  is a finite non empty family of vector spaces over a field  $F$  then  $\langle V_1, 1_{V_1} \rangle$  forms a tensor product.

**Proof.**

1.  $V_1$  is trivially a vector space over  $F$
2.  $1_{V_1} \in \text{Hom}(V_1; V_1)$  trivially
3.  $\mathcal{S}(1_{V_1}(\prod_{i \in \{1\}} V_i)) = \mathcal{S}(1_{V_1}(V_1)) = \mathcal{S}(V_1) = V_1$
4. If  $\varphi \in \text{Hom}(V_1, U)$  then  $\varphi \in \text{Hom}(V_1, U)$  satisfies  $\varphi = \varphi \circ 1_{V_1}$   $\square$

Next we prove a theorem that specifies the zero elements in tensor product considered as a vector space

**Theorem 10.379.** If  $\{V_i\}_{i \in \{1, \dots, m\}}$  is a finite non empty family of vector spaces over a field  $F$  with multiplicity zero then if  $(v_1, \dots, v_m) \in \prod_{i \in \{1, \dots, m\}} V_i$  we have  $(v_1 \otimes \dots \otimes v_m) = 0$  if and only if  $\exists i \in \{1, \dots, m\}$  with  $v_i = 0$

**Proof.**

( $\Rightarrow$ ) This is proved by contradiction, so let  $v_1 \otimes \dots \otimes v_m = 0$  with  $\forall i \in \{1, \dots, m\}$  we have  $v_i \neq 0$  then by 10.178 there exists linear mappings  $\varphi_i: V_i \rightarrow F$  with  $\varphi_i(v_i) \neq 0$ . This defines a multilinear map  $\varphi: \prod_{i \in \{1, \dots, m\}} V_i \rightarrow F$  defined by  $\varphi(v_1, \dots, v_m) = \prod_{i \in \{1, \dots, m\}} \varphi_i(v_i)$ , by the definition of a tensor space there exist then a linear mapping  $h: V_1 \otimes \dots \otimes V_m \rightarrow F$  such that  $f \circ \otimes = \varphi$ . So  $f(v_1 \otimes \dots \otimes v_m) = \varphi(v_1, \dots, v_m) = 1 \underset{v_1 \otimes \dots \otimes v_m = 0 \text{ and } h \text{ is linear}}{\Rightarrow} 0 = 1$  a contradiction as  $F$  has multiplicity of zero. So we must conclude that  $\exists i \in \{1, \dots, m\}$  such that  $v_i = 0$

( $\Leftarrow$ ) If  $\exists i \in \{1, \dots, m\}$  with  $v_i = 0$  then  $v_i = 0 \cdot v_i$  so that  $v_1 \otimes \dots \otimes v_m = v_1 \otimes \dots \otimes v_{i-1} \otimes (0 \cdot v_i) \otimes v_{i+1}, \dots, \otimes v_m = 0 \cdot (v_1 \otimes \dots \otimes v_m) = 0$   $\square$

**Theorem 10.380.** If  $\{V_i\}_{i \in \{1, \dots, n\}}$  is a finite non empty family of vector spaces over a field  $F$  and  $V_1 \otimes \dots \otimes V_n$  a tensor product then  $\forall x \in V_1 \otimes \dots \otimes V_n$  there exist a  $\{x_{1,i} \otimes \dots \otimes x_{n,i}\}_{i \in k}$  such that  $x = \sum_{i \in \{1, \dots, k\}} (x_{1,i} \otimes \dots \otimes x_{n,i})$  where  $k$  is minimal.

**Proof.** As  $V_1 \otimes \dots \otimes V_n = \mathcal{S}(\otimes(\prod_{i \in \{1, \dots, n\}} V_i))$  we have that for every  $x \in V_1 \otimes \dots \otimes V_n$  there exists a  $\{\lambda_i\}_{i \in \{1, \dots, k\}}$  and  $\{y_{1,i} \otimes \dots \otimes y_{n,i}\}_{i \in \{1, \dots, k\}}$  such that  $x = \sum_{i \in \{1, \dots, k\}} \lambda_i \cdot (y_{1,i} \otimes \dots \otimes y_{n,i})$  then using the multilinearity of  $\otimes$  and the fact that  $V_1$  is a vector space we have that  $\lambda_i \cdot (y_{1,i} \otimes \dots \otimes y_{n,i}) = (\lambda_i \cdot y_{1,i}) \otimes \dots \otimes y_{n,i} \in \otimes(\prod_{i \in \{1, \dots, n\}} V_i)$  so if  $x_{1,i} \otimes \dots \otimes x_{n,i} = (\lambda_i \cdot y_{1,i}) \otimes \dots \otimes y_{n,i} \in \otimes(\prod_{i \in \{1, \dots, n\}} V_i)$  we have then that  $x = \sum_{i \in \{1, \dots, k\}} x_{1,i} \otimes \dots \otimes x_{n,i}$ . So there exists a  $\{x_{1,i} \otimes \dots \otimes x_{n,i}\}_{i \in \{1, \dots, k\}}$  such that  $x = \sum_{i \in \{1, \dots, k\}} x_{1,i} \otimes \dots \otimes x_{n,i}$  proving that  $N = \{m \in \mathbb{N}_0 \mid \text{there exists a } \{x_{1,i} \otimes \dots \otimes x_{n,i}\}_{i \in \{1, \dots, m\}} \text{ such that } x = \sum_{i \in \{1, \dots, m\}} x_{1,i} \otimes \dots \otimes x_{n,i}\} \neq \emptyset$  so that  $k = \min(N)$  (see 4.52) exists so that there exists a  $\{x_{1,i} \otimes \dots \otimes x_{n,i}\}_{i \in \{1, \dots, k\}}$  so that  $x = \sum_{i \in \{1, \dots, k\}} (x_{1,i} \otimes \dots \otimes x_{n,i})$  proving the theorem.  $\square$

We proceed now to prove that there exists always a tensor product.

Remember that given a set  $X$  and a field  $\langle F, +, \cdot \rangle$  we can using 10.171 create a vector space  $\langle \mathcal{F}(X, F), +, \cdot \rangle$  over the field  $\langle F, +, \cdot \rangle$  called the **free vector space over  $X$**  such that  $\{\delta_a\}_{a \in X}$  forms a basis of  $\langle \mathcal{F}(X, F), +, \cdot \rangle$ . Or if  $F$  has characterization 0 we have that  $X$  is embedded in  $\mathcal{F}(X, F)$  by the injection  $\delta$  and forms the basis of  $\mathcal{F}(X, F)$ .

Given a finite family of vector spaces  $\{V_i\}_{i \in \{1, \dots, n\}}$  over a field  $F$  we define now the tensor product  $V_1 \otimes \dots \otimes V_n$  over the field  $F$  using the following steps.

First we form the free vector space over  $\prod_{i \in \{1, \dots, n\}} V_i$  (We consider this as just a set with no additional structure).

**Definition 10.381.** If  $\{V_i\}_{i \in \{1, \dots, n\}}$  is a finite non empty family of vector spaces over a field  $F$  define then  $\mathcal{N}(\prod_{i \in \{1, \dots, n\}} V_i, F) \subseteq \mathcal{F}(\prod_{i \in \{1, \dots, n\}} V_i, F)$  by  $\mathcal{N}(\prod_{i \in \{1, \dots, n\}} V_i, F) = \bigcup_{i \in \{1, \dots, n\}} \{\delta_{(x_1, \dots, x_{i-1}, \alpha \cdot x + \beta \cdot y)} - \alpha \cdot \delta_{(x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_n)} - \beta \cdot \delta_{(x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_n)} \mid (x_1, \dots, x_{i-1}, \alpha \cdot x + \beta \cdot y, x_{i+1}) \in \prod_{i \in \{1, \dots, n\}} V_i \wedge \alpha, \beta \in F\}$

We are now ready to define  $V_1 \otimes \dots \otimes V_n$

**Definition 10.382.** If  $\{V_i\}_{i \in \{1, \dots, n\}}$  is a finite family of vector spaces over a field  $F$  take then the free vector space  $\langle \mathcal{F}(\prod_{i \in \{1, \dots, n\}} V_i, +, \cdot) \rangle$  over the field  $F$  and the sub vector space  $\mathcal{S}(\mathcal{N}(\prod_{i \in \{1, \dots, n\}} V_i, F))$  then  $\langle V_1 \otimes \dots \otimes V_n, +, \cdot \rangle$  is defined to by the factor space (see 10.193)  $\langle \mathcal{F}(\prod_{i \in \{1, \dots, n\}} V_i, +, \cdot) / \mathcal{S}(\mathcal{N}(\prod_{i \in \{1, \dots, n\}} V_i, F)), +, \cdot \rangle$ .

**Definition 10.383.** If  $\{V_i\}_{i \in \{1, \dots, n\}}$  is a finite family of vector spaces over a field  $F$  then we define  $\otimes: \prod_{i \in \{1, \dots, n\}} \rightarrow V_1 \otimes \dots \otimes V_n$  by  $(v_1, \dots, v_n) \rightarrow v_1 \otimes \dots \otimes v_n = \pi(\delta(v_1, \dots, v_n))$  so that  $\otimes = \pi \circ \delta$  where we have defined

1.  $\delta: \prod_{i \in \{1, \dots, n\}} V_i \rightarrow \mathcal{F}(\prod_{i \in \{1, \dots, n\}} V_i)$  by  $(x_1, \dots, x_n) \rightarrow \delta_{(x_1, \dots, x_n)}$  and  $\delta_{(x_1, \dots, x_n)}(r_1, \dots, s) = \begin{cases} 1 & \text{if } (x_1, \dots, x_n) = (r_1, \dots, r_n) \\ 0 & \text{if } (x_1, \dots, x_n) \neq (r_1, \dots, r_n) \end{cases}$  (see 10.171)

2.  $\pi: \mathcal{F}(\prod_{i \in \{1, \dots, n\}} V_i, F) \rightarrow \mathcal{F}(\prod_{i \in \{1, \dots, n\}} V_i, F) / \mathcal{S}(\mathcal{N}(\prod_{i \in \{1, \dots, n\}} V_i, F))$  is the canonical surjection  $x \rightarrow \sim[x] = \{y | x - y \in \mathcal{S}(\mathcal{N}(\prod_{i \in \{1, \dots, n\}} V_i, F))\}$  which is linear by the definition of sum and scalar product on the quotient space (see 10.193)

we call  $v_1 \otimes \dots \otimes v_n$  the tensor product of the vectors  $v_1, \dots, v_n$

We prove now that the tensor product exist.

**Theorem 10.384.** *If  $\{V_i\}_{i \in \{1, \dots, n\}}$  is a non empty finite family of vector spaces over a field  $F$  then  $\otimes: \prod_{i \in \{1, \dots, n\}} V_i \rightarrow V_1 \otimes \dots \otimes V_n$  satisfies the axioms of a tensor product*

1.  $\otimes$  is multilinear or in other words  $\otimes \in \text{Hom}(V_1, \dots, V_n; V_1 \otimes \dots \otimes V_n)$  (in a simple notation this means  $v_1 \otimes \dots \otimes v_{i-1} \otimes \alpha \cdot x + \beta \cdot y \otimes v_{i+1} \otimes \dots \otimes v_n = \alpha \cdot v_1 \otimes \dots \otimes v_{i-1} \otimes x \otimes v_{i+1} \otimes \dots \otimes v_n + \beta \cdot v_1 \otimes \dots \otimes v_{i-1} \otimes y \otimes v_{i+1} \otimes \dots \otimes v_n$ )
2.  $\mathcal{S}(\otimes(\prod_{i \in \{1, \dots, n\}} V_i)) = V_1 \otimes \dots \otimes V_n$
3. **(Universal factorization property)** If  $U$  is any vector space over  $F$  and  $\varphi \in \text{Hom}(V_1, \dots, V_n; U)$  is arbitrary then there exist a  $h \in \text{Hom}(V_1 \otimes \dots \otimes V_n, U)$  (a linear mapping) such that  $\varphi = h \circ \otimes$

Or in other words  $\langle V_1 \otimes \dots \otimes V_n, \otimes \rangle$  is a tensor product of  $\{V_i\}_{i \in \{1, \dots, n\}}$ .

### Proof.

1. If  $(v_1, \dots, v_{i-1}, \alpha \cdot x + \beta \cdot y, v_{i+1}, \dots, v_n) \in \prod_{i \in \{1, \dots, n\}} V_i$  then  $\delta(v_1, \dots, v_{i-1}, \alpha \cdot x + \beta \cdot y, v_{i+1}, \dots, v_n) - \alpha \cdot \delta(v_1, \dots, v_{i-1}, x, v_i, \dots, v_n) - \beta \cdot \delta(v_1, \dots, v_{i-1}, y, v_{i+1}, \dots, v_n) \in \mathcal{N}(\prod_{i \in \{1, \dots, n\}} V_i, F) \subseteq \mathcal{S}(\mathcal{N}(\prod_{i \in \{1, \dots, n\}} V_i, F))$  so that  $0 = \pi(\delta(v_1, \dots, v_{i-1}, \alpha \cdot x + \beta \cdot y, v_{i+1}, \dots, v_n) - \alpha \cdot \delta(v_1, \dots, v_{i-1}, x, v_i, \dots, v_n) - \beta \cdot \delta(v_1, \dots, v_{i-1}, y, v_{i+1}, \dots, v_n)) = \pi(\delta(v_1, \dots, v_{i-1}, \alpha \cdot x + \beta \cdot y, v_{i+1}, \dots, v_n)) - \alpha \cdot \pi(\delta(v_1, \dots, v_{i-1}, x, v_i, \dots, v_n)) - \beta \cdot \pi(\delta(v_1, \dots, v_{i-1}, y, v_{i+1}, \dots, v_n)) \Rightarrow 0 = \otimes(v_1, \dots, v_{i-1}, \alpha \cdot x + \beta \cdot y, v_{i+1}, \dots, v_n) - \alpha \cdot \otimes(v_1, \dots, v_{i-1}, x, v_{i+1}, \dots, v_n) - \beta \cdot \otimes(v_1, \dots, v_{i-1}, y, v_{i+1}, \dots, v_n)$  giving that  $\otimes(v_1, \dots, v_{i-1}, \alpha \cdot x + \beta \cdot y, v_{i+1}, \dots, v_n) = \alpha \cdot \otimes(v_1, \dots, v_{i-1}, x, v_{i+1}, \dots, v_n) + \beta \cdot \otimes(v_1, \dots, v_{i-1}, y, v_{i+1}, \dots, v_n)$  so that  $\otimes \in \text{Hom}(V_1, \dots, V_n, V_1 \otimes \dots \otimes V_n)$
2. As we already have that  $\mathcal{S}(\otimes(\prod_{i \in \{1, \dots, n\}} V_i)) \subseteq V_1 \otimes \dots \otimes V_n$  we must only prove that  $V_1 \otimes \dots \otimes V_n \subseteq \mathcal{S}(\otimes(\prod_{i \in \{1, \dots, n\}} V_i))$ . So let  $x \in V_1 \otimes \dots \otimes V_n$  then as  $\pi$  is a surjection there exists a  $y \in \mathcal{F}(\prod_{i \in \{1, \dots, n\}} V_i, F)$  such that  $\pi(y) = x$ . Now as  $\delta(\prod_{i \in \{1, \dots, n\}} V_i)$  forms a basis of  $\mathcal{F}(\prod_{i \in \{1, \dots, n\}} V_i, F)$  (see 10.171) there exists a  $\{\alpha_i\}_{i \in \{1, \dots, n\}}$  and a  $\{\delta(v_1^{(i)}, \dots, v_n^{(i)})\}_{i \in \{1, \dots, n\}}$  (see 10.134) such that  $y = \sum_{i \in \{1, \dots, n\}} \alpha_i \cdot \delta(v_1^{(i)}, \dots, v_n^{(i)})$  and thus  $x = \pi(\sum_{i \in \{1, \dots, n\}} \alpha_i \cdot \delta(v_1^{(i)}, \dots, v_n^{(i)})) = \sum_{i \in \{1, \dots, n\}} \alpha_i \cdot \pi(\delta(v_1^{(i)}, \dots, v_n^{(i)})) = \sum_{i \in \{1, \dots, n\}} \alpha_i \cdot v_1^{(i)} \otimes \dots \otimes v_n^{(i)} \in \mathcal{S}(\otimes(\prod_{i \in \{1, \dots, n\}} V_i))$
3. As  $\{\delta_x | z \in \prod_{i \in \{1, \dots, n\}} V_i\}$  is by 10.171 a basis for  $\mathcal{F}(\prod_{i \in \{1, \dots, n\}} V_i, F)$  we have if  $x \in \mathcal{F}(\prod_{i \in \{1, \dots, n\}} V_i, F)$  that there exists a unique  $\{\alpha_z^{(x)}\}_{z \in \prod_{i \in \{1, \dots, n\}} V_i}$  with finite support such that  $x = \sum_{z \in \prod_{i \in \{1, \dots, n\}} V_i} \alpha_z^{(x)} \cdot \delta_z$ . Define now  $\gamma: \mathcal{F}(\prod_{i \in \{1, \dots, n\}} V_i, F) \rightarrow U$  by  $x \rightarrow \sum_{z \in \prod_{i \in \{1, \dots, n\}} V_i} \alpha_z^{(x)} \cdot \varphi(z)$  then we have that
  - a. **( $\gamma \circ \delta = \varphi$ )** If  $x \in \prod_{i \in \{1, \dots, n\}} V_i$  then  $\delta(x) = \delta_x$  so that  $(\gamma \circ \delta)(x) = \gamma(\delta(x)) = 1 \cdot f(x) = \varphi(x)$
  - b. **( $\gamma$  is linear)** as  $\alpha \cdot x + \beta \cdot y = \alpha \cdot \sum_{z \in \prod_{i \in \{1, \dots, n\}} V_i} \alpha_z^{(x)} + \beta \cdot \sum_{z \in \prod_{i \in \{1, \dots, n\}} V_i} \alpha_z^{(y)} = \sum_{z \in \prod_{i \in \{1, \dots, n\}} V_i} (\alpha \cdot \alpha_z^{(x)} + \beta \cdot \alpha_z^{(y)})$  so that  $\gamma(\alpha \cdot x + \beta \cdot y) = \sum_{z \in \prod_{i \in \{1, \dots, n\}} V_i} (\alpha \cdot \alpha_z^{(x)} + \beta \cdot \alpha_z^{(y)}) \cdot \varphi(z) = \alpha \cdot \sum_{z \in \prod_{i \in \{1, \dots, n\}} V_i} \alpha_z^{(x)} \cdot \varphi(z) + \beta \cdot \sum_{z \in \prod_{i \in \{1, \dots, n\}} V_i} \alpha_z^{(y)} \cdot \varphi(z) = \alpha \cdot \gamma(x) + \beta \cdot \gamma(y)$

- c.  $(\mathcal{S}(\mathcal{N}(\prod_{i \in \{1, \dots, n\}} V_i, F)) \subseteq \ker(\gamma))$  If  $z \in \mathcal{N}(\prod_{i \in \{1, \dots, n\}} V_i, F)$  then there  $\exists i \in \{1, \dots, n\}, \exists (x_1, \dots, x_{i-1}, \alpha \cdot x + \beta \cdot y, x_{i+1}, \dots, x_n) \in \prod_{i \in \{1, \dots, n\}} V_i$  and a  $\alpha, \beta \in F$  such that  $z = \delta_{(x_1, \dots, x_{i-1}, \alpha \cdot x + \beta \cdot y, x_{i+1}, \dots, x_n)} - \alpha \cdot \delta_{(x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_n)} - \beta \cdot \delta_{(x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_n)}$  so that  $\gamma(z) = 1 \cdot \varphi(x_1, \dots, x_{i-1}, \alpha \cdot x + \beta \cdot y, x_{i+1}, \dots, x_n) - \alpha \cdot \varphi(x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_n) - \beta \cdot \varphi(x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_n) \underset{\varphi \text{ is multilinear}}{\equiv} 0$ . If now  $z \in \mathcal{S}(\mathcal{N}(\prod_{i \in \{1, \dots, n\}} V_i, F))$  then there exists a  $\{\alpha_x\}_{x \in \mathcal{N}(\prod_{i \in \{1, \dots, n\}} V_i, F)}$  with finite support such that  $z = \sum_{x \in \mathcal{N}(\prod_{i \in \{1, \dots, n\}} V_i, F)} \alpha_x \cdot x$  so that we have  $\gamma(x) = \gamma\left(\sum_{x \in \mathcal{N}(\prod_{i \in \{1, \dots, n\}} V_i, F)} \alpha_x \cdot x\right) \underset{\gamma \text{ is linear}}{\equiv} \sum_{x \in \mathcal{N}(\prod_{i \in \{1, \dots, n\}} V_i, F)} \alpha_x \cdot \underset{x \in \mathcal{N}(\prod_{i \in \{1, \dots, n\}} V_i, F)}{\underset{=}{\gamma(x)}} = 0$  so we have  $\mathcal{S}(\mathcal{N}(\prod_{i \in \{1, \dots, n\}} V_i, F)) \subseteq \ker(\gamma)$ .

Using 10.225 and (b,c) we have then the existence of a linear function  $h: \mathcal{F}(\prod_{i \in \{1, \dots, n\}} V_i, F) / \mathcal{S}(\mathcal{N}(\prod_{i \in \{1, \dots, n\}} V_i, F)) = V_1 \otimes \dots \otimes V_n \rightarrow U$  such that  $\gamma = h \circ \pi$ . Using (a) we have then  $\varphi \underset{(a)}{=} \gamma \circ \delta = (h \circ \pi) \circ \delta = h \circ (\pi \circ \delta) \underset{\otimes = \pi \circ \delta}{=} h \circ \otimes$

□

We prove now some useful theorems about tensor products.

**Theorem 10.385.** If  $\{V_i\}_{i \in \{1, \dots, n\}}$  is a finite non empty family of vector spaces over a field  $F$  with characterization 0,  $\langle V_1 \otimes \dots \otimes V_n, \otimes \rangle$  is a tensor product and  $(x_1, \dots, x_n) \in \prod_{i \in \{1, \dots, n\}} V_i$  then we have  $x_1 \otimes \dots \otimes x_n = 0$  if and only if  $\exists i \in \{1, \dots, n\}$  such that  $x_i = 0$

**Proof.**

1.  $(\Rightarrow)$  If  $x_1 \otimes \dots \otimes x_n = 0$  and assume that  $\forall i \in \{1, \dots, n\}$  we have  $x_i = 0$  then by 10.178 there exists  $\forall i \in \{1, \dots, n\}$  a  $L_i \in \text{Hom}(V_i, F)$  such that  $L_i(x_i) = 1$  then we have by 10.233 that  $\prod_{i \in \{1, \dots, n\}} L_i \in \text{Hom}(V_1, \dots, V_n, F)$  then there exists a linear mapping  $h: V_1 \otimes \dots \otimes V_n \rightarrow F$  such that  $(\prod_{i \in \{1, \dots, n\}} L_i) = h \circ \otimes$  and thus  $0 = h(0) = h(x_1 \otimes \dots \otimes x_n) = (h \circ \otimes)(x_1, \dots, x_n) = (\prod_{i \in \{1, \dots, n\}} L_i)(x_1, \dots, x_n) = \prod_{i \in \{1, \dots, n\}} L_i(x_i) = \prod_{i \in \{1, \dots, n\}} 1 = 1$  giving that  $1 = 0$  which is a contradiction as  $F$  is of characterization 0. So  $\exists i \in \{1, \dots, n\}$  such that  $x_i = 0$
2.  $(\Leftarrow)$  If  $i \in \{1, \dots, n\}$  is such that  $x_i = 0$  then  $x_i = 0 \cdot x_i$  so that we have  $(x_1 \otimes \dots, \otimes x_n) = \otimes(x_1, \dots, x_n) = \otimes(x_1, \dots, x_{n-1}, 0 \cdot x_i, x_i, \dots, x_n) = 0 \cdot (x_1, \dots, x_n) = 0$

Next we consider the problem of defining a tensor product of sub spaces  $W_i \subseteq V_i$  in such a way that  $W_1 \otimes \dots \otimes W_n \subseteq V_1 \otimes \dots \otimes V_n$ , this is solved in the next theorem

**Theorem 10.386.** Let  $\{V_i\}_{i \in \{1, \dots, n\}}$  a finite non empty family of vector spaces over a field  $F$  and let  $\{W_i\}_{i \in \{1, \dots, n\}}$  a family of sub spaces /  $\forall i \in \{1, \dots, n\}$  we have  $W_i$  is a sub space of  $V_i$ . If  $\langle P, \nu \rangle$  is a tensor product of  $\{V_i\}_{i \in \{1, \dots, n\}}$  then if we define  $\langle Q, \mu \rangle$  by (note that  $\prod_{i \in \{1, \dots, n\}} W_i \subseteq \prod_{i \in \{1, \dots, n\}} V_i$ )

1.  $Q = \mathcal{S}(\mu(\prod_{i \in \{1, \dots, n\}} W_i))$
2.  $\mu = \nu|_{\prod_{i \in \{1, \dots, n\}} W_i}: \prod_{i \in \{1, \dots, n\}} W_i \rightarrow Q$

then we have that  $\langle Q, \mu \rangle$  is a tensor product such that  $Q \subseteq P$  and trivially by the definition of  $\mu$  that  $\forall (x_1, \dots, x_n) \in \prod_{i \in \{1, \dots, n\}} W_i$  we have  $\mu(x_1, \dots, x_n) = \nu(x_1, \dots, x_n)$ .

**Proof.** If  $i \in \{1, \dots, n\}$  then if  $(x_1, \dots, x_{i-1}, \alpha \cdot x + \beta \cdot y, x_{i+1}, \dots, x_n) \in \prod_{i \in \{1, \dots, n\}} W_i$  then  $\mu(x_1, \dots, x_{i-1}, \alpha \cdot x + \beta \cdot y, x_{i+1}, \dots, x_n) = \nu(x_1, \dots, x_{i-1}, \alpha \cdot x + \beta \cdot y, x_{i+1}, \dots, x_n) \underset{\nu \text{ is multilinear}}{\equiv} \alpha \cdot \nu(x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_n) + \beta \cdot \nu(x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_n) = \alpha \cdot \mu(x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_n) + \beta \cdot \mu(x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_n)$ , so we have

$$\mu \in \text{Hom}(W_1, \dots, W_n, Q). \quad (10.151)$$

Further we have by definition that

$$\mathcal{S}\left(\mu\left(\prod_{i \in \{1, \dots, n\}} W_i\right)\right) = Q. \quad (10.152)$$

Now if  $U$  is a vector space over the field  $F$  and  $\psi \in \text{Hom}(W_1, \dots, W_n; U)$  then using 10.246 there exists a  $\varphi \in \text{Hom}(V_1, \dots, V_n; U)$  such that  $\varphi|_{\prod_{i \in \{1, \dots, n\}} W_i} = \psi$ . As  $\langle P, \nu \rangle$  is a tensor product there exists a  $h \in \text{Hom}(P, U)$  such that  $h \circ \nu = \varphi$ , define now  $k = h|_Q: Q \subseteq P \rightarrow U$  then we have  $\forall (w_1, \dots, w_n) \in \prod_{i \in \{1, \dots, n\}} W_i$  that

$$\begin{aligned}
 (k \circ \mu)(w_1, \dots, w_n) &= k(\mu(w_1, \dots, w_n)) \\
 &= k(\nu(w_1, \dots, w_n)) \\
 &\stackrel{\mu = \nu|_{\prod_{i \in \{1, \dots, n\}} W_i}}{=} h(\nu(w_1, \dots, w_n)) \\
 &\stackrel{\nu(\prod_{i \in \{1, \dots, n\}} W_i) = Q \wedge k = h|_Q}{=} h(\nu(w_1, \dots, w_n)) \\
 &\stackrel{h \circ \nu = \varphi}{=} \varphi(w_1, \dots, w_n) \\
 &\stackrel{\varphi|_{\prod_{i \in \{1, \dots, n\}} W_i} = \psi}{=} \psi(w_1, \dots, w_n)
 \end{aligned}$$

proving that

$$k \circ \mu = \psi \quad (10.153)$$

Using 10.151, 10.152 and 10.153 we have that  $\langle Q, \mu \rangle$  is indeed a tensor product.  $\square$

In other words the above theorem means using the  $\otimes$  notation that if  $\{V_i\}_{i \in \{1, \dots, n\}}$  is a family of vector spaces with  $\{W_i\}_{i \in \{1, \dots, n\}}$  a family of sub spaces then if  $\langle V_1 \otimes \dots \otimes V_n, \otimes \rangle$  is a tensor product of  $\{V_i\}_{i \in \{1, \dots, n\}}$  then the restriction of  $\otimes$  defined by  $\otimes|_{\prod_{i \in \{1, \dots, n\}} W_i}: \prod_{i \in \{1, \dots, n\}} W_i \rightarrow \otimes(\prod_{i \in \{1, \dots, n\}} W_i)$   $\stackrel{\text{definition}}{=} W_1 \otimes \dots \otimes W_n$  forms a tensor product noted by  $\langle W_1 \otimes \dots \otimes W_n, \otimes \rangle$  such that  $W_1 \otimes \dots \otimes W_n \subseteq V_1 \otimes \dots \otimes V_n$  and  $\forall (x_1, \dots, x_n) \in \prod_{i \in \{1, \dots, n\}} W_i$  we have  $x_1 \otimes \dots \otimes x_n \in W_1 \otimes \dots \otimes W_n \subseteq V_1 \otimes \dots \otimes V_n$

**Theorem 10.387.** *Let  $\{V_i\}_{i \in \{1, \dots, 2\}}$  be vector spaces over a field  $F$  and let  $z \in V_1 \otimes V_2$ . Then  $\{k \in \mathbb{N}_0 \mid \exists \{u_i\}_{i \in \{1, \dots, k\}} \text{ a family in } V_1 \text{ and } \exists \{v_i\}_{i \in \{1, \dots, k\}} \text{ such that } z = \sum_{i \in \{1, \dots, k\}} u_i \otimes v_i\} \neq \emptyset$  and thus its minimum exist (see 4.52). If  $k_z = \min(\{k \in \mathbb{N}_0 \mid \exists \{u_i\}_{i \in \{1, \dots, k\}} \text{ a family in } V_1 \text{ and } \exists \{v_i\}_{i \in \{1, \dots, k\}} \text{ such that } z = \sum_{i \in \{1, \dots, k\}} u_i \otimes v_i\})$  then  $\{u_i\}_{i \in \{1, \dots, k\}}, \{v_i\}_{i \in \{1, \dots, k\}}$  are linear independent sets.*

**Proof.** As by definition  $\mathcal{S}(\otimes(V_1 \times V_2)) = V_1 \otimes V_2$  we have that  $\forall v \in V_1 \otimes V_2$  there exists by 10.134 finite families  $\{v'_i \otimes u_i\}_{i \in \{1, \dots, n\}}$  in  $\otimes(V_1 \times V_2)$  and  $\{\alpha_i\}_{i \in \{1, \dots, n\}}$  in  $F$  such that  $v = \sum_{i \in \{1, \dots, n\}} \alpha_i \cdot (v'_i \otimes u_i) = \sum_{i \in \{1, \dots, n\}} v_i \otimes u_i$  where  $v_i = \alpha_i \cdot v'_i \in V_1$  so that  $\{k \in \mathbb{N}_0 \mid \exists \{u_i\}_{i \in \{1, \dots, k\}} \text{ a family in } V_1 \text{ and } \exists \{v_i\}_{i \in \{1, \dots, k\}} \text{ such that } z = \sum_{i \in \{1, \dots, k\}} u_i \otimes v_i\} \neq \emptyset$ . Let now  $k = \min(\{k \in \mathbb{N}_0 \mid \exists \{u_i\}_{i \in \{1, \dots, k\}} \text{ a family in } V_1 \text{ and } \exists \{v_i\}_{i \in \{1, \dots, k\}} \text{ such that } z = \sum_{i \in \{1, \dots, k\}} u_i \otimes v_i\})$ . If now  $z = \sum_{i \in \{1, \dots, k\}} u_i^{(1)} \otimes u_i^{(2)}$  then we have that:

1. If  $\{u_i\}_{i \in \{1, \dots, k\}}$  is dependent then there exists a  $i_0 \in \{1, \dots, k\}$  such that  $u_{i_0} = \sum_{i \in \{1, \dots, k\} \setminus \{i_0\}} c_i \cdot u_i$  then

$$\begin{aligned}
 z &= \sum_{i \in \{1, \dots, k\}} u_i \otimes v_i \\
 &= \sum_{i \in \{1, \dots, k\} \setminus \{i_0\}} u_i \otimes v_i + u_{i_0} \otimes v_{i_0} \\
 &= \sum_{i \in \{1, \dots, k\} \setminus \{i_0\}} u_i \otimes v_i + \left( \sum_{i \in \{1, \dots, k\} \setminus \{i_0\}} c_i \cdot u_i \right) \times v_{i_0} \\
 &= \sum_{i \in \{1, \dots, k\} \setminus \{i_0\}} u_i \otimes v_i + \sum_{i \in \{1, \dots, k\} \setminus \{i_0\}} (c_i \cdot (u_i \otimes v_{i_0})) \\
 &= \sum_{i \in \{1, \dots, k\} \setminus \{i_0\}} (u_i \otimes v_i + c_i \cdot (u_i \otimes v_{i_0})) \\
 &= \sum_{i \in \{1, \dots, k\} \setminus \{i_0\}} (u_i \otimes (v_i + c_i \cdot v_{i_0}))
 \end{aligned}$$

If we define now the bijection  $b: \{1, \dots, k-1\} \rightarrow \{1, \dots, k\} \setminus \{i_0\}$  by  $b(i) = \begin{cases} i & \text{if } i < i_0 \\ i+1 & \text{if } i_0 \leq i \leq k-1 \end{cases}$  and  $\{u'_i \otimes v'_i\}_{i \in \{1, \dots, k-1\}}$  by  $u'_i \otimes v'_i = u_{b_i} \otimes (v_{b_i} + c_{b_i} \cdot v_{i_0})$  then we have  $v = \sum_{i \in \{1, \dots, k-1\}} u'_i \otimes v'_i$  contradicting the minimality of  $k$ .

2. If  $\{v_{i_i}\}_{i \in \{1, \dots, k\}}$  is dependent then there exists a  $i_0 \in \{1, \dots, k\}$  such that  $v_{i_0} = \sum_{i \in \{1, \dots, k\} \setminus \{i_0\}} c_i \cdot v_i$  then

$$\begin{aligned} z &= \sum_{i \in \{1, \dots, k\}} u_i \otimes v_i \\ &= \sum_{i \in \{1, \dots, k\} \setminus \{i_0\}} u_i \otimes v_i + u_{i_0} \otimes v_{i_0} \\ &= \sum_{i \in \{1, \dots, k\} \setminus \{i_0\}} u_i \otimes v_i + u_{i_0} \otimes \left( \sum_{i \in \{1, \dots, k\} \setminus \{i_0\}} c_i \cdot v_i \right) \\ &= \sum_{i \in \{1, \dots, k\} \setminus \{i_0\}} u_i \otimes v_i + \sum_{i \in \{1, \dots, k\} \setminus \{i_0\}} (c_i \cdot (u_{i_0} \otimes v_i)) \\ &= \sum_{i \in \{1, \dots, k\} \setminus \{i_0\}} (u_i \otimes v_i + c_i \cdot (u_{i_0} \otimes v_i)) \\ &= \sum_{i \in \{1, \dots, k\} \setminus \{i_0\}} ((u_i + c_i \cdot u_{i_0}) \otimes v_i) \end{aligned}$$

If we define now the bijection  $b: \{1, \dots, k-1\} \rightarrow \{1, \dots, k\} \setminus \{i_0\}$  by  $b(i) = \begin{cases} i & \text{if } i < i_0 \\ i+1 & \text{if } i_0 \leq i \leq k-1 \end{cases}$  and  $\{u'_i \otimes v'_i\}_{i \in \{1, \dots, k-1\}}$  by  $u'_i \otimes v'_i = (u_{b_i} + c_{b_i} \cdot u_{i_0}) \otimes v_{b_i}$  then we have  $v = \sum_{i \in \{1, \dots, k-1\}} u'_i \otimes v'_i$  contradicting the minimality of  $k$ .

So we have proved that  $\{u_i\}_{i \in \{1, \dots, k\}}$  and  $\{v_i\}_{i \in \{1, \dots, k\}}$  are linear dependent.  $\square$

**Theorem 10.388.** Let  $\{V_i\}_{i \in \{1, \dots, m+1\}}$  be vector spaces over a field  $F$  and  $V_1 \otimes \dots \otimes V_{m+1}$  is a tensor product then for the tensor product  $V_1 \otimes \dots \otimes V_m$  there exists a tensor product of  $V_1 \otimes \dots \otimes V_m$  and  $V_{m+1}$  such that  $V_1 \otimes \dots \otimes V_{m+1} = (V_1 \otimes \dots \otimes V_m) \otimes V_{m+1}$

**Proof.** To avoid confusion in notation we write  $\langle V_1 \otimes \dots \otimes V_{m+1}, \nu_{m+1} \rangle$  for the tensor product of  $\{V_i\}_{i \in \{1, \dots, m+1\}}$   $\langle V_1 \otimes \dots \otimes V_m, \nu_m \rangle$  for the tensor product of  $\{V_i\}_{i \in \{1, \dots, m\}}$  Using 10.384 we have also a tensor product of  $V_1 \otimes \dots \otimes V_m$  and  $V_{m+1}$  that we call  $\langle (V_1 \otimes \dots \otimes V_m) \otimes V_{m+1}, \nu \rangle$ . To prove the theorem we must construct a new tensor product  $\langle V_1 \otimes \dots \otimes V_{m+1}, \tau \rangle$  of  $V_1 \otimes \dots \otimes V_m$  and  $V_{m+1}$ . Given  $v_{m+1} \in V_{m+1}$  define

$$\beta_{v_m} \in \text{Hom}(V_1, \dots, V_m, V_1 \otimes \dots \otimes V_{m+1}) \text{ by } \beta_{v_m}(v_1, \dots, v_m) = \nu_{m+1}(v_1, \dots, v_m, v_{m+1}) \quad (10.154)$$

which is indeed multilinear because of the multilinearity of the tensor product  $\nu_{m+1}$ . Also because of the multilinearity of  $\nu_{m+1}$  we have for  $\alpha, \beta \in F$  and  $\nu_{m+1}, \nu'_{m+1} \in V_{m+1}$  that

$$\beta_{\alpha \cdot \nu_{m+1} + \beta \cdot \nu'_{m+1}} = \alpha \cdot \beta_{\nu_{m+1}} + \beta \cdot \beta_{\nu'_{m+1}} \quad (10.155)$$

By the definition of the tensor product  $\langle V_1 \otimes \dots \otimes V_m, \nu_m \rangle$  we have the existence of a linear mapping

$$g_{v_m} \in \text{Hom}(V_1 \otimes \dots \otimes V_m, V_1 \otimes \dots \otimes V_{m+1})$$

$$g_{v_{m+1}} \circ \nu_m = \beta_{\nu_m} \Rightarrow g_{v_{m+1}}(\nu_m(v_1, \dots, v_m)) = \nu_{m+1}(v_1, \dots, v_m, v_{m+1}) \quad (10.156)$$

Take now the multilinear mapping  $\varphi$

$$\varphi \in \text{Hom}(V_1, \dots, V_{m+1}, (V_1 \otimes \dots \otimes V_m) \otimes V_{m+1}) \text{ defined by}$$

$$\varphi(v_1, \dots, v_{m+1}) = \nu(\nu_m(v_1, \dots, v_m), v_{m+1}) \quad (10.157)$$

Then by the definition of the tensor product  $\langle V_1 \otimes \dots \otimes V_{m+1}, \nu_{m+1} \rangle$  there exists a linear mapping

$$h \in \text{Hom}(V_1 \otimes \dots \otimes V_{m+1}, (V_1 \otimes \dots \otimes V_m) \otimes V_{m+1})$$

$$h \circ \nu_{m+1} = \varphi \Rightarrow h(\nu_{m+1}(v_1, \dots, v_{m+1})) = \nu(\nu_m(v_1, \dots, v_m), v_{m+1}) \quad (10.158)$$

Define now the multilinear mapping  $\psi$

$$\psi: (V_1 \otimes \dots \otimes V_m) \times V_{m+1} \rightarrow V_1 \otimes \dots \otimes V_{m+1} \text{ by } \psi(z, v_{m+1}) = g_{v_{m+1}}(z) \quad (10.159)$$

Then using 10.156 we have that  $\psi$  is multilinear so

$$\psi \in \text{Hom}((V_1 \otimes \dots \otimes V_m), V_{m+1}, V_1 \otimes \dots \otimes V_{m+1}) \quad (10.160)$$

Using the definition of the tensor product  $\langle (V_1 \otimes \dots \otimes V_m) \otimes V_{m+1}, \nu \rangle$  there exist then a linear mapping

$$k \in \text{Hom}((V_1 \otimes \dots \otimes V_m) \otimes V_{m+1}, V_1 \otimes \dots \otimes V_{m+1}) \text{ with } k \circ \nu = \psi$$

So we have if  $z = \nu_m(v_1, \dots, v_m)$  that

$$\begin{aligned} k(\nu(\nu_m(v_1, \dots, v_m), v_{m+1})) &= k(\nu(z, v_{m+1})) \\ &= \psi(z, v_{m+1}) \\ &= g_{v_{m+1}}(z) \\ &= g_{v_{m+1}}(\nu_m(v_1, \dots, v_m)) \\ &\stackrel{10.156}{=} \nu_{m+1}(v_1, \dots, v_m, v_{m+1}) \end{aligned} \quad (10.161)$$

Consider then the linear mapping  $k \circ h: V_1 \otimes \dots \otimes V_{m+1} \rightarrow V_1 \otimes \dots \otimes V_{m+1}$  then we have if  $z \in V_1 \otimes \dots \otimes V_{m+1}$  that  $z = \sum_{i \in \{1, \dots, r\}} \nu_{m+1}(v_{i,1}, \dots, v_{i,m+1})$  and

$$\begin{aligned} (k \circ h)(z) &= k\left(h\left(\sum_{i \in \{1, \dots, r\}} \nu_{m+1}(v_{i,1}, \dots, v_{i,m+1})\right)\right) \\ &= \sum_{i \in \{1, \dots, r\}} k(h(\nu_{m+1}(v_{i,1}, \dots, v_{i,m+1}))) \\ &\stackrel{10.158}{=} \sum_{i \in \{1, \dots, r\}} k(\nu(\nu_m(v_{i,1}, \dots, v_{i,m}), v_{i,m+1})) \\ &\stackrel{10.161}{=} \sum_{i \in \{1, \dots, r\}} \nu_{m+1}(v_{i,1}, \dots, v_{i,m+1}) \\ &= z \end{aligned}$$

Proving that

$$k \circ h = 1_{V_1 \otimes \dots \otimes V_{m+1}} \quad (10.162)$$

Likewise we have for the linear mapping  $h \circ k: (V_1 \otimes \dots \otimes V_m) \otimes V_{m+1} \rightarrow (V_1 \otimes \dots \otimes V_m) \otimes V_{m+1}$  that if  $z \in (V_1 \otimes \dots \otimes V_m) \otimes V_{m+1}$  then  $z = \sum_{i \in \{1, \dots, r\}} \nu\left(\sum_{j \in \{1, \dots, l\}} \nu_m(v_{i,j,1}, \dots, v_{i,j,m}), v_{i,m+1}\right)$

$$\begin{aligned} (h \circ k)(z) &= h\left(k\left(\sum_{i \in \{1, \dots, r\}} \nu\left(\sum_{j \in \{1, \dots, l\}} \nu_m(v_{i,j,1}, \dots, v_{i,j,m}), v_{i,m+1}\right)\right)\right) \\ &= \sum_{i \in \{1, \dots, r\}} h\left(k\left(\nu\left(\sum_{j \in \{1, \dots, l\}} \nu_m(v_{i,j,1}, \dots, v_{i,j,m}), v_{i,m+1}\right)\right)\right) \\ &\stackrel{\nu \text{ is multilinear}}{=} \sum_{i \in \{1, \dots, r\}} h\left(k\left(\sum_{j \in \{1, \dots, l\}} \nu(\nu_m(v_{i,j,1}, \dots, v_{i,j,m}), v_{i,m+1})\right)\right) \\ &= \sum_{i \in \{1, \dots, r\}} \left( \sum_{j \in \{1, \dots, l\}} h(k(\nu(\nu_m(v_{i,j,1}, \dots, v_{i,j,m}), v_{i,m+1}))) \right) \\ &\stackrel{10.161}{=} \sum_{i \in \{1, \dots, r\}} \left( \sum_{j \in \{1, \dots, l\}} h(\nu_{m+1}(\nu_{m+1}(v_{i,j,1}, \dots, v_{i,j,m}, v_{i,m+1}))) \right) \\ &\stackrel{10.158}{=} \sum_{i \in \{1, \dots, r\}} \nu\left(\sum_{j \in \{1, \dots, l\}} \nu(\nu_m(v_{i,j,1}, \dots, v_{i,j,m}), v_{i,m+1})\right) \\ &= z \end{aligned}$$

Proving that

$$h \circ k = 1_{(V_1 \otimes \dots \otimes V_m) \otimes V_{m+1}} \quad (10.163)$$

Define now

$$\tau: (V_1 \otimes \dots \otimes V_m) \times V_{m+1} \rightarrow V_1 \otimes \dots \otimes V_{m+1} \text{ by } \tau(z, v_{m+1}) = k(\nu(z, v_{m+1})) \quad (10.164)$$

Then by the linearity of  $k$  and the multilinearity of  $\nu$  we have that  $\tau$  is multilinear or

$$\tau \in \text{Hom}((V_1 \otimes \dots \otimes V_m), V_{m+1}, V_1 \otimes \dots \otimes V_{m+1}) \quad (10.165)$$

We finish by proving that  $\langle V_1 \otimes \dots \otimes V_{m+1}, \tau \rangle$  is a tensor product of  $V_1 \otimes \dots \otimes V_m$  and  $V_{m+1}$ .

1.  $V_1 \otimes \dots \otimes V_{m+1}$  is a vector space space over the field  $F$  (by the definition of  $\langle V_1 \otimes \dots \otimes V_{m+1}, \nu_{m+1} \rangle$ )
2. By 10.177 we have  $\tau \in \text{Hom}((V_1 \otimes \dots \otimes V_m), V_{m+1}, V_1 \otimes \dots \otimes V_{m+1})$
3. We have as  $\tau(V_1 \otimes \dots \otimes V_m \times V_{m+1}) \subseteq V_1 \otimes \dots \otimes V_{m+1}$  that  $\mathcal{S}(\tau(V_1 \otimes \dots \otimes V_m \times V_m)) \subseteq \mathcal{S}(V_1 \otimes \dots \otimes V_{m+1}) = V_1 \otimes \dots \otimes V_{m+1} \Rightarrow \tau(V_1 \otimes \dots \otimes V_m \times V_{m+1}) \subseteq V_1 \otimes \dots \otimes V_{m+1}$ . If  $v \in V_1 \otimes \dots \otimes V_{m+1}$  then  $v = \sum_{i \in \{1, \dots, l\}} \nu_{m+1}(v_1, \dots, v_{m+1}) \stackrel{10.162}{=} \sum_{i \in \{1, \dots, l\}} k(h(\nu_{m+1}(v_1, \dots, v_{m+1}))) \stackrel{10.158}{=} \sum_{i \in \{1, \dots, l\}} k(\nu(\nu_m(v_1, \dots, v_m), v_{m+1})) \stackrel{10.176}{=} \sum_{i \in \{1, \dots, l\}} \tau(\nu_m(v_1, \dots, v_m), v_{m+1}) \in \mathcal{S}(\tau(V_1 \otimes \dots \otimes V_m \times V_{m+1}))$  proving that  $V_1 \otimes \dots \otimes V_{m+1} \subseteq \mathcal{S}(\tau(V_1 \otimes \dots \otimes V_m \times V_{m+1}))$ . This proves that  $\mathcal{S}(\tau(V_1 \otimes \dots \otimes V_m \times V_{m+1})) = V_1 \otimes \dots \otimes V_{m+1}$
4. If  $\theta \in \text{Hom}(V_1 \otimes \dots \otimes V_m, V_{m+1}, U)$  then by the definition of  $\langle (V_1 \otimes \dots \otimes V_m) \otimes V_{m+1}, \nu \rangle$  there exists a  $\xi \in \text{Hom}((V_1 \otimes \dots \otimes V_m) \otimes V_{m+1}, U)$  such that  $\xi \circ \nu = \theta$ . Define now  $\lambda = \xi \circ h: V_1 \otimes \dots \otimes V_{m+1} \rightarrow U$  which is linear because  $\xi$  and  $h$  are linear, so that  $\lambda \in \text{Hom}(V_1 \otimes \dots \otimes V_{m+1}, U)$ . Further we have if  $(z, v_{m+1}) \in V_1 \otimes \dots \otimes V_m \times V_{m+1}$

$$\begin{aligned} (\lambda \circ \tau)(z, v_{m+1}) &= \lambda(\tau(z, v_{m+1})) \\ &= \lambda(k(\nu(z, v_{m+1}))) \\ &\stackrel{10.163}{=} \xi(h(k(\nu(z, v_{m+1})))) \\ &= \xi(\nu(z, v_{m+1})) \\ &= \theta(z, v_{m+1}) \end{aligned}$$

proving that  $\lambda \circ \tau = \theta$  □

**Lemma 10.389.** Let  $\{X_i\}_{i \in \{1, \dots, n\}}$  be a finite family of finite sets then  $\prod_{i \in \{1, \dots, n\}} X_i$  is a finite set and  $\#(\prod_{i \in \{1, \dots, n\}} X_i) = \prod_{i \in \{1, \dots, n\}} \#(X_i)$

**Proof.** First note that if  $m > 2$  then we have that  $\prod_{i \in \{1, \dots, m\}} X_i$  is bijective with  $(\prod_{i \in \{1, \dots, m-1\}} X_i) \times X_m$  with  $b: \prod_{i \in \{1, \dots, m\}} X_i \rightarrow (\prod_{i \in \{1, \dots, m-1\}} X_i) \times X_m$  where  $b(x_1, \dots, x_m) = ((x_1, \dots, x_{m-1}), x_m)$

1. **(injectivity)** If  $b(x_1, \dots, x_m) = b(y_1, \dots, y_m) \Rightarrow ((x_1, \dots, x_{m-1}), x_m) = ((y_1, \dots, y_{m-1}), y_m) \Rightarrow x_m = y_m \wedge (x_1, \dots, x_{m-1}) = (y_1, \dots, y_{m-1}) \Rightarrow (x_1, \dots, x_m) = (y_1, \dots, y_m)$
2. **(surjectivity)** If  $((x_1, \dots, x_{m-1}), y) \in (\prod_{i \in \{1, \dots, m-1\}} X_i) \times X_m$  then  $b(x_1, \dots, x_{m-1}, y) = ((x_1, \dots, x_{m-1}), y)$

Next we proceed by induction so let  $S = \{n \in \mathbb{N}_0 \mid \text{if } \{X_i\}_{i \in \{1, \dots, n\}} \text{ be a finite family of finite sets then } \prod_{i \in \{1, \dots, n\}} \text{ is finite with } \#(\prod_{i \in \{1, \dots, n\}} X_i) = \prod_{i \in \{1, \dots, n\}} \#(X_i)\}$  then we have

1. **(n=1)** then  $\prod_{i \in \{1, \dots, 1\}} X_i$  is bijective with  $X_1$  by using the bijection  $b: \prod_{i \in \{1, \dots, 1\}} X_i \rightarrow X_1$  where  $b(x) = x(1)$ 
  - a. **(injectivity)** If  $b(x) = b(y) = x(1) = y(1) \Rightarrow \forall i \in \{1, \dots, 1\}$  we have  $x(i) = y(i) \Rightarrow x = y$
  - b. **(surjectivity)** If  $x \in X_1$  take then  $x' \in \prod_{i \in \{1, \dots, 1\}} X_i$  by  $\forall i \in \{1, \dots, 1\}$  we have  $x'(i) = x'(1) = x$  so that  $b(x') = x'(1) = x$

So we have proved that  $\#(\prod_{i \in \{1, \dots, 1\}} X_i) = \#(X_1) = \prod_{i \in \{1, \dots, 1\}} \#(X_i)$  proving that  $1 \in S$

2. ( $n \in S$ ) then for  $n+1$  we have that  $\prod_{i \in \{1, \dots, n+1\}} X_i$  is bijective with  $(\prod_{i \in \{1, \dots, n\}} X_i) \times X_{n+1}$  which by 5.44 is finite and for which we have  $\#((\prod_{i \in \{1, \dots, n\}} X_i) \times X_{n+1}) = \#(\prod_{i \in \{1, \dots, n\}} X_i) \cdot \#(X_{n+1})$ . So we have  $\#(\prod_{i \in \{1, \dots, n+1\}} X_i) = \#(\prod_{i \in \{1, \dots, n\}} X_i) \cdot \#(X_{n+1}) \prod_{i \in \{1, \dots, n\}} \#(X_i) \cdot \#(X_{n+1}) = \prod_{i \in \{1, \dots, n+1\}} \#(X_i)$  and thus  $n+1 \in S$

Using mathematical induction we have then proved our theorem.  $\square$

**Definition 10.390.** Let  $\{n_i\}_{i \in \{1, \dots, k\}}$  be a family in  $\mathbb{N}_0$  then  $\Gamma(n_1, \dots, n_k) = \prod_{i \in \{1, \dots, k\}} \{1, \dots, n_i\}$ . Using the above lemma we have then that  $\#(\Gamma(k_1, \dots, k_m)) = \prod_{i \in \{1, \dots, m\}} k_i$

**Example 10.391.**  $\Gamma(2, 3) = \{(1, 1), (1, 2), (1, 3), (2, 1), (2, 2), (2, 3)\}$

**Theorem 10.392.** Let  $\{V_i\}_{i \in \{1, \dots, m\}}$  be vector spaces over a field  $F$ ,  $V$  a vector space over the field  $F$  and  $\varphi: \prod_{i \in \{1, \dots, m\}} V_i \rightarrow V$  a multilinear mapping,  $\{k_i\}_{i \in \{1, \dots, m\}}$  a family in  $\mathbb{N}_0$  and  $\forall i \in \{1, \dots, m\}$  a family  $\{w_{i,j}\}_{j \in \{1, \dots, k_i\}}$  in  $V_i$  then  $\varphi(\sum_{j \in \{1, \dots, k_1\}} w_{1,j}, \dots, \sum_{j \in \{1, \dots, k_m\}} w_{m,j}) = \sum_{\gamma \in \Gamma} \varphi(w_{1,\gamma(1)}, \dots, w_{m,\gamma(m)})$

**Proof.** We prove this by induction on let  $S = \{m \in \mathbb{N}_0 \mid \text{the assertion of the theorem is true}\}$  then we have

1. ( $m = 1$ ) In this case we have that  $\Gamma(k_1) = \{1, \dots, k_1\}^{\{1\}} = \{f: \{1\} \rightarrow \{1, \dots, k_1\}\}$  if  $\gamma \in \Gamma(k_1)$  and we can trivially make a bijection from  $\tau: \{1, \dots, k_1\} \rightarrow \Gamma(k_1)$  by  $i \mapsto \tau(i)$  where  $(\tau(i))(1) = i$
- (injectivity)** If  $\tau(i) = \tau(j) \Rightarrow (\tau(i))(1) = (\tau(j))(1) \Rightarrow i = j$
  - (surjectivity)** If  $\gamma \in \Gamma(k_1)$  take then  $i = \gamma(1) \in \{1, \dots, k_1\}$  then we have that  $\forall j \in \{1\}$  that  $(\tau(i))(j) = (\tau(i))(1) = i = \gamma(1) = \gamma(j) \Rightarrow \gamma = \tau(i)$

As  $\varphi(\sum_{j \in \{1, \dots, k_1\}} w_{1,j}) \underset{\varphi \text{ is multilinear}}{=} \sum_{j \in \{1, \dots, k_1\}} \varphi(w_{1,j})$   $\sum_{j \in \{1, \dots, k_1\}} \varphi(w_{1,\tau(j)(1)}) \underset{10.43}{=} \sum_{\gamma \in \Gamma(k_1)} \varphi(w_{1,\gamma(1)})$  proving that  $1 \in S$

2. ( $m \in S$ ) then if we take  $m+1$  and  $\varphi: \prod_{i \in \{1, \dots, m+1\}} V_i \rightarrow V$  then if we define (see 10.240)  $\varphi_{\dots \sum_{j \in \{1, \dots, k_{m+1}\}} w_{m+1,j}}: \prod_{i \in \{1, \dots, m\}} V_i \rightarrow V$  by  $\varphi_{\dots \sum_{j \in \{1, \dots, k_{m+1}\}} w_{m+1,j}}(x_1, \dots, x_m) = \varphi(x_1, \dots, x_m, \sum_{j \in \{1, \dots, k_{m+1}\}} w_{m+1,j})$  then we have that by 10.240 that  $\varphi_{\dots \sum_{j \in \{1, \dots, k_{m+1}\}} w_{m+1,j}}$  is multilinear and as  $m \in S$  we have then that  $\varphi(\sum_{j \in \{1, \dots, k_1\}} w_{1,j}, \dots, \sum_{j \in \{1, \dots, k_{m+1}\}} w_{m+1,j}) = \varphi_{\dots \sum_{j \in \{1, \dots, k_{m+1}\}} w_{m+1,j}}(\sum_{j \in \{1, \dots, k_1\}} w_{1,j}, \dots, \sum_{j \in \{1, \dots, k_m\}} w_{m,j}) = \sum_{\gamma \in \Gamma(k_1, \dots, k_m)} \varphi_{\dots \sum_{j \in \{1, \dots, k_{m+1}\}} w_{m+1,j}}(w_{1,\gamma(1)}, \dots, w_{m,\gamma(m)}) = \sum_{\gamma \in \Gamma(k_1, \dots, k_m)} \varphi(w_{1,\gamma(1)}, \dots, w_{m,\gamma(m)}, \sum_{j \in \{1, \dots, k_{m+1}\}} w_{m+1,j}) \underset{\varphi \text{ is multilinear}}{=} \sum_{\gamma \in \Gamma(k_1, \dots, k_m)} (\sum_{j \in \{1, \dots, k_{m+1}\}} \varphi(w_{1,\gamma(1)}, \dots, w_{m,\gamma(m)}, w_{m+1,j}))$  or

$$\varphi\left(\sum_{j \in \{1, \dots, k_1\}} w_{1,j}, \dots, \sum_{j \in \{1, \dots, k_{m+1}\}} w_{m+1,j}, w_{m,\gamma(m)}, w_{m+1,j}\right) = \sum_{\gamma \in \Gamma(k_1, \dots, k_m)} \left( \sum_{j \in \{1, \dots, k_{m+1}\}} \varphi(w_{1,\gamma(1)}, \dots, w_{m,\gamma(m)}, w_{m+1,j}) \right) \quad (10.166)$$

Define now  $\chi: \Gamma(k_1, \dots, k_m) \times \{1, \dots, k_{m+1}\} \rightarrow \Gamma(k_1, \dots, k_{m+1})$  by  $(\gamma, i) \mapsto \chi(\gamma, i)$  by  $\forall j \in \{1, \dots, m+1\}$  we have  $\chi(\gamma, i)(j) = \begin{cases} \gamma(j) & \text{if } j \in \{1, \dots, m\} \\ i & \text{if } j = m+1 \end{cases}$ . We prove now that  $\chi$  is a bijection

- (injectivity)** If  $\chi(\gamma, i) = \chi(\gamma', i')$  then  $\forall j \in \{1, \dots, m\}$  we have  $\chi(\gamma, i)(j) = \chi(\gamma', i')(j) \Rightarrow \gamma(j) = \gamma'(j)$  and thus we have  $\gamma = \gamma'$ , also we have  $\chi(\gamma, i)(m+1) = \chi(\gamma', i')(m+1) \Rightarrow i = i'$ . So finally we must have  $(\gamma, i) = (\gamma', i')$  proving injectivity.

2. **(surjectivity)** If  $\gamma \in \Gamma(k_1, \dots, k_{m+1})$  taken then  $\gamma|_{\{1, \dots, m\}} \in \Gamma(k_1, \dots, k_m)$  and  $\gamma(m+1) \in \{1, \dots, k_{m+1}\}$  then we have if  $j \in \{1, \dots, m\}$  that  $\chi(\gamma|_{\{1, \dots, m\}}, \gamma(m+1))(j) = \gamma|_{\{1, \dots, m\}}(j) = \gamma(j)$  and if  $j = m+1$  then we have  $\chi(\gamma|_{\{1, \dots, m\}}, \gamma(m+1)) = \gamma(m+1)$ . This proves that  $\chi(\gamma|_{\{1, \dots, m\}}, \gamma(m+1)) = \gamma$  and thus surjectivity.  $\square$

Using  $\chi$  we have then that  $\sum_{(\gamma, i) \in \Gamma(k_1, \dots, k_m) \times \{1, \dots, k_{m+1}\}} \varphi(w_{1, \gamma(1)}, \dots, w_{m, \gamma(m)}, w_{m+1, i}) = \sum_{(\gamma, i) \in \Gamma(k_1, \dots, k_m) \times \{1, \dots, k_{m+1}\}} \varphi(w_{1, \chi(\gamma, i)(1)}, \dots, w_{m, \chi(\gamma, i)(m)}, w_{m+1, \chi(\gamma, i)(m+1)})$  or

$$\sum_{(\gamma, i) \in \Gamma(k_1, \dots, k_m) \times \{1, \dots, k_{m+1}\}} \varphi(w_{1, \gamma(1)}, \dots, w_{m, \gamma(m)}, w_{m+1, i}) = \sum_{\gamma \in \Gamma(k_1, \dots, k_{m+1})} \varphi(w_{1, \gamma(1)}, \dots, w_{m, \gamma(m)}, w_{m+1, \gamma(m+1)}) \quad (10.167)$$

Given now a  $\gamma \in \Gamma(k_1, \dots, k_m)$  construct then the trivially bijective function  $\rho_\gamma: \{\gamma\} \times \{1, \dots, k_{m+1}\} \rightarrow \{1, \dots, k_{m+1}\}$  by  $(\gamma, i) \rightarrow \rho_\gamma(\gamma, i) = i$

Now as  $\Gamma(k_1, \dots, k_m) \times \{1, \dots, k_{m+1}\} = \bigcup_{\gamma \in \Gamma(k_1, \dots, k_m)} \{\gamma\} \times \{1, \dots, k_{m+1}\}$  and also that trivially the union is pairwise disjoint then we have by 10.44 and that  $\sum_{(\gamma, i) \in \Gamma(k_1, \dots, k_m) \times \{1, \dots, k_{m+1}\}} \varphi(w_{1, \gamma(1)}, \dots, w_{m, \gamma(m)}, w_{m+1, i}) = \sum_{\gamma \in \Gamma(k_1, \dots, k_m)} \left( \sum_{(\kappa, i) \in \{\gamma\} \times \{1, \dots, k_{m+1}\}} \varphi(w_{1, \kappa(1)}, \dots, w_{m, \kappa(m)}, w_{m+1, i}) \right) = \sum_{\gamma \in \Gamma(k_1, \dots, k_m)} \left( \sum_{(\kappa, i) \in \{\gamma\} \times \{1, \dots, k_{m+1}\}} \varphi(w_{1, \gamma(1)}, \dots, w_{m, \gamma(m)}, w_{m+1, \rho_\gamma(\kappa, i)}) \right)$  or

$$\sum_{(\gamma, i) \in \Gamma(k_1, \dots, k_m) \times \{1, \dots, k_{m+1}\}} \varphi(w_{1, \gamma(1)}, \dots, w_{m, \gamma(m)}, w_{m+1, i}) = \sum_{\gamma \in \Gamma(k_1, \dots, k_m)} \left( \sum_{j \in \{1, \dots, k_{m+1}\}} \varphi(w_{1, \gamma(1)}, \dots, w_{m, \gamma(m)}, w_{m+1, j}) \right) \quad (10.168)$$

Finally we have by 10.166, 10.167 and the above 10.168 that  $\varphi(\sum_{j \in \{1, \dots, k_1\}} w_{1, j}, \dots, \sum_{j \in \{1, \dots, k_{m+1}\}} w_{m+1, j}) = \sum_{\gamma \in \Gamma(k_1, \dots, k_{m+1})} \varphi(w_{1, \gamma(1)}, \dots, w_{m+1, \gamma(m+1)})$  proving that  $m+1 \in S$ .

By mathematical induction we have then proved the theorem.

**Lemma 10.393.** Let  $\{V_i\}_{i \in \{1, \dots, n\}}$  be vector spaces over a field  $F$  such that  $\forall i \in \{1, \dots, n\}$  there exists a family  $\{W_{i,j}\}_{j \in \{1, \dots, k_i\}}$  of sub spaces of  $V_i$  such that  $V_i = \sum_{j \in \{1, \dots, k_i\}} W_{i,j}$  (see 10.359) then we have that  $V_1 \otimes \dots \otimes V_n = \sum_{\gamma \in \Gamma(k_1, \dots, k_n)} (W_{1, \gamma(1)} \otimes \dots \otimes W_{n, \gamma(n)})$  (where the  $\otimes$  on the  $W_{i,j}$  is defined by 10.386 based on  $\otimes$  of the  $V_i$ 's)

**Proof.** First as  $W_{1, \gamma(1)} \otimes \dots \otimes W_{n, \gamma(n)} \subseteq V_1 \otimes \dots \otimes V_n$  (see 10.386)  $\forall \gamma \in \Gamma(k_1, \dots, k_n)$  we have if  $v \in \sum_{\gamma \in \Gamma(k_1, \dots, k_n)} (W_{1, \gamma(1)} \otimes \dots \otimes W_{n, \gamma(n)})$  then  $v = \sum_{\gamma \in \Gamma(k_1, \dots, k_n)} v_\gamma$  where  $\forall \gamma \in \Gamma(k_1, \dots, k_n)$  we have  $v_\gamma \in W_{1, \gamma(1)} \otimes \dots \otimes W_{n, \gamma(n)} \subseteq V_1 \otimes \dots \otimes V_n$  so that  $v = \sum_{\gamma \in \Gamma(k_1, \dots, k_n)} v_\gamma \in V_1 \otimes \dots \otimes V_n$  proving that

$$\sum_{\gamma \in \Gamma(k_1, \dots, k_n)} (W_{1, \gamma(1)} \otimes \dots \otimes W_{n, \gamma(n)}) \subseteq V_1 \otimes \dots \otimes V_n. \quad (10.169)$$

Now if  $(v_1, \dots, v_n) \in \prod_{i \in \{1, \dots, n\}} V_i$  then  $\forall i \in \{1, \dots, n\}$  there exists a  $\{w_{i,j}\}_{j \in \{1, \dots, k_i\}}$  such that  $v_i = \sum_{j \in \{1, \dots, k_i\}} w_{i,j}$ . As  $\otimes$  is multilinear we have by 10.392 that  $v_1 \otimes \dots \otimes v_n = \otimes(v_1, \dots, v_n) = \sum_{\gamma \in \Gamma(k_1, \dots, k_n)} \otimes(v_{1, \gamma(1)}, \dots, v_{n, \gamma(n)}) = \sum_{\gamma \in \Gamma(k_1, \dots, k_n)} w_{1, \gamma(1)} \otimes \dots \otimes w_{n, \gamma(n)} \in \sum_{\gamma \in \Gamma(k_1, \dots, k_n)} (W_{1, \gamma(1)} \otimes \dots \otimes W_{n, \gamma(n)})$  proving that  $\otimes(\prod_{i \in \{1, \dots, n\}} V_i) \subseteq \sum_{\gamma \in \Gamma(k_1, \dots, k_n)} (W_{1, \gamma(1)} \otimes \dots \otimes W_{n, \gamma(n)})$ . From  $V_1 \otimes \dots \otimes V_n = \mathcal{S}(\otimes(\prod_{i \in \{1, \dots, n\}} V_i)) \subseteq_{10.138} \mathcal{S}(\sum_{\gamma \in \Gamma(k_1, \dots, k_n)} (W_{1, \gamma(1)} \otimes \dots \otimes W_{n, \gamma(n)}))$   $\stackrel{10.365 \text{ and } 10.136}{=} \sum_{\gamma \in \Gamma} (W_{1, \gamma(1)} \otimes \dots \otimes W_{n, \gamma(n)})$  which together with 10.169 gives that

$$V_1 \otimes \dots \otimes V_n = \sum_{\gamma \in \Gamma(k_1, \dots, k_n)} (W_{1, \gamma(1)} \otimes \dots \otimes W_{n, \gamma(n)}) \quad \square$$

**Lemma 10.394.** Let  $\{V_i\}_{i \in \{1,2\}}$  be vector spaces over a field  $F$  (with characterization 0) and  $V_1 = W_1 \oplus W_2$  (or  $V_2 = W_1 \oplus W_2$ ) then  $V_1 \otimes V_2 = (W_1 \otimes V_2) \oplus (W_2 \otimes V_2)$  (or  $V_1 \otimes V_2 = (V_1 \otimes W_1) \oplus (V_2 \otimes W_2)$ )

**Proof.**

1. ( $V_1 = W_1 \oplus W_2$ ) By 10.393 we have already proved that  $V_1 \otimes V_2 = (W_1 \otimes V_2) + (W_2 \otimes V_2)$ , so the only thing left to prove is that  $(W_1 \otimes V_2) \cap (W_2 \otimes V_2) = \{0\}$ . We do this by contradiction, so assume that  $x \in (W_1 \otimes V_2) \cap (W_2 \otimes V_2)$  with  $x \neq 0$  then using 10.380 there exists families  $\{u_i\}_{i \in \{1, \dots, k\}}$  of elements of  $W_1$ ,  $\{v_i\}_{i \in \{1, \dots, r\}}$  of elements of  $W_2$  and  $\{x_i\}_{i \in \{1, \dots, k\}}$ ,  $\{y_i\}_{i \in \{1, \dots, r\}}$  of elements of  $V_2$  such that

$$\begin{aligned} x &= \sum_{i \in \{1, \dots, k\}} u_i \otimes x_i \in W_1 \otimes V_2 \\ &= \sum_{i \in \{1, \dots, r\}} v_i \otimes y_i \in W_2 \otimes V_2 \end{aligned} \tag{10.170}$$

in which  $r, k$  are minimal. Using 10.387 we have then that  $\{u\}_{i \in \{1, \dots, k\}}$ ,  $\{v_i\}_{i \in \{1, \dots, r\}}$ ,  $\{x_i\}_{i \in \{1, \dots, k\}}$ ,  $\{y_i\}_{i \in \{1, \dots, r\}}$  are linear independent sets. As  $\{x_i\}_{i \in \{1, \dots, k\}}$  is linear independent we have using 10.180 the existence of a linear mapping  $h: V_2 \rightarrow F$  such that  $h(x_1) = 1$  and  $\forall i \in \{2, \dots, k\}$  we have  $h(x_i) = 0$ . Choose a arbitrary linear mapping  $f: V_1 \rightarrow F$  then we have that  $f \cdot h: V_1 \times V_2 \rightarrow F$  is a multilinear mapping and thus by the definition of the tensor product there exists a unique linear  $g: V_1 \otimes V_2 \rightarrow F$  such that

$$g \circ \otimes = f \cdot h \tag{10.171}$$

Using 10.170 we have then

$$\begin{aligned} g(x) &= g\left(\sum_{i \in \{1, \dots, k\}} u_i \otimes x_i\right) \\ &= \sum_{i \in \{1, \dots, k\}} g(u_i \otimes x_i) \\ &= \sum_{i \in \{1, \dots, k\}} (g \circ \otimes)(u_i, x_i) \\ &= \sum_{i \in \{1, \dots, k\}} f(u_i) \cdot h(x_i) \\ &= f(u_1) \end{aligned} \tag{10.172}$$

But we have also

$$\begin{aligned} g(x) &= g\left(\sum_{i \in \{1, \dots, r\}} v_i \otimes y_i\right) \\ &= \sum_{i \in \{1, \dots, r\}} g(v_i \otimes y_i) \\ &= \sum_{i \in \{1, \dots, r\}} f(v_i) \cdot h(y_i) \\ &\stackrel{f \text{ is linear}}{=} f\left(\sum_{i \in \{1, \dots, r\}} h(y_i) \cdot v_i\right) \end{aligned}$$

so that we have for a arbitrary linear function  $f: V_1 \rightarrow F$  that  $f(u_1) = f\left(\sum_{i \in \{1, \dots, r\}} h(y_i) \cdot v_i\right)$  which gives by 10.179 that  $W_1 \ni u_1 = \sum_{i \in \{1, \dots, r\}} h(y_i) \cdot v_i \in W_2$ . This gives that  $u_1 \in W_1 \cap W_2 = \{0\}$  or  $u_1 = 0$  contradicting the linear Independence of  $\{u_i\}_{i \in \{1, \dots, k\}}$ . So we conclude that our assumption of  $x \neq 0$  is wrong and thus we must have that  $(W_1 \otimes V_2) \cap (W_2 \otimes V_2) = \{0\}$ .

2. ( $\mathbf{V}_2 = \mathbf{W}_1 \oplus \mathbf{W}_2$ ) This is proved in a similar way as (1). By 10.393 we have already proved that  $V_1 \otimes V_2 = (V_1 \otimes W_2) + (V_1 \otimes W_2)$ , so the only thing left to prove is that  $(V_1 \otimes W_1) \cap (V_1 \otimes W_2) = \{0\}$ . We do this by contradiction, so assume that  $x \in (V_1 \otimes W_1) \cap (V_1 \otimes W_2)$  with  $x \neq 0$  then using 10.380 there exists families  $\{u_i\}_{i \in \{1, \dots, k\}}$ ,  $\{v_i\}_{i \in \{1, \dots, r\}}$  of elements of  $V_1$ ,  $\{x_i\}_{i \in \{1, \dots, k\}}$  a family in  $W_1$  and  $\{y_i\}_{i \in \{1, \dots, r\}}$  a family in  $W_2$  such that

$$\begin{aligned} x &= \sum_{i \in \{1, \dots, k\}} u_i \otimes x_i \in V_1 \otimes W_1 \\ &= \sum_{i \in \{1, \dots, r\}} v_i \otimes y_i \in V_1 \otimes W_2 \end{aligned} \tag{10.173}$$

in which  $r, k$  are minimal. Using 10.387 we have then that  $\{u\}_{i \in \{1, \dots, k\}}$ ,  $\{v\}_{i \in \{1, \dots, r\}}$ ,  $\{x\}_{i \in \{1, \dots, k\}}$ ,  $\{y\}_{i \in \{1, \dots, r\}}$  are linear independent sets. As  $\{u_i\}_{i \in \{1, \dots, k\}}$  is linear independent we have using 10.180 the existence of a linear mapping  $h: V_1 \rightarrow F$  such that  $h(u_1) = 1$  and  $\forall i \in \{2, \dots, k\}$  we have  $h(u_i) = 0$ . Choose a arbitrary linear mapping  $f: V_2 \rightarrow F$  then we have that  $h \cdot f: V_1 \times V_2 \rightarrow F$  is a multilinear mapping and thus by the definition of the tensor product there exists a unique linear  $g: V_1 \otimes V_2 \rightarrow F$  such that

$$g \circ \otimes = h \cdot f \tag{10.174}$$

Using 10.173 we have then

$$\begin{aligned} g(x) &= g\left(\sum_{i \in \{1, \dots, k\}} u_i \otimes x_i\right) \\ &= \sum_{i \in \{1, \dots, k\}} g(u_i \otimes x_i) \\ &= \sum_{i \in \{1, \dots, k\}} (g \circ \otimes)(u_i, x_i) \\ &= \sum_{i \in \{1, \dots, k\}} h(u_i) \cdot f(x_i) \\ &= f(x_1) \end{aligned} \tag{10.175}$$

But we have also

$$\begin{aligned} g(x) &= g\left(\sum_{i \in \{1, \dots, r\}} v_i \otimes y_i\right) \\ &= \sum_{i \in \{1, \dots, r\}} g(v_i \otimes y_i) \\ &= \sum_{i \in \{1, \dots, r\}} h(v_i) \cdot f(y_i) \\ &\stackrel{f \text{ is linear}}{=} f\left(\sum_{i \in \{1, \dots, r\}} h(v_i) \cdot y_i\right) \end{aligned}$$

so that we have for a arbitrary linear function  $f: V_1 \rightarrow F$  that  $f(x_1) = f\left(\sum_{i \in \{1, \dots, r\}} h(v_i) \cdot y_i\right)$  which gives by 10.179 that  $W_1 \ni x_1 = \sum_{i \in \{1, \dots, r\}} h(v_i) \cdot y_i \in W_2$ . This gives that  $x_1 \in W_1 \cap W_2 = \{0\}$  or  $x_1 = 0$  contradicting the linear Independence of  $\{u_i\}_{i \in \{1, \dots, k\}}$ . So we conclude that our assumption of  $x \neq 0$  is wrong and thus we must have that  $(V_1 \otimes W_1) \cap (V_1 \otimes W_2) = \{0\}$ .  $\square$

**Lemma 10.395.** *Let  $V_1, V_2$  be vector spaces over a field  $F$  (with characterization zero) such that  $V_1 = \sum_{i \in I}^\oplus W_i$  (or  $V_2 = \sum_{i \in I}^\oplus W_i$ ) then we have that  $V_1 \otimes V_2 = \sum_{i \in \{1, \dots, m\}}^\oplus (W_i \otimes V_2)$  (or  $V_1 \otimes V_2 = \sum_{i \in \{1, \dots, m\}}^\oplus (V_1 \otimes W_i)$ )*

**Proof.** Consider the two cases of the theorem

1.  $(V_1 = \sum_{i \in I}^{\oplus} W_i)$  We prove this by induction, so define  $S = \{m \in \mathbb{N}_0 \mid \text{if } W = \sum_{i \in \{1, \dots, m\}}^{\oplus} W_i, V \text{ are vector spaces over } F \text{ then } W \otimes V = \sum_{i \in \{1, \dots, m\}}^{\oplus} (W_i \otimes V)\}$  we have then that
  - a.  $(m = 1)$  Here we have  $W = \sum_{i \in \{1\}}^{\oplus} W_i = W_1$  (see 10.367) and thus  $W \otimes V = W_1 \otimes V \stackrel{10.367}{=} \sum_{i \in \{1\}} (W_i \otimes V)$  proving that  $1 \in S$
  - b.  $(m \in S)$  Consider now  $m+1$  then using 10.370 we have that  $W = \sum_{i \in \{1, \dots, m+1\}}^{\oplus} W_i = (\sum_{i \in \{1, \dots, m\}}^{\oplus} W_i) \oplus W_{m+1}$  and using the previous lemma we have then that  $W \otimes V = ((\sum_{i \in \{1, \dots, m\}}^{\oplus} W_i) \otimes V) \oplus (W_{m+1} \otimes V) \stackrel{m \in S}{=} (\sum_{i \in \{1, \dots, m\}}^{\oplus} (W_i \otimes V)) \oplus (W_{m+1} \otimes V) \stackrel{10.370}{=} \sum_{i \in \{1, \dots, m+1\}}^{\oplus} (W_i \otimes V)$  proving that  $m+1 \in S$

Using induction we have then that  $S = \mathbb{N}_0$  proving the first case (using  $V_1 = W, V_2 = V$ ).

2.  $(V_2 = \sum_{i \in I}^{\oplus} W_i)$  We prove this by induction, so define  $S = \{m \in \mathbb{N}_0 \mid \text{if } V = \sum_{i \in \{1, \dots, m\}}^{\oplus} W_i, W \text{ are vector spaces over } F \text{ then } W \otimes V = \sum_{i \in \{1, \dots, m\}}^{\oplus} (W \otimes W_i)\}$  we have then that
  - a.  $(m = 1)$  Here we have  $V = \sum_{i \in \{1\}}^{\oplus} W_i = W_1$  (see 10.367) and thus  $W \otimes V = W \otimes W_1 \stackrel{10.367}{=} \sum_{i \in \{1\}} (W \otimes W_i)$  proving that  $1 \in S$
  - b.  $(m \in S)$  Consider now  $m+1$  then using 10.370 we have that  $V = \sum_{i \in \{1, \dots, m+1\}}^{\oplus} W_i = (\sum_{i \in \{1, \dots, m\}}^{\oplus} W_i) \oplus W_{m+1}$  and using the previous lemma we have then that  $W \otimes V = (W \otimes (\sum_{i \in \{1, \dots, m\}}^{\oplus} W_i)) \oplus (W \otimes W_{m+1}) \stackrel{m \in S}{=} (\sum_{i \in \{1, \dots, m\}}^{\oplus} (W \otimes W_i)) \oplus (W \otimes W_{m+1}) \stackrel{10.370}{=} \sum_{i \in \{1, \dots, m+1\}}^{\oplus} (W \otimes W_i)$  proving that  $m+1 \in S$

Using induction we have then that  $S = \mathbb{N}_0$  proving the second case (if  $W = V_1, V = V_2$ )  $\square$

**Theorem 10.396.** Let  $V, W$  be vector spaces such that  $V = \sum_{i \in \{1, \dots, k\}}^{\oplus} W_i, W = \sum_{i \in \{1, \dots, l\}}^{\oplus} U_i$ . Then  $V \otimes W = \sum_{\gamma \in \Gamma(k, l)}^{\oplus} W_{\gamma(1)} \otimes U_{\gamma(2)}$

**Proof.** Again this is proved by induction so given  $k \in \mathbb{N}_0$  and let  $V = \sum_{i \in \{1, \dots, k\}}^{\oplus} W_i$  let  $S_k = \{m \in \mathbb{N}_0 \mid \text{if } W = \sum_{i \in \{1, \dots, m\}}^{\oplus} U_i \text{ then } V \otimes W = \sum_{\gamma \in \Gamma(k, m)}^{\oplus} W_{\gamma(1)} \otimes U_{\gamma(2)}\}$  then we have

1.  $(m = 1)$  Then  $W = \sum_{i \in \{1\}}^{\oplus} U_i = U_1$  and using the above lemma we have then that  $V \otimes W = \sum_{i \in \{1, \dots, k\}}^{\oplus} W_i \otimes U_1$ , if we then consider the trivial bijection  $\tau: \{1, \dots, k\} \rightarrow \Gamma(k, 1)$  by  $\tau(i)(1) = i$  and  $\tau(i)(2) = 1$  then  $V \otimes W = \sum_{i \in \{1, \dots, k\}}^{\oplus} W_{\tau(i)(1)} \otimes U_{\tau(i)(2)} \stackrel{10.369}{=} \sum_{\gamma \in \Gamma(k, 1)}^{\oplus} W_{\gamma(1)} \otimes U_{\gamma(2)}$  proving that  $1 \in S_k$
2.  $(m \in S)$  Then  $W = \sum_{i \in \{1, \dots, m+1\}}^{\oplus} U_i \stackrel{10.370}{=} (\sum_{i \in \{1, \dots, m\}}^{\oplus} U_i) \oplus U_{m+1}$  so that by the above lemma we have  $V \otimes W = (V \otimes (\sum_{i \in \{1, \dots, m\}}^{\oplus} U_i)) \oplus (V \otimes U_{m+1}) = ((\sum_{i \in \{1, \dots, k\}}^{\oplus} W_i) \oplus (\sum_{i \in \{1, \dots, m\}}^{\oplus} U_i)) \oplus ((\sum_{i \in \{1, \dots, k\}}^{\oplus} W_i) \otimes U_{m+1}) \stackrel{m \in S \text{ and previous lemma}}{=} (\sum_{\gamma \in \Gamma(k, m)}^{\oplus} W_{\gamma(1)} \otimes U_{\gamma(2)}) \oplus (\sum_{i \in \{1, \dots, k\}}^{\oplus} W_i \otimes U_{m+1})$  proving that

$$V \otimes W = \left( \sum_{\gamma \in \Gamma(k, m)}^{\oplus} W_{\gamma(1)} \otimes U_{\gamma(2)} \right) \oplus \left( \sum_{i \in \{1, \dots, k\}}^{\oplus} W_i \otimes U_{m+1} \right) \quad (10.176)$$

Define the trivial bijection  $\kappa: \{1, \dots, k\} \rightarrow \{1, \dots, k\} \times \{m+1\}$  by  $\kappa(i)(1) = i$  and  $\kappa(i)(2) = m+1$  then we have  $\sum_{i \in \{1, \dots, k\}}^{\oplus} W_i \otimes U_{m+1} = \sum_{i \in \{1, \dots, k\}}^{\oplus} W_{\kappa(i)(1)} \otimes U_{\kappa(i)(2)} \stackrel{10.369}{=} \sum_{\gamma \in \{1, \dots, k\} \times \{m+1\}}^{\oplus} W_{\gamma(1)} \otimes U_{\gamma(2)}$  giving

$$\sum_{i \in \{1, \dots, k\}}^{\oplus} W_i \otimes U_{m+1} = \sum_{\gamma \in \{1, \dots, k\} \times \{m+1\}}^{\oplus} W_{\gamma(1)} \otimes U_{\gamma(2)} \quad (10.177)$$

As we have trivially that  $\Gamma(k, m+1) = \{1, \dots, k\} \times \{1, \dots, m+1\} = \{1, \dots, k\} \times \{1, \dots, m\} \cup \{1, \dots, k\} \times \{m+1\} = \Gamma(k, m) \cup \{1, \dots, k\} \times \{m+1\}$  where the union is disjoint. Then using 10.176, 10.177 and 10.373 we have that  $V \otimes W = (\sum_{\gamma \in \Gamma(k, m)}^{\oplus} W_{\gamma(1)} \otimes U_{\gamma(2)}) \oplus \sum_{\gamma \in \{1, \dots, k\} \times \{m+1\}} W_{\gamma(1)} \otimes U_{\gamma(2)} = \sum_{\gamma \in \Gamma(k, m+1)} W_{\gamma(1)} \otimes U_{\gamma(2)}$ . This proves that  $m+1 \in S$ . Using induction we have then that  $S_k = \mathbb{N}_0$ . So as  $l \in \mathbb{N}_0 = S_k$  we have  $V \otimes W = \sum_{\gamma \in \Gamma(k, l)}^{\oplus} W_{\gamma(1)} \otimes U_{\gamma(2)}$ .  $\square$

**Theorem 10.397.** Let  $\{V_i\}_{i \in \{1, \dots, m\}}$  be vector spaces over a field  $F$  such that  $\forall i \in \{1, \dots, m\}$  there exists a family  $\{W_{i,j}\}_{j \in \{1, \dots, k_i\}}$  of sub spaces of  $V_i$  such that  $V_i = \sum_{j \in \{1, \dots, k_i\}}^{\oplus} W_{i,j}$  (see  $\oplus_{i \in \{1, \dots, n\}} V_i$ ) then we have that  $V_1 \otimes \dots \otimes V_n = \sum_{\gamma \in \Gamma(k_1, \dots, k_n)}^{\oplus} (W_{1,\gamma(1)} \otimes \dots \otimes W_{m,\gamma(m)})$  (where the  $\otimes$  on the  $W_{i,j}$  is defined by 10.386 based on  $\otimes$  of the  $V_i$ 's)

**Proof.** We prove this by induction so let  $S = \{m \in \mathbb{N}_0 \mid \text{If } \{V_i\}_{i \in \{1, \dots, m\}} \text{ be vector spaces over a field } F \text{ such that } \forall i \in \{1, \dots, m\} \text{ there exists a family } \{W_{i,j}\}_{j \in \{1, \dots, k_i\}} \text{ of subspaces of } V_i \text{ such that } V_i = \sum_{j \in \{1, \dots, k_i\}}^{\oplus} W_{i,j} \text{ then } V_1 \otimes \dots \otimes V_n = \sum_{\gamma \in \Gamma(k_1, \dots, k_n)}^{\oplus} (W_{1,\gamma(1)} \otimes \dots \otimes W_{m,\gamma(m)})\}$  then

1. ( $m = 1$ ) Then we have  $V_1 \otimes \dots \otimes V_1 = V_1 = \sum_{j \in \{1, \dots, k_1\}}^{\oplus} W_{1,j}$ , using the trivial bijection  $\kappa: \{1, \dots, k_1\} \rightarrow \Gamma(k_1)$  by  $\kappa(i)(1) = i$  then  $\sum_{j \in \{1, \dots, k_1\}}^{\oplus} W_{1,\kappa(j)(1)} = \sum_{\gamma \in \Gamma(k_1)} W_{1,\gamma(1)} = \sum_{\gamma \in \Gamma(k_1)} W_{1,\gamma(1)} \otimes \dots \otimes W_{1,\gamma(1)}$  proving that  $1 \in S$
2. ( $m \in S$ ) Then we have for  $m+1$  that  $V_1 \otimes \dots \otimes V_{m+1} = (V_1 \otimes \dots \otimes V_m) \otimes V_{m+1}$  (we can find a tensor product of  $V_1 \otimes \dots \otimes V_m$  and  $V_{m+1}$  satisfying the equation).  $V_1 \otimes \dots \otimes V_m = \sum_{i \in S}^{\oplus} \sum_{\gamma \in \Gamma(k_1, \dots, k_m)}^{\oplus} W_{1,\gamma(1)} \otimes \dots \otimes W_{m,\gamma(m)}$  and thus

$$\begin{aligned}
 (V_1 \otimes \dots \otimes V_m) \otimes V_{m+1} &= \left( \sum_{\gamma \in \Gamma(k_1, \dots, k_m)}^{\oplus} W_{1,\gamma(1)} \otimes \dots \otimes W_{m,\gamma(m)} \right) \otimes V_{m+1} \\
 &= \left( \sum_{\gamma \in \Gamma(k_1, \dots, k_m)}^{\oplus} ((W_{1,\gamma(1)} \otimes \dots \otimes W_{m,\gamma(m)})) \right) \otimes \\
 &\quad \sum_{i \in \{1, \dots, k_{m+1}\}}^{\oplus} W_{m+1,i} \\
 &\stackrel{10.395}{=} \sum_{i \in \{1, \dots, k_{m+1}\}}^{\oplus} \left( \left( \sum_{\gamma \in \Gamma(k_1, \dots, k_m)}^{\oplus} ((W_{1,\gamma(1)} \otimes \dots \otimes W_{m,\gamma(m)})) \right) \otimes \right. \\
 &\quad \left. W_{m+1,i} \right) \\
 &\stackrel{10.395}{=} \sum_{i \in \{1, \dots, k_{m+1}\}}^{\oplus} \left( \sum_{\gamma \in \Gamma(k_1, \dots, k_m)}^{\oplus} (W_{1,\gamma(1)} \otimes \dots \otimes W_{m,\gamma(m)}) \otimes \right. \\
 &\quad \left. W_{m+1,i} \right) \\
 &\stackrel{10.388}{=} \sum_{i \in \{1, \dots, k_{m+1}\}}^{\oplus} \left( \sum_{\gamma \in \Gamma(k_1, \dots, k_m)}^{\otimes} W_{1,\gamma(1)} \otimes \dots \otimes W_{m,\gamma(m)} \otimes W_{m+1,i} \right) \\
 \end{aligned} \tag{10.178}$$

If we define now  $\{1, \dots, k_{m+1}\} \times \Gamma(k_1, \dots, k_m) \xrightarrow{\text{disjoint union}} \bigcup_{i \in \{1, \dots, k_m\}} \{i\} \times \Gamma(k_1, \dots, k_m)$  then 10.178 becomes

$$(V_1 \otimes \dots \otimes V_m) \otimes V_{m+1} \underset{10.374}{=} \sum_{(i, \gamma) \in \{1, \dots, k_{m+1}\} \times \Gamma(k_1, \dots, k_m)} W_{1, \gamma(1)} \otimes \dots \otimes W_{m, \gamma(m)} \otimes W_{m+1, i} \quad (10.179)$$

Define the function  $\chi: \{1, \dots, k_{m+1}\} \times \Gamma(k_1, \dots, k_m) \rightarrow \Gamma(k_1, \dots, k_{m+1})$  where  $\chi(i, \gamma)(j) = \begin{cases} \gamma(j) & \text{if } j \leq m \\ i & \text{if } j = m+1 \end{cases}$  then we have

- a. **(injectivity)** If  $\chi(i, \gamma) = \chi(i', \gamma')$  then we have  $\chi(i, \gamma)(m+1) = \chi(i', \gamma')(m+1) \Rightarrow i = i'$ , also if  $j \in \{1, \dots, m\}$  then  $\chi(i, \gamma)(j) = \chi(i', \gamma')(j) \Rightarrow \gamma(j) = \gamma'(j) \Rightarrow \gamma = \gamma'$  so that  $(i, \gamma) = (i', \gamma')$
- b. **(surjectivity)** If  $\gamma \in \Gamma(k_1, \dots, k_{m+1})$  construct then  $(\gamma(m+1), \gamma|_{\{1, \dots, m\}})$  and then we have if  $j \in \{1, \dots, m+1\}$  that  $\chi(\gamma(m+1)) = \begin{cases} \gamma|_{\{1, \dots, m\}}(j) & \text{if } j \leq m \\ \gamma(m+1) & \text{if } j = m+1 \end{cases} = \begin{cases} \gamma(j) & \text{if } j \leq m \\ \gamma(m+1) & \text{if } j = m+1 \end{cases} = \gamma(j) \Rightarrow \chi(\gamma(m+1), \gamma|_{\{1, \dots, m\}}) = \gamma$

so  $\chi$  is a bijection and we can rewrite 10.179 as follows

$$\begin{aligned} V_1 \otimes \dots \otimes V_{m+1} &= (V_1 \otimes \dots \otimes V_m) \otimes V_{m+1} \\ &= \sum_{(i, \gamma) \in \{1, \dots, k_{m+1}\} \times \Gamma(k_1, \dots, k_m)} W_{1, \chi(1, \gamma)(1)} \otimes \dots \otimes W_{m, \chi(m, \gamma)(m)} \otimes \\ &\quad W_{m+1, \chi(m+1, \gamma)(m+1)} \\ &\underset{10.369}{=} \sum_{\gamma \in \Gamma(k_1, \dots, k_{m+1})} W_{1, \gamma} \otimes \dots \otimes W_{m+1, \gamma(m+1)} \end{aligned}$$

This proves that  $m+1 \in S$

By induction we have then that  $S = \mathbb{N}_0$  proving the theorem  $\square$

We use now the above theorem to construct the basis of a tensor product out of the basis of the factors.

First we prove the following lemma

**Lemma 10.398.** Let  $\{V_i\}_{i \in \{1, \dots, m\}}$  be a family of one dimensional vector spaces ( $\forall i \in \{1, \dots, m\}$  we have that  $\exists e_i \in V_i, e_i \neq 0$  such that  $V_i = \mathcal{S}(\{e_i\})$ ) then every tensor product  $V_1 \otimes \dots \otimes V_m = \mathcal{S}(\{e_1 \otimes \dots \otimes e_m\})$

**Proof.** First we prove that  $\langle P, \nu \rangle$  where  $P = \mathcal{S}(\{e_1 \otimes \dots \otimes e_m\})$  and  $\nu: \prod_{i \in \{1, \dots, m\}} \mathcal{S}(\{e_i\}) \rightarrow \mathcal{S}(\{e_1 \otimes \dots \otimes e_m\})$  is defined by  $\nu(\lambda_1 \cdot e_1, \dots, \lambda_m \cdot e_m) = (\prod_{i \in \{1, \dots, m\}} \lambda_i) \cdot (e_1 \otimes \dots \otimes e_m)$  is a tensor product of  $\{V_i\}_{i \in \{1, \dots, m\}}$

1.  $P$  is a vector space by 10.132
2. If  $\alpha, \beta \in F$  then  $\nu(\lambda_1 \cdot e_1, \dots, \lambda_{i-1} \cdot e_{i-1}, \alpha \cdot (\lambda'_i \cdot e_i) + \beta \cdot (\lambda''_i \cdot e_i), \lambda_{i+1} \cdot e_{i+1}, \dots, \lambda_m \cdot e_m) = \nu(\lambda_1 \cdot e_1, \dots, \lambda_{i-1} \cdot e_{i-1}, (\alpha \cdot \lambda'_i + \beta \cdot \lambda''_i) \cdot e_i, \lambda_{i+1} \cdot e_{i+1}, \dots, \lambda_m \cdot e_m) = (\prod_{i \in \{1, \dots, m\}} \lambda_i) \cdot (e_1 \otimes \dots \otimes e_m)$  where  $\{\lambda_i\}_{i \in \{1, \dots, m\}}$  is defined by  $\lambda_j = \begin{cases} \lambda_j & \text{if } j \neq i \\ \alpha \cdot \lambda'_i + \beta \cdot \lambda''_i & \text{if } j = i \end{cases}$  and then by 10.232 we have that  $(\prod_{i \in \{1, \dots, m\}} \lambda_i) = \alpha \cdot (\prod_{i \in \{1, \dots, m\}} (\lambda_1, \dots, \lambda_{i-1}, \lambda'_i, \lambda_{i+1}, \dots, \lambda_m)_i) + \beta \cdot (\prod_{i \in \{1, \dots, m\}} (\lambda_1, \dots, \lambda_{i-1}, \lambda''_i, \lambda_{i+1}, \dots, \lambda_m)_i) \Rightarrow \nu(\lambda_1 \cdot e_1, \dots, \lambda_{i-1} \cdot e_{i-1}, \alpha \cdot (\lambda'_i \cdot e_i) + \beta \cdot (\lambda''_i \cdot e_i), \lambda_{i+1} \cdot e_{i+1}, \dots, \lambda_m \cdot e_m) = \alpha \cdot (\prod_{i \in \{1, \dots, m\}} (\lambda_1, \dots, \lambda_{i-1}, \lambda'_i, \lambda_{i+1}, \dots, \lambda_m)_i) \cdot (e_1 \otimes \dots \otimes e_m) + \beta \cdot (\prod_{i \in \{1, \dots, m\}} (\lambda_1, \dots, \lambda_{i-1}, \lambda''_i, \lambda_{i+1}, \dots, \lambda_m)_i) \cdot (e_1 \otimes \dots \otimes e_m) = \alpha \cdot \nu(\lambda_1 \cdot e_1, \dots, \lambda_{i-1} \cdot e_{i-1}, \lambda'_i \cdot e_i, \lambda_{i+1} \cdot e_{i+1}, \dots, \lambda_m \cdot e_m) + \beta \cdot \nu(\lambda_1 \cdot e_1, \dots, \lambda_{i-1} \cdot e_{i-1}, \lambda''_i \cdot e_i, \lambda_{i+1} \cdot e_{i+1}, \dots, \lambda_m \cdot e_m)$  proving that  $\nu$  is multilinear  $\Rightarrow \nu \in \text{Hom}(\mathcal{S}(\{e_1\}), \dots, \mathcal{S}(\{e_m\}), \mathcal{S}(\{e_1 \otimes \dots \otimes e_m\}))$

3. As by definition of  $\nu$  we have  $\nu(\prod_{i \in \{1, \dots, m\}} \mathcal{S}(\{e_i\})) \subseteq \mathcal{S}(\{e_1 \otimes \dots \otimes e_m\})$  we have  $\mathcal{S}(\nu(\prod_{i \in \{1, \dots, m\}} \mathcal{S}(\{e_i\}))) \subseteq \mathcal{S}(\mathcal{S}(\{e_1 \otimes \dots \otimes e_m\})) = \mathcal{S}(\{e_1 \otimes \dots \otimes e_m\}) = P \Rightarrow \mathcal{S}(\nu(\prod_{i \in \{1, \dots, m\}} \mathcal{S}(\{e_i\}))) \subseteq P$ . If  $v \in P = \mathcal{S}(\{e_1 \otimes \dots \otimes e_m\}) \Rightarrow v = \lambda \cdot (e_1 \otimes \dots \otimes e_m)$  and  $\nu(\lambda \cdot e_1, 1 \cdot e_2, \dots, 1 \cdot e_m) = v \Rightarrow P \subseteq \nu(\prod_{i \in \{1, \dots, m\}} \mathcal{S}(\{e_i\})) \Rightarrow P \underset{P \text{ is a vectorspace}}{=} \mathcal{S}(P) \subseteq \mathcal{S}(\nu(\prod_{i \in \{1, \dots, m\}} \mathcal{S}(\{e_i\})))$  proving that  $P = \mathcal{S}(\nu(\prod_{i \in \{1, \dots, m\}} \mathcal{S}(\{e_i\})))$
4. If  $\varphi \in \text{Hom}(\mathcal{S}(\{e_1\}), \dots, \mathcal{S}(\{e_m\}), U)$  define then  $h: \mathcal{S}(\{e_1 \otimes \dots \otimes e_m\}) \rightarrow U$  by  $\lambda \cdot (e_1 \otimes \dots \otimes e_m) \rightarrow h(\lambda \cdot (e_1 \otimes \dots \otimes e_m)) = \lambda \cdot \varphi(e_1, \dots, e_m)$  which is trivially linear. Then we have  $\varphi(\lambda_1 \cdot e_1, \dots, \lambda_m \cdot e_m) \underset{10.241}{=} (\prod_{i \in \{1, \dots, m\}} \lambda_i) \cdot \varphi(e_1, \dots, e_m) = h((\prod_{i \in \{1, \dots, m\}} \lambda_i) \cdot (e_1 \otimes \dots \otimes e_m)) = h(\nu(\lambda_1 \cdot e_1, \dots, \lambda_m \cdot e_m))$  proving that  $h \circ \nu = \varphi$ .

So we have proved that  $\langle P, \nu \rangle$  is a tensor product of  $\{V_i\}_{i \in \{1, \dots, m\}}$  which is obviously one dimensional as by 10.379  $e_1 \otimes \dots \otimes e_m = \nu(e_1, \dots, e_m) \neq 0$ . If now  $\langle Q, \mu \rangle$  is another tensor product of  $\{V_i\}_{i \in \{1, \dots, m\}}$  then by 10.376 we have that  $Q$  is isomorphic with  $P$ , using 10.222 we have then that  $Q$  is also one dimensional. So there exist a  $v_0 \in Q$  such that  $Q = \mathcal{S}(\{v_0\})$ . Now as by 10.379  $e_1 \otimes \dots \otimes e_m = \mu(e_1, \dots, e_m) \neq 0$  we have

1.  $x \in \mathcal{S}(e_1 \otimes \dots \otimes e_m) \Rightarrow x = \lambda \cdot (e_1 \otimes \dots \otimes e_m) \Rightarrow x \in Q$  so  $\mathcal{S}(e_1 \otimes \dots \otimes e_m) \subseteq Q$
2. if  $x \in Q$  then  $x = \lambda \cdot v_0$ , now as  $0 \neq e_1 \otimes \dots \otimes e_m$  there exists a  $\tau \neq 0$  such that  $e_1 \otimes \dots \otimes e_m = \tau \cdot v_0 \Rightarrow v_0 = \frac{1}{\tau} \cdot e_1 \otimes \dots \otimes e_m$  so that  $x = (\lambda \cdot \frac{1}{\tau}) \cdot e_1 \otimes \dots \otimes e_m$  proving that  $Q \subseteq \mathcal{S}(e_1 \otimes \dots \otimes e_m)$

(1) and (2) gives finally that  $Q = \mathcal{S}(\{e_1 \otimes \dots \otimes e_m\})$  for all tensor spaces  $\square$

**Theorem 10.399.** Let  $\{V_i\}_{i \in \{1, \dots, m\}}$  be finite dimensional vector spaces over a field  $F$  with  $\forall i \in I$   $\{e_{i,j}\}_{j \in \{1, \dots, k_i\}}$  a basis of  $V_i$  then  $\{e_{1,\gamma(1)} \otimes \dots \otimes e_{m,\gamma(m)}\}_{\gamma \in \Gamma(k_1, \dots, k_m)}$  is a basis of  $V_1 \otimes \dots \otimes V_m$ . So the dimension of  $V_1 \otimes \dots \otimes V_m$  is  $\prod_{i \in \{1, \dots, m\}} k_i$

**Proof.** First remark that from 10.153 and the definition of a direct sum it follows that  $\forall i \in I$ ,  $\forall j \in \{1, \dots, k_i\}$  we have that  $\mathcal{S}(\{e_{i,j}\}) = \{\lambda \cdot e_{i,j} \mid \lambda \in F\}$  and using 10.153 we have then that  $V_i = \sum_{j \in \{1, \dots, k_i\}}^{\oplus} \mathcal{S}(e_{i,j})$ . Using 10.397 we have then that  $V_1 \otimes \dots \otimes V_m = \sum_{\gamma \in \Gamma(k_1, \dots, k_m)}^{\otimes} (\mathcal{S}(\{e_{1,\gamma(1)}\}) \otimes \dots \otimes \mathcal{S}(\{e_{m,\gamma(m)}\}))$ . As  $\forall i \in I, \forall j \in \{1, \dots, k_i\}$  we have  $e_{i,j} \neq 0$  (they are part of a basis) we use the previous lemma to prove that  $\mathcal{S}(\{e_{1,\gamma(1)}\}) \otimes \dots \otimes \mathcal{S}(\{e_{m,\gamma(m)}\}) = \mathcal{S}(e_{1,\gamma(1)} \otimes \dots \otimes e_{m,\gamma(m)})$ . This gives  $V_1 \otimes \dots \otimes V_m = \sum_{\gamma \in \Gamma(k_1, \dots, k_m)}^{\otimes} \mathcal{S}(e_{1,\gamma(1)} \otimes \dots \otimes e_{m,\gamma(m)})$  proving by 10.153 and the definition of a direct sum that  $\{e_{1,\gamma(1)} \otimes \dots \otimes e_{m,\gamma(m)}\}_{\gamma \in \Gamma(k_1, \dots, k_m)}$  forms a basis of  $V_1 \otimes \dots \otimes V_m$ .  $\square$

Using the above theorem we can then define the components of a the tensor product of finite dimensional spaces.

**Definition 10.400.** Let  $\{V_i\}_{i \in \{1, \dots, m\}}$  be finite dimensional vector spaces over a field  $F$  with  $\forall i \in I$   $\{e_{i,j}\}_{j \in \{1, \dots, k_i\}}$  a basis of  $V_i$  then by the previous theorem we have that  $\{e_{1,\gamma(1)} \otimes \dots \otimes e_{m,\gamma(m)}\}_{\gamma \in \Gamma(k_1, \dots, k_m)}$  is a basis of  $V_1 \otimes \dots \otimes V_m$  so that we can find for a  $T \in V_1 \otimes \dots \otimes V_m$  a  $\{T_{\gamma}\}_{\gamma \in \Gamma(k_1, \dots, k_m)}$  so that  $T = \sum_{\gamma \in \Gamma(k_1, \dots, k_m)} T_{\gamma} \cdot (e_{1,\gamma(1)} \otimes \dots \otimes e_{m,\gamma(m)})$  the family  $\{T_{\gamma}\}_{\gamma \in \Gamma(k_1, \dots, k_m)}$  of  $\prod_{i \in I} k_i$  elements of  $F$  is called the family of components of the tensor  $T$ . As a element  $\gamma \in \Gamma(k_1, \dots, k_m)$  is written as  $(\gamma(1), \dots, \gamma(m)) = (\gamma_1, \dots, \gamma_m)$  where  $\forall i \in \{1, \dots, m\}$  we have  $\gamma(i) = \gamma_i \in \{1, \dots, k_i\}$  we can write the components also as  $\{T_{\gamma_1, \dots, \gamma_m}\}_{\gamma \in \Gamma(k_1, \dots, k_m)}$

**Definition 10.401.** If  $V$  is a vector space over a field  $F$  then if  $p, q \in \mathbb{N}_0$  then  $V_q^p = V_1 \otimes \dots \otimes V_m$  where  $m = q + p$  and  $\forall i \in \{1, \dots, p\}$  we have  $V_i = V$  and  $\forall i \in \{p+1, \dots, p+q = m\}$   $V_i = V^*$  (see 10.188). Elements of  $V_q^p$  are called  $p$ -contra-variant,  $q$ -co-variant tensors. As a shorthand we write

$$V_q^p = \overbrace{V \otimes \dots \otimes V}^p \otimes \overbrace{V^* \otimes \dots \otimes V^*}^q. \text{ If } p, q = 0 \text{ then we define } V_0^0 = F$$

**Theorem 10.402.** Let  $V$  be a vector space over a field  $F$  then if  $p, q \in \mathbb{N}_0$ . For  $s \in \{1, \dots, p\}$ ,  $t \in \{p+1, \dots, p+q\}$  define  $\varphi_{s,t}: \overbrace{V \times \dots \times V}^p \times \overbrace{V^* \times \dots \times V^*}^q \rightarrow V_{q-1}^{p-1}$  defined by  $\varphi_{s,t}(v_1, \dots, v_p, v_{p+1}, \dots, v_{p+q}) = v_t(v_s) \cdot v_1 \otimes \dots \otimes v_{s-1} \otimes v_{s+1} \otimes \dots \otimes v_p \otimes v_{p+1} \otimes \dots \otimes v_{t-1} \otimes v_{t+1} \otimes \dots \otimes v_{p+q}$  then  $\varphi_{s,t}$  is multilinear and thus  $\varphi_{s,t} \in \text{Hom}\left(\overbrace{V, \dots, V}^p, \overbrace{V^*, \dots, V^*}^q, V_{q-1}^{p-1}\right)$ . By definition of the tensor product there exists a mapping  $c_{s,t}: V_q^p \rightarrow V_{q-1}^{p-1}$  such that  $c_{s,t} \circ \otimes = \varphi_{s,t}$  or  $c_{s,t}(v_1 \otimes \dots \otimes v_p \otimes v_{p+1} \otimes \dots \otimes v_{p+q}) = v_t(v_s) \cdot v_1 \otimes \dots \otimes v_{s-1} \otimes v_{s+1} \otimes \dots \otimes v_p \otimes v_{p+1} \otimes \dots \otimes v_{t-1} \otimes v_{t+1} \otimes \dots \otimes v_{p+q}$ .  $c_{s,t}$  is called a contraction mapping

**Proof.** To prove that  $\varphi_{s,t}$  is linear take  $i \in \{1, \dots, p+q\}$  then we have either

1. ( $i = s$ ) then  $\varphi_{s,t}(v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_{p+q}) = \varphi_{s,t}(v_1, \dots, v_{s-1}, \alpha \cdot v + \beta \cdot v', v_{s+1}, \dots, v_p, v_{p+1}, \dots, v_{p+q}) = v_t(\alpha \cdot v + \beta \cdot v') \cdot v_1 \otimes \dots \otimes v_{s-1} \otimes v_{s+1} \otimes \dots \otimes v_p \otimes v_{p+1} \otimes \dots \otimes v_{t-1} \otimes v_{t+1} \otimes \dots \otimes v_{p+q} \stackrel{v_t \in V^*}{=} \alpha \cdot v_t(v) \cdot v_1 \otimes \dots \otimes v_{s-1} \otimes v_{s+1} \otimes \dots \otimes v_p \otimes v_{p+1} \otimes \dots \otimes v_{t-1} \otimes v_{t+1} \otimes \dots \otimes v_{p+q} + \beta \cdot v_t(v') \cdot v_{-1} \otimes \dots \otimes v_{-(s-1)} \otimes v_{-(s+1)} \otimes \dots \otimes v_{-p} \otimes v_{-(p+1)} \otimes \dots \otimes v_{-(t-1)} \otimes v_{-(t+1)} \otimes \dots \otimes v_{-(p+q)} = \alpha \cdot \varphi_{s,t}(v_1, \dots, v_{s-1}, v, v_{s+1}, \dots, v_p, v_{p+1}, \dots, v_{p+q}) + \beta \cdot \varphi_{s,t}(v_1, \dots, v_{s-1}, v', v_{s+1}, \dots, v_p, v_{p+1}, \dots, v_{p+q}) = \alpha \cdot \varphi_{s,t}(v_1, \dots, v_{i-1}, v, v_{i+1}, \dots, v_{p+q}) + \beta \cdot \varphi_{s,t}(v_1, \dots, v_{i-1}, v', v_{i+1}, \dots, v_{p+q})$
2. ( $i = t$ ) then  $\varphi_{s,t}(v_1, \dots, v_{i-1}, \alpha \cdot v + \beta \cdot v', v_{i+1}, \dots, v_{p+q}) = \varphi_{s,t}(v_1, \dots, v_p, v_{p+1}, \dots, v_{t-1}, \alpha \cdot v + \beta \cdot v', v_{t+1}, \dots, v_{p+q}) = (\alpha \cdot v + \beta \cdot v')(v_s) \cdot v_1 \otimes \dots \otimes v_{s-1} \otimes v_{s+1} \otimes \dots \otimes v_p \otimes v_{p+1} \otimes \dots \otimes v_{t-1} \otimes v_{t+1} \otimes \dots \otimes v_{p+q} = \alpha \cdot v(v_s) \cdot v_1 \otimes \dots \otimes v_{s-1} \otimes v_{s+1} \otimes \dots \otimes v_p \otimes v_{p+1} \otimes \dots \otimes v_{t-1} \otimes v_{t+1} \otimes \dots \otimes v_{p+q} + \beta \cdot v'(v_s) \cdot v_1 \otimes \dots \otimes v_{s-1} \otimes v_{s+1} \otimes \dots \otimes v_p \otimes v_{p+1} \otimes \dots \otimes v_{t-1} \otimes v_{t+1} \otimes \dots \otimes v_{p+q} = \alpha \cdot \varphi_{s,t}(v_1, \dots, v_p, v_{p+1}, \dots, v_{t-1}, v, v_{t+1}, \dots, v_{p+q}) + \beta \cdot \varphi_{s,t}(v_1, \dots, v_p, v_{p+1}, \dots, v_{t-1}, v', v_{t+1}, \dots, v_{p+q}) = \alpha \cdot \varphi_{s,t}(v_1, \dots, v_{i-1}, v, v_{i+1}, \dots, v_{p+q}) + \beta \cdot \varphi_{s,t}(v_1, \dots, v_{i-1}, v', v_{i+1}, \dots, v_{p+q})$
3. ( $i \neq s, t$ ) then we have the following sub-cases to consider as  $s < t$ 
  - a. ( $i < s < t$ )  $\varphi_{s,t}(v_1, \dots, v_{i-1}, \alpha \cdot v + \beta \cdot v', v_{i+1}, \dots, v_{p+q}) = v_t(v_s) \cdot v_1 \otimes \dots \otimes v_{i-1} \otimes (\alpha \cdot v + \beta \cdot v') \otimes v_{i+1} \otimes \dots \otimes v_{s-1} \otimes v_{s+1} \otimes \dots \otimes v_{t-1} \otimes v_{t+1} \otimes \dots \otimes v_{p+q} \stackrel{\otimes \text{ is multilinear}}{=} \alpha \cdot v_t(v_s) \cdot v_1 \otimes \dots \otimes v_{i-1} \otimes v \otimes v_{i+1} \otimes \dots \otimes v_{s-1} \otimes v_{s+1} \otimes \dots \otimes v_{t-1} \otimes v_{t+1} \otimes \dots \otimes v_{p+q} + \beta \cdot v_t(v_s) \cdot v_1 \otimes \dots \otimes v_{i-1} \otimes v' \otimes v_{i+1} \otimes \dots \otimes v_{s-1} \otimes v_{s+1} \otimes \dots \otimes v_{t-1} \otimes v_{t+1} \otimes \dots \otimes v_{p+q} = \alpha \cdot \varphi_{s,t}(v_1, \dots, v_{i-1}, v, v_{i+1}, \dots, v_{p+q}) + \beta \cdot \varphi_{s,t}(v_1, \dots, v_{i-1}, v', v_{i+1}, \dots, v_{p+q})$
  - b. ( $s < i < t$ )  $\varphi_{s,t}(v_1, \dots, v_{i-1}, \alpha \cdot v + \beta \cdot v', v_{i+1}, \dots, v_{p+q}) = v_t(v_s) \cdot v_1 \otimes \dots \otimes v_{s-1} \otimes v_{s+1} \otimes \dots \otimes v_{i-1} \otimes (\alpha \cdot v + \beta \cdot v') \otimes \dots \otimes v_{t-1} \otimes v_{t+1} \otimes \dots \otimes v_{p+q} \stackrel{\otimes \text{ is multilinear}}{=} \alpha \cdot v_t(v_s) \cdot v_1 \otimes \dots \otimes v_{s-1} \otimes v_{s+1} \otimes \dots \otimes v_{i-1} \otimes v \otimes v_{i+1} \otimes \dots \otimes v_{t-1} \otimes v_{t+1} \otimes \dots \otimes v_{p+q} + \beta \cdot v_t(v_s) \cdot v_1 \otimes \dots \otimes v_{s-1} \otimes v_{s+1} \otimes \dots \otimes v_{i-1} \otimes v' \otimes v_{i+1} \otimes \dots \otimes v_{t-1} \otimes v_{t+1} \otimes \dots \otimes v_{p+q} = \alpha \cdot \varphi_{s,t}(v_1, \dots, v_{i-1}, v, v_{i+1}, \dots, v_{p+q}) + \beta \cdot \varphi_{s,t}(v_1, \dots, v_{i-1}, v', v_{i+1}, \dots, v_{p+q})$
  - c. ( $s < t < i$ )  $\varphi_{s,t}(v_1, \dots, v_{i-1}, \alpha \cdot v + \beta \cdot v', v_{i+1}, \dots, v_{p+q}) = v_t(v_s) \cdot v_1 \otimes \dots \otimes v_{s-1} \otimes v_{s+1} \otimes \dots \otimes v_{t-1} \otimes v_{t+1} \otimes \dots \otimes v_{i-1} \otimes (\alpha \cdot v + \beta \cdot v') \otimes v_{i+1} \otimes \dots \otimes v_{p+q} \stackrel{\otimes \text{ is multilinear}}{=} \alpha \cdot v_t(v_s) \cdot v_1 \otimes \dots \otimes v_{s-1} \otimes v_{s+1} \otimes \dots \otimes v_{t-1} \otimes v_{t+1} \otimes \dots \otimes v_{i-1} \otimes v \otimes v_{i+1} \otimes \dots \otimes v_{p+q} + \beta \cdot v_t(v_s) \cdot v_1 \otimes \dots \otimes v_{s-1} \otimes v_{s+1} \otimes \dots \otimes v_{t-1} \otimes v_{t+1} \otimes \dots \otimes v_{i-1} \otimes v' \otimes v_{i+1} \otimes \dots \otimes v_{p+q} = \alpha \cdot \varphi_{s,t}(v_1, \dots, v_{i-1}, v, v_{i+1}, \dots, v_{p+q}) + \beta \cdot \varphi_{s,t}(v_1, \dots, v_{i-1}, v', v_{i+1}, \dots, v_{p+q})$

(1),(2) and (3) proves then that  $\varphi_{s,t}$  is multilinear.  $\square$

# Chapter 11

## Topology

### 11.1 Topological spaces

**Definition 11.1.** A topological space  $\langle X, \mathcal{T} \rangle$  is a set  $X$  together with  $\mathcal{T} \subseteq \mathcal{P}(X)$  (a set of subsets of  $X$ ) such that:

1.  $\emptyset \in \mathcal{T}$
2.  $X \in \mathcal{T}$
3.  $\forall U, B \in \mathcal{T}$  we have  $U \cap V \in \mathcal{T}$
4. If  $\{U_i\}_{i \in I}$  is a family of sets in  $\mathcal{T}$  then  $\bigcup_{i \in I} U_i \in \mathcal{T}$

We call  $\mathcal{T}$  the topology of  $X$  and elements of  $\mathcal{T}$  open sets.

**Example 11.2.**  $\langle \emptyset, \{\emptyset\} \rangle$  is a trivially a topological space

**Example 11.3.**  $\langle X, \mathcal{P}(X) \rangle$  is a trivially a topological space.

**Theorem 11.4.** Let  $\langle X, \mathcal{T} \rangle$  be a topological space and let  $\{U_i\}_{i \in I}$  be a finite family of open sets then  $\bigcap_{i \in I} U_i$  is open.

**Proof.** As  $I$  is finite we have the following possibilities

1. ( $I = \emptyset$ ) then if  $x \in X$  we have  $\forall i \in I | x \in U_i$  is fulfilled vacuously so  $\bigcap_{i \in I} U_i = X \in \mathcal{T}$
2. ( $I \neq \emptyset$ ) then there exists a  $n \in \mathbb{N}_0$  and a bijection  $i: \{1, \dots, n\} \rightarrow I$  so that  $\bigcap_{k \in I} U_k = \bigcap_{j \in \{1, \dots, n\}} U_{i(j)}$ . We prove the rest by induction so let  $X = \{n \in \mathbb{N}_0 | \text{if } \{U_{i(j)}\}_{j \in \{1, \dots, n\}} \text{ is open then } \bigcap_{j \in \{1, \dots, n\}} U_{i(j)} \text{ is open}\}$  then:
  - a. If  $n = 1$  then  $\bigcap_{j \in \{1\}} U_{i(j)} = U_{i(1)} \in \mathcal{T}$
  - b. If  $n \in X$  then  $\bigcup_{j \in \{1, \dots, n+1\}} U_{i(j)} \stackrel{1.101}{=} \left( \bigcup_{j \in \{1, \dots, n\}} U_{i(j)} \right) \cup U_{i(n+1)}$  which is open because of  $n \in X$  and (3) of the definition of a topology.  $\square$

**Theorem 11.5.** Let  $\langle X, \mathcal{T} \rangle$  be a topological space then we have  $U \in \mathcal{T}$  iff  $\forall x \in U$  we have  $\exists V_x \in \mathcal{T}$  such that  $x \in V_x \subseteq U$

**Proof.**

1. ( $\Rightarrow$ ) If  $U \in \mathcal{T}$  then we have either
  - a. ( $U = \emptyset$ ) then  $\forall x \in \emptyset \models \exists V_x \in \mathcal{T} \vdash x \in V_x \subseteq \emptyset$  is satisfied vacuously
  - b. ( $U \neq \emptyset$ ) then  $\forall x \in U$  we have  $x \in U \subseteq U$
2. ( $\Leftarrow$ ) By the hypothesis we have that  $\forall x \in U$  the set  $\mathcal{A}_x = \{V \in \mathcal{T} | x \in V \subseteq U\}$  is not empty, so by the axiom of choice (see 2.210) that there exists a function  $j: U \rightarrow \bigcup_{x \in U} \mathcal{A}_x$  and this defines then a family  $\{j(x)\}_{x \in U}$  of open sets such that  $x \in j(x) \subseteq U$  so that  $U = \bigcup_{x \in U} j(x)$  is open by the definition of a topology.  $\square$

We define now the subspace topology

**Theorem 11.6.** Let  $\langle X, \mathcal{T} \rangle$  be a topological space and  $Y \subseteq X$  then if we define  $\mathcal{T}_Y = \{U \cap Y \mid U \in \mathcal{T}\}$  we have that  $\langle Y, \mathcal{T}_Y \rangle$  forms a topological space.  $\mathcal{T}_Y$  is called the subspace topology induced by  $\mathcal{T}$  on  $Y$ .

**Proof.**

1.  $\emptyset = \emptyset \cap Y \in \mathcal{T}_Y$
2.  $Y = X \cap Y \in \mathcal{T}_Y$
3. If  $U_1, U_2 \in \mathcal{T}_Y$  then there exists  $V_1, V_2 \in \mathcal{T}$  such that  $U_1 = V_1 \cap Y, U_2 = V_2 \cap Y$  then  $U_1 \cap U_2 = (V_1 \cap Y) \cap (V_2 \cap Y) = (V_1 \cap V_2) \cap (Y \cap Y) = (V_1 \cap V_2) \cap Y \in \mathcal{T}_Y$  as  $V_1 \cap V_2 \in \mathcal{T}$
4. If  $\{U_i\}_{i \in I}$  is a family of open sets in  $\mathcal{T}$  then  $\forall i \in I$  there exists a  $V_i$  such that  $U_i = V_i \cap Y$ , this defines a family  $\{V_i\}_{i \in I}$  then  $\bigcup_{i \in I} U_i = \bigcup_{i \in I} (V_i \cap Y) \stackrel{1.101}{=} (\bigcup_{i \in I} V_i) \cap Y \in \mathcal{T}_Y$  as  $\bigcup_{i \in I} V_i \in \mathcal{T}$   $\square$

**Remark 11.7.** If  $\langle X, \mathcal{T} \rangle$  is a topological space and  $U \in \mathcal{T}$  then if  $V \in \mathcal{T}_U$  then  $\exists V' \in \mathcal{T}$  such that  $V = V' \cap U \in \mathcal{T}$ . So in the case of  $U \in \mathcal{T}$  (or  $U$  is open) we have  $\mathcal{T}_U \subseteq \mathcal{T}$  (all open sets in the subspace topology induced on a open set are open in the inducing topology).

The following theorem proves essentially that the subspace topology of a subspace topology is a subspace topology.

**Theorem 11.8.** Let  $\langle X, \mathcal{T} \rangle$  be a topological space and  $Z \subseteq Y \subseteq X$  then  $(\mathcal{T}_Y)_Z = \mathcal{T}_Z$

**Proof.** First we have

$$\begin{aligned} U \in (\mathcal{T}_Y)_Z &\Rightarrow \exists V \in \mathcal{T}_Y \text{ such that } U = V \cap Z \\ &\Rightarrow \exists W \in \mathcal{T} \text{ such that } V = W \cap Y \\ &\Rightarrow U = (W \cap Z) \cap Y = W \cap (Z \cap Y) \\ &\stackrel{Z \subseteq Y}{\Rightarrow} U = W \cap Z \\ &\Rightarrow U \in \mathcal{T}_Z \end{aligned}$$

Also we have

$$\begin{aligned} U \in \mathcal{T}_Z &\Rightarrow \exists W \in \mathcal{T} \text{ such that } U = W \cap Z \\ &\stackrel{Z \subseteq Y}{\Rightarrow} U = W \cap (Y \cap Z) = (W \cap Y) \cap Z \\ &\stackrel{W \cap Y \in \mathcal{T}_Y}{\Rightarrow} U \in (\mathcal{T}_Y)_Z \\ &\square \end{aligned}$$

**Definition 11.9.** Let  $\langle X, \mathcal{T} \rangle$  be a topological space and  $A \in X$  then  $A^\circ = \bigcup_{U \in \{V \in \mathcal{T} \mid V \subseteq A\}} U$  is called the inner set of  $A$  (note that  $\emptyset \in \{V \in \mathcal{T} \mid V \subseteq A\}$ ). Note that  $A^\circ$  is open (see (4) in the definition of a topology) and that  $A^\circ \subseteq A$ .

**Theorem 11.10.** Let  $\langle X, \mathcal{T} \rangle$  be a topological space and  $A \in X$  then if  $U$  is a open set such that  $U \subseteq A$  then  $U \subseteq A^\circ$ . So  $A^\circ$  is the biggest open set contained in  $A$ .

**Proof.** If  $U$  is open and  $U \subseteq A$  then  $U \in \{V \in \mathcal{T} \mid V \subseteq A\}$  so that  $U \subseteq \bigcup_{V \in \{V \in \mathcal{T} \mid V \subseteq A\}} V = A^\circ$   $\square$

### 11.1.1 Closed sets

**Definition 11.11.** Let  $\langle X, \mathcal{T} \rangle$  be a topological space then  $\mathcal{C} \subseteq \mathcal{P}(X)$  the set of closed sets is defined by  $\mathcal{C} = \{A \in \mathcal{P}(X) \mid X \setminus A \in \mathcal{T}\}$

**Note 11.12.** If  $\langle X, \mathcal{T} \rangle$  is a topological space and  $A$  is a closed set then  $X \setminus A$  is open

**Proof.** If  $A$  is open then  $\exists U$  open such that  $A = X \setminus U = X \cap U^c$  and then  $X \setminus A = X \cap A^c = X \cap (X \cap U^c)^c = X \cap (X^c \cup U) = X \cap U = U$  open  $\square$

The next theorem proves that we can also define a topology using closed sets

**Theorem 11.13.** *Let  $\langle X, \mathcal{T} \rangle$  be a topological space then the set of closed sets  $\mathcal{C}$  satisfies the following:*

1.  $\emptyset \in \mathcal{C}$
2.  $X \in \mathcal{C}$
3. If  $A, B \in \mathcal{C}$  then  $A \cup B \in \mathcal{C}$
4. If  $I$  is finite and  $\{A_i\}_{i \in I} \subseteq \mathcal{C}$  is a finite family of closed sets then  $\bigcup_{i \in I} A_i \in \mathcal{C}$
5. If  $\{A_i\}_{i \in I}$  is a family in  $\mathcal{C}$  (a family of closed sets) then  $\bigcap_{i \in I} A_i \in \mathcal{C}$

Furthermore when  $\mathcal{C} \subseteq \mathcal{P}(X)$  is a set of subsets of  $X$  satisfying 1,2,3 and 4 of the above then  $\mathcal{T} = \{X \setminus A \mid A \in \mathcal{C}\}$  then  $\langle X, \mathcal{T} \rangle$  is a topology with the set of closed sets  $\mathcal{C}$ .

**Proof.**

1.  $X \setminus \emptyset = X \in \mathcal{T} \Rightarrow \emptyset \in \mathcal{C}$
2.  $X \setminus X = \emptyset \in \mathcal{T} \Rightarrow X \in \mathcal{C}$
3. If  $A, B \in \mathcal{C}$  then  $X \setminus (A \cup B) \stackrel{1.28}{=} (X \setminus A) \cap (X \setminus B) \in \mathcal{T}$  (as  $X \setminus A, X \setminus B \in \mathcal{T}$  and the definition of a topology).
4. If  $\{A_i\}_{i \in I} \subseteq \mathcal{C}$  then  $\forall i \in I$  we have that  $X \setminus A_i$  is open. As  $I$  is finite we have by 11.4 that  $\bigcap_{i \in I} (X \setminus A_i)$  is open, hence  $X \setminus (\bigcap_{i \in I} (X \setminus A_i))$  is closed. Finally  $X \setminus (\bigcap_{i \in I} (X \setminus A_i)) \stackrel{1.102}{=} \bigcup_{i \in I} (X \setminus (X \setminus A_i)) \stackrel{1.19}{=} \bigcup_{i \in I} A_i$
5. If  $\{A_i\}_{i \in I}$  is a family in  $\mathcal{C}$  then  $X \setminus (\bigcap_{i \in I} A_i) \stackrel{1.102}{=} \bigcup_{i \in I} (X \setminus A_i) \in \mathcal{T}$  (as  $\forall i \in I$  we have  $X \setminus A_i \in \mathcal{T}$ )

Assume now that  $\mathcal{C} \subseteq \mathcal{P}(X)$  fulfills 1,2,3 and 4 and define  $\mathcal{T} = \{U \in \mathcal{P}(X) \mid X \setminus U \in \mathcal{C}\}$  then

1.  $\emptyset = X \setminus X \Rightarrow \emptyset \in \mathcal{T}$
2.  $X = X \setminus \emptyset \Rightarrow X \in \mathcal{T}$
3. If  $U, V \in \mathcal{T}$  then  $\exists A, B \in \mathcal{C}$  such that  $U = X \setminus A, V = X \setminus B$  and  $U \cap V = (X \setminus A) \cap (X \setminus B) = X \setminus (A \cup B) \in \mathcal{T}$  (as  $A \cup B \in \mathcal{C}$ )
4. If  $\{U_i\}_{i \in I}$  is a family in  $\mathcal{T}$  then  $\forall i \in I$  there exists  $A_i \in \mathcal{C}$  such that  $U_i = X \setminus A_i$  and thus  $\bigcup_{i \in I} U_i = \bigcup_{i \in I} (X \setminus A_i) \stackrel{1.102}{=} X \setminus (\bigcap_{i \in I} A_i)$

Now if  $A$  is closed in  $\mathcal{T}$  then  $X \setminus A \in \mathcal{T}$  so there exists a  $C \in \mathcal{C}$  such that  $X \setminus A = X \setminus C$  we have then  $A \stackrel{A \subseteq X \wedge 1.28}{=} X \setminus (X \setminus A) \stackrel{X \setminus A = X \setminus C}{=} X \setminus (X \setminus C) \stackrel{C \subseteq X \wedge 1.28}{=} C$  and thus the set of closed sets of  $\mathcal{T}$  is a subset of  $\mathcal{C}$ . If  $A \in \mathcal{C}$  then by definition  $X \setminus A \in \mathcal{T}$  so  $A$  is closed in  $\mathcal{T}$  and thus  $\mathcal{C}$  is a subset of the set of closed sets in  $\mathcal{T}$ . These two last remarks proves that  $\mathcal{C}$  is indeed the set of closed sets of  $\mathcal{T}$ .  $\square$

**Theorem 11.14.** *Let  $\langle X, \mathcal{T} \rangle$  be a topological space and  $A \subseteq X$  then the set of closed sets of  $\langle A, \mathcal{T}_A \rangle$  is  $\{C \cap A \mid C \text{ is closed in } \langle X, \mathcal{T} \rangle\}$*

**Proof.** If  $B$  is closed in  $\langle A, \mathcal{T}_A \rangle$  then  $A \setminus B \in \mathcal{T}_A \Rightarrow \exists U \in \mathcal{T}$  such that  $(A \setminus B) = U \cap A$  then  $B \stackrel{B \subseteq A \wedge 1.28}{=} B = A \setminus (A \setminus B) = A \setminus (U \cap A) = (A \setminus U) \cup (A \setminus A) = (A \setminus U) \cup \emptyset = A \setminus U = A \cap U^c \stackrel{A \subseteq X}{=} (A \cap X) \cap U^c = A \cap (X \cap U^c) = A \cap (X \setminus U) = (X \setminus U) \cap A \in \{C \cap A \mid C \text{ is closed in } \mathcal{T}\}$ . If  $B \in \{C \cap A \mid C \text{ is closed in } \langle X, \mathcal{T} \rangle\}$  then there exists a  $C \subseteq X$  such that  $X \setminus C \in \mathcal{T}$  and  $B = C \cap A$ . Now  $A \setminus B = A \setminus (C \cap A) = (A \setminus C) \cup (A \setminus A) = A \setminus C \stackrel{A \subseteq X}{=} (A \cap X) \setminus C = (A \cap X) \cap C^c = A \cap (X \cap C^c) = A \cap (X \setminus C) \in \mathcal{T}_A$  as  $X \setminus C \in \mathcal{T}$  so we have that  $B$  is closed in  $\langle A, \mathcal{T}_A \rangle$ .  $\square$

**Definition 11.15.** *Let  $\langle X, \mathcal{T} \rangle$  be a topological space and  $A \subseteq X$  then  $\bar{A} = \bigcap_{C \in \{C \mid C \text{ is closed and } A \subseteq C\}} C$  is closed (as a intersection of a family of closed sets), contains  $A$  and is called the closure of  $A$  (note that  $X \in \{C \mid C \text{ is closed and } A \subseteq C\} \Rightarrow \{C \mid C \text{ is closed and } A \subseteq C\} \neq \emptyset$ )*

**Theorem 11.16.** *Let  $\langle X, \mathcal{T} \rangle$  be a topological space and  $A \subseteq X$  then  $\bar{A}$  is the smallest closed set containing  $A$*

**Proof.** If  $C$  is a closed set such that  $A \subseteq C$  then  $C \in \{C \mid C \text{ is closed and } A \subseteq C\}$  and thus  $\bigcap_{B \in \{C \mid C \text{ is closed and } A \subseteq C\}} B \subseteq C$   $\square$

**Theorem 11.17.** If  $\langle X, \mathcal{T} \rangle$  is a topological space then  $A \subseteq X$  is closed if and only if  $A = \bar{A}$

**Proof.**

- ( $\Rightarrow$ ) If  $A$  is closed then as  $A \subseteq A$  we have that  $A \subseteq \bar{A} \subseteq A \Rightarrow A = \bar{A}$
- ( $\Leftarrow$ ) If  $A = \bar{A}$  then as  $\bar{A}$  is closed we have that  $A$  is closed

**Definition 11.18.** Let  $\langle X, \mathcal{T} \rangle$  be a topological space and  $A \subseteq X$  then  $x \in X$  is a limit point or accumulation point of  $A$  iff  $\forall U \in \mathcal{T}$  with  $x \in U$  we have  $\emptyset \neq (A \setminus \{x\}) \cap U$ . The set of all limit points of  $A$  is called the derived set of  $A$  and is denoted by  $A'$ .

**Theorem 11.19.** Let  $\langle X, \mathcal{T} \rangle$  be a topological space and  $A \subseteq X$  then  $\bar{A} = A \cup A'$

**Proof.** If  $x \in \bar{A}$  then we have two cases

1. ( $x \in A$ ) then trivially  $x \in A \cup A'$
2. ( $x \notin A$ ) then if  $x \notin A'$  then there exists a  $U \in \mathcal{T}$  with  $x \in U$  such that  $\emptyset \neq (A \setminus \{x\}) \cap U \underset{x \in A \Rightarrow \bar{A} = A \setminus \{x\}}{=} A \cap U \Rightarrow \forall a \in A \subseteq X$  we have  $a \notin U \wedge a \in X \Rightarrow A \subseteq X \setminus U$  which means as  $X \setminus U$  is closed that  $\bar{A} \subseteq X \setminus U \Rightarrow x \notin U$  contradicting  $x \in U$ . So we must have  $x \in A' \Rightarrow x \in A \cup A'$

So we have proved that  $\bar{A} \subseteq A \cup A'$ .

If now  $x \in A \cup A'$  then we have either

1. ( $x \in A$ ) then as  $A \subseteq \bar{A}$  we have  $x \in \bar{A}$
2. ( $x \notin A$ ) then we must have  $x \in A'$ . If  $x \notin \bar{A}$  then  $x \in X \setminus \bar{A}$  which is open and thus we have  $\emptyset \neq (A \setminus \{x\}) \cap (X \setminus \bar{A}) \underset{x \notin A \Rightarrow \bar{A} = A \setminus \{x\}}{=} A \cap (X \setminus \bar{A})$  but then there exists a  $y \in A \wedge y \notin \bar{A} \underset{y \in A \subseteq \bar{A}}{\Rightarrow} y \in \bar{A} \wedge y \in \bar{A}$  a contradiction. So we must have  $x \in \bar{A}$

This proves then  $A \cup A' \subseteq \bar{A}$  and thus finally  $\bar{A} = A \cup A'$   $\square$

**Corollary 11.20.** Let  $\langle X, \mathcal{T} \rangle$  be a topological space and  $A \subseteq X$  then  $x \in \bar{A}$  if and only if  $\forall U \in \mathcal{T}$  with  $x \in U$  we have  $U \cap A \neq \emptyset$ . Or in other words  $\bar{A} = \{x \in X \mid \forall U \in \mathcal{T} \text{ with } x \in U \text{ we have } A \cap U \neq \emptyset\}$

**Proof.** If  $x \in \bar{A} = A \cup A'$  then either

1. ( $x \in A$ ) then trivially  $\forall U \in \mathcal{T}$  with  $x \in U$  we have  $x \in U \cap A \Rightarrow U \cap A \neq \emptyset$
2. ( $x \in A'$ ) then  $\forall U \in \mathcal{T}$  with  $x \in U$  we have  $\emptyset \neq (A \setminus \{x\}) \cap U \Rightarrow U \cap A \neq \emptyset$

so in all cases we have  $\forall U \in \mathcal{T}$  with  $x \in U$  we have  $U \cap A \neq \emptyset$ .

On the other hand if  $x \in X$  is such that  $\forall U \in \mathcal{T}$  with  $x \in U$  we have that  $U \cap A \neq \emptyset$ . If now  $x \notin \bar{A}$  then  $x \in X \setminus \bar{A}$  a open set so that we have  $\emptyset \neq (X \setminus \bar{A}) \cap A = (X \cap \bar{A}^c) \cap A = X \cap (\bar{A}^c \cap A) \underset{A \subseteq \bar{A}}{=} X \cap \emptyset = \emptyset \Rightarrow \emptyset \neq \emptyset$  a contradiction. So we must have that  $x \in \bar{A}$   $\square$

As  $A$  is closed if and only if  $A = \bar{A}$  (see 11.17) we have the following

**Corollary 11.21.** Let  $\langle X, \mathcal{T} \rangle$  be a topological space and  $A \subseteq X$  then  $A$  is closed if and only if  $A = \{x \in X \mid \forall U \in \mathcal{T} \text{ with } x \in U \text{ we have } A \cap U \neq \emptyset\}$

**Definition 11.22.** Let  $X$  be a set and  $\mathcal{T}_1, \mathcal{T}_2$  two topologies for  $X$  then  $\mathcal{T}_2$  is finer than  $\mathcal{T}_1$  iff  $\mathcal{T}_1 \subseteq \mathcal{T}_2$ . So if  $\mathcal{T}_1$  is finer than  $\mathcal{T}_2$  and  $\mathcal{T}_2$  is finer than  $\mathcal{T}_1$  then  $\mathcal{T}_1 = \mathcal{T}_2$ .

### 11.1.2 Basis of a topological space

**Definition 11.23.** Let  $\langle X, \mathcal{T} \rangle$  be a topological space then  $\mathcal{B} \subseteq \mathcal{T}$  is a basis for the topology  $\mathcal{T}$  if every open set is the union of a family in  $\mathcal{B}$ . In other words  $\mathcal{B}$  is a basis for  $\mathcal{T}$  iff  $\mathcal{T} = \{U \in \mathcal{P}(X) \mid \exists \{B_i\}_{i \in I} \text{ in } \mathcal{B} \text{ such that } U = \bigcup_{i \in I} U_i\}$

**Example 11.24.** If  $U = \emptyset$  then  $\{B_i\}_{i \in \emptyset}$  (a empty family) is such that  $U = \bigcup_{i \in \emptyset} B_i$  [if  $x \in \bigcup_{i \in \emptyset} B_i$  then  $\exists i \in \emptyset$  such that  $x \in B_i$  which is impossible for the empty set, so  $\emptyset = \bigcup_{i \in \emptyset} B_i$ ].

**Theorem 11.25.** Let  $\langle X, \mathcal{T} \rangle$  be a topological space with a basis  $\mathcal{B}$  then if  $A \subseteq X$  we have that  $\mathcal{B}_A = \{B \cap A \mid B \in \mathcal{B}\}$  is a basis for the subspace topology  $\mathcal{T}_A$  of  $A$

**Proof.** Let  $U \in \mathcal{T}_A$  then there exists a  $V \in \mathcal{T}$  such that  $U = V \cap A$ . By definition of a basis there exists a family  $\{B_i\}_{i \in I}$  in  $\mathcal{B}$  (meaning  $\forall i \in I$  we have  $B_i \in \mathcal{B}$ ) such that  $V = \bigcup_{i \in I} B_i$  and thus  $U = V \cap A = (\bigcup_{i \in I} B_i) \cap A = \bigcup_{i \in I} (B_i \cap A)$  so  $V = \bigcup_{i \in I} (B_i \cap A)$  where  $\{B_i \cap A\}_{i \in I}$  is a family in  $\mathcal{B}_A$ . Finally as  $\mathcal{B} \subseteq \mathcal{T}$  we have  $\forall B \in \mathcal{B}$  that  $B \cap A \in \mathcal{T}_A$  and thus that  $\mathcal{B}_A \subseteq \mathcal{T}_A$ .  $\square$

**Theorem 11.26.** Let  $\langle X, \mathcal{T} \rangle$  be a topological space and let  $\mathcal{B} \subseteq \mathcal{T}$  then  $\mathcal{B}$  is a basis for  $\mathcal{T}$  if and only if  $\forall U \in \mathcal{T}$  and  $\forall x \in U$  we have that  $\exists B \in \mathcal{B}$  such that  $x \in B \subseteq U$

**Proof.**

$(\Rightarrow)$  If  $U \in \mathcal{T}$  then there exists a  $\{B_i\}_{i \in I}$  in  $\mathcal{B}$  such that  $U = \bigcup_{i \in I} B_i$  so that if  $x \in U$  then there exists a  $i \in I$  with  $x \in B_i \subseteq \bigcup_{i \in I} B_i \subseteq U$  where of course  $B_i \in \mathcal{B}$

$(\Leftarrow)$  If  $U \in \mathcal{T}$  then  $\forall x \in U$  there exists a  $B_x \in \mathcal{B}$  such that  $x \in B_x \subseteq U$  this forms a family  $\{B_x\}_{x \in U}$  in  $\mathcal{B}$  such that  $\bigcup_{x \in U} B_x = U$  proving that  $\mathcal{B}$  is a basis for  $\mathcal{T}$   $\square$

**Theorem 11.27.** Let  $\langle X, \mathcal{T} \rangle$  be a topological space and let  $\mathcal{B}$  be a basis for  $\mathcal{T}$  then

1.  $\forall x \in X$  we have  $\exists B \in \mathcal{B}$  with  $x \in B$
2.  $\forall B_1, B_2 \in \mathcal{B}$  then if  $x \in B_1 \cap B_2$  then  $\exists B_3 \in \mathcal{B}$  such that  $x \in B_3 \subseteq B_1 \cap B_2$

**Proof.**

1. As  $X \in \mathcal{T}$  there exists a  $\{B_i\}_{i \in I}$  in  $\mathcal{B}$  such that  $X = \bigcup_{i \in I} B_i$  and thus if  $x \in X$  there exists a  $i \in I$  such that  $x \in B_i \in \mathcal{B}$
2. If  $x \in B_1 \cap B_2$  then as  $B_1, B_2 \in \mathcal{B} \subseteq \mathcal{T} \Rightarrow B_1 \cap B_2$  there exists a  $\{C_i\}_{i \in I}$  in  $\mathcal{B}$  such that  $B_1 \cap B_2 = \bigcup_{i \in I} C_i$  so if  $x \in B_1 \cap B_2$  then there exists a  $i \in I$  such that  $x \in C_i \subseteq \bigcup_{i \in I} C_i = B_1 \cap B_2$  and thus  $x \in C_i \subseteq B_1 \cap B_2$ .  $\square$

The above show the necessary condition that a basis must full fill, the following shows that any set of subset full filling the above conditions can be the basis of a topology.

**Theorem 11.28.** Let  $X$  be as set and let  $\mathcal{B} \subseteq \mathcal{P}(X)$  be a set of subsets of  $X$  full filling

1.  $\forall x \in X$  there exists a  $B \in \mathcal{B}$  such that  $x \in B$
2.  $\forall B_1, B_2 \in \mathcal{B}$  then if  $x \in B_1 \cap B_2$  there exists a  $B_3 \in \mathcal{B}$  such that  $x \in B_3 \subseteq B_1 \cap B_2$

then

$$\mathcal{T} = \{U \in \mathcal{P}(X) \mid \forall x \in U \text{ we have } \exists B \in \mathcal{B} \text{ such that } x \in B \subseteq U\}$$

is a topology for  $X$  that has  $\mathcal{B}$  as its basis.  $\mathcal{B}$  is called the generating basis for  $\mathcal{T}$  and  $\mathcal{T}$  is called the topology generated by  $\mathcal{B}$

**Proof.** First we prove that  $\mathcal{T}$  is a topology for  $X$

- a)  $\emptyset \in \mathcal{T}$  (because every element in the empty set full fills vacuously every condition).
- b)  $X \in \mathcal{T}$  for if  $x \in X$  then by (1) there exists a  $B \in \mathcal{B}$  such that  $x \in B \subseteq X$
- c) If  $\{U_i\}_{i \in I}$  is a family in  $\mathcal{T}$  then if  $x \in \bigcup_{i \in I} U_i$  there exists a  $i \in I$  such that  $x \in U_i \in \mathcal{T}$  so there exists a  $B \in \mathcal{B}$  such that  $x \in B \subseteq U_i \subseteq \bigcup_{i \in I} U_i$  and thus  $\bigcup_{i \in I} U_i \in \mathcal{T}$
- d) If  $U_1, U_2 \in \mathcal{T}$  then if  $x \in U_1 \cap U_2 \Rightarrow x \in U_1 \wedge x \in U_2$  so there exists  $B_1, B_2 \in \mathcal{B}$  such that  $x \in B_1 \subseteq U_1 \wedge x \in B_2 \subseteq U_2 \Rightarrow x \in B_1 \cap B_2 \subseteq U_1 \cap U_2$  and by (2) there exists a  $B_3$  such that  $x \in B_3 \subseteq B_1 \cap B_2 \subseteq U_1 \cap U_2$  proving that  $U_1 \cap U_2 \in \mathcal{T}$

So  $\langle X, \mathcal{T} \rangle$  is a topological space.

Also if  $B \in \mathcal{B}$  then if  $\forall x \in B$  there exists a  $B \in \mathcal{B}$  such that  $x \in B \subseteq B$  and thus  $B \in \mathcal{T}$  giving  $\mathcal{B} \subseteq \mathcal{T}$ .

Next if  $U \in \mathcal{T}$  and if  $x \in U$  then by definition there exists a  $B \in \mathcal{B}$  such that  $x \in B \subseteq U$  which by 11.26 means that  $\mathcal{B}$  is a basis for  $\mathcal{T}$ .  $\square$

The following theorem shows how to create a topology out of a set of topologies

**Theorem 11.29.** *Let  $X$  be a set and  $\{\langle X_i, \mathcal{T}_i \rangle\}_{i \in I}$  a family of topological spaces such that  $\forall i \in I$  we have  $X_i \subseteq X$ ,  $X = \bigcup_{i \in I} X_i$  and  $\forall i, j \in I$ ,  $U \in \mathcal{T}_i$ ,  $V \in \mathcal{T}_j$  if  $x \in U \cap V$  then  $\exists W \in \mathcal{T}_i \cap \mathcal{T}_j$  such that  $x \in W \subseteq U \cap V$ . Then  $\mathcal{B} = \bigcup_{i \in I} \mathcal{T}_i$  full fills the requirement for a generating basis (see 11.28) so that  $\langle X, \mathcal{T} \rangle$  with  $\mathcal{T} = \{U \subseteq X \mid \forall x \in U \text{ we have } \exists B \in \mathcal{B} \text{ such that } x \in B \subseteq U\}$  the generated topology. Also  $\forall i \in I$  we have that  $\mathcal{T}_i = \mathcal{T}_{X_i}$  (the subspace topology for  $X_i$ ). We call this procedure to create a topology from subspace topologies **stitching topologies**.*

**Proof.** First we prove that the requirements of 11.28 are satisfied:

1. If  $x \in X = \bigcup_{i \in I} X_i \Rightarrow \exists i \in I \mid x \in X_i \underset{\mathcal{T}_i \text{ is a topology}}{\Rightarrow} \exists U \in \mathcal{T}_i \subseteq \mathcal{B}$  such that  $x \in U$ .
2. If  $B_1, B_2 \in \mathcal{B}$  and  $x \in B_1 \cap B_2$  then  $\exists i, j \in I$  with  $B_1 \in \mathcal{T}_i$ ,  $B_2 \in \mathcal{T}_j$  then  $\exists W \in \mathcal{T}_i \cap \mathcal{T}_j \subseteq \mathcal{B}$  such that  $x \in W \subseteq B_1 \cap B_2$

By 11.28 we have then that  $\langle X, \mathcal{T} \rangle$  is a topological vector space.

Now if  $i \in I$  and  $U \in \mathcal{T}_{X_i}$  then there exists a  $V \in \mathcal{T}$  such that  $U = V \cap X_i$ . Then  $\forall x \in U$  we have  $x \in V$  and thus  $\exists B_x \in \mathcal{B}$  such that  $x \in B_x \subseteq V$ , but then there exists a  $i_x \in I$  such that  $B_x \in \mathcal{T}_{i_x}$ . We have then that  $V = \bigcup_{x \in V} B_x$  where  $B_x \in \mathcal{T}_{i_x}$  and thus  $U = (\bigcup_{x \in V} B_x) \cap X_i = \bigcup_{x \in V} (B_x \cap X_i)$ . If now  $y \in B_x \cap X_i$  where  $B_x \in \mathcal{T}_{i_x}$  and  $X_i \in \mathcal{T}_i$  then by the hypothesis there exists a  $W_y \in \mathcal{T}_{i_x} \cap \mathcal{T}_i \subseteq \mathcal{T}_i$  such that  $y \in W_y \subseteq B_x \cap X_i$  so that  $B_x \cap X_i = \bigcup_{y \in B_x \cap X_i} W_y \in \mathcal{T}_i$  [as  $\mathcal{T}_i$  is a topology and  $W_y \in \mathcal{T}_i$ ]. So as  $B_x \cap X_i \in \mathcal{T}_i$  we have that  $U = V \cap X_i = \bigcup_{x \in V} (B_x \cap X_i) \in \mathcal{T}_i \Rightarrow U \in \mathcal{T}_i$  giving

$$\mathcal{T}_{X_i} \subseteq \mathcal{T}_i \quad (11.1)$$

If now  $U \in \mathcal{T}_i \subseteq \mathcal{B} \subseteq \mathcal{T}$  then as  $U \subseteq X_i$  we have  $U = U \cap X_i \in \mathcal{T}_{X_i}$  so that we have

$$\mathcal{T}_i \subseteq \mathcal{T}_{X_i} \quad (11.2)$$

Using 11.1 and 11.2 we have finally  $\mathcal{T}_i = \mathcal{T}_{X_i}$   $\square$

**Theorem 11.30.** *Let  $\langle X, \mathcal{T} \rangle$  be a topological space and  $A \subseteq X$  then*

1.  $x \in \bar{A} \Leftrightarrow \forall U \in \mathcal{T} \mid x \in U \text{ we have } U \cap A \neq \emptyset$
2. If  $\mathcal{B}$  is a basis of the topology  $\mathcal{T}$  then  $x \in \bar{A} \Leftrightarrow \forall U \in \mathcal{B} \text{ with } x \in U \text{ we have } U \cap A \neq \emptyset$

**Proof.**

1. This is already proved in 11.20
2. If  $x \in \bar{A}$  then if  $U \in \mathcal{B} \subseteq \mathcal{T}$  is such that  $x \in U$  we have by (1) that  $U \cap A \neq \emptyset$ . If  $x \in X$  is such that  $\forall B \in \mathcal{B} \mid x \in B$  we have  $B \cap A \neq \emptyset$  then if  $U \in \mathcal{T}$  is such that  $x \in U$  there exists a  $B \in \mathcal{B}$  such that  $x \in B \subseteq U$  and then  $\emptyset \neq B \cap A \subseteq U \cap A \Rightarrow U \cap A \neq \emptyset$  so that by (1) we have  $x \in \bar{A}$ .  $\square$

**Theorem 11.31.** *Let  $\langle X, \mathcal{T} \rangle$  be a topological space then if  $A \subseteq X$  is closed we have*

1.  $x \in A \Leftrightarrow \forall U \in \mathcal{T} \mid x \in U \text{ we have } U \cap A \neq \emptyset$
2. If  $\mathcal{B}$  is a basis of the topology  $\mathcal{T}$  then  $x \in A \Leftrightarrow \forall U \in \mathcal{B} \text{ with } x \in U \text{ we have } U \cap A \neq \emptyset$

**Proof.** This is trivial using 11.17 and the previous theorem 11.19.  $\square$

**Theorem 11.32.** *Let  $X$  be a set and  $\mathcal{T}_1, \mathcal{T}_2$  two topologies on  $X$  with basis  $\mathcal{B}_1, \mathcal{B}_2$  then the following is equivalent*

1.  $\mathcal{T}_2$  is finer than  $\mathcal{T}_1$
2.  $\forall x \in X, \forall B \in \mathcal{B}_1 \mid x \in B \text{ there } \exists B' \in \mathcal{B}_2 \text{ such that } x \in B' \subseteq B$

**Proof.**

1. **(1  $\Rightarrow$  2)** then if  $x \in X$  and  $B \in \mathcal{B}_1 \subseteq \mathcal{T}_1 \subseteq \mathcal{T}_2$  such that  $x \in B$  then there exists a  $B' \in \mathcal{B}_2$  such that  $x \in B' \subseteq B$ .
2. **(2  $\Rightarrow$  1)** Let  $U \in \mathcal{T}_1$  then  $\forall x \in U$  there exists a  $B_x \in \mathcal{B}_1$  such that  $x \in B_x \subseteq U$  and thus by (2)  $\exists B'_x \in \mathcal{B}_2$  such that  $x \in B'_x \subseteq B_x \subseteq U$  and thus  $U = \bigcup_{x \in U} B'_x \in \mathcal{T}_2$  proving that  $\mathcal{T}_1 \subseteq \mathcal{T}_2$   $\square$

**Theorem 11.33.** Let  $X$  be a set and  $\mathcal{S} \subseteq \mathcal{P}(X)$  then the set  $\mathcal{B} = \{B \in \mathcal{P}(S) \mid \exists \{S_i\}_{i \in I} \text{ a finite non empty family in } \mathcal{S} \text{ such that } B = \bigcap_{i \in I} S_i\} \cup \{X\}$  satisfies the conditions to become a basis (see 11.28) and generates thus a topology  $\mathcal{T} = \{U \in \mathcal{P}(X) \mid \forall x \in U \text{ there } \exists B \in \mathcal{B} \text{ such that } x \in B \subseteq U\}$ . This topology is called the topology generated by the sub-base  $\mathcal{S}$ .  $\mathcal{B}$  is called the basis generated by the sub basis.

**Proof.** We have to show that  $\mathcal{B}$  satisfies the conditions from 11.28

1. Let  $x \in X$  then  $x \in X \in \mathcal{B}$
2. Let  $x \in B_1 \cap B_2$  where  $B_1, B_2 \in \mathcal{B}$  then there exists finite families  $\{S_i\}_{i \in I_1}$  and  $\{T_i\}_{i \in I_2}$  such that  $B_1 = \bigcap_{i \in I_1} S_i$  and  $B_2 = \bigcap_{i \in I_2} T_i$ . Form then the family  $\{R_{(i,0)}\}_{i \in (I_1 \times \{0\}) \cup (I_2 \times \{1\})}$  by  $R_{(i,0)} = S_i$  and  $R_{(i,1)} = T_i$  then we have

$$\begin{aligned}
 x \in \bigcap_{i \in (I_1 \times \{0\}) \cup (I_2 \times \{1\})} R_i &\Leftrightarrow \forall (i, j) \in (I_1 \times \{0\}) \cup (I_2 \times \{1\}) \text{ we have } x \in R_{(i,j)} \\
 &\Leftrightarrow [\forall (i, j) \in I_1 \times \{0\} \text{ we have } x \in R_{(i,j)}] \wedge [\forall (i, j) \in I_2 \times \{1\} \text{ we have } x \in R_{(i,j)}] \\
 &\Leftrightarrow [\forall i \in I_1 \text{ we have } x \in R_{(i,0)}] \wedge [\forall i \in I_2 \text{ we have } x \in R_{(i,1)}] \\
 &\Leftrightarrow [\forall i \in I_1 \text{ we have } x \in S_i] \wedge [\forall i \in I_2 \text{ we have } x \in T_i] \\
 &\Leftrightarrow x \in \bigcap_{i \in I_1} S_i \wedge x \in \bigcap_{i \in I_2} T_i \\
 &\Leftrightarrow x \in B_1 \cap B_2
 \end{aligned}$$

This proves that  $B_1 \cap B_2 = \bigcap_{i \in (I_1 \times \{0\}) \cup (I_2 \times \{1\})} R_i \in \mathcal{B}$  and thus we have found a  $B = B_1 \cap B_2 \in \mathcal{B}$  such that  $x \in B \subseteq B_1 \cap B_2$ .  $\square$

**Definition 11.34.** Let  $I$  be a non empty set,  $\{\langle X_i, \mathcal{T}_i \rangle\}_{i \in I}$  be a family of topological spaces then the box topology  $\mathcal{T}_{\text{box}}$  on  $\prod_{i \in I} X_i$  is the topology generated by the basis  $\mathcal{B} = \{\prod_{i \in I} U_i \mid \forall i \in I \text{ we have } U_i \in \mathcal{T}_i\}$ .

**Proof.** We have of course prove that 11.28 is satisfied.

1. If  $x \in \prod_{i \in I} X_i$  then as  $\forall i \in I$  we have  $X_i \in \mathcal{T}_i \Rightarrow \prod_{i \in I} X_i \in \mathcal{B}$
2. Let  $U_1 = \prod_{i \in I} V_i \in \mathcal{B}$ ,  $U_2 = \prod_{i \in I} W_i \in \mathcal{B}$  then  $\forall i \in I$  we have  $V_i, W_i \in \mathcal{T}_i \Rightarrow V_i \cap W_i \in \mathcal{T}_i$  we have then  $U_1 \cap U_2 = \prod_{i \in I} (V_i \cap W_i) \in \mathcal{B}$  [as  $V_i \cap W_i \in \mathcal{T}_i$ ]  $\square$

**Theorem 11.35.** Let  $I$  be a non empty set,  $\{\langle X_i, \mathcal{T}_i \rangle\}_{i \in I}$  be a family of topological spaces so that  $\forall i \in I$  we have that  $\mathcal{B}_i$  is a basis of  $\mathcal{T}_i$ . Then  $\mathcal{B} = \{\prod_{i \in I} B_i \mid \forall i \in I \models B_i \in \mathcal{B}_i\}$  is a basis for the box topology on  $\prod_{i \in I} X_i$ .

**Proof.** Let  $U \in \mathcal{T}_{\text{box}}$  and  $x \in U$  then as  $\{\prod_{i \in I} U_i \mid \forall i \in I \models U_i \in \mathcal{T}_i\}$  we have by 11.26 that there exists a  $\{U_i\}_{i \in I}$  where  $\forall i \in I \models U_i \in \mathcal{T}_i$  such that  $x \in \prod_{i \in I} U_i \subseteq U$ . Now  $\forall i \in I$  we have  $x_i \in U_i$  and thus as  $\mathcal{B}_i$  is a basis for  $\mathcal{T}_i$  there exists a  $B \in \mathcal{B}_i$  such that  $x_i \in B \subseteq U_i$ . So if we define  $\forall i \in I \mathcal{A}_i = \{B \in \mathcal{B}_i \mid x_i \in B \subseteq U_i\}$  then  $\mathcal{A}_i \neq \emptyset$  and thus by the axiom of choice (see 2.210) there exists a choice function  $B: I \rightarrow \bigcup_{i \in I} \mathcal{A}_i$  such that  $\forall i \in I$  we have  $B_i \in \mathcal{A}_i$  and thus  $x_i \in B_i \subseteq U_i$  and  $B_i \in \mathcal{B}_i$ . This defines a family  $\{B_i\}_{i \in I}$  such that  $\forall i \in I$  we have  $B_i \in \mathcal{B}_i$  and  $x_i \in B_i \subseteq U_i$  and thus by 2.84 and 2.89 we have  $x \in \prod_{i \in I} B_i \subseteq \prod_{i \in I} U_i \subseteq U$  where  $\prod_{i \in I} B_i \in \mathcal{B}$ , using 11.26 again we have then proved that  $\mathcal{B}$  is a basis for the box topology.  $\square$

**Definition 11.36.** Let  $I$  be a non empty set,  $\{\langle X_i, \mathcal{T}_i \rangle\}_{i \in I}$  be a family of topological spaces then the product topology  $\mathcal{T}_{\text{product}}$  on  $\prod_{i \in I} X_i$  is defined by the topology generated by the sub base  $\mathcal{S} = \{\pi_i^{-1}(V) | i \in I \wedge V \in \mathcal{T}_i\}$

**Theorem 11.37.** Let  $I$  be a non empty set,  $\{\langle X_i, \mathcal{T}_i \rangle\}_{i \in I}$  be a family of topological spaces then the product topology  $\mathcal{T}_{\text{product}}$  on  $X = \prod_{i \in I} X_i$  has as its basis  $\mathcal{B} = \{\prod_{i \in I} U_i | \text{where } \{U_i\}_{i \in I} \text{ is a family such that } \exists A \text{ finite } \subseteq I \text{ such that } \forall i \in I \setminus A \text{ we have } U_i = X_i \text{ and } \forall i \in A \text{ we have } U_i \in \mathcal{T}_i\}$

**Proof.** First  $\mathcal{S} = \{\pi_i^{-1}(V) | i \in I, V \in \mathcal{T}_i\}$  is the generating sub basis of  $\mathcal{T}_{\text{product}}$  and the basis  $\mathcal{B}'$  for  $\mathcal{T}_{\text{product}}$  is given by  $\mathcal{B}' = \{B \in \mathcal{P}(S) | \exists \{S_i\}_{i \in J} \text{ a finite non empty family in } \mathcal{S} \text{ such that } B = \bigcap_{i \in J} S_i\} \cup \{X\}$ . We have then:

1. ( $B \in \mathcal{B}'$ ) then we have either

- ( $B = X$ ) then  $B = \prod_{i \in I} X_i$  and if we take  $A = \emptyset$  (which is finite) then  $\forall i \in A$  we have  $X_i \in \mathcal{T}_i$  vacuously and  $\forall i \in I \setminus A$  we have  $X_i \in \mathcal{T}_i$  and thus  $B \in \mathcal{B}'$
- ( $\exists \{S_i\}_{i \in J} \text{ a finite non empty family in } \mathcal{S} \text{ such that } B = \bigcap_{i \in J} S_i$ ) as  $\forall i \in J$  we have  $S_i \in \mathcal{S}$  there exists a  $k_i \in I$  and  $U_i \in \mathcal{T}_{k_i}$  such that  $S_i = \pi_{k_i}^{-1}(U_i)$ . Now if we take  $\{B_{k_i, k}\}_{k \in I}$  where  $B_{k_i, k} = \begin{cases} U_i & \text{if } k = k_i \\ X_k & \text{if } k \neq k_i \end{cases} \Rightarrow B_{k_i, k} \in \mathcal{T}_k$  [if  $k = k_i$  then  $B_{k_i, k} = U_i \in \mathcal{T}_{k_i} = \mathcal{T}_k$  and if  $k \neq k_i$  then  $B_{k_i, k} = X_k \in \mathcal{T}_k$ ]. Then we have  $S_i = \pi_{k_i}^{-1}(U_i) = \prod_{k \in I} B_{k_i, k}$  [If  $x \in \pi_{k_i}^{-1}(U_i)$  then  $\pi_{k_i}(x) \in U_i = B_{k_i, k_i}$  and  $\forall k \in I \setminus \{k_i\}$  we have  $\pi_k(x) \in X_k$  so that  $x \in \prod_{k \in I} B_{k_i, k}$ , if  $x \in \prod_{k \in I} B_{k_i, k}$  then  $\pi_{k_i}(x) \in B_{k_i, k_i} = U_i \Rightarrow x \in \pi_{k_i}^{-1}(U_i)$ ] So  $B = \bigcap_{i \in J} (\prod_{k \in I} B_{k_i, k}) \stackrel{2.100}{=} \prod_{k \in I} (\bigcap_{i \in J} B_{k_i, k}) = \prod_{k \in I} C_k$  where  $\forall k \in I$  we have that  $C_k = \bigcap_{i \in J} B_{k_i, k}$ . As  $J$  is finite and  $B_{k_i, k} \in \mathcal{T}_k$  we have that  $C_k \in \mathcal{T}_k$ . Define now  $A \subseteq I$  by  $A = \{k_i | i \in J\}$  which is finite as  $J$  is finite. If  $k \in A$  we have  $C_k \in \mathcal{T}_k$  and if  $k \in I \setminus A$  then  $\forall i \in J$  we have  $k_i \neq k \Rightarrow B_{k_i, k} = X_k \Rightarrow C_k = X_k$ . This proves that  $B \in \mathcal{B}$ .
- ( $B \in \mathcal{B}$ ) then there exists a family  $\{U_i\}_{i \in I}$  and a finite  $A \subseteq I$  such that  $\forall i \in A$  we have  $U_i \in \mathcal{T}_i$  and  $\forall i \in I \setminus A$  we have  $U_i = X_i$  with  $B = \prod_{i \in I} U_i$ . Define now  $\{U_{i,j}\}_{i \in A \times I}$  by  $A_{i,j} = \begin{cases} U_i & \text{if } i = j \\ X_j & \text{if } i \neq j \end{cases}$  then  $U_i = \bigcap_{j \in A} A_{j,i}$  [if  $x \in U_i \subseteq X_i$  then  $\forall j \in A$  we have either  $i = j \Rightarrow x \in U_i = A_{i,i}$  or if  $i \neq j$  then  $x \in X_i = A_{j,i}$  so that  $x \in \bigcap_{j \in A} A_{j,i}$ , if  $x \in \bigcap_{j \in A} A_{j,i}$  then if  $i \in I \setminus A$  we have  $\forall j \in A$  that  $j \neq i$  so that  $A_{j,i} = X_i = U_i$  and  $x \in U_i$  and if  $i \in A$  we have  $x \in A_{i,i} = U_i$ ]. So  $B = \prod_{i \in I} (\bigcap_{j \in A} A_{j,i}) \stackrel{2.100}{=} \bigcap_{j \in A} (\prod_{i \in I} A_{j,i})$ . Now  $\forall j \in A$  we have  $\prod_{i \in I} A_{j,i} = \pi_j^{-1}(U_i)$  [if  $x \in \prod_{i \in I} A_{j,i}$  then  $\pi_j(x) \in A_{j,j} = U_j \Rightarrow x \in \pi_j^{-1}(U_j)$ , if  $x \in \pi_j^{-1}(U_j)$  then  $x \in X$  such that  $\pi_j(x) \in U_j = A_{j,j}$  and if  $i \neq j$  we have  $\pi_i(x) \in X_i = A_{j,i}$  so that  $x \in \prod_{i \in I} A_{j,i}$ ] so that  $B = \bigcap_{j \in A} \pi_j^{-1}(U_j)$  proving that  $B \in \mathcal{B}'$   $\square$

**Corollary 11.38.** Let  $I$  be a non empty set,  $\{\langle X_i, \mathcal{T}_i \rangle\}_{i \in I}$  be a family of topological spaces then the product topology  $\mathcal{T}_{\text{product}}$  on  $X = \prod_{i \in I} X_i$  has as its basis  $\mathcal{B} = \{\bigcap_{i \in A} \pi_i^{-1}(U_i) | A \subseteq I \wedge A \text{ is finite and } \forall i \in A \text{ we have that } U_i \text{ is open in } X_i\}$

**Proof.** Take  $\mathcal{B}$  the basis of  $X$  defined in the previous theorem (see 11.37) then  $B \in \mathcal{B}$  if and only if  $\exists A \subseteq I$ ,  $A$  finite and  $B = \prod_{i \in I} U_i$  where  $U_i = X_i$  if  $i \in I \setminus A$  and  $U_i$  is open in  $X_i$  if  $i \in A$ . We prove now that  $B = \bigcap_{i \in A} \pi_i^{-1}(U_i)$

- If  $x \in \prod_{i \in I} U_i$  then if  $i \in A$  we have that  $\pi_i(x) \in U_i$  or  $x \in \pi_i^{-1}(U_i) \Rightarrow x \in \bigcap_{i \in A} \pi_i^{-1}(U_i)$
- If  $x \in \bigcap_{i \in A} \pi_i^{-1}(U_i)$  then  $\forall i \in A$  we have  $x \in \pi_i^{-1}(U_i) \Rightarrow x_i = \pi_i(x) \in U_i$ . If  $i \in I \setminus A$  then we have trivially that  $\pi_i(x) \in X_i = U_i$  so that  $x \in \prod_{i \in I} U_i$   $\square$

Using the above theorem it is now easy to prove that in the finite case the box topology is equivalent with the product topology.

**Theorem 11.39.** Let  $\{\langle X_i, \mathcal{T}_i \rangle_{i \in I}\}$  be a family of topological spaces then  $\mathcal{T}_{\text{box}}$  on  $X = \prod_{i \in I} X_i$  is finer than  $\mathcal{T}_{\text{product}}$  on  $X$ . If  $I$  is finite then  $\mathcal{T}_{\text{box}} = \mathcal{T}_{\text{product}}$

**Proof.** Let  $\mathcal{B}_{\text{product}} \stackrel{11.37}{=} \{\prod_{i \in I} U_i \mid \text{where } \{U_i\}_{i \in I} \text{ is a family such that } \exists A \text{ finite } \subseteq I \text{ such that } \forall i \in I \setminus A \text{ we have } U_i = X_i \text{ and } \forall i \in A \text{ we have } U_i \in \mathcal{T}_i\}$  be the basis for the product topology and let  $\mathcal{B}_{\text{box}} = \{\prod_{i \in I} U_i \mid \forall i \in I \text{ we have } U_i \in \mathcal{T}_i\}$  be the basis of the box topology. Then if  $B \in \mathcal{B}_{\text{product}}$  there exists a family  $\{U_i\}_{i \in I}$  and a finite  $A \subseteq I$  such that  $\forall i \in A$  we have  $U_i \in \mathcal{T}_i$ ,  $\forall i \in I \setminus A$  we have  $U_i = X_i \in \mathcal{T}_i$  and  $B = \prod_{i \in I} U_i$  so that  $B \in \mathcal{B}_{\text{box}}$ , this proves that  $\mathcal{B}_{\text{product}} \subseteq \mathcal{B}_{\text{box}}$  and thus by 11.32 we have that  $\mathcal{T}_{\text{box}}$  is finer than  $\mathcal{T}_{\text{product}}$  or  $\mathcal{T}_{\text{product}} \subseteq \mathcal{T}_{\text{box}}$ . If  $I$  is finite then if  $B \in \mathcal{B}_{\text{box}}$  then there exists a family  $\{U_i\}_{i \in I}$  such that  $\forall i \in I$  finite we have  $U_i \in \mathcal{T}_i$  and  $\forall i \in I \setminus I = \emptyset$  we have  $U_i = X_i$  vacuously with  $B = U_i$  proving that  $B \in \mathcal{B}_{\text{product}}$ . This proves that  $\mathcal{B}_{\text{product}} = \mathcal{B}_{\text{box}}$  and thus that  $\mathcal{T}_{\text{product}} = \mathcal{T}_{\text{box}}$ .  $\square$

**Theorem 11.40.** Let  $\{\langle X_i, \mathcal{T}_i \rangle_{i \in I}\}$  be a family of topological spaces and let  $\forall i \in I A_i \subseteq X_i$  be equipped with the subspace topology  $\mathcal{T}_{A_i}$  then the subspace topology of  $\prod_{i \in I} A_i \subseteq \prod_{i \in I} X_i$  is the same as the product topology of  $\{\langle A_i, \mathcal{T}_{A_i} \rangle_{i \in I}\}$

**Proof.** Let  $\mathcal{B}_1$  be the base of the product topology on  $\prod_{i \in I} X_i$  then the subspace topology on  $\prod_{i \in I} A_i$  is defined by  $\mathcal{B}_1 = \{B \cap (\prod_{i \in I} A_i) \mid B \in \mathcal{B}_1\}$  and let  $\mathcal{B}_2$  be a basis for the product topology of  $\prod_{i \in I} A_i$ .

1. If  $U \in \mathcal{B}_1 \stackrel{11.38}{=} U = (\prod_{i \in I} U_i) \cap (\prod_{j \in I} A_j)$  where  $U_i = X_i$  except for a finite subset of  $I$  where  $U_i \in \mathcal{T}_i \Rightarrow U = \prod_{i \in I} (U_i \cap A_i)$  where  $(U_i \cap A_i) = A_i$  (if  $U_i = X_i$ ) except for a finite subset of  $I$  where  $(U_i \cap A_i) \in \mathcal{T}_{A_i}$  and thus  $U \in \mathcal{B}_2$
2. If  $U \in \mathcal{B}_2$  then  $B = \prod_{i \in I} U_i$  where  $U_i = A_i = A_i \cap A_i$  except for a finite subset of  $I$  where  $U_i = W_i \cap A_i$  ( $W_i \in \mathcal{T}_i$ ) and thus if  $V_i = X_i$  except for a finite subset where  $V_i = W_i$  then  $B = \prod_{i \in I} (V_i \cap A_i) = (\prod_{i \in I} V_i) \cap (\prod_{i \in I} A_i)$  where  $V_i = A_i$  except for a finite subset of  $I$  where  $V_i \in \mathcal{T}_i \Rightarrow B \in \mathcal{B}_1$

now as we just have proved that  $\mathcal{B}_1 = \mathcal{B}_2$  we use 11.32 to prove that both topologies are the same.  $\square$

### 11.1.3 Dense sets

**Definition 11.41.** Let  $\langle X, \mathcal{T} \rangle$  be a topological space then  $A \subseteq X$  is a dense subset if  $\bar{A} = X$

**Definition 11.42.** A topological space  $\langle X, \mathcal{T} \rangle$  is a Baire space if for every family  $\{A_i\}_{i \in \mathbb{N}_0}$  with  $\forall i \in \mathbb{N}_0$  we have that  $A_i$  is closed and  $A_i^\circ = \emptyset$ , we have that  $(\bigcup_{i \in \mathbb{N}_0} A_i)^\circ = \emptyset$ . In other words the union of any family of closed sets with empty interior has also empty interior.

**Theorem 11.43.** Let  $\langle X, \mathcal{T} \rangle$  be a topological space,  $A \subseteq X$  then  $A^\circ = \emptyset$  if and only if  $X \setminus A$  is dense in  $A$

**Proof.**

1.  $(\Rightarrow)$  If  $A^\circ = \emptyset$  then if  $x \in X$  and  $U$  open with  $x \in U$  we have  $U \not\subseteq A$  (otherwise  $x \in A^\circ \Rightarrow U \cap (X \setminus A) \neq \emptyset \Rightarrow \bar{X \setminus A} = X$  proving that  $X \setminus A$  is dense in  $X$ ).
2.  $(\Leftarrow)$  Assume  $X \setminus A$  is dense in  $A$  and that  $x \in A^\circ$  which as  $A^\circ$  is open means by the density of  $X \setminus A$  in  $X$  means that  $A^\circ \cap (X \setminus A) \neq \emptyset$  so there exists a  $y \notin A$  and  $y \in A^\circ \subseteq A$  which is a contradiction. This means that  $A^\circ = \emptyset$ .  $\square$

**Theorem 11.44.** Let  $\langle X, \mathcal{T} \rangle$  be a topological space then  $\langle X, \mathcal{T} \rangle$  is a Baire space if and only if for every  $\{U_i\}_{i \in \mathbb{N}_0}$  with  $\forall i \in \mathbb{N}_0$  we have that  $U_i$  is open and dense in  $X$  then  $\bigcap_{i \in \mathbb{N}_0} U_i$  is dense in  $X$ . In other words the intersection of every sequence of open dense sets in  $X$  is dense in  $X$ .

### Proof.

1.  $(\Rightarrow)$  Let  $\{U_i\}_{i \in \mathbb{N}_0}$  be a family of open dense sets in  $X$  define then  $\{A_i\}_{i \in \mathbb{N}_0}$  where  $\forall i \in \mathbb{N}_0$  we have  $A_i = X \setminus U_i$  a closed set. As  $U_i = X \setminus (X \setminus U_i) = X \setminus A_i$  is by assumption dense in  $X$  we have by the previous theorem that  $\forall i \in \mathbb{N}_0$  we have  $A_i^\circ = \emptyset$ . By definition of a Baire space we have then that  $(\bigcup_{i \in \mathbb{N}_0} A_i)^\circ = \emptyset$  and thus by the previous theorem we have  $X \setminus (\bigcup_{i \in \mathbb{N}_0} A_i)$  is dense in  $X$ . Also  $X \setminus (\bigcup_{i \in \mathbb{N}_0} A_i) = X \cap (\bigcup_{i \in \mathbb{N}_0} A_i)^c = X \cap (\bigcap_{i \in \mathbb{N}_0} A_i^c) = \bigcap_{i \in \mathbb{N}_0} X \cap A_i^c = \bigcap_{i \in \mathbb{N}_0} (X \setminus A_i) = \bigcap_{i \in \mathbb{N}_0} U_i$  so that  $\bigcap_{i \in \mathbb{N}_0} U_i$  is dense in  $X$ .
2.  $(\Leftarrow)$  Let  $\{A_i\}_{i \in \mathbb{N}_0}$  a family of closed sets with  $A_i^\circ = \emptyset$  for all  $i \in \mathbb{N}_0$ , define then  $\{U_i\}_{i \in \mathbb{N}_0}$  by  $\forall i \in \mathbb{N}_0 U_i = X \setminus A_i$ . As  $A_i$  is closed we have that  $U_i$  is open and by the previous theorem we have that  $U_i$  is dense in  $X$ . By the assumption we have then that  $\bigcap_{i \in \mathbb{N}_0} U_i$  is dense in  $X$ . Using the previous theorem we have then that  $\emptyset = (X \setminus \bigcap_{i \in \mathbb{N}_0} U_i)^\circ = (X \cap (\bigcap_{i \in \mathbb{N}_0} U_i)^c)^\circ = (X \cap (\bigcup_{i \in \mathbb{N}_0} U_i^c))^\circ = (\bigcup_{i \in \mathbb{N}_0} (X \cap U_i^c))^\circ = (\bigcup_{i \in \mathbb{N}_0} (X \cap (X \cap A_i^c)^c))^\circ = (\bigcup_{i \in \mathbb{N}_0} (X \cap (X^c \cup A_i)))^\circ = (\bigcup_{i \in \mathbb{N}_0} (X \cap A_i))^\circ \underset{A_i \subseteq X}{=} (\bigcup_{i \in \mathbb{N}_0} A_i)^\circ$  proving that  $\langle X, \mathcal{T} \rangle$  is Baire.  $\square$

## 11.2 Metric Spaces

**Definition 11.45.** A pseudo metric space  $\langle X, d \rangle$  is a non empty set together with a function  $d: X \times X \rightarrow \mathbb{R}$  [d is called the pseudo metric] such that  $\forall x, y, z \in X$  we have

1.  $d(x, x) = 0$
2.  $d(x, y) = d(y, x)$
3.  $d(x, z) \leq d(x, y) + d(y, z)$

If also  $d(x, y) = 0 \Rightarrow x = y$  then we call d a metric and  $\langle X, d \rangle$  a metric space

**Theorem 11.46.** If  $\langle X, d \rangle$  is a pseudo metric space then  $\forall x, y \in X$  we have  $d(x, y) \geq 0$  (distance are positive).

**Proof.** If  $x, y \in X$  then  $0 = d(x, x) \leq d(x, y) + d(y, x) = d(x, y) + d(x, y) = 2 \cdot d(x, y) \Rightarrow 0 \leq d(x, y)$   $\square$

**Definition 11.47.** Let  $\langle X, d_X \rangle$  and  $\langle Y, d_Y \rangle$  be pseudo metric spaces and  $f: X \rightarrow Y$  a **partial function** (should not be a function but can be a function if  $\text{dom}(f) = X$ ). Given a  $x \in X$  and a  $y \in Y$  we say that  $\lim_{h \rightarrow x} f(h) = y$  if and only if  $\forall \varepsilon > 0$  ( $\varepsilon \in \mathbb{R}$ ) there exists a  $\delta > 0$  ( $\delta \in \mathbb{R}$ ) such that  $\forall x' \in \text{dom}(f)$  with  $0 < d_X(x, x') < \varepsilon$  we have that  $d_Y(y, f(x')) < \varepsilon$

**Note 11.48.** Note that neither  $x$  should be in  $\text{dom}(f)$  or  $y$  should be in  $\text{range}(f)$

**Definition 11.49.** Given a pseudo metric space  $\langle X, d \rangle$  and  $\varepsilon \in \mathbb{R}_+$  then  $B_d(x, \varepsilon) = \{y \in X \mid d(x, y) < \varepsilon\}$  is called the open ball centered around  $x$  with radius  $\varepsilon$ .

**Definition 11.50.** Given a pseudo metric space  $\langle X, d \rangle$  and  $\varepsilon \in \mathbb{R}_+$  then  $\bar{B}_d(x, \varepsilon) = \{y \in X \mid d(x, y) \leq \varepsilon\}$  is called the closed ball centered around  $x$  with radius  $\varepsilon$ .

**Remark 11.51.** Until further notice we adopt the usual convention that  $\varepsilon > 0$  means the same as  $\varepsilon \in \mathbb{R}_+$

**Theorem 11.52.** Given a pseudo metric space  $\langle X, d \rangle$  then if  $x \in B_d(x_1, \varepsilon_1) \cap B_d(x_2, \varepsilon_2)$  then  $\exists \varepsilon > 0$  such that  $x \in B_d(x, \varepsilon) \subseteq B_d(x_1, \varepsilon_1) \cap B_d(x_2, \varepsilon_2)$

**Proof.** Let  $x \in B_d(x_1, \varepsilon_1) \cap B_d(x_2, \varepsilon_2) \Rightarrow d(x_1, x) < \varepsilon_1 \wedge d(x_2, x) < \varepsilon_2$ , take then  $\varepsilon = \min(\varepsilon_1 - d(x_1, x), \varepsilon_2 - d(x_2, x)) > 0$  then if  $y \in B_d(x, \varepsilon)$  we have  $d(x, y) < \varepsilon$  so for  $i = 1, 2$  we have  $d(x_i, y) \leq d(x_i, x) + d(x, y) < d(x_i, x) + \varepsilon \leq d(x_i, x) + (\varepsilon_i - d(x_i, x)) = \varepsilon_i \Rightarrow y \in B_d(x_i, \varepsilon_i)$   $\square$

As  $y \in B_d(x, \varepsilon)$  means  $y \in B_d(x, \varepsilon) \cap B_d(x, \varepsilon)$  we can use the previous theorem to prove the following corollary.

**Corollary 11.53.** Given a pseudo metric space  $\langle X, d \rangle$  and  $y \in B_d(x, \varepsilon)$  then  $\exists \delta > 0$  such that  $B_d(y, \delta) \subseteq B_d(x, \varepsilon)$

**Theorem 11.54.** Given a pseudo metric space  $\langle X, d \rangle$  we have that  $\mathcal{B} = \{B_d(x, \varepsilon) | \varepsilon \in \mathbb{R}_+ \wedge x \in X\}$  satisfies the requirements for a generating basis (see 11.28), the generated topology on  $X$  is called the topology generated by the pseudo metric and is noted by  $\mathcal{T}_d$ .

**Proof.** To see if  $\mathcal{B}$  satisfies the requirements of 11.28 note that:

1.  $\forall x \in X$  we have  $x \in B_d(x, 1) \in \mathcal{B}$
2.  $\forall B_1, B_2 \in \mathcal{B}$  with  $x \in B_1 \cap B_2$  we have  $B_1 = B_d(x_1, \varepsilon_1)$ ,  $B_2 = B_d(x_2, \varepsilon_2)$  then as  $x \in B_d(x_1, \varepsilon_1) \cap B_d(x_2, \varepsilon_2)$  we have by 11.52 the existence of a  $\delta$  such that  $x \in B_d(x, \delta) \subseteq B_d(x_1, \varepsilon_1) \cap B_d(x_2, \varepsilon_2)$  so we found a  $B_3 = B_d(x, \delta) \in \mathcal{B}$  such that  $x \in B_3 \subseteq B_1 \cap B_2$   $\square$

**Corollary 11.55.** Given a pseudo metric space  $\langle X, d \rangle$  then  $U$  is a open set in the metric topology if and only if  $\forall x \in U$  there exists a  $\delta > 0$  such that  $x \in B_d(x, \delta)$

**Proof.**

1.  $(\Rightarrow)$  If  $x \in U$  open in the metric topology then by the definition of a basis there exists a  $y \in X$  and a  $\varepsilon > 0$  such that  $x \in B_d(y, \varepsilon) \subseteq U$ . Using 11.53 there exists a  $\delta > 0$  such that  $x \in B_d(x, \delta) \subseteq B_d(y, \varepsilon) \subseteq U$
2.  $(\Leftarrow)$  If for every  $x \in U$  there exists a  $\delta_x > 0$  such that  $x \in B_d(x, \delta_x) \subseteq U \Rightarrow U = \bigcup_{x \in U} B_d(x, \delta_x)$  which is open as  $B_d(x, \delta_x)$  is open by the definition of a basis.  $\square$

**Theorem 11.56.** Let  $\langle X, d \rangle$  be a pseudo metric space then closed balls are indeed closed

**Proof.** Let  $\bar{B}_d(x, \varepsilon)$  be a closed ball and let  $y \in X \setminus \bar{B}_d(x, \varepsilon) \Rightarrow d(x, y) > \varepsilon$  take then  $\delta = d(x, y) - \varepsilon > 0$ , now let  $z \in B_d(y, \delta)$  then assume that  $d(z, x) \leq \varepsilon \Rightarrow d(x, y) \leq d(x, z) + d(z, y) \leq \varepsilon + d(z, y) < \varepsilon + \delta = \varepsilon + d(x, y) - \varepsilon = d(x, y) \Rightarrow d(x, y) < d(x, y)$  a contradiction, so  $d(z, x) > \varepsilon \Rightarrow z \in X \setminus \bar{B}_d(x, \varepsilon) \Rightarrow B_d(y, \delta) \subseteq X \setminus \bar{B}_d(x, \varepsilon)$   $\square$

**Theorem 11.57.** Let  $\langle X, d \rangle$  be a (pseudo) metric space and  $A \subseteq X$  then the subspace topology on  $A$  is generated by the (pseudo) metric  $d|_{A \times A} : A \times A \rightarrow \mathbb{R}$   $|d|_{A \times A}$  is the restriction of  $d$  to  $A \times A$

**Proof.**

First we prove that  $d|_{A \times A}$  is a pseudo metric so let  $x, y, z \in A$  then we have

1.  $d|_{A \times A}(x, x) = d(x, x) = 0$
2.  $d|_{A \times A}(x, y) = d(x, y) = d(y, x) = d|_{A \times A}(y, x)$
3.  $d|_{A \times A}(x, y) = d(x, y) \leq d(x, z) + d(y, z) = d|_{A \times A}(x, z) + d|_{A \times A}(z, y)$
4. If  $d$  is a metric then we have if  $d|_{A \times A}(x, y) = 0 \Rightarrow d(x, y) = 0$   $\xrightarrow{\text{dis a metric}} x = y$

Second  $B_{d|_{A \times A}}(x, \delta) = \{y \in A | d|_{A \times A}(x, y) < \delta\} = \{y \in A | d(x, y) < \delta\} = \{y \in X | d(x, y) < \delta\} \cap A = B_d(x, \delta) \cap A$ . Let now  $\mathcal{T}$  the topology generated by  $d$  and  $\mathcal{T}_A$  the subspace topology. Then if  $V \in \mathcal{T}_A$  there exists a  $U \in \mathcal{T}$  such that  $V = U \cap A$ , if now  $x \in V$  then  $x \in U$  and by 11.55 there exists a  $\delta_x > 0$  such that  $x \in B_d(x, \delta_x) \subseteq U$  and thus  $x \in B_{d|_{A \times A}}(x, \delta_x) = B_d(x, \delta_x) \cap A \subseteq U \cap A = V$  so we have  $x \in B_{d|_{A \times A}}(x, \delta_x) \subseteq V$  and thus by 11.55 we conclude that  $V$  is open in the topology generated by  $d|_{A \times A}$ . Assume now that  $V$  is open in the topology generated by  $d|_{A \times A}$  on  $A$  then  $V \subseteq A$  and  $\forall x \in V$  there exists a  $\delta_x > 0$  such that  $x \in B_{d|_{A \times A}}(x, \delta_x) \subseteq V$  and thus  $x \in B_{d|_{A \times A}}(x, \delta_x) = B_d(x, \delta_x) \cap A \subseteq V$  so that  $V = \bigcup_{x \in V} B_{d|_{A \times A}}(x, \delta_x)$ . Take now the open set  $U = \bigcup_{x \in V} B_d(x, \delta_x)$  in  $\mathcal{T}$  then  $U \cap A = (\bigcup_{x \in V} B_d(x, \delta_x)) \cap A = \bigcup_{x \in V} (B_d(x, \delta_x) \cap A) = V$  proving that  $V$  is open in the subspace topology.  $\square$

**Definition 11.58.** Two pseudo metrics  $d_1, d_2$  on a set  $X$  are equivalent iff they generate the same topology.

**Theorem 11.59.** Let  $d_1, d_2$  be two metrics on a set  $X$  and let  $\mathcal{T}_1, \mathcal{T}_2$  the generated metric topologies then  $\mathcal{T}_2$  is finer then  $\mathcal{T}_1$  if and only if  $\forall x \in X, \forall \varepsilon > 0$  there  $\exists \delta > 0$  such that  $x \in B_{d_2}(x, \delta) \subseteq B_{d_1}(x, \varepsilon)$

**Proof.**

1.  $(\Rightarrow)$  If  $x \in X$  and  $\varepsilon > 0$  then by 11.32 for  $x \in B_{d_1}(x, \varepsilon)$  there exists a  $y \in X, \delta' > 0$  such that  $x \in B_{d_2}(y, \delta') \subseteq B_{d_1}(x, \varepsilon)$ , using 11.53 there exists a  $\delta > 0$  such that  $x \in B_{d_2}(x, \delta) \subseteq B_{d_2}(y, \delta') \subseteq B_{d_1}(x, \varepsilon)$
2.  $(\Leftarrow)$  If  $y \in B_{d_1}(x, \varepsilon)$  then using 11.53 there exists a  $\delta' > 0$  such that  $y \in B_{d_1}(y, \delta') \subseteq B_{d_1}(x, \varepsilon)$  and by the hypothesis there exists a  $\delta > 0$  such that  $y \in B_{d_2}(y, \delta) \subseteq B_{d_1}(y, \delta') \subseteq B_{d_1}(x, \varepsilon)$  proving by 11.32 that  $\mathcal{T}_2$  is finer then  $\mathcal{T}_1$   $\square$

**Definition 11.60.** Let  $\langle X, d_X \rangle$  and  $\langle Y, d_Y \rangle$  be two pseudo metric spaces then a function  $\varphi: X \rightarrow Y$  is called a isometry iff

1.  $\varphi$  is a bijective map
2.  $\forall x, y \in X$  we have  $d_X(x, y) = d_Y(\varphi(x), \varphi(y))$

**Theorem 11.61.** Let  $\langle X, d_X \rangle, \langle Y, d_Y \rangle$  be two pseudo metric spaces and  $\varphi: X \rightarrow Y$  a isometry then  $\varphi^{-1}$  is a isometry.

**Proof.** Because 2.41 we have that  $\varphi^{-1}$  is a bijection and if  $x, y \in Y$  then  $d_Y(x, y) = d_Y(\varphi(\varphi^{-1}(x)), \varphi(\varphi^{-1}(y)))$   $\underset{\text{disisometry}}{=} d_X(\varphi^{-1}(x), \varphi^{-1}(y))$   $\square$

**Theorem 11.62.** Let  $\langle X, d_X \rangle, \langle Y, d_Y \rangle, \langle Z, d_Z \rangle$  be metric spaces and let  $\varphi: X \rightarrow Y$  and  $\psi: Y \rightarrow Z$  be isometries then  $\psi \circ \varphi: X \rightarrow Z$  is a isometry

**Proof.** Using 2.49 we have that  $\psi \circ \varphi: X \rightarrow Z$  is a bijection and  $d_X(x, y) = d_Y(\varphi(x), \varphi(y)) = d_Z(\psi(\varphi(x)), \psi(\varphi(y))) = d_Z((\psi \circ \varphi)(x), (\psi \circ \varphi)(y))$   $\square$

**Theorem 11.63.** Let  $\langle X, d_X \rangle, \langle Y, d_Y \rangle$  be two metric spaces with a isometry  $\varphi: X \rightarrow Y$  between them then  $\varphi(B_{d_X}(x, \delta)) = B_{d_Y}(\varphi(x), \delta)$  and  $\varphi^{-1}(B_{d_Y}(y, \delta)) = B_{d_X}(\varphi^{-1}(y), \delta)$

**Proof.**

1. If  $y \in \varphi(B_{d_X}(x, \delta)) \Rightarrow y = \varphi(z) \wedge z \in B_{d_X}(x, \delta) \Rightarrow y = \varphi(z) \wedge d_X(x, z) < \delta \Rightarrow y = \varphi(z) \wedge d_Y(\varphi(x), \varphi(z)) < \delta \Rightarrow d_Y(\varphi(x), y) < \delta \Rightarrow y \in B_{d_Y}(\varphi(x), \delta)$ . If  $y \in B_{d_Y}(\varphi(x), \delta) \Rightarrow d_Y(\varphi(x), y) < \delta \Rightarrow d_Y(\varphi(x), \varphi(\varphi^{-1}(y))) < \delta \Rightarrow d_X(x, \varphi^{-1}(y)) < \delta \Rightarrow \varphi^{-1}(y) \in B_{d_X}(x, \delta) \Rightarrow y \in \varphi(B_{d_X}(x, \delta))$ . This proves that  $\varphi(B_{d_X}(x, \delta)) = B_{d_Y}(\varphi(x), \delta)$ .
2. Let  $x \in \varphi^{-1}(B_{d_Y}(y, \delta)) \Rightarrow \varphi(x) \in B_{d_Y}(y, \delta) \Rightarrow d_Y(y, \varphi(x)) < \delta \Rightarrow d_Y(\varphi(\varphi^{-1}(y)), \varphi(x)) \Rightarrow d_X(\varphi^{-1}(y), x) < \delta \Rightarrow x \in B_{d_X}(\varphi^{-1}(y), \delta)$ . If  $x \in B_{d_X}(\varphi^{-1}(y), \delta) \Rightarrow d_X(\varphi^{-1}(y), x) < \delta \Rightarrow d_Y(\varphi(\varphi^{-1}(y)), \varphi(x)) < \delta \Rightarrow \varphi(x) \in B_{d_Y}(y, \delta) \Rightarrow x \in \varphi^{-1}(B_{d_Y}(y, \delta))$ . This proves that  $\varphi^{-1}(B_{d_Y}(y, \delta)) = B_{d_X}(\varphi^{-1}(y), \delta)$   $\square$

**Theorem 11.64.** Let  $\langle X, d_X \rangle, \langle Y, d_Y \rangle$  be two metric spaces with a isometry  $\varphi: X \rightarrow Y$  between them . If  $\mathcal{T}_X, \mathcal{T}_Y$  be the two metric topologies on  $X, Y$  then  $\mathcal{T}_X = \{\varphi^{-1}(V) | V \in \mathcal{T}_Y\} = \{U \subseteq X | \varphi(U) \in \mathcal{T}_Y\}$  and  $\mathcal{T}_Y = \{\varphi(U) | U \in \mathcal{T}_X\} = \{V \subseteq Y | \varphi^{-1}(V) \in \mathcal{T}_X\}$ .

**Proof.** First we prove that

$$\{\varphi^{-1}(V) | V \in \mathcal{T}_Y\} = \{U \subseteq X | \varphi(U) \in \mathcal{T}_Y\}, \quad (11.3)$$

so if  $U \in \{\varphi^{-1}(V) | V \in \mathcal{T}_Y\}$  there exists a  $V \in \mathcal{T}_Y$  such that  $U = \varphi^{-1}(V) \subseteq X$  and as  $\varphi$  is a bijection we have  $\varphi(U) = V \in \mathcal{T}_Y$  or  $U \in \{U \subseteq X | \varphi(U) \in \mathcal{T}_Y\}$  also if  $U \in \{U \subseteq X | \varphi(U) \in \mathcal{T}_Y\}$  then  $U \subseteq X \wedge \varphi(U) \in \mathcal{T}_Y$  and as  $\varphi$  is a bijection we have  $U = \varphi^{-1}(\varphi(U))$  where  $\varphi(U) \in \mathcal{T}_Y$  so that  $U \in \{\varphi^{-1}(V) | V \in \mathcal{T}_Y\}$ .

Second we prove that  $\mathcal{T}_X = \{\varphi^{-1}(V) | V \in \mathcal{T}_Y\}$ . So if  $U \in \mathcal{T}_X$  take then  $V = \varphi(U)$  and  $y \in V$  so that  $\exists x \in U$  such that  $y = \varphi(x)$ . Using 11.55 there exists a  $\delta > 0$  such that  $x \in B_{d_X}(x, \delta) \subseteq U$  and thus  $y = \varphi(x) \in \varphi(B_{d_X}(x, \delta)) \subseteq \varphi(U) = V \xrightarrow{11.63} y \in B_{d_Y}(\varphi(x), \delta) \subseteq V \xrightarrow{y = \varphi(x)} y \in B_{d_Y}(y, \delta) \subseteq V$  which by 11.55 means that  $V$  is open in  $\mathcal{T}_Y$  and as  $\varphi$  is bijective we have  $U = \varphi^{-1}(V)$  where  $V \in \mathcal{T}_Y$  such that  $U \in \{\varphi^{-1}(V) | V \in \mathcal{T}_Y\}$  or

$$\mathcal{T}_X \subseteq \{\varphi^{-1}(V) | V \in \mathcal{T}_Y\} \quad (11.4)$$

Next if  $U \in \{\varphi^{-1}(V) | V \in \mathcal{T}_Y\}$  then there exists a  $V \in \mathcal{T}_Y$  such that  $U = \varphi^{-1}(V)$  or as  $\varphi$  is a bijection we have  $\varphi(U) = V$ . If now  $x \in U$  then  $\varphi(x) \in \varphi(U) = V \in \mathcal{T}_Y$  and using 11.55 there exists a  $\delta > 0$  such that  $\varphi(x) \in B_{d_Y}(\varphi(x), \delta) \subseteq V = \varphi(U)$  so that  $x \in \varphi^{-1}(B_{d_Y}(\varphi(x), \delta)) \subseteq \varphi^{-1}(\varphi(U)) \xrightarrow{\varphi \text{ is a bijection}} U \xrightarrow{11.63} x \in B_{d_X}(\varphi^{-1}(\varphi(x)), \delta) \subseteq U \Rightarrow x \in B_{d_X}(x, \delta) \subseteq U$  proving by 11.55 that  $U \in \mathcal{T}_X$  so that we have

$$\{\varphi^{-1}(V) | V \in \mathcal{T}_Y\} \subseteq \mathcal{T}_X \quad (11.5)$$

Using 11.4 and 11.5 we have then

$$\mathcal{T}_X = \{\varphi^{-1}(V) | V \in \mathcal{T}_Y\} \quad (11.6)$$

As for the rest note that  $\Phi = \varphi^{-1}$  is a isometry from  $Y \rightarrow X$  (see 11.61) so by 11.6 and 11.3 we have that  $\mathcal{T}_Y = \{\Phi^{-1}(U) | U \in \mathcal{T}_X\} = \{V \subseteq Y | \Phi(V) \in \mathcal{T}_X\} \xrightarrow{\Phi = \varphi^{-1} \Rightarrow \Phi^{-1} = \varphi} \{\varphi(U) | U \in \mathcal{T}_X\} = \{V \subseteq Y | \varphi^{-1}(V) \in \mathcal{T}_X\}$   $\square$

**Definition 11.65.** Let  $\langle X, d \rangle$  be a pseudo metric space and  $A \subseteq X$  a subset, then  $A$  is called **bounded** iff  $\exists M \in \mathbb{R}$  such that  $\forall x, y \in X$  we have  $d(x, y) \leq M$ . If  $A$  is bounded then  $\sup(\{d(x, y) | x, y \in A\})$  exists ( $\mathbb{R}, \leq$  is conditional complete (see 9.43)) and is called the **diameter** of  $A$  noted as  $\text{diam}(A)$ .

**Example 11.66.** Let  $\langle X, d \rangle$  be a metric space then every closed and open ball is bounded

**Proof.** Let  $\overline{B_d}(a, \delta)$  be a closed ball then if  $M = 2 \cdot \delta$  we have if  $x, y \in \overline{B_d}(a, \delta)$  that  $d(x, y) \leq d(x, a) + d(a, y) \leq \delta + \delta = 2 \cdot \delta$ . As  $B_d(a, \delta) \subseteq \overline{B_d}(a, \delta)$  we have also that  $\overline{B_d}(a, \delta)$  is bounded.  $\square$

**Theorem 11.67.** Let  $\{\langle X_i, d_i \rangle\}_{i \in I}$  be a finite family of (pseudo) metric spaces then  $d: (\prod_{i \in I} X_i) \times (\prod_{i \in I} X_i) \rightarrow \mathbb{R}$  defined by  $(x, y) \rightarrow d(x, y) = \max(\{d_i(\pi_i(x), \pi_i(y)) | i \in I\})$  (see 5.50) is a (pseudo) metric space and the topology generated is the product topology  $\mathcal{T}_{\text{product}}$  and the box topology  $\mathcal{T}_{\text{box}}$ . The metric  $d$  is called the **product metric**.

**Proof.** First we have to prove that  $d$  is indeed a (pseudo) metric

1.  $d(x, x) = \max(\{d_i(\pi_i(x), \pi_i(x)) | i \in I\}) = 0$  as  $\forall i \in I$  we have  $d_i(\pi_i(x), \pi_i(x)) = 0$
2.  $d(x, y) = \max(\{d_i(\pi_i(x), \pi_i(y)) | i \in I\}) = \max(\{d_i(\pi_i(y), \pi_i(x)) | i \in I\}) = d(y, x)$
3.  $d(x, y) = \max(\{d_i(\pi_i(x), \pi_i(y)) | i \in I\}) \xrightarrow{2.168} \max(\{d_i(\pi_i(x), \pi_i(z)) + d_i(\pi_i(z), \pi_i(y)) | i \in I\}) \leq \max(\{\max(\{d_i(\pi_i(x), \pi_i(z)) | i \in I\}) + d_i(\pi_i(z), \pi_i(y)) | i \in I\}) \leq \max(\max(\{d_i(\pi_i(x), \pi_i(z)) | i \in I\}) + \max(\{d_i(\pi_i(z), \pi_i(y)) | i \in I\})) = \max(\{d_i(\pi_i(x), \pi_i(z))\}) + \max(\{d_i(\pi_i(z), \pi_i(y)) | i \in I\}) = d(x, z) + d(y, z)$
4. If  $\{\langle X_i, d_i \rangle\}_{i \in I}$  are metric spaces then if  $d(x, y) = 0$  we have  $\forall i \in I$  that  $d_i(\pi_i(x), \pi_i(y)) = 0 \Rightarrow \forall i \in I$  we have  $\pi_i(x) = \pi_i(y) \Rightarrow x = y$

Next as  $I$  is finite we have by 11.39 that  $\mathcal{T}_{\text{product}} = \mathcal{T}_{\text{box}}$ . The generating basis for the box topology is  $\mathcal{B}_{\text{box}} = \{\prod_{i \in I} U_i | U_i \text{ open in the metric topology } \mathcal{T}_i \text{ of } \langle X_i, d_i \rangle\}$ , the generating basis for the metric topology is  $\mathcal{B}_d = \{B_d(x, \delta) | x \in X_i, \delta > 0\}$ . Let  $x \in X$  and  $B \in \mathcal{B}_{\text{box}}$  with  $x \in B$  then  $B = \prod_{i \in I} U_i$  where  $U_i \in \mathcal{T}_i$ . So if  $x \in B$  then  $\forall i \in I$  we have  $\pi_i(x) \in U_i$  so that  $\exists \delta_i > 0$  such that  $\pi_i(x) \in B_{d_i}(\pi_i(x), \delta_i) \subseteq U_i$ , take now  $\delta = \min(\{\delta_i | i \in I\}) > 0$  then if  $y \in B_d(x, \delta)$  we have  $d(x, y) < \delta \Rightarrow \forall i \in I$  we have  $d_i(\pi_i(x), \pi_i(y)) < \delta \leq \delta_i \Rightarrow \forall i \in I$  we have  $\pi_i(y) \in B_{d_i}(\pi_i(x), \delta_i) \subseteq U_i \Rightarrow y \in \prod_{i \in I} U_i = B$  so that we have proved that  $x \in B_d(x, \delta) \subseteq B$  or

$$\forall x \in X, \forall B \in \mathcal{B}_{\text{box}} \text{ there } \exists B' \in \mathcal{B}_d \text{ such that } x \in B' \subseteq B \quad (11.7)$$

If now  $x \in X$  and  $B \in \mathcal{B}_d$  with  $x \in B$  then  $\exists \delta' > 0, y \in X$  such that  $x \in B_d(y, \delta') = B$  which means that  $\exists \delta > 0$  such that  $x \in B_d(x, \delta) \subseteq B_d(y) = B$ . Take now  $\prod_{i \in I} B_{d_i}(\pi_i(x), \delta) \in \mathcal{B}_{\text{box}}$  then if  $z \in \prod_{i \in I} B_{d_i}(\pi_i(x), \delta)$  then  $\forall i \in I$  we have  $\pi_i(z) \in B_{d_i}(\pi_i(x), \delta) \Rightarrow d_i(\pi_i(x), \pi_i(z)) < \delta \Rightarrow d(x, z) = \max(\{d_i(\pi_i(x), \pi_i(z)) | i \in I\}) < \delta \Rightarrow z \in B_d(x, \delta) \subseteq B$ . So we have proved that  $x \in \prod_{i \in I} B_{d_i}(\pi_i(x), \delta) \subseteq B$

$$\forall x \in X, \forall B \in \mathcal{B}_d \text{ there } \exists B' \in \mathcal{B}_{\text{box}} \text{ such that } x \in B' \subseteq B \quad (11.8)$$

Using then 11.32 we have then that the metric topology is equal to the box topology and thus also the product topology.  $\square$

### 11.3 Normed spaces

**Definition 11.68.** A pseudo normed space  $\langle X, \|\cdot\| \rangle$  is a (real or complex) vector space  $X$  together with a function  $\|\cdot\|: X \rightarrow \mathbb{R}$  such that

1.  $\forall x \in X$  we have  $\|x\| \geq 0$
2.  $\forall x \in X, \forall \alpha \in \mathbb{R}(\mathbb{C})$  we have  $\|\alpha \cdot x\| = |\alpha| \cdot \|x\|$
3.  $\forall x, y \in X$  we have  $\|x + y\| \leq \|x\| + \|y\|$

if also  $\|x\| = 0 \Rightarrow x = 0$  then  $\langle X, \|\cdot\| \rangle$  is called a normed space.

**Note 11.69.** Let  $\langle X, \|\cdot\| \rangle$  be a normed space then if  $x = 0$  we have  $\|x\| = 0$

**Proof.** If  $x = 0$  then  $x = 0 \cdot x \Rightarrow \|x\| = \|0 \cdot x\| = |0| \cdot \|x\| = 0$   $\square$

**Theorem 11.70.** Let  $\langle X, \|\cdot\| \rangle$  be a pseudo normed space then

1.  $\forall x, y$  we have

$$\|x\| - \|y\| \leq \|x + y\| \quad (11.9)$$

2. If  $I$  is a finite set,  $\{a_i\}_{i \in I}$  then

$$\left\| \sum_{i \in I} a_i \right\| \leq \sum_{i \in I} \|a_i\| \quad (11.10)$$

**Proof.**

1.  $\|x\| = \|x + y - y\| \leq \|x + y\| + \|y\| = \|x + y\| + \|y\| \Rightarrow \|x\| - \|y\| \leq \|x + y\|$  and  $\|y\| = \|x + y - x\| \leq \|x + y\| + \|x\| = \|x + y\| + \|x\| \Rightarrow \|y\| - \|x\| \leq \|x + y\|$ . From this it follows that  $\|x\| - \|y\| \leq \|x + y\|$ .

2. If  $I = \emptyset$  then we have that  $\sum_{i \in I} a_i = 0$  so that trivially  $\|\sum_{i \in I} a_i\| \leq \sum_{i \in I} \|a_i\|$ . If  $I \neq \emptyset$  we first prove that  $\|\sum_{i=1}^n a_i\| \leq \sum_{i=1}^n \|a_i\|$  by induction. So let  $S = \{n \in \mathbb{N}_0 | \text{If } \{a_i\}_{i \in \{1, \dots, n\}} \subseteq X \text{ then } \sum_{i=1}^n a_i \leq \sum_{i=1}^n \|a_i\|\}$  then we have

**n = 1.** then if  $\{a_i\}_{i \in \{1, \dots, 1\}} \subseteq X$  then  $\|\sum_{i=1}^1 a_i\| = \|a_1\| = \sum_{i=1}^1 \|a_i\|$  proving that  $1 \in S$

**n ∈ S ⇒ n + 1 ∈ S.** if  $\{a_i\}_{i \in \{1, \dots, n+1\}} \subseteq X$  then  $\|\sum_{i=1}^{n+1} a_i\| = \|a_{n+1} + \sum_{i=1}^n a_i\| \leq \|a_{n+1}\| + \|\sum_{i=1}^n a_i\| \leq n \in S \|a_{n+1}\| + \sum_{i=1}^n \|a_i\| = \sum_{i=1}^{n+1} \|a_i\|$  proving that  $n + 1 \in S$

Finally if  $I$  is a finite non empty set then there exists a  $n \in \mathbb{N}_0$  and a bijection  $b: \{1, \dots, n\} \rightarrow I$  such that  $\sum_{i \in I} a_i = \sum_{i=1}^n a_{b_i}$  so that  $\|\sum_{i \in I} a_i\| = \|\sum_{i=1}^n a_{b_i}\| \leq \sum_{i=1}^n \|a_{b_i}\| = \sum_{i \in I} \|a_i\|$   $\square$

Every (pseudo) normed space defines a (pseudo) metric space

**Theorem 11.71.** Let  $\langle X, \|\cdot\| \rangle$  be a (pseudo) normed space then  $\langle X, d_{\|\cdot\|} \rangle$  where  $d_{\|\cdot\|}(x, y) = \|x - y\|$  is a (pseudo) metric space. The generated topology is called the (pseudo) normed topology and its basis is  $\mathcal{B} = \{B_{d_{\|\cdot\|}}(x, \delta) | x \in X, \delta > 0\}$  where  $B_{d_{\|\cdot\|}}(x, \delta) = \{y | \|x - y\| < \delta\} \stackrel{\text{defined}}{=} B_{\|\cdot\|}(x, \delta)$ . We note the generated topology by  $\mathcal{T}_{\|\cdot\|}$ .

**Proof.**

1.  $d_{|||}(x, y) = \|x - y\| \geq 0$
2.  $d_{|||}(x, y) = \|x - y\| = \|(-1) \cdot (y - x)\| = |-1| \cdot \|y - x\| = \|y - x\| = d(y, x)$
3.  $d_{|||}(x, z) = \|x - z\| = \|x - y + y - z\| \leq \|x - y\| + \|y - z\| = d_{|||}(x, y) + d_{|||}(y, z)$
4. If  $|||$  is normed then if  $d(x, y) = 0 \Rightarrow \|x - y\| = 0 \Rightarrow x - y = 0 \Rightarrow x = y$  □

**Proposition 11.72.**  $\langle \mathbb{R}, || \rangle$  and  $\langle \mathbb{C}, || \rangle$  where  $|x|$  is the absolute value of  $x$  forms a normed space. In  $\langle \mathbb{R}, || \rangle$  the basis  $\mathcal{B}$  of the normed topology is given by

$$\begin{aligned}\mathcal{B} &= \{B_{||}(x, \varepsilon) \mid x \in \mathbb{R} \wedge \varepsilon \in \mathbb{R}_+\} \\ &= \{]x - \varepsilon, x + \varepsilon[ \mid x \in \mathbb{R} \wedge \varepsilon \in \mathbb{R}_+\} \\ &= \{]a, b[ \mid a, b \in \mathbb{R} \wedge a < b\}\end{aligned}$$

**Proof.** From 8.66 and 9.33 it follows that  $||$  is a norm so that we can apply 11.71. For the case of  $\langle \mathbb{R}, || \rangle$  its basis is  $\mathcal{B} = \{B_{||}(x, \varepsilon) \mid x \in \mathbb{R} \wedge \varepsilon \in \mathbb{R}_+\}$ . First if  $x \in \mathbb{R}$  and  $\varepsilon \in \mathbb{R}_+$  then

$$\begin{aligned}y \in B_{||}(x, \varepsilon) &\Leftrightarrow |x - y| < \varepsilon \\ &\Leftrightarrow x - y < \varepsilon \wedge y - x < \varepsilon \\ &\Leftrightarrow x - \varepsilon < y \wedge y < x + \varepsilon \\ &\Leftrightarrow y \in ]x - \varepsilon, x + \varepsilon[\end{aligned}$$

proving that

$$B_{||}(x, \varepsilon) = ]x - \varepsilon, x + \varepsilon[$$

So

$$\mathcal{B} = \{]x - \varepsilon, x + \varepsilon[ \mid x \in \mathbb{R} \wedge \varepsilon \in \mathbb{R}_+\}$$

Further if  $B \in \{]x - \varepsilon, x + \varepsilon[ \mid x \in \mathbb{R} \wedge \varepsilon \in \mathbb{R}_+\}$  then as  $x = \varepsilon, x + \varepsilon \in \mathbb{R}$  and  $x - \varepsilon < x + \varepsilon$  we have that  $B \in \{]a, b[ \mid a, b \in \mathbb{R} \wedge a < b\}$  proving that

$$\mathcal{B}_{||} \subseteq \mathcal{B}$$

If now  $B \in \{]a, b[ \mid a, b \in \mathbb{R} \wedge a < b\}$  then there exists  $a, b \in \mathbb{R}$  with  $a < b$  such that  $B = ]a, b[$ . Take then  $\varepsilon = \frac{b-a}{2} > 0$  and  $x = a + \varepsilon$  then we have

$$\begin{aligned}y \in ]x - \varepsilon, x + \varepsilon[ &\Leftrightarrow x - \varepsilon < y \wedge y < x + \varepsilon \\ &\Leftrightarrow a + \varepsilon - \varepsilon < y \wedge y < a + \varepsilon + \varepsilon \\ &\Leftrightarrow a < y \wedge y < a + 2 \cdot \varepsilon \\ &\Leftrightarrow a < y \wedge y < a + b - a \\ &\Leftrightarrow a < y \wedge y < b \\ &\Leftrightarrow y \in ]a, b[\end{aligned}$$

proving that  $B = ]x - \varepsilon, x + \varepsilon[ \in \mathcal{B}_{||}$  and thus that

$$\mathcal{B} \subseteq \mathcal{B}_{||}$$

□

**Lemma 11.73.** Let  $I$  be a generalized interval [see 2.185 and 2.187] then the following intervals are open

1.  $]a, b[$  where  $a, b \in \mathbb{R}$  with  $]a, b[ \neq \emptyset$  [so that  $a < b$ ]
2.  $]-\infty, a[$  where  $a \in \mathbb{R}$
3.  $]a, \infty[$  where  $a \in \mathbb{R}$
4.  $\mathbb{R} = ]-\infty, \infty[$
5.  $I = \emptyset$

Furthermore these are the only open generalized intervals.

**Proof.** First we have to prove that the other forms of generalized intervals [see 2.187] are not open.

$]-\infty, a] \wedge a \in \mathbb{R}$ . Assume that  $]-\infty, a]$  is open then as  $\mathcal{B} = \{]x, y[, x, y \in \mathbb{R}, x < y\}$  is a basis and  $a \in ]-\infty, a]$  there exists a  $]x, y[$  with  $a \in ]x, y[ \subseteq ]-\infty, a]$  so that  $a < y \leq a$  a contradiction.

$[a, b] \wedge a, b \in \mathbb{R} \wedge [a, b] \neq \emptyset$ . Assume that  $[a, b]$  is open then as  $\mathcal{B} = \{]x, y[, x, y \in \mathbb{R}, x < y\}$  is a basis and  $b \in [a, b]$  there exists a  $]x, y[$  with  $b \in ]x, y[ \subseteq [a, b]$  so that  $b < y \leq b$  a contradiction.

$[a, b] \wedge a, b \in \mathbb{R} \wedge [a, b] \neq \emptyset$ . Assume that  $[a, b]$  is open then as  $\mathcal{B} = \{]x, y[, x, y \in \mathbb{R}, x < y\}$  is a basis and  $b \in [a, b]$  there exists a  $]x, y[$  with  $b \in ]x, y[ \subseteq [a, b]$  so that  $b < y \leq b$  a contradiction.

$[a, \infty[ \wedge a \in \mathbb{R}$ . Assume that  $[a, \infty[$  is open then as  $\mathcal{B} = \{]x, y[, x, y \in \mathbb{R}, x < y\}$  is a basis and  $a \in [a, \infty[$  there exists a  $]x, y[$  such that  $a \in ]x, y[ \subseteq [a, \infty[$  so that  $a < x < a$  a contradiction.

$[a, b[ \wedge a, b \in \mathbb{R} \wedge [a, b] \neq \emptyset$ . Assume that  $[a, b[$  is open then as  $\mathcal{B} = \{]x, y[, x, y \in \mathbb{R}, x < y\}$  is a basis and  $a \in [a, b[$  there exists a  $]x, y[$  such that  $a \in ]x, y[ \subseteq [a, b[$  so that  $a < x < a$  a contradiction.

As we have in all cases a contradiction we have proved that the above sets can not be open. Next we must prove that the remaining sets are indeed open. We must prove this for the following sets

$[a, b[ \wedge a, b \in \mathbb{R} \wedge [a, b] \neq \emptyset$ . then as  $\mathcal{B} = \{]x, y[, x, y \in \mathbb{R}, x < y\}$  is a basis we have that  $]a, b[$  is a open set.

$]-\infty, a[ \wedge a \in \mathbb{R}$ . Let  $b \in ]-\infty, a[$  then  $b < a$  so that  $b - 1 < b < a \Rightarrow b \in ]b - 1, a[ \subseteq ]-\infty, a[$  from which it follows that  $]-\infty, a[$  is a open set.

$[a, \infty[ \wedge a \in \mathbb{R}$ . Let  $b \in [a, \infty[$  then  $a < b$  so that  $a < b < b + 1 \Rightarrow b \in ]a, b + 1[ \subseteq [a, \infty[$  from which it follows that  $[a, \infty[$  is a open set.

$\mathbb{R} = ]-\infty, \infty[$ . This follows from the definition of a topology

$\emptyset$ . This follows from the definition of a topology.  $\square$

**Corollary 11.74.** As a application of the above we have that

1.  $\forall a \in \mathbb{R} [a, \infty[$  is closed
2.  $\forall a \in \mathbb{R} ]-\infty, a]$  is closed
3.  $\forall a, b \in \mathbb{R}$  with  $a \leq b$   $[a, b]$  is closed

**Proof.**

1. First

$$\begin{aligned} x \in \mathbb{R} \setminus ]-\infty, a[ &\Leftrightarrow x \in \mathbb{R} \wedge \neg(x < a) \\ &\Leftrightarrow x \in \mathbb{R} \wedge a \leq x \\ &\Leftrightarrow x \in [a, \infty[ \end{aligned}$$

proving that  $[a, \infty[ = \mathbb{R} \setminus ]-\infty, a[$  which is closed as by 11.73  $]-\infty, a[$  is open.

2. First

$$\begin{aligned} x \in \mathbb{R} \setminus ]a, \infty[ &\Leftrightarrow x \in \mathbb{R} \wedge \neg(a < x) \\ &\Leftrightarrow x \in \mathbb{R} \wedge (x \leq a) \\ &\Leftrightarrow x \in ]-\infty, a] \end{aligned}$$

proving that  $]-\infty, a] = \mathbb{R} \setminus ]a, \infty[$  which is closed as by 11.73  $]a, \infty[$  is open.

3. First

$$\begin{aligned} x \in [a, \infty[ \cap ]-\infty, b] &\Leftrightarrow x \in [a, \infty[ \wedge x \in ]-\infty, b] \\ &\Leftrightarrow x \in \mathbb{R} \wedge a \leq x \wedge x \leq b \\ &\Leftrightarrow x \in [a, b] \end{aligned}$$

proving that  $[a, b] = [a, \infty[ \cap ]-\infty, b]$  which is by (1) and (2) is the intersection of two closed sets hence closed.  $\square$

**Proposition 11.75.** *Let  $\langle \mathbb{R}, || \rangle$  be the normed space of the real numbers and  $U$  a non empty open set then there exists a countable set  $\mathcal{I}$  and a disjoint family of open generalized intervals [see 2.185]  $\{I_i\}_{i \in \mathcal{I}}$  such that  $U = \bigcup_{i \in \mathbb{N}_0} ]a_i, b_i[$ .*

**Proof.** Let  $x, y \in U$  then we say that  $x \sim y$  iff  $\min(x, y), \max(x, y) \subseteq U$  which defines a relation on  $U$ . we prove now that this is a equivalence relation.

**reflexivity.** Let  $x \in U$  then  $\min(x, x), \max(x, x) = ]x, x[ = \emptyset \subseteq U$  proving that  $x \sim x$ .

**symmetry.** Let  $x, y \in U$  be such that  $x \sim y$  then  $\min(x, y), \max(x, y) \subseteq U$  so that  $\min(y, x), \max(y, x) = ]\min(x, y), \max(x, y)[ \subseteq U$  proving that  $y \sim x$ .

**transitivity.** Let  $x, y, z \in U$  be such that  $x \sim y$  and  $y \sim z$  then we may always assume that  $x \leq y \leq z$  [otherwise exchange  $x, y, z$ ] so that  $]x, y[ = \min(x, y), \max(x, y)[, ]y, z[ = \min(y, z), \max(y, z)[ \subseteq U$ . Then we have to consider the following cases:

$x = y = z$ . then  $]x, z[ = \emptyset \subseteq U$

$x = y < z$ . then  $]x, z[ = ]y, z[ \subseteq U$

$x < y = z$ . then  $]x, z[ = ]x, y[ \subseteq U$

$x < y < z$ . then as  $y \in U$  and  $]x, z[ = ]x, y[ \cup \{y\} \cup ]y, z[ \subseteq U$

proving transitivity.

Define then  $\mathcal{U} = \{\sim[x] | x \in U\}$  the set of equivalence classes defined by  $\sim$  then using 2.121 we have that

$$U = \bigcup_{I \in \mathcal{U}} I \text{ and } \forall I, J \in \mathcal{U} \text{ with } I \neq J \text{ we have } I \cap J = \emptyset \quad (11.11)$$

Let now  $I \in \mathcal{U}$  then there exists a  $x \in U$  such that  $I = \sim[x]$ . Let  $y \in I \subseteq U$  then as  $y \in U$  there exists a  $]a, b[$  with  $y \in ]a, b[ \subseteq U$ . Then  $\forall z \in ]a, b[$  we have either

$y \leq z$ . then  $a < y \leq z < b$  so that  $\min(y, z), \max(y, z) = ]y, z[ \subseteq ]a, b[ \subseteq U$

$z < y$ . then  $a < z < y < b$  so that  $\min(y, z), \max(y, z) = ]z, y[ \subseteq ]a, b[ \subseteq U$

proving that  $y \sim z$ . As further from  $y \in I = \sim[x]$  we have  $x \sim y$  it follows by transitivity that  $x \sim z$  proving that  $z \sim \sim[x]$ . Hence  $y \in ]a, b[ \subseteq \sim[x]$  which prove that

$$I = \sim[x] \text{ is open} \quad (11.12)$$

Further if  $r, s \in I = \sim[x]$  with  $r < s$  then  $x \sim r \wedge s \sim x \Rightarrow s \sim r$  hence  $]r, s[ \subseteq U$ . If  $t \in ]r, s[$  then  $]r, t[ \subseteq ]r, s[ \subseteq U$  giving  $r \sim t$  which as  $x \sim r$  gives  $x \sim t$  proving that  $t \in I = \sim[x]$  so that  $]r, s[ \subseteq \sim[x] = I$ . Using 2.188 we have that  $I$  is a generalized interval and as  $I$  is open we have by 11.73 that

$$\forall I \in \mathcal{U} \text{ we have either } I = ]x, y[ \vee I = ]-\infty, x[ \vee ]x, \infty[ \vee ]-\infty, \infty[ \quad (11.13)$$

Now if  $I \in \mathcal{U}$  then  $\exists x \in U$  with  $I = \sim[x]$  so that  $I \neq \emptyset$ . We have the following cases to consider for  $\mathbb{Q} \cap I$

$I = ]a, b[ \wedge a, b \in \mathbb{R}$ . then as  $I \neq \emptyset$  we have  $a < b$ , so using the density of  $\mathbb{Q}$  in  $\mathbb{R}$  [see 9.58] there exists a  $q \in \mathbb{Q}$  so that  $a < q < b$  hence  $q \in ]a, b[$ , which proves that  $\mathbb{Q} \cap I \neq \emptyset$ .

$I = ]-\infty, a[ \wedge a \in \mathbb{R}$ . then  $a - 1 < a$  and using 9.58 there exists a  $q \in \mathbb{Q}$  such that  $a - 1 < q < a$  hence  $q \in I$  proving that  $\mathbb{Q} \cap I \neq \emptyset$ .

$I = ]a, \infty[ \wedge a \in \mathbb{R}$ . then  $a < a + 1$  and using 9.58 there exists a  $q \in \mathbb{Q}$  such that  $a < q < a + 1$  hence  $q \in I$  proving that  $\mathbb{Q} \cap I \neq \emptyset$ .

$I = ]-\infty, \infty[$ . then as  $\mathbb{Q} \subseteq \mathbb{R} = ]-\infty, \infty[$  and  $1 \in \mathbb{Q}$  we have that  $\mathbb{Q} \cap I \neq \emptyset$

hence

$$\forall I \in \mathcal{U} \text{ we have that } \mathbb{Q} \cap I \neq \emptyset \text{ or that } \{\mathbb{Q} \cap I\}_{I \in \mathcal{U}} \text{ is a family of non empty sets} \quad (11.14)$$

Using the Axiom of Choice [see 2.210] there exists a function  $f: \mathcal{U} \rightarrow \bigcup_{I \in \mathcal{U}} (\mathbb{Q} \cap I) \subseteq \mathbb{Q}$  such that  $f(I) \in \mathbb{Q} \cap I$ . Further if  $I, J \in \mathcal{U}$  with  $f(I) = f(J)$  then  $I = J$  [for if  $I \neq J$  then by 11.11 we have that  $I \cap J = \emptyset$  contradicting  $f(I) = f(J) \in (\mathbb{Q} \cap I) \cap (\mathbb{Q} \cap J) = \mathbb{Q} \cap \emptyset = \emptyset$ ]. So  $f$  is injective and as  $\mathbb{Q}$  is denumerable [see 7.40] we have by 5.67 that

$$\mathcal{U} \text{ is countable and } U = \bigcup_{I \in \mathcal{U}} I \quad (11.15)$$

Proving the theorem.  $\square$

**Corollary 11.76.** *Let  $\langle \mathbb{R}, \|\cdot\| \rangle$  be the normed space of the real numbers and  $U$  a non empty open set then there exists a countable family  $\{[a_i, b_i]\}_{i \in I}$  such that  $U = \bigcup_{i \in I} [a_i, b_i]$*

**Proof.** Using the previous proposition 11.75 there exists a countable set  $I$  and a family  $\{I_i\}_{i \in I}$  of open generalized intervals. Let  $i \in I$  then by 11.73 we have either

$I_i = [a_i, b_i] \wedge a_i, b_i \in \mathbb{R} \wedge a_i \leq b_i$ . Then if we take  $J_i = \{1\}$  and  $\{[a_{(i,j)}, b_{(i,j)}]\}_{j \in \{1\}}$  where  $a_{(i,1)} = a_i$  and  $b_{(i,1)} = b_i$  then  $J_i$  is countable and  $I_i = \bigcup_{j \in J_i} [a_{(i,j)}, b_{(i,j)}]$

$I_i = ]-\infty, a_i] \wedge a_i \in \mathbb{R}$ . Then take  $J_i = \mathbb{N}_0$  and  $\{[a_{(i,j)}, b_{(i,j)}]\}_{j \in J_i}$  by  $a_{(i,j)} = -i$  and  $b_{(i,j)} = a_i$  then  $J_i$  is countable and  $]-\infty, a_i[ = \bigcup_{n \in \mathbb{N}_0} ]-n, a_i[ = \bigcup_{j \in J_i} [a_{(i,j)}, b_{(i,j)}]$

$I_i = ]a_i, \infty[ \wedge a_i \in \mathbb{R}$ . Then take  $J_i = \mathbb{N}_0$  and  $\{[a_{(i,j)}, b_{(i,j)}]\}_{j \in J_i}$  by  $a_{(i,j)} = a_i$  and  $b_{(i,j)} = j$  then  $J_i$  is countable and  $]a_i, \infty[ = \bigcup_{n \in \mathbb{N}_0} ]a_i, n[ = \bigcup_{j \in J_i} [a_{(i,j)}, b_{(i,j)}]$

$I_i = ]-\infty, \infty[$ . Then take  $J_i = \mathbb{N}_0$  and  $\{[a_{(i,j)}, b_{(i,j)}]\}_{j \in J_i}$  where  $a_{(i,j)} = -j$  and  $b_{(i,j)} = j$  then  $J_i$  is countable and  $]-\infty, \infty[ = \bigcup_{n \in \mathbb{N}_0} ]-n, n[ = \bigcup_{j \in J_i} [a_{(i,j)}, b_{(i,j)}]$

proving that

$$\forall i \in I \text{ there exists a countable } J_i \text{ such that } I_i = \bigcup_{j \in J_i} [a_{(i,j)}, b_{(i,j)}] \quad (11.16)$$

Now we have

$$\begin{aligned} x \in \bigcup_{r \in \bigcup_{i \in I} \{i\} \times J_i} [a_r, b_r] &\Rightarrow x \in \bigcup_{r \in \bigcup_{i \in I} \{i\} \times J_i} [a_{(r_1, r_2)}, b_{(r_1, r_2)}] \\ &\Rightarrow \exists r \in \bigcup_{i \in I} \{i\} \times J_i \text{ so that } x \in [a_{(r_1, r_2)}, b_{(r_1, r_2)}] \\ &\Rightarrow \exists i \in I \text{ so that } r \in \{i\} \times J_i \Rightarrow r_1 = i \wedge r_2 \in J_i \\ &\stackrel{11.16}{\Rightarrow} [a_{r_1, r_2}, b_{r_1, r_2}] = [a_{1, r_2}, b_{i, r_2}] \subseteq I_i \\ &\Rightarrow x \in I_i \\ &\Rightarrow x \in \bigcup_{i \in I} I_i \end{aligned}$$

and

$$\begin{aligned} x \in \bigcup_{i \in I} I_i &\Rightarrow \exists i \in I \text{ so that } x \in I_i \\ &\stackrel{11.16}{\Rightarrow} \exists j \in J_i \text{ so that } x \in [a_{(i,j)}, b_{(i,j)}] \\ &\Rightarrow r = (i, j) \in \{i\} \times J_i \text{ with } x \in [a_{(r_1, r_2)}, b_{(r_1, r_2)}] \\ &\Rightarrow x \in \bigcup_{r \in \bigcup_{i \in I} \{i\} \times J_i} [a_{(r_1, r_2)}, b_{(r_1, r_2)}] \\ &\Rightarrow x \in \bigcup_{r \in \bigcup_{i \in I} \{i\} \times J_i} [a_r, b_r] \end{aligned}$$

which proves that

$$U = \bigcup_{i \in I} I_i = \bigcup_{r \in \bigcup_{i \in I} \{i\} \times J_i} [a_r, b_r] \quad (11.17)$$

Now using 5.70 we have  $\forall i \in I$  that  $\{i\} \times J_i$  is countable and using 5.74 we have that

$$\bigcup_{i \in I} \{i\} \times J_i \text{ is countable} \quad (11.18)$$

Finally 11.17 and 11.18 proves the theorem.  $\square$

**Theorem 11.77.** Let  $\langle X, \|\cdot\| \rangle$  be a normed space,  $A \subseteq X$  then  $A$  is bounded if and only if  $\exists M \in \mathbb{R}$  such that  $\forall x \in A \quad \|x\| \leq M$

**Proof.**

$\Rightarrow$ . If  $A = \emptyset$  then  $\forall x \in A \quad \|x\| < 1$  is vacuously satisfied. If  $A \neq \emptyset$  then there exists a  $x_0 \in A$ , further as  $A$  is bounded there exists a  $M' \in \mathbb{R}$  such that  $\forall x, y \in A$  we have  $\|x - y\| = d_{\|\cdot\|}(x, y) \leq M'$ . Take  $M = M + \|x_0\|$  then  $\forall x \in A$  we have  $\|x\| = \|x - x_0 + x_0\| \leq \|x - x_0\| + \|x_0\| \leq M' + \|x_0\| = M$ .

$\Leftarrow$ . Assume that  $\exists M' \in \mathbb{R}$  such that  $\forall x \in A$  we have  $\|x\| \leq M'$  so if we take  $M = 2 \cdot M'$  we have  $\forall x, y \in A$  that  $\|x - y\| \leq \|x\| + \|y\| \leq M' + M' = 2 \cdot M' = M$   $\square$

**Proposition 11.78.** Let  $\langle X, \|\cdot\| \rangle$  be a (pseudo) normed space with generated topology  $\mathcal{T}_{\|\cdot\|}$  and  $A \subseteq X$  then if we define  $\|\cdot\|_A: A \rightarrow \mathbb{R}$  by  $\|x\|_A = \|\cdot\|$  then  $\langle A, \|\cdot\|_A \rangle$  is a (pseudo) normed space and  $\mathcal{T}_{\|\cdot\|_A}$  is the subspace topology on  $A$  induced by  $\mathcal{T}_{\|\cdot\|}$ .

**Proof.** First we prove that  $\|\cdot\|_A$  is a pseudo norm on  $A$

1.  $\forall x \in A$  we have  $\|x\|_A = \|x\| \geq 0$
2.  $\forall x \in A$  and  $\forall \alpha \in \mathbb{R}(\mathbb{C})$  we have  $\|\alpha \cdot x\|_A = \|\alpha \cdot x\| = \alpha \cdot \|x\| = \alpha \cdot \|x\|_A$
3.  $\forall x, y \in A$  we have  $\|x + y\|_A = \|x + y\| \leq \|x\| + \|y\| = \|x\|_A + \|y\|_A$

Further if  $\|\cdot\|$  is a norm then if  $\|x\|_A = 0$  we have  $\|x\| = 0$  proving that  $x = 0$  proving that in this case  $\|\cdot\|_A$  is also a norm. Next as  $\forall x, y \in A$  we have that  $(d_{\|\cdot\|})_{|A \times A}(x, y) = d_{\|\cdot\|}(x, y) = \|x - y\| = \|x - y\|_A = d_{\|\cdot\|_A}(x, y)$  it follows that  $(d_{\|\cdot\|})_{|A \times A} = d_{\|\cdot\|_A}$ . Applying then 11.57 proves that  $\mathcal{T}_{\|\cdot\|_A}$  is the subspace topology on  $A$  induced by  $\mathcal{T}_{\|\cdot\|}$ .  $\square$

**Theorem 11.79.** Let  $\langle X, \|\cdot\| \rangle$  be a normed vector space over  $\mathbb{K} = (\mathbb{C} \text{ or } \mathbb{R})$  then  $\forall \alpha \neq 0, \alpha \in \mathbb{K}$  and  $\forall x \in X$  we have (see 10.122 for the definitions)

1.  $\forall M \subseteq X$  we have  $\overline{\alpha \cdot M} = \alpha \cdot \bar{M}$
2.  $\forall M$  open in  $X$  we have that  $\alpha \cdot M$  is open
3.  $\forall M$  open in  $X$  we have that  $x + M$  is open
4.  $\forall M$  open in  $X$  and  $A \subseteq X$  we have that  $A + M$  is open
5.  $\forall M \subseteq X$  we have  $\bar{M} - \bar{M} \subseteq \overline{M - M}$

**Proof.**

1. If  $x \in \overline{\alpha \cdot M}$  then as  $\alpha \neq 0$  we can define  $z = \frac{1}{\alpha} \cdot x$  so that  $x = \alpha \cdot z$ . If now  $z \in U$  open then there exists a  $\delta > 0$  such that  $z \in B_{\|\cdot\|}(z, \delta) \subseteq U$ . Now as  $x \in B_{\|\cdot\|}(x, |\alpha| \cdot \delta)$  a open set and  $x \in \overline{\alpha \cdot M}$  there exists by 11.20 a  $m \in M$  such that  $\alpha \cdot m \in B_{\|\cdot\|}(x, |\alpha| \cdot \delta)$  or  $\|x - \alpha \cdot m\| < |\alpha| \cdot \delta \Rightarrow \|\alpha \cdot z - \alpha \cdot m\| < |\alpha| \cdot \delta \Rightarrow |\alpha| \cdot \|z - m\| < |\alpha| \cdot \delta \Rightarrow \|z - m\| < \delta \Rightarrow m \in B_{\|\cdot\|}(z, \delta) \subseteq U$  so we have  $U \cap M \neq \emptyset$ . So using 11.20 we have that  $z \in \bar{M}$  or that  $x \in \alpha \cdot \bar{M}$ , this proves that

$$\overline{\alpha \cdot M} \subseteq \alpha \cdot \bar{M} \quad (11.19)$$

If  $x \in \alpha \cdot \bar{M}$  and  $x \in U$  a open set, then there exists a  $\delta > 0$  such that  $x \in B_{\|\cdot\|}(x, \delta) \subseteq U$ , now as  $x \in \alpha \cdot \bar{M}$  there exists a  $z \in \bar{M}$  such that  $x = \alpha \cdot z$  and as  $z \in B_{\|\cdot\|}\left(z, \frac{\delta}{|\alpha|}\right)$  a open set we have by 11.20 that there exist a  $m \in M$  such that  $m \in B_{\|\cdot\|}\left(z, \frac{\delta}{|\alpha|}\right)$  or  $\|z - m\| < \frac{\delta}{|\alpha|} \Rightarrow |\alpha| \cdot \|z - m\| < \delta \Rightarrow \|\alpha \cdot z - \alpha \cdot m\| < \delta \Rightarrow \|x - \alpha \cdot m\| < \delta \Rightarrow \alpha \cdot m \in B_{\|\cdot\|}(x, \delta) \xrightarrow{\alpha \cdot m \in \alpha \cdot M} U \cap (\alpha \cdot M) \neq \emptyset$  proving by 11.20 that  $x \in \overline{\alpha \cdot M}$ , which proves that

$$\alpha \cdot \bar{M} \subseteq \overline{\alpha \cdot M} \quad (11.20)$$

Using 11.19 and 11.20 we have that

$$\alpha \cdot \bar{M} = \overline{\alpha \cdot M}$$

2. If  $x \in \alpha \cdot M$  then  $x = \alpha \cdot m$  where  $m \in M$ . As  $M$  is open there exists a  $\delta > 0$  such that  $m \in B_{\|\cdot\|}(m, \delta) \subseteq M$ . So if  $z \in B_{\|\cdot\|}(x, |\alpha| \cdot \delta)$  then we have  $\|x - z\| < |\alpha| \cdot \delta \Rightarrow \frac{1}{|\alpha|} \cdot \|x - z\| < \delta \Rightarrow \left\| \frac{1}{\alpha} \cdot x - \frac{1}{\alpha} \cdot z \right\| < m \Rightarrow \left\| m - \frac{1}{\alpha} \cdot z \right\| < \delta \Rightarrow \frac{1}{\alpha} \cdot z \in B_{\|\cdot\|}(m, \delta) \subseteq M \Rightarrow z \in \alpha \cdot M$  so that  $x \in B_{\|\cdot\|}(x, |\alpha| \cdot \delta) \subseteq \alpha \cdot M$  proving that  $\alpha \cdot M$  is open.
3. Let  $z \in x + M$  then  $z - x \in M$  which is open and there exists a  $\delta > 0$  such that  $(z - x) \in B_{\|\cdot\|}(z - x, \delta) \subseteq M$ . If now  $y \in B_{\|\cdot\|}(z, \delta)$  then  $\|(y - x) - (z - x)\| = \|z - y\| < \delta$  or  $y - x \in B_{\|\cdot\|}(z - x, \delta) \subseteq M \Rightarrow y \in x + M$  proving that  $z \in B_{\|\cdot\|}(z, \delta) \subseteq x + M$  and thus that  $x + M$  is open.
4. This follows from (3) as  $A + M = \bigcup_{x \in A} (x + M)$  a union of open sets which is open by definition.
5. If  $x \in \bar{M} - \bar{M}$  then  $x = y_1 - y_2$ ,  $y_1, y_2 \in \bar{M}$  and suppose that  $x \in U$  a open set. So there exists a  $\delta > 0$  such that  $x \in B_{\|\cdot\|}(x, \delta) \subseteq U$ . As  $y_1, y_2 \in \bar{M}$  we have by the openness of  $B_{\|\cdot\|}\left(y_1, \frac{\delta}{2}\right)$ ,  $B_{\|\cdot\|}\left(y_2, \frac{\delta}{2}\right)$  and 11.20 that there exists  $z_1 \in B_{\|\cdot\|}\left(y_1, \frac{\delta}{2}\right) \cap M$ ,  $z_2 \in B_{\|\cdot\|}\left(y_2, \frac{\delta}{2}\right) \cap M$ . Take then  $z = z_1 - z_2 \in M - M$  then  $\|x - z\| = \|y_1 - z_1 + y_2 - z_2\| \leq \|y_1 - z_1\| + \|y_2 - z_2\| < \frac{\delta}{2} + \frac{\delta}{2} = \delta$  proving that  $z \in B_{\|\cdot\|}(x, \delta) \cap (M - M) \subseteq U \cap (M - M)$  so by 11.20 we have  $x \in \overline{M - M}$ .  $\square$

**Theorem 11.80.** Let  $\|\cdot\|_1, \|\cdot\|_2$  be two norms on a set  $X$  and  $\mathcal{T}_1, \mathcal{T}_2$  the generated normed topologies. Then  $\mathcal{T}_2$  is finer than  $\mathcal{T}_1$  iff there exists a  $M > 0$  such that  $\forall x \in X$  we have  $\|x\|_1 \leq M \cdot \|x\|_2$

**Proof.** We use 11.59 to prove this theorem

1.  $(\Rightarrow)$  As  $0 \in B_{\|\cdot\|_1}(0, 1)$  there exists by 11.59 a  $\delta > 0$  such that  $0 \in B_{\|\cdot\|_2}(0, \delta) \subseteq B_{\|\cdot\|_1}(0, 1)$  so if  $\|z\|_2 < \delta \Rightarrow z \in B_{\|\cdot\|_2}(0, \delta) \subseteq B_{\|\cdot\|_1}(0, 1) \Rightarrow \|z\|_1 < 1$  so if  $\|z\|_2 < \delta$  then we have  $\|z\|_1 < 1$ . Take now  $M = \frac{2}{\delta}$  then we have two cases to consider for  $x \in X$ 
  - a.  $(x = 0)$  then  $\|x\|_1 = 0 \leq 0 = M \cdot 0 = M \cdot \|x\|_2$
  - b.  $(x \neq 0)$  for  $z = \frac{1}{M \cdot \|x\|_2}x$  we have  $\|z\|_2 = \frac{1}{M \cdot \|x\|_2} \cdot \|x\|_2 = \frac{1}{M} = \frac{\delta}{2} < \delta \Rightarrow \|z\|_1 < 1 \Rightarrow \frac{1}{M \cdot \|x\|_2} \|x\|_1 < 1 \Rightarrow \|x\|_1 < M \cdot \|x\|_2 \Rightarrow \|x\|_1 \leq M \cdot \|x\|_2$
2.  $(\Leftarrow)$  If  $x \in X$  and  $\varepsilon > 0$  so that  $x \in B_{\|\cdot\|_1}(x, \varepsilon)$  taken then  $\delta = \frac{\varepsilon}{M} > 0$  (division is OK as  $M > 0$  by the hypothesis). If  $y \in B_{\|\cdot\|_2}(x, \delta)$  then  $\|x - y\|_2 < \delta$  and we have  $\|x - y\|_1 \leq M \cdot \|x - y\|_2 < M \cdot \delta = M \cdot \frac{\varepsilon}{M} = \varepsilon \Rightarrow y \in B_{\|\cdot\|_1}(x, \varepsilon)$  or  $x \in B_{\|\cdot\|_2}(x, \delta) \subseteq B_{\|\cdot\|_1}(x, \varepsilon)$  proving by 11.59 that  $\mathcal{T}_2$  is finer than  $\mathcal{T}_1$ .  $\square$

**Lemma 11.81.** Let  $A \subseteq \mathbb{R}$  then given  $\alpha \in \mathbb{R}$  with  $\alpha > 0$  then if  $\max(A)$  exists we have that  $\max(\alpha \cdot A)$  exists and  $\max(\alpha \cdot A) = \alpha \cdot \max(A)$  where  $\alpha \cdot A = \{\alpha \cdot a \mid a \in A\}$

**Proof.** As  $\max(A)$  exists we have that  $\max(A) \in A$  and  $\forall a \in A$  we have  $a \leq \max(A)$ . If now  $b \in \alpha \cdot A$  then there exists a  $a \in A$  such that  $b = \alpha \cdot a$ , from  $a \leq \max(A)$  and  $\alpha > 0$  we have that  $b = \alpha \cdot a \leq \alpha \cdot \max(A)$ . This together with the fact that  $\max(A) \in A \Rightarrow \alpha \cdot \max(A) \in \alpha \cdot A$  proves that  $\max(\alpha \cdot A)$  exists and is equal to  $\alpha \cdot \max(A)$ .  $\square$

**Theorem 11.82.** Let  $\{\langle X_i, \|\cdot\|_i \rangle\}_{i \in I}$ ,  $I$  finite be a finite family of normed spaces then  $\|\cdot\|: \prod_{i \in I} X_i \rightarrow \mathbb{R}$  defined by  $\|x\| = \max(\{\|\pi_i(x)\|_i \mid i \in I\})$  is a norm and the generated topology on  $\prod_{i \in I} X_i$  coincidences with the product topology  $\mathcal{T}_{\text{product}}$  (see 11.37) (using the norm topologies of  $X_i$ ). The norm  $\|\cdot\|$  is called the maximum norm.

**Proof.** First we prove that  $\|\cdot\|$  is a norm (note that by 5.50 the maximum is always defined)

1.  $\|x\| = \max(\{\|\pi_i(x)\|_i \mid i \in I\}) \geq 0$  as  $\forall i \in I$  we have  $\|\pi_i(x)\|_i \geq 0$
2.  $\|\alpha \cdot x\| = \max(\{\|\pi_i(\alpha \cdot x)\|_i \mid i \in I\}) \stackrel{\pi_i(\alpha \cdot x) = (\alpha \cdot x)(i) = \alpha \cdot x(i) = \alpha \cdot \pi_i(x)}{=} \max(\{\|\alpha \cdot \pi_i(x)\|_i \mid i \in I\}) = \max(\{|\alpha| \cdot \|\pi_i(x)\|_i \mid i \in I\}) \stackrel{11.81}{=} |\alpha| \cdot \max(\{\|\pi_i(x)\|_i \mid i \in I\}) = |\alpha| \cdot \|x\|$

3.  $\|x + y\| = \max (\{\|\pi_i(x + y)\|_i | i \in I\}) \underset{\pi_i(x+y)=(x+y)(i)=x(i)+y(i)=\pi_i(x)+\pi_i(y)}{=} \max (\{\|\pi_i(x) + \pi_i(y)\|_i | i \in I\}) \leq \max (\{\|\pi_i(x)\|_i + \|\pi_i(y)\|_i | i \in I\}) \leq \max (\{\max (\{\pi_i(x)_i | i \in I\}) + \max (\{\|\pi_i(x)\|_i | i \in I\})\}) = \max (\{\|\pi_i(x)\|_i | i \in I\}) + \max (\{\|\pi_i(x)\|_i | i \in I\}) = \|x\| + \|y\|$
4. If  $\|x\| = 0 \Rightarrow \max (\{\|\pi_i(x)\|_i | i \in I\}) = 0 \Rightarrow \forall i \in I \text{ we have } \|\pi_i(x)\|_i = 0 \Rightarrow \pi_i(x) = 0 \Rightarrow \forall i \in I \models x(i) = 0 \Rightarrow x = 0$

Next using 11.67 we find that the product topology is generated by  $d(x, y) = \max (\{d_{\|\cdot\|_i}(\pi_i(x), \pi_i(y)) | i \in I\}) = \max (\{\|\pi_i(x) - \pi_i(y)\|_i | i \in I\}) = \|x - y\|$  so  $\|\cdot\|$  defines indeed the metric that generates the product topology.  $\square$

**Example 11.83.** The product topology of  $\mathbb{R}^n = \prod_{i \in \{1, \dots, n\}} \mathbb{R}$  is generated by the maximum norm  $\|\cdot\|$  where  $\|x\| = \max (\{|x_i| | i \in \{1, \dots, n\}\})$

**Definition 11.84.** Two norms on a vector space are equivalent if they generate the same topology

**Theorem 11.85.** Let  $X$  be a vector space and  $\|\cdot\|_1, \|\cdot\|_2$  two norms on  $X$  then these norms are equivalent iff there exists  $M_1, M_2 \in \mathbb{R}$  with  $0 < M_1, M_2$  such that  $\forall x \in X$  we have  $M_1 \cdot \|x\|_1 \leq \|x\|_2 \leq M_2 \cdot \|x\|_1$

**Proof.** Let  $\mathcal{T}_1, \mathcal{T}_2$  be the topologies generated by the norms  $\|\cdot\|_1, \|\cdot\|_2$

1.  $(\Rightarrow)$  If  $\|\cdot\|_1, \|\cdot\|_2$  are equivalent then  $\mathcal{T}_1$  is finer than  $\mathcal{T}_2$  and  $\mathcal{T}_2$  is finer than  $\mathcal{T}_1$  so by 11.80 there exists a  $M'_1, M'_2 > 0$  such that  $\forall x \in X$  we have  $\|x\|_1 \leq \|x\|_2 \cdot M'_1$  and  $\|x\|_2 \leq \|x\|_1 \cdot M'_2$   $\Rightarrow \frac{1}{M'_1} \cdot \|x\|_1 \leq \|x\|_2 \leq \|x\|_1 \cdot M'_2$  which given  $M_1 = \frac{1}{M'_1}, M_2 = M'_2$  proves that  $\|x\|_1 \cdot M_1 \leq \|x\|_2 \leq \|x\|_1 \cdot M_2$
2.  $(\Leftarrow)$  If  $\forall x \in X \quad \|x\|_1 \cdot M_1 \leq \|x\|_2 \leq \|x\|_1 \cdot M_2$  then we have  $\forall x \in X$  that  $\|x\|_1 \leq \|x\|_2 \cdot \frac{1}{M_1}$  and  $\|x\|_2 \leq \|x\|_1 \cdot M_2$  which by 11.80 means that  $\mathcal{T}_1$  is finer than  $\mathcal{T}_2$  and  $\mathcal{T}_2$  is finer than  $\mathcal{T}_1$  so that  $\mathcal{T}_1 = \mathcal{T}_2$   $\square$

Using the above it is easy to prove that all norms in  $\mathbb{R}$  or  $\mathbb{C}$  are equivalent.

**Example 11.86.** All norms on  $\mathbb{R}$  and  $\mathbb{C}$  are equivalent with the norm  $\|\cdot\|$  (if we consider  $\langle \mathbb{R}, +, \cdot \rangle$  a real vector space as a field and  $\langle \mathbb{C}, +, \cdot \rangle$  a complex vector space as a field).

**Proof.** Let  $\|\cdot\|$  be a norm on  $\mathbb{R}(\mathbb{C})$  then  $\forall x \in \mathbb{R}(\mathbb{C})$  we have  $\|x\| = \|x \cdot 1\| = |x| \cdot \|1\|$  hence we have if we take  $M = \|1\|$  that  $M \cdot |x| = \|x\|$  or  $M \cdot |x| \leq \|x\| \leq M \cdot |x|$  proving that  $\|x\|$  is equivalent with  $\|\cdot\|$ .  $\square$

**Definition 11.87.** Let  $\langle X, \|\cdot\|_X \rangle$  and  $\langle Y, \|\cdot\|_Y \rangle$  be normed spaces then a isometry is a bijective function  $\varphi: X \rightarrow Y$  such that  $\forall x \in X$  we have that  $\|x\|_X = \|\varphi(x)\|_Y$

**Theorem 11.88.** Let  $\langle X, \|\cdot\|_X \rangle$  and  $\langle Y, \|\cdot\|_Y \rangle$  be normed spaces and  $\varphi: X \rightarrow Y$  a isometry then  $\varphi^{-1}: Y \rightarrow X$  is also a isometry

**Proof.** First because of 2.41 we have that  $\varphi^{-1}: Y \rightarrow X$  is also a bijection. Second if  $y \in Y$  then  $\|\varphi^{-1}(y)\|_X = \|\varphi(\varphi^{-1}(y))\|_Y = \|y\|_Y$   $\square$

**Theorem 11.89.** Let  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle$  and  $\langle Z, \|\cdot\|_Z \rangle$  be three normed spaces,  $\varphi: X \rightarrow Y$  and  $\psi: Y \rightarrow Z$  be isometries then  $\psi \circ \varphi: X \rightarrow Z$  is a isometry

**Proof.** By 2.49  $\psi \circ \varphi: X \rightarrow Z$  is a bijection. Further  $\forall x \in X$  we have  $\|(\psi \circ \varphi)(x)\|_Z = \|\psi(\varphi(x))\|_Z \underset{\psi \text{ is isometry}}{=} \|\varphi(x)\|_Y \underset{\varphi \text{ is isometry}}{=} \|x\|_X$   $\square$

**Theorem 11.90.** Let  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle$  be normed spaces and  $\varphi: X \rightarrow Y$  a linear isometry then if  $\mathcal{T}_X, \mathcal{T}_Y$  are the topologies generated by  $\|\cdot\|_X, \|\cdot\|_Y$  then  $\mathcal{T}_X = \{\varphi^{-1}(V) | V \in \mathcal{T}_Y\} = \{U \subseteq X | \varphi(U) \in \mathcal{T}_Y\}$  and  $\mathcal{T}_Y = \{V \subseteq Y | \varphi^{-1}(V) \in \mathcal{T}_X\} = \{\varphi(U) | U \in \mathcal{T}_X\}$ . In other words the topologies are equivalent (see 11.64)

**Proof.** The topologies  $T_X$ ,  $T_Y$  are generated by definition by the metrics  $d_X$ ,  $d_Y$  where  $d_X(x, y) = \|x - y\|_X$  and  $d_Y(x', y') = \|x' - y'\|_Y$ . If  $x, y \in X$  then  $d_Y(\varphi(x), \varphi(y)) = \|\varphi(x) - \varphi(y)\|_Y \stackrel{\varphi \text{ is linear}}{=} \|\varphi(x - y)\|_Y = \|x - y\|_X = d_X(x, y)$  proving that  $\varphi$  is also a isometry in the metric sense. Using 11.64 proves then the theorem.  $\square$

## 11.4 Inner product spaces

**Definition 11.91.** A real inner product space  $\langle X, \langle \cdot, \cdot \rangle \rangle$  is a vector space  $X$  over the field  $\mathbb{R}$  together with a mapping  $\langle \cdot, \cdot \rangle: X \times X \rightarrow \mathbb{R}$  (the inner product) satisfying the following:

1.  $\forall x, y \in X$  we have  $\langle x, y \rangle = \langle y, x \rangle$
2.  $\forall x, y, z \in X$  and  $\forall \alpha, \beta \in \mathbb{R}$  we have  $\langle \alpha \cdot x + \beta \cdot y, z \rangle = \alpha \cdot \langle x, z \rangle + \beta \cdot \langle y, z \rangle$

**Note 11.92.** Because of (1) we have also  $\langle z, \alpha \cdot x + \beta \cdot y \rangle = \alpha \cdot \langle z, x \rangle + \beta \cdot \langle z, y \rangle$ , so that  $\langle \cdot, \cdot \rangle$  is multi-linear.

3.  $\forall x \in X$  we have  $\langle x, x \rangle \geq 0$
4.  $\forall x \in X$  we have  $\langle x, x \rangle = 0 \Rightarrow x = 0$

**Example 11.93.** Let  $n \in \mathbb{N}_0$  then if we define  $\langle \cdot, \cdot \rangle: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$  by  $\langle (x_1, \dots, x_n), (y_1, \dots, y_n) \rangle = \sum_{i=1}^n x_i \cdot y_i$  then  $\langle \mathbb{R}^n, \langle \cdot, \cdot \rangle \rangle$  is a real inner product space

**Proof.**

1.  $\langle (x_1, \dots, x_n), (y_1, \dots, y_n) \rangle = \sum_{i=1}^n x_i \cdot y_i = \sum_{i=1}^n y_i \cdot x_i = \langle (y_1, \dots, y_n), (x_1, \dots, x_n) \rangle$
2.  $\langle \alpha \cdot (x_1, \dots, x_n) + \beta \cdot (y_1, \dots, y_n), (z_1, \dots, z_n) \rangle = \langle (\alpha \cdot x_1 + \beta \cdot y_1, \dots, \alpha \cdot x_n + \beta \cdot y_n), (z_1, \dots, z_n) \rangle = \sum_{i=1}^n (\alpha \cdot x_i + \beta \cdot y_i) \cdot z_i = \sum_{i=1}^n (\alpha \cdot x_i \cdot z_i + \beta \cdot y_i \cdot z_i) = \alpha \cdot \sum_{i=1}^n x_i \cdot z_i + \beta \cdot \sum_{i=1}^n y_i \cdot z_i = \alpha \cdot \langle (x_1, \dots, x_n), (z_1, \dots, z_n) \rangle + \beta \cdot \langle (y_1, \dots, y_n), (z_1, \dots, z_n) \rangle$
3.  $\langle (x_1, \dots, x_n), (x_1, \dots, x_n) \rangle = \sum_{i=1}^n x_i^2 \geq 0$
4. If  $0 = \langle (x_1, \dots, x_n), (x_1, \dots, x_n) \rangle = \sum_{i=1}^n x_i^2$  then if there exists a  $i \in \{1, \dots, n\}$  such that  $x_i \neq 0$  then  $0 < x_i^2$  hence  $0 < x_i^2 + \sum_{j \in \{1, \dots, n\} \setminus \{i\}} x_j^2 = \sum_{j \in \{i\}} x_i^2 + \sum_{j \in \{1, \dots, n\} \setminus \{i\}} x_j^2 = \sum_{i=1}^n x_i^2 = 0$  a contradiction so we have  $\forall i \in \{1, \dots, n\}$  that  $x_i = 0$  hence we have  $(x_1, \dots, x_n) = 0$   $\square$

**Definition 11.94.** A complex inner product space  $\langle X, \langle \cdot, \cdot \rangle \rangle$  is a vector space  $X$  over the field  $\mathbb{C}$  together with a mapping  $\langle \cdot, \cdot \rangle: X \times X \rightarrow \mathbb{C}$  satisfying the following:

1.  $\forall x, y \in X$  we have  $\langle x, y \rangle = \overline{\langle y, x \rangle}$

**Note 11.95.** Because of (1) we have  $\langle x, x \rangle = \overline{\langle x, x \rangle}$  or  $\langle x, x \rangle$  is real.

2.  $\forall x, y, z \in X$  and  $\alpha, \beta \in \mathbb{C}$  we have  $\langle \alpha \cdot x + \beta \cdot y, z \rangle = \alpha \cdot \langle x, z \rangle + \beta \cdot \langle y, z \rangle$

**Note 11.96.** Because of (1) we have  $\langle z, \alpha \cdot x + \beta \cdot y \rangle = \overline{\langle \alpha \cdot x + \beta \cdot y, z \rangle} = \overline{\alpha \cdot \langle x, z \rangle + \beta \cdot \langle y, z \rangle} = \bar{\alpha} \cdot \langle z, x \rangle + \bar{\beta} \cdot \langle z, y \rangle$

3.  $\forall x \in X$  we have  $\langle x, x \rangle \geq 0$  (make sense since  $\langle x, x \rangle$  is real)

4.  $\forall x \in X$  if  $\langle x, x \rangle = 0 \Rightarrow x = 0$

**Example 11.97.** Let  $n \in \mathbb{N}_0$  then if we define  $\langle \cdot, \cdot \rangle: \mathbb{C}^n \times \mathbb{C}^n \rightarrow \mathbb{C}$  by  $\langle (x_1, \dots, x_n), (y_1, \dots, y_n) \rangle = \sum_{i=1}^n x_i \cdot \bar{y}_i$  then  $\langle \mathbb{C}^n, \langle \cdot, \cdot \rangle \rangle$  is a complex inner product space

**Proof.**

1.  $\langle (x_1, \dots, x_n), (y_1, \dots, y_n) \rangle = \sum_{i=1}^n x_i \cdot \bar{y}_i = \sum_{i=1}^n \bar{x}_i \cdot y_i = \overline{\sum_{i=1}^n y_i \cdot \bar{x}_i} = \overline{\langle y, x \rangle}$

2.  $\langle \alpha \cdot (x, \dots, x_n) + \beta \cdot (y_1, \dots, y_n), (z_1, \dots, z_n) \rangle = \langle (\alpha \cdot x_1 + \beta \cdot y_1, \dots, \alpha \cdot x_n + \beta \cdot y_n), (z_1, \dots, z_n) \rangle = \sum_{i=1}^n (\alpha \cdot x_i + \beta \cdot y_i) \cdot z_i = \sum_{i=1}^n (\alpha \cdot x_i \cdot z_i + \beta \cdot y_i \cdot z_i) = \alpha \cdot \sum_{i=1}^n x_i \cdot z_i + \beta \cdot \sum_{i=1}^n y_i \cdot z_i = \alpha \cdot \langle (x_1, \dots, x_n), (z_1, \dots, z_n) \rangle + \beta \cdot \langle (y_1, \dots, y_n), (z_1, \dots, z_n) \rangle$
3.  $\langle (x_1, \dots, x_n), (x_1, \dots, x_n) \rangle = \sum_{i=1}^n x_i \cdot x_i = \sum_{i=1}^n |x_i|^2 \geq 0$
4. If  $0 = \langle (x_1, \dots, x_n), (x_1, \dots, x_n) \rangle = \sum_{i=1}^n |x_i|^2$  then if there exists a  $i \in \{1, \dots, n\}$  such that  $x_i \neq 0$  then  $0 < |x_i|^2$  hence  $0 < |x_i|^2 + \sum_{j \in \{1, \dots, n\} \setminus \{i\}} |x_j|^2 = \sum_{j \in \{i\}} |x_j|^2 + \sum_{j \in \{1, \dots, n\} \setminus \{i\}} |x_j|^2 = \sum_{i=1}^n |x_i|^2 = 0$  a contradiction so we have  $\forall i \in \{1, \dots, n\}$  that  $x_i = 0$  hence we have  $(x_1, \dots, x_n) = 0$   $\square$

**Lemma 11.98.** Let  $\langle X, \langle \cdot, \cdot \rangle \rangle$  be a real (complex) inner product space  $n \in \mathbb{N}_0$ ,  $y \in X$  and  $\{x_i\}_{i \in \{1, \dots, n\}} \subseteq X$  then  $\langle \sum_{i=1}^n x_i, y \rangle = \sum_{i=1}^n \langle x_i, y \rangle$  and  $\langle y, \sum_{i=1}^n x_i \rangle = \sum_{i=1}^n \langle y, x_i \rangle$

**Proof.** The first part is easily proved by induction so let  $\mathcal{S} = \{n \in \mathbb{N}_0 \mid \forall \{x_i\}_{i \in \{1, \dots, n\}} \subseteq X \text{ we have } \langle \sum_{i=1}^n x_i, y \rangle = \sum_{i=1}^n \langle x_i, y \rangle\}$  then we have

**$n = 1$ .** this is true as for  $\{x_i\}_{i \in \{1, \dots, 1\}}$  we have  $\langle \sum_{i=1}^1 x_i, y \rangle = \langle x_1, y \rangle = \sum_{i=1}^1 \langle x_i, y \rangle$

**$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$ .** let  $\{x_i\}_{i \in \{1, \dots, n+1\}}$  then we have  $\langle \sum_{i=1}^{n+1} x_i, y \rangle = \langle x_{n+1} + (\sum_{i=1}^n x_i), y \rangle = \langle x_{n+1}, y \rangle + \langle \sum_{i=1}^n x_i, y \rangle = \langle x_{n+1}, y \rangle + \sum_{i=1}^n \langle x_i, y \rangle = \sum_{i=1}^{n+1} \langle x_i, y \rangle$  proving that  $n+1 \in \mathcal{S}$   $\square$

The second part is equally easy proved by induction so let  $\mathcal{S} = \{n \in \mathbb{N}_0 \mid \forall \{x_i\}_{i \in \{1, \dots, n\}} \subseteq X \text{ we have } \langle y, \sum_{i=1}^n x_i \rangle = \sum_{i=1}^n \langle y, x_i \rangle\}$  then we have

**$n = 1$ .** this is true as for  $\{x_i\}_{i \in \{1, \dots, 1\}}$  we have  $\langle y, \sum_{i=1}^1 x_i \rangle = \langle y, x_1 \rangle = \sum_{i=1}^1 \langle y, x_i \rangle$

**$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$ .** let  $\{x_i\}_{i \in \{1, \dots, n+1\}}$  then we have  $\langle y, \sum_{i=1}^{n+1} x_i \rangle = \langle y, x_{n+1} + (\sum_{i=1}^n x_i) \rangle = \langle y, x_{n+1} \rangle + \langle y, \sum_{i=1}^n x_i \rangle = \langle y, x_{n+1} \rangle + \sum_{i=1}^n \langle y, x_i \rangle = \sum_{i=1}^{n+1} \langle y, x_i \rangle$  proving that  $n+1 \in \mathcal{S}$

**Theorem 11.99.** Let  $\langle X, \langle \cdot, \cdot \rangle \rangle$  be a real (complex) inner product space then  $\forall x \in X$  we have  $\langle x, 0 \rangle = \langle 0, x \rangle = 0$

**Proof.**  $\langle 0, x \rangle = \langle 0 \cdot x, x \rangle = |0| \cdot \langle x, x \rangle = 0 \cdot \langle x, x \rangle = 0$ . For the real case we have then  $\langle x, 0 \rangle = \langle 0, x \rangle = 0$  and for the complex case we have  $\langle x, 0 \rangle = \overline{\langle 0, x \rangle} = \bar{0} = 0$   $\square$

**Corollary 11.100.** Let  $\langle X, \langle \cdot, \cdot \rangle \rangle$  be a real (complex) inner product space then if  $\forall x \in X$  we have  $\langle x, y \rangle = \langle x, z \rangle$  we conclude that  $y = z$

**Proof.** By assumption we have  $\langle z - y, y \rangle = \langle z - y, z \rangle = \langle z - y, y + (z - y) \rangle = \langle z - y, y \rangle + \langle z - y, z - y \rangle \Rightarrow \langle z - y, z - y \rangle = 0 \Rightarrow z - y = 0 \Rightarrow z = y$   $\square$

**Lemma 11.101.** If  $\{s_i\}_{i \in \{1, \dots, n\}}$  is a family of elements  $\mathbb{R}$  ( $\mathbb{C}_{\mathbb{R}}$ ) such that  $\forall i \in \{1, \dots, n\} \models s_i \geq 0$  then  $\sum_{i \in \{1, \dots, n\}} s_i = 0$  implies that  $\forall i \in \{1, \dots, n\}$  we have  $s_i = 0$

**Proof.** We prove this by induction on  $n$ . So let  $S = \{n \in \mathbb{N}_0 \mid \text{if } \{s_i\}_{i \in \{1, \dots, n\}} \text{ is such that } s_i \geq 0 \text{ then from } \sum_{i \in \{1, \dots, n\}} s_i = 0 \text{ it follows that } \forall i \in \{1, \dots, n\} \text{ we have } s_i = 0\}$ , we have then

1. If  $n = 1$  then we have  $0 = \sum_{i \in \{1, \dots, 1\}} s_i = s_1 \Rightarrow \forall i \in \{1, \dots, 1\}$  we have  $s_i = 0$  so that  $1 \in S$
2. Assume that  $n \in S$  if now  $\{s_i\}_{i \in \{1, \dots, n+1\}}$  is such that  $s_i \geq 0$  then if  $0 = \sum_{i \in \{1, \dots, n+1\}} s_i = \sum_{i \in \{1, \dots, n\}} s_i + s_{n+1} \stackrel{9.42, 9.29}{\Rightarrow} s_{n+1} = 0, \sum_{i \in \{1, \dots, n\}} s_i = 0 \stackrel{n \in S}{\Rightarrow} s_{n+1} = 0, \forall i \in \{1, \dots, n\}$  we have  $s_i \Rightarrow \forall i \in \{1, \dots, n+1\} \models s_i \Rightarrow n+1 \in S$

Using induction we have then proved the theorem.  $\square$

**Theorem 11.102.** Let  $X$  be a finite dimensional real vector space, given a basis  $\{e_i\}_{i \in \{1, \dots, n\}}$  define  $\forall x, y \in X \langle x, y \rangle = \sum_{i \in \{1, \dots, n\}} x_i \cdot y_i$  where  $x = \sum_{i \in \{1, \dots, n\}} x_i \cdot e_i$  and  $y = \sum_{i \in \{1, \dots, n\}} y_i$  are the unique expansions of  $x, y$  in  $\{e_i\}_{i \in \{1, \dots, n\}}$ . We have then that  $\langle X, \langle \cdot, \cdot \rangle \rangle$  is a real inner product space.

**Proof.**

1. If  $x, y \in X$  then  $\langle x, y \rangle = \sum_{i \in \{1, \dots, n\}} x_i \cdot y_i = \sum_{i \in \{1, \dots, n\}} y_i \cdot x_i = \langle y, x \rangle$
2. If  $x, y, z \in X$  and  $\alpha, \beta \in \mathbb{R}$  then  $\alpha \cdot x + \beta \cdot y = \alpha \cdot \sum_{i \in \{1, \dots, n\}} x_i + \beta \cdot \sum_{i \in \{1, \dots, n\}} y_i = \sum_{i \in \{1, \dots, n\}} (\alpha \cdot x_i + \beta \cdot y_i)$  so that  $\langle \alpha \cdot x + \beta \cdot y, z \rangle = \sum_{i \in \{1, \dots, n\}} (\alpha \cdot x_i + \beta \cdot y_i) \cdot z_i = \sum_{i \in \{1, \dots, n\}} (\alpha \cdot (x_i \cdot z_i) + \beta \cdot (y_i \cdot z_i)) = \alpha \cdot \sum_{i \in \{1, \dots, n\}} x_i \cdot z_i + \beta \cdot \sum_{i \in \{1, \dots, n\}} y_i \cdot z_i = \alpha \cdot \langle x, z \rangle + \beta \cdot \langle y, z \rangle$
3. If  $x \in X$  then  $\langle x, x \rangle = \sum_{i \in \{1, \dots, n\}} x_i \cdot x_i = \sum_{i \in \{1, \dots, n\}} x_i^2 \geq 0$  (using 9.41 and 10.61).
4. If  $x \in X$  let  $0 = \langle x, x \rangle = \sum_{i \in \{1, \dots, n\}} x_i^2$  then if  $x \neq 0$  there exists a  $i \in \{1, \dots, n\}$  such that  $x_i \neq 0$  then we have  $x_i^2 > 0$  (see 9.41) and thus by 10.61 we have  $\langle x, x \rangle = \sum_{i \in \{1, \dots, n\}} x_i^2 > 0$  contradicting  $\langle x, x \rangle = 0$  so we must conclude that  $x = 0$   $\square$

**Lemma 11.103.** If  $\{x_i\}_{i \in \{1, \dots, n\}}$  is a family in  $\mathbb{C}$  then  $\overline{\sum_{i \in \{1, \dots, n\}} x_i} = \sum_{i \in \{1, \dots, n\}} \bar{x}_i$

**Proof.** We prove this by induction, so let  $S = \{n \in \mathbb{N}_0 \mid \{x_i\}_{i \in \{1, \dots, n\}}$  is a family in  $\mathbb{C}$  then  $\overline{\sum_{i \in \{1, \dots, n\}} x_i} = \sum_{i \in \{1, \dots, n\}} \bar{x}_i\}$  then we have

1. If  $n = 0$  then  $\overline{\sum_{i \in \{0, \dots, 0\}} x_i} = \bar{x}_i = \sum_{i \in \{0, \dots, 0\}} \bar{x}_i \Rightarrow 0 \in S$
2. If  $n \in S$  then  $\overline{\sum_{i \in \{1, \dots, n+1\}} x_i} = \overline{\sum_{i \in \{1, \dots, n\}} x_i + x_{n+1}} = \overline{\sum_{i \in \{1, \dots, n\}} x_i} + \overline{x_{n+1}} = \sum_{i \in \{1, \dots, n\}} \bar{x}_i + \bar{x}_{n+1} = \sum_{i \in \{1, \dots, n+1\}} \bar{x}_i$  so that  $n+1 \in S$ .  $\square$

**Theorem 11.104.** Let  $X$  be a real (complex) finite dimensional vector space, given a basis  $\{e_i\}_{i \in \{1, \dots, n\}}$  define  $\forall x, y \in X$   $\langle x, y \rangle = \sum_{i \in \{1, \dots, n\}} x_i \cdot \bar{y}_i$  where  $x = \sum_{i \in \{1, \dots, n\}} x_i \cdot e_i$  and  $y = \sum_{i \in \{1, \dots, n\}} y_i \cdot e_i$  are the unique expansions of  $x, y$  in  $\{e_i\}_{i \in \{1, \dots, n\}}$ . We have then that  $\langle X, \langle \rangle \rangle$  is a complex inner product space

**Proof.**

1. If  $x, y \in X$  then  $\langle x, y \rangle = \sum_{i \in \{1, \dots, n\}} x_i \cdot \bar{y}_i = \sum_{i \in \{1, \dots, n\}} \overline{x_i \cdot \bar{y}_i} = \sum_{i \in \{1, \dots, n\}} \bar{x}_i \cdot \bar{\bar{y}_i} = \overline{\sum_{i \in \{1, \dots, n\}} x_i \cdot y_i} = \overline{\sum_{i \in \{1, \dots, n\}} y_i \cdot \bar{x}_i} = \langle y, x \rangle$
2. If  $x, y, z \in X$  and  $\alpha, \beta \in \mathbb{C}$  then  $\langle \alpha \cdot x + \beta \cdot y, z \rangle = \sum_{i \in \{1, \dots, n\}} (\alpha \cdot x_i + \beta \cdot y_i) \cdot \bar{z}_i = \sum_{i \in \{1, \dots, n\}} (\alpha \cdot (x_i \cdot \bar{z}_i) + \beta \cdot (y_i \cdot \bar{z}_i)) = \alpha \cdot \sum_{i \in \{1, \dots, n\}} x_i \cdot \bar{z}_i + \beta \cdot \sum_{i \in \{1, \dots, n\}} y_i \cdot \bar{z}_i = \alpha \cdot \langle x, z \rangle + \beta \cdot \langle y, z \rangle$
3. If  $x \in X$  then  $\langle x, x \rangle = \sum_{i \in \{1, \dots, n\}} x_i \cdot \bar{x}_i = \sum_{i \in \{1, \dots, n\}} (|x_i|^2, 0) \geq 0$  (using 9.31 and 10.61)
4. If  $x = 0$  and  $\langle x, x \rangle = 0$  then  $\sum_{i \in \{1, \dots, n\}} (|x_i|^2, 0) = 0$  and using 10.61 we have  $\forall i \in \{1, \dots, n\}$  that  $(|x_i|^2, 0) = 0 \Rightarrow |x_i|^2 = 0$   $\xrightarrow{9.31} x_i = 0$  proving that  $x = 0$ .  $\square$

**Definition 11.105.** Let  $\langle X, \langle \rangle \rangle$  be a real (complex) inner product space then the inner product norm is defined as  $\forall x \in X$  we define  $\|\cdot\|: X \rightarrow \mathbb{R}$  by  $\|x\| = \sqrt{\langle x, x \rangle}$

**Lemma 11.106.** Let  $\langle X, \langle \rangle \rangle$  be a real (complex) inner product space then  $\|0\| = 0$

**Proof.**  $\|0\| = \sqrt{\langle 0, 0 \rangle} = \sqrt{0} = 0$   $\square$

**Theorem 11.107. (Schwartz's inequality)** If  $\langle X, \langle \rangle \rangle$  is a real (complex) inner space and  $\|\cdot\|$  the inner product norm then  $\forall x, y \in X$  we have  $|\langle x, y \rangle| \leq \|x\| \cdot \|y\|$ . Equality holds if and only if  $x$  and  $y$  are linearly dependent.

**Proof.** We have the following cases to consider

1. ( $x = 0$ ) then  $x, y$  are linear dependent,  $\langle 0, y \rangle \xrightarrow{11.99} 0$  and because of the previous lemma we have  $\|x\| = \|0\| = 0$  so that we have  $|\langle x, y \rangle| = |\langle 0, y \rangle| = 0 = 0 \cdot \|y\| = \|x\| \cdot \|y\|$

2. ( $y = 0$ ) then  $x, y$  are linear dependent,  $\langle x, 0 \rangle \stackrel{11.99}{=} 0$  and because of the previous lemma we

have  $\|y\| = \|0\| = 0$  so that we have  $|\langle x, y \rangle| = |\langle x, 0 \rangle| = 0 = \|x\| \cdot 0 = \|x\| \cdot \|y\|$

3. ( $x \neq 0 \wedge y \neq 0$ ) then  $\langle y, y \rangle \neq 0$  so that we can define  $\alpha = \frac{-\langle x, y \rangle}{\langle y, y \rangle} = -\frac{\langle x, y \rangle}{\|y\|^2}$

a. (**real inner product**)  $0 \leq \langle x + \alpha \cdot y, x + \alpha \cdot y \rangle = \langle x, x \rangle + 2 \cdot \alpha \cdot \langle x, y \rangle + \alpha^2 \cdot$

$$\langle y, y \rangle = \|x\|^2 - 2 \cdot \frac{\langle x, y \rangle^2}{\|y\|^2} + \frac{\langle x, y \rangle^2}{\|y\|^4} \cdot \|y\|^2 = \|x\|^2 - 2 \cdot \frac{\langle x, y \rangle^2}{\|y\|^2} + \frac{\langle x, y \rangle^2}{\|y\|^2} = \|x\|^2 -$$

$$\frac{\langle x, y \rangle^2}{\|y\|^2} \xrightarrow{\text{multiply by } \|y\|^2} 0 \leq \|x\|^2 \cdot \|y\|^2 - \langle x, y \rangle^2 \Rightarrow \langle x, y \rangle^2 \leq \|x\|^2 \cdot \|y\|^2 = (\|x\| \cdot \|y\|)^2 \xrightarrow{9.72} \langle x, y \rangle \leq \|x\| \cdot \|y\|. \text{ Now assume that } x, y \text{ are linearly dependent then we have either}$$

i. ( $x = \beta \cdot y$ ) then  $\langle x, y \rangle^2 = \beta \cdot \langle x, y \rangle \cdot \langle y, y \rangle = \langle x, x \rangle \cdot \langle y, y \rangle = \|x\|^2 \cdot \|y\|^2 \xrightarrow{9.72} |\langle x, y \rangle| = \|x\| \cdot \|y\|$

ii. ( $y = \beta \cdot x$ ) then  $\langle x, y \rangle^2 = \beta \cdot \langle x, x \rangle \cdot \langle x, y \rangle = \langle x, x \rangle \cdot \langle y, y \rangle = \|x\|^2 \cdot \|y\|^2 \xrightarrow{9.72} |\langle x, y \rangle| = \|x\| \cdot \|y\|$

Assume now that  $|\langle x, y \rangle| = \|x\| \cdot \|y\|$  then we have that  $\langle y, y \rangle \cdot x - \langle x, y \rangle \cdot y, \langle y, y \rangle \cdot x - \langle x, y \rangle \cdot y \rangle = \langle y, y \rangle^2 \cdot \langle x, x \rangle - 2 \cdot \langle y, y \rangle \cdot \langle x, y \rangle^2 + \langle x, y \rangle^2 \cdot \langle y, y \rangle = \|y\|^4 \cdot \|x\|^2 - \|y\|^2 \cdot \langle x, y \rangle^2 = \|y\|^4 - \|y\|^2 \cdot \|x\|^2 \cdot \|y\|^2 = 0$  proving that  $\langle y, y \rangle \cdot x = \langle x, y \rangle \cdot y \Rightarrow x = \frac{\langle x, y \rangle}{\langle y, y \rangle} \cdot y$  and linear dependency

b. (**complex inner product**)  $0 \leq \langle x + \alpha \cdot y, x + \alpha \cdot y \rangle = \langle x, y \rangle + \bar{\alpha} \cdot \langle x, y \rangle + \alpha \cdot \langle y, x \rangle + \alpha \cdot \langle y, y \rangle = \|x\|^2 - \frac{\langle x, y \rangle}{\|y\|^2} \cdot \langle x, y \rangle - \frac{\langle x, y \rangle}{\|y\|^2} \cdot \langle x, y \rangle + |\alpha|^2 \cdot \|y\|^2 = \|x\|^2 - 2 \cdot \frac{|\langle x, y \rangle|^2}{\|y\|^2} + \frac{|\langle x, y \rangle|^2}{\|y\|^4} \cdot$

$$\|y\|^2 = \|x\|^2 - \frac{|\langle x, y \rangle|^2}{\|y\|^2} \Rightarrow 0 \leq \|x\|^2 \cdot \|y\|^2 - |\langle x, y \rangle|^2 \Rightarrow |\langle x, y \rangle|^2 \leq \|x\|^2 \cdot \|y\|^2 \xrightarrow{9.72} |\langle x, y \rangle| \leq \|x\| \cdot \|y\|. \text{ Now assume that } x, y \text{ are linearly dependent then we have either}$$

i. ( $x = \beta \cdot y$ ) then  $\langle x, y \rangle^2 = \beta \cdot \langle x, y \rangle \cdot \langle y, y \rangle = \langle x, x \rangle \cdot \langle y, y \rangle = \|x\|^2 \cdot \|y\|^2 \xrightarrow{9.72} |\langle x, y \rangle| = \|x\| \cdot \|y\|$

ii. ( $y = \beta \cdot x$ ) then  $\langle x, y \rangle^2 = \beta \cdot \langle x, x \rangle \cdot \langle x, y \rangle = \langle x, x \rangle \cdot \langle y, y \rangle = \|x\|^2 \cdot \|y\|^2 \xrightarrow{9.72} |\langle x, y \rangle| = \|x\| \cdot \|y\|$

Assume now that  $|\langle x, y \rangle| = \|x\| \cdot \|y\|$  then  $\langle y, y \rangle \cdot x - \langle x, y \rangle \cdot y, \langle y, y \rangle \cdot x - \langle x, y \rangle \cdot y \rangle = \langle y, y \rangle^2 \cdot \langle x, x \rangle - \langle y, y \rangle \cdot \langle x, y \rangle \cdot \langle x, y \rangle - \langle x, y \rangle \cdot \langle y, y \rangle \cdot \langle y, x \rangle + \langle x, y \rangle \cdot \langle x, y \rangle \cdot \langle y, y \rangle = \|y\|^4 \cdot \|x\|^2 - \|y\|^2 \cdot |\langle x, y \rangle|^2 - \|y\|^2 \cdot |\langle x, y \rangle|^2 + \|y\|^2 \cdot |\langle x, y \rangle|^2 = \|y\|^4 \cdot \|x\|^2 - 2 \cdot \|y\|^2 \cdot$

$$\|x\|^2 \cdot \|y\|^2 + \|y\|^2 \cdot \|x\|^2 \cdot \|y\|^2 = 0 \text{ meaning that } \langle y, y \rangle \cdot x = \langle x, y \rangle \cdot y \Rightarrow x = \frac{\langle x, y \rangle}{\langle y, y \rangle} \cdot y \text{ and linear dependency. } \square$$

**Theorem 11.108.** Let  $\langle X, \langle \rangle \rangle$  be a inner product space and  $\|\cdot\|$  the inner product norm then  $\forall x, y \in X$  we have  $\|x + y\| \leq \|x\| + \|y\|$

**Proof.** We make a distinction between the real case and the complex case

1. (**real inner product**)  $\|x + y\|^2 = \langle x + y, x + y \rangle = \langle x, x \rangle + 2 \cdot \langle x, y \rangle + \langle y, y \rangle \leq \|x\|^2 + 2 \cdot |\langle x, y \rangle| + \|y\|^2 \leq \|x\|^2 + 2 \cdot \|x\| \cdot \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2 \xrightarrow{9.72} \|x + y\| \leq \|x\| + \|y\|$

2. (**complex inner product**)  $\|x + y\|^2 = \langle x + y, x + y \rangle = \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle = \|x\|^2 + \langle x, y \rangle + \langle x, y \rangle + \|y\|^2 = \|x\|^2 + 2 \cdot \text{Re}(\langle x, y \rangle) + \|y\|^2 \leq 9.33 \|x\|^2 + 2 \cdot |\langle x, y \rangle| + \|y\|^2 \leq \|x\|^2 + 2 \cdot \|x\| \cdot \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2 \xrightarrow{9.72} \|x + y\| \leq \|x\| + \|y\| \text{ } \square$

**Theorem 11.109.** Let  $\langle X, \langle \rangle \rangle$  be a inner product space and  $\|\cdot\|$  the inner product norm then  $\langle X, \|\cdot\| \rangle$  forms a normed space

**Proof.**

1.  $\|x\| = \sqrt{\langle x, x \rangle} \geq 0$  (or  $\text{Re}(\sqrt{\langle x, x \rangle}) > 0$  in the complex case)

2.  $\|\alpha \cdot x\| = \sqrt{\langle \alpha \cdot x, \alpha \cdot x \rangle} \cdot \sqrt{\alpha^2 \cdot \langle x, x \rangle} = |\alpha| \cdot \sqrt{\langle x, x \rangle} = |\alpha| \cdot \|x\|$  (or  $\text{Re}(\sqrt{\langle \alpha \cdot x, \alpha \cdot x \rangle}) = \text{Re}(\sqrt{\alpha^2 \cdot \langle x, x \rangle}) = \text{Re}(\alpha \cdot \sqrt{\langle x, x \rangle}) = \alpha \cdot \text{Re}(\sqrt{\langle x, x \rangle}) = \alpha \cdot \|x\|$  in the complex case)

3. Using 11.108 we have  $\|x + y\| \leq \|x\| + \|y\|$
4. If  $\|x\| = 0 \Rightarrow \sqrt{\langle x, x \rangle} = 0 \Rightarrow \langle x, x \rangle = 0 \Rightarrow x = 0$  (or  $\operatorname{Re}(\sqrt{\langle x, x \rangle}) = 0 \Rightarrow \sqrt{\langle x, x \rangle} = 0 \Rightarrow \langle x, x \rangle = 0 \Rightarrow x = 0$  in the complex case).  $\square$

**Example 11.110.** Let  $n \in \mathbb{N}_0$  then  $\langle \mathbb{R}^n, \|\cdot\|_e \rangle$  where  $\|(x_1, \dots, x_n)\| = \sqrt{\sum_{i=1}^n x_i^2}$  is a normed space

**Proof.** This follows directly from the above theorem and 11.93.  $\square$

We can extend the above to general finite dimensional spaces

**Example 11.111.** If  $X$  is a finite dimensional real (complex) space with the canonical inner product defined by the basis  $\{e_i\}_{i \in \{1, \dots, n\}}$  then  $\forall x \in X$  we have  $\|x\| = \sqrt{\sum_{i \in \{1, \dots, n\}} x_i^2} = \sqrt{\sum_{i \in \{1, \dots, n\}} |x_i|^2}$  (or  $\|x\| = \operatorname{Re}(\sqrt{\sum_{i \in \{1, \dots, n\}} x_i \cdot \bar{x}_i})$   $\stackrel{10.177}{=} \sqrt{\sum_{i \in \{1, \dots, n\}} \operatorname{Re}(x_i \cdot \bar{x}_i)} = \sqrt{\sum_{i \in \{1, \dots, n\}} |x_i|^2}$  in the complex case) as the inner product norm (here  $x = \sum_{i \in \{1, \dots, n\}} x_i \cdot e_i$  is the unique expansion of

The following shows that  $\mathbb{R}^2$  and  $\mathbb{C}$  are isometric

**Example 11.112.** Let  $\langle \mathbb{R}^2, \|\cdot\|_e \rangle$  be the normed space using the euclidean norm (see 11.110),  $\langle \mathbb{C}, \|\cdot\| \rangle$  be the normed space using the canonical norm then  $\mathcal{C}: \mathbb{R}^2 \rightarrow \mathbb{C}$  defined by  $\mathcal{C}((x, y)) = x + i \cdot y$  is a isometry

**Proof.** Using 10.182 we have that  $\mathcal{C}$  is a isomorphims hence a bijection. Further  $|\mathcal{C}((x, y))| = |x + i \cdot y| = \sqrt{x^2 + y^2} = \|(x, y)\|_e$   $\square$

In fact finite dimensional real (complex) spaces of the same dimensionality are isometric using the inner product norm.

**Theorem 11.113.** Let  $X, Y$  be finite dimensional real vector spaces of the same dimensionality with basis  $\{e_i\}_{i \in \{1, \dots, n\}}$ ,  $\{f_i\}_{i \in \{1, \dots, n\}}$  then both spaces are isomorphic and isometric, so that the topologies generated by there inner product norm are equivalent (see 11.90)

**Proof.** Define  $\varphi: X \rightarrow Y$  by  $x = \sum_{i \in \{1, \dots, n\}} x_i \cdot e_i \rightarrow \varphi(x) = \sum_{i \in \{1, \dots, n\}} x_i \cdot f_i$  then we have

1. **(injectivity)** If  $\varphi(x) = \varphi(y)$  then  $\sum_{i \in \{1, \dots, n\}} x_i \cdot f_i = \sum_{i \in \{1, \dots, n\}} y_i \cdot f_i$  then by uniqueness of the expansion in a basis we must have  $\forall i \in \{1, \dots, n\}$  that  $x_i = y_i \Rightarrow x = y$
2. **(surjectivity)** If  $y \in Y$  then  $y = \sum_{i \in \{1, \dots, n\}} y_i \cdot f_i$  and  $\varphi(\sum_{i \in \{1, \dots, n\}} y_i \cdot e_i) = \sum_{i \in \{1, \dots, n\}} y_i \cdot f_i = y$

Also  $\varphi$  is multilinear for if  $x = \sum_{i \in \{1, \dots, n\}} x_i \cdot e_i$ ,  $y = \sum_{i \in \{1, \dots, n\}} y_i \cdot e_i$  then  $\varphi(\alpha \cdot x + \beta \cdot y) = \varphi(\alpha \cdot \sum_{i \in \{1, \dots, n\}} x_i \cdot e_i + \beta \cdot \sum_{i \in \{1, \dots, n\}} y_i \cdot e_i) = \varphi(\sum_{i \in \{1, \dots, n\}} (\alpha \cdot x_i + \beta \cdot y_i) \cdot e_i) = \sum_{i \in \{1, \dots, n\}} (\alpha \cdot x_i + \beta \cdot y_i) \cdot f_i = \alpha \cdot \sum_{i \in \{1, \dots, n\}} x_i \cdot f_i + \beta \cdot \sum_{i \in \{1, \dots, n\}} y_i \cdot f_i = \alpha \cdot \varphi(x) + \beta \cdot \varphi(y)$

Finally we have if  $x = \sum_{i \in \{1, \dots, n\}} x_i \cdot e_i$  then  $\|\varphi(x)\| = \|\sum_{i \in \{1, \dots, n\}} x_i \cdot f_i\| = \sqrt{\sum_{i \in \{1, \dots, n\}} |x_i|^2} = \|x\|$   $\square$

**Definition 11.114.** Let  $\langle X, \langle \cdot, \cdot \rangle \rangle$  be a real (complex) inner vector space then  $x, y \in X$  are **orthogonal** iff  $\langle X, y \rangle = 0$

**Definition 11.115.** Let  $\langle X, \langle \cdot, \cdot \rangle \rangle$  be a real (complex) inner vector space then a non empty family  $\{e_i\}_{i \in I}$  on  $X$  is **orthonormal** iff  $\forall i, j \in I$  we have  $\langle e_i, e_j \rangle = \delta_{i,j}$  (hence  $\|e_i\| = 1$ )

**Example 11.116.** Let  $n \in \mathbb{N}_0$  then we have that

1. Given  $\langle \mathbb{R}^n, \langle \cdot, \cdot \rangle \rangle$  (see 11.93) then the basis  $\{\varepsilon\}_{i \in \{1, \dots, n\}}$  (see 10.164) is orthonormal

2. Given  $\langle \mathbb{C}^n, \langle \rangle \rangle$  (see 11.97) then the basis  $\{\varepsilon\}_{i \in \{1, \dots, n\}}$  (see 10.164) is orthonormal

**Proof.**

1. Given  $i, j \in \{1, \dots, n\}$  we have that  $\langle \varepsilon_i, \varepsilon_j \rangle = \sum_{k=1}^n (\varepsilon_i)_k \cdot (\varepsilon_j)_k = \sum_{k=1}^n \delta_{i,k} \cdot \delta_{j,k} = \delta_{i,j}$
2. Given  $i, j \in \{1, \dots, n\}$  we have that  $\langle \varepsilon_i, \varepsilon_j \rangle = \sum_{k=1}^n (\varepsilon_i)_k \cdot \overline{(\varepsilon_j)_k} = \sum_{k=1}^n \delta_{i,k} \cdot \overline{\delta_{j,k}} = \delta_{i,j}$   $\square$

Every orthonormal family is linear independent (the opposite is of course not always true)

**Theorem 11.117.** Let  $\langle X, \langle \rangle \rangle$  be a real (complex) inner product vector space then every orthonormal family  $\{e_i\}_{i \in I}$  is linear independent.

**Proof.** Let  $\{a_i\}_{i \in I}$  is a family in  $\mathbb{R}(\mathbb{C})$  with finite support such that  $0 = \sum_{i \in I} a_i \cdot e_i$  then  $\forall j \in I$  we have that  $0 = \langle 0, e_j \rangle = \langle \sum_{i \in I} a_i \cdot e_i, e_j \rangle$ . Now as  $\{a_i\}_{i \in I}$  has finite support and thus  $\{a_i \cdot e_i\}_{i \in I}$  has finite support there exists a  $n \in \mathbb{N}$  and a bijection  $\varphi: \{0, \dots, n\} \rightarrow I$  such that  $\sum_{i \in I} a_i \cdot e_i = \sum_{i=0}^n a_{\varphi(i)} \cdot e_{\varphi(i)}$  hence  $0 = \langle \sum_{i \in I} a_i \cdot e_i, e_j \rangle = \langle \sum_{i=0}^n a_{\varphi(i)} \cdot e_{\varphi(i)}, e_j \rangle \stackrel{11.98}{=} \sum_{i=0}^n \langle a_{\varphi(i)} \cdot e_{\varphi(i)}, e_j \rangle = \sum_{i=0}^n a_{\varphi(i)} \cdot \delta_{\varphi(i), j} = \sum_{i \in I} a_i \cdot \delta_{i,j} = \sum_{i \in I \setminus \{j\}} a_i \cdot \delta_{i,j} + \sum_{i \in \{j\}} a_i \cdot \delta_{i,j} = 0 + a_j = a_j$  proving that  $\forall j \in I$  we have  $a_j = 0$ .  $\square$

**Theorem 11.118.** Let  $\langle X, \langle \rangle \rangle$  be a real (complex) inner product space with a orthonormal basis  $\{e_i\}_{i \in \{1, \dots, n\}}$ ,  $L \in \text{Hom}(X)$  then  $\forall i, j \in \{1, \dots, n\}$   $\mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}})_{j,i} = \langle L(e_i), e_j \rangle$

**Proof.**

$$\begin{aligned} \langle L(e_i), e_j \rangle &= \left\langle \sum_{k=1}^n \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}})_{k,i} \cdot e_k, e_j \right\rangle \\ &= \sum_{k=1}^n \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}})_{k,i} \cdot \langle e_k, e_j \rangle \\ &= \sum_{k=1}^n \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}})_{k,i} \cdot \delta_{k,j} \\ &= \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}})_{j,i} \end{aligned}$$

$\square$

It turns out that every finite dimensional real (complex) inner vector space has a orthonormal basis, for this we use the Gram-Schmidt procedure.

**Theorem 11.119.** Let  $\langle X, \langle \rangle \rangle$  be a real (complex) inner vector space and  $\{e_i\}_{i \in \{1, \dots, n\}}$ ,  $n \in \mathbb{N}_0$  a linear independent family of vectors in  $X$  then if we define  $\{f_i\}_{i \in \{1, \dots, n\}}$  by  $f_1 = \frac{e_1}{\|e_1\|}$  and  $\forall i \in \{2, \dots, n\}$   $\hat{e}_i = \frac{e_i - \sum_{j=1}^{i-1} \langle e_i, \hat{e}_j \rangle \cdot \hat{e}_j}{\|e_i - \sum_{j=1}^{i-1} \langle e_i, \hat{e}_j \rangle \cdot \hat{e}_j\|}$  then  $\{\hat{e}_i\}_{i \in \{1, \dots, n\}}$  is well defined (no division by 0), orthonormal and  $\mathcal{S}(\{e_i | i \in \{1, \dots, j\}\}) = \mathcal{S}(\{\hat{e}_i | i \in \{1, \dots, j\}\})$  for  $j \in \{1, \dots, n\}$  (the spans are equal).

**Proof.** We will use induction to prove this so let  $\mathcal{F} = \{n \in \mathbb{N}_0 | \text{if } \{e_i\}_{i \in \{1, \dots, n\}} \text{ then } \{\hat{e}_i\}_{i \in \{1, \dots, n\}}$  is well defined, orthonormal and  $j \in \{1, \dots, n\}$   $\mathcal{S}(\{e_i | i \in \{1, \dots, j\}\}) = \mathcal{S}(\{\hat{e}_i | i \in \{1, \dots, j\}\})\}$  then we have

**1  $\in \mathcal{F}$ .** As  $\{e_i\}_{i \in \{1, \dots, 1\}}$  is linear independent we have that  $e_1 \neq 0 \Rightarrow \|e_1\| \neq 0$  so  $\hat{e}_1 = \frac{e_1}{\|e_1\|}$  is well defined  $\langle \hat{e}_1, \hat{e}_1 \rangle = \left\langle \frac{e_1}{\|e_1\|}, \frac{e_1}{\|e_1\|} \right\rangle = \frac{\langle e_1, e_1 \rangle}{\|e_1\|^2} = \frac{\|e_1\|^2}{\|e_1\|^2} = 1$  and  $\mathcal{S}(\{e_1\}) = \{\alpha \cdot e_1 | \alpha \in \mathbb{R}(\mathbb{C})\} = \left\{ \alpha \cdot \frac{e_1}{\|e_1\|} | \alpha \in \mathbb{R}(\mathbb{C}) \right\} = \mathcal{S}(\{\hat{e}_1\})$  proving that  $1 \in \mathcal{F}$

**$n \in \mathcal{F} \Rightarrow n+1 \in \mathcal{F}$ .** Let  $\{e_i\}_{i \in \{1, \dots, n+1\}}$  be a linear independent set, then as  $n \in \mathcal{F}$  we have that

$$\mathcal{S}(\{e_i | i \in \{1, \dots, n\}\}) = \mathcal{S}(\{\hat{e}_i | i \in \{1, \dots, n\}\}) \quad \forall j \in \{1, \dots, n\} \quad (11.21)$$

Now as  $\{e_i\}_{i \in \{1, \dots, n+1\}}$  is linear independent we have that  $e_{n+1} \notin \mathcal{S}(\{e_i | i \in \{1, \dots, n\}\}) = \mathcal{S}(\{\hat{e}_i | i \in \{1, \dots, n\}\})$  so that  $e_{n+1} \neq \sum_{k=1}^n \langle e_i, \hat{e}_k \rangle \cdot \hat{e}_k$  and thus that  $\|e_{n+1} - \sum_{k=1}^n \langle e_{n+1}, \hat{e}_k \rangle \cdot \hat{e}_k\|$  so that  $\hat{e}_{n+1}$  is well defined (no division by zero). Further from the definition of  $\hat{e}_{n+1}$  it follows that  $e_{n+1} = \|e_{n+1} - \sum_{k=1}^{n+1} \langle e_{n+1}, \hat{e}_k \rangle \cdot \hat{e}_k\| \cdot \hat{e}_{n+1} + \sum_{k=1}^n \langle e_{n+1}, \hat{e}_k \rangle \cdot \hat{e}_k$  proving that  $e_{n+1} \in \mathcal{S}(\{\hat{e}_i | i \in \{1, \dots, n+1\}\})$  hence as  $\{e_i | i \in \{1, \dots, n\}\} \subseteq \mathcal{S}(\{e_i | i \in \{1, \dots, n\}\}) = \mathcal{S}(\{\hat{e}_i | i \in \{1, \dots, n\}\})$  we have that  $\{e_i | i \in \{1, \dots, n+1\}\} \subseteq \mathcal{S}(\{\hat{e}_i | i \in \{1, \dots, n\}\}) \subseteq \mathcal{S}(\{\hat{e}_i | i \in \{1, \dots, n+1\}\})$  we have that  $\mathcal{S}(\{e_i | i \in \{1, \dots, n+1\}\}) \subseteq \mathcal{S}(\mathcal{S}(\{\hat{e}_i | i \in \{1, \dots, n+1\}\})) \stackrel{10.137}{=} \mathcal{S}(\{\hat{e}_i | i \in \{1, \dots, n+1\}\})$  proving

$$\mathcal{S}(\{e_i | i \in \{1, \dots, n+1\}\}) \subseteq \mathcal{S}(\{\hat{e}_i | i \in \{1, \dots, n+1\}\}) \quad (11.22)$$

Next for  $i, j \in \{1, \dots, n+1\}$  we have to consider the following cases

$i, j \in \{1, \dots, n\}$ . then as  $n \in \mathcal{F}$  we have  $\langle \hat{e}_i, \hat{e}_j \rangle = \delta_{i,j}$

$i = n+1, j \in \{1, \dots, n\}$ . then

$$\begin{aligned} \langle \hat{e}_i, \hat{e}_j \rangle &= \left\langle \frac{e_i - \sum_{k=1}^{i-1} \langle e_i, \hat{e}_k \rangle \cdot \hat{e}_k}{\|e_i - \sum_{k=1}^{i-1} \langle e_i, \hat{e}_k \rangle \cdot \hat{e}_k\|}, \hat{e}_j \right\rangle \\ &= \frac{1}{\|e_i - \sum_{k=1}^{i-1} \langle e_i, \hat{e}_k \rangle \cdot \hat{e}_k\|} \cdot \left( \langle e_i, \hat{e}_j \rangle - \sum_{k=1}^{i-1} \langle e_i, \hat{e}_k \rangle \cdot \langle \hat{e}_k, \hat{e}_j \rangle \right) \\ &\stackrel{n \in \mathcal{F}}{=} \frac{1}{\|e_i - \sum_{k=1}^{i-1} \langle e_i, \hat{e}_k \rangle \cdot \hat{e}_k\|} \cdot \left( \langle e_i, \hat{e}_j \rangle - \sum_{k=1}^{i-1} \langle e_i, \hat{e}_k \rangle \cdot \delta_{k,j} \right) \\ &= \frac{1}{\|e_i - \sum_{k=1}^{i-1} \langle e_i, \hat{e}_k \rangle \cdot \hat{e}_k\|} \cdot (\langle e_i, \hat{e}_j \rangle - \langle e_i, \hat{e}_j \rangle) \\ &= 0 \end{aligned}$$

$i \in \{1, \dots, n\} \wedge j = n+1$ . then  $\langle \hat{e}_i, \hat{e}_j \rangle = \overline{\langle \hat{e}_i, \hat{e}_j \rangle}$  previous case  $= 0$

$i = j = n+1$ . then

$$\begin{aligned} \langle \hat{e}_i, \hat{e}_j \rangle &= \langle \hat{e}_{n+1}, \hat{e}_{n+1} \rangle \\ &= \left\langle \frac{e_{n+1} - \sum_{k=1}^n \langle e_{n+1}, \hat{e}_k \rangle \cdot \hat{e}_k}{\|e_{n+1} - \sum_{k=1}^n \langle e_{n+1}, \hat{e}_k \rangle \cdot \hat{e}_k\|}, \frac{e_{n+1} - \sum_{k=1}^n \langle e_{n+1}, \hat{e}_k \rangle \cdot \hat{e}_k}{\|e_{n+1} - \sum_{k=1}^n \langle e_{n+1}, \hat{e}_k \rangle \cdot \hat{e}_k\|} \right\rangle \\ &= \frac{1}{\|e_{n+1} - \sum_{k=1}^n \langle e_{n+1}, \hat{e}_k \rangle \cdot \hat{e}_k\|^2} \cdot \left\langle e_{n+1} - \sum_{k=1}^n \langle e_{n+1}, \hat{e}_k \rangle \cdot \hat{e}_k, e_{n+1} - \sum_{k=1}^n \langle e_{n+1}, \hat{e}_k \rangle \cdot \hat{e}_k \right\rangle \\ &= \frac{\|e_{n+1} - \sum_{k=1}^{n-1} \langle e_{n+1}, \hat{e}_k \rangle \cdot \hat{e}_k\|^2}{\|e_{n+1} - \sum_{k=1}^n \langle e_{n+1}, \hat{e}_k \rangle \cdot \hat{e}_k\|^2} \\ &= 1 \end{aligned}$$

Proving that  $\{\hat{e}_i\}_{i \in \{1, \dots, n+1\}}$  is orthonormal. As we have just proved that  $\{\hat{e}_i\}_{i \in \{1, \dots, n+1\}}$  is orthonormal we have by 11.117 that  $\{\hat{e}_i\}_{i \in \{1, \dots, n+1\}}$  is a linear independent set and as  $\mathcal{S}(\{\hat{e}_i | i \in \{1, \dots, n+1\}\})$  is a vector space (see 10.132) we have that  $\mathcal{S}(\{\hat{e}_i | i \in \{1, \dots, n+1\}\})$  is a  $n+1$  dimensional vector space. As by 11.22 we have that  $\{e_i | i \in \{1, \dots, n+1\}\}$  is a linear independent subset of  $\mathcal{S}(\{\hat{e}_i | i \in \{1, \dots, n+1\}\})$  we have using 10.209 that  $\{e_i\}_{i \in \{1, \dots, n+1\}}$  is a basis for  $\mathcal{S}(\{\hat{e}_i | i \in \{1, \dots, n+1\}\})$  hence we have  $\mathcal{S}(\{e_i | i \in \{1, \dots, n+1\}\}) = \mathcal{S}(\{\hat{e}_i | i \in \{1, \dots, n+1\}\})$  and using 11.21 we conclude that

$$\forall j \in \{1, \dots, n+1\} \mathcal{S}(\{e_i | i \in \{1, \dots, n+1\}\}) = \mathcal{S}(\{\hat{e}_i | i \in \{1, \dots, n+1\}\}) \quad (11.23)$$

which finally proves that  $n+1 \in \mathcal{F}$  □

**Corollary 11.120.** Let  $\langle X, \langle \rangle \rangle$  be a finite dimensional inner vector space with dimension  $n \in \mathbb{N}_0$  (so a non trivial vector space) then we have a orthonormal basis  $\{e_i\}_{i \in \{1, \dots, n\}}$  on  $X$

**Proof.** As  $X$  is a  $n$ -dimensional vector space then there exists a linear independent family  $\{f_i\}_{i \in \{1, \dots, n\}}$  (the basis of  $X$ ) such that  $\mathcal{S}(\{f_i | i \in \{1, \dots, n\}\}) = X$ , using the Gram-Schmidt procedure we find a orthonormal family (linear independent by 11.117)  $\{e_i\}_{i \in \{1, \dots, n\}}$  such that  $\mathcal{S}(\{e_i | i \in \{1, \dots, n\}\}) = \mathcal{S}(\{f_i | i \in \{1, \dots, n\}\}) = X$ . So  $\{e_i\}_{i \in \{1, \dots, n\}}$  is a orthonormal basis of  $X$ .  $\square$

**Theorem 11.121.** Let  $\langle X, \langle \rangle \rangle$  be a finite dimensional inner vector space with dimension  $n \in \mathbb{N}_0$  and  $\{e_i\}_{i \in \{1, \dots, n\}}$  a orthonormal family then  $\{e_i\}_{i \in \{1, \dots, n\}}$  is a basis of  $X$

**Proof.** Let  $\{\alpha_i\}_{i \in \{1, \dots, n\}} \subseteq \mathbb{R}(\mathbb{C})$  such that  $\sum_{i=1}^n \alpha_i \cdot e_i = 0$  then  $\forall j \in \{1, \dots, n\}$  we have  $0 = \langle \sum_{i=1}^n \alpha_i \cdot e_i, e_j \rangle = \sum_{i=1}^n \alpha_i \cdot \langle e_i, e_j \rangle = \sum_{i=1}^n \alpha_i \cdot \delta_{i,j} = \alpha_j$ , hence we have that  $\{e_i\}_{i \in \{1, \dots, n\}}$  is linear independent. Using 10.205 proves then that  $\{e_i\}_{i \in \{1, \dots, n\}}$  forms a basis of  $X$ .  $\square$

**Theorem 11.122. (Riesz Representation Theorem)** Let  $\langle X, \langle \rangle \rangle$  be a finite dimensional real (complex) inner product space and  $\varphi \in \text{Hom}(X, \mathbb{R})$  (or  $\text{Home}(X, \mathbb{C})$ ) then there exists a **unique**  $u \in X$  such that  $\forall x \in X$  we have  $\varphi(x) = \langle x, u \rangle$

**Proof.** Using 11.120) we have that there exists a orthonormal basis  $\{e_i\}_{i \in \{1, \dots, n\}}$  of  $X$ . Hence if  $x \in X$  there exists a unique  $\{x_i\}_{i \in \{1, \dots, n\}}$  such that  $x = \sum_{i=1}^n x_i \cdot e_i$ . Then  $\varphi(x) = \varphi(\sum_{i=1}^n x_i \cdot e_i) = \sum_{i=1}^n x_i \cdot \varphi(e_i)$ . Take then  $u = \sum_{i=1}^n \varphi(e_i) \cdot e_i$  then

$$\begin{aligned} \langle x, u \rangle &= \left\langle x, \sum_{i=1}^n \varphi(e_i) \cdot e_i \right\rangle \\ &= \sum_{i=1}^n \varphi(e_i) \cdot \langle x, e_i \rangle \\ &= \sum_{i=1}^n \varphi(e_i) \cdot \left\langle \sum_{k=1}^n x_k \cdot e_k, e_i \right\rangle \\ &= \sum_{i=1}^n \varphi(e_i) \cdot \sum_{k=1}^n x_k \cdot \langle e_k, e_i \rangle \\ &= \sum_{i=1}^n \varphi(e_i) \cdot \left( \sum_{k=1}^n x_k \cdot \delta_{k,i} \right) \\ &= \sum_{i=1}^n x_i \cdot \varphi(e_i) \\ &= \varphi(x) \end{aligned}$$

proving the theorem.  $\square$

Riesz Representation Theorem allows us to define the adjoint of a operator

**Theorem 11.123.** Let  $\langle X, \langle \rangle_X \rangle, \langle Y, \langle \rangle_Y \rangle$  be finite dimensional real (complex) inner product spaces and  $L \in \text{Hom}(X, Y)$  then  $L^*: Y \rightarrow X$  is defined by  $L^*(y)$  is the unique number such that  $\langle L(x), y \rangle_Y = \langle x, L^*(y) \rangle_X \forall x \in X$ .  $L^*$  is called the adjoint of  $L$ .

**Proof.** First we have to prove that given  $y \in Y$  such a number  $L^*(y)$  exists. To do this given  $y \in Y$  define  $\varphi_y: X \rightarrow \mathbb{R}(\mathbb{C})$  by  $\varphi_y(x) = \langle L(x), y \rangle_Y$ . For  $x_1, x_2 \in X, \alpha, \beta \in \mathbb{R}(\mathbb{C})$  we have then that  $\varphi_y(\alpha \cdot x_1 + \beta \cdot y_1) = \langle L(\alpha \cdot x_1 + \beta \cdot y_1), y \rangle_Y = \langle \alpha \cdot L(x_1) + \beta \cdot L(x_2), y \rangle_Y = \alpha \cdot \langle L(x_1), y \rangle_Y + \beta \cdot \langle L(x_2), y \rangle_Y = \alpha \cdot \varphi_y(x_1) + \beta \cdot \varphi_y(x_2)$  proving that  $\varphi_y \in \text{Hom}(X, \mathbb{R})(\text{or } \text{Hom}(Y, \mathbb{C}))$ . Using the Reisz Representation Theorem (see 11.122) there exists then a **unique**  $L^*(y) \in X$  such that  $\forall x \in X$  such that  $\forall x \in X$  we have  $\varphi_y(x) = \langle x, L^*(y) \rangle_X$  or using the definition of  $\varphi_y(x)$  we have  $\langle x, L^*(y) \rangle_X = \langle L(x), y \rangle_Y$ .  $\square$

**Theorem 11.124.** Let  $\langle X, \langle \cdot \rangle_X \rangle, \langle Y, \langle \cdot \rangle_Y \rangle$  be finite dimensional real (complex) inner product spaces and  $L \in \text{Hom}(X, Y)$  then  $(L^*)^*: X \rightarrow Y$  is equal to  $L: X \rightarrow Y$

**Proof.** As  $L: X \rightarrow Y$  we have  $L^*: Y \rightarrow X$  and thus  $(L^*)^*: X \rightarrow Y$ . Let  $x \in X, y \in Y$  then we have two cases to consider

**$X$  is a real space.**

$$\begin{aligned} \langle y, (L^*)^*(x) \rangle_Y &= \langle L^*(y), x \rangle_X \\ &= \langle x, L^*(y) \rangle_X \\ &= \langle L(x), y \rangle_Y \\ &= \langle y, L(x) \rangle_X \end{aligned}$$

**$X$  is a complex space.**

$$\begin{aligned} \langle y, (L^*)^*(x) \rangle_Y &= \langle L^*(y), x \rangle_X \\ &= \overline{\langle x, L^*(y) \rangle_X} \\ &= \overline{\langle L(x), y \rangle_Y} \\ &= \langle y, L(x) \rangle_Y \end{aligned}$$

So we have in all cases that  $\langle y, (L^*)^*(x) \rangle_Y = \langle y, L(x) \rangle_Y$  proving by definition of the adjoint that  $(L^*)^* = L$   $\square$

The operator  $*$  is linear as proved in the following operator

**Theorem 11.125.** Let  $\langle X, \langle \cdot \rangle_X \rangle, \langle Y, \langle \cdot \rangle_Y \rangle$  be finite dimensional real (complex) inner product spaces,  $L_1, L_2 \in \text{Hom}(X, Y)$  and  $\alpha, \beta \in \mathbb{R}$  then  $(\alpha \cdot L_1 + \beta \cdot L_2)^* = \alpha \cdot L_1^* + \beta \cdot L_2^*$

**Proof.** Given  $x \in X, y \in Y$  then C

$$\begin{aligned} \langle x, (\alpha \cdot L_1 + \beta \cdot L_2)^*(y) \rangle &= \langle (\alpha \cdot L_1 + \beta \cdot L_2)(x), y \rangle \\ &= \langle \alpha \cdot L_1(x) + \beta \cdot L_2(x), y \rangle \\ &= \alpha \cdot \langle L_1(x), y \rangle + \beta \cdot \langle L_2(x), y \rangle \\ &= \alpha \cdot \langle x, L_1^*(y) \rangle + \beta \cdot \langle x, L_2^*(y) \rangle \\ &\stackrel{\alpha, \beta \in \mathbb{R}}{=} \langle x, \alpha \cdot L_1^*(y) + \beta \cdot L_2^*(y) \rangle \\ &= \langle x, \alpha \cdot L_1^*(y) + \beta \cdot L_2^*(y) \rangle \\ &= \langle x, (\alpha \cdot L_1^* + \beta \cdot L_2^*)(y) \rangle \end{aligned}$$

proving by the definition of the adjoint that  $(\alpha \cdot L_1 + \beta \cdot L_2)^* = \alpha \cdot L_1^* + \beta \cdot L_2^*$   $\square$

**Theorem 11.126.** Let  $\langle X, \langle \cdot \rangle_X \rangle, \langle Y, \langle \cdot \rangle_Y \rangle$  and  $\langle Z, \langle \cdot \rangle_Z \rangle$  be finite dimensional real (complex) inner product spaces and  $L_1 \in \text{Hom}(X, Y), L_2 \in \text{Hom}(Y, Z)$  then  $(L_2 \circ L_1): X \rightarrow Z$  has a adjoint  $(L_2 \circ L_1)^*: Z \rightarrow X$  defined by  $(L_2 \circ L_1)^* = L_1^* \circ L_2^*$

**Proof.** Let  $x \in X, z \in Z$  then

$$\begin{aligned} \langle x, (L_2 \circ L_1)^*(z) \rangle_X &= \langle (L_2 \circ L_1)(x), z \rangle_Z \\ &= \langle L_2(L_1(x)), z \rangle_Z \\ &= \langle L_1(x), L_2^*(z) \rangle_Y \\ &= \langle x, L_1^*(L_2^*(z)) \rangle_X \\ &= \langle x, (L_1^* \circ L_2^*)(z) \rangle \end{aligned}$$

which by definition means that  $(L_2 \circ L_1)^* = L_1^* \circ L_2^*$   $\square$

To see the relation between the adjoint of a linear operator and its matrix we define the concept of the transpose of a matrix and the conjugate transpose of a matrix.

**Definition 11.127.** Let  $n, m \in \mathbb{N}_0$ ,  $\mathcal{F}$  a field and  $M \in \mathcal{M}(n \times m, \mathcal{F})$  then  $M^T$  is defined by  $(M^T)_{i,j} = M_{j,i} \forall i, j \in \{1, \dots, n\} \times \{1, \dots, m\}$  (see 10.298)

**Definition 11.128.** Let  $n, m \in \mathbb{N}_0$ ,  $\mathcal{F} = \mathbb{R}$  or  $\mathcal{F} = \mathbb{C}$ ,  $M \in \mathcal{M}(n \times m, \mathcal{F})$  then  $M^H$  is defined by  $(M^H)_{i,j} = \overline{M_{j,i}} \forall i, j \in \{1, \dots, n\} \times \{1, \dots, m\}$ . Note that in the case  $\mathcal{F} = \mathbb{R}$  we have that  $M^H = M^T$

**Definition 11.129. (Symmetric Matrix)** Let  $n, m \in \mathbb{N}_0$ ,  $\mathcal{F}$  a field then  $M \in \mathcal{M}(n \times m, \mathcal{F})$  is symmetric if  $M^T = M$

**Definition 11.130. (Hermitian Matrix)** Let  $n, m \in \mathbb{N}_0$ ,  $\mathcal{F}$  a field then  $M \in \mathcal{M}(n \times m, \mathcal{F})$  is hermitian if  $M^H = M$ . Note that in the real case a symmetric matrix is Hermitian.

**Definition 11.131.** Let  $n \in \mathbb{N}_0$ ,  $\mathcal{F} = \mathbb{R}$  (or  $\mathbb{C}$ ) and  $M \in \mathcal{M}(n \times b, \mathcal{F})$  then  $M$  is **unitary** if  $U^H \cdot U = E = U \cdot U^H$

**Theorem 11.132.** Let  $\langle X, \langle \cdot \rangle_X \rangle$ ,  $\langle Y, \langle \cdot \rangle_Y \rangle$  be finite dimensional real (complex) inner product spaces and  $L \in \text{Hom}(X, Y)$  then  $L^*: Y \rightarrow X$  is linear hence  $L^* \in \text{Hom}(Y, X)$ . Further if  $\{e_i\}_{i \in \{1, \dots, n\}}$  is a orthonormal basis of  $X$  and  $\{f_i\}_{i \in \{1, \dots, m\}}$  is a orthonormal basis of  $Y$  then  $\forall (i, j) \in \{1, \dots, n\} \times \{1, \dots, m\}$  we have  $\mathcal{M}(L^*, \{f_k\}_{k \in \{1, \dots, m\}}, \{e_k\}_{k \in \{1, \dots, n\}})_{i,j} = \mathcal{M}(L, \{e_k\}_{k \in \{1, \dots, n\}}, \{f_k\}_{k \in \{1, \dots, m\}})_{j,i}^H \stackrel{\text{definition}}{=} \overline{\mathcal{M}(L, \{e_k\}_{k \in \{1, \dots, n\}}, \{f_k\}_{k \in \{1, \dots, m\}})_{j,i}}$  in the complex case and  $\mathcal{M}(L^*, \{f_k\}_{k \in \{1, \dots, m\}}, \{e_k\}_{k \in \{1, \dots, n\}})_{i,j} = \mathcal{M}(L, \{e_k\}_{k \in \{1, \dots, n\}}, \{f_k\}_{k \in \{1, \dots, m\}})_{j,i} =$  in the real case.

**Proof.** We have to make a distinction between the real and complex case.

**Real Case.** Then we have that  $\forall x \in X, \forall \alpha, \beta \in \mathbb{R}$  and  $y_1, y_2 \in Y$

$$\begin{aligned} \langle x, \alpha \cdot L^*(y_1) + \beta \cdot L^*(y_2) \rangle_X &= \alpha \cdot \langle x, L^*(y_1) \rangle + \beta \cdot \langle x, L^*(y_2) \rangle_X \\ &= \alpha \cdot \langle L(x), y_1 \rangle_Y + \beta \cdot \langle L(x), y_2 \rangle_Y \\ &= \langle L(x), \alpha \cdot y_1 + \beta \cdot y_2 \rangle_Y \\ &= \langle x, L^*(\alpha \cdot y_1 + \beta \cdot y_2) \rangle_X \end{aligned}$$

hence using 11.100 we have that  $L^*(y_1) + \beta \cdot L^*(y_2) = L^*(\alpha \cdot y_1 + \beta \cdot y_2)$  which proves linearity. Take now  $M = \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{1, \dots, m\}})$  and  $M^* = \mathcal{M}(L^*, \{f_i\}_{i \in \{1, \dots, m\}}, \{e_i\}_{i \in \{1, \dots, n\}})$  then by definition (see 10.302) we have for  $i \in \{1, \dots, n\}, j \in \{1, \dots, m\}$  that

$$\begin{aligned} \langle e_i, L^*(f_j) \rangle &\stackrel{10.302}{=} \left\langle e_i, \sum_{k=1}^n M_{k,j}^* \cdot e_k \right\rangle \\ &= \sum_{k=1}^n M_{k,j}^* \langle e_i, e_k \rangle \\ &= \sum_{k=1}^n M_{k,j}^* \cdot \delta_{i,k} \\ &= M_{i,j}^* \end{aligned}$$

Using the above we have then that

$$\begin{aligned} M_{i,j}^* &= \langle e_i, L^*(f_j) \rangle \\ &= \langle L(e_i), f_j \rangle \\ &\stackrel{10.302}{=} \left\langle \sum_{j=1}^m M_{k,i} \cdot f_k, f_j \right\rangle \\ &= \sum_{j=1}^m M_{k,i} \cdot \langle f_k, f_j \rangle \\ &= \sum_{j=1}^m M_{k,i} \cdot \delta_{k,j} \\ &= M_{j,i} \end{aligned}$$

proving that  $\mathcal{M}(L, \{e_k\}_{k \in \{1, \dots, n\}}, \{f_k\}_{k \in \{1, \dots, m\}})_{i,j} = \mathcal{M}(L^*, \{f_k\}_{k \in \{1, \dots, m\}}, \{e_k\}_{k \in \{1, \dots, n\}})_{j,i}$  where  $(i, j) \in \{1, \dots, n\} \times \{1, \dots, m\}$

**Complex Case.** Then we have that  $\forall x \in X, \forall \alpha, \beta \in \mathbb{C}$  and  $y_1, y_2 \in Y$

$$\begin{aligned} \langle x, \alpha \cdot L^*(y_1) + \beta \cdot L^*(y_2) \rangle_X &= \bar{\alpha} \cdot \langle x, L^*(y_1) \rangle + \bar{\beta} \cdot \langle x, L^*(y_2) \rangle_X \\ &= \bar{\alpha} \cdot \langle L(x), y_1 \rangle_Y + \bar{\beta} \cdot \langle L(x), y_2 \rangle_Y \\ &= \langle L(x), \alpha \cdot y_1 + \beta \cdot y_2 \rangle_Y \\ &= \langle x, L^*(\alpha \cdot y_1 + \beta \cdot y_2) \rangle_X \end{aligned}$$

hence using 11.100 we have that  $L^*(y_1) + \beta \cdot L^*(y_2) = L^*(\alpha \cdot y_1 + \beta \cdot y_2)$  which proves linearity. Take now  $M = \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{1, \dots, n\}})$  and  $M^* = \mathcal{M}(L^*, \{f_i\}_{i \in \{1, \dots, m\}}, \{e_i\}_{i \in \{1, \dots, n\}})$  then by definition (see 10.302) we have for  $i \in \{1, \dots, n\}, j \in \{1, \dots, m\}$  that

$$\begin{aligned} \langle L^*(f_j), e_i \rangle &\stackrel{10.302}{=} \left\langle \sum_{k=1}^m M_{k,j}^* \cdot e_k, e_i \right\rangle \\ &= \sum_{k=1}^m M_{k,j}^* \langle e_k, e_i \rangle \\ &= \sum_{k=1}^m M_{k,j}^* \cdot \delta_{k,i} \\ &= M_{i,j}^* \end{aligned}$$

Using the above we have then that

$$\begin{aligned} M_{i,j}^* &= \frac{\langle L^*(f_j), e_i \rangle}{\langle e_i, L^*(f_j) \rangle} \\ &= \frac{\langle L(e_i), f_j \rangle}{\overline{\left\langle \sum_{k=1}^m M_{k,i} \cdot f_k, f_j \right\rangle}} \\ &\stackrel{10.302}{=} \frac{\overline{\sum_{k=1}^m M_{k,i} \cdot \langle f_k, f_j \rangle}}{\overline{\sum_{k=1}^m M_{k,i}}} \\ &= \frac{\sum_{k=1}^m M_{k,i} \cdot \delta_{k,j}}{\overline{\sum_{k=1}^m M_{k,i}}} \\ &= \overline{M_{j,i}} \end{aligned}$$

proving that  $\forall (i, j) \in \{1, \dots, n\} \times \{1, \dots, m\}$  we have  $\mathcal{M}(L^*, \{f_k\}_{k \in \{1, \dots, m\}}, \{e_k\}_{k \in \{1, \dots, n\}})_{i,j} = \overline{\mathcal{M}(L, \{e_k\}_{k \in \{1, \dots, n\}}, \{f_k\}_{k \in \{1, \dots, m\}})_{j,i}}$   $\square$

**Corollary 11.133.** Let  $\langle X, \langle \rangle \rangle$  be a finite dimensional real inner product space,  $L \in \text{Hom}(X)$  then  $\det(L) = \det(L^*)$

**Proof.** Let  $\{e_i\}_{i \in \{1, \dots, n\}}$  be a basis in  $X$  then  $\det(L) \stackrel{10.321}{=} \det(\mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}})) \stackrel{10.320}{=} \det(\mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}})^T) = \det(\mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}})^H) \stackrel{11.132}{=} \det(L^*)$   $\square$

## 11.5 Continuity

**Definition 11.134.** Let  $\langle X, \mathcal{T}_X \rangle$  and  $\langle Y, \mathcal{T}_Y \rangle$  be topological spaces and let  $f: X \rightarrow Y$  be a function between  $X$  and  $Y$  then  $f$  is continuous at  $x$  iff  $\forall V \in \mathcal{T}_Y \vdash f(x) \in V$  we have  $\exists U \in \mathcal{T}_X \vdash x \in U \wedge f(U) \subseteq V$

**Definition 11.135.** Let  $\langle X, \mathcal{T}_X \rangle$  and  $\langle Y, \mathcal{T}_Y \rangle$  be topological spaces and let  $f: X \rightarrow Y$  be a function between  $X$  and  $Y$  then  $f$  is continuous if  $\forall x \in X$   $f$  is continuous at  $x$

**Theorem 11.136.** Let  $\langle X, \mathcal{T}_X \rangle$  and  $\langle Y, \mathcal{T}_Y \rangle$  be topological spaces with topological basis  $\mathcal{B}_X$  and  $\mathcal{B}_Y$  then for a function  $f: X \rightarrow Y$  we have the following equivalences:

1.  $f$  is continuous at  $x$
2.  $\forall V \in \mathcal{B}_Y \vdash f(x) \in V$  we have  $\exists U \in \mathcal{T}_X \vdash x \in U \wedge f(U) \subseteq V$
3.  $\forall V \in \mathcal{B}_Y \vdash f(x) \in V$  we have  $\exists U \in \mathcal{B}_X \vdash x \in U \wedge f(U) \subseteq V$

**Proof.**

1.  $(1 \Rightarrow 2)$  This is trivial as  $\mathcal{B}_Y \subseteq \mathcal{T}_Y$
2.  $(2 \Rightarrow 3)$  Let  $V \in \mathcal{B}_Y$  then by (2) there exists a  $U' \in \mathcal{T}_X$  with  $x \in U' \wedge f(U') \subseteq V$ , by the definition of a basis  $\exists U \in \mathcal{B}_X$  with  $x \in U \subseteq U' \Rightarrow f(U) \subseteq f(U') \subseteq V$
3.  $(3 \Rightarrow 1)$  Let  $V \in \mathcal{T}_Y$  with  $f(x) \in V$  then by the definition of a basis there exists a  $W \in \mathcal{B}_Y$  with  $f(x) \in W \subseteq V$  and thus there  $\exists U \in \mathcal{B}_X \subseteq \mathcal{T}_X$  such that  $x \in U \wedge f(U) \subseteq W \subseteq V$  proving continuity  $\square$

**Theorem 11.137.** Let  $\langle X, \mathcal{T}_X \rangle$  and  $\langle Y, \mathcal{T}_Y \rangle$  be topological spaces and let  $f: X \rightarrow Y$  be a continuous function between  $X$  and  $Y$  then if  $x \in A \subseteq X$  we have that  $f|_A: A \rightarrow Y$  is continuous at  $x$  (using the subspace topology of  $A$ ).

**Proof.** If  $V \in \mathcal{T}_Y$  with  $f_A(x) \in V$  then we have  $f(x) = f|_A(x) \in V$  and by continuity at  $x$  there exists a  $W \in \mathcal{T}_X$  with  $x \in W \wedge f(W) \subseteq V \Rightarrow x \in W \cap A \subseteq W \Rightarrow f(W \cap A) \subseteq f(W) \subseteq V$   $\square$

**Theorem 11.138.** Let  $\langle X, \mathcal{T}_X \rangle$  and  $\langle Y, \mathcal{T}_Y \rangle$  be topological spaces,  $f: X \rightarrow Y$ ,  $f(X) \subseteq A$  then  $f: X \rightarrow Y$  is continuous on  $X$  iff  $f: X \rightarrow A$  is continuous (using the subspace topology of  $A$ )

**Proof.** If  $f: X \rightarrow Y$  is continuous then if  $x \in U$  and  $V$  open in  $A$  with  $f(x) \in V$  then there exists a  $W \in \mathcal{T}_Y$  such that  $V = A \cap W \Rightarrow f(x) \in W$  so by continuity of  $f: X \rightarrow Y$  there exists a  $U \in \mathcal{T}_X$  with  $x \in U$  and  $f(U) \subseteq W \Rightarrow f(U) \cap A \subseteq W \cap A = V$  proving that  $f: X \rightarrow A$  is continuous in the subspace topology at  $x$  and as  $x$  is chosen arbitrary  $f: X \rightarrow A$  is continuous. If  $f: X \rightarrow A$  is continuous then if  $x \in X$  and  $V \in \mathcal{T}_Y$  such that  $f(x) \in V$  then as  $f(x) \in A$  we have  $f(x) \in V \cap A$  so by continuity of  $f: X \rightarrow A$  we have that  $\exists U \in \mathcal{T}_X$  with  $x \in U$  and  $f(U) \subseteq V \cap A \subseteq V$  proving that  $f: X \rightarrow Y$  is continuous at  $x$  and as  $x$  is chosen arbitrary it is  $f: X \rightarrow Y$  is continuous.  $\square$

**Definition 11.139.** Let  $\langle X, \mathcal{T}_X \rangle$  and  $\langle Y, \mathcal{T}_Y \rangle$  be topological spaces and let  $f: X \rightarrow Y$  then  $f$  is continuous if and only if  $\forall x \in X \vdash f$  is continuous at  $x$ . The set of continuous functions between  $X$  and  $Y$  is noted as  $\mathcal{C}(X, Y)$ .

**Example 11.140.** Let  $\langle X, \mathcal{T}_X \rangle$  and  $\langle Y, \mathcal{T}_Y \rangle$  be topological functions,  $c \in Y$  and  $C_{X,c}: X \rightarrow Y$  defined by  $C_c(x) = c$  is continuous

**Proof.**  $\forall x \in X$  we have that if  $c = C_{X,c}(x) \in V \in \mathcal{T}_Y$  then if  $y \in X \in \mathcal{T}_X$  we have  $C_{X,c}(y) = c \in V \Rightarrow C_{X,c}(X) \subseteq V$   $\square$

**Example 11.141.** Let  $\langle X, \mathcal{T} \rangle$  be a topological space then  $\mathbb{1}_X: X \rightarrow X$  defined by  $x \rightarrow x$  is continuous

**Proof.**  $\forall x \in X$  we have if  $x = \mathbb{1}_X(x) \in V \in \mathcal{T}$  then if  $y \in V$  then  $\mathbb{1}_X(y) = y \in V \Rightarrow \mathbb{1}_X(V) = V \subseteq V$   $\square$

**Theorem 11.142.** Let  $\langle X, \mathcal{T}_X \rangle$  and  $\langle Y, \mathcal{T}_Y \rangle$  be topological spaces and let  $f: X \rightarrow Y$  be a function between  $X$  and  $Y$  then the following are equivalent

1.  $f$  is continuous
2.  $\forall V \in \mathcal{T}_Y$  we have  $f^{-1}(V) \in \mathcal{T}_X$
3.  $\forall A \subseteq X$  we have  $f(\bar{A}) \subseteq \bar{f(A)}$

4.  $\forall F \subseteq Y \vdash F$  is closed we have that  $f^{-1}(F)$  is closed

**Proof.**

1. **(1  $\Rightarrow$  2)** Let  $V \in \mathcal{T}_Y$  then either  $f^{-1}(V) = \emptyset \in \mathcal{T}_X$  or  $\forall x \in f^{-1}(V)$  we have by continuity that  $\exists U_x \in \mathcal{T}_X \vdash x \in U_x \wedge f(U_x) \subseteq V \xrightarrow{2.58} U_x \subseteq f^{-1}(V) \Rightarrow f^{-1}(V) = \bigcup_{x \in f^{-1}(V)} U_x \in \mathcal{T}_X$
2. **(2  $\Rightarrow$  1)** If  $x \in X$  then if  $f(x) \in V \subseteq \mathcal{T}_Y$  then  $x \in U = f^{-1}(V) \in \mathcal{T}_X$  and  $f(U) \subseteq V$
3. **(1  $\Rightarrow$  3)** Let  $y \in f(\bar{A})$  then  $\exists x \in \bar{A}$  such that  $y = f(x)$ , if now  $y \in V \in \mathcal{T}_Y$  then by continuity  $\exists U \in \mathcal{T}(x)$  such that  $x \in U$  and  $f(U) \subseteq V$ , as  $x \in \bar{A}$  we have by 11.19 that  $U \cap A \neq \emptyset \Rightarrow \emptyset \neq f(U \cap A) \subseteq_{2.62} f(U) \cap f(A) \subseteq V \cap f(A) \Rightarrow y \in f(A)$
4. **(3  $\Rightarrow$  2 and thus  $\Rightarrow$  1)** Let  $A$  be a set in  $Y$  then as  $f(\bar{A}) \subseteq \overline{f(A)}$  we have by 2.58 that  $\bar{A} \subseteq f^{-1}(\overline{f(A)})$ . So if  $B \subseteq Y$  is closed we have  $\bar{B} = B$  and if we take  $A = f^{-1}(B)$  then  $\overline{f^{-1}(B)} = \bar{A} \subseteq f^{-1}(\overline{f(A)}) = f^{-1}(f(f^{-1}(B))) \subseteq \overline{f^{-1}(B)} = f^{-1}(B)$ , as we have also trivially  $f^{-1}(B) \subseteq \overline{f^{-1}(B)}$  it follows that  $f^{-1}(B) = \overline{f^{-1}(B)}$  and thus  $f^{-1}(B)$  is closed. Now if  $V \in \mathcal{T}_Y$  then  $Y \setminus V$  is closed and as  $X \setminus f^{-1}(V) \stackrel{2.58}{=} f^{-1}(Y \setminus V)$  is closed we have that  $f^{-1}(V)$  is open.
5. **(2  $\Rightarrow$  4)** If  $F$  is closed then we have that  $Y \setminus F$  is open and by (2) we have that  $f^{-1}(Y \setminus U)$  is open so that  $X \setminus f^{-1}(Y \setminus F)$  is closed. Now  $X \setminus f^{-1}(Y \setminus F) \stackrel{2.58}{=} f^{-1}(Y \setminus (Y \setminus F)) = f^{-1}(F)$  so that  $f^{-1}(F)$  is closed.
6. **(4  $\Rightarrow$  2)** If  $U$  is open in  $Y$  then  $Y \setminus U$  is closed and by (4) we have that  $f^{-1}(Y \setminus U)$  is closed and thus  $X \setminus f^{-1}(Y \setminus U)$  is open. Finally  $X \setminus f^{-1}(Y \setminus U) \stackrel{2.58}{=} f^{-1}(Y \setminus (Y \setminus U)) = f^{-1}(U)$  proving that  $f^{-1}(U)$  is open.  $\square$

We prove now that continuity is essential a local feature

**Theorem 11.143.** Let  $\langle X, \mathcal{T}_X \rangle, \langle Y, \mathcal{T}_Y \rangle$  be topological spaces and  $f: X \rightarrow Y$  a function between  $X \rightarrow Y$  then  $f$  is continuous iff  $\forall x \in X$  there exists a  $U \in \mathcal{T}_X$  with  $x \in U$  such that  $f|_U: U \rightarrow Y$  is continuous in the subspace topology of  $U$

**Proof.**

1.  $(\Rightarrow)$  This follows from 11.137.
2.  $(\Leftarrow)$  Let  $x \in X$  then there exists a  $U \in \mathcal{T}_X$  with  $x \in U$  such that  $f|_U: U \rightarrow Y$  is continuous. So if  $V$  is open in  $Y$  containing  $f|_U(x) = f(x)$  then there exists a  $W'$  open in  $U$  such that  $f(W') = f|_U(W') \subseteq V$ , as  $W' \stackrel{\text{definition of subspace topology}}{=} W \cap U$  is open in  $X$  (as  $W, U$  are open) this means that  $f$  is continuous at  $x$  and as  $x \in X$  was chosen arbitrary we have that  $f$  is continuous.  $\square$

**Definition 11.144.** Let  $\langle X, \mathcal{T}_X \rangle, \langle Y, \mathcal{T}_Y \rangle$  be topological spaces then a function  $f: X \rightarrow Y$  is open iff  $\forall U \in \mathcal{T}_X$  we have that  $f(U) \in \mathcal{T}_Y$  (every image of a open set is open)

**Theorem 11.145.** Let  $\langle X, \mathcal{T}_X \rangle, \langle Y, \mathcal{T}_Y \rangle$  be topological spaces then for a open injective function  $f: X \rightarrow Y$  we have that  $f^{-1}: f(X) \rightarrow X$  is continuous (see 2.21)

**Proof.** Let  $U$  be open in  $X$  then as  $f$  is open we have that  $f(U)$  is open and  $(f^{-1})^{-1}(U) \stackrel{2.58}{=} f(U)$  is open in  $X$  and as  $f(U) \subseteq f(X)$  we have that  $f(U) \cap f(X) = f(U)$  so that  $f(U) = (f^{-1})^{-1}(U)$  is open in  $f(X)$  proving continuity of  $f^{-1}: f(X) \rightarrow X$   $\square$

**Theorem 11.146.** Let  $\langle X, \mathcal{T}_X \rangle, \langle Y, \mathcal{T}_Y \rangle, \langle Z, \mathcal{T}_Z \rangle$  be topological spaces and  $f: X \rightarrow Y, g: Y \rightarrow Z$  be open functions then  $g \circ f: X \rightarrow Z$  is a open function.

**Proof.** This is trivial for if  $U \in \mathcal{T}_X$  then  $f(U)$  is open and thus  $g(f(U)) = (g \circ f)(U)$  is open.  $\square$

**Theorem 11.147.** Let  $\langle X, \mathcal{T}_X \rangle, \langle Y, \mathcal{T}_Y \rangle, \langle Z, \mathcal{T}_Z \rangle$  be topological spaces and  $f: X \rightarrow Y, g: Y \rightarrow Z$  be continuous functions then  $g \circ f: X \rightarrow Z$  is continuous

**Proof.** And let  $W$  be open in  $Z$  then by continuity of  $g$  we have that  $g^{-1}(W)$  is open and thus by continuity of  $f$  we have that  $(g \circ f)^{-1}(W) = f^{-1}(g^{-1}(W))$  is open.  $\square$

**Theorem 11.148.** Let  $\{\langle X_i, \mathcal{T}_i \rangle\}_{i \in I}$  be a family of topological spaces and let  $\langle \prod_{i \in I} X_i, \mathcal{T} \rangle$  be the product of the family equipped with the product topology (see 11.36) then  $\forall i \in I$  we have that the projection map  $\pi_i: \prod_{j \in I} X_j \rightarrow X_i$  defined by  $\pi_i(x_1, \dots, x_n) = x_i$  is open and continuous

**Proof.**

- I. **(continuity)** If  $U \in \mathcal{T}_i$  then by the definition of the product topology  $\pi_i^{-1}(U)$  is a element of the generating sub-basis, thus a element of the basis and finally a element of the topology.
- II. **(open)** Let  $U$  be open in the product topology and let  $t \in \pi_i(U)$  then  $\exists x \in U$  such that  $t = \pi_i(x) = x(i) = x_i$ , as  $x \in U$  open we use 11.37 to find a family  $\{A_i^{(t)}\}_{i \in I}$  and a finite  $A \subseteq I$  such that  $\forall j \in I$  we have  $A_j^{(t)} = \begin{cases} A_j^{(t)} \text{ is open if } j \in A \\ A_j^{(t)} = X_j \text{ if } j \in I \setminus A \Rightarrow A_j^{(t)} \text{ is open} \end{cases}$  so that  $x \in \prod_{i \in I} A_i^{(t)} \subseteq U$  so that  $t = \pi_i(x) \subseteq \pi_i(\prod_{i \in I} A_i^{(t)}) \subseteq A_i^{(t)} \subseteq \pi_i(U)$  where  $A_i$  is open. Using 11.5 we have then that  $\pi_i(U)$  is open.  $\square$

**Theorem 11.149.** Let  $\langle X, \mathcal{T}_X \rangle$  be a topological space,  $\{X_i, \mathcal{T}_i\}_{i \in I}$  be a family of topological spaces and  $\langle \prod_{i \in I} X_i, \mathcal{T} \rangle$  where  $\mathcal{T}$  is the product topology. A function  $f: X \rightarrow \prod_{i \in I} X_i$  is continuous iff  $\forall i \in I$  we have that  $f_i = \pi_i \circ f$  is continuous

**Proof.**

- 1.  $(\Rightarrow)$  As  $f$  is continuous and  $\forall i \in I$  we have  $\pi_i$  is continuous (see 11.148) we have by 11.147 that  $f_i = \pi_i \circ f$  is continuous.
- 2.  $(\Leftarrow)$  Let  $x \in X$  and  $f(x) \in V$  which is open in the product topology (see 11.36), then by the definition of the product topology there exists a finite family  $\{U_i^{(x)}\}_{i \in J}$  of open sets such that  $f(x) \in \bigcap_{i \in J} \pi_i^{-1}(U_i^{(x)}) \subseteq V \Rightarrow x \in f^{-1}(\bigcap_{i \in J} \pi_i^{-1}(U_i^{(x)})) \subseteq f^{-1}(V) \xrightarrow{2.62} x \in \bigcap_{i \in J} f^{-1}(\pi_i^{-1}(U_i^{(x)})) = \bigcap_{i \in J} (\pi_i \circ f)^{-1}(U_i^{(x)}) \subseteq f^{-1}(V)$ . Now as  $\pi_i \circ f$  is continuous we have that  $W_x = \bigcap_{i \in J} (\pi_i \circ f)^{-1}(U_i^{(x)})$  is open and  $x \in W_x \subseteq f^{-1}(V)$ . Using 11.5 we have then that  $f^{-1}(V)$  is open.  $\square$

**Theorem 11.150.** Let  $\langle X_1, \mathcal{T}_1 \rangle$ ,  $\langle X_2, \mathcal{T}_2 \rangle$ ,  $\langle Y, \mathcal{T}_Y \rangle$  be topological spaces,  $\langle X_1 \times X_2, \mathcal{T} \rangle$  where  $\mathcal{T}$  is the product topology (= the box topology by 11.39). If now  $f: X_1 \times X_2 \rightarrow Y$  is a continuous function then the following functions are also continuous:

1.  $\forall x \in X_1 f_1(x): X_2 \rightarrow Y$  defined by  $t \rightarrow f_1(x)(t) = f(x, t)$
2.  $\forall x \in X_2 f_2(x): X_1 \rightarrow Y$  defined by  $t \rightarrow f_2(x)(t) = f(t, x)$

**Proof.**

- 1. Let  $V \subseteq Y$  be a open set then by continuity of  $f$  we have that  $f^{-1}(V)$  is open in  $X_1 \times X_2$ . Take now  $y \in f_1(x)^{-1}(V)$  then  $f(x, y) = f_1(x)(y) \in V \Rightarrow (x, y) \in f^{-1}(V) \xrightarrow{f^{-1}(V) \text{ is open product topology is box topology}} \exists U \in \mathcal{T}_1 \wedge \exists W \in \mathcal{T}_2$  such that  $(x, y) \in U \times W \subseteq f^{-1}(V) \Rightarrow x \in U \wedge y \in W$ . Now if  $t \in W \Rightarrow (x, t) \in U \times W \subseteq f^{-1}(V) \Rightarrow f_1(x)(t) = f(x, t) \in V \Rightarrow t \in f_1(x)^{-1}(V) \Rightarrow W \subseteq f_1(x)^{-1}(V)$ . So we have that  $y \in W \subseteq f_1(x)^{-1}(V)$  which by 11.5 proves that  $f_1(x)^{-1}(V)$  is open.
- 2. Let  $V \subseteq Y$  be a open set then by continuity of  $f$  we have that  $f^{-1}(V)$  is open in  $X_1 \times X_2$ . Take now  $y \in f_2(x)^{-1}(V)$  then  $f(y, x) = f_2(x)(y) \in V \Rightarrow (y, x) \in f^{-1}(V) \xrightarrow{f^{-1}(V) \text{ is open product topology is box topology}} \exists U \in \mathcal{T}_1 \wedge \exists W \in \mathcal{T}_2$  such that  $(y, x) \in U \times W \subseteq f^{-1}(V) \Rightarrow y \in U \wedge x \in W$ . Now if  $t \in U \Rightarrow (t, x) \in U \times W \subseteq f^{-1}(V) \Rightarrow f_2(x)(t) = f(t, x) \in V \Rightarrow t \in f_2(x)^{-1}(V) \Rightarrow U \subseteq f_2(x)^{-1}(V)$ . So we have that  $y \in U \subseteq f_2(x)^{-1}(V)$  which by 11.5 proves that  $f_2(x)^{-1}(V)$  is open.  $\square$

We now prove the condition to have a continuous function in metric spaces. From now on we assume that for the metric spaces we use the metric topology.

**Theorem 11.151.** *Let  $\langle X, d_X \rangle$ ,  $\langle Y, d_Y \rangle$  be metric spaces and  $f: X \rightarrow Y$  a function then  $f$  is continuous in the metric topology at  $x \in X$  iff  $\forall \varepsilon \in \mathbb{R}_+$  we have that  $\exists \delta \in \mathbb{R}_+$  such that  $\forall y \in X \mid d_X(x, y) < \delta$  we have  $d_Y(f(x), f(y)) < \varepsilon$*

**Proof.**

1.  $(\Rightarrow)$  If  $f$  is continuous at  $x$  then if  $\varepsilon \in \mathbb{R}_+$  we have that  $f(x) \in B_{d_Y}(f(x), \varepsilon)$  which is open so there exists a  $U$  open in  $X$  such that  $x \in U$  and  $f(U) \subseteq B_{d_Y}(f(x), \varepsilon)$ . As  $U$  is open in the metric topology there exists a  $\delta \in \mathbb{R}_+$  such that  $x \in B_{d_X}(x, \delta) \subseteq U \Rightarrow f(B_{d_X}(x, \delta)) \subseteq f(U) \subseteq B_{d_Y}()$
2.  $(\Leftarrow)$  Let  $x \in X$  then to prove continuity we can use 11.136 so if  $\varepsilon \in \mathbb{R}_+$  and  $f(x) \in B_{d_Y}(f(x), \varepsilon)$ . Now by the hypothesis there exists a  $\delta \in \mathbb{R}_+$  such that if  $d_X(x, y) < \delta \Rightarrow d_Y(f(x), f(y)) < \varepsilon$  or if  $y \in B_{d_X}(x, \delta) \Rightarrow d_X(x, y) < \delta \Rightarrow d_Y(f(x), f(y)) \Rightarrow f(y) \in B_{d_Y}(f(x), \varepsilon) \Rightarrow f(B_{d_X}(x, \delta)) \subseteq B_{d_Y}(y, \delta)$   $\square$

**Corollary 11.152.** *Let  $\langle X, d_X \rangle$ ,  $\langle Y, d_Y \rangle$  be metric space,  $A \subseteq X$  then  $f: A \rightarrow Y$  is continuous (using the subspace topology on  $A$  based on the metric topology of  $X$ ) at  $x \in A$  iff  $\forall \varepsilon \in \mathbb{R}_+$  there exists a  $\delta \in \mathbb{R}_+$  such that  $\forall y \in A$  with  $d_X(x, y) < \delta$  we have  $d_Y(f(x), f(y))$*

**Proof.** Using 11.57 we have that the sub space topology is generated by the restricted metric  $d_{X|A \times A}: A \times A \rightarrow \mathbb{R}$  defined by  $d_{X|A \times A}(x, y) = d_A(x, y)$ . Using the previous theorem we have then that  $f: A \rightarrow Y$  is continuous iff  $\forall \varepsilon \in \mathbb{R}_+$  there exists a  $\delta \in \mathbb{R}_+$  such that  $\forall y \in A$  with  $d_{X|A \times A}(x, y) < \varepsilon$  we have  $d_Y(f(x), f(y)) < \varepsilon$  which as  $d_{X|A \times A}(x, y) = d_X(x, y)$  is equivalent with  $\forall \varepsilon \in \mathbb{R}_+$  there exists a  $\delta \in \mathbb{R}_+$  such that  $\forall y \in A$  with  $d_X(x, y) < \varepsilon$  we have  $d_Y(f(x), f(y)) < \varepsilon$   $\square$

**Theorem 11.153.** *Let  $\langle X, d_X \rangle$  and  $\langle Y, d_Y \rangle$  be metric spaces and  $f: X \rightarrow Y$  a function then  $f$  is open if and only if  $\forall x \in X, \forall \delta > 0$  we have that there exists a  $\delta' > 0$  such that  $f(x) \in B_{d_Y}(f(x), \delta') \subseteq f(B_{d_X}(x, \delta))$*

**Proof.**

1.  $(\Rightarrow)$  Because  $f$  is open we have that if  $x \in B_{d_X}(x, \delta)$  then  $f(B_{d_X}(x, \delta))$  is open and thus as  $f(x) \in f(B_{d_X}(x, \delta))$  there exists a  $\delta'$  such that  $f(x) \in B_{d_Y}(f(x), \delta') \subseteq f(B_{d_X}(x, \delta))$
2.  $(\Leftarrow)$  Let  $U \subseteq X$ ,  $U$  open and let  $y \in f(U)$  then there exists a  $x \in U$  such that  $f(x) = y$  then by the hypothesis there exists a  $\delta_y$  with  $y = f(x) \in B_{d_Y}(f(x), \delta_y) = B_{d_Y}(y, \delta_y) \subseteq f(B_{d_X}(x, \delta)) \subseteq f(U) \Rightarrow f(U) = \bigcup_{y \in f(U)} B_{d_Y}(y, \delta_y)$  and thus  $f(U)$  is open  $\square$

**Definition 11.154.** *Let  $\langle X, d_X \rangle$ ,  $\langle Y, d_Y \rangle$  be metric spaces then a function  $f: X \rightarrow Y$  is **uniform continuous** in  $K \subseteq X$  iff  $\forall \varepsilon \in \mathbb{R}_+$  we have that  $\exists \delta \in \mathbb{R}_+$  such that  $\forall x, y \in K$  with  $d_X(x, y) < \varepsilon$  we have that  $d_Y(f(x), f(y)) < \varepsilon$*

**Theorem 11.155.** *Let  $\langle X, \|\cdot\|_X \rangle$ ,  $\langle Y, \|\cdot\|_Y \rangle$  be normed spaces then a function  $f: X \rightarrow Y$  is continuous at  $x$  in the norm topologies iff  $\forall \varepsilon \in \mathbb{R}_+$  there exists a  $\delta \in \mathbb{R}_+$  such that if  $\|x - y\|_X < \delta$  then  $\|f(x) - f(y)\|_Y < \varepsilon$*

**Proof.** This follows from 11.151 and the fact that  $d_X(x, y) = \|x - y\|_X$  and  $d_Y(f(x), f(y)) = \|f(x) - f(y)\|$   $\square$

**Example 11.156.** Let  $\langle \mathbb{C}, \|\cdot\| \rangle$  the normed space using the complex norm and  $\langle \mathbb{R}, \|\cdot\| \rangle$  the normed space using the absolute value.

1.  $\text{Re}: \mathbb{C} \rightarrow \mathbb{R}$  defined by  $x \rightarrow \text{Re}(x)$  is continuous
2.  $\text{Img}: \mathbb{C} \rightarrow \mathbb{R}$  defined by  $x \rightarrow \text{Img}(x)$  is continuous

**Proof.** Take  $z \in \mathbb{C}$  and take

- Given  $\varepsilon \in \mathbb{R}_+$  take  $\delta = \varepsilon$  then if  $|z - y| < \varepsilon$  we have

$$|\operatorname{Re}(z) - \operatorname{Re}(y)| = |\operatorname{Re}(z - y)| <_{9.33} |z - y| < \varepsilon$$

- Given  $\varepsilon \in \mathbb{R}_+$  take  $\delta = \varepsilon$  then if  $|z - y| < \varepsilon$  we have

$$|\operatorname{Img}(z) - \operatorname{Img}(y)| = |\operatorname{Img}(z - y)| <_{9.33} |z - y| < \varepsilon$$

□

We can extend the above theorem to a subspace of a normed space.

**Theorem 11.157.** Let  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle$  be normed spaces,  $A \subseteq X$   $A \subseteq X$  then  $f: A \rightarrow Y$  is continuous (using the subspace topology on  $A$  based on the normed topology of  $X$ ) at  $x \in A$  iff  $\forall \varepsilon \in \mathbb{R}_+$  there exists a  $\delta \in \mathbb{R}_+$  such that  $\forall y \in A$  with  $\|x - y\|_X < \delta$  we have  $\|f(x) - f(y)\| < \varepsilon$

**Proof.** This follows from 11.152 and the fact that  $d_X(x, y) = \|x - y\|_X$  and  $d_Y(f(x), f(y)) = \|f(x) - f(y)\|$  □

**Definition 11.158.** Let  $\langle X, \|\cdot\|_X \rangle$  and  $\langle Y, \|\cdot\|_Y \rangle$  be normed spaces then a function  $f: X \rightarrow Y$  is **Lipschitz continuous** if  $\forall x, y \in X$  we have  $\|f(x) - f(y)\|_Y \leq \|x - y\|_X$

**Theorem 11.159.** Let  $\langle X, \|\cdot\|_X \rangle$  and  $\langle Y, \|\cdot\|_Y \rangle$  be normed spaces then a Lipschitz continuous function  $f: X \rightarrow Y$  is uniform continuous and thus continuous.

**Proof.** Let  $\varepsilon > 0$  take  $d_{\|\cdot\|_X}(x, y) < \varepsilon$ , then we have  $d_{\|\cdot\|_Y}(f(x) - f(y)) = \|f(x) - f(y)\|_Y \leq \|x - y\|_X = d_{\|\cdot\|_X}(x, y) < \varepsilon \Rightarrow d_{\|\cdot\|_Y}(f(x), f(y))$  □

**Definition 11.160.** Let  $\langle X, \|\cdot\| \rangle$  be a normed space,  $\emptyset \neq S \subseteq X$  a non-empty set then the **set distance function**  $\delta_S: X \rightarrow \mathbb{R}$  is defined by  $x \rightarrow \delta_S(x) = \inf(\{\|x - y\| \mid y \in S\})$ . Note that if  $x \in S$  then  $\delta_S(x) = 0$ .

**Theorem 11.161.** Let  $\langle X, \|\cdot\| \rangle$  be a normed space,  $\emptyset \neq S \subseteq X$  a non-empty set then  $\delta_S: X \rightarrow \mathbb{R}$  is Lipschitz continuous and thus continuous.

**Proof.** Let  $x, y \in X, \varepsilon > 0$  then there exists a  $y_\varepsilon \in S$  such that  $\inf(\{\|x - y\| \mid y \in S\}) \leq \|x - y_\varepsilon\| < \inf(\{\|x - y\| \mid y \in S\}) + \varepsilon \Rightarrow \delta_S(x) \leq \|x - y_\varepsilon\| < \delta_S(x) + \varepsilon$ . Then  $\delta_S(y) - \delta_S(x) \leq \|y - y_\varepsilon\| - \|x - y_\varepsilon\| + \varepsilon \leq \|y - x\| + \|x - y_\varepsilon\| - \|x - y_\varepsilon\| + \varepsilon = \|y - x\| + \varepsilon \Rightarrow \delta_S(y) - \delta_S(x) \leq \|y - x\| + \varepsilon$ . If now  $\|y - x\| < \delta_S(y) - \delta_S(x)$  take then  $\varepsilon = \frac{\delta_S(y) - \delta_S(x) - \|y - x\|}{2} > 0$  then we have  $\delta_S(y) - \delta_S(x) \leq \|y - x\| + \varepsilon \leq \|y - x\| + \frac{\delta_S(y) - \delta_S(x) - \|y - x\|}{2} < \|y - x\| + \delta_S(y) - \delta_S(x) - \|y - x\| = \delta_S(y) - \delta_S(x)$  giving the contradiction  $\delta_S(y) - \delta_S(x) < \delta_S(y) = \delta_S(x)$ , so we must have

$$\delta_S(y) - \delta_S(x) \leq \|y - x\|$$

Exchange  $x, y$  gives then also

$$\delta_S(x) - \delta_S(y) \leq \|x - y\| = \|y - x\|$$

proving that  $|\delta_S(y) - \delta_S(x)| \leq \|y - x\|$  □

**Theorem 11.162.** Let  $\langle X, \|\cdot\| \rangle$  be a normed space then  $+: X \times X \rightarrow X$  defined by  $(x, y) \rightarrow x + y$ ,  $\cdot: X \times \mathbb{R}(\mathbb{C}) \rightarrow X$  defined by  $(x, \alpha) \rightarrow \alpha \cdot x$  and  $\|\cdot\|: X \rightarrow \mathbb{R}$  are continuous in the norm topology and product topologies of  $X \times X$  and  $X \times \mathbb{R}(\mathbb{C})$

**Proof.**

- (+) Let  $x + y \in U$  (open in the norm topology) then  $\exists \varepsilon \in \mathbb{R}_+$  such that  $x + y \in B_{d_{\|\cdot\|}}(x + y, \varepsilon) \subseteq U$ . Take now  $z' = (x', y') \in B_{d_{\|\cdot\|}}(x, \frac{\varepsilon}{2}) \times B_{d_{\|\cdot\|}}(y, \frac{\varepsilon}{2})$  (which is open in  $X \times X$ ) then  $\|x + y - (x' + y')\| = \|(x - x') + (y - y')\| \leq \|x - x'\| + \|y - y'\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \Rightarrow x' + y' \in B_{d_{\|\cdot\|}}(x + y, \varepsilon) \subseteq U \Rightarrow +\left(B_{d_{\|\cdot\|}}(x, \frac{\varepsilon}{2}) \times B_{d_{\|\cdot\|}}(y, \frac{\varepsilon}{2})\right) \subseteq U$  proving that  $+$  is continuous at  $(x, y)$ . As  $(x, y) \in X \times X$  has been chosen arbitrary we have that  $+$  is continuous.

2. (.) Let  $\lambda \cdot x \in U$  (open in  $X$ ) then there exists a  $\varepsilon \in \mathbb{R}_+$  such that  $\lambda \cdot x \in B_{d_{\|\cdot\|}}(\lambda \cdot x, \varepsilon) \subseteq U$ . Let know  $z' = (x', \lambda') \in B_{d_{\|\cdot\|}}(x, \delta_1) \times B_{d_{\|\cdot\|}}(\lambda, \delta_2)$  which is open in  $X \times \mathbb{R}(\mathbb{C})$  then  $\|\lambda \cdot x - \lambda' \cdot x'\| = \|\lambda \cdot x - \lambda' \cdot x + \lambda' \cdot x - \lambda' \cdot x'\| \leq \|\lambda \cdot x - \lambda' \cdot x\| + \|\lambda' \cdot x - \lambda' \cdot x'\| = |\lambda - \lambda'| \cdot \|x\| + |\lambda'| \cdot \|x - x'\| < \delta_2 \cdot \|x\| + |\lambda'| \cdot \delta_1 = \delta_2 \cdot \|x\| + \delta_1 \cdot |\lambda' - \lambda + \lambda| \leq \delta_2 \cdot \|x\| + \delta_1 \cdot (|\lambda' - \lambda| + |\lambda|) < \delta_2 \cdot \|x\| + \delta_1 \cdot (\delta_2 + |\lambda|)$  giving

$$\|\lambda \cdot x - \lambda' \cdot x'\| < \delta_2 \cdot \|x\| + \delta_1 \cdot (\delta_2 + |\lambda|) \quad (11.24)$$

consider now the following cases

a. ( $\|x\| = 0$ ) take then  $\delta_2 = 1$  and  $\delta_1 = \frac{\varepsilon}{1+|\lambda|}$  then  $\delta_2 \cdot \|x\| + \delta_1 \cdot (\delta_2 + |\lambda|) = \frac{\varepsilon}{1+|\lambda|} \cdot (1 + |\lambda|) = \varepsilon \Rightarrow \|\lambda \cdot x - \lambda' \cdot x'\| < \varepsilon$

b. ( $\|x\| \neq 0$ ) take then  $\delta_1 = \frac{\varepsilon}{2 \cdot (1+|\lambda|)} \wedge \delta_2 = \min\left(1, \frac{\varepsilon}{2 \cdot \|x\|}\right)$  then  $\delta_2 \cdot \|x\| + \delta_1 \cdot (\delta_2 + |\lambda|) \leq \frac{\varepsilon}{2 \cdot \|x\|} \cdot \|x\| + \frac{\varepsilon}{2 \cdot (1+|\lambda|)} \cdot (1+|\lambda|) = \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \Rightarrow \|\lambda \cdot x - \lambda' \cdot x'\| < \varepsilon$

So with the correct choice of  $\delta_1, \delta_2$  we have  $\|\lambda \cdot x - \lambda' \cdot x'\| < \varepsilon \Rightarrow \lambda' \cdot x' \in B_{d_{\|\cdot\|}}(\lambda \cdot x, \varepsilon) \subseteq U$  if  $(x', \lambda') \in B_{d_{\|\cdot\|}}(x, \delta_1) \cdot B_{d_{\|\cdot\|}}(\lambda, \delta_2)$  or  $(B_{d_{\|\cdot\|}}(x, \delta_1) \cdot B_{d_{\|\cdot\|}}(\lambda, \delta_2)) \subseteq U$  proving that  $\cdot$  is continuous at  $(x, \lambda)$ . As  $(x, \lambda)$  was chosen arbitrary we have that  $\cdot$  is continuous.

3. ( $\|\cdot\|$ ) Let  $x \in X$  and let  $\varepsilon \in \mathbb{R}_+$  then if  $\|x - y\| < \varepsilon \Rightarrow \|\|x\| - \|y\|\| \leq \|x - y\| < \varepsilon$  proving by 11.155 that  $\|\cdot\|$  is continuous at  $x$ . As  $x$  is chosen arbitrary we have that  $\|\cdot\|$  is continuous.  $\square$

**Corollary 11.163.** Let  $\langle X, \|\cdot\| \rangle$  be a normed space then we have

1.  $\forall x \in X$  the function  $\tau_x: X \rightarrow X$  defined by  $\tau_x(y) = x + y$  is continuous
2.  $\forall \alpha \in \mathbb{R}(\mathbb{C})$  the function  $\mu_\alpha: X \rightarrow X$  defined by  $\mu_\alpha(y) = \alpha \cdot y$  is continuous
3.  $\forall x \in X$  the function  $\nu_x: \mathbb{R}(\mathbb{C}) \rightarrow X$  defined by  $\nu_x(\alpha) = \alpha \cdot x$  is continuous

**Proof.** This follows from the continuity of  $+: X \times X \rightarrow X$  and  $\cdot: X \times \mathbb{R}(\mathbb{C}) \rightarrow X$  and 11.150.  $\square$

**Theorem 11.164.** When  $\langle \mathbb{R}^n, \|\cdot\| \rangle$  is equipped with product topology generated by the maximum norm  $\|\cdot\|$  (see 11.82) and  $\|\cdot\|^*$  is another norm on  $\mathbb{R}^n$  then the function  $\|\cdot\|^*: \mathbb{R}^n \rightarrow \mathbb{R}$  defined by  $x \mapsto \|x\|^*$  is continuous.

**Proof.** Let  $\{\mathcal{E}_i\}_{i \in \{1, \dots, n\}}$  be the canonical basis on  $\mathbb{R}^n$  (see 10.163) then  $\forall x = (x_1, \dots, x_n) \in \mathbb{R}^n$  we have  $x = \sum_{i=1}^n x_i \cdot \mathcal{E}_i \Rightarrow \|x\|^* = \|\sum_{i=1}^n x_i \cdot \mathcal{E}_i\|^* \leq \sum_{i=1}^n \|x_i \cdot \mathcal{E}_i\|^* = \sum_{i=1}^n |x_i| \cdot \|\mathcal{E}_i\|^* \leq \max(\{|x_i| : i \in \{1, \dots, n\}\}) \cdot \sum_{i=1}^n \|\mathcal{E}_i\|^* = \|x\| \cdot \sum_{i=1}^n \|\mathcal{E}_i\|^* = A \cdot \|x\|$  where  $A = \sum_{i=1}^n \|\mathcal{E}_i\|^* \geq 0$ . So we have

$$\|x\|^* \leq A \cdot \|x\| \text{ where } A \geq 0 \quad (11.25)$$

If now  $x \in \mathbb{R}^n$  and  $\varepsilon \in \mathbb{R}_+$  take then  $\delta = \frac{\varepsilon}{A+1}$  then if  $\|x - y\| < \delta \Rightarrow \|x - y\|^* \leq A \cdot \|x - y\| < A \cdot \delta = A \cdot \frac{\varepsilon}{A+1} \leq \varepsilon$  (as  $0 \leq A$ ) proving that  $\|\cdot\|^*$  is continuous at  $x$ . As  $x$  was chosen arbitrary we have proved that  $\|\cdot\|^*$  is continuous.  $\square$

**Definition 11.165.** Let  $\langle X, \mathcal{T}_X \rangle, \langle Y, \mathcal{T}_Y \rangle$  be topological spaces then a function  $f: X \rightarrow Y$  is a **homeomorphism** iff  $f$  is a bijection and  $f, f^{-1}$  are continuous.  $\langle X, \mathcal{T}_X \rangle$  and  $\langle Y, \mathcal{T}_Y \rangle$  are then said to be **homeomorphic**.

**Note 11.166.** Using 11.145 it is easy to see that  $f: X \rightarrow Y$  is a homeomorphism if it is a bijection and if it is open and continuous.

**Theorem 11.167.** Let  $\langle X, \mathcal{T}_X \rangle, \langle Y, \mathcal{T}_Y \rangle$  be topological spaces,  $f: X \rightarrow Y$  a homeomorphism and  $\emptyset \neq A \subseteq X$  then  $f|_A: A \rightarrow f(A)$  is a homeomorphism (using the subspace topologies of  $\mathcal{T}_X$  and  $\mathcal{T}_Y$  on  $A$  and  $f(A)$ ).

**Proof.**

1. **(bijectivity)** First using 2.52 and the fact that  $f$  is bijective and thus injective we have that  $f|_A: A \rightarrow f(A)$  is a bijection.

2. **(continuity)** Let  $x \in A$  then if  $V$  is a open set in  $f(A)$  containing  $f|_A(x) = f(x)$  there exists a  $W \in \mathcal{T}_Y$  such that  $V = f(A) \cap W$ . As  $f$  is continuous at  $x$  there exists a  $U \in \mathcal{T}_X$  with  $x \in U$  such that  $f(U) \subseteq W$ . So we have  $x \in A \cap U$  (which is open by definition in  $A$ ) and  $f|_A(A \cap U) \underset{A \cap U \subseteq A}{=} f(A \cap U) \underset{f \text{ is injective and 2.62}}{=} f(A) \cap f(U) \subseteq f(A) \cap W = V$  proving that  $f|_A$  is continuous at  $x$ . As  $x \in A$  was choosen arbitrary we conclude that  $f|_A$  is continuous.
3. **(open)** Let  $U$  open in  $A$  then there exists a  $W \in \mathcal{T}_X$  such that  $U = A \cap W$ . We have then that  $f|_A(U) \underset{U \subseteq A}{=} f(U) = f(A \cap W) \underset{f \text{ is injective and 2.62}}{=} f(A) \cap f(W)$  as  $f$  is open we have that  $f(W) \in \mathcal{T}_Y \Rightarrow f|_A(U) = f(A) \cap W$  is open in the subspace topology of  $f(A)$ .  $\square$

**Theorem 11.168.** Let  $\langle X, \mathcal{T}_X \rangle, \langle Y, \mathcal{T}_Y \rangle$  be topological spaces then if  $f: X \rightarrow Y$  is a homeomorphism we have that  $f^{-1}: Y \rightarrow X$  is a homeomorphism

**Proof.** As  $f$  is a bijection we have by 2.41 that  $f^{-1}$  is a bijection. By definition we have that  $f^{-1}$  is continuous and by 2.44 we have that  $(f^{-1})^{-1} = f$  is continuous proving that  $f^{-1}$  is a homeomorphism.  $\square$

**Theorem 11.169.** Let  $\langle X, \mathcal{T}_X \rangle, \langle Y, \mathcal{T}_Y \rangle$  and  $\langle Z, \mathcal{T}_Z \rangle$  be topological spaces and  $f: X \rightarrow Y, g: Y \rightarrow Z$  be homeomorphism's then  $g \circ f: X \rightarrow Z$  is a homeomorphism.

**Proof.** Using 2.49 we have that  $g \circ f$  is a bijection, also using 11.166 we have that  $g, f$  are continuous and open. Using 11.147, 11.146 we have then that  $g \circ f$  is open and continuous, which by 11.166 means that  $g \circ f$  is a homeomorphism.  $\square$

**Theorem 11.170.** Let  $\langle X, \mathcal{T} \rangle$  be a topological space,  $Y$  a set and  $f: X \rightarrow Y$  a bijection. Define then  $\mathcal{T}_{f, \mathcal{T}} = \{f(U) | U \in \mathcal{T}\}$  forms a topology on  $Y$  so that  $f$  is a homeomorphism. Further if  $g: X \rightarrow Y$  is such that  $f = g \circ h$  where  $h: X \rightarrow X$  is a homeomorphism then  $g$  is a homeomorphism and  $\mathcal{T}_{f, \mathcal{T}} = \mathcal{T}_{g, \mathcal{T}}$

**Proof.**

1. **( $\mathcal{T}_{f, \mathcal{T}}$  is a topology)** We have to full fill all the axioms of a topology
  - a.  $\emptyset = f(\emptyset) \in \mathcal{T}_{f, \mathcal{T}}$
  - b.  $Y \underset{f \text{ is bijection thus surjective}}{=} f(X) \in \mathcal{T}_{f, \mathcal{T}}$
  - c. If  $\{V_i\}_{i \in I}$  is a family of sets in  $\mathcal{T}_{f, \mathcal{T}}$  then  $\forall i \in I$  there exists a  $U_i \in \mathcal{T}$  such that  $V_i = f(U_i)$  so that  $\bigcup_{i \in I} V_i = \bigcup_{i \in I} f(U_i) \underset{2.62}{=} f(\bigcup_{i \in I} U_i) \in \mathcal{T}_{f, \mathcal{T}}$
  - d. Let  $V_1, V_2 \in \mathcal{T}_{f, \mathcal{T}}$  then  $\exists U_1, U_2 \in \mathcal{T}$  such that  $V_1 = f(U_1), V_2 = f(U_2)$  then  $V_1 \cap V_2 = f(U_1) \cap f(U_2) \underset{2.62}{=} f(U_1 \cap U_2) \in \mathcal{T}_{f, \mathcal{T}}$
2. **( $f$  is a homeomorphism)** If  $U \in \mathcal{T}$  then  $f(U) \in \mathcal{T}_{f, \mathcal{T}}$  so that  $f$  is open. If  $V \in \mathcal{T}_{f, \mathcal{T}}$  then  $\exists U \in \mathcal{T}$  such that  $V = f(U) \Rightarrow f^{-1}(V) = f^{-1}(f(U)) \underset{2.56}{=} U \in \mathcal{T}$  proving that  $f$  is continuous. As  $f$  is a bijection we have that  $f$  is a homeomorphism.
3. **( $\mathcal{T}_{f, \mathcal{T}} = \mathcal{T}_{g, \mathcal{T}}$ )** First as  $h$  is a homeomorphism and thus a bijection we have that  $h^{-1}: X \rightarrow X$  is a homeomorphism (see 11.168) so that  $f \circ h^{-1} = (g \circ h) \circ h^{-1} = g \circ (h \circ h^{-1}) = g \circ 1_X = g$  so that  $g$  is a homeomorphism (see 11.169) and  $\mathcal{T}_{g, \mathcal{T}}$  is defined. Next if  $U \in \mathcal{T}_{f, \mathcal{T}}$  then  $\exists V \in \mathcal{T}$  such that  $U = f(V) = f(h^{-1}(h(V))) = (f \circ h^{-1})(h(V)) = g(h(V)) \in \mathcal{T}_{g, \mathcal{T}}$  as  $h$  is a homeomorphism so that  $g(V) \in \mathcal{T}$ . Finally if  $U \in \mathcal{T}_{g, \mathcal{T}}$  then  $\exists V \in \mathcal{T}$  such that  $U = g(V) = g(h(h^{-1}(V))) = (g \circ h)(h^{-1}(V)) = f(h^{-1}(V)) \in \mathcal{T}_{f, \mathcal{T}}$  as  $h^{-1}(V) \in \mathcal{T}$  because  $h$  is a homeomorphism.  $\square$

**Theorem 11.171.** Let  $X$  be a set,  $\langle Y, \mathcal{T} \rangle$  a topological space and  $f: X \rightarrow Y$  a function then  $\mathcal{T}_{f, \mathcal{T}}^{-1} = \{f^{-1}(U) | U \in \mathcal{T}\}$  is a topology on  $X$  called the inverse induced topology. Furthermore  $f$  is a continuous function between  $X$  and  $Y$  if  $X$  is equipped with the inverse induced topology.

**Proof.**

1.  $\emptyset = f^{-1}(\emptyset) \in \mathcal{T}_{f, \mathcal{T}}^{-1}$

2.  $X = f^{-1}(Y) \in \mathcal{T}_{f, \mathcal{T}}^{-1}$
3. Let  $\{V_i\}_{i \in I}$  be a family of sets with  $\forall i \in I$  we have  $V_i \in \mathcal{T}_{f, \mathcal{T}}^{-1}$  so that  $\exists U_i \in \mathcal{T}$  with  $V_i = f^{-1}(U_i)$  then  $\bigcup_{i \in I} V_i = \bigcup_{i \in I} f^{-1}(U_i) \stackrel{2.62}{=} f^{-1}(\bigcup_{i \in I} U_i) \in \mathcal{T}_{f, \mathcal{T}}^{-1}$  [as  $\bigcup_{i \in I} U_i \in \mathcal{T}$ ].
4. If  $V_1, V_2 \in \mathcal{T}_{f, \mathcal{T}}^{-1}$  then  $\exists U_1, U_2 \in \mathcal{T}$  such that  $f^{-1}(U_1) = V_1, f^{-1}(U_2) = V_2$ . So  $V_1 \cap V_2 = f^{-1}(U_1) \cap f^{-1}(U_2) \stackrel{2.62}{=} f^{-1}(U_1 \cap U_2) \in \mathcal{T}_{f, \mathcal{T}}^{-1}$  [as  $U_1 \cap U_2 \in \mathcal{T}$ ]

As by definition we have if  $U \in \mathcal{T}$  that  $f^{-1}(U) \in \mathcal{T}_{f, \mathcal{T}}^{-1}$  so that  $f$  is continuous.  $\square$

**Theorem 11.172.** Let  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle$  be normed vector spaces over  $\mathbb{R}(\mathbb{C})$  and let  $f: X \rightarrow Y$  be a linear isometry then  $f$  is a homeomorphism.

**Proof.** As  $f$  is a isometry it is a bijection. To prove continuity take  $\varepsilon \in \mathbb{R}_+$  and  $x \in X$  then if  $\|x - x'\|_X < \varepsilon$  we have that  $\varepsilon > \|x - x'\|_X = \|f(x - x')\|_Y = \|f(x) - f(x')\|_Y$  so that by 11.155  $f$  is continuous. As  $f^{-1}$  is a isometry by 11.88 and by 10.183 is also linear we can apply the above reasoning to prove that  $f^{-1}$  is continuous.  $\square$

**Example 11.173.** Let  $\langle \mathbb{R}^2, \|\cdot\|_e \rangle$  be equiped with the euclidean norm (see 11.110),  $\langle \mathbb{C}, \|\cdot\| \rangle$  the complex space with the canonical norm then  $\mathcal{C}: \mathbb{R}^2 \rightarrow \mathbb{C}$  is a homeomorphism

**Proof.** Using 10.182 we have that  $\mathcal{C}$  is a isomorphism and by 11.112 a isometry and thus a linear isometry. Applying then the above theorem (see 11.172 we have that  $\mathcal{C}$  is a homeomorphism.  $\square$

**Theorem 11.174.** Let  $\langle X, \|\cdot\| \rangle$  be a finite  $n$ -dimensional normed space over  $\mathbb{R}(\mathbb{C})$  then there exist a norm  $\|\cdot\|_X$  on  $\mathbb{R}^n(\mathbb{C}^n)$  such that the function  $\varphi: X \rightarrow \mathbb{R}^n(\mathbb{C}^n)$  defined by  $x = \sum_{i \in \{1, \dots, n\}} x_i \cdot e_i \rightarrow \varphi(x) = (x_1, \dots, x_n)$  (where  $\{e_i\}_{i \in \{1, \dots, n\}}$  is a basis in  $X$ ) is a isometry, isomorphism and by the previous theorem they also a homeomorphism.

**Proof.** First we prove that  $\varphi$  is a bijection

1. **(injectivity)** If  $\varphi(x) = \varphi(x')$  where  $x = \sum_{i \in \{1, \dots, n\}} x_i \cdot e_i, x' = \sum_{i \in \{1, \dots, n\}} x'_i \cdot e_i$  are the unique expansions in the basis  $\{e_i\}_{i \in I}$  then  $(x_1, \dots, x_n) = (x'_1, \dots, x'_n) \Rightarrow \forall i \in I \models x_i = x'_i$  and thus  $x = x'$ .
2. **(surjectivity)** If  $(x_1, \dots, x_n) \in \mathbb{R}^n(\mathbb{C}^n)$  then  $\varphi(\sum_{i \in \{1, \dots, n\}} x_i \cdot e_i) = (x_1, \dots, x_n)$

Next to prove linearity if  $x = \sum_{i \in \{1, \dots, n\}} x_i \cdot e_i, y = \sum_{i \in \{1, \dots, n\}} y_i \cdot e_i$  and  $\alpha, \beta \in \mathbb{R}(\mathbb{C})$  then  $\alpha \cdot x + \beta \cdot y = \alpha \cdot \sum_{i \in \{1, \dots, n\}} x_i \cdot e_i + \beta \cdot \sum_{i \in \{1, \dots, n\}} y_i \cdot e_i = \sum_{i \in \{1, \dots, n\}} (\alpha \cdot x_i + \beta \cdot y_i) \Rightarrow \varphi(\alpha \cdot x + \beta \cdot y) = \alpha \cdot \varphi(x) + \beta \cdot \varphi(y)$ .

Define now the norm  $\|\cdot\|_X$  on  $\mathbb{R}^n(\mathbb{C}^n)$  by  $\|(x_1, \dots, x_n)\|_X = \|\sum_{i \in \{1, \dots, n\}} x_i \cdot e_i\|$  then  $\|\cdot\|$  is indeed a norm:

1.  $\|(x_1, \dots, x_n)\|_X = \|\sum_{i \in \{1, \dots, n\}} x_i \cdot e_i\| \geq 0$
2.  $\|\alpha \cdot (x_1, \dots, x_n)\|_X = \|(\alpha \cdot x_1, \dots, \alpha \cdot x_n)\|_X = \|\sum_{i \in \{1, \dots, n\}} (\alpha \cdot x_i) \cdot e_i\|_X = \|\alpha \cdot \sum_{i \in \{1, \dots, n\}} x_i \cdot e_i\| = |\alpha| \cdot \|\sum_{i \in \{1, \dots, n\}} x_i \cdot e_i\| = |\alpha| \cdot \|x\|$
3.  $\|(x_1, \dots, x_n) + (y_1, \dots, y_n)\|_X = \|(x_1 + y_1, \dots, x_n + y_n)\|_X = \|\sum_{i \in \{1, \dots, n\}} (x_i + y_i)\| = \|\sum_{i \in \{1, \dots, n\}} x_i + \sum_{i \in \{1, \dots, n\}} y_i\| \leq \|\sum_{i \in \{1, \dots, n\}} x_i\| + \|\sum_{i \in \{1, \dots, n\}} y_i\| = \|(x_1, \dots, x_n)\|_X + \|(y_1, \dots, y_n)\|_X$

If now  $x = \sum_{i \in \{1, \dots, n\}} x_i \cdot e_i$  then  $\|\varphi(x)\|_X = \|(x_1, \dots, x_n)\|_X = \|\sum_{i \in \{1, \dots, n\}} x_i \cdot e_i\| = \|x\|$  proving that  $\varphi$  is a isometry.  $\square$

**Theorem 11.175.** Let  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle$  be normed spaces over  $\mathbb{R}(\mathbb{C})$ ,  $U \subseteq X$  a open set,  $f: U \rightarrow Y, g: U \rightarrow Y$  be continuous functions and  $\alpha \in \mathbb{R}(\mathbb{C})$  then  $f + g: U \rightarrow Y$  and  $\alpha \cdot f: U \rightarrow Y$  are continuous.

**Proof.** We use 11.155 to prove continuity at every point  $x \in U$  and thus continuity. So take  $\varepsilon \in \mathbb{R}_+$  and  $x \in X$

1. By continuity of  $f, g$  there exists a  $\delta_1, \delta_2 \in \mathbb{R}_+$  such that if  $\|x - x'\|_X < \delta_1$  then  $\|f(x) - f(x')\|_Y < \frac{\varepsilon}{2}$  and if  $\|x - x'\|_X < \delta_2$  then  $\|f(x) - f(x')\|_Y < \frac{\varepsilon}{2}$ . Take now  $\delta = \min(\delta_1, \delta_2)$  then if  $\|x - x'\|_X < \delta \Rightarrow \|x - x'\|_X < \delta_1, \delta_2 \Rightarrow \|(f+g)(x) - (f+g)(x')\|_Y = \|f(x) - f(x') + g(x) - g(x')\|_Y \leq \|f(x) - f(x')\|_Y + \|g(x) - g(x')\|_Y = \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$
2. Consider the following cases for  $\alpha \in \mathbb{R}(\mathbb{C})$ 
  - a. ( $\alpha = 0$ ) then if  $\|x - x'\|_X < 1$  we have that  $\|0 \cdot f(x) - 0 \cdot f(x')\|_Y = \|0 - 0\|_Y = 0 < \varepsilon$  proving the continuity of  $\alpha \cdot f$
  - b. ( $\alpha \neq 0$ ) by continuity there exists a  $\delta \in \mathbb{R}_+$  such that if  $\|x - x'\|_X < \delta$  then  $|f(x) - f(x')| < \frac{\varepsilon}{|\alpha|}$  so that  $|\alpha \cdot f(x) - \alpha \cdot f(x')| = |\alpha \cdot (f(x) - f(x'))| = |\alpha| \cdot |f(x) - f(x')| < |\alpha| \cdot \frac{\varepsilon}{|\alpha|} = \varepsilon$

Proving that  $f + g$  and  $\alpha \cdot f$  are continuous.  $\square$

Using induction we can extend the above to a finite set of functions

**Theorem 11.176.** Let  $n \in \mathbb{N}_0$ ,  $\langle X, \|\cdot\|_X \rangle$ ,  $\langle Y, \|\cdot\|_Y \rangle$  be normed spaces over  $\mathbb{R}(\mathbb{C})$  and  $\{f_i|X \rightarrow Y\}_{i \in \{1, \dots, n\}}$  a finite family of continuous functions then  $\sum_{i=1}^n f_i: X \rightarrow Y$  defined by  $(\sum_{i=1}^n f_i)(x) = \sum_{i=1}^n f_i(x)$  is continuous.

**Proof.** Let  $\mathcal{S} = \{n \in \mathbb{N}_0 | \text{let } \{f_i: X \rightarrow Y\}_{i \in \{1, \dots, n\}} \text{ be a family continuous functions then } \sum_{i=1}^n f_i \text{ is continuous}\}$  then we have :

**1  $\in \mathcal{S}$ .** As  $(\sum_{i=1}^1 f_i)(x) = \sum_{i=1}^1 f_i(x) = f_1(x)$  we have that  $\sum_{i=1}^1 f_i = f_1$  which is continuous hence  $1 \in \mathcal{S}$

**$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$ .** Let  $\{f_i: X \rightarrow Y\}_{i \in \{1, \dots, n+1\}}$  a family of continuous functions then we have  $(\sum_{i=1}^n f_i)(x) = \sum_{i=1}^{n+1} f_i(x) = f_{n+1}(x) + \sum_{i=1}^n f_i(x) = (f_{n+1} + \sum_{i=1}^n f_i)(x)$  proving that  $\sum_{i=1}^{n+1} f_i = f_{n+1} + \sum_{i=1}^n f_i$  which as  $f_{n+1}$  is continuous and  $\sum_{i=1}^n f_i$  is continuous (because  $n \in \mathcal{S}$ ) means by the previous theorem that  $\sum_{i=1}^{n+1} f_i$  is continuous. Hence  $n+1 \in \mathcal{S}$ .  $\square$

**Theorem 11.177.** Let  $\langle X, \|\cdot\|_X \rangle$  be a normed spaces over  $\mathbb{R}(\mathbb{C})$  and  $f: X \rightarrow \mathbb{R}(\mathbb{C})$ ,  $g: X \rightarrow \mathbb{R}(\mathbb{C})$  be continuous functions  $f \cdot g$  is continuous.

**Proof.** Let  $x \in X$  and  $\varepsilon \in \mathbb{R}_+$ . By continuity of  $f, g$  there exists a  $\delta_1 \in \mathbb{R}_+$  such that if  $\|x - x'\|_X < \delta_1$  then  $|f(x) - f(x')| < 1 \Rightarrow |f(x')| \leq |f(x') - f(x) + f(x)|_Y \leq |f(x') - f(x)| + |f(x)| \leq 1 + |f(x)|$ . Also choose a  $\delta_2$  such that if  $\|x - x'\|_X < \delta_2$  then  $|f(x) - f(x')| < \frac{\varepsilon}{2(1 + |f(x)|)}$ . Now choose  $\delta_3 \in \mathbb{R}_+$  such that if  $\|x - x'\|_X < \delta_3$  then  $|g(x) - g(x')| < \frac{\varepsilon}{2(1 + |f(x)|)}$ . Take then  $\delta = \min(\delta_1, \delta_2, \delta_3)$  then if  $\|x - x'\|_X < \delta$  we have

$$\begin{aligned} |f(x) \cdot g(x) - f(x') \cdot g(x')| &= |f(x) \cdot g(x) - f(x') \cdot g(x) + f(x') \cdot g(x) - f(x') \cdot g(x')| \\ &\leq |f(x) \cdot g(x) - f(x') \cdot g(x)|_Y + \|f(x') \cdot g(x) - f(x') \cdot g(x')\|_Y \\ &\leq |g(x)| \cdot |f(x) - f(x')| + |f(x')| \cdot |g(x) - g(x')| \\ &< |g(x)| \cdot \frac{\varepsilon}{2(1 + |f(x)|)} + (1 + |f(x)|) \cdot \frac{\varepsilon}{2(1 + |f(x)|)} \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \end{aligned}$$

proving that  $f \cdot g$  is continuous at  $x$ . As  $x \in X$  was chosen arbitrary we have that  $f \cdot g$  is continuous.  $\square$

**Corollary 11.178.** Let  $n \in \mathbb{N}$ ,  $\langle X, \|\cdot\|_X \rangle$  be a normed spaces over  $\mathbb{R}(\mathbb{C})$  and  $f: X \rightarrow \mathbb{R}(\mathbb{C})$  then  $f^n: X \rightarrow \mathbb{R}(\mathbb{C})$  defined by  $f^n(x) = (f(x))^n$  is continuous

**Proof.** Let  $\mathcal{S} = \{n \in \mathbb{N} | f^n: X \rightarrow \mathbb{R}(\mathbb{C}) \text{ is a continuous function}\}$  then we have

**0  $\in \mathcal{S}$ .** If  $n = 0$  then  $f^0: X \rightarrow \mathbb{R}(\mathbb{C})$  is  $C_1$  which is continuous by 11.140.

**0 ∈ S ⇒ n + 1 ∈ S.** As  $\forall x \in X$  we have that  $f^{n+1}(x) = (f(x))^{n+1} = f(x) \cdot (f(x))^n = (f \cdot f^n)(x)$  which is continuous because  $f^n$  is continuous as  $n \in S$ ,  $f$  is continuous and the above theorem.  $\square$

Using this corollary on the identity function  $1_X$  proves the following corollary.

**Corollary 11.179.** *The mapping  $.^n: \mathbb{R} \rightarrow \mathbb{R}$  defined by  $x \rightarrow x^n$  where  $n \in \mathbb{N}$  is continuous in the normed topology defined by  $\| \cdot \|$*

## 11.6 Linear maps and continuity

In this chapter we assume that all the normed spaces are defined over the same field  $\mathbb{K}$  (which is either  $\mathbb{R}$  or  $\mathbb{C}$ ).

**Theorem 11.180.** *Let  $\langle X, \| \cdot \|_X \rangle, \langle Y, \| \cdot \|_Y \rangle$  be normed vector spaces over the real (complex) field  $\mathbb{K}$  and  $L: X \rightarrow Y$  a linear mapping ( $L \in \text{Hom}(X, Y)$ ) then the following are equivalent*

1.  $L$  is continuous
2.  $L$  is continuous at  $0 \in X$
3.  $\exists k \in \mathbb{R}_+$  such that  $\forall x \in X \vdash \|x\|_X = 1$  we have  $\|L(x)\|_Y \leq k$
4.  $\exists k \in \mathbb{R}_+$  such that  $\forall x \in X$  we have  $\|L(x)\|_Y \leq k \cdot \|x\|_X$

**Proof.**

1. **(1 ⇒ 2)** If  $L$  is continuous then by definition it is continuous at  $0 \in X$
2. **(2 ⇒ 3)** Take  $1 \in \mathbb{R}_+$  then using 11.155 there exists a  $\delta \in \mathbb{R}_+$  such that if  $x \in X$  and  $\|x\|_X = \|0 - x\| < \delta \Rightarrow \|L(x)\|_Y = \|L(x - 0)\|_Y = \|L(x) - L(0)\|_Y = \|L(0) - L(x)\|_Y < 1$ . Now  $\forall x \in X \vdash \|x\|_X = 1$  take then  $0 < \delta' < \delta$  (see 9.58) then  $\|\delta' \cdot x\|_X = |\delta'| \cdot \|x\|_X = \delta' \cdot 1 = \delta' < \delta \Rightarrow \delta' \cdot \|L(x)\|_Y = |\delta'| \cdot \|L(x)\|_Y = \|\delta' \cdot L(x)\|_Y = \|L(\delta' \cdot x)\|_Y < 1 \xrightarrow{\delta' > 0} \|L(x)\|_Y < \frac{1}{\delta'}$ . So if we take  $k = \frac{1}{\delta'}$  then  $\forall x \in X$  with  $\|x\|_X = 1$  we have  $\|L(x)\|_Y < k \Rightarrow \|L(x)\|_Y < k$
3. **(3 ⇒ 4)** Let  $k \in \mathbb{R}_+$  be such that  $\forall x \in X$  with  $\|x\|_X = 1$  we have  $\|L(x)\|_Y \leq k$ . Take then  $x \in X$  then we have the following possibilities
  - a.  $(\|x\|_X = 0)$  then  $x = 0$  so that  $L(x) = 0 \Rightarrow \|x\|_X = 0 \wedge \|L(x)\|_Y = 0 \Rightarrow \|x\|_X = 0 \leq 0 = k \cdot 0 = \|L(x)\|_X$
  - b.  $(\|x\|_X \neq 0)$  then we have that  $\left\| \frac{1}{\|x\|_X} \cdot x \right\|_X = \left| \frac{1}{\|x\|_X} \right| \cdot \|x\|_X = \frac{1}{\|x\|_X} \cdot \|x\|_X = 1$  so that  $\frac{1}{\|x\|_X} \cdot \|L(x)\|_Y = \left| \frac{1}{\|x\|_X} \right| \cdot \|L(x)\|_Y = \left\| \frac{1}{\|x\|_X} \cdot L(x) \right\|_Y = \left\| L \left( \frac{1}{\|x\|_X} \cdot x \right) \right\|_Y \leq k \Rightarrow \|L(x)\|_Y \leq k \cdot \|x\|_X$
4. **(4 ⇒ 1)** Let  $k \in \mathbb{R}_+$  such that  $\forall x \in X$  we have  $\|L(x)\|_Y \leq k \cdot \|x\|_X$ . If now  $x \in X$  and  $\varepsilon > 0$  take then  $\delta = \frac{\varepsilon}{k}$  then if  $y \in X$  with  $\|x - y\|_X < \delta \Rightarrow \|L(x) - L(y)\|_Y = \|L(x - y)\|_Y \leq k \cdot \|x - y\|_Y < k \cdot \frac{\varepsilon}{k} = \varepsilon$  proving that  $L$  is continuous at  $x$ . As  $x$  is chosen arbitrary we have that  $L$  is continuous.  $\square$

**Corollary 11.181.** *Let  $\langle \mathbb{R}^n, \| \cdot \|_n \rangle, \langle \mathbb{R}^m, \| \cdot \|_m \rangle$  be equipped with the maximum norm then every linear function  $L: \mathbb{R}^n \rightarrow \mathbb{R}^m$  is continuous*

**Proof.** Let  $\{\mathcal{E}_i\}_{i \in \{1, \dots, n\}}$  be the canonical basis of  $\mathbb{R}^n$  (see 10.163) then  $\forall x \in \mathbb{R}^n$  we have that  $x = (x_1, \dots, x_n) = \sum_{i \in \{1, \dots, n\}} x_i \cdot \mathcal{E}_i$  where  $x_i = \pi_i(x)$  so that  $\|L(x)\|_m = \left\| \sum_{i \in \{1, \dots, n\}} \pi_i(x) \cdot \mathcal{E}_i \right\|_m \leq \sum_{i \in \{1, \dots, n\}} \|\pi_i(x) \cdot \mathcal{E}_i\| = \sum_{i \in \{1, \dots, n\}} |\pi_i(x)| \cdot \|\mathcal{E}_i\| \leq k' \cdot \sum_{i \in \{1, \dots, n\}} |\pi_i(x)|$  where  $k' = \max(\{\|L(\mathcal{E}_i)\|_m | i \in \{1, \dots, n\}\})$ . As  $\max(\{|\pi_i(x)| | i \in \{1, \dots, n\}\}) = \|x\|_n$  we have then  $\|L(x)\|_m \leq k' \cdot n \cdot \|x\|_n$  or if we take  $k = k' \cdot n \in \mathbb{R}_+$  we have  $\|L(x)\|_m \leq k \cdot \|x\|_n$  proving continuity by the previous theorem.  $\square$

**Example 11.182.** Let  $\langle X, \|\cdot\| \rangle$  be a normed spaces then  $\mathbb{1}_X: X \rightarrow X$  defined by  $\mathbb{1}_X(x) = x$  is linear and continuous

**Proof.** Linearity is trivial as  $\mathbb{1}_X(\alpha \cdot x + \beta \cdot y) = \alpha \cdot x + \beta \cdot y = \alpha \cdot \mathbb{1}_X(x) + \beta \cdot \mathbb{1}_X(y)$ . Further as  $\|\mathbb{1}_X(x)\| = \|x\| = 1 \cdot \|x\|$  it follows from 11.180.  $\square$

**Corollary 11.183.** Let  $\langle \mathbb{R}^n, \|\cdot\|_n \rangle$ ,  $\langle \mathbb{R}^n, \|\cdot\|_n \rangle$  be equipped with the maximum norm then every linear isomorphism  $L: \mathbb{R}^n \rightarrow \mathbb{R}^n$  is a homeomorphism (see 11.165)

**Proof.** Using the previous corollary (see 11.181) we have that  $L$  is continuous. As  $L^{-1}: \mathbb{R}^n \rightarrow \mathbb{R}^n$  is also a linear isomorphism we have that  $L^{-1}$  is also continuous.  $\square$

**Theorem 11.184.** Let  $\langle X, \|\cdot\|_X \rangle$ ,  $\langle Y, \|\cdot\|_Y \rangle$  be normed spaces and let  $L: X \rightarrow Y$  be a continuous function then if we define  $A_{L,s} = \{k \in \mathbb{R}_+ | \forall x \in X \vdash \|x\|_X = 1 \text{ we have } \|L(x)\|_Y \leq k\}$ ,  $A_{L,r} = \{k \in \mathbb{R}_+ | \forall x \in X \vdash \|L(x)\|_Y \leq k \cdot \|x\|_X\}$  we have that

1.  $A_{L,s} = A_{L,r}$
2.  $\inf(A_{L,s}) = \inf(A_{L,r})$  exists (is finite) and  $\inf(A_{L,s}) = \inf(A_{L,r}) \geq 0$

**Proof.**

1. If  $k \in A_{L,s}$  then  $\forall x \in X$  we have the following cases to consider
  - a. ( $\|x\|_X = 0$ ) then  $x = 0$  and thus  $\|L(x)\|_Y = \|L(0)\|_Y = \|0\|_Y = 0 \leq k \cdot 0 = k \cdot \|0\|_X = k \cdot \|x\|_X$
  - b. ( $\|x\|_X \neq 0$ ) then  $\left\| \frac{1}{\|x\|_X} \right\|_X = \left| \frac{1}{\|x\|_X} \right| \cdot \|x\|_X = \frac{1}{\|x\|_X} \cdot \|x\|_X = 1$  so that as  $k \in A_s$  we have that  $\frac{1}{\|x\|_X} \cdot \|L(x)\|_Y = \left\| \frac{1}{\|x\|_X} \cdot L(x) \right\|_Y = \left\| L \left( \frac{1}{\|x\|_X} \right) \right\|_Y \leq k \Rightarrow \|L(x)\|_Y \leq k \cdot \|x\|_X$

So in all cases we have  $\|L(x)\|_Y \leq k \cdot \|x\|_X$  proving that  $x \in A_{L,r}$ .

2. If  $k \in A_{L,r}$  then  $\forall x \in X$  with  $\|x\|_X = 1$  we have that  $\|L(x)\|_Y \leq k \cdot \|x\|_X = k \cdot 1 = k$  so that  $k \in A_{L,s}$

Using (1) and (2) we have proved that  $A_{L,s} = A_{L,r}$ .

As  $L$  is a continuous linear mapping we have by 11.180 the existence of a  $k \in A_{L,r}$  so that  $\emptyset \neq A_{L,r} = A_{L,s}$ . As  $\forall k \in A_{L,r}$  we have that  $0 < k$  we see that  $A_{L,r}, A_{L,s}$  is bounded below by 0. Using then the fact that the reals are conditionally complete (see 9.43) and 2.180 we have that  $\inf(A_{L,r}) = \inf(A_{L,s})$  exists and  $\inf(A_{L,s}) = \inf(A_{L,r}) \geq 0$  (by the definition of a infimum).  $\square$

The above theorem means that the following definition makes sense

**Definition 11.185. (norm of linear continuous function)** Let  $\langle X, \|\cdot\|_X \rangle$ ,  $\langle Y, \|\cdot\|_Y \rangle$  be normed spaces and let  $L: X \rightarrow Y$  be a continuous function then  $\|L\| = \inf(A_{L,s}) = \inf(A_{L,r})$  is called the operator norm.

**Theorem 11.186.** Let  $\langle X, \|\cdot\|_X \rangle$ ,  $\langle Y, \|\cdot\|_Y \rangle$  be normed spaces over the real (complex) field  $\mathbb{K}$  and let  $L, L_1, L_2: X \rightarrow Y$  be linear continuous functions then we have

1.  $\forall x \in X$  we have  $\|L(x)\|_Y \leq \|L\| \cdot \|x\|_Y$
2.  $L_1 + L_2$  is linear and continuous and  $\|L_1 + L_2\| \leq \|L_1\| + \|L_2\|$
3. If  $L = C_0$  ( $C_0: X \rightarrow Y$  where  $C_0(x) = 0$ ) then  $\|L\| = 0$
4.  $\forall \alpha \in \mathbb{K}$  we have that  $\alpha \cdot L$  is continuous and  $\|\alpha \cdot L\| = |\alpha| \cdot \|L\|$

**Proof.**

1. If  $x \in X$  then we have to consider the following cases

- a. ( $x = 0$ ) then  $\|L(x)\|_Y = \|L(0)\|_Y = \|0\|_Y = 0 \leq \|L\| \cdot 0 = \|L\| \cdot \|0\|_Y = \|L\| \cdot \|x\|_X \Rightarrow \|L(x)\|_Y \leq \|L\| \cdot \|x\|_X$

- b. ( $\mathbf{x} \neq 0$ ) we proceed now by contradiction, so assume that  $\|L\| \cdot \|x\|_X < \|L(x)\|_Y \Rightarrow \|L\| < \frac{\|L(x)\|_Y}{\|x\|_X}$  then by 2.171 there exists a  $k \in A_{L,r} = \{k \in \mathbb{R}_+ \mid \forall x \in X \models \|L(x)\|_Y \leq k \cdot \|x\|_X\}$  such that  $\|L\| \leq k < \frac{\|L(x)\|_Y}{\|x\|_X} \Rightarrow k \cdot \|x\| < \|L(x)\|_Y$  contradicting  $k \in A_{L,r}$ . So we must have that  $\|L(x)\|_Y \leq \|L\| \cdot \|x\|_X$
2. Because the sum of continuous (linear) functions is continuous (linear) (see 11.175 and 10.187) we have that  $L_1 + L_2$  is linear and continuous. Using (1) we have then that  $\forall x \in X$  then  $\|L_1(x)\|_Y \leq \|L_1\| \cdot \|x\|_X$  and  $\|L_2(x)\|_Y \leq \|L_2\| \cdot \|x\|_X$  so that  $\|(L_1 + L_2)(x)\|_Y = \|L_1(x) + L_2(x)\|_Y \leq \|L_1(x)\|_Y + \|L_2(x)\|_Y \leq \|L_1\| \cdot \|x\|_X + \|L_2\| \cdot \|x\|_X = (\|L_1\| + \|L_2\|) \cdot \|x\|_X$  so that  $\|L_1\| + \|L_2\| \in A_{L_1+L_2,r} \Rightarrow \|L_1 + L_2\| = \inf(A_{L_1+L_2,r}) \leq \|L_1\| + \|L_2\|$
3. If  $L = C_0$  then  $\forall x \in X$  we have  $L(x) = 0 \Rightarrow \|L(x)\|_Y = 0 \leq 0 = 0 \cdot \|x\|_X$  so that  $0 \in A_{C_0,r}$  and thus  $0 \leq \|C_0\| = \inf(A_{C_0,r}) \leq 0 \Rightarrow \|C_0\| = 0$
4. Using 11.175 and 10.187 we have that  $\alpha \cdot L$  is a continuous linear mapping. Because of (1) we have  $\forall x \in X$  that  $\|L(x)\|_Y \leq \|L\| \cdot \|x\|_X$  so that  $\|\alpha \cdot L(x)\|_Y = |\alpha| \cdot \|L(x)\|_Y \leq |\alpha| \cdot (\|L\| \cdot \|x\|_X) = (|\alpha| \cdot \|L\|) \cdot \|x\|_X$  so that  $|\alpha| \cdot \|L\| \in A_{\alpha \cdot L,r} \Rightarrow \|\alpha \cdot L\| \leq |\alpha| \cdot \|L\|$ . We consider now two cases for  $\alpha \in \mathbb{K}$
- a. ( $\alpha = 0$ ) then  $\alpha \cdot L = 0 \cdot L = C_0$  and thus  $\|\alpha \cdot L\| = \|C_0\| = 0 = |0| \cdot \|L\|$  so we have  $\|\alpha \cdot L\| = |\alpha| \cdot \|L\|$
- b. ( $\alpha \neq 0$ ) then  $|\alpha| \neq 0$  then  $\forall x \in X$  we have  $\|\alpha \cdot L(x)\|_Y = |\alpha| \cdot \|L(x)\|_Y$  so that  $\|L(x)\|_Y = \frac{\|\alpha \cdot L(x)\|_Y}{|\alpha|} \leq \frac{\|\alpha \cdot L\| \cdot \|x\|_X}{|\alpha|}$  so that  $\frac{\|\alpha \cdot L\|}{|\alpha|} \in A_{L,r} \Rightarrow \|L\| \leq \frac{\|\alpha \cdot L\|}{|\alpha|} \Rightarrow |\alpha| \cdot \|L\| \leq \|\alpha \cdot L\|$  this together with  $\|\alpha \cdot L\| \leq |\alpha| \cdot \|L\|$   $\square$

**Theorem 11.187.** Let  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle$  be normed spaces over the real (complex) field  $\mathbb{K}$ , let  $L: X \rightarrow Y$  be a continuous function and let  $X$  be non trivial (so there exists a  $x \in X$  with  $x \neq 0$ ) then we have if  $A_{L,t} = \{\|L(x)\|_Y \mid x \in X \text{ with } \|x\|_X = 1\}$  and  $A_{L,u} = \{\|L(x)\|_Y \mid x \in X \text{ with } \|x\|_X \leq 1\}$  that  $\|L\| = \sup(A_{L,t}) = \sup(A_{L,u})$

**Proof.** If  $k \in A_{L,t}$  then  $\exists x \in X$  with  $\|x\|_X = 1 \leq 1$  such that  $k = \|L(x)\|_Y \Rightarrow k \in A_{L,u}$  proving that

$$A_{L,t} \subseteq A_{L,u} \quad (11.26)$$

As  $\exists x_0 \in X$  with  $x_0 \neq 0 \Rightarrow \left\| \frac{1}{\|x_0\|_X} \cdot x_0 \right\|_X = 1 \leq 1$  so that  $\frac{1}{\|x_0\|_X} \cdot x \in A_{L,t}, A_{L,u}$  and thus

$$\emptyset \neq A_{L,t} \wedge \emptyset \neq A_{L,u} \quad (11.27)$$

Now if  $k \in A_{L,u}$  then  $\forall x \in X$  with  $\|x\|_X \leq 1$  we have that  $k = \|L(x)\|_Y$  and as  $\|L(x)\|_Y \leq \|L\| \cdot \|x\|_X \leq \|L\|$  we see that  $\|L\|$  is an upper bound of  $A_{L,u}$  and by 11.26 also of  $A_{L,t}$ . Using 9.43 we have that  $\sup(A_{L,u})$  and  $\sup(A_{L,t})$  exists and that

$$\sup(A_{L,t}) \leq \sup(A_{L,u}) \leq \|L\| \quad (11.28)$$

(using also 11.26 and 2.175). Now  $\forall x \in X$  with  $\|x\|_X = 1$  we have  $\|L(x)\|_Y \in A_{L,t} \Rightarrow \|L(x)\|_Y \leq \sup(A_{L,t}) \Rightarrow \sup(A_{L,t}) \in A_{L,s} \Rightarrow \|L\| = \inf(A_{L,s}) \leq \sup(A_{L,t})$  which together with 11.28 gives that  $\|L\| = \sup(A_{L,t}) = \sup(A_{L,u})$   $\square$

**Example 11.188.** If  $\{\langle X_i, \|\cdot\|_i \rangle\}_{i \in \{1, \dots, n\}}, n \in \mathbb{N}_0$  be a family of normed spaces,  $\langle \prod_{i \in \{1, \dots, n\}} X_i, \|\cdot\| \rangle$  is the product space with the maximum norm then  $\forall i \in \{1, \dots, n\}$  we have that  $\pi_i: \prod_{i \in \{1, \dots, n\}} X_i \rightarrow X_i$  is a continuous linear mapping with  $\|\pi_i\| < 1$ .

**Proof.**

1. (**linearity**) I  $x, y \in \prod_{i \in \{1, \dots, n\}}, \alpha, \beta \in \mathbb{K}$  then  $\pi_i(\alpha \cdot x + \beta \cdot y) = (\alpha \cdot x + \beta \cdot y)_i = \alpha \cdot x_i + \beta \cdot y_i = \alpha \cdot \pi_i(x) + \beta \cdot \pi_i(y)$  proving that  $\pi_i$  is linear.

2. **(continuity)** If  $x \in \prod_{i \in \{1, \dots, n\}} X_i$  then  $\|\pi_i(x)\|_i = \|x_i\|_i \leq \{\|x_i\|_i \mid i \in \{1, \dots, n\}\} = \|x\| \leq 1 \cdot \|x\|$  proving that  $\pi_i$  is continuous with  $\|\pi_i\| \leq 1$  as  $\|\pi_i\| = \inf(\{k \in \mathbb{R}_+ \mid \forall x \in X \models \|\pi_i(x)\|_i \leq k \cdot \|x\|_i\}) \leq 1$   $\square$

**Definition 11.189.** Let  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle$  be normed spaces over the real (complex) field  $\mathbb{K}$  then  $L(X, Y) = \{L \in \text{Hom}(X, Y) \mid L \text{ is continuous}\} \subseteq \text{Hom}(X, Y)$  is the set of all linear continuous functions between  $X$  and  $Y$ .

**Theorem 11.190.** Let  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle$  be normed spaces over the real (complex) field  $\mathbb{K}$  then  $\langle L(X, Y), +, \cdot \rangle$  forms a subspace of  $\text{Hom}(X, Y)$  (and is thus a vector space). Furthermore if we define  $\|\cdot\|_{L(X, Y)}: L(X, Y) \rightarrow \mathbb{R}_+$  by  $L \mapsto \|L\|_{L(X, Y)} = \|L\| = \sup(A_{L,t}) = \sup(A_{L,u})$   $\stackrel{11.187}{=} \inf(A_{L,s}) = \inf(A_{L,r})$  then  $\langle L(X, Y), \|\cdot\|_{L(X, Y)} \rangle$  forms a normed vector space over  $\mathbb{K}$

**Proof.** If  $\alpha, \beta \in \mathbb{K}$  and  $L_1, L_2 \in L(X, Y)$  then by 11.186 we have that  $\alpha \cdot L_1 + \beta \cdot L_2 \in L(X, Y)$ . Also using 11.186 we have that  $\|\cdot\|$  is a pseudo norm. To prove that it is a norm assume that  $\|L\| = 0$  then  $\forall x \in X$  we have  $\|L(x)\|_Y \leq \|L\| \cdot \|x\|_X = 0 \cdot \|x\|_X = 0 \Rightarrow \|L(x)\|_Y = 0 \Rightarrow L(x) = 0$  and thus  $L = C_0 = 0$  (the zero function).  $\square$

**Theorem 11.191.** Let  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle$  and  $\langle Z, \|\cdot\|_Z \rangle$  be normed spaces over the real (complex) field and  $L_1: X \rightarrow Y, L_2: Y \rightarrow Z$  be linear continuous functions then  $L_2 \circ L_1: X \rightarrow Z$  is linear and continuous, so that  $L_2 \circ L_1 \in L(X, Z)$ . Further we have that  $\|L_2 \circ L_1\|_{L(X, Z)} \leq \|L_1\|_{L(X, Y)} \cdot \|L_2\|_{L(Y, Z)}$

**Proof.** By 11.147 and 10.184 we have that  $L_2 \circ L_1$  is linear and continuous. Now  $\forall x \in X$  we have that  $\|(L_2 \circ L_1)(x)\|_Z = \|L_2(L_1(x))\|_Z \leq \|L_2\|_{L(Y, Z)} \cdot \|L_1(x)\|_Y \leq \|L_2\|_{L(Y, Z)} \cdot \|L_1\|_{L(X, Y)} \cdot \|x\|_X$  meaning that  $\|L_1\| \circ \|L_2\| \in A_{L_2 \circ L_1, r}$  and thus that  $\|L_2 \circ L_1\| = \inf(A_{L_2 \circ L_1}) \leq \|L_1\| \cdot \|L_2\|$   $\square$

As  $\|L\| = \{\|L(x)\|_Y \mid x \in X \text{ with } \|x\|_X = 1\}$  we see that  $\|\cdot\|$  is dependent on the norm chosen for  $Y$  and  $X$ . We prove now that from the point of the generated topologies there is no difference.

**Theorem 11.192.** Let  $X, Y$  be vector spaces over the field  $\mathbb{K}$  and let  $\|\cdot\|_{X_1}, \|\cdot\|_{X_2}$  be two equivalent norms on  $X$  and let  $\|\cdot\|_{Y_1}, \|\cdot\|_{Y_2}$  be two equivalent norms on  $Y$ . Then if  $\|\cdot\|_1$  is a norm on  $L(X, Y)$  based on  $\|\cdot\|_{X_1}, \|\cdot\|_{Y_1}$  and  $\|\cdot\|_2$  is a norm based on  $\|\cdot\|_{X_2}, \|\cdot\|_{Y_2}$  then  $\|\cdot\|_1, \|\cdot\|_2$  are equivalent norms. So the topology of  $L(X, Y)$  only depends on the topologies of  $X, Y$

**Proof.** First given  $L \in L(X, Y)$  then if  $A_{1,L} = \{k \mid \forall x \in X \text{ we have } \|L(x)\|_{Y_1} \leq k \cdot \|x\|_{X_1}\}$   $\|L\|_1 = \inf(A_{1,L})$ , if  $A_{2,L} = \{k \mid \forall x \in X \text{ we have } \|L(x)\|_{Y_2} \leq k \cdot \|x\|_{X_2}\}$  then  $\|L\|_2 = \inf(A_{2,L})$ . As  $\|\cdot\|_{X_1}, \|\cdot\|_{X_2}$  are equivalent and  $\|\cdot\|_{Y_1}, \|\cdot\|_{Y_2}$  are equivalent there exists (see 11.84)  $\alpha_X > 0, \alpha_Y > 0, \beta_X > 0, \beta_Y > 0$  such that  $\forall x \in X, y \in Y$  we have  $\alpha_X \cdot \|x\|_{X_2} \leq \|x\|_{X_1} \leq \beta_X \cdot \|x\|_{X_2}$  and  $\alpha_Y \cdot \|y\|_{Y_2} \leq \|y\|_{Y_1} \leq \beta_Y \cdot \|y\|_{Y_2}$ . Then we have :

1. If  $k \in A_{1,L}$  then  $\forall x \in X$  we have  $\|L(x)\|_{Y_1} \leq k \cdot \|x\|_{X_1} \Rightarrow \alpha_Y \cdot \|L(x)\|_{Y_2} \leq \alpha_Y \cdot \|L(x)\|_{Y_1} \leq k \cdot \alpha_Y \cdot \|x\|_{X_1} \leq (k \cdot \beta_X) \cdot \|x\|_{X_2} \Rightarrow \|L(x)\|_{Y_2} \leq \frac{k \cdot \beta_X}{\alpha_Y} \cdot \|x\|_{X_2} \Rightarrow \frac{k \cdot \beta_X}{\alpha_Y} \in A_{2,L} \Rightarrow \inf(A_{2,L}) \leq k \cdot \frac{\beta_X}{\alpha_Y} \Rightarrow \frac{\alpha_Y}{\beta_X} \cdot \inf(A_{2,L}) \leq k$ . Assume now that  $\inf(A_{1,L}) < \frac{\alpha_Y}{\beta_X} \cdot \inf(A_{2,L}) \Rightarrow \exists k \in A_{1,L} \text{ with } \inf(A_{1,L}) \leq k < \frac{\alpha_Y}{\beta_X} \cdot \inf(A_{2,L}) \leq k$  a contradiction so that  $\frac{\alpha_Y}{\beta_X} \cdot \|L\|_{X_2} = \frac{\alpha_Y}{\beta_X} \cdot \inf(A_{2,L}) \leq \inf(A_{1,L}) = \|L\|_1$
2. If  $k \in A_{2,L}$  then  $\forall x \in X$  we have  $\|L(x)\|_{Y_2} \leq k \cdot \|x\|_{X_2} \Rightarrow \frac{1}{\beta_Y} \cdot \|L(x)\|_{Y_1} \leq \frac{1}{\beta_Y} \cdot \|L(x)\|_{Y_2} \leq k \cdot \frac{1}{\beta_Y} \cdot \|x\|_{X_2} \leq \left(\frac{1}{\alpha_X} \cdot k\right) \cdot \|x\|_{X_1} \Rightarrow \|L(x)\|_{Y_1} \leq \frac{\beta_Y \cdot k}{\alpha_X} \cdot \|x\|_{X_1} \Rightarrow \frac{k \cdot \beta_Y}{\alpha_X} \in A_{1,L} \Rightarrow \inf(A_{1,L}) \leq k \cdot \frac{\beta_Y}{\alpha_X} \Rightarrow \frac{\alpha_X}{\beta_Y} \cdot \inf(A_{1,L}) \leq k$ . Assume now that  $\inf(A_{2,L}) < \frac{\alpha_X}{\beta_Y} \cdot \inf(A_{1,L}) \Rightarrow \exists k \in A_{2,L} \text{ such that } \inf(A_{2,L}) \leq k < \frac{\alpha_X}{\beta_Y} \cdot \inf(A_{1,L}) \leq k$  a contradiction so that  $\frac{\alpha_X}{\beta_Y} \cdot \inf(A_{1,L}) \leq \inf(A_{2,L}) \Rightarrow \inf(A_{1,L}) \leq \frac{\beta_Y}{\alpha_X} \cdot \inf(A_{2,L}) \Rightarrow \|L\|_1 \leq \frac{\beta_Y}{\alpha_X} \cdot \|L\|_2$

Using (1) and (2) we have  $\frac{\alpha_Y}{\beta_X} \cdot \|L\|_2 \leq \|L\|_1 \leq \frac{\beta_Y}{\alpha_X} \cdot \|L\|_2$  which as  $0 < \frac{\alpha_Y}{\beta_X}, \frac{\beta_Y}{\alpha_X}$  and  $L$  is chosen arbitrary proves that  $\|\cdot\|_1$  and  $\|\cdot\|_2$  are equivalent.  $\square$

**Theorem 11.193.** Let  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle$  be normed vector spaces on  $\mathbb{K}$  then a linear function  $L: X \rightarrow Y$  is open if and only if there exists  $\delta > 0$  such that  $B_{\|\cdot\|_Y}(0, \delta) \subseteq L(B_{\|\cdot\|_X}(0, 1))$

**Proof.**

1.  $(\Rightarrow)$  This trivial using 11.153
2.  $(\Leftarrow)$  Using 11.153 we only have to prove that  $\forall x \in X$  and  $\delta > 0$  there exists a  $\delta' > 0$  such that  $L(x) \subseteq B_{\|\cdot\|_Y}(L(x), \delta') \subseteq L(B_{\|\cdot\|_X}(x, \delta))$ . Now as we now by the hypothesis that there exists a  $\delta'' > 0$  such that  $B_{\|\cdot\|_Y}(0, \delta'') \subseteq L(B_{\|\cdot\|_X}(0, 1))$  we have:

$$\begin{aligned}
 y \in B_{\|\cdot\|_Y}\left(L(x), \frac{\delta''}{\delta}\right) &\Rightarrow \|y - L(x)\|_Y < \frac{\delta''}{\delta} \\
 &\Rightarrow \|\delta \cdot (y - L(x))\|_Y < \delta'' \\
 &\Rightarrow \delta \cdot (y - L(x)) \in B_{\|\cdot\|_Y}(0, \delta'') \\
 &\Rightarrow \delta \cdot (y - L(x)) \in L(B_{\|\cdot\|_X}(0, 1)) \\
 &\Rightarrow \exists x_0 \in X \text{ with } \|x_0\|_X < 1 \text{ such that } L(x_0) = \delta \cdot (y - L(x)) \\
 &\Rightarrow y = \frac{L(x_0)}{\delta} + L(x) \text{ take } x_1 = \frac{x_0}{\delta} \\
 &\Rightarrow \|x_1\| < \delta \wedge y = L(x_1) + L(x) = L(x_1 + x) \text{ take now } x_2 = x_1 + x \\
 &\Rightarrow \|x_2 - x\| < \delta \text{ and } y = L(x_2) \\
 &\Rightarrow y \in L(B_{\|\cdot\|_Y}(x, \delta))
 \end{aligned}$$

□

## 11.7 Multilinear mappings and continuity

In this chapter we assume that all the normed spaces are defined over the same field  $\mathbb{K}$  (which is either  $\mathbb{R}$  or  $\mathbb{C}$ ).

**Theorem 11.194.** Let  $\{\langle X_i, \|\cdot\|_i \rangle\}_{i \in \{1, \dots, n\}}$  be normed vector spaces over the real (complex) field  $\mathbb{K}$ ,  $\langle Y, \|\cdot\|_Y \rangle$  a normed vector space over the field  $\mathbb{K}$  and  $L: \prod_{i \in \{1, \dots, n\}} X_i \rightarrow Y$  a multi-linear function then we have the following equivalences (using the product topology on  $\prod_{i \in \{1, \dots, n\}} X_i$ ) and the norm topology on  $Y$ .

1.  $L$  is continuous
2.  $L$  is continuous at  $0 \in \prod_{i \in \{1, \dots, n\}} X_i$
3.  $\exists k \in \mathbb{R}_+$  such that  $\forall x \in \prod_{i \in \{1, \dots, n\}} X_i$  with  $\forall i \in \{1, \dots, n\}$  we have that  $\|x_i\|_i = 1$  then  $\|L(x)\|_Y \leq k$
4.  $\exists k \in \mathbb{R}_+$  such that  $\forall x \in \prod_{i \in \{1, \dots, n\}} X_i$  we have  $\|L(x)\|_Y \leq k \cdot \prod_{i \in \{1, \dots, n\}} \|x_i\|_i$

**Proof.** First note that by 11.82 then product topology is generated by the norm  $\|x\| = \max(\|\pi_i(x)\|_i | i \in \{1, \dots, n\}) = \max(\|x_i\|_i | i \in \{1, \dots, n\})$

1.  $(1 \Rightarrow 2)$  this is trivial
2.  $(2 \Rightarrow 3)$  using 11.155 given  $\varepsilon = 1$  the exists a  $\delta \in \mathbb{R}_+$  such that  $\forall x \in \prod_{i \in \{1, \dots, n\}} X_i$  with  $\|x\| = \|x - 0\| < 1$  we have  $\|L(x)\|_Y = \|L(x) - L(0)\|_Y < 1$ . As  $0 < \delta$  there exists a  $0 < \delta' < \delta$  take then  $k = \frac{1}{\delta'}$ . Then  $\forall x \in \prod_{i \in \{1, \dots, n\}} X_i$  with  $\|x\| = \max\{\|x_i\|_i | i \in \{1, \dots, n\}\} = 1$  then  $\|\delta' \cdot x\| = |\delta'| \cdot \|x\| = \delta' \cdot \|x\| < \delta \cdot 1 = \delta \Rightarrow \delta' \cdot \|L(x)\|_Y = |\delta'| \cdot \|L(x)\|_Y = \|\delta' \cdot L(x)\|_Y = \|L(\delta' \cdot x)\|_Y < 1 \Rightarrow \|L(x)\|_Y < \frac{1}{\delta'} = k$ . So  $\forall x \in \prod_{i \in \{1, \dots, n\}} X_i$  with  $\|x\| = 1$  we have  $\|L(x)\|_Y \leq k$

3. **(3  $\Rightarrow$  4)** Let  $k \in \mathbb{R}_+$  such that  $\forall x \in \prod_{i \in \{1, \dots, n\}} X_i$  with  $\|x\| = 1$  we have  $\|L(x)\|_Y \leq k$  then  $\forall x \in \prod_{i \in \{1, \dots, n\}} X_i$  we have to consider the following cases
- $(\exists i \in \{1, \dots, n\} \vdash x_i = 0)$  then by 10.242 we have  $L(x) = 0 \Rightarrow \|L(x)\|_Y = 0 \leq k \cdot \prod_{i \in \{1, \dots, n\}} \|x_i\|_i$
  - $(\forall i \in \{1, \dots, n\} \models x_i \neq 0)$  define then  $y = (y_1, \dots, y_n)$  by  $y_i = \frac{1}{\|x_i\|_i}$  then  $\|y_i\|_i = 1$  and thus  $\frac{1}{\prod_{i \in \{1, \dots, n\}} \|x_i\|_i} \cdot \|L(x_1, \dots, x_n)\|_Y \stackrel{10.55}{=} \left( \prod_{i \in \{1, \dots, n\}} \frac{1}{\|x_i\|_i} \right) \cdot \|L(x_1, \dots, x_n)\|_Y = \left\| \prod_{i \in \{1, \dots, n\}} \frac{1}{\|x_i\|_i} \cdot L(x_1, \dots, x_n) \right\|_Y = \left\| L\left( \frac{1}{\|x_1\|_1} \cdot x_1, \dots, \frac{1}{\|x_n\|_n} \cdot x_n \right) \right\|_Y = \|L(y)\|_Y \leq k \Rightarrow \|L(x)\|_Y \leq k \cdot \prod_{i \in \{1, \dots, n\}} \|x_i\|_i$
4. **(4  $\Rightarrow$  1)** Let  $x \in \prod_{i \in \{1, \dots, n\}} X_i$  and  $\varepsilon \in \mathbb{R}_+$  take then  $0 < \delta < \min(1, \frac{\varepsilon}{k})$  ( $k$  from the condition in (4)) then if  $\|x - y\| < \delta$  then we have  $\forall i \in \{1, \dots, n\}$  we have  $\|x_i - y_i\|_i = \|(x - y)_i\|_i < \delta$  so that  $\|L(x) - L(y)\|_Y = \|L(x - y)\|_Y \leq k \cdot \prod_{i \in \{1, \dots, n\}} \|x_i - y_i\|_i < k \cdot \delta \leq \varepsilon$  [as  $\forall i \in \{1, \dots, n\}$  we have  $\|x_i - y_i\| < \delta < 1$  and 10.63] proving continuity.  $\square$

**Example 11.195.** Let  $\langle \mathbb{K}, \| \cdot \| \rangle$  be the complex (real space) equipped with the absolute norm,  $n \in \mathbb{N}_0$  then

- $\forall (x_1, \dots, x_n) \in \mathbb{K}^n$  we have  $\left| \prod_{i \in \{1, \dots, n\}} x_i \right| = \prod_{i \in \{1, \dots, n\}} |x_i|$
- $g: \mathbb{K}^n \rightarrow \mathbb{K}$  defined by  $g(x_1, \dots, x_n) = \prod_{i \in \{1, \dots, n\}} x_i$  is multilinear and continuous
- Let  $\langle X, \| \cdot \| \rangle$  be a normed space and  $z \in X$  then  $h: \mathbb{K}^n \rightarrow X$  defined by  $z \cdot \prod_{i \in \{1, \dots, n\}} x_i$  is multilinear and continuous

**Proof.**

- We prove this by induction on  $n$  so let  $S = \{n \in \mathbb{N}_0 \mid \text{if } (x_1, \dots, x_n) \in \mathbb{K}^n \text{ then } \left| \prod_{i \in \{1, \dots, n\}} x_i \right| = \prod_{i \in \{1, \dots, n\}} |x_i|\}$  then we have

**1  $\in \mathcal{S}$ .** If  $(x_1) \in \mathbb{K}^n$  then  $\left| \prod_{i \in \{1, \dots, 1\}} x_i \right| = |x_1| = \prod_{i \in \{1, \dots, 1\}} |x_i|$  proving that  $1 \in \mathcal{S}$

**$n \in \mathcal{S} \Rightarrow n + 1 \in \mathcal{S}$ .** then if  $(x_1, \dots, x_{n+1}) \in \mathbb{K}^{n+1}$  then

$$\begin{aligned} \left| \prod_{i \in \{1, \dots, n+1\}} x_i \right| &= \left| x_{n+1} \cdot \prod_{i \in \{1, \dots, n\}} x_i \right| \\ &= |x_{n+1}| \cdot \left| \prod_{i \in \{1, \dots, n\}} x_i \right| \\ &\stackrel{n \in \mathcal{S}}{=} |x_{n+1}| \cdot \prod_{i \in \{1, \dots, n\}} |x_i| \\ &= \prod_{i \in \{1, \dots, n+1\}} |x_i| \end{aligned}$$

proving that  $n + 1 \in \mathcal{S}$

- We have to prove multilinearity and continuity

**multilinearity.** This follows from 10.232.

**continuity.** Let  $(x_1, \dots, x_n) \in \mathbb{K}^n$  then we have  $|g(x_1, \dots, x_n)| = \left| \prod_{i \in \{1, \dots, n\}} x_i \right| = 1 \cdot \prod_{i \in \{1, \dots, n\}} |x_i|$  proving that  $g$  is continuous by 11.194.

- We have to prove multilinearity and continuity

**multilinearity.** For  $n$  we have to consider the following cases:

**$n = 1$ .** Then we have for  $\alpha, \beta \in \mathbb{K}$  and  $(y_1) = (\alpha \cdot x + \beta \cdot y) \in \mathbb{K}^1$  that  $z \cdot \prod_{i \in \{1, \dots, 1\}} y_i = z \cdot y_1 = z \cdot (\alpha \cdot x + \beta \cdot y) = \alpha \cdot (z \cdot x) + \beta \cdot (z \cdot y) = \alpha \cdot h(x) + \beta \cdot h(y)$

**1 < n.** Let  $\alpha, \beta \in \mathbb{K}$  and  $(y_1, \dots, y_n) = (x_1, \dots, x_{i-1}, \alpha \cdot x + \beta \cdot y, x_{i+1}, \dots, x_n)$  then

$$\begin{aligned} h((x_1, \dots, x_{i-1}, \alpha \cdot x + \beta \cdot y, x_{i+1}, \dots, x_n)) &= \\ z \cdot \prod_{i \in \{1, \dots, n\}} y_i &= \\ z \cdot \left( \left( \prod_{i \in \{1, \dots, n\} \setminus \{1\}} y_i \right) \cdot (\alpha \cdot x + \beta \cdot y) \right) &= \\ \alpha \cdot \left( z \cdot \left( x \cdot \prod_{i \in \{1, \dots, n\} \setminus \{1\}} y_i \right) \right) + \beta \cdot \left( z \cdot \left( y \cdot \prod_{i \in \{1, \dots, n\} \setminus \{1\}} y_i \right) \right) &= \\ \alpha \cdot h(x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_n) + \beta \cdot h(x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_n) \end{aligned}$$

**continuity.** Then we have for  $(x_1, \dots, x_n) \in \mathbb{K}^n$  that

$$\begin{aligned} \|h(x_1, \dots, x_n)\| &= \left\| z \cdot \prod_{i \in \{1, \dots, n\}} x_i \right\| \\ &= \left| \prod_{i \in \{1, \dots, n\}} x_i \right| \cdot \|y\| \\ &\stackrel{(1)}{=} \|y\| \cdot \prod_{i \in \{1, \dots, n\}} |x_i| \end{aligned}$$

proving continuity by 11.194.  $\square$

**Theorem 11.196.** Let  $\{\langle X_i, \|\cdot\|_i \rangle\}_{i \in \{1, \dots, n\}}$  be normed vector spaces over the real (complex) field  $\mathbb{K}$ ,  $\langle Y, \|\cdot\|_Y \rangle$  a normed vector space over the field  $\mathbb{K}$  and  $L: \prod_{i \in \{1, \dots, n\}} X_i \rightarrow Y$  a multi-linear function then if we define

$$\begin{aligned} M_{L,s} &= \left\{ k \in \mathbb{R}_+ \mid \forall x \in \prod_{i \in \{1, \dots, n\}} X_i \mid \forall i \in \{1, \dots, n\} \text{ we have } \|x_i\|_i = 1 \text{ then we have } \|L(x)\|_Y \leq k \right\} \\ M_{L,r} &= \left\{ k \in \mathbb{R}_+ \mid \forall x \in \prod_{i \in \{1, \dots, n\}} X_i \text{ we have } \|L(x)\|_Y \leq k \cdot \prod_{i \in \{1, \dots, n\}} \|x_i\|_i \right\} \end{aligned}$$

then  $M_{L,s} = M_{L,r}$  and  $\inf(M_{L,s}) = \inf(M_{L,r})$  exists and is finite (and obviously  $\geq 0$ )

**Proof.** First prove that  $M_{L,s} = M_{L,r}$

1. If  $k \in M_{L,s}$  and  $x \in \prod_{i \in \{1, \dots, n\}} X_i$  then we have two cases :

a.  $(\exists i \in \{1, \dots, n\} \vdash \|x_i\|_i = 0)$  then  $x_i = 0$  and thus using 10.242 we have that  $L(x) = 0 \Rightarrow \|L(x)\|_Y = 0 = k \cdot 0 \stackrel{10.55}{=} k \cdot \prod_{i \in \{1, \dots, n\}} \|x_i\|_i \Rightarrow k \in M_{L,r}$

b.  $(\forall i \in \{1, \dots, n\} \models \|x_i\|_i \neq 0)$  then if we define  $y \in \prod_{i \in \{1, \dots, n\}} X_i$  by  $y_i = \frac{1}{\|x_i\|_i}$  then  $\|y_i\|_i = 1$  and thus  $\frac{1}{\prod_{i \in \{1, \dots, n\}} \|x_i\|_i} \cdot \|L(x)\|_Y \stackrel{10.55}{=} \left( \prod_{i \in \{1, \dots, n\}} \frac{1}{\|x_i\|_i} \right) \cdot \|L(x)\|_Y = \left\| \left( \prod_{i \in \{1, \dots, n\}} \frac{1}{\|x_i\|_i} \right) \cdot L(x) \right\|_Y = \|L(y)\|_Y \leq k \Rightarrow \|L(x)\|_Y \leq k \cdot \prod_{i \in \{1, \dots, n\}} \|x_i\|_i \Rightarrow k \in M_{L,r}$

2. If  $k \in M_{L,r}$  and  $x \in \prod_{i \in \{1, \dots, n\}} X_i$  with  $\|x_i\|_i = 1 \forall i \in I$  then  $\|L(x)\|_Y \leq k \cdot \prod_{i \in \{1, \dots, n\}} \|x_i\|_i = k \cdot 1 = k \Rightarrow k \in M_{L,s}$

Next by 11.194 there exists a  $k \in \mathbb{R}_+$  such that  $\forall x \in \prod_{i \in \{1, \dots, n\}} X_i$  we have  $\|L(x)\|_Y \leq k \cdot \prod_{i \in \{1, \dots, n\}} \|x_i\|_i$  and thus  $k \in M_{L,s} \Rightarrow M_{L,s} \neq \emptyset$ . Using then the fact that the reals are conditionally complete (see 9.43) and 2.180 we have that  $\inf(M_{L,r}) = \inf(M_{L,s})$  exists and  $\inf(M_{L,s}) = \inf(M_{L,r}) \geq 0$  (by the definition of a infimum).  $\square$

**Definition 11.197.** Let  $\{\langle X_i, \|\cdot\|_i \rangle\}_{i \in \{1, \dots, n\}}$  be normed vector spaces over the real (complex) field  $\mathbb{K}$ ,  $\langle Y, \|\cdot\|_Y \rangle$  a normed vector space over the field  $\mathbb{K}$  and  $L: \prod_{i \in \{1, \dots, n\}} X_i \rightarrow Y$  a multi-linear function then we define  $\|L\| = \inf(M_{L,r}) = \inf(M_{L,s})$  to be the operator norm.

**Theorem 11.198.** Let  $\{\langle X_i, \|\cdot\|_i \rangle\}_{i \in \{1, \dots, n\}}$  be normed vector spaces over the real (complex) field  $\mathbb{K}$ ,  $\langle Y, \|\cdot\|_Y \rangle$  a normed vector space over the field  $\mathbb{K}$  and  $L, L_1, L_2: \prod_{i \in \{1, \dots, n\}} X_i \rightarrow Y$  a multi-linear function then we have

1.  $\forall x \in \prod_{i \in \{1, \dots, n\}} X_i$  we have  $\|L(x)\|_Y \leq \|L\| \cdot \prod_{i \in \{1, \dots, n\}} \|x_i\|_i$
2. If  $\forall x \in \prod_{i \in \{1, \dots, n\}} X_i$  we have  $\|L(x)\|_Y \leq A \cdot \prod_{i \in \{1, \dots, n\}} \|x_i\|_i$  then  $\|L\| \leq A$
3. If  $\forall x \in \prod_{i \in \{1, \dots, n\}} X_i$  with  $\|x_i\| = 1$  we have  $\|L(x)\|_Y \leq A$  then  $\|L\| \leq A$
4.  $L_1 + L_2$  is a continuous multilinear mapping and  $\|L_1 + L_2\| \leq \|L_1\| + \|L_2\|$
5. If  $L = C_0$  (the constant 0 function) then  $\|L\| = 0$
6. If  $\|L\| = 0$  then  $L = C_0$
7.  $\forall \alpha \in \mathbb{K}$  we have that  $\alpha \cdot L$  is a continuous multilinear mapping and  $\|\alpha \cdot L\| = |\alpha| \cdot \|L\|$

**Proof.**

1. Let  $x \in \prod_{i \in \{1, \dots, n\}} X_i$  then we have the following possibilities
  - a.  $(\exists i \in \{1, \dots, n\} \vdash x_i = 0)$  then by 10.242 we have  $L(x) = 0 \Rightarrow \|L(x)\|_Y = 0 \leq \|L\| \cdot \prod_{i \in \{1, \dots, n\}} \|x_i\|_i$
  - b.  $(\forall i \in \{1, \dots, n\} \models x_i \neq 0)$  then  $\prod_{i \in \{1, \dots, n\}} \|x_i\|_i > 0$ , assume now that  $\exists x$  so that  $\|L\| \cdot \prod_{i \in \{1, \dots, n\}} \|x_i\|_i < \|L(x)\|_Y \Rightarrow \|L\| < \frac{1}{\prod_{i \in \{1, \dots, n\}} \|x_i\|_i} \cdot \|L(x)\|_Y$  as  $\|L\| = \inf(M_{L,r})$  there exists a  $k \in M_{L,r}$  such that  $\|L\| \leq k < \frac{1}{\prod_{i \in \{1, \dots, n\}} \|x_i\|_i} \cdot \|L(x)\|_Y \leq \frac{1}{\prod_{i \in \{1, \dots, n\}} \|x_i\|_i} \cdot k \cdot \prod_{i \in \{1, \dots, n\}} \|x_i\|_i = k$  giving the contradiction  $k < k$ . We conclude thus that  $\forall x \in \prod_{i \in \{1, \dots, n\}} X_i$  with  $x_i \neq 0 \forall i \in \{1, \dots, n\}$  we have  $\|L(x)\|_Y \leq \|L\| \cdot \prod_{i \in \{1, \dots, n\}} \|x_i\|_i$
2. This follows from the fact that  $\|L\| = \inf(M_{L,r}) = \inf(\{k \in \mathbb{R}_+ \mid \forall x \in \prod_{i \in \{1, \dots, n\}} X_i \text{ we have } \|L(x)\|_Y \leq k \cdot \prod_{i \in \{1, \dots, n\}} \|x_i\|_i\})$
3. This followed from the fact  $\|L\| = \inf(M_{L,s}) = \inf(k \in \mathbb{R}_+ \mid \forall x \in \prod_{i \in \{1, \dots, n\}} X_i \vdash \forall i \in \{1, \dots, n\} \text{ we have } \|x_i\| = 1 \text{ then we have } \|L(x)\|_Y \leq k)$
4. Using 10.235 we have that  $L_1 + L_2$  is multilinear and using 11.175 we have that  $L_1 + L_2$  is continuous. If now  $x \in \prod_{i \in \{1, \dots, n\}} X_i$  then  $\|L_1(x) + L_2(x)\|_Y \leq \|L_1(x)\|_Y + \|L_2(x)\|_Y \leq \|L_1\| \cdot \prod_{i \in \{1, \dots, n\}} \|x_i\|_i + \|L_2\| \cdot \prod_{i \in \{1, \dots, n\}} \|x_i\|_i = (\|L_1\| + \|L_2\|) \cdot \prod_{i \in \{1, \dots, n\}} \|x_i\|_i \Rightarrow \|L_1 + L_2\| \leq \|L_1\| + \|L_2\| \in M_{L_1 + L_2, r} \Rightarrow \|L_1 + L_2\| = \inf(M_{L_1 + L_2, r}) \leq \|L_1\| + \|L_2\|$
5. If  $L = C_0$  then  $\forall x \in \prod_{i \in \{1, \dots, n\}} X_i$  that  $\|L(x)\|_Y = \|C_0(x)\|_Y = \|0\|_Y = 0 \leq 0 \cdot \prod_{i \in \{1, \dots, n\}} \|x_i\|_i$  proving that  $0 \in M_{L,r} \Rightarrow 0 \leq \|L\| = \inf(M_{L,r}) \leq 0 \Rightarrow \|L\| = 0$
6. If  $\|L\| = 0$  then if  $x \in \prod_{i \in \{1, \dots, n\}} X_i$  we have  $\|L(x)\|_Y = 0 \Rightarrow L(x) = 0 \Rightarrow L = C_0$
7. Using 10.235 we have that  $\alpha \cdot L$  is multilinear and using 11.175 we have that  $\alpha \cdot L$  is continuous. If  $x \in \prod_{i \in \{1, \dots, n\}} X_i$  then  $\|\alpha \cdot L(x)\|_Y = |\alpha| \cdot \|L(x)\|_Y \leq |\alpha| \cdot \|L\| \cdot \prod_{i \in \{1, \dots, n\}} \|x_i\|_i \Rightarrow |\alpha| \cdot \|L\| \in M_{\alpha \cdot L, r} \Rightarrow \|\alpha \cdot L\| = \inf(M_{\alpha \cdot L, r}) \leq |\alpha| \cdot \|L\|$  so

$$\|\alpha \cdot L\| \leq |\alpha| \cdot \|L\| \quad (11.29)$$

Assume now that  $\|\alpha \cdot L\| < |\alpha| \cdot \|L\|$  then consider the cases

- a.  $(\alpha = 0)$  then  $\alpha \cdot L = C_0$  and thus  $0 = 0 \cdot \|L\| = |\alpha| \cdot \|L\| > \|\alpha \cdot L\| = \|C_0\| \stackrel{(3)}{=} 0 \Rightarrow 0 > 0$  a contradiction.
- b.  $(\alpha \neq 0)$  then  $|\alpha| \neq 0$  so that  $\|\alpha \cdot L\| < |\alpha| \cdot \|L\|$  as  $\|\alpha \cdot L\| = \inf(A_{\alpha \cdot L, r})$  there exist a  $k \in A_{\alpha \cdot L, r}$  such that  $\|\alpha \cdot L\| \leq k < |\alpha| \cdot \|L\|$  so if  $x \in \prod_{i \in \{1, \dots, n\}} X_i$  then  $|\alpha| \cdot \|L(x)\|_Y = \|\alpha \cdot L(x)\|_Y \leq k \cdot \prod_{i \in \{1, \dots, n\}} \|x_i\|_i \Rightarrow \|L(x)\|_Y \leq \frac{k}{|\alpha|} \cdot \prod_{i \in \{1, \dots, n\}} \|x_i\|_i \Rightarrow \frac{k}{|\alpha|} \in M_{L,r} \Rightarrow \|L\| \leq \frac{k}{|\alpha|} < \frac{|\alpha| \cdot \|L\|}{|\alpha|} = \|L\|$  giving the contradiction  $\|L\| < \|L\|$ .

as the assumption turns up a contradiction in all cases we must have  $|\alpha| \cdot \|L\| \leq \|\alpha \cdot L\|$  which together with 11.29 means that  $\|\alpha \cdot L\| = |\alpha| \cdot \|L\|$   $\square$

From the above theorem we have then trivially the following theorem

**Theorem 11.199.** *Let  $\{\langle X_i, \|\cdot\|_i \rangle\}_{i \in \{1, \dots, n\}}$  be normed vector spaces over the real (complex) field  $\mathbb{K}$ ,  $\langle Y, \|\cdot\|_Y \rangle$  a normed vector space over the real (complex) field  $\mathbb{K}$  then if  $L(X_1, \dots, X_n; Y) = \{L \in \text{Hom}(X_1, \dots, X_n; Y) | L \text{ is continuous}\} \subseteq \text{Hom}(X_1, \dots, X_n; Y)$  [The set of multilinear mappings between  $\prod_{i \in \{1, \dots, n\}} X_i$  and  $Y$ ] then  $L(X_1, \dots, X_n; Y)$  is a subspace of  $\text{Hom}(X_1, \dots, X_n; Y)$  and  $\langle L(X_1, \dots, X_n; Y), \|\cdot\| \rangle$  ( $\|\cdot\|$  the operator norm) forms a normed vector space.*

**Theorem 11.200.** *Let  $\{\langle X_i, \|\cdot\|_i \rangle\}_{i \in \{1, \dots, n\}}$  be normed vector spaces over the real (complex) field  $\mathbb{K}$ ,  $\langle Y, \|\cdot\|_Y \rangle$ ,  $\langle Z, \|\cdot\|_Z \rangle$  normed vector spaces over the real (complex) field  $\mathbb{K}$  then we have if  $L \in L(Y, Z)$  and  $K \in L(X_1, \dots, X_n; Y)$  then  $L \circ K \in L(X_1, \dots, X_n; Z)$  with  $\|L \circ K\|_{L(X_1, \dots, X_n; Z)} \leq \|L\|_{L(Y, Z)} \cdot \|K\|_{L(X_1, \dots, X_n; Y)}$*

**Proof.** First multilinearity follows from 10.236, continuity follows from the fact that the composition of continuous functions is continuous. Finally if  $x = (x_1, \dots, x_n) \in \prod_{i \in \{1, \dots, n\}} X_i$  then  $\|(L \circ K)(x_1, \dots, x_n)\|_Z = \|L(K(x_1, \dots, x_n))\|_Z \leq_{L \text{ is linear}} \|L\|_{L(X_1, \dots, X_n; Y)} \cdot \|K(x_1, \dots, x_n)\|_Y \leq_{K \text{ is multilinear}} \|L\|_{L(Y, Z)} \cdot \|K\|_{L(X_1, \dots, X_n; Y)} \cdot \prod_{i \in \{1, \dots, n\}} \|x_i\|_i$   $\stackrel{11.198}{\Rightarrow} \|L \circ K\|_{L(X_1, \dots, X_n; Z)} \leq \|L\|_{L(Y, Z)} \cdot \|K\|_{L(X_1, \dots, X_n; Y)}$   $\square$

We prove now that the operator norm depends only on the topology of  $\langle X_i, \|\cdot\|_i \rangle$  and  $\langle Y, \|\cdot\|_Y \rangle$  in the following theorem.

**Theorem 11.201.** *Let  $\{\langle X_i, \|\cdot\|_{1,i} \rangle\}_{i \in \{1, \dots, n\}}$ ,  $\{\langle X_i, \|\cdot\|_{2,i} \rangle\}_{i \in \{1, \dots, n\}}$ ,  $\langle Y, \|\cdot\|_{1,Y} \rangle$ ,  $\langle Y, \|\cdot\|_{2,Y} \rangle$  be normed vector spaces over the real (complex) field  $\mathbb{K}$ , with  $\forall i \in \{1, \dots, n\} \|\cdot\|_{1,i}$  is equivalent with  $\|\cdot\|_{2,i}$  and  $\|\cdot\|_{1,Y}$  are equivalent then if  $\|\cdot\|_1$  is the operator norm based on  $\{\langle X_i, \|\cdot\|_{1,i} \rangle\}$ ,  $\langle Y, \|\cdot\|_1 \rangle$  and  $\|\cdot\|_2$  is the operator norm based on  $\{\langle X_i, \|\cdot\|_{2,i} \rangle\}$ ,  $\langle Y, \|\cdot\|_2 \rangle$  then  $\|\cdot\|_1$  and  $\|\cdot\|_2$  are equivalent.*

**Proof.** Given  $L \in L(X_1, \dots, X_n; Y)$  let  $A_{1,L} = \{k \in \mathbb{R}_+ | \forall x \in \prod_{i \in \{1, \dots, n\}} X_i \text{ we have } \|L(x)\|_{Y_1} \leq k \cdot \prod_{i \in \{1, \dots, n\}} \|x_i\|_{1,i}\}$  and  $A_{2,L} = \{k \in \mathbb{R}_+ | \forall x \in \prod_{i \in \{1, \dots, n\}} X_i \text{ we have } \|L(x)\|_{Y_2} \leq k \cdot \prod_{i \in \{1, \dots, n\}} \|x_i\|_{2,i}\}$  then  $\|L\|_1 = \inf(A_{1,L})$  and  $\|L\|_2 = \inf(A_{2,L})$ . By the hypothesis and the definition of equivalence of norms (see 11.85) there exists  $\{(\alpha_i, \beta_i)\}_{i \in \{1, \dots, n\}}$  and  $\alpha, \beta > 0$  such that  $\forall i \in \{1, \dots, n\}$  we have  $\alpha_i, \beta_i > 0$  and  $\forall x \in X_i$  that  $\alpha_i \cdot \|x\|_{2,i} \leq \|x\|_{1,i} \leq \beta_i \cdot \|x\|_{1,i}$ . Also  $\forall y \in Y$  we have  $\alpha \cdot \|y\|_{2,Y} \leq \|y\|_{1,Y} \leq \beta \cdot \|y\|_{2,Y}$ . Then we have

$$1. \forall k \in A_{1,L} \text{ we have } \forall x \in \prod_{i \in \{1, \dots, n\}} X_i \text{ that } \|L(x)\|_{1,Y} \leq k \cdot \prod_{i \in \{1, \dots, n\}} \|x_i\|_{1,i} \stackrel{\beta_i > 0, \|\cdot\|_{1,i} \geq 0}{\Rightarrow} \alpha \cdot \|L(x)\|_{2,Y} \leq \|L(x)\|_1 \leq k \cdot \prod_{i \in \{1, \dots, n\}} \|x_i\|_{1,i} \leq (k \cdot \prod_{i \in \{1, \dots, n\}} \beta_i) \cdot$$

$$\prod_{i \in \{1, \dots, n\}} \|x\|_{2,i} \stackrel{\alpha > 0}{\Rightarrow} \|L(x)\|_{2,Y} \leq \left( \frac{\prod_{i \in \{1, \dots, n\}} \beta_i}{\alpha} \cdot k \right) \cdot \prod_{i \in \{1, \dots, n\}} \|x_i\|_{2,i} \text{ proving that } \left( \frac{\prod_{i \in \{1, \dots, n\}} \beta_i}{\alpha} \cdot k \right) \in A_{2,L} \text{ so that } \inf(A_{2,L}) \leq \left( \frac{\prod_{i \in \{1, \dots, n\}} \beta_i}{\alpha} \cdot k \right) \Rightarrow \frac{\alpha}{\prod_{i \in \{1, \dots, n\}} \beta_i} \cdot \inf(A_{2,L}) \leq k. \text{ Assume that } \inf(A_{1,L}) < \frac{\alpha}{\prod_{i \in \{1, \dots, n\}} \beta_i} \cdot \inf(A_{2,L}) \text{ then } \exists k \in A_{1,L} \text{ such that } \inf(A_{1,L}) \leq k < \frac{\alpha}{\prod_{i \in \{1, \dots, n\}} \beta_i} \cdot \inf(A_{2,L}) \leq k \text{ a contradiction so that } \frac{\alpha}{\prod_{i \in \{1, \dots, n\}} \beta_i} \cdot \|L\|_2 = \frac{\alpha}{\prod_{i \in \{1, \dots, n\}} \beta_i} \cdot \inf(A_{2,L}) \leq \inf(A_{1,L}) = \|L\|_1$$

$$2. \forall k \in A_{2,L} \text{ we have } \forall x \in \prod_{i \in \{1, \dots, n\}} X_i \text{ that } \|L(x)\|_{2,Y} \leq k \cdot \prod_{i \in \{1, \dots, n\}} \|x_i\|_{2,i} \stackrel{\frac{1}{\beta} \cdot \|L(x)\|_{1,Y} \leq \|L(x)\|_{2,Y} \leq k \cdot \prod_{i \in \{1, \dots, n\}} \|x_i\|_{2,i}}{\Rightarrow} \frac{k}{\prod_{i \in \{1, \dots, n\}} \alpha_i} \cdot \prod_{i \in \{1, \dots, n\}} \|x_i\|_{1,i} \Rightarrow \|L(x)\|_{1,Y} \leq \left( \frac{\beta}{\prod_{i \in \{1, \dots, n\}} \alpha_i} \right) \cdot k \cdot \prod_{i \in \{1, \dots, n\}} \|x_i\|_{1,i} \text{ proving that } k \cdot \frac{\beta}{\prod_{i \in \{1, \dots, n\}} \alpha_i} \in A_{1,L} \text{ so that } \inf(A_{1,L}) \leq k \cdot \frac{\beta}{\prod_{i \in \{1, \dots, n\}} \alpha_i} \stackrel{\beta > 0}{\Rightarrow} \frac{\prod_{i \in \{1, \dots, n\}} \alpha_i}{\beta} \cdot \inf(A_{1,L}) \leq k. \text{ Assume that } \inf(A_{2,L}) < \frac{\prod_{i \in \{1, \dots, n\}} \alpha_i}{\beta} \cdot \inf(A_{1,L}) \Rightarrow \exists k \in A_{2,L} \text{ such that } \inf(A_{2,L}) \leq k < \frac{\prod_{i \in \{1, \dots, n\}} \alpha_i}{\beta} \cdot \inf(A_{1,L}) \leq k \text{ a contradiction so that } \frac{\prod_{i \in \{1, \dots, n\}} \alpha_i}{\beta} \cdot \|L\|_1 = \frac{\prod_{i \in \{1, \dots, n\}} \alpha_i}{\beta} \cdot \inf(A_{1,L}) \leq \inf(A_{2,L}) = \|L\|_2$$

From (1),(2) and the fact that  $\frac{\alpha}{\prod_{i \in \{1, \dots, n\}} \beta_i} > 0$ ,  $\frac{\beta}{\prod_{i \in \{1, \dots, n\}} \alpha_i} > 0$  we have  $\frac{\alpha}{\prod_{i \in \{1, \dots, n\}} \beta_i} \cdot \|L\|_2 \leq \|L\|_1 \leq \frac{\beta}{\prod_{i \in \{1, \dots, n\}} \alpha_i} \cdot \|L\|_2$  proving the equivalency of  $\|\cdot\|_1$  and  $\|\cdot\|_2$ .  $\square$

**Notation 11.202.** If  $\{\langle X_i, \|\cdot\|_i \rangle\}_{i \in \{1, \dots, n\}}, n \in \mathbb{N}_0$  be family of normed spaces, where  $\forall i \in \{1, \dots, n\}$  we have  $\langle X_i, \|\cdot\|_i \rangle = \langle X, \|\cdot\|_X \rangle$  a normed space,  $\langle Y, \|\cdot\|_Y \rangle$  then  $\langle L(X_1, \dots, X_n; Y), \|\cdot\| \rangle$  is noted as  $\langle L(X^n; Y), \|\cdot\| \rangle$  where  $L \in L(X^n; Y) \Rightarrow L: X^n = \prod_{i \in \{1, \dots, n\}} X \rightarrow Y$  is multilinear and continuous and has norm  $\|L\|$ . Note that  $L(X^1; Y) = L(X, Y)$  by definition.

**Theorem 11.203.** Let  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle$  and  $\langle Z, \|\cdot\|_Z \rangle$  be normed spaces over the real (complex) field  $\mathbb{K}$  then the composition function  $\circ: L(Y, Z) \times L(X, Y) \rightarrow L(X, Z)$  defined by  $\circ(L_2, L_1) = L_2 \circ L_1$  is multilinear and continuous (using the canonical norms) with  $\|\circ\| = 1$ . So  $\circ \in L(L(Y, Z), L(X, Y); L(X, Z))$

**Proof.** First note that the composition of continuous linear mappings is again linear and continuous so that if  $(L_2, L_1) \in L(Y, Z) \times L(X, Y)$  then  $\circ(L_2, L_1) = L_2 \circ L_1 \in L(X, Z)$  proving that  $\circ$  is indeed a mapping from  $L(Y, Z) \times L(X, Y)$  to  $L(X, Z)$ .

To prove multilinearity note:

1. Let  $\alpha, \beta \in \mathbb{K}, L_1, L_2 \in L(Y, Z)$  and  $L \in L(X, Y)$  then  $\forall x \in X$  we have  $\circ(\alpha \cdot L_1 + \beta \cdot L_2, L)(x) = ((\alpha \cdot L_1 + \beta \cdot L_2) \circ L)(x) = (\alpha \cdot L_1 + \beta \cdot L_2)(L(x)) = \alpha \cdot L_1(L(x)) + \beta \cdot L_2(L(x)) = \alpha \cdot \circ(L_1, L)(x) + \beta \cdot \circ(L_2, L)(x)$  so that  $\circ(\alpha \cdot L_1 + \beta \cdot L_2) = \alpha \cdot \circ(L_1, L) + \beta \cdot \circ(L_2, L)$
2. Let  $\alpha, \beta \in \mathbb{K}, L \in L(Y, Z)$  and  $L_1, L_2 \in L(X, Y)$  then  $\forall x \in X$  we have  $\circ(L, \alpha \cdot L_1 + \beta \cdot L_2)(x) = L((\alpha \cdot L_1 + \beta \cdot L_2)(x)) = L(\alpha \cdot L_1(x) + \beta \cdot L_2(x)) = \alpha \cdot L(L_1(x)) + \beta \cdot L(L_2(x)) = \alpha \cdot \circ(L, L_1)(x) + \beta \cdot \circ(L, L_2)(x) = (\alpha \cdot \circ(L, L_1) + \beta \cdot \circ(L, L_2))(x)$  so that  $\circ(L, \alpha \cdot L_1 + \beta \cdot L_2) = \alpha \cdot \circ(L, L_1) + \beta \cdot \circ(L, L_2)$

Finally to prove continuity note  $\|\cdot\|_1, \|\cdot\|_2, \|\cdot\|_3$  as the operator norm of  $L(Y, Z), L(X, Y), L(X, Z)$  then  $\forall (L_1, L_2) \in L(Y, Z) \times L(X, Y)$  then using 11.191 we have that  $\|L_1 \circ L_2\|_3 \leq \|L_1\|_1 \cdot \|L_2\|_2 = 1 \cdot \prod_{i \in \{1, \dots, 2\}} \|L_i\|_i$  proving that  $\circ$  is continuous and that  $\|\circ\| \leq 1$   $\square$

**Definition 11.204.** Let  $n \in \mathbb{N}_0 \{\langle X_i, \|\cdot\|_i \rangle\}_{i \in \{1, \dots, n\}}$  a finite family of normed spaces,  $\langle Y, \|\cdot\|_Y \rangle$  a normed space then  $L^n(X_1, \dots, X_n; Y)$  is defined by

1. If  $n = 1$  then  $L^n(X_1, \dots, X_1; Y) = L(X_1, Y)$
2. If  $n > 1$  then  $L^n(X_1, \dots, X_n; Y) = L(X_1, L^{n-1}(X_2, \dots, X_n; Y))$

**Example 11.205.**

1.  $L^1(X_1; Y) = L(X_1; Y)$
2.  $L^2(X_1, X_2; Y) = L(X_1, L(X_2, Y))$
3.  $L^3(X_1, \dots, X_3) = L(X_1, L^2(X_2, X_3; Y)) = L(X_1, L(X_2, L^1(X_3, \dots, X_3; Y))) = L(X_1, L(X_2, L(X_3, Y)))$
4. ...

**Definition 11.206.** Let  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle$ , be normed vector spaces,  $n \in \mathbb{N}_0$  then  $L^n(X; Y)$  is defined as  $L^n(X_1, \dots, X_n; Y)$  where the family  $\{X_i\}_{i \in \{1, \dots, n\}}$  is defined by  $\forall i \in \{1, \dots, n\} X_i = X$

1.  $n = 1 \Rightarrow L^1(X; Y) = L(X, Y)$
2.  $n > 1 \Rightarrow L^n(X; Y) = L(X, L^{n-1}(X; Y))$

**Example 11.207.**

1.  $L^1(X; Y) = L(X, Y)$
2.  $L^2(X; Y) = L(X, L^1(X; Y)) = L(X, L(X, Y))$
3.  $L^3(X; Y) = L(X, L^2(X; Y)) = L(X, L(X, L(X, Y)))$
4. ...

**Theorem 11.208.** Let  $n \in \mathbb{N}_0 \{\langle X_i, \|\cdot\|_i \rangle\}_{i \in \{1, \dots, n+1\}}$  a finite family of normed spaces,  $\langle Y, \|\cdot\|_Y \rangle$  a normed space then  $L^n(X_1, \dots, X_n; L(X_{n+1}, Y)) = L^{n+1}(X_1, \dots, X_{n+1}; Y)$

**Proof.** We prove this by induction, so let  $S = \{n \in \mathbb{N}_0 \mid \text{for } \{\langle X_i, \|\cdot\|_i \rangle\}_{i \in \{1, \dots, n+1\}} \text{ we have } L^n(X_1, \dots, X_n; L(X_{n+1}, Y)) = L^{n+1}(X_1, \dots, X_{n+1}; Y)\}$  then:

1. If  $n = 1$  then  $L^1(X_1, \dots, X_1; L(X_2, Y)) = L(X_1, L(X_2, Y)) = L^2(X_1, \dots, X_2; Y)$  proving  $1 \in S$
2. If  $n \in S$  then  $L^{n+1}(X_1, \dots, X_{n+1}; L(X_{n+2}, Y)) = L(X_1, L^n(X_2, \dots, X_{n+1}; L(X_{n+2}, Y))) \stackrel{n \in S}{=} L(X_1, L^{n+1}(X_2, \dots, X_{n+2}; Y)) \stackrel{\text{def}}{=} L^{(n+1)+1}(X_1, \dots, X_{n+2}; Y) \Rightarrow n+1 \in S$

using mathematical induction proves the theorem.  $\square$

A easy corrolary of the above theorem is then

**Corollary 11.209.** Let  $n \in \mathbb{N}_0$ ,  $\langle X, \|\cdot\|_X \rangle$  a normed space,  $\langle Y, \|\cdot\|_Y \rangle$  a normed vector space and  $n \in \mathbb{N}_0$  then  $L^n(X; L(X, Y)) = L^{n+1}(X; Y)$

**Definition 11.210.** Let  $n \in \mathbb{N}_0$   $\{\langle X_i, \|\cdot\|_i \rangle\}_{i \in \{1, \dots, n+1\}}$  a finite family of normed spaces,  $\langle Y, \|\cdot\|_Y \rangle$  a normed vector space,  $m \in \mathbb{N}_0$  with  $m \leq n$  then if  $L \in L^n(X_1, \dots, X_n; Y)$  and  $x = (x_1, \dots, x_m) \in \prod_{i \in \{1, \dots, m\}} X_i$  we define  $L(x_1: \dots: x_m)$  recursively by

1.  $m = 1$  then  $L(x_1: \dots: x_1) = L(x_1)$
2.  $m > 1$  then  $L(x_1: \dots: x_m) = (L(x_1: \dots: x_{m-1}))(x_m)$

To show that the above is well defined we have the following theorem

**Theorem 11.211.** Let  $n \in \mathbb{N}_0$   $\{\langle X_i, \|\cdot\|_i \rangle\}_{i \in \{1, \dots, n\}}$  a finite family of normed spaces,  $\langle Y, \|\cdot\|_Y \rangle$  a normed vector space,  $m \in \mathbb{N}_0$  with  $m \leq n$  then if  $L \in L^n(X_1, \dots, X_n; Y)$  and  $(x_1, \dots, x_m) \in \prod_{i \in \{1, \dots, m\}} X_i$  then if  $m < n$  we have that  $L(x_1: \dots: x_m) \in L^{n-m}(X_{m+1}, \dots, X_n; Y)$  and if  $m = n$  then  $L(x_1: \dots: x_m) \in Y$

**Proof.** We have two cases to consider

1. ( $m < n$ ) we proceed by induction so let  $S_n = \{m \in \mathbb{N}_0 \mid \text{if } m < n \text{ then } L(x_1: \dots: x_m) \in L^{n-m}(X_{m+1}, \dots, X_n; Y)\}$  then we must consider the following cases
  - a. if  $m = 1$  then if  $1 < n$  we have  $L(x_1: \dots: x_m) = L(x_1: \dots: x_1) = L(x_1)$  and as  $L \in L^n(X_1, \dots, X_n; Y) = L(X_1, L^{n-1}(X_2, \dots, X_n; Y))$  we have that  $L(x_1) \in L^{n-1}(X_2, \dots, X_n; Y)$  proving that  $1 \in S_n$
  - b. if  $m \in S_n$  and  $m+1 < n \Rightarrow m < n$  then  $L(x_1: \dots: x_m) \in L^{n-m}(X_{m+1}, \dots, X_n; Y) \stackrel{m < n \Rightarrow n-m > 1}{=} L(X_{m+1}, L^{(n-m)-1}(X_{m+2}, \dots, X_n; Y))$  so that if  $m+1 < n$  then  $L(x_1: \dots: x_{m+1}) = L(x_1: \dots: x_m)(x_{m+1}) \in L^{(n-m)-1}(X_{m+2}, \dots, X_n; Y) = L^{n-(m+1)}(X_{m+1}, \dots, X_n; Y)$  proving that  $m+1 \in S_n$

By mathematical induction we have then that  $S_n = \mathbb{N}_0$  so if  $m < n \Rightarrow m \in S_n \wedge m < n$  so that  $L(x_1: \dots: x_m) \in L^{n-m}(X_{m+1}, \dots, X_n; Y)$

2. ( $m = n$ ) Two cases must be considered

- a. ( $n = 1$ ) then  $L(x_1: \dots: x_1) = L(x_1) \in Y$  as  $L \in L^1(X_1, \dots, X_1; Y) = L(X_1, Y)$
- b. ( $n > 1$ ) then as  $n-1 < n$  we have by (1)  $L(x_1: \dots: x_{n-1}) \in L^{n-(n-1)}(X_n, \dots, X_n; Y) = L^1(X_n, \dots, X_n; Y) = L(X_n, Y)$  so that  $L(x_1, \dots, x_n) = L(x_1, \dots, x_{n-1})(x_n) \in Y$   $\square$

The next theorem shows that  $(x_1: \dots: x_n)$  behaves the same as  $(x_1, \dots, x_n)$  functions.

**Lemma 11.212.** Let  $n \in \mathbb{N}_0$   $\{\langle X_i, \|\cdot\|_i \rangle\}_{i \in \{1, \dots, n\}}$  a finite family of normed spaces,  $\langle Y, \|\cdot\|_Y \rangle$  a normed vector space and  $L, L' \in L^n(X_1, \dots, X_n; Y)$  then if  $\forall (x_1, \dots, x_n) \in \prod_{i \in \{1, \dots, n\}} X_i$  we have  $L(x_1: \dots: x_n) = L'(x_1: \dots: x_n)$  then  $L = L'$ .

**Proof.** We prove this by induction. So let  $S_n = \{m \in \mathbb{N}_0 \mid \text{if } m \leq n \text{ and } \forall (x_1, \dots, x_m) \in \prod_{i \in \{1, \dots, m\}} X_i \text{ we have } L(x_1: \dots: x_m) = L'(x_1: \dots: x_m) \text{ then } L = L'\}$  then we have

1. if  $m = 1$  and  $\forall (x_1) \in \prod_{i \in \{1, \dots, 1\}} X_i$  we have  $L(x_1) = L'(x_1)$  we have  $\forall x \in X_1$  that  $(x) \in \prod_{i \in \{1, \dots, m\}} X_i$  and thus  $L(x) = L'(x)$  so that  $L = L'$  and thus  $1 \in S_n$

2. if  $m \in S_n$  then if  $m+1 \leq n$  and we have  $\forall (x_1, \dots, x_{m+1}) \in \prod_{i \in \{1, \dots, m+1\}} X_i$  we have  $L(x_1: \dots: x_{m+1}) = L'(x_1: \dots: x_{m+1}) \Rightarrow L(x_1: \dots: x_m)(x_{m+1}) = L'(x_1: \dots: x_m)(x_{m+1})$   $\xrightarrow{x_{m+1} \text{ is chosen arbitrary}} L(x_1: \dots: x_m) = L'(x_1: \dots: x_m) \underset{m \in S_n}{=} L = L'$  proving that  $m+1 \in S_n$   $\square$

Finally we show that  $L(x_1)(x_2: \dots: x_n) = L(x_1: \dots: x_n)$

**Theorem 11.213.** Let  $n \in \mathbb{N}_0 \setminus \{1\}$ ,  $\{\langle X_i, \|\cdot\|_i \rangle\}_{i \in \{1, \dots, n\}}$  a finite family of normed spaces,  $\langle Y, \|\cdot\|_Y \rangle$  a normed vector space and  $L \in L^n(X_1, \dots, X_n; Y)$  then  $\forall (x_1, \dots, x_n) \in \prod_{i \in \{1, \dots, n\}} X_i$  we have  $(L(x_1))(x_2: \dots: x_n) = L(x_1: \dots: x_n)$

**Proof.** We proof this by induction so let  $S = \{m \in \{2, \dots\} \mid \text{if } m \leq n \text{ then } (L(x_1))(x_2: \dots: x_m) = L(x_1: \dots: x_m)\}$  then we have

1. If  $m = 2$  then if  $m \leq n$  we have  $(L(x_1))(x_2: \dots: x_m) = (L(x_1))(x_2: \dots: x_2) = (L(x_1))(x_2) = L(x_1: \dots: x_2)$  proving that  $2 \in S$
2. If  $m \in S$  then if  $m+1 \leq n$  we have  $(L(x_1))(x_2: \dots: x_{m+1}) = ((L(x_1))(x_2: \dots: x_m))(x_{m+1}) \underset{m \in S}{=} (L(x_1: \dots: x_m))(x_{m+1}) = L(x_1: \dots: x_{m+1})$  proving that  $m+1 \in S$

So by induction we have that  $S = \{2, \dots\}$  so if  $m = n > 1 \Rightarrow m \in \{2, \dots\} = S \Rightarrow L(x_1)(x_2: \dots: x_n) = L(x_1)(x_2: \dots: x_m) = L(x_1: \dots: x_m) = L(x_1: \dots: x_n)$ .  $\square$

**Corollary 11.214.** Let  $Y$  a vector space over  $\mathbb{K}$ .  $n \in \mathbb{N}_0$ ,  $\{v_i\}_{i \in \{1, \dots, n\}} \subseteq \mathbb{K}$  /where  $\mathbb{K}$  is the field  $\mathbb{R}$  or  $\mathbb{C}$ / and  $L \in L^n(\mathbb{K}; Y)$  then  $L(v_1: \dots: v_n) = L\left(\underbrace{1: \dots: 1}_n\right) \cdot \prod_{i \in \{1, \dots, n\}} v_i$

**Proof.** We prove this by induction so let  $\mathcal{S} = \left\{ n \in \mathbb{N}_0 \mid \text{if } L \in L^n(\mathbb{K}; Y), \{v_i\}_{i \in \{1, \dots, n\}} \subseteq \mathbb{K} \text{ then } L(v_1: \dots: v_n) = L\left(\underbrace{1: \dots: 1}_n \cdot \prod_{i \in \{1, \dots, n\}} v_i\right) \right\}$  then we have:

$1 \in \mathcal{S}$ . If  $\{v_i\}_{i \in \{1, \dots, 1\}}$  then  $v_1 = v_1 = v_1 \cdot 1$  hence if  $L \in L^1(\mathbb{K}; Y) = L(\mathbb{K}, Y)$  so my linearity we have  $L(v_1) = L(v_1 \cdot 1) = L(1) \cdot v_1 = L\left(\underbrace{1: \dots: 1}_1 \cdot \prod_{i \in \{1, \dots, 1\}} v_i\right)$  proving that  $1 \in \mathcal{S}$

$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$ . Let  $L \in L^n(\mathbb{K}, Y)$  and  $\{v_i\}_{i \in \{1, \dots, n+1\}} \subseteq \mathbb{K}$  then we have

$$\begin{aligned}
 L(v_1: \dots: v_{n+1}) &\stackrel{11.213}{=} L(v_1)(v_2, \dots, v_{n+1}) \\
 &= L(v_1)(v_{(1+1)}, \dots, v_{(n+1)}) \\
 &\stackrel{n \in \mathcal{S} \wedge L(v_1) \in L^n(\mathbb{K}; Y)}{=} L(v_1)\left(\underbrace{1: \dots: 1}_n\right) \cdot \prod_{i \in \{1, \dots, n\}} v_{i+1} \\
 &\stackrel{\text{linearity}}{=} L(1)\left(\underbrace{1: \dots: 1}_n\right) \cdot \left(\prod_{i \in \{1, \dots, 1\}} v_i\right) \cdot \prod_{i \in \{2, \dots, n+1\}} v_i \\
 &\stackrel{11.213}{=} L\left(\underbrace{1: \dots: 1}_{n+1}\right) \cdot \prod_{i \in \{1, \dots, n+1\}} v_i
 \end{aligned}$$

proving that  $n+1 \in \mathcal{S}$ .  $\square$

**Lemma 11.215.** Let  $n \in \mathbb{N}_0 \setminus \{1\}$ ,  $\{\langle X_i, \|\cdot\|_i \rangle\}_{i \in \{1, \dots, n\}}$  a finite family of normed spaces,  $\langle Y, \|\cdot\|_Y \rangle$  a normed vector space  $\alpha, \beta \in \mathbb{K}$ ,  $L_1, L_2 \in L^n(X_1, \dots, X_n; Y)$ ,  $x_1, \dots, x_n \in X$  then  $(\alpha \cdot L_1 + \beta \cdot L_2)(x_1: \dots: x_n) = \alpha \cdot L_1(x_1: \dots: x_n) + \beta \cdot L_2(x_1: \dots: x_n)$

**Proof.** We prove this by induction so let  $S_n = \{m \in \mathbb{N}_0 \mid \text{if } m < n \text{ then } (\alpha \cdot L_1 + \beta \cdot L_2)(x_1: \dots: x_n) = \alpha \cdot L_1(x_1: \dots: x_n) + \beta \cdot L_2(x_1: \dots: x_n)\}$  then

1. if  $m = 1$   $(\alpha \cdot L_1 + \beta \cdot L_2)(x_1: \dots: x_1) = (\alpha \cdot L_1 + \beta \cdot L_2)(x_1) = \alpha \cdot L_1(x_1) + \beta \cdot L_2(x_1) = \alpha \cdot L_1(x_1: \dots: x_1) + \beta \cdot L_2(x_1: \dots: x_1)$  so  $1 \in S_n$ .
2. if  $m \in S$  then  $(\alpha \cdot L_1 + \beta \cdot L_2)(x_1: \dots: x_{n+1}) = ((\alpha \cdot L_1 + \beta \cdot L_2)(x_1: \dots: x_n))(x_{n+1}) \underset{m \in S}{=} (\alpha \cdot L_1(x_1: \dots: x_n) + \beta \cdot L_2(x_1: \dots: x_n))(x_{n+1}) = \alpha \cdot L_1(x_1: \dots: x_n)(x_{n+1}) + \beta \cdot L_2(x_1: \dots: x_n)(x_{n+1}) = \alpha \cdot L_1(x_1: \dots: x_{n+1}) + \beta \cdot L_2(x_1: \dots: x_{n+1})$

Using induction we have  $S_n = \mathbb{N}_0$  proving the theorem.  $\square$

**Lemma 11.216.** Let  $n \in \mathbb{N}_0 \setminus \{1\}$ ,  $\{\langle X_i, \|\cdot\|_i \rangle\}_{i \in \{1, \dots, n\}}$  a finite family of normed spaces,  $\langle Y, \|\cdot\|_Y \rangle$  a normed vector,  $L \in L^n(X_1, \dots, X_n; Y)$  and  $(x_1, \dots, x_n) \in \prod_{i \in \{1, \dots, n\}} X_i$  then  $\|L(x_1: \dots: x_n)\|_Y \leq \|L\| \cdot \prod_{i \in \{1, \dots, n\}} \|x_i\|_i$

**Proof.** We must consider the following cases for  $n$

1. ( $n = 1$ ) Then as  $L \in L^1(X_1; Y) = L(X_1, Y)$  we have  $\forall (x_1) \in \prod_{i \in \{1, \dots, 1\}} X_i$  that  $\|L(x_1)\|_Y \leq \|L\| \cdot \|x_1\|_1 = \|L\| \cdot \prod_{i \in \{1, \dots, 1\}} \|x_i\|_i$  proving the theorem in the case of  $n = 1$
2. ( $n > 1$ ) then if we note the norms in  $L^{n-i}(X_{i+1}, \dots, X_n; Y)$ ,  $i \leq n$  by  $\|\cdot\|_{(i)}$  and define  $T_n = \{m \in \mathbb{N}_0 \mid m \leq n-1 \text{ then } \forall (x_1, \dots, x_m) \in \prod_{i \in \{1, \dots, m\}} X_m \text{ we have } \|L(x_1: \dots: x_m)\|_{(m)} \leq \|L\| \cdot \prod_{i \in \{1, \dots, m\}} \|x_i\|_i\}$  then we have:
  - a. If  $m = 1$  then as  $n \geq 2$  we have  $m < n-1$  and  $\|L(x_1)\|_{(1)} \leq \|L\| \cdot \|x_1\|_1 = \|L\| \cdot \prod_{i \in \{1, \dots, 1\}} \|x_i\|_i \Rightarrow 1 \in T_n$
  - b. If  $m \in T_n$  then if  $m+1 \leq n-1$  and  $(x_1, \dots, x_{m+1}) \in \prod_{i \in \{1, \dots, m+1\}} X_i$  then  $\|L(x_1: \dots: x_{m+1})\|_{(m+1)} = \|L(x_1: \dots: x_m)(x_{m+1})\|_{(m+1)} \leq \|L(x_1: \dots: x_m)\|_{(m)} \cdot \|x_{m+1}\|_{m+1} \leq_{m \in T_n} (\|L\| \cdot \prod_{i \in \{1, \dots, m\}} \|x_i\|_i) = \|L\| \cdot \prod_{i \in \{1, \dots, m+1\}} \|x_i\|_i$  proving that  $m+1 \in T_n$

Using mathematical induction we have that  $T_n = \mathbb{N}_0$  so for  $n-1 \in \mathbb{N}_0 = T_n$  we have as  $n-1 \leq n-1$  that  $\|L(x_1: \dots: x_{n-1})\|_{(n-1)} \leq \|L\| \cdot \prod_{i \in \{1, \dots, n-1\}} \|x_i\|_i$  so that  $\|L(x_1: \dots: x_n)\|_Y = \|L(x_1: \dots: x_{n-1})(x_n)\| \leq \|L(x_1: \dots: x_{n-1})\|_{(n-1)} \cdot \|x_n\|_n \leq (\|L\| \cdot \prod_{i \in \{1, \dots, n-1\}} \|x_i\|_i) \cdot \|x_n\|_n = \|L\| \cdot \prod_{i \in \{1, \dots, n\}} \|x_i\|_i$  proving the theorem for  $n > 1$   $\square$

**Theorem 11.217.** Let  $\langle \mathbb{K}, \|\cdot\| \rangle$ ,  $\langle Y, \|\cdot\|_Y \rangle$  be normed spaces (where  $\mathbb{K} = \mathbb{C}$  or  $\mathbb{R}$  with the absolute norm),  $n \in \mathbb{N}_0$  and  $L \in L^n(\mathbb{K}, Y)$  then  $\|L\|_{L^n(\mathbb{K}; Y)} \leq \|L(1: \dots: 1)\|_Y$

**Proof.** We prove this by induction on  $n$ , denote  $\|\cdot\|_i = \|\cdot\|_{L^i(\mathbb{K}; Y)}$  the operator norm in  $L^i(\mathbb{K}; Y)$  and let  $S = \{n \in \mathbb{N}_0 \mid L \in L^n(\mathbb{K}, Y)\}$  then  $\|L\|_i \leq \|L(1: \dots: 1)\|_Y$  then we have:

1. If  $n = 1$  then  $L \in L^1(\mathbb{K}; Y) = L(\mathbb{K}, Y)$  then if  $|x_1| = 1$  we have as  $x_1 = x_1 \cdot 1$  that  $\|L(x_1)\|_Y = \|x_1 \cdot L(1)\|_Y = |x_1| \cdot \|L(1)\|_Y = \|L(1)\|_Y \leq \|L(1)\|_Y$  and as  $\|L\|_1 = \inf\{k \in \mathbb{R}_+ \mid \forall x \in \mathbb{K} \text{ such that } |x| = 1 \text{ we have } \|L(x)\|_Y \leq k\}$  that  $\|L\|_1 \leq \|L(1)\|_Y$  proving that  $1 \in S$
2. If  $n \in S$  then as  $L^{n+1}(\mathbb{K}; Y) = L(\mathbb{K}, L^n(\mathbb{K}; Y))$  we have if  $L \in L^{n+1}(\mathbb{K}; Y)$  that  $L(1) \in L^n(\mathbb{K}; Y)$  so that as  $n \in S$  we have that

$$\|L(1)\|_n \leq \left\| L(1) \underbrace{(1: \dots: 1)}_n \right\|_Y = \left\| L \underbrace{(1: \dots: 1)}_{n+1} \right\|_Y \quad (11.30)$$

now if  $x \in \mathbb{K}$  is such that  $|x| = 1$  then  $\|L(x)\|_Y = \|L(x \cdot 1)\|_Y = \|x \cdot L(1)\|_Y = |x| \cdot \|L(1)\|_Y = \|L(1)\|_Y$  which as  $\|L\|_{n+1} = \inf\{k \in \mathbb{R}_+ \mid \forall x \in \mathbb{K} \text{ such that } |x| = 1 \text{ we have } \|L(x)\|_Y \leq k\}$  we have that  $\|L\|_{n+1} \leq \|L(1)\|_Y$  proving using 11.31 that  $\|L\|_{n+1} \leq \left\| L \underbrace{(1: \dots: 1)}_{n+1} \right\|_Y$  and thus that  $n+1 \in S$

Mathematical induction then proves that  $S = \mathbb{N}_0$  proving the theorem.  $\square$

**Definition 11.218.** Let  $n \in \mathbb{N}_0$ ,  $\{\langle X_i, \|\cdot\|_i \rangle\}_{i \in \{1, \dots, n\}}$  a finite family of normed spaces,  $\langle Y, \|\cdot\|_Y \rangle$  a normed vector then given  $L \in L^n(X_1, \dots, X_n; Y)$  we define  $\mathcal{P}_n(L): \prod_{i \in \{1, \dots, n\}} X_i \rightarrow Y$  by  $\mathcal{P}_n(L)(x_1, \dots, x_n) = L(x_1: \dots: x_n)$ . Note that if  $n = 1$  then  $L^1(X_1, \dots, X_1; Y) = L(X_1, Y) = L(X_1, \dots, X_1; Y)$  and then  $\mathcal{P}_n(L) = 1_{L(X_1, Y)}$ .

**Theorem 11.219.** Let  $n \in \mathbb{N}_0$ ,  $\{\langle X_i, \|\cdot\|_i \rangle\}_{i \in \{1, \dots, n+1\}}$  a finite family of normed spaces,  $\langle Y, \|\cdot\|_Y \rangle$  a normed vector then given  $L \in L^{n+1}(X_1, \dots, X_{n+1}; Y) \stackrel{11.208}{=} L^n(X_1, \dots, X_n; L(X_{n+1}, Y))$  and  $(x_1, \dots, x_{n+1}) \in \prod_{i \in \{1, \dots, n+1\}} X_i$  then  $\mathcal{P}_{n+1}(L)(x_1, \dots, x_{n+1}) = \mathcal{P}_n(L)(x_1, \dots, x_n)(x_{n+1})$

**Proof.**  $\mathcal{P}_{n+1}(L)(x_1, \dots, x_{n+1}) = L(x_1: \dots: x_{n+1}) = L(x_1: \dots: x_n)(x_{n+1}) = \mathcal{P}_n(L)(x_{n+1})$   $\square$

**Theorem 11.220.** Let  $n \in \mathbb{N}_0$ ,  $\{\langle X_i, \|\cdot\|_i \rangle\}_{i \in \{1, \dots, n+1\}}$  a finite family of normed spaces,  $\langle Y, \|\cdot\|_Y \rangle$  a normed vector and  $L \in L^n(X_1, \dots, X_n; Y)$  then  $\mathcal{P}_n(L)$  is multilinear and continuous (so that  $\mathcal{P}_n(L) \in L(X_1, \dots, X_n; Y)$  where  $\|\mathcal{P}_n(L)\|_0 \leq \|L\|$  (where  $\|\cdot\|$  is the operator norm of  $L^n(X_1, \dots, X_n; Y) = \begin{cases} L(X_1, Y) & \text{if } n=1 \\ L(X_1, L^{n-1}(X_2, \dots, X_n; Y)) & \text{if } n > 1 \end{cases}$ ) and  $\|\cdot\|_0$  the operator norm on  $L(X_1, \dots, X_n; Y)$ .

**Proof.** We divide the proof in different phases

1. **(multilinearity)** Let  $S = \{m \in \mathbb{N}_0 \mid \text{if } m \leq n \text{ then if } i \in \{1, \dots, m\} \text{ we have } \forall \alpha, \beta \in \mathbb{K}, \forall x, y \in X_i, \forall z \in \prod_{j \in \{1, \dots, n\} \setminus \{i\}} X_j \text{ that } \mathcal{P}_m(L)(z_1, \dots, z_{i-1}, \alpha \cdot x + \beta \cdot y, z_{i+1}, \dots, z_m) = \alpha \cdot \mathcal{P}_m(L)(z_1, \dots, z_{i-1}, x, z_{i+1}, \dots, z_m) + \beta \cdot \mathcal{P}_m(z_1, \dots, z_{i-1}, y, z_{i+1}, \dots, z_m)\}$  then we have

a. If  $m = 1$  (so  $m \leq n$ ) then as  $\mathcal{P}_1(L) = L$  and as  $L$  is linear we have  $1 \in S$

b. If  $m \in S$  then if  $m+1 \leq n$  we have two cases to consider for  $i \in \{1, \dots, m+1\}$

i. **( $i \leq m$ )** then as  $m \leq n$  we have  $\mathcal{P}_{m+1}(L)(z_1, \dots, z_{i-1}, \alpha \cdot x + \beta \cdot y, z_{i+1}, \dots, z_{m+1}) \stackrel{11.219}{=} \mathcal{P}_m(L)(z_1, \dots, z_{i-1}, \alpha \cdot x + \beta \cdot y, z_{i+1}, \dots, z_m)(z_{m+1}) \stackrel{m \in S \wedge m < n}{=} (\alpha \cdot \mathcal{P}_m(L)(z_1, \dots, z_{i-1}, x, z_{i+1}, \dots, z_m) + \beta \cdot \mathcal{P}_m(z_1, \dots, z_{i-1}, y, z_{i+1}, \dots, z_m))(z_{m+1}) = \alpha \cdot \mathcal{P}_m(L)(z_1, \dots, z_{i-1}, x, z_{i+1}, \dots, z_m)(z_{m+1}) + \beta \cdot \mathcal{P}_m(z_1, \dots, z_{i-1}, y, z_{i+1}, \dots, z_m)(z_{m+1}) = \alpha \cdot \mathcal{P}_{m+1}(L)(z_1, \dots, z_{i-1}, x, z_{i+1}, \dots, z_{m+1}) + \beta \cdot \mathcal{P}_{m+1}(z_1, \dots, z_{i-1}, y, z_{i+1}, \dots, z_{m+1})$

ii. **( $i = m+1$ )** then  $\mathcal{P}_{m+1}(L)(z_1, \dots, z_{i-1}, \alpha \cdot x + \beta \cdot y, z_{i+1}, \dots, z_n) = \mathcal{P}_{m+1}(L)(z_1, \dots, z_m, \alpha \cdot x + \beta \cdot y) \stackrel{11.219}{=} \mathcal{P}_m(L)(z_1, \dots, z_m)(\alpha \cdot x + \beta \cdot y) = L(z_1: \dots: z_m)(\alpha \cdot x + \beta \cdot y) \stackrel{L(z_1: \dots: z_m) \in L^{n-m}(X_{m+1}, \dots, X_n; Y)}{=} \alpha \cdot L(z_1: \dots: z_m)(x) + \beta \cdot L(z_1: \dots: z_m)(y) = \alpha \cdot L(z_1: \dots: z_m; x) + \beta \cdot L(z_1: \dots: z_m; y) = \alpha \cdot \mathcal{P}_{m+1}(L)(z_1, \dots, z_{i-1}, x, z_{i+1}, \dots, z_{m+1}) + \beta \cdot \mathcal{P}_{m+1}(z_1, \dots, z_{i-1}, y, z_{i+1}, \dots, z_{m+1})$

(i) and (2) proves then that  $m+1 \in S$

By mathematical induction we have then that  $S = \mathbb{N}_0$  so as  $n \in \mathbb{N}_0 = S$  and  $n \leq n$  we have proved multilinearity.

2. **(continuity)** If  $(x_1, \dots, x_n) \in \prod_{i \in \{1, \dots, n\}} X_i$  then we have by 11.216 that  $\|\mathcal{P}(L)(x_1, \dots, x_n)\|_Y = \|L(x_1: \dots: x_n)\|_Y \leq \|L\| \cdot \prod_{i \in \{1, \dots, n\}} \|x_i\|_i$  proving that as  $\mathcal{P}(L)$  is multilinear that  $\mathcal{P}(L)$  is continuous and that  $\|\mathcal{P}(L)\| \leq \|L\|$  (see 11.198)  $\square$

**Theorem 11.221.** Let  $n \in \mathbb{N}_0 \setminus \{1\}$ ,  $\{\langle X_i, \|\cdot\|_i \rangle\}_{i \in \{1, \dots, n\}}$  a finite family of normed spaces,  $\langle Y, \|\cdot\|_Y \rangle$  a normed vector and  $L \in L^n(X_1, \dots, X_n; Y)$  then  $\mathcal{P}_n: L^n(X_1, \dots, X_n; Y) \rightarrow L(X_1, \dots, X_n; Y)$  defined by  $\mathcal{P}_n(L)(x_1, \dots, x_n) = L(x_1: \dots: x_n)$  (see 11.218) is a norm preserving isomorphism.

**Remark 11.222.** As  $\mathcal{P}_n$  is norm preserving we have that  $\forall L \in L^n(X_1, \dots, X_n; Y)$  that  $\|\mathcal{P}(L)\|_{L^n(X_1, \dots, X_n; Y)} = \|L\|_{L(X_1, \dots, X_n; Y)}$  which together with the linearity of  $\mathcal{P}$  proves by 11.180 continuity so that that  $\mathcal{P}_n \in L(L^n(X_1, \dots, X_n; Y), L(X_1, \dots, X_n; Y))$ . Further using 11.184 we have that  $\|\mathcal{P}_n\|_{L(L^n(X_1, \dots, X_n; Y), L(X_1, \dots, X_n; Y))} \leq 1$ .

**Proof.** The proof is done in different steps

1. **(injectivity)** If  $\mathcal{P}_n(L) = \mathcal{P}_n(L')$  then  $\forall x_1, \dots, x_n \in X$  we have  $L(x_1: \dots: x_n) = \mathcal{P}_n(L) = \mathcal{P}_n(L') = L'(x_1: \dots: x_n) \stackrel{11.212}{\Rightarrow} L = L'$  proving injectivity.
2. **(surjectivity)** Let  $S = \{n \in \mathbb{N}_0 \mid \{\langle X_i, \|\cdot\|_i \rangle\}_{i \in \{1, \dots, n+1\}}$  is a family of normed spaces and  $\langle Y, \|\cdot\|_Y \rangle$  is a normed space then  $\mathcal{P}_n: L^n(X_1, \dots, X_n; Y) \rightarrow L(X_1, \dots, X_n; Y)$  is a surjection} then we have
  - a. If  $n = 1$  then  $L^1(X_1; Y) = L(X_1, Y) = L(X_1, \dots, X_1; Y)$  and  $\mathcal{P}_1 = 1_{L(X_1, \dots, X_1; Y)}$  which is clearly surjective so that  $1 \in S$

- b. If  $n \in S$  then let  $L \in L(X_1, \dots, X_{n+1}; Y)$  then  $\forall (x_1, \dots, x_n) \in \prod_{i \in \{1, \dots, n\}} X_i$  define then  $L_1: \prod_{i \in \{1, \dots, n\}} X_i \rightarrow L(X_{n+1}, Y)$  by  $(x_1, \dots, x_n) \mapsto L_1(x_1, \dots, x_n)$  where  $L_1(x_1, \dots, x_n)(x) = L(x_1, \dots, x_n, x)$ . Of course we must prove that  $L_1(x_1, \dots, x_n) \in L(X_{n+1}, Y)$
- (linearity)** If  $\alpha, \beta \in \mathbb{K}$  and  $x, y \in X_{n+1}$  then  $L_1(x_1, \dots, x_n)(\alpha \cdot x + \beta \cdot y) = L(x_1, \dots, x_n, \alpha \cdot x + \beta \cdot y) \stackrel{L \text{ is multilinear}}{=} \alpha \cdot L(x_1, \dots, x_n, x) + \beta \cdot L(x_1, \dots, x_n, y) = \alpha \cdot L_1(x_1, \dots, x_n)(x) + \beta \cdot L_1(x_1, \dots, x_n)(y)$  proving that  $L_1(x_1, \dots, x_n)$  is linear.
  - (continuity)** Let  $x \in X_{n+1}$  then  $\|L_1(x_1, \dots, x_n)(x_{n+1})\|_Y = \|L(x_1, \dots, x_n, x)\|_Y \leq (\|L\| \cdot \prod_{i \in \{1, \dots, n\}} \|x_i\|_i) \cdot \|x\|$  proving continuity.

Further the mapping  $L_1$  itself is multilinear and continuous

- (multilinearity)** If  $i \in \{1, \dots, n\}$ ,  $x, y \in X_i$ ,  $\alpha, \beta \in \mathbb{K}$ ,  $(z_1, \dots, z_n) \in \prod_{j \in \{1, \dots, n\} \setminus \{i\}} X_j$  and  $w \in X_{n+1}$  then  $L_1(z_1, \dots, z_{i-1}, \alpha \cdot x + \beta \cdot y, z_{i+1}, \dots, z_n)(w) = L(z_1, \dots, z_{i-1}, \alpha \cdot x + \beta \cdot y, z_{i+1}, \dots, z_n, w) \stackrel{L \text{ is multilinear}}{=} \alpha \cdot L(z_1, \dots, z_{i-1}, x, z_{i+1}, \dots, z_n, w) + \beta \cdot L(z_1, \dots, z_{i-1}, y, z_{i+1}, \dots, z_n, w) = \alpha \cdot L_1(z_1, \dots, z_{i-1}, x, z_{i+1}, \dots, z_n)(w) + \beta \cdot L_1(z_1, \dots, z_{i-1}, y, z_{i+1}, \dots, z_n)(w)$  proving as  $w$  is choosen arbitrary we have that  $L_1(z_1, \dots, z_{i-1}, \alpha \cdot x + \beta \cdot y, z_{i+1}, \dots, z_n) = \alpha \cdot L_1(z_1, \dots, z_{i-1}, x, z_{i+1}, \dots, z_n) + \beta \cdot L_1(z_1, \dots, z_{i-1}, y, z_{i+1}, \dots, z_n)$
- (continuity)**  $\forall (x_1, \dots, x_n) \in \prod_{i \in \{1, \dots, n\}} X_i$ ,  $\forall x \in X_{n+1}$  we have that  $\|L_1(x_1, \dots, x_n)(x)\|_Y = \|L(x_1, \dots, x_n, x)\|_Y \leq (\|L\| \cdot \prod_{i \in \{1, \dots, n\}} \|x_i\|_i) \cdot \|x\|_{n+1}$  so that  $\|L_1(x_1, \dots, x_n)\|_{L(X_{n+1}, Y)} \leq \|L\| \cdot \prod_{i \in \{1, \dots, n\}} \|x_i\|$  and thus proving that  $L_1$  is continuous.

So we conclude that  $L_1 \in L(X_1, \dots, X_n; L(X_{n+1}, Y))$  then as  $n \in S$  there exists a  $L_2 \in L^n(X_1, \dots, X_n; L(X_{n+1}, Y)) \stackrel{11.208}{=} \mathcal{P}_n(L_2) = L_1$ . Then  $\forall (x_1, \dots, x_{n+1}) \in \prod_{i \in \{1, \dots, n+1\}} X_i$  we have  $L(x_1, \dots, x_{n+1}) = L_1(x_1, \dots, x_n)(x_{n+1}) = (\mathcal{P}_n(L_2)(x_1, \dots, x_n))(x_{n+1}) = L_2(x_1, \dots, x_n)(x_{n+1}) = L_2(x_1, \dots, x_{n+1}) = \mathcal{P}_{n+1}(L_2)(x_1, \dots, x_{n+1})$  proving that  $L = \mathcal{P}_{n+1}(L_2)$  and thus that  $\mathcal{P}_{n+1}$  is surjective.

- ( $\mathcal{P}_n$  is linear)** So let  $\alpha, \beta \in \mathbb{K}$  and  $L_1, L_2 \in L^n(X_1, \dots, X_n; Y)$  then  $\forall (x_1, \dots, x_n) \in \prod_{i \in \{1, \dots, n\}} X_i$  we have  $(\mathcal{P}_n(\alpha \cdot L_1 + \beta \cdot L_2))(x_1, \dots, x_n) = (\alpha \cdot L_1 + \beta \cdot L_2)(x_1, \dots, x_n) \stackrel{11.215}{=} \alpha L_1(x_1, \dots, x_n) + \beta \cdot L_2(x_1, \dots, x_n) = \alpha \cdot \mathcal{P}_n(L_1)(x_1, \dots, x_n) + \beta \cdot \mathcal{P}_n(L_2)(x_1, \dots, x_n) = (\alpha \cdot \mathcal{P}_n(L_1) + \beta \cdot \mathcal{P}_n(L_2))(x_1, \dots, x_n)$  proving that  $\mathcal{P}_n(\alpha \cdot L_1 + \beta \cdot L_2) = \alpha \cdot \mathcal{P}_n(L_1) + \beta \cdot \mathcal{P}_n(L_2)$
- (Norm preserving)** By the previous theorem we have already that if  $L \in L^n(X_1, \dots, X_n; Y)$  that  $\|\mathcal{P}_n(L)\|_{L(X_1, \dots, X_n; Y)} \leq \|L\|$ . We prove now by induction that  $\|L\| \leq \|\mathcal{P}_n(L)\|_{L(X_1, \dots, X_n; Y)} \leq \|L\|$ . Define given  $i \in \{0, \dots, n\}$  that  $\|\cdot\|_{(i)} = \begin{cases} \|\cdot\|_Y & \text{if } i = 0 \\ \|\cdot\|_{L^i(X_{n-(i-1)}, \dots, X_n; Y)} & \text{if } i > 0 \end{cases}$  and take then  $S_n = \{m \in \mathbb{N} \mid \text{if } m < n \text{ then } \forall (x_1, \dots, x_{n-m}) \in \prod_{i \in \{1, \dots, n-m\}} X_i \text{ with } \forall i \in \{1, \dots, n-m\} \mid \|x_i\|_i = 1 \text{ we have } \|L(x_1, \dots, x_{n-m})\|_{(m)} \leq \|\mathcal{P}_n(L)\|_{L(X_1, \dots, X_n; Y)}\}$  then we have :
  - If  $m = 0$  then  $m < n$  and  $n - m = n$  so if  $(x_1, \dots, x_n) \in \prod_{i \in \{1, \dots, n-m\}} X_i$  with  $\forall i \in \{1, \dots, n\} \mid \|x_i\|_i = 1$  we have  $\|L(x_1, \dots, x_n)\|_0 = \|L(x_1, \dots, x_n)\|_Y = \|\mathcal{P}_n(L)(x_1, \dots, x_n)\|_Y \leq \|\mathcal{P}_n(L)\|_{L(X_1, \dots, X_n; Y)} \cdot \prod_{i \in \{1, \dots, n\}} \|x_i\|_i = \|\mathcal{P}_n(L)\|_{L(X_1, \dots, X_n; Y)}$  proving that  $0 \in S_n$
  - If  $m \in S_n$  then if  $m+1 < n \Rightarrow 0 < n - (m+1)$  we have  $\forall (x_1, \dots, x_{n-(m+1)}) \in \prod_{i \in \{1, \dots, n-(m+1)\}} X_i$  with  $\|x_i\|_i = 1$  if  $i \in \{1, \dots, n - (m+1)\}$  we have using 11.187 that  $\|L(x_1, \dots, x_{n-(m+1)})\|_{(m+1)} = \sup(\{\|L(x_1, \dots, x_{n-(m+1)})(x)\|_{(m)} \mid x \in X_{n-m} \text{ with } \|x\|_{n-m} = 1\}) = \sup(\{\|L(x_1, \dots, x_{n-(m+1)})(x)\|_{(m)} \mid x \in X_{n-m} \text{ with } \|x\|_{n-m} = 1\}) \leq \|\mathcal{P}_n(L)\|_{L(X_1, \dots, X_n; Y)}$  [as  $m \in S$  we have that as  $\forall z = (x_1, \dots, x_{n-(m+1)}, x) \in \prod_{i \in \{1, \dots, n-m\}} X_i$  that  $\|z_i\|_i = 1$  that  $\|L(x_1, \dots, x_{n-(m+1)}, x)\|_{(m)} = \|L(z_1, \dots, z_{n-m})\|_{(m)} \leq \|\mathcal{P}_n(L)\|_{L(X_1, \dots, X_n; Y)}$ ] so that  $m+1 \in S_n$ .

By induction we have then that  $S_n = \mathbb{N} \Rightarrow n-1 \in S$  so as  $n-1 < n$  we have that  $\forall x \in X_1 \models \|x\|_1 = 1 \Rightarrow \forall (x) \in \prod_{i \in \{1, \dots, n-(n-1)\}} X_i$  with  $\|x\|_1 = 1$  that  $\|L(x_1)\|_{(n-1)} = \|L(x_1: \dots: x_1)\|_{(n-1)} \leq \|\mathcal{P}_n(L)\|$ . As  $\|L\| = \sup(\|L(x)\|_{(n-1)} | x \in X_1 \models \|x\|_1 = 1)$  we have that  $\|L\| \leq \|\mathcal{P}_n(L)\|$ . This together with  $\|\mathcal{P}_n(L)\|_{L(X_1, \dots, X_n; Y)} \leq \|L\|$  proves that  $\|\mathcal{P}_n(L)\|_{L(X_1, \dots, X_n; Y)} = \|L\|$ .  $\square$

The above theorem lets us to identify  $L^n(X; Y)$  with  $L(X^n; Y)$  which will be used for higher order differentiation.

**Definition 11.223.** Let  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle, \langle Z, \|\cdot\|_Z \rangle$  be normed spaces,  $n \in \mathbb{N}_0$  and  $L \in L(Y, Z)$  then given  $K \in L^n(X; Y)$  we have  $\mathcal{P}_{n,Y}(K) \in L(X^n; Y) \Rightarrow L \circ \mathcal{P}_{n,Y}(K) \in L(X^n; Z) \Rightarrow \mathcal{P}_{n,Z}^{-1}(L \circ \mathcal{P}_{n,Y}(K)) \in L^n(X; Z)$  we define then  $L \bullet_n K = \mathcal{P}_{n,Z}^{-1}(L \circ \mathcal{P}_{n,Y}(K))$ . If  $n=1$  then  $L \bullet_1 K = L \circ K$

**Theorem 11.224.** Let  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle, \langle Z, \|\cdot\|_Z \rangle$  be normed spaces and  $L \in L(Y, Z)$  then given  $n \in \mathbb{N}_0$  then  $K \in L^n(X; Y)$  we have if  $x_1, \dots, x_n \in X$  that  $(L \bullet_n K)(x_1: \dots: x_n) = L(K(x_1: \dots: x_n))$

**Proof.**  $(L \bullet_n K)(x_1: \dots: x_n) = \mathcal{P}_{n,Z}^{-1}(L \circ \mathcal{P}_{n,Y}(K))(x_1: \dots: x_n) \stackrel{\text{definition of } \mathcal{P}_{n,Z}}{=} (L \circ \mathcal{P}_{n,Y}(K))(x_1, \dots, x_n) = L(\mathcal{P}_{n,Y}(K)(x_1, \dots, x_n)) = L(K(x_1: \dots: x_n))$   $\square$

**Theorem 11.225.** Let  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle, \langle Z, \|\cdot\|_Z \rangle$  be normed spaces and  $L \in L(Y, Z)$  then given  $\text{ev}_{L,n}: L^n(X; Y) \rightarrow L^n(X; Z)$  defined by  $K \rightarrow L \bullet_n K$  is a linear continuous function.

**Proof.**

1. **(Linearity)** Let  $K_1, K_2 \in L^n(X; Y)$  and  $\alpha, \beta \in \mathbb{K}$  then we have that  $\text{ev}_L(\alpha \cdot K_1 + \beta \cdot K_2) = L \bullet_n (\alpha \cdot K_1 + \beta \cdot K_2) = \mathcal{P}_{n,Z}^{-1}(L \circ \mathcal{P}_{n,Y}(\alpha \cdot K_1 + \beta \cdot K_2)) \stackrel{\mathcal{P}_{n,Y} \text{ is linear (see 11.221)}}{=} \mathcal{P}_{n,Z}^{-1}(L \circ (\alpha \cdot \mathcal{P}_{n,Y}(K_1) + \beta \cdot \mathcal{P}_{n,Y}(K_2))) \stackrel{L \text{ is linear}}{=} \mathcal{P}_{n,Z}^{-1}(\alpha \cdot (L \circ \mathcal{P}_{n,Y}(K_1)) + \beta \cdot (L \circ \mathcal{P}_{n,Y}(K_2))) \stackrel{\mathcal{P}_{n,Z} \text{ is a linear isomorphism + 10.183}}{=} \alpha \cdot \mathcal{P}_{n,Z}^{-1}(L \circ \mathcal{P}_{n,Y}(K_1)) + \beta \cdot \mathcal{P}_{n,Z}^{-1}(L \circ \mathcal{P}_{n,Y}(K_2)) = \alpha \cdot \text{ev}_L(K_1) + \beta \cdot \text{ev}_L(K_2)$
2. **(continuity)** Let  $K \in L^n(X, Y)$  then  $\|\text{ev}_L(K)\|_{L^n(X; Y)} = \|\mathcal{P}_{n,Z}^{-1}(L \circ \mathcal{P}_{n,Y}(K))\|_{L^n(X; Z)} \stackrel{\mathcal{P}_{n,Z} \text{ is norm preserving}}{=} \|\mathcal{P}_{n,Z}(\mathcal{P}_{n,Z}^{-1}(L \circ \mathcal{P}_{n,Y}(K)))\|_{L(X^n; Z)} = \|L \circ \mathcal{P}_{n,Y}(K)\|_{L(X^n; Z)} \stackrel{\text{11.200}}{\leq} \|L\|_{L(Y, Z)} \cdot \|\mathcal{P}_{n,Y}(K)\| \stackrel{\mathcal{P}_{n,Y} \text{ is norm preserving}}{=} \|L\|_{L(X, Z)} \cdot \|K\|_{L^n(X; Y)}$  proving continuity.  $\square$

The next theorem shows how  $\text{ev}_{L,n}$  can be used to go one level higher

**Theorem 11.226.** Let  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle, \langle Z, \|\cdot\|_Z \rangle$  be normed spaces and  $L \in L(Y, Z)$  then given  $n \in \mathbb{N}_0$  we have if  $K \in L^{n+1}(X; Y)$  that  $\text{ev}_{L,n} \circ K = L \bullet_{n+1} K$

**Proof.** Let  $x_1, \dots, x_{n+1} \in X$  then  $K(x_1) \in L(X, L^n(X; Y))$  so that  $(\text{ev}_{L,n} \circ K)(x_1) = \text{ev}_L(K(x_1)) \stackrel{\text{definition of } \text{ev}_{L,n}}{=} L \bullet_n K(x_1)$  so that  $(\text{ev}_{L,n} \circ K)(x_1: \dots: x_{n+1}) = (\text{ev}_{L,n} \circ K)(x_1)(x_2: \dots: x_n) = (L \bullet_n K(x_1)) \stackrel{\text{11.224}}{=} L(K(x_1)(x_2: \dots: x_{n+1})) = L(K(x_1: \dots: x_{n+1})) \stackrel{\text{11.224}}{=} (L \bullet_{n+1} K)(x_1: \dots: x_{n+1})$  so that  $(\text{ev}_{L,n} \circ K)(x_1: \dots: x_{n+1}) = (L \bullet_{n+1} K)(x_1: \dots: x_{n+1})$  so as  $x_1, \dots, x_n \in X$  is choosen arbitrary we have by 11.212 that  $\text{ev}_{L,n} \circ K = L \bullet_{n+1} K$ .  $\square$

## 11.8 Separation

**Definition 11.227.** A topological space  $\langle X, \mathcal{T} \rangle$  is **Hausdorff** if  $\forall x, y \in X$  with  $x \neq y$  there exists a  $U, V \in \mathcal{T}$  with  $x \in U \wedge y \in V \wedge U \cap V = \emptyset$

**Example 11.228.** A metric space  $\langle X, d \rangle$  is Hausdorff in the metric topology

**Proof.** If  $x, y \in X$  with  $x \neq y$  then  $\varepsilon = d(x, y) > 0$  define then  $x \in B_d(x, \frac{\varepsilon}{2})$  and  $y \in B_d(y, \frac{\varepsilon}{2})$ . If now  $z \in B_d(x, \frac{\varepsilon}{2}) \cap B_d(y, \frac{\varepsilon}{2}) \Rightarrow \varepsilon = d(x, y) \leq d(x, z) + d(y, z) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \Rightarrow \varepsilon < \varepsilon$  a contradiction so that  $B_d(x, \frac{\varepsilon}{2}) \cap B_d(y, \frac{\varepsilon}{2}) = \emptyset$   $\square$

**Theorem 11.229.** *Every finite subset of a Hausdorff space  $\langle X, \mathcal{T} \rangle$  is closed.*

**Proof.** First note that the finite set  $\emptyset$  is closed as  $X \setminus \emptyset = X$  is open. Second we prove that if  $x \in X$  then  $\{x\}$  is closed or equivalent that  $X \setminus \{x\}$  is open, so let  $y \in X \setminus \{x\} \Rightarrow y \neq x \Rightarrow \exists U, V$  open with  $x \in U, y \in V$  and  $U \cap V = \emptyset \Rightarrow x \notin V \Rightarrow y \in V \subseteq X \setminus \{x\} \Rightarrow X \setminus \{x\}$  is open. If now  $F$  is a finite non empty set then there exists a bijection  $b: \{0, \dots, n\} \rightarrow F$  so that  $F = \bigcup_{i \in \{1, \dots, n\}} \{b(i)\} \Rightarrow F$  is closed as a finite set of closed sets is closed.  $\square$

**Definition 11.230.** *A topological space  $\langle X, \mathcal{T} \rangle$  is **regular** if for every closed set  $A$  and  $\forall x \in X$  with  $x \notin A$  there exists disjoint open sets  $U, V \in \mathcal{T}$  with  $x \in U, A \subseteq V$  and  $U \cap V = \emptyset$*

**Theorem 11.231.** *A metric space  $\langle X, d \rangle$  is regular*

**Proof.** Let  $A$  be a closed set and  $x \notin A$  then as  $x \in X \setminus A$  (which is open) there exists a  $U$  open with  $x \in U \subseteq X \setminus A$  so that  $x \in U$  and  $U \cap A = \emptyset$ . As  $x \in U$  open there exist a  $\delta \in \mathbb{R}_+$  such that  $x \in B_d(x, \delta) \subseteq U \Rightarrow B_d(x, \delta) \cap A = \emptyset$ . Given  $a \in A$  assume that  $z \in B_d(x, \frac{\delta}{2}) \cap B_d(a, \frac{\delta}{2}) \Rightarrow d(x, z) < \frac{\delta}{2} \wedge d(z, a) < \frac{\delta}{2} \Rightarrow d(x, a) \leq d(x, z) + d(z, a) < \frac{\delta}{2} + \frac{\delta}{2} = \delta \Rightarrow a \in B_d(x, \delta) \Rightarrow a \in B_d(x, \delta) \cap A = \emptyset$  a contradiction so we must have that  $B_d(x, \frac{\delta}{2}) \cap B_d(a, \frac{\delta}{2}) = \emptyset$  and thus as  $A \subseteq \bigcup_{a \in A} B_d(a, \frac{\delta}{2}) = V$  a open set with  $V \cap B_d(x, \frac{\delta}{2}) = \bigcup_{a \in A} (B_d(a, \frac{\delta}{2}) \cap B_d(x, \frac{\delta}{2})) = \bigcup_{a \in A} \emptyset = \emptyset$  so  $A \subseteq V$  a open set that does not intersect the open set  $B_d(x, \frac{\delta}{2})$  containing  $x$ .  $\square$

**Theorem 11.232.** *Let  $\langle X, \mathcal{T} \rangle$  be a regular topological space and  $U$  a open set with  $x \in U$  then there exists a open set  $V$  with  $x \in V \wedge V \subseteq U$*

**Proof.** Take  $x \in U$  then as  $x \notin X \setminus U$  a closed set we have by regularity a  $V, W$  open with  $x \in V, X \setminus U \subseteq W$  and  $V \cap W = \emptyset \Rightarrow V \subseteq X \setminus W \xrightarrow{X \setminus W \text{ is closed}} \bar{V} \subseteq X \setminus W \xrightarrow{X \setminus U \subseteq W} \bar{V} \subseteq X \setminus (X \setminus U) = U$  and  $x \in V$ .  $\square$

**Definition 11.233.** *A topological space  $\langle X, \mathcal{T} \rangle$  is **normal** if for every pair of disjoint closed sets  $A, B$  there exists a pair of disjoint open sets  $U, V$  with  $A \subseteq U \wedge B \subseteq V$*

**Theorem 11.234.** *A normal topological space where every singleton is closed is regular, a regular space where every singleton is closed is Hausdorff*

**Proof.**

1. Let  $\langle X, \mathcal{T} \rangle$  be a normal space so that  $\forall x \in X$  we have  $\{x\}$  is closed. If  $A$  is a closed set and  $x \notin A$  then as  $\{x\}$  is closed there exists by normality open sets  $U, V$  with  $A \subseteq U, \{x\} \subseteq V \Rightarrow x \in V$  such that  $U \cap V = \emptyset$  proving that  $\langle X, \mathcal{T} \rangle$  is regular.
2. Let  $\langle X, \mathcal{T} \rangle$  be a regular space so that  $\forall x \in X$  we have  $\{x\}$  is closed. If now  $x, y \in X$  with  $x \neq y \Rightarrow y \notin \{x\}$  a closed set so that by regularity there exists disjoint open sets  $U, V$  with  $y \in U$  and  $\{x\} \subseteq V \Rightarrow x \in V$  proving Hausdorff.  $\square$

**Definition 11.235.** *Let  $\langle X, \mathcal{T} \rangle$  be a topological space and  $x \in X$  then a set  $A$  is a neighborhood of  $x$  iff there exists a open set  $U$  such that  $x \in U \subseteq A$ . If  $A$  is also a open set then it is called a open neighborhood.*

**Definition 11.236.** *Let  $\langle X, \mathcal{T} \rangle$  be a topological space,  $x \in X$  then a fundamental system of neighborhoods of  $x$  is a set  $\mathcal{N}$  of neighborhoods of  $x$  such that for every neighborhood  $A$  of  $x$  there exists a  $N \in \mathcal{N}$  such that  $x \in N \subseteq A$*

**Definition 11.237.** *A topological space  $\langle X, \mathcal{T} \rangle$  is first countable if every element of  $X$  has a countable fundamental system of neighborhoods.*

**Theorem 11.238.** *A metric space  $\langle X, d \rangle$  is first countable*

**Proof.** Given  $x \in X$  define the countable set  $\mathcal{N}_x = \{B_d(x, \frac{1}{n}) | n \in \mathbb{N}_0\}$  then if  $A$  is a neighborhood of  $x$  there exists a open set  $U$  with  $x \in U \subseteq A$ . So there exists a  $\varepsilon \in \mathbb{R}_+$  such that  $x \in B_d(x, \varepsilon)$ . Using 9.56 there exists a  $n \in \mathbb{N}_0$  such that  $0 < \frac{1}{n} < \varepsilon$  so that  $x \in B_d(x, \frac{1}{n}) \subseteq B_d(x, \varepsilon) \subseteq U \subseteq A$  so that  $\mathcal{N}_x$  is a fundamental system of neighborhoods of  $x$ .  $\square$

**Definition 11.239.** A topological space  $\langle X, \mathcal{T} \rangle$  is second countable if it has a countable basis

**Theorem 11.240.** A second countable topological space is first countable

**Proof.** Let  $\langle X, \mathcal{T} \rangle$  be a second countable topological space and let  $\mathcal{B} \subseteq \mathcal{T}$  be the countable basis of  $\mathcal{T}$  then if  $x \in X$  and  $A$  a neighborhood of  $x$  then there exists a  $U \in \mathcal{T}$  such that  $x \in U \subseteq A$ . As  $\mathcal{B}$  is a basis there exists a  $B \in \mathcal{B}$  such that  $x \in B \subseteq U \subseteq A$ . So  $\mathcal{B}$  is a fundamentally system of neighborhoods of every element of  $x$ .  $\square$

## 11.9 Topological Vector Spaces

**Definition 11.241.** A topological vector space is a vector space  $\langle X, +, \cdot \rangle$  over  $\mathbb{K}$  together with a topology  $\mathcal{T}$  on  $X$  such that

1.  $+: X \times X \rightarrow X$  defined by  $(x, y) \mapsto +x + y$  is continuous
2.  $\cdot: F \times X \rightarrow X$  defined by  $(\alpha, x) \mapsto \alpha \cdot x$  is continuous

using the product topology of  $X \times X$  and  $\mathbb{K} \times X$  (using the topology of  $\langle \mathbb{K}, \|\cdot\| \rangle$  and of  $X$ ).

**Example 11.242.** Using 11.162 we have that a normed vector space is a topological vector space.

**Definition 11.243.** Let  $X, Y$  be topological vector spaces then a function  $L: X \rightarrow Y$  is a toplinear isomorphism iff

1.  $L$  is a linear isomorphism, meaning that
  - a.  $L$  is a bijection
  - b.  $L(x + y) = L(x) + L(y)$
  - c.  $L(\alpha \cdot x) = \alpha \cdot L(x)$
2.  $L$  and  $L^{-1}$  are continuous

**Theorem 11.244.** Let  $\langle X, +, \cdot \rangle$  be a topological vector space then the functions

1.  $+_x: X \rightarrow X$  defined by  $y \mapsto +_x(y) = x + y$
2.  $\cdot_\alpha: X \rightarrow X$  defined by  $y \mapsto \cdot_\alpha(y) = \alpha \cdot y$

are continuous.

**Proof.**

1. Let  $U$  be open and take  $y \in (+_x)^{-1}(U) = \{y \in X | x + y \in U\}$  then  $(x, y) \in +^{-1}(U)$  and by continuity of  $+$  (and the definition of the product topology) there exists open  $V_1, V_2$  with  $(x, y) \in V_1 \times V_2$  such that  $+ (V_1 \times V_2) \subseteq U \Rightarrow V_1 + V_2 \subseteq U$  and thus as  $x \in V_1$  we have  $+_x(V_2) = \{x\} + V_2 \subseteq V_1 + V_2 \subseteq U$  proving continuity of  $+_x$ .
2. Let  $U$  be open and take  $x \in \cdot_\alpha(U) = \{x \in X | \alpha \cdot x \in U\}$  then  $(\alpha, x) \in \cdot^{-1}(U)$  and by continuity of  $\cdot$  there exists open  $V_1, V_2$  with  $(\alpha, x) \in V_1 \times V_2$  such that  $\cdot (V_1 \times V_2) = V_1 \cdot V_2 \subseteq U$  and thus  $\cdot_\alpha(V_2) = \{\alpha\} \cdot V_2 \subseteq V_1 \cdot V_2 \subseteq U$  proving continuity.  $\square$

## 11.10 Compact Spaces

**Definition 11.245.** A topological space  $\langle X, \mathcal{T} \rangle$  is compact iff for every family of  $\{U_i\}_{i \in I}$  of open sets such that  $X = \bigcup_{i \in I} U_i$  [ $\{U_i\}_{i \in I}$  cover's  $X$ ] there exists a finite  $J \subseteq I$  such that  $X = \bigcup_{i \in J} U_i$

**Definition 11.246.** Let  $\langle X, \mathcal{T} \rangle$  be a topological space then a subset  $C \subseteq X$  is compact in  $X$  iff  $\langle C, \mathcal{T}_C \rangle$  is a compact topological space (where  $\mathcal{T}_C$  is the subspace topology)

**Theorem 11.247.** Let  $\langle X, \mathcal{T} \rangle$  be a topological space then a subset  $C \subseteq X$  is compact in  $X$  if and only if for every family  $\{U_i\}_{i \in I}$  of open sets such that  $C \subseteq \bigcup_{i \in I} U_i$  there exists a finite  $J \subseteq I$  such that  $C \subseteq \bigcup_{i \in J} U_i$

**Proof.**

1.  $(\Rightarrow)$  If  $\{U_i\}_{i \in I}$  is a collection of open sets such that  $C \subseteq \bigcup_{i \in I} U_i$  then  $C = \bigcup_{i \in I} (C \cap U_i)$  where  $\forall i \in I$  we have that  $U_i \cap C \in \mathcal{T}_C$  so that by the fact that  $\langle C, \mathcal{T}_C \rangle$  is compact there exists a finite  $J \subseteq I$  such that  $C = \bigcup_{i \in J} (U_i \cap C) \Rightarrow C \subseteq \bigcup_{i \in J} U_i$
2.  $(\Leftarrow)$  Assume that  $\{V_i\}_{i \in I}$  is a collection of open sets in  $\mathcal{T}_C$  such that  $C = \bigcup_{i \in I} V_i$  then by the definition of the subspace topology we have  $\forall i \in I$  there exists a  $U_i$  such that  $V_i = U_i \cap C$  so that  $C = \bigcup_{i \in I} (U_i \cap C) \subseteq \bigcup_{i \in I} U_i$  so by the hypothesis there exists a finite  $J \subseteq I$  such that  $C \subseteq \bigcup_{i \in J} U_i \Rightarrow C = \bigcup_{i \in J} (U_i \cap C) = \bigcup_{i \in J} V_i$  proving that  $\langle C, \mathcal{T}_C \rangle$  is a compact topological space.  $\square$

**Example 11.248.** If  $\langle X, \mathcal{T} \rangle$  is a topological space then  $\emptyset$  is a compact subset of  $X$

**Proof.** If  $\{U_i\}_{i \in I}$  is a non empty family of open sets covering  $\emptyset$  then we have for  $I$  either

$I = \emptyset$ . then  $J = \emptyset = I$  is finite so  $\{U_i\}_{i \in J}$  is a finite covering of  $\emptyset$

$I \neq \emptyset$ . then  $\exists i \in I$  if we take  $J = \{i\} \subseteq I$  then  $J$  is finite and  $\emptyset \subseteq E_i = \bigcup_{j \in \{i\}} U_i$   $\square$

**Theorem 11.249.** Let  $\langle X, \mathcal{T} \rangle$  be a topological space and  $\{C_i\}_{i \in \{1, \dots, n\}}$  be a finite family of compact subsets of  $X$  then  $\bigcup_{i \in \{1, \dots, n\}} C_i$  is compact.

**Proof.** Let  $\bigcup_{i \in I} U_i$  is a open cover of  $\bigcup_{i \in \{1, \dots, n\}} C_i$  then  $\forall i \in \{1, \dots, n\}$  we have  $C_i \subseteq \bigcup_{j \in I} U_j$   $C_i$  is compact  $\Rightarrow \exists$  finite  $I_i \subseteq I \vdash C_i \subseteq \bigcup_{j \in I_i} U_j$  so that  $\bigcup_{i \in \{1, \dots, n\}} C_i \subseteq \bigcup_{i \in \{1, \dots, n\}} (\bigcup_{j \in I_i} U_j)$ . If now  $x \in \bigcup_{i \in \{1, \dots, n\}} (\bigcup_{j \in I_i} U_j)$  there exists a  $i \in \{1, \dots, n\}$  such that  $x \in \bigcup_{j \in I_i} U_j \Rightarrow \exists i \in \{1, \dots, n\}$  such that  $\exists j \in I_i$  with  $x \in U_j \Rightarrow \exists j \in \bigcup_{i \in \{1, \dots, n\}} I_i \subseteq I \vdash x \in U_j$  proving that  $\bigcup_{i \in \{1, \dots, n\}} (\bigcup_{j \in I_i} U_j) \subseteq \bigcup_{j \in \bigcup_{i \in \{1, \dots, n\}} I_i} U_j$  and thus we have  $\bigcup_{i \in \{1, \dots, n\}} C_i \subseteq \bigcup_{j \in \bigcup_{i \in \{1, \dots, n\}} I_i} U_i$  where  $\bigcup_{i \in \{1, \dots, n\}} I_i$  is a finite subset of  $I$ .  $\square$

**Theorem 11.250.** Let  $\langle X, \mathcal{T} \rangle$  be a topological space and let  $A \subseteq X$  then if  $C \subseteq A$  is compact in  $\langle A, \mathcal{T}_A \rangle$  iff it is compact in  $\langle X, \mathcal{T} \rangle$

**Proof.**

$\Rightarrow$ . Let  $C$  be compact in  $\langle X, \mathcal{T} \rangle$ . Let  $\{U_i\}_{i \in I}$  be a set of open sets in  $\mathcal{T}$  with  $C \subseteq \bigcup_{i \in I} U_i \Rightarrow C = C \cap A \subseteq (\bigcup_{i \in I} U_i) \cap A = \bigcup_{i \in I} (U_i \cap A)$   $C$  is compact in  $\langle A, \mathcal{T}_A \rangle \Rightarrow \exists J \subseteq I$  with  $J$  finite and  $C \subseteq \bigcup_{i \in J} (U_i \cap A) \subseteq \bigcup_{i \in J} U_i$  so that  $C$  is compact in  $\langle X, \mathcal{T} \rangle$

$\Leftarrow$ . Let  $C$  be compact in  $\langle A, \mathcal{T}_A \rangle$  then if  $C \subseteq \bigcup_{i \in I} U_i$  (a open cover of  $A$  in  $\mathcal{T}$ ) then  $C = C \cap A \subseteq \bigcup_{i \in I} (U_i \cap A)$  (a open cover of  $A$  in  $\mathcal{T}_A$ ), so by compactness we have that there exists a finite set  $J \subseteq I$  such that  $C \subseteq \bigcup_{i \in J} (U_i \cap A) \subseteq \bigcup_{i \in J} U_i$  proving that  $C$  is compact in  $\mathcal{T}$ .  $\square$

**Theorem 11.251.** Let  $\langle X, d \rangle$  be a metric space and let  $C \subseteq X$  be a compact set then  $C$  is bounded (see 11.65)

**Proof.** Let  $C$  be a compact subset of  $X$  then  $\forall c \in C$  take  $B_d(c, 1) \ni c$  which is open. So we have  $C \subseteq \bigcup_{c \in C} B_d(c, 1)$  and thus by compactness there exists a finite set  $\{c_1, \dots, c_n\} \subseteq C$  such that  $C \subseteq \bigcup_{i \in \{1, \dots, n\}} B_d(c_i, 1)$ . Take then  $M = 2 + \max(\{d(c_i, c_j) | (i, j) \in \{1, \dots, n\} \times \{1, \dots, n\}\})$  (which exists because of 5.50 and 5.44). If now  $x, y \in C \Rightarrow \exists i, j \in \{1, \dots, n\}$  such that  $x \in B_d(c_i, 1)$  and  $y \in B_d(c_j, 1)$  and thus  $d(x, y) \leq d(x, c_i) + d(c_i, y) \leq d(x, c_i) + d(c_i, c_j) + d(c_j, y) < 1 + d(c_i, c_j) + 1 = 2 + d(c_i, c_j) \leq M$  proving that  $C$  is bounded.  $\square$

**Corollary 11.252.** Let  $\langle X, \|\cdot\| \rangle$  be a normed space and let  $C \subseteq X$  be a compact set then  $\exists M \in \mathbb{R}$  such that  $\forall x \in C$  we have that  $\|x\| \leq M$

**Proof.** Using the previous theorem [see 11.251] we have that  $C$  is bounded hence using 11.77 there exists a  $M \in \mathbb{R}$  such that  $\forall x \in C$  we have  $\|x\| \leq M$ .  $\square$

**Theorem 11.253.** Let  $\langle X, \mathcal{T}_X \rangle$ ,  $\langle Y, \mathcal{T}_Y \rangle$  be topological spaces and  $f: X \rightarrow Y$  a continuous map then if  $C \subseteq X$  is compact then  $f(C)$  is compact.

**Proof.** Let  $\{V_i\}_{i \in I}$  be a family of open sets in  $Y$  so that  $f(C) \subseteq \bigcup_{i \in I} V_i$  then  $\forall x \in C$  we have  $f(x) \in f(C)$  and thus  $\exists i \in I$  such that  $f(x) \in V_i$  so that by continuity of  $f$  there exists a  $U_i$  open in  $X$  with  $x \in U_i$  and  $f(x) \in f(U_i) \subseteq V_i$  so that  $C \subseteq \bigcup_{i \in I} U_i$ . By compactness of  $C$  there exist a finite  $J \subseteq I$  so that  $C \subseteq \bigcup_{i \in J} U_i \stackrel{2.58}{\Rightarrow} f(C) \subseteq f(\bigcup_{i \in J} U_i) \stackrel{2.62}{=} \bigcup_{i \in J} f(U_i) \subseteq \bigcup_{i \in J} V_i$  proving that  $f(C)$  is compact.  $\square$

**Theorem 11.254.** Let  $\langle X, \mathcal{T} \rangle$  be a Hausdorff topological space then every compact subspace is closed

**Proof.** Let  $C$  be a compact subset of  $X$  then if  $y \in X \setminus C$  we have that  $\forall x \in C$  there exists open sets  $U_x, V_x$  such that  $x \in U_x$ ,  $y \in V_x$  and  $U_x \cap V_x \neq \emptyset$ . So  $C \subseteq \bigcup_{x \in X} U_x$  and by compactness there exists a finite set  $F \subseteq X$  such that  $C \subseteq \bigcup_{x \in F} U_x$ . Now  $y \in \bigcap_{z \in F} V_z$  which is open (because  $F$  is finite) and  $(\bigcap_{z \in F} V_z) \cap C \subseteq (\bigcap_{z \in F} V_z) \cap (\bigcup_{x \in F} U_x) = \bigcup_{x \in F} ((\bigcap_{z \in F} V_z) \cap U_x) \subseteq \bigcup_{x \in F} (V_x \cap U_x) = \bigcup_{x \in F} \emptyset = \emptyset$  so that  $y \in \bigcap_{z \in F} V_z \subseteq X \setminus C$  proving that  $X \setminus C$  is open and thus that  $C$  is closed.  $\square$

**Theorem 11.255.** Let  $\langle X, \mathcal{T} \rangle$  be a topological space then every closed subset of a compact subset of  $X$  is compact.

**Proof.** Let  $C \subseteq X$  be compact and let  $A \subseteq C$  be a closed set and let  $\{U_i\}_{i \in I}$  be a family of open sets covering  $A$  (so  $A \subseteq \bigcup_{i \in I} U_i$ ) then as  $C = A \cup (C \setminus A) \subseteq A \cup (X \setminus A)$  we have that if  $\mathcal{W} = \{U_i \mid i \in I\} \cup \{X \setminus A\}$  then  $\{\mathcal{W}\}_{W \in \mathcal{W}}$  covers  $C$ . By compactness we have that a finite subset  $\mathcal{V}$  of  $\mathcal{W}$  covers  $C$  and thus  $A$  so  $A \subseteq \bigcup_{U \in \mathcal{V}} U$ . If now  $A \not\subseteq \bigcup_{U \in \mathcal{V} \setminus \{X \setminus A\}} U$  then there exists a  $a \in A$  with  $a \in \bigcup_{U \in \mathcal{V} \setminus \{X \setminus A\}} U$  which as  $A \subseteq \bigcup_{U \in \mathcal{V}} U$  means that  $a \in X \setminus A$  contradicting that  $a \in A$  so we must have that  $A \subseteq \bigcup_{U \in \mathcal{V} \setminus \{X \setminus A\}} U$ . If we define now  $\{\mathcal{I}_U\}_{U \in \mathcal{V} \setminus \{X \setminus A\}}$  by  $\mathcal{I}_U = \{i \in I \mid U_i = U\}$  then as  $\mathcal{V} \setminus \{X \setminus A\} \subseteq \mathcal{W} \setminus \{X \setminus A\} = \{U_i \mid i \in I\}$  we have that  $\mathcal{I}_U \neq \emptyset$ . Using the axiom of choice (see 2.210) we have then that there exist a function  $c: \mathcal{V} \setminus \{X \setminus A\} \rightarrow \bigcup_{U \in \mathcal{V} \setminus \{X \setminus A\}} \mathcal{I}_U \subseteq I$  with  $c(U) \in \mathcal{I}_U$ . So we have then the surjection  $c: \mathcal{V} \setminus \{X \setminus A\} \rightarrow c(\mathcal{V} \setminus \{X \setminus A\}) \subseteq I$  which means that because  $\mathcal{V} \setminus \{X \setminus A\}$  is finite that  $c(\mathcal{V} \setminus \{X \setminus A\})$  is finite. Now if  $a \in A$  then as  $A \subseteq \bigcup_{U \in \mathcal{V} \setminus \{X \setminus A\}} U$  there exist a  $U \in \mathcal{V} \setminus \{X \setminus A\}$  such that  $a \in U \stackrel{c(U) \in \mathcal{I}_U}{=} U_{c(U)} \subseteq \bigcup_{i \in c(\mathcal{V} \setminus \{X \setminus A\})} U_i$  proving that  $A \subseteq \bigcup_{i \in c(\mathcal{V} \setminus \{X \setminus A\})} U_i$  and thus that  $A$  is compact.  $\square$

**Lemma 11.256.** Given  $\langle \mathbb{R}, \|\cdot\| \rangle$  the normed space with the canonical norm  $\|\cdot\|$ ,  $S$  a non empty closed set such that  $\sup(S)$  [or  $\inf(S)$ ] exists then  $\sup(S) \in S$  [ $\inf(S) \in S$ ]

**Proof.** First if  $\sup(S) \in U \in \mathcal{T}_{\|\cdot\|}$  then there exists a  $\varepsilon > 0$  such that  $\sup(S) \in B_{\|\cdot\|}(\sup(S), \varepsilon) \subseteq U$ . As  $\sup(S) - \varepsilon < \sup(S)$  there exists using the definition of the supremum a  $y \in S$  such that  $\sup(S) - \varepsilon < y \leq \sup(S) \Rightarrow y \in B_{\|\cdot\|}(\sup(S), \varepsilon) \subseteq U$  proving that  $U \cap S = \emptyset$  hence by 11.21 and the closure of  $S$  we have  $\sup(S) \in S$ .

Second if  $\inf(S) \in U \in \mathcal{T}_{\|\cdot\|}$  then there exists a  $\varepsilon > 0$  such that  $\inf(S) \in B_{\|\cdot\|}(\inf(S), \varepsilon) \subseteq U$ . As  $\inf(S) - \varepsilon < \inf(S)$  there exists using the definition of the infimum a  $y \in S$  such that  $\inf(S) \leq y < \inf(S) + \varepsilon \Rightarrow y \in B_{\|\cdot\|}(\inf(S), \varepsilon) \subseteq U$  proving that  $U \cap S = \emptyset$  hence by 11.21 and the closure of  $S$  we have  $\inf(S) \in S$ .  $\square$

**Theorem 11.257. (Extreme value Theorem)** Let  $\langle X, \mathcal{T} \rangle$  be a topological space,  $C$  a non empty compact subset and  $f: X \rightarrow \mathbb{R}$  a continuous function then there exists a  $m, M \in X$  such that  $\forall x \in X$  we have  $f(m) \leq f(x) \leq f(M)$

**Proof.** As  $X$  is compact we have by 11.253 that  $f(C)$  is a compact set in  $\mathbb{R}$ , so by 11.251  $f(C)$  is bounded. Hence there exists a  $M$  such that  $\forall x, y \in f(C)$  we have  $|f(x) - f(y)| \leq M$ . As  $C \neq \emptyset$  there exists a  $x_0 \in C$  so that  $\forall x \in C$  we have  $|f(x) - f(x_0)| \leq M \Rightarrow f(x_0) - M \leq f(x) \leq f(x_0) + M$ . Using the conditional completeness of the reals we have that  $\inf(f(C))$  and  $\sup(f(C))$  exists. Using the previous theorem we have then that  $\inf(f(C)), \sup(f(C)) \in f(C)$  so that  $\exists m, M \in C$  such that  $f(m) = \inf(f(C)) \wedge f(M) = \sup(f(C))$ . Hence if  $x \in C$  then  $f(x) < f(C)$  and thus  $f(m) = \inf(f(C)) \leq f(x) \leq \sup(f(C)) = f(M)$ .  $\square$

**Theorem 11.258.** Let  $\langle X, \mathcal{T} \rangle$  be a Hausdorff compact topological space then it is regular and normal.

**Proof.**

1. **(regular)** Let  $A$  be a closed set and  $x \in X$  with  $x \notin A$  then because of Hausdorff we have that  $\forall a \in A$  we have the existence of disjoint open sets  $V_a, U_a$  with  $x \in V_a, a \in U_a$  and  $V_a \cap U_a = \emptyset$  then  $A \subseteq \bigcup_{a \in A} U_a$  and because  $A$  is compact (closed sets in a compact space are compact) there exist a finite  $B \subseteq A$  such that  $A \subseteq \bigcup_{a \in B} U_a$ . We have then that  $V = \bigcap_{a \in B} V_a$  is a open set (because  $B$  is finite),  $a \in V$  and  $V \cap A = (\bigcap_{b \in B} V_b) \cap A \subseteq (\bigcap_{b \in B} V_b) \cap (\bigcup_{a \in B} U_a) = \bigcup_{a \in B} ((\bigcap_{b \in B} V_b) \cap U_a) \subseteq \bigcup_{a \in B} (V_a \cap U_a) = \emptyset \Rightarrow V \cap A = \emptyset$  proving regularity.
2. **(normal)** Let  $A, B$  be two disjoint closed sets then  $\forall a \in A$  we have using (1) the existence of disjoint open sets  $V_a, U_a$  so that  $B \subseteq V_a$  and  $a \in U_a$  so that  $A \subseteq \bigcup_{a \in A} U_a$ . From the compactness of  $A$  (because closed sets in a compact space are compact) there exist a finite set  $F \subseteq A$  so that  $A \subseteq \bigcup_{a \in F} U_a$  then  $B \subseteq \bigcap_{a \in F} V_a$  and if we take  $U = \bigcup_{a \in F} U_a$  (which is open) and  $V = \bigcap_{a \in F} V_a$  (which is open as  $F$  is finite) then  $A \subseteq U, B \subseteq V$  and  $V \cap U = (\bigcap_{b \in B} V_b) \cap (\bigcup_{a \in F} U_a) = \bigcup_{a \in B} ((\bigcap_{b \in B} V_b) \cap U_a) \subseteq \bigcup_{a \in B} (V_a \cap U_a) = \emptyset \Rightarrow U \cap V = \emptyset$  proving normality.  $\square$

**Theorem 11.259.** Let  $\langle X, \mathcal{T} \rangle$  be a topological space with a basis  $\mathcal{B}$  then a subset  $C \subseteq X$  is compact iff every cover of  $C$  by sets in the base contains a finite cover.

**Proof.**

1.  $(\Rightarrow)$  This is trivial from  $\mathcal{B} \subseteq \mathcal{T}$
2.  $(\Leftarrow)$  Let  $\{U_i\}_{i \in I}$  be a family of open sets such that  $C \subseteq \bigcup_{i \in I} U_i$  then  $\forall x \in C$  there exists a  $i_x \in I$  such that  $x \in U_{i_x}$  and thus there exist a  $B_x \in \mathcal{B}$  such that  $x \in B_x \subseteq U_x$ . So we have  $C \subseteq \bigcup_{x \in C} B_x$  and by the hypothesis there exists a finite  $\{x_1, \dots, x_n\} \subseteq C$  such that  $C \subseteq \bigcup_{x \in \{x_1, \dots, x_n\}} B_x$   $\overset{B_x \subseteq U_{i_x}}{\Rightarrow} \bigcup_{x \in \{x_1, \dots, x_n\}} U_{i_x} = \bigcup_{i \in \{i_{x_1}, \dots, i_{x_n}\}} U_{i_x}$  proving as  $\{i_{x_1}, \dots, i_{x_n}\}$  is finite that  $C$  is compact.  $\square$

**Definition 11.260.** Let  $\langle X, \mathcal{T} \rangle$  be a topological space then a subset  $C$  is called sequential compact if every infinite  $B \subseteq C$  has a limit point (see 11.18) that is in  $C$ .

**Theorem 11.261.** A compact subspace of a topological space  $\langle X, \mathcal{T} \rangle$  is sequential compact

**Proof.** We proceed by contradiction. Let  $C$  be a compact subset and let  $B \subseteq C$  be a infinite subset that does not have a limit point in  $C$ . Then  $\forall x \in C$  there exists a open  $U_x$  with  $x \in U_x$  and  $U_x \cap (B \setminus \{x\}) = \emptyset$ . If  $y \in U_x \cap B$  then if  $y \neq x \Rightarrow y \notin \{x\} \Rightarrow y \in (U_x \cap B) \setminus \{x\} = U_x \cap (B \setminus \{x\})$  contradicting  $U_x \cap (B \setminus \{x\}) = \emptyset$  so we must have  $y = x$  or  $U_x \cap B \subseteq \{x\}$ . Now  $C \subseteq \bigcup_{x \in C} U_x$  and as  $C$  is compact there exists a  $F \subseteq C$  such that  $C \subseteq \bigcup_{x \in F} U_x$  and thus  $B \overset{B \subseteq C}{\subseteq} B \cap C = B \cap (\bigcup_{x \in F} U_x) = \bigcup_{x \in F} (B \cap U_x) \subseteq \bigcup_{x \in F} \{x\} = F$  which is finite contradicting that  $B$  is infinite.  $\square$

**Theorem 11.262.** Let  $\langle X, d \rangle$  be a metric space equipped with the metric topology. Let  $K \subseteq X$  be a sequential compact subset of  $X$  and  $\{U_i\}_{i \in I}$  be a open cover of  $K$  then  $\exists \delta \in \mathbb{R}_+$  such that  $\forall x \in K$  there  $\exists i_x \in I$  such that  $x \in B_d(x, \delta) \subseteq U_{i_x}$

**Proof.** We prove this by contradiction. So assume that  $\forall \delta \in \mathbb{R}_+$  there  $\exists x \in K$  such that  $\forall i \in I$  we have  $B_d(x, \delta) \not\subseteq U_i$ . In particular this means that (using the isomorphism  $i_{\mathbb{Q}_N}: \mathbb{N} \rightarrow \mathbb{N}_{\mathbb{R} \subseteq \mathbb{R}}$  (see 8.56) and noting  $i_{\mathbb{Q}_N}(n)$  by  $n$ ) we have  $\forall n \in \mathbb{N}_0$  there exists a  $x_n \in K$  such that  $\forall i \in I$  we have  $B_d(x_n, \frac{1}{n}) \not\subseteq U_i$ . This defines a sequence  $\{x_i\}_{i \in \mathbb{N}_0}$  of elements in  $K$  such that  $B_d(x_n, \frac{1}{n}) \not\subseteq U_i \forall i \in I, \forall n \in \mathbb{N}_0$ . Define now the set  $O = \{x_n | n \in \mathbb{N}_0\} \subseteq K$ . If this set was finite then using 5.52 there exists a  $x_N \in O$  such that  $\forall n \in \mathbb{N}_0$  there exists a  $m \in \mathbb{N}_0$  with  $m \geq n$  and  $x_m = x_N$ . Now as  $x_N \in O \subseteq K$   $\{U_i\}_{i \in I}$  is a open cover of  $K \Rightarrow \exists i \in I$  with  $x_N \in U_i \Rightarrow \exists \varepsilon \in \mathbb{R}_+$  such that  $x_N \subseteq B_d(x_N, \varepsilon) \subseteq U_i$ .

Take then  $M \in \mathbb{N}_{\mathbb{R}}$  such that  $0 < \frac{1}{M} < \varepsilon$  (see 9.56) then by the above there exists a  $m \in \mathbb{N}_0$  with  $m \geq M$  (and thus  $\frac{1}{M} \geq \frac{1}{m}$ ) so that  $x_N = x_m$ , from this it follows that  $x_m \in B_d(x_m, \frac{1}{m}) = B_d(x_N, \frac{1}{m}) \subseteq B_d(x_N, \frac{1}{M}) \subseteq B_d(x_N, \varepsilon) \subseteq U_i$  contradicting the fact that  $B_d(x_m, \frac{1}{m}) \not\subseteq U_i \forall i \in I$ . We must thus conclude that  $O$  is infinite. As  $K$  is sequential compact we have thus a limit point  $k \in K$  for  $O$ . Given that  $K \subseteq \bigcup_{i \in I} U_i$  there exists a  $i_0 \in I$  such that  $k \in U_{i_0} \Rightarrow \exists \varepsilon_0 \in \mathbb{R}_+$  such that  $k \in B_d(k, \varepsilon_0) \subseteq U_{i_0}$ . The set  $P = \{n \in \mathbb{N}_0 | x_n \in B_d(k, \frac{\varepsilon_0}{2})\}$  is a infinite set of natural numbers [if it was finite we could take  $\varepsilon = \min(\frac{\varepsilon_0}{2}, \min(\{d(k, x_i) | i \in P\}))$   $\min(\{d(k, x_i) | i \in P\})$  exists if  $P$  is finite] then  $k \in B_d(k, \varepsilon) \subseteq B_d(k, \frac{\varepsilon_0}{2})$  and  $\forall i \in \mathbb{N}_0$  we have either  $i \notin P$  so that then  $x_i \notin B_d(k, \frac{\varepsilon_0}{2}) \supseteq B_d(k, \varepsilon) \Rightarrow x_i \notin B_d(k, \varepsilon)$  or  $i \in P$  and then as  $d(k, x_i) \geq \varepsilon$  we have  $x_i \notin B_d(k, \varepsilon)$ . So  $\forall i \in \mathbb{N}_0$  we have  $x_i \notin B_d(k, \varepsilon) \Rightarrow O \cap B_d(k, \varepsilon) = \emptyset$  contradicting the fact that  $k$  is the limit point of  $O$ . So we must have that  $P$  is infinite]. As  $P \subseteq \mathbb{N}_0$  is infinite there exists a  $m \in P$  such that  $\frac{1}{m} < \frac{\varepsilon_0}{2}$  [using 9.56 there exist a  $N \in \mathbb{N}_0$  such that  $\frac{1}{N} < \frac{\varepsilon_0}{2}$  and as  $P$  is infinite we have that  $P \not\subseteq S_N = \{1, \dots, N-1\}$  so that  $\exists m \in P$  with  $m \in \{1, \dots, N-1\} \Rightarrow m > N-1 \Rightarrow m \geq N \Rightarrow \frac{1}{m} \leq \frac{1}{N} < \frac{\varepsilon_0}{2}$ , as  $m \in P$  we have that  $x_m \in B_d(k, \frac{\varepsilon_0}{2}) \Rightarrow d(k, x_m) < \frac{\varepsilon_0}{2}$ . Now if  $z \in B_d(x_m, \frac{1}{m}) \Rightarrow d(x_m, z) < \frac{1}{m} \Rightarrow d(k, z) \leq d(k, x_m) + d(x_m, z) < \frac{\varepsilon_0}{2} + \frac{1}{m} < \frac{\varepsilon_0}{2} + \frac{\varepsilon_0}{2} = \varepsilon_0 \Rightarrow z \in B_d(k, \varepsilon_0) \Rightarrow B_d(x_m, \frac{1}{m}) \subseteq B_d(k, \varepsilon_0) \subseteq U_{i_0}$  contradicting the fact that  $\forall i \in I$  we have  $\forall n \in \mathbb{N}_0$  that  $B_d(x_n, \frac{1}{n}) \not\subseteq U_i$ .  $\square$ ]

**Theorem 11.263.** Let  $\langle X, d \rangle$  be a metric space with the metric topology and  $K \subseteq X$  then the following are equivalent

1.  $K$  is compact
2.  $K$  is sequential compact

**Proof.**

1.  $(\Rightarrow)$  This is already proved in 11.261
2.  $(\Leftarrow)$  Let  $K$  be sequential compact and  $\{U_i\}_{i \in I}$  a cover of  $K$ . If  $K = \emptyset$  then it is trivially compact so lets assume that  $K \neq \emptyset$ . Using 11.262 there exists a  $\delta \in \mathbb{R}_+$  such that  $\forall x \in K$  there exists a  $i \in I$  such that  $x \in B_d(x, \delta) \subseteq U_i$ . If we prove now that

$$\exists F \text{ (finite)} \subseteq K \text{ such that } K \subseteq \bigcup_{x \in F} B_d(x, \delta) \quad (11.31)$$

then  $\forall f \in F$  there exists a  $j_f \in I$  such that  $x_{j_f} \in B_d(x_{j_f}, \delta) \subseteq U_{j_f}$  defining a function  $\sigma: F \rightarrow I$  given by  $f \rightarrow j_f$  (by the Axiom of Choice 2.211) so that  $K \subseteq \bigcup_{f \in F} B_d(x_{\sigma(f)}, \delta) \subseteq \bigcup_{f \in F} U_{\sigma(f)} = \bigcup_{j \in P} U_j$  where  $P = \sigma(F) \subseteq I$  is finite, proving compactness. We prove 11.31 by contradiction. So assume that

$$\forall F \text{ (finite)} \subseteq K \text{ we have that } K \not\subseteq \bigcup_{x \in F} B_d(x, \delta) \quad (11.32)$$

Given a family  $\{x_i\}_{i \in \{0, \dots, n\}}$  we have that the set  $\{x_i | i \in \{0, \dots, n\}\}$  is finite so that there exists a  $\rho(\{x_i\}_{i \in \{0, \dots, n\}}) \in K \setminus (\bigcup_{i \in \{0, \dots, n\}} B_d(x_i, \delta))$ , using the Axiom of Choice (see 2.211) this defines a function  $\rho: \mathcal{M} = \{\{x_i\}_{i \in \{0, \dots, n\}} | n \in \mathbb{N} \wedge \forall i \in \{0, \dots, n\} \text{ we have } x_i \in K\} \rightarrow K$ . As  $K \neq \emptyset$  there exists a  $x_0 \in K$  so that using recursion (see 5.21) there exists a sequence  $f: \mathbb{N} \rightarrow K$  such that

- a.  $f(0) = x_0$
- b.  $f(n+1) = \rho(\{f(i)\}_{i \in \{0, \dots, n\}}) \in K \setminus (\bigcup_{i \in \{0, \dots, n\}} B_d(f(i), \delta))$

Or if  $n \in \mathbb{N}_0$  then  $x_n \notin K \setminus (\bigcup_{i \in \{0, \dots, n-1\}} B_d(f(i), \delta))$ . If now  $i \neq j$  then we have  $i < j$  (for the other case just switch  $i$  and  $j$ ) then  $f(i) \notin B_d(f(j), \delta) \Rightarrow d(f(i), f(j)) \geq \delta > 0 \Rightarrow f(i) \neq f(j)$  proving that  $f$  is injective. Using 5.49 we have then that  $O = f(\mathbb{N}) \subseteq K$  is finite so it has a accumulation point  $x \in K$  (because of sequential compactness). Take then  $B_d(x, \frac{\delta}{2})$  then

there exists a  $i \in \mathbb{N}$  such that  $f(i) \neq x \Rightarrow d(x, f(i)) > 0$  and  $f(i) \in B_d(x, \frac{\delta}{2})$ , if we take then  $\varepsilon = \min\left(\frac{\delta}{4}, \frac{d(x, f(i))}{2}\right)$  then  $d(x, f(i)) > \varepsilon \Rightarrow f(i) \notin B_d(x, \varepsilon)$ . If  $m \neq i$  and  $f(m) \in B_d(x, \varepsilon)$   $\delta \leq d(f(m), f(i)) \leq d(f(m), x) + d(x, f(i)) < \varepsilon + \frac{\delta}{2} < \frac{\delta}{2} + \frac{\delta}{2} = \delta$  giving the contradiction  $\delta < \delta$  so we must have that  $f(m) \in B_d(x, \varepsilon)$  this together with  $f(i) \in B_d(x, \varepsilon)$  proves that  $O = f(\mathbb{N}) \cap B_d(x, \varepsilon) = \emptyset$  contradicting the fact that  $x$  is a accumulation point of  $O$ . So 11.32 is false proving 11.31 and thus the theorem.  $\square$

**Definition 11.264.** Let  $X$  be a set then  $K \subseteq \mathcal{P}(X)$ , a set of subsets of  $X$ , is a **compact class** if for any sequence  $\{K_i\}_{i \in \mathbb{N}_0}$  of elements in  $K$  with  $\bigcap_{i \in \mathbb{N}_0} K_i = \emptyset$  we have that  $\exists N \in \mathbb{N}_0$  such that  $\bigcap_{i \in \{1, \dots, N\}} K_i = \emptyset$

**Example 11.265.** Let  $\langle X, \mathcal{T} \rangle$  be a Hausdorff topological space and  $K$  a set of compact sets in  $X$  then  $K$  is a countable compact class.

**Proof.** We prove this by contradiction so let  $\{K_i\}_{i \in \mathbb{N}_0}$  a countable family of elements in  $K$  with  $\bigcap_{i \in \mathbb{N}_0} K_i = \emptyset$  such that  $\forall n \in \mathbb{N}_0$  we have  $\bigcap_{i \in \{1, \dots, n\}} K_i \neq \emptyset$ . Define then  $\{E_i\}_{i \in \mathbb{N}_0}$  by  $E_i = \bigcap_{j=1}^i K_j$  then as  $K_j$  is closed (see 11.254) we have that  $\{X \setminus E_i\}_{i \in \mathbb{N}_0}$  is a family of open sets. If now  $x \in K_1$  then as  $\bigcap_{i \in \mathbb{N}_0} K_i = \emptyset$  there exists a  $n \in \mathbb{N}_0$  such that  $x \notin K_n \Rightarrow x \in X \setminus K_n$  which proves that  $K_1 \subseteq \bigcup_{i \in \mathbb{N}_0} (K \setminus E_i)$ , using compactness of  $K_1$  there exists then a finite  $\{i_1, \dots, i_n\} \subseteq \mathbb{N}_0$  such that  $K_1 \subseteq \bigcup_{i \in \{i_1, \dots, i_n\}} (K \setminus E_i)$ . Take now  $m = \max(i_1, \dots, i_n)$  then as  $\forall j \in \{i_1, \dots, i_n\}$  we have  $E_m \subseteq E_j \Rightarrow K \setminus E_j \subseteq K \setminus E_m$  so that  $K_1 \subseteq K \setminus E_m$ . As we assume that  $E_m = \bigcap_{j=1}^m K_j \neq \emptyset$  there exists a  $y \in E_m \subseteq K_1$  so that  $y \in K_1 \subseteq K \setminus E_m \Rightarrow y \in E_m \wedge y \notin E_m$  a contradiction. So we must have that there exists a  $N \in \mathbb{N}_0$  with  $\bigcap_{i \in \{1, \dots, N\}} K_i = \emptyset$ .  $\square$

**Theorem 11.266.** Let  $[a_i, b_i]$ ,  $a_i, b_i \in \mathbb{R}, i \in \mathbb{N}$   $a_i \leq b_i$  be a decreasing sequence (so  $[a_{i+1}, b_{i+1}] \subseteq [a_i, b_i]$ ) of closed bounded intervals in  $\mathbb{R}$  then

1.  $\exists a, b \in \mathbb{R}, a \leq b$  such that  $\bigcap_{i \in \mathbb{N}} [a_i, b_i] = [a, b]$
2. If the sequence  $\{b_i - a_i\}_{i \in \mathbb{N}}$  has the limit 0 (see 11.322 for a definition of a limit) then there exists a number  $a$  such that  $\bigcap_{i \in \mathbb{N}} [a_i, b_i] = [a]$

**Proof.** First we prove by induction that  $\forall i, j \in \mathbb{N}$  with  $i \leq j$  we have  $a_i \leq a_j \leq b_j \leq b_i$  so let  $S_i = \{n \in \mathbb{N} \mid a_i \leq a_{i+n} \leq b_{i+n} \leq b_i\}$  then we have

1. if  $n = 0$  then  $a_i \leq a_{i+0} \leq b_{i+0} \leq b_i$  proving that  $0 \in S$
2. If  $n \in \mathbb{N}$  then  $[a_{i+(n+1)}, b_{i+(n+1)}] \subseteq [a_{i+n}, b_{i+n}] \underset{n \in S}{\Rightarrow} a_i \leq a_{i+n} \leq a_{i+(n+1)} \leq b_{i+(n+1)} \leq b_{i+n} \leq b_i$  proving that  $n+1 \in S$

Using induction we have then that  $S = \mathbb{N}$  so if  $i \leq j$  then  $j - i \in \mathbb{N} = S \Rightarrow a_i \leq a_{i+(j-i)} = a_j \leq b_j = b_{i+(j-i)} \leq b_i$ . Specially we have that  $\forall i \in \mathbb{N}$  that  $a_0 \leq a_i \leq b_i \leq b_0$ .

1.  $A = \{a_i \mid i \in \mathbb{N}\}$  is then bounded above by  $b_0$  and  $B = \{b_i \mid i \in \mathbb{N}_0\}$  is bounded below by  $a_0$  so using 9.43 and 2.180 there exists a  $a = \sup(A)$  and  $b = \inf(B)$ . As  $\forall i, j \in \mathbb{N}$  we have either  $i \leq j$  so that  $a_i \leq a_j \leq b_j$  or  $j \leq i$  then  $a_i \leq b_i \leq b_j$  so in all case we have  $a_i \leq b_j$  and thus  $a_i \leq \inf(B)$  and  $a = \sup(A) \leq \inf(B) = b \Rightarrow a \leq b$ . And  $\forall i \in \mathbb{N}$  we have by the definition of inf and sup that  $a_i \leq a \leq b \leq b_i$  so that  $[a, b] \subseteq \bigcap_{i \in \mathbb{N}} [a_i, b_i]$ . Now assume that  $x \in \bigcap_{i \in \mathbb{N}} [a_i, b_i]$  then  $\forall i \in \mathbb{N}$  we have  $a_i \leq x \leq b_i$  so  $x$  is a upper bound for  $A$  and a lower bound for  $B$  so that  $a \leq x \leq b \Rightarrow x \in [a, b]$  proving that  $\bigcap_{i \in \mathbb{N}} [a_i, b_i] = [a, b]$
2. By (1) we have that there exists a  $a, b \in \mathbb{R}$  with  $\bigcap_{i \in \mathbb{N}} [a_i, b_i] = [a, b]$ . Suppose now that  $a < b$  then  $\forall i \in \mathbb{N}$  we have  $a_i \leq a < b \leq b_i \Rightarrow 0 < b - a \leq b_i - a \leq b_i - a_i$ . Now as by hypothesis we have that  $\{b_i - a_i\}_{i \in \mathbb{N}}$  has limit 0 then given  $\varepsilon = b - a > 0$  choose  $N \in \mathbb{N}$  such that  $\forall i \geq N$  we have  $b_i - a_i = |b_i - a_i| < \varepsilon \Rightarrow \varepsilon = b - a \leq b_N - a_N < \varepsilon$  giving the contradiction  $\varepsilon < \varepsilon$  so we must have that  $a = b$  but then  $\bigcap_{i \in \mathbb{N}} [a_i, b_i] = [a, a] = \{a\}$   $\square$

**Theorem 11.267.** Let  $[a, b], a \leq b, a, b \in \mathbb{R}$  be a closed and bounded interval in  $\mathbb{R}$  then  $[a, b]$  is a compact subset in  $\mathbb{R}$  (using the norm topology generated by  $\|\cdot\|$ )

**Proof.** The proof is divided in two cases as  $a \leq b$

1. **( $a = b$ )** then  $[a, b] = \{a\}$  and if  $\bigcup_{i \in I} U_i$  is a covering of  $\{a\}$  then there exists a  $i \in I$  such that  $a \in U_i \Rightarrow [a, b] = \{a\} \subseteq \bigcup_{j \in \{i\}} U_j$  proving as  $\{a\}$  is finite compactness
2. **( $a < b$ )** Here we proceed by contradiction and assume that  $[a, b]$  is not compact. Then there exists by 11.72 and 11.259 a family  $\{(a_i, b_i)\}_{i \in I}$  of open intervals such that  $[a, b] \subseteq \bigcup_{i \in I} (a_i, b_i)$  that does not contain a finite cover. If we define now  $\mathcal{A} = \{[x, y] | x, y \in \mathbb{R} \wedge a \leq x < y \leq b\}$  and  $\mathcal{M} = \{[x_i, y_i] | i \in \{0, \dots, n\} \wedge [x_i, y_i] \in \mathcal{A}\}$ . Then if  $\{[x_i, y_i] | i \in \{0, \dots, n\}\} \in \mathcal{M}$  we have that  $[x_n, y_n]$  is not covered by a finite subset of  $\{(a_i, b_i)\}_{i \in I}$  and as  $[x_n, y_n] \in \mathcal{A}$  we have that  $a \leq x_n < y_n \leq b$ , take now  $c = x_n + \frac{y_n - x_n}{2}$ . Then if  $[x_i, c], [c, y_i]$  are both covered by finite sub-cover of  $\{(a_i, b_i)\}_{i \in I}$  there exists finite  $A, B \subseteq I$  such that  $[x_n, c] \subseteq \bigcup_{i \in A} (a_i, b_i) \wedge [c, y_n] \subseteq \bigcup_{i \in B} (a_i, b_i) \Rightarrow [x_n, y_n] \subseteq \bigcup_{i \in A \cup B} (a_i, b_i)$  meaning that  $[x_i, y_i]$  would be covered by a finite sub-cover of  $\{(a_i, b_i)\}_{i \in I}$  a contradiction. So either  $[x_n, c]$  or  $[c, y_n]$  is not covered by a finite sub-cover of  $\{(a_i, b_i)\}_{i \in I}$ . Consider then the two cases to construct a  $[x, y]$ 
  - a. **( $[x_n, c]$  is not covered by a finite sub-cover of  $\{(a_i, b_i)\}_{i \in I}$ )** take then  $[x, y] = [x_n, c] \subseteq [x_n, y_n]$ ,  $[x, y]$  is not covered by a finite sub-cover of  $\{(a_i, b_i)\}_{i \in I}$  and  $y - x = x_n + \frac{y_n - x_n}{2} - x_n = \frac{y_n - x_n}{2}$  and  $a \leq x_n \leq x < y \leq y_n \leq b$
  - b. **( $[x_n, c]$  is covered by a finite sub-cover of  $\{(a_i, b_i)\}_{i \in I}$ )** then  $[c, y_n]$  is not covered by a finite sub-cover of  $\{(a_i, b_i)\}_{i \in I}$ . Take then  $[x, y] = [c, y_n] \subseteq [x_n, y_n]$  and  $y - x = y_n - (x_n + \frac{y_n - x_n}{2}) = (y_n - x_n) - \frac{y_n - x_n}{2} = \frac{y_n - x_n}{2}$  and  $a \leq x_n \leq x < y < y_n \leq b$

this defines a function  $\delta: \mathcal{M} \rightarrow \mathcal{A}$  mapping  $\{[x_i, y_i] | i \in \{0, \dots, n\}\} \rightarrow \delta(\{[x_i, y_i] | i \in \{0, \dots, n\}\}) = [x, y]$  such that  $[x - y] \in \mathcal{A}$  and  $[x, y] \subseteq [x_n, y_n]$ . Using recursion and the fact that  $[a, b] \in \mathcal{A}$  (see 5.24) we have found a sequence  $\{[x_i, y_i]\}_{i \in \mathbb{N}}$  such that  $\forall i \in \mathbb{N}$  we have that  $[x_i, y_i]$  is not covered by a finite sub-cover of  $\{(a_i, b_i)\}_{i \in I}$ ,  $a \leq x_i < y_i \leq b$ ,  $[x_0, y_0] = [a, b]$  and  $\forall i \in \mathbb{N}$  we have that  $[x_{i+1}, y_{i+1}] \subseteq [x_i, y_i]$  and  $y_i - x_i = \frac{b - a}{2^i}$ . Using the previous theorem and the fact that  $\left\{\frac{b - a}{2^i}\right\}_{i \in \mathbb{N}}$  converges to 0, we have  $\exists c \in \mathbb{R}$  such that  $\{c\} = \bigcap_{i \in \mathbb{N}} [x_i, y_i] \subseteq [x_0, y_0] = [a, b]$ . As  $c \in [a, b]$  there exists a  $i \in I$  such that  $c \in (a_i, b_i) \Rightarrow a_i < c < b_i$ . Take now  $\varepsilon = \frac{1}{2} \cdot \min(c - a_i, b_i - c) > 0$  and choose a  $j \in \mathbb{N}$  such that  $y_j - x_j = |y_j - x_j| = \frac{b - a}{2^j} < \varepsilon$   $\left\{\frac{b - a}{2^i}\right\}_{i \in \mathbb{N}}$  converges to 0 (see 11.330). Then  $c \in [x_j, y_j] \Rightarrow x_j \leq c \leq y_j \Rightarrow 0 \leq c - x_j \leq y_j - x_j < \varepsilon \leq c - a_i \Rightarrow -x_j < -a_i \Rightarrow a_i < x_j$  also  $y_j - c < y_j - x_j < \varepsilon \leq b_i - c \Rightarrow y_j < b_i$  which proves that  $a_i < x_j < y_j < b_i \Rightarrow [x_j, y_j] \subseteq (a_i, b_i)$  proving that  $[x_j, y_j]$  is covered by a finite sub-cover of  $\{(a_i, b_i)\}_{i \in \mathbb{N}}$  which is a contradiction. So the assumption that  $[a, b]$  is not compact leads to a contradiction and so  $[a, b]$  is compact.  $\square$

**Theorem 11.268.** Let  $\langle X, d_X \rangle, \langle Y, d_Y \rangle$  be metric spaces  $K \subseteq X$  a compact subset and  $f: X \rightarrow Y$  a continuous function then  $f$  is uniform continuous in  $K$  (see 11.154)

**Proof.** Let  $\varepsilon > 0$  then by continuity of  $f$  at  $x \in K$  (see 11.151) there exists a  $\delta_x \in \mathbb{R}_+$  such that if  $d_X(x, y) < \delta_x$  then  $d_Y(f(x), f(y)) < \frac{\varepsilon}{2}$ . The family  $\{B_{d_X}(x, \delta_x)\}_{x \in K}$  covers  $K$  so by compactness there exist a finite set  $\{x_1, \dots, x_n\}$  such that  $K \subseteq \bigcup_{x \in \{x_1, \dots, x_n\}} B_{d_X}\left(x, \frac{\delta(x)}{2}\right)$ . Take now  $\delta = \min\left(\left\{\frac{\delta(x)}{2} | x \in \{x_1, \dots, x_n\}\right\}\right)$ . Then if  $x, y \in K$  with  $d_X(x, y) < \delta$  then there exists a  $x_i \in \{x_1, \dots, x_n\}$  such that  $x \in B_{d_X}(x_i, \delta(x_i))$  so that  $d_X(x_i, y) \leq d_X(x_i, x) + d_X(x, y) < \frac{\delta(x_i)}{2} + \delta \leq \frac{\delta(x_i)}{2} + \frac{\delta(x_i)}{2} = \delta(x_i)$ , so we have that  $d_Y(f(x), f(y)) \leq d(f(x), f(x_i)) + d(f(x_i), f(y)) \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$   $\square$

### 11.10.1 Filter bases in Topological spaces

To prove Tychonoff's theorem, that essentially says that any product of non empty compact topological spaces is compact in the product topology, we must develop a whole theoretical framework of filter bases.

**Definition 11.269.** Let  $\langle X, \mathcal{T} \rangle$  be a topological space. A filter base  $\mathfrak{U}$  in  $X$  is a nonempty family  $\mathfrak{U} = \{A_\alpha\}_{\alpha \in \mathbb{U}}$  of subsets of  $X$  ( $\mathbb{U} \neq \emptyset$ ) such that

1.  $\forall \alpha \in \mathbb{U}$  we have that  $A_\alpha \neq \emptyset$
2.  $\forall \alpha, \beta \in \mathbb{U}$  we have  $\exists \gamma \in \mathbb{U}$  such that  $A_\gamma \subseteq A_\alpha \cap A_\beta$  (this of course implies then that the intersection is non empty)

Using mathematical induction it is then easy to prove the following theorem

**Theorem 11.270.** Let  $\langle X, \mathcal{T} \rangle$  be a topological space and  $\mathfrak{U} = \{A_\alpha\}_{\alpha \in \mathbb{U}}$  a filter base then for each finite  $\{\alpha_1, \dots, \alpha_n\} \subseteq \mathbb{U}$  there exists a  $\gamma \in \mathbb{U}$  such that  $A_\gamma \subseteq \bigcap_{\alpha \in \{\alpha_1, \dots, \alpha_n\}} A_\alpha$

**Proof.** Let  $S = \{n \in \mathbb{N}_0 \mid \text{if } \{\alpha_1, \dots, \alpha_n\} \subseteq \mathbb{U} \text{ then } \exists \gamma \in \mathbb{U} \text{ such that } A_\gamma \subseteq \bigcap_{\alpha \in \mathbb{U}} A_\alpha\}$  then we have

1. if  $n = 1$  then if  $\{\alpha_1\} \subseteq \mathbb{U}$  then  $A_{\alpha_1} = \bigcup_{\alpha \in \{\alpha_1\}} A_\alpha$  so that  $1 \in S$
2. if  $n \in S$  then if  $\{\alpha_1, \dots, \alpha_{n+1}\} \subseteq \mathbb{U}$  then as  $\{\alpha_1, \dots, \alpha_n\} \subseteq \mathbb{U}$  and  $n \in S$  there exists a  $\gamma_1 \in \mathbb{U}$  such that  $A_{\gamma_1} \subseteq \bigcap_{\alpha \in \{\alpha_1, \dots, \alpha_n\}} A_\alpha$ , as  $\mathfrak{U}$  is a filter base we can then find a  $\gamma \in \mathbb{U}$  such that  $A_\gamma \subseteq A_{\gamma_1} \cap A_{n+1} \subseteq (\bigcap_{\alpha \in \{\alpha_1, \dots, \alpha_n\}} A_\alpha) \cap A_{n+1} = \bigcap_{\alpha \in \{\alpha_1, \dots, \alpha_{n+1}\}} A_\alpha$  and thus we have  $n+1 \in S$

Using mathematical induction we have then  $S = \mathbb{N}_0$  proving the theorem.  $\square$

**Definition 11.271.** Let  $\langle X, \mathcal{T} \rangle$  be a topological space,  $x \in X$  then  $\mathfrak{U}(x) = \{U \mid U \in \mathcal{T} \wedge x \in U\}$  /the set of open sets containing  $x$ / is called the **neighborhood filter base** of  $x$

**Theorem 11.272.** Let  $\langle X, \mathcal{T} \rangle$  be a topological space,  $x \in X$  then  $\{U\}_{U \in \mathfrak{U}(x)}$  (see 2.105) is a filter-base, which is also noted as  $\mathfrak{U}(x)$

**Proof.**

1. If  $U \in \mathfrak{U}(x)$  then  $x \in U \Rightarrow U \neq \emptyset$
2. If  $U, V \in \mathfrak{U}(x) \Rightarrow x \in U \cap V$  and as  $U \cap V$  is open we have that  $U \cap V \in \mathfrak{U}(x)$  so we found a  $W = U \cap V \in \mathfrak{U}(x)$  such that  $W = U \cap V \subseteq U \cap V$   $\square$

**Example 11.273.** If  $\langle X, \mathcal{T} \rangle$  and  $\emptyset \neq A \subseteq X$  then  $\{A\}_{A \in \{A\}}$  is trivially a filter base.

**Theorem 11.274.** Let  $\langle X, \mathcal{T} \rangle$  be a topological space and let  $\mathfrak{U} = \{A_\alpha\}_{\alpha \in \mathbb{U}}$  and  $\mathfrak{W} = \{B_\beta\}_{\beta \in \mathbb{W}}$  then

1.  $\mathfrak{U} \cup \mathfrak{W} \stackrel{\text{defined}}{=} \{A_\alpha \cup B_\beta\}_{(\alpha, \beta) \in \mathbb{U} \times \mathbb{W}}$  is a filter base
2. If  $\forall \alpha \in \mathbb{U}, \forall \beta \in \mathbb{W}$  we have that  $A_\alpha \cap B_\beta \neq \emptyset$  then  $\mathfrak{U} \cap \mathfrak{W} \stackrel{\text{defined}}{=} \{A_\alpha \cap B_\beta\}_{(\alpha, \beta) \in \mathbb{U} \times \mathbb{W}}$  is a filter base

**Proof.**

1. ( $\mathfrak{U} \cup \mathfrak{W}$  is a filter base)
  - a.  $\forall (\alpha, \beta) \in \mathbb{U} \times \mathbb{W}$  we have  $A_\alpha \neq \emptyset \wedge B_\beta \neq \emptyset \Rightarrow A_\alpha \cup B_\beta \neq \emptyset$
  - b.  $\forall (\alpha_1, \beta_1), (\alpha_2, \beta_2) \in \mathbb{U} \times \mathbb{W}$  there exists a  $\gamma_1 \in \mathbb{U}, \gamma_2 \in \mathbb{W}$  such that  $A_{\gamma_1} \subseteq A_{\alpha_1} \cap A_{\alpha_2} \wedge B_{\gamma_2} \subseteq B_{\beta_1} \cap B_{\beta_2}$  then  $\gamma = (\gamma_1, \gamma_2) \in \mathbb{U} \times \mathbb{W}$  and  $A_{\gamma_1} \cup B_{\gamma_2} \subseteq (A_{\alpha_1} \cap A_{\alpha_2}) \cup (B_{\beta_1} \cap B_{\beta_2}) \subseteq A_{\alpha_1} \cup B_{\beta_1} (A_{\alpha_2} \cup B_{\beta_2}) \Rightarrow A_{\gamma_1} \cup B_{\gamma_2} \subseteq (A_{\alpha_1} \cup B_{\beta_1}) \cap (A_{\alpha_2} \cup B_{\beta_2})$
2. ( $\mathfrak{U} \cap \mathfrak{W}$  is a filter base (if  $\forall (\alpha, \beta) \in \mathbb{U} \times \mathbb{W}$  we have that  $A_\alpha \cap B_\beta \neq \emptyset$ ))
  - a.  $\forall (\alpha, \beta) \in \mathbb{U} \times \mathbb{W}$  we have by the hypothesis that  $A_\alpha \cap B_\beta \neq \emptyset$
  - b.  $\forall (\alpha_1, \beta_1), (\alpha_2, \beta_2) \in \mathbb{U} \times \mathbb{W}$  there exists a  $\gamma_1 \in \mathbb{U}, \gamma_2 \in \mathbb{W}$  such that  $A_{\gamma_1} \subseteq A_{\alpha_1} \cap A_{\alpha_2} \wedge B_{\gamma_2} \subseteq B_{\beta_1} \cap B_{\beta_2} \Rightarrow \gamma = (\gamma_1, \gamma_2) \in \mathbb{U} \times \mathbb{W}$  and  $A_{\gamma_1} \cap B_{\gamma_2} \subseteq (A_{\alpha_1} \cap A_{\alpha_2}) \cap (B_{\beta_1} \cap B_{\beta_2}) = (A_{\alpha_1} \cap B_{\beta_1}) \cap (A_{\alpha_2} \cap B_{\beta_2})$   $\square$

**Theorem 11.275.** Let  $\langle X, \mathcal{T} \rangle$  be a topological space and  $\mathfrak{U} = \{A_\alpha\}_{\alpha \in \mathbb{U}}$  be a filter base then it has the **finite intersection property**, that is  $\forall \mathcal{B} \subseteq \mathbb{U}, \mathcal{B}$  is finite we have  $\bigcap_{\beta \in \mathcal{B}} A_\beta \neq \emptyset$

**Proof.** Using 11.270 there exists a  $\gamma \in \mathcal{B}$  such that  $\emptyset \neq A_\gamma \subseteq \bigcap_{\beta \in \mathcal{B}} A_\beta$   $\square$

**Definition 11.276.** Let  $\mathfrak{U} = \{A_\alpha\}_{\alpha \in \mathbb{U}}$  be a filter base in a topological space  $\langle X, \mathcal{T} \rangle$  then we define

1.  $\mathfrak{U}$  converges to  $y \in X$  written as  $\mathfrak{U} \rightarrow y$  iff  $\forall U \in \mathcal{T}$  with  $y \in U$  we have that  $\exists \alpha \in \mathbb{U}$  with  $A_\alpha \subseteq U$
2.  $\mathfrak{U}$  accumulates at  $y \in X$  (or  $y$  is an accumulation point of  $\mathfrak{U}$ ) written as  $\mathfrak{U} \succ y$  iff  $\forall U \in \mathcal{T}$  with  $y \in U$  we have  $\forall \alpha \in \mathbb{U}$  that  $U \cap A_\alpha \neq \emptyset$

**Remark 11.277.** If  $\mathfrak{U} \succ y$  then  $\forall \alpha \in \mathbb{U}$  we have  $[\forall U \in \mathcal{T} \vdash y \in U \text{ we have } U \cap A_\alpha \neq \emptyset] \Leftrightarrow [y \in \overline{A_\alpha}]$  so  $\mathfrak{U} \succ y$  is equivalent with  $y \in \bigcap_{\alpha \in \mathbb{U}} \overline{A_\alpha}$

**Example 11.278.** Let  $\langle X, \mathcal{T} \rangle$  be a topological space then  $\forall x \in X$  we have  $\mathfrak{U}(x) \rightarrow x$

**Proof.** If  $U \in \mathcal{T}$  and  $x \in U$  then by definition  $U \in \mathfrak{U}(x)$  so that we found a  $U \in \mathfrak{U}(x)$  with  $U \subseteq U$   $\square$

**Definition 11.279.** Let  $\mathfrak{U} = \{A_\alpha\}_{\alpha \in \mathbb{U}}, \mathfrak{W} = \{B_\beta\}_{\beta \in \mathbb{W}}$  be two filter bases on a topological space  $\langle X, \mathcal{T} \rangle$  then  $\mathfrak{W}$  is subordinate to  $\mathfrak{U}$  noted by  $\mathfrak{W} \gg \mathfrak{U}$  (or  $\mathfrak{U} \ll \mathfrak{W}$ ) iff  $\forall \alpha \in \mathbb{U}$  we have that  $\exists \beta \in \mathbb{W}$  such that  $B_\beta \subseteq A_\alpha$

**Definition 11.280.** Let  $\mathfrak{U} = \{A_\alpha\}_{\alpha \in \mathbb{U}}, \mathfrak{W} = \{B_\beta\}_{\beta \in \mathbb{W}}$  be two filter bases on a topological space  $\langle X, \mathcal{T} \rangle$  then  $\mathfrak{U} \subseteq \mathfrak{W}$  iff  $\forall \alpha \in \mathbb{U}$  there exists a  $\beta \in \mathbb{W}$  such that  $A_\alpha = B_\beta$

**Theorem 11.281.** Let  $\mathfrak{U} = \{A_\alpha\}_{\alpha \in \mathbb{U}}, \mathfrak{W} = \{B_\beta\}_{\beta \in \mathbb{W}}$  be two filter bases on a topological space  $\langle X, \mathcal{T} \rangle$  then the following hold:

1. If  $\mathfrak{U} \subseteq \mathfrak{W}$  then  $\mathfrak{W} \gg \mathfrak{U}$
2. If  $\mathfrak{W} \gg \mathfrak{U}$  then  $\forall \beta \in \mathbb{W}$  we have that  $\forall \alpha \in \mathbb{U}$  that  $A_\alpha \cap B_\beta \neq \emptyset$
3.  $\mathfrak{U} \rightarrow x$  if and only if  $\mathfrak{U} \gg \mathfrak{U}(x)$
4. If  $\mathfrak{W} = \{A_\alpha\}_{\alpha \in \mathbb{W}}, \mathfrak{W}' = \{A'_\alpha\}_{\alpha \in \mathbb{W}'}, \mathfrak{W}'' = \{A''_\alpha\}_{\alpha \in \mathbb{W}''}$  are filter bases such that  $\mathfrak{W} \gg \mathfrak{W}'$  and  $\mathfrak{W}' \gg \mathfrak{W}''$  then  $\mathfrak{W} \gg \mathfrak{W}''$

**Proof.**

1. Let  $\alpha \in \mathbb{U}$  then as  $\mathfrak{U} \subseteq \mathfrak{W}$  there exists a  $b \in \mathbb{W}$  such that  $A_\alpha = B_b$  so that  $B_b \subseteq A_\alpha$  so that by definition we have  $\mathfrak{W} \gg \mathfrak{U}$
2. This is proved by contradiction, so assume that  $\exists \alpha \in \mathbb{U}$  and  $\exists \beta \in \mathbb{W}$  such that  $A_\alpha \cap B_\beta = \emptyset$  then since  $\mathfrak{W} \gg \mathfrak{U}$   $\exists \gamma \in \mathbb{W}$  such that  $B_\gamma \subseteq A_\alpha$  and then there exists a  $\delta \in \mathbb{W}$  such that  $B_\delta \subseteq B_\gamma \cap B_\beta \subseteq A_\alpha \cap B_\beta$  which as  $B_\delta \neq \emptyset$  means that  $A_\alpha \cap B_\beta \neq \emptyset$  a contradiction.
3.  $\mathfrak{U} \rightarrow x \Leftrightarrow \forall U$  open with  $x \in U$  we have that  $\exists \alpha \in \mathbb{U}$  such that  $A_\alpha \subseteq U \Leftrightarrow \forall U \in \mathfrak{U}(x)$  we have  $\exists \alpha \in \mathbb{U}$  such that  $A_\alpha \subseteq U \Leftrightarrow \mathfrak{U} \gg \mathfrak{U}(x)$
4. Let  $\alpha'' \in \mathbb{W}''$  then because  $\mathfrak{W}' \gg \mathfrak{W}''$  there exists a  $\alpha' \in \mathbb{W}'$  such that  $A'_{\alpha'} \subseteq A''_{\alpha''}$ , because  $\mathfrak{W} \gg \mathfrak{W}'$  there exists a  $\alpha \in \mathbb{U}$  such that  $A_\alpha \subseteq A'_{\alpha'} \subseteq A''_{\alpha''}$  and thus  $\mathfrak{W} \gg \mathfrak{W}''$   $\square$

**Theorem 11.282.** Let  $\langle X, \mathcal{T} \rangle$  be a topological space then  $X$  is Hausdorff  $\Leftrightarrow$  each filter base converges to exactly one point.

**Proof.**

1. ( $\Rightarrow$ ) Assume that  $X$  is Hausdorff,  $\mathfrak{U} = \{A_\alpha\}_{\alpha \in \mathbb{U}}$  so that  $\mathfrak{U} \rightarrow x$  and  $\mathfrak{U} \rightarrow y$  and  $x \neq y$  then there exists  $U, V$  open with  $x \in U \wedge y \in V \wedge U \cap V = \emptyset$ . From  $\mathfrak{U} \rightarrow x$  and  $\mathfrak{U} \rightarrow y$  there exists  $\alpha, \beta \in \mathbb{U}$  such that  $A_\alpha \subseteq U \wedge A_\beta \subseteq V \Rightarrow \emptyset \neq A_\alpha \cap A_\beta \subseteq U \cap V = \emptyset$  a contradiction. So we must have that  $x = y$
2. ( $\Leftarrow$ ) Assume that  $X$  is not Hausdorff then  $\exists x, y \in X$   $x \neq y$  such that  $\forall U, V$  open with  $x \in U \wedge y \in V$  we have  $U \cap V \neq \emptyset$  then by 11.274 we have that  $\mathfrak{W} = \mathfrak{U}(x) \cap \mathfrak{U}(y)$  is a filter base. If  $x \in U$  ( $U$  open)  $\Rightarrow U \in \mathfrak{U}(x)$  and as  $X \in \mathfrak{U}(y)$  then  $U = U \cap X \in \mathfrak{W}$  and thus  $\mathfrak{W} \rightarrow x$ , if  $y \in V$  ( $V$  open)  $\Rightarrow V \in \mathfrak{U}(y)$  and as  $X \in \mathfrak{U}(x)$  we have that  $V = X \cap V \in \mathfrak{W}$  proving that  $\mathfrak{W} \rightarrow y$ . By the hypothesis we have then that  $x = y$  contradicting  $x \neq y$ . So  $X$  must be Hausdorff.  $\square$

**Theorem 11.283.** Let  $\langle X, \mathcal{T} \rangle$  be a topological space and  $\mathfrak{U} = \{A_\alpha\}_{\alpha \in \mathbb{U}}$ ,  $\mathfrak{W} = \{B_\beta\}_{\beta \in \mathbb{W}}$  be filter bases then

1. If  $\mathfrak{U} \rightarrow x$  then  $\mathfrak{U} \succ x$  and if  $X$  is also Hausdorff then if  $\mathfrak{U} \succ y \Rightarrow x = y$
2. Let  $\mathfrak{W} \gg \mathfrak{U}$  then
  - a.  $\mathfrak{U} \rightarrow x \Rightarrow \mathfrak{W} \rightarrow x$
  - b.  $\mathfrak{W} \succ x \Rightarrow \mathfrak{U} \succ x$

**Proof.**

1. Let  $U$  be open with  $x \in U$  then  $\exists \alpha \in \mathbb{U}$  such that  $A_\alpha \subseteq U$ . Given  $\beta \in \mathbb{U}$  then as (see 11.275)  $\emptyset \neq A_\alpha \cap A_\beta \subseteq U \cap A_\beta \Rightarrow \mathfrak{U} \succ x$ . Now if  $X$  is Hausdorff and let  $\mathfrak{U} \succ y$  with  $x \neq y$  then  $\exists U, V$  open such that  $x \in U, y \in V$  and  $U \cap V = \emptyset$  then  $\exists \alpha \in \mathbb{U}$  such that  $A_\alpha \subseteq U$  (as  $\mathfrak{U} \rightarrow x$ ) so that  $A_\alpha \cap V \subseteq U \cap V = \emptyset$  contradicting  $\mathfrak{U} \succ y$  so we must have  $x = y$ .
2. Assume that  $\mathfrak{W} \gg \mathfrak{U}$  then
  - a. Assume that  $\mathfrak{U} \rightarrow x$  then if  $U$  is open with  $x \in U$  then  $\exists \alpha \in \mathbb{U}$  such that  $A_\alpha \subseteq U$  and as  $\mathfrak{W} \gg \mathfrak{U}$  then  $\exists \beta \in \mathbb{W}$  such that  $A_\beta \subseteq A_\alpha \subseteq U$  and thus  $\mathfrak{W} \rightarrow x$
  - b. Assume that  $\mathfrak{W} \succ x$  then if  $U$  is open such that  $x \in U$  then  $\forall \beta \in \mathbb{W}$  we have  $U \cap B_\beta = \emptyset$ , now if  $\alpha \in \mathbb{U}$  then as  $\mathfrak{W} \gg \mathfrak{U}$  there  $\exists \gamma \in \mathbb{W}$  such that  $B_\gamma \subseteq A_\alpha \Rightarrow \emptyset \neq U \cap B_\gamma \subseteq U \cap A_\alpha \Rightarrow \mathfrak{U} \succ x$   $\square$

**Theorem 11.284.** Let  $\langle X, \mathcal{T} \rangle$  be a topological space and let  $\mathfrak{U} = \{A_\alpha\}_{\alpha \in \mathbb{U}}$  be a filter-base then

1.  $\mathfrak{U} \rightarrow x$  if and only if  $\forall \mathfrak{W} = \{B_\beta\}_{\beta \in \mathbb{W}}$  (a filter base) with  $\mathfrak{W} \gg \mathfrak{U}$  we have that there exists a filter-base  $\mathfrak{C} = \{C_\gamma\}_{\gamma \in \mathbb{C}}$  with  $\mathfrak{C} \gg \mathfrak{W}$  with  $\mathfrak{C} \rightarrow x$
2.  $\mathfrak{U} \succ x$  if and only if  $\exists \mathfrak{W} = \{B_\beta\}_{\beta \in \mathbb{B}}$  (a filter base) with  $\mathfrak{W} \gg \mathfrak{U}$  such that  $\mathfrak{W} \rightarrow x$

**Proof.**

1.
  - a.  $(\Rightarrow)$  Assume  $\mathfrak{U} \rightarrow x$  then if  $\mathfrak{W} \gg \mathfrak{U} \xrightarrow{11.283} \mathfrak{W} \rightarrow x$  and as trivially we have  $\mathfrak{W} \gg \mathfrak{W}$  we can use  $\mathfrak{C} = \mathfrak{W}$
  - b.  $(\Leftarrow)$  Assume that  $\mathfrak{U} \not\rightarrow x$  ( $\mathfrak{U}$  does not converge to  $x$ ) then  $\exists U$  open with  $x \in U$  such that  $\forall \alpha \in \mathbb{U}$  we have  $A_\alpha \not\subseteq U \Rightarrow A_\alpha \cap (X \setminus U) \neq \emptyset$  then  $\mathfrak{W} = \{A_\alpha \cap (X \setminus U)\}_{\alpha \in \mathbb{U}}$  forms a filter basis (see 11.273 and 11.274) and we have trivially  $\mathfrak{W} \gg \mathfrak{U}$ . Then by the hypothesis  $\exists \mathfrak{C} = \{C_\gamma\}_{\gamma \in \mathbb{C}}$  such that  $\mathfrak{C} \gg \mathfrak{W}$  and  $\mathfrak{C} \rightarrow x \xrightarrow{11.283} \mathfrak{W} \succ x$  which means that  $\forall \alpha \in \mathbb{U}$  we have that  $(A_\alpha \cap (X \setminus U)) \cap U \neq \emptyset$  which is a contradiction because  $(A_\alpha \cap (X \setminus U)) \cap U \subseteq (X \setminus U) \cap U = \emptyset$ . So we must have that  $\mathfrak{U} \rightarrow x$
2.
  - a.  $(\Rightarrow)$  Assume that  $\mathfrak{U} \succ x$  then if  $U$  is open with  $x \in U$  then  $\forall \alpha \in \mathbb{U}$  we have that  $A_\alpha \cap U \neq \emptyset$ , so using 11.271 and 11.274 we have then that  $\mathfrak{W} = \mathfrak{U} \cap \mathfrak{U}(x) = \{A_\alpha \cap U\}_{(\alpha, U) \in \mathbb{U} \times \mathfrak{U}(x)}$  is a filter base. As  $\forall \alpha \in \mathbb{U}$  we have that  $A_\alpha \cap U \subseteq A_\alpha$  so that  $\mathfrak{W} \gg \mathfrak{U}$ . As by the definition of a filter base  $\mathfrak{U} \neq \emptyset$  there exists a  $\alpha \in \mathbb{U}$  and then we have that  $A_\alpha \cap U \subseteq U$  proving that  $\mathfrak{W} \rightarrow x$
  - b.  $(\Leftarrow)$  If there exists a  $\mathfrak{W} \gg \mathfrak{U}$  with  $\mathfrak{W} \rightarrow x$  then using 11.283 we have  $\mathfrak{W} \succ x$  and using 11.283 again we have then  $\mathfrak{U} \succ x$   $\square$

**Definition 11.285.** Let  $\langle X, \mathcal{T} \rangle$  be a topological space and  $A \subseteq X$  then a filter base  $\mathfrak{U} = \{A_\alpha\}_{\alpha \in \mathbb{U}}$  is said to be *on*  $A$  if and only if  $\forall \alpha \in \mathbb{U}$  we have  $A_\alpha \subseteq A$

**Theorem 11.286.** Let  $\langle X, \mathcal{T} \rangle$  be a topological space and let  $A \subseteq X$  then we have the following equivalence  $y \in \bar{A} \Leftrightarrow \exists \mathfrak{U} = \{A_\alpha\}_{\alpha \in \mathbb{U}}$  which is a filter base on  $A$  with  $\mathfrak{U} \rightarrow x$

**Proof.**

1.  $(\Rightarrow)$  If  $y \in \bar{A}$  then  $\forall U$  open with  $y \in U$  we have  $A \cap U \neq \emptyset$ , using 11.273, 11.271 and 11.274 we have that  $\mathfrak{W} = \{A\} \cap \mathfrak{U}(x) = \{A \cap U\}_{U \in \mathfrak{U}(x)}$  is a filter base, which is on  $A$  [as  $A \cap U \subseteq A$ ]. If  $U$  is open with  $x \in U$  then  $U \in \mathfrak{U}(x)$  and  $A \cap U \subseteq U$  proving that  $\mathfrak{W} \rightarrow x$
2.  $(\Leftarrow)$  If  $\mathfrak{U} = \{A_\alpha\}_{\alpha \in \mathbb{U}}$  is a filter base on  $A$  with  $\mathfrak{U} \rightarrow x$  then if  $U$  is open and  $x \in U$  open then there exists a  $\alpha \in \mathbb{U}$  such that  $A_\alpha \subseteq U$  and as  $A_\alpha \neq \emptyset$  and  $A_\alpha \subseteq A$  we have that  $\emptyset \neq A_\alpha = A_\alpha \cap A \subseteq U \cap A \Rightarrow x \in \bar{A}$   $\square$

**Theorem 11.287.** Let  $\langle X, \mathcal{T} \rangle$ ,  $\langle Y, \mathcal{S} \rangle$  be topological spaces and let  $f: X \rightarrow Y$  be a function. If  $\mathfrak{U} = \{A_\alpha\}_{\alpha \in \mathbb{U}}$  is a filter base in  $X$  then  $f(\mathfrak{U}) = \{f(A_\alpha)\}_{\alpha \in \mathbb{U}}$  is a filter base in  $Y$  define

**Proof.** First as  $\mathbb{U} \neq \emptyset$  and  $\forall \alpha \in \mathbb{U}$  we have  $A_\alpha \neq \emptyset \Rightarrow f(A_\alpha) \neq \emptyset \Rightarrow \{f(A_\alpha)\}_{\alpha \in \mathbb{U}}$  is a non empty family of non empty sets. Second if  $\alpha, \beta \in \mathbb{U}$  then there exists a  $\gamma \in \mathbb{U}$  such that  $A_\gamma \subseteq A_\alpha \cap A_\beta \Rightarrow f(A_\gamma) \subseteq f(A_\alpha \cap A_\beta) \subseteq f(A_\alpha) \cap f(A_\beta)$  so that  $\{f(A_\alpha)\}_{\alpha \in \mathbb{U}}$  is a filter base.  $\square$

**Theorem 11.288.** Let  $\langle X, \mathcal{T} \rangle$ ,  $\langle Y, \mathcal{S} \rangle$  be topological spaces and let  $f: X \rightarrow Y$  be a function then  $f$  is continuous at  $x$  if and only if  $f(\mathfrak{U}(x)) \rightarrow f(x)$

**Proof.**

1.  $(\Rightarrow)$  Assume that  $f$  is continuous at  $x$  and let  $V$  open with  $f(x) \in V$  then  $\exists U$  open with  $x \in U$  such that  $f(U) \subseteq V \Rightarrow f(\mathfrak{U}(x)) \rightarrow f(x)$
2.  $(\Leftarrow)$  Let  $V$  be open with  $f(x) \in V$  then as  $f(\mathfrak{U}(x)) \rightarrow f(x)$  and  $\mathfrak{U}(x) = \{U\}_{U \in \mathfrak{U}(x)}$  there exists a  $U \in \mathfrak{U}(x) \Rightarrow x \in U$ ,  $U$  open such that  $f(U) \subseteq V \Rightarrow f$  is continuous at  $x$   $\square$

**Theorem 11.289.** Let  $\langle X, \mathcal{T} \rangle$ ,  $\langle Y, \mathcal{S} \rangle$  be topological spaces and let  $f: X \rightarrow Y$  be a function then  $f$  is continuous on  $X$  if and only if  $\forall x \in X$  and for every filter base  $\mathfrak{U} = \{A_\alpha\}_{\alpha \in \mathbb{U}}$  in  $X$  with  $\mathfrak{U} \rightarrow x$  we have  $f(\mathfrak{U}) \rightarrow f(x)$

**Proof.**

1.  $(\Rightarrow)$  Assume that  $f$  is continuous and let  $\mathfrak{U} \rightarrow x$  then by 11.281 we have  $\mathfrak{U} \gg \mathfrak{U}(x)$  so if  $U$  is open and  $x \in U$  then  $\exists \alpha \in \mathbb{U}$  such that  $A_\alpha \subseteq U \Rightarrow f(A_\alpha) \subseteq f(U)$  proving that  $f(\mathfrak{U}) \gg f(\mathfrak{U}(x))$ . As  $f$  is continuous we have by 11.288 that  $f(\mathfrak{U}(x)) \rightarrow f(x)$  so using 11.283 and  $f(\mathfrak{U}) \gg f(\mathfrak{U}(x))$  we have  $f(\mathfrak{U}) \rightarrow f(x)$
2.  $(\Leftarrow)$  Let  $A \subseteq X$  and let  $y \in f(\bar{A})$  then  $\exists x \in \bar{A}$  such that  $y = f(x)$  so that by 11.286 there exists a filter base  $\mathfrak{U} = \{A_\alpha\}_{\alpha \in \mathbb{A}}$  on  $A$  so that  $\mathfrak{U} \rightarrow x$ . As  $\mathfrak{U}$  is a filter base on  $A$  we have  $\forall a \in \mathbb{U}$  that  $A_\alpha \subseteq A$  so that  $f(A_\alpha) \subseteq f(A)$  and thus  $f(\mathfrak{U})$  is on  $f(A)$ . This last together with  $f(\mathfrak{U}) \rightarrow f(x)$  gives by 11.286 that  $f(x) \in f(A)$  or  $f(\bar{A}) \subseteq f(A)$  and thus by 11.142 that  $f$  is continuous.  $\square$

**Theorem 11.290.** Let  $\{\langle X_i, \mathcal{T}_i \rangle\}_{i \in I}$  be a family of topological spaces then a filter-base  $\mathfrak{U} = \{A_\alpha\}_{\alpha \in \mathbb{U}}$  in  $X = \prod_{i \in I} X_i$  converges to  $y \in \prod_{i \in I} X_i$  (using the product topology) if and only if  $\forall i \in I$  we have that  $\pi_i(\mathfrak{U}) \rightarrow \pi_i(y) = y(i) = y_i$

**Proof.**

1.  $(\Rightarrow)$  Let  $\mathfrak{U} \rightarrow y$  then as  $\pi_i$  is continuous (if we use the product topology) we have by 11.289 that  $\pi_i(\mathfrak{U}) \rightarrow \pi_i(y)$
2.  $(\Leftarrow)$  Let  $y \in X = \prod_{i \in I} X_i$  and  $U$  open in  $X$  with  $y \in U$  then there exists a  $B$  in the basis of the product topology such that  $y \in B \subseteq U$ . By the definition of the basis of a product topology there exists a finite set  $\mathcal{A} \subseteq I$  such that  $B = \bigcap_{i \in \mathcal{A}} \pi_i^{-1}(U_i)$  where  $U_i$  is a open set in  $X_i$  (see 11.38). So  $\forall i \in \mathcal{A}$  we have that  $\pi_i(y) \in U_i$ . As we have by the hypothesis that  $\forall i \in I$  we have that  $\pi_i(\mathfrak{U}) \rightarrow \pi_i(y)$  there exists  $\forall i \in \mathcal{A}$  a  $\alpha_i \in \mathbb{U}$  such that  $\pi_i(A_{\alpha_i}) \subseteq U_i \Rightarrow A_{\alpha_i} \subseteq \pi_i^{-1}(U_i) \Rightarrow \bigcap_{i \in \mathcal{A}} A_{\alpha_i} \subseteq \bigcap_{i \in \mathcal{A}} \pi_i^{-1}(U_i) = B$ . Using 11.270 and the finiteness of  $\mathcal{A}$  there exists a  $\gamma \in \mathbb{U}$  such that  $A_\gamma \subseteq \bigcap_{i \in \mathcal{A}} A_{\alpha_i} \subseteq B \subseteq U \Rightarrow \mathfrak{U} \rightarrow y$   $\square$

**Definition 11.291.** A filter base  $\mathfrak{U}$  in a topological space  $\langle X, \mathcal{T} \rangle$  is called **maximal** or a **ultra filter** if  $\forall \mathfrak{W}$ ,  $\mathfrak{W}$  a filter base in  $X$  with  $\mathfrak{W} \gg \mathfrak{U}$  we have that  $\mathfrak{U} \gg \mathfrak{W}$ .

**Theorem 11.292.** A filter base  $\mathfrak{U}$  in a topological space  $\langle X, \mathcal{T} \rangle$  is maximal if and only if  $\forall A \subseteq X$  one and only one of the two sets  $A$  and  $X \setminus A$  contains a member of  $\mathfrak{U}$ .

**Proof.**

1. ( $\Rightarrow$ ) Assume that  $\mathfrak{U} = \{A_\alpha\}_{\alpha \in \mathbb{U}}$  is maximal and let  $A \subseteq X$  then if there exists a  $\alpha, \beta \in \mathbb{U}$  such that  $A_\alpha \subseteq A$  and  $A_\beta \subseteq X \setminus A$  then as by definition of a filter base we have  $A_\alpha \cap A_\beta \neq \emptyset$  we have that  $\emptyset \neq A_\alpha \cap A_\beta \subseteq A_\alpha \cap X \setminus A = \emptyset$  a contradiction, so we can not have a  $\alpha, \beta \in \mathbb{U}$  such that  $A_\alpha \subseteq A \wedge A_\beta \subseteq X \setminus A$ . Assume now that  $\forall \alpha \in \mathbb{U}$  we have that  $A_\alpha \subseteq A$  then we have that  $\forall \alpha \in \mathbb{U}$  that  $(X \setminus A) \cap A_\alpha \neq \emptyset$  so  $\mathfrak{W} = \{(X \setminus A) \cap A_\alpha\}_{\alpha \in \mathbb{U}}$  forms a filter base (see 11.273 and 11.274) and we have  $\mathfrak{W} \gg \mathfrak{U}$  [if  $\alpha \in \mathbb{U}$  then  $(X \setminus A) \cap A_\alpha \subseteq A_\alpha$ ] so that by maximality we must have that  $\mathfrak{U} \gg \mathfrak{W}$ . So  $\forall \alpha \in \mathbb{U}$  there exists a  $\beta \in \mathbb{U}$  so that  $A_\beta \subseteq (X \setminus A) \cap A_\alpha \Rightarrow A_\beta \subseteq X \setminus A$
2. ( $\Leftarrow$ ) Let  $\mathfrak{U} = \{A_\alpha\}_{\alpha \in \mathbb{U}}$  be a filter base such that  $\forall A \subseteq X$  one and only one of the two sets  $A, X \setminus A$  contains a member of  $\mathfrak{U}$ . Let then  $\mathfrak{W} = \{B_\beta\}_{\beta \in \mathbb{W}}$  such that  $\mathfrak{W} \gg \mathfrak{U}$ . If  $\beta \in \mathbb{W}$  then [by taking  $B_\beta = A$  in the above) we have the existence of a  $\alpha \in A$  such that either  $A_\alpha \subseteq B_\beta$  or  $A_\alpha \subseteq X \setminus B_\beta$ . The last case would mean that  $A_\alpha \cap B_\beta = \emptyset$  contradicting 11.281 so that we are left with  $A_\alpha \subseteq B_\beta$  proving that  $\mathfrak{U} \gg \mathfrak{W}$  and thus maximality.  $\square$

**Theorem 11.293.** Let  $\mathfrak{U} = \{A_\alpha\}_{\alpha \in \mathbb{U}}$  be any filter base in a topological space  $\langle X, \mathcal{T} \rangle$  then there exists a maximal filter base  $\mathfrak{W}$  in  $X$  with  $\mathfrak{W} \gg \mathfrak{U}$

**Proof.** Let  $\mathcal{A} = \{\mathfrak{W} \mid \mathfrak{W} \text{ is a filterbase such that } \mathfrak{W} \gg \mathfrak{U}\}$  then we have that  $\mathcal{A} \neq \emptyset$  because  $\mathfrak{U} \gg \mathfrak{U}$  and thus  $\mathfrak{U} \in \mathcal{A}$ . We define now a pre-order (see 2.135)  $\leqslant$  on  $\mathcal{A}$  defined by  $\mathfrak{W} \leqslant \mathfrak{W}' \Leftrightarrow \mathfrak{W}' \gg \mathfrak{W}$  this is a pre-order because

1. (reflexivity)  $\mathfrak{W} \gg \mathfrak{W} \Rightarrow \mathfrak{W} \leqslant \mathfrak{W}$
2. (transitivity) Assume that  $\mathfrak{W} \leqslant \mathfrak{W}' \wedge \mathfrak{W}' \leqslant \mathfrak{W}'' \Rightarrow \mathfrak{W}' \gg \mathfrak{W} \wedge \mathfrak{W}'' \gg \mathfrak{W}' \stackrel{11.281}{\Rightarrow} \mathfrak{W}'' \gg \mathfrak{W} \Rightarrow \mathfrak{W} \leqslant \mathfrak{W}''$

Take now  $\mathcal{C} \subseteq \mathcal{A}$  be any chain in the pre-ordered set  $\langle \mathcal{A}, \leqslant \rangle$  and prove that it has a upper bound in  $\mathcal{A}$ . We have two cases to consider:

1. ( $\mathcal{C} = \emptyset$ ) Then we have vacuous that  $\forall \mathfrak{C} \in \mathcal{C}$  we have that  $\mathfrak{C} \leqslant \mathfrak{U}$  so  $\mathfrak{U} \in \mathcal{A}$  is a upper bound of  $\mathcal{C}$
2. ( $\mathcal{C} \neq \emptyset$ ) Define  $\mathfrak{B} = \{A \mid \exists \mathfrak{C} = \{C_\gamma\}_{\gamma \in \mathbb{C}} \in \mathcal{C} \vdash \exists \gamma \in \mathbb{C} \text{ with } A = C_\gamma\}$  then  $\mathfrak{B} = \{B\}_{B \in \mathfrak{B}}$  is a filter base because
  - a. If  $B \in \mathfrak{B}$  then  $\exists \mathfrak{C} = \{C_\gamma\}_{\gamma \in \mathbb{C}} \in \mathcal{C}$  such that  $\exists \gamma \in \mathbb{C}$  with  $B = C_\gamma \neq \emptyset$
  - b. If  $B_1, B_2 \in \mathfrak{B}$  then there exists  $\mathfrak{C}_1 = \{C_\gamma^1\}_{\gamma \in \mathbb{C}_1}, \mathfrak{C}_2 = \{C_\gamma^2\}_{\gamma \in \mathbb{C}_2}, \mathfrak{C}_1, \mathfrak{C}_2 \in \mathcal{C}$  such that  $\exists \gamma_1 \in \mathbb{C}_1 \wedge \exists \gamma_2 \in \mathbb{C}_2$  such that  $B_1 = C_{\gamma_1}^1 \wedge B_2 = C_{\gamma_2}^2$ . Now as  $\mathcal{C}$  is a chain we have either:
    - i. ( $\mathfrak{C}_1 \leqslant \mathfrak{C}_2 \Rightarrow \mathfrak{C}_2 \gg \mathfrak{C}_1 \Rightarrow \exists \gamma \in \mathbb{C}_2$  such that  $C_\gamma^2 \subseteq C_{\gamma_1}^1 = B_1$  and because  $\mathfrak{C}_2$  is a filter base there exists a  $\beta \in \mathbb{C}_2$  such that  $C_\beta^2 \subseteq C_\gamma^2 \cap C_{\gamma_2}^2 \subseteq B_1 \cap B_2$  so that we have found a  $B_3 = C_\beta^2 \in \mathfrak{B}$  such that  $B_3 \subseteq B_1 \cap B_2$ )
    - ii. ( $\mathfrak{C}_2 \leqslant \mathfrak{C}_1 \Rightarrow \mathfrak{C}_1 \gg \mathfrak{C}_2 \Rightarrow \exists \gamma \in \mathbb{C}_1$  such that  $C_\gamma^1 \subseteq C_{\gamma_2}^2 = B_2$  and because  $\mathfrak{C}_1$  is a filter base there exists a  $\beta \in \mathbb{C}_1$  such that  $C_\beta^1 \subseteq C_\gamma^1 \cap C_{\gamma_2}^2 \subseteq B_2 \cap B_1 = B_1 \cap B_2$  so that we have found a  $B_3 = C_\beta^1 \in \mathfrak{B}$  such that  $B_3 \subseteq B_1 \cap B_2$ )

Using (a) and (b) we have proved that  $\mathfrak{B}$  is a filter base. Now as  $\mathcal{C} \neq \emptyset$  there exists a  $\mathfrak{C} = \{C_\gamma\}_{\gamma \in \mathbb{C}} \in \mathcal{C} \subseteq \mathcal{A}$  so that  $\mathfrak{C} \gg \mathfrak{U}$  and thus if  $\alpha \in \mathbb{U}$  we have a  $\gamma \in \mathbb{C}$  such that  $C_\gamma \subseteq A_\alpha$  and as by definition we have that  $C_\gamma \in \mathfrak{B}$  it follows that  $\mathfrak{B} \gg \mathfrak{U}$  so that  $\mathfrak{B} \in \mathcal{A}$ . Finally  $\forall \mathfrak{C} = \{C_\gamma\}_{\gamma \in \mathbb{C}} \in \mathcal{C}$  we have that  $\forall \gamma \in \mathbb{C}$  that  $C_\gamma \subseteq C_\gamma$  and because  $C_\gamma \in \mathfrak{B}$  (by definition of  $\mathfrak{B}$ ) we have that  $\mathfrak{B} \gg \mathfrak{C}$  or  $\mathfrak{C} \leqslant \mathfrak{B}$ . This completes the proof that  $\mathfrak{B}$  is a upper bound in  $\mathcal{A}$  of  $\mathcal{C}$ . Using 2.235 there exists a maximum element  $\mathfrak{M}$  of  $\mathcal{A}$  in the pre-order on  $\mathcal{A}$ . This means (see 2.164) that for every  $\mathfrak{A} \in \mathcal{A}$  with  $\mathfrak{A} \gg \mathfrak{M} \Rightarrow \mathfrak{M} \leqslant \mathfrak{A}$  we have  $\mathfrak{A} \leqslant \mathfrak{M} \Rightarrow \mathfrak{M} \gg \mathfrak{A}$ . Also if  $\mathfrak{Q}$  is a filter base in  $X$  with  $\mathfrak{Q} \gg \mathfrak{M}$  then as  $\mathfrak{M} \gg \mathfrak{U}$  we have that  $\mathfrak{Q} \gg \mathfrak{U}$  so that  $\mathfrak{Q} \in \mathcal{A}$  and thus by the previous that  $\mathfrak{M} \gg \mathfrak{Q}$  proving that  $\mathfrak{M}$  is a maximal filter base (or ultra filter) with  $\mathfrak{M} \gg \mathfrak{U}$ .  $\square$

**Theorem 11.294.** Let  $\langle X, \mathcal{T} \rangle$  be a topological space and  $\mathfrak{U} = \{A_\alpha\}_{\alpha \in \mathbb{U}}$  a maximal filter base then  $\mathfrak{U} \succ x \Leftrightarrow \mathfrak{U} \rightarrow x$  for every  $x \in X$

**Proof.** Because of 11.283 we must only prove that  $\mathfrak{U} \succ x \Rightarrow \mathfrak{U} \rightarrow x$  for a maximal filter base. So let  $x \in U$ ,  $U$  open then by 11.292 there exists a  $\alpha \in \mathbb{U}$  such that either  $A_\alpha \subseteq U$  or  $A_\alpha \subseteq X \setminus U$ . Now as  $\mathfrak{U} \succ x$  we must have that  $A_\alpha \cap U \neq \emptyset$  contradicting  $A_\alpha \subseteq X \setminus U$  so we must have that  $A_\alpha \subseteq U$  or  $\mathfrak{U} \rightarrow x$   $\square$

**Theorem 11.295.** Let  $\langle X, \mathcal{T}_X \rangle$ ,  $\langle Y, \mathcal{T}_Y \rangle$  be topological spaces and let  $\mathfrak{U} = \{A_\alpha\}_{\alpha \in \mathbb{U}}$  is a maximal filter base in  $X$  then for any function  $f: X \rightarrow Y$  we have that  $f(\mathfrak{U})$  is a maximal filter base in  $Y$

**Proof.** By 11.287 we have that  $f(\mathfrak{U}) = \{f(A_\alpha)\}_{\alpha \in \mathbb{U}}$  is a filter base. To prove maximality we use 11.292, so let  $A \subseteq Y$  and take  $f^{-1}(A)$  and  $X \setminus f^{-1}(A)$  then by maximality of  $\mathfrak{U}$  there exists a  $\alpha \in \mathbb{U}$  such that  $A_\alpha \subseteq f^{-1}(A)$  or  $A_\alpha \subseteq X \setminus f^{-1}(A)$  so that we have either  $f(A_\alpha) \subseteq f(f^{-1}(A)) = A$  or  $f(A_\alpha) \subseteq f(X \setminus f^{-1}(A)) \subseteq Y \setminus A$  proving that  $f(\mathfrak{U})$  is maximal.  $\square$

All this hard work is to prove Tychonoffs theorem on the product of compact spaces. The following theorem creates the relation between compactness and filter bases.

**Theorem 11.296.** Let  $\langle X, \mathcal{T} \rangle$  be a topological space then the following are equivalent

1.  $X$  is compact
2. For each family  $\{A_\alpha\}_{\alpha \in \mathcal{A}}$  of closed sets in  $X$  with  $\bigcap_{\alpha \in \mathcal{A}} A_\alpha = \emptyset$  we have the existence of a finite sub-family  $\{A_\beta\}_{\beta \in \mathcal{B}}$  [where  $\mathcal{B} = \{\alpha_1, \dots, \alpha_n\} \subseteq \mathcal{A}$ ] such that  $\bigcap_{\beta \in \mathcal{B}} A_\beta = \emptyset$  (or in other words if  $\{A_\alpha\}_{\alpha \in \mathcal{A}}$  is a family such that for every finite  $\mathcal{B} \subseteq \mathcal{A}$  we have that  $\bigcap_{\alpha \in \mathcal{B}} A_\alpha \neq \emptyset$  (the finite intersection property) then we must have  $\bigcap_{\alpha \in \mathcal{A}} A_\alpha \neq \emptyset$ )
3. Each filter base  $\{A_\alpha\}_{\alpha \in \mathbb{U}}$  in  $X$  has at least one accumulation point
4. Each maximal filter base in  $X$  converges

**Proof.**

1. **(1  $\Rightarrow$  2)** Let  $\{A_\alpha\}_{\alpha \in \mathcal{A}}$  be a family of closed sets with  $\bigcap_{\alpha \in \mathcal{A}} A_\alpha = \emptyset$  then  $X = X \setminus \emptyset = X \setminus (\bigcap_{\alpha \in \mathcal{A}} A_\alpha) = \bigcup_{\alpha \in \mathcal{A}} (X \setminus A_\alpha)$  and because  $X \setminus A_\alpha$  is open  $\forall \alpha \in \mathcal{A}$  we have by compactness the existence of a finite  $\mathcal{B} \subseteq \mathcal{A}$  such that  $X = \bigcup_{\beta \in \mathcal{B}} (X \setminus A_\beta) \Rightarrow \emptyset = X \setminus X = X \setminus (\bigcup_{\beta \in \mathcal{B}} (X \setminus A_\beta)) = \bigcap_{\beta \in \mathcal{B}} (X \setminus (X \setminus A_\beta)) = \bigcap_{\beta \in \mathcal{B}} A_\beta$
2. **(2  $\Rightarrow$  1)** Let  $\{U_\alpha\}_{\alpha \in \mathcal{A}}$  be a family of open sets such that  $X = \bigcup_{\alpha \in \mathcal{A}} U_\alpha \Rightarrow \emptyset = X \setminus X = X \setminus (\bigcup_{\alpha \in \mathcal{A}} U_\alpha) = \bigcap_{\alpha \in \mathcal{A}} (X \setminus U_\alpha)$  and as we have that  $\forall \alpha \in \mathcal{A}$  that  $X \setminus U_\alpha$  is closed we have by the hypothesis that  $\exists \mathcal{B} \subseteq \mathcal{A}$ ,  $\mathcal{B}$  finite such that  $\emptyset = \bigcap_{\beta \in \mathcal{B}} (X \setminus U_\beta) \Rightarrow X \setminus \emptyset = X \setminus (\bigcap_{\beta \in \mathcal{B}} (X \setminus U_\beta)) = \bigcup_{\beta \in \mathcal{B}} (X \setminus (X \setminus U_\beta)) = \bigcup_{\beta \in \mathcal{B}} U_\beta$  proving compactness
3. **(2  $\Rightarrow$  3)** Let  $\mathfrak{U} = \{A_\alpha\}_{\alpha \in \mathbb{U}}$  be a filter base, then  $\forall \mathcal{B} \subseteq \mathbb{U} \vdash \mathcal{B}$  is finite we have that  $\emptyset \neq \bigcap_{\beta \in \mathcal{B}} A_\beta$  (see 11.270) then  $\emptyset \neq \bigcap_{\beta \in \mathcal{B}} A_\beta \subseteq \bigcap_{\beta \in \mathcal{B}} \overline{A_\beta}$  and then by (2) we have  $\emptyset \neq \bigcap_{\alpha \in \mathbb{U}} \overline{A_\alpha} \Rightarrow \exists x \in \bigcap_{\alpha \in \mathbb{U}} \overline{A_\alpha}$  and then by 11.277 we have that  $x$  is a accumulation point of  $\mathfrak{U}$ .
4. **(3  $\Rightarrow$  4)** Let  $\mathfrak{U}$  be a maximal filter base then by (3) there exists a  $x \in X$  such that  $\mathfrak{U} \succ x$ , and then using 11.294 we have  $\mathfrak{U} \rightarrow x$
5. **(4  $\Rightarrow$  3)** Let  $\mathfrak{U}$  be a filter base then using 11.293 there exist a maximal filter base  $\mathfrak{W} \gg \mathfrak{U}$  and then by (4) we have then  $\exists x \in X$  such that  $\mathfrak{W} \rightarrow x \stackrel{11.294}{\Rightarrow} \mathfrak{W} \succ x$  and then using 11.283 we have  $\mathfrak{U} \succ x$
6. **(3  $\Rightarrow$  2)** Let  $\{F_\alpha\}_{\alpha \in \mathcal{A}}$  be a family of closed sets such that  $\forall \mathcal{B} \subseteq \mathcal{A}$ ,  $\mathcal{B}$  finite we have  $\bigcap_{\beta \in \mathcal{B}} F_\beta \neq \emptyset$  then if  $\mathfrak{U} = \{\bigcap_{\beta \in \mathcal{B}} A_\beta \mid \mathcal{B} \subseteq \mathcal{A}, \mathcal{B}$  finite} we form the family  $\mathfrak{U} = \{A\}_{A \in \mathfrak{U}}$ . We prove now that  $\mathfrak{U}$  is a filter base, first of all we have by the hypothesis that  $\bigcap_{\beta \in \mathcal{B}} F_\beta \neq \emptyset$ , also if  $B_1, B_2 \in \mathfrak{U}$  then  $\exists \mathcal{B}_1, \mathcal{B}_2 \subseteq \mathcal{A}$ ,  $\mathcal{B}_1, \mathcal{B}_2$  finite and thus  $\mathcal{B}_1 \cup \mathcal{B}_2$  is finite such that  $B_1 \cap B_2 = (\bigcap_{\beta \in \mathcal{B}_1} F_\beta) \cap (\bigcap_{\beta \in \mathcal{B}_2} F_\beta) = \bigcap_{\beta \in \mathcal{B}_1 \cup \mathcal{B}_2} F_\beta \in \mathfrak{U}$  [if  $x \in (\bigcap_{\beta \in \mathcal{B}_1} F_\beta) \cap (\bigcap_{\beta \in \mathcal{B}_2} F_\beta)$  then  $\forall \beta \in \mathcal{B}_1$ ,  $\forall \alpha \in \mathcal{B}_1$  we have  $x \in F_\beta \wedge x \in F_\alpha \Rightarrow$  if  $\gamma \in \mathcal{B}_1 \cup \mathcal{B}_2$  then  $x \in F_\gamma$ , if  $x \in \bigcap_{\gamma \in \mathcal{B}_1 \cup \mathcal{B}_2} F_\gamma$  then if  $\gamma \in \mathcal{B}_1 \cup \mathcal{B}_2$  we have  $x \in F_\gamma$  so that if  $\alpha \in \mathcal{B}_1$  and  $\beta \in \mathcal{B}_1$  then  $x \in (\bigcap_{\beta \in \mathcal{B}_1} F_\beta) \cap (\bigcap_{\beta \in \mathcal{B}_2} F_\beta)$ ] so by (3)  $\exists x \in X$  such that  $\mathfrak{U} \succ x$  or using 11.277 we have  $x \in \bigcap_{B \in \mathfrak{U}} \bar{B}$  so  $\forall B \in \mathfrak{U}$  we have  $x \in \bar{B}$  now as  $\forall \alpha \in \mathcal{A}$  we have  $\{A_\alpha\}$  is finite and thus  $F_\alpha = \bigcap_{\beta \in \{A_\alpha\}} F_\beta \in \mathfrak{U} \Rightarrow x \in \overline{F_\alpha} \stackrel{A_\alpha \text{ is closed}}{=} F_\alpha \Rightarrow x \in \bigcap_{\alpha \in \mathcal{A}} F_\alpha \Rightarrow \bigcap_{\alpha \in \mathcal{A}} F_\alpha \neq \emptyset$   $\square$

### 11.10.2 Product of compact spaces

Let's use now filter bases to prove Tychonoff's theorem

**Theorem 11.297. (Tychonoff's)** *Let  $\{\langle X_i, \mathcal{T}_i \rangle\}_{i \in I}$  be a family non empty topological spaces then  $\prod_{i \in I} X_i$  is compact in the product topology if and only if  $\forall i \in I$  we have that  $X_i$  is compact*

**Proof.**

1.  $(\Rightarrow)$  As the projection map  $\pi_i: \prod_{j \in I} X_j \rightarrow X_i$  is continuous (see 11.148) we have using 11.253 that  $X_i = \pi_i(\prod_{j \in I} X_j)$  is compact.
2.  $(\Leftarrow)$  Assume that  $\forall i \in I$  we have that  $X_i$  is compact. Let  $\mathfrak{U}$  be a maximum filter base in  $\prod_{i \in I} X_i$  then using 11.287 we have that  $\forall j \in I$  that  $\pi_j(\mathfrak{U})$  is a maximum filter base. Using the compactness of  $X_i$  we have using 11.296 that  $\exists x_i \in X_i$  such that  $\pi_i(\mathfrak{U}) \rightarrow x_i$ . Define then  $x \in \prod_{j \in I} X_j$  by  $x(i) = x_i$  so that  $\pi_i(\mathfrak{U}) \rightarrow \pi_i(x)$ . Using 11.290 we have then that  $\mathfrak{U} \rightarrow x$  which by 11.296 means that  $\prod_{i \in I} X_i$  is convergent.  $\square$

**Corollary 11.298.** *Let  $\{\langle X_i, \mathcal{T}_i \rangle\}_{i \in I}$  be a family of non-empty topological spaces then if  $\forall i \in I$  we have that  $C_i \subseteq X_i$  then  $\prod_{i \in I} C_i$  is compact if and only if  $\forall i \in I$  we have that  $C_i$  is compact (using the subspace topology).*

**Proof.** First note that the subspace topology of  $\prod_{i \in I} C_i$  is equal to the product topology of the subspace topologies (see 11.40). As by definition a subset is compact is if it is compact in the subspace topology we can use the previous theorem to prove this theorem.  $\square$

**Example 11.299.** Let  $\langle \mathbb{R}^n, \mathcal{T} \rangle$  be the set of real tuple's together with the product topology (based on the trivial normed topology on  $\langle \mathbb{R}, || \cdot || \rangle$ ) then  $\prod_{i \in \{1, \dots, n\}} [a_i, b_i]$  is compact.

**Proof.** Using 11.267 we have  $\forall i \in \{1, \dots, n\}$  that  $[a_i, b_i]$  is a compact set in  $\mathbb{R}$  (using the norm topology generated in  $\mathbb{R}$ , so using the previous corollary to Tychonoff's theorem we have that  $\prod_{i \in \{1, \dots, n\}} [a_i, b_i]$  is compact.  $\square$

**Theorem 11.300. (Heine-Borel)** *Let  $\langle \mathbb{R}^n, \mathcal{T} \rangle$  be the set of real tuple's together with the product topology (based on the trivial normed topology on  $\langle \mathbb{R}, || \cdot || \rangle$ ) then a subset of  $C \subseteq \mathbb{R}^n$  is compact if and only if  $C$  is bounded and closed.*

**Proof.**

1.  $(\Rightarrow)$  Using 11.251 and 11.67 we have that  $C$  is bounded. As by 11.228 and 11.67 we have that  $\langle \mathbb{R}^n, \mathcal{T} \rangle$  is Hausdorff and using 11.254 that  $C$  is closed.
2.  $(\Leftarrow)$  Assume now that  $C$  is closed and bounded. If  $C$  is empty then it is trivially compact so for the rest of the proof we assume that  $C$  is not empty. Thus there exists a  $c \in C$ . If we use the canonical norm  $\|x\| = \max(\{\pi_i(x) \mid i \in \{1, \dots, n\}\})$  and the metric  $d(x, y) = \|x - y\| = \max(\{\|\pi_i(x) - \pi_i(y)\| \mid i \in \{1, \dots, n\}\}) = \max(\{\|\pi_i(x) - \pi_i(y)\| \mid i \in \{1, \dots, n\}\})$  then we have as  $C$  is bounded that there exists a  $M > 0$  such that  $\forall y \in C$  we have that  $d(c, y) \leq M$  or  $\forall i \in \{1, \dots, n\}$  we have that  $|\pi_i(c) - \pi_i(y)| \leq M \Rightarrow \pi_i(y) \in [\pi_i(c) - M, \pi_i(c) + M]$  and thus  $C \subseteq \prod_{i \in \{1, \dots, n\}} [\pi_i(c) - M, \pi_i(c) + M]$  which is compact by 11.267 and the corollary to Tychonoff's theorem. As  $C = C \cap (\prod_{i \in \{1, \dots, n\}} [\pi_i(c) - M, \pi_i(c) + M])$  is closed in the subspace topology of the compact set  $\prod_{i \in \{1, \dots, n\}} [\pi_i(c) - M, \pi_i(c) + M]$  (as  $C$  is closed) we can use 11.255 to prove that  $C$  is compact.  $\square$

**Theorem 11.301.** *Let  $\langle X, \mathcal{T} \rangle$  be a topological space and  $f: X \rightarrow \mathbb{R}$  be a continuous real function (using the canonical topology on  $\mathbb{R}$ ). Then for every compact  $C \subseteq X$  there exists a  $M_C > 0$  such that  $\forall x \in C$  we have  $|f(x)| \leq M_C$*

**Proof.** First if  $C = \emptyset$  then the theorem is vacuously satisfied so assume that  $C \neq \emptyset$ . Using 11.253 it follows that  $f(C)$  is compact so that by Heine-Borel (see 11.300) we find that  $f(C)$  is bounded and closed. So  $\exists k > 0$  so that  $\forall y_1, y_2 \in f(C)$  we have that  $|y_1 - y_2| \leq k$  and thus if  $x_1, x_2 \in C \Rightarrow f(x_1), f(x_2) \in f(C) \Rightarrow |f(x_1) - f(x_2)| \leq k$ . As  $C \neq \emptyset$  there exists a  $c_0 \in C$ , take then  $M_C = k + |f(c_0)|$  then  $\forall c \in C$  we have  $|f(c)| \leq |f(c) - f(c_0)| + |f(c_0)| \leq k + |f(c)| = M_C$ .  $\square$

**Theorem 11.302.** Let  $\langle X, \mathcal{T} \rangle$  be a topological space and  $f: X \rightarrow \mathbb{R}$  be a continuous function. Then for every non empty compact subset  $C \subseteq X$  there exists a  $x_1, x_2 \in C$  such that  $\forall c \in C$  we have that  $f(x_1) \leq f(c) \leq f(x_2)$ .

**Proof.** By 11.253 we have that  $f(C)$  is compact so that by 11.300  $f(C)$  is bounded and closed. Using then 11.301 there exists a  $M_C$  such that  $\forall c \in C$  we have  $|f(c)| \leq M_C \Rightarrow -M_C \leq f(c) \leq M_C$ . By the lower (upper) bound property of the real's (see 9.43 and 2.180) and the fact that  $f(C) \neq \emptyset$  there exists a  $M_1 = \inf(f(C))$  and  $M_2 = \sup(f(C))$ . If now  $U \subseteq \mathbb{R}$  is a open set with  $M_1 \in U$  (or  $M_2 \in U$ ) then there exists a open ball  $]a, b[$  such that  $M_1 \in ]a, b[ \Rightarrow a < M_1 < b$  (or  $M_2 \in ]a, b[ \subseteq U \Rightarrow a < M_2 < b$ ) and by the definition of  $\inf$  (or  $\sup$ ) we have the existence of a  $y \in f(C)$  such that  $a < M_1 \leq y < b \Rightarrow y \in U$  (or  $a < y \leq M_2 < b \Rightarrow y \in U$ ) proving that  $y \in f(C) \cap U \Rightarrow f(C) \cap U \neq \emptyset$ . Using 11.20 we have then that  $M_1, M_2 \in \overline{f(C)}$   $\underset{f(C) \text{ is compact}}{=} \underset{f(C) \text{ is closed}}{=} f(C)$  so that there exists

a  $c_1, c_2 \in C$  such that  $M_1 = f(c_1), M_2 = f(c_2)$  and taking in account the definition of  $\inf$  and  $\sup$  we have then that  $\forall c \in C$  we have  $f(c_1) = M_1 = \inf(f(C)) \leq f(c) \leq \sup(f(C)) = M_2 = f(c_2) \Rightarrow f(c_1) \leq f(c) \leq f(c_2)$ .  $\square$

**Theorem 11.303.** All norms on  $\mathbb{R}^n$  are equivalent (see 11.85)

**Proof.** Assume that  $\|\cdot\|$  is the maximum norm in  $\mathbb{R}^n$  and  $\|\cdot\|^*$  another norm in  $\mathbb{R}^n$  then by 11.164 we have that  $\|\cdot\|^*: \mathbb{R}^n \rightarrow \mathbb{R}$  is a continuous mapping (using the maximum norm topology on  $\mathbb{R}^n$ ). Define then  $S^{n-1} = \{x \in \mathbb{R}^n \mid \|x\| = 1\}$  which is closed because  $\{1\}$  is closed [as a metric space is Hausdorff (see 11.228) and every finite set in a Hausdorff space is closed (see 11.229)] and  $S^{n-1} = \|\cdot\|^{-1}(\{1\})$  together with 11.142.  $S^{n-1}$  is also bound because if  $x, y \in S^{n-1}$  then  $\|x - y\| \leq \|x\| + \|y\| \leq 1 + 1 = 2$ . According to the Heine Borell theorem (see 11.300) and the fact that the product topology on  $\mathbb{R}^n$  is the topology generated by the maximum norm (see 11.82) we have then that  $S^{n-1}$  is compact in  $\mathbb{R}^n$  using the maximum norm. Using the previous theorem (see 11.302) there exists a  $x_1, x_2 \in S^{n-1}$  such that  $\forall x \in S^{n-1}$  we have  $\|x_1\|^* \leq \|x\|^* \leq \|x_2\|^* \Rightarrow \alpha \leq \|x\|^* \leq \beta$  if we define  $\alpha = \|x_1\|^* > 0$  ( $x_1 \neq 0$  as  $\|x_1\| = 1$ ) and  $\beta = \|x_2\|^* > 0$  ( $x_2 \neq 0$  as  $\|x_2\| = 1$ ). If now  $x \in \mathbb{R}^n$  then we have either

1.  $(x = 0) \Rightarrow \|x\| = 0 = \|x\|^* \Rightarrow \alpha \cdot \|x\| \leq \|x\|^* \leq \beta \cdot \|x\|$
2.  $(x \neq 0)$  then  $\|x\| \neq 0$  and  $\left\| \frac{1}{\|x\|} \cdot x \right\| = \frac{\|x\|}{\|x\|} = 1 \Rightarrow \alpha \leq \left\| \frac{1}{\|x\|} \cdot x \right\|^* \leq \beta \Rightarrow \alpha \leq \frac{1}{\|x\|} \cdot \|x\| \leq \beta \Rightarrow \alpha \cdot \|x\| \leq \|x\|^* \leq \beta \cdot \|x\|$

Using 11.85 we find that  $\|\cdot\|$  and  $\|\cdot\|^*$  are equivalent, so every norm generates the same topology as  $\|\cdot\|$  and thus all the norms are equivalent.  $\square$

**Theorem 11.304.** Let  $\langle X, \|\cdot\| \rangle$  be a normed space,  $n \in \mathbb{N}_0$  such that  $\varphi: \mathbb{R}^n \rightarrow X$  a isomorphism then if  $C \subseteq X$  that  $C$  is compact if and only if  $C$  is closed and bounded

**Proof.** Define  $\|\cdot\|_n: \mathbb{R}^n \rightarrow X$  by  $\|x\|_n = \|\varphi(x)\|$  then we have

1. If  $x \in \mathbb{R}^n$  then  $\|x\|_n = \|\varphi(x)\| \geq 0$
2. If  $\alpha \in \mathbb{R}$ ,  $x \in \mathbb{R}^n$  then  $\|\alpha \cdot x\|_n = \|\varphi(\alpha \cdot x)\| \underset{\varphi \text{ is linear}}{=} \|\alpha \cdot \varphi(x)\| \underset{\|\cdot\| \text{ is a norm}}{=} |\alpha| \cdot \|\varphi(x)\| = |\alpha| \cdot \|x\|_n$
3. If  $x, y \in \mathbb{R}^n$  then  $\|x + y\|_n = \|\varphi(x + y)\| \underset{\varphi \text{ is linear}}{=} \|\varphi(x) + \varphi(y)\| \leq \|\varphi(x)\| + \|\varphi(y)\| = \|x\|_n + \|y\|_n$
4. If  $\|x\|_n = 0$  then  $\|\varphi(x)\| = 0 \underset{\|\cdot\| \text{ is a norm}}{\Rightarrow} \varphi(x) \underset{\varphi \text{ is a isomorphism}}{=} x = 0$

The above proves that  $\|\cdot\|_n$  is a norm an by construction a isometry, using 11.172 we have then that  $\varphi$  is a homeomorphism and isometry between the normed spaces  $\langle \mathbb{R}^n, \|\cdot\|_n \rangle$  and  $\langle X, \|\cdot\| \rangle$ . Let now  $C \subseteq X$  be a closed set then as  $\varphi$  is a homeomorphism we have that  $(\varphi^{-1})(C)$  is closed. Further as  $C$  is bounded we have that there exists a  $0 < M$  such that  $\forall x, y \in C$  we have  $\|x - y\| \leq M$ . Hence if  $x', y' \in (\varphi^{-1})(C)$  we have that  $x' = (\varphi^{-1})(x), y' = (\varphi^{-1})(y)$  where  $x, y \in C$  so that  $x = \varphi(x')$ ,  $y = \varphi(y')$  hence  $\|x' - y'\|_2 = \|\varphi(x' - y')\| = \|\varphi(x') - \varphi(y')\| = \|x - y\| \leq M$  proving that  $(\varphi^{-1})(C)$  is closed and bounded in the topology generated by  $\|\cdot\|_n$  on  $\mathbb{R}^n$ . Using the previous theorem (11.303) we have that, if  $\mathcal{T}$  is the product topology generated by the maximum norm (see 11.82), that:

If  $C \subseteq X$  is closed and bounded in  $\langle X, \|\cdot\| \rangle$  then  $(\varphi^{-1})(C)$  is closed and bounded in  $\langle \mathbb{R}^n, \mathcal{T} \rangle$  (11.33)

Using the Heine Borell theorem (see 11.300) we have then that  $(\varphi^{-1})(C)$  is compact and as  $\varphi$  is a homeomorphism and thus continuous we have by 11.253 that  $C = \varphi((\varphi^{-1})(C))$  is compact. Hence we conclude that

$$\text{If } C \subseteq X \text{ is closed and bounded in } \langle X, \|\cdot\| \rangle \text{ then } C \text{ is compact in } \langle X, \|\cdot\| \rangle \quad (11.34)$$

If  $C$  is compact then as  $\varphi: \mathbb{R}^n \rightarrow X$  is a homeomorphism between  $\langle \mathbb{R}^n, \|\cdot\|_n \rangle$  and  $\langle X, \|\cdot\| \rangle$  we have that  $\varphi^{-1}: X \rightarrow \mathbb{R}^n$  is continuous so that by 11.253 we have that  $(\varphi^{-1})(C)$  is compact. As by the previous theorem (11.303) we have that the topology generated by  $\|\cdot\|_n$  is the same as the product topology generated by the maximum norm (see 11.82) it follows from the Heine Borell theorem (see 11.300) we have that  $(\varphi^{-1})(C)$  is closed and bounded. As  $\varphi^{-1}$  is continuous we have  $C = \varphi((\varphi^{-1})(C)) = (\varphi^{-1})^{-1}((\varphi^{-1})(C))$  is closed. Further as  $\varphi^{-1}(C)$  is bounded there exists  $M > 0$  such that  $\forall x, y \in \varphi^{-1}(C)$  we have  $\|x - y\|_n < M$ , hence if  $x, y \in C$  then  $\varphi^{-1}(x), \varphi^{-1}(y) \in (\varphi^{-1})(C) \Rightarrow \|\varphi^{-1}(x) - \varphi^{-1}(y)\|_n \leq M$ . So  $\|x - y\| = \|\varphi(\varphi^{-1}(x)) - \varphi(\varphi^{-1}(y))\| = \|\varphi(\varphi^{-1}(x) - \varphi^{-1}(y))\| = \|\varphi^{-1}(x) - \varphi^{-1}(y)\|_n \leq M$  proving that

$$\text{If } C \subseteq X \text{ is compact in } \langle X, \|\cdot\| \rangle \text{ then } C \text{ is closed and bounded in } \langle X, \|\cdot\| \rangle \quad (11.35) \quad \square$$

We can use this to prove that closed bounded sets in  $\mathbb{C}$  are compact

**Corollary 11.305.** *Let  $\langle \mathbb{C}, \|\cdot\| \rangle$  be the complex space with a norm then if  $C \subseteq \mathbb{C}$  we have that  $C$  is compact if and only if  $C$  is closed and bounded*

**Proof.** Using 10.182 we have that there exists a isomorphism  $\mathcal{C}: \mathbb{R}^2 \rightarrow \mathbb{C}$  between the vector spaces  $\langle \mathbb{R}^2, +, \cdot \rangle$  and  $\langle \mathbb{C}, +, \cdot \rangle$ . The theorem is then prove by applying the previous theorem (see 11.304).  $\square$

**Theorem 11.306.** *Let  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle$  be finite dimensional spaces then every linear function  $L: X \rightarrow Y$  is continuous.*

**Proof.** Using 11.174 there exists  $\varphi_1: X \rightarrow \mathbb{R}^n, \varphi_2: Y \rightarrow \mathbb{R}^m$  and norms  $\|\cdot\|_1, \|\cdot\|_2$  on  $\mathbb{R}^n$  and  $\mathbb{R}^m$  where  $\varphi_1, \varphi_2$  are isometries, homeomorphism's and isomorphism's (using the norms  $\|\cdot\|_1, \|\cdot\|_2$ ). So given a linear map  $L: X \rightarrow Y$  we have that  $L' = \varphi_2 \circ L \circ \varphi_1^{-1}: \mathbb{R}^n \rightarrow \mathbb{R}^m$  is linear (as isomorphism's are linear). By 11.181 the linear mapping  $L': \mathbb{R}^n \rightarrow \mathbb{R}^m$  is continuous if we use the maximum norms on  $\mathbb{R}^n, \mathbb{R}^m$ , by the equivalence of the norms of  $\mathbb{R}^n, \mathbb{R}^m$  (see 11.303) we have also that  $L'$  is continuous if we use the norms  $\|\cdot\|_1, \|\cdot\|_2$ . Then as  $L = \varphi_2^{-1} \circ L' \circ \varphi_1$  we have that  $L$  is continuous as  $\varphi_2^{-1}, L'$  and  $\varphi_1$  are continuous.  $\square$

**Theorem 11.307.** *Every finite  $n$ -dimensional normed space  $\langle X, \|\cdot\|_X \rangle$  is isomorphic and homeomorphic with every other  $n$ -dimensional normed space  $\langle Y, \|\cdot\|_Y \rangle$ .*

**Proof.** Using 11.174 there exists  $\varphi_1: X \rightarrow \mathbb{R}^n, \varphi_2: Y \rightarrow \mathbb{R}^n$  and norms  $\|\cdot\|_1, \|\cdot\|_2$  on  $\mathbb{R}^n$  and  $\mathbb{R}^n$  where  $\varphi_1, \varphi_2$  are isometries, homeomorphism's and isomorphism's (using the norms  $\|\cdot\|_1, \|\cdot\|_2$ ). Using 11.303 we have that  $\|\cdot\|_1$  and  $\|\cdot\|_2$  must be equivalent so that  $\varphi_1, \varphi_2$  are also homeomorphism's using the topology generated by  $\|\cdot\|_1$  so that  $\varphi_2^{-1} \circ \varphi_1: X \rightarrow Y$  is a homeomorphism as a composition of homeomorphism (see 11.169 and 11.168) and also a isomorphism (by 10.184 and 10.183).  $\square$

## 11.11 Convergence

### 11.11.1 Sequences

**Notation 11.308.** *Given a  $m \in \mathbb{N}$  we note by  $\{m, \dots, \infty\}$  the set  $\{n \in \mathbb{N} | n \geq m\}$ . For example  $\{0, \dots, \infty\} = \mathbb{N}$  and  $\{1, \dots, \infty\} = \mathbb{N}_0$ .*

**Definition 11.309.** *Let  $X$  be a set and  $k \in \mathbb{N}$  then  $\{x_i\}_{i \in \{k, \dots, \infty\}} \subseteq X$  is called a s sequence (here we extend the concept of sequences as families of the form  $\{x_i\}_{i \in \mathbb{N}}$ )*

**Definition 11.310.** Let  $\langle X, \leq \rangle$  be a partially ordered set and  $\{x_i\}_{i \in \{k, \dots, \infty\}} \subseteq X$  a sequence such that  $\forall i \in \{k, \dots, \infty\}$  we have  $x_i \leq x_{i+1}$  [or  $x_{i+1} \leq x_i$ ] then  $\{x_i\}_{i \in \{k, \dots, \infty\}}$  is called a **increasing (decreasing) sequence**.

The following lemma will be used a lot of times when dealing with increasing (decreasing) sequences.

**Lemma 11.311.** Let  $\langle X, \leq \rangle$  be a partially ordered set and  $\{x_i\}_{i \in \{k, \dots, \infty\}}$  a increasing (decreasing) sequence then  $\forall n, m \in \{k, \dots, \infty\}$  with  $n \leq m$  we have  $x_n \leq x_m$  ( $x_m \leq x_n$ )

**Proof.** This is proved by induction so let  $n \in \{k, \dots, \infty\}$  and take  $\mathcal{S}_n = \{i \in \mathbb{N} \mid x_n \leq x_{n+i}\}$  then we have

$$0 \in \mathcal{S}_n. \text{ as } x_n = x_{n+0} \leq x_{n+0}$$

$$i \in \mathcal{S} \Rightarrow i+1 \in \mathcal{S}_n. \text{ as } i \in \mathcal{S}_n \text{ we have } x_n \leq x_{n+i} \text{ and as } x_{n+i} \leq x_{(n+i)+1} \text{ we have } x_n \leq x_{n+(i+1)}$$

proving that  $i+1 \in \mathcal{S}_n$

Using mathematical induction we have then that  $\mathcal{S}_n = \mathbb{N}$ . If now  $n \leq m$  then  $m = n + (m - n)$  and as  $(m - n) \in \mathbb{N} = \mathcal{S}_n$  we have  $x_n \geq x_{n+(m-n)} = x_m$   $\square$

**Definition 11.312.** If  $\langle X, \mathcal{T} \rangle$  is a topological space then a sequence  $\{x_i\}_{i \in \{m, \dots, \infty\}}$  in  $X$  is convergent to  $x \in X$  if  $\forall U \in \mathcal{T}$  with  $x \in U$  there exists a  $N \in \{m, \dots, \infty\}$  such that  $\forall n \in \mathbb{N}$  with  $n \geq N$  we have  $x_n \in U$ . We call  $x$  a limit of the sequence  $\{x_i\}_{i \in \{m, \dots, \infty\}}$ .

**Theorem 11.313.** If  $\langle X, \mathcal{T} \rangle$  is a Hausdorff topological space then a convergent sequence  $\{x_i\}_{i \in \{m, \dots, \infty\}}$  in  $X$  has a unique limit. We note this limit as  $\lim_{i \rightarrow \infty} x_i$

**Proof.** Suppose that  $x, y$  are two limits of  $\{x_i\}_{i \in \{m, \dots, n\}}$  with  $x \neq y$  then by the Hausdorff property there exists open sets  $U, V$  with  $x \in U \wedge y \in V \wedge U \cap V = \emptyset$ . By the definition of convergence there exists  $N_x, N_y \in \{m, \dots, \infty\}$  such that if  $n \geq N_x$  then  $x_n \in U$  and if  $k \geq N_y$  then  $x_n \in V$  so if  $n \geq \max(N_x, N_y)$  then  $x_n \in U \wedge x_n \in V \Rightarrow x_n \in U \cap V = \emptyset$  a contradiction, so we must have that  $x = y$ .  $\square$

The following theorem shows that convergence is not dependent on where a sequence start.

**Theorem 11.314.** If  $\langle X, \mathcal{T} \rangle$  is a topological space and  $\{x_i\}_{i \in \{m, \dots, \infty\}}$  a sequence in  $X$  then

1. If  $\{x_i\}_{i \in \{m, \dots, n\}}$  converges to  $x$  we have that  $\forall n \geq m$  that the subsequence  $\{x_i\}_{i \in \{n, \dots, \infty\}}$  converges to  $x$
2. If there exists a  $m \geq n$  such that  $\{x_i\}_{i \in \{n, \dots, \infty\}}$  converges to  $x$  then  $\{x_i\}_{i \in \{m, \dots, \infty\}}$  converges to  $x$

**Proof.**

1. Given a  $U$  open containing  $x$  there exists by the hypothese a  $N' \in \{m, \dots, \infty\}$  such that  $\forall i \geq N'$  we have  $x_i \in U$ . Take then  $N = \max(n, N')$  then  $N \in \{n, \dots, \infty\}$  and  $i \geq N$  such that  $x_i \in U$  proving that  $\{x_i\}_{i \in \{n, \dots, \infty\}}$  converges to  $x$ .
2. If  $\{x_i\}_{i \in \{n, \dots, \infty\}}$  converges to  $x$  where  $n \geq m$  then there exists a  $N \in \{n, \dots, \infty\} \Rightarrow N \in \{m, \dots, \infty\}$  such that if  $i \geq N$  we have  $x_i \in U$  proving that  $\{x_i\}_{i \in \{m, \dots, \infty\}}$  converges to  $x$ .  $\square$

For metric spaces we define a limit of a sequence as follows

**Definition 11.315.** If  $\langle X, d \rangle$  is a pseudo-metric space then a sequence  $\{x_i\}_{i \in \{m, \dots, \infty\}}$  in  $X$  is convergent to  $x \in X$  if  $\forall \varepsilon \in \mathbb{R}_+$  there exists a  $N \in \{m, \dots, \infty\}$  such that if  $n \geq N$  then  $d(x, x_n) < \varepsilon$ .  $x$  is called a limit of  $\{x_i\}_{i \in \{m, \dots, \infty\}}$ .

The following theorem proves that both definitions are the same.

**Theorem 11.316.** *If  $\langle X, d \rangle$  is a pseudo-metric space and  $\{x_i\}_{i \in \{m, \dots, \infty\}}$  is a sequence then this sequence is convergent to  $x$  in the topological space sense if and only if it is convergent in the pseudo-metric sense to  $x$ .*

**Proof.**

1.  $(\Rightarrow)$  Let  $x$  be the limit of  $\{x_i\}_{i \in \{m, \dots, \infty\}}$  in the topological space sense then if  $\varepsilon \in \mathbb{R}_+$  we have for the open set  $B_d(x, \varepsilon)$  that  $x \in B_d(x, \varepsilon)$  and there exists a  $N \in \{m, \dots, \infty\}$  such that if  $n \geq N$  we have  $x_n \in B_d(x, \varepsilon) \Rightarrow d(x, x_n) < \varepsilon$  proving that  $x$  is a limit of  $\{x_i\}_{i \in \{m, \dots, \infty\}}$  in the pseudo-metric sense.
2.  $(\Leftarrow)$  Let  $x$  be the limit of  $\{x_i\}_{i \in \{m, \dots, \infty\}}$  in the metric space then if  $U$  is a open set such that  $x \in U$  then by the definition of the metric topology there exists a  $\varepsilon \in \mathbb{R}_+$  such that  $x \in B_d(x, \varepsilon) \subseteq U$ . Using the convergence in the pseudo-metric sense there exists a  $N \in \{m, \dots, \infty\}$  such that if  $n \geq N$  then we have  $d(x, x_n) < \varepsilon \Rightarrow x_n \in B_d(x, \varepsilon) \subseteq U \Rightarrow x_n \in U$ . Proving that  $x$  is the limit of  $\{x_i\}_{i \in \{m, \dots, \infty\}}$  in the topological sense.  $\square$

Note that the limits of a sequence do not have to be unique, however the following theorem proves that in the case of a metric space the limit is unique.

**Theorem 11.317.** *If  $\langle X, d \rangle$  is a metric space then if the sequence  $\{x_i\}_{i \in \{m, \dots, \infty\}}$  is convergent it has a unique limit*

**Proof.** This is trivial as  $\langle X, d \rangle$  is a Hausdorff (see 11.228) and Hausdorff spaces has a unique limit (see 11.313).  $\square$

**Definition 11.318.** *If  $\langle X, d \rangle$  is a metric space and  $\{x_i\}_{i \in \{m, \dots, \infty\}}$  is convergent then we note the unique limit as  $\lim_{i \rightarrow \infty} x_i$ .*

**Definition 11.319.** *If  $\langle X, d \rangle$  is a pseudo-metric space then  $\{x_i\}_{i \in \{m, \dots, \infty\}}$  is a Cauchy sequence if  $\forall \varepsilon \in \mathbb{R}_+$  there exists a  $N \in \mathbb{N}_0$  such that  $\forall n, k \geq N$  we have  $d(x_n, x_k) < \varepsilon$*

**Theorem 11.320.** *If  $\langle X, d \rangle$  is a pseudo-metric space then  $\{x_i\}_{i \in \{m, \dots, \infty\}}$  is a Cauchy sequence iff  $\forall \varepsilon \in \mathbb{R}_+$  there exists a  $N \in \{m, \dots, \infty\}$  such that  $\forall n \in \mathbb{N}$  we have that  $d(x_N, x_{N+n}) < \varepsilon$*

**Proof.**

1.  $(\Rightarrow)$  If  $\{x_i\}_{i \in \mathbb{N}_0}$  is Cauchy then given  $\varepsilon \in \mathbb{R}_+$  there exist a  $N \in \{m, \dots, \infty\}$  such that for  $n, k \geq N$  we have  $d(x_n, x_k) < \varepsilon$  so if  $n \in \mathbb{N}$  then  $N \geq n$  and  $N+n \geq N$  and thus we have  $d(x_N, x_{N+n})$
2.  $(\Leftarrow)$  Let  $\varepsilon \in \mathbb{R}_+$  and find then a  $N \in \{m, \dots, \infty\}$  such that if  $n \in \mathbb{N}$  then  $d(x_N, x_{N+n}) < \frac{\varepsilon}{2}$ . If now  $n, k \geq N$  then  $n-N, k-N \in \mathbb{N}$  so that  $d(x_n, x_k) = d(x_{(n-N)+N}, x_{(k-N)+N}) \leq d(x_{(n-N)+N}, x_N) + d(x_N, x_{(k-N)+N}) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$   $\square$

**Theorem 11.321.** *If  $\langle X, d \rangle$  is a pseudo-metric space and  $\{x_i\}_{i \in \{m, \dots, \infty\}}$  is convergent then it is Cauchy.*

**Proof.** If  $\{x_i\}_{i \in \{m, \dots, \infty\}}$  is convergent to  $x$  then if  $\varepsilon \in \mathbb{R}_+$  there exists a  $N \in \{m, \dots, \infty\}$  such that if  $n \geq N$  that  $d(x, x_n) < \frac{\varepsilon}{2}$ . So if  $n, m \geq N$  then  $d(x_n, x_m) \leq d(x_n, x) + d(x, x_m) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$  proving that it is Cauchy.  $\square$

We have defined the convergence in a metric space, lets create equivalent definitions in a pseudo normed space.

**Definition 11.322.** *If  $\langle X, \|\cdot\| \rangle$  is a pseudo-normed space then a sequence  $\{x_i\}_{i \in \{m, \dots, \infty\}}$  of elements of  $X$  is convergent to  $x \in X$  if  $\forall \varepsilon \in \mathbb{R}_+$  we have that  $\exists N \in \{m, \dots, \infty\}$  so that  $\forall n \geq N$  we have  $\|x - x_n\| < \varepsilon$*

**Example 11.323.** If  $\langle X, \|\cdot\| \rangle$  is a pseudo normed space,  $x \in X$  the if we have  $\{x_n\}_{n \in \{m, \dots, \infty\}}$  defined by  $x_n = x$  then  $\lim_{n \rightarrow \infty} x_n = x$

**Proof.** Let  $\varepsilon > 0$  then if  $n \geq m$  we have  $\|x_n - x_m\| = \|x - x\| = 0 < \varepsilon$   $\square$

**Theorem 11.324.** Let  $\langle X, \|\cdot\| \rangle$  be a speudo normed space,  $F$  a closed space and  $\{x_i\}_{i \in \{m, \dots, \infty\}} \subseteq F$  with a limit  $x$  then  $x \in F$

**Proof.** Assume that  $x \in X \setminus F$ , a open set, then there exists a  $\varepsilon > 0$  such that  $x \in B_{\|\cdot\|}(x, \varepsilon) \subseteq X \setminus F$ . As  $x$  is the limit of  $\{x_i\}_{i \in \{m, \dots, \infty\}}$  there excists a  $M \in \{m, \dots, \infty\}$  such that  $\|x - x_M\| < \varepsilon$  so that  $x_M \in B_{\|\cdot\|}(x, \varepsilon) \subseteq X \setminus F$  contradicting  $x_M \in F$ . So we must have that  $x \in F$ .  $\square$

The next theorem proves that the definition of convergence in a normed space is equivalent to convergence in the associated metric space (and thus by 11.316 equivalent with convergence in a topological space).

**Theorem 11.325.** If  $\langle X, \|\cdot\| \rangle$  is a pseudo-normed space and  $\langle X, d \rangle$  the associated metric space (where  $d: X \times X \rightarrow \mathbb{R}$  is defined by  $d(x, y) = \|x - y\|$  (see 11.71)) then  $\{x_i\}_{i \in \{m, \dots, \infty\}}$  is convergent to  $x$  in the normed sense iff it is convergent in the metric sense to  $x$ .

**Proof.**

1.  $(\Rightarrow)$  Let  $\{x_i\}_{i \in \{m, \dots, \infty\}}$  converges to  $x$  in the normed sense, then  $\forall \varepsilon \in \mathbb{R}_+$  there exists a  $N \in \{m, \dots, \infty\}$  such that  $\forall n \geq N$  we have  $\|x - x_n\| < \varepsilon \Rightarrow d(x, x_n) < \varepsilon$  proving that  $\{x_i\}_{i \in \{m, \dots, \infty\}}$  converges to  $x$  in the metric sense.
2.  $(\Leftarrow)$  Let  $\{x_i\}_{i \in \{m, \dots, \infty\}}$  converges to  $x$  in the metric sense, then  $\forall \varepsilon \in \mathbb{R}_+$  there exists a  $N \in \{m, \dots, \infty\}$  such that  $\forall n \geq N$  we have  $d(x, x_n) < \varepsilon \Rightarrow \|x - x_n\| < \varepsilon$  proving that  $\{x_i\}_{i \in \{m, \dots, \infty\}}$  converges to  $x$  in the normed sense.  $\square$

**Theorem 11.326.** If  $\langle X, \|\cdot\| \rangle$  is a normed space then if  $\{x_i\}_{i \in \{m, \dots, \infty\}}$  is convergent it has a unique limit. This unique limit is noted (like in the metric case) as  $\lim_{i \rightarrow \infty} x_i$

**Proof.** This follows from the fact that a normed space is a metric space and in metric spaces limits are unique (see 11.317)  $\square$

**Definition 11.327.** If  $\langle X, \|\cdot\| \rangle$  is a pseudo-normed space then  $\{x_i\}_{i \in \{m, \dots, \infty\}}$  is a Cauchy sequence if  $\forall \varepsilon \in \mathbb{R}_+$  there exists a  $N \in \{m, \dots, \infty\}$  such that if  $n, k \geq N$  then  $\|x_n - x_k\| < \varepsilon$

Again it is easy proved that the definition of Cauchy in a normed space is equivalent with the definition in the associated metric space

**Theorem 11.328.** If  $\langle X, \|\cdot\| \rangle$  is a pseudo-normed space and  $\langle X, d \rangle$  the associated metric space (where  $d: X \times X \rightarrow \mathbb{R}$  is defined by  $d(x, y) = \|x - y\|$  (see 11.71)) then  $\{x_i\}_{i \in \{m, \dots, \infty\}}$  is Cauchy in the metric sense iff it is Cauchy in the metric sense

**Proof.** We have

$$\begin{aligned} \{x_i\}_{i \in \{m, \dots, \infty\}} \text{ is Cauchy (metric sense)} &\Leftrightarrow \forall \varepsilon \in \mathbb{R}_+ \models \exists N \in \{m, \dots, \infty\} \vdash \forall n, k \geq N \text{ we have } \|x_n - x_k\| < \varepsilon \\ &\stackrel{d(x, y) = \|x - y\|}{\Leftrightarrow} \forall \varepsilon \in \mathbb{R}_+ \models \exists N \in \{m, \dots, \infty\} \vdash \forall n, k \geq N \text{ we have } d(x_n, x_k) < \varepsilon \\ &\Leftrightarrow \{x_i\}_{i \in \{m, \dots, \infty\}} \text{ is Cauchy (normed sense)} \end{aligned} \quad \square$$

**Theorem 11.329.** If  $\langle X, \|\cdot\| \rangle$  is a pseudo-normed space then if  $\{x_i\}_{i \in \mathbb{N}_0}$  is convergent then it is Cauchy.

**Proof.** This follows from the fact that  $\langle X, \|\cdot\| \rangle$  forms a metric space and 11.321.  $\square$

**Example 11.330.** If  $a \in \mathbb{R}, b \in \mathbb{R}$  with  $0 \leq b < 1$  then  $\{a \cdot b^n\}_{n \in \mathbb{N}}$  converges to 0.

**Proof.** Two cases must be considered:

1. ( $b = 0$ ) then  $a \cdot b^n = 0$  so  $\{a \cdot b^n\}_{n \in \mathbb{N}}$  converges to 0
2. ( $0 < b$ ) Let  $\varepsilon \in \mathbb{R}_+$  then by 9.67 there exists an  $N \in \mathbb{N}$  such that if  $n \geq N$  then we have that  $0 < b^n < \frac{\varepsilon}{a} \Rightarrow 0 < a \cdot b^n < \varepsilon \Rightarrow |a \cdot b^n - 0| < \varepsilon$  proving that  $\{a \cdot b^n\}_{n \in \mathbb{N}}$  converges to 0.  $\square$

**Example 11.331.** Let  $a \in \mathbb{R}$  then  $\left\{ \frac{1}{a+i} \right\}_{i \in \mathbb{N}_0}$  converges to 0

**Proof.** Let  $\varepsilon \in \mathbb{R}_+$  then using 9.56 there exists a  $N \in \mathbb{N}_0$  such that  $\max\left(\frac{1}{\varepsilon} - a, -a\right) < N$  so if  $N \leq i$  then  $\frac{1}{\varepsilon} - a < i \Rightarrow \frac{1}{\varepsilon} < i + a$  and  $-a < i \Rightarrow 0 < i + a$  so that  $\frac{1}{a+i} < \varepsilon$  proving that  $\left| \frac{1}{a+i} - 0 \right| = \left| \frac{1}{a+i} \right| = \frac{1}{a+i} < \varepsilon$  which proves that  $\lim_{i \rightarrow \infty} \frac{1}{a+i} = 0$   $\square$

**Example 11.332.** Let  $a \in \mathbb{R}, b \in \mathbb{R}_+$  then  $\left\{ \frac{a+i}{b+i} \right\}_{i \in \mathbb{N}}$  converges to 1

**Proof.** First let  $i \in \mathbb{N}$  then  $\frac{a+i}{b+i} - 1 = \frac{a+i-(b+i)}{b+i} = \frac{a-b}{b+i}$ . Let  $\varepsilon \in \mathbb{R}_+$  then using 9.56 there exists a  $N \in \mathbb{N}_0$  such that  $\frac{|a-b|}{\varepsilon} - b < N$  then if  $N \leq i$  we have  $\frac{|a-b|}{\varepsilon} - b < i \Rightarrow \frac{|a-b|}{\varepsilon} < b + i \Rightarrow \frac{|a-b|}{b+i} < \varepsilon$  so that  $\left| \frac{a+i}{b+i} - 1 \right| = \left| \frac{a-b}{b+i} \right| = \frac{|a-b|}{b+i} < \varepsilon$  proving that  $\lim_{i \rightarrow \infty} \frac{a+i}{b+i} = 1$   $\square$

**Theorem 11.333.** Let  $\{x_i\}_{i \in \{k, \dots, \infty\}} \subseteq \mathbb{C}$  be a convergent sequence of complex numbers then  $\{\bar{x}_i\}_{i \in \{k, \dots, \infty\}}$  is convergent and  $\lim_{i \rightarrow \infty} \bar{x}_i = \overline{\lim_{i \rightarrow \infty} x_i}$

**Proof.** Let  $\varepsilon > 0$  then there exists a  $N \in \{k, \dots, \infty\}$  such that  $\forall i \geq N$  we have  $|x_i - x| < \varepsilon$ . Now  $|\bar{x}_i - \bar{x}| = |\overline{x_i - x}| = |x_i - x| < \varepsilon$   $\square$

**Theorem 11.334.** Let  $\langle X, d_X \rangle, \langle Y, d_Y \rangle$  be metric spaces,  $f: X \rightarrow Y$  a function and  $x \in X$  then we have the following equivalence

$f$  is continuous at  $x \Leftrightarrow$  for every convergent sequence  $\{x_n\}_{n \in \mathbb{N}_0}$  with limit  $x$  we have that  $\{f(x_n)\}_{n \in \mathbb{N}_0}$  converges with limit  $f(x)$

**Proof.**

$\Rightarrow$ . Let  $\varepsilon > 0$  then by continuity of  $f$  there exists a  $\delta(x) > 0$  such that if  $d_X(x, y) < \delta(x)$  then  $d_Y(f(x), f(y)) < \varepsilon$ . As  $\lim_{n \rightarrow \infty} x_n = x$  there exists a  $N \in \{k, \dots, \infty\}$  such that if  $n \geq N$  then  $d_X(x, x_n) < \delta(x) \Rightarrow d_Y(f(x), f(x_n)) < \varepsilon$  proving that  $\lim_{n \rightarrow \infty} f(x_n) = f(x)$

$\Leftarrow$ . Assume that  $f$  is not continuous at  $x$  then there exists a  $\varepsilon_x > 0$  such that  $\forall \delta > 0$  there exists a  $y \in X$  with  $d_X(x, y) < \delta$  and  $d_Y(f(x), f(y)) > \varepsilon_x$ . So  $\forall n \in \mathbb{N}_0$  we have  $\exists x_n \in X$  with  $d_X(x, x_n) < \frac{1}{N}$  and  $d_Y(f(x), f(x_n)) > \varepsilon_x$ . Take now  $\varepsilon > 0$  then by the Archimedean property of the real numbers (see 9.56) there exists a  $N_\varepsilon \in \mathbb{N}_0$  such that  $0 < \frac{1}{N_\varepsilon} < \varepsilon$ , if now  $n \geq N_\varepsilon$  then we have  $d_X(x, x_n) < \frac{1}{n} \leq \frac{1}{N_\varepsilon} < \varepsilon$ , which proves that  $\lim_{n \rightarrow \infty} x_n = x$ . As we have however  $\forall n \in \mathbb{N}_0$  that  $d_Y(f(x), f(x_n)) \geq \varepsilon_x$  we do not have that  $\lim_{n \rightarrow \infty} f(x_n) = f(x)$  contradicting our hypothesis, so  $f$  must be continuous at  $x$ .  $\square$

**Lemma 11.335.** Let  $\langle X, \|\cdot\| \rangle$  be a normed space and  $\{x_i\}_{i \in \{k, \dots, \infty\}}$  be a Cauchy sequence then there exists a  $K \in \mathbb{R}_0^+$  so that  $\forall i \in \{k, \dots, \infty\}$  we have  $\|x_i\| < K$

**Proof.** Take  $\varepsilon = 1$  then because of Cauchy there exists a  $N \in \{k, \dots, \infty\}$  such that  $\forall n, m \geq N$  we have  $\|x_n - x_m\| < 1$  and thus if  $n \geq N$  we have  $\|x_n\| \leq \|x_n - x_N\| + \|x_N\| < 1 + \|x_N\|$  so if we take  $K = \max(\|x_1\|, \dots, \|x_{N-1}\|, \|x_N\| + 1)$  then we have  $\forall i \in \{k, \dots, \infty\}$  that  $\|x_i\| < K$ .  $\square$

**Corollary 11.336.** Let  $\langle X, \|\cdot\| \rangle$  be a normed space and  $\{x_i\}_{i \in \{k, \dots, \infty\}}$  be a convergent sequence then there exists a  $K \in \mathbb{R}_0^+$  so that  $\forall i \in \{k, \dots, \infty\}$  we have  $\|x_i\| < K$ .

**Proof.** If  $\{x_n\}_{n \in \{k, \dots, \infty\}}$  converges then by 11.329  $\{x_n\}_{n \in \{k, \dots, \infty\}}$  is Cauchy. Using the previous lemma [see 11.335] we have then that there exists a  $K \in \mathbb{R}_0^+$  so that  $\forall i \in \{k, \dots, \infty\}$  we have  $\|x_i\| < K$ .  $\square$

## 11.11.2 Complete spaces

**Definition 11.337.** A pseudo metric space  $\langle X, d \rangle$  is complete if every Cauchy sequence is convergent. A normed space  $\langle X, \|\cdot\| \rangle$  is complete if every Cauchy sequence is convergent (this is the same as being Cauchy in the metric sense). A complete normed space is called a **Banach** space.

**Proposition 11.338.** Let  $\{\langle X_i, \|\cdot\|_i \rangle\}_{i \in \{1, \dots, n\}}$  be a finite family of Banach spaces then  $\langle \prod_{i \in \{1, \dots, n\}} X_i, \|\cdot\| \rangle$  [where  $\|\cdot\|$  is the maximum norm [see 11.82]] is a Banach space

**Proof.** Using 11.82 we have that  $\langle \prod_{i \in \{1, \dots, n\}} X_i, \|\cdot\| \rangle$  is a normed space. Let  $\{x_i\}_{i \in \mathbb{N}_0}$  be a Cauchy sequence. Take  $\varepsilon > 0$  then there exists a  $N \in \mathbb{N}_0$  such that  $\forall k, l \geq N$  we have  $\|x_k - x_l\| < \varepsilon$ . Hence  $\forall i \in \{1, \dots, n\}$  we have that  $\|(x_k)_i - (x_l)_i\| \leq \max(\{\|(x_k)_i - (x_l)_i\| : i \in \{1, \dots, n\}\}) = \|x_k - x_l\| < \varepsilon$  proving that  $\{(x_j)_i\}_{j \in \mathbb{N}_0}$  is a Cauchy sequence. Using the completeness of  $X_i \lim_{j \rightarrow \infty} (x_j)_i = y_i$  exists hence give  $\varepsilon$  there exists a  $N_i$  such that if  $k \geq N_i$  then  $\|(x_k)_i - y_i\| < \varepsilon$ . Take  $y = (y_1, \dots, y_n) \in \prod_{i \in \{1, \dots, n\}} X_i$  and  $N = \max(\{N_i : i \in \{1, \dots, n\}\})$  then if  $k \geq N$  we have  $\|x_k - y\| = \max(\{\|(x_k)_i - y_i\| : i \in \{1, \dots, n\}\}) < \varepsilon$ . Proving that  $\{x_i\}_{i \in \mathbb{N}_0}$  converges and that  $\langle \prod_{i \in \{1, \dots, n\}} X_i, \|\cdot\| \rangle$  is a Banach space.  $\square$

**Definition 11.339.** Let  $X$  be a topological space,  $\langle Y, \|\cdot\| \rangle$  a normed space then a function  $f: X \rightarrow Y$  is **bounded** iff  $f(X)$  is bounded (see 11.65). In other words if  $\exists M \in \mathbb{R}$  such that  $\forall x, y \in X$  we have  $\|f(x) - f(y)\| \leq M$ . Let  $\mathcal{B}(X, Y) \subseteq Y^X$  be defined by  $\mathcal{B}(X, Y) = \{f \in Y^X \mid f \text{ is bounded}\}$  the set of bounded functions from  $X$  to  $Y$ .

**Theorem 11.340.** Let  $X$  be a non empty topological space,  $\langle Y, \|\cdot\| \rangle$  a normed space then a function  $f: X \rightarrow Y$  is **bounded** iff  $\exists M \in \mathbb{R}$  such that  $\forall x \in X$  we have  $\|f(x)\| \leq M$

**Proof.**

$(\Rightarrow)$  As  $f$  is bounded there exists a  $M \in \mathbb{R}$  such that  $\forall x, y \in X$  we have  $\|f(x) - f(y)\| \leq M$ . As  $X \neq \emptyset$  there exists a  $x_0 \in X$  let now  $x \in X$  then  $\|f(x)\| = \|f(x) - f(x_0) + f(x_0)\| \leq \|f(x_0)\| + \|f(x) - f(x_0)\| \leq \|f(x_0)\| + M$

$(\Leftarrow)$  If  $\forall x \in X$  we have  $\|f(x)\| \leq M$  then if  $x, y \in X$  we have  $\|f(x) - f(y)\| \leq \|f(x)\| + \|f(y)\| \leq 2M$   $\square$

**Theorem 11.341.** Let  $X$  be a set,  $\langle Y, \|\cdot\| \rangle$  a normed space over a field  $\mathbb{K}$  then  $\mathcal{B}(X, Y)$  is a vector space (using the canonical sum and scalar product of functions defined in 10.115)

**Proof.** As  $Y$  is a vector space we have by 10.115 that  $Y^X$  forms a vector space. Now if  $\alpha, \beta \in \mathbb{K}$  and  $f, g \in \mathcal{B}(X, Y)$  then there exists  $M_f, M_g \in \mathbb{R}$  such that  $\forall x, y \in X$  we have  $\|f(x) - f(y)\| \leq M_f$ ,  $\|g(x) - g(y)\| \leq M_g$ . We have then that  $\|\alpha \cdot f(x) + \beta \cdot g(x) - \alpha \cdot f(y) - \beta \cdot g(y)\| \leq \|\alpha \cdot (f(x) - f(y)) + \beta \cdot (g(x) - g(y))\| \leq |\alpha| \cdot \|f(x) - f(y)\| + |\beta| \cdot \|g(x) - g(y)\| \leq |\alpha| \cdot M_f + |\beta| \cdot M_g$  proving that  $\alpha \cdot f + \beta \cdot g \in \mathcal{B}(X, Y)$ . Using 10.111 we have then that  $\mathcal{B}(X, Y)$  forms a vector space.  $\square$

**Definition 11.342.** Let  $X$  be a non empty set,  $\langle Y, \|\cdot\| \rangle$  a normed space over a field  $\mathbb{K}$   $f, g \in \mathcal{B}(X, Y)$  then  $\|f\|_s$  is defined by  $\sup(\{\|f(x)\| : x \in X\})$

**Theorem 11.343.** Let  $X$  be a non empty topological space,  $\langle Y, \|\cdot\| \rangle$  a normed space over a field  $\mathbb{K}$  then  $\|\cdot\|_s: \mathcal{B}(X, Y) \rightarrow \mathbb{R}_+$  is well defined and forms a norm. So  $\langle \mathcal{B}(X, Y), \|\cdot\|_s \rangle$  is a normed space.

**Proof.** As  $X \neq \emptyset$  there exists a  $x \in X$  then if  $y \in X$  we have that  $\|f(y)\| \leq \|f(x) - f(y)\| + \|f(x)\| \leq M_f + \|f(x)\|$  proving that  $\forall t \in \{\|f(y)\| \mid y \in X\}$  we have that  $t \leq M_f + \|f(x)\|$  so that  $M_f + \|f(x)\|$  forms a upper bound of  $\{\|f(y)\| \mid y \in X\}$  and thus by 9.43 we have that  $\sup(\{\|f(y)\| \mid y \in X\})$  exists. Now to proof that  $\|\cdot\|_s$  is a norm

1.  $\forall f \in \mathcal{B}(X, Y) \vdash \|f\|_s \geq 0$  is trivial by the definition.
2. If  $f \in \mathcal{B}(X, Y)$  and  $\alpha \in \mathbb{K}$  then  $\{\|\alpha f(x)\| \mid x \in X\} = \{|\alpha| \cdot \|f(x)\| \mid x \in X\} = |\alpha| \cdot \{\|f(x)\| \mid x \in X\} \Rightarrow \|f\|_s = \sup(|\alpha| \cdot \{\|f(x)\| \mid x \in X\}) \stackrel{9.45}{=} |\alpha| \cdot \sup(\{\|f(x)\| \mid x \in X\}) = |\alpha| \cdot \|f\|_s$
3. If  $f, g \in \mathcal{B}(X, Y)$  then

$$\begin{aligned}
 \|f + g\|_s &= \sup(\{\|(f + g)(x)\| \mid x \in X\}) \\
 &= \sup(\{\|f(x) + g(x)\| \mid x \in X\}) \\
 &\leq \sup(\{\|f(x)\| + \|g(x)\| \mid x \in X\}) \text{ (see 2.176)} \\
 &\leq \sup(\{\|f(x)\| \mid x \in X\} + \{\|g(x)\| \mid x \in X\}) \text{ (see 2.175)} \\
 &\stackrel{9.46}{=} \sup(\{\|f(x)\| \mid x \in X\}) + \sup(\{\|g(x)\| \mid x \in X\}) \\
 &= \|f\|_s + \|g\|_s
 \end{aligned}$$

□

**Theorem 11.344.** Let  $X$  be a non empty set,  $\langle Y, \|\cdot\| \rangle$  a Banach space over a field  $\mathbb{K}$  then  $\langle \mathcal{B}(X, Y), \|\cdot\|_s \rangle$  is a Banach space.

**Proof.** Let  $\{f_n\}_{n \in \{k, \dots, \infty\}}$  be a Cauchy sequence in  $\mathcal{B}(X, Y)$  then given  $\varepsilon > 0$  there exists a  $N \in \mathbb{N}$  such that  $\forall n, m$  we have  $\sup(\{\|f_n(x) - f_m(x)\| \mid x \in X\}) = \|f_n - f_m\|_s < \varepsilon$  so that  $\forall x \in X$  we have  $\|f_n(x) - f_m(x)\| \leq \sup(\{\|f_n(x) - f_m(x)\| \mid x \in X\}) < \varepsilon$  or  $\forall x \in \{f_n(x)\}_{n \in \{k, \dots, \infty\}}$  is a Cauchy sequence so as  $Y$  is a Banach space there exists a  $f(x) \in Y$  such that  $\lim_{n \rightarrow \infty} f_n(x) = f(x)$ .

So we can define a function  $f: X \rightarrow Y$  defined by  $x \mapsto f(x)$ . We prove now that  $f$  is bounded. So as  $1 > 0$  there exists a  $N(x) \in \mathbb{N}$  such that if  $n \geq N(x)$  then  $\|f_n(x) - f(x)\| < 1 \Rightarrow \|f(x)\| \leq \|f_n(x)\| + \|f_n(x) - f(x)\| < \|f_n(x)\| + 1$  of  $\forall n \geq N(x)$  we have  $\|f(x)\| < 1 + \|f_n(x)\|$ . Now as  $\{f_n\}_{n \in \{k, \dots, \infty\}}$  is Cauchy we have that  $\exists N \in \mathbb{N}$  such that if  $n, m \geq N$  then  $\|f_n - f_m\|_s < 1$  so if  $m \geq N$  then  $\|f_m - f_N\|_s < 1 \Rightarrow \|f_m\|_s \leq \|f_N\|_s + \|f_m - f_N\|_s < \|f_N\|_s + 1$  or if  $m \geq N$  and  $x \in X$  then  $\|f_m(x)\| \leq \sup(\{\|f_m(x)\| \mid x \in X\}) = \|f_m\|_s < \|f_N\|_s + 1$ . Now given  $x \in X$  if  $l \geq \max(N(x), N)$  then  $\|f(x)\| < 1 + \|f_l(x)\| < 1 + (\|f_N\|_s + 1) = 2 + \|f_N\|_s$  proving by 11.340 that  $f$  is a bounded function and thus that  $f \in \mathcal{B}(X, Y)$ .

Next we prove that  $\lim_{n \rightarrow \infty} f_n = f$ , so let  $\varepsilon > 0$  then if  $x \in X$  there exists a  $N(x)$  such that if  $n \geq N(x)$  then  $\|f_n(x) - f(x)\| < \frac{\varepsilon}{2}$ . As  $\{f_n\}_{n \in \{k, \dots, \infty\}}$  there exist a  $N$  such that  $\|f_n - f_N\| < \frac{\varepsilon}{2}$  if  $n \geq N$  so that  $\forall x \in X$  we have  $\|f_n(x) - f_N(x)\| \leq \sup(\{\|f_n(x) - f_N(x)\| \mid x \in X\}) = \|f_n - f_N\| < \frac{\varepsilon}{2}$ . Take then  $x \in X$  and  $m \geq N$  and  $n \geq \max(N(x), N)$  then we have that  $\|f_m(x) - f(x)\| \leq \|f_m(x) - f_N(x)\| + \|f_N(x) - f(x)\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$  proving that  $\forall x \in X$  we have if  $m \geq N$  that  $\|f_m(x) - f(x)\| < \varepsilon$  proving that  $\varepsilon$  is a upper bound of  $\{\|f_m(x) - f(x)\| \mid x \in X\}$  or  $\|f_m - f\|_s = \sup(\{\|f_m(x) - f(x)\| \mid x \in X\}) < \varepsilon$  proving that  $\lim_{n \rightarrow \infty} f_n = f$  and thus that  $\mathcal{B}(X, Y)$  is Banach. □

**Definition 11.345.** Let  $X$  be a set,  $\langle Y, \|\cdot\| \rangle$  a metric space be topological spaces and  $\{f_n\}_{n \in \{k, \dots, \infty\}}$  a sequence of function from  $X \rightarrow Y$  then  $\{f_n\}_{n \in \{k, \dots, \infty\}}$  **converges uniformly** to a function  $f: X \rightarrow Y$  if  $\forall \varepsilon \in \mathbb{R} \vdash \varepsilon > 0$  there exists a  $N \in \mathbb{N}$  such that  $\forall n \geq N$  and  $\forall x \in X$  we have  $\|f_n(x) - f(x)\| < \varepsilon$

As a example of uniform convergence we use the set of bounded functions

**Theorem 11.346.** Let  $X$  be a non empty set,  $\langle Y, \|\cdot\| \rangle$  a normed space then if  $\{f_n\}_{n \in \{k, \dots, \infty\}}$  is a sequence of bounded functions then convergence in  $\mathcal{B}(X, Y)$  is equivalent with uniform convergence.

**Proof.**

1.  $(\Rightarrow)$  Given  $\varepsilon > 0$  there exists a  $N \in \mathbb{N}$  such that if  $n \geq N$  then  $\sup(\{\|f_n(x) - f(x)\| \mid x \in X\}) = \|f_n - f\|_s < \varepsilon$  sup is a upperbound  $\Rightarrow \forall x \in X$  we have  $\|f_n(x) - f(x)\| < \varepsilon$  proving uniform convergence.
2.  $(\Leftarrow)$  Given  $\varepsilon > 0$  there exists a  $N \in \mathbb{N}$  such that if  $n \geq N$  we have  $\forall x \in X$  that  $\|f(x) - f_n(x)\| < \varepsilon$  proving that  $\varepsilon$  is a upperbound of  $\{\|f_n(x) - f(x)\| \mid x \in X\}$  so that  $\|f_n - f\|_s = \sup(\{\|f_n(x) - f(x)\| \mid x \in X\}) < \varepsilon$  proving that  $\lim_{n \rightarrow \infty} f_n = f$ .  $\square$

**Theorem 11.347.** Let  $\langle X, \|\cdot\|_X \rangle$  be a normed space,  $\langle Y, \|\cdot\|_Y \rangle$  a normed space then if  $\{f_n\}_{n \in \{k, \dots, \infty\}}$  is a sequence of continuous functions from  $X$  to  $Y$  uniformly converging to a function  $f: X \rightarrow Y$  then  $f$  is continuous.

**Proof.** Take  $x \in X$  and  $\varepsilon > 0$ , from uniform convergence there exists a  $N \in \{k, \dots, \infty\}$  such that  $\forall n \geq N$  we have  $\forall x \in X$  that  $\|f_n(x) - f(x)\|_Y < \frac{\varepsilon}{3}$ . Now as  $\forall i \in \{k, \dots, \infty\}$  we have that  $f_N$  is continuous  $\delta_N > 0$  such that if  $\|x - y\|_X < \delta_N$  then  $\|f_N(x) - f_N(y)\|_Y < \frac{\varepsilon}{3}$ . Let  $\|x - y\|_X < \delta_N$  then  $\|f(x) - f(y)\|_Y \leq \|f(x) - f_N(x)\| + \|f_N(x) - f_N(y)\| + \|f_N(y) - f(y)\|_Y \leq \|f(x) - f_N(x)\|_Y + \|f_N(x) - f_N(y)\|_Y + \|f_N(y) - f(y)\|_Y < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon$  proving that  $f$  is continuous.  $\square$

**Theorem 11.348.** Let  $\langle X, d \rangle$  is complete and  $A \subseteq X$  is closed then  $\langle A, d_A \rangle$  is complete.

**Proof.** Let  $\{x_n\}_{n \in \{k, \dots, \infty\}}$  be a Cauchy sequence in  $A$  then trivially  $\{x_n\}_{n \in \{k, \dots, \infty\}}$  is a Cauchy sequence in  $X$  so there exists a  $x \in X$  such that  $\lim_{n \rightarrow \infty} x_n = x$ . So if  $U$  is a open set with  $x \in U$  then there exists a  $\varepsilon > 0$  such that  $x \in B_d(x, \varepsilon) \subseteq U$  and as  $\lim_{n \rightarrow \infty} x_n = x$  there exists a  $N \in \{k, \dots, \infty\}$  so that  $d(x, x_n) < \varepsilon$  is  $n \geq N$  proving that  $x_N \in B_d(x, \varepsilon) \subseteq U \Rightarrow U \cap A \neq \emptyset$  proving that  $x \in A$  (see 11.21) and thus that  $\{x_n\}_{n \in \{k, \dots, \infty\}}$  is convergent in  $S$ .  $\square$

**Theorem 11.349.** Let  $\langle X, \|\cdot\| \rangle$  be a normed space and let  $\{x_i\}_{i \in \{k, \dots, \infty\}}$  be a convergent sequence with  $\lim_{n \rightarrow \infty} x_n$  then if there exists a  $m \in \{k, \dots, \infty\}$  such that  $\forall i \in \{m, \dots, \infty\}$  we have  $\|x_i\| \leq A \in \mathbb{R}$  (or  $A \leq \|x_i\|$ ) then we have that  $\|x\| \leq A$  (or  $A \leq \|x\|$ )

**Proof.** Suppos that there exists a  $m \in \{k, \dots, \infty\}$  such that

1.  $(\forall i \in \{m, \dots, \infty\} \text{ we have } \|x_i\| \leq A)$  We proceed by contradiction so assume that  $A < \|x\|$  then there exists a  $\varepsilon = \|x\| - A$  with  $\varepsilon < 0$  so there exists a  $N \in \{k, \dots, \infty\}$  such that  $\|x - x_{\max(N, m)}\| < \varepsilon$  and thus  $\|x\| \leq \|x - x_{\max(N, m)}\| + \|x_{\max(N, m)}\| < \varepsilon + \|x_{\max(N, m)}\| \leq \varepsilon + A = \|x\| - A + A = \|x\|$  giving the contradiction  $\|x\| < \|x\|$ .
2.  $(\forall i \in \{m, \dots, \infty\} \text{ we have } A \leq \|x_i\|)$  We proceed by contradiction so assume that  $\|x\| < A$  then  $\varepsilon = A - \|x\| > 0$  and there exists a  $N \in \{k, \dots, \infty\}$  such that  $\|x_{\max(N, m)} - x\| < \varepsilon$  then  $A \leq \|x_{\max(N, m)}\| \leq \|x - x_{\max(N, m)}\| + \|x\| < \varepsilon + \|x\| = A - \|x\| + \|x\| = A$  leading to the contradiction  $A < A$  so we must have  $A \leq \|x\|$ .  $\square$

**Theorem 11.350.** Let  $\langle X, \|\cdot\| \rangle$  be a normed space on  $\mathbb{K}$ ,  $\{x_i\}_{i \in \{k, \dots, n\}}$  a convergent sequence in  $X$ ,  $\{s_i\}_{i \in \{k, \dots, \infty\}}$  a convergent sequence in  $\mathbb{R}$  with  $\forall i \in \{k, \dots, \infty\} \|x_i\| < s_i$  (so  $0 \leq s_N$ ) then  $\|\lim_{i \rightarrow \infty} x_i\| < \lim_{i \rightarrow \infty} s_i$

**Proof.** Let  $x = \lim_{i \rightarrow \infty} x_i$ ,  $s = \lim_{i \rightarrow \infty} s_i$  (using the previous theorem we must then have that  $0 \leq s$  as  $0 \leq \|x_i\| \leq s$ ) and assume that  $s < \|x\|$  so that  $\varepsilon = \|x\| - s > 0$  and find  $N_1, N_2 \in \{k, \dots, \infty\}$  such that  $\|x_n - x\| < \frac{\varepsilon}{2}$  if  $n \geq N_1$  and  $|s_m - s| < \frac{\varepsilon}{2}$  then if  $N = \max(N_1, N_2)$  we have  $\|x\| \leq \|x - x_N\| + \|x_N\| < \frac{\varepsilon}{2} + \|x_N\| \leq \frac{\varepsilon}{2} + s_N \leq \frac{\varepsilon}{2} + |s_N - s| + |s| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} + |s| = \varepsilon + |s| \underset{0 \leq s \Rightarrow s = |s|}{=} \varepsilon + s = \|x\| - s + s = \|x\|$  leading to the contradiction  $\|x\| < \|x\| \Rightarrow \|x\| \leq s$ .  $\square$

**Theorem 11.351.** Let  $\langle X, \|\cdot\| \rangle$  be a normed space and let  $\{x_n\}_{n \in \{k, \dots, \infty\}}, \{y_n\}_{n \in \{k, \dots, \infty\}}$  be convergent sequences with  $\forall n \in \{k, \dots, \infty\}$  we have  $\|x_n\| \leq \|y_n\|$  then  $\|\lim_{n \rightarrow \infty} x_n\| \leq \|\lim_{n \rightarrow \infty} y_n\|$

**Proof.** Let  $x = \lim_{n \rightarrow \infty} x_n$ ,  $y = \lim_{n \rightarrow \infty} y_n$  and assume that  $\|y\| < \|x\|$  so that  $0 < \|x\| - \|y\| = \varepsilon$  and find  $N_1, N_2 \in \{k, \dots, \infty\}$  such that  $\|x_N - x\| < \frac{\varepsilon}{2}$ ,  $\|y_N - y\| < \frac{\varepsilon}{2}$  if  $n \geq N_1, m \geq N_2$  then if  $N = \max(N_1, N_2)$  we have  $\|x\| \leq \|x_N - x\| + \|x_N\| \leq \|x_N - x\| + \|y_N\| \leq \|x_N - x\| + \|y_N - y + y\| \leq \|x_N - x\| + \|y_N - y\| + \|y\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} + \|y\| = \varepsilon + \|y\| = \|x\| - \|y\| + \|y\| = \|x\| \Rightarrow \|x\| < \|x\|$  a contradiction so we must have that  $\|x\| \leq \|y\|$   $\square$

**Theorem 11.352.** Let  $\langle X, \|\cdot\| \rangle$  be a normed space,  $a \in X$  and  $\{x_n\}_{n \in \{k, \dots, \infty\}}$  a sequence in  $X$  then  $\{x_n\}_{n \in \mathbb{N}_0}$  convergence if and only if  $\{x_n + a\}_{n \in \mathbb{N}_0}$  converges. Further if we have convergence then  $\lim_{n \rightarrow \infty} (x_n + a) = \left( \lim_{n \rightarrow \infty} x_n \right) + a$

**Proof.**

1. Assume that  $\{x_n\}_{n \in \{k, \dots, \infty\}}$  is convergent with  $\lim_{n \rightarrow \infty} x_n = x$  then if  $\varepsilon > 0$  there exists a  $N_\varepsilon \in \{k, \dots, \infty\}$  such that  $\forall n \geq N_\varepsilon$  we have  $\|x_n - x\| < \varepsilon$ . Take then  $\varepsilon > 0$  then for  $n \geq N_\varepsilon$  we have  $\|(x_n + a) - (x + a)\| = \|x_n - x\| < \varepsilon$  which proves that  $\{x_n + a\}_{n \in \mathbb{N}_0}$  converge with  $\lim_{n \rightarrow \infty} (x_n + a) = x + a$
2. Assume that  $\{x_n + a\}_{n \in \{k, \dots, \infty\}}$  is convergent with  $\lim_{n \rightarrow \infty} (x_n + a) = y$  then by (1)  $\{(x_n + a) + (-a)\}_{n \in \mathbb{N}_0} = \{x_n\}_{n \in \mathbb{N}_0}$  converges with  $\lim_{n \rightarrow \infty} x_n = \lim_{n \rightarrow \infty} ((x_n + a) - a) \stackrel{(1)}{=} \lim_{n \rightarrow \infty} (x_n + a) - a = y - a$  so that  $\lim_{n \rightarrow \infty} (x_n + a) = \left( \lim_{n \rightarrow \infty} x_n \right) + a$ .  $\square$

**Theorem 11.353.** Let  $\langle X, \|\cdot\| \rangle$  be a normed space and given two sequences  $\{x_n\}_{n \in \{k, \dots, \infty\}}$ ,  $\{y_n\}_{n \in \{k, \dots, \infty\}}$  be convergent sequences with  $\lim_{n \rightarrow \infty} x_n = x$  and  $\lim_{n \rightarrow \infty} y_n = y$  then  $\{x_n + y_n\}_{n \in \{k, \dots, \infty\}}$  is convergent with  $\lim_{n \rightarrow \infty} (x_n + y_n) = x + y$ . Second  $\forall \alpha \in \mathbb{K}$  ( $\mathbb{K} = \mathbb{C}$  or  $\mathbb{R}$  the field of scalars in  $X$ ) we have  $\lim_{n \rightarrow \infty} (\alpha \cdot x_n) = \alpha \cdot x$ . Third if  $z \in X$  then  $\lim_{n \rightarrow \infty} (z + x_n) = z + \lim_{n \rightarrow \infty} x_n$ .

**Proof.**

1. Given  $\varepsilon > 0$  find a  $N_1, N_2 \in \{k, \dots, \infty\}$  such that  $n \in \{k, \dots, \infty\}$  with  $n \geq N_1, N_2$  we have  $\|x_n - x\| < \frac{\varepsilon}{2}$ ,  $\|y_n - y\| < \frac{\varepsilon}{2}$  so that if  $n \in \{k, \dots, \infty\}$  with  $k \geq N = \max(N_1, N_2)$  we have  $\|(x_n + y_n) - (x + y)\| \leq \|x_n - x\| + \|y_n - y\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$
2. We have two cases to consider
  - a. ( $\alpha = 0$ ) then  $\alpha \cdot x_n = 0$  and it is trivial to prove that  $\lim_{n \rightarrow \infty} (\alpha \cdot x_n) = 0 = \alpha \cdot \lim_{n \rightarrow \infty} x_n$
  - b. ( $\alpha \neq 0$ ) then  $|\alpha| \neq 0$  and given  $\varepsilon > 0$  find a  $N \in \{k, \dots, \infty\}$  so that  $\forall n \in \{k, \dots, \infty\}$  with  $n \geq N$  we have  $\|x_n - x\| < \frac{\varepsilon}{|\alpha|}$  so that  $\|\alpha \cdot x_n - \alpha \cdot x\| = |\alpha| \cdot \|x_n - x\| < \frac{\varepsilon}{|\alpha|} \cdot |\alpha| = \varepsilon$
3. Given  $\varepsilon > 0$  then there exists a  $N \in \{k, \dots, \infty\}$  such that if  $n \geq N$ ,  $n \in \{k, \dots, \infty\}$  that  $\|x_n - x\| < \varepsilon$  then  $\|(z + x_n) - (z + x)\| = \|x_n - x\| < \varepsilon$ .  $\square$

**Theorem 11.354.** Let  $\{x_i\}_{i \in \{k, \dots, \infty\}} \subseteq [0, \infty[$  be a convergent sequence of positive numbers then  $0 \leq \lim_{i \rightarrow \infty} x_i$

**Proof.** Assume that  $x = \lim_{i \rightarrow \infty} x_i < 0$  then  $0 < -x = \varepsilon$  then there exists a  $N \in \{k, \dots, \infty\}$  such that  $|x_N - x| < \frac{\varepsilon}{2}$ . As  $0 \leq x_N, x < 0 \Rightarrow 0 < -x$  we have that  $0 < x_N - x$  so that  $x_N - x = |x_N - x| < \frac{\varepsilon}{2} \Rightarrow x_N < \frac{\varepsilon}{2} + x = \frac{\varepsilon}{2} - \varepsilon = -\frac{\varepsilon}{2} < 0$  contradicting  $0 \leq x_N$  so we must have that  $0 \leq \lim_{i \rightarrow \infty} x_i$   $\square$

**Corollary 11.355.** Let  $\{x_i\}_{i \in \{k, \dots, \infty\}}, \{y_i\}_{i \in \{k, \dots, \infty\}} \subseteq \mathbb{R}$  be convergent sequences with  $\forall i \in \{k, \dots, \infty\} x_i \leq y_i$  then  $\lim_{i \rightarrow \infty} x_i \leq \lim_{i \rightarrow \infty} y_i$

**Proof.** Using the assumption we have that  $\{y_i - x_i\}_{i \in \{k, \dots, \infty\}} \subseteq [0, \infty[$  hence using the previous theorem we have that  $\lim_{i \rightarrow \infty} y_i - \lim_{i \rightarrow \infty} x_i = \lim_{i \rightarrow \infty} (y_i - x_i) \geq 0$  hence  $\lim_{i \rightarrow \infty} y_i \geq \lim_{i \rightarrow \infty} x_i$ .  $\square$

**Theorem 11.356.** Let  $X$  be a set with two equivalent compatible norms  $\|\cdot\|_1, \|\cdot\|_2$  then

1.  $\{x_i\}_{i \in \{k, \dots, \infty\}}$  is convergent to  $x \in X$  in  $\|\cdot\|_1$  if and only if it is convergent to  $x$  in  $\|\cdot\|_2$
2.  $\{x_i\}_{i \in \{k, \dots, \infty\}}$  is Cauchy in  $\|\cdot\|_1$  if and only if it is Cauchy in  $\|\cdot\|_2$

**Proof.** First note that we only have to prove  $\Rightarrow$  (as  $\Leftarrow$  follows from applying the theorem in the opposite direction. Next as  $\|\cdot\|_1$  and  $\|\cdot\|_2$  are equivalent there exists  $M_1 > 0, M_2 > 0$  such that  $\forall x \in X$  we have  $M_1 \cdot \|x\|_1 \leq \|x\|_2 \leq M_2 \cdot \|x\|_1$ . We have then

1. **(Convergence)** Let  $\varepsilon > 0$  then by convergence in  $\|\cdot\|_1$  there exists a  $N \in \{k, \dots, \infty\}$  such that  $\forall n \geq N$  we have  $\|x - x_n\| < \frac{\varepsilon}{M_2}$  so that  $\|x - x_n\|_2 \leq M_2 \cdot \|x - x_n\|_1 < M_2 \cdot \frac{\varepsilon}{M_2} = \varepsilon$
2. **(Cauchy)** Let  $\varepsilon > 0$  then by the Cauchy property in  $\|\cdot\|_1$  there exists a  $N \in \{k, \dots, \infty\}$  such that  $\forall n, m \geq N$  we have  $\|x_n - x_m\| < \frac{\varepsilon}{M_2}$  and thus  $\|x_n - x_m\|_2 \leq M_2 \cdot \|x_n - x_m\|_1 < M_2 \cdot \frac{\varepsilon}{M_2} = \varepsilon$   $\square$

**Theorem 11.357.** Let  $\langle X, d \rangle$  be a metric space,  $\{x_i\}_{i \in \{k, \dots, \infty\}}$  a sequence and  $x$  a limit point of  $\{x_i | i \in \{k, \dots, \infty\}\}$  then  $\forall N \in \{k, \dots, \infty\}, \varepsilon > 0$  there exists a  $n \in \{k, \dots, \infty\}$  with  $n > N$  such that  $d(x, x_n) < \varepsilon$

**Proof.** Given  $\varepsilon > 0$  take then  $A_\varepsilon = \{i \in \{k, \dots, \infty\} | 0 < d(x, x_i) < \varepsilon\} \subseteq \{k, \dots, \infty\}$ . We prove now that  $A_\varepsilon$  must be finite by contradiction. So assume that  $A_\varepsilon$  is finite. Then as  $x$  is a limit point of  $\{x_i | i \in \{k, \dots, \infty\}\}$  there exists a  $x_i \in \{x_i | i \in \{k, \dots, \infty\}\}$  with  $x_i \neq x$  and  $x \in B_d(x, \varepsilon) \Rightarrow 0 < d(x, x_i) < \varepsilon \Rightarrow i \in A_\varepsilon \Rightarrow A_\varepsilon \neq \emptyset$ . As we assumed that  $A_\varepsilon$  is finite, we have that  $B_\varepsilon = \{d(x, x_i) | i \in A_\varepsilon\}$  is finite (as  $f: A_\varepsilon \rightarrow B_\varepsilon, i \rightarrow f(i) = d(x, x_i)$  is a surjection) and we have by 5.50 the existence of a  $\delta = \min(B_\varepsilon) = \min(d(x, x_i) | i \in A_\varepsilon)$ . We must have that  $\delta > 0$  (as  $\forall i \in A_\varepsilon$  we have  $d(x, x_i) > 0$  by definition of  $A_\varepsilon$ ). Again by the fact that  $x$  is the limit point of  $\{x_i | i \in \{k, \dots, \infty\}\}$  that there exists a  $x_j \in \{x_i | i \in \{k, \dots, \infty\}\}$  such that  $x_j \neq x$  and  $x_j \in B_d(x, \delta) \Rightarrow 0 < d(x, x_j) < \delta \Rightarrow j \in A_\varepsilon \Rightarrow d(x, x_j) \in B_\varepsilon = \{d(x, x_i) | i \in A_\varepsilon\} \Rightarrow \delta \leq d(x, x_j) < \delta \Rightarrow \delta < \delta$  a contradiction. So we must have that  $A_\varepsilon$  is infinite.

If the theorem is not true then  $\exists N_0 \in \{k, \dots, \infty\}$  such that  $\forall n > N_0$  that  $d(x, x_n) \geq \varepsilon \Rightarrow n \notin A_\varepsilon$  and thus  $A_\varepsilon \subseteq \{k, \dots, N_0\}$  so  $A_\varepsilon$  would be finite contradicting the fact that  $A_\varepsilon$  is infinite.  $\square$

**Theorem 11.358.** Let  $\langle X, d \rangle$  be a metrics space then every Cauchy sequence has a maximum of one limit point.

**Proof.** Assume that  $x, y$  are different limit points of the Cauchy sequence  $\{x_i\}_{i \in \{k, \dots, \infty\}}$ . Then as  $x \neq y$  we have that  $\varepsilon = d(x, y) > 0$ . Now by the Cauchy condition we can find a  $N \in \{k, \dots, \infty\}$  such that  $\forall i, j \geq N$  we have  $d(x_i, x_j) < \frac{\varepsilon}{3}$ . By 11.357 there exists  $n, m \in \{k, \dots, \infty\}$  such that  $n, m > N$  and  $d(x, x_n) < \frac{\varepsilon}{2}, d(x, x_m) < \frac{\varepsilon}{2} \Rightarrow \varepsilon = d(x, y) \leq d(x, x_n) + d(x_n, x_m) + d(x_m, y) < \frac{\varepsilon}{2} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon \Rightarrow \varepsilon < \varepsilon$  a contradiction so we must have that  $x = y$ .  $\square$

### 11.11.3 Examples of complete spaces

**Theorem 11.359.** Every compact metric space  $\langle X, d \rangle$  is complete.

**Proof.** Let  $\{x_i\}_{i \in \{k, \dots, \infty\}}$  be a arbitrary Cauchy sequence and let  $S = \{x_i | i \in \{k, \dots, \infty\}\}$  then we must consider the following cases:

1. **(S is finite)** In this case there exists a  $N \in \{k, \dots, \infty\}$  such that  $\forall i, j \geq N$  we have  $x_i = x_j$
2. [We prove this by contradiction, so assume that  $\forall N \in \{k, \dots, \infty\}$  there exists  $i, j \geq N$  such that  $x_i \neq x_j$  so that  $A = \{d(x, y) | x, y \in S \wedge x \neq y\}$  is not empty (take  $N = k$  then there exists a  $i, j \geq k$  with  $x_i \neq x_j$ ) and finite (as by 5.44  $S \times S$  is finite and  $f: S \times S \rightarrow A$  defined by  $(x, y) \rightarrow d(x, y)$  is a surjection). By 5.50  $\varepsilon = \min(A)$  exists and is  $> 0$  (for if  $d \in A \Rightarrow d = d(x, y) > 0$  as  $x \neq y$ ). By the Cauchy option there exists then a  $N_1 \in \{k, \dots, \infty\}$  such that if  $i', j' \geq N_1$  then  $d(x_i, x_j) < \varepsilon$  and by the hypothesis there exists a  $i, j \geq N_1$  such that  $x_i \neq x_j \Rightarrow d(x_i, x_j) \in A \Rightarrow \varepsilon \leq d(x, y) < \varepsilon \Rightarrow \varepsilon < \varepsilon$  a contradiction]. Given a  $\varepsilon > 0$  take  $n \geq N$  then  $x_n = x_N \Rightarrow d(x_n, x_N) = 0 < \varepsilon$  proving that  $\{x_i\}_{i \in \{k, \dots, \infty\}}$  converges to  $x_N$ .]

3. ( **$S$  is infinite**) Then using 11.263 there exists limit points for  $S$  which because of the previous theorem must be unique (see 11.358), let's call this limit point  $x$ . Take now  $\varepsilon > 0$  then by the Cauchy property there exist a  $N \in \{k, \dots, \infty\}$  such that  $\forall i, j \geq N$  we have  $d(x_i, x_j) < \frac{\varepsilon}{2}$ , by 11.357 there exists a  $n \in \{k, \dots, \infty\}$  such that  $n \geq N$  and  $d(x, x_n) < \frac{\varepsilon}{2}$  so that if  $m \geq N$  then  $d(x, x_m) \leq d(x, x_n) + d(x_n, x_m) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$  proving convergence to  $x$ .  $\square$

**Theorem 11.360.** *The normed space  $\langle \mathbb{R}, \|\cdot\| \rangle$  of the real numbers together with the default metric is complete.*

**Proof.** Let  $\{x_i\}_{i \in \{k, \dots, \infty\}}$  be a Cauchy sequence in  $\mathbb{R}$  so that there exists a  $N \in \{k, \dots, \infty\}$  such that  $\forall i, j \geq N$  that  $|x_i - x_j| < 1$ . Take then  $A = \max(|x_1|, \dots, |x_{N-1}|, |x_N| + 1)$  and  $[-A, A]$ . If  $i \geq N$  we have  $|x_i| \leq |x_i - x_N| + |x_N| < 1 + |x_N| \leq A \Rightarrow x_i \in [-A, A]$  and if  $i < N \Rightarrow |x_i| \leq A$  so that  $\forall i \in \{k, \dots, \infty\}$  we have  $x_i \in [-A, A]$ . By 11.267  $[-A, A]$  is compact in  $\mathbb{R}$  (and thus compact in the subspace topology) and by the previous theorem (see 11.359)  $[-A, A]$  is complete in the subspace topology and thus as  $\{x_n\}_{n \in \{k, \dots, \infty\}}$  is trivially Cauchy in the subspace topology there exists a  $x \in [-A, A]$  such that  $\exists M \in \{k, \dots, \infty\}$  such that  $\forall m \geq M$  we have  $|x - x_m| = |x - x_m|_A < \varepsilon$  proving convergence to  $x$ .  $\square$

**Corollary 11.361.** *The normed space  $\langle \mathbb{C}, \|\cdot\|_c \rangle$  of the complex numbers together with the complex norm is complete. Here  $\|\cdot\|_c$  is the canonical norm in the complex space in contrast with the absolute value  $\|\cdot\|$ .*

**Proof.** Let  $\{z_i\}_{i \in \{k, \dots, \infty\}}$  be a Cauchy sequence, where  $\forall i \in \{k, \dots, \infty\}$  we have  $z_i = x_i + i \cdot y_i$ . Then  $\forall n, m \in \{k, \dots, \infty\}$  we have

$$\begin{aligned} |z_m - z_n| &= \sqrt{(x_m - x_n)^2} \\ &\leq_{9.72} \sqrt{(x_m - x_n)^2 + (y_m - y_n)^2} \\ &= |z_m - z_n|_c \end{aligned}$$

and like wise

$$\begin{aligned} |y_m - y_n| &= \sqrt{(y_m - y_n)^2} \\ &\leq_{9.72} \sqrt{(x_m - x_n)^2 + (y_m - y_n)^2} \\ &= |z_m - z_n|_c \end{aligned}$$

proving that  $\forall n, m \in \{k, \dots, \infty\} \models |x_m - x_n|, |y_m - y_n| \leq |z_m - z_n|_c$ . Given a  $\varepsilon > 0$  there exists a  $N \in \{k, \dots, \infty\}$  so that  $\forall n, m \geq N$  we have that  $|z_n - z_m|_c < \varepsilon \Rightarrow |x_n - x_m|, |y_n - y_m| < \varepsilon$ . So we have that  $\{x_i\}_{i \in \{k, \dots, \infty\}}, \{y_i\}_{i \in \{k, \dots, \infty\}}$  is Cauchy and as  $\langle \mathbb{R}, \|\cdot\| \rangle$  is complete  $x = \lim_{i \rightarrow \infty} x_i, y = \lim_{i \rightarrow \infty} y_i$  exists. So given a  $\varepsilon > 0$  there exists  $N_1, N_2 \in \{k, \dots, \infty\}$  such that if  $k \geq N_1, l \geq N_2$  that  $|x_n - x| < \frac{\varepsilon}{2}, |y_n - y| < \frac{\varepsilon}{2}$ . Then  $|z_m - z_n|_c = |(x_m - x_n) + i \cdot (y_m - y_n)| \leq |x_m - x_n|_c + |y_m - y_n|_c = |x_m - x_n| + |y_m - y_n| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$ . This proves that  $\lim_{i \rightarrow \infty} z_i$  exists and that  $\langle \mathbb{C}, \|\cdot\|_c \rangle$  is complete.  $\square$

**Theorem 11.362.** *Let  $\langle \mathbb{R}^n, \|\cdot\| \rangle$  be the normed space based on the maximum (product) norm (see 11.82)  $\|x\| = \max(\{|x_i| \mid i \in \{1, \dots, n\}\})$  then  $\langle \mathbb{R}^n, \|\cdot\| \rangle$  is a Banach space. As all norms on  $\mathbb{R}^n$  are equivalent (see 11.303) we have that  $\langle \mathbb{R}^n, \|\cdot\| \rangle$  is a Banach space for every norm  $\|\cdot\|$  on  $\mathbb{R}^n$ .*

**Proof.** If  $\{x_i\}_{i \in \{k, \dots, \infty\}}$  is a Cauchy sequence then by the previous lemma there exist a  $K \in \mathbb{R}_0^+$  such that  $\forall i \in \{k, \dots, \infty\}$  we have  $\|x_i\| < K \Rightarrow \|x_i\| \leq K$  or if we take  $\overline{B_{d_{\|\cdot\|}}}(0, K) = \{x \in \mathbb{R}^n \mid \|x\| \leq K\}$  then  $\forall i \in \{k, \dots, \infty\}$  we have  $x_i \in B$ .  $\overline{B_{d_{\|\cdot\|}}}(0, K)$  is closed (see 11.56) and trivially bounded [if  $x, y \in \overline{B_{d_{\|\cdot\|}}}(0, K)$  then  $d_{\|\cdot\|}(x, y) = \|x - y\| \leq \|x\| + \|y\| \leq K + K = 2 \cdot K$ ] so by the Heine-Borel theorem (11.300) we have that it is compact and thus by 11.359 is complete. So the sequence  $\{x_i\}_{i \in \{k, \dots, \infty\}}$  in  $\overline{B_{d_{\|\cdot\|}}}(0, K)$  being Cauchy in  $\overline{B_{d_{\|\cdot\|}}}(0, K)$  has a limit  $x$  which is trivially also the limit of  $\{x_i\}_{i \in \{k, \dots, \infty\}}$ .  $\square$

**Theorem 11.363.** *Every finite dimensional normed vector space is a Banach space*

**Proof.** Let  $\langle X, \|\cdot\| \rangle$  be a  $n$ -dimensional normed vector space then using 11.174 there exists a norm  $\|\cdot\|$  on  $\mathbb{R}^n$  and a isometry (also isomorphism)  $\varphi$  between  $X$  and  $\mathbb{R}^n$ . If then  $\{x_i\}_{i \in \{k, \dots, \infty\}}$  is a Cauchy sequence in  $X$  we have for  $\varepsilon > 0$  the existence of a  $N \in \{k, \dots, \infty\}$  such that if  $n, m \geq N$  that  $\|x_n - x_m\| < \varepsilon$ . Take now the sequence  $\{\varphi(x_i)\}_{i \in \{k, \dots, \infty\}}$  then if  $n, m \geq N$  we have that  $\|\varphi(x_n) - \varphi(x_m)\| = \|\varphi(x_n - x_m)\|_{\varphi \text{ is a isometry}} = \|x_n - x_m\| < \varepsilon$  proving that  $\{\varphi(x_i)\}_{i \in \{k, \dots, \infty\}}$  is Cauchy, using the previous theorem we have that  $\langle \mathbb{R}^n, \|\cdot\| \rangle$  is a Banach space and so  $\{\varphi(x_i)\}_{i \in \{k, \dots, \infty\}}$  has a limit  $y \in \mathbb{R}^n$  or in other words if  $\varepsilon > 0$  then there exists a  $N \in \{k, \dots, \infty\}$  such that if  $n \geq N$  then  $\|\varphi(x_i) - y\| < \varepsilon$ . If we then take  $x = \varphi^{-1}(y)$  then if  $n \geq N$  we have that  $\|x_n - x\| = \|\varphi(x_n - x)\| = \|\varphi(x_n) - \varphi(x)\| = \|\varphi(x_n) - \varphi(\varphi^{-1}(y))\| = \|\varphi(x_n) - y\| < \varepsilon$ . This finally proves that  $\langle X, \|\cdot\| \rangle$  is a Banach space.  $\square$

**Theorem 11.364.** Let  $\{\langle X_i, \|\cdot\|_i \rangle\}_{i \in \{1, \dots, n\}}$  be a family of normed spaces and let  $\langle Y, \|\cdot\| \rangle$  be a Banach space then  $L(X_1, \dots, X_n; Y)$  is a Banach space using the operator norm.

**Proof.** Let  $\{L_i\}_{i \in \{k, \dots, \infty\}}$  be a Cauchy sequence in  $L(X_1, \dots, X_n; Y)$  and take  $(x_1, \dots, x_n) \in \prod_{i \in \{k, \dots, \infty\}} X_i$  then we have for  $p, q \in \{k, \dots, \infty\}$

$$\|L_p(x_1, \dots, x_n) - L_q(x_1, \dots, x_n)\| = \|(L_p - L_q)(x_1, \dots, x_n)\| \leq \|L_p - L_q\| \cdot \prod_{i \in \{1, \dots, n\}} \|x_i\|_i \quad (11.36)$$

Take now the family  $\{L_p(x_1, \dots, x_n)\}_{p \in \{k, \dots, \infty\}}$  and take  $\varepsilon > 0$  then we have two possible cases:

1. ( $\prod_{i \in \{1, \dots, n\}} \|x_i\|_i = 0$ ) then by 11.36 we have that  $\|L_p(x_1, \dots, x_n) - L_q(x_1, \dots, x_n)\| = 0 < \varepsilon$  so if  $p, q \geq k \in \{k, \dots, \infty\}$  then  $\|L_p(x_1, \dots, x_n) - L_q(x_1, \dots, x_n)\| < \varepsilon$
2. ( $\prod_{i \in \{1, \dots, n\}} \|x_i\|_i \neq 0$ ) take then  $\varepsilon' = \frac{\varepsilon}{\prod_{i \in \{1, \dots, n\}} \|x_i\|_i}$  and as  $\{L_p\}_{p \in \{k, \dots, \infty\}}$  is Cauchy there exists a  $N \in \{k, \dots, \infty\}$  such that if  $p, q \geq N$  then  $\|L_p - L_q\| < \varepsilon' \Rightarrow \|L_p(x_1, \dots, x_n) - L_q(x_1, \dots, x_n)\| \leq \|L_p - L_q\| \cdot \prod_{i \in \{1, \dots, n\}} \|x_i\|_i < \varepsilon' \cdot \prod_{i \in \{1, \dots, n\}} \|x_i\|_i = \varepsilon$

from (1) and (2) we conclude then that

$$\{L_p(x_1, \dots, x_n)\}_{p \in \{k, \dots, \infty\}} \text{ is a Cauchy sequence in } Y \quad (11.37)$$

As  $\langle Y, \|\cdot\| \rangle$  is a Banach space it follows from 11.37 that  $\{L_p(x_1, \dots, x_n)\}_{p \in \{k, \dots, \infty\}}$  has a limit (depending of course on  $(x_1, \dots, x_n)$ ) let's note this limit as  $L(x_1, \dots, x_n)$  so we have

$$L(x_1, \dots, x_n) = \lim_{q \rightarrow \infty} L_p(x_1, \dots, x_n) \quad (11.38)$$

Define now the function  $L: \prod_{i \in \{1, \dots, n\}} X_i \rightarrow Y$  by  $(x_1, \dots, x_n) \mapsto L(x_1, \dots, x_n)$ . We prove now that  $L \in L(X_1, \dots, X_n; Y)$ :

1. (**multilinearity**) Given  $\alpha, \beta \in \mathbb{K}$ ,  $(x_1, \dots, x_n) \in \prod_{i \in \{1, \dots, n\}} X_i$ ,  $i \in \{1, \dots, n\}$  and  $x, y \in X_i$  then we have that  $L(x_1, \dots, x_{i-1}, \alpha \cdot x + \beta \cdot y, x_{i+1}, \dots, x_n) = \lim_{p \rightarrow \infty} L_p(x_1, \dots, x_{i-1}, \alpha \cdot x + \beta \cdot y, x_{i+1}, \dots, x_n)$   $\stackrel{L_p \text{ is multilinear}}{=} \lim_{p \rightarrow \infty} (\alpha \cdot L_p(x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_n) + \beta \cdot L_p(x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_n)) \stackrel{11.353}{=} \alpha \cdot \lim_{p \rightarrow \infty} L_p(x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_n) + \beta \cdot \lim_{p \rightarrow \infty} L_p(x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_n) = \alpha \cdot L(x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_n) + \beta \cdot L(x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_n)$
2. (**continuity**) Take the family  $\{\|L_p\|\}_{p \in \{k, \dots, \infty\}}$  in  $\mathbb{R}$  then if  $\varepsilon > 0$  we have from the fact that  $\{L_p\}_{p \in \{k, \dots, \infty\}}$  is Cauchy there exists a  $N \in \{k, \dots, \infty\}$  such that if  $p, q \geq N$  then  $\|L_p - L_q\| < \varepsilon \Rightarrow \|\|L_p\| - \|L_q\|\| \leq \|L_p - L_q\| < \varepsilon$  proving that  $\{\|L_p\|\}_{p \in \{k, \dots, \infty\}}$  is Cauchy in  $\langle \mathbb{R}, \|\cdot\| \rangle$  and as  $\langle \mathbb{R}, \|\cdot\| \rangle$  is complete (see 11.360) there exists a  $A \in \mathbb{R}$  with  $A = \lim_{p \rightarrow \infty} \|L_p\|$ . Using 11.351 and  $\|L_p(x_1, \dots, x_n)\| \leq \|L_p\| \cdot \prod_{i \in \{1, \dots, n\}} \|x_i\|_i$  we have that  $\|L(x_1, \dots, x_n)\| = \lim_{p \rightarrow \infty} \|L_p(x_1, \dots, x_n)\| \leq \lim_{p \rightarrow \infty} (\|L_p\| \cdot \prod_{i \in \{1, \dots, n\}} \|x_i\|_i) = \left( \lim_{p \rightarrow \infty} \|L_p\| \right) \cdot \prod_{i \in \{1, \dots, n\}} \|x_i\|_i = A \cdot \prod_{i \in \{1, \dots, n\}} \|x_i\|_i$  proving that  $L$  is indeed continuous and that

$$\|L\| \leq A = \lim_{p \rightarrow \infty} \|L_p\| \quad (11.39)$$

So the only thing left to prove is that  $\{L_q\}_{q \in \{k, \dots, \infty\}}$  converges to  $L$ . Let  $\varepsilon > 0$  then for every  $(x_1, \dots, x_n) \in \prod_{i \in \{1, \dots, n\}} X_i$  with  $\forall i \in \{1, \dots, n\} \|x_i\| = 1$  we have by the fact that  $\{L_q\}_{q \in \{1, \dots, \infty\}}$  is Cauchy there exists a  $N_1 \in \{k, \dots, \infty\}$  such that if  $p, q \geq N_1$  then  $\|L_p - L_q\| < \frac{\varepsilon}{2}$  so that  $\|L_p(x_1, \dots, x_n) - L_q(x_1, \dots, x_n)\| = \|(L_p - L_q)(x_1, \dots, x_n)\| \leq \|L_p - L_q\| \cdot \prod_{i \in \{1, \dots, n\}} \|x_i\|_i = \|L_p - L_q\| < \frac{\varepsilon}{4}$ . Further using 11.38 there exists a  $N_2(x_1, \dots, x_n) \in \{k, \dots, \infty\}$  such that if  $r \geq N_2(x_1, \dots, x_n)$  then  $\|L(x_1, \dots, x_n) - L_r(x_1, \dots, x_n)\| < \frac{\varepsilon}{4}$ . So that if  $p \geq N_1$  then we have  $\|(L - L_p)(x_1, \dots, x_n)\| = \|L(x_1, \dots, x_n) - L_p(x_1, \dots, x_n)\| \leq \|L(x_1, \dots, x_n) - L_s(x_1, \dots, x_n)\| + \|L_s(x_1, \dots, x_n) - L_p(x_1, \dots, x_n)\| < \frac{\varepsilon}{4} + \frac{\varepsilon}{4} = \frac{\varepsilon}{2}$  if we choose  $s \geq \max(N_1, N_2(x_1, \dots, x_n))$  proving that if  $p \geq N_1$  then  $\|(L - L_p)(x_1, \dots, x_n)\| < \frac{\varepsilon}{2}$  proving by 11.198 that  $\|L - L_p\| \leq \frac{\varepsilon}{2} < \varepsilon$  and thus that  $\lim_{p \rightarrow \infty} L_p = L$ .  $\square$

**Corollary 11.365.** Let  $\langle X, \|\cdot\|_X \rangle$  be a normed space,  $\langle Y, \|\cdot\|_Y \rangle$  a Banach space,  $n \in \mathbb{N}_0$  then  $\langle L^n(X; Y), \|\cdot\|_{L^n(X; Y)} \rangle$  is a Banach space (using the operator norms)

**Proof.** Let  $S = \{n \in \mathbb{N}_0 \mid \text{If } \langle X, \|\cdot\|_X \rangle \text{ is a normed space, } \langle Y, \|\cdot\|_Y \rangle \text{ a Banach space then } L^n(A; Y) \text{ is a Banach space}\}$  then:

1. If  $n = 1$  then  $L^1(X; Y) = L(X, Y) = L(X; Y)$  is a Banach space by the above theorem, so that  $1 \in S$ .
  2. If  $n \in S$  then  $L^n(X; Y)$  is a Banach space (if  $\langle X, \|\cdot\|_X \rangle$  is a normed space,  $\langle Y, \|\cdot\|_Y \rangle$  a Banach space) so that by (1)  $L^{n+1}(X; Y) = L(X; L^n(X; Y))$  is a Banach space proving that  $n+1 \in S$
- Induction proves then the theorem.  $\square$

For the real numbers we have for increasing [or decreasing] sequences the following result

**Lemma 11.366.** Let  $\{x_i\}_{i \in \{k, \dots, \infty\}} \subseteq \mathbb{R}$  then we have

1. If  $\{x_i\}_{i \in \{k, \dots, \infty\}}$  is increasing [ $\forall i \in \{k, \dots, \infty\}$  we have  $x_i \leq x_{i+1}$ ] then
  - a. If  $\sup(\{x_i \mid i \in \{k, \dots, \infty\}\})$  exists then  $\lim_{i \rightarrow \infty} x_i$  exists and  $\lim_{i \rightarrow \infty} x_i = \sup(\{x_i \mid i \in \{k, \dots, \infty\}\})$
  - b. If  $\lim_{i \rightarrow \infty} x_i$  exists then  $\sup(\{x_i \mid i \in \{k, \dots, \infty\}\})$  exists and  $\lim_{i \rightarrow \infty} x_i = \sup(\{x_i \mid i \in \{k, \dots, \infty\}\})$
2. If  $\{x_i\}_{i \in \{k, \dots, \infty\}}$  is decreasing [ $\forall i \in \{k, \dots, \infty\}$  we have  $x_{i+1} \leq x_i$ ] then
  - a. If  $\inf(\{x_i \mid i \in \{k, \dots, \infty\}\})$  exists then  $\lim_{i \rightarrow \infty} x_i$  exists and  $\lim_{i \rightarrow \infty} x_i = \inf(\{x_i \mid i \in \{k, \dots, \infty\}\})$
  - b. If  $\lim_{i \rightarrow \infty} x_i$  exists then  $\inf(\{x_i \mid i \in \{k, \dots, \infty\}\})$  exists and  $\lim_{i \rightarrow \infty} x_i = \inf(\{x_i \mid i \in \{k, \dots, \infty\}\})$

**Proof.**

1. Assume that  $\{x_i\}_{i \in \{k, \dots, \infty\}}$  is increasing then
  - a. Let  $x = \sup(\{x_i \mid i \in \{k, \dots, \infty\}\})$  and take  $\varepsilon > 0$  then  $x - \varepsilon < x < x + \varepsilon$  and by the definition of a supremum there exists a  $N \in \{k, \dots, \infty\}$  such that  $x - \varepsilon < x_N \leq x < x + \varepsilon$ . If now  $n \geq N$  we have (see 11.311) that  $x_N \leq x_n \leq x \Rightarrow x - \varepsilon < x_n \leq x < x + \varepsilon \Rightarrow |x - x_n| < \varepsilon$  proving that  $\lim_{i \rightarrow \infty} x_i = x = \sup(\{x_i \mid i \in \{k, \dots, \infty\}\})$
  - b. As  $\lim_{i \rightarrow \infty} x_i = x$  we have for  $1 > 0$  that there exists a  $N \in \{k, \dots, \infty\}$  such that  $\forall i \geq N$  we have  $|x_i - x| < 1 \Rightarrow x_i - x < 1 \Rightarrow x_i < 1 + x$ . Hence  $\forall i \in \{k, \dots, \infty\}$  we have  $x_i \leq \max(\{x_i \mid i \in \{k, \dots, N\}\} \cup \{1 + x\})$ . Using 9.43 we have then that  $\sup(\{x_i \mid i \in \{k, \dots, \infty\}\})$  exists. Let  $y = \sup(\{x_i \mid i \in \{k, \dots, \infty\}\})$  then using (a) that we have just proved we get that  $y = \lim_{i \rightarrow \infty} x_i = x$  proving using the uniqueness of a limit that  $\sup(\{x_i \mid i \in \{k, \dots, \infty\}\}) = x$

2. Assume that  $\{x_i\}_{i \in \{k, \dots, \infty\}}$  is decreasing then

- Let  $x = \inf(\{x_i | i \in \{k, \dots, \infty\}\})$  and take  $\varepsilon > 0$  then  $x - \varepsilon < x < x + \varepsilon$  and by the definition of a infimum there exists a  $N \in \{k, \dots, \infty\}$  such that  $x - \varepsilon < x \leq x_N < x + \varepsilon$ . If now  $n \geq N$  we have (see 11.311) that  $x \leq x_n \leq x_N \Rightarrow x - \varepsilon < x \leq x_n < x + \varepsilon \Rightarrow |x - x_n| < \varepsilon$  proving that  $\lim_{i \rightarrow \infty} x_i = x = \inf(\{x_i | i \in \{k, \dots, \infty\}\})$
- As  $\lim_{i \rightarrow \infty} x_i = x$  we have for  $1 > 0$  that there exists a  $N \in \{k, \dots, \infty\}$  such that  $\forall i \geq N$  we have  $|x_i - x| < 1 \Rightarrow x - x_i < 1 \Rightarrow x - 1 < x_i$ . Hence  $\forall i \in \{k, \dots, \infty\}$  we have  $\min(\{x_i | i \in \{k, \dots, N\} \cup \{x - 1\}\}) \leq x_i$ . Using 9.43 we have then that  $\inf(\{x_i | i \in \{k, \dots, \infty\}\})$  exists. Let  $y = \inf(\{x_i | i \in \{k, \dots, \infty\}\})$  then using (a) that we have just proved we get that  $y = \lim_{i \rightarrow \infty} x_i = x$  proving using the uniqueness of a limit that  $\inf(\{x_i | i \in \{k, \dots, \infty\}\}) = x$   $\square$

**Example 11.367.** Let we have that

- $\lim_{n \rightarrow \infty} \frac{1}{n} = 0$
- $\inf(\{\frac{1}{n} | n \in \mathbb{N}_0\}) = 0$

**Proof.**

- This is proved in 11.331.
- As  $\forall n \in \mathbb{N}_0$  we have  $\frac{1}{n+1} \leq \frac{1}{n} \{\frac{1}{n}\}_{n \in \mathbb{N}_0}$  is a decreasing sequence, hence by the previous lemma we have  $\inf(\{\frac{1}{n} | n \in \mathbb{N}_0\}) = \lim_{n \rightarrow \infty} \frac{1}{n} = 0$ .  $\square$

#### 11.11.4 Limit infimum and limit supremum in $\mathbb{R}$

For this subsection we use the normed space  $\langle \mathbb{R}, \|\cdot\| \rangle$  based on the absolute value in  $\mathbb{R}$

**Definition 11.368.** Let  $\{x_n\}_{n \in \{k, \dots, \infty\}} \subseteq \mathbb{R}$  then we say

- $\limsup_{n \rightarrow \infty} x_i$  exists iff  $\forall n \in \{k, \dots, \infty\}$   $\sup(\{x_i | i \in \{n, \dots, \infty\}\})$  exists and  $\limsup_{n \rightarrow \infty} (\{x_i | i \in \{n, \dots, \infty\}\})$  exists. In that case  $\limsup_{n \rightarrow \infty} x_i = \limsup_{n \rightarrow \infty} (\{x_i | i \in \{n, \dots, \infty\}\})$ .
- $\liminf_{n \rightarrow \infty} x_i$  exists iff  $\forall n \in \{k, \dots, \infty\}$   $\inf(\{x_i | i \in \{n, \dots, \infty\}\})$  exists and  $\liminf_{n \rightarrow \infty} (\{x_i | i \in \{n, \dots, \infty\}\})$  exists. In that case  $\liminf_{n \rightarrow \infty} x_i = \liminf_{n \rightarrow \infty} (\{x_i | i \in \{n, \dots, \infty\}\})$ .

**Proposition 11.369.** Let  $\{x_i\}_{i \in \{k, \dots, \infty\}} \subseteq \mathbb{R}$  then

- $\limsup_{n \rightarrow \infty} x_i$  exists if and only if  $\forall n \in \{k, \dots, \infty\}$   $\sup(\{x_i | i \in \{n, \dots, \infty\}\})$  exists and  $\inf(\{\sup(\{x_i | i \in \{n, \dots, \infty\}\}) | n \in \{k, \dots, \infty\}\})$  exists. Further if  $\limsup_{n \rightarrow \infty} x_i$  exists then  $\limsup_{n \rightarrow \infty} x_i = \inf(\{\sup(\{x_i | i \in \{n, \dots, \infty\}\}) | n \in \{k, \dots, \infty\}\})$ .
- $\liminf_{n \rightarrow \infty} x_i$  exists if and only if  $\forall n \in \{k, \dots, \infty\}$   $\inf(\{x_i | i \in \{n, \dots, \infty\}\})$  exists and  $\sup(\{\inf(\{x_i | i \in \{n, \dots, \infty\}\}) | n \in \{k, \dots, \infty\}\})$  exists. Further if  $\liminf_{n \rightarrow \infty} x_i$  exists then  $\liminf_{n \rightarrow \infty} x_i = \sup(\{\inf(\{x_i | i \in \{n, \dots, \infty\}\}) | n \in \{k, \dots, \infty\}\})$

**Proof.**

- If  $\forall n \in \{k, \dots, \infty\}$   $\sup(\{x_i | i \in \{n, \dots, \infty\}\})$  exists then  $\forall n \in \{k, \dots, \infty\}$  we have  $\sup(\{x_i | i \in \{n, \dots, \infty\}\}) \leq \sup(\{x_i | i \in \{n+1, \dots, \infty\}\})$ . Applying then 11.366 proves (1)
- If  $\forall n \in \{k, \dots, \infty\}$   $\inf(\{x_i | i \in \{n, \dots, \infty\}\})$  exists then  $\forall n \in \{k, \dots, \infty\}$  we have  $\inf(\{x_i | i \in \{n, \dots, \infty\}\}) \leq \inf(\{x_i | i \in \{n+1, \dots, \infty\}\})$ . Applying then 11.366 proves (1)  $\square$

**Example 11.370.** Let  $x \in \mathbb{R}$  then if we define  $\{x_i\}_{i \in \{k, \dots, \infty\}}$  by  $x_i = x$  then  $\liminf_{n \rightarrow \infty} x_i, \limsup_{n \rightarrow \infty} x_i$  exists and  $x = \liminf_{n \rightarrow \infty} x_i = \limsup_{n \rightarrow \infty} x_i$

**Proof.** Let  $n \in \{k, \dots, \infty\}$  then  $\sup(\{x_i | i \in \{n, \dots, \infty\}\}) = \sup(\{x\}) = x$  and  $\inf(\{x_i | i \in \{1, \dots, n\}\}) = \inf(\{x\}) = x$  and thus  $\inf(\{\sup(\{x_i | i \in \{n, \dots, \infty\}\}) | n \in \{k, \dots, \infty\}\}) = \inf(\{x\}) = x$  and  $\sup(\{\inf(\{x_i | i \in \{n, \dots, \infty\}\}) | n \in \{k, \dots, \infty\}\}) = \sup(\{x\}) = x$  which by the previous proposition [see 11.369] proves that  $\liminf_{n \rightarrow \infty} x_i, \limsup_{n \rightarrow \infty} x_i$  exists and  $x = \liminf_{n \rightarrow \infty} x_i = \limsup_{n \rightarrow \infty} x_i$   $\square$

**Proposition 11.371.** Let  $\{x_i\}_{i \in \{k, \dots, \infty\}} \subseteq \mathbb{R}$  be such that  $\liminf_{n \rightarrow \infty} x_i$  and  $\limsup_{n \rightarrow \infty} x_i$  exists then  $\liminf_{n \rightarrow \infty} x_n \leq \limsup_{n \rightarrow \infty} x_n$

**Proof.** Fix  $n \in \{k, \dots, \infty\}$  then  $\forall m \in \{k, \dots, \infty\}$  we have either

$m \in \{k, \dots, n-1\}$ . then  $m \leq n$  so that  $\{x_i | i \in \{n, \dots, \infty\}\} \subseteq \{x_i | i \in \{m, \dots, \infty\}\}$  hence  $\inf(\{x_i | i \in \{m, \dots, \infty\}\}) \leq \inf(\{x_i | i \in \{n, \dots, \infty\}\}) \leq x_n \leq \sup(\{x_i | i \in \{n, \dots, \infty\}\})$

$m \in \{n, \dots, \infty\}$ . then  $\inf(\{x_i | i \in \{m, \dots, \infty\}\}) \leq x_n \leq \sup(\{x_i | i \in \{n, \dots, \infty\}\})$

so we have

$$\forall m \in \{k, \dots, \infty\} \text{ we have } \inf(\{x_i | i \in \{m, \dots, \infty\}\}) \leq x_n \leq \sup(\{x_i | i \in \{n, \dots, \infty\}\})$$

and thus

$$\sup(\{\inf(\{x_i | i \in \{m, \dots, \infty\}\}) | m \in \{k, \dots, \infty\}\}) \leq \sup(\{x_i | i \in \{n, \dots, \infty\}\})$$

giving that

$$\sup(\{\inf(\{x_i | i \in \{m, \dots, \infty\}\}) | m \in \{k, \dots, \infty\}\}) \leq \inf(\{\sup(\{x_i | i \in \{n, \dots, \infty\}\}) | n \in \{k, \dots, \infty\}\})$$

using then 11.369 we have

$$\liminf_{n \rightarrow \infty} x_n \leq \limsup_{n \rightarrow \infty} x_n$$

proving the theorem.  $\square$

**Proposition 11.372.** Let  $\{x_i\}_{i \in \{k, \dots, \infty\}} \subseteq \mathbb{R}$  then  $\lim_{n \rightarrow \infty} x_i$  exists if and only if  $\limsup_{n \rightarrow \infty} x_i$  exists and  $\liminf_{n \rightarrow \infty} x_i$  exists and  $\liminf_{n \rightarrow \infty} x_i = \limsup_{n \rightarrow \infty} x_i$ . If  $\lim_{n \rightarrow \infty} x_i$  exists then  $\lim_{n \rightarrow \infty} x_i = \liminf_{n \rightarrow \infty} x_i = \limsup_{n \rightarrow \infty} x_i$ .

**Proof.**

1. Assume that  $x = \lim_{n \rightarrow \infty} x_i$  exists then by 11.336 there exists a  $K \in \mathbb{R}_0^+$  so that  $\forall i \in \{k, \dots, \infty\}$  we have  $|x_i| < K$  so that  $-x_i, x_i < K$  giving

$$\forall i \in \{k, \dots, \infty\} \text{ we have } -K < x_i < K \quad (11.40)$$

Let  $n \in \{k, \dots, \infty\}$  then  $\{x_i | i \in \{n, \dots, \infty\}\}$  is bounded above by  $K$  and below by  $-K$ . Using the conditional completeness of  $\mathbb{R}$  [see 9.43] we have

$$\forall n \in \{k, \dots, \infty\} \quad \sup(\{x_i | i \in \{n, \dots, \infty\}\}) \text{ and } \inf(\{x_i | i \in \{n, \dots, \infty\}\}) \text{ exists} \quad (11.41)$$

and using the definition of the supremum and infimum we have

$$-K \leq \sup(\{x_i | i \in \{n, \dots, \infty\}\}) \text{ and } \inf(\{x_i | i \in \{n, \dots, \infty\}\}) \leq K \quad (11.42)$$

Using the above we have that  $\{\sup(\{x_i | i \in \{n, \dots, \infty\}\}) | n \in \{k, \dots, \infty\}\}$  is bounded below by  $-K$  and  $\{\inf(\{x_i | i \in \{n, \dots, \infty\}\}) | n \in \{k, \dots, \infty\}\}$  is bounded above by  $K$  hence we have that

$$\inf(\{\sup(\{x_i | i \in \{n, \dots, \infty\}\}) | n \in \{k, \dots, \infty\}\}) \text{ and } \sup(\{\inf(\{x_i | i \in \{n, \dots, \infty\}\}) | n \in \{k, \dots, \infty\}\}) \text{ exists} \quad (11.43)$$

Using then 11.369 we have that

$$\limsup_{n \rightarrow \infty} x_n \text{ and } \liminf_{n \rightarrow \infty} x_n \text{ exists} \quad (11.44)$$

Assume now that  $x < \sup(\{\inf(\{x_i | i \in \{n, \dots, \infty\}\}) | n \in \{k, \dots, \infty\}\})$  then there exists a  $l \in \{k, \dots, \infty\}$  such that  $x < \inf(\{x_i | i \in \{l, \dots, \infty\}\})$ . Take  $\varepsilon = \inf(\{x_i | i \in \{l, \dots, \infty\}\}) - x$  then  $\varepsilon > 0$ . Hence there exists a  $N \in \{k, \dots, \infty\}$  such that  $\forall i \in \{N, \dots, \infty\}$  we have  $|x - x_i| < \varepsilon$ . Take  $M = \max(l, N)$  then we have  $x - \varepsilon < x_M < x + \varepsilon = x + \inf(\{x_i | i \in \{l, \dots, \infty\}\}) - x = \inf(\{x_i | i \in \{l, \dots, \infty\}\}) \leq x_M$  leading to the contradiction  $x_M < x_M$ . Hence we must have

$$\sup(\{\inf(\{x_i | i \in \{n, \dots, \infty\}\}) | n \in \{k, \dots, \infty\}\}) < x = \lim_{n \rightarrow \infty} x_n \quad (11.45)$$

Assume now that  $\inf(\{\sup(\{x_i | i \in \{n, \dots, \infty\}\}) | n \in \{k, \dots, \infty\}\}) < x$  then  $\exists l \in \{k, \dots, \infty\}$  such that  $\sup(\{x_i | i \in \{l, \dots, \infty\}\}) < x$ . Take  $\varepsilon = x - \sup(\{x_i | i \in \{l, \dots, \infty\}\})$  then  $0 < \varepsilon$  and there exists a  $N \in \{k, \dots, \infty\}$  such that  $\forall i \in \{N, \dots, \infty\}$  we have  $|x - x_i| < \varepsilon$ . Take  $M = \max(N, l)$  then we have  $x + \varepsilon > x_M > x - \varepsilon = x - x + \sup(\{x_i | i \in \{l, \dots, \infty\}\}) = \sup(\{x_i | i \in \{l, \dots, \infty\}\}) \geq x_M$  leading to the contradiction  $x_M > x_M$ . Hence we must have

$$\lim_{n \rightarrow \infty} x_n = x \leq \sup(\{\inf(\{x_i | i \in \{n, \dots, \infty\}\}) | n \in \{k, \dots, \infty\}\}) \quad (11.46)$$

Finally we have  $\lim_{n \rightarrow \infty} x_n \leq \sup(\{\inf(\{x_i | i \in \{n, \dots, \infty\}\}) | n \in \{k, \dots, \infty\}\}) \leq \inf(\{\sup(\{x_i | i \in \{n, \dots, \infty\}\}) | n \in \{k, \dots, \infty\}\}) \leq \lim_{n \rightarrow \infty} x_i$  which proves that

$$\lim_{n \rightarrow \infty} x_i = \liminf_{n \rightarrow \infty} x_i = \limsup_{n \rightarrow \infty} x_i$$

2. Assume that  $\limsup_{n \rightarrow \infty} x_i$  exists and  $\liminf_{n \rightarrow \infty} x_i$  exists and  $\liminf_{n \rightarrow \infty} x_i = \limsup_{n \rightarrow \infty} x_i$ . Then using 11.369 we have that  $x = \liminf_{n \rightarrow \infty} x_i = \limsup_{n \rightarrow \infty} x_i = \sup(\{\inf(\{x_i | i \in \{n, \dots, \infty\}\}) | n \in \{k, \dots, \infty\}\}) \leq \inf(\{\sup(\{x_i | i \in \{n, \dots, \infty\}\}) | n \in \{k, \dots, \infty\}\})$ . Take  $\varepsilon > 0$  then we have as  $x - \varepsilon < x < x + \varepsilon$  that there exists  $N_1, N_2 \in \{k, \dots, \infty\}$  such that  $x - \varepsilon < \inf(\{x_i | i \in \{N_1, \dots, \infty\}\})$  and  $\sup(\{x_i | i \in \{N_2, \dots, \infty\}\}) < x + \varepsilon$ . Take  $N = \max(N_1, N_2)$  then  $\forall i \in \{N, \dots, \infty\}$  we have  $x - \varepsilon < \inf(\{x_i | i \in \{N, \dots, \infty\}\}) \leq x_i \leq \sup(\{x_i | i \in \{N, \dots, \infty\}\}) < x + \varepsilon$  proving that  $|x - x_i| < \varepsilon$ . So  $\lim_{n \rightarrow \infty} x_i$  exists and is equal to  $\liminf_{n \rightarrow \infty} x_i$  and  $\limsup_{n \rightarrow \infty} x_i$ .  $\square$

**Lemma 11.373.** Let  $\{x_i\}_{i \in \{k, \dots, \infty\}}, \{y_i\}_{i \in \{k, \dots, \infty\}} \subseteq \mathbb{R}$  the we have

1. If  $\sup(\{x_i | i \in \{k, \dots, \infty\}\}), \sup(\{y_i | i \in \{k, \dots, \infty\}\})$  exists then  $\sup(\{x_i + y_i | i \in \{k, \dots, \infty\}\})$  exists and  $\sup(\{x_i + y_i | i \in \{k, \dots, \infty\}\}) \leq \sup(\{x_i | i \in \{k, \dots, \infty\}\}) + \sup(\{y_i | i \in \{k, \dots, \infty\}\})$
2. If  $\inf(\{x_i | i \in \{k, \dots, \infty\}\}), \inf(\{y_i | i \in \{k, \dots, \infty\}\})$  exists then  $\inf(\{x_i + y_i | i \in \{k, \dots, \infty\}\})$  exists and  $\inf(\{x_i + y_i | i \in \{k, \dots, \infty\}\}) \geq \inf(\{x_i | i \in \{k, \dots, \infty\}\}) + \inf(\{y_i | i \in \{k, \dots, \infty\}\})$

**Proof.**

- Take  $i \in \{k, \dots, \infty\}$  then  $x_i + y_i \leq \sup(\{x_i | i \in \{k, \dots, \infty\}\}) + \sup(\{y_i | i \in \{k, \dots, \infty\}\})$  we have that  $\{x_i + y_i | i \in \{k, \dots, \infty\}\}$  is bounded above by  $\sup(\{x_i | i \in \{k, \dots, \infty\}\}) + \sup(\{y_i | i \in \{k, \dots, \infty\}\})$ , as  $\mathbb{R}$  is conditionally complete we have that  $\sup(\{x_i + y_i | i \in \{k, \dots, \infty\}\})$  exists and

$$\sup(\{x_i + y_i | i \in \{k, \dots, \infty\}\}) \leq \sup(\{x_i | i \in \{k, \dots, \infty\}\}) + \sup(\{y_i | i \in \{k, \dots, \infty\}\})$$

- Take  $i \in \{k, \dots, \infty\}$  then  $\inf(\{x_i | i \in \{k, \dots, \infty\}\}) + \inf(\{y_i | i \in \{k, \dots, \infty\}\}) \leq x_i + y_i$  so that  $\{x_i + y_i | i \in \{k, \dots, \infty\}\}$  is bounded below by  $\inf(\{x_i | i \in \{k, \dots, \infty\}\}) + \inf(\{y_i | i \in \{k, \dots, \infty\}\})$ , as  $\mathbb{R}$  is conditionally complete we have that  $\inf(\{x_i + y_i | i \in \{k, \dots, \infty\}\})$  exists and

$$\inf(\{x_i | i \in \{k, \dots, \infty\}\}) + \inf(\{y_i | i \in \{k, \dots, \infty\}\}) \leq \inf(\{x_i + y_i | i \in \{k, \dots, \infty\}\})$$

□

**Proposition 11.374.** Take  $\{x_i\}_{i \in \{k, \dots, \infty\}}$  then we have

1. If  $\liminf_{n \rightarrow \infty} x_i$  exists then  $\limsup_{n \rightarrow \infty} (-x_i)$  exists and  $\liminf_{n \rightarrow \infty} x_i = -\limsup_{n \rightarrow \infty} (-x_i)$
2. If  $\limsup_{n \rightarrow \infty} x_i$  exists then  $\liminf_{n \rightarrow \infty} (-x_i)$  exists and  $\limsup_{n \rightarrow \infty} x_i = -\liminf_{n \rightarrow \infty} (-x_i)$
3. If  $\limsup_{n \rightarrow \infty} x_i$  exists and  $\alpha \in [0, \infty[$  then  $\limsup_{n \rightarrow \infty} (\alpha \cdot x_i)$  exists and  $\limsup_{n \rightarrow \infty} (\alpha \cdot x_i) = \alpha \cdot \limsup_{n \rightarrow \infty} x_i$
4. If  $\liminf_{n \rightarrow \infty} x_i$  exists and  $\alpha \in [0, \infty[$  then  $\liminf_{n \rightarrow \infty} (\alpha \cdot x_i)$  exists and  $\liminf_{n \rightarrow \infty} (\alpha \cdot x_i) = \alpha \cdot \liminf_{n \rightarrow \infty} x_i$
5. If  $x \in \mathbb{R}$  and  $\liminf_{n \rightarrow \infty} x_i$  exists then for  $\liminf_{n \rightarrow \infty} (x_i + x)$  exists and  $\liminf_{n \rightarrow \infty} (x_i + x) = \liminf_{n \rightarrow \infty} x_i + x$
6. If  $x \in \mathbb{R}$  and  $\limsup_{n \rightarrow \infty} x_i$  exists then for  $\{x_i + x\}_{i \in \{k, \dots, \infty\}}$   $\limsup_{n \rightarrow \infty} (x_i + x)$  exists and  $\limsup_{n \rightarrow \infty} (x_i + x) = \limsup_{n \rightarrow \infty} x_i + x$

**Proof.**

1. As  $\liminf_{n \rightarrow \infty} x_i$  exists we have

$$\forall n \in \{k, \dots, \infty\} \inf(\{x_i | i \in \{n, \dots, \infty\}\}) \text{ exists} \quad (11.47)$$

Using 9.44 we have then that  $\sup(\{-x_i | i \in \{n, \dots, \infty\}\})$  exists and

$$\forall n \in \mathbb{N}_0 \in \{k, \dots, \infty\} \sup(\{-x_i | i \in \{n, \dots, \infty\}\}) = -\inf(\{x_i | i \in \{n, \dots, \infty\}\}) \quad (11.48)$$

Again using the existence of  $\liminf_{n \rightarrow \infty} x_i$  we have the existence of  $\liminf_{n \rightarrow \infty} (-x_i)$ , further we have

$$\begin{aligned} \liminf_{n \rightarrow \infty} (\{x_i | i \in \{n, \dots, \infty\}\}) &\stackrel{11.48}{=} \lim_{n \rightarrow \infty} (-\sup(\{-x_i | i \in \{n, \dots, \infty\}\})) \\ &\stackrel{11.353}{=} -\lim_{n \rightarrow \infty} (\sup(\{-x_i | i \in \{n, \dots, \infty\}\})) \\ &\stackrel{\text{def}}{=} -\limsup_{n \rightarrow \infty} (-x_i) \end{aligned}$$

proving (1).

2. As  $\limsup_{n \rightarrow \infty} x_i$  exists we have

$$\forall n \in \{k, \dots, \infty\} \sup(\{x_i | i \in \{n, \dots, \infty\}\}) \text{ exists} \quad (11.49)$$

Using 9.44 we have then that  $\inf(\{-x_i | i \in \{n, \dots, \infty\}\})$  exists and

$$\forall n \in \mathbb{N}_0 \in \{k, \dots, \infty\} \sup(\{-x_i | i \in \{n, \dots, \infty\}\}) = -\inf(\{x_i | i \in \{n, \dots, \infty\}\}) \quad (11.50)$$

Again using the existence of  $\limsup_{n \rightarrow \infty} x_i$  we have the existence of  $\limsup_{n \rightarrow \infty} (-x_i)$ , further we have

$$\begin{aligned} \limsup_{n \rightarrow \infty} (\{x_i | i \in \{n, \dots, \infty\}\}) &\stackrel{11.50}{=} \lim_{n \rightarrow \infty} (-\inf(\{-x_i | i \in \{n, \dots, \infty\}\})) \\ &\stackrel{11.353}{=} -\lim_{n \rightarrow \infty} (\inf(\{-x_i | i \in \{n, \dots, \infty\}\})) \\ &\stackrel{\text{def}}{=} -\liminf_{n \rightarrow \infty} (-x_i) \end{aligned}$$

proving (1).

3. As  $\limsup_{n \rightarrow \infty} x_i$  exists we have

$$\forall n \in \{k, \dots, \infty\} \sup(\{x_i | i \in \{n, \dots, \infty\}\}) \text{ exists}$$

Using 9.45 we have that  $\sup(\{\alpha \cdot x_i | i \in \{n, \dots, \infty\}\})$  exists and is equal to  $\alpha \cdot \sup(\{x_i | i \in \{n, \dots, \infty\}\})$ . So as  $\limsup_{n \rightarrow \infty}(\{x_i | i \in \{n, \dots, \infty\}\})$  exists we have by 11.353 that  $\lim_{n \rightarrow \infty}(\alpha \cdot \sup(\{x_i | i \in \{n, \dots, \infty\}\}))$  exists. Hence using 11.353 we have

$$\begin{aligned}\alpha \cdot \limsup_{n \rightarrow \infty} x_n &= \alpha \cdot \limsup_{n \rightarrow \infty}(\{\alpha \cdot x_i | i \in \{n, \dots, \infty\}\}) \\ &\stackrel{11.353}{=} \lim_{n \rightarrow \infty}(\alpha \cdot \sup(\{x_i | i \in \{n, \dots, \infty\}\})) \\ &= \lim_{n \rightarrow \infty}(\sup(\{\alpha \cdot x_i | i \in \{n, \dots, \infty\}\})) \\ &= \limsup_{n \rightarrow \infty}(\alpha \cdot x_n)\end{aligned}$$

4. As  $\liminf_{n \rightarrow \infty} x_i$  exists we have

$$\forall n \in \{k, \dots, \infty\} \inf(\{x_i | i \in \{n, \dots, \infty\}\}) \text{ exists}$$

Using 9.45 we have that  $\inf(\{\alpha \cdot x_i | i \in \{n, \dots, \infty\}\})$  exists and is equal to  $\alpha \cdot \inf(\{x_i | i \in \{n, \dots, \infty\}\})$ . So as  $\liminf_{n \rightarrow \infty}(\{x_i | i \in \{n, \dots, \infty\}\})$  exists we have by 11.353 that  $\lim_{n \rightarrow \infty}(\alpha \cdot \inf(\{x_i | i \in \{n, \dots, \infty\}\}))$  exists. Hence using 11.353 we have

$$\begin{aligned}\alpha \cdot \liminf_{n \rightarrow \infty} x_n &= \alpha \cdot \limsup_{n \rightarrow \infty}(\{\alpha \cdot x_i | i \in \{n, \dots, \infty\}\}) \\ &\stackrel{11.353}{=} \lim_{n \rightarrow \infty}(\alpha \cdot \sup(\{x_i | i \in \{n, \dots, \infty\}\})) \\ &= \lim_{n \rightarrow \infty}(\sup(\{\alpha \cdot x_i | i \in \{n, \dots, \infty\}\})) \\ &= \limsup_{n \rightarrow \infty}(\alpha \cdot x_n)\end{aligned}$$

5. As  $\liminf_{n \rightarrow \infty} x_i$  exists we have that  $\forall n \in \{k, \dots, \infty\}$ . Using 9.47 we have that  $\sup(\{x_i + x | i \in \{n, \dots, \infty\}\})$  exists and  $\sup(\{x_i + x | i \in \{n, \dots, \infty\}\}) = \sup(\{x_i | i \in \{n, \dots, \infty\}\}) + x$ . Hence using 11.352 we have

$$\begin{aligned}\lim_{n \rightarrow \infty}(\sup(\{x_i + x | i \in \{n, \dots, \infty\}\})) &= \lim_{n \rightarrow \infty}(\sup(\{x_i | i \in \{n, \dots, \infty\}\}) + x) \\ &\stackrel{11.352}{=} \limsup_{n \rightarrow \infty}(\{x_i | i \in \{n, \dots, \infty\}\}) + x \\ &= \limsup_{n \rightarrow \infty} x_i + x\end{aligned}$$

6. As  $\liminf_{n \rightarrow \infty} x_i$  exists we have that  $\forall n \in \{k, \dots, \infty\}$ . Using 9.49 we have that  $\inf(\{x_i + x | i \in \{n, \dots, \infty\}\})$  exists and  $\inf(\{x_i + x | i \in \{n, \dots, \infty\}\}) = \inf(\{x_i | i \in \{n, \dots, \infty\}\}) + x$ . Hence using 11.352 we have

$$\begin{aligned}\lim_{n \rightarrow \infty}(\inf(\{x_i + x | i \in \{n, \dots, \infty\}\})) &= \lim_{n \rightarrow \infty}(\inf(\{x_i | i \in \{n, \dots, \infty\}\}) + x) \\ &\stackrel{11.352}{=} \liminf_{n \rightarrow \infty}(\{x_i | i \in \{n, \dots, \infty\}\}) + x \\ &= \liminf_{n \rightarrow \infty} x_i + x\end{aligned}$$

□

### 11.11.5 Series

The idea of series is to introduce the concept of infinite sums. Note that, in contrast with finite sums, infinite sums in general do not have the properties of commutativity and associativity. We will define infinite sums as a limit of finite sums. We use the following definitions:

**Definition 11.375. (Serie)** Let  $\langle X, \|\cdot\| \rangle$  be a normed space and  $\{x_i\}_{i \in \{k, \dots, \infty\}} \subseteq X$  a sequence then we note:

1.  $\sum_{i=k}^{\infty} x_i$  is another notation for  $\{x_i\}_{i \in \{k, \dots, \infty\}}$
2.  $\sum_{i=k}^{\infty} x_i$  is convergent if  $\lim_{n \rightarrow \infty}(\sum_{i=k}^n x_i)$  exists.
3. If  $\sum_{i=k}^{\infty} x_i$  converges then we note  $\lim_{n \rightarrow \infty}(\sum_{i=k}^n x_i)$  by  $\sum_{i=k}^{\infty} x_i$

**Theorem 11.376.** Let  $\langle X, \|\cdot\| \rangle$  be a normed space then if  $\sum_{i=k}^{\infty} x_i$  is convergent it follows that  $\lim_{n \rightarrow \infty} x_n = 0$  and  $\exists K \in \mathbb{R}_+$  such that  $\forall i \in \{k, \dots, \infty\}$  we have  $\|x_i\| \leq K$

**Proof.** Given  $\varepsilon > 0$  then because  $\{\sum_{i=k}^n x_i\}_{n \in \{k, \dots, \infty\}}$  has a limit it must be Cauchy so there exists a  $N \in \{k, \dots, \infty\}$  such that if  $n \geq N + 1$  and thus  $n, n - 1 \geq N$  then  $\|x_n - 0\| = \|x_n\| = \|\sum_{i=k}^n x_i - \sum_{i=k}^{n-1} x_i\| < \varepsilon$  proving that  $\lim_{n \rightarrow \infty} x_n = 0$ . Further as  $\lim_{n \rightarrow \infty} x_i$  exists we have by 11.321 that  $\{x_i\}_{i \in \{k, \dots, \infty\}}$  is Cauchy hence using 11.335 there exists a  $K \in \mathbb{R}_+$  such that  $\forall i \in \{k, \dots, \infty\}$  we have  $\|x_i\| \leq K$ .  $\square$

**Theorem 11.377.** Let  $\langle X, \|\cdot\| \rangle$  be a Banach space  $\sum_{i=k}^{\infty} x_i$  converge if and only if  $\forall \varepsilon > 0$  there exists a  $N \in \{k, \dots, \infty\}$  such that  $\forall n, m \geq N \vdash n \leq m$  we have  $\|\sum_{i=n}^m x_i\| < \varepsilon$

**Proof.** Take  $\{s_i\}_{i \in \{k, \dots, \infty\}}$  defined by  $s_i = \sum_{j=k}^i x_j$  then

$\Rightarrow$ . If  $\sum_{i=k}^{\infty} x_i$  converges we have that  $\lim_{i \rightarrow \infty} s_i$  exists, take  $\varepsilon > 0$ , then using 11.321 there exists a  $N' \in \{k, \dots, \infty\}$  such that  $\forall n, m \geq N'$  we have  $\|s_m - s_n\| < \varepsilon$ . Take  $n, m \geq N = N' + 1$  with  $n \leq m$  then  $N' \leq n - 1, m$  and  $\|\sum_{i=n}^m x_i\| \stackrel{10.17}{=} \|\sum_{i=k}^m x_i - \sum_{i=k}^{n-1} x_i\| = \|s_m - s_n\| \leq \varepsilon$

$\Leftarrow$ . Let  $\varepsilon > 0$  then there exists a  $N \in \{k, \dots, \infty\}$  such that  $\forall n, m \geq N$  with  $n \leq m$  we have  $\|\sum_{i=n}^m x_i\| < \varepsilon$ , as  $\sum_{i=n+1}^m x_i = \sum_{i=k}^m x_i - \sum_{i=k}^n x_i = s_m - s_n$  we conclude that  $\|s_m - s_n\| = \|\sum_{i=n+1}^m x_i\| < \varepsilon$ . From this it follows by definition that  $\{s_i\}_{i \in \{k, \dots, \infty\}}$  is Cauchy and as  $X$  is a Banach space and thus complete,  $\lim_{i \rightarrow \infty} s_i$  exists proving that  $\sum_{i=k}^{\infty} x_i$  converges.  $\square$

**Theorem 11.378.** Let  $\langle X, \|\cdot\| \rangle$  be a normed space,  $i \in \mathbb{N}$ ,  $\{x_i\}_{i \in \{k+l, \dots, \infty\}}$  then  $\sum_{i=k}^{\infty} x_{i+l}$  is convergent if and only if  $\sum_{i=k+l}^{\infty} x_i$  is convergent. Further  $\sum_{i=k}^{\infty} x_{i+l} = \sum_{i=k+l}^{\infty} x_i$

**Proof.**

$\Rightarrow$ . Suppose  $\sum_{i=k}^{\infty} x_{i+l}$  converges to  $x$ . Let  $\varepsilon > 0$  then there exists a  $N' \in \{k, \dots, \infty\}$  such that  $\forall n \geq N'$  we have  $\|\sum_{i=k}^n x_{i+l} - x\| < \varepsilon$ . Take then  $N = N' + l$  then if  $n \geq N = N' + l$  we have that  $n - l \geq N'$ . Now  $\|\sum_{i=k+l}^n x_i - x\| = \|\sum_{i=k}^{n-l} x_{i+l} - x\| < \varepsilon$  (as  $n - l \geq N'$ ) proving that  $\sum_{i=k+l}^{\infty} x_i$  converges and  $\sum_{i=k+l}^{\infty} x_i = \sum_{i=k}^{\infty} x_{i+l}$

$\Leftarrow$ . Suppose  $\sum_{i=k+l}^{\infty} x_i$  converges to  $x$ . Let  $\varepsilon > 0$  then there exists a  $N' \in \{k+l, \dots, \infty\}$  such that  $\forall n \geq N'$  we have  $\|\sum_{i=k+l}^n x_i - x\| < \varepsilon$ . Take  $N = N' - l$  then if  $n \geq N$  we have that  $n + l \geq N'$ . Now  $\|\sum_{i=k}^n x_{i+l} - x\| = \|\sum_{i=k+l}^{n+l} x_i - x\| < \varepsilon$  (as  $n + l \geq N'$ ) proving that  $\sum_{i=k}^{\infty} x_i$  converges and  $\sum_{i=k}^{\infty} x_i = \sum_{i=k+l}^{\infty} x_i$ .  $\square$

**Theorem 11.379.** Let  $\langle X, \|\cdot\| \rangle$  be a normed space and  $\{x_n\}_{n \in \{k, \dots, \infty\}}$  a sequence and  $m \in \{k, \dots, \infty\}$  then  $\sum_{i=k}^{\infty} x_i$  is convergent if and only if  $\sum_{i=m+1}^{\infty} x_i$  is a convergent series (so  $\forall m \in \{k, \dots, \infty\}$   $\sum_{i=m}^{\infty} x_i$  converges). Also if any of the two series is convergent then we have  $\sum_{i=k}^{\infty} x_i = (\sum_{i=k}^m x_i) + (\sum_{i=m+1}^{\infty} x_i)$

**Proof.**

1. ( $\Rightarrow$ ) Given  $\varepsilon > 0$  then  $\exists N \in \{k, \dots, \infty\}$  such that if  $n \geq N$  we have  $\|\sum_{i=k}^n x_i - \sum_{i=k}^{\infty} x_i\| < \varepsilon$ . Now take  $N' = N + (m - k) + 1 \in \{m + 1, \dots, \infty\}$  [as  $N \geq k$ ] and if  $n \geq N' \geq N$  [as  $m \geq k$ ] then  $\|\sum_{i=m+1}^n x_i - (\sum_{i=k}^{\infty} x_i - \sum_{i=k}^m x_i)\| = \|\sum_{i=m+1}^n x_i + \sum_{i=k}^m x_i - \sum_{i=k}^{\infty} x_i\| = \|\sum_{i=k}^n x_i - \sum_{i=k}^{\infty} x_i\| < \varepsilon$  proving that  $\sum_{i=m+1}^{\infty} x_i$  exists and is equal to  $\sum_{i=k}^{\infty} x_i - \sum_{i=k}^m x_i$

2. ( $\Leftarrow$ ) Given  $\varepsilon > 0$  then  $\exists N \in \{m + 1, \dots, \infty\}$  such that if  $n \geq N$  we have  $\|\sum_{i=m+1}^n x_i - \sum_{i=m+1}^{\infty} x_i\| < \varepsilon$ . Then also  $N \in \{k, \dots, \infty\}$  and  $\|\sum_{i=k}^n x_i - (\sum_{i=k}^m x_i + \sum_{i=m+1}^{\infty} x_i)\| = \|\sum_{i=k}^n x_i - \sum_{i=m+1}^{\infty} x_i\| < \varepsilon$  proving that  $\sum_{i=k}^{\infty} x_i$  exist and is equal to  $\sum_{i=k}^m x_i + \sum_{i=m+1}^{\infty} x_i$ .  $\square$

**Theorem 11.380.** Let  $\langle X, \|\cdot\| \rangle$  be a normed space and  $\sum_{i=k}^{\infty} x_i$  convergent serie then

$$\lim_{N \rightarrow \infty} \sum_{i=N}^{\infty} x_i = 0$$

**Proof.** Let  $\varepsilon > 0$  then as  $\sum_{i=k}^{\infty} x_i$  converges there exists a  $N \geq k$  so that  $\forall m-1 \geq N$  we have  $\|\sum_{i=k}^{\infty} x_i - \sum_{i=k}^{m-1} x_i\| < \varepsilon$  then  $\|\sum_{i=m}^{\infty} x_i - 0\| = \|\sum_{i=m}^{\infty} x_i\| \stackrel{\text{see previous theorem 11.379}}{=} \|\sum_{i=k}^{\infty} x_i - \sum_{i=k}^{m-1} x_i\| < \varepsilon$  proving our theorem.  $\square$

**Theorem 11.381.** Let  $\langle X, \|\cdot\| \rangle$  be a normed space over the field  $\mathbb{K}$ ,  $\alpha, \beta \in \mathbb{K}$  and  $\sum_{i=k}^{\infty} x_i$ ,  $\sum_{i=k}^{\infty} y_i$  convergent series then  $\sum_{i=k}^{\infty} (\alpha \cdot x_i + \beta \cdot y_i)$  is convergent and  $\sum_{i=k}^{\infty} (\alpha \cdot x_i + \beta \cdot y_i) = \alpha \cdot \sum_{i=k}^{\infty} x_i + \beta \cdot \sum_{i=k}^{\infty} y_i$ .

**Proof.** Given  $n \in \{k, \dots, \infty\}$  we have that  $\sum_{i=k}^n (\alpha \cdot x_i + \beta \cdot y_i) \stackrel{10.58, 10.52}{=} \alpha \cdot \sum_{i=k}^n x_i + \beta \cdot \sum_{i=k}^n y_i$ , using 11.353 we have then that  $\lim_{n \rightarrow \infty} \sum_{i=k}^n (\alpha \cdot x_i + \beta \cdot y_i)$  exists and is equality to  $\alpha \cdot \lim_{n \rightarrow \infty} \sum_{i=k}^n x_i + \beta \cdot \lim_{n \rightarrow \infty} \sum_{i=k}^n y_i$ .  $\square$

### 11.11.6 Series of positive numbers

Series are in general not commutative but series of positive numbers (and later absolute convergent series) are commutative.

**Lemma 11.382.** Let  $\sum_{i=k}^{\infty} x_i$  be such that  $\forall i \in \{k, \dots, \infty\}$  we have  $x_i \in [0, \dots, \infty[$  then the following are equivalent

1.  $\sum_{i=k}^{\infty} x_i$  is convergent and  $\sum_{i=k}^{\infty} x_i = x$
2.  $\sup(\{\sum_{i=k}^n x_i \mid n \in \{k, \dots, \infty\}\})$  exists and  $x = \sup(\{\sum_{i=k}^{\infty} x_i \mid n \in \{k, \dots, \infty\}\})$
3.  $\sup(\{\sum_{n \in K} x_i \mid K \subseteq \mathcal{P}(\{k, \dots, \infty\}) \wedge K \text{ is finite}\})$  exists and  $x = \sup(\{\sum_{n \in K} x_i \mid K \subseteq \mathcal{P}(\{k, \dots, \infty\}) \wedge K \text{ is finite}\})$
4.  $\forall \varepsilon > 0$  there exists a  $n \in \{k, \dots, \infty\}$  such that  $\forall p \in \mathbb{N}_0$  and  $l \geq n$  we have  $\sum_{i=l+1}^{l+p} x_i < \varepsilon$
5.  $\forall \varepsilon > 0$  there exists a finite  $K \subseteq \{k, \dots, \infty\}$  such that  $\forall H \subseteq \{k, \dots, \infty\}$  with  $H$  finite and  $K \cap H = \emptyset$  we have  $\sum_{i \in H} x_i < \varepsilon$

**Proof.** First we have that If  $s \in \{\sum_{i=k}^n x_i \mid n \in \{k, \dots, \infty\}\}$  then  $\exists n \in \{k, \dots, \infty\}$  so that  $s = \sum_{i=k}^n x_i = \sum_{i \in \{k, \dots, n\}} x_i \in \{\sum_{i \in K} x_i \mid K \subseteq \mathcal{P}(\{k, \dots, \infty\}) \wedge K \text{ is finite}\}$  proving that

$$\left\{ \sum_{i=k}^n x_i \mid n \in \{k, \dots, \infty\} \right\} \in \left\{ \sum_{i \in K} x_i \mid K \subseteq \mathcal{P}(\{k, \dots, \infty\}) \wedge K \text{ is finite} \right\} \quad (11.51)$$

**1  $\Leftrightarrow$  2.** Given  $n \in \{k, \dots, \infty\}$  we have that  $\sum_{i=k}^{n+1} x_i = x_{n+1} + \sum_{i=k}^n x_i \geq \sum_{i=k}^n x_i$  proves that  $\{\sum_{i=k}^n x_i\}_{n \in \{k, \dots, \infty\}}$  is a increasing sequence. the previous lemma (see 11.366) proves then that **1  $\Leftrightarrow$  2**

**2  $\Rightarrow$  3.** Take  $s \in \{\sum_{n \in K} x_i \mid K \subseteq \mathcal{P}(\{k, \dots, \infty\}) \wedge K \text{ is finite}\}$  then there exists a finite  $K \subseteq \{k, \dots, \infty\}$  so that  $s = \sum_{i \in K} x_i$ . Take then  $n = \max(K)$  then as  $i \in K \Rightarrow k \leq i \leq n = \{k, \dots, n\}$  proving that  $K \subseteq \{k, \dots, n\}$ . So  $s = \sum_{i \in K} x_i \leq \sum_{i \in K} x_i + \sum_{i \in \{1, \dots, n\} \setminus K} x_i = \sum_{i \in \{1, \dots, n\}} x_i = \sum_{i=1}^n x_i \leq \sup(\{\sum_{i=k}^n x_i \mid n \in \{k, \dots, \infty\}\})$  proving by 9.43 that

$$\sup \left( \left\{ \sum_{i \in K} x_i \mid K \subseteq \mathcal{P}(\{k, \dots, \infty\}) \wedge K \text{ is finite} \right\} \right) \text{ exists} \quad (11.52)$$

exists and that

$$\sup \left( \left\{ \sum_{i \in K} x_i \mid K \subseteq \mathcal{P}(\{k, \dots, \infty\}) \wedge K \text{ is finite} \right\} \right) \leq \sup \left( \left\{ \sum_{i=k}^n x_i \mid n \in \{k, \dots, \infty\} \right\} \right) \quad (11.53)$$

Using 11.51 we have that  $\sup(\{\sum_{i=k}^n x_i | n \in \{k, \dots, \infty\}\}) \leq \sup(\{\sum_{i \in K} x_i | K \subseteq \mathcal{P}(\{k, \dots, \infty\}) \wedge K \text{ is finite}\})$  which together with 11.53 proves that

$$\sup\left(\left\{\sum_{i \in K} x_i | K \in \mathcal{P}(\{k, \dots, \infty\}) \wedge K \text{ is finite}\right\}\right) = \sup\left(\left\{\sum_{i=k}^n x_i | n \in \{k, \dots, \infty\}\right\}\right)$$

**3  $\Rightarrow$  2.** Let  $s \in \{\sum_{i=k}^n x_i | n \in \{k, \dots, \infty\}\}$  then from 11.51 we have that  $s \in \{\sum_{i \in K} x_i | K \subseteq \mathcal{P}(\{k, \dots, \infty\}) \wedge K \text{ is finite}\}$  so that  $s \leq \sup(\{\sum_{i \in K} x_i | K \subseteq \mathcal{P}(\{k, \dots, \infty\}) \wedge K \text{ is finite}\})$  hence by 9.43 we have that

$$\sup\left(\left\{\sum_{i=k}^n x_i | n \in \{k, \dots, \infty\}\right\}\right) \text{ exists}$$

Applying then (2) proves the rest.

**1  $\Rightarrow$  2.** As  $\sum_{i=k}^{\infty} x_i$  converges we have that  $\lim_{n \rightarrow \infty} (\sum_{i=k}^n x_i)$  exists so that  $\{\sum_{i=k}^n x_i\}_{n \in \{k, \dots, \infty\}}$  is Cauchy (see 11.321) hence there exists a  $n \in \{k, \dots, \infty\}$  such that  $\forall r, s \geq n$  we have  $|\sum_{i=k}^r x_i - \sum_{i=k}^s x_i| < \varepsilon$ . Then if  $l \geq n$  and  $p \in \mathbb{N}_0$  we have  $n \leq l \leq l+p$  so that  $\sum_{i=l+1}^{l+p} x_i = \sum_{i=k}^{l+p} x_i - \sum_{i=k}^l x_i \leq \sum_{i=k}^{l+p} x_i \leq \sum_{i=k}^r x_i \leq \sum_{i=k}^{l+p} x_i$  so that  $|\sum_{i=k}^{l+p} x_i - \sum_{i=k}^l x_i| < \varepsilon$

**2  $\Rightarrow$  1.** Take  $\varepsilon > 0$  then there exists a  $n \in \{k, \dots, \infty\}$  such that  $\forall l \geq n$  and  $p \in \mathbb{N}_0$  we have that  $\sum_{i=l+1}^{l+p} x_i < \varepsilon$ . Take now  $r, s \geq n$  then we may assume without lossing generality that  $r \geq s$ . If now  $r = s$  then  $|\sum_{i=k}^r x_i - \sum_{i=k}^s x_i| = 0 < \varepsilon$  and if  $s < r$  then  $|\sum_{i=k}^r x_i - \sum_{i=k}^s x_i| = |\sum_{i=s+1}^r x_i| = \sum_{i=s+1}^r x_i \leq \sum_{i=s+1}^{s+p} x_i < \varepsilon$  proving that  $\{\sum_{i=k}^n x_i\}_{n \in \{k, \dots, \infty\}}$  is Cauchy. As  $\mathbb{R}$  is complete  $\{\sum_{i=k}^n x_i\}_{n \in \{k, \dots, \infty\}}$  converges hence  $\sum_{i=k}^{\infty} x_i$  converges

**4  $\Rightarrow$  5.** Let  $\varepsilon > 0$  then there exists a  $n \in \{k, \dots, \infty\}$  such that  $\forall p \in \mathbb{N}_0$  and  $l \geq n$  we have  $\sum_{i=l+1}^{l+p} x_i < \varepsilon$ . Take  $K = \{k, \dots, n\}$  then if  $H \subseteq \{k, \dots, \infty\}$  is finite with  $H \cap K = \emptyset$  then we have either

$$K = \emptyset. \text{ then } \sum_{i \in K} x_i = 0 < \varepsilon$$

$$K \neq \emptyset. \text{ then } n < \min(K) \leq \max(K) \text{ so that } K \subseteq \{n+1, \dots, n+p\} \text{ where } p = \max(K) - n \in \mathbb{N}_0 \text{ so that } \sum_{i \in K} x_i \leq \sum_{i \in K} x_i + \sum_{i \in \{n+1, \dots, n+p\}} x_i = \sum_{i=n+1}^{n+p} x_i < \varepsilon$$

proving in all cases  $\sum_{i \in K} x_i < \varepsilon$

**5  $\Rightarrow$  4.** Let  $\varepsilon > 0$  then there exists a finite  $K \subseteq \{k, \dots, \infty\}$  such that  $\forall H \subseteq \{k, \dots, \infty\}$  with  $H$  finite and  $K \cap H = \emptyset$  we have  $\sum_{i \in H} x_i < \varepsilon$ . Take  $K' = K \cup \{k\}$  then  $\emptyset \neq K'$  so that  $n = \max(K')$  exists. If now  $n \leq l$  and  $p \in \mathbb{N}_0$  then  $K \cap \{l+1, \dots, l+p\} \subseteq K' \cap \{l+1, \dots, l+p\} = \emptyset$  hence  $\sum_{i=l+1}^{l+p} x_i = \sum_{i \in \{l+1, \dots, l+p\}} x_i < \varepsilon$   $\square$

**Corollary 11.383.** Let  $\sum_{i=k}^{\infty} x_i$  be convergent serie such that  $\forall i \in \{k, \dots, \infty\}$  we have  $0 < x_i$  then  $0 < \sum_{i=k}^{\infty} x_i$

**Proof.**  $0 < x_k \leq \sup(\{\sum_{i=k}^n x_i | n \in \{k, \dots, \infty\}\}) = \sum_{i=k}^{\infty} x_i$   $\square$

The above equivalences can be used to prove the commutativity of series of positive numbers

**Theorem 11.384.** Let  $\{x_i\}_{i \in \{k, \dots, \infty\}} \subseteq [0, \dots, \infty[$  and  $\beta: \{k, \dots, \infty\} \rightarrow \{k, \dots, \infty\}$  a bijection then if  $\sum_{i=k}^{\infty} x_i$  exists we have that  $\sum_{i=k}^{\infty} x_{\beta(i)}$  exists and  $\sum_{i=k}^{\infty} x_i = \sum_{i=k}^{\infty} x_{\beta(i)}$

**Proof.** If  $s \in \{\sum_{i \in K} x_i | K \subseteq \mathcal{P}(\{k, \dots, \infty\}) \wedge K \text{ is finite}\}$  then there exists a  $K \subseteq \{k, \dots, \infty\}$  such that  $s = \sum_{i \in K} x_i$ . As  $(\beta^{-1})|_{\beta(K)}: \beta(K) \rightarrow K$  is a bijection we have using 10.43 that  $\sum_{i \in K} x_i = \sum_{i \in \beta^{-1}(K)} x_{\beta(i)} \in \{\sum_{i \in K} x_{\beta(i)} | K \subseteq \mathcal{P}(\{k, \dots, \infty\}) \wedge K \text{ is finite}\}$  [As  $\beta^{-1}(K) \subseteq \{k, \dots, \infty\}$  is finite] so we have that

$$\left\{ \sum_{i \in K} x_i | K \subseteq \mathcal{P}(\{k, \dots, \infty\}) \wedge K \text{ is finite} \right\} \subseteq \left\{ \sum_{i \in K} x_{\beta(i)} | K \subseteq \mathcal{P}(\{k, \dots, \infty\}) \wedge K \text{ is finite} \right\} \quad (11.54)$$

If  $s \in \{\sum_{i \in K} x_{\beta(i)} | K \subseteq \mathcal{P}(\{k, \dots, \infty\}) \wedge K \text{ is finite}\}$  then there exists a  $K \subseteq \{k, \dots, \infty\}$  such that  $s = \sum_{i \in K} x_{\beta(i)}$  then as  $\beta_K: K \rightarrow \beta(K)$  is a bijection we can use 10.43 again giving  $\sum_{i \in K} x_{\beta(i)} = \sum_{i \in \beta(K)} x_i \in \{\sum_{i \in K} x_i | K \subseteq \mathcal{P}(\{k, \dots, \infty\}) \wedge K \text{ is finite}\}$  [as  $\beta(K) \subseteq \{k, \dots, \infty\}$  is finite]. So we have proved that  $\{\sum_{i \in K} x_{\beta(i)} | K \subseteq \mathcal{P}(\{k, \dots, \infty\}) \wedge K \text{ is finite}\} \subseteq \{\sum_{i \in K} x_i | K \subseteq \mathcal{P}(\{k, \dots, \infty\}) \wedge K \text{ is finite}\}$  which using 11.54 gives

$$\left\{ \sum_{i \in K} x_i | K \subseteq \mathcal{P}(\{k, \dots, \infty\}) \wedge K \text{ is finite} \right\} \subseteq \left\{ \sum_{i \in K} x_{\beta(i)} | K \subseteq \mathcal{P}(\{k, \dots, \infty\}) \wedge K \text{ is finite} \right\} \quad (11.55)$$

As  $x = \sum_{i=k}^{\infty} x_i$  exists we have using 11.382 that  $x = \sup(\{\sum_{i \in K} x_i | K \subseteq \mathcal{P}(\{k, \dots, \infty\}) \wedge K \text{ is finite}\}) = \sup(\{\sum_{i \in K} x_{\beta(i)} | K \subseteq \mathcal{P}(\{k, \dots, \infty\}) \wedge K \text{ is finite}\}) = \sum_{i=k}^{\infty} x_{\beta(i)}$   $\square$

The above theorem allows us to extend the definition of a serie of positive numbers. Later we will use this extension also for absolute convergent series.

**Definition 11.385.** Let  $I$  be a countable set,  $\{x_i\}_{i \in I} \subseteq [0, \infty[$  a countable family of positive numbers then we define convergence of  $\sum_{i \in I} x_i$  as follows

1. If  $I$  is finite then  $\sum_{i \in I} x_i$  converges and  $\sum_{i \in I} x_i$  is as defined in 10.36). Using 10.39 we have then that  $\sum_{i \in I} x_i = \sum_{i=0}^n x_{\beta(i)}$  where  $\beta: \{0, \dots, n\} \rightarrow I$  is a bijection.
2. If  $I$  is infinite then  $\sum_{i \in I} x_i$  converges if  $\sum_{i=0}^{\infty} x_{\beta(i)}$  converges where  $\beta: \mathbb{N} \rightarrow I$  is a bijection (note that as  $I$  is infinite and countable such a bijection exists). We denote then  $\sum_{i \in I} x_i$  by  $\sum_{i=0}^{\infty} x_{\beta(i)}$

we must of course prove that this definition is independent of the choice of  $\beta$

**Proof.**

1. This follows from 10.39
2. Let  $\alpha: \mathbb{N} \rightarrow I$  be another bijection then  $\alpha^{-1} \circ \beta: \mathbb{N} \rightarrow \mathbb{N}$  is a bijection. So using 11.384 we have that  $\sum_{i=0}^{\infty} x_{\alpha(\alpha^{-1}(\beta(i)))}$  is convergent and  $\sum_{i=0}^{\infty} x_{\alpha(i)} = \sum_{i=0}^{\infty} x_{\alpha(\alpha^{-1}(\beta(i)))} = \sum_{i=0}^{\infty} x_{\beta(i)} = s$   $\square$

**Lemma 11.386.** Let  $I, J$  be countable sets,  $I \subseteq J$  and  $\{x_i\}_{i \in J} \subseteq [0, \infty[$  such that  $\sum_{i \in J} x_i$  exists then we have that  $\sum_{i \in I} x_i$  and  $\sum_{i \in I} x_i \leq \sum_{i \in J} x_i$

**Proof.** We must consider the following cases for  $J$

**$J$  is finite.** then as  $I \subseteq J$  we have that  $I$  is also finite and  $\sum_{i \in I} x_i = \sum_{i \in J \setminus I} x_i + \sum_{i \in I} x_i \geq \sum_{i \in I} x_i$

**$J$  is infinite.** By definition there exists a bijection  $\beta: \mathbb{N} \rightarrow J$  such that  $\sum_{i \in J} x_i = \sum_{i=0}^{\infty} x_{\beta(i)}$ . For  $I$  we have either

**$I$  is finite.** As  $I$  is finite and  $\beta$  is a bijection we have that  $\beta^{-1}(I) \subseteq \mathbb{N}$  is finite so that  $\sum_{i \in I} x_i = \sum_{i \in \beta^{-1}(I)} x_{\beta(i)} \leq \sup(\{\sum_{i \in K} x_{\varphi(i)} | K \subseteq \mathbb{N} \wedge K \text{ is finite}\}) = \sum_{i=0}^{\infty} x_{\varphi(i)} = \sum_{i \in J} x_i$  proving that  $\sum_{i \in I} x_i \leq \sum_{i \in J} x_i$

**$J$  is infinite.** Then there exists a bijection  $\alpha: \mathbb{N} \rightarrow I \subseteq J$ . Take a finite  $K \subseteq \mathbb{N}$  then  $\alpha(K) \subseteq J \wedge \alpha(K)$  is finite so that  $\beta^{-1}(\alpha(K))$  is finite. As we have that

$$\begin{aligned} \sum_{i \in K} x_{\alpha(i)} &\stackrel{10.43}{=} \sum_{i \in \alpha(K)} x_i \\ &\stackrel{10.43}{=} \sum_{i \in \beta^{-1}(\alpha(K))} x_{\beta(i)} \\ &\leq \sup \left( \left\{ \sum_{i \in L} x_{\beta(i)} | L \subseteq \mathbb{N} \wedge L \text{ is finite} \right\} \right) \\ &\leq \sum_{i=0}^{\infty} x_{\beta(i)} \end{aligned}$$

Proving by 9.43 that  $\sup(\{\sum_{i \in K} x_{\alpha(i)} | K \subseteq N \wedge K \text{ is finite}\})$  exists and  $\sup(\{\sum_{i \in K} x_{\alpha(i)} | K \subseteq N \wedge K \text{ is finite}\}) \leq \sum_{i=0}^{\infty} x_{\beta(i)}$  proving that  $\sum_{i=0}^{\infty} x_{\alpha(i)}$  exists and  $\sum_{i=0}^{\infty} x_{\alpha(i)} \leq \sum_{i=0}^{\infty} x_{\beta(i)}$ . So by definition we have that  $\sum_{i \in I} x_i$  exists and  $\sum_{i \in I} x_i \leq \sum_{i \in J} x_i$ .  $\square$

**Lemma 11.387.** Let  $I$  be countable set and  $\{x_i\}_{i \in I} \subseteq [0, \infty[$  then we have the following equivalences

1.  $\sum_{i \in I} x_i$  converges to  $s$
2.  $\sup(\{\sum_{i \in K} x_i | K \subseteq I \wedge K \text{ is finite}\})$  exists and  $\sup(\{\sum_{i \in K} x_i | K \subseteq I \wedge K \text{ is finite}\}) = s$

**Proof.** As  $I$  is countable we have two possibilities

**$I$  is finite.** then  $\sum_{i \in I} x_i$  converges by definition and as  $\forall K \subseteq I$  we have  $K$  is finite and  $\sum_{i \in K} x_i \leq \sum_{i \in I} x_i + \sum_{i \in I \setminus K} x_i \stackrel{10.45}{=} \sum_{i \in I} x_i$  so that  $\sum_{i \in I} x_i = \max(\{\sum_{i \in K} x_i | K \subseteq I \wedge K \text{ is finite}\}) = \sup(\{\sum_{i \in K} x_i | K \subseteq I \wedge K \text{ is finite}\})$  (which of course exists)

**$I$  is infinite.** Then there exists a bijection  $\beta: \mathbb{N} \rightarrow I$ . If now  $s \in \{\sum_{i \in K} x_{\beta(i)} | K \subseteq \mathbb{N} \wedge K \text{ is finite}\}$  then  $s = \sum_{i \in K} x_{\beta(i)} \stackrel{10.43}{=} \sum_{i \in \beta(K)} x_i \in \{\sum_{i \in K} x_i | K \subseteq I \wedge K \text{ is finite}\}$ , also if  $s \in \{\sum_{i \in K} x_i | K \subseteq I \wedge K \text{ is finite}\}$  then  $s = \sum_{i \in K} x_i \stackrel{10.43}{=} \sum_{i \in \beta^{-1}(K)} x_{\beta(i)} \in \{\sum_{i \in K} x_{\beta(i)} | K \subseteq \mathbb{N} \wedge K \text{ is finite}\}$  proving that

$$\left\{ \sum_{i \in K} x_{\beta(i)} | K \subseteq \mathbb{N} \wedge K \text{ is finite} \right\} = \left\{ \sum_{i \in K} x_i | K \subseteq I \wedge K \text{ is finite} \right\} \quad (11.56)$$

Now

$$\begin{aligned} \sum_{i \in I} x_i \text{ converges to } s &\Leftrightarrow \text{definition } \sum_{i=0}^{\infty} x_{\beta(i)} \text{ converges to } s \\ &\Leftrightarrow \text{11.382 } \sup_s \left( \left\{ \sum_{i \in K} x_{\beta(i)} | K \subseteq \mathbb{N} \wedge K \text{ is finite} \right\} \right) \text{ exists and is equal to } s \\ &\Leftrightarrow \text{11.56 } \sup_s \left( \left\{ \sum_{i \in K} x_i | K \subseteq I \wedge K \text{ is finite} \right\} \right) \text{ exists and is equal to } s \end{aligned}$$

$\square$

**Example 11.388.** Let  $I$  be a countable set and  $\{x_i\}_{i \in I}$  be such that  $\forall i \in I$  we have  $x_i = 0$  then  $\sum_{i \in I} x_i$  converges to 0  $\square$

**Proof.** If  $K \subseteq I$  finite then we have as  $\forall i \in K$   $x_i = 0$  that  $\sum_{i \in K} x_i \stackrel{10.37}{=} 0$  so that  $\sup(\{\sum_{i \in K} x_i | K \subseteq I \wedge K \text{ is finite}\}) = 0$  which by the previous lemma (see 11.387) proves the example.  $\square$

Next we proceed to prove that series of positive numbers are associative. First we need a little lemma concerning supremums in  $\mathbb{R}$ .

**Lemma 11.389.** Let  $n \in \mathbb{N}$  then if  $\{A_i\}_{i \in \{0, \dots, n\}}$  be such that  $\forall i \in \{0, \dots, n\} \models A_i \subseteq \mathbb{R}$  and  $\exists s \in \mathbb{R}$  so that  $\forall (x_0, \dots, x_n) \in \prod_{i \in \{0, \dots, n\}} A_i$  we have  $\sum_{i=0}^n x_i \leq s$  then  $\forall i \in \{0, \dots, n\} \sup(A_i)$  exists and  $\sum_{i=0}^n \sup(A_i) \leq s$

**Proof.** We prove this by induction so let  $\mathcal{S} = \{n \in \mathbb{N} | \text{if } \{A_i\}_{i \in \{0, \dots, n\}} \subseteq \mathbb{R} \text{ be such that } \exists s \in \mathbb{R} \text{ so that } \forall (x_0, \dots, x_n) \in \prod_{i \in \{0, \dots, n\}} A_i \text{ we have } \sum_{i=0}^n x_i \leq s \text{ then } \forall i \in \{0, \dots, n\} \sup(A_i) \text{ exists and } \sum_{i=0}^n \sup(A_i) \leq s\}$  then

**0  $\in \mathcal{S}$ .** Let  $(x_0) \in \prod_{i \in \{0, \dots, 0\}} A_i$  then  $x_0 = \sum_{i=0}^0 x_i \leq s$  so that  $\sup(A_0)$  exists and  $\sum_{i=0}^0 \sup(A_i) = \sup(A_0) \leq s$

**$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$ .** Let  $\{A_i\}_{i \in \{0, \dots, n+1\}} \subseteq \mathbb{R}$  be such that  $\exists s \in \mathbb{R}$  so that  $\forall (x_0, \dots, x_{n+1}) \in \prod_{i \in \{0, \dots, n+1\}} A_i$  we have  $\sum_{i=0}^{n+1} x_i \leq s$ . Take then  $x \in A_{n+1}$  and  $(x_0, \dots, x_n) \in \prod_{i \in \{0, \dots, n\}} A_i$  then we have as  $(x_0, \dots, x_n, x) \in \prod_{i \in \{0, \dots, n+1\}} A_i$  that  $\sum_{i=0}^n x_i + x = \sum_{i=0}^{n+1} x_i \leq s$  so that  $x \leq s - \sum_{i=0}^n x_i$  hence by 9.43 we have that  $\sup(A_{n+1})$  exists and  $\sup(A_{n+1}) \leq s - \sum_{i=0}^n x_i \Rightarrow \sum_{i=0}^n x_i \leq s - \sup(A_{n+1})$ . As  $n \in \mathcal{S}$  we have then that  $\sum_{i=0}^n \sup(A_i) \leq s - \sup(A_{n+1}) \Rightarrow \sum_{i=0}^n \sup(A_i) + \sup(A_{n+1}) \leq s$  proving that  $\sum_{i=0}^{n+1} \sup(A_i) \leq s$  and hence  $n+1 \in \mathcal{S}$   $\square$

**Lemma 11.390.** Let  $I$  be a countable set,  $\sum_{i \in I} x_i$  a convergent serie with  $\{x_i\}_{i \in I} \subseteq [0, \infty[$ ,  $n \in \mathbb{N}$  and  $\{N_i\}_{i \in \{0, \dots, n\}}$  such that  $\forall i, j \in \{0, \dots, n\} \vdash i \neq j$  we have  $N_i \cap N_j = \emptyset$  and  $I = \bigcup_{i \in \{0, \dots, n\}} N_i$  then  $\forall i \in \{0, \dots, n\}$  we have that  $\forall i \in \{0, \dots, n\} \sum_{j \in N_i} x_i$  converges and  $\sum_{i \in I} x_i = \sum_{i=0}^n (\sum_{j \in N_i} x_j)$

**Proof.** First as  $\forall i \in \{0, \dots, n\} N_i \subseteq I$  it follows from the convergence of  $\sum_{i \in I} x_i$  and lemma 11.386 that

$$\forall i \in \{0, \dots, n\} \models \sum_{i \in N_i} x_i \text{ converges} \quad (11.57)$$

Second as  $\sum_{i \in I} x_i \stackrel{11.387}{=} \sup(\{\sum_{i \in K} x_i \mid K \subseteq I \wedge K \text{ finite}\})$  we have given  $\varepsilon > 0$  that there exists a  $K \subseteq I$  finite so that  $\sum_{i \in I} x_i - \varepsilon < \sum_{i \in K} x_i$  hence

$$\sum_{i \in I} x_i < \sum_{i \in K} x_i + \varepsilon \quad (11.58)$$

Now given  $i \in \{0, \dots, n\}$  we have that  $N_i \cap K$  is finite,  $K = I \cap K = K \cap (\bigcup_{i \in \{0, \dots, n\}} N_i) = \bigcup_{i \in \{0, \dots, n\}} (N_i \cap K)$  and  $\forall i, j \in \{0, \dots, n\}$  we have  $(N_i \cap K) \cap (N_j \cap K) = (N_i \cap N_j) \cap K = \emptyset$ . So

$$\begin{aligned} \sum_{i \in K} x_i &\stackrel{10.45}{=} \sum_{i=0}^n \left( \sum_{j \in N_i \cap K} x_j \right) \\ &\leq \sum_{i=0}^n \left( \sum_{j \in N_i} x_j \right) \text{(see 11.386, 11.57 and } N_i \cap K \subseteq N_i) \end{aligned}$$

Using 11.58 together with the above gives then that  $\sum_{i \in I} x_i < \sum_{i=0}^n (\sum_{j \in N_i} x_j) + \varepsilon$  which as  $\varepsilon > 0$  was arbitrary choosen means by 9.57 that

$$\sum_{i \in I} x_i \leq \sum_{i=0}^n \left( \sum_{j \in N_i} x_j \right) \quad (11.59)$$

Take now  $(s_0, \dots, s_n) \in \prod_{i \in \{0, \dots, n\}} \{\sum_{i \in K} \mid K \subseteq N_i \wedge K \text{ finite}\}$  then there exists a family  $\{K_i\}_{i \in \{0, \dots, n\}}$  of sets with  $\forall i \in \{0, \dots, n\} K_i \subseteq N_i$  (so the sets are finite and pairwise disjoint) and  $s_i = \sum_{j \in K_i} x_j$ . Hence  $\sum_{i=0}^n s_i = \sum_{i=0}^n (\sum_{j \in K_i} x_j) \stackrel{10.45}{=} \sum_{j \in \bigcup_{i \in \{0, \dots, n\}} K_i} x_i \leq \sum_{i \in I} x_i$  (as  $\bigcup_{i \in \{0, \dots, n\}} K_i \subseteq \bigcup_{i \in \{0, \dots, n\}} N_i = I$  together with 11.386). Hence using 11.389 we have then  $\sum_{i=0}^n \sup(\{\sum_{j \in K_i} x_j \mid K_i \subseteq N_i\}) \leq \sum_{i \in I} x_i$ , so using 11.387 we have that

$$\sum_{i=0}^n \left( \sum_{j \in N_i} x_j \right) \leq \sum_{i \in I} x_i \quad (11.60)$$

Using 11.59 and 11.60 we have

$$\sum_{i \in I} x_i = \sum_{i=0}^n \left( \sum_{j \in N_i} x_j \right) \quad \square$$

**Lemma 11.391.** Let  $I$  be a countable set,  $\sum_{i \in I} x_i$  a convergent serie with  $\{x_i\}_{i \in I} \subseteq [0, \infty[$ ,  $\{N_i\}_{i \in \mathbb{N}}$  be such that  $\mathbb{N} = \bigcup_{i \in \mathbb{N}} N_i$  and  $\forall i, j \in \mathbb{N} \vdash i \neq j \ N_i \cap N_j = \emptyset$  then we have that  $\forall i \in \mathbb{N} \sum_{j \in N_i} x_j$  converges and  $\sum_{i=0}^{\infty} (\sum_{j \in N_i} x_j)$  converges to  $\sum_{i \in \mathbb{N}} x_i$

**Proof.** First using 11.386 we have as  $\forall i \in \mathbb{N} \ N_i \subseteq I$  that

$$\forall i \in \mathbb{N} \models \sum_{j \in N_i} x_j \text{ converges} \quad (11.61)$$

Let  $n \in \mathbb{N}$  then  $K_n = \bigcup_{i \in \{0, \dots, n\}} N_i \subseteq I$  so by 11.386 we have that  $\sum_{i \in K_n} x_i$  converges and  $\sum_{i \in K_n} x_i \leq \sum_{i \in I} x_i$ . Further by 5.66 we have that  $K_n$  is countable and as  $\{N_i\}_{i \in \{0, \dots, n\}}$  is pairwise disjoint we can use the previous lemma (see 11.390) giving  $\sum_{i=0}^n (\sum_{j \in N_i} x_j) = \sum_{i \in K_n} x_i \leq \sum_{i \in I} x_i$ . As this is true for every  $n \in \mathbb{N}$  we have by 9.43 that  $\sup(\{\sum_{i=0}^n (\sum_{j \in N_i} x_j) | n \in \mathbb{N}\})$  exists and  $\sup(\{\sum_{i=0}^n (\sum_{j \in N_i} x_j) | n \in \mathbb{N}\}) \leq \sum_{i \in I} x_i$ . Using then 11.382 we conclude that

$$\sum_{i=0}^{\infty} \left( \sum_{j \in N_i} x_j \right) \text{ converges and } \sum_{i=0}^{\infty} \left( \sum_{j \in N_i} x_j \right) \leq \sum_{i \in I} x_i \quad (11.62)$$

For the opposite equality, let  $\varepsilon > 0$  then as by 11.387 we have that  $\sum_{i \in I} x_i = \sup(\{\sum_{i \in K} x_i | K \subseteq I \wedge K \text{ is finite}\})$  there exists a finite  $K$  such that  $\sum_{i \in I} x_i - \varepsilon < \sum_{i \in K} x_i$  hence

$$\sum_{i \in I} x_i < \sum_{i \in K} x_i + \varepsilon \quad (11.63)$$

Take now  $J = \{i \in \mathbb{N} | K \cap N_i \neq \emptyset\} \subseteq I$  so that by 5.66  $J$  is countable. Then  $\bigcup_{i \in J} (K \cap N_i) \subseteq K$  and if  $x \in K = K \cap \mathbb{N} = K \cap (\bigcup_{i \in \mathbb{N}} N_i) = \bigcup_{i \in \mathbb{N}} (K \cap N_i)$  there exists a  $i \in \mathbb{N}$  such that  $x \in K \cap N_i \Rightarrow K \cap N_i \neq \emptyset$  hence  $x \in K \cap N_i$  where  $i \in J$ . So we have that

$$K = \bigcup_{i \in J} (K \cap N_i) \quad (11.64)$$

Now if  $J$  is infinite then as  $\{K \cap N_i\}_{i \in J}$  is a family of pairwise disjoint non empty sets we would by 5.73 have that  $K$  is infinite a contradiction. So  $J$  is finite, then we have either

$J = \emptyset$ . then if  $\forall i \in \mathbb{N}$  we have  $K \cap N_i = \emptyset \Rightarrow \emptyset = \bigcup_{i \in \mathbb{N}} (K \cap N_i) = K \cap (\bigcup_{i \in \mathbb{N}} N_i) = K \cap \mathbb{N} = K$  so that  $\sum_{i \in I} x_i < 0 + \varepsilon$  which as  $\varepsilon$  is chosen arbitrary means  $\sum_{i \in I} x_i \leq 0 \leq \sum_{i=0}^{\infty} (\sum_{j \in N_i} x_j)$

$J \neq \emptyset$ . then as  $J$  is finite  $m = \max(J)$  exists. If now  $x \in \bigcup_{i \in \{0, \dots, m\}} (K \cap N_i)$  then  $\exists i \in \{0, \dots, m\}$  so that  $x \in K \cap N_i \Rightarrow K \cap N_i \neq \emptyset \Rightarrow i \in J \Rightarrow x \in \bigcup_{i \in J} (K \cap N_i)$ , also if  $x \in \bigcup_{i \in J} (K \cap N_i)$  then  $\exists i \in J \subseteq \{0, \dots, m\}$  such that  $x \in K \cap N_i \Rightarrow x \in \bigcup_{i \in \{0, \dots, m\}} (K \cap N_i)$ . So we have

$$K = \bigcup_{i \in J} (K \cap N_i) = \bigcup_{i \in \{0, \dots, m\}} (K \cap N_i) \quad (11.65)$$

As  $\{K \cap N_i\}_{i \in \{0, \dots, m\}}$  is pairwise disjoint we have using the above and the previous lemma (see 11.390) that  $\sum_{i \in K} x_i \stackrel{11.390}{=} \sum_{i=0}^m (\sum_{j \in (K \cap N_i)} x_j) \leq \sum_{i=0}^m (\sum_{j \in N_i} x_j) \leq \sum_{i=0}^{\infty} (\sum_{j \in N_i} x_j)$ . Hence by 11.63 we have that  $\sum_{i \in I} x_i < \sum_{i=0}^{\infty} (\sum_{j \in N_i} x_j) + \varepsilon$  proving that  $\sum_{i \in I} x_i \leq \sum_{i=0}^{\infty} (\sum_{j \in N_i} x_j)$

As we have in all cases that  $\sum_{i \in I} x_i \leq \sum_{i=0}^{\infty} (\sum_{j \in N_i} x_j)$  we can use 11.62 to prove that

$$\sum_{i \in I} x_i = \sum_{i=0}^{\infty} \left( \sum_{j \in N_i} x_j \right) \quad \square$$

We can easily extend the above to the following general case of countable families

**Theorem 11.392.** Let  $I, K$  be countable sets,  $I \neq \emptyset$ ,  $\{K_i\}_{i \in I}$  a family of countable sets such that  $K = \bigcup_{i \in I} K_i$  and  $\forall i, j \in I$   $K_i \cap K_j = \emptyset$  then for every  $\{x_i\}_{i \in I} \subseteq [0, \infty[$  such that  $\sum_{i \in K} x_i$  converges we have that  $\forall i \in I$   $\sum_{j \in K_i} x_j$  converges and  $\sum_{i \in I} (\sum_{j \in K_i} x_j)$  converges to  $\sum_{i \in K} x_i$

**Proof.** As  $I$  is countable we have two possible cases to consider

**$I$  is finite.** then there exists a  $n \in \mathbb{N}$  and a bijection  $\beta: \{0, \dots, n\} \rightarrow I$ . Define now  $\{L_i\}_{i \in \{0, \dots, n\}}$  by  $L_i = K_{\beta(i)}$  then we have  $\bigcup_{i \in \{0, \dots, n\}} L_i = \bigcup_{i \in \{0, \dots, n\}} K_{\beta(i)} \stackrel{2.68}{=} \bigcup_{i \in I} K_i = K$ , further  $\forall i, j \in \{0, \dots, n\}$  with  $i \neq j$  we have as  $\beta$  is a bijection that  $\beta(i) = \beta(j)$  so that  $L_i \cap L_j = K_{\beta(i)} \cap K_{\beta(j)} = \emptyset$ . So we can use 11.389 giving that  $\forall i \in \{0, \dots, n\} \sum_{j \in L_i} x_j$  converges and  $\sum_{i \in \{0, \dots, n\}} (\sum_{j \in L_i} x_j) = \sum_{i \in K} x_i$ . Finally  $\forall i \in I$  we have that  $K_i = K_{\beta(\beta^{-1}(i))} = L_{\beta^{-1}(i)}$  so that  $\sum_{j \in K_i} x_j$  converges and  $\sum_{i \in I} (\sum_{j \in K_i} x_j)$   $\stackrel{\text{definition (see 11.385)}}{=} \sum_{i \in \{0, \dots, n\}} (\sum_{j \in L_i} x_j) = \sum_{i \in K} x_i$ .

**I is infinite.** there exists a bijection  $\beta: \mathbb{N} \rightarrow I$ . Define now  $\{L_i\}_{i \in \mathbb{N}}$  by  $L_i = K_{\beta(i)}$  then we have  $\bigcup_{i \in \mathbb{N}} L_i = \bigcup_{i \in \mathbb{N}} K_{\beta(i)} \stackrel{2.68}{=} \bigcup_{i \in I} K_i = K$ , further  $\forall i, j \in \mathbb{N}$  with  $i \neq j$  we have as  $\beta$  is a bijection that  $\beta(i) = \beta(j)$  so that  $L_i \cap L_j = K_{\beta(i)} \cap K_{\beta(j)} = \emptyset$ . So we can use the previous lemma (see 11.391) which gives that  $\forall i \in \mathbb{N}$  we have  $\sum_{j \in L_i} x_j$  converges and  $\sum_{i=0}^{\infty} (\sum_{j \in L_i} x_j)$  converges to  $\sum_{i \in K} x_i$ . Finally  $\forall i \in I$  we have that  $K_i = K_{\beta(\beta^{-1}(i))} = L_{\beta^{-1}(i)}$  so that  $\sum_{j \in K_i} x_j$  converges and  $\sum_{i \in I} (\sum_{j \in K_i} x_j) \stackrel{\text{definition (see 11.385)}}{=} \sum_{i=0}^{\infty} (\sum_{j \in K_i} x_j) = \sum_{i \in K} x_i$ .  $\square$

### 11.11.7 Absolute convergent series

**Theorem 11.393. (Dominant Convergence)** Let  $\langle X, \|\cdot\| \rangle$  be a Banach space  $\sum_{i=k}^{\infty} x_i$  a serie such that there exists a convergent serie  $\sum_{i=k}^{\infty} s_i$  such that  $\forall i \in \{k, \dots, \infty\}$  we have  $s_i \in \mathbb{R} \wedge \|x_i\| \leq s_i$  (so  $s_i \geq 0$ ) then  $\sum_{i=k}^{\infty} x_i$  converges and  $\|\sum_{i=k}^{\infty} x_i\| \leq \sum_{i=k}^{\infty} s_i$

**Proof.** First we prove that  $\{\sum_{i=k}^n x_i\}_{n \in \{k, \dots, \infty\}}$  is Cauchy. So let  $\varepsilon > 0$  then as  $\{\sum_{i=k}^n\}_{n \in \{k, \dots, \infty\}}$  converges it is Cauchy and thus  $\exists N$  such that if  $n, m \geq N$  then  $|\sum_{i=k}^n s_i - \sum_{i=k}^m s_i| < \varepsilon$ . Consider now the following sub cases for  $n, m$ :

1. ( $n = m$ ) then  $\|\sum_{i=k}^n x_i - \sum_{i=k}^m x_i\| = 0 < \varepsilon$
2. ( $n < m$ ) then  $\|\sum_{i=k}^n x_i - \sum_{i=k}^m x_i\| = \|\sum_{i=k}^m x_i - \sum_{i=k}^n x_i\| = \|\sum_{i=n+1}^m x_i\| \stackrel{11.70}{\leq} \sum_{i=n+1}^m \|x_i\| \stackrel{10.62}{\leq} \sum_{i=n+1}^m |s_i| \stackrel{s_i \geq 0}{=} \sum_{i=n+1}^m s_i \stackrel{0 \leq \sum_{i=n+1}^m s_i}{=} |\sum_{i=n+1}^m s_i| = |\sum_{i=k}^n s_i - \sum_{i=k}^m s_i| < \varepsilon$
3. ( $n > m$ ) then  $\|\sum_{i=k}^n x_i - \sum_{i=k}^m x_i\| = \|\sum_{i=m+1}^n x_i\| \stackrel{11.70}{\leq} \sum_{i=m+1}^n \|x_i\| \stackrel{10.62}{\leq} \sum_{i=m+1}^n |s_i| = \sum_{i=m+1}^n s_i \stackrel{0 \leq \sum_{i=m+1}^n s_i}{=} |\sum_{i=m+1}^n s_i| = |\sum_{i=k}^n s_i - \sum_{i=k}^m s_i| < \varepsilon$  (as  $s_i \geq 0 \Rightarrow |s_i| = s_i$ )  $\square$

So in all cases we have  $\|\sum_{i=k}^n x_i - \sum_{i=k}^m x_i\| < \varepsilon$  proving that  $\{\sum_{i=k}^n\}_{n \in \{k, \dots, \infty\}}$  is Cauchy. as  $\langle X, \|\cdot\| \rangle$  is a Banach space and thus complete we have that  $\{\sum_{i=k}^n\}_{n \in \{k, \dots, \infty\}}$  converges and thus that  $\sum_{i=k}^{\infty} x_i$  exist.

Second if  $n \in \{k, \dots, \infty\}$  then we have  $\|\sum_{i=k}^n x_i\| \leq \sum_{i=k}^n \|x_i\| \leq \sum_{i=k}^n |x_i| = \sum_{i=k}^n s_i$  so that by 11.351 we have that  $\|\sum_{i=k}^{\infty} x_i\| = \left\| \lim_{n \rightarrow \infty} \sum_{i=k}^n x_i \right\| \stackrel{11.350}{\leq} \lim_{n \rightarrow \infty} \sum_{i=k}^n s_i$ .

**Corollary 11.394.** Let  $\langle X, \|\cdot\| \rangle$  is a normed Banach space,  $\{x_i\}_{i \in \{k, \dots, \infty\}}$  a sequence so that  $\sum_{i=k}^{\infty} \|x_i\|$  converges then  $\sum_{i=k}^{\infty} x_i$  converges and  $\|\sum_{i=k}^{\infty} x_i\| \leq \sum_{i=k}^{\infty} \|x_i\|$

**Proof.** This follows from the previous theorem as  $\{\|x_i\|\}_{i \in \{k, \dots, \infty\}}$  is a sequence in  $\mathbb{R}$  and  $\forall i \in \{k, \dots, \infty\}$  we have  $\|x_i\| \leq \|x_i\|$ .  $\square$

the above theorem motivates the definition of absolute convergence of a serie.

**Definition 11.395. (Absolute Convergence)** Let  $\langle X, \|\cdot\| \rangle$  be a normed Banach space then  $\sum_{i=k}^{\infty} x_i$  is **absolute convergent** if and only if  $\sum_{i=k}^{\infty} \|x_i\|$  is convergent,

**Note 11.396.** Using 11.394 we have that absolute convergence implies convergence.

We set now out to prove that absolute convergent series in a Banach space are commutative and associative. First we need some theorems and lemmas.

**Theorem 11.397.** Let  $\langle X, \|\cdot\| \rangle$  be a normed space then we have the following equivalences for  $\sum_{i=k}^{\infty} x_i$

1.  $\sum_{i=k}^{\infty} x_i$  is absolute convergent
2.  $\forall \varepsilon > 0$  there exists a  $n \in \{k, \dots, \infty\}$  such that  $\forall n \leq l$  and  $p \in \mathbb{N}_0$  we have  $\sum_{i=l+1}^{l+p} \|x_i\| < \varepsilon$

3.  $\forall \varepsilon > 0$  there exists a finite  $K \subseteq \{k, \dots, \infty\}$  such that  $\forall H \subseteq \{k, \dots, \infty\}$  with  $H$  finite and  $K \cap H = \emptyset$  we have  $\sum_{i \in H} \|x_i\| < \varepsilon$

**Proof.** As  $\{\|x_i\|\}_{i \in \{k, \dots, \infty\}}$  is a sequence of positive numbers and absolute convergence means that  $\sum_{i=k}^{\infty} \|x_i\|$  converges we can use 11.382 to prove the theorem.  $\square$

**Lemma 11.398.** Let  $\langle X, \|\cdot\|\rangle$  be a normed space,  $\varepsilon > 0$  and  $\sum_{i=k}^{\infty} x_i$  a absolute convergent serie then there exists a finite  $K \subseteq \{k, \dots, \infty\}$  such that  $\forall \varepsilon > 0$  we have

1.  $\forall H \subseteq \{k, \dots, \infty\}$  with  $H$  finite and  $K \subseteq H$  we have  $\|\sum_{i=k}^{\infty} x_i - \sum_{i \in H} x_i\| < \varepsilon$
2.  $\forall L \subseteq \{k, \dots, \infty\}$  with  $L$  finite and  $K \cap L = \emptyset$  we have  $\sum_{i \in L} \|x_i\| < \varepsilon$

**Proof.** Let  $\varepsilon > 0$ . As absolute convergence implies convergence (see 11.396) there exists a  $n \in \{k, \dots, \infty\}$  so that

$$\forall l \geq n \text{ we have } \left\| \sum_{i=k}^l x_i - \sum_{i=k}^{\infty} x_i \right\| < \frac{\varepsilon}{2} \quad (11.66)$$

Using absolute convergence and the previous theorem (see 11.397) there exists a finite  $K' \subseteq \{k, \dots, \infty\}$  such that  $\forall L \subseteq \{k, \dots, \infty\}$  with  $L$  finite and  $K' \cap L = \emptyset$  we have  $\sum_{i \in L} \|x_i\| < \frac{\varepsilon}{2}$ . Take now  $K = \{k, \dots, n\} \cup K'$  then we have

$$\{k, \dots, n\} \subseteq K \wedge \forall L \subseteq \{k, \dots, \infty\} \vdash L \text{ finite and } K \cap L = \emptyset \text{ we have } \sum_{i \in L} \|x_i\| < \frac{\varepsilon}{2} < \varepsilon \quad (11.67)$$

Let now  $H \subseteq \{k, \dots, \infty\}$  with  $K \subseteq H$  and  $H$  finite. Let  $p = \max(H) \Rightarrow H \subseteq \{k, \dots, p\}$  then

$$\begin{aligned} \left\| \sum_{i \in H} x_i - \sum_{i=k}^{\infty} x_i \right\| &= \left\| \sum_{i \in \{k, \dots, p\}} x_i - \sum_{i \in \{k, \dots, p\} \setminus H} x_i - \sum_{i=k}^{\infty} x_i \right\| \\ &= \left\| \sum_{i=k}^p x_i - \sum_{i \in \{k, \dots, p\} \setminus H} x_i - \sum_{i=k}^{\infty} x_i \right\| \\ &\leq \left\| \sum_{i=k}^p x_i - \sum_{i=k}^{\infty} x_i \right\| + \left\| \sum_{i \in \{k, \dots, p\} \setminus H} x_i \right\| \\ &< \left\| \sum_{i=k}^p x_i - \sum_{i=k}^{\infty} x_i \right\| + \frac{\varepsilon}{2} \text{ as } K \cap \{k, \dots, p\} \setminus H \subseteq \{k, \dots, p\} \setminus K = \emptyset \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \text{ as } \{k, \dots, n\} \subseteq K \subseteq H \Rightarrow n \leq \max(H) = p \text{ and 11.66} \\ &= \varepsilon \end{aligned}$$

which proves (1). (2) follows from 11.67.  $\square$

**Theorem 11.399.** Let  $\langle X, \|\cdot\|\rangle$  be a Banach space,  $k, m, n \in \mathbb{N} \vdash m \leq n$  and  $\{x_{i,j}\}_{(i,j) \in \{k, \dots, \infty\} \times \{m, \dots, n\}}$  a family in  $X$  then if  $\forall j \in \{m, \dots, n\}$  we have that  $\sum_{i=k}^{\infty} x_{i,j}$  exists then  $\sum_{i=k}^{\infty} (\sum_{j=m}^n x_{i,j})$  exists and is equal to  $\sum_{j=m}^n (\sum_{i=k}^{\infty} x_{i,j})$

**Proof.** As  $\forall j \in \{m, \dots, n\}$  we have that  $b_j = \sum_{i=k}^{\infty} x_{i,j}$  exists there is given a  $\varepsilon > 0$  a  $N_j \geq 0$  such that if  $l \geq N_j$  then  $\|\sum_{i=k}^l x_{i,j} - b_j\| < \frac{\varepsilon}{n-m+1}$ . then we have that if  $l \geq \max(\{b_j | j \in \{m, \dots, n\}\})$  that  $\|\sum_{i=k}^l (\sum_{j=m}^n x_{i,j}) - \sum_{j=m}^n b_j\| \stackrel{10.12}{=} \|\sum_{j=m}^n (\sum_{i=k}^l x_{i,j}) - \sum_{j=m}^n b_j\| = \|\sum_{j=m}^n (\sum_{i=k}^l x_{i,j} - b_j)\| \leq 11.70 \sum_{j=m}^n \|\sum_{i=k}^l x_{i,j} - b_j\| < \sum_{j=m}^n \frac{\varepsilon}{n-m+1} < \varepsilon$  proving that  $\sum_{i=k}^{\infty} (\sum_{j=m}^n x_{i,j}) = \sum_{j=m}^n (\sum_{i=k}^{\infty} x_{i,j})$ .  $\square$

We are now ready to prove commutativity for absolute convergent series, first we prove that permutations conserves convergence.

**Theorem 11.400.** Let  $\langle X, \|\cdot\| \rangle$  be a Banach space and  $\sum_{i=k}^{\infty} x_i$  a absolute convergent serie then for every permutation  $\beta: \{k, \dots, \infty\} \rightarrow \{k, \dots, \infty\}$  we have that  $\sum_{i=k}^{\infty} x_{\beta(i)}$  is absolute convergent. If  $\langle X, \|\cdot\| \rangle$  is a Banach space then it follows from 11.396 that  $\sum_{i=k}^{\infty} x_{\beta(i)}$  is convergent.

**Proof.** As  $\sum_{i=k}^{\infty} x_i$  is Absolute convergent we have that for the sequence  $\{\|x_i\|\}_{i \in \{k, \dots, \infty\}}$  of positive numbers that  $\sum_{i=k}^{\infty} \|x_i\|$  converges. Using 11.384 it follows then that  $\sum_{i=k}^{\infty} \|x_{\beta(i)}\|$  is convergent.  $\square$

Next we must prove that the sum is independent of the bijection.

**Theorem 11.401.** Let  $\langle X, \|\cdot\| \rangle$  be a Banach space and  $\sum_{i=k}^{\infty} x_i$  a absolute convergent serie,  $\alpha, \beta: \{k, \dots, \infty\} \rightarrow \{k, \dots, \infty\}$  two bijections then  $\sum_{i=k}^{\infty} x_{\alpha(i)} = \sum_{i=k}^{\infty} x_{\beta(i)}$  (the previous theorem ensures convergence). Note that by taking  $\alpha = 1_{\{k, \dots, \infty\}}$  we have that  $\sum_{i=k}^{\infty} x_i = \sum_{i=k}^{\infty} x_{\beta(i)}$ .

**Proof.** Let  $s = \sum_{i=k}^{\infty} x_{\alpha(i)}$  and  $t = \sum_{i=k}^{\infty} x_{\beta(i)}$  and assume that  $s \neq t \Rightarrow 0 < \|s - t\|$ . Then using 11.398 there exists finite  $K_{\alpha}, K_{\beta} \subseteq \{k, \dots, \infty\}$  such that  $\forall H_{\alpha}, H_{\beta}$  finite with  $K_{\alpha} \subseteq H_{\alpha}$  and  $K_{\beta} \subseteq H_{\beta}$  we have

$$\left\| \sum_{i \in H_{\alpha}} x_{\alpha(i)} - s \right\|, \left\| \sum_{i \in H_{\beta}} x_{\beta(i)} - t \right\| < \frac{\|s - t\|}{4} \quad (11.68)$$

Define now  $P_{\alpha\beta} = K_{\alpha} \bigcup \alpha^{-1}(\beta(K_{\beta}))$  and  $P_{\beta\alpha} = K_{\beta} \bigcup \beta^{-1}(\alpha(K_{\alpha}))$  then as  $\alpha, \beta$  are bijections and  $K_{\alpha}, K_{\beta}$  are finite we have that  $P_{\alpha\beta}, P_{\beta\alpha}$  are finite and  $H_{\alpha} \subseteq P_{\alpha\beta}, H_{\beta} \subseteq P_{\beta\alpha}$  so that by 11.68

$$\left\| \sum_{i \in P_{\alpha\beta}} x_{\alpha(i)} - s \right\|, \left\| \sum_{i \in P_{\beta\alpha}} x_{\beta(i)} - t \right\| < \frac{\|s - t\|}{4} \quad (11.69)$$

Now

$$\begin{aligned} \sum_{i \in P_{\alpha\beta}} x_{\alpha(i)} &= \sum_{i \in K_{\alpha} \bigcup \alpha^{-1}(\beta(K_{\beta}))} x_{\alpha(i)} \\ &\stackrel{10.43}{=} \sum_{i \in \alpha(K_{\alpha} \bigcup \alpha^{-1}(\beta(K_{\beta})))} x_i \\ &= \sum_{i \in \alpha(K_{\alpha}) \bigcup \beta(K_{\beta})} x_i \\ &= \sum_{i \in \beta(K_{\beta}) \bigcup \beta^{-1}(\alpha(K_{\alpha}))} x_i \\ &\stackrel{10.43}{=} \sum_{i \in K_{\beta} \bigcup \beta^{-1}(\alpha(K_{\alpha}))} x_{\beta(i)} \\ &= \sum_{i \in P_{\beta\alpha}} x_{\beta(i)} \end{aligned}$$

proving that

$$\sum_{i \in P_{\alpha\beta}} x_{\alpha(i)} = \sum_{i \in P_{\beta\alpha}} x_{\beta(i)} \quad (11.70)$$

Now using the above we have

$$\begin{aligned} \|s - t\| &= \left\| s - \sum_{i \in P_{\alpha\beta}} x_{\alpha(i)} + \sum_{i \in P_{\alpha\beta}} x_{\alpha(i)} - t \right\| \\ &\leq \left\| \sum_{i \in P_{\alpha\beta}} x_{\alpha(i)} - s \right\| + \left\| \sum_{i \in P_{\alpha\beta}} x_{\alpha(i)} - t \right\| \\ &\stackrel{11.69}{<} \frac{\|s - t\|}{4} + \frac{\|s - t\|}{4} \\ &< \frac{\|s - t\|}{2} \\ &< \|s - t\| \end{aligned}$$

a contradiction. So we must have that  $s = t$ .  $\square$

The above theorem lets us to extend the definition of a absolute convergent serie to a absolute convergent summable serie.

**Definition 11.402. (Absolute convergent serie of a countable family)** Let  $\langle X, \|\cdot\| \rangle$  be a Banach space, and  $I$  a countable set,  $\{x_i\}_{i \in I} \subseteq X$  a countable family then we define absolute convergence of  $\sum_{i \in I} x_i$  as follows

1. If  $I$  is finite then  $\sum_{i \in I} x_i$  converges and  $\sum_{i \in I} x_i$  is as defined in 10.36). Using 10.39 we have then that  $\sum_{i \in I} x_i = \sum_{i=0}^n x_{\beta(i)}$  where  $\beta: \{0, \dots, n\} \rightarrow I$  is a bijection.
2. If  $I$  is infinite then  $\sum_{i \in I} x_i$  converges if  $\sum_{i=0}^{\infty} x_{\beta(i)}$  is absolute convergent where  $\beta: \mathbb{N} \rightarrow I$  is a bijection (note that as  $I$  is infinite and countable such a bijection exists). We denote then  $\sum_{i \in I} x_i$  by  $\sum_{i=0}^{\infty} x_{\beta(i)}$  (which is well defined as absolute convergence implies convergence (see 11.396))

we must of course prove that this definition is independent of the choice of  $\beta$

**Proof.**

1. This follows from 10.39
2. Let  $\alpha: \mathbb{N} \rightarrow I$  be another bijection then  $\alpha^{-1} \circ \beta: \mathbb{N} \rightarrow \mathbb{N}$  is a bijection. So using 11.401 and absolute convergence we have that  $\sum_{i=0}^{\infty} x_{\alpha(\alpha^{-1}(\beta(i)))}$  is convergent and  $\sum_{i=0}^{\infty} x_{\alpha(i)} = \sum_{i=0}^{\infty} x_{\alpha(\alpha^{-1}(\beta(i)))} = \sum_{i=0}^{\infty} x_{\beta(i)} = s$   $\square$

Using the above definition we can extend 10.43

**Theorem 11.403.** Let  $\langle X, \|\cdot\| \rangle$  be a normed space  $I$  a countable set,  $\beta: I \rightarrow J$  a bijection (so that  $J$  is also countable) and  $\sum_{j \in J} x_j$  a absolute convergent space then  $\sum_{i \in I} x_{\beta(i)}$  is absolute convergent and  $\sum_{i \in I} x_i = \sum_{j \in J} x_j$

**Proof.** As  $I$  is countable we must consider two cases

**$I$  is finite.** then  $\sum_{i \in I} x_{\beta(i)}$  is absolute convergent by definition and  $\sum_{i \in I} x_{\beta(i)} = \sum_{i \in I} x_i$   $\square$

**$I$  is infinite.** then there exists a bijection  $\alpha: \mathbb{N} \rightarrow I$  so that  $\beta \circ \alpha: \mathbb{N} \rightarrow J$  is a bijection. By definition we have then that  $\sum_{i=0}^{\infty} x_{\beta(\alpha(i))} = \sum_{i=0}^{\infty} x_{(\beta \circ \alpha)(i)}$  is absolute convergent. So we have that  $\sum_{i \in I} x_{\beta(i)}$  is absolute convergent and  $\sum_{i \in I} x_{\beta(i)} = \sum_{i=1}^{\infty} x_{\beta(\alpha(i))} = \sum_{j \in J} x_j$   $\square$

**Lemma 11.404.** Let  $\langle X, \|\cdot\| \rangle$  be a normed space and  $I, J$  set such that there exists a bijection  $\beta: I \rightarrow J$  then  $\{\sum_{i \in H} \|x_{\beta(i)}\| \mid H \subseteq J \wedge H \text{ is finite}\} = \{\sum_{i \in H} \|x_i\| \mid H \subseteq I \wedge H \text{ is finite}\}$

**Proof.**

$$\begin{aligned}
 s \in \left\{ \sum_{i \in H} \|x_{\beta(i)}\| \mid H \subseteq J \wedge H \text{ is finite} \right\} &\Leftrightarrow \exists H \text{ finite } \wedge H \subseteq J \text{ we have } s = \sum_{i \in H} \|x_{\beta(i)}\| \\
 &\Leftrightarrow \exists H \text{ finite } \wedge H \subseteq J \text{ we have } s = \sum_{i \in \beta(H)} \|x_i\| \quad \text{10.43} \\
 &\stackrel{\beta \text{ is bijective}}{\Leftrightarrow} \exists H \text{ finite } \wedge H \subseteq I \text{ we have } s = \sum_{i \in H} \|x_i\| \\
 &\Leftrightarrow s \in \left\{ \sum_{i \in H} \|x_i\| \mid H \subseteq I \wedge H \text{ is finite} \right\}
 \end{aligned}$$

$\square$

**Theorem 11.405.** Let  $\langle X, \|\cdot\| \rangle$  be a Banach space,  $I$  a countable set then  $\sum_{i \in I} x_i$  is absolute convergent if and only if  $\sup(\{\sum_{i \in H} \|x_i\| \mid H \subseteq I \wedge H \text{ is finite}\})$  exists.

**Proof.**

$\Rightarrow$ . as  $I$  is countable we must consider two cases

**$I$  is finite.** then  $\forall x \in \{\sum_{i \in H} \|x_i\| \mid H \subseteq I \wedge H \text{ is finite}\}$  we have that  $\exists H$  finite with  $H \subseteq I$  so that  $x = \sum_{i \in H} \|x_i\| \leq \sum_{i \in I} \|x_i\| \in \{\sum_{i \in H} \|x_i\| \mid H \subseteq I \wedge H \text{ is finite}\}$  so that  $\sup(\{\sum_{i \in H} \|x_i\| \mid H \subseteq I \wedge H \text{ is finite}\}) = \sum_{i \in I} \|x_i\|$  exists

**$I$  is infinite.** then there exists a bijection  $\beta: \mathbb{N} \rightarrow I$  so that  $\sum_{i=0}^{\infty} x_i$  is absolute convergent meaning that  $\sum_{i=0}^{\infty} \|x_{\beta(i)}\|$  is convergent which by 11.387 gives that  $\sup(\{\sum_{i \in H} \|x_{\beta(i)}\| \mid H \subseteq I \wedge H \text{ is finite}\})$  exists. Using 11.404 we have then that  $\sup(\{\sum_{i \in H} \|x_i\| \mid H \subseteq I \wedge H \text{ is finite}\})$  exists.

$\Leftarrow$ . for  $I$  countable we have either

**$I$  is finite.** then by definition  $\sum_{i \in I} x_i$  is absolute convergent

**$I$  is infinite.** then as  $I$  is countable there exists a bijection  $\beta: \mathbb{N} \rightarrow I$ . By the hypothese  $\sup(\{\sum_{i \in H} \|x_i\| \mid H \subseteq I \wedge H \text{ is finite}\})$  exists so by 11.404  $\sup(\{\sum_{i \in \mathbb{N}} \|x_{\beta(i)}\|\})$  exists which by 11.387 means the  $\sum_{i=0}^{\infty} \|x_i\|$  is convergent or that  $\sum_{i \in I} x_i$  is absolute convergent.  $\square$

We can now extend theorem 11.397 to this new definition of absolute convergence.

**Theorem 11.406.** Let  $\langle X, \|\cdot\| \rangle$  be a Banach space,  $I$  a countable set then  $\sum_{i \in I} x_i$  is absolute convergent if and only if  $\forall \varepsilon > 0$  there exists a finite  $H \subseteq I$  such that  $\forall K \subseteq I$  with  $K \cap H = \emptyset$  we have  $\sum_{i \in K} \|x_i\| < \varepsilon$

**Proof.**

$\Rightarrow$ . Take  $\varepsilon > 0$ , as  $I$  is countable we must consider two cases

**$I$  is finite.** take  $H = I$  then if  $K \subseteq I$  with  $K \cap H = \emptyset$  then  $K = K \cap I = K \cap H = \emptyset$  so that  $\sum_{i \in H} \|x_i\| = 0 < \varepsilon$

**$I$  is infinite.** as  $\sum_{i \in I} x_i$  is absolute convergent there exists a bijection  $\beta: \mathbb{N} \rightarrow I$  such that  $\sum_{i=0}^{\infty} \|x_{\beta(i)}\|$  is absolute convergent. Using 11.397 there exists a finite  $H' \subseteq \mathbb{N}$  such that  $\forall K \subseteq \mathbb{N}$  with  $K \cap H' = \emptyset$  we have  $\sum_{i \in K} \|x_{\beta(i)}\| < \varepsilon$ . Take  $H = \beta(H')$  then  $H$  is finite and if  $K \subseteq I$  with  $K \cap H = \emptyset$  we have  $\emptyset = \beta^{-1}(K \cap H) = \beta^{-1}(K) \cap \beta^{-1}(H) = \beta^{-1}(K) \cap H'$  so that  $\sum_{i \in K} \|x_i\| \stackrel{10.43}{=} \sum_{i \in \beta^{-1}(K)} \|x_{\beta(i)}\| < \varepsilon$ .

$\Leftarrow$ . as  $I$  is countable we must consider two cases

**$I$  is finite.** by definition we have then that  $\sum_{i \in I} x_i$  is absolute convergent

**$I$  is infinite.** then there exists a bijection  $\beta: \mathbb{N} \rightarrow I$ . Let  $\varepsilon > 0$ , by the hypothese there exist a finite  $H' \subseteq I$  such that  $\forall K \subseteq H'$  with  $K \cap H' = \emptyset$  we have  $\sum_{i \in K} \|x_i\| < \varepsilon$ . Take  $H = \beta^{-1}(H')$  then if  $K \subseteq \mathbb{N}$  with  $K \cap H = \emptyset$  we have that  $\emptyset = \beta(K \cap H) = \beta(K) \cap \beta(H) = \beta(K) \cap H'$  so that  $\sum_{i \in K} \|x_{\beta(i)}\| \stackrel{10.43}{=} \sum_{i \in \beta(K)} \|x_i\| < \varepsilon$ .  $\square$

We have now the following extension of the dominant convergence theorem.

**Lemma 11.407. (Dominant Convergence)** Let  $\langle X, \|\cdot\| \rangle$  be a Banach space,  $I$  a countable set,  $\{x_i\}_{i \in I}$  a countable family such that there exists a  $\{s_i\}_{i \in I} \subseteq [0, \dots, \infty[$  such that  $\sum_{i \in I} s_i$  is convergent and  $\forall i \in I$  we have  $\|x_i\| \leq s_i$  then  $\sum_{i \in I} x_i$  is absolute convergent and  $\|\sum_{i \in I} x_i\| \leq \sum_{i \in I} s_i$

**Proof.** As  $I$  is countable we must consider two possibilities

**$I$  is finite.** then by definition  $\sum_{i \in I} x_i$  converges and  $\|\sum_{i \in I} x_i\| \leq \sum_{i \in I} \|x_i\| \leq \sum_{i \in I} s_i$

**$I$  is infinite.** then by definition (see 11.385) there exists a bijection  $\beta: \mathbb{N} \rightarrow I$  such that  $\sum_{i=0}^{\infty} s_{\beta(i)}$  is convergent. As  $\forall i \in \mathbb{N}$  we have that  $\|\|x_{\beta(i)}\|\| = \|x_{\beta(i)}\| \leq s_{\beta(i)}$  we have using 11.393 that  $\sum_{i=0}^{\infty} \|x_i\|$  is convergent and also that  $\sum_{i=0}^{\infty} x_i$  is convergent with  $\|\sum_{i=0}^{\infty} x_{\beta(i)}\| \leq \sum_{i=0}^{\infty} s_i$ . So by definition  $\sum_{i \in I} x_i$  is absolute convergent and  $\|\sum_{i \in I} x_i\| = \|\sum_{i=0}^{\infty} x_{\beta(i)}\| \leq \sum_{i=0}^{\infty} s_{\beta(i)} = \sum_{i \in I} x_i$ .  $\square$

We can extend now 11.398 to absolute convergent series of families

**Lemma 11.408.** Let  $\langle X, \|\|\rangle$  be a Banach space,  $I$  a countable set then  $\sum_{i \in I} x_i$  is absolute convergent then  $\forall \varepsilon > 0$  there exists a finite  $K \subseteq I$  such that

1.  $\forall H$  finite with  $K \subseteq H \subseteq I$  we have that  $\|\sum_{i \in I} x_i - \sum_{i \in H} x_i\| < \varepsilon$
2.  $\forall L \subseteq I$  with  $L$  finite and  $K \cap L = \emptyset$  we have  $\sum_{i \in L} \|x_i\| < \varepsilon$

**Proof.** As  $I$  is countable we have two cases to consider

**$I$  is finite.** take then  $K = I$  so if  $K \subseteq H \subseteq I$  then  $H = I$  hence  $\|\sum_{i \in I} x_i - \sum_{i \in I} x_i\| = \|\sum_{i \in I} x_i - \sum_{i \in H} x_i\| = 0 < \varepsilon$ , further if  $L \subseteq I$  with  $K \cap L = \emptyset$  then  $L = K \cap L = \emptyset$  so that  $\sum_{i \in L} \|x_i\| = 0 < \varepsilon$

**$I$  is infinite.** then by definition there exists a  $\beta: \mathbb{N} \rightarrow I$  such that  $\sum_{i=0}^{\infty} x_{\beta(i)}$  is absolute convergent. Using 11.398 there exists then a  $K' \subseteq \mathbb{N}$  such  $\forall \varepsilon > 0$  we have

1.  $\forall H$  finite with  $K' \subseteq H \subseteq \mathbb{N}$  we have  $\|\sum_{i=0}^{\infty} x_{\beta(i)} - \sum_{i \in H} x_{\beta(i)}\| < \varepsilon$
2.  $\forall L \subseteq \mathbb{N}$  with  $L$  finite and  $K' \cap L = \emptyset$  we have  $\sum_{i \in L} \|x_{\beta(i)}\| < \varepsilon$

Define now  $K = \beta(K')$  then  $K \subseteq I$  and  $K$  is finite. If  $H$  is a finite set such that  $K \subseteq H \subseteq I$   $\xrightarrow[2.58]{} \beta^{-1}(K) \subseteq \beta^{-1}(H) \subseteq \beta^{-1}(I) \Rightarrow K' \subseteq \beta^{-1}(H) \subseteq \mathbb{N}$  hence using (1) we have  $\|\sum_{i=0}^{\infty} x_{\beta(i)} - \sum_{i \in \beta^{-1}(H)} x_{\beta(i)}\| < \varepsilon$ . As  $\sum_{i=0}^{\infty} x_{\beta(i)}$   $\stackrel{\text{definition}}{=} \sum_{i \in I} x_i$  and  $\sum_{i \in \beta^{-1}(H)} x_{\beta(i)} \stackrel{10.43}{=} \sum_{i \in K} x_i$  we must have

$$\forall H \text{ finite with } K \subseteq H \subseteq I \text{ we have } \left\| \sum_{i \in I} x_i - \sum_{i \in H} x_i \right\| < \varepsilon$$

Further if  $L$  is finite with  $L \subseteq I$  and  $K \cap L = \emptyset$  then we have  $\beta^{-1}(L) \subseteq \beta^{-1}(I) = \mathbb{N}$  and  $\beta^{-1}(K) \cap \beta^{-1}(L) = \beta^{-1}(\emptyset) = \emptyset \Rightarrow K' \cap \beta^{-1}(L) = \emptyset$  so by (2) we have  $\sum_{i \in \beta^{-1}(L)} \|x_{\beta(i)}\| < \varepsilon$ . As  $\sum_{i \in \beta^{-1}(L)} \|x_{\beta(i)}\| \stackrel{10.43}{=} \sum_{i \in L} \|x_i\|$  we conclude that

$$\forall L \subseteq I \text{ with } L \text{ finite and } K \cap L = \emptyset \text{ we have } \sum_{i \in L} \|x_i\| < \varepsilon \quad \square$$

We have the usually properties absolute convergent series of a countable family for scalar product and sum.

**Theorem 11.409.** Let  $\langle X, \|\|\rangle$  be a Banach space over a field  $\mathbb{K}$  and  $I$  a countable set then we have

1. If  $\alpha \in \mathbb{K}$ ,  $\sum_{i \in I} x_i$  a absolute convergent serie then  $\sum_{i \in I} (\alpha \cdot x_i)$  is absolute convergent and  $\sum_{i \in I} (\alpha \cdot x_i) = \alpha \cdot \sum_{i \in I} x_i$
2. If  $\sum_{i \in I} x_i$ ,  $\sum_{i \in I} y_i$  are absolute convergent series then  $\sum_{i \in I} (x_i + y_i)$  is absolute convergent and  $\sum_{i \in I} (x_i + y_i) = \sum_{i \in I} x_i + \sum_{i \in I} y_i$

**Proof.** As  $I$  is countable we have either

**$I$  is finite.** then (1) follows from 10.125 and (2) follows from 10.52

**$I$  is infinite.**

1. then there exists as bijection  $\beta: \mathbb{N} \rightarrow I$  such that  $\sum_{i=0}^{\infty} \|x_{\beta(i)}\|$  is convergent. Hence by 11.381 we have that  $\sum_{i=0}^{\infty} \|\alpha \cdot x_i\| = \sum_{i=0}^{\infty} |\alpha| \cdot \|x_{\beta(i)}\|$  is convergent, proving that  $\sum_{i \in I} (\alpha \cdot x_i)$  is absolute convergent (and thus convergent). Further we have  $\sum_{i \in I} (\alpha \cdot x_i) \stackrel{\text{definition}}{=} \sum_{i=0}^{\infty} (\alpha \cdot x_{\beta(i)}) \stackrel{11.381}{=} \alpha \cdot \sum_{i=0}^{\infty} x_{\beta(i)} \stackrel{\text{definition}}{=} \alpha \cdot \sum_{i \in I} x_i$

2. then there exists as bijection  $\beta: \mathbb{N} \rightarrow I$  such that  $\sum_{i=0}^{\infty} \|x_{\beta(i)}\|, \sum_{i=0}^{\infty} \|y_{\beta(i)}\|$  are convergent. Hence by 11.381 we have that  $\sum_{i=0}^{\infty} \|x_{\beta(i)} + y_{\beta(i)}\|$  is convergent, proving that  $\sum_{i \in I} (x_i + y_i)$  is absolute convergent (and thus convergent). Further we have  $\sum_{i \in I} (x_i + y_i) \stackrel{\text{definition}}{=} \sum_{i=0}^{\infty} (x_{\beta(i)} + y_{\beta(i)}) \stackrel{11.381}{=} \sum_{i=0}^{\infty} x_{\beta(i)} + \sum_{i=0}^{\infty} y_{\beta(i)} \stackrel{\text{definition}}{=} \sum_{i \in I} x_i + \sum_{i \in I} y_i$   $\square$

Next we set out to prove associativity of absolute convergent series

**Theorem 11.410.** *Let  $\langle X, \|\cdot\| \rangle$  be a Banach space,  $I, K$  countable sets,  $\{N_i\}_{i \in I}$  a family of sets such that  $\forall i, j \in I \vdash \bigcup_{i \in I} N_i = K$ ,  $\sum_{i \in K} x_i$  a absolute convergent serie then  $\forall i \in I$  we have that  $\sum_{j \in N_i} x_j$  is absolute convergent and further  $\sum_{i \in I} (\sum_{j \in N_i} x_j)$  is absolute convergent and  $\sum_{i \in K} x_i = \sum_{i \in I} (\sum_{j \in N_i} x_j)$*

**Proof.** As  $\sum_{i \in K} x_i$  is absolute convergent we have that  $\sum_{i \in K} \|x_i\|$  is a convergent serie of positive numbers. So we can use the associativity of convergent series of positive numbers (see 11.392) from which it follows that  $\forall i \in I \vdash \sum_{j \in N_i} \|x_j\|$  converges and  $\sum_{i \in I} (\sum_{j \in N_i} \|x_j\|)$  converges. Using dominant convergence (see 11.407) we have that

$$\forall i \in I \text{ that } \sum_{j \in N_i} x_j \text{ is absolute convergent and } \left\| \sum_{j \in N_i} x_j \right\| \leq \sum_{j \in N_i} \|x_j\| \quad (11.71)$$

Applying then dominant convergence on the last equality of the above we have that

$$\sum_{i \in I} \left( \sum_{j \in N_i} x_j \right) \text{ is absolute convergent} \quad (11.72)$$

Take now  $s = \sum_{i \in K} x_i$  and  $t = \sum_{i \in I} (\sum_{j \in N_i} x_j)$ . Let  $\varepsilon > 0$  then using the previous lemma (see 11.408) there exists a finite  $G \subseteq K$  such that

$$\forall H \text{ finite with } G \subseteq H \subseteq K \text{ we have } \left\| \sum_{i \in H} x_i - s \right\| < \frac{\varepsilon}{3} \quad (11.73)$$

Define now  $J = \{i \in I \mid N_i \cap G \neq \emptyset\} \subseteq I$ . If  $x \in G \subseteq K = \bigcup_{i \in I} N_i$  there exists a  $i \in I$  such that  $x \in G \cap N_i \Rightarrow i \in J \Rightarrow x \in \bigcup_{j \in J} (N_i \cap G)$  proving that  $G \subseteq \bigcup_{j \in J} (N_i \cap G) \subseteq G$  so that  $G = \bigcup_{j \in J} (N_i \cap G)$ . As  $J \subseteq K$  a countable set we have by 5.66 that  $J$  is countable. If now  $J$  is infinite we have as  $\{N_i \cap G\}_{i \in J}$  is a pairwise disjoint countable family of non empty sets by 5.73 we would have that  $G$  is infinite a contradiction. Hence  $J$  is finite and thus  $\bigcup_{i \in J} (N_i \cap G)$  is finite (as a finite union of finite sets (see 5.39)).

$$J = \{i \in I \mid N_i \cap G \neq \emptyset\} \text{ is finite and } G = \bigcup_{i \in J} (N_i \cap G) \quad (11.74)$$

As  $\sum_{i \in I} (\sum_{j \in N_i} x_j) = s$  we can apply the previous lemma (see 11.408) again to find a finite  $L \subseteq I$  such that  $\forall H$  finite with  $L \subseteq H \subseteq I$  we have  $\|\sum_{i \in H} (\sum_{j \in N_i} x_j) - t\| < \frac{\varepsilon}{3}$ . Take then  $M = J \cup L$  then  $M$  is finite and  $L \subseteq M \subseteq I$  so that

$$\left\| \sum_{i \in M} \left( \sum_{j \in N_i} x_j \right) - t \right\| < \frac{\varepsilon}{3} \quad (11.75)$$

As  $\forall i \in M$  we have that  $\sum_{j \in N_i} x_j$  is absolute convergent we have by using the previous lemma (see 11.408) that there exists a finite  $M_i \subseteq N_i$  such that

$$\forall H \text{ finite with } M_i \subseteq H \subseteq N_i \text{ we have } \left\| \sum_{j \in H} x_j - \sum_{j \in M_i} x_j \right\| < \frac{\varepsilon}{3 \cdot \#(M)} \quad (11.76)$$

Now  $\forall i \in M$  define  $K_i = M_i \bigcup (N_i \cap G)$  a finite set then we have that  $M_i \subseteq K_i \subseteq N_i$  so by 11.76 we have

$$\forall i \in M \text{ we have } \left\| \sum_{j \in K_i} x_j - \sum_{j \in N_i} x_j \right\| < \frac{\varepsilon}{3 \cdot \#(M)} \quad (11.77)$$

Now as  $G = \bigcup_{j \in J} (G \cap N_j) \subseteq \bigcup_{j \in M} (G \cap N_j) \subseteq \bigcup_{j \in M} (J_j \bigcup (G \cap N_j)) = \bigcup_{j \in M} K_j \subseteq_{J_j \subseteq N_j} N_i$  and as  $M$  is finite we have that  $\bigcup_{j \in M} K_j$  is finite. So using 11.73 we have that

$$\left\| \sum_{i \in \bigcup_{j \in M} K_j} x_i - s \right\| < \frac{\varepsilon}{3} \quad (11.78)$$

Let's calculate now  $\|s - t\|$

$$\begin{aligned} \|s - t\| &= \left\| s - \sum_{i \in \bigcup_{j \in M} K_j} x_i - \sum_{i \in \bigcup_{j \in M} K_j} x_i - t \right\| \\ &\leq \left\| s - \sum_{i \in \bigcup_{j \in M} K_j} x_i \right\| + \left\| \sum_{i \in \bigcup_{j \in M} K_j} x_i - t \right\| \\ &\stackrel{11.78}{<} \frac{\varepsilon}{3} + \left\| \sum_{i \in \bigcup_{j \in M} K_j} x_i - t \right\| \\ &\stackrel{10.46 + \text{if } i \neq j \Rightarrow K_i \cap K_j = \emptyset}{=} \frac{\varepsilon}{3} + \left\| \sum_{j \in M} \left( \sum_{j \in K_j} x_j \right) - t \right\| \\ &= \frac{\varepsilon}{3} + \left\| \sum_{i \in M} \left( \sum_{j \in K_i} x_j \right) - \sum_{i \in M} \left( \sum_{j \in N_i} x_j \right) + \sum_{i \in M} \left( \sum_{j \in N_i} x_j \right) - t \right\| \\ &\leq \frac{\varepsilon}{3} + \left\| \sum_{i \in M} \left( \sum_{j \in K_i} x_j \right) - \sum_{i \in M} \left( \sum_{j \in N_i} x_j \right) \right\| + \left\| \sum_{i \in M} \left( \sum_{j \in N_i} x_j \right) - t \right\| \\ &\stackrel{11.75}{<} \frac{\varepsilon}{3} + \left\| \sum_{i \in M} \left( \sum_{j \in K_i} x_j \right) - \sum_{i \in M} \left( \sum_{j \in N_i} x_j \right) \right\| + \frac{\varepsilon}{3} \\ &= \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \left\| \sum_{i \in M} \left( \sum_{j \in K_i} x_j - \sum_{j \in N_i} x_j \right) \right\| \\ &\stackrel{M \text{ is finite}}{\leq} \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \sum_{i \in M} \left\| \sum_{j \in K_i} x_j - \sum_{j \in N_i} x_j \right\| \\ &\stackrel{11.77}{\leq} \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \sum_{i \in M} \left( \frac{\varepsilon}{3 \cdot \#(M)} \right) \\ &= \varepsilon \end{aligned}$$

So given  $\varepsilon > 0$  we have that  $\|s - t\| < \varepsilon$  from which it follows that  $\|s - t\| \leq 0$  and thus  $s = t$  or using the definitions of  $s$  and  $t$  that

$$\sum_{i \in K} x_i = \sum_{i \in I} \left( \sum_{j \in N_i} x_j \right) \quad \square$$

As a application of associativity we proof now distributivity for absolute convergent series in  $\mathbb{R}$

**Theorem 11.411.** *Let  $I, J$  be countable sets,  $\langle \mathbb{R}, \|\cdot\| \rangle$  the set of the reals with the absolute value as norm (which is a Banach space by 11.360).  $\sum_{i \in I} x_i$  and  $\sum_{i \in J} y_i$  be absolute convergent series of countable families then  $\sum_{(i,j) \in I \times J} (x_i \cdot y_j)$  is absolute convergent and  $\sum_{(i,j) \in I \times J} (x_i \cdot y_j) = (\sum_{i \in I} x_i) \cdot (\sum_{j \in J} y_j)$  [note that by 5.70  $I \times J$  is countable]. Here  $\sum_{(i,j) \in I \times J} (x_i \cdot y_j)$  is a shorthand for  $\sum_{\alpha \in I \times J} (x_{\pi_1(\alpha)} \cdot y_{\pi_2(\alpha)})$*

**Proof.** As  $\sum_{i \in I} x_i$  and  $\sum_{j \in J} y_j$  are absolute convergent we have that  $s = \sup(\{\sum_{i \in I} \|x_i\| \mid H \subseteq I \wedge H \text{ finite}\})$  and  $t = \sup(\{\sum_{j \in J} \|y_j\| \mid H \subseteq J \wedge H \text{ finite}\})$  exists (see 11.405). Take now a finite  $H \subseteq I \times J$  then as  $\pi_1: H \rightarrow \pi_1(H)$  and  $\pi_2: H \rightarrow \pi_2(H)$  are surjective we have that  $\pi_1(H), \pi_2(H)$  are finite (see 5.48). If  $(i, j) \in H$  then  $i = \pi_1(i) \in \pi_1(H)$  and  $j = \pi_2(i, j) \in \pi_2(H)$  so that  $(i, j) \in \pi_1(H) \times \pi_2(H)$  proving that  $H \subseteq \pi_1(H) \times \pi_2(H)$ . Hence  $\sum_{(i, j) \in H} (\|x_i\| \cdot \|y_j\|) \leq \sum_{(i, j) \in \pi_1(H) \times \pi_2(H)} \stackrel{10.59}{=} \sum_{(i, j) \in H} (\|x_i\| \cdot \|y_j\|)$

$$\begin{aligned} \sum_{(i, j) \in H} \|x_i \cdot y_j\| &= \sum_{(i, j) \in H} (\|x_i\| \cdot \|y_j\|) \\ &\leq \sum_{(i, j) \in \pi_1(H) \times \pi_2(H)} (\|x_i\| \cdot \|y_j\|) \\ &\stackrel{10.59}{=} \left( \sum_{i \in \pi_1(H)} \|x_i\| \right) \cdot \left( \sum_{j \in \pi_2(H)} \|y_j\| \right) \\ &\stackrel{9.41}{\leq} s \cdot t \end{aligned}$$

So  $\sup(\{\sum_{(i, j) \in H} \|x_i \cdot y_j\| \mid H \subseteq I \times J \wedge H \text{ finite}\})$  exists (see 9.43) and thus by 11.405 we have that

$$\sum_{(i, j) \in I \times J} (x_i \cdot y_j) \text{ is absolute convergent} \quad (11.79)$$

Next we prove that

$$I \times J = \bigcup_{i \in I} (\{i\} \times J) \text{ and } \forall i, j \in I \vdash \text{we have } \{i\} \times J \bigcap \{j\} \times J = \emptyset \quad (11.80)$$

First if  $(x, y) \in \bigcup_{i \in I} (\{i\} \times J)$  then  $\exists i \in I$  such that  $(x, y) \in \{i\} \times J \Rightarrow x = i \in I \wedge y \in J = (x, y) \in I \times J$ . Second if  $(x, y) \in I \times J$  then  $x \in I \wedge j \in J \Rightarrow (x, y) \in \{x\} \times J \wedge x \in I \Rightarrow (x, y) \in \bigcup_{i \in I} \{i\} \times J$  proving  $I \times J = \bigcup_{i \in I} (\{i\} \times J)$ . Also if  $i \neq j$  and  $(x, y) \in \{i\} \times J \bigcap \{j\} \times J$  then  $i = x = j$  contradicting  $i \neq j$ . this proves our assertion.

Let  $i \in I$  and take  $(x, y), (x', y') \in \{i\} \times J$  such that  $\pi_2((x, y)) = \pi_2((x', y'))$  then  $x = i = x'$  and  $y = y'$  so that  $(x, y), (x', y')$ . Further if  $j \in J$  then  $(i, j) \in \{i\} \times J$  and  $\pi_2((i, j)) = j$ . So we have that

$$\forall i \in I \ \pi_2: \{i\} \times J \rightarrow J \text{ is a bijection} \quad (11.81)$$

$$\begin{aligned} \sum_{\alpha \in I \times J} (x_{\pi_1(\alpha)} \cdot y_{\pi_2(\alpha)}) &\stackrel{11.80 \text{ and associativity (see 11.410)}}{=} \sum_{i \in I} \left( \sum_{\alpha \in \{i\} \times J} (x_{\pi_1(\alpha)} \cdot y_{\pi_2(\alpha)}) \right) \\ &\stackrel{a \in \{i\} \times J \Rightarrow \pi_1(\alpha) = i}{=} \sum_{i \in I} \left( \sum_{\alpha \in \{i\} \times J} (x_i \cdot y_{\pi_2(\alpha)}) \right) \\ &\stackrel{11.409}{=} \sum_{i \in I} \left( x_i \cdot \sum_{\alpha \in \{i\} \times J} y_{\pi_2(\alpha)} \right) \\ &\stackrel{11.81 \text{ together with 11.403}}{=} \sum_{i \in I} \left( x_i \cdot \sum_{j \in J} y_j \right) \\ &\stackrel{11.409}{=} \left( \sum_{i \in I} x_i \right) \cdot \left( \sum_{j \in J} y_j \right) \end{aligned}$$

□

### 11.11.8 Properties of complete spaces

**Definition 11.412.** Let  $\langle S, d \rangle$  be a metric space then a function  $f: S \rightarrow S$  is a **contraction** if  $\exists \lambda \in [0, 1[ \subseteq \mathbb{R}$  such that  $\forall x, y \in S$  we have  $d(f(x), f(y)) \leq \lambda \cdot d(x, y)$

**Lemma 11.413.** Let  $\langle S, d \rangle$  be a metric space and  $f: S \rightarrow S$  a contraction then  $f$  is continuous

**Proof.** Let  $\lambda \in [0, 1[$  be such that  $d(f(x), f(y)) \leq \lambda \cdot d(x, y)$  then given a  $x \in S$  and a  $y \in S$  such that  $d(x, y) < \varepsilon$  we have  $d(f(x), f(y)) \leq \lambda \cdot d(x, y) < \lambda \cdot d(x, y) < d(x, y)$ . □

**Lemma 11.414.** Let  $\lambda \in \mathbb{R}$  such that  $\lambda \neq 1$  then if  $k \in \mathbb{N}_0$  and  $n \in \mathbb{N}$  we have  $\sum_{i=1}^k \lambda^{n+(i-1)} = \lambda^n \cdot \frac{1-\lambda^k}{1-\lambda}$ . So if we take  $n=0$  then  $\sum_{i=1}^k \lambda^{(i-1)} = \frac{1-\lambda^k}{1-\lambda}$  and if  $n=1$  then  $\sum_{i=1}^k \lambda^i = \lambda \cdot \frac{1-\lambda^k}{1-\lambda}$

**Proof.** We proof this by induction so let  $S = \left\{ k \in \mathbb{N}_0 \mid \sum_{i=1}^k \lambda^{n+(i-1)} = \lambda^n \cdot \frac{1-\lambda^k}{1-\lambda} \right\}$  then we have:

1. If  $k=1$  then  $\sum_{i=1}^1 \lambda^{n+(i-1)} = \lambda^n = \lambda^n \cdot \frac{1-\lambda}{1-\lambda} = \lambda^n \cdot \frac{1-\lambda^1}{1-\lambda}$  so that  $1 \in S$
2. Assume that  $k \in S$  then we have

$$\begin{aligned} \sum_{i=1}^{k+1} \lambda^{n+(i-1)} &= \left( \sum_{i=1}^k \lambda^{n+(i-1)} \right) + \lambda^{n+((k+1)-1)} \\ &= \left( \sum_{i=1}^k \lambda^{n+(i-1)} \right) + \lambda^{n+k} \\ &\stackrel{k \in S}{=} \lambda^n \cdot \frac{1-\lambda^k}{1-\lambda} + \lambda^{n+k} \\ &= \frac{\lambda^n \cdot (1-\lambda^k) + (1-\lambda) \cdot \lambda^{n+k}}{1-\lambda} \\ &= \frac{\lambda^n \cdot (1-\lambda^k + (1-\lambda) \cdot \lambda^k)}{1-\lambda} \\ &= \frac{\lambda^n \cdot (1-\lambda^k + \lambda^k - \lambda^{k+1})}{1-\lambda} \\ &= \lambda^n \cdot \frac{1-\lambda^{k+1}}{1-\lambda} \end{aligned}$$

proving that  $k+1 \in S$

□

**Lemma 11.415.** Let  $\lambda \in \mathbb{R}$  with  $0 < \lambda < 1$  then  $\sum_{i=1}^{\infty} \lambda^i$  converges and  $\sum_{i=1}^{\infty} \lambda^i = \frac{\lambda}{1-\lambda}$

**Proof.** Given a  $k \in \mathbb{N}_0$  we have that  $\left| \left( \sum_{i=1}^k \lambda^i \right) - \frac{\lambda}{1-\lambda} \right| = \left| (\lambda \cdot \sum_{i=1}^k \lambda^{i-1}) - \frac{\lambda}{1-\lambda} \right|$  previous lemma with  $n=0$   $\left| \lambda \cdot \left( \lambda^0 \cdot \frac{1-\lambda^k}{1-\lambda} \right) - \frac{\lambda}{1-\lambda} \right| = \left| \frac{\lambda}{1-\lambda} \cdot (1-\lambda^k - 1) \right| = \left| \frac{\lambda^{k+1}}{1-\lambda} \right| = \frac{\lambda}{1-\lambda} \cdot \lambda^k$ . Now using 11.330 given  $\varepsilon > 0$  there exists a  $N \in \{1, \dots, \infty\}$  such that  $\frac{\lambda}{1-\lambda} \cdot \lambda^k = \left| \frac{\lambda}{1-\lambda} \cdot \lambda^k - 0 \right| < \varepsilon$  if  $k \geq N \Rightarrow \left| \left( \sum_{i=1}^k \lambda^i \right) - \frac{\lambda}{1-\lambda} \right| < \varepsilon$  if  $k \geq N$

□

**Theorem 11.416. (The Banach Fixed Point Theorem)** Let  $\langle S, d \rangle$  be a complete metric space and  $f: S \rightarrow S$  any **contraction** then  $\exists! x_0 \in S$  such that  $f(x_0) = x_0$  ( $f$  has exactly one fixed point  $x_0$ ). If  $x \in S$  then  $\lim_{n \rightarrow \infty} f^{(n)}(x) \rightarrow x_0$  where  $f^n$  is recursively defined by

$$\begin{aligned} f^{(1)} &= f \\ f^{(n+1)} &= f \circ f^{(n)} \end{aligned}$$

**Proof.** First as  $f$  is a contraction there exists a  $\lambda \in [0, 1]$  such that  $d(f(x), f(y)) \leq \lambda \cdot d(x, y)$ . As  $S$  is not empty there exists a  $x \in S$ . We prove now by induction that  $\forall n \in \mathbb{N}_0$  we have  $d(f^{(n)}(x), f^{(n+1)}(x)) \leq \lambda^n \cdot d(x, f(x))$ . So take then  $P = \{n \in \mathbb{N}_0 \mid d(f^{(n)}(x), f^{(n+1)}(x)) \leq \lambda^n \cdot d(x, f(x))\}$  then we have:

1. If  $n=1$  then  $d(f^{(n)}(x), f^{(n+1)}(x)) = d(f^{(1)}(x), f^{(2)}(x)) = d(f(x), f(f(x))) \leq \lambda \cdot d(x, f(x)) = \lambda^1 \cdot d(x, f(x))$  proving that  $1 \in P$
2. If  $n \in P$  then we have  $d(f^{(n+1)}(x), f^{((n+1)+1)}(x)) = d(f(f^{(n)}(x)), f(f^{(n+1)}(x))) \leq \lambda \cdot d(f^{(n)}(x), f^{(n+1)}(x)) \leq \lambda \cdot \lambda^n \cdot d(x, f(x)) = \lambda^n \cdot d(x, f(x))$  proving that  $n+1 \in P$

Next we prove by induction that given  $n, k \in \mathbb{N}_0$  that  $d(f^{(n)}(x), f^{(n+k)}(x)) \leq (\sum_{i=1}^k \lambda^{n+(i-1)}) \cdot d(x, f(x))$ . So let  $Q_n = \{k \in \mathbb{N}_0 \mid d(f^{(n)}(x), f^{(n+k)}(x)) \leq (\sum_{i=1}^k \lambda^{n+(i-1)}) \cdot d(x, f(x))\}$  then we have:

1. If  $k=1$  then  $d(f^{(n)}(x), f^{(n+1)}(x)) \leq \lambda^n \cdot d(x, f(x)) = (\sum_{i=1}^1 \lambda^{n+(i-1)}) \cdot d(x, f(x))$  proving that  $1 \in Q_n$

2. If  $k \in Q_n$  then

$$\begin{aligned}
 d(f^{(n)}(x), f^{(n+(k+1))}(x)) &\leq d(f^{(n)}(x), f^{(n+k)}(x)) + d(f^{(n+k)}(x), f^{(n+(k+1))}(x)) \\
 &\leq \left( \sum_{i=1}^k \lambda^{n+(i-1)} \right) \cdot d(x, f(x)) + d(f^{(n+k)}(x), f^{((n+k)+1)}(x)) \\
 &\leq \left( \sum_{i=1}^k \lambda^{n+(i-1)} \right) \cdot d(x, f(x)) + \lambda^{n+k} \cdot d(x, f(x)) \\
 &= \left( \left( \sum_{i=1}^k \lambda^{n+(i-1)} \right) + \lambda^{n+((k+1)-1)} \right) \cdot d(x, f(x)) \\
 &= \left( \sum_{i=1}^{k+1} \lambda^{n+(i-1)} \right) \cdot d(x, f(x))
 \end{aligned}$$

proving that  $k+1 \in Q_n$

So  $d(f^{(n)}(x), f^{(n+k)}(x)) \leq (\sum_{i=1}^{k+1} \lambda^{n+(i-1)}) \cdot d(x, f(x)) \stackrel{11.414}{=} \lambda^n \cdot \frac{1-\lambda^{k+1}}{1-\lambda} \cdot d(x, f(x)) <_{\lambda^k > 0} \frac{\lambda^n}{1-\lambda} \cdot d(x, f(x))$  proving

$$d(f^{(n)}(x), f^{(n+k)}(x)) < \frac{\lambda^n}{1-\lambda} \cdot d(x, f(x)) \quad (11.82)$$

Next we prove that  $\{f^{(n)}(x)\}_{n \in \{1, \dots, \infty\}}$  is Cauchy (using 11.320) so let  $\varepsilon > 0$  be given then using 11.82 and 11.330 there exists a  $N \in \{1, \dots, \infty\}$  such that if  $k \geq N$  then  $d(f^{(n)}(x), f^{(n+k)}(x)) < \varepsilon$  proving the Cauchy property.

As  $S$  is complete there exists a  $x_0 \in S$  such that

$$\lim_{n \rightarrow \infty} f^{(n)}(x) = x_0 \quad (11.83)$$

Using the fact that a contraction mapping is continuous (11.413), 11.83 and 11.334 we have

$$\lim_{n \rightarrow \infty} f^{(n+1)}(x) = \lim_{n \rightarrow \infty} f(f^{(n)}(x)) = f(x_0) \quad (11.84)$$

Assume now that  $f(x_0) \neq x_0$  then  $d(f(x_0), x_0) > 0$  by 11.83 there exists a  $N_1$  such that  $d(f^{(n)}(x_0)) < \frac{d(f(x_0), x_0)}{2}$  if  $n \geq N_1$ , likewise from 11.84 there exists a  $N_2$  such that  $d(f^{(n+1)}(x_0), f(x_0)) < \frac{d(f(x_0), x_0)}{2}$ . Let  $n \geq \max(N_1, N_2 + 1) \Rightarrow n \geq N_1, n - 1 \geq N_2$  then  $d(f(x_0), x_0) \leq d(f(x_0), f^{((n-1)+1)}(x_0)) + d(f^{((n-1)+1)}(x_0), x_0) = d(f(x_0), f^{((n-1)+1)}(x_0)) + d(f^{(n)}(x_0), x_0) < \frac{d(f(x_0), x_0)}{2} + \frac{d(f(x_0), x_0)}{2} = d(f(x_0), x_0)$  giving the contradiction  $d(f(x_0), x_0) < d(f(x_0), x_0)$ . So we must have that  $f(x_0) = x_0$ .

Finally to prove uniqueness assume that there is also a  $x_1 \in S$  such that  $f(x_1) = x_1$ , then if  $x_1 \neq x_0$  then  $0 < d(x_1, x_0) = d(f(x_0), f(x_1)) < \lambda d(x_0, x_1) < d(x_0, x_1)$  as  $\lambda \in [0, 1[ \Rightarrow d(x_1, x_0) < d(x_1, x_0)$  a contradiction so  $x_0 = x_1$   $\square$

**Lemma 11.417.** If  $\{A_i\}_{i \in \mathbb{N}_0}$  is a family of sets such that  $\forall i \in \mathbb{N}_0 \setminus \{1\}$  we have that  $A_i \subseteq A_{i-1}$  then for every  $n, m \in \mathbb{N}_0$  with  $n \leq m$  we have  $A_m \subseteq A_n$

**Proof.** This is easy proved by induction. So let  $S_n = \{k \in \mathbb{N} \mid A_{n+k} \subseteq A_n\}$  then we have

1. If  $k = 0$  then  $A_{n+k} = A_n \subseteq A_n$  proving that  $0 \in S_n$
2. If  $k \in S_n$  then we have that  $A_{n+(k+1)} = A_{(n+k)+1} \subseteq A_{((n+k)+1)-1} = A_{n+k} \subseteq A_n \Rightarrow k+1 \in S_n$

Using induction we have then that  $S_n = \mathbb{N}$ . If now  $n \leq m$  then  $m - n \in \mathbb{N} = S_n$  so that  $A_m = A_{n+(m-n)} \subseteq A_n$   $\square$

**Lemma 11.418.** Let  $\langle X, d \rangle$  be a complete metric space and let  $\{A_i\}_{i \in \mathbb{N}_0}$  be a sequence of non empty closed sets with  $\forall i \in \mathbb{N}_0 \setminus \{1\}$  we have  $A_i \subseteq A_{i-1}$  and  $\lim_{n \rightarrow \infty} \text{diam}(A_i) = 0$  then  $\bigcap_{i \in \mathbb{N}_0} A_i$  is not empty.

**Proof.** Using the axiom of choice (see 2.210) there exists a function  $c: \mathbb{N}_0 \rightarrow \bigcup_{i \in \mathbb{N}_0} A_i$  such that  $\forall i \in \mathbb{N}_0$  we have  $c_i \in A_i$  defining the family  $\{c(i)\}_{i \in \mathbb{N}_0}$ . Now if  $\varepsilon > 0$  we use  $\lim_{n \rightarrow \infty} \text{diam}(A_i) = 0$  to find a  $N \in \mathbb{N}_0$  such that  $\forall n \geq N$  we have  $\text{diam}(A_n) = |\text{diam}(A_n) - 0| < \varepsilon$ . If now  $n, m \geq N$  then using the above lemma we have that  $A_n, A_m \subseteq A_N$  so that  $c(n), c(m) \in A_N$  and thus  $d(c(n), c(m)) \leq \text{diam}(A_n) < \varepsilon$  proving that  $\{c(i)\}_{i \in \mathbb{N}_0}$  is Cauchy. As  $X$  is complete there exists a  $x \in X$  such that  $\lim_{i \rightarrow \infty} c(i) = x$ . If now  $n \in \mathbb{N}_0$  and  $U$  a open set with  $x \in U$  then there exists a  $\varepsilon > 0$  such that  $x \in B_d(x, \varepsilon) \subseteq U$ . From the definition of  $x$  there exists then a  $N \in \mathbb{N}_0$  such that if  $m \geq N$  we have  $d(x, c(m)) < \varepsilon \Rightarrow c(m) \in B_d(x, \varepsilon) \subseteq U$ . So if we take  $q = \max(N, n)$  then  $q \geq n \Rightarrow c(q) \in A_n$  and  $q \geq N \Rightarrow c(q) \in U$  proving that  $A_n \cap U \neq \emptyset$  so that  $x \in A = \overline{A_n}_{A \text{ is closed}} = A$ . As  $n$  was chosen arbitrary we have that  $\forall n \in \mathbb{N}_0$  that  $x \in A_n$  or  $x \in \bigcap_{n \in \mathbb{N}_0} A_n$ .  $\square$

**Theorem 11.419. (Baire Category Theorem)** *If  $X$  is a compact Hausdorff topological space or if  $\langle X, d \rangle$  is a complete metric space then  $X$  is a Baire space (11.42).*

**Proof.** First note that because of 11.231 and 11.258 we have that  $X$  is regular in both cases. Let  $\mathcal{T}$  be the topology of  $\langle X, d \rangle$  and take then  $\mathcal{A} = \{U \in \mathcal{T} | U \neq \emptyset\}$ . Let  $\{A_i\}_{i \in \mathbb{N}_0}$  be a sequence of closed sets of  $X$  with  $\forall i \in \mathbb{N}_0$  we have  $A_i^\circ = \emptyset$ , we must show then that  $\left(\bigcup_{i \in \mathbb{N}_0} A_i\right)^\circ = \emptyset$ . Take  $U_0 \in \mathcal{A}$  a non empty open set arbitrarily.

Consider the set  $A_1$  then because  $A_1^\circ = \emptyset$  we have  $U_0 \not\subseteq A_1 \Rightarrow \exists y \in U_0$  such that  $y \notin A_1$ , using regularity there exists open sets  $V_1, U'_1$  such that  $y \in U'_1 \wedge A_1 \subseteq V_1 \wedge U'_1 \cap V_1 = \emptyset$ . From  $U'_1 \cap V_1 = \emptyset$  we deduce that  $U'_1 \subseteq X \setminus V_1$  which is closed so that  $\overline{U'_1} \subseteq X \setminus V_1$  and thus  $\overline{U'_1} \cap V_1 = \emptyset \Rightarrow \overline{U'_1} \cap A_1 \subseteq \overline{U'_1} \cap V_1 = \emptyset$ . Using 11.232 there exists a  $U''_1$  such that  $y \in U''_1$  and  $\overline{U''_1} \subseteq U_0$ . If we take now  $U_1 = U'_1 \cap U''_1 \cap U_0 \ni y \Rightarrow U_1 \in \mathcal{A}$  [in case of a complete metric space we take  $U_1 = U'_1 \cap U''_1 \cap B_d(y, \frac{1}{2}) \ni y \Rightarrow U_1 \in \mathcal{A}$ ] giving

$$\overline{U_1} \cap A_1 = \emptyset \wedge \overline{U_1} \subseteq U_0 \text{ and in case of a metric space } \text{diam}(\overline{U_1}) \leq 1 \quad (11.85)$$

Define now  $\mathcal{N}_1 = \{\{U_i\}_{i \in \{1, \dots, n\}} | n \in \mathbb{N}_0 \wedge \forall i \in \{1, \dots, n\} \text{ we have } U_i \in \mathcal{A} \wedge \text{we have } \overline{U_i} \cap A_i = \emptyset \wedge \overline{U_i} \subseteq U_{i-1} \text{ and in case of a metric space } \text{diam}(\overline{U_i}) \leq \frac{1}{i}\} \subseteq \mathcal{M} = \{\{U_i\}_{i \in \{0, \dots, n\}} | n \in \mathbb{N} \wedge \forall i \in \{0, \dots, n\} \text{ we have } U_i \in \mathcal{A}\}$ . Using 11.85 we have then

$$\{U_1\}_{i \in \{1, \dots, 1\}} \in \mathcal{N}_1 \quad (11.86)$$

Define now  $\rho: \mathcal{N}_1 \rightarrow \mathcal{A}$  as follows. Take  $\{U_i\}_{i \in \{1, \dots, n\}} \in \mathcal{N}_1$  then as  $U_n$  is not empty and  $A_{n+1}^\circ = \emptyset$  we can not have  $U_n \subseteq A_{n+1} \Rightarrow \exists y \in U_n$  such that  $y \in A_{n+1}$ , using regularity we can find a  $U'_{n+1}, V_{n+1} \in \mathcal{T}$  such that  $y \in U'_{n+1}, A_{n+1} \subseteq V_{n+1}$  and  $U'_{n+1} \cap V_{n+1} = \emptyset$ . So  $U'_{n+1} \subseteq X \setminus V_{n+1}$  a closed set so that  $\overline{U'_{n+1}} \subseteq X \setminus V_{n+1}$  giving  $\overline{U'_{n+1}} \cap V_{n+1} = \emptyset$ . Using 11.232 there exists a  $U''_{n+1} \in \mathcal{T}$  with  $y \in U''_{n+1}$  and  $\overline{U''_{n+1}} \subseteq U'_{n+1}$ . Taken now  $\rho(\{U_i\}_{i \in \{1, \dots, n\}}) = U'_{n+1} \cap U''_{n+1} \cap U_n \ni y$  [in case of a metric space we take  $\rho(\{U_i\}_{i \in \{1, \dots, n\}}) = U'_{n+1} \cap U''_{n+1} \cap U_n \cap B_d(y, \frac{1}{2 \cdot (n+1)}) \ni y$  so that we have that

$$\rho(\{U_i\}_{i \in \{1, \dots, n\}}) \in \mathcal{A} \wedge \overline{\rho(\{U_i\}_{i \in \{1, \dots, n\}})} \cap A_{n+1} = \emptyset \wedge \overline{\rho(\{U_i\}_{i \in \{1, \dots, n\}})} \subseteq U_n \text{ and in case of a metric space } \text{diam}(\overline{\rho(\{U_i\}_{i \in \{1, \dots, n\}})}) \leq \frac{1}{n} \quad (11.87)$$

So if we define  $\{U'_i\}_{i \in \{1, \dots, n+1\}}$  by  $U'_i = \begin{cases} U_i & \text{if } i \in \{1, \dots, n\} \\ \rho(\{U_i\}_{i \in \{1, \dots, n\}}) & \text{if } i = n+1 \end{cases}$  then we have

$$\{U'_i\}_{i \in \{1, \dots, n+1\}} \in \mathcal{N}_1 \quad (11.88)$$

Using 5.26 there exists a function  $U: \mathbb{N}_0 \rightarrow \mathcal{A}$  such that

1.  $U(1) = U_1$
2.  $\forall n \in \mathbb{N}_0$  we have  $U(n+1) = \rho(\{U(i)\}_{i \in \{1, \dots, n\}})$
3.  $\forall n \in \mathbb{N}_0$  we have  $\{U(i)\}_{i \in \{1, \dots, n\}} \in \mathcal{N}_1$

This defines the sequence  $\{U(i)\}_{i \in \mathbb{N}_0}$  in  $\mathcal{A}$  (the set of non empty open sets) so that:

$$\forall i \in \mathbb{N}_0 \text{ we have } \overline{U(i)} \cap A_i = \emptyset, \overline{U_i} \subseteq U_{i-1}, \text{ in case of a metric space } \text{diam}(\overline{U(i)}) \leq \frac{1}{i} \quad (11.89)$$

We prove now that  $\bigcap_{i \in \mathbb{N}_0} \bar{U}_i \neq \emptyset$  consider the two cases:

1. ( **$X$  is compact Hausdorff**) Then we prove by induction that  $\emptyset \neq \bar{U}_n \subseteq \bigcap_{i \in \{1, \dots, n\}} \bar{U}_i$ . So let  $\mathcal{S} = \{n \in \mathbb{N}_0 \mid \bar{U}_n \subseteq \bigcap_{i \in \{1, \dots, n\}} \bar{U}_i\}$  then we have:

a. If  $n = 1$  then  $\bar{U}_1 \subseteq \bar{U}_1 = \bigcap_{i \in \{1, \dots, 1\}} \bar{U}_i$  so that  $1 \in \mathcal{S}$

b. If  $n \in \mathcal{S}$  then  $\bar{U}_{n+1} \subseteq U_n \subseteq \bar{U}_n \Rightarrow \bar{U}_{n+1} \subseteq \bar{U}_n \cap \bar{U}_{n+1} \subseteq (\bigcap_{i \in \{1, \dots, n\}} \bar{U}_i) \cap \bar{U}_{n+1} = \bigcap_{i \in \{1, \dots, n+1\}} \bar{U}_i$  proving that  $n+1 \in \mathcal{S}$ .

Mathematical induction proves then that  $S = \mathbb{N}_1$ . This proves that  $\{\bar{U}_i\}_{i \in \mathbb{N}_0}$  has the finite intersection property so that by 11.296 we have that  $\bigcap_{i \in \mathbb{N}_0} \bar{U}_i \neq \emptyset$

2. ( **$\langle X, d \rangle$  is a metric space**) As  $\text{diam}(\bar{U}_i) \leq \frac{1}{i}$  we have that  $\lim_{i \rightarrow \infty} \text{diam}(\bar{U}_i) = 0$  so by the previous lemma and  $\bar{U}_i \subseteq U_{i-1} \subseteq \bar{U}_{i-1}$  we have again that  $\bigcap_{i \in \mathbb{N}_0} \bar{U}_i \neq \emptyset$ .

Take now  $x \in \bigcap_{i \in \mathbb{N}_0} \bar{U}_i$  then  $x \in \bar{U}_1 \subseteq U_0 \Rightarrow x \in U_0$  and  $\forall i \in \mathbb{N}_0$  we have  $\bar{U}_i \cap A_i = \emptyset$  so that  $x \notin A_i$  so that  $x \notin \bigcup_{i \in \mathbb{N}_0} A_i$ . As  $U_0$  was chosen arbitrary we have proved that  $(\bigcup_{i \in \mathbb{N}_0} A_i)^\circ = \emptyset$  for if  $y \in (\bigcup_{i \in \mathbb{N}_0} A_i)^\circ$  then  $(\bigcup_{i \in \mathbb{N}_0} A_i)^\circ \in \mathcal{A}$  and we have then that  $y \notin \bigcup_{i \in \mathbb{N}_0} A_i$  contradicting the fact that  $y \in (\bigcup_{i \in \mathbb{N}_0} A_i)^\circ \subseteq \bigcup_{i \in \mathbb{N}_0} A_i \Rightarrow y \in \bigcup_{i \in \mathbb{N}_0} A_i$ .  $\square$

We use the Baire category theorem to prove the open mapping theorem.

**Theorem 11.420. (Open Mapping Theorem)** *Let  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle$  be Banach spaces over  $\mathbb{K}$  and  $L: X \rightarrow Y$  a surjective continuous linear function then  $L$  is also a open function.*

**Proof.** First if  $\|L\| = 0$  then we have  $\forall x \in X$  that  $\|L(x)\|_Y \leq \|L\| \cdot \|x\| = 0 \Rightarrow L(x) = 0$  so as  $L$  is a surjection we have that  $Y = \{0\}$  and thus also  $X = \{0\}$  so that the only open set in  $X$  is  $\{0\}$  and  $L(\{0\}) = \{0\} = Y$  is open proving that  $L$  is a open function. Let's now take the case that  $\|L\| > 0$ . We split the proof in little sub-proofs.

1. First we show that given  $\delta > 0$  and  $B_{\|\cdot\|_X}(0, \delta)$  a open ball around 0 then there exists  $\lambda > 0$  with  $B_{\|\cdot\|_Y}(0, \lambda) \subseteq \overline{L(B_{\|\cdot\|_X}(0, \delta))}$

**Proof.** Take  $\delta > 0$  and  $x \in X$  then given the Archimedean order property of  $\mathbb{R}$  (see 9.55) there exists a  $n \in \mathbb{N}_0$  such that  $\frac{2}{\delta} \cdot \|x\| < n \cdot 1 = n$  so that  $\|x\| < \frac{n \cdot \delta}{2}$  so if  $y = \frac{x}{n} \Rightarrow x = y \cdot n \Rightarrow \|n \cdot y\| < \frac{n \cdot \delta}{2} \Rightarrow \|y\| < \frac{\delta}{2} \Rightarrow x \in n \cdot B_{\|\cdot\|_X}\left(0, \frac{\delta}{2}\right) \Rightarrow x \in \bigcup_{n \in \mathbb{N}_0} \left(n \cdot B_{\|\cdot\|_X}\left(0, \frac{\delta}{2}\right)\right)$  proving that

$$X = \bigcup_{n \in \mathbb{N}_0} n \cdot B_{\|\cdot\|_X}\left(0, \frac{\delta}{2}\right) \quad (11.90)$$

As  $L$  is surjective we have that  $Y = L(X) = \bigcup_{n \in \mathbb{N}_0} L\left(n \cdot B_{\|\cdot\|_X}\left(0, \frac{\delta}{2}\right)\right) = \bigcup_{n \in \mathbb{N}_0} n \cdot L\left(B_{\|\cdot\|_X}\left(0, \frac{\delta}{2}\right)\right) \subseteq \bigcup_{n \in \mathbb{N}_0} n \cdot \overline{L\left(B_{\|\cdot\|_X}\left(0, \frac{\delta}{2}\right)\right)}$  giving

$$Y = \bigcup_{n \in \mathbb{N}_0} n \cdot \overline{L\left(B_{\|\cdot\|_X}\left(0, \frac{\delta}{2}\right)\right)} \quad (11.91)$$

Note that that  $0 \in Y = Y^\circ = \overline{\left(\bigcup_{n \in \mathbb{N}_0} n \cdot \overline{L\left(B_{\|\cdot\|_X}\left(0, \frac{\delta}{2}\right)\right)}\right)} \Rightarrow \bigcup_{n \in \mathbb{N}_0} n \cdot \overline{L\left(B_{\|\cdot\|_X}\left(0, \frac{\delta}{2}\right)\right)} \neq \emptyset$  we must as  $Y$  is a Banach space (thus a complete metric space) using the Baire Category Theorem (11.419) there must exists a  $n \in \mathbb{N}_0$  such that  $\left(\overline{n \cdot L\left(B_{\|\cdot\|_X}\left(0, \frac{\delta}{2}\right)\right)}\right)^\circ \neq \emptyset$ . So  $\exists z \in \left(\overline{n \cdot L\left(B_{\|\cdot\|_X}\left(0, \frac{\delta}{2}\right)\right)}\right)^\circ$  which is open so there exists a  $\beta > 0$  such that  $z \in B_{\|\cdot\|_Y}(z, \beta) \subseteq \left(\overline{n \cdot L\left(B_{\|\cdot\|_X}\left(0, \frac{\delta}{2}\right)\right)}\right)^\circ \subseteq \overline{n \cdot L\left(B_{\|\cdot\|_X}\left(0, \frac{\delta}{2}\right)\right)} = \overline{n \cdot L\left(B_{\|\cdot\|_X}\left(0, \frac{\delta}{2}\right)\right)}$  so that

$$\frac{1}{n} \cdot B_{\|\cdot\|_Y}(z, \beta) \subseteq \overline{L\left(B_{\|\cdot\|_X}\left(0, \frac{\delta}{2}\right)\right)} \quad (11.92)$$

As  $B_{\|\cdot\|_X}\left(0, \frac{\delta}{2}\right) - B_{\|\cdot\|_X}\left(0, \frac{\delta}{2}\right) \subseteq B_{\|\cdot\|_Y}(0, \delta)$  [If  $x \in B_{\|\cdot\|_X}\left(0, \frac{\delta}{2}\right) - B_{\|\cdot\|_X}\left(0, \frac{\delta}{2}\right) \Rightarrow x = x_1 = x_2$  with  $\|x_1\|_X < \frac{\delta}{2}$ ,  $\|x_2\|_X < \frac{\delta}{2} \Rightarrow \|x\|_X \leq \|x_1\|_X + \|x_2\|_X < \frac{\delta}{2} + \frac{\delta}{2} = \delta$ ] we have  $L\left(B_{\|\cdot\|_X}\left(0, \frac{\delta}{2}\right)\right) - L\left(B_{\|\cdot\|_X}\left(0, \frac{\delta}{2}\right)\right) = L\left(B_{\|\cdot\|_X}\left(0, \frac{\delta}{2}\right) - B_{\|\cdot\|_X}\left(0, \frac{\delta}{2}\right)\right) = L\left(B_{\|\cdot\|_X}\left(0, \frac{\delta}{2}\right) - B_{\|\cdot\|_X}\left(0, \frac{\delta}{2}\right)\right) \subseteq L(B_{\|\cdot\|_X}(0, \delta))$  giving

$$L\left(B_{\|\cdot\|_X}\left(0, \frac{\delta}{2}\right)\right) - L\left(B_{\|\cdot\|_X}\left(0, \frac{\delta}{2}\right)\right) \subseteq L(B_{\|\cdot\|_X}(0, \delta)) \quad (11.93)$$

Further  $\bigcup_{x \in \frac{1}{n}B_{\|\cdot\|_Y}(z, \beta)} (x - \frac{1}{n}B_{\|\cdot\|_Y}(z, \beta)) \subseteq \frac{1}{n}B_{\|\cdot\|_Y}(z, \beta) - \frac{1}{n}B_{\|\cdot\|_Y}(z, \beta) \subseteq_{11.92} \frac{L\left(B_{\|\cdot\|_X}\left(0, \frac{\delta}{2}\right)\right)}{L\left(B_{\|\cdot\|_X}\left(0, \frac{\delta}{2}\right)\right)} - \frac{L\left(B_{\|\cdot\|_X}\left(0, \frac{\delta}{2}\right)\right)}{L\left(B_{\|\cdot\|_X}\left(0, \frac{\delta}{2}\right)\right)} \subseteq_{11.79} L\left(B_{\|\cdot\|_X}\left(0, \frac{\delta}{2}\right)\right) - L\left(B_{\|\cdot\|_X}\left(0, \frac{\delta}{2}\right)\right) \subseteq_{11.93} \overline{L(B_{\|\cdot\|_X}(0, \delta))}$  so that we have for  $U = \bigcup_{x \in \frac{1}{n}B_{\|\cdot\|_Y}(z, \beta)} (x - \frac{1}{n}B_{\|\cdot\|_Y}(z, \beta))$

$$U = \bigcup_{x \in \frac{1}{n}B_{\|\cdot\|_Y}(z, \beta)} \left(x - \frac{1}{n}B_{\|\cdot\|_Y}(z, \beta)\right) \subseteq \overline{L(B_{\|\cdot\|_X}(0, \delta))} \quad (11.94)$$

Now using 11.79 again we have that  $x - \frac{1}{n}B_{\|\cdot\|_Y}(z, \beta)$  is open so that  $U$  being the union of open sets must be open, as  $0 = z - z \in U$  there exists a  $\lambda > 0$  such that  $0 \in B_{\|\cdot\|_Y}(0, \lambda) \subseteq U$  or using 11.94

$$B_{\|\cdot\|_Y}(0, \lambda) \subseteq \overline{L(B_{\|\cdot\|_X}(0, \delta))} \quad (11.95)$$

what we set out to prove.  $\square$

2. Next we refine (1) to show that

$$\text{If } \sigma_0 > 0 \text{ then } \exists \lambda > 0 \text{ such that } B_{\|\cdot\|_Y}(0, \lambda) \subseteq L(B_{\|\cdot\|_X}(0, 2 \cdot \sigma_0)) \quad (11.96)$$

**Proof.** Using the Archimedean property of the reals (see 9.56) we find a  $N \in \mathbb{N}_0$  such that  $\frac{1}{N} < \sigma_0$ . Define then  $\forall n \in \mathbb{N}_0 \sigma_n = \left(\frac{1}{N+1}\right)^n$ , we have then that  $\sum_{i=1}^{\infty} \sigma^i = \sum_{i=1}^{\infty} \left(\frac{1}{N+1}\right)^i \stackrel{\frac{1}{N+1} < 1 \text{ and 11.415}}{=} \frac{\frac{1}{N+1}}{1 - \frac{1}{N+1}} = \frac{\frac{1}{N+1}}{\frac{N+1-1}{N+1}} = \frac{\frac{1}{N+1}}{\frac{N}{N+1}} = \frac{1}{N} < \sigma_0$  or

$$\sum_{i=1}^{\infty} \sigma_i = \frac{1}{N} < \sigma_0 \quad (11.97)$$

Now  $\forall n \in \mathbb{N} = \mathbb{N}_0 \cup \{0\}$  we have using (1) the existence of a  $\lambda'_n > 0$  such that  $B_{\|\cdot\|_Y}(0, \lambda'_n) \subseteq \overline{L(B_{\|\cdot\|_X}(0, \sigma_n))}$ . Define  $0 < \lambda_n = \min\left(\lambda'_n, \frac{1}{n+1}\right) < \frac{1}{n+1}$  then trivially we have  $B_{\|\cdot\|_Y}(0, \lambda_n) \subseteq \overline{L(B_{\|\cdot\|_X}(0, \sigma_n))}$  and using 11.330 and 11.350 we have that  $\lim_{n \rightarrow \infty} \lambda_n = 0$ . So we have found a sequence  $\{\lambda_i\}_{i \in \mathbb{N}}$  such that

$$\lim_{i \rightarrow \infty} \lambda_i \wedge \forall n \in \mathbb{N} \text{ we have } B_{\|\cdot\|_Y}(0, \lambda_n) \subseteq \overline{L(B_{\|\cdot\|_X}(0, \sigma_n))} \quad (11.98)$$

Taken now  $y \in B_{\|\cdot\|_Y}(0, \sigma_0)$ . Then  $y \in B_{\|\cdot\|_Y}(0, \lambda_0) \Rightarrow y \in \overline{L(B_{\|\cdot\|_X}(0, \sigma_0))}$  so there exists a  $y_0 \in B_{\|\cdot\|_Y}(0, \lambda_1) \cap L(B_{\|\cdot\|_X}(0, \sigma_0))$  [as  $y \in B_{\|\cdot\|_Y}(0, \lambda_1)$  a open set] thus  $\exists x_0 \in B_{\|\cdot\|_X}(0, \sigma_0)$  such that  $y_0 = L(x_0) \in B_{\|\cdot\|_Y}(0, \lambda_1)$ . This gives

$$\exists x_0 \in B_{\|\cdot\|_X}(0, \sigma_0) \text{ with } \|y - L(x_0)\|_Y < \lambda_1 \quad (11.99)$$

Take now  $\mathcal{N}_0 = \{\{x_i\}_{i \in \{0, \dots, n\}} | n \in \mathbb{N} \wedge \forall i \in \{0, \dots, n\} \text{ we have } x_i \in B_{\|\cdot\|_X}(0, \sigma_i) \wedge \|y - L(\sum_{j=0}^i x_j)\|_Y < \lambda_{i+1}\}$  then we have

$$\{x_0\}_{i \in \{0, \dots, n\}} \in \mathcal{N}_0 \quad (11.100)$$

If  $\{x_i\}_{i \in \{1, \dots, n\}} \in \mathcal{N}_0$  then as  $\|y - L(\sum_{j=0}^n x_i)\|_Y < \lambda_{n+1}$  we have that  $y - L(\sum_{j=0}^n x_j) \in B_{\|\cdot\|_Y}(0, \lambda_{n+1}) \subseteq \overline{L(B_{\|\cdot\|_X}(0, \sigma_{n+1}))}$  so that there exists a  $y_{n+1} \in B_{\|\cdot\|_Y}(y - L(\sum_{j=0}^n x_j), \lambda_{n+2}) \cap L(B_{\|\cdot\|_X}(0, \sigma_{n+1})) \Rightarrow \exists x_{n+1} \in B_{\|\cdot\|_X}(0, \sigma_{n+1})$  such that  $y_{n+1} = L(x_{n+1}) \in B_{\|\cdot\|_Y}(y - L(\sum_{j=0}^n x_j), \lambda_{n+2})$ . Then  $\|y - L(\sum_{j=0}^{n+1} x_j)\|_Y = \|(y - L(\sum_{j=0}^n x_j)) - L(x_{n+1})\|_Y < \lambda_{n+1} = \lambda_{(n+1)+1}$ . This defines a function  $\rho: \mathcal{N}_0 \rightarrow X$  such that  $x_{n+1} = \rho(\{x_i\}_{i \in \{0, \dots, n\}})$  and  $\{x_i\}_{i \in \{0, \dots, n+1\}} \in \mathcal{N}_0$ . Using recursion (see 5.26) this defines a sequence (or function from  $\mathbb{N} \rightarrow X$ )  $\{x_i\}_{i \in \mathbb{N}}$  such that

$$\forall i \in \mathbb{N} \text{ we have } x_i \in B_{\|\cdot\|_X}(0, \sigma_i) \wedge \left\| y - L\left(\sum_{j=0}^i x_j\right) \right\|_Y < \lambda_{i+1} \quad (11.101)$$

As  $x_i \in B_{\|\cdot\|_X}(0, \sigma_i)$  we have  $\|x_i\|_X < \sigma_i$  so that using 11.393 we have that  $\sum_{i=0}^{\infty} x_i$  converges and  $\|\sum_{i=0}^{\infty} x_i\|_X \leq \sum_{i=0}^{\infty} \sigma_i \stackrel{11.379}{=} \sigma_0 + \sum_{i=1}^{\infty} \sigma_i < \stackrel{11.97}{\sigma} + \sigma = 2 \cdot \sigma$  so if  $x = \sum_{i=0}^{\infty} x_i$  then  $\|x\|_X < 2 \cdot \sigma$  and thus

$$x = \sum_{j=0}^{\infty} x_j \in B_{\|\cdot\|_X}(0, 2 \cdot \sigma_0) \quad (11.102)$$

Assume now that  $\|y - L(x)\|_Y > 0$  then we can take  $\varepsilon = \|y - L(x)\|_Y > 0$  so that as  $\lim_{i \rightarrow \infty} \lambda_i = 0$  and  $0 < \lambda_i$  that there exists a  $N_1 \in \mathbb{N}$  such that if  $n \geq N_1$  then  $\lambda_n = |\lambda_n - 0| < \frac{\varepsilon}{2}$ . As  $x = \sum_{i=0}^{\infty} x_i$  there exists a  $N_2 \in \mathbb{N}$  such that if  $n \geq N_2$  then  $\|x - \sum_{i=0}^n x_i\|_X < \frac{\varepsilon}{2 \cdot \|L\|}$ . If  $n \geq \max(N_1, N_2)$  then  $\varepsilon = \|y - L(x)\|_Y = \|y - L(\sum_{i=0}^n x_i) + L(\sum_{i=0}^n x_i) - L(x)\|_Y \leq \|y - L(\sum_{i=0}^n x_i)\|_Y + \|L(\sum_{i=0}^n x_i) - L(x)\|_Y = \|y - L(\sum_{i=0}^n x_i)\|_Y + \|L(\sum_{i=0}^n x_i - x)\|_Y < \lambda_{n+1} + \|L\| \cdot \|\sum_{i=0}^n x_i - x\| < \frac{\varepsilon}{2} + \|L\| \cdot \frac{\varepsilon}{2 \cdot \|L\|} = \varepsilon$  giving the contradiction  $\varepsilon < \varepsilon$ . So we must have that  $\|y - L(x)\|_Y = 0 \Rightarrow y = L(x) \stackrel{11.102}{\Rightarrow} y \in L(B_{\|\cdot\|_X}(0, 2 \cdot \sigma_0))$ . As  $y \in B_{\|\cdot\|_Y}(0, \sigma_0)$  was chosen arbitrary we have that  $B_{\|\cdot\|_Y}(0, \sigma_0) \subseteq L(B_{\|\cdot\|_X}(0, 2 \cdot \sigma_0))$ , which taking  $\lambda = \sigma_0$  proves our assertion.

3. Now take  $B_{\|\cdot\|_X}(0, 1) = B_{\|\cdot\|_X}(0, 2 \cdot \frac{1}{2})$  we find using (2) the existence of a  $\lambda > 0$  such that  $B_{\|\cdot\|_Y}(0, \lambda) \subseteq L(B_{\|\cdot\|_X}(0, 1))$  which using 11.193 means that  $L$  is open.  $\square$   $\square$

**Corollary 11.421.** Let  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle$  be Banach spaces over  $\mathbb{K}$  and  $L: X \rightarrow Y$  a continuous linear isomorphism then  $L^{-1}$  is continuous.

**Proof.** This is trivial using the open function theorem (11.420) and 11.145.  $\square$

## 11.12 Integration in Banach space

**Note 11.422.** In the following we assume that with  $[a, b]$  we mean a closed and bounded interval in  $\mathbb{R}$  that is not empty (or  $a < b$ )

**Definition 11.423.**  $\mathbb{N}_1 = \{1, \dots\} = \mathbb{N}_0 \setminus \{1\}$

**Definition 11.424.** A partition  $\mathcal{P}$  of  $[a, b]$  is a family  $\{t_i\}_{i \in \{1, \dots, n\}}$   $n \in \mathbb{N}_1$  such that  $t_1 = a, t_n = b$  and  $\forall i \in \{1, \dots, n-1\}$  we have  $t_i < t_{i+1}$ .

We prove now some trivial facts for a partition of a interval

**Theorem 11.425.** Let  $[a, b] \subseteq \mathbb{R}$  and  $\mathcal{P} = \{t_i\}_{i \in \{1, \dots, n\}}$  a partition of  $[a, b]$  then  $[a, b] = \bigcup_{i \in \{1, \dots, n-1\}} [t_i, t_{i+1}]$

**Proof.** As  $\forall i \in \{1, \dots, n\}$  we have  $t_i \in [a, b] \Rightarrow a \leq t_i \leq b$  so that  $\forall i \in \{1, \dots, n-1\}$  we have  $a \leq t_i < t_{i+1} \leq b \Rightarrow [t_i, t_{i+1}] \subseteq [a, b] \Rightarrow (\bigcup_{i \in \{1, \dots, n-1\}} [t_i, t_{i+1}]) \subseteq [a, b]$ . Take now  $x \in [a, b]$  and take  $M_x = \{i \in \{1, \dots, n-1\} | t_i \leq x\}$  which is finite and as  $t_1 = a \leq x \Rightarrow 1 \in M_x$  is not empty, so  $m = \max(M_x)$  exist. As  $m \in M_x$  we have  $t_m \leq x$ , if  $t_{m+1} < x \Rightarrow t_{m+1} \leq x \Rightarrow m+1 \in M_x \Rightarrow m+1 \leq m$  a contradiction so  $x \leq t_{m+1}$ . So  $x \in [t_m, t_{m+1}] \subseteq \bigcup_{i \in \{1, \dots, n-1\}} [t_i, t_{i+1}] \Rightarrow [a, b] \subseteq \bigcup_{i \in \{1, \dots, n-1\}} [t_i, t_{i+1}]$   $\square$

**Lemma 11.426.** Let  $\{t_i\}_{i \in \{1, \dots, n\}}$ ,  $n \in \mathbb{N}$  be such that  $\forall i \in \{1, \dots, n-1\}$  we have  $t_i < t_{i+1}$  then

1.  $\forall i, j \in \{1, \dots, n\}$  with  $i < j$  we have  $t_i < t_j$
2. If  $t_i = t_j$  then  $i = j$
3. If  $i \in \{1, \dots, n-1\}$ ,  $k \in \{1, \dots, n\}$  and  $t_k \in [t_i, t_{i+1}]$  then  $k = i$  or  $k = i+1$

**Proof.**

1. We prove this by induction, so given  $i \in \{1, \dots, n\}$  let  $\mathcal{S}_i = \{k \in \mathbb{N}_0 \mid \text{if } i+k \leq n \text{ then } t_i < t_{i+k}\}$  then:

- a. If  $k = 1$  then if  $i+k = i+1 \leq n$  we have  $i \in \{1, \dots, n-1\}$  and by the hypothesis we have  $t_i < t_{i+1} = t_{i+k}$  so  $1 \in \mathcal{S}_i$
- b. If  $k \in \mathcal{S}_i$  then for  $k+1$  we have if  $i+(k+1) \leq n \Rightarrow (i+k)+1 \leq n$  that  $t_{i+k} < t_{(i+k)+1}$  (hypothesis) and  $t_i < t_{i+k}$  ( $k \in \mathcal{S}_i$ ) so that  $t_i < t_{(i+k)+1} = t_{i+(k+1)}$  proving that  $k+1 \in \mathcal{S}_i$

Using induction we have that  $\mathcal{S}_i = \mathbb{N}_0$  so if  $i, j \in \{1, \dots, n\}$  with  $i < j$  then  $k = j - i \in \mathbb{N}_0 \in \mathcal{S}_i$  and  $i+k = j \leq n$  so that  $t_i < t_{i+k} = t_j$ .

2. If  $t_i = t_k$  then if  $i < k$  we have by (1)  $t_i < t_k$  a contradiction, if  $k < i$  we have by (1)  $t_k < t_i$  a contradiction. So the only possibility left is  $i = k$
3. If  $t_k \in [t_i, t_{i+1}]$  then  $t_i \leq t_k \leq t_{i+1}$ . If now  $k < i$  then using (1) we have  $t_k < t_i$  contradicting  $t_i \leq t_k$ , if  $i+1 < k$  we have using (1)  $t_{i+1} < t_k$  contradicting  $t_k \leq t_{i+1}$ . So we must have  $i \leq k \leq i+1$  which means  $i = k$  or  $k = i+1$   $\square$

**Theorem 11.427.** Let  $\mathcal{P}_1 = \{t_i^1\}_{i \in \{1, \dots, n_1\}}$ ,  $\mathcal{P}_2 = \{t_i^2\}_{i \in \{1, \dots, n_2\}}$  be two partitions of  $[a, b]$  then there exists a unique partition  $\mathcal{P} = \{t_i\}_{i \in \{1, \dots, n\}}$  of  $[a, b]$  such that  $\{t_i \mid i \in \{1, \dots, n\}\} = \{t_i^1 \mid i \in \{1, \dots, n_1\}\} \cup \{t_i^2 \mid i \in \{1, \dots, n_2\}\}$ . This partition of  $[a, b]$  is noted by  $\mathcal{P}_1 \boxplus \mathcal{P}_2$ .

**Proof.** Define  $P = \{t_i^1 \mid i \in \{1, \dots, n_1\}\} \cup \{t_i^2 \mid i \in \{1, \dots, n_2\}\} \subseteq [a, b]$  which is a finite set. Using 5.53 there exists a unique bijection  $t: \{1, \dots, \#(P)\} \rightarrow P$  such that  $\forall i \in \{1, \dots, \#(P)-1\}$  we have  $t(i) < t(i+1)$ . As  $a, b \in P$  there exists a  $i, j \in \{1, \dots, \#(P)\}$  such that  $t(i) = a$  and  $t(j) = b$ . If  $i > 1$  then  $a = t(i) > t(1) \geq a$  a contradiction so we must have  $i = 1$ . If  $j < \#(P)$  then  $b = t(j) < t(\#(P)) \leq b$  a contradiction so we must have  $j = \#(P)$ . If we then take  $n = \#(P)$  and  $\mathcal{P} = \{t(i)\}_{i \in \{1, \dots, n\}}$  then we have that  $t(1) = a, t(n) = b$  and  $\forall i \in \{1, \dots, n-1\}$  that  $t(i) = t(i+1)$  proving our theorem.  $\square$

**Lemma 11.428.** Let  $\mathcal{P}_1 = \{t_i^1\}_{i \in \{1, \dots, n_1\}}$ ,  $\mathcal{P}_2 = \{t_i^2\}_{i \in \{1, \dots, n_2\}}$  be two partitions of  $[a, b]$  with  $\mathcal{P} = \mathcal{P}_1 \boxplus \mathcal{P}_2 = \{t_i\}_{i \in \{1, \dots, n\}}$  then for every  $i \in \{1, \dots, n-1\}$  there exists unique  $i_1 \in \{1, \dots, n_1-1\}$ ,  $i_2 \in \{1, \dots, n_2-1\}$  such that  $[t_i, t_{i+1}] \subseteq [t_{i_1}^1, t_{i_1+1}^1]$  and  $[t_i, t_{i+1}] \subseteq [t_{i_2}^2, t_{i_2+1}^2]$ .

**Proof.** We prove this for  $i_1$  the proof for  $i_2$  is similar (by switching  $1 \leftrightarrow 2$ ). First as  $i \in \{1, \dots, n-1\}$  we have  $t_i \neq b = t_{n_1} = t_{n_2}$ , then as  $t_i \in \{t_j \mid j \in \{1, \dots, n\}\} = \{t_j^1 \mid j \in \{1, \dots, n_1\}\} \cup \{t_j^2 \mid j \in \{1, \dots, n_2\}\}$  we have either

1.  $(t_1 \in \{t_j^1 \mid j \in \{1, \dots, n_1-1\}\})$  then there exists a  $k \in \{1, \dots, n_1-1\}$  such that  $t_i = t_k^1$ , by the definition of  $\mathcal{P} = \mathcal{P}_1 \boxplus \mathcal{P}_2$  there exists a  $l \in \{1, \dots, n\}$  such that  $t_l = t_{k+1}^1$ . Suppose now that  $t_{k+1}^1 < t_{i+1} \Rightarrow t_l < t_{i+1}$  then  $t_i = t_k^1 < t_{k+1}^1 = t_l < t_{i+1} \Rightarrow t_l \in [t_i, t_{i+1}] \stackrel{11.426}{\Rightarrow} l = i$  or  $l = i+1$ , as  $t_i < t_l$  we must have  $i \neq l$  so that  $l = i+1$  but this contradict  $t_l < t_{i+1}$ . The conclusion is thus that  $t_{i+1} \leq t_{k+1}^1 \Rightarrow t_k^1 = t_i < t_{i+1} \leq t_{k+1}^1 \Rightarrow [t_i, t_{i+1}] \subseteq [t_k^1, t_{k+1}^1]$  so we take  $i_1 = k$  to get  $[t_i, t_{i+1}] \subseteq [t_{i_1}^1, t_{i_1+1}^1]$ .
2.  $(t_1 \in \{t_j^2 \mid j \in \{1, \dots, n_2-1\}\})$  then there exists a  $k \in \{1, \dots, n_2-1\}$  such that  $t_k^2 = t_i$ . Define then  $B = \{j \in \{1, \dots, n_1\} \mid t_j^1 \leq t_i\}$  then as  $t_1^1 = a \leq a = t_1$  we have  $1 \in B \Rightarrow B \neq \emptyset$  and as  $B$  is finite (subset of a finite set) there exist a  $m = \max(B) \in \{1, \dots, n_1\}$ . If  $m = n_1$  then  $b = t_m^1 = t_{n_1}^1 = b \leq t_i < b$  a contradiction so  $m \in \{1, \dots, n_1-1\}$ . Then we have by the definition of a maximum and  $B$  that  $t_m^1 \leq t_i < t_{m+1}^1$ . Assume now that  $t_{m+1}^1 < t_{i+1}$  then as by the definition of  $\mathcal{P}$  there exists a  $l \in \{1, \dots, n\}$  such that  $t_l = t_{m+1}^1$  we have from  $t_i < t_{m+1}^1 = t_l < t_{i+1} \Rightarrow t_l \in [t_i, t_{i+1}] \stackrel{11.426}{\Rightarrow} l = i$  or  $l = i+1$  which is impossible as  $t_i < t_l < t_{i+1}$ , we are thus forced to conclude that  $t_{i+1} \leq t_{m+1}^1$ . This gives  $t_m^1 \leq t_i < t_{i+1} \leq t_{m+1}^1 \Rightarrow [t_i, t_{i+1}] \subseteq [t_m^1, t_{m+1}^1]$  giving if we take  $i_1 = m$  that  $[t_i, t_{i+1}] \subseteq [t_m^1, t_{m+1}^1]$

Next we prove uniqueness so assume that there exists a  $i_1, i'_1$  such that  $[t_i, t_{i+1}] \subseteq [t_{i_1}^1, t_{i_1+1}^1], [t_{i_1}^1, t_{i_1+1}^1]$  and assume that  $i_1 \neq i'_1$  we have then two cases

1. ( $i_1 < i'_1$ ) then  $t_{i_1}^1 < t_{i'_1}^1$  and so  $t_{i_1}^1 < t_{i'_1}^1 \leq t_i < t_{i+1} \leq t_{i_1+1}^1 \Rightarrow t_{i'_1}^1 \in [t_{i_1}^1, t_{i_1+1}^1] \xrightarrow{11.426} i'_1 = i_1$  or  $i'_1 = i_1 + 1$  which as  $t_{i_1}^1 < t_{i'_1}^1$  means that  $i'_1 = i_1 + 1$  but then  $t_{i_1+1}^1 = t_{i'_1}^1 \leq t_i < t_{i_1+1}^1$  a contradiction.
2. ( $i'_1 < i_1$ ) then  $t_{i'_1}^1 < t_{i_1}^1$  and so  $t_{i'_1}^1 < t_{i_1}^1 \leq t_i < t_{i+1} \leq t_{i_1+1}^1 \Rightarrow t_{i_1}^1 \in [t_{i'_1}^1, t_{i_1+1}^1] \xrightarrow{11.426} i_1 = i'_1$  or  $i_1 = i'_1 + 1$  which as  $t_{i'_1}^1 < t_{i_1}^1$  means that  $i_1 = i'_1 + 1$  but then  $t_{i_1+1}^1 = t_{i'_1}^1 \leq t_i < t_{i_1+1}^1$  a contradiction.

so we must conclude that  $i_1 = i'_1$   $\square$

**Lemma 11.429.** Let  $\mathcal{P}_1 = \{t_i^1\}_{i \in \{1, \dots, n_1\}}$ ,  $\mathcal{P}_2 = \{t_i^2\}_{i \in \{1, \dots, n_2\}}$  be partitions of  $[a, b]$  and  $\mathcal{P} = \mathcal{P}_1 \boxplus \mathcal{P}_2 = \{t_i\}_{i \in \{1, \dots, n\}}$  then using the previous lemma for every  $i \in \{0, \dots, n-1\}$  there exists unique  $i_1 \in \{1, \dots, n_1-1\}$ ,  $i_2 \in \{1, \dots, n_2-1\}$  so that  $[t_i, t_{i+1}] \subseteq [t_{i_1}^1, t_{i_1+1}^1], [t_{i_2}^2, t_{i_2+1}^2]$ . This defines functions  $i_1: \{1, \dots, n-1\} \rightarrow \{1, \dots, n_1-1\}$ ,  $i_2: \{1, \dots, n-1\} \rightarrow \{1, \dots, n_2-1\}$  such that  $[t_i, t_{i+1}] \subseteq [t_{i_1(i)}^1, t_{i_1(i)+1}^1], [t_{i_2(i)}^2, t_{i_2(i)+1}^2]$ . We have then additional that

1.  $i_1, i_2$  are surjective
2.  $\forall i \in \{1, \dots, n_1-1\}$  we have that
  - a.  $i_1^{-1}(\{i\}) = \{m_i^1, M_i^1\}$  where  $t_{m_i^1} = t_i^1$  and  $t_{M_i^1+1} = t_{i+1}^1$
  - b.  $i_2^{-1}(\{i\}) = \{m_i^2, M_i^2\}$  where  $t_{m_i^2} = t_i^2$  and  $t_{M_i^2+1} = t_{i+1}^2$
3. If  $i \neq j$  then  $i_1^{-1}(\{i\}) \cap i_1^{-1}(\{j\}) = \emptyset$  and  $i_2^{-1}(\{i\}) \cap i_2^{-1}(\{j\}) = \emptyset$

**Proof.** We prove this for  $i_1$  (the proof for  $i_2$  is similar by interchanging  $1 \leftrightarrow 2$ )

1. Let  $k \in \{1, \dots, n_1-1\}$  then for  $t_k^1$  there exists a  $i \in \{1, \dots, n\}$  such that  $t_k^1 = t_i^1$  and as  $k < n_1 \Rightarrow t_i = t_k^1 < t_{n_1}^1 = b = t_n$  we must have  $i \in \{1, \dots, n-1\}$ . Now assume that  $t_{k+1}^1 < t_{i+1}$  then as there exists a  $j \in \{1, \dots, n\}$  such that  $t_j = t_{k+1}^1$  we have then  $t_i = t_k^1 < t_{k+1}^1 = t_j < t_{i+1} \Rightarrow t_j \in [t_i, t_{i+1}] \xrightarrow{11.426 \text{ and } t_i < t_j} j = i+1$  but then  $t_{i+1} = t_j = t_{k+1}^1 < t_{i+1}$  a contradiction. So we must have  $t_{i+1} \leq t_{k+1}^1$  or  $t_k^1 = t_i < t_{i+1} \leq t_{k+1}^1 \Rightarrow [t_i, t_{i+1}] \subseteq [t_k^1, t_{k+1}^1]$  our  $i_1(i) = k$  proving surjectivity.
2. Let  $i \in \{1, \dots, n_1-1\}$ , as  $i_1$  is surjective we have that  $\emptyset \neq i_1^{-1}(\{i\}) \subseteq \{1, \dots, n-1\}$  so there exists  $m_i^1 = \min(i_1^{-1}(\{i\}))$ ,  $M_i^1 = \max(i_1^{-1}(\{i\}))$ . We prove now that  $i_1^{-1}(\{i\}) = \{m_i^1, M_i^1\}$ 
  - a.  $i_1^{-1}(\{i\}) \subseteq \{m_i^1, M_i^1\}$  (this is trivial by the definition of minimum and maximum)
  - b. If  $k \in \{m_i^1, M_i^1\}$  then  $m_i^1 \leq k \leq M_i^1 \Rightarrow t_{m_i^1} \leq t_k \leq t_{M_i^1+1} \Rightarrow t_{m_i^1+1} \leq t_{M_i^1+1}$ . As we have  $[t_{m_i^1}, t_{m_i^1+1}], [t_{M_i^1}, t_{M_i^1+1}] \subseteq [t_i^1, t_{i+1}^1]$  we have  $t_i^1 \leq t_{m_i^1} \leq t_k < t_{k+1} \leq t_{M_i^1+1} \leq t_{i+1}^1 \Rightarrow [t_k, t_{k+1}] \subseteq [t_i^1, t_{i+1}^1] \Rightarrow i_1(k) = i \Rightarrow k \in i_1^{-1}(\{i\})$ . Proving  $\{m_i^1, M_i^1\} \subseteq i_1^{-1}(\{i\})$

As  $i_1^{-1}(\{i\}) = \{m_i^1, \dots, M_i^1\}$  we have  $m_i^1, M_i^1 \in i_1^{-1}(\{i\}) \Rightarrow i_1(m_i^1) = i, i_1(M_i^1) = i$  or

$$[t_{m_i^1}, t_{m_i^1+1}], [t_{M_i^1}, t_{M_i^1+1}] \subseteq [t_i^1, t_{i+1}^1] \quad (11.103)$$

From 11.103 it follows that  $t_i^1 \leq t_{m_i^1}$ , assume now that  $t_i^1 < t_{m_i^1}$  then by the definition of  $\mathcal{P}_1 \boxplus \mathcal{P}_2$  there exists a  $l \in \{1, \dots, n\}$  such that  $t_l = t_i^1$  then  $t_l = t_i^1 < t_{m_i^1}$  so that  $l < m_i^1$  or  $l+1 < m_i^1 + 1 \Rightarrow t_i^1 = t_l < t_{l+1} < t_{m_i^1+1} \leq t_{i+1}^1 \Rightarrow [t_l, t_{l+1}] \subseteq [t_i^1, t_{i+1}^1]$  proving that  $i_1(l) = i \Rightarrow l \in i_1^{-1}(\{i\}) = \{m_i^1, \dots, M_i^1\}$  proving that  $m_i^1 \leq l < m_i^1$  a contradiction. So we must have  $t_i^1 = t_{m_i^1}$ . From 11.103 it follows that  $t_{M_i^1+1} \leq t_{i+1}^1$ , assume now that  $t_{M_i^1+1} < t_{i+1}^1$  then by the definition of  $\mathcal{P}_1 \boxplus \mathcal{P}_2$  there exists a  $l \in \{1, \dots, n\}$  so that  $t_{i+1}^1 = t_l$  then  $t_{M_i^1+1} < t_{i+1}^1 = t_l$  so that  $M_i^1 + 1 < l \Rightarrow M_i^1 < l-1$  thus  $t_i^1 \leq t_{M_i^1} < t_{l-1} < t_l = t_{i+1}^1 \Rightarrow [t_{l-1}, t_l] \subseteq [t_i^1, t_{i+1}^1]$  proving that  $i_1(l-1) = i \Rightarrow l-1 \in i_1^{-1}(\{i\}) = \{m_i^1, \dots, M_i^1\} \Rightarrow l-1 \leq M_i^1 < l-1$  a contradiction. So we must have that  $t_{M_i^1+1} = t_{i+1}^1$

3. Assume that  $k \in i_1^{-1}(i) \cap i_1^{-1}(j), i \neq j$  then  $i_1(k) = i$  and  $i_1(k) = j$  so that by definition of a function we must have  $i = j$ .  $\square$

**Definition 11.430.** Let  $[a, b] \subseteq \mathbb{R}$  and  $\mathcal{P} = \{t_i\}_{i \in \{1, \dots, n\}}$  be a partition of  $[a, b]$  then the norm  $\mu(\mathcal{P})$  of the partition is defined by  $\mu(\mathcal{P}) = \max(\{|t_{i+1} - t_i| \mid i \in \{1, \dots, n-1\}\})$   $\underset{t_i < t_{i+1}}{=} \max(\{|t_{i+1} - t_i| \mid i \in \{1, \dots, n-1\}\})$

**Definition 11.431.** Let  $[a, b] \subseteq \mathbb{R}$  and  $\mathcal{P} = \{t_i\}_{i \in \{1, \dots, n\}}$  be a partition of  $[a, b]$  then a **tag** on a partition is a family  $\{s_i\}_{i \in \{1, \dots, n-1\}}$  such that  $\forall i \in \{1, \dots, n-1\}$  we have  $s_i \in [t_i, t_{i+1}]$ . A tagged partition  $\mathbb{P}$  is a pair of a partition and a tag on this partition, so  $\mathbb{P} = (\mathcal{P}, s) = (\{t_i\}_{i \in \{1, \dots, n\}}, \{s_i\}_{i \in \{1, \dots, n-1\}})$  such that  $\forall i \in \{1, \dots, n\}$  we have  $t_i < t_{i+1} \wedge s_i \in [t_i, t_{i+1}]$  and  $t_i = a, t_n = b$ . The norm of a tagged partition is the norm of its partition so  $\mu(\mathbb{P}) = \mu((\mathcal{P}, \{s_i\}_{i \in \{1, \dots, n-1\}})) = \mu(\mathcal{P})$

**Definition 11.432.** Let  $[a, b] \subseteq \mathbb{R}$ ,  $\mathbb{P} = \{\{t_i\}_{i \in \{1, \dots, n\}}, \{s_i\}_{i \in \{1, \dots, n-1\}}\}$  is a tagged partition on  $[a, b]$ ,  $\langle X, \|\cdot\| \rangle$  a normed real vector space and  $f: [a, b] \rightarrow X$  a function then a Riemann sum of  $f$  using the tagged partition, noted by  $\mathcal{S}(f, \mathbb{P})$  is defined by  $\mathcal{S}(f, \mathbb{P}) = \sum_{i=1}^{n-1} f(s_i) \cdot (t_{i+1} - t_i)$ .

**Theorem 11.433.** Let  $[a, b] \subseteq \mathbb{R}$ ,  $\mathbb{P} = \{\{t_i\}_{i \in \{1, \dots, n\}}, \{s_i\}_{i \in \{1, \dots, n-1\}}\}$  a tagged partition on  $[a, b]$ ,  $f, g: [a, b] \rightarrow \mathbb{R}$  a function such that  $\forall x \in [a, b]$  we have  $f(x) \leq g(x)$  then  $\mathcal{S}(f, \mathbb{P}) \leq \mathcal{S}(g, \mathbb{P})$

**Proof.** As  $\forall i \in \{1, \dots, n\}$  we have  $0 \leq t_{i+1} - t_i, f(s_i) \leq g(s_i)$  it follows that  $f(s_i) \cdot (t_{i+1} - t_i) \leq g(s_i) \cdot (t_{i+1} - t_i)$  so that  $\mathcal{S}(f, \mathbb{P}) = \sum_{i=1}^{n-1} f(s_i) \cdot (t_{i+1} - t_i) = \sum_{i=1}^{n-1} g(s_i) \cdot (t_{i+1} - t_i) = \mathcal{S}(g, \mathbb{P})$   $\square$

**Lemma 11.434.** Let  $[a, b] \subseteq \mathbb{R}$ ,  $\langle X, \|\cdot\| \rangle$  a normed real vector space and  $f: [a, b] \rightarrow X$  a function,  $\mathbb{P}_1 = (\{t_i^1\}_{i \in \{1, \dots, n_1\}}, \{s_i^1\}_{i \in \{1, \dots, n_1-1\}})$ ,  $\mathbb{P}_2 = (\{t_i^2\}_{i \in \{1, \dots, n_2\}}, \{s_i^2\}_{i \in \{1, \dots, n_2-1\}})$  tagged partitions of  $[a, b]$  then if  $\mathcal{P} = \{t_i\}_{i \in \{1, \dots, n\}} = \{t_i^1\}_{i \in \{1, \dots, n_1\}} \boxplus \{t_i^2\}_{i \in \{1, \dots, n_2\}}$  we have

1.  $\mathcal{S}(f, \mathbb{P}_1) = \sum_{i=1}^{n-1} f(s_{i_1(i)}^1) \cdot (t_{i+1} - t_i)$
2.  $\mathcal{S}(f, \mathbb{P}_2) = \sum_{i=1}^{n-1} f(s_{i_2(i)}^2) \cdot (t_{i+1} - t_i)$
3.  $\mathcal{S}(f, \mathbb{P}_1) - \mathcal{S}(f, \mathbb{P}_2) = \sum_{i \in \{1, \dots, n\}} (f(s_{i_1(i)}^1) - f(s_{i_2(i)}^2))$

( $i_1, i_2$  as defined in 11.429)

**Proof.** First we prove that for  $j = 1, 2$  we have  $\mathcal{S}(f, \mathbb{P}_j) = \sum_{i=1}^{n-1} f(s_{i_j(i)}^j) \cdot (t_{i+1} - t_i)$ . First as by 11.429 we have that  $i_j$  is surjective so that  $\{1, \dots, n-1\} = \bigcup_{k \in \{1, \dots, n_j-1\}} i_j^{-1}(\{k\})$  and  $\forall k, l \in \{1, \dots, n_j-1\}$  with  $k \neq l$  we have  $i_j^{-1}(\{k\}) \cap i_j^{-1}(\{l\}) = \emptyset$  and  $i_j^{-1}(\{k\}) = \{m_k^j, \dots, M_k^j\}$  and  $t_{m_k^j}^j = t_k$  and  $t_{M_k^j+1}^j = t_{k+1}$ , so if  $i \in \{m_k^j, \dots, M_k^j\}$  then  $k = i_j(i)$ . Using 10.46 we have  $\sum_{i=1}^n f(s_{i_j(i)}^j) \cdot (t_{i+1} - t_i) = \sum_{i \in \{1, \dots, n\}} f(s_{i_j(i)}^j) \cdot (t_{i+1} - t_i) = \sum_{k \in \{1, \dots, n_j-1\}} \left( \sum_{i \in \{m_k^j, \dots, M_k^j\}} f(s_{i_j(i)}^j) \cdot (t_{i+1} - t_i) \right) = \sum_{k \in \{1, \dots, n_j-1\}} \left( \sum_{i \in \{m_k^j, \dots, M_k^j\}} f(s_k^j) \cdot (t_{i+1} - t_i) \right) \underset{10.125}{=} \sum_{k \in \{1, \dots, n_j-1\}} f(s_k^j) \cdot \left( \sum_{i=m_k^j}^{M_k^j} (t_{i+1} - t_i) \right) \underset{10.27}{=} \sum_{k \in \{1, \dots, n_j-1\}} f(s_k^j) \cdot (t_{M_k^j+1}^j - t_{m_k^j}^j) = \sum_{k=1}^{n_j-1} f(s_k^j) \cdot (t_{k+1}^j - t_k^j) = \mathcal{S}(f, \mathbb{P}_j)$ .

Finally  $\mathcal{S}(f, \mathbb{P}_1) - \mathcal{S}(f, \mathbb{P}_2) = \left( \sum_{i=1}^n f(s_{i_1(i)}^1) \cdot (t_{i+1} - t_i) \right) - \left( \sum_{i=1}^n f(s_{i_2(i)}^2) \cdot (t_{i+1} - t_i) \right) = \sum_{i=1}^n (f(s_{i_1(i)}^1) - f(s_{i_2(i)}^2)) \cdot (t_{i+1} - t_i)$   $\square$

**Lemma 11.435.** Let  $[a, b] \subseteq \mathbb{R}$ ,  $\langle X, \|\cdot\| \rangle$  a normed real vector space and  $f: [a, b] \rightarrow X$  a continuous function then for every  $\varepsilon > 0$  there exists a  $\delta < 0$  such that  $\|\mathcal{S}(f, \mathbb{P}_1) - \mathcal{S}(f, \mathbb{P}_2)\| < \varepsilon$  for all  $\mathbb{P}_1, \mathbb{P}_2$  such that  $\mu(\mathbb{P}_1) < \delta, \mu(\mathbb{P}_2) < \delta$

**Proof.** Using the fact that  $[a, b]$  is compact in  $\langle \mathbb{R}, \|\cdot\| \rangle$  (see 11.267),  $f$  is continuous and 11.268 we have that  $f$  is uniform continuous on  $[a, b]$ . So there exists a  $\delta > 0$  such that  $\forall x, x' \in [a, b]$  with  $|x - x'| < 2 \cdot \delta$  we have  $\|f(x) - f(x')\| < \frac{\varepsilon}{b-a}$ . Take now  $\mathbb{P}_1, \mathbb{P}_2$  such that  $\mu(\mathbb{P}_1) < \delta, \mu(\mathbb{P}_2) < \delta$  then by using the previous lemma we have

$$\|\mathcal{S}(f, \mathbb{P}_1) - \mathcal{S}(f, \mathbb{P}_2)\| = \left\| \sum_{i=1}^{n-1} (f(s_{i_1(i)}^1) - f(s_{i_2(i)}^2)) \cdot (t_{i+1} - t_i) \right\| \quad (11.104)$$

As  $t_i \in [t_i, t_{i+1}] \subseteq [t_{i_1(i)}, t_{i_1(i)+1}] \cap [t_{i_2(i)}, t_{i_2(i)+1}]$  we have as also  $s_{i_1(i)}^1 \in [t_{i_1(i)}, t_{i_1(i)+1}]$ ,  $s_{i_2(i)}^2 \in [t_{i_2(i)}, t_{i_2(i)+1}]$  that  $|s_{i_1(i)}^1 - s_{i_2(i)}^2| \leq |s_{i_1(i)}^1 - t_i| + |t_i - s_{i_2(i)}^2| \leq \mu(\mathbb{P}_1) + \mu(\mathbb{P}_2) < \delta + \delta = 2 \cdot \delta$  so that

$$\|f(s_{i_1(i)}^1) - f(s_{i_2(i)}^2)\| < \frac{\varepsilon}{b-a} \quad (11.105)$$

Using 11.104 and 11.105 we have  $\|\mathcal{S}(f, \mathbb{P}_1) - \mathcal{S}(f, \mathbb{P}_2)\| \leq \sum_{i=1}^{n-1} \|f(s_{i_1(i)}^1) - f(s_{i_2(i)}^2)\| |t_{i+1} - t_i| \leq \frac{\varepsilon}{b-a} \cdot \sum_{i=1}^{n-1} |t_{i+1} - t_i| \stackrel{10.27}{=} \frac{\varepsilon}{b-a} \cdot \sum_{i=1}^{n-1} (t_{i+1} - t_i) \stackrel{10.27}{=} \frac{\varepsilon}{b-a} \cdot (t_n - t_1) = \frac{\varepsilon}{b-a} \cdot (b-a) = \varepsilon$  what we set out to prove.  $\square$

**Lemma 11.436.** *Let  $[a, b] \subseteq \mathbb{R}$ ,  $a < b$ ,  $\langle X, \|\cdot\| \rangle$  a real normed space and  $f: [a, b] \rightarrow X$  a function then if a  $I \in X$  satisfies the following condition*

1.  $\forall \varepsilon > 0$  there exists a  $\delta > 0$  such that for every Riemann  $\mathcal{S}(f, \mathbb{P})$  with  $\mu(\mathbb{P}) < \delta$  (where  $\mathbb{P}$  is a tagged partition of  $[a, b]$ ) we have  $\|I - \mathcal{S}(f, \mathbb{P})\| < \varepsilon$

then  $I$  is unique.

**Proof.** Suppose that  $I, I'$  satisfies (1) and  $I \neq I'$  then  $\varepsilon = \|I - I'\| > 0$  then by (1) there exists a  $\delta > 0$  such that if  $\mu(\mathbb{P}) < \delta$  then  $\|I - \mathcal{S}(f, \mathbb{P})\| < \frac{\varepsilon}{2}$ ,  $\|I' - \mathcal{S}(f, \mathbb{P})\| < \frac{\varepsilon}{2}$ . This means  $\varepsilon = \|I - I'\| \leq \|I - \mathcal{S}(f, \mathbb{P})\| + \|\mathcal{S}(f, \mathbb{P}) - I'\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \Rightarrow \varepsilon < \varepsilon$  a contradiction. The only conclusion left is that  $I = I'$ .  $\square$

**Definition 11.437.** *Let  $[a, b] \subseteq \mathbb{R}$ ,  $\langle X, \|\cdot\| \rangle$  and  $f: [a, b] \rightarrow X$  a function so that there exists a  $I(f) \in X$  with  $\forall \varepsilon > 0$  there exists a  $\delta > 0$  such that for every Riemann sum  $\mathcal{S}(f, \mathbb{P})$  with  $\mu(\mathbb{P}) < \delta$  (where  $\mathbb{P}$  is a tagged partition of  $[a, b]$ ) we have  $\|I(f) - \mathcal{S}(f, \mathbb{P})\| < \varepsilon$ .  $f$  is then called **Riemann Integrable** and  $I(f)$  is called the **Riemann Integral of  $f$**  and noted by  $\int_a^b f$ . The set of integrable functions on  $[a, b]$  is noted as  $\mathcal{L}([a, b], X)$ . The integral  $\int$  is then a function  $\int_a^b: \mathcal{L}([a, b], X) \rightarrow X$  defined by  $f \rightarrow \int_a^b f$*

**Note 11.438.** Sometimes we note a function  $f: [a, b] \rightarrow Y$  defined by  $x \rightarrow f(x)$  where  $f(x)$  is in general a expression depending on  $x$ , in this case we can note  $\int_a^b f$  by  $\int_a^b f(x) dx$ , this notation is especially useful if we have for example a function  $f: [a, b] \times X \rightarrow Y$  defined by  $(t, x) \rightarrow f(t, x)$  then if we define for a  $x \in X$   $g_x: [a, b] \rightarrow Y$  by  $t \rightarrow g_x(t) = f(t, x)$  then we can write  $\int_a^b g_x$  as  $\int_a^b f(t, x) dt$  which avoids the extra definition of  $g_x$ . Another benefit is that if we have  $f: [a, b] \rightarrow \mathbb{K}$  with  $x \rightarrow f(x) = \frac{\cos(x) + 1}{x+1}$  then again we can write for  $\int_a^b f$  the expression  $\int_a^b \frac{\cos(x) + 1}{x+1} dx$ . On the other hand we have introduced a extra indeterminism as  $\int_a^b f(x) dx$  is the same as  $\int_a^b f(y) dy \dots$ , this is the reason that we avoid this notation for the rest of this chapter. However if needed we can fall back on this notation if we feel a need to it.

**Theorem 11.439.** *Let  $[a, b] \subseteq \mathbb{R}$ ,  $\langle X, \|\cdot\| \rangle$  a normed real space then we have that there exists a family of tagged partitions on  $[a, b]$   $\{\mathbb{P}_i\}_{i \in \mathbb{N}_0}$  such that  $\lim_{i \rightarrow \infty} \mu(\mathbb{P}_i) = 0$*

**Proof.** Let  $k \in \mathbb{N}_0$  define then  $\{t_i^{(k)}\}_{i \in \{1, \dots, k+1\}}$  by  $t_i^{(k)} = (i-1) \cdot \frac{b-a}{k} + a$  then we have  $t_1^{(k)} = (1-1) \cdot \frac{b-a}{k} + a = a$  and  $t_{k+1}^{(k)} = (k+1-1) \cdot \frac{b-a}{k} + a = \frac{k \cdot (b-a) + a \cdot k}{k} = \frac{k \cdot b - a \cdot k + a \cdot k}{k} = b$ , further  $t_{i+1}^{(k)} - t_i^{(k)} = ((i+1)-1) \cdot \frac{b-a}{k} + a - (i-1) \cdot \frac{b-a}{k} - a = i \cdot \frac{b-a}{k} - i \cdot \frac{b-a}{k} + \frac{b-a}{k} = \frac{b-a}{k} > 0$ , so that using 11.330 we have  $\lim_{k \rightarrow \infty} (t_{i+1}^{(k)} - t_i^{(k)}) = 0$ . Define now  $\mathbb{P}_k = \{\{t_i^{(k)}\}_{i \in \{1, \dots, k+1\}}, \{t_i^{(k)}\}_{i \in \{1, \dots, k\}}\}$  then we have  $\mu(\mathbb{P}_k) = \frac{b-a}{k}$  and thus for  $\{\mathbb{P}_k\}_{k \in \mathbb{N}_0}$  we have  $\lim_{i \rightarrow \infty} \mu(\mathbb{P}_k) = 0$   $\square$

**Theorem 11.440.** *Let  $[a, b] \subseteq \mathbb{R}$ ,  $X$ ,  $f: [a, b] \rightarrow X$  a **continuous** function then the following are equivalent*

1.  $f$  is Riemann integrable with integral  $\int_a^b f$
2. There exists a  $I \in X$  such that for every family of tagged partitions  $\{\mathbb{P}_i\}_{i \in \mathbb{N}_0}$  with  $\lim_{i \rightarrow \infty} \mu(\mathbb{P}_i) = 0$  we have  $\lim_{i \rightarrow \infty} \mathcal{S}(f, \mathbb{P}_i) = I$ . We have then that  $I = \int_a^b f$

**Proof.**

1.  $(\Rightarrow)$  Let  $f$  be Riemann integrable with integral  $\int_a^b f$  and take  $\{\mathbb{P}_i\}_{i \in \mathbb{N}_0}$  such that  $\lim_{i \rightarrow \infty} \mu(\mathbb{P}_i) = 0$ . Then if  $\varepsilon > 0$  there exists a  $\delta > 0$  such that if  $\mathbb{P}$  is such that  $\mu(\mathbb{P}) < \delta$  then  $\|\mathcal{S}(f, \mathbb{P}) - \int_a^b f\| < \varepsilon$ . Now as  $\lim_{i \rightarrow \infty} \mu(\mathbb{P}_i) = 0$  then there exists a  $N \in \mathbb{N}_0$  such that if  $n \geq N$  then  $\mu(\mathbb{P}_n) = |\mu(\mathbb{P}_n) - 0| < \delta \Rightarrow \|\mathcal{S}(f, \mathbb{P}_n) - \int_a^b f\| < \varepsilon$  proving that  $\lim_{i \rightarrow \infty} \mathcal{S}(f, \mathbb{P}_i) = \int_a^b f$
2.  $(\Leftarrow)$  First by 11.439 there exists a  $\{\mathbb{P}_i\}_{i \in \mathbb{N}_0}$  with  $\lim_{i \rightarrow \infty} \mu(\mathbb{P}_i) = 0$ , by the hypothesis we have then that  $I = \lim_{i \rightarrow \infty} \mathcal{S}(f, \mathbb{P}_i)$  exists. Take now  $\varepsilon > 0$  then there exists a  $N_1 \in \mathbb{N}_0$  such that if  $n \geq N_1$  then  $\|\mathcal{S}(f, \mathbb{P}_n) - I\| < \frac{\varepsilon}{2}$ , using 11.435 there exists a  $\delta > 0$  such that if  $\mathbb{P}', \mathbb{P}''$  are two tagged partitions with  $\mu(\mathbb{P}') < \delta, \mu(\mathbb{P}'') < \delta$  then  $\|\mathcal{S}(f, \mathbb{P}') - \mathcal{S}(f, \mathbb{P}'')\| < \frac{\varepsilon}{2}$ , as  $\lim_{i \rightarrow \infty} \mu(\mathbb{P}_i) = 0$  there exists also a  $N_2$  so that if  $n \geq N_2$  then  $\mu(\mathbb{P}_n) = |\mu(\mathbb{P}_n) - 0| < \delta$ . Take now  $N = \max(N_1, N_2) \geq N_1, N_2$  then  $\mu(\mathbb{P}_N) < \delta$  and  $\|\mathcal{S}(f, \mathbb{P}_N) - I\| < \frac{\varepsilon}{2}$  and assume that for a partition  $\mathbb{P}$  we have  $\mu(\mathbb{P}) < \delta$  then  $\|\mathcal{S}(f, \mathbb{P}) - I\| = \|\mathcal{S}(f, \mathbb{P}) - \mathcal{S}(f, \mathbb{P}_N)\| + \|\mathcal{S}(f, \mathbb{P}_N) - I\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$ . Proving that  $f$  is integrable and as we have uniqueness by 11.436 we must have  $I = \int_a^b f$ .  $\square$

**Theorem 11.441.** Let  $f, g \in \mathcal{L}([a, b], \mathbb{R})$  such that  $\forall x \in [a, b]$  we have  $f(x) \leq g(x)$  then  $\int_a^b f \leq \int_a^b g$  then  $\int_a^b f \leq \int_a^b g$

**Proof.** Using 11.439 there exists a  $\{\mathbb{P}_i\}_{i \in \mathbb{N}_0}$  such that  $\lim_{i \rightarrow \infty} \mu(\mathbb{P}_i) = 0$  then we have

$$\begin{aligned} \int_a^b f &\stackrel{11.440}{=} \lim_{i \rightarrow \infty} \mathcal{S}(f, \mathbb{P}_i) \\ &\stackrel{\text{11.355 and 11.433}}{\leq} \lim_{i \rightarrow \infty} \mathcal{S}(g, \mathbb{P}_i) \\ &\stackrel{11.440}{=} \int_a^b g \end{aligned}$$

$\square$

We prove now that  $\mathcal{L}([a, b], X)$  is not empty if  $X$  is a Banach space and  $f$  is continuous.

**Theorem 11.442.** Let  $[a, b] \subseteq \mathbb{R}$ ,  $\langle X, \|\cdot\| \rangle$  a real normed Banach space and  $f: [a, b] \rightarrow X$  a continuous function then  $f$  is integrable. In other words we have that  $\mathcal{C}([a, b], X) \subseteq \mathcal{L}([a, b], X)$

**Proof.** Let  $\{\mathbb{P}_i\}_{i \in \mathbb{N}_0}$  be a family of tagged partitions on  $[a, b]$  with  $\lim_{i \rightarrow \infty} \mu(\mathbb{P}_i) = 0$ . Given a  $\varepsilon > 0$  we can find by 11.435 there exists a  $\delta > 0$  such that if  $\mu(\mathbb{P}) < \delta, \mu(\mathbb{P}') < \delta$  then  $\|\mathcal{S}(f, \mathbb{P}) - \mathcal{S}(f, \mathbb{P}')\| < \varepsilon$ , we can then find a  $N \in \mathbb{N}_0$  such that if  $i \geq N$  we have  $\mu(\mathbb{P}_i) = |\mu(\mathbb{P}_i) - 0| < \delta$ . So if  $n, m \geq N$  then  $\mu(\mathbb{P}_n), \mu(\mathbb{P}_m) < \delta$  and thus  $\|\mathcal{S}(f, \mathbb{P}_n) - \mathcal{S}(f, \mathbb{P}_m)\| < \varepsilon$  proving that  $\{\mathcal{S}(f, \mathbb{P}_i)\}_{i \in \mathbb{N}_0}$  is a Cauchy sequence in  $X$ . As  $X$  is a Banach space we have that  $\{\mathcal{S}(f, \mathbb{P}_i)\}_{i \in \mathbb{N}_0}$  converges to a limit  $I$ . Using then the previous theorem (see 11.440) we have then that  $f$  is Integrable with  $\lim_{i \rightarrow \infty} \mathcal{S}(f, \mathbb{P}_i) = \int_a^b f$   $\square$

Up to now we have always assumed that  $a < b$  let's now remove this restriction

**Definition 11.443.** Let  $[a, b] \subseteq \mathbb{R}$  with  $a \leq b$ ,  $\langle X, \|\cdot\| \rangle$  a real normed space then  $f: [a, b] \rightarrow X$  is Riemann Integrable with integral  $\int_a^b f$  if and only if

1. If  $a < b$  then we use the definition of 11.437
2. If  $a = b$  then  $f$  is always integrable and  $\int_a^b f = 0$

From now on if we say  $[a, b] \in \mathbb{R}$  then  $a \leq b$  (so  $a = b$  is back allowed)

**Theorem 11.444.** Let  $[a, b] \subseteq \mathbb{R}, a \leq b$  and  $\langle X, \|\cdot\| \rangle$  a real normed space then  $\int_a^b: \mathcal{L}([a, b], X) \rightarrow X$  is linear. In other words  $\forall \alpha, \beta \in \mathbb{R}$  and  $\forall f, g \in \mathcal{L}([a, b], X)$  we have  $\alpha \cdot f + \beta \cdot g \in \mathcal{L}([a, b], X)$  and  $\int_a^b (\alpha \cdot f + \beta \cdot g) = \alpha \cdot \int_a^b f + \beta \cdot \int_a^b g$

**Proof.** Consider the two cases for the integral

1. **( $a = b$ )** Then the integral exists always and  $\int_a^b (\alpha \cdot f + \beta \cdot g) = 0 + 0 = \alpha \cdot 0 + \beta \cdot 0 = \alpha \cdot \int_a^b f + \beta \cdot \int_a^b g$
2. **( $a < b$ )** Let  $\mathbb{P} = (\{t_i\}_{i \in \{1, \dots, n\}}, \{s_i\}_{i \in \{1, \dots, n-1\}})$  a tagged partition of  $[a, b]$  then  $\mathcal{S}(\alpha \cdot f + \beta \cdot g, \mathbb{P}) = \sum_{i=1}^n (\alpha \cdot f + \beta \cdot g)(s_i) \cdot (t_{i+1} - t_i) = \sum_{i=1}^n (\alpha \cdot f(s_i) + \beta \cdot g(s_i)) \cdot (t_{i+1} - t_i) = \alpha \cdot \sum_{i=1}^n f \cdot (t_{i+1} - t_i) + \beta \cdot \sum_{i=1}^n g \cdot (t_{i+1} - t_i) = \alpha \cdot \mathcal{S}(f, \mathbb{P}) + \beta \cdot \mathcal{S}(g, \mathbb{P})$  giving

$$\mathcal{S}(\alpha \cdot f + \beta \cdot g, \mathbb{P}) = \alpha \cdot \mathcal{S}(f, \mathbb{P}) + \beta \cdot \mathcal{S}(g, \mathbb{P}) \quad (11.106)$$

Let  $\varepsilon > 0$  then as  $f, g$  is integrable there exists a  $\delta > 0$  such that if  $\mu(\mathbb{P}) < \delta$  then  $\|\int_a^b f - \mathcal{S}(f, \mathbb{P})\| < \frac{\varepsilon}{2 \cdot (|\alpha| + 1)}$  and  $\|\int_a^b g - \mathcal{S}(g, \mathbb{P})\| < \frac{\varepsilon}{2 \cdot (|\beta| + 1)}$  and thus  $\|(\alpha \cdot \int_a^b f + \beta \cdot \int_a^b g) - \mathcal{S}(\alpha \cdot f + \beta \cdot g, \mathbb{P})\| \stackrel{11.106}{=} \|\alpha \cdot \int_a^b f + \beta \cdot \int_a^b g - \alpha \cdot \mathcal{S}(f, \mathbb{P}) - \beta \cdot \mathcal{S}(g, \mathbb{P})\| \leq \|\alpha \cdot (\int_a^b f - \mathcal{S}(f, \mathbb{P}))\| + \|\beta \cdot (\int_a^b g - \mathcal{S}(g, \mathbb{P}))\| = |\alpha| \cdot \|\int_a^b f - \mathcal{S}(f, \mathbb{P})\| + |\beta| \cdot \|\int_a^b g - \mathcal{S}(g, \mathbb{P})\| < |\alpha| \cdot \frac{\varepsilon}{2 \cdot (|\alpha| + 1)} + |\beta| \cdot \frac{\varepsilon}{2 \cdot (|\beta| + 1)} < \varepsilon$  proving that  $\int_a^b (\alpha \cdot f + \beta \cdot g) = \alpha \cdot \int_a^b f + \beta \cdot \int_a^b g$ .  $\square$

**Theorem 11.445.** Let  $[a, b] \subseteq \mathbb{R}, a \leq b$  and  $\langle X, \|\cdot\| \rangle$  a real normed space and  $f \in \mathcal{L}([a, b], X)$  then we have

1.  $\forall \varphi \in L(X, \mathbb{R})$  (a continuous linear function) we have that  $\varphi \circ f$  is integrable and  $\int_a^b (\varphi \circ f) = \varphi(\int_a^b f)$
2. If  $\|f\|$  is continuous then  $f, \|f\|$  are integrable and  $\|\int_a^b f\| \leq \int_a^b \|f\|$
3. Given  $c \in X$  the constant function  $c$  defined by  $x \mapsto c(x) = c$  is integrable and  $\int_a^b c = c \cdot (b - a)$
4. If  $f$  is continuous and  $m \in \mathbb{R}$  is such that  $\forall x \in [a, b]$  we have  $\|f(x)\| < m$  then  $\|\int_a^b f\| \leq m \cdot (b - a)$

**Proof.** We divide the proof in the cases of  $a = b$  and  $a < b$

1. **( $a = b$ )**
  - a. If  $\varphi \in L(X, \mathbb{R})$  then we have  $\int_a^a \varphi \circ f = 0 \stackrel{10.175}{=} \varphi(0) = \varphi(\int_a^a f)$
  - b.  $\|\int_a^a f\| = \|0\| = 0 \leq 0 = \int_a^a \|f\|$
  - c.  $\int_a^a c = 0 = c \cdot 0 = c \cdot (a - a)$
  - d.  $\|\int_a^a f\| = \|0\| \leq 0 = m \cdot 0 = m \cdot (a - a)$
2. **( $a < b$ )**
  - a. If  $\mathbb{P} = (\{t_i\}_{i \in \{1, \dots, n\}}, \{s_i\}_{i \in \{1, \dots, n-1\}})$  then we have  $\mathcal{S}(\varphi \circ f, \mathbb{P}) = \sum_{i=1}^{n-1} \varphi(f(s_i)) \cdot (t_{i+1} - t_i) \stackrel{\varphi \text{ is linear}}{=} \varphi(\sum_{i=1}^{n-1} f(s_i) \cdot (t_{i+1} - t_i)) = \varphi(\mathcal{S}(f, \mathbb{P}))$ , so we have

$$\mathcal{S}(\varphi \circ f, \mathbb{P}) = \varphi(\mathcal{S}(f, \mathbb{P})) \quad (11.107)$$

Now as  $\varphi$  is continuous, we have that given a  $\varepsilon > 0$  there exists a  $\delta' > 0$  such that if  $\|x\| = \|x - 0\| < \delta'$  then  $|\varphi(x)| < \varepsilon$ . As  $f$  is integrable there exists a  $\delta > 0$  such that if  $\mu(\mathbb{P}) < \delta$  then  $\|\int_a^b f - \mathcal{S}(f, \mathbb{P})\| < \delta'$  and thus  $|\varphi(\int_a^b f) - \mathcal{S}(\varphi \circ f, \mathbb{P})| \stackrel{11.107}{=} |\varphi(\int_a^b f) - \varphi(\mathcal{S}(f, \mathbb{P}))| \stackrel{\varphi \text{ is linear}}{=} |\varphi((\int_a^b f) - \mathcal{S}(f, \mathbb{P}))| < \varepsilon$  proving that  $\varphi \circ f$  is integrable with integral  $\varphi(\int_a^b f)$ .

- b. As  $f$  is continuous we have that  $f$  is integrable, further given  $x \in X$  and  $\varepsilon > 0$  a  $\delta > 0$  so that if  $\|x - y\| < \delta$  then  $\|f(x) - f(y)\| < \varepsilon \Rightarrow \|\|f(x)\| - \|f(y)\|\| \leq \|f(x) - f(y)\| < \varepsilon$  proving that  $\|f\|$  is continuous and thus integrable. Second given  $\mathbb{P} = (\{t_i\}_{i \in \{1, \dots, n\}}, \{s_i\}_{i \in \{1, \dots, n-1\}})$  we have  $\|\mathcal{S}(f, \mathbb{P})\| = \|\sum_{i=1}^{n-1} f(s_i) \cdot (t_{i+1} - t_i)\| \leq \sum_{i=1}^{n-1} \|f(s_i) \cdot (t_{i+1} - t_i)\| = \sum_{i=1}^{n-1} \|f(s_i)\| \cdot (t_{i+1} - t_i)$  proving that

$$\|\mathcal{S}(f, \mathbb{P})\| \leq \mathcal{S}(\|f\|, \mathbb{P}) \quad (11.108)$$

This means that if  $\{\mathbb{P}_i\}_{i \in \mathbb{N}_0}$  is such that  $\lim_{i \rightarrow \infty} \mu(\mathbb{P}_i) = 0$  then  $\| \int_a^b f \| = \left\| \lim_{i \rightarrow \infty} \mathcal{S}(f, \mathbb{P}_i) \right\|$   $\| \cdot \|$  is continuous and  $\| \mathcal{S}(f, \mathbb{P}) \| \leq 11.108$  and  $11.351 \lim_{i \rightarrow \infty} \mathcal{S}(\|f\|, \mathbb{P}_i) = \int_a^b \|f\|$ .

- c. If  $c$  is the constant function then we have for  $\mathbb{P} = (\{t_i\}_{i \in \{1, \dots, n\}}, \{s_i\}_{i \in \{1, \dots, n-1\}})$  that  $\mathcal{S}(c, \mathbb{P}) = \sum_{i=1}^{n-1} c(s_i) \cdot (t_{i+1} - t_i) = \sum_{i=1}^{n-1} c \cdot (t_{i+1} - t_i) = c \cdot \sum_{i=1}^{n-1} (t_{i+1} - t_i) \stackrel{10.27}{=} c \cdot (t_n - t_1) = c \cdot (b - a)$
- d. First  $\mathbb{P} = (\{t_i\}_{i \in \{1, \dots, n\}}, \{s_i\}_{i \in \{1, \dots, n-1\}})$  is a tagged partition on  $[a, b]$  then we have  $\| \mathcal{S}(f, \mathbb{P}) \| = \left\| \sum_{i=1}^{n-1} f(s_i) \cdot (t_{i+1} - t_i) \right\| \leq \sum_{i=1}^{n-1} \|f(s_i)\| \cdot (t_{i+1} - t_i) \leq \sum_{i=1}^{n-1} m \cdot (t_{i+1} - t_i) = m \cdot \sum_{i=1}^{n-1} (t_{i+1} - t_i) \stackrel{10.27}{=} m \cdot (b - a)$  proving that

$$\| \mathcal{S}(f, \mathbb{P}) \| \leq m \cdot (b - a) \quad (11.109)$$

If now  $\{\mathbb{P}_i\}_{i \in \mathbb{N}_0}$  is such that  $\lim_{i \rightarrow \infty} \mu(\mathbb{P}_i) = 0$  then  $\| \int_a^b f \| = \left\| \lim_{i \rightarrow \infty} \mathcal{S}(f, \mathbb{P}_i) \right\|$   $\| \cdot \|$  is continuous and  $\| \mathcal{S}(f, \mathbb{P}_i) \| \leq 11.109$  and  $11.349 m \cdot (b - a)$   $\square$

**Lemma 11.446.** Let  $c \in [a, b] \subseteq \mathbb{R}$ ,  $a < c < b$  and let  $\mathbb{P}_1 = (\{t_i^1\}_{i \in \{1, \dots, n_1\}}, \{s_i^2\}_{i \in \{1, \dots, n_1-1\}})$  be a tagged partition on  $[a, c]$ ,  $\mathbb{P}_2 = (\{t_i^2\}_{i \in \{1, \dots, n_2\}}, \{s_i^2\}_{i \in \{1, \dots, n_2-1\}})$  a tagged partition on  $[c, b]$  then  $\mathbb{P} = \mathbb{P}_1 \cup \mathbb{P}_2$  defined by  $\mathbb{P} = (\{t_i\}_{i \in \{1, \dots, n_1+n_2-1\}}, \{s_i\}_{i \in \{1, \dots, n_1+n_2-2\}})$  where  $t_i = \begin{cases} t_i^1 & \text{if } i \in \{1, \dots, n_1\} \\ t_{i-n_1+1}^2 & \text{if } i \in \{n_1+1, \dots, n_1+n_2-1\} \end{cases}$  and  $s_i = \begin{cases} s_i^1 & \text{if } i \in \{1, \dots, n_1-1\} \\ s_{i-n_1+1}^2 & \text{if } i \in \{n_1, \dots, n_1+n_2-2\} \end{cases}$  then  $\mathbb{P}$  is a tagged partition of  $[a, b]$  with  $\mu(\mathbb{P}) = \max(\mu(\mathbb{P}_1), \mu(\mathbb{P}_2))$ . If  $\langle X, \| \cdot \| \rangle$  is a normed space and  $f: [a, b] \rightarrow X$  a function then  $\mathcal{S}(f, \mathbb{P}) = \mathcal{S}(f, \mathbb{P}_1) + \mathcal{S}(f, \mathbb{P}_2)$

**Proof.** To prove this note:

1.  $t_1 = t_1^1 = a$
2.  $t_{n_1+n_2-1} < n_2 \Rightarrow 2 \leq n_2 \Rightarrow 2 + (n_1 - 1) \leq n_2 \Rightarrow n_1 + 1 \leq n_1 + n_2 - 1 \Rightarrow t_{(n_1+n_2-1)-n_1+1}^2 = t_{n_2}^2 = b$
3. If  $i \in \{1, \dots, n_1+n_2-2\}$  then we have either
  - a. ( $i \in \{1, \dots, n_1-1\}$ ) then  $s_i = s_i^1 \in [t_i^1, t_{i+1}^1] \stackrel{1 \leq i, i+1 \leq n_1}{=} [t_i, t_{i+1}]$
  - b. ( $i = n_1$ ) then  $s_{n_1} = s_{(n_1-n_1+1)}^2 = s_1^2 \in [t_1^2, t_{1+1}^2] = [c, t_1^2] = [t_{n_1}^1, t_2^2] = [t_{n_1}, t_{(n_1+1)-n_1+1}^2] = [t_{n_1}, t_{n_1+1}]$
  - c. ( $i \in \{n_1+1, \dots, n_1+n_2-1\}$ ) then  $s_i = s_{i-n_1+1}^2 \in [t_{i-n_1+1}^2, t_{(i-n_1+1)+1}^2] = [t_i, t_{i+1}]$
4.  $\{|t_{i+1} - t_i| \mid i \in \{1, \dots, n_1+n_2-2\}\} = \{|t_{i+1} - t_i| \mid i \in \{1, \dots, n_1-1\}\} \cup \{|t_{i+1} - t_i| \mid i \in \{n_1\}\} \cup \{|t_{i+1} - t_i| \mid i \in \{n_1+1, \dots, (n_1+n_2-1)-1\}\} = \{|t_{i+1}^1 - t_i^1| \mid i \in \{1, \dots, n_1-1\}\} \cup \{|t_{(n_1+1)-n_1+1}^2 - t_{n_1}^1| \mid i \in \{1, \dots, n_1+n_2-2\}\} = \{|t_{i+1}^1 - t_i^1| \mid i \in \{1, \dots, n_1-1\}\} \cup \{|t_2^2 - c| \mid i \in \{1, \dots, n_1-1\}\} \cup \{|t_{(j+(n_1+1)-2)-n_1+1+1}^2 - t_{(j+(n_1+1)-2)-n_1+1}^2| \mid j \in \{2, \dots, n_2-1\}\} = \{|t_{i+1}^1 - t_i^1| \mid i \in \{1, \dots, n_1-1\}\} \cup \{|t_2^2 - t_1^2| \mid i \in \{1, \dots, n_1-1\}\} \cup \{|t_{j+1}^2 - t_{j+1}^2| \mid j \in \{2, \dots, n_2-1\}\} = \{|t_{i+1}^1 - t_i^1| \mid i \in \{1, \dots, n_1-1\}\} \cup \{|t_{j+1}^2 - t_{j+1}^2| \mid j \in \{1, \dots, n_2-1\}\}$  so that  $\mu(\mathbb{P}) = \max(\{|t_{i+1} - t_i| \mid i \in \{1, \dots, n_1+n_2-2\}\}) = \max(\{\max(\{|t_{i+1}^1 - t_i^1| \mid i \in \{1, \dots, n_1-1\}\}), \max(\{|t_{j+1}^2 - t_{j+1}^2| \mid j \in \{1, \dots, n_2-1\}\})\}) = \max(\mu(\mathbb{P}_1), \mu(\mathbb{P}_2))$
5.  $\mathcal{S}(f, \mathbb{P}) = \sum_{i=1}^{(n_1+n_2-1)-1} f(s_i) \cdot (t_{i+1} - t_i) = (\sum_{i=1}^{n_1-1} f(s_i) \cdot (t_{i+1} - t_i)) + (\sum_{i=n_1}^{n_1} f(s_i) \cdot (t_{i+1} - t_i)) + (\sum_{i=n_1+1}^{(n_1+n_2-1)-1} f(s_{i-n_1+1}^2) \cdot (t_{(i-n_1+1)+1}^2 - t_{(i-n_1+1)}^2)) = (\sum_{i=1}^{n_1-1} f(s_i^1) \cdot (t_{i+1}^1 - t_i^1)) + (f(s_{n_1}) \cdot (t_{n_1+1} - t_{n_1})) + (f(s_1^2) \cdot (t_{(n_1-n_1+1)+1}^2 - t_{n_1}^1)) + (\sum_{j=2}^{n_2-1} f(s_{j+(n_1+1)-2-n_1+1}^2) \cdot (t_{((j+(n_1+1)-2)-n_1+1)+1}^2 - t_{(j+(n_1+1)-2)-n_1+1}^2)) = (\sum_{i=1}^{n_1-1} f(s_i^1) \cdot (t_{i+1}^1 - t_i^1)) + (f(s_{n_1}) \cdot (t_2^2 - c)) + (\sum_{j=2}^{n_2-1} f(s_j^2) \cdot (t_{j+1}^2 - t_j^2)) = (\sum_{i=1}^{n_1-1} f(s_i^1) \cdot (t_{i+1}^1 - t_i^1)) + (f(s_{n_1}) \cdot (t_2^2 - t_1^2)) + (\sum_{j=2}^{n_2-1} f(s_j^2) \cdot (t_{j+1}^2 - t_j^2)) = (\sum_{i=1}^{n_1-1} f(s_i^1) \cdot (t_{i+1}^1 - t_i^1)) + (\sum_{j=1}^{n_2-1} f(s_j^2) \cdot (t_{j+1}^2 - t_j^2)) = \mathcal{S}(f, \mathbb{P}_1) + \mathcal{S}(f, \mathbb{P}_2) \quad \square$

**Theorem 11.447.** Let  $[a, b] \subseteq \mathbb{R}$ ,  $a \leq b$ ,  $\langle X, \|\cdot\| \rangle$  a real metric space and  $c \in [a, b]$  and  $f \in \mathcal{C}([a, b], X)$  then  $f \in \mathcal{C}([a, c], X)$  and  $f \in \mathcal{C}([c, b], X)$  so that  $f: [a, b] \rightarrow X$ ,  $f: [a, c] \rightarrow X$  and  $f: [c, b] \rightarrow X$  are integrable and  $\int_a^c f + \int_c^b f = \int_a^b f$

**Proof.** We have the following cases to consider:

1. ( $a = b$ ) Then  $a = c = b$  so that  $\int_a^b f = \int_a^a = 0 = 0 + 0 = \int_a^a f + \int_a^a f = \int_a^c f + \int_c^b f$
2. ( $a < b \wedge a = c$ ) Then  $\int_a^b f = 0 + \int_a^b f = \int_a^a f + \int_a^b f = \int_a^c f + \int_c^b f$
3. ( $a < b \wedge c = b$ ) Then  $\int_a^b f = \int_a^b f + 0 = \int_a^b f + \int_b^b = \int_a^c f + \int_c^b f$
4. ( $a < c \wedge c < b$ ) As  $f: [a, c] \rightarrow X$  and  $f: [c, b] \rightarrow X$  are continuous we have that they are integrable so that by 11.440 there exists a  $\{\mathbb{P}_i^1\}_{i \in \mathbb{N}_0}$  (tagged partitions on  $[a, c]$ ) and  $\{\mathbb{P}_i^2\}_{i \in \mathbb{N}_0}$  (tagged partitions on  $[c, b]$ ) with  $\lim_{i \rightarrow \infty} \mu(\mathbb{P}_i^1) = 0$ ,  $\lim_{i \rightarrow \infty} \mu(\mathbb{P}_i^2) = 0$  so that  $\lim_{i \rightarrow \infty} \mathcal{S}(f, \mathbb{P}_i^1) = \int_a^c f$  and  $\lim_{i \rightarrow \infty} \mathcal{S}(f, \mathbb{P}_i^2) = \int_c^b f$ . Take then  $\{\mathbb{P}_i^1 \cup \mathbb{P}_i^2\}_{i \in \mathbb{N}_0}$  then by 11.446 we have that  $\{\mathbb{P}_i^1 \cup \mathbb{P}_i^2\}_{i \in \{1, \dots, n\}}$  is a family of tagged partitions of  $[a, b]$ . Also if  $\varepsilon > 0$  then  $\exists N_1, N_2$  such that if  $i \geq N_1$  then  $\mu(\mathbb{P}_i^1) < \frac{\varepsilon}{2}$ , if  $i \geq N_2$  then  $\mu(\mathbb{P}_i^2) < \frac{\varepsilon}{2}$ . So if  $i \geq N = \max(N_1, N_2)$  then  $|\mu(\mathbb{P}_i^1 \cup \mathbb{P}_i^2) - 0| = |\mu(\mathbb{P}_i^1 \cup \mathbb{P}_i^2)| \stackrel{11.446}{=} |\mu(\mathbb{P}_i^1) + \mu(\mathbb{P}_i^2)| \leq |\mu(\mathbb{P}_i^1)| + |\mu(\mathbb{P}_i^2)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$  proving that  $\lim_{i \rightarrow \infty} \mu(\mathbb{P}_i^1 \cup \mathbb{P}_i^2) = 0$ . Next  $\int_a^c f + \int_c^b f = \lim_{i \rightarrow \infty} \mathcal{S}(f, \mathbb{P}_i^1) + \lim_{i \rightarrow \infty} \mathcal{S}(f, \mathbb{P}_i^2) \stackrel{11.353}{=} \lim_{i \rightarrow \infty} (\mathcal{S}(f, \mathbb{P}_i^1) + \mathcal{S}(f, \mathbb{P}_i^2)) \stackrel{11.446}{=} \lim_{i \rightarrow \infty} \mathcal{S}(f, \mathbb{P}_i^1 \cup \mathbb{P}_i^2) \stackrel{11.446}{=} \int_a^b f$

□

**Theorem 11.448.** Let  $[a, b] \subseteq \mathbb{R}$ ,  $a \leq b$ ,  $\langle X, \|\cdot\|_X \rangle$  a real normed space,  $\langle Y, \|\cdot\|_Y \rangle$  a real Banach space (so that by 11.364  $L(X, Y)$  is a real Banach space,  $f \in \mathcal{C}([a, b], L(X, Y))$ ). If we define  $\forall x \in X$   $f(\cdot)(x): [a, b] \rightarrow Y$  by  $f(\cdot)(x)(t) = f(t)(x)$  then we have  $(\int_a^b f)(x) = \int_a^b f(\cdot)(x)$

**Proof.** First note that given a  $x \in X$  the function  $\varphi_x: L(X, Y) \rightarrow Y$  defined by  $\varphi_x(f) = f(x)$  is continuous as  $\|\varphi_x(f) - \varphi_x(g)\|_Y = \|f(x) - g(x)\|_X = \|(f - g)(x)\| \leq \|f - g\| \cdot \|x\|_X$  so if  $\varepsilon > 0$  take then  $\delta = \frac{\varepsilon}{\|x\|_X + 1}$  so that if  $\|f - g\| < \delta$  then  $\|\varphi_x(f) - \varphi_x(g)\| < \frac{\|x\|_X \cdot \varepsilon}{\|x\|_X + 1} < \varepsilon$ . Second  $\varphi_x$  is linear as is proved by  $\varphi_x(\alpha \cdot f + \beta \cdot g) = (\alpha \cdot f + \beta \cdot g)(x) = \alpha \cdot f(x) + \beta \cdot g(x) = \alpha \cdot \varphi_x(f) + \beta \cdot \varphi_x(g)$ . Note that if  $f \in \mathcal{C}([a, b], L(X, Y))$  we have that  $\forall t \in [a, b]$  that  $f(\cdot)(x)(t) = f(t)(x) = \varphi_x(f(t)) = (\varphi_x \circ f)(t)$  so that  $f(\cdot)(x) = \varphi_x \circ f: [a, b] \rightarrow Y$  is continuous as  $\varphi_x$  and  $f$  is continuous. From the continuity it follows that  $f(\cdot)(x)$  is integrable and thus that  $\int_a^b f(\cdot)(x)$  exists. Thirdly if  $\mathbb{P} = (\{t_i\}_{i \in \{1, \dots, n\}}, \{s_i\}_{i \in \{1, \dots, n-1\}})$  is a tagged partition of  $[a, b]$  then  $\mathcal{S}(f(\cdot)(x), \mathbb{P}) = \mathcal{S}(\varphi_x \circ f, \mathbb{P}) = \sum_{i=1}^n (\varphi_x \circ f)(s_i) \cdot (t_{i+1} - t_i) = \sum_{i=1}^n \varphi_x(f(s_i)) \cdot (t_{i+1} - t_i) = \varphi_x(\sum_{i=1}^n f(s_i) \cdot (t_{i+1} - t_i)) = \varphi_x(\mathcal{S}(f, \mathbb{P}))$ . So if  $\{\mathbb{P}_i\}_{i \in \mathbb{N}_0}$  is a family of tagged partitions with  $\lim_{i \rightarrow \infty} \mu(\mathbb{P}_i) = 0$  then  $\int_a^b f(\cdot)(x) = \lim_{i \rightarrow \infty} \mathcal{S}(f(\cdot)(x), \mathbb{P}_i) = \lim_{i \rightarrow \infty} \varphi_x(\mathcal{S}(f, \mathbb{P}_i)) \stackrel{11.334}{=} \varphi_x(\lim_{i \rightarrow \infty} \mathcal{S}(f, \mathbb{P}_i)) = \varphi_x(\int_a^b f) = (\int_a^b f)(x)$

□

## 11.13 Connected Sets

**Definition 11.449.** A topological space  $\langle X, \mathcal{T} \rangle$  is **connected** if  $\forall U_1, U_2 \in \mathcal{T}$  with  $U_1 \cap U_2 = \emptyset$  and  $X = U_1 \cup U_2$  we have that either  $U_1 = \emptyset$  or  $U_2 = \emptyset$  (in other words  $X$  is not the union of two non empty disjoint open sets). The topological space is **disconnected** if it is not connected, or in other words there exists open  $U_1, U_2 \in \mathcal{T}$  with  $U_1 \neq \emptyset \neq U_2$  and  $X = U_1 \cup U_2$

**Theorem 11.450.** Let  $\langle X, \mathcal{T} \rangle$  be a topological space then the following are equivalent

1.  $\emptyset$  and  $X$  are the only subsets of  $X$  that are open and closed
2.  $X$  is connected
3.  $\forall A_1, A_2$  closed sets with  $A_1 \cap A_2 = \emptyset$  and  $X = A_1 \cup A_2$  we have either  $A_1 = \emptyset$  or  $A_2 = \emptyset$

**Proof.** We make in this proof heavy use of 1.29!

1. **(1  $\Rightarrow$  2)** Assume that  $U_1, U_2$  are disjoint open sets with  $X = U_1 \cup U_2$  then  $U_2 = X \setminus U_1$  which is closed. According to (1) we must then have that either  $U_2 = \emptyset$  or  $U_2 = X$  and then  $U_1 = X \setminus U_2 = X \setminus X = \emptyset$  so  $U_1 = \emptyset$  proving that  $X$  is connected.
2. **(2  $\Rightarrow$  3)** Assume that we have two disjoint closed sets  $A_1, A_2$  with  $X = A_1 \cup A_2$  then  $A_1 = X \setminus A_2$  is open and  $A_2 = X \setminus A_1$  is open, so by the connectedness of  $X$  we must have either  $A_1 = \emptyset$  or  $A_2 = \emptyset$ .
3. **(3  $\Rightarrow$  1)** Let  $A \subseteq X$  and  $A$  is open and closed then  $X \setminus A$  is closed,  $A \cap (X \setminus A) = \emptyset$  and  $X = A \cup (X \setminus A)$ . By (3) we must have either  $A = \emptyset$  or  $X \setminus A = \emptyset \Rightarrow A = X$ .  $\square$

**Definition 11.451.** Let  $\langle X, \mathcal{T} \rangle$  be a topological space and let  $A \subseteq X$  then  $A$  is connected if it is connected in the subspace topology

**Theorem 11.452.** Let  $\langle \mathbb{R}, \|\cdot\| \rangle$  be the real normed space then  $\forall a, b \in \mathbb{R}$  with  $a \leq b$  we have that  $[a, b]$  is a connected subset.

**Proof.** First as  $[a, b] = \left[ \frac{a+b}{2} - \frac{b-a}{2}, \frac{a+b}{2} + \frac{b-a}{2} \right] = \bar{B}_{\|\cdot\|} \left( \frac{a+b}{2}, \frac{b-a}{2} \right)$  which is closed by 11.56 so that  $[a, b]$  is a closed subset of  $\mathbb{R}$ . So if  $A \subseteq [a, b]$  is closed in the subspace topology of  $[a, b]$  then  $A = A' \cap [a, b]$  with  $A'$  is closed in  $\mathbb{R}$  so that  $A$  is also closed in  $\mathbb{R}$ . Assume now that  $[a, b]$  is not connected then by the previous theorem we can find disjoint closed sets (closed in  $[a, b]$  and thus also in  $\mathbb{R}$ )  $A_1, A_2$  with  $A_1, A_2 \neq \emptyset$  such that  $[a, b] = A_1 \cup A_2$ . We can always assume that  $b \in A_2$  (otherwise switch  $1 \leftrightarrow 2$ ). As  $A_1 \subseteq A_1 \cup A_2 = [a, b]$  we have that  $A_1$  is bounded above by  $b$  and is not empty so by order conditional completeness of the reals (see 9.43)  $c = \sup(A_1)$  exists. If now there exists a  $x, y \in \mathbb{R}$  with  $c \in ]x, y[$  then  $x < c < y$  and by the definition of a supremum there exists a  $a' \in A_1$  such that  $x < a' \leq c < y \Rightarrow a' \in ]x, y[ \Rightarrow x, y \cap A_1 \neq \emptyset \Rightarrow c \in A_1$ . Since  $b$  is an upper bound of  $A_1$ ,  $b \in A_2$  and  $A_1 \cap A_2 = \emptyset \Rightarrow b \notin A_1$  we must have  $c < b \Rightarrow b - c > 0$ . Now if  $x \in ]c, b[$  then  $x > c$  and thus  $x \notin A_1$  ( $c = \sup(A_1) \Rightarrow ]c, b[ \subseteq [a, b] \setminus A_1 = A_2$ ). If  $c \in ]x, y[$  then if  $\varepsilon = \min(y - c, b - c) > 0$  then for  $c + \frac{\varepsilon}{2}$  we have  $c < c + \frac{\varepsilon}{2} < c + \varepsilon \leq y, b \Rightarrow c + \frac{\varepsilon}{2} \in ]c, y[ \cap ]c, b[ \subseteq ]x, y[ \cap A_2$  proving that  $]x, y[ \cap A_2 \neq \emptyset$  and as  $A_2$  is closed we must have that  $c \in A_2$ . So we have finally that  $c \in A_1 \cap A_2 = \emptyset$  a contradiction. So our assumption that  $[a, b]$  is not connected is wrong and thus  $[a, b]$  is connected what we set out to prove.  $\square$

**Theorem 11.453.** Let  $\langle \mathbb{R}, \|\cdot\| \rangle$  be the a normed space then the following are equivalent for  $I \subseteq \mathbb{R}$

1.  $I$  is connected
2.  $I$  is a generalized interval

**Proof.**

1. **(1  $\Rightarrow$  2)** We proceed by contradiction so assume that  $I$  is connected but not a generalized interval. Using the properties of a generalized interval (see 2.188) we have that  $\exists x, y \in I$  with  $x < y$  and  $\exists z \in \mathbb{R} \setminus I$  such that  $x < z < y$  then  $U_1 = ]-\infty, z[ \cap I$  contains  $x$  and is open in  $I$ , also  $U_2 = ]z, \infty[ \cap I$  contains  $y$  and is open in  $I$ . Then as  $z \in I$  we have  $I \subseteq ]-\infty, \infty[ \setminus \{z\} = ]-\infty, z[ \cup ]z, \infty[ \Rightarrow I = I \cap (]-\infty, z[ \cup ]z, \infty[) = U_1 \cup U_2$ , as also  $U_1 \cap U_2 \subseteq ]-\infty, z[ \cap ]z, \infty[ = \emptyset \Rightarrow U_1 \cap U_2 = \emptyset$  we have that  $I$  is connected, contradicting our assumption that  $I$  is not connected.  $I$  must thus be connected.
2. **(2  $\Leftarrow$  1)** We prove this also by contradiction. So assume that  $I$  is a generalized interval and that  $I$  is not connected. Then we have open sets  $U_1, U_2 \subseteq \mathbb{R}$  such that  $I = (U_1 \cap I) \cup (U_2 \cap I)$ ,  $\emptyset = (U_1 \cap I) \cap (U_2 \cap I) = U_1 \cap U_2 \cap I$  and  $U_1 \cap I \neq \emptyset \neq U_2 \cap I$ . So  $\exists x \in U_1 \cap I$  and  $\exists y \in U_2 \cap I$  and as  $\emptyset = (U_1 \cap I) \cap (U_2 \cap I)$  we must have  $x \neq y$ , assume that  $x < y$  (if  $y < x$  interchange the role of  $U_1$  and  $U_2$ ) then using 2.188 we have  $[x, y] \subseteq I \Rightarrow [x, y] = [x, y] \cap I = [x, y] \cap (U_1 \cap I) \cup (U_2 \cap I) = (U_1 \cap I \cap [x, y]) \cup (U_2 \cap I \cap [x, y]) = (U_1 \cap [x, y]) \cup (U_2 \cap [x, y]) \Rightarrow [x, y] = (U_1 \cap [x, y]) \cup (U_2 \cap [x, y])$ , as also  $x \in U_1 \cap [x, y]$ ,  $y \in U_2 \cap [x, y]$  and  $(U_1 \cap [x, y]) \cap (U_2 \cap [x, y]) \subseteq (U_1 \cap I) \cap (U_2 \cap I) = \emptyset$  we have that  $[x, y]$  is covered by two non empty disjoint open sets and thus not connected, contradicting the fact that  $[x, y]$  is connected by the previous theorem.  $\square$

**Theorem 11.454.** *Let  $\langle X, \mathcal{T}_X \rangle$  be a connected topological space,  $\langle Y, \mathcal{T}_Y \rangle$  a topological space and  $f: X \rightarrow Y$  a continuous function then  $f(X)$  is connected.*

**Proof.** We proceed by contradiction. Assume that  $f(X)$  is not connected then we have open sets  $V_1, V_2$  in  $Y$  such that  $f(X) = (f(X) \cap V_1) \cup (f(X) \cap V_2) \subseteq V_1 \cup V_2$ ,  $\exists y_1 \in f(X) \cap V_1$ ,  $\exists y_2 \in f(X) \cap V_2$  and  $(f(X) \cap V_1) \cap (f(X) \cap V_2) = \emptyset$ . The  $\exists x_1, x_2 \in X$  such that  $f(x_1) = y_1 \in V_1$ ,  $f(x_2) = y_2 \in V_2$  thus  $x_1 \in f^{-1}(V_1) = U_1$ ,  $x_2 \in f^{-1}(V_2) = U_2$  where  $U_1, U_2$  are open because of continuity of  $f$ . Further if  $x \in X \Rightarrow f(x) \in f(X) \subseteq V_1 \cup V_2 \Rightarrow x \in f^{-1}(V_1 \cup V_2) \subseteq f^{-1}(V_1) \cup f^{-1}(V_2) = U_1 \cup U_2$  proving that  $X \subseteq U_1 \cup U_2$ , as trivially  $U_1 \cup U_2 \subseteq X$  we have  $X = U_1 \cup U_2$ . Finally assume that  $y \in U_1 \cap U_2 \Rightarrow f(y) \in f(U_1) \cap f(U_2) \subseteq V_1 \cap V_2 = \emptyset$  a contradiction so that we have  $U_1 \cap U_2 = \emptyset$ . So we have proved that  $X$  is covered by non-empty disjoint open sets contradicting the fact that  $X$  is connected.  $\square$

**Theorem 11.455.** *Let  $f: [a, b] \rightarrow \mathbb{R}$  a continuous function then  $f([a, b]) = [c, d]$  (so  $f$  retains its minimum and maximum settings).*

**Proof.** As  $f$  is continuous and  $[a, b]$  is compact we have by 11.253 that  $f([a, b])$  is compact and thus by 11.251 it is bounded. As by the previous theorem  $f([a, b])$  is connected we have by 11.453 that  $f([a, b])$  is a generalized interval and as it is bounded there exists a  $c, d \in \mathbb{R}$  such that  $f([a, b]) = [c, d]$ .  $\square$

**Corollary 11.456. (Intermediate Value Theorem)** *Let  $f: [a, b] \rightarrow \mathbb{R}$  be a continuous function then if  $y$  is such that either  $f(a) < y < f(b)$  or  $f(b) < y < f(a)$  we have that there exists a  $x \in [a, b]$  such that  $f(x) = y$*

**Proof.** Using the previous theorem we have that  $f([a, b]) = [c, d]$  so that  $f(a), f(b) \in [c, d]$

**$f(a) < y < f(b)$ .** then  $y \in [f(a), f(b)]$  and  $f(a) < f(b)$  where  $f(a), f(b) \in [c, d]$  so by the properties of the generalized interval (see 2.188 (2)) we have that  $[f(a), f(b)] \subseteq [c, d] = f([a, b])$  hence  $y \in f([a, b])$  and thus  $\exists x \in [a, b]$  such that  $y = f(x)$ .

**$f(b) < y < f(a)$ .** then  $y \in [f(b), f(a)]$  and  $f(b) < f(a)$  where  $f(a), f(b) \in [c, d]$  so by the properties of the generalized interval (see 2.188 (2)) we have that  $[f(b), f(a)] \subseteq [c, d] = f([a, b])$  hence  $y \in f([a, b])$  and thus  $\exists x \in [a, b]$  such that  $y = f(x)$ .  $\square$



# Chapter 12

## Fundamental theorem of algebra and the Spectral theorems

### 12.1 Fundamental theorem of algebra

#### 12.1.1 Polynomials

To prove the fundamental theorem of algebra which says that every non constant complex polynomial has a zero we need first a lot of definitions and lemmas.

**Definition 12.1.** A function  $p: \mathbb{C} \rightarrow \mathbb{C}$  is a complex polynomial if there exists a  $n \in \mathbb{N}$  and a  $\{a_i\}_{i \in \{0, \dots, n\}} \subseteq \mathbb{C}$  such that  $p(z) = \sum_{i=0}^n a_i \cdot z^i$ . The set of complex polynomials is noted by  $\mathcal{P}$

Next we prove that complex polynomials are uniquely defined by its coefficients.

**Lemma 12.2.** Let  $p: \mathbb{C} \rightarrow \mathbb{C}$  be a polynomial with coefficients  $\{a_i\}_{i \in \{1, \dots, n\}}, n \in \mathbb{N}$  such that  $\forall x \in \mathbb{R} \subseteq \mathbb{C}$  we have  $p(x) = 0$  then  $\forall i \in \{1, \dots, n\}$  we have  $a_i = 0$

**Proof.** Assume that  $\exists i \in \{0, \dots, n\}$  such that  $a_i \neq 0$  then  $\mathcal{N} = \{i \in \{0, \dots, n\} | a_i \neq 0\} \neq \emptyset$  so that  $m = \max(\mathcal{N})$  and  $\forall x \in \mathbb{R}$  we have  $0 = p(x) = \sum_{i=1}^n a_i \cdot x^i = \sum_{i=1}^m a_i \cdot x^i$  and  $a_m \neq 0$ . We have then for  $m$  the following cases to consider

**$m = 0$ .** then  $0 = p(1) = a_0$  a contradiction

**$0 < m$ .** then let  $x = \frac{\sum_{i=0}^{m-1} |a_i|}{|a_m|} + 1$  (defined as  $a_m \neq 0$ ) then  $1 \leq x$  and thus by 9.31 we have  $\forall j \in \{0, \dots, m-1\}$  we have  $x^j \leq x^{m-1}$ , so that

$$\begin{aligned}
 \left| \sum_{i=0}^{m-1} a_i \cdot x^i \right| &\stackrel{\text{9.31}}{\leq} \sum_{i=0}^{m-1} |a_i| \cdot x^i \\
 &= \sum_{i=0}^{m-1} |a_i| \cdot |x|^i \\
 &\leq \sum_{i=0}^{m-1} |a_i| \cdot |x|^{m-1} \\
 &= |x|^{m-1} \cdot \sum_{i=0}^{m-1} |a_i| \\
 &= |x|^{m-1} \cdot |a_m| \cdot (x-1) \\
 &< |x|^m \cdot |a_m|
 \end{aligned}$$

hence  $\sum_{i=0}^{m-1} a_i \cdot x^i \neq -a_m \cdot x^m \Rightarrow \sum_{i=0}^m a_i \cdot x^i = \sum_{i=0}^{m-1} a_i \cdot x^i + a_m \cdot x^m \neq 0$  contradicting  $\sum_{i=0}^m a_i \cdot z^i = 0$ .

As in all cases we have a contradiction the assumption is wrong and thus  $\forall i \in \{0, \dots, n\}$  we have  $a_i = 0$   $\square$

**Corollary 12.3.** *Let  $p: \mathbb{C} \rightarrow \mathbb{C}$  be a polynomial such that  $\forall z \in \mathbb{C}$  we have that  $p(z) = \sum_{i=0}^n a_i \cdot z^i = \sum_{i=0}^n b_i \cdot z^i$  then  $\forall i \in \{0, \dots, n\}$  we have  $a_i = b_i$*

**Proof.** As  $\forall z \in \mathbb{C}$  we have that  $\sum_{i=0}^n a_i \cdot z^i = \sum_{i=0}^n b_i \cdot z^i \Rightarrow \sum_{i=1}^n (a_i - b_i) \cdot z^i = 0$  we have by the previous lemma that  $\forall i \in \{0, \dots, n\}$  we have  $a_i - b_i \Rightarrow a_i = b_i$   $\square$

**Definition 12.4.** *Let  $p: \mathbb{C} \rightarrow \mathbb{C}$  be a polynomial defined by  $p(x) = \sum_{i=1}^n a_i \cdot x^i$  then if we define  $\mathcal{A} = \{i \mid a_i \neq 0\}$  we define  $\text{ord}(p)$  by  $\text{ord}(p) = \begin{cases} 0 & \text{if } \mathcal{A} = \emptyset \\ \max(\mathcal{A}) & \text{if } \mathcal{A} \neq \emptyset \end{cases}$*

**Theorem 12.5.** *Given the set of polynomials then we have that  $\text{ord}: \mathcal{P} \rightarrow \mathbb{N}$  defined by  $\text{ord}(p)$  is a function. Furthermore there exists a unique  $\{c_i\}_{i \in \{0, \dots, \text{ord}(p)\}}$  such that  $\forall z \in \mathbb{C}$  we have  $p(z) = \sum_{i=0}^{\text{ord}(p)} c_i \cdot z^i$*

**Proof.** We must proof that the value of the function is independent of the representation  $p$ , so let  $p: \mathbb{C} \rightarrow \mathbb{C}$  be such that  $p(z) = \sum_{i=0}^n a_i \cdot z^i = \sum_{i=0}^m b_i \cdot z^i$  (where we may assume without loosing generality that  $n \leq m$ ) then if we define  $\{a'_i\}_{i \in \{1, \dots, m\}}$  by  $a'_i = \begin{cases} a_i & \text{if } i \in \{0, \dots, n\} \\ 0 & \text{if } i \in \{n+1, \dots, m\} \end{cases}$  then  $\forall z \in \mathbb{C}$  we have  $p(z) = \sum_{i=0}^m a'_i \cdot z^i = \sum_{i=0}^n b_i \cdot z^i$ . Using 12.3 we conclude then that  $\forall i \in \{1, \dots, m\}$  we have  $a'_i = b_i$  hence  $\forall i \in \{1, \dots, n\}$  we have  $a_i = b_i$  and  $\forall i \in \{n+1, \dots, m\}$  we have  $b_i = 0$ . Hence  $\{i \in \{0, \dots, m\} \mid a_i \neq 0\} = \{i \in \{0, \dots, n\} \mid b_i \neq 0\}$  so that  $\min(\{i \in \{0, \dots, m\} \mid a_i \neq 0\}) = \min(\{i \in \{0, \dots, n\} \mid b_i \neq 0\})$  proving that  $\text{ord}(p)$  is independent of the representation of  $p$ . For the last part take  $\{c_i\}_{i \in \{0, \dots, \text{ord}(p)\}}$  defined by  $c_i = a_i$  for every  $i \in \{0, \dots, \text{ord}(p)\}$ .  $\square$

**Definition 12.6.** *A polynomial  $p \in \mathcal{P}$  is **non constant** if  $\text{ord}(p) \in \mathbb{N}_0$  (in other words if there exists a  $\{a_i\}_{i \in \{1, \dots, n\}}$  with  $a_n \neq 0$  such that  $\forall z \in \mathbb{C}$  we have  $p(z) = \sum_{i=0}^n a_i \cdot z^i$*

We show next that the product of a polynomial is again a polynomial, first we need a little lemma.

**Lemma 12.7.** *Let  $n, m \in \mathbb{N}_0$  then  $\{0, \dots, n\} \times \{0, \dots, m\} = \bigcup_{i \in \{0, \dots, n+m\}} \{(r, s) \in \{0, \dots, n\} \times \{0, \dots, m\} \mid r+s=i\}$  and  $\forall i, j \in \{0, \dots, n+m\}$  with  $i \neq j$  we have  $\{(r, s) \in \{0, \dots, n\} \times \{0, \dots, m\} \mid r+s=i\} \cap \{(r, s) \in \{0, \dots, n\} \times \{0, \dots, m\} \mid r+s=j\} = \emptyset$*

**Proof.** Given  $i \in \{0, \dots, n+m\}$  define  $I_i = \{(r, s) \in \{0, \dots, n\} \times \{0, \dots, m\} \mid r+s=i\} \subseteq \{0, \dots, n\} \times \{0, \dots, m\}$ . So trivially we have that  $\bigcup_{i \in \{0, \dots, n+m\}} I_i \subseteq \{0, \dots, n\} \times \{0, \dots, m\}$ , further  $(k, l) \in \{0, \dots, n\} \times \{0, \dots, m\}$  then  $k+l \leq n+m$  hence  $(k, l) \in I_{k+l} \subseteq \bigcup_{i \in \{0, \dots, n+m\}} I_i$  proving that

$$\{0, \dots, n\} \times \{0, \dots, m\} = \bigcup_{i \in \{0, \dots, n+m\}} \{(r, s) \in \{0, \dots, n\} \times \{0, \dots, m\} \mid r+s=i\}$$

Further if  $i, j \in \{0, \dots, n+m\}$  with  $i = j$  we have for  $(r, s) \in \{(r, s) \in \{0, \dots, n\} \times \{0, \dots, m\} \mid r+s=i\} \cap \{(r, s) \in \{0, \dots, n\} \times \{0, \dots, m\} \mid r+s=j\}$  we have  $r+s=i \neq j = r+s$  a contradiction so we have that  $\{(r, s) \in \{0, \dots, n\} \times \{0, \dots, m\} \mid r+s=i\} \cap \{(r, s) \in \{0, \dots, n\} \times \{0, \dots, m\} \mid r+s=j\} = \emptyset$ .  $\square$

**Theorem 12.8.** *Let  $p, q \in \mathcal{P}$  be two polynomials defined by  $p(z) = \sum_{i=0}^n a_i \cdot z^i$  and  $q(z) = \sum_{i=0}^m b_i \cdot z^i$  then  $p \cdot q: \mathbb{C} \rightarrow \mathbb{C}$  defined by  $(p \cdot q)(z) = p(z) \cdot q(z)$  is also a polynomial defined by  $(p \cdot q)(z) = \sum_{i=0}^{n+m} (\sum_{(k, l) \in \{(r, s) \in \{0, \dots, n\} \times \{0, \dots, m\} \mid r+s=i\}} a_k \cdot b_l) \cdot z^i$ . Further if  $\text{ord}(p) = n \in \mathbb{N}_0$  and  $\text{ord}(q) = m \in \mathbb{N}_0$  then  $\text{ord}(p \cdot q) = n+m \in \mathbb{N}_0$*

**Proof.** Let  $z \in \mathbb{C}$  then

$$\begin{aligned}
 p(z) \cdot q(z) &= \left( \sum_{i=0}^n a_i \cdot z^i \right) \cdot \left( \sum_{j=0}^m b_j \cdot z^j \right) \\
 &= \left( \sum_{i \in \{0, \dots, n\}} a_i \cdot z^i \right) \cdot \left( \sum_{j \in \{0, \dots, m\}} b_j \cdot z^j \right) \\
 &\stackrel{10.60}{=} \sum_{(i,j) \in \{0, \dots, n\} \times \{0, \dots, m\}} a_i \cdot b_j \cdot z^{i+j} \\
 &= \sum_{(i,j) \in \{0, \dots, n\} \times \{0, \dots, m\}} a_i \cdot b_j \cdot z^{i+j} \\
 &\stackrel{10.46 \text{ togehter with 12.7}}{=} \sum_{i \in \{0, \dots, n+m\}} \left( \sum_{(k,l) \in \{(r,s) \in \{0, \dots, n\} \times \{0, \dots, m\} \mid r+s=i\}} a_k \cdot b_l \cdot z^{k+l} \right) \\
 &= \sum_{i \in \{0, \dots, n+m\}} \left( \sum_{(k,l) \in \{(r,s) \in \{0, \dots, n\} \times \{0, \dots, m\} \mid r+s=i\}} a_k \cdot b_l \cdot z^i \right) \\
 &= \sum_{i \in \{0, \dots, n+m\}} \left( \sum_{(k,l) \in \{(r,s) \in \{0, \dots, n\} \times \{0, \dots, m\} \mid r+s=i\}} a_k \cdot b_l \right) \cdot z^i \\
 &= \sum_{i=0}^{n+m} \left( \sum_{(k,l) \in \{(r,s) \in \{0, \dots, n\} \times \{0, \dots, m\} \mid r+s=i\}} a_k \cdot b_l \right) \cdot z^i
 \end{aligned}$$

Finally if  $\text{ord}(p), \text{ord}(q) \in \mathbb{N}_0$  then  $a_n, a_m \neq 0$  hence  $a_n \cdot a_m \neq 0$ , as  $\{(r, s) \in \{0, \dots, n\} \times \{0, \dots, m\} \mid r+s=n+m\} = \{(n, m)\}$  we have that the coefficient of  $z^{n+m}$  is  $a_n \cdot a_m \neq 0$  hence  $\text{ord}(p \cdot q) = n+m$ .  $\square$

**Definition 12.9.** Let  $n \in \mathbb{N}_0$  and  $\{p_i\}_{i \in \{1, \dots, n\}} \subseteq \mathcal{P}$  a finite non empty family of polynomials then  $\prod_{i=1}^n p_i: \mathbb{C} \rightarrow \mathbb{C}$  is defined by  $(\prod_{i=1}^n p_i)(z) = \prod_{i=1}^n p_i(z)$

Using mathematical induction we can then prove that  $\prod_{i=1}^n p_i$  is also a polynomial

**Lemma 12.10.** Let  $n \in \mathbb{N}_0$  and  $\{p_i\}_{i \in \{1, \dots, n\}} \subseteq \mathcal{P}$  a finite non empty family of polynomials then  $\prod_{i=1}^n p_i: \mathbb{C} \rightarrow \mathbb{C}$  is a polynomial with  $\text{ord}(\prod_{i=1}^n p_i) = \sum_{i=1}^n \text{ord}(p_i)$

**Proof.** Let  $\mathcal{S} = \{n \in \mathbb{N}_0 \mid \text{if } \{p_i\}_{i \in \{1, \dots, n\}} \subseteq \mathcal{P} \text{ then } \prod_{i=1}^n p_i: \mathbb{C} \rightarrow \mathbb{C} \text{ is a polynomial with } \text{ord}(\prod_{i=1}^n p_i) = \sum_{i=1}^n \text{ord}(p_i)\}$  then we have

**1  $\in \mathcal{S}$ .** then  $\forall z \in \mathbb{C}$  we have  $(\prod_{i=1}^1 p_i)(z) = p_1(z)$  proving that  $\prod_{i=1}^1 p_i = p_1$  a polynomial and  $\text{ord}(\prod_{i=1}^1 p_i) = \text{ord}(p_1) = \sum_{i=1}^1 \text{ord}(p_i)$  hence  $1 \in \mathcal{S}$

**$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$ .** then  $\forall z \in \mathbb{C}$  we have that  $(\prod_{i=1}^{n+1} p_i)(z) = \prod_{i=1}^{n+1} p_i(z) = p_{n+1}(z) \cdot (\prod_{i=1}^n p_i)(z)$  proving that  $\prod_{i=1}^{n+1} p_i = p_{n+1} \cdot \prod_{i=1}^n p_i$  which as  $n \in \mathcal{S}$  is a product of polynomials and thus by 12.8 is a polynomial, further

$$\begin{aligned}
 \text{ord}\left(\prod_{i=1}^{n+1} p_i\right) &= \text{ord}\left(p_{n+1} \cdot \prod_{i=1}^n p_i\right) \\
 &\stackrel{12.8}{=} \text{ord}(p_{n+1}) + \text{ord}\left(\prod_{i=1}^n p_i\right) \\
 &\stackrel{n \in \mathcal{S}}{=} \text{ord}(p_{n+1}) + \sum_{i=1}^n \text{ord}(p_i) \\
 &= \sum_{i=1}^{n+1} \text{ord}(p_i)
 \end{aligned}$$

hence  $n+1 \in \mathcal{S}$   $\square$

### 12.1.2 Divergent sequences

We need also the concept of divergent limits to prove the fundamental theorem of algebra

**Definition 12.11.** Let  $f: \mathbb{R} \rightarrow \mathbb{R}$  be a function then

1.  $\lim_{x \rightarrow \infty} f(x) = \infty$  if  $\forall M \in \mathbb{R}$  there exists a  $N \in \mathbb{R}$  such that  $\forall x \geq N$  we have  $f(x) \geq M$
2.  $\lim_{x \rightarrow \infty} f(x) = -\infty$  if  $\forall M \in \mathbb{R}$  there exists a  $N \in \mathbb{R}$  such that  $\forall x \geq N$  we have  $f(x) \leq M$

**Note 12.12.** Let  $f: \mathbb{R} \rightarrow \mathbb{R}$  be a function and consider then  $-f: \mathbb{R} \rightarrow \mathbb{R}$  defined by  $(-f)(x) = -f(x)$

1. If  $\lim_{x \rightarrow \infty} f(x) = \infty$  then  $\lim_{x \rightarrow \infty} (-f)(x) = -\infty$
2. If  $\lim_{x \rightarrow \infty} f(x) = -\infty$  then  $\lim_{x \rightarrow \infty} (-f)(x) = \infty$

**Proof.**

1. Take  $M \in \mathbb{R}$  then there exists a  $N$  such that  $\forall x \geq N$  we have  $-M \leq f(x) \Rightarrow (-f)(x) \leq M$  proving that  $\lim_{x \rightarrow \infty} (-f)(x) = -\infty$
2. Take  $M \in \mathbb{R}$  then there exists a  $N$  such that  $\forall x \geq N$  we have  $f(x) \leq -M \Rightarrow M \leq (-f)(x)$  proving that  $\lim_{x \rightarrow \infty} (-f)(x) = \infty$   $\square$

**Lemma 12.13.** Let  $n \in \mathbb{N}_0$ ,  $\{a_i\}_{i \in \{0, \dots, n\}} \subseteq \mathbb{R}$  then

1. If  $a_n > 0$  we have  $\lim_{x \rightarrow \infty} (\sum_{i=0}^n a_i \cdot x^i) = \infty$
2. If  $a_n < \infty$  we have  $\lim_{x \rightarrow \infty} (\sum_{i=0}^n a_i \cdot x^i) = -\infty$

**Proof.**

1. we prove this by induction so let  $\mathcal{S} = \left\{ n \in \mathbb{N}_0 \mid \forall \{a_i\}_{i \in \{0, \dots, n\}} \subseteq \mathbb{R} \vdash a_n > 0 \text{ we have } \lim_{x \rightarrow \infty} (\sum_{i=0}^n a_i \cdot x^i) = \infty \right\}$  then we have  
 $1 \in \mathcal{S}$ . Let  $M \in \mathbb{R}$  and take  $N = \frac{M - a_0}{a_1}$  then if  $x \geq N$  we have  $M - a_0 = a_1 \cdot N \leq a_1 \cdot x = a_1 \cdot x_1$  hence  $M \leq a_1 \cdot x_1 + a_0 = \sum_{i=0}^1 a_i \cdot x^i$  proving that  $1 \in \mathcal{S}$   
 $n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$ . Let  $\{a_i\}_{i \in \{0, \dots, n+1\}} \subseteq \mathbb{R} \vdash a_{n+1} > 0$  then  $\sum_{i=0}^{n+1} a_i \cdot x^i = (\sum_{i=1}^{n+1} a_i \cdot x^i) + a_0 = x \cdot (\sum_{i=1}^{n+1} a_i \cdot x^{i-1}) + a_0 = x \cdot (\sum_{i=0}^n a_{i+1} \cdot x^i) + a_0$ , hence if we define  $\{b_i\}_{i \in \{0, \dots, n\}}$  by  $b_i = a_{i+1}$ , then we have

$$\sum_{i=1}^{n+1} a_i \cdot x^i = x \cdot \left( \sum_{i=0}^n b_i \cdot x^i \right) + a_0 \wedge b_n = a_{n+1} > 0$$

Take  $M \in \mathbb{R}$  then as  $n \in \mathcal{S}$  there exists a  $N' \in \mathbb{R}$  such that if  $N \leq x$  we have that  $\max(1, M - a_0) \leq \sum_{i=0}^n b_i \cdot x^i$ . Take then  $N = \max(N', 1)$  then we have  $1 \leq x$  and as  $0 < \max(1, M - a_0)$  we have  $\max(1, M - a_0) \leq x \cdot \max(1, M - a_0)$ . As  $0 < 1 \leq x$  then we have  $x \cdot \max(1, M - a_0) \leq x \cdot (\sum_{i=0}^n b_i \cdot x^i)$ . So we conclude that  $M - a_0 = \max(1, M - a_0) \leq x \cdot (\sum_{i=0}^n b_i \cdot x^i)$  and thus  $M \leq (\sum_{i=0}^n b_i \cdot x^i) + a_0 = \sum_{i=1}^{n+1} a_i \cdot x^i$ . This proves that  $n+1 \in \mathcal{S}$

2. If  $a_n < 0$  then  $0 < (-a_n)$  and we have by (1) that  $\lim_{x \rightarrow \infty} (\sum_{i=0}^n (-a_i) \cdot x^i) = \infty$  and thus by 12.12 we have that  $\lim_{x \rightarrow \infty} (\sum_{i=0}^n a_i \cdot x^i) = \lim_{x \rightarrow \infty} (-(\sum_{i=0}^n a_i \cdot x^i)) = -\infty$   $\square$

### 12.1.3 Properties of $\mathbb{C}$ needed for the fundamental theorem

We define now a norm on  $\mathbb{C}$  that we will use instead of the classical norm on  $\mathbb{C}$  that includes a square root.

**Lemma 12.14.**  $\| \cdot \|: \mathbb{C} \rightarrow \mathbb{R}$  by  $\|z\| = |\operatorname{Re}(z)| + |\operatorname{Im}(z)|$  is a norm

**Proof.** We have

1. If  $z \in \mathbb{C}$  then trivially  $0 \leq \|z\|$

2. If  $\alpha \in \mathbb{R}$  then for  $z = x + i \cdot y$  we have that  $\|\alpha \cdot z\| = \|\alpha \cdot x + i \cdot \alpha \cdot y\| = |\alpha \cdot x| + |\alpha \cdot y| = |\alpha| \cdot (|x| + |y|) = \alpha \cdot \|z\|$
3. If  $z_1 = x_1 + i \cdot y_1, z_2 = x_2 + i \cdot y_2 \in \mathbb{C}$  then  $\|z_1 + z_2\| = \|(x_1 + i \cdot y_1) + (x_2 + i \cdot y_2)\| = \|(x_1 + x_2) + i \cdot (y_1 + y_2)\| = |x_1 + x_2| + |y_1 + y_2| \leq |x_1| + |x_2| + |y_1| + |y_2| = \|z_1\| + \|z_2\|$
4. Take  $z = z + i \cdot y \in \mathbb{C}$  such that  $\|z\| = 0$  then  $|x| + |y| = 0 \xrightarrow[0 \leq |x|, |y|]{} |x| = 0 = |y| \Rightarrow x = 0 \wedge y = 0 \Rightarrow z = 0$   $\square$

The above norm has the following properties

**Lemma 12.15.** *We have the following properties for the above norm*

1. If  $z \in \mathbb{R}$  then  $\|z\| = |z|$
2.  $\forall z \in \mathbb{C}$  we have that  $\|\bar{z}\| = \|z\|$
3.  $\forall z, w \in \mathbb{C}$  we have that  $\frac{\|z\| \cdot \|w\|}{2} \leq \|z \cdot w\| \leq \|z\| \cdot \|w\|$
4.  $\forall z, w \in \mathbb{C}$  we have  $\|\operatorname{Re}(z \cdot w)\| \leq \|z\| \cdot \|w\|$
5.  $\forall n \in \mathbb{N}_0, z \in \mathbb{C}$  then  $\|z^n\| \leq \|z\|^n$
6.  $\forall n \in \mathbb{N}_0, z \in \mathbb{C}$  then  $\|\operatorname{Re}(z^n)\| \leq \|z\|^n$
7.  $\forall n \in \mathbb{N}_0, z \in \mathbb{C}$  then  $\frac{\|z\|^n}{2^n} \leq \|z^n\|$

**Proof.**

1. If  $z \in \mathbb{R}$  then  $\operatorname{Img}(z) = 0$  and  $\operatorname{Re}(z) = z$  thus  $\|z\| = |\operatorname{Re}(z)| + |\operatorname{Img}(z)| = |z| + 0 = |z|$
2. If  $z = x + i \cdot y$  then we have  $\|z\| = |x| + |y| = |x| + |-y| = \|x - i \cdot y\| = \|\bar{z}\|$
3. First we have given  $a, b, c, d \in \mathbb{R}$  that  $0 \leq (|a| - |b|)^2 = a^2 + b^2 - 2 \cdot |a| \cdot |b|$  proving that  $2 \cdot |a| \cdot |b| \leq a^2 + b^2$  hence  $(|a| + |b|)^2 = a^2 + b^2 + 2 \cdot |a| \cdot |b| \leq 2 \cdot (a^2 + b^2)$ . Similar we have that  $0 \leq (|c| - |d|)^2 = c^2 + d^2 - 2 \cdot |c| \cdot |d|$  proving that  $2 \cdot |c| \cdot |d| \leq c^2 + d^2$  hence  $(|c| + |d|)^2 = c^2 + d^2 + 2 \cdot |c| \cdot |d| \leq 2 \cdot (c^2 + d^2)$ . So  $(|a| + |b|)^2 \cdot (|c| + |d|)^2 \leq 2 \cdot (a^2 + b^2) \cdot (c^2 + d^2)$  and  $2 \cdot (|c| + |d|)^2 \cdot (a^2 + b^2) \leq 4 \cdot (c^2 + d^2)$  proving that

$$(|a| + |b|)^2 \cdot (|c| + |d|)^2 \leq 4 \cdot (a^2 + b^2) \cdot (c^2 + d^2) \quad (12.1)$$

Now

$$\begin{aligned} (a \cdot c - b \cdot d)^2 + (a \cdot d + b \cdot c)^2 &= a^2 \cdot c^2 + b^2 \cdot d^2 - 2 \cdot a \cdot c \cdot b \cdot d + a^2 \cdot d^2 + b^2 \cdot c^2 + 2 \cdot a \cdot d \cdot b \cdot c \\ &= a^2 \cdot c^2 + a^2 \cdot d^2 + b^2 \cdot d^2 + b^2 \cdot c^2 \\ &= a^2 \cdot (c^2 + d^2) + b^2 \cdot (c^2 + d^2) \\ &= (a^2 + b^2) \cdot (c^2 + d^2) \end{aligned}$$

Using 12.1 we have then

$$\begin{aligned} (|a| + |b|)^2 \cdot (|c| + |d|)^2 &\leq 4 \cdot ((a \cdot c - b \cdot d)^2 + (a \cdot d + b \cdot c)^2) \\ &\leq 4 \cdot ((a \cdot c - b \cdot d)^2 + (a \cdot d + b \cdot c)^2 + 2 \cdot (a \cdot c - b \cdot d) \cdot (a \cdot d + b \cdot c)) \\ &= 4 \cdot ((a \cdot c - b \cdot d) + (a \cdot d + b \cdot c))^2 \\ &= 4 \cdot |(a \cdot c - b \cdot d) + (a \cdot d + b \cdot c)|^2 \end{aligned}$$

proving taking the square root (see 9.72)  $(|a| + |b|) \cdot (|c| + |d|) \leq 2 \cdot |(a \cdot c - b \cdot d) + (a \cdot d + b \cdot c)| \leq |a \cdot c - b \cdot d| + |a \cdot d + b \cdot c|$

$$(|a| + |b|) \cdot (|c| + |d|) \leq 2 \cdot |a \cdot c - b \cdot d| + |a \cdot d + b \cdot c| \quad (12.2)$$

So if  $z = a + i \cdot b$  and  $w = c + i \cdot d$  then  $z \cdot w = (a \cdot c - b \cdot d) + i \cdot (a \cdot d + b \cdot c)$  so that  $\|z \cdot w\| = |a \cdot c - b \cdot d| + |a \cdot d + b \cdot c| \geq_{12.2} \frac{1}{2} \cdot (|a| + |b|) \cdot (|c| + |d|) = \frac{1}{2} \cdot \|z\| \cdot \|w\|$  proving that

$$\frac{1}{2} \cdot \|z\| \cdot \|w\| \leq \|z \cdot w\| \quad (12.3)$$

Further we have

$$\begin{aligned} |a \cdot c - b \cdot d| + |a \cdot d + b \cdot c| &\leq |a| \cdot |c| + |b| \cdot |d| + |a| \cdot |d| + |b| \cdot |c| \\ &= |a| \cdot (|c| + |d|) + |b| \cdot (|c| + |d|) \\ &= (|a| + |b|) \cdot (|c| + |d|) \end{aligned}$$

proving

$$|a \cdot c - b \cdot d| + |a \cdot d + b \cdot c| \leq (|a| + |b|) \cdot (|c| + |d|) \quad (12.4)$$

hence

$$\begin{aligned} \|z \cdot w\| &= |a \cdot c - b \cdot d| + |a \cdot d + b \cdot c| \\ &\leq_{12.4} (|a| + |b|) \cdot (|c| + |d|) \\ &= \|z\| \cdot \|w\| \end{aligned}$$

proving

$$\|z \cdot w\| \leq \|z\| \cdot \|w\| \quad (12.5)$$

(3) is thus proved by 12.3 and 12.5.

4. If  $z = a + i \cdot b$  and  $w = c + i \cdot d$  then

$$\begin{aligned} \|\operatorname{Re}(z \cdot w)\| &\stackrel{(1)}{=} |\operatorname{Re}((a \cdot c - b \cdot d) + (a \cdot d + b \cdot c))| \\ &= |a \cdot c - b \cdot d| \\ &\leq |a \cdot c - b \cdot d| + |a \cdot d + b \cdot c| \\ &= \|z \cdot w\| \\ &\leq_{(3)} \|z\| \cdot \|w\| \end{aligned}$$

5. This trivially proved by induction so let  $\mathcal{S} = \{n \in \mathbb{N}_0 \mid \|z^n\| \leq \|z\|^n\}$  then we have

**1**  $\in \mathcal{S}$ . then  $\|z^1\| = \|z\| = \|z\|^1$  proving that  $1 \in \mathcal{S}$

**n**  $\in \mathcal{S} \Rightarrow \mathbf{n+1} \in \mathcal{S}$ . then  $\|z^{n+1}\| = \|z \cdot z^n\| \leq_{(3)} \|z\| \cdot \|z\|^n = \|z\|^{n+1}$  proving that  $n+1 \in \mathcal{S}$

6. Take  $z = x + i \cdot y$  then we must consider two cases for  $n$

**n = 1**. then  $\|\operatorname{Re}(z^1)\| = \|x\| \stackrel{(1)}{=} |x| \leq |x| + |y| = \|z\|$

**n > 1**. then  $1 \leq n-1$  and  $\|\operatorname{Re}(z^n)\| = \|\operatorname{Re}(z \cdot z^{n-1})\| \leq \|z\| \cdot \|z^{n-1}\| \leq_{(3)} \|z\| \cdot \|z\|^{n-1} = \|z\|^n$

7. This is proved by induction so let  $\mathcal{S} = \left\{ n \in \mathbb{N} \mid \frac{\|z\|^n}{2^n} \leq \|z^n\| \right\}$  then we have

**1**  $\in \mathcal{S}$ . then as  $\frac{1}{2} < 1 \Rightarrow \frac{\|z\|}{2^1} = \frac{1}{2} \cdot \|z\| \leq 1 \cdot \|z\| = \|z\|^1$  we have that  $1 \in \mathcal{S}$

**n**  $\in \mathcal{S} \Rightarrow \mathbf{n+1} \in \mathcal{S}$ . we have  $\frac{\|z\| \cdot \|z^n\|}{2} \leq_{(3)} \|z \cdot z^n\| = \|z^{n+1}\|$  as  $n \in \mathcal{S}$  we have  $\frac{\|z\|^n}{2^n} \leq \|z^n\|$  and thus  $\frac{\|z\|^{n+1}}{2^{n+1}} = \frac{\|z\|}{2} \cdot \frac{\|z\|^n}{2^n} \leq \frac{\|z\|}{2} \cdot \|z^n\| \leq \|z^{n+1}\|$  proving that  $n+1 \in \mathcal{S}$   $\square$

In the proof of the fundamental theorem of algebra we use the extrem value theorem so we need continuity using the above norm. This is proved in the following theorem.

**Theorem 12.16.** *The following functions are continuous in the normed space  $(\mathbb{C}, \|\cdot\|)$*

1. Given  $\alpha \in \mathbb{C}$  then function  $(\alpha \cdot) : \mathbb{C} \rightarrow \mathbb{C}$  defined by  $(\alpha \cdot)(x) = \alpha \cdot x$  is continuous
2. Given  $n \in \mathbb{N}$  we have that  $(1_{\mathbb{C}})^n : \mathbb{C} \rightarrow \mathbb{C}$  is continuous
3. Given  $n \in \mathbb{N}_0$ ,  $\{\alpha_i\}_{i \in \mathbb{N} \setminus \{0, \dots, n\}}$  the function  $p : \mathbb{C} \rightarrow \mathbb{C}$  defined by  $p(x) = \sum_{i=0}^n \alpha_i \cdot x^i$  is continuous
4. The function  $\langle \cdot \rangle : \mathbb{C} \rightarrow \mathbb{R}$  defined by  $\langle z \rangle = z \cdot \bar{z} = |z|^2$  is continuous
5. Given  $n \in \mathbb{N}_0$ ,  $\{\alpha_i\}_{i \in \mathbb{N}_0}$  the function  $p \cdot \bar{p} : \mathbb{C} \rightarrow \mathbb{R}$  defined by  $(p \cdot \bar{p})(x) = (\sum_{i=0}^n \alpha_i \cdot x^i) \cdot (\sum_{i=0}^n \bar{\alpha}_i \cdot \bar{x}^i)$  is continuous
6. The function  $\bar{\cdot} : \mathbb{C} \rightarrow \mathbb{C}$  defined by  $\bar{\cdot}(z) = \bar{z}$  is continuous
7. Given  $z \in \mathbb{C}$  then the function  $(\cdot z) : \mathbb{R} \rightarrow \mathbb{C}$  defined by  $(\cdot z)(\alpha) = \alpha \cdot z$  is continuous
8. If  $p : \mathbb{C} \rightarrow \mathbb{C}$  is a polynomial and  $z \in \mathbb{C}$  then  $p_z : \mathbb{R} \rightarrow \mathbb{C}$  defined by  $p_z(\alpha) = p(\alpha \cdot z)$  is continuous.
9. The function  $\operatorname{Re} : \mathbb{C} \rightarrow \mathbb{R}$  defined by  $x \rightarrow \operatorname{Re}(x)$  is continuous

**Proof.**

1. As  $(\alpha \cdot) = \alpha \cdot 1_{\mathbb{C}}$  this follows from the fact that  $1_{\mathbb{C}}$  is continuous (see 11.141) and 11.175
2. As  $1_{\mathbb{C}}$  is a continuous function (see 11.141) this follows from 11.178.
3. If  $\{\alpha_i\}_{i \in \{0, \dots, n\}}$  we have that  $\forall i \in \{1, \dots, n\}$  that  $(\alpha_i \cdot) \circ (1_{\mathbb{C}})^i$  is continuous as the composition of two continuous functions (see (1) and (2). Hence usng 11.176 we have that  $\sum_{i=1}^n (\alpha_i \cdot) \circ (1_{\mathbb{C}})^i$  is continuous. As  $\forall x \in \mathbb{C}$  we have  $(\sum_{i=1}^n (\alpha_i \cdot) \circ (1_{\mathbb{C}})^i)(x) = \sum_{i=1}^n ((\alpha_i \cdot) \circ (1_{\mathbb{C}})^i) = \sum_{i=1}^n \alpha_i \cdot x^i = p(x)$  we have that  $p = \sum_{i=1}^n (\alpha_i \cdot) \circ (1_{\mathbb{C}})^i$  is continuous.
4. First let  $z, \delta \in \mathbb{C}$  then

$$\begin{aligned}
 |\langle z + \delta \rangle - \langle z \rangle| &= \|\langle z + \delta \rangle - \langle z \rangle\| \\
 &= \|(z + \delta) \cdot \overline{(z + \delta)} - z \cdot \bar{z}\| \\
 &= \|(z + \delta) \cdot (\bar{z} + \bar{\delta}) - z \cdot \bar{z}\| \\
 &= \|z \cdot \bar{z} + z \cdot \bar{\delta} + \delta \cdot \bar{z} + \delta \cdot \bar{\delta} - z \cdot \bar{z}\| \\
 &= \|z \cdot \bar{\delta} + \delta \cdot \bar{z} + \delta \cdot \bar{\delta}\| \\
 &\leq \|\|z\| \cdot \|\bar{\delta}\| + \|\delta\| \cdot \|\bar{z}\| + \|\delta\| \cdot \|\bar{\delta}\|\| \\
 &\stackrel{12.15 (3)}{\leq} \|\|z\| \cdot \|\bar{\delta}\| + \|\delta\| \cdot \|\bar{z}\| + \|\delta\| \cdot \|\bar{\delta}\|\| \\
 &\stackrel{12.15 (2)}{\leq} \|\|z\| \cdot \|\delta\| + \|\delta\| \cdot \|\bar{z}\| + \|\delta\| \cdot \|\bar{\delta}\|\| \\
 &= \|\delta\| \cdot (2 \cdot \|z\| + \|\delta\|) \\
 \end{aligned}$$

So given  $\varepsilon > 0$  we have if we choose  $\delta$  such that  $\|\delta\| < \min\left(1, \frac{\varepsilon}{2 \cdot \|z\| + 1}\right)$  we have that

$$\begin{aligned}
 \|\langle z + \delta \rangle - \langle z \rangle\| &< \|\delta\| \cdot (2 \cdot \|z\| + 1) \\
 &< \frac{\varepsilon}{2 \cdot \|z\| + 1} \cdot (2 \cdot \|z\| + 1) \\
 &= \delta
 \end{aligned}$$

proving continuity of  $\langle \rangle$ .

5. This follows as  $p \cdot \bar{p} = \langle \rangle \circ p$  a composition of continuous functions by (3) and (4)
6. Let  $z \in \mathbb{C}$  and take  $\varepsilon > 0$  then if  $z' \in \mathbb{C}$  is such that  $\|z - z'\| < \varepsilon$  we have that  $\|{}^-(z) - {}^-(z')\| = \|\bar{z} - \bar{z}'\| = \|\overline{z - z'}\| \stackrel{12.15(2)}{=} \|z - z'\| < \varepsilon$
7. Let  $z \in \mathbb{C}$ ,  $\alpha \in \mathbb{R}$ , given  $\varepsilon > 0$  define  $\delta = \frac{\varepsilon}{\|z\| + 1} > 0$  then we have  $\forall \alpha' \in \mathbb{R}$  with  $|\alpha - \alpha'| < \delta$  that  $\|(\cdot z)(\alpha) - (\cdot z)(\alpha')\| = \|\alpha \cdot z - \alpha' \cdot z\| = \|(\alpha - \alpha') \cdot z\| = |\alpha - \alpha'| \cdot \|z\| < \frac{\varepsilon}{\|z\| + 1} \cdot \|z\| < \varepsilon$  proving that  $(\cdot z)$  is continuous (see 11.157).
8. As  $p \circ (\cdot z) = p$  is the composition of two continuous functions (see (3) and (7)) we have that  $p \circ (\cdot z)$  is continuous.
9. Let  $z \in \mathbb{C}$  then given  $\varepsilon$  we have if  $\delta = \varepsilon$  that  $\forall z' \in \mathbb{C}$  with  $\|z - z'\| < \delta$  that  $|\operatorname{Re}(z) - \operatorname{Re}(z')| = |\operatorname{Re}(z - z')| \leq |\operatorname{Re}(z - z')| + |\operatorname{Im}(z - z')| = \|z - z'\| < \delta = \varepsilon$  proving continuity.  $\square$

We need in the proof also to calculate the value of  $(1+i)^{4 \cdot n+2}$  and  $(1-i)^{4 \cdot n+2}$

**Lemma 12.17.** Let  $n \in \mathbb{N}$  then  $(1+i)^{4 \cdot n+2} = 2 \cdot (-4)^n \cdot i$  and  $(1-i)^{4 \cdot n+2} = -2 \cdot (-4)^n \cdot i$

**Proof.** We use induction to prove this

1. Take  $\mathcal{S} = \{n \in \mathbb{N} \mid (1+i)^{4 \cdot n+2} = 2 \cdot (-4)^n \cdot i\}$  then we have  
 $0 \in \mathcal{S}$ .  $(1+i)^{4 \cdot 0+2} = (1+i)^2 = 1 + 2 \cdot i + i^2 = 2 \cdot i = 2 \cdot (-4)^0 \cdot i$  proving that  $0 \in \mathcal{S}$   
 $n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$ .

$$\begin{aligned}
 (1+i)^{4 \cdot (n+1)+2} &= (1+i)^{4 \cdot n+2+4} \\
 &= (1+i)^{4 \cdot n+2}(1+i)^4 \\
 &\stackrel{n \in \mathcal{S}}{=} 2 \cdot (-4)^n \cdot i \cdot (1+i)^2(1+i)^2 \\
 &= 2 \cdot (-4)^n \cdot i \cdot (2 \cdot i)(2 \cdot i) \\
 &= 2 \cdot (-4)^n \cdot i \cdot (-4) \\
 &= 2 \cdot (-4)^{n+1} \cdot i
 \end{aligned}$$

proving that  $n+1 \in \mathcal{S}$

2. Take  $\mathcal{S} = \{n \in \mathbb{N} \mid (1-i)^{4 \cdot n+2} = -2 \cdot (-4)^n \cdot i\}$  then we have

**0 ∈ S.**  $(1-i)^{4 \cdot 0+2} = (1-i)^2 = 1 - 2 \cdot i + (-i)^2 = -2 \cdot i = -2 \cdot (-4)^0 \cdot i$  proving that  $0 \in \mathcal{S}$   
 $n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$ .

$$\begin{aligned} (1-i)^{4 \cdot (n+1)+2} &= (1-i)^{4 \cdot n+2+4} \\ &= (1-i)^{4 \cdot n+2}(1-i)^4 \\ &\stackrel{n \in \mathcal{S}}{=} -2 \cdot (-4)^n \cdot i \cdot (1-i)^2(1-i)^2 \\ &= -2 \cdot (-4)^n \cdot i \cdot (-2 \cdot i)(-2 \cdot i) \\ &= -2 \cdot (-4)^n \cdot i \cdot (-4) \\ &= -2 \cdot (-4)^{n+1} \cdot i \end{aligned}$$

proving that  $n+1 \in \mathcal{S}$

□

The binomial expansion of the power of a sum is used in the theorem so we have to define the factorial and the binomial constants and prove the binomial formula.

**Definition 12.18.** Given  $n \in \mathbb{N}$  we have that  $n!$  is defined recursively by  $n! = \begin{cases} 1 & \text{if } n=0 \\ n \cdot (n-1)! & \text{if } n > 1 \end{cases}$

**Definition 12.19.** Given  $n \in \mathbb{N}$  and  $k \in \{0, \dots, n\}$  then  $\binom{n}{k} = \frac{n!}{k! \cdot (n-k)!}$

**Note 12.20.** If  $n \in \mathbb{N}$  then  $\binom{n}{0} = 1$  and  $\binom{n}{n} = 1$

**Proof.**  $\binom{n}{0} = \frac{n!}{0! \cdot (n-0)!} = \frac{n!}{1 \cdot n!} = 1$  and  $\binom{n}{n} = \frac{n!}{n! \cdot (n-n)!} = \frac{1}{0!} = 1$

□

**Theorem 12.21.** If  $n \in \mathbb{N}_0$  and  $0 < k \leq n$  then  $\binom{n}{k} + \binom{n}{k-1} = \binom{n+1}{k}$

**Proof.**

$$\begin{aligned} \binom{n}{k} + \binom{n}{k-1} &= \frac{n!}{k! \cdot (n-k)!} + \frac{n!}{(k-1)! \cdot (n-(k-1))!} \\ &= \frac{n! \cdot ((n+1)-k)}{k! \cdot (n-k)! \cdot ((n+1)-k)} + \frac{n! \cdot k}{k \cdot (k-1)! \cdot (n+1-k)} \\ &= \frac{n! \cdot ((n+1)-k)}{k! \cdot ((n+1)-k)!} + \frac{n! \cdot k}{k! \cdot ((n+1)-k)!} \\ &= \frac{n! \cdot (n+1)}{k! \cdot ((n+1)-k)!} \\ &= \frac{(n+1)!}{k! \cdot ((n+1)-k)!} \\ &= \binom{n+1}{k} \end{aligned}$$

□

**Theorem 12.22. (Binomial formula)** Let  $F$  be a field,  $n \in \mathbb{N}_0$ ,  $a, b \in F$  then  $(a+b)^n = \sum_{k=0}^n \left( \binom{n}{k} \cdot a^k \cdot b^{n-k} \right)$

**Proof.** We prove this theorem by mathematical induction so let  $\mathcal{S} = \left\{ n \in \mathbb{N}_0 \mid (a+b)^n = \sum_{k=0}^n \left( \binom{n}{k} \cdot a^k \cdot b^{n-k} \right) \right\}$  then we have

**1 ∈ S.**  $\sum_{k=0}^1 \left( \binom{1}{k} \cdot a^k \cdot b^{(n-k)} \right) = \binom{1}{0} \cdot a^0 \cdot b^{(1-0)} + \binom{1}{1} \cdot a^1 \cdot b^{(1-1)} \stackrel{12.20}{=} b+a = (a+b)^1$  proving that  $1 \in \mathcal{S}$ .

$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$ .

$$\begin{aligned}
(a+b)^{n+1} &= (a+b)^n \cdot (a+b) \\
&\stackrel{n \in \mathcal{S}}{=} \left( \sum_{k=0}^n \left( \binom{n}{k} \cdot a^k \cdot b^{(n-k)} \right) \right) \cdot (a+b) \\
&= \sum_{k=0}^n \left( \binom{n}{k} \cdot a^{k+1} \cdot b^{(n-k)} \right) + \sum_{k=0}^n \left( \binom{n}{k} \cdot a^k \cdot b^{(n-k+1)} \right) \\
&= \sum_{k=0}^n \left( \binom{n}{k} \cdot a^{k+1} \cdot b^{((n+1)-(k+1))} \right) + \sum_{k=0}^n \left( \binom{n}{k} \cdot a^k \cdot b^{((n+1)-k)} \right) \\
&= \sum_{k=1}^{n+1} \left( \binom{n}{k-1} \cdot a^k \cdot b^{((n+1)-k)} \right) + \sum_{k=0}^n \left( \binom{n}{k} \cdot a^k \cdot b^{((n+1)-k)} \right) \\
&= \binom{n}{n} \cdot a^{n+1} \cdot b^{((n+1)-(n+1))} + \sum_{k=1}^n \left( \binom{n}{k-1} \cdot a^k \cdot b^{((n+1)-k)} \right) + \sum_{k=1}^n \left( \binom{n}{k} \cdot a^k \cdot b^{((n+1)-k)} \right) \\
&\quad + \binom{0}{0} \cdot a^0 \cdot b^{((n+1)-0)} \\
&= a^{n+1} \cdot b^0 + \sum_{k=1}^n \left( \left( \binom{n}{k-1} + \binom{n}{k} \right) \cdot a^k \cdot b^{((n+1)-k)} \right) + a^0 \cdot b^{n+1} \\
&\stackrel{12.21}{=} \binom{n+1}{0} \cdot a^{n+1} \cdot b^{((n+1)-(n+1))} + \sum_{k=1}^n \left( \binom{n+1}{k} \cdot a^k \cdot b^{((n+1)-k)} \right) + \binom{n+1}{0} \cdot a^0 \cdot b^{n+1} \\
&= \sum_{k=0}^{n+1} \left( \binom{n+1}{k} \cdot a^k \cdot b^{((n+1)-k)} \right)
\end{aligned}$$

proving that  $n+1 \in \mathcal{S}$

□

**Corollary 12.23.** Let  $p: \mathbb{C} \rightarrow \mathbb{C}$  is a non constant polynomial and  $z_0 \in \mathbb{C}$  then  $p_{\rightarrow z_0}: \mathbb{C} \rightarrow \mathbb{C}$  defined by  $p_{\rightarrow z_0}(z) = p(z + z_0)$  is a non constant polynomial of order  $n$

**Proof.** We prove this by induction so let  $\mathcal{S} = \{n \in \mathbb{N} \mid p \text{ is a polynomial with } \text{ord}(p) \leq n \text{ then } p_{\rightarrow z_0} \text{ is a polynomial of order } n\}$  then we have

$1 \in \mathcal{S}$ . If  $p$  is a polynomial of order 1 then  $p(z) = a_1 \cdot z + a_0$  hence  $p_{\rightarrow z_0}(z) = p(z + z_0) = a_1 \cdot (z + z_0) + a_0 = a_1 \cdot z + (a_1 \cdot z_0 + a_0)$  a polynomial of order 1 proving that  $1 \in \mathcal{S}$

$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$ . Let  $p$  be a polynomial with  $\text{ord}(p) \leq n+1$  then we have either

**ord( $p$ )  $\leq n$ .** then as  $n \in \mathcal{S}$  we have  $p_{\rightarrow z_0}$  is a polynomial with order  $n$

**ord( $p$ )  $= n+1$ .** then there exists a  $\{a_i\}_{i \in \{1, \dots, n+1\}} \subseteq \mathbb{C}$  with  $a_{n+1} \neq 0$  such that

$$\begin{aligned}
p(z) &= \sum_{i=0}^{n+1} a_i \cdot z^i \\
&= \sum_{i=1}^{n+1} a_i \cdot z^i + a_0 \\
&= \sum_{i=1}^{n+1} z \cdot (a_i \cdot z^{i-1}) + a_0 \\
&= z \cdot \sum_{i=1}^{n+1} a_i \cdot z^{i-1} + a_0 \\
&= z \cdot \sum_{i=0}^n a_{i+1} \cdot z^i + a_0
\end{aligned}$$

$$\begin{aligned}
&= z \cdot \sum_{i=0}^n b_i \cdot z^i + a_0 \text{ where } \{b_i\}_{i \in \{0, \dots, n\}} \text{ is defined by } b_i = a_{i+1} \\
&= z \cdot q(z) + a_0 \text{ where } q \text{ is a polynomial with } \text{ord}(q) = n \text{ (as } b_n = a_{n+1} \neq 0\text{)}
\end{aligned}$$

As  $n \in \mathcal{S}$  we have that  $q_{\rightarrow z_0}$  is a polynomial of order  $n$  so that there exists a  $\{c_i\}_{i \in \{0, \dots, n\}} \subseteq \mathbb{C}$  such that  $c_n \neq 0$  and  $q(z + z_0) = q_{\rightarrow z_0}(z) = \sum_{i=0}^n c_i \cdot z^i$  hence we have

$$\begin{aligned}
p_{\rightarrow z_0}(z) &= (z + z_0) \cdot q(z + z_0) + a_0 \\
&= (z + z_0) \cdot q_{\rightarrow z_0}(z) + a_0 \\
&= (z + z_0) \cdot \sum_{i=0}^n c_i \cdot z^i + a_0 \\
&= \sum_{i=0}^n c_i \cdot z^{i+1} + \sum_{i=0}^n (z_0 \cdot c_i) \cdot z^i + a_0 \\
&= \sum_{i=1}^{n+1} c_{i-1} \cdot z^i + \sum_{i=1}^n (z_0 \cdot c_i) \cdot z^i + (z_0 \cdot c_0 + a_0) \\
&= c_n \cdot z^{n+1} + \sum_{i=1}^n (c_{i-1} + z_0 \cdot c_i) \cdot z^i + (z_0 \cdot c_0 + a_0) \\
&= \sum_{i=0}^{n+1} d_i \cdot z^i \quad \text{where } \{d_i\}_{i \in \{0, \dots, n+1\}} \text{ is defined by } d_i = \\
&\quad \begin{cases} c_n & \text{if } i = n+1 \\ (c_{i-1} + z_0 \cdot c_i) & \text{if } i \in \{1, \dots, n\} \\ z_0 \cdot c_0 + a_0 & \text{so that } d_{n+1} \neq 0 \end{cases}
\end{aligned}$$

which proved that  $p_{\rightarrow z_0}$  is a polynomial of order  $n+1$  proving that  $n+1 \in \mathcal{S}$   $\square$

**Lemma 12.24.** *Let  $p: \mathbb{C} \rightarrow \mathbb{C}$  be a non constant polynomial of order  $n \in \mathbb{N}_0$  then there exists a  $1 \leq k \leq n$  and a polynomial  $q$  of order  $n-k$  such that  $p(z) = p(0) + z^k \cdot q(z)$  and  $q(0) \neq 0$*

**Proof.** As  $p$  is a polynomial of order  $n$  there exists a  $\{a_i\}_{i \in \{0, \dots, n\}} \subseteq \mathbb{C}$  with  $a_n \neq 0$  so that  $p(z) = \sum_{i=0}^n a_i \cdot z^i$ . Hence  $p(0) = \sum_{i=0}^n a_i \cdot 0^i = a_0$  further let  $k = \min\{i \in \{1, \dots, n\} | a_i \neq 0\}$  (which is defined as  $a_n \neq 0$ ) then  $p(z) = \sum_{i=0}^n a_i \cdot z^i = \sum_{i=1}^n a_i \cdot z^i + a_0 = \sum_{i=k}^n a_i \cdot z^i + p(0) = \sum_{i=0}^{n-k} a_{i+k} \cdot z^{i+k} + p(0) = z^k \cdot \sum_{i=0}^{n-k} a_{i+k} \cdot z^i + p(0) = p(0) + q(z)$  where  $q(z) = \sum_{i=0}^{n-k} b_i$  and  $\{b_i\}_{i \in \{0, \dots, n-k\}}$  is defined by  $b_i = a_{i+k}$ . Finally  $q(0) = b_0 = a_k \neq 0$ .  $\square$

**Lemma 12.25.** *Let  $n \in \mathbb{N}_0$  and  $\{a_i\}_{i \in \{1, \dots, n\}} \subseteq \mathbb{C}$  then  $\overline{\sum_{i=1}^n a_i} = \sum_{i=1}^n \bar{a}_i$*

**Proof.** This is easy proved by induction so let  $S = \{n \in \mathbb{N}_0 | \forall \{a_i\}_{i \in \{1, \dots, n\}} \subseteq \mathbb{C} \text{ we have } \overline{\sum_{i=1}^n a_i} = \sum_{i=1}^n \bar{a}_i\}$  then we have

**1  $\in S$ .** this follows from  $\overline{\sum_{i=1}^1 a_i} = \bar{a}_1 = \sum_{i=1}^1 \bar{a}_i$  for  $\{a_i\}_{i \in \{1, \dots, 1\}} \subseteq \mathbb{C}$

**$n \in S \Rightarrow n+1 \in S$ .** we have for  $\{a_i\}_{i \in \{1, \dots, n+1\}}$  that

$$\begin{aligned}
\overline{\sum_{i=1}^{n+1} a_i} &= \overline{a_{n+1} + \sum_{i=1}^n a_i} \\
&= \overline{\bar{a}_{n+1}} + \overline{\sum_{i=1}^n a_i} \\
&\stackrel{n \in S}{=} \overline{\bar{a}_{n+1}} + \sum_{i=1}^n \bar{a}_i \\
&= \sum_{i=1}^{n+1} \bar{a}_i
\end{aligned}$$

proving that  $n+1 \in \mathcal{S}$

□

The following lemma shows how we can split up a set  $\{1, \dots, 2 \cdot n\}$  in a set of even and odd numbers.

**Lemma 12.26.** *Let  $n \in \mathbb{N}_0$  then  $\{0, \dots, 2 \cdot n\} = \{2 \cdot k \mid k \in \{0, \dots, n\}\} \sqcup \{2 \cdot k + 1 \mid k \in \{0, \dots, n-1\}\}$  and  $\{0, \dots, 2 \cdot n-1\} = \{2 \cdot k \mid k \in \{0, \dots, n-1\}\} \sqcup \{2 \cdot k + 1 \mid k \in \{0, \dots, n-1\}\}$*

**Proof.** We prove this by induction so let  $\mathcal{S} = \{n \in \mathbb{N}_0 \mid \{0, \dots, 2 \cdot n\} = \{2 \cdot k \mid k \in \{0, \dots, n\}\} \cup \{2 \cdot k + 1 \mid k \in \{1, \dots, n-1\}\}\}$  then we have

**1  $\in \mathcal{S}$ .** then  $\{0, 1, 2\} = \{0, 2\} \cup \{1\} = \{2 \cdot k \mid k \in \{0, 1\}\} \cup \{2 \cdot k + 1 \mid k \in \{0\}\}$  proving that  $1 \in \mathcal{S}$

**$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$ .**

$$\begin{aligned} \{1, \dots, 2 \cdot (n+1)\} &= \{1, \dots, 2 \cdot n\} \bigcup \{2 \cdot n + 1, 2 \cdot (n+1)\} \\ &\stackrel{n \in \mathcal{S}}{=} \{2 \cdot k \mid k \in \{0, \dots, n\}\} \bigcup \{2 \cdot k + 1 \mid k \in \{0, \dots, n-1\}\} \bigcup \{2 \cdot n + 1\} \bigcup \{2 \cdot (n+1)\} \\ &= \{2 \cdot k \mid k \in \{0, \dots, n+1\}\} \bigcup \{2 \cdot k + 1 \mid k \in \{0, \dots, n\}\} \\ &= \{2 \cdot k \mid k \in \{0, \dots, n+1\}\} \bigcup \{2 \cdot k + 1 \mid k \in \{0, \dots, (n+1)-1\}\} \end{aligned}$$

proving that  $n+1 \in \mathcal{S}$ .

It remains to prove that  $\{2 \cdot k \mid k \in \{0, \dots, n\}\} \cap \{2 \cdot k + 1 \mid k \in \{0, \dots, n-1\}\} = \emptyset$  so assume that  $l \in \{2 \cdot k \mid k \in \{0, \dots, n\}\} \cap \{2 \cdot k + 1 \mid k \in \{0, \dots, n-1\}\}$  then there exists a  $k \in \{0, \dots, n\}$  and a  $k' \in \{0, \dots, n-1\}$  such that  $2 \cdot k = l = 2 \cdot k' + 1$  which means that a even number is equal to a odd number which is not possible (see 6.56). Hence  $\{2 \cdot k \mid k \in \{0, \dots, n\}\} \cap \{2 \cdot k + 1 \mid k \in \{0, \dots, n-1\}\} = \emptyset$ .

For the second part let  $\mathcal{S} = \{n \in \mathbb{N}_0 \mid \{0, \dots, 2 \cdot n-1\} = \{2 \cdot k \mid k \in \{0, \dots, n-1\}\} \cup \{2 \cdot k + 1 \mid k \in \{0, \dots, n-1\}\}\}$  then we have

**1  $\in \mathcal{S}$ .** then  $\{0, 1\} = \{0\} \cup \{1\} = \{2 \cdot k \mid k \in \{0\}\} \cup \{2 \cdot k + 1 \mid k \in \{0\}\}$  proving that  $0 \in \mathcal{S}$

**$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$ .**

$$\begin{aligned} \{1, \dots, 2 \cdot (n+1)-1\} &= \{1, \dots, 2 \cdot n+1\} \\ &= \{1, \dots, 2 \cdot n-1\} \bigcup \{2 \cdot n+1\} \bigcup \{2 \cdot n\} \\ &\stackrel{n \in \mathcal{S}}{=} \{2 \cdot k \mid k \in \{0, \dots, n-1\}\} \bigcup \{2 \cdot k + 1 \mid k \in \{0, \dots, n-1\}\} \bigcup \{2 \cdot n+1\} \bigcup \{2 \cdot n\} \\ &= \{2 \cdot k \mid k \in \{0, \dots, n\}\} \bigcup \{2 \cdot k + 1 \mid k \in \{0, \dots, n\}\} \\ &= \{2 \cdot k \mid k \in \{0, \dots, (n+1)-1\}\} \bigcup \{2 \cdot k + 1 \mid k \in \{0, \dots, (n+1)-1\}\} \end{aligned}$$

proving that  $n+1 \in \mathcal{S}$

If  $l \in \{2 \cdot k \mid k \in \{0, \dots, n-1\}\} \cap \{2 \cdot k + 1 \mid k \in \{0, \dots, n-1\}\}$  then there exists a  $k \in \{0, \dots, n-1\}$  and a  $k' \in \{0, \dots, n-1\}$  such that  $2 \cdot k = l = 2 \cdot k' + 1$  which means that a even number is equal to a odd number which is not possible (see 6.56). Hence  $\{2 \cdot k \mid k \in \{0, \dots, n-1\}\} \cap \{2 \cdot k + 1 \mid k \in \{0, \dots, n-1\}\} = \emptyset$  □

**Lemma 12.27.** *Let  $n \in \mathbb{N}_0$  then if we define  $A = \{(k, l) \in \{0, \dots, n\}^2 \mid k < l\}$  and  $\forall m \in \{1, \dots, 2 \cdot n-1\}$   $B_m = \{(k, l) \in A \mid k + l = m\}$  then  $\bigsqcup_{m \in \{1, \dots, n-1\}} B_m = A$*

**Proof.** As  $\forall m \in \{1, \dots, 2 \cdot n - 1\}$  we have that  $B_m \subseteq A$  (by definition) we must have that  $\bigcup_{m \in \{1, \dots, 2 \cdot n - 1\}} B_m \subseteq A$ . If  $(k, l) \in A$  we have  $0 \leq k, l \leq n$  and as  $k < l$  we have we have that  $k + l < l + l \leq 2 \cdot n$  so that  $k + l \leq 2 \cdot n - 1$ , further from  $0 \leq k < l$  we have  $0 < l \Rightarrow 0 \leq k < k + l \Rightarrow 1 \leq k + l$  hence  $(k, l) \in B_{k+l}$  proving that  $A \subseteq \bigcup_{m \in \{1, \dots, n-1\}} B_m$ . And thus taking in account the previous that

$$A = \bigcup_{m \in \{1, \dots, 2 \cdot n - 1\}} B_m$$

If now  $m, m' \in \{1, \dots, n-1\}$  with  $m \neq m'$  then if  $(k, l) \in B_m \cap B_{m'}$  we have that  $m = k + l = m'$  contradicting that  $m \neq m'$  so we have that

$$A = \bigcup_{m \in \{1, \dots, 2 \cdot n - 1\}} B_m \quad \square$$

**Lemma 12.28.** Let  $n \in \mathbb{N}_0$  we have for  $k = 2 \cdot n$  and  $\zeta = \left(1 + \frac{i}{k}\right)^2$  that  $\operatorname{Re}(\zeta^k) < 0 < \operatorname{Img}(\zeta^k)$

**Proof.** We have

$$\begin{aligned} \left(1 + \frac{i}{k}\right)^{2 \cdot k} &= \left(\frac{i}{k} + 1\right)^{2 \cdot k} \\ &= \sum_{l=0}^{2 \cdot k} \left( \binom{2 \cdot k}{l} \left(\frac{i}{k}\right)^l \cdot 1^{2 \cdot k - l} \right) \\ &= \sum_{l \in \{0, \dots, 2 \cdot k\}} \left( \binom{2 \cdot k}{l} \left(\frac{i}{k}\right)^l \right) \\ &\stackrel{12.26}{=} \sum_{l \in \{2 \cdot j \mid j \in \{0, \dots, k\}\} \sqcup \{2 \cdot j + 1 \mid j \in \{0, \dots, k-1\}\}} \left( \binom{2 \cdot k}{l} \left(\frac{i}{k}\right)^l \right) \\ &= \left( \sum_{l \in \{2 \cdot j \mid j \in \{0, \dots, k\}\}} \left( \binom{2 \cdot k}{l} \left(\frac{i}{k}\right)^l \right) \right) + \left( \sum_{l \in \{2 \cdot j + 1 \mid j \in \{0, \dots, k-1\}\}} \left( \binom{2 \cdot k}{l} \left(\frac{i}{k}\right)^l \right) \right) \\ &= \left( \sum_{j \in \{0, \dots, k\}} \left( \binom{2 \cdot k}{2 \cdot j} \left(\frac{i}{k}\right)^{2 \cdot j} \right) \right) + \left( \sum_{j \in \{0, \dots, k-1\}} \left( \binom{2 \cdot k}{2 \cdot j + 1} \left(\frac{i}{k}\right)^{2 \cdot j + 1} \right) \right) \end{aligned}$$

Further as  $k = 2 \cdot n$

$$\begin{aligned} \sum_{j \in \{0, \dots, k\}} \left( \binom{2 \cdot k}{2 \cdot j} \left(\frac{i}{k}\right)^{2 \cdot j} \right) &= \sum_{j \in \{0, \dots, 2 \cdot n\}} \left( \binom{2 \cdot k}{2 \cdot j} \left(\frac{i}{k}\right)^{2 \cdot j} \right) \\ &\stackrel{12.26}{=} \sum_{j \in \{2 \cdot l \mid l \in \{0, \dots, n\}\} \sqcup \{2 \cdot l + 1 \mid l \in \{0, \dots, n-1\}\}} \left( \binom{2 \cdot k}{2 \cdot j} \left(\frac{i}{k}\right)^{2 \cdot j} \right) \\ &= \left( \sum_{j \in \{2 \cdot l \mid l \in \{0, \dots, n\}\}} \left( \binom{2 \cdot k}{2 \cdot j} \left(\frac{i}{k}\right)^{2 \cdot j} \right) \right) + \\ &\quad \left( \sum_{j \in \{2 \cdot l + 1 \mid l \in \{0, \dots, n-1\}\}} \left( \binom{2 \cdot k}{2 \cdot j} \left(\frac{i}{k}\right)^{2 \cdot j} \right) \right) \\ &= \left( \sum_{l \in \{0, \dots, n\}} \left( \binom{2 \cdot k}{2 \cdot 2 \cdot l} \left(\frac{i}{k}\right)^{2 \cdot 2 \cdot l} \right) \right) + \\ &\quad \left( \sum_{l \in \{0, \dots, n-1\}} \left( \binom{2 \cdot k}{2 \cdot (2 \cdot l + 1)} \left(\frac{i}{k}\right)^{2 \cdot (2 \cdot l + 1)} \right) \right) \end{aligned}$$

$$\begin{aligned}
&= \left( \sum_{l \in \{0, \dots, n\}} \left( \binom{2 \cdot k}{4 \cdot l} \left( \frac{i}{k} \right)^{4 \cdot l} \right) \right) + \\
&\quad \left( \sum_{l \in \{0, \dots, n-1\}} \left( \binom{2 \cdot k}{4 \cdot l+2} \left( \frac{i}{k} \right)^{4 \cdot l+2} \right) \right) \\
&\stackrel{9.22}{=} \left( \sum_{l \in \{0, \dots, n\}} \left( \binom{2 \cdot k}{4 \cdot l} \cdot \frac{1}{k^{4 \cdot l}} \right) \right) - \left( \sum_{l \in \{0, \dots, n-1\}} \left( \binom{2 \cdot k}{4 \cdot l+2} \cdot \frac{1}{k^{4 \cdot l+2}} \right) \right) \\
&= \left( \binom{2 \cdot k}{4 \cdot 0} \frac{1}{k^{4 \cdot 0}} + \binom{2 \cdot k}{4} \cdot \frac{1}{k^4} - \binom{2 \cdot k}{4 \cdot 0+2} \cdot \frac{1}{k^{4 \cdot 0+2}} \right) + \\
&\quad \sum_{l \in \{2, \dots, n\}} \left( \binom{2 \cdot k}{4 \cdot l} \cdot \frac{1}{k^{4 \cdot l}} \right) - \sum_{l \in \{1, \dots, n-1\}} \left( \binom{2 \cdot k}{4 \cdot l+2} \cdot \frac{1}{k^{4 \cdot l+2}} \right) \\
&= 1 - \binom{2 \cdot k}{2} \cdot \frac{1}{k^2} + \binom{2 \cdot k}{4} \cdot \frac{1}{k^4} + \sum_{l \in \{2, \dots, n\}} \left( \binom{2 \cdot k}{4 \cdot l} \cdot \frac{1}{k^{4 \cdot l}} \right) - \\
&\quad \sum_{l \in \{1, \dots, n-1\}} \left( \binom{2 \cdot k}{4 \cdot l+2} \cdot \frac{1}{k^{4 \cdot l+2}} \right) \\
&= 1 - \binom{2 \cdot k}{2} \cdot \frac{1}{k^2} + \binom{2 \cdot k}{4} \cdot \frac{1}{k^4} + \sum_{l \in \{1, \dots, n-1\}} \left( \binom{2 \cdot k}{4 \cdot (l+1)} \cdot \frac{1}{k^{4 \cdot (l+1)}} \right) - \\
&\quad \sum_{l \in \{1, \dots, n-1\}} \left( \binom{2 \cdot k}{2 \cdot (2 \cdot l+1)} \cdot \frac{1}{k^{2 \cdot (2 \cdot l+1)}} \right) \\
&= 1 - \binom{2 \cdot k}{2} \cdot \frac{1}{k^2} + \binom{2 \cdot k}{4} \cdot \frac{1}{k^4} + \sum_{l \in \{1, \dots, n-1\}} \left( \binom{2 \cdot k}{2 \cdot (2 \cdot l+1)+2} \cdot \frac{1}{k^{2 \cdot (2 \cdot l+1)+2}} \right) - \\
&\quad \sum_{l \in \{2 \cdot l+1 \mid l \in \{1, \dots, n-1\}\}} \left( \binom{2 \cdot k}{2 \cdot j} \cdot \frac{1}{k^{2 \cdot j}} \right) \\
&= 1 - \binom{2 \cdot k}{2} \cdot \frac{1}{k^2} + \binom{2 \cdot k}{4} \cdot \frac{1}{k^4} + \sum_{j \in \{2 \cdot l+1 \mid l \in \{1, \dots, n-1\}\}} \left( \binom{2 \cdot k}{2 \cdot j+2} \cdot \frac{1}{k^{2 \cdot j+2}} \right) - \\
&\quad \sum_{l \in \{2 \cdot l+1 \mid l \in \{1, \dots, n-1\}\}} \left( \binom{2 \cdot k}{2 \cdot j} \cdot \frac{1}{k^{2 \cdot j}} \right) \\
&= 1 - \binom{2 \cdot k}{2} \cdot \frac{1}{k^2} + \binom{2 \cdot k}{4} \cdot \frac{1}{k^4} + \sum_{j \in \{2 \cdot l+1 \mid l \in \{1, \dots, n-1\}\}} \left( \binom{2 \cdot k}{2 \cdot j+2} \cdot \frac{1}{k^{2 \cdot j+2}} - \binom{2 \cdot k}{2 \cdot j} \cdot \frac{1}{k^{2 \cdot j}} \right) \\
&\sum_{j \in \{0, \dots, k-1\}} \left( \binom{2 \cdot k}{2 \cdot j+1} \left( \frac{i}{k} \right)^{2 \cdot j+1} \right) = \sum_{j \in \{0, \dots, 2 \cdot n-1\}} \left( \binom{2 \cdot k}{2 \cdot j+1} \left( \frac{i}{k} \right)^{2 \cdot j+1} \right)
\end{aligned}$$

$$\begin{aligned}
&= \sum_{j \in \{2 \cdot l \mid l \in \{0, \dots, n-1\}\}} \left( \binom{2 \cdot k}{2 \cdot j + 1} \left( \frac{i}{k} \right)^{2 \cdot j + 1} \right) + \\
&\quad \sum_{j \in \{2 \cdot l + 1 \mid l \in \{0, \dots, n-1\}\}} \left( \binom{2 \cdot k}{2 \cdot j + 1} \left( \frac{i}{k} \right)^{2 \cdot j + 1} \right) \\
&= \sum_{l \in \{0, \dots, n-1\}} \left( \binom{2 \cdot k}{2 \cdot (2 \cdot l) + 1} \left( \frac{i}{k} \right)^{2 \cdot (2 \cdot l) + 1} \right) + \\
&\quad \sum_{l \in \{0, \dots, n-1\}} \left( \binom{2 \cdot k}{2 \cdot (2 \cdot l + 1) + 1} \left( \frac{i}{k} \right)^{2 \cdot (2 \cdot l + 1) + 1} \right) \\
&= \sum_{l \in \{0, \dots, n-1\}} \left( \binom{2 \cdot k}{4 \cdot l + 1} \left( \frac{i}{k} \right)^{4 \cdot l + 1} \right) + \\
&\quad \sum_{l \in \{0, \dots, n-1\}} \left( \binom{2 \cdot k}{4 \cdot l + 3} \left( \frac{i}{k} \right)^{4 \cdot l + 3} \right) \\
&\stackrel{9.22}{=} \sum_{l \in \{0, \dots, n-1\}} \left( \binom{2 \cdot k}{4 \cdot l + 1} \cdot \frac{i}{k^{4 \cdot l + 1}} \right) + \\
&\quad \sum_{l \in \{0, \dots, n-1\}} \left( \binom{2 \cdot k}{4 \cdot l + 3} \cdot \frac{-i}{k^{4 \cdot l + 3}} \right) \\
&= i \cdot \sum_{l \in \{0, \dots, n-1\}} \left( \binom{2 \cdot k}{4 \cdot l + 1} \cdot \frac{1}{k^{4 \cdot l + 1}} - \binom{2 \cdot k}{4 \cdot l + 3} \cdot \frac{1}{k^{4 \cdot l + 3}} \right) \\
&= i \cdot \sum_{l \in \{0, \dots, n-1\}} \left( \binom{2 \cdot k}{2 \cdot (2 \cdot l + 1) - 1} \cdot \frac{1}{k^{2 \cdot (2 \cdot l + 1) - 1}} - \right. \\
&\quad \left. \binom{2 \cdot k}{2 \cdot (2 \cdot l + 1) + 1} \cdot \frac{1}{k^{2 \cdot (2 \cdot l + 1) + 1}} \right) \\
&= i \cdot \sum_{j \in \{2 \cdot l + 1 \mid l \in \{0, \dots, n-1\}\}} \left( \binom{2 \cdot k}{2 \cdot j - 1} \cdot \frac{1}{k^{2 \cdot j - 1}} - \right. \\
&\quad \left. \binom{2 \cdot k}{2 \cdot j + 1} \cdot \frac{1}{k^{2 \cdot j + 1}} \right)
\end{aligned}$$

To summarize we have that

$$\begin{aligned}
\operatorname{Re} \left( \left( 1 + \frac{i}{k} \right)^{2 \cdot k} \right) &= 1 - \binom{2 \cdot k}{2} \cdot \frac{1}{k^2} + \binom{2 \cdot k}{4} \cdot \frac{1}{k^4} + \sum_{j \in \{2 \cdot l + 1 \mid l \in \{0, \dots, n-1\}\}} \left( \binom{2 \cdot k}{2 \cdot j + 2} \cdot \frac{1}{k^{2 \cdot j + 2}} - \right. \\
&\quad \left. \binom{2 \cdot k}{2 \cdot j} \cdot \frac{1}{k^{2 \cdot j}} \right) \tag{12.6}
\end{aligned}$$

$$\operatorname{Img} \left( \left( 1 + \frac{i}{k} \right)^{2 \cdot k} \right) = \sum_{j \in \{2 \cdot l + 1 \mid l \in \{0, \dots, n-1\}\}} \left( \binom{2 \cdot k}{2 \cdot j - 1} \cdot \frac{1}{k^{2 \cdot j - 1}} - \binom{2 \cdot k}{2 \cdot j + 1} \cdot \frac{1}{k^{2 \cdot j + 1}} \right) \tag{12.7}$$

Further we have that as  $2 \leq k \Rightarrow 4 \leq 2 \cdot k$

$$\begin{aligned}
\binom{2 \cdot k}{2} &= \frac{(2 \cdot k)!}{2! \cdot (2 \cdot k - 2)!} \\
&= \frac{(2 \cdot k) \cdot (2 \cdot k - 1) \cdot (2 \cdot k - 2)!}{2 \cdot (2 \cdot k - 2)!} \\
&= k \cdot (2 \cdot k - 1) \\
&= 2 \cdot k^2 - k
\end{aligned}$$

$$\begin{aligned}
\binom{2 \cdot k}{4} &= \frac{(2 \cdot k)!}{4! \cdot (2 \cdot k - 4)!} \\
&= \frac{(2 \cdot k)! \cdot (2 \cdot k - 3) \cdot (2 \cdot k - 2)}{4 \cdot 3 \cdot 2! \cdot (2 \cdot k - 4)! \cdot (2 \cdot k - 3) \cdot (2 \cdot k - 2)} \\
&= \frac{(2 \cdot k)! \cdot (2 \cdot k - 3) \cdot (2 \cdot k - 2)}{4 \cdot 3 \cdot 2! \cdot (2 \cdot k - 2)!} \\
&= \frac{(2 \cdot k - 3) \cdot (2 \cdot k - 2)}{12} \cdot \binom{2 \cdot k}{2} \\
&= \binom{2 \cdot k}{2} \cdot \frac{(2 \cdot k - 3) \cdot (k - 1)}{6} \\
&= \binom{2 \cdot k}{2} \cdot \frac{2 \cdot k^2 - 2 \cdot k - 3 \cdot k + 3}{6} \\
&= \binom{2 \cdot k}{2} \cdot \frac{2 \cdot k^2 - 5 \cdot k + 3}{6}
\end{aligned}$$

Hence

$$\begin{aligned}
1 - \binom{2 \cdot k}{2} \cdot \frac{1}{k^2} + \binom{2 \cdot k}{4} \cdot \frac{1}{k^4} &= 1 - \binom{2 \cdot k}{2} \cdot \left[ \frac{1}{k^2} - \frac{2 \cdot k^2 - 5 \cdot k + 3}{6 \cdot k^4} \right] \\
&= 1 - \binom{2 \cdot k}{2} \cdot \frac{1}{k^2} \cdot \left[ 1 - \frac{2 \cdot k^2 - 5 \cdot k + 3}{6 \cdot k^2} \right] \\
&= 1 - \frac{2 \cdot k^2 - k}{k^2} \cdot \left[ 1 - \frac{1}{3} + \frac{5 \cdot k - 3}{6 \cdot k^2} \right] \\
&= 1 - \left( 2 - \frac{1}{k} \right) \cdot \left( \frac{2}{3} + \frac{5 \cdot k - 3}{6 \cdot k^2} \right)
\end{aligned}$$

As  $0 < 2 \leq k$  we have  $\frac{1}{k} \leq \frac{1}{2} \Rightarrow -\frac{1}{2} \leq -\frac{1}{k}$  hence  $(2 - \frac{1}{k}) \geq (2 - \frac{1}{2}) = \frac{3}{2} \Rightarrow -\frac{3}{2} \geq (2 - \frac{1}{k})$  further as  $2 \leq k \Rightarrow 10 \leq 5 \cdot k \Rightarrow 0 < 7 \leq 5 \cdot k - 3$  hence  $0 \leq \frac{2}{3} + \frac{5 \cdot k - 3}{6 \cdot k^2}$  so  $-(2 - \frac{1}{k}) \cdot \left( \frac{2}{3} + \frac{5 \cdot k - 3}{6 \cdot k^2} \right) \leq -\frac{3}{2} \cdot \left( \frac{2}{3} + \frac{5 \cdot k - 3}{6 \cdot k^2} \right)$  from which we have

$$\begin{aligned}
1 - \left( 2 - \frac{1}{k} \right) \cdot \left( \frac{2}{3} + \frac{5 \cdot k - 3}{6 \cdot k^2} \right) &\leq 1 - \frac{3}{2} \cdot \left( \frac{2}{3} + \frac{5 \cdot k - 3}{6 \cdot k^2} \right) \\
&= 1 - 1 - \frac{3}{2} \cdot \frac{5 \cdot k - 3}{6 \cdot k^2} \\
&= \frac{5 \cdot k - 3}{4 \cdot k^2} \\
&< 0
\end{aligned}$$

So we have

$$1 - \binom{2 \cdot k}{2} \cdot \frac{1}{k^2} + \binom{2 \cdot k}{4} \cdot \frac{1}{k^4} < 0 \quad (12.8)$$

Further for  $j \in \{2 \cdot l + 1 \mid l \in \{1, \dots, n - 1\}\}$  we have  $1 \leq j \leq 2 \cdot (n - 1) + 1 = 2 \cdot n - 1 = k - 1 < k$  hence

$$\begin{aligned}
\binom{2 \cdot k}{2 \cdot j + 2} \cdot \frac{1}{k^{2 \cdot j + 2}} - \binom{2 \cdot k}{2 \cdot j} \cdot \frac{1}{k^{2 \cdot j}} &= \frac{(2 \cdot k)!}{(2 \cdot j + 2)! \cdot (2 \cdot k - 2 \cdot j - 2)! \cdot k^{2 \cdot j + 2}} - \\
&\quad \frac{(2 \cdot k)!}{(2 \cdot j)! \cdot (2 \cdot k - 2 \cdot j)! \cdot k^{2 \cdot j}} \\
&= \frac{(2 \cdot k)!}{k^{2 \cdot j}} \cdot \left( \frac{1}{(2 \cdot j + 2)! \cdot (2 \cdot k - 2 \cdot j - 2)! \cdot k^2} - \right. \\
&\quad \left. \frac{1}{(2 \cdot j)! \cdot (2 \cdot k - 2 \cdot j)!} \right)
\end{aligned}$$

$$\begin{aligned}
&= \frac{(2 \cdot k)!}{k^{2 \cdot j}} \cdot \left( \frac{1}{(2 \cdot j + 2) \cdot (2 \cdot j + 1) \cdot (2 \cdot j)! \cdot (2 \cdot k - 2 \cdot j - 2)! \cdot k^2} - \right. \\
&\quad \left. \frac{1}{(2 \cdot j)! \cdot (2 \cdot k - 2 \cdot j)!} \right) \\
&= \frac{(2 \cdot k)!}{(2 \cdot j)! \cdot k^{2 \cdot j}} \cdot \left( \frac{1}{(2 \cdot j + 2) \cdot (2 \cdot j + 1) \cdot (2 \cdot k - 2 \cdot j - 2)! \cdot k^2} - \right. \\
&\quad \left. \frac{1}{(2 \cdot k - 2 \cdot j)!} \right) \\
&= \frac{(2 \cdot k)!}{(2 \cdot j)! \cdot k^{2 \cdot j}} \cdot \left( \frac{1}{(2 \cdot j + 2) \cdot (2 \cdot j + 1) \cdot (2 \cdot k - 2 \cdot j - 2)! \cdot k^2} - \right. \\
&\quad \left. \frac{1}{(2 \cdot k - 2 \cdot j) \cdot (2 \cdot k - 2 \cdot j - 1) \cdot (2 \cdot k - 2 \cdot j - 2)!} \right) \\
&= \frac{(2 \cdot k)!}{(2 \cdot j)! \cdot (2 \cdot k - 2 \cdot j - 2)! \cdot k^{2 \cdot j}} \cdot \left( \frac{1}{(2 \cdot j + 2) \cdot (2 \cdot j + 1) \cdot k^2} - \right. \\
&\quad \left. \frac{1}{(2 \cdot k - 2 \cdot j) \cdot (2 \cdot k - 2 \cdot j - 1)} \right)
\end{aligned}$$

proving that

$$\left( \binom{2 \cdot k}{2 \cdot j + 2} \cdot \frac{1}{k^{2 \cdot j + 2}} - \binom{2 \cdot k}{2 \cdot j} \cdot \frac{1}{k^{2 \cdot j}} \right) = \frac{(2 \cdot k)!}{(2 \cdot j)! \cdot (2 \cdot k - 2 \cdot j - 2)! \cdot k^{2 \cdot j}} \cdot \left( \frac{1}{(2 \cdot j + 2) \cdot (2 \cdot j + 1) \cdot k^2} - \right. \\
\left. \frac{1}{(2 \cdot k - 2 \cdot j) \cdot (2 \cdot k - 2 \cdot j - 1)} \right) \quad (12.9)$$

Now as  $0 \leq j < k$  we have

$$0 < (2 \cdot j + 2) \cdot (2 \cdot j + 1) \cdot k^2, (2 \cdot k - 2 \cdot j) \cdot (2 \cdot k - 2 \cdot j - 1) \quad (12.10)$$

and

$$\begin{aligned}
&(2 \cdot j + 2) \cdot (2 \cdot j + 1) \cdot k^2 - (2 \cdot k - 2 \cdot j) \cdot (2 \cdot k - \\
&2 \cdot j - 1) = 2 \cdot [(j + 1) \cdot (2 \cdot j + 1) \cdot k^2 - (k - 1) \cdot (2 \cdot k - 2 \cdot j - 1)] \\
&= 2 \cdot [2 \cdot j^2 \cdot k^2 + j \cdot k^2 + 2 \cdot j \cdot k^2 + k^2 - [2 \cdot k^2 - \\
&2 \cdot k \cdot j - k - 2 \cdot k + 2 \cdot j + 1]] \\
&= 2 \cdot [2 \cdot j^2 \cdot k^2 + 3 \cdot j \cdot k^2 + k^2 - 2 \cdot k^2 + 2 \cdot k \cdot j + k + \\
&2 \cdot k - 2 \cdot j - 1] \\
&= 2 \cdot [2 \cdot j^2 \cdot k^2 + 3 \cdot j \cdot k^2 - k^2 + 2 \cdot k \cdot j + k + 2 \cdot k - \\
&2 \cdot j - 1] \\
&> 2 \cdot [2 \cdot j^2 \cdot k^2 + 3 \cdot k^2 - k^2 + 2 \cdot k \cdot j + k + 2 \cdot k - \\
&2 \cdot j - 1] \\
&= 2 \cdot [2 \cdot j^2 \cdot k^2 + 2 \cdot k^2 + 2 \cdot k \cdot j + (k - 1) + 2 \cdot (k - j)] \\
&> 0
\end{aligned}$$

hence using the above, 9.41 and 12.10 we have that  $\left( \frac{1}{(2 \cdot j + 2) \cdot (2 \cdot j + 1) \cdot k^2} - \right. \left. \frac{1}{(2 \cdot k - 2 \cdot j) \cdot (2 \cdot k - 2 \cdot j - 1)} \right) < 0$  so that by 12.9 we have

$$\left( \binom{2 \cdot k}{2 \cdot j + 2} \cdot \frac{1}{k^{2 \cdot j + 2}} - \binom{2 \cdot k}{2 \cdot j} \cdot \frac{1}{k^{2 \cdot j}} \right) < 0 \quad (12.11)$$

Combining then 12.8 and 12.11 on 12.6 proves that

$$\operatorname{Re}\left(\left(1+\frac{i}{k}\right)^{2 \cdot k}\right) < 0 \quad (12.12)$$

Now to extmiate the imaginair part note that

$$\begin{aligned} \binom{2 \cdot k}{2 \cdot j - 1} \cdot \frac{1}{k^{2 \cdot j - 1}} - \binom{2 \cdot k}{2 \cdot j + 1} \cdot \frac{1}{k^{2 \cdot j + 1}} &= \frac{1}{k^{2 \cdot j - 1}} \cdot \left( \binom{2 \cdot k}{2 \cdot j - 1} - \binom{2 \cdot k}{2 \cdot j + 1} \cdot \frac{1}{k^2} \right) \\ &= \frac{1}{k^{2 \cdot j - 1}} \cdot \left( \frac{(2 \cdot k)!}{(2 \cdot j - 1)! \cdot (2 \cdot k - 2 \cdot j + 1)!} - \frac{(2 \cdot k)!}{(2 \cdot j + 1)! \cdot (2 \cdot k - 2 \cdot j - 1)!} \cdot \frac{1}{k^2} \right) \\ &= \frac{(2 \cdot k)!}{k^{2 \cdot j - 1}} \cdot \left( \frac{1}{(2 \cdot j - 1)! \cdot (2 \cdot k - 2 \cdot j + 1)!} - \frac{1}{(2 \cdot j + 1)! \cdot (2 \cdot k - 2 \cdot j - 1)! \cdot k^2} \right) \\ &= \frac{(2 \cdot k)!}{k^{2 \cdot j - 1}} \cdot \left( \frac{1}{(2 \cdot j - 1)! \cdot (2 \cdot k - 2 \cdot j + 1)!} - \frac{1}{(2 \cdot j + 1) \cdot (2 \cdot j) \cdot (2 \cdot j - 1)! \cdot (2 \cdot k - 2 \cdot j - 1)! \cdot k^2} \right) \\ &= \frac{(2 \cdot k)!}{k^{2 \cdot j - 1} \cdot (2 \cdot j - 1)!} \cdot \left( \frac{1}{(2 \cdot k - 2 \cdot j + 1)!} - \frac{1}{(2 \cdot j + 1) \cdot (2 \cdot j) \cdot (2 \cdot k - 2 \cdot j - 1)! \cdot k^2} \right) \\ &= \frac{(2 \cdot k)!}{k^{2 \cdot j - 1} \cdot (2 \cdot j - 1)!} \cdot \left( \frac{1}{(2 \cdot k - 2 \cdot j + 1) \cdot (2 \cdot k - 2 \cdot j)!} - \frac{1}{(2 \cdot j + 1) \cdot (2 \cdot j) \cdot (2 \cdot k - 2 \cdot j - 1)! \cdot k^2} \right) \\ &= \frac{(2 \cdot k)!}{k^{2 \cdot j - 1} \cdot (2 \cdot j - 1)!} \cdot \left( \frac{1}{(2 \cdot k - 2 \cdot j + 1) \cdot (2 \cdot k - 2 \cdot j)!} - \frac{1}{(2 \cdot j + 1) \cdot (2 \cdot j) \cdot (2 \cdot k - 2 \cdot j - 1)!} \right) \end{aligned}$$

Proving that

$$\begin{aligned} \binom{2 \cdot k}{2 \cdot j - 1} \cdot \frac{1}{k^{2 \cdot j - 1}} - \binom{2 \cdot k}{2 \cdot j + 1} \cdot \frac{1}{k^{2 \cdot j + 1}} &= \left( \frac{1}{(2 \cdot k - 2 \cdot j + 1) \cdot (2 \cdot k - 2 \cdot j)} - \frac{1}{(2 \cdot j + 1) \cdot (2 \cdot j) \cdot k^2} \right) \end{aligned} \quad (12.13)$$

Now we have using  $1 < j < k$  that  $2 \cdot k - 2 \cdot j > 0$  hence

$$0 < (2 \cdot k - 2 \cdot j + 1) \cdot (2 \cdot k - 2 \cdot j), (2 \cdot j + 1) \cdot (2 \cdot j) \cdot k^2 \quad (12.14)$$

further taking in account that form  $0 < 1 \leq j$  we have by 9.41 that  $j^2 \geq 1$

$$\begin{aligned} (2 \cdot j + 1) \cdot (2 \cdot j) \cdot k^2 - (2 \cdot k - 2 \cdot j + 1) \cdot (2 \cdot k - 2 \cdot j) &= 2 \cdot [(2 \cdot j + 1) \cdot j \cdot k^2 - (2 \cdot k - 2 \cdot j + 1) \cdot (k - j)] \\ &= 2 \cdot [2 \cdot j^2 \cdot k^2 + j \cdot k^2 - [2 \cdot k^2 - 2 \cdot k \cdot j + k - 2 \cdot k \cdot j - 2 \cdot j^2 - j]] \end{aligned}$$

$$\begin{aligned}
&= 2 \cdot [2 \cdot j^2 \cdot k^2 + j \cdot k^2 - 2 \cdot k^2 + 2 \cdot k \cdot j - k + 2 \cdot k \cdot j + \\
&\quad 2 \cdot j^2 + j] \\
&= 2 \cdot [2 \cdot j^2 \cdot k^2 + j \cdot k^2 - 2 \cdot k^2 + 4 \cdot k \cdot j - k + 2 \cdot j^2 + \\
&\quad j] \\
&\geq 2 \cdot [2 \cdot k^2 + j \cdot k^2 - 2 \cdot k^2 + 4 \cdot k - k + 2 \cdot j^2 + j] \\
&= 2 \cdot [j \cdot k^2 \cdot k + 2 \cdot j^2 + j] \\
&> 0
\end{aligned}$$

Hence using the above together with 12.14 gives  $\frac{1}{(2 \cdot k - 2 \cdot j + 1) \cdot (2 \cdot k - 2 \cdot j)} - \frac{1}{(2 \cdot j + 1) \cdot (2 \cdot j) \cdot k^2} = \frac{(2 \cdot j + 1) \cdot (2 \cdot j) \cdot k^2 - (2 \cdot k - 2 \cdot j + 1) \cdot (2 \cdot k - 2 \cdot j)}{(2 \cdot k - 2 \cdot j + 1) \cdot (2 \cdot k - 2 \cdot j) \cdot (2 \cdot j + 1) \cdot (2 \cdot j) \cdot k^2} > 0$  which using 12.13 proves that  $\left(\frac{2 \cdot k}{2 \cdot j - 1}\right) \cdot \frac{1}{k^{2 \cdot j - 1}} - \left(\frac{2 \cdot k}{2 \cdot j + 1}\right) \cdot \frac{1}{k^{2 \cdot j + 1}} > 0$ . Substituting this in 12.7 gives

$$\text{Img}\left(\left(1 + \frac{i}{k}\right)^{2 \cdot k}\right) > 0 \quad (12.15)$$

The theorem is then proved by 12.12 and 12.15  $\square$

#### 12.1.4 Proof of the fundamental theorem of algebra

After all this work we are ready for the fundamental theorem of algebra.

**Theorem 12.29.** *Let  $n \in \mathbb{N}_0$ ,  $\{a_i\}_{i \in \{0, \dots, n\}} \subseteq \mathbb{C}$  with  $a_n \neq 0$  and  $p: \mathbb{C} \rightarrow \mathbb{C}$  defined by  $p(z) = \sum_{i=1}^n a_i \cdot z^i$  ( $p$  is a non constant polynomial) then there exists a  $z_0 \in \mathbb{C}$  such that  $p(z_0) = 0$  (in other words every non constant polynomial has a zero point)*

**Proof.** First define  $p \cdot \bar{p}: \mathbb{C} \rightarrow \mathbb{R}$  by  $(p \cdot \bar{p})(z) = p(z) \cdot \overline{p(z)}$  now given  $z \in \mathbb{C}$  we have

$$\begin{aligned}
(p \cdot \bar{p})(z) &= p(z) \cdot \overline{p(z)} \\
&\stackrel{12.25}{=} \left( \sum_{i \in \{0, \dots, n\}} a_i \cdot z^i \right) \cdot \overline{\left( \sum_{i \in \{0, \dots, n\}} a_i \cdot z^i \right)} \\
&= \left( \sum_{i \in \{0, \dots, n\}} a_i \cdot z^i \right) \cdot \left( \sum_{i \in \{0, \dots, n\}} \overline{a_i \cdot z^i} \right) \\
&= \left( \sum_{i \in \{0, \dots, n\}} a_i \cdot z^i \right) \cdot \left( \sum_{i \in \{0, \dots, n\}} \bar{a}_i \cdot \bar{z}^i \right) \\
&\stackrel{10.60}{=} \sum_{(i, j) \in \{0, \dots, n\}^2} a_i \cdot z^i \cdot \bar{a}_j \cdot \bar{z}^j \\
&= \sum_{(i, j) \in \{(k, l) \in \{0, \dots, n\}^2 \mid k=l\} \sqcup \{(k, l) \in \{0, \dots, n\}^2 \mid k < l\} \sqcup \{(k, l) \in \{0, \dots, n\}^2 \mid l < k\}} a_i \cdot z^i \cdot \bar{a}_j \cdot \bar{z}^j \\
&= \sum_{(i, j) \in \{(k, l) \in \{0, \dots, n\}^2 \mid k=l\}} a_i \cdot z^i \cdot \bar{a}_j \cdot \bar{z}^j + \sum_{(i, j) \in \{(k, l) \in \{0, \dots, n\}^2 \mid k < l\}} a_i \cdot z^i \cdot \bar{a}_j \cdot \bar{z}^j + \\
&\quad \sum_{(i, j) \in \{(k, l) \in \{0, \dots, n\}^2 \mid l < k\}} a_i \cdot z^i \cdot \bar{a}_j \cdot \bar{z}^j \\
&= \sum_{i \in \{0, \dots, n\}} a_i \cdot z^i \cdot \bar{a}_i \cdot \bar{z}^i + \sum_{(i, j) \in \{(k, l) \in \{0, \dots, n\}^2 \mid k < l\}} a_i \cdot z^i \cdot \bar{a}_j \cdot \bar{z}^j + \\
&\quad \sum_{(j, i) \in \{(k, l) \in \{0, \dots, n\}^2 \mid l < k\}} a_j \cdot z^j \cdot \bar{a}_i \cdot \bar{z}^i \\
&= \sum_{i \in \{0, \dots, n\}} a_i \cdot z^i \cdot \bar{a}_i \cdot \bar{z}^i + \sum_{(i, j) \in \{(k, l) \in \{0, \dots, n\}^2 \mid k < l\}} a_i \cdot z^i \cdot \bar{a}_j \cdot \bar{z}^j + \\
&\quad \sum_{(i, j) \in \{(l, k) \in \{0, \dots, n\}^2 \mid k < l\}} a_j \cdot z^j \cdot \bar{a}_i \cdot \bar{z}^i
\end{aligned}$$

$$\begin{aligned}
&= \sum_{i \in \{0, \dots, n\}} a_i \cdot z^i \cdot \bar{a}_i \cdot \bar{z}^i + \sum_{(i, j) \in \{(k, l) \in \{0, \dots, n\}^2 \mid k < l\}} (a_i \cdot z^i \cdot \bar{a}_j \cdot \bar{z}^j + a_j \cdot z^j \cdot \bar{a}_i \cdot \bar{z}^i) \\
&= \sum_{i \in \{0, \dots, n\}} a_i \cdot z^i \cdot \bar{a}_i \cdot \bar{z}^i + \sum_{(i, j) \in \{(k, l) \in \{0, \dots, n\}^2 \mid k < l\}} (a_i \cdot z^i \cdot \bar{a}_j \cdot \bar{z}^j + a_j \cdot z^j \cdot \bar{a}_i \cdot \bar{z}^i) \\
&= \sum_{i \in \{0, \dots, n\}} a_i \cdot z^i \cdot \bar{a}_i \cdot \bar{z}^i + \sum_{(i, j) \in \{(k, l) \in \{0, \dots, n\}^2 \mid k < l\}} (a_i \cdot z^i \cdot \bar{a}_j \cdot \bar{z}^j + \overline{a_i \cdot z^i \cdot \bar{a}_j \cdot z^j}) \\
&= \sum_{i \in \{0, \dots, n\}} a_i \cdot z^i \cdot \bar{a}_i \cdot \bar{z}^i + 2 \cdot \sum_{(i, j) \in \{(k, l) \in \{0, \dots, n\}^2 \mid k < l\}} \operatorname{Re}(a_i \cdot z^i \cdot \bar{a}_j \cdot \bar{z}^j) \\
&= \sum_{i \in \{0, \dots, n\}} |a_i|^2 \cdot |z^i|^2 + 2 \cdot \sum_{(i, j) \in \{(k, l) \in \{0, \dots, n\}^2 \mid k < l\}} \operatorname{Re}(a_i \cdot z^i \cdot \bar{a}_j \cdot \bar{z}^j)
\end{aligned}$$

So we have that

$$(p \cdot \bar{p})(z) = \sum_{i \in \{0, \dots, n\}} |a_i|^2 \cdot |z^i|^2 + 2 \cdot \sum_{(i, j) \in \{(k, l) \in \{0, \dots, n\}^2 \mid k < l\}} \operatorname{Re}(a_i \cdot z^i \cdot \bar{a}_j \cdot \bar{z}^j) \quad (12.16)$$

Using the above we have

$$\begin{aligned}
(p \cdot \bar{p})(z) &= p(z) \cdot \overline{p(z)} \\
&\stackrel{0 \leq p(z) \cdot \overline{p(z)} \in \mathbb{R}}{=} |p(z) \cdot \overline{p(z)}| \\
&\stackrel{p(z) \cdot \overline{p(z)} \in \mathbb{R} \text{ and 12.15}}{=} \|p(z) \cdot \overline{p(z)}\| \\
&= \left\| \sum_{i \in \{0, \dots, n\}} |a_i|^i \cdot |z^i|^2 + 2 \cdot \sum_{(i, j) \in \{(k, l) \in \{0, \dots, n\}^2 \mid k < l\}} \operatorname{Re}(a_i \cdot z^i \cdot \bar{a}_j \cdot \bar{z}^j) \right\| \\
&\stackrel{\geq 11.70}{\geq} \left\| \sum_{i \in \{0, \dots, n\}} |a_i|^{2 \cdot i} \cdot |z^i|^2 \right\| - \left\| 2 \cdot \sum_{(i, j) \in \{(k, l) \in \{0, \dots, n\}^2 \mid k < l\}} \operatorname{Re}(a_i \cdot z^i \cdot \bar{a}_j \cdot \bar{z}^j) \right\| \\
&\geq \left\| \sum_{i \in \{0, \dots, n\}} |a_i|^{2 \cdot i} \cdot |z^i|^2 \right\| - 2 \cdot \left\| \sum_{(i, j) \in \{(k, l) \in \{0, \dots, n\}^2 \mid k < l\}} \operatorname{Re}(a_i \cdot z^i \cdot \bar{a}_j \cdot \bar{z}^j) \right\| \\
&= \sum_{i \in \{0, \dots, n\}} |a_i|^{2 \cdot i} \cdot |z^i|^2 - 2 \cdot \left\| \sum_{(i, j) \in \{(k, l) \in \{0, \dots, n\}^2 \mid k < l\}} \operatorname{Re}(a_i \cdot z^i \cdot \bar{a}_j \cdot \bar{z}^j) \right\| \\
&\quad \text{as } [0 \leq \sum_{i \in \{0, \dots, n\}} |a_i|^i \cdot |z|^i \in \mathbb{R} \text{ and 12.15 (1)}] \\
&\geq |a_n|^{2 \cdot n} \cdot |z^n|^2 - 2 \cdot \left\| \sum_{(i, j) \in \{(k, l) \in \{0, \dots, n\}^2 \mid k < l\}} \operatorname{Re}(a_i \cdot z^i \cdot \bar{a}_j \cdot \bar{z}^j) \right\| \\
&\stackrel{\geq 12.15 (3)}{\geq} |a_n|^{2 \cdot n} \cdot |z^n|^2 - 2 \cdot \sum_{(i, j) \in \{(k, l) \in \{0, \dots, n\}^2 \mid k < l\}} \|a_i \cdot \bar{a}_j\| \cdot \|z^i\| \cdot \|z^j\| \\
&\stackrel{\geq 12.15 (5)}{\geq} |a_n|^{2 \cdot n} \cdot |z^n|^2 - 2 \cdot \sum_{(i, j) \in \{(k, l) \in \{0, \dots, n\}^2 \mid k < l\}} \|a_i \cdot \bar{a}_j\| \cdot \|z\|^i \cdot \|z\|^j \\
&\stackrel{0 \leq |z^n|^2 \text{ and } \in \mathbb{R} 12.15}{=} |a_n|^{2 \cdot n} \cdot \||z^n|^2\| - 2 \cdot \sum_{(i, j) \in \{(k, l) \in \{0, \dots, n\}^2 \mid k < l\}} \|a_i \cdot \bar{a}_j\| \cdot \|z\|^{i+j} \\
&= |a_n|^{2 \cdot n} \cdot \|z^n \cdot \bar{z}^n\| - 2 \cdot \sum_{(i, j) \in \{(k, l) \in \{0, \dots, n\}^2 \mid k < l\}} \|a_i \cdot \bar{a}_j\| \cdot \|z\|^{i+j} \\
&\stackrel{\geq 12.15 (3)}{\geq} |a_n|^{2 \cdot n} \frac{\|z^n\| \cdot \|\bar{z}^n\|}{2} - 2 \cdot \sum_{(i, j) \in \{(k, l) \in \{0, \dots, n\}^2 \mid k < l\}} \|a_i \cdot \bar{a}_j\| \cdot \|z\|^{i+j}
\end{aligned}$$

$$\begin{aligned}
&\stackrel{12.15(2)}{=} |a_n|^{2 \cdot n} \frac{\|z^n\|^2}{2} - 2 \cdot \sum_{(i,j) \in \{(k,l) \in \{0, \dots, n\}^2 \mid k < l\}} \|a_i \cdot \bar{a_j}\| \cdot \|z\|^{i+j} \\
&\stackrel{12.15(7)}{\geq} |a_n|^{2 \cdot n} \frac{1}{2} \cdot \left( \frac{\|z\|^n}{2^n} \right)^2 - 2 \cdot \sum_{(i,j) \in \{(k,l) \in \{0, \dots, n\}^2 \mid k < l\}} \|a_i \cdot \bar{a_j}\| \cdot \|z\|^{i+j} \\
&= |a_n|^{2 \cdot n} \frac{\|z\|^{2 \cdot n}}{2^{2 \cdot n+1}} - 2 \cdot \sum_{(i,j) \in \{(k,l) \in \{0, \dots, n\}^2 \mid k < l\}} \|a_i \cdot \bar{a_j}\| \cdot \|z\|^{i+j}
\end{aligned}$$

proving that

$$(p \cdot \bar{p})(z) \geq |a_n|^{2 \cdot n} \frac{\|z\|^{2 \cdot n}}{2^{2 \cdot n+1}} - 2 \cdot \sum_{(i,j) \in \{(k,l) \in \{0, \dots, n\}^2 \mid k < l\}} \|a_i \cdot \bar{a_j}\| \cdot \|z\|^{i+j} \quad (12.17)$$

Further by 12.27 that  $A = \{(k, l) \in \{0, \dots, n\}^2 \mid k < l\} = \bigsqcup_{m \in \{1, \dots, 2 \cdot n-1\}} B_m$  where  $B_m = \{(k, l) \in A \mid k + l = m\}$  hence

$$\begin{aligned}
\sum_{(i,j) \in \{(k,l) \in \{0, \dots, n\}^2 \mid k < l\}} \|a_i \cdot \bar{a_j}\| \cdot \|z\|^{i+j} &= \sum_{(i,j) \in \bigsqcup_{m \in \{1, \dots, 2 \cdot n-1\}} B_m} \|a_i \cdot \bar{a_j}\| \cdot \|z\|^{i+j} \\
&\stackrel{10.46}{=} \sum_{m \in \{1, \dots, 2 \cdot n-1\}} \left( \sum_{(i,j) \in B_m} \|a_i \cdot \bar{a_j}\| \cdot \|z\|^{i+j} \right) \\
&= \sum_{m \in \{1, \dots, 2 \cdot n-1\}} \left( \sum_{(i,j) \in B_m} \|a_i \cdot \bar{a_j}\| \cdot \|z\|^m \right) \\
&= \sum_{m=1}^{2 \cdot n-1} \left( \sum_{(i,j) \in B_m} \|a_i \cdot \bar{a_j}\| \right) \|z\|^m \\
&= \sum_{m=1}^{2 \cdot n-1} b_m \cdot \|z\|^m
\end{aligned}$$

where  $b_m = \sum_{(i,j) \in B_m} \|a_i \cdot \bar{a_j}\|$ . Using the above and 12.17 gives then that  $(p \cdot \bar{p})(z) \geq |a_n|^{2 \cdot n} \frac{\|z\|^{2 \cdot n}}{2^{2 \cdot n+1}} - 2 \cdot \sum_{m=1}^{2 \cdot n-1} b_m \|z\|^m$  so if we define  $\{c_i\}_{i \in \{0, \dots, n\}}$  by  $c_i = \begin{cases} -2 \cdot b_m & \text{if } i \in \{1, \dots, 2 \cdot n-1\} \\ \frac{|a_n|^{2 \cdot n}}{2^{2 \cdot n+1}} & \text{if } i = 2 \cdot n \end{cases}$  then

$$c_n > 0 \wedge (p \cdot \bar{p})(z) \geq \sum_{m \in \{1, \dots, 2 \cdot n\}} c_i \cdot \|z\|^i \quad (12.18)$$

As  $c_n > 0$  we have by 12.13 that  $\lim_{x \rightarrow \infty} (\sum_{i=1}^{2 \cdot n} c_i \cdot x^i) = \infty$  hence there exists a  $R' \in \mathbb{R}$  such that if  $x \geq R'$  then  $\sum_{i=1}^{2 \cdot n} c_i \cdot x^i \geq (p \cdot \bar{p})(0) + 1$  so if  $x \geq \max(R', 1) > 0 \Rightarrow x \geq R'$  we have  $\sum_{i=1}^{2 \cdot n} c_i \cdot x^i > (p \cdot \bar{p})(0)$ . Hence by taking  $R = \max(R', 1)$  we have proved that

$$\exists R > 0 \text{ so that if } \|z\| \geq R \text{ then } (p \cdot \bar{p})(z) > (p \cdot \bar{p})(0) \quad (12.19)$$

Now as  $\overline{B_{|||}}(0, R) = \{x \in \mathbb{C} \mid \|x\| \leq R\}$  is a closed and bounded set in  $\langle \mathbb{C}, \|\cdot\| \rangle$  (see 11.56 and 11.66) we have by 11.305 that  $\overline{B_{|||}}(0, R)$  is compact. By the extreme value theorem (see 11.257) and the fact that  $p \cdot \bar{p}$  is continuous (see 12.16 (5) we have that there exists a  $z_0 \in \overline{B_{|||}}(0, R)$  such that  $(p \cdot \bar{p})(z) \geq (p \cdot \bar{p})(z_0)$  for every  $z \in \overline{B_{|||}}(0, R)$ . As  $0 \in \overline{B_{|||}}(0, R)$  we have also  $(p \cdot \bar{p})(0) \geq (p \cdot \bar{p})(z_0)$ . Hence if  $z \in \mathbb{C}$  then we have either  $z \leq \mathbb{R}$  and thus by the above we have  $(p \cdot \bar{p})(z) \geq (p \cdot \bar{p})(z_0)$  or  $R < z$  and then by 12.19 we have  $(p \cdot \bar{p})(z) \geq (p \cdot \bar{p})(0) \geq (p \cdot \bar{p})(z_0)$ , so we reach the conclusion that

$$\forall z \in \mathbb{C} \text{ we have } (p \cdot \bar{p})(z) \geq (p \cdot \bar{p})(z_0) \quad (12.20)$$

If we define now

$$r: \mathbb{C} \rightarrow \mathbb{C} \text{ by } r(z) = p(z + z_0) \quad (12.21)$$

then  $r(0) = p(0 + z_0) = p(z_0)$  and  $(r \cdot \bar{r})(z) = r(z) \cdot \bar{r}(z) = p(z + z_0) \cdot \bar{p}(z + z_0) \geq (p(z_0) \cdot \bar{p}(z_0)) = r(0) \cdot \bar{r}(0) = (r \cdot \bar{r})(0)$  so using 12.20 we have that

$$\forall z \in \mathbb{C} \text{ we have } r(z) \cdot \bar{r}(z) - r(0) \cdot \bar{r}(0) \geq 0 \quad (12.22)$$

By 12.23 we have that  $r$  is a non constant polynomial of order  $n$  hence using 12.24 we have

$$\exists k \in \{1, \dots, n\} \text{ and a polynomial } q \text{ of order } n-k \text{ such that } q(0) \neq 0 \wedge r(z) = z^k \cdot q(z) + r(0) \quad (12.23)$$

Then we have that

$$\begin{aligned} r(z) \cdot \overline{r(z)} - r(0) \cdot \overline{r(0)} &= (z^k \cdot q(z) + r(0)) \cdot \overline{z^k \cdot q(z) + r(0)} - r(0) \cdot \overline{r(0)} \\ &= (z^k \cdot q(z) + r(0)) \cdot (z^k \cdot \overline{q(z)} + \overline{r(0)}) - r(0) \cdot \overline{r(0)} \\ &= z^k \cdot q(z) \cdot \overline{z^k \cdot q(z)} + z^k \cdot q(z) \cdot \overline{r(0)} + r(0) \cdot \overline{z^k \cdot q(z)} + r(0) \cdot \overline{r(0)} - r(0) \cdot \overline{r(0)} \\ &= z^k \cdot q(z) \cdot \overline{z^k \cdot q(z)} + z^k \cdot q(z) \cdot \overline{r(0)} + z^k \cdot q(z) \cdot \overline{r(0)} \\ &= z^k \cdot q(z) \cdot \overline{z^k \cdot q(z)} + 2 \cdot \operatorname{Re}(z^k \cdot q(z) \cdot \overline{r(0)}) \end{aligned}$$

Using 12.22 we have then that

$$z^k \cdot q(z) \cdot \overline{z^k \cdot q(z)} + 2 \cdot \operatorname{Re}(z^k \cdot q(z) \cdot \overline{r(0)}) \geq 0 \quad (12.24)$$

If now  $x \in \mathbb{C}$  and  $\delta \in \mathbb{R}$  with  $0 < \delta$  then substituting  $\delta \cdot x$  for  $z$  in the above gives

$$\begin{aligned} 0 &\leq (\delta \cdot x)^k \cdot q(\delta \cdot x) \cdot \overline{(\delta \cdot x)^k \cdot q(\delta \cdot x)} + 2 \cdot \operatorname{Re}((\delta \cdot x)^k \cdot q(\delta \cdot x) \cdot \overline{r(0)}) \\ &= \delta^k \cdot x^k \cdot q(\delta \cdot x) \cdot \delta^k \cdot \overline{x^k \cdot q(\delta \cdot x)} + \delta^k \cdot 2 \cdot \operatorname{Re}(x^k \cdot q(\delta \cdot x) \cdot \overline{r(0)}) \\ &= \delta^k \cdot [x^k \cdot q(\delta \cdot x) \cdot \delta^k \cdot \overline{x^k \cdot q(\delta \cdot x)} + 2 \cdot \operatorname{Re}(x^k \cdot q(\delta \cdot x) \cdot \overline{r(0)})] \\ &= \delta^k \cdot [\delta^k \cdot |x^k \cdot q(\delta \cdot x)|^2 + 2 \cdot \operatorname{Re}(x^k \cdot q(\delta \cdot x) \cdot \overline{r(0)})] \end{aligned}$$

hence dividing by  $0 < \delta^k$  we have then

$$\forall r > 0, \forall x \in \mathbb{C} \text{ that } 0 \leq \delta^k \cdot |x^k \cdot q(\delta \cdot x)|^2 + 2 \cdot \operatorname{Re}(x^k \cdot q(\delta \cdot x) \cdot \overline{r(0)}) \quad (12.25)$$

If we consider the function  $f_1: \mathbb{R} \rightarrow \mathbb{R}$  defined by  $f_1(\delta) = \operatorname{Re}(x^k \cdot q(\delta \cdot x) \cdot \overline{r(0)})$  then using the definitions in 12.16 we have that  $f_1(\delta) = \operatorname{Re}(x^k \cdot q(\delta \cdot x) \cdot \overline{r(0)}) = (\operatorname{Re} \circ (\cdot(x^k \cdot r(0))) \circ q_x)(\delta)$  hence  $f_1$  is continuous as the product of continuous functions (see 12.16). Further the function  $f_2: \mathbb{R} \rightarrow \mathbb{R}$  defined by  $f_2(\delta) = |x^k \cdot q(\delta \cdot x)|^2 = (\diamond \circ (\cdot(x^k)) \circ q_z)(\delta)$  is continuous as the composition of continuous functions (see 12.16). Hence  $f_3: \mathbb{R} \rightarrow \mathbb{R}$  defined by  $f_3(\delta) = \delta^k \cdot |x^k \cdot q(\delta \cdot x)|^2 = \delta^k \cdot f_2(\delta)$  is continuous as the product of continuous function and  $f_4: \mathbb{R} \rightarrow \mathbb{R}$  defined by  $f_4(\delta) = 2 \cdot \operatorname{Re}(x^k \cdot q(\delta \cdot x) \cdot \overline{r(0)})$  is continuous as the product of a constant (thus continuous function) and a continuous function. So finally  $f: \mathbb{R} \rightarrow \mathbb{R}$  defined by  $f(\delta) = \delta^k \cdot |x^k \cdot q(\delta \cdot x)|^2 + 2 \cdot \operatorname{Re}(x^k \cdot q(\delta \cdot x) \cdot \overline{r(0)}) = f_3(\delta) + f_4(\delta)$  is continuous as the sum of continuous functions. To summarize

$$\forall x \in \mathbb{C} \models f: \mathbb{R} \rightarrow \mathbb{R} \text{ defined by } f(\delta) = \delta^k \cdot |x^k \cdot q(\delta \cdot x)|^2 + 2 \cdot \operatorname{Re}(x^k \cdot q(\delta \cdot x) \cdot \overline{r(0)}) \text{ is continuous} \quad (12.26)$$

Using 12.25 we have  $\forall \delta > 0$  that  $0 \leq f(\delta)$ . Assume now that  $f(0) < 0$  then if we take  $\varepsilon = -f(0)$  there exists by continuity of  $f$  a  $\zeta > 0$  such that for  $0 < \frac{\zeta}{2} < \zeta$  we have  $|f\left(\frac{\zeta}{2}\right) - f(0)| < -f(0)$  and as  $f(0) < 0 \leq f\left(\frac{\zeta}{2}\right)$  we have  $f\left(\frac{\zeta}{2}\right) - f(0) = |f\left(\frac{\zeta}{2}\right) - f(0)| < -f(0)$  proving that  $0 \leq f\left(\frac{\zeta}{2}\right) < 0$  a contradiction. So we must have that  $0 \leq f(0)$  giving using the definition of  $f$  that

$$\forall x \in \mathbb{C} \text{ we have } 0 \leq \operatorname{Re}(x^k \cdot q(0) \cdot \overline{r(0)}) = \operatorname{Re}(x^k \cdot (a + i \cdot b)) \text{ where } q(0) \cdot \overline{r(0)} = a + i \cdot b \quad (12.27)$$

Then for  $k \in \{1, \dots, n\}$  we have the following cases

**$k$  is odd.** then if  $x = 1$  we have  $0 \leq \operatorname{Re}(1^k \cdot (a + i \cdot b)) = \operatorname{Re}(a + i \cdot b) = a$ , if  $x = -1$  we have  $0 \leq \operatorname{Re}((-1)^k \cdot (a + i \cdot b)) = \operatorname{Re}(-a - i \cdot b) = -a$  from which it follows that

$$a = 0 \quad (12.28)$$

As  $k$  is odd then we have that  $k = 2 \cdot m + 1$  and for  $m$  we have again two cases

**$m = \text{even}$ .** then  $m = 2 \cdot l \Rightarrow k = 4 \cdot l + 1$  hence if  $x = i$  we have by 9.22 that  $x^k = i^k = i$  and thus  $0 \leq \operatorname{Re}(i \cdot a - b) = -b$  and if  $x = -i$  then  $x^k = (-i)^k = (-1)^k \cdot i^k = -i$  and thus  $0 \leq \operatorname{Re}(-i \cdot a + b) = b$  proving that  $b = 0$

**$m = \text{odd}$ .** then  $m = 2 \cdot l + 1 \Rightarrow k = 4 \cdot l + 3$  hence if  $x = i$  we have by 9.22 that  $x^k = i^k = -i$  and thus  $0 \leq \operatorname{Re}(-i \cdot a + b) = b$  and if  $x = -i$  then  $x^k = (-i)^k = (-1)^k \cdot i^k = (-1) \cdot (-i) = i$  and thus  $0 \leq \operatorname{Re}(i \cdot a - b) = -b$  proving that  $b = 0$

So in all cases we have that  $b = 0$  which together with 12.28 we have that  $q(0) \cdot \overline{r(0)} = 0$  we have as  $q(0) \neq 0$  (see 12.23) that  $\overline{r(0)} = 0$  hence we have

$$r(0) = 0$$

**$k$  is even.** As  $k$  is even we have  $k = 2 \cdot m$  and then we have two cases for  $m$  to consider

**$m$  is odd.** then  $m = 2 \cdot l + 1$  and  $k = 4 \cdot l + 2$ , using 9.22 we have then that  $i^k = -1$  so if  $x = i \Rightarrow x^k = -1$  we have  $0 \leq \operatorname{Re}(-a - i \cdot b) = -a$ , if  $x = 1$  then  $x^k = 1$  so that  $0 \leq \operatorname{Re}(a + i \cdot b) = a$  which proves that  $a = 0$ . Hence if we take  $x = (1 \pm i)$  we have  $0 \leq \operatorname{Re}(x^k \cdot (a + i \cdot b)) = \operatorname{Re}((1 \pm i)^{4 \cdot l + 2} \cdot i \cdot b) \stackrel{12.17}{=} \operatorname{Re}(\pm 2 \cdot (-4)^l \cdot i \cdot b) = \operatorname{Re}(\mp 2 \cdot (-4)^l \cdot b) = \mp 2 \cdot (-4)^l \cdot b$  proving  $0 \leq b \vee 0 \leq -b$  hence  $b = 0$ . So we have  $q(0) \cdot \overline{r(0)} = 0 \stackrel{q(0) \neq 0}{\Rightarrow} \overline{r(0)} = 0$  and thus

$$r(0) = 0$$

**$m$  is even.** then  $m = 2 \cdot l$  and  $k = 4 \cdot l$  if  $x = 1$  we have  $0 \leq \operatorname{Re}(1^k \cdot (a + i \cdot b)) = \operatorname{Re}(a + i \cdot b) = a$  proving that

$$0 \leq a \tag{12.29}$$

Use now  $x = \left(1 + \frac{i}{2 \cdot l}\right)$  we have then that  $x^k = \left(1 + \frac{i}{2 \cdot l}\right)^{4 \cdot l} = \left(\left(1 + \frac{i}{2 \cdot l}\right)^2\right)^k$  then using 12.28 we have that  $\operatorname{Re}(x^k) < 0 < \operatorname{Im}(x^k)$ . As  $x^k \in \mathbb{C}$  we have a  $c, d \in \mathbb{R}$  with  $x^k = c + i \cdot d$  and  $c = \operatorname{Re}(x^k) < 0 < \operatorname{Im}(x^k) = d$ . Hence using 12.27 we have  $0 \leq \operatorname{Re}((c + i \cdot d) \cdot (a + i \cdot b)) = \operatorname{Re}((c \cdot a - d \cdot b) + i \cdot (c \cdot b + d \cdot a)) = c \cdot a - d \cdot b$  proving

$$0 \leq a \cdot c - b \cdot d \tag{12.30}$$

If we take now  $y = \bar{x}$  so that  $y^k = \bar{x}^k \stackrel{9.9}{=} \bar{x^k} = c - i \cdot d$  by 12.27 we have then that  $0 \leq \operatorname{Re}(y^k \cdot (a + i \cdot b)) = \operatorname{Re}((c - i \cdot d) \cdot (a + i \cdot b)) = \operatorname{Re}((c \cdot a + d \cdot b) + i \cdot (c \cdot b - d \cdot a)) = c \cdot a + d \cdot b$  proving

$$0 \leq a \cdot c + b \cdot d \tag{12.31}$$

Summing 12.30 and 12.31 gives that  $0 \leq a \cdot c - b \cdot d + a \cdot c + b \cdot d = 2 \cdot a \cdot c \Rightarrow 0 \leq a \cdot c$  and as  $c < 0$  we have  $a \leq 0$  which by 12.29 proves that  $a = 0$ . Substituting this in 12.30 and 12.31 proves that  $0 \leq -b \cdot d \Rightarrow b \cdot d \leq 0$  and  $0 \leq b \cdot d$  hence  $b \cdot d = 0$  and as  $0 < d$  we have  $b = 0$ . So we conclude that  $q(0) \cdot \overline{r(0)} = 0 \stackrel{q(0) \neq 0}{\Rightarrow} \overline{r(0)} = 0$  and thus

$$r(0) = 0$$

hence for all cases of  $k$  we have that  $r(0) = 0$ . As by definition (see 12.21)  $0 = r(0) = p(0 + z_0) = p(z_0)$  we have proved finally that  $p(z_0) = 0$  and we have find a zero of the polynomial  $p$  proving the fundamental theorem of algebra.  $\square$

**Theorem 12.30.** Let  $p$  be a polynomial with real coefficients and  $z \in \mathbb{C}$  is such that  $p(z) = 0$  then  $p(\bar{z}) = 0$

**Proof.** If  $p$  is defined by the coefficients  $\{a_i\}_{i \in \{0, \dots, n\}} \subseteq \mathbb{R}$  then we have as  $p(z) = 0$  that

$$\begin{aligned} 0 &= \bar{0} \\ &= \overline{p(z)} \\ &= \overline{\sum_{i=0}^n a_i \cdot z^i} \\ &\stackrel{12.25}{=} \sum_{i=0}^n \overline{a_i \cdot z^i} \end{aligned}$$

$$\begin{aligned}
&= \sum_{i=0}^n \bar{a}_i \cdot \bar{z}^i \\
&\stackrel{9.9}{=} \sum_{i=0}^n \bar{a}_i \cdot \bar{z}^i \\
&= \sum_{i=0}^n a_i \cdot \bar{z}^i \\
&= p(\bar{z})
\end{aligned}$$

proving the theorem.  $\square$

**Lemma 12.31.** *Let  $p$  be a polynomial then  $p(x_0) = 0$  if and only if there exists a polynomial  $q$  such that  $\forall x \in \mathbb{C}$  we have  $p(x) = (x - x_0) \cdot q(x)$  and  $\text{ord}(q) = \text{ord}(p) - 1$*

**Proof.**

$\Rightarrow$ . Suppose  $p(x_0) = 0$  then we have to consider the following cases

**ord( $p$ ) = 0.** then  $p(x) = \sum_{i=0}^0 a_0 \cdot x^i = a_0$  hence as  $a_0 = p(x_0) = 0$  we have  $a_0 = 0$  so if we define  $q$  by  $q = C_0$  we have  $p(x) = a_0 = 0 = (x - x_0) \cdot p(x) = (x - x_0) \cdot 0$

**ord( $p$ ) = 1.** then  $p(x) = \sum_{i=0}^1 a_i \cdot x^i = a_1 \cdot x + a_0$  with  $a_1 \neq 0$ , hence as  $0 = p(x_0) = a_1 \cdot x_0 + a_0$  we have  $a_0 = -a_1 \cdot x_0$  hence  $p(x) = a_1 \cdot x - a_1 \cdot x_0 = (x - x_0) \cdot a_1 = (x - x_0) \cdot q(x)$  where  $q(x) = a_1 \neq 0$  so that  $\text{ord}(q) = 0$

**ord( $p$ ) > 1.** then if  $n = \text{ord}(p)$  we have  $p(x) = \sum_{i=0}^n a_i \cdot x^i$  and  $a_n \neq 0$  so that  $0 = p(x_0) = \sum_{i=0}^n a_i \cdot x^i = \sum_{i=1}^n a_i \cdot (x_0)^i + a_0 \Rightarrow a_0 = -\sum_{i=1}^n a_i \cdot (x_0)^i$  so that

$$\begin{aligned}
p(x) &= \sum_{i=0}^n a_i \cdot x^i \\
&= \sum_{i=1}^n a_i \cdot x^i + a_0 \\
&= \sum_{i=1}^n a_i \cdot x^i + \sum_{i=1}^n a_i \cdot (x_0)^i \\
&= \sum_{i=1}^n a_i \cdot (x^i - (x_0)^i) \\
&= (x - x_0) \cdot \sum_{i=1}^n a_i \cdot (x^i - (x_0)^{i-1}) \\
&= (x - x_0) \cdot \sum_{i=0}^{n-1} a_{i+1} \cdot (x^i - (x_0)^i) \\
&= (x - x_0) \cdot \left( \sum_{i=1}^{n-1} a_{i+1} \cdot x^i + \left( a_1 - \sum_{i=0}^{n-1} a_{i+1} \cdot (x_0)^i \right) \right) \\
&= (x - x_0) \cdot q(x)
\end{aligned}$$

where  $q = \sum_{i=0}^{n-1} b_i \cdot x^i$  and  $b_i = \begin{cases} (a_1 - \sum_{j=0}^{n-1} a_{j+1} \cdot (x_0)^j) & \text{if } i = 0 \\ a_{i+1} & \text{if } i \in \{1, \dots, n-1\} \end{cases}$  as  $b_{n-1} = a_n \neq 0$  we have also  $\text{ord}(q) = n - 1 = \text{ord}(p) - 1$

$\Leftarrow$ . If  $p(x) = (x - x_0) \cdot q(x)$  then  $p(x_0) = (x_0 - x_0) \cdot q(x_0) = 0 \cdot q(x_0) = 0$   $\square$

**Theorem 12.32.** *Let  $p$  be a non constant polynomial then there exists a  $0 \neq c \in \mathbb{C}$ ,  $\{x_i\}_{i \in \{1, \dots, \text{ord}(p)\}}$  such that  $\forall x \in \mathbb{C}$  we have  $p(x) = c \cdot \prod_{j=1}^{\text{ord}(p)} (x - x_i)$ . Furthermore we have that  $\{z \in \mathbb{C} | p(z) = 0\} = \{x_i | i \in \{1, \dots, \text{ord}(p)\}\}$*

**Proof.** We prove this by induction on the order of the polynomial let  $\mathcal{S} = \{n \in \mathbb{N}_0 \mid \text{if } p \text{ is a non constant polynomial with } \text{ord}(p) = n \text{ then } \exists c, \{x_i\}_{i \in \{1, \dots, \text{ord}(p)\}} \text{ such that } p(x) = c \cdot \prod_{j=1}^{\text{ord}(p)} (x - x_j)\}$ , then we have

**$n = 1$ .** If  $p$  is a non constant polynomial with  $\text{ord}(p) \leq 1$  then  $p(x) = a_1 \cdot x + a_0$  where  $a_1 \neq 0$  hence  $a_1 \cdot \left(x - \left(-\frac{a_0}{a_1}\right)\right) = a_1 \cdot x + a_0$  proving that  $1 \in \mathcal{S}$

**$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$ .** If  $p$  is a polynomial of order  $n+1$  then using the fundamental theorem of algebra (see 12.29) there exists a  $x_{n+1} \in \mathbb{C}$  such that  $p(x_1) = 0$ , applying then the previous lemma (see 12.31) there exists a  $q$  with  $\text{ord}(q) = n$  such that  $p(x) = (x - x_{n+1}) \cdot q(x)$ . As  $n \in \mathcal{S}$  we have that  $\exists c \in \mathbb{C}, c \neq 0, \{x_i\}_{i \in \{1, \dots, n\}} \subseteq \mathbb{C}$  such that  $q(x) = c \cdot \prod_{i=1}^n (x - x_i)$  hence  $p(x) = c \cdot (x - x_{n+1}) \cdot \prod_{i=1}^n (x - x_i) = c \cdot \prod_{i=1}^{n+1} (x - x_i)$  proving that  $n+1 \in \mathcal{S}$ .

Further we have if  $x \in \{x_i \mid i \in \{1, \dots, \text{ord}(p)\}\}$  there exists a  $i \in \{1, \dots, \text{ord}(P)\}$  such that  $x = x_i$  hence  $p(x) = p(x_i) = c \cdot \prod_{j \in \{1, \dots, \text{ord}(p)\}} (x_i - x_j) = c \cdot (x_i - x_i) \cdot \prod_{j \in \{1, \dots, \text{ord}(p)\} \setminus \{i\}} (x_i - x_j) = 0$  proving that  $x \in \{z \mid p(z) = 0\}$ . If  $x \in \{z \mid p(z) = 0\}$  then  $p(x) = 0$  hence  $0 = c \cdot \prod_{i=1}^{\text{ord}(p)} (x - x_i) \underset{c \neq 0}{\Rightarrow} \prod_{i=1}^{\text{ord}(p)} (x - x_i) = 0$ , if now for  $\forall i \in \{1, \dots, \text{ord}(p)\}$  we have that  $x \neq x_i$  we have by 10.55 that  $\prod_{i=1}^{\text{ord}(p)} (x - x_i) \neq 0$  a contradiction so there exists a  $i \in \{1, \dots, \text{ord}(p)\}$  such that  $x = x_i$  and thus  $x \in \{x_i \mid i \in \{1, \dots, \text{ord}(p)\}\}$ .  $\square$

**Theorem 12.33.** Let  $p \in \mathcal{P}$  be a non constant polynomial then there exists a  $m, M \in \mathbb{N}$ ,  $c \in \mathbb{R}$   $\{\lambda_i\}_{i \in \{1, \dots, m\}} \subseteq \mathbb{R}$ ,  $\{b_i\}_{i \in \{1, \dots, M\}} \subseteq \mathbb{R}$  and  $\{c_i\}_{i \in \{1, \dots, M\}} \subseteq \mathbb{R}$  such that  $\forall i \in \{1, \dots, M\}$  we have  $b_i^2 < 4 \cdot c_i$  and  $\forall z \in \mathbb{C}$  we have  $p(z) = c \cdot (\prod_{i=1}^m (z - \lambda_i)) \cdot (\prod_{i=1}^M (z^2 + b_i \cdot x + c_i))$ .

**Proof.** We prove this by induction so let  $\mathcal{S} = \{n \in \mathbb{N} \mid \text{let } p \in \mathcal{P} \text{ with } \text{ord}(p) \leq n \text{ and real coefficients then } p(z) = (\prod_{i=1}^m (z - \lambda_i)) \cdot (\prod_{i=1}^M (z^2 + b_i \cdot x + c_i)) \forall z \in \mathbb{C} \text{ where } \{\lambda_i\}_{i \in \{1, \dots, m\}} \subseteq \mathbb{R}, \{b_i\}_{i \in \{1, \dots, M\}} \subseteq \mathbb{R} \text{ and } \{c_i\}_{i \in \{1, \dots, M\}} \subseteq \mathbb{R} \text{ such that } \forall i \in \{1, \dots, M\} \text{ we have } b_i^2 < 4 \cdot c_i\}$  then

**$1 \in \mathcal{S}$ .** If  $\text{ord}(p) = 1$  then  $\forall z \in \mathbb{C}$  we have  $p(z) = a' \cdot z + c' = a' \cdot \left(z - \left(-\frac{c'}{a'}\right)\right)$  where  $a', c' \in \mathbb{R}$  and  $a' \neq 0$  hence if we take  $c = a'$ ,  $\{\lambda_i\}_{i \in \{1, \dots, 1\}}$  defined by  $\lambda_1 = -\frac{c'}{a'}$  and  $\{b_i\}_{i \in \{1, \dots, 0\}} = \{c_i\}_{i \in \{1, \dots, 0\}} = \emptyset$  so that  $\forall i \in \{1, \dots, 0\} = \emptyset$   $b_i^2 < 4 \cdot c_i$  is satisfied vacuously. This proves that  $1 \in \mathcal{S}$

**$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$ .** So let  $p \in \mathcal{P}$  with real coefficients then if  $\text{ord}(p) < n+1$  then  $\text{ord}(p) \leq n$  and as  $n \in \mathcal{S}$  we have that  $p(z) = (\prod_{i=1}^m (z - \lambda_i)) \cdot (\prod_{i=1}^M (z^2 + b_i \cdot x + c_i))$ , hence we must only consider the case with  $\text{ord}(p) = n+1$ . As  $p$  is non constant we have using the previous theorem that there exists a  $\{\alpha_i\}_{i \in \{1, \dots, \text{ord}(p)\}}$  such that  $\forall z \in \mathbb{C}$  we have  $p(z) = c \cdot \prod_{i \in \{1, \dots, \text{ord}(p)\}} (z - \alpha_i)$ . If now  $\forall i \in \{1, \dots, \text{ord}(p) = n+1\}$  we have  $\alpha_i \in \mathbb{R}$  then  $n+1 \in \mathcal{S}$  by taking  $m = \text{ord}(p) = n$  and  $M = 0$ . So let assume that  $\exists i \in \{1, \dots, \text{ord}(p) = n+1\}$  such that  $\alpha_i \in \mathbb{C} \setminus \mathbb{R}$  then  $p(\bar{\alpha}_i) = 0$ . Using 12.30 we have that  $p(\bar{\alpha}_i) = 0$  and thus by 12.32 there exists a  $j \in \{1, \dots, \text{ord}(p) = n+1\}$  such that  $\bar{\alpha}_i = \alpha_j$  hence

$$\begin{aligned} p(z) &= c \cdot \prod_{k=1}^{n+1} (z - \alpha_k) \\ &= c \cdot \prod_{k \in \{1, \dots, n+1\}} (z - \alpha_k) \\ &= c \cdot \left( \prod_{k \in \{1, \dots, n+1\} \setminus \{i, j\}} (z - \alpha_k) \right) \cdot \prod_{k \in \{i, j\}} (z - \alpha_k) \\ &= c \cdot \left( \prod_{k \in \{1, \dots, n+1\} \setminus \{i, j\}} (z - \alpha_k) \right) \cdot (z - \alpha_i) \cdot (z - \alpha_j) \\ &= c \cdot \left( \prod_{k \in \{1, \dots, n+1\} \setminus \{i, j\}} (z - \alpha_k) \right) \cdot (z - \alpha_i) \cdot (z - \alpha_j) \end{aligned}$$

$$\begin{aligned}
&= c \cdot \left( \prod_{k \in \{1, \dots, n+1\} \setminus \{i, j\}} (z - \alpha_k) \right) \cdot (z - \alpha_i) \cdot (z - \bar{\alpha}_i) \\
&= c \cdot \left( \prod_{k \in \{1, \dots, n+1\} \setminus \{i, j\}} (z - \alpha_k) \right) \cdot (z^2 - (\alpha_i + \bar{\alpha}_i) \cdot z + \alpha_i \cdot \bar{\alpha}_i) \\
&= c \cdot \left( \prod_{k \in \{1, \dots, n+1\} \setminus \{i, j\}} (z - \alpha_k) \right) \cdot (z^2 - 2 \cdot \operatorname{Re}(\alpha_i) \cdot z + |\alpha_i|^2)
\end{aligned}$$

Hence if we define  $q(z) = c \cdot \left( \prod_{k \in \{1, \dots, \operatorname{ord}(p)\} \setminus \{i, j\}} (x - \alpha_k) \right)$  (which is a polynomial with  $\operatorname{ord}(q) = n+1-2 = n-1 < n$  (see 12.10)), so

$$p(z) = q(z) \cdot (z^2 + b \cdot z + c) \text{ where } q \in \mathcal{P}, \operatorname{ord}(q) < n+1 \text{ and } b = -2 \cdot \operatorname{Re}(\alpha_i) \in \mathbb{R}, c = |\alpha_i|^2 \in \mathbb{R} \quad (12.32)$$

Now as  $\alpha_i \in \mathbb{C} \setminus \mathbb{R}$  we have that  $0 < \operatorname{Img}(\alpha_i)$  giving  $(\operatorname{Re}(\alpha_i))^2 < (\operatorname{Re}(\alpha_i))^2 + (\operatorname{Img}(\alpha_i))^2 = |\alpha_i|^2$  hence  $b^2 = 4 \cdot \operatorname{Re}(\alpha_i)^2 < 4 \cdot |\alpha_i|^2 = 4 \cdot c$  proving that

$$b^2 < 4 \cdot c \quad (12.33)$$

As  $q$  is a polynomial of order  $n-1$  we have that  $q(z) = \sum_{k=0}^{n-1} d_k \cdot z^k$  Now given  $x \in \mathbb{R}$  we have  $x^2 + b \cdot x + c = \left( x + \frac{b}{2} \right)^2 + \left( c - \frac{b^2}{4} \right) > 0$  (see above) hence using 12.32 we have that  $\forall x \in \mathbb{R}$  that  $q(x) = \frac{p(x)}{x^2 + b \cdot x + c} \in \mathbb{R}$  (as  $p$  has real coefficients). So  $\forall x \in \mathbb{R}$  we have  $0 = \operatorname{Img}(q(x)) = \operatorname{Img}(\sum_{k=0}^{n-1} d_k \cdot z^k) = \sum_{k=0}^{n-1} \operatorname{Img}(d_k \cdot x^k) \underset{x \in \mathbb{R}}{=} \sum_{k=0}^{n-1} \operatorname{Img}(d_k) \cdot z^k$  hence using 12.2 we have that  $\forall k \in \{0, \dots, n-1\} \operatorname{Img}(d_k) = 0$  proving that  $q$  has real coefficients. As  $n \in \mathcal{S}$  there exists then  $m, M \in \mathbb{N}$ ,  $\{\lambda_k\}_{k \in \{1, \dots, m\}} \subseteq \mathbb{R}$ ,  $\{b'_k\}_{k \in \{1, \dots, M\}} \subseteq \mathbb{R}$  and  $\{c'_k\}_{k \in \{1, \dots, M\}} \subseteq \mathbb{R}$  such that  $\forall k \in \{1, \dots, M\} b'^2_k \leq 4 \cdot c'_k$  and  $\forall z \in \mathbb{C}$  we have  $q(z) = (\prod_{k=1}^m (z - \lambda_k)) \cdot (\prod_{k=1}^M (z^2 + b'_k \cdot z + c'_k))$ . Hence if we define  $\{b_k\}_{k \in \{1, \dots, M+1\}}$ ,  $\{c_k\}_{k \in \{1, \dots, M+1\}} \subseteq \mathbb{R}$  by  $b_k = \begin{cases} b'_k & \text{if } k \in \{1, \dots, M\} \\ b_{M+1} & = b \end{cases}$  and  $c_k = \begin{cases} c'_k & \text{if } k \in \{1, \dots, M\} \\ c_{M+1} & = c \end{cases}$  then  $\forall k \in \{1, \dots, M+1\}$  we have  $b_k^2 < 4 \cdot c_k$  and  $\forall z \in \mathbb{C}$  that  $p(z) = (\prod_{k=1}^m (z - \lambda_k)) \cdot (\prod_{k=1}^{M+1} (z^2 + b_k \cdot z + c_k))$ , finally proving that  $n+1 \in \mathcal{S}$   $\square$

## 12.2 Spectral theorem

### 12.2.1 Polynomials and operators

We use now the fundamental theorem of algebra to prove the spectral theorem of linear algebra. First we define the concept of operators

**Definition 12.34. (Linear operator)** Let  $\langle X, +, \cdot \rangle$  be a vector space then  $L: X \rightarrow X$  is a **linear operator on  $X$**  if  $L \in \operatorname{Hom}(X, X)$  (see 10.186). The set of all linear operators on  $X$  is called  $\operatorname{Hom}(X)$ .

**Theorem 12.35.** Let  $\langle X, +, \cdot \rangle$  be a non trivial finite dimensional vector space, then for  $L \in \operatorname{Hom}(X)$  we have the following equivalences

1.  $L$  is a isomorphism
2.  $L$  is injective
3.  $L$  is surjective

**Proof.**

1  $\Rightarrow$  2. This is trivial as a isomorphism is bijective and thus injective

**1  $\Rightarrow$  3.** This is trivial as a isomorphism is bijective and thus surjective

**2  $\Rightarrow$  1.** As  $L$  is injective we have by 10.222 that  $\text{rank}(L) = \dim(X)$  and by 10.222 again (using the fact that  $L: X \Rightarrow X$ ) that  $L$  is a isomorphism.

**3  $\Rightarrow$  1.** As  $L$  is surjective we have by 10.219 that  $\text{rank}(L) = \dim(X)$  and by 10.222 (using the fact that  $L: X \Rightarrow X$ ) that  $L$  is a isomorphism.  $\square$

**Definition 12.36.** Let  $\langle X, +, \cdot \rangle$  a vector space and  $\{L_i: X \rightarrow X\}_{i \in \{1, \dots, n\}}$  a family of operators between  $X$  then  $\dot{\prod}_{i=1}^n L_i: X \rightarrow X$  is defined recursively by  $(\dot{\prod}_{i=1}^n L_i)(x) = \begin{cases} L_1(x) & \text{if } n = 1 \\ L_n((\dot{\prod}_{i=1}^{n-1} L_i)(x)) & \text{if } n > 1 \end{cases}$

**Definition 12.37.** Let  $X$  be a vector space and  $L \in \text{Hom}(X)$  then given  $n \in \mathbb{N}$  we define  $L^n: X \rightarrow X$  recursively by  $L^n = \begin{cases} 1_X & \text{if } n = 0 \\ L \circ L^{n-1} & \text{if } n > 0 \end{cases}$

**Note 12.38.** As a application of the above definition we have we have that  $L^0 = 1_X$  and  $L^1 = L$

The power of operators behaves as a normal powerfunction.

**Lemma 12.39.** Let  $\langle X, +, \cdot \rangle$  be a vector space and  $L \in \text{Hom}(X)$  then given  $n, m \in \mathbb{N}$  we have  $L^n \circ L^m = L^{n+m}$

**Proof.** We prove this by induction on  $n$  so let  $\mathcal{S} = \{n \in \mathbb{N} \mid L^n \circ L^m = L^{n+m}\}$  then we have

**0  $\in$   $\mathcal{S}$ .** this follows from  $L^0 \circ L^m = 1_X \circ L^m = L^m = L^{0+m}$

**$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$ .** then  $L^{n+1} \circ L^m = (L \circ L^n) \circ L^m \stackrel{\text{composition is associative}}{=} L \circ (L^n \circ L^m) \stackrel{n \in \mathcal{S}}{=} L \circ L^{n+m} = L^{(n+1)+m}$  proving that  $n+1 \in \mathcal{S}$   $\square$

The power of a linear operator is again a linear operator

**Lemma 12.40.** Let  $\langle X, +, \cdot \rangle$  be a vector space over a field  $\mathcal{F}$ ,  $L \in \text{Hom}(X)$  and  $n \in \mathbb{N}$  then  $L^n \in \text{Hom}(X)$  ( $L^n$  is a linear operator)

**Proof.** Again this easily proved by induction and define  $\mathcal{S} = \{n \in \mathbb{N} \mid L^n \text{ is linear}\}$  then we have

**0  $\in$   $\mathcal{S}$ .** then  $L^0 = 1_X$  which is linear so that  $0 \in \mathcal{S}$

**$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$ .** As  $L^{n+1} = L \circ L^n$  which is linear as  $L$  is linear,  $L^n$  is linear and the composition of linear maps is linear (see 10.184) proving that  $n+1 \in \mathcal{S}$   $\square$

Using the definition of the power of a operator we can define a polynomial of linear operator

**Definition 12.41.** Let  $n \in \mathbb{N}$ ,  $\langle X, +, \cdot \rangle$  a complex vector space and  $p: \mathbb{C} \rightarrow \mathbb{C}$  a polynomial (see 12.1) defined by  $\{\alpha_i\}_{i \in \{1, \dots, n\}}$  and  $L \in \text{Hom}(X)$  a linear operator then  $p(L): X \rightarrow X$  is defined by  $p(E)(x) = \sum_{i=0}^n L^i(x)$ .  $p(L)$  is called a polynomial operator.

We show now that the product of polynomial functions is equal to the composition of the operator polynomials.

**Lemma 12.42.** Let  $\langle X, +, \cdot \rangle$  a complex vector space and  $p, q: \mathbb{C} \rightarrow \mathbb{C}$  polynomials and  $L \in \text{Hom}(X)$  then  $(p \cdot q)(L) = p(L) \circ q(L)$

**Proof.** Let  $p$  be defined  $p(z) = \sum_{i=0}^n a_i \cdot z^i$  and  $q$  defined by  $q(z) = \sum_{i=0}^m b_i \cdot z^i$  we have by 12.8 that

$$(p \cdot q)(z) = \sum_{i=0}^{n+m} \left( \sum_{(k, l) \in \{(r, s) \in \{0, \dots, n\} \times \{0, \dots, m\} \mid r+s=i\}} a_k \cdot b_l \right) \cdot z^i \quad (12.34)$$

Next we calculate  $p(L) \circ q(L)$  first we prove that given  $x \in X$  we have

$$(p(L) \circ q(L))(x) = \sum_{(k, l) \in \{0, \dots, n\} \times \{0, \dots, m\}} a_k \cdot b_l \cdot L^{k+l}(x) \quad (12.35)$$

To prove this note that

$$\begin{aligned}
(p(L) \circ q(L))(x) &= p(L)(q(L)(x)) \\
&= \sum_{i=0}^n a_i \cdot L^i(q(L)(x)) \\
&= \sum_{i=1}^n a_i \cdot L^i \left( \sum_{j=0}^m b_j \cdot L^j(x) \right) \\
&\stackrel{L^i \text{ is linear by 12.40}}{=} \sum_{i=0}^n L^i \left( a_i \cdot \sum_{j=0}^m b_j \cdot L^j(x) \right) \\
&= \sum_{i=0}^n L^i \left( \sum_{j=0}^m a_i \cdot b_j \cdot L^j(x) \right) \\
&\stackrel{10.176 \text{ and 12.40}}{=} \sum_{i=0}^n \left( \sum_{j=0}^m a_i \cdot b_j \cdot L^i(L^j(x)) \right) \\
&= \sum_{i=0}^n \left( \sum_{j=0}^m a_i \cdot b_j \cdot (L^i \circ L^j)(x) \right) \\
&\stackrel{12.39}{=} \sum_{i=0}^n \left( \sum_{j=0}^m a_i \cdot b_j \cdot L^{i+j}(x) \right) \\
&= \sum_{i \in \{0, \dots, n\}} \left( \sum_{j \in \{0, \dots, m\}} a_i \cdot b_j \cdot L^{i+j}(x) \right) \\
&\stackrel{10.47}{=} \sum_{(k, l) \in \{0, \dots, n\} \times \{0, \dots, m\}} a_k \cdot b_l \cdot L^{k+l}(x)
\end{aligned}$$

which proves 12.35. Now using 12.7 we have that  $\{0, \dots, n\} \times \{0, \dots, m\}$  is a disjoint union of the sets  $\{(r, s) \subseteq \{0, \dots, n\} \times \{0, \dots, m\} \mid r + s = i\}$  where  $i \in \{0, \dots, n+m\}$ , so using 12.35 together with 10.46 we have

$$\begin{aligned}
(p(L) \circ q(L))(x) &= \sum_{i \in \{0, \dots, n+m\}} \left( \sum_{(k, l) \in \{(r, s) \subseteq \{0, \dots, n\} \times \{0, \dots, m\} \mid r + s = i\}} a_k \cdot b_l \cdot L^{k+l}(x) \right) \\
&= \sum_{i \in \{0, \dots, n+m\}} \left( \sum_{(k, l) \in \{(r, s) \subseteq \{0, \dots, n\} \times \{0, \dots, m\} \mid r + s = i\}} a_k \cdot b_l \cdot L^i(x) \right) \\
&= \sum_{i \in \{0, \dots, n+m\}} \left( \sum_{(k, l) \in \{(r, s) \subseteq \{0, \dots, n\} \times \{0, \dots, m\} \mid r + s = i\}} a_k \cdot b_l \right) \cdot L^i(x) \\
&\stackrel{12.34}{=} (p \cdot q)(L)(x)
\end{aligned}$$

which as  $x \in X$  is chosen arbitrary proves that  $(p \cdot q)(L) = p(L) \circ q(L)$   $\square$

We extend the above by recursion to the composition of  $n$  linear operators

**Corollary 12.43.** *Let  $n \in \mathbb{N}_0$ ,  $\langle X, +, \cdot \rangle$  a complex vector space and  $L \subseteq \text{Hom}(X)$  and  $\{p_i\}_{i \in \{1, \dots, n\}} \subseteq \mathcal{P}$  a finite non empty family of polynomials (hence by 12.10  $\prod_{i=1}^n p_i$  is a polynomial) then  $(\prod_{i=1}^n p_i)(L) = \prod_{i=1}^n p_i(L)$  proving that  $1 \in \mathcal{S}$*

**Proof.** This is easily proved by induction using the previous lemma. So let  $\mathcal{S} = \{n \in \mathbb{N}_0 \mid \text{If } \{p_i\}_{i \in \{1, \dots, n\}} \subseteq \mathcal{P} \text{ then } (\prod_{i=1}^n p_i)(L) = \prod_{i=1}^n p_i(L)\}$  then we have

$$1 \in \mathcal{S}. \text{ then } (\prod_{i=1}^1 p_i)(L) = p_1(L) = \prod_{i=1}^1 p_i(L)$$

$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$ . if  $\{p_i\}_{i \in \{1, \dots, n+1\}} \subseteq \mathcal{P}$  then

$$\begin{aligned} \left( \prod_{i=1}^{n+1} p_i \right)(L) &= \left( p_{n+1} \cdot \prod_{i=1}^n p_i \right)(L) \\ &\stackrel{12.42}{=} p_{n+1}(L) \circ \left( \prod_{i=1}^n p_i \right)(L) \\ &\stackrel{n \in \mathcal{S}}{=} p_{n+1}(L) \circ \left( \prod_{i=1}^n p_i(L) \right) \\ &= \prod_{i=1}^{n+1} p_i(L) \end{aligned}$$

proving that  $n+1 \in \mathcal{S}$

□

**Definition 12.44.** Let  $\langle X, +, \cdot \rangle$  be a complex vector space,  $L \in \text{Hom}(X)$  then  $\mathcal{P}[L] = \{p(L) | p \in \mathcal{P}\}$  where  $\mathcal{P}$  is the set of all complex polynomials. (see 12.1)

Using the composition as a additive operator we can form a additive semi group

**Theorem 12.45.** Let  $\langle X, +, \cdot \rangle$  be a complex vector space,  $L \in \text{Hom}(X)$  then  $\langle \mathcal{P}[L], \circ \rangle$  is a abelian semi group

**Proof.** First if  $p_1(L), p_2(L) \in \mathcal{P}[L]$  then  $p_1(L) \circ p_2(L) \stackrel{12.42}{=} (p_1 \cdot p_2)(L) \in \mathcal{P}[L]$  we have that  $\circ: \mathcal{P}[L] \times \mathcal{P}[L] \rightarrow \mathcal{P}[L]$  further we have that  $\circ$  as the composition of functions is associative,  $1_X = \sum_{i=0}^0 1 \cdot L^i$  is the neutral element and  $p_1(L) \circ p_2(L) = (p_1 \cdot p_2)(L) = (p_2 \cdot p_1)(L) = p_2(L) \circ p_1(L)$  □

Combining the above with the commutativity of a sum for semi groups (see 10.26) we have the following

**Theorem 12.46.** Let  $\langle X, +, \cdot \rangle$  be a complex vector space,  $L \in \text{Hom}(X)$ ,  $n \in \mathbb{N}_0$ ,  $\{p_i\}_{i \in \{1, \dots, n\}} \subseteq \mathcal{P}$  and  $\sigma: \{1, \dots, n\} \rightarrow \{1, \dots, n\}$  a permutation then  $\prod_{i=1}^n p_i(L) = \prod_{i=1}^n p_{\sigma(i)}(L)$

We are now read to prove the following theorem

**Theorem 12.47.** Let  $\langle X, +, \cdot \rangle$  be a complex vector space,  $L \in \text{Hom}(X)$  and  $p \in \mathcal{P}$  a polynomial with  $\text{ord}(p) \in \mathbb{N}_0$  (a non constant polynomial) then there exists a  $0 \neq c \in \mathbb{C}$  and  $\{\lambda_i\}_{i \in \{1, \dots, n\}}$  such that  $p(L) = c \cdot \prod_{i=1}^n (L - \lambda_i \cdot L)$

**Proof.** Using 12.32 there exists a  $0 \neq c \in \mathbb{C}$  and  $\{\lambda_i\}_{i \in \{1, \dots, n\}}$  such that  $p(z) = c \cdot \prod_{i=1}^n (z - \lambda_i) = C_c(z) \cdot \prod_{i=1}^n (z - \lambda_i)$ , now  $\forall i \in \{1, \dots, n\}$  we have  $z - \lambda_i = \sum_{j=0}^1 c_{i,j} \cdot z^j$  where  $c_{i,0} = -\lambda_i$  and  $c_{i,1} = 1$  so that  $(z - \lambda_i)(L) = c_{i,1} \cdot L^1 + c_{i,0} L^0 = L - \lambda_i \cdot 1_X$ . As  $c$  is the polynomial  $C_c$  defined by  $\sum_{i=0}^0 d_i \cdot z$  where  $d_i = 0$  so that  $C_c(L) = c \cdot 1_X$ . Hence we have that

$$\begin{aligned} p(z)(L) &\stackrel{12.42}{=} c_c(L) \circ \left( \prod_{i=1}^n (z - \lambda_i) \right)(L) \\ &\stackrel{12.43}{=} c \cdot 1_X \circ \left( \prod_{i=1}^n (z - \lambda_i)(L) \right) \\ &= c \cdot \prod_{i=1}^n (L - \lambda_i \cdot 1_X) \end{aligned}$$

□

**Lemma 12.48.** Let  $\langle X, +, \cdot \rangle$  be a vector space over a field  $\mathcal{F}$ ,  $n \in \mathbb{N}_0$  and  $\{L_i\}_{i \in \{1, \dots, n\}} \subseteq \text{Hom}(X)$  such that  $\forall i \in \{1, \dots, n\}$  we have that  $\ker(L_i) = \{0\}$  (meaning by 10.213 that  $L_i$  is regular (=injective)) then  $\ker(\prod_{i=1}^n L_i) = \{0\}$

**Proof.** We prove this by induction so let  $\mathcal{S} = \{n \in \mathbb{N}_0 \mid \text{if } \{L_i\}_{i \in \{1, \dots, n\}} \in \text{Hom}(X) \text{ with } \forall i \in \{1, \dots, n\} \vdash \ker(L_i) = \{0\} \text{ then } \ker(\dot{\prod}_{i=1}^n L_i) = \{0\}\}$  then we have

$$1 \in \mathcal{S}, \text{ then } \dot{\prod}_{i=1}^1 L_i = L_1 \text{ then } \ker(\dot{\prod}_{i=1}^1 L_i) = \ker(L_1)$$

$$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}. \text{ then assume that } x \in \ker(\dot{\prod}_{i=1}^{n+1} L_i) \text{ then we have } 0 = (\dot{\prod}_{i=1}^{n+1} L_i)(x) = (L_{n+1} \circ \dot{\prod}_{i=1}^n L_i)(x) = L_{n+1}((\dot{\prod}_{i=1}^n L_i)(x)) \text{ and as } \ker(L_{n+1}) = \{0\} \text{ we have that } (\dot{\prod}_{i=1}^n L_i)(x) = 0 \text{ so that as } n \in \mathcal{S} \text{ we conclude that } x = 0 \text{ hence } \ker(\dot{\prod}_{i=1}^{n+1} L_i) = \{0\} \text{ proving that } n+1 \in \mathcal{S} \quad \square$$

### 12.2.2 Eigenvectors and eigen values

**Definition 12.49.** Let  $\langle X, +, \cdot \rangle$  be a vector space over a field  $\mathcal{F}$ ,  $L \in \text{Hom}(X)$  then  $x$  is a **eigen-vector** of  $L$  with **eigenvalue**  $\lambda$  iff  $x \neq 0$  and  $L(x) = \lambda \cdot x$

**Remark 12.50.** The condition that  $x \neq 0$  is needed because we have  $L \cdot 0 = 0 = \lambda \cdot 0$  for every  $\lambda \in \mathcal{F}$  so the concept of a special eigen value does not make sense.

In a complex finite dimensional vector space every linear operator has at least one eigen vector as is expressed in the following theorem.

**Theorem 12.51.** Let  $\langle X, +, \cdot \rangle$  be a non trivial complex finite dimensional non trivial vector space then every  $L \in \text{Hom}(X)$  has a eigen vector with a eigen value.

**Proof.** Let  $n = \dim(X) \in \mathbb{N}_0$  then as  $X$  is non trivial there exists a  $x \in X$  with  $x \neq 0$ , consider then the family  $\{L^i(x)\}_{i \in \{0, \dots, n\}}$  of  $n+1$  vectors in  $X$  which as  $\dim(X)$  is  $n$  must be linear dependent. Hence there exists a  $\{a_i\}_{i \in \{0, \dots, n\}}$  not all equal to 0 such that  $\sum_{i=0}^n a_i \cdot L^i(x) = 0$ . If we take then  $m = \max(\{i \in \{0, \dots, n\} \mid a_i \neq 0\})$  then we have that

$$\sum_{i=0}^m a_i \cdot L^i(x) = 0 \wedge a_m \neq 0 \quad (12.36)$$

If now  $m = 0$  then  $0 = \sum_{i=0}^m a_i \cdot L^i(x) = a_0 \cdot L^0(x) = a_0 \cdot x = 0 \xrightarrow{a_0 \neq 0} x = 0$  a contradiction hence we have  $m > 0$  and  $p \in \mathcal{P}$  defined by  $p(z) = \sum_{i=0}^m a_i \cdot z^i$  is a non constant polynomial. Applying then 12.47 we find that there exists a  $0 \neq c \in \mathbb{C}$  and  $\{\lambda_i\}_{i \in \{0, \dots, m\}} \subseteq \mathbb{C}$  we have that

$$\sum_{i=0}^m a_i \cdot L^i = c \cdot \prod_{i=1}^m (L - \lambda_i \cdot 1_X) \quad (12.37)$$

So we have

$$\begin{aligned} 0 &\stackrel{12.36}{=} \sum_{i=0}^m a_i \cdot L^i(x) \\ &\stackrel{12.37}{=} c \cdot \left( \prod_{i=1}^m (L - \lambda_i \cdot 1_X) \right)(x) \end{aligned}$$

Hence as  $v \neq 0$  we have that  $c \neq 0$  we have that  $(\prod_{i=1}^m (L - \lambda_i \cdot 1_X))(v) = 0$  and as  $v \neq 0$  we have using 12.48 that there exists a  $i \in \{1, \dots, m\}$  such that  $\ker(L - \lambda_i \cdot 1_X) \neq \{0\}$  and thus there exists a  $0 \neq w \in \ker(L - \lambda_i \cdot 1_X)$  or in other word there exists a  $0 \neq w \in X$  such that  $(L - \lambda \cdot 1_X)(w) = 0 \Rightarrow L(w) - \lambda \cdot w = 0 \Rightarrow L(w) = \lambda \cdot w$  proving the theorem.  $\square$

**Definition 12.52.** Let  $\langle X, +, \cdot \rangle$  be a vector space,  $Y \subseteq X$  a sub vector space (see 10.110) and  $T \in \text{Hom}(X)$  then  $X$  is **invariant** under  $L$  if  $L(Y) \subseteq Y$  [in other words  $\forall u \in Y$  we have  $L(u) \in Y$ ]

**Definition 12.53.** Let  $n \in \mathbb{N}_0$ ,  $\mathcal{F}$  be a field then  $M \in \mathcal{M}(n \times n, \mathcal{F})$  (the set of all the  $n$  by  $n$  square matrices) is a **upper triangular matrix** if  $\forall i, j \in \{1, \dots, n\}$  with  $i > j$  we have  $M_{i,j} = 0$

A special case of a upper triangular matrix is a diagonal matrix.

**Definition 12.54.** Let  $n \in \mathbb{N}_0$ ,  $\mathcal{F}$  be a field then  $M \in \mathcal{M}(n \times n, \mathcal{F})$  (the set of all the  $n$  by  $n$  square matrices) is a **diagonal matrix** if  $\forall i, j \in \{1, \dots, n\}$  with  $i \neq j$  we have  $M_{i,j} = 0$

The following theorem gives equivalent conditions for the matrix of a linear operator to be triangular

**Theorem 12.55.** Let  $\langle X, +, \cdot \rangle$  be a non trivial vector space (hence  $n = \dim(V) \in \mathbb{N}_0$ ),  $\{e_i\}_{i \in \{1, \dots, n\}}$  a basis for  $X$ ,  $L \in \text{Hom}(X)$  a linear operator the following are equivalent

1.  $\mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}})$  is upper triangular
2.  $\forall i \in \{1, \dots, n\}$  we have  $L(e_i) \in \mathcal{S}(\{e_k | k \in \{1, \dots, i\}\})$  (here  $\mathcal{S}$  is the span of  $\{e_k\}_{k \in \{1, \dots, i\}}$  see 10.130)
3.  $\forall i \in \{1, \dots, n\}$  we have that  $\mathcal{S}(\{e_k | k \in \{1, \dots, i\}\})$  is invariant under  $L$

**Proof.** Let  $M = \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}})$

**1  $\Rightarrow$  2.** Let  $i \in \{1, \dots, n\}$  then we have  $L(e_i) = \sum_{k=1}^n M_{k,i} \cdot e_k = \sum_{k \in \{1, \dots, i\}} M_{k,i} \cdot e_k + \sum_{k \in \{i+1, \dots, n\}} M_{k,i} \cdot e_k = \sum_{k \in \{1, \dots, i\}} M_{k,i} \cdot e_k \in \mathcal{S}(\{e_k | k \in \{1, \dots, i\}\})$

**2  $\Rightarrow$  1.** As  $\forall i \in \{1, \dots, n\}$  we have that  $L(e_i) \in \mathcal{S}(\{e_k | k \in \{1, \dots, i\}\})$  we have that  $L(e_i) = \sum_{k=1}^i L_{k,i} \cdot e_k = \sum_{k=1}^i M_{k,i} e_k$  where  $M_{k,i} = \begin{cases} 0 & \text{if } k > i \\ L_{k,i} & \text{if } k \in \{1, \dots, i\} \end{cases}$  so using the definition of  $\mathcal{M}(L, \{e_k\}_{k \in \{1, \dots, n\}}, \{e_k\}_{k \in \{1, \dots, n\}})$  (see 10.302) we have that  $\mathcal{M}(L, \{e_j\}_{j \in \{1, \dots, n\}}, \{e_j\}_{j \in \{1, \dots, n\}})_{k,i} = M_{k,i}$  hence  $\mathcal{M}(L, \{e_j\}_{j \in \{1, \dots, n\}}, \{e_j\}_{j \in \{1, \dots, n\}})_{k,i} = 0$  if  $k > i$  proving upper triangularity.

**2  $\Rightarrow$  3.** Given  $i \in \{1, \dots, n\}$  take  $x \in \mathcal{S}(\{e_k | k \in \{1, \dots, i\}\}) \Rightarrow x = \sum_{k=1}^i x_k \cdot e_k$  then as  $\forall l \in \{1, \dots, k\}$  we have  $\{e_k | k \in \{1, \dots, l\}\} \subseteq \{e_k | k \in \{1, \dots, i\}\} \xrightarrow{10.138} \mathcal{S}(\{e_k | k \in \{1, \dots, l\}\}) \subseteq \mathcal{S}(\{e_k | k \in \{1, \dots, i\}\})$  hence  $\forall l \in \{1, \dots, i\}$  we have  $L(e_l) \in \mathcal{S}(\{e_k | k \in \{1, \dots, i\}\})$ . So  $L(x) = L(\sum_{k=1}^i x_k \cdot e_k) = \sum_{k=1}^i x_k \cdot L(e_k) \in \mathcal{S}(\mathcal{S}(\{e_k | k \in \{1, \dots, i\}\})) \xrightarrow{10.137} \mathcal{S}(\{e_k | k \in \{1, \dots, i\}\})$  proving that  $\mathcal{S}(\{e_k | k \in \{1, \dots, i\}\})$  is invariant under  $L$ .

**3  $\Rightarrow$  2.** Given  $i \in \{1, \dots, n\}$  we have as  $\mathcal{S}(\{e_k | k \in \{1, \dots, i\}\})$  is invariant under  $L$  that  $L(e_i) \in \mathcal{S}(\{e_k | k \in \{1, \dots, i\}\})$ .  $\square$

**Theorem 12.56.** Let  $\langle X, +, \cdot \rangle$  be a complex finite dimensional non trivial vector space with  $\dim(X) = n \in \mathbb{N}_0$  and  $L \in \text{Hom}(X)$  then there exists a basis  $\{e_i\}_{i \in \{1, \dots, n\}}$  such that  $\mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}})$  is upper triangular.

**Proof.** We proof this by induction so let  $\mathcal{F} = \{n \in \mathbb{N}_0\}$  if  $X$  is a  $m$ -dimensional complex vector space with  $m \leq n$   $L \in \text{Hom}(X)$  then there exists a basis  $\{e_i\}_{i \in \{1, \dots, n\}}$  such that  $\mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}})$  then

**1  $\in \mathcal{F}$ .** If  $\dim(X) = 1$  then there exists a  $\{e_i\}_{i \in \{1, \dots, 1\}}$  such that if  $L \in \text{Hom}(X)$  we have  $L(e_1) = \sum_{i=1}^1 \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, 1\}}, \{e_i\}_{i \in \{1, \dots, 1\}})_{i,1} \cdot e_i = \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, 1\}}, \{e_i\}_{i \in \{1, \dots, 1\}})$  which as its a  $1 \times 1$  matrix is upper triangular.

**$n \in \mathcal{F} \Rightarrow n+1 \in \mathcal{F}$ .** Assume that  $\langle X, +, \cdot \rangle$  is a  $n+1$ -dimensional complex vector spaces and that  $L \in \text{Hom}(X)$  is a linear operator. As  $L$  has a eigenvector  $x$  with eigenvalue  $\lambda$  (see 12.51) so that  $(L - \lambda \cdot I_X)(x) = 0$  hence  $0 \neq x \in \ker(L - \lambda \cdot I_X) \Rightarrow \ker(L - \lambda \cdot I_X) \neq \{0\}$  proving that  $\dim(\ker(L - \lambda \cdot I_X)) > 0$ . Hence for  $(L - \lambda \cdot I_X) \in \text{Hom}(X)$  we have that  $\dim(X) \xrightarrow{10.220} \dim(\ker(L - \lambda \cdot I_X)) + \text{rank}(L - \lambda \cdot I_X)$  proving that  $\dim((L - \lambda \cdot I_X)(X)) < \dim(X)$ . Hence

$$\text{If } U = (L - \lambda \cdot I_X)(X) \text{ and } \dim(U) < \dim(X) \quad (12.38)$$

Now if  $u \in U$  we have  $L(u) = L(u) - \lambda \cdot u + \lambda \cdot u = (L - \lambda \cdot 1_X)(u) + \lambda \cdot u \in U$  (as  $U$  is a vector space by 10.216) hence

$$L(U) \subseteq U \quad (12.39)$$

So using the above we have that  $L|_U: U \rightarrow U$  is in  $\text{Hom}(U)$  where  $\dim(U) < \dim(X) = n + 1 \Rightarrow \dim(U) \leq n$  hence as  $n \in \mathcal{F}$  we have that  $\exists \{u_i\}_{i \in \{1, \dots, \dim(U)\}}$  a basis for  $U$  such that  $\mathcal{M}(L|_U, \{u_i\}_{i \in \{1, \dots, \dim(U)\}}, \{u_i\}_{i \in \{1, \dots, \dim(U)\}})$  is upper triangular and by 12.55 we have

$$\forall j \in \{1, \dots, \dim(U)\} \quad L(u_j) = (L|_U)(u_j) \in \mathcal{S}(\{u_k | k \in \{1, \dots, j\}\}) \quad (12.40)$$

As  $\{u_i\}_{i \in \{1, \dots, \dim(U)\}}$  is linear independent we can extend  $\{u_i\}_{i \in \{1, \dots, \dim(U)\}}$  by 10.205 to a basis  $\{w_i\}_{i \in \{1, \dots, n+1\}}$  of  $X$ . Hence using 12.40 we have

$$\forall j \in \{1, \dots, \dim(U)\} \quad L(w_j) \in \mathcal{S}(\{w_k | k \in \{1, \dots, j\}\}) \quad (12.41)$$

Now  $\forall j \in \{\dim(U) + 1, \dots, n + 1\}$  we have  $L(w_j) = (L - \lambda \cdot 1_X)(w_j) + \lambda \cdot 1_X(w_j) = (L - \lambda \cdot 1_X)(w_j) + \lambda \cdot w_j$ . As  $(L - \lambda \cdot 1_X)(w_j) \in \text{range}(L - \lambda \cdot 1_X) = U = \mathcal{S}(\{u_k | k \in \{1, \dots, \dim(U)\}\}) = \mathcal{S}(\{w_k | k \in \{1, \dots, \dim(U)\}\}) \subseteq_{10.138} \mathcal{S}(\{w_k | k \in \{1, \dots, j\}\})$  and  $w_j \in \mathcal{S}(\{w_k | k \in \{1, \dots, j\}\})$  we have as  $\mathcal{S}(\{w_k | k \in \{1, \dots, j\}\})$  is a vector space (see 10.132) we conclude that  $L(w_j) \in \mathcal{S}(\{w_k | k \in \{1, \dots, j\}\})$ . To summarize we have that

$$\forall j \in \{\dim(U) + 1, \dots, n + 1\} \quad L(w_j) \in \mathcal{S}(\{w_k | k \in \{1, \dots, j\}\}) \quad (12.42)$$

From 12.41 and 12.42 we conclude that  $\forall j \in \{1, \dots, n + 1\}$  we have that  $L(w_j) \in \mathcal{S}(\{w_k | k \in \{1, \dots, j\}\})$  hence using 12.55 we have that  $\mathcal{M}(L, \{w_i\}_{i \in \{1, \dots, n+1\}}, \{w_i\}_{i \in \{1, \dots, n+1\}})$  is upper triangular, proving that  $n + 1 \in \mathcal{F}$   $\square$

**Theorem 12.57. (Schur's Theorem)** *Let  $\langle X, \langle \cdot, \cdot \rangle \rangle$  be a real (complex) finite dimensional inner product space with  $\dim(X) = n \in \mathbb{N}_0$  and  $L \in \text{Hom}(X)$  then  $L$  has a orthonormal basis  $\{e_i\}_{i \in \{1, \dots, n\}}$  such that  $\mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}})$  is upper triangular.*

**Proof.** Using the previous theorem there exists a basis  $\{f_i\}_{i \in \{1, \dots, n\}}$  such that  $\mathcal{M}(L, \{f_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{1, \dots, n\}})$  is diagonal and using 12.55 we have that  $\forall j \in \{1, \dots, n\}$  we have that  $L(\mathcal{S}\{f_i | i \in \{1, \dots, n\}\}) \subseteq \mathcal{S}\{f_i | i \in \{1, \dots, n\}\}$ . Using the Gram-Schmidt procedure (see 11.119) we find a orthonormal basis  $\{e_i\}_{i \in \{1, \dots, n\}}$  such that  $\forall j \in \{1, \dots, n\}$  we have that  $\mathcal{S}(\{f_i | i \in \{1, \dots, n\}\}) = \mathcal{S}(e_i | i \in \{1, \dots, n\})$  hence  $L(\{e_i | i \in \{1, \dots, n\}\}) = L(\{f_i | i \in \{1, \dots, n\}\}) \subseteq \{f_i | i \in \{1, \dots, n\}\} = \{e_i | i \in \{1, \dots, n\}\}$  hence by 12.55 we have that  $\mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}})$  is upper triangular.  $\square$

**Definition 12.58.** *Let  $\langle X, \langle \cdot, \cdot \rangle \rangle$  be a real (complex) finite dimensional inner product space then  $L \in \text{Hom}(X)$  is **self-adjoint** iff  $L = L^*$  (or using the definition of the adjoint of a operator (see 11.123) we have  $\forall x \in X$  that  $\langle L(x), y \rangle = \langle x, L^*(y) \rangle = \langle x, L(y) \rangle$ )*

**Theorem 12.59.** *Let  $\langle X, \langle \cdot, \cdot \rangle \rangle$  be a real (complex) finite dimensional inner product space with a basis  $\{e_i\}_{i \in \{1, \dots, n\}}$  and  $L \in \text{Hom}(X)$  then the following are equivalent*

1.  $L$  is self-adjoint
2.  $\mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}})$  is Hermitian

**Proof.**

**1  $\Rightarrow$  2.**

$$\begin{aligned} \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}) &\stackrel{L^* = L}{=} \mathcal{M}(L^*, \{e_i\}_{i \in \{1, \dots, n\}}) \\ &\stackrel{11.132}{=} \mathcal{M}(L^*, \{e_i\}_{i \in \{1, \dots, n\}})^H \end{aligned}$$

**2  $\Rightarrow$  1.** Take  $x = \sum_{i=1}^n x_i \cdot e_i \in X$  then we have for  $i \in \{1, \dots, n\}$  that

$$\begin{aligned} (L(x))_i &= \sum_{k=1}^n \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}})_{i,k} \cdot x_k \\ &= \sum_{k=1}^n \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}})_{i,k}^H \cdot x_k \\ &\stackrel{11.132}{=} \sum_{k=1}^n \mathcal{M}(L^*, \{e_i\}_{i \in \{1, \dots, n\}})_{i,k} \cdot x_k \\ &= (L^*(x))_i \end{aligned}$$

□

**Corollary 12.60.** Let  $\langle X, \langle \cdot, \cdot \rangle \rangle$  be a real finite dimensional inner product space  $L \in \text{Hom}(X)$  such that there exists a basis  $\{e_i\}_{i \in \{1, \dots, n\}}$  so that  $\mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}})$  is diagonal then  $L$  is self-adjoint.

**Proof.** This follows from the fact that  $\forall i, j \in \{1, \dots, n\}$  we have  $(\mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}))^H)_{i,j} = (\mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}})^*)_{j,i} \stackrel{\text{realspace}}{=} \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}})_{j,i} = \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}})_{i,i} \cdot \delta_{j,i} = \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}})_{i,j}$  so that  $\mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}})$  is Hermitian and thus by the above theorem we conclude that  $L$  is self-adjoint. □

**Example 12.61.** Let  $\langle X, \langle \cdot, \cdot \rangle \rangle$  be a real (complex) finite dimensional inner product space then  $1_X \in \text{Hom}(X)$  is self-adjoint as  $\forall x, y \in X$  we have  $\langle 1_X(x), y \rangle = \langle x, y \rangle = \langle x, 1_X(y) \rangle$

**Theorem 12.62.** Let  $\langle X, \langle \cdot, \cdot \rangle \rangle$  be a real (complex) finite dimensional inner product space and  $L \in \text{Hom}(X)$  a self-adjoint operator then every eigenvalue of  $L$  is real.

**Proof.** In the real case the theorem is trivial, so let's consider the complex case. Let  $x \in X$  be a eigen vector of  $L$  with eigenvalue  $\lambda$  then  $x \neq 0$  and  $L(x) = \lambda \cdot x$ . Hence

$$\begin{aligned} \lambda \cdot \|x\|^2 &= \lambda \cdot \langle x, x \rangle \\ &= \langle \lambda \cdot x, x \rangle \\ &= \langle L(x), x \rangle \\ &\stackrel{L \text{ is self-adjoint}}{=} \langle L^*(x), x \rangle \\ &= \langle x, L(x) \rangle \\ &= \langle x, \lambda \cdot x \rangle \\ &= \bar{\lambda} \cdot \langle x, x \rangle \\ &= \bar{\lambda} \cdot \|x\|^2 \end{aligned}$$

Hence as  $\|x\|^2 \neq 0$  we have  $\lambda \in \mathbb{R}$ . □

**Theorem 12.63.** Let  $\langle X, \langle \cdot, \cdot \rangle \rangle$  be a real finite dimensional inner product space and  $U \in \text{Hom}(X)$  a unitary operator then  $\det(U^*) = \det(U) = \pm 1$

**Proof.** We have  $1 = \det(1_X) = \det(U^* \circ U) \stackrel{10.282}{=} \det(U^*) \cdot \det(U) \stackrel{11.133}{=} \det(U)^2 = \det(U^*)^2$  proving that  $\det(U) = \pm 1$  and  $\det(U^*) = \pm 1$  □

**Definition 12.64. (Unitary operator)** Let  $\langle X, \langle \cdot, \cdot \rangle \rangle$  be a real (complex) finite dimensional inner product space then  $L \in \text{Hom}(X)$  is **unitary** if  $L^* \circ L = 1_X$ .

Every unitary operator has a unitary matrix in every basis as is expressed in the following theorem.

**Theorem 12.65.** Let  $n \in \mathbb{N}_0$ ,  $\langle X, \langle \cdot, \cdot \rangle \rangle$  be a real (complex) finite dimensional inner product space,  $\{e_i\}_{i \in \{1, \dots, n\}}$  a basis in  $X$  and  $U \in \text{Hom}(X)$  then we have the following equivalences

1.  $U$  is unitary

2.  $\mathcal{M}(U, \{e_i\}_{i \in \{1, \dots, n\}})$  is a unitary matrix

**Proof.**

**1  $\Rightarrow$  2.**

$$\begin{aligned}
 & \mathcal{M}(U, \{e_i\}_{i \in \{1, \dots, n\}})^H \cdot \mathcal{M}(U, \{e_i\}_{i \in \{1, \dots, n\}}) \\
 & \stackrel{11.132}{=} \mathcal{M}(U^*, \{e_i\}_{i \in \{1, \dots, n\}}) \cdot \mathcal{M}(U, \{e_i\}_{i \in \{1, \dots, n\}}) \\
 & \stackrel{10.307}{=} \mathcal{M}(U^* \circ U, \{e_i\}_{i \in \{1, \dots, n\}}) \\
 & = \mathcal{M}(1_X, \{e_i\}_{i \in \{1, \dots, n\}}) \\
 & = E \\
 & = \mathcal{M}(1_X, \{e_i\}_{i \in \{1, \dots, n\}}) \\
 & = \mathcal{M}(U \circ U^*, \{e_i\}_{i \in \{1, \dots, n\}}) \\
 & \stackrel{10.307}{=} \mathcal{M}(U, \{e_i\}_{i \in \{1, \dots, n\}}) \cdot \mathcal{M}(U^*, \{e_i\}_{i \in \{1, \dots, n\}}) \\
 & \stackrel{11.132}{=} \mathcal{M}(U, \{e_i\}_{i \in \{1, \dots, n\}}) \cdot \mathcal{M}(U, \{e_i\}_{i \in \{1, \dots, n\}})^H
 \end{aligned}$$

**2  $\Rightarrow$  1.** Let  $x \in X$  then we have that  $x = \sum_{i=1}^n x_i \cdot e_i$  and

$$\begin{aligned}
 ((U^* \circ U)(x))_i &= \sum_{k=1}^n \mathcal{M}(U^* \circ U, \{e_i\}_{i \in \{1, \dots, n\}})_{i,k} \cdot x_k \\
 &\stackrel{10.307}{=} \sum_{k=1}^n (\mathcal{M}(U^*, \{e_i\}_{i \in \{1, \dots, n\}}) \cdot \mathcal{M}(U, \{e_i\}_{i \in \{1, \dots, n\}}))_{i,k} \cdot x_k \\
 &\stackrel{11.132}{=} \sum_{k=1}^n (\mathcal{M}(U, \{e_i\}_{i \in \{1, \dots, n\}})^H \cdot \mathcal{M}(U, \{e_i\}_{i \in \{1, \dots, n\}}))_{i,k} \cdot x_k \\
 &= \sum_{k=1}^n E_{i,k} \cdot x_k \\
 &= \sum_{k=1}^n \delta_{i,k} \cdot x_k \\
 &= x_i
 \end{aligned}$$

proving that  $U^* \circ U = E$

□

a unary operator then  $\mathcal{M}(U, \{e_i\}_{i \in \{1, \dots, n\}})$  is a unitary matrix

**Proof.**

$$\begin{aligned}
 \mathcal{M}(U, \{e_i\}_{i \in \{1, \dots, n\}})^H \cdot \mathcal{M}(U, \{e_i\}_{i \in \{1, \dots, n\}}) &\stackrel{11.132}{=} \mathcal{M}(U^*, \{e_i\}_{i \in \{1, \dots, n\}}) \cdot \mathcal{M}(U, \{e_i\}_{i \in \{1, \dots, n\}}) \\
 &\stackrel{10.307}{=} \mathcal{M}(U^* \circ U, \{e_i\}_{i \in \{1, \dots, n\}}) \\
 &= \mathcal{M}(1_X, \{e_i\}_{i \in \{1, \dots, n\}}) \\
 &= E \\
 &= \mathcal{M}(1_X, \{e_i\}_{i \in \{1, \dots, n\}}) \\
 &= \mathcal{M}(U \circ U^*, \{e_i\}_{i \in \{1, \dots, n\}}) \\
 &\stackrel{10.307}{=} \mathcal{M}(U, \{e_i\}_{i \in \{1, \dots, n\}}) \cdot \mathcal{M}(U^*, \{e_i\}_{i \in \{1, \dots, n\}}) \\
 &\stackrel{11.132}{=} \mathcal{M}(U, \{e_i\}_{i \in \{1, \dots, n\}}) \cdot \mathcal{M}(U, \{e_i\}_{i \in \{1, \dots, n\}})^H
 \end{aligned}$$

□

**Example 12.66.** Let  $\langle X, \langle \rangle \rangle$  be a real (complex) finite dimensional inner product space then  $1_X \in \text{Hom}(X)$  is unitary.

**Proof.**  $1_X \circ (1_X)^* = (1_X)^*$  12.61  $= 1_X$  □

**Theorem 12.67.** Let  $\langle X, \langle \rangle \rangle$  be a real (complex) finite dimensional inner product space then the following are equivalent for  $L \in \text{Hom}(X)$

1.  $L$  is unitary
2.  $\forall x, y \in X$  we have  $\langle L(x), L(y) \rangle = \langle x, y \rangle$

**Proof.**

**1  $\Rightarrow$  2.** As  $L$  is unitary we have  $\forall x, y \in X$  that  $\langle x, y \rangle = \langle L^*(L(x)), y \rangle$

$$\begin{aligned} \langle x, y \rangle &= \langle 1_X(x), y \rangle \\ &= \langle (L^* \circ L)(x), y \rangle \\ &= \langle L^*(L(x)), y \rangle \\ &= \langle L(x), (L^*)^*(y) \rangle \\ &\stackrel{11.124}{=} \langle L(x), L(y) \rangle \end{aligned}$$

**2  $\Rightarrow$  1.** We have  $\forall x, y \in X$  that

$$\begin{aligned} \langle x, (L^* \circ L)(y) \rangle &= \langle x, L^*(L(y)) \rangle \\ &= \langle L(x), L(y) \rangle \\ &= \langle x, y \rangle \\ &= \langle x, 1_X(y) \rangle \end{aligned}$$

Hence using 11.100 we have that  $\forall y \in X$  we have that  $(L^* \circ L)(y) = 1_X(y)$  so that  $L^* \circ L = 1_X$ . □

**Example 12.68.** Let  $\langle X, \langle \rangle \rangle$  be a real (complex) finite dimensional inner product space then  $I: X \rightarrow X$  defined by  $x \rightarrow -x = (-1) \cdot x$  is a unitary operator.

**Proof.** First as  $I(\alpha \cdot x + \beta \cdot y) = -(\alpha \cdot x + \beta \cdot y) = \alpha \cdot (-x) + \beta \cdot (-y) = \alpha \cdot I(x) + \beta \cdot I(y)$  we have that  $I \in \text{Hom}(X)$ . Further  $\langle I(x), I(y) \rangle = \langle (-1) \cdot x, (-1) \cdot y \rangle = (-1) \cdot (-1) \cdot \langle x, y \rangle = \langle x, y \rangle$  □

**Theorem 12.69.** Let  $\langle X, \langle \rangle \rangle$  be a real (complex) finite dimensional inner product space and  $L \in \text{Hom}(X)$  a unitary operator then we have

1.  $L \circ L^* = 1_X$
2.  $L^* = L^{-1}$ ,  $L = (L^*)^{-1}$  and  $L^*$ ,  $L^{-1}$  are unitary
3.  $L$  is self-adjoint
4.  $L$  is a linear isometry (in the normed space generated by the inner product)

**Proof.**

1.  $L^* \circ L = (L \circ L^*)^* = (1_X)^*$  12.61  $= 1_X$
2. From the definition and (1) we have that  $L^* \circ L = 1_X = L \circ L^*$  proving that  $L^*$  is unitary and  $L^* = L^{-1}$ ,  $L = (L^*)^{-1}$ .
3. This is trivial as in (2) we saw that  $L^* \circ L = 1_X = L \circ L^*$
4. Using (2) we have that  $L$  is a bijection and  $\forall x \in X$  we have  $\|L(x)\| = \sqrt{\langle L(x), L(x) \rangle}$  12.67  $= \sqrt{\langle x, x \rangle} = \|x\|$  proving that  $L$  is a linear isomorphism. □

The following theorem shows that every mapping of one basis to a orthonormal basis is

**Theorem 12.70.** Let  $\langle X, \langle \rangle \rangle$  be a real (complex) finite dimensional inner product space with two basis  $\{e_i\}_{i \in \{1, \dots, n\}}$ ,  $\{f_i\}_{i \in \{1, \dots, n\}}$  then for the mappings  $U, V \in \text{Hom}(X)$  defined by  $\forall i \in \{1, \dots, n\}$   $U(e_i) = f_i$  and  $V(f_i) = e_i$  ( $U, V$  are called coordinate transformations) then we have that

1.  $U \circ V = 1_X = V \circ U$  (or equivalently  $U = V^{-1}$  and  $V = U^{-1}$ )
2.  $\mathcal{M}(V, \{e_i\}_{i \in \{1, \dots, n\}}) = \mathcal{M}(U, \{e_i\}_{i \in \{1, \dots, n\}})^{-1}$
3. If  $\{e_i\}_{i \in \{1, \dots, n\}}$ ,  $\{f_i\}_{i \in \{1, \dots, n\}}$  are orthonormal then  $U$  is unitary (see 11.130)
4. If  $\{e_i\}_{i \in \{1, \dots, n\}}$ ,  $\{f_i\}_{i \in \{1, \dots, n\}}$  are orthonormal then  $\mathcal{M}(U, \{e_i\}_{i \in \{1, \dots, n\}})^{-1} = \mathcal{M}(U, \{e_i\}_{i \in \{1, \dots, n\}})^H$  (proving that  $\mathcal{M}(U, \{e_i\}_{i \in \{1, \dots, n\}})$  is unitary)
5. If  $L \in \text{Hom}(X)$  then we have that

$$\mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}) = \mathcal{M}(U, \{e_i\}_{i \in \{1, \dots, n\}}) \cdot \mathcal{M}(L, \{f_i\}_{i \in \{1, \dots, n\}}) \cdot \mathcal{M}(U, \{e_i\}_{i \in \{1, \dots, n\}})^{-1}$$

or bringing over the matrices and applying the inverses

$$\mathcal{M}(L, \{f_i\}_{i \in \{1, \dots, n\}}) = \mathcal{M}(U, \{e_i\}_{i \in \{1, \dots, n\}})^{-1} \cdot \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}) \cdot \mathcal{M}(U, \{e_i\}_{i \in \{1, \dots, n\}})$$

**Proof.**

1. Let  $x \in X$  then  $x = \sum_{i=1}^n x_i \cdot e_i = \sum_{i=1}^n x'_i \cdot f_i$  and we have that

$$\begin{aligned} (U \circ V)(x) &= \left( U \left( V \left( \sum_{i=1}^n x'_i \cdot f_i \right) \right) \right) \\ &= \left( U \left( \sum_{i=1}^n x'_i \cdot V(f_i) \right) \right) \\ &= U \left( \sum_{i=1}^n x'_i \cdot e_i \right) \\ &= \sum_{i=1}^n x'_i \cdot U(e_i) \\ &= \sum_{i=1}^n x'_i \cdot f_i \\ &= x \\ &= 1_X(x) \end{aligned}$$

proving that  $U \circ V = 1_X$ . Like wise we have

$$\begin{aligned} (V \circ U)(x) &= \left( V \left( U \left( \sum_{i=1}^n x_i \cdot e_i \right) \right) \right) \\ &= \left( V \left( \sum_{i=1}^n x_i \cdot U(e_i) \right) \right) \\ &= V \left( \sum_{i=1}^n x_i \cdot f_i \right) \\ &= \sum_{i=1}^n x_i \cdot V(f_i) \\ &= \sum_{i=1}^n x_i \cdot e_i \\ &= x \\ &= 1_X(x) \end{aligned}$$

proving that  $V \circ U = 1_X$

2. This follows from the fact that

$$\begin{aligned}
 \mathcal{M}(U, \{e_i\}_{i \in \{1, \dots, n\}}) \cdot \mathcal{M}(V, \{e_i\}_{i \in \{1, \dots, n\}}) &\stackrel{10.307}{=} \mathcal{M}(U \circ V, \{1, \dots, n\}) \\
 &\stackrel{(1)}{=} \mathcal{M}(1_X, \{1, \dots, n\}) \\
 &= E \\
 &= \mathcal{M}(1_X, \{1, \dots, n\}) \\
 &\stackrel{(1)}{=} \mathcal{M}(V \circ U, \{1, \dots, n\}) \\
 &\stackrel{10.307}{=} \mathcal{M}(V, \{e_i\}_{i \in \{1, \dots, n\}}) \cdot \mathcal{M}(U, \{e_i\}_{i \in \{1, \dots, n\}})
 \end{aligned}$$

3. Let  $x, y \in X$  then  $x = \sum_{i=1}^n x_i \cdot e_i = \sum_{i=1}^n x'_i \cdot f_i$  and  $y = \sum_{i=1}^n y_i \cdot e_i = \sum_{i=1}^n y'_i \cdot f_i$  then we have

$$\begin{aligned}
 \langle U(x), U(y) \rangle &= \left\langle U\left(\sum_{i=1}^n x_i \cdot e_i\right), U\left(\sum_{k=1}^n y_k \cdot e_k\right) \right\rangle \\
 &= \left\langle \sum_{i=1}^n x_i \cdot U(e_i), \left(\sum_{k=1}^n y_k \cdot U(e_k)\right) \right\rangle \\
 &= \left\langle \sum_{i=1}^n x_i \cdot f_i, \left(\sum_{k=1}^n y_k \cdot f_k\right) \right\rangle \\
 &= \sum_{i=1}^n x_i \cdot \left( \left\langle f_i, \sum_{k=1}^n y_k \cdot f_k \right\rangle \right) \\
 &= \sum_{i=1}^n \left( \sum_{k=1}^n x_i \cdot \bar{y}_i \cdot \langle f_i, f_k \rangle \right) \\
 &\stackrel{\text{orthornormality } \{f_i\}_{i \in \{1, \dots, n\}}}{=} \sum_{i=1}^n \left( \sum_{k=1}^n x_i \cdot \bar{y}_i \cdot \delta_{i,k} \right) \\
 &\stackrel{\text{orthonormality of } \{e_i\}_{i \in \{1, \dots, n\}}}{=} \sum_{i=1}^n \left( \sum_{k=1}^n x_i \cdot \bar{y}_i \cdot \langle e_i, e_k \rangle \right) \\
 &= \sum_{i=1}^n x_i \cdot \left( \sum_{k=1}^n \langle e_i, y_k \cdot e_k \rangle \right) \\
 &= \sum_{i=1}^n x_i \cdot \left\langle e_i, \sum_{k=1}^n y_k \cdot e_k \right\rangle \\
 &= \left\langle \sum_{i=1}^n x_i \cdot e_i, \sum_{k=1}^n y_k \cdot e_k \right\rangle \\
 &= \langle x, y \rangle
 \end{aligned}$$

Using 12.67 we have then that  $U$  is a unitary operator.

4. This follows from

$$\begin{aligned}
 \mathcal{M}(U, \{e_i\}_{i \in \{1, \dots, n\}})^H \cdot \mathcal{M}(U, \{e_i\}_{i \in \{1, \dots, n\}}) &\stackrel{10.340}{=} \mathcal{M}(U^*, \{e_i\}_{i \in \{1, \dots, n\}}) \cdot \mathcal{M}(U, \{e_i\}_{i \in \{1, \dots, n\}}) \\
 &\stackrel{10.307}{=} \mathcal{M}(U^* \circ U, \{1, \dots, n\}) \\
 &\stackrel{(1)}{=} \mathcal{M}(1_X, \{1, \dots, n\}) \\
 &= E \\
 &= \mathcal{M}(1_X, \{1, \dots, n\})
 \end{aligned}$$

$$\begin{aligned}
&\stackrel{(1)}{=} \mathcal{M}(U \circ U^*, \{1, \dots, n\}) \\
&= \mathcal{M}(U, \{e_i\}_{i \in \{1, \dots, n\}}) \cdot \mathcal{M}(U^*, \{e_i\}_{i \in \{1, \dots, n\}}) \\
&\stackrel{10.340}{=} \mathcal{M}(U, \{e_i\}_{i \in \{1, \dots, n\}}) \cdot \mathcal{M}(U, \{e_i\}_{i \in \{1, \dots, n\}})^H
\end{aligned}$$

5. Given  $i \in \{1, \dots, n\}$  we have that

$$\begin{aligned}
L(e_i) &= L(V(f_i)) \\
&= L\left(\sum_{k=1}^n \mathcal{M}(V, \{f_i\}_{i \in \{1, \dots, n\}})_{k,j} \cdot f_k\right) \\
&= \sum_{k=1}^n \mathcal{M}(V, \{f_i\}_{i \in \{1, \dots, n\}})_{k,j} \cdot L(f_k) \\
&= \sum_{k=1}^n \left( \mathcal{M}(V, \{f_i\}_{i \in \{1, \dots, n\}})_{k,j} \cdot \sum_{l=1}^n \mathcal{M}(L, \{f_i\}_{i \in \{1, \dots, n\}})_{l,k} \cdot f_l \right) \\
&= \sum_{k=1}^n \left( \mathcal{M}(V, \{f_i\}_{i \in \{1, \dots, n\}})_{k,j} \cdot \sum_{l=1}^n \mathcal{M}(L, \{f_i\}_{i \in \{1, \dots, n\}})_{l,k} \cdot (U(e_l)) \right) \\
&= \sum_{k=1}^n \left( \mathcal{M}(V, \{f_i\}_{i \in \{1, \dots, n\}})_{k,j} \cdot \sum_{l=1}^n \mathcal{M}(L, \{f_i\}_{i \in \{1, \dots, n\}})_{l,k} \cdot \left( \sum_{r=1}^n \mathcal{M}(U, \{e_i\}_{i \in \{1, \dots, n\}})_{r,l} \cdot e_r \right) \right) \\
&= \sum_{k=1}^n \left( \mathcal{M}(V, \{f_i\}_{i \in \{1, \dots, n\}})_{k,j} \cdot \sum_{l=1}^n \left( \sum_{r=1}^n \mathcal{M}(L, \{f_i\}_{i \in \{1, \dots, n\}})_{l,k} \cdot \mathcal{M}(U, \{e_i\}_{i \in \{1, \dots, n\}})_{r,l} \cdot e_r \right) \right) \\
&\stackrel{10.47}{=} \sum_{k=1}^n \left( \mathcal{M}(V, \{f_i\}_{i \in \{1, \dots, n\}})_{k,j} \cdot \sum_{r=1}^n \left( \sum_{l=1}^n \mathcal{M}(L, \{f_i\}_{i \in \{1, \dots, n\}})_{l,k} \cdot \mathcal{M}(U, \{e_i\}_{i \in \{1, \dots, n\}})_{r,l} \cdot e_r \right) \right) \\
&= \sum_{k=1}^n \left( \mathcal{M}(V, \{f_i\}_{i \in \{1, \dots, n\}})_{k,j} \cdot \sum_{r=1}^n (\mathcal{M}(U, \{e_i\}_{i \in \{1, \dots, n\}}) \cdot \mathcal{M}(L, \{f_i\}_{i \in \{1, \dots, n\}}))_{r,k} \cdot e_r \right) \\
&= \sum_{k=1}^n \left( \sum_{r=1}^n \mathcal{M}(V, \{f_i\}_{i \in \{1, \dots, n\}})_{k,j} \cdot (\mathcal{M}(U, \{e_i\}_{i \in \{1, \dots, n\}}) \cdot \mathcal{M}(L, \{f_i\}_{i \in \{1, \dots, n\}}))_{r,k} \cdot e_r \right) \\
&\stackrel{10.47}{=} \sum_{r=1}^n \left( \sum_{k=1}^n \mathcal{M}(V, \{f_i\}_{i \in \{1, \dots, n\}})_{k,j} \cdot (\mathcal{M}(U, \{e_i\}_{i \in \{1, \dots, n\}}) \cdot \mathcal{M}(L, \{f_i\}_{i \in \{1, \dots, n\}}))_{r,k} \cdot e_r \right) \\
&= \sum_{r=1}^n (\mathcal{M}(U, \{e_i\}_{i \in \{1, \dots, n\}}) \cdot \mathcal{M}(L, \{f_i\}_{i \in \{1, \dots, n\}}) \cdot \mathcal{M}(V, \{f_i\}_{i \in \{1, \dots, n\}}))_{r,j} \cdot e_r \\
&\stackrel{10.314}{=} \sum_{r=1}^n (\mathcal{M}(U, \{e_i\}_{i \in \{1, \dots, n\}}) \cdot \mathcal{M}(L, \{f_i\}_{i \in \{1, \dots, n\}}) \cdot \mathcal{M}(U, \{\text{fe}_i\}_{i \in \{1, \dots, n\}})^{-1})_{r,j} \cdot e_r
\end{aligned}$$

hence we have by definition of the matrix of a linear function proved that

$$\mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}) = \mathcal{M}(U, \{e_i\}_{i \in \{1, \dots, n\}}) + \mathcal{M}(L, \{f_i\}_{i \in \{1, \dots, n\}}) + \mathcal{M}(U, \{e_i\}_{i \in \{1, \dots, n\}})^{-1} \quad (12.43) \quad \square$$

**Definition 12.71. (Normal operator)** Let  $\langle X, \langle \cdot, \cdot \rangle \rangle$  be a real (complex) finite dimensional inner product space then  $L \in \text{Hom}(X)$  is **normal** if  $L \circ L^* = L^* \circ L$

**Lemma 12.72.** Let  $\langle X, \langle \cdot, \cdot \rangle \rangle$  be a real (complex) finite dimensional inner product space then every self-adjoint operator is normal. Hence every unitary operator is normal (see 12.69).

**Proof.** This is trivial as  $L \circ L^* = L \circ L = L^* \circ L$   $\square$

**Theorem 12.73.** Let  $\langle X, \langle \cdot, \cdot \rangle \rangle$  be a complex inner product vector space,  $L \in \text{Hom}(X)$  then if  $\forall x \in X$  we have  $\langle L(x), x \rangle = 0$  we have  $L = 0$

**Proof.** Let  $x, y \in X$  then we have

$$\begin{aligned} \langle L(x+y), x+y \rangle - \langle L(x-y), x-y \rangle &= \langle L(x) + L(y), x+y \rangle - \langle L(x) - L(y), x-y \rangle \\ &= \langle L(x), x \rangle + \langle L(x), y \rangle + \langle L(y), x \rangle + \langle L(y), y \rangle - \langle L(x), x \rangle + \\ &\quad \langle L(x), y \rangle + \langle L(y), x \rangle - \langle L(y), y \rangle \\ &= 2 \cdot \langle L(x), y \rangle + 2 \cdot \langle L(y), x \rangle \end{aligned}$$

and

$$\begin{aligned} \langle L(x+i \cdot y), x+i \cdot y \rangle - \langle L(x-i \cdot y), x-i \cdot y \rangle &= \langle L(x) + i \cdot L(y), x+i \cdot y \rangle - \langle L(x) - i \cdot L(y), x-i \cdot y \rangle \\ &= \langle L(x), x \rangle + \langle L(x), i \cdot y \rangle + \langle i \cdot L(y), x \rangle + \langle i \cdot L(y), i \cdot y \rangle - \\ &\quad \langle L(x), x \rangle + \langle L(x), i \cdot y \rangle + \langle i \cdot L(y), x \rangle - \langle i \cdot L(y), i \cdot y \rangle \\ &= \langle L(x), i \cdot y \rangle + \langle i \cdot L(y), x \rangle + \langle L(x), i \cdot y \rangle + \langle i \cdot L(y), x \rangle - \\ &\quad x \rangle \\ &= -2 \cdot i \cdot \langle L(x), y \rangle + 2 \cdot i \cdot \langle L(y), x \rangle \end{aligned}$$

Hence we have

$$\begin{aligned} (\langle L(x+y), x+y \rangle - \langle L(x-y), x-y \rangle) + i \cdot \\ (\langle L(x+i \cdot y), x+i \cdot y \rangle - \langle L(x-i \cdot y), x-i \cdot y \rangle) &= 2 \cdot \langle L(x), y \rangle + 2 \cdot \langle L(y), x \rangle + 2 \cdot \langle L(x), y \rangle - 2 \cdot \\ &\quad \langle L(y), x \rangle \\ &= 4 \cdot \langle L(x), y \rangle \end{aligned}$$

proving that

$$\begin{aligned} \langle L(x), y \rangle &= \frac{\langle L(x+i \cdot y), x+i \cdot y \rangle - \langle L(x-i \cdot y), x-i \cdot y \rangle}{4} + i \cdot \frac{\langle L(x+i \cdot y), x+i \cdot y \rangle - \langle L(x-i \cdot y), x-i \cdot y \rangle}{4} \\ &= \frac{0-0}{4} + i \cdot \frac{0-0}{4} \\ &= 0 \end{aligned} \quad .$$

Hence if we take  $y = L(x)$  we have that  $\|L(x)\| = \langle L(x), L(x) \rangle = 0$  so that  $L(x) = 0$  and as  $x$  is choosen arbitray we have  $L = 0$ .  $\square$

**Theorem 12.74.** Let  $\langle X, \langle \cdot, \cdot \rangle \rangle$  be a real (complex) finite dimensional inner product space and  $L \in \text{Hom}(X)$  a self-adjoint operator then  $L = 0$  if and only if  $\forall x \in X$  we have  $\langle L(x), x \rangle = 0$

**Proof.** If  $\langle X, \langle \rangle \rangle$  is a complex space the theorem is proved by the previous theorem (see 12.73) (even with assumption of self-adjointness) so we must only prove the case of a real inner product space. Let  $x, y \in X$  then

$$\begin{aligned}
 \langle L(x+y), x+y \rangle - \langle L(x-y), x-y \rangle &= \langle L(x) + L(y), x+y \rangle - \langle L(x) - L(y), x-y \rangle \\
 &= \langle L(x), x \rangle + \langle L(x), y \rangle + \langle L(y), x \rangle + \langle L(y), y \rangle - (\langle L(x), x \rangle - \langle L(x), y \rangle - \langle L(y), x \rangle + \langle L(y), y \rangle) \\
 &= \langle L(x), x \rangle + \langle L(x), y \rangle + \langle L(y), x \rangle + \langle L(y), y \rangle - \langle L(x), x \rangle + \langle L(x), y \rangle + \langle L(y), x \rangle - \langle L(y), y \rangle \\
 &= \langle L(x), y \rangle + \langle L(y), x \rangle + \langle L(x), y \rangle + \langle L(y), x \rangle \\
 &= 2 \cdot \langle L(x), y \rangle + 2 \cdot \langle L(y), x \rangle \\
 &\stackrel{X \text{ is real}}{=} 2 \cdot \langle L(x), y \rangle + 2 \cdot \langle x, L(y) \rangle \\
 &= 2 \cdot \langle L(x), y \rangle + 2 \cdot \langle L^*(x), y \rangle \\
 &\stackrel{L=L^*}{=} 2 \cdot \langle L(x), y \rangle + 2 \cdot \langle L(x), y \rangle \\
 &= 4 \cdot \langle L(x), y \rangle
 \end{aligned}$$

hence  $\langle L(x), y \rangle = \frac{\langle L(x+y), x+y \rangle - \langle L(x-y), x-y \rangle}{4} = \frac{0-0}{4} = 0$ . So taking  $y=L(x)$  we have that  $\|L(x)\| = \langle L(x), L(x) \rangle = 0$  proving as  $x$  is choosen arbitrary that  $L=0$ .  $\square$

**Lemma 12.75.** Let  $\langle X, \langle \rangle \rangle$  be a real (complex) finite dimensional inner product space together with the inner product norm  $\|\cdot\|$  (see 11.105) then  $L \in \text{Hom}(X)$  is normal if and only if  $\forall x \in X$  we have  $\|L(x)\| = \|L^*(x)\|$

**Proof.** First we have

$$\begin{aligned}
 (L \circ L^* + L^* \circ L)^* &\stackrel{11.125}{=} (L \circ L^*)^* + (L^* \circ L)^* \\
 &\stackrel{11.126}{=} L^* \circ (L^*)^* + (L^*)^* \circ L^* \\
 &\stackrel{11.124}{=} L^* \circ L + L \circ L^* \\
 &= L \circ L^* + L^* \circ L
 \end{aligned}$$

proving that

$$L \circ L^* + L^* \circ L \text{ is self adjoint} \quad (12.44)$$

$$\begin{aligned}
 L \text{ is normal} &\Leftrightarrow L \circ L^* = L^* \circ L \\
 &\Leftrightarrow L \circ L^* - L^* \circ L = 0 \\
 &\stackrel{12.44 \text{ and } 12.74}{\Leftrightarrow} \forall x \in X \models \langle (L \circ L^* - L^* \circ L)(x), x \rangle = 0 \\
 &\Leftrightarrow \forall x \in X \models \langle (L \circ L^*(x))x, x \rangle = \langle (L^* \circ L)(x), x \rangle \\
 &\Leftrightarrow \forall x \in X \models \langle L(L^*(x)), x \rangle = \langle L^*(L(x)), x \rangle \\
 &\Leftrightarrow \forall x \in X \models \langle L^*(x), L^*(x) \rangle = \langle L(x), (L^*)^*(x) \rangle \\
 &\stackrel{11.124}{\Leftrightarrow} \forall x \in X \models \langle L^*(x), L^*(x) \rangle = \langle L(x), L(x) \rangle \\
 &\Leftrightarrow \forall x \in X \models \|L^*(x)\| = \|L(x)\|
 \end{aligned}$$

proving the theorem.  $\square$

**Theorem 12.76.** Let  $\langle X, \langle \rangle \rangle$  be a real (complex) finite dimensional inner product space with  $\dim(X) = n \in \mathbb{N}_0$  and  $L \in \text{Hom}(X)$  then the following are equivalent

1. There exists a orthonormal basis  $\{e_i\}_{i \in \{1, \dots, n\}}$  of  $X$  such that  $\mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}})$  is diagonal
2.  $X$  has a orthonormal basis  $\{e_i\}_{i \in \{1, \dots, n\}}$  where  $\forall i \in \{1, \dots, n\}$   $e_i$  is a eigen vector of  $L$

**Proof.**

**1  $\Rightarrow$  2.** As there exists a basis  $\{e_i\}_{i \in \{1, \dots, n\}}$  such that  $M = \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}})$  is diagonal we have that  $\forall i \in [1, \dots, n]$  that  $L(e_i) = \sum_{k=1}^n M_{k,i} \cdot e_k = M_{i,i} \cdot e_i \Rightarrow e_i$  is a eigenvector with eigen value  $M_{i,i}$ .

**2  $\Rightarrow$  1.** As  $\{e_i\}_{i \in \{1, \dots, n\}}$  is a set of eigen vectors with eigen values  $\lambda_i$  we have if we define  $M$  by  $M_{i,j} = \delta_{i,j} \cdot \lambda_i$  so that  $L(e_i) = \lambda_i \cdot e_i = \sum_{k=1}^n \delta_{k,i} \cdot \lambda_i \cdot e_k = \sum_{k=1}^n M_{k,j} \cdot e_k$  proving that  $\mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}}) = M$  which is by construction diagonal.  $\square$

**Theorem 12.77. (Complex Spectral Theorem)** Let  $\langle X, \langle \rangle \rangle$  be a complex finite dimensional inner product space with  $\dim(X) = n \in \mathbb{N}_0$  and  $L \in \text{Hom}(X)$  then the following are equivalent

1.  $L$  is normal
2. There exists a orthonormal basis  $\{e_i\}_{i \in \{1, \dots, n\}}$  of  $X$  such that  $\mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}})$  is diagonal.
3.  $X$  has a orthonormal basis  $\{e_i\}_{i \in \{1, \dots, n\}}$  where  $\forall i \in \{1, \dots, n\}$   $e_i$  is a eigen vector of  $L$

**Proof.**

**1  $\Rightarrow$  2.** As  $L$  is normal we have using Schur's theorem (see 12.57) there exists a orthonormal basis  $\{e_i\}_{i \in \{1, \dots, n\}}$  of  $X$  such that  $M = \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}})$  is a upper triangluar matrix. Define now  $\mathcal{S} = \{m \in \mathbb{N}_0 \mid \text{If } m \leq n \text{ then } \forall i \in \{1, \dots, m\} \wedge \forall j \in \{i+1, \dots, n\} \text{ we have that } M_{i,j} = 0\}$  then we have

**1  $\in \mathcal{S}$ .** As  $M$  is upper triangular we have that  $L(e_1) = \sum_{k=1}^n M_{k,1} \cdot e_k = M_{1,1} \cdot e_1$  so that  $\|L(e_1)\|^2 = \langle M_{1,1} \cdot e_1, M_{1,1} \cdot e_1 \rangle = M_{1,1} \cdot \overline{M_{1,1}} \cdot \langle e_1, e_1 \rangle = |M_{1,1}|^2$  proving that

$$\|L(e_1)\|^2 = |M_{1,1}|^2 \quad (12.45)$$

Further if  $M^* = \mathcal{M}(L^*, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}})$  then using 11.132 we have that  $M_{i,j} = \overline{M_{j,i}}$  so that

$$\begin{aligned} \|L^*(e_1)\|^2 &= \langle L(e_1), L(e_1) \rangle \\ &= \left\langle \sum_{k=1}^n M_{k,1}^* \cdot e_k, \sum_{l=1}^n M_{l,1}^* \cdot e_l \right\rangle \\ &= \sum_{k=1}^n M_{k,1}^* \cdot \left\langle e_k, \sum_{l=1}^n M_{l,1}^* \cdot e_l \right\rangle \\ &= \sum_{k=1}^n M_{k,1}^* \cdot \left( \sum_{l=1}^n M_{l,1}^* \cdot \langle e_k, e_l \rangle \right) \\ &= \sum_{k=1}^n M_{k,1}^* \cdot \left( \sum_{l=1}^n M_{l,1}^* \cdot \delta_{k,l} \right) \\ &= \sum_{k=1}^n M_{k,1}^* \cdot \overline{M_{k,1}} \\ &= \sum_{k=1}^n \overline{M_{1,k}} \cdot \overline{\overline{M_{1,k}}} \\ &= \sum_{k=1}^n \overline{M_{1,k}} \cdot M_{1,k} \\ &= \sum_{k=1}^n |M_{1,k}|^2 \end{aligned}$$

proving that

$$\|L^*(e_1)\|^2 = \sum_{k=1}^n |M_{1,k}|^2 \quad (12.46)$$

As by 12.75 we have  $\|L(e_1)\| = \|L^*(e_1)\| \stackrel{12.45 \text{ and } 12.46}{\Rightarrow} |M_{1,1}|^2 = |M_{1,1}|^2 + \sum_{k=2}^n |M_{1,k}|^2$  proving that  $0 = \sum_{k=2}^n |M_{1,k}|^2$  so that  $\forall k \in \{2, \dots, n\}$  we have that  $|M_{1,k}| = 0$  or  $\forall i \in \{1\}$  we have  $\forall j \in \{i+1, \dots, n\}$  that  $M_{1,j} = 0$ . Hence  $1 \in \mathcal{S}$ .

$m \in \mathcal{S} \Rightarrow m+1 \in \mathcal{S}$ . If  $n < m+1$  then  $m+1 \in \mathcal{S}$  so we have to consider the case that  $m+1 \leq n$ . As  $m \in \mathcal{S}$  we have

$$\forall i \in \{1, \dots, m\} \text{ and } \forall j \in \{i+1, \dots, n\} \text{ we have } M_{i,j} = 0 \quad (12.47)$$

Now  $L(e_{m+1}) = \sum_{k=1}^n M_{k,m+1} \cdot e_k$   $M$  is upper triangular  $\sum_{k=1}^{m+1} M_{k,m+1} \cdot e_k = M_{m+1,m+1} \cdot e_{m+1} + \sum_{k=1}^m M_{k,m+1} \cdot e_k \stackrel{12.47}{=} M_{m+1,m+1} \cdot e_{m+1}$  hence  $\|L(e_{m+1})\|^2 = \langle L(e_{m+1}), L(e_{m+1}) \rangle = \langle M_{m+1,m+1} \cdot e_{m+1}, M_{m+1} \rangle = M_{m+1,m+1} \cdot \overline{M_{m+1,m+1}} \cdot \langle e_{m+1}, e_{m+1} \rangle$ . So we have

$$\|L(e_{m+1})\|^2 = |M_{m+1,m+1}|^2 \quad (12.48)$$

Further if  $M^* = \mathcal{M}(L^*, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}})$  then using 11.132 we have that  $M_{i,j} = \overline{M_{j,i}}$  so that

$$\begin{aligned} \|L^*(e_{m+1})\|^2 &= \langle L(e_{m+1}), L(e_{m+1}) \rangle \\ &= \left\langle \sum_{k=1}^n M_{k,m+1}^* \cdot e_k, \sum_{l=1}^n M_{l,m+1}^* \cdot e_l \right\rangle \\ &= \sum_{k=1}^n M_{k,m+1}^* \cdot \left\langle e_k, \sum_{l=1}^n M_{k,m+1}^* \cdot e_l \right\rangle \\ &= \sum_{k=1}^n M_{k,m+1}^* \cdot \left( \sum_{l=1}^n M_{k,m+1}^* \cdot \langle e_k, e_l \rangle \right) \\ &= \sum_{k=1}^n M_{k,m+1}^* \cdot \left( \sum_{l=1}^n M_{k,m+1}^* \cdot \delta_{k,l} \right) \\ &= \sum_{k=1}^n M_{k,m+1}^* \cdot \overline{M_{k,m+1}} \\ &= \sum_{k=1}^n \overline{M_{m+1,k}} \cdot \overline{M_{m+1,k}} \\ &= \sum_{k=1}^n \overline{M_{m+1,k}} \cdot M_{m+1,k} \\ &= \sum_{k=1}^n |M_{m+1,k}|^2 \end{aligned}$$

proving that

$$\|L^*(e_{m+1})\|^2 = \sum_{k=1}^n |M_{m+1,k}|^2 \quad (12.49)$$

As by 12.75 we have  $\|L(e_{m+1})\| = \|L^*(e_{m+1})\| \stackrel{12.48 \text{ and } 12.49}{\Rightarrow} |M_{m+1,m+1}|^2 = |M_{m+1,m+1}|^2 + \sum_{k \in \{1, \dots, n\} \setminus \{m+1\}} |M_{m+1,k}|^2$  proving that  $0 = \sum_{k \in \{1, \dots, n\} \setminus \{m+1\}} |M_{m+1,k}|^2$  so that  $\forall k \in \{1, \dots, n\} \setminus \{m+1\}$  we have that  $|M_{m+1,k}| = 0$ . Hence  $\forall j \in \{(m+1)+1, \dots, n\}$  we have  $M_{m+1,j} = 0$  and using 12.47 we conclude that  $\forall i \in \{1, \dots, m+1\}$  we have  $\forall j \in \{i+1, \dots, n\}$  we have  $M_{i,j} = 0$ .

Mathematical induction proves that

$$\mathcal{S} = \mathbb{N}_0 \quad (12.50)$$

Hence if  $i, j \in \{1, \dots, n\}$  with  $i \neq j$  we have either

$i < j$ . As  $i \in \mathbb{N}_0 = \mathcal{S}$  we have as  $i \leq n$  and  $j \in \{i+1, \dots, n\}$  that  $M_{i,j} = 0$

$j < i$ . As  $M$  is upper diagonal we have  $M_{i,j} = 0$

proving that  $M = \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}})$  is diagonal.

**2  $\Rightarrow$  1.** As there exists a basis  $\{e_i\}_{i \in \{1, \dots, n\}}$  such that  $M = \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}})$  is diagonal and  $M^* = \mathcal{M}(L^*, \{e_i\}_{i \in \{1, \dots, n\}}, \{e\})$  is defined by  $M_{i,j}^* = \overline{M_{j,i}}$  (see 10.340) we have that  $M^*$  is also diagonal. The we have  $\forall j \in \{1, \dots, n\}$  that

$$\begin{aligned} (L \circ L^*)(e_j) &= L(L^*(e_j)) \\ &= L\left(\sum_{k=1}^n M_{k,j}^* \cdot e_k\right) \\ &= L(M_{j,j}^* \cdot e_j) \\ &= M_{j,j}^* \cdot L(e_j) \\ &= M_{j,j}^* \cdot \left(\sum_{k=1}^n M_{k,j} \cdot e_k\right) \\ &= M_{j,j}^* \cdot M_{j,j} \cdot e_j \\ &= M_{i,j} \cdot M_{i,j}^* \cdot e_j \\ &= M_{j,j} \cdot \sum_{k=1}^n M_{k,j}^* \cdot e_k \\ &= M_{j,j} \cdot L^*(e_j) \\ &= L^*(M_{j,j} \cdot e_j) \\ &= L^*\left(\sum_{k=1}^n M_{k,j} \cdot e_k\right) \\ &= L^*(L(e_j)) \end{aligned}$$

proving that

$$\forall j \in \{1, \dots, n\} \text{ we have } (L \circ L^*)(e_j) = (L^* \circ L)(e_j) \quad (12.51)$$

Hence if  $x \in X$  we have that  $x = \sum_{i=1}^n x_i \cdot e_i$  and thus

$$\begin{aligned} (L \circ L^*)(x) &= (L \circ L^*)\left(\sum_{i=1}^n x_i \cdot e_i\right) \\ &= \sum_{i=1}^n x_i \cdot (L \circ L^*)(e_i) \\ &= \sum_{i=1}^n x_i \cdot (L^* \circ L)(e_i) \\ &= (L^* \circ L)\left(\sum_{i=1}^n x_i \cdot e_i\right) \\ &= (L^* \circ L)(x) \end{aligned}$$

proving that  $L \circ L^* = L^* \circ L$  and thus that  $L$  is normal.

**2  $\Leftrightarrow$  3.** This follows from 12.76.  $\square$

**Corollary 12.78.** Let  $\langle X, \langle \rangle \rangle$  be a complex finite dimensional inner product space with  $\dim(X) = n \in \mathbb{N}_0$ ,  $\{e_i\}_{i \in \{1, \dots, n\}}$  a orthonormal base,  $L \in \text{Hom}(X)$  a normal operator then there exists a unary operator  $U \in \text{Hom}(X)$  such that  $\mathcal{M}(U^* \circ L \circ U, \{e_i\}_{i \in \{1, \dots, n\}})$  is diagonal.

**Proof.** Using the Complex spectral theorem 12.77 there exists a basis  $\{f_i\}_{i \in \{1, \dots, n\}}$  such that  $\mathcal{M}(L, \{f_i\}_{i \in \{1, \dots, n\}})$  is diagonal. Take now the unitary mapping  $U \in \text{Hom}(X)$  defined by  $\forall i \in \{1, \dots, n\} U(e_i) = f_i$  (see 12.70) then we have  $\forall i, j \in \{1, \dots, n\}$  that

$$\begin{aligned}
 \mathcal{M}(U^* \circ L \circ U, \{e_i\}_{i \in \{1, \dots, n\}})_{i,j} &\stackrel{11.118}{=} \langle (U^* \circ L \circ U)(e_j), e_i \rangle \\
 &= \langle U^*(L(U(e_j))), e_i \rangle \\
 &= \langle U^*(L(f_j)), e_i \rangle \\
 &= \langle L(f_j), (U^*)^*(e_i) \rangle \\
 &= \langle L(f_j), U(e_i) \rangle \\
 &= \langle L(f_j), f_i \rangle \\
 &\stackrel{11.118}{=} \mathcal{M}(L, \{f_i\}_{i \in \{1, \dots, n\}})_{i,j} \\
 &= \mathcal{M}(L, \{f_i\}_{i \in \{1, \dots, n\}})_{i,i} \cdot \delta_{i,j}
 \end{aligned}$$

□

We need the following theorem to prove the spectral theorem in the real case which is the analogue to the conditions for quadratic equations to be solvable.

**Theorem 12.79.** *Let  $\langle X, \langle \rangle \rangle$  be a finite dimensional real (complex) inner product space then if  $L \in \text{Hom}(X)$  is a self-adjoint linear operator and  $b, c \in \mathbb{R}$  such that  $b^2 < 4 \cdot c$  then  $L^2 + b \cdot L + c \cdot 1_X$  is a isomorphism.*

**Proof.** As by 12.40 we have that  $L^2 + b \cdot L + c \cdot 1_X$  is linear so that using 12.35 we must only prove that  $L^2 + b \cdot L + c \cdot 1_X$  is injective. Now  $\forall x \in X$  with  $x \neq 0$  we have using the Cauchy Schwarz inequality (see 11.107) that  $b \cdot \langle L(x), x \rangle = \langle b \cdot L(x), x \rangle \leq \|b \cdot L(x)\| \cdot \|x\| = |b| \cdot \|L(x)\| \cdot \|x\|$  giving by multiplying by  $-1$  that

$$-|b| \cdot \|L(x)\| \cdot \|x\| \leq b \cdot \langle L(x), x \rangle \quad (12.52)$$

$$\begin{aligned}
 \langle (L^2 + b \cdot L + c \cdot 1_X)(x), x \rangle &= \langle L^2(x) + b \cdot L(x) + c \cdot x, x \rangle \\
 &= \langle L^2(x), x \rangle + b \cdot \langle L(x), x \rangle + c \cdot \langle x, x \rangle \\
 &\stackrel{T \text{ is self-adjoint}}{=} \langle L(x), L(x) \rangle + b \cdot \langle L(x), x \rangle + c \cdot \langle x, x \rangle \\
 &= \|L(x)\|^2 + b \cdot \langle L(x), x \rangle + c \cdot \|x\|^2 \\
 &\stackrel{12.52}{\geq} \|L(x)\|^2 - |b| \cdot \|L(x)\| \cdot \|x\| + c \cdot \|x\|^2
 \end{aligned}$$

proving

$$\langle (L^2 + b \cdot L + c \cdot 1_X)(x), x \rangle \geq \|L(x)\|^2 - |b| \cdot \|L(x)\| \cdot \|x\| + c \cdot \|x\|^2 \quad (12.53)$$

Next

$$\begin{aligned}
 \left( \|L(x)\| - \frac{|b| \cdot \|x\|}{2} \right)^2 + \left( c - \frac{b^2}{4} \right) \cdot \|x\|^2 &= \|L(x)\|^2 - |b| \cdot \|L(x)\| \cdot \|x\| + \frac{|b|^2 \cdot \|x\|^2}{4} + c \cdot \|x\|^2 - \frac{b^2 \cdot \|x\|^2}{4} \\
 &\stackrel{b^2 = |b|^2}{=} \|L(x)\|^2 - |b| \cdot \|L(x)\| \cdot \|x\| + c \cdot \|x\|^2 \\
 &\stackrel{14.217}{\leq} \langle (L^2 + b \cdot L + c \cdot 1_X)(x), x \rangle
 \end{aligned}$$

Now as  $0 \leq \left( \|L(x)\| - \frac{|b| \cdot \|x\|}{2} \right)^2$  and  $b^2 < 4 \cdot c \Rightarrow 0 < c - \frac{b^2}{4} \Rightarrow 0 < \left( c - \frac{b^2}{4} \right) \cdot \|x\|^2$  we have by the above that

$$0 < \langle (L^2 + b \cdot L + c \cdot 1_X)(x), x \rangle \quad (12.54)$$

From the above we have then that  $\forall x \in X \setminus \{0\} (L^2 + b \cdot L + c \cdot 1_X)(x) \neq 0$  hence  $\ker(L^2 + b \cdot L + c \cdot 1_X) = \{0\}$  and thus by 10.213 we have that  $L^2 + b \cdot L + c \cdot 1_X$  is injective. □

We are now ready to prove that self adjoint operators have a eigenvector with a eigenvalue.

**Theorem 12.80.** *Let  $\langle X, \langle \rangle \rangle$  be a non trivial real (complex) finite dimensional inner product space and  $L \in \text{Hom}(X)$  is a self adjoint operator then  $L$  has a eigenvector with a eigenvalue which is real by 12.62.*

**Proof.** First the theorem is already proved in the complex case (see 12.51) so we have only to consider the real case. Let  $n = \dim(X) \in \mathbb{N}_0$  and choose  $x \in X$  with  $x \neq 0$ . Then as  $\{L^i(x)\}_{i \in \{0, \dots, n\}}$  consists of  $n+1$  vectors the family is linear dependent and thus there exists a  $\{a_i\}_{i \in \{1, \dots, n\}} \subseteq \mathbb{R}$  such that  $\sum_{i=0}^n a_i \cdot L^i(x) = 0$  with  $\{i \in \{0, \dots, n\} \mid a_i \neq 0\} \neq \emptyset$  so  $N = \max(\{i \in \{1, \dots, n\}\})$  exists and we have

$$\sum_{i=0}^N a_i \cdot L^i(x) = 0 \text{ with } a_N \neq 0 \text{ where } N \in \{0, \dots, n\} \quad (12.55)$$

If  $N = 0$  then  $0 = \sum_{i=0}^0 a_i \cdot L^i(x) = a_0 \cdot L^0(x) = a_0 \cdot x \xrightarrow{a_0 \neq x} x = 0$  a contradiction, so we may assume that  $N \neq 0$ .

$$N \neq 0 \quad (12.56)$$

Hence the polynomial  $p$  with real coefficients defined by  $p(z) = \sum_{i=0}^N a_i \cdot z^i$  is a non constant polynomial. Using 12.33 there exists a  $m, M \in \mathbb{N}$ ,  $\{\lambda_i\}_{i \in \{1, \dots, m\}} \subseteq \mathbb{R}$ ,  $\{b_i\}_{i \in \{1, \dots, M\}}$ ,  $\{c_i\}_{i \in \{1, \dots, M\}} \subseteq \mathbb{R}$  so that  $p(z) = (\prod_{i=1}^m (z - \lambda_i)) \cdot (\prod_{i=1}^M (z^2 + b_i \cdot z + c_i))$  and  $\forall i \in \{1, \dots, M\}$  we have  $b_i^2 < 4 \cdot c_i$ . Define now  $\{p_i\}_{i \in \{1, \dots, m+M\}}$  by  $p_i(z) = \begin{cases} (z - \lambda_i) & \text{if } i \in \{1, \dots, m\} \\ (z^2 + b_{i-m} \cdot z + c_{i-m}) & \text{if } i \in \{m+1, \dots, m+M\} \end{cases}$  then  $\forall z \in \mathbb{C}$  we have

$$\begin{aligned} \prod_{i \in \{1, \dots, m+M\}} p_i(z) &= \prod_{i \in \{1, \dots, m+M\}} p_i(z) \\ &= \left( \prod_{i \in \{1, \dots, m\}} p_i(z) \right) \cdot \left( \prod_{i \in \{m+1, \dots, m+M\}} p_i(z) \right) \\ &= \left( \prod_{i=1}^m (z - \lambda_i) \right) \cdot \left( \prod_{i=m+1}^{m+M} z^2 + b_{i-m} \cdot z + c_{i-m} \right) \\ &= \left( \prod_{i=1}^m (z - \lambda_i) \right) \cdot \left( \prod_{i=1}^M z^2 + b_i \cdot z + c_i \right) \\ &= p(z) \end{aligned}$$

to summarize we have

$$\forall z \in \mathbb{C} \text{ we have } p(z) = \prod_{i \in \{1, \dots, m+M\}} p_i(z) \quad (12.57)$$

Using 12.43 we have then that

$$\sum_{i=0}^N a_i \cdot L^i = p(L) = \prod_{i=1}^{m+M} p_i(L) \quad (12.58)$$

Now as by 12.55 we have that  $p(L)(x) = \sum_{i=0}^N a_i \cdot L^i(x) = 0$  hence as  $x \neq 0$  we have that  $p(L)$  is not regular. Using 12.58 together with 12.48 proves that there exists a  $i \in \{1, \dots, m+M\}$  such that  $p_i(L)$  is not regular. Now if  $i \in \{m+1, \dots, M+1\}$  we have  $p_i(L) = L^2 + a_i \cdot L + c_i$  which as  $b_i^2 < 4 \cdot c_i$  is regular by 12.79, so we must have that  $i \in \{1, \dots, m\}$  hence  $p_i(L) = L - \lambda_i \cdot 1_X$  is not regular, so there exists a  $y \in X$  with  $y \neq 0$  such that  $L(y) - \lambda_i \cdot 1_X(y) = 0 \Rightarrow L(y) = \lambda_i \cdot y$ . So we have found our eigenvector with eigenvalue  $\lambda$ .  $\square$

**Definition 12.81.** Let  $\langle X, \langle \cdot, \cdot \rangle \rangle$  be a real (complex) inner product vector space and  $U \subseteq X$  then  $U^\perp = \{x \in X \mid \forall y \in U \text{ we have } \langle x, y \rangle = 0\}$ .  $U$  is called the orthogonal complement of  $U$ .

**Theorem 12.82.** Let  $\langle X, \langle \cdot, \cdot \rangle \rangle$  be a real (complex) inner vector space then

1. If  $U \subseteq X$  then  $U^\perp$  is a sub space of  $X$
2.  $\{0\}^\perp = X$
3.  $X^\perp = \{0\}$
4. If  $U \subseteq X$  then  $U \cap U^\perp \subseteq \{0\}$

5. If  $U \subseteq X$  is a subspace then  $U \cap U^\perp = \{0\}$

6. If  $U, W \subseteq X$  with  $U \subseteq W$  then  $W^\perp \subseteq U^\perp$

**Proof.**

1. If  $x, y \in U^\perp$  and  $\alpha, \beta \in \mathbb{R}(\mathbb{C})$  then if  $u \in U$  we have  $\langle \alpha \cdot x + \beta \cdot y, u \rangle = \alpha \cdot \langle x, u \rangle + \beta \cdot \langle y, u \rangle = 0$  so  $U^\perp$  is a sub space of  $X$
2. As  $\forall x \in X$  we have  $\langle X, 0 \rangle = 0$  we have that  $\{0\}^\perp = X$
3. First  $\forall x \in X$  we have  $\langle 0, x \rangle = 0$  we have  $0 \in X^\perp$ , if  $y \in X^\perp$  then  $\forall x \in X$  we have  $\langle y, x \rangle = 0$  hence  $\langle y, y \rangle = 0$  proving that  $y = 0$  so that  $X^\perp = \{0\}$
4. If  $x \in U \cap U^\perp$  then  $x \in U$  and  $x \in U^\perp$  so that  $\langle x, x \rangle = 0$  proving that  $x = 0$  so that  $U \cap U^\perp \subseteq \{0\}$
5. If  $U$  is a subspace then  $0 \in U$ , as by (1)  $U^\perp$  is also a sub space we have that  $0 \in U^\perp$  so using (4) we have that  $0 \in U \cap U^\perp \subseteq \{0\}$  proving that  $U \cap U^\perp = \{0\}$
6. If  $x \in W^\perp$  then  $\forall y \in W$  we have  $y \in W$  so that  $\langle x, y \rangle = 0$  hence  $x \in U^\perp$  □

**Theorem 12.83.** Let  $\langle X, \langle \cdot, \cdot \rangle \rangle$  be a real(complex) inner product space,  $U$  a finite dimensional sub space of  $X$  then  $X = U \oplus U^\perp$

**Proof.** We have the following cases to consider for  $U$

**$U$  is trivial.** then  $U = \{0\}$  we have by the previous theorem that  $U^\perp = X$  so that  $U \oplus U^\perp = \{0\} \oplus X = X$

**$U$  is not trivial.** then as  $U$  is dinite dimensional there exists a orthonormal basis  $\{e_i\}_{i \in \{1, \dots, n\}}$  for  $U$  (see 11.120). Let  $x \in X$  then we have that  $x = x + y$  where  $x = \sum_{i=1}^n \langle x, e_i \rangle \cdot e_i$  and  $y = x - \sum_{i=1}^n \langle x, e_i \rangle \cdot e_i$ . If now  $u \in U$  then we have that  $u = \sum_{i=1}^n u \cdot e_i$  and

$$\begin{aligned}
 \langle y, u \rangle &= \left\langle x - \sum_{i=1}^n \langle x, e_i \rangle \cdot e_i, \sum_{k=1}^n u_k \cdot e_k \right\rangle \\
 &= \sum_{k=1}^n \bar{u}_k \cdot \left\langle x - \sum_{i=1}^n \langle x, e_i \rangle \cdot e_i, e_k \right\rangle \\
 &= \sum_{k=1}^n \bar{u}_k \cdot \left( \langle x, e_k \rangle - \left\langle \sum_{i=1}^n \langle x, e_i \rangle \cdot e_i, e_k \right\rangle \right) \\
 &= \sum_{k=1}^n \bar{u}_k \cdot \left( \langle x, e_k \rangle - \sum_{i=1}^n (\langle x, e_i \rangle \cdot \langle e_i, e_k \rangle) \right) \\
 &= \sum_{k=1}^n \bar{u}_k \cdot \left( \langle x, e_k \rangle - \sum_{i=1}^n (\langle x, e_i \rangle \cdot \delta_{i,k}) \right) \\
 &= \sum_{k=1}^n \bar{u}_k \cdot (\langle x, e_k \rangle - \langle x, e_k \rangle) \\
 &= 0
 \end{aligned}$$

proving that  $y \in U^\perp$  and as  $x$  is trivially a element of  $U$  we have that  $x \in U + U^\perp$ . So  $X = U + U^\perp$  and as by the previous theorem  $U \cap U^\perp = \{0\}$  we have by 10.368) that  $X = U \oplus U^\perp$  □

**Theorem 12.84.** Let  $\langle X, \langle \cdot, \cdot \rangle \rangle$  be a real (complex) finite dimensional space,  $L \in \text{Hom}(X)$  a self adjoint operator and  $U \subseteq X$  a sub space of  $X$  so that  $L(U) \subseteq U$  Then

1.  $L(U^\perp) \subseteq U^\perp$

2.  $T|_U \in \text{Hom}(U)$  is self adjoint in  $\langle U, \langle \rangle_{|U \times U} \rangle$
3.  $T|_{U^\perp} \in \text{Hom}(U^\perp)$  is self adjoint in  $\langle U^\perp, \langle \rangle_{U^\perp \times U^\perp} \rangle$

**Proof.**

1. If  $y \in L(U^\perp)$  then there exists a  $x \in U^\perp$  such that  $y = L(x)$ , if now  $u \in U$  we have that  $\langle y, u \rangle = \langle L(x), u \rangle \underset{\text{L is self adjoint}}{=} \langle x, L(u) \rangle \underset{x \in U^\perp \wedge L(u) \in U}{=} 0$  proving that  $y = L(x) \in U^\perp$ . Hence  $L(U^\perp) \subseteq U^\perp$
2. If  $u, v \in U$  then as  $T|_U(u) = T(u) \in U$  and  $T|_U(v) = T(v)$  so that  $\langle T|_U(u), v \rangle_{U \times U} = \langle T|_U(u), v \rangle = \langle T(u), v \rangle = \langle u, T(v) \rangle = \langle u, T|_U(v) \rangle = \langle u, T|_U(n) \rangle_{|U \times U}$  proving that  $T|_U$  is self adjoint in  $\langle U, \langle \rangle_{U \times U} \rangle$
3. This follows from (2) by replacing  $U$  by  $U^\perp$  and the fact that  $U^\perp$  is also a subspace in  $X$  (see 12.82).  $\square$

**Theorem 12.85. (Real Spectral Theorem)** Let  $\langle X, \langle \rangle \rangle$  be a real finite dimensional inner product space with  $\dim(X) = n \in \mathbb{N}_0$  and  $L \in \text{Hom}(X)$  then the following are equivalent

1.  $L$  is self-adjoint
2. There exists a orthonormal basis  $\{e_i\}_{i \in \{1, \dots, n\}}$  for  $X$  of eigenvectors of  $L$ .
3. There exists a basis for  $X$  that  $\mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{e_i\}_{i \in \{1, \dots, n\}})$  is diagonal.

**Proof.**

**1  $\Rightarrow$  2.** We prove first by induction on the dimension of  $X$  that  $X$  has a orthonormal basis of eigenvectors of  $T$ . So let  $\mathcal{F} = \{n \in \mathbb{N}_0 \mid \text{if } \dim(X) = n \text{ and } L \in \text{Hom}(X) \text{ is self-adjoint then there exist a orthonormal basis consisting of eigenvectors of } T\}$  then we have:

**1  $\in$   $\mathcal{F}$ .** If  $\dim(X) = 1$  then  $X$  has a basis  $\{f_i\}_{i \in \{1, \dots, 1\}}$  consisting of one vector  $f_1$ , take then  $e_1 = \frac{f_1}{\|f_1\|}$  then  $\forall x \in X$  there exists a  $\alpha_1 \in \mathbb{R}$  such that  $x = \alpha_1 \cdot f_1 = \frac{\alpha_1 \cdot \|f_1\|}{\|f_1\|} \cdot f_1 = \alpha_1 \cdot e_1$  proving that  $\{e_i\}_{i \in \{1, \dots, 1\}}$  is a orthonormal basis of  $X$ . Hence as  $L(e_1) \in X$  there exists a  $\lambda \in \mathbb{R}$  such that  $L(e_1) = \lambda \cdot e_1$  proving that  $e_1$  is a eigenvector of  $X$ . So  $1 \in \mathcal{F}$

**$n \in \mathcal{F} \Rightarrow n+1 \in \mathcal{F}$ .** Let  $\dim(X) = n+1$  and  $L \in \text{Hom}(X)$  is a self-adjoint operator.

Using 12.80 there exists a eigen vector  $x \in X$  such that  $L(x) = \lambda \cdot x$ , as by definition of eigenvectors  $x \neq 0$  we can define  $e_{n+1} = \frac{x}{\|x\|}$  so that  $L(e_{n+1}) = \frac{1}{\|x\|} \cdot L(x) = \frac{\lambda}{\|x\|} \cdot x = \lambda \cdot e_{n+1}$ . So we have

$$L \text{ has a eigenvector } e_{n+1} \text{ with } \|e_{n+1}\| = 1 \quad (12.59)$$

Define now  $U = \mathcal{S}\{e_{n+1}\} = \{\lambda \cdot e_{n+1} \mid \lambda \in \mathbb{R}\}$  which is a vector space with  $\dim(U) = 1$ . Take the vector space  $U^\perp$  (see 12.82) then by 12.83 we have that  $X = U \oplus U^\perp$  hence using 10.372 we get  $n+1 = \dim(X) = \dim(U) + \dim(U^\perp) = 1 + \dim(U)$ . So  $\dim(U^\perp) = n$ . As by the previous theorem (see 12.84) that  $L|_{U^\perp}$  is self-adjoint, hence using the fact that  $n \in \mathcal{F}$  there exists a orthonormal basis  $\{e_i\}_{i \in \{1, \dots, n\}}$  for  $U^\perp$  of eigen vectors of  $L|_{U^\perp}$  with eigenvalues  $\{\lambda_i\}_{i \in \{1, \dots, n\}}$ . Take now  $\{e_i\}_{i \in \{1, \dots, n+1\}}$  then for  $i, j \in \{1, \dots, n+1\}$  with  $i \neq j$  then we have the following cases:

- $i, j \in \{1, \dots, n\}$ . as  $\{e_i\}_{i \in \{1, \dots, n\}}$  is orthonormal we have that  $\langle e_i, e_j \rangle = 0$
- $i = n+1, j \in \{1, \dots, n\}$ . as  $e_j \in U^\perp$  and  $e_i \in U$  we have that  $\langle e_i, e_j \rangle = 0$
- $i \in \{1, \dots, n\}, j = n+1$ . as  $e_i \in U^\perp$  and  $e_j \in U$  we have that  $\langle e_i, e_j \rangle = 0$

The above proves that  $\{e_i\}_{i \in \{1, \dots, n+1\}}$  is orthonormal and by 11.121 that it is also a basi of  $C$ . Finally  $\forall i \in \{1, \dots, n\}$  we have that  $L(e_i) \underset{e_i \in U^\perp}{=} L|_{U^\perp}(e_i) = \lambda_i \cdot e_i$  proving together with 12.59 that  $\{e_i\}_{i \in \{1, \dots, n+1\}}$  is a orthonormal basis for  $X$  of eigenvectors of  $L$ . This proves that  $n+1 \in \mathcal{F}$ .

**2  $\Leftrightarrow$  3.** This follows from 12.76.

**2⇒1.** Let  $\{e_i\}_{i \in \{1, \dots, n\}}$  be a basis  $\{e\}_{i \in \{1, \dots, n\}}$  for  $X$  of eigen vectors for  $L$  with eigenvalues  $\{\lambda_i\}_{i \in \{1, \dots, n\}}$ . Then if  $x \in X$  we have that  $x = \sum_{i=1}^n x_i \cdot e_i$ ,  $y = \sum_{i=1}^n y_i \cdot e_i$

$$\begin{aligned}
\langle L(x), y \rangle &= \left\langle L\left(\sum_{i=1}^n x_i \cdot e_i\right), \sum_{j=1}^n y_j \cdot e_j \right\rangle \\
&= \left\langle \sum_{i=1}^n x_i \cdot L(e_i), \sum_{j=1}^n y_j \cdot e_j \right\rangle \\
&= \sum_{i=1}^n x_i \left\langle \lambda_i \cdot e_i, \sum_{j=1}^n y_j \cdot e_j \right\rangle \\
&= \sum_{i=1}^n x_i \cdot \lambda_i \left\langle e_i, \sum_{j=1}^n y_j \cdot e_j \right\rangle \\
&= \sum_{i=1}^n \left( x_i \cdot \lambda_i \cdot \sum_{j=1}^n y_j \cdot \langle e_i, e_j \rangle \right) \\
&= \sum_{i=1}^n \lambda_i \cdot x_i \cdot y_i \\
\langle x, L(y) \rangle &= \left\langle \sum_{i=1}^n x_i \cdot e_i, L\left(\sum_{j=1}^n y_j \cdot e_j\right) \right\rangle \\
&= \sum_{i=1}^n x_i \cdot \left\langle e_i, \sum_{j=1}^n y_j \cdot L(e_j) \right\rangle \\
&= \sum_{i=1}^n x_i \cdot \left( \sum_{j=1}^n y_j \cdot \lambda_j \cdot \langle e_i, e_j \rangle \right) \\
&= \sum_{i=1}^n \lambda_i \cdot x_i \cdot y_i \\
&= \langle L(x), y \rangle
\end{aligned}$$

proving that  $L$  is self-adjoint □

**Corollary 12.86.** Let  $\langle X, \langle \cdot, \cdot \rangle \rangle$  be a real finite dimensional inner product space with  $\dim(X) = n \in \mathbb{N}_0$ ,  $\{e_i\}_{i \in \{1, \dots, n\}}$  a orthonormal base,  $L \in \text{Hom}(X)$  a adjoint operator then there exists a unitary operator  $U \in \text{Hom}(X)$  such that  $\mathcal{M}(U^* \circ L \circ U, \{e_i\}_{i \in \{1, \dots, n\}})$  is diagonal.

**Proof.** Using the Real spectral theorem 12.85 there exists a basis  $\{f_i\}_{i \in \{1, \dots, n\}}$  such that  $\mathcal{M}(L, \{f_i\}_{i \in \{1, \dots, n\}})$  is diagonal. Take now the unitary mapping  $U \in \text{Hom}(X)$  defined by  $\forall i \in \{1, \dots, n\}$   $U(e_i) = f_i$  (see 12.70) then we have  $\forall i, j \in \{1, \dots, n\}$  that

$$\begin{aligned}
\mathcal{M}(U^* \circ L \circ U, \{e_i\}_{i \in \{1, \dots, n\}})_{i,j} &\stackrel{11.118}{=} \langle (U^* \circ L \circ U)(e_j), e_i \rangle \\
&= \langle U^*(L(U(e_j))), e_i \rangle \\
&= \langle U^*(L(f_j)), e_i \rangle \\
&= \langle L(f_j), (U^*)^*(e_i) \rangle \\
&= \langle L(f_j), U(e_i) \rangle \\
&= \langle L(f_j), f_i \rangle \\
&\stackrel{11.118}{=} \mathcal{M}(L, \{f_i\}_{i \in \{1, \dots, n\}})_{i,j} \\
&= \mathcal{M}(L, \{f_i\}_{i \in \{1, \dots, n\}})_{i,i} \cdot \delta_{i,j}
\end{aligned}$$

□

**Theorem 12.87.** *Let  $n \in \mathbb{N}_0$  then we have we have*

1. *If  $M \in \mathcal{M}(n \times n, \mathbb{R})$  such that  $M$  is symmetric then there exists a unitary matrix  $U$  such that  $U^T \cdot M \cdot U \underset{U \text{ is unitary}}{=} U^{-1} \cdot M \cdot U$  is diagonal*
2. *If  $M \in \mathcal{M}(n \times n, \mathbb{C})$  such that  $M$  is Hermitian then there exists a unitary matrix  $U$  such that  $U^T \cdot M \cdot U \underset{U \text{ is unitary}}{=} U^{-1} \cdot M \cdot U$  is diagonal.*

**Proof.**

1. Take the inner product vector space  $\langle \mathbb{R}^n, \langle \rangle \rangle$  with the orthonormal basis  $\{\varepsilon_i\}_{i \in \{1, \dots, n\}}$  (see 11.116) and define  $L: \mathbb{R}^n \rightarrow \mathbb{R}^n$  by  $L(\varepsilon_i) = \sum_{k=1}^n M_{k,i} \cdot \varepsilon_k$  so that  $\mathcal{M}(L, \{\varepsilon_i\}_{i \in \{1, \dots, n\}}) = M$ . As  $M$  is symmetric we have in the real case that  $M$  is Hermitian and thus by 12.59 that  $L$  is self-adjoint. Using the Real Spectral theorem (see 12.85) there exists a orthonormal basis  $\{f_i\}_{i \in \{1, \dots, n\}}$  such that  $\mathcal{M}(L, \{f_i\}_{i \in \{1, \dots, n\}})$  is diagonal. Using 12.70 we have then that  $\mathcal{M}(L, \{f_i\}_{i \in \{1, \dots, n\}}) = \mathcal{M}(U, \{\varepsilon_i\}_{i \in \{1, \dots, n\}})^{-1} \cdot \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}) \cdot \mathcal{M}(U, \{\varepsilon_i\}_{i \in \{1, \dots, n\}}) = U^{-1} \cdot M \cdot U$  where  $U = \mathcal{M}(U, \{\varepsilon_i\}_{i \in \{1, \dots, n\}})$  is unitary (see 12.70 and the fact that  $\{\varepsilon_i\}_{i \in \{1, \dots, n\}}$  and  $\{f_i\}_{i \in \{1, \dots, n\}}$  are orthonormal).
2. Take the inner product vector space  $\langle \mathbb{C}^n, \langle \rangle \rangle$  with the orthonormal basis  $\{\varepsilon_i\}_{i \in \{1, \dots, n\}}$  (see 11.116) and define  $L: \mathbb{C}^n \rightarrow \mathbb{C}^n$  defined by  $L(\varepsilon_i) = \sum_{k=1}^n M_{k,i} \cdot \varepsilon_k$  so that  $\mathcal{M}(L, \{\varepsilon_i\}_{i \in \{1, \dots, n\}}) = M$ . As  $M$  is Hermitian we have by 12.59 that  $L$  is self-adjoint and thus rormal (see 12.72). Using the Complex Spectral theorem (see 12.85) there exists a orthonormal basis  $\{f_i\}_{i \in \{1, \dots, n\}}$  such that  $\mathcal{M}(L, \{f_i\}_{i \in \{1, \dots, n\}})$  is diagonal. Using 12.70 we have then that  $\mathcal{M}(L, \{f_i\}_{i \in \{1, \dots, n\}}) = \mathcal{M}(U, \{\varepsilon_i\}_{i \in \{1, \dots, n\}})^{-1} \cdot \mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}) \cdot \mathcal{M}(U, \{\varepsilon_i\}_{i \in \{1, \dots, n\}}) = U^{-1} \cdot M \cdot U$  where  $U = \mathcal{M}(U, \{\varepsilon_i\}_{i \in \{1, \dots, n\}})$  is unitary (see 12.70 and the fact that  $\{\varepsilon_i\}_{i \in \{1, \dots, n\}}$  and  $\{f_i\}_{i \in \{1, \dots, n\}}$  are orthonormal).  $\square$

# Chapter 14

## Differentiability in normed vector spaces

We assume in this chapter that, unless stated otherwise, all the normed and Banach spaces are using the same field  $\mathbb{K}$  where  $\mathbb{K}$  is either  $\mathbb{C}$  or  $\mathbb{R}$ . Further we use the following notation for constant functions.

**Notation 14.1.** Let  $X, Y$  be sets,  $y \in Y$  then  $C_{X,Y,y}: X \rightarrow Y$  is defined  $C_{X,Y,y}(x) = y$  [see 2.15]  
For example if  $\langle X, \|\cdot\|_X \rangle$  and  $\langle Y, \|\cdot\|_Y \rangle$  are normed spaces and  $U \subseteq X$  then we have

1.  $C_{X,Y,y}$  is the function from  $X$  to  $Y$  defined by  $C_{X,Y,y}(x) = y \in Y$
2.  $C_{X,Y,0}$  is the function from  $X$  to  $Y$  defined by  $C_{X,Y,0}(x) = 0$  where  $0$  is the neutral element of  $Y$
3.  $C_{U,Y,y}$  is the function from  $U$  to  $Y$  defined by  $C_{U,Y,y}(x) = y \in Y$
4.  $C_{U,Y,0}$  is the function from  $U$  to  $Y$  defined by  $C_{U,Y,0}(x) = 0$  where  $0$  is the neutral element of  $Y$
5. Let  $n \in \mathbb{N}_0$  then  $C_{X,L^n(X;Y),0}$  is the function from  $X$  to  $Y$  defined by  $C_{X,L^n(X;Y),0}(x) = 0$  where  $0$  is the neutral element in  $L^n(X, Y)$
6. Let  $n \in \mathbb{N}_0$  then  $C_{U,L^n(X;Y),0}$  is the function from  $U$  to  $Y$  defined by  $C_{U,L^n(X;Y),0}(x) = 0$  where  $0$  is the neutral element in  $L^n(X, Y)$

### 14.1 Differentiability

#### 14.1.1 Differential

A function is differentiable at a point if we can approximate the function with a continuous linear function in a small enough neighborhood around the point. To ensure that we don't evaluate the function outside its domain we define given a set  $U$  the set  $U_x$ .

**Definition 14.2.** Let  $\langle X, \|\cdot\| \rangle$  be a normed vector space and let  $U \subseteq X$ ,  $x \in U$  then  $U_x = \{h \in X \mid x + h \in U\}$

**Proposition 14.3.** Let  $\langle X, \|\cdot\| \rangle$  be a normed space,  $\delta > 0$  and  $x \in X$  then  $B(0, \delta) = B(x, \delta)_x$

**Proof.**

$$\begin{aligned} h \in B(x, \delta)_x &\Leftrightarrow x + h \in B(x, \delta) \\ &\Leftrightarrow h \in X \wedge \|x + h - x\| < \delta \\ &\Leftrightarrow h \in X \wedge \|h\| < \delta \\ &\Leftrightarrow h \in B_{\|\cdot\|}(0, \delta) \end{aligned}$$

□

**Theorem 14.4.** Let  $\langle X, \|\cdot\| \rangle$  be a normed vector space,  $U \subseteq X$ ,  $x \in U$  then  $U_x = (-x) + U = U - x$

**Proof.** First

$$\begin{aligned} h \in U_x &\Leftrightarrow x + h \in U \\ &\Leftrightarrow \exists y \in U \text{ with } x + h = y \\ &\Leftrightarrow \exists y \in U \text{ with } h = (-x) + y \\ &\Leftrightarrow h \in (-x) + U \end{aligned}$$

Second

$$\begin{aligned} h \in (-x) + U &\Leftrightarrow \exists y \in U \text{ with } h = (-x) + y \\ &\Leftrightarrow \exists y \in U \text{ with } h = y - x \\ &\Leftrightarrow h \in U - x \end{aligned}$$

□

**Theorem 14.5.** Let  $\langle X, \|\cdot\| \rangle$  be a normed vector space and let  $U \subseteq X$  open,  $x \in U$  then  $U_x$  is open.

**Proof.** This follows directly from the above and 11.79

□

Let's now define the differentiability of a function.

**Definition 14.6.** Let  $\langle X, \|\cdot\|_X \rangle$ ,  $\langle Y, \|\cdot\|_Y \rangle$  be real normed spaces,  $U \subseteq X$  a open set,  $V \subseteq Y$  then a function  $f: U \rightarrow V$  is differentiable at  $x \in U$  if and only if there exists a continuous linear function  $L: X \rightarrow Y$  (or  $L \in L(X, Y)$ ) such that  $\forall \varepsilon > 0$  there exists a  $\delta > 0$  such that for all  $h \in U_x$  with  $\|h\|_X < \delta$  we have  $\|f(x + h) - f(x) - L(h)\|_Y \leq \varepsilon \cdot \|h\|_X$

Note that  $h \in U_x$  ensures that  $x + h \in U = \text{dom}(f)$  so our definition is well defined. Next we will prove that the linear approximation of  $f$  is unique.

**Theorem 14.7.** Let  $\langle X, \|\cdot\|_X \rangle$ ,  $\langle Y, \|\cdot\|_Y \rangle$  be real normed space,  $U \subseteq X$  a open set,  $V \subseteq Y$ ,  $x \in U$  and  $f: U \rightarrow V$  differentiable at  $x$  then the  $L$  used in the above definition is unique

**Proof.** Let  $h \in X$  then we have the following cases to consider:

**$h = 0$ .** then  $\|L_1(h) - L_2(h)\|_Y = \|L_1(0) - L_2(0)\|_Y = \|0 - 0\|_Y = 0$

**$h \neq 0$ .** Let  $\varepsilon > 0$  then for  $i \in \{1, 2\}$  there exists  $\delta_i > 0$  such that for  $h \in U_x$  and  $\|h\|_X < \delta_i$  we have  $\|f(x + h) - f(x) - L_i(h)\|_Y \leq \frac{\varepsilon}{2} \cdot \|h\|_X$ . Further as  $U_x$  is open there exists a  $\delta_3 > 0$  such that  $B_{\|\cdot\|_X}(x, \delta_3) \subseteq U_x$ . Take  $\delta = \min(\delta_1, \delta_2, \delta_3)$  and define  $h' = \frac{\delta}{2 \cdot \|h\|_X} \cdot h$  then  $\|h'\|_X = \frac{\delta}{2 \cdot \|h\|_X} \cdot \|h\|_X = \frac{\delta}{2} < \delta_1, \delta_2, \delta_3$  so that  $h' \in U_x$  and

$$\begin{aligned} \|L_1(h') - L_2(h')\|_Y &= \|((f(x + h') - f(x) - L_2(h')) - (f(x + h') - f(x) - L_1(h'))\|_Y \\ &\leq \frac{\varepsilon}{2} \cdot \|h'\|_X + \frac{\varepsilon}{2} \cdot \|h'\|_X = \varepsilon \cdot \|h'\|_X \end{aligned} \tag{14.1}$$

hence

$$\begin{aligned} \frac{\delta}{2 \cdot \|h\|_X} \cdot \|L_1(h) - L_2(h)\|_Y &\stackrel{L_1, L_2 \text{ are linear}}{=} \left\| L_2 \left( \frac{\delta}{2 \cdot \|h\|_X} \cdot h \right) - L_1 \left( \frac{\delta}{2 \cdot \|h\|_X} \cdot h \right) \right\|_Y \\ &= \|L_2(h') - L_1(h')\|_Y \\ &\stackrel{14.1}{\leq} \varepsilon \cdot \|h'\|_X \end{aligned}$$

As  $\varepsilon$  was chosen arbitrary we have by 9.57 that  $0 \leq \frac{\delta}{2 \cdot \|h\|_X} \cdot \|L_1(h) - L_2(h)\|_Y \leq 0$  proving that  $\|L_1(h) - L_2(h)\|_Y = 0$ .

So we have for  $h \in X$  that  $\|L_1(h) - L_2(h)\|_Y = 0 \Rightarrow L_1(h) - L_2(h) = 0 \Rightarrow L_1(h) = L_2(h)$  which proves that  $L_1 = L_2$ .

□

The unique continuous linear approximation of a function at a point is called the differential of the function at this point.

**Definition 14.8.** Let  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle$  be real normed spaces,  $U \subseteq X$  an open set,  $V \subseteq Y$ ,  $x \in U$  and  $f: U \rightarrow V$  differentiable at  $x$  then the unique [by the previous theorem]  $L \in L(X, Y)$  such that  $\forall \varepsilon > 0$  there exists a  $\delta > 0$  so that if  $h \in U_x$  and  $\|h\|_X < \delta$  then  $\|f(x+h) - f(x) - L(h)\|_Y \leq \varepsilon \cdot \|h\|_X$  is called the **differential of  $f$  at  $x$**  and is noted as  $Df(x)$ . In other words  $f$  is differentiable at  $x$  with differential  $Df(x)$  if and only there exists a  $Df(x) \in L(X, Y)$  [which is unique] such that  $\forall \varepsilon > 0$  there exists a  $\delta > 0$  such that if  $h \in U_x$  and  $\|h\|_X < \delta$  then  $\|f(x+h) - f(x) - Df(x)(h)\|_Y \leq \varepsilon \cdot \|h\|_X$

The next theorem shows that differentiation is independent of the chosen norm (as long as the norms are equivalent).

**Theorem 14.9.** Let  $X, Y$  be vector spaces over  $\mathbb{K}$  with equivalent norms  $\|\cdot\|_{X_1}, \|\cdot\|_{X_2}$  on  $X$ , equivalent norms  $\|\cdot\|_{Y_1}, \|\cdot\|_{Y_2}$  on  $Y$ ,  $U \subseteq X$  an open set [as norms are equivalent the topology on  $X, Y$  is the same for the two norms hence  $U$  is open in both topologies],  $V \subseteq Y$  and  $f: U \rightarrow V$  a function then if  $f$  is differentiable at  $x \in U$  with differential  $Df(x)$  [using norms  $\|\cdot\|_{X_1}, \|\cdot\|_{Y_1}$ ] we have that  $f$  is differentiable at  $x \in U$  with differential  $Df(x)$  [using norms  $\|\cdot\|_{X_2}, \|\cdot\|_{Y_2}$ ]

**Proof.** From the equivalence of the norms we have by 11.85 that there exists  $\alpha_X, \beta_X, \alpha_Y, \beta_Y > 0$  such that  $\forall x \in X$  and  $\forall y \in Y$  we have

$$\alpha_X \cdot \|x\|_{X_2} \leq \|x\|_{X_1} \leq \beta_X \cdot \|x\|_{X_2} \text{ and } \alpha_Y \cdot \|y\|_{Y_1} \leq \|y\|_{Y_2} \leq \beta_Y \cdot \|y\|_{Y_1} \quad (14.2)$$

Assume that  $f: U \rightarrow V$  is differentiable at  $x$  with differential  $Df(x)$  using the norms  $\|\cdot\|_{X_1}, \|\cdot\|_{Y_1}$ . Then for  $\varepsilon > 0$  there exists a  $\delta'$  such that

$$\|f(x+h) - f(x) - Df(x)(h)\|_{Y_1} \leq \frac{\varepsilon}{\beta_Y \cdot \beta_X} \cdot \|h\|_{X_1} \quad (14.3)$$

Take  $\delta = \frac{\delta'}{\beta_X}$  then  $\forall h \in U_x$  with  $\|h\|_{X_2} < \delta$  we have  $\beta_X \cdot \|h\|_{X_2} < \delta' \Rightarrow \|h\|_{X_1} < \delta'$  so using 14.3 we have that  $\|f(x+h) - f(x) - Df(x)(h)\|_{Y_1} \leq \frac{\varepsilon}{\beta_Y \cdot \beta_X} \cdot \|h\|_{X_1}$  so that  $\beta_Y \cdot \|f(x+h) - f(x) - Df(x)(h)\|_{Y_1} \leq \frac{\varepsilon}{\beta_X} \cdot \|h\|_{X_1} \leq \frac{\varepsilon}{\beta_X} \cdot \|h\|_{X_2}$  or using 14.2 we have

$$\|f(x+h) - f(x) - Df(x)(h)\|_{Y_2} \leq \varepsilon \cdot \|h\|_{X_2}$$

which together with the fact that  $Df(x)$  is also a continuous function between  $X$  and  $Y$  [as the topologies are the same] proves that  $f$  has also a differential  $Df(x)$  at  $x$  using the norms  $\|\cdot\|_{X_2}$  and  $\|\cdot\|_{Y_2}$ .  $\square$

We now prove that differentiability of  $f: U \rightarrow V$  at  $x$  is essential a local property and is not dependent on  $U$  but just on the value of  $f$  near  $x$ . Further differentiability is only dependent on the range of a function

**Theorem 14.10.** Let  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle$  be normed spaces,  $U \subseteq X$  an open set,  $V \subseteq Y$ ,  $x \in U$  and  $f: U \rightarrow V$  then

1. If  $f$  is differentiable at  $x$  then
  - a.  $\forall W$  open with  $x \in W \subseteq U$  we have that  $f|_W: W \rightarrow Y$  is differentiable at  $x$
  - b.  $\forall W \subseteq Y$  with  $\text{range}(f) \subseteq W$  we have that  $f: U \rightarrow W$  is differentiable at  $x$
2. If there exists a  $W \subseteq X$  open with  $x \in W \subseteq U$  such that  $f|_W: W \rightarrow V$  is differentiable at  $x$  then  $f: U \rightarrow V$  is differentiable at  $x$ .
3. If there exists a  $W \subseteq Y$  with  $\text{range}(f) \subseteq W$  such that  $f: U \rightarrow W$  is differentiable at  $x$  then  $f: U \rightarrow V$  is differentiable at  $x$ .

Also  $Df(x) = D(f|_V)(x)$  in (1), (2) and (3).

**Proof.**

1. As  $f$  is differentiable we have given  $\varepsilon > 0$  a  $\delta > 0$  so that  $\forall h \in U_x$  with  $\|h\|_X < \delta$  we have  $\|f(x+h) - f(x) - Df(x)(h)\|_Y \leq \varepsilon \cdot \|h\|_X$ .
  - a. Now if  $h \in W_x \Rightarrow x+h \in W \subseteq U \Rightarrow h \in U_x \Rightarrow W_x \subseteq U_x$ . So if  $h \in W_x$  with  $\|h\|_X < \delta$  we have  $\|f|_W(x+h) - f|_W(x) - Df(x)(h)\|_Y = \|f(x+h) - f(x) - Df(x)(h)\|_Y \leq \varepsilon \cdot \|h\|_X$  proving that  $f|_W: W \rightarrow Z$  is differentiable with differential  $Df(x)$  and thus  $D(f|_W)(x) = Df(x)$

- b. Note that differentiability is a property of the graph of a function. Hence if  $\text{range}(f) \subseteq W \subseteq Y$  then as the graph of  $f: U \rightarrow V$  and  $f: U \rightarrow W$  is  $f$  we have that  $f: U \rightarrow W$  is differentiable at  $x$  with the same differential  $Df(x)$
2. First as  $x \in W$  open there exists a  $\delta_1 > 0$  such that  $x \in B_{\| \cdot \|_X}(x, \delta_1) \subseteq W$ , so if  $\|h\|_X \leq \delta_1$  then  $\|x + h - x\|_X = \|h\|_X < \delta_1 \Rightarrow x + h \in B_{\| \cdot \|_X}(x, \delta_1) \subseteq W$  or in other words
- $$\exists \delta_1 > 0 \text{ such that if } \|h\|_X \leq \delta_1 \Rightarrow h \in W_x \quad (14.4)$$
- Take now  $\varepsilon > 0$ , then by differentiability of  $f|_W$  there exists a  $\delta_2 > 0$  such that if  $h \in W_x$  and  $\|h\|_X < \delta_2$  then  $\|f|_W(x + h) - f|_W(x) - D(f|_W)(x)(h)\|_Y \leq \varepsilon \cdot \|h\|_X \Rightarrow \|f(x + h) - f(x) - D(f)(x)(h)\|_Y \leq \varepsilon \cdot \|h\|_X$  or in other words
- $$\text{if } h \in W_x \text{ and } \|h\|_X < \delta_2 \text{ then } \|f(x + h) - f(x) - D(f)(x)(h)\|_Y \leq \varepsilon \cdot \|h\|_X \quad (14.5)$$
- Take now  $\delta = \min(\delta_1, \delta_2)$  then if  $h \in U_x$  and  $\|h\|_X < \delta$  then  $\|h\|_X < \delta_1 \stackrel{14.4}{\Rightarrow} h \in W_x$  and as  $\|h\|_X < \delta_2$  we have by 14.5 that  $\|f(x + h) - f(x) - D(f)(x)(h)\|_Y \leq \varepsilon \cdot \|h\|_X$  proving that  $f$  is differentiable at  $x$  with differential  $D(f)(x)$  thus  $Df(x) = D(f)(x)$ .
3. Note that differentiability is a property of the graph of a function. Hence if  $\text{range}(f) \subseteq W \subseteq Y$  then as the graph of  $f: U \rightarrow V$  and  $f: U \rightarrow W$  is  $f$  we have that  $f: U \rightarrow V$  is differentiable at  $x$  with the same differential  $Df(x)$ .  $\square$

Let's now look at some alternative definitions for differentiability at a point.

**Definition 14.11.** Let  $\langle X, \| \cdot \|_X \rangle, \langle Y, \| \cdot \|_Y \rangle$  be normed spaces,  $U \subseteq X$  a open set and  $x \in U$  then a  $\varepsilon$ -mapping at  $x$  is a function  $\varepsilon: U_x \rightarrow Y$  which is continuous at  $0 \in U_x$  [using the subspace topology on  $U_x$ ] and for which  $\varepsilon(0) = 0$

We prove now the following equivalent alternative definitions of differentiability

**Theorem 14.12.** Let  $\langle X, \| \cdot \|_X \rangle, \langle Y, \| \cdot \|_Y \rangle$  be normed spaces,  $U \subseteq X$  a open set,  $V \subseteq Y$ ,  $x \in U$  and  $f: U \rightarrow V$  a function then the following are equivalent

1.  $f$  is differentiable at  $x$  with differential  $Df(x)$
2. There exists a  $Df(x) \in (X, Y)$  and a  $\varepsilon$ -mapping  $\varepsilon: U_x \rightarrow Y$  such that  $\forall h \in U_x$  we have  $f(x + h) - f(x) - Df(x)(h) = \varepsilon(h) \cdot \|h\|_X$
3. There exists a  $Df(x) \in L(X, Y)$  such that  $\forall \varepsilon > 0$  there exists a  $\delta > 0$  such that  $\forall h \in U_x$  with  $0 < \|h\| < \varepsilon$  we have  $\frac{\|f(x + h) - f(x) - Df(x)(h)\|_Y}{\|h\|_X} < \varepsilon$
4. There exists a  $Df(x) \in L(X, Y)$  such that  $\forall \{h_i\}_{i \in \mathbb{N}_0}$  with  $\lim_{h \rightarrow 0} h_i = 0$  and  $\forall i \in \mathbb{N}_0 \|h_i\|_X > 0$  and  $h_i \in U_x$  we have that  $\lim_{n \rightarrow \infty} \frac{\|f(x + h_i) - f(x) - Df(x)(h_i)\|_Y}{\|h_i\|_X} = 0$

**Proof.**

**1  $\Rightarrow$  2.** Define  $\varepsilon: U_x \rightarrow Y$  by  $\varepsilon(h) = \begin{cases} 0 & \text{if } h = 0 \\ \frac{f(x + h) - f(x) - Df(x)(h)}{\|h\|_X} & \text{if } h \neq 0 \end{cases}$  then we have that  $\varepsilon(h) \cdot \|h\|_X = \begin{cases} 0 & \text{if } h = 0 \\ f(x + h) - f(x) - Df(x)(h) & \text{if } h \neq 0 \end{cases} = f(x + h) - f(x) - Df(x) \cdot \|h\|_X$ . Further take  $\zeta > 0$  then by definition there exists a  $\delta > 0$  such that for  $h \in U_x$  with  $\|h\|_X < \delta$  we have that  $\|f(x + h) - f(x) - Df(x)(h)\|_Y \leq \frac{\zeta}{2} \cdot \|h\|_X < \zeta \cdot \|h\|_X$  so that

$$\begin{aligned} \|\varepsilon(h) - \varepsilon(0)\|_Y &= \|\varepsilon(h)\|_Y \\ &= \begin{cases} \|0\|_Y < \zeta & \text{if } h = 0 \\ \frac{\|f(x + h) - f(x) - Df(x)(h)\|_Y}{\|h\|_X} & \text{if } h \neq 0 \end{cases} < \zeta \end{aligned}$$

which using 11.78 and 11.155 proves that  $\varepsilon$  is continuous at 0 in the subspace topology on  $U_x$ .

**2  $\Rightarrow$  3.** Let  $\varepsilon: U_x \rightarrow Y$  be the  $\varepsilon$ -mapping such that  $f(x+h) - f(x) - Df(x)(h) = \varepsilon(h) \cdot \|h\|_X$   $\forall h \in U_x$ . Then as  $\varepsilon$  is continuous at 0 and  $\varepsilon(0) = 0$  there exists given a  $\zeta > 0$  a  $\delta > 0$  such that if  $h \in U_x$  with  $\|h\|_X < \delta$  we have  $\|\varepsilon(h)\|_Y = \|\varepsilon(h) - 0\|_Y = \|\varepsilon(h) - \varepsilon(0)\|_Y < \zeta$  so that if additional  $0 < \|h\|_X$  we have  $\frac{\|f(x+h) - f(x) - Df(x)(h)\|_Y}{\|h\|_X} = \frac{\|\varepsilon(h)\|_Y \cdot \|h\|_X}{\|h\|_X} = \|\varepsilon(h)\|_X < \zeta$ .

**3  $\Rightarrow$  1.** Let  $\varepsilon > 0$ , find a  $\delta > 0$  such that for  $h \in U_x$  with  $0 < \|h\|_X < \delta$  we have  $\frac{\|f(x+h) - f(x) - Df(x)(h)\|_Y}{\|h\|_Y} < \varepsilon$ . Now  $\forall h \in U_x$  with  $\|h\|_X < \delta$  we have two cases to consider

$\|h\|_X = 0$ . then  $h = 0$  so that  $\|f(x+h) - f(x) - Df(x)(h)\|_Y = 0 \leq 0 = \varepsilon \cdot 0 = \varepsilon \cdot \|h\|_X$

$\|h\|_X \neq 0$ . then  $\frac{\|f(x+h) - f(x) - Df(x)(h)\|_Y}{\|h\|_X} < \varepsilon \Rightarrow \|f(x+h) - f(x) - Df(x)(h)\|_Y \leq \varepsilon \cdot \|h\|_X$

so in all cases we have  $\|f(x+h) - f(x) - Df(x)(h)\|_Y \leq \varepsilon \cdot \|h\|_X$  proving differentiability at  $x$ .

**3  $\Rightarrow$  4.** Let  $\{h_i\}_{i \in \mathbb{N}_0}$  be a sequence of strict positive real numbers with limit 0. Take now  $\varepsilon > 0$  then by (3) there exists a  $\delta_\varepsilon$  such that  $\forall h \in U_x$  with  $0 < \|h\|_X < \delta_\varepsilon$  we have  $\frac{\|f(x+h) - f(x) - Df(x)(h)\|_Y}{\|h\|_X} < \varepsilon$ . As  $\lim_{n \rightarrow \infty} h_n = 0$  there exists a  $N \in \mathbb{N}_0$  such that  $\forall n \geq N$  we have  $\|h_n\| < \delta_\varepsilon$ , then as we have assumed that  $\forall n \in \mathbb{N}_0$   $0 < h_n \in U_x$ , we have if  $n \geq N$  that  $\frac{\|f(x+h_n) - f(x) - Df(x)(h_n)\|_Y}{\|h_n\|_X} < \varepsilon$  proving that

$$\lim_{n \rightarrow \infty} \frac{\|f(x+h_n) - f(x) - Df(x)(h_n)\|_Y}{\|h_n\|_X} = 0$$

**4  $\Rightarrow$  3.** Assume that (3) is false then there exists a  $\varepsilon_0 > 0$  such that  $\forall \delta > 0$  there exists a  $h \in U_x$  with  $0 < \|h\|_X < \delta$  so that  $\frac{\|f(x+h) - f(x) - Df(x)(h)\|_Y}{\|h\|_X} > \varepsilon_0$ . So for every  $n \in \mathbb{N}_0$  there exists a  $h_n \in U_x$  with  $0 < \|h_n\|_X < \frac{1}{n}$  so that  $\frac{\|f(x+h_n) - f(x) - Df(x)(h_n)\|_Y}{\|h_n\|_X} > \varepsilon_0$ . Now if  $\delta > 0$  there exists by the Archimedean property [see 9.56] a  $N_\delta \in \mathbb{N}_0$  such that  $\frac{1}{N_\delta} < \delta$  so if  $n \geq N_\delta$  we have  $\|h_n - 0\|_X = \|h_n\|_X < \frac{1}{n} \leq \frac{1}{N_\delta} < \delta$  proving that  $\lim_{n \rightarrow \infty} h_n = 0$ . Now by the hypothesis (4) we must have that  $\lim_{n \rightarrow \infty} \frac{\|f(x+h_n) - f(x) - Df(x)(h_n)\|_Y}{\|h_n\|_X} = 0$  and thus there exists a  $N \in \mathbb{N}_0$  such that for  $n \geq \max(N, N_\delta)$  we have  $\frac{\|f(x+h_n) - f(x) - Df(x)(h_n)\|_Y}{\|h_n\|_X} < \varepsilon_0$  contradicting  $\frac{\|f(x+h_n) - f(x) - Df(x)(h_n)\|_Y}{\|h_n\|_X} > \varepsilon_0$ . As the assumption gives a contradiction we must have that (3) is true.  $\square$

**Remark 14.13.** If  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle$  are finite dimensional spaces then by 11.306 we have that every linear map between  $X$  and  $Y$  is continuous, so in the the previous theorem and the definition of differentiability we can leave out the requirement for the continuity of  $Df(x)$  in the finite dimensional case.

Using the alternative definitions of differentiability [see 14.12] and the locality of the differential [see 14.10] we have then the following criteria for differentiability.

**Corollary 14.14.** Let  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle$  be normed spaces,  $U \subseteq X$  a open set,  $x \in U$ ,  $V \subseteq Y$  and  $f: U \rightarrow V$  a function then the following are equivalent

1.  $f$  is differentiable at  $x$  with differential  $Df(x)$
2. There exists a open set  $W \subseteq U$ ,  $x \in W$ , a  $Df(x) \in L(X, Y)$  and a  $\varepsilon$ -mapping  $\varepsilon: W_x \rightarrow Y$  such that  $\forall h \in W_x$  we have  $f(x+h) - f(x) - Df(x)(h) = \varepsilon(h) \cdot \|h\|_X$
3. There exists a open set  $W \subseteq U$ ,  $x \in W$ , a  $Df(x) \in L(X, Y)$  such that  $\forall \varepsilon > 0$  there exists a  $\delta > 0$  such that  $\forall h \in W_x$  with  $0 < \|h\| < \varepsilon$  we have  $\frac{\|f(x+h) - f(x) - Df(x)(h)\|_Y}{\|h\|_X} < \varepsilon$

4. There exists a open set  $W \subseteq U$ ,  $x \in W$ , a  $Df(x) \in L(X, Y)$  such that  $\forall \{h_i\}_{i \in \mathbb{N}_0}$  with  $\lim_{h_i \rightarrow 0} h_i = 0$  and  $\forall i \in \mathbb{N}_0 \ \|h_i\| > 0$  and  $h_i \in W$  we have that  $\lim_{n \rightarrow \infty} \frac{\|f(x + h_i) - f(x) - Df(x)(h_i)\|_Y}{\|h_i\|_X} = 0$

The following theorem is a consequence of the fact that the differential is a continuous map.

**Theorem 14.15.** Let  $\langle X, \|\cdot\|_X \rangle$ ,  $\langle Y, \|\cdot\|_Y \rangle$  be normed spaces,  $U \subseteq X$  a open set,  $V \subseteq Y$ ,  $x \in U$  and  $f: U \rightarrow V$  a function that is differentiable at  $x$  then  $f$  is continuous at  $x$ .

**Proof.** Let  $W$  be a open subset of  $Y$  with  $f(x) \in W$  then we have

$$\exists \delta_1 > 0 \text{ such that } f(x) \subseteq B_{\|\cdot\|_Y}(f(x), \delta_1) \subseteq W \quad (14.6)$$

From the alternative definition of differentiability at  $x$  (see 14.12) there exists a  $\varepsilon$ -mapping  $\varepsilon: U_x \rightarrow Y$  such that  $\forall h \in U_x$  we have  $f(x + h) - f(x) - Df(x)(h) = \varepsilon(h) \cdot \|h\|_X$ . As  $\varepsilon$  is continuous at 0 and  $\varepsilon(0) = 0$  there exists a  $\delta'_2 > 0$  such that if  $\|h\|_X < \delta$  then  $\|\varepsilon(h)\|_Y < \frac{\delta_1}{2}$ . Take now  $\delta_2 = \min(\delta'_2, 1)$  then we have  $\|\varepsilon(h)\|_Y \cdot \|h\|_X < \|\varepsilon(h)\|_Y < \frac{\delta_1}{2}$  if  $\|h\|_X < \delta_2$  giving

$$\text{If } h \in U_x \text{ and } \|h\|_X < \delta_2 \text{ then } \|f(x + h) - f(x) - Df(x)(h)\|_Y < \frac{\delta_1}{2} \quad (14.7)$$

Using the continuity of the linear map  $Df(x)$  and  $Df(x)(0) = 0$  we have also

$$\exists \delta_3 > 0 \text{ such that if } \|h\|_X < \delta_3 \text{ then } \|Df(x)(h)\|_Y < \frac{\delta_1}{2} \quad (14.8)$$

As  $U$  is open and  $x \in U$  we have

$$\exists \delta_4 > 0 \text{ such that } x \in B_{\|\cdot\|_X}(x, \delta_4) \subseteq U \quad (14.9)$$

Define now  $\delta = \min(\delta_1, \delta_2, \delta_3, \delta_4)$ . If  $y \in B_{\|\cdot\|_X}(x, \delta) \subseteq B_{\|\cdot\|_X}(x, \delta_4)$  then by the above  $y \in U$  so that  $(y - x) + x = y \in U$  or  $y - x \in U_x$  and as  $\|y - x\|_X < \delta < \delta_2$  we have by 14.7 that  $\|f(x + (y - x)) - f(x) - Df(x)(y - x)\|_Y < \frac{\delta_1}{2}$  giving

$$\forall y \in B_{\|\cdot\|_X}(x, \delta) \text{ we have } \|f(y) - f(x) - Df(x)(y - x)\|_Y < \frac{\delta_1}{2} \quad (14.10)$$

Further if  $y \in B_{\|\cdot\|_X}(x, \delta)$  we have also that  $\|y - x\|_X < \delta_3$  so that by 14.8 we have

$$\forall y \in B_{\|\cdot\|_X}(x, \delta) \text{ we have } \|Df(x)(y - x)\|_Y < \frac{\delta_1}{2} \quad (14.11)$$

Finally if  $y \in B_{\|\cdot\|_X}(x, \delta)$  then we have

$$\begin{aligned} \|f(y) - f(x)\|_Y &= \|f(y) - f(x) - Df(x)(y - x) + Df(x)(y - x)\|_Y \\ &\leq \|f(y) - f(x) - Df(x)(y - x)\|_Y + \|Df(x)(y - x)\|_Y \\ &\stackrel{14.10.14.11}{<} \frac{\delta_1}{2} + \frac{\delta_1}{2} = \delta \end{aligned}$$

which proves that  $f$  is continuous at  $x$ .  $\square$

Let's study now some trivial examples of functions differentiable at a point.

**Example 14.16. (Differential of a constant function)** Let  $\langle X, \|\cdot\|_X \rangle$ ,  $\langle Y, \|\cdot\|_Y \rangle$  be normed spaces,  $y \in Y$ ,  $U \subseteq X$ ,  $V \subseteq Y$  open and  $C_{U, V, y}: U \rightarrow V$  defined by  $C_{U, V, y}(x) = y$  then  $C_{U, V, y}: U \rightarrow V$  is differentiable at every  $x \in U$  and  $DC_{U, V, y}(x) = C_{X, Y, 0}$  where  $C_{X, Y, 0}: X \rightarrow Y$  is defined by  $C_{X, Y, 0}(x) = 0$

**Proof.** Let  $x \in U$  and take  $\varepsilon > 0$  then for  $h \in U_x$  with  $\|h\|_X < 1$  we have  $\|C_{U, V, y}(x + h) - C_{U, V, y}(x) - C_{X, Y, 0}(h)\|_Y = \|y - y - 0\|_Y = \|0\|_Y = 0 \leq \varepsilon \cdot \|h\|_X$ . Also as  $\forall x \in X$  we have  $\|C_{X, Y, 0}(x)\|_Y = \|0\|_Y = 0 \leq \varepsilon \cdot \|x\|_X$  it follows that  $C_{X, Y, 0} \in L(X, Y)$ .  $\square$

**Example 14.17.** Let  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle$  be normed spaces then  $\forall L \in L(X, Y)$  we have  $\forall x \in X$  that  $L$  is differentiable at  $x$  with differential  $DL(x) = L$

**Proof.** If  $\varepsilon > 0$  and  $h \in X_x = X$  with  $\|h\|_X < 1$  then  $\|L(x+h) - L(x) - L(h)\|_Y = \|L(x) + L(h) - L(x) - L(h)\|_Y = \|0\|_Y = 0 \leq \varepsilon \cdot \|h\|_X$  proving that  $L$  is differentiable at every  $x \in U$  with  $DL(x) = L$ .  $\square$

### 14.1.2 Derivative of a function

**Definition 14.18.** Given  $\langle \mathbb{K}, \|\cdot\| \rangle$  (where  $\mathbb{K} = \mathbb{R}$  or  $\mathbb{C}$ ),  $\langle Y, \|\cdot\| \rangle$  a normed space,  $U \subseteq \mathbb{K}$  a open set,  $V \subseteq Y$ ,  $x \in U$  and a function  $f: U \rightarrow V$  then  $f$  has a **derivate** at  $x$  noted by  $f'(x)$  if and only if  $\forall \varepsilon > 0$  there exists a  $\delta > 0$  such that  $\forall h \in U_x$  with  $0 < |h| < \delta$  we have  $\left\| \frac{f(x+h) - f(x)}{h} - f'(x) \right\| \leq \varepsilon$

The above notation suggests that the derivate of a function, if it exists, is unique. The following theorem shows that this is indeed so and also the relation with the differential of a function.

**Theorem 14.19.** Given  $\langle \mathbb{K}, \|\cdot\| \rangle$  (where  $\mathbb{K} = \mathbb{R}$  or  $\mathbb{C}$ ),  $\langle Y, \|\cdot\| \rangle$  a normed space,  $U \subseteq \mathbb{K}$  a open set,  $V \subseteq Y$ ,  $x \in U$  and a function  $f: U \rightarrow V$  then  $f$  is differentiable at  $x$  with differential  $Df(x)$  if and only if  $f$  has a derivate  $f'(x)$  at  $x$ .

Furthermore if  $f$  is differentiable [or equivalent has a derivate] then  $Df(x): \mathbb{K} \rightarrow Y$  is defined by  $h \rightarrow Df(x)(h) = f'(x) \cdot h$  which means that  $f'(x) = Df(x)(1)$ .

Note that as the differential is unique this theorem also proves that the derivate must be unique.

#### Proof.

$\Rightarrow$ . If  $f$  is differentiable with differential  $Df(x)$ , take then  $f'(x) = Df(x)(1)$  then as  $Df(x) \in L(\mathbb{K}, Y)$  we have  $\forall h \in \mathbb{K}$  that  $Df(x)(h) = Df(x)(h \cdot 1) = h \cdot Df(x)(1) = h \cdot f'(x)$ . Using 14.12 (3) we have that given a  $\varepsilon > 0$  there exists a  $\delta > 0$  such that  $\forall h \in U_x$  with  $0 < |h| < \delta$  we have  $\frac{\|f(x+h) - f(x) - Df(x)(h)\|}{|h|} < \varepsilon \Rightarrow \left\| \frac{f(x+h) - f(x) - f'(x) \cdot h}{h} \right\| < \varepsilon \Rightarrow \left\| \frac{f(x+h) - f(x)}{h} - f'(x) \right\| < \varepsilon$  proving that  $f'(x) = Df(x)(1)$  is indeed a derivate of  $f$  at  $x$  and that  $Df(x)(h) = f'(x) \cdot h$

$\Leftarrow$ . Define  $Df(x): \mathbb{K} \rightarrow Y$  by  $h \rightarrow f'(x) \cdot h$  where  $f'(x)$  is a derivate of  $f$ , which is trivially linear and is also continuous [as  $\|Df(x)(h)\|_Y = \|f'(x) \cdot h\|_Y = \|f'(x)\| \cdot |h|$ ].

Further given  $\varepsilon > 0$  there exists a  $\delta > 0$  such that  $\forall h \in U_x$  with  $0 < |h| < \delta$  we have  $\left\| \frac{f(x+h) - f(x)}{h} - f'(x) \right\| \leq \varepsilon \Rightarrow \frac{\|f(x+h) - f(x) - Df(x)(h)\|}{|h|} = \left\| \frac{f(x+h) - f(x) - f'(x) \cdot h}{h} \right\| = \left\| \frac{f(x+h) - f(x)}{h} - f'(x) \right\| \leq \varepsilon$  which proves by 14.12 (3) that  $f$  is differentiable at  $x$  with  $Df(x)(h) = f'(x) \cdot h$ .  $\square$

As the differentiability of a function is a local property the above theorem shows that this is also true for the existence of the derivate of a function.

**Corollary 14.20.** Given  $\langle \mathbb{K}, \|\cdot\| \rangle$  (where  $\mathbb{K} = \mathbb{R}$  or  $\mathbb{C}$ ),  $\langle Y, \|\cdot\| \rangle$  a normed space,  $U \subseteq \mathbb{K}$ ,  $V \subseteq Y$ ,  $x \in U$  a open set, a function  $f: U \rightarrow V$  and  $W \subseteq U$  a open set with  $x \in W$  then  $f$  has a derivate at  $x$  if and only if  $f|_W$  has a derivate at  $x$  and  $(f|_W)'(x) = f'(x)$

**Proof.** This follows from 14.10 and the above theorem.  $\square$

### 14.1.3 Properties of the differential

**Theorem 14.21.** Let  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle$  be normed spaces over  $\mathbb{K}$ ,  $U \subseteq X$  a open set,  $V_f, V_g \subseteq Y$  and  $f: U \rightarrow V_f, g: U \rightarrow V_g$  be functions differentiable at  $x \in U$  then

1.  $f + g: U \rightarrow V_f + V_g$  is differentiable at  $x$  with  $D(f+g)(x) = Df(x) + Dg(x)$
2.  $\forall \alpha \in \mathbb{K}$  then  $\alpha \cdot f: U \rightarrow \alpha \cdot V_f$  is differentiable at  $x$  with  $D(\alpha \cdot f)(x) = \alpha \cdot Df(x)$

**Proof.**

1. Let  $\varepsilon > 0$  then as  $f, g$  are differentiable at  $x$  there exists  $\delta_1, \delta_2 > 0$  such that if  $h \in U_x$  we have if  $\|h\|_X < \delta_1$  then  $\|f(x+h) - f(x) - Df(x)(h)\|_Y \leq \frac{\varepsilon}{2} \cdot \|h\|_X$  and if  $\|h\|_X < \delta_2$  then  $\|g(x+h) - g(x) - Dg(x)(h)\|_Y \leq \frac{\varepsilon}{2} \cdot \|h\|_X$ . So if  $h \in U_x$  and  $\|h\|_X < \delta = \min(\delta_1, \delta_2)$  we have that  $\|(f+g)(x+h) - (f+g)(x) - (Df(x) + Dg(x))(h)\|_Y = \|f(x+h) - f(x) - Df(x)(h) + g(x+h) - g(x) - Dg(x)(h)\|_Y \leq \|f(x+h) - f(x) - Df(x)(h)\|_Y + \|g(x+h) - g(x) - Dg(x)(h)\|_Y \leq \frac{\varepsilon}{2} \cdot \|h\|_X + \frac{\varepsilon}{2} \cdot \|h\|_X = \varepsilon \cdot \|h\|_X$  proving that  $f+g$  is differentiable with differential  $Df(x) + Dg(x)$ .
2. Let  $\varepsilon > 0$  then by the differentiability of  $f$  at  $x$  there exists a  $\delta > 0$  such that if  $h \in U_x$  and  $\|h\|_X < \delta$  we have  $\|f(x+h) - f(x) - Df(x)(h)\|_Y \leq \frac{\varepsilon}{|\alpha|+1} \cdot \|h\|_X$ . So  $\|(\alpha \cdot f)(x+h) - (\alpha \cdot f)(x) - \alpha \cdot Df(x)(h)\|_Y = |\alpha| \cdot \|f(x+h) - f(x) - Df(x)(h)\|_Y \leq |\alpha| \cdot \frac{\varepsilon}{|\alpha|+1} \cdot \|h\|_X < \varepsilon \cdot \|h\|_X$ . Proving that  $\alpha \cdot f$  is indeed differentiable at  $x \in U$  with differential  $\alpha \cdot Df(x)$ .  $\square$

**Theorem 14.22. (The chain rule)** Let  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle, \langle Z, \|\cdot\|_Z \rangle$  be normed spaces,  $U \subseteq X$ ,  $V \subseteq Y$  be open sets,  $R \subseteq Z$ ,  $f: U \rightarrow R$ ,  $g: V \rightarrow S$  functions,  $x \in U \cap f^{-1}(V)$  such that  $f$  is differentiable at  $x$  and  $g$  is differentiable at  $f(x)$  then  $g \circ f: U \cap f^{-1}(V) \rightarrow S$  is differentiable at  $x$  with  $D(g \circ f)(x) = Dg(f(x)) \circ Df(x)$

**Proof.** First as  $f: U \rightarrow Y$  is differentiable at  $x$  we have by 14.15 that  $f$  is continuous at  $x$  so that  $f^{-1}(V)$  is open proving that  $U \cap f^{-1}(V)$  is open [needed for differentiability to make sense]. Using the alternative definition for differentiability (see 14.14) there exists  $\varepsilon$ -mapping  $\varepsilon_f: U_x \rightarrow Y$  and  $\varepsilon_g: V_{f(x)} \rightarrow Y$  such that

$$\forall h \in (U \cap f^{-1}(V))_x \text{ we have } f(x+h) - f(x) - Df(x)(h) = \varepsilon_f(h) \cdot \|h\|_X \quad (14.12)$$

and

$$\forall k \in V_{f(x)} \text{ we have } g(f(x)+k) - g(f(x)) - Dg(f(x))(k) = \varepsilon_g(k) \cdot \|k\|_Y \quad (14.13)$$

If  $h \in (U \cap f^{-1}(V))_x$  then  $x+h \in U \cap f^{-1}(V)$  so that  $(f(x+h) - f(x)) + f(x) = f(x+h) \in V$  proving that

$$\forall h \in (U \cap f^{-1}(V))_x \text{ we have that } f(x+h) - f(x) \in V_{f(x)} \quad (14.14)$$

Combining the above with 14.13 gives then that for  $h \in (U \cap f^{-1}(V))_x$  we have

$$\begin{aligned} \varepsilon_g(f(x+h) - f(x)) \cdot \|f(x+h) - f(x)\|_Y &= g(f(x) + f(x+h) - f(x)) - g(f(x)) - Dg(f(x))(f(x+h) - f(x)) \\ &= g(f(x+h)) - g(f(x)) - Dg(f(x))(f(x+h) - f(x)) \\ &\stackrel{14.12}{=} g(f(x+h)) - g(f(x)) - Dg(f(x))(\varepsilon_f(h) \cdot \|h\|_X + Df(x)(h)) \\ &= g(f(x+h)) - g(f(x)) - Dg(f(x))(\varepsilon_f(h) \cdot \|h\|_X + Dg(f(x))(Df(x)(h))) \\ &= (g \circ f)(x+h) - (g \circ f)(x) - Dg(f(x))(\varepsilon_f(h) \cdot \|h\|_X + Dg(f(x))(Df(x)(h))) \end{aligned}$$

so that  $(g \circ f)(x+h) - (g \circ f)(x) - Dg(f(x))(Df(x)(h)) = \varepsilon_g(f(x+h) - f(x)) \cdot \|f(x+h) - f(x)\|_Y + Dg(f(x))(\varepsilon_f(h) \cdot \|h\|_X)$  hence

$$(g \circ f)(x+h) - (g \circ f)(x) - (Dg(f(x)) \circ Df(x)(h)) = \varepsilon_g(f(x+h) - f(x)) \cdot \|f(x+h) - f(x)\|_Y + Dg(f(x))(\varepsilon_f(h) \cdot \|h\|_X) \quad (14.15)$$

Define now  $\zeta: (U \cap f^{-1}(V))_x \rightarrow Z$  by

$$\zeta(h) = \begin{cases} \varepsilon_g(f(x+h) - f(x)) \cdot \frac{\|f(x+h) - f(x)\|_Y}{\|h\|_X} + Dg(f(x))(\varepsilon_f(h)) & \text{if } h \neq 0 \\ 0 & \text{if } h = 0 \end{cases}$$

then as for  $h=0$  we have that  $(g \circ f)(x+h) - (g \circ f)(x) - (Dg(f(x)) \circ Df(x))(h) = (g \circ f)(x) - (g \circ f)(x) - (Dg(f(x)) \circ Df(x))(0) = 0$  we have by 14.15 that

$$(g \circ f)(x+h) - (g \circ f)(x) - (Dg(f(x)) \circ Df(x))(h) = \zeta(h) \cdot \|h\|_X \quad (14.16)$$

If we prove that  $\zeta$  is continuous at 0 and thus a  $\varepsilon$ -mapping we have by 14.14 and the fact that  $Dg(f(x)) \circ Df(x) \in L(X, Z)$  (see 11.200) that  $g \circ f: U \cap f^{-1}(V) \rightarrow Z$  is differentiable at  $x$  with differential  $Dg(f(x)) \circ Df(x)$ .

To prove continuity of  $\zeta$  take  $\varepsilon > 0$ . By the continuity of  $\varepsilon_f$  at 0 there exists a  $\delta_1 > 0$  such that if  $\|h\|_X < \delta_1$  then  $\|\varepsilon_f(h)\| < 1$  and using 14.12 we have if  $h$  is also a element of  $(U \cap f^{-1}(V))_x$  that  $\|f(x+h) - f(x)\|_Y \leq \|\varepsilon_f(h) \cdot h\|_X + \|Df(x)(h)\|_Y \leq \|\varepsilon_f(h)\| \cdot \|h\|_X + \|Df(x)\| \cdot \|h\|_X < \|h\|_X + \|Df(x)\| \cdot \|h\|_X$  so that

$$\forall h \in (U \cap f^{-1}(V))_x \text{ with } 0 < \|h\|_X < \delta_1 \text{ we have } \frac{\|f(x+h) - f(x)\|_Y}{\|h\|_X} < (1 + \|Df(x)\|) \quad (14.17)$$

By continuity of  $\varepsilon_g$  at 0 there exists a  $\delta_2 > 0$

$$\forall k \in V_{f(x)} \text{ with } \|k\|_Y < \delta_2 \text{ we have } \|\varepsilon_g(k)\|_Z < \frac{\varepsilon}{2 \cdot (1 + \|Df(x)\|)} \quad (14.18)$$

As  $f$  is differentiable at  $x$  and thus by 14.15 continuous at  $x$  there exists a  $\delta_3 > 0$  such that if  $h \in (U \cap f^{-1}(V))_x$   $\xrightarrow{14.14} f(x+h) - f(x) \in V_{f(x)}$  and  $\|h\|_X < \delta_3$  then  $\|f(x+h) - f(x)\|_Y < \delta_2 \xrightarrow{14.18} \|\varepsilon_g(f(x+h) - f(x))\|_Z < \frac{\varepsilon}{2 \cdot (1 + \|Df(x)\|)}$ . Using this with 14.17 gives

$$\text{If } h \in (U \cap f^{-1}(V))_x \text{ with } 0 < \|h\|_X < \min(\delta_1, \delta_3) \text{ we have } \|\varepsilon_g(f(x+h) - f(x))\|_X \cdot \frac{\|f(x+h) - f(x)\|_Y}{\|h\|_X} < \frac{\varepsilon}{2} \quad (14.19)$$

As  $Dg(f(x)) \in L(Y, Z)$  we have that  $\|Dg(f(x))(\varepsilon_f(h))\|_Z \leq \|Dg(f(x))\| \cdot \|\varepsilon_f(h)\|_Y$  and as  $\varepsilon_f$  is continuous at 0 there exists a  $\delta_4$  such that if  $\|h\|_X < \delta_4$  then  $\|\varepsilon_f(h)\|_Y < \frac{\varepsilon}{2 \cdot (1 + \|Dg(f(x))\|)}$  and thus

$$\text{If } \|h\|_X < \delta_4 \text{ then } \|Dg(f(x))(\varepsilon_f(h))\|_Z < \frac{\varepsilon}{2} \quad (14.20)$$

Finally if we take  $\delta = \min(\delta_1, \delta_3, \delta_4)$  we have for  $h \in (U \cap f^{-1}(V))_x$  with  $\|h\|_X < \delta$  that either

$h = \mathbf{0}$ . then  $\|\zeta(h) - 0\|_Z = \|0 - 0\|_X = 0 < \varepsilon$

$h \neq \mathbf{0}$ . then  $0 < \|h\|_X < \delta$  and thus

$$\begin{aligned} \|\zeta(h) - 0\|_Z &= \|\zeta(h)\|_Z \\ &\stackrel{\text{def of } \zeta}{=} \left\| \varepsilon_g(f(x+h) - f(x)) \cdot \frac{\|f(x+h) - f(x)\|_Y}{\|h\|_X} + Dg(f(x))(\varepsilon_f(h)) \right\|_Z \\ &\leq \left\| \varepsilon_g(f(x+h) - f(x)) \cdot \frac{\|f(x+h) - f(x)\|_Y}{\|h\|_X} \right\|_Z + \|Dg(f(x))(\varepsilon_f(h))\|_Z \\ &\stackrel{14.19}{<} \frac{\varepsilon}{2} + \|Dg(f(x))(\varepsilon_f(h))\|_Z \\ &\stackrel{14.20}{<} \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\ &= \varepsilon \\ &= \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \text{ by 14.19 and 14.20.} \end{aligned}$$

So in all cases we have that  $\|\zeta(h) - 0\|_Z < \varepsilon$  proving continuity of  $\zeta$  at 0 and thus that  $\zeta$  is a  $\varepsilon$ -mapping.  $\square$

The above theorem is a generic form of the Chain Rule more know in the form below.

**Corollary 14.23.** *Let  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle, \langle Z, \|\cdot\|_Z \rangle$  be normed spaces over  $\mathbb{K}$ ,  $U \subseteq X$  a open set,  $V \subseteq Y$  a open set,  $R \subseteq Z$ ,  $f: U \rightarrow R$ ,  $g: V \rightarrow S$  functions with  $f(U) \subseteq V$ , such that  $f$  is differentiable at  $x \in U$  and  $g$  is differentiable at  $f(x) \in V$  then  $g \circ f: U \rightarrow S$  is differentiable at  $x$  with  $D(g \circ f)(x) = Dg(f(x)) \circ Df(x)$*

**Proof.** As  $f(U) \subseteq V$  we have that  $U \subseteq f^{-1}(V)$  proving that  $U = U \cap f^{-1}(V)$ , using then the previous theorem we have that  $g \circ f: U \rightarrow S$  is differentiable at  $x \in U$  with  $D(g \circ f)(x) = Df(g(x)) \circ Df(x)$   $\square$

The chain rule can be translated to the concept of derivates

**Corollary 14.24.** *Let  $\mathbb{K}$  be the field  $\mathbb{R}$  or  $\mathbb{C}$ ,  $U, V$  open sets in  $\mathbb{K}$ ,  $f: U \rightarrow \mathbb{K}$ ,  $g: V \rightarrow \mathbb{R}$  functions such that for  $x \in U \cap f^{-1}(V)$   $f$  has a derivate on  $x$  and  $g$  has a derivate on  $f(x)$  then  $g \circ f$  has a derivate on  $x$  and  $(g \circ f)'(x) = g'(f(x)) \cdot f'(x)$*

**Proof.** Using 14.19 we have that  $f$  is differentiable at  $x$  with  $f'(x) = Df(x)(1)$  and  $g$  is differentiable at  $f(x)$  with  $Dg(f(x))(1)$ . By the chain rule (see 14.22) we have that  $g \circ f$  is differentiable at  $x$  with  $D(g \circ f)(x) = Dg(f(x)) \circ Df(x)$ . Using then 14.19 again we have that  $g \circ f$  has a derivate  $(g \circ f)'(x)$  at  $x$  with

$$\begin{aligned} (g \circ f)'(x) &= (Dg(f(x)) \circ Df(x))(1) \\ &= (Dg(f(x)))(Df(x)(1)) \\ &= (Dg(f(x)))(f'(x)) \\ &\stackrel{14.19}{=} g'(f(x)) \cdot f'(x) \end{aligned}$$

$\square$

#### 14.1.4 Partial differentiation

**Definition 14.25.** *Given a finite family  $\{X_i\}_{i \in \{1, \dots, n\}}$  of sets then given a  $i \in \{1, \dots, n\}$  and  $x \in \prod_{j \in \{1, \dots, n\} \setminus \{i\}} X_j$  we define  $(i \rightarrow x): X_i \rightarrow \prod_{i \in \{1, \dots, n\}} X_i$  by  $t \mapsto (i \rightarrow x)(t)$  where  $\forall j \in \{1, \dots, n\}$   $((i \rightarrow x)(t))_j = \begin{cases} x_j & \text{if } j \in \{1, \dots, n\} \setminus \{i\} \\ t & \text{if } j = i \end{cases}$ .*

**Remark 14.26.** Essentially  $(i \rightarrow x)(t)$  substitute the  $i$ -the element of  $x$  by  $t$ , so we have that  $(i \rightarrow (x_1, \dots, x_i, x_i, x_{i+1}, \dots, x_n))(t) = (x_1, \dots, x_{i-1}, t, x_i, \dots, x_n)$

We have the following extension of the Kronecker Delta function;

**Definition 14.27.** *Let  $\{X_i\}_{i \in \{1, \dots, n\}}$  a family of vector spaces then given  $i, j$  and  $x \in X_i$  then  $\delta_{i,j}$ :  $X_i \rightarrow \bigcup_{j \in \{1, \dots, n\}} X_j$  is defined by  $\delta_{i,j}(x) = \begin{cases} x & \text{if } i = j \\ 0_j & \text{if } i \neq j \end{cases}$  where  $0_j$  is the neutral element in  $X_j$*

**Proposition 14.28.** *Given a finite family  $\{X_i\}_{i \in \{1, \dots, n\}}$  of sets,  $i \in \{1, \dots, n\}$ ,  $U \subseteq \prod_{i \in \{1, \dots, n\}} X_i$  and  $x \in U$  then  $(i \rightarrow x)^{-1}(U) \subseteq \pi_i(U)$*

**Proof.** If  $y \in (i \rightarrow x)^{-1}(U)$  then  $(i \rightarrow x)(y) \in U$  so that  $y = ((i \rightarrow x)(y))_i = \pi_i((i \rightarrow x)(y)) \in \pi_i(U)$  proving that

$$(i \rightarrow x)^{-1}(U) \subseteq \pi_i(U) \tag{14.21} \quad \square$$

**Proposition 14.29.** *Let  $\{X_i\}_{i \in \{1, \dots, n\}}$  be a finite family of vector spaces over the field  $\mathbb{K}$ ,  $x \in \prod_{i \in \{1, \dots, n\}} X_i$  then  $\forall i \in \{1, \dots, n\}$  we have*

1.  $(i \rightarrow x)(x_i) = x$
2.  $(i \rightarrow x)(t) - (i \rightarrow x)(s) = (i \rightarrow 0)(t - s)$
3.  $(i \rightarrow x)(t) = (i \rightarrow 0)(t) + (i \rightarrow x)(0)$
4. If  $y \in \prod_{i \in \{1, \dots, n\}} X_i$  then  $y + (i \rightarrow 0)(t) = (i \rightarrow y)(t + y_i)$

**Proof.**

1.  $((i \rightarrow x)(x_i))_i = \begin{cases} \stackrel{=} {x_i} & \text{if } i = i \\ x_i & \text{if } i \in \{1, \dots, n\} \setminus \{i\} \end{cases}$

2. Let  $k \in \{1, \dots, n\}$  then

$$\begin{aligned}
 ((i \rightarrow x)(t) - (i \rightarrow x)(s))_k &= ((i \rightarrow x)(t))_k - ((i \rightarrow x)(s))_k \\
 &= \begin{cases} x_k \text{ if } k \in \{1, \dots, n\} \setminus \{i\} \\ t \text{ if } k = i \end{cases} - \begin{cases} x_k \text{ if } k \in \{1, \dots, n\} \setminus \{i\} \\ s \text{ if } k = i \end{cases} \\
 &= \begin{cases} 0_i \text{ if } k \in \{1, \dots, n\} \setminus \{i\} \\ t - s \text{ if } k = i \end{cases} \\
 &= (i \rightarrow 0)(s - t)
 \end{aligned}$$

3.  $(i \rightarrow x)(t) - (i \rightarrow x)(0) \stackrel{(2)}{=} (i \rightarrow 0)(t - 0) = (i \rightarrow 0)(t)$  giving  $(i \rightarrow x)(t) = (i \rightarrow 0)(t) + (i \rightarrow x)(0)$

4. Let  $k \in \{1, \dots, n\}$  then

$$\begin{aligned}
 (y + (i \rightarrow 0)(t))_k &= y_k + ((i \rightarrow 0)(t))_k \\
 &= y_k + \begin{cases} 0 \text{ if } k \in \{1, \dots, n\} \setminus \{1\} \\ t \text{ if } k = 1 \end{cases} \\
 &= \begin{cases} y_k \text{ if } k \in \{1, \dots, n\} \setminus \{1\} \\ y_k + t \text{ if } k = 1 \end{cases} \\
 &= (i \rightarrow y)(y_k + t)
 \end{aligned}$$

□

**Theorem 14.30.** Given a finite family  $\{X_i\}_{i \in \{1, \dots, n\}}$  of vector spaces over the field  $\mathbb{K}$  then given  $a \in \{1, \dots, n\}$  and  $0 \in \prod_{j \in \{1, \dots, n\} \setminus \{i\}} X_j$ , where  $0_j \in X_j$  is the neutral element of  $X_j$  then

1.  $\forall i, k \in \{1, \dots, n\}$  and  $x \in X_i$  then  $((i \rightarrow 0)(x))_k = \delta_{i,k}(x)$
2. Let  $i \in \{1, \dots, n\}$  and  $x \in X_i$  then  $\|\delta_{i,j}(x)\|_j \leq \|x\|_i \forall j \in \{1, \dots, n\}$
3.  $\forall i \in \{1, \dots, n\}$  we have that  $(i \rightarrow 0) \in L(X_i, \prod_{i \in \{1, \dots, n\}} X_i)$ , further  $\forall x \in X_i$  we have  $\|(0 \rightarrow i)(x)\| = \|x\|_i$  [using the maximum norm  $\|\cdot\|$  on  $\prod_{i \in \{1, \dots, n\}} X_i$ ]
4.  $\forall x \in \prod_{j \in \{1, \dots, n\}} X_j$  and  $i \in \{1, \dots, n\}$  we have that  $(i \rightarrow x)$  is continuous

**Proof.**

1. Let  $i \in \{1, \dots, n\}$  then for  $k \in \{1, \dots, n\}$  we have either

$$\mathbf{k} \neq \mathbf{i}. \text{ then } ((i \rightarrow 0)(x)) = \begin{cases} 0_k \text{ if } k \in \{1, \dots, n\} \setminus \{i\} \\ x \text{ if } k = i \end{cases} = 0_k = \delta_{i,k}(x)$$

$$\mathbf{k} = \mathbf{i}. \text{ then } ((i \rightarrow 0)(x)) = \begin{cases} 0_k \text{ if } k \in \{1, \dots, n\} \setminus \{i\} \\ x \text{ if } k = i \end{cases} = x = \delta_{i,k}(x)$$

2. Let  $i \in \{1, \dots, n\}$  and  $x \in X_i$  then for  $j \in \{1, \dots, n\}$  we have either:

$$\mathbf{j} = \mathbf{i}. \text{ then } \|\delta_{i,j}(x)\|_j = \|x\|_i$$

$$\mathbf{j} \neq \mathbf{i}. \text{ then } \|\delta_{i,j}(x)\|_j = \|0_j\|_j = 0 \leq \|x\|_i$$

so in all cases we have  $\|\delta_{i,j}(x)\|_j = \begin{cases} 0 \text{ if } j \neq i \\ \|x_j\| \text{ if } j = i \end{cases} \leq \|x\|_i$

3. We must prove that  $(i \rightarrow 0)$  is linear and continuous. Let  $i \in \{1, \dots, n\}$  then

- a. Let  $\alpha, \beta \in \mathbb{K}$  and  $x, y \in X_i$  then given  $k \in \{1, \dots, n\}$  we have

$$\begin{aligned}
 ((i \rightarrow 0)(\alpha \cdot x + \beta \cdot y))_k &= \begin{cases} 0 \text{ if } k \neq i \\ \alpha \cdot x + \beta \cdot y \text{ if } k = i \end{cases} \\
 &= \begin{cases} \alpha \cdot 0 + \beta \cdot 0 \text{ if } k \neq i \\ \alpha \cdot x + \beta \cdot y \text{ if } k = i \end{cases} \\
 &= \alpha \cdot \begin{cases} 0 \text{ if } k \neq i \\ x \text{ if } k = i \end{cases} + \beta \cdot \begin{cases} 0 \text{ if } k \neq i \\ y \text{ if } k = i \end{cases} \\
 &= (\alpha \cdot (i \rightarrow 0)(x))_k + (\beta \cdot (i \rightarrow 0)(y))_k \\
 &= (\alpha \cdot (i \rightarrow 0)(x) + \beta \cdot (i \rightarrow 0)(y))_k
 \end{aligned}$$

proving that  $(i \rightarrow 0)(\alpha \cdot x + \beta \cdot y) = \alpha \cdot (i \rightarrow 0)(x) + \beta \cdot (i \rightarrow 0)(y)$ . So  $(i \rightarrow 0)$  is linear.

b. Let  $x \in X_i$  then we have

$$\begin{aligned} \|(i \rightarrow 0)(x)\| &= \max(\{\|((i \rightarrow 0)(x))_j\|_{j \in \{1, \dots, n\}} \mid j \in \{1, \dots, n\}\}) \\ &\stackrel{(1)}{=} \max(\{\|\delta_{i,j}(x)\|_j \mid j \in \{1, \dots, n\}\}) \\ &\stackrel{(2)}{=} \max\left(\left\{\begin{array}{l} 0 \text{ if } i \neq j \\ \|x\|_i \text{ if } i = j \end{array} \mid j \in \{1, \dots, n\} \right\}\right) \\ &\stackrel{0 \leq \|x\|_i}{=} \|x\|_i \end{aligned}$$

which by 11.180 proves that  $(i \rightarrow 0) \in L(X_i, \prod_{j \in \{1, \dots, n\}} X_j)$  and  $\|(i \rightarrow 0)(x)\| = \|x\|_i$

4. Let  $\varepsilon > 0$ ,  $x \in \prod_{j \in \{1, \dots, n\}} X_j$ ,  $s \in X_i$  take then  $t \in X_i$  such that  $\|s - t\|_i < \varepsilon$  then we have

$$\begin{aligned} \|(i \rightarrow x)(s) - (i \rightarrow x)(t)\| &\stackrel{14.29 \text{ (2)}}{=} \|(0 \rightarrow x)(s - t)\| \\ &\stackrel{(3)}{=} \|s - t\|_i \\ &< \varepsilon \end{aligned}$$

proving continuity.  $\square$

**Theorem 14.31.** *Given a finite family  $\{X_i\}_{i \in \{1, \dots, n\}}$  of vector spaces over the field  $\mathbb{K}$  then for every  $x \in \prod_{i \in \{1, \dots, n\}} X_i$  we have  $x = \sum_{i=1}^n (i \rightarrow 0)(x_i)$*

**Proof.** Let  $k \in \{1, \dots, n\}$  then

$$\begin{aligned} \left( \sum_{i=1}^n (i \rightarrow 0)(x_i) \right)_k &\stackrel{10.10}{=} \sum_{i=1}^n ((i \rightarrow 0)(x_i))_k \\ &\stackrel{14.30}{=} \sum_{i=1}^n \delta_{i,k}(x_i) \\ &= x_k \end{aligned}$$

proving that  $\sum_{i=1}^n (i \rightarrow 0)(x_i) = x$ .  $\square$

**Theorem 14.32.** *Given a finite family  $\{\langle X_i, \|\cdot\|_i \rangle\}_{i \in \{1, \dots, n\}}$  of vector spaces over the field  $\mathbb{K}$ ,  $X = \prod_{i \in \{1, \dots, n\}} X_i$  the product space together with the maximum norm  $\|\cdot\|$  [see 11.82] [so that  $\langle X, \|\cdot\|\rangle$  forms a normed vector space over  $\mathbb{K}$ ] then given  $i \in \{1, \dots, n\}$ ,  $t \in X_i$  and  $x \in \prod_{j \in \{1, \dots, n\} \setminus \{i\}} X_j$  we have that  $(i \rightarrow x): X_i \rightarrow \prod_{j \in \{1, \dots, n\}} X_j$  is differentiable at  $t$  with  $D(i \rightarrow x)(t) = (i \rightarrow 0)$ .*

**Proof.** Let  $i \in \{1, \dots, n\}$  then using 14.30 we have that

$$(i \rightarrow 0) \in L\left(X_i, \prod_{j \in \{1, \dots, n\}} X_j\right)$$

Take now  $\varepsilon > 0$ ,  $t \in X_i$  then if  $h \in (X_i)_t = X_i$  with  $\|h\|_i < 1$  we have using 14.29 that  $(i \rightarrow x)(t + h) - (i \rightarrow x)(t) = (i \rightarrow 0)(h)$ . Hence  $\|(i \rightarrow x)(t + h) - (i \rightarrow x)(t) - (i \rightarrow 0)(h)\| = \|0\| = 0 \leq \varepsilon$  proving as  $(i \rightarrow 0) \in L(X_i, \prod_{i \in \{1, \dots, n\}} X_i)$  [see 14.30] that  $(i \rightarrow x)$  is differentiable at  $t$  with differential  $(i \rightarrow 0)$ .  $\square$

**Corollary 14.33.** *Let  $\{\langle X_i, \|\cdot\|_i \rangle\}_{i \in \{1, \dots, n\}}$  be a finite family of normed vector space over the field  $\mathbb{K}$ ,  $U \subseteq \prod_{i \in \{1, \dots, n\}} X_i$ ,  $x \in U$  and  $i \in \{1, \dots, n\}$  then*

1.  $(i \rightarrow x)^{-1}(U)$  is a open set in  $\prod_{i \in \{1, \dots, n\}} X_i$
2.  $\pi_i(U)$  is a open set in  $X_i$
3.  $(i \rightarrow x)|_{(i \rightarrow x)^{-1}(U)}: (i \rightarrow x)^{-1}(U) \rightarrow \prod_{j \in \{1, \dots, n\}} X_j$  is differentiable at  $t \in U$  with  $D(i \rightarrow x)|_{(i \rightarrow x)^{-1}(U)}(t) = (i \rightarrow 0)$

**Proof.**

1. This follows from 14.30 (4)
2. This follows from the fact that  $\pi_i$  is open [see 11.148]
3. This follows from 14.32, (2) and 14.10

□

**Definition 14.34.** Let  $n \in \mathbb{N}_0 \setminus \{1\}$   $\{\langle X_i, \|\cdot\|_i \rangle\}_{i \in \{1, \dots, n\}}$  be a finite family of normed spaces,  $\langle Y, \|\cdot\|_Y \rangle$  a normed space,  $X = \langle \prod_{i \in \{1, \dots, n\}} X_i, \|\cdot\| \rangle$  the normed space equipped with the maximum norm [see 11.82],  $U \subseteq \prod_{i \in \{1, \dots, n\}} X_i$  a open set,  $V \subseteq Y$ ,  $x \in U$ ,  $f: U \rightarrow V$  a function and  $i \in \{1, \dots, n\}$  then  $f: U \rightarrow Y$  is  $i$ -partial differentiable at  $x$  with  $i$ -partial differential  $D_i f(x) \in L(X_i, Y)$  if  $f \circ (i \rightarrow x)|_{(i \rightarrow x)^{-1}(U)}: (i \rightarrow x)^{-1}(U) \rightarrow V$  is differentiable at  $x_i$  with  $D_i f(x) = D(f \circ (i \rightarrow x)|_{(i \rightarrow x)^{-1}(U)})(x_i)$  [This make sense as by 14.33  $(i \rightarrow x)^{-1}(U)$  is open in  $X_i$  and  $(i \rightarrow x)(x_i) \stackrel{14.29}{=} x \in U$  so that  $x_i \in (i \rightarrow x)^{-1}(U)$ ].

The next theorem shows that the partial differentials exists if a function is differentiable. Note that the opposite is not always true.

**Theorem 14.35.** Let  $n \in \mathbb{N}_0 \setminus \{1\}$   $\{\langle X_i, \|\cdot\|_i \rangle\}_{i \in \{1, \dots, n\}}$  be a finite family of normed spaces,  $\langle Y, \|\cdot\|_Y \rangle$  a normed space,  $\langle \prod_{i \in \{1, \dots, n\}} X_i, \|\cdot\| \rangle$  the normed space equipped with the maximum norm [see 11.82],  $U \subseteq \prod_{i \in \{1, \dots, n\}} X_i$  a open set,  $V \subseteq Y$ ,  $x \in U$  and  $f: U \rightarrow V$  differentiable at  $x \in U$  then

1.  $\forall i \in \{1, \dots, n\}$   $f$  is  $i$ -partial differentiable at  $x$  with  $D_i f(x) = Df(x) \circ (i \rightarrow 0)$
2.  $Df(x) = \sum_{i=1}^n D_i f(x) \circ \pi_i$  [or  $\forall h \in \prod_{i \in \{1, \dots, n\}} X_i$  we have  $Df(x)(h) = \sum_{i=1}^n D_i f(x)(h_i)$  [or  $Df(x) = \sum_{i=1}^n D_i f(x) \circ \pi_i$ ]

**Proof.**

1. Using 14.33 we have that  $(i \rightarrow x)|_{\pi_i(U)}: \pi_i(U) \rightarrow \prod_{i \in \{1, \dots, n\}} X_i$  is differentiable at  $x_i$  with  $D(i \rightarrow x)|_{\pi_i(U)}(x_i) = (i \rightarrow 0)$ . Using the chain rule 14.22 we have that  $f \circ (i \rightarrow x)|_{(i \rightarrow x)^{-1}(U)}: (i \rightarrow x)^{-1}(U) \rightarrow Y$  is differentiable at  $x_i$  with

$$\begin{aligned} D_i f(x) &\stackrel{\text{def}}{=} D(f \circ (i \rightarrow x)|_{(i \rightarrow x)^{-1}(U)})(x_i) \\ &= Df((i \rightarrow x)|_{(i \rightarrow x)^{-1}(U)}(x_i)) \circ D(i \rightarrow x)|_{(i \rightarrow x)^{-1}(U)}(x_i) \\ &\stackrel{14.29}{=} Df(x) \circ D(i \rightarrow x)|_{(i \rightarrow x)^{-1}(U)}(x_i) \\ &= Df(x) \circ (i \rightarrow 0) \end{aligned}$$

2. Let  $h \in X$  then we have

$$\begin{aligned} \sum_{i=1}^n D_i f(x)(h_i) &= \sum_{i=1}^n Df(x)((i \rightarrow 0)(h_i)) \\ &\stackrel{Df(x) \text{ is linear}}{=} Df(x) \left( \sum_{i=1}^n (i \rightarrow 0)(h_i) \right) \\ &\stackrel{14.31}{=} Df(x)(h) \end{aligned}$$

□

The following shows a extension of the chain rule to partial differentials

**Theorem 14.36.** Let  $n \in \mathbb{N}_0 \setminus \{1\}$ ,  $\{\langle X_i, \|\cdot\|_i \rangle\}_{i \in \{1, \dots, n\}}$  a finite family of normed spaces,  $\langle Y, \|\cdot\|_Y \rangle$ ,  $\langle Z, \|\cdot\|_Z \rangle$  normed spaces,  $\langle \prod_{i \in \{1, \dots, n\}} X_i, \|\cdot\| \rangle$  be the normed space with the maximum norm  $\|\cdot\|$  [see 11.82],  $U \subseteq \prod_{i \in \{1, \dots, n\}} X_i$  open,  $R \subseteq Y$ ,  $S \subseteq Z$   $i \in \{1, \dots, n\}$  and  $f: U \rightarrow R$  a function with a  $i$ -partial differential on  $U$ ,  $f(U) \subseteq V$  a open set in  $Y$ ,  $g: V \rightarrow S$  a differentiable function on  $V$  then  $g \circ f$  has a  $i$ -partial differential with  $D_i(g \circ f)(x) = Dg(f(x)) \circ D_i f(x)$

**Proof.** As  $(g \circ f) \circ (i \rightarrow x)|_{(i \rightarrow x)^{-1}(U)} = g \circ (f \circ (i \rightarrow x)|_{(i \rightarrow x)^{-1}(U)})$  and  $g$  and  $(f \circ (i \rightarrow x)|_{(i \rightarrow x)^{-1}(U)})$  are differentiable it follows from the chain rule 14.83 that  $(g \circ f) \circ (i \rightarrow x)|_{(i \rightarrow x)^{-1}(U)}$  is differentiable and

$$\begin{aligned}
 D_i(g \circ f)(x) &= D((g \circ f) \circ (i \rightarrow x)|_{(i \rightarrow x)^{-1}(U)})(x_i) \\
 &\stackrel{14.22}{=} D(g \circ (f \circ (i \rightarrow x)|_{(i \rightarrow x)^{-1}(U)}))(x_i) \\
 &= Dg((f \circ (i \rightarrow x)|_{(i \rightarrow x)^{-1}(U)})(x_i)) \circ D(f \circ (i \rightarrow x)|_{(i \rightarrow x)^{-1}(U)})(x_i) \\
 &= Dg(f(i \rightarrow x)|_{(i \rightarrow x)^{-1}(U)}(x_i)) \circ D(f \circ (i \rightarrow x)|_{(i \rightarrow x)^{-1}(U)})(x_i) \\
 &\stackrel{14.29}{=} Dg(f(x)) \circ D(f \circ (i \rightarrow x)|_{(i \rightarrow x)^{-1}(U)})(x_i) \\
 &= Dg(f(x)) \circ D_i f(x)
 \end{aligned}$$

□

The partial differential has many of the same properties as a differential as is expressed in the following theorem.

**Theorem 14.37.** Let  $n \in \mathbb{N}_0 \setminus \{1\}$ ,  $\{X_i, \|\cdot\|_i\}_{i \in \{1, \dots, n\}}$  a finite family of normed spaces,  $\langle Y, \|\cdot\|_Y \rangle$  a normed space, then we have for  $i \in \{1, \dots, n\}$ ,  $U \subseteq \prod_i X_i$  and  $x \in \prod_{i \in \{1, \dots, n\}} X_i$  that

1. If  $V_f, V_g \subseteq Y$  and  $f: U \rightarrow V_f$ ,  $g: U \rightarrow V_g$  is such that  $D_i f(x)$ ,  $D_i g(x)$  exists then for  $f + g: U \rightarrow V_f + V_g$   $D_i(f + g)(x)$  exists and  $D_i(f + g)(x) = D_i f(x) + D_i g(x)$
2. If  $V \subseteq Y$  and  $f: U \rightarrow V$  is such that  $D_i f(x)$  exists and  $\alpha \in \mathbb{K}$  then for  $\alpha \cdot f: U \rightarrow \alpha \cdot V$   $D_i(\alpha \cdot f)$  exists and  $D_i(\alpha \cdot f)(x) = \alpha \cdot D_i f(x)$
3. If  $V \subseteq Y$  and  $f: U \rightarrow V$  is such that  $f_i \circ (i \rightarrow x)$  is a constant function then  $D_i f(x)$  exists and  $D_i f(x) = C_{X_i, Y, 0}$
4. If  $f: \prod_{i \in \{1, \dots, n\}} X_i \rightarrow Y$  is such that  $f_i \circ (i \rightarrow x) \in L(X_i, Y)$  then  $D_i f(x)$  exists and  $D_i f(x) = f_i \circ (i \rightarrow x)$

**Proof.** The proof makes use of the fact that  $D f_i(x) = D(f \circ (i \rightarrow x))(x_i)$  and the properties of the differential.

1. By definition we have that  $D_i f(x) = D(f \circ (i \rightarrow x)|_{(i \rightarrow x)^{-1}(U)})(x_i)$  and  $D_i g(x) = D(g \circ (i \rightarrow x)|_{(i \rightarrow x)^{-1}(U)})(x_i)$  exist. Using 14.21 we have then that  $D(f \circ (i \rightarrow x)|_{(i \rightarrow x)^{-1}(U)} + g \circ (i \rightarrow x)|_{(i \rightarrow x)^{-1}(U)})(x_i)$  exist and  $D(f \circ (i \rightarrow x)|_{(i \rightarrow x)^{-1}(U)} + g \circ (i \rightarrow x)|_{(i \rightarrow x)^{-1}(U)})(x_i) = D(f \circ (i \rightarrow x)|_{(i \rightarrow x)^{-1}(U)})(x_i) + D(g \circ (i \rightarrow x)|_{(i \rightarrow x)^{-1}(U)})(x_i) = D_i f(x) + D_i g(x)$ . As  $\forall y \in U$

$$\begin{aligned}
 (f \circ (i \rightarrow x)|_{(i \rightarrow x)^{-1}(U)} + g \circ (i \rightarrow x)|_{(i \rightarrow x)^{-1}(U)})(y) &= f((i \rightarrow x)(y)) + g((i \rightarrow x)(y)) \\
 &= (f + g)((i \rightarrow x)(y)) \\
 &= ((f + g) \circ (i \rightarrow x)|_{(i \rightarrow x)^{-1}(U)})(y)
 \end{aligned}$$

It follows that  $f \circ (i \rightarrow x)|_{(i \rightarrow x)^{-1}(U)} + g \circ (i \rightarrow x)|_{(i \rightarrow x)^{-1}(U)} = (f + g) \circ (i \rightarrow x)|_{(i \rightarrow x)^{-1}(U)}$  proving that  $D_i(f + g)(x)$  exist and that  $D_i(f + g)(x) = D_i f(x) + D_i g(x)$ .

2. By definition we have that  $D_i f(x) = D(f \circ (i \rightarrow x)|_{(i \rightarrow x)^{-1}(U)})(x_i)$  exist. Using 14.21 we have then that  $D(\alpha \cdot (f \circ (i \rightarrow x)|_{(i \rightarrow x)^{-1}(U)}))(x_i)$  exist and  $D(\alpha \cdot (f \circ (i \rightarrow x)|_{(i \rightarrow x)^{-1}(U)}))(x_i) = \alpha \cdot D(f \circ (i \rightarrow x)|_{(i \rightarrow x)^{-1}(U)})(x_i) = \alpha \cdot D_i f(x)$ . As  $\forall y \in U$

$$\begin{aligned}
 ((\alpha \cdot f) \circ (i \rightarrow x)|_{(i \rightarrow x)^{-1}(U)})(y) &= (\alpha \cdot f)((i \rightarrow x)(y)) \\
 &= \alpha \cdot f((i \rightarrow x)(y)) \\
 &= \alpha \cdot (f \circ (i \rightarrow x)|_{(i \rightarrow x)^{-1}(U)})(y)
 \end{aligned}$$

It follows that  $(\alpha \cdot f) \circ (i \rightarrow x)|_{(i \rightarrow x)^{-1}(U)} = \alpha \cdot (f \circ (i \rightarrow x)|_{(i \rightarrow x)^{-1}(U)})$  proving that  $D_i(\alpha \cdot f)(x)$  exists and  $D_i(\alpha \cdot f)(x) = \alpha \cdot D_i f(x)$ .

3. As  $f \circ (i \rightarrow x)|_{(i \rightarrow x)^{-1}(U)}$  is constant on  $U$  we have by 14.16 that  $D(f \circ (i \rightarrow x)|_{(i \rightarrow x)^{-1}(U)})(x_i)$  exist and  $D(f \circ (i \rightarrow x)|_{(i \rightarrow x)^{-1}(U)})(x_i) = C_{X_i, Y, 0}$  proving that  $D_i f(x) = C_{X_i, Y, 0}$ .

4. As  $f \circ (i \rightarrow x) \in L(X_i, Y)$  we have by 14.17 that  $D(f \circ (i \rightarrow x))(x_i) = f \circ (i \rightarrow x)$  proving that  $D_i f(x)$  exists and  $D_i f(x) = f \circ (i \rightarrow x)$   $\square$

Just like we have defined the derivate of a function, we define now the partial derivate of a function.

**Definition 14.38.** Let  $n \in \mathbb{N}_0 \setminus \{1\}$ ,  $\langle \mathbb{K}^n, \|\cdot\| \rangle$  be the vector space  $\mathbb{K}^n$  equipped with the maximum norm,  $\langle Y, \|\cdot\|_Y \rangle$  a normed space,  $U \subseteq \mathbb{K}^n$  a open set,  $V \subseteq Y$ ,  $x \in U$ ,  $i \in \{1, \dots, n\}$  and  $f: \mathbb{K}^n \rightarrow V$  then  $f$  has a  **$i$ -partial derivate at  $x$**  noted by  $\partial_i f(x) \in Y$  if and only if  $\forall \varepsilon > 0$  there exists a  $\delta > 0$  such that if  $0 < |h| < \delta$  and  $h \in ((i \rightarrow x)^{-1}(U))_x$  then  $\left\| \frac{f((i \rightarrow x)(x_i + h)) - f(x)}{h} - \partial_i f(x) \right\|_Y < \varepsilon$ .

The above definition suggest that the partial derivate is unique which is indeed true as is expressed in the following theorem which also show the relation with the partial differential.

**Theorem 14.39.** Let  $n \in \mathbb{N}_0 \setminus \{1\}$ ,  $\langle \mathbb{K}^n, \|\cdot\| \rangle$  be the vector space  $\mathbb{K}^n$  equipped with the maximum norm,  $\langle Y, \|\cdot\|_Y \rangle$  a normed space,  $U \subseteq \mathbb{K}^n$  a open set,  $V \subseteq Y$ ,  $x \in U$ ,  $i \in \{1, \dots, n\}$  and  $f: \mathbb{K}^n \rightarrow V$  then  $f$  has a  $i$ -partial derivate at  $x$  if and only if  $f$  has a  $i$ -partial differential. Further if the  $i$ -partial derivate exists [or the  $i$ -partial differential exists] at  $x$  then  $\partial_i f(x) = D_i f(x)(1)$  so that  $D_i f(x)(h) = D_i f(x)(h \cdot 1) = h \cdot D_i f(x)(1)$ . Note as the  $i$ -partial differential is unique this proves also that the  $i$ -partial derivate is unique.

**Proof.**

$\Rightarrow$ . Assume that  $\partial_i f(x)$  exist. Then given  $\varepsilon > 0$  there exists a  $\delta > 0$  such that  $\forall h \in ((i \rightarrow x)^{-1}(U))_x$  with  $0 < |h| < \delta$  we have that

$$\left\| \frac{f((i \rightarrow x)(x_i + h)) - f(x)}{h} - \partial_i f(x) \right\|_Y < \varepsilon$$

As  $\left\| \frac{f((i \rightarrow x)_{(i \rightarrow x)^{-1}(U)}(x_i + h)) - f((i \rightarrow x)_{(i \rightarrow x)^{-1}(U)}(x_i))}{h} - \partial_i f(x) \right\|_Y \stackrel{14.29}{=} \left\| \frac{f((i \rightarrow x)(x_i + h)) - f(x)}{h} - \partial_i f(x) \right\|_Y < \varepsilon$  proving that the derivate of  $f \circ (i \rightarrow x)_{(i \rightarrow x)^{-1}(U)}$  at  $x_i$  exists and  $(f \circ (i \rightarrow x)_{(i \rightarrow x)^{-1}(U)})'(x_i) = \partial_i f(x)$ . Using 14.19 we have that  $D(f \circ (i \rightarrow x)_{(i \rightarrow x)^{-1}(U)})(x_i)$  exists and  $D(f \circ (i \rightarrow x)_{(i \rightarrow x)^{-1}(U)})(x_i)(1) = (f \circ (i \rightarrow x))'(x_i) = \partial_i f(x)$ . Using the definition of the partial differential (see 14.34) it follows that

$$D_i f(x) \text{ exists and } D_i f(x)(1) = \partial_i f(x)$$

$\Leftarrow$ . Assume that  $D_i f(x)$  exists at  $x$  then by definition  $D(f \circ (i \rightarrow x)_{(i \rightarrow x)^{-1}(U)})(x_i)$  exists. Using 14.19 we have that  $(f \circ (i \rightarrow x)_{(i \rightarrow x)^{-1}(U)})'(x_i)$  exists so that given  $\varepsilon > 0$  there exists a  $\delta > 0$  such that  $\forall h \in ((i \rightarrow x)^{-1}(U))_x$  with  $0 < |h| < \delta$  we have  $\|(f \circ (i \rightarrow x)_{(i \rightarrow x)^{-1}(U)}(x_i + h)) - (f \circ (i \rightarrow x)_{(i \rightarrow x)^{-1}(U)}(x_i)) - (f \circ (i \rightarrow x)_{(i \rightarrow x)^{-1}(U)})'(x_i)\|_Y < \varepsilon$  which as  $(f \circ (i \rightarrow x)_{(i \rightarrow x)^{-1}(U)}(x_i + h)) = f((i \rightarrow x)(x_i + h))$ ,  $f \circ (i \rightarrow x)_{(i \rightarrow x)^{-1}(U)}(x_i) = f((i \rightarrow x)(x_i))$  proves that  $\partial_i f(x)$  exists.  $\square$

We have then the equivalence of theorem 14.35

**Theorem 14.40.** Let  $n \in \mathbb{N}_0 \setminus \{1\}$ ,  $\langle \mathbb{K}^n, \|\cdot\| \rangle$  be the vector space  $\mathbb{K}^n$  equipped with the maximum norm based on the absolute value norm,  $\langle Y, \|\cdot\|_Y \rangle$  a normed space,  $V \subseteq Y$ ,  $U \subseteq \mathbb{K}^n$  a open set,  $x \in U$  and  $f: U \rightarrow V$  a function that is differentiable at  $x$  then

1.  $\forall i \in \{1, \dots, n\}$   $f$  has a partial derivate  $\partial_i f(x)$  and the partial differential  $D_i f(x) \in L(\mathbb{K}, Y)$  is defined by  $k \rightarrow k \cdot \partial_i f(x)$  [so  $D_i f(x)(1) = \partial_i f(x)$ ]

2. If  $\{e_i\}_{i \in \{1, \dots, n\}}$  is the canonical base on  $\mathbb{K}^n$  (defined by  $(e_i)_j = \delta_{i,j}$ ) then  $\partial_i f(x) = D f(x)(e_i)$  and  $D f(x)(h) = \sum_{i \in \{1, \dots, n\}} h_i \cdot D f(x)(e_i) = \sum_{i \in \{1, \dots, n\}} h_i \cdot \partial_i f(x)$ . So if we use matrix

notation and define  $h = \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix}$  and  $\partial f(x) = (\partial_1 f(x), \dots, \partial_n f(x))$  then we have  $D f(x)(h) = \partial f(x) \cdot h$  (using the matrix product).

**Proof.**

1. Let  $i \in \{1, \dots, n\}$  then by 14.35  $D_i f(x)$  exists so using 14.39 we have that  $\partial_i f(x)$  exists and  $\forall h \in \mathbb{K} D_i f(x)(h) = h \cdot \partial_i f(x)$  or  $\partial_i f(x) = D_i f(x)(1)$
2. Using 14.35 we have  $\forall i \in \{1, \dots, n\}$  that  $D_i f(x) = Df(x) \circ (i \rightarrow 0)$ . As  $((i \rightarrow 0)(1))_k = \delta_{i,k}(1) = \begin{cases} 1 & \text{if } i = k \\ 0 & \text{if } i \neq k \end{cases} = \delta_{i,k}$  proving that  $((i \rightarrow 0)(1)) = e_i$ . Hence it follows that  $\partial_i f(x) = D_i f(x)(1) = Df(x)((i \rightarrow 0)(1)) = Df(x)(e_i)$ . So

$$\partial_i f(x) = Df(x)(e_i) \quad (14.22)$$

Finally as  $h = \sum_{i=1}^n h_i \cdot e_i$  we have by the linearity of  $Df(x)$  that  $Df(x)(h) = Df(x)(\sum_{i=1}^n h_i \cdot e_i) = \sum_{i=1}^n h_i \cdot Df(x)(e_i) \stackrel{14.22}{=} \sum_{i=1}^n h_i \cdot \partial_i f(x)$  proving that

$$Df(x)(h) = \sum_{i=1}^n h_i \cdot \partial_i f(x)$$

□

**Definition 14.41.** Let  $n \in \mathbb{N}_0 \setminus \{1\}$ ,  $X, \{Y_i\}_{i \in \{1, \dots, n\}}$  be sets and  $\{f_i: X \rightarrow Y_i\}_{i \in \{1, \dots, n\}}$  a family of functions,  $V \subseteq \prod_{i \in \{1, \dots, n\}} Y_i$  then  $(f_1, \dots, f_n): X \rightarrow \prod_{i \in \{1, \dots, n\}} Y_i$  is defined by  $(f_1, \dots, f_n)(x) = (f_1(x), \dots, f_n(x))$

**Theorem 14.42.** Let  $X, \{Y_i\}_{i \in \{1, \dots, n\}}$  be sets and  $f: X \rightarrow \prod_{i \in \{1, \dots, n\}} Y_i$  then  $f = (\pi_1 \circ f, \dots, \pi_n \circ f)$

**Proof.** Let  $x \in X$  then  $\forall i \in \{1, \dots, n\} (f(x))_i = \pi_i(f(x)) = (\pi_i \circ f)(x)$  so that  $f(x) = ((\pi_1 \circ f)(x), \dots, (\pi_n \circ f)(x)) \stackrel{\text{def}}{=} ((\pi_1 \circ f), \dots, (\pi_n \circ f))(x)$  proving that  $f = (\pi_1 \circ f, \dots, \pi_n \circ f)$ . □

**Theorem 14.43.** Let  $n \in \mathbb{N}_0 \setminus \{1\}$ ,  $\langle X, \|\cdot\|_X \rangle$  be a normed space,  $\{\langle Y_i, \|\cdot\|_i \rangle\}_{i \in \{1, \dots, n\}}$  a family of normed spaces,  $\langle \prod_{i \in \{1, \dots, n\}} Y_i, \|\cdot\| \rangle$  the normed spaces equipped with the maximum norm [see 11.82],  $V \subseteq \prod_{i \in \{1, \dots, n\}} Y_i$ ,  $U \subseteq X$  open,  $x \in U$  and  $f: U \rightarrow V$  a function. Then we have that  $f$  is differentiable at  $x$  if and only if  $\forall i \in \{1, \dots, n\} \pi_i \circ f$  is differentiable at  $x$ . Further if  $f$  is differentiable at  $x$  [or equivalently  $\pi_i \circ f$  is differentiable at  $x$ ] then  $\pi_i \circ Df(x) = D(\pi_i \circ f)(x) \forall i \in \{1, \dots, n\}$  or equivalently  $Df(x) = (D(\pi_1 \circ f)(x), \dots, D(\pi_n \circ f)(x))$ .

**Proof.**

⇒. As  $f$  is differentiable at  $x$  there exists given  $\varepsilon > 0$  a  $\delta > 0$  such that  $\forall h \in U_x$  we have that  $\|f(x + h) - f(x) - Df(x)(h)\| \leq \varepsilon \cdot \|h\|_X$ . Using the definition of the maximum norm [see 11.82] we have that

$$\begin{aligned} \|f(x + h) - f(x) - Df(x)(h)\| &= \\ \max(\{\|\pi_i(f(x + h) - f(x) - Df(x)(h))\|_i | i \in \{1, \dots, n\}\}) &= \\ \max(\{\|\pi_i(f(x + h)) - \pi_i(f(x)) - \pi_i(Df(x)(h))\|_i | i \in \{1, \dots, n\}\}) &= \\ \max(\{\|(\pi_i \circ f)(x + h) - (\pi_i \circ f)(x) - (\pi_i \circ Df(x))(h)\|_i | i \in \{1, \dots, n\}\}) \end{aligned}$$

proving that  $\forall i \in \{1, \dots, n\} \|(\pi_i \circ f)(x + h) - (\pi_i \circ f)(x) - (\pi_i \circ Df(x))(h)\|_i < \varepsilon \cdot \|h\|_X$ . Hence

$\forall i \in \{1, \dots, n\} \pi_i \circ f$  is differentiable at  $x$  with  $D(\pi_i \circ f)(x) = \pi_i \circ Df(x)$

⇐. As  $\forall i \in \{1, \dots, n\} \pi_i \circ f$  is differentiable at  $x \in U$  there exists given  $\varepsilon > 0$  there exists a  $\delta_i > 0$  such that  $\forall h \in U_x$  with  $\|h\|_X < \delta_i$  we have  $\|(\pi_i \circ f)(x + h) - (\pi_i \circ f)(x) - D(\pi_i \circ f)(x)(h)\|_i \leq \varepsilon \cdot \|h\|_X$ . Define  $Df(x): X \rightarrow \prod_{i \in \{1, \dots, n\}} Y_i$  by  $Df(x) = (D(\pi_1 \circ f)(x), \dots, D(\pi_n \circ f)(x))$  then we have  $\forall \alpha, \beta \in \mathbb{K}$  and  $h_1, h_2 \in X$  that

$$\begin{aligned} Df(x)(\alpha \cdot h_1 + \beta \cdot h_2) &= \\ (D(\pi_1 \circ f)(x)(\alpha \cdot h_1 + \beta \cdot h_2), \dots, D(\pi_n \circ f)(x)(\alpha \cdot h_1 + \beta \cdot h_2)) &= \\ (\alpha \cdot D(\pi_1 \circ f)(x)(h_1) + \beta \cdot D(\pi_1 \circ f)(x)(h_2), \dots, \alpha \cdot D(\pi_n \circ f)(x)(h_1) + \beta \cdot D(\pi_n \circ f)(x)(h_2)) &= \\ \alpha \cdot (D(\pi_1 \circ f)(x)(h_1), \dots, D(\pi_n \circ f)(x)(h_1)) + \beta \cdot (D(\pi_1 \circ f)(x)(h_2), \dots, D(\pi_n \circ f)(x)(h_2)) \end{aligned}$$

proving that  $Df(x)$  is linear. Further if  $h \in X$  we have that

$$\begin{aligned}\|Df(x)(h)\| &= \max(\{\|(Df(x)(h))_i\|_i | i \in \{1, \dots, n\}\}) \\ &= \max(\{\|D(\pi_i \circ f)(x)(h)\|_i | i \in \{1, \dots, n\}\}) \\ &\leq \max(\{\|D(\pi_i \circ f)\|_{L(X, Y_i)} \cdot \|h\|_X | i \in \{1, \dots, n\}\}) \\ &= \max(\{\|D(\pi_i \circ f)\|_{L(X, Y_i)} | i \in \{1, \dots, n\}\}) \cdot \|h\|_X\end{aligned}$$

which prove by 11.180 that

$$Df(x) \in L\left(X, \prod_{i \in \{1, \dots, n\}} Y_i\right) \quad (14.23)$$

Take  $\delta = \min(\{\delta_i | i \in \{1, \dots, n\}\})$  then we have if  $h \in U_x$  with  $\|h\|_X < \delta$  that

$$\begin{aligned}\|f(x+h) - f(x) - Df(x)(h)\| &= \\ \max(\{\|\pi_i(f(x+h) - f(x) - Df(x)(h))\|_i | i \in \{1, \dots, n\}\}) &= \\ \max(\{\|\pi_i(f(x+h)) - \pi_i(f(x)) - \pi_i(Df(x)(h))\|_i | i \in \{1, \dots, n\}\}) &= \\ \max(\{\|\pi_i(f(x+h)) - \pi_i(f(x)) - \pi_i(Df(x)(h))\|_i | i \in \{1, \dots, n\}\}) &= \\ \max(\{(\pi_i \circ f)(x+h) - (\pi_i \circ f)(x) - (\pi_i \circ Df(x))(h)\|_i | i \in \{1, \dots, n\}\}) &< \varepsilon\end{aligned}$$

proving that

$f$  is differentiable at  $x$

□

**Corollary 14.44.** Let  $n \in \mathbb{N}_0 \setminus \{1\}$   $\{\langle X_i, \|\cdot\|_i \rangle\}_{i \in \{1, \dots, n\}}$  be a finite family of normed spaces,  $\langle Y, \|\cdot\|_Y \rangle, \langle Z, \|\cdot\|_Z \rangle$  normed spaces,  $\langle \prod_{i \in \{1, \dots, n\}} X_i, \|\cdot\| \rangle$  the normed space equipped with the maximum norm [see 11.82],  $U \subseteq \prod_{i \in \{1, \dots, n\}} X_i$  a open set,  $V \subseteq Y$  a open set,  $R \subseteq Z$  and  $f: U \rightarrow R$  and  $g: V \rightarrow Z$  functions and  $x \in U \cap f^{-1}(V)$  such that  $f$  is differentiable at  $x$  and  $g$  is differentiable at  $f(x)$  then  $g \circ f: U \cap f^{-1}(V) \rightarrow Z$  [so that by the chain rule 14.22]  $g \circ f$  is differentiable at  $x$  then  $D(g \circ f)(x) = \sum_{i=1}^n D_i g(f(x)) \circ D(\pi_i \circ f)(x)$ .

**Proof.** Let  $h \in X$  then

$$\begin{aligned}D(g \circ f)(x)(h) &\stackrel{14.22}{=} (Dg(f(x)) \circ Df(x))(h) \\ &= Dg(f(x))(Df(x)(h)) \\ &\stackrel{14.35}{=} \sum_{i=1}^n D_i g(f(x))((Df(x)(h))_i) \\ &= \sum_{i=1}^n D_i g(f(x))(\pi_i(Df(x)(h))) \\ &= \sum_{i=1}^n D_i g(f(x))((\pi_i \circ Df(x))(h)) \\ &\stackrel{14.43}{=} \sum_{i=1}^n D_i g(f(x))((D(\pi_i \circ f)(x))(h)) \\ &= \sum_{i=1}^n (D_i g(f(x)) \circ D(\pi_i \circ f)(x))(h) \\ &= \left( \sum_{i=1}^n D_i g(f(x)) \circ D(\pi_i \circ f)(x) \right)(h)\end{aligned}$$

proving that

$$D(g \circ f)(x) = \sum_{i=1}^n D_i g(f(x)) \circ D(\pi_i \circ f)(x)$$

□

**Theorem 14.45. (Jacobian matrix)** Let  $n \in \mathbb{N}_0 \setminus \{1\}$ ,  $\langle \mathbb{K}^n, \|\cdot\|_n \rangle$ ,  $\langle \mathbb{K}^m, \|\cdot\|_m \rangle$  be two normed spaces based on  $\langle \mathbb{K}, \|\cdot\| \rangle$  [with the maximum norm see 11.82], a open set  $U \subseteq \mathbb{K}^n$ ,  $V \subseteq \mathbb{K}^m$ ,  $x \in U$  and  $f: U \rightarrow V$  a differentiable function then  $\forall j \in \{1, \dots, m\}$  we have

$$\pi_j(Df(x)(h)) = (Df(x)(h))_j = Df_j(x)(h) = \sum_{i=1}^n \partial_i(\pi_j \circ f)(x) \cdot h_i$$

If we define the matrix  $\left[ \frac{\partial(f_1, \dots, f_m)(x)}{\partial(1, \dots, n)} \right] = \begin{pmatrix} \partial_1(\pi_1 \circ f)(x) & \dots & \partial_n(\pi_1 \circ f)(x) \\ \vdots & \ddots & \vdots \\ \partial_1(\pi_m \circ f)(x) & \dots & \partial_n(\pi_m \circ f)(x) \end{pmatrix}$  and  $h = \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix}$  then we can write using matrix multiplication  $Df(x)(h) = \left[ \frac{\partial(f_1, \dots, f_m)(x)}{\partial(1, \dots, n)} \right] \cdot h$ . The matrix  $\left[ \frac{\partial(f_1, \dots, f_m)(x)}{\partial(1, \dots, n)} \right]$  is called the Jacobian matrix of  $f$  at  $x$ .

**Proof.** Using 14.43 we have

$$Df(x)(h) = (D(\pi_1 \circ f)(x)(h), \dots, D(\pi_m \circ f)(x)(h)) \quad (14.24)$$

Let  $j \in \{1, \dots, m\}$  then as  $\pi_j \circ f: U \rightarrow Y$  is differentiable we have by 14.40 that

$$D(\pi_j \circ f)(x)(h) = \sum_{i=1}^n h_i \cdot \partial_i(\pi_j \circ f)(x) \quad (14.25)$$

Further

$$\begin{aligned} \left( \left[ \frac{\partial(f_1, \dots, f_m)(x)}{\partial(1, \dots, n)} \right] \cdot \begin{bmatrix} h_1 \\ \vdots \\ h_n \end{bmatrix} \right)_j &= \sum_{i=1}^n \left[ \frac{\partial(f_1, \dots, f_m)(x)}{\partial(1, \dots, n)} \right]_{j,i} \cdot h_i \\ &= \sum_{i=1}^n \partial_i(\pi_j \circ f)(x) \cdot h_i \\ &\stackrel{14.25}{=} D(\pi_j \circ f)(x)(h) \\ &\stackrel{14.25}{=} \pi_j(Df(x)(h)) \end{aligned}$$

proving that

$$\left[ \frac{\partial(f_1, \dots, f_m)(x)}{\partial(1, \dots, n)} \right] \cdot \begin{bmatrix} h_1 \\ \vdots \\ h_n \end{bmatrix} = Df(x)(h) \quad \square$$

**Example 14.46.** Let  $n \in \mathbb{N}_0 \setminus \{1\}$ ,  $\langle \mathbb{K}^n, \|\cdot\|_n \rangle$ ,  $\langle \mathbb{K}^m, \|\cdot\|_m \rangle$  be two vector spaces based on  $\langle \mathbb{K}, \|\cdot\| \rangle$  (with the maximum norm),  $\{e_i\}_{i \in \{1, \dots, n\}}$ ,  $\{f_i\}_{i \in \{1, \dots, m\}}$  the canonical bases [see 10.163] and  $L: \mathbb{K}^n \rightarrow \mathbb{K}^m$  a linear mapping with matrix [see 10.302]  $\mathcal{M}(L, \{e_i\}_{i \in \{1, \dots, n\}}, \{f_i\}_{i \in \{1, \dots, m\}}) = \{L_{i,j}\}_{(i,j) \in \{1, \dots, m\} \times \{1, \dots, n\}}$  in the canonical basis. Then  $\forall x \in \mathbb{K}^n$   $L$  is differentiable at  $x$  with Jacobian  $\left[ \frac{\partial f(1, \dots, m)(x)}{\partial(1, \dots, n)} \right] = L$

**Proof.** As  $L$  is linear it is differentiable at  $x$  [see 14.17] and  $\forall j \in \{1, \dots, m\}$   $\pi_j$  is linear we have that  $\pi_j \circ L$  is linear. Hence given  $i \in \{1, \dots, n\}$ ,  $j \in \{1, \dots, m\}$  we have that

$$\begin{aligned} \left[ \frac{\partial f(1, \dots, m)(x)}{\partial (1, \dots, n)} \right]_{j,i} &= \partial_i(\pi_j \circ L)(x) \\ &\stackrel{14.40}{=} (D(\pi_j \circ L)(x))(e_i) \\ &\stackrel{14.17}{=} (\pi_j \circ L)(e_i) \\ &= (L(e_i))_j \\ &\stackrel{10.302}{=} L_{j,i} \end{aligned}$$

proving that  $\frac{\partial(L_1, \dots, L_m)(x)}{\partial(1, \dots, n)} = L$  □

**Theorem 14.47.** Let  $\langle X_1, \|\cdot\|_1 \rangle$ ,  $\langle X_2, \|\cdot\|_2 \rangle$ ,  $\langle Y, \|\cdot\|_Y \rangle$  be normed spaces, and  $\langle X_1 \times X_2, \|\cdot\| \rangle$  equipped with the maximum norm,  $x = (x_1, x_2) \in X_1 \times X_2$  and  $L \in L(X_1, X_2; Y)$  [a bi-linear continuous mapping] then  $L: X_1 \times X_2 \rightarrow Y$  is differentiable at  $x$  and  $DL(x_1, x_2) = L(x_1, *) + L(*, x_2)$  where  $L(x_1,*)(h_1, h_2) = L(x_1, h_2)$  and  $L(*, x_2)(h_1, h_2) = L(h_1, x_2)$

**Proof.** First we prove that  $L(x_1, *) + L(*, x_2)$  is linear.

$$\begin{aligned} (L(x_1, *) + L(*, x_2))(\alpha \cdot (h_1, h_2) + \beta \cdot (g_1, g_2)) &= (L(x_1, *) + L(*, x_2))(\alpha \cdot h_1 + \beta \cdot g_1, \alpha \cdot h_2 + \beta \cdot g_2) \\ &= L(x_1, \alpha \cdot h_2 + \beta \cdot g_2) + L(\alpha \cdot h_1 + \beta \cdot g_1, x_2) \\ &= \alpha \cdot L(x_1, h_2) + \alpha \cdot L(h_1, x_2) + \beta \cdot L(x_1, g_2) + \beta \cdot L(g_1, x_2) \\ &= \alpha \cdot (L(x_1, *) + L(*, x_2))(h_1, h_2) + \beta \cdot (L(x_1, *) + L(*, x_2))(g_1, g_2) \end{aligned}$$

Second we prove that  $L(x_1, *) + L(*, x_2)$  is continuous let  $(h_1, h_2) \in X_1 \times X_2$  with  $\|(h_1, h_2)\| = 1$  then (see definition of maximum norm) we have  $\|h_1\|_1 \leq 1$  and  $\|h_2\|_2 \leq 1$ , so

$$\begin{aligned} \|(L(x_1, *) + L(*, x_2))(h_1, h_2)\|_Y &= \|L(x_1, h_2) + L(h_1, x_2)\|_Y \\ &\leq \|L(x_1, h_2)\|_Y + \|L(h_1, x_2)\|_Y \\ &\stackrel{11.198}{\leq} \|L\| \cdot \|x_1\|_1 \cdot \|h_2\|_2 + \|L\| \cdot \|x_2\|_2 \cdot \|h_1\|_1 \\ &\leq \|L\| \cdot (\|x_1\|_1 + \|x_2\|_2) \end{aligned}$$

which by 11.194 means that  $L(x_1,*)(h_1, h_2) = L(x_1, h_2)$  is continuous.

Finally let  $\varepsilon > 0$  and take  $0 < \delta < \frac{\varepsilon}{\|L\| + 1}$  then if  $h = (h_1, h_2) \in (X_1 \times X_2)_x = X_1 \times X_2$  with  $\|(h_1, h_2)\| < \delta \Rightarrow \|h_1\|_1, \|h_2\|_2 \leq \|(h_1, h_2)\| < \frac{\varepsilon}{\|L\| + 1}$  then

$$\begin{aligned} \|L((x_1, x_2) + (h_1, h_2)) - L((x_1, x_2)) - (L(x_1, *) + L(*, x_2))(h_1, h_2)\|_Y &= \\ \|L(x_1 + h_1, x_2 + h_2) - L(x_1, x_2) - L(x_1, h_2) - L(h_1, x_2)\|_Y &= \\ \|L(x_1, x_2 + h_2) + L(h_1, x_2 + h_2) - L(x_1, x) - L(x_1, h_2) - L(h_1, x_2)\|_Y &= \\ \|L(x_1, x_2) + L(x_1, h_2) + L(h_1, x_2) + L(h_1, h_2) - L(x_1, x_2) - L(x_1, h_2) - L(h_1, x_2)\|_Y &= \\ \|L(h_1, h_2)\|_Y &\leq \\ \|L\| \cdot \|h_1\|_1 \cdot \|h_2\|_2 &< \\ \|L\| \cdot \frac{\varepsilon}{\|L\| + 1} \cdot \|h_2\|_2 &< \\ \varepsilon \cdot \|h_2\|_2 &< \varepsilon \cdot \|(h_1, h_2)\| \end{aligned}$$

proving that  $DL(x, y) = L(x_1, *) + L(*, x_2)$  □

We can use the above theorem to find the differential of a product of differentiable functions.

**Corollary 14.48.** *Let  $\langle X, \|\cdot\|_X \rangle$  be a normed space,  $\langle \mathbb{K}, \|\cdot\| \rangle$ ,  $U \subseteq \mathbb{K}$  open with  $x \in U$  and  $f, g: U \rightarrow \mathbb{K}$  two maps differentiable at  $x$  then  $f \cdot g: U \rightarrow \mathbb{K}$  is differentiable at  $x$  with  $D(f \cdot g)(x) = g(x) \cdot Df(x) + f(x) \cdot Dg(x)$ .*

**Proof.** Define  $h: X \rightarrow (\mathbb{K}, \mathbb{K})$  by  $x \mapsto (f(x), g(x))$  which is differentiable at  $x \in U$  as  $h_1 = \pi_1 \circ h = f$  and  $h_2 = \pi_2 \circ h = g$  are differentiable (see 14.43). Also  $p: \mathbb{K} \times \mathbb{K} \rightarrow \mathbb{K}$  defined by  $p(x, y) = x \cdot y$  is clearly bi-linear and continuous [because  $|p(x, y)| = |x| \cdot |y|$  and 11.194]. Using the previous theorem  $p$  is differentiable at  $(x, y)$  with  $Dp(x, y) = p(x, *) + p(*, y) = x \cdot * + * \cdot y$ . Now as  $f \cdot g = p \circ h$  we have by the chain rule (see 14.22) that  $f \cdot g$  is differentiable and

$$\begin{aligned} D(f \cdot g)(x) &= Dp(h(x)) \circ Dh(x) \\ &= Dp(f(x), g(x)) \circ Dh(x) \\ &\stackrel{14.47 \text{ and } 14.43}{=} (f(x) \cdot * + * \cdot g(x)) \circ (Df(x), Dg(x)) \end{aligned}$$

so that

$$D(f \cdot g)(x)(h) = (f(x) \cdot * + * \cdot g(x))(Df(x)(h), Dg(x)(h)) = f(x) \cdot Dg(x)(h) + g(x) \cdot Df(x)(h)$$

proving that

$$D(f \cdot g)(x) = f(x) \cdot Dg(x) + g(x) \cdot Df(x) \quad \square$$

If we take in the above theorem  $\langle X, \|\cdot\|_X \rangle = \langle \mathbb{K}, \|\cdot\| \rangle$  and 14.19 the above corollary gives

**Corollary 14.49.** *Take the normed space  $\langle \mathbb{K}, \|\cdot\| \rangle$ ,  $U \subseteq \mathbb{K}$  a open set,  $x \in U$  and  $f, g: U \rightarrow \mathbb{K}$  functions that have a derivate at  $x$  then  $f \cdot g$  has a derivate at  $x$  and  $(f \cdot g)'(x) = f(x) \cdot g'(x) + g(x) \cdot f'(x)$*

**Proof.** Using 14.19 we have that  $f, g$  are differentiable and thus by the above theorem we have that  $f \cdot g$  is differentiable and thus by 14.19 has a derivate. Also using 14.19 we have  $(f \cdot g)'(x) = D(f \cdot g)(x)(1) = g(x) \cdot Df(x)(1) + f(x) \cdot Dg(x)(1) = g(x) \cdot f'(x) + f(x) \cdot g'(x)$  □

We can now calculate the differential of a power

**Corollary 14.50.** *Take the normed space  $\langle \mathbb{K}, \|\cdot\| \rangle$ ,  $U \subseteq \mathbb{K}$  a open set,  $x \in U$ ,  $z \in \mathbb{N}_0$  define then  $*^z: \mathbb{K} \rightarrow \mathbb{K}$  by  $x \mapsto x^z$  then  $*^z$  has a derivate at  $x \in \mathbb{K}$  given by  $(*^z)'(x) = z \cdot *^{z-1}(x)$ . If  $z \in \{-z \mid z \in \mathbb{N}_0\}$  then  $*^z: \mathbb{K} \setminus \{0\} \rightarrow \mathbb{K}$  defined by  $x \mapsto x^z = \frac{1}{x^{-z}}$  has a derivate at  $x \in \mathbb{K} \setminus \{0\}$  [a open set] given by  $(*^z)'(x) = z \cdot *^{z-1}(x)$*

**Proof.** We have the following cases to consider for  $z$

$z \in \mathbb{N}_0$ . We proceed here by induction so let  $S = \{z \in \mathbb{N}_0 \mid *^z \text{ has a derivate } z \cdot *^{z-1}\}$  then we have

$1 \in S$ . If  $z = 1$  then  $*^1: \mathbb{K} \rightarrow \mathbb{K}$  is linear and continuous so that  $(*^1)'(x) = D(*^1)(x)(1) = *^1(1) = 1^1 = 1 = 1 \cdot x^0 = 1 \cdot x^{z-1}$  so that  $1 \in S$

$z \in S \Rightarrow z + 1 \in S$ . Assume now that  $z \in S$  then we have that  $(*^{z+1})(x) = x^{z+1} = x \cdot x^z$  so that  $*^{z+1} = *^1 \cdot *^z$  and then using the previous corollary [see 14.49] we have

$$\begin{aligned} (*^{z+1})'(x) &= (*^1 \cdot *^z)'(x) \\ &= *^1(x) \cdot (*^z)'(x) + (*^z)(x) \cdot (*^1)'(x) \\ &\stackrel{1 \in S}{=} x \cdot (*^z)'(x) + x^z \cdot 1 \\ &\stackrel{z \in S}{=} x \cdot z \cdot (*^{z-1})(x) + x^z \\ &= x \cdot z \cdot x^{z-1} + x^z \\ &= (z+1) \cdot x^z \\ &= (z+1) \cdot (*^{(z+1)-1})(x) \end{aligned}$$

proving that  $z + 1 \in S$

$z \in \{-z \mid z \in \mathbb{N}_0\}$ . again we prove this by induction so let  $S = \{z \in \mathbb{N}_0 \mid *^{-z} \text{ has a derivate } (-z) \cdot *^{(-z-1)}\}$  then we have

$1 \in S$ . If  $z = 1$  then  $*^{-1} : \mathbb{K} \setminus \{0\} \rightarrow \mathbb{K}$  is defined by  $x \mapsto \frac{1}{x}$ . If  $x \in \mathbb{K} \setminus \{0\}$  and  $0 < h$  so that  $x + h \in \mathbb{K} \setminus \{0\}$  then

$$\begin{aligned}
 \frac{\frac{1}{x+h} - \frac{1}{x}}{h} + \frac{1}{x^2} &= \frac{\frac{x}{x \cdot (x+h)} - \frac{x+h}{x \cdot (x+h)}}{h} + \frac{1}{x^2} \\
 &= \frac{\frac{-h}{x \cdot (x+h)}}{h} + \frac{1}{x^2} \\
 &= \frac{-1}{x \cdot (x+h)} + \frac{1}{x^2} \\
 &= \frac{1}{x} \cdot \left( \frac{-1}{x+h} + \frac{1}{x} \right) \\
 &= \frac{1}{x} \cdot \frac{-x+x+h}{x \cdot (x+h)} \\
 &= \frac{h}{x^2 \cdot (x+h)}
 \end{aligned} \tag{14.26}$$

Given  $\varepsilon > 0$  take then  $0 < \delta = \min \left( \frac{\varepsilon \cdot |x|^3}{2}, \frac{|x|}{2} \right)$  with  $0 < |h| < \delta$  and  $x + h \in \mathbb{K} \setminus \{0\}$  [meaning  $h \in (\mathbb{K} \setminus \{0\})_x$ ]. Then we have  $|h| < \frac{|x|}{2}$  giving  $-\frac{|x|}{2} < -|h| \Rightarrow \frac{|x|}{2} = |x| - \frac{|x|}{2} < |x| - |h| \leq |x + h|$  [because  $|x| = |x + h - h| \leq |x + h| + |h| \Rightarrow |x| - |h| \leq |x + h|$ ] and thus  $\frac{|x|}{2} < |x + h|$  or  $\frac{1}{|x+h|} < \frac{2}{|x|}$ . So  $\left| \frac{h}{x^2 \cdot (x+h)} \right| = \frac{|h|}{|x|^2 \cdot |x+h|} < \frac{2 \cdot |h|}{|x| \cdot |x|^2} = 2 \cdot \frac{|h|}{|x|^3} < 2 \cdot \frac{\varepsilon \cdot |x|^3}{|x|^3 \cdot 2} = \varepsilon$  [as  $|h| < \delta = \frac{\varepsilon \cdot |x|^3}{2}$ ]. So if  $0 < \delta = \min \left( \frac{\varepsilon \cdot |x|^3}{2}, \frac{|x|}{2} \right)$  we have  $\left| \frac{*^{-1}(x+h) - *^{-1}(x)}{h} - (-1 \cdot *^{(-2)}(x)) \right| = \left| \frac{\frac{1}{x+h} - \frac{1}{x}}{h} + \frac{1}{x^2} \right| \stackrel{14.26}{=} \left| \frac{h}{x^2 \cdot (x+h)} \right| < \varepsilon$  proving that  $(*^{-1})'(x) = -1 \cdot (*^{-1-1})(x)$  or  $1 \in S$ .

$z \in S \Rightarrow z + 1 \in S$ . If  $z \in S$  then  $(*^{-1})'(x) = \frac{1}{*^{z+1}(x)} = \frac{1}{x^{z+1}} = \frac{1}{x \cdot x^z} = (*^{-1})(x) \cdot (*^{-z})(x)$  so that  $(*^{-1})'(*^{-1}) = (*^{-1}) \cdot (*^{-z})$ . Using the previous corollary [see 14.49] we have

$$\begin{aligned}
 ((*^{-1})'(*^{-1}))'(x) &= (*^{-1})(x) \cdot (*^{-z})'(x) + (*^{-z})(x) \cdot (*^{-1})'(x) \\
 &\stackrel{z \in S \text{ and } z \in S}{=} x^{-1} \cdot (-z) \cdot (*^{-z-1})(x) + x^{-z} \cdot (-1) \cdot (*^{-2})(x) \\
 &= x^{-1} \cdot (-z) \cdot x^{-z-1} - x^{-z} \cdot x^{-2} \\
 &= -z \cdot x^{-z-2} - x^{-z-2} \\
 &= (-z-1) \cdot x^{-z-2} \\
 &= ((-z-1)) * x^{-(z+1)-1} \\
 &= ((-z-1)) \cdot (*^{-(z+1)-1})(x)
 \end{aligned}$$

proving that  $z + 1 \in S$

□

## 14.2 Higher order differentiability

We define higher order differentiability using recursion as follows

**Definition 14.51.** Let  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle$  be normed spaces,  $U \subseteq X$  a open set,  $V \subseteq Y$ ,  $n \in \mathbb{N}_0$  and  $f: U \rightarrow V$  a function then  $f$  is  $n$ -times differentiable on  $U$  with  $n$ -the differential  $D^n f: U \rightarrow L^n(X; Y)$  [see 11.206] if one of the following is correct:

1. If  $n = 1$  then  $f: U \rightarrow V$  must be differentiable  $\forall x \in U$  and  $D^1 f: U \rightarrow L(X, Y)$  is defined by  $x \rightarrow D^1 f(x) = Df(x)$
2. If  $1 < n$  then  $f$  must be  $(n-1)$ -times differentiable on  $U$  and  $D^{n-1} f: U \rightarrow L^{n-1}(X; Y)$  must be differentiable at every  $x \in U$ .  $D^n f: U \rightarrow L^n(X; Y)$  is defined by  $x \rightarrow D^n f(x) = D(D^{n-1} f)(x) \in L(X, L(X^{n-1}; Y)) = L^n(X; Y)$  (see 11.206)

**Remark 14.52.** We have seen by 11.221 that there exists a norm preserving isomorphism  $\mathcal{P}_n: L^n(X; Y) \rightarrow L(X^n; Y)$  allowing us to consider  $\mathcal{P}_n(D^n f(x)) \in L(X^n; Y)$  as a multi-linear mapping where  $\mathcal{P}_n(D^n f(x))(x_1, \dots, x_n) = D^n f(x)(x_1, \dots, x_n)$ . To simplify notation when we write  $D^n f(x)(h_1, \dots, h_n)$  we actually means  $\mathcal{P}_n(D^n f(x))(x_1, \dots, x_n)$  where

$$\begin{aligned} \mathcal{P}_n(D^n f(x))(x_1, \dots, x_n) &= D^n f(x)(h_1, \dots, h_n) \\ &\stackrel{11.213}{=} \begin{cases} (D^n f(x)(h_1))(h_2, \dots, h_n) & \text{if } n > 1 \\ D^1 f(x)(h_1) = Df(x)(h_1) & \text{if } n = 1 \end{cases} \end{aligned}$$

or using a shorter notation

$$\mathcal{P}_n(D^n f(x))(h_1, \dots, h_n) = D^n f(x)(h_1, \dots, h_n) = (\dots(D^n f(x)(h_1))(h_2), \dots)(h_n)$$

**Definition 14.53.** Let  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle$  be normed spaces,  $U \subseteq X$  a open set,  $V \subseteq Y$ ,  $n \in \mathbb{N}_0$  and  $f: U \rightarrow V$  a function then  $f$  is  $\infty$ -differentiable if  $\forall n \in \mathbb{N}_0$  we have that  $f$  is  $n$ -differentiable.

We next define the concept of differentiable classes

**Definition 14.54.** Let  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle$  be normed spaces,  $U \subseteq X$  a open set,  $V \subseteq Y$ ,  $n \in \mathbb{N}$  and  $f: U \rightarrow V$  a function then  $f$  is of class  $C^n$  if one of the following is satisfied

1. If  $n = 0$  then  $f$  must be continuous on  $U$
2. If  $1 < n$  then  $f$  must be  $n$ -times differentiable on  $U$  and  $D^n f: U \rightarrow L^n(X; Y)$  must be continuous on  $U$

**Definition 14.55.** Let  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle$  be normed spaces,  $U \subseteq X$  a open set,  $V \subseteq Y$  and  $f: U \rightarrow V$  a function then  $f$  is of class  $C^\infty$  if  $\forall n \in \mathbb{N}$  we have that  $f$  is of class  $C^n$ . Functions of class  $C^\infty$  are also called **smooth** functions.

As for derivates we use the following definition

**Definition 14.56.** Let  $\langle Y, \|\cdot\|_Y \rangle$  be a normed space and  $\langle \mathbb{K}, \|\cdot\| \rangle$  be the normed space  $\mathbb{K}$  equipped with the canonical norm,  $U \subseteq \mathbb{K}$  a open set,  $V \subseteq Y$ ,  $n \in \mathbb{N}_0$  and  $f: U \rightarrow V$  a function then  $f$  has a  $n$ -the derivate  $f^{(n)}$  on  $U$  if one of the following is satisfied.

1. If  $n = 1$  we must have  $\forall x \in U$  that  $f$  has a derivate at  $x$  and we define  $f^{(1)}: U \rightarrow Y$  by  $x \rightarrow f^{(1)}(x) = f'(x)$
2. If  $n > 1$  we must have that  $f$  has a  $(n-1)$ -the derivate on  $U$  and  $\forall x \in U$  we have that  $f^{(n-1)}: U \rightarrow Y$  has a derivate at  $x$  and we define  $f^{(n)}: U \rightarrow Y$  by  $x \rightarrow f^n(x) = (f^{(n-1)}(x))'$

**Lemma 14.57.** Let  $\langle \mathbb{K}, \|\cdot\| \rangle$  be the vector space  $\mathbb{K}$  using the canonical norm,  $\langle Y, \|\cdot\|_Y \rangle$  a normed space,  $U \subseteq \mathbb{K}$  a open set and  $f: U \subseteq \mathbb{K} \rightarrow L(\mathbb{K}, Y)$  then if we define  $f(\star)(1): U \rightarrow Y$  by  $t \rightarrow (f(\star)(1))(t) = f(t)(1)$  we have that  $f$  is continuous at  $x \in U$  if and only if  $f(\star)(1)$  is continuous at  $x \in U$

**Proof.**

$\Rightarrow$ . Let  $x \in U$  then as  $f$  is continuous on  $U$ ,  $f$  is continuous at  $x$ , so  $\forall \varepsilon > 0$  there exists a  $\delta > 0$  such that if  $|x - y| < \delta \Rightarrow \|f(x) - f(y)\|_{L(\mathbb{K}, Y)} < \varepsilon$  then

$$\begin{aligned} \|(f(\star)(1))(x) - (f(\star)(1))(y)\|_Y &= \|f(x)(1) - f(y)(1)\|_Y \\ &= \|(f(x) - f(y))(1)\|_Y \\ &\leq \|f(x) - f(y)\|_{L(\mathbb{K}, Y)} \cdot |1| \\ &= \|f(x) - f(y)\|_{L(\mathbb{K}, Y)} \\ &< \varepsilon \end{aligned}$$

proving continuity of  $f(\star)(1)$  at  $x$ . As  $x$  was chosen arbitrary we have that  $f(\star)(1)$  is continuous on  $U$ .

$\Leftarrow$ . Let  $x \in U$  then as  $f(\star)(1)$  is continuous on  $U$  we have that  $f(\star)(1)$  is continuous at  $x$ , so given  $\varepsilon > 0$  there exists a  $\delta > 0$  such that if  $|x - y| < \delta \Rightarrow \|(f(\star)(1))(x) - (f(\star)(1))(y)\|_Y < \varepsilon$ . Then as

$$\begin{aligned} \|(f(x) - f(y))(r)\|_Y &= \|(f(x) - f(y))(r \cdot 1)\|_Y \\ &= \|r \cdot (f(x) - f(y))(1)\|_Y \\ &= |r| \cdot \|(f(x) - f(y))(1)\|_Y \\ &= |r| \|f(x)(1) - f(y)(1)\|_Y \\ &= |r| \|(f(\star)(1))(x) - (f(\star)(1))(y)\|_Y \\ &< \varepsilon \cdot |r| \end{aligned}$$

proving, as  $f(x) - f(y) \in L[\mathbb{K}, Y]$  by definition of the operator norm [see 11.185] that  $\|(f(x) - f(y))(r)\|_Y < \varepsilon$ . So  $f$  is continuous at  $x$ .  $\square$

**Theorem 14.58.** Let  $\langle \mathbb{K}, \|\cdot\| \rangle$  be the vector space  $\mathbb{K}$  using the canonical norm,  $\langle Y, \|\cdot\|_Y \rangle$  a normed space,  $U \subseteq \mathbb{K}$  a open set,  $V \subseteq Y$  and  $f: U \subseteq \mathbb{K} \rightarrow V$  a function then the following are equivalent:

1.  $f$  is  $C^1$
2.  $\forall t \in U$  we have that  $f'(t)$  exists and  $f': U \rightarrow V$  defined by  $t \rightarrow f'(t)$  is continuous on  $U$  (or  $f$  has a 1-derivate that is continuous).

further we have in case of (1) or equivalently (2) that  $f' = Df(\star)(1)$

**Proof.**

**1  $\Rightarrow$  2.** As  $f$  is  $C^1$  we have that  $f$  is differentiable on  $U$  and  $Df: x \rightarrow Df(x) \in L(\mathbb{K}, Y)$  is continuous. Then using 14.19 we have that  $\forall x \in U$   $f'(x)$  exists and  $f'(x) = Df(x)(1) = (Df(\star)(1))(x)$  proving that  $f' = Df(\star)(1)$ . We can use now the previous lemma to find that  $f'$  is continuous on  $U$  because continuity of  $Df$ .

**2  $\Rightarrow$  1.** As  $\forall x \in U$   $f'(x)$  exists we use 14.19 again to find the existence of  $Df(x)$  where  $f'(x) = Df(x)(1)$  and using the previous lemma and the continuity of  $f' = Df(\star)(1)$  we find that  $Df: x \rightarrow Df(x)$  is continuous.  $\square$

We extend now the above case to the more general case of  $n \in \mathbb{N}_0$ .

**Theorem 14.59.** Let  $\langle \mathbb{K}, \|\cdot\| \rangle$  be the real (complex space) equipped with the canonical norm,  $\langle Y, \|\cdot\|_Y \rangle$  a Banach space,  $U \subseteq \mathbb{K}$  open,  $V \subseteq Y$ ,  $n \in \mathbb{N}_0$  and  $f: U \rightarrow V$  a function then  $f$  is  $C^n$  if and only if the  $n$ -the derivate  $f^{(n)}$  exists and is continuous. Furthermore if  $x \in U$  then  $D^n f(x) \left( \underbrace{1: \dots : 1}_n \right) = f^{(n)}(x)$ , hence using 11.214  $D^n f(x)(v_1: \dots : v_n) = f^n(x) \cdot \prod_{i \in \{1, \dots, n\}} v_i$  where  $\{v_i\}_{i \in \{1, \dots, n\}} \subseteq \mathbb{K}$ .

**Proof.** We prove this by induction so let  $S = \left\{ n \in \mathbb{N}_0 \mid f \text{ is } C^n \text{ at } x \in U \text{ if and only if } f^n \text{ exists and is continuous. Further } D^n f(x) \left( \underbrace{1: \dots: 1}_n \right) = f^{(n)} \right\}$  then we have:

$n \in \mathcal{S}$ . If  $n = 1$  then by 14.58 we have that  $f$  is  $C^1$  if and only if  $f^{(1)} = f'$  exists on  $U$  and is continuous. Further  $f^{(1)}(x) = f'(x) = Df(x)(1) = D^1 f(x)(1: \dots: 1)$  proving that  $1 \in S$

$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$ . If  $n \in S$  take then  $n+1$  then we have to prove the following statements.

**$f$  is  $C^{n+1} \Rightarrow f^{(n+1)}$  exists and is continuous.** Then given  $x \in U$  we have by definition that  $D^{n+1} f(x) = D(D^n f)(x)$  exists, so given  $\varepsilon > 0$  there exists a  $\delta > 0$  such that if  $0 < |h| < \delta$  and  $h \in U_x \Rightarrow x + h \in U$  then

$$\|D^n f(x + h) - D^n f(x) - D^{n+1} f(x)(h)\|_{L^n(X; Y)} < \varepsilon \cdot |h| \quad (14.27)$$

Further we have  $D^n f(x + h), D^n f(x), D^{n+1} f(x)(h) \in L^n(X; Y)$  hence as  $L^n(X; Y)$  is a vector space it follows that  $(D^n f(x + h) - D^n f(x) - D^{n+1} f(x)(h)) \in L^n(X; Y)$ . Hence

$$\begin{aligned} & \left\| (D^n f(x + h) - D^n f(x) - D^{n+1} f(x)(h)) \left( \underbrace{1: \dots: 1}_n \right) \right\|_Y \stackrel{11.216}{\leq} \\ & \|D^n f(x + h) - D^n f(x) - D^{n+1} f(x)(h)\|_{L^n(X; Y)} \cdot \prod_{i=1}^n |1| = \\ & \|D^n f(x + h) - D^n f(x) - D^{n+1} f(x)(h)\|_{L^n(X; Y)} <_{14.27} \varepsilon \cdot |h| \quad (14.28) \end{aligned}$$

Next

$$\begin{aligned} & (D^n f(x + h) - D^n f(x) - D^{n+1} f(x)(h)) \left( \underbrace{1: \dots: 1}_n \right) \stackrel{11.215}{=} \\ & D^n f(x + h) \left( \underbrace{1: \dots: 1}_n \right) - D^n f(x) \left( \underbrace{1: \dots: 1}_n \right) - D^{n+1} f(x)(h) \left( \underbrace{1: \dots: 1}_n \right) \stackrel{n \in \mathcal{S}}{=} \\ & f^{(n)}(x + h) - f^{(n)}(x) - D^{n+1} f(x)(h) \left( \underbrace{1: \dots: 1}_n \right) = \\ & f^{(n)}(x + h) - f^{(n)}(x) - h \cdot D^{n+1} f(x)(1) \left( \underbrace{1: \dots: 1}_n \right) \stackrel{11.213}{=} \\ & f^{(n)}(x + h) - f^{(n)}(x) - h \cdot D^{n+1} f(x) \left( \underbrace{1: \dots: 1}_{n+1} \right) \quad (14.29) \end{aligned}$$

Hence

$$\begin{aligned} & \left\| \frac{f^{(n)}(x + h) - f^{(n)}(x) - D^{n+1} f(x) \left( \underbrace{1: \dots: 1}_{n+1} \right)}{h} \right\|_Y = \\ & \left\| \frac{f^{(n)}(x + h) - f^{(n)}(x) - h \cdot D^{n+1} f(x) \left( \underbrace{1: \dots: 1}_{n+1} \right)}{h} \right\|_Y \stackrel{14.29}{=} \\ & \left\| \frac{(D^n f(x + h) - D^n f(x) - D^{n+1} f(x)(h)) \left( \underbrace{1: \dots: 1}_n \right)}{h} \right\|_Y = \\ & \left\| \frac{(D^n f(x + h) - D^n f(x) - D^{n+1} f(x)(h)) \left( \underbrace{1: \dots: 1}_n \right)}{|h|} \right\|_Y <_{14.28} \frac{\varepsilon \cdot |h|}{|h|} = \varepsilon \end{aligned}$$

which proves that

$$f^{(n+1)}(x) \text{ exists and } f^{(n+1)}(x) = (f^{(n)}(x))'(x) = D^{n+1}f(x) \left( \underbrace{1: \dots: 1}_{n+1} \right) \quad (14.30)$$

To prove continuity note that  $D^{n+1}f$  is continuous on  $U$ , so if  $x \in U$  and  $\varepsilon > 0$  there exists a  $\delta > 0$  such that if  $|x - y| < \delta$  and  $y \in U$  then  $\|D^{n+1}f(x) - D^{n+1}f(y)\|_{L^{n+1}(X; Y)} < \varepsilon$ . Hence

$$\begin{aligned} \|f^{(n+1)}(x) - f^{(n+1)}(y)\|_Y &\stackrel{14.30}{=} \left\| D^{n+1}f(x) \left( \underbrace{1: \dots: 1}_{n+1} \right) - D^{n+1}f(y) \left( \underbrace{1: \dots: 1}_{n+1} \right) \right\|_Y \\ &= \left\| (D^{n+1}f(x) - D^{n+1}f(y)) \left( \underbrace{1: \dots: 1}_{n+1} \right) \right\|_Y \\ &\leqslant_{11.216} \|D^{n+1}f(x) - D^{n+1}f(y)\|_{L^{n+1}(X; Y)} \cdot \prod_{i \in \{1, \dots, n+1\}} |1| \\ &= \|D^{n+1}f(x) - D^{n+1}f(y)\|_{L^{n+1}(X; Y)} \\ &< \varepsilon \end{aligned}$$

proving that  $f^{(n+1)}$  is continuous at  $x \in U$ . Hence we have that

$$f^{(n+1)} \text{ exists and is continuous with } f^{(n+1)}(x) = D^{n+1}f(x) \left( \underbrace{1: \dots: 1}_{n+1} \right) \quad (14.31)$$

**$f^{(n+1)}$  exists and is continuous  $\Rightarrow f$  is  $C^{n+1}$ .** Then  $\forall x \in U f^{(n)}(x)$  exists and has a derivate  $(f^{(n)}(x))' = f^{(n+1)}(x)$ . So given  $\varepsilon > 0$  there exists a  $\delta > 0$  such that if  $0 < |h| < \delta$  and  $h \in U_x$  then  $\left\| \frac{f^{(n)}(x+h) - f^{(n)}(x)}{h} - f^{(n+1)}(x) \right\|_Y < \varepsilon$ , giving after multiplying both sides by  $h$

$$\|f^{(n)}(x+h) - f^{(n)}(x) - h \cdot f^{(n+1)}(x)\|_Y < \varepsilon \cdot |h| \quad (14.32)$$

Now

$$\begin{aligned} \left\| D^n f(x+h) \left( \underbrace{1: \dots: 1}_n \right) - D^n f(x) \left( \underbrace{1: \dots: 1}_n \right) - h \cdot f^{(n+1)}(x) \right\|_Y &\stackrel{n \in \mathcal{S}}{=} \\ \|f^{(n)}(x+h) - f^{(n)}(x) - h \cdot f^{(n+1)}(x)\|_Y &<_{14.32} \varepsilon \cdot |h| \end{aligned} \quad (14.33)$$

Define now  $k_x: \mathbb{K}^{n+1} \rightarrow Y$  by  $k_x(v_1, \dots, v_{n+1}) = f^{(n+1)}(x) \cdot \prod_{i \in \{1, \dots, n+1\}} v_i$  then by 11.195 we have that  $k_x$  is multi-linear and continuous so that  $k_x \in L(\mathbb{K}^{n+1}, Y)$ . Define then  $g_x \in L^{n+1}(\mathbb{K}; Y)$  by  $g_x = (\mathcal{P}_{n+1})^{-1}(k_x)$  so that

$$\begin{aligned} g_x(v_1: \dots: v_{n+1}) &= k_x(v_1, \dots, v_{n+1}) \\ &= f^{(n+1)}(x) \cdot \prod_{i \in \{1, \dots, n+1\}} v_i \end{aligned} \quad (14.34)$$

Further

$$\begin{aligned} g_x(h) \left( \underbrace{1: \dots: 1}_n \right) &\stackrel{11.213}{=} g_x \left( h: \underbrace{1: \dots: 1}_n \right) \\ &\stackrel{14.34}{=} f^{(n+1)}(x) \cdot \prod_{i \in \{1, \dots, n+1\}} h_i \text{ where } h_i = \begin{cases} h \text{ if } i = 1 \\ 1 \text{ if } i \in \{2, \dots, n+1\} \end{cases} \\ &= f^{(n+1)}(x) \cdot \left( \left( \prod_{i \in \{1\}} h_i \right) \cdot \prod_{i \in \{2, \dots, n\}} h_i \right) \\ &= f^{(n+1)}(x) \cdot (h \cdot 1) \\ &= f^{(n+1)}(x) \cdot h \end{aligned} \quad (14.35)$$

Hence

$$\begin{aligned}
 & \|D^n f(x+h) - D^n f(x) - g_x(h)\|_{L^n(\mathbb{K}, Y)} \stackrel{\text{11.217}}{\leq} \\
 & \left\| (D^n f(x+h) - D^n f(x) - g_x(h)) \left( \underbrace{1: \dots: 1}_n \right) \right\|_Y = \\
 & \left\| D^n f(x+h)(1: \dots: 1) - D^n f(x)(1: \dots: 1) - g_x(h) \left( \underbrace{1: \dots: 1}_n \right) \right\|_Y \stackrel{\text{14.35}}{=} \\
 & \|D^n f(x+h)(1: \dots: 1) - D^n f(x)(1: \dots: 1) - f^{(n+1)}(x) \cdot h\|_Y <_{\text{14.33}} \varepsilon \cdot |h|
 \end{aligned}$$

proving as  $g_x \in L^{n+1}(\mathbb{K}; Y) = L(L^n(\mathbb{K}; Y), Y)$  that  $D(D^n f)(x)$  exists and  $D(D^n f)(x) = g_x(h)$ . So

$$\forall x \in U \text{ we have } D^{n+1}f(x) \text{ exists and } D^{n+1}f(x) = g_x \quad (14.36)$$

Let  $x \in U$  then as  $f^{(n+1)}$  is continuous at  $x$  we have that given  $\varepsilon > 0$  there exist a  $\delta > 0$  so that if  $|x - y| < \delta$  and  $y \in U$  then  $\|f^{(n+1)}(x) - f^{(n+1)}(y)\|_Y < \varepsilon$ . Define now  $g: U \rightarrow L^{n+1}(\mathbb{K}; Y)$  by  $g(x) = g_x$  then  $D^{(n+1)}f = g$  and

$$\begin{aligned}
 & \|D^{n+1}f(x) - D^{n+1}f(y)\|_{L^{n+1}(\mathbb{K}; Y)} = \\
 & \|g(x) - g(y)\|_{L^{n+1}(\mathbb{K}; Y)} \stackrel{\text{11.217}}{\leq} \\
 & \left\| (g(x) - g(y)) \left( \underbrace{1: \dots: 1}_{n+1} \right) \right\|_Y = \\
 & \left\| g(x) \left( \underbrace{1: \dots: 1}_{n+1} \right) - g(y) \left( \underbrace{1: \dots: 1}_{n+1} \right) \right\|_Y = \\
 & \left\| g_x \left( \underbrace{1: \dots: 1}_{n+1} \right) - g_y \left( \underbrace{1: \dots: 1}_{n+1} \right) \right\|_Y \stackrel{\text{14.34}}{=} \\
 & \left\| f^{(n+1)}(x) \cdot \prod_{i \in \{1, \dots, n+1\}} 1 - f^{(n+1)}(y) \cdot \prod_{i \in \{1, \dots, n+1\}} 1 \right\|_Y = \\
 & \|f^{(n+1)}(x) - f^{(n+1)}(y)\|_Y < \varepsilon
 \end{aligned}$$

Proving that  $D^{n+1}f$  is continuous and thus that

$$f \text{ is of class } C^{n+1} \quad (14.37)$$

hence

$$n+1 \in \mathcal{S}$$

By mathematical induction we have then  $S = \mathbb{N}_0$  proving our theorem.  $\square$

**Example 14.60.** Let  $\langle Y, \|\cdot\| \rangle$  a normed space then  $\sigma: \mathbb{K} \rightarrow Y$  defined by  $\sigma(t) = x + t \cdot y$  is  $C^1$  [and is thus continuous] with  $\forall t \in \mathbb{K} \sigma'(t) = y$  so that  $\sigma' = C_{\mathbb{K}, Y, y}$

**Proof.** Take  $t \in \mathbb{K}$ . Let  $\varepsilon > 0$  take  $\delta = 1$  then if  $0 < |h| < \delta$  we have

$$\begin{aligned}
 \left\| \frac{\sigma(t+h) - \sigma(t)}{h} - y \right\| &= \left\| \frac{x + (t+h) \cdot y - (x + t \cdot y)}{h} - y \right\| \\
 &= \left\| \frac{h \cdot y}{h} - y \right\| \\
 &= 0
 \end{aligned}$$

Which proves that  $\sigma'(t) = y$  so that  $\sigma' = C_{\mathbb{K}, Y, y}$  which is continuous as a constant function. The example is then proved by the previous theorem [see 14.58]  $\square$

**Theorem 14.61.** Let  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle$  be normed spaces,  $U \subseteq X$  a open set,  $V \subseteq Y$  and  $f: U \rightarrow V$  a function then  $f$  is of class  $C^\infty$  iff  $f$  is  $\infty$ -times differentiable.

**Proof.**

$\Rightarrow$ . If  $\forall n \in \mathbb{N}$  we have that  $f$  is of class  $C^n$  then by definition we have that  $f$  is  $n$ -times differentiable proving that  $f$  is  $\infty$ -times differentiable.

$\Leftarrow$ . If  $n=0$  then as  $f$  is differentiable at every  $x \in U$  we have by 14.15 that  $f$  is continuous at  $x$  and thus that  $f$  is continuous on  $U$  proving that  $f$  is of class  $C^0$ . If  $n \in \mathbb{N}_0$  we have that  $f$  is  $n$ -times differentiable and  $f$  is  $(n+1)$ -times differentiable which as  $\forall x \in U$  we have  $D^{n+1}f(x) = D(D^n f)(x)$  proves by 14.15 that  $D^n f$  is continuous at  $x$  and thus that  $D^n f$  is continuous on  $U$  proving that  $f$  is of class  $C^n$ . So we have that  $f$  is of class  $C^\infty$ .  $\square$

**Theorem 14.62.** Let  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle$  be normed spaces,  $U \subseteq X$  a open set,  $V \subseteq Y$ ,  $n \in \mathbb{N}$  and  $f: U \rightarrow V$  a function that is  $n$ -times differentiable then we have that  $\forall m \in \{1, \dots, n\}$   $f$  is  $m$ -times differentiable

**Proof.** We proceed by induction so let  $S_n = \{i \in \mathbb{N} \mid \text{if } 0 \leq i < n \text{ then } f \text{ is } (n-i)\text{-times differentiable}\}$  we have then

$0 \in S_n$ . If  $i=0 \Rightarrow n-i=n$  and as by the hypothesis  $f$  is  $n$ -times differentiable we have  $0 \in S_n$

$i \in S \Rightarrow i+1 \in S$ . If  $i \in S_n$  and  $0 \leq i+1 < n$  [so that  $1 < (n-i)$ ] then as  $0 \leq i < n$  and  $i \in S$  we have that  $f$  is  $(n-i)$ -times differentiable on  $U$ , so by definition [as  $1 < (n-1)$ ] we have that  $f$  is  $((n-i)-1)$ -times differentiable and as  $((n-i)-1) = (n-(i+1))$  we have that  $i+1 \in S_n$

By mathematical induction we have then that  $S_n = \mathbb{N}$ . So if  $m \in \{1, \dots, n\}$  take  $i = n-m \in \{0, \dots, n-1\}$  then as  $i \in \mathbb{N} = S_n$  and  $0 \leq i < n$  that  $f$  is  $(n-i)$ -times differentiable which proves our theorem as  $m = n-i$ .  $\square$

**Theorem 14.63.** Let  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle$  be normed spaces,  $U \subseteq X$  a open set,  $V \subseteq Y$ ,  $n \in \mathbb{N}$  and  $f: U \rightarrow V$  a function of class  $C^n$  then  $\forall i \in \{0, \dots, n\}$  we have that  $f$  is of class  $C^i$

**Proof.** We have to consider the following cases for  $n$ :

$n=0$ . Then  $\{0, \dots, n\} = \{0\}$  which makes the theorem trivial

$n=1$ . Then  $f$  is of class  $C^1$  and thus  $f$  is differentiable everywhere on  $U$  so that by 14.15  $f$  is continuous on  $U$  and thus  $C^0$ .

$1 < n$ . Two cases must be considered:

$m=n$ . then by the hypothesis we have that  $f$  is of class  $C^m$

$m < n$ . Then by the previous theorem  $f$  is  $m$ -times and  $(m+1)$ -times differentiable on  $U$ , so that  $D^m f: U \rightarrow L^m(X; Y)$  is differentiable at every  $x \in U$ , the  $(m+1)$ -times differentiability together with 14.15 means that  $D^m f$  is continuous on  $U$  and thus that  $f$  is  $C^m$ .  $\square$

**Proposition 14.64.** Let  $n, m \in \mathbb{N}_0$ ,  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle$  be normed spaces,  $U \subseteq X$  a open set,  $V \subseteq Y$  and  $f: U \rightarrow V$  be of class  $C^{n+m}$  then  $D^n f$  is of class  $C^m$  and  $D^m(D^n f) = D^{n+m} f$  [notice that by 14.63  $f$  is of class  $C^m$  so that  $D^n f$  is defined]

**Proof.** We prove this by induction so let  $S_n = \{m \in \mathbb{N}_0 \mid \text{if } f: U \rightarrow V \text{ is of class } C^{n+m} \text{ then } D^n f \text{ is of class } C^m \text{ and } D^{n+m} f = D^m(D^n f)\}$  then we have

$1 \in S_n$ . If  $f$  is of class  $C^{n+1}$  then  $D^{n+1} f$  exists on  $U$  and is continuous, so that by definition  $D^{n+1} f = D(D^n f)$ , hence  $D^n f$  is differentiable with  $D^1(D^n f) = D(D^n f)$  a continuous function proving that  $1 \in S_n$ .

$m \in S_n \Rightarrow m+1 \in S_n$ . If  $f$  is of class  $C^{n+(m+1)}$  then by definition  $D^{n+(m+1)} f$  exists on  $U$  and is continuous hence

$$D(D^{n+m} f) \text{ exists and is continuous with } D^{n+(m+1)} f = D(D^{n+m} f) \quad (14.38)$$

Using 14.63  $f$  is of class  $C^{n+m}$  hence as  $m \in \mathcal{S}_n$  we have that  $D^n f$  is of class  $C^m$  and  $D^{n+m} f = D^m(D^n f)$ . So

$$\begin{aligned} D^{n+(m+1)} f &\stackrel{14.38}{=} D(D^{n+m} f) \\ &= D(D^m(D^n f)) \\ &= D^{m+1}(D^n f) \end{aligned}$$

proving using 14.38 that  $D^{m+1}(D^n f)$  is continuous and thus that  $m+1 \in \mathcal{S}_n$   $\square$

**Corollary 14.65.** Let  $n \in \mathbb{N}_0$  with  $n \geq 2$ ,  $\langle X, \|\cdot\|_X \rangle$ ,  $\langle Y, \|\cdot\|_Y \rangle$  be normed spaces,  $U \subseteq X$  a open set and  $f: U \rightarrow Y$  be of class  $C^n$  then  $\forall k \in \{1, \dots, n-1\}$  we have that  $D^{n-k} f$  is of class  $C^k$  with  $D^k(D^{n-k} f) = D^n f$

**Proof.** Take  $m = n - k$  then  $m \in \mathbb{N}_0$  so that  $n = m + k$  hence by 14.64 we have that  $D^m f = D^{n-k} f$  is of class  $C^k$  and  $D^k(D^m f) = D^{m+k} f = D^n f$ .  $\square$

The next theorem is very usable in induction proofs.

**Theorem 14.66.** Let  $\langle X, \|\cdot\|_X \rangle$ ,  $\langle Y, \|\cdot\|_Y \rangle$  be normed spaces,  $U \subseteq X$  a open set,  $V \subseteq Y$ ,  $n \in \mathbb{N}_0 \setminus \{1\}$  and  $f: U \rightarrow V$  a function then  $f$  is  $n$ -times differentiable if and only if  $D^1 f$  exists and is  $(n-1)$ -times differentiable. Also if  $f$  is  $n$ -times differentiable [or equivalent  $D^1 f$  exists and is  $(n-1)$ -times differentiable] then  $D^{n-1}(D^1 f) = D^n f$ .

**Proof.**

$\Rightarrow$ . First as  $n \in \mathbb{N}_0 \setminus \{1\}$  we have that  $1 \in \{1, \dots, n\}$  so using 14.62 we have that  $f$  is 1-times differentiable, hence  $D f = D^1 f$  exists on  $U$ .

For the rest we use induction, so let  $\mathcal{S} = \{n \in \{2, \dots, \infty\} \mid \text{if } f \text{ is } n\text{-times differentiable then } D^1 f \text{ is } (n-1)\text{-times differentiable and } D^n f = D^{n-1}(D^1 f)\}$  then we have

**2  $\in \mathcal{S}$ .** If  $n = 2$  then as  $f$  is 2-times differentiable we have by definition that  $f$  is 1-times differentiable and  $D^2 f = D(D^1 f) = D^1(D^1 f) = D^{2-1}(D^1 f)$  proving that  $2 \in \mathcal{S}$

**$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$ .** Take  $n \in \mathcal{S}$  then if  $f$  is  $n+1$  differentiable we have by definition that  $f$  is  $n$ -times differentiable and  $D^{(n+1)-1} f = D^n f$  is differentiable on  $U$ . Further as  $n \in \mathcal{S}$  we have that  $D^1 f$  is  $(n-1)$ -times differentiable and  $D^{(n-1)}(D^1 f) = D^n f$  so as  $D^n f$  is differentiable we have that  $D^1 f$  is  $n$ -times differentiable and  $D(D^n f) = D(D^{(n-1)}(D^1 f)) = D^n(D^1 f) = D^{(n+1)-1}(D^1 f)$ . Hence  $n+1 \in \mathcal{S}$

Using mathematical induction we have then

if  $f$  is  $n$ -times differentiable with  $n \in \mathbb{N}_0 \setminus \{1\}$  then  $D^1 f$  is  $(n-1)$ -times differentiable and  $D^n f = D^{n-1}(D^1 f)$   $\quad (14.39)$

$\Leftarrow$ . This is proved by induction so take  $\mathcal{S} = \{n \in \{2, \dots, \infty\} \mid \text{if } D^1 f \text{ exists and is } (n-1)\text{-times differentiable then } f \text{ is } n\text{-times differentiable}\}$  then we have

**2  $\in \mathcal{S}$ .** As  $D^1 f$  exists and is  $(2-1)$ -times differentiable, it follows that  $D^1 f$  is differentiable and thus by definition,  $f$  is 2-times differentiable.

**$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$ .** Assume that  $D^1 f$  is  $((n+1)-1)$ -times differentiable then  $D^1 f$  is  $n$ -times differentiable. By definition  $D^1 f$  is then  $(n-1)$ -times differentiable. From  $n \in \mathcal{S}$  it follows then that  $f$  is  $n$ -times differentiable and using 14.39 we have that  $D^n f = D^{n-1}(D^1 f)$ . As  $D^1 f$  is  $n$ -times differentiable we have by definition that the differential of  $D^{n-1}(D^1 f) = D^n f$  exists. So  $f$  is  $(n+1)$ -times differentiable proving that  $n+1 \in \mathcal{S}$ .  $\square$

Next we extend the above to show that it also works for differentiable classes

**Theorem 14.67.** Let  $\langle X, \|\cdot\|_X \rangle$ ,  $\langle Y, \|\cdot\|_Y \rangle$  be normed spaces,  $U \subseteq X$  a open set,  $V \subseteq Y$ ,  $n \in \mathbb{N}_0$  and  $f: U \rightarrow V$  a function then  $f$  is of class  $C^n$  if and only if  $D^1 f$  exists and is of class  $C^{n-1}$

**Proof.** The following cases must be considered

**$n = 1$ .** If  $f$  is of class  $C^1$  then by definition  $D^1f = Df$  is continuous and thus of class  $C^0$ . If  $D^1f$  exists and is of class  $C^0$  then  $f$  is 1-times differentiable with  $D^1f$  continuous and thus of class  $C^1$ .

**$n > 1$ .**

$\Rightarrow$ . If  $f$  is of class  $C^n$  then  $f$  is  $n$ -times differentiable and  $D^n f$  is continuous, using the previous theorem we have then that  $D^1 f$  is  $(n-1)$  times differentiable and  $D^{n-1}(D^1 f) = D^n f$  which is continuous so that  $D^1 f$  is of class  $C^{n-1}$

$\Leftarrow$ . If  $D^1 f$  is of class  $C^{n-1}$  then it is  $n-1$  times differentiable and continuous so again by the previous theorem we have then that  $f$  is  $n$ -times differentiable and as  $D^n f = D^{n-1}(D^1 f)$  which is continuous we have that  $f$  is of class  $C^n$ .  $\square$

**Corollary 14.68.** Let  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle$  be normed spaces,  $U \subseteq X$  a open set,  $V \subseteq Y$ ,  $n \in \mathbb{N}_0$  and  $f: U \rightarrow V$  a function then  $f$  is of class  $C^\infty$  if and only if  $D^1 f$  exists and is of class  $C^\infty$

**Proof.** Assume that  $D^1 f$  exists and is of class  $C^\infty$  then for  $n \in \mathbb{N}_0$  we have either

**$n = 0$ .** then as  $D^1 f$  exists we have by 14.15 that  $f$  is continuous on  $U$  hence  $f$  is of class  $C^\infty$

**$1 \leq n$ .** then as  $D^1 f$  is  $C^\infty$  we have by definition that  $D^1 f$  is of class  $C^{n-1}$  so using 14.67 we have that  $f$  is of class  $C^n$

So  $\forall n \in \mathbb{N}_0$  we have that  $f$  is of class  $C^n$  proving that  $f$  is of class  $C^\infty$ . On the other side if  $f$  is of class  $C^\infty$  then given  $n \in \mathbb{N}$  we have that  $f$  is of class  $C^{n+1}$  and using 14.67 we have that  $D^1 f$  exists and  $D^1 f$  is of class  $C^{(n+1)-1} = C^n$  proving that  $D^1 f$  is of class  $C^\infty$ .  $\square$

**Theorem 14.69.** Let  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle$  be normed spaces,  $U \subseteq X$  a open set,  $V \subseteq Y$ ,  $n \in \mathbb{N}_0$  and  $f: U \rightarrow V$  a function such that  $f$  is  $n$ -times differentiable on  $U$  then  $\forall W \subseteq U$ ,  $W$  open we have  $f|_W$  is  $n$ -times differentiable on  $W$  and  $(D^n f)|_W = D^n(f|_W)$

**Proof.** We prove this by induction. So let  $\mathcal{S} = \{n \in \mathbb{N}_0 \mid \text{if } f \text{ is } n\text{-times differentiable on } U \text{ then for all open } W \subseteq U \text{ we have that } f|_W \text{ is } n\text{-times differentiable with } (D^n f)|_W = D^n(f|_W)\}$  then we have

**$1 \in \mathcal{S}$ .** If  $f$  is 1-times differentiable on  $U$  then for a open  $W \subseteq U$  we have that  $\forall x \in W D^1 f(x) = Df(x)$  exists. Using 14.10 we have that  $\forall x \in W f|_W$  is differentiable at  $x$  with  $D(f|_W)(x) = Df(x) \underset{x \in W}{=} (Df)|_W(x)$ , proving that  $f|_W$  is 1-times differentiable on  $W$  and  $D^1(f|_W) = (D^1 f)|_W$ .

**$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$ .** Assume that  $f$  is  $(n+1)$ -times differentiable on  $U$  then by definition  $f$  is  $n$ -times differentiable and  $\forall x \in U$  we have that  $D^n f$  is differentiability at  $x$  with  $D(D^n f)(x) = D^{n+1} f(x)$ . Let  $x \in W \subseteq U$  then by 14.10  $(D^n f)|_W(x)$  is differentiable at  $x$  and

$$\begin{aligned} (D^{n+1} f)|_W(x) &\underset{x \in W}{=} (D(D^n f))|_W(x) \\ &\underset{14.10}{=} (D(D^n f)|_W)(x) \\ &\underset{n \in \mathcal{S}}{=} (D(D^n f|_W))(x) \\ &\underset{\text{def}}{=} D^{n+1}(f|_W)(x) \end{aligned}$$

proving that  $(D^{n+1} f)|_W = D^{n+1}(f|_W)$  and thus

$$n+1 \in \mathcal{S}$$

Induction proves then the theorem.  $\square$

Next we prove that the above theorem is also true for classes

**Theorem 14.70.** Let  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle$  be normed spaces,  $U \subseteq X$  a open set,  $V \subseteq Y$ ,  $n \in \mathbb{N}$  and  $f: U \rightarrow V$  a function such that  $f$  is of class  $C^n$  then if  $V \subseteq U$  is open we have that  $f|_W$  is of class  $C^n$ . Further if  $n > 0$  then  $D^n(f|_W) = (D^n f)|_W$

**Proof.** We have to consider the following cases for  $n \in \mathbb{N}$

**$n = 0$ .** This follows from 11.137.

**$n > 0$ .** Using the previous theorem we have that if  $V$  is open and  $V \subseteq U$  then  $f|_V$  is  $n$ -times differentiable with  $D^n(f|_V) = (D^n f)|_V$  which is continuous on  $V$  by 11.137.  $\square$

**Theorem 14.71.** Let  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle$  be normed spaces,  $U \subseteq X, V_f, V_g \subseteq Y, n \in \mathbb{N}_0$  and  $f: U \rightarrow V_f, g: U \rightarrow V_g$  functions then

1. If  $f, g$  are  $n$ -times differentiable functions then  $f + g: U \rightarrow V_f + V_g$  is  $n$ -times differentiable and  $D^n(f + g) = D^n f + D^n g$
2. If  $f$  is  $n$ -times differentiable then  $\forall \alpha \in \mathbb{K}$  we have that  $\alpha \cdot f: U \rightarrow \alpha \cdot V_f$  is  $n$ -times differentiable with  $D^n(\alpha \cdot f) = \alpha \cdot D^n f$

**Proof.**

1. We prove this by induction so let  $\mathcal{S} = \{n \in \mathbb{N}_0 \mid \text{if } f, g \text{ are } n\text{-times differentiable then } f + g \text{ is } n\text{-times differentiable with } D^n(f + g) = D^n f + D^n g\}$  then we have

**$1 \in \mathcal{S}$ .** Then as  $f, g$  are 1-times differentiable on  $U$  we have  $\forall x \in U$  that  $Df(x)$  and  $Dg(x)$  exists, using 14.21 it follows that  $D(f + g)(x)$  exists and  $D(f + g)(x) = Df(x) + Dg(x)$ . Hence  $f + g$  is 1-times differentiable with  $D^1(f + g) = D^1(f) + D^1(g)$  proving that  $1 \in \mathcal{S}$ .

**$n \in \mathcal{S} \Rightarrow n + 1 \in \mathcal{S}$ .** If  $f, g$  are  $(n + 1)$ -times differentiable then  $f, g$  are  $n$ -times differentiable and  $D^n f, D^n g$  are differentiable at  $x \forall x \in U$ . Then for  $x \in U$

$$\begin{aligned} D^{n+1}(f + g)(x) &= D(D^n(f + g))(x) \\ &\stackrel{n \in \mathcal{S}}{=} D(D^n f + D^n g)(x) \\ &\stackrel{14.21}{=} D(D^n f)(x) + D(D^n g)(x) \\ &\stackrel{\text{def}}{=} D^{n+1}f(x) + D^{n+1}g(x) \\ &= (D^{n+1}f + D^{n+1}g)(x) \end{aligned}$$

proving that  $D^{n+1}(f + g) = D^{n+1}f + D^{n+1}g$ . So we have

$$n + 1 \in \mathcal{S}$$

2. We prove this by induction so let  $\mathcal{S} = \{n \in \mathbb{N}_0 \mid \text{if } f \text{ is } n\text{-times differentiable then } \alpha \cdot f \text{ is } n\text{-times differentiable with } D^n(\alpha \cdot f) = \alpha \cdot D^n f\}$  then we have

**$1 \in \mathcal{S}$ .** then as  $f$  1-times differentiable on  $U$  we have  $\forall x \in U$  that  $Df(x)$  exists, using 14.21 we have that  $D(\alpha \cdot f)(x)$  exists and  $D(\alpha \cdot f)(x) = \alpha \cdot Df(x)$ . Hence  $\alpha \cdot f$  is 1-times differentiable with  $D^1(\alpha \cdot f) = \alpha \cdot D^1(f)$  proving that  $1 \in \mathcal{S}$ .

**$n \in \mathcal{S} \Rightarrow n + 1 \in \mathcal{S}$ .** if  $f$  is  $(n + 1)$ -times differentiable then  $f$  is  $n$ -times differentiable and  $D^n f$  are differentiable  $\forall x \in U$ . Then for  $x \in U$

$$\begin{aligned} D^{n+1}(\alpha \cdot f)(x) &= D(D^n(\alpha \cdot f))(x) \\ &\stackrel{n \in \mathcal{S}}{=} D(\alpha \cdot D^n f)(x) \\ &\stackrel{14.21}{=} \alpha \cdot D(D^n f)(x) \\ &= \alpha \cdot D^{n+1}f(x) \\ &= (\alpha \cdot D^{n+1}f)(x) \end{aligned}$$

proving that  $D^{n+1}(\alpha \cdot f) = \alpha \cdot D^{n+1}f$ , giving

$$n + 1 \in \mathcal{S}$$

$\square$

**Theorem 14.72.** Let  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle$  be normed spaces,  $U \subseteq X, V_f, V_g \subseteq Y, n \in \mathbb{N}_0$  and  $f: U \rightarrow V_f, g: U \rightarrow V_g$  functions then

1. If  $f, g$  are of class  $C^n$  then  $f + g: U \rightarrow V_f + V_g$  and  $\alpha \cdot f$  are  $C^n$
2. If  $f, g$  are of class  $C^\infty$  then  $\alpha \cdot f$  and  $f + g$  are of class  $C^\infty$

**Proof.**

1. The following cases must be considered for  $n \in \mathbb{N}_0$

**$n = 0$ .** As  $f, g$  are  $C^0$  we have that  $f, g$  are continuous on  $U$  hence by 11.175 we have that  $f + g$  and  $\alpha \cdot f$  are continuous. So by definition  $f + g$  and  $\alpha \cdot f$  are of class  $C^0$ .

**$n > 0$ .** Using the previous theorem we have already that  $f + g$  and  $\alpha \cdot f$  are  $n$ -times differentiable with  $D^n(f + g) = D^n f + D^n g$  and  $D^n(\alpha \cdot f) = \alpha \cdot D^n f$ . As  $D^n f, D^n g$  are also continuous we have by 11.175 that  $D^n(f + g) = D^n f + D^n g$  and  $D^n(\alpha \cdot f) = \alpha \cdot D^n f$  are continuous proving that  $f + g, \alpha \cdot f$  are of class  $C^n$

2. If  $f, g$  are of class  $C^\infty$  then  $\forall n \in \mathbb{N}_0$   $f, g$  are of class  $C^n$ , hence by (1) it follows that  $\alpha \cdot f$  and  $f + g$  are of class  $C^n$  proving that  $\alpha \cdot f$  and  $f + g$  are of class  $C^\infty$ .  $\square$

Using mathematical induction it is easy to extend the above theorems to a finite sum.

**Theorem 14.73.** Let  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle$  be normed spaces  $U \subseteq X$  a open set,  $k \in \mathbb{N}_0, n \in \mathbb{N}$   $\{f_i: U \rightarrow Y\}_{i \in \{1, \dots, k\}}$  a finite set of functions from which are  $n$ -times differentiable (in which case  $n > 0$ ) [or  $C^n$ ] then  $\sum_{i \in \{1, \dots, k\}} f_i: U \rightarrow Y$  is  $n$ -times differentiable [or  $C^n$ ].

**Proof.** We prove this by induction so let  $S = \{k \in \mathbb{N}_0 \mid \text{if } \{f_i: U \rightarrow Y\}_{i \in \{1, \dots, k\}} \text{ is a finite set of functions from } U \text{ to } Y \text{ which are } n\text{-times differentiable } (n > 0) \text{ [or } C^n\text{] then } \sum_{i \in \{1, \dots, k\}} f_i: U \rightarrow Y \text{ is } n\text{-times differentiable [or } C^n\text{]}\}$  then we have:

**$1 \in S$ .** Then  $\sum_{i \in \{1, \dots, 1\}} f_i = f_1: U \rightarrow Y$  which is 1-times differentiable [or  $C^1$ ] so that  $1 \in S$

**$k \in S \Rightarrow k + 1 \in S$ .** If  $\{f_i: U \rightarrow Y\}_{i \in \{1, \dots, k+1\}}$  are  $n$ -times differentiable functions [or  $C^n$  functions] then we have, as  $k \in S$ , that  $\sum_{i=1}^k f_i$  is  $n$ -times differentiable [or  $C^n$ ]. As  $\sum_{i=1}^{k+1} f_i = (\sum_{i=1}^k f_i) + f_{k+1}$  is the sum of  $n$ -times differentiable [or  $C^n$ ] function we have by 14.71 [or 14.72] that  $\sum_{i=1}^{k+1} f_i$  is  $n$ -times differentiable [or  $C^n$ ]. Hence  $n + 1 \in S$   $\square$

**Theorem 14.74.** Let  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle$  be normed spaces  $U \subseteq X$  open,  $y \in Y, y \in V \subseteq Y$  and  $C_{U,V,y}: U \rightarrow V$  a constant function then  $C_{U,V,y}$  is  $C^\infty$  and we have  $\forall n \in \mathbb{N}_0$  that  $D^n C_{U,V,y} = C_{U,L^n(X,Y),0}$

**Proof.** Let  $n \in \mathbb{N}$  then we have either

**$n = 0$ .** Then as  $C_{U,Y,y}$  is continuous (see 11.140) we have by definition that  $C_{U,Y,y}$  is of class  $C^0$

**$n \in \mathbb{N}_0$ .** We proceed then by induction, let  $S = \{n \in \mathbb{N}_0 \mid \text{then } D^n C_{U,Y,y} \text{ exists and } D^n C_{U,Y,y} = C_{U,L^n(X,Y),0}\}$  then we have

**$1 \in S$ .**  $\forall x \in U$  we have by 14.16 that  $D^1 C_{U,Y,y}(x) = D C_{U,Y,y}(x) = C_{X,Y,0}$  [0 the neutral element in  $Y$ ], as  $C_{X,Y,0}$  is the neutral element [see 10.115] in  $L(X, Y) = L^1(X; Y)$  we have that  $D^1 C_{U,Y,y}$  exists and is equal to  $C_{U,L^1(X;Y),0}$  proving that  $1 \in S$ .

**$n \in S \Rightarrow n + 1 \in S$ .** As  $n \in S$  we have  $D^n C_{U,Y,y} = C_{U,L^n(X;Y),0}$  so that  $D^{n+1}(C_{U,Y,y})(x) = D(D^n C_{U,Y,y})(x) = D(C_{U,L^n(X;Y),0})(x) \stackrel{14.16}{=} C_{X,L^n(X;Y),0}$ . As  $C_{X,L^n(X;Y),0}$  is the neutral element in  $L(X, L^n(X;Y)) \stackrel{11.204}{=} L^{n+1}(X; Y)$  [see 10.115] we have that  $D^{n+1}(C_{U,Y,y})$  exists and is equal to  $C_{U,L^{n+1}(X;Y),0}$ . Which proves that  $n + 1 \in S$

So  $\forall n \in \mathbb{N}$  we have that  $C_{U,Y,y}$  is of class  $C^n$  proving that  $C_{U,Y,y}$  is of class  $C^\infty$  and  $\forall n \in \mathbb{N}_0 D^n C_{U,Y,y} = C_{U,L^n(X,Y),0}$ .  $\square$

**Theorem 14.75.** Let  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle$  be normed spaces over  $\mathbb{K}$  and  $L \in L(X, Y)$  a linear continuous map then we have

1.  $L$  is of class  $C^\infty$
2.  $\forall x \in X D^1 L(x) = L$  [so  $D^1 L = C_{X,L(X,Y),L}$ ]

3.  $\forall n \in \mathbb{N}_0 \setminus \{1\}$  then  $D^n L = C_{X, L^n(X; Y), 0}$

**Proof.** Let  $n \in \mathbb{N}_0$  then we have the following possible cases to consider

**n = 0.** As  $L$  is continuous  $L$  is of class  $C^0$

**n = 1.** Using 14.17 we have  $\forall x \in X$  that  $DL(x) = L$  exists, proving (2) and that  $L$  is of class  $C^1$

**$n \in \mathbb{N}_0 \setminus \{1\}$ .** Using (1) it follows that  $D^1 L = C_{X, L(X; Y), L}$ . Next

$$\begin{aligned} D^n L &\stackrel{14.66}{=} D^{n-1}(D^1 L) \\ &\stackrel{(2)}{=} D^{n-1}(C_{X, L(X; L), L}) \\ &\stackrel{14.74}{=} C_{X, L^{n-1}(X, L(X; Y)), 0} \\ &\stackrel{11.209}{=} C_{X, L^n(X; Y), 0} \end{aligned}$$

which proves (3).

As  $\forall n \in \mathbb{N}_0$   $D^n L$  exists we have using 14.61 that  $L$  is of class  $C^\infty$ .  $\square$

**Corollary 14.76.** Given a finite family  $\{\langle X_i, \|\cdot\|_i \rangle\}_{i \in \{1, \dots, n\}}$ ,  $n \in \mathbb{N}_0 \setminus \{1\}$  of normed vector spaces then given  $i \in \{1, \dots, n\}$  and  $x = (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n) \in \prod_{i \in \{1, \dots, n\} \setminus \{i\}} X_i = X$  then  $(i \rightarrow x)$ :  $X_i \rightarrow \prod_{j \in \{1, \dots, n\}} X_j$  [see 14.25] is  $C^\infty$ . Further  $D(i \rightarrow x) = (i \rightarrow 0)$ .

**Proof.** As  $(x \rightarrow i)(t) \stackrel{14.29}{=} (i \rightarrow x)(0) + (i \rightarrow 0)(t)$  a sum of a constant function [which is of class  $C^\infty$  by 14.74] and a linear function [see 14.30] [which is of class  $C^\infty$  by 14.75] and is thus  $C^\infty$  [see 14.72]. Further  $D(x \rightarrow i)(t) = D((i \rightarrow x) + (i \rightarrow 0)(t)) = C_{X_i, X, 0} + (i \rightarrow 0) = (i \rightarrow 0)$ .  $\square$

**Proposition 14.77.** Let  $\langle X, \|\cdot\|_X \rangle$  be a normed space and  $\tau_{x_0}: X \rightarrow X$  defined by  $x \rightarrow \tau_{x_0}(x) = x + x_0$  (translation function) then we have

1.  $\tau_{x_0}$  is  $C^\infty$
2.  $D^1 \tau_{x_0} = C_{X, L(X; Y), \mathbb{1}_X}$
3.  $\forall n \in \mathbb{N}_0 \setminus \{1\}$  we have  $D^n \tau_{x_0} = C_{X, L^n(X; Y), 0}$

**Proof.**

1. Let  $x \in X$  then  $(\mathbb{1}_X + C_{X, X, x_0})(x) = \mathbb{1}_X(x) + C_{X, X, x_0}(x) = x + x_0 = \tau_{x_0}(x)$  proving that

$$\tau_{x_0} = \mathbb{1}_X + C_{X, X, x_0} \quad (14.40)$$

As  $\mathbb{1}_X \in L(X, X)$  [see 11.182] we have by 14.75, 14.74 and 14.72 that  $\tau_{x_0}$  is of class  $C^\infty$ .

2. Let  $x \in X$  then

$$\begin{aligned} D^1(\tau_{x_0})(x) &= D(\mathbb{1}_X + C_{X, X, x_0})(x) \\ &\stackrel{14.71}{=} D^1(\mathbb{1}_X)(x) + D^1(C_{X, X, x_0})(x) \\ &\stackrel{14.75 \text{ and } 14.74}{=} \mathbb{1}_X(x) + C_{X, L(X; Y), 0}(x) \\ &= x + 0 \\ &= \mathbb{1}_X(x) \end{aligned}$$

proving that  $D^1(\tau_{x_0}) = C_{X, L(X; X), \mathbb{1}_X}$ .

3. Let  $x \in \mathbb{N}_0 \setminus \{1\}$  we have

$$\begin{aligned} D^n(\tau_{x_0})(x) &= D^n(\mathbb{1}_X + C_{X, X, x_0})(x) \\ &\stackrel{14.71}{=} D^n(\mathbb{1}_X)(x) + D^n(C_{X, X, x_0})(x) \\ &\stackrel{14.75 \text{ and } 14.74}{=} C_{X, L^n(X; Y), 0}(x) + C_{X, L^n(X; Y), 0}(x) \\ &= 0 + 0 \\ &= C_{X, L^n(X; Y), 0}(x) \end{aligned}$$

proving that  $D^n(\tau_{x_0}) = C_{X, L^n(X; Y), 0}$ .  $\square$

**Theorem 14.78.** Let  $\langle X_1, \|\cdot\|_1 \rangle, \langle X_2, \|\cdot\|_2 \rangle$  and  $\langle Y, \|\cdot\|_Y \rangle$  be normed spaces.  $L \in L(X_1, X_2; Y)$  a continuous bi-linear mapping from  $X_1 \times X_2 \rightarrow Y$  then  $L$  is  $C^\infty$ . Further we have

1.  $D^1 L: X_1 \times X_2 \rightarrow L(X_1 \times X_2, Y)$  is defined by  $D^1 L(x, y) = L(x, *) + L(*, y)$  where  $L(x, *) + L(*, y) \in L(X_1 \times X_2, Y)$  is defined by  $(r, s) \rightarrow (L(x, *) + L(*, y))(r, s) = L(x, s) + L(r, y)$
2.  $D^2 L = C_{X_1 \times X_2, L^2(X_1 \times X_2; Y), D^1 L}$
3. If  $n > 2$  then  $D^n L = C_{X_1 \times X_2, L^n(X_1 \times X_2, Y), 0}$

**Proof.** We have to consider the following cases for  $n \in \mathbb{N}_0$

**$n = 0$ .** As  $L$  is continuous we have that  $L$  is of class  $C^0$

**$n = 1$ .** Using the differential of a bi-linear mapping [see 14.47] we have that  $L$  is differentiable at every  $(x, y) \in X_1 \times X_2$  with  $D^1 L(x, y) = DL(x, y) = L(x, *) + L(*, y)$  so that  $D^1 L$  is defined by  $D^1 L(x, y) = L(x, *) + L(*, y)$

**$n = 2$ .** We prove now that  $D^1 L: X_1 \times X_2 \rightarrow L(X_1 \times X_2, Y)$  is linear and continuous. If  $\alpha, \beta \in \mathbb{K}$  and  $(r, s), (x_1, y_1), (x_2, y_2) \in X_1 \times X_2$  then

$$\begin{aligned} D^1 L((\alpha \cdot (x_1, y_1) + \beta \cdot (x_2, y_2)))(r, s) &= D^1 L(\alpha \cdot x_1 + \beta \cdot x_2, \alpha \cdot y_1 + \beta \cdot y_2)(r, s) \\ &= (L(\alpha \cdot x_1 + \beta \cdot x_2, *) + L(*, \alpha \cdot y_1 + \beta \cdot y_2))(r, s) \\ &= L(\alpha \cdot x_1 + \beta \cdot x_2, s) + L(r, \alpha \cdot y_1 + \beta \cdot y_2) \\ &= \alpha \cdot L(x_1, s) + \alpha \cdot L(r, y_1) + \beta \cdot L(x_2, s) + \beta \cdot L(r, y_2) \\ &= \alpha \cdot (L(x_1, *) + L(*, y_1))(r, s) + \beta \cdot (L(x_2, *) + L(*, y_2))(r, s) \\ &= \alpha \cdot D^1 L(x_1, y_1)(r, s) + \beta \cdot D^1 L(x_2, y_2)(r, s) \end{aligned}$$

proving that  $D^1 L((\alpha \cdot (x_1, y_1) + \beta \cdot (x_2, y_2))) = \alpha \cdot D^1 L(x_1, y_1) + \beta \cdot D^1 L(x_2, y_2)$ , hence

$$D^1 L \text{ is linear} \tag{14.41}$$

Further we have that

$$\begin{aligned} \|D_1 L(x, y)(r, s)\|_Y &= \|L(x, s) + L(r, y)\|_Y \\ &\leq \|L(x, s)\|_Y + \|L(r, y)\|_Y \\ &\stackrel{11.198}{\leq} \|L\| \cdot (\|x\|_{X_1} \cdot \|s\|_{X_2}) + \|L\| \cdot (\|y\|_{X_2} \cdot \|r\|_{X_1}) \\ &= \|L\| \cdot (\|x\|_{X_1} \cdot \|s\|_{X_1} + \|y\|_{X_2} \cdot \|r\|_{X_2}) \\ &\leq \|L\| \cdot ((\|x\|_{X_1} + \|y\|_{X_2}) \cdot \max\{\|r\|_{X_1}, \|s\|_{X_2}\}) \\ &= (\|L\| \cdot (\|x\|_X + \|y\|_Y)) \cdot \|(r, s)\|_{X_1 \times X_2} \end{aligned}$$

Using 11.198 we have then that

$$\begin{aligned} \|D_1 L(x, y)\|_{L(X_1 \times X_2, Y)} &\leq \|L\| \cdot (\|x\|_{X_1} + \|y\|_{X_2}) \\ &\leq \|L\| \cdot 2 \cdot \max\{\|x\|_X, \|y\|_Y\} \\ &= 2 \cdot \|L\| \cdot \|(x, y)\|_{X_1 \times X_2} \end{aligned}$$

which proves that

$$D^1 L \text{ is continuous} \tag{14.42}$$

Using 14.41 and 14.42 it follows then that

$$D^1 L \in L(X_1 \times X_2, L(X_1 \times X_2, Y))$$

Using then 14.75 we have that  $D^1 L$  is differentiable and  $D^2 L = D^1(D^1 L) = C_{X_1 \times X_2, L(X_1 \times X_2, L(X_1 \times X_2, Y)), D^1 L} = C_{X_1 \times X_2, L^2(X_1 \times X_2; Y), D^1 L}$  proving (2).

**$n > 2$ .** We have

$$\begin{aligned} D^n L &\stackrel{14.66}{=} D^{n-1}(D^1 L) \\ &\stackrel{14.75}{=} C_{X_1 \times X_2, L^{n-1}(X_1 \times X_2, L(X_1 \times X_2, L(X_1 \times X_2, Y))), 0} \\ &\stackrel{11.209}{=} C_{X_1 \times X_2, L^n(X_1 \times X_2, Y), 0} \end{aligned}$$

proving (3)  $\square$

**Theorem 14.79.** Let  $\{\langle X_i, \|\cdot\|_i \rangle\}_{i \in \{1, \dots, n\}}$  be a finite family of normed spaces and  $i \in \{1, \dots, n\}$  then the projection map  $\pi_i: \prod_{j \in \{1, \dots, n\}} X_j \rightarrow X_i$  is  $C^\infty$  with  $D^1 \pi_i = C_{\prod_{j \in \{1, \dots, n\}} X_j, X_i, \pi_i}$

**Proof.** This is trivial as the projection maps are linear and continuous (see 11.148) so that we can use 14.75.  $\square$

**Theorem 14.80.** Let  $\langle X, \|\cdot\|_X \rangle$  be a normed space,  $\{\langle Y_i, \|\cdot\|_i \rangle\}_{i \in \{1, \dots, n\}}$  a family of normed spaces and  $\|\cdot\|$  the maximum norm defined on  $\prod_{i \in \{1, \dots, n\}} Y_i$  then given  $U \subseteq X$ ,  $V \subseteq Y = \prod_{i \in \{1, \dots, n\}} Y_i$  open sets and  $r \in \mathbb{N}_0$  we have that  $f: U \rightarrow V$  is  $r$ -times differentiable if and only if  $\forall i \in \{1, \dots, n\}$  we have  $\pi_i \circ f: U \rightarrow \pi_i(V)$  is  $r$ -times differentiable and then  $\pi_i \circ D^r f = D^r(\pi_i \circ f)$  (or equivalently  $D^r f = (D^r(\pi_1 \circ f), \dots, D^r(\pi_n \circ f))$ )

**Proof.** We prove this by induction so let  $S = \{r \in \mathbb{N}_0 \mid \text{Let } f: U \rightarrow \prod_{i \in \{1, \dots, n\}} Y_i \text{ then } f \text{ is } r\text{-times differentiable if and only if } \forall i \in \{1, \dots, n\} \pi_i \circ f \text{ is } r\text{-times differentiable and then } \pi_i \circ D^r f = D^r(\pi_i \circ f)\}$  then:

$1 \in S$ . If  $r = 1$  then using 14.43 we find that  $\forall x \in U$  that  $f$  is differentiable at  $x \in U$  iff  $\forall i \in \{1, \dots, n\} \pi_i \circ f$  is differentiable at  $x$  and that  $\pi_i \circ Df(x) = D(\pi_i \circ f)(x)$  so that  $1 \in S$

$r \in S \Rightarrow r + 1 \in S$ . Let  $r \in S$  then we have

$\Rightarrow$ . Take  $i \in \{1, \dots, n\}$  and assume that  $f$  is  $r + 1$  times differentiable. By definition  $f$  is then  $r$ -times differentiable, as  $r \in S$  we have that  $\pi_i \circ D^1 f$  is  $r$ -times differentiable and  $D^r(\pi_i \circ f) = \pi_i \circ D^r f$ . Further as  $f$  is  $r + 1$  times differentiable  $D^r f$  is 1-times differentiable, so using the fact that  $1 \in S$  we have that  $\pi_i \circ D^r f$  is differentiable with  $D^1(\pi_i \circ D^r f) = \pi_1 \circ D^1(D^r f) = \pi_i \circ D^{r+1} f$ . Hence  $D^r(\pi_i \circ f)$  is differentiable (proving that  $\pi_i \circ f$  is  $(r + 1)$ -times differentiable) with  $D^{r+1}(\pi_i \circ f) = D^1(D^r(\pi_i \circ f)) = \pi_i \circ D^{r+1} f$ . So we have that  $r + 1 \in S$ .

$\Leftarrow$ . Assume that  $\pi_i \circ f$  is  $(r + 1)$ -times differentiable and  $\pi_i \circ D^{r+1} f = D^{r+1}(\pi_i \circ f)$ , then  $D^1(\pi_i \circ f)$  is  $r$ -times differentiable by 14.66 and thus as  $r \in S$  we have that  $D^r f$  is  $r$ -times differentiable, using again 14.66 we have then that  $f$  is  $(r + 1)$ -times differentiable.  $\square$

**Theorem 14.81.** Let  $\langle X, \|\cdot\|_X \rangle$  be a normed space,  $\{\langle Y_i, \|\cdot\|_i \rangle\}_{i \in \{1, \dots, n\}}$  a family of normed spaces and  $\|\cdot\|$  the maximum norm defined on  $\prod_{i \in \{1, \dots, n\}} Y_i$  then given  $U \subseteq X$ ,  $V \subseteq \prod_{i \in \{1, \dots, n\}} Y_i$  open sets and  $r \in \mathbb{N}$  we have that  $f: U \rightarrow V$  is  $C^r$  iff  $\pi_i \circ f$  is  $C^r \forall i \in \{1, \dots, n\}$

**Proof.** We have the following cases to consider for  $r \in \mathbb{N}_0$

1. ( $r = 0$ ) This follows from 11.149.

2. ( $r > 0$ )

$\Rightarrow$ . If  $f$  is  $C^r$  then  $f$  is  $r$ -times differentiable, using the previous theorem we have then  $\forall i \in \{1, \dots, n\}$  that  $\pi_i \circ f$  is  $r$ -times differentiable with  $D^r(\pi_i \circ f) = \pi_i \circ D^r f$ , as  $\pi_i$  and  $D^r f$  is continuous we have that  $\pi_i \circ f$  is  $C^r$

$\Leftarrow$ . If  $\forall i \in \{1, \dots, n\}$  we have  $\pi_i \circ f$  is  $C^r$  then  $\pi_i \circ f$  is  $r$ -times differentiable, so by the previous theorem we have that  $f$  is  $r$ -times differentiable and also by the previous theorem we must then have that  $\pi_i \circ D^r f = D^r(\pi_i \circ f)$  proving that  $\pi_i \circ D^r f$  is continuous and by 11.149 that  $D^r f$  is continuous. So we have that  $f$  is  $C^r$ .  $\square$

**Lemma 14.82.** Let  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle, \langle Z, \|\cdot\|_Z \rangle$  be normed spaces,  $R \subseteq Y$ ,  $S \subseteq Z$ , open set  $U \subseteq X$ , open set  $V \subseteq Y$  and  $f: U \rightarrow R$ ,  $g: V \rightarrow S$  functions then we have

1. If  $f$  and  $g$  are 1-times differentiable then

a.  $g \circ f: U \cap f^{-1}(V) \rightarrow S$  is 1-times differentiable

b.  $(\circ): L(Y, Z) \times L(X, Y) \rightarrow L(X, Z)$  defined by  $(\circ)(L, K) = L \circ K$  is of class  $C^\infty$

- c.  $h: U \rightarrow L(Y, Z) \times L(X, Y)$  defined by  $h(x) = (D^1g(f(x)), D^1f(x))$  is continuous if  $f, g$  are of class  $C^1$
- d.  $D^1(g \circ f) = D(g \circ f) = (\circ) \circ h, \pi_1 \circ h = (D^1g) \circ f$  and  $\pi_2 \circ h = D^1f$
- 2. If  $f$  and  $g$  are of class  $C^1$  then  $g \circ f$  is of class  $C^1$

**Proof.**

1. As  $f$  and  $g$  are 1-times differentiable we have  $\forall x \in U$  that  $Df(x)$  and  $Dg(f(x))$ , applying then the chain rule [see 14.22] gives

$$g \circ f \text{ is 1-times differentiable and } \forall x \in U \text{ we have } D^1(g \circ f)(x) = D^1g(f(x)) \circ Df(x) \quad (14.43)$$

Using 11.203 we have that  $(\circ)$  is continuous and bi-linear so by 14.78 we have that

$$(\circ) \text{ is of class } C^\infty \quad (14.44)$$

If  $x \in U$  then

$$\begin{aligned} (\pi_1 \circ h)(x) &= \pi_1(D^1g(f(x)), D^1f(x)) \\ &= D^1g(f(x)) \\ &= ((D^1g) \circ f)(x) \\ (\pi_2 \circ h)(x) &= \pi_2(D^1g(f(x)), D^1f(x)) \\ &= D^1f(x) \end{aligned}$$

proving that

$$\pi_1 \circ h = (D^1g) \circ f \text{ and } \pi_2 \circ h = D^1f \quad (14.45)$$

If now  $f, g$  are of class  $C^1$  then  $D^1g, D^1f$  and  $f$  are continuous so that  $\pi_1 \circ h$  and  $\pi_2 \circ h$  will be continuous. Using then 11.149 gives

$$\text{If } f, g \text{ are over class } C^1 \text{ then } h \text{ is continuous} \quad (14.46)$$

Let  $x \in U$  then

$$\begin{aligned} ((\circ) \circ h)(x) &= (\circ)(h(x)) \\ &= (\circ)(D^1g(f(x)), D^1f(x)) \\ &= D^1g(f(x)) \circ D^1f(x) \\ &\stackrel{14.43}{=} D^1(g \circ f)(x) \end{aligned}$$

giving

$$D^1(g \circ f) = D(g \circ f) = (\circ) \circ h$$

2. If  $f, g$  are of class  $C^1$  then  $f, g$  are 1-times differentiable so that by (1)  $g \circ f$  is 1-times differentiable. Further using 14.44 and 14.45 we have that  $D(g \circ f)$  is continuous so that

$$g \circ f \text{ is of class } C^1$$

□

**Theorem 14.83. (General Chain Rule)** Let  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle, \langle Z, \|\cdot\|_Z \rangle$  be normed spaces,  $R \subseteq Y, S \subseteq Z$ , open set  $U \subseteq X$ , open set  $V \subseteq Y$  and  $f: U \rightarrow R, g: V \rightarrow S$  functions then we have

1.  $\forall n \in \mathbb{N}_0$  if  $f, g$  are  $n$ -times differentiable then  $g \circ f: U \cap f^{-1}(V) \rightarrow Z$  is  $n$ -times differentiable
2.  $\forall n, m \in \mathbb{N}_0$  if  $f$  is  $n$ -times differentiable and  $g$  is  $m$ -times differentiable then  $g \circ f: U \cap f^{-1}(V) \rightarrow Z$  is  $\min(n, m)$ -times differentiable.
3.  $\forall n \in \mathbb{N}$  if  $f, g$  are of class  $C^n$  then  $g \circ f: U \cap f^{-1}(V) \rightarrow Z$  is of class  $C^n$ .
4.  $\forall n, m \in \mathbb{N}$  if  $f$  is of class  $C^n$  and  $g$  is of class  $C^m$  then  $g \circ f: U \cap f^{-1}(V) \rightarrow Z$  is of class  $C^{\min(n, m)}$
5. If  $f, g$  are of class  $C^\infty$  then  $g \circ f$  is of class  $C^\infty$

**Proof.**

1. We prove this by induction so let  $\mathcal{S} = \{n \in \mathbb{N}_0 \mid \text{if } \langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle, \langle Z, \|\cdot\|_Z \rangle \text{ are normed spaces, } R \subseteq Y, S \subseteq Y, \text{ open } U \subseteq X, \text{ open } V \subseteq Y \text{ and } f: U \rightarrow R, g: V \rightarrow S \text{ are } n\text{-times differentiable then } g \circ f: U \cap f^{-1}(V) \rightarrow S \text{ is } n\text{-times differentiable}\}$  then we have

**1  $\in \mathcal{S}$ .** This follows from the previous lemma [see 14.82] now

**$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$ .** Assume that  $f, g$  are  $(n+1)$ -times differentiable. Then using 14.66 it follows that  $D^1g, D^1f$  are  $n$ -times differentiable and by 14.62  $f$  is  $n$ -times differentiable. Take now  $h$  as in 14.82 then  $\pi_1 \circ h = (D^1g) \circ f$  is  $n$ -times differentiable [as  $n \in \mathcal{S}$ ] and  $\pi_2 \circ h = D^1f$  is  $n$ -times differentiable. Using 14.80 it follows that  $h$  is  $n$ -times differentiable. As  $(\circ)$  is of class  $C^\infty$  hence  $(\circ)$   $n$ -times differentiable we have as  $n \in \mathcal{S}$  that  $(\circ) \circ h$  is  $n$ -times differentiable. So that  $D^1(g \circ f) \stackrel{14.82}{=} (\circ) \circ h$  is  $n$ -times differentiable. Applying then 14.66 proves that  $g \circ f$  is  $n$ -times differentiable proving that  $n+1 \in \mathcal{S}$ .

2. Using 14.62 we have that  $f, g$  are  $\min(n, m)$ -times differentiable. So using (1) it follows that  $g \circ f$  is  $\min(n, m)$ -times differentiable.

3. We prove this by induction so let  $\mathcal{S} = \{n \in \mathbb{N} \mid \text{if } \langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle, \langle Z, \|\cdot\|_Z \rangle \text{ are normed spaces, } R \subseteq Y, S \subseteq Y, \text{ open } U \subseteq X, \text{ open } V \subseteq Y \text{ and } f: U \rightarrow R, g: V \rightarrow S \text{ are of class } C^n \text{ then } g \circ f \text{ is of class } C^n\}$  then we have

**0  $\in \mathcal{S}$ .** If  $f, g$  are of class  $C^0$  then  $f, g$  are continuous so that  $g \circ f$  is continuous [see 11.147]

**1  $\in \mathcal{S}$ .** This follows from the previous lemma [see 14.82]

**$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$ .** Assume that  $f$  and  $g$  are of class  $C^{n+1}$ . Then using 14.67 we have that  $D^1f, D^1g$  are of class  $C^n$  and by 14.63 that  $f$  is of class  $C^n$ . Take now  $h$  as in 14.82 then  $\pi_1 \circ h = (D^1g) \circ f$  is of class  $C^n$  [as  $n \in \mathcal{S}$ ] and  $\pi_2 \circ h = D^1f$  is of class  $C^n$ . Using 14.81 it follows that  $h$  is of class  $C^n$ . As  $(\circ)$  is of class  $C^\infty$  it follows that  $(\circ)$  is of class  $C^n$ . As  $n \in \mathcal{S}$  we have that  $(\circ) \circ h$  is of class  $C^n$ , hence  $D^1(g \circ f) \stackrel{14.82}{=} (\circ) \circ h$  is of class  $C^n$ . Applying then 14.67 proves that  $g \circ f$  is of class  $C^{n+1}$  proving that  $n+1 \in \mathcal{S}$ .

4. Using 14.63 we have that  $f, g$  are of class  $C^{\min(n, m)}$  hence using (3) we have that  $g \circ f$  is of class  $C^{\min(n, m)}$ .

5. If  $f, g$  are of class  $C^\infty$  then  $\forall n \in \mathbb{N} f, g$  are of class  $C^n$  hence by (3)  $g \circ f$  is of class  $C^n$  which proves that  $g \circ f$  is of class  $C^\infty$ .  $\square$

**Lemma 14.84.** Let  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle$  be normed spaces then the evaluation function  $\text{ev}: L(X, Y) \times X \rightarrow Y$  defined by  $(L, v) \rightarrow \text{ev}(L, v) = L(v)$  is bi-linear and continuous hence of class  $C^\infty$  [see 14.78] and  $D^1\text{ev}(L, v) = \text{ev}(L, *) + \text{ev}(*, v)$  (by 14.78).

**Proof.**

1. **(bi-linearity)**
  - a. If  $\alpha, \beta \in \mathbb{K}, v \in X$  and  $L_1, L_2 \in L(X, Y)$  then  $\text{ev}(\alpha \cdot L_1 + \beta \cdot L_2, v) = (\alpha \cdot L_1 + \beta \cdot L_2)(v) = \alpha \cdot L_1(v) + \beta \cdot L_2(v) = \alpha \cdot \text{ev}(L_1, v) + \beta \cdot \text{ev}(L_2)$
  - b. If  $\alpha, \beta \in \mathbb{K}, v_1, v_2 \in X$  and  $L \in L(X, Y)$  then  $\text{ev}(L, \alpha \cdot v_1 + \beta \cdot v_2) = L(\alpha \cdot v_1 + \beta \cdot v_2) = \alpha \cdot L(v_1) + \beta \cdot L(v_2) = \alpha \cdot \text{ev}(L, v_1) + \beta \cdot \text{ev}(L, v_2)$
2. **(continuity)**  $\|\text{ev}(L, v)\|_Y = \|L(v)\|_Y \leq \|L\|_{L(X, Y)} \cdot \|v\|_X \leq 1 \cdot (\|L\|_{L(X, Y)} \cdot \|v\|_X)$  proving that  $\text{ev}$  is multi linear with  $\|\text{ev}\| \leq 1$  [see 11.194]  $\square$

**Lemma 14.85.** Let  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle$  be normed spaces,  $U \subseteq X$  open,  $V \subseteq Y$ ,  $n \in \mathbb{N}_0, n > 1$  and  $f: U \rightarrow V$  a  $n$ -times differentiable function [or  $f$  is of class  $C^n$ ] then given  $h \in X$  we have that

$$D^1f(*)(h): U \rightarrow Y \text{ defined by } (D^1f(*)(h))(x) = D^1f(x)(h)$$

is  $(n-1)$ -times differentiable [or is of class  $C^{n-1}$ ]. Further  $\forall k, h, x \in X$  we have

$$D^1(D^1f(*)(h))(x)(k) = D^2f(x)(k)(h).$$

**Proof.** Given  $v \in X$  define  $\varphi_v: U \rightarrow L(X, Y) \times X$  by  $\varphi_v(x) = (D^1f(x), v)$ . Then we have that  $\pi_1 \circ \varphi_v = D^1f$  which is  $(n-1)$ -times differentiable [see 14.66] [or  $C^{n-1}$  by 14.67] and  $\pi_2 \circ \varphi_v = C_{U, X, v}$  which is  $C^\infty$  [see 14.74] so that  $\pi_2 \circ \varphi_v$  is  $(n-1)$ -times differentiable [or  $C^{n-1}$ ] Applying 14.80 [or 14.81] proves that

$$\varphi_v \text{ is } (n-1)\text{-times differentiable [or } \varphi_v \text{ is } C^{n-1} \text{]} \quad (14.47)$$

As  $n > 1$  we have that  $n-1 \geq 1$  so that by 14.62  $\varphi_v$  is 1-times differentiable hence using 14.80 we have that  $\pi_1 \circ D\varphi_v = D(\pi_1 \circ \varphi_v) = D(D^1f) = D^2f$  and  $\pi_2 \circ D\varphi_v = D(\pi_2 \circ \varphi_v) = D(C_{U, X, v}) \stackrel{14.74}{=} C_{U, L(X, X), 0}$  which proves that

$$D^1\varphi_v: U \rightarrow L(X, L(X, Y) \times X) \text{ is defined by } D^1\varphi_v(x) = (D^2f(x), C_{X, X, 0}) \quad (14.48)$$

Take now  $h, x \in X$  then we have

$$\begin{aligned} D^1f(*)(h)(x) &= D^1f(x)(h) \\ &= \text{ev}(D^1f(x), h) \\ &= \text{ev}(\varphi_h(x)) \\ &= (\text{ev} \circ \varphi_h)(x) \end{aligned}$$

proving that

$$D^1f(*)(h) = \text{ev} \circ \varphi_h \quad (14.49)$$

As  $\text{ev}$  is of class  $C^\infty$  [see 14.84] and  $\varphi_v$  is  $(n-1)$ -times differentiable [see 14.47] we can use the above together with the chain rule [see 14.83] to prove that

$$D^1f(*)(h) \text{ is } (n-1)\text{-times differentiable [or } C^{n-1} \text{]} \quad (14.50)$$

and

$$\begin{aligned} D^1(D^1f(*)(h))(x) &= D(D^1f(*)(h))(x) \\ &\stackrel{14.49}{=} D((\text{ev} \circ \varphi_h)(x)) \\ &\stackrel{14.22}{=} D\text{ev}(\varphi_h(x)) \circ D\varphi_h(x) \\ &= D\text{ev}((D^1f(x), h)) \circ D\varphi_h(x) \\ &\stackrel{14.84}{=} (\text{ev}(D^1f(x), *) + \text{ev}(*, h)) \circ D\varphi_h(x) \\ &\stackrel{14.48}{=} (\text{ev}(D^1f(x), *) + \text{ev}(*, h)) \circ (D^2f(x), C_{X, X, 0}) \end{aligned} \quad (14.51)$$

Let  $k \in X$  then

$$\begin{aligned} ((\text{ev}(D^1f(x), *) + \text{ev}(*, h)) \circ (D^2f(x), \\ C_{X, X, 0}))(k) &= (\text{ev}(D^1f(x), *) + \text{ev}(*, h))(D^2f(x), C_{X, X, 0})(k) \\ &= (\text{ev}(D^1f(x), *) + \text{ev}(*, h))(D^2f(x)(k), \\ C_{X, X, 0}(k)) \\ &= (\text{ev}(D^1f(x), *) + \text{ev}(*, h))(D^2f(x)(k), 0) \\ &= (\text{ev}(D^1f(x), 0) + \text{ev}(D^2f(x)(k), h)) \\ &= D^1f(x)(0) + D^2f(x)(k)(h) \\ &= 0 + D^2f(x)(k)(h) \\ &= D^2f(x)(k)(h) \end{aligned}$$

which using 14.51 proves that

$$D^1(D^1f(*)(h))(x)(k) = D^2f(x)(k)(h) \quad (14.52)$$

The lemma is then proved by 14.50 and 14.52 □

**Lemma 14.86.** Let  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle$  be normed spaces,  $n \in \mathbb{N}_0$ ,  $f: U \rightarrow Y$  a  $n$ -times differentiable function [or  $C^n$  function] then  $\Delta f: X \times X \rightarrow Y$  defined by  $\Delta f(x, h) = D^1 f(x)(h)$  is  $(n-1)$ -times differentiable [or  $C^{n-1}$ ]

**Proof.** Define  $\varphi: U \times X \rightarrow L(X, Y) \times X$  by  $\varphi(x, v) = (D^1 f(x), v)$  then

$$\begin{aligned} (\pi_1 \circ \varphi)(x, v) &= \pi_1(\varphi(x, v)) \\ &= D^1 f(x) \\ &= D^1 f(\pi_1(x, v)) \\ &= ((D^1 f) \circ \pi_1)(x, v) \\ (\pi_2 \circ \varphi)(x, v) &= v \\ &= \pi_2(x, v) \end{aligned}$$

proving that

$$\pi_1 \circ \varphi = D^1 f \circ \pi_1 \text{ and } \pi_2 \circ \varphi = \pi_2$$

As  $D^1 f$  is  $(n-1)$ -times differentiable by 14.66 [or  $C^{n-1}$  by 14.67] and  $\pi_1$  is  $C^\infty$  [see 14.79] we have by the chain rule [see 14.83] that  $\pi_1 \circ \varphi$  is  $(n-1)$ -times differentiable [or  $C^{n-1}$ ]. This together with the fact that  $\pi_2$  is  $C^\infty$  [see 14.79] proves that

$$\varphi \text{ is } (n-1)\text{-times differentiable [or } C^{n-1} \text{]} \quad (14.53)$$

Now

$$\begin{aligned} (\text{ev} \circ \varphi)(x, v) &= \text{ev}(\varphi(x, v)) \\ &= \text{ev}(D^1 f(x), v) \\ &= D^1 f(x)(v) \\ &= \Delta f(x, v) \end{aligned}$$

Proving that  $\text{ev} \circ \varphi = \Delta f$ . As  $\varphi$  is  $(n-1)$ -times differentiable [or  $C^{n-1}$ ] and  $\text{ev}$  is  $C^\infty$  [see 14.84] we have by the generalized chain rule [see 14.83] that  $\Delta f$  is  $(n-1)$ -times differentiable [or of class  $C^{n-1}$ ].  $\square$

**Theorem 14.87.** Let  $\langle X, \|\cdot\|_X \rangle$  be a normed space,  $\langle \mathbb{K}, \|\cdot\| \rangle$  be  $\mathbb{K}$  equipped with the canonical norm,  $U \subseteq X$  and  $n \in \mathbb{N}_0$  [or  $n \in \mathbb{N}$ ] and  $f, g: U \rightarrow \mathbb{K}$  are functions  $n$ -times differentiable [or of class  $C^n$ ] then  $f \cdot g: U \rightarrow \mathbb{K}$  is  $n$ -times differentiable [or of class  $C^r$ ]

**Proof.** First if  $n = 0$  then as  $f, g$  are  $C^0$  and thus continuous we have by 11.177 that  $f \cdot g$  is continuous and thus  $C^0$ .

Second given  $n > 0$  define  $\varphi: U \rightarrow \mathbb{K} \times \mathbb{K}$  by  $\varphi(x) = (f(x), g(x))$  then as  $\pi_1 \circ \varphi = f$  and  $\pi_2 \circ \varphi = g$  we have that  $\varphi$  is  $n$ -times differentiable [or  $C^n$ ] (see 14.80 [or 14.81]). As  $(\cdot): \mathbb{K} \times \mathbb{K} \rightarrow \mathbb{K}$  defined by  $(x, y) \rightarrow x \cdot y$  is trivially bi-linear and continuous [see 11.195] we have by 14.78 that  $(\cdot)$  is  $C^\infty$ . Now as  $((\cdot) \circ \varphi)(x) = (\cdot)(\varphi(x)) = (\cdot)(f(x), g(x)) = f(x) \cdot g(x)$  proving that  $(\cdot) \circ \varphi = f \cdot g$ . By the generalized chain rule it follows that  $f \cdot g$  is  $n$ -times differentiable [or  $C^n$ ].  $\square$

### 14.3 Differentiability on general sets

**Definition 14.88.** Let  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle$  be normed spaces  $S \subseteq X, V \subseteq Y$  and  $f: S \rightarrow V$  a function then  $f$  differentiable on  $S$  if there exists a  $U$  open such that  $S \subseteq U \subseteq X$  and a  $f^U: U \rightarrow Y$  with  $(f^U)|_S = f$  that is differentiable on  $U$ .

Also  $f$  is  $C^r$  if there exists a  $U$  open such that  $S \subseteq U \subseteq X$  and a  $f^U: U \rightarrow Y$  with  $(f^U)|_S = f$  such that  $f^U$  is  $C^r$ .

In other words  $f$  is differentiable [or  $C^r$ ] if it can be extended to a differentiable [or  $C^r$ ] function on a open set containing  $S$ .

**Theorem 14.89.** Let  $\langle X, \|\cdot\|_X \rangle$ ,  $\langle Y, \|\cdot\|_Y \rangle$ ,  $\langle Z, \|\cdot\|_Z \rangle$  be normed spaces  $S \subseteq X$ ,  $T \subseteq Y$ ,  $r \in \mathbb{N}$ ,  $A \subseteq Y$ ,  $B \subseteq Z$  and  $f: S \rightarrow A$ ,  $g: T \rightarrow B$  functions of class  $C^r$  then  $g \circ f: S \cap f^{-1}(T)$  is  $C^r$  on  $S$ .

**Proof.** By definition there exists a  $U$  open in  $X$  and a  $V$  open in  $Y$  such that  $S \subseteq U$ ,  $T \subseteq V$  and  $C^r$  functions  $f^U: U \rightarrow X$ ,  $g^V: V \rightarrow Z$  such that  $(f^U)|_S = f$  and  $(g^V)|_T = g$ . Take then  $W = U \cap (f^U)^{-1}(V)$  which is open because  $f^U$  is continuous [being of class  $C^r$ ]. As  $f^U(S \cap f^{-1}(T)) \subseteq f^U(f^{-1}(T)) \underset{f^{-1}(T) \subseteq S}{=} f(f^{-1}(T)) \subseteq S \subseteq V \Rightarrow S \subseteq (f^U)^{-1}(V) \underset{S \subseteq U}{\Rightarrow} S \subseteq U \cap (f^U)^{-1}(V) = W$ . Define now  $f^W = (f^U)|_W: W \rightarrow Y$  then  $f^W$  is  $C^r$  [see 14.70] and as  $S \subseteq W \subseteq U$   $f^W$  is still a extension of  $f$ . Now  $f^W(W) \underset{W \subseteq U}{=} f^U(W) \subseteq f^U((f^U)^{-1}(V)) \subseteq V$  so that by using the chain rule (14.83) we have that  $f^W \circ g^V$  is  $C^r$  and as  $(g^V \circ f^U)|_S = g \circ f$  and  $S \subseteq W$  we have that  $g \circ f$  is  $C^r$ .  $\square$

**Definition 14.90.** Let  $\langle \mathbb{R}, \|\cdot\| \rangle$  be the real space with the canonical norm,  $\langle Y, \|\cdot\|_Y \rangle$  a normed space,  $[a, b] \subseteq \mathbb{R}$ ,  $a < b$  and  $f: [a, b] \rightarrow Y$  then we define the right or left derivate by

1. If  $x \in [a, b[$  then  $f'_+(x)$  is the right derivate of  $f$  at  $x$  if and only if  $\forall \varepsilon > 0$  there exists a  $\delta > 0$  such that for all  $h$  with  $0 < h < \delta$  and  $x + h \in [a, b]$  we have  $\left\| \frac{f(x+h) - f(x)}{h} - f'_+(x) \right\|_Y < \varepsilon$
2. If  $x \in ]a, b]$  then  $f'_-(x)$  is the left derivate of  $f$  at  $x$  if and only if  $\forall \varepsilon > 0$  there exists a  $\delta > 0$  such that for all  $h$  with  $0 < h < \delta$  and  $x - h \in [a, b]$  we have  $\left\| \frac{f(x-h) - f(x)}{-h} - f'_-(x) \right\|_Y = \left\| \frac{f(x) - f(x-h)}{h} - f'_-(x) \right\|_Y < \varepsilon$

**Lemma 14.91.** Let  $\langle \mathbb{R}, \|\cdot\| \rangle$  be the real space with the canonical norm,  $[a, b] \subseteq \mathbb{R}$ ,  $a < b$  and  $f: [a, b] \rightarrow Y$  then

1. If  $x \in [a, b[$  then  $f'_+(x)$  if it exists is unique
2. If  $x \in ]a, b]$  then  $f'_-(x)$  if it exists is unique
3. If  $x \in ]a, b[$  and  $f'_-(x)$ ,  $f'_+(x)$  both exists and  $f'_-(x) = f'_+(x)$  then  $(f|_{]a, b[})'(x)$  exists and  $(f|_{]a, b[})'(x) = f'_-(x) = f'_+(x)$ . Also if  $(f|_{]a, b[})'(x)$  exists then  $f'_-(x)$ ,  $f'_+(x)$  exists and  $f_+(x) = f_-(x)$ .
4. If  $[a, b] \subseteq U$  open and  $f^U: U \rightarrow Y$  is a function with  $(f^U)|_{[a, b]} = f$  for which  $(f^U)'(x)$  exists  $\forall x \in U$  then  $f'(x)$  exists on  $]a, b[$  with  $(f^U)'(x) = f'(x)$ ,  $f'_+(a)$ ,  $f'_-(b)$  exists with  $f'_+(a) = (f^U)'(a)$  and  $(f^U)'(b) = f'_-(b)$

**Proof.**

1. Assume that there exists two different right derivate  $d, d'$  at  $x \in [a, b[$  define then  $\varepsilon = \|d' - d\|_Y > 0$ , then there exists  $\delta, \delta' > 0$  such that if  $0 < h < \delta$  and  $x + h \in [a, b]$  that  $\left\| \frac{f(x+h) - f(x)}{h} - d \right\|_Y < \frac{\varepsilon}{2}$  and if  $0 < h < \delta'$  and  $x + h \in [a, b]$  then  $\left\| \frac{f(x+h) - f(x)}{h} - d' \right\|_Y < \frac{\varepsilon}{2}$ . As  $x \in [a, b[$  we have  $a \leq x < b$  so there exists a  $\delta'' > 0$  such that  $a < x + \delta'' < b$ , take  $k = \min(\delta'', \delta', \delta)$  then  $0 < k < \delta, \delta'$  and  $a < x + k \leq x + \delta'' < b \Rightarrow x + k \in [a, b]$ . Hence

$$\begin{aligned} \|d' - d\|_Y &= \left\| \frac{f(x+k) - f(x)}{k} - d + d' - \frac{f(x+k) - f(x)}{k} \right\|_Y \\ &\leq \left\| \frac{f(x+k) - f(x)}{k} - d \right\|_Y + \left\| \frac{f(x+k) - f(x)}{k} - d' \right\|_Y \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\ &= \varepsilon \\ &= \|d' - d\|_Y \end{aligned}$$

a contradiction. So we must have that  $d = d'$  an the right derivate is unique.

2. Assume that there exists two different left derivates  $d, d'$  at  $x \in ]a, b]$  define then  $\varepsilon = \|d' - d\|_Y > 0$ , then there exists  $\delta, \delta' > 0$  such that if  $0 < h < \delta$  and  $x - h \in [a, b]$  that  $\left\| \frac{f(x-h) - f(x)}{-h} - d \right\|_Y < \frac{\varepsilon}{2}$  and if  $0 < h < \delta'$  and  $x - h \in [a, b]$  then  $\left\| \frac{f(x-h) - f(x)}{-h} - d' \right\|_Y < \frac{\varepsilon}{2}$ . As  $x \in ]a, b]$  we have  $a < x \leq b$  so there exists a  $\delta'' > 0$  such that  $a < x - \delta'' < x \leq b$ , take  $k = \min(\delta'', \delta', \delta)$  then  $0 < k < \delta, \delta'$  and  $a < x - \delta'' \leq x - k \leq b \Rightarrow x - k \in [a, b]$ . Hence

$$\begin{aligned} \|d' - d\|_Y &= \left\| \frac{f(x-h) - f(x)}{-h} - d + d' - \frac{f(x-h) - f(x)}{-h} - d \right\|_Y \\ &\leq \left\| \frac{f(x-h) - f(x)}{-h} - d \right\|_Y + \left\| \frac{f(x-h) - f(x)}{-h} - d' \right\|_Y \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\ &= \varepsilon \\ &= \|d - d'\|_Y \end{aligned}$$

a contradiction. So we must have that  $d = d'$  an the right derivate is unique.

3. Let  $d = f'_+(x) = f'_-(x)$  and let  $\varepsilon > 0$  then there exists  $\delta', \delta'' > 0$  such that if  $0 < h < \delta'$  and  $x + h \in [a, b]$  then  $\left\| \frac{f(x+h) - f(x)}{h} - d \right\|_Y < \varepsilon$  and if  $0 < h < \delta''$  and  $x - h \in [a, b]$  then  $\left\| \frac{f(x-h) - f(x)}{-h} - d \right\|_Y < \varepsilon$ . Take then  $0 < |h| < \min(\delta', \delta'')$  such that  $x + h \in ]a, b[$  then two cases exists:

**$h > 0$ .** then as  $0 < |h| < \min(\delta', \delta'') \Rightarrow 0 < h < \delta', x + h \in ]a, b[$  and thus  $\left\| \frac{f_{[a,b]}(x+h) - f_{[a,b]}(x)}{h} - d \right\|_Y = \left\| \frac{f(x+h) - f(x)}{h} - d \right\|_Y < \varepsilon$

**$h < 0$ .** then  $h = -|h|$  hence  $x - |h| \in ]a, b[$  and  $h = -|h|$  thus  $\left\| \frac{f_{[a,b]}(x+h) - f_{[a,b]}(x)}{h} - d \right\|_Y = \left\| \frac{f(x+h) - f(x)}{h} - d \right\|_{Y, h=-|h|} = \left\| \frac{f(x-|h|) - f(x)}{-|h|} - d \right\|_Y < \varepsilon$

So we have in all cases that  $\left\| \frac{f_{[a,b]}(x+h) - f_{[a,b]}(x)}{h} - d \right\|_Y < \varepsilon$  proving that  $(f_{[a,b]})'(x) = d = f'_-(x) = f'_+(x)$ .

On the other hand if  $(f_{[a,b]})'(x)$  exists then given  $\varepsilon > 0$  there exists a  $\delta' > 0$  such that if  $0 < |h| < \delta'$  and  $x + h \in ]a, b[$  then we have  $\left\| \frac{f(x+h) - f(x)}{h} - (f_{[a,b]})'(x) \right\|_Y = \left\| \frac{f_{[a,b]}(x+h) - f_{[a,b]}(x)}{h} - d \right\|_Y < \varepsilon$ . Also as  $x \in ]a, b[$  we have  $0 < x - a, b - x$  and there exists a  $\delta''$  such that  $0 < \delta'' < \min(x - a, b - x)$  so if  $0 < h < \delta''$  then  $x - h > x - \delta'' > x - (x - a) = a$  and  $x + h < x + \delta'' < x + b - x = b$  giving  $x - h, x + h \in ]a, b[$  so if  $0 < h < \delta = \min(\delta', \delta'')$  then we have

a.  $x + h \in [a, b]$  and  $\left\| \frac{f(x+h) - f(x)}{h} - (f_{[a,b]})'(x) \right\| < \varepsilon$  proving that  $f'_+(x)$  exists and is equal to  $f'(x)$ .

b.  $x - h \in [a, b]$  and as  $0 < -h = |h| < \delta$  we have  $\left\| \frac{f(x-h) - f(x)}{-h} - (f_{[a,b]})'(x) \right\| < \varepsilon$  proving that  $f'_-(x)$  exists and is equal to  $f'(x)$ .

4. We must consider the following cases:

**$x \in ]a, b[$ .** Let  $\varepsilon > 0$  then as  $(f^U)'(x)$  exists there exists a  $\delta > 0$  such that if  $0 < |h| < \delta$  with  $h \in U_x$  then  $\left\| \frac{f^U(x+h) - f^U(x)}{h} - (f^U)'(x) \right\|_Y < \varepsilon$ . So if  $0 < |h| < \delta$  with  $x \in (]a, b[)_x \Rightarrow x + h \in ]a, b[ \subseteq U \Rightarrow h \in U_x$  then  $\left\| \frac{f(x+h) - f(x)}{h} - (f^U)'(x) \right\|_Y = \left\| \frac{f^U(x+h) - f^U(x)}{h} - (f^U)'(x) \right\|_Y < \varepsilon$  proving that  $f'(x)$  exists and is equal to  $(f^U)'(x)$

**$x = a$ .** Then as  $(f^U)'$  exists on  $U$  and  $[a, b] \subseteq U$  given  $\varepsilon > 0$  there exists a  $\delta > 0$  such that if  $0 < |h| < \delta$  and  $h \in U_a$  then  $\left\| \frac{f^U(a+h) - f^U(a)}{h} - (f^U)'(a) \right\|_Y < \varepsilon$ . So if  $0 < h < \delta$  and  $a + h \in [a, b] \subseteq U \Rightarrow h \in U_a$  then we have  $\left\| \frac{f(a+h) - f(a)}{h} - (f^U)'(a) \right\|_Y = \left\| \frac{f^U(a+h) - f^U(a)}{h} - (f^U)'(a) \right\|_Y < \varepsilon$  proving that  $f'_+(a)$  exists and  $f'_+(a) = (f^U)'(a)$ .

**$x = b$ .** Then as  $(f^U)'$  exists on  $U$  and  $[a, b] \subseteq U$  given  $\varepsilon > 0$  there exists a  $\delta > 0$  such that if  $0 < |h| < \delta$  and  $h \in U_a$  then  $\left\| \frac{f^U(a+h) - f^U(a)}{h} - (f^U)'(a) \right\|_Y < \varepsilon$ . So if  $0 < -h < \delta \Rightarrow 0 < |-h| < \delta$  with  $b + (-h) \in [a, b] \subseteq U \Rightarrow -h \in U_b$  then we have  $\left\| \frac{f^U(b-h) - f^U(b)}{-h} - (f^U)'(b) \right\|_Y = \left\| \frac{f^U(b+(-h)) - f^U(b)}{-h} - (f^U)'(b) \right\|_Y < \varepsilon$  proving that  $f'_-(b)$  exists and  $(f^U)'(b) = f'_-(b)$ .  $\square$

The above theorem leads to the following definition

**Definition 14.92.** Let  $\langle \mathbb{R}, \|\cdot\| \rangle$  be the real space with the canonical norm,  $\langle Y, \|\cdot\|_Y \rangle$  a normed space,  $V \subseteq Y$ ,  $[a, b] \subseteq \mathbb{R}$ ,  $a < b$ ,  $f: [a, b] \rightarrow V$  and  $x \in [a, b]$  then  $f$  has a derivate  $f'$  at  $x$  on  $[a, b]$  if and only if either  $x = a$  and  $f'_+(a)$  exists,  $x = b$  and  $f'_-(b)$  exists or  $x \in ]a, b[$  and  $f'_+(x)$ ,  $f'_-(x)$  exists with  $f'_+(x) = f'_-(x)$ . Further if  $f'$  exists at  $x$  then  $f'(x) = \begin{cases} f'_+(a) & \text{if } x = a \\ f'_-(b) & \text{if } x = b \\ (f|_{]a, b[})' & \text{if } x \neq a, b \end{cases}$

**Theorem 14.93.** Let  $\langle \mathbb{R}, \|\cdot\| \rangle$  be the real space with the canonical norm,  $\langle Y, \|\cdot\|_Y \rangle$  a normed space,  $V \subseteq Y$ ,  $[a, b] \subseteq \mathbb{R}$ ,  $a < b$  then the following are equivalent for  $f: [a, b] \rightarrow V$

1.  $f$  is  $C^1$  on  $[a, b]$
2. We have the following
  - a.  $\forall t \in ]a, b[$  we have the existence of the derivate  $(f|_{]a, b[})'(t)$
  - b. There exists a right, left derivate  $f'_+(a)$ ,  $f'_-(b)$  at  $a, b$
  - c. Let  $f': [a, b] \rightarrow Y$  defined by  $x \rightarrow f'(x) = \begin{cases} f'_+(a) & \text{if } x = a \\ (f|_{]a, b[})'(x) & \text{if } x \neq a, b \\ f'_-(b) & \text{if } x = b \end{cases}$  then  $f'$  is continuous on  $[a, b]$

Further if  $f$  is  $C^1$  then  $\forall U$  open,  $f^U: U \rightarrow Y$  with  $[a, b] \subseteq U$  and  $(f^U)|_{[a, b]} = f$  we have that  $((f^U)')|_{[a, b]} = f'$

**Proof.** The proof is simple but elaborated.

**1  $\Rightarrow$  2.** Assume that  $f$  is  $C^1$  then there exists a  $U \subseteq \mathbb{R}$  be a open set with  $[a, b] \subseteq U$ , a function  $f^U: U \rightarrow \mathbb{R}$  with  $(f^U)|_{[a, b]} = f$  such that  $f^U$  is  $C^1$ . Then using 14.58 we have that

$$\forall t \in U (f^U)'(t) \text{ exists and } (f^U)': U \rightarrow \mathbb{R} \text{ is continuous} \quad (14.54)$$

We proceed now to prove 2.a, 2.b and 2.c:

- a) As  $]a, b[$  is a open set,  $]a, b[ \subseteq U$  and  $(f^U)|_{]a, b[} = f|_{]a, b[}$  we have by 14.54 and 14.91 (4) that

$$\forall x \in ]a, b[ (f|_{]a, b[})'(x) \text{ and } (f|_{]a, b[})'(x) = (f^U)'(x). \quad (14.55)$$

- b) As  $[a, b] \subseteq U$  and  $(f^U)|_{[a, b]} = f$  we have by 14.54 and 14.91 (4) that

$$f'_+(a) \text{ exists and that } f'_+(a) = (f^U)'(a) \quad (14.56)$$

and

$$f'_-(b) \text{ exists and that } f'_-(b) = (f^U)'(b) \quad (14.57)$$

- c) Let  $x \in [a, b]$  then we have either

$$\mathbf{x = a.} \text{ then } f'(x) = f'_+(a) \stackrel{(14.56)}{=} (f^U)'(a) = (f^U)'(x)$$

$$\mathbf{x = b.} \text{ then } f'(x) = f'_-(b) \stackrel{(14.57)}{=} (f^U)'(a) = (f^U)'(x)$$

$$\mathbf{x \in ]a, b[.} \text{ then } f'(x) = (f|_{]a, b[})'(x) \stackrel{(14.55)}{=} (f^U)'(x)$$

which proves that

$$f' = ((f^U)')|_{[a, b]} \quad (14.58)$$

Using 11.137 and the continuity of  $(f^U)'$  [see (14.54)] it follow that

$$f' \text{ is continuous in the subspace topology on } [a, b] \quad (14.59)$$

**2⇒1.** Let  $f: [a, b] \rightarrow Y$  be a function satisfying 2.a, 2.b and 2.c then  $[a, b] \subseteq ]a-1, b+1[$  a open set. Define then  $f^{[a-1, b+1]}: ]a-1, b+1[ \rightarrow Y$  by

$$f^{[a-1, b+1]}(x) = \begin{cases} f(x) & \text{if } x \in [a, b] \\ f(a) + f'_+(a) \cdot (x-a) & \text{if } x \in ]a-1, a[ \\ f(b) + f'_-(b) \cdot (x-b) & \text{if } x \in ]b, b+1[ \end{cases}$$

then we have trivial

$$(f^{[a-1, b+1]})_{|[a, b]} = f \quad (14.60)$$

Next we prove that  $f^{[a-1, b+1]}$  has a derivative for every element in  $]a-1, b+1[$ . Let  $x \in ]a-1, b+1[$  then we must consider the following case:

**x = a.** Let  $\varepsilon > 0$  then by the definition of  $f'_+(a)$  there exists a  $\delta' > 0$  such that if  $0 < h < \delta'$  with  $x+h \in [a, b]$  we have  $\left\| \frac{f(a+h) - f(x)}{h} - f'_+(a) \right\|_Y < \varepsilon$ . Take  $\delta = \min(\delta', b-a) > 0$ . If  $0 < |h| < \delta$  with  $h \in ]a-1, b+1[ \Rightarrow a+h \in ]a-1, b+1[$  we have either

**h < 0.** then  $a+h \in ]a-1, a[$  so that

$$\begin{aligned} & \left\| \frac{f^{[a-1, b+1]}(a+h) - f^{[a-1, b+1]}(a)}{h} - f'_+(a) \right\|_Y = \\ & \left\| \frac{f(a) + f'_+(a) \cdot (a+h-a) - f(a)}{h} - f'_+(a) \right\|_Y = \\ & \left\| \frac{f'_+(a) \cdot h}{h} - f'_+(a) \right\|_Y = 0 < \varepsilon \end{aligned}$$

**0 < h.** then  $0 < h < \delta \leq b-a \Rightarrow a < a+h < (b-a) + a$  proving that  $a+h \in ]a, b[ \subseteq [a, b]$  so that

$$\begin{aligned} & \left\| \frac{f^{[a-1, b+1]}(a+h) - f^{[a-1, b+1]}(a)}{h} - f'_+(a) \right\|_Y = \\ & \left\| \frac{f(a+h) - f(a)}{h} - f'_+(a) \right\|_Y < \varepsilon \text{ [as } h < \delta'] \end{aligned}$$

so in all cases we have that  $\left\| \frac{f^{[a-1, b+1]}(a+h) - f^{[a-1, b+1]}(a)}{h} - f'_+(a) \right\|_Y < \varepsilon$  which proves that  $(f^{[a-1, b+1]})'(x)$  exists and is equal to  $f'_+(a)$ .

**x = b.** Let  $\varepsilon > 0$  then by the definition of  $f'_-(b)$  there exists a  $\delta' > 0$  such that if  $0 < h < \delta'$  with  $b-h \in [a, b]$  we have  $\left\| \frac{f(b-h) - f(b)}{-h} - f'_-(b) \right\|_Y < \varepsilon$ . Take then  $\delta = \min(\delta', b-a)$ . If  $0 < |h| < \delta$  with  $h \in ]a-1, b+1[ \Rightarrow b+h \in ]a-1, b+1[$  we have either

**h < 0.** then  $-h = |h| < \delta \leq b-a \Rightarrow a-b < h \Rightarrow a < b+h \Rightarrow b+h \in [a, b]$  hence

$$\begin{aligned} & \left\| \frac{f^{[a-1, b+1]}(b+h) - f^{[a-1, b+1]}(b)}{h} - f'_-(b) \right\|_Y = \\ & \left\| \frac{f(b+h) - f(b)}{h} - f'_-(b) \right\|_Y = \\ & \left\| \frac{f(b-|h|) - f(b)}{-|h|} - f'_-(b) \right\|_Y < \varepsilon \text{ [as } |h| < \delta'] \end{aligned}$$

**0 < h.** then as  $b+h \in ]b, b+1[$  we have

$$\begin{aligned} & \left\| \frac{f^{[a-1, b+1]}(b+h) - f^{[a-1, b+1]}(b)}{h} - f'_-(b) \right\|_Y = \\ & \left\| \frac{f(b) + f'_-(b) \cdot (b+h-b) - f(b)}{h} - f'_-(b) \right\|_Y = \\ & \left\| \frac{f'_-(b) \cdot h}{h} - f'_-(b) \right\|_Y = 0 < \varepsilon \end{aligned}$$

so in all cases we have that  $\left\| \frac{f^{[a-1,b+1]}(b+h) - f^{[a-1,b+1]}(b)}{h} - f'_-(b) \right\|_Y < \varepsilon$  which proves that  $(f^{[a-1,b+1]})'(x)$  exists and is equal to  $f'_-(b)$ .

$x \in ]a, b[$ . By 2.a.  $(f_{[a,b]})'(x)$  exists so given  $\varepsilon > 0$  there exists a  $\delta' > 0$  such that if  $0 < |h| < \delta'$  and  $h \in ]a, b[ \Rightarrow x+h \in ]a, b[$  we have  $\left\| \frac{f_{[a,b]}(x+h) - f_{[a,b]}(x)}{h} - (f_{[a,b]})'(x) \right\|_Y < \varepsilon$ . Take now  $\delta = \min(\delta', x-a, b-x) > 0$  then if  $0 < |h| < \delta$  we have  $h < b-x \Rightarrow x+h < b$  and  $-h < x-a \Rightarrow a < x+h$  proving that  $x+h \in ]a, b[$

$$\begin{aligned} & \left\| \frac{f^{[a-1,b+1]}(x+h) - f^{[a-1,b+1]}(x)}{h} - (f_{[a,b]})'(x) \right\|_Y \quad x+h, x \in \overline{a, b} \subseteq [a, b] \\ & \left\| \frac{f(x+h) - f(x)}{h} - (f_{[a,b]})'(x) \right\|_Y \quad x \in \overline{a, b} \\ & \left\| \frac{f_{[a,b]}(x+h) - f_{[a,b]}(x)}{h} - (f_{[a,b]})'(x) \right\|_Y < \varepsilon \end{aligned}$$

which proves that  $(f^{[a-1,b+1]})'(x)$  exists and is equal to  $(f_{[a,b]})'(x)$

$x \in ]a-1, a[$ . Let  $\varepsilon > 0$  take  $\delta = a-x > 0$  then if  $0 < |h| < \delta$  and  $h \in ]a-1, b+1[$  then  $x+h \leq x+|h| < x+\delta = a$  and  $x+h \in ]a-1, b+1[$  so that  $x+h \in ]a-1, a[$  so that

$$\begin{aligned} & \left\| \frac{f^{[a-1,b+1]}(x+h) - f^{[a-1,b+1]}(x)}{h} - f'_+(a) \right\|_Y = \\ & \left\| \frac{f(a) + f'_+(a) \cdot (x+h-a) - [f(a) + f'_+(a) \cdot (x-a)]}{h} - f'_+(a) \right\|_Y = \\ & \left\| \frac{f'_+(a) \cdot h}{h} - f'_+(a) \right\|_Y = 0 < \varepsilon \end{aligned}$$

which proves that  $(f^{[a-1,b+1]})'(x)$  exists and is equal to  $f'_+(a)$ .

$x \in ]b, b+1[$ . Let  $\varepsilon > 0$  take  $\delta = x-b > 0$  then if  $0 < |h| < \delta$  and  $h \in ]a-1, b+1[$  then  $-h \leq |h| < \delta = x-b \Rightarrow b < x+h$  and  $x+h \in ]a-1, b+1[$  so that  $x+h \in ]b, b+1[$ . So that

$$\begin{aligned} & \left\| \frac{f^{[a-1,b+1]}(x+h) - f^{[a-1,b+1]}(x)}{h} - f'_-(b) \right\|_Y = \\ & \left\| \frac{f(b) + f'_-(b) \cdot (x+h-b) - [f(b) + f'_-(b) \cdot (x-b)]}{h} - f'_-(b) \right\|_Y = \\ & \left\| \frac{f'_-(b) \cdot h}{h} - f'_-(b) \right\|_Y = 0 < \varepsilon \end{aligned}$$

which proves that  $(f^{[a-1,b+1]})'(x)$  exists and is equal to  $f'_-(b)$

Summarizing we have that  $\forall x \in ]a-1, b+1[$

$$(f^{[a-1,b+1]})'(x) \text{ exists and } (f^{[a-1,b+1]})'(x) = \begin{cases} f'_+(a) & \text{if } x \in ]a-1, a[ \\ f'_-(b) & \text{if } x \in ]b, b+1[ \\ (f_{[a,b]})'(x) & \text{if } x \in ]a, b[ \end{cases} \quad (14.61)$$

Now to prove continuity of  $(f^{[a-1,b+1]})'$  on  $]a-1, b+1[$ . Let  $x \in ]a-1, b+1[$  then we have to consider the following cases:

**$x = a$ .** As  $f'$  is continuous on  $[a, b]$  we have  $\forall \varepsilon > 0$  that there exists a  $\delta' > 0$  such that  $\forall y \in [a, b] \cap ]a-\delta', a+\delta'[$  we have  $\|f'(y) - f'(a)\|_Y < \varepsilon$ . Take  $\delta = \min(\delta, b-a) > 0$  then if  $y \in ]a-\delta, a+\delta[ \cap ]a-1, b+1[$  we have the following cases to consider

**$y \leq a$ .** then  $y \in ]a-\delta, a] \cap ]a-1, b+1[$

$$\begin{aligned} & \|(f^{[a-1,b+1]})'(y) - (f^{[a-1,b+1]})'(a)\|_Y \stackrel{(14.61)}{=} \|f'_+(a) - f'_+(a)\|_Y \\ & = 0 < \varepsilon \end{aligned}$$

$\mathbf{a} < \mathbf{y}$ . then as  $y \in ]a - \delta, a + \delta[ \Rightarrow a < y < a + \delta \leq a + b - a = b$  so that  $y \in ]a, b[$  and  $y \in ]a - \delta', a + \delta'[$  hence

$$\begin{aligned} \|(f^{[a-1, b+1]})'(y) - (f^{[a-1, b+1]})'(a)\|_Y &\stackrel{(14.61)}{=} \|(f_{[a, b]})'(y) - f'_+(a)\|_Y \\ &\stackrel{\text{def}}{=} \|f'(y) - f'(a)\|_Y \\ &< \varepsilon \end{aligned}$$

Hence  $(f^{[a-1, b+1]})'$  is continuous at  $a$

$\mathbf{x} = \mathbf{b}$ . As  $f'$  is continuous on  $[a, b]$  we have  $\forall \varepsilon > 0$  that there exists a  $\delta' > 0$  such that  $\forall y \in [a, b] \cap ]b - \delta', b + \delta'[$  we have  $\|f'(y) - f'(b)\|_Y < \varepsilon$ . Take  $\delta = \min(\delta', b - a) > 0$  then if  $y \in ]b - \delta, b + \delta[ \cap ]a - 1, b + 1[$  we have the following cases to consider

$\mathbf{b} \leq \mathbf{y}$ . then  $y \in [b, b + 1] \cap ]a - 1, b + 1[$

$$\begin{aligned} \|(f^{[a-1, b+1]})'(y) - (f^{[a-1, b+1]})'(b)\|_Y &\stackrel{(14.61)}{=} \|f'_-(b) - f'_-(b)\|_Y \\ &= 0 < \varepsilon \end{aligned}$$

$\mathbf{y} < \mathbf{b}$ . then as  $y \in ]b - \delta, b + \delta[ \Rightarrow y > b - \delta \geq b - (b - a) = a$  so that  $y \in ]a, b[$  and  $y \in ]a - \delta', a + \delta'[$  hence

$$\begin{aligned} \|(f^{[a-1, b+1]})'(y) - (f^{[a-1, b+1]})'(b)\|_Y &\stackrel{(14.61)}{=} \|(f_{[a, b]})'(y) - f'_-(b)\|_Y \\ &\stackrel{\text{def}}{=} \|f'(y) - f'(b)\|_Y \\ &< \varepsilon \end{aligned}$$

Hence  $f'$  is continuous at  $b$

$\mathbf{x} \in ]a - 1, a[$ . Take  $\delta = \min(a - x, x - (a - 1)) > 0$  then for  $y \in ]x - \delta, x + \delta[ \cap ]a - 1, b + 1[$  we have  $y \in ]x - (x - (a - 1)), x + (a - x)[ = ]a - 1, a[$ . So

$$\begin{aligned} \|(f^{[a-1, b+1]})'(y) - (f^{[a-1, b+1]})'(x)\|_Y &\stackrel{(14.61)}{=} \|f'_+(a) - f'_+(a)\|_Y \\ &= 0 < \varepsilon \end{aligned}$$

Hence  $(f^{[a-1, b+1]})'$  is continuous at  $x$ .

$\mathbf{x} \in ]b, b + 1[$ . Take  $\delta = \min(b + 1 - x, x - b) > 0$  then for  $y \in ]x - \delta, x + \delta[ \cap ]a - 1, b + 1[$  we have  $y \in ]x - (x - b), x + (b + 1 - x)[ = ]b, b + 1[$ . So

$$\begin{aligned} \|(f^{[a-1, b+1]})'(y) - (f^{[a-1, b+1]})'(x)\|_Y &\stackrel{(14.61)}{=} \|f'_-(b) - f'_-(b)\|_Y \\ &= 0 < \varepsilon \end{aligned}$$

Hence  $(f^{[a-1, b+1]})'$  is continuous at  $x$ .

$\mathbf{x} \in ]a, b[$ . Let  $\varepsilon > 0$  As  $f'$  is continuous at  $x$  there exists a  $\delta' > 0$  such that for  $y \in ]x - \delta', x + \delta'[ \cap [a, b]$  we have  $\|f'(x) - f'(y)\|_Y < \varepsilon$ . Take then  $\delta = \min(x - a, b - x, \delta') > 0$  then if  $y \in ]x - \delta, x + \delta[ \subseteq ]x - (x - a), x + (b - x)[ = ]a, b[$  we have

$$\begin{aligned} \|(f^{[a-1, b+1]})'(y) - (f^{[a-1, b+1]})'(x)\|_Y &\stackrel{(14.61)}{=} \|(f_{[a, b]})'(y) - (f_{[a, b]})'(x)\|_Y \\ &\stackrel{\text{def}}{=} \|f'(y) - f'(x)\|_Y \\ &< \varepsilon \end{aligned}$$

Hence  $(f^{[a-1, b+1]})'$  is continuous at  $x$ .

The above proves that  $\forall x \in ]a - 1, b + 1[$   $(f^{[a-1, b+1]})'$  is continuous at  $x$ . So by definition  $f^{[a-1, b+1]}$  is  $C^1$ . As further  $[a, b] \subseteq ]a - 1, b + 1[$  a open set,  $(f^{[a-1, b+1]})|_{[a, b]} = f$  it follow by definition that

$f$  is of class  $C^1$

proving (1).

The final part of the theorem is (14.58).  $\square$

**Example 14.94.** Let  $\langle \mathbb{R}, \|\cdot\| \rangle$  be the real normed space,  $\langle Y, \|\cdot\|_Y \rangle$  a normed space over  $\mathbb{R}$ ,  $x, y \in Y$  then  $\sigma: [a, b] \rightarrow Y$  defined by  $\sigma(t) = x + t \cdot y$  is  $C^1$  and  $\sigma'(t) = y$  if  $t \in ]a, b[$  together with  $\sigma'_+(a) = y$ ,  $\sigma'_-(b) = y$

**Proof.** Define  $\sigma^{\mathbb{R}}: \mathbb{R} \rightarrow Y$  by  $\sigma^{\mathbb{R}}(t) = x + t \cdot y$ , using 14.60 we have then that  $\sigma^{\mathbb{R}}$  is  $C^1$  with  $(\sigma^{\mathbb{R}})' = C_{\mathbb{R}, Y, y}$ . As  $\mathbb{R}$  is open,  $[a, b] \subseteq \mathbb{R}$  and  $(\sigma^{\mathbb{R}})|_{[a, b]} = \sigma$  we have by definition that

$$\sigma \text{ is } C^1$$

Using the previous theorem [see 14.93] we have then that

$$\begin{aligned}\sigma'_+(a) &= ((\sigma^{\mathbb{R}})')|_{[a, b]}(a) = y \\ \sigma'_-(b) &= ((\sigma^{\mathbb{R}})')|_{[a, b]}(b) = y \\ \forall t \in ]a, b[ \sigma'(t) &= ((\sigma^{\mathbb{R}})')|_{[a, b]}(t) = y\end{aligned}$$

□

## 14.4 Intermediate value and mean value theorems

**Definition 14.95.** Let  $X$  be a topological space,  $U \subseteq X$  open  $f: U \rightarrow \mathbb{R}$  then  $x \in U$  is a **local weak minimum (maximum) of  $f$**  if there exists a  $\delta > 0$  such that  $B_\delta(x) \subseteq U$  and  $\forall y \in B_\delta(x)$  we have  $f(x) \leq f(y)$  [or  $f(y) \leq f(x)$ ]. A local minimum or maximum is also called a **local extremum**.

**Theorem 14.96.** Let  $\langle \mathbb{R}, \|\cdot\| \rangle$  be the real normed space,  $U \subseteq \mathbb{R}$  and  $f: U \rightarrow \mathbb{R}$  be a function and  $x \in U$  a local extremum then if  $f'(x)$  exists we have  $f'(x) = 0$ .

**Proof.** We have two cases to consider

**local weak maximum.** Then there exists a  $\delta > 0$  such that  $]x - \delta, x + \delta[ \subseteq U$  and  $\forall y \in ]x - \delta, x + \delta[$  we have  $f(y) \leq f(x)$ . Assume that  $f'(x) \neq 0$  then we have either

**$0 < f'(x)$ .** Take  $\varepsilon = f'(x) > 0$  then by definition there exists a  $\delta' > 0$  such that if  $0 < |h| < \delta'$  and  $h \in U_x$  then  $\left| \frac{f(x+h) - f(x)}{h} - f'(x) \right| < \varepsilon = f'(x)$ . Let  $h = \frac{\min(\delta, \delta')}{2}$  then  $0 < h < \delta$ ,  $\delta' < h < x + h < x + \delta \Rightarrow x + h \in U \Rightarrow h \in U_x$  so that  $f'(x) \leq f'(x) - \frac{f(x+h) - f(x)}{h} \leq \left| f'(x) - \frac{f(x+h) - f(x)}{h} \right| < \varepsilon = f'(x)$  a contradiction.

**$f(x) < 0$ .** Take  $\varepsilon = -f'(x) > 0$  then by definition there exists a  $\delta' > 0$  such that if  $0 < |h| < \delta'$  and  $h \in U_x$  then  $\left| \frac{f(x+h) - f(x)}{h} - f'(x) \right| < \varepsilon = -f'(x)$ . Let  $h = \frac{\min(\delta, \delta')}{2}$  then  $0 < h < \delta$ ,  $\delta' < h < x - h < x \Rightarrow x - h \in U \Rightarrow -h \in U_x$  so that  $-f'(x) \leq f'(x) - \frac{f(x-h) - f(x)}{-h} \leq \left| f'(x) - \frac{f(x-h) - f(x)}{-h} \right| < \varepsilon = -f'(x)$  a contradiction.

So as in all cases we have a contradiction.

**local weak minimum.** Then there exists a  $\delta > 0$  such that  $]x - \delta, x + \delta[ \subseteq U$  and  $\forall y \in ]x - \delta, x + \delta[$  we have  $f(x) \leq f(y)$ . Assume that  $f'(x) \neq 0$  then we have either

**$0 < f'(x)$ .** Take  $\varepsilon = f'(x) > 0$  then by definition there exists a  $\delta' > 0$  such that if  $0 < |h| < \delta'$  and  $h \in U_x$  then  $\left| \frac{f(x+h) - f(x)}{h} - f'(x) \right| < \varepsilon = f'(x)$ . Let  $h = \frac{\min(\delta, \delta')}{2}$  then  $0 < h < \delta$ ,  $\delta' < h < x - h < x \Rightarrow x - h \in U \Rightarrow -h \in U_x$  so that  $f'(x) \leq f'(x) - \frac{f(x-h) - f(x)}{-h} \leq \left| f'(x) - \frac{f(x-h) - f(x)}{-h} \right| < \varepsilon = f'(x)$  a contradiction.

**$f'(x) < 0$ .** Take  $\varepsilon = -f'(x) > 0$  then by definition there exists a  $\delta' > 0$  such that if  $0 < |h| < \delta'$  and  $h \in U_x$  then  $\left| \frac{f(x+h) - f(x)}{h} - f'(x) \right| < \varepsilon = -f'(x)$ . Let  $h = \frac{\min(\delta, \delta')}{2}$  then  $0 < h < \delta$ ,  $\delta' < h < x + h < x \Rightarrow x + h \in U \Rightarrow h \in U_x$  so that  $-f'(x) \leq f'(x) - \frac{f(x+h) - f(x)}{h} \leq \left| f'(x) - \frac{f(x+h) - f(x)}{h} \right| < \varepsilon = -f'(x)$  a contradiction.

So in all cases we have a contradiction.

As in all cases we reach a contradiction it follows that the hypothesis  $f'(x) \neq 0$  is false hence

$$f'(x) = 0$$

□

**Theorem 14.97. (Rolle's theorem)** Let  $[a, b] \subseteq \mathbb{R}$ ,  $a < b$ , the normed space  $\langle \mathbb{R}, \|\cdot\| \rangle$  [which is a Banach space by 11.360] and  $f \in \mathcal{C}([a, b], \mathbb{R})$  such that  $f(a) = f(b) = 0$  and  $f$  is differentiable on  $]a, b[$  then there exists a  $\zeta \in ]a, b[$  such that  $0 = f'(\zeta)$ .

**Proof.** Using 11.455 there exists a  $c, d \in \mathbb{R}$  such that  $f([a, b]) = [c, d]$  then as  $c \leq f(a) = f(b) = 0 \leq d \Rightarrow c \leq 0 \leq d$  we have the following cases to consider:

**$c = d = 0$ .** Then  $f$  is constant and  $\forall \zeta \in ]a, b[$  we have  $f'(\zeta) = Df(\zeta)(1) = 0$  [see 14.16]

**$c < 0$ .** As  $f(a) = f(b) = 0$  and  $f([a, b]) = [c, d]$  there exists a  $\zeta \in ]a, b[$  such that  $f(\zeta) = c > 0 = f(a) = f(b)$  hence  $\zeta \in ]a, b[$ . Further  $\forall x \in ]a, b[$  we have  $f(x) \in [c, d] \Rightarrow f(\zeta) = c \leq f(x)$  proving that  $\zeta$  is a local minimum. Using the previous theorem (14.96) it follows that  $f'(\zeta) = 0$ .

**$0 < d$ .** As  $f(a) = f(b) = 0$  and  $f([a, b]) = [c, d]$  there exists a  $\zeta \in ]a, b[$  such that  $f(\zeta) = d > 0 = f(a) = f(b)$  hence  $\zeta \in ]a, b[$ . Further  $\forall x \in ]a, b[$  we have  $f(x) \in [c, d] \Rightarrow f(x) \leq d = f(\zeta)$  proving that  $\zeta$  is a local maximum. Using the previous theorem (14.96) it follows that  $f'(\zeta) = 0$ . □

**Theorem 14.98. (Lagrange's theorem)** Let  $[a, b] \subseteq \mathbb{R}$ ,  $a < b$   $\langle \mathbb{R}, \|\cdot\| \rangle$  and  $f \in \mathcal{C}([a, b], \mathbb{R})$  which have a derivative everywhere in  $]a, b[$  then there exists a  $\zeta \in ]a, b[$  such that  $f(b) - f(a) = f'(\zeta) \cdot (b - a)$

**Proof.** Define the function  $g: [a, b] \rightarrow \mathbb{R}$  by  $g(x) = f(x) - \left( f(a) + \frac{f(b) - f(a)}{b - a} \cdot (x - a) \right)$  which is differentiable in  $]a, b[$  and continuous on  $[a, b]$  [because constant functions are  $C^\infty$ , linear functions in a finite dimensional spaces are continuous [see 11.181] and thus  $C^\infty$  [see 14.75] and the sum of  $C^\infty$  functions is  $C^\infty$ ]. Now  $g(a) = f(a) - \left( f(a) + \frac{f(b) - f(a)}{b - a} \cdot (a - a) \right) = f(a) - f(a) = 0$  and also  $g(b) = f(b) - \left( f(a) + \frac{f(b) - f(a)}{b - a} \cdot (b - a) \right) = f(b) - f(a) - f(b) + f(a) = 0$ . Applying Rolle's theorem we have the existence of a  $\zeta \in ]a, b[$  such that  $0 = g'(\zeta) = 0 \Rightarrow 0 = g'(\zeta) = f'(\zeta) - \frac{f(b) - f(a)}{b - a} \Rightarrow f(b) - f(a) = f'(\zeta) \cdot (b - a)$ . □

**Corollary 14.99.** Let  $U \subseteq \mathbb{R}$  be a open set,  $f: U \rightarrow \mathbb{R}$  such that  $f$  has a derivative  $\forall x \in U$ ,  $\emptyset \neq A \subseteq U$  such that  $\forall x \in A$  we have  $0 < f'(x)$  [or  $0 \leq f'(x)$ ] then  $f$  is strictly increasing [or increasing] on  $A$  and if  $\forall x \in A$  we have that  $f'(x) < 0$  [or  $f'(x) \leq 0$ ] then  $f$  is strictly decreasing [or decreasing] on  $A$

**Proof.** First as  $\forall x \in U$   $f'(x)$  exists we have by 14.19 that  $f$  is differentiable at  $x$  and thus continuous at  $x$  [see 14.15]. So if  $x, y \in A$  with  $x < y$  we can apply Lagrange's theorem [see 14.98] so that  $f(y) - f(x) = f'(\varepsilon) \cdot (y - x)$  where  $\varepsilon \in ]x, y[$ .

Assume that  $\forall z \in A$   $0 < f'(z)$  then  $f'(\varepsilon) > 0$ ,  $y - x > 0$  so that  $f(y) - f(x) = f'(\varepsilon) \cdot (y - x) > 0 \Rightarrow f(y) > f(x)$  proving that  $f$  is strictly increasing on  $A$ .

Assume  $\forall z \in A$   $0 \leq f'(z)$  then  $f'(\varepsilon) \geq 0$ ,  $y - x > 0$  so that  $f(y) - f(x) = f'(\varepsilon) \cdot (y - x) \geq 0 \Rightarrow f(y) \geq f(x)$  proving that  $f$  is increasing on  $A$ .

Assume  $\forall z \in A$  we have  $f'(z) < 0$  [or  $f'(z) \leq 0$ ] then we have that  $(-f)'(z) = -f'(z) > 0$  [or  $(-f)'(z) = -f'(z) \geq 0$ ] so that we have that  $-f$  is strictly increasing (or increasing) proving that  $f$  is strictly decreasing (decreasing). □

**Definition 14.100.** Let  $U \subseteq \mathbb{R}$  be a open set,  $f: U \rightarrow \mathbb{R}$  a function  $a, b \in U$  with  $a < b$  and  $[a, b] \subseteq U$  then  $f$  is convex [or concave] on  $[a, b]$  if  $\forall x \in [a, b]$  we have  $f(a) + \frac{f(b) - f(a)}{b - a} \cdot (x - a) \geq f(x)$  (or  $f(a) + \frac{f(b) - f(a)}{b - a} \cdot (x - a) \leq f(x)$ ). Essentially  $f$  is convex (concave) if a straight line between  $f(a)$  and  $f(b)$  lies above [below] the curve  $f: [a, b] \rightarrow \mathbb{R}$ .

We can easily rewrite the this definition as follows

**Lemma 14.101.** Let  $U \subseteq \mathbb{R}$  be a open set,  $f: U \rightarrow \mathbb{R}$  a function  $a, b \in U$  with  $a < b$  then  $f$  is convex [or concave] on  $[a, b]$  if  $\forall x \in [a, b]$  we have  $\frac{f(a) \cdot (b - x) + f(b) \cdot (x - a)}{b - a} \geq f(x)$  [or  $\frac{f(a) \cdot (b - x) + f(b) \cdot (x - a)}{b - a} \leq f(x)$ ]

**Proof.** this follows from the following

$$\begin{aligned}
 f(a) + \frac{f(b) - f(a)}{b - a} \cdot (x - a) &= \frac{f(a) \cdot (b - a) + (f(b) - f(a)) \cdot (x - a)}{b - a} \\
 &= \frac{f(a) \cdot b - f(a) \cdot a + f(b) \cdot x - f(b) \cdot a - f(a) \cdot x + f(a) \cdot a}{b - a} \\
 &= \frac{f(a) \cdot (b - x) + f(b) \cdot (x - a)}{b - a} \\
 &\quad \square
 \end{aligned}$$

**Lemma 14.102.** Let  $U \subseteq \mathbb{R}$  be a open set,  $f: U \rightarrow \mathbb{R}$  a function,  $a, b \in U$  with  $a < b$  and  $[a, b] \subseteq U$  then  $f$  is convex [or concave] on  $[a, b]$  if  $\forall \alpha, \beta$  with  $0 \leq \alpha, \beta$  and  $\alpha + \beta = 1$  we have that  $f(\alpha \cdot a + \beta \cdot b) \geq \alpha \cdot f(a) + \beta \cdot f(b)$  [or  $f(\alpha \cdot a + \beta \cdot b) \leq \alpha \cdot f(a) + \beta \cdot f(b)$ ]

**Proof.** If  $x \in [a, b]$  then  $0 \leq \frac{x-a}{b-a} = \beta$ ,  $\frac{b-x}{b-a} = \alpha$  and  $\alpha + \beta = \beta + \alpha = \frac{x-a+b-x}{b-a} = 1$  and

$$\begin{aligned}
 \alpha \cdot a + \beta \cdot b &= \frac{(b - x) \cdot a + (x - a) \cdot b}{b - a} \\
 &= \frac{b \cdot a - x \cdot a + x \cdot b - a \cdot b}{b - a} \\
 &= \frac{x \cdot (b - a)}{b - a} \\
 &= x
 \end{aligned}$$

Further if  $\exists \alpha, \beta \vdash 0 \leq \alpha, \beta \wedge \alpha + \beta = 1$  then

$$\begin{aligned}
 \alpha \cdot a + \beta \cdot b &\geq_{0 \leq \beta \wedge \alpha < b} \alpha \cdot a + \beta \cdot a \\
 &= \alpha \cdot a + (1 - \alpha) \cdot a \\
 &= a \\
 \alpha \cdot a + \beta \cdot b &\leq_{0 \leq \alpha \wedge a < b} \alpha \cdot b + \beta \cdot b \\
 &= \alpha \cdot b + (1 - \alpha) \cdot b \\
 &= b
 \end{aligned}$$

Hence we have

$$x \in [a, b] \Leftrightarrow \exists \alpha, \beta \vdash 0 \leq \alpha, \beta \wedge \alpha + \beta = 1 \text{ such that } x = \alpha \cdot a + \beta \cdot b \quad (14.62)$$

So if  $x \in [a, b]$  there exists a  $\alpha, \beta$  such that  $0 \leq \alpha, \beta$  and  $\alpha + \beta = 1$

$$\begin{aligned}
 \frac{f(a) \cdot (b - x) + f(b) \cdot (x - a)}{b - a} &\stackrel{(14.62)}{=} \frac{f(a) \cdot (b - \alpha \cdot a - \beta \cdot b) + f(b) \cdot (\alpha \cdot a + \beta \cdot b - a)}{b - a} \\
 &= \frac{f(a) \cdot (b \cdot (1 - \beta) - \alpha \cdot a) + f(b) \cdot (\beta \cdot b - a \cdot (1 - \alpha))}{b - a} \\
 &= \frac{f(a) \cdot (b \cdot \alpha - \alpha \cdot a) + f(b) \cdot (\beta \cdot b - a \cdot \beta)}{b - a} \\
 &= \frac{f(a) \cdot (b - a) \cdot \alpha + f(b) \cdot (b - a) \cdot \beta}{b - a} \\
 &= \alpha \cdot f(a) + \beta \cdot f(b)
 \end{aligned}$$

Hence as  $f(x) = f(\alpha \cdot a + \beta \cdot b)$  we have that  $f$  is convex [concave] if  $\alpha \cdot f(a) + \beta \cdot f(b) \geq f(\alpha \cdot a + \beta \cdot b)$  (or  $\alpha \cdot f(a) + \beta \cdot f(b) \leq f(\alpha \cdot a + \beta \cdot b)$ ).  $\square$

The next theorem shows the relation between the second derivate and convex and concave functions.

**Theorem 14.103.** Let  $U \subseteq \mathbb{R}$  be a open set,  $f: U \rightarrow \mathbb{R}$  a function  $a, b \in U$  with  $a < b$  and  $[a, b] \subseteq U$  such that  $\forall x \in U$  we have that  $f''(x)$  exists and  $\forall x \in ]a, b[$  we have  $f''(x) \geq 0$  [or  $f''(x) \leq 0$ ] then  $f$  is convex (or concave) on  $[a, b]$ .

**Proof.** Let  $x \in [a, b]$  then we have either

$$x = a. \text{ then } \frac{f(a) \cdot (b-x) + f(b) \cdot (x-a)}{b-a} = f(a) = f(x)$$

$$x = b. \text{ then } \frac{f(a) \cdot (b-x) + f(b) \cdot (x-a)}{b-a} = f(b) = f(x)$$

$x \in ]a, b[$ . Then

$$\begin{aligned} \frac{(x-a)}{(b-a)} + \frac{(b-x)}{(b-a)} &= \frac{x-a+b-x}{b-a} \\ &= 1 \end{aligned}$$

so by multiplying both sides by  $f(x)$  we have  $f(x) = \frac{x-a}{b-a} \cdot f(x) + \frac{b-x}{b-a} \cdot f(x)$  so that

$$\begin{aligned} \frac{f(a) \cdot (b-x) + f(b) \cdot (x-a)}{b-a} - f(x) &= \frac{b-x}{b-a} \cdot f(a) + \frac{x-a}{b-a} \cdot f(b) - \frac{x-a}{b-a} \cdot f(x) - \frac{b-x}{b-a} \cdot f(x) \\ &= \frac{(x-a)}{b-a} \cdot (f(b) - f(x)) - \frac{b-x}{b-a} \cdot (f(x) - f(a)) \\ &= \frac{(x-a) \cdot (b-x)}{b-a} \cdot \frac{f(b) - f(x)}{b-x} - \frac{(b-x) \cdot (x-a)}{b-a} \cdot \frac{f(x) - f(a)}{x-a} \\ &= \frac{(x-a) \cdot (b-x)}{b-a} \left[ \frac{f(b) - f(x)}{b-x} - \frac{f(x) - f(a)}{x-a} \right] \end{aligned} \quad (14.63)$$

Now as  $\forall x \in U f''(x)$  exist we have that  $f'(x)$  exists and that  $f$  is continuous at  $x$  so using Lagrange's theorem (see 14.98) there exists a  $y_1 \in ]x, b[$  and  $y_2 \in ]a, x[$  such that  $\frac{f(b) - f(x)}{(b-x)} = f'(y_1)$  and  $\frac{f(x) - f(a)}{(x-a)} = f'(y_2)$  which using the above give

$$\frac{f(a) \cdot (b-x) + f(b) \cdot (x-a)}{b-a} - f(x) = \frac{(x-a) \cdot (b-x)}{b-a} \cdot (f'(y_1) - f'(y_2)) \quad (14.64)$$

As  $\forall x \in U f''(x)$  exists we have that  $f'(x)$  exists is continuous, further from  $y_1 \in ]x, b[$  and  $y_2 \in ]a, x[$  we have that  $y_2 < y_1$ . Using 14.98 there exists a  $z \in ]y_2, y_1[$  such that  $\frac{f'(y_1) - f'(y_2)}{(y_1 - y_2)} = (f')'(z) = f''(x)$  which by substituting in 14.64 gives

$$\begin{aligned} \frac{f(a) \cdot (b-x) + f(b) \cdot (x-a)}{b-a} - f(x) &= \frac{(x-a) \cdot (b-x) \cdot (y_1 - y_2)}{b-a} \cdot f''(x) \\ &= \alpha \cdot f''(x) \end{aligned}$$

where  $\alpha = \frac{(x-a) \cdot (b-x) \cdot (y_1 - y_2)}{b-a} > 0$ . Hence if  $f''(x) \geq 0$  [or  $f''(x) \leq 0$ ] then we have  $\frac{f(a) \cdot (b-x) + f(b) \cdot (x-a)}{b-a} - f(x) \geq 0$  [or  $\leq 0$ ], proving that

$$\frac{f(a) \cdot (b-x) + f(b) \cdot (x-a)}{b-a} \geq f(x) \text{ [or } \leq f(x)]$$

So in all cases we have  $\frac{f(a) \cdot (b-x) + f(b) \cdot (x-a)}{b-a} \geq f(x)$  if  $f''(x) \geq 0$  or  $\frac{f(a) \cdot (b-x) + f(b) \cdot (x-a)}{b-a} \leq f(x)$  if  $f''(x) \leq 0$  proving that  $f$  is convex [or concave].  $\square$

We can use convexity or concavity as a condition to see if a extremum is a local minimum or a local maximum on a interval.

**Proposition 14.104.** *Let  $U \subseteq \mathbb{R}$  be a open set,  $f: U \rightarrow \mathbb{R}$  such that  $\forall x \in U f''(x)$  exists and  $\exists x_0 \in U$  such that  $f'(x_0) = 0$  then we have*

1. *If there exists a open  $V$  such that  $x_0 \in V \subseteq U$  and  $\forall x \in V f''(x) \geq 0$  then  $x_0$  is a local weak minimum of  $f$*

2. If there exists a open  $V$  such that  $x_0 \in V \subseteq U$  and  $\forall x \in V f''(x) \leq 0$  then  $x_0$  is a local weak maximum of  $f$

**Proof.**

1. As  $V$  is open there exists a  $\delta > 0$  such that  $]x_0 - \delta, x_0 + \delta[ \subseteq V$  hence

$$\forall x \in ]x_0 - \delta, x_0 + \delta[ \text{ we have } f''(x) \geq 0$$

Using the fact that  $f''(x) = (f')'(x)$  we have by the above and 14.99 that  $f'$  is increasing on  $]x_0 - \delta, x_0 + \delta[$  hence we have

$$\forall x \in ]x_0 - \delta, x_0] f'(x) \leq f'(x_0) = 0 \text{ and } \forall x \in [x_0, x_0 + \delta[ 0 = f'(x_0) \leq f'(x)$$

So using 14.99 we have that  $f$  is decreasing on  $]x_0 - \delta, x_0] \Rightarrow \forall x \in ]x_0 - \delta, x_0]$  we have  $f(x) \geq f(x_0)$  and  $f$  is increasing on  $[x_0, x_0 + \delta[ \Rightarrow \forall x \in [x_0, x_0 + \delta[$  we have  $f(x_0) \leq f(x)$ . Hence

$$\forall x \in ]x_0 - \delta, x_0 + \delta[ \text{ we have } f(x_0) \leq f(x) \text{ or } x_0 \text{ is a local weak minimum}$$

2. As  $V$  is open there exists a  $\delta > 0$  such that  $]x_0 - \delta, x_0 + \delta[ \subseteq V$  hence

$$\forall x \in ]x_0 - \delta, x_0 + \delta[ \text{ we have } f''(x) \leq 0$$

Using the fact that  $f''(x) = (f')'(x)$  we have by the above and 14.99 that  $f'$  is decreasing on  $]x_0 - \delta, x_0 + \delta[$  hence we have

$$\forall x \in ]x_0 - \delta, x_0] f'(x) \geq f'(x_0) = 0 \text{ and } \forall x \in [x_0, x_0 + \delta[ 0 = f'(x_0) \geq f'(x)$$

So using 14.99 we have that  $f$  is increasing on  $]x_0 - \delta, x_0] \Rightarrow \forall x \in ]x_0 - \delta, x_0]$  we have  $f(x) \leq f(x_0)$  and  $f$  is decreasing on  $[x_0, x_0 + \delta[ \Rightarrow \forall x \in [x_0, x_0 + \delta[$  we have  $f(x_0) \geq f(x)$ . Hence

$$\forall x \in ]x_0 - \delta, x_0 + \delta[ \text{ we have } f(x_0) \leq f(x) \text{ or } x_0 \text{ is a local weak maximum}$$

□

Actually we can extend the above to maximum and minimums on a interval as follows

**Corollary 14.105.** Let  $U \subseteq \mathbb{R}$  be a open set,  $f: U \rightarrow \mathbb{R}$  a function such that  $\forall x \in U f''(x)$  exists and  $f''(x) \geq 0$  [or  $f''(x) \leq 0$ ] then if  $a, b \in U$  with  $a \leq b$ ,  $[a, b] \subseteq U$  such that  $\exists x_0 \in [a, b]$  with  $f'(x_0) = 0$  we have that  $\forall x \in [a, b] f(x) \geq f(x_0)$  ( $x_0$  is a minimum of  $f$  on  $[a, b]$ ) [or  $\forall x \in [a, b] f(x) \leq f(x_0)$  ( $x_0$  is a maximum of  $f$  on  $[a, b]$ )]

**Proof.** If  $a = b \Rightarrow [a, b] = \{x_0\}$  so that the theorem is trivial, so we may assume that  $a < b$  in the proof. We have then to consider the following cases:

$0 \leq f''(x)$ . From the previous corollary [see 14.104] we conclude that  $x_0$  is a local weak minimum of  $f$  hence there exists a  $\delta > 0$  such that

$$\forall x \in ]x_0 - \delta, x_0 + \delta[ \subseteq U \text{ we have } f(x_0) \leq f(x) \quad (14.65)$$

Assume that there exists a  $y \in [a, b]$  such that  $f(y) < f(x_0)$  then we can not have  $y = x_0$  so we must have either:

$y < x_0$ . Then  $\max(x_0 - \delta, y) < x_0$  and there exists a  $z$  such that  $x_0 - \delta, y < z < x_0$  [see 9.56]. Applying 14.65 we have that

$$f(z) \geq f(x_0) \quad (14.66)$$

Using 14.103 we have that  $f$  is convex on  $[y, x_0]$ . So we have by 14.101 that

$$\frac{f(y) \cdot (x_0 - z) + f(x_0) \cdot (z - y)}{x_0 - y} \geq f(z) \quad (14.67)$$

As  $0 < x_0 - z$  and  $f(y) < f(x_0)$  [by assumption] it follows that

$$f(y) \cdot (x_0 - z) < f(x_0) \cdot (x_0 - z) \quad (14.68)$$

Hence we have

$$\begin{aligned} f(x_0) &= \frac{f(x_0) \cdot (x_0 - z) + f(x_0) \cdot (z - y)}{x_0 - y} \\ &>_{(14.68)} \frac{f(y) \cdot (x_0 - z) + f(x_0) \cdot (z - y)}{x_0 - y} \\ &\geq_{(14.67)} f(z) \\ &\geq_{(14.66)} f(x_0) \end{aligned}$$

a contradiction.

**$x_0 < y$ .** Then  $x_0 < \min(y, x_0 + \delta)$  and there exists a  $z$  such that  $x_0 < z < y, x_0 + \delta$  [see 9.56]. Applying (14.65) we have that

$$f(x_0) \leq f(z) \quad (14.69)$$

Using 14.103 we have that  $f$  is convex on  $[x_0, y]$ . So we have by 14.101 that

$$\frac{f(x_0) \cdot (y - z) + f(y) \cdot (z - x_0)}{y - x_0} \geq f(z) \quad (14.70)$$

As  $z - x_0 > 0$  and  $f(y) < f(x_0)$  [by assumption] we have

$$f(y) \cdot (z - x_0) < f(x_0) \cdot (z - x_0) \quad (14.71)$$

Hence we have

$$\begin{aligned} f(x_0) &= \frac{f(x_0) \cdot (y - z) + f(x_0) \cdot (z - x_0)}{y - x_0} \\ &>_{(14.71)} \frac{f(x_0) \cdot (y - z) + f(y) \cdot (z - x_0)}{y - x_0} \\ &\geq_{(14.70)} f(z) \\ &\geq_{14.69} f(x_0) \end{aligned}$$

a contradiction.

As we reach in all cases a contradiction we must have that  $\forall x \in [a, b] f(x_0) \leq f(x)$  proving that  $x_0$  is a minimum of  $f$  on  $[a, b]$ .

**$f''(x) \leq 0$ .** From the previous corollary [see 14.104] we conclude that  $x_0$  is a local weak maximum of  $f$  hence there exists a  $\delta > 0$  such that

$$\forall x \in ]x_0 - \delta, x_0 + \delta[ \subseteq U \text{ we have } f(x_0) \geq f(x) \quad (14.72)$$

Assume that there exists a  $y \in [a, b]$  such that  $f(y) > f(x_0)$  then we can not have  $y = x_0$  so we must have either:

**$y < x_0$ .** Then  $\max(x_0 - \delta, y) < x_0$  and there exists a  $z$  such that  $x_0 - \delta, y < z < x_0$  [see 9.56]. Applying 14.72 we have that

$$f(z) \leq f(x_0) \quad (14.73)$$

Using 14.103 we have that  $f$  is concave on  $[y, x_0]$ . So we have by 14.101 that

$$\frac{f(y) \cdot (x_0 - z) + f(x_0) \cdot (z - y)}{x_0 - y} \leq f(z) \quad (14.74)$$

As  $0 < x_0 - z$  and  $f(y) > f(x_0)$  [by assumption] it follows that

$$f(y) \cdot (x_0 - z) > f(x_0) \cdot (x_0 - z) \quad (14.75)$$

Hence we have

$$\begin{aligned} f(x_0) &= \frac{f(x_0) \cdot (x_0 - z) + f(x_0) \cdot (z - y)}{x_0 - y} \\ &<_{14.75} \frac{f(y) \cdot (x_0 - z) + f(x_0) \cdot (z - y)}{x_0 - y} \\ &\leq_{(14.74)} f(x) \\ &\leq_{(14.73)} f(x_0) \end{aligned}$$

a contradiction.

$x_0 < y$ . Then  $x_0 < \min(y, x_0 + \delta)$  and there exists a  $z$  such that  $x_0 < z < y, x_0 + \delta$  [see 9.56]. Applying (14.72) we have that

$$f(x_0) \geq f(z) \quad (14.76)$$

Using 14.103 we have that  $f$  is concave on  $[x_0, y]$ . So we have that

$$\frac{f(x_0) \cdot (y - z) + f(y) \cdot (z - x_0)}{y - x_0} \leq f(z) \quad (14.77)$$

As  $0 < z - x_0$  and  $f(y) > f(x_0)$  [by assumption] we have

$$f(x_0) \cdot (z - x_0) < f(y) \cdot (z - x_0) \quad (14.78)$$

Hence we have

$$\begin{aligned} f(x_0) &= \frac{f(x_0) \cdot (y - z) + f(x_0) \cdot (z - x_0)}{y - x_0} \\ &\stackrel{(14.78)}{<} \frac{f(x_0) \cdot (y - z) + f(y) \cdot (z - x_0)}{y - x_0} \\ &\stackrel{(14.77)}{\leq} f(z) \\ &\stackrel{(14.76)}{\leq} f(x_0) \end{aligned}$$

A contradiction.

As we reach in all cases a contradiction we must have that  $\forall x \in [a, b]$  we have  $f(x_0) \geq f(x)$  proving that  $x_0$  is a maximum of  $f$  on  $[a, b]$ .  $\square$

**Lemma 14.106.** Let  $\langle X, \|\cdot\| \rangle$  be a Banach space,  $\langle \mathbb{R}, \|\cdot\| \rangle$  the real normed space,  $[a, b] \subseteq \mathbb{R}, a < b$   $f: [a, b] \rightarrow X$  a continuous function and  $F: [a, b] \rightarrow X$  defined by  $F(x) = \int_a^x f$  [see 11.437 for the definition of the Riemann Integral and 11.442 for the existence of the Riemann Integral]. Then  $F$  is  $C^1$  on  $[a, b]$  and  $F' = f$  [see 14.93 for a definition of ' on  $[a, b]$ ]

**Proof.** We have the following cases to consider for  $x \in [a, b]$

$x = a$ . Let  $\varepsilon > 0$  then as  $f$  is continuous at  $a$  there exists a  $\delta > 0$  such that if  $|y - a| < \delta$  and  $y \in [a, b]$  then  $\|f(y) - f(a)\| < \frac{\varepsilon}{2}$ . So if  $0 < h < \delta$  and  $a + h \in [a, b]$  we have  $\forall y \in [a, a + h] \Rightarrow a \leq y \leq a + h$  that  $|y - a| = y - a < a + h - a = h < \delta$  so that

$$\forall y \in [a, a + h] \text{ we have } \|f(y) - f(a)\| < \frac{\varepsilon}{2} \quad (14.79)$$

Further

$$\begin{aligned} \left\| \frac{F(a + h) - F(a)}{h} - f(a) \right\| &= \left\| \frac{\int_a^{a+h} f - \int_a^a f}{h} - f(a) \right\| \\ &\stackrel{11.443}{=} \left\| \frac{\int_a^{a+h} f}{h} - f(a) \right\| \\ &= \left\| \frac{(\int_a^{a+h} f) - h \cdot f(a)}{h} \right\| \\ &\stackrel{11.445}{=} \left\| \frac{\int_a^{a+h} f - \int_a^{a+h} f(a)}{h} \right\| \\ &= \left\| \frac{\int_a^{a+h} (f - f(a))}{h} \right\| \\ &\stackrel{(14.79) \text{ and } 11.445}{\leq} \frac{\varepsilon \cdot h}{2 \cdot h} \\ &< \varepsilon \end{aligned}$$

So we have that  $F'_+(a) = f(a)$

**$x = b$ .** Let  $\varepsilon > 0$  then as  $f$  is continuous at  $a$  there exists a  $\delta > 0$  such that if  $|y - b| < \delta$  and  $y \in [a, b]$  then  $\|f(y) - f(b)\| < \frac{\varepsilon}{2}$ . So if  $0 < h < \delta$  and  $b - h \in [a, b]$  we have  $\forall y \in [b - h, b] \Rightarrow b - h \leq y \leq b$  that  $|b - y| = b - y \leq b - (b - h) = h < \delta$  so that

$$\forall y \in [b - h, b] \text{ we have } \|f(y) - f(b)\| < \frac{\varepsilon}{2} \quad (14.80)$$

Further

$$\begin{aligned} \left\| \frac{F(b-h) - F(b)}{-h} - f(b) \right\| &= \left\| \frac{\int_a^{b-h} f - \int_a^b f}{-h} - f(b) \right\| \\ &\stackrel{11.447}{=} \left\| \frac{-\int_{b-h}^b f}{-h} - f(b) \right\| \\ &= \left\| \frac{\int_{b-h}^b f}{h} - f(b) \right\| \\ &= \left\| \frac{(\int_{b-h}^b f) - h \cdot f(b)}{h} \right\| \\ &\stackrel{11.445}{=} \left\| \frac{\int_{b-h}^b (f - f(b))}{h} \right\| \\ &\stackrel{(14.80) \text{ and } 11.445}{\leq} \frac{\varepsilon \cdot h}{2 \cdot h} \\ &< \varepsilon \end{aligned}$$

Hence we have that  $F'_-(b) = f(b)$

**$x \in ]a, b[$ .** Let  $\varepsilon > 0$  then as  $f$  is continuous at  $x$  there exists a  $\delta > 0$  such that if  $|y - x| < \delta$  then  $\|f(y) - f(x)\| < \frac{\varepsilon}{2}$ . Now if  $h$  is such that  $0 < |h| < \delta$  and  $x + h \in [a, b]$  then we have the following cases to consider

**$0 < h$ .** If  $y \in [x, x + h] \Rightarrow x \leq y \leq x + h$  we have that  $|y - x| = y - x \leq x + h - x = h < \delta$  so that

$$\forall y \in [x, x + h] \text{ we have } \|f(y) - f(x)\| < \frac{\varepsilon}{2} \quad (14.81)$$

Further

$$\begin{aligned} \left\| \frac{F(x+h) - F(x)}{h} - f(x) \right\| &= \left\| \frac{\int_a^{x+h} f - \int_a^x f}{h} - f(x) \right\| \\ &\stackrel{11.447}{=} \left\| \frac{\int_x^{x+h} f}{h} - f(x) \right\| \\ &= \left\| \frac{(\int_x^{x+h} f) - h \cdot f(x)}{h} \right\| \\ &\stackrel{11.445}{=} \left\| \frac{\int_x^{x+h} (f - f(x))}{h} \right\| \\ &\stackrel{(14.81) \text{ and } 11.445}{\leq} \frac{\varepsilon \cdot h}{2 \cdot h} \\ &< \varepsilon \end{aligned}$$

**$h < 0$ .** In this case  $h = -|h|$ . If  $y \in [x+h, x] = [x-|h|, x] \Rightarrow x-|h| \leq y \leq x$  so that  $|y-x| = x-y = x - (x-|h|) = |h| < \delta$  so that

$$\forall y \in [x+h, x] \text{ we have } \|f(y) - f(x)\| \leq \frac{\varepsilon}{2} \quad (14.82)$$

Further

$$\begin{aligned} \left\| \frac{F(x+h) - F(x)}{h} - f(x) \right\| &= \left\| \frac{F(x-|h|) - F(x)}{-|h|} - f(x) \right\| \\ &= \left\| \frac{\int_a^{x-|h|} f - \int_a^x f}{-|h|} - f(x) \right\| \\ &\stackrel{11.447}{=} \left\| \frac{-\int_{x-|h|}^x f}{-|h|} - f(x) \right\| \\ &= \left\| \frac{\int_{x-|h|}^x f}{|h|} - f(x) \right\| \\ &= \left\| \frac{(\int_{x-|h|}^x f) - h \cdot f(x)}{|h|} \right\| \\ &= \left\| \frac{\int_{x-|h|}^x (f - f(x))}{|h|} \right\| \\ &\stackrel{\leq (14.82) \text{ and } 11.445}{<} \frac{\varepsilon \cdot h}{2 \cdot h} \\ &< \varepsilon \end{aligned}$$

As in all cases we have  $\left\| \frac{F(x+h) - F(x)}{h} - f(x) \right\| < \varepsilon$  it follows that  $F'(x) = f(x)$ .

Considering all the cases  $x=a, x=b$  and  $x \in ]a, b[$  proves that  $F' = f$ . As also  $f$  is continuous on  $[a, b]$  we have by 14.93 that  $F$  is  $C^1$ .  $\square$

We prove now the classical fundamental theorem of Calculus.

**Theorem 14.107. (Fundamental Theorem of Calculus (Classical))** Let  $\langle X, \|\cdot\| \rangle$  be a Banach space,  $\langle \mathbb{R}, \|\cdot\| \rangle$  the real normed space,  $[a, b] \subseteq \mathbb{R}, a < b$   $f: [a, b] \rightarrow X$  a  $C^1$  function on  $[a, b]$  then we have that  $f(b) - f(a) = \int_a^b f'$

**Proof.** First note that as  $f'$  is defined and continuous on  $[a, b]$  [see 14.93] and that  $\langle X, \|\cdot\| \rangle$  is a Banach space we have that  $\forall x \in [a, b] \int_a^x f'$  is well defined [see 11.447]. So the function  $g: [a, b] \rightarrow X$  defined by  $g(x) = f(a) + \int_a^x f'$  is well defined. Using the previous lemma we have then that  $g$  is  $C^1$  and  $g'(x) = (f(a) + \int_a^x f')' = 0 + f'(x) = f'(x) \Rightarrow g' = f'$  proving that

$$g' = f' \quad (14.83)$$

Note also that  $g(a) = f(a) + \int_a^a f' = f(a) + 0 = f(a)$  or

$$f(a) = g(a) \quad (14.84)$$

Consider now  $\varphi: ]a, b[ \rightarrow \mathbb{R}$  defined by  $\varphi(x) = \|f(x) - g(x)\|$ . Given  $x \in ]a, b[$  choose a  $0 < \delta$  such that  $B_{\|\cdot\|}(x, \delta) \subseteq ]a, b[$  then if  $0 < |h| < \delta$  we have that  $|x+h-x|=|h|<\delta \Rightarrow x+h \in B_{\|\cdot\|}(x, \delta) \Rightarrow x+h \in ]a, b[$  or

$$\text{If } 0 < |h| < \delta \Rightarrow x+h \in ]a, b[ \quad (14.85)$$

So we have that

$$\begin{aligned}
\left| \frac{\varphi(x+h) - \varphi(x)}{h} \right| &= \left| \frac{\|f(x+h) - g(x+h)\| - \|f(x) - g(x)\|}{h} \right| \\
&= \left| \frac{\|f(x+h) - g(x+h)\| - \|g(x) - f(x)\|}{h} \right| \\
&\stackrel{\text{11.70}}{\leq} \frac{\|f(x+h) - g(x+h) + (g(x) - f(x))\|}{|h|} \\
&= \left\| \frac{(f(x+h)) - f(x) - (g(x+h) - g(x))}{h} \right\| \\
&\stackrel{f' = g' \text{ (see 14.83)}}{=} \left\| \frac{f(x+h) - f(x)}{h} - f'(x) - \left( \frac{g(x+h) - g(x)}{h} - g'(x) \right) \right\| \\
&\leq \left\| \frac{f(x+h) - f(x)}{h} - f'(x) \right\| + \left\| \frac{g(x+h) - g(x)}{h} - g'(x) \right\|
\end{aligned}$$

Now take  $\varepsilon > 0$  then there exists a  $\delta', \delta''$  such that if  $0 < |h| < \delta', \delta''$  then  $\left\| \frac{f(x+h) - f(x)}{h} - f'(x) \right\| < \frac{\varepsilon}{2}$  and  $\left\| \frac{g(x+h) - g(x)}{h} - g'(x) \right\| < \frac{\varepsilon}{2}$  so if  $0 < |h| < \delta''' = \min(\delta, \delta', \delta'') > 0$  we have  $\left\| \frac{\varphi(x+h) - \varphi(x)}{h} - 0 \right\| < \varepsilon$  proving that

$$\forall x \in ]a, b] \models \varphi'(x) = 0 \quad (14.86)$$

Given  $x, y \in [a, b]$  with  $x < y$  we have by Lagrange's Theorem (see 14.98) that there exists a  $t \in ]x, y[$  such that  $\varphi(x) - \varphi(y) = \varphi'(t) \cdot (x - y) = 0 \cdot (x - y) = 0$  proving that  $\varphi$  is constant on  $[a, b]$ . As  $\varphi(a) = \|f(a) - g(a)\| = 0$  we have that  $\forall x \in [a, b] \models \varphi(x) = 0 \Rightarrow f(x) = g(x)$  so that  $f(b) = g(b) = f(a) + \int_a^b f' \Rightarrow f(b) - f(a) = \int_a^b f'$  proving the theorem.  $\square$

**Theorem 14.108. (Mean Value Theorem (first version))** *Let  $\langle X, \|\cdot\|_X \rangle$  be a normed vector space,  $a$  open set  $U \subseteq X$ ,  $\langle \mathbb{R}, \|\cdot\| \rangle$  the real normed space and  $f: U \rightarrow \mathbb{R}$  a function that is  $C^1$  on  $U$ . If  $x, y \in U$  such that  $\forall t \in [0, 1]$  we have  $(1-t) \cdot x + t \cdot y = x + t \cdot (y - x) \in U$  then there exists a  $\xi \in [0, 1]$  such that  $f(y) - f(x) = Df(z)(y - x)$  where  $z = x + \xi \cdot (y - x)$*

**Proof.** Define  $\sigma: [0, 1] \rightarrow U$  by  $\sigma(t) = x + t \cdot (y - x)$  then  $\sigma(0) = x$ ,  $\sigma(1) = y$ . Define then  $g = f \circ \sigma: [0, 1] \rightarrow \mathbb{R}$ . This function is continuous and differentiable on  $[0, 1]$  [and thus also on  $]0, 1[$ ] [as  $\sigma$  is  $C^1$  on  $[0, 1]$  by 14.94,  $f$  is  $C^1$  on  $U$ ,  $\sigma[0, 1] \subseteq U$  and 14.89]. By the Lagrange theorem [see 14.98] there exists a  $\xi \in ]0, 1[$  such that

$$g(1) - g(0) = g'(\xi) \cdot (1 - 0) = g'(\xi) \quad (14.87)$$

Using the chain rule we have that

$$\begin{aligned}
g'(\xi) &= Dg(\xi)(1) \\
&= D(f \circ \sigma)(\xi)(1) \\
&= (Df(\sigma(\xi)) \circ (D\sigma(\xi)))(1) \\
&= Df(\sigma(\xi))(D\sigma(\xi)(1)) \\
&= Df(\sigma(\xi))(\sigma'(\xi)) \\
&\stackrel{14.94}{=} Df(\sigma(\xi))(y - x)
\end{aligned} \quad (14.88)$$

Take  $z = x + \xi \cdot (y - x) = \sigma(\xi)$  then we have

$$\begin{aligned}
f(y) - f(x) &= f(\sigma(1)) - f(\sigma(0)) \\
&= g(1) - g(0) \\
&\stackrel{(14.87)}{=} g'(\xi) \\
&\stackrel{(14.88)}{=} Df(\sigma(\xi))(y - x) \\
&= Df(z) \cdot (y - x)
\end{aligned}$$

proving the theorem.  $\square$

**Definition 14.109.** A subset  $C$  of a vector space  $X$  over  $\mathbb{R}$  is convex if for every  $t \in [0, 1]$  we have  $\forall x, y \in C$  that  $(1-t) \cdot x + t \cdot y = x + t \cdot (y - x) \in C$

**Theorem 14.110.** Let  $\langle X, \|\cdot\| \rangle$  be a normed space,  $x \in X$  and  $\delta > 0$  then  $B_{\|\cdot\|}(x, \delta)$  and  $\bar{B}_{\|\cdot\|}(x, \delta)$  are convex.

**Proof.** Let  $t \in [0, 1]$  then for  $x, y, z \in X$  we have

$$\begin{aligned} \|y + t \cdot (z - y) - x\| &= \|y + t \cdot (z - y) - x + t \cdot x - t \cdot x\| \\ &= \|y \cdot 1 + t \cdot z - t \cdot y - x \cdot 1 + t \cdot x - t \cdot x\| \\ &= \|y \cdot (1-t) - x \cdot (1-t) + t \cdot z - t \cdot x\| \\ &= \|(y - x) \cdot (1-t) + t \cdot (z - x)\| \\ &\leq (1-t) \cdot \|y - x\| + t \cdot \|z - x\| \end{aligned} \tag{14.89}$$

So if  $y, z \in B_{\|\cdot\|}(x, \delta)$  then

$$\begin{aligned} \|y + t \cdot (z - y) - x\| &\leq_{14.89} (1-t) \cdot \|y - x\| + t \cdot \|z - x\| \\ &< (1-t) \cdot \delta + t \cdot \delta \\ &= \delta \end{aligned}$$

proving that  $y + t \cdot (z - y) \in B_{\|\cdot\|}(x, \delta)$ . If  $y, z \in \bar{B}_{\|\cdot\|}(x, \delta)$  then

$$\begin{aligned} \|y + t \cdot (z - y) - x\| &\leq_{14.89} (1-t) \cdot \|y - x\| + t \cdot \|z - x\| \\ &\leq (1-t) \cdot \delta + t \cdot \delta \\ &= \delta \end{aligned}$$

proving that  $y + t \cdot (z - y) \in \bar{B}_{\|\cdot\|}(x, \delta)$ .  $\square$

**Theorem 14.111. (Mean Value Theorem (generalized))** Let  $\langle X, \|\cdot\|_X \rangle$  be a normed space over  $\mathbb{R}$  and let  $\langle Y, \|\cdot\|_Y \rangle$  be a Banach space over  $\mathbb{R}$ ,  $V \subseteq Y$ ,  $U \subseteq X$  a open set and  $f: U \rightarrow V$  a  $C^1$  function. Assume that there is a convex subset  $C \subseteq U$  and a constant  $k \in \mathbb{R}$  such that  $\forall x \in C$  we have  $\|Df(x)\|_{L(X, Y)} \leq k$  then  $\forall x, y \in C$  we have  $\|f(y) - f(x)\|_Y \leq k \cdot \|y - x\|_X$ .

**Proof.** Let  $x, y \in C$ . Define  $\sigma: \mathbb{R} \rightarrow X$  by  $\sigma(t) = x + t \cdot (y - x)$  then by definition of a convex set we have  $\sigma([0, 1]) \subseteq C \subseteq U \Rightarrow [0, 1] \subseteq \sigma^{-1}(U)$  [open because  $\sigma$  is continuous]. Define then  $g = f \circ \sigma: \sigma^{-1}(U) = \sigma^{-1}(U) \cap \mathbb{R} = \sigma^{-1}(U) \rightarrow Y$  which is of class  $C^1$  [as  $\sigma$  is of class  $C^1$  by 14.60,  $f$  is of class  $C^1$  together with 14.83]. Using the chain rule [see 14.22] we have  $\forall t \in f^{-1}(V)$  that  $Dg(t) = Df(\sigma(t)) \circ D\sigma(t)$  so that

$$\begin{aligned} g'(t) &= (Dg(t))(1) \\ &= (Df(\sigma(t)) \circ D\sigma(t))(1) \\ &= Df(\sigma(t))(D\sigma(t)(1)) \\ &= Df(\sigma(t))(\sigma'(t)) \\ &\stackrel{14.94}{=} Df(\sigma(t))(y - x) \end{aligned} \tag{14.90}$$

Next if we take  $h = g|_{[0,1]}: [0, 1] \rightarrow Y$  then using 14.93 we have that  $h$  is  $C^1$  and  $h' = ((g')|_{[0,1]}(t))$ . So if  $t \in [0, 1]$  then

$$\begin{aligned} \|h'(t)\|_Y &= \|((g')|_{[0,1]}(t))\|_Y \\ &= \|g'(t)\|_Y \\ &\stackrel{14.90}{=} \|Df(\sigma(t))(y - x)\|_Y \\ &\leq \|Df(\sigma(t))\|_{L(X, Y)} \cdot \|y - x\|_X \\ &\leq k \cdot \|y - x\|_Y \end{aligned} \tag{14.91}$$

$$\begin{aligned}
\|f(y) - f(x)\|_Y &= \|f(x + 1 \cdot (y - x)) + f(x + 0 \cdot (y - x))\|_Y \\
&= \|(f \circ \sigma)(1) - (f \circ \sigma)(0)\|_Y \\
&= \|h(1) - h(0)\|_Y \\
&\stackrel{14.107}{=} \left\| \int_0^1 h' \right\|_Y \\
&\stackrel{11.445 \text{ and } (14.91)}{\leq} (1 - 0) \cdot k \cdot \|y - x\|_Y \\
&= k \cdot \|y - x\|_Y
\end{aligned}$$

which proves the theorem.  $\square$

**Definition 14.112.** Let  $X$  be a vector space,  $a, b \in X$  then the line segment connecting  $a$  to  $b$  is the set  $L_{a,b} = \{t \cdot a + (1-t) \cdot b \mid t \in [0, 1]\} = \{b + t \cdot (a - b)\} \subseteq X$ . Note that  $a, b \in L_{a,b}$  [take  $t = 0, 1$ ]

**Example 14.113.**  $[0, 1] = L_{0,1}$  in  $\mathbb{R}$

**Proof.** If  $x \in L_{0,1}$  then there exists a  $t \in [0, 1]$  such that  $x = t \cdot 0 + (1-t) \cdot 1 = (1-t) \Rightarrow 0 \leq 1-t \leq 1 \Rightarrow L_{0,1} \subseteq [0, 1]$ . If  $x \in [0, 1]$  then  $0 \leq x \leq 1 \Rightarrow 0 \leq 1-x \leq 1 \Rightarrow x = (1-x) \cdot 0 + (1-(1-x)) \cdot 1 \in L_{0,1} \Rightarrow x \in L_{0,1} \Rightarrow [0, 1] \subseteq L_{0,1}$ .  $\square$

**Lemma 14.114.** Let  $\langle X, \|\cdot\| \rangle$  be a normed space then  $a, b \in X$  then  $L_{a,b}$  is convex, compact and bounded.

**Proof.** If  $x, y \in L_{a,b}$  then there exists  $t_x, t_y \in [0, 1]$  such that  $x = t_x \cdot a + (1-t_x) \cdot b$ ,  $y = t_y \cdot a + (1-t_y) \cdot b$ . Then we have for  $t \in [0, 1]$

$$\begin{aligned}
x + t \cdot (y - x) &= t_x \cdot a + (1-t_x) \cdot b + t \cdot (t_y \cdot a + (1-t_y) \cdot b - t_x \cdot a - (1-t_x) \cdot b) \\
&= t_x \cdot a + b - t_x \cdot b + t \cdot t_y \cdot a + t \cdot b - t \cdot t_y \cdot b - t \cdot t_x \cdot a - t \cdot b + t \cdot t_x \cdot b \\
&= (t_x + t \cdot t_y - t \cdot t_x) \cdot a + (1-t_x + t - t \cdot t_y - t + t \cdot t_x) \cdot b \\
&= (t_x + t \cdot t_y - t \cdot t_x) \cdot a + (1-t_x - t \cdot t_y + t \cdot t_x) \cdot b \\
&= (t_x + t \cdot (t_y - t_x)) \cdot a + (1 - (t_x + t \cdot (t_y - t_x))) \cdot b \\
&= s \cdot a + (1-s) \cdot b \quad [\text{where } s = t_x + t \cdot (t_y - t_x) = t_x - t \cdot t_x + t \cdot t_y]
\end{aligned}$$

Now as  $0 \leq t, t_x, t_y \leq 1$  we have that  $t_x \cdot t \leq t_x \Rightarrow 0 \leq t_x - t_x \cdot t \Rightarrow 0 \leq t_x - t \cdot t_x + t \cdot t_y = t_x + t \cdot (t_y - t_x) \leq t_x + (t_y - t_x) = t_y \leq 1$  proving that  $s \in [0, 1]$  and thus that  $x + t \cdot (y - x) \in L_{a,b}$ . So

$L_{a,b}$  is convex

Define  $\sigma: [0, 1] \rightarrow X$  by  $\sigma(t) = t \cdot a + (1-t) \cdot b$ . Given  $\varepsilon > 0$ ,  $s \in [0, 1]$  take  $\delta = \frac{\varepsilon}{\|a - b\|_Y + 1} > 0$  then if  $|s - t| < \delta$  and  $t \in [0, 1]$  we have

$$\begin{aligned}
\|\sigma(t) - \sigma(s)\|_Y &= \|t \cdot a + (1-t) \cdot b - s \cdot a - (1-s) \cdot b\|_Y \\
&= \|(t-s) \cdot a + (s-t) \cdot b\|_Y \\
&= |t-s| \cdot \|a-b\|_Y < \\
&< \delta \cdot \|a-b\|_Y \\
&< \varepsilon
\end{aligned}$$

proving that

$\sigma$  is continuous

As clearly  $L_{a,b} = \sigma([0, 1])$  we have using 11.253 and 11.267 that  $L_{a,b} = \sigma([0, 1])$  is compact, and by 11.251 that  $L_{a,b}$  is bounded.  $\square$

**Corollary 14.115.** Let  $\langle X, \|\cdot\|_X \rangle$  be a normed space over  $\mathbb{R}$  and let  $\langle Y, \|\cdot\|_Y \rangle$  be a Banach space over  $\mathbb{R}$ ,  $V \subseteq Y$ ,  $U \subseteq X$  a open set,  $x, y \in X$  such that  $L_{x,y} \subseteq U$  and  $f: U \rightarrow V$  a  $C^1$  function then  $\sup(\{\|Df(\xi)\|_{L(X,Y)} \mid \xi \in L_{x,y}\})$  exists and  $\|f(x) - f(y)\|_Y \leq \sup(\{\|Df(\xi)\|_{L(X,Y)} \mid \xi \in L_{x,y}\}) \cdot \|x - y\|_X$

**Proof.** As  $L_{x,y}$  is compact by the previous lemma and  $Df$  is continuous [as  $f$  is  $C^1$ ] we have that  $Df(L_{x,y})$  is compact [see 11.253] and thus by 11.251 bounded. So there exists a  $M \geq 0$  such that  $\forall L_1, L_2 \in Df(L_{x,y})$  we have  $\|L_1 - L_2\| \leq M$ . So if  $\xi \in L_{x,y}$  we have  $\|Df(\xi)\|_{L(X,Y)} = \|Df(\xi) - Df(x) + Df(x)\|_{L(X,Y)} \leq \|Df(\xi) - Df(x)\|_{L(X,Y)} + \|Df(x)\|_{L(X,Y)} \leq M + \|Df(x)\|_{L(X,Y)}$  proving that  $\{\|Df(\xi)\|_{L(X,Y)} | \xi \in L_{x,y}\}$  is bounded above and thus by 9.43 that  $\sup(\{\|Df(\xi)\|_{L(X,Y)} | \xi \in L_{x,y}\})$  exists and is finite. Finally as  $L_{x,y}$  is convex,  $x, y \in L_{x,y}$  and  $L_{x,y} \subseteq U$  we have by the Mean Value Theorem (14.111) that  $\|f(x) - f(y)\|_Y \leq \sup(\{\|Df(\xi)\|_{L(X,Y)} | \xi \in L_{x,y}\}) \cdot \|x - y\|_X$ .  $\square$

**Corollary 14.116.** Let  $\langle X, \|\cdot\|_X \rangle$  be a normed space,  $\langle Y, \|\cdot\|_Y \rangle$  be a Banach space, a open set  $U \subseteq X$ , a open ball  $B_{\|\cdot\|_X}(x, \delta) \subseteq U$  and a function  $f: U \rightarrow Y$  a function such that  $\forall y \in B_{\|\cdot\|_X}(x, \delta)$  we have that  $f$  is differentiable at  $y$  then  $\forall y, z \in B_{\|\cdot\|_X}(x, \delta)$  we have that  $\sup(\{\|Df(y + t \cdot (z - y))\|_{L(X,Y)} | t \in [0, 1]\})$  exists and

$$\|f(z) - f(y)\|_Y \leq \|z - y\|_X \cdot \sup(\{\|Df(y + t \cdot (z - y))\|_{L(X,Y)} | t \in [0, 1]\})$$

**Proof.** Let  $y, z \in B_{\|\cdot\|_X}(x, \delta)$  then by 14.110 we have that  $L_{z,y} = \{y + t \cdot (z - y) | t \in [0, 1]\} \subseteq B_{\|\cdot\|_X}(x, \delta) \subseteq U$  so that using 14.115  $\sup(\{\|Df(\xi)\|_{L(X,Y)} | \xi \in L_{z,y}\})$  exists and  $\|f(z) - f(y)\|_Y \leq \sup(\{\|Df(\xi)\|_{L(X,Y)} | \xi \in L_{z,y}\}) \cdot \|z - y\|_X$ . As

$$\begin{aligned} r \in \{\|Df(\xi)\|_{L(X,Y)} | \xi \in L_{y,z}\} &\Leftrightarrow \exists \xi \in L_{y,z} \text{ with } r = \|Df(\xi)\|_{L(X,Y)} \\ &\Leftrightarrow \exists t \in [0, 1] \text{ with } r = \|Df(y + t \cdot (z - y))\|_{L(X,Y)} \\ &\Leftrightarrow \{\|Df(y + t \cdot (z - y))\|_{L(X,Y)} | t \in [0, 1]\} \end{aligned}$$

it follows that  $\sup(\{\|Df(y + t \cdot (z - y))\|_{L(X,Y)} | t \in [0, 1]\})$  exists and  $\|f(z) - f(y)\|_Y \leq \sup(\{\|Df(y + t \cdot (z - y))\|_{L(X,Y)} | t \in [0, 1]\}) \cdot \|z - y\|_X$ .  $\square$

**Theorem 14.117. (Second Mean Value Theorem)** Let  $\langle X, \|\cdot\|_X \rangle$  be a normed space over  $\mathbb{R}$ ,  $\langle Y, \|\cdot\|_Y \rangle$  a Banach space over  $\mathbb{R}$ ,  $V \subseteq Y$ ,  $a, b \in X$ ,  $U \subseteq X$  a open set such that  $L_{a,b} \subseteq U$  and  $f: U \rightarrow V$  a  $C^1$  function then  $\forall x \in U$  we have  $\|f(a) - f(b) - Df(x)(a - b)\|_Y \leq \|a - b\|_X \cdot \sup(\{\|Df(\xi) - Df(x)\|_{L(X,Y)} | \xi \in L_{a,b}\})$ .

**Proof.** Define  $g: U \rightarrow Y$  by  $g(\xi) = f(\xi) - Df(x)(\xi)$  which is  $C^1$  as  $g$  is  $C^1$  and  $Df(x)$  is  $C^\infty$  [as it is a continuous linear function]. Further if  $\xi \in U$  then  $Dg(\xi) = Df(\xi) - Df(x)$  [as  $Df(x)$  is linear and continuous] Now using the Mean Value Theorem (see 14.115) we have  $\|g(b) - g(a)\|_Y \leq \sup(\{\|Dg(\xi)\|_{L(X,Y)} | \xi \in L_{a,b}\}) \cdot \|b - a\|_X \Rightarrow \|f(b) - Df(x)(b) - f(a) + Df(x)(a)\|_Y \leq \sup(\{\|Df(\xi) - Df(x)\|_{L(X,Y)} | \xi \in L_{a,b}\}) \Rightarrow \|f(b) - f(a) - Df(x)(b - a)\|_Y \leq \sup(\{\|Df(\xi) - Df(x)\|_{L(X,Y)} | \xi \in L_{a,b}\})$ .  $\square$

There is another example of a theorem named Mean value theorem (see Dieudonne), first we need a lemma.

**Lemma 14.118.** Let  $\langle Y, \|\cdot\| \rangle$  be a normed space,  $V \subseteq Y$ ,  $I = [a, b] \subseteq \mathbb{R}$  a non empty interval and  $f: I \rightarrow V$ ,  $x \in [a, b]$  such that  $f'(x)$  exists [see 14.92] then given  $\varepsilon > 0$  there exists a  $\lambda > 0$  such that if  $x \leq \xi < \lambda$  then  $\|f(\xi) - f(x) - f'(x) \cdot (\xi - x)\| \leq \varepsilon \cdot (\xi - x)$

**Proof.** Let  $\varepsilon > 0$ . Then by definition [see 14.92] we have that  $f'_+(x)$  exists so there exists a  $\delta > 0$  such that if  $0 < h < \delta$  and  $x + h \in [a, b]$  we have  $\left\| \frac{f(x+h) - f(x)}{h} - f'_+(x) \right\| < \varepsilon$ . Take then  $\lambda = x + \delta$  and let  $x \leq \xi < \lambda$  then we have either:

$\xi = x$ . then

$$\begin{aligned} \|f(\xi) - f(x) - f'_+(x) \cdot (\xi - x)\| &= \|f(x) - f(x) - f'_+(x) \cdot (x - x)\| \\ &= 0 \\ &\leq \varepsilon \cdot 0 \\ &= \varepsilon \cdot (\xi - x) \end{aligned}$$

$x < \xi$ . then  $0 < \xi - x < \lambda - x = \delta$  so that

$$\begin{aligned} & \left\| \frac{f(x + (\xi - x)) - f(x)}{(\xi - x)} - f'_+(x) \right\| < \varepsilon \Rightarrow \\ & \left\| \frac{f(\xi) - f(x)}{\xi - x} - f'_+(x) \right\| \leq \varepsilon \quad 0 < \xi - x \\ & \left\| (\xi - x) \cdot \left( \frac{f(\xi) - f(x)}{\xi - x} - f'_+(x) \right) \right\| \leq \varepsilon \cdot (\xi - x) \Rightarrow \\ & \|f(\xi) - f(x) - f'_+(x) \cdot (\xi - x)\| \leq \varepsilon \cdot (\xi - x) \end{aligned}$$

Hence we have in both cases that

$$\|f(\xi) - f(x) - f'(x) \cdot (\xi - x)\| \leq \varepsilon \cdot (\xi - x) \quad \square$$

**Theorem 14.119.** Let  $I = [a, b] \subseteq \mathbb{R}$  be a non empty interval,  $\langle Y, \|\cdot\| \rangle$  a normed space,  $V \subseteq Y$ ,  $f: I \rightarrow V$ ,  $\varphi: I \rightarrow \mathbb{R}$  continuous functions such that there exists a denumerable  $E \subseteq [a, b]$  such that  $\forall x \in [a, b] \setminus E$   $f$  and  $\varphi$  have derivatives  $f'(x)$  and  $\varphi'(x)$  [see 14.92] with  $\|f'(x)\| \leq \varphi'(x)$  then  $\|f(b) - f(a)\| \leq \varphi(b) - \varphi(a)$

**Proof.** As  $E$  is denumerable there exists a bijection  $\rho: \mathbb{N} \rightarrow E$ . Take now  $\varepsilon > 0$  and define  $\xi(\beta)$  and  $A$  as follows:

$$\forall \beta \in [a, b] \quad \mathcal{E}(\beta) = \begin{cases} 0 & \text{if } \{i \in \mathbb{N} \mid \rho(i) < \beta\} = \emptyset \\ \sum_{n \in \{i \in \mathbb{N} \mid \rho(i) < \beta\}} \frac{1}{2^n} & \text{if } \{i \in \mathbb{N} \mid \rho(i) < \beta\} \neq \emptyset \end{cases} \quad (14.92)$$

[note that by 9.56  $\{i \in \mathbb{N} \mid \sigma(i) < \beta\}$  is finite as there exists a  $n \in \mathbb{N}_0$  such that  $\beta \leq n$  hence  $\{i \in \mathbb{N}_0 \mid \sigma(i) < \beta\} \subseteq \{1, \dots, n\}$ ] and

$$A_\varepsilon = \{\beta \in [a, b] \mid \forall \gamma \in [a, \beta] \models \|f(\gamma) - f(a)\| \leq \varphi(\gamma) - \varphi(a) + \varepsilon \cdot (\gamma - a) + \varepsilon \cdot \mathcal{E}(\gamma)\} \subseteq [a, b] \quad (14.93)$$

Then as  $\|f(a) - f(a)\| = 0 = \varphi(a) - \varphi(a)$ ,  $\varepsilon \cdot (a - a) = 0$  and  $\mathcal{E}(a) = 0$  [as  $\rho(\mathbb{N}) = E \subseteq [a, b] \Rightarrow \forall i \in \mathbb{N} a \leq \rho(i)$  so that  $\{i \in \mathbb{N} \mid \rho(i) < a\} = \emptyset$ ] it follows that

$$a \in A_\varepsilon \Rightarrow A_\varepsilon \neq \emptyset \quad (14.94)$$

Also if  $\beta \in A_\varepsilon$  then if  $\xi \in [a, \beta]$  we have by the definition of  $A_\varepsilon$  that also  $\xi \in A_\varepsilon$  proving together with  $\beta \in A_\varepsilon$  that

$$\forall \beta \in A_\varepsilon \text{ we have } [a, \beta] \subseteq A_\varepsilon \quad (14.95)$$

Take now

$$\sigma = \sup(A_\varepsilon)$$

[which exists because from  $A_\varepsilon \subseteq [a, b]$  it follows that  $A_\varepsilon$  is bounded above by  $b$ ,  $A_\varepsilon \neq \emptyset$  and the fact that  $\mathbb{R}$  is conditional complete [see 9.43]] then if  $\xi \in [a, \sigma]$  there exists a  $\beta \in A_\varepsilon$  such that  $a \leq \xi < \beta \leq \sigma \Rightarrow [a, \xi] \subseteq A_\varepsilon \Rightarrow \xi \in A_\varepsilon$  proving that

$$[a, \sigma] \subseteq A_\varepsilon \quad (14.96)$$

Further for the  $\beta$  and  $\xi$  we have from  $\beta \in A_\varepsilon$  and  $\xi \in [a, \beta]$  that  $\|f(\xi) - f(a)\| \leq \varphi(\xi) - \varphi(a) + \varepsilon \cdot (\xi - a) + \varepsilon \cdot \mathcal{E}(\xi)$  proving as we have chosen  $\xi \in [a, \sigma]$  that also  $\sigma \in A_\varepsilon$ . So we have that

$$[a, \sigma] \subseteq A_\varepsilon \quad (14.97)$$

If now  $x \in A_\varepsilon \setminus [a, \sigma]$  then as  $A_\varepsilon \subseteq [a, b]$  we have  $a \leq x$  and thus must have  $\sigma < x$  contradicting the fact that  $\sigma$  as a supremum is an upper bound of  $A_\varepsilon$  so that  $A_\varepsilon \setminus [a, \sigma] = \emptyset$ . Hence if  $x \in A_\varepsilon$  then we must have that  $x \in [a, \sigma]$  proving that  $A_\varepsilon \subseteq [a, \sigma]$  which together with 14.97 gives

$$[a, \sigma] = A_\varepsilon \text{ where } \sigma = \sup(A_\varepsilon) \quad (14.98)$$

Now if  $\alpha, \beta \in [a, b]$  with  $\alpha < \beta$  then if  $n \in \{i \in \mathbb{N} \mid \rho(i) < \alpha\}$  we have that  $\rho(n) < \alpha < \beta \Rightarrow n \in \{i \in \mathbb{N} \mid \rho(i) < \beta\}$  proving that  $\{i \in \mathbb{N} \mid \rho(i) < \alpha\} \subseteq \{i \in \mathbb{N} \mid \rho(i) < \beta\}$ . We must now consider two cases

$\{i \in \mathbb{N} \mid \rho(i) < \alpha\} = \emptyset$ . then  $\xi(\alpha) = 0 \leq \xi(\beta)$

$$\{i \in \mathbb{N} \mid \rho(i) < \alpha\} \neq \emptyset. \text{ then } \{i \in \mathbb{N} \mid \rho(i) < \beta\} \neq \emptyset \text{ and } \xi(\alpha) = \sum_{n \in \{i \in \mathbb{N} \mid \rho(i) < \alpha\}} \frac{1}{2^n} \leq \sum_{n \in \{i \in \mathbb{N} \mid \rho(i) < \beta\}} \frac{1}{2^n} = \xi(\beta)$$

so that  $\mathcal{E}(\alpha) \leq \mathcal{E}(\beta)$  proving

$$\forall \alpha, \beta \in [a, b] \text{ with } \alpha \leq \beta \text{ we have } \mathcal{E}(\alpha) \leq \mathcal{E}(\beta) \quad (14.99)$$

Now let  $\xi > 0$ , chosen arbitrary, then by the continuity of  $\varphi$  there exists a  $\delta_\varphi > 0$  such that if  $\beta \in [\sigma - \delta_\varphi, \sigma + \delta_\varphi] \cap [a, b]$  then  $|\varphi(\beta) - \varphi(\sigma)| < \frac{\xi}{2}$ . So if  $\beta \in [a, \sigma] \cap [\sigma - \delta_\varphi, \sigma + \delta_\varphi] \cap [a, b]$  then  $|\varphi(\beta) - \varphi(\sigma)| < \frac{\xi}{2} \Rightarrow \varphi(\beta) - \varphi(a) = \varphi(\sigma) - \varphi(a) + \varphi(\beta) - \varphi(\sigma) \leq \varphi(\sigma) - \varphi(a) + |\varphi(\beta) - \varphi(\sigma)| < \varphi(\sigma) - \varphi(a) + \frac{\xi}{2}$ . Hence

$$\forall \beta \in [a, \sigma] \cap [\sigma - \delta_\varphi, \sigma] \text{ we have } \varphi(\beta) - \varphi(a) < \varphi(\sigma) - \varphi(a) + \frac{\xi}{2} \quad (14.100)$$

As  $f$  is continuous there exists a  $\delta_f > 0$  such that if  $\beta \in [\sigma - \delta_f, \sigma + \delta_f] \cap [a, b]$  then  $\|f(\sigma) - f(\beta)\| < \frac{\xi}{2}$ . So if  $\beta \in [a, \sigma] \cap [\sigma - \delta_f, \sigma]$  then  $\beta \in [\sigma - \delta_f, \sigma + \delta_f] \cap [a, b]$  so that  $\|f(\beta) - f(a)\| < \frac{\xi}{2}$  and

$$\begin{aligned} \|f(\sigma) - f(a)\| &\leq \|f(\beta) - f(a) + f(\sigma) - f(\beta)\| \\ &\leq \|f(\beta) - f(a)\| + \|f(\sigma) - f(\beta)\| \\ &< \|f(\beta) - f(a)\| + \frac{\xi}{2} \\ &\leq_{\text{def. of } A_\varepsilon, \sigma \in A_\varepsilon, a \leq \beta < \sigma} \varphi(\beta) - \varphi(a) + \varepsilon \cdot (\beta - \alpha) + \varepsilon \cdot \mathcal{E}(\beta) + \frac{\xi}{2} \end{aligned} \quad (14.101)$$

So if  $\beta \in [a, \sigma] \cap [\sigma - \min(\delta_\varphi, \delta_f), \sigma]$  then

$$\begin{aligned} \|f(\sigma) - f(a)\| &<_{(14.101)} \varphi(\beta) - \varphi(a) + \varepsilon \cdot (\beta - \alpha) + \varepsilon \cdot \mathcal{E}(\beta) + \frac{\xi}{2} \\ &<_{(14.100)} \varphi(\sigma) - \varphi(a) + \frac{\xi}{2} + \varepsilon \cdot (\beta - \alpha) + \varepsilon \cdot \mathcal{E}(\beta) + \frac{\xi}{2} \\ &\leq_{(14.99) \wedge \beta < \sigma} \varphi(\sigma) - \varphi(a) + \varepsilon \cdot (\beta - \alpha) + \varepsilon \cdot \mathcal{E}(\sigma) + \xi \\ &<_{\beta < \sigma} \varphi(\sigma) - \varphi(a) + \varepsilon \cdot (\sigma - \alpha) + \varepsilon \cdot \mathcal{E}(\sigma) + \xi \end{aligned}$$

As  $\xi$  was chosen arbitrary we have using 9.57 that

$$\|f(\sigma) - f(a)\| \leq \varphi(\sigma) - \varphi(a) + \varepsilon \cdot (\sigma - a) + \varepsilon \cdot \mathcal{E}(\sigma) \quad (14.102)$$

Suppose now that  $\sigma < b$  then we have for  $\sigma$  to possibilities in relation to  $E$

**$\sigma \notin E$ .** Then as  $\varphi, f$  have a derivate at  $\sigma$  and  $\sigma < b$  we have by 14.118 that there exists  $\lambda_\varphi, \lambda_f > 0$  such that for  $\sigma \leq \xi < \lambda_\varphi < b$  we have  $|\varphi(\xi) - \varphi(\sigma) - \varphi'(\sigma) \cdot (\xi - \sigma)| \leq \frac{\varepsilon}{2} \cdot (\xi - \sigma)$  and for  $\sigma \leq \xi < \lambda_f < b$  we have  $\|f(\xi) - f(\sigma) - f'(\sigma) \cdot (\xi - \sigma)\| \leq \frac{\varepsilon}{2} \cdot (\xi - \sigma)$  so that if  $\sigma \leq \xi < \min(\lambda_\varphi, \lambda_f) < b$  we have

$$\begin{aligned} \varphi'(\sigma) \cdot (\xi - \sigma) &= (\varphi'(\sigma) \cdot (\xi - \sigma) - (\varphi(\xi) - \varphi(\sigma)) + (\varphi(\xi) - \varphi(\sigma))) \\ &\leq |\varphi'(\sigma) \cdot (\xi - \sigma) - (\varphi(\xi) - \varphi(\sigma))| + \varphi(\xi) - \varphi(\sigma) \\ &= |\varphi(\xi) - \varphi(\sigma) - \varphi'(\sigma) \cdot (\xi - \sigma)| + \varphi(\xi) - \varphi(\sigma) \\ &\leq \varphi(\xi) - \varphi(\sigma) + \frac{\varepsilon}{2} \cdot (\xi - \sigma) \\ \|f(\xi) - f(\sigma)\| &= \|f(\xi) - f(\sigma) - f'(\sigma) \cdot (\xi - \sigma) + f'(\sigma) \cdot (\xi - \sigma)\| \\ &\leq \|f(\xi) - f(\sigma) - f'(\sigma) \cdot (\xi - \sigma)\| + \|f'(\sigma) \cdot (\xi - \sigma)\| \\ &\leq \frac{\varepsilon}{2} \cdot (\xi - \sigma) + \|f'(\sigma)\| \cdot (\xi - \sigma) \\ &\leq \frac{\varepsilon}{2} \cdot (\xi - \sigma) + \varphi'(\sigma) \cdot (\xi - \sigma) \text{ (by the hypothesis of the theorem)} \\ &\leq \frac{\varepsilon}{2} \cdot (\xi - \sigma) + \varphi(\xi) - \varphi(\sigma) + \frac{\varepsilon}{2} \cdot (\xi - \sigma) \\ &= \varphi(\xi) - \varphi(\sigma) + \varepsilon \cdot (\xi - \sigma) \end{aligned}$$

proving

$$\|f(\xi) - f(\sigma)\| \leq \varphi(\xi) - \varphi(\sigma) + \varepsilon \cdot (\xi - \sigma) \quad (14.103)$$

Next

$$\begin{aligned}
\|f(\xi) - f(a)\| &= \|f(\xi) - f(\sigma) + f(\sigma) - f(a)\| \\
&\leq \|f(\xi) - f(\sigma)\| + \|f(\sigma) - f(a)\| \\
&\stackrel{\text{14.103}}{\leq} \varphi(\xi) - \varphi(\sigma) + \varepsilon \cdot (\xi - \sigma) + \|f(\sigma) - f(a)\| \\
&\stackrel{\text{14.102}}{\leq} \varphi(\xi) - \varphi(\sigma) + \varepsilon \cdot (\xi - \sigma) + \varphi(\sigma) - \varphi(a) + \varepsilon \cdot (\sigma - a) + \varepsilon \cdot \mathcal{E}(\sigma) \\
&= \varphi(\xi) - \varphi(a) + \varepsilon \cdot (\xi - a) + \varepsilon \cdot \mathcal{E}(\sigma) \\
&\stackrel{\text{14.99 and } \sigma \leq \xi}{\leq} \varphi(\xi) - \varphi(a) + \varepsilon \cdot (\xi - a) + \varepsilon \cdot \mathcal{E}(\xi)
\end{aligned}$$

But this proves that  $\min(\lambda_\varphi, \lambda_f) \in A_\varepsilon$  which is a contradiction as  $\sup(A_\varepsilon) = \sigma < \min(\lambda_\varphi, \lambda_f)$ . So this case leads to a contradiction.

$\sigma \in E$ . then there exists a  $m \in \mathbb{N}$  such that  $\sigma = \rho(m)$ . By the continuity of  $f, \varphi$  at  $\sigma$  together with  $\sigma < b$  there exists  $\mu_\varphi, \mu_f$  such that if  $\sigma < \xi < \mu_\varphi < b$  then  $|\varphi(\xi) - \varphi(\sigma)| \leq \frac{\varepsilon}{2} \cdot \frac{1}{2^m}$  and if  $\sigma < \xi < \mu_f < b$  then  $\|f(\xi) - f(\sigma)\| < \frac{\varepsilon}{2} \cdot \frac{1}{2^m}$ . So if  $\sigma < \xi < \min(\mu_\varphi, \mu_f) < b$  then we have

$$\begin{aligned}
\|f(\xi) - f(a)\| &= \|f(\xi) - f(\sigma) + f(\sigma) - f(a)\| \\
&\leq \|f(\xi) - f(\sigma)\| + \|f(\sigma) - f(a)\| \\
&\leq \frac{\varepsilon}{2} \cdot \frac{1}{2^m} + \|f(\sigma) - f(a)\| \\
&\stackrel{\text{14.102}}{\leq} \frac{\varepsilon}{2} \cdot \frac{1}{2^m} + \varphi(\sigma) - \varphi(a) + \varepsilon \cdot (\sigma - a) + \varepsilon \cdot \mathcal{E}(\sigma) \\
&= \frac{\varepsilon}{2} \cdot \frac{1}{2^m} + \varphi(\xi) - \varphi(a) + \varphi(\sigma) - \varphi(\xi) + \varepsilon \cdot (\sigma - a) + \varepsilon \cdot \mathcal{E}(\sigma) \\
&\leq \frac{\varepsilon}{2} \cdot \frac{1}{2^m} + \varphi(\xi) - \varphi(a) + |\varphi(\sigma) - \varphi(\xi)| + \varepsilon \cdot (\sigma - a) + \varepsilon \cdot \mathcal{E}(\sigma) \\
&\leq \frac{\varepsilon}{2} \cdot \frac{1}{2^m} + \varphi(\xi) - \varphi(a) + \frac{\varepsilon}{2} \cdot \frac{1}{2^m} + \varepsilon \cdot (\sigma - a) + \varepsilon \cdot \mathcal{E}(\sigma) \\
&= \varphi(\xi) - \varphi(a) + \varepsilon \cdot (\sigma - a) + \varepsilon \cdot \frac{1}{2^m} + \varepsilon \cdot \sum_{n \in \{i \in \mathbb{N}_0 \mid \rho(i) < \sigma\}} \frac{1}{2^n} \\
&\stackrel{\sigma = \rho(m)}{=} \varphi(\xi) - \varphi(a) + \varepsilon \cdot (\sigma - a) + \varepsilon \cdot \sum_{n \in \{i \in \mathbb{N}_0 \mid \rho(i) \leq \sigma\}} \frac{1}{2^n} \\
&\stackrel{\xi < \sigma}{\leq} \varphi(\xi) - \varphi(a) + \varepsilon \cdot (\sigma - a) + \varepsilon \cdot \sum_{n \in \{i \in \mathbb{N}_0 \mid \rho(i) < \xi\}} \frac{1}{2^n} \\
&\stackrel{\rho(m) = \sigma < \xi \Rightarrow \{i \in \mathbb{N}_0 \mid \rho(i) < \xi\} \neq \emptyset}{\leq} \varphi(\xi) - \varphi(a) + \varepsilon \cdot (\xi - a) + \varepsilon \cdot \mathcal{E}(\xi)
\end{aligned}$$

proving again that  $\min(\mu_\varphi, \mu_f) \in A_\varepsilon$  which is a contradiction as  $\sup(A_\varepsilon) = \sigma < \min(\mu_\varphi, \mu_f)$

As in all cases we have a contradiction we must have that  $\sigma = b$  so by (14.94)  $b \in A_\varepsilon$ . Hence by the definition of  $A$  we have  $\|f(b) - f(a)\| \leq \varphi(b) - \varphi(a) + \varepsilon \cdot (\sigma - a) + \varepsilon \cdot \mathcal{E}(b)$ . As  $\varepsilon > 0$  is arbitrary chosen we have by 9.57 that

$$\|f(b) - f(a)\| \leq \varphi(b) - \varphi(a) \quad \square$$

**Corollary 14.120.** Let  $[a, b] \subseteq \mathbb{R}$  be a non empty interval,  $\langle Y, \|\cdot\|\rangle$  a normed space,  $f: [a, b] \rightarrow Y$  a continuous function such that there exists a denumerable set  $E \subseteq [a, b]$  so that  $\forall x \in [a, b] \setminus E$   $f'(x)$  exists [see 14.92] and  $\|f'(x)\| \leq M$  then  $\|f(b) - f(a)\| \leq M \cdot (b - a)$

**Proof.** Define  $\varphi: \mathbb{R} \rightarrow \mathbb{R}$  by  $\varphi(x) = M \cdot (x - a)$  then  $\varphi'(x) = M$  so that by 14.91, 14.92  $\varphi|_{[a, b]}: [a, b] \rightarrow \mathbb{R}$  has  $\forall x \in [a, b]$  the derivate  $\varphi'(x) = M$ . So  $\forall x \in [a, b] \setminus E$  we have that  $\|f'(x)\| \leq \varphi'(x)$  and thus by the previous theorem [see 14.119] we have  $\|f(b) - f(a)\| \leq \varphi(b) - \varphi(a) = M \cdot (b - a) - M \cdot (b - a) = M \cdot (b - a)$   $\square$

**Corollary 14.121.** Let  $[a, b] \subseteq \mathbb{R}$  be a non empty interval,  $\langle Y, \|\cdot\| \rangle$  a normed space,  $V \subseteq Y$  and  $f: [a, b] \rightarrow V$  a continuous function so that  $\forall x \in [a, b]$   $f'(x)$  exists [see 14.92] and  $\|f'(x)\| \leq M$  then  $\|f(b) - f(a)\| \leq M \cdot (b - a)$

**Proof.** As  $\forall n \in \mathbb{N}$  we have that  $0 < \frac{1}{n+1} < 1 \underset{0 < b-a}{\Rightarrow} 0 < (b-a) \cdot \frac{1}{n+1} < (b-a) \Rightarrow a < a + (b-a) \cdot \frac{1}{n+1} < b$  we have that  $E = \left\{ a + (b-a) \cdot \frac{1}{n+1} \mid n \in \mathbb{N} \right\} \subseteq [a, b]$  is a denumerable set and for  $\forall x \in [a, b] \setminus E$  we have that  $f'(x)$  exists [as it exists on  $[a, b]$ ] and  $\|f'(x)\| \leq M$  so that by the previous corollary we have that  $\|f(b) - f(a)\| \leq M \cdot (b - a)$   $\square$

## 14.5 Symmetry of Higher Order Differentials

**Lemma 14.122.** Let  $\langle X, \|\cdot\|_X \rangle$  be a normed space,  $\langle Y, \|\cdot\|_Y \rangle$  be a Banach space,  $V \subseteq Y$ ,  $U \subseteq X$  open and  $f: U \rightarrow V$  of class  $C^2$  then  $\forall x \in U$  and  $v, w \in X$  we have that  $D^2 f(x)(v)(w) = D^2 f(x)(w)(v)$

**Proof.** Take  $x \in U$  then as  $U$  is open there exists a  $r > 0$  such that  $x \in B_{\|\cdot\|_X}(x, r) \subseteq U$ . Take  $v, w \in X$  such that  $\|v\|_X < \frac{r}{4}$ ,  $\|w\|_X < \frac{r}{4}$  and define  $\varphi, \psi: \mathbb{R} \rightarrow X$  by  $\varphi(t) = x + t \cdot v$ ,  $\psi(t) = x + w + t \cdot v$ . If  $t \in ]-2, 2[$  then we have  $\|\varphi(t) - x\|_X = \|x + t \cdot v - x\|_X = |t| \cdot \|v\|_X \leq 2 \cdot \frac{r}{4} < r$  and  $\|\psi(t) - x\|_X = \|x + w + t \cdot v - x\|_X \leq |t| \cdot \|v\|_X + \|w\|_X < 2 \cdot \frac{r}{4} + \frac{r}{4} < r$  proving that

$$\varphi(]-2, 2[) \subseteq U \wedge \psi(]-2, 2[) \subseteq U \quad (14.104)$$

Define  $g_{v,w}: ]-2, 2[ \rightarrow Y$  by  $g_{v,w} = f \circ \psi|_{]-2, 2[} - f \circ \varphi|_{]-2, 2[}$  [which is well defined because of 14.104]. As  $\varphi, \psi$  are  $C^1$  [see 14.60 and 14.10] with

$$(\varphi|_{]-2, 2[})' = \varphi' = C_{\mathbb{R}, X, v} = \psi' = (\psi|_{]-2, 2[})' \quad (14.105)$$

we have by 14.83 that  $g_{v,w}$  is of class  $C^1$ . As  $L_{0,1} \stackrel{14.113}{=} [0, 1] \subseteq ]-2, 2[$  we can use the second mean value theorem [see 14.117] giving  $\|g_{v,w}(1) - g_{v,w}(0) - Dg_{v,w}(0)(1 - 0)\|_Y \leq \sup(\{\|Dg_{v,w}(\xi) - Dg_{v,w}(0)\|_{L(X,Y)} \mid \xi \in [0, 1]\})$ . As  $g'_{v,w}(0) = Dg_{v,w}(0)(1)$  we have

$$\|g_{v,w}(1) - g_{v,w}(0) - g'_{v,w}(0)\|_Y \leq \sup(\{\|Dg_{v,w}(\xi) - Dg_{v,w}(0)\|_{L(X,Y)} \mid \xi \in [0, 1]\}) \quad (14.106)$$

Now  $\forall h \in \mathbb{R}$  we have

$$\begin{aligned} \|(Dg_{v,w}(\xi) - Dg_{v,w}(0))(h)\|_Y &= \|Dg_{v,w}(\xi)(h) - Dg_{v,w}(0)(h)\|_Y \\ &\stackrel{14.19}{=} \|g'_{v,w}(\xi) \cdot h - g'_{v,w}(0) \cdot h\|_Y \\ &= \|(g'_{v,w}(\xi) - g'_{v,w}(0)) \cdot h\|_Y \\ &= \|g'_{v,w}(\xi) - g'_{v,w}(0)\|_Y \cdot |h| \end{aligned}$$

proving by the definition of the norm of a linear continuous function [see 11.185] that  $\|Dg_{v,w}(\xi) - Dg_{v,w}(0)\|_{L(X,Y)} \leq \|g'_{v,w}(\xi) - g'_{v,w}(0)\|_Y$ . Using 14.106 it follows then that

$$\|g_{v,w}(1) - g_{v,w}(0) - g'_{v,w}(0)\|_Y \leq \sup(\{\|g'_{v,w}(\xi) - g'_{v,w}(0)\|_Y \mid \xi \in [0, 1]\}) \quad (14.107)$$

Using the Chain rule we have that for  $\xi \in [0, 1]$

$$\begin{aligned} g'_{v,w}(\xi) &= Dg_{v,w}(\xi)(1) \\ &= [D(f \circ \psi|_{]-2, 2[})(\xi) - D(f \circ \varphi|_{]-2, 2[})(1)] \\ &= [Df(\psi|_{]-2, 2[}(\xi)) \circ D\psi|_{]-2, 2[}(\xi) - Df(\varphi|_{]-2, 2[}(0)) \circ D\varphi|_{]-2, 2[}(0)](1) \\ &= Df(\psi|_{]-2, 2[}(\xi))(D\psi|_{]-2, 2[}(\xi)(1)) - Df(\varphi|_{]-2, 2[}(0))(D\varphi|_{]-2, 2[}(0)(1)) \\ &= Df(\psi|_{]-2, 2[}(\xi))(D\psi|_{]-2, 2[}(\xi)(1)) - Df(\varphi|_{]-2, 2[}(0))(D\varphi|_{]-2, 2[}(0)(1)) \\ &= Df(\psi|_{]-2, 2[}(\xi))((\psi|_{]-2, 2[})'(\xi)) - Df(\varphi|_{]-2, 2[}(0))((\varphi|_{]-2, 2[})'(\xi)) \\ &\stackrel{14.105}{=} Df(\psi|_{]-2, 2[}(\xi))(v) - Df(\varphi|_{]-2, 2[}(0))(v) \\ &= Df(x + \xi \cdot v + w)(v) - Df(x + \xi \cdot v)(v) \end{aligned} \quad (14.108)$$

As  $f$  is  $C^2$  we have that  $Df$  is  $C^1$  hence given a  $\varepsilon > 0$  there exists a  $\delta > 0$  such that if  $\|h\|_X < \delta$  and  $h \in U_x$  then  $\|Df(x+h) - Df(x) - D^2f(x)(h)\|_{L(X,Y)} < \varepsilon \cdot \|h\|_X$ . Take now  $v, w \in X$  such that  $\|v\|_X, \|w\|_X < \min\left(\frac{r}{4}, \frac{\delta}{2}\right)$  then  $\forall \xi \in [0, 1]$  we have that  $\|\xi \cdot v + w\|_X \leq |\xi| \cdot \|v\|_X + \|w\|_X \leq 1 \cdot \frac{\delta}{2} + \frac{\delta}{2} = \delta$  and  $x + \xi \cdot v + w \in U \Rightarrow \xi \cdot v + w \in U_x$ , also  $\|\xi \cdot v\|_X \leq \frac{\delta}{2} < \delta$  and  $x + \xi \cdot v \in U \Rightarrow \xi \cdot v \in U_x$ . So that for  $\xi \in [0, 1]$  we have

$$\|Df(x + \xi \cdot v + w) - Df(x) - D^2f(x) \cdot (\xi \cdot v + w)\|_{L(X,Y)} < \varepsilon \cdot \|\xi \cdot v + w\|_X \quad (14.109)$$

$$\|Df(x + \xi \cdot v) - Df(x) - D^2f(x) \cdot (\xi \cdot v)\|_{L(X,Y)} < \varepsilon \cdot \|\xi \cdot v\|_X \quad (14.110)$$

Also

$$D^2f(x)(\xi \cdot v + w)(v) - D^2f(x)(\xi \cdot v)(v) \underset{D^2f(x) \text{ is linear}}{=} D^2f(x)(w)(v) \quad (14.111)$$

Now

$$\begin{aligned} & \|g'_{v,w}(\xi) - D^2f(x)(w)(v)\|_Y && \stackrel{14.108}{=} \\ & \|Df(x + \xi \cdot v + w)(v) - Df(x + \xi \cdot v)(v) - D^2f(x)(w)(v)\|_Y && \stackrel{14.111}{=} \\ & \|Df(x + \xi \cdot v + w)(v) - D^2f(x)(\xi \cdot v + w)(v) + D^2f(x)(\xi \cdot v)(v) - Df(x + \xi \cdot v)(v)\|_Y && = \\ & \|Df(x + \xi \cdot v + w)(v) - Df(x)(v) - D^2f(x)(\xi \cdot v + w)(v) - [Df(x + \xi \cdot v)(v) - Df(x)(v) - D^2f(x)(\xi \cdot v)(v)]\|_Y && \leq \\ & \|Df(x + \xi \cdot v + w)(v) - Df(x)(v) - D^2f(x)(\xi \cdot v + w)(v)\|_Y + \|Df(x + \xi \cdot v)(v) - Df(x)(v) - D^2f(x)(\xi \cdot v)(v)\|_Y && = \\ & \|(Df(x + \xi \cdot v + w) - Df(x) - D^2f(x)(\xi \cdot v + w))(v)\|_Y + \|(Df(x + \xi \cdot v) - Df(x) - D^2f(x)(\xi \cdot v))(v)\|_Y && \leq \\ & \|Df(x + \xi \cdot v + w) - Df(x) - D^2f(x)(\xi \cdot v + w)\|_{L(X,Y)} \cdot \|v\|_X + \|Df(x + \xi \cdot v) - Df(x) - D^2f(x)(\xi \cdot v)\|_{L(X,Y)} \cdot \|v\|_X && <_{14.109 \text{ and } 14.110} \\ & \varepsilon \cdot \|\xi \cdot v + w\|_X \cdot \|v\|_X + \varepsilon \cdot \|\xi \cdot v\|_X \cdot \|v\|_X && \leq \\ & \varepsilon \cdot |\xi| \|v\|_X \cdot \|v\|_X + \varepsilon \cdot \|w\|_X \cdot \|v\|_X + \varepsilon \cdot |\xi| \cdot \|v\|_X \cdot \|v\|_X && \leq_{\xi \in [0,1]} \\ & \varepsilon \cdot \|v\|_X \cdot \|v\|_X + \varepsilon \cdot \|v\|_X \cdot \|v\|_X + \varepsilon \cdot \|w\|_X \cdot \|v\|_X && = \\ & \varepsilon \cdot \|v\|_X \cdot (2 \cdot \|v\|_X + \|w\|_X) && \leq \\ & 2 \cdot \varepsilon \cdot \|v\|_X \cdot (\|v\|_X + \|w\|_X) && \end{aligned}$$

proving that

$$\forall \xi \in [0, 1] \text{ we have } \|g'_{v,w}(\xi) - D^2f(x)(w)(v)\|_Y \leq 2 \cdot \varepsilon \cdot \|v\|_X \cdot (\|v\|_X + \|w\|_X) \quad (14.112)$$

Further for  $t \in [0, 1]$  we have

$$\begin{aligned} \|g'_{v,w}(t) - g'_{v,w}(0)\|_Y &= \|g'_{v,w}(t) - D^2f(x)(w)(v) - (g'_{v,w}(0) - D^2f(x)(w)(v))\|_Y \\ &\leq \|g'_{v,w}(t) - D^2f(x)(w)(v)\|_Y + \|g'_{v,w}(0) - D^2f(x)(w)(v)\|_Y \\ &\leq_{t,0 \in [0,1] \wedge (14.112)} 2 \cdot \varepsilon \cdot \|v\|_X \cdot (\|v\|_X + \|w\|_X) + 2 \cdot \varepsilon \cdot \|v\|_X \cdot (\|v\|_X + \|w\|_X) \\ &\leq 4 \cdot \varepsilon \cdot \|v\|_X \cdot (\|v\|_X + \|w\|_X) \end{aligned} \quad (14.113)$$

Define the function  $h: [-2, 2] \rightarrow \mathbb{R}$  by  $h(\xi) = \|g'_{v,w}(\xi) - g'_{v,w}(0)\|_Y$ . As  $g'_{v,w}$  is continuous [see 14.58] and  $\|\cdot\|: Y \rightarrow \mathbb{R}$  defined by  $\|\cdot\|_Y = \|h\|_Y$  is continuous [see 11.164] it follows that  $h$  is continuous. Using 11.302 there exists a  $\xi_0 \in [0, 1]$  so that  $\forall t \in [0, 1]$  we have  $h(t) \leq h(\xi_0)$  so that  $h(\xi_0) \leq \sup(h([0, 1])) \leq h(\xi_0)$  giving

$$h(\xi_0) = \sup(h([0, 1])) \text{ or } \|g'_{v,w}(\xi_0) - g'_{v,w}(0)\|_Y = \sup(\{\|g'_{v,w}(t) - g'_{v,w}(0)\|_Y \mid t \in [0, 1]\}) \quad (14.114)$$

Next

$$\begin{aligned}
& \|g_{v,w}(1) - g_{v,w}(0) - D^2f(x)(w)(v)\|_Y = \\
& \|g_{v,w}(1) - g_{v,w}(0) - g'_{v,w}(0) + g'_{v,w}(0) - g'_{v,w}(\xi_0) + g'_{v,w}(\xi_0) - D^2f(x)(w)(v)\|_Y \leqslant \\
& \|g_{v,w}(1) - g_{v,w}(0) - g'_{v,w}(0)\|_Y + \|g'_{v,w}(0) - g'_{v,w}(\xi_0)\|_Y + \|g'_{v,w}(\xi_0) - D^2f(x)(w)(v)\|_Y \leqslant_{(14.107)} \\
& \sup(\{\|g'_{v,w}(\xi) - g'_{v,w}(0)\|_Y \mid \xi \in [0, 1]\}) + \|g'_{v,w}(0) - g'_{v,w}(\xi_0)\|_Y + \|g'_{v,w}(\xi_0) - D^2f(x)(w)(v)\|_Y \\
& =_{14.114} \\
& \|g'_{v,w}(\xi_0) - g'_{v,w}(0)\|_Y + \|g'_{v,w}(0) - g'_{v,w}(\xi_0)\|_Y + \|g'_{v,w}(\xi_0) - D^2f(x)(w)(v)\|_Y = \\
& 2 \cdot \|g'_{v,w}(\xi_0) - g'_{v,w}(0)\|_Y + \|g'_{v,w}(\xi_0) - D^2f(x)(w)(v)\|_Y \leqslant_{14.112} \\
& 2 \cdot \|g'_{v,w}(\xi_0) - g'_{v,w}(0)\|_Y + 2 \cdot \varepsilon \cdot \|v\|_X \cdot (\|v\|_X + \|w\|_X) \leqslant_{14.113} \\
& 8 \cdot \varepsilon \cdot \|v\|_X \cdot (\|v\|_X + \|w\|_X) + 2 \cdot \varepsilon \cdot \|v\|_X \cdot (\|v\|_X + \|w\|_X) = \\
& 10 \cdot \varepsilon \cdot \|v\|_X \cdot (\|v\|_X + \|w\|_X)
\end{aligned}$$

proving that

$$\|g_{v,w}(1) - g_{v,w}(0) - D^2f(x)(w)(v)\|_Y \leqslant 10 \cdot \varepsilon \cdot \|v\|_X \cdot (\|v\|_X + \|w\|_X) \quad (14.115)$$

If we define now  $v' = w$  and  $w' = v$  then we still have  $\|v'\|_X, \|w'\|_X < \min(\frac{r}{4}, \frac{\rho}{2})$  so that by 14.115  $\|g_{v',w'}(1) - g_{v',w'}(0) - D^2f(x)(w')(v')\|_Y \leqslant 0 \cdot \varepsilon \cdot \|v'\|_X \cdot (\|v'\|_X + \|w'\|_X)$  giving by using the definition of  $v', w'$  that

$$\|g_{w,v}(1) - g_{w,v}(0) - D^2f(x)(v)(w)\|_Y \leqslant 10 \cdot \varepsilon \cdot \|w\|_X \cdot (\|w\|_X + \|v\|_X) \quad (14.116)$$

As

$$\begin{aligned}
g_{v,w}(1) - g_{v,w}(0) &= f(x + v \cdot 1 + w) - f(x + v \cdot 1) - f(x + v \cdot 0 + w) + f(x + v \cdot 0) \\
&= f(x + v + w) - f(x + v) - f(x + w) + f(x) \\
&= f(x + w \cdot 1 + v) - f(x + w \cdot 1) - f(x + w \cdot 0 + v) + f(x + w \cdot 0) \\
&= g_{w,v}(1) - g_{w,v}(0)
\end{aligned}$$

so applying then (14.116) we have

$$\|g_{v,w}(1) - g_{v,w}(0) - D^2f(x)(v)(w)\|_Y \leqslant 10 \cdot \varepsilon \cdot \|w\|_X \cdot (\|v\|_X + \|w\|_X) \quad (14.117)$$

Next we have

$$\begin{aligned}
& \|D^2f(x)(v, w) - D^2f(x)(w, v)\|_Y = \\
& \|g_{v,w}(1) - g_{v,w}(0) - D^2f(x)(w, v) - (g_{v,w}(1) - g_{v,w}(0) - D^2f(x)(v)(w))\|_Y \leqslant \\
& \|g_{v,w}(1) - g_{v,w}(0) - D^2f(x)(w, v)\|_Y + \|(g_{v,w}(1) - g_{v,w}(0) - D^2f(x)(v)(w))\|_Y \leqslant_{14.115 \text{ and } 14.117} \\
& 10 \cdot \varepsilon \cdot \|v\|_X \cdot (\|v\|_X + \|w\|_X) + 10 \cdot \varepsilon \cdot \|w\|_X \cdot (\|v\|_X + \|w\|_X) \leqslant \\
& \varepsilon \cdot (10 \cdot (\|v\|_X \cdot (\|v\|_X + \|w\|_X) + \|w\|_X \cdot (\|v\|_X + \|w\|_X)))
\end{aligned}$$

As  $\varepsilon$  was chosen arbitrary we have by 9.57 that  $\|D^2f(x)(v, w) - D^2f(x)(w, v)\|_Y = 0$  proving finally that

$$D^2f(x)(v, w) = D^2f(w, v) \quad \square$$

To extend the above to  $n$ -the differentials we have to state and prove different lemmas. The first lemma shows that we only have to prove that a differential, as a member of  $L^n(X; Y)$ , is invariant under specific transposition, to prove that it is invariant under general permutations.

**Lemma 14.123.** *Let  $n \in \mathbb{N}_0 \setminus \{1\}$ ,  $\langle X, \|\cdot\|_X \rangle$  and  $\langle Y, \|\cdot\|_Y \rangle$  normed sets and  $L \in L^n(X; Y)$  is such that  $\forall i \in \{1, \dots, n-1\}$  and  $\forall x_1, \dots, x_n \in X$   $L(x_1: \dots: x_n) = L(x_{(i \leftrightarrow n, i+1)(1)}: \dots: x_{(i \leftrightarrow n, i+1)(n)})$  then we have that  $\forall \sigma \in P_n$  [see 10.64] and  $\forall x_1, \dots, x_n \in X$  that  $L(x_1: \dots: x_n) = L(x_{\sigma(1)}: \dots: x_{\sigma(n)})$*

**Proof.** Using 11.221 we have that  $\mathcal{P}_n: L^n(X; Y) \rightarrow L(X^n; Y)$  defined by  $\mathcal{P}(L)(x_1, \dots, x_n) = L(x_1; \dots; x_n)$  is a norm preserving isomorphism. Then as  $\forall i \in \{1, \dots, n-1\}$  we have  $L(x_1: \dots: x_n) = L(x_{(i \leftrightarrow n, i+1)(1)}: \dots: x_{(i \leftrightarrow n, i+1)(n)})$  it follows that  $\mathcal{P}_n(L)(x_1, \dots, x_n) = L(x_1: \dots: x_n) = L(x_{(i \leftrightarrow n, i+1)(1)}: \dots: x_{(i \leftrightarrow n, i+1)(n)}) = \mathcal{P}_n(L)(x_{(i \leftrightarrow n, i+1)(1)}, \dots, x_{(i \leftrightarrow n, i+1)(n)})$ . Let  $\sigma \in P_n$  then using 10.73 we have that  $\mathcal{P}_n(L)(x_1, \dots, x_n) = \mathcal{P}_{n+1}(x_{\sigma(1)}, \dots, x_{\sigma(n)})$ . Hence  $L(x_1: \dots: x_n) = \mathcal{P}_n(L)(x_1, \dots, x_n) = \mathcal{P}_n(x_{\sigma(1)}, \dots, x_{\sigma(n)}) = L(x_{\sigma(1)}, \dots, x_{\sigma(n)})$ .  $\square$

Next we prove that the evaluation operator is linear and continuous hence of class  $C^\infty$

**Lemma 14.124.** *Let  $n \in \mathbb{N}_0$ ,  $\langle X, \|\cdot\|_X \rangle$ ,  $\langle Y, \|\cdot\|_Y \rangle$  normed spaces and define given  $(v_1, \dots, v_n) \in X^n$   $e_{v_1, \dots, v_n}: L^n(X; Y) \rightarrow Y$  by  $e_{v_1, \dots, v_n}(L) = L(v_1: \dots: v_n)$ . Then we have that  $e_{v_1, \dots, v_n}$  is linear and continuous. So by 14.75  $e_{v_1, \dots, v_n}$  is  $C^\infty$  and  $De_{v_1, \dots, v_n}(L) = e_{v_1, \dots, v_n}$  and if  $k \geq 2$  then  $D^k e_{v_1, \dots, v_n}(L) = C_{X, L^n(X; Y), 0}$  [the constant mapping to 0]*

**Proof.** First let  $L_1, L_2 \in L^n(X; Y)$  and  $\alpha, \beta \in \mathbb{K}$  then we have that  $e_{v_1, \dots, v_n}(\alpha \cdot L_1 + \beta \cdot L_2) = (\alpha \cdot L_1 + \beta \cdot L_2)(v_1: \dots: v_n) \stackrel{11.215}{=} \alpha \cdot L_1(v_1: \dots: v_n) + \beta \cdot L_2(v_1: \dots: v_n) = \alpha \cdot e_{v_1, \dots, v_n}(L_1) + \beta \cdot e_{v_1, \dots, v_n}(L_2)$  proving that

$$e_{v_1, \dots, v_n} \text{ is linear}$$

Second we have fro  $L \in L^n(X; Y)$   $\|e_{v_1, \dots, v_n}(L)\|_Y = \|L(v_1: \dots: v_n)\|_Y \leq 11.216 \|L\| \cdot \prod_{i \in \{1, \dots, n\}} \|v_i\|_X$  proving that  $\|e_{v_1, \dots, v_n}\| \leq \prod_{i \in \{1, \dots, n\}} \|v_i\|_X$  and thus that

$$e_{v_1, \dots, v_n} \text{ is continuous}$$

The rest of the theorem follows from 14.75 □

The next lemma shows how we can relate  $D^n f(x)$  to  $D^{n-1} f(x)$  which will be needed for our proof by mathematical induction.

**Lemma 14.125.** *Let  $n \in \mathbb{N}_0$  with  $n \geq 2$ ,  $\langle X, \|\cdot\|_X \rangle$ ,  $\langle Y, \|\cdot\|_Y \rangle$  be normed spaces,  $V \subseteq Y$ ,  $U \subseteq X$  open,  $v_1, \dots, v_{n-1} \in X^{n-1}$  and  $f: U \rightarrow V$  a function of class  $C^n$  then if  $g: U \rightarrow Y$  is defined by  $g(x) = D^{n-1} f(x)(v_1: \dots: v_{n-1})$  then  $g$  is  $C^1$  and  $Dg(x)(v) = D^n f(x)(v: v_1: \dots: v_{n-1})$*

**Proof.** Note that

$$\begin{aligned} g(x) &= D^{n-1} f(x)(v_1: \dots: v_{n-1}) \\ &\stackrel{14.124}{=} e_{v_1, \dots, v_{n-1}}(D^{n-1} f(x)) \\ &= ((e_{v_1, \dots, v_{n-1}} \circ D^{n-1} f))(x) \end{aligned}$$

proving that

$$g = e_{v_1, \dots, v_{n-1}} \circ D^{n-1} f \tag{14.118}$$

Using then the chain rule [see 14.83], the fact that  $D^{n-1} f$  is of class  $C^1$  [see 14.65] with 14.124 we have that

$$g \text{ is of class } C^1 \tag{14.119}$$

further for  $x \in U$  we have

$$\begin{aligned} Dg(x) &\stackrel{14.118}{=} D(e_{v_1, \dots, v_{n-1}} \circ D^{n-1} f)(x) \\ &\stackrel{14.22}{=} De_{v_1, \dots, v_{n-1}}(D^{n-1} f(x)) \circ D(D^{n-1} f)(x) \\ &= De_{v_1, \dots, v_{n-1}}(D^{n-1} f(x)) \circ D^n f(x) \\ &\stackrel{14.124}{=} e_{v_1, \dots, v_{n-1}} \circ D^n f(x) \end{aligned}$$

hence

$$\begin{aligned} Dg(x)(v) &= (e_{v_1, \dots, v_{n-1}} \circ D^n f(x))(v) \\ &= e_{v_1, \dots, v_{n-1}}(D^n f(x)(v)) \\ &= (D^n f(x)(v))(v_1: \dots: v_{n-1}) \\ &\stackrel{11.213}{=} D^n f(x)(v: v_1: \dots: v_{n-1}) \end{aligned} \tag{14.120}$$

The lemma is then proved by 14.119 and 14.120. □

**Lemma 14.126.** *Let  $n \in \mathbb{N}_0 \vdash n > 2$ ,  $\langle X, \|\cdot\|_X \rangle$ ,  $\langle Y, \|\cdot\|_Y \rangle$  be normed spaces,  $V \subseteq Y$ ,  $U \subseteq X$  open,  $v_1, \dots, v_{n-2} \in X^{n-2}$  and  $f: U \rightarrow V$  a function of class  $C^n$ . Let  $g: U \rightarrow Y$  defined by  $x \rightarrow g(x) = D^{n-2} f(x)(v_1: \dots: v_{n-2})$  then  $g$  is  $C^2$  and  $D^2 g(x)(v)(u) = D^n f(x)(v: u: v_1: \dots: v_{n-2})$*

**Proof.** Note that

$$\begin{aligned} g(x) &= D^{n-2}f(x)(v_1: \dots: v_{n-2}) \\ &= e_{v_1, \dots, v_{n-2}}(D^{n-2}f(x)) \\ &= (e_{v_1, \dots, v_{n-2}} \circ D^{n-2}f)(x) \end{aligned}$$

proving that

$$g = (e_{v_1, \dots, v_{n-2}} \circ D^{n-2}f) \quad (14.121)$$

Using then the Chain rule [see 14.83], the fact that  $D^{n-2}f$  is of class  $C^2$  [by 14.65] we have that

$$g \text{ is of class } C^2 \quad (14.122)$$

Further

$$\begin{aligned} Dg(x) &= D((e_{v_1, \dots, v_{n-2}} \circ D^{n-2}f))(x) \\ &= De_{v_1, \dots, v_{n-2}}(D^{n-2}f(x)) \circ D(D^{n-2}f)(x) \\ &\stackrel{14.124}{=} e_{v_1, \dots, v_{n-2}} \circ D(D^{n-2}f)(x) \\ &= e_{v_1, \dots, v_{n-2}} \circ D^{n-1}f(x) \end{aligned}$$

so that if  $u \in X$  we have

$$\begin{aligned} Dg(x)(u) &= e_{v_1, \dots, v_{n-2}}(D^{n-1}f(x)(u)) \\ &= (D^{n-1}f(x)(u))(v_1, \dots, v_{n-2}) \\ &\stackrel{11.213}{=} D^{n-1}f(x)(u: v_1: \dots, v_{n-2}) \\ &\stackrel{14.124}{=} e_{u, v_1, \dots, v_{n-2}}(D^{n-1}f(x)) \\ &= (e_{u, v_1, \dots, v_{n-2}} \circ D^{n-1}f)(x) \end{aligned} \quad (14.123)$$

Define now  $Dg(*)u: U \rightarrow Y$  by  $(Dg(*)u)(x) = Dg(x)(u)$  then using 14.123 we have that  $(Dg(*)u)(x) = (e_{u, v_1, \dots, v_{n-2}} \circ D^{n-1}f)(x)$  proving that

$$Dg(*)u = e_{u, v_1, \dots, v_{n-2}} \circ D^{n-1}f \quad (14.124)$$

Using the chain rule [see 14.83] and the fact that  $D^{n-1}f$  is  $C^1$  [see 14.65] we have that  $Dg(*)u$  is  $C^1$  and

$$\begin{aligned} D(Dg(*)u)(x) &\stackrel{14.123}{=} D(e_{u, v_1, \dots, v_{n-2}} \circ D^{n-1}f)(x) \\ &= De_{u, v_1, \dots, v_{n-2}}(D^{n-1}f(x)) \circ D(D^{n-1}f)(x) \\ &= De_{u, v_1, \dots, v_{n-2}}(D^{n-1}f(x)) \circ D^n f(x) \\ &\stackrel{14.124}{=} e_{u, v_1, \dots, v_{n-2}} \circ D^n f(x) \end{aligned} \quad (14.125)$$

Next

$$\begin{aligned} D^2g(x)(v)(u) &\stackrel{14.85}{=} D(Dg(*)u)(x)(v) \\ &\stackrel{14.125}{=} (e_{u, v_1, \dots, v_{n-2}} \circ D^n f(x))(v) \\ &= e_{u, v_1, \dots, v_{n-2}}(D^n f(x)(v)) \\ &= (D^n f(x)(v))(u, v_1, \dots, v_{n-2}) \\ &\stackrel{11.213}{=} D^n f(x)(v: u: v_1, \dots, v_{n-2}) \end{aligned}$$

proving the last part of the lemma.  $\square$

**Lemma 14.127.** Let  $n \in \mathbb{N}_0$  with  $1 < n$ ,  $\langle X, \|\cdot\|_X \rangle$  a normed space.  $\langle Y, \|\cdot\|_Y \rangle$  a Banach space,  $V \subseteq Y$ ,  $U \subseteq X$  and  $f: U \rightarrow V$  a function of class  $C^n$  then  $\forall x \in U$  and  $\forall i \in \{1, \dots, n-1\}$  we have that  $D^n f(x)(u_1: \dots: u_n) = D^n f(x)(u_{(i \leftrightarrow n, i+1)(1)}: \dots: u_{(i \leftrightarrow n, i+1)(n)})$  where  $(u_1, \dots, u_n) \in X^n$

**Proof.** We prove this by induction on  $n$  so let  $\mathcal{S} = \{n \in \{2, \dots, \infty\} \mid \text{if } f: U \rightarrow V \text{ is of class } C^n \text{ then } \forall x \in U \wedge \forall i \in \{1, \dots, n-1\} \text{ we have } D^n f(x)(u_1: \dots: u_n) = D^n f(x)(u_{(i \leftrightarrow n-i+1)(1)}: \dots: u_{(i \leftrightarrow n-i+1)(n)}) \text{ where } (u_1, \dots, u_n \in X^n)\}$  then we have

$2 \in \mathcal{S}$ . If  $n=2$  and  $f: U \rightarrow Y$  is of class  $C^2$  then if  $i \in \{1, \dots, 2-1\} = \{1\}$  we have that  $i=1$  and

$$\begin{aligned} D^2 f(x)(u_1: u_2) &= D^2 f(x)(u_1)(u_2) \\ &\stackrel{14.122}{=} D^2 f(x)(u_2)(u_1) \\ &= D^2 f(x)(u_2: \dots: u_1) \\ &= D^2 f(x)(u_{(1 \leftrightarrow 2)(1)}: \dots: u_{(1 \leftrightarrow 2)(2)}) \end{aligned}$$

proving that  $2 \in \mathcal{S}$ .

$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$ . Let  $i \in \{1, \dots, (n+1)-1\} = \{1, \dots, n\}$  and let  $f: U \rightarrow V$  be of class  $C^{n+1}$  then by 14.65 Take  $x \in U$  and  $(u_1, \dots, u_{n+1}) \in X^{n+1}$ . For  $i$  we have the following cases to consider

**i = 1.** Then  $(i \leftrightarrow_{n+1} i+1)(1) = 2 \wedge (i \leftrightarrow_{n+1} i+1)(2) = 1$  and  $(i \leftrightarrow_{n+1} i+1)(k) = k$  if  $k \in \{3, \dots, n+1\}$ . Define then  $g: U \rightarrow Y$  by  $g(x) = D^{(n+1)-2} f(x)(u_3: \dots: u_{n+1}) = D^{n-1} f(x)(u_3: \dots: u_{n+1})$  then we have that  $g$  is  $C^2$  [see 14.126]

$$\begin{aligned} D^{n+1} f(x)(u_1: u_2: \dots: u_{n+1}) &\stackrel{14.126}{=} D^2 g(x)(u_1)(u_2) \\ &\stackrel{14.122}{=} D^2 g(x)(u_2)(u_1) \\ &\stackrel{14.126}{=} D^{n+1} f(x)(u_2: u_1: \dots: u_{n+1}) \\ &= D^{n+1} f(x)(u_{(i \leftrightarrow_{n+1} i+1)(1)}: \dots: u_{(i \leftrightarrow_{n+1} i+1)(n+1)}) \end{aligned}$$

proving that  $n+1 \in \mathcal{S}$  in this case.

**$1 < i$ .** Then  $i \in \{2, \dots, n\}$  so  $\forall k \in \{1, \dots, n-1\}$   $(i-1 \leftrightarrow_n i)(k)+1 = 1 + \begin{cases} i-1 & \text{if } k=i \\ i & \text{if } k=i-1 \\ k & \text{if } k \neq i-1, i \end{cases} = \begin{cases} i & \text{if } k=i \\ i+1 & \text{if } k=i-1 \\ k+1 & \text{if } k \neq i, i-1 \end{cases}$  and  $(i \leftrightarrow_{n+1} i+1)(k+1) = \begin{cases} i & \text{if } k+1=i+1 \\ i+1 & \text{if } k+1=i \\ k+1 & \text{if } k+1 \neq i, i+1 \end{cases} = \begin{cases} i & \text{if } k=i \\ i+1 & \text{if } k=i-1 \\ k+1 & \text{if } k \neq i-1, i \end{cases}$  proving that

$$\forall k \in \{1, \dots, n-1\} \text{ we have } (i \leftrightarrow_n i-1)(k)+1 = (i \leftrightarrow i+1)(k+1) \quad (14.126)$$

Using 14.63 it follows that  $f$  is of class  $C^n$ , further if we define  $(y_1, \dots, y_n)$  by  $y_i = u_{i+1}$  then

$$\begin{aligned} D^n f(x)(u_2: \dots, u_{n+1}) &= D^n f(x)(y_1: \dots: y_n) \\ &\stackrel{n \in \mathcal{S}}{=} D^n f(x)(y_{(i-1 \leftrightarrow_n i)(1)}: \dots: y_{(i-1 \leftrightarrow_n i)(n)}) \\ &= D^n f(x)(u_{(i-1 \leftrightarrow_n i)(1)+1}, \dots, u_{(i-1 \leftrightarrow_n i)(1)+1}) \\ &\stackrel{14.126}{=} D^n f(x)(u_{(i \leftrightarrow_{n+1} i+1)(2)}: \dots: u_{(i \leftrightarrow_{n+1} i+1)(n+1)}) \quad (14.127) \end{aligned}$$

Define now  $g: U \rightarrow Y$  by  $g(x) = D^n f(x)(u_2: \dots: u_{n+1})$  and  $h: U \rightarrow Y$  by  $h(x) = D^n f(x)(u_{(i \leftrightarrow_{n+1} i+1)(2)}: \dots: u_{(i \leftrightarrow_{n+1} i+1)(n+1)})$  then by 14.127 we have that  $g = h$ . Finally

$$\begin{aligned} D^{n+1} f(x)(u_1: \dots: u_{n+1}) &\stackrel{1 < i < i+1}{=} D^{n+1} f(x)(u_{(i \leftrightarrow_{n+1} i+1)(1)}: u_2: \dots: u_{n+1}) \\ &\stackrel{14.125}{=} Dg(x)(u_{(i \leftrightarrow_{n+1} i+1)(1)}) \\ &= Dh(x)(u_{(i \leftrightarrow_{n+1} i+1)(1)}) \\ &\stackrel{14.125}{=} D^{n+1} f(x)(u_{(i \leftrightarrow_{n+1} i+1)(1)}: \dots: u_{(i \leftrightarrow_{n+1} i+1)(n+1)}) \end{aligned}$$

proving in this case that  $n+1 \in \mathcal{S}$ .

So in all cases we have  $n+1 \in \mathcal{S}$ .

Mathematical induction proves then the theorem.  $\square$

Now we are finally ready to prove the symmetric nature of the differential

**Theorem 14.128.** Let  $n \in \mathbb{N}_0$ ,  $\langle X, \|\cdot\|_X \rangle$  be a normed space,  $\langle Y, \|\cdot\|_Y \rangle$  a Banach space,  $V \subseteq Y$ ,  $U \subseteq X$  open and  $f: U \rightarrow Y$  is  $C^n$  then  $\forall x \in U$ ,  $\sigma \in P_n$  and  $v = (v_1, \dots, v_n) \in X$  we have  $D^n f(x)(v_{\sigma(1)} : \dots : v_{\sigma(n)}) = D^n f(x)(v_1 : \dots : v_n)$

**Proof.** This follows from 14.127 and 14.123.  $\square$

## 14.6 Higher Order Partial Differentiation

**Definition 14.129.** Let  $n \in \mathbb{N}_0 \setminus \{1\}$ ,  $m \in \mathbb{N}_0$ ,  $\{\langle X_i, \|\cdot\|_i \rangle\}_{i \in \{1, \dots, n\}}$  a finite family of normed spaces,  $\langle Y, \|\cdot\|_Y \rangle$  a normed space,  $V \subseteq Y$ ,  $\langle \prod_{i \in \{1, \dots, n\}} X_i, \|\cdot\| \rangle$  the normed space  $X = \prod_{i \in \{1, \dots, n\}} X_i$  with the maximum norm  $\|\cdot\|$  [see 11.82],  $U \subseteq X$  a open set,  $\{i_j\}_{j \in \{1, \dots, m\}} \subseteq \{1, \dots, n\}$  and  $f: U \rightarrow V$  a function then we say that  $f$  has a higher order partial differential  $D_{i_m, \dots, i_1}^m f$  of order  $m$  on  $U$  recursively as follows

1. If  $m = 1$  then  $\forall x \in U$  the partial differential  $D_{i_1} f(x)$  must exists and  $D_{i_1, \dots, i_1}^1 f: U \rightarrow L^1(X_{i_1}, \dots, X_{i_1}; Y) = L(X_{i_1}; Y)$  is defined by  $(D_{i_1, \dots, i_1}^1 f)(x) = D_{i_1} f(x)$
2. If  $m > 1$  then the partial differential  $D_{i_{m-1}, \dots, i_1}^{m-1} f \in L^{m-1}(X_{i_{m-1}}, \dots, X_{i_1}; Y)$  must exists on  $U$  and  $\forall x \in U$   $D_{i_m} (D_{i_{m-1}, \dots, i_1}^{m-1} f)(x)$  exists and  $D_{i_1, \dots, i_m}^m f: U \rightarrow L(X_{i_m}, L^{m-1}(X_{i_{m-1}}, \dots, X_{i_1}; Y)) = L^m(X_m, \dots, X_{i_1}; Y)$  is defined by  $D_{i_1, \dots, i_m}^m f(x) = D_{i_m} (D_{i_1, \dots, i_m}^m f)(x)$

It will turn out that if a function is  $m$ -times differentiable then all the higher order partial differentials of order  $m$  exist. Also if a function is of class  $C^m$  then all the higher order partial differentials of order  $m$  will be continuous. Note that the existence of partial differentials of order  $m$  does not automatically implies that a function is  $m$ -times differentiable. First we need the following lemma to prove the continuity of the partial differentials.

**Lemma 14.130.** Let  $n \in \mathbb{N}_0 \setminus \{1\}$ ,  $\{\langle X_i, \|\cdot\|_i \rangle\}_{i \in \{1, \dots, n\}}$  a finite family of normed spaces,  $\langle Y, \|\cdot\|_Y \rangle$  a normed space,  $\langle \prod_{i \in \{1, \dots, n\}} X_i, \|\cdot\| \rangle$  the normed space  $X = \prod_{i \in \{1, \dots, n\}} X_i$  with the maximum norm  $\|\cdot\|$  [see 11.82],  $U \subseteq X$  a open set and  $f: U \rightarrow L(X, Y)$  a continuous function then  $\forall i \in \{1, \dots, n\}$  we have that  $f(\star) \circ (i \rightarrow 0): U \rightarrow L(X_i, Y)$  defined by  $(f(\star) \circ (i \rightarrow 0))(x) = f(x) \circ (i \rightarrow 0)$  is continuous.

**Proof.** Let  $x \in U$ . First note that as  $f(x) \in L(X, Y)$  and  $(i \rightarrow 0) \in L(X_i, X)$  [see 14.30] we have by 11.191 that  $(f(\star) \circ (i \rightarrow 0))(x) = f(x) \circ (i \rightarrow 0) \in L(X)$ .

Take  $\varepsilon > 0$  then as  $f$  is continuous there exists a  $\delta > 0$  such that  $\forall y \in X$  with  $\|x - y\| < \delta$  we have that  $\|f(x) - f(y)\|_{L(X, Y)} < \varepsilon$ . Next if  $h \in X_i$  we have

$$\begin{aligned} \|((f(\star) \circ (i \rightarrow 0))(x) - (f(\star) \circ (i \rightarrow 0))(y))(h)\|_Y &= \|(f(x) \circ (i \rightarrow 0) - f(y) \circ (i \rightarrow 0))(h)\|_Y \\ &= \|f(x)((i \rightarrow 0)(h)) - f(y)((i \rightarrow 0)(h))\|_Y \\ &= \|(f(x) - f(y))((i \rightarrow 0)(h))\|_Y \\ &\leq \|f(x) - f(y)\|_{L(X, Y)} \cdot \|(i \rightarrow 0)(h)\| \\ &\stackrel{14.30}{=} \|f(x) - f(y)\|_{L(X, Y)} \cdot \|h\|_i \end{aligned}$$

proving by 11.185 that  $\|((f(\star) \circ (i \rightarrow 0))(x) - (f(\star) \circ (i \rightarrow 0))(y))(h)\|_{L(X_i, Y)} \leq \|f(x) - f(y)\|_{L(X, Y)} < \varepsilon$  which proves continuity.  $\square$

Next we prove that 1-times differentiability implies the existence of partial differentials of order 1.

**Lemma 14.131.** Let  $n \in \mathbb{N}_0 \setminus \{1\}$ ,  $\{\langle X_i, \|\cdot\|_i \rangle\}_{i \in \{1, \dots, n\}}$  a finite family of normed spaces,  $\langle Y, \|\cdot\|_Y \rangle$  a normed space,  $\langle \prod_{i \in \{1, \dots, n\}} X_i, \|\cdot\| \rangle$  the normed space with the maximum norm  $\|\cdot\|$  [see 11.82],  $U \subseteq \prod_{i \in \{1, \dots, n\}} X_i$  a open set and  $f: U \rightarrow Y$  a 1 times differentiable function then  $\forall i \in \{1, \dots, n\}$  we have that  $D_{i_1}^1 f: U \rightarrow L^1(X_i, \dots, X_i; Y) = L(X_i, Y)$  exists [and is continuous if  $f$  is also  $C^1$ ], further  $\forall x \in U$  we have  $D_{i_1, \dots, i_1}^1 f(x) = D^1 f(x) \circ (i \rightarrow 0)$  and  $D f(x) \stackrel{\text{def}}{=} D^1 f(x) = \sum_{i \in \{1, \dots, n\}} D_{i_1, \dots, i_1}^1 f(x) \circ \pi_i$ .

**Proof.** Let  $i \in \{1, \dots, n\}$  then as  $\forall x \in U D f(x)$  exists we have by 14.35 that  $\forall x \in U D_{i_1} f(x)$  exists,  $D_{i_1} f(x) = D f(x) \circ (i \rightarrow 0)$  and  $D f(x) = \sum_{i \in \{1, \dots, n\}} (D_{i_1} f(x) \circ \pi_i)$ . Hence by definition we have that  $D_{i_1, \dots, i_1}^1 f: U \rightarrow L^1(X_i; Y)$  exists with  $D_{i_1, \dots, i_1}^1 f(x) = D f(x) \circ (i \rightarrow 0) = (D f(\star) \circ (i_1 \rightarrow 0))(x)$  and  $D^1 f(x) = \sum_{i \in \{1, \dots, n\}} D_{i_1, \dots, i_1}^1 f(x) \circ \pi_i$ .

Further as  $f$  is  $C^1$  we have that  $Df$  is continuous so using 14.130 we have that  $D_{i_1, \dots, i_m}^1 f = Df(*) \circ (i \rightarrow 0)$  is continuous.  $\square$

Finally we use induction to prove that  $m$ -times differentiability implies the existance of partial differentials of order  $m$ .

**Theorem 14.132.** Let  $n \in \mathbb{N}_0 \setminus \{1\}$ ,  $m \in \mathbb{N}_0$ ,  $\{X_i, \|\cdot\|_i\}_{i \in \{1, \dots, n\}}$  a finite family of normed spaces,  $\langle Y, \|\cdot\|_Y \rangle$  a normed space,  $V \subseteq Y$ ,  $\langle X, \|\cdot\| \rangle$  the normed space  $X = \prod_{i \in \{1, \dots, n\}} X_i$  with the maximum norm  $\|\cdot\|$  [see 11.82], a open set  $U \subseteq X$  and  $f: U \rightarrow V$  a  $m$ -times differentiable function then  $\forall \{i_j\}_{j \in \{1, \dots, m\}} \subseteq \{1, \dots, n\}$  we have that the higher order partial differential  $D_{i_m, \dots, i_1}^m f$  exists and  $\forall x \in U$  we have  $\forall h \in \prod_{j \in \{1, \dots, m\}} X_{i_j}$  that  $D_{i_m, \dots, i_1}^m f(x)(h_m: \dots: h_1) = D^m f(x)((i_m \rightarrow 0)(h_m): \dots: (i_1 \rightarrow 0)(h_1))$ . Further if  $f$  is of class  $C^m$  then  $D_{i_m, \dots, i_1}^m f$  is continuous.

**Proof.** We prove this by induction so let  $\mathcal{S} = \{m \in \mathbb{N}_0 \mid \text{if } f: X \rightarrow V \text{ is } m\text{-time differentiable then } \forall \{i_j\}_{j \in \{1, \dots, m\}} \subseteq \{1, \dots, n\} D_{i_m, \dots, i_1}^m f \text{ exists and } \forall x \in U \text{ we have } \forall h \in \prod_{j \in \{1, \dots, m\}} X_{i_j} \text{ that } D_{i_m, \dots, i_1}^m f(x)(h_m: \dots: h_1) = D^m f(x)((i_m \rightarrow 0)(h_m): \dots: (i_1 \rightarrow 0)(h_1))\}$ . Further if  $f$  is of class  $C^m$  then  $D_{i_m, \dots, i_1}^m f$  is continuous} then we have:

**1  $\in \mathcal{S}$ .** Using 14.131 we have that  $D_{i_1, \dots, i_1}^1 f$  exists and is continuous if  $f$  is  $C^1$ . Further  $\forall x \in U$  and  $h = (h_1) \in \prod_{j \in \{1, \dots, 1\}} X_{i_j}$  we have

$$\begin{aligned} D_{i_1, \dots, i_1}^1 f(x)(h_1: \dots: h_1) &\stackrel{14.131}{=} (D^1 f(x) \circ (i_1 \rightarrow 0))(h_1) \\ &= D^1 f(x)((i_1 \rightarrow 0)(h_1)) \\ &\stackrel{\text{def}}{=} D^1 f(x)((i_1 - 0)(h_1): \dots: (i_1 - 0)(h_1)) \end{aligned}$$

proving that  $1 \in \mathcal{S}$ .

**$m \in \mathcal{S} \Rightarrow m+1 \in \mathcal{S}$ .** Assume that  $f$  is  $(m+1)$ -times differentiable and take  $\{i_j\}_{j \in \{1, \dots, m+1\}} \subseteq \{1, \dots, n\}$ . Define for  $L \in L^m(X; Y)$  and  $(h_1, \dots, h_m) \in \prod_{j \in \{1, \dots, m\}} X_{i_j}$

$$\Psi(L)(h_m, \dots, h_1) = L((i_m \rightarrow 0)(h_m): \dots: (i_1 \rightarrow 0)(h_1))$$

If  $\alpha, \beta \in \mathbb{K}$ ,  $j \in \{1, \dots, m\}$  and  $k, l \in X_j$  then

$$\begin{aligned} &\Psi(L)\left(h_m, \dots, \underbrace{\alpha \cdot k + \beta \cdot l}_{j}, \dots, h_1\right) = \\ &L\left((i_m \rightarrow 0)(h_m): \dots, \underbrace{(i_j \rightarrow 0)(\alpha \cdot k + \beta \cdot l)}_j, \dots: (i_1 \rightarrow 1)(h_1)\right) \stackrel{11.218}{=} \\ &\mathcal{P}_n(L)\left((i_m \rightarrow 0)(h_m), \dots, \underbrace{(i_j \rightarrow 0)(\alpha \cdot k + \beta \cdot l)}_j, \dots, (i_1 \rightarrow 1)(h_1)\right) \stackrel{14.30}{=} \\ &\mathcal{P}_n(L)\left((i_m \rightarrow 0)(h_m), \dots, \underbrace{\alpha \cdot (i_j \rightarrow 0)(k) + \beta \cdot (i_j \rightarrow 0)(l)}_j, \dots, (i_1 \rightarrow 1)(h_1)\right) \stackrel{11.220}{=} \\ &\alpha \cdot \mathcal{P}_n(L)\left((i_m \rightarrow 0)(h_m), \dots, \underbrace{(i_j \rightarrow 0)(k)}_j, \dots, (i_1 \rightarrow 1)(h_1)\right) + \beta \cdot \mathcal{P}_n(L)\left((i_m \rightarrow 0)(h_m), \dots, \right. \\ &\left. \underbrace{(i_j \rightarrow 0)(l)}_j, \dots, (i_1 \rightarrow 1)(h_1)\right) = \\ &\alpha \cdot L\left((i_m \rightarrow 0)(h_m): \dots, \underbrace{(i_j \rightarrow 0)(k)}_j, \dots: (i_1 \rightarrow 1)(h_1)\right) + \beta \cdot L\left((i_m \rightarrow 0)(h_m): \dots, \right. \\ &\left. \underbrace{(i_j \rightarrow 0)(l)}_j, \dots: (i_1 \rightarrow 1)(h_1)\right) \stackrel{11.218}{=} \\ &\alpha \cdot \Psi(L)\left(h_m, \dots, \underbrace{k}_j, \dots, h_1\right) + \beta \cdot \Psi(L)\left(h_m, \dots, \underbrace{l}_j, \dots, h_1\right) \end{aligned}$$

proving that

$$\Psi(L) \text{ is multilinear} \quad (14.128)$$

Further we have for  $(h_1, \dots, m) \in \prod_{j \in \{1, \dots, m\}} X_{i_j}$  that

$$\begin{aligned} \|\Psi(L)(h_m, \dots, h_1)\|_Y &= \|L((i_m \rightarrow 0)(h_m); \dots; (i_1 \rightarrow 0)(h_1))\|_Y \\ &\stackrel{11.216}{\leq} \|L\|_{L^m(X; Y)} \cdot \prod_{j \in \{m, \dots, 1\}} \|(i_j \rightarrow 0)(h_j)\| \\ &\stackrel{14.30}{=} \|L\|_{L(X^m; Y)} \cdot \prod_{j \in \{m, \dots, 1\}} \|h_j\|_j \end{aligned}$$

proving by 11.194 and 11.198 that

$$\Psi(L) \in L(X_m, \dots, X_1; Y) \text{ with } \|\Psi(L)\|_{L(X_m, \dots, X_1; Y)} \leq \|L\|_{L(X^m; Y)} \quad (14.129)$$

We can now use the above and  $\mathcal{P}_n^{-1}: L(X_m, \dots, X_1; Y) \rightarrow L^m(X_m, \dots, X_1; Y)$  to define

$$\Phi: L^m(X; Y) \rightarrow L^m(X_m, \dots, X_1; Y) \text{ by } \Phi = \mathcal{P}_n^{-1} \circ \Psi \quad (14.130)$$

Then we have if  $\alpha, \beta \in \mathbb{K}$ ,  $L_1, L_2 \in L^m(X; Y)$  that  $\forall h \in \prod_{j \in \{1, \dots, m\}} X_{i_j}$

$$\begin{aligned} \Psi(\alpha \cdot L_1 + \beta \cdot L_2)(h_m, \dots, h_1) &= \\ (\alpha \cdot L_1 + \beta \cdot L_2)((i_m \rightarrow 0)(h_m); \dots; (i_1 \rightarrow 0)(h_1)) &\stackrel{11.215}{=} \\ \alpha \cdot L_1((i_m \rightarrow 0)(h_m); \dots; (i_1 \rightarrow 0)(h_1)) + \beta \cdot L_2((i_m \rightarrow 0)(h_m); \dots; (i_1 \rightarrow 0)(h_1)) &= \\ (\alpha \cdot \Psi(L_1) + \beta \cdot \Psi(L_2))(h_m; \dots; h_1) \end{aligned}$$

proving that  $\Psi(\alpha \cdot L_1 + \beta \cdot L_2) = \alpha \cdot \Psi(L_1) + \beta \cdot \Psi(L_2)$  and thus as  $\mathcal{P}_n^{-1}$  is linear that  $\Phi$  is linear.

Further if  $L \in L^m(X; Y)$  we have

$$\begin{aligned} \|\Phi(L)\|_{L^m(X_m, \dots, X_1; Y)} &= \|\mathcal{P}_n^{-1}(\Psi(L))\|_{L^m(X_m, \dots, X_1; Y)} \\ &\stackrel{11.221}{=} \|\Psi(L)\|_{L(X_m, \dots, X_1; Y)} \\ &\stackrel{14.129}{\leq} \|L\|_{L(X^m; Y)} \end{aligned}$$

proving by 11.180 that  $\Phi$  is continuous. So we have that

$$\Phi \in L(L^m(X; Y), L^m(X_m, \dots, X_1; Y)) \quad (14.131)$$

Let  $x \in U$  we have as  $D^m f(x) \in L^m(X; Y)$  that for  $h \in \prod_{j \in \{1, \dots, m\}} X_{i_j}$

$$\begin{aligned} \Phi(D^m f(x))(h_m; \dots; h_1) &= \mathcal{P}_n^{-1}(\Psi(D^m f(x)))(h_m; \dots; h_1) \\ &= \Psi(D^m f(x))(h_m, \dots, h_1) \\ &= D^m(f(x))((i_m \rightarrow 0)(h_m); \dots; (i_1 \rightarrow 0)(h_1)) \\ &\stackrel{m \in \mathcal{S}}{=} (D_{i_m, \dots, i_1}^m f(x))(h_m; \dots; h_1) \end{aligned}$$

proving by 11.212 that  $\Phi(D^m f(x)) = D_{i_m, \dots, i_1}^m f(x)$  or that

$$D_{i_m, \dots, i_1}^m f = \Phi \circ D^m f \quad (14.132)$$

Take  $x \in U$  then, as  $(i_{m+1} \rightarrow x)$  is  $C^\infty$  [see 14.76],  $\Phi$  is  $C^\infty$  [see 14.131 and 14.75],  $f$  is  $(m+1)$ -times differentiable so that  $D^m f$  is 1-times differentiable, we have by the chain rule [see 14.83] that  $\Phi \circ D^m f \circ (i_{m+1} \rightarrow x)$  is differentiable at  $x_{m+1}$ . So by 14.132  $D((D_{i_m, \dots, i_1}^m f) \circ (i_{m+1} \rightarrow x))(x_{i_{m+1}})$  exists which proves that  $D_{i_{m+1}}(D_{i_m, \dots, i_1}^m f)(x)$  exists. Hence

$$\forall x \in U \text{ we have that } D_{i_{m+1}, \dots, i_1}^{m+1} f(x) \text{ exists} \quad (14.133)$$

Next

$$\begin{aligned}
 D_{i_{m+1}, \dots, i_1}^{m+1} f(x) &= D((\Phi \circ D^m f) \circ (i_{m+1} \rightarrow x))(x_{i_{m+1}}) \\
 &\stackrel{14.22}{=} D(\Phi \circ D^m f)((i_{m+1} \rightarrow x)(x_{i_{m+1}})) \circ D(i_{m+1} \rightarrow x)(x_{i_{m+1}}) \\
 &\stackrel{14.76}{=} D(\Phi \circ D^m f)((i_{m+1} \rightarrow x)(x_{i_{m+1}})) \circ (i_{m+1} \rightarrow 0) \\
 &\stackrel{14.29}{=} D(\Phi \circ D^m f)(x) \circ (i_{m+1} \rightarrow 0) \\
 &\stackrel{14.22}{=} D\Phi((D^m f(x))) \circ D(D^m f)(x) \circ (i_{m+1} \rightarrow 0) \\
 &\stackrel{14.131 \text{ and } 14.75}{=} \Phi \circ D^{m+1} f(x) \circ (i_{m+1} \rightarrow 0)
 \end{aligned}$$

proving that

$$D_{i_{m+1}, \dots, i_1}^{m+1} f(x) = \Phi \circ D^{m+1} f(x) \circ (i_{m+1} \rightarrow 0) \quad (14.134)$$

Take  $h \in \prod_{j \in \{1, \dots, m+1\}} X_{i_j}$  then

$$\begin{aligned}
 D_{i_{m+1}, \dots, i_1}^{m+1} f(x)(h_{m+1}: \dots: h_1) &\stackrel{14.134}{=} \\
 (\Phi \circ D^{m+1} f(x) \circ (i_{m+1} \rightarrow 0))(h_{m+1}: \dots: h_1) &\stackrel{11.213}{=} \\
 (\Phi \circ D^{m+1} f(x) \circ (i_{m+1} \rightarrow 0))(h_{m+1})(h_m: \dots: h_1) &= \\
 \Phi(D^{m+1} f(x)((i_{m+1} \rightarrow 0)(h_{m+1}))(h_m: \dots: h_1)) &\stackrel{(1)}{=} \\
 (D^{m+1} f(x)((i_{m+1} \rightarrow 0)(h_{m+1}))((i_m \rightarrow 0)(h_m): \dots: (i_1 \rightarrow 0)(h_1))) &\stackrel{11.213}{=} \\
 D^{m+1} f(x)((i_{m+1} \rightarrow 0)(h_{m+1}), \dots, (i_1 \rightarrow 0)(h_1))
 \end{aligned}$$

where (1) is because of  $D^{m+1} f(x)((i_{m+1} \rightarrow 0)(h_{m+1})) \in L^m(X; Y)$ . Proving that

$$D_{i_{m+1}, \dots, i_1}^{m+1} f(x)(h_{m+1}: \dots: h_1) = D^{m+1} f(x)((i_{m+1} \rightarrow 0)(h_{m+1}): \dots: (i_1 \rightarrow 0)(h_1)) \quad (14.135)$$

If  $f$  is of class  $C^{m+1}$  then  $D^{m+1} f: U \rightarrow L^{m+1}(X; Y) = L(X, L^m(X, Y))$  is continuous hence using 14.130 we have that  $D^{(m+1)} f(*) \circ (i_{m+1} \rightarrow 0)$  is continuous. As  $\Phi$  is continuous it follows that  $\Phi \circ (D^{(m+1)} f(*) \circ (i_{m+1} \rightarrow 0))$  is continuous. As  $(\Phi \circ D^{(m+1)} f(*) \circ (i_{m+1} \rightarrow 0))(x) = \Phi \circ D^{(m+1)} f(x) \circ (i_{m+1} \rightarrow 0) \stackrel{14.134}{=} D_{i_{m+1}, \dots, i_1}^{m+1} f(x)$  it follows that  $\Phi \circ D^{(m+1)} f(*) \circ (i_{m+1} \rightarrow 0) = D_{i_{m+1}, \dots, i_1}^{m+1} f$ . So

$$\text{if } f \text{ is } C^{m+1} \text{ then } D_{i_{m+1}, \dots, i_1}^{m+1} f \text{ is continuous} \quad (14.136)$$

Finally 14.133, 14.135 and 14.136 proves that  $n+1 \in \mathcal{S}$ .

Mathematical induction completes then the proof.  $\square$

Next we show that the order of partial differentiation does not matter.

**Corollary 14.133.** Let  $n \in \mathbb{N}_0 \setminus \{1\}$ ,  $m \in \mathbb{N}_0$ ,  $\{X_i, \|\cdot\|_i\}_{i \in \{1, \dots, n\}}$  a finite family of normed spaces,  $\langle Y, \|\cdot\|_Y \rangle$  a normed spaces,  $V \subseteq Y$ ,  $\langle X, \|\cdot\| \rangle$  the normed space  $X = \prod_{i \in \{1, \dots, n\}} X_i$  with the maximum norm  $\|\cdot\|$  [see 11.82], a open set  $U \subseteq X$  and  $f: U \rightarrow V$  a  $m$ -times differentiable function and  $\{i_j\}_{j \in \{1, \dots, m\}} \subseteq \{1, \dots, n\}$  then  $\forall \sigma \in P_n$  [see 10.66] we have that  $D_{i_m, \dots, i_1}^m f(x)(h_m: \dots: h_1) = D_{i_{\sigma(m)}, \dots, i_{\sigma(1)}}^m f(x)(h_{\sigma(m)}: \dots: h_{\sigma(1)})$  for  $h \in \prod_{j \in \{1, \dots, m\}} X_{i_j}$ .

**Proof.** Let  $x \in U$  and  $h \in \prod_{j \in \{1, \dots, m\}} X_{i_j}$  then

$$\begin{aligned}
 (D_{i_{\sigma(m)}, \dots, i_{\sigma(1)}}^m f(x))(h_{\sigma(m)}: \dots: h_{\sigma(1)}) &\stackrel{14.132}{=} \\
 D^m f(x)((i_{\sigma(m)} \rightarrow 0)(h_{\sigma(m)}): \dots: (i_{\sigma(1)} \rightarrow 0)(h_{\sigma(1)})) &\stackrel{14.128}{=} \\
 D^m f(x)((i_m \rightarrow 0)(h_m): \dots: (i_1 \rightarrow 0)(h_1)) &= \\
 D_{i_m, \dots, i_1}^m f(h_m: \dots: h_1) &= \\
 \square
 \end{aligned}$$

We prove now that if the partial differentials of a function exists and are continuous then the function is differentiable.

**Theorem 14.134.** Let  $n \in \mathbb{N}_0 \setminus \{1\}$ ,  $m \in \mathbb{N}_0$ ,  $\{\langle X_i, \|\cdot\|_i \rangle\}_{i \in \{1, \dots, n\}}$  a finite family of normed spaces,  $\langle Y, \|\cdot\|_Y \rangle$  a Banach space,  $\langle X, \|\cdot\| \rangle$  the normed space  $X = \prod_{i \in \{1, \dots, n\}} X_i$  with the maximum norm  $\|\cdot\|$  [see 11.82 and 11.338], a open set  $U \subseteq X$ ,  $V \subseteq Y$  and  $f: U \rightarrow V$  such that  $\forall i \in \{1, \dots, n\}$   $D_i f$  exists on  $U$  and is continuous then  $f$  is of class  $C^1$  with  $Df(x)(h) = \sum_{i=1}^n D_i f(x)(h_i)$  where  $h \in X$  or  $Df(x) = \sum_{i=1}^n D_i f(x) \circ \pi_i$

**Proof.** Let  $\varepsilon > 0$ . Given  $h \in X$  and  $i \in \{0, \dots, n\}$  define  $h^i = \begin{cases} \sum_{j=1}^i (j \rightarrow 0)(h_j) & \text{if } i \in \{1, \dots, n\} \\ 0 & \text{if } i = 0 \end{cases}$ . Using 14.31 we have then that

$$h = h^n \quad (14.137)$$

Given  $i \in \{1, \dots, n\}$  and  $k \in \{1, \dots, n\}$  we calculate the value of  $(h^i)_k$ . We have for  $k$  the following cases :

$k \in \{1, \dots, i\}$ . then

$$\begin{aligned} (h^i)_k &= \left( \sum_{j=1}^i (j \rightarrow 0)(h_j) \right)_k \\ &= \sum_{j=1}^i ((j \rightarrow 0)(h_j))_k \\ &\stackrel{14.30}{=} \sum_{j=1}^i \delta_{j,k}(h_j) \\ &= h_k \end{aligned}$$

$k \in \{i+1, \dots, n\}$ . then

$$\begin{aligned} (h^i)_k &= \left( \sum_{j=1}^i (j \rightarrow 0)(h_j) \right)_k \\ &= \sum_{j=1}^i ((j \rightarrow 0)(h_j))_k \\ &\stackrel{14.30}{=} \sum_{j=1}^i \delta_{j,k}(h_j) \\ &\forall j \in \{1, \dots, n\} k \neq j \quad \sum_{j=1}^i 0_k \text{ [neutral element of } X_k] \\ &= 0_k \end{aligned}$$

proving that  $(h^i)_k = \begin{cases} h_k & \text{if } k \in \{1, \dots, i\} \\ 0_k & \text{if } k \in \{i+1, \dots, n\} \end{cases}$ . So  $\|(h^i)_k\|_k = \begin{cases} \|h_k\|_k & \text{if } k \in \{1, \dots, i\} \\ \|0_k\|_k = 0 & \text{if } k \in \{i+1, \dots, n\} \end{cases} \leq \|h_k\|_k$  hence  $\forall i \in \{0, \dots, n\}$

$$\begin{aligned} \|h^i\| &= \max(\{\|(h^i)_k\|_k \mid k \in \{1, \dots, n\}\}) \\ &\leq \max(\{\|h_k\|_k \mid k \in \{1, \dots, n\}\}) \\ &= \|h\| \end{aligned} \quad (14.138)$$

Let  $x \in U$  then as  $U$  is open there exists a  $\delta_1 > 0$  such that  $x \in B_{\|\cdot\|}(x, \delta_1) \subseteq U$ . Define now

$$g: B_{\|\cdot\|}(x, \delta_1) \rightarrow Y \text{ by } g(y) = f(y) - f(x) - \sum_{i=1}^n D_i f(x)(y_i - x_i) \quad (14.139)$$

Then  $g(x) = f(x) - f(x) - \sum_{i=1}^n D_i f(x)(x_i - x_i) = 0 + \sum_{i=1}^n 0 = 0$  proving

$$g(x) = 0 \quad (14.140)$$

Now define  $L: X \rightarrow Y$  by  $L(y) = \sum_{i=1}^n D_i f(x)(y_i)$  so that using 14.139

$$g = f - C_{X, Y, f(x)} - L - C_{X, Y, \sum_{i=1}^n D_i f(x)(x_i)} \quad (14.141)$$

Then if  $\alpha, \beta \in \mathbb{K}$  and  $y, z \in X$  we have

$$\begin{aligned} L(\alpha \cdot y + \beta \cdot z) &= \sum_{i=1}^n D_i f(x)(\alpha \cdot y_i + \beta \cdot z_i) \\ &= \sum_{i=1}^n (\alpha \cdot D_i f(x)(y_i) + \beta \cdot D_i f(x)(z_i)) \\ &= \alpha \cdot \sum_{i=1}^n D_i f(x)(y_i) + \beta \cdot \sum_{i=1}^n D_i f(x)(z_i) \\ &= \alpha \cdot L(y) + \beta \cdot L(z) \end{aligned}$$

proving that  $L$  is linear. Further

$$\begin{aligned} \|L(y)\|_Y &= \left\| \sum_{i=1}^n D_i f(x)(y_i) \right\|_Y \\ &\leq \sum_{i=1}^n \|D_i f(x)(y_i)\|_Y \\ &\leq \sum_{i=1}^n (\|D_i f(x)\|_{L(X_i, Y)} \cdot \|y_i\|_i) \\ &\leq \sum_{i=1}^n (\|D_i f(x)\|_{L(X_i, Y)} \cdot \|y\|) \\ &\leq \left( \sum_{i=1}^n \|D_i f(x)\|_{L(X_i, Y)} \right) \cdot \|y\| \end{aligned}$$

So using 11.180 we have that

$$L \in L(X, Y) \quad (14.142)$$

hence  $L$  is  $C^\infty$  [see 14.75] with  $DL(y) = L$  [see 14.17] and thus by 14.35 that

$$D_i L(y) \text{ exists with } D_i L(y) = L \circ (i \rightarrow 0) \quad (14.143)$$

Further if  $k \in X_i$  then

$$\begin{aligned} (L \circ (i \rightarrow 0))(k) &= L((i \rightarrow 0)(k)) \\ &= \sum_{j=1}^n D_j f(x)((i \rightarrow 0)(k))_j \\ &\stackrel{14.30}{=} \sum_{j=1}^n D_j f(x) \cdot \delta_{i,j}(k) \\ &= D_i f(x)(k) \end{aligned}$$

proving that  $L \circ (i \rightarrow 0) = D_i f(x)$  which if we substitute this in 14.143 proves that

$$D_i L(y) = D_i f(x)$$

The above together with 14.141 and 14.37 proves that

$$\forall y \in B_{\|\cdot\|}(x, \delta_1) \text{ and } \forall i \in \{1, \dots, n\} D_i g(y) \text{ exists and } D_i g(y) = D_i f(y) - D_i f(x) \quad (14.144)$$

Also as for  $i \in \{1, \dots, n\}$   $D_i f$  is continuous there exists a  $\delta_{2,i} > 0$  such that if  $y \in B_{\|\cdot\|}(x, \delta_{2,i}) \subseteq B_{\|\cdot\|}(x, \delta_1)$  then  $\|D_i g(y)\|_{L(X, Y)} \stackrel{14.144}{=} \|D_i f(y) - D_i f(x)\|_{L(X, Y)} < \frac{\varepsilon}{n}$  hence if we take  $\delta_2 = \min(\{\delta_{2,i} | i \in \{1, \dots, n\}\})$  we have

$$\forall h \in X \text{ with } \|h\| < \delta_2 \text{ we have } \|D_i g(x+h)\| = \|D_i f(x+h) - D_i f(x)\| < \frac{\varepsilon}{n} \quad (14.145)$$

Now

$$\begin{aligned} g(x+h) &\stackrel{14.137}{=} g(x+h^n) - g(x+h^0) \\ &\stackrel{10.27}{=} \sum_{i=1}^n (g(x+h^i) - g(x+h^{i-1})) \end{aligned}$$

so that

$$\|g(x+h)\|_Y \leq \sum_{i=1}^n \|g(x+h^i) - g(x+h^{i-1})\|_Y \quad (14.146)$$

Now if  $\|h\| < \delta_2$  then  $\forall i \in \{0, \dots, n\}$  we have by 14.138 that  $\|h^i\| \leq \|h\| < \delta_2$  so that  $\forall i \in \{1, \dots, n\}$   $x+h^i, x+h^{i-1} \in B_{\|\cdot\|}(x, \delta_2)$  hence we have using a consequence of the main value theorem [see 14.116]

$$\begin{aligned} &\|g(x+h^i) - g(x+h^{i-1})\|_Y \leq \\ &\|x+h^i - x - h^{i-1}\| \cdot \sup(\{\|Dg((x+h^i) + t \cdot ((x+h^{i-1}) - (x+h^i)))\|_{L(X,Y)} | t \in [0, 1]\}) \\ &\|h^i - h^{i-1}\| \cdot \sup(\{\|Dg((x+h^{i-1}) + t \cdot ((x+h^i) - (x+h^{i-1})))\|_{L(X,Y)} | t \in [0, 1]\}) = \\ &\|h^i - h^{i-1}\| \cdot \sup(\{\|Dg(x+(h^{i-1} + t \cdot (h^i - h^{i-1})))\|_{L(X,Y)} | t \in [0, 1]\}) = \end{aligned} \quad (14.147)$$

Further

$$\begin{aligned} \|h^{i-1} + t \cdot (h^i - h^{i-1})\| &= \|(1-t) \cdot h^{i-1} + t \cdot h^i\| \\ &\leq_{t \in [0,1]} (1-t) \cdot \|h^{i-1}\| + t \cdot \|h^i\| \\ &\leq_{14.138} (1-t) \cdot \|h\| + t \cdot \|h\| \\ &= \|h\| \\ &< \delta_2 \end{aligned}$$

hence using 14.145 we have that  $\|Dg(a+(h^{i-1} + t \cdot (h^i - h^{i-1})))\|_{L(X,Y)} < \frac{\varepsilon}{n}$  so that by 14.147

$$\begin{aligned} \|g(x+h^i) - g(x+h^{i-1})\|_Y &< \|h^i - h^{i-1}\| \cdot \frac{\varepsilon}{n} \\ &= \left\| \sum_{j=1}^i (j \rightarrow 0)(h_j) - \sum_{j=1}^{i-1} (j \rightarrow 0)(h_j) \right\| \cdot \frac{\varepsilon}{n} \\ &= \|(i \rightarrow 0)(h_i)\| \cdot \frac{\varepsilon}{n} \\ &\stackrel{14.30}{=} \|h_i\|_i \cdot \frac{\varepsilon}{n} \\ &\leq \|h\| \cdot \frac{\varepsilon}{n} \end{aligned}$$

Substituting the above in 14.146 we have

$$\begin{aligned} \|g(x+h)\|_Y &< \sum_{i=1}^n \left( \|h\| \cdot \frac{\varepsilon}{n} \right) \\ &< \varepsilon \cdot \|h\| \end{aligned}$$

giving using the definition of  $g$  that

$$\left\| f(x+h) - f(x) - \sum_{i=1}^n D_i f(x)(h_i) \right\|_Y < \varepsilon \cdot \|h\| \text{ if } \|h\| < \delta_2$$

Which proves as  $L \in L(X, Y)$  [see 14.142] with  $L(h) = \sum_{i=1}^n D_i f(x)(h_i) = (\sum_{i=1}^n D_i f(x) \circ \pi_i)(h)$  that

$$Df(x) \text{ exists and } Df(x) = \sum_{i=1}^n D_i f(x) \circ \pi_i \quad (14.148)$$

Let  $x \in U$  then if  $y \in B_{\|\cdot\|}(x, \delta_2)$  we have

$$\begin{aligned}
 \|(Df(x) - Df(y))(h)\|_Y &= \left\| \left( \sum_{i=1}^n D_i f(x) - \sum_{i=1}^n D_i f(y) \right)(h) \right\|_Y \\
 &\leq \left\| \sum_{i=1}^n D_i f(x) - \sum_{i=1}^n D_i f(y) \right\|_{L(X, Y)} \cdot \|h\| \\
 &= \left\| \sum_{i=1}^n (D_i f(x) - D_i f(y)) \right\|_{L(X, Y)} \cdot \|h\| \\
 &\leq \sum_{i=1}^n \|D_i f(x) - D_i f(y)\|_{L(X, Y)} \cdot \|h\| \\
 &\stackrel{<_{14.145}}{\leq} \varepsilon \cdot \|h\|
 \end{aligned}$$

which proves by 11.185 that

$$\begin{aligned}
 \|Df(x) - Df(y)\|_{L(X, Y)} &\leq \sum_{i=1}^n \|D_i f(x) - D_i f(y)\|_{L(X, Y)} \\
 &\stackrel{<_{14.145}}{\leq} \varepsilon
 \end{aligned}$$

So  $Df$  is continuous on  $U$  and thus

$$f \text{ is of class } C^1 \quad \square$$

Combining 14.131 with theorem 14.134 gives the following result

**Theorem 14.135.** *Let  $n \in \mathbb{N}_0 \setminus \{1\}$ ,  $m \in \mathbb{N}_0$ ,  $\langle X_i, \|\cdot\|_i \rangle_{i \in \{1, \dots, n\}}$  a finite family of normed spaces,  $\langle Y, \|\cdot\|_Y \rangle$  a Banach space,  $V \subseteq Y$ ,  $\langle X, \|\cdot\| \rangle$  the normed space  $X = \prod_{i \in \{1, \dots, n\}} X_i$  with the maximum norm  $\|\cdot\|$  [see 11.82 and 11.338], a open set  $U \subseteq X$  and  $f: U \rightarrow V$  a function then the following are equivalent*

1.  $f$  is of class  $C^1$
  2.  $\forall i \in \{1, \dots, n\}$   $D_i^1 f = D_i f$  exists on  $U$  and is continuous
- Further if either (1) or (2) is true then  $\forall x \in U Df(x) = \sum_{i=1}^n D_i f(x) \circ \pi_i$

Let's now define partial derivative of higher order.

**Definition 14.136.** *Let  $n \in \mathbb{N}_0 \setminus \{1\}$ ,  $m \in \mathbb{N}_0$ ,  $\langle \mathbb{K}^n, \|\cdot\| \rangle$  be the vector space  $\mathbb{K}^n$  with the maximum norm,  $\langle Y, \|\cdot\|_Y \rangle$  a normed space,  $V \subseteq Y$ ,  $U \subseteq \mathbb{K}^n$  a open set,  $f: U \rightarrow Y$  then given  $\{i_j\}_{j \in \{1, \dots, m\}} \subseteq \{1, \dots, n\}$  we say that  $f$  has a **partial derivative**  $\partial_{i_m, \dots, i_1}^m f$  on  $U$  iff*

1. If  $m = 1$  then  $\partial_{i_1, \dots, i_1}^1 f$  exist on  $U$  if  $\forall x \in U f$  has a  $i_1$ -partial derivate at  $x$  [see 14.38] and  $\partial_{i_1, \dots, i_1}^1 f: U \rightarrow Y$  is defined by  $\partial_{i_1, \dots, i_1}^1 f(x) = \partial_i f(x)$ .
2. If  $m > 1$  then  $\partial_{i_m, \dots, i_1}^m f$  exists on  $U$  if  $\partial_{i_{m-1}, \dots, i_1}^{m-1} f$  exists on  $U$  and  $\forall x \in U$  the  $i_m$ -partial derivate  $\partial_{i_m}(\partial_{i_{m-1}, \dots, i_1}^{m-1} f)(x)$  exists. We define then  $\partial_{i_m, \dots, i_1}^m f: U \rightarrow Y$  by  $\partial_{i_m, \dots, i_1}^m f(x) = \partial_{i_1 m}(\partial_{i_{m-1}, \dots, i_1}^{m-1} f)(x)$

Just as there is a relation between the differential of a function on  $\mathbb{K}$  and its derivative we have also a relation between the partial differentials of a higher order on a function on  $\mathbb{K}^n$  and its partial derivatives.

**Theorem 14.137.** *Let  $n \in \mathbb{N}_0 \setminus \{1\}$ ,  $m \in \mathbb{N}_0$ ,  $\langle \mathbb{K}^n, \|\cdot\| \rangle$  be the vector space  $\mathbb{K}^n$  with the maximum norm,  $\langle Y, \|\cdot\|_Y \rangle$  a normed space,  $V \subseteq Y$ , a open set  $U \subseteq \mathbb{K}^n$ ,  $\{i_j\}_{j \in \{1, \dots, m\}} \subseteq \{1, \dots, n\}$  and  $f: U \rightarrow V$  a function such that  $D_{i_m, \dots, i_1}^m f$  exists on  $U$  then  $\partial_{i_m, \dots, i_1}^m f$  exists on  $U$  [and is continuous if  $D_{i_m, \dots, i_1}^m f$  is continuous] with  $\partial_{i_m, \dots, i_1}^m f(x) = (D_{i_m, \dots, i_1}^m f(x)) \left( \underbrace{1: \dots: 1}_m \right)$*

**Proof.** We prove this by induction so let  $\mathcal{S} = \left\{ m \in \mathbb{N}_0 \mid \{i_j\}_{j \in \{1, \dots, m\}} \subseteq \{1, \dots, n\}, f: U \rightarrow V \text{ such that } D_{i_m, \dots, i_n}^m f \text{ exist on } U \text{ then } \partial_{i_m, \dots, i_1}^m f \text{ exists on } U \text{ [and is continuous if } D_{i_m, \dots, i_1}^m f \text{ is continuous] with } \partial_{i_m, \dots, i_1}^m f(x) = (D_{i_m, \dots, i_1}^m f(x)) \left( \underbrace{1: \dots: 1}_r \right) \right\}$  then we have

**1  $\in \mathcal{S}$ .** Take  $\{i_j\}_{j \in \{1\}} \subseteq \{1, \dots, n\}$  let  $f: U \rightarrow V$  be such that  $D_{i_1}^1 f = D_{i_1} f$  exists on  $U$ . Using 14.39 we have that  $\forall x \in U \partial_{i_1} f(x)$  exists and  $\partial_{i_1} f(x) = D_{i_1} f(x)(1)$ , proving that  $\partial_{i_1, \dots, i_1}^1 f(x)$  exists and

$$\partial_{i_1, \dots, i_1}^1 f(x) = D_{i_1, \dots, i_1}^1 f(x) \left( \underbrace{1: \dots: 1}_1 \right) \quad (14.149)$$

Also we have that

$$\begin{aligned} \left( e_{\underbrace{1, \dots, 1}_1} \circ D_{i_1, \dots, i_1}^1 f \right) (x) &\stackrel{14.124}{=} e_{\underbrace{1, \dots, 1}_1} (D_{i_1, \dots, i_1}^1 f(x)) \\ &= D_{i_1, \dots, i_1}^1 f(x) \left( \underbrace{1: \dots: 1}_1 \right) \\ &\stackrel{14.149}{=} \partial_{i_1, \dots, i_1}^1 f(x) \end{aligned}$$

proving that  $e_{\underbrace{1, \dots, 1}_1} \circ D_{i_1, \dots, i_1}^1 f = \partial_{i_1, \dots, i_1}^1 f$ , so as  $e_{\underbrace{1, \dots, 1}_1}$  is continuous by 14.124 we have

$$\text{If } D_{i_1, \dots, i_1}^1 f \text{ is continuous then } \partial_{i_1, \dots, i_1}^1 f \text{ is continuous} \quad (14.150)$$

Finally 14.149 and 14.150 proves that  $1 \in \mathcal{S}$

**$m \in \mathcal{S} \Rightarrow m+1 \in \mathcal{S}$ .** Take  $\{i_j\}_{j \in \{1, \dots, m+1\}} \subseteq \{1, \dots, n\}$  and  $f: U \rightarrow V$  such that  $D_{i_{m+1}, \dots, i_1}^{m+1} f$  exists on  $U$ . By definition we have then that  $D_{i_m, \dots, i_1}^m f$  exists on  $U$  and has a  $i_m$ -partial differential with  $D_{i_{m+1}}(D_{i_m, \dots, i_1}^m f) = D_{i_{m+1}, \dots, i_1}^{m+1} f(x)$ . As  $m \in \mathcal{S}$  we have that

$$\partial_{i_m, \dots, i_1} f \text{ exist and } \forall x \in U \quad \partial_{i_m, \dots, i_1} f(x) = D_{i_m, \dots, i_1}^1 f(x) \left( \underbrace{1: \dots: 1}_m \right) \quad (14.151)$$

From  $e_{\underbrace{1, \dots, 1}_m} (D_{i_m, \dots, i_1}^m f(x)) = D_{i_m, \dots, i_1}^m f(x) \left( \underbrace{1: \dots: 1}_m \right) \stackrel{14.151}{=} \partial_{i_m, \dots, i_1}^m f(x)$  it follows that  $\partial_{i_m, \dots, i_1}^m f = e_{\underbrace{1, \dots, 1}_m} \circ D_{i_m, \dots, i_1}^m f$ . As  $e_{\underbrace{1, \dots, 1}_m}$  is  $C^\infty$  [see 14.124] we have using the chain rule for partial differentials [see 14.36] that  $\partial_{i_m, \dots, i_1}^m f$  has a  $i_{m+1}$ -partial differential with

$$\begin{aligned} D_{i_{m+1}}(\partial_{i_m, \dots, i_1}^m f)(x) &= D_{i_{m+1}} \left( e_{\underbrace{1, \dots, 1}_m} \circ D_{i_m, \dots, i_1}^m f \right) (x) \\ &\stackrel{14.36}{=} D \left( e_{\underbrace{1, \dots, 1}_m} (D_{i_m, \dots, i_1}^m f(x)) \right) \circ D_{i_{m+1}}(D_{i_m, \dots, i_1}^m f)(x) \\ &\stackrel{14.124}{=} e_{\underbrace{1, \dots, 1}_m} \circ D_{i_{m+1}}(D_{i_m, \dots, i_1}^m f)(x) \\ &= e_{\underbrace{1, \dots, 1}_m} \circ D_{i_{m+1}, \dots, i_1}^{m+1} f(x) \\ &= D_{i_{m+1}, \dots, i_1}^{m+1} f(x) \left( \underbrace{1: \dots: 1}_m \right) \end{aligned} \quad (14.152)$$

Applying 14.39 proves then that  $\partial_{i_{m+1}}(\partial_{i_m, \dots, i_1}^m f)$  exists on  $U$  with

$$\begin{aligned} \partial_{i_{m+1}, \dots, i_1}^{m+1} f(x) &\stackrel{\text{def}}{=} \partial_{m+1}(\partial_{i_m, \dots, i_1}^m f)(x) \\ &\stackrel{14.39}{=} D_{i_{m+1}}(\partial_{i_m, \dots, i_1}^m f)(x)(1) \\ &\stackrel{14.152}{=} D_{i_{m+1}, \dots, i_1}^{m+1} f(x) \left( \underbrace{1: \dots: 1}_m \right)(1) \\ &\stackrel{\text{def}}{=} D_{i_{m+1}, \dots, i_1}^{m+1} f(x) \left( \underbrace{1: \dots: 1}_{m+1} \right) \end{aligned} \quad (14.153)$$

Further as  $e_{\underbrace{1:\dots:1}_m}(D_{i_{m+1},\dots,i_1}^{m+1}f)(x) = D_{i_{m+1},\dots,i_1}^{m+1}f(x)\left(\underbrace{1:\dots:1}_{m+1}\right) \stackrel{14.153}{=} \partial_{i_{m+1},\dots,i_1}^{m+1}f(x)$  it follows that  $e_{\underbrace{1:\dots:1}_m} \circ D_{i_{m+1},\dots,i_1}^{m+1}f = \partial_{i_{m+1},\dots,i_1}^{m+1}f$ . Hence as  $e_{\underbrace{1:\dots:1}_m}$  is continuous by 14.124 we have

$$\partial_{i_{m+1},\dots,i_1}^{m+1}f \text{ is continuous if } D_{i_{m+1},\dots,i_1}^{m+1}f \text{ is continuous} \quad (14.154)$$

Finally using 14.153 and 14.154 it follows that  $m+1 \in \mathcal{S}$   $\square$

**Corollary 14.138.** Let  $n \in \mathbb{N}_0 \setminus \{1\}$ ,  $m \in \mathbb{N}_0$ ,  $\langle \mathbb{K}^n, \|\cdot\| \rangle$  be the vector space  $\mathbb{K}^n$  with the maximum norm,  $\langle Y, \|\cdot\|_Y \rangle$  a normed space,  $V \subseteq Y$ , a open set  $U \subseteq \mathbb{K}^n$ ,  $f: U \rightarrow V$  a function that is  $m$ -differentiable then  $\forall \{i_j\}_{j \in \{1,\dots,m\}} \subseteq \{1,\dots,n\}$   $\partial_{i_m,\dots,i_1}^m f$  exists on  $U$  and is continuous if  $f$  is of class  $C^m$ . Further  $\forall x \in U$  we have  $\partial_{i_m,\dots,i_1}^m f(x) = D_{i_m,\dots,i_1}^m f(x)\left(\underbrace{1:\dots:1}_m\right) = D^m f(x)(e_{i_m}:\dots:e_{i_1})$  where  $\{e_i\}_{i \in \{1,\dots,n\}}$  is the canonical base on  $\mathbb{K}^n$  [defined by  $(e_i)_j = \delta_{i,j}$ ]

**Proof.** Using 14.132 we have that  $D_{i_m,\dots,i_1}^m f$  exists on  $U$  with  $D_{i_m,\dots,i_1}^m f(x)(h_m:\dots:h_1) = D^m f(x)((i_m \rightarrow 0)(h_m):\dots:(i_1 \rightarrow 0)(h_1))$  [and is continuous if  $f$  is of class  $C^m$ ]. Using the previous theorem [see 14.137] we have then that  $\partial_{i_m,\dots,i_1}^m f$  exists on  $U$  [and is continuous if  $f$  is of class  $C^m$ ] with

$$\begin{aligned} \partial_{i_m,\dots,i_1}^m f(x) &\stackrel{14.137}{=} D_{i_m,\dots,i_1}^m f(x)\left(\underbrace{1:\dots:1}_n\right) \\ &\stackrel{14.132}{=} D^m f(x)((i_m \rightarrow 0)(1):\dots:(i_1 \rightarrow 0)(1)) \\ &= D^m f(x)(e_{i_m}:\dots:e_{i_1}) \left[ \text{as } ((i_j \rightarrow 0)(1))_k \stackrel{14.30}{=} \delta_{j,k}(1) = \delta_{j,k} \right] \end{aligned} \quad \square$$

## 14.7 Inverse function theorem

Remember the concept of a toplinear isomorphism [see 11.243] a open, continuous linear isomorphism. Using this definition it is trivial to prove the next proposition.

**Proposition 14.139.** Let  $\langle X, \|\cdot\|_X \rangle$ ,  $\langle Y, \|\cdot\|_Y \rangle$  be normed spaces then  $L \in L(X, Y)$  is a toplinear isomorphism [see 11.243] if also  $L^{-1}$  is continuous.

**Proof.** This is trivial using the definition [see 11.243].  $\square$

In Banach spaces every linear continuous isomorphism turns out to be a toplinear isomorphism.

**Corollary 14.140.** If  $\langle X, \|\cdot\|_X \rangle$ ,  $\langle Y, \|\cdot\|_Y \rangle$  are Banach spaces then if  $L: X \rightarrow Y$  is a linear continuous isomorphism [or  $L \in L(X, Y)$  and is a isomorphism] then  $L$  is a toplinear isomorphism.

**Proof.** This is a consequence of 11.421.  $\square$

**Definition 14.141.** If  $\langle X, \|\cdot\| \rangle$  is a normed space then  $\mathcal{GL}(X) = \{L \in L(X, X) | L \text{ is toplinear}\} \subseteq L(X, X)$

**Theorem 14.142.** If  $\langle X, \|\cdot\| \rangle$  is a Banach space then  $\mathcal{GL}(X)$  is a group with the composition operator as the group operator. The identity operator  $\mathbb{1}_X$  is the neutral element and the inverse of  $L \in \mathcal{GL}(X)$  is  $L^{-1}$ ,

**Proof.** If  $L_1, L_2 \in \mathcal{GL}(X)$  then  $L_1 \circ L_2$  is also a isomorphism that is linear and continuous. As  $X$  is a Banach space we have using 11.421 that  $(L_1 \circ L_2)^{-1}$  is continuous. So  $L_1 \circ L_2 \in \mathcal{GL}(X)$  proving that  $\circ: X \times X \rightarrow X$  defined by  $(L_1, L_2) \rightarrow L_1 \circ L_2$  is well defined.

1. **(identity)** Consider  $\mathbb{1}_X: X \rightarrow X$  which is evidently linear, continuous and a isomorphism with a continuous inverse. If  $L \in \mathcal{GL}(X)$  then  $L \circ \mathbb{1}_X = L = \mathbb{1}_X \circ L$

2. **(inverse element)** If  $L \in GL(X)$  then by definition  $L^{-1} \in GL(X)$  and  $L \circ L^{-1} = 1_X = L^{-1} \circ L$
3. **(associativity)** This is trivial as composition of functions is associative.  $\square$

**Definition 14.143.** Let  $\langle X, \|\cdot\| \rangle$  be a normed space,  $H: X \rightarrow X$  a continuous linear function and  $n \in \mathbb{N}_0$  then we define  $H^n: X \rightarrow X$  recursively as follows

$$\begin{aligned} H^1 &= H \\ H^{n+1} &= H \circ H^n \end{aligned}$$

**Note 14.144.** Let  $\langle X, \|\cdot\| \rangle$  be a normed space,  $H \in L(X, X)$  then  $-H \in L(X, X)$  and  $\forall n \in \mathbb{N}_0$  we have  $(-H)^n = (-1)^n \cdot H^n$

**Proof.** Let  $S = \{n \in \mathbb{N}_0 \mid \text{if } H \in L(X, X) \text{ then } (-H)^n = (-1)^n \cdot H\}$  then:

**1  $\in S$ .** Let  $x \in X$  then

$$\begin{aligned} (-H)^1(x) &= (-H(x)) \\ &= (-1)^1 \cdot H(x) \\ &= ((-1)^1 \cdot H)(x) \end{aligned}$$

proving that  $(-H)^1 = (-1)^1 \cdot H$  so that  $1 \in S$

**$n \in S \Rightarrow n+1 \in S$ .** then

$$\begin{aligned} (-H)^{n+1}(x) &= ((-H) \circ (-H)^n)(x) \\ &= (-H((-H)^n(x))) \\ &= (-1) \cdot (H((-H)^n(x))) \\ &\stackrel{n \in S}{=} -1 \cdot (H((-1)^n \cdot H^n(x))) \\ &= (-1) \cdot (-1)^n (H(H^n(x))) \\ &= (-1)^{n+1} \cdot H^{n+1}(x) \\ &= ((-1)^{n+1} \cdot H^{n+1})(x) \end{aligned}$$

proving that  $n+1 \in S$ .  $\square$

**Lemma 14.145.** Let  $\langle X, \|\cdot\|_X \rangle$  be a normed vector space and  $H \in L(X, X)$ ,  $n \in \mathbb{N}_0$  then if  $\|\cdot\|$  is the norm on  $L(X, X)$  we have

1.  $H^n: X \rightarrow X$  is linear and continuous with  $\forall x \in X \models \|H^n(x)\|_X \leq \|H\|^n \|x\|_X$  [or using 11.184  $\|H^n\| \leq \|H\|^n$ ]
2.  $\sum_{i=1}^n H^i$  is linear and continuous with  $\forall x \in X \models \|(\sum_{i=1}^n H^i)(x)\|_X \leq (\sum_{i=1}^n \|H\|^i) \cdot \|x\|_X$  [or using 11.184  $\|\sum_{i=1}^n H^i\| \leq \sum_{i=1}^n \|H\|^i$ ]

**Proof.**

1. We use induction so let  $S = \{n \in \mathbb{N}_0 \mid H^n \text{ is linear and continuous with } \forall x \in X \ \|H^n(x)\|_X \leq \|H\|^n \cdot \|x\|\}$  then we have

**1  $\in S$ .** First  $H^1(\alpha \cdot x + \beta \cdot y) = H(\alpha \cdot x + \beta \cdot y) \stackrel{H \text{ is linear}}{=} \alpha \cdot H(x) + \beta \cdot H(y) = \alpha \cdot H^1(x) + \beta \cdot H^1(y)$  proving linearity of  $H^1$ . Secondly  $\|H^1(x)\|_X = \|H(x)\|_X \leq_{H \in L(X, X)} \|H\| \cdot \|x\|_X$  proving continuity. Hence  $1 \in S$

**$n \in S \Rightarrow n+1 \in S$ .** Then

$$\begin{aligned} H^{n+1}(\alpha \cdot x + \beta \cdot y) &= (H \circ H^n)(\alpha \cdot x + \beta \cdot y) \\ &= H(H^n(\alpha \cdot x + \beta \cdot y)) \\ &\stackrel{n \in S}{=} H(\alpha \cdot H^n(x) + \beta \cdot H^n(y)) \\ &= \alpha \cdot H(H^n(x)) + \beta \cdot H(H^n(y)) \\ &= \alpha \cdot H^{n+1}(x) + \beta \cdot H^{n+1}(y) \end{aligned}$$

and

$$\begin{aligned}
 \|H^{n+1}(x)\|_X &= \|H(H^n(x))\|_X \\
 &\leq \|H\| \cdot \|H^n(x)\|_X \\
 &\leq_{n \in \mathcal{S}} \|H\| \cdot \|H\|^n \cdot \|x\|_X \\
 &= \|H\|^{n+1} \cdot \|x\|_X
 \end{aligned}$$

proving that  $n+1 \in \mathcal{S}$

2. First note that

$$\begin{aligned}
 \left( \sum_{i=1}^n H^i \right) (\alpha \cdot x + \beta \cdot y) &\stackrel{\text{def}}{=} \sum_{i=1}^n H^i(\alpha \cdot x + \beta \cdot y) \\
 &\stackrel{(1)}{=} \sum_{i=1}^n (\alpha \cdot H^i(x) + \beta \cdot H^i(y)) \\
 &= \alpha \cdot \sum_{i=1}^n H^i(x) + \beta \cdot \sum_{i=1}^n H^i(y) \\
 &= \left( \alpha \cdot \sum_{i=1}^n H^i \right) (x) + \beta \cdot \left( \sum_{i=1}^n H^i \right) (y)
 \end{aligned}$$

and secondary

$$\begin{aligned}
 \left\| \left( \sum_{i=1}^n H^i \right) (x) \right\|_X &= \left\| \sum_{i=1}^n H^i(x) \right\|_X \\
 &\leq \sum_{i=1}^n \|H^i(x)\| \\
 &\leq_{(1)} \sum_{i=1}^n (\|H^i\| \cdot \|x\|_X) \\
 &= \left( \sum_{i=1}^n \|H^i\| \right) \cdot \|x\|_X
 \end{aligned}$$

□

**Definition 14.146.** Let  $\langle X, \|\cdot\|_X \rangle$  be a normed space and  $H \in L(X, X)$ . Then  $\sum_{i=1}^{\infty} H^i$  converges in  $L(X, Y)$  if  $\forall x \in X$  we have that  $\sum_{i=1}^{\infty} H^i(x)$  converges in  $X$ . Further if  $\sum_{i=1}^{\infty} H^i$  converges then  $\sum_{i=1}^{\infty} H^i: X \rightarrow Y$  is defined by  $(\sum_{i=1}^{\infty} H^i)(x) = \sum_{i=1}^{\infty} H^i(x)$ .

**Lemma 14.147.** Let  $\langle X, \|\cdot\|_X \rangle$  be a normed space  $L, H \in L(X, Y)$  and  $\sum_{i=1}^{\infty} H^i$  converges then  $\forall x \in X$  we have  $\lim_{i \rightarrow \infty} (L(x) + \sum_{i=1}^n H^i(x)) = (L + \sum_{i=1}^{\infty} H^i)(x)$

**Proof.**

$$\begin{aligned}
 \left( L + \sum_{i=1}^{\infty} H^i \right) (x) &= L(x) + \left( \sum_{i=1}^{\infty} H^i \right) (x) \\
 &\stackrel{\text{definition}}{=} L(x) + \lim_{n \rightarrow \infty} \left( \sum_{i=1}^n H^i(x) \right) \\
 &\stackrel{11.353}{=} \lim_{n \rightarrow \infty} \left( L(x) + \sum_{i=1}^n H^i(x) \right)
 \end{aligned}$$

□

**Theorem 14.148.** Let  $\langle X, \|\cdot\|_X \rangle$  be a Banach space and  $H \in B_{\|\cdot\|}(0, 1) = \{H \in L(X, X) \mid \|H\| < 1\}$  then  $\mathbb{1}_X - H \in \mathcal{GL}(X)$ . Further  $\sum_{i=1}^{\infty} H^i$  converges in  $L(X, X)$  and  $(\mathbb{1}_X - H)^{-1} = \mathbb{1}_X + \sum_{i=1}^{\infty} H^i$

**Proof.** Using the previous lemma (see 14.145) gives

$$\forall x \in X, n \in \mathbb{N}_0 \text{ we have } \|H^n(x)\|_X \leq \|H\|^n \cdot \|x\|_X \quad (14.155)$$

Using the fact that  $\|H\| < 1$  and 11.415 it follows that  $\sum_{i=1}^{\infty} \|H\|^i$  is convergent and  $\sum_{i=1}^{\infty} \|H\|^i = \frac{\|H\|}{1 - \|H\|}$ . Further using 11.381 we have that  $\sum_{i=1}^{\infty} (\|H\|^i \cdot \|x\|_X)$  is convergent to  $\frac{\|H\| \cdot \|x\|_X}{1 - \|H\|}$ . Next using the fact that  $X$  is a Banach space, 14.155 and dominant convergence [see 11.393] we have  $\forall x \in X$  that

$$\sum_{i=1}^{\infty} H^i(x) \text{ is convergent and } \left\| \sum_{i=1}^{\infty} H^i(x) \right\|_X \leq \frac{\|H\| \cdot \|x\|_X}{1 - \|H\|} \quad (14.156)$$

Further by 11.376 we have further

$$\lim_{n \rightarrow \infty} H^n(x) = 0 \quad (14.157)$$

So we can define the function  $(\mathbb{1}_X + \sum_{i=1}^{\infty} H^i): X \rightarrow X$  by

$$\forall x \in X \models \left( \mathbb{1}_X + \sum_{i=1}^{\infty} H^i \right)(x) = x + \sum_{i=1}^{\infty} H^i(x) \quad (14.158)$$

We show now that  $(\mathbb{1}_X + \sum_{i=1}^{\infty} H^i)$  is linear

$$\begin{aligned} \left( \mathbb{1}_X + \sum_{i=1}^{\infty} H^i \right)(\alpha \cdot x + \beta \cdot y) &= \alpha \cdot x + \beta \cdot y + \sum_{i=1}^{\infty} H^i(\alpha \cdot x + \beta \cdot y) \\ &\stackrel{14.145}{=} \alpha \cdot x + \beta \cdot y + \sum_{i=1}^{\infty} (\alpha \cdot H^i(x) + \beta \cdot H^i(y)) \\ &\stackrel{11.381}{=} \alpha \cdot x + \beta \cdot y + \alpha \cdot \sum_{i=1}^{\infty} H^i(x) + \beta \cdot \sum_{i=1}^{\infty} H^i(y) \\ &= \alpha \cdot \left( \mathbb{1}_X + \sum_{i=1}^{\infty} H^i \right)(x) + \beta \cdot \left( \mathbb{1}_X + \sum_{i=1}^{\infty} H^i \right)(y) \end{aligned}$$

Next we prove continuity of  $(\mathbb{1}_X + \sum_{i=1}^{\infty} H^i)$

$$\begin{aligned} \left\| \left( \mathbb{1}_X + \sum_{i=1}^{\infty} H^i \right)(x) \right\|_X &= \left\| \mathbb{1}_X(x) + \sum_{i=1}^{\infty} H^i(x) \right\|_X \\ &\leq \|x\|_X + \left\| \sum_{i=1}^{\infty} H^i(x) \right\|_X \\ &\stackrel{14.156}{\leq} \|x\|_X + \frac{\|H\|}{1 - \|H\|} \cdot \|x\|_X \\ &= \left( 1 + \frac{\|H\|}{1 - \|H\|} \right) \cdot \|x\|_X \\ &= \frac{1}{1 - \|H\|} \cdot \|x\|_X \end{aligned}$$

proving by 11.180 and 11.184 that

$$\left( \mathbb{1}_X + \sum_{i=1}^{\infty} H^i \right) \in L(X, X) \wedge \left\| \left( \mathbb{1}_X + \sum_{i=1}^{\infty} H^i \right) \right\| \leq \frac{1}{1 - \|H\|} \quad (14.159)$$

As  $\mathbb{1}_X \in L(X, X)$  we have that

$$\sum_{i=1}^{\infty} H^i = \left( \mathbb{1}_X + \sum_{i=1}^{\infty} H^i \right) - \mathbb{1}_X \in L(X, X) \quad (14.160)$$

Next we prove by induction on  $n$  that

$$(\mathbb{1}_X - H) \circ \left( \mathbb{1}_X + \sum_{i=1}^n H^i \right) = (\mathbb{1}_X - H^{n+1}) \quad (14.161)$$

**Proof.** Let  $S = \{n \in \mathbb{N}_0 \mid (\mathbb{1}_X - H) \circ (\mathbb{1}_X + \sum_{i=1}^n H^i) = (\mathbb{1}_X - H^{n+1})\}$  then we have:

**1**  $\in \mathcal{S}$ .  $\forall x \in X$  we have

$$\begin{aligned} \left( (\mathbb{1}_X - H) \circ \left( \mathbb{1}_X + \sum_{i=1}^1 H^i \right) \right) (x) &= (\mathbb{1}_X - H) \left( \left( \mathbb{1}_X + \sum_{i=1}^1 H^i \right) (x) \right) \\ &= (\mathbb{1}_X - H) \left( \mathbb{1}_X(x) + \sum_{i=1}^1 H^i(x) \right) \\ &= (\mathbb{1}_X - H)(x + H(x)) \\ &= x + H(x) - H(x) - H^2(x) \\ &= (\mathbb{1}_X - H^2)(x) \end{aligned}$$

proving that  $1 \in S$ .

**n**  $\in \mathcal{S} \Rightarrow n + 1 \in \mathcal{S}$ . Then  $\forall x \in X$  we have

$$\begin{aligned} \left( (\mathbb{1}_X - H) \circ \left( \mathbb{1}_X + \sum_{i=1}^{n+1} H^i \right) \right) &= \\ (\mathbb{1}_X - H) \left( \left( \mathbb{1}_X + \sum_{i=1}^{n+1} H^i \right) (x) \right) &= \\ (\mathbb{1}_X - H) \left( \left( \mathbb{1}_X + \sum_{i=1}^n H^i \right) (x) + H^{n+1}(x) \right) &\stackrel{\text{linearity of } \mathbb{1}_X - H}{=} \\ (\mathbb{1}_X - H) \left( \left( \mathbb{1}_X + \sum_{i=1}^n H^i \right) (x) \right) + (\mathbb{1}_X - H)(H^{n+1}(x)) &\stackrel{n \in S}{=} \\ (\mathbb{1}_X - H^{n+1})(x) + H^{n+1}(x) - H^{(n+1)+1}(x) &= \\ (\mathbb{1}_X - H^{(n+1)+1})(x) &= \end{aligned}$$

proving that  $n + 1 \in S$ .

Using induction we have that  $S = \mathbb{N}_0$  so that 14.161 is valid.  $\square$

Next by induction again we prove that

$$\left( \mathbb{1}_X + \sum_{i=1}^n H^i \right) \circ (\mathbb{1}_X - H) = \mathbb{1}_X - H^{n+1} \quad (14.162)$$

**Proof.** Let  $S = \{n \in \mathbb{N}_0 \mid (\mathbb{1}_X + \sum_{i=1}^n H^i) \circ (\mathbb{1}_X - H) = \mathbb{1}_X - H^{n+1}\}$  then we have :

**1**  $\in \mathcal{S}$ .  $\forall x \in X$

$$\begin{aligned} \left( \left( \mathbb{1}_X + \sum_{i=1}^n H^i \right) \circ (\mathbb{1}_X - H) \right) (x) &= \left( \mathbb{1}_X + \sum_{i=1}^1 H^i \right) ((\mathbb{1}_X - H)(x)) \\ &= \left( \mathbb{1}_X + \sum_{i=1}^1 H^i \right) (x - H(x)) \\ &= (x - H(x) + H(x - H(x))) \\ &= x - H(x) + H(x) - H^2(x) \\ &= x - H^2(x) \\ &= (\mathbb{1}_X - H^2)(x) \end{aligned}$$

proving that  $1 \in S$ .

$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$ . Then we have  $\forall x \in X$  we have

$$\begin{aligned}
\left( \left( \mathbb{1}_X + \sum_{i=1}^{n+1} H^i \right) \circ (\mathbb{1}_X - H) \right) (x) &= \left( \mathbb{1}_X + \sum_{i=1}^{n+1} H^i \right) ((\mathbb{1}_X - H)(x)) \\
&= \left( \mathbb{1}_X + \sum_{i=1}^{n+1} H^i \right) (x - H(x)) \\
&= \left( \mathbb{1}_X + \sum_{i=1}^n H^i + H^{n+1} \right) (x - H(x)) \\
&= \left( \mathbb{1}_X + \sum_{i=1}^n H^i \right) (x - H(x)) + H^{n+1}(x - H(x)) \\
&= \left( \left( \mathbb{1}_X + \sum_{i=1}^n H^i \right) \circ (\mathbb{1}_X - H) \right) (x) + H^{n+1}(x - H(x)) \\
&\stackrel{n \in S}{=} (\mathbb{1}_X - H^{n+1})(x) + H^{n+1}(x) - H^{(n+1)+1}(x) \\
&= x - H^{n+1}(x) + H^{n+1}(x) - H^{(n+1)+1}(x) \\
&= x - H^{(n+1)+1}(x) \\
&= (\mathbb{1}_X - H^{(n+1)+1})(x)
\end{aligned}$$

proving that  $n+1 \in S$ .

Using induction we have that  $S = \mathbb{N}_0$  so that 14.162 is indeed valid.  $\square$

Now if  $\varepsilon > 0$ ,  $x \in X$  then by 14.157 there exists a  $N \in \mathbb{N}_0$  such that if  $n \geq N$  then  $\|H^{n+1}(x)\|_X < \varepsilon$ , so using 14.161 and 14.162 we have

$$\begin{aligned}
\left\| (\mathbb{1}_X - H) \left( \left( \mathbb{1}_X + \sum_{i=1}^n H^i \right) (x) \right) - \mathbb{1}_X(x) \right\|_X &\stackrel{14.161}{=} \|\mathbb{1}_X(x) - H^{n+1}(x) - \mathbb{1}_X(x)\|_X \\
&= \|-H^{n+1}(x)\|_X \\
&= \|H^{n+1}(x)\|_X \\
&< \varepsilon
\end{aligned}$$

and

$$\begin{aligned}
\left\| \left( \mathbb{1}_X + \sum_{i=1}^n H^i \right) ((\mathbb{1}_X - H)(x)) - \mathbb{1}_X(x) \right\|_X &\stackrel{14.162}{=} \|\mathbb{1}_X(x) - H^{n+1}(x) - \mathbb{1}_X(x)\|_X \\
&= \|-H^{n+1}(x)\|_X \\
&= \|H^{n+1}(x)\|_X \\
&< \varepsilon
\end{aligned}$$

proving that

$$\forall x \in X \models \lim_{n \rightarrow \infty} \left[ (\mathbb{1}_X - H) \left( \left( \mathbb{1}_X + \sum_{i=1}^n H^i \right) (x) \right) \right] = \mathbb{1}_X(x) \quad (14.163)$$

and

$$\forall x \in X \models \lim_{n \rightarrow \infty} \left[ \left( \mathbb{1}_X + \sum_{i=1}^n H^i \right) ((\mathbb{1}_X - H)(x)) \right] = \mathbb{1}_X(x) \quad (14.164)$$

Now as  $(\mathbb{1}_X - H)$  is continuous we have

$$\begin{aligned} \mathbb{1}_X(x) &\stackrel{14.163}{=} \lim_{n \rightarrow \infty} \left[ (\mathbb{1}_X - H) \left( \left( \mathbb{1}_X + \sum_{i=1}^n H^i \right) (x) \right) \right] \\ &\stackrel{11.334}{=} (\mathbb{1}_X - H) \left( \lim_{n \rightarrow \infty} \left( \left( \mathbb{1}_X + \sum_{i=1}^n H^i \right) (x) \right) \right) \\ &= (\mathbb{1}_X - H) \left( \lim_{n \rightarrow \infty} \left( \mathbb{1}_X(x) + \sum_{i=1}^n H^i(x) \right) \right) \\ &\stackrel{11.353}{=} (\mathbb{1}_X - H) \left( \mathbb{1}_X(x) + \lim_{n \rightarrow \infty} \sum_{i=1}^n H^i(x) \right) \\ &= (\mathbb{1}_X - H) \left( \mathbb{1}_X(x) + \sum_{i=1}^{\infty} H^i(x) \right) \\ &\stackrel{\text{def}}{=} (\mathbb{1}_X - H) \left( \mathbb{1}_X(x) + \left( \sum_{i=1}^{\infty} H^i \right) (x) \right) \\ &= (\mathbb{1}_X - H) \left( \left( \mathbb{1}_X + \sum_{i=1}^{\infty} H^i \right) (x) \right) \\ &= \left( (\mathbb{1}_X - H) \circ \left( \mathbb{1}_X + \sum_{i=1}^{\infty} H^i \right) \right) (x) \end{aligned}$$

proving that

$$\mathbb{1}_X = (\mathbb{1}_X - H) \circ \left( \mathbb{1}_X + \sum_{i=1}^{\infty} H^i \right) \quad (14.165)$$

Also

$$\begin{aligned} \mathbb{1}_X(x) &\stackrel{14.164}{=} \lim_{n \rightarrow \infty} \left[ \left( \mathbb{1}_X + \sum_{i=1}^n H^i \right) ((\mathbb{1}_X - H)(x)) \right] \\ &= \lim_{n \rightarrow \infty} \left[ (\mathbb{1}_X((\mathbb{1}_X - H)(x))) + \left( \sum_{i=1}^n H^i \right) ((\mathbb{1}_X - H)(x)) \right] \\ &= \lim_{n \rightarrow \infty} \left[ (\mathbb{1}_X((\mathbb{1}_X - H)(x))) + \sum_{i=1}^n H^i((\mathbb{1}_X - H)(x)) \right] \\ &\stackrel{14.147}{=} \left( \mathbb{1}_X((\mathbb{1}_X - H)(x)) + \lim_{n \rightarrow \infty} \sum_{i=1}^n H^i((\mathbb{1}_X - H)(x)) \right) \\ &= \left( \mathbb{1}_X((\mathbb{1}_X - H)(x)) + \sum_{i=1}^{\infty} H^i((\mathbb{1}_X - H)(x)) \right) \\ &\stackrel{\text{def}}{=} \left( \mathbb{1}_X((\mathbb{1}_X - H)(x)) + \left( \sum_{i=1}^{\infty} H^i \right) ((\mathbb{1}_X - H)(x)) \right) \\ &= \left( \mathbb{1}_X + \sum_{i=1}^{\infty} H^i \right) ((\mathbb{1}_X - H)(x)) \\ &= \left[ \left( \mathbb{1}_X + \sum_{i=1}^{\infty} H^i \right) \circ (\mathbb{1}_X - H) \right] (x) \end{aligned}$$

proving that

$$\mathbb{1}_X = \left( \mathbb{1}_X + \sum_{i=1}^{\infty} H^i \right) \circ (\mathbb{1}_X - H) \quad (14.166)$$

Using 14.165 and 14.166 we have that  $\mathbb{1}_X - H$  is a bijection with  $(\mathbb{1}_X - H)^{-1} = \mathbb{1}_X + \sum_{i=1}^{\infty} H^i$ . Further as  $\mathbb{1}_X, H \in L(X, X)$  we have that  $\mathbb{1}_X - H$  is linear [hence a isomorphism] and continuous. Using 14.140 it follows that  $(\mathbb{1}_X - H)$  is toplinear. This together with 14.160 proves that

$$\sum_{i=1}^{\infty} H^i \in L(X, X) \wedge \mathbb{1}_X - H \in \text{GL}(X) \wedge (\mathbb{1}_X - H)^{-1} = \mathbb{1}_X + \sum_{i=1}^{\infty} H^i \quad \square$$

**Corollary 14.149.** Let  $\langle X, \|\cdot\| \rangle$  be a Banach space,  $H \in B_{\|\cdot\|}(0, 1)$  then  $\mathbb{1}_X + H \in \text{GL}(X)$ ,  $\sum_{i=1}^{\infty} (-1)^i \cdot H^i$  converges in  $L(X, X)$  with  $(\mathbb{1}_X + H)^{-1} = \mathbb{1}_X + \sum_{i=1}^{\infty} (-1)^i \cdot H^i$

**Proof.** This follows from the previous theorem [see 14.149] and 14.144 by using  $-H$  and the fact that  $\|-H\| = \|H\| < 1$ .  $\square$

**Theorem 14.150.** Let  $\langle X, \|\cdot\| \rangle$  be a Banach space then we have

1.  $\text{GL}(X)$  is a open set in  $L(X, X)$
2. The function  $\tau: \text{GL}(X) \rightarrow L(X, X)$  defined by  $\tau(L) = L^{-1}$  is differentiable of class  $C^\infty$  and  $D\tau(L) \in L(L(X, X), L(L(X, X), L(X, X)))$  is defined by  $D\tau(L)(H) = -L^{-1} \circ H \circ L^{-1}$

**Proof.**

1. Let  $L \in \text{GL}(X)$  then for any  $H \in L(X, X)$  we can write

$$\begin{aligned} H &= L + (H - L) \\ &= L \circ \mathbb{1}_X + (L \circ L^{-1}) \circ (H - L) \\ &\stackrel{11.203}{=} L \circ (\mathbb{1}_X + L^{-1} \circ (H - L)) \end{aligned} \quad (14.167)$$

Now according to 11.191 we have  $\|L^{-1} \circ (H - L)\| \leq \|L^{-1}\| \cdot \|H - L\|$ . Take  $\delta = \max(1, \|L^{-1}\|) > 0$  then if  $\|H - L\| < \frac{1}{\delta}$  we have  $\|L^{-1} \circ (H - L)\| < \|L^{-1}\| \cdot \frac{1}{\delta} \leq 1$  and using the previous corollary [see 14.149] we have then that  $(\mathbb{1}_X + L^{-1} \circ (H - L)) \in \text{GL}(X)$  and as  $L \in \text{GL}(X)$  we have by 14.142 that  $H = L \circ (\mathbb{1}_X + L^{-1} \circ (H - L)) \in \text{GL}(X)$ . So if  $H \in B_{\|\cdot\|}(L, \frac{1}{\delta})$  then  $\|H - L\| < \frac{1}{\delta}$  giving  $H \in \text{GL}(X)$  meaning that  $B_{\|\cdot\|}(L, \frac{1}{\delta}) \subseteq \text{GL}(X)$  proving that

$\text{GL}(X)$  is a open set in  $L(X, X)$

2. If  $H \in B_{\|\cdot\|}(\mathbb{1}_X, 1)$  then  $\|H - \mathbb{1}_X\| < 1$  so that by 14.149  $H = (H - \mathbb{1}_X) + \mathbb{1}_X \in \text{GL}(X)$  proving that

$$B_{\|\cdot\|}(\mathbb{1}_X, 1) \subseteq \text{GL}(X) \quad (14.168)$$

Let  $H \in B_{\|\cdot\|}(0, 1)$  then using 14.149 we have that  $\sum_{i=1}^{\infty} (-1)^i \cdot \|H\|^i$  exists and using 11.379 we have

$$\sum_{i=2}^n (-1)^i \cdot H^i \text{ converges in } L(X, X) \text{ so } \sum_{i=2}^n (-1)^i \cdot H^i \in L(X, X) \quad (14.169)$$

Define now

$$\varepsilon(H) = \begin{cases} C_{X, X, 0} \text{ if } H = 0 \\ \frac{1}{\|H\|} \cdot \sum_{i=2}^{\infty} (-1)^i \cdot H^i \text{ if } H \neq 0 \end{cases}$$

Given  $x, y \in X$  and  $\alpha, \beta \in \mathbb{K}$  we have

$$\begin{aligned}
 \varepsilon(H)(\alpha \cdot x + \beta \cdot y) &= \\
 \begin{cases} C_{X,X,0}(\alpha \cdot x + \beta \cdot y) & \text{if } H = 0 \\ \frac{1}{\|H\|} \cdot (\sum_{i=2}^{\infty} (-1)^i \cdot H^i)(\alpha \cdot x + \beta \cdot y) & \text{if } H \neq 0 \end{cases} & C_{X,0} \in \overline{\overline{L}}(X, X) \\
 \begin{cases} \alpha \cdot C_{X,X,0}(x) + \beta \cdot C_{X,X,0}(y) & \text{if } H = 0 \\ \frac{1}{\|H\|} \cdot (\sum_{i=2}^{\infty} (-1)^i \cdot H^i)(\alpha \cdot x + \beta \cdot y) & \text{if } H \neq 0 \end{cases} & \stackrel{14.169}{=} \\
 \begin{cases} \alpha \cdot C_{X,X,0}(x) + \beta \cdot C_{X,X,0}(y) & \text{if } H = 0 \\ \frac{\alpha}{\|H\|} \cdot (\sum_{i=2}^{\infty} (-1)^i \cdot H^i)(x) + \frac{\beta}{\|H\|} \cdot (\sum_{i=2}^{\infty} (-1)^i \cdot H^i)(y) & \text{if } H \neq 0 \end{cases} & = \\
 \alpha \cdot \varepsilon(H)(x) + \beta \cdot \varepsilon(H)(y) &
 \end{aligned}$$

and for  $x \in X$

$$\begin{aligned}
 \|\varepsilon(H)(x)\|_X &= \begin{cases} \|C_{X,X,0}(x)\|_X & \text{if } H = 0 \\ \left\| \frac{1}{\|H\|} \cdot (\sum_{i=2}^{\infty} (-1)^i \cdot H^i)(x) \right\|_X & \text{if } H \neq 0 \end{cases} \\
 &= \begin{cases} 0 & \text{if } H = 0 \\ \left\| \frac{1}{\|H\|} \cdot (\sum_{i=2}^{\infty} (-1)^i \cdot H^i)(x) \right\|_X & \text{if } H \neq 0 \end{cases} \\
 &= \begin{cases} \|x\|_X \cdot 0 & \text{if } H = 0 \\ \frac{1}{\|H\|} \cdot \left\| \frac{1}{\|H\|} \cdot (\sum_{i=2}^{\infty} (-1)^i \cdot H^i) \right\| \cdot \|x\|_X & \text{if } H \neq 0 \end{cases} \\
 &\leq_{14.169} \|x\|_X \cdot \begin{cases} 0 & \text{if } H = 0 \\ \frac{1}{\|H\|} \cdot \left\| \sum_{i=2}^{\infty} (-1)^i \cdot H^i \right\| & \text{if } H \neq 0 \end{cases}
 \end{aligned}$$

proving by 11.180 and 11.184 that

$$\forall H \in \mathbb{B}_{\|\cdot\|}(0, 1) \models \varepsilon(H) \in L(X, X) \wedge \|\varepsilon(H)\| \leq \begin{cases} 0 & \text{if } H = 0 \\ \frac{1}{\|H\|} \cdot \left\| \sum_{i=2}^{\infty} (-1)^i \cdot H^i \right\| & \text{if } H \neq 0 \end{cases} \quad (14.170)$$

hence as  $B_{\|\cdot\|}(0, 1) \stackrel{14.3}{=} B_{\|\cdot\|}(\mathbb{1}_X, 1)_{\mathbb{1}_X}$  we can define the function  $\varepsilon$

$$\varepsilon: B_{\|\cdot\|}(\mathbb{1}_X, 1)_{\mathbb{1}_X} = B_{\|\cdot\|}(0, 1) \rightarrow L(X, X) \text{ by } \varepsilon(H) = \begin{cases} C_{X,X,0} & \text{if } H = 0 \\ \frac{1}{\|H\|} \cdot \sum_{i=2}^{\infty} (-1)^i \cdot H^i & \text{if } H \neq 0 \end{cases} \quad (14.171)$$

Next for  $H \in B_{\|\cdot\|}(0, 1) = B_{\|\cdot\|}(\mathbb{1}_X, 1)_{\mathbb{1}_X}$

$$\begin{aligned}
 \sum_{i=2}^{\infty} \|H\|^i &\stackrel{11.379}{=} \left( \sum_{i=1}^{\infty} \|H\|^i \right) - \|H\| \\
 &\stackrel{11.415 \text{ and } \|H\| < 1}{=} \frac{\|H\|}{1 - \|H\|} - \|H\| \\
 &= \frac{\|H\| - (1 - \|H\|) \cdot \|H\|}{1 - \|H\|} \\
 &= \frac{\|H\|^2}{1 - \|H\|} \quad (14.172)
 \end{aligned}$$

Further, as using theorem 14.145 we have  $\|(-1)^i \cdot H^i\| = \|H^i\| \leq \|H\|^i$ , we can apply dominant convergence [see 11.393] on 14.172 giving  $\|\sum_{i=2}^{\infty} (-1)^i \cdot H^i\| \leq \frac{\|H\|^2}{1 - \|H\|}$ . So using 14.170

$$\forall H \in B_{\|\cdot\|}(\mathbb{1}_X, 1) \text{ we have } \|\varepsilon(H)\| \leq \frac{\|H\|}{1 - \|H\|} \quad (14.173)$$

Consider  $L \in \mathcal{GL}(X)$  then we have the following cases to consider

$L = \mathbb{1}_X$ . If  $H \in B_{\|\cdot\|}(\mathbb{1}_X, 1) \underset{14.3}{=} B_{\|\cdot\|}(0, 1)$  then

$$\begin{aligned} \tau(\mathbb{1}_X + H) - \tau(\mathbb{1}_X) &= (\mathbb{1}_X + H)^{-1} + \mathbb{1}_X^{-1} \\ &= (\mathbb{1}_X + H)^{-1} + \mathbb{1}_X \\ &\underset{14.149}{=} \mathbb{1}_X + \sum_{i=1}^{\infty} (-1)^i \cdot H^i - \mathbb{1}_X \\ &= \sum_{i=1}^{\infty} (-1)^i \cdot H^i \\ &\underset{11.379}{=} -(-1)^1 \cdot H^1 + \sum_{i=2}^{\infty} (-1)^i \cdot H^i \\ &= -H + \sum_{i=2}^{\infty} (-1)^i \cdot H^i \end{aligned} \quad (14.174)$$

So as  $\tau(\mathbb{1}_X + 0) - \tau(\mathbb{1}_X) = 0$  we have using the above

$$\begin{aligned} \tau(\mathbb{1}_X + H) - \tau(\mathbb{1}_X) - (-H) &\underset{14.174}{=} \begin{cases} C_{X,X,0} \text{ if } H = 0 \\ \sum_{i=2}^{\infty} (-1)^i \cdot H^i \text{ if } H \neq 0 \end{cases} \\ &= \begin{cases} \|H\| \cdot C_{X,X,0} \\ \|H\| \cdot \frac{1}{\|H\|} \cdot \sum_{i=2}^{\infty} (-1)^i \cdot H^i \end{cases} \end{aligned}$$

which as trivially  $-H = -\mathbb{1}_X \circ H \circ \mathbb{1}_X$  we have gives

$$\tau(\mathbb{1}_X + H) - \tau(\mathbb{1}_X) - (-\mathbb{1}_X \circ H \circ \mathbb{1}_X) = \varepsilon(H) \cdot \|H\| \quad (14.175)$$

Given  $\xi > 0$  choose  $0 < \|H\| < \min\left(\frac{\xi}{2}, \frac{1}{2}\right)$  then we have

$$\begin{aligned} \|\varepsilon(H) - \varepsilon(0)\| &= \|\varepsilon(H)\| \\ &\underset{14.173}{\leq} \frac{\|H\|}{1 - \|H\|} \\ &< \frac{\|H\|}{1 - \frac{1}{2}} \\ &= 2 \cdot \|H\| \\ &< 2 \cdot \frac{\xi}{2} \\ &= \xi \end{aligned}$$

which proves that

$$\varepsilon \text{ is continuous at } 0 \quad (14.176)$$

and thus that  $\varepsilon$  is a  $\varepsilon$ -mapping [see 14.11]. Define  $\Psi: L(X, X) \rightarrow L(X, X)$  by  $\Psi(H) = -H = \mathbb{1}_X \circ H \circ \mathbb{1}_X$  so that trivially  $\Psi \in L(L(X, X), L(X, X))$ . As by 14.175  $\tau(\mathbb{1}_H + H) - \tau(\mathbb{1}_X) - (-\mathbb{1}_X \circ H \circ \mathbb{1}_X) = \varepsilon(H) \cdot \|H\|$  we have by 14.14 that

$$D\tau(\mathbb{1}_X) \text{ exist and } D\tau(\mathbb{1}_X) = \Psi \text{ or } D\tau(\mathbb{1}_X)(H) = -\mathbb{1}_X \circ H \circ \mathbb{1}_X \quad (14.177)$$

$L \neq \mathbb{1}_X$ . Take  $\delta = \min\left(\frac{1}{1 + \|L^{-1}\|}, 1\right)$  [well defined as  $L \in GL(X)$ ] then if  $H \in L(X, X)$  with  $\|H\| < \delta$  we have

$$\begin{aligned}\|H \circ L^{-1}\| &\leq_{11.191} \|H\| \cdot \|L^{-1}\| \\ &< \delta \cdot \|L^{-1}\| \\ &\leq \frac{1}{1 + \|L^{-1}\|} \cdot \|L^{-1}\| \\ &< 1\end{aligned}$$

hence using the previous corollary [see 14.149] we get

$$\forall H \in B_{\|\cdot\|}(0, \delta) \text{ we have that } \mathbb{1}_X + H \circ L^{-1} \in GL(X) \quad (14.178)$$

As  $L \in GL(X)$  we have by 14.142 that  $(\mathbb{1}_X + H \circ L^{-1}) \circ L \in GL(X)$  and as  $(\mathbb{1}_X + H \circ L^{-1}) \circ L \underset{11.203}{=} \mathbb{1}_X \circ L + H \circ (L^{-1} \circ L) = L + H$  it follows that

$$\forall H \in B_{\|\cdot\|}(0, \delta) \text{ we have } L + H = (\mathbb{1}_X + H \circ L^{-1}) \circ L \in GL(X) \quad (14.179)$$

If now  $H \in B_{\|\cdot\|}(L, \delta)$  then  $\|L - H\| < \delta \Rightarrow L - H \in B_{\|\cdot\|}(0, \delta)$  giving  $(L - H) + H \in GL(X)$  hence

$$B_{\|\cdot\|}(L, \delta) \subseteq GL(X) \quad (14.180)$$

If  $H \in B_{\|\cdot\|}(L, \delta)_L \underset{14.3}{=} B_{\|\cdot\|}(0, \delta)$  then

$$\begin{aligned}\|H \circ L^{-1}\| &\leq_{11.191} \|H\| \cdot \|L^{-1}\| \\ &< \delta \cdot \|L^{-1}\| \\ &\leq \frac{1}{1 + \|L^{-1}\|} \cdot \|L^{-1}\| \\ &< 1\end{aligned} \quad (14.181)$$

proving as  $\text{dom}(\varepsilon) \underset{14.171}{=} B_{\|\cdot\|}(0, 1)$

$$\forall H \in B_{\|\cdot\|}(L, \delta)_L \text{ we have that } \varepsilon(H \circ L^{-1}) \text{ is defined} \quad (14.182)$$

Further

$$\begin{aligned}\tau(L + H) - \tau(H) &= (L + H)^{-1} - L^{-1} \\ &\underset{14.179}{=} ((\mathbb{1}_X + H \circ L^{-1}) \circ L)^{-1} - L^{-1} \\ &\underset{2.49}{=} L^{-1} \circ (\mathbb{1}_X + H \circ L^{-1})^{-1} - L^{-1} \\ &= L^{-1} \circ (\mathbb{1}_X + H \circ L^{-1})^{-1} - L^{-1} \circ \mathbb{1}_X \\ &= L^{-1} \circ ((\mathbb{1}_X + H \circ L^{-1})^{-1} - \mathbb{1}_X) \\ &= L^{-1} \circ (\tau(\mathbb{1}_X + H \circ L^{-1}) - \tau(\mathbb{1}_X)) \\ &\underset{14.181, 14.175}{=} L^{-1} \circ (-\mathbb{1}_X \circ (H \circ L^{-1}) \circ \mathbb{1}_X + \varepsilon(H \circ L^{-1}) \cdot \|H \circ L^{-1}\|) \\ &= L^{-1} \circ (-H \circ L^{-1} + \varepsilon(H \circ L^{-1}) \cdot \|H \circ L^{-1}\|) \\ &\underset{11.203}{=} -L^{-1} \circ H \circ L^{-1} + L^{-1} \circ \varepsilon(H \circ L^{-1}) \cdot \|H \circ L^{-1}\|\end{aligned}$$

giving

$$\tau(L + H) - \tau(H) - (-L^{-1} \circ H \circ L^{-1}) = L^{-1} \circ \varepsilon(H \circ L^{-1}) \cdot \|H \circ L^{-1}\| \quad (14.183)$$

Define now given  $H \in B_{\|\cdot\|}(L, \delta)_L$

$$\varepsilon'(H) = \begin{cases} C_{X, X, 0} \text{ if } H = 0 \\ L^{-1} \circ \varepsilon(H \circ L^{-1}) \cdot \frac{\|H \circ L^{-1}\|}{\|H\|} \text{ if } H \neq 0 \end{cases}$$

Given  $x, y \in X$ ,  $\alpha, \beta \in \mathbb{K}$  we have as  $L^{-1} \circ \varepsilon(H \circ L^{-1}) \cdot \frac{\|H \circ L^{-1}\|}{\|H\|} \in L(X, X)$  that

$$\begin{aligned} \varepsilon'(H)(\alpha \cdot x + \beta \cdot y) &= \\ &\begin{cases} C_{X, X, 0}(\alpha \cdot x + \beta \cdot y) \text{ if } H = 0 \\ \left( L^{-1} \circ \varepsilon(H \circ L^{-1}) \cdot \frac{\|H \circ L^{-1}\|}{\|H\|} \right)(\alpha \cdot x + \beta \cdot y) \text{ if } H \neq 0 \end{cases} = \\ &\begin{cases} \alpha \cdot C_{X, X, 0}(x) + \beta \cdot C_{X, X, 0}(y) \text{ if } H = 0 \\ \alpha \cdot \left( L^{-1} \circ \varepsilon(H \circ L^{-1}) \cdot \frac{\|H \circ L^{-1}\|}{\|H\|} \right)(x) + \beta \cdot \left( L^{-1} \circ \varepsilon(H \circ L^{-1}) \cdot \frac{\|H \circ L^{-1}\|}{\|H\|} \right)(y) \text{ if } H \neq 0 \end{cases} = \\ &\alpha \cdot \varepsilon(H)(x) + \beta \cdot \varepsilon(H)(y) \end{aligned}$$

and

$$\begin{aligned} \|\varepsilon'(H)(x)\|_X &= \left\| \begin{cases} C_{X, X, 0}(x) \text{ if } H = 0 \\ \left( L^{-1} \circ \varepsilon(H \circ L^{-1}) \cdot \frac{\|H \circ L^{-1}\|}{\|H\|} \right)(x) \text{ if } H \neq 0 \end{cases} \right\|_X \\ &= \left\| \begin{cases} 0 \text{ if } H = 0 \\ \left( L^{-1} \circ \varepsilon(H \circ L^{-1}) \cdot \frac{\|H \circ L^{-1}\|}{\|H\|} \right)(x) \text{ if } H \neq 0 \end{cases} \right\|_X \\ &= \begin{cases} \|0\|_X \text{ if } H = 0 \\ \left\| \left( L^{-1} \circ \varepsilon(H \circ L^{-1}) \cdot \frac{\|H \circ L^{-1}\|}{\|H\|} \right)(x) \right\|_X \text{ if } H \neq 0 \end{cases} \\ &\stackrel{11.191}{\leq} \|x\|_X \cdot \begin{cases} 0 \text{ if } H = 0 \\ \|L^{-1}\| \cdot \|\varepsilon(H \circ L^{-1})\| \cdot \frac{\|H \circ L^{-1}\|}{\|H\|} \text{ if } H \neq 0 \end{cases} \\ &\stackrel{11.191}{\leq} \|x\|_X \cdot \begin{cases} 0 \text{ if } H = 0 \\ \|L^{-1}\| \cdot \|\varepsilon(H \circ L^{-1})\| \cdot \frac{\|H\| \cdot \|L^{-1}\|}{\|H\|} \text{ if } H \neq 0 \end{cases} \\ &= \|x\|_X \cdot \begin{cases} 0 \text{ if } H = 0 \\ \|L^{-1}\| \cdot \|\varepsilon(H \circ L^{-1})\| \cdot \|L^{-1}\| \text{ if } H \neq 0 \end{cases} \\ &\leq \|x\|_X \cdot \|L^{-1}\|^2 \cdot \|\varepsilon(H \circ L^{-1})\| \end{aligned}$$

proving by 11.180 and 11.184 that

$$\varepsilon'(H) \in L(X, X) \text{ and } \|\varepsilon'(H)\| \leq \|L^{-1}\|^2 \cdot \|\varepsilon(H \circ L^{-1})\| \quad (14.184)$$

This allows us to define

$$\varepsilon': B_{\|\cdot\|}(L, \delta)_L \rightarrow L(X, X) \text{ by } \varepsilon'(H) = \begin{cases} C_{X, X, 0} \text{ if } H = 0 \\ L^{-1} \circ \varepsilon(H \circ L^{-1}) \cdot \frac{\|H \circ L^{-1}\|}{\|H\|} \text{ if } H \neq 0 \end{cases} \quad (14.185)$$

and

$$\forall B \in B_{\|\cdot\|}(0, \delta) \quad \|\varepsilon'(H)\| \leq \|L^{-1}\|^2 \cdot \|\varepsilon(H \circ L^{-1})\| \quad (14.186)$$

Let  $\xi > 0$  then as  $\varepsilon$  is continuous at 0 [see 14.176] there exists a  $\delta_1 > 0$  such that if  $\|H\| < \delta_1$  then  $\|\varepsilon(H)\| = \|\varepsilon(H) - 0\| < \frac{\xi}{\|L^{-1}\|^2 + 1}$ . Take  $\delta_2 = \frac{\delta_1}{\|L^{-1}\| + 1}$  then if  $0 < \|H\| < \delta_2$  we have  $\|H \circ L^{-1}\| \stackrel{11.191}{\leq} \|H\| \cdot \|L^{-1}\| < \frac{\delta_1}{\|L^{-1}\| + 1} \cdot \|L^{-1}\| < \delta_1$  hence  $\|\varepsilon(H \circ L^{-1})\| < \frac{\xi}{\|L^{-1}\|^2 + 1}$ . So if  $0 < \|H\| < \min(\delta, \delta_2)$  we have

$$\begin{aligned} \|\varepsilon'(H)\| &\stackrel{H \neq 0, 14.184}{\leq} \|\varepsilon(H \circ L^{-1})\| \cdot \|L^{-1}\|^2 \\ &< \frac{\xi}{\|L^{-1}\|^2 + 1} \cdot \|L^{-1}\|^2 \\ &< \xi \end{aligned}$$

which proves that  $\varepsilon'$  is continuous at 0 and as  $\varepsilon'(0) = 0$  that

$$\varepsilon' \text{ is a } \varepsilon\text{-mapping [see 14.11]} \quad (14.187)$$

Define now  $\Phi: L(X, X) \rightarrow L(X, X)$  by  $\Phi(H) = -L^{-1} \circ H \circ L$ . Then we have for  $\alpha, \beta \in \mathbb{K}$  and  $H_1, H_2 \in L(X, X)$  that

$$\begin{aligned} \Phi(\alpha \cdot H_1 + \beta \cdot H_2) &= -L^{-1} \circ (\alpha \cdot H_1 + \beta \cdot H_2) \circ L^{-1} \\ &\stackrel{11.203}{=} -L^{-1} \circ (\alpha \cdot H_1 \circ L^{-1} + \beta \cdot H_2 \circ L^{-1}) \\ &\stackrel{11.203}{=} \alpha \cdot (-L^{-1} \circ H_1 \circ L^{-1}) + \beta \cdot (-L^{-1} \circ H_2 \circ L^{-1}) \\ &= \alpha \cdot \Phi(H_1) + \beta \cdot \Phi(H_2) \end{aligned}$$

and for  $H \in L(X, X)$

$$\|\Phi(H)\| \leq_{11.191} \|L^{-1}\|^2 \cdot \|H\|$$

proving by 11.180 that

$$\Phi \in L(L(X, X), L(X, X)) \quad (14.188)$$

Combining the above with 14.183, 14.185 gives

$$\tau(L + H) - \tau(L) - \Phi(H) = \varepsilon'(H) \cdot \|H\|$$

which using 14.187, 14.188 on 14.14 proves that

$$D\tau(L) \text{ exists and } D\tau(L) = \Phi \text{ or } D\tau(L)(H) = -L^{-1} \circ H \circ L^{-1} \quad (14.189)$$

Combining 14.177 and 14.189 proves finally that

$$\forall L \in GL(X) \ D\tau(L) \text{ exists and } D\tau(L)(H) = -L^{-1} \circ H \circ L^{-1} \quad (14.190)$$

To prove that  $\tau$  is  $C^\infty$  we write  $D\tau$  as a composition of  $C^\infty$  mappings and  $\tau$ . Define  $\chi: GL(X) \rightarrow GL(X) \times GL(X)$  by  $\chi(L) = (L, L)$  then we have that  $\pi_1 \circ \chi = \mathbb{1}_{GL(X)}$  and  $\pi_2 \circ \chi = \mathbb{1}_{GL(X)}$  are  $C^\infty$ . Hence using 14.81 we have that

$\chi$  is of class  $C^\infty$

Let  $L, L' \in GL(X)$  define  $\Phi_{L, L'}: L(X, X) \rightarrow L(X, X)$  by  $\Phi_{L, L'}(H) = -L \circ H \circ L'$  [which is well defined as  $GL(X) \in L(X, X)$  so that  $-L \circ H \circ L' \in L(X, X)$ ]. Now if  $\alpha, \beta \in \mathbb{K}$  and  $H_1, H_2 \in L(X, Y)$  then we have

$$\begin{aligned} \Phi_{L, L'}(\alpha \cdot H_1 + \beta \cdot H_2) &= -L \circ (\alpha \cdot H_1 + \beta \cdot H_2) \circ L' \\ &= -L \circ ((\alpha \cdot H_1 + \beta \cdot H_2) \circ L') \\ &\stackrel{11.203}{=} (-L) \circ (\alpha \cdot (H_1 \circ L') + \beta \cdot (H_2 \circ L')) \\ &\stackrel{11.203}{=} \alpha \cdot ((-L) \circ H_1 \circ L') + \beta \cdot ((-L) \circ H_2 \circ L') \\ &= \alpha \cdot \Phi_{L, L'}(H_1) + \beta \cdot \Phi_{L, L'}(H_2) \end{aligned}$$

and if  $H \in L(X, X)$  that

$$\begin{aligned} \|\Phi_{L, L'}(H)\| &= \| -L \circ H \circ L' \| \\ &\leq_{11.191} \| -L \| \cdot \| L' \| \cdot \| H \| \\ &= \| L \| \cdot \| L' \| \cdot \| H \| \end{aligned}$$

proving by 11.180 and 11.184 that

$$\Phi_{L, L'} \in L(L(X, X), L(X, X)) \text{ with } \|\Phi_{L, L'}\|_{L(L(X, X), L(X, X))} \leq \|L\| \cdot \|L'\| \quad (14.191)$$

Define then  $\rho: \mathcal{GL}(X) \times \mathcal{GL}(X) \rightarrow L(L(X, X) \times L(X, X))$  by  $\rho(L, L') = \Phi_{L, L'}$ . Then we have for  $\alpha, \beta \in \mathbb{K}, L, L_1, L_2 \in \mathcal{GL}(X)$  that  $\forall H \in L(X, X)$  we have

$$\begin{aligned}
 \rho(\alpha \cdot L_1 + \beta \cdot L_2, L)(H) &= \Phi_{\alpha \cdot L_1 + \beta \cdot L_2, L}(H) \\
 &= -(\alpha \cdot L_1 + \beta \cdot L_2) \circ H \circ L \\
 &= ((-\alpha) \cdot L_1 + (-\beta) \cdot L_2) \circ H \circ L \\
 &\stackrel{11.203}{=} (-\alpha \cdot (L_1 \circ H \circ L) + (-\beta) \cdot (L_2 \circ H \circ L)) \\
 &\stackrel{11.203}{=} \alpha \cdot (-L_1 \circ H \circ L) + \beta \cdot (-L_2 \circ H \circ L) \\
 &= \alpha \cdot \Phi_{L_1, L}(H) + \beta \cdot \Phi_{L_2, L}(H) \\
 &= \alpha \cdot \rho(L_1, L)(H) + \beta \cdot \rho(L_2, L)(H) \\
 &= (\alpha \cdot \rho(L_1, L) + \beta \cdot \rho(L_2, L))(H)
 \end{aligned}$$

and

$$\begin{aligned}
 \rho(L, \alpha \cdot L_1 + \beta \cdot L_2)(H) &= \Phi_{L, \alpha \cdot L_1 + \beta \cdot L_2}(H) \\
 &= (-L \circ H) \circ (\alpha \cdot L_1 + \beta \cdot L_2) \\
 &\stackrel{11.203}{=} (-L \circ H) \circ (\alpha \cdot L_1) + (-L \circ H) \circ (\beta \cdot L_2) \\
 &\stackrel{11.203}{=} \alpha \cdot (-L \circ H \circ L_1) + \beta \cdot (-L \circ H \circ L_2) \\
 &= \alpha \cdot \Phi_{L, L_1}(H) + \beta \cdot \Phi_{L, L_2}(H) \\
 &= \alpha \cdot \rho(L, L_1)(H) + \beta \cdot \rho(L, L_2)(H) \\
 &= (\alpha \cdot \rho(L, L_1) + \beta \cdot \rho(L, L_2))(H)
 \end{aligned}$$

which proves that  $\rho(\alpha \cdot L_1 + \beta \cdot L_2, L) = \alpha \cdot \rho(L_1, L) + \beta \cdot \rho(L_2, L)$  and  $\rho(L, \alpha \cdot L_1 + \beta \cdot L_2) = \alpha \cdot \rho(L, L_1) + \beta \cdot \rho(L, L_2)$ . So we have that

$$\rho \text{ is multilinear [and thus bilinear]} \quad (14.192)$$

Further if  $(L_1, L_2) \in \mathcal{GL}(X)$  then

$$\begin{aligned}
 \|\rho(L_1, L_2)\|_{L(L(X, X), L(X, X))} &= \|\Phi_{L_1, L_2}\|_{L(L(X, X), L(X, X))} \\
 &\stackrel{14.191}{\leq} \|L\| \cdot \|L'\|
 \end{aligned}$$

which by 11.194 proves that

$$\rho \text{ is continuous} \quad (14.193)$$

As continuous bilinear mappings are  $C^\infty$  [see 14.78] we have using 14.192 and 14.193 that

$$\rho \text{ is of class } C^\infty \quad (14.194)$$

Take now  $L \in \mathcal{GL}(X)$  then  $\forall H \in L(X, X)$

$$\begin{aligned}
 ((\rho \circ \chi \circ \tau)(L))(H) &= (\rho(\chi(\tau(L))))(H) \\
 &= (\rho(\chi(L^{-1}))(H)) \\
 &= (\rho(L^{-1}, L^{-1}))(H) \\
 &= \Phi_{L^{-1}, L^{-1}}(H) \\
 &= -L^{-1} \circ H \circ L^{-1} \\
 &\stackrel{14.190}{=} D\tau(L)(H)
 \end{aligned}$$

which proves that  $(\rho \circ \chi \circ \tau)(L) = D\tau(L)$  so that

$$\rho \circ \chi \circ \tau = D\tau \quad (14.195)$$

We use now induction to prove that  $\forall n \in \mathbb{N}$   $\tau$  is of class  $C^n$  and thus by definition  $C^\infty$ . So let  $S = \{n \in \mathbb{N} \mid \tau \text{ is } C^n\}$  then we have

**0 ∈ S.** As by 14.190  $D\tau(L)$  exists for every  $L \in U$  we have by 14.15 that  $\tau$  is continuous on  $U$  hence  $\tau$  is of class  $C^0$  proving that  $0 \in S$

$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$ . As  $n \in \mathcal{S}$  we have that  $\tau$  is of class  $C^n$  and as  $\rho, \mathcal{X}$  are of class  $C^\infty$  we have by the chain rule [see 14.83] that  $\rho \circ \mathcal{X} \circ \tau$  is of class  $C^n$ . Using 14.195 we have then that  $D\tau$  is of class  $C^n$ . Hence using 14.67 it follows that  $\tau$  is of class  $C^{n+1}$  proving that  $n+1 \in \mathcal{S}$ .

Mathematical induction proves then that  $\mathcal{S} = \mathbb{N}$  and thus that

$$\tau \text{ is of class } C^\infty$$

□

**Definition 14.151.** Let  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle$  be normed spaces  $U \subseteq X$  a open set in  $X$ ,  $V \subseteq Y$  a open set and  $f: U \rightarrow V$  a function. Then we say that

1.  $f$  is a diffeomorphism iff
  - a.  $f: U \rightarrow V$  is a bijection
  - b.  $f: U \rightarrow V$  is differentiable on  $U$
  - c.  $f^{-1}: V \rightarrow U$  is differentiable on  $V$
2.  $f$  is a diffeomorphism of class  $C^n$  [ $n \in \mathbb{N}$ ] iff
  - a.  $f: U \rightarrow V$  is a bijection
  - b.  $f: U \rightarrow V$  is of class  $C^n$
  - c.  $f^{-1}: V \rightarrow U$  is of class  $C^n$

**Example 14.152.** Let  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle$  be normed spaces then a homeomorphism  $f: X \rightarrow Y$  is a diffeomorphism of class  $C^0$

**Example 14.153.** Let  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle$  be Banach spaces and  $L: X \rightarrow Y$  a linear continuous isomorphism [or equivalence  $L \in L(X, Y)$  and  $L$  is a bijection] then  $L$  is diffeomorphism of class  $C^\infty$

**Proof.** First by assumption we have that  $L$  is a bijection. Using 14.140 we have that  $L$  is toplinear hence  $L$  and  $L^{-1}$  are continuous. Further as by 10.183  $L^{-1}$  is also linear we have by 14.75 that  $L$  and  $L^{-1}$  are  $C^\infty$  hence  $L$  is a diffeomorphism of class  $C^\infty$ . □

**Proposition 14.154.** Let  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle$  be normed spaces,  $r \in \mathbb{N}$ ,  $U, W \subseteq X, V \subseteq Y$  open sets with  $W \subseteq U$  and  $f: U \rightarrow V$  a diffeomorphism of class  $C^r$  then  $f|_W: W \rightarrow f(W)$  is a diffeomorphism of class  $C^r$ .

**Proof.** As  $f$  is of class  $C^r$  we have by 14.70 that  $f|_W$  is of class  $C^r$ . Also as  $f$  is open [as  $f^{-1}$  is continuous] we have that  $f(W)$  is open in  $V$ . Further as  $(f^{-1})|_{f(W)} = (f|_W)^{-1}$  and  $f^{-1}$  is  $C^r$  we have using 14.70 that  $(f|_W)^{-1}$  is  $C^r$ . □

**Note 14.155.** Let  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle$  be normed spaces  $U \subseteq X$  a open set,  $V \subseteq Y$  a open set and a function  $f: U \rightarrow V$  then using 2.43 we have that  $f$  is a diffeomorphism [of class  $C^r$ ] if and only if  $f$  is differentiable [of class  $C^r$ ] and there exists a function  $g: V \rightarrow U$  such that  $g$  is differentiable [of class  $C^r$ ] and  $f \circ g = \mathbb{1}_V$  and  $g \circ f = \mathbb{1}_U$ .

**Theorem 14.156.** Let  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle$  be normed spaces,  $U \subseteq X$  a open set,  $V \subseteq Y$  a open set and  $f: U \rightarrow V$  a diffeomorphism and  $g = f^{-1}: V \rightarrow U$  then if we take  $x \in U$  with  $y = f(x)$  we have  $Dg(y) \circ Df(x) = \mathbb{1}_X$  and  $Df(x) \circ Dg(y) = \mathbb{1}_Y$  proving that  $Df(x)$  and  $Dg(y)$  are toplinear isomorphisms.

**Proof.** Let  $y = f(x) \Rightarrow g(y) = x$  then we have

1. From  $f \circ g = \mathbb{1}_V$  and  $\mathbb{1}_Y = D\mathbb{1}_V(x) = D(f \circ g)(y) = Df(g(y)) \circ Dg(y) = Df(x) \circ Dg(y)$
2. From  $g \circ f = \mathbb{1}_U$  and  $\mathbb{1}_X = D\mathbb{1}_U(x) = D(g \circ f)(x) = Dg(f(x)) \circ Df(x) = Dg(y) \circ Df(x)$

from which it follows that  $Df(x)$  and  $Df(x)$  are bijections with  $(Df(x))^{-1} = Dg(y) = D(f^{-1})(y)$  and  $(Dg(y))^{-1} = Df(x)$ . As  $Df(x) \in L(X, Y)$  and  $Dg(y) \in L(Y, X)$  we have by 14.139 that  $Df(x)$  and  $Dg(y)$  are toplinear isomorphisms.  $\square$

**Theorem 14.157.** *Let  $\langle X, \|\cdot\|_X \rangle$ ,  $\langle Y, \|\cdot\|_Y \rangle$ ,  $\langle Z, \|\cdot\|_Z \rangle$  be normed spaces,  $U \subseteq X$  an open set,  $V \subseteq Y$  an open set,  $W \subseteq Z$  an open set and  $f: U \rightarrow V$  a diffeomorphism [of class  $C^r$ ],  $g: V \rightarrow W$  a diffeomorphism [of class  $C^r$ ] then  $f \circ g: U \rightarrow W$  is a diffeomorphism [of class  $C^r$ ]*

**Proof.** Using 2.49 it follows that  $f \circ g$  is a bijection with  $(f \circ g)^{-1} = g^{-1} \circ f^{-1}$ . Further as  $f, g$  are diffeomorphisms [of class  $C^r$ ]. We have by the Chain Rule (see 14.83) that  $f \circ g$  is differentiable on  $U$  [or  $C^r$ ] and  $(f \circ g)^{-1}$  is differentiable on  $W$ . Hence we have that  $f \circ g$  is a diffeomorphism.  $\square$

**Definition 14.158.** *Let  $\langle X, \|\cdot\|_X \rangle$ ,  $\langle Y, \|\cdot\|_Y \rangle$  be normed spaces,  $U \subseteq X$ ,  $V \subseteq Y$  then a function  $f: U \rightarrow V$  is a local diffeomorphism [of class  $C^r$ ] at  $x \in U$  if there exists a open set  $U_x \subseteq U$ ,  $V_{f(x)} \subseteq V$  with  $x \in U_x \subseteq U$ ,  $f(x) \subseteq V_{f(x)}$  and  $f|_{U_x}: U_x \rightarrow V_{f(x)}$  is a diffeomorphism [of class  $C^r$ ]*

We state now the Inverse function theorem

**Theorem 14.159. (The Inverse Function Theorem)** *Let  $\langle X, \|\cdot\|_X \rangle$ ,  $\langle Y, \|\cdot\|_Y \rangle$  be Banach spaces over  $\mathbb{R}$ ,  $U \subseteq X$  an open set,  $V \subseteq Y$   $f: U \rightarrow V$  a differentiable function of class  $C^r$ ,  $r \in \mathbb{N}_0$ . Assume that for a  $x_0 \in U$  we have that  $Df(x_0): X \rightarrow Y$  is a isomorphism then  $f$  is a local diffeomorphism of class  $C^r$  at  $x_0$ . [In other words there exists open sets  $U_{x_0} \subseteq U \subseteq X$ ,  $V_{f(x_0)} \subseteq V$  with  $x_0 \in U_{x_0}$ ,  $f(x_0) \subseteq V_{f(x_0)}$  such that  $f|_{U_{x_0}}: U_{x_0} \rightarrow V_{f(x_0)}$  is a diffeomorphism].*

**Proof.** The proof is quite long and is done by proving several propositions and conjectures. First note that by the definition of the differential [see 14.8] we have that  $Df(x_0)$  is a linear continuous isomorphism so by 14.140 and the fact that  $X, Y$  are Banach spaces we have that

$$Df(x_0) \text{ is a toplinear isomorphism} \quad (14.196)$$

First we prove a more specific version of the the Inverse Function Theorem where  $V = Y = X$  that we will extend later to the more general case.

**Proposition 14.160.** *Let  $\langle X, \|\cdot\|_X \rangle$  be a Banach space over  $\mathbb{R}$ ,  $r \in \mathbb{N}_0$ ,  $0 \in U \subseteq X$  an open set and  $f: U \rightarrow X$  a differentiable function of class  $C^r$  such that  $Df(0) = \mathbb{1}_X$  and  $f(0) = 0$  then  $f$  is a local diffeomorphism at 0.*

**Proof.** Define  $T = \mathbb{1}_X - f: U \rightarrow X$  so that  $T(x) = \mathbb{1}_X(x) - f(x)$  then  $T(0) = 0$  and  $T$  is of class  $C^r$  [as it is the sum of a  $C^\infty$  and a  $C^r$  function [see 14.72]] so

$$D^1T \text{ is continuous on } U \text{ [as } 1 \leq r \Rightarrow T \text{ is of class } C^1] \quad (14.197)$$

Next  $D^1T(0) = D^1\mathbb{1}_X(0) - D^1f(0) = \mathbb{1}_X - \mathbb{1}_X = 0$  [ $0 \in L(X, X)$  the neutral element in  $L(X, Y)$ ] giving

$$D^1T(0) = 0 \quad (14.198)$$

As  $D^1T$  is continuous at 0 there exists a  $\rho'$  with  $0 \in B_{\|\cdot\|_X}(0, \rho') \subseteq U$  such that  $\forall x \in B_{\|\cdot\|_X}(0, \rho')$  we have that  $D^1T(x) \in B_{\|\cdot\|}(0, \frac{1}{2}) \Rightarrow \|D^1T(x)\| < \frac{1}{2}$ . Take then  $\rho = \frac{\rho'}{2}$  then

$$\bar{B}_{\|\cdot\|_X}(0, \rho) \subseteq B_{\|\cdot\|}(0, \rho') \subseteq U \quad (14.199)$$

so that we have

$$\text{if } \|x\|_X \leq \rho \text{ [or } x \in B_{\|\cdot\|_X}(0, \rho)] \text{ then } \|D^1T(x)\| < \frac{1}{2} < 1 \quad (14.200)$$

Now as  $D^1T(x) = \mathbb{1}_X - D^1f(x) \Rightarrow D^1f(x) = \mathbb{1}_X - D^1T(x)$  we can apply 14.148 to get

$$\text{If } \|x\|_X \leq \rho \text{ [or } x \in B_{\|\cdot\|_X}(0, \rho)] \text{ then } Df(x) \in GL(X) \text{ or } Df(x) \text{ is a toplinear isomorphism} \quad (14.201)$$

Now using the fact that  $\bar{B}_{\|\cdot\|_X}(0, \rho)$  is convex [see 14.110] and the Mean Value theorem [see 14.111] together with 14.200 we find that if  $x \in \bar{B}_{\|\cdot\|_X}(0, \delta)$  then  $\|T(x)\|_Y = \|T(x) - T(0)\|_X \leq \frac{1}{2} \cdot \|x - 0\|_X = \frac{1}{2} \cdot \|x\|_X$  giving

$$\text{If } x \in \bar{B}_{\|\cdot\|_X}(0, \rho) \text{ then } \|T(x)\|_X \leq \frac{1}{2} \cdot \|x\|_X \quad (14.202)$$

Next we prove the following conjectures to finish the proposition.

**Conjecture 14.161.** *Let  $0 < \sigma \leq \rho$  then  $\forall y \in \bar{B}_{\|\cdot\|_Y}(0, \frac{\sigma}{2})$  there exists a unique determined  $x \in \bar{B}_{\|\cdot\|_X}(0, \sigma)$  such that  $f(x) = y$*

**Proof.** First note that  $\bar{B}_{\|\cdot\|_X}(0, \sigma), \bar{B}_{\|\cdot\|_X}(0, \frac{\sigma}{2})$  are closed [see 11.56] and thus by 11.348 and the completeness of  $X$  are complete. Given  $y \in \bar{B}_{\|\cdot\|_X}(0, \frac{\sigma}{2})$  and  $x \in B_{\|\cdot\|_X}(0, \sigma)$  define  $T_y(x) = y + T(x)$  then we have  $\|T_y(x)\|_X = \|y + T(x)\|_X \leq \|y\|_X + \|T(x)\|_X \leq_{y \in \bar{B}_{\|\cdot\|_X}(0, \frac{\sigma}{2})} \frac{\sigma}{2} + \|T(x)\|_X \leq_{14.202} \frac{\sigma}{2} + \frac{\|x\|_X}{2} \leq \frac{\sigma}{2} + \frac{\sigma}{2} = \sigma$  so that we can define the following function

$$T_y = C_{\bar{B}_{\|\cdot\|_X}(0, \sigma), X, y} + T: \bar{B}_{\|\cdot\|_X}(0, \sigma) \rightarrow \bar{B}_{\|\cdot\|_X}(0, \sigma) \text{ by } x \mapsto T_y(x) \quad (14.203)$$

Note that  $T_y$  is  $C^r$  as it is the sum of the  $C^\infty$  function  $C_{\bar{B}_{\|\cdot\|_X}(0, \sigma), X, y}$  and the  $C^r$  function  $T$  with  $D^1 T_y = 0 + D^1 T = D^1 T$  hence  $\forall x \in \bar{B}_{\|\cdot\|_X}(0, \sigma) \subseteq_{\sigma \leq \rho} \bar{B}_{\|\cdot\|_X}(0, \rho)$  we have by 14.200 that  $D^1 T_y(x) < \frac{1}{2}$ . Using the Mean Value Theorem 14.111 gives

$$\forall x_1, x_2 \in \bar{B}_{\|\cdot\|_X}(0, \sigma) \text{ we have } \|T_y(x_1) - T_y(x_2)\|_X \leq \frac{1}{2} \cdot \|x_1 - x_2\|_X \quad (14.204)$$

proving that  $T_y$  is a contraction [see 11.412]. Using the fact that  $\bar{B}_{\|\cdot\|_X}(0, \sigma)$  is a Banach space [using 11.348 and the fact that  $\bar{B}_{\|\cdot\|_X}(0, \sigma)$  is closed] and the Banach Fixed Point theorem [see 11.416] there exists a unique  $x \in \bar{B}_{\|\cdot\|_X}(0, \sigma)$  such that  $T_y(x) = x \Leftrightarrow x = y + x - f(x) \Leftrightarrow y = f(x)$  proving the conjecture.  $\square$

**Conjecture 14.162.** *There exists a  $\chi$  with  $0 < \chi < \rho$  such that  $f|_{U_\chi}: U_\chi \rightarrow B_{\|\cdot\|_X}(0, \frac{\chi}{2})$  is a bijective function where  $U_\chi = B_{\|\cdot\|_X}(0, \chi) \cap f^{-1}(B_{\|\cdot\|_X}(0, \frac{\chi}{2}))$  a open set [as  $f$  is continuous].*

**Proof.** Take  $\chi = \frac{\rho}{2}$  then  $0 < \chi < \rho$  then  $f|_{U_\chi}$  is a bijection because

**surjectivity.** Using the previous conjecture [see 14.161] we have if  $y \in B_{\|\cdot\|_X}(0, \frac{\chi}{2}) \subseteq \bar{B}_{\|\cdot\|_X}(0, \frac{\chi}{2})$  that there exists a unique  $x \in \bar{B}_{\|\cdot\|_X}(0, \chi)$  such that  $f(x) = y$ . We prove now by contradiction that

$$\|x\|_X \neq \chi$$

**Proof.** So assume that  $\|x\|_X = \chi$  take then  $\varepsilon = \frac{\chi}{2} - \|y\|_X > 0$  then by the continuity of  $f$  [as  $f$  is  $C^r$ ] there exists a  $\tau' > 0$  such that if  $\|x - x'\|_X < \tau'$  then  $\|f(x) - f(x')\|_X < \varepsilon$ . Take now  $\tau = \min(\tau', \frac{\rho}{4})$  and  $x_1 = x \cdot \left(1 + \frac{\tau}{\chi}\right)$  then we have  $\|x_1 - x\|_X = \frac{\tau}{\chi} \cdot \|x\|_X \stackrel{\|x\|_X = \chi}{=} \tau \leq \tau'$ . So that

$$\begin{aligned} \|f(x_1)\|_X &\leq \|f(x) - f(x_1)\|_X + \|f(x)\|_X \\ &< \varepsilon + \|f(x)\|_X \\ &\stackrel{f(x)=y}{=} \varepsilon + \|y\|_X \\ &= \frac{\chi}{2} - \|y\|_X + \|y\|_X \\ &= \frac{\chi}{2} \end{aligned}$$

proving  $f(x_1) \in \bar{B}_{\|\cdot\|_X}(0, \frac{\chi}{2})$  which as  $0 < \chi < \rho$  means that by the previous conjecture [see 14.161] there exists a  $x_2 \in \bar{B}_{\|\cdot\|_X}(0, \chi)$  such that  $f(x_1) = f(x_2)$ . As  $\|x_1\|_X = \|x\|_X \cdot \left(1 + \frac{\tau}{\chi}\right) = \chi + \tau < \frac{\rho}{2} + \frac{\rho}{4} = \frac{3}{4} \cdot \rho < \rho$  giving

$$x_1 \in \bar{B}_{\|\cdot\|_X}(0, \rho) \quad (14.205)$$

As  $x_2 \in \bar{B}_{\|\cdot\|_X}(0, \chi) = \bar{B}_{\|\cdot\|_X}(0, \frac{\rho}{2}) \subseteq \bar{B}_{\|\cdot\|_X}(0, \rho)$  giving

$$x_2 \in \bar{B}_{\|\cdot\|_X}(0, \rho) \wedge x_2 \in \bar{B}_{\|\cdot\|_X}(0, \chi) \quad (14.206)$$

Also  $f(x_2) = f(x_1) \in \bar{B}_{\|\cdot\|_X}(0, \frac{\chi}{2}) = \bar{B}_{\|\cdot\|_X}(0, \frac{\rho}{4}) \subseteq \bar{B}_{\|\cdot\|_X}(0, \frac{\rho}{2})$  proving

$$f(x_1) = f(x_2) \in \bar{B}_{\|\cdot\|_X}\left(0, \frac{\rho}{2}\right) \quad (14.207)$$

Using the fact that  $0 < \rho \leq \rho$ , 14.205, 14.206, 14.207 the previous conjecture [see 14.161] [especially the uniqueness] we reach the conclusion  $x_1 = x_2$  but then  $\|x_2\|_X \leq \chi < \chi \cdot \left(1 + \frac{\tau}{\chi}\right)_{\|\cdot\|_X} = \chi \cdot \left(1 + \frac{\tau}{\chi}\right)_{\|\cdot\|_X} = \|x_1\|_X \Rightarrow \|x_2\|_X < \|x_1\|_X$  contradicting  $x_1 = x_2$ . Hence we reach the conclusion that  $\|x\|_X \neq \chi$   $\square$

So as  $\|x\|_X \neq \chi$  and  $x \in \bar{B}_{\|\cdot\|_X}(0, \chi)$  we must have that  $x \in B_{\|\cdot\|_X}(0, \chi) \xrightarrow{f(x)=y \in \bar{B}_{\|\cdot\|_X}(0, \frac{\chi}{2})} x \in U_\chi$  proving surjectivity.

**injectivity.** Assume that there exists a  $x_1, x_2 \in U_\chi \subseteq B_{\|\cdot\|_X}(0, \chi) \subseteq \bar{B}_{\|\cdot\|_X}(0, \chi)$  such that  $f(x_1) = f(x_2)$ . As  $f(x_1), f(x_2) \in f(U_\chi) \subseteq B_{\|\cdot\|_X}(0, \frac{\chi}{2})$ , Then as  $0 < \chi < \rho$  we can use the uniqueness in the previous conjecture [see 14.161] to conclude that  $x_1 = x_2$ .  $\square$

Now by the above conjecture we have a  $\chi$  with  $0 < \chi < \rho$  such that  $f|_{U_\chi}: U_\chi \rightarrow B_{\|\cdot\|_X}(0, \frac{\chi}{2})$  is a bijection. Hence the inverse function

$$g = (f|_{U_\chi})^{-1}: B_{\|\cdot\|_X}\left(0, \frac{\chi}{2}\right) \rightarrow U_\chi \quad (14.208)$$

will then exists. We will now prove that  $g$  is continuous and differentiable

**Conjecture 14.163.**  $g$  is continuous on  $B_{\|\cdot\|_X}\left(0, \frac{\chi}{2}\right)$

**Proof.** If  $x \in U_\chi \subseteq \bar{B}_{\|\cdot\|_X}(0, \rho)$  a convex set then  $x = x - f(x) + f(x) = T(x) + f(x)$  so if  $x_1, x_2 \in U_\chi$  then

$$\begin{aligned} \|x_1 - x_2\|_X &= \|T(x_1) + f(x_1) - T(x_2) - f(x_2)\|_X \\ &\leq \|T(x_1) - T(x_2)\|_X + \|f(x_1) - f(x_2)\|_X \end{aligned}$$

so that by using 14.200, the convexity of  $\bar{B}_{\|\cdot\|_X}(0, \rho)$  [see 14.110] and the Mean Value Theorem (see 14.111) we have  $\|x_1 - x_2\|_X \leq \frac{1}{2} \cdot \|x_1 - x_2\|_X + \|f(x_1) - f(x_2)\|_X$  giving

$$\forall x_1, x_2 \in U_\chi \text{ we have } \|x_1 - x_2\|_X \leq 2 \cdot \|f(x_1) - f(x_2)\|_X \quad (14.209)$$

If now  $y_1, y_2 \in B_{\|\cdot\|_X}\left(0, \frac{\chi}{2}\right)$  then  $g(y_1), g(y_2) \in U_\chi$  [see 14.208], so using 14.209 we have  $\|g(y_1) - g(y_2)\|_X \leq 2 \cdot \|f(g(y_1)) - f(g(y_2))\|_X = 2 \cdot \|y_1 - y_2\|_X$  giving

$$\text{If } y_1, y_2 \in B_{\|\cdot\|_X}\left(0, \frac{\chi}{2}\right) \text{ then } \|g(y_1) - g(y_2)\|_X \leq 2 \cdot \|y_1 - y_2\|_X \quad (14.210)$$

So if  $\varepsilon > 0$  and  $y \in B_{\|\cdot\|_X}\left(0, \frac{\chi}{2}\right)$  then for all  $y' \in B_{\|\cdot\|_X}\left(0, \frac{\chi}{2}\right)$  with  $\|y - y'\|_X < \frac{\varepsilon}{2}$  we have  $\|g(y) - g(y')\|_X \leq 2 \cdot \frac{\varepsilon}{2} = \varepsilon$  proving that  $g$  is continuous on  $B_{\|\cdot\|_X}\left(0, \frac{\chi}{2}\right)$ .  $\square$

Next we prove that  $g$  is differentiable.

**Conjecture 14.164.**  $g$  is differentiable on  $B_{\|\cdot\|_X}\left(0, \frac{\chi}{2}\right)$  and  $Dg(y) = (Df(g(y)))^{-1}$

**Proof.** If  $y \in B_{\|\cdot\|_X}\left(0, \frac{\chi}{2}\right)$  we have that  $g(y) \in U_\chi \xrightarrow{14.162} g(y) \in B_{\|\cdot\|_X}(0, \chi) \subseteq B_{\|\cdot\|_X}(0, \rho) \subseteq 14.199 U$ . Further we have by 14.201 that  $Df(g(y))$  is a toplinear isomorphism so that  $Df(g(y))^{-1}$  is continuous and linear. Take now  $h \in (B_{\|\cdot\|_X}\left(0, \frac{\chi}{2}\right))_y \Rightarrow y + h \in B_{\|\cdot\|_X}\left(0, \frac{\chi}{2}\right)$  then for  $k$  defined by

$$k = g(y + h) - g(y) \quad (14.211)$$

we have  $k + g(y) = g(y + h) \in g(B_{\|\cdot\|_X}(0, \frac{\chi}{2}))$   $\stackrel{14.208}{=} U_X \subseteq B_{\|\cdot\|_X}(0, \rho) \subseteq_{14.199} U \Rightarrow k \in (B_{\|\cdot\|_X}(0, \rho))_{g(y)} \subseteq U_{g(y)}$ . As  $f$  is differentiable at  $g(y)$  [for  $g(y) \in U$ ] there exists a  $\varepsilon$ -mapping  $\zeta: U_{g(y)} \rightarrow X$  such that

$$f(g(y) + k) - f(g(y)) - Df(g(y))(k) = \zeta(k) \cdot \|k\|_X \quad (14.212)$$

So using  $f(g(y) + k) = f(g(y + h)) = y + h$  and  $f(g(y)) = y$  it follows that  $f(g(y) + k) - f(g(y)) = h$ , hence using 14.212 we have

$$h = Df(g(y))(k) + \zeta(k) \cdot \|k\|_X \quad (14.213)$$

Applying  $Df(g(y))^{-1}$  to both sides of 14.213 and using linearity we have

$$\begin{aligned} Df(g(y))^{-1}(h) &= Df(g(y))^{-1}(Df(g(y))(k) + \zeta(k) \cdot \|k\|_X) \\ &= Df(g(y))^{-1}(Df(g(y))(k)) + Df(g(y))^{-1}(\zeta(k)) \cdot \|k\|_X \\ &= k + Df(g(y))^{-1}(\zeta(k)) \cdot \|k\|_X \\ &= g(y + h) - g(y) + Df(g(y))^{-1}(\zeta(k)) \cdot \|k\|_X \\ &= g(y + h) - g(y) + Df(g(y))^{-1}(\zeta(g(y + h) - g(y))) \cdot \|g(y + h) - g(y)\|_X \end{aligned} \quad (14.214)$$

Giving

$$g(y + h) - g(y) - Df(g(y))^{-1}(h) = -Df(g(y))^{-1}(\zeta(g(y + h) - g(y))) \cdot \|g(y + h) - g(y)\|_X \quad (14.215)$$

Define

$$\varepsilon: \left( B_{\|\cdot\|_X}(0, \frac{\chi}{2}) \right)_y \rightarrow X \text{ by } \varepsilon(h) = \begin{cases} 0 & \text{if } h = 0 \\ -Df(g(y))^{-1}(\zeta(g(y + h) - g(y))) \cdot \frac{\|g(y + h) - g(y)\|_X}{\|h\|_X} & \text{if } h \neq 0 \end{cases}$$

We have then, using the fact that  $g(y + 0) - g(y) - Df(g(y))^{-1}(0) = 0$ , it follows from 14.215

$$g(y + h) - g(y) - Df(g(y))^{-1}(h) = \varepsilon(h) \cdot \|h\|_X \quad (14.216)$$

Note that in the proof of the previous conjecture [see 14.210] we have as  $y + h, y \in B_{\|\cdot\|_X}(0, \frac{\chi}{2})$  that  $\|g(y + h) - g(y)\|_X \leq 2 \cdot \|y + h - y\|_X = 2 \cdot \|h\|_X$  hence if  $h \neq 0$  then

$$\frac{\|g(y + h) - g(y)\|_X}{\|h\|_X} \leq 2 \quad (14.217)$$

So let  $\gamma < 0$  then by continuity of  $\zeta$  at 0 there exists a  $\delta' > 0$  such that if  $\|z\|_X < \delta'$  then  $\|\zeta(z)\|_X < \gamma$ . By continuity of  $g$  at  $y$  there exists a  $\delta_\gamma > 0$  such that if  $\|h\|_X < \delta_\gamma$  then  $\|g(y + h) - g(y)\|_X < \delta'$  so that  $\|\zeta(g(y + h) - g(y))\|_X < \gamma$ . In other words we have

$$\forall \gamma > 0 \text{ there exists } \delta_\gamma > 0 \text{ such that if } \|h\|_X < \gamma \text{ then } \|\zeta(g(y + h) - g(y))\|_X < \gamma \quad (14.218)$$

Take now  $\mu > 0$  and take  $\gamma = \frac{\mu}{2 \cdot (\|Df(g(y))^{-1}\| + 1)} > 0$  then there exists by the above a  $\delta_\gamma > 0$  such that if  $0 < \|h\|_X < \delta_\gamma$  then

$$\|\zeta(g(y + h) - g(y))\|_X < \frac{\mu}{2 \cdot (\|Df(g(y))^{-1}\| + 1)} \quad (14.219)$$

Using the definition of  $\varepsilon$  we have that

$$\begin{aligned} \|\varepsilon(h) - \varepsilon(0)\|_X &= \|\varepsilon(h)\|_X \\ &= \frac{\|g(y + h) - g(y)\|_X}{\|h\|_X} \cdot \|Df(g(y))^{-1}(\zeta(g(y + h) - g(y)))\|_X \\ &\leq_{14.217} 2 \cdot \|Df(g(y))^{-1}(\zeta(g(y + h) - g(y)))\|_X \\ &\leq 2 \cdot \|Df(g(y))^{-1}\| \cdot \|\zeta(g(y + h) - g(y))\|_X \\ &<_{14.219} 2 \cdot \|Df(g(y))^{-1}\| \cdot \frac{\mu}{2 \cdot (\|Df(g(y))^{-1}\| + 1)} \\ &< \mu \end{aligned}$$

Proving as  $\|\varepsilon(0) - \varepsilon(0)\|_X = 0 < \mu$  that  $\varepsilon$  is continuous at 0 and thus a  $\varepsilon$ -mapping. This together with the linearity, continuity of  $Df(g(y))^{-1}$  and 14.216 proves that  $g$  is differentiable at  $y$  with differential  $Dg(y) = Df(g(y))^{-1}$ . This concludes the proof of the conjecture 14.164.  $\square$

Next we have to prove that  $g$  is of class  $C^r$  to prove the proposition 14.160. To do this take  $\tau: GL(X, X) \rightarrow L(X, X)$  defined by  $L \rightarrow \tau(L) = L^{-1}$  [see 14.150] then we have if  $y \in B_{\|\cdot\|_X}(0, \frac{\chi}{2})$  that  $(\tau \circ Df \circ g)(y) = \tau(Df(g(y))) = Df(g(y))^{-1} \underset{14.164}{=} Dg(y)$  proving that

$$\tau \circ Df \circ g = Dg \quad (14.220)$$

Next we use induction to prove that  $g$  is  $C^r$ , so let  $S = \{n \in \mathbb{N}_0 \mid \text{if } n \leq r \text{ then } g \text{ is } C^n\}$  then we have

$1 \in S$ . Then using 14.150 we have that  $\tau$  is  $C^\infty$  and thus  $\tau$  is continuous, by assumption  $Df$  is continuous and we have proved that  $g$  is continuous [see 14.163] so we must have that  $Dg$  is continuous, proving that  $g$  is  $C^1$  and thus that  $1 \in S$ .

$n \in S \Rightarrow n+1 \in S$ . Then if  $n+1 \leq r$  we have that  $g$  is  $C^n$ ,  $\tau$  is  $C^\infty$  and  $f$  is  $C^r$  so that  $Dg = \tau \circ Df \circ g$  is  $C^n$  hence  $g$  is  $C^{n+1}$  proving that  $n+1 \in S$ .

Using induction we have then that  $S = \mathbb{N}_0$  so as  $r \in \mathbb{N}_0 = S$  and  $r \leq r$  we have  $g$  is  $C^r$ . This proves that

$$f|_{U_\chi}: U_\chi \rightarrow B_{\|\cdot\|_X}\left(0, \frac{\chi}{2}\right) \text{ is a diffeomorphism of class } C^r \quad (14.221)$$

Further as  $f(0) = 0 \in B_{\|\cdot\|_X}\left(0, \frac{\chi}{2}\right)$  we have  $0 \in f^{-1}\left(B_{\|\cdot\|_X}\left(0, \frac{\chi}{2}\right)\right)$  and as  $U_\chi = B_{\|\cdot\|_X}(0, \chi) \cap f^{-1}\left(B_{\|\cdot\|_X}\left(0, \frac{\chi}{2}\right)\right)$  we have that  $0 \in U_\chi$ . To summarize

$$f(0) \in B_{\|\cdot\|_X}\left(0, \frac{\chi}{2}\right) \text{ and } 0 \in U_\chi \quad (14.222)$$

From 14.221 and 14.222 it follows that  $f$  is a local  $C^r$  diffeomorphism at 0 which proves this proposition [see 14.160].  $\square$

Next we must extend proposition 14.160 to the more general case.

**Proposition 14.165.** *Let  $\langle X, \|\cdot\|_X \rangle, \langle Y, \|\cdot\|_Y \rangle$  be Banach spaces over  $\mathbb{R}$ ,  $U \subseteq X$  an open set,  $f: U \rightarrow Y$  a differentiable function of class  $C^r$ ,  $r \in \mathbb{N}_0$ . Assume that for a  $0 \in U$  we have that  $f(0) = 0$  and that  $Df(0): X \rightarrow Y$  is a isomorphism then  $f$  is a local diffeomorphism of class  $C^r$  at 0.*

**Proof.** As  $Df(0)$  is continuous, linear and a isomorphism we have by 14.140 that  $Df(0)$  is toplinear so that  $Df(0)^{-1}: Y \rightarrow X$  is defined, linear and continuous and thus by 14.75 is  $C^\infty$ . Define

$$h = Df(0)^{-1} \circ f: U \rightarrow X$$

which by the above and the chain rule [see 14.83] is differentiable on  $U$  of class  $C^r$ . Then

$$\begin{aligned} h(0) &= (Df(0)^{-1} \circ f)(0) \\ &= Df(0)^{-1}(f(0)) \\ &\stackrel{f(0)=0}{=} Df(0)^{-1}(0) \\ &\stackrel{\text{linearity}}{=} 0 \end{aligned} \quad (14.223)$$

Also

$$\begin{aligned} Dh(0) &= D(Df(0)^{-1})(f(0)) \circ Df(0) \\ &= Df(0)^{-1} \circ Df(0) \\ &= \mathbb{1}_X \end{aligned} \quad (14.224)$$

So by 14.160

$$\exists U_0, V_0' \text{ open in } X \text{ with } 0 \in U_0, 0 = h(0) \in V_0' \text{ and } h|_{U_0}: U_0 \rightarrow V_0' \text{ is a diffeomorphism of class } C^r \quad (14.225)$$

As  $Df(0)$  is continuous, linear and a isomorphism we have by 14.140 that  $Df(0)$  is toplinear and thus open. Hence

$$V_0 = Df(0)(V_0') \text{ is open in } X$$

Further as  $Df(0)$  is linear, continuous and a isomorphism we have by 14.153 that  $Df(0)$  is a diffeomorphism of class  $C^\infty$ . Applying then 14.154 gives

$$Df(0)|_{V'_0}: V'_0 \xrightarrow{\text{def}} Df(0)(V'_0) = V_0 \text{ is a diffeomorphism of class } C^\infty \quad (14.226)$$

Next as  $0 \in V'_0$  and  $Df(0)|_{V'_0}(0) = Df(0) = 0$  it follows that

$$0 \in V_0$$

Take now

$$g = Df(0)|_{V'_0} \circ h|_{U_0}: U_0 \rightarrow V_0$$

then as  $g$  is a composition of a diffeomorphism of class  $C^r$  [see 14.225] and a  $C^\infty$  diffeomorphism [by 14.226] we have using 14.157 that  $g$  is a diffeomorphism of class  $C^r$ . If now  $x \in U_0$  then we have

$$\begin{aligned} g(x) &= (Df(0)|_{V'_0} \circ h|_{U_0})(x) \\ &= Df(0)|_{V'_0}(h|_{U_0}(x)) \\ &= Df(0)((Df(0)^{-1} \circ f)(x)) \\ &= Df(0)(Df(0)^{-1}(f(x))) \\ &= f(x) \end{aligned}$$

proving that

$$f|_{U_0} = g$$

and as  $g$  is a diffeomorphism of class  $C^r$  it follows that  $f|_{U_0}$  is a diffeomorphism of class  $C^r$ . This with the fact that  $0 \in U_0$  and  $f(0) = 0 \in V_0$  proves that  $f$  is a local diffeomorphism at 0.  $\square$

Finally we extend the above proposition to the general case, so lets repeat the Inverse Function Theorem

**Proposition 14.166.** *Let  $\langle X, \|\cdot\|_X \rangle$ ,  $\langle Y, \|\cdot\|_Y \rangle$  be Banach spaces over  $\mathbb{R}$ ,  $V \subseteq Y$ ,  $U \subseteq X$  a open set,  $f: U \rightarrow V$  a differentiable function of class  $C^r$ ,  $r \in \mathbb{N}_0$ . Assume that for a  $x_0 \in U$  we have that  $Df(x_0): X \rightarrow Y$  is a isomorphism then  $f$  is a local diffeomorphism of class  $C^r$  at  $x_0$ .*

**Proof.** First define

$$\text{Given } y \in Y \text{ define } t_y^Y: Y \rightarrow Y \text{ by } t_y^Y(x) = x + y \quad (14.227)$$

and

$$\text{Given } y \in X \text{ define } t_y^X(x) = x + y \quad (14.228)$$

Then as  $(t_y^Y)^{-1} = t_{-y}^Y$  and  $(t_y^X)^{-1} = t_{-y}^X$  we have using 14.77 that

$$\forall y \in Y t_y^Y \text{ is a diffeomorphism of class } C^\infty \text{ and } Dt_y^Y = C_{Y, L(Y, Y), \mathbb{1}_Y} \quad (14.229)$$

and

$$\forall y \in X t_y^X \text{ is a diffeomorphism of class } C^\infty \text{ and } Dt_y^X = C_{X, L(X, X), \mathbb{1}_X} \quad (14.230)$$

Consider then

$$h = t_{-f(x_0)}^Y \circ f \circ t_{x_0}^X: t_{-x_0}^X(U) \rightarrow Y \quad (14.231)$$

then using the Chain Rule [see 14.22] we have that  $h$  is of class  $C^r$  and

$$\begin{aligned} Dh(0) &= D(t_{-f(x_0)}^Y \circ f \circ t_{x_0}^X)(0) \\ &= D(t_{-f(x_0)}^Y \circ f)(t_{x_0}^X(0)) \circ Dt_{x_0}^X(0) \\ &= D(t_{-f(x_0)}^Y \circ f)(t_{x_0}^X(0)) \circ \mathbb{1}_X \\ &= D(t_{-f(x_0)}^Y \circ f)(x_0) \\ &= D(t_{-f(x_0)}^Y)(f(x_0)) \circ Df(x_0) \\ &= \mathbb{1}_Y \circ Df(x_0) \\ &= Df(x_0) \end{aligned}$$

Summarizing we have

$$h \text{ is of class } C^r \text{ and } Dh(0) = Df(x_0) \text{ a isomorphism} \quad (14.232)$$

Further we have

$$\begin{aligned} h(0) &= (t_{-f(x_0)}^Y \circ f \circ t_{x_0}^X)(0) \\ &= t_{-f(x_0)}^Y(f(t_{x_0}^X(0))) \\ &= t_{-f(x_0)}^Y(f(x_0)) \\ &= 0 \end{aligned}$$

So all the conditions for proposition 14.165 are met for  $h$  hence there exists open sets  $U'_0 \subseteq X$  and  $V'_0 \subseteq Y$  such that  $0 \in U'_0$ ,  $0 \in V'_0$  and

$$h|_{U'_0}: U'_0 \rightarrow V'_0 \text{ is a diffeomorphism of class } C^r \quad (14.233)$$

Define now  $U_{x_0} = t_{x_0}^X(U'_0)$  and  $V_{f(x_0)} = t_{f(x_0)}^Y(V'_0)$  [which are open sets because  $t_{x_0}^X$  and  $t_{f(x_0)}^Y$  are continuous and open]

$$g = t_{f(x_0)}^Y \circ h|_{U'_0} \circ t_{-x_0}^X: U_{x_0} \rightarrow V_{f(x_0)}$$

Then as  $t_{-x_0}^X$ ,  $t_{f(x_0)}^Y$  are  $C^\infty$  diffeomorphisms [see 14.229 and 14.230] and  $h|_{U'_0}$  is a diffeomorphism of class  $C^r$  we have by 14.157 that

$$g \text{ is a diffeomorphism of class } C^r \quad (14.234)$$

Further if  $x \in U_{x_0}$  then we have

$$\begin{aligned} g(x) &= (t_{f(x_0)}^Y \circ h|_{U'_0} \circ t_{-x_0}^X)(x) \\ &= t_{f(x_0)}^Y(h|_{U'_0}(t_{-x_0}^X(x))) \\ &= t_{f(x_0)}^Y(h|_{U'_0}(x - x_0)) \\ &= t_{f(x_0)}^Y(h(x - x_0)) \\ &= t_{f(x_0)}^Y((t_{-f(x_0)}^Y \circ f \circ t_{x_0}^X)(x - x_0)) \\ &= t_{f(x_0)}^Y(t_{-f(x_0)}^Y(f(t_{x_0}^X(x - x_0)))) \\ &= t_{f(x_0)}^Y(t_{-f(x_0)}^Y(f(x))) \\ &= t_{f(x_0)}^Y(f(x) - f(x_0)) \\ &= f(x) \end{aligned}$$

proving that

$$g = f|_{U_{x_0}} \text{ is a diffeomorphism of class } C^r \quad (14.235)$$

Next from  $0 \in U'_0$  we have that  $x_0 = t_{x_0}^X(0) \in t_{x_0}^X(U'_0) = U_{x_0}$  and from  $0 \in V'_0$  we have  $f(x_0) = t_{f(x_0)}^Y(0) \in t_{f(x_0)}^Y(V'_0) = V_{f(x_0)}$  giving

$$x_0 \in U_{x_0} \text{ and } f(x_0) \in V_{f(x_0)} \quad (14.236)$$

Finally from 14.235 and 14.236 proves that

$$f \text{ is a local diffeomorphism of class } C^r \text{ at } x_0$$

□

The above proposition proves finally the Inverse Function Theorem

□

**Corollary 14.167.** *Given  $\langle \mathbb{K}, \|\cdot\| \rangle$  (where  $\mathbb{K} = \mathbb{R}$  or  $\mathbb{C}$ ), a normed space,  $U \subseteq \mathbb{K}$  a open set,  $x \in U$  a open set,  $V \subseteq \mathbb{K}$  and a function  $f: U \rightarrow V$  such that  $\forall x \in U$  we have that  $f'(x)$  exists, is continuous and for  $x_0 \in U$  we have that  $f'(x_0) \neq 0$  then there exists open sets  $U', V'$  with  $x_0 \in U' \subseteq U$  and  $f(x_0) \in V' \subseteq V$  such that  $f|_{U'}: U' \rightarrow V'$  is a bijection and  $(f|_{U'})^{-1}: V' \rightarrow U'$  has a derivative  $((f|_{U'})^{-1})'(x)$  at every  $x \in V'$  such that  $((f|_{U'})^{-1})'(x) \cdot f'((f|_{U'})^{-1}(x)) = 1$ .*

**Proof.** Using 14.58 we have from the hypothese that

$$f \text{ is } C^1 \text{ and } f' = Df(*)(1) \quad (14.237)$$

As  $f'(x_0) \neq 0$  we can define the function  $g: \mathbb{K} \rightarrow \mathbb{K}$  by  $g(t) = \frac{t}{f'(x_0)}$  then we have  $\forall t \in \mathbb{K}$  that

$$\begin{aligned} (Df(x_0) \circ g)(t) &= Df(x_0)(g(t)) \\ &= Df(x_0)\left(\frac{t}{f'(x_0)}\right) \\ &= \frac{t}{f'(x_0)} \cdot Df(x_0)(1) \\ &\stackrel{14.237}{=} \frac{t}{f'(x_0)} \cdot f'(x_0)(1) \\ &= t \\ &= \mathbb{1}_{\mathbb{K}}(t) \\ (g \circ Df(x_0))(t) &= g(Df(x_0)(t)) \\ &= \frac{Df(x_0)(t)}{f'(x_0)} \\ &= \frac{Df(x_0)(t)}{f'(x_0)} \\ &= \frac{t \cdot Df(x_0)(1)}{f'(x_0)} \\ &\stackrel{14.237}{=} \frac{t}{f'(x_0)} \\ &= \mathbb{1}_{\mathbb{K}}(t) \end{aligned}$$

proving that  $Df(x_0) \circ g = \mathbb{1}_{\mathbb{K}} = g \circ Df(x_0)$  or

$$Df(x_0) \text{ is a bijection with } Df(x_0)^{-1} = g \quad (14.238)$$

Using the Inverse Function Theorem [see 14.159] on 14.237 and 14.238 proves that  $f$  is a local diffeomorphism of class  $C^1$ . Hence there exists a  $U', V'$  open with  $U' \subseteq U, V' \subseteq V$  such that

$$x_0 \in U' \wedge f(x_0) \in V' \text{ and } f|_{U'}: U' \rightarrow V' \text{ is a diffeomorphism} \quad (14.239)$$

Let  $x \in V'$  then  $(f|_{U'})^{-1}(x)$  exists, using 14.156 we have  $D((f|_{U'})^{-1})(x) \circ Df|_{U'}((f|_{U'})^{-1}(x)) = \mathbb{1}_{\mathbb{K}}$  hence

$$\begin{aligned} 1 &= \mathbb{1}_{\mathbb{K}}(1) \\ &= (D((f|_{U'})^{-1})(x) \circ Df|_{U'}((f|_{U'})^{-1}(x)))(1) \\ &= D((f|_{U'})^{-1})(x)(Df|_{U'}((f|_{U'})^{-1}(x))(1)) \\ &= D((f|_{U'})^{-1})(x)((f|_{U'})'((f|_{U'})^{-1}(x))) \\ &= (f|_{U'})'((f|_{U'})^{-1}(x)) \cdot D((f|_{U'})^{-1})(x)(1) \\ &= (f|_{U'})'((f|_{U'})^{-1}(x)) \cdot ((f|_{U'})^{-1})'(x) \end{aligned}$$

proving that

$$((f|_{U'})^{-1})'((f|_{U'})^{-1}(x)) \cdot (f|_{U'})'((f|_{U'})^{-1}(x)) = 1$$

Using now the fact that the derivative is local [see 14.20] we have

$$((f|_{U'})^{-1})'((f|_{U'})^{-1}(x)) \cdot f'((f|_{U'})^{-1}(x)) = 1$$

□

A import conclusion of the Inverse Function Theorem is the below Implicit Function Theorem. First we need a little lemma.

**Lemma 14.168.** Let  $X, Y, Z$  be vector spaces over the field  $\mathbb{K}$ ,  $L \in \text{Hom}(X, Z)$  and  $H \in \text{Hom}(Y, Z)$  such that  $H^{-1} \in \text{Hom}(Z, Y)$  exists then  $T: X \times Y \rightarrow X \times Z$  defined by  $(x, y) \mapsto (x, L(x) + H(y))$  is a isomorphism with inverse  $S: X \times Z \rightarrow X \times Y$  defined by  $S(x, y) = (x, -H^{-1}(L(x)) + H^{-1}(y))$  so that  $T$

**Proof.** Take  $(x_1, y_1), (x_2, y_2) \in X \times Y$  and  $\alpha, \beta \in \mathbb{K}$  then

$$\begin{aligned} T(\alpha \cdot (x_1, y_1) + \beta \cdot (x_2, y_2)) &= T(\alpha \cdot x_1 + \beta \cdot x_2, \alpha \cdot y_1 + \beta \cdot y_2) \\ &= (\alpha \cdot x_1 + \beta \cdot x_2, L(\alpha \cdot x_1 + \beta \cdot x_2) + H(\alpha \cdot y_1 + \beta \cdot y_2)) \\ &= (\alpha \cdot x_1 + \beta \cdot x_2, \alpha \cdot L(x_1) + \beta \cdot L(x_2) + \alpha \cdot H(y_1) + \beta \cdot H(y_2)) \\ &= (\alpha \cdot x_1, \alpha \cdot L(x_1) + \alpha \cdot H(y_1)) + (\beta \cdot x_2, \beta \cdot L(y_1) + \beta \cdot H(y_2)) \\ &= \alpha \cdot (x_1, L(x_1) + H(y_1)) + \beta \cdot (x_2, L(y_1) + H(y_2)) \\ &= \alpha \cdot T(x_1, y_1) + \beta \cdot T(x_2, y_2) \end{aligned}$$

proving that

$T$  is linear

First if  $x \in X$ ,  $y \in Y$  and  $z \in Z$  then

$$\begin{aligned} T(S(x, z)) &= T(x, -H^{-1}(L(x)) + H^{-1}(z)) \\ &= (x, L(x) + H(-H^{-1}(L(x)) + H^{-1}(z))) \\ &= (x, L(x) + -H(H^{-1}(L(x))) + H(H^{-1}(z))) \\ &= (x, L(x) - L(x) + z) \\ &= (x, z) \\ S(T(x, y)) &= S(x, L(x) + H(y)) \\ &= (x, -H^{-1}(L(x)) + H^{-1}(L(x) + H(y))) \\ &= (x, -H^{-1}(L(x)) + H^{-1}(L(x)) + H^{-1}(H(y))) \\ &= (x, y) \end{aligned}$$

proving that  $T$  is a bijection with inverse  $S$ .  $\square$

**Theorem 14.169. (Implicit Function Theorem)** Let  $\langle X, \|\cdot\|_X \rangle$ ,  $\langle Y, \|\cdot\|_Y \rangle$  and  $\langle Z, \|\cdot\|_Z \rangle$  be Banach spaces,  $U \subseteq X$ ,  $V \subseteq Y$  be open sets, and let  $f: U \times V \rightarrow Z$  be a differentiable function of class  $C^r$ ,  $r \in \mathbb{N}_0$ . Assume that  $\exists x_0 \in U$ ,  $\exists y_0 \in V$  such that  $D_2f(x_0, y_0): Y \rightarrow Z$  is a isomorphism then

1. There exists a open set  $W_0 \subseteq U \times Z$  and a open set  $V_0 \subseteq V$  with  $(x_0, f(x_0, y_0)) \in W_0$  such that there exists a  $C^r$  map  $g: W_0 \rightarrow V_0$  such that  $g(x_0, f(x_0, y_0)) = y_0$  and  $\forall (x, z) \in W_0$   $f(x, g(x, z)) = z$ .
2. If  $h: W_0 \rightarrow V_0$  such that  $f(x, g(x, z)) = z$  then  $h = g$  [so  $g$  is unique]
3. Further  $\forall (x, z) \in W_0$ 
  - a.  $[D_2f(x, g(x, z))]^{-1}$  exist
  - b.  $D_1g(x, z) = -[D_2f(x, g(x, z))]^{-1} \circ D_1f(x, g(x, z))$
  - c.  $D_2g(x, z) = [D_2f(x, g(x, z))]^{-1}$

**Proof.** First as we will have different project functions [see 11.148] we will use the following notations for the projection functions

$$\begin{aligned} \pi_1^{X \times Y}: X \times Y \rightarrow X &\text{ defined by } \pi_1^{X \times Y}(x, y) = x \\ \pi_2^{X \times Y}: X \times Y \rightarrow Y &\text{ defined by } \pi_2^{X \times Y}(x, y) = y \\ \pi_1^{X \times Z}: X \times Z \rightarrow X &\text{ defined by } \pi_1^{X \times Z}(x, y) = x \\ \pi_2^{X \times Z}: X \times Z \rightarrow Z &\text{ defined by } \pi_2^{X \times Z}(x, y) = y \end{aligned}$$

Note that by 14.79, 14.70 we have that

$$\pi_1^{X \times Y}, \pi_2^{X \times Y}, \pi_1^{X \times Z} \text{ and } \pi_2^{X \times Z} \text{ and their restrictions to open sets are of class } C^\infty \quad (14.240)$$

Define now

$$\varphi: U \times V \rightarrow X \times Z \text{ by } \varphi(x, y) = (x, f(x, y)) \quad (14.241)$$

then  $\forall (x, y) \in U \times V$  we have that

$$\begin{aligned} (\pi_1^{X \times Z} \circ \varphi)(x, y) &= \pi_1^{X \times Z}(\varphi(x, y)) \\ &= \pi_1^{X \times Z}(x, f(x, y)) \\ &= x \\ &= \pi_1^{X \times Y}(x, y) \end{aligned}$$

proving that

$$\pi_1^{X \times Z} \circ \varphi = (\pi_1^{X \times Y})_{|U \times V} \text{ so by 14.240 } \pi_1^{X \times Z} \circ \varphi \text{ is } C^\infty \quad (14.242)$$

Hence

$$\begin{aligned} D^1(\pi_1^{X \times Z} \circ \varphi)(x, y) &= D^1((\pi_1^{X \times Y})_{|U \times V})(x, y) \\ &\stackrel{14.79, 14.70}{=} \pi_1^{X \times Y} \end{aligned} \quad (14.243)$$

Further we have that

$$\begin{aligned} (\pi_2^{X \times Z} \circ \varphi)(x, y) &= \pi_2^{X \times Z}(\varphi(x, y)) \\ &= \pi_2^{X \times Z}(x, f(x, y)) \\ &= f(x, y) \end{aligned}$$

proving that

$$\pi_2^{X \times Z} \circ \varphi = f \text{ a } C^r \text{ function} \quad (14.244)$$

and

$$\begin{aligned} D^1(\pi_2^{X \times Z} \circ \varphi) &= D^1f \\ &\stackrel{14.131}{=} D_1f(x) \circ \pi_1^{X \times Y} + D_2f(x) \circ \pi_2^{X \times Y} \end{aligned} \quad (14.245)$$

Using 14.81 on 14.242 and 14.244 gives

$$\varphi \text{ is of class } C^r \text{ and } \pi_1^{X \times Y} \circ D^1\varphi(x, y) \stackrel{14.80}{=} \pi_1^{X \times Y} \text{ and } \pi_2^{X \times Y} \circ D^1\varphi(x, y) \stackrel{14.80}{=} D_1f(x) \circ \pi_1^{X \times Y} + D_2f(x) \circ \pi_2^{X \times Y} \quad (14.246)$$

or

$$D^1\varphi(x, y)(r, s) \stackrel{14.80}{=} (r, D_1f(x, y)(r) + D_2f(x, y)(s)) \quad (14.247)$$

Using the above for  $(x_0, y_0) \in U \times V$  gives  $D^1\varphi(x_0, y_0)(x, y) = (x, D_1f(x_0, y_0)(x) + D_2f(x_0, y_0)(y))$ . Using the lemma 14.168 and the fact that  $D_2f(x_0, y_0)$  is a isomorphism proves that  $D^1\varphi(x_0, y_0)$  is a isomorphism. Using the Inverse Function Theorem [see 14.159] there exists open sets  $\widetilde{U}_{(x_0, y_0)} \subseteq U \times V$ ,  $\widetilde{V}_{\varphi(x_0, y_0)} \subseteq X \times Z$  with  $(x_0, y_0) \in \widetilde{U}_{(x_0, y_0)}$  and  $(x_0, f(x_0, y_0)) = \varphi(x_0, y_0) \in \widetilde{V}_{\varphi(x_0, y_0)}$  such that

$$\varphi|_{\widetilde{U}_{(x_0, y_0)}}: \widetilde{U}_{(x_0, y_0)} \rightarrow \widetilde{V}_{\varphi(x_0, y_0)} \text{ is a diffeomorphisme of class } C^r \quad (14.248)$$

Using the definition of the product topology on  $X \times Y$  [see 11.34 and 11.39] there exist open sets  $U_0 \subseteq X$ ,  $V_0 \subseteq Y$  such that  $(x_0, y_0) \in U_0 \times V_0 \subseteq \widetilde{U}_{(x_0, y_0)}$ . Define then  $W_0 = \varphi(U_0 \times V_0)$  then using 14.154 together with 14.248 and the fact that  $(\varphi|_{\widetilde{U}_{(x_0, y_0)}})|_{U_0 \times V_0} \stackrel{2.31}{=} \varphi|_{U_0 \times V_0}$  gives

$$\varphi|_{U_0 \times V_0}: U_0 \times V_0 \rightarrow W_0 \text{ is a diffeomorphisme of class } C^r \quad (14.249)$$

where

$$W_0 \stackrel{\text{def}}{=} \varphi(U_0 \times V_0) \subseteq U \times Z \text{ and } \varphi|_{U_0 \times V_0}(x_0, y_0) = (x_0, f(x_0, y_0)) \in W_0 \quad (14.250)$$

Also if  $y \in V_0$  then as  $x_0 \in U_0$  we have  $(x_0, y) \in U_0 \times V_0 \subseteq \widetilde{U}_{(x_0, y_0)} \subseteq U \times V$  so that  $y \in V_0$  proving that

$$V_0 \subseteq V \quad (14.251)$$

So  $(\varphi|_{U_0 \times V_0})^{-1}: W_0 \rightarrow U_0 \times V_0$  exists and is  $C^r$ . Hence if  $(x, z) \in W_0$  then

$$\begin{aligned} (x, z) &= \varphi|_{U_0 \times V_0}((\varphi|_{U_0 \times V_0})^{-1}(x, z)) \\ &= \varphi((\varphi|_{U_0 \times V_0})^{-1}(x, z)) \\ &= \varphi(\pi_1^{X \times Y}((\varphi|_{U_0 \times V_0})^{-1}(x, z)), \pi_2^{X \times Y}((\varphi|_{U_0 \times V_0})^{-1}(x, z))) \\ &= (\pi_1^{X \times Y}((\varphi|_{U_0 \times V_0})^{-1}(x, z)), f(\pi_1^{X \times Y}((\varphi|_{U_0 \times V_0})^{-1}(x, z)), \pi_2^{X \times Y}((\varphi|_{U_0 \times V_0})^{-1}(x, z)))) \end{aligned}$$

so that

$$x = \pi_1^{X \times Y}((\varphi|_{U_0 \times V_0})^{-1}(x, z)) \in \pi_1^{X \times Y}(U_0 \times V_0) = U_0 \wedge z = f(x, \pi_2^{X \times Y}((\varphi|_{U_0 \times V_0})^{-1}(x, z))) \quad (14.252)$$

So if we take in account that  $V_0 = \pi_2^{X \times Y}(U_0 \times V_0)$  we can define

$$g: W_0 \rightarrow V_0 \text{ by } g(x, z) = \pi_2^{X \times Y}((\varphi|_{U_0 \times V_0})^{-1}(x, z)) \quad (14.253)$$

Further

$$\forall (x, z) \in W_0 \text{ we have } f(x, g(x, z)) \underset{14.252 \text{ and } 14.253}{=} z \quad (14.254)$$

As from 14.250 we have  $\varphi|_{U_0 \times V_0}(x_0, y_0) = (x_0, f(x_0, y_0))$  it follows that  $(\varphi|_{U_0 \times V_0})^{-1}(x_0, f(x_0, y_0)) = (x_0, y_0) \Rightarrow y_0 = \pi_2^{X \times Y}((\varphi|_{U_0 \times V_0})^{-1}(x_0, f(x_0, y_0))) = g(x_0, f(x_0, y_0))$  proving that

$$g(x_0, f(x_0, y_0)) = y_0 \quad (14.255)$$

Next as  $\pi_2^{X \times Y}$  is  $C^\infty$  and  $(\varphi|_{U_0 \times V_0})^{-1}$  is  $C^r$  we have that

$$g \text{ is } C^r \quad (14.256)$$

So that point (1) of the theorem is proved.

To prove point (2) take  $h: W_0 \rightarrow V_0$  such that  $\forall (x, z) \in W_0 f(x, h(x, z)) = z$ . Let  $(x, z) \in W_0$  then we have  $f(x, g(x, z)) = z = f(x, h(x, z)) \Rightarrow g(x, z) = h(x, z)$  proving that  $g = h$ .

Now to prove (3.a) note that by 14.249 and 14.156 we have  $\forall (x, y) \in U_0 \times V_0$  that  $D(\varphi|_{U_0 \times V_0})(x, y)$  is bijective or using the fact that  $D(\varphi|_{U_0 \times V_0})(x, y) \underset{14.10}{=} D\varphi(x, y)$  we have

$$\forall (x, y) \in U_0 \times V_0 \quad D\varphi(x, y) \text{ is bijective} \quad (14.257)$$

Take  $r, s \in Y$  such that  $D_2 f(x, y)(r) = D_2 f(x, y)(s)$  then

$$\begin{aligned} D\varphi(x, y)(0, r) &\underset{14.247}{=} (0, D_1 f(x, y)(0) + D_2 f(x, y)(r)) \\ &= (0, D_1 f(x, y)(0) + D_2 f(x, y)(s)) \\ &\underset{14.247}{=} D\varphi(x, y)(0, s) \end{aligned}$$

proving as  $D\varphi(x, y)$  is bijective that  $(0, r) = (0, s) \Rightarrow r = s$  giving that

$$D_2 f(x, y) \text{ is injective} \quad (14.258)$$

If  $z \in Z$  then as  $D\varphi(x, y)$  is bijective there exists a  $(r, s) \in X \times Z$  such that  $D\varphi(r, s) = (0, z) \Rightarrow (r, D_1 f(x, y)(r) + D_2 f(x, y)(s)) = (0, z)$  hence  $r = 0$  and  $D_1 f(x, y)(0) + D_2 f(x, y)(s) = z$  or  $D_2 f(x, y)(s) = z$ . So  $D_2 f(x, y)$  is surjective which together with 14.258 proves

$$\forall (x, y) \in U_0 \times V_0 \text{ we have } D_2 f(x, y) \text{ is bijective} \quad (14.259)$$

Take now  $(x, z) \in W_0$  then using 14.252 we have  $x \in U_0$  and as by 14.253  $g(x, z) \in V_0$  we have by the above that

$$\forall (x, y) \in W_0 \quad D_2 f(x, g(x, z)) \text{ is bijective so that } [D_2 f(x, g(x, z))]^{-1} \text{ exists} \quad (14.260)$$

proving (3.a).

Now to calculate the differentials of  $g$  define

$$\psi: W_0 \rightarrow X \times Z \text{ by } \psi(x, z) = (x, g(x, z)) \quad (14.261)$$

then for  $f \circ \psi: W_0 \rightarrow Z$  we have  $(f \circ \psi)(x, z) = f(\psi(x, z)) = f(x, g(x, z)) \underset{14.254}{=} z = \pi_2^{X \times Z}(x, z)$  proving that

$$f \circ \psi = \pi_2^{X \times Z} \quad (14.262)$$

Now as  $((f \circ \psi) \circ (1 \rightarrow (x, z)))(t) = \pi_2^{X \times Z}((1 \rightarrow (x, z))(t)) = \pi_2^{X \times Z}((t, z)) = z = C_{(1 \rightarrow (x, z))^{-1}(W_0), Z, z}$  we have that  $D_1(f \circ \psi)(x, z) = D(((f \circ \psi) \circ (1 \rightarrow (x, z))))(x) = 0$  so that

$$D_1(f \circ \psi)(x, z) = 0 \text{ [the neutral element in } L(X, Z)] \quad (14.263)$$

Also

$$\begin{aligned} (\pi_1^{X \times Z} \circ (\psi \circ (1 \rightarrow (x, z))))(t) &= \pi_1^{X \times Z}(\psi((1 \rightarrow (x, z))(t))) \\ &= \pi_1^{X \times Z}(\psi(t, z)) \\ &= \pi_1^{X \times Z}(t, g(t, z)) \\ &= t \\ &= \mathbb{1}_X(t) \end{aligned}$$

and

$$\begin{aligned} (\pi_2^{X \times Z} \circ (\psi \circ (1 \rightarrow (x, z))))(t) &= \pi_2^{X \times Z}(\psi((1 \rightarrow (x, z))(t))) \\ &= \pi_2^{X \times Z}(\psi(t, z)) \\ &= \pi_2^{X \times Z}(t, g(t, z)) \\ &= g(t, z) \\ &= g((1 \rightarrow (x, z))(t)) \\ &= (g \circ (1 \rightarrow (x, z)))(t) \end{aligned}$$

proving that

$$\pi_1^{X \times Z} \circ (\psi \circ (1 \rightarrow (x, z))) = \mathbb{1}_X \quad (14.264)$$

$$\pi_2^{X \times Z} \circ (\psi \circ (1 \rightarrow (x, z))) = g \circ (1 \rightarrow (x, z)) \quad (14.265)$$

We can now recalculate  $D_1(f \circ \psi)(x, z)$  as follows

$$\begin{aligned} (D_1(f \circ \psi)(x, z))(h) &\stackrel{14.36}{=} \\ (Df(\psi(x, z)) \circ D_1\psi(x, z))(h) &\stackrel{\text{def}}{=} \\ (Df(\psi(x, z)) \circ D(\psi \circ (1 \rightarrow (x, z)))(x))(h) &\stackrel{14.264, 14.265, 14.43}{=} \\ (Df(\psi(x, z)) \circ (D\mathbb{1}_X(x), D(g \circ (1 \rightarrow (x, z)))(x)))(h) &\stackrel{\text{def}}{=} \\ (Df(\psi(x, z)) \circ (\mathbb{1}_X, D_1g(x, z)))(h) &= \\ (Df(\psi(x, z))(h, D_1g(x, z)(h))) &\stackrel{14.35}{=} \\ D_1f(\psi(x, z))(h) + D_2f(\psi(x, z))(D_1g(x, z)(h)) &= \\ D_1f(\psi(x, z))(h) + (D_2f(\psi(x, z)) \circ D_1g(x, z))(h) &= \\ (D_1f(\psi(x, z)) + D_2f(\psi(x, z)) \circ D_1g(x, z))(h) &\stackrel{\text{def of } \psi}{=} \\ (D_1f(x, g(x, z)) + D_2f(x, g(x, z)) \circ D_1g(x, z))(h) & \end{aligned}$$

Proving that  $D_1(f \circ \psi)(x, z) = D_1f(x, g(x, z)) + D_2f(x, g(x, z)) \circ D_1g(x, z)$  which as by 14.263  $D_1(f \circ \psi)(x, z) = 0$  gives

$$D_2f(x, g(x, z)) \circ D_1g(x, z) = -D_1f(x, g(x, z))$$

Using 14.260 gives

$$\forall (x, z) \in W_o \quad D_1g(x, z) = -[D_2f(x, g(x, z))]^{-1} \circ D_1f(x, g(x, z)) \quad (14.266)$$

proving (3.b) of the theorem.

To prove (3.c) note that

$$\begin{aligned} (\pi_2^{X \times Z} \circ (2 \rightarrow (x, z)))(t) &= \pi_2^{X \times Z}((2 \rightarrow (x, z))(t)) \\ &= \pi_2^{X \times Z}(x, t) \\ &= t \\ &= \mathbb{1}_Z(t) \end{aligned} \quad (14.267)$$

So

$$\begin{aligned} D_2(f \circ \psi)(x, z) &\stackrel{14.262}{=} D(\pi_2^{X \times Z} \circ (2 \rightarrow (x, z)))(z) \\ &\stackrel{14.267}{=} D(\mathbb{1}_Z)(z) \\ &= \mathbb{1}_Z \end{aligned} \quad (14.268)$$

Further

$$\begin{aligned} (\psi \circ (2 \rightarrow (x, z)))(t) &= \psi((2 \rightarrow (x, z))(t)) \\ &= \psi(x, t) \\ &\stackrel{14.261}{=} (x, g(x, t)) \\ &= (C_{Z, X, x}(t), (g \circ (2 \rightarrow (x, z)))(t)) \end{aligned}$$

proving that

$$\pi_1^{X \times Z} \circ (\psi \circ (2 \rightarrow (x, z))) = C_{Z, X, x} \text{ and } \pi_2^{X \times Z} \circ (\psi \circ (2 \rightarrow (x, z))) = g \circ (2 \rightarrow (x, z)) \quad (14.269)$$

We calculate now  $D_2(f \circ \psi)(x, z)$  in another way

$$\begin{aligned} (D_2(f \circ \psi)(x, z))(h) &\stackrel{14.36}{=} \\ (Df(\psi(x, z)) \circ D_2\psi(x, z))(h) &\stackrel{\text{def}}{=} \\ Df(\psi(x, z)) \circ D((\psi \circ (2 \rightarrow (x, z)))(x, z))(h) &\stackrel{14.269, 14.43}{=} \\ (Df(\psi(x, z)) \circ (DC_{Z, X, x}, D(g \circ (2 \rightarrow (x, z)))(x)))(h) &= \\ (Df(\psi(x, z)) \circ (0, D_2g(x, z)))(h) &= \\ Df(\psi(x, z))(0, D_2g(x, z)(h)) &\stackrel{14.35}{=} \\ D_1f(\psi(x, z))(0) + D_2f(\psi(x, z))(D_2g(x, z)(h)) &= \\ D_2f(\psi(x, z))(D_2g(x, z)(h)) &= \\ (D_2f(\psi(x, z)) \circ D_2g(x, z))(h) &\stackrel{14.261}{=} \\ (D_2f(x, g(x, z)) \circ D_2g(x, z)) & \end{aligned}$$

proving that  $D_2(f \circ \psi)(x, z) = D_2f(x, g(x, z)) \circ D_2g(x, z)$  or using 14.268  $\mathbb{1}_Z = D_2f(x, g(x, z)) \circ D_2g(x, z)$  giving using 14.260 that

$$\forall (x, z) \in W_0 \text{ we have } D_2g(x, z) = [D_2f(x, g(x, z))]^{-1} \quad (14.270)$$

proving (3.c) of the theorem.  $\square$

A trivial consequence of the implicit theorem is

**Theorem 14.170.** Let  $\langle X, \|\cdot\|_X \rangle$ ,  $\langle Y, \|\cdot\|_Y \rangle$  and  $\langle Z, \|\cdot\|_Z \rangle$  be Banach spaces,  $U \subseteq X$ ,  $V \subseteq Y$  be open sets and let  $f: U \times V \rightarrow Z$  be a differentiable function of class  $C^r$ ,  $r \in \mathbb{N}_0$ . Assume that  $\exists x_0 \in U$ ,  $\exists y_0 \in V$  such that  $D_2f(x_0, y_0): Y \rightarrow Z$  is a isomorphism. Then there exists a open set  $W$  with  $x_0 \in W \subseteq U$  and a function  $g: W \rightarrow V$  such that  $g(x_0) = y_0$  and  $\forall x \in W$  we have  $f(x, g(x)) = f(x_0, z_0)$ . Further  $\forall x \in W$   $(D_2f(x, g(x)))^{-1}$  exist and  $Dg(x) = -(D_2f(x, g(x)))^{-1} \circ D_1f(x, g(x))$

**Proof.** Using the Implicit Function Theorem [see 14.169] there exists a open set  $W_0 \subseteq U \times Z$  with  $(x_0, f(x_0, y_0)) \in W_0$ , a open set  $V_0 \subseteq V$  and a  $C^r$  function  $g': W_0 \rightarrow V_0 \subseteq V$  such that

$$g'(x_0, f(x_0, y_0)) = y_0 \quad (14.271)$$

and  $\forall (x, y) \in W_0$  we have

$$[D_2f(x, g'(x, z))]^{-1} \text{ exists} \quad (14.272)$$

$$f(x, g'(x, z)) = z \quad (14.273)$$

$$D_1g'(x, z) = -[D_2f(x, g'(x, z))]^{-1} \circ D_1f(x, g'(x, z)) \quad (14.274)$$

$$D_2g'(x, z) = [D_2f(x, g'(x, z))]^{-1} \quad (14.275)$$

Using the definition of the product topology on  $X \times Y$  [see 11.34 and 11.39] there exist open sets there exists a open set  $W$  in  $X$  and a open set  $R$  in  $Z$  such that

$$W \times R \subseteq U \times Z \text{ and } (x_0, f(x_0, y_0)) \in W \times R$$

then  $(x_0, f(x_0, y_0)) \in W \times R$  and  $\forall x \in W$  we have that  $(x, f(x_0, y_0)) \in U \times Z$  so that the following function is well defined

$$g: W \rightarrow V \text{ by } g(x) = g'(x, f(x_0, y_0)) \quad (14.276)$$

then we have

$$g(x_0) = g'(x_0, f(x_0, y_0)) = y_0 \quad (14.277)$$

Further  $f(x, g(x)) = f'(x, g'(x, f(x_0, y_0))) \underset{14.273}{=} f(x_0, y_0)$  hence

$$f(x, g(x)) = f(x_0, y_0) \quad (14.278)$$

Further

$$\begin{aligned} (g' \circ (1 \rightarrow (x_0, f(x_0, y_0))))(x) &= g'((1 \rightarrow (x_0, f(x_0, y_0))(x))) \\ &= g'(x, f(x_0, y_0)) \\ &= g(x) \end{aligned}$$

which proves that

$$g = g' \circ (1 \rightarrow (x_0, f(x_0, y_0)))$$

Using now the Chain Rule [see 14.83] we have using the fact that  $(1 \rightarrow (x_0, f(x_0, y_0)))$  is  $C^\infty$  and  $g'$  is  $C^r$  that

$$g \text{ is of class } C^r$$

and for  $x \in W$

$$\begin{aligned} Dg(x) &= D(g' \circ (1 \rightarrow (x_0, f(x_0, y_0))))(x) \\ &\stackrel{\text{def}}{=} D_1g'(x, f(x_0, y_0)) \\ &= -[D_2f(x, g'(x, f(x_0, y_0)))]^{-1} \circ D_1f(x, g'(x, f(x_0, y_0))) \\ &= -[D_2f(x, g(x))]^{-1} \circ D_1f(x, g(x)) \end{aligned} \quad (14.279)$$

The theorem is then proved by 14.276, 14.277 and 14.278.  $\square$

# Chapter 17

## The extended real numbers

### 17.1 Extented reals

In measure theory we have to work with series of positive real numbers that will not always convert. By extending the set of the real numbers we are able to deal with these series, in addition the supremum and infimum of non empty sets will always exists. The disadvantage that we will encounter is that the sum of extended reals is not always well defined.

First we extend the set of the reals with two non real numbers  $\infty$  and  $-\infty$ .

**Lemma 17.1.** *There exists at least two sets that are not a element of  $\mathbb{R}$ .*

**Proof.** Using the definition of the real numbers [see 8.1 and 8.2] we have that  $\emptyset \notin \mathbb{R}$  and  $\mathbb{Q} \notin \mathbb{R}$ .  $\square$

Using the above lemma and the fact that a set is by definition a element the following definition of extended reals make sense

**Definition 17.2.** *The set  $\bar{\mathbb{R}} = \mathbb{R} \cup \{\infty, -\infty\}$  where  $\infty, -\infty \notin \mathbb{R}$  and  $\infty \neq -\infty$*

The original real numbers are now called finite real numbers in the set of the extended reals

**Definition 17.3.**  *$x \in \bar{\mathbb{R}}$  is finite iff  $x \in \mathbb{R}$  (or in other words  $x \neq \infty, -\infty$ )*

We extend now the order relation  $\leq_{\mathbb{R}}$  on  $\mathbb{R}$  [see 9.27] to a order relation on  $\bar{\mathbb{R}}$  that will make  $\langle \bar{\mathbb{R}}, \leq \rangle$  a fully ordered set. By doing so we will show that as a benefit every non empty set in  $\bar{\mathbb{R}}$  has a supremum and a infimum.

**Definition 17.4.** *Given the fully ordered set  $\langle \bar{\mathbb{R}}, \leq \rangle$  [see 9.27] define  $\leq \subseteq \bar{\mathbb{R}} \times \bar{\mathbb{R}}$  by*

1.  $-\infty \leq \infty$
2.  $-\infty \leq -\infty$
3.  $\infty \leq \infty$
4.  $\forall x \in \mathbb{R}$  we have  $-\infty \leq x$  and  $x \leq \infty$
5.  $\forall x, y \in \mathbb{R}$  with  $x \leq_{\mathbb{R}} y$  we have  $x \leq y$

**Note 17.5.** As  $\{-\infty, \infty\} \cap \mathbb{R} = \emptyset$  and  $-\infty \neq \infty$  we have

1.  $-\infty \neq \infty$
2.  $\forall x \in \mathbb{R}$  we have  $-\infty < x$  and  $x < \infty$

**Theorem 17.6.**  *$\langle \bar{\mathbb{R}}, \leq \rangle$  is fully ordered*

**Proof.**

**reflexitivity.** The following cases occurs for  $x \in \bar{\mathbb{R}}$

$x = \infty$ . then by definition  $x \leq x$

$x = -\infty$ . then by definition  $x \leq x$

$x \in \mathbb{R}$ . then as  $x \leq_{\mathbb{R}} x \Rightarrow x \leq x$

proving reflexivity.

**anti-symmetry.** Let  $x, y \in \overline{\mathbb{R}}$  with  $x \leq y \wedge y \leq x$  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  $y \not\leq x$  this case will not apply.

$x \in \mathbb{R} \wedge y = \infty$ . then as by definition  $y \not\leq x$  this case does not apply

$x = \infty \wedge y = -\infty$ . then as by definition  $x \not\leq y$  this case does not apply

$x = -\infty \wedge y = -\infty$ . then  $x = y$

$x \in \mathbb{R} \wedge y = -\infty$ . then as by definition  $x \not\leq y$  the case does not apply

$x = \infty \wedge y \in \mathbb{R}$ . then as by definition  $x \not\leq y$  this case does not apply

$x = -\infty \wedge y \in \mathbb{R}$ . then as by definition  $y \not\leq x$  this case does not apply

$x \in \mathbb{R} \wedge y \in \mathbb{R}$ . then by definition  $x \leq y$  and  $y \leq x$  from which it follows that  $x = y$

so in all the cases where  $x \leq y \wedge y \leq x$  we have  $x = y$

**transitivity.** Let  $x, y, z \in \overline{\mathbb{R}}$  with  $x \leq y \wedge y \leq z$  then the following cases must be considered for  $x, y, z \in \overline{\mathbb{R}}$ :

$x = \infty \wedge y = \infty \wedge z = \infty$ . then  $x \leq z$

$x = -\infty \wedge y = \infty \wedge z = \infty$ . then  $x \leq z$

$x \in \mathbb{R} \wedge y = \infty \wedge z = \infty$ . then  $x \leq z$

$x = \infty \wedge y = -\infty \wedge z = \infty$ . then as  $x \not\leq y$  this case does not count

$x = -\infty \wedge y = -\infty \wedge z = \infty$ . then  $x \leq z$

$x \in \mathbb{R} \wedge y = -\infty \wedge z = \infty$ . then as  $x \not\leq y$  this case does not count

$x = \infty \wedge y \in \mathbb{R} \wedge z = \infty$ . then as  $x \not\leq y$  this case does not count

$x = -\infty \wedge y \in \mathbb{R} \wedge z = \infty$ . then  $x \leq z$

$x \in \mathbb{R} \wedge y \in \mathbb{R} \wedge z = \infty$ . then  $x \leq z$

$x = \infty \wedge y = \infty \wedge z = -\infty$ . then as  $y \not\leq z$  this case does not count

$x = -\infty \wedge y = \infty \wedge z = -\infty$ . then as  $y \not\leq z$  this case does not count

$x \in \mathbb{R} \wedge y = \infty \wedge z = -\infty$ . then as  $y \not\leq z$  this case does not count

$x = \infty \wedge y = -\infty \wedge z = -\infty$ . then as  $x \not\leq y$  this case does not count

$x = -\infty \wedge y = -\infty \wedge z = -\infty$ . then  $x \leq z$

$x \in \mathbb{R} \wedge y = -\infty \wedge z = -\infty$ . then as  $x \not\leq y$  this case does not count

$x = \infty \wedge y \in \mathbb{R} \wedge z = -\infty$ . then as  $y \not\leq z$  this case does not count

$x = -\infty \wedge y \in \mathbb{R} \wedge z = -\infty$ . then as  $y \not\leq z$  this case does not count

$x \in \mathbb{R} \wedge y \in \mathbb{R} \wedge z = -\infty$ . then as  $y \not\leq z$  this case does not count

$x = \infty \wedge y = \infty \wedge z \in \mathbb{R}$ . then as  $y \not\leq z$  this case does not count

$x = -\infty \wedge y = \infty \wedge z \in \mathbb{R}$ . then as  $y \not\leq z$  this case does not count

$x \in \mathbb{R} \wedge y = \infty \wedge z \in \mathbb{R}$ . then as  $y \not\leq z$  this case does not count

$x = \infty \wedge y = -\infty \wedge z \in \mathbb{R}$ . then as  $x \not\leq y$  this case does not count

$x = -\infty \wedge y = -\infty \wedge z \in \mathbb{R}$ . then  $x \leq z$

$x \in \mathbb{R} \wedge y = -\infty \wedge z \in \mathbb{R}$ . then as  $x \not\leq y$  this case does not count

$x = \infty \wedge y \in \mathbb{R} \wedge z \in \mathbb{R}$ . then as  $x \not\leq y$  this case does not count

$x = -\infty \wedge y \in \mathbb{R} \wedge z \in \mathbb{R}$ . then  $x \leq z$

$x \in \mathbb{R} \wedge y \in \mathbb{R} \wedge z \in \mathbb{R}$ . then  $x \leq z$

so in all cases that count we have  $x \leq z$

**fully-ordered.** The following cases must be considered for  $x, y \in \bar{\mathbb{R}}$ :

$x = \infty \wedge y = \infty$ . then  $x \leq y$

$x = -\infty \wedge x = \infty$ . then  $x \leq y$

$x \in \mathbb{R} \wedge y = \infty$ . then  $x \leq y$

$x = \infty \wedge y = -\infty$ . then  $y \leq x$

$x = -\infty \wedge y = -\infty$ . then  $x \leq y$

$x \in \mathbb{R} \wedge y = -\infty$ . then  $y \leq x$

$x = \infty \wedge y \in \mathbb{R}$ . then  $y \leq x$

$x = -\infty \wedge y \in \mathbb{R}$ . then  $x \leq y$

$x \in \mathbb{R} \wedge y \in \mathbb{R}$ . then either  $x \leq_R y \Rightarrow x \leq y$  or  $y \leq_R x \Rightarrow y \leq x$

so in all possible cases we have either  $x \leq y$  or  $y \leq x$

□

**Proposition 17.7.** Let  $x, y \in \bar{\mathbb{R}}$  such that  $x < y$  then  $x \neq \infty$  and  $y \neq -\infty$

**Proof.** Assume that  $x = \infty$  then by definition  $y \leq x \xrightarrow{x < y} y < y$  a contradiction so we must have that  $x \neq \infty$ . Assume that  $y = -\infty$  then by definiton  $y \leq x \xrightarrow{x < y} y < y$  a contradiction so we must have that  $y \neq -\infty$ .

□

Some of the properties of the real numbers (like the density of the rational numbers in the set of real numbers still apply for extended real numbers).

**Proposition 17.8.** Let  $x, y \in \bar{\mathbb{R}}$  with  $x < y$  then  $n \in \mathbb{N}_0$  such that  $x + \frac{1}{n} < y$

**Proof.** As  $x < y$  we can not have  $y = -\infty$  [see 17.7i], so we must only consider the following cases for  $x, y$ :

$y \in \mathbb{R}$ . As  $x < y$  we can not have  $x = \infty$  [see 17.7], so we must have for  $x$  either

$x \in \mathbb{R}$ . then using 9.56 there exists a  $n \in \mathbb{N}_0$  such that  $x < x + \frac{1}{n} < y$

$x = -\infty$ . then for  $n = 2 \in \mathbb{N}_0$  we have  $x + \frac{1}{n} = -\infty + \frac{1}{2} = -\infty < y$

$y = \infty$ . As  $x < y$  we can not have  $x = \infty$  [see 17.7], so we must have for  $x$  either

$x \in \mathbb{R}$ . then for  $n = 2 \in \mathbb{N}_0$  we have  $x + \frac{1}{n} < \infty = y$

$x = -\infty$ . then for  $n = 2 \in \mathbb{N}_0$  we have  $x + \frac{1}{n} = -\infty + \frac{1}{2} = -\infty < y$

□

We have the following analog to the density of the rational numbers in  $\mathbb{R}$

**Theorem 17.9.** If  $x, y \in \bar{\mathbb{R}}$  with  $x < y$  then  $\exists z \in \mathbb{Q}$  such that  $x < z < y$

**Proof.** As  $x < y$  we have by 17.7 that  $x \neq -\infty$ ,  $y \neq \infty$  and  $x \neq y$  So we have to consider only the following cases for  $x$  and  $y$ :

$x \in \mathbb{R} \wedge y = \infty$ . then as  $x < x + 1$  there exists by 9.58 a  $z \in \mathbb{Q}$  such that  $x < z < x + 1 < \infty = y \Rightarrow x < z < y$ .

$x = -\infty \wedge y = \infty$ . take then  $z = 0 \in \mathbb{Q} \Rightarrow x < z < y$ .

$x \in \mathbb{R} \wedge y \in \mathbb{R}$ . using 9.58 there exists then a  $z \in \mathbb{Q}$  such that  $x < z < y$ .

$x = -\infty \wedge y \in \mathbb{R}$ . then as  $y - 1 < y$  there exists by 9.58 a  $z \in \mathbb{Q}$  such that  $-\infty < y - 1 < z < y$  so that  $x < z < y$ .

□

We look now at the infimum and supremum on the extended reals. To avoid confusion with the infimum and supremum on the real numbers and the extended real numbers we introduce the following notation.

**Notation 17.10.** Let  $\emptyset \neq A \subseteq \mathbb{R}$  be a non empty real set then  $\sup_{\mathbb{R}}(A), \inf_{\mathbb{R}}(A)$  is the supremum and infimum of  $A$  in  $\langle \mathbb{R}, \leq_{\mathbb{R}} \rangle$  (where  $\leq_{\mathbb{R}}$  is the canonical order in  $\mathbb{R}$  see 8.36)

**Notation 17.11.** Let  $\emptyset \neq A \subseteq \overline{\mathbb{R}}$  be a non empty extended real set then  $\sup(A), \inf(A)$  is the supremum and infimum of  $A$  in  $\langle \overline{\mathbb{R}}, \leq \rangle$

We prove now that in  $\langle \overline{\mathbb{R}}, \leq \rangle$  every non empty set has a supremum and infimum, which the primary reason that we introduced the set of the extended real numbers.

**Theorem 17.12.** Let  $\emptyset \neq A \subseteq \overline{\mathbb{R}}$  then  $\sup(A)$  and  $\inf(A)$  exist. Further more we have for  $A$  the following exclusive cases possible

1. If  $A = \{-\infty\}$  then  $\sup(A) = -\infty$
2. If  $\infty \in A$  then  $\sup(A) = \infty$
3. If  $\infty \notin A \neq \{\infty\}$  then we have
  - a. If  $\exists u \in \mathbb{R} \vdash \forall x \in A$  we have  $x \leq u$  then  $\sup(A) = \sup_{\mathbb{R}}(A \cap \mathbb{R})$
  - b. If  $\exists u \in \mathbb{R} \vdash \forall x \in A$  we have  $x \leq u$  then  $\sup(A) = \infty$
4. If  $A = \{\infty\}$  then  $\inf(A) = \infty$
5. If  $-\infty \in A$  then  $\inf(A) = -\infty$
6. If  $-\infty \notin A \neq \{\infty\}$  then we have
  - a. If  $\exists l \in \mathbb{R} \vdash \forall x \in A$  we have  $l \leq x$  then  $\inf(A) = \inf_{\mathbb{R}}(A \cap \mathbb{R})$
  - b. If  $\forall l \in \mathbb{R} \vdash \exists x \in A$  such that  $l \not\leq x$  then  $\inf(A) = -\infty$

**Proof.** Lets first consider the supremum then we must consider the following cases:

$A = \{-\infty\}$ . then as  $-\infty \leq -\infty$ ,  $-\infty$  is trivially a upper bound, further if  $u \in \overline{\mathbb{R}}$  is another upper bound then as  $-\infty \leq u$  we have that

$$\sup(A) = -\infty$$

$\infty \in A$ . then trivially  $\infty$  is a upper bound of  $A$  [actually of every set in  $\mathbb{R}$ ]. Further if  $u$  is a another upper bound of  $A$  we must have that  $\infty \leq u$  [ $\infty \in A$ ] so that

$$\sup(A) = \infty$$

$\infty \notin A \neq \{-\infty\}$ . Now we consider the following cases for  $A$ :

$\exists u \in \mathbb{R} \vdash \forall x \in A$  we have  $x \leq u$ . then  $\forall x \in A \cap \mathbb{R} \subseteq A$  we have  $x \leq u \xrightarrow{x, u \in \mathbb{R}} x \leq_{\mathbb{R}} u$  so that  $A \cap \mathbb{R}$  is bounded above [using  $\leq_{\mathbb{R}}$ . Hence as  $\mathbb{R}$  is conditional complete [see 9.43]  $S = \sup_{\mathbb{R}}(A \cap \mathbb{R})$  exists. Let  $x \in A$  then we have as  $\infty \notin A$  either:

$x \in A \cap \mathbb{R}$ . then as a supremum is a upper bound we have  $x \leq_{\mathbb{R}} S \Rightarrow x \leq S$

$x = -\infty$ . then  $x \leq S$

So we have that  $S$  is a upper bound of  $A$  [using  $\leq$ ]. Let  $U$  be another upper bound of  $A$  [using  $\leq$ ] then we have either:

$U \in \mathbb{R}$ . in this case as  $U$  is also a upper bound of  $A \cap \mathbb{R}$  [using  $\leq$ ] we have  $\forall a \in A \cap \mathbb{R}$  that  $a \leq u \xrightarrow{a, U \in \mathbb{R}} a \leq U$ . It follows that  $U$  is another upper bound of  $A \cap \mathbb{R}$  [using  $\leq_{\mathbb{R}}$ ] so that  $S \leq_{\mathbb{R}} U$ , hence  $S \leq U$ .

$U = \infty$ . then trivially  $S \leq U$

$U = -\infty$ . then  $\forall x \in A$  we have  $-\infty \leq x \leq -\infty$  contradicting  $A \neq \{-\infty\}$  so this will not occur.

Hence in all valid cases we have  $S \leq U$  which proves that

$$\sup(A) = \sup_{\mathbb{R}}(A)$$

$\forall u \in \mathbb{R} \vdash \exists x \in A \text{ such that } x \not\leq u$ . As  $\forall x \in \bar{\mathbb{R}}$  we have  $x \leq \infty$ ,  $\infty$  is clearly a upper bound of  $A$  [using  $\leq$ ]. Let  $u$  be another upper bound of  $A$  [using  $\leq$ ] and assume that  $\infty \not\leq u$ , so that  $u < \infty$ , then we must have either:

$u = -\infty$ . then  $\forall x \in A$  we have  $-\infty \leq x \leq -\infty$  contradicting  $A \neq \{-\infty\}$ .

$u \in \mathbb{R}$ . then  $\exists x \in A$  such that  $x \not\leq u$  contradicting the fact that  $u$  is a upper bound.

As in all cases we have a contradcition we must have that  $u \leq \infty$  which proves that

$$\sup(A) = \infty$$

So in all possible cases for  $A$  we have that

$$\sup(A) \text{ exists} \quad (17.1)$$

Now for the existance of the infimum we must consider the following cases:

$A = \{\infty\}$ . then as  $\infty \leq \infty$ ,  $\infty$  is a lower bound of  $A$ , further if  $l \in \bar{\mathbb{R}}$  is another lower bound then  $l \leq \infty$  which proves that

$$\inf(A) = \infty$$

$-\infty \in A$ . then  $-\infty$  is a lower bound of  $A$  [actually of every set in  $\mathbb{R}$ ]. If  $l$  is another lower bound of  $A$  then we must have  $l \leq -\infty$  so that

$$\inf(A) = -\infty$$

$-\infty \notin A \wedge A \neq \{\infty\}$ . Then we must consider the following possible cases:

$\exists l \in \mathbb{R} \vdash \forall x \in A \text{ we have } l \leq x$ . then  $\forall x \in A \cap \mathbb{R} \subseteq A$  we have  $l \leq x \Rightarrow l \leq_{\mathbb{R}} x$  so that  $l$  is a lower bound of  $A \cap \mathbb{R}$  [using  $\leq_{\mathbb{R}}$ ]. Hence as  $\mathbb{R}$  is conditional complete [see 9.43]  $I = \inf_{\mathbb{R}}(A \cap \mathbb{R})$  exists. Let  $x \in A$  then we have as  $-\infty \notin A$  either:

$x \in A \cap \mathbb{R}$ . then by the definition of a infimum we have  $I \leq_{\mathbb{R}} x \Rightarrow I \leq x$

$x = \infty$ . then  $I \leq x$

So it follows that  $I$  is a upper bound of  $A$  [using  $\leq$ ]. Let  $L$  be another lower bound of  $A$  [using  $\leq$ ] then we have either

$L \in \mathbb{R}$ . then  $L$  is also a lower bound of  $A \cap \mathbb{R}$  [using  $\leq$ ], so that  $\forall a \in A \cap \mathbb{R}$  we have  $L \leq a \Rightarrow L \leq_{\mathbb{R}} a$ , proving that  $L$  is also a lower bound of  $A \cap \mathbb{R}$  [using  $\leq_{\mathbb{R}}$ ] so that  $L \leq_{\mathbb{R}} I$ , from which it follows that  $L \leq I$

$L = \infty$ . then  $\forall a \in A$  we have  $\infty \leq a \leq \infty \Rightarrow a = \infty$  contradicting  $A \neq \{\infty\}$ , so this case does not apply.

$L = -\infty$ . then  $L \leq I$

So in all valid cases we have  $L \leq I$  which proves that

$$\inf(A) = \inf_{\mathbb{R}}(A)$$

$\forall l \in \mathbb{R} \vdash \exists x \in A \text{ such that } l \not\leq x$ . as  $\forall x \in \bar{\mathbb{R}}$  we have  $-\infty \leq x$ ,  $-\infty$  is a lower bound of  $A$  [using  $\leq$ ]. Let  $l$  be another lower bound of  $A$  [using  $\leq$ ] and assume that  $l \not\leq -\infty$  so that  $-\infty < l$ . then we must have either

$l = \infty$ . then  $\forall x \in A$  we have  $\infty \leq x \leq \infty$  contradicting  $A = \{\infty\}$

$l \in \mathbb{R}$ . then  $\exists x \in A$  usch that  $l \not\leq x$  contradicting the fact that  $l$  is a lower bound,

As in all cases we have a contradiction we must have that  $l \leq -\infty$  which proves that

$$\inf(A) = -\infty$$

So in all possible cases for  $A$  we have that

$$\inf(A) \text{ exists} \quad (17.2) \quad \square$$

**Corollary 17.13.** *If  $A \subseteq \mathbb{R}$  then  $\sup(A) \neq -\infty$  and  $\inf(A) \neq \infty$*

**Proof.** For  $A \subseteq \mathbb{R}$  we have  $\infty \notin A \neq \{\infty\}$  and  $-\infty \notin A \neq \{-\infty\}$  so that using 17.12 we have  $\sup(A) = \infty \vee \sup(A) \in \mathbb{R}$  and  $\inf(A) = -\infty \vee \inf(A) \in \mathbb{R}$ . So  $\sup(A) \neq -\infty$  and  $\inf(A) \neq \infty$ .  $\square$

The next theorem shows the necessary and sufficient constraint for the supremum and infimum to be not finite.

**Corollary 17.14.** *Let  $\emptyset \neq A \subseteq \bar{\mathbb{R}}$  then*

1.  $\sup(A) = \infty$  if and only if  $\infty \in A$  or  $\forall u \in \mathbb{R}$  there exists  $a \in A \cap \mathbb{R}$  with  $a \not\leq u$  [in other words  $A \cap \mathbb{R}$  is not bounded above using  $\leq_{\mathbb{R}}$ ]
2.  $\inf(A) = -\infty$  if and only if  $-\infty \in A$  or  $\forall l \in \mathbb{R}$  there exists  $a \in A \cap \mathbb{R}$  with  $l \not\leq a$  [in other words  $A \cap \mathbb{R}$  is not bounded below using  $\leq_{\mathbb{R}}$ ]
3.  $\sup(A) = -\infty$  if and only if  $A = \{-\infty\}$
4.  $\inf(A) = \infty$  if and only if  $A = \{\infty\}$

**Proof.**

1. This follows from 17.12 (2), {3.b})
2. This follows from 17.12 (5), {6.b})
3. This follows from 17.12 (5), {6.b})
4. This follows from 17.12 (5), {6.b})

$\square$

We show now that the supremum (infimum) in  $\bar{\mathbb{R}}$  is an extension of the supremum and infimum in  $\mathbb{R}$  (if they exists)

**Theorem 17.15.** *Let  $\emptyset \neq A \subseteq \bar{\mathbb{R}}$  then*

1. *If  $A \subseteq ]-\infty, \infty]$  and  $\sup(A) < \infty$  then  $A \subseteq \mathbb{R}$ ,  $\sup_{\mathbb{R}}(A)$  exists and  $\sup(A) = \sup_{\mathbb{R}}(A)$*
2. *If  $A \subseteq [-\infty, \infty[$  and  $-\infty < \inf(A)$  then  $A \subseteq \mathbb{R}$ ,  $\inf_{\mathbb{R}}(A)$  exists and  $\inf(A) = \inf_{\mathbb{R}}(A)$*

**Proof.**

1. If  $\infty \in A$  then by the above theorem (see 17.14) we have that  $\sup(A) = \infty$  a contradiction so we have

$$A \subseteq \mathbb{R}$$

Further using 17.12 we have as  $\sup(A) < \infty$  that  $\sup(A) = \sup_{\mathbb{R}}(A)$ .

2. If  $-\infty \in A$  then by the above theorem (see 17.14) we have that  $\sup(A) = \infty$  a contradiction so we have

$$A \subseteq \mathbb{R}$$

Further using 17.12 we have as  $-\infty < \inf(A)$  that  $\inf(A) = \inf_{\mathbb{R}}(A)$ .  $\square$

**Note 17.16.** The condition  $A \subseteq ]-\infty, \infty]$  in (1) above or  $A \subseteq [-\infty, \infty[$  in (2) is needed for if  $A = \{-\infty\}$  then  $A \not\subseteq \mathbb{R}$  and  $\sup(A) = -\infty \neq \sup_{\mathbb{R}}(A)$  and if  $A = \{\infty\}$  then  $A \subseteq \mathbb{R}$  and  $\inf(A) = \infty \neq \inf_{\mathbb{R}}(A)$ .

**Convention 17.17.** For the rest of the document, if  $A \subseteq \mathbb{R}$  then  $\sup(A) < \infty$  means  $\sup_{\mathbb{R}}(A)$  exists and  $-\infty < \inf(A)$  means  $\inf_{\mathbb{R}}(A)$  exists. Where  $\sup_{\mathbb{R}}(A) \in \mathbb{R}$  and  $\inf_{\mathbb{R}}(A) \in \mathbb{R}$ . This is of course consistent with the above theorem.

To avoid any confusion we note the sum, product and division on  $\mathbb{R}$  as  $+_{\mathbb{R}}$ ,  $\cdot_{\mathbb{R}}$  and  $/_{\mathbb{R}}$ . We will extend the algebraic operations on  $\mathbb{R}$  to  $\overline{\mathbb{R}}$ , however it will not be possible to define  $+$  fully on  $\overline{\mathbb{R}} \times \overline{\mathbb{R}}$ .

**Definition 17.18.** On  $\overline{\mathbb{R}}$  we define the  $+$  and  $\cdot$  operations for the following cases

1. The function  $(+): (\overline{\mathbb{R}} \times \overline{\mathbb{R}}) \setminus \{(\infty, -\infty), (-\infty, \infty)\} \rightarrow \overline{\mathbb{R}}$  is defined by  $(+)(x, y) = x + y$  where  $+$  is defined as follows:

$$x + y = \begin{cases} \infty & \text{if } x = \infty \wedge y \in \overline{\mathbb{R}} \setminus \{-\infty\} \\ -\infty & \text{if } x = -\infty \wedge y \in \overline{\mathbb{R}} \setminus \{\infty\} \\ \infty & \text{if } x \in \overline{\mathbb{R}} \setminus \{\infty\} \wedge y = \infty \\ -\infty & \text{if } x \in \overline{\mathbb{R}} \setminus \{-\infty\} \wedge y = -\infty \\ x +_{\mathbb{R}} y & \text{if } x \in \mathbb{R} \wedge y \in \mathbb{R} \end{cases}$$

Note: Using the above definition it is easy to see that if  $x + y$  exists then  $y + x$  exists and  $x + y = y + x$ .

2. The function  $(\cdot): \overline{\mathbb{R}} \times \overline{\mathbb{R}} \rightarrow \overline{\mathbb{R}}$  is defined by  $(\cdot)(x, y)$  where  $x \cdot y$  is defined as follows:

$$x \cdot y = \begin{cases} \infty & \text{if } x = \infty \wedge 0 < y \\ -\infty & \text{if } x = \infty \wedge y < 0 \\ 0 & \text{if } x = \infty \wedge y = 0 \\ -\infty & \text{if } x = -\infty \wedge 0 < y \\ \infty & \text{if } x = -\infty \wedge y < 0 \\ 0 & \text{if } x = -\infty \wedge y = 0 \\ \infty & \text{if } 0 < x \wedge y = \infty \\ -\infty & \text{if } x < 0 \wedge y = \infty \\ 0 & \text{if } x = 0 \wedge y = \infty \\ -\infty & \text{if } 0 < x \wedge y = -\infty \\ \infty & \text{if } x < 0 \wedge y = -\infty \\ 0 & \text{if } x = 0 \wedge y = -\infty \\ x \cdot_{\mathbb{R}} y & \text{if } x \in \mathbb{R} \wedge y \in \mathbb{R} \end{cases}$$

3. We define  $(-): \overline{\mathbb{R}} \rightarrow \overline{\mathbb{R}}$  by  $(-)(x) = (-1) \cdot x$  and note  $(-)(x)$  as  $-x$ . Using (2) we have then that

$$-x = \begin{cases} \infty & \text{if } x = -\infty \\ -\infty & \text{if } x = \infty \\ -x & \stackrel{\text{def}}{=} (-1) \cdot_{\mathbb{R}} x \end{cases}$$

4. Let  $(x, y) \in (\overline{\mathbb{R}} \times \overline{\mathbb{R}}) \setminus \{(\infty, \infty), (-\infty, -\infty)\}$  then  $x - y = x + (-y) = x + ((-1) \cdot y)$

5. Let  $x \in \overline{\mathbb{R}} \setminus \{0\}$  then

$$\frac{1}{x} = \begin{cases} 1 /_{\mathbb{R}} x & \text{if } x \in \mathbb{R} \setminus \{0\} \\ 0 & \text{if } x \in \{-\infty, \infty\} \end{cases}$$

**Note 17.19.** The product of two extended numbers is always defined, however the sum is not defined for the cases  $\infty + (-\infty)$  and  $-\infty + \infty$ .

**Lemma 17.20.** Let  $x, y \in \overline{\mathbb{R}}$  we have

1. If  $0 < x \wedge 0 < y$  then  $0 < x \cdot y$
2. If  $x < 0 \wedge y < 0$  then  $0 < x \cdot y$
3. If  $x < 0 \wedge 0 < y$  then  $x \cdot y < 0$
4. If  $0 < x \wedge y < 0$  then  $x \cdot y < 0$
5. If  $x, y \in \overline{\mathbb{R}}$  with  $0 < x \leq y$  then  $\frac{1}{y} \leq \frac{1}{x}$

**Proof.** Take  $\mathbb{R}_+ = \{x \in \mathbb{R} | 0 < x\}$  and  $\mathbb{R}_- = \{x \in \mathbb{R} | x < 0\}$

1. If  $0 < x \wedge 0 < y$  then we have either

$$x = \infty \wedge y = \infty. \text{ then } x \cdot y = \infty \Rightarrow 0 < x \cdot y$$

$$x = \infty \wedge y \in \mathbb{R}_+. \text{ then } x \cdot y = \infty \Rightarrow 0 < x \cdot y$$

$$x \in \mathbb{R}_+ \wedge y = \infty. \text{ then } x \cdot y = \infty \Rightarrow 0 < x \cdot y$$

$$x \in \mathbb{R}_+ \wedge y \in \mathbb{R}_+. \text{ then } 0 < x \cdot_{\mathbb{R}} y = x \cdot y$$

2. If  $x < 0 \wedge y < 0$  then we have either then we have either

$$x = -\infty \wedge y = -\infty. \text{ then } x \cdot y = \infty \Rightarrow 0 < x \cdot y$$

$$x = -\infty \wedge y \in \mathbb{R}_-. \text{ then } x \cdot y = \infty \Rightarrow 0 < x \cdot y$$

$$x \in \mathbb{R}_- \wedge y = -\infty. \text{ then } x \cdot y = \infty \Rightarrow 0 < x \cdot y$$

$$x \in \mathbb{R}_- \wedge y \in \mathbb{R}_-. \text{ then } 0 < x \cdot_{\mathbb{R}} y = x \cdot y$$

3. If  $x < 0 \wedge 0 < y$  then we have either then we have either

$$x = -\infty \wedge y = \infty. \text{ then } x \cdot y = -\infty \Rightarrow x \cdot y < 0$$

$$x = -\infty \wedge y \in \mathbb{R}_+. \text{ then } x \cdot y = -\infty \Rightarrow x \cdot y < 0$$

$$x \in \mathbb{R}_- \wedge y = \infty. \text{ then } x \cdot y = -\infty \Rightarrow x \cdot y < 0$$

$$x \in \mathbb{R}_- \wedge y \in \mathbb{R}_+. \text{ then } x \cdot y = x \cdot_{\mathbb{R}} y < 0$$

4. If  $0 < x \wedge y < 0$  then we have either then we have either

$$x = \infty \wedge y = -\infty. \text{ then } x \cdot y = -\infty \Rightarrow 0 < x \cdot y$$

$$x = \infty \wedge y \in \mathbb{R}_-. \text{ then } x \cdot y = -\infty \Rightarrow 0 < x \cdot y$$

$$x \in \mathbb{R}_+ \wedge y = -\infty. \text{ then } x \cdot y = -\infty \Rightarrow 0 < x \cdot y$$

$$x \in \mathbb{R}_+ \wedge y \in \mathbb{R}_-. \text{ then } x \cdot y = x \cdot_{\mathbb{R}} y < 0$$

5. As  $0 < x \leq y$  we have for  $x, y$  to consider the following cases

$$x = \infty \wedge y = \infty. \text{ then } \frac{1}{y} = 0 = \frac{1}{x} \text{ proving } \frac{1}{y} \leq \frac{1}{x}$$

$$x \in \mathbb{R} \setminus \{0\} \wedge y = \infty. \text{ then as } 0 < x \text{ we have } \frac{1}{x} = 1/x > 0 = \frac{1}{\infty} = \frac{1}{y} \text{ proving } \frac{1}{y} \leq \frac{1}{x}$$

$$x \in \mathbb{R} \setminus \{0\} \wedge y \in \mathbb{R} \setminus \{0\}. \text{ then } \frac{1}{y} = 1/y \leq 1/x = \frac{1}{x} \text{ proving } \frac{1}{y} \leq \frac{1}{x}$$

□

**Proposition 17.21.** *We have the following properties for the product of extended real numbers*

1.  $\forall x, y \in \overline{\mathbb{R}}$  we have

$$a. x \cdot y = 0 \text{ if and only if } x = 0 \text{ or } y = 0$$

$$b. x \cdot y \in \mathbb{R} \setminus \{0\} \text{ if and only if } x, y \in \mathbb{R}$$

$$c. x \cdot y = \infty \text{ if and only if}$$

$$(x = \infty \wedge 0 < y) \vee (x = -\infty \wedge y < 0) \\ \vee (0 < x \wedge y = \infty) \\ \vee (x < 0 \wedge y = -\infty)$$

$$d. x \cdot y = -\infty \text{ if and only if}$$

$$(x = -\infty \wedge 0 < y) \vee (x = \infty \wedge y < 0) \\ \vee (0 < x \wedge y = -\infty) \\ \vee (x < 0 \wedge y = \infty)$$

2.  $\forall x, y \in \overline{\mathbb{R}}$  we have  $x \cdot y = y \cdot x$

3.  $\forall x, y, z \in \bar{\mathbb{R}}$  we have  $(x \cdot y) \cdot z = x \cdot (y \cdot z)$

4.  $\forall x \in \bar{\mathbb{R}}$  we have  $1 \cdot x = x$

**Proof.**

1.

a.

$\Rightarrow$ . If  $x \cdot y = 0$  then we must consider the following cases for  $x$

$x \in \{-\infty, \infty\}$ . then using the definition of  $(\cdot)$  we must have  $y = 0$

$x \in \mathbb{R}$ . If  $y \in \{-\infty, \infty\}$  then using the definition of the  $\cdot$  we must have that  $x = 0$ , if  $y \in \mathbb{R}$  then  $x \cdot y = x \cdot_{\mathbb{R}} y$  so that we must have  $y = 0$  or  $x = 0$

so in all cases we have either  $x = 0$  or  $y = 0$

$\Leftarrow$ . this trivially follows from the definition of the product.

b. this trivially follows from the definition of the product

c. this trivially follows from the definition of the product

d. this trivially follows from the definition of the product

2. For  $x \cdot y$  we have either

$x \cdot y = \infty$ . then using (1.c) we have either

$(x = \infty \wedge 0 < y)$ . then  $y \cdot x = \infty = x \cdot y$

$(x = -\infty \wedge y < 0)$ . then  $y \cdot x = \infty = x \cdot y$

$(0 < x \wedge y = \infty)$ . then  $y \cdot x = \infty = x \cdot y$

$(x < 0 \wedge y = -\infty)$ . then  $y \cdot x = \infty = x \cdot y$

$x \cdot y = -\infty$ . then using (1.d) we have either

$(x = -\infty \wedge 0 < y)$ . then  $y \cdot x = -\infty = x \cdot y$

$(x = \infty \wedge y < 0)$ . then  $y \cdot x = -\infty = x \cdot y$

$(0 < x \wedge y = -\infty)$ . then  $y \cdot x = -\infty = x \cdot y$

$(x < 0 \wedge y = \infty)$ . then  $y \cdot x = -\infty = x \cdot y$

$x \cdot y = 0$ . then using (1.a) we have either  $x = 0$  so that  $y \cdot x = 0 = x \cdot y$  or  $y = 0$  so that  $y \cdot x = 0 = x \cdot y$

$x \cdot y \in \mathbb{R} \setminus \{0\}$ . then using (1.b) we have  $x, y \in \mathbb{R}$  so that  $y \cdot x = y \cdot_{\mathbb{R}} x = x \cdot_{\mathbb{R}} y = x \cdot y$

3. For  $x \cdot y$  we have the following possibilities to consider :

$x \cdot y = \infty$ . then using (1.c) we have either

$(x = \infty \wedge 0 < y)$ . then for  $z$  we must have either

$0 < z$ . then  $(x \cdot y) \cdot z = \infty$ , further using 17.20  $0 < y \cdot z$  so that  $x \cdot (y \cdot z) = \infty$  which proves that  $(x \cdot y) \cdot z = x \cdot (y \cdot z)$ .

$z < 0$ . then  $(x \cdot y) \cdot z = -\infty$ , further using 17.20  $y \cdot z < 0$  so that  $x \cdot (y \cdot z) = -\infty$  which proves that  $(x \cdot y) \cdot z = x \cdot (y \cdot z)$

$z = 0$ . then  $(x \cdot y) \cdot z = 0 = x \cdot (y \cdot z)$

$(x = -\infty \wedge y < 0)$ . then for  $z$  we must have either

$0 < z$ . then  $(x \cdot y) \cdot z = \infty$ , further using 17.20  $y \cdot z < 0$  so that  $x \cdot (y \cdot z) = \infty$  which proves that  $(x \cdot y) \cdot z = x \cdot (y \cdot z)$ .

$z < 0$ . then  $(x \cdot y) \cdot z = -\infty$ , further using 17.20  $0 < y \cdot z$  so that  $x \cdot (y \cdot z) = -\infty$  which proves that  $(x \cdot y) \cdot z = x \cdot (y \cdot z)$

$z = 0$ . then  $(x \cdot y) \cdot z = 0 = x \cdot (y \cdot z)$

$(0 < x \wedge y = \infty)$ . then for  $z$  we must have either

$0 < z$ . then  $(x \cdot y) \cdot z = \infty$ , further  $y \cdot z = \infty$  so that  $x \cdot (y \cdot z) = \infty$  which proves that  $(x \cdot y) \cdot z = x \cdot (y \cdot z)$ .

$z < 0$ . then  $(x \cdot y) \cdot z = -\infty$ , further  $y \cdot z = -\infty$  so that  $x \cdot (y \cdot z) = -\infty$  which proves that  $(x \cdot y) \cdot z = x \cdot (y \cdot z)$

$z = 0$ . then  $(x \cdot y) \cdot z = 0 = x \cdot (y \cdot z)$

$(x < 0 \wedge y = -\infty)$ . then for  $z$  we must have either

$0 < z$ . then  $(x \cdot y) \cdot z = \infty$ , further  $y \cdot z = -\infty$  so that  $x \cdot (y \cdot z) = \infty$  which proves that  $(x \cdot y) \cdot z = x \cdot (y \cdot z)$ .

$z < 0$ . then  $(x \cdot y) \cdot z = -\infty$ , further  $y \cdot z = \infty$  so that  $x \cdot (y \cdot z) = -\infty$  which proves that  $(x \cdot y) \cdot z = x \cdot (y \cdot z)$

$z = 0$ . then  $(x \cdot y) \cdot z = 0 = x \cdot (y \cdot z)$

$x \cdot y = -\infty$ . then using (1.d) we have either

$(x = -\infty \wedge 0 < y)$ . then for  $z$  we must have either

$0 < z$ . then  $(x \cdot y) \cdot z = -\infty$ , further using 17.20  $0 < y \cdot z$  so that  $x \cdot (y \cdot z) = -\infty$  which proves that  $(x \cdot y) \cdot z = x \cdot (y \cdot z)$

$z < 0$ . then  $(x \cdot y) \cdot z = \infty$ , further using 17.20  $y \cdot z < 0$  so that  $x \cdot (y \cdot z) = \infty$

$z = 0$ . then  $(x \cdot y) \cdot z = 0 = x \cdot (y \cdot z)$

$(x = \infty \wedge y < 0)$ . then for  $z$  we must have either

$0 < z$ . then  $(x \cdot y) \cdot z = -\infty$ , further using 17.20  $y \cdot z < 0$  so that  $x \cdot (y \cdot z) = -\infty$  proving that  $(x \cdot y) \cdot z = x \cdot (y \cdot z)$

$z < 0$ . then  $(x \cdot y) \cdot z = \infty$ , further using 17.20  $0 < y \cdot z$  so that  $x \cdot (y \cdot z) = \infty$  proving that  $(x \cdot y) \cdot z = x \cdot (y \cdot z)$

$z = 0$ . then  $(x \cdot y) \cdot z = 0 = x \cdot (y \cdot z)$

$(0 < x \wedge y = -\infty)$ . then for  $z$  we must have either

$0 < z$ . then  $(x \cdot y) \cdot z = -\infty$ , further  $y \cdot z = -\infty$  so that  $x \cdot (y \cdot z) = -\infty$  which proves that  $(x \cdot y) \cdot z = x \cdot (y \cdot z)$

$z < 0$ . then  $(x \cdot y) \cdot z = \infty$ , further  $y \cdot z = \infty$  so that  $x \cdot (y \cdot z) = \infty$  which proves that  $(x \cdot y) \cdot z = x \cdot (y \cdot z)$

$z = 0$ . then  $(x \cdot y) \cdot z = 0 = x \cdot (y \cdot z)$

$(x < 0 \wedge y = \infty)$ . then for  $z$  we must have either

$0 < z$ . then  $(x \cdot y) \cdot z = -\infty$ , further  $y \cdot z = \infty$  so that  $x \cdot (y \cdot z) = -\infty$  which proves that  $(x \cdot y) \cdot z = x \cdot (y \cdot z)$

$z < 0$ . then  $(x \cdot y) \cdot z = \infty$ , further  $y \cdot z = -\infty$  so that  $x \cdot (y \cdot z) = \infty$  which proves that  $(x \cdot y) \cdot z = x \cdot (y \cdot z)$

$z = 0$ . then  $(x \cdot y) \cdot z = 0 = x \cdot (y \cdot z)$

$x \cdot y = 0$ . then using (1.a) we have either

$x = 0$ . then  $(x \cdot y) \cdot z = 0 \cdot z = 0 \cdot (y \cdot z) = x \cdot (y \cdot z)$

$y = 0$ . then  $(x \cdot y) \cdot z = 0 \cdot z = x \cdot (0 \cdot z) = x \cdot (y \cdot z)$

$x \cdot y \in \mathbb{R}_+$ . then by (1.b)  $x, y \in \mathbb{R}$  so that using 17.20 we have

$x, y \in \mathbb{R}_+$ . for  $z$  we must then consider the following cases:

$z = \infty$ . then as by 17.20  $0 < x \cdot y$  we have  $(x \cdot y) \cdot z = \infty = x \cdot (y \cdot z)$

$z = -\infty$ . then as by 17.20  $0 < x \cdot y$  we have  $(x \cdot y) \cdot z = -\infty = x \cdot (y \cdot z)$

$z \in \mathbb{R}$ . then  $(x \cdot y) \cdot z = (x \cdot_{\mathbb{R}} y) \cdot_{\mathbb{R}} z = x \cdot_{\mathbb{R}} (y \cdot_{\mathbb{R}} z) = x \cdot (y \cdot z)$

$x, y \in \mathbb{R}_-$ . for  $z$  we must then consider the following cases:

$z = \infty$ . then as by 17.20  $0 < x \cdot y$  we have  $(x \cdot y) \cdot z = \infty = x \cdot (y \cdot z)$

$z = -\infty$ . then as by 17.20  $0 < x \cdot y$  we have  $(x \cdot y) \cdot z = -\infty = x \cdot (y \cdot z)$

$z \in \mathbb{R}$ . then  $(x \cdot y) \cdot z = (x \cdot_{\mathbb{R}} y) \cdot_{\mathbb{R}} z = x \cdot_{\mathbb{R}} (y \cdot_{\mathbb{R}} z) = x \cdot (y \cdot z)$

$x, y \in \mathbb{R}_-$ . then by (1.b)  $x, y \in \mathbb{R}$  so that using 17.20 we have

$x \in \mathbb{R}_+ \wedge y \in \mathbb{R}_-$ . for  $z$  we must then consider the following cases:

$z = \infty$ . then as by 17.20  $x \cdot y < 0$  we have  $(x \cdot y) \cdot z = -\infty = x \cdot (y \cdot z)$

$z = -\infty$ . then as by 17.20  $x \cdot y < 0$  we have  $(x \cdot y) \cdot z = \infty = x \cdot (y \cdot z)$

$z \in \mathbb{R}$ . then  $(x \cdot y) \cdot z = (x \cdot_{\mathbb{R}} y) \cdot_{\mathbb{R}} z = x \cdot_{\mathbb{R}} (y \cdot_{\mathbb{R}} z) = x \cdot (y \cdot z)$

$x \in \mathbb{R}_- \wedge y \in \mathbb{R}_+$ . for  $z$  we must then consider the following cases:

$z = \infty$ . then as by 17.20  $x \cdot y < 0$  we have  $(x \cdot y) \cdot z = -\infty = x \cdot (y \cdot z)$

$z = -\infty$ . then as by 17.20  $x \cdot y < 0$  we have  $(x \cdot y) \cdot z = \infty = x \cdot (y \cdot z)$

$z \in \mathbb{R}$ . then  $(x \cdot y) \cdot z = (x \cdot_{\mathbb{R}} y) \cdot_{\mathbb{R}} z = x \cdot_{\mathbb{R}} (y \cdot_{\mathbb{R}} z) = x \cdot (y \cdot z)$

4. For  $x$  we have either

$x \in \mathbb{R}$ . then  $1 \cdot x = 1 \cdot_{\mathbb{R}} x = x$

$x = \infty$ . then  $1 \cdot x = \infty = x$

$x = -\infty$ . then  $1 \cdot x = -\infty = x$

□

**Theorem 17.22.** If  $x, y \in \overline{\mathbb{R}}$  then we have

1. If  $-x = -y$  then  $x = y$

2. If  $x \leq y$  then  $-y \leq -x$

3. If  $x < y$  then  $-y < -x$

**Proof.**

1. For  $x$  we have to consider the following cases to consider

$x \in \mathbb{R}$ . then  $-x = (-1) \cdot_{\mathbb{R}} x \in \mathbb{R}$  so that  $-y \in \mathbb{R}$ . If  $y = \infty$  or  $y = -\infty$  then  $-y \notin \mathbb{R}$ , so we must have that  $y \in \mathbb{R}$ , but then we have that  $-y = (-1) \cdot_{\mathbb{R}} y$ . Hence we have  $(-1) \cdot_{\mathbb{R}} x = (-1) \cdot_{\mathbb{R}} y \Rightarrow x = y$

$x = \infty$ . then  $-x = -\infty$  so that  $-y = -\infty$ . If  $y \in \mathbb{R}$  then  $-y \in \mathbb{R}$  and if  $y = -\infty$  then  $-y = \infty$ , so we must have that  $y = \infty$  from which it follows that  $x = y$ .

$x = -\infty$ . then  $-x = \infty$  so that  $-y = \infty$ . If  $y \in \mathbb{R}$  then  $-y \in \mathbb{R}$  and if  $y = \infty$  then  $-y = -\infty$ , so we must have that  $y = -\infty$  from which it follows that  $x = y$ .

2. If  $x \leq y$  then we have to consider the following possibilities for  $x$

$x \in \mathbb{R}$ . then as  $x \leq y$  we have to consider the following possibilities for  $y$

$y \in \mathbb{R}$ . then using the properties of the reals we have  $(-y) \leq (-x)$

$y = \infty$ . then  $-y = -\infty < -x \in \mathbb{R} \Rightarrow (-y) \leq (-x)$

$x = \infty$ . then as  $x \leq y$  we have  $y = \infty$  so that  $-y = -\infty \leq -\infty = -x$  proving  $(-y) \leq (-x)$

$x = -\infty$ . then as  $x \leq y$  we have to consider the following possibilities

$y \in \mathbb{R}$ . then  $-x = \infty$  so as  $-y \in \mathbb{R}$   $-y \leq \infty = -x$  proving  $(-y) \leq (-x)$

$y = \infty$ . then  $-y = -\infty < \infty = -x$  proving  $(-y) \leq (-x)$

$y = -\infty$ . then  $-y = \infty = -x$  proving  $(-y) \leq (-x)$

3. If  $x < y$  then  $x \neq y$  and  $x < y$ , using (2) we have then that  $(-y) \leq -x$  and using (1) that  $-x \neq -y$  proving that  $-y < -x$   $\square$

**Corollary 17.23.** *If  $x, y \in \bar{\mathbb{R}}$  with  $x \leq y$  then for  $\lambda \in \bar{\mathbb{R}}$  we have*

1. *If  $0 \leq \lambda$  then  $\lambda \cdot x \leq \lambda \cdot y$*
2. *If  $\lambda \leq 0$  then  $\lambda \cdot y \leq \lambda \cdot x$*

**Proof.**

1. For  $0 \leq \lambda$  we have the following cases to consider for  $\lambda$

$\lambda = 0$ . then  $\lambda \cdot x = 0 = \lambda \cdot y \Rightarrow \lambda \cdot x \leq \lambda \cdot y$

$\lambda \in \mathbb{R}_+$ . then for  $x$  we have either

$x = -\infty$ . then  $\lambda \cdot x = -\infty \leq \lambda \cdot y$

$x \in \mathbb{R}$ . then as  $-\infty < x$  we must have that  $-\infty < y$ , so for  $y$  to consider either

$y = \infty$ . then  $\lambda \cdot y = \infty$  so that  $\lambda \cdot x \leq \lambda \cdot y$

$y \in \mathbb{R}$ . then as  $x \leq y$  we have using 9.41 that  $\lambda \cdot x \leq \lambda \cdot y$  proving  $\lambda \cdot x \leq \lambda \cdot y$ .

$x = \infty$ . then  $y \underset{x \leq y}{=} \infty$  hence  $\lambda \cdot x = \infty = \lambda \cdot y \Rightarrow \lambda \cdot x \leq \lambda \cdot y$

$\lambda = \infty$ . then for  $x$  we have either

$x = -\infty$ . then  $\lambda \cdot x = -\infty \leq \lambda \cdot y$

$x \in \mathbb{R}_-$ . then  $\lambda \cdot x = -\infty \leq \lambda \cdot y$

$x = 0$ . then as  $0 = x \leq y$  we have for  $y$  either

$y = 0$ . then  $\lambda \cdot x = 0 = \lambda \cdot y \Rightarrow \lambda \cdot x \leq \lambda \cdot y$

$y \in \mathbb{R}_+$ . then  $\lambda \cdot x = 0 \leq \infty = \lambda \cdot y \Rightarrow \lambda \cdot x \leq \lambda \cdot y$

$y = \infty$ . then  $\lambda \cdot x = 0 \leq \infty = \lambda \cdot y \Rightarrow \lambda \cdot x \leq \lambda \cdot y$

$x \in \mathbb{R}_+$ . then as  $0 < x \leq y$  we have either

$y \in \mathbb{R}_+$ . then  $\lambda \cdot x = \infty = \lambda \cdot y \Rightarrow \lambda \cdot x \leq \lambda \cdot y$

$y = \infty$ . then  $\lambda \cdot x = \infty = \lambda \cdot y \Rightarrow \lambda \cdot x \leq \lambda \cdot y$

$x = \infty$ . then  $y = \infty$  and  $\lambda \cdot x = \infty = \lambda \cdot y \Rightarrow \lambda \cdot x \leq \lambda \cdot y$

2. As  $\lambda \leq 0$  we have by 17.22 that  $0 \leq (-\lambda)$  then by (1)  $(-\lambda) \cdot x \leq (-\lambda) \cdot y$ , Using 17.22 we have that  $((-\lambda) \cdot y) \leq ((-\lambda) \cdot x)$ . As

$$((-\lambda) \cdot y) \underset{\text{def}}{=} ((-1) \cdot (((-1) \cdot \lambda) \cdot y)) \underset{17.21}{=} ((-1) \cdot (-1))(\lambda \cdot y) = \lambda \cdot y$$

and

$$((-\lambda) \cdot x) \underset{\text{def}}{=} ((-1) \cdot (((-1) \cdot \lambda) \cdot x)) \underset{17.21}{=} ((-1) \cdot (-1))(\lambda \cdot x) = \lambda \cdot x$$

it follows that

$$\lambda \cdot y < \lambda \cdot x$$

$\square$

We have to be very carefull if we do sums on elements of  $\bar{\mathbb{R}}$  for example if  $x = 10, y = 4$  and  $z = \infty$  then  $x + z = y + z = \infty$  however  $x \neq y$ , also  $x + z \leq y + z$  however  $x \not\leq y$ . The following theorems shows what is possible.

**Theorem 17.24.** *If  $x, y \in \bar{\mathbb{R}}$  and  $z \in \mathbb{R}$  then we have*

1. *if  $x + z = y + z$  we have  $x = y$*
2. *if  $x + z \leq y + z$  we have  $x \leq y$*
3. *if  $x + z < y + z$  we have  $x < y$*

4. if  $x + z = y$  we have  $x = y - z \stackrel{\text{def}}{=} y + (-z)$  [where  $y - z$  is well defined]
5. if  $x + z \leq y$  we have  $x \leq y - z \stackrel{\text{def}}{=} y + (-z)$  [where  $y - z$  is well defined]
6. if  $x + z < y$  we have  $x < y - z \stackrel{\text{def}}{=} y + (-z)$  [where  $y - z$  is well defined]
7. if  $x \leq y + z$  we have  $x + (-z) \leq y$
8. if  $x < y + z$  we have  $x + (-z) < y$

**Proof.**

1. Assume  $x + z = y + z$  and consider the following possibilities for  $x, y$ 
  - $x \in \mathbb{R} \wedge y \in \mathbb{R}$ . then  $x +_{\mathbb{R}} z = y +_{\mathbb{R}} z$  so that  $x = y$
  - $x \in \mathbb{R} \wedge y = \infty$ . then  $x + z \in \mathbb{R}$  and  $y + z = \infty$  giving the contradiction  $x + z \neq y + z$  so this case does not apply.
  - $x \in \mathbb{R} \wedge y = -\infty$ . then  $x + z \in \mathbb{R}$  and  $y + z = -\infty$  giving the contradiction  $x + z \neq y + z$  so this case does not apply.
  - $x = \infty \wedge y \in \mathbb{R}$ . then  $x + z = \infty$  and  $y + z \in \mathbb{R}$  giving the contradiction  $x + z \neq y + z$  so this case does not apply.
  - $x = \infty \wedge y = \infty$ . then  $x = y$ .
  - $x = \infty \wedge y = -\infty$ . then  $x + z = \infty \neq -\infty = y + z$  so this does not apply.
  - $x = -\infty \wedge y \in \mathbb{R}$ . then  $x + z = -\infty$  and  $y + z \in \mathbb{R}$  giving the contradiction  $x + z \neq y + z$  so this does not apply.
  - $x = -\infty \wedge y = \infty$ . then  $x + z = -\infty \neq y + z = \infty$  so this does not apply.
  - $x = -\infty \wedge y = -\infty$ . then  $x = y$

so in all cases that apply we have proved that  $x = y$

2. Assume  $x + y \leq y + z$  and consider the following possibilities for  $x, y$ 
  - $x \in \mathbb{R} \wedge y \in \mathbb{R}$ . then using the properties of the reals we have  $x \leq y$
  - $x \in \mathbb{R} \wedge y = \infty$ . then  $x \leq y$
  - $x \in \mathbb{R} \wedge y = -\infty$ . then  $x + z \not\leq -\infty = y + z$  so this does not apply.
  - $x = \infty \wedge y \in \mathbb{R}$ . then  $\infty = x + z \not\leq y + z$  so this does not apply.
  - $x = \infty \wedge y = \infty$ . then  $x \leq y$
  - $x = \infty \wedge y = -\infty$ . then  $x + z = \infty \not\leq -\infty = y + z$  so this does not apply.
  - $x = -\infty \wedge y \in \mathbb{R}$ . then  $x < y \Rightarrow x \leq y$
  - $x = -\infty \wedge y = \infty$ . then  $x < y \Rightarrow x \leq y$
  - $x = -\infty \wedge y = -\infty$ . then  $x = y \Rightarrow x \leq y$

so in all cases that apply we have  $x \leq y$

3. Assume that  $x + z < y + z$ 
  - $x \in \mathbb{R} \wedge y \in \mathbb{R}$ . then as  $x + z < y + z$  we have using the properties of the real numbers that  $x < y$
  - $x \in \mathbb{R} \wedge y = \infty$ . then  $x < y$
  - $x \in \mathbb{R} \wedge y = -\infty$ . then  $x + z \in \mathbb{R}$  and  $y + z = -\infty$  so that  $x + z \not\leq y + z$  so this does not apply.
  - $x = \infty \wedge y \in \mathbb{R}$ . then  $x + z = \infty$  and  $y + z \in \mathbb{R}$  giving the contradiction  $x + z \not\leq y + z$  so this does not apply.
  - $x = \infty \wedge y = \infty$ . then  $x + z = y + z \Rightarrow x + z \not\leq y + z$  so this does not apply.
  - $x = \infty \wedge y = -\infty$ . then  $\infty = x + z \not\leq y + z = -\infty$  so this does not apply.

$x = -\infty \wedge y \in \mathbb{R}$ . then  $x < y$

$x = -\infty \wedge y = \infty$ . then  $x < y$

$x = -\infty \wedge y = -\infty$ . then  $x + z - \infty \not< -\infty = y + z$  so this does not apply.

So in all cases that apply we have  $x < y$ .

4. If  $x + z = y$  we have for  $x$  the following possibilities

$x \in \mathbb{R}$ . then  $y \in \mathbb{R}$  and we have  $x = y +_{\mathbb{R}} (-z) = y + (-z)$

$x = \infty$ . then  $y = \infty$  and we have  $x = \infty = \infty + (-z) = y + (-z)$

$x = -\infty$ . then  $y = -\infty$  and we have  $x = -\infty = -\infty + (-z) = y + (-z)$

so in all cases we have  $x = y + (-z)$

5. If  $x + z \leq y$  we have for  $x$  the following possibilities

$x \in \mathbb{R}$ . then for  $y$  we have either

$y \in \mathbb{R}$ . then  $x \leq y +_{\mathbb{R}} (-z) = y + (-z)$

$y = \infty$ . then  $x \leq \infty = \infty + (-z) \Rightarrow x \leq y + (-z)$

$x = \infty$ . then  $x + z = \infty \Rightarrow y = \infty$  so that  $x = \infty = \infty + (-z) = y + (-z) \Rightarrow x \leq y + (-z)$

$x = -\infty$ . then  $x + (-z) = -\infty$  and we have the following possibilities for  $y$

$y \in \mathbb{R}$ . then  $x = -\infty < y \Rightarrow x \leq y + (-z)$

$y = \infty$ . then  $x = -\infty < \infty = y + (-z) \Rightarrow x \leq y + (-z)$

$y = -\infty$ . then  $x = -\infty = -\infty + (-z) = y + (-z) \Rightarrow x \leq y + (-z)$

6. If  $x + z < y$  we have for  $x$  the following possibilities

$x \in \mathbb{R}$ . then we have as  $x + z \not< -\infty$  for  $y$  either

$y \in \mathbb{R}$ . then by the properties of the reals we have  $x < y + (-z)$

$y = \infty$ . then  $x < \infty = \infty + (-z) = y + (-z) \Rightarrow x < y + (-z)$

$x = \infty$ . then  $x + z = \infty < y$  a contradiction so this does not apply

$x = -\infty$ . then we have as  $x + z \not< -\infty$  the following possibilities for  $y$

$y \in \mathbb{R}$ . then  $x = -\infty < y + (-z) \Rightarrow x < y + (-z)$

$y = \infty$ . then  $x = -\infty < \infty = \infty + (-z) = y + (-z) \Rightarrow x < y + (-z)$

so in all cases we have  $x < y + (-z)$

7. If  $x \leq y + z$  then for  $x$  we have the following possibilities for  $x$

$x \in \mathbb{R}$ . then for  $y$  we have the following possibilities

$y \in \mathbb{R}$ . then  $x + (-z) \leq y$

$y = \infty$ . then  $x + (-z) \leq \infty = y \Rightarrow x + (-z) \leq y$

$y = -\infty$ . then as  $x \in \mathbb{R}$  that  $x \not\leq -\infty = y + z$  a contradiction, so this case does not apply.

$x = \infty$ . then  $y + z = \infty$  so that  $y = \infty$  and  $x + (-z) = \infty = y \Rightarrow x + (-z) \leq y$

$x = -\infty$ . then  $x + (-z) = -\infty \leq y$  proving that  $x + (-z) \leq y$

so in all cases we have  $x + (-z) \leq y$

8. If  $x < y + z$  then we have the following possibilities for  $x$

$x \in \mathbb{R}$ . then we have the following possibilities for  $y$

$y \in \mathbb{R}$ . then by the properties of the real numbers we have  $x + (-z) < y$

$y = \infty$ . then as  $x + (-z) \in \mathbb{R}$  we have  $x + (-z) < y$

$y = -\infty$ . then as  $x \in \mathbb{R}$  we have  $x \not\leq -\infty = y + z$  so this case does not apply

$x = -\infty$ . then we have the following possibilities for  $y$

$y \in \mathbb{R}$ . then as  $x + (-z) = -\infty < y$

$y = \infty$ . then  $x + (-z) = -\infty < \infty = y$

$y = -\infty$ . then  $x = -\infty \not< -\infty = y + z$  so this case does not apply.

$x = \infty$ . then  $x = \infty \not< y + z$  so this case does not apply

so in all cases we have  $x + (-z) < y$

□

**Theorem 17.25.** *If  $x, y, z \in \bar{\mathbb{R}}$  such that  $x \leq y$  then  $x + z \leq y + z$  assuming that the sums are well defined. If  $x < y$  and  $z \in \mathbb{R}$  then  $x + z < y + z$ .*

**Proof.** We must consider the following possible cases (excluding the undefined sums and the cases where  $x \not\leq y$ )

$x \in \mathbb{R} \wedge y \in \mathbb{R} \wedge z \in \mathbb{R}$ . then using the properties of  $\mathbb{R}$  we have  $x + z \leq y + z$ .

$x \in \mathbb{R} \wedge y \in \mathbb{R} \wedge z = \infty$ . then  $x + y \leq \infty = y + z$ .

$x \in \mathbb{R} \wedge y \in \mathbb{R} \wedge z = -\infty$ . then  $x + z = -\infty = y + z \Rightarrow x + z \leq y + z$ .

$x \in \mathbb{R} \wedge y = \infty \wedge z \in \mathbb{R}$ . then  $x + z \leq \infty = y + z$ .

$x \in \mathbb{R} \wedge y = \infty \wedge z = \infty$ . then  $x + z = \infty = y + z \Rightarrow x + u \leq y + z$ .

$x \in \mathbb{R} \wedge y = \infty \wedge z = -\infty$ . as  $y + z$  is not well defined this does not apply.

$x \in \mathbb{R} \wedge y = -\infty \wedge z \in \mathbb{R}$ . as  $x \not\leq y$  this does not apply.

$x \in \mathbb{R} \wedge y = -\infty \wedge z = \infty$ . as  $x \not\leq y$  this does not apply.

$x \in \mathbb{R} \wedge y = -\infty \wedge z = -\infty$ . as  $x \not\leq y$  this does not apply.

$x = \infty \wedge y \in \mathbb{R} \wedge z \in \mathbb{R}$ . as  $x \not\leq y$  this does not apply.

$x = \infty \wedge y \in \mathbb{R} \wedge z = \infty$ . as  $x \not\leq y$  this does not apply.

$x = \infty \wedge y = \infty \wedge z \in \mathbb{R}$ . then  $x + z = \infty = y + z \Rightarrow x + z \leq y + z$ .

$x = \infty \wedge y = \infty \wedge z = \infty$ .  $x + z = \infty = y + z \Rightarrow x + z \leq y + z$ .

$x = \infty \wedge y = \infty \wedge z = -\infty$ . as  $y + z$  is not well defined this does not apply.

$x = \infty \wedge y = -\infty \wedge z \in \mathbb{R}$ . as  $x \not\leq y$  this does not apply.

$x = \infty \wedge y = -\infty \wedge z = \infty$ . as  $x \not\leq y$  this does not apply.

$x = \infty \wedge y = -\infty \wedge z = -\infty$ . as  $x \not\leq y$  this does not apply.

$x = -\infty \wedge y \in \mathbb{R} \wedge z \in \mathbb{R}$ . then  $x + z = -\infty \leq y + z$

$x = -\infty \wedge y \in \mathbb{R} \wedge z = \infty$ . then  $x + z = -\infty < \infty = y + z \Rightarrow x + z \leq y + z$ .

$x = -\infty \wedge y \in \mathbb{R} \wedge z = -\infty$ . then  $x + z = -\infty = y + z \Rightarrow x + z \leq y + z$

$x = -\infty \wedge y = \infty \wedge z \in \mathbb{R}$ . then  $x + z = -\infty < \infty = y + z \Rightarrow x + z \leq y + z$ .

$x = -\infty \wedge y = \infty \wedge z = \infty$ . as  $x + z$  is not well defined this does not apply.

$x = -\infty \wedge y = \infty \wedge z = -\infty$ . as  $y + z$  is not well defined this does not apply.

$x = -\infty \wedge y = -\infty \wedge z \in \mathbb{R}$ . then  $x + z = -\infty = y + z \Rightarrow x + z \leq y + z$ .

$x = -\infty \wedge y = -\infty \wedge z = \infty$ . as  $y + z$  is not well defined this does not apply.

$x = -\infty \wedge y = -\infty \wedge z = -\infty$ . then  $x + z = -\infty = y + z \Rightarrow x + z \leq y + z$ .

proving the first part of the theorem. For the second part if  $x < y \Rightarrow x \leq y$  first part  $\Rightarrow x + z \leq y + z$ .

Assume now that  $x + z = y + z$  then as  $z \in \mathbb{R}$  we have by the previous theorem [see 17.24] that  $x = y$  contradicting  $x < y$  so we must have that  $x + z < y + z$ . □

**Corollary 17.26.** Let  $x, y \in \bar{\mathbb{R}}$  with  $0 \leq x \wedge 0 < y$  then  $0 \leq x + y$

**Proof.** As  $0 < x$  we have by 17.25 that  $y \leq x + y$  which as  $0 \leq y$  proves  $0 \leq x + y$   $\square$

**Theorem 17.27.** Let  $x, y, z \in \bar{\mathbb{R}}$  with  $0 \leq y, z = x + y$  [where  $x + y$  is well defined] then we have  $x \leq z$

**Proof.** As  $x + y$  is well defined we have to consider the following cases (taking in account that  $0 \leq y$ )

$x, y \in \mathbb{R}$ . as  $0 \leq y$  we have  $x = x + 0 \leq x + y = z \Rightarrow x \leq z$

$x = \infty, y \in \mathbb{R}$ . as  $z = x + y = \infty = x$  we have  $x \leq z$

$x = -\infty, y \in \mathbb{R}$ . then  $x \leq z$

$x \in \mathbb{R}, y = \infty$ . then  $z = x + \infty = \infty$  so that  $x \leq z$

$x = \infty, y = \infty$ . then  $z = \infty + \infty = \infty$  so that  $x \leq z$   $\square$

Although  $\langle \bar{\mathbb{R}}, + \rangle$  is not a group we can prove that  $\langle \bar{\mathbb{R}}_+, + \rangle$  where  $\bar{\mathbb{R}}_+ = \{x \in \bar{\mathbb{R}} | x \geq 0\}$  is a abelian semi-group [see 3.1] which is important because then all the theorems about generalized sums and semi-groups are valid.

**Theorem 17.28.**  $\langle \bar{\mathbb{R}}_+, + \rangle$  where  $\bar{\mathbb{R}}_+ = \{x \in \bar{\mathbb{R}} | x \geq 0\}$  forms a abelian semi-group

**Proof.** First if  $x, y \in \bar{\mathbb{R}}_+$  then we have [as  $-\infty < 0$ ] either

$x = \infty \vee y = \infty$ . then  $x + y = \infty \in \bar{\mathbb{R}}_+$

$x, y \in \mathbb{R}_+$ . then as  $0 \leq x, y$  we have that  $0 \leq x \leq x + y \Rightarrow x + y \in \bar{\mathbb{R}}_+$

so that  $+: \bar{\mathbb{R}}_+ \times \bar{\mathbb{R}}_+ \rightarrow \bar{\mathbb{R}}_+$  defined by  $(x, y) \rightarrow x + y$  is indeed a function.

**Neutral Element.** First  $0 \in \bar{\mathbb{R}}_+$ , next let  $x \in \bar{\mathbb{R}}_+$  then we have [as  $-\infty < 0$ ] either

$x \in \mathbb{R}_+$ . then  $x + 0 = 0 + x = x$  [as  $\langle \mathbb{R}, + \rangle$  is a group with neutral element 0]

$x = \infty$ . then  $\infty + 0 = \infty = \infty + 0$

so  $0 \in \bar{\mathbb{R}}$  is a neutral element.

**Associativity.** Let  $x, y, z \in \bar{\mathbb{R}}$  then we have [as  $-\infty < 0$ ] either

$\exists z \in \{x, y, z\} \vdash z = \infty$ . then  $x + (y + z) = \infty = (x + y) + z$

$x, y, z \in \mathbb{R}_+$ . then  $x + (y + z) = (x + y) + z$  [as  $\langle \mathbb{R}, + \rangle$  is a group]

**Commutativity.** Let  $x, y \in \bar{\mathbb{R}}_+$  then as  $[-\infty < 0]$  we have either

$x = \infty \vee y = \infty$ . and then  $x + y = \infty = y + x$

$x, y \in \mathbb{R}_+$ . then  $x + y = y + x$  [as  $\langle \mathbb{R}, + \rangle$  is a Abelian group]  $\square$

We have also the distributive law in  $\bar{\mathbb{R}}$  as is proved in the following theorem, but we have to be carefull because for example  $\infty \cdot (1 - 3) = \infty \cdot (-2) = -\infty$  but  $\infty \cdot 1 + \infty \cdot (-3) = \infty + (-\infty)$  which is not defined. So we must restrict the distributive law.

**Theorem 17.29.** Let  $\alpha \in \mathbb{R}$ ,  $x, y \in \bar{\mathbb{R}}$  then if  $x + y$  is defined we have  $\alpha \cdot (x + y) = \alpha \cdot x + \alpha \cdot y$

**Proof.** As  $x + y$  is well defined we have to consider only the following cases

$x \in \mathbb{R} \wedge y \in \mathbb{R}$ . then  $\alpha \cdot (x + y) = \alpha \cdot_{\mathbb{R}} (x +_{\mathbb{R}} y) = \alpha \cdot_{\mathbb{R}} x +_{\mathbb{R}} \alpha \cdot_{\mathbb{R}} y = \alpha \cdot x + \alpha \cdot y$

$x \in \mathbb{R} \wedge y = \infty$ . then for  $\alpha$  we have either

$\alpha = 0$ . then  $\alpha \cdot (x + y) = 0 = 0 + 0 = \alpha \cdot x + \alpha \cdot y$ .

$0 < \alpha$ . then  $\alpha \cdot (x + y) = \alpha \cdot \infty = \infty = \alpha \cdot x + \infty = \alpha \cdot x + \alpha \cdot y$

$\alpha < 0$ . then  $\alpha \cdot (x + y) = \alpha \cdot \infty = -\infty = \alpha \cdot x + (-\infty) = \alpha \cdot x + \alpha \cdot y$

$x \in \mathbb{R} \wedge y = -\infty$ . then for  $\alpha$  we have either

$\alpha = 0$ . then  $\alpha \cdot (x + y) = 0 = 0 + 0 = \alpha \cdot x + \alpha \cdot y$ .

**0 <  $\alpha$ .** then  $\alpha \cdot (x + y) = \alpha \cdot (-\infty) = -\infty = \alpha \cdot x + (-\infty) = \alpha \cdot x + \alpha \cdot y$

**$\alpha < 0$ .** then  $\alpha \cdot (x + y) = \alpha \cdot (-\infty) = \infty = \alpha \cdot x + \infty = \alpha \cdot x + \alpha \cdot y$

$x = \infty \wedge y \in \mathbb{R}$ . then for  $\alpha$  we have either

**$\alpha = 0$ .** then  $\alpha \cdot (x + y) = 0 = 0 + 0 = \alpha \cdot x + \alpha \cdot y$ .

**0 <  $\alpha$ .** then  $\alpha \cdot (x + y) = \alpha \cdot \infty = \infty = \infty + \alpha \cdot y = \alpha \cdot x + \alpha \cdot y$

**$\alpha < 0$ .** then  $\alpha \cdot (x + y) = \alpha \cdot \infty = -\infty = -\infty + \alpha \cdot y = \alpha \cdot x + \alpha \cdot y$

$x = \infty \wedge y = \infty$ . then for  $\alpha$  we have either

**$\alpha = 0$ .** then  $\alpha \cdot (x + y) = 0 = 0 + 0 = \alpha \cdot x + \alpha \cdot y$ .

**0 <  $\alpha$ .** then  $\alpha \cdot (x + y) = \alpha \cdot \infty = \infty = \infty + \infty = \alpha \cdot x + \alpha \cdot y$

**$\alpha < 0$ .** then  $\alpha \cdot (x + y) = \alpha \cdot \infty = -\infty = -\infty + (-\infty) = \alpha \cdot x + \alpha \cdot y$

$x = -\infty \wedge y \in \mathbb{R}$ . then for  $\alpha$  we have either

**$\alpha = 0$ .** then  $\alpha \cdot (x + y) = 0 = 0 + 0 = \alpha \cdot x + \alpha \cdot y$ .

**0 <  $\alpha$ .** then  $\alpha \cdot (x + y) = \alpha \cdot (-\infty) = -\infty = -\infty + \alpha \cdot y = \alpha \cdot x + \alpha \cdot y$

**$\alpha < 0$ .** then  $\alpha \cdot (x + y) = \alpha \cdot (-\infty) = \infty = \infty + \alpha \cdot y = \alpha \cdot x + \alpha \cdot y$

$x = -\infty \wedge y = -\infty$ . then for  $\alpha$  we have either

**$\alpha = 0$ .** then  $\alpha \cdot (x + y) = 0 = 0 + 0 = \alpha \cdot x + \alpha \cdot y$ .

**0 <  $\alpha$ .** then  $\alpha \cdot (x + y) = \alpha \cdot (-\infty) = -\infty = -\infty + (-\infty) = \alpha \cdot x + \alpha \cdot y$

**$\alpha < 0$ .** then  $\alpha \cdot (x + y) = \alpha \cdot (-\infty) = \infty = -\infty + \infty = \alpha \cdot x + \alpha \cdot y$

So in all cases we have  $\alpha \cdot (x + y) = \alpha \cdot x + \alpha \cdot y$  □

The next two theorems show how the supremum and infimum behaves when a sum or product is involved.

**Theorem 17.30.** Let  $\emptyset \neq A \subseteq \bar{\mathbb{R}}$  then we have

1. If  $y \in \mathbb{R}$  then

a.  $\sup(A) + y$  is well defined and  $\sup(\{x + y \mid x \in A\}) = \sup(A) + y$

b.  $\inf(A) + y$  is well defined and  $\inf(\{x + y \mid x \in A\}) = \inf(A) + y$

2. If  $y \in [0, \infty]$  and  $A \subseteq [0, \infty]$  then  $\sup(A) + y$  is well defined and  $\sup(A + y) = \sup(A) + y$

3. If  $y \in [-\infty, 0]$  and  $A \subseteq [-\infty, 0]$  then  $\inf(A) + y$  is well defined and  $\inf(A + y) = \inf(A) + y$

**Proof.**

1. As  $y \in \mathbb{R}$  we have that  $\forall x \in \bar{\mathbb{R}}$   $x + y$  is well defined and  $\sup(A) + y$ ,  $\inf(A) + y$  is well defined.

a.  $\forall x \in A$  we have  $x \leq \sup(A)$  so that by 17.25 we have  $x + y \leq \sup(A) + y$  so that

$$\sup(\{x + y \mid x \in A\}) \leq \sup(A) + y \quad (17.3)$$

Assume now that  $\sup(\{x + y \mid y \in A\}) < \sup(A) + y$   $\underset{y \in \mathbb{R} \text{ and 17.25}}{\Rightarrow} \sup(\{x + y \mid y \in A\}) - y < \sup(A)$  so by the definition of the supremum there exists a  $z \in A$  such that  $\sup(\{x + y \mid x \in A\}) - y < z \leq \sup(A)$   $\underset{y \in \mathbb{R} \text{ and 17.24}}{\Rightarrow} \sup(\{x + y \mid x \in A\}) < z + y$  a contradiction [as  $z + y \in \{x + y \mid y \in A\}$ ]. So we must have that  $\sup(A) + y \leq \sup(\{x + y \mid x \in A\})$  which together with 17.3 gives

$$\sup(\{x + y \mid x \in A\}) = \sup(A) + y$$

b.  $\forall x \in A$  we have  $\inf(A) \leq x$  so that by 17.25 we have  $\inf(A) + y \leq x + y$  proving that

$$\inf(A) + y \leq \inf(\{x + y \mid x \in A\}) \quad (17.4)$$

Assume now that  $\inf(A) + y < \inf(\{x + y \mid x \in A\})$  then using 17.24 we have  $\inf(A) < \inf(\{x + y \mid x \in A\}) - y$  so by the definition of the infimum there exists a  $z \in A$  such that  $\inf(A) \leq z < \inf(\{x + y \mid x \in A\}) - y \underset{y \in \mathbb{R} \text{ and 17.24}}{\Rightarrow} z + y < \inf(\{x + y \mid x \in A\})$  a contradiction [as  $z + y \in \{x + y \mid x \in A\}$ ]. So we must have that  $\inf(\{x + y \mid x \in A\}) \leq \inf(A) + y$  proving that

$$\inf(A) + y = \inf(\{x + y \mid x \in A\})$$

2. For  $y \in [0, \infty]$  we have either

**$y \in \mathbb{R}$ .** then by (1.a) we have that  $\sup(A) + y$  is well defined and  $\sup(\{x + y \mid x \in A\}) = \sup(A) + y$

**$y = \infty$ .** then as  $A \subseteq [0, \infty]$  we have  $\forall a \in A$  that  $a + y = \infty$  so that  $\{a + y \mid a \in A\} = \{\infty\}$ , hence  $\sup(A + y) = \infty$ , further  $0 \leq \sup(A) \Rightarrow \sup(A) \neq -\infty \Rightarrow \sup(A) + y$  is well defined and  $\sup(A) + y = \infty$  which proves that

$$\sup(A + y) = \infty = \sup(A) + y$$

3. For  $y \in [-\infty, 0]$  we have either

**$y \in \mathbb{R}$ .** then by 1.b we have that  $\inf(A) + y$  is well defined and  $\inf(\{x + y \mid x \in A\}) = \inf(A) + y$

**$y = -\infty$ .** then as  $A \subseteq [-\infty, 0]$  we have  $\forall a \in A$   $a + y = -\infty$  so that  $\{a + y \mid a \in A\} = \{-\infty\}$ , hence  $\inf(A + y) = -\infty$ , further  $\inf(A) \leq 0 \Rightarrow \inf(A) + y$  is well defined and  $\inf(A) + y = -\infty$  which proves that

$$\inf(A + y) = -\infty = \inf(A) + y$$

□

**Theorem 17.31.** Let  $\emptyset \neq A \subseteq \bar{\mathbb{R}}$  be a non empty set then

1.  $-\sup(A) = \inf(\{-x \mid x \in A\})$
2.  $-\inf(A) = \sup(\{-x \mid x \in A\})$
3. If  $\alpha \in [0, \infty[$  then  $\alpha \cdot \sup(A) = \sup(\{\alpha \cdot x \mid x \in A\})$
4. If  $\alpha \in [0, \infty[$  then  $\alpha \cdot \inf(A) = \inf(\{\alpha \cdot x \mid x \in A\})$

**Proof.**

1. If  $x \in A$  then  $x \leq \sup(A) \underset{17.22}{\Rightarrow} -x \leq -\sup(A)$  so that

$$-\sup(A) \leq \inf(\{-x \mid x \in A\}) \tag{17.5}$$

Assume now that  $-\sup(A) < \inf(\{-x \mid x \in A\})$  then using 17.22 we have  $-\inf(\{-x \mid x \in A\}) < \sup(A)$  so by the definition of the supremum there exists a  $z \in A$  such that  $-\inf(\{-x \mid x \in A\}) < z \leq \sup(A) \underset{17.22}{\Rightarrow} -z < \inf(\{-x \mid x \in A\})$  which as  $-z \in \{-x \mid x \in A\}$  is a contradiction.

So we must have  $\inf(\{-x \mid x \in A\}) \leq -\sup(A)$  which together with 17.5 proves that

$$-\sup(A) = \inf(\{-x \mid x \in A\})$$

2. Take  $B = \{-x \mid x \in A\}$  then  $A = \{-x \mid x \in B\}$  and  $-\inf(A) = -\inf(\{-x \mid x \in B\}) \underset{(1)}{=} -(-\sup(B)) = \sup(\{-x \mid x \in A\})$  proving

$$-\inf(A) = \sup(\{-x \mid x \in A\})$$

3. We must consider two cases for  $\alpha$

**$\alpha = 0$ .** In this case  $\{\alpha \cdot x \mid x \in A\} = \{0\}$  and thus  $\sup(\{\alpha \cdot x \mid x \in A\}) = \sup(\{0\}) = 0 = 0 \cdot \sup(A) = \alpha \cdot \sup(A)$

**$\alpha \neq 0$ .**  $\forall x \in A$  we have  $x \leq \sup(A) \underset{\alpha \geq 0}{\Rightarrow} \alpha \cdot x \leq \alpha \cdot \sup(A)$  giving

$$\sup(\{\alpha \cdot x \mid x \in A\}) \leq \alpha \cdot \sup(A) \tag{17.6}$$

Assume now that  $\sup(\{\alpha \cdot x \mid x \in A\}) < \alpha \cdot \sup(A) \Rightarrow \frac{1}{\alpha} \cdot \sup(\{\alpha \cdot x \mid x \in A\}) < \sup(A)$  hence by the definition of the supremum there exists a  $z \in A$  such that  $\frac{1}{\alpha} \cdot \sup(\{\alpha \cdot x \mid x \in A\}) < z \leq \sup(A) \Rightarrow \sup(\{\alpha \cdot x \mid x \in A\}) < \alpha \cdot z$  a contradiction [as  $\alpha \cdot z \in \{\alpha \cdot x \mid x \in A\}$ ]. So we must have that  $\alpha \cdot \sup(A) \leq \sup(\{\alpha \cdot x \mid x \in A\})$  which together with 17.6 gives

$$\sup(\{\alpha \cdot x \mid x \in A\}) = \alpha \cdot \sup(A)$$

4. We must consider two cases for  $\alpha$

**$\alpha = 0$ .** In this case  $\{\alpha \cdot x \mid x \in A\} = \{0\}$  and thus  $\inf(\{\alpha \cdot x \mid x \in A\}) = \inf(\{0\}) = 0 = 0 \cdot \inf(A) = \alpha \cdot \inf(A)$

**$\alpha \neq 0$ .**  $\forall x \in A$  we have  $\inf(A) \leq x \underset{\alpha \neq 0}{\Rightarrow} \alpha \cdot \inf(A) \leq \alpha \cdot x$  giving

$$\alpha \cdot \inf(A) \leq \inf(\{\alpha \cdot x \mid x \in A\}) \quad (17.7)$$

Assume now that  $\alpha \cdot \inf(A) < \inf(\{\alpha \cdot x \mid x \in A\}) < \inf(A) < \frac{1}{\alpha} \cdot \inf(\{\alpha \cdot x \mid x \in A\})$  hence by the definition of the infimum there exists a  $z \in A$  such that  $\inf(A) \leq z < \frac{1}{\alpha} \cdot \inf(\{\alpha \cdot x \mid x \in A\}) \Rightarrow \alpha \cdot z < \inf(\{\alpha \cdot x \mid x \in A\})$  a contradiction as  $[\alpha \cdot z \in \{\alpha \cdot x \mid x \in A\}]$ . So we must have that  $\inf(\{\alpha \cdot x \mid x \in A\}) \leq \inf(A)$  which together with 17.7 gives

$$\inf(\{\alpha \cdot x \mid x \in A\}) = \alpha \cdot \inf(A)$$

□

## 17.2 Topology on $\bar{\mathbb{R}}$

**Definition 17.32.** Let  $x \in \bar{\mathbb{R}}$  then  $|x| = \begin{cases} x & \text{if } 0 \leq x \\ -x & \text{if } x < 0 \end{cases}$ , || is called the absolute value on the extended reals.

**Note 17.33.** From the definition it is clear that  $x \leq |x|$

The absolute value on the extended reals has similar properties as the absolute value on the real numbers. However we must take in account that the sum is not always defined.

**Theorem 17.34.** The absolute value has the following properties

1.  $\forall x \in \bar{\mathbb{R}}$  we have that  $|x| = |x|_r$  (where we use the special notation  $||_r$  to indicate the absolute value on the real numbers (see 8.65))
2.  $|x| = 0 \Leftrightarrow x = 0$
3.  $\forall x, y \in \bar{\mathbb{R}}$  we have  $|x \cdot y| = |x| \cdot |y|$
4.  $\forall x, y \in \bar{\mathbb{R}}$  such that  $x + y$  is well defined we have that  $|x + y| \leq |x| + |y|$

**Proof.**

1. If  $x \in \mathbb{R}$  then  $|x| = \begin{cases} x & \text{if } x \geq 0 \\ -x & \text{if } x < 0 \end{cases} \underset{8.65}{=} |x|_r$
2. First  $|0| \underset{0 \leq 0}{=} 0$  by definition, second if  $|x| = 0$  then if  $x \neq 0$  we have either  $x < 0 \Rightarrow 0 = |x| = -x \neq 0$  a contradiction or  $0 < x \Rightarrow 0 = |x| = x \neq 0$  again a contradiction so we must have  $x = 0$
3. We must consider the following cases for  $x, y$

**$0 \leq x \wedge 0 \leq y$ .** then by 17.23 we have that  $0 \leq x \cdot y$  so that  $|x \cdot y| = x \cdot y = |x| \cdot |y|$

**$0 \leq x \wedge y \leq 0$ .** then by 17.23 we have that  $x \cdot y \leq 0$  so that  $|x \cdot y| = -(x \cdot y) = x \cdot (-y) = |x| \cdot |y|$

**$x \leq 0 \wedge 0 \leq y$ .** then by 17.23 we have that  $x \cdot y \leq 0$  so that  $|x \cdot y| = -(x \cdot y) = (-x) \cdot y = |x| \cdot |y|$

$x \leq 0 \wedge y \leq 0$ . then by 17.23 we have that  $0 \leq x \cdot y$  so that  $|x \cdot y| = x \cdot y = (-|x|) \cdot (-|y|) = (-1) \cdot (-1) \cdot |x| \cdot |y| = |x| \cdot |y|$

4. We have to consider the following (valid) cases for  $x, y$

$x \in \mathbb{R} \wedge y \in \mathbb{R}$ . then  $|x + y| = |x + y|_r \leq |x_r| + |y|_r = |x| + |y|$

$x \in \mathbb{R} \wedge y = \infty$ . then  $|x + y| = |\infty| = \infty \leq \infty = |x| + \infty = |x| + |\infty|$

$x \in \mathbb{R} \wedge y = -\infty$ . then  $|x + y| = |-\infty| = \infty \leq \infty = |x| + \infty = |x| + |y|$

$x = \infty \wedge y \in \mathbb{R}$ . then  $|x + y| = |\infty| = \infty \leq \infty = \infty + |y| = |x| + |y|$

$x = \infty \wedge y = \infty$ . then  $|x + y| = |\infty| = \infty \leq \infty = \infty + \infty = |x| + |y|$

$x = -\infty \wedge y \in \mathbb{R}$ . then  $|x + y| = |-\infty| = \infty \leq \infty = \infty + |y| = |x| + |y|$

$x = -\infty \wedge y = -\infty$ . then  $|x + y| = |-\infty| = \infty \leq \infty = |x| + |y|$   $\square$

To define the topology on  $\overline{\mathbb{R}}$  we first define some sets of sets that we use to create a generating basis for the topology (see 11.28).

**Notation 17.35.** In what follows  $\varepsilon > 0$  means  $\varepsilon \in \mathbb{R} \wedge \varepsilon > 0$

**Definition 17.36.** We define the following sets

1.  $\mathcal{B}_{\mathbb{R}} = \{[a, b] | a, b \in \mathbb{R} \wedge a < b\}$  [note this is by 11.72 the generating base of  $\mathcal{T}_{\mathbb{R}}$ ]
2.  $\mathcal{B}_{\infty} = \{[x, \infty] | x \in \mathbb{R}\}$
3.  $\mathcal{B}_{-\infty} = \{[-\infty, x] | x \in \mathbb{R}\}$
4.  $\mathcal{B}_{\overline{\mathbb{R}}} = \mathcal{B}_{\mathbb{R}} \cup \mathcal{B}_{\infty} \cup \mathcal{B}_{-\infty} \subseteq \mathcal{P}(\overline{\mathbb{R}})$

To define the topology  $\overline{\mathbb{R}}$  we first prove that  $\mathcal{B}_{\overline{\mathbb{R}}}$  is a generating basis on  $\overline{\mathbb{R}}$  [see 11.28]

**Lemma 17.37.**  $\mathcal{B}_{\overline{\mathbb{R}}}$  is a generating basis for  $\overline{\mathbb{R}}$  [see 11.28]

**Proof.**

1. If  $x \in \overline{\mathbb{R}}$  then we have either:

$x \in \mathbb{R}$ . then  $x \in ]x - 1, x + 1[ \in \mathcal{B}_{\overline{\mathbb{R}}}$

$x = \infty$ . then  $x \in ]0, \infty] \in \mathcal{B}_{\overline{\mathbb{R}}}$

$x = -\infty$ . then  $x \in [-\infty, 0[ \in \mathcal{B}_{\overline{\mathbb{R}}}$

proving that

$$\forall x \in \overline{\mathbb{R}} \text{ there exists a } B \in \mathcal{B}_{\overline{\mathbb{R}}} \text{ such that } x \in B \quad (17.8)$$

2. If  $A, B \in \mathcal{B}_{\overline{\mathbb{R}}}$  and  $x \in A \cap B$

$A, B \in \mathcal{B}_{\mathbb{R}}$ . then  $\exists a, b, c, d \in \mathbb{R}$  with  $A = [a, b], B = [c, d]$ . As  $x \in A \cap B$  we have that  $a < x < b \wedge c < x < d$ . So if we take  $e = \max(a, c), f = \min(b, d)$  then  $a, c \leq e < x < f \leq b, d$  or if we define  $C = [e, f] \in \mathcal{B}_{\overline{\mathbb{R}}}$  that  $x \in C \subseteq A \cap B$

$A \in \mathcal{B}_{\mathbb{R}} \wedge B \in \mathcal{B}_{\infty}$ . then  $\exists a, b, d \in \mathbb{R}$  such that  $A = [a, b], B = [d, \infty]$ . As  $x \in A \cap B$  we have that  $a < x < b \wedge d < x$ . So if we take  $e = \max(a, d)$  then  $a, d \leq e < x < b$ , or if we define  $C = [e, b] \in \mathcal{B}_{\overline{\mathbb{R}}}$  that  $x \in C \subseteq A \cap B$

$A \in \mathcal{B}_{\mathbb{R}} \wedge B \in \mathcal{B}_{-\infty}$ . then  $\exists a, b, d \in \mathbb{R}$  such that  $A = [a, b], B = [-\infty, d]$ . As  $x \in A \cap B$  we have that  $a < x < b \wedge x < d$ . So if we take  $e = \min(b, d)$  then  $a < x < e \leq b, d$ , or if we define  $C = [a, e] \in \mathcal{B}_{\overline{\mathbb{R}}}$  that  $x \in C \subseteq A \cap B$

$A \in \mathcal{B}_{\infty} \wedge B \in \mathcal{B}_{\mathbb{R}}$ . interchange  $A$  and  $B$  then we have the case  $A \in \mathcal{B}_{\mathbb{R}} \wedge B \in \mathcal{B}_{\infty}$  so there exists a  $C \in \mathcal{B}_{\overline{\mathbb{R}}}$  such that  $x \in C \subseteq B \cap A = A \cap B$

$A \in \mathcal{B}_{\infty} \wedge B \in \mathcal{B}_{\infty}$ . then  $\exists a, b \in \mathbb{R}$  such that  $A = [a, \infty], B = [b, \infty]$ . As  $x \in A \cap B$  we have that  $a < x \wedge b < x$ . So if we take  $c = \max(a, b)$  then  $a, b \leq c < x$  or if we define  $C = [c, \infty] \in \mathcal{B}_{\overline{\mathbb{R}}}$  that  $x \in C \subseteq A \cap B$

$A \in \mathcal{B}_\infty \wedge B \in \mathcal{B}_{-\infty}$ . then  $\exists a, b \in \mathbb{R}$  with  $A = ]a, \infty]$ ,  $B = [-\infty, b[$ . As  $x \in A \cap B$  we have that  $a < x \wedge x < b$ . So if we define  $C = ]a, b[ \in \mathcal{B}_{\bar{\mathbb{R}}}$  that  $x \in C \subseteq A \cap B$

$A \in \mathcal{B}_{-\infty} \wedge B \in \mathcal{B}_{\mathbb{R}}$ . interchange  $A$  and  $B$  then we have the case  $A \in \mathcal{B}_{\mathbb{R}} \wedge B \in \mathcal{B}_{-\infty}$  so there exists a  $C \in \mathcal{B}_{\bar{\mathbb{R}}}$  such that  $x \in C \subseteq B \cap A = A \cap B$

$A \in \mathcal{B}_{-\infty} \wedge B \in \mathcal{B}_\infty$ . interchange  $A$  and  $B$  then we have the case  $A \in \mathcal{B}_\infty \wedge B \in \mathcal{B}_{-\infty}$  so there exists a  $C \in \mathcal{B}_{\bar{\mathbb{R}}}$  such that  $x \in C \subseteq B \cap A = A \cap B$

$A \in \mathcal{B}_{-\infty} \wedge B \in \mathcal{B}_{-\infty}$ . then  $\exists a, b \in \mathbb{R}$  such that  $A = [-\infty, a[$ ,  $B = [-\infty, b[$ . As  $x \in A \cap B$  we have that  $x < a \wedge x < b$ . Take  $c = \min(a, b)$  then  $x < c \leq a, b$  or if we define  $C = [-\infty, c[ \in \mathcal{B}_{\bar{\mathbb{R}}}$  that  $x \in C \subseteq A \cap B$ .

this proves that

$$\forall A, B \in \mathcal{B}_{\bar{\mathbb{R}}} \text{ with } x \in A \cap B \text{ there exists } aC \in \mathcal{B}_{\bar{\mathbb{R}}} \text{ such that } x \in C \subseteq A \cap B \quad (17.9)$$

Finally using the definition of a generating basis [see 11.28] together with 17.8 and 17.9 proves the lemma.  $\square$

We can use now the above generating basis on  $\bar{\mathbb{R}}$  to define the topology on  $\bar{\mathbb{R}}$

**Definition 17.38.**  $\mathcal{T}_{\bar{\mathbb{R}}} = \{U \subseteq \bar{\mathbb{R}} \mid \forall x \in U \models \exists B \in \mathcal{B}_{\bar{\mathbb{R}}} \text{ such that } x \in B \subseteq U\}$  is a topology called the canonical topology on  $\bar{\mathbb{R}}$

The canonical topology on  $\bar{\mathbb{R}}$  has the following properties.

**Proposition 17.39.** Let  $\mathcal{T}_{\bar{\mathbb{R}}}$  be the canonical topology on  $\bar{\mathbb{R}}$  and  $\mathcal{T}_{\mathbb{R}}$  the canonical topology on  $\mathbb{R}$  [see 11.72] then  $\mathcal{T}_{\mathbb{R}} \subseteq \mathcal{T}_{\bar{\mathbb{R}}}$  and  $\mathcal{T}_{\mathbb{R}} = \{U \cap \mathbb{R} \mid U \in \mathcal{T}_{\bar{\mathbb{R}}}\}$ .

**Proof.** By definition we have  $\mathcal{T}_{\bar{\mathbb{R}}} = \{U \subseteq \bar{\mathbb{R}} \mid \forall x \in U \models \exists B \in \mathcal{B}_{\bar{\mathbb{R}}} \text{ with } x \in B \subset -U\}$  and  $\mathcal{T}_{\mathbb{R}} = \{U \subseteq \mathbb{R} \mid \forall x \in U \models \exists B \in \mathcal{B}_{\mathbb{R}}\}$ . If  $U \in \mathcal{T}_{\mathbb{R}}$  then  $\forall x \in U \exists B_x \in \mathcal{B}_{\mathbb{R}}$  such that  $x \in B_x \subseteq U$ , so  $U = \bigcup_{x \in U} B_x$ . As  $B_x \in \mathcal{B}_{\mathbb{R}} \subseteq \mathcal{B}_{\bar{\mathbb{R}}} \subseteq \mathcal{T}_{\bar{\mathbb{R}}}$  it follows that  $U \in \mathcal{T}_{\bar{\mathbb{R}}}$ . So

$$\mathcal{T}_{\mathbb{R}} \subseteq \mathcal{T}_{\bar{\mathbb{R}}} \quad (17.10)$$

Now if  $U \in \mathcal{T}_{\mathbb{R}}$  then  $U \subseteq \mathbb{R} \Rightarrow U = U \cap \mathbb{R}$  and as  $U \in \mathcal{T}_{\bar{\mathbb{R}}}$  we have that  $U \in \{U \cap \mathbb{R} \mid U \in \mathcal{T}_{\bar{\mathbb{R}}}\}$  proving that

$$\mathcal{T}_{\mathbb{R}} \subseteq \{U \cap \mathbb{R} \mid U \in \mathcal{T}_{\bar{\mathbb{R}}}\} \quad (17.11)$$

If  $U \in \{U \cap \mathbb{R} \mid U \in \mathcal{T}_{\bar{\mathbb{R}}}\}$  then there exists a  $V \in \mathcal{T}_{\bar{\mathbb{R}}}$  such that  $U = V \cap \mathbb{R}$ . If  $x \in U$  then  $x \in V, x \in \mathbb{R}$  and there exists a  $B_x \in \mathcal{B}_{\bar{\mathbb{R}}}$  such that  $x \in B_x \subseteq V$ . We have now the following cases to consider for  $B_x$ :

$B_x \in \mathcal{B}_{\mathbb{R}}$ . then if we define  $B'_x = B_x$  we have  $x \in B'_x \subseteq V \subseteq U$  where  $B'_x \in \mathcal{T}_{\mathbb{R}}$

$B_x \in \mathcal{B}_\infty$ . then  $\exists a \in \mathbb{R}$  such that  $x \in B_x = ]a, \infty]$ . Take  $B'_x = ]a, x+1[ \in \mathcal{T}_{\mathbb{R}}$  then  $x \in B'_x \subseteq B_x \subseteq V \subseteq U$

$B_x \in \mathcal{B}_{-\infty}$ . then  $\exists a \in \mathbb{R}$  such that  $x \in B_x = [-\infty, a[$ . Take  $B'_x = ]x-1, a[ \in \mathcal{T}_{\mathbb{R}}$  then  $x \in B'_x \subseteq B_x \subseteq V \subseteq U$

so in all cases there exists a  $B'_x \in \mathcal{T}_{\mathbb{R}}$  such that  $x \in B'_x \subseteq U$  which as  $x \in U$  was chosen arbitrary proves that  $U \in \mathcal{T}_{\mathbb{R}}$ . So we have that  $\{U \cap \mathbb{R} \mid U \in \mathcal{T}_{\bar{\mathbb{R}}}\} \subseteq \mathcal{T}_{\mathbb{R}}$ , combining this with 17.11 proves that

$$\mathcal{T}_{\mathbb{R}} = \{U \cap \mathbb{R} \mid U \in \mathcal{T}_{\bar{\mathbb{R}}}\} \quad \square$$

**Theorem 17.40.** We have the following closed sets in  $\mathcal{T}_{\bar{\mathbb{R}}}$  where  $x \in \mathbb{R}$

1.  $[-\infty, x]$
2.  $[x, \infty]$
3.  $\{x\}$
4.  $[-\infty, \infty]$

5.  $\{-\infty, \infty\}$
6.  $\{-\infty\}$
7.  $\{\infty\}$

**Proof.**

1. As  $y \in \bar{\mathbb{R}} \setminus [-\infty, x] \Leftrightarrow y \in \neg(-\infty \leq y \wedge y \leq x) \Leftrightarrow (y < -\infty \vee x < y) \Leftrightarrow x < y \Leftrightarrow y \in ]x, \infty]$  we have that  $\bar{\mathbb{R}} \setminus [-\infty, x] = ]x, \infty]$  by definition a open set, hence  $[-\infty, x]$  is closed.
2. As  $y \in \bar{\mathbb{R}} \setminus [x, \infty] \Leftrightarrow y \in \neg(x \leq y \wedge y \leq \infty) \Leftrightarrow (y < x \vee \infty < y) \Leftrightarrow y < x \Leftrightarrow y \in [-\infty, x[$  we have that  $\bar{\mathbb{R}} \setminus [x, \infty] = [-\infty, x[$  by definiton a open set, hence  $[x, \infty]$  is a closed set.
3.  $\{x\} = [-\infty, x] \cap [x, \infty]$  by (1) and (2) a intersection of closed sets and thus closed.
4. This is trivial as  $\bar{\mathbb{R}} = [-\infty, \infty]$ ,
5. As  $\bar{\mathbb{R}} \setminus \{-\infty, \infty\} = \mathbb{R} \in \mathcal{T}_{\mathbb{R}} \subseteq \mathcal{T}_{\bar{\mathbb{R}}}$  we have that  $\bar{\mathbb{R}} \setminus \{-\infty, \infty\}$  is open so that  $\{-\infty, \infty\}$  is closed.
6. This follows from  $\{-\infty\} = \{-\infty, \infty\} \cap [-\infty, 0]$  a intersection of closed sets.
7. This follows from  $\{\infty\} = \{-\infty, \infty\} \cap [0, \infty]$  a intersection of closed sets.  $\square$

We have the following characterization of open sets in  $\bar{\mathbb{R}}$ .

**Theorem 17.41.** *Let  $x \in \bar{\mathbb{R}}$  and  $x \in U$  a open set then we have*

1. *If  $x \in \mathbb{R}$  then there exists a  $\varepsilon > 0$  such that  $x \in ]x - \varepsilon, x + \varepsilon[ \subseteq U$*
2. *If  $x = \infty$  then there exists a  $\varepsilon > 0$  such that  $x \in ]\varepsilon, \infty] \subseteq U$*
3. *If  $x = -\infty$  then there exists a  $\varepsilon > 0$  such that  $x \in [-\infty, -\varepsilon[ \subseteq U$*

**Proof.** Let  $x \in \bar{\mathbb{R}}$  and  $x \in U$  a open set then

1. If  $x \in \mathbb{R}$  we have, as  $\mathcal{B}_{\bar{\mathbb{R}}}$  is a basis for  $\mathcal{T}_{\bar{\mathbb{R}}}$ , the following possible cases to consider
  - $\exists \delta > 0, y \in \mathbb{R} \vdash x \in ]y - \delta, y + \delta[$ . then  $y - \delta < x < y + \delta$  take then  $\varepsilon = \min(x - (y - \delta), (y + \delta) - x) > 0$ . Then  $x - \varepsilon > x - (x - (y - \delta)) = y - \delta$  and  $x + \varepsilon < x + (y + \delta) - x = y + \delta$  so that  $x \in ]x - \varepsilon, x + \varepsilon[ \subseteq ]y - \delta, y + \delta[ \subseteq U$  proving that  $x \in ]x - \varepsilon, x + \varepsilon[ \subseteq U$ .
  - $\exists \delta \in \mathbb{R} \vdash x \in ]\delta, \infty]$ . then  $\delta < x < \infty$ , take then  $\varepsilon = x - \delta > 0$  then we have  $x - \varepsilon = x - x + \delta = \delta$  and  $x + \varepsilon < \infty$  so that  $x \in ]x - \varepsilon, x + \varepsilon[ \subseteq ]\delta, \infty] \subseteq U$
  - $\exists \delta \in \mathbb{R} \vdash x \in [-\infty, \delta[$ . then  $-\infty < x < \delta$ , take then  $\varepsilon = \delta - x > 0$  then we have  $x + \varepsilon < \delta$  and  $-\infty < x - \varepsilon$  so that  $x \in ]x - \varepsilon, x + \varepsilon[ \subseteq [-\infty, \delta[ \subseteq U$ .
2. If  $x = \infty$  then we have, as  $\mathcal{B}_{\bar{\mathbb{R}}}$  is a basis for  $\mathcal{T}_{\bar{\mathbb{R}}}$ , the following possible cases to consider
  - $\exists \delta > 0, y \in \mathbb{R} \vdash x \in ]y - \delta, y + \delta[$ . then  $\infty < y + \delta < \infty$  a contradiction so this will never happen.
  - $\exists \delta \in \mathbb{R} \vdash x \in ]\delta, \infty]$ . then as  $1 < \infty = x = \infty$  we have, if we take  $\varepsilon = \max(\delta, 1) > 0$ , that  $\delta \leq \varepsilon$  so that  $x \in ]\varepsilon, \infty] \subseteq ]\delta, \infty] \subseteq U$
  - $\exists \delta \in \mathbb{R} \vdash x \in [-\infty, \delta[$ . then  $\infty = x < \delta < \infty$  a contradiction so this will not apply
3. If  $x = -\infty$  then we have as  $\mathcal{B}_{\bar{\mathbb{R}}}$  is a basis for  $\mathcal{T}_{\bar{\mathbb{R}}}$  the following possible cases
  - $\exists \delta > 0, y \in \mathbb{R} \vdash x \in ]y - \delta, y + \delta[$ . then  $-\infty < y - \delta < x = -\infty$  a contradiction.
  - $\exists \delta \in \mathbb{R} \vdash x \in ]\delta, \infty]$ . then  $-\infty < \delta < x = -\infty$  a contradiction.
  - $\exists \delta \in \mathbb{R} \vdash x \in [-\infty, \delta[$ . then as  $-\infty < -1$  we have if we take  $\varepsilon = -\min(-1, \delta) > 0$  that  $-\varepsilon \leq \delta$  so that  $x = -\infty \in [-\infty, -\varepsilon[ \subseteq [-\infty, \delta[ \subseteq U$   $\square$

We show now that continuous functions to  $\mathbb{R}$  are continuous in  $\bar{\mathbb{R}}$ .

**Theorem 17.42.** *Let  $\langle X, \mathcal{T} \rangle$  be a topological space and  $f: X \rightarrow \mathbb{R}$  a continuous function (using  $\mathcal{T}_{\mathbb{R}}$ ) then  $f: X \rightarrow \bar{\mathbb{R}}$  is continuous in  $\mathcal{T}_{\bar{\mathbb{R}}}$ .*

**Proof.** Let  $x \in X$  and  $f(x) \in U$  then as  $f(x) \in \mathbb{R}$  we have by the above theorem a  $\varepsilon > 0$  such that  $]f(x) - \varepsilon, f(x) + \varepsilon[ \subseteq U$ . As  $f: X \rightarrow \mathbb{R}$  is continuous there exists a  $V \in \mathcal{T}$  such that  $f(V) \subseteq ]f(x) - \varepsilon, f(x) + \varepsilon[ \subseteq U$  proving that  $f: X - \bar{\mathbb{R}}$  is continuous.  $\square$

**Theorem 17.43.** If  $U \in \mathcal{T}_{\bar{\mathbb{R}}}$  then if  $\delta \in \mathbb{R}$  we have that

1.  $\delta + U = \{\delta + x | x \in U\} \in \mathcal{T}_{\bar{\mathbb{R}}}$ .
2. If  $\delta \neq 0$  we have  $\delta \cdot U = \{\delta \cdot x | x \in U\} \in \mathcal{T}_{\bar{\mathbb{R}}}$ .

**Proof.**

1. If  $x \in \delta + U$  then there exists a  $u \in U$  such that  $x = \delta + u$ , we have now for  $u$  the following possibilities

$u \in \mathbb{R}$ . then using 17.41 there exists a  $\varepsilon > 0$  such that  $u \in ]u - \varepsilon, u + \varepsilon[ \subseteq U$ , then  $u - \varepsilon < u < u + \varepsilon \Rightarrow (u - \varepsilon) + \delta < u + \delta < (u + \varepsilon) + \delta$  proving that  $x = u + \delta \in ](u - \varepsilon) + \delta, (u + \varepsilon) + \delta[ = ](u + \delta) - \varepsilon, (u + \delta) + \varepsilon[$ . Also if  $y \in ](u + \delta) - \varepsilon, (u + \delta) + \varepsilon[$  then  $(u + \delta) - \varepsilon < y < (u + \delta) + \varepsilon \Rightarrow u - \varepsilon < y - \delta < u + \varepsilon \Rightarrow y - \delta \in ]u - \varepsilon, u + \varepsilon[ \subseteq U$  so that  $y = \delta + (y - \delta) \in \delta + U$  proving that  $x \in ](u + \delta) - \varepsilon, (u + \delta) + \varepsilon[ \subseteq \delta + U$ . This proves that  $\exists B \in \mathcal{B}_{\bar{\mathbb{R}}}$  such that  $x \in B \subseteq \delta + U$ .

$u = \infty$ . then using 17.41 there exists a  $\varepsilon > 0$  such that  $u = \infty \in ]\varepsilon, \infty[ \subseteq U$ , take now  $]\varepsilon + \delta, \infty]$  then  $x = \delta + \infty = \infty \in ]\varepsilon + \delta, \infty]$ . If now  $y \in ]\varepsilon + \delta, \infty]$  then we have either  $y = \infty = \delta + \infty = \delta + u \in \delta + U$  or  $y \in \mathbb{R}$  so that  $\varepsilon + \delta < y < \infty \Rightarrow \varepsilon < y - \delta < \infty \Rightarrow y - \delta \in ]\varepsilon, \infty[ \subseteq U \Rightarrow y = \delta + (y - \delta) \in \delta + U$  proving that  $x \in ]\varepsilon + \delta, \infty[ \subseteq \delta + U$ . This proves that  $\exists B \in \mathcal{B}_{\bar{\mathbb{R}}}$  such that  $x \in B \subseteq \delta + U$ .

$u = -\infty$ . then using 17.41 there exists a  $\varepsilon > 0$  such that  $u = -\infty \in [-\infty, -\varepsilon[ \subseteq U$ , take now  $[-\infty, -\varepsilon + \delta]$  then  $x = \delta + (-\infty) = -\infty \in [-\infty, -\varepsilon + \delta]$ . If now  $y \in [-\infty, -\varepsilon + \delta]$  we have either  $y = -\infty = \delta + (-\infty) = \delta + u \in \delta + U$  or  $y \in \mathbb{R}$  so that  $-\infty < y < -\varepsilon + \delta \Rightarrow -\infty < y - \delta < -\varepsilon \Rightarrow y - \delta \in [-\infty, -\varepsilon[ \subseteq U \Rightarrow y = \delta + (y - \delta) \in \delta + U$  proving that  $x \in [-\infty, -\varepsilon + \delta] \subseteq \delta + U$ . This proves that  $\exists B \in \mathcal{B}_{\bar{\mathbb{R}}}$  such that  $x \in B \subseteq \delta + U$ .

As in all possible cases we have  $\forall x \in \delta + U$  that  $\exists B \in \mathcal{B}_{\bar{\mathbb{R}}}$  such that  $x \in B \subseteq \delta + U$  it follows that  $\delta + U$  is a open set.

2. If  $x \in \delta \cdot U$  then there exists a  $u \in U$  such that  $x = \delta \cdot u$  we have then for  $u$  the following possibilities

$u \in \mathbb{R}$ . then using 17.41 there exists a  $\varepsilon > 0$  such that  $u \in ]u - \varepsilon, u + \varepsilon[ \subseteq U$  so that  $u - \varepsilon < u < u + \varepsilon$  consider then the following possibilities for  $\delta \neq 0$ :

$0 < \delta$ . then  $\delta \cdot u - \delta \cdot \varepsilon < \delta \cdot u < \delta \cdot u + \delta \cdot \varepsilon$  so that  $x = \delta \cdot u \in ]\delta \cdot u - \delta \cdot \varepsilon, \delta \cdot u + \delta \cdot \varepsilon[ = ]u \cdot \delta - \rho, u \cdot \delta + \rho[$  where  $\rho = \delta \cdot \varepsilon > 0$ . If now  $y \in ]u \cdot \delta - \rho, u \cdot \delta + \rho[ \Rightarrow u \cdot \delta - \rho < y < u \cdot \delta + \rho \Rightarrow \frac{u \cdot \delta}{\delta} - \frac{\rho}{\delta} < \frac{y}{\delta} < \frac{u \cdot \delta}{\delta} + \frac{\rho}{\delta} \Rightarrow u - \varepsilon < \frac{y}{\delta} < u + \varepsilon \Rightarrow \frac{y}{\delta} \in ]u - \varepsilon, u + \varepsilon[ \subseteq U \Rightarrow y = \delta \cdot \frac{y}{\delta} \in \delta \cdot U$  proving that  $x \in ]\delta \cdot u - \rho, \delta \cdot u + \rho[ \subseteq \delta \cdot U$ . This proves that  $\exists B \in \mathcal{B}_{\bar{\mathbb{R}}}$  such that  $x \in B \subseteq \delta \cdot U$

$\delta < 0$ . then  $\delta \cdot u + \delta \cdot \varepsilon < \delta \cdot u < \delta \cdot u - \delta \cdot \varepsilon \Rightarrow x = \delta \cdot u \in ]\delta \cdot u + \delta \cdot \varepsilon, \delta \cdot u - \delta \cdot \varepsilon[ = ]\delta \cdot u - \rho, \delta \cdot u + \rho[$  where  $\rho = -\delta \cdot \varepsilon > 0$ . If now  $y \in ]u \cdot \delta - \rho, u \cdot \delta + \rho[ \Rightarrow u \cdot \delta - \rho < y < u \cdot \delta + \rho \Rightarrow \frac{u \cdot \delta}{\delta} + \frac{\rho}{\delta} < \frac{y}{\delta} < \frac{u \cdot \delta}{\delta} - \frac{\rho}{\delta} \Rightarrow u - \varepsilon < \frac{y}{\delta} < u + \varepsilon \Rightarrow \frac{y}{\delta} \in ]u - \varepsilon, u + \varepsilon[ \subseteq U$  so that  $y = \delta \cdot \frac{y}{\delta} \in \delta \cdot U$  proving that  $x \in ]u - \varepsilon, u + \varepsilon[ \subseteq \delta \cdot U$ . This proves that  $\exists B \in \mathcal{B}_{\bar{\mathbb{R}}}$  such that  $x \in B \subseteq \delta \cdot U$

$u = \infty$ . then using 17.41 there exists a  $\varepsilon > 0$  such that  $u \in ]\varepsilon, \infty[ \subseteq U$ , we have to consider the following cases for  $\delta$

$0 < \delta$ . then  $x = \delta \cdot u = \infty \in ]\varepsilon \cdot \delta, \infty[$ . If  $y \in ]\varepsilon \cdot \delta, \infty[$  then if  $y = \infty = \delta \cdot u$  we have  $y = \delta \cdot u \in \delta \cdot U$  and if  $y \in \mathbb{R}$  we have that  $\varepsilon \cdot \delta < y < \infty$  so that  $\varepsilon < \frac{y}{\delta} < \infty \Rightarrow \frac{y}{\delta} \in ]\varepsilon, \infty[ \subseteq U \Rightarrow y = \delta \cdot \frac{y}{\delta} \in \delta \cdot U$ . So we have proved that  $x \in ]\varepsilon \cdot \delta, \infty[ \subseteq U$  and thus that  $\exists B \in \mathcal{B}_{\bar{\mathbb{R}}}$  such that  $x \in B \subseteq \delta \cdot U$

**$\delta < 0$ .** then  $x = \delta \cdot u = -\infty \in [-\infty, \varepsilon \cdot \delta]$ . If  $y \in [-\infty, \varepsilon \cdot \delta]$  then if  $y = -\infty = \delta \cdot u$  we have  $y = \delta \cdot u \in \delta \cdot U$  and if  $y \in \mathbb{R}$  we have that  $-\infty < y < \varepsilon \cdot \delta$  so that  $\varepsilon < \frac{y}{\delta} < \infty \Rightarrow \frac{y}{\delta} \in [\varepsilon, \infty] \subseteq U \Rightarrow y = \delta \cdot \frac{y}{\delta} \in \delta \cdot U$  So we have proved that  $x \in [-\infty, \varepsilon \cdot \delta] \subseteq \delta \cdot U$  and thus that  $\exists B \in \mathcal{B}_{\overline{\mathbb{R}}}$  such that  $x \in B \subseteq \delta \cdot U$

**$u = -\infty$ .** then using 17.41 there exists a  $\varepsilon > 0$  such that  $u \in [-\infty, -\varepsilon] \subseteq U$ , we have to consider the following cases for  $\delta$

**$0 < \delta$ .** then  $x = \delta \cdot u = -\infty \in [-\infty, -\varepsilon \cdot \delta]$ . If  $y \in [-\infty, -\varepsilon \cdot \delta]$  then if  $y = -\infty$  we have  $y = \delta \cdot (-\infty) = \delta \cdot u \in \delta \cdot U$  and if  $y \in \mathbb{R}$  we have that  $-\infty < y < -\varepsilon \cdot \delta$  so that  $-\infty < \frac{y}{\delta} < -\varepsilon \Rightarrow \frac{y}{\delta} \in [-\infty, -\varepsilon] \subseteq U \Rightarrow y = \delta \cdot \frac{y}{\delta} \in \delta \cdot U$ . So we have proved that  $x \in [-\infty, -\varepsilon \cdot \delta] \subseteq U$  and thus that  $\exists B \in \mathcal{B}_{\overline{\mathbb{R}}}$  such that  $x \in B \subseteq \delta \cdot U$

**$\delta < 0$ .** then  $x = \delta \cdot u = \infty \in [\varepsilon \cdot \delta, \infty]$ . If  $y \in [\varepsilon \cdot \delta, \infty]$  then if  $y = \infty$  we have  $y = \delta \cdot \infty = \delta \cdot u \in \delta \cdot U$  and if  $y \in \mathbb{R}$  we have that  $\varepsilon \cdot \delta < y < \infty$  so that  $-\infty < \frac{y}{\delta} < \varepsilon \Rightarrow \frac{y}{\delta} \in [-\infty, \varepsilon] \subseteq U \Rightarrow y = \delta \cdot \frac{y}{\delta} \in \delta \cdot U$ . So we have proved that  $x \in [\varepsilon \cdot \delta, \infty] \subseteq \delta \cdot U$  and thus that  $\exists B \in \mathcal{B}_{\overline{\mathbb{R}}}$  such that  $x \in B \subseteq \delta \cdot U$

As in all possible cases we have  $\forall x \in \delta \cdot U \ \exists B \in \mathcal{B}_{\overline{\mathbb{R}}}$  such that  $x \in B \subseteq \delta \cdot U$  which proves that  $\delta \cdot U$  is open.  $\square$

**Note 17.44.** The restrictions in the above theorem are needed, for example  $U = ]-1, 1[$  is a open set but  $0 \cdot U = 0 \cdot ]-1, 1[ = \{0\}$  a closed set. Also  $\infty + U = \infty + ]-1, 1[ = \{\infty\}$  a closed set.

**Theorem 17.45.**  $\mathcal{T}_{\overline{\mathbb{R}}}$  is Hausdorff (see 11.227).

**Proof.** Let  $x_1, x_2 \in \overline{\mathbb{R}}$  with  $x_1 \neq x_2$  then we have to check the following cases

$x_1 \in \mathbb{R} \wedge x_2 \in \mathbb{R}$ . then we have either

**$x_1 < x_2$ .** Take then  $\delta = \frac{x_2 - x_1}{2} > 0$  then  $x_1 \in ]x_1 - \delta, x_2 + \delta[ \in \mathcal{T}_{\overline{\mathbb{R}}}$ ,  $x_2 \in ]x_2 - \delta, x_2 + \delta[ \in \mathcal{T}_{\overline{\mathbb{R}}}$ , as  $]x_1 - \delta, x_1 + \delta[ = ]x_1 - \delta, \frac{x_1 + x_2}{2}[, ]x_2 - \delta, x_2 + \delta[ = ]\frac{x_1 + x_2}{2}, x_2 + \delta[$  we have  $\emptyset = ]x_1 - \delta, x_1 + \delta[ \cap ]x_2 - \delta, x_2 + \delta[$ .

**$x_2 < x_1$ .** Take then  $\delta = \frac{x_1 - x_2}{2} > 0$  then  $x_1 \in ]x_1 - \delta, x_2 + \delta[ \in \mathcal{T}_{\overline{\mathbb{R}}}$ ,  $x_2 \in ]x_2 - \delta, x_2 + \delta[ \in \mathcal{T}_{\overline{\mathbb{R}}}$ , as  $]x_2 - \delta, x_2 + \delta[ = ]x_2 - \delta, \frac{x_1 + x_2}{2}[, ]x_1 - \delta, x_1 + \delta[ = ]\frac{x_1 + x_2}{2}, x_1 + \delta[$  we have  $\emptyset = ]x_1 - \delta, x_1 + \delta[ \cap ]x_2 - \delta, x_2 + \delta[$ .

**$x_1 \in \mathbb{R} \wedge x_2 = \infty$ .** then  $x_1 \in ]x_1 - 1, x_1 + 1[ \in \mathcal{T}_{\overline{\mathbb{R}}}$  and  $\infty \in ]x_1 + 1, \infty] \in \mathcal{T}_{\overline{\mathbb{R}}}$  with  $]x_1 - 1, x_1 + 1[ \cap ]x_1 + 1, \infty] = \emptyset$

**$x_1 \in \mathbb{R} \wedge x_2 = -\infty$ .** then  $x_1 \in ]x_1 - 1, x_1 + 1[ \in \mathcal{T}_{\overline{\mathbb{R}}}$  and  $-\infty \in [-\infty, x_1 - 1[ \in \mathcal{T}_{\overline{\mathbb{R}}}$  with  $]x_1 - 1, x_1 + 1[ \cap [-\infty, x_1 - 1[ = \emptyset$

**$x_1 = \infty \wedge x_2 \in \mathbb{R}$ .** then  $\infty \in ]x_2 + 1, \infty] \in \mathcal{T}_{\overline{\mathbb{R}}}$  and  $x_2 \in ]x_2 - 1, x_2 + 1[ \in \mathcal{T}_{\overline{\mathbb{R}}}$  with  $]x_2 + 1, \infty] \cap ]x_2 - 1, x_2 + 1[ = \emptyset$

**$x_1 = \infty \wedge x_2 = -\infty$ .** then  $\infty \in ]0, \infty] \in \mathcal{T}_{\overline{\mathbb{R}}}$  and  $-\infty \in [-\infty, 0[$  with  $]0, \infty] \cap [-\infty, 0[ = \emptyset$

**$x_1 = -\infty \wedge x_2 \in \mathbb{R}$ .** then  $-\infty \in [-\infty, x_2 - 1[ \in \mathcal{T}_{\overline{\mathbb{R}}}$  and  $x_2 \in ]x_2 - 1, x_2 + 1[ \in \mathcal{T}_{\overline{\mathbb{R}}}$  with  $[-\infty, x_2 - 1[ \cap ]x_2 - 1, x_2 + 1[ = \emptyset$

**$x_1 = -\infty \wedge x_2 = \infty$ .** then  $-\infty \in [-\infty, 0[ \in \mathcal{T}_{\overline{\mathbb{R}}}$  and  $\infty \in ]0, \infty] \in \mathcal{T}_{\overline{\mathbb{R}}}$  with  $[-\infty, 0[ \cap ]0, \infty] = \emptyset$

So in all cases there exists a  $V_1, V_2 \in \mathcal{T}_{\overline{\mathbb{R}}}$  with  $x \in V_1 \wedge x \in V_2 \wedge V_1 \cap V_2 = \emptyset$  proving that  $\mathcal{T}_{\overline{\mathbb{R}}}$  is Hausdorff.  $\square$

### 17.3 Sequences in $\overline{\mathbb{R}}$

Note that the sum in  $\overline{\mathbb{R}}$  is not always defined so if we want to see how the supremum and infimum behaves on sums we have to be carefull, that is the reason that we restrict ourself to positive extended reals in many cases.

**Notation 17.46.** Given  $k \in \mathbb{N}_0$  we note  $\{i \in \mathbb{N}_0 \mid k \leq i\}$  as  $\{k, \dots, \infty\}$

**Definition 17.47.** Let  $k \in \mathbb{N}_0$ ,  $\{x_i\}_{i \in \{k, \dots, \infty\}}$  be a sequence in a partial ordered set  $\langle S, \leq \rangle$  then we say that  $\{x_i\}_{i \in \{k, \dots, \infty\}}$  is a increasing (decreasing) sequence if  $\forall i \in \{k, \dots, \infty\}$  we have  $x_i \leq x_{i+1}$  ( $x_{i+1} \leq x_i$ )

**Lemma 17.48.** Let  $k \in \mathbb{N}_0$ ,  $\{x_i\}_{i \in \{k, \dots, \infty\}}$  be a increasing (or decreasing) sequence then  $\forall n, m \in \{k, \dots, \infty\}$  with  $n \leq m$  we have  $x_n \leq x_m$  ( $x_m \leq x_n$ )

**Proof.** The proof is trivial using mathematical induction. Let  $\{x_i\}_{i \in \{k, \dots, \infty\}}$  be a increasing (or decreasing) sequence  $n \in \{k, \dots, \infty\}$  and define  $S_n = \{m \in \mathbb{N} \mid x_n \leq x_{n+m} \text{ [or } x_{n+m} \leq x_n]\}$  then we have

$0 \in S$ . this is trivial as  $x_n \leq x_n = x_{n+0}$  [or  $x_{n+0} = x_n \leq x_n$ ]

**if  $m \in S \Rightarrow m + 1 \in S$ .** from  $m \in S$  we have  $x_n \leq x_{n+m} \leq x_{n+m+1} \Rightarrow x_n \leq x_{n+m+1}$  [or  $x_{n+m+1} \leq x_{n+m} \leq x_n \Rightarrow x_{n+m+1} \leq x_n$ ] proving  $m + 1 \in S$

Let now  $n, m \in \{k, \dots, \infty\}$  with  $n \leq m$  [or  $m \leq n$ ] then  $i = m - n \in S_n \Rightarrow x_n \leq x_{n+(m-n)} = x_m$  [or  $x_m = x_{n+(m-n)} \leq x_n$ ]  $\square$

Recap the following result for increasing/decreasing sequences of real numbers [see 11.366]

**Lemma 17.49.** Let  $k \in \mathbb{N}_0$ ,  $\{x_i\}_{i \in \{k, \dots, \infty\}}$  be a increasing [decreasing] sequence of real numbers then we have

1. If  $\{x_i \mid i \in \{k, \dots, \infty\}\}$  has a supremum [infimum] then  $\lim_{i \rightarrow \infty} x_i$  exists and  $\lim_{i \rightarrow \infty} x_i = \sup(\{x_i \mid i \in \{k, \dots, \infty\}\})$
2. If  $\lim_{i \rightarrow \infty} x_i$  exists then  $\sup(\{x_i \mid i \in \{k, \dots, \infty\}\}) / \inf(\{x_i \mid i \in \{k, \dots, \infty\}\})$  exists and  $\sup(\{x_i \mid i \in \{k, \dots, \infty\}\}) = \lim_{i \rightarrow \infty} x_i / \inf(\{x_i \mid i \in \{k, \dots, \infty\}\}) = \lim_{i \rightarrow \infty} x_i$

**Proof.** This is already proved in 11.366  $\square$

We want to extend the above lemma to the extended real numbers. As the supremum and the infimum in  $\bar{\mathbb{R}}$  always exists this suggest that the limit always exists for increasing and decreasing sequences of extended reals. However we must then first extend also the limit of sequences to the set of the extended real numbers.

**Theorem 17.50.** Let  $k \in \mathbb{N}_0$ ,  $\{x_i\}_{i \in \{n, \dots, \infty\}}$  be a sequence of extended reals then

1.  $\inf(\{\sup(\{x_i \mid i \in \{k, \dots, \infty\}\}) \mid k \in \{n, \dots, \infty\}\})$  exists
2.  $\sup(\{\inf(\{x_i \mid i \in \{k, \dots, \infty\}\}) \mid k \in \{n, \dots, \infty\}\})$  exists
3.  $\forall m \geq n$  we have  

$$\inf(\{\sup(\{x_i \mid i \in \{k, \dots, \infty\}\}) \mid k \in \{n, \dots, \infty\}\}) = \inf(\{\sup(\{x_i \mid i \in \{k, \dots, \infty\}\}) \mid k \in \{m, \dots, \infty\}\})$$
4.  $\forall m \geq n$  we have  

$$\sup(\{\inf(\{x_i \mid i \in \{k, \dots, \infty\}\}) \mid k \in \{n, \dots, \infty\}\}) = \sup(\{\inf(\{x_i \mid i \in \{k, \dots, \infty\}\}) \mid k \in \{m, \dots, \infty\}\})$$
5.  $\forall m \in \mathbb{N}_0$  we have
  - a.  $\inf(\{\sup(\{x_i \mid i \in \{k, \dots, \infty\}\}) \mid k \in \{n, \dots, \infty\}\}) = \inf(\{\sup(\{x_{i+m} \mid i \in \{k, \dots, \infty\}\}) \mid k \in \{n, \dots, \infty\}\})$
  - b.  $\sup(\{\inf(\{x_i \mid i \in \{k, \dots, \infty\}\}) \mid k \in \{n, \dots, \infty\}\}) = \sup(\{\inf(\{x_{i+m} \mid i \in \{k, \dots, \infty\}\}) \mid k \in \{n, \dots, \infty\}\})$

**Proof.**

1. The existence follows from the fact that in  $\bar{\mathbb{R}}$  every non empty set has a sup and inf [see 17.12]

2. The existence follows from the fact that in  $\bar{\mathbb{R}}$  every non empty set has a sup and inf [see 17.12]

3. Given  $k \in \{n, \dots, \infty\}$  take  $S_k = \sup(\{x_i | i \in \{k, \dots, \infty\}\})$  then  $\{S_k | k \in \{m, \dots, \infty\}\} \subseteq \{S_k | k \in \{n, \dots, \infty\}\}$  proving using 2.175 that  $\inf(S_k | k \in \{n, \dots, \infty\}) \leq \inf(S_k | k \in \{m, \dots, \infty\})$  proving that

$$\inf(\{\sup(\{x_i | i \in \{k, \dots, \infty\}\}) | k \in \{n, \dots, \infty\}\}) \leq \inf(\{\sup(\{x_i | i \in \{k, \dots, \infty\}\}) | k \in \{m, \dots, \infty\}\}) \quad (17.12)$$

For the opposite inclusion, given  $k \in \{n, \dots, \infty\}$ , take  $l = \max(k, m)$ , then  $k \leq l$  and  $l \in \{m, \dots, \infty\}$ . We have then that  $\{x_i | i \in \{l, \dots, \infty\}\} \subseteq \{x_i | i \in \{k, \dots, \infty\}\}$  so that using 2.175  $S_l = \sup(\{x_i | i \in \{l, \dots, \infty\}\}) \leq \sup(\{x_i | i \in \{k, \dots, \infty\}\}) = S_k$ . So for every  $S_k \in \{S_i | i \in \{n, \dots, \infty\}\}$  we find a  $S_l \in \{S_i | i \in \{m, \dots, \infty\}\}$  such that  $S_l \leq S_k$ . Using 2.176 we have then that  $\inf(S_i | i \in \{m, \dots, \infty\}) \leq \inf(S_i | i \in \{n, \dots, \infty\})$  or  $\inf(\{\sup(\{x_i | i \in \{k, \dots, \infty\}\}) | k \in \{m, \dots, \infty\}\}) \leq \inf(\{\sup(\{x_i | i \in \{k, \dots, \infty\}\}) | k \in \{n, \dots, \infty\}\})$ . Hence using 17.12 we have

$$\inf(\{\sup(\{x_i | i \in \{k, \dots, \infty\}\}) | k \in \{n, \dots, \infty\}\}) = \inf(\{\sup(\{x_i | i \in \{k, \dots, \infty\}\}) | k \in \{m, \dots, \infty\}\})$$

4. Given  $k \in \{n, \dots, \infty\}$  take  $I_k = \inf(\{x_i | i \in \{k, \dots, \infty\}\})$  then  $\{I_k | k \in \{m, \dots, \infty\}\} \subseteq \{I_k | k \in \{n, \dots, \infty\}\}$  proving using 2.175 that  $\sup(\{S_k | k \in \{m, \dots, \infty\}\}) \leq \sup(\{S_k | k \in \{n, \dots, \infty\}\})$ , hence

$$\sup(\{\inf(\{x_i | i \in \{k, \dots, \infty\}\}) | k \in \{m, \dots, \infty\}\}) \leq \sup(\{\inf(\{x_i | i \in \{k, \dots, \infty\}\}) | k \in \{n, \dots, \infty\}\}) \quad (17.13)$$

For the opposite inclusion, given  $k \in \{n, \dots, \infty\}$  take  $l = \max(k, m)$  then  $k \leq l$  and  $l \in \{m, \dots, \infty\}$ . We have then that  $\{x_i | i \in \{l, \dots, \infty\}\} \subseteq \{x_i | i \in \{k, \dots, \infty\}\}$  so that using 2.175 we have  $I_k = \inf(\{x_i | i \in \{k, \dots, \infty\}\}) \leq \inf(\{x_i | i \in \{l, \dots, \infty\}\}) = I_l$ . So for every  $I_k \in \{I_i | i \in \{n, \dots, \infty\}\}$  we can find a  $I_l \in \{I_i | i \in \{m, \dots, \infty\}\}$  such that  $I_k \leq I_l$ . Using 2.176 we have then that  $\sup(\{I_i | i \in \{n, \dots, \infty\}\}) \leq \sup(\{I_i | i \in \{m, \dots, \infty\}\})$  or  $\sup(\{\inf(\{x_i | i \in \{k, \dots, \infty\}\}) | k \in \{n, \dots, \infty\}\}) \leq \sup(\{\inf(\{x_i | i \in \{k, \dots, \infty\}\}) | k \in \{m, \dots, \infty\}\})$ . Hence using 17.13 we have that

$$\sup(\{\inf(\{x_i | i \in \{k, \dots, \infty\}\}) | k \in \{n, \dots, \infty\}\}) = \sup(\{\inf(\{x_i | i \in \{k, \dots, \infty\}\}) | k \in \{m, \dots, \infty\}\})$$

5. Given  $k \in \{n+m, \dots, \infty\}$  define  $A_k = \{x_i | i \in \{k, \dots, \infty\}\}$  and for  $k \in \{n, \dots, \infty\}$   $B_k = \{x_{i+m} | i \in \{k, \dots, \infty\}\}$ . Then we have for  $k \in \{n+m, \dots, \infty\}$

$$\begin{aligned} x \in A_k &\Rightarrow \exists i \in \{k, \dots, \infty\} \text{ such that } x = x_i \\ &\Rightarrow k \leq i \Rightarrow k - m \leq i - m \wedge x_i = x_{(i-m)+m} \\ &\Rightarrow x_i \in B_{k-m} \\ &\Rightarrow x \in B_{k-m} \\ x \in B_{k-m} &\Rightarrow \exists i \in \{k-m, \dots, \infty\} \text{ such that } x = x_{i+m} \\ &\Rightarrow k - m \leq i \Rightarrow k \leq i + m \\ &\Rightarrow x \in A_k \end{aligned}$$

proving that

$$\forall k \in \{n+m, \dots, \infty\} \text{ we have that } A_k = B_{k-m} \quad (17.14)$$

On the other hand if  $k \in \{n, \dots, \infty\}$  then we have

$$\begin{aligned} x \in B_k &\Rightarrow \exists i \in \{k, \dots, \infty\} \text{ such that } x = x_{i+m} \\ &\Rightarrow k + m \leq i + m \\ &\Rightarrow x \in A_{k+m} \\ x \in A_{k+m} &\Rightarrow \exists i \in \{k+m, \dots, \infty\} \text{ such that } x = x_i \\ &\Rightarrow k + m \leq i \Rightarrow k \leq i - m \text{ and } x_i = x_{(i-m)+m} \\ &\Rightarrow x \in B_k \end{aligned}$$

proving that

$$\forall k \in \{n, \dots, \infty\} \text{ we have that } B_k = A_{k+m} \quad (17.15)$$

Now if  $x \in \{\sup(A_k) | k \in \{n+m, \dots, \infty\}\}$  then  $\exists k \in \{n+m, \dots, \infty\}$  such that  $x = \sup(A_k) \stackrel{17.14}{=} \sup(B_{k-m})$  where  $k-m \in \{n, \dots, \infty\}$ , hence  $x \in \{\sup(B_k) | k \in \{n, \dots, \infty\}\}$  proving that  $\{\sup(A_k) | k \in \{n+m, \dots, \infty\}\} \subseteq \{\sup(B_k) | k \in \{n, \dots, \infty\}\}$ . Using 2.175 it follows that:

$$\inf(\{\sup(B_k) | k \in \{n, \dots, \infty\}\}) \leq \inf(\{\sup(A_k) | k \in \{n+m, \dots, \infty\}\}) \quad (17.16)$$

If  $x \in \{\sup(B_k) | k \in \{n, \dots, \infty\}\}$  then  $\exists k \in \{1, \dots, n\}$  such that  $x = \sup(B_k) \stackrel{17.15}{=} \sup(A_{k+m})$  where  $k+m \in \{n+m, \dots, \infty\}$ , hence  $x \in \{\sup(A_k) | k \in \{n+m, \dots, \infty\}\}$  proving that  $\{\sup(B_k) | k \in \{n, \dots, \infty\}\} \subseteq \{\sup(A_k) | k \in \{n+m, \dots, \infty\}\}$ . Using 2.175 it follows that:

$$\inf(\{\sup(A_k) | k \in \{n+m, \dots, \infty\}\}) \leq \inf(\{\sup(B_k) | k \in \{n, \dots, \infty\}\}) \quad (17.17)$$

Combining 17.16 and 17.17 together with the definition of  $A_k, B_k$  we have

$$\inf(\{\sup(\{x_i | i \in \{k, \dots, \infty\}\}) | k \in \{n+m, \dots, \infty\}\}) = \inf(\{\sup(x_{i+m} | i \in \{k, \dots, \infty\}) | k \in \{n, \dots, \infty\}\}) \quad (17.18)$$

Using then (3) on the above proves:

$$\inf(\{\sup(\{x_i | i \in \{k, \dots, \infty\}\}) | k \in \{n, \dots, \infty\}\}) = \inf(\{\sup(x_{i+m} | i \in \{k, \dots, \infty\}) | k \in \{n, \dots, \infty\}\}) \quad (17.19)$$

Now if  $x \in \{\inf(A_k) | k \in \{n+m, \dots, \infty\}\}$  then  $\exists k \in \{n+m, \dots, \infty\}$  such that  $x = \inf(A_k) \stackrel{17.14}{=} \inf(B_{k-m})$  where  $k-m \in \{n, \dots, \infty\}$ , hence  $x \in \{\inf(B_k) | k \in \{n, \dots, \infty\}\}$  proving that  $\{\inf(A_k) | k \in \{n+m, \dots, \infty\}\} \subseteq \{\inf(B_k) | k \in \{n, \dots, \infty\}\}$ . Applying then 2.175 gives

$$\sup(\{\inf(A_k) | k \in \{n+m, \dots, \infty\}\}) \leq \sup(\{\inf(B_k) | k \in \{n, \dots, \infty\}\}) \quad (17.20)$$

If  $x \in \{\inf(B_k) | k \in \{n, \dots, \infty\}\}$  then  $\exists k \in \{1, \dots, n\}$  such that  $x = \inf(B_k) \stackrel{17.15}{=} \inf(A_{k+m})$  where  $k+m \in \{n+m, \dots, \infty\}$ , hence  $x \in \{\inf(A_k) | k \in \{n+m, \dots, \infty\}\}$  proving that  $\{\inf(B_k) | k \in \{n, \dots, \infty\}\} \subseteq \{\inf(A_k) | k \in \{n+m, \dots, \infty\}\}$ . Applying then 2.175 gives:

$$\sup(\{\inf(B_k) | k \in \{n, \dots, \infty\}\}) \leq \sup(\{\inf(A_k) | k \in \{n+m, \dots, \infty\}\}) \quad (17.21)$$

Combining 17.20 and 17.21 together with the definition of  $A_k, B_k$  we have

$$\inf(\{\sup(\{x_i | i \in \{k, \dots, \infty\}\}) | k \in \{n+m, \dots, \infty\}\}) = \inf(\{\sup(x_{i+m} | i \in \{k, \dots, \infty\}) | k \in \{n, \dots, \infty\}\}) \quad (17.22)$$

Using then (3) on the above proves

$$\inf(\{\sup(\{x_i | i \in \{k, \dots, \infty\}\}) | k \in \{n, \dots, \infty\}\}) = \inf(\{\sup(x_{i+m} | i \in \{k, \dots, \infty\}) | k \in \{n, \dots, \infty\}\}) \quad (17.23)$$

Finally 5 (a) is proved by 17.19 and 5 (b) by 17.23 □

The above theorem motivates the following definition

**Definition 17.51.** Let  $k \in \mathbb{N}_0$ ,  $\{x_i\}_{i \in \{k, \dots, \infty\}}$  be a sequence of extended reals then

$$\liminf_{i \rightarrow \infty} x_i = \sup(\{\inf(\{x_i | i \in \{l, \dots, \infty\}\}) | l \in \{k, \dots, \infty\}\})$$

and

$$\limsup_{i \rightarrow \infty} x_i = \inf(\{\sup(\{x_i | i \in \{l, \dots, \infty\}\}) | l \in \{k, \dots, \infty\}\}).$$

**Note 17.52.** In the above above definition the starting index  $k$  is not mentioned in the notation  $\limsup_{i \rightarrow \infty} x_i$  and  $\liminf_{i \rightarrow \infty} x_i$  but this is ok because of the previous theorem [see 17.50] we can use every number in  $\{k, \dots, \infty\}$  as the starting index.

**Example 17.53.** Let  $k \in \mathbb{N}_0$  and  $\{x_n\}_{n \in \{k, \dots, n\}}$  defined by  $x_n = x \in \mathbb{R}$  then

$$x = \liminf_{n \rightarrow \infty} x_n = \limsup_{n \rightarrow \infty} x_n$$

**Proof.**

$$\begin{aligned}
 \liminf_{n \rightarrow \infty} x_n &= \sup (\{\inf (\{x_i | i \in \{n, \dots, \infty\}\}) | n \in \{k, \dots, \infty\}\}) \\
 &= \sup (\{\inf (\{x\}) | n \in \{k, \dots, \infty\}\}) \\
 &= \sup (\{x | n \in \{k, \dots, \infty\}\}) \\
 &= x \\
 &= \inf (\{x | n \in \{k, \dots, \infty\}\}) \\
 &= \inf (\{\sup (\{x\}) | n \in \{k, \dots, \infty\}\}) \\
 &= \inf (\{\sup (\{x_i | i \in \{n, \dots, \infty\}\}) | n \in \{k, \dots, \infty\}\}) \\
 &= \limsup_{n \rightarrow \infty} x_n
 \end{aligned}$$

□

Using this definition and the previous theorem [see 17.50] we have the following:

**Theorem 17.54.** Let  $k \in \mathbb{N}_0$ ,  $\{x_i\}_{i \in \{k, \dots, \infty\}}$  be a sequence of extended reals then we have

1.  $\liminf_{i \rightarrow \infty} x_i$  exists
2.  $\limsup_{i \rightarrow \infty} x_i$  exists
3.  $\forall n \in \mathbb{N}_0$  we have that
  - a.  $\liminf_{i \rightarrow \infty} x_i = \liminf_{i \rightarrow \infty} x_{i+n}$
  - b.  $\limsup_{i \rightarrow \infty} x_i = \limsup_{i \rightarrow \infty} x_{i+n}$

**Proposition 17.55.** Let  $k \in \mathbb{N}_0$ ,  $\{x_i\}_{i \in \{k, \dots, \infty\}}$ ,  $\{y_i\}_{i \in \{k, \dots, \infty\}}$  be sequences of extended reals such that  $\forall i \in \{k, \dots, \infty\} x_i \leq y_i$  then

1.  $\liminf_{n \rightarrow \infty} x_n \leq \liminf_{n \rightarrow \infty} y_n$
2.  $\limsup_{n \rightarrow \infty} x_n \leq \limsup_{n \rightarrow \infty} y_n$

**Proof.**

1. Let  $l \in \{k, \dots, \infty\}$  then using 2.176 we have that  $\inf (\{x_i | i \in \{l, \dots, \infty\}\}) \leq \inf (\{y_i | i \in \{l, \dots, \infty\}\})$ . Applying 2.176 again we have that

$$\begin{aligned}
 \liminf_{n \rightarrow \infty} x_n &\stackrel{\text{def}}{=} \sup (\{\inf (\{x_i | i \in \{l, \dots, \infty\}\}) | l \in \{k, \dots, \infty\}\}) \\
 &\leq \sup (\{\inf (\{y_i | i \in \{l, \dots, \infty\}\}) | l \in \{k, \dots, \infty\}\}) \\
 &= \liminf_{n \rightarrow \infty} y_n
 \end{aligned}$$

2. Let  $l \in \{k, \dots, \infty\}$  then using 2.176 we have that  $\sup (\{x_i | i \in \{l, \dots, \infty\}\}) \leq \sup (\{y_i | i \in \{l, \dots, \infty\}\})$ . Applying 2.176 again we have that

$$\begin{aligned}
 \limsup_{n \rightarrow \infty} x_n &\stackrel{\text{def}}{=} \inf (\{\sup (\{x_i | i \in \{l, \dots, \infty\}\}) | l \in \{k, \dots, \infty\}\}) \\
 &\leq \inf (\{\sup (\{y_i | i \in \{l, \dots, \infty\}\}) | l \in \{k, \dots, \infty\}\}) \\
 &= \limsup_{n \rightarrow \infty} y_n
 \end{aligned}$$

□

**Proposition 17.56.** Let  $k \in \mathbb{N}_0$ ,  $\{x_i\}_{i \in \{k, \dots, \infty\}}$  be a sequence of extended reals and  $x \in \overline{\mathbb{R}}$  then we have

1. If  $\forall i \in \{k, \dots, \infty\}$  we have  $x \leq x_i$  then we have

- a.  $x \leq \liminf_{n \rightarrow \infty} x_n$

- b.  $x \leq \limsup_{n \rightarrow \infty} x_n$
2. If  $\forall i \in \{k, \dots, \infty\}$  we have  $x_i \leq x$  then we have
- $\liminf_{n \rightarrow \infty} x_n \leq x$
  - $\limsup_{n \rightarrow \infty} x_n \leq x$

**Proof.**

1.

- a. Let  $l \in \{k, \dots, \infty\}$  then  $\forall i \in \{l, \dots, \infty\} \subseteq \{k, \dots, \infty\}$  we have  $x \leq x_i$ , hence it follows that  $x \leq \inf(\{x_i | i \in \{l, \dots, \infty\}\})$ . As  $\inf(\{x_i | i \in \{l, \dots, \infty\}\}) \leq \sup(\{\inf(\{x_i | i \in \{l, \dots, \infty\}\}) | l \in \{k, \dots, \infty\}\}) = \liminf_{n \rightarrow \infty} x_n$  we have that

$$x \leq \liminf_{n \rightarrow \infty} x_n$$

- b. Let  $l \in \{k, \dots, \infty\}$  then  $\forall i \in \{l, \dots, \infty\} \subseteq \{k, \dots, \infty\}$  we have  $x \leq x_i \leq \sup(\{x_i | i \in \{l, \dots, \infty\}\})$ , so that  $x \leq \inf(\{\sup(\{x_i | i \in \{l, \dots, \infty\}\}) | l \in \{k, \dots, \infty\}\}) = \limsup_{n \rightarrow \infty} x_n$ . Hence

$$x \leq \limsup_{n \rightarrow \infty} x_n$$

2.

- a. Let  $l \in \{k, \dots, \infty\}$  then  $\forall i \in \{l, \dots, \infty\} \subseteq \{k, \dots, \infty\}$  we have  $\inf(\{x_i | i \in \{l, \dots, \infty\}\}) \leq x_i \leq x$ , so that  $\liminf_{n \rightarrow \infty} x_n = \sup(\{\inf(\{x_i | i \in \{l, \dots, n\}\}) | l \in \{k, \dots, \infty\}\}) \leq x$ . So that

$$\liminf_{n \rightarrow \infty} x_n \leq x$$

- b. Let  $l \in \{k, \dots, \infty\}$  then  $\forall i \in \{l, \dots, \infty\} \subseteq \{k, \dots, \infty\}$   $x_i \leq x$ , so it follows that  $\sup(\{x_i | i \in \{l, \dots, \infty\}\}) \leq x$ , hence  $\inf(\{\sup(\{x_i | i \in \{l, \dots, \infty\}\}) | l \in \{k, \dots, \infty\}\}) \leq x$ . Using the definition of  $\limsup_{n \rightarrow \infty} x_n$  it follows then that

$$\limsup_{n \rightarrow \infty} x_n \leq x \quad \square$$

**Theorem 17.57.** Let  $k \in \mathbb{N}_0$ ,  $\{x_i\}_{i \in \{k, \dots, \infty\}}$  be a sequence of extended reals,  $x \in \mathbb{R}$ , such that  $\{x_i + x\}_{i \in \{k, \dots, \infty\}}$  is well defined then

- $\liminf_{i \rightarrow \infty} (x_i + x) = \left( \liminf_{i \rightarrow \infty} x_i \right) + x$
- $\limsup_{i \rightarrow \infty} (x_i + x) = \left( \limsup_{i \rightarrow \infty} x_i \right) + x$

**Proof.**

1.

$$\begin{aligned} \liminf_{i \rightarrow \infty} (x_i + x) &= \sup(\{\inf(\{x_i + x | i \in \{l, \dots, \infty\}\}) | l \in \{k, \dots, \infty\}\}) \\ &\stackrel{17.30}{=} \sup(\{\inf(\{x_i | i \in \{l, \dots, \infty\}\}) + x | l \in \{k, \dots, \infty\}\}) \\ &\stackrel{17.30}{=} \sup(\{\inf(\{x_i | i \in \{l, \dots, \infty\}\}) | l \in \{k, \dots, \infty\}\}) + x \\ &= \left( \liminf_{i \rightarrow \infty} x_i \right) + x \end{aligned}$$

2.

$$\begin{aligned} \limsup_{i \rightarrow \infty} (x_i + x) &= \inf(\{\sup(\{x_i + x | i \in \{l, \dots, \infty\}\}) | l \in \{k, \dots, \infty\}\}) \\ &\stackrel{17.30}{=} \inf(\{\sup(\{x_i | i \in \{l, \dots, \infty\}\}) + x | l \in \{k, \dots, \infty\}\}) \\ &\stackrel{17.30}{=} \inf(\{\sup(\{x_i | i \in \{l, \dots, \infty\}\}) | l \in \{k, \dots, \infty\}\}) + x \\ &= \left( \limsup_{i \rightarrow \infty} x_i \right) + x \end{aligned}$$

□

**Theorem 17.58.** Let  $k \in \mathbb{N}_0$ ,  $\{x_i\}_{i \in \{k, \dots, n\}}$  be a sequence of extended reals then

$$1. \liminf_{i \rightarrow \infty} (-x_i) = -\limsup_{i \rightarrow \infty} x_i$$

$$2. \limsup_{i \rightarrow \infty} (-x_i) = -\liminf_{i \rightarrow \infty} x_i$$

**Proof.**

1.

$$\begin{aligned} \liminf_{i \rightarrow \infty} (-x_i) &= \sup (\{\inf (\{-x_i \mid i \in \{l, \dots, \infty\}\}) \mid l \in \{k, \dots, \infty\}\}) \\ &\stackrel{17.31}{=} \sup (\{-\sup (\{x_i \mid i \in \{l, \dots, \infty\}\}) \mid l \in \{k, \dots, \infty\}\}) \\ &\stackrel{17.31}{=} -\inf (\{\sup (\{x_i \mid i \in \{l, \dots, \infty\}\}) \mid l \in \{k, \dots, \infty\}\}) \\ &= -\limsup_{i \rightarrow \infty} x_i \end{aligned}$$

2.

$$\begin{aligned} \limsup_{i \rightarrow \infty} (-x_i) &\stackrel{(1)}{=} -\liminf_{i \rightarrow \infty} (-(-x_i)) \\ &= -\liminf_{i \rightarrow \infty} x_i \end{aligned}$$

□

**Lemma 17.59.** Let  $k \in \mathbb{N}_0$ ,  $\{x_i\}_{i \in \{k, \dots, \infty\}}$ ,  $\{y_i\}_{i \in \{k, \dots, \infty\}} \subseteq \mathbb{R}$  sequences of reals [so that the sums are well defined] then

1. If  $\forall i \in \{k, \dots, \infty\}$  we have  $x_i \leq x_{i+1}$  and  $y_i \leq y_{i+1}$  then

$$\sup (\{x_i \mid i \in \{k, \dots, \infty\}\}) + \sup (\{y_i \mid i \in \{k, \dots, \infty\}\}) = \sup (\{x_i + y_i \mid i \in \{k, \dots, \infty\}\})$$

2. If  $\forall i \in \{k, \dots, \infty\}$  we have  $x_{i+1} \leq x_i$  and  $y_{i+1} \leq y_i$  then

$$\inf (\{x_i \mid i \in \{k, \dots, \infty\}\}) + \inf (\{y_i \mid i \in \{k, \dots, \infty\}\}) = \inf (\{x_i + y_i \mid i \in \{k, \dots, \infty\}\})$$

**Proof.**

1. As  $\forall i \in \{k, \dots, \infty\}$  we have  $x_i \leq \sup (\{x_j \mid j \in \{k, \dots, \infty\}\})$  and  $y_i \leq \sup (\{y_j \mid j \in \{k, \dots, \infty\}\})$  it follows that  $\forall i \in \{k, \dots, \infty\}$  that  $x_i + y_i \leq \sup (\{x_j \mid j \in \{k, \dots, \infty\}\}) + \sup (\{y_j \mid j \in \{k, \dots, \infty\}\})$  hence

$$\sup (\{x_i + y_i \mid i \in \{k, \dots, \infty\}\}) \leq \sup (\{x_i \mid i \in \{k, \dots, \infty\}\}) + \sup (\{y_i \mid i \in \{k, \dots, \infty\}\}) \quad (17.24)$$

Take now  $l \in \{k, \dots, \infty\}$  then for  $i \in \{k, \dots, \infty\}$  we have either

**i** < **l**. then  $x_l + y_i \leq x_l + y_l \leq \sup (\{x_j + y_j \mid j \in \{k, \dots, \infty\}\})$  so that  $y_i \leq \sup (\{x_j + y_j \mid j \in \{k, \dots, \infty\}\}) - x_l$

**l** ≤ **i**.  $x_l + y_i \leq x_i + y_i \leq \sup (\{x_j + y_j \mid j \in \{k, \dots, \infty\}\})$  so that  $y_i \leq \sup (\{x_j + y_j \mid j \in \{k, \dots, \infty\}\}) - x_l$

as in all cases we have that  $y_i \leq \sup (\{x_j + y_j \mid j \in \{k, \dots, \infty\}\}) - x_l$  it follows that

$$\forall l \in \{k, \dots, \infty\} \text{ we have } \sup (\{y_i \mid i \in \{k, \dots, \infty\}\}) \leq \sup (\{x_j + y_j \mid j \in \{k, \dots, \infty\}\}) - x_l \quad (17.25)$$

We have as  $\{x_i\}_{i \in \{k, \dots, n\}} \subseteq \mathbb{R}$  by 17.13] only the following possibilities for  $\sup (\{y_i \mid i \in \{k, \dots, \infty\}\})$

**sup** ( $\{y_i \mid i \in \{k, \dots, \infty\}\}$ ) =  $\infty$ . then by 17.25 we have  $\sup (\{x_j + y_j \mid j \in \{k, \dots, \infty\}\}) = \infty$  so that  $\sup (\{x_i \mid i \in \{k, \dots, \infty\}\}) + \sup (\{y_i \mid i \in \{k, \dots, \infty\}\}) \leq \sup (\{x_j + y_j \mid j \in \{k, \dots, \infty\}\})$ .

$\sup(\{y_i | i \in \{k, \dots, \infty\}\}) \in \mathbb{R}$ . then by 17.25 we have  $\forall l \in \{k, \dots, \infty\}$  that  $x_l \leq \sup(\{x_j + y_j | j \in \{k, \dots, \infty\}\}) - \sup(\{y_i | i \in \{k, \dots, \infty\}\})$  proving that  $\sup(\{x_i | i \in \{k, \dots, \infty\}\}) \leq \sup(\{x_j + y_j | j \in \{k, \dots, \infty\}\}) - \sup(\{y_i | i \in \{k, \dots, \infty\}\})$  and thus  $\sup(\{x_i | i \in \{k, \dots, \infty\}\}) + \sup(\{y_i | i \in \{k, \dots, \infty\}\}) \leq \sup(\{x_j + y_j | j \in \{k, \dots, \infty\}\})$

As in all cases we have  $\sup(\{x_i | i \in \{k, \dots, \infty\}\}) + \sup(\{y_i | i \in \{k, \dots, \infty\}\}) \leq \sup(\{x_j + y_j | j \in \{k, \dots, \infty\}\})$  it follows using 17.24 that

$$\sup(\{x_i | i \in \{k, \dots, \infty\}\}) + \sup(\{y_i | i \in \{k, \dots, \infty\}\}) = \sup(\{x_j + y_j | j \in \{k, \dots, \infty\}\})$$

2. Define  $\{\bar{x}_i\}_{i \in \{k, \dots, n\}}, \{\bar{y}_i\}_{i \in \{k, \dots, \infty\}}$  by  $\bar{x}_i = -x_i$  and  $\bar{y}_i = -y_i$  then  $\forall i \in \{k, \dots, \infty\}$  we have  $\bar{x}_i \leq \bar{x}_{i+1}$  and  $\bar{y}_i \leq \bar{y}_{i+1}$ . Hence using (1) we have that  $\sup(\{\bar{x}_i | i \in \{k, \dots, \infty\}\}) + \sup(\{\bar{y}_i | i \in \{k, \dots, \infty\}\}) = \sup(\{\bar{x}_i + \bar{y}_i | i \in \{k, \dots, \infty\}\})$  or  $\sup(\{-x_i | i \in \{k, \dots, \infty\}\}) + \sup(\{-y_i | i \in \{k, \dots, \infty\}\}) = \sup(\{-(x_i + y_i) | i \in \{k, \dots, \infty\}\})$ . Using 17.31 we have then that

$$\inf(\{x_i | i \in \{k, \dots, \infty\}\}) + \inf(\{y_i | i \in \{k, \dots, \infty\}\}) = \inf(\{x_i + y_i | i \in \{k, \dots, \infty\}\})$$

□

**Theorem 17.60.** Let  $k \in \mathbb{N}_0$  then  $\liminf_{i \rightarrow \infty}$  and  $\limsup_{i \rightarrow \infty}$  have the following properties

1. If  $\{x_i\}_{i \in \{k, \dots, \infty\}} \subseteq \overline{\mathbb{R}}$  is a sequence of extended reals and  $0 \leq \alpha \in \mathbb{R}$  then
  - a.  $\liminf_{i \rightarrow \infty}(\alpha \cdot x_i) = \alpha \cdot \liminf_{i \rightarrow \infty} x_i$
  - b.  $\limsup_{i \rightarrow \infty}(\alpha \cdot x_i) = \alpha \cdot \limsup_{i \rightarrow \infty} x_i$
2. If  $\{x_i\}_{i \in \{k, \dots, n\}}, \{y_i\}_{i \in \{k, \dots, n\}} \subseteq \mathbb{R}$  be sequences of real numbers such that  $(\liminf_{i \rightarrow \infty} x_i) + (\liminf_{i \rightarrow \infty} y_i)$  is well defined then
  - a.  $(\liminf_{i \rightarrow \infty} x_i) + (\liminf_{i \rightarrow \infty} y_i) \leq \liminf_{i \rightarrow \infty} (x_i + y_i)$
  - b.  $\limsup_{i \rightarrow \infty} (x_i + y_i) \leq (\liminf_{i \rightarrow \infty} x_i) + (\liminf_{i \rightarrow \infty} y_i)$

**Proof.**

1.

$$\begin{aligned} \liminf_{i \rightarrow \infty}(\alpha_i \cdot x) &= \sup(\{\inf(\{\alpha \cdot x_i | i \in \{l, \dots, \infty\}\}) | l \in \{k, \dots, \infty\}\}) \\ &\stackrel{17.31}{=} \sup(\{\alpha \cdot \inf(\{x_i | i \in \{l, \dots, \infty\}\}) | l \in \{k, \dots, \infty\}\}) \\ &\stackrel{17.31}{=} \alpha \cdot \sup(\{\inf(\{x_i | i \in \{l, \dots, \infty\}\}) | l \in \{k, \dots, \infty\}\}) \\ &= \alpha \cdot \liminf_{i \rightarrow \infty} x_i \\ \limsup_{i \rightarrow \infty}(\alpha_i \cdot x) &= \inf(\{\sup(\{\alpha \cdot x_i | i \in \{l, \dots, \infty\}\}) | l \in \{k, \dots, \infty\}\}) \\ &\stackrel{17.31}{=} \inf(\{\alpha \cdot \sup(\{x_i | i \in \{l, \dots, \infty\}\}) | l \in \{k, \dots, \infty\}\}) \\ &\stackrel{17.31}{=} \alpha \cdot \inf(\{\sup(\{x_i | i \in \{l, \dots, \infty\}\}) | l \in \{k, \dots, \infty\}\}) \\ &= \alpha \cdot \limsup_{i \rightarrow \infty} x_i \end{aligned}$$

2.

- a. Let  $l \in \{k, \dots, \infty\}$  then we have  $\forall i \in \{l, \dots, \infty\}$  that  $\inf(\{x_j | j \in \{l, \dots, \infty\}\}) \leq x_i \in \mathbb{R}$  and  $\inf(\{y_j | j \in \{l, \dots, \infty\}\}) \leq y_i \in \mathbb{R}$  so that  $\inf(\{x_j | j \in \{l, \dots, \infty\}\}) + \inf(\{y_j | j \in \{l, \dots, \infty\}\})$  is well defined and  $\inf(\{x_j | j \in \{l, \dots, \infty\}\}) + \inf(\{y_j | j \in \{l, \dots, \infty\}\}) \leq x_i + y_i$  So  $\inf(\{x_j | j \in \{l, \dots, \infty\}\}) + \inf(\{y_j | j \in \{l, \dots, \infty\}\}) \leq \inf(\{x_j + y_j | j \in \{l, \dots, \infty\}\})$  hence we have

$$\sup(\{\inf(\{x_j | j \in \{l, \dots, \infty\}\}) + \inf(\{y_j | j \in \{l, \dots, \infty\}\}) | l \in \{k, \dots, \infty\}\}) \leq \sup(\{\inf(\{x_j + y_j | j \in \{l, \dots, \infty\}\}) | l \in \{k, \dots, \infty\}\}) = \liminf_{i \rightarrow \infty} (x_i + y_i) \quad (17.26)$$

Now as  $\forall l \in \{k, \dots, \infty\}$   $\inf(\{x_j | j \in \{l, \dots, \infty\}\}) \leq \inf(\{x_j | j \in \{l+1, \dots, \infty\}\})$  and  $\inf(\{y_j | j \in \{l, \dots, \infty\}\}) \leq \inf(\{y_j | j \in \{l+1, \dots, \infty\}\})$  we can use 17.59 giving  $\sup(\{\inf(\{x_j | j \in \{l, \dots, \infty\}\}) | l \in \{k, \dots, \infty\}\}) + \sup(\{\inf(\{y_j | j \in \{l, \dots, \infty\}\})\}) = \sup(\{\inf(\{x_j | j \in \{l, \dots, \infty\}\}) + \inf(\{y_j | j \in \{l, \dots, \infty\}\}) | l \in \{k, \dots, \infty\}\})$ . Applying this together with 17.26 and the definition of  $\liminf_{i \rightarrow \infty}$  we have finally

$$\liminf_{i \rightarrow \infty} x_i + \liminf_{i \rightarrow \infty} y_i \leq \liminf_{i \rightarrow \infty} (x_i + y_i)$$

- b. Define  $\{\bar{x}_i\}_{i \in \{k, \dots, \infty\}}$ ,  $\{\bar{y}_i\}_{i \in \{k, \dots, \infty\}}$  by  $\bar{x}_i = -x_i$ ,  $\bar{y}_i = -y_i$  then using (1) we have that  $\liminf_{i \rightarrow \infty} \bar{x}_i + \liminf_{i \rightarrow \infty} \bar{y}_i \leq \liminf_{i \rightarrow \infty} (\bar{x}_i + \bar{y}_i) \Rightarrow \liminf_{i \rightarrow \infty} (-x_i) + \liminf_{i \rightarrow \infty} (-y_i) \leq \liminf_{i \rightarrow \infty} (-x_i + y_i)$ , using 17.58 we have  $-\limsup_{i \rightarrow \infty} x_i + \left(-\limsup_{i \rightarrow \infty} y_i\right) \leq -\limsup_{i \rightarrow \infty} (x_i + y_i)$ . Hence

$$\limsup_{i \rightarrow \infty} (x_i + y_i) \leq \limsup_{i \rightarrow \infty} x_i + \limsup_{i \rightarrow \infty} y_i \quad \square$$

The following theorem shows the relation between both types of limits.

**Theorem 17.61.** Let  $k \in \mathbb{N}_0$ ,  $\{x_i\}_{i \in \{k, \dots, \infty\}} \subseteq \bar{\mathbb{R}}$  then  $\liminf_{i \rightarrow \infty} x_i \leq \limsup_{i \rightarrow \infty} x_i$

**Proof.** Fix  $n \in \{k, \dots, \infty\}$  then  $\forall m \in \{k, \dots, \infty\}$  we have either

$m \in \{k, \dots, n-1\}$ . then  $m \leq n$  so that  $\{x_i | i \in \{n, \dots, \infty\}\} \subseteq \{x_i | i \in \{m, \dots, \infty\}\}$  hence  $\inf(\{x_i | i \in \{m, \dots, \infty\}\}) \leq \inf(\{x_i | i \in \{n, \dots, \infty\}\}) \leq x_n \leq \sup(\{x_i | i \in \{n, \dots, \infty\}\})$

$m \in \{n, \dots, \infty\}$ . then  $\inf(\{x_i | i \in \{m, \dots, \infty\}\}) \leq x_n \leq \sup(\{x_i | i \in \{n, \dots, \infty\}\})$

So we have

$$\forall m \in \{k, \dots, \infty\} \text{ that } \inf(\{x_i | i \in \{m, \dots, \infty\}\}) \leq x_n \leq \sup(\{x_i | i \in \{n, \dots, \infty\}\})$$

and thus

$$\sup(\{\inf(\{x_i | i \in \{m, \dots, \infty\}\}) | m \in \{k, \dots, \infty\}\}) \leq \sup(\{x_i | i \in \{n, \dots, \infty\}\})$$

giving that

$$\sup(\{\inf(\{x_i | i \in \{m, \dots, \infty\}\}) | m \in \{k, \dots, \infty\}\}) \leq \inf(\{\sup(\{x_i | i \in \{n, \dots, \infty\}\}) | n \in \{k, \dots, \infty\}\})$$

proving

$$\liminf_{n \rightarrow \infty} x_n \leq \limsup_{n \rightarrow \infty} x_n \quad \square$$

**Theorem 17.62.** Let  $k \in \mathbb{N}_0$ ,  $\{x_i\}_{i \in \{k, \dots, n\}}$  be a sequence of **real numbers** then we have

1. If  $\lim_{i \rightarrow \infty} x_i$  exists (in  $\mathbb{R}$ ) then  $\liminf_{i \rightarrow \infty} x_i = \limsup_{i \rightarrow \infty} x_i = \lim_{i \rightarrow \infty} x_i$
2. If  $\liminf_{i \rightarrow \infty} x_i, \limsup_{i \rightarrow \infty} x_i \in \mathbb{R}$  and  $\liminf_{i \rightarrow \infty} x_i = \limsup_{i \rightarrow \infty} x_i$  then  $\lim_{i \rightarrow \infty} x_i$  exists and  $\liminf_{i \rightarrow \infty} x_i = \limsup_{i \rightarrow \infty} x_i = \lim_{i \rightarrow \infty} x_i$

**Proof.**

1. Let  $\lim_{i \rightarrow \infty} x_i = x \in \mathbb{R}$ . Assume that  $\liminf_{i \rightarrow \infty} x_i < x$  then we have either

$\liminf_{i \rightarrow \infty} x_i = -\infty$ . Assume that  $-\infty \neq \inf(\{x_i | i \in \{k, \dots, \infty\}\})$  then we have  $-\infty < \inf(\{x_i | i \in \{k, \dots, \infty\}\}) \leq \sup(\{\inf(\{x_i | i \in \{l, \dots, \infty\}\}) | l \in \{k, \dots, \infty\}\}) = \liminf_{i \rightarrow \infty} x_i = -\infty$  a contradiction. So we must have

$$\inf(\{x_i | i \in \{k, \dots, \infty\}\}) = -\infty \quad (17.27)$$

As  $\lim_{i \rightarrow \infty} x_i$  exists we have that  $\{x_i\}_{i \in \{k, \dots, \infty\}}$  is bounded [see 11.336], so there exists a  $K \in \bar{\mathbb{R}}$  such that  $\forall i \in \{k, \dots, \infty\}$  we have  $|x_i| < K$  or

$$\forall i \in \{k, \dots, \infty\} \text{ we have } -K < x_i < K \quad (17.28)$$

Using 17.27 and the fact that  $-\infty < -K$  there exists a  $i \in \{k, \dots, \infty\}$  with  $x_i < -K$  contradicting 17.28.

**lim inf  $x_i \in \mathbb{R}$ .** Take now  $\varepsilon = x - \liminf_{i \rightarrow \infty} x_i > 0$ . As  $\lim_{i \rightarrow \infty} x_i$  exists there exists a  $N \in \{k, \dots, \infty\}$  such that

$$\forall n \in \{N, \dots, \infty\} \text{ we have } |x_n - x| < \frac{\varepsilon}{2} \quad (17.29)$$

. Now

$$\begin{aligned} \inf(\{x_i | i \in \{N, \dots, \infty\}\}) &\leq \\ \sup(\{\inf(\{x_i | i \in \{l, \dots, \infty\}\}) | l \in \{k, \dots, \infty\}\}) &= \\ \liminf_{i \rightarrow \infty} x_i &= x - \varepsilon \\ &< x - \frac{\varepsilon}{2} \end{aligned}$$

so that by the definition of a infimum there exists a  $M \in \{N, \dots, \infty\}$  such that  $x_M < x - \frac{\varepsilon}{2}$ . Hence  $|x_M - x| = x - x_M > \frac{\varepsilon}{2}$  which contradicts 17.29.

As we in all cases have a contradiction. the assumption that  $\liminf_{i \rightarrow \infty} x_i < x$  must be wrong so we must have that

$$x \leq \liminf_{i \rightarrow \infty} x_i \quad (17.30)$$

Assume now that  $x < \limsup_{i \rightarrow \infty} x_i$ . Then for  $\limsup_{i \rightarrow \infty} x_i$  we have either

**lim sup  $x_i = \infty$ .** Assume that  $\sup(\{x_i | i \in \{k, \dots, \infty\}\}) \neq \infty$  then we have  $\infty = \inf(\{\sup(\{x_i | i \in \{l, \dots, \infty\}\}) | l \in \{k, \dots, \infty\}\}) \leq \sup(\{x_i | i \in \{k, \dots, \infty\}\}) < \infty$  a contradiction. So we must have that

$$\sup(\{x_i | i \in \{k, \dots, \infty\}\}) = \infty \quad (17.31)$$

As  $\lim_{i \rightarrow \infty} x_i$  exists we have that  $\{x_i\}_{i \in \{k, \dots, \infty\}}$  is bounded [see 11.336], so there exists a  $K \in \mathbb{R}$  such that  $\forall i \in \{k, \dots, \infty\}$  we have  $|x_i| < K$  or

$$\forall i \in \{k, \dots, \infty\} \text{ we have } -K < x_i < K \quad (17.32)$$

As  $K < \infty$  there exists by the definition of the supremum and 17.31 that  $\exists i \in \{k, \dots, n\}$  such that  $K < x_i$  contradicting 17.32.

**lim sup  $x_i \in \mathbb{R}$ .** Take  $\varepsilon = \limsup_{i \rightarrow \infty} x_i - x > \varepsilon$ . As  $\lim_{i \rightarrow \infty} x_i$  exists there exists a  $N \in \{k, \dots, \infty\}$  such that

$$\forall n \in \{N, \dots, \infty\} \text{ we have } |x_n - x| < \frac{\varepsilon}{2} \quad (17.33)$$

Now

$$\begin{aligned} \sup(\{x_i | i \in \{N, \dots, \infty\}\}) &\geq \\ \inf(\{\sup(\{x_i | i \in \{N, \dots, \infty\}\}) | l \in \{k, \dots, \infty\}\}) &= \\ \limsup_{i \rightarrow \infty} x_i &= x + \varepsilon \\ &> x + \varepsilon \end{aligned}$$

so using the definiton of the supremum there exists a  $M \in \{N, \dots, \infty\}$  such that  $x_M > x + \frac{\varepsilon}{2}$ . Hence  $|x_M - x| > \frac{\varepsilon}{2}$  contradicting 17.33

As we have in all cases a contradiction the assumption  $x < \limsup_{i \rightarrow \infty} x_i$  must be wrong, hence

$$\limsup_{i \rightarrow \infty} x_i \leq x \quad (17.34)$$

Finally we have  $\limsup_{i \rightarrow \infty} x_i \leq_{17.34} x \leq_{17.30} \liminf_{i \rightarrow \infty} x_i \leq_{17.61} \limsup_{i \rightarrow \infty} x_i$  proving that

$$\liminf_{i \rightarrow \infty} x_i \in \mathbb{R} \wedge \limsup_{i \rightarrow \infty} x_i \in \mathbb{R} \wedge \lim_{i \rightarrow \infty} x_i = \liminf_{i \rightarrow \infty} x_i = \limsup_{i \rightarrow \infty} x_i$$

2. Assume that  $\liminf_{i \rightarrow \infty} x_i = \limsup_{i \rightarrow \infty} x_i = x \in \mathbb{R}$ . Take  $\varepsilon > 0$  then we have  $x - \varepsilon < x = \liminf_{i \rightarrow \infty} x_i = \sup(\{\inf(\{x_i | i \in \{k, \dots, \infty\}\}) | k \in \{1, \dots, \infty\}\})$  so there exists a  $M \in \{1, \dots, \infty\}$  such that  $x - \varepsilon < \inf(\{x_i | i \in \{M, \dots, \infty\}\}) \leq x$  and thus  $\forall n \geq M$  we have  $x - \varepsilon < x_n$  so we have

$$\exists M \in \{1, \dots, \infty\} \text{ such that } \forall n \geq M \text{ we have } x - \varepsilon < x_n \quad (17.35)$$

As also  $\inf(\{\sup(\{x_i | i \in \{k, \dots, \infty\}\}) | k \in \{1, \dots, \infty\}\}) = \limsup_{i \rightarrow \infty} x_i = x < x + \varepsilon$  there exists a  $N \in \{1, \dots, \infty\}$  such that  $x \leq \sup(\{x_i | i \in \{N, \dots, \infty\}\}) < x + \varepsilon$  giving

$$\exists N \in \{1, \dots, \infty\} \text{ such that } \forall n \geq N \text{ we have } x_n < x + \varepsilon \quad (17.36)$$

So using 17.35, 17.36 we have if  $n \geq \max(N, M)$  that  $x - \varepsilon < x_n < x + \varepsilon \Rightarrow |x - x_n| < \varepsilon$  proving that  $\lim_{i \rightarrow \infty} x_i = x$ .  $\square$

Motivated by the above theorem we can extend the notation of a limit on  $\mathbb{R}$  to  $\overline{\mathbb{R}}$  as follows

**Definition 17.63.** Let  $k \in \mathbb{N}_0$ ,  $\{x_i\}_{i \in \{k, \dots, \infty\}}$  be a sequence of extended reals then  $\lim_{i \rightarrow \infty} x_i$  exists iff  $\liminf_{i \rightarrow \infty} x_i = \limsup_{i \rightarrow \infty} x_i$  and  $\lim_{i \rightarrow \infty} x_i$  is defined as  $\lim_{i \rightarrow \infty} x_i = \liminf_{i \rightarrow \infty} x_i = \limsup_{i \rightarrow \infty} x_i$ .

**Note 17.64.** By 17.52 the limit, if it exists, is independent of the value of  $k$  (as is expressed in the notation).

Using the above definition and 17.62 and proves that the limit of the real numbers is the same as the limit in the extended reals as is expressed in the following proposition. We note  $\lim_{i \rightarrow \infty} x_i$  as the limit in  $\mathbb{R}$  to clarify the distinction between the two limits.

**Proposition 17.65.** Let  $k \in \mathbb{N}_0$  and  $\{x_i\}_{i \in \{k, \dots, \infty\}}$  a sequence of real numbers then  $\lim_{i \rightarrow \infty} x_i = x$  if and only if  $\lim_{i \rightarrow \infty} x_i = x$

**Proof.** See 17.62 and the definition.  $\square$

**Remark 17.66.** To avoid excessive notation we use the convention that if we use  $\lim_{i \rightarrow \infty} x_i = x$  then we automatically assume that the limit exists and is equal to  $x$ .

**Example 17.67.** Let  $\{x_i\}_{i \in \{k, \dots, \infty\}}$  be defined by  $x_i = x \in \overline{\mathbb{R}}$  then  $\lim_{i \rightarrow \infty} x_i = x$

**Proof.**

$$\begin{aligned} \sup(\{\inf(\{x_i | i \in \{n, \dots, \infty\}\}) | n \in \{k, \dots, \infty\}\}) &= \sup(\{\inf(\{x\})\}) \\ &= \sup(\{x\}) \\ &= x \\ &= \inf(\{x\}) \\ &= \inf(\{\sup(\{x\})\}) \\ &= \inf(\{\sup(\{x_i | i \in \{n, \dots, \infty\}\}) | n \in \{k, \dots, \infty\}\}) \end{aligned}$$

proving that  $\lim_{i \rightarrow \infty} x_i = x$ .  $\square$

The limit of a sequence of extended reals is independent of a translation of the index

**Proposition 17.68.** Let  $\{x_i\}_{i \in \{k, \dots, \infty\}}$  be a sequence of extended reals and  $n \in \mathbb{N}_0$  then if  $\lim_{i \rightarrow \infty} x_i$  exists we have that  $\lim_{i \rightarrow \infty} x_{i+n}$  exists and  $\lim_{i \rightarrow \infty} x_i = \lim_{i \rightarrow \infty} x_{i+n}$

**Proof.** As  $\lim_{i \rightarrow \infty} x_i$  exists we have by definition that  $\lim_{i \rightarrow \infty} x_i = \liminf_{i \rightarrow \infty} x_i = \limsup_{i \rightarrow \infty} x_i$ . Next we have  $\liminf_{i \rightarrow \infty} x_{i+n} = \liminf_{i \rightarrow \infty} x_i = \lim_{i \rightarrow \infty} x_i = \limsup_{i \rightarrow \infty} x_i = \limsup_{i \rightarrow \infty} x_{i+n}$ .  $\square$

**Theorem 17.69.** Let  $x \in \bar{\mathbb{R}}$ ,  $\{x_i\}_{i \in \mathbb{N}_0}$  be such that  $\forall i \in \mathbb{N}_0$  we have that  $x + x_i$  is well defined and  $\lim_{i \rightarrow \infty} x_i$  exists then  $\lim_{i \rightarrow \infty} (x_i + x)$  exists and  $\lim_{i \rightarrow \infty} (x + x_i) = x + \lim_{i \rightarrow \infty} x_i$

**Proof.** By assumption we have that  $\liminf_{i \rightarrow \infty} x_i = \limsup_{i \rightarrow \infty} x_i = \lim_{i \rightarrow \infty} x_i$ . Now  $\liminf_{i \rightarrow \infty} (x + x_i) \stackrel{17.57}{=} x + \liminf_{i \rightarrow \infty} x_i = x + \lim_{i \rightarrow \infty} x_i = x + \limsup_{i \rightarrow \infty} x_i \stackrel{17.57}{=} \limsup_{i \rightarrow \infty} (x + x_i)$  proving that

$$\lim_{i \rightarrow \infty} (x + x_i) = x + \lim_{i \rightarrow \infty} x_i \quad \square$$

In the theorem 17.62 we have proved that if a sequence of reals converges in  $\mathbb{R}$  then it converges also in  $\bar{\mathbb{R}}$ . We intend now to prove that increasing (decreasing) sequences in  $\bar{\mathbb{R}}$  converges and that the limit is equal to the supremum and ibfnum of the sequence [as in  $\mathbb{R}$ ]. To do this we need the following lemma.

**Lemma 17.70.** Let  $\{x_i\}_{i \in \{n, \dots, \infty\}}$  be a increasing [or decreasing] sequence of extended real numbers then if  $k \in \{n, \dots, \infty\}$  we have  $\sup(\{x_i | i \in \{n, \dots, \infty\}\}) = \sup(\{x_i | i \in \{k, \dots, \infty\}\})$  /or  $\inf(\{x_i | i \in \{n, \dots, \infty\}\}) = \inf(\{x_i | i \in \{k, \dots, \infty\}\})$

**Proof.** Let  $k \in \{n, \dots, \infty\}$

**increasing sequence.** As  $\{x_i | i \in \{k, \dots, \infty\}\} \subseteq \{x_i | i \in \{n, \dots, \infty\}\}$  we have by 2.175 that

$$\sup(\{x_i | i \in \{k, \dots, \infty\}\}) \leq \sup(\{x_i | i \in \{n, \dots, \infty\}\})$$

Next if  $x \in \{x_i | i \in \{n, \dots, \infty\}\}$  there  $\exists m \in \{n, \dots, \infty\}$  such that  $x = x_m$  we have for  $m$  the following possibilities

$m \in \{n, \dots, k-1\}$ . then  $x = x_m \leq_{k-1 \leq k} x_k \in \{x_i | i \in \{k, \dots, \infty\}\}$

$m \in \{k, \dots, \infty\}$ . then  $x \leq x_m \in \{x_i | i \in \{k, \dots, \infty\}\}$

so that by 2.176 we have

$$\sup(\{x_i | i \in \{n, \dots, \infty\}\}) \leq \sup(\{x_i | i \in \{k, \dots, \infty\}\}).$$

**decreasing sequence.** As  $\{x_i | i \in \{k, \dots, \infty\}\} \subseteq \{x_i | i \in \{n, \dots, \infty\}\}$  we must have by 2.175

$$\inf(\{x_i | i \in \{n, \dots, \infty\}\}) \leq \inf(\{x_i | i \in \{k, \dots, \infty\}\}).$$

Next if  $x \in \{x_i | i \in \{n, \dots, \infty\}\}$  then  $\exists m \in \{n, \dots, \infty\}$  such that  $x = x_m$  and we have the following cases for  $m$

$m \in \{n, \dots, k-1\}$ . then  $x = x_m \geq x_k \in \{x_i | i \in \{k, \dots, \infty\}\}$

$m \in \{k, \dots, \infty\}$ . then  $x = x_m \in \{x_i | i \in \{k, \dots, \infty\}\}$

so by 2.176 we have

$$\inf(\{x_i | i \in \{k, \dots, \infty\}\}) \leq \inf(\{x_i | i \in \{n, \dots, \infty\}\}). \quad \square$$

**Theorem 17.71.** Let  $k \in \mathbb{N}_0$ ,  $\{x_i\}_{i \in \{k, \dots, \infty\}}$  be a increasing [or decreasing] sequence of extended reals then  $\lim_{i \rightarrow \infty} x_i$  exists and  $\sup(\{x_i | i \in \{k, \dots, \infty\}\}) = \lim_{i \rightarrow \infty} x_i$  /or  $\inf(\{x_i | i \in \{k, \dots, \infty\}\}) = \lim_{i \rightarrow \infty} x_i$

**Proof.**

**increasing sequence.** First

$$\begin{aligned} \limsup_{i \rightarrow \infty} x_i &= \inf(\{\sup(\{x_i | i \in \{n, \dots, \infty\}\}) | n \in \{k, \dots, \infty\}\}) \\ &\stackrel{17.70}{=} \inf(\{\sup(\{x_i | i \in \{k, \dots, \infty\}\}) | n \in \{k, \dots, \infty\}\}) \\ &= \inf(\{\sup(\{x_i | i \in \{k, \dots, \infty\}\})\}) \\ &= \sup(\{x_i | i \in \{k, \dots, \infty\}\}) \end{aligned} \tag{17.37}$$

Next as  $\{x_i\}_{i \in \{k, \dots, \infty\}}$  is increasing we have for  $n \in \{k, \dots, \infty\}$  that  $\forall i \in \{n, \dots, \infty\}$  that  $x_n \leq x_i$  so that  $\inf(\{x_i | i \in \{n, \dots, \infty\}\}) \leq x_n \leq \inf(\{x_i | i \in \{n, \dots, \infty\}\})$  proving that

$$\inf(\{x_i | i \in \{n, \dots, \infty\}\}) = x_n \quad (17.38)$$

$$\begin{aligned} \liminf_{i \rightarrow \infty} x_i &= \sup(\{\inf(\{x_i | i \in \{n, \dots, \infty\}\}) | n \in \{k, \dots, \infty\}\}) \\ &\stackrel{17.38}{=} \sup(\{x_n | n \in \{k, \dots, \infty\}\}) \end{aligned} \quad (17.39)$$

Using 17.37 and 17.39 we have that  $\liminf_{i \rightarrow \infty} x_i = \limsup_{i \rightarrow \infty} x_i = \sup(\{x_i | i \in \{k, \dots, \infty\}\})$  proving that

$$\lim_{i \rightarrow \infty} x_i \text{ exists and } \lim_{i \rightarrow \infty} x_i = \sup(\{x_i\})$$

**decreasing sequence.** Define  $\{\bar{x}_i\}_{i \in \{k, \dots, \infty\}}$  by  $\bar{x}_i = -x_i$  then as  $\{x_i\}_{i \in \{k, \dots, \infty\}}$  is decreasing we have that  $\{\bar{x}_i\}_{i \in \{k, \dots, \infty\}}$  is increasing. So using 17.37 and 17.39 we have that  $\limsup_{i \rightarrow \infty} \bar{x}_i = \liminf_{i \rightarrow \infty} \bar{x}_i = \sup(\{\bar{x}_i | i \in \{k, \dots, \infty\}\})$ . Applying then 17.58 and 17.31 we have  $-\liminf_{i \rightarrow \infty} x_i = -\limsup_{i \rightarrow \infty} x_i = -\inf(\{x_i | i \in \{k, \dots, \infty\}\})$  proving that  $\limsup_{i \rightarrow \infty} x_i = \liminf_{i \rightarrow \infty} x_i = \inf(\{x_i | i \in \{k, \dots, \infty\}\})$  giving

$$\lim_{i \rightarrow \infty} x_i \text{ exists and } \lim_{i \rightarrow \infty} x_i = \inf(\{x_i | i \in \{k, \dots, \infty\}\}) \quad \square$$

The above theorem motivates the following definition

**Definition 17.72.** Let  $\{x_n\}_{n \in \{k, \dots, \infty\}}$  be a sequence of extended reals then if we say that  $x_n \uparrow x$  we mean that  $\{x_n\}_{n \in \{k, \dots, \infty\}}$  is increasing and that  $\lim_{n \rightarrow \infty} x_n = x$ . Likewise  $x_n \downarrow x$  means that  $\{x_n\}_{n \in \{k, \dots, \infty\}}$  is decreasing and that  $\lim_{n \rightarrow \infty} x_n = x$ . Note that using 17.71 we have that  $x_n \uparrow x$  is equivalent with  $x = \sup(\{x_n | n \in \{k, \dots, \infty\}\})$  and  $x_n \downarrow x$  is equivalent with  $x = \inf(\{x_n | n \in \{k, \dots, \infty\}\})$

The idea of a limit in the real numbers that the sequence approaches its limit the higher the index is (as is expressed in the  $\varepsilon$  definition of a limit in the real numbers), we show now a alternative definition of the limit in  $\overline{\mathbb{R}}$  that embed this idea.

**Theorem 17.73.** Let  $\{x_i\}_{i \in \{k, \dots, \infty\}}$  be a sequence of extended reals then we have the following equivalences

1.  $\lim_{i \rightarrow \infty} x_i = \infty$  if and only  $\forall C \in \mathbb{R}_+$  there exists a  $N \in \{k, \dots, \infty\}$  such that  $\forall n \in \{N, \dots, \infty\}$  we have  $C < x_n$
2.  $\lim_{i \rightarrow \infty} x_i = -\infty$  if and only  $\forall C \in \mathbb{R}_+$  there exists a  $N \in \{k, \dots, \infty\}$  such that  $\forall n \in \{N, \dots, \infty\}$  we have  $x_n < -C$
3.  $\lim_{i \rightarrow \infty} x_i = x \in \mathbb{R}$  if and only  $\forall \varepsilon \in \mathbb{R}_+$  there exists a  $N \in \{k, \dots, \infty\}$  such that  $\forall n \in \{N, \dots, \infty\}$  we have that  $x - \varepsilon < x_n < x + \varepsilon$

**Note 17.74.** as  $x, \varepsilon \in \mathbb{R}$  we have that  $x - \varepsilon < x_n < x + \varepsilon$  is equivalent with  $x_n \in \mathbb{R} \wedge |x - x_n| < \varepsilon$

**Proof.**

1.

$\Rightarrow$ . If  $\lim_{i \rightarrow \infty} x_i = \infty$  then  $\liminf_{i \rightarrow \infty} x_i = \sup(\{\inf(\{x_i | i \in \{n, \dots, \infty\}\}) | n \in \{k, \dots, \infty\}\}) = \infty$ , using 17.14 we have then either:

$\infty \in \{\inf(\{x_i | i \in \{n, \dots, \infty\}\}) | n \in \{k, \dots, \infty\}\}$ . So there exists a  $N \in \{k, \dots, \infty\}$  such that  $\inf(\{x_i | i \in \{n, \dots, \infty\}\}) = \infty$ . Using 17.14 we have that  $\forall n \in \{N, \dots, \infty\}$  we have  $x_n = \infty$ . Hence if  $C \in \mathbb{R}_+$  then  $\forall n \in \{N, \dots, \infty\}$  we have  $C < x_n$ .

$\forall C \in \mathbb{R} \exists x \in \{\inf(\{x_i | i \in \{n, \dots, \infty\}\}) | n \in \{k, \dots, \infty\}\} \cap \mathbb{R}$  with  $C < x$ . then given  $C \in \mathbb{R}_+$  there exists a  $N \in \{k, \dots, \infty\}$  such that  $C < \inf(\{x_i | i \in \{N, \dots, \infty\}\})$ . Hence  $\forall n \in \{N, \dots, \infty\}$  we have  $C < \inf(\{x_i | i \in \{N, \dots, \infty\}\}) \leq x_n$

So in all cases we have that

$$\forall C \in \mathbb{R}_+ \text{ there } \exists N \in \{k, \dots, \infty\} \text{ such that } \forall n \geq N \text{ we have } C < x_n$$

$\Leftarrow$ . Assume that  $\liminf_{i \rightarrow \infty} x_i < \infty$ . Take  $C = \max(\liminf_{i \rightarrow \infty} x_i, 1) \in \mathbb{R}_+$  then there exists a  $N \in \{k, \dots, \infty\}$  such that  $\forall n \in \{N, \dots, \infty\}$  we have  $C < x_n$ . Hence  $C \leq \inf(\{x_i | i \in \{N, \dots, \infty\}\}) \leq \sup(\{\inf(\{x_i | i \in \{n, \dots, \infty\}\}) | n \in \{k, \dots, \infty\}\}) = \liminf_{i \rightarrow \infty} x_i$  leading to the contradiction  $\liminf_{i \rightarrow \infty} x_i < \liminf_{i \rightarrow \infty} x_i$ . Hence  $\infty \leq \liminf_{i \rightarrow \infty} x_i \leq \limsup_{i \rightarrow \infty} x_i$  proving that  $\liminf_{i \rightarrow \infty} x_i = \limsup_{i \rightarrow \infty} x_i = \infty$  or that

$$\lim_{i \rightarrow \infty} x_i = \infty$$

2.

$\Rightarrow$ . If  $\lim_{i \rightarrow \infty} x_i = -\infty$  then  $\limsup_{i \rightarrow \infty} x_i = \inf(\{\sup(\{x_i | i \in \{n, \dots, \infty\}\}) | n \in \{k, \dots, \infty\}\}) = -\infty$ , using 17.14 we have then either:

$-\infty \in \{\sup(\{x_i | i \in \{n, \dots, \infty\}\}) | n \in \{k, \dots, \infty\}\}$ . So there exists a  $N \in \{k, \dots, \infty\}$  such that  $\sup(\{x_i | i \in \{N, \dots, \infty\}\}) = -\infty$ . Using 17.14 we have that  $\forall n \in \{N, \dots, \infty\}$  we have  $x_n = \infty$ . Hence if  $C \in \mathbb{R}_+$  then  $\forall n \in \{N, \dots, \infty\}$  we have  $x_n < -C$ .

$\forall C \in \mathbb{R} \exists x \in \{\sup(\{x_i | i \in \{n, \dots, \infty\}\}) | n \in \{k, \dots, \infty\}\} \cap \mathbb{R}$  with  $x < C$ . then given  $C \in \mathbb{R}_+$  there exists a  $N \in \{k, \dots, \infty\}$  such that  $\sup(\{x_i | i \in \{N, \dots, \infty\}\}) < -C$ . Hence  $\forall n \in \{N, \dots, \infty\}$  we have  $x_n \leq \sup(\{x_i | i \in \{N, \dots, \infty\}\}) < -C$

So in all cases we have that

$$\forall C \in \mathbb{R}_+ \text{ there } \exists N \in \{k, \dots, \infty\} \text{ such that } \forall n \geq N \text{ we have } x_n < -C$$

$\Leftarrow$ . Assume that  $-\infty < \limsup_{i \rightarrow \infty} x_i$ . Take  $C = -\min(\limsup_{i \rightarrow \infty} x_i, -1) \in \mathbb{R}_+$  then there exists a  $N \in \{k, \dots, \infty\}$  such that  $\forall n \in \{N, \dots, \infty\}$  we have  $x_n < -C$ . Hence  $\inf(\{\sup(\{x_i | i \in \{n, \dots, \infty\}\}) | n \in \{k, \dots, \infty\}\}) \leq \sup(\{x_i | i \in \{N, \dots, \infty\}\}) \leq -C$  leading to the contradiction  $\limsup_{i \rightarrow \infty} x_i < \limsup_{i \rightarrow \infty} x_i$ . Hence  $\liminf_{i \rightarrow \infty} x_i \leq \limsup_{i \rightarrow \infty} x_i \leq -\infty$  proving that  $\liminf_{i \rightarrow \infty} x_i = \limsup_{i \rightarrow \infty} x_i = -\infty$  or that

$$\lim_{i \rightarrow \infty} x_i = -\infty$$

3.

$\Rightarrow$ . If  $\lim_{i \rightarrow \infty} x_i = x \in \mathbb{R}$  then  $\sup(\{\inf(\{x_i | i \in \{n, \dots, \infty\}\}) | n \in \{k, \dots, \infty\}\}) = x = \inf(\{\sup(\{x_i | i \in \{n, \dots, \infty\}\}) | n \in \{k, \dots, \infty\}\})$ . Let  $\varepsilon \in \mathbb{R}_+$  then we have

a.  $x - \varepsilon < \sup(\{\inf(\{x_i | i \in \{n, \dots, \infty\}\}) | n \in \{k, \dots, \infty\}\}) = x$  so that  $\exists N_1 \in \{k, \dots, \infty\}$  such that  $x - \varepsilon < \inf(\{x_i | i \in \{N_1, \dots, \infty\}\}) \leq x \Rightarrow \forall n \geq N_1$  we have  $x - \varepsilon < \inf(\{x_i | i \in \{N_1, \dots, \infty\}\}) \leq x_n \Rightarrow x - \varepsilon \leq x_n$

b.  $x = \inf(\{\sup(\{x_i | i \in \{n, \dots, \infty\}\}) | n \in \{k, \dots, \infty\}\}) < x + \varepsilon$  so that  $\exists N_2 \in \{k, \dots, \infty\}$  such that  $x \leq \sup(\{x_i | i \in \{N_2, \dots, \infty\}\}) < x + \varepsilon \Rightarrow \forall n \geq N_2$  we have  $x_n \leq \sup(\{x_i | i \in \{N_2, \dots, \infty\}\}) < x + \varepsilon \Rightarrow x_n < x + \varepsilon$

Using (a) en (b) we have if  $n \geq \max(N_1, N_2)$  that  $x - \varepsilon < x_n < x + \varepsilon \Rightarrow |x - x_n| < \varepsilon$

$\Leftarrow$ . Given  $\varepsilon \in \mathbb{R}_+$  there exists a  $N_\varepsilon$  such that  $\forall n \in \{N_\varepsilon, \dots, \infty\}$  we have  $x - \varepsilon < x_n < x + \varepsilon$ . Hence we have that  $\sup(\{x_i | i \in \{N_\varepsilon, \dots, \infty\}\}) \leq x + \varepsilon$  and  $x - \varepsilon \leq \inf(\{x_i | i \in \{N_\varepsilon, \dots, \infty\}\})$  from which it follows that  $\inf(\{\sup(\{x_i | i \in \{n, \dots, \infty\}\}) | n \in \{k, \dots, \infty\}\}) \leq x + \varepsilon$  and  $x - \varepsilon \leq \sup(\{\inf(\{x_i | i \in \{n, \dots, \infty\}\}) | n \in \{k, \dots, \infty\}\})$ . Using the definition of  $\liminf_{i \rightarrow \infty} x_i$  and  $\limsup_{i \rightarrow \infty} x_i$  we have:

$$\varepsilon - \infty < x - \varepsilon \leq \liminf_{i \rightarrow \infty} x_i \leq \limsup_{i \rightarrow \infty} x_i \leq x + \varepsilon < \infty \quad (17.40)$$

As  $\varepsilon$  was chosen arbitrary we have by 9.57 that  $x \leq \liminf_{i \rightarrow \infty} \leq \limsup_{i \rightarrow \infty} \leq x$  proving that  $x = \liminf_{i \rightarrow \infty} = \limsup_{i \rightarrow \infty}$ . Hence we have

$$\lim_{i \rightarrow \infty} x_i \text{ exists and } \lim_{i \rightarrow \infty} x_i = x$$

□

**Corollary 17.75.** Let  $\{x_n\}_{n \in \mathbb{N}_0}$  be defined by  $x_n = n$  then  $\lim_{n \rightarrow \infty} x_n$  exists and is equal to  $\infty$ . In other words  $\lim_{n \rightarrow \infty} n = \infty$

**Proof.** Let  $C \in \mathbb{R}_+$  then by 9.56 there exist a  $n \in \mathbb{N}_0$  such that  $C < n$  which by the previous theorem 17.73 proves that  $\lim_{n \rightarrow \infty} x_n = \infty$  □

**Corollary 17.76.** Let  $\{x_i\}_{i \in \{k, \dots, \infty\}}$  be a sequence of extended reals with  $\lim_{i \rightarrow \infty} x_i = x$  then  $\lim_{i \rightarrow \infty} |x_i| = |x|$

**Proof.** We have three cases to consider for  $x$

$x \in \mathbb{R}$ . using the previous theorem [see 17.73] we have  $\forall \varepsilon \in \mathbb{R}_+$  that  $\exists N \in \{k, \dots, \infty\}$  such that  $\forall n \geq N$  we have  $x_n \in \mathbb{R} \wedge |x - x_n| < \varepsilon \Rightarrow |x| - |x_n| \leq |x - x_n| < \varepsilon$  proving by the previous theorem again that  $\lim_{i \rightarrow \infty} |x_i| = |x|$ .

$x = \infty$ . using the previous theorem (see 17.73) we have  $\forall C \in \mathbb{R}_+$  that there exists a  $N \in \{k, \dots, \infty\}$  such that  $\forall n \geq N$  we have  $C \leq x_n \Rightarrow C \leq |x_n|$  proving by the previous theorem that  $\lim_{i \rightarrow \infty} x_n = \infty = |x|$

$x = -\infty$ . using the previous theorem [see 17.73] we have  $\forall C \in \mathbb{R}_+$  that there exists a  $N \in \{k, \dots, \infty\}$  such that  $\forall n \geq N$  we have  $x_n \leq -C \Rightarrow -x_n \geq C \Rightarrow -x_n \leq |x_n|$  proving by the previous theorem that  $\lim_{i \rightarrow \infty} x_n = -\infty = |x|$  □

We show now that the limit in  $\overline{\mathbb{R}}$  has similar properties as the limit in  $\mathbb{R}$

**Theorem 17.77.** The limit in  $\overline{\mathbb{R}}$  has the following properties

1. Let  $A \in \overline{\mathbb{R}}$ ,  $\{x_i\}_{i \in \{k, \dots, \infty\}}$  a sequence of extended reals with a limit then
  - a. If  $\forall i \in \{k, \dots, \infty\}$   $x_i \leq A$  then  $\lim_{i \rightarrow \infty} x_i \leq A$
  - b. If  $\forall i \in \{k, \dots, \infty\}$   $A \leq x_i$  then  $A \leq \lim_{i \rightarrow \infty} x_i$
2. Let  $\{x_i\}_{i \in \{k, \dots, \infty\}}$ ,  $\{y_i\}_{i \in \{k, \dots, \infty\}}$  be sequences of extended reals with a limit such that  $\forall i \in \{k, \dots, \infty\}$  we have  $x_i \leq y_i$  then  $\lim_{i \rightarrow \infty} x_i \leq \lim_{i \rightarrow \infty} y_i$
3. If  $\{x_i\}_{i \in \{k, \dots, \infty\}}$  has a limit and  $\alpha \in \mathbb{R}$  then  $\{\alpha \cdot x_i\}_{i \in \{k, \dots, \infty\}}$  has a limit and  $\lim_{i \rightarrow \infty} (\alpha \cdot x_i) = \alpha \cdot \lim_{i \rightarrow \infty} x_i$
4. Let  $\{x_i\}_{i \in \{k, \dots, \infty\}} \subseteq \overline{\mathbb{R}}$  and  $\{y_i\}_{i \in \{k, \dots, \infty\}} \subseteq \overline{\mathbb{R}}$  be sequences of extended reals with limits such that  $(\lim_{i \rightarrow \infty} x_i) + (\lim_{i \rightarrow \infty} y_i)$  is well defined and there exists a  $n \in \{k, \dots, \infty\}$  such that  $\forall i \in \{n, \dots, \infty\}$  we have that  $x_i + y_i$  is well defined then for  $\{x_i + y_i\}_{i \in \{n, \dots, \infty\}}$  we have  $\lim_{i \rightarrow \infty} (x_i + y_i) = (\lim_{i \rightarrow \infty} x_i) + (\lim_{i \rightarrow \infty} y_i)$
5. Let  $\{x_i\}_{i \in \{k, \dots, \infty\}} \subseteq \overline{\mathbb{R}}$  and  $\{y_i\}_{i \in \{k, \dots, \infty\}} \subseteq \overline{\mathbb{R}}$  be sequences of extended reals with limits such that for  $x = \lim_{i \rightarrow \infty} x_i$  and  $y = \lim_{i \rightarrow \infty} y_i$  we have if  $x = 0$  then  $y \neq \infty, -\infty$  and if  $y = 0$  then  $x \neq \infty, -\infty$  then  $\lim_{i \rightarrow \infty} (x_i \cdot y_i) = (\lim_{i \rightarrow \infty} x_i) \cdot (\lim_{i \rightarrow \infty} y_i)$
6. Let  $\{x_i\}_{i \in \{k, \dots, \infty\}}$ ,  $\{y_i\}_{i \in \{k, \dots, \infty\}}$  be sequences of extended reals with limits then we have
  - a.  $\lim_{i \rightarrow \infty} (\min(x_i, y_i)) = \min(\lim_{i \rightarrow \infty} x_i, \lim_{i \rightarrow \infty} y_i)$

$$b. \lim_{i \rightarrow \infty} (\max(x_i, y_i)) = \max \left( \lim_{i \rightarrow \infty} x_i, \lim_{i \rightarrow \infty} y_i \right)$$

**Proof.**

1.

- a. If  $\forall i \in \{k, \dots, \infty\}$  we have  $x_i \leq A$  then  $\sup(\{x_i | i \in \{k, \dots, \infty\}\}) \leq A \Rightarrow \forall n \in \{k, \dots, \infty\}$  we have  $\sup(\{x_i | i \in \{n, \dots, \infty\}\}) \leq \sup(\{x_i | i \in \{k, \dots, \infty\}\}) \leq A$  and thus  $\inf(\{\sup(\{x_i | i \in \{n, \dots, \infty\}\}) | n \in \{k, \dots, \infty\}\}) \leq A \stackrel{\text{definition of limit}}{\Rightarrow} \lim_{i \rightarrow \infty} x_i \leq A$
- b. If  $\forall i \in \{k, \dots, \infty\}$  we have  $A \leq x_i$  then  $A \leq \inf(\{x_i | i \in \{k, \dots, \infty\}\}) \Rightarrow \forall n \in \{k, \dots, \infty\}$  we have  $A \leq \inf(\{x_i | i \in \{n, \dots, \infty\}\}) \leq \inf(\{x_i | i \in \{n, \dots, \infty\}\})$  so that  $A \leq \sup(\{\inf(\{x_i | i \in \{n, \dots, \infty\}\}) | n \in \{k, \dots, \infty\}\}) \stackrel{\text{definition of limit}}{\Rightarrow} A \leq \lim_{i \rightarrow \infty} x_i$ .

2. As  $\forall i \in \{k, \dots, \infty\}$  we have  $x_i \leq y_i$ , we have using 17.55 that  $\limsup_{n \rightarrow \infty} x_n \leq \limsup_{n \rightarrow \infty} y_n$   
 $\Rightarrow \lim_{i \rightarrow \infty} x_i \leq \lim_{i \rightarrow \infty} y_i$

3. Consider the following possible cases for  $\alpha \in \mathbb{R}$

**$\alpha = 0$ .** then  $\forall i \in \{k, \dots, \infty\}$  we have  $\alpha \cdot x_i = 0$  so that  $\forall n \in \{k, \dots, \infty\}$  we have  $\sup(\{\alpha \cdot x_i | i \in \{n, \dots, \infty\}\}) = \inf(\{\alpha \cdot x_i | i \in \{n, \dots, \infty\}\}) = 0$  so that  $\inf(\{\sup(\{\alpha \cdot x_i | i \in \{n, \dots, \infty\}\}) | n \in \{k, \dots, \infty\}\}) = 0 = \sup(\{\inf(\{\alpha \cdot x_i | i \in \{n, \dots, \infty\}\}) | n \in \{k, \dots, \infty\}\}) \Rightarrow \lim_{i \rightarrow \infty} (\alpha \cdot x_i) = 0 = \alpha \cdot \lim_{i \rightarrow \infty} x_i$  proving the theorem in this case.

**$\alpha \in \mathbb{R} \setminus \{0\}$ .** then we have the following possibilities for  $\alpha$

**$\alpha > 0$ .** consider now the following possibilities for  $\lim_{i \rightarrow \infty} x_i$

**$\lim_{i \rightarrow \infty} x_i = \infty$ .** Let  $C \in \mathbb{R}_+$  then using 17.73 there exists a  $N \in \{k, \dots, \infty\}$  such that  $\forall n \geq N$  we have  $x_n > \frac{C}{\alpha} \Rightarrow \alpha \cdot x_n > C$  proving that  $\lim_{i \rightarrow \infty} (\alpha \cdot x_i) = \infty = \alpha \cdot \infty = \alpha \cdot \lim_{i \rightarrow \infty} x_i$  and thus  $\lim_{i \rightarrow \infty} (\alpha \cdot x_i) = \alpha \cdot \lim_{i \rightarrow \infty} x_i$

**$\lim_{i \rightarrow \infty} x_i = -\infty$ .** Let now  $C \in \mathbb{R}_+$  then there exists a  $N \in \{k, \dots, \infty\}$  such that  $\forall n \geq N$  we have  $x_n < -\frac{C}{\alpha} \Rightarrow \alpha \cdot x_n < -C \Rightarrow \lim_{i \rightarrow \infty} (\alpha \cdot x_i) = -\infty = \alpha \cdot (-\infty) = \alpha \cdot \lim_{i \rightarrow \infty} x_i$  and thus  $\lim_{i \rightarrow \infty} (\alpha \cdot x_i) = \alpha \cdot \lim_{i \rightarrow \infty} x_i$

**$\lim_{i \rightarrow \infty} x_i \in \mathbb{R}$ .** Take  $x = \lim_{i \rightarrow \infty} x_i$ , let now  $\varepsilon \in \mathbb{R}_+$  then there exists a  $N \in \{k, \dots, \infty\}$  such that  $\forall n \geq N$  we have  $|x - x_n| < \frac{\varepsilon}{\alpha} \Rightarrow x - \frac{\varepsilon}{\alpha} < x_n < x + \frac{\varepsilon}{\alpha} \Rightarrow \alpha \cdot x - \varepsilon < \alpha \cdot x_n < \alpha \cdot x + \varepsilon \Rightarrow |\alpha \cdot x - \alpha \cdot x_n| < \varepsilon$  proving that  $\lim_{i \rightarrow \infty} (\alpha \cdot x_i) = \alpha \cdot x = \alpha \cdot \lim_{i \rightarrow \infty} x_i$

**$\alpha < 0$ .** consider now the following possibilities for  $\lim_{i \rightarrow \infty} x_i$

**$\lim_{i \rightarrow \infty} x_i = \infty$ .** Let now  $C \in \mathbb{R}_+$  then there exists a  $N \in \{k, \dots, \infty\}$  such that  $\forall n \geq N$  we have  $x_n > \frac{C}{-\alpha} \Rightarrow \alpha \cdot x_n < -C \Rightarrow \lim_{i \rightarrow \infty} (\alpha \cdot x_i) = -\infty = \alpha \cdot \infty = \alpha \cdot \lim_{i \rightarrow \infty} x_i$  proving that  $\lim_{i \rightarrow \infty} (\alpha \cdot x_i) = \alpha \cdot \lim_{i \rightarrow \infty} x_i$

**$\lim_{i \rightarrow \infty} x_i = -\infty$ .** Let now  $C \in \mathbb{R}_+$  then there exists a  $N \in \{k, \dots, \infty\}$  such that  $\forall n \geq N$  we have  $x_n < \frac{C}{\alpha} \Rightarrow C < \alpha \cdot x_n \Rightarrow \lim_{i \rightarrow \infty} x_i = \infty = \alpha \cdot (-\infty) = \alpha \cdot \lim_{i \rightarrow \infty} x_i$  proving that  $\lim_{i \rightarrow \infty} (\alpha \cdot x_i) = \alpha \cdot \lim_{i \rightarrow \infty} x_i$

**$\lim_{i \rightarrow \infty} x_i \in \mathbb{R}$ .** Take  $x = \lim_{i \rightarrow \infty} x_i$  and let  $\varepsilon > 0$  then there exists a  $N \in \{k, \dots, \infty\}$  such that  $|x - x_n| < \frac{\varepsilon}{-\alpha} \Rightarrow x + \frac{\varepsilon}{\alpha} < x_n < x - \frac{\varepsilon}{\alpha} \Rightarrow \alpha \cdot x - \varepsilon < \alpha \cdot x_n < \alpha \cdot x + \varepsilon \Rightarrow |\alpha \cdot x - \alpha \cdot x_n| < \varepsilon$  proving that  $\lim_{i \rightarrow \infty} (\alpha \cdot x_i) = \alpha \cdot x = \alpha \cdot \lim_{i \rightarrow \infty} x_i$

So in all cases we have that  $\lim_{i \rightarrow \infty} (\alpha \cdot x_i) = \alpha \cdot \lim_{i \rightarrow \infty} x_i$

4. We have that  $\left( \lim_{i \rightarrow \infty} x_i \right) + \left( \lim_{i \rightarrow \infty} y_i \right)$  is only well defined in the following cases

**$\lim_{i \rightarrow \infty} x_i = \infty = \lim_{i \rightarrow \infty} y_i$ .** then there exists  $N_1, N_2 \in \{k, \dots, \infty\}$  such that  $\forall i \geq N_1$  we have  $1 < x_i$  and  $\forall i \geq N_2$  we have  $1 < y_i$  so if  $n = \max(N_1, N_2)$  we have  $\forall i \geq n$  that  $x_i + y_i$  is well defined. Let now  $C \in \mathbb{R}_+$  then there exists  $M_1, M_2 \in \{k, \dots, \infty\}$  such that  $\forall i \geq M_1$  we have  $x_i > \frac{C}{2}$  and  $\forall i \geq M_2$  we have  $y_i > \frac{C}{2}$  so that if  $n \geq \max(n, M_1, M_2)$  we have  $x_i + y_i > \frac{C}{2} + \frac{C}{2} = C$  proving that for  $\{x_i + y_i\}_{i \in \{n, \dots, \infty\}}$  we have  $\lim_{i \rightarrow \infty} (x_i + y_i) = \infty = \infty + \infty = \left( \lim_{i \rightarrow \infty} x_i \right) + \left( \lim_{i \rightarrow \infty} y_i \right)$

**$\lim_{i \rightarrow \infty} x_i = -\infty = \lim_{i \rightarrow \infty} y_i$ .** then there exists  $N_1, N_2 \in \{k, \dots, \infty\}$  such that  $\forall i \geq N_1$  we have  $x_i < -1$  and  $\forall i \geq N_2$  we have  $y_i < -1$  so if  $n = \max(N_1, N_2)$  we have  $\forall i \geq n$  that  $x_i + y_i$  is well defined. Let now  $C \in \mathbb{R}_+$  then there exists  $M_1, M_2 \in \{k, \dots, \infty\}$  such that  $\forall i \geq M_1$  we have  $x_i < \frac{-C}{2}$  and  $\forall i \geq M_2$  we have  $y_i < \frac{-C}{2}$  so that if  $n \geq \max(n, M_1, M_2)$  we have  $x_i + y_i < \frac{-C}{2} + \frac{-C}{2} = -C$  proving that for  $\{x_i + y_i\}_{i \in \{n, \dots, \infty\}}$  we have  $\lim_{i \rightarrow \infty} (x_i + y_i) = -\infty = -\infty + (-\infty) = \left( \lim_{i \rightarrow \infty} x_i \right) + \left( \lim_{i \rightarrow \infty} y_i \right)$

**$\lim_{i \rightarrow \infty} x_i = \infty \wedge \lim_{i \rightarrow \infty} y_i \in \mathbb{R}$ .** Take  $y = \lim_{i \rightarrow \infty} y_i$  then there exists a  $N_1 \in \{k, \dots, \infty\}$  such that  $\forall i \geq N_1$  we have  $|y - y_i| < 1 \Rightarrow -\infty < y - 1 < y_i$ . Also there exists a  $N_2 \in \{k, \dots, \infty\}$  such that  $\forall i \geq N_2$  we have  $x_i > 1 > -\infty$ . So if  $n = \max(N_1, N_2)$  then we have  $\forall i \geq n$  that  $x_i + y_i$  is well defined and  $y_i > y - 1$ . Take now  $C \in \mathbb{R}_+$  then there exists a  $M \in \{k, \dots, n\}$  such that  $\forall i \geq M$  we have  $x_i > \max(C - (y - 1), 1)$  then if  $i \geq N = \max(n, M)$  we have that  $x_i > C - (y - 1)$  and  $y_i > y - 1$  so that  $x_i + y_i > C$  proving that  $\lim_{i \rightarrow \infty} (x_i + y_i) = \infty = \infty + y = \left( \lim_{i \rightarrow \infty} x_i \right) + \left( \lim_{i \rightarrow \infty} y_i \right)$

**$\lim_{i \rightarrow \infty} x_i = -\infty \wedge \lim_{i \rightarrow \infty} y_i \in \mathbb{R}$ .** Take  $y = \lim_{i \rightarrow \infty} y_i$  then there exists a  $N_1 \in \{k, \dots, \infty\}$  such that  $\forall i \geq N_1$  we have  $|y - y_i| < 1 \Rightarrow y_i < y + 1 < \infty$ . Also there exists a  $N_2 \in \{k, \dots, \infty\}$  such that  $\forall i \geq N_2$  we have  $x_i < -1 < \infty$ . So if  $n = \max(N_1, N_2)$  then we have  $\forall i \geq n$  that  $x_i + y_i$  is well defined and  $y_i < y + 1$ . Take now  $C \in \mathbb{R}_+$  then there exists a  $M \in \{k, \dots, n\}$  such that  $\forall i \geq M$  we have  $x_i < \min(-C - (y + 1), -1)$  then if  $i \geq N = \max(n, M)$  we have that  $x_i < -C - (y + 1)$  and  $y_i < y + 1$  so that  $x_i + y_i < -C$  proving that  $\lim_{i \rightarrow \infty} (x_i + y_i) = -\infty = -\infty + y = \left( \lim_{i \rightarrow \infty} x_i \right) + \left( \lim_{i \rightarrow \infty} y_i \right)$

**$\lim_{i \rightarrow \infty} x_i \in \mathbb{R} \wedge \lim_{i \rightarrow \infty} y_i = \infty$ .** This is equivalent with the case  $\lim_{i \rightarrow \infty} x_i = \infty \wedge \lim_{i \rightarrow \infty} y_i \in \mathbb{R}$  if we interchange  $\{x_i\}_{i \in \{k, \dots, \infty\}}$  and  $\{y_i\}_{i \in \{k, \dots, \infty\}}$

**$\lim_{i \rightarrow \infty} x_i \in \mathbb{R} \wedge \lim_{i \rightarrow \infty} y_i = -\infty$ .** This is equivalent with the case  $\lim_{i \rightarrow \infty} x_i = -\infty \wedge \lim_{i \rightarrow \infty} y_i \in \mathbb{R}$  if we interchange  $\{x_i\}_{i \in \{k, \dots, \infty\}}$  and  $\{y_i\}_{i \in \{k, \dots, \infty\}}$

**$\lim_{i \rightarrow \infty} x_i \in \mathbb{R} \wedge \lim_{i \rightarrow \infty} y_i \in \mathbb{R}$ .** Let  $\lim_{i \rightarrow \infty} x_i = x$  and  $\lim_{i \rightarrow \infty} y_i = y$ , then there exists a  $N_1, N_2 \in \{k, \dots, \infty\}$  such that if  $i \geq N_1$  we have  $|x_i - x| < 1 \Rightarrow x - 1 < x_i < x + 1 \Rightarrow x_i \in \mathbb{R}$  and if  $i \geq N_2$  we have  $|y - y_i| < 1 \Rightarrow y - 1 < y_i < y + 1 \Rightarrow y_i \in \mathbb{R}$ . Take  $n = \max(N_1, N_2)$  then if  $i \geq n$  we have that  $x_i + y_i$  is well defined. Let now  $\varepsilon \in \mathbb{R}_+$  then there exists  $M_1, M_2 \in \{k, \dots, \infty\}$  such that  $\forall i \geq M_1$  we have  $|x - x_i| < \frac{\varepsilon}{2} \Rightarrow x - \frac{\varepsilon}{2} < x_i < x + \frac{\varepsilon}{2}$  and  $\forall i \geq M_2$  we have  $|y - y_i| < \frac{\varepsilon}{2} \Rightarrow y - \frac{\varepsilon}{2} < y_i < y + \frac{\varepsilon}{2}$ . So if  $i \geq \max(M_1, M_2, n)$  we have  $(x - \frac{\varepsilon}{2}) + (y - \frac{\varepsilon}{2}) < x_i + y_i < (x + \frac{\varepsilon}{2}) + (y + \frac{\varepsilon}{2}) \Rightarrow (x + y) - \varepsilon < x_i + y_i < (x + y) + \varepsilon \Rightarrow |(x + y) - (x_i + y_i)| < \varepsilon$  proving that  $\lim_{i \rightarrow \infty} (x_i + y_i) = x + y = \left( \lim_{i \rightarrow \infty} x_i \right) + \left( \lim_{i \rightarrow \infty} y_i \right)$

So in all cases we have that  $\exists n \in \{k, \dots, \infty\}$  such that  $\forall i \geq n$  we have  $x_i + y_i$  is well defined, and for  $\{x_i + y_i\}_{i \in \{k, \dots, \infty\}}$  we have  $\lim_{i \rightarrow \infty} (x_i + y_i) = \left( \lim_{i \rightarrow \infty} x_i \right) + \left( \lim_{i \rightarrow \infty} y_i \right)$ .

5. Let  $x = \lim_{i \rightarrow \infty} x_i$  and  $y = \lim_{i \rightarrow \infty} y_i$  we have if  $x = 0$  then  $y \neq \infty, -\infty$  and if  $y = 0$  then  $x \neq \infty, -\infty$  so we must consider the following remaining cases:

$x = \infty \wedge y = \infty$ . Let  $C \in \mathbb{R}_+$  then there exists  $N_1, N_2 \in \{k, \dots, \infty\}$  such that  $\forall i \geq N_1$ ,  $\forall j \geq N_2$  we have  $1 \leq x_i$  and  $C \leq y_j$ . Take  $n = \max(N_1, N_2)$  then if  $i \geq n$  we have  $1 \leq x_i \wedge C \leq y_i \Rightarrow C \leq C \cdot x_i \wedge C \cdot x_i \leq x_i \cdot y_i \Rightarrow C \leq x_i \cdot y_i$ . This proves that  $\lim_{i \rightarrow \infty} (x_i \cdot y_i) = \infty = \infty \cdot \infty = (\lim_{i \rightarrow \infty} x_i) \cdot (\lim_{i \rightarrow \infty} y_i)$

$x = \infty \wedge y = -\infty$ . Let  $C \in \mathbb{R}_+$  then there exists  $N_1, N_2 \in \{k, \dots, \infty\}$  such that  $\forall i \geq N_1$ ,  $\forall j \geq N_2$  we have  $1 \leq x_i$  and  $y_j \leq -C$ . Take  $n = \max(N_1, N_2)$  then if  $i \geq n$  we have  $1 \leq x_i \wedge y_i \leq -C \Rightarrow C \leq C \cdot x_i \wedge y_i \cdot x_i \leq -C \cdot x_i \Rightarrow -C \cdot x_i \leq -C \wedge y_i \cdot x_i \leq -C \cdot x_i \Rightarrow x_i \cdot y_i \leq -C$ . This proves that  $\lim_{i \rightarrow \infty} (x_i \cdot y_i) = -\infty = \infty \cdot -\infty = (\lim_{i \rightarrow \infty} x_i) \cdot (\lim_{i \rightarrow \infty} y_i)$

$x = \infty \wedge y \in \mathbb{R} \setminus \{0\}$ . We have then for  $y$  two possibilities:

$0 < y$ . Let  $C \in \mathbb{R}_+$  take  $\varepsilon = \frac{y}{2} > 0$  then  $\exists N_1, N_2 \in \{k, \dots, \infty\}$  such that  $\forall i \geq N_1$ ,  $\forall j \geq N_2$  we have  $y - \varepsilon < y_i < y + \varepsilon \wedge \frac{2 \cdot C}{y} \leq x_i$ . Take  $n = \max(N_1, N_2)$  then if  $n \leq i$  we have  $\frac{y}{2} < y_i \wedge \frac{2 \cdot C}{y} \leq x_i \xrightarrow{0 < \frac{2 \cdot C}{y} \wedge 0 < \frac{y}{2} < y_i} \left(\frac{2 \cdot C}{y}\right) \cdot \frac{y}{2} \leq \frac{2 \cdot C}{y} y_i \wedge \frac{2 \cdot C}{y} \cdot y_i \leq x_i \cdot y_i \Rightarrow C \leq \frac{2 \cdot C}{y} \cdot y_i \wedge \frac{2 \cdot C}{y} \cdot y_i \leq x_i \cdot y_i \Rightarrow C \leq x_i \cdot y_i$ . So we have  $\lim_{i \rightarrow \infty} (x_i \cdot y_i) = \infty = x \cdot \infty = (\lim_{i \rightarrow \infty} x_i) \cdot (\lim_{i \rightarrow \infty} y_i)$

$y < 0$ . Let  $\{y'_i\}_{i \in \{k, \dots, \infty\}}$  defined by  $y'_i = -y_i$  then  $\lim_{i \rightarrow \infty} y'_i = \lim_{i \rightarrow \infty} (-y_i) = -y > 0$  so by the previous case we have that  $-\lim_{i \rightarrow \infty} (x_i \cdot y_i) = \lim_{i \rightarrow \infty} (x_i \cdot (-y_i)) = (\lim_{i \rightarrow \infty} x_i) \cdot (\lim_{i \rightarrow \infty} (-y_i)) = -(\lim_{i \rightarrow \infty} x_i) \cdot (\lim_{i \rightarrow \infty} y_i)$  proving that  $\lim_{i \rightarrow \infty} (x_i \cdot y_i) = (\lim_{i \rightarrow \infty} x_i) \cdot (\lim_{i \rightarrow \infty} y_i)$

$x = -\infty \wedge y = \infty$ . this is the same as  $x = \infty \wedge y = -\infty$  with  $x$  and  $y$  exchanged.

$x = -\infty \wedge y = -\infty$ . Define  $\{x'_i\}_{i \in \{k, \dots, \infty\}}$ ,  $\{y'_i\}_{i \in \{k, \dots, \infty\}}$  by  $x'_i = -x_i \wedge y'_i = -y_i$  then we have  $\lim_{i \rightarrow \infty} x'_i = -\lim_{i \rightarrow \infty} x_i = -x > 0 \wedge \lim_{i \rightarrow \infty} y'_i = -\lim_{i \rightarrow \infty} y_i = -y > 0$ . As we have already proved this case we have  $\lim_{i \rightarrow \infty} (x_i \cdot y_i) = \lim_{i \rightarrow \infty} (-x_i \cdot (-y_i)) = (\lim_{i \rightarrow \infty} -x_i) \cdot (\lim_{i \rightarrow \infty} (-y_i)) = (\lim_{i \rightarrow \infty} x_i) \cdot (\lim_{i \rightarrow \infty} y_i)$  proving that  $\lim_{i \rightarrow \infty} (x_i \cdot y_i) = (\lim_{i \rightarrow \infty} x_i) \cdot (\lim_{i \rightarrow \infty} y_i)$

$x = -\infty \wedge y \in \mathbb{R} \setminus \{0\}$ . As  $\lim_{i \rightarrow \infty} (-x_i) = \infty$  we can use the case  $x = \infty \wedge y \in \mathbb{R} \setminus \{0\}$  giving  $\lim_{i \rightarrow \infty} ((-x_i) \cdot y_i) = (\lim_{i \rightarrow \infty} (-x_i)) \cdot \lim_{i \rightarrow \infty} y_i \Rightarrow -(\lim_{i \rightarrow \infty} (x_i \cdot y_i)) = -(\lim_{i \rightarrow \infty} x_i) \cdot (\lim_{i \rightarrow \infty} y_i) \Rightarrow \lim_{i \rightarrow \infty} (x_i \cdot y_i) = (\lim_{i \rightarrow \infty} x_i) \cdot (\lim_{i \rightarrow \infty} y_i)$

$x \in \mathbb{R} \setminus \{0\} \wedge y = \infty$ . this is the same as  $x = \infty \wedge y \in \mathbb{R} \setminus \{0\}$  with  $x, y$  interchanged.

$x \in \mathbb{R} \setminus \{0\} \wedge y = -\infty$ . this is the same as  $x = -\infty \wedge y \in \mathbb{R} \setminus \{0\}$  with  $x, y$  interchanged.

$x \in \mathbb{R} \wedge y \in \mathbb{R}$ . Take  $\varepsilon > 0$ . First there exists a  $N_1 \in \{k, \dots, \infty\}$  such that  $\forall i \geq N_1$  we have  $|x_i - x| < 1 \Rightarrow |x_i| \leq |x_i - x| + |x| < 1 + |x|$ . Find then  $N_2 \in \{k, \dots, \infty\}$  such that  $\forall i \geq N_2$  we have  $|x_i - x| < \frac{\varepsilon}{2 \cdot (1 + |y|)}$  and a  $N_3 \in \{k, \dots, \infty\}$  such that  $|y_i - y| < \frac{\varepsilon}{2 \cdot (1 + |x|)}$ . Take now  $n = \max(N_1, N_2, N_3)$  then if  $i \geq n$  we have  $|x_i \cdot y_i - x \cdot y| = |x_i \cdot y_i - x_i \cdot y + x_i \cdot y - x \cdot y| = |x_i \cdot (y_i - y) + (x_i - x) \cdot y| \leq |x_i| \cdot |y_i - y| + |y| \cdot |x_i - x| < (1 + |x|) \cdot \frac{\varepsilon}{2 \cdot (1 + |x|)} + \frac{\varepsilon}{2 \cdot (1 + |y|)} \cdot |y| = \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$  proving that  $\lim_{i \rightarrow \infty} (x_i \cdot y_i) = x \cdot y = (\lim_{i \rightarrow \infty} x_i) \cdot (\lim_{i \rightarrow \infty} y_i)$ .

6. Let  $x = \lim_{i \rightarrow \infty} x_i$  and  $y = \lim_{i \rightarrow \infty} y_i$  then we have the following cases to consider for  $x, y$

$x = \infty \wedge y = \infty$ . then given  $C \in \mathbb{R}_+$  there exists  $N_1, N_2 \in \{k, \dots, \infty\}$  such that  $\forall i \geq N_1, \forall j \geq N_2$  we have  $C \leq x_i, C \leq y_j$ . Take  $n = \max(N_1, N_2)$  then  $\forall i \geq n$  we have  $C \leq \max(x_i, y_i)$  and  $C \leq \min(x_i, y_i)$  proving that

$$\text{a. } \lim_{i \rightarrow \infty} (\min(x_i, y_i)) = \infty = \min(\infty, \infty) = \min\left(\lim_{i \rightarrow \infty} x_i, \lim_{i \rightarrow \infty} y_i\right)$$

$$\text{b. } \lim_{i \rightarrow \infty} (\max(x_i, y_i)) = \infty = \max(\infty, \infty) = \max\left(\lim_{i \rightarrow \infty} x_i, \lim_{i \rightarrow \infty} y_i\right)$$

$x = \infty \wedge y = -\infty$ . then given  $C \in \mathbb{R}_+$  there exists  $N_1, N_2 \in \{k, \dots, \infty\}$  such that  $\forall i \geq N_1, \forall j \geq N_2$  we have  $C \leq x_i, y_j \leq -C$ . Take  $N = \max(N_1, N_2)$  then  $\forall i \in \mathbb{N}$  we have  $C \leq x_i \leq \max(x_i, y_i)$  and  $\min(x_i, y_i) \leq y_i \leq -C$  proving that

$$\text{a. } \lim_{i \rightarrow \infty} (\min(x_i, y_i)) = -\infty = \min(\infty, -\infty) = \min\left(\lim_{i \rightarrow \infty} x_i, \lim_{i \rightarrow \infty} y_i\right)$$

$$\text{b. } \lim_{i \rightarrow \infty} (\max(x_i, y_i)) = \infty = \max(\infty, -\infty) = \max\left(\lim_{i \rightarrow \infty} x_i, \lim_{i \rightarrow \infty} y_i\right)$$

$x = \infty \wedge y \in \mathbb{R}$ . then given  $C \in \mathbb{R}_+$  and  $\varepsilon > 0$  there exists a  $N_1, N_2 \in \{k, \dots, \infty\}$  such that  $\forall i \geq N_1, \forall j \geq N_2$  we have  $\max(C, |y| + \varepsilon) \leq x_i \wedge |y - y_i| < \varepsilon$ . If we take  $N = \max(N_1, N_2)$  then we have  $\forall i \geq N$  that  $C \leq x_i \wedge |y - y_i| < \varepsilon \wedge y < y + \varepsilon \leq |y| + \varepsilon \leq x_i$ . Hence

$$\text{a. } |\min(x_i, y_i) - x| = |y_i - x| < \varepsilon \text{ proving that } \lim_{i \rightarrow \infty} (\min(x_i, y_i)) = y = \min(\infty, y) = \min\left(\lim_{i \rightarrow \infty} x_i, \lim_{i \rightarrow \infty} y_i\right)$$

$$\text{b. } C \leq x_i = \max(x_i, y_i) \text{ proving that } \lim_{i \rightarrow \infty} (\max(x_i, y_i)) = \infty = \max(\infty, y) = \max\left(\lim_{i \rightarrow \infty} x_i, \lim_{i \rightarrow \infty} y_i\right)$$

$x = -\infty \wedge y = \infty$ . this reduces to the case  $x = \infty \wedge y = -\infty$  if we interchange  $x$  and  $y$

$x = -\infty \wedge y = -\infty$ . then given  $C \in \mathbb{R}_+$  there exists  $N_1, N_2 \in \{k, \dots, \infty\}$  such that  $\forall i \geq N_1, \forall j \geq N_2$  we have  $x_i \leq -C, y_j \leq -C$ . Take  $N = \max(N_1, N_2)$  then  $\forall i \geq N$  we have  $\max(x_i, y_i) \leq -C$  and  $\min(x_i, y_i) \leq -C$  proving that

$$\text{a. } \lim_{i \rightarrow \infty} (\min(x_i, y_i)) = -\infty = \min(-\infty, -\infty) = \min\left(\lim_{i \rightarrow \infty} x_i, \lim_{i \rightarrow \infty} y_i\right)$$

$$\text{b. } \lim_{i \rightarrow \infty} (\max(x_i, y_i)) = -\infty = \max(-\infty, -\infty) = \max\left(\lim_{i \rightarrow \infty} x_i, \lim_{i \rightarrow \infty} y_i\right)$$

$x = -\infty \wedge y \in \mathbb{R}$ . then given  $C \in \mathbb{R}_+$  and  $\varepsilon > 0$  there exists  $N_1, N_2 \in \{k, \dots, \infty\}$  such that  $\forall i \geq N_1, \forall j \geq N_2$  we have  $x_i \leq -\max(C, -y + \varepsilon)$  and  $|y - y_i| < \varepsilon$ . Take  $N = \max(N_1, N_2)$  then we have  $\forall i \geq N$  that  $x_i \leq -C \wedge |y - y_i| < \varepsilon \wedge x_i \leq y - \varepsilon < y_i$ , hence

$$\text{a. } \min(x_i, y_i) = x_i \leq -C \text{ proving that } \lim_{i \rightarrow \infty} \min(x_i, y_i) = -\infty = \min(-\infty, y) = \min\left(\lim_{i \rightarrow \infty} x_i, \lim_{i \rightarrow \infty} y_i\right)$$

$$\text{b. } |\max(x_i, y_i) - y| = |y_i - y| < \varepsilon \text{ proving that } \lim_{i \rightarrow \infty} (\min(x_i, y_i)) = y = \max(-\infty, y) = \max\left(\lim_{i \rightarrow \infty} x_i, \lim_{i \rightarrow \infty} y_i\right)$$

$x \in \mathbb{R} \wedge y = \infty$ . this reduce to the case  $x = \infty \wedge y \in \mathbb{R}$  if we interchange  $x$  and  $y$

$x \in \mathbb{R} \wedge y = -\infty$ . this reduce to the case  $x = -\infty \wedge y \in \mathbb{R}$  if we interchange  $x$  and  $y$

$x \in \mathbb{R} \wedge y \in \mathbb{R}$ . then given  $\varepsilon > 0$  there exists  $N_1, N_2 \in \{k, \dots, \infty\}$  such that  $\forall i \geq N_1, \forall j \geq N_2$  we have  $|x_i - x| < \varepsilon \wedge |y_j - y| < \varepsilon \Leftrightarrow x - \varepsilon < x_i < x + \varepsilon \wedge y - \varepsilon < y_j < y + \varepsilon$ . Take then  $N = \max(N_1, N_2)$  then we have  $\forall i \geq N$  that

$$\text{a. } \min(x, y) - \varepsilon = \min(x - \varepsilon, y - \varepsilon) < x_i, y_i \Rightarrow \min(x, y) - \varepsilon < \min(x_i, y_i) \text{ and } \min(x_i, y_i) < x + \varepsilon, y + \varepsilon \Rightarrow \min(x_i, y_i) < \min(x + \varepsilon, y + \varepsilon) = \min(x, y) + \varepsilon \text{ proving that } |\min(x_i, y_i) - \min(x, y)| < \varepsilon \text{ giving } \lim_{i \rightarrow \infty} (\min(x_i, y_i)) = \min(x, y) = \min\left(\lim_{i \rightarrow \infty} x_i, \lim_{i \rightarrow \infty} y_i\right)$$

$$\text{b. } x - \varepsilon, y - \varepsilon < \max(x_i, y_i) \Rightarrow \max(x, y) - \varepsilon = \max(x - \varepsilon, y - \varepsilon) < \max(x_i, y_i) \text{ and } x_i, y_i < \max(x + \varepsilon, y + \varepsilon) = \max(x, y) + \varepsilon \text{ proving that } |\max(x_i, y_i) - \max(x, y)| < \varepsilon \text{ giving } \lim_{i \rightarrow \infty} (\max(x_i, y_i)) = \max(x, y) = \max\left(\lim_{i \rightarrow \infty} x_i, \lim_{i \rightarrow \infty} y_i\right) \quad \square$$

**Note 17.78.** The condition that  $\alpha$  in the above theorem is real is required as the following example shows. The sequence  $\left\{\frac{(-1)^i}{i}\right\}_{i \in \{1, \dots, \infty\}}$  converges to 0 but  $\left\{\infty \cdot \frac{(-1)^i}{i}\right\}_i = \{(-1)^i \cdot \infty\}_{i \in \{1, \dots, \infty\}}$  does not converges.

**Note 17.79.** The extra conditions in (5) are needed, for example take  $\{x_i\}_{i \in \mathbb{N}_0}$  defined by  $x_i = \frac{1}{i}$  and  $\{y_i\}_{i \in \mathbb{N}_0}$  defined by  $y_i = i^2$  then  $\lim_{i \rightarrow \infty} x_i = 0$  and  $\lim_{i \rightarrow \infty} y_i = \infty$  but  $\lim_{i \rightarrow \infty} (x_i \cdot y_i) = \lim_{i \rightarrow \infty} i = \infty \neq \infty \cdot 0 = 0$

To prove that  $\lim_{i \rightarrow \infty} \left( \frac{1}{x_i} \right) = \frac{1}{\lim_{i \rightarrow \infty} x_i}$  we first need a lemma

**Lemma 17.80.** Let  $\{x_i\}_{i \in \{k, \dots, \infty\}} \subseteq \mathbb{R}$  be a sequence of real numbers such that  $\forall i \in \{k, \dots, \infty\}$  we have  $x_i \neq 0$  and  $0 \neq \lim_{i \rightarrow \infty} x_i \in \mathbb{R}$  then  $\inf(\{|x_i| \mid i \in \{k, \dots, \infty\}\}) > 0$

**Proof.** Take  $x = \lim_{i \rightarrow \infty} x_i$  and define  $\varepsilon = \frac{|x|}{2} > 0$ , as  $x \in \mathbb{R}$  there exists a  $N \in \{k+1, \dots, \infty\}$  such that  $\forall i \geq N$  we have  $|x_i - x| < \varepsilon$ . Assume that  $\exists i \geq N$  such that  $|x_i| < \varepsilon$  then we have  $|x| = |x - x_i + x_i| \leq |x - x_i| + |x_i| < \varepsilon + \varepsilon = 2 \cdot \varepsilon = |x| \Rightarrow |x| < |x|$  a contradiction. So we must have

$$\forall i \geq N \models |x_i| \geq \varepsilon \quad (17.41)$$

If we take  $M' = \min(\{|x_i| \mid i \in \{k, \dots, N\}\}) > 0$  [as  $\forall i \in \{k, \dots, \infty\}$  we have  $x_i \neq 0$ ] and  $M = \min(M', \varepsilon) > 0$  then  $\forall i \in \{k, \dots, \infty\}$  we have either

$i \in \{k, \dots, N\}$ . then  $|x_i| \geq \varepsilon \geq M$

$i \in \{N+1, \dots, \infty\}$ .  $|x_i| \geq M' \geq M$

giving  $|x_i| \geq M > 0$  so that  $\inf(\{|x_i| \mid i \in \{k, \dots, \infty\}\}) \geq M > 0$ .  $\square$

**Theorem 17.81.** Let  $\{x_i\}_{i \in \{k, \dots, \infty\}}$  be a sequence of extended reals such that  $\lim_{i \rightarrow \infty} x_i \neq 0$  then there exists a  $n \in \{k, \dots, \infty\}$  such that  $\forall i \in \{n, \dots, \infty\}$  we have  $x_i \neq 0$  and for  $\{x_i\}_{i \in \{n, \dots, \infty\}}$  we have that  $\lim_{i \rightarrow \infty} \frac{1}{x_i} = \frac{1}{\lim_{i \rightarrow \infty} x_i}$

**Proof.** Let  $x = \lim_{i \rightarrow \infty} x_i$  then we must consider the following cases for  $x$

$x = \infty$ . First there exists a  $n \in \{k, \dots, \infty\}$  such that  $\forall i \in \{k, \dots, \infty\}$  we have  $1 \leq x_i \Rightarrow \forall i \in \{n, \dots, \infty\}$  we have  $x_i \neq 0$ . Take now  $\varepsilon > 0$  then by the Archimedean property of the reals [see 9.56] there exists a  $C \in \mathbb{N}_0$  such that  $0 < \frac{1}{C} < \varepsilon$ . As  $x = \infty$  there exists a  $N \in \{n, \dots, \infty\}$  such that  $\forall i \geq N$  we have  $x_i \geq C > 0$ . For  $x_i$  we have then either

$$x_i = \infty. \text{ then } \frac{1}{x_i} = 0 \text{ so that } \left| \frac{1}{x_i} - 0 \right| = 0 < \varepsilon$$

$$x_i < \infty. \text{ then } 0 < \frac{1}{x_i} \leq \frac{1}{C} < \varepsilon \Rightarrow \left| \frac{1}{x_i} - 0 \right| = \frac{1}{x_i} < \varepsilon$$

$$\text{which proves that } \lim_{i \rightarrow \infty} \frac{1}{x_i} = 0 = \frac{1}{\infty} = \frac{1}{\lim_{i \rightarrow \infty} x_i}$$

$x = -\infty$ . First there exists a  $n \in \{k, \dots, \infty\}$  such that  $\forall i \in \{k, \dots, \infty\}$  we have  $x_i \leq -1 \Rightarrow \forall i \in \{n, \dots, \infty\}$  we have  $x_i \neq 0$ . Take now  $\varepsilon > 0$  then by the Archimedean property of the reals (see 9.56) there exists a  $C \in \mathbb{N}_0$  such that  $0 < \frac{1}{C} < \varepsilon \Rightarrow -\varepsilon < -\frac{1}{C} < 0$ . As  $x = -\infty$  there exists a  $N \in \{n, \dots, \infty\}$  such that  $\forall i \geq N$  we have  $x_i \leq -C < 0 \Rightarrow 0 < C \leq -x_i$ . For  $x_i$  we have then either

$$x_i = -\infty. \text{ then } \frac{1}{x_i} = 0 \text{ so that } \left| \frac{1}{x_i} - 0 \right| = 0 < \varepsilon$$

$$-\infty < x_i. \text{ then } 0 < \frac{1}{-x_i} \leq \frac{1}{C} < \varepsilon \text{ so that } \left| \frac{1}{-x_i} - 0 \right| = \left| \frac{1}{x_i} \right| = \frac{1}{-x_i} < \varepsilon$$

$$\text{which proves that } \lim_{i \rightarrow \infty} \frac{1}{x_i} = 0 = \frac{1}{-\infty} = \frac{1}{\lim_{i \rightarrow \infty} x_i}$$

$x \in \mathbb{R} \setminus \{0\}$ . As  $x \neq 0$  we have that  $0 < |x|$  take  $\delta = \frac{|x|}{2} > 0$  then there exists a  $n \in \{k, \dots, \infty\}$  such that  $\forall i \geq n$  we have  $|x_i - x| < \delta = \frac{|x|}{2}$ , if now there exists a  $i \geq n$  such that  $x_i = 0$  then we have  $|x| < \frac{|x|}{2}$  a contradiction and if  $|x_i| = \infty$  we have  $\infty < \frac{|x|}{2} \in \mathbb{R}$  also a contradiction..

So  $\forall i \in \{n, \dots, \infty\}$  we have  $0 \neq x_i \in \mathbb{R}$ . Using the above lemma 17.80 we have that  $m = \inf(\{|x_i| \mid i \in \{n, \dots, \infty\}\}) > 0$  hence  $\forall i \in \{n, \dots, \infty\}$  we have that  $|x_i| \geq m > 0$ . Take now  $\varepsilon > 0$  then there exists a  $N \in \{n, \dots, \infty\}$  such that  $\forall i \geq N$  we have  $|x - x_i| < \varepsilon \cdot |x| \cdot m$ , then

$$\left| \frac{1}{x_i} - \frac{1}{x} \right| = \left| \frac{x_i - x}{x_i \cdot x} \right| = \frac{|x_i - x|}{|x_i| \cdot |x|} \leq \frac{|x_i - x|}{m \cdot |x|} < \frac{\varepsilon \cdot |x| \cdot m}{|x| \cdot m} = \varepsilon \text{ proving that } \lim_{i \rightarrow \infty} \left( \frac{1}{x_i} \right) = \frac{1}{\lim_{i \rightarrow \infty} x_i}$$

□

**Example 17.82.** We have that  $\lim_{n \rightarrow \infty} n = \infty$  and  $\lim_{n \rightarrow \infty} \frac{1}{n} = 0$

**Proof.** As  $\{n\}_{n \in \mathbb{N}_0}$  is increasing  $\lim_{n \rightarrow \infty} n = \sup(\{n | n \in \mathbb{N}_0\}) = \sup(\mathbb{N}_0)$  which exists. Assume that  $\sup(\mathbb{N}_0) < \infty$  then there exists a  $n \in \mathbb{N}_0$  such that  $\sup(\mathbb{N}_0) < n \leq \sup(\mathbb{N}_0)$  leading to the contradiction  $\sup(\mathbb{N}_0) < \sup(\mathbb{N}_0)$  hence we have  $\lim_{n \rightarrow \infty} n = \infty$ . Using the previous theorem 17.81 we have then  $\lim_{n \rightarrow \infty} \left(\frac{1}{n}\right) = \frac{1}{\lim_{n \rightarrow \infty} n} = \frac{1}{\infty} = 0$ . □

**Theorem 17.83.** Let  $\{x_i\}_{i \in \{k, \dots, \infty\}} \subseteq \bar{\mathbb{R}}$  with  $\lim_{i \rightarrow \infty} x_i = x \in \mathbb{R}$  then  $\lim_{i \rightarrow \infty} |x - x_i| = 0$ . Further if  $x \in \mathbb{R}$  is such that  $\lim_{i \rightarrow \infty} |x - x_i| = 0$  then  $\lim_{i \rightarrow \infty} x_i = x$

**Proof.** First as  $x \in \mathbb{R}$  we have that  $\forall i \in \mathbb{N}_0 x - x_i$  is well defined. Also by 17.73 we have that  $\forall \varepsilon \in \mathbb{R}_+$  there exists a  $N \in \{k, \dots, \infty\}$  such that  $\forall n \geq N$  we have  $|x_n - x| < \varepsilon \Rightarrow ||x_n - x| - 0| = |x_n - x| < \varepsilon$  proving by 17.73 again that  $\lim_{i \rightarrow \infty} |x_i - x| = 0$ . If  $\lim_{i \rightarrow \infty} |x_i - x| = 0 \in \mathbb{R}$  then by 17.73 we have that  $\forall \varepsilon \in \mathbb{R}_+$  there  $N \in \{k, \dots, \infty\}$  such that  $\forall n \geq N$  we have  $||x_n - x| - 0| < \varepsilon \Rightarrow |x_n - x| < \varepsilon$  proving by 17.73 again that  $\lim_{i \rightarrow \infty} x_i = x = 0$  □

**Note 17.84.** The condition that  $x \in \mathbb{R}$  in the above definition is needed. For example take  $\{x_i\}_{i \in \mathbb{N}_0}$  defined by  $x_i = i$  then  $\lim_{i \rightarrow \infty} x_i = \infty$  however  $\lim_{i \rightarrow \infty} |x_i - x| = \lim_{i \rightarrow \infty} |x_i - \infty| = \infty \neq 0$

### 17.3.1 Series in $\bar{\mathbb{R}}_+$

To avoid any problems with non defined sums in  $\bar{\mathbb{R}}$  [like  $-\infty + \infty$  and  $+\infty + (-\infty)$ ] we restrict ourself to  $\bar{\mathbb{R}}_+ = \{x \in \bar{\mathbb{R}} | x \geq 0\}$  so that by 17.28  $\langle \bar{\mathbb{R}}_+, + \rangle$  is a Abelian semi-group. All the theorems, definitions and propositions for Abelian semi groups in section 10.1 will then apply. In particular we can use the definition of a finite sum in 10.3 to define  $\sum_{i=0}^n x_i$  where  $\forall i \in \{0, \dots, n\}$  we have  $x_i \in \bar{\mathbb{R}}_+$ . So using 10.6 we have the following:

**Theorem 17.85.** Let  $n \in \mathbb{N}$  and  $\{x_i\}_{i \in \{0, \dots, n\}}$  be a finite family of non negative extended reals  $/\forall i \in \{0, \dots, n\} 0 \leq x_i/$  then  $\sum_{i=0}^n x_i = \begin{cases} x_0 \text{ if } n=0 \\ \sum_{i=0}^{n-1} x_i + x_n \text{ if } n > 0 \end{cases}$

Using 10.13 and 10.15 we have also in general

**Theorem 17.86.** Let  $n, m \in \mathbb{N}$  with  $n \leq m$  and  $\{x_i\}_{i \in \{n, \dots, m\}}$  a finite family of non negative extended reals then  $\sum_{i=n}^m x_i \stackrel{\text{def}}{=} \sum_{i=0}^{m-n} x_{i+n} = \begin{cases} x_n \text{ if } m=n \\ \sum_{i=n}^{m-1} x_i + x_m \text{ if } m > n \end{cases}$

Using the above we prove that the sum of a finite family of non negative extended reals is non negative as expressed in the following theorem.

**Theorem 17.87.** Let  $k, n \in \mathbb{N}$  with  $k \leq n$  and  $\{x_i\}_{i \in \{k, \dots, n\}}$  be a finite family of non negative extended reals then  $0 \leq \sum_{i=k}^n x_i$  and more general if  $\{x_i\}_{i \in I}$  is a finite family of non negative extended reals then  $0 \leq \sum_{i \in I} x_i$  [see 10.36]

**Proof.** We start by induction so let  $\mathcal{S} = \{n \in \{k, \dots, \infty\} | \text{for every family of non negatieve extended reals } \{x_i\}_{i \in \{k, \dots, n\}} \text{ we have } 0 \leq \sum_{i=k}^n x_i\}$ . Then we have

$k \in \mathcal{S}$ . As  $0 \leq x_k = \sum_{i=k}^k x_i$

$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$ . Let  $\{x_i\}_{i \in \{k, \dots, n+1\}}$  a family of non negative extended reals then  $\sum_{k=1}^{n+1} x_i = (\sum_{i=k}^n x_i) + x_{n+1} \geq 0$  [as  $n \in \mathcal{S} \Rightarrow 0 \leq \sum_{i=k}^n x_i, 0 \leq x_{n+1}$  and 17.26] proving  $n+1 \in \mathcal{S}$ .

For the last part of the theorem if  $I$  is a finite family we have a bijection  $\sigma: \{1, \dots, n\} \rightarrow I$  so that  $\sum_{i \in I} x_i \stackrel{\text{def}}{=} \sum_{i=1}^n x_{\sigma(i)} \geq 0$   $\square$

As a application of the above theorem we prove that series are increasing based on there domain.

**Theorem 17.88.**  $\{x_i\}_{i \in I}$  is a finite family of extended non negative reals then  $\forall J \subseteq I$  we have  $\sum_{i \in J} x_i \leq \sum_{i \in I} x_i$ . So if  $\{x_i\}_{i \in \{1, \dots, n\}}$  is a finite family of non negative extended reals then  $\forall m \in \{1, \dots, n\}$  we have  $\sum_{i=1}^m x_i \leq \sum_{i=1}^n x_i$ .

**Proof.** As  $I = (I \setminus J) \cup J$  and  $(I \setminus J) \cap I = \emptyset$  we have by 10.45 that  $\sum_{i \in I} x_i = \sum_{i \in (I \setminus J)} x_i + \sum_{i \in J} x_i \stackrel{0 \leq \sum_{i \in I \setminus J} x_i \text{ and 17.27}}{\Rightarrow} \sum_{i \in J} x_i \leq \sum_{i \in I} x_i$   $\square$

**Theorem 17.89.** Let  $n, m \in \mathbb{N}$  with  $n \leq m$  and  $\{x_i\}_{i \in \{n, \dots, m\}}$  be a finite family of non negative reals then if  $\alpha \in \mathbb{R}$  we have  $\sum_{i=n}^m (\alpha \cdot x_i) = \alpha \cdot \sum_{i=n}^m x_i$

**Proof.** We prove this by induction so let  $S = \{m \in \{n, \dots, \infty\} \mid \text{For every family of non negative extended real numbers } \{x_i\}_{i \in \{n, \dots, m\}} \text{ we have } \sum_{i=n}^m (\alpha \cdot x_i) = \alpha \cdot \sum_{i=n}^m x_i\}$  then

$m \in S$ . If  $\{x_i\}_{i \in \{m, \dots, m\}}$  is a finite family of non negative extended reals then  $\sum_{i=m}^m (\alpha \cdot x_i) = \alpha \cdot x_m = \alpha \cdot (\sum_{i=m}^m x_i)$  proving that  $m \in S$

$m \in S \Rightarrow m + 1 \in S$ . If  $\{x_i\}_{i \in \{n, \dots, m+1\}}$  is a finite family of non negative extended reals then  $\sum_{i=n}^{m+1} (\alpha \cdot x_i) = (\sum_{i=n}^m (\alpha \cdot x_i)) + \alpha \cdot x_{m+1} \stackrel{m \in S}{=} \alpha \cdot (\sum_{i=1}^m x_i) + \alpha \cdot x_{m+1} \stackrel{17.29}{=} \alpha \cdot (\sum_{i=1}^{m+1} x_i)$  proving that  $m + 1 \in S$   $\square$

The next theorem shows that the finite sum of limits is the limit of finite sums.

**Theorem 17.90.** If  $k, n \in \mathbb{N}$  and  $\{\{x_{i,j}\}_{j \in \{k, \dots, \infty\}}\}_{i \in \{l, \dots, n\}}$  is a finite family of sequences of non negative extended reals such that  $\forall i \in \{l, \dots, n\}$  we have that  $\{x_{i,j}\}_{j \in \{k, \dots, \infty\}}$  has a limit then  $\lim_{j \rightarrow \infty} (\sum_{i=l}^n x_{i,j}) = \sum_{i=l}^n \left( \lim_{j \rightarrow \infty} x_{i,j} \right)$

**Proof.** We use mathematical induction to prove this, so let  $S = \{n \in \{l, \dots, \infty\} \mid \text{if } \{\{x_{i,j}\}_{j \in \{k, \dots, \infty\}}\}_{i \in \{l, \dots, n\}} \text{ is such that } \forall i \in \{l, \dots, n\} \models \{x_{i,j}\}_{j \in \{k, \dots, \infty\}}$  has a limit then  $\lim_{j \rightarrow \infty} (\sum_{i=l}^n x_{i,j}) = \sum_{i=l}^n \left( \lim_{j \rightarrow \infty} x_{i,j} \right)\}$ . Then :

$l \in S$ . If  $\{\{x_{i,j}\}_{j \in \{k, \dots, \infty\}}\}_{i \in \{l, \dots, l\}}$  is such that  $\forall i \in \{l, \dots, l\} \models \{x_{i,j}\}_{j \in \{k, \dots, \infty\}}$  has a limit then  $\{\sum_{i=l}^l x_{i,j}\}_{j \in \{k, \dots, \infty\}} = \{x_{l,j}\}_{j \in \{k, \dots, \infty\}}$  has a limit and  $\lim_{j \rightarrow \infty} (\sum_{i=l}^l x_{i,j}) = \lim_{j \rightarrow \infty} x_{l,j} = \sum_{i=l}^l \left( \lim_{j \rightarrow \infty} x_{i,j} \right)$  proving that  $l \in S$

$n \in S \Rightarrow n + 1 \in S$ . If  $\{\{x_{i,j}\}_{j \in \{k, \dots, \infty\}}\}_{i \in \{l, \dots, n+1\}}$  is such that  $\forall i \in \{l, \dots, n+1\} \models \{x_{i,j}\}_{j \in \{k, \dots, \infty\}}$  has a limit, then as  $n \in S$  we have that  $\{\sum_{i=l}^n x_{i,j}\}_{j \in \{k, \dots, \infty\}}$  has a limit. Then  $\{\sum_{i=l}^{n+1} x_{i,j}\}_{j \in \{k, \dots, \infty\}} = \{(\sum_{i=l}^n x_{i,j}) + x_{n+1,j}\}_{j \in \{k, \dots, \infty\}}$  has a limit (using 17.77 (4)) and

$$\begin{aligned} \lim_{j \rightarrow \infty} \left( \left( \sum_{i=l}^n x_{i,j} \right) + x_{n+1,j} \right) &\stackrel{17.77 (4)}{=} \lim_{j \rightarrow \infty} \left( \sum_{i=l}^n x_{i,j} \right) + \lim_{j \rightarrow \infty} x_{n+1,j} \\ &\stackrel{n \in S}{=} \sum_{i=l}^n \left( \lim_{j \rightarrow \infty} x_{i,j} \right) + \lim_{j \rightarrow \infty} x_{n+1,j} \\ &= \sum_{i=l}^{n+1} \left( \lim_{j \rightarrow \infty} x_{i,j} \right) \end{aligned}$$

proving that  $n+1 \in \mathcal{S}$ .

Using mathematical induction it follows that  $\mathcal{S} = \{l, \dots, \infty\}$  which completes the proof.  $\square$

**Theorem 17.91.** Let  $n, m \in \mathbb{N}$  with  $n \leq m$  and  $\{x_i\}_{i \in \{n, \dots, m\}}$  be a finite family of extended non negative reals then we have

1. If  $\infty \in \{x_i | i \in \{n, \dots, m\}\}$  then  $\sum_{i=n}^m x_i = \infty$
2. If  $\infty \notin \{x_i | i \in \{n, \dots, m\}\}$  then  $\sum_{i=n}^m x_i < \infty$

**Proof.**

1. We prove this by induction so let  $\mathcal{S} = \{m \in \{n, \dots, \infty\} | \text{Let } \{x_i\}_{i \in \{n, \dots, m\}} \subseteq \bar{\mathbb{R}}_+ \text{ with } \infty \in \{x_n, \dots, x_m\} \text{ then } \sum_{i=n}^m x_i = \infty\}$  then we have

$n \in \mathcal{S}$ . then  $\infty \in \{x_n, \dots, x_n\} \Rightarrow x_n = \infty$  and  $\sum_{i=n}^n x_i = x_n = \infty$  proving that  $n \in \mathcal{S}$

$m \in \mathcal{S} \Rightarrow m+1 \in \mathcal{S}$ . let  $\infty \in \{x_n, \dots, x_{m+1}\}$  then we have either

$x_{m+1} = \infty$ . then  $\sum_{i=n}^{m+1} x_i = \sum_{i=n}^m x_i + x_{m+1} = \sum_{i=n}^m x_i + \infty = \infty$

$\infty \in \{x_n, \dots, x_m\}$ . then  $\sum_{i=n}^{m+1} x_i = \sum_{i=n}^m x_i + x_{m+1} \underset{n \in \mathcal{S}}{=} \infty + x_{m+1} = \infty$

so in all cases we have  $\sum_{i=n}^{m+1} x_i = \infty$  proving that  $m+1 \in \mathcal{S}$

Using mathematical induction we have then that  $\forall m \in \{n, \dots, \infty\}$  we have  $\sum_{i=n}^m x_i = \infty$

2. We prove this also by induction so let  $\mathcal{S} = \{m \in \{n, \dots, \infty\} | \text{Let } \{x_i\}_{i \in \{n, \dots, m\}} \subseteq \bar{\mathbb{R}}_+ \text{ with } \infty \notin \{x_n, \dots, x_m\} \text{ then } \sum_{i=n}^m x_i < \infty\}$  then we have

$n \in \mathcal{S}$ . then as  $\infty \notin \{x_n, \dots, x_n\} \Rightarrow x_n < \infty$  and thus  $\sum_{i=n}^n x_i = x_n < \infty$  proving that  $n \in \mathcal{S}$

$m \in \mathcal{S} \Rightarrow m+1 \in \mathcal{S}$ . then from  $\infty \notin \{x_n, \dots, x_{m+1}\}$  we have  $\infty \notin \{x_n, \dots, x_m\} \wedge x_{m+1} \neq \infty$ . As  $m \in \mathcal{S}$  we have that  $\sum_{i=n}^{m+1} x_i < \infty$  which as  $\sum_{i=n}^{m+1} x_i = \sum_{i=n}^m x_i + x_{m+1}$  proves that  $m+1 \in \mathcal{S}$

By mathematical induction we have that  $\mathcal{S} = \{n, \dots, \infty\}$  proving our theorem.  $\square$

We extend now the sum of constant values [see 10.30 for the real case].

**Proposition 17.92.** Let  $n, k \in \mathbb{N}$  with  $k \leq n$  and  $\{x_i\}_{i \in \{k, \dots, n\}}$  be a finite family with  $\forall i \in \{k, \dots, n\}$   $x_i = c \in \bar{\mathbb{R}}_+$  then  $\sum_{i=k}^n x_i = (n - k + 1) \cdot c$ . More general if  $\{x_i\}_{i \in I}$  is a family with finite support such that  $\forall i \in I$  we have  $x_i = c$  then  $\sum_{i \in I} x_i = \#(I) \cdot c$

**Proof.** We prove the first part by induction so let  $S = \{n \in \{k, \dots, \infty\} | \text{if } \{x_i\}_{i \in \{k, \dots, n\}}$  is a family with  $x_i = c$  then  $\sum_{i=k}^n x_i = n \cdot c\}$  then we have

$k \in S$ . If  $\{x_i\}_{i \in \{k, \dots, k\}}$  is such that  $x_i = c$  then  $\sum_{i=k}^k x_i = x_k = c = (k - k + 1) \cdot c$  proving that  $k \in S$

$n \in S \Rightarrow n+1 \in S$ . take  $\{x_i\}_{i \in \{k, \dots, n+1\}}$  with  $x_i = c$  then we have  $\sum_{i=k}^{n+1} x_i = (\sum_{i=k}^n x_i) + x_{n+1} \underset{n \in S}{=} (n - k + 1) \cdot c + x_{n+1} = (n - k + 1) \cdot c + c$ . We have now two cases either  $c = \infty$  then  $((n + 1) - k + 1) \cdot \infty = \infty = (n - k + 1) \cdot c + c$  or  $c < \infty$  then  $(n - k + 1) \cdot c + c = ((n + 1) - k + 1) \cdot c$ . So  $\sum_{i=1}^{n+1} x_i = ((n + 1) - k + 1) \cdot c$  proving  $n + 1 \in S$

Mathematical induction completes then the proof of the first part of the proposition.

For the second part either we have  $c = 0$  and then  $\text{support}(\{x_i\}_{i \in I}) = \emptyset$  and thus by definition  $\sum_{i \in I} x_i = 0 = \#(I) \cdot 0 = \#(I) \cdot c$  or  $c > 0$  and then  $\text{support}(\{x_i\}_{i \in I}) = I$ , so there exists a bijection  $b: \{1, \dots, \#(I)\} \rightarrow I$  and  $\sum_{i \in I} x_i = \sum_{i=1}^{\#(I)} x_{b(i)} = (\#(I) - 1 + 1) \cdot c = \#(I) \cdot c$ .  $\square$

**Theorem 17.93.** Let  $k \in \mathbb{N}_0$ ,  $\{x_i\}_{i \in \{k, \dots, \infty\}} \subseteq \bar{\mathbb{R}}_+$  [a sequence of non negative extended reals] then

1.  $\forall n \in \{k, \dots, \infty\}$  we have  $\sum_{i=k}^n x_i \leq \sum_{i=k}^{n+1} x_i$

$$2. \lim_{n \rightarrow \infty} (\sum_{i=k}^n x_i) \text{ exists and } \lim_{n \rightarrow \infty} (\sum_{i=k}^n x_i) = \sup (\{\sum_{i=k}^n x_i | n \in \{k, \dots, \infty\}\})$$

**Proof.** This follows from 17.88 and 17.71  $\square$

The above theorem allows us to define the infinite sum of all sequences of non negative extended real numbers.

**Definition 17.94.** Let  $\{x_i\}_{i \in \{n, \dots, \infty\}}$  be a sequence of non negative extended real numbers then we define  $\sum_{i=k}^{\infty} x_i$  by

$$\sum_{i=k}^{\infty} x_i \stackrel{\text{def}}{=} \lim_{n \rightarrow \infty} \left( \sum_{i=k}^n x_i \right) \stackrel{17.93}{=} \sup \left( \left\{ \sum_{i=k}^n x_i | n \in \{k, \dots, \infty\} \right\} \right)$$

**Theorem 17.95.** Let  $n \in \mathbb{N}$ ,  $\{x_i\}_{i \in \{n, \dots, \infty\}}$  be a sequence of non negative extended real numbers abd  $\alpha \in \mathbb{R}$ ,  $0 \leq \alpha$  then

$$\sum_{i=n}^{\infty} (\alpha \cdot x_i) = \alpha \cdot \sum_{i=n}^{\infty} x_i$$

**Proof.** This follows from  $\sum_{i=n}^{\infty} (\alpha \cdot x_i) = \lim_{k \rightarrow \infty} (\sum_{i=n}^k \alpha \cdot x_i) \stackrel{17.89}{=} \lim_{k \rightarrow \infty} (\alpha \cdot \sum_{i=n}^k x_i) \stackrel{17.77}{=} \alpha \cdot \lim_{k \rightarrow \infty} (\sum_{i=n}^k x_i) = \alpha \cdot \sum_{i=k}^{\infty} x_i$   $\square$

**Theorem 17.96.** Let  $k \in \mathbb{N}$  and  $\{x_i\}_{i \in \{k, \dots, \infty\}}$ ,  $\{y_i\}_{i \in \{k, \dots, \infty\}}$  be sequences of non negative extended real numbers then

$$\sum_{i=k}^{\infty} (x_i + y_i) = \sum_{i=k}^{\infty} x_i + \sum_{i=k}^{\infty} y_i$$

**Proof.** Let  $n \in \{k, \dots, \infty\}$  then

$$\begin{aligned} \sum_{i=k}^{\infty} (x_i + y_i) &= \lim_{n \rightarrow \infty} \sum_{i=k}^n (x_i + y_i) \\ &\stackrel{10.52}{=} \lim_{n \rightarrow \infty} \left( \sum_{i=k}^n x_i + \sum_{i=k}^n y_i \right) \\ &\stackrel{17.77}{=} \lim_{n \rightarrow \infty} \sum_{i=k}^n x_i + \lim_{n \rightarrow \infty} \sum_{i=k}^n y_i \\ &= \sum_{i=k}^{\infty} x_i + \sum_{i=k}^{\infty} y_i \end{aligned} \quad \square$$

Series preserve the order relation in the set of extended reals as the following theorem shows.

**Theorem 17.97.** Let  $\{x_i\}_{i \in \{k, \dots, \infty\}}$ ,  $\{y_i\}_{i \in \{k, \dots, \infty\}}$  be sequences of positive extended reals such that  $\forall i \in \{k, \dots, \infty\}$  we have  $x_i \leq y_i$  then we have that  $\sum_{i=k}^{\infty} x_i \leq \sum_{i=k}^{\infty} y_i$

**Proof.** Note that  $\forall l \in \{k, \dots, \infty\}$  we have by 10.32 that  $\sum_{i=k}^l x_i \leq \sum_{i=k}^l y_i$  and thus by 2.176 we have that  $\sum_{i=k}^{\infty} x_i \stackrel{17.94}{=} \sup (\{\sum_{i=k}^l x_i | l \in \{k, \dots, \infty\}\}) \leq \sup (\{\sum_{i=k}^l y_i | l \in \{k, \dots, \infty\}\}) \stackrel{17.94}{=} \sum_{i=k}^{\infty} y_i$   $\square$

The following two theorems are intuitive very clear but still needs some proof.

**Theorem 17.98.** Let  $k \in \mathbb{N}$  and  $\{x_i\}_{i \in \{k, \dots, \infty\}}$  be a sequence of non negative extended reals such that  $\sum_{i=k}^{\infty} x_i < \infty$  then  $\forall i \in \{k, \dots, \infty\}$  we have  $x_i < \infty$ .

**Proof.** As  $\sup (\sum_{i=k}^n x_i | n \in \{k, \dots, \infty\}) \stackrel{17.94}{=} \sum_{i=k}^{\infty} x_i < \infty$  we have, using 17.14, that  $\forall n \in \{k, \dots, \infty\} \models \sum_{i=k}^n x_i < \infty$  which as  $x_n \leq \sum_{i=k}^n x_i$  proves that  $x_i < \infty$ .  $\square$

**Theorem 17.99.** Let  $k \in \mathbb{N}$  and  $\{x_i\}_{i \in \{k, \dots, \infty\}}$  a sequence of non negative extended reals such that  $\exists m \in \{k, \dots, \infty\}$  with  $\forall i \in \{m+1, \dots, \infty\}$  we have  $x_i = 0$  then  $\sum_{i=k}^{\infty} x_i = \sum_{i=k}^m x_i$

**Proof.** As  $\sum_{i=k}^{\infty} x_i = \sup(\{\sum_{i=k}^n x_i \mid n \in \{k, \dots, \infty\}\}) \geq \sum_{i=k}^m x_i$  it follows that  $\sum_{i=k}^{\infty} x_i \leq \sum_{i=1}^{\infty} x_i$ . For the opposite inequality if  $n \in \{k, \dots, \infty\}$  we have either

$$n \in \{k, \dots, m\}. \text{ then } \sum_{i=k}^n x_i \leq \text{17.88} \sum_{i=k}^m x_i$$

$$n \in \{m+1, \dots, \infty\}. \text{ then } \sum_{i=k}^n x_i = \sum_{i=k}^m x_i + \sum_{i=m+1}^n x_i = \sum_{i=k}^m x_i$$

hence  $\sum_{i=k}^{\infty} x_i = \sup(\{\sum_{i=k}^n x_i \mid n \in \{k, \dots, \infty\}\}) \leq \sum_{i=k}^m x_i$ .  $\square$

We show now that the finite sum of a family of extended positive reals is independent of the order of the terms.

**Theorem 17.100.** Let  $n, m \in \mathbb{N}$  with  $n \leq m$  and  $\{x_i\}_{i \in \{n, \dots, m\}}$  a finite family of non negative extended reals and  $\sigma: \{n, \dots, m\} \rightarrow \{n, \dots, m\}$  a bijection then  $\sum_{i=n}^m x_i = \sum_{i=n}^m x_{\sigma(i)}$

**Proof.** This follows from the fact that  $\langle \bar{\mathbb{R}}_+, + \rangle$  is a Abelian semi-group and  $\sum_{i=n}^m x_i \stackrel{\text{17.86}}{=} \sum_{i=0}^{m-n} x_{i+n} \stackrel{\text{10.26}}{=} \sum_{i=0}^{m-n} x_{\sigma(i+n)} = \sum_{i=n}^m x_{\sigma(i)}$   $\square$

We extend now the above to the infinite sum of non negative real numbers.

**Theorem 17.101.** Let  $k \in \mathbb{N}$  and  $\{x_i\}_{i \in \{k, \dots, \infty\}}$  a sequence of non negative extended real numbers then  $\sum_{i=k}^{\infty} x_i = \sum_{i=k}^{\infty} x_{\sigma(i)}$  for every permutation  $\sigma: \{k, \dots, \infty\} \rightarrow \{k, \dots, \infty\}$

**Proof.** Let  $\sigma: \{k, \dots, \infty\} \rightarrow \{k, \dots, \infty\}$  be a bijection. Take  $s \in \{\sum_{i=1}^n x_{\sigma(i)} \mid n \in \{k, \dots, \infty\}\}$  then  $\exists n \in \{k, \dots, \infty\}$  such that  $s = \sum_{i=k}^n x_{\sigma(i)}$ . Define  $m_n = \max(\{\sigma(i) \mid i \in \{k, \dots, n\}\})$  then  $\forall i \in \{k, \dots, n\}$  we have  $\sigma(i) \in \{k, \dots, m_n\}$  or

$$\{\sigma(i) \mid i \in \{k, \dots, n\}\} \subseteq \{k, \dots, m_n\} \quad (17.42)$$

so that

$$\begin{aligned} \sum_{i=k}^n x_{\sigma(i)} &\stackrel{\text{10.36}}{=} \sum_{i \in \{\sigma(i) \mid i \in \{k, \dots, n\}\}} x_i \\ &\leq \text{17.88 and 17.42} \sum_{i \in \{k, \dots, m_n\}} x_i \\ &\stackrel{\text{10.36}}{=} \sum_{i=k}^{m_n} x_i \end{aligned}$$

so using 2.176 and the fact that  $\sum_{i=k}^{m_n} x_i \in \{\sum_{i=1}^n x_i \mid n \in \{k, \dots, \infty\}\}$  it follows that  $\sup(\{\sum_{i=1}^n x_{\sigma(i)} \mid n \in \{k, \dots, \infty\}\}) \leq \sup(\{\sum_{i=1}^n x_i \mid n \in \{k, \dots, \infty\}\})$ . Hence using the definition of infinite sums [see 17.94] we have

$$\text{If } \sigma: \{k, \dots, \infty\} \rightarrow \{k, \dots, \infty\} \text{ is a bijection then } \sum_{k=1}^{\infty} x_{\sigma(i)} \leq \sum_{k=1}^{\infty} x_i \quad (17.43)$$

For the opposite inequality let  $\sigma: \{k, \dots, \infty\} \rightarrow \{k, \dots, \infty\}$  be a a bijection and take the bijection  $\tau = \sigma^{-1}$  then we have

$$\begin{aligned} \sum_{k=1}^{\infty} x_i &= \sum_{k=1}^{\infty} x_{\tau(\sigma(i))} \\ &\leq \text{17.43 } \tau \text{ is a bijection} \sum_{k=1}^{\infty} x_{\sigma(i)} \end{aligned}$$

which using 17.43 proves that

$$\text{If } \sigma: \{k, \dots, \infty\} \rightarrow \{k, \dots, \infty\} \text{ is a bijection } \sum_{i=k}^{\infty} x_{\sigma(i)} = \sum_{i=k}^{\infty} x_i \quad \square$$

The next theorem shows how we can remove zeroes out of a infinite sum of extended reals.

**Theorem 17.102.** Let  $k \in \mathbb{N}$  and  $\{x_i\}_{i \in \{k, \dots, \infty\}}$  be a family of extended positive reals and  $S = \{i \in \{k, \dots, n\} | x_i \neq 0\}$  then we have the following possibilities

**$S = \emptyset$ .** then  $\sum_{i=k}^{\infty} x_i = 0$

**$\emptyset \neq S$  is finite.** then there exists a bijection  $\beta: \{1, \dots, n\} \rightarrow S$  such that  $\sum_{i=k}^{\infty} x_i = \sum_{i=1}^n x_{\beta(i)}$

**$S$  is infinite countable.** then there exists a bijection  $\beta: \mathbb{N}_0 \rightarrow S$  such that  $\sum_{i=k}^{\infty} x_i = \sum_{i=1}^{\infty} x_{\beta(i)}$

**Proof.** As  $S \subseteq \{k, \dots, \infty\}$  we have that

$$\forall i \in \{k, \dots, \infty\} \setminus S \models x_i = 0 \quad (17.44)$$

For  $S$  we have the following cases to consider

**$S = \emptyset$ .** Then  $\forall i \in \{k, \dots, \infty\}$  we have  $x_i = 0$ . so that  $\sum_{i=k}^{\infty} x_i = \sup(\{\sum_{i=k}^n x_i | n \in \{k, \dots, \infty\}\}) = \sup(\{0\}) = 0$

**$\emptyset \neq S$  is finite.** Then there exists a  $n \in \mathbb{N}_0$  and a bijection  $\beta: \{1, \dots, n\} \rightarrow S$ . Take  $m = \max(S) \subseteq \{k, \dots, \infty\}$  then if  $i > m$  we have  $i \notin S$  and thus by 17.44  $x_i = 0$  giving

$$\forall i > m \text{ we have } x_i = 0 \quad (17.45)$$

Further  $S \subseteq \{k, \dots, m\}$  so that

$$\{k, \dots, m\} = (\{k, \dots, m\} \setminus S) \bigcup S \text{ and } \{k, \dots, m\} \setminus S \bigcap S = \emptyset \quad (17.46)$$

Now for  $l \in \{m, \dots, \infty\}$  we have

$$\begin{aligned} \sum_{i=k}^l x_i &= \sum_{i=k}^m x_i + \sum_{i=m+1}^l x_i \\ &\stackrel{17.45}{=} \sum_{i=k}^m x_i \\ &= \sum_{i \in \{k, \dots, m\}} x_i \\ &\stackrel{10.44 \text{ and } 17.46}{=} \sum_{i \in \{k, \dots, m\} \setminus S} x_i + \sum_{i \in S} x_i \\ &\stackrel{17.44}{=} 0 + \sum_{i \in S} x_i \\ &= \sum_{i \in \beta(\{1, \dots, n\})} x_i \\ &\stackrel{10.43}{=} \sum_{i \in \{1, \dots, n\}} x_{\beta(i)} \\ &= \sum_{i=1}^n x_{\beta(i)} \end{aligned} \quad (17.47)$$

Further

$$\begin{aligned} \sum_{i=k}^{\infty} x_i &\stackrel{17.94}{=} \sup \left( \left\{ \sum_{i=k}^n x_i \mid n \in \{k, \dots, \infty\} \right\} \right) \\ &\stackrel{17.70}{=} \sup \left( \left\{ \sum_{i=k}^n x_i \mid n \in \{m, \dots, \infty\} \right\} \right) \\ &\stackrel{17.47}{=} \sup \left( \left\{ \sum_{i=1}^n x_{\beta(i)} \right\} \right) \\ &= \sum_{i=1}^n x_{\beta(i)} \end{aligned}$$

**$S$  is infinite countable.** Then there exists a bijection  $\beta: \mathbb{N}_0 \rightarrow S$ . Take  $n \in \{k, \dots, \infty\}$  then for  $\{k, \dots, n\}$  we have to consider the following cases:

$\forall i \in \{k, \dots, n\} \models x_i = 0$ . then  $\sum_{i=1}^n x_i = 0 \leq \sum_{i=k}^n x_{\beta(i)}$

$\exists i \in \{k, \dots, n\} \models x_i \neq 0$ . then  $i \in \{k, \dots, n\} \cap S$  and as  $\beta$  is surjective there exists a  $j \in \{1, \dots, \infty\}$  such that  $\beta(j) = i \in \{k, \dots, n\} \cap S$  proving that  $\emptyset \neq \beta^{-1}(\{k, \dots, n\} \cap S)$ . As  $\beta^{-1}(\{k, \dots, n\} \cap S)$  is also finite  $m = \max(\beta^{-1}(\{k, \dots, n\} \cap S))$  exists, then  $\beta^{-1}(\{k, \dots, n\} \cap S) \subseteq \{1, \dots, m\}$  so that

$$\{k, \dots, n\} \cap S \subseteq \beta(1, \dots, m) \quad (17.48)$$

Futher we have that

$$\{k, \dots, n\} = (\{k, \dots, n\} \setminus S) \bigcup (\{k, \dots, n\} \cap S) \quad (17.49)$$

and

$$(\{k, \dots, n\} \setminus S) \cap (\{k, \dots, n\} \cap S) = \emptyset \quad (17.50)$$

So

$$\begin{aligned} \sum_{i=k}^n x_i &= \sum_{i \in \{k, \dots, n\}} x_i \\ &\stackrel{10.44 \text{ and } 17.49, 17.50}{=} \sum_{i \in \{k, \dots, n\} \setminus S} x_i + \sum_{i \in \{k, \dots, n\} \cap S} x_i \\ &= 0 + \sum_{i \in \{k, \dots, n\} \cap S} x_i \\ &\leq \sum_{i \in \{k, \dots, n\} \cap S} x_i + \sum_{i \in \beta(\{1, \dots, m\}) \setminus (\{k, \dots, n\} \cap S)} x_i \\ &\stackrel{10.44 \text{ and } 17.48}{=} \sum_{i \in \beta(\{1, \dots, m\})} x_i \\ &\stackrel{10.43}{=} \sum_{i \in \{1, \dots, m\}} x_{\beta(i)} \\ &= \sum_{i=1}^m x_{\beta(i)} \end{aligned}$$

So  $\forall n \in \{k, \dots, \infty\}$  there exists a  $m \in \{1, \dots, \infty\}$  such that  $\sum_{i=1}^n x_i \leq \sum_{i=1}^m x_i$  proving using 2.176 that

$$\sup \left( \left\{ \sum_{i=k}^n x_i \mid n \in \{k, \dots, \infty\} \right\} \right) \leq \sup \left( \left\{ \sum_{i=1}^m x_{\beta(i)} \mid m \in \{1, \dots, \infty\} \right\} \right) \quad (17.51)$$

For the opposite inequality, take  $n \in \{1, \dots, \infty\}$  then  $\beta(\{1, \dots, n\})$  is non empty and finite so that  $m = \max(\beta(\{1, \dots, n\})) \subseteq S \subseteq \{k, \dots, \infty\}$  exists giving

$$\beta(\{1, \dots, n\}) \subseteq \{k, \dots, m\} \quad (17.52)$$

hence

$$\begin{aligned} \sum_{i=1}^n x_{\beta(i)} &= \sum_{i \in \{1, \dots, n\}} x_{\beta(i)} \\ &\stackrel{10.43}{=} \sum_{i \in \beta(\{1, \dots, n\})} x_i \\ &\leq \sum_{i \in \{k, \dots, m\} \setminus \beta(\{1, \dots, n\})} x_i + \sum_{i \in \beta(\{k, \dots, m\})} x_i \\ &\stackrel{10.44 \text{ and } 17.52}{=} \sum_{i \in \{k, \dots, m\}} x_i \\ &= \sum_{i=k}^m x_i \end{aligned}$$

Hence  $\forall n \in \{1, \dots, \infty\}$  there exists a  $m \in \{k, \dots, \infty\}$  such that  $\sum_{i=1}^n x_i \leq \sum_{i=k}^m x_i$  proving using 2.176 that

$$\sup \left( \left\{ \sum_{i=1}^n x_{\beta(i)} \mid n \in \{1, \dots, \infty\} \right\} \right) \leq \sup \left( \left\{ \sum_{i=k}^m x_i \mid n \in \{k, \dots, \infty\} \right\} \right)$$

Finally using the above together with 17.51 we have

$$\sum_{i=k}^{\infty} x_i = \sup \left( \left\{ \sum_{i=k}^n x_i \mid n \in \{k, \dots, \infty\} \right\} \right) = \sup \left( \left\{ \sum_{i=1}^n x_{\beta(i)} \mid n \in \{1, \dots, \infty\} \right\} \right) = \sum_{i=1}^{\infty} x_{\beta(i)}$$

□

The next theorem shows that we can interchange infinite sums.

**Theorem 17.103.** *Let  $k, l \in \mathbb{N}$*

1. *Let  $n, m \in \mathbb{N}$  with  $k \leq n \wedge l \leq m$  and  $\{x_{(i,j)}\}_{(i,j) \in \{k, \dots, n\} \times \{l, \dots, m\}}$  a family of extended non negative real numbers then*

$$\sum_{i=k}^n \left( \sum_{j=l}^m x_{(i,j)} \right) = \sum_{j=l}^m \left( \sum_{i=k}^n x_{(i,j)} \right)$$

2. *Let  $n \in \mathbb{N}$  with  $l \leq n$  and  $\{x_{(i,j)}\}_{(i,j) \in \{k, \dots, \infty\} \times \{l, \dots, n\}}$  a family of extended non negative real numbers then*

$$\sum_{i=k}^{\infty} \left( \sum_{j=l}^n x_{(i,j)} \right) = \sum_{j=l}^n \left( \sum_{i=k}^{\infty} x_{(i,j)} \right)$$

3. *Let  $n \in \mathbb{N}$  with  $l \leq n$  and  $\{x_{(i,j)}\}_{(i,j) \in \{k, \dots, n\} \times \{l, \dots, \infty\}}$  a family of extended non negative real numbers then*

$$\sum_{i=k}^n \left( \sum_{j=l}^{\infty} x_{(i,j)} \right) = \sum_{j=l}^{\infty} \left( \sum_{i=k}^n x_{(i,j)} \right)$$

4. *Let  $\{x_{(i,j)}\}_{(i,j) \in \{k, \dots, \infty\} \times \{l, \dots, \infty\}}$  be a family of extended non negative real numbers then*

$$\sum_{i=k}^{\infty} \left( \sum_{j=l}^{\infty} x_{(i,j)} \right) = \sum_{j=l}^{\infty} \left( \sum_{i=k}^{\infty} x_{(i,j)} \right)$$

**Proof.**

1. This follows from

$$\begin{aligned} \sum_{i=k}^n \left( \sum_{j=l}^m x_{(i,j)} \right) &\stackrel{10.40}{=} \sum_{i \in \{k, \dots, n\}} \left( \sum_{j \in \{l, \dots, m\}} x_{(i,j)} \right) \\ &\stackrel{10.47}{=} \sum_{j \in \{l, \dots, n\}} \left( \sum_{i \in \{k, \dots, m\}} x_{(i,j)} \right) \\ &= \sum_{j=l}^m \left( \sum_{i=k}^n x_{(i,j)} \right) \end{aligned}$$

2. We prove this by induction so let  $\mathcal{S} = \{n \in \mathbb{N} \mid \text{if } \{x_{(i,j)}\}_{(i,j) \in \{k, \dots, \infty\} \times \{l, \dots, l+n\}} \subseteq [0, \infty] \text{ then } \sum_{i=k}^{\infty} (\sum_{j=l}^{l+n} x_{(i,j)}) = \sum_{j=l}^{l+n} (\sum_{i=k}^{\infty} x_{(i,j)})\}$  then we have

**0 ∈ S.** If  $n=0$  then

$$\begin{aligned} \sum_{i=k}^{\infty} \left( \sum_{j=l}^{l+n} x_{(i,j)} \right) &= \sum_{i=k}^{\infty} \left( \sum_{j=l}^l x_{(i,j)} \right) \\ &= \sum_{i=k}^{\infty} x_{(i,l)} \\ &= \sum_{j=l}^l \left( \sum_{i=k}^{\infty} x_{(i,j)} \right) \\ &= \sum_{j=l}^{l+n} \left( \sum_{i=k}^{\infty} x_{(i,j)} \right) \end{aligned}$$

proving that  $0 \in \mathcal{S}$

**n ∈ S → n + 1 ∈ S.** Let  $\{x_{(i,j)}\}_{(i,j) \in \{k, \dots, \infty\} \times \{l, \dots, l+(n+1)\}} \subseteq [0, \infty]$  then we have

$$\begin{aligned} \sum_{i=k}^{\infty} \left( \sum_{j=l}^{l+n+1} x_{(i,j)} \right) &= \sum_{i=k}^{\infty} \left( \sum_{j=l}^{l+n} x_{(i,j)} + x_{(i,l+n+1)} \right) \\ &\stackrel{17.96}{=} \sum_{i=k}^{\infty} \left( \sum_{j=l}^{l+n} x_{(i,j)} \right) + \sum_{i=k}^{\infty} x_{(i,l+n+1)} \\ &\stackrel{n \in \mathcal{S}}{=} \sum_{j=l}^{l+n} \left( \sum_{i=k}^{\infty} x_{(i,j)} \right) + \sum_{i=k}^{\infty} x_{(i,l+n+1)} \\ &= \sum_{j=1}^{l+n+1} \left( \sum_{i=k}^{\infty} x_{(i,j)} \right) \end{aligned}$$

which proves that  $n+1 \in \mathcal{S}$

Using mathematical induction we have then that  $\mathcal{S} = \mathbb{N}$ . If  $l \leq n$  then  $n-l \in \mathbb{N} = \mathcal{S}$  so that  $\sum_{i=k}^{\infty} (\sum_{j=l}^n x_{(i,j)}) = \sum_{i=k}^{\infty} (\sum_{j=l}^{l+(n-l)} x_{(i,j)}) = \sum_{j=l}^{l+(n-l)} (\sum_{i=k}^{\infty} x_{(i,j)}) = \sum_{j=l}^n (\sum_{i=k}^{\infty} x_{(i,j)})$  proving (2).

3. Define  $\{y_{(i,j)}\}_{(i,j) \in \{k, \dots, n\} \times \{l, \dots, \infty\}}$  by  $y_{(i,j)} = x_{(j,i)}$  then we have

$$\begin{aligned} \sum_{i=k}^n \left( \sum_{j=l}^{\infty} x_{(i,j)} \right) &= \sum_{i=k}^n \left( \sum_{j=l}^{\infty} y_{(j,i)} \right) \\ &\stackrel{(2)}{=} \sum_{j=l}^{\infty} \left( \sum_{i=k}^n y_{(j,i)} \right) \\ &= \sum_{j=l}^{\infty} \left( \sum_{i=k}^n x_{(i,j)} \right) \end{aligned}$$

4. Let  $n \in \{l, \dots, \infty\}$  then

$$\begin{aligned} \sum_{i=k}^n \left( \sum_{j=l}^{\infty} x_{(i,j)} \right) &\stackrel{(3)}{=} \sum_{j=l}^{\infty} \left( \sum_{i=k}^n x_{(i,j)} \right) \\ &\leqslant_{17.97} \sum_{j=l}^{\infty} \left( \sup \left( \left\{ \sum_{i=k}^n x_{(i,j)} \mid n \in \{k, \dots, \infty\} \right\} \right) \right) \\ &= \sum_{j=1}^{\infty} \left( \sum_{i=k}^{\infty} x_{(i,j)} \right) \end{aligned}$$

hence  $\sup(\{\sum_{i=k}^n (\sum_{j=1}^{\infty} x_{(i,j)}) | n \in \{k, \dots, \infty\}\}) \leq \sum_{j=1}^{\infty} (\sum_{i=k}^{\infty} x_{(i,j)})$  proving that

$$\sum_{i=k}^{\infty} \left( \sum_{j=l}^{\infty} x_{(i,j)} \right) \leq \sum_{j=1}^{\infty} \left( \sum_{i=k}^{\infty} x_{(i,j)} \right) \quad (17.53)$$

Further take  $m \in \{l, \dots, \infty\}$  then

$$\begin{aligned} \sum_{j=1}^m \left( \sum_{i=k}^{\infty} x_{(i,j)} \right) &\stackrel{(2)}{=} \sum_{i=k}^{\infty} \left( \sum_{j=l}^m x_{(i,j)} \right) \\ &\leq_{17.97} \sum_{i=k}^{\infty} \left( \sup \left( \left\{ \sum_{j=l}^m x_{(i,j)} | m \in \{l, \dots, \infty\} \right\} \right) \right) \\ &= \sum_{i=k}^{\infty} \left( \sum_{j=l}^{\infty} x_{(i,j)} \right) \end{aligned}$$

hence  $\sup(\{\sum_{j=1}^m (\sum_{i=k}^{\infty} x_{(i,j)}) | m \in \{l, \dots, \infty\}\}) \leq \sum_{i=k}^{\infty} (\sum_{j=l}^{\infty} x_{(i,j)})$  proving that  $\sum_{j=1}^{\infty} (\sum_{i=k}^{\infty} x_{(i,j)}) \leq \sum_{i=k}^{\infty} (\sum_{j=l}^{\infty} x_{(i,j)})$ . Combining this with 17.53 proves that

$$\sum_{i=k}^{\infty} \left( \sum_{j=l}^{\infty} x_{(i,j)} \right) = \sum_{j=1}^{\infty} \left( \sum_{i=k}^{\infty} x_{(i,j)} \right)$$

□

Every finite sum of a denumerable sums can be written as a denumerable sum.

**Theorem 17.104.** Let  $n \in \{l, \dots, \infty\}$  and  $\{x_{(i,j)}\}_{(i,j) \in \{k, \dots, n\} \times [l, \dots, \infty]}$  be a family of non negative extended real then if

$$\sigma: \{p, \dots, \infty\} \rightarrow \{k, \dots, n\} \times \{l, \dots, \infty\}$$

is a bijection (which exist by 5.64) then

$$\sum_{j=l}^{\infty} \left( \sum_{i=k}^n x_{(i,j)} \right) \stackrel{17.103}{=} \sum_{i=k}^n \left( \sum_{j=l}^{\infty} x_{(i,j)} \right) = \sum_{i=p}^{\infty} x_{\sigma(i)} = \sum_{i=p}^{\infty} x_{(\sigma(i)_1, \sigma(i)_2)}$$

**Proof.** Let  $m \in \{p, \dots, \infty\}$  and define  $N_m = \max(\{\sigma(i)_2 | i \in \{p, \dots, m\}\})$  then we have  $\sigma(\{p, \dots, m\}) \subseteq \{k, \dots, n\} \times \{l, \dots, N_m\}$  so that

$$\begin{aligned} \sum_{i=p}^m x_{\sigma(i)} &\stackrel{10.40}{=} \sum_{i \in \{p, \dots, m\}} x_{\sigma(i)} \\ &\stackrel{10.43}{=} \sum_{(i,j) \in \sigma(\{p, \dots, m\})} x_{(i,j)} \\ &\leq_{17.88} \sum_{(i,j) \in \{k, \dots, n\} \times \{l, \dots, N_m\}} x_{(i,j)} \\ &\stackrel{10.47}{=} \sum_{i \in \{k, \dots, n\}} \left( \sum_{j \in \{l, \dots, N_m\}} x_{(i,j)} \right) \end{aligned}$$

$$\begin{aligned}
&\stackrel{10.40}{=} \sum_{i=k}^n \left( \sum_{j=l}^{N_m} x_{(i,j)} \right) \\
&\leqslant_{10.32} \sum_{i=k}^n \sup \left( \left\{ \sum_{j=l}^r x_{(i,j)} \mid m \in \{l, \dots, \infty\} \right\} \right) \\
&= \sum_{i=k}^n \left( \sum_{j=l}^{\infty} x_{(i,j)} \right)
\end{aligned}$$

From the above it follows that  $\sum_{i=p}^{\infty} x_{\sigma(i)} = \sup(\{\sum_{i=p}^m x_{\sigma(i)} \mid m \in \mathbb{N}_0\}) \leq \sum_{i=k}^n (\sum_{j=l}^{\infty} x_{(i,j)})$  so we have

$$\sum_{i=p}^{\infty} x_{\sigma(i)} \leq \sum_{i=k}^n \left( \sum_{j=l}^{\infty} x_{(i,j)} \right) \quad (17.54)$$

For the opposite inequality let  $m \in \{l, \dots, \infty\}$  and take  $N_m = \max(\sigma^{-1}(\{k, \dots, n\} \times \{l, \dots, m\}))$  then if  $i \in \sigma^{-1}(\{k, \dots, n\} \times \{l, \dots, m\}) \Rightarrow p \leq i \leq N_m$  proving that  $\sigma^{-1}(\{k, \dots, n\} \times \{l, \dots, m\}) \subseteq \{p, \dots, N_m\}$ . So  $\{k, \dots, n\} \times \{l, \dots, m\} \stackrel{\sigma \text{ is a bijection}}{=} \sigma(\sigma^{-1}(\{k, \dots, n\} \times \{l, \dots, m\})) \subseteq \sigma(\{p, \dots, N_m\})$ . It follows that

$$\begin{aligned}
&\sum_{i=k}^n \left( \sum_{j=l}^m x_{(i,j)} \right) \stackrel{10.40}{=} \sum_{i \in \{k, \dots, n\}} \left( \sum_{j \in \{l, \dots, m\}} x_{(i,j)} \right) \\
&\stackrel{10.47}{=} \sum_{(i,j) \in \{k, \dots, n\} \times \{l, \dots, m\}} x_{(i,j)} \\
&\leqslant_{17.88} \sum_{(i,j) \in \sigma(\{p, \dots, N_m\})} x_{(i,j)} \\
&\stackrel{10.43}{=} \sum_{i \in \{p, \dots, N_m\}} x_{\sigma(i)} \\
&= \sum_{i=p}^{N_m} x_{\sigma(i)} \\
&\leq \sup \left( \left\{ \sum_{i=p}^m x_{\sigma(i)} \mid m \in \{p, \dots, \infty\} \right\} \right) \\
&= \sum_{i=p}^{\infty} x_{\sigma(i)} \quad (17.55)
\end{aligned}$$

Then we have

$$\begin{aligned}
&\sum_{i=k}^n \left( \sum_{j=l}^{\infty} x_{(i,j)} \right) \stackrel{17.103}{=} \sum_{j=l}^{\infty} \left( \sum_{i=k}^n x_{(i,j)} \right) \\
&= \sup \left( \left\{ \left( \sum_{j=l}^m \left( \sum_{i=k}^n x_{(i,j)} \right) \right) \mid m \in \{l, \dots, \infty\} \right\} \right) \\
&\stackrel{17.103}{=} \sup \left\{ \left( \sum_{i=k}^n \left( \sum_{j=l}^m x_{(i,j)} \right) \right) \mid m \in \{l, \dots, \infty\} \right\} \\
&\leqslant_{17.55} \sum_{i=p}^{\infty} x_{\sigma(i)}
\end{aligned}$$

proving together with 17.54 that

$$\sum_{i=1}^{\infty} x_{\sigma(i)} = \sum_{i=k}^n \left( \sum_{j=l}^{\infty} x_{(i,j)} \right)$$

and thus the theorem.  $\square$

Next we extend the above theorem for arbitrary finite sums of different size.

**Theorem 17.105.** Let  $k \in \mathbb{N}$  and  $\{x_{(i,j)}\}_{(i,j) \in \bigsqcup_{i \in \{k, \dots, \infty\}} \{(i, m_i), \dots, (i, n_i)\}} \subseteq [0, \infty]$  [where  $\forall i \in \{k, \dots, \infty\}$  we have  $m_i \leq n_i$ ] then if

$$\sigma: \{l, \dots, \infty\} \rightarrow \bigsqcup_{i \in \{k, \dots, \infty\}} \{(i, m_i), \dots, (i, n_i)\}$$

is a bijection [which must exists by 5.73] we have

$$\sum_{i=l}^{\infty} x_{\sigma(i)} = \sum_{i=k}^{\infty} \left( \sum_{j=m_i}^{n_i} x_{(i,j)} \right)$$

**Proof.** First let  $i \in \{k, \dots, \infty\}$  and define the bijection  $\tau_i: \{m_i, \dots, n_i\} \rightarrow \{(i, m_i), \dots, (i, n_i)\}$  by  $\tau_i(j) = (i, j)$  so that

$$\begin{aligned} \sum_{(i,j) \in \{(i, m_i), \dots, (i, n_i)\}} x_{(i,j)} &\stackrel{10.43}{=} \sum_{j \in \{m_i, \dots, n_i\}} x_{\tau_i(j)} \\ &= \sum_{j \in \{m_i, \dots, n_i\}} x_{(i,j)} \\ &= \sum_{j=m_i}^{n_i} x_{(i,j)} \end{aligned} \tag{17.56}$$

Let  $m \in \{l, \dots, \infty\}$  define then  $N_m = \max(\{\sigma(i)_1 | i \in \{l, \dots, m\}\})$ . If  $i \in \{l, \dots, m\}$  then as  $\sigma(i) \in \bigsqcup_{i \in \{k, \dots, \infty\}} \{(i, m_i), \dots, (i, n_i)\}$  there exists a  $j \in \{k, \dots, \infty\}$  such that  $\sigma(i) \in \{(j, m_j), \dots, (j, n_j)\}$  so as additional  $\sigma(i)_1 = j \in \{k, \dots, N_m\}$  it follows that

$$\sigma(\{l, \dots, m\}) \subseteq \bigsqcup_{i \in \{k, \dots, N_m\}} \{(i, m_i), \dots, (i, n_i)\} \text{ [a pairwise disjoint union]} \tag{17.57}$$

Next

$$\begin{aligned} \sum_{i=1}^m x_{\sigma(i)} &\stackrel{10.40}{=} \sum_{i \in \{l, \dots, m\}} x_{\sigma(i)} \\ &\stackrel{10.43}{=} \sum_{(i,j) \in \sigma(\{l, \dots, m\})} x_{(i,j)} \\ &\stackrel{\text{17.57 and 17.88}}{=} \sum_{(i,j) \in \bigsqcup_{r \in \{k, \dots, N_m\}} \{(r, m_r), \dots, (r, n_r)\}} x_{(i,j)} \\ &\stackrel{10.45}{=} \sum_{r \in \{k, \dots, N_m\}} \left( \sum_{(i,j) \in \{(r, m_r), \dots, (r, n_r)\}} x_{(i,j)} \right) \\ &= \sum_{r=k}^{N_m} \left( \sum_{(i,j) \in \{(r, m_r), \dots, (r, n_r)\}} x_{(i,j)} \right) \\ &\stackrel{17.56}{=} \sum_{r=k}^{N_m} \left( \sum_{j=m_r}^{n_r} x_{(r,j)} \right) \end{aligned}$$

$$\begin{aligned}
&\stackrel{17.88}{\leq} \sup \left( \left\{ \sum_{r=k}^n \left( \sum_{j=m_r}^{n_r} x_{(r,j)} \right) \mid n \in \{k, \dots, \infty\} \right\} \right) \\
&= \sum_{r=k}^{\infty} \left( \sum_{j=m_r}^{n_r} x_{(r,j)} \right)
\end{aligned}$$

So  $\sup (\{\sum_{i=l}^n x_{\sigma(i)} \mid n \in \{l, \dots, \infty\}\}) \leq \sum_{r=k}^{\infty} (\sum_{j=m_r}^{n_r} x_{(r,j)})$

$$\sum_{i=l}^{\infty} x_{\sigma(i)} \leq \sum_{r=k}^{\infty} \left( \sum_{j=m_r}^{n_r} x_{(r,j)} \right) \tag{17.58}$$

For the opposite inequality take  $m \in \{k, \dots, \infty\}$  and take  $N_m = \max(\sigma^{-1}(\bigsqcup_{i \in \{k, \dots, m\}} \{(i, m_i), \dots, (i, n_i)\}))$ . Let  $i \in \sigma^{-1}(\bigsqcup_{j \in \{k, \dots, m\}} \{(j, m_j), \dots, (j, n_j)\})$  then  $i \in \{l, \dots, N_m\}$  so that  $\sigma^{-1}(\bigsqcup_{j \in \{k, \dots, m\}} \{(j, m_j), \dots, (j, n_j)\}) \subseteq \{l, \dots, N_m\}$ , giving as  $\sigma$  is a bijection that

$$\bigsqcup_{j \in \{k, \dots, m\}} \{(j, m_j), \dots, (j, n_j)\} \subseteq \sigma(\{l, \dots, N_m\}) \tag{17.59}$$

Next

$$\begin{aligned}
&\sum_{r=k}^m \left( \sum_{j=m_r}^{n_r} x_{(r,j)} \right) \stackrel{17.56}{=} \sum_{r=k}^m \left( \sum_{(i,j) \in \{(r, m_r), \dots, (r, n_r)\}} x_{(i,j)} \right) \\
&\stackrel{10.45}{=} \sum_{(i,j) \in \bigsqcup_{r \in \{k, \dots, m\}} \{(r, m_r), \dots, (r, n_r)\}} x_{(i,j)} \\
&\stackrel{\text{17.88 together with 17.59}}{\leq} \sum_{(i,j) \in \sigma(\{l, \dots, N_m\})} x_{(i,j)} \\
&\stackrel{10.43}{=} \sum_{i \in \{l, \dots, N_m\}} x_{\sigma(i)} \\
&= \sum_{i=l}^{N_m} x_{\sigma(i)} \\
&\leq \sup \left( \left\{ \sum_{i=l}^n x_{\sigma(i)} \mid n \in \mathbb{N}_0 \right\} \right) \\
&= \sum_{i=l}^{\infty} x_{\sigma(i)}
\end{aligned}$$

Using the definition of a supremum we have then  $\sum_{i=k}^{\infty} (\sum_{j=m_i}^{n_i} x_{(i,j)}) = \sup (\{\sum_{i=1}^m (\sum_{j=m_i}^{n_i} x_{(i,j)}) \mid m \in \mathbb{N}_0\}) \leq \sum_{i=l}^{\infty} x_{\sigma(i)}$ , this together with 17.58 proves

$$\sum_{i=l}^{\infty} x_{\sigma(i)} = \sum_{i=k}^{\infty} \left( \sum_{j=m_i}^{n_i} x_{(i,j)} \right)$$

□

**Note 17.106.** To simplify or terminology we write  $x_{i,j}$  instead of  $x_{(i,j)}$  from now on, unless want express that  $(i, j)$  is a couple.

The next theorem shows that a denumerable sum of denumerable sums can be written as one single denumerable sum.

**Theorem 17.107.** Let  $k, l \in \mathbb{N}$  and  $\{\{x_{i,j}\}_{i \in \{k, \dots, \infty\}}\}_{j \in \{l, \dots, \infty\}} \subseteq [0, \infty]$  a sequence of sequences of extended non negative reals and  $\sigma: \mathbb{N}_0 \rightarrow \{k, \dots, \infty\} \times \{l, \dots, \infty\}$  a bijection [which exists as by 5.59  $\{k, \dots, \infty\} \times \{l, \dots, \infty\}$  is denumerable] then

$$\sum_{i \in \mathbb{N}_0} x_{\sigma(i)_1, \sigma(i)_2} = \sum_{i=k}^{\infty} \left( \sum_{j=l}^{\infty} x_{i,j} \right) \stackrel{17.103}{=} \sum_{j=l}^{\infty} \left( \sum_{i=k}^{\infty} x_{i,j} \right)$$

**Proof.** Let  $n \in \mathbb{N}_0$  take then  $N_n = \max(\{\sigma(i)_1 | i \in \{1, \dots, n\}\})$ ,  $M_n = \max(\{\sigma(i)_2 | i \in \{1, \dots, n\}\})$  then we have that  $\sigma(\{1, \dots, n\}) \subseteq \{k, \dots, N_n\} \times \{l, \dots, M_n\}$  and thus

$$\begin{aligned} \sum_{i=1}^n x_{\sigma(i)_1, \sigma(i)_2} &\stackrel{10.40}{=} \sum_{i \in \{1, \dots, n\}} x_{\sigma(i)_1, \sigma(i)_2} \\ &\stackrel{10.43}{=} \sum_{(i,j) \in \sigma(\{1, \dots, n\})} x_{i,j} \\ &\stackrel{17.88}{\leq} \sum_{(i,j) \in \{k, \dots, N_n\} \times \{l, \dots, M_n\}} x_{i,j} \\ &\stackrel{10.47}{=} \sum_{i \in \{k, \dots, N_n\}} \left( \sum_{j \in \{l, \dots, M_n\}} x_{i,j} \right) \\ &\stackrel{10.40}{=} \sum_{i=k}^{N_n} \left( \sum_{j=l}^{M_n} x_{i,j} \right) \\ &\leq \sum_{i=k}^{N_n} \sup \left( \left\{ \sum_{j=1}^m x_{i,j} \mid m \in \mathbb{N}_0 \right\} \right) \\ &\stackrel{17.94}{=} \sum_{i=1}^{N_n} \left( \sum_{j=1}^{\infty} x_{i,j} \right) \\ &\leq \sup \left( \left\{ \sum_{i=1}^r \left( \sum_{j=1}^{\infty} x_{i,j} \right) \mid r \in \mathbb{N}_0 \right\} \right) \\ &\stackrel{17.94}{=} \sum_{i=1}^{\infty} \left( \sum_{j=1}^{\infty} x_{i,j} \right) \end{aligned}$$

So  $\sup(\{\sum_{i=1}^n x_{\sigma(i)_1, \sigma(i)_2} \mid n \in \mathbb{N}_0\}) \leq \sum_{i=1}^{\infty} (\sum_{j=1}^{\infty} x_{i,j})$  or using 17.94

$$\sum_{i=1}^{\infty} x_{\sigma(i)_1, \sigma(i)_2} \leq \sum_{i=k}^{\infty} \left( \sum_{j=l}^{\infty} x_{i,j} \right) \quad (17.60)$$

Let now take  $n \in \{k, \dots, \infty\}$ ,  $m \in \{l, \dots, \infty\}$  and take  $N = \max(\sigma^{-1}(\{k, \dots, n\} \times \{l, \dots, m\}))$  then if  $i \in \sigma^{-1}(\{k, \dots, n\} \times \{l, \dots, m\})$  we have  $1 \leq i \leq N$  proving that  $\sigma^{-1}(\{k, \dots, n\} \times \{l, \dots, m\}) \subseteq \{1, \dots, N\}$ . Hence as  $\sigma$  is a bijection it follows that  $\{k, \dots, n\} \times \{l, \dots, m\} \subseteq \sigma(\{1, \dots, N\})$ . So

$$\begin{aligned} \sum_{i=k}^n \left( \sum_{j=l}^m x_{i,j} \right) &\stackrel{10.40}{=} \sum_{i \in \{k, \dots, n\}} \left( \sum_{j \in \{l, \dots, m\}} x_{i,j} \right) \\ &\stackrel{10.47}{=} \sum_{(i,j) \in \{k, \dots, n\} \times \{l, \dots, m\}} x_{i,j} \\ &\stackrel{17.88}{\leq} \sum_{(i,j) \in \sigma(\{1, \dots, N\})} x_{i,j} \end{aligned}$$

$$\begin{aligned}
&\stackrel{10.43}{=} \sum_{i \in \{1, \dots, N\}} x_{\sigma(i)_1, \sigma(i)_2} \\
&= \sum_{i=1}^N x_{\sigma(i)_1, \sigma(i)_2} \\
&\leq \sup \left( \left\{ \sum_{i=1}^r x_{\sigma(i)_1, \sigma(i)_2} \mid r \in \mathbb{N}_0 \right\} \right) \\
&\stackrel{17.94}{=} \sum_{i=1}^{\infty} x_{\sigma(i)_1, \sigma(i)_2}
\end{aligned} \tag{17.61}$$

Further

$$\begin{aligned}
\sum_{i=k}^{\infty} \left( \sum_{j=l}^m x_{i,j} \right) &\stackrel{17.94}{=} \sup \left( \left\{ \sum_{i=1}^n \left( \sum_{j=1}^m x_{i,j} \right) \mid n \in \mathbb{N}_0 \right\} \right) \\
&\leq \sum_{i=1}^{\infty} x_{\sigma(i)_1, \sigma(i)_2}
\end{aligned} \tag{17.62}$$

Next we have

$$\begin{aligned}
\sum_{j=l}^m \left( \sum_{i=k}^{\infty} x_{i,j} \right) &\stackrel{17.103}{=} \sum_{i=k}^{\infty} \left( \sum_{j=l}^m x_{i,j} \right) \\
&\leq \sum_{i=1}^{\infty} x_{\sigma(i)_1, \sigma(i)_2}
\end{aligned} \tag{17.63}$$

So that

$$\begin{aligned}
\sum_{i=k}^{\infty} \left( \sum_{j=l}^{\infty} x_{i,j} \right) &\stackrel{17.103}{=} \sum_{j=l}^{\infty} \left( \sum_{i=k}^{\infty} x_{i,j} \right) \\
&\stackrel{17.94}{=} \sup \left( \left\{ \sum_{j=l}^r \left( \sum_{i=k}^{\infty} x_{i,j} \right) \mid r \in \{l, \dots, \infty\} \right\} \right) \\
&\leq \sum_{i=1}^{\infty} x_{\sigma(i)_1, \sigma(i)_2}
\end{aligned}$$

Finally combining the above with 17.60 proves

$$\sum_{i=1}^{\infty} x_{\sigma(i)_1, \sigma(i)_2} = \sum_{i=1}^{\infty} \left( \sum_{j=1}^{\infty} x_{i,j} \right) \quad \square$$

The following definition allows us to sum from  $-\infty$  to  $+\infty$

**Definition 17.108.** Let  $\{x_i\}_{i \in \mathbb{Z}}$  then  $\sum_{i=-\infty}^{\infty} x_i = \sum_{i=0}^{\infty} x_i + \sum_{i=1}^{\infty} x_{-i}$

# Chapter 18

## The exponential function

### 18.1 Power series

#### 18.1.1 Definition and convergence domains

In this chapter we limit ourselves to series of real numbers so as  $\langle \mathbb{R}, +, 0 \rangle$  is a field we have that finite sums are always defined (see 10.3) So we can use the definition of a convergent (divergent) serie in  $\mathbb{R}$  (see 11.375).

**Theorem 18.1.** Let  $\{x_i\}_{i \in \{k, \dots, \infty\}} \subseteq \mathbb{R}$  then  $\sum_{i=k}^{\infty} x_i$  converges if and only  $\limsup_{n \rightarrow \infty} (\sum_{i=k}^n x_i) = \liminf_{n \rightarrow \infty} (\sum_{i=k}^{\infty} x_i) \in \mathbb{R}$  in which case we have  $\sum_{i=k}^{\infty} x_i = \limsup_{n \rightarrow \infty} (\sum_{i=k}^n x_i) = \liminf_{n \rightarrow \infty} (\sum_{i=k}^{\infty} x_i) \in \mathbb{R}$

**Proof.** This follows from 11.375 and 17.62  $\square$

**Definition 18.2. (Power Serie)** Let  $z \in \mathbb{Z}$  and  $\{a_n\}_{n \in \mathbb{N}} \subseteq \mathbb{C}$  then  $\sum_{i=0}^{\infty} a_i \cdot z^i$  is called a power serie.

**Theorem 18.3.** Let  $\{a_n\}_{n \in \mathbb{N}} \subseteq \mathbb{C}$  be a sequence then  $a_0 = \sum_{i=0}^{\infty} a_i \cdot 0^i$  and  $|a_0| = \sum_{i=0}^{\infty} |a_i \cdot 0^i| = \sum_{i=0}^{\infty} |a_i| \cdot 0^i$ . Let  $n \in \mathbb{N}$  then  $\sum_{i=1}^n |a_i \cdot 0^i| = \sum_{i=1}^n |a_i| \cdot |0|^i = |a_0| \cdot |0|^0 + \sum_{i=1}^{\infty} |a_i| \cdot |0|^i = |a_0| + \sum_{i=1}^{\infty} |a_i| \cdot 0 \cdot 0^{i-1} = |a_0| + 0 \cdot \sum_{i=1}^{\infty} a_i \cdot 0^{i-1} = |a_0|$ .

**Proof.** Let  $n \in \mathbb{N}$  then  $\sum_{i=1}^n a_i \cdot 0^i = a_0 \cdot 0^0 + \sum_{i=1}^{\infty} a_i \cdot 0^i = a_0 + \sum_{i=1}^{\infty} a_i \cdot 0 \cdot 0^{i-1} = a_0 + 0 \cdot \sum_{i=1}^{\infty} a_i \cdot 0^{i-1} = a_0$ .  $\square$

**Definition 18.4.** Let  $\sum_{i=0}^{\infty} a_i \cdot z^i$  be a power serie then we define

1.  $\mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i} = \{z \in \mathbb{C} \mid \sum_{i=0}^{\infty} a_i \cdot z^i \text{ is convergent}\}$ ,  $\mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}$  is called the **convergence domain** of  $\sum_{i=0}^{\infty} a_i \cdot z^i$
2.  $\overline{\mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}} = \{z \in \mathbb{C} \mid \sum_{i=1}^{\infty} a_i \text{ is absolute convergent}\}$ ,  $\overline{\mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}}$  is called the **absolute convergence domain** of  $\sum_{i=0}^{\infty} a_i \cdot z^i$

**Theorem 18.5.** Let  $\sum_{i=0}^{\infty} a_i \cdot z^i$  be a power serie then  $0 \in \overline{\mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}} \subseteq \mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}$

**Proof.** Using 18.3 we have that  $0 \in \overline{\mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}}$ , further as absolute convergence implies convergence (see 11.394) we have that  $\overline{\mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}} \subseteq \mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}$   $\square$

**Theorem 18.6.** Let  $\sum_{i=0}^{\infty} a_i \cdot z^i$  be a power serie then for  $0 \neq z_0 \in \mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}$  we have  $\forall \rho$  with  $0 \leq \rho < |z_0|$  that  $\sum_{i=0}^{\infty} |a_n| \cdot \rho^n$  is convergent.

**Proof.** Let  $0 \leq \rho < |z_0|$  then we have for  $\delta$  either

$\rho = 0$ . Using 18.3 we have that  $\sum_{i=0}^{\infty} |a_n| \cdot \rho^n$  is convergent.

$0 < \rho$ . As  $z_0 \in \mathcal{D}_{\{a_i\}_{i \in \mathbb{N}}}$  we have that  $\sum_{i=0}^{\infty} a_n \cdot z_0^i$  is convergent, so using 11.376 we have that  $\lim_{n \rightarrow \infty} a_n \cdot z_0^n = 0$ , hence by 11.329  $\{a_n \cdot z_0^n\}_{n \in \mathbb{N}}$  is Cauchy and thus bounded (see 11.335) proving

$$\exists K \in \mathbb{R} \text{ such that } \forall i \in \mathbb{N} \models |a_n \cdot z^n| \leq K \quad (18.1)$$

Now  $\forall n \in \mathbb{N}$  we have  $|a_n| \cdot \rho^n \underset{0 \neq |z_0|}{=} |a_n| \cdot |z_0|^n \cdot \left(\frac{\rho}{|z_0|}\right)^n = |a_n z_0^n| \cdot q^n$  with  $q = \left|\frac{\rho}{|z_0|}\right| < 1$  and using 18.1 we have then that

$$\forall n \in \mathbb{N} \models |a_n| \cdot \rho^n \leq K \cdot q^n \text{ where } q < 1 \quad (18.2)$$

Using 11.415 we have that  $\sum_{i=0}^{\infty} \rho^i$  is convergent and thus using 11.381 we have that  $\sum_{i=0}^{\infty} K \cdot \rho^i$  is convergent. Using 18.2 on the above allows us to apply dominant convergence (see 11.393) to prove that  $\sum_{i=0}^{\infty} |a_n| \cdot \rho^n$  is convergent.  $\square$

**Corollary 18.7.** Let  $\sum_{i=0}^{\infty} a_i \cdot z^i$  be a power serie,  $0 \neq z_0 \in \mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}$  then  $B_{||}(0, |z_0|) \subseteq \overline{\mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}} \subseteq \mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}$

**Proof.** Take  $z \in B_{||}(0, |z_0|)$  then for  $z$  we have either

$z = 0$ . then we have  $z = 0 \in \overline{\mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}}$

$z \neq 0$ . then  $0 < |z| < |z_0| \Rightarrow \exists \rho \vdash 0 < |z| < \rho < |z_0|$ . Using the previous theorem we have that  $\sum_{i=0}^{\infty} |a_n| \cdot \rho^n$  is convergent. Finally as  $\forall i \in \mathbb{N}$  we have  $|a_i \cdot z^i| = |a_i| \cdot |z|^i < |a_i| \cdot \rho^i$  it follows from dominant convergence (see 11.393) that  $\sum_{i=1}^{\infty} |a_i \cdot z^i|$  is convergent, hence  $\sum_{i=0}^{\infty} a_i \cdot z^i$  is absolute convergent and thus  $z \in \overline{\mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}}$ .  $\square$

From the above it follows that  $0 \in [0, \infty[ \cap \mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i} \Rightarrow \emptyset \neq [0, \infty[ \cap \mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}$  which allows us to define the following

**Definition 18.8.** Let  $\sum_{i=0}^{\infty} a_i \cdot z^i$  be a power serie then  $R_{\sum_{i=0}^{\infty} a_i \cdot z^i} = \sup(\mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i} \cap [0, \infty[)$  is called the **conversion radius**

The above corollary is then summarized as follows

**Corollary 18.9.** Let  $\{a_i\}_{i \in \mathbb{N}} \subseteq \mathbb{C}$  be a sequence then we have for  $R_{\sum_{i=0}^{\infty} a_i \cdot z^i}$

1. If  $0 \leq R_{\sum_{i=0}^{\infty} a_i \cdot z^i} < \infty$  then  $B_{||}(0, R_{\sum_{i=0}^{\infty} a_i \cdot z^i}) \subseteq \overline{\mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}} \subseteq \mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i} \subseteq \overline{B_{||}(0, R_{\sum_{i=0}^{\infty} a_i \cdot z^i})}$
2. If  $R_{\sum_{i=0}^{\infty} a_i \cdot z^i} = \infty$  then  $\overline{\mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}} = \mathbb{C} = \mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}$
3. If  $\overline{\mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}} = \mathbb{C} = \mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}$  then  $R_{\sum_{i=0}^{\infty} a_i \cdot z^i} = \infty$  and  $B_{||}(0, R_{\sum_{i=0}^{\infty} a_i \cdot z^i}) = \overline{B_{||}(0, R_{\sum_{i=0}^{\infty} a_i \cdot z^i})}$

**Proof.**

1. For  $0 \leq R_{\sum_{i=0}^{\infty} a_i \cdot z^i} < \infty$  we have either

$R_{\sum_{i=0}^{\infty} a_i \cdot z^i} = 0$ . if  $0 \neq z \in \mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}$  then  $0 < |z|$  so there exists a  $\rho \in \mathbb{R}$  such that  $0 < r < |z|$  proving that  $0 < r \in [0, \infty[ \cap B_{||}(0, |z|) \subseteq_{18.7} [0, \infty[ \cap \mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}$ . From this it follows that  $R_{\sum_{i=0}^{\infty} a_i \cdot z^i} = \sup(\mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i} \cap [0, \infty[) > 0$  a contradiction. So we have that  $B_{||}(0, R_{\sum_{i=0}^{\infty} a_i \cdot z^i}) = \emptyset \subseteq \overline{\mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}} \subseteq \mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i} = \{0\} = \overline{\mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}} \subseteq \mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i} = \{0\} = \overline{B_{||}(0, R_{\sum_{i=0}^{\infty} a_i \cdot z^i})}$ .

$0 < R_{\sum_{i=0}^{\infty} a_i \cdot z^i} < \infty$ . Let  $z \in B_{||}(0, R_{\sum_{i=0}^{\infty} a_i \cdot z^i})$  then we have either

$z = 0$ . then  $z = 0 \in \overline{\mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}}$

$z \neq 0$ . then  $0 < |z| < R_{\sum_{i=0}^{\infty} a_i \cdot z^i} = \sup(\mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i} \cap [0, \infty[)$  so there exists a  $z' \in \mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i} \cap [0, \infty[$  such that  $0 < |z| < z' \leq R_{\sum_{i=0}^{\infty} a_i \cdot z^i}$ . Hence  $z \in B_{||}(0, z') = B_{||}(0, |z'|) \subseteq_{18.7} \overline{\mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}}$

so in all cases we have  $z \in B_{||}(0, R_{\{a_i\}_{i \in \mathbb{N}}})$  proving that

$$B_{||}(0, R_{\{a_i\}_{i \in \mathbb{N}}}) \subseteq \overline{\mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}} \subseteq \mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i} \quad (18.3)$$

If  $z \notin \overline{B_{||}(0, R_{\sum_{i=0}^{\infty} a_i \cdot z^i})}$  then  $R_{\sum_{i=0}^{\infty} a_i \cdot z^i} < |z|$  then  $\exists r \in [0, \infty[$  such that  $R_{\sum_{i=0}^{\infty} a_i \cdot z^i} < r < |z|$ . If now  $z \in \mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}$  then by 18.7 we have that  $r \in \mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i} \cap [0, \infty[$  hence  $r \leq \sup(\mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i} \cap [0, \infty[) = R_{\sum_{i=0}^{\infty} a_i \cdot z^i} < r$  a contradiction. So we must have that  $\mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i} \subseteq \overline{B_{||}(0, R_{\sum_{i=0}^{\infty} a_i \cdot z^i})}$  proving using 18.3 that

$$B_{||}(0, R_{\sum_{i=0}^{\infty} a_i \cdot z^i}) \subseteq \overline{\mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}} \subseteq \mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i} \subseteq \overline{B_{||}(0, R_{\sum_{i=0}^{\infty} a_i \cdot z^i})}$$

2. If  $z \in \mathbb{C}$  then as  $R_{\sum_{i=0}^{\infty} a_i \cdot z^i} = \infty$  there exists by 17.14 a  $r \in \mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i} \cap [0, \infty[$  such that  $|z| < r = |r| \Rightarrow z \in B_{||}(0, |r|) \subseteq_{18.7} \overline{\mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}}$  proving that  $\mathbb{C} \subseteq \overline{\mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}} \subseteq \mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i} \subseteq \mathbb{C}$  giving

$$\overline{\mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}} = \mathbb{C} = \mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}$$

3. As by (1) we have  $B_{||}(0, R_{\sum_{i=0}^{\infty} a_i \cdot z^i}) \subseteq \mathbb{C} \subseteq \overline{B_{||}(0, R_{\sum_{i=0}^{\infty} a_i \cdot z^i})}$  we must have that  $\overline{B_{||}(0, R_{\sum_{i=0}^{\infty} a_i \cdot z^i})} = \mathbb{C}$ . If now  $R_{\sum_{i=0}^{\infty} a_i \cdot z^i} < \infty$  then as  $R_{\sum_{i=0}^{\infty} a_i \cdot z^i} + 1 \in \mathbb{C} = \overline{B_{||}(0, R_{\sum_{i=0}^{\infty} a_i \cdot z^i})}$  so that  $R_{\sum_{i=0}^{\infty} a_i \cdot z^i} < R_{\sum_{i=0}^{\infty} a_i \cdot z^i} + 1 \leq R_{\sum_{i=0}^{\infty} a_i \cdot z^i}$  giving the contradiction  $R_{\sum_{i=0}^{\infty} a_i \cdot z^i} < R_{\sum_{i=0}^{\infty} a_i \cdot z^i}$  so we must have that  $R_{\sum_{i=0}^{\infty} a_i \cdot z^i} = \infty$  which means by (2) that also  $\mathbb{C} = \mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}$ .  $\square$

We can simplify the above corollary if we extend the definition of a open/closed ball in  $\mathbb{C}$  so that the radius can be infinite.

**Definition 18.10.** Let  $\delta \in [0, \infty]$  and  $z_0 \in \mathbb{C}$  then  $B_{||}(z_0, \delta) = \{z \in \mathbb{C} \mid |z - z_0| < \rho\}$  and  $B_{||}(z_0, \delta) = \{z \in \mathbb{C} \mid |z - z_0| \leq \delta\}$  then we have

1. If  $\delta = 0$  then  $B_{||}(z_0, 0) = \emptyset$  a open set and  $\overline{B_{||}(z_0, \delta)} = \{z_0\}$  a closed set
2. If  $\delta < \infty$  then  $B_{||}(z_0, \delta)$  is the classical open ball and  $\overline{B_{||}(z_0, \delta)}$  is the classical closed ball.
3. If  $\delta = \infty$  then if  $z \in \mathbb{C}$  we have  $|z - z_0| < |z - z_0| + 1 < \infty$  so that  $B_{||}(z_0, \delta) = \mathbb{C} = \overline{B_{||}(z_0, \delta)}$  a open/closed set

we have then the following

**Lemma 18.11.** Let  $\delta, \delta' \in [0, \infty]$  and  $z_0 \in \mathbb{C}$  then if  $B_{||}(z_0, \delta) \subseteq \overline{B_{||}(z_0, \delta')}$  then  $\delta \leq \delta'$

**Proof.** We have the following possibilities to consider for  $\delta'$

$\delta' \in ]0, \infty[$ . assume then that  $\delta' < \delta$  then there exists a  $\rho \in \mathbb{R}$  such that  $0 < \delta' < \rho < \delta$  hence  $\delta' < |\rho| = \rho < \delta \Rightarrow \rho \in B_{||}(z_0, \delta) \Rightarrow \rho \in B_{||}(z_0, \delta') \Rightarrow \delta' < \rho \leq \delta' \Rightarrow \delta' < \delta'$  a contradiction so we must have that  $\delta \leq \delta'$

$\delta' = 0$ . then  $B_{||}(z_0, \delta) \subseteq \overline{B_{||}(z_0, \delta')} = \{0\}$  if now  $0 < \delta$  then there exists a  $0 < \rho < \delta \Rightarrow |\rho| = \rho < \delta \Rightarrow \rho \in B_{||}(z_0, \delta) \subseteq \{0\} \Rightarrow 0 < \rho = 0$  a contradiction so we must have  $\delta \leq 0 = \delta'$

$\delta' = \infty$ . then  $\delta \leq \infty = \delta'$

So in all cases we have  $\delta \leq \delta'$ .  $\square$

The above corollary condences then to

**Corollary 18.12.** Let  $\sum_{i=0}^{\infty} a_i \cdot z^i$  be a power serie then we have that  $B_{||}(0, R_{\sum_{i=0}^{\infty} a_i \cdot z^i}) \subseteq \overline{\mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}} \subseteq \mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i} \subseteq \overline{B_{||}(0, R_{\sum_{i=0}^{\infty} a_i \cdot z^i})}$  where  $B_{||}(0, R_{\sum_{i=0}^{\infty} a_i \cdot z^i})$  is open and  $\overline{B_{||}(0, R_{\sum_{i=0}^{\infty} a_i \cdot z^i})}$  is closed.

### 18.1.2 Convergence criteria

**Lemma 18.13.** Let  $n \in \mathbb{N}$  and  $\{z_i\}_{i \in \{0, \dots, n\}} \subseteq \mathbb{C}$ ,  $\{\alpha_i\}_{i \in \{0, \dots, n\}} \subseteq \mathbb{C}$  and  $\{s_i\}_{i \in \{0, \dots, n\}}$  defined by  $s_i = \sum_{j=0}^i z_j$  then  $\sum_{i=0}^n \alpha_i \cdot z_i = \sum_{i=0}^{n-1} s_i \cdot (\alpha_i - \alpha_{i+1}) + \alpha_n \cdot s_n$

**Proof.** We prove this by induction so let  $\mathcal{S} = \{n \in \mathbb{N} \mid \{z_i\}_{i \in \{0, \dots, n\}} \subseteq \mathbb{C}, \{\alpha_i\}_{i \in \{0, \dots, n\}} \subseteq \mathbb{C} \text{ and } \{s_i\}_{i \in \{0, \dots, n\}}$  defined by  $s_i = \sum_{j=0}^i z_j$  then  $\sum_{i=0}^n \alpha_i \cdot z_i = \sum_{i=0}^{n-1} s_i \cdot (\alpha_i - \alpha_{i+1}) + \alpha_n \cdot s_n\}$  then we have

$1 \in \mathcal{S}$ . for

$$\begin{aligned} \sum_{i=0}^{n-1} s_i \cdot (\alpha_i - \alpha_{i+1}) + \alpha_n \cdot s_n &\stackrel{n=1}{=} \sum_{i=0}^0 s_i \cdot (\alpha_i - \alpha_{i+1}) + \alpha_1 \cdot s_1 \\ &= s_0 \cdot (\alpha_0 - \alpha_1) + \alpha_1 \cdot s_1 \\ &= z_0 \cdot (\alpha_0 - \alpha_1) + \alpha_1 \cdot (z_0 + z_1) \\ &= \alpha_0 \cdot z_0 + \alpha_1 \cdot z_1 \\ &= \sum_{i=0}^1 \alpha_i \cdot z_i \end{aligned}$$

proving that  $1 \in \mathcal{S}$

$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$ . We have

$$\begin{aligned}
 \sum_{i=0}^{n+1} \alpha_i \cdot z_i &= \left( \sum_{i=0}^n \alpha_i \cdot z_i \right) + \alpha_{n+1} \cdot z_{n+1} \\
 &\stackrel{n \in \mathcal{S}}{=} \left( \sum_{i=0}^{n-1} s_i \cdot (\alpha_i - \alpha_{i+1}) + \alpha_n \cdot s_n \right) + \alpha_{n+1} \cdot z_{n+1} \\
 &= \left( \sum_{i=0}^{n-1} s_i \cdot (\alpha_i - \alpha_{i+1}) \right) + (\alpha_n \cdot s_n - \alpha_{n+1} \cdot s_n) + (\alpha_{n+1} \cdot z_{n+1} + \alpha_{n+1} \cdot s_n) \\
 &= \left( \sum_{i=0}^n s_i \cdot (\alpha_i - \alpha_{i+1}) \right) + (\alpha_{n+1} \cdot z_{n+1} + \alpha_{n+1} \cdot s_n) \\
 &= \left( \sum_{i=0}^n s_i \cdot (\alpha_i - \alpha_{i+1}) \right) + \alpha_{n+1} \cdot \left( z_{n+1} + \sum_{i=0}^n s_i \right) \\
 &= \left( \sum_{i=0}^n s_i \cdot (\alpha_i - \alpha_{i+1}) \right) + \alpha_{n+1} \cdot s_{n+1}
 \end{aligned}$$

proving that  $n+1 \in \mathcal{S}$

□

**Lemma 18.14. (Inequality of Abel)** Let  $n \in \mathbb{N}$   $\{z_i\}_{i \in \{0, \dots, n\}} \subseteq \mathbb{C}$ ,  $\{\alpha_i\}_{i \in \{0, \dots, n\}} \subseteq [0, \infty[$  such that  $\forall i \in \{0, \dots, n-1\}$  we have  $\alpha_{i+1} \leq \alpha_i$ ,  $\sigma = \sum_{i=0}^n \alpha_i \cdot z_i$ ,  $\{s_k\}_{k \in \{0, \dots, n\}}$  defined by  $s_k = \sum_{i=0}^k z_i$  and  $\mu \in [0, \infty[$ , then if  $\forall k \in \{0, \dots, n\}$  we have  $|s_k| \leq \mu$  it follows that  $|\sigma| \leq \alpha_0 \cdot \mu$

**Proof.** We have for  $n$  two cases to consider

$$\begin{aligned}
 n = 0. \text{ then } |\sigma| &= |\sum_{i=0}^n \alpha_i \cdot z_i| = |\alpha_0 \cdot z_0| = |\alpha_0 \cdot \sum_{i=0}^0 z_i| = |\alpha_0 \cdot s_0| = |\alpha_0| \cdot |s_0| \stackrel{0 \leq \alpha_0}{=} \alpha_0 \cdot |s_0| \leq \alpha_0 \cdot \mu \\
 \text{proving that } |\sigma| &\leq \alpha_0 \cdot \mu
 \end{aligned}$$

$n \neq 0$ . then  $n \in \mathbb{N}$  so we can use the previous lemma

$$\begin{aligned}
 |\sigma| &= \left| \sum_{i=0}^n \alpha_i \cdot z_i \right| \\
 &\stackrel{\text{previous lemma}}{=} \left| \sum_{i=0}^{n-1} s_i \cdot (\alpha_i - \alpha_{i+1}) + \alpha_n \cdot s_n \right| \\
 &\leq \sum_{i=0}^{n-1} |s_i \cdot (\alpha_i - \alpha_{i+1})| + |\alpha_n \cdot s_n| \\
 &\stackrel{\alpha_{i+1} \leq \alpha_i, 0 \leq \alpha_n}{=} \sum_{i=0}^{n-1} |s_i| \cdot (\alpha_i - \alpha_{i+1}) + \alpha_n \cdot s_n \\
 &\leq \sum_{i=0}^{n-1} \mu \cdot (\alpha_i - \alpha_{i+1}) + \alpha_n \cdot s_n \\
 &= \mu \cdot \sum_{i=0}^{n-1} (\alpha_i - \alpha_{i+1}) + \alpha_n \cdot \mu \\
 &= \mu \cdot \left( \sum_{i=0}^{n-1} (\alpha_i - \alpha_{i+1}) + \alpha_n \right) \\
 &\stackrel{10.28}{=} \mu \cdot (\alpha_0 - \alpha_n + \alpha_n) \\
 &= \mu \cdot \alpha_0
 \end{aligned}$$

which proves the theorem.  $\square$

The above lemma will be used to prove the convergence criteria of Abel Dirichlet

**Theorem 18.15. (Abel Dirichlet)** Let  $\{z_i\}_{i \in \mathbb{N}} \subseteq \mathbb{C}$ ,  $\{s_n\}_{n \in \mathbb{N}}$  defined by  $s_n = \sum_{i=0}^n z_i$ ,  $\{\varepsilon_i\}_{i \in \mathbb{N}} \subseteq [0, \infty[$  such that  $\forall i \in \mathbb{N}$  we have  $\varepsilon_{i+1} \leq \varepsilon_i$  then we have

1. If  $\exists C \in \mathbb{R}$  such that  $\forall i \in \mathbb{N}$  we have  $|s_i| \leq C$  and  $\lim_{i \rightarrow \infty} \varepsilon_i = 0$  then  $\sum_{i=0}^{\infty} \varepsilon_i \cdot z_i$  is convergent
2. If  $\sum_{i=0}^{\infty} z_i$  is convergent then  $\sum_{i=0}^{\infty} \varepsilon_i \cdot z_i$  is convergent

**Proof.**

1. Given  $n, m \in \mathbb{N} \vdash n \leq m$  define  $\{x_i^{(n)}\}_{i \in \{0, \dots, m-n\}}$  by  $x_i^{(n)} = z_{n+i}$ , define  $\{s_k^{(n)}\}_{k \in \{0, \dots, m-n\}}$  by  $s_k^{(n)} = \sum_{i=0}^k x_i^{(n)}$  and define  $\{\alpha_i\}_{i \in \{0, \dots, m-n\}}$  by  $\alpha_i^{(n)} = \varepsilon_{n+i}$  then we have that

$$\forall i \in \{0, \dots, m-n-1\} \models \alpha_{i+1}^{(n)} \leq \alpha_i^{(n)} \text{ and } \{\{\alpha_i\}_{i \in \{0, \dots, m-n\}}\} \subseteq [0, \infty[ \quad (18.4)$$

Further giving  $k \in \{0, \dots, m-n\}$  we have  $s_k^{(n)} = \sum_{i=0}^k x_i^{(n)} = \sum_{i=0}^{n+k} z_{n+i} = \sum_{i=n}^{n+k} z_i \stackrel{10.17}{=} \sum_{i=0}^{n+k} z_i - \sum_{i=0}^{n-1} z_i = s_{n+k} - s_{n-1}$  which in addition to  $|s_{n+k} - s_{n-1}| \leq |s_{n+k}| + |s_n| \leq 2 \cdot C$  proves that

$$\forall k \in \{0, \dots, m-n\} \text{ we have } s_k^{(n)} = s_{n+k} - s_{n-1} \text{ and } |s_k^{(n)}| \leq 2 \cdot C \quad (18.5)$$

Using the the inequality of Abele (see 18.14) we have then that  $|\sum_{i=0}^{m-n} \alpha_i \cdot x_i^{(n)}| \leq \alpha_0 \cdot (2 \cdot C)$ , and as  $\sum_{i=0}^{m-n} \alpha_i \cdot x_i^{(n)} = \sum_{i=0}^{m-n} \varepsilon_{n+i} \cdot z_{n+i} = \sum_{i=n}^m \varepsilon_i \cdot z_i$  we have that

$$\left| \sum_{i=n}^m \varepsilon_i \cdot z_i \right| \leq 2 \cdot C \cdot \varepsilon_n \quad (18.6)$$

Take now  $\varepsilon > 0$  then as  $\lim_{n \rightarrow \infty} \varepsilon_n = 0$  there exists a  $N \in \mathbb{N}$  such that  $\forall n \geq N$  we have that  $\varepsilon_n = |\varepsilon_n - 0| < \frac{\varepsilon}{2 \cdot C}$ . So we conclude using 18.6 that

$$\forall \varepsilon > 0 \text{ there } \exists N \in \mathbb{N} \text{ such that } \left| \sum_{i=N}^m \varepsilon_i \cdot z_i \right| < \varepsilon$$

which by the Cauchy condition for series (see 11.377) and the fact that  $\langle \mathbb{C}, \|\cdot\| \rangle$  proves that  $\sum_{i=0}^{\infty} \varepsilon_i \cdot z_i$  converges.

2. From the hypothese it follows that  $\{\varepsilon_i\}_{i \in \mathbb{N}}$  is a decreasing sequence, so using 17.49 we have that  $\alpha = \lim_{i \rightarrow \infty} \varepsilon_i$  exists so that  $\lim_{i \rightarrow \infty} (\varepsilon_i - \alpha) = 0$ . From (1) it follows then that  $\sum_{i=1}^n (\varepsilon_i - \alpha) \cdot z_i$  converges. As by the hypothese  $\sum_{i=0}^{\infty} z_i$  converges  $\stackrel{11.381}{\Rightarrow} \sum_{i=0}^{\infty} \alpha \cdot z_i$  converges we have then by 11.381 that  $\sum_{i=0}^{\infty} ((\varepsilon_i - \alpha) \cdot z_i + \alpha \cdot z_i)$  converges proving that  $\sum_{i=0}^{\infty} \varepsilon_i \cdot z_i$   $\square$

**Lemma 18.16.** Let  $\{x_i\}_{i \in \{k, \dots, \infty\}}$ ,  $\{y_i\}_{i \in \{k, \dots, \infty\}} \subseteq [0, \infty[$  be such that  $\forall i \in \{k, \dots, \infty\}$  we have  $\frac{x_{i+1}}{y_{i+1}} \leq \frac{x_i}{y_i}$  then if  $\sum_{i=k}^{\infty} y_i$  is convergent then  $\sum_{i=k}^{\infty} x_i$  is convergent.

**Proof.** Take  $\alpha = \frac{y_k}{x_k}$  then by the hypothese we have  $\forall i \in \{k, \dots, \infty\}$  that  $\frac{x_k}{y_k} \leq \alpha \stackrel{0 < y_k}{\Rightarrow} x_k \leq \alpha \cdot y_k$   $\stackrel{0 < x_k}{\Rightarrow} |x_k| \leq \alpha \cdot y_k$ . As  $\sum_{i=k}^{\infty} \alpha \cdot y_i$  is convergent (see 11.381) we have by the dominant convergence (see 11.393) that  $\sum_{i=k}^{\infty} x_i$  is convergent.  $\square$

**Theorem 18.17. (d'Alembert)** Let  $\{x_i\}_{i \in \{k, \dots, \infty\}} \subseteq [0, \infty[$  and take  $\alpha = \liminf_{i \rightarrow \infty} \frac{x_{i+1}}{x_i}$  and  $\beta = \limsup_{i \rightarrow \infty} \frac{x_{i+1}}{x_i}$  then we have

1. If  $\beta < 1$  then  $\sum_{i=k}^{\infty} x_i$  is convergent

2. If  $1 < \alpha$  then  $\sum_{i=k}^{\infty} x_i$  is divergent (meaning  $\sum_{i=k}^{\infty} x_i$  is not convergent)

**Proof.**

- As  $\beta < 1$  there exists a  $\rho$  such that  $0 \leq \beta < \rho < 1$  then we have  $\inf \left( \left\{ \sup \left( \left\{ \frac{x_{i+1}}{x_i} \mid i \in \{n, \dots, \infty\} \right\} \right) \mid n \in \{k, \dots, \infty\} \right\} \right) < \rho$  so there exists a  $n \in \{k, \dots, \infty\}$  such that  $\sup \left( \left\{ \frac{x_{i+1}}{x_i} \mid i \in \{n, \dots, \infty\} \right\} \right) < \rho$  hence

$$\forall i \in \{n, \dots, \infty\} \text{ that } \frac{x_{i+1}}{x_i} < \rho \stackrel{0 < \rho}{\equiv} \frac{\rho^{i+1}}{\rho^i} \quad (18.7)$$

Now as  $\rho < 1$  we have by 11.415 that  $\sum_{i=1}^{\infty} \rho^i$  is convergent and thus using 11.379 we have

$$\sum_{i=n}^{\infty} \rho^i \quad (18.8)$$

Using lemma 18.16 we have by 18.7 and 18.8 that  $\sum_{i=n}^{\infty} x_i$  is convergent. Applying 11.379 again we have that  $\sum_{i=k}^{\infty} x_i$  is convergent.

- If  $\alpha < 1$  then we have  $1 < \sup \left( \left\{ \inf \left( \left\{ \frac{x_{i+1}}{x_i} \mid i \in \{n, \dots, \infty\} \right\} \right) \mid n \in \{k, \dots, \infty\} \right\} \right)$  hence there exists a  $n \in \{k, \dots, \infty\}$  such that  $1 < \inf \left( \left\{ \frac{x_{i+1}}{x_i} \mid i \in \{n, \dots, \infty\} \right\} \right)$  hence  $\forall i \in \{n, \dots, \infty\}$  we have that  $1 < \frac{x_{i+1}}{x_i} \Rightarrow x_i < x_{i+1}$  so  $\{x_i\}_{i \in \{n, \dots, \infty\}}$  is a increasing sentence. Using 17.71 we have that  $\lim_{n \rightarrow \infty} x_i = \sup(\{x_i \mid i \in \{n, \dots, \infty\}\}) \geq x_k > 0 \Rightarrow \lim_{i \rightarrow \infty} x_i \neq 0$ . If now  $\sum_{k=1}^{\infty} x_i$  is convergent then we have by 11.376 that  $\lim_{i \rightarrow \infty} x_i = 0$  a contradiction proving that  $\sum_{k=1}^{\infty} x_i$  is divergent.  $\square$

**Corollary 18.18.** Let  $z \in \mathbb{C}$  then  $\sum_{n=0}^{\infty} \frac{|z|^n}{n!}$  and  $\sum_{n=0}^{\infty} \frac{z^n}{n!}$  converges and  $\sum_{n=0}^{\infty} \frac{0^i}{i!} = 1$ . So we have  $B_{\mathbb{C}}(0, R_{\sum_{i=0}^{\infty} \frac{z^i}{i!}}) = B_{\mathbb{C}}(0, R_{\sum_{i=0}^{\infty} \frac{z^i}{i!}}) = \mathcal{D}_{\sum_{n=0}^{\infty} \frac{z^n}{n!}} = \mathbb{C} = \overline{\mathcal{D}_{\sum_{n=0}^{\infty} \frac{z^n}{n!}}}$

**Proof.** Let  $z \in \mathbb{C}$  then we have either

$z = 0$ . then using 18.3 we have that  $\sum_{n=0}^{\infty} \frac{|z|^n}{n!}$  and  $\sum_{n=0}^{\infty} \frac{z^n}{n!}$  converges

$z \neq 0$ . then we have  $\forall n \in \{0, \dots, \infty\}$  that  $0 < \frac{|z|^n}{n!}$  and  $\frac{\frac{|z|^{n+1}}{(n+1)!}}{\frac{|z|^n}{n!}} = \frac{|z|}{n+1}$ . Using the consequence of the Archimedean property of the reals (see 8.62) there exists a  $k \in \{0, \dots, \infty\}$  such that  $2 \cdot |z| < k$  so that  $\forall n \geq k$  we have  $|z| < k \Rightarrow \frac{|z|}{k} < \frac{1}{2}$  proving that  $\sup \left( \left\{ \frac{\frac{|z|^{n+1}}{(n+1)!}}{\frac{|z|^n}{n!}} \mid n \in \{k, \dots, \infty\} \right\} \right) \leq \frac{1}{2}$  hence  $\inf \left( \left\{ \sup \left( \left\{ \frac{\frac{|z|^{n+1}}{(n+1)!}}{\frac{|z|^n}{n!}} \mid n \in \{k, \dots, \infty\} \right\} \right) \mid n \in \{0, \dots, \infty\} \right\} \right) \leq \inf \left( \left\{ \sup \left( \left\{ \frac{\frac{|z|^{n+1}}{(n+1)!}}{\frac{|z|^n}{n!}} \mid n \in \{k, \dots, \infty\} \right\} \right) \mid n \in \{0, \dots, \infty\} \right\} \right) \leq \frac{1}{2} < 1$ . Applying then d'Alembert (see 18.17) we conclude that  $\sum_{i=0}^{\infty} \frac{|z|^i}{i!}$  converges. Using dominant convergence (see 11.393) we have then that  $\sum_{n=0}^{\infty} \frac{z^n}{n!}$  is convergent.  $\square$

Using the above corollary the following definition is well defined

**Definition 18.19. (Exponential Function)**  $\exp: \mathbb{C} \rightarrow \mathbb{C}$  is defined by  $z \rightarrow \sum_{n=0}^{\infty} \frac{z^n}{n!}$

### 18.1.3 Differentiation of power series

**Definition 18.20.** Let  $\{a_i\}_{i \in \mathbb{N}}$  be a sequence and  $D_{\sum_{i=0}^{\infty} a_i \cdot z^i}$  its convergence domain then we can define  $\sum_{n=0}^{\infty} a_n \cdot (.)^n: D_{\{a_i\}_{i \in \mathbb{N}}} \rightarrow \mathbb{C}$  by  $z \rightarrow (\sum_{n=0}^{\infty} a_n \cdot (.)^n)(z) = \sum_{n=0}^{\infty} a_n \cdot z^n$

If  $0 < R_{\{a_i\}_{i \in \mathbb{N}}}$  then  $0 \in B_{||}(0, R_{\sum_{i=0}^{\infty} a_i \cdot z^i}) \subseteq D_{\{a_i\}_{i \in \mathbb{N}}}$  so that it makes sense to talk about the derivative of  $\sum_{n=0}^{\infty} a_n \cdot z^n$  at 0 on  $B_{||}(0, R_{\{a_i\}_{i \in \mathbb{N}}})$ . If we would differentiate the different terms of  $\sum_{i=0}^{\infty} a_n \cdot z^n$  in an informal way we would have  $\sum_{i=0}^{\infty} n \cdot a_n \cdot z^{n-1} = \sum_{i=1}^{\infty} n \cdot a_n \cdot z^{n-1} = \sum_{i=0}^{\infty} (n+1) \cdot a_{n+1} \cdot z^n$ . Of course we must find out that  $\sum_{i=0}^{\infty} (n+1) \cdot a_{n+1} \cdot z^n$  is convergent and what its convergence domain is before even proving that this is the derivative of  $\sum_{i=0}^{\infty} a_i \cdot z^i$ .

**Definition 18.21.** Let  $\sum_{i=0}^{\infty} a_i \cdot z^i$  be a power serie then  $\sum_{i=0}^{\infty} (i+1) \cdot a_{i+1} \cdot z^i$  is called the *derivated power serie of*  $\sum_{i=0}^{\infty} a_i \cdot z^i$

**Lemma 18.22.** Let  $\sum_{i=0}^{\infty} a_i \cdot z^i$  be a powerserie then

1.  $\mathcal{D}_{\sum_{i=0}^{\infty} (i+1) \cdot a_{i+1} \cdot z^i} \subseteq \mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}$
2. If  $0 < R_{\sum_{i=0}^{\infty} a_i \cdot z^i}$  then  $R_{\sum_{i=0}^{\infty} (i+1) \cdot a_{i+1} \cdot z^i} = R_{\sum_{i=0}^{\infty} a_i \cdot z^i}$

**Proof.**

1. Let  $z \in \mathcal{D}_{\sum_{i=0}^{\infty} (i+1) \cdot a_{i+1} \cdot z^i}$  then  $\sum_{i=0}^{\infty} (i+1) \cdot a_{i+1} \cdot z^i$  is convergent. Now for the sequence  $\left\{ \frac{1}{i+1} \right\}_{i \in \mathbb{N}}$  we have that  $\forall i \in \mathbb{N}$  that  $\frac{1}{(i+1)+1} < \frac{1}{(i+1)}$  so using Abel Dirichlet (see 18.15) we have that  $\sum_{i=0}^{\infty} a_{i+1} \cdot z^i = \sum_{i=0}^{\infty} \frac{i+1}{i+1} \cdot a_{i+1} \cdot z^i$  is convergent. Using 11.381 we have then that  $\sum_{i=0}^{\infty} a_{i+1} \cdot z^{i+1} = \sum_{i=0}^{\infty} z \cdot a_{i+1} \cdot z^i$  is convergent. As  $a_0 \cdot z^0 + \sum_{i=0}^{\infty} a_{i+1} \cdot z^{i+1} \stackrel{11.378}{=} a_0 \cdot z^0 + \sum_{i=1}^{\infty} a_i \cdot z^i \stackrel{11.379}{=} \sum_{i=0}^{\infty} a_i \cdot z^i$  we have that  $\sum_{i=0}^{\infty} a_i \cdot z^i$  converges proving that  $z \in \mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i}$
2. First as  $B_{||}(0, R_{\sum_{i=0}^{\infty} (i+1) \cdot a_{i+1} \cdot z^i}) \subseteq \mathcal{D}_{\sum_{i=0}^{\infty} (i+1) \cdot a_{i+1} \cdot z^i} \subseteq \mathcal{D}_{\sum_{i=0}^{\infty} a_i \cdot z^i} \subseteq \overline{B_{||}(0, R_{\sum_{i=0}^{\infty} a_i \cdot z^i})}$  so using 18.11 we have that

$$R_{\sum_{i=0}^{\infty} (i+1) \cdot a_{i+1} \cdot z^i} \leq R_{\sum_{i=0}^{\infty} a_i \cdot z^i} \quad (18.9)$$

Let  $0 < z \in \mathbb{C}$  be such that  $|z| < R_{\sum_{i=0}^{\infty} a_i \cdot z^i}$  then  $\exists \rho \in \mathbb{R}$  such that  $0 < |z| < \rho < R_{\sum_{i=0}^{\infty} a_i \cdot z^i}$ . From  $\rho < R_{\sum_{i=0}^{\infty} a_i \cdot z^i}$  it follows that  $\sum_{i=0}^{\infty} a_n \cdot \rho^n$  converges. Using 11.376 we have then that  $\exists K \in \mathbb{R}_+$  such that

$$\forall n \in \mathbb{N} \models |a_n| \cdot \rho^n = |a_n \cdot \rho^n| < K \quad (18.10)$$

Now  $\forall n \in \mathbb{N}$  we have

$$\begin{aligned} |(n+1) \cdot a_{n+1} \cdot z^n| &= (n+1) \cdot |a_{n+1}| \cdot |z|^n \\ &= \frac{n+1}{\rho} \cdot \left( \frac{|z|}{\rho} \right)^n \cdot a_{n+1} \cdot \rho^{n+1} \\ &\stackrel{18.10}{\leq} K \cdot \frac{n+1}{\rho} \cdot \left( \frac{|z|}{\rho} \right)^n \\ &= K \cdot \frac{n+1}{\rho} \cdot q^n \text{ where } 0 < q = \frac{|z|}{\rho} < 1 \end{aligned}$$

proving that

$$\forall n \in \mathbb{N} \models |(n+1) \cdot a_{n+1} \cdot z^n| \leq K \cdot \frac{n+1}{\rho} \cdot q^n \text{ where } 0 < q < 1 \quad (18.11)$$

Now for the serie  $\sum_{n=0}^{\infty} (n+1) \cdot q^n$  we have that  $\frac{((n+1)+1) \cdot q^{n+1}}{(n+1) \cdot q^n} = \frac{n+2}{n+1} \cdot q$ . As by 11.332 we have that  $\lim_{n \rightarrow \infty} \frac{n+2}{n+1} = 1 \Rightarrow \lim_{n \rightarrow \infty} \left( \frac{n+2}{n+1} \cdot q \right) = q < 1$  we can use d'Alembert (see 18.17) proving to conclude that  $\sum_{i=0}^{\infty} (n+1) \cdot q^n$  is convergent. Using then dominant convergence (see 11.393) together with 18.11 we conclude then that  $\sum_{i=0}^{\infty} (n+1) \cdot a_{n+1} \cdot z^n$  is convergent. So we have proved that  $\forall x \in \mathbb{C}$  with  $0 < |z| < R_{\sum_{i=0}^{\infty} a_i \cdot z^i}$  we have  $z \in \mathcal{D}_{\sum_{i=0}^{\infty} (i+1) \cdot a_{i+1} \cdot z^i}$  and as  $0 \in \mathcal{D}_{\sum_{i=0}^{\infty} (i+1) \cdot a_{i+1} \cdot z^i}$  we conclude that  $B_{||}(0, R_{\sum_{i=0}^{\infty} a_i \cdot z^i}) \subseteq \mathcal{D}_{\sum_{i=0}^{\infty} (i+1) \cdot a_{i+1} \cdot z^i} \subseteq \overline{B_{||}(0, R_{\sum_{i=0}^{\infty} (i+1) \cdot a_{i+1} \cdot z^i})}$ . From 18.11 it follows then that  $R_{\sum_{i=0}^{\infty} a_i \cdot z^i} \leq R_{\sum_{i=0}^{\infty} (i+1) \cdot a_{i+1} \cdot z^i}$  which together with 18.9 proves finally that

$$R_{\sum_{i=0}^{\infty} (i+1) \cdot a_{i+1} \cdot z^i} = R_{\sum_{i=0}^{\infty} a_i \cdot z^i} \quad \square$$

We now prove that the derivated power serie of a power serie is indeed its derivative.

**Theorem 18.23.** Let  $\sum_{i=0}^{\infty} a_i \cdot z^i$  be a power serie with  $0 < R_{\sum_{i=0}^{\infty} a_i \cdot z^i}$  and  $z \in \mathbb{C}$  then  $\sum_{i=0}^{\infty} a_i \cdot (.)^i$ :  $B_{\parallel}(0, R_{\sum_{i=0}^{\infty} a_i \cdot z^i}) \rightarrow \mathbb{C}$  has a derivate (see 14.18) at every  $z \in B_{\parallel}(0, R_{\sum_{i=0}^{\infty} a_i \cdot z^i})$  with  $(\sum_{i=0}^{\infty} a_i \cdot (.)^i)'(z) = (\sum_{i=0}^{\infty} (i+1) \cdot a_{i+1} \cdot (.)^i)(z)$

**Proof.** Define  $f = \sum_{i=0}^{\infty} a_i \cdot (.)^i$  and  $g = \sum_{i=0}^{\infty} (i+1) \cdot a_{i+1} \cdot (.)^i$ . Let  $\varepsilon > 0$  and take  $z$  such that  $|z| < R_{\sum_{i=0}^{\infty} a_i \cdot z^i} \Rightarrow z \in B_{\parallel}(0, R_{\sum_{i=0}^{\infty} a_i \cdot z^i})$ . As  $B_{\parallel}(0, R_{\sum_{i=0}^{\infty} a_i \cdot z^i})$  is open there exists a  $0 < \rho < R_{\sum_{i=0}^{\infty} a_i \cdot z^i}$  such that  $B_{\parallel}(z, \rho) \subseteq B_{\parallel}(0, R_{\sum_{i=0}^{\infty} a_i \cdot z^i})$ , if now  $h \in \mathbb{C}$  is choosen such that  $|h| < \min(\rho, \rho - |z|)$  then  $|h+z| \leq |h| + |z| < \rho < R_{\sum_{i=0}^{\infty} a_i \cdot z^i}$  so that  $f(z)$ ,  $f(z+h)$  and  $g(z)$  are well defined (see also 18.22) and  $|h+z-z| = |h| < \rho$  so that  $h \in B_{\parallel}(z, \rho)$  (needed for the definition of a derivative)

$$\begin{aligned}
\frac{f(z+h) - f(z)}{h} - g(z) &= \frac{1}{h} \cdot (f(z+h) - f(z) - h \cdot g(z)) \\
&= \frac{1}{h} \cdot \left( \sum_{i=0}^{\infty} a_i \cdot (z+h)^i - \sum_{i=0}^{\infty} a_i \cdot z^i - h \cdot \sum_{i=0}^{\infty} (i+1) \cdot a_{i+1} \cdot z^i \right) \\
&= \frac{1}{h} \cdot \left( \sum_{i=0}^{\infty} a_i \cdot \left( \sum_{j=0}^i \binom{i}{j} \cdot h^j \cdot z^{i-j} \right) - \sum_{i=0}^{\infty} a_i \cdot z^i - h \cdot \sum_{i=0}^{\infty} (i+1) \cdot a_{i+1} \cdot z^i \right. \\
&\quad \left. z^i \right) \\
&= \frac{1}{h} \cdot \left( \sum_{i=0}^{\infty} a_i \cdot \left( \left( \sum_{j=0}^i \binom{i}{j} \cdot h^j \cdot z^{i-j} \right) - z^i \right) - h \cdot \sum_{i=0}^{\infty} (i+1) \cdot a_{i+1} \cdot z^i \right) \\
&= \frac{1}{h} \cdot \left( \sum_{i=1}^{\infty} a_i \cdot \left( \left( \sum_{j=0}^i \binom{i}{j} \cdot h^j \cdot z^{i-j} \right) - z^i \right) + a_0 \cdot \left( \left( \sum_{j=0}^0 \binom{0}{j} \cdot h^j \cdot z^{0-j} \right) - z^0 \right) - h \cdot \sum_{i=0}^{\infty} (i+1) \cdot a_{i+1} \cdot z^i \right) \\
&= \frac{1}{h} \cdot \left( \sum_{i=1}^{\infty} a_i \cdot \left( \left( \sum_{j=0}^i \binom{i}{j} \cdot h^j \cdot z^{i-j} \right) - z^i \right) + a_0 \cdot \left( \left( \binom{0}{0} \cdot h^0 \cdot z^0 - 1 \right) - h \cdot \sum_{i=0}^{\infty} (i+1) \cdot a_{i+1} \cdot z^i \right) \right) \\
&= \frac{1}{h} \cdot \left( \sum_{i=1}^{\infty} a_i \cdot \left( \left( \sum_{j=0}^i \binom{i}{j} \cdot h^j \cdot z^{i-j} \right) - z^i \right) - h \cdot \sum_{i=0}^{\infty} (i+1) \cdot a_{i+1} \cdot z^i \right) \\
&= \frac{1}{h} \cdot \left( \sum_{i=1}^{\infty} a_i \cdot \left( \left( \sum_{j=1}^i \binom{i}{j} \cdot h^j \cdot z^{i-j} + \binom{i}{0} \cdot h^0 \cdot z^{i-0} \right) - z^i \right) - h \cdot \sum_{i=0}^{\infty} (i+1) \cdot a_{i+1} \cdot z^i \right) \\
&\quad h \cdot \sum_{i=0}^{\infty} (i+1) \cdot a_{i+1} \cdot z^i \\
&= \frac{1}{h} \cdot \left( \sum_{i=1}^{\infty} a_i \cdot \left( \left( \sum_{j=1}^i \binom{i}{j} \cdot h^j \cdot z^{i-j} + z^i \right) - z^i \right) - h \cdot \sum_{i=0}^{\infty} (i+1) \cdot a_{i+1} \cdot z^i \right)
\end{aligned}$$

$$\begin{aligned}
&= \frac{1}{h} \cdot \left( \sum_{i=1}^{\infty} a_i \cdot \left( \sum_{j=1}^i \binom{i}{j} \cdot h^j \cdot z^{i-j} \right) - h \cdot \sum_{i=0}^{\infty} (i+1) \cdot a_{i+1} \cdot z^i \right) \\
&= \frac{1}{h} \cdot \left( \sum_{i=2}^{\infty} a_i \cdot \left( \sum_{j=1}^i \binom{i}{j} \cdot h^j \cdot z^{i-j} \right) + a_1 \cdot \left( \sum_{j=1}^1 \binom{i}{j} \cdot h^j \cdot z^{i-j} \right) - \right. \\
&\quad \left. h \cdot \sum_{i=1}^{\infty} (i+1) \cdot a_{i+1} \cdot z^i - h \cdot (0+1) \cdot a_{0+1} \cdot z^0 \right) \\
&= \frac{1}{h} \cdot \left( \sum_{i=2}^{\infty} a_i \cdot \left( \sum_{j=1}^i \binom{i}{j} \cdot h^j \cdot z^{i-j} \right) + a_1 \cdot \left( \begin{matrix} 1 \\ 1 \end{matrix} \right) \cdot h^1 \cdot z^{1-1} - h \cdot \sum_{i=1}^{\infty} (i+1) \cdot a_{i+1} \cdot z^i - h \cdot a_1 \cdot z^0 \right) \\
&= \frac{1}{h} \cdot \left( \sum_{i=2}^{\infty} a_i \cdot \left( \sum_{j=1}^i \binom{i}{j} \cdot h^j \cdot z^{i-j} \right) - h \cdot \sum_{i=1}^{\infty} (i+1) \cdot a_{i+1} \cdot z^i \right) \\
&\stackrel{11.378}{=} \frac{1}{h} \cdot \left( \sum_{i=2}^{\infty} a_i \cdot \left( \sum_{j=1}^i \binom{i}{j} \cdot h^j \cdot z^{i-j} \right) - h \cdot \sum_{i=2}^{\infty} i \cdot a_i \cdot z^{i-1} \right) \\
&= \frac{1}{h} \cdot \left( \sum_{i=2}^{\infty} a_i \cdot \left( \left( \sum_{j=1}^i \binom{i}{j} \cdot h^j \cdot z^{i-j} \right) - i \cdot h \cdot z^{i-1} \right) \right) \\
&= \sum_{i=2}^{\infty} a_i \cdot \left( \left( \sum_{j=1}^i \binom{i}{j} \cdot h^{j-1} \cdot z^{i-j} \right) - i \cdot z^{i-1} \right) \\
&= \sum_{i=2}^{\infty} a_i \cdot \left( \left( \sum_{j=2}^i \binom{i}{j} \cdot h^{j-1} \cdot z^{i-j} \right) + \left( \begin{matrix} i \\ 1 \end{matrix} \right) \cdot h^{1-1} \cdot z^{i-1} - i \cdot z^{i-1} \right) \\
&= \sum_{i=2}^{\infty} a_i \cdot \left( \left( \sum_{j=2}^i \binom{i}{j} \cdot h^{j-1} \cdot z^{i-j} \right) + i \cdot z^{i-1} - i \cdot z^{i-1} \right) \\
&= \sum_{i=2}^{\infty} a_i \cdot \left( \sum_{j=2}^i \binom{i}{j} \cdot h^{j-1} \cdot z^{i-j} \right)
\end{aligned}$$

to summarize

$$\frac{f(z+h) - f(z)}{h} - h \cdot g(z) = \sum_{i=2}^{\infty} a_i \cdot \left( \sum_{j=2}^i \binom{i}{j} \cdot h^{j-1} \cdot z^{i-j} \right) \quad (18.12)$$

Now let  $i \in \{2, \dots, \infty\}$  then we have

$$\begin{aligned}
\left| \sum_{j=2}^i \binom{i}{j} \cdot h^{j-1} \cdot z^{i-j} \right| &= |h| \cdot \left| \sum_{j=2}^i \binom{i}{j} \cdot h^{j-2} \cdot z^{i-j} \right| \\
&\leq |h| \cdot \left| \sum_{j=2}^i \binom{i}{j} \cdot |h^{j-2}| \cdot |z^{i-j}| \right| \\
&= |h| \cdot \left| \sum_{j=2}^i \frac{i!}{j! \cdot (i-j)!} \cdot |h^{j-2}| \cdot |z^{i-j}| \right|
\end{aligned}$$

$$\begin{aligned}
& \stackrel{i, j \geq 2}{=} |h| \cdot \left| \sum_{j=2}^i \frac{i \cdot (i-1) \cdot (i-2)!}{j \cdot (j-1) \cdot (j-2)! \cdot (i-j)!} \cdot |h^{j-2}| \cdot |z^{i-j}| \right| \\
& \leqslant_{j \geq 2 \Rightarrow j \cdot (j-1) \geq 1} |h| \cdot \sum_{j=2}^i \frac{i \cdot (i-1) \cdot (i-2)!}{(j-2)! \cdot (i-j)!} \cdot |h^{j-2}| \cdot |z^{i-j}| \\
& = |h| \cdot i \cdot (i-1) \cdot \sum_{j=2}^i \frac{(i-2)!}{(j-2)! \cdot ((i-2)-(j-2))!} \cdot |h^{j-2}| \cdot |z^{(i-2)-(j-2)}| \\
& = |h| \cdot i \cdot (i-1) \cdot \sum_{j=0}^{i-2} \frac{(i-2)!}{j! \cdot ((i-2)-j)!} \cdot |h^j| \cdot |z^{(i-2)-j}| \\
& = |h| \cdot i \cdot (i-1) \cdot (|z| + |h|)^{i-2}
\end{aligned}$$

Using the above with 18.12 gives

$$\left| \frac{f(z+h) - f(z)}{h} - g(z) \right| \leq |h| \cdot \sum_{i=2}^{\infty} (i \cdot (i-1) \cdot a_i \cdot (|z| + |h|)^{i-2}) \quad (18.13)$$

Now as  $|z| + |h| < \rho < R_{\sum_{i=0}^{\infty} a_i \cdot z^i}$  we have using 18.6 that  $\sum_{i=0}^{\infty} a_i \cdot (|z| + |h|)^i$  converges. From 18.22 it follows then that  $\sum_{i=0}^{\infty} (i+1) \cdot a_{i+1} \cdot (|z| + |h|)^i$  converges and applying 18.22 we have that  $\sum_{i=0}^{\infty} (i+1) \cdot (i+2) \cdot a_{i+2} \cdot (|z| + |h|)^i$  converges. As  $\sum_{i=0}^{\infty} (i+1) \cdot (i+2) \cdot a_{i+2} \cdot (|z| + |h|)^i \stackrel{11.378}{=} \sum_{i=2}^{\infty} (i-1) \cdot i \cdot a_i \cdot (|z| + |h|)^{i-2}$  proving that

$$0 < \sum_{i=2}^{\infty} (i \cdot (i-1) \cdot a_i \cdot (|z| + |h|)^{i-2}) \in \mathbb{R} \quad (18.14)$$

If we now add the extra condition on  $h$  that  $0 < |h| < \min \left( \rho, \rho - |z|, \frac{\varepsilon}{\sum_{i=2}^{\infty} (i \cdot (i-1) \cdot a_i \cdot (|z| + |h|)^{i-2})} \right)$  then by 18.3 we have

$$\left| \frac{f(z+h) - f(z)}{h} - g(z) \right| < \varepsilon$$

which proves our theorem.  $\square$

We can now prove what the derivative of the exponential function is.

**Corollary 18.24.** *The exponential function  $\exp: \mathbb{C} \rightarrow \mathbb{C}$  defined by  $\exp(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!}$  has a derivative equal to itself at every point in  $\mathbb{C}$ , in other words  $\forall z \in \mathbb{C}$  that  $\exp'(z) = \exp(z)$*

**Proof.** First by 18.18 we have that  $\mathcal{D}_{\sum_{n=0}^{\infty} \frac{z^n}{n!}} = \mathbb{C}$ . So if  $z \in \mathbb{C}$  we have that using the above theorem that  $\exp'(z)$  exists and that  $\exp'(z) = \sum_{i=0}^{\infty} (i+1) \cdot \frac{z^i}{(i+1)!} = \sum_{i=0}^{\infty} \frac{z^i}{i!} = \exp(z)$   $\square$

**Corollary 18.25.** *The exponential function  $\exp: \mathbb{C} \rightarrow \mathbb{C}$  defined by  $\exp(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!}$  is continuous on  $\mathbb{C}$*

**Proof.** Let  $z \in \mathbb{C}$  then by the previous corollary  $\exp$  has a derivative at  $z$  and thus  $\exp$  is differentiable at  $z$  (see 14.19). As differentiability implies continuity (see 14.15) we have that  $\exp$  is continuous at  $z$  hence  $\exp$  is continuous on  $\mathbb{C}$ .  $\square$

Next we show that the exponential function behaves as a power which motivates the other notation for  $\exp(z)$  as  $e^z$ .

**Theorem 18.26.** *Let  $x, y \in \mathbb{C}$  then  $\exp(x+y) = \exp(x) \cdot \exp(y)$*

**Proof.** Let  $x, y \in \mathbb{C}$  then as  $\mathcal{D}_{\sum_{n=0}^{\infty} \frac{z^n}{n!}} = \mathbb{C} = \overline{\mathcal{D}_{\sum_{n=0}^{\infty} \frac{z^n}{n!}}}$  we have that  $\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$  and  $\exp(y) = \sum_{n=0}^{\infty} \frac{y^n}{n!}$  are absolute convergent

$$\begin{aligned} \exp(x) \cdot \exp(y) &= \left( \sum_{n=0}^{\infty} \frac{x^n}{n!} \right) \cdot \left( \sum_{n=0}^{\infty} \frac{y^n}{n!} \right) \\ &\stackrel{11.402}{=} \left( \sum_{n \in \mathbb{N}} \frac{x^n}{n!} \right) \cdot \left( \sum_{n \in \mathbb{N}} \frac{y^n}{n!} \right) \\ &\stackrel{11.411}{=} \sum_{(i,j) \in \mathbb{N} \times \mathbb{N}} \left( \frac{x^i}{i!} \cdot \frac{y^j}{j!} \right) \{ \text{absolute convergent} \} \end{aligned}$$

giving

$$\exp(x) \cdot \exp(y) = \sum_{(i,j) \in \mathbb{N} \times \mathbb{N}} \left( \frac{x^i}{i!} \cdot \frac{y^j}{j!} \right) a \text{ absolute convergent serie} \quad (18.15)$$

Let  $n \in \mathbb{N}$  define then  $N_n = \{(i, j) \in \mathbb{N} \times \mathbb{N} \mid i + j = n\}$ , as  $N_n \subseteq \mathbb{N} \times \mathbb{N}$  we have  $\bigcup_{n \in \mathbb{N}} N_n \subseteq \mathbb{N} \times \mathbb{N}$  and if  $(i, j) \in \mathbb{N} \times \mathbb{N}$  then either  $(i, j) \in N_{i+j}$  proving that  $\mathbb{N} \times \mathbb{N} \subseteq \bigcup_{n \in \mathbb{N}} N_n$ . Further we have that if  $n, n' \in \mathbb{N}$  with  $n = n'$  then if  $(i, j) \in N_n \cap N_{n'} \Rightarrow i + j = n \wedge i + j = n' \Rightarrow n = n'$  a contradiction so we must have  $N_n \cap N_{n'} = \emptyset$ . Summarized

$$\mathbb{N} \times \mathbb{N} = \bigsqcup_{i \in \mathbb{N}} N_n \quad (18.16)$$

Further we have that  $\beta: \{0, \dots, n\} \rightarrow N_n$  defined by  $\beta(i) = (i, n - i)$  is a bijection

**injectivity.** if  $\beta(i) = \beta(j)$  then  $(i, n - i) = (j, n - j) \Rightarrow i = j$

**surjectivity.** if  $(i, j) \in N_n$  then  $i + j = n \Rightarrow j = n - i \Rightarrow (i, j) = (i, n - i) = \beta(i)$

Using then the associativity of absolute convergent series we have

$$\begin{aligned} \sum_{(i,j) \in \mathbb{N} \times \mathbb{N}} \left( \frac{x^i}{i!} \cdot \frac{y^j}{j!} \right) &\stackrel{11.410}{=} \sum_{n \in \mathbb{N}} \left( \sum_{(i,j) \in N_n} \left( \frac{x^i}{i!} \cdot \frac{y^j}{j!} \right) \right) \\ &\stackrel{11.403}{=} \sum_{n \in \mathbb{N}} \left( \sum_{i \in \{0, \dots, n\}} \left( \frac{x^{\beta(i)_1} \cdot y^{\beta(i)_2}}{\beta(i)_1! \cdot \beta(i)_2!} \right) \right) \\ &= \sum_{n \in \mathbb{N}} \left( \sum_{i \in \{0, \dots, n\}} \left( \frac{x^i \cdot y^{(n-i)}}{i! \cdot j!} \right) \right) \\ &= \sum_{n \in \mathbb{N}} \frac{1}{n!} \left( \sum_{i \in \{0, \dots, n\}} \left( \frac{n! \cdot x^i \cdot y^{(n-i)}}{i! \cdot n - i!} \right) \right) \\ &= \sum_{n \in \mathbb{N}} \frac{1}{n!} \left( \sum_{i \in \{0, \dots, n\}} \left( \binom{n}{i} \cdot x^i \cdot y^{(n-i)} \right) \right) \\ &= \sum_{n \in \mathbb{N}} \frac{(x+y)^n}{n!} \\ &= \exp(x+y) \end{aligned}$$

which by 18.15 proves that

$$\exp(x+y) = \exp(x) \cdot \exp(y) \quad \square$$

**Theorem 18.27.**  $\exp: \mathbb{C} \rightarrow \mathbb{C}$  satisfies

1.  $\exp(0) = 1$
2.  $\forall z \in \mathbb{C}$  we have  $\exp(z) \neq 0$

3.  $\forall z \in \mathbb{C}$  we have  $\exp(-z) = \frac{1}{\exp(z)}$
4.  $\forall z \in \mathbb{R}$  we have  $0 < \exp(z)$  in other words  $\exp(\mathbb{R}) = \mathbb{R}_+$
5.  $\forall z \in \mathbb{R}$  with  $x > 0$  we have  $\exp(x) > 1$
6.  $\forall x, y \in \mathbb{R}$  with  $x > y$  we have  $\exp(x) > \exp(y)$  /exp is strictly increasing on  $\mathbb{R}$

**Proof.**

1.  $\exp(0) = \sum_{i=0}^{\infty} \frac{z^i}{i!} \stackrel{18.3}{=} \frac{z^0}{0!} = 1$
2. Assume that  $\exists z \in \mathbb{C}$  such that  $\exp(z) = 0$  then as  $1 = \exp(0) = \exp(z + (-z)) \stackrel{18.26}{=} \exp(z) \cdot \exp(-z) = 0 \cdot \exp(-z) = 0$  leading to the contradiction  $1 = 0$ . Hence we have  $\forall z \in \mathbb{C}$  that  $\exp(z) \neq 0$
3.  $1 = \exp(0) = \exp(z + (-z)) \stackrel{18.26}{=} \exp(z) \cdot \exp(-z) \stackrel{\exp(z) \neq 0}{\Rightarrow} \exp(-z) = \frac{1}{\exp(z)}$
4. If  $z \in \mathbb{R}$  then we have 3 cases to consider for  $z$ 
  - $z \in \mathbb{R}_+$ . then  $\forall i \in \mathbb{N}$  we have  $0 < \frac{z^i}{i!}$  (see 9.69) so that by 11.383  $0 < \sum_{i=0}^{\infty} \frac{z^i}{i!} = \exp(z)$
  - $z = 0$ . then  $0 < 1 = \exp(z)$
  - $z \in \mathbb{R}_-$ . then  $-z \in \mathbb{R}_+$  so that  $0 < \frac{1}{\exp(z)} = \exp(-z)$
5. If  $z > 0$  then by 9.69  $\forall i \in \mathbb{N}$  we have  $0 < z^i \Rightarrow 0 < \frac{z^i}{i!}$  so that using 11.383
$$0 < \sum_{i=1}^{\infty} \frac{z^i}{i!} \tag{18.17}$$
So  $\exp(z) = \sum_{i=0}^{\infty} \frac{z^i}{i!} = \frac{z^0}{0!} + \sum_{i=1}^{\infty} \frac{z^i}{i!} = 1 + \sum_{i=1}^{\infty} \frac{z^i}{i!} > 1$
6. Let  $x > y$  then  $x - y > 0$  so by 18.27(5) we have  $\exp(x) = \exp(y + (x - y)) = \exp(y) \cdot \exp(x - y) > \exp(y)$   $\square$

**Corollary 18.28.** Let  $x, y \in \mathbb{R}$  with  $x < y$  then exp is convex on  $[x, y]$

**Proof.** This follows from the fact that  $\forall x \in \mathbb{R}$  we have that  $\exp''(x) = (\exp'(x))' \stackrel{18.24}{=} \exp'(x) \stackrel{18.24}{=} \exp(x) > 0$  and the condition for convexity (see 14.103).  $\square$

**Definition 18.29. (Euler's number)**  $e = \exp(1)$

We can now see how for natural numbers exp behaves as a power of  $e$

**Theorem 18.30.**

1.  $\forall n \in \mathbb{N}$  we have  $\exp(n) = e^n$
2.  $\forall n \in \mathbb{N}$  we have  $\exp(-n) = \frac{1}{e^n} \stackrel{\text{definition}}{=} e^{-n}$

**Proof.**

1. We use mathematical induction so let  $\mathcal{S} = \{n \in \mathbb{N} \mid \exp(n) = e^n\}$  then we have
  - $0 \in \mathcal{S}$ .  $\exp(0) = 1 = e^0$
  - $n \in \mathcal{S} \Rightarrow n + 1 \in \mathcal{S}$ . then  $\exp(n + 1) \stackrel{18.26}{=} \exp(n) \cdot \exp(1) \stackrel{n \in \mathcal{S}}{=} e^n \cdot e = e^{n+1}$  proving that  $n + 1 \in \mathcal{S}$
2.  $\exp(-n) \stackrel{18.27}{=} \frac{1}{\exp(n)} \stackrel{(1)}{=} \frac{1}{e^n} e^{-n}$   $\square$

**Theorem 18.31.** Let  $y \in \mathbb{R}_+$  then there exists a  $x \in \mathbb{R}$  such that  $y < \exp(x)$

**Proof.** As  $0 < 1$  we have by 18.27 that  $1 = \exp(0) < \exp(1) = e$ . Using 9.65 there exists a  $n \in \mathbb{N}$  such that  $y < e^n \stackrel{\text{previous theorem}}{=} \exp(n)$   $\square$

Next we want to define the Napierian logarithm which is the inverse of exp restricted to  $\mathbb{R}$ , so we must prove that  $\exp|_{\mathbb{R}}: \mathbb{R} \rightarrow \mathbb{R}_+$  is a bijection.

**Theorem 18.32.**  $\exp: \mathbb{R} \rightarrow \mathbb{R}_+$  is a bijection where  $\exp(\mathbb{R}_-) = ]0, 1[$  and  $\exp(\mathbb{R}_+) = ]1, \infty[$

**Proof.**

**injectivity.** Let  $\exp(x) = \exp(y)$  and assume that  $x \neq y$  then we may assume that  $x > y$ , as  $\exp$  is strictly increasing (see 18.27 (6)) we have  $\exp(x) > \exp(y)$  a contradiction so we must have that  $x = y$ .

**surjectivity.** Let  $y \in \mathbb{R}_+ = ]0, \infty[$  then we must consider 3 cases

**$y = 1$ .** then  $y = 1 = \exp(0)$

**$1 < y$ .** then by 18.31 there exists a  $b \in \mathbb{R}$  such that  $y < \exp(b)$  so that we have  $\exp(0) < y < \exp(b)$ . As  $\exp$  is continuous on  $\mathbb{R}$  (see 18.25) and thus on  $[0, b]$  we have by the intermediate value theorem (see 11.456) that  $\exists x \in [0, b]$  such that  $\exp(x) = y$

**$0 < y < 1$ .** So  $1 < \frac{1}{y}$  and we can apply the previous case to find a  $x \in \mathbb{R}$  such that  $\exp(x) = \frac{1}{y}$  hence  $\exp(-x) = \frac{1}{\exp(x)} = y$

for the remaining of the theorem we have if  $x \in \mathbb{R}_+$  by 18.27 (5) that  $1 < \exp(x) \Rightarrow \exp(x) \in ]1, \infty[$  so that

$$\exp(\mathbb{R}_+) \subseteq ]1, \infty[ \quad (18.18)$$

If  $x \in \mathbb{R}_-$  then  $-x \in \mathbb{R}_+$  so that  $1 < \exp(-x) = \frac{1}{\exp(x)}$  or  $1 < \frac{1}{\exp(x)} \Rightarrow \exp(x) < 1$  proving that  $\exp(x) \in ]0, 1[$  so that

$$\exp(\mathbb{R}_-) \subseteq ]0, 1[ \quad (18.19)$$

If now  $y \in ]1, \infty[ \subseteq \mathbb{R}_+$  then as  $\exp$  is a surjection there exists a  $x \in \mathbb{R}$  such that  $\exp(x) = y$  assume  $x \in \mathbb{R}_-$  then by 18.19 we have  $y = \exp(x) \in ]0, 1[$  not  $]1, \infty[$  so we must have  $x \in \mathbb{R}_+$  proving that  $]1, \infty[ \subseteq \exp(\mathbb{R}_+)$  which together with 18.18 gives

$$\exp(\mathbb{R}_+) = ]1, \infty[$$

If now  $y \in ]0, 1[ \subseteq \mathbb{R}_+$  then as  $\exp$  is a surjection there exists a  $x \in \mathbb{R}$  such that  $\exp(x) = y$  assume  $x \in \mathbb{R}_+$  then by 18.18 we have  $y = \exp(x) \in ]1, \infty[$  not  $]0, 1[$  so we must have  $x \in \mathbb{R}_-$  proving that  $]0, 1[ \subseteq \exp(\mathbb{R}_-)$  which together with 18.19 gives

$$\exp(\mathbb{R}_-) = ]0, 1[$$

□

The above definition allows us to define the Napierian logarithm

**Definition 18.33. (Napierian logarithm)**  $\log: \mathbb{R}_+ \rightarrow \mathbb{R}$  is defined by  $\log = \exp^{-1}$

Based on the properties of  $\exp$  we have the following properties of the logarithm

**Theorem 18.34.**  $\log: \mathbb{R}_+ \rightarrow \mathbb{R}$  has the following properties

1.  $\forall x \in \mathbb{R}_+$  we have that  $\exp(\log(x)) = x$
2.  $\forall x \in \mathbb{R}$  we have that  $\log(\exp(x)) = x$
3.  $\log(e) = 1$
4.  $\log(1) = 0$
5.  $\forall x, y \in \mathbb{R}_+$  we have  $\log(x \cdot y) = \log(x) + \log(y)$
6.  $\forall x \in \mathbb{R}_+$  we have  $\log\left(\frac{1}{x}\right) = -\log(x)$
7.  $\forall x \in \mathbb{R}_+$   $\log$  has a derivative  $\log'(x) = \frac{1}{x}$
8.  $\log: \mathbb{R}_+ \rightarrow \mathbb{R}$  is strictly increasing and  $\forall x < 1$  we have  $\log(x) < 0$

**Proof.** This follows easily from the definition of  $\log$  as the inverse of  $\exp$  on  $\mathbb{R}_+$

1. Let  $x \in \mathbb{R}$  then  $\exp(\log(x)) = \exp(\exp^{-1}(x)) = x$
2. Let  $x \in \mathbb{R}$  then  $\log(\exp(x)) = \exp^{-1}(\exp(x)) = x$

3.  $\log(e) = \log(\exp(1)) \stackrel{(2)}{=} 1$
4. As  $\exp(0) = 1$  we have that  $\log(1) = \exp^{-1}(\exp(0)) = 0$
5. Let  $x, y \in \mathbb{R}_+$  then  $\exp(\log(x) + \log(y)) = \exp(\log(x)) \cdot \exp(\log(y)) \stackrel{(1)}{=} x \cdot y$  so that  $\log(x \cdot y) = \log(\exp(\log(x) + \log(y))) \stackrel{(2)}{=} \log(x) + \log(y)$
6. Let  $x \in \mathbb{R}_+$  then  $x = \exp(\log(x))$  so that  $\exp(-\log(x)) = \frac{1}{x}$  and thus  $-\log(x) = \log(\exp(-\log(x))) = \log\left(\frac{1}{x}\right)$
7. Take  $x \in \mathbb{R}_+$  and take  $y = \log(x)$ . As  $\exp(y) > 0$  we have by a consequence of the inverse function theorem (see 14.167) that there exists a open  $U \subseteq \mathbb{R}_+, V \subseteq \mathbb{R}$  with  $y \in U$  and  $x = \exp(\log(x)) = \exp(y) \in V$  such that  $\exp|_U: U \rightarrow V$  has a inverse  $\exp|_U^{-1}: V \rightarrow U$  such that  $\forall z \in V$  we have  $(\exp|_U^{-1})'(z) \cdot \exp(\exp^{-1}(z))$ . As  $\forall w \in U$  we have that  $(\exp|_U \circ \log|_V)(w) = \exp|_U(\log|_V(w)) = \exp(\log(w)) = w = 1_V(w)$  and  $\forall w \in U$  we have that  $(\log|_V \circ \exp|_U)(w) = \log(\exp(w)) = w = 1_U(w)$  we conclude that  $\exp|_U^{-1} = \log|_V$ . So we have that  $\forall z \in V$  that  $1 = (\log|_V)'(z) \cdot \exp(\log(z)) \stackrel{14.20}{=} \log'(z) \cdot z$ . More specific as  $x \in V \subseteq \mathbb{R}_+$  we proved that  $\log'(x) = \frac{1}{x}$
8. Let  $x, y \in \mathbb{R}_+$  with  $x < y$  then if  $\log(y) \leq \log(x)$  we have if  $\exp$  is strictly increasing we have  $y = \exp(\log(y)) \leq \exp(\log(x)) = x$  contradicting  $x < y$  so that  $\log(x) < \log(y)$ . Further if  $x < 1$  then  $\log(x) < \log(1) = 0$ .  $\square$

Once we have introduced the log function and the definition  $e^x = \exp(x)$  (which was expired by the fact that  $\exp(n) = e^n = \underbrace{e \dots e}_{n \text{ times}}$  (see 18.30) we can now define a general definiton of the power of a real number to a real number.

**Definition 18.35.** Let  $a \in ]0, \infty[$  be a positive number and  $x \in \mathbb{R}$  then  $a^{[x]} = \exp(x \cdot \log(a)) \in \mathbb{R}$

**Remark 18.36.** The restrictions in the above definition are needed for  $\log(a)$  is only defined if  $a \in \mathbb{R}_+ = ]0, \infty[$ , further if for example if  $x = \frac{1}{2}$  then we will see that  $a^{\frac{1}{2}} = \sqrt{a}$  which is not defined on the negative real numbers.

Next we prove now that we have the classical properties of our generalized power. TODO add  $a^{x+y} = a^x \cdot a^y$

**Theorem 18.37.** We have the following properties for the generalized power

1.  $\forall n \in \mathbb{N}$  and  $\forall a \in ]a, \infty[$  we have that  $a^{[n]} = \underbrace{a \dots a}_{n \text{ times}}$  (we have a conflict of notation here so we use  $\underbrace{a \dots a}$  to note the classical power on  $\mathbb{R}$ )
2.  $\forall x \in \mathbb{R}$  we have  $1^{[x]} = 1$
3.  $\forall x \in \mathbb{R}$  we have  $e^{[x]} = \exp(x)$
4.  $\forall x \in \mathbb{R}$  we have  $\frac{a^{[x]}}{a} = a^{[x-1]}$
5.  $\forall x, y \in \mathbb{R}$  and  $\forall a \in ]0, \infty[$  we have  $(a^{[x]})^{[y]} = a^{[x \cdot y]}$
6.  $\forall x \in \mathbb{R}$  and  $\forall a \in ]0, \infty[$  we have  $a^{[-x]} = \frac{1}{a^{[x]}}$
7.  $\forall a, b \in ]0, \infty[$  and  $\forall x \in \mathbb{R}$  we have  $(a \cdot b)^{[x]} = a^{[x]} \cdot b^{[x]}$
8.  $\forall a \in ]0, \infty[$  and  $\forall x, y \in \mathbb{R}$  we have  $a^{[x+y]} = a^{[x]} \cdot a^{[y]}$
9.  $\forall x \in \mathbb{R}$  and  $\forall a \in ]0, \infty[$  we have  $\log(a^{[x]}) = x \cdot \log(a)$
10. Let  $a \in ]a, \infty[$  and define  $a^{(\cdot)}: \mathbb{R} \rightarrow \mathbb{R}_+$  by  $a^{(\cdot)}(x) = a^{[x]}$  then  $\forall x \in \mathbb{R}$  we have  $(a^{(\cdot)})'(x) = \log(a) \cdot a^{[x]}$
11. Let  $x \in \mathbb{R}$  and define  $(\cdot)^{[x]}: \mathbb{R}_+ \rightarrow \mathbb{R}_+$  by  $(\cdot)^{[x]}(y) = y^{[x]}$  then  $\forall y \in \mathbb{R}$  we have that  $((\cdot)^{[x]})'(y) = x \cdot y^{[x-1]}$

**Proof.**

1. We proof this by induction so let  $\mathcal{S} = \left\{ n \in \mathbb{N} \mid a^{[n]} = \underbrace{a \dots a}_{n \text{ times}} \right\}$  then we have

$0 \in \mathcal{S}$ .  $a^{[0]} = \exp(0 \cdot \log(a)) = \exp(0) \stackrel{18.27}{=} 1 = \underbrace{a \dots a}_{0 \text{ times}}$  proving that  $n \in \mathbb{N}$

$n \in \mathcal{S} \Rightarrow n + 1 \in \mathcal{S}$ . then

$$\begin{aligned} a^{[n+1]} &= \exp((n+1) \cdot \log(a)) \\ &= \exp(n \cdot \log(a) + \log(a)) \\ &\stackrel{18.26}{=} \exp(n \cdot \log(a)) \cdot \exp(\log(a)) \\ &\stackrel{n \in \mathcal{S}}{=} \underbrace{a \dots a}_{n \text{ times}} \cdot \exp(\log(a)) \\ &= \underbrace{a \dots a}_{n \text{ times}} \cdot a \\ &= \underbrace{a \dots a}_{n+1 \text{ times}} \end{aligned}$$

proving that  $n + 1 \in \mathcal{S}$

2. Let  $x \in \mathbb{R}$  then  $x^{[1]} = \exp(\log(1) \cdot x) \stackrel{18.34}{=} \exp(0 \cdot x) = \exp(0) \stackrel{18.27}{=} 1$

3.  $e^{[x]} = \exp(x \cdot \log(e)) \stackrel{18.34}{=} \exp(x \cdot 1) = \exp(x)$

4.

$$\begin{aligned} (a^{[x]})^{[y]} &= (\exp(x \cdot \log(a)))^{[y]} \\ &= \exp(y \cdot \log(\exp(x \cdot \log(a)))) \\ &= \exp(y \cdot (x \cdot \log(a))) \\ &= \exp((x \cdot y) \cdot \log(a)) \\ &= a^{[x] \cdot [y]} \end{aligned}$$

5.

$$\begin{aligned} \frac{a^{[x]}}{a} &= \frac{\exp(\log(a) \cdot x)}{\exp(\log(a))} \\ &\stackrel{18.27}{=} \exp(-\log(a)) \cdot \exp(\log(a) \cdot x) \\ &= \exp(\log(a) \cdot x - \log(a)) \\ &= \exp(\log(a) \cdot (x - 1)) \\ &= a^{[x-1]} \end{aligned}$$

6. Let  $x \in \mathbb{R}$  then  $a^{[-x]} = \exp(-x \cdot \log(a)) \stackrel{18.27}{=} \exp\left(\frac{1}{\exp(x \cdot \log(a))}\right) = \frac{1}{a^{[x]}}$

7. Let  $b \in ]0, \infty[$  and  $x \in \mathbb{R}$  then  $(a \cdot b)^{[x]} = \exp(\log(a \cdot b) \cdot x) \stackrel{18.34}{=} \exp((\log(a) + \log(b)) \cdot x) = \exp(\log(a) \cdot x + \log(b) \cdot x) \stackrel{18.26}{=} \exp(\log(a) \cdot x) \cdot \exp(\log(b) \cdot x) = a^{[x]} \cdot b^{[x]}$

8. Let  $x, y \in \mathbb{R}$  we have

$$\begin{aligned} a^{[x+y]} &= \exp((x+y) \cdot \log(a)) \\ &= \exp(x \cdot \log(a) + y \cdot \log(a)) \\ &\stackrel{18.26}{=} \exp(x \cdot \log(a)) \cdot \exp(y \cdot \log(a)) \\ &= a^{[x]} \cdot a^{[y]} \end{aligned}$$

9. Let

10.  $\log(a^{[x]}) = \log(\exp(x \cdot \log(a))) = x \cdot \log(a)$

11. If we define  $f: \mathbb{R} \rightarrow \mathbb{R}$  by  $f(x) = \log(a) \cdot x$  then we have that  $(\exp \circ f)(x) = \exp(f(x)) = \exp(\log(a)) = a^{[x]} = (a^{(\cdot)})(x)$  proving that  $\exp \circ f = a^{(\cdot)}$ . As  $\forall x \in \mathbb{R}$  we have that  $\exp'(x)$  and  $f'(x)$  exists we have by 14.24 that  $(a^{(\cdot)})'(x) = \exp'(f(x)) \cdot f'(x) = \exp(f(x)) \cdot \log(a) = \log(a) \cdot \exp(\log(a) \cdot x) = \log(a) \cdot a^x$

12. Let  $x \in \mathbb{R}$  and define  $f: \mathbb{R}_+ \rightarrow \mathbb{R}$  by  $f(y) = \log(y) \cdot x$  then  $\forall y \in \mathbb{R}$  we have  $f'(y) = x \cdot \log(y) \stackrel{18.34(7)}{=} \frac{x}{y}$ . Further as  $(\exp \circ f)(y) = \exp(f(y)) = \exp(\log(y) \cdot x) = a^{[x]} = ((\cdot)^x)(y)$  we conclude that  $(\cdot)^x = \exp \circ f$ . So using 14.24 we have  $\forall x \in \mathbb{R}$  that  $((\cdot)^x)'(y) = \exp(f(y)) \cdot f'(y) = \exp(f(y)) \cdot \frac{x}{y} = x \cdot \frac{\exp(\log(y) \cdot x)}{y} = x \cdot \frac{y^{[x]}}{x} \stackrel{(4)}{=} x \cdot y^{[x-1]}$   $\square$

Using the above theorem (9) we see that the function  $a^{(\cdot)}: \mathbb{R}_+ \rightarrow \mathbb{R}_+$  has a derivative and thus by 14.19 and 14.15 we have that  $a^{(\cdot)}$  is continuous at every  $x \in \mathbb{R}_+ = ]0, \infty[$  this can be extended to  $[0, \infty[$  using the subspace topology on  $[0, \infty[$  if we extend the definition of a power to  $[0, \infty[$  which can be done if we restrict the power to strict positive numbers.

**Definition 18.38.** Let  $p \in ]0, \infty[$  and  $a \in [0, \infty[$  then we define  $a^p = \begin{cases} a^{[p]} = \exp(p \cdot \log(a)) \text{ if } a \in ]0, \infty[ \\ 0 \text{ if } a = 0 \end{cases}$

Note that the requirement  $0 < p$  is needed, for if  $a \in ]0, \infty[$  then  $a^0 = \exp(\log(a) \cdot 0) = 1$  and  $0^0 = 0$  so that the function  $[\cdot]^0: [0, \infty[ \rightarrow \mathbb{R}$  defined by  $[\cdot]^0(x)$  is discontinuous at 0. Further if  $p < 0$  then  $0^p = 0^{-p} = \frac{1}{0^p}$  is not well defined. The new definition  $a^p$  has similar properties as  $a^{[p]}$  as is expressed in the following theorem.

**Theorem 18.39.** Let  $a \in [0, \infty[ = \mathbb{R}_+$  then we that

1.  $\forall p \in ]0, \infty[$  we have  $a^p = 0 \Leftrightarrow a = 0$
2.  $\forall n \in \mathbb{N}$  we have that  $a^n = \underbrace{a \dots a}_{n \text{ times}}$  (we have a conflict of notation here so we use  $\underbrace{a \dots a}_n$  to note the classical power on  $\mathbb{R}$ )
3.  $\forall x \in ]0, \infty[ = \mathbb{R}_+$  we have  $1^x = 1$
4.  $\forall x \in ]0, \infty[$   $e^x = \exp(x)$
5.  $\forall x \in ]1, \infty[$  and  $a \neq 0$  then  $\frac{a^x}{a} = a^{x-1}$
6.  $\forall x, y \in ]0, \infty[$  we have  $(a^x)^y = a^{x \cdot y}$
7.  $\forall a, b \in [0, \infty[$  and  $\forall x \in \mathbb{R}$  we have  $(a \cdot b)^x = a^x \cdot b^x$
8.  $\forall a \in [0, \infty[$  and  $\forall x, y \in \mathbb{R}$  we have  $a^{x+y} = a^x \cdot a^y$
9.  $\forall a \in ]0, \infty[$  and  $\forall x \in \mathbb{R}$  we have  $a^{-x} = \frac{1}{a^x}$
10.  $\forall p \in ]0, \infty[$  we have that  $[\cdot]^p: [0, \infty] \rightarrow \mathbb{R}$  defined by  $[\cdot]^p(x) = x^p$  is strictly increasing

**Proof.**

1. If  $a^p = 0$  if  $a \neq 0$  we have  $a^p = \exp(p \cdot \log(a)) > 0$  a contradiction so we must have  $a = 0$ . The opposite equivalence follows from  $0^p = 0$
2. If  $n \in \mathbb{N}$  then if  $a \in ]0, \infty[$  we have  $a^n = a^{[n]} \stackrel{18.37(1)}{=} \underbrace{a \dots a}_{n \text{ times}}$  and  $0^n = 0 = \underbrace{0 \dots 0}_{1 \text{ time}}$
3. Let  $x \in ]0, \infty[$  then as  $1 \in ]0, \infty[$  we have that  $1^x = 1^{[x]} \stackrel{18.37(2)}{=} 1$
4. Let  $x \in ]0, \infty[$  then as  $e \neq 0$  we have  $e^x = e^{[x]} \stackrel{18.37(3)}{=} \exp(x)$
5. Let  $x \in ]1, \infty[$  and  $a \neq 0$  then  $\frac{a^x}{a} = \frac{a^{[x]}}{a} \stackrel{18.37}{=} a^{[x-1]} \stackrel{0 < x-1}{=} a^{x-1}$
6.  $\forall x, y \in ]0, \infty[$  we have for  $a \in [a, \infty[$  that
  - $a = \mathbf{0}$ . then  $(0^x)^y = 0^y = 0 = 0^{x \cdot y}$
  - $a \neq \mathbf{0}$ . then  $a^x = a^{[x]} \neq 0$  (see (1)) so that  $(a^x)^y = (a^{[x]})^{[y]} \stackrel{18.37}{=} a^{[x \cdot y]} = a^{x \cdot y}$
7. We have for  $a, b$  either
  - $a = \mathbf{0} \wedge b = \mathbf{0}$ . then  $(a \cdot b)^x = 0^x = 0 = 0 \cdot 0 = 0^x \cdot 0^x = a^x \cdot b^x$
  - $a = \mathbf{0} \wedge b \neq \mathbf{0}$ . then  $(a \cdot b)^x = 0^x = 0 \cdot b^x = 0^x \cdot b^x = a^x \cdot b^x$
  - $a \neq \mathbf{0} \wedge b = \mathbf{0}$ . then  $(a \cdot b)^x = 0^x = a^x \cdot 0^x = a^x \cdot 0^x = a^x \cdot b^x$
  - $a \neq \mathbf{0} \wedge b \neq \mathbf{0}$ .  $(a \cdot b)^x \stackrel{a \cdot b \neq 0}{=} (a \cdot b)^{[x \cdot y]} \stackrel{18.37(7)}{=} a^{[x]} \cdot b^{[x]} = a^x \cdot b^x$

8. We have for  $a$  either

$$\mathbf{a = 0.} \text{ then } a^{x+y} = 0^{x+y} = 0 = 0 \cdot 0 = 0^x \cdot 0^y = a^x \cdot a^y$$

$$\mathbf{a \neq 0.} \text{ then } a^{x+y} = a^{[x+y]} \stackrel{18.37(8)}{=} a^{[x]} \cdot a^{[y]} = a^x \cdot a^y$$

$$9. \text{ As } a \in ]0, \infty[ \text{ we have } a^{-x} = a^{[-x]} \stackrel{18.37(6)}{=} \frac{1}{a^{[x]}} = \frac{1}{a^x}$$

10. Let  $x, y \in [0, \infty[$  with  $x < y$  then we have for  $x$  either

$$\mathbf{x = 0.} \text{ then } 0 < y \text{ and } [.]^p(x) = [.]^p(0) = 0 < \exp(p \cdot \log(x)) = [.]^p(x)$$

$$\mathbf{0 < x.} \text{ then } \log(x) <_{18.34(8)} 0 <_p p \cdot \log(x) < p \cdot \log(y) \stackrel{18.27(6)}{\Rightarrow} \exp(p \cdot \log(x)) < \exp(p \cdot \log(y)) \\ \text{so that } [.]^p(x) = \exp(p \cdot \log(x)) < \exp(p \cdot \log(y)) = [.]^p(y) \quad \square$$

**Theorem 18.40.** Let  $p \in ]0, \infty[$  then  $[.]^p: [0, \infty[ \rightarrow \mathbb{R}$  defined by  $[.]^p(x) = x^p$  then we have  $[.]^p$  is continuous  $\forall x \in [0, \infty[$  where we use the canonical topology on  $\mathbb{R}$  based on the norm  $\| \cdot \|$  and the subspace topology on  $[0, \infty[$  (see 11.6)

**Proof.** Let  $x \in [0, \infty[$  the we can divide the proof in two pieces

$\mathbf{x \neq 0.}$  Let  $[.]^p(x) \in B_{\|}(x, \varepsilon)$  then as  $(.)^p: ]0, \infty[ \rightarrow \mathbb{R}$  is continuous  $((.)^p)'(x)$  exists together with 14.19 and 14.15) there exists a  $B_{\|}(x, \delta_1)$  such that  $\forall y \in B_{\|}(x, \delta_1) \cap ]0, \infty[$  we have  $[.]^p(y) = y^p = (.)^p(y) \in B_{\|}(x, \varepsilon)$ . Further as  $x \neq 0$  we have  $0 < x$  take then  $\delta = \min(x, \delta_1)$ . If  $0 \in B_{\|}(x, \delta)$  then  $x = x - 0 < \delta \leq x$  a contradiction so that  $0 \notin B_{\|}(x, \delta)$  which proves that  $B_{\|}(x, \delta) \cap ]0, \infty[ = B_{\|}(x, \delta) \cap [0, \infty[$ . Finally  $\forall y \in B_{\|}(x, \delta) \cap [0, \infty[ = B_{\|}(x, \delta) \cap ]0, \infty[ \subseteq B_{\|}(x, \delta_1) \cap ]0, \infty[$  we have that  $[.]^p(y) \in B_{\|}(x, \varepsilon)$  proving that  $[.]^p$  is continuous at  $x$ .

$\mathbf{x = 0.}$  Let  $\varepsilon > 0$  then as  $\exp: \mathbb{R} \rightarrow \mathbb{R}_+$  is a bijection there exists a  $\delta_1 \in \mathbb{R}$  such that  $\exp(\delta_1) = \varepsilon$ .

Further as  $\log: \mathbb{R}^+ \rightarrow \mathbb{R}$  is a bijection there exists a  $\delta < \mathbb{R}_+$  such that  $\log(\delta) = \frac{\delta_1}{p}$ . Hence if  $y \in B_{\|}(0, \delta) \cap [0, \infty[$  then  $y < \delta$  so that  $\log(y) < \log(\delta)$  [as  $\log$  is strictly increasing], as  $0 < p$  if follows that  $p \cdot \log(y) < p \cdot \log(\delta) \stackrel{\text{exp is increasing}}{\Rightarrow} \exp(p \cdot \log(y)) < \exp(p \cdot \log(\delta)) = \exp\left(p \cdot \frac{\delta_1}{p}\right) = \exp(\delta_1) = \varepsilon$ . Finally as  $y^p = \exp(p \cdot \log(y)) > 0$  it follows that  $y^p \in B_{\|}(0, \varepsilon)$  which proves that  $[.]^p$  is continuous at 0  $\square$

Now that we have extended the concept of powers we can prove Young' inequality. First we need a little lemma.

**Lemma 18.41.** Let  $p \in ]1, \infty[$  and  $f_p: [0, \infty[ \rightarrow \mathbb{R}$  defined by  $f_p(x) = (1 - x^{1-p}) + (p-1) \cdot (1-x)$  then  $x = 1 \Leftrightarrow f_p(x) = 0$

**Proof.**

$$\Rightarrow. \text{ If } x = 1 \text{ then } f_p(1) = (1 - 1^{1-p}) + (p-1) \cdot (1-1) \stackrel{18.39(2)}{=} (1-1) + (p-1) \cdot (1-1) = 0$$

$$\Leftarrow. \text{ Let } x \in ]0, \infty[ \text{ then } f'_p(x) = -(1-p) \cdot x^{-p} - (p-1) = (p-1) \cdot (x^{-p} - 1) \text{ and } f''_p(x) = (p-1) \cdot (-p) \cdot x^{-p} \stackrel{0 < p-1, -p < 0, 0 < x^{-p}}{\Rightarrow} f''_p(x) < 0 \text{ proving by 14.99 that } f' \text{ is strictly decreasing on } ]0, \infty[$$

$$f'_p(1) = (p-1) \cdot (1^{-p} - 1) = (p-1) \cdot (1-1) = 0 \quad (18.20)$$

and

$$f' \text{ is strictly decreasing} \quad (18.21)$$

Assume now that  $\exists x \in ]0, \infty[$  such that  $x \neq 1$  and  $f_p(x) = 0$  we have either

$\mathbf{x < 1.}$  then as  $f_p(x) = 0 = f_p(1)$  we have by the theorem of Rolle (see 14.97) that  $\exists z \in ]x, 1[$  such that  $f'_p(z) = 0$ . As  $f'$  is strictly increasing we must have that  $f'_p(z) < f'_p(1) = 0$  giving  $0 < 0$  a contradiction.

$\mathbf{1 < x.}$  hen as  $f_p(1) = 0 = f_p(0)$  we have by the theorem of Rolle (see 14.97) that  $\exists z \in ]1, z[$  such that  $f'_p(z) = 0$ . As  $f'$  is strictly increasing we must have that  $0 = f'_p(1) < f'_p(z)$  giving  $0 < 0$  a contradiction.

As we have in both cases a contradiction we must have that  $\forall x \in ]0, \infty[$  with  $x \neq 1$  that  $f_p(x) \neq 0$ . So if  $f_p(x) = 0$  we must have  $x = 1$   $\square$

**Theorem 18.42. (Young's inequality)** Let  $p, q \in ]0, \infty[$  such that  $\frac{1}{p} + \frac{1}{q} = 1$  and  $a, b \in ]0, \infty[$  then we have that  $a \cdot b \leq \frac{a^p}{p} + \frac{b^q}{q}$ . Further if  $p \in ]0, \infty[$  then we have that  $a \cdot b = \frac{a^p}{p} + \frac{b^q}{q} \Leftrightarrow b = a^{p-1} \Leftrightarrow a^p = b^q$

**Proof.**

$$\begin{aligned}
 a \cdot b &= \exp(\log(a \cdot b)) \\
 &\stackrel{18.34}{=} \exp(\log(a) + \log(b)) \\
 &= \exp\left(\frac{1}{p} \cdot p \cdot \log(a) + \frac{1}{q} \cdot q \cdot \log(b)\right) \\
 &\stackrel{18.39}{=} \exp\left(\frac{1}{p} \cdot \log(a^p) + \frac{1}{q} \cdot \log(b^q)\right) \\
 &\leq \exp \text{ is convex (see 18.28) and 14.102} \quad \frac{1}{p} \cdot \exp(\log(a^p)) + \frac{1}{q} \cdot \exp(\log(b^q)) \\
 &= \frac{1}{p} \cdot a^p + \frac{1}{q} \cdot b^q \\
 &= \frac{a^p}{p} + \frac{b^q}{q}
 \end{aligned}$$

proving that

$$a \cdot b \leq \frac{a^p}{p} + \frac{b^q}{q} \quad (18.22)$$

Further we have that

$$\begin{aligned}
 \frac{1}{p} + \frac{1}{q} = 1 &\Rightarrow \frac{p}{p} + \frac{p}{q} = p \\
 &\Rightarrow 1 + \frac{p}{q} = p \\
 &\Rightarrow \frac{p}{q} = (p-1) \\
 &\Rightarrow \frac{q}{p} + \frac{q}{q} = q \\
 &\Rightarrow 1 + \frac{q}{p} = q \\
 &\Rightarrow \frac{q}{p} = (q-1)
 \end{aligned}$$

so we have by the hypothese that

$$\frac{p}{q} = (p-1) \wedge \frac{q}{p} = (q-1) \text{ and thus } 1 = (p-1) \cdot (q-1) \quad (18.23)$$

Next we assume that additionally  $p \in ]1, \infty[$ . First we have

$$\begin{aligned}
 a^p = b^q &\Leftrightarrow_{0 < q} (a^p)^{\frac{1}{q}} = (b^q)^{\frac{1}{q}} \\
 &\stackrel{18.39 (5)}{\Leftrightarrow} a^{\frac{p}{q}} = b^{\frac{q}{q}} \\
 &\Leftrightarrow a^{\frac{p}{q}} = b^1 \\
 &\Leftrightarrow a^{\frac{p}{q}} = b \\
 &\stackrel{18.23}{\Leftrightarrow} a^{p-1} = b
 \end{aligned}$$

and

$$\begin{aligned}
 a^p = b^q &\Leftrightarrow_{0 < p} (a^p)^{\frac{1}{p}} = (b^q)^{\frac{1}{p}} \\
 &\stackrel{18.39 (5)}{\Leftrightarrow} a^{\frac{p}{p}} = b^{\frac{q}{p}} \\
 &\Leftrightarrow a^1 = b^{\frac{q}{p}} \\
 &\Leftrightarrow a = b^{\frac{q}{p}} \\
 &\stackrel{18.23}{\Leftrightarrow} a = b^{q-1}
 \end{aligned}$$

so that we have

$$a^p = a^q \Leftrightarrow a^{p-1} = b \Leftrightarrow a = b^{\frac{q}{p}} \Leftrightarrow b = a^{\frac{p}{q}} \Leftrightarrow a = b^{q-1} \quad (18.24)$$

$\Leftarrow$ . If  $a = b^{p-1}$  then we have

$$\begin{aligned} \frac{a^p}{p} + \frac{b^q}{q} &= \frac{a^p \cdot q + b^q \cdot p}{p \cdot q} \\ &\stackrel{18.24}{=} \frac{q \cdot (b^{\frac{q}{p}})^p + b^q \cdot p}{p \cdot q} \\ &= \frac{q \cdot b^{\frac{q}{p} \cdot p} + p \cdot b^q}{p \cdot q} \\ &= \frac{q \cdot b^q + p \cdot b^q}{p \cdot q} \\ &= \frac{q + p \cdot b^q}{p \cdot q} \\ &= \left( \frac{1}{p} + \frac{1}{q} \right) \cdot b^q \\ &= b^q \\ &= b \cdot b^{q-1} \\ &\stackrel{18.24}{=} a \cdot b \end{aligned}$$

proving that  $\frac{a^p}{p} + \frac{b^q}{q} = a \cdot b$

$\Rightarrow$ . Suppose that  $a \cdot b = \frac{a^p}{p} + \frac{b^q}{q}$  then we have

$$\begin{aligned} a \cdot b &= \frac{q \cdot a^p + p \cdot b^q}{p \cdot q} \\ &\stackrel{\frac{1}{p} + \frac{1}{q} = 1 \Rightarrow \frac{p+q}{p \cdot q} = 1 \Rightarrow \frac{1}{p \cdot q} = \frac{1}{p+q}}{=} \frac{q \cdot a^p + p \cdot b^q}{p + q} \end{aligned}$$

so that

$$\begin{aligned} p + q = \frac{q \cdot a^p + p \cdot b^q}{a \cdot b} &\Rightarrow p + q = q \cdot \frac{a^{p-1}}{b} + p \cdot \frac{b^{q-1}}{a} \\ &\Rightarrow p \cdot \left( 1 - \frac{b^{q-1}}{a} \right) = q \cdot \left( \frac{a^{p-1}}{b} - 1 \right) \\ &\Rightarrow \left( \frac{a^{p-1}}{b} - 1 \right) = \frac{p}{q} \cdot \left( 1 - \frac{b^{q-1}}{a} \right) \\ &\stackrel{18.23}{\Rightarrow} \left( \frac{a^{p-1}}{b} - 1 \right) = (p-1) \cdot \left( 1 - \frac{b^{q-1}}{a} \right) \\ &\stackrel{18.39(6)}{\Rightarrow} \left( \frac{b^{-1}}{a^{-(p-1)}} - 1 \right) = (p-1) \cdot \left( 1 - \frac{b^{q-1}}{a} \right) \\ &\stackrel{18.23}{\Rightarrow} \left( \frac{b^{-(p-1) \cdot (q-1)}}{a^{-(p-1)}} - 1 \right) = (p-1) \cdot \left( 1 - \frac{b^{q-1}}{a} \right) \\ &\Rightarrow \left( \frac{b^{(1-p) \cdot (q-1)}}{a^{(1-p)}} - 1 \right) = (p-1) \cdot \left( 1 - \frac{b^{q-1}}{a} \right) \\ &\Rightarrow \left( \left( \frac{b^{q-1}}{a} \right)^{1-p} - 1 \right) = (p-1) \cdot \left( 1 - \frac{b^{q-1}}{a} \right) \\ &\Rightarrow 0 = \left( 1 - \left( \frac{b^{q-1}}{a} \right)^{1-p} \right) + (p-1) \cdot \left( 1 - \frac{b^{q-1}}{a} \right) \end{aligned}$$

so if we define  $f_p: ]0, \infty[ \rightarrow \mathbb{R}$  by  $f_p(x) = (1 - x^{1-p}) + (p-1) \cdot (1-x)$  we have by the above that  $f_p\left(\frac{b^{q-1}}{a}\right) = 0$ . Using the previous lemma (see 18.41) we conclude then that  $\frac{b^{q-1}}{a} = 1$  or that  $a = b^{q-1}$  or using 18.23 that

$$b = a^{p-1} \quad \square$$

Next we examine the relation between the exponential map and goniometric functions, for that we must consider the behavior of  $\exp$  on the space of complex numbers.

**Theorem 18.43.**  $\exp: \mathbb{C} \rightarrow \mathbb{C}$  has the following properties

1.  $\forall z \in \mathbb{C}$  we have that  $\overline{\exp(z)} = \exp(\bar{z})$
2.  $\forall z \in \mathbb{C}$  we have  $|\exp(z)|^2 = \exp(\operatorname{Re}(z))$
3.  $\forall z = x + i \cdot y \in \mathbb{C}$  we have that  $|\exp(z)| = \exp(x)$  and  $|\exp(i \cdot y)| = \exp(0) = 1$

**Proof.**

1.

$$\begin{aligned} \exp(\bar{z}) &= \sum_{i=0}^{\infty} \frac{\bar{z}^i}{i!} \\ &= \lim_{n \rightarrow \infty} \left( \sum_{i=0}^n \frac{\bar{z}^i}{i!} \right) \\ &= \lim_{n \rightarrow \infty} \left( \sum_{i=0}^n \frac{\overline{z^i}}{i!} \right) \\ &= \lim_{n \rightarrow \infty} \overline{\sum_{i=0}^{\infty} \frac{z^i}{i!}} \\ &\stackrel{11.333}{=} \overline{\lim_{n \rightarrow \infty} \left( \sum_{i=0}^{\infty} \frac{z^i}{i!} \right)} \\ &= \exp(z) \end{aligned}$$

2.

$$\begin{aligned} |\exp(z)|^2 &= \exp(z) \cdot \overline{\exp(z)} \\ &\stackrel{(1)}{=} \exp(z) \cdot \exp(\bar{z}) \\ &= \exp(z + \bar{z}) \\ &= \exp(2 \cdot \operatorname{Re}(z)) \\ &= \exp(\operatorname{Re}(z))^2 \end{aligned}$$

Now as  $|\exp(z)|, \exp(\operatorname{Re}(z)) \in \mathbb{R}_+$  we have by 9.71 that  $|\exp(z)| = \exp(\operatorname{Re}(z))$

3. If  $x = x + i \cdot y$  then  $|\exp(z)| \stackrel{(2)}{=} \exp(\operatorname{Re}(z)) = \exp(x)$ . Further  $|\exp(i \cdot y)| = \exp(\operatorname{Re}(i \cdot y)) = \exp(0) = 1$   $\square$

We can define the trigonometric functions in terms of the exponential function

**Definition 18.44. (trigonometric functions)**

1.  $\cos: \mathbb{C} \rightarrow \mathbb{C}$  is defined by  $\cos = \frac{\exp(i \cdot z) + \exp(-i \cdot z)}{2}$
2.  $\sin: \mathbb{C} \rightarrow \mathbb{C}$  is defined by  $\sin = \frac{\exp(i \cdot z) - \exp(-i \cdot z)}{2 \cdot i}$

**Theorem 18.45.** We have the following properties for the trigonometric functions

1.  $\cos: \mathbb{C} \rightarrow \mathbb{C}$  and  $\sin: \mathbb{C} \rightarrow \mathbb{C}$  are continuous on  $\mathbb{C}$

2.  $\forall z \in \mathbb{C}$  we have that  $\cos'(z) = -\sin(z)$  and  $\sin'(z) = \cos(z)$
3.  $\cos(0) = 1$
4.  $\sin(0) = 0$
5.  $\forall z \in \mathbb{C}$  we have
  - a.  $\cos(-z) = \cos(z)$
  - b.  $\sin(-z) = -\sin(z)$
6.  $\forall z \in \mathbb{C}$  we have  $\cos(z)^2 + \sin(z)^2 = 1$
7.  $\forall x, y \in \mathbb{C}$  we have
  - a.  $\cos(x+y) = \cos(x) \cdot \cos(y) - \sin(x) \cdot \sin(y)$
  - b.  $\sin(x+y) = \sin(x) \cdot \cos(y) + \cos(x) \cdot \sin(y)$
8.  $\forall x \in \mathbb{R}$  we have
  - a.  $\cos(x) \in \mathbb{R}$
  - b.  $\sin(x) \in \mathbb{R}$
9.  $\forall x \in \mathbb{C}$  we have  $e^{i \cdot x} = \cos(x) + i \cdot \sin(x)$  (which by (8) means that for  $x \in \mathbb{R}$  we have  $\operatorname{Re}(i \cdot x) = \cos(x)$  and  $\operatorname{Im}(e^{i \cdot x}) = \sin(x)$ )
10.  $\forall x \in \mathbb{R}$  we have  $|e^{i \cdot x}| = 1$

**Proof.**

1. This follows as the product of a scalar with a continuous function is continuous and the sum of continuous functions is continuous.
2. Let  $z \in \mathbb{C}$  then

$$\begin{aligned}
 \cos'(z) &= \frac{1}{2} \cdot (i \cdot \exp(i \cdot z) + (-i) \cdot \exp(-i \cdot z)) \\
 &= \frac{i \cdot i \cdot \exp(i \cdot z) + i \cdot (-i) \cdot \exp(-i \cdot z)}{2 \cdot i} \\
 &= \frac{-\exp(i \cdot z) + \exp(-i \cdot z)}{2 \cdot i} \\
 &= -\sin(z)
 \end{aligned}$$

and

$$\begin{aligned}
 \sin'(z) &= \frac{1}{2 \cdot i} \cdot (i \cdot \exp(i \cdot z) - (-i) \cdot \exp(-i \cdot z)) \\
 &= \frac{\exp(i \cdot z) + \exp(-i \cdot z)}{2} \\
 &= \cos(z)
 \end{aligned}$$

3.

$$\begin{aligned}
 \cos(0) &= \frac{\exp(i \cdot 0) + \exp(-i \cdot 0)}{2} \\
 &= \frac{1+1}{2} \\
 &= 1
 \end{aligned}$$

4.

$$\begin{aligned}
 \sin(0) &= \frac{\exp(i \cdot 0) - \exp(-i \cdot 0)}{2 \cdot i} \\
 &= \frac{1-1}{2 \cdot i} \\
 &= 0
 \end{aligned}$$

5.

a.

$$\begin{aligned}
 \cos(-z) &= \frac{\exp(i \cdot (-z)) + \exp(-i \cdot (-z))}{2} \\
 &= \frac{\exp(i \cdot z) + \exp(-i \cdot z)}{2} \\
 &= \cos(z)
 \end{aligned}$$

b.

$$\begin{aligned}
 \sin(-z) &= \frac{\exp(i \cdot (-z)) - \exp(-i \cdot (-z))}{2 \cdot i} \\
 &= \frac{\exp(-i \cdot z) - \exp(i \cdot z)}{2 \cdot i} \\
 &= -\sin(z)
 \end{aligned}$$

6.

$$\begin{aligned}
 \cos(z)^2 + \sin(z)^2 &= \frac{(\exp(i \cdot z) + \exp(-i \cdot z))^2}{4} + \frac{(\exp(i \cdot z) - \exp(-i \cdot z))^2}{-4} \\
 &= \frac{1}{4} \cdot ((\exp(i \cdot z))^2 + (\exp(-i \cdot z))^2 + 2 \cdot \exp(i \cdot z) \cdot \exp(-i \cdot z) - (\exp(i \cdot z))^2 - \\
 &\quad (\exp(-i \cdot z))^2 + 2 \cdot \exp(i \cdot z) \cdot \exp(-i \cdot z)) \\
 &= \frac{1}{4} \cdot (4 \cdot \exp(i \cdot z) \cdot \exp(-i \cdot z)) \\
 &= \exp(i \cdot z) \cdot \exp(-i \cdot z) \\
 &= \exp(i \cdot z - i \cdot z) \\
 &= \exp(0) \\
 &= 1
 \end{aligned}$$

7.

a.

$$\begin{aligned}
 \cos(x) \cdot \cos(y) - \sin(x) \cdot \sin(y) &= \frac{\exp(i \cdot x) + \exp(-i \cdot x)}{2} \cdot \frac{\exp(i \cdot y) + \exp(-i \cdot y)}{2} - \\
 &\quad \frac{\exp(i \cdot x) - \exp(-i \cdot x)}{2 \cdot i} \cdot \frac{\exp(i \cdot y) - \exp(-i \cdot y)}{2 \cdot i} \\
 &= \frac{1}{4} \cdot (\exp(i \cdot x) \cdot \exp(i \cdot y) + \exp(i \cdot x) \cdot \exp(-i \cdot y) + \\
 &\quad \exp(-i \cdot x) \cdot \exp(i \cdot y) + \exp(-i \cdot x) \cdot \exp(-i \cdot y)) + \\
 &\quad \frac{1}{4} \cdot (\exp(i \cdot x) \cdot \exp(i \cdot y) - \exp(i \cdot x) \cdot \exp(-i \cdot y) - \\
 &\quad \exp(-i \cdot x) \cdot \exp(i \cdot y) + \exp(-i \cdot x) \cdot \exp(-i \cdot y)) \\
 &= \frac{1}{4} \cdot (\exp(i \cdot x) \cdot \exp(i \cdot y) + \exp(i \cdot x) \cdot \exp(-i \cdot y) + \\
 &\quad \exp(-i \cdot x) \cdot \exp(i \cdot y) + \exp(-i \cdot x) \cdot \exp(-i \cdot y) + \\
 &\quad \exp(i \cdot x) \cdot \exp(i \cdot y) - \exp(i \cdot x) \cdot \exp(-i \cdot y) - \\
 &\quad \exp(-i \cdot x) \cdot \exp(i \cdot y) + \exp(-i \cdot x) \cdot \exp(-i \cdot y)) \\
 &= \frac{1}{4} \cdot (2 \cdot \exp(i \cdot x) \cdot \exp(i \cdot y) + 2 \cdot \exp(-i \cdot x) \cdot \exp(-i \cdot y)) \\
 &= \frac{1}{2} \cdot (\exp(i \cdot x + i \cdot y) + \exp(-i \cdot x - i \cdot y)) \\
 &= \frac{\exp(i \cdot (x + y)) + \exp(-i \cdot (x + y))}{2} \\
 &= \cos(x + y)
 \end{aligned}$$

b.

$$\begin{aligned}
\sin(x) \cdot \cos(y) + \cos(x) \cdot \sin(y) &= \frac{\exp(i \cdot x) - \exp(-i \cdot x)}{2 \cdot i} \cdot \frac{\exp(i \cdot y) + \exp(-i \cdot y)}{2} + \\
&\quad \frac{\exp(i \cdot x) + \exp(-i \cdot x)}{2} \cdot \frac{\exp(i \cdot y) - \exp(-i \cdot y)}{2 \cdot i} \\
&= \frac{1}{4 \cdot i} \cdot (\exp(i \cdot x) \cdot \exp(i \cdot y) + \exp(i \cdot x) \cdot \exp(-i \cdot y) - \exp(-i \cdot x) \cdot \exp(i \cdot y) - \exp(-i \cdot x) \cdot \exp(-i \cdot y) + \exp(i \cdot x) \cdot \exp(i \cdot y) - \exp(i \cdot x) \cdot \exp(-i \cdot y) + \exp(-i \cdot x) \cdot \exp(i \cdot y) - \exp(-i \cdot x) \cdot \exp(-i \cdot y)) \\
&= \frac{1}{4 \cdot i} \cdot (2 \cdot \exp(i \cdot x) \cdot \exp(i \cdot y) - 2 \cdot \exp(-i \cdot x) \cdot \exp(-i \cdot y)) \\
&= \frac{\exp(i \cdot x + i \cdot y) - \exp(-i \cdot x - i \cdot y)}{2 \cdot i} \\
&= \frac{\exp(i \cdot (x + y)) - \exp(-i \cdot (x + y))}{2 \cdot i} \\
&= \sin(x + y)
\end{aligned}$$

8. Let  $x \in \mathbb{R}$  then

a.

$$\begin{aligned}
\cos(x) &= \frac{\exp(i \cdot x) + \exp(-i \cdot x)}{2} \\
&= \frac{\exp(i \cdot x) + \exp(\overline{i \cdot x})}{2} \\
&\stackrel{18.43 \ (1)}{=} \frac{\exp(i \cdot x) + \overline{\exp(i \cdot x)}}{2} \\
&= \frac{2 \cdot \operatorname{Re}(\exp(i \cdot x))}{2} \\
&= \operatorname{Re}(\exp(i \cdot x)) \\
&\in \mathbb{R}
\end{aligned}$$

b.

$$\begin{aligned}
\sin(x) &= \frac{\exp(i \cdot x) - \exp(-i \cdot x)}{2 \cdot i} \\
&= \frac{\exp(i \cdot x) - \exp(\overline{i \cdot x})}{2 \cdot i} \\
&= \frac{\exp(i \cdot x) - \exp(\overline{i \cdot x})}{2 \cdot i} \\
&= \frac{2 \cdot i \cdot \operatorname{Img}(\exp(i \cdot x))}{2 \cdot i} \\
&= \operatorname{Img}(\exp(i \cdot x)) \\
&\in \mathbb{R}
\end{aligned}$$

9. Let  $x \in \mathbb{R}$  then

$$\begin{aligned}
\cos(x) + i \cdot \sin(x) &= \frac{\exp(i \cdot x) + \exp(-i \cdot x)}{2} + i \cdot \frac{\exp(i \cdot x) - \exp(-i \cdot x)}{2 \cdot i} \\
&= \frac{1}{2} \cdot (\exp(i \cdot x) + \exp(-i \cdot x) + \exp(i \cdot x) - \exp(-i \cdot x)) \\
&= \exp(i \cdot x)
\end{aligned}$$

10. Let  $x \in \mathbb{R}$  then

$$\begin{aligned} |e^{i \cdot x}| &\stackrel{(9)}{=} |\cos(x) + i \cdot \sin(x)| \\ &\stackrel{\cos(x), \sin(x) \in \mathbb{R}}{=} \sqrt{\cos^2(x) + \sin^2(x)} \\ &\stackrel{(6)}{=} 1 \\ &\square \end{aligned}$$

The next theorem shows that  $\cos: \mathbb{R} \rightarrow \mathbb{R}$  has a zero, this will allow us to define  $\pi$  and show that sin and cos are periodically.

**Theorem 18.46.**  $\exists x_0 \in [0, \infty[$  such that  $\cos(x) = 0$

**Proof.** We proof this by contradiction so let assume that  $\forall x \in [0, \infty[$  we have that  $\cos(x) \neq 0$ . If now there exists a  $x \in [0, \infty[$  such that  $\cos(x) \leq 0 < 1 \stackrel{18.45 (3)}{=} \cos(0) \Rightarrow 0 \in [\cos(x), 1]$  we have by the continuity of exp and the intermediate value theorem (see 11.456) that there exists a  $y \in [x, 1]$  such that  $\cos(y) = 0$  contradicting the assumption that  $\cos(x) \neq 0$ . So we conclude that  $\forall x \in [0, \infty[$  we have  $0 < \cos(x)$ , as  $\sin'(x) = \cos(x)$  we conclude that  $\forall x \in [0, \infty[$   $\sin'(x) > 0$  hence by 14.99 we have that

$$\sin: [0, \infty[ \rightarrow [0, \infty[ \text{ is strictly increasing} \quad (18.25)$$

As  $\sin(0) \stackrel{18.44 (4)}{=} 0$  we conclude that

$$\forall x \in [0, \infty[ \text{ that } 0 < \sin(x) \quad (18.26)$$

Further we have by 18.45 (6) that  $\forall x \in \mathbb{R}$  that  $\cos^2(x) + \sin^2(x) = 1$  so that  $|\cos(x)|^2 = \cos^2(x) = 1 - \sin^2(x) \leq 1$  so as  $0 \leq |\cos(x)|$  and  $\sqrt{\cdot}$  is increasing (see 9.72) we have that  $|\cos(x)| = \sqrt{|\cos(x)|^2} \leq \sqrt{1} = 1$  so that

$$\forall x \in \mathbb{R} \text{ we have } \cos(x), -\cos(x) \leq 1 \quad (18.27)$$

We have then for  $0 < x < y$  that if we take  $C_{\sin(x)}: [0, x] \rightarrow \mathbb{R}$  defined by  $C_{\sin(x)}(t) = \sin(t)$  so that  $\forall t \in [x, y]$  we have  $C_{\sin(x)}(t) \stackrel{18.25}{=} \sin(x) < \sin(t)$

$$\begin{aligned} \sin(x) \cdot (y - x) &\stackrel{11.445}{=} \int_x^y C_{\sin(x)} \\ &\leq C_{\sin(x)}(t) \leq \sin(t) \text{ and } 11.441 \\ &\stackrel{\text{fundamental theorem of calculus (14.107)}}{=} -\int_x^y \cos' \\ &\stackrel{\cos' = -\sin}{=} -(\cos(y) - \cos(x)) \\ &= \cos(x) - \cos(y) \\ &\stackrel{18.27}{\leq} 2 \end{aligned}$$

proving that

$$\forall x, y \in \mathbb{R} \text{ with } 0 < x < y \text{ we have } (y - x) \leq \frac{2}{\sin(x)} \quad (18.28)$$

Take now  $0 < x$  and take  $y = \frac{2}{\sin(x)} + x + 1$  then as  $0 < \sin(x)$  we have that  $0 < x < y$  and  $y - x = \frac{2}{\sin(x)} + 1 > \frac{2}{\sin(x)} \geq 18.28 (y - x)$  a contradiction. So our assumption is wrong and ther must exists a  $x \in [0, \infty[$  such that  $\cos(x) = 0$   $\square$

**Corollary 18.47.**  $\{x \in [0, \infty[ | \cos(x) = 0\} = \cos^{-1}(\{0\}) \cap [0, \infty[$  has a minimum

**Proof.** Using the previous theorem we have that  $\cos^{-1}(\{0\}) \cap [0, \infty[ \neq \emptyset$  so that by the conditionally completeness of the real numbers (see 9.43)  $x_0 = \cos^{-1}(\{0\})$  exists. Further as  $\{0\}$  is closed (see 11.229),  $[0, \infty[$  is closed (see 11.74) and the continuity of  $\cos$  it follows that  $\cos^{-1}(\{0\}) \cap [0, \infty[$  is closed. Assume now that  $x_0 \notin \cos^{-1}(\{0\}) \cap [0, \infty[$  then  $x_0 \in \mathbb{R} \setminus (\cos^{-1}(\{0\}) \cap [0, \infty[)$  a open set so that  $\exists \delta > 0$  such that  $]x_0 - \delta, x_0 + \delta[ = B_{\mathbb{R}}(x_0, \delta) \subseteq \mathbb{R} \setminus (\cos^{-1}(\{0\}) \cap [0, \infty[)$  proving that  $]x_0 - \delta, x_0 + \delta[ \cap \cos^{-1}(\{0\}) \cap [0, \infty[ = \emptyset$ . As  $x_0 < x_0 + \delta$  there exists by the definition of a infimum a  $y \in \cos^{-1}(\{0\}) \cap [0, \infty[$  such that  $x_0 \leq y < x_0 + \delta$  proving that  $(\cos^{-1}(\{0\}) \cap [0, \infty[) \cap ]x_0 - \delta, x_0 + \delta[ \neq \emptyset$  a contradiction. So we must have that  $x_0 \in \cos^{-1}(\{0\}) \cap [0, \infty[$  which as  $\forall y \in \cos^{-1}(\{0\}) \cap [0, \infty[$  we have  $x_0 \leq y$  means that  $x_0 = \min(\cos^{-1}(\{0\}) \cap [0, \infty[)$   $\square$

The above corollary allows us to define  $\pi$  for the trigonometric functions.

**Definition 18.48.** ( $\pi$ )  $\pi = 2 \cdot \min(\cos^{-1}(\{0\}) \cap [0, \infty[)$  so that  $\frac{\pi}{2} = \min(\cos^{-1}(\{0\}) \cap [0, \infty[)$  hence  $\cos(\frac{\pi}{2}) = 0$  and  $\forall x \in [0, \frac{\pi}{2}[$  we have  $\cos(\frac{\pi}{2}) \neq 0$ . Note that as  $\cos(0) = 1$  we have that  $0 < \frac{\pi}{2}$ .

Next we prove that  $\pi$  and the trigonometric functions have the usual properties

**Theorem 18.49.** We have

1.  $\forall x \in [0, \frac{\pi}{2}[$  we have  $0 < \cos(x)$  (hence as  $\cos(\frac{\pi}{2}) = 0$  we have  $\forall x \in [0, \frac{\pi}{2}]$  that  $0 \leq \cos(x)$ )
2.  $\forall x \in ]0, \frac{\pi}{2}]$  we have  $0 < \sin(x)$  (hence as  $\sin(0) = 0$  we have  $\forall x \in [0, \frac{\pi}{2}]$  that  $0 \leq \sin(x)$ )

**Proof.**

1. Assume that  $\exists y \in [0, \frac{\pi}{2}[$  such that  $\cos(y) \leq 0$  then  $\cos(y) < 1 \stackrel{18.45 \text{ (3)}}{=} \cos(0)$  so using the intermediate value theorem (see 11.456) there exists a  $z \in [0, y]$  such that  $\cos(z) = 0$ . Hence as  $y < \frac{\pi}{2}$  we have found a  $z \in [0, \frac{\pi}{2}[$  such that  $\cos(z) = 0$  contradicting  $\frac{\pi}{2} = \min(\cos^{-1}(\{0\}) \cap [0, \infty[)$ . So we must have that  $\forall y \in [0, \frac{\pi}{2}[$  that  $0 < \cos(y)$ )
2. As  $\forall y \in [0, \frac{\pi}{2}[$  we have  $\sin'(x) \stackrel{18.45 \text{ (2)}}{=} \cos(x) > 0$  (by (1)) we have by 14.99 that  $\sin$  is strictly increasing on  $[0, \frac{\pi}{2}[$  so that

$$\forall x \in ]0, \frac{\pi}{2}[ \text{ we have } 0 < \sin(x) \quad (18.29)$$

More special we have that for  $0 < \frac{\pi}{4} < \frac{\pi}{2}$  that  $0 < \sin(\frac{\pi}{4})$ , then as  $\forall x \in [\frac{\pi}{4}, \frac{\pi}{2}]$  we have by (2) that  $\sin'(x) = \cos(x) \geq 0$  that by 14.99  $\sin(x)$  is increasing on  $\square$

**Theorem 18.50.** We have the following properties for  $\pi$  and the trigonometric functions

1.  $\cos(0) = 1$
2.  $\cos(\frac{\pi}{2}) = 0$
3.  $\sin(0) = 0$
4.  $\sin(\frac{\pi}{2}) = 1$
5.  $\cos(\pi) = -1$
6.  $\sin(\pi) = 0$
7.  $\cos(2 \cdot \pi) = 1$
8.  $\sin(2 \cdot \pi) = 0$
9.  $\forall x \in \mathbb{R}$  we have
  - a.  $\sin(x + 2 \cdot \pi) = \sin(x)$
  - b.  $\cos(x + 2 \cdot \pi) = \cos(x)$
10.  $\forall x \in \mathbb{R}$  we have  $\exp(i \cdot (x + 2 \cdot \pi)) = \exp(i \cdot x)$  (or using the power notation  $e^{i \cdot (x + 2 \cdot \pi)} = e^{i \cdot x}$ )
11.  $\exp(i \cdot \pi) = -1$  (or using the power notation  $e^{i \cdot \pi} = -1$ ) (the most beautifull equation in the world because it shows how three fundamental constants ( $\pi, e$  and  $1$ ) are related)

**Proof.**

1. This was already proved in 18.45 (3)
2. This follows from the definition of  $\frac{\pi}{2}$
3. This was already proved in 18.45 (4)
4. Using 18.45 (6) we have that  $\cos^2\left(\frac{\pi}{2}\right) + \sin^2\left(\frac{\pi}{2}\right) = 1 \xrightarrow{(2)} \sin^2\left(\frac{\pi}{2}\right) = 1$  so that  $\sin\left(\frac{\pi}{2}\right) = 1 \vee \sin\left(\frac{\pi}{2}\right) = -1$ , as  $\sin\left(\frac{\pi}{2}\right) > 1$  (see 18.49) it follows that  $\sin\left(\frac{\pi}{2}\right) = 1$
- 5.

$$\begin{aligned}
 \cos(\pi) &= \cos\left(\frac{\pi}{2} + \frac{\pi}{2}\right) \\
 &\stackrel{18.45 \text{ (7)}}{=} \cos\left(\frac{\pi}{2}\right) \cdot \cos\left(\frac{\pi}{2}\right) - \sin\left(\frac{\pi}{2}\right) \cdot \sin\left(\frac{\pi}{2}\right) \\
 &\stackrel{(2) \text{ and (4)}}{=} 0 \cdot 0 - 1 \cdot 1 \\
 &= -1
 \end{aligned}$$

6.

$$\begin{aligned}
 \sin(\pi) &= \sin\left(\frac{\pi}{2} + \frac{\pi}{2}\right) \\
 &\stackrel{18.45 \text{ (7)}}{=} \sin\left(\frac{\pi}{2}\right) \cdot \cos\left(\frac{\pi}{2}\right) + \cos\left(\frac{\pi}{2}\right) \cdot \sin\left(\frac{\pi}{2}\right) \\
 &\stackrel{(2) \text{ and (4)}}{=} 1 \cdot 0 - 0 \cdot 1 \\
 &= 0
 \end{aligned}$$

7.

$$\begin{aligned}
 \cos(2 \cdot \pi) &= \cos(\pi + \pi) \\
 &= \cos(\pi) \cdot \cos(\pi) - \sin(\pi) \cdot \sin(\pi) \\
 &\stackrel{(5) \text{ and (6)}}{=} (-1) \cdot (-1) - 0 \cdot 0 \\
 &= 1
 \end{aligned}$$

8.

$$\begin{aligned}
 \sin(2 \cdot \pi) &= \sin(\pi + \pi) \\
 &\stackrel{18.45 \text{ (7)}}{=} \sin(\pi) \cdot \cos(\pi) + \cos(\pi) \cdot \sin(\pi) \\
 &\stackrel{(5) \text{ and (6)}}{=} 0 \cdot (-1) + (-1) \cdot 0 \\
 &= 0
 \end{aligned}$$

9. Let  $x \in \mathbb{R}$  then

$$\begin{aligned}
 \cos(x + 2 \cdot \pi) &\stackrel{18.45 \text{ (7)}}{=} \cos(x) \cdot \cos(2 \cdot \pi) - \sin(x) \cdot \sin(2 \cdot \pi) \\
 \text{a.} \quad &\stackrel{(7) \text{ and (8)}}{=} \cos(x) \cdot 1 - \sin(x) \cdot 0 \\
 &= \cos(x) \\
 \sin(x + 2 \cdot \pi) &\stackrel{18.45 \text{ (7)}}{=} \sin(x) \cdot \cos(2 \cdot \pi) + \cos(x) \cdot \sin(2 \cdot \pi) \\
 \text{b.} \quad &\stackrel{(7) \text{ and (8)}}{=} \sin(x) \cdot 1 + \cos(x) \cdot 0 \\
 &= \sin(x)
 \end{aligned}$$

10. Let  $x \in \mathbb{R}$  then

$$\begin{aligned}
 \exp(i \cdot (x + 2 \cdot \pi)) &= \cos(x + 2 \cdot \pi) + i \cdot \sin(x + 2 \cdot \pi) \\
 &\stackrel{(9)}{=} \cos(x) + i \cdot \sin(x) \\
 &= \exp(i \cdot x)
 \end{aligned}$$

11.

$$\begin{aligned}\exp(i \cdot \pi) &= \cos(\pi) + i \cdot \sin(\pi) \\ &\stackrel{(5)(6)}{=} -1 + i \cdot 0 \\ &= -1 \\ &\quad \square\end{aligned}$$



# Chapter 19

## Measure Theory

### 19.1 Prerequisites

We have the following useful [and maybe trivial] set relations

**Proposition 19.1.** *Let  $A, B, X$  be sets with  $A, B \subseteq X$  then  $A \setminus B = (X \setminus B) \cap A$*

**Proof.**

$$\begin{aligned} x \in A \setminus B &\Leftrightarrow x \in A \wedge x \notin B \\ &\stackrel{A \subseteq X}{\Leftrightarrow} x \in A \wedge x \in X \wedge x \notin B \\ &\Leftrightarrow x \in A \wedge x \in X \setminus B \\ &\Leftrightarrow x \in (X \setminus B) \cap A \end{aligned}$$

□

**Proposition 19.2.** *Let  $A, B, C$  be sets then  $(A \setminus B) \cap C = (A \cap C) \setminus (B \cap C)$*

**Proof.** If  $x \in (A \setminus B) \cap C$  then  $x \in A \wedge x \notin B \wedge x \in C$  so that  $x \in (A \cap C)$  and  $x \notin B \cap C$  proving that

$$(A \setminus B) \cap C \subseteq (A \cap C) \setminus (B \cap C)$$

Further for the opposite inclusion

$$\begin{aligned} x \in (A \cap C) \setminus (B \cap C) &\Rightarrow x \in A \wedge x \in C \wedge x \notin (B \cap C) \\ &\Rightarrow (x \in A \wedge x \in C) \wedge \neg(x \in B \cap C) \\ &\Rightarrow (x \in A \wedge x \in C) \wedge (x \notin B \vee x \notin C) \\ &\Rightarrow (x \in A \wedge x \in C \notin B) \vee (x \in A \wedge x \in C \wedge x \notin C) \\ &\Rightarrow x \in A \wedge x \in C \notin B \\ &\Rightarrow x \in (A \setminus B) \cap C \end{aligned}$$

proving that

$$(A \cap C) \setminus (B \cap C) \subseteq (A \setminus B) \cap C$$

□

**Proposition 19.3.** *Let  $\{A_i\}_{i \in I}$  a family of sets and let  $I = K \bigcup L$  then*

1.  $\bigcup_{i \in I} A_i = (\bigcup_{i \in K} A_i) \bigcup (\bigcup_{i \in L} A_i)$
2.  $\bigcap_{i \in I} A_i = (\bigcap_{i \in K} A_i) \cap (\bigcap_{i \in L} A_i)$

**Proof.**

1. As

$$\begin{aligned} x \in \bigcup_{i \in I} A_i &\Leftrightarrow \exists i \in I \vdash x \in A_i \\ &\Leftrightarrow (\exists i \in K \vdash x \in A_i) \vee (\exists i \in L \vdash x \in A_i) \\ &\Leftrightarrow x \in \left( \bigcup_{i \in I} A_i \right) \bigcup \left( \bigcup_{i \in L} A_i \right) \end{aligned}$$

proving that  $\bigcup_{i \in I} A_i = (\bigcup_{i \in I} A_i) \bigcup (\bigcup_{i \in L} A_i)$ .

2.

$$\begin{aligned} x \in \bigcap_{i \in I} A_i &\Leftrightarrow \exists i \in I \vdash x \in A_i \\ &\Leftrightarrow (\exists i \in K \vdash x \in A_i) \wedge (\exists i \in L \vdash x \in A_i) \\ &\Leftrightarrow x \in \left( \bigcap_{i \in I} A_i \right) \bigcap \left( \bigcap_{i \in L} A_i \right) \end{aligned}$$

proving that  $\bigcap_{i \in I} A_i = (\bigcap_{i \in I} A_i) \bigcap (\bigcap_{i \in L} A_i)$ .  $\square$

**Proposition 19.4.** Let  $\{A_i\}_{i \in I}$  be a family of sets and  $f: J \rightarrow I$  a surjection then

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

**Proof.**

1. If  $x \in \bigcup_{i \in I} A_i$  then  $\exists i \in I$  such that  $x \in A_i$ . And as  $f$  is a surjection there exists a  $j \in J$  such that  $i = f(j)$  hence  $x \in A_{f(j)} \subseteq \bigcup_{j \in J} A_{f(j)}$  proving that

$$\bigcup_{i \in I} A_i \subseteq \bigcup_{j \in J} A_{f(j)}$$

Further if  $x \in \bigcup_{j \in J} A_{f(j)}$  then there exists a  $j \in J$  such that  $x \in A_{f(j)}$ , as  $f(j) \in I$  it follows that  $x \in \bigcup_{i \in I} A_i$  proving that

$$\bigcup_{j \in J} A_{f(j)} \subseteq \bigcup_{i \in I} A_i$$

2. If  $x \in \bigcap_{i \in I} A_i$  then  $\forall i \in I$  we have  $x \in A_i$  then if  $j \in J$   $f(j) \in I$  so that  $x \in A_{f(j)}$  proving that  $x \in \bigcap_{j \in J} A_{f(j)}$ , hence

$$\bigcap_{i \in I} A_i \subseteq \bigcap_{j \in J} A_{f(j)}$$

If  $x \in \bigcap_{j \in J} A_{f(j)}$  then  $\forall j \in J$  we have that  $x \in A_{f(j)}$ , then if  $i \in I$  there exists a  $k \in J$  such that  $i = f(k)$  so that  $x \in A_{f(k)} = A_i$  proving that  $x \in \bigcap_{i \in I} A_i$  hence

$$\bigcap_{j \in J} A_{f(j)} \subseteq \bigcap_{i \in I} A_i$$

$\square$

**Proposition 19.5.** Let  $\{A_i\}_{i \in \mathbb{N}_0}$  be a sequence of sets then if we define  $\{B_i\}_{i \in \mathbb{N}_0}$  by  $B_i = \bigcup_{j \in \{1, \dots, i\}} A_j$  then  $\bigcup_{i \in \mathbb{N}_0} A_i = \bigcup_{i \in \mathbb{N}_0} B_i$  and  $\forall i \in \mathbb{N}_0$  we have  $B_i \subseteq B_{i+1}$

**Proof.** First if  $i \in \mathbb{N}_0$  and  $x \in B_i$  there exists a  $j \in \{1, \dots, i\}$  such that  $x \in A_j \subseteq \bigcup_{j \in \{1, \dots, i+1\}} A_j = B_{i+1}$  proving that

$$\forall i \in \mathbb{N}_0 B_i \subseteq B_{i+1} \tag{19.1}$$

Second if  $x \in \bigcup_{i \in \mathbb{N}_0} A_i$  then  $\exists i \in \mathbb{N}_0$  such that  $x \in A_i$ . So as  $i \in \{1, \dots, i\}$  we have that  $x \in \bigcup_{j \in \{1, \dots, i\}} A_i = B_i$  which proves that  $x \in \bigcup_{i \in \mathbb{N}_0} B_i$ , or

$$\bigcup_{i \in \mathbb{N}_0} A_i \subseteq \bigcup_{i \in \mathbb{N}_0} B_i \quad (19.2)$$

Third if  $x \in \bigcup_{i \in \mathbb{N}_0} B_i$  then  $\exists i \in \mathbb{N}_0$  such that  $x \in B_i = \bigcup_{j \in \{1, \dots, i\}} A_j$ , so there exists a  $j \in \{1, \dots, i\} \subseteq \mathbb{N}_0$  with  $x \in A_j$ , proving that  $x \in \bigcup_{i \in \mathbb{N}_0} A_i$ . Hence

$$\bigcup_{i \in \mathbb{N}_0} B_i \subseteq \bigcup_{i \in \mathbb{N}_0} A_i \quad (19.3)$$

The theorem is then proved by 19.1, 19.2 and 19.3.  $\square$

**Theorem 19.6.** Let  $\{A_i\}_{i \in \mathbb{N}_0}$  be a countable family of sets then if we define  $\{B_i\}_{i \in \mathbb{N}_0}$  by  $B_i = \begin{cases} A_1 & \text{if } i = 1 \\ A_i \setminus \bigcup_{k \in \{1, \dots, i-1\}} A_k & \text{if } i \in \mathbb{N}_0 \setminus \{1\} \end{cases}$  then

1.  $\forall i \in \mathbb{N}_0$  we have  $B_i \subseteq A_i$
2.  $\forall i, j \in \mathbb{N}_0$  with  $i \neq j$  we have  $B_i \cap B_j = \emptyset$
3.  $\bigcup_{i \in \mathbb{N}_0} A_i = \bigcup_{i \in \mathbb{N}_0} B_i$

in other words we have

$$\bigcup_{i \in \mathbb{N}_0} A_i = \bigsqcup_{i \in \mathbb{N}_0} B_i$$

**Proof.** Let  $i, j \in \mathbb{N}_0$  with  $i \neq j$  then we may always assume that  $i < j$  [otherwise exchange  $i$  with  $j$ ]. If now  $x \in B_i \cap B_j$  then we have the following cases to consider:

**$i = 1$ .** then  $x \in B_1 = A_1$  and  $x \in B_j = A_j \setminus \bigcup_{k \in \{1, \dots, j-1\}} A_k$  so that  $\forall k \in \{1, \dots, j-1\}$  we have  $x \notin A_k$  giving as  $1 < j \Rightarrow i \leq j-1$  that  $x \notin A_1$  a contradiction.

**$1 < i$ .** then  $x \in B_i = A_i \setminus \bigcup_{k \in \{1, \dots, i-1\}} A_k$  so that  $x \in A_i$ , as  $i < j$  we have  $i \leq j-1$  hence  $x \in \bigcup_{k \in \{1, \dots, j-1\}} A_k$  contradicting  $x \in A_j \setminus \bigcup_{k \in \{1, \dots, j-1\}} A_k$ .

Hence

$$\forall i, j \in \mathbb{N}_0 \text{ with } i \neq j \text{ we have } B_i \cap B_j = \emptyset \quad (19.4)$$

As  $\forall i \in \mathbb{N}_0$  we have  $B_i = \begin{cases} A_1 & \text{if } i = 1 \\ A_i \setminus \bigcup_{k \in \{1, \dots, i-1\}} A_k & \text{if } i \in \mathbb{N}_0 \setminus \{1\} \end{cases} \subseteq \begin{cases} A_1 & \text{if } i = 1 \\ A_i & \text{if } i \in \mathbb{N}_0 \setminus \{1\} \end{cases} \subseteq A_i$  we have that

$$\forall i \in \mathbb{N}_0 \text{ we have } B_i \subseteq A_i \quad (19.5)$$

and

$$\bigcup_{i \in \mathbb{N}_0} B_i \subseteq \bigcup_{i \in \mathbb{N}_0} A_i \quad (19.6)$$

If  $x \in \bigcup_{i \in \mathbb{N}_0} A_i$  then  $\exists i \in \mathbb{N}_0$  such that  $x \in A_i$ , hence  $i \in \{j \in \{1, \dots, i\} | x \in A_j\}$ , so that  $m = \min(\{j \in \{1, \dots, i\} | x \in A_j\})$  exists. We must consider the following cases for  $m$ :

**$m = 1$ .** then  $x \in A_1 = B_1 \subseteq \bigcup_{i \in \mathbb{N}_0} B_i$

**$1 < m$ .** then  $1 \leq m-1$  and  $\forall j \in \{1, \dots, m-1\}$  we have  $x \notin A_j$  hence  $x \in A_m \setminus \bigcup_{k \in \{1, \dots, m-1\}} A_k = B_m \subseteq \bigcup_{i \in \mathbb{N}_0} B_i$

proving that  $\bigcup_{i \in \mathbb{N}_0} A_i \subseteq \bigcup_{i \in \mathbb{N}_0} B_i$ . Combining this with 19.6 proves

$$\bigcup_{i \in \mathbb{N}_0} B_i = \bigcup_{i \in \mathbb{N}_0} A_i \quad (19.7)$$

The theorem is then proved by 19.4, 19.5 and 19.7.  $\square$

For a finite family of sets we have the following

**Proposition 19.7.** *Let  $X$  be a set and  $\mathcal{A} \subseteq \mathcal{P}(X)$  then we have*

1. *If  $\forall A, B \in \mathcal{A}$  we have  $A \cap B \in \mathcal{A}$  then if  $\{A_i\}_{i \in I} \subseteq \mathcal{A}$  with  $I$  finite and non empty it follows that  $\bigcap_{i \in I} A_i \in \mathcal{A}$ .*
2. *If  $\forall A, B \in \mathcal{A}$  we have  $A \cup B \in \mathcal{A}$  then if  $\{A_i\}_{i \in I} \subseteq \mathcal{A}$  with  $I$  finite and not empty it follows that  $\bigcup_{i \in I} A_i \in \mathcal{A}$ .*

**Proof.**

1. As  $I$  is finite and non empty there exists a  $n \in \mathbb{N}_0$  and a bijection  $\varphi: \{1, \dots, n\} \rightarrow I$  then using 19.4 we have

$$\bigcap_{i \in I} A_i = \bigcap_{i \in \{1, \dots, n\}} A_{\varphi(i)} \quad (19.8)$$

The rest is proved by induction so let  $\mathcal{S} = \{i \in \mathbb{N}_0 \mid \text{if } i \leq n \text{ then } \bigcap_{j \in \{1, \dots, i\}} A_{\varphi(j)} \in \mathcal{A}\}$  then

**1  $\in \mathcal{S}$ .** As  $\bigcap_{i \in \{1, \dots, n\}} A_i = A_1 \in \mathcal{A}$  we have that  $i \in \mathcal{S}$

**$i \in \mathcal{S} \Rightarrow i+1 \in \mathcal{S}$ .** If  $i+1 \leq n \Rightarrow i < n$  then we have  $A_{i+1}, \bigcap_{j \in \{1, \dots, i\}} A_j \in \mathcal{S}$  so that  $\bigcup_{j \in \{1, \dots, i+1\}} A_j = (\bigcap_{j \in \{1, \dots, n\}} A_j) \cup A_{i+1} \in \mathcal{A}$  proving that  $i+1 \in \mathcal{A}$ .

Mathematical induction and the fact that  $n \leq n$  proves then that  $\bigcap_{i \in \{1, \dots, n\}} A_i \in \mathcal{A}$ , combining this with 19.8 proves (1),

2. As  $I$  is finite and non empty there exists a  $n \in \mathbb{N}_0$  and a bijection  $\varphi: \{1, \dots, n\} \rightarrow I$  then using 19.4 we have

$$\bigcup_{i \in \{1, \dots, n\}} A_i = \bigcup_{i \in I} A_i \quad (19.9)$$

The rest is proved by induction so let  $\mathcal{S} = \{i \in \mathbb{N}_0 \mid \text{if } i \leq n \text{ then } \bigcup_{j \in \{1, \dots, i\}} A_{\varphi(j)} \in \mathcal{A}\}$  then

**1  $\in \mathcal{S}$ .** As  $\bigcup_{i \in \{1, \dots, n\}} A_i = A_1 \in \mathcal{A}$  we have that  $i \in \mathcal{S}$

**$i \in \mathcal{S} \Rightarrow i+1 \in \mathcal{S}$ .** If  $i+1 \leq n \Rightarrow i < n$  then we have  $A_{i+1}, \bigcup_{j \in \{1, \dots, i\}} A_j \in \mathcal{S}$  so that  $\bigcup_{j \in \{1, \dots, i+1\}} A_j = (\bigcup_{j \in \{1, \dots, n\}} A_j) \cup A_{i+1} \in \mathcal{A}$  proving that  $i+1 \in \mathcal{A}$ .

Mathematical induction and the fact that  $n \leq n$  proves then that  $\bigcup_{i \in \{1, \dots, n\}} A_i \in \mathcal{A}$ , combining this with 19.9 proves (2),  $\square$

## 19.2 Algebras and $\sigma$ -algebras

**Definition 19.8. (Algebra)** *Let  $X$  be a set and  $\mathcal{A} \subseteq \mathcal{P}(X)$  a collection of subsets of  $X$  then  $\mathcal{A}$  is a algebra on  $X$  if*

1.  $X \in \mathcal{A}$
2.  $\forall A \in \mathcal{A}$  we have that  $X \setminus A \in \mathcal{A}$
3.  $\forall A, B \in \mathcal{A}$  we have that  $A \cup B \in \mathcal{A}$

Note that we have the following equivalent definition of a algebra

**Proposition 19.9.** *Let  $X$  be a set and  $\mathcal{A} \subseteq \mathcal{P}(X)$  then we have that the following are equivalent*

1.  $\mathcal{A}$  is a algebra

2.  $\mathcal{A}$  satisfies

- a.  $\emptyset \in \mathcal{A}$
- b.  $\forall A \in \mathcal{A}$  we have  $X \setminus A \in \mathcal{A}$
- c.  $\forall A, B \in \mathcal{A}$  we have  $A \cup B \in \mathcal{A}$

3.  $\mathcal{A}$  satisfies

- a.  $\emptyset \in \mathcal{A}$
- b.  $\forall A \in \mathcal{A}$  we have  $X \setminus A \in \mathcal{A}$
- c.  $\forall A, B \in \mathcal{A}$  we have  $A \cap B \in \mathcal{A}$

4.  $\mathcal{A}$  satisfies

- a.  $X \in \mathcal{A}$
- b.  $\forall A \in \mathcal{A}$  we have  $X \setminus A \in \mathcal{A}$
- c.  $\forall A, B \in \mathcal{A}$  we have  $A \cup B \in \mathcal{A}$

**Proof.**

1  $\Rightarrow$  2. Assume that  $\mathcal{A}$  is a algebra then

- a)  $X \in \mathcal{A} \xrightarrow{\mathcal{A} \text{ is a algebra}} \emptyset = X \setminus X \in \mathcal{A}$  proving that  $\emptyset \in \mathcal{A}$
- b)  $\forall A \in \mathcal{A}$  we have as  $\mathcal{A}$  is a algebra that  $X \setminus A \in \mathcal{A}$
- c)  $\forall A, B \in \mathcal{A}$  we have as  $\mathcal{A}$  is a algebra that  $A \cup B \in \mathcal{A}$

2  $\Rightarrow$  3. Assume that (2) is valid then

- a) From (2.a) it follows that  $\emptyset \in \mathcal{A}$
- b) Let  $A \in \mathcal{A}$  then from (2.b) we have that  $X \setminus A \in \mathcal{A}$
- c) Let  $A, B \in \mathcal{A}$  then by (2.b) we have that  $X \setminus A, X \setminus B \in \mathcal{A}$  hence we have

$$X \setminus (A \cap B) \stackrel{1.28}{=} (X \setminus A) \cup (X \setminus B) \in \mathcal{A} \text{ [because (2.c)]}$$

Hence

$$A \cap B \stackrel{1.19}{=} (X \setminus (X \setminus (A \cap B))) \in \mathcal{A} \text{ [because (2.b)]}$$

3  $\Rightarrow$  4. Assume that (3) is valid then

- a) From (3.a) it follows that  $\emptyset \in \mathcal{A}$
- b) Let  $A \in \mathcal{A}$  then from (2.b) we have that  $X \setminus A \in \mathcal{A}$
- c) Let  $A, B \in \mathcal{A}$  then by (3.b) we have that  $X \setminus A, X \setminus B \in \mathcal{A}$  hence we have

$$X \setminus (A \cup B) \stackrel{1.28}{=} (X \setminus A) \cap (X \setminus B) \in \mathcal{A} \text{ because (3.c)]}$$

hence

$$A \cup B \stackrel{1.19}{=} (X \setminus (X \setminus (A \cup B))) \text{ [because (3.b)]}$$

4  $\Rightarrow$  1. Assume that (4) is valid then

- a) Then as  $\emptyset \in \mathcal{A}$  we have by (4.b) that  $X = X \setminus \emptyset \in \mathcal{A}$
- b) Let  $A \in \mathcal{A}$  then from (4.b) we have that  $X \setminus A \in \mathcal{A}$
- c) Let  $A, B \in \mathcal{A}$  then from (4.c) we have  $A \cup B \in \mathcal{A}$

which proves that  $\mathcal{A}$  is a algebra. □

Using induction we can prove that a algebra is stable under finite unions and intersections.

**Proposition 19.10.** *Let  $X$  be a set and  $\mathcal{A}$  a algebra on  $X$  then we have*

1. *If  $I$  is a finite set and  $\{A_i\}_{i \in I} \subseteq \mathcal{A}$  then  $\bigcup_{i \in I} A_i \in \mathcal{A}$*
2. *If  $I$  is a finite non empty set and  $\{A_i\}_{i \in I} \subseteq \mathcal{A}$  then  $\bigcap_{i \in I} A_i \in \mathcal{A}$*
3. *If  $A, B \in \mathcal{A}$  then  $A \setminus B \in \mathcal{A}$*

**Proof.**

1. If  $I$  is finite then either  $I = \emptyset$  in which case we have that  $\bigcup_{i \in I} A_i = \emptyset \in \mathcal{A}$  or using 19.7 we have that  $\bigcup_{i \in I} A_i$ .
2. This follows from 19.7.
3. If  $A, B \in \mathcal{A}$  then  $X \setminus B \in \mathcal{A}$  so that  $A \cap (X \setminus B) \in \mathcal{A}$ , further  $A \cap (X \setminus B) \stackrel{1.28}{=} (A \cap X) \setminus B = A \setminus B$  proving that  $A \setminus B \in \mathcal{A}$   $\square$

**Definition 19.11. ( $\sigma$ -algebra)** *Let  $X$  be a set and  $\mathcal{A} \subseteq \mathcal{P}(X)$  then  $\mathcal{A}$  is a sigma algebra [noted as  $\sigma$ -algebra] on  $X$  if*

1.  $X \in \mathcal{A}$
2.  $\forall A \in \mathcal{A}$  we have  $X \setminus A \in \mathcal{A}$
3.  $\forall \{A_i\}_{i \in \mathbb{N}_0}$  we have  $\bigcup_{i \in \mathbb{N}_0} A_i \in \mathcal{A}$

**Example 19.12.** Let  $X$  be a set then  $\mathcal{P}(X)$  is a  $\sigma$ -algebra on  $X$

**Proof.** This trivial but lets write this out

1. As  $X \subseteq X$  we have  $X \in \mathcal{P}(X)$
2. If  $A \in \mathcal{P}(X)$  then  $X \setminus A \subseteq X$  proving that  $X \setminus A \in \mathcal{P}(X)$
3. If  $\{A_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{P}(X)$  then  $\forall i \in \mathbb{N}_0$  we have that  $A_i \subseteq X$  hence  $\bigcup_{i \in \mathbb{N}_0} A_i \subseteq X$  proving that  $\bigcup_{i \in \mathbb{N}_0} A_i \in \mathcal{P}(X)$   $\square$

**Definition 19.13. (Measurable Space)** *Let  $X$  be a non empty set set and  $\mathcal{A}$  a  $\sigma$ -algebra then we call  $\langle X, \mathcal{A} \rangle$  a measurable space.*

Just like for algebras we have the following equivalent definitions.

**Proposition 19.14.** *Let  $X$  be a set and  $\mathcal{A} \subseteq \mathcal{P}(X)$  then we have that the following are equivalent*

1.  $\mathcal{A}$  is a  $\sigma$ -algebra
2.  $\mathcal{A}$  satisfies
  - a.  $\emptyset \in \mathcal{A}$
  - b.  $\forall A \in \mathcal{A}$  we have  $X \setminus A \in \mathcal{A}$
  - c.  $\forall \{A_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{A}$  we have  $\bigcup_{i \in \mathbb{N}_0} A_i \in \mathcal{A}$
3.  $\mathcal{A}$  satisfies
  - a.  $\emptyset \in \mathcal{A}$
  - b.  $\forall A \in \mathcal{A}$  we have  $X \setminus A \in \mathcal{A}$
  - c.  $\forall \{A_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{A}$  we have  $\bigcap_{i \in \mathbb{N}_0} A_i \in \mathcal{A}$
4.  $\mathcal{A}$  satisfies
  - a.  $X \in \mathcal{A}$
  - b.  $\forall A \in \mathcal{A}$  we have  $X \setminus A \in \mathcal{A}$
  - c.  $\forall \{A_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{A}$  we have  $\bigcup_{i \in \mathbb{N}_0} A_i \in \mathcal{A}$

5.  $\mathcal{A}$  satisfies

- a.  $\emptyset \in \mathcal{A}$
- b.  $\forall A \in \mathcal{A}$  we have  $X \setminus A \in \mathcal{A}$
- c.  $\forall \{A_i\}_{i \in I} \subseteq \mathcal{A}$  with  $I$  countable we have  $\bigcup_{i \in I} A_i \in \mathcal{A}$

6.  $\mathcal{A}$  satisfies

- a.  $\emptyset \in \mathcal{A}$
- b.  $\forall A \in \mathcal{A}$  we have  $X \setminus A \in \mathcal{A}$
- c.  $\forall \{A_i\}_{i \in I} \subseteq \mathcal{A}$  with  $I$  countable we have  $\bigcap_{i \in I} A_i \in \mathcal{A}$

**Proof.**

**1  $\Rightarrow$  2.** Assume that  $\mathcal{A}$  is a algebra then

- a)  $X \in \mathcal{A} \xrightarrow{\mathcal{A} \text{ is a algebra}} \emptyset = X \setminus X \in \mathcal{A}$  proving that  $\emptyset \in \mathcal{A}$
- b)  $\forall A \in \mathcal{A}$  we have as  $\mathcal{A}$  is a algebra that  $X \setminus A \in \mathcal{A}$
- c)  $\forall \{A_i\}_{i \in \mathbb{N}_0} \in \mathcal{A}$  we have as  $\mathcal{A}$  is a algebra that  $\bigcup_{i \in \mathbb{N}_0} A_i \in \mathcal{A}$

**2  $\Rightarrow$  3.** Assume that (2) is valid then

- a) From (2.a) it follows that  $\emptyset \in \mathcal{A}$
- b) Let  $A \in \mathcal{A}$  then from (2.b) we have that  $X \setminus A \in \mathcal{A}$
- c) Let  $\{A_i\}_{i \in \mathbb{N}_0} \in \mathcal{A}$  then by (2.b) we have that  $\forall i \in \mathbb{N}_0 X \setminus A_i \in \mathcal{A}$ , hence we have

$$X \setminus \left( \bigcap_{i \in \mathbb{N}_0} A_i \right) \stackrel{1.102}{=} \bigcup_{i \in \mathbb{N}_0} (X \setminus A_i) \in \mathcal{A} \quad [\text{because (2.c)}]$$

Hence

$$\bigcap_{i \in \mathbb{N}_0} A_i \stackrel{1.19}{=} \left( X \setminus \left( X \setminus \left( \bigcap_{i \in \mathbb{N}_0} A_i \right) \right) \right) \in \mathcal{A} \quad [\text{because (2.b)}]$$

**3  $\Rightarrow$  4.** Assume that (3) is valid then

- a) From (3.a) it follows that  $\emptyset \in \mathcal{A}$  hence  $X = X \setminus \emptyset \in \mathcal{A}$  by (3.b)
- b) Let  $A \in \mathcal{A}$  then from (3.b) we have that  $X \setminus A \in \mathcal{A}$
- c) Let  $\{A_i\}_{i \in \mathbb{N}_0} \in \mathcal{A}$  then by (3.b) we have that  $\forall i \in \mathbb{N}_0 X \setminus A_i \in \mathcal{A}$  hence we have

$$X \setminus \left( \bigcup_{i \in \mathbb{N}_0} A_i \right) = \bigcap_{i \in \mathbb{N}_0} (X \setminus A_i) \stackrel{1.28}{\in} \mathcal{A} \quad [\text{because (3.c)}]$$

hence

$$\bigcup_{i \in \mathbb{N}_0} (X \setminus A_i) \stackrel{1.19}{=} \left( X \setminus \left( X \setminus \left( \bigcup_{i \in \mathbb{N}_0} (X \setminus A_i) \right) \right) \right) \in \mathcal{A} \quad [\text{because (3.b)}]$$

**4  $\Rightarrow$  1.** Assume that (4) is valid then

- a) As  $\emptyset \in \mathcal{A}$  we have by (4.b) that  $X = X \setminus \emptyset \in \mathcal{A}$
- b) Let  $A \in \mathcal{A}$  then from (4.b) we have that  $X \setminus A \in \mathcal{A}$
- c) Let  $\{A_i\}_{i \in \mathbb{N}_0} \in \mathcal{A}$  then from (4.c) we have  $\bigcup_{i \in \mathbb{N}_0} A_i \in \mathcal{A}$

which proves that  $\mathcal{A}$  is a  $\sigma$ -algebra.

**1  $\Rightarrow$  5.** Assume that (1) is valid then

- a) By (1.a) we have that  $\emptyset \in \mathcal{A}$
- b) By (1.b) we have that  $\forall A \in \mathcal{A} X \setminus A \in \mathcal{A}$

c) Let  $\{A_i\}_{i \in I} \subseteq \mathcal{A}$  with  $I$  countable then we have either

$$I = \emptyset. \text{ then } \bigcup_{i \in I} A_i = \emptyset \in \mathcal{A}$$

**$I$  is finite.** then there exists a  $n \in \mathbb{N}_0$  and a bijection  $\beta: \{1, \dots, n\} \rightarrow I$ . Define then  $\{B_i\}_{i \in \mathbb{N}_0}$  with  $B_i = \begin{cases} A_{\beta(i)} & \text{if } i \in \{1, \dots, n\} \\ \emptyset & \text{if } i \in \mathbb{N}_0 \setminus \{1, \dots, n\} \end{cases} \in \mathcal{A}$  then  $\{B_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{A}$  so that by

(1.c)  $\bigcup_{i \in \mathbb{N}_0} B_i \in \mathcal{A}$ . Further if  $x \in \bigcup_{i \in I} A_i$  then  $\exists i \in I$  such that  $x \in A_i$ , as  $\beta$  is a bijection there exists a  $j \in \{1, \dots, n\}$  with  $i = \beta(j)$  so that  $x \in A_i = A_{\beta(j)} = B_j \subseteq \bigcup_{i \in \{1, \dots, n\}} B_i$ , proving  $\bigcup_{i \in I} A_i \subseteq \bigcup_{i \in \mathbb{N}_0} B_i$ . Also as  $\forall i \in \mathbb{N}_0$  we have  $B_i = \begin{cases} A_{\beta(i)} & \text{if } i \in \{1, \dots, n\} \\ \emptyset & \text{if } i \in \mathbb{N}_0 \setminus \{1, \dots, n\} \end{cases} \subseteq \bigcup_{j \in I} A_j$  it follows that  $\bigcup_{i \in \mathbb{N}_0} B_i \subseteq \bigcup_{i \in I} A_i$  so that  $\bigcup_{i \in I} A_i = \bigcup_{i \in \mathbb{N}_0} B_i \in \mathcal{A}$ .

**$I$  is denumerable.** then there exists a bijection  $\beta: \mathbb{N}_0 \rightarrow I$ . Define then  $\forall i \in \mathbb{N}_0 B_i = A_{\beta(i)} \in \mathcal{A}$  so that by (1.c)  $\bigcup_{i \in \mathbb{N}_0} B_i \in \mathcal{A}$ . Further if  $x \in \bigcup_{i \in I} A_i$  then  $\exists i \in I$  such that  $x \in A_i$ , as  $\beta$  is a bijection there exists a  $j \in \{1, \dots, n\}$  with  $i = \beta(j)$  so that  $x \in A_i = A_{\beta(j)} = B_j \subseteq \bigcup_{i \in \{1, \dots, n\}} B_j$ , proving  $\bigcup_{i \in I} A_i \subseteq \bigcup_{i \in \mathbb{N}_0} B_i$ . Also as  $\forall i \in \mathbb{N}_0$  we have  $B_i = A_{\beta(i)} \subseteq \bigcup_{j \in I} A_j$  so that  $\bigcup_{i \in I} A_i = \bigcup_{j \in \mathbb{N}_0} B_j \in \mathcal{A}$ .

**5  $\Rightarrow$  1.**

- a) By (5.a) we have that  $\emptyset \in \mathcal{A}$
- b) By (5.b) we have that  $\forall A \in \mathcal{A} X \setminus A \in \mathcal{A}$
- c) If  $\{A_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{A}$  then as  $\mathbb{N}_0$  is countable we have by (5.c) that  $\bigcup_{i \in I} A_i \in \mathcal{A}$ .

**5  $\Rightarrow$  6.**

- a) By (5.a) we have that  $\emptyset \in \mathcal{A}$
- b) By (5.b) we have that  $\forall A \in \mathcal{A} X \setminus A \in \mathcal{A}$
- c) If  $\{A_i\}_{i \in I} \subseteq \mathcal{A}$  then  $\forall i \in I$  we have  $X \setminus A_i \in \mathcal{A}$  so that by (5.c)  $\bigcup_{i \in I} (X \setminus A_i) \in \mathcal{A}$ . Further as  $\bigcap_{i \in I} A_i = X \setminus (X \setminus \bigcap_{i \in I} A_i) \stackrel{1.19}{=} X \setminus (\bigcup_{i \in I} (X \setminus A_i)) \in \mathcal{A}$  which by (5.b) proves that  $\bigcap_{i \in I} A_i \in \mathcal{A}$ .

**6  $\Rightarrow$  5.**

- a) By (6.a) we have that  $\emptyset \in \mathcal{A}$
- b) By (6.b) we have that  $\forall A \in \mathcal{A} X \setminus A \in \mathcal{A}$
- c) If  $\{A_i\}_{i \in I} \subseteq \mathcal{A}$  then  $\forall i \in I$  we have  $X \setminus A_i \in \mathcal{A}$  so that by (6.b)  $\bigcap_{i \in I} (X \setminus A_i) \in \mathcal{A}$ . Further as  $\bigcup_{i \in I} A_i = X \setminus (X \setminus \bigcup_{i \in I} A_i) \stackrel{1.19}{=} X \setminus (\bigcap_{i \in I} (X \setminus A_i)) \in \mathcal{A}$  which by (5.b) proves that  $\bigcup_{i \in I} A_i \in \mathcal{A}$ .  $\square$

Note that a  $\sigma$ -algebra is a special case of a algebra as expressed in the following proposition.

**Proposition 19.15.** Let  $X$  be a set and  $\mathcal{A} \subseteq \mathcal{P}(X)$  a  $\sigma$ -algebra then  $\mathcal{A}$  is a algebra.

**Proof.** Let  $A, B \in \mathcal{A}$  define then  $\{A_i\}_{i \in \mathbb{N}_0}$  by  $A_i = \begin{cases} A & \text{if } i = 1 \\ B & \text{if } i = 2 \\ \emptyset & \text{if } i \in \mathbb{N}_0 \setminus \{1, 2\} \end{cases}$  then

$$\begin{aligned} A \bigcup B &= A_1 \bigcup A_2 \bigcup \emptyset \\ &= \left( \bigcup_{i \in \{1, 2\}} A_i \right) \bigcup \left( \bigcup_{i \in \mathbb{N}_0 \setminus \{1, 2\}} A \right) \\ &\stackrel{19.3}{=} \bigcup_{i \in \mathbb{N}_0} A_i \\ &\in \mathcal{A} \text{ [as } \mathcal{A} \text{ is a } \sigma\text{-algebra]} \end{aligned}$$

$\square$

We show now how given a set  $\mathcal{S}$  of subsets of  $X$  it is possible to generate the smallest  $\sigma$ -algebra on  $X$  containing  $\mathcal{S}$ . First we need a little lemma.

**Lemma 19.16.** *Let  $X$  be a set,  $\Sigma$  a non empty set of  $\sigma$ -algebras on  $X$  then  $\bigcap_{\mathcal{A} \in \Sigma} \mathcal{A}$  is a  $\sigma$ -algebra on  $X$*

**Proof.**

1. As  $\forall \mathcal{A} \in \Sigma \ \mathcal{A}$  is a  $\sigma$ -algebra on  $X$  we have  $X \in \mathcal{A}$  hence  $X \in \bigcap_{\mathcal{A} \in \Sigma} \mathcal{A}$
2. If  $C \in \bigcap_{\mathcal{A} \in \Sigma} \mathcal{A}$  then  $\forall \mathcal{A} \in \Sigma \ C \in \mathcal{A}$   $\mathcal{A}$  is a  $\sigma$ -algebra on  $X \setminus C \in \mathcal{A}$  proving that  $X \setminus C \in \bigcap_{\mathcal{A} \in \Sigma} \mathcal{A}$
3. If  $\{C_i\}_{i \in \mathbb{N}_0} \subseteq \bigcap_{\mathcal{A} \in \Sigma} \mathcal{A}$  then  $\forall \mathcal{A} \in \Sigma \ \{C_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{A}$   $\mathcal{A}$  is a  $\sigma$ -algebra  $\bigcup_{i \in \mathbb{N}_0} C_i \subseteq \mathcal{A}$  proving that  $\bigcup_{i \in \mathbb{N}_0} C_i \in \bigcap_{\mathcal{A} \in \Sigma} \mathcal{A}$   $\square$

**Theorem 19.17.** *Let  $X$  be a set and  $\mathcal{A} \subseteq \mathcal{P}(X)$  then there exists a **unique**  $\sigma$ -algebra  $\mathcal{X}$  on  $X$  with  $\mathcal{A} \subseteq \mathcal{X}$  such that if  $\mathcal{B}$  is another  $\sigma$ -algebra on  $X$  with  $\mathcal{A} \subseteq \mathcal{B}$  then  $\mathcal{X} \subseteq \mathcal{B}$ . We call this unique smallest  $\sigma$ -algebra containing  $\mathcal{A}$  the  $\sigma$ -algebra generated by  $\mathcal{A}$  and note it as  $\sigma[\mathcal{A}]$ .*

**Proof.** Define  $\Sigma$  as  $\Sigma = \{\mathcal{C} \subseteq \mathcal{P}(X) | \mathcal{C}$  is a  $\sigma$ -algebra and  $\mathcal{A} \subseteq \mathcal{C}\}$  then using the previous lemma [see 19.16] we have that

$$\mathcal{X} = \bigcap_{\mathcal{C} \in \Sigma} \mathcal{C} \text{ is a } \sigma\text{-algebra} \quad (19.10)$$

Further as  $\forall \mathcal{C} \in \Sigma$  we have by definition that  $\mathcal{A} \subseteq \mathcal{C}$  it follows that

$$\mathcal{A} \subseteq \bigcap_{\mathcal{C} \in \Sigma} \mathcal{C} = \mathcal{X} \quad (19.11)$$

Next if  $\mathcal{B}$  is a  $\sigma$ -algebra with  $\mathcal{A} \subseteq \mathcal{B}$  then by definition  $\mathcal{B} \in \Sigma$  hence

$$\mathcal{X} = \bigcap_{\mathcal{C} \in \Sigma} \mathcal{C} \subseteq \mathcal{B} \quad (19.12)$$

Finally if there is another  $\sigma$ -algebra  $\mathcal{R}$  such that  $\mathcal{A} \subseteq \mathcal{R}$  and for every  $\sigma$ -algebra  $\mathcal{B}$  with  $\mathcal{A} \subseteq \mathcal{B}$  we have  $\mathcal{R} \subseteq \mathcal{B}$  then by 19.10 and 19.11 we have that  $\mathcal{R} \subseteq \mathcal{X}$ , further using 19.12 we have that  $\mathcal{X} \subseteq \mathcal{R}$ . Hence we have that  $\mathcal{X} = \mathcal{R}$  proving the uniqueness of  $\mathcal{X}$ ,  $\square$

We introduce now a import set of  $\sigma$ -algebras the Borel algebras.

**Definition 19.18. (Borel algebra)** *Let  $\langle X, \mathcal{T} \rangle$  be a topological space then the  $\sigma$ -algebra generated by the topology  $\mathcal{T}$  is called a **Borel algebra** and noted by  $\mathcal{B}[X, \mathcal{T}]$  where  $\mathcal{B}[X, \mathcal{T}] = \sigma[\mathcal{T}]$ .*

**Example 19.19.** Let  $\langle \mathbb{R}, \mathcal{T}_{\mathbb{R}} \rangle$  be the topological space where  $\mathcal{T}_{\mathbb{R}}$  is the generated by the absolute value norm [see 11.72] then  $\mathcal{B}[\mathbb{R}, \mathcal{T}_{\mathbb{R}}]$  is noted as  $\mathcal{B}[\mathbb{R}]$ .

**Example 19.20.** Let  $n \in \mathbb{N}_0 \setminus \{1\}$  and  $\langle \mathbb{R}^n, \mathcal{T}_{\mathbb{R}^n} \rangle$  be the topological space where  $\mathcal{T}_{\mathbb{R}^n}$  is generated by the maximum norm  $\|x\| = \max(\{|x_i| | i \in \{1, \dots, n\}\})$  [see 11.83] then  $\mathcal{B}[\mathbb{R}^n, \mathcal{T}_{\mathbb{R}^n}]$  is noted as  $\mathcal{B}[\mathbb{R}^n]$ .

**Theorem 19.21.** *Let  $\langle X, \mathcal{T} \rangle$  be a topological space and let  $\mathcal{C} = \{C \subseteq X | C \text{ is closed}\}$  then  $\mathcal{B}[X, \mathcal{T}] = \sigma[\mathcal{C}]$*

**Proof.** First if  $C \in \mathcal{C}$  then there exists a  $U \in \mathcal{T}$  such that  $C = X \setminus U \in \sigma[\mathcal{T}] = \mathcal{B}[X, \mathcal{T}]$  proving that  $C \subseteq \mathcal{B}[X, \mathcal{T}]$  giving

$$\sigma[\mathcal{C}] \subseteq \mathcal{B}[X, \mathcal{T}]$$

On the other hand we have if  $U \in \mathcal{T}$  then  $X \setminus U \in \mathcal{C}$  so that  $U = X \setminus (X \setminus U) \in \sigma[\mathcal{C}]$  proving that  $\mathcal{T} \subseteq \sigma[\mathcal{C}]$  giving as  $\mathcal{B}[X, \mathcal{T}] = \sigma[\mathcal{T}]$  that

$$\mathcal{B}[X, \mathcal{T}] \subseteq \sigma[\mathcal{C}] \quad \square$$

### 19.2.1 Borel algebra on $\mathbb{R}$

First let's the concept of intervals in  $\mathbb{R}$

**Definition 19.22.** *We define the following set of bounded intervals [see 2.187 and 2.188]:*

1.  $\mathcal{I}_{[]} = \{[a, b] | a, b \in \mathbb{R} \wedge [a, b] \neq \emptyset\}$  (the set of non empty closed bounded intervals)
2.  $\mathcal{I}_{[[]} = \{[a, b[ | a, b \in \mathbb{R} \wedge [a, b[ \neq \emptyset\}$  (the set of non empty right open bounded intervals)
3.  $\mathcal{I}_{[]]} = \{]a, b] | a, b \in \mathbb{R} \wedge ]a, b] \neq \emptyset\}$  (the set of non empty left open bounded intervals)
4.  $\mathcal{I}_{[[]]} = \{]a, b[ | a, b \in \mathbb{R} \wedge ]a, b[ \neq \emptyset\}$  (the set of non empty open bounded intervals)
5.  $\mathcal{I} = \mathcal{I}_{[]} \cup \mathcal{I}_{[[]} \cup \mathcal{I}_{[]]} \cup \mathcal{I}_{[[]]}$  (the set of all non empty bounded intervals)

**Note 19.23.** If  $b < a$  then  $[a, b] = [a, b[ = ]a, b] = ]a, b[ = \emptyset$  and if  $a = b$  then  $[a, b[ = ]a, b] = ]a, b[ = \emptyset$ . These intervals are well defined but not part of  $\mathcal{I}_{[]} \cup \mathcal{I}_{[[]} \cup \mathcal{I}_{[]]}$ .

**Lemma 19.24.**  $\mathcal{I}_{[]}, \mathcal{I}_{[[]}, \mathcal{I}_{[]]}, \mathcal{I}_{[[]]}$  are pairwise disjoint

**Proof.** We have to check the following intersections to prove this (taking in account commutativity)

$\mathcal{I}_{[]} \cap \mathcal{I}_{[[]}$ . If  $I \in \mathcal{I}_{[]} \cap \mathcal{I}_{[[]}$  then  $I = [a, b] = [c, d[$ . As  $[a, b] \neq \emptyset$  we have by 2.184 that  $a \leq b$  so that  $b \in [a, b] = [c, d[$  hence  $c \leq b < d$ . Using 9.56 there exists a  $x \in \mathbb{R}$  such that  $c \leq b < x < d$  proving that  $x \in [c, d[ = [a, b] \Rightarrow x \leq b \Rightarrow b < b$  a contradiction. Hence  $\mathcal{I}_{[]} \cap \mathcal{I}_{[[]} = \emptyset$

$\mathcal{I}_{[]} \cap \mathcal{I}_{[]]}$ . If  $I \in \mathcal{I}_{[]} \cap \mathcal{I}_{[]]}$  then  $I = [a, b] = ]c, d]$ . As  $[a, b] \neq \emptyset$  we have by 2.184 that  $a \leq b$  so that  $a \in [a, b] = ]c, d]$  hence  $c < a \leq d$ . Using 9.56 there exists a  $x \in \mathbb{R}$  such that  $c < x < a \leq d$  proving that  $x \in ]c, d] = [a, b]$  hence  $a \leq x \Rightarrow a < a$  a contradiction. Hence  $\mathcal{I}_{[]} \cap \mathcal{I}_{[]]} = \emptyset$

$\mathcal{I}_{[]} \cap \mathcal{I}_{[[]}$ . If  $I \in \mathcal{I}_{[]} \cap \mathcal{I}_{[[]}$  then  $I = [a, b] = ]c, d[$ . As  $[a, b] \neq \emptyset$  we have by 2.184 that  $a \leq b$  so that  $b \in [a, b] = ]c, d[$  hence  $c < b < d$ . Using 9.56 there exists a  $x \in \mathbb{R}$  such that  $c < b < x < d$  so that  $x \in ]c, d[ = [a, b]$  hence  $x \leq b \Rightarrow b < b$  a contradiction. Hence  $\mathcal{I}_{[]} \cap \mathcal{I}_{[[]} = \emptyset$ .

$\mathcal{I}_{[[]} \cap \mathcal{I}_{[]]}$ . If  $I \in \mathcal{I}_{[[]} \cap \mathcal{I}_{[]]}$  then  $I = [a, b[ = ]c, d]$ . As  $[a, b[ \neq \emptyset$  we have by 2.184  $a < b$  so that  $a \in [a, b[ = ]c, d]$  hence  $c < a \leq d$ . Using 9.56 there exists a  $x \in \mathbb{R}$  such that  $c < x < a \leq d$  proving that  $x \in ]c, d] = [a, b[$  hence  $a \leq x \Rightarrow a < a$  a contradiction. Hence  $\mathcal{I}_{[[]} \cap \mathcal{I}_{[]]} = \emptyset$ .

$\mathcal{I}_{[[]} \cap \mathcal{I}_{[[]}$ . If  $I \in \mathcal{I}_{[[]} \cap \mathcal{I}_{[[]}$  then  $I = [a, b[ = ]c, d[$ . As  $[a, b[ \neq \emptyset$  we have by 2.184 that  $a < b$  so that  $a \in [a, b[ = ]c, d[$  hence  $c < a < d$ . Using 9.56 there exists a  $x \in \mathbb{R}$  such that  $c < x < a < d$  proving that  $x \in ]c, d[ = [a, b[$  hence  $a \leq x \Rightarrow a < a$  a contradiction. Hence  $\mathcal{I}_{[[]} \cap \mathcal{I}_{[[]} = \emptyset$ .  $\square$

Next we have some basic equalities for intervals

**Lemma 19.25.** *Let  $a, b \in \mathbb{R}$  the we have*

1.  $]a, b] = \bigcap_{n \in \mathbb{N}_0} ]a, b + \frac{1}{n}[$
2.  $[a, b[ = \bigcap_{n \in \mathbb{N}_0} ]a - \frac{1}{n}, b[$
3.  $[a, b] = \bigcap_{n \in \mathbb{N}_0} ]a - \frac{1}{n}, b]$
4.  $[a, b] = \bigcap_{n \in \mathbb{N}} [a, b + \frac{1}{n}[$
5.  $[a, b] = \bigcap_{n \in \mathbb{N}_0} ]a - \frac{1}{n}, b + \frac{1}{n}[$
6.  $]-\infty, a] = \bigcap_{n \in \mathbb{N}_0} ]-\infty, a + \frac{1}{n}[$
7.  $[a, \infty[ = \bigcap_{n \in \mathbb{N}_0} ]a - \frac{1}{n}, \infty[$

8.  $]a, b] = \bigcup_{n \in \mathbb{N}_0} [a + \frac{1}{n}, b]$
9.  $[a, b[ = \bigcup_{n \in \mathbb{N}_0} [a, b - \frac{1}{n}]$
10.  $]a, b[ = \bigcup_{n \in \mathbb{N}_0} ]a, b - \frac{1}{n}]$
11.  $]a, b[ = \bigcup_{n \in \mathbb{N}_0} [a + \frac{1}{n}, b[$
12.  $]a, b[ = \bigcup_{n \in \mathbb{N}_0} [a + \frac{1}{n}, b - \frac{1}{n}]$
13.  $]-\infty, a[ = \bigcup_{n \in \mathbb{N}_0} ]-\infty, a - \frac{1}{n}]$
14.  $]a, \infty[ = \bigcup_{n \in \mathbb{N}_0} [a + \frac{1}{n}, \infty[$
15.  $[a, b[ = [a, \infty[ \setminus [b, \infty[$
16.  $]a, b] = ]-\infty, b] \setminus ]-\infty, a]$

**Proof.**

1. Let  $x \in ]a, b]$  then  $a < x \leq b$  and as  $\forall n \in \mathbb{N}_0 b < b + \frac{1}{n}$  we have  $a < x < b + \frac{1}{n}$  so it follows that  $x \in ]a, b + \frac{1}{n}[$ . Hence

$$]a, b] \subseteq \bigcap_{n \in \mathbb{N}_0} ]a, b + \frac{1}{n}[ \quad (19.13)$$

If  $x \in \bigcap_{n \in \mathbb{N}_0} ]a, b + \frac{1}{n}[$  then  $\forall n \in \mathbb{N}_0 x \in ]a, b + \frac{1}{n}[$  or  $a < x < b + \frac{1}{n}$  which by 9.57 proves that  $a < x \leq b$  hence  $x \in ]a, b]$ . So it follows that  $\bigcap_{n \in \mathbb{N}_0} ]a, b + \frac{1}{n}[ \subseteq ]a, b]$  which by 19.13 proves that

$$]a, b] = \bigcap_{n \in \mathbb{N}_0} ]a, b + \frac{1}{n}[$$

2. Let  $x \in [a, b[$  then  $a \leq x < b$  and as  $\forall n \in \mathbb{N}_0$  we have  $a - \frac{1}{n} < a$  it follows that  $x \in ]a - \frac{1}{n}, b[$ . Hence

$$[a, b[ \subseteq \bigcap_{n \in \mathbb{N}_0} ]a - \frac{1}{n}, b[ \quad (19.14)$$

If  $x \in \bigcap_{n \in \mathbb{N}_0} ]a - \frac{1}{n}, b[$  then  $\forall n \in \mathbb{N}_0$  we have  $a - \frac{1}{n} < x < b \Rightarrow a < x + \frac{1}{n} \wedge x < b$  which by 9.57 proves that  $a \leq x < b$  hence  $x \in [a, b[$ . So  $\bigcap_{n \in \mathbb{N}_0} ]a - \frac{1}{n}, b[ \subseteq [a, b[$  which by 19.14

$$[a, b[ = \bigcap_{n \in \mathbb{N}_0} ]a - \frac{1}{n}, b[$$

3. Let  $x \in [a, b]$  then  $a \leq x < b$  and as  $\forall n \in \mathbb{N}_0$  we have  $a - \frac{1}{n} < a \Rightarrow x \in ]a - \frac{1}{n}, b]$ . Hence

$$[a, b] \subseteq \bigcap_{n \in \mathbb{N}_0} ]a - \frac{1}{n}, b] \quad (19.15)$$

If  $x \in \bigcap_{n \in \mathbb{N}_0} ]a - \frac{1}{n}, b]$  then  $\forall n \in \mathbb{N}_0$  then  $a - \frac{1}{n} < x \leq b$  or  $a < x + \frac{1}{n}$  which by 9.57 proves that  $a \leq x$  giving  $x \in [a, b]$ . So we have  $\bigcap_{n \in \mathbb{N}_0} ]a - \frac{1}{n}, b] \subseteq [a, b]$  which combined with 19.15 gives

$$[a, b] = \bigcap_{n \in \mathbb{N}_0} ]a - \frac{1}{n}, b]$$

4. Let  $x \in [a, b]$  then  $a \leq x \leq b$  then  $\forall n \in \mathbb{N}_0$  we have  $a \leq x \leq b < b + \frac{1}{n}$ . Which proves that

$$[a, b] \subseteq \bigcap_{n \in \mathbb{N}_0} [a, b + \frac{1}{n}[ \quad (19.16)$$

If  $x \in \bigcap_{n \in \mathbb{N}_0} [a, b + \frac{1}{n}]$  then  $\forall n \in \mathbb{N}_0$  we have  $a \leq x < b + \frac{1}{n}$ . Using by 9.57 we have  $x \leq b$  proving that  $x \in [a, b]$ . Hence  $\bigcap_{n \in \mathbb{N}_0} [a, b + \frac{1}{n}] \subseteq [a, b]$  which combined with 19.16 gives

$$[a, b] = \bigcap_{n \in \mathbb{N}_0} \left[ a, b + \frac{1}{n} \right]$$

5. Let  $x \in [a, b]$  then  $a \leq x \leq b$  and as  $\forall n \in \mathbb{N}_0$  we have  $a - \frac{1}{n} < x < b + \frac{1}{n}$  it follows that  $x \in \left] a - \frac{1}{n}, b + \frac{1}{n} \right[$ . Hence

$$[a, b] \subseteq \bigcap_{n \in \mathbb{N}_0} \left] a - \frac{1}{n}, b + \frac{1}{n} \right[ \quad (19.17)$$

Let  $x \in \bigcap_{n \in \mathbb{N}_0} \left] a - \frac{1}{n}, b + \frac{1}{n} \right[$  then  $\forall n \in \mathbb{N}_0$  we have  $a - \frac{1}{n} < x \wedge x < b + \frac{1}{n} \Rightarrow a < x + \frac{1}{n} \wedge x < b + \frac{1}{n}$ . Using 9.57 we have  $a \leq x \wedge x \leq b$  or  $x \in [a, b]$ . Hence  $\bigcap_{n \in \mathbb{N}_0} \left] a - \frac{1}{n}, b + \frac{1}{n} \right[ \subseteq [a, b]$  which together with 19.17 proves

$$[a, b] = \bigcap_{n \in \mathbb{N}_0} \left] a - \frac{1}{n}, b + \frac{1}{n} \right[$$

6. Let  $x \in ]-\infty, a]$  then  $x \leq a$  and as  $\forall n \in \mathbb{N}_0$  we have  $x < a + \frac{1}{n}$  it follows that  $x \in ]-\infty, a + \frac{1}{n}]$ . Hence

$$]-\infty, a] \subseteq \bigcap_{n \in \mathbb{N}_0} \left] -\infty, a + \frac{1}{n} \right[ \quad (19.18)$$

Let  $x \in \bigcap_{n \in \mathbb{N}_0} \left] -\infty, a + \frac{1}{n} \right[$  then  $\forall n \in \mathbb{N}_0$  we have  $x < a + \frac{1}{n}$  which by 9.57 gives  $x \leq a$  or  $x \in ]-\infty, a]$ . Hence  $\bigcap_{n \in \mathbb{N}_0} \left] -\infty, a + \frac{1}{n} \right[ \subseteq ]-\infty, a]$  which together with 19.18

$$\bigcap_{n \in \mathbb{N}_0} \left] -\infty, a + \frac{1}{n} \right[ = ]-\infty, a]$$

7. Let  $x \in [a, \infty[$  then  $a \leq x$  so that  $\forall n \in \mathbb{N}_0$  we have  $a - \frac{1}{n} < x$  giving  $x \in \left] a - \frac{1}{n}, \infty \right[$  proving that

$$[a, \infty[ \subseteq \bigcap_{n \in \mathbb{N}_0} \left] a - \frac{1}{n}, \infty \right[ \quad (19.19)$$

Let  $x \in \bigcap_{n \in \mathbb{N}_0} \left] a - \frac{1}{n}, \infty \right[$  then  $\forall n \in \mathbb{N}_0$  we have  $a - \frac{1}{n} < x \Rightarrow a < x + \frac{1}{n}$  which by 9.57 proves that  $x \in [a, \infty[$ . Hence  $\bigcap_{n \in \mathbb{N}_0} \left] a - \frac{1}{n}, \infty \right[ \subseteq [a, \infty]$  which together with 19.19 proves that

$$[a, \infty[ = \bigcap_{n \in \mathbb{N}_0} \left] a - \frac{1}{n}, \infty \right[$$

8. Let  $x \in ]a, b]$  then  $a < x \leq b$  and using 9.56 there exists a  $n \in \mathbb{N}_0$  such that  $a < a + \frac{1}{n} < x \leq b$  proving that  $x \in \left[ a + \frac{1}{n}, b \right]$ . Hence

$$]a, b] \subseteq \bigcup_{n \in \mathbb{N}_0} \left[ a + \frac{1}{n}, b \right] \quad (19.20)$$

Let  $x \in \bigcup_{n \in \mathbb{N}_0} \left[ a + \frac{1}{n}, b \right]$  there exists a  $n \in \mathbb{N}_0$  such that  $x \in \left[ a + \frac{1}{n}, b \right] \Rightarrow a < a + \frac{1}{n} \leq x \leq b$ . Hence  $x \in ]a, b]$  proving that  $\bigcup_{n \in \mathbb{N}_0} \left[ a + \frac{1}{n}, b \right] \subseteq ]a, b]$  which by 19.20 gives

$$]a, b] = \bigcup_{n \in \mathbb{N}_0} \left[ a + \frac{1}{n}, b \right]$$

9. Let  $x \in [a, b[$  then  $a \leq x < b$  and using 9.56 there exists a  $n \in \mathbb{N}_0$  such that  $a \leq x < x + \frac{1}{n} < b \Rightarrow a \leq x \wedge x < b - \frac{1}{n}$  proving that  $x \in \left[ a, b - \frac{1}{n} \right]$ . So

$$a, b[ \subseteq \bigcup_{n \in \mathbb{N}_0} \left[ a, b - \frac{1}{n} \right] \quad (19.21)$$

If  $x \in \bigcup_{n \in \mathbb{N}_0} [a, b - \frac{1}{n}]$  then  $\exists n \in \mathbb{N}_0$  such that  $x \in [a, b - \frac{1}{n}] \Rightarrow a \leq x \leq b - \frac{1}{n} < b$  proving that  $x \in [a, b]$ . Hence  $\bigcup_{n \in \mathbb{N}_0} [a, b - \frac{1}{n}] \subseteq [a, b]$  which combined with 19.21 gives

$$[a, b] = \bigcup_{n \in \mathbb{N}_0} \left[ a, b - \frac{1}{n} \right]$$

10. Let  $x \in ]a, b[$  then  $a < x < b$  and using 9.56 there exists a  $n \in \mathbb{N}_0$  such that  $a < x < x + \frac{1}{n} < b$  hence  $a < x < b - \frac{1}{n}$  proving that  $x \in ]a, b - \frac{1}{n}].$  So

$$]a, b[ \subseteq \bigcup_{n \in \mathbb{N}_0} \left[ a, b - \frac{1}{n} \right] \quad (19.22)$$

If  $x \in \bigcup_{n \in \mathbb{N}_0} ]a, b - \frac{1}{n}]$  then  $\exists n \in \mathbb{N}_0$  such that  $a < x \leq b - \frac{1}{n} < b \Rightarrow x \in ]a, b[$  proving that  $\bigcup_{n \in \mathbb{N}_0} ]a, b - \frac{1}{n}] \subseteq ]a, b[.$  Combining this with 19.22 gives

$$]a, b[ = \bigcup_{n \in \mathbb{N}_0} \left[ a, b - \frac{1}{n} \right]$$

11. Let  $x \in ]a, b[$  then  $a < x < b$  and using 9.56 there exists a  $n \in \mathbb{N}_0$  such that  $a < a + \frac{1}{n} < x < b$  proving that  $x \in [a + \frac{1}{n}, b[.$  Hence

$$]a, b[ \subseteq \bigcup_{n \in \mathbb{N}_0} \left[ a + \frac{1}{n}, b \right] \quad (19.23)$$

If  $x \in \bigcup_{n \in \mathbb{N}_0} [a + \frac{1}{n}, b[$  then  $\exists n \in \mathbb{N}_0$  such that  $a < a + \frac{1}{n} < x < b \Rightarrow x \in ]a, b[$  proving that  $\bigcup_{n \in \mathbb{N}_0} [a + \frac{1}{n}, b[ \subseteq ]a, b[.$  Combining this with 19.23 gives

$$]a, b[ = \bigcup_{n \in \mathbb{N}_0} \left[ a + \frac{1}{n}, b \right]$$

12. Let  $x \in ]a, b[$  then  $a < x < b$  and using 9.56 there exists  $n_1, n_2 \in \mathbb{N}_0$  such that  $a < a + \frac{1}{n_1} < x < x + \frac{1}{n_2} < b \Rightarrow a + \frac{1}{n_1} < x < b - \frac{1}{n_2}$ . Take  $m = \max(n_1, n_2)$  then  $\frac{1}{m} \leq \frac{1}{n_1}, \frac{1}{n_2}$  so that  $a + \frac{1}{m} \leq a + \frac{1}{n_1} < x < b - \frac{1}{m} \leq b - \frac{1}{n_2}$  proving that  $x \in [a + \frac{1}{m}, b - \frac{1}{m}]$ . Hence

$$]a, b[ \subseteq \bigcup_{n \in \mathbb{N}_0} \left[ a + \frac{1}{m}, b - \frac{1}{m} \right] \quad (19.24)$$

If  $x \in \bigcup_{n \in \mathbb{N}_0} [a + \frac{1}{n}, b - \frac{1}{n}]$  then  $\exists n \in \mathbb{N}_0$  such that  $x \in [a + \frac{1}{n}, b - \frac{1}{n}] \Rightarrow a < a + \frac{1}{n} \leq x < b - \frac{1}{n} < b$  proving that  $a < x < b$  or  $x \in ]a, b[.$  So  $\bigcup_{n \in \mathbb{N}_0} [a + \frac{1}{n}, b - \frac{1}{n}] \subseteq ]a, b[$  which together with 19.24 gives

$$]a, b[ = \bigcup_{n \in \mathbb{N}_0} \left[ a + \frac{1}{m}, b - \frac{1}{m} \right]$$

13. Let  $x \in ]-\infty, a[$  then  $x < a$  and using 9.56 there exists a  $n \in \mathbb{N}_0$  such that  $x + \frac{1}{n} < a \Rightarrow x < a - \frac{1}{n} \Rightarrow x \in ]-\infty, a - \frac{1}{n}]$  proving that

$$]-\infty, a[ \subseteq \bigcup_{n \in \mathbb{N}_0} \left[ -\infty, a - \frac{1}{n} \right] \quad (19.25)$$

If  $x \in \bigcup_{n \in \mathbb{N}_0} ]-\infty, a - \frac{1}{n}]$  then  $\exists n \in \mathbb{N}_0$  such that  $x \in ]-\infty, a - \frac{1}{n}] \Rightarrow x \leq a - \frac{1}{n} < a$  giving  $x \in ]-\infty, a[$  and proving that  $\bigcup_{n \in \mathbb{N}_0} ]-\infty, a[ \subseteq ]-\infty, a[.$  Using this together with 19.25 proves

$$]-\infty, a[ = \bigcup_{n \in \mathbb{N}_0} \left[ -\infty, a - \frac{1}{n} \right]$$

14. Let  $x \in ]a, \infty[$  then  $a < x$  and using 9.56 there exists a  $n \in \mathbb{N}_0$  such that  $a + \frac{1}{n} < x \Rightarrow x \in [a + \frac{1}{n}, \infty[$ . Hence

$$]a, \infty[ \subseteq \bigcup_{n \in \mathbb{N}_0} \left[ a + \frac{1}{n}, \infty \right[ \quad (19.26)$$

If  $x \in \bigcup_{n \in \mathbb{N}_0} [a + \frac{1}{n}, \infty[$  then  $\exists n \in \mathbb{N}_0$  such that  $x \in [a + \frac{1}{n}, \infty[ \Rightarrow a < a + \frac{1}{n} \leq x$  hence  $x \in ]a, \infty[$ . This proves that  $\bigcup_{n \in \mathbb{N}_0} [a + \frac{1}{n}, \infty[ \subseteq ]a, \infty[$ , combining this with 19.26 gives

$$]a, \infty[ = \bigcup_{n \in \mathbb{N}_0} \left[ a + \frac{1}{n}, \infty \right[$$

15. Let  $x \in [a, b[$  then  $a \leq x \wedge x < b \Rightarrow a \leq x \wedge \neg(b \leq x)$  so that  $x \in [a, \infty[$  and  $x \notin [b, \infty[$  proving that

$$[a, b[ \subseteq [a, \infty[ \setminus [b, \infty[ \quad (19.27)$$

If  $x \in [a, \infty[ \setminus [b, \infty[$  then  $a \leq x \wedge \neg(b \leq x)$  or  $a \leq x$  and  $x < b$  proving that  $[a, \infty[ \setminus [b, \infty[ \subseteq [a, b[$ . Combining this with 19.27 proves that

$$[a, b[ = [a, \infty[ \setminus [b, \infty[$$

16. Let  $x \in ]a, b]$  then  $a < x \wedge x \leq b$  then  $\neg(x \leq a) \wedge x \leq b$  so that  $x \notin ]-\infty, a] \wedge x \in ]-\infty, b]$  proving that  $x \in ]-\infty, b[ \setminus ]-\infty, a]$ . For the opposite inclusion let  $x \in ]-\infty, b[ \setminus ]-\infty, a]$  then  $x \leq b$  and  $\neg(x \leq a)$  proving that  $x \leq b$  and  $a < x$  or  $x \in ]a, b]$ . Hence

$$]a, b] = ]-\infty, b[ \setminus ]-\infty, a]$$

□

**Proposition 19.26.** *We have the following equivalent definitions for  $\mathcal{B}[\mathbb{R}]$*

$$\begin{aligned} \mathcal{B}[\mathbb{R}] &= \sigma[\mathcal{I}_{[]}]] \\ &= \sigma[\mathcal{I}_{[]}]] \\ &= \sigma[\mathcal{I}_{[]}]] \\ &= \sigma[\mathcal{I}_{[]}]] \\ &= \sigma[\{[a, \infty[ | a \in \mathbb{R}\}] \\ &= \sigma[\{[a, \infty[ | a \in \mathbb{R}\}] \\ &= \sigma[\{]-\infty, a] | a \in \mathbb{R}\}] \\ &= \sigma[\{]-\infty, a[ | a \in \mathbb{R}\}] \\ &= \sigma[\{C \subseteq \mathbb{R} | C \text{ is closed}\}] \end{aligned}$$

**Proof.** Let  $U \in \mathcal{T}_{\mathbb{R}}$  then there exist a countable family  $\{[a_i, b_i]\}_{i \in I}$  with  $\forall i \in I \ \emptyset \neq [a_i, b_i]$  such that  $U = \bigcup_{i \in I} [a_i, b_i]$  [see 11.76]. As  $\{[a_i, b_i]\}_{i \in I} \subseteq \mathcal{I}_{[]}$  we have using 19.14 that  $\bigcup_{i \in I} [a_i, b_i] \in \sigma[\mathcal{I}_{[]}]]$  proving that  $\mathcal{T}_{\mathbb{R}} \subseteq \sigma[\mathcal{I}_{[]}]]$ . Hence using 19.17 we have

$$\mathcal{B}[\mathbb{R}] = \sigma[\mathcal{T}_{\mathbb{R}}] \subseteq \sigma[\mathcal{I}_{[]}]] \quad (19.28)$$

If  $]a, b[ \in \mathcal{I}_{[]}$  then using the previous lemma [see 19.25] we have  $]a, b[ = \bigcup_{n \in \mathbb{N}_0} [a, b - \frac{1}{n}]$ . For  $n \in \mathbb{N}_0$  we have  $\begin{cases} [a, b - \frac{1}{n}] = \emptyset \in \mathcal{I}_{[]} \\ \emptyset \neq [a, b - \frac{1}{n}] \in \mathcal{I}_{[]} \end{cases} \in \sigma[\mathcal{I}_{[]}]]$  so that  $\bigcup_{n \in \mathbb{N}_0} [a, b - \frac{1}{n}] \in \sigma[\mathcal{I}_{[]}]]$ . Hence  $\mathcal{I}_{[]} \subseteq \sigma[\mathcal{I}_{[]}]]$  which by 19.17 proves

$$\sigma[\mathcal{I}_{[]}]] \subseteq \sigma[\mathcal{I}_{[]}]] \quad (19.29)$$

If  $]a, b[ \in \mathcal{I}_{[]}]]$  then using the previous lemma [see 19.25] we have  $]a, b[ = \bigcup_{n \in \mathbb{N}_0} [a + \frac{1}{n}, b]$ . For  $n \in \mathbb{N}_0$  we have  $\begin{cases} [a + \frac{1}{n}, b] = \emptyset \in \mathcal{I}_{[]} \\ \emptyset \neq [a + \frac{1}{n}, b] \in \mathcal{I}_{[]} \end{cases} \in \sigma[\mathcal{I}_{[]}]]$  so that  $]a, b[ = \bigcup_{n \in \mathbb{N}_0} [a + \frac{1}{n}, b] \in \sigma[\mathcal{I}_{[]}]]$ . Hence  $\mathcal{I}_{[]} \subseteq \sigma[\mathcal{I}_{[]}]]$  which by 19.17 proves that

$$\sigma[\mathcal{I}_{[]}]] \subseteq \sigma[\mathcal{I}_{[]}]] \quad (19.30)$$

Now if  $[a, b] \in \mathcal{I}_{[]} \text{ [so that } a \leq b]$  then by the previous lemma [see 19.25] we have that  $[a, b] = \bigcap_{n \in \mathbb{N}_0} [a, b + \frac{1}{n}]$ . As  $\forall n \in \mathbb{N}_0 \emptyset \neq [a, b + \frac{1}{n}] \in \mathcal{I}_{[]} \subseteq \sigma[\mathcal{I}_{[]}]$  it follows that  $[a, b] = \bigcap_{n \in \mathbb{N}_0} [a, b + \frac{1}{n}] \in \sigma[\mathcal{I}_{[]}]$ . So  $\mathcal{I}_{[]} \subseteq \sigma[\mathcal{I}_{[]}]$  and using 19.17 we have

$$\sigma[\mathcal{I}_{[]} \subseteq \sigma[\mathcal{I}_{[]}] \quad (19.31)$$

If  $[a, b] \in \mathcal{I}_{[]} \text{ [so that } a < b]$  then by the previous lemma [see 19.25]  $[a, b] = \bigcap_{n \in \mathbb{N}_0} [a - \frac{1}{n}, b]$ . Now as  $\forall n \in \mathbb{N}_0$  we have that  $[a - \frac{1}{n}, b] \in \mathcal{T}_{\mathbb{R}} \subseteq \sigma[\mathcal{T}_{\mathbb{R}}]$  [see 11.72] giving  $\bigcap_{n \in \mathbb{N}_0} [a - \frac{1}{n}, b] \in \sigma[\mathcal{T}_{\mathbb{R}}]$ . So  $\mathcal{I}_{[]} \subseteq \sigma[\mathcal{T}_{\mathbb{R}}]$  proving by 19.17 that

$$\sigma[\mathcal{I}_{[]} \subseteq \sigma[\mathcal{T}_{\mathbb{R}}] \stackrel{\text{def}}{=} \mathcal{B}[\mathbb{R}] \quad (19.32)$$

Next combining 19.28, 19.29, 19.30, 19.31 and 19.32 proves

$$\sigma[\mathcal{I}_{[]} = \sigma[\mathcal{I}_{[]} = \sigma[\mathcal{I}_{[]} = \sigma[\mathcal{I}_{[]} = \mathcal{B}[\mathbb{R}] \quad (19.33)$$

Next if  $[a, b] \in \mathcal{I}_{[]} \text{ then using 19.25 we have that } [a, b] = [a, \infty[ \setminus [b, \infty[ \in \sigma[\{[a, \infty[ | a \in \mathbb{R}\}]$ . So  $\mathcal{I}_{[]} \subseteq \sigma[\{[a, \infty[ | a \in \mathbb{R}\}]$  giving by 19.17

$$\mathcal{B}[\mathbb{R}] \stackrel{19.33}{=} \sigma[\mathcal{I}_{[]} \subseteq \sigma[\{[a, \infty[ | a \in \mathbb{R}\}] \quad (19.34)$$

If  $[a, \infty[ \in \{[a, \infty[ | a \in \mathbb{R}\}$  then using 19.25 we have that  $[a, \infty[ = \bigcap_{n \in \mathbb{N}_0} [a - \frac{1}{n}, \infty[ \in \sigma[\{[a, \infty[ | a \in \mathbb{R}\}]$ . So  $\{[a, \infty[ | a \in \mathbb{R}\} \subseteq \sigma[\{[a, \infty[ | a \in \mathbb{R}\}]$  giving by 19.17 that

$$\sigma[\{[a, \infty[ | a \in \mathbb{R}\}] \subseteq \sigma[\{[a, \infty[ | a \in \mathbb{R}\}] \quad (19.35)$$

Next if  $[a, \infty[ \in \{[a, \infty[ | a \in \mathbb{R}\}$  then by 11.73  $[a, \infty[$  is open so that  $\{[a, \infty[ | a \in \mathbb{R}\} \subseteq \mathcal{T}_{\mathbb{R}} \subseteq \sigma[\mathcal{T}_{\mathbb{R}}]$ . Hence using 19.17 we have

$$\sigma[\{[a, \infty[ | a \in \mathbb{R}\}] \subseteq \sigma[\mathcal{T}_{\mathbb{R}}] = \mathcal{B}[\mathbb{R}] \quad (19.36)$$

Combining 19.34, 19.35 and 19.36 gives

$$\sigma[\{[a, \infty[ | a \in \mathbb{R}\}] = \sigma[\{[a, \infty[ | a \in \mathbb{R}\}] = \mathcal{B}[\mathbb{R}] \quad (19.37)$$

Next if  $[-\infty, a] \in \{[-\infty, a] | a \in \mathbb{R}\}$  then by 19.25 we have that  $[-\infty, a] = \mathbb{R} \setminus [a, \infty[ \in \sigma[\{[a, \infty[ | a \in \mathbb{R}\}]$ . So using 19.17 we have that

$$\sigma[\{[-\infty, a] | a \in \mathbb{R}\}] \subseteq \sigma[\{[a, \infty[ | a \in \mathbb{R}\}] \quad (19.38)$$

If  $[b, a] \in \mathcal{I}_{[]}$  then by 19.25 we have that  $[b, a] = [-\infty, b] \setminus [-\infty, a] \in \sigma[\{[-\infty, a] | a \in \mathbb{R}\}]$  proving that  $\mathcal{I}_{[]} \subseteq \sigma[\{[-\infty, a] | a \in \mathbb{R}\}]$ . So by 19.17 again we have

$$\mathcal{B}[\mathbb{R}] \stackrel{19.33}{=} \sigma[\mathcal{I}_{[]} \subseteq \sigma[\{[-\infty, a] | a \in \mathbb{R}\}] \quad (19.39)$$

Now if  $[-\infty, a] \in \{[-\infty, a] | a \in \mathbb{R}\}$  then by 19.25  $[-\infty, a] = \bigcap_{n \in \mathbb{N}_0} [-\infty, a + \frac{1}{n}] \in \sigma[\{[-\infty, a] | a \in \mathbb{R}\}]$  proving that  $\{[-\infty, a] | a \in \mathbb{R}\} \subseteq \sigma[\{[-\infty, a] | a \in \mathbb{R}\}]$ . By 19.17 we have then

$$\sigma[\{[-\infty, a] | a \in \mathbb{R}\}] \subseteq \sigma[\{[-\infty, a] | a \in \mathbb{R}\}] \quad (19.40)$$

If  $[-\infty, a] \in \{[-\infty, a] | a \in \mathbb{R}\}$  then by 11.73  $[-\infty, a]$  is open so that  $\{[-\infty, a] | a \in \mathbb{R}\} \subseteq \mathcal{T}_{\mathbb{R}} \subseteq \sigma[\mathcal{T}_{\mathbb{R}}]$ , hence using 19.17 we have that

$$\sigma[\{[-\infty, a] | a \in \mathbb{R}\}] \subseteq \sigma[\mathcal{T}_{\mathbb{R}}] = \mathcal{B}[\mathbb{R}] \quad (19.41)$$

Combining 19.39, 19.40 and 19.41 gives then

$$\sigma[\{[-\infty, a] | a \in \mathbb{R}\}] = \sigma[\{[-\infty, a] | a \in \mathbb{R}\}] = \mathcal{B}[\mathbb{R}] \quad (19.42)$$

Next using 19.21 we have

$$\mathcal{B}[\mathbb{R}] = \sigma[\mathcal{T}_{\mathbb{R}}] = \sigma[\{C \subseteq \mathbb{R} | C \text{ is closed}\}] \quad (19.43)$$

Summarizing 19.33, 19.37, 19.42 and 19.43 we have

$$\begin{aligned}
 \mathcal{B}[\mathbb{R}] &= \sigma[\mathcal{I}_1] \\
 &= \sigma[\mathcal{I}_2] \\
 &= \sigma[\mathcal{I}_3] \\
 &= \sigma[\mathcal{I}_4] \\
 &= \sigma[\{[a, \infty] | a \in \mathbb{R}\}] \\
 &= \sigma[\{]a, \infty] | a \in \mathbb{R}\}] \\
 &= \sigma[\{]-\infty, a] | a \in \mathbb{R}\}] \\
 &= \sigma[\{]-\infty, a[ | a \in \mathbb{R}\}] \\
 &= \sigma[\{C \subseteq \mathbb{R} | C \text{ is closed}\}]
 \end{aligned}$$

□

### 19.2.2 Borel algebra on $\overline{\mathbb{R}}$

**Notation 19.27.** Consider the topology on the extended real numbers  $\mathcal{T}_{\overline{\mathbb{R}}}$  [see 17.38] then the Borel algebra on  $\overline{\mathbb{R}}$  is defined  $\sigma[\mathcal{T}_{\overline{\mathbb{R}}}]$  and noted as  $\mathcal{B}[\overline{\mathbb{R}}]$

We formulate now some alternative definitions of the Borel algebra on  $\overline{\mathbb{R}}$ . First we need the following set relations.

**Lemma 19.28.**

1.  $\forall a \in \mathbb{R}$  we have  $[a, \infty] = \bigcap_{n \in \mathbb{N}_0} [a - \frac{1}{n}, \infty]$
2.  $\{\infty\} = \bigcap_{n \in \mathbb{N}_0} ]n, \infty]$
3.  $\{-\infty\} = \bigcap_{n \in \mathbb{N}_0} [-\infty, -n[$

**Proof.**

1. Let  $n \in \mathbb{N}_0$  then if  $x \in [a, \infty]$   $a \leq x \leq \infty \Rightarrow a - \frac{1}{n} < a \leq x$  proving  $[a, \infty] \subseteq [a - \frac{1}{n}, \infty]$ . Hence

$$[a, \infty] \subseteq \bigcap_{n \in \mathbb{N}_0} [a - \frac{1}{n}, \infty] \quad (19.44)$$

Take  $x \in \bigcap_{n \in \mathbb{N}_0} [a - \frac{1}{n}, \infty]$  then  $\forall n \in \mathbb{N}_0$  we have  $x \in [a - \frac{1}{n}, \infty]$ . Assume that  $x < a$  then using 9.56 there exists a  $n \in \mathbb{N}_0$  such that  $x + \frac{1}{n} < a \Rightarrow x < a - \frac{1}{n}$ , hence  $x \notin [a - \frac{1}{n}, \infty]$  a contradiction. So we must have that  $a \leq x$  and as always  $x \leq \infty$  we have  $x \in [a, \infty]$ . So  $\bigcap_{n \in \mathbb{N}_0} [a - \frac{1}{n}, \infty] \subseteq [a, \infty]$  which combined with 19.44 gives

$$[a, \infty] = \bigcap_{n \in \mathbb{N}_0} [a - \frac{1}{n}, \infty]$$

2. As  $\forall n \in \mathbb{N}_0$  we have  $n < \infty \leq \infty$  it follows that  $\infty \in \bigcap_{n \in \mathbb{N}_0} ]n, \infty]$ , hence

$$\{\infty\} \subseteq \bigcap_{n \in \mathbb{N}_0} ]n, \infty] \quad (19.45)$$

Let  $x \in \bigcap_{n \in \mathbb{N}_0} ]n, \infty]$  then  $\forall n \in \mathbb{N}_0$  we have that  $n < x \leq \infty$ . Assume that  $x \neq \infty$  then  $x < \infty$  and as  $-\infty < n < x$  we have also that  $x \neq -\infty$  hence  $x \in \mathbb{R}$ , then using 9.56 there exists a  $n \in \mathbb{N}_0$  such that  $x < n$  a contradiction. So we must have that  $x = \infty$  which proves that  $\bigcap_{n \in \mathbb{N}_0} ]n, \infty] \subseteq \{\infty\}$  and combining this with 19.45 gives

$$\{\infty\} = \bigcap_{n \in \mathbb{N}_0} ]n, \infty]$$

3. As  $\forall n \in \mathbb{N}_0$  we have  $-\infty \leq -\infty < -n$  it follows that  $-\infty \in [-\infty, n[$ , hence

$$\{-\infty\} \subseteq \bigcap_{n \in \mathbb{N}_0} [-\infty, -n[ \quad (19.46)$$

Let  $x \in \bigcap_{n \in \mathbb{N}_0} [-\infty, -n[$  then  $\forall n \in \mathbb{N}_0$  we have  $-\infty \leq x < -n$ , Assume that  $x \neq -\infty$  then  $-\infty < x$  and as  $x < -n < \infty$  we have also  $x \neq \infty$  hence  $x \in \mathbb{R}$ . Using 9.56 there exists a  $n \in \mathbb{N}_0$  such that  $-x < n$  or  $-n < x$  a contradiction. So we must have that  $x = -\infty$  which proves that  $\bigcap_{n \in \mathbb{N}_0} [-\infty, n] \subseteq \{-\infty\}$ , combining this with 19.46 gives

$$\{-\infty\} = \bigcap_{n \in \mathbb{N}_0} [-\infty, -n] \quad \square$$

**Corollary 19.29.** For  $\sigma(\{[a, \infty] | a \in \mathbb{R}\})$  we have the following

1.  $\forall a \in \mathbb{R}$  we have that  $[a, \infty] \in \sigma(\{[a, \infty] | a \in \mathbb{R}\})$
2.  $\forall a \in \mathbb{R}$  we have that  $[-\infty, a[ \in \sigma(\{[a, \infty] | a \in \mathbb{R}\})$
3.  $\forall a, b \in \mathbb{R}$  with  $a \leq b$  we have that  $]a, b[ \in \sigma(\{[a, \infty] | a \in \mathbb{R}\})$
4.  $\forall U \in \mathcal{T}_{\mathbb{R}}$  we have that  $U \in \sigma(\{[a, \infty] | a \in \mathbb{R}\})$
5.  $\{\infty\} \in \sigma(\{[a, \infty] | a \in \mathbb{R}\})$
6.  $\{-\infty\} \in \sigma(\{[a, \infty] | a \in \mathbb{R}\})$

**Proof.**

1. Let  $a \in \mathbb{R}$  then  $\{]a - \frac{1}{n}, \infty[\}_{n \in \mathbb{N}_0} \subseteq \{[a, \infty] | a \in \mathbb{R}\} \subseteq \sigma(\{[a, \infty] | a \in \mathbb{R}\})$ , hence  $\bigcup_{n \in \mathbb{N}_0} ]a - \frac{1}{n}, \infty[ \in \sigma(\{[a, \infty] | a \in \mathbb{R}\})$ . As  $[a, \infty] \underset{19.28}{=} \bigcup_{n \in \mathbb{N}_0} ]a - \frac{1}{n}, \infty[$  we have that

$$[a, \infty] \in \sigma(\{[a, \infty] | a \in \mathbb{R}\})$$

2. Let  $a \in \mathbb{R}$  then

$$\begin{aligned} x \in \overline{\mathbb{R}} \setminus [a, \infty] &\Leftrightarrow x \in [-\infty, \infty] \wedge x \notin [a, \infty] \\ &\Leftrightarrow (-\infty \leq x \wedge x \leq \infty) \wedge \neg(a \leq x \leq \infty) \\ &\Leftrightarrow (-\infty \leq x \wedge x \leq \infty) \wedge (x < a \vee \infty < x) \\ &\Leftrightarrow -\infty \leq x \wedge x \leq \infty \wedge x < a \\ &\Leftrightarrow -\infty \leq x \wedge x < a \\ &\Leftrightarrow x \in [-\infty, a[ \end{aligned}$$

proving that  $[-\infty, a[ = \overline{\mathbb{R}} \setminus [a, \infty]$  which as by (1)  $[a, \infty] \in \sigma(\{[a, \infty] | a \in \mathbb{R}\})$  proves

$$[-\infty, a[ \in \sigma(\{[a, \infty] | a \in \mathbb{R}\})$$

3. Let  $a, b \in \mathbb{R}$  with  $a \leq b$  then

$$\begin{aligned} x \in ]a, \infty] \bigcap [-\infty, b[ &\Leftrightarrow x \in ]a, \infty] \wedge x \in [-\infty, b[ \\ &\Leftrightarrow a < x \wedge x \leq \infty \wedge -\infty \leq x \wedge x < b \\ &\Leftrightarrow a < x \wedge x < b \\ &\Leftrightarrow x \in ]a, b[ \end{aligned}$$

proving that  $]a, b[ = ]a, \infty] \bigcap [-\infty, b[$ . As  $]a, \infty] \in \sigma(\{[a, \infty] | a \in \mathbb{R}\})$  and by (2)  $[-\infty, b[ \in \sigma(\{[a, \infty] | a \in \mathbb{R}\})$  we have that  $]a, \infty] \bigcap [-\infty, b[ \in \sigma(\{[a, \infty] | a \in \mathbb{R}\})$ . So

$$]a, b[ \in \sigma(\{[a, \infty] | a \in \mathbb{R}\})$$

4. Let  $U$  be a open set in  $\mathbb{R}$  then there exists a sequence  $\{]a_i, b_i[\}_{i \in \mathbb{N}_0}$  such that  $U = \bigcup_{i \in \mathbb{N}_0} ]a_i, b_i[$  [see 11.76]. As by (3) we have that  $\{]a_i, b_i[\}_{i \in \mathbb{N}_0} \subseteq \sigma(\{[a, \infty] | a \in \mathbb{R}\})$  [which is a  $\sigma$ -algebra] it follows that  $\bigcup_{i \in \mathbb{N}_0} ]a_i, b_i[ \in \sigma(\{[a, \infty] | a \in \mathbb{R}\})$ . So

$$U \in \sigma(\{[a, \infty] | a \in \mathbb{R}\})$$

5. Using 19.28 we have that  $\{\infty\} = \bigcap_{n \in \mathbb{N}_0} [n, \infty]$  which, as  $\{[n, \infty]\}_{n \in \mathbb{N}_0} \in \sigma(\{[a, \infty] | a \in \mathbb{R}\})$ , proves that

$$\{\infty\} \in \sigma(\{[a, \infty] | a \in \mathbb{R}\})$$

6. Using 19.28 we have that  $\{-\infty\} = \bigcap_{n \in \mathbb{N}_0} [-\infty, -n]$  which, as by (2)  $\{[-\infty, -n] | n \in \mathbb{N}_0\} \subseteq \sigma(\{[a, \infty] | a \in \mathbb{R}\})$ , proves that

$$\{-\infty\} \in \sigma(\{[a, \infty] | a \in \mathbb{R}\}) \quad \square$$

We can now find equivalent definitions of the Borel algebra's on the extended real numbers.

**Lemma 19.30.**  $\sigma[\mathcal{T}_{\bar{\mathbb{R}}}] = \sigma(\{[a, \infty] | a \in \mathbb{R}\})$

**Proof.** Let  $I \in \{[a, \infty] | a \in \mathbb{R}\}$  then there exists a  $a \in \mathbb{R}$  such that  $I = [a, \infty]$ , using 17.36 and 17.37 we have that  $I \in \mathcal{B}_{\infty} \in \mathcal{T}_{\bar{\mathbb{R}}}$ . So  $\{[a, \infty] | a \in \mathbb{R}\} \subseteq \mathcal{T}_{\bar{\mathbb{R}}} \subseteq \sigma[\mathcal{T}_{\bar{\mathbb{R}}}]$ . Hence

$$\sigma(\{[a, \infty] | a \in \mathbb{R}\}) \subseteq \sigma[\mathcal{T}_{\bar{\mathbb{R}}}] \quad (19.47)$$

Let  $U \in \mathcal{T}_{\bar{\mathbb{R}}}$  then by 17.39 we have that  $U \cap \mathbb{R} \in \mathcal{T}_{\mathbb{R}}$ . From  $\bar{\mathbb{R}} = \mathbb{R} \cup \{-\infty\} \cup \{\infty\}$  it follows that

$$U = U \cap \bar{\mathbb{R}} = (U \cap \mathbb{R}) \cup (U \cap \{-\infty\}) \cup (U \cap \{\infty\}) \quad (19.48)$$

Now as  $U \cap \mathbb{R} \in \mathcal{T}_{\mathbb{R}}$  we have by 19.29 that

$$U \cap \mathbb{R} \in \sigma(\{[a, \infty] | a \in \mathbb{R}\}) \quad (19.49)$$

Also  $U \cap \{-\infty\}$  is either  $\emptyset$  or  $\{-\infty\}$  and as by 19.29  $\{-\infty\} \in \sigma(\{[a, \infty] | a \in \mathbb{R}\})$  we have that

$$U \cap \{-\infty\} \in \sigma(\{[a, \infty] | a \in \mathbb{R}\}) \quad (19.50)$$

Finally  $U \cap \{\infty\}$  is either  $\emptyset$  or  $\{\infty\}$  and as by 19.29  $\{\infty\} \in \sigma(\{[a, \infty] | a \in \mathbb{R}\})$  it follows that

$$U \cap \{\infty\} \in \sigma(\{[a, \infty] | a \in \mathbb{R}\}) \quad (19.51)$$

From 19.48, 19.49, 19.50 and 19.51 it follows that  $U \in \sigma(\{[a, \infty] | a \in \mathbb{R}\})$ . Hence  $\mathcal{T}_{\bar{\mathbb{R}}} \subseteq \sigma(\{[a, \infty] | a \in \mathbb{R}\})$  [a sigma algebra] so that  $\sigma[\mathcal{T}_{\bar{\mathbb{R}}}] \subseteq \sigma(\{[a, \infty] | a \in \mathbb{R}\})$ , combining this with 19.47 gives

$$\sigma(\{[a, \infty] | a \in \mathbb{R}\}) = \sigma[\mathcal{T}_{\bar{\mathbb{R}}}] \quad (19.52) \quad \square$$

**Lemma 19.31.**  $\sigma[\mathcal{T}_{\bar{\mathbb{R}}}] = \{U \in \mathcal{P}(\bar{\mathbb{R}}) | U \cap \mathbb{R} \in \mathcal{B}[\mathbb{R}]\}$  where  $\mathcal{B}[\mathbb{R}] = \sigma[\mathcal{T}_{\mathbb{R}}]$

**Proof.** First we prove that

$$\{U \in \mathcal{P}(\bar{\mathbb{R}}) | U \cap \mathbb{R} \in \mathcal{B}[\mathbb{R}]\} \text{ is a } \sigma\text{-algebra on } \bar{\mathbb{R}} \quad (19.53)$$

**Proof.**

1. As  $\bar{\mathbb{R}} \cap \mathbb{R} = \mathbb{R} \in \mathcal{B}[\mathbb{R}]$  it follows that  $\bar{\mathbb{R}} \in \{U \in \mathcal{P}(\bar{\mathbb{R}}) | U \cap \mathbb{R} \in \mathcal{B}[\mathbb{R}]\}$
2. Let  $U \in \{U \in \mathcal{P}(\bar{\mathbb{R}}) | U \cap \mathbb{R} \in \mathcal{B}[\mathbb{R}]\}$  then  $(\bar{\mathbb{R}} \setminus U) \cap \mathbb{R} = (\bar{\mathbb{R}} \cap \mathbb{R}) \setminus (U \cap \mathbb{R}) = \mathbb{R} \setminus (U \cap \mathbb{R}) \in \mathcal{B}[\mathbb{R}] \in \mathcal{B}[\mathbb{R}]$  [as  $\mathcal{B}[\mathbb{R}]$  is a  $\sigma$ -algebra and  $U \cap \mathbb{R} \in \mathcal{B}[\mathbb{R}]$ ].
3. Let  $\{U_i\}_{i \in \mathbb{N}_0} \subseteq \{U \in \mathcal{P}(\bar{\mathbb{R}}) | U \cap \mathbb{R} \in \mathcal{B}[\mathbb{R}]\}$  then  $\forall i \in \mathbb{N}_0 U_i \cap \mathbb{R} \in \mathcal{B}[\mathbb{R}]$ , further  $(\bigcup_{i \in \mathbb{N}_0} U_i) \cap \mathbb{R} = \bigcup_{i \in \mathbb{N}_0} (U_i \cap \mathbb{R}) \in \mathcal{B}[\mathbb{R}]$ .  $\square$

Take  $U \in \mathcal{T}_{\bar{\mathbb{R}}}$  then by 17.39 we have that  $U \cap \mathbb{R} \in \mathcal{T}_{\mathbb{R}} \subseteq \sigma[\mathcal{T}_{\mathbb{R}}] \stackrel{\text{def}}{=} \mathcal{B}[\mathbb{R}]$  hence  $U \in \{U \in \mathcal{P}(\bar{\mathbb{R}}) | U \cap \mathbb{R} \in \mathcal{B}[\mathbb{R}]\}$ . This proves that  $\mathcal{T}_{\bar{\mathbb{R}}} \subseteq \{U \in \mathcal{P}(\bar{\mathbb{R}}) | U \cap \mathbb{R} \in \mathcal{B}[\mathbb{R}]\}$  and as  $\{U \in \mathcal{P}(\bar{\mathbb{R}}) | U \cap \mathbb{R} \in \mathcal{B}[\mathbb{R}]\}$  is a  $\sigma$ -algebra we have

$$\sigma[\mathcal{T}_{\bar{\mathbb{R}}}] \subseteq \{U \in \mathcal{P}(\bar{\mathbb{R}}) | U \cap \mathbb{R} \in \mathcal{B}[\mathbb{R}]\} \quad (19.54)$$

Now for the opposite inclusion. Define

$$\mathcal{S} = \{U \cap \mathbb{R} : U \in \sigma[\mathcal{T}_{\bar{\mathbb{R}}}\}] \subseteq \mathbb{R} \quad (19.55)$$

then as  $\mathcal{T}_{\bar{\mathbb{R}}} \subseteq \sigma[\mathcal{T}_{\bar{\mathbb{R}}}]$  we have that

$$\mathcal{S} \supseteq \{U \cap \mathbb{R} : U \in \mathcal{T}_{\bar{\mathbb{R}}}\} \stackrel{17.39}{=} \mathcal{T}_{\mathbb{R}} \quad (19.56)$$

Further as  $\mathbb{R} \in \mathcal{T}_{\mathbb{R}} \subseteq \mathcal{T}_{\bar{\mathbb{R}}}$  [see 17.39] we have for  $U \in \sigma[\mathcal{T}_{\bar{\mathbb{R}}}]$  that  $U \cap \mathbb{R} \in \sigma[\mathcal{T}_{\bar{\mathbb{R}}}]$  so that

$$\mathcal{S} \subseteq \sigma[\mathcal{T}_{\bar{\mathbb{R}}}] \quad (19.57)$$

Further we prove that

$$\mathcal{S} \text{ is a } \sigma\text{-algebra on } \mathbb{R} \quad (19.58)$$

**Proof.**

1. As  $\mathbb{R} = \mathbb{R} \cap \mathbb{R}$  and  $\mathbb{R} \in \mathcal{T}_{\mathbb{R}} \subseteq \mathcal{T}_{\bar{\mathbb{R}}} \in \sigma[\mathcal{T}_{\bar{\mathbb{R}}}]$  we have that  $\mathbb{R} \in \mathcal{S}$
2. If  $A \in \mathcal{S}$  then  $\exists V \in \sigma[\mathcal{T}_{\bar{\mathbb{R}}}]$  such that  $A = V \cap \mathbb{R}$ . Then  $\mathbb{R} \setminus A = \mathbb{R} \setminus (V \cap \mathbb{R}) = (\mathbb{R} \setminus V) \cup (\mathbb{R} \setminus \mathbb{R}) = \mathbb{R} \setminus V \in \mathcal{T}_{\bar{\mathbb{R}}}$  [as  $\mathbb{R}, V \in \mathcal{T}_{\bar{\mathbb{R}}}$ ], further  $(\mathbb{R} \setminus A) \cap \mathbb{R} = \mathbb{R} \setminus A$ . So  $\mathbb{R} \setminus A \in \mathcal{S}$ .
3. If  $\{A_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{S}$  then  $\forall i \in \mathbb{N}_0$  then exists a  $V_i \in \mathcal{T}_{\bar{\mathbb{R}}}$  such that  $A_i = V_i \cap \mathbb{R}$ . Then  $\bigcup_{i \in \mathbb{N}_0} A_i = \bigcup_{i \in \mathbb{N}_0} (V_i \cap \mathbb{R}) = (\bigcup_{i \in \mathbb{N}_0} V_i) \cap \mathbb{R}$  where  $\bigcup_{i \in \mathbb{N}_0} V_i \in \mathcal{T}_{\bar{\mathbb{R}}}$ . So  $\bigcup_{i \in \mathbb{N}_0} A_i \in \mathcal{S}$   $\square$

Combining the above [19.58] with 19.56 gives  $\sigma(\mathcal{T}_{\mathbb{R}}) \subseteq \mathcal{S}$ . Using then 19.57 together with the fact that  $\mathcal{B}[\mathbb{R}] \stackrel{\text{def}}{=} \sigma[\mathcal{T}_{\bar{\mathbb{R}}}]$  result in

$$\mathcal{B}[\mathbb{R}] \subseteq \mathcal{S} \subseteq \sigma[\mathcal{T}_{\bar{\mathbb{R}}}] \quad (19.59)$$

Take  $A \in \{U \in \mathcal{P}[\bar{\mathbb{R}}] : U \cap \mathbb{R} \in \mathcal{B}[\mathbb{R}]\}$  then  $A = A \cap \bar{\mathbb{R}} = A \cap (\mathbb{R} \cup \{-\infty\}, \{\infty\})$  giving

$$A = (A \cap \mathbb{R}) \cap (A \cap \{-\infty\}) \cap (A \cap \{\infty\}) \quad (19.60)$$

where

1.  $A \cap \mathbb{R} \in \mathcal{B}[\mathbb{R}] \stackrel{19.59}{\Rightarrow} A \cap \mathbb{R} \in \sigma[\mathcal{T}_{\bar{\mathbb{R}}}]$
2.  $A \cap \{-\infty\}$  is either  $\emptyset \in \mathcal{T}_{\bar{\mathbb{R}}}$  or  $\{-\infty\} \in \sigma[\mathcal{T}_{\bar{\mathbb{R}}}]$  [as  $-\infty$  is closed by 17.40 and closed sets are part of  $\sigma[\mathcal{T}_{\bar{\mathbb{R}}}]$  by 19.26]
3.  $A \cap \{\infty\}$  is either  $\emptyset \in \mathcal{T}_{\bar{\mathbb{R}}}$  or  $\{\infty\} \in \sigma[\mathcal{T}_{\bar{\mathbb{R}}}]$  [as  $\infty$  is closed by 17.40 and closed sets are part of  $\sigma[\mathcal{T}_{\bar{\mathbb{R}}}]$  by 19.26]

From the above we conclude that  $\{U \in \mathcal{P}[\bar{\mathbb{R}}] : U \cap \mathbb{R} \in \mathcal{B}[\mathbb{R}]\} \subseteq \sigma[\mathcal{T}_{\bar{\mathbb{R}}}]$  which combined with 19.54 proves that

$$\sigma[\mathcal{T}_{\bar{\mathbb{R}}}] = \{U \in \mathcal{P}[\bar{\mathbb{R}}] : U \cap \mathbb{R} \in \mathcal{B}[\mathbb{R}]\} \quad \square$$

To summarize the two previous lemma's we have

**Theorem 19.32.** *The following equivalences for the Borel algebra on  $\bar{\mathbb{R}}$  exists*

$$\begin{aligned} \mathcal{B}[\bar{\mathbb{R}}] &\stackrel{\text{def}}{=} \sigma[\mathcal{T}_{\bar{\mathbb{R}}}] \\ &\stackrel{19.21}{=} \sigma[\{C \mid C \text{ is closed in } \mathcal{T}_{\bar{\mathbb{R}}}\}] \\ &= \sigma[\{]a, \infty] \mid a \in \mathbb{R}\}] \\ &= \sigma[\{U \subseteq \mathcal{P}[\bar{\mathbb{R}}] \mid U \cap \mathbb{R} \in \mathcal{B}[\mathbb{R}]\}] \end{aligned}$$

**Proof.** Use 19.30 and 19.31.  $\square$

### 19.2.3 Borel algebra on $\mathbb{R}^n$

We extend first the concept of generalized intervals on  $\mathbb{R}^n$ . First we define  $\leq$  on  $\mathbb{R}^n$  (note that this is not a order relation, we use this only to have a simpler definition of intervals).

**Definition 19.33.** Let  $n \in \mathbb{N}_0$  [see 4.7] and  $a, b \in \mathbb{R}^n$  then we say that

1.  $a \leq b$  iff  $\forall i \in \{1, \dots, n\}$  we have  $a_i \leq b_i$
2.  $a < b$  iff  $\forall i \in \{1, \dots, n\}$  we have  $a_i < b_i$

Next we can define intervals on  $\mathbb{R}^n$

**Definition 19.34.** Let  $n \in \mathbb{N}_0$ ,  $a, b \in \mathbb{R}^n$  such that  $a \leq b$  then we have

$$\begin{aligned}
 [a, b] &= \{x \in \mathbb{R}^n \mid a \leq x \wedge x \leq b\} \\
 &= \{x \in \mathbb{R}^n \mid \forall i \in \{1, \dots, n\} \text{ we have } a_i \leq x_i \wedge x_i \leq b_i\} \\
 &= \{x \in \mathbb{R}^n \mid \forall i \in \{1, \dots, n\} \text{ we have } x_i \in [a_i, b_i]\} \\
 &= \prod_{i \in \{1, \dots, n\}} [a_i, b_i] \\
 [a, b[ &= \{x \in \mathbb{R}^n \mid a \leq x \wedge x < b\} \\
 &= \{x \in \mathbb{R}^n \mid \forall i \in \{1, \dots, n\} \text{ we have } a_i \leq x_i \wedge x_i < b_i\} \\
 &= \{x \in \mathbb{R}^n \mid \forall i \in \{1, \dots, n\} \text{ we have } x_i \in [a_i, b_i[\} \\
 &= \prod_{i \in \{1, \dots, n\}} [a_i, b_i[ \\
 ]a, b] &= \{x \in \mathbb{R}^n \mid a < x \wedge x \leq b\} \\
 &= \{x \in \mathbb{R}^n \mid \forall i \in \{1, \dots, n\} \text{ we have } a_i < x_i \wedge x_i \leq b_i\} \\
 &= \{x \in \mathbb{R}^n \mid \forall i \in \{1, \dots, n\} \text{ we have } x_i \in ]a_i, b_i]\} \\
 &= \prod_{i \in \{1, \dots, n\}} ]a_i, b_i] \\
 ]a, b[ &= \{x \in \mathbb{R}^n \mid a < x \wedge x < b\} \\
 &= \{x \in \mathbb{R}^n \mid \forall i \in \{1, \dots, n\} \text{ we have } a_i < x_i \wedge x_i < b_i\} \\
 &= \{x \in \mathbb{R}^n \mid \forall i \in \{1, \dots, n\} \text{ we have } x_i \in ]a_i, b_i[\} \\
 &= \prod_{i \in \{1, \dots, n\}} ]a_i, b_i[ \\
 ]-\infty, a] &= \{x \in \mathbb{R}^n \mid x \leq a\} \\
 &= \{x \in \mathbb{R}^n \mid \forall i \in \{1, \dots, n\} \text{ we have } x_i \leq a_i\} \\
 &= \{x \in \mathbb{R}^n \mid \forall i \in \{1, \dots, n\} \text{ we have } x_i \in ]-\infty, a_i]\} \\
 &= \prod_{i \in \{1, \dots, n\}} ]-\infty, a_i] \\
 ]-\infty, a[ &= \{x \in \mathbb{R}^n \mid x < a\} \\
 &= \{x \in \mathbb{R}^n \mid \forall i \in \{1, \dots, n\} \text{ we have } x_i < a_i\} \\
 &= \{x \in \mathbb{R}^n \mid \forall i \in \{1, \dots, n\} \text{ we have } x_i \in ]-\infty, a_i[\} \\
 &= \prod_{i \in \{1, \dots, n\}} ]-\infty, a_i[ \\
 [a, \infty[ &= \{x \in \mathbb{R}^n \mid a \leq x\} \\
 &= \{x \in \mathbb{R}^n \mid \forall i \in \{1, \dots, n\} \text{ we have } a_i \leq x_i\} \\
 &= \{x \in \mathbb{R}^n \mid \forall i \in \{1, \dots, n\} \text{ we have } x_i \in [a_i, \infty[\} \\
 &= \prod_{i \in \{1, \dots, n\}} [a_i, \infty[ \\
 ]a, \infty[ &= \{x \in \mathbb{R}^n \mid a < x\} \\
 &= \{x \in \mathbb{R}^n \mid \forall i \in \{1, \dots, n\} \text{ we have } a_i < x_i\} \\
 &= \{x \in \mathbb{R}^n \mid \forall i \in \{1, \dots, n\} \text{ we have } x_i \in ]a_i, \infty[\} \\
 &= \prod_{i \in \{1, \dots, n\}} ]a_i, \infty[
 \end{aligned}$$

**Proposition 19.35.** Let  $n \in \mathbb{N}_0$  and  $a, b \in \mathbb{R}^n$  with  $a \leq b$  then we have

1.  $[a, b] = \emptyset \Leftrightarrow \exists i \in \{1, \dots, n\}$  such that  $a_i = b_i$
2.  $]a, b] = \emptyset \Leftrightarrow \exists i \in \{1, \dots, n\}$  such that  $a_i = b_i$
3.  $]a, b[ = \emptyset \Leftrightarrow \exists i \in \{1, \dots, n\}$  such that  $a_i = b_i$

**Proof.**

1.  $\Rightarrow$ . If  $[a, b] = \emptyset$  then by 5.103  $\exists i \in \{1, \dots, n\}$  such that  $[a_i, b_i] = \emptyset$ . If  $a_i < b_i$  then by 9.56 there exists a  $x \in \mathbb{R}$  such that  $a_i < x < b_i \Rightarrow x \in [a_i, b_i]$  contradicting  $[a_i, b_i] = \emptyset$ . So we must have  $a_i = b_i$ .  
 $\Leftarrow$ . If  $a_i = b_i$  then if  $x \in [a_i, b_i]$  we have  $a_i \leq x \wedge x < a_i$  leading to the contradiction  $a_i < a_i$  so  $[a_i, b_i] = \emptyset$
2.  $\Rightarrow$ . If  $]a, b] = \emptyset$  then by 5.103  $\exists i \in \{1, \dots, n\}$  such that  $]a_i, b_i] = \emptyset$ . If  $a_i < b_i$  then by 9.56 there exists a  $x \in \mathbb{R}$  such that  $a_i < x < b_i \Rightarrow x \in ]a_i, b_i]$  contradicting  $]a_i, b_i] = \emptyset$ . So we must have  $a_i = b_i$ .  
 $\Leftarrow$ . If  $a_i = b_i$  then if  $x \in ]a_i, b_i]$  we have  $a_i < x \wedge x \leq a_i$  leading to the contradiction  $a_i < a_i$  so  $[a_i, b_i] = \emptyset$
3.  $\Rightarrow$ . If  $]a, b[ = \emptyset$  then by 5.103  $\exists i \in \{1, \dots, n\}$  such that  $]a_i, b_i[ = \emptyset$ . If  $a_i < b_i$  then by 9.56 there exists a  $x \in \mathbb{R}$  such that  $a_i < x < b_i \Rightarrow x \in ]a_i, b_i[$  contradicting  $]a_i, b_i[ = \emptyset$ . So we must have  $a_i = b_i$ .  
 $\Leftarrow$ . If  $a_i = b_i$  then if  $x \in ]a_i, b_i[$  we have  $a_i < x \wedge x < a_i$  leading to the contradiction  $a_i < a_i$  so  $[a_i, b_i] = \emptyset$   $\square$

Another simplifying notation is a overloading of the plus operator

**Definition 19.36.** Let  $n \in \mathbb{N}_0$ ,  $x \in \mathbb{R}^n$  and  $y \in \mathbb{R}$  then  $x + y \in \mathbb{R}^n$  is defined by  $(x + y)_i = x_i + y$  for all  $i \in \{1, \dots, n\}$ .

To simplify notation further we have the following definitions.

**Definition 19.37.** Let  $n \in \mathbb{N}_0$  and  $a, b \in \mathbb{R}$  then we define the following sets

1.  $[a, b]^n = \{x \in \mathbb{R}^n | a \leq x_i \leq b\}$
2.  $[a, b[^n = \{x \in \mathbb{R}^n | a \leq x_i < b\}$
3.  $]a, b]^n = \{x \in \mathbb{R}^n | a < x_i \leq b\}$
4.  $]a, b[^n = \{x \in \mathbb{R}^n | a < x_i < b\}$
5.  $]-\infty, a]^n = \{x \in \mathbb{R}^n | x_i \leq a\}$
6.  $]-\infty, a[^n = \{x \in \mathbb{R}^n | x_i < a\}$
7.  $[a, \infty[^n = \{x \in \mathbb{R}^n | a \leq x_i\}$
8.  $]a, \infty[^n = \{x \in \mathbb{R}^n | a < x_i\}$

Using the intervals in  $\mathbb{R}^n$  we can define the following set of half-open rectangles

**Definition 19.38.** Let  $n \in \mathbb{N}_0$  then we define  $\mathcal{R}^n$  by  $\mathcal{R}^n = \{[a, b] | a, b \in \mathbb{R}^n \wedge a \leq b\}$

The following proposition allows us to define unique begin and endpoint for elements of  $\mathcal{R}^n$ .

**Theorem 19.39.** Let  $n \in \mathbb{N}_0$  and  $\emptyset \neq R \in \mathcal{R}^n$  then there exists unique  $a, b \in \mathbb{R}^n$  such that  $R = [a, b[$

**Proof.** If  $I \in \mathcal{R}^n$  then there exists by definition of  $\mathcal{R}^n$  a  $a, b \in \mathbb{R}^n$  such that  $I = [a, b]$ . Assume that there also exists a  $a', b' \in \mathbb{R}^n$  such that  $R = [a', b']$ . Using 5.104 and the fact that  $R \neq \emptyset$  we must have then that  $\forall i \in \{1, \dots, n\}$  we have that  $[a_i, b_i] = [a'_i, b'_i]$ . As  $R \neq \emptyset$  we have by 19.35 that  $[a_i, b_i] \neq \emptyset$  so that using 2.189  $a_i = a'_i \wedge b_i = b'_i$ . Hence  $a = a' \wedge b = b'$  proving uniqueness.  $\square$

The above theorem allows us to define the boundaries of elements of  $\mathcal{R}^n$  and the volume of these elements.

**Definition 19.40.** Let  $n \in \mathbb{N}_0$  then we define:

1. begin:  $\mathcal{R}^n \setminus \{\emptyset\} \rightarrow \mathbb{R}^n$  by  $\text{begin}(I) = a$  where  $I = [a, b]$
2. end:  $\mathcal{R}^n \setminus \{\emptyset\} \rightarrow \mathbb{R}^n$  by  $\text{begin}(I) = b$  where  $I = [a, b]$
3.  $v^n: \mathcal{R}^n \rightarrow [0, \infty[$  by  $v^n(I) = \begin{cases} 0 & \text{if } I = \emptyset \\ \prod_{i=1}^n (b_i - a_i) & \text{where } I = [a, b] \end{cases}$

**Note 19.41.** If  $R \in \mathcal{R}^n$  then there exists a  $a, b \in \mathbb{R}^n$  with  $a \leq b$  such that  $R = [a, b]$  and  $v^n(R) = \prod_{i=1}^n (b_i - a_i)$

**Proof.** If  $R \in \mathcal{R}^n$  then by definition  $R = [a, b]$  where  $a, b \in \mathbb{R}^n$  and  $a \leq b$ . We have now two cases to check for  $R$ :

**$R = \emptyset$ .** Then if  $\forall i \in \{1, \dots, n\}$  we have  $a_i < b_i$  we have that  $a_i < \frac{a_i + b_i}{2} < b_i$  so that  $\left(\frac{a_1 + b_1}{2}, \dots, \frac{a_n + b_n}{2}\right) \in [a, b] = R$  contradicting  $R = \emptyset$ . So  $\exists i \in \{1, \dots, n\}$  such that  $b_i < a_i$  which as  $a \leq b \Rightarrow a_i \leq b_i$  means that  $a_i = b_i$  thence  $\prod_{i=1}^n (b_i - a_i) = 0 = v^n(R)$

**$R \neq \emptyset$ .** then by definition  $v^n(R) = \prod_{i=1}^n (b_i - a_i)$   $\square$

**Lemma 19.42.** Let  $a, b, c, d \in \mathbb{R}$  with  $a < b$  then if  $[a, b] \subseteq [c, d]$  we have  $c \leq a < b \leq d$

**Proof.** As  $a < b$  we have that  $\delta_1 = \frac{b-a}{2} > 0$ .

Assume that  $a < c$  then  $\delta_2 = \frac{c-a}{2} > 0$ . Take  $\delta = \min(\delta_1, \delta_2) > 0$  then  $a < a + \delta \leq a + \delta_1 = a + \frac{b-a}{2} = \frac{a+b}{2} < \frac{b+b}{2} = b$  proving that  $a + \delta \in [a, b]$ . Further  $a + \delta < a + \delta_2 = a + \frac{c-a}{2} = \frac{a+c}{2} < \frac{c+c}{2} = c$  proving that  $a + \delta \notin [c, d]$  contradicting  $a + \delta \in [a, b] \subseteq [c, d]$ . Hence we have

$$c \leq a$$

Assume that  $d < b$  then  $\delta_3 = \frac{b-d}{2} > 0$  take  $\gamma = \min(\delta_1, \delta_3) > 0$  then for  $b > b - \gamma > b - \delta_1 = b - \frac{b-a}{2} = \frac{b+a}{2} > \frac{a+a}{2} = a$  proving that  $b - \gamma \in [a, b]$ . Further  $b - \gamma > b - \delta_3 = b - \frac{b-d}{2} = \frac{b+d}{2} > \frac{d+d}{2} = d$  proving that  $b - \gamma \notin [c, d]$  contradicting  $b - \gamma \in [a, b] \subseteq [c, d]$ . Hence we have

$$b \leq d$$

$\square$

**Lemma 19.43.** Let  $n \in \mathbb{N}_0$  then if  $R, Q \in \mathcal{R}^n$  with  $R \subseteq Q$  we have  $v^n(R) \leq v^n(Q)$

**Proof.** For  $R$  we have to consider the following cases to consider:

**$R = \emptyset$ .** then  $v^n(R) = 0 \leq v^n(Q)$

**$R \neq \emptyset$ .** then there exists a  $x \in R$ . Further as  $R, Q \in \mathcal{R}^n$  there exists  $a, b, c, d \in \mathbb{R}^n$  with  $a \leq b$  and  $c \leq d$  such that  $R = [a, b]$  and  $Q = [c, d]$ . As  $x \in R$  we have that  $\forall i \in \{1, \dots, n\}$   $x_i \in [a_i, b_i]$ .

Assume that  $\exists j \in \{1, \dots, n\}$  such that  $[a_j, b_j] \not\subseteq [c_j, d_j]$  then there exists a  $r \in [a_j, b_j]$  such that  $r \notin [c_j, d_j]$ . Define  $y \in \mathbb{R}^n$  by  $y_i = \begin{cases} x_i & \text{if } i \in \{1, \dots, n\} \setminus \{j\} \\ r & \text{if } j = i \end{cases}$  then  $\forall i \in \{1, \dots, n\}$   $y_i \in [a_j, b_j]$

so that  $y \in R \xrightarrow{R \subseteq Q} y \in Q$ . Hence  $r = y_j \in [c_j, d_j]$  a contradiction. So  $\forall i \in \{1, \dots, n\}$  we have

$x_i \in [a_i, b_i] \subseteq [c_i, d_i]$  or  $a_i \leq x_i < b_i$  and using 19.42 that  $c_i \leq a_i < b_i \leq d_i$ . From which it follows that  $0 < b_i - a_i \leq d_i - c_i$  and finally that  $v^n(R) = \prod_{i=1}^n (b_i - a_i) \leq \prod_{i=1}^n (d_i - c_i) = v^n(Q)$ .  $\square$

Now to prove that the Borel algebra on  $\mathbb{R}^n$  can be generated from  $\mathcal{R}^n$  we must be able to write every open set in  $\mathbb{R}^n$  as a denumerable union of rectangles in  $\mathbb{R}^n$ . To be able to do this we construct a denumerable subset of  $\mathcal{I}_{[]}^n$  and  $\mathcal{R}^n$  called Dyadic intervals and Dyadic cubes.

**Definition 19.44. (Dyadic intervals)** Let  $m \in \mathbb{N}$  then we define  $\mathcal{D}_m$  by

$$\mathcal{D}_m = \left\{ \left[ \frac{z}{2^m}, \frac{z+1}{2^m} \right] \mid z \in \mathbb{Z} \right\}$$

the union of all  $\mathcal{D}_m$  is noted as  $\mathcal{D}$  so

$$\mathcal{D} = \bigcup_{m \in \mathbb{N}} \mathcal{D}_m$$

and called the set of Dyadic intervals. Dyadic intervals have the following properties

**Proposition 19.45.** Let  $k, l \in \mathbb{N}$  with  $k \leq l$  then for any  $\left[ \frac{r}{2^l}, \frac{r+1}{2^l} \right] \in \mathcal{D}_l$  and  $\left[ \frac{q}{2^k}, \frac{q+1}{2^k} \right] \in \mathcal{D}_k$  we have

$$\begin{aligned} \left[ \frac{r}{2^l}, \frac{r+1}{2^l} \right] \cap \left[ \frac{q}{2^k}, \frac{q+1}{2^k} \right] \neq \emptyset &\Rightarrow \left[ \frac{r}{2^l}, \frac{r+1}{2^l} \right] \subseteq \left[ \frac{q}{2^k}, \frac{q+1}{2^k} \right] \\ \left[ \frac{r}{2^l}, \frac{r+1}{2^l} \right] \cap \left[ \frac{q}{2^k}, \frac{q+1}{2^k} \right] = \emptyset &\Rightarrow \left[ \frac{r}{2^l}, \frac{r+1}{2^l} \right] \not\subseteq \left[ \frac{q}{2^k}, \frac{q+1}{2^k} \right] \end{aligned}$$

or in other words. We have either  $\left[ \frac{r}{2^l}, \frac{r+1}{2^l} \right] \cap \left[ \frac{q}{2^k}, \frac{q+1}{2^k} \right] = \emptyset$  or  $\left[ \frac{r}{2^l}, \frac{r+1}{2^l} \right] \subseteq \left[ \frac{q}{2^k}, \frac{q+1}{2^k} \right]$  but not both.

**Proof.** Take  $I = \left[ \frac{q}{2^l}, \frac{q+1}{2^l} \right]$  and  $J = \left[ \frac{r}{2^k}, \frac{r+1}{2^k} \right]$ . As  $k \leq l$  we have  $0 \leq l - k = n$ . Next

$$\begin{aligned} J &= \left[ \frac{r}{2^k}, \frac{r+1}{2^k} \right] \\ &= \left[ \frac{r}{2^k} \cdot \frac{2^{l-k}}{2^{l-k}}, \frac{r+1}{2^k} \cdot \frac{2^{l-k}}{2^{l-k}} \right] \\ &= \left[ \frac{r \cdot 2^{l-k}}{2^l}, \frac{r \cdot 2^{l-k} + 2^{l-k}}{2^l} \right] \\ &= \left[ \frac{r \cdot 2^n}{2^l}, \frac{r \cdot 2^n + 2^n}{2^l} \right] \end{aligned} \tag{19.61}$$

Assume that  $I \cap J \neq \emptyset$  so that  $\exists x \in I \cap J$  giving

$$\frac{q}{2^l} \leq x < \frac{q+1}{2^l} \wedge \frac{r \cdot 2^n}{2^l} \leq x < \frac{r \cdot 2^n + 2^n}{2^l} \tag{19.62}$$

If now  $\frac{q+1}{2^l} \leq \frac{r \cdot 2^n}{2^l}$  then  $x < \frac{q+1}{2^l} \leq \frac{r \cdot 2^n}{2^l} \leq x$  giving the contradiction  $x < x$  so we must have

$$\frac{r \cdot 2^n}{2^l} < \frac{q+1}{2^l} \tag{19.63}$$

Further if  $\frac{r \cdot 2^n + 2^n}{2^l} \leq \frac{q}{2^l}$  then  $x < \frac{r \cdot 2^n + 2^n}{2^l} \leq \frac{q}{2^l} \leq x$  giving the contradiction  $x < x$  again so we must have that

$$\frac{q}{2^l} < \frac{r \cdot 2^n + 2^n}{2^l} \tag{19.64}$$

Multiplying 19.63 and 19.64 by  $2^l$  gives  $r \cdot 2^n < q + 1 \wedge q < r \cdot 2^n + 2^n$ . As  $r \cdot 2^n, q + 1, r \cdot 2^n + 2^n \in \mathbb{Z}$  we have by 9.54 that  $r \cdot 2^n + 1 \leq q + 1 \wedge q + 1 \leq r \cdot 2^n + 2^n \Rightarrow r \cdot 2^n \leq q \wedge q + 1 \leq r \cdot 2^n + 2^n$ . Dividing by  $2^l$  gives:

$$\frac{r \cdot 2^n}{2^l} \leq \frac{q}{2^l} \wedge \frac{q+1}{2^l} \leq \frac{r \cdot 2^n + 2^n}{2^l} \tag{19.65}$$

Next if  $x \in I$  then  $\frac{q}{2^l} \leq x < \frac{q+1}{2^l} \stackrel{19.65}{\Rightarrow} \frac{r \cdot 2^n}{2^l} \leq x < \frac{r \cdot 2^n + 2^n}{2^l}$  proving using 19.61 that  $x \in J$ . Hence we have

$$\text{If } I \cap J \neq \emptyset \text{ then } I \subseteq J$$

Finally if  $I \cap J = \emptyset$  then if  $I \subseteq J$  we have  $\emptyset = I \cap J = I$  a contradiction as  $\frac{q}{2^l} \in I$ . Hence we have either  $I \cap J = \emptyset$  or  $I \subseteq J$  but not both.  $\square$

**Definition 19.46. (Dyadic Cubes)** Let  $n \in \mathbb{N}_0$ ,  $m \in \mathbb{N}$  and  $z \in \mathbb{Z}^n$  then  $[\frac{z}{2^m}, \frac{z+1}{2^m}] \in \mathcal{R}^n$  is called a Dyadic cube of order  $m$ . The set of Dyadic cubes of order  $m$  is noted as  $\mathcal{D}_m^n$ . So  $\mathcal{D}_m^n = \{[\frac{z}{2^m}, \frac{z+1}{2^m}] \mid z \in \mathbb{Z}^n\}$ . The set of all Dyadic cubes noted by  $\mathcal{D}^n$  hence  $\mathcal{D}^n = \bigcup_{m \in \mathbb{N}} \mathcal{D}_m^n \subseteq \mathcal{R}^n$ . SS

**Note 19.47.** As  $\frac{z}{2^m} < \frac{z+1}{2^m}$  we have  $\frac{z}{2^m} \in [\frac{z}{2^m}, \frac{z+1}{2^m}]$  proving that Dyadic cubes are always non empty.

**Lemma 19.48.** Let  $n \in \mathbb{N}_0$  then we have

1.  $\forall m \in \mathbb{N}$  and  $\forall R, Q \in \mathcal{D}_m^n$  with  $R \neq Q$  we have  $R \cap Q = \emptyset$
2. If  $k, l \in \mathbb{N}$  with  $k \leq l$  then if  $R \in \mathcal{D}_l^n$  and  $Q \in \mathcal{D}_k^n$  we have

$$\begin{aligned} R \cap Q \neq \emptyset &\Rightarrow R \subseteq Q \\ R \cap Q = \emptyset &\Rightarrow R \subseteq Q \end{aligned}$$

3.  $\forall m \in \mathbb{N}$  we have  $\mathbb{R}^n = \bigcup_{R \in \mathcal{D}_m^n} R$  [a pairwise disjoint union]

4.  $\forall m \in \mathbb{N}$  and  $\forall R \in \mathcal{D}_m^n$  we have  $v^n(R) = 2^{-n \cdot m} < 1$

**Proof.**

1. Let  $m \in \mathbb{N}$  and  $R, Q \in \mathcal{D}_m^n$  with  $R \neq Q$  then  $R = [\frac{r}{2^m}, \frac{r+1}{2^m}] = \prod_{i \in \{1, \dots, n\}} [\frac{r_i}{2^m}, \frac{r_i+1}{2^m}]$  and  $Q = [\frac{q}{2^m}, \frac{q+1}{2^m}] = \prod_{i \in \{1, \dots, n\}} [\frac{q_i}{2^m}, \frac{q_i+1}{2^m}]$ , where  $r, w \in \mathbb{Z}^n$ . As  $R \neq Q$  there exists a  $i \in \mathbb{N}_1$  such that  $[\frac{r_i}{2^m}, \frac{r_i+1}{2^m}] \neq [\frac{q_i}{2^m}, \frac{q_i+1}{2^m}]$ , Hence we must have that  $r_i \neq q_i$  and we can assume that  $r_i < q_i \Rightarrow r_i + 1 \leq q_i$  or  $\frac{r_i+1}{2^m} \leq \frac{q_i}{2^m}$ . Assume that  $x \in [\frac{r_i}{2^m}, \frac{r_i+1}{2^m}] \cap [\frac{q_i}{2^m}, \frac{q_i+1}{2^m}]$  then  $x \leq \frac{r_i+1}{2^m} \leq \frac{q_i}{2^m} < x$  leading to the contradiction  $x < x$ . So we must have that  $[\frac{r_i}{2^m}, \frac{r_i+1}{2^m}] \cap [\frac{q_i}{2^m}, \frac{q_i+1}{2^m}] = \emptyset$  from which we have [see 5.103]

$$R \cap Q = \emptyset$$

2. Let  $k, l \in \mathbb{N}$  with  $k \leq l$  and  $R \in \mathcal{D}_l^n$ ,  $Q \in \mathcal{D}_k^n$  then there exists a  $r, q \in \mathbb{Z}^n$  with  $R = [\frac{r}{2^l}, \frac{r+1}{2^l}]$  and  $Q = [\frac{q}{2^k}, \frac{q+1}{2^k}]$ . Assume  $R \not\subseteq Q$  then we have either

$$\begin{aligned} \exists i \in \{1, \dots, n\} \models [\frac{r_i}{2^l}, \frac{r_i+1}{2^l}] \cap [\frac{q_i}{2^k}, \frac{q_i+1}{2^k}] &= \emptyset. \text{ then using 5.103 we have } R \cap Q = \emptyset \\ \forall i \in \{1, \dots, n\} \models [\frac{r_i}{2^l}, \frac{r_i+1}{2^l}] \cap [\frac{q_i}{2^k}, \frac{q_i+1}{2^k}] &\neq \emptyset. \text{ then using 19.45 we have } \forall i \in \{1, \dots, n\} \\ \text{that } [\frac{r_i}{2^l}, \frac{r_i+1}{2^l}] &\subseteq [\frac{q_i}{2^k}, \frac{q_i+1}{2^k}]. \text{ Hence using 5.103 again it follows that } R \subseteq Q \text{ contradicting } R \not\subseteq Q. \text{ So this case does not apply.} \end{aligned}$$

So we conclude that if  $R \not\subseteq Q$  then  $R \cap Q = \emptyset$  or using contra position we have

$$\text{If } R \cap Q \neq \emptyset \text{ then } R \subseteq Q$$

On the other hand if  $R \cap Q = \emptyset$  and  $R \subseteq Q$  then  $\emptyset \neq R = R \cap Q = \emptyset$  a contradiction, so

$$\text{If } R \cap Q = \emptyset \text{ then } R \not\subseteq Q$$

3. Let  $m \in \mathbb{N}$  and  $x \in \mathbb{R}^n$  then  $\forall i \in \{1, \dots, n\}$  we have  $x_i \in \mathbb{R}$ . Using 9.56 (2) there exists a  $z_i \in \mathbb{Z}$  such that  $z_i \leq x_i \cdot 2^m < z_i + 1$  or  $\frac{z_i}{2^m} \leq x_i < \frac{z_i+1}{2^m}$ . So if take  $z = (z_1, \dots, z_m)$  then  $x \in \prod_{i \in \{1, \dots, n\}} [\frac{z_i}{2^m}, \frac{z_i+1}{2^m}] = [\frac{z}{2^m}, \frac{z+1}{2^m}] \in \mathcal{D}_m^n$  proving that  $\mathbb{R}^n \subseteq \bigcup_{R \in \mathcal{D}_m^n} R$ . As  $\forall R \in \mathcal{D}_m^n$  we have  $R \subseteq \mathbb{R}^n$  we have

$$\mathbb{R}^n = \bigcup_{R \in \mathcal{D}_m^n} R$$

Using (1) we have if  $D, D' \in \mathcal{D}_m^n$  with  $D \neq D'$  that  $D \cap D' = \emptyset$  so that

$$\mathbb{R}^n = \bigsqcup_{D \in \mathcal{D}_m^n} D$$

4. If  $R \in \mathcal{R}_m^n$  then  $\exists z \in \mathbb{Z}^n$  such that  $R = \prod_{i \in \{1, \dots, n\}} \left[ \frac{z_i}{2^m}, \frac{z_i + 1}{2^m} \right]$  so that

$$v^n(R) = \prod_{i=1}^n \left( \frac{z_i + 1 - z_i}{2^m} \right) = \prod_{i=1}^n \frac{1}{2^m} = \frac{1}{2^{n \cdot m}}$$

□

One of the reasons that we introduce the Dyadic cubes as a subset of  $\mathcal{R}^n$  is that  $\mathcal{D}_m^n$  is a denumerable set in contrast with  $\mathcal{R}^n$  which we need giving the definition of a  $\sigma$ -algebra. This is proved in the next lemma.

**Lemma 19.49.** *Let  $n \in \mathbb{N}_0$*

1.  $\forall m \in \mathbb{N} \mathcal{D}_m^n$  is denumerable
2.  $\mathcal{D}^n$  is denumerable

**Proof.**

1. Let  $m \in \mathbb{N}$  then as  $\mathbb{Z}$  is denumerable [see 8.54] we have by 5.102 that  $\mathbb{Z}^n$  is also denumerable.

Define now  $\beta: \mathbb{Z}^n \rightarrow \mathcal{D}_m^n$  by  $\beta(z) = \left[ \frac{z}{2^m}, \frac{z+1}{2^m} \right]$  then  $\beta$  is a bijection.

**injectivity.** If  $\beta(z) = \beta(w)$  then  $\left[ \frac{z}{2^m}, \frac{z+1}{2^m} \right] = \left[ \frac{w}{2^m}, \frac{w+1}{2^m} \right]$ . Using the fact that a Dyadic cube is not empty together with 19.39 gives  $\frac{z}{2^m} = \frac{w}{2^m}$  proving that  $z = w$ .

**surjectivity.** This follows from the definition of  $\mathcal{D}_m^n$

So  $\mathcal{D}_m^n$  is denumerable.

2. Using 5.60 it follows that  $\mathcal{D}^n = \bigcup_{m \in \mathbb{N}} \mathcal{D}_m^n$  is denumerable

□

We are now finally ready to prove that every open set in  $\mathbb{R}^n$  can be written as a countable union of pairwise disjoint Dyadic cubes.

**Lemma 19.50.** *Let  $n \in \mathbb{N}_0$  and  $\mathcal{T}_{\mathbb{R}^n}$  the canonical topology on  $\mathbb{R}^n$  [see 11.83] then  $\forall U \in \mathcal{T}_{\mathbb{R}^n}$  with  $U \neq \emptyset$  we have that there exists a pairwise disjoint family  $\{D_i\}_{i \in \mathbb{N}} \subseteq \mathcal{D}^n \subseteq \mathcal{R}^n$  such that  $U = \bigcup_{i \in \mathbb{N}} D_i$*

**Proof.** Let  $\emptyset \neq U \in \mathcal{T}_{\mathbb{R}^n}$  be a non empty open set in  $\mathbb{R}^n$ . Given  $m \in \mathbb{N}$  define

$$\mathcal{S}_m = \{D \in \mathcal{D}_m^n \mid D \subseteq U\} \subseteq \mathcal{D}_m^n \quad (19.66)$$

and recursively

$$\mathcal{T}_m = \begin{cases} S_0 \text{ if } m = 0 \\ \{D \in \mathcal{S}_m \mid \forall i \in \{1, \dots, m-1\} \text{ we have } \forall P \in \mathcal{T}_i \text{ that } P \cap D = \emptyset\} \text{ if } m > 0 \end{cases} \quad (19.67)$$

Using the above definitions we have then that

$$\mathcal{T}_m \subseteq \mathcal{S}_m \subseteq \mathcal{D}_m^n \subseteq \mathcal{D}^n \quad (19.68)$$

Finally we define

$$\mathcal{T} = \bigcup_{m \in \mathbb{N}} \mathcal{T}_m \quad (19.69)$$

If  $D \in \mathcal{T}$  then  $\exists i \in \mathbb{N}$  such that  $D \in \mathcal{T}_i \subseteq \mathcal{S}_i$  proving by 19.66 that  $D \subseteq U$  hence

$$\bigcup_{D \in \mathcal{T}} D \subseteq U \quad (19.70)$$

For the opposite inclusion take  $x \in U$  then as  $U$  is open there exists a  $\varepsilon > 0$  such that  $x \in B_{|||}(x, \varepsilon) \subseteq U$ . Using the Archimedean property [see 9.56] there exists a  $n_x \in \mathbb{N}_0$  such that  $0 < \frac{1}{n_x} < \varepsilon$ . Using 19.48 (3) there exists a  $D_x = \prod_{i \in \{1, \dots, n\}} \left[ \frac{z_i}{2^{n_x}}, \frac{z_i + 1}{2^{n_x}} \right] \in \mathcal{D}_{n_x}^n$  such that  $x \in D_x$  proving that

$$\forall i \in \{1, \dots, n\} \text{ we have } \frac{z_i}{2^{n_x}} \leq x_i < \frac{z_i + 1}{2^{n_x}} \quad (19.71)$$

If also  $y \in D_x$  then

$$\forall i \in \{1, \dots, n\} \text{ we have } \frac{z_i}{2^{n_x}} \leq y_i < \frac{z_i + 1}{2^{n_x}} \quad (19.72)$$

Next

$$\begin{aligned}
 y_i - x_i &< \frac{z_i + 1}{2^{n_x}} - x_i \\
 &\leq \frac{z_i + 1}{2^{n_x}} - \frac{z_i}{2^{n_x}} \\
 &= \frac{1}{2^{n_x}} \\
 &< \varepsilon \\
 x_i - y_i &< \frac{z_i + 1}{2^{n_x}} - y_i \\
 &\leq \frac{z_i + 1}{2^{n_x}} - \frac{z_i}{2^{n_x}} \\
 &= \frac{1}{2^{n_x}} \\
 &< \varepsilon
 \end{aligned}$$

proving that  $\forall i \in \{1, \dots, n\} |y_i - x_i| < 0$ . Hence  $\|x - y\| = \max(\{|y_i - x_i| | i \in \{1, \dots, n\}\}) < \varepsilon$  proving that  $D_x \subseteq B_{\|(x, \varepsilon)} \subseteq U$ . Hence

$$x \in D_x \in \mathcal{S}_{n_x} \quad (19.73)$$

Define now

$$\mathcal{K}_x = \{i \in \{0, \dots, n_x\} | \exists D \in \mathcal{S}_i \vdash x \in D\} \quad (19.74)$$

then by 19.73 we have that  $n_x \in \mathcal{K}_x$  so that  $\mathcal{K}_x \neq \emptyset$ . Hence

$$m_x = \min(\mathcal{K}_x) \text{ exists} \quad (19.75)$$

We must now consider two cases for  $m_x$ :

**$m_x = 0$ .** then  $\exists D \in \mathcal{S}_0 \stackrel{19.67}{=} \mathcal{T}_0 \subseteq \mathcal{T}$  so that  $x \in D$

**$0 < m_x$ .** then  $\exists D \in \mathcal{S}_{m_x} \subseteq \mathcal{D}_{m_x}^n$  such that  $x \in D$  and  $\forall i \in \{0, \dots, m_x - 1\}$  we have  $\forall R \in \mathcal{S}_i$  that  $x \notin R$ . Let  $i \in \{0, \dots, m_x - 1\}$  and assume that  $\exists R \in \mathcal{S}_i \subseteq \mathcal{D}_i^n$  such that  $R \cap D \neq \emptyset$  then by 19.48 (2) we have  $D \subseteq R$ . But then as  $x \in D$  we reach the contradiction  $x \in R \wedge x \notin R$ . Hence  $\forall i \in \{0, \dots, m_x - 1\}$  we have  $\forall R \in \mathcal{S}_i$  that  $R \cap D = \emptyset$ . So by the definition 19.67 it follows that  $D \in \mathcal{T}_{m_x} \subseteq \mathcal{T}$ . Hence  $\exists D \in \mathcal{T}$  such that  $x \in D$ .

So in all cases there exists a  $D \in \mathcal{T}$  with  $x \in D$  proving that  $x \in \bigcup_{D \in \mathcal{T}} D$  which as  $x \in U$  was chosen arbitrary gives that  $U \subseteq \bigcup_{D \in \mathcal{T}} D$ . Combining this with 19.70 proves

$$U = \bigcup_{D \in \mathcal{T}} D \quad (19.76)$$

Now let  $D, E \in \mathcal{T}$  with  $D \neq E$  then there exists  $k, l \in \mathbb{N}$  such that  $D \in \mathcal{T}_k \subseteq \mathcal{D}_k^n$  and  $E \in \mathcal{T}_l \subseteq \mathcal{D}_l^n$ . For  $k, l$  we have the following cases to consider:

**$k = l$ .** then using 19.48 (1) we have  $D \cap E = \emptyset$

**$k < l$ .** then using 19.67 we have  $D \cap E = \emptyset$

**$l < k$ .** then using 19.67 we have  $D \cap E = \emptyset$

Combining the above with 19.76 gives

$$U = \bigsqcup_{D \in \mathcal{T}} D \quad (19.77)$$

Assume now that  $\mathcal{T}$  is finite. As  $U \neq \emptyset$  we must have that  $\mathcal{T} \neq \emptyset$  so that by the finiteness of  $\mathcal{T}$   $m = \min((\text{begin}(D))_1 | D \in \mathcal{T})$  is defined. Hence there exists a  $R = [a, b] \in \mathcal{T}$  such that  $a_1 = m$ . As  $a \in R \subseteq U$  a open set there exists a  $\varepsilon > 0$  such that  $a \in B_{\|(a, \varepsilon)} \subseteq U$ . Given that  $a - \frac{\varepsilon}{2} \cdot e_1 \subseteq B_{\|(a, \varepsilon)} \subseteq U$  [where  $(e_1)_i = \begin{cases} 1 & \text{if } i = 1 \\ 0 & \text{if } i \in \{2, \dots, n\} \end{cases}$  so that  $a - (a - \frac{\varepsilon}{2} \cdot e_1) = \begin{cases} \frac{1}{2} \cdot \varepsilon & \text{if } i = 1 \\ 0 & \text{if } i \in \{2, \dots, n\} \end{cases}$ ], there exists a  $D = [a', b'] \in \mathcal{T}$  such that  $a - \frac{\varepsilon}{2} \cdot e_1 \in D$ . But then  $a'_1 = (\text{begin}(D))_1 \leq a_1 - \frac{\varepsilon}{2} < a_1 = m$  contradicting  $m \leq \min((\text{begin}(D))_1 | D \in \mathcal{T}) \leq a'_1 < m$ . So we must have that

$$\mathcal{T} \text{ is infinite} \quad (19.78)$$

As  $\mathcal{T} \subseteq \mathcal{D}^n$  is a denumerable set [see 19.49] it follows from the above and 5.55 it follows that

$$\mathcal{T} \text{ is denumerable} \quad (19.79)$$

Hence there exist a bijection  $\beta: \mathbb{R} \rightarrow \mathcal{T}$ . Define  $\{D_i\}_{i \in \mathbb{N}}$  by  $D_i = \beta(i)$  then we have using 2.68

$$U = \bigsqcup_{i \in \mathbb{N}} D_i$$

proving the lemma.  $\square$

**Lemma 19.51.** *Let  $n \in \mathbb{N}_0$  then  $\forall R \in \mathcal{R}^n$  with  $R \neq \emptyset$  there exists a sequence  $\{U_i\}_{i \in \{1, \dots, n\}} \subseteq \mathcal{T}_{\mathbb{R}^n}$  of open sets such that  $R = \bigcap_{i \in \mathbb{N}} U_i$*

**Proof.** As  $R \in \mathcal{R}^n$  there exists a  $a, b \in \mathbb{R}^n$  with  $a \leq b$  such that  $R = [a, b] = \prod_{i \in \{1, \dots, n\}} [a_i, b_i]$ . Define then  $\{U_i\}_{i \in \mathbb{N}}$  by  $U_i = \prod_{j \in \{1, \dots, n\}} [a_j - \frac{1}{i}, b_j]$ . As  $[a_j - \frac{1}{i}, b_j] \in \mathcal{T}_{\mathbb{R}}$  and  $\mathcal{T}_{\mathbb{R}^n}$  is the product topology based on  $\mathcal{T}_{\mathbb{R}}$  [see 11.83] it follows that  $\{U_i\}_{i \in \mathbb{N}} \subseteq \mathcal{T}_{\mathbb{R}^n}$ . Let  $i \in \mathbb{N}$  and  $x \in \mathbb{R}$  then we have  $\forall j \in \{1, \dots, n\}$  that  $x_j \in [a_j, b_j] \Rightarrow a_j \leq x_j < b_j \Rightarrow a_j - \frac{1}{i} < a_j \leq x_j < b_j$ , proving using 5.103 that  $R \subseteq U_i$ . Hence we have

$$R \subseteq \bigcap_{i \in \mathbb{N}} U_i \quad (19.80)$$

For the opposite inequality let  $x \in \bigcap_{i \in \mathbb{N}} U_i$  then  $\forall i \in \mathbb{N}$  we have  $x \in U_i$  so that  $\forall j \in \{1, \dots, n\}$  we have  $x_j \in [a_j - \frac{1}{i}, b_j] \Rightarrow a_j - \frac{1}{i} < x_j < b_j$ . Assume that  $\exists j \in \{1, \dots, n\}$  such that  $x_j < a_j$  then using 9.56 there exists a  $i \in \mathbb{N}$  such that  $x_j < x_j + \frac{1}{i} < a_j \Rightarrow x_j < a_j - \frac{1}{i}$  contradicting the fact that  $a_j - \frac{1}{i} < x_j$ . Hence  $\forall j \in \{1, \dots, n\}$  we have  $a_j \leq x_j < b_j$  proving that  $x \in R$  so that  $\bigcap_{i \in \mathbb{N}} U_i \subseteq R$ . This together with 19.80 proves

$$R = \bigcap_{i \in \mathbb{N}} U_i \quad \square$$

We are finally ready to proved a alternative definition for  $\mathcal{B}[\mathbb{R}^n]$ .

**Theorem 19.52.** *Let  $n \in \mathbb{N}_n$  then we have*

$$\begin{aligned} \sigma[\mathcal{R}^n] &= \mathcal{B}[\mathbb{R}^n] \text{ [where } \mathcal{B}[\mathbb{R}^n] \stackrel{\text{def}}{=} \mathcal{B}[\mathbb{R}^n, \mathcal{T}_{\mathbb{R}^n}] = \sigma[\mathcal{T}_{\mathbb{R}^n}] \text{]} \\ &\stackrel{19.21}{=} \mathcal{S}[\{C \subseteq \mathbb{R}^n \mid C \text{ is closed in } \mathcal{T}_{\mathbb{R}^n}\}] \end{aligned}$$

**Proof.** Let  $U \in \mathcal{T}_{\mathbb{R}}$  then we have either:

$U = \emptyset$ . then  $U \in \sigma[\mathcal{R}^n]$

$U \neq \emptyset$ . then using 19.50 there exists a  $\{R_i\}_{i \in \mathbb{N}} \subseteq \mathcal{R}^n$  such that  $U = \bigsqcup_{i \in \mathbb{N}} R_i$  which as  $\sigma[\mathcal{R}^n]$  is a  $\sigma$ -algebra proves that  $U \in \sigma[\mathcal{R}^n]$

which proves that  $\mathcal{T}_{\mathbb{R}} \subseteq \sigma[\mathcal{R}^n]$ . Hence we have

$$\sigma[\mathcal{T}_{\mathbb{R}}] \subseteq \sigma[\mathcal{R}^n] \quad (19.81)$$

For the opposite inclusion let  $R \in \mathcal{R}^n$  then by the previous lemma [see 19.51] there exists a  $\{U_i\}_{i \in \mathbb{N}} \subseteq \mathcal{T}_{\mathbb{R}^n}$  such that  $R = \bigcap_{i \in \mathbb{N}} U_i$ . As  $\sigma[\mathcal{T}_{\mathbb{R}^n}]$  is a  $\sigma$ -algebra it follows that  $R \in \sigma[\mathcal{T}_{\mathbb{R}^n}]$  we have that  $\mathcal{R}^n \subseteq \sigma[\mathcal{T}_{\mathbb{R}^n}]$ . Hence  $\sigma[\mathcal{R}^n] \subseteq \sigma[\mathcal{T}_{\mathbb{R}^n}]$  which by 19.81 proves that

$$\sigma[\mathcal{R}^n] = \sigma[\mathcal{T}_{\mathbb{R}^n}] \quad \square$$

**Definition 19.53.** *Let  $n \in \mathbb{N}$  then given  $i \in \{1, \dots, n\}$  and  $a \in \mathbb{R}$   $]-\infty, a[i]$  is defined by  $]-\infty, a[i] = \{x \in \mathbb{R}^n \mid x_i < a\}$ . The set  $]-\infty, a[i]$  is called a lower half open space. The set of all lower half open spaces is noted by  $\mathcal{H}^n$  so  $\mathcal{H}^n = \{]-\infty, a[i] \mid i \in \{1, \dots, n\} \wedge a \in \mathbb{R}^n\}$*

We prove now that  $\mathcal{H}^n$  is a collection of open sets.

**Lemma 19.54.** Let  $n \in \mathbb{N}_0$  then  $\forall H \in \mathcal{H}^n$  we have that  $H \in \mathcal{T}_{\mathbb{R}^n}$  [the canonical topology on  $\mathbb{R}^n$ ]

**Proof.** Let  $H \in \mathcal{H}^n$  then there exists a  $i \in \{1, \dots, n\}$  and a  $a \in \mathbb{R}$  such that  $H = \{x \in \mathbb{R}^n | x_i < a\}$ . Take  $x \in H$  then  $x_i < a$  so that  $\varepsilon = a - x_i > 0$ . Let  $y \in B_{\|\cdot\|_{\max}}(x, \varepsilon)$  then  $y_i - x_i < |y_i - x_i| \leq \max\{|y_i - x_i| | i \in \{1, \dots, n\}\} < \varepsilon = a - x_i$  proving that  $y_i < a$ . Hence  $y \in B_{\|\cdot\|_{\max}}(x, \varepsilon) \subseteq H$  which proves that  $H$  is open.  $\square$

**Definition 19.55.** Let  $n \in \mathbb{N}^n$  then given  $i \in \{1, \dots, n\}$  and  $a \in \mathbb{R}$  we define  $]-\infty, a]_i$  as  $]-\infty, a] = \{x \in \mathbb{R} | x_i \leq a\}$ . The set of half spaces  $\mathcal{H}^n$  is defined as  $\mathcal{H}^n = \{]-\infty, a] | a \in \mathbb{R} \wedge i \in \mathbb{N}_1\}$

Next we show that every non empty set in  $\mathcal{R}^n$  can be written as a finite intersection of sets in  $\sigma[\mathcal{H}^n]$ .

**Lemma 19.56.** Let  $n \in \mathbb{N}_0$  and  $a, b \in \mathbb{R}^n$  with  $a \leq b$  then  $[a, b] = \bigcap_{i \in \{1, \dots, n\}} (]-\infty, b_i] \setminus ]-\infty, a_i[)$

**Proof.** Let  $x \in [a, b]$  then if  $i \in \{1, \dots, n\}$  we have that  $a_i \leq x_i \wedge x_i < b_i$  or  $\neg(x_i < a_i) \wedge x_i < b_i$ , hence  $x \in ]-\infty, b_i] \wedge x \notin ]-\infty, b_i[$ . So  $\forall i \in \{1, \dots, n\}$  we have  $x \in (]-\infty, b_i] \setminus ]-\infty, a_i[)$  proving that

$$[a, b] \subseteq \bigcap_{i \in \{1, \dots, n\}} (]-\infty, b_i] \setminus ]-\infty, a_i[) \quad (19.82)$$

For the opposite inclusion let  $x \in \bigcap_{i \in \{1, \dots, n\}} (]-\infty, b_i] \setminus ]-\infty, a_i[)$  then  $\forall i \in \{1, \dots, n\}$  we have  $x \in ]-\infty, b_i] \wedge x \notin ]-\infty, b_i[$  hence  $x_i < b_i \wedge \neg(x_i < a_i) \Rightarrow x_i < b_i \wedge a_i \leq x_i$  proving that  $x \in [a, b]$ . So  $\bigcap_{i \in \{1, \dots, n\}} (]-\infty, b_i] \setminus ]-\infty, a_i[) \subseteq [a, b]$  which together with 19.82 proves

$$[a, b] = \bigcap_{i \in \{1, \dots, n\}} (]-\infty, b_i] \setminus ]-\infty, a_i[) \quad \square$$

**Lemma 19.57.** let  $n \in \mathbb{N}_0$  then we have that  $\mathcal{B}[\mathbb{R}^n] = \sigma[\mathcal{H}^n]$

**Proof.** If  $H \in \mathcal{H}^n$  then  $H$  is open [see 19.54], so that  $H \in \sigma[\mathcal{T}_{\mathbb{R}^n}] \stackrel{19.52}{=} \mathcal{B}[\mathbb{R}^n]$ . Hence  $\mathcal{H}^n \subseteq \mathcal{B}[\mathbb{R}^n]$  giving

$$\sigma[\mathcal{H}^n] \subseteq \mathcal{B}[\mathbb{R}^n] \quad (19.83)$$

Let  $R \in \mathcal{R}^n$  then  $\exists a, b \in \mathbb{R}^n$  with  $a \leq b$  such that  $R = [a, b]$ , using the previous lemma [see 19.56] we have then that  $R = [a, b] = \bigcap_{i \in \{1, \dots, n\}} (]-\infty, b_i] \setminus ]-\infty, a_i[) \in \sigma[\mathcal{H}^n]$ . So  $\sigma[\mathcal{R}^n] \subseteq \sigma[\mathcal{H}^n]$  which as  $\sigma[\mathcal{R}] \stackrel{19.52}{=} \mathcal{B}[\mathbb{R}^n]$  proves that  $\mathcal{B}[\mathbb{R}^n] \subseteq \sigma[\mathcal{H}^n]$ . Combining this with 19.83 gives

$$\sigma[\mathcal{H}^n] = \mathcal{B}[\mathbb{R}^n] \quad \square$$

We can summarize the above equivalences of the Borel algebra on  $\mathbb{R}^n$  [see 19.52 and 19.57] in the following theorem.

**Theorem 19.58.**  $\forall n \in \mathbb{N}_0$  we have

$$\begin{aligned} \mathcal{B}[\mathbb{R}^n] &= \sigma[\{C \subseteq \mathbb{R}^n | C \text{ is closed in } \mathcal{T}_{\mathbb{R}^n}\}] \\ &= \sigma[\mathcal{R}^n] \\ &= \sigma[\mathcal{H}^n] \end{aligned}$$

The following relation between open and closed sets will be usefull later

**Proposition 19.59.** Let  $n \in \mathbb{N}_0$  and consider the canonical topological space  $\langle \mathbb{R}^n, \mathcal{T}_{\mathbb{R}^n} \rangle$  based on the maximum norm then we have

1. If  $F$  is a closed set then  $\exists \{U_i\}_{i \in \mathbb{N}_0}$  of open sets such that  $F = \bigcap_{i \in \mathbb{N}_0} U_i$
2. If  $U$  is a open set then  $\exists (F_i)_{i \in \mathbb{N}_0}$  of closed sets such that  $U = \bigcup_{i \in \mathbb{N}_0} F_i$

### Proof.

1. Let  $F$  be closed. Given  $x \in F$  define the open set  $U_{n,x}$  by  $U_{n,x} = B_{\|\cdot\|_{\max}}(x, \frac{1}{n})$  and define then the open set  $U_n = \bigcup_{x \in F} U_{n,x}$ , then trivially we have that  $F \subseteq U_n$ . Hence

$$F \subseteq \bigcap_{n \in \mathbb{N}_0} U_n \quad (19.84)$$

Assume that  $x \in \bigcap_{n \in \mathbb{N}_0} U_n$  then  $\forall n \in \mathbb{N}_0$  we have that  $x \in U_n = \bigcup_{y \in F} U_{n,y}$  so that there exists a  $f_n \in F$  such that  $x \in U_{n,f_n}$  or  $\|x - f_n\|_{\max} < \frac{1}{n}$ . Hence if we take  $\{f_n\}_{n \in \mathbb{N}_0} \subseteq F$  then if  $\varepsilon > 0$  we have [see 9.56] that there exists a  $M \in \mathbb{N}_0$  such that  $\frac{1}{M} < \varepsilon$ . So if  $n \geq M$  then  $\|x - f_n\|_{\max} \leq \frac{1}{M} < \varepsilon$  which proves that  $\lim_{n \rightarrow \infty} f_n = x$ . Hence using 11.324 we have that  $x \in F$  which proves that  $\bigcap_{n \in \mathbb{N}_0} U_n \subseteq F$ , combine this with 19.84 gives

$$F = \bigcap_{n \in \mathbb{N}_0} U_n$$

2. Let  $U$  be a open set then  $X \setminus U$  is closed and using (1) there exists a family  $\{U_i\}_{i \in \mathbb{N}_0}$  of open sets such that  $X \setminus U = \bigcup_{i \in \mathbb{N}_0} U_i$ . Hence

$$\begin{aligned} U &\stackrel{1.19}{=} X \setminus (X \setminus U) \\ &= X \setminus \left( \bigcup_{i \in \mathbb{N}_0} U_i \right) \\ &\stackrel{1.100}{=} \bigcap_{i \in \mathbb{N}_0} (X \setminus U_i) \\ &= \bigcap_{i \in \mathbb{N}_0} F_i \text{ where } F_i = X \setminus U_i \text{ are closed sets} \end{aligned}$$

□

## 19.3 Measures

### 19.3.1 Measure definition and properties

**Definition 19.60.** Let  $\langle X, \mathcal{A} \rangle$  be a measurable space [see 19.13] and  $\mu: \mathcal{A} \rightarrow [0, \infty]$  a function then we say that:

1.  $\mu$  is **finite additive** if for all pairwise disjoint  $\{A_i\}_{i \in \{1, \dots, n\}} \subseteq \mathcal{A}$ ,  $n \in \mathbb{N}_0$  we have that

$$\mu \left( \bigsqcup_{i \in \{1, \dots, n\}} A_i \right) = \sum_{i=1}^n \mu(A_i)$$

2.  $\mu$  is **finite sub additive** if for all  $\{A_i\}_{i \in \{1, \dots, n\}} \subseteq \mathcal{A}$ ,  $n \in \mathbb{N}_0$  we have that

$$\mu \left( \bigcup_{i \in \{1, \dots, n\}} A_i \right) \leq \sum_{i=1}^n \mu(A_i)$$

3.  $\mu$  is **countable additive** if for all pairwise disjoint  $\{A_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{A}$ ,  $n \in \mathbb{N}_0$  we have that

$$\mu \left( \bigsqcup_{i \in \mathbb{N}_0} A_i \right) = \sum_{i=1}^{\infty} \mu(A_i)$$

4.  $\mu$  is **countable sub additive** if for all  $\{A_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{A}$  we have that

$$\mu \left( \bigcup_{i \in \mathbb{N}_0} A_i \right) \leq \sum_{i=1}^{\infty} \mu(A_i)$$

**Definition 19.61.** Let  $\langle X, \mathcal{A} \rangle$  be a measurable space and  $\mu: \mathcal{A} \rightarrow [0, \infty]$  a function then we say that  $\mu: \mathcal{A} \rightarrow [0, \infty]$  is a **measure** on  $\mathcal{A}$  if and only if

1.  $\mu(\emptyset) = 0$
2. For every pairwise disjoint  $\{A_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{A}$  we have

$$\mu\left(\bigsqcup_{i \in \mathbb{N}_0} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i)$$

[or  $\mu$  is countable additive]

A triple  $\langle X, \mathcal{A}, \mu \rangle$  where  $\langle X, \mathcal{A} \rangle$  is a measurable space and  $\mu: \mathcal{A} \rightarrow [0, \infty]$  a **measure** is called a **measure space**.

**Proposition 19.62.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space and let  $\mathcal{B}$  be a  $\sigma$ -algebra on  $X$  such that  $\mathcal{B} \subseteq \mathcal{A}$  then  $\langle X, \mathcal{B}, \mu|_{\mathcal{B}} \rangle$  is a measure space

**Proof.** As  $\mathcal{B}$  is already a  $\sigma$ -algebra on  $X$  we only have to proof that  $\mu|_{\mathcal{B}}$  is a measure on  $\mathcal{B}$ .

1. As  $\emptyset \in \mathcal{B} \subseteq \mathcal{A}$  we have  $\mu|_{\mathcal{B}}(\emptyset) = \mu(\emptyset) = 0$
2. If  $\{A_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{B} \subseteq \mathcal{A}$  is a pairwise disjoint family then  $\mu|_{\mathcal{B}}(\bigsqcup_{i \in \mathbb{N}_0} A_i) = \mu(\bigsqcup_{i \in \mathbb{N}_0} A_i) = \sum_{i=1}^{\infty} \mu(A_i) = \sum_{i=1}^{\infty} \mu|_{\mathcal{B}}(A_i)$   $\square$

**Note 19.63.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measurable space then if  $\mu: \mathcal{A} \rightarrow [0, \infty]$  is countable additive then  $\mu$  is additive.

**Proof.** Let  $n \in \mathbb{N}_0$ ,  $\{A_i\}_{i \in \{1, \dots, n\}} \subseteq \mathcal{A}$  such that  $\forall i, j \in \{1, \dots, n\}$  with  $i \neq j$  we have  $A_i \cap A_j = \emptyset$ . Define now  $\{B_i\}_{i \in \mathbb{N}_0}$  by  $B_i = \begin{cases} A_i & \text{if } i \in \{1, \dots, n\} \\ \emptyset & \text{if } i \in \mathbb{N}_0 \setminus \{1, \dots, n\} \end{cases}$ . If  $k, l \in \mathbb{N}_0$  with  $k \neq l$  then we have either

$k, l \in \{1, \dots, n\}$ . then  $B_k \cap B_l = A_k \cap A_l = \emptyset$

$k \notin \{1, \dots, n\}$ . then  $B_k \cap B_l = \emptyset \cap B_l = \emptyset$

$l \notin \{1, \dots, n\}$ . then  $B_k \cap B_l = B_k \cap \emptyset = \emptyset$

So we have by countable additivity that

$$\begin{aligned} \mu\left(\bigsqcup_{i \in \{1, \dots, n\}} A_i\right) &\stackrel{1.103}{=} \mu\left(\bigcup_{i \in \mathbb{N}_0} B_i\right) \\ &= \sum_{i=1}^{\infty} \mu(B_i) \\ &\stackrel{17.99}{=} \sum_{i=1}^n \mu(B_i) \\ &= \sum_{i=1}^n \mu(A_i) \end{aligned}$$

proving the note.  $\square$

**Theorem 19.64.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space then

1.  $\forall A, B \in \mathcal{A}$  with  $A \subseteq B$  we have  $\mu(A) \leq \mu(B)$
2.  $\forall A, B \in \mathcal{A}$  with  $A \subseteq B$  and  $\mu(A) < \infty$  we have  $\mu(B \setminus A) = \mu(B) - \mu(A)$

**Proof.**

1. As  $A \cap (B \setminus A) = \emptyset$  and  $B = A \cup (B \setminus A)$  it follows by the above [see 19.63] that

$$\mu(B \setminus A) + \mu(A) = \mu(B)$$

further as  $0 \leq \mu(B \setminus A)$  we have by 17.27 that  $\mu(A) \leq \mu(B)$

2. If in addition to (1)  $\mu(A)$  is finite it follows from 17.24 that

$$\mu(B \setminus A) = \mu(B) - \mu(A) \quad \square$$

Next we prove that countable additivity implies countable sub additivity.

**Lemma 19.65.** Let  $\langle X, \mathcal{A} \rangle$  be a measurable space then if  $\{A_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{A}$  then there exists a pairwise disjoint  $\{B_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{A}$  such that  $\forall i \in \mathbb{N}_0 B_i \subseteq A_i$  and  $\bigcup_{i \in \mathbb{N}_0} A_i = \bigsqcup_{i \in \mathbb{N}_0} B_i$

**Proof.** Define  $\{B_i\}_{i \in \mathbb{N}_0}$  by  $B_i = \begin{cases} A_1 & \text{if } i = 1 \\ X_i \setminus \bigcup_{j \in \{1, \dots, i-1\}} A_j & \text{if } i \in \mathbb{N}_0 \setminus \{1\} \end{cases}$  then by the properties of a  $\sigma$ -algebra it follows that  $\{B_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{A}$ . Also using 19.6 we have that  $\bigcup_{i \in \mathbb{N}_0} A_i = \bigsqcup_{i \in \mathbb{N}_0} B_i$  and  $\forall i \in \mathbb{N}_0$  we have  $B_i \subseteq A_i$ .  $\square$

**Theorem 19.66.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space then for every  $\{A_i\}_{i \in \mathbb{N}_0}$  we have that  $\mu(\bigcup_{i \in \mathbb{N}_0} A_i) \leq \sum_{i=1}^{\infty} \mu(A_i)$ .

**Proof.** Using 19.65 there exists  $\{B_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{A}$  such that  $\bigcup_{i \in \mathbb{N}_0} A_i = \bigsqcup_{i \in \mathbb{N}_0} B_i$  and  $\forall i \in \mathbb{N}_0 B_i \subseteq A_i$ . So

$$\begin{aligned} \mu\left(\bigcup_{i \in \mathbb{N}_0} A_i\right) &= \mu\left(\bigsqcup_{i \in \mathbb{N}_0} B_i\right) \\ &= \sum_{i=1}^{\infty} \mu(B_i) \\ &\leq \text{17.97 and 19.64} \sum_{i=1}^{\infty} \mu(A_i) \end{aligned}$$

$\square$

**Theorem 19.67.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space then

1.  $\forall \{A_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{A}$  with  $\forall i \in \mathbb{N}_0 A_i \subseteq A_{i+1}$  we have  $\mu(\bigcup_{i \in \mathbb{N}_0} A_i) = \lim_{i \rightarrow \infty} \mu(A_i) = \sup(\{\mu(A_i) | i \in \mathbb{N}_0\})$
2.  $\forall \{A_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{A}$  with  $\forall i \in \mathbb{N}_0 A_{i+1} \subseteq A_i$  and  $\exists k \in \{1, \dots, n\}$  such that  $\mu(A_k) < \infty$  we have that  $\mu(\bigcap_{i \in \mathbb{N}_0} A_i) = \lim_{i \rightarrow \infty} \mu(A_i)$  [or as  $\mu(A_{i+1}) \leq \mu(A_i)$  we have by 17.71 that  $\mu(\bigcap_{i \in \mathbb{N}_0} A_i) = \inf(\{\mu(A_i) | i \in \mathbb{N}_0\})$ ]

**Proof.**

1. Define  $\{B_i\}_{i \in \mathbb{N}_0}$  by  $B_i = \begin{cases} A_1 & \text{if } i = 1 \\ A_i \setminus A_{i-1} & \text{if } i \in \mathbb{N}_0 \setminus \{1\} \end{cases} \in \mathcal{A}$ . Let  $i, j \in \mathbb{N}_0$  with  $i \neq j$  then we may always assume that  $i < j$  [otherwise exchange  $i$  and  $j$ ]. If  $x \in B_i \cap B_j$  then we have either   
**i = 1.** then  $x \in B_1 = A_1$  and  $x \in B_j = A_j \setminus A_{j-1} \Rightarrow x \notin A_{j-1}$ . As  $i < j \Rightarrow 1 \leq j-1$  we have that  $A_1 \subseteq A_{j-1}$  so that we reach the contradiction that  $x \in A_1$  and  $x \notin A_1$ .   
**1 < i.** then  $i < j \Rightarrow i \leq j-1$  so that  $A_i \subseteq A_{j-1}$ . Now as  $x \in B_i = A_i \setminus A_{i-1}$  we have that  $x \in A_i \subseteq A_{j-1}$  which contradict that  $x \in B_j = A_j \setminus A_{j-1}$ .

so we conclude that

$$\forall i, j \in \mathbb{N}_0 \text{ with } i \neq j \text{ we have } B_i \cap B_j = \emptyset \quad (19.85)$$

Next if  $n \in \mathbb{N}_0$  we have  $\forall i \in \{1, \dots, n\}$  that  $B_i = \begin{cases} A_1 \subseteq A_1 & \text{if } i = 1 \\ A_i \setminus A_{i-1} \subseteq A_i & \text{if } i \in \mathbb{N}_0 \setminus \{1\} \end{cases} \subseteq A_i \subseteq A_n$  it follows that

$$\bigcup_{i \in \{1, \dots, n\}} B_i \subseteq A_n \wedge \bigcup_{i \in \mathbb{N}_0} B_i \subseteq \bigcup_{i \in \mathbb{N}_0} A_i \quad (19.86)$$

Further if  $x \in A_n$  then  $n \in \{i \in \{1, \dots, n\} | x \in A_i\}$  so that  $m = \min(\{i \in \{1, \dots, n\} | x \in A_i\})$  exists. We have now two cases to consider for  $m$ :

**$m = 1$ .** then  $x \in A_1 = B_1 \subseteq \bigcup_{i \in \{1, \dots, n\}} B_i$

**$1 < m$ .** then  $x \notin A_{m-1}$  and  $x \in A_m$  so that  $x \in A_m \setminus A_{m-1} = B_m \subseteq \bigcup_{i \in \{1, \dots, n\}} B_i$

proving that  $A_n \subseteq \bigcup_{i \in \{1, \dots, n\}} B_i$  which by 19.86 gives

$$\forall n \in \mathbb{N}_0 \text{ we have } A_n = \bigcup_{i \in \{1, \dots, n\}} B_i \quad (19.87)$$

From the above it follows that  $A_n \subseteq \bigcup_{i \in \{1, \dots, n\}} B_i \subseteq \bigcup_{i \in \mathbb{N}_0} B_i$  which by 19.85 and 19.86 gives

$$\bigcup_{i \in \mathbb{N}_0} A_i = \bigcup_{i \in \mathbb{N}_0} B_i \quad (19.88)$$

Next we have

$$\begin{aligned} \mu\left(\bigcup_{i \in \mathbb{N}_0} A_i\right) &= \mu\left(\bigcup_{i \in \mathbb{N}_0} B_i\right) \\ \text{μ is a measure} &\quad \sum_{i=1}^{\infty} \mu(B_i) \\ &\stackrel{\text{def}}{=} \lim_{n \rightarrow \infty} \left( \sum_{i=1}^n \mu(B_i) \right) \\ \text{μ is additive by 19.63} &\quad \lim_{n \rightarrow \infty} \left( \mu\left(\bigcup_{i \in \{1, \dots, n\}} B_i\right) \right) \\ &\stackrel{19.87}{=} \lim_{n \rightarrow \infty} \mu(A_n) \end{aligned}$$

proving (1). Finally as  $\forall i \in \mathbb{N}_0$  we have  $A_i \subseteq A_{i+1}$  it follows that  $\mu(A_i) \leq \mu(A_{i+1})$ . Applying then 17.71 gives

$$\mu\left(\bigcup_{i \in \mathbb{N}_0} A_i\right) = \lim_{i \rightarrow \infty} \mu(A_i) = \sup(\{\mu(A_i) | i \in \mathbb{N}_0\})$$

2. First we consider the case where  $k = 1$ . So let  $\mu(A_1) < \infty$  then as  $\forall i \in \mathbb{N}_0$  we have  $1 \leq i$  it follows that  $A_i \subseteq A_1 \Rightarrow \mu(A_i) \leq \mu(A_1) < \infty$  proving that

$$\forall i \in \mathbb{N}_0 \text{ we have } \mu(A_i) < \infty \quad (19.89)$$

Define now  $\{B_i\}_{i \in \mathbb{N}_0}$  by  $B_i = A_1 \setminus A_i$  then we have if  $x \in B_i$  that  $x \in A_1 \wedge x \notin A_i \supseteq A_{i+1} \Rightarrow x \notin A_{i+1}$  proving that  $x \in B_{i+1}$  hence

$$\forall i \in \mathbb{N}_0 \text{ we have } B_i \subseteq B_{i+1} \quad (19.90)$$

So we may use (1) giving

$$\mu\left(\bigcup_{i \in \mathbb{N}_0} B_i\right) = \lim_{i \rightarrow \infty} \mu(B_i) \quad (19.91)$$

Now

$$\begin{aligned} A_1 \setminus \left( \bigcup_{i \in \mathbb{N}_0} B_i \right) &\stackrel{1.102}{=} \bigcap_{i \in \mathbb{N}_0} (A_1 \setminus B_i) \\ &= \bigcap_{i \in \mathbb{N}_0} (A_1 \setminus (A_1 \setminus A_i)) \\ &\stackrel{1.19}{=} \bigcap_{i \in \mathbb{N}_0} A_i \end{aligned} \quad (19.92)$$

As  $\forall i \in \mathbb{N}_0 B_i = A_1 \setminus A_i \subseteq A_1$  we have that  $\mu(B_i) \leq \mu(A_1) < \infty$  and  $\bigcup_{i \in \mathbb{N}_0} B_i \subseteq A_1$  giving  $\mu(\bigcup_{i \in \mathbb{N}_0} B_i) \leq \mu(A_1) < \infty$ . Then

$$\begin{aligned} \mu\left(\bigcap_{i \in \mathbb{N}_0} A_i\right) &\stackrel{19.92}{=} \mu\left(A_1 \setminus \left(\bigcup_{i \in \mathbb{N}_0} B_i\right)\right) \\ &\stackrel{19.64}{=} \mu(A_1) - \mu\left(\bigcup_{i \in \mathbb{N}_0} B_i\right) \\ &\stackrel{19.91}{=} \mu(A_1) - \lim_{i \rightarrow \infty} \mu(B_i) \\ &= \mu(A_1) - \lim_{i \rightarrow \infty} (\mu(A_1) - \mu(A_i)) \\ &\stackrel{17.69 \text{ and } 17.77}{=} \mu(A_1) - \mu(A_1) + \lim_{i \rightarrow \infty} \mu(A_i) \\ &= \lim_{i \rightarrow \infty} \mu(A_i) \end{aligned}$$

Hence we have

$$\begin{aligned} \forall \{A_i\}_{i \in \mathbb{N}_0} \text{ with } \mu(A_1) < \infty \text{ and } \forall i \in \mathbb{N}_0 A_{i+1} \subseteq A_i \text{ we have } \mu\left(\bigcap_{i \in \mathbb{N}_0} A_i\right) = \\ \lim_{i \rightarrow \infty} \mu(A_i) \end{aligned} \tag{19.93}$$

We consider the general case  $1 < k$ . Let now  $\{C_i\}_{i \in \mathbb{N}_0}$  be such that  $\forall i \in C_{i+1} \subseteq C_i$  and  $\exists k \in \mathbb{N}_0 \setminus \{1\}$  such that  $\mu(C_k) < \infty$ . Define then  $\{A_i\}_{i \in \mathbb{N}_0}$  by  $A_i = C_{i+(k-1)}$  then  $\mu(A_1) = \mu(C_{1+(k-1)}) = \mu(C_k) < \infty$  and  $\forall i \in \mathbb{N}_0$  we have  $A_{i+1} = C_{(i+(k-1))+1} \subseteq C_{i+(k-1)} = A_i$ . Hence using 19.93 we have  $\mu(\bigcap_{i \in \mathbb{N}_0} A_i) = \lim_{i \rightarrow \infty} \mu(A_i) = \lim_{i \rightarrow \infty} \mu(C_{i+(k-1)}) \stackrel{17.68}{=} \lim_{i \rightarrow \infty} \mu(C_i)$  proving that

$$\mu\left(\bigcap_{i \in \mathbb{N}_0} A_i\right) = \lim_{i \rightarrow \infty} \mu(C_i) \tag{19.94}$$

Further if  $x \in \bigcap_{i \in \mathbb{N}_0} C_i$  then  $\forall i \in \mathbb{N}_0$  we have as  $i + (k-1) \in \mathbb{N}_0$  that  $x \in C_{i+(k-1)} = A_i$  proving that  $x \in \bigcap_{i \in \mathbb{N}_0} A_i$  hence

$$\bigcap_{i \in \mathbb{N}_0} C_i \subseteq \bigcap_{i \in \mathbb{N}_0} A_i \tag{19.95}$$

If now  $x \in \bigcap_{i \in \mathbb{N}_0} A_i$  then  $\forall i \in \mathbb{N}_0$  we have  $x \in A_i$ , then for  $j \in \mathbb{N}_0$  we have either

$j \in \{k, \dots, \infty\}$ . then as  $j - (k-1) \in \mathbb{N}_0$  we have  $x \in A_{j-(k-1)} = C_{j-(k-1)+(k-1)} = C_j$   
[so also  $x \in C_k$ ]

$j \in \{1, \dots, k-1\}$ . then as  $j < k$  we have  $C_k \subseteq C_j \xrightarrow{x \in C_k} x \in C_j$

hence  $x \in \bigcap_{j \in \mathbb{N}_0} C_j$ . Combining this with 19.95 gives  $\bigcap_{i \in \mathbb{N}_0} C_i = \bigcap_{i \in \mathbb{N}_0} A_i$ . Finally using this on 19.94 we have

$$\mu\left(\bigcap_{j \in \mathbb{N}_0} C_j\right) = \lim_{j \rightarrow \infty} \mu(C_j)$$

proving the theorem for the case  $1 < k$ .

□

We have the following inverse of the above theorem that provides methods to check if set function is a measure.

**Theorem 19.68.** *Let  $\langle X, \mathcal{A} \rangle$  be a measurable space and  $\mu: \mathcal{A} \rightarrow [0, \infty]$  be such that*

1.  $\mu(\emptyset)$

2. If  $\{A_i\}_{i \in \{1, \dots, n\}} \subseteq \mathcal{A}$  is such that  $\forall i, j \in \{1, \dots, n\}$  we have  $i \neq j \Rightarrow A_i \cap A_j = \emptyset$  then  $\mu(\bigcup_{i \in \{1, \dots, n\}} A_i) = \sum_{i=1}^n \mu(A_i)$ .

then we have

1.  $\mu$  is a measure if for every family  $\{A_i\}_{i \in \mathbb{N}} \subseteq \mathcal{A}$  with  $\forall i \in \mathbb{N} \Rightarrow A_i \subseteq A_{i+1}$  we have  $\mu(\bigcup_{i \in \mathbb{N}} A_i) = \lim_{i \rightarrow \infty} \mu(A_i)$ .
2.  $\mu$  is a measure if for every family  $\{A_i\}_{i \in \mathbb{N}} \subseteq \mathcal{A}$  with  $\forall i \in \mathbb{N} \Rightarrow A_{i+1} \subseteq A_i$  and  $\bigcap_{i \in \mathbb{N}} A_i = \emptyset$  we have  $\lim_{i \rightarrow \infty} \mu(A_i) = 0$ .

**Proof.**

1. Let  $\{A_i\}_{i \in \mathbb{N}} \subseteq \mathcal{A}$  such that  $\forall i, j \in \mathbb{N}$  with  $i \neq j$  we have  $A_i \cap A_j = \emptyset$  and define  $\{B_i\}_{i \in \mathbb{N}}$  by  $B_i = \bigcup_{j \in \{1, \dots, i\}} A_j$ . Using 19.5 we have

$$\bigcup_{i \in \mathbb{N}} A_i = \bigcup_{i \in \mathbb{N}} B_i \text{ and } \forall i \in \mathbb{N} \Rightarrow B_i \subseteq B_{i+1} \quad (19.96)$$

Next

$$\begin{aligned} \mu\left(\bigcup_{i \in \mathbb{N}} A_i\right) &\stackrel{19.96}{=} \mu\left(\bigcup_{i \in \mathbb{N}} B_i\right) \\ &\stackrel{\text{assumption}}{=} \lim_{i \rightarrow \infty} \mu(B_i) \\ &= \lim_{i \rightarrow \infty} \mu\left(\bigcup_{j \in \{1, \dots, i\}} A_j\right) \\ &\stackrel{\text{finite additivity}}{=} \lim_{i \rightarrow \infty} \left(\sum_{j=1}^i \mu(A_j)\right) \\ &= \sum_{j=1}^{\infty} \mu(A_j) \end{aligned}$$

proving that  $\mu$  is countable additive and thus that  $\mu$  is a measure.

2. Let  $\{A_i\}_{i \in \mathbb{N}} \subseteq \mathcal{A}$  be such that  $\forall i, j \in \mathbb{N}$  with  $i \neq j$  we have  $A_i \cap A_j = \emptyset$  and define  $\{B_i\}_{i \in \mathbb{N}}$  by  $B_i = \bigcup_{j \in \{i, \dots, \infty\}} A_j$ . Then  $\forall i \in \mathbb{N}$  we have  $B_{i+1} = \bigcup_{j \in \{i+1, \dots, \infty\}} A_j \stackrel{1.99}{\subseteq} \bigcup_{j \in \{i, \dots, \infty\}} A_j = B_i$ . Further if  $x \in \bigcap_{i \in \mathbb{N}} B_i$  then given  $i \in \mathbb{N}$  we have that  $x \in B_i \Rightarrow \exists j \in \{i, \dots, \infty\}$  such that  $x \in A_j$  but then  $\forall k \in \{j+1, \dots, \infty\} \Rightarrow k \neq j$  so that  $x \notin A_k$  hence  $x \notin B_{j+1}$  contradicting that  $x \in \bigcap_{i \in \mathbb{N}} B_i$ . To summarize we have

$$\forall i \in \mathbb{N} \text{ that } B_{i+1} \subseteq B_i \text{ and } \bigcap_{i \in \mathbb{N}} B_i = \emptyset$$

Using the hypotheses we have then that  $\lim_{i \rightarrow \infty} \mu(B_i) = 0$  or using 17.69

$$\lim_{i \rightarrow \infty} \mu(B_{i+1}) = 0 \quad (19.97)$$

Let  $k \in \mathbb{N}$ . If  $x \in (\bigcup_{i \in \{1, \dots, k\}} A_i) \cap (\bigcup_{i \in \{k+1, \dots, \infty\}} A_i)$  then  $\exists i \in \{1, \dots, k\}$  and  $\exists j \in \{k+1, \dots, \infty\}$  such that  $x \in A_i \wedge x \in A_j \Rightarrow x \in A_i \cap A_j$ . As  $i \leq k < k+1 \leq j$  we have  $i \neq j$  giving the contradiction  $A_i \cap A_j = \emptyset$ . Hence

$$\left(\bigcup_{i \in \{1, \dots, k\}} A_i\right) \cap \left(\bigcup_{i \in \{k+1, \dots, \infty\}} A_i\right) = \emptyset \quad (19.98)$$

and

$$\bigcup_{i \in \mathbb{N}} A_i \stackrel{1.99}{=} \left(\bigcup_{i \in \{1, \dots, k\}} A_i\right) \cup \left(\bigcup_{i \in \{k+1, \dots, \infty\}} A_i\right) \quad (19.99)$$

So as  $\mu$  is finite additive we have

$$\begin{aligned}
 \mu\left(\bigcup_{i \in \mathbb{N}} A_i\right) &= \mu\left(\bigcup_{i \in \{1, \dots, k\}} A_i\right) + \mu\left(\bigcup_{i \in \{k+1, \dots, \infty\}} A_i\right) \\
 &= \mu\left(\bigcup_{i \in \{1, \dots, k\}} A_i\right) + \mu(B_{k+1}) \\
 &= \left(\sum_{i=1}^k \mu(A_i)\right) + \mu(B_{k+1})
 \end{aligned} \tag{19.100}$$

So that

$$\begin{aligned}
 \sum_{i=1}^{\infty} \mu(A_i) &\stackrel{19.97}{=} \left(\sum_{i=1}^{\infty} \mu(A_i)\right) + \lim_{k \rightarrow \infty} \mu(B_{k+1}) \\
 &= \lim_{k \rightarrow \infty} \left(\sum_{i=1}^k \mu(A_i)\right) + \lim_{k \rightarrow \infty} \mu(B_{k+1}) \\
 &\stackrel{17.77}{=} \lim_{k \rightarrow \infty} \left(\sum_{i=1}^k \mu(A_i) + \mu(B_{k+1})\right) \\
 &\stackrel{19.100}{=} \lim_{k \rightarrow \infty} \mu\left(\bigsqcup_{i \in \mathbb{N}} A_i\right) \\
 &\stackrel{17.67}{=} \mu\left(\bigsqcup_{i \in \mathbb{N}} A_i\right)
 \end{aligned}$$

□

### 19.3.2 Carathéodory theorem

The Carathéodory theorem allows us to generate a  $\sigma$ -algebra and a measure on this  $\sigma$ -algebra. First we define the concept of a outer measure that has enough properties to use as a basis to generate a  $\sigma$ -algebra and measure.

**Definition 19.69.** Let  $X$  be a set then  $\varphi: \mathcal{P}(X) \rightarrow [0, \infty]$  is a **outer measure** if

1.  $\varphi(\emptyset) = 0$
2. If  $A, B \in \mathcal{P}(X)$  with  $A \subseteq B$  we have  $\varphi(A) \leq \varphi(B)$
3. If  $\{A_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{P}(X)$  then  $\varphi(\bigcup_{i \in \mathbb{N}_0} A_i) \leq \sum_{i=1}^{\infty} \mu(A_i)$  (**semi-additivity**)

**Lemma 19.70.** Let  $X$  be a set and  $\varphi: \mathcal{P}(X) \rightarrow [0, \infty]$  a outer measure then  $\forall \{A_i\}_{i \in \{1, \dots, n\}}, n \in \mathbb{N}_0$  we have  $\varphi(\bigcup_{i \in \{1, \dots, n\}} A_i) \leq \sum_{i=1}^n \varphi(A_i)$

**Proof.** Define  $\{B_i\}_{i \in \mathbb{N}}$  by  $B_i = \begin{cases} A_i & \text{if } i \in \{1, \dots, n\} \\ \emptyset & \text{if } i \in \mathbb{N}_0 \setminus \{1, \dots, n\} \end{cases}$  then we have

$$\begin{aligned}
 \varphi\left(\bigcup_{i \in \{1, \dots, n\}} A_i\right) &\stackrel{1.103}{=} \varphi\left(\bigcup_{i \in \mathbb{N}_0} B_i\right) \\
 &\leq \sum_{i=1}^{\infty} \mu(B_i) \\
 &\stackrel{17.99}{=} \sum_{i=1}^n \mu(B_i) \\
 &= \sum_{i=1}^n \varphi(A_i)
 \end{aligned}$$

□

We show now a way of generating a outer measure based on semi-additive functions

**Theorem 19.71.** *Let  $X$  be a set,  $\mathcal{A} \subseteq \mathcal{P}(X)$  and  $\mu: \mathcal{A} \rightarrow [0, \infty]$  such that*

1.  $\emptyset \in \mathcal{A}$
2.  $\exists \{A_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{A}$  such that  $X \subseteq \bigcup_{i \in \mathbb{N}_0} A_i$
3.  $\mu(\emptyset) = 0$
4. If  $A \in \mathcal{A}$  then  $\forall \{A_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{A}$  with  $A \subseteq \bigcup_{i \in \mathbb{N}_0} A_i$  we have  $\mu(A) \leq \sum_{i=1}^{\infty} \mu(A_i)$

then if we define

$$\varphi: \mathcal{P}(X) \rightarrow [0, \infty] \text{ by } \varphi(A) = \inf \left( \left\{ \sum_{i=1}^{\infty} \mu(A_i) \mid \{A_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{A} \text{ such that } A \subseteq \bigcup_{i \in \mathbb{N}} A_i \right\} \right)$$

we have that

- $\varphi$  is a outer measure
- $\forall A \in \mathcal{A}$  we have  $\varphi(A) = \mu(A)$

**Proof.** First we must prove that  $\varphi$  is well defined. Define

$$\forall A \in \mathcal{P}(X) \quad \mathcal{C}_A = \left\{ \sum_{i=1}^{\infty} \mu(A_i) \mid \{A_i\}_{i \in \mathbb{N}} \subseteq \mathcal{A} \text{ such that } A \subseteq \bigcup_{i \in \mathbb{N}} A_i \right\} \quad (19.101)$$

Given  $A \in \mathcal{P}(X)$  there exists using (2) a  $\{A_i\}_{i \in \mathbb{N}} \subseteq \mathcal{A}$  such that  $X \subseteq \bigcup_{i \in \mathbb{N}} A_i$  which as  $A \subseteq X$  proves that  $\mathcal{C}_A \neq \emptyset$  so that by 17.12  $\inf(\mathcal{C}_A)$  exists and as 0 is a lower bound for  $\mathcal{C}_A$  we have

$$\forall A \in \mathcal{A} \quad \varphi(A) = \inf(\mathcal{C}_A) \text{ exists and } \varphi(A) \in [0, \infty] \quad (19.102)$$

Next we check the conditions for a outer measure:

- i. Define  $\{A_i\}_{i \in \mathbb{N}_0}$  by  $A_i = \emptyset$  then by (1)  $\{A_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{A}$  and clearly  $\emptyset \subseteq \emptyset = \bigcup_{i \in \mathbb{N}_0} A_i$ , further  $\sum_{i=1}^{\infty} \mu(A_i) = \sum_{i=1}^{\infty} \mu(\emptyset) = 0$ . So  $0 \in \mathcal{C}_\emptyset$  hence  $0 \leq \inf(\mathcal{C}_\emptyset) \leq 0$  proving that

$$\varphi(\emptyset) \quad (19.103)$$

- ii. Let  $A, B \subseteq \mathcal{P}(X)$  with  $A \subseteq B$  then if  $x \in \mathcal{C}_B$  there exists a  $\{B_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{A}$  such that  $B \subseteq \bigcup_{i \in \mathbb{N}_0} B_i$  such that  $x = \sum_{i=1}^{\infty} \mu(B_i)$  then as  $A \subseteq B$  we have  $A \subseteq \bigcup_{i \in \mathbb{N}_0} B_i$  so that  $\mathcal{C}_B \subseteq \mathcal{C}_A$ . Using 2.175 we have that  $\inf(\mathcal{C}_A) \leq \inf(\mathcal{C}_B)$  proving that

$$\varphi(A) \leq \varphi(B) \text{ if } A \subseteq B \quad (19.104)$$

- iii. Let  $\{A_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{P}(X)$  and note  $A = \bigcup_{i \in \mathbb{N}_0} A_i$ . Take  $\varepsilon > 0$  and  $n \in \mathbb{N}_0$  then as  $\varphi(A_n) = \inf(\mathcal{C}_{A_n})$  there exists a  $\{I_{n,i}\}_{i \in \mathbb{N}_0} \subseteq \mathcal{A}$  such that

$$\varphi(A_n) \leq \sum_{i=1}^{\infty} \mu(I_{n,i}) < \varphi(A_n) + \frac{\varepsilon}{2^n} \text{ and } A_n \subseteq \bigcup_{i \in \mathbb{N}_0} I_{n,i} \quad (19.105)$$

As  $\mathbb{N}_0 \times \mathbb{N}_0$  is denumerable (see 5.59) there exists a bijection  $\beta: \mathbb{N}_0 \rightarrow \mathbb{N}_0 \times \mathbb{N}_0$ . If now  $x \in A$  then  $\exists n \in \mathbb{N}_0$  such that  $x \in A_n$ , as  $A_n \subseteq \bigcup_{i \in \mathbb{N}_0} I_{n,i}$  there exists a  $i \in \mathbb{N}_0$  such that  $x \in I_{n,i}$ . As  $\beta$  is surjective there exists then a  $k \in \mathbb{N}_0$  such that  $\beta(k) = (n, i)$  hence  $x \in I_{\beta(k)_1, \beta(k)_2} \subseteq \bigcup_{k \in \mathbb{N}} I_{\beta(k)_1, \beta(k)_2}$  proving that

$$A \subseteq \bigcup_{k \in \mathbb{N}_0} I_{\beta(k)_1, \beta(k)_2} \quad (19.106)$$

Next

$$\begin{aligned}
\varphi(A) &= \inf(\mathcal{C}_A) \\
&= \inf\left(\left\{\sum_{i=1}^{\infty} \mu(B_i) \mid \{B_i\}_{i \in \mathbb{N}} \subseteq \mathcal{A} \text{ such that } A \subseteq \bigcup_{i \in \mathbb{N}_0} B_i\right\}\right) \\
&\stackrel{19.106}{\leq} \sum_{k \in \mathbb{N}}^{\infty} \mu(I_{\beta(k)_1, \beta(k)_2}) \\
&\stackrel{17.107}{=} \sum_{i=1}^{\infty} \left( \sum_{j=1}^{\infty} \mu(I_{i,j}) \right) \\
&\stackrel{19.105}{<} \sum_{i=1}^{\infty} \left( \varphi(A_i) + \frac{1}{2^i} \right) \\
&\stackrel{17.96}{=} \sum_{i=1}^{\infty} \varphi(A_i) + \sum_{i=1}^{\infty} \frac{1}{2^i} \\
&\stackrel{11.415}{=} \sum_{i=1}^{\infty} \varphi(A_i) + \varepsilon \cdot \frac{\frac{1}{2}}{1 - \frac{1}{2}} \\
&= \sum_{i=1}^{\infty} \varphi(A_i) + \varepsilon
\end{aligned}$$

As  $\varepsilon$  is chosen arbitrary we have using 9.57 that

$$\varphi(A) \leq \sum_{i=1}^{\infty} \varphi(A_i) \quad (19.107)$$

Using 19.103, 19.104 and 19.107 it follows that

$$\varphi \text{ is a outer measure on } X \quad (19.108)$$

Now for the last part of the theorem. Let  $A \in \mathcal{A}$  and define  $\{A_i\}_{i \in \mathbb{N}_0}$  by  $A_i = \begin{cases} A & \text{if } i = 1 \\ \emptyset & \text{if } i \in \mathbb{N}_0 \setminus \{1\} \end{cases}$  then by (1)  $\{A_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{A}$  and  $A \subseteq A = \bigcup_{i \in \mathbb{N}_0} A_i$  so that  $\mu(A) \stackrel{(3)}{=} \mu(A) + \sum_{i=2}^{\infty} \mu(\emptyset) = \sum_{i=1}^{\infty} \mu(A_i) \in \mathcal{C}_A$ . Hence we have

$$\varphi(A) = \inf(\mathcal{C}_A) \leq \mu(A) \quad (19.109)$$

Further if  $\{A_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{A}$  such that  $A \subseteq \bigcup_{i \in \mathbb{N}_0} A_i$  we have by (4) that  $\mu(A) \leq \sum_{i=1}^{\infty} \mu(A_i) \in \mathcal{C}_A$  so that  $\mu(A) \leq \inf(\mathcal{C}_A) = \varphi(A)$  proving by 19.109 that

$$\forall A \in \mathcal{A} \text{ we have } \mu(A) = \varphi(A) \quad \square$$

Note that a outer measure is not a measure, it is not defined on a  $\sigma$ -algebra of  $X$  but on the set of all subsets of  $X$ , further  $\varphi$  is not countable additive but only countable sub additive. To be able to define a  $s$ -algebra we introduce the concept of a measurable set.

**Definition 19.72.** Let  $X$  be a set and  $\varphi: \mathcal{P}(X) \rightarrow [0, \infty]$  a outer measure then  $B \in \mathcal{P}(X)$  is  **$\varphi$ -measurable** if  $\forall A \in \mathcal{P}(X)$  we have  $\varphi(A) = \varphi(A \cap B) + \varphi(A \setminus B)$ . The set of all  $\varphi$ -measurable sets is noted as  $\mathcal{M}_\varphi[X]$ . So  $\mathcal{M}_\varphi[X] = \{B \in \mathcal{P}(X) \mid B \text{ is } \varphi\text{-measurable}\}$

**Lemma 19.73.** Let  $X$  be a set and  $\varphi: \mathcal{P}(X) \rightarrow [0, \infty]$  a outer measure then  $B \in \mathcal{P}(X)$  is  $\varphi$ -measurable if  $\forall A \in \mathcal{P}(X)$  we have  $\varphi(A) \geq \varphi(A \cap B) + \varphi(A \setminus B)$ .

**Proof.** Let  $A \in \mathcal{P}(X)$  then as  $A = (A \cap B) \cup (A \setminus B)$  we have by 19.70 that  $\varphi(A) \leq \varphi(A \cap B) + \varphi(A \setminus B)$  it follows from the assumption  $\varphi(A) \geq \varphi(A \cap B) + \varphi(A \setminus B)$  that  $\varphi(A) = \varphi(A \cap B) + \varphi(A \setminus B)$ .  $\square$

**Proposition 19.74.** *Let  $X$  be a set,  $\varphi$  a outer measure on  $X$  and  $B \subseteq X$  then*

1. *If  $\varphi(B) = 0$  then  $B \in \mathcal{M}_\varphi[X]$*
2. *If  $\varphi(X \setminus B)$  then  $B \in \mathcal{M}_\varphi[X]$*

**Proof.**

1. Let  $A \in X$  then as  $A \cap B \subseteq B$  we have  $\varphi(A \cap B) \leq \varphi(B) = 0$  proving that  $\varphi(A \cap B) = 0$ . Further as  $A \setminus B \subseteq A$  we have  $\varphi(A \setminus B) \leq \varphi(A)$ . Hence  $\varphi(A \cap B) + \varphi(A \setminus B) = 0 + \varphi(A \setminus B) \leq \varphi(A)$  which by 19.73 proves that  $B \in \mathcal{M}_\varphi[X]$
2. Let  $A \in X$  then as  $A \setminus B \subseteq X \setminus B$  we have  $\varphi(A \setminus B) \leq \varphi(X \setminus B) = 0$  proving that  $\varphi(A \setminus B) = 0$ . Further as  $A \cap B \subseteq A$  we have  $\varphi(A \cap B) \leq \varphi(A)$ . Hence  $\varphi(A \cap B) + \varphi(A \setminus B) = \varphi(A \setminus B) + 0 \leq \varphi(A)$  which by 19.73 proves that  $B \in \mathcal{M}_\varphi[X]$   $\square$

We are now ready to prove the Carathéodory theorem that shows how we can construct a measure space using a outer measure.

**Theorem 19.75. (Carathéodory theorem)** *Let  $X$  be a set,  $\varphi: \mathcal{P}(X) \rightarrow [0, \infty]$  a outer measure then*

1.  $\mathcal{M}_\varphi[X]$  is a  $\sigma$ -algebra on  $X$
2.  $\varphi|_{\mathcal{M}_\varphi[X]}$  is a  $\sigma$ -measure

hence  $\langle X, \mathcal{M}_\varphi[X], \varphi|_{\mathcal{M}_\varphi[X]} \rangle$  is a measure space.

**Proof.**

1. Let  $A \subseteq X$  then

$$\begin{aligned} \varphi(\emptyset \cap B) + \varphi(A \setminus \emptyset) &= \varphi(\emptyset) + \varphi(A) \\ &= 0 + \varphi(A) \\ &= \varphi(A) \end{aligned}$$

proving that

$$\emptyset \in \mathcal{M}_\varphi[X] \quad (19.110)$$

If  $B \in \mathcal{M}_\varphi[X]$  then

$$\begin{aligned} \varphi(A \cap (X \setminus B)) + \varphi(A \setminus (X \setminus B)) &\stackrel{1.28(10)}{=} \varphi(A \cap (X \setminus B)) + \varphi(A \cap B) \\ &\stackrel{1.28(9)}{=} \varphi((A \cap X) \setminus B) + \varphi(A \cap B) \\ &\stackrel{A \subseteq X}{=} \varphi(A \setminus B) + \varphi(A \cap B) \\ &\stackrel{B \in \mathcal{M}_\varphi[X]}{=} \varphi(A) \end{aligned}$$

proving that

$$\forall B \in \mathcal{M}_\varphi[X] \text{ we have } X \setminus B \in \mathcal{M}_\varphi[X] \quad (19.111)$$

Let  $B_1, B_2 \in \mathcal{M}_\varphi[X]$  and  $A \subseteq X$  then

$$\begin{aligned} \varphi(A \cap (B_1 \cup B_2)) &= \varphi((A \cap (B_1 \cup B_2)) \cap B_1) + \varphi((A \cap (B_1 \cup B_2)) \setminus B_1) \\ &= \varphi(A \cap B_1) + \varphi((A \cap (B_1 \cup B_2)) \setminus B_1) \\ &\stackrel{1.28(9)}{=} \varphi(A \cap B_1) + \varphi(A \cap ((B_1 \cup B_2) \setminus B_1)) \\ &\stackrel{1.28(8)}{=} \varphi(A \cap B_1) + \varphi(A \cap ((B_1 \setminus B_1) \cup (B_2 \setminus B_1))) \\ &= \varphi(A \cap B_1) + \varphi(A \cap (B_2 \setminus B_1)) \end{aligned} \quad (19.112)$$

Next

$$\begin{aligned}
 & \varphi(A \bigcap (B_1 \bigcup B_2)) + \varphi(A \setminus (B_1 \bigcup B_2)) && \stackrel{19.112}{=} \\
 & \varphi(A \bigcap B_1) + \varphi(A \bigcap (B_2 \setminus B_1)) + \varphi(A \setminus (B_1 \bigcup B_2)) && \stackrel{1.28(4)}{=} \\
 & \varphi(A \bigcap B_1) + \varphi(A \bigcap (B_2 \setminus B_1)) + \varphi((A \setminus B_1) \setminus B_2) && \stackrel{1.28(9)}{=} \\
 & \varphi(A \bigcap B_1) + \varphi((A \setminus B_1) \bigcap B_2) + \varphi((A \setminus B_1) \setminus B_1) && \stackrel{B_1 \in \bar{\mathcal{M}}_\varphi[X]}{=} \\
 & \varphi(A \bigcap B_1) + \varphi(A \setminus B_1) && = \\
 & \varphi(A)
 \end{aligned}$$

proving that

$$\forall B_1, B_2 \in \mathcal{M}_\varphi[X] \text{ we have } B_1 \bigcup B_2 \in \mathcal{M}_\varphi[X] \quad (19.113)$$

Hence using 19.110, 19.111 and 19.113 on 19.9 it follows that

$$\mathcal{M}_\varphi[X] \text{ is a algebraa}$$

Next we show by induction that  $\forall \{B_i\}_{i \in \mathbb{N}} \subseteq \mathcal{M}_\varphi[X]$  pairwise disjoint and  $A \subseteq X$  we have  $\forall n \in \mathbb{N}$  that

$$\varphi(A) = \sum_{i=1}^n \varphi(A \bigcap B_i) + \varphi\left(A \setminus \bigsqcup_{i=1}^n B_i\right) \quad (19.114)$$

**Proof.** Let  $\mathcal{S} = \{n \in \mathbb{N} \mid \varphi(A) = \sum_{i=1}^n \varphi(A \bigcap B_i) + \varphi(A \setminus \bigsqcup_{i=1}^n B_i)\}$  then we have

**1**  $\in \mathcal{S}$ . As  $B_1 \in \mathcal{M}_\varphi[X]$  we have

$$\begin{aligned}
 \varphi(A) &= \varphi(A \bigcap B_1) + \varphi(A \setminus B_1) \\
 &= \sum_{i=1}^1 \varphi(A \bigcap B_i) + \varphi\left(A \setminus \bigsqcup_{i=1}^1 B_i\right)
 \end{aligned}$$

proving that  $1 \in \mathcal{S}$

**n**  $\in \mathcal{S} \Rightarrow \mathbf{n+1} \in \mathcal{S}$ . If  $x \in A \bigcap B_{n+1}$  then  $x \in A$  and  $x \in B_{n+1} \Rightarrow x \notin B_i \forall i \in \{1, \dots, n\} \Rightarrow x \notin \bigsqcup_{i=1}^n B_i$  which proves that  $x \in (A \setminus \bigsqcup_{i=1}^n B_i) \bigcap B_{n+1}$ . Hence  $A \bigcap B_{n+1} \subseteq (A \setminus \bigsqcup_{i=1}^n B_i) \bigcap B_{n+1}$ . As also  $(A \setminus \bigsqcup_{i=1}^n B_i) \bigcap B_{n+1} \subseteq A \bigcap B_{n+1}$  it follows that

$$\left(A \setminus \bigsqcup_{i=1}^n B_i\right) \bigcap B_{n+1} = A \bigcap B_{n+1} \quad (19.115)$$

As  $B_{n+1} \in \mathcal{M}_\varphi[X]$  we have

$$\begin{aligned}
 \varphi\left(A \setminus \bigsqcup_{i=1}^n B_i\right) &= \varphi\left(\left(A \setminus \bigsqcup_{i=1}^n B_i\right) \bigcap B_{n+1}\right) + \varphi\left(\left(A \setminus \bigsqcup_{i=1}^n B_i\right) \setminus B_{n+1}\right) \\
 &\stackrel{19.115}{=} \varphi(A \bigcap B_{n+1}) + \varphi\left(\left(A \setminus \bigsqcup_{i=1}^n B_i\right) \setminus B_{n+1}\right) \\
 &\stackrel{1.28(4)}{=} \varphi(A \bigcap B_{n+1}) + \varphi\left(A \setminus \left(\left(\bigsqcup_{i=1}^n B_i\right) \bigsqcup B_{n+1}\right)\right) \\
 &= \varphi(A \bigcap B_{n+1}) + \varphi\left(A \setminus \bigsqcup_{i=1}^{n+1} B_i\right) \quad (19.116)
 \end{aligned}$$

Finally

$$\begin{aligned}
 \varphi(A) &\stackrel{n \in \mathcal{S}}{=} \sum_{i=1}^n \varphi(A \cap B_i) + \varphi\left(A \setminus \bigsqcup_{i=1}^n B_i\right) \\
 &\stackrel{19.116}{=} \sum_{i=1}^n \varphi(A \cap B_i) + \varphi(A \cap B_{n+1}) + \varphi\left(A \setminus \bigsqcup_{i=1}^{n+1} B_i\right) \\
 &= \sum_{i=1}^{n+1} \varphi(A \cap B_i) + \varphi\left(A \setminus \bigsqcup_{i=1}^{n+1} B_i\right)
 \end{aligned}$$

which proves that  $n+1 \in \mathcal{S}$ .

Mathematical induction proves then 19.114 □

Using the fact that  $\forall n \in \mathbb{N} A \setminus \bigsqcup_{i \in \mathbb{N}} B_i \subseteq A \setminus \bigsqcup_{i=1}^n B_i$  it follows that  $\varphi(A \setminus \bigsqcup_{i \in \mathbb{N}} B_i) \leq \varphi(A \setminus \bigsqcup_{i=1}^n B_i)$  hence

$$\begin{aligned}
 \sum_{i=1}^n \varphi(A \cap B_i) + \varphi\left(A \setminus \bigsqcup_{i \in \mathbb{N}} B_i\right) &\leq \sum_{i=1}^n \varphi(A \cap B_i) + \varphi\left(A \setminus \bigsqcup_{i=1}^n B_i\right) \\
 &\stackrel{19.114}{=} \varphi(A)
 \end{aligned} \tag{19.117}$$

Next

$$\begin{aligned}
 \sum_{i=1}^{\infty} \varphi(A \cap B_i) + \varphi\left(A \setminus \bigsqcup_{i \in \mathbb{N}} B_i\right) &= \lim_{n \rightarrow \infty} \left( \sum_{i=1}^n \varphi(A \cap B_i) \right) + \varphi\left(A \setminus \bigsqcup_{i \in \mathbb{N}} B_i\right) \\
 &\stackrel{17.69}{=} \lim_{n \rightarrow \infty} \left( \sum_{i=1}^n \varphi(A \cap B_i) + \varphi\left(A \setminus \bigsqcup_{i \in \mathbb{N}} B_i\right) \right) \\
 &\leq_{17.77, 19.117} \varphi(A)
 \end{aligned} \tag{19.118}$$

As  $A \cap (\bigsqcup_{i \in \mathbb{N}} B_i) \stackrel{1.101}{=} \bigcup_{i \in \mathbb{N}} (A \cap B_i)$  we have by the definition of a outer measure [see 19.69] that  $\varphi(A \cap (\bigsqcup_{i \in \mathbb{N}} B_i)) \leq \sum_{i=1}^{\infty} \varphi(A \cap B_i)$  so that by 19.118

$$\varphi\left(A \cap \left(\bigsqcup_{i \in \mathbb{N}} B_i\right)\right) + \varphi\left(A \setminus \bigsqcup_{i \in \mathbb{N}} B_i\right) \leq \varphi(A) \tag{19.119}$$

which by 19.73 and the fact that  $A \subseteq X$  was arbitrary chosen proves that

$$\forall \{B_i\}_{i \in \mathbb{N}} \subseteq \mathcal{M}_{\varphi}[X] \text{ pairwise disjoint we have that } \bigsqcup_{i \in \mathbb{N}} B_i \in \mathcal{M}_{\varphi}[X] \tag{19.120}$$

Combining the above with the fact that we have proven that  $\mathcal{M}_{\varphi}[X]$  is a algebra we have

$$\mathcal{M}_{\varphi}[X] \text{ is a } \sigma\text{-algebra on } X$$

2. Let  $\{B_i\}_{i \in \mathbb{N}} \subseteq \mathcal{M}_{\varphi}[X]$  pairwise disjoint then for  $A = \bigsqcup_{i=1}^{\infty} B_i$  we have that

$$\begin{aligned}
 \sum_{i=1}^{\infty} \varphi(B_i) &= \sum_{i=1}^{\infty} \varphi\left(\left(\bigsqcup_{i \in \mathbb{N}} B_i\right) \cap B_i\right) + 0 \\
 &= \sum_{i=1}^{\infty} \varphi\left(\left(\bigsqcup_{i \in \mathbb{N}} B_i\right) \cap B_i\right) + \varphi\left(\left(\bigsqcup_{i \in \mathbb{N}} B_i\right) \setminus \bigsqcup_{i \in \mathbb{N}} B_i\right) \\
 &= \sum_{i=1}^{\infty} \varphi(A \cap B_i) + \varphi\left(A \setminus \bigsqcup_{i \in \mathbb{N}} B_i\right) \\
 &\leq_{19.118} \varphi(A) \\
 &= \varphi\left(\bigsqcup_{i \in \mathbb{N}} B_i\right)
 \end{aligned}$$

Applying this with the semi-additivity of a outer measure [see 19.69] and the fact that  $\varphi(\emptyset)$  we have that

$$\varphi|_{\mathcal{M}_\varphi[A]}: \mathcal{M}_\varphi[X] \rightarrow [0, \infty] \text{ is a } \sigma\text{-measure}$$

□

### 19.3.3 The Carathéodory theorem and the set of real numbers

First we define the concept of the length of bounded intervals [see 19.22]

**Definition 19.76.** *The length function  $l: \mathcal{I} \rightarrow [0, \infty]$  is defined as follows [see also 19.22]*

$$l: \mathcal{I} \rightarrow [0, \infty] \text{ where } l(I) = \begin{cases} 0 & \text{if } I = \emptyset \\ b - a & \text{if } I = [a, b] \in \mathcal{I}_{[]} \\ b - a & \text{if } I = [a, b] \in \mathcal{I}_{[[]} \\ b - a & \text{if } I = ]a, b] \in \mathcal{I}_{[]]} \\ b - a & \text{if } I = ]a, b[ \in \mathcal{I}_{[[]} \end{cases}$$

We must of course prove that the function  $l$  is well defined and is not dependent on the boundary points.

**Proof.** Let  $I \in \mathcal{I}$  then we have either

$$I = \emptyset. \text{ then } l(\emptyset) = 0$$

$I \neq \emptyset$ . then we have either as  $\mathcal{I}_{[]}, \mathcal{I}_{[[]}, \mathcal{I}_{[]]}$  and  $\mathcal{I}_{[[]}$  are pairwise disjoint [see 19.24] we have for  $I$  either of the following exclusive cases to consider:

$I \in \mathcal{I}_{[]}$ . If  $I = [a, b] = [c, d]$  then by 2.189  $a = c \wedge b = d$  hence  $b - a = d - c$

$I \in \mathcal{I}_{[[]}$ . If  $I = [a, b] = [c, d]$  then by 2.189  $a = c \wedge b = d$  hence  $b - a = d - c$

$I \in \mathcal{I}_{[]]}$ . If  $I = ]a, b] = ]c, d]$  then by 2.189  $a = c \wedge b = d$  hence  $b - a = d - c$

$I \in \mathcal{I}_{[[]}$ . If  $I = ]a, b] = ]c, d]$  then by 2.189  $a = c \wedge b = d$  hence  $b - a = d - c$

so  $l$  is well defined on  $\mathcal{I}_{[]} \cup \mathcal{I}_{[[]} \cup \mathcal{I}_{[]]} \cup \mathcal{I}_{[[]}$  □

**Definition 19.77.** *Let  $A \in \mathcal{P}(\mathbb{R})$  then  $\mathcal{C}_A = \{\{I_i\}_{i \in \mathbb{N}} \mid \{I_i\}_{i \in \mathbb{N}} \subseteq \mathcal{I}_{[[]} \text{ and } A \subseteq \bigcup_{i \in \mathbb{N}} I_i\}$*

**Lemma 19.78.**  $\mathbb{R} = \bigcup_{i \in \mathbb{N}_0} ]-i, i[ = \bigcup_{i \in \mathbb{N}_0} [-i, i[ = \bigcup_{i \in \mathbb{N}_0} [-i, i]$

**Proof.** As  $\forall i \in \mathbb{N}_0$  we have  $] -i, i[ \subseteq [-i, i] \subseteq \mathbb{R}$  it follows that

$$\bigcup_{i \in \mathbb{N}_0} ]-i, i[ \subseteq \bigcup_{i \in \mathbb{N}_0} [-i, i] \subseteq \mathbb{R}$$

Let  $x \in \mathbb{R}$  then by 9.56 there exists a  $n \in \mathbb{Z}$  such that  $n \leq x < n + 1$ . We have then either

$0 < n$ . then as  $n < n + 1$  we have  $-(n + 1) < -n \leq 0 \leq n$  so that  $-(n + 1) < x < n + 1$  proving that  $x \in \bigcup_{i \in \mathbb{N}_0} ]-i, i[$

$n < 0$ . then then  $-n = |n| \Rightarrow -(|n| + 1) = -|n| - 1 < -|n| \leq x < n + 1 \leq |n + 1| < |n| + 1$  proving that  $x \in \bigcup_{i \in \mathbb{N}_0} ]-i, i[$

So we have

$$\mathbb{R} \subseteq \bigcup_{i \in \mathbb{N}_0} ]-i, i[ \subseteq \bigcup_{i \in \mathbb{N}_0} [-i, i] \subseteq \bigcup_{i \in \mathbb{N}_0} [-i, i]$$

□

Now to be able to use theorem 19.71 we prove the following lemma:

**Lemma 19.79.** *If  $I \in \mathcal{I}_{[[]} \cup \{\emptyset\}$  and  $\{I_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{I}_{[[]} \cup \{\emptyset\}$  is such that  $I \subseteq \bigcup_{i \in \mathbb{N}_0} I_i$  then*

$$l(I) \leq \sum_{i=1}^{\infty} l(I_i)$$

**Proof.** We have the following cases to consider for  $I$

$I = \emptyset$ . then trivially  $l(I) = l(\emptyset) = 0 \leq \sum_{i=1}^{\infty} l(I_i)$

$I \neq \emptyset$ . then  $\exists a, b \in \mathbb{R}$  with  $a < b$  such that  $I = [a, b]$ . Further for every  $i \in \mathbb{N}_0$  there exists a  $a_i, b_i \in \mathbb{R}$  with  $I_i = [a_i, b_i]$  [note: if  $I_i = \emptyset$  we can take  $a_i = b_i = 1$ ]. Define now  $\forall x \in [a, b]$

$$\{I_i^x\}_{i \in \mathbb{N}_0} \text{ by } I_i^x = \begin{cases} [a_i, \min(x, b_i)] & \text{if } x \in I_i = [a_i, b_i] \\ \emptyset & \text{if } x \notin I_i \end{cases} \quad (19.121)$$

and use this to define the following set

$$A = \left\{ x \in \mathbb{R} \mid x \in [a, b] \wedge x - a \leq \sum_{i=1}^{\infty} l(I_i^x) \right\} \subseteq [a, b] \quad (19.122)$$

As  $a \in [a, b]$  and  $a - a = 0 \leq \sum_{i=1}^{\infty} l(I_i^x)$  it follows that  $a \in A \Rightarrow A \neq \emptyset$ , further  $A \subseteq [a, b]$  and is thus bounded above by  $b$ , so we have using 9.43 that

$$c = \sup(A \in \mathbb{R}) \text{ exists and } a \leq c \leq b \quad (19.123)$$

Define  $B = \{\sum_{i=1}^{\infty} l(I_i^x) \mid x \in A\}$  then given  $x \in A$  we have that  $x - a \leq \sum_{i=1}^{\infty} l(I_i^x) \in B$  so that  $x - a \leq \sup(B)$  [note that  $A \neq \emptyset \Rightarrow B \neq \emptyset$  so that by 17.12  $\sup(B)$  exists (although it can be infinite)]. As  $a \in \mathbb{R}$  we have by 17.24 that  $x \leq \sup(B) + a$ . Hence  $c = \sup(A) \leq \sup(B) + a$  or using 17.24

$$c - a \leq \sup(B) = \sup \left( \left\{ \sum_{i=1}^{\infty} l(I_i^x) \mid x \in A \right\} \right) \quad (19.124)$$

Let  $x \in A$  and  $i \in \mathbb{N}_0$  then we have for  $I_i^x$  either:

$x \notin I_i$ . then  $I_i^x = \emptyset$  so that  $l(I_i^x) = l(\emptyset) = 0 \leq l(I_i^c)$

$x \in I_i$ . then  $a_i \leq x < b_i$ , as  $x \leq \sup(A) = c$  we have that  $a_i \leq \min(x, b_i) \leq \min(c, b_i) \leq b_i$ .

It follows that  $l(I_i^x) = \min(x, b_i) - a_i \leq \min(c, b_i) - a_i = l(I_i^c)$ .

So in all cases we have  $l(I_i^x) \leq l(I_i^c)$  giving  $\sum_{i=1}^{\infty} l(I_i^x) \leq \sum_{i=1}^{\infty} l(I_i^c)$  [see 17.97]. Taking in account 19.124 we have then that

$$c - a \leq \sup \left( \left\{ \sum_{i=1}^{\infty} l(I_i^x) \mid x \in A \right\} \right) \leq \sum_{i=1}^{\infty} l(I_i^c) \quad (19.125)$$

which as  $c \in [a, b]$  [see 19.123] proves that

$$c \in A \quad (19.126)$$

Assume now that  $c < b$  then  $c \in [a, b] \subseteq \bigcup_{i \in \mathbb{N}} I_i$  so that  $\exists k \in \mathbb{N}_0$  such that  $c \in I_k \Rightarrow I_k \neq \emptyset \Rightarrow c \in [a_k, b_k]$  giving

$$c < \min(b, b_k) = m \quad (19.127)$$

Given  $i \in \mathbb{N}_0$  we have either

$c \notin I_i$ . then  $l(I_i^c) = l(\emptyset) = 0 \leq l(I_i^c)$

$c \in I_i$ . then  $l(I_i^c) = \min(b_i, c) - a_i \leq \min(b_i, m) - a_i = l(I_i^m)$

so that we have

$$\forall i \in \mathbb{N}_0 \text{ we have } l(I_i^c) \leq l(I_i^m) \quad (19.128)$$

Further  $l(I_k^m) \underset{I_k \neq \emptyset}{=} \min(m, b_k) - a_k \underset{19.127 \Rightarrow m \leq b_k}{=} m - a_k$  and  $l(I_k^c) \underset{I_k \neq \emptyset}{=} \min(c, b_k) - a_k \underset{19.127 \Rightarrow c \leq b_k}{=} c - a_k$  giving

$$l(I_k^m) = l(I_k^c) + m - c \quad (19.129)$$

Define now  $\{c_i\}_{i \in \mathbb{N}_0}$  by  $c_i = \begin{cases} m - c & \text{if } i = k \\ 0 & \text{if } i \in \mathbb{N}_0 \setminus \{k\} \end{cases}$  then

$$\forall i \in \mathbb{N}_0 \text{ we have } l(I_i^c) + c_i = \begin{cases} l(I_k^c) + m - c & \underset{19.129}{=} l(I_k^m) \text{ if } i = k \\ l(I_i^c) \underset{19.128}{\leq} l(I_i^m) & \text{if } i \in \mathbb{N}_0 \setminus \{k\} \end{cases} \leq l(I_i^m) \quad (19.130)$$

So

$$\begin{aligned} \sum_{i=1}^{\infty} l(I_i^c) + m - c &= \sum_{i=1}^{\infty} l(I_i^c) + \sum_{i=1}^{\infty} c_i \\ &= \sum_{i=1}^{\infty} (l(I_i^c) + c_i) \\ &\stackrel{19.130}{\leq} \sum_{i=1}^{\infty} l(I_i^m) \end{aligned}$$

As  $m - c \in \mathbb{R}$  we have that  $\sum_{i=1}^{\infty} l(I_i^c) \leq \sum_{i=1}^{\infty} l(I_i^m) + c - m$  which using 19.125 gives  $c - a \leq \sum_{i=1}^{\infty} l(I_i^m) + c - m$  proving that

$$m - a \leq \sum_{i=1}^{\infty} l(I_i^m) \quad (19.131)$$

From  $c \in [a_k, b_k]$  and  $a \leq c$  [see 19.123] we have  $a < b_k \leq \min(b_k, b) \stackrel{19.127}{=} m = \min(b_k, b) \leq b$  it follows that  $m \in [a, b]$ . Combining this with 19.131 and the definition of  $A$  [see 19.122] it follows that  $m \in A$ . Hence  $m \leq \sup(A) = c$  contradicting  $c < m$  [see 19.127]. So the assumption that  $c < b$  is wrong so that  $b \leq c$  which using 19.123 proves that

$$c = b \quad (19.132)$$

From 19.126 and the definition of  $A$  we have that  $c - a \leq \sum_{i=1}^{\infty} l(I_i^c) \stackrel{19.132}{\Rightarrow} b - a \leq \sum_{i=1}^{\infty} l(I_i^b)$  proving

$$b - a \leq \sum_{i=1}^{\infty} l(I_i^b) \quad (19.133)$$

Finally if  $i \in \mathbb{N}_0$  then we have either

$$\mathbf{b} \notin \mathbf{I}_i. \text{ then } l(I_i^b) = \emptyset \leq l(I_i)$$

$$\mathbf{b} \in \mathbf{I}_i. \text{ then } a_i \leq b < b_i \text{ then } l(I_i^b) = \min(b, b_i) - a_i \leq b_i - a_i = l(I_i)$$

So in all cases we have that  $l(I_i^b) \leq l(I_i)$  which using 19.133 and 17.97 proves

$$b - a \leq \sum_{i=1}^{\infty} l(I_i)$$

□

We are finally able to define the outer measure on  $\mathbb{R}$

**Corollary 19.80.**  $\varphi: \mathcal{P}(\mathbb{R}) \rightarrow [0, \infty]$  defined by  $\varphi(A) = \inf(\{\sum_{i=1}^{\infty} l(I_i) \mid \{I_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{I}[\bigcup \{\emptyset\}]\})$  is a outer measure and  $\forall I \in \mathcal{I}[\bigcup \{\emptyset\}]$  we have that  $\varphi(I) = l(I)$

**Proof.** This follows from 19.78, 19.79 and 19.71. □

We can now use the Carathéodory method to define a measure space on  $\mathbb{R}$  that is a extension of  $\mathcal{B}[\mathbb{R}]$  [the Borel sets on  $\mathbb{R}$ ] and  $l$  the length of intervals on  $\mathbb{R}$ .

**Theorem 19.81. (Lebesgue Measure Space)** Define

$$\varphi: \mathcal{P}(\mathbb{R}) \rightarrow [0, \infty] \text{ by } \varphi(A) = \inf \left( \left\{ \sum_{i=1}^{\infty} l(I_i) \mid \{I_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{I}[\bigcup \{\emptyset\}] \right\} \right)$$

$$\begin{aligned} \mathcal{L}[\mathbb{R}] &= \{L \in \mathcal{P}(\mathbb{R}) \mid \forall A \in \mathcal{P}(\mathbb{R}) \varphi(A) = \varphi(A \bigcap L) + \varphi(A \setminus L)\} = \{L \in \mathcal{P}(\mathbb{R}) \mid L \text{ is } \varphi\text{-measurable}\} = \\ &= \mathcal{M}_{\varphi}[\mathbb{R}] \end{aligned}$$

$$\lambda: \mathcal{L}[\mathbb{R}] \rightarrow [0, \infty] \text{ defined by } \lambda(L) = \varphi(L) = \varphi|_{\mathcal{L}[\mathbb{R}]}$$

then  $\langle \mathbb{R}, \mathcal{L}[\mathbb{R}], \lambda \rangle$  is a measure space. Further we have  $\mathcal{B}[\mathbb{R}] \subseteq \mathcal{L}[\mathbb{R}]$  and  $\forall I \in \mathcal{I} \lambda(I) = l(I)$  where  $\mathcal{I} = \bigcup \mathcal{I}_{[]} \bigcup \mathcal{I}_{[\cdot]} \bigcup \mathcal{I}_{[\cdot]} \bigcup \{\emptyset\}$ .

**Proof.** (1) Using 19.75 and the above corollary [see 19.80] we have that

$$\langle \mathbb{R}, \mathcal{L}[\mathbb{R}], \lambda \rangle \text{ is a measure space} \quad (19.134)$$

Next we must prove that  $\mathcal{B}[\mathbb{R}] \subseteq \mathcal{L}[\mathbb{R}]$ , to do this we make use of the fact that  $\mathcal{B}[\mathbb{R}] = \sigma\{[-\infty, a] \mid a \in \mathbb{R}\}$  [see 19.26]. Let  $[-\infty, c] \in \{[-\infty, a] \mid a \in \mathbb{R}\}$ . Take  $A \subseteq \mathbb{R}$  then we have for  $\varphi$  the following cases to consider:

$\varphi(A) = \infty$ . Then  $\varphi(A \cap [-\infty, c]) + \varphi(A \setminus [-\infty, c]) \leq \varphi(A)$

$\varphi(A) \in \mathbb{R}$ . Let  $\varepsilon > 0$  then by the definition of  $\varphi$  as a infimum there exists a  $\{[a_i, b_i]\}_{i \in \mathbb{N}_0} \subseteq \mathcal{I}_{[']}$  with  $A \subseteq \bigcup_{i \in \mathbb{N}_0} [a_i, b_i]$  such that

$$\sum_{i=1}^{\infty} (b_i - a_i) = \sum_{i=1}^{\infty} l([a_i, b_i]) < \varphi(A) + \varepsilon \quad (19.135)$$

Take  $i \in \mathbb{N}_0$  then we have either

$a_i \leq c$ . In this cases we have for every  $x \in [a_i, b_i] \mid x < b_i \leq c$  proving that  $[a_i, b_i] \subseteq [-\infty, c]$  so that  $[a_i, b_i] \cap [-\infty, c] = [a_i, b_i]$  and  $[a_i, b_i] \setminus [-\infty, c] = \emptyset$ . Using then 19.70 we have

$$\begin{aligned} \varphi([a_i, b_i]) &= \varphi([a_i, b_i] \cap [-\infty, c]) \\ &= \varphi([a_i, b_i] \cap [-\infty, c]) + 0 \\ &= \varphi([a_i, b_i] \cap [-\infty, c]) + \varphi([a_i, b_i] \setminus [-\infty, c]) \end{aligned}$$

proving

$$\varphi([a_i, b_i]) = \varphi([a_i, b_i] \cap [-\infty, c]) + \varphi([a_i, b_i] \setminus [-\infty, c])$$

$c < b_i$ . Then we have for  $a_i$  either

$a_i \leq c$ . then

$$\begin{aligned} x \in [a_i, b_i] \cap [-\infty, c] &\Leftrightarrow a_i \leq x \wedge x < b_i \wedge x < c \\ &\stackrel{c < b_i}{\Leftrightarrow} a_i \leq x \wedge x < c \\ &\Leftrightarrow x \in [a_i, c] \end{aligned}$$

proving that  $[a_i, b_i] \cap [-\infty, c] = [a_i, c]$  [or  $\emptyset$  if  $a_i = c$ ] Hence

$$\varphi([a_i, b_i] \cap [-\infty, c]) = \varphi([a_i, c]) \stackrel{19.80}{=} l([a_i, c]) = c - a_i \quad (19.136)$$

Further

$$\begin{aligned} x \in [a_i, b_i] \setminus [-\infty, c] &\Leftrightarrow a_i \leq x \wedge x < b_i \wedge \neg(x < c) \\ &\Leftrightarrow a_i \leq x \wedge x < b_i \wedge c \leq x \\ &\stackrel{a_i \leq c}{\Leftrightarrow} x < b_i \wedge c \leq x \\ &\Leftrightarrow x \in [c, b_i] \end{aligned}$$

proving that  $[a_i, b_i] \setminus [-\infty, c] = [c, b_i]$ . Hence

$$\varphi([a_i, b_i] \setminus [-\infty, c]) = \varphi([c, b_i]) \stackrel{19.80}{=} l([c, b_i]) = b_i - c \quad (19.137)$$

Next

$$\begin{aligned} \varphi([a_i, b_i]) &\stackrel{19.80}{=} b_i - a_i \\ &= b_i - c + c - a_i \\ &\stackrel{19.136 \text{ and } 19.137}{=} \varphi([a_i, b_i] \cap [-\infty, c]) + \varphi([a_i, b_i] \setminus [-\infty, c]) \end{aligned}$$

proving

$$\varphi([a_i, b_i]) = \varphi([a_i, b_i[\cap ]-\infty, c]) + \varphi([a_i, b_i[\setminus ]-\infty, c])$$

**c**  $\leq a_i$ . If  $x \in [a_i, b_i[\cap ]-\infty, c[$  then  $a_i \leq x \wedge x < b_i \wedge x < c \leq a_i$  giving the contradiction  $a_i < a_i$ . Hence  $[a_i, b_i[\cap ]-\infty, c] = \emptyset$  so that  $[a_i, b_i[\setminus ]-\infty, c] = [a_i, b_i[$ . So

$$\begin{aligned}\varphi([a_i, b_i]) &= \varphi([a_i, b_i[\setminus ]-\infty, c]) \\ &= \varphi([a_i, b_i[\setminus ]-\infty, c]) + 0 \\ &= \varphi([a_i, b_i[\setminus ]-\infty, c]) + \varphi([a_i, b_i[\cap ]-\infty, c])\end{aligned}$$

proving again

$$\varphi([a_i, b_i]) = \varphi([a_i, b_i[\cap ]-\infty, c]) + \varphi([a_i, b_i[\setminus ]-\infty, c])$$

So in all cases we have

$$l([a_i, b_i]) \underset{19.80}{=} \varphi([a_i, b_i]) = \varphi([a_i, b_i[\cap ]-\infty, c]) + \varphi([a_i, b_i[\setminus ]-\infty, c]) \quad (19.138)$$

Next as  $A \subseteq \bigcup_{i \in \mathbb{N}_0} [a_i, b_i[$  we have that

$$\begin{aligned}A \setminus ]-\infty, c[ &\subseteq \left( \bigcup_{i \in \mathbb{N}_0} [a_i, b_i[ \right) \setminus ]-\infty, c[ \\ &\underset{1.102}{=} \bigcup_{i \in \mathbb{N}_0} ([a_i, b_i[\setminus ]-\infty, c])\end{aligned} \quad (19.139)$$

and

$$\begin{aligned}A \cap ]-\infty, c[ &\subseteq \left( \bigcup_{i \in \mathbb{N}_0} [a_i, b_i[ \right) \cap ]-\infty, c[ \\ &\underset{1.101}{=} \bigcup_{i \in \mathbb{N}_0} ([a_i, b_i[\cap ]-\infty, c])\end{aligned} \quad (19.140)$$

Next using 19.69, 19.139 and 19.140 we have

$$\begin{aligned}\varphi(A \cap ]-\infty, c]) + \varphi(A \setminus ]-\infty, c]) &\leq \underset{19.69, 19.139 \text{ and } 19.140}{\varphi(A)} \\ \varphi\left(\bigcup_{i \in \mathbb{N}_0} ([a_i, b_i[\cap ]-\infty, c])\right) + \varphi\left(\bigcup_{i \in \mathbb{N}_0} ([a_i, b_i[\setminus ]-\infty, c])\right) &\leq_{19.69} \\ \sum_{i=1}^{\infty} \varphi([a_i, b_i[\cap ]-\infty, c]) + \sum_{i=1}^{\infty} \varphi([a_i, b_i[\setminus ]-\infty, c]) &\underset{17.96}{=} \\ \sum_{i=1}^{\infty} (\varphi([a_i, b_i[\cap ]-\infty, c]) + \varphi([a_i, b_i[\setminus ]-\infty, c])) &\underset{19.138}{=} \\ \sum_{i=1}^{\infty} l([a_i, b_i[) &\underset{19.135}{<} \\ \varphi(A) + \varepsilon &\end{aligned}$$

This proves by 9.57 that  $\varphi(A \cap ]-\infty, c]) + \varphi(A \setminus ]-\infty, c]) \leq \varphi(A)$ , hence by 19.73 we have that  $]-\infty, c[$  is  $\varphi$ -measurable. Hence  $\{]-\infty, c[ \mid c \in \mathbb{R}\} \subseteq \mathcal{L}[\mathbb{R}]$ . Using 19.17 it follows that  $\sigma\{]-\infty, c[ \mid c \in \mathbb{R}\} \subseteq \mathcal{L}[\mathbb{R}]$ . Finally by 19.26 it follows that

$$\mathcal{B}[\mathbb{R}] \subseteq \mathcal{L}[\mathbb{R}] \quad (19.141)$$

Next we have to prove that  $\forall I \in \mathcal{I}$  we have  $l(I) = \varphi(I)$ . Let  $I \in \mathcal{I}$  then we have either:

$I \in \mathcal{I}_{[\![} \cup \{\emptyset\} \mathcal{I}_{[\![}$ . Then  $\exists a, b \in \mathbb{R}$  with  $a \leq b$  such that  $I = [a, b]$ . Hence

$$\varphi(I) = \varphi([a, b]) \underset{19.80}{=} l(I) = l([a, b]) = b - a \quad (19.142)$$

$I \in \mathcal{I}_{[\![}$ . Then  $\exists a, b \in \mathbb{R}$  with  $a \leq b$  such that  $I = [a, b]$ . As  $[a, b] \supseteq [a, b] \in \mathcal{I}_{[\![}$  we have that

$$l(I) = b - a = l([a, b]) \underset{19.142}{=} \varphi([a, b]) \leq \varphi([a, b]) \quad (19.143)$$

For the opposite inequality, let  $\varepsilon > 0$  and define  $\{J_i\}_{i \in \mathbb{N}_0}$  by  $J_i = \begin{cases} [a - \frac{\varepsilon}{4}, b + \frac{\varepsilon}{4}] & \text{if } i = 1 \\ [0, \frac{\varepsilon}{2^i}] & \text{if } i \in \mathbb{N}_0 \setminus \{1\} \end{cases}$  then we have  $[a, b] \subseteq [a - \frac{\varepsilon}{4}, b + \frac{\varepsilon}{4}] \subseteq \bigcup_{i \in \mathbb{N}_0} J_i$  and  $\{J_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{I}_{[\![}$ . So

$$\begin{aligned} \varphi([a, b]) &= \inf \left( \sum_{i=1}^{\infty} l(I_i) \mid \{I_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{I}_{[\![} \wedge [a, b] \subseteq \bigcup_{i \in \mathbb{N}_0} I_i \right) \\ &\leq \sum_{i=1}^{\infty} l(J_i) \\ &= l(J_1) + \sum_{i=2}^{\infty} l(J_i) \\ &= b + \frac{\varepsilon}{4} - \left( a - \frac{\varepsilon}{4} \right) + \sum_{i=2}^{\infty} \frac{1}{2^i} \\ &= b - a + \frac{\varepsilon}{2} + \sum_{i=2}^{\infty} \frac{1}{2^i} \\ &= b - a + \sum_{i=1}^{\infty} \frac{1}{2^i} \\ &\stackrel{11.415}{=} b - a + \frac{\varepsilon}{2} \\ &< b - a + \varepsilon \\ &= l([a, b]) + \varepsilon \end{aligned}$$

As  $\varepsilon > 0$  was chosen arbitrary it follows from 9.57 that  $\varphi([a, b]) \leq l([a, b])$  which combined with 19.143 proves

$$\varphi(I) = \varphi([a, b]) = l([a, b]) = l(I) \quad (19.144)$$

$I \in \mathcal{I}_{[a, b]}$ . Then  $\exists a, b \in \mathbb{R}$  with  $a < b$  such that  $I = [a, b]$ . As  $]a, b] \subseteq [a, b]$  we have

$$\varphi(]a, b]) \leq \varphi([a, b]) \quad (19.145)$$

Take  $\varepsilon$  such that  $0 < \varepsilon < b - a$  then  $a < a + \varepsilon < a + (b - a) = b$ . Hence if  $a \leq x \leq b$  then either  $a \leq x \leq a + \varepsilon$  or  $a + \varepsilon < x \leq b$  or  $[a, b] = [a, a + \varepsilon] \cup ]a + \varepsilon, b]$ . So that

$$\begin{aligned} \varphi([a, b]) &= \varphi([a, a + \varepsilon] \cup ]a + \varepsilon, b]) \\ &\leq_{19.70} \varphi([a, a + \varepsilon]) + \varphi(]a + \varepsilon, b]) \\ &\stackrel{19.144}{=} l([a, a + \varepsilon]) + \varphi(]a + \varepsilon, b]) \\ &= a + \varepsilon - a + \varphi(]a + \varepsilon, b]) \\ &= \varepsilon + \varphi(]a + \varepsilon, b]) \\ &\leq_{19.69} \varepsilon + \varphi(]a, b]) \end{aligned}$$

Using 9.57 we have then that  $\varphi([a, b]) \leq \varphi(]a, b])$  which combined with 19.145 proves that  $\varphi(]a, b]) = \varphi([a, b]) \underset{19.144}{=} l([a, b]) = b - a = l([a, b])$ . So we have

$$\varphi(I) = \varphi(]a, b]) = l(]a, b]) = b - a = l(I) \quad (19.146)$$

$I \in \mathcal{I}_{[]}^*$ . Then there exists a  $a, b \in \mathbb{R}$  with  $a < b$  such that  $I = ]a, b[$ . As  $]a, b[ \subseteq ]a, b]$  we have

$$\varphi(]a, b[) \subseteq \varphi(]a, b]) \quad (19.147)$$

Take  $\varepsilon$  such that  $0 < \varepsilon < b - a$  then  $a = b - (b - a) < b - \varepsilon < b$  so that  $]a, b[ = ]a, b - \varepsilon[ \cup ]b - \varepsilon, b]$ . So

$$\begin{aligned} \varphi(]a, b[) &= \varphi(]a, b - \varepsilon[ \cup ]b - \varepsilon, b]) \\ &\leq \varphi(]a, b - \varepsilon[) + \varphi(]b - \varepsilon, b]) \\ &\stackrel{19.144}{=} \varphi(]a, b - \varepsilon[) + l(]b - \varepsilon, b]) \\ &= \varphi(]a, b - \varepsilon[) + b - (b - \varepsilon) \\ &= \varphi(]a, b - \varepsilon[) + \varepsilon \\ &\leq \varphi(]a, b[) + \varepsilon \end{aligned}$$

Using 9.57 we have then that  $\varphi(]a, b[) \leq \varphi(]a, b])$  which together with 19.147 proves that  $\varphi(]a, b[) = \varphi(]a, b]) \stackrel{19.146}{=} b - a = l(]a, b[)$  proving that

$$\varphi(I) = \varphi(]a, b[) = l(]a, b[) = l(I) \quad (19.148)$$

Using all the cases above we conclude that

$$\forall I \in \mathcal{I} = \mathcal{I}_{[]} \cup \mathcal{I}_{[]}^* \cup \mathcal{I}_{[]}^* \cup \mathcal{I}_{[]}^* \cup \{\emptyset\} \text{ we have } \varphi(I) = l(I) \quad (19.149)$$

The theorem is then proved by 19.134, 19.141 and 19.149.  $\square$

### 19.3.4 The Carathéodory theorem and $\mathbb{R}^n$

Just as in the real case we use theorem 19.71 to define a outer measure on  $\mathbb{R}^n$ . first we must prove that  $\mathbb{R}^n$  can be covered by sets in  $\mathcal{R}^n$ .

**Lemma 19.82.** *Let  $n \in \mathbb{N}_0$  then  $\mathbb{R}^n = \bigcup_{i \in \mathbb{N}_0} ]-i, i[^n = \bigcup_{i \in \mathbb{N}_0} [-i, i[^n = \bigcup_{i \in \mathbb{N}_0} [-i, i]$*

**Proof.** First as  $\forall i \in \mathbb{N}_0 ]-i, i[^n \subseteq ]-i, i[^n \subseteq \mathbb{R}^n$  we have trivially that

$$\bigcup_{i \in \mathbb{N}_0} ]-i, i[^n \subseteq \bigcup_{i \in \mathbb{N}_0} [-i, i[^n \subseteq \bigcup_{i \in \mathbb{N}_0} [-i, i] \subseteq \mathbb{R}^n.$$

For the opposite inclusion let  $x \in \mathbb{R}^n$  then  $\forall i \in \{1, \dots, n\}$   $x_i \in \mathbb{R} \stackrel{19.78}{=} \bigcup_{j \in \mathbb{N}_0} [-j, j[$  so there exists  $j_i \in \mathbb{N}_0$  such that  $-j_i < x_i < j_i$ . Take  $k = \max(\{j_i | i \in \{1, \dots, n\}\}) \in \mathbb{N}_0$  then we have that  $-k \leq -j_i < x_j < j_i \leq k$  proving that  $x \in ]-k, k[^n$ . Hence

$$\mathbb{R}^n \subseteq \bigcup_{i \in \mathbb{N}_0} ]-i, i[^n \subseteq \bigcup_{i \in \mathbb{N}_0} [-i, i[^n \subseteq \bigcup_{i \in \mathbb{N}_0} [-i, i] \quad \square$$

Next we have to prove that the volume is semi-additive on  $\mathcal{R}^n$ .

**Lemma 19.83.** *Let  $n \in \mathbb{N}_0$ ,  $R, \{R_i\}_{i \in \mathbb{N}} \subseteq \mathcal{R}^n$  such that  $R \subseteq \bigcup_{i \in \mathbb{N}} R_i$  then  $v^n(R) \leq \sum_{i=1}^{\infty} v^n(R_i)$  [see 19.40]*

**Proof.** First if  $R = \emptyset$  then  $v^n(R) = 0 \leq \sum_{i=1}^{\infty} v^n(R_i)$  which proves the theorem in this case. So we must only prove the case for

$$R \neq \emptyset \quad (19.150)$$

We prove this by induction so define

$$\mathcal{S} = \left\{ n \in \mathbb{N}_0 | \forall R \in \mathcal{R}^n, \{R_i\}_{i \in \mathbb{N}} \subseteq \mathcal{R}^n \text{ with } R \subseteq \bigcup_{i \in \mathbb{N}} R_i \text{ we have } v^n(R) \leq \sum_{i=1}^n v^n(R_i) \right\} \quad (19.151)$$

then we have:

**1 ∈ S.** As  $R, \{R_i\}_{i \in \mathbb{N}} \in \mathcal{R}^1$  there exists  $a, b \in \mathcal{R}^1$  and  $\{a^i\}_{i \in \mathbb{N}}, \{b^i\}_{i \in \mathbb{N}} \subseteq \mathcal{R}^1$  such that  $R = [a, b[$  and  $R_i = [a^i, b^i[$ . Let  $x \in [a_1, b_1[$  then  $\underbrace{(x, \dots, x)}_1 \in [a, b[ \subseteq \bigcup_{i \in \mathbb{N}} [a^i, b^i[$  so that  $\exists k \in \mathbb{N}$  such that  $\underbrace{(x, \dots, x)}_1 \in [a^k, b^k[ \Rightarrow x \in [(a^k)_1, (b^k)_1[$  proving that

$$[a_1, b_1[ \subseteq \bigcup_{i \in \mathbb{N}} [(a^i)_1, (b^i)_1[ \quad (19.152)$$

Using 19.79 we have

$$b_1 - a_1 = l([a_1, b_1[) \leq \sum_{i=1}^{\infty} l([(a^i)_1, (b^i)_1[) = \sum_{i=1}^{\infty} ((b^i)_1 - (a^i)_1) \quad (19.153)$$

Finally

$$\begin{aligned} v^1(R) &= \prod_{i=1}^1 (b_i - a_i) \\ &= (b_1 - a_1) \\ &\stackrel{19.153}{\leq} \sum_{i=1}^{\infty} ((b^i)_1 - (a^i)_1) \\ &= \sum_{i=1}^{\infty} \left( \prod_{j=1}^1 ((b^i)_j - (a^i)_j) \right) \\ &= \sum_{i=1}^{\infty} v^1([b^i, a^i[) \end{aligned}$$

proving that  $1 \in \mathcal{S}$ .

**n ∈ S ⇒ n + 1 ∈ S.** Let  $R \in \mathcal{R}^{n+1}, \{R_i\}_{i \in \mathbb{N}} \subseteq \mathcal{R}^{n+1}$  such that  $R \subseteq \bigcup_{i \in \mathbb{N}} R_i$ . then  $R = [a, b[$ ,  $a \in \mathbb{R}^{n+1}$  with  $a \leq b$  and  $\forall i \in \mathbb{N} R_i = [a^i, b^i[, a^i, b^i \in \mathbb{R}^{n+1}$  with  $a_i \leq b_i$ . Further as  $R \neq \emptyset$  we must have  $\forall i \in \{1, \dots, n\}$  that  $a_i < b_i$  so that  $0 < \prod_{i=1}^n (b_i - a_i)$ . Define now  $V$  by

$$V = \prod_{i=1}^n (b_i - a_i) > 0 \quad (19.154)$$

then

$$\begin{aligned} v^{n+1}(R) &= \prod_{i=1}^{n+1} (b_i - a_i) \\ &= \left( \prod_{i=1}^n (b_i - a_i) \right) \cdot (b_{n+1} - a_{n+1}) \\ &\stackrel{19.154}{=} V \cdot (b_{n+1} - a_{n+1}) \end{aligned} \quad (19.155)$$

Given  $i \in \mathbb{N}, r \in \mathbb{R}$  define  $a^{i,r}, b^{i,r} \in \mathbb{R}^{n+1}$  by

$$(a^{i,r})_j = \begin{cases} (a^i)_j & \text{if } j \in \{1, \dots, n\} \\ \min((a^i)_{n+1}, r) & \text{if } j = n+1 \end{cases}, (b^{i,r})_j = \begin{cases} (b^i)_j & \text{if } j \in \{1, \dots, n\} \\ \min((b^i)_{n+1}, r) & \text{if } j = n+1 \end{cases} \quad (19.156)$$

Then  $(a^i)_j = \begin{cases} (a^i)_j & \text{if } j \in \{1, \dots, n\} \\ \min((a^i)_{n+1}, r) & \text{if } j = n+1 \end{cases} \leq_{a^i} b^i \begin{cases} (b^i)_j & \text{if } j \in \{1, \dots, n\} \\ \min((b^i)_{n+1}, r) & \text{if } j = n+1 \end{cases} = (b_j^{i,r})$  proving that

$$a^{i,r} \leq b^{i,r} \quad (19.157)$$

Now for

$$H^{i,r} = \{x \in R_i \mid x_{n+1} < r\} \subseteq R_i \quad (19.158)$$

we have either:

$$(a^i)_{n+1} \leq r.$$

$$\begin{aligned} x \in H^{i,r} &\Leftrightarrow x \in R_i \wedge x_{n+1} < r \\ &\Leftrightarrow (\forall j \in \{1, \dots, n+1\} (a^i)_j \leq x_j < (b^i)_j) \wedge x_{n+1} < r \\ &\Leftrightarrow (\forall j \in \{1, \dots, n\} (a^i)_j \leq x_j < (b^i)_j) \wedge (a^i)_{n+1} \leq x_{n+1} < (b^i)_{n+1} \wedge x_{n+1} < r \\ &\stackrel{(a^i)_{n+1} \leq r}{\Leftrightarrow} (\forall j \in \{1, \dots, n\} (a^i)_j \leq x_j < (b^i)_j) \wedge \min((a^i)_{n+1}, r) \leq x_{n+1} < \min((b^i)_{n+1}, r) \\ &\Leftrightarrow \forall j \in \{1, \dots, n+1\} (a^{i,r})_j \leq x_j < (b^{i,r})_j \\ &\Leftrightarrow x \in [a^{i,r}, b^{i,r}[ \end{aligned}$$

proving that

$$H^{i,r} = [b^{i,r}, a^{i,r}[$$

$r < (a^i)_{n+1}$ . If  $x \in H^{i,r}$  then  $r < (a^i)_{n+1} \leq x_{n+1} < r$  a contradiction so that  $H^{i,r} = \emptyset$ . If  $x \in [a^{i,r}, b^{i,r}[$  then  $r = \min((a^i)_{n+1}, r) \leq x_{n+1} < \min((b^i)_{n+1}, r) = r$  a contradiction so that  $[a^{i,r}, b^{i,r}[ = \emptyset$  hence

$$H^{i,r} = [b^{i,r}, a^{i,r}[$$

So in all cases we have  $H^{i,r} = [(a^{i,r})_{n+1}, (b^{i,r})_{n+1}[$  and as  $a^{i,r} \leq b^{i,r}$  [see 19.157] we have that  $H^{i,r} \in \mathcal{R}^{n+1}$ . To summarize

$$\forall i \in \mathbb{N}, r \in \mathbb{R} \text{ we have } H^{i,r} = [a^{i,r}, b^{i,r}[ \in \mathcal{R}^{n+1} \quad (19.159)$$

Next if  $r, s \in \mathbb{R}$  with  $r \leq s$  then if  $x \in H^{i,r}$  we have  $x \in R_i$  and  $x_{n+1} < r \leq s$  so that  $x \in H^{i,s}$  proving that  $H^{i,r} \subseteq H^{i,s}$ . Applying 19.43 gives

$$\forall r, s \in \mathbb{R} \text{ with } r \leq s \text{ and } \forall i \in \mathbb{N} \text{ we have } v^{n+1}(H^{i,r}) \leq v^{n+1}(H^{i,s}) \quad (19.160)$$

Further

$$\begin{aligned} v^{n+1}(H^{i,r}) &\stackrel{19.159}{=} v^{n+1}([b^{i,r}, a^{i,r}[) \\ &= \prod_{j=1}^{n+1} ((b^{i,r})_j - (a^{i,r})_j) \\ &= ((b^{i,r})_{n+1} - (a^{i,r})_{n+1}) \cdot \prod_{j=1}^n ((b^{i,r})_j - (a^{i,r})_j) \\ &= ((b^{i,r})_{n+1} - (a^{i,r})_{n+1}) \cdot \prod_{j=1}^n ((b^i)_j - (a^i)_j \quad (19.161) \end{aligned}$$

Let  $\varepsilon > 0$  and define

$$A_\varepsilon = \left\{ x \in [a_{n+1}, b_{n+1}] \mid V \cdot (x - a_{n+1}) \leq (1 + \varepsilon) \cdot \sum_{i=1}^{\infty} v^{n+1}(H^{i,x}) \right\} \subseteq [a_{n+1}, b_{n+1}] \quad (19.162)$$

As  $a_{n+1} \in [a_{n+1}, b_{n+1}]$  and  $V \cdot (a_{n+1} - a_{n+1}) = 0 \leq \sum_{i=1}^{\infty} v^{n+1}(H^{i,a_{n+1}})$  it follows that  $a_{n+1} \in A_\varepsilon$  hence  $\emptyset \neq A_\varepsilon \subseteq [a_{n+1}, b_{n+1}]$ . Using the conditional completeness of the real numbers [see 9.43] we have that

$$\gamma = \sup(A_\varepsilon) \text{ exists and } a_{n+1} \leq \gamma \leq b_{n+1} \quad (19.163)$$

For  $x \in A_\varepsilon$  we have now the following

$$\begin{aligned} V \cdot (x - a_{n+1}) &\leq_{\text{def } A_\varepsilon} (1 + \varepsilon) \cdot \sum_{i=1}^{\infty} v^{n+1}(H^{i,x}) \\ &\leq_{x \leq \lambda \text{ and } 19.160} (1 + \varepsilon) \cdot \sum_{i=1}^{\infty} v^{n+1}(H^{i,\gamma}) \quad (19.164) \end{aligned}$$

Next

$$\begin{aligned}
 V \cdot (\gamma - a_{n+1}) &= V \cdot (\sup(A_\varepsilon) - a_{n+1}) \\
 &\stackrel{17.30}{=} V \cdot (\sup(\{x - a_{n+1} \mid x \in A_\varepsilon\})) \\
 &\stackrel{17.30 \text{ and } 0 < V}{=} \sup(\{V \cdot (x - a_{n+1}) \mid x \in A_\varepsilon\}) \\
 &\stackrel{19.164}{\leq} (1 + \varepsilon) \cdot \sum_{i=1}^{\infty} v^{n+1}(H^{i, \gamma})
 \end{aligned} \tag{19.165}$$

proving that

$$\gamma \in A_\varepsilon \tag{19.166}$$

From 19.163 we have that  $a_{n+1} \leq \gamma \leq b_{n+1}$ , we show now by contradiction that  $\gamma = b_{n+1}$ . So assume that  $\gamma < b_{n+1}$ . Then

$$\gamma \in [a_{n+1}, b_{n+1}[ \tag{19.167}$$

Define now

$$J = \{x \in \mathbb{R}^n \mid (x, \dots, x_n, \gamma) \in R\} \text{ and } \forall i \in \mathbb{N} J_i = \{x \in \mathbb{R}^n \mid (x_1, \dots, x_n, \gamma) \in R_i\} \tag{19.168}$$

Define using  $a, b \in \mathbb{R}^{n+1}$   $a' = (a_1, \dots, a_n), b' = (b_1, \dots, b_n) \in \mathbb{R}^n$  then as  $a \leq b \forall j \in \{1, \dots, n\}$   $(a')_j = a_j \leq b_j = (b')_j$  proving  $a' \leq b'$ . Further

$$\begin{aligned}
 x \in [a', b'[ &\Leftrightarrow \forall j \in \{1, \dots, n\} a_j = (a')_j \leq x_j < (b')_j = b_j \\
 &\stackrel{19.167}{\Leftrightarrow} (\forall j \in \{1, \dots, n\} a_j = (a')_j \leq x_j < (b')_j = b_j) \wedge a_{n+1} \leq \gamma < b_{n+1} \\
 &\Leftrightarrow (x_1, \dots, x_n, \gamma) \in R \\
 &\Leftrightarrow x \in J
 \end{aligned}$$

As  $a' \leq b'$  we have also  $J = [a', b'[ \in \mathcal{R}^n$  and  $v^n(J) = \prod_{j=1}^n ((b')_j - (a')_j) = \prod_{j=1}^n (b_j - a_j) \stackrel{19.154}{=} V$ . Summarizing the above we have

$$J = [a', b'[ \in \mathcal{R}^n \text{ and } v^n(J) = V \tag{19.169}$$

Further for  $i \in \mathbb{N}$  define  $\bar{a}^i, \bar{b}^i \in \mathbb{R}^n$  by  $(\bar{a}^i)_j = (a^i)_j$  and  $(\bar{b}^i)_j = (b^i)_j$ . Then  $\forall j \in \{1, \dots, n\}$  we have  $(\bar{a}^i)_j = (a^i)_j \leq (b^i)_j = (\bar{b}^i)_j$  proving that

$$\bar{a}^i \leq \bar{b}^i \tag{19.170}$$

If  $J_i \neq \emptyset$  then  $\exists x \in \mathbb{R}^n$  such that  $(x_1, \dots, x_n, \gamma) \in R_i$  hence  $(a^i)_{n+1} \leq \gamma < (b^i)_{n+1}$ . Further we have

$$\begin{aligned}
 x \in J_i &\Leftrightarrow (x_1, \dots, x_n, \gamma) \in R_i \\
 &\Leftrightarrow \forall j \in \{1, \dots, n\} (a^i)_j \leq x_j < (b^i)_j \text{ and } (a^i)_{n+1} \leq \gamma < (b^i)_{n+1} \\
 &\stackrel{(a^i)_{n+1} \leq \gamma < (b^i)_{n+1}}{\Leftrightarrow} \forall j \in \{1, \dots, n\} (a^i)_j \leq x_j < (b^i)_j \\
 &\Leftrightarrow \forall j \in \{1, \dots, n\} (\bar{a}^i)_j \leq x_j < (\bar{b}^i)_j \\
 &\Leftrightarrow x \in [\bar{a}^i, \bar{b}^i[
 \end{aligned}$$

which proves that  $J_i = [\bar{a}^i, \bar{b}^i[$ . As  $\bar{a}^i \leq \bar{b}^i$  [see 19.170] we have  $J_i \in \mathcal{R}^n$  and  $v^n(J_i) = \prod_{j=1}^n ((\bar{b}^i)_j - (\bar{a}^i)_j) = \prod_{j=1}^n ((b^i)_j - (a^i)_j)$ . To summarize we have:

If  $J_i = \emptyset \Rightarrow J_i \in \mathcal{R}^n$  and if  $J_i \neq \emptyset$  we have  $J_i = [\bar{a}^i, \bar{b}^i[ \in \mathcal{R}^n$ ,  $(a^i)_{n+1} \leq \gamma < (b^i)_{n+1}$  and  $v^n(J_i) = \prod_{j=1}^n ((b^i)_j - (a^i)_j)$  (19.171)

If  $x \in J$  then  $(x_1, \dots, x_n, \gamma) \in R \subseteq \bigcup_{i \in \mathbb{N}} R_i$  so that  $\exists i \in \mathbb{N}$  such that  $(x_1, \dots, x_n, \gamma) \in R_i$  proving that  $x \in J_i$ . So

$$J \subseteq \bigcup_{i \in \mathbb{N}} J_i \tag{19.172}$$

As  $n \in \mathcal{S}$  [the induction hypothesis] we have by 19.169, 19.171 and 19.172 that

$$V \underset{19.169}{=} v^n(J) \leq \sum_{i=1}^{\infty} v^n(J_i) \underset{17.94}{=} \sup \left( \left\{ \sum_{i=1}^m v^n(J_i) \mid m \in \mathbb{N} \right\} \right) \quad (19.173)$$

As  $\varepsilon > 0$  we have  $\frac{1}{1+\varepsilon} < 1$  so that by multiplying with  $0 < V$  it follows that  $\frac{V}{1+\varepsilon} < V \leq \sup \left( \left\{ \sum_{i=1}^m v^n(J_i) \mid m \in \mathbb{N} \right\} \right)$ . Using the definition of a supremum we have then that  $\frac{V}{1+\varepsilon} < \sum_{i=1}^m v^n(J_i)$  giving

$$\exists m \in \mathbb{N} \text{ such that } V < (1+\varepsilon) \cdot \sum_{i=1}^m v^n(J_i) \quad (19.174)$$

Define now

$$\chi = \min \left( \{b_{n+1}\} \bigcup \{(b^i)_{n+1} \mid i \in \{1, \dots, m\} \wedge J_i \neq \emptyset\} \right) \quad (19.175)$$

As we have assumed that  $\gamma < b_{n+1}$  and by 19.171  $\gamma < (b^i)_{n+1}$  if  $J_i \neq \emptyset$  it follows that

$$a_{n+1} \leq \gamma < \chi \leq b_{n+1} \text{ [note that } \gamma < \chi \text{ will be crucial further in the proof]} \quad (19.176)$$

Let  $i \in \{1, \dots, m\}$  then for  $i \in \{1, \dots, m\}$  we have for  $J_i$  either

$J_i = \emptyset$ . then

$$\begin{aligned} v^{n+1}(H^{i,\gamma}) + (\chi - \gamma) \cdot v^n(J_i) &= v^{n+1}(H^{i,\gamma}) + 0 \\ &= v^{n+1}(H^{i,\gamma}) \\ &\leq \underset{19.160, 19.176}{v^{n+1}(H^{i,\chi})} \end{aligned}$$

$J_i \neq \emptyset$ . then  $(a^i)_{n+1} \leq \gamma < (b^i)_{n+1}$  [see 19.171] and  $\gamma < \chi$  [see 19.176] so that  $(a^i)_{n+1} < \chi$ . From this it follows that

$$(a^{i,\chi})_{n+1} = \min((a^i)_{n+1}, \chi) = (a^i)_{n+1} = \min((a^i)_{n+1}, \gamma) = (a^{i,\gamma})_{n+1} \quad (19.177)$$

$$(b^{i,\chi})_{n+1} = \min((b^i)_{n+1}, \chi) \underset{19.175}{=} \chi \quad (19.178)$$

$$(b^{i,\gamma})_{n+1} = \min((b^i)_{n+1}, \gamma) = \gamma \quad (19.179)$$

Next

$$\begin{aligned} v^{n+1}(H^{i,\chi}) &\underset{19.161}{=} ((b^{i,\chi})_{n+1} - (a^{i,\chi})_{n+1}) \cdot \prod_{j=1}^n ((b^i)_j - (a^i)_j) \\ &\underset{19.177, 19.178}{=} (\chi - (a^i)_{n+1}) \cdot \prod_{j=1}^n ((b^i)_j - (a^i)_j) \\ &= (\chi - \gamma + \gamma - (a^i)_{n+1}) \cdot \prod_{j=1}^n ((b^i)_j - (a^i)_j) \\ &= (\chi - \gamma) \cdot \prod_{j=1}^n ((b^i)_j - (a^i)_j) + (\gamma - (a^i)_{n+1}) \cdot \prod_{j=1}^n ((b^i)_j - (a^i)_j) \\ &\underset{19.171}{=} (\chi - \gamma) \cdot v^n(J_i) + (\gamma - (a^i)_{n+1}) \cdot \prod_{j=1}^n ((b^i)_j - (a^i)_j) \\ &\underset{19.177}{=} (\chi - \gamma) \cdot v^n(J_i) + ((b^{i,\gamma})_j - (a^{i,\gamma})_j) \cdot \prod_{j=1}^n ((b^i)_j - (a^i)_j \\ &\underset{19.161}{=} (\chi - \gamma) \cdot v^n(J_i) + v^{n+1}(H^{i,\gamma}) \end{aligned}$$

So in all cases we have

$$\forall i \in \{1, \dots, m\} \text{ we have } v^{n+1}(H^{i,\gamma}) + (\chi - \gamma) \cdot v^n(J_i) \leq v^{n+1}(H^{i,\chi}) \quad (19.180)$$

Further we have

$$\begin{aligned}
V \cdot (\chi - a_{n+1}) &= V \cdot (\chi - \gamma + \gamma - a_{n+1}) \\
&= V \cdot (\gamma - a_{n+1}) + V \cdot (\chi - \gamma) \\
&\stackrel{19.165}{\leq} (1 + \varepsilon) \cdot \sum_{i=1}^{\infty} v^{n+1}(H^{i,\gamma}) + V \cdot (\chi - \gamma) \\
&\stackrel{19.174}{\leq} (1 + \varepsilon) \cdot \sum_{i=1}^{\infty} v^{n+1}(H^{i,\gamma}) + (1 + \varepsilon) \cdot (\chi - \gamma) \cdot \sum_{i=1}^m v^n(J_i) \\
&= (1 + \varepsilon) \cdot \sum_{i=m+1}^{\infty} v^{n+1}(H^{i,\gamma}) + (1 + \varepsilon) \cdot \sum_{i=1}^m v^{n+1}(H^{i,\gamma}) + (1 + \varepsilon) \cdot (\chi - \gamma) \cdot \sum_{i=1}^m v^n(J_i) \\
&= (1 + \varepsilon) \cdot \sum_{i=m+1}^{\infty} v^{n+1}(H^{i,\gamma}) + (1 + \varepsilon) \cdot \left( \sum_{i=1}^m v^{n+1}(H^{i,\gamma}) + (\chi - \gamma) \cdot \sum_{i=1}^m v^n(J_i) \right) \\
&= (1 + \varepsilon) \cdot \sum_{i=m+1}^{\infty} v^{n+1}(H^{i,\gamma}) + (1 + \varepsilon) \cdot \sum_{i=1}^m (v^{n+1}(H^{i,\gamma}) + (\chi - \gamma) \cdot v^n(J_i)) \\
&\stackrel{19.180}{\leq} (1 + \varepsilon) \cdot \sum_{i=m+1}^{\infty} v^{n+1}(H^{i,\gamma}) + (1 + \varepsilon) \cdot \sum_{i=1}^m v^{n+1}(H^{i,\chi}) \\
&\stackrel{19.160, 19.176}{\leq} (1 + \varepsilon) \cdot \sum_{i=m+1}^{\infty} v^{n+1}(H^{i,\chi}) + (1 + \varepsilon) \cdot \sum_{i=1}^m v^{n+1}(H^{i,\chi}) \\
&= (1 + \varepsilon) \cdot \sum_{i=1}^{\infty} v^{n+1}(H^{i,\chi})
\end{aligned}$$

which using 19.176 and the definition of  $A_\varepsilon$  proves that  $\chi \in A_\varepsilon$ . So  $\chi \leq \sup(A_\varepsilon) = \gamma < \chi$  [see 19.176] giving the contradiction  $\chi < \chi$ . Hence the assumption  $\gamma < b_{n+1}$  is wrong giving  $b_{n+1} \leq \gamma \leq 19.163 b_{n+1}$  proving that

$$b_{n+1} = \gamma \in A_\varepsilon \quad (19.181)$$

Using the definition of  $A_\varepsilon$  we have that  $V \cdot (b_{n+1} - a_{n+1}) \leq (1 + \varepsilon) \cdot \sum_{i=1}^{\infty} v^{n+1}(H^{i,b_{n+1}})$ . As  $H^{i,b_{n+1}} \subseteq R_i$  [see 19.158] we have using 19.43 that

$$V \cdot (b_{n+1} - a_{n+1}) \leq (1 + \varepsilon) \cdot \sum_{i=1}^{\infty} v^{n+1}(R_i) \quad (19.182)$$

As  $v^{n+1}(R) = V \cdot (b_{n+1} - a_{n+1})$  [see 19.155] we have that  $v^{n+1}(R) \leq (1 + \varepsilon) \cdot \sum_{i=1}^{\infty} v^{n+1}(R_i)$ . Given that  $\varepsilon > 0$  was chosen arbitrary we have by 9.57 that

$$v^{n+1}(R) \leq \sum_{i=1}^{\infty} v^{n+1}(R_i)$$

which proves that  $n+1 \in \mathcal{S}$ . □

We can now define a outer measure based on the volume of half open intervals

**Definition 19.84.** Let  $n \in \mathbb{N}_0$  then  $\varphi^n: \mathcal{P}(\mathbb{R}^n) \rightarrow [0, \infty]$  is defined by

$$\varphi^n(R) = \inf \left( \left\{ \sum_{i=1}^{\infty} v^n(R_i) \mid \{R_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{R}^n \text{ such that } R \subseteq \bigcup_{i \in \mathbb{N}_0} R_i \right\} \right)$$

**Theorem 19.85.** Let  $n \in \mathbb{N}_0$  then  $\varphi^n : \mathcal{P}(\mathbb{R}^n) \rightarrow [0, \infty]$  is a outer measure. Further  $\forall A \in \mathcal{R}^n$  we have that  $v^n(A) = \varphi^n(A)$

**Proof.** We use 19.71 to prove this. Let's first check the conditions for the theorem.

1.  $\emptyset = \prod_{i \in \{1, \dots, n\}} [i, i] \in \mathcal{R}^n$
2. Using 19.82 there exists a  $\{R_i\}_{i \in \mathbb{N}} \subseteq \mathcal{R}^n$  such that  $\mathbb{R}^n = \bigcup_{i \in \mathbb{N}} R_i$
3.  $v^n(\emptyset) = 0$  by definition
4. Using 19.83 we have  $\forall R \in \mathcal{R}^n, \forall \{A_i\}_{i \in \mathbb{N}} \subseteq \mathcal{R}^n$  with  $R \subseteq \bigcup_{i \in \mathbb{N}} R_i$  that  $v^n(R) \leq \sum_{i=1}^{\infty} v^n(R_i)$

Applying then theorem 19.71 proves this theorem.  $\square$

Applying the above on the Carathéodory theorem (see 19.75, 19.72 and 19.73) allows us to define a  $\sigma$ -algebra and measure on  $\mathbb{R}^n$

**Definition 19.86.** Let  $n \in \mathbb{N}_0$  then  $\langle \mathbb{R}^n, \mathcal{L}[\mathbb{R}^n], \lambda^n \rangle$  is a measure space where

$$\begin{aligned} \mathcal{L}[\mathbb{R}^n] &= \mathcal{M}_{\varphi^n}[\mathbb{R}^n] \\ &= \{B \in \mathcal{P}(\mathbb{R}^n) \mid \forall A \in \mathcal{P}^n(\mathbb{R}^n) \text{ we have } \varphi^n(A) = \varphi^n(A \cap B) + \varphi^n(A \setminus B)\} \\ &= \{B \in \mathcal{P}(\mathbb{R}^n) \mid \forall A \in \mathcal{P}^n(\mathbb{R}^n) \text{ we have } \varphi^n(A) \geq \varphi^n(A \cap B) + \varphi^n(A \setminus B)\} \end{aligned}$$

and

$$\lambda^n = (\varphi^n)|_{\mathcal{L}[\mathbb{R}^n]} : \mathcal{L}[\mathbb{R}^n] \rightarrow [0, \infty]$$

If  $A \in \mathcal{L}[\mathbb{R}^n]$  then we set that  $A$  is Lebesgue measurable on  $\mathbb{R}^n$ .

Just as with the real numbers we proceed now to prove that the Borel algebra defined by the open sets in  $\mathbb{R}^n$  is included in  $\mathcal{L}[\mathbb{R}^n]$ . To do this we first prove that half spaces are Lebesgue measurable in  $\mathbb{R}^n$ .

**Lemma 19.87.** Let  $n \in \mathbb{N}_0, i \in \{1, \dots, n\}$  and  $x \in \mathbb{R}$  then  $]-\infty, x[i]$  defined by  $]-\infty, x[i] = \{y \in \mathbb{R}^n \mid y_i < x\}$  [19.53] is Lebesgue measurable on  $\mathbb{R}^n$  or in other words  $]-\infty, x[i] \in \mathcal{L}[\mathbb{R}^n]$ .

**Proof.** First we prove that

$$\forall R \in \mathcal{R}^n \text{ we have } \varphi^n(R) = \varphi^n(R \cap ]-\infty, x[i]) + \varphi^n(R \setminus ]-\infty, x[i]) \quad (19.183)$$

**Proof.** Consider the following exclusive cases for  $R$  and  $]-\infty, x[i]$

$R \subseteq ]-\infty, x[i]$ . then

$$\begin{aligned} \varphi^n(R \cap ]-\infty, x[i]) + \varphi^n(R \setminus ]-\infty, x[i]) &= \varphi^n(R) + \varphi^n(\emptyset) \\ &\stackrel{19.69 \text{ and } 19.85}{=} \varphi^n(R) + 0 \\ &= \varphi^n(R) \end{aligned}$$

$R \not\subseteq ]-\infty, x[i] \wedge R \cap ]-\infty, x[i] = \emptyset$ . then

$$\begin{aligned} \varphi^n(R \cap ]-\infty, x[i]) + \varphi^n(R \setminus ]-\infty, x[i]) &= \varphi^n(\emptyset) + \varphi^n(R) \\ &\stackrel{19.69 \text{ and } 19.85}{=} v^n(R) \end{aligned}$$

$R \not\subseteq ]-\infty, x[i] \wedge R \cap ]-\infty, x[i] \neq \emptyset$ . then  $R \neq \emptyset$  so there exists  $a, b \in \mathbb{R}^n$  with  $a < b$  such that  $R = [a, b]$ . Consider now the following cases for  $x$ :

**$b_i \leq x$ .** let  $y \in R$  then  $a_i \leq y_i < b_i \leq x$  so that  $y \in ]-\infty, x[i]$  proving that  $R \subseteq ]-\infty, x[i]$  which contradicts the assumption  $R \not\subseteq ]-\infty, x[i]$ . So we must have  $x < b_i$

**$x \leq a_i$ .** as  $R \cap ]-\infty, x[i] \neq \emptyset$  there exists a  $y \in R \cap ]-\infty, x[i]$  so that  $a_i \leq y_i < b_i \wedge y_i < x_i \leq a_i$  giving the contradiction  $a_i < a_i$ . So we must have  $a_i < x$ .

From the above we conclude that

$$a_i < x < b_i \quad (19.184)$$

Next define  $\bar{b}$  by  $(\bar{b})_k = \begin{cases} x & \text{if } k = i \\ b_k & \text{if } k \in \{1, \dots, n\} \setminus \{i\} \end{cases}$  then using 19.184 together with  $a < b$  we have that

$$a < \bar{b} \quad (19.185)$$

$$\begin{aligned} y \in R \bigcap ]-\infty, x[i] &\Leftrightarrow (\forall j \in \{1, \dots, n\} \setminus \{i\} \text{ we have } a_j \leq y_j < b_j) \wedge a_i \leq y_i < b_i \wedge y_i < x \\ &\Leftrightarrow_{x < b_j} (\forall j \in \{1, \dots, n\} \setminus \{i\} \text{ we have } a_j \leq y_j < b_j) \wedge a_i \leq y_i < x \\ &\Leftrightarrow \forall j \in \{1, \dots, n\} \text{ we have } a_j \leq y_j < (\bar{b})_j \\ &\Leftrightarrow y \in [a, \bar{b}[ \end{aligned}$$

Hence we have that  $R \bigcap ]-\infty, x[i] = [a, \bar{b}[ \in \mathcal{R}^n$  [using 19.185] so that  $v^n(R \bigcap ]-\infty, x[i])$  exists and

$$\begin{aligned} \varphi^n(R \bigcap ]-\infty, x[i]) &\stackrel{19.80}{=} v^n(R \bigcap ]-\infty, x[i]) \\ &= \prod_{j \in \{1, \dots, n\}} ((\bar{b})_j - a_j) \\ &= \left( \prod_{j \in \{1, \dots, n\} \setminus \{i\}} (b_j - a_j) \right) \cdot (x - a_i) \end{aligned} \quad (19.186)$$

Further we have if we define  $\bar{a}$  by  $(\bar{a})_k = \begin{cases} x & \text{if } k = i \\ a_k & \text{if } k \in \{1, \dots, n\} \setminus \{i\} \end{cases}$  then using 19.184 and  $a < b$  we have that

$$\bar{a} < b \quad (19.187)$$

and

$$\begin{aligned} y \in R \setminus ]-\infty, x[i] &\Leftrightarrow (\forall j \in \{1, \dots, n\} \setminus \{i\} a_j \leq y_j < b_j) \wedge a_i \leq y_i < b_i \wedge x \leq y_i \\ &\Leftrightarrow_{a_i < x} (\forall j \in \{1, \dots, n\} \setminus \{i\} a_j \leq y_j < b_j) \wedge x \leq y_i < b_i \\ &\Leftrightarrow \forall j \in \{1, \dots, n\} (\bar{a})_j \leq y_j < b_j \\ &\Leftrightarrow y \in [\bar{a}, b[ \end{aligned}$$

So we have that  $R \setminus ]-\infty, x[i] = [\bar{a}, b[ \in \mathcal{R}^n$  [using 19.187] hence  $v^n(R \setminus ]-\infty, x[i])$  exists and

$$\begin{aligned} \varphi^n(R \setminus ]-\infty, x[i]) &\stackrel{19.85}{=} v^n(R \setminus ]-\infty, x[i]) \\ &= \prod_{j \in \{1, \dots, n\}} (b_i - (\bar{a})_j) \\ &= \left( \prod_{j \in \{1, \dots, n\} \setminus \{i\}} (b_j - a_j) \right) \cdot (b_i - x) \end{aligned} \quad (19.188)$$

Hence

$$\begin{aligned} \varphi^n(R \bigcap ]-\infty, x[i]) + \varphi^n(R \setminus ]-\infty, x[i]) &\stackrel{19.186 \text{ and } 19.188}{=} \\ \left( \prod_{j \in \{1, \dots, n\} \setminus \{i\}} (b_j - a_j) \right) \cdot (x - a_i) + \left( \prod_{j \in \{1, \dots, n\} \setminus \{i\}} (b_j - a_j) \right) \cdot (b_i - x) &= \\ \left( \prod_{j \in \{1, \dots, n\} \setminus \{i\}} (b_j - a_j) \right) (x - a_i + b_i - x) &= \\ \left( \prod_{j \in \{1, \dots, n\} \setminus \{i\}} (b_j - a_j) \right) (b_i - a) &= \\ v^n(R) &\stackrel{19.85}{=} \\ \varphi^n(R) \end{aligned}$$

So in all cases we have  $\varphi^n(R \bigcap ]-\infty, x[i]) + \varphi^n(R \setminus ]-\infty, x[i]) = \varphi^n(R)$  which proves 19.183  $\square$

We are now ready to prove that  $]-\infty, x[i]$  is Lebesgue measurable. Let  $A \in \mathcal{P}(\mathbb{R}^n)$  and take  $\varepsilon > 0$  then from the definition of  $\varphi^n$  and the infimum there exists a  $\{R_j\}_{j \in \mathbb{N}} \subseteq \mathcal{R}^n$  with  $R \subseteq \bigcup_{j \in \mathbb{N}} R_j$  such that

$$\varphi^n(A) \leq \sum_{j=1}^{\infty} v^n(R_j) < \varphi^n(A) + \varepsilon \quad (19.189)$$

As

$$\begin{aligned} A \cap ]-\infty, x[i] &\subseteq \left( \bigcup_{j \in \mathbb{N}} R_j \right) \cap ]-\infty, x[i] \\ &\stackrel{1.101}{=} \bigcup_{j \in \mathbb{N}} (R_j \cap ]-\infty, x[i]) \end{aligned}$$

and

$$\begin{aligned} A \setminus \bigcup_{j \in \mathbb{N}} (R_j \cap ]-\infty, x[i]) &\subseteq \left( \bigcup_{j \in \mathbb{N}} R_j \right) \setminus ]-\infty, x[i] \\ &\stackrel{1.102}{=} \bigcup_{j \in \mathbb{N}} (R_j \setminus ]-\infty, x[i]) \end{aligned}$$

we have [see 19.69] that

$$\begin{aligned} \varphi^n(A \cap ]-\infty, x[i]) + \varphi^n(A \setminus ]-\infty, x[i]) &\leq \varphi^n\left(\bigcup_{j \in \mathbb{N}} (R_j \cap ]-\infty, x[i])\right) + \varphi^n\left(\bigcup_{j \in \mathbb{N}} (R_j \setminus ]-\infty, x[i])\right) \\ &\stackrel{19.69}{\leq} \sum_{j=1}^{\infty} \varphi^n(R_j \cap ]-\infty, x[i]) + \sum_{j=1}^{\infty} \varphi^n(R_j \setminus ]-\infty, x[i]) \\ &\stackrel{17.96}{=} \sum_{j=1}^{\infty} (\varphi^n(R_j \cap ]-\infty, x[i]) + \varphi^n(R_j \setminus ]-\infty, x[i])) \\ &\stackrel{19.183}{\leq} \sum_{j=1}^{\infty} \varphi^n(R_j) \\ &\stackrel{19.189}{\leq} \varphi^n(A) + \varepsilon \end{aligned}$$

which as  $\varepsilon$  was chosen arbitrary proves by 9.57 that

$$\varphi^n(A \cap ]-\infty, x[i]) + \varphi^n(A \setminus ]-\infty, x[i]) \leq \varphi^n(A)$$

Applying 19.86 proves then that

$$]-\infty, x[i] \in \mathcal{L}[\mathbb{R}^n] \quad \square$$

**Lemma 19.88.** *Let  $n \in \mathbb{N}_0$  then*

$$\begin{aligned} \mathcal{B}[\mathbb{R}^n] &= \sigma[\{C \mid C \text{ is closed in } \mathcal{T}_{\mathbb{R}^n}\}] \\ &= \sigma[\mathcal{R}^n] \\ &= \sigma[\mathcal{H}^n] \\ &\subseteq \mathcal{L}[\mathbb{R}^n] \end{aligned}$$

$$\forall R \in \mathcal{R}^n \text{ we have } v^n(R) = \lambda^n(R)$$

and

$$\forall R \in \mathcal{B}[\mathbb{R}^n] \text{ we have } \varphi^n(R) = \lambda^n(R)$$

**Proof.** Using 19.87 and the definition  $\mathcal{H}^n$  [see 19.53] we have that  $\mathcal{H}^n \subseteq \mathcal{L}[\mathbb{R}^n]$  hence  $\sigma[\mathcal{H}^n] \subseteq \mathcal{L}[\mathbb{R}^n]$ . Using 19.58 the first part of the theorem is proved. The rest of the theorem follows from 19.86 and 19.85.  $\square$

**Corollary 19.89.** *Let  $n \in \mathbb{N}_0$  then  $\langle \mathbb{R}^n, \mathcal{B}[\mathbb{R}^n], (\lambda^n)_{|\mathcal{B}[\mathbb{R}^n]} \rangle$  is a measure space*

**Proof.** This follows trivially from 19.62 and the above (see 19.88)  $\square$

### 19.3.5 Properties of $\langle \mathbb{R}^n, \mathcal{L}[\mathbb{R}^n], \lambda^n \rangle$ and $\langle \mathbb{R}^n, \mathcal{B}[\mathbb{R}^n], (\lambda^n)_{|\mathcal{B}[\mathbb{R}^n]} \rangle$

**Lemma 19.90.** Let  $n \in \mathbb{N}_0$ ,  $a, b \in \mathbb{R}^n$  with  $a \leq b$  and  $\varepsilon > 0$  then there exists a  $a', b' \in \mathbb{R}^n$  such that  $[a, b] \subseteq [a', b']$  and  $v^n([a', b']) = \prod_{i \in \{1, \dots, n\}} (b'_i - a'_i) \leq v^n([a, b]) + \varepsilon$

**Proof.** First if  $[a, b] = \emptyset$  define  $a', b'$  by  $a'_i = 0, b'_i = (\frac{\varepsilon}{2})^{\frac{1}{n}}$ . Then  $a' < b'$ ,  $[a, b] = \emptyset \subseteq [a', b']$  and  $\prod_{i=1}^n (b'_i - a'_i) = \prod_{i=1}^n (\frac{\varepsilon}{2})^{\frac{1}{n}} = \left(\left(\frac{\varepsilon}{2}\right)^{\frac{1}{n}}\right)^n \stackrel{18.39}{=} \frac{\varepsilon}{2} < \varepsilon = 0 + \varepsilon = v^n([a, b]) + \varepsilon$  which proves the theorem for the case  $[a, b] = \emptyset$ . So we are left with the case where  $[a, b] \neq \emptyset$ . In this case we must have  $\forall i \in \{1, \dots, n\}$  that  $a_i < b_i$  [otherwise  $[a, b] = \emptyset$ ].

We prove the theorem by induction so let  $\mathcal{S} = \{n \in \mathbb{N}_0 \mid \forall \varepsilon > 0, \forall a, b \in \mathbb{R}^n \text{ with } a < b \text{ there exists } a', b' \in \mathbb{R}^n \text{ with } [a, b] \subseteq [a', b'] \text{ and } \prod_{i=1}^n (b'_i - a'_i) \leq v^n([a, b]) + \varepsilon\}$  then we have:

**1  $\in \mathcal{S}$ .** Let  $a, b \in \mathbb{R}^1$  with  $a < b$  and  $\varepsilon > 0$  and take  $a' = a - \varepsilon, b' = b$  then  $[a, b] \subseteq [a - \varepsilon, b] = [a', b']$  and  $\prod_{i=1}^1 (b'_i - a'_i) = b' - a' = b - (a - \varepsilon) = (b - a) + \varepsilon = v^1([a, b]) + \varepsilon$  which proves that  $1 \in \mathcal{S}$ .

**$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$ .** Take  $a = (a_1, \dots, a_{n+1}), (b_1, \dots, b_{n+1}) \in \mathbb{R}^{n+1}$  with  $a < b$  then  $\bar{a} = (a_1, \dots, a_n), \bar{b} = (b_1, \dots, b_n) \in \mathbb{R}^n$  with  $\bar{a} < \bar{b}$ . Take  $\varepsilon > 0$ . As  $n \in \mathcal{S}$  there exists a  $\bar{a}' = (a'_1, \dots, a'_n), \bar{b}' = (b'_1, \dots, b'_n)$  with  $\bar{a}' < \bar{b}'$  such that  $[\bar{a}, \bar{b}] \subseteq [\bar{a}', \bar{b}']$  and for

$$\delta_1 = \min\left(\frac{\varepsilon}{3 \cdot (b_{n+1} - a_{n+1})}, 1\right) \quad (19.190)$$

we have

$$\prod_{i=1}^n (b'_i - a'_i) \leq v^n([\bar{a}, \bar{b}]) + \delta_1 = \prod_{i=1}^n (b_i - a_i) + \delta_1 \quad (19.191)$$

Take

$$\delta_2 = \min\left(\frac{\varepsilon}{3}, \frac{\varepsilon}{3 \cdot \prod_{i=1}^n (b_i - a_i)}\right) \quad (19.192)$$

and define  $a'_{n+1} = a_{n+1} - \delta_2, b'_{n+1} = b_{n+1}$  and  $a' = (a'_1, \dots, a'_{n+1}), b' = (b'_1, \dots, b'_{n+1})$ . If  $x \in [a, b]$  we have for  $i \in \{1, \dots, n+1\}$  either

**$i \in \{1, \dots, n\}$ .**  $x_i \in [a_i, b_i] = [\bar{a}_i, \bar{b}_i] \stackrel{[\bar{a}, \bar{b}] \subseteq [\bar{a}', \bar{b}']}{\Rightarrow} x_i \in [\bar{a}'_i, \bar{b}'_i] = [a'_i, b'_i]$

**$i = n+1$ .**  $x_{n+1} \in [a_{n+1}, b_{n+1}] \Rightarrow x_{n+1} \in [a_{n+1} - \delta_2, b_{n+1}] = [a'_{n+1}, b'_{n+1}]$

which proves

$$[a, b] \subseteq [a', b'] \quad (19.193)$$

Further

$$\begin{aligned} & \prod_{i=1}^{n+1} (b'_i - a'_i) = \\ & \left( \prod_{i=1}^n (b'_i - a'_i) \right) \cdot (b'_{n+1} - a'_{n+1}) \stackrel{19.191}{\leq} \\ & \left( \prod_{i=1}^n (b_i - a_i) + \delta_1 \right) \cdot (b'_{n+1} - a'_{n+1}) = \\ & \left( \prod_{i=1}^n (b_i - a_i) + \delta_1 \right) \cdot ((b_{n+1} - a_{n+1}) + \delta_2) = \\ & \left( \prod_{i=1}^n (b_i - a_i) \right) \cdot (b_{n+1} - a_{n+1}) + \delta_2 \cdot \prod_{i=1}^n (b_i - a_i) + \delta_1 \cdot (b_{n+1} - a_{n+1}) + \delta_1 \cdot \delta_2 \stackrel{19.190, 19.192}{<} \\ & \left( \prod_{i=1}^n (b_i - a_i) \right) \cdot (b_{n+1} - a_{n+1}) + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \\ & \left( \prod_{i=1}^n (b_i - a_i) \right) \cdot (b_{n+1} - a_{n+1}) + \varepsilon \end{aligned}$$

which together with 19.193 proves that  $n+1 \in \mathcal{S}_n$   $\square$

**Proposition 19.91.** *Let  $n \in \mathbb{N}_0$  and  $A \in \mathcal{L}[\mathbb{R}^n]$  /A is a Lebesgue measurable set/ then*

$$\lambda^n(A) = \inf(\{\lambda^n(U) | U \in \mathcal{T}_{\mathbb{R}^n} \wedge A \subseteq U\}) = \inf(\{\lambda^n(U) | U \text{ open in } \mathcal{T}_{\mathbb{R}^n} \wedge A \subseteq U\})$$

*Note: As  $\mathbb{R}^n$  is open and  $\mathcal{T}_{\mathbb{R}^n} \subseteq \mathcal{L}[\mathbb{R}^n]$  we have that  $\inf(\{\lambda^n(U) | U \text{ open in } \mathcal{T}_{\mathbb{R}^n} \wedge A \subseteq U\})$  is well defined.*

**Proof.** First for all  $U \in \mathcal{T}_{\mathbb{R}^n}$  with  $A \subseteq U$  we have  $\lambda^n(A) \leq \lambda^n(U)$  so that  $\lambda^n(A)$  is a lower bound of  $\{\lambda^n(U) | U \in \mathcal{T}_{\mathbb{R}^n} \wedge A \subseteq U\}$  hence

$$\lambda^n(A) \leq \inf(\{\lambda^n(U) | U \in \mathcal{T}_{\mathbb{R}^n} \wedge A \subseteq U\}) \quad (19.194)$$

Now for the inverse inequality, consider the following possible cases for  $\lambda(A)$

$\lambda(A) = \infty$ . then trivially

$$\inf(\{\lambda^n(U) | U \in \mathcal{T}_{\mathbb{R}^n} \wedge A \subseteq U\}) \leq \infty = \lambda(A)$$

$\lambda(A) < \infty$ . Let  $\varepsilon > 0$  then as  $\lambda^n(A) \underset{19.86, 19.84}{=} \inf(\{\sum_{i=1}^{\infty} v^n(R_i) | \{R_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{R}^n \wedge A \subseteq \bigcup_{i \in \mathbb{N}_0} R_i\})$  there exists a  $\{R_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{R}^n$  with  $A \subseteq \bigcup_{i \in \mathbb{N}_0} R_i$  and

$$\sum_{i=1}^{\infty} v^n(R_i) < \lambda^n(A) + \frac{\varepsilon}{2} \quad (19.195)$$

As  $\forall i \in \mathbb{N}_0$  we have  $R_i = [a_i, b_i]$  where  $a_i, b_i \in \mathbb{R}^n$  with  $a_i \leq b_i$  we can use 19.90 to find  $[a'_i, b'_i] \in \mathcal{T}_{\mathbb{R}^n}$  such that  $[a_i, b_i] \subseteq [a'_i, b'_i]$  and  $v^n([a'_i, b'_i]) \leq v^n([a_i, b_i]) + \frac{\varepsilon}{2^{i+1}} = v^n(R_i) + \frac{\varepsilon}{2^{i+1}}$ . So

$$\begin{aligned} \sum_{i=1}^{\infty} v^n([a'_i, b'_i]) &\leq \sum_{i=1}^{\infty} \left( v^n(R_i) + \frac{\varepsilon}{2^{i+1}} \right) \\ &= \sum_{i=1}^{\infty} v^n(R_i) + \sum_{i=1}^{\infty} \frac{\varepsilon}{2^{i+1}} \\ &= \sum_{i=1}^{\infty} v^n(R_i) + \frac{\varepsilon}{2} \cdot \sum_{i=1}^{\infty} \frac{1}{2^i} \\ &\underset{11.415}{=} \sum_{i=1}^{\infty} v^n(R_i) + \frac{1}{2} \cdot \varepsilon \\ &\underset{19.195}{<} \lambda^n(A) + \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\ &< \lambda^n(A) + \varepsilon \end{aligned} \quad (19.196)$$

Take  $U = \bigcup_{i \in \mathbb{N}_0} [a'_i, b'_i] \in \mathcal{T}_{\mathbb{R}^n}$  then  $A \subseteq \bigcup_{i \in \mathbb{N}_0} R_i \subseteq \bigcup_{i \in \mathbb{N}_0} [a'_i, b'_i] = U$  giving

$$\lambda^n(U) \in \{\lambda^n(U) | U \in \mathcal{T}_{\mathbb{R}^n} \wedge A \subseteq U\} \quad (19.197)$$

As  $U = \bigcup_{i \in \mathbb{N}_0} [a'_i, b'_i] \subseteq \bigcup_{i \in \mathbb{N}_0} [a'_i, b'_i]$  we have

$$\begin{aligned} \lambda^n(U) &\leq \lambda^n\left(\bigcup_{i \in \mathbb{N}_0} [a'_i, b'_i]\right) \\ &\underset{19.66}{\leq} \sum_{i=1}^{\infty} \lambda^n([a'_i, b'_i]) \\ &\underset{19.88}{=} \sum_{i=1}^{\infty} v^n([a'_i, b'_i]) \\ &\underset{19.196}{\leq} \lambda^n(A) + \varepsilon \end{aligned} \quad (19.198)$$

From 19.197 and 19.198 it follows that  $\inf(\{\lambda^n(U) | U \in \mathcal{T}_{\mathbb{R}^n} \wedge A \subseteq U\}) \leq \lambda^n(A) + \varepsilon$  which as  $\varepsilon$  was arbitrary chosen that

$$\inf(\{\lambda^n(U) | U \in \mathcal{T}_{\mathbb{R}^n} \wedge A \subseteq U\}) \leq \lambda^n(A)$$

Hence in all cases we have  $\inf(\{\lambda^n(U) | U \in \mathcal{T}_{\mathbb{R}^n} \wedge A \subseteq U\}) \leq \lambda^n(A)$  which combined with 19.194 proves that

$$\lambda^n(A) = \inf(\{\lambda^n(U) | U \in \mathcal{T}_{\mathbb{R}^n} \wedge A \subseteq U\})$$

□

**Proposition 19.92.** *Let  $n \in \mathbb{N}_0$  and  $A \in \mathcal{L}[\mathbb{R}^n]$  [A is a Lebesgue measurable set] then*

$$\lambda^n(A) = \sup(\{\lambda^n(K) : K \subseteq A \wedge K \text{ is compact in } \mathcal{T}_{\mathbb{R}}\})$$

*Note: As compact spaces are closed in  $\mathbb{R}^n$  [see 11.254 and 11.228] we have by 19.88 that every compact sets is Lebesgue measurable, also  $\emptyset$  is a compact set with  $\emptyset \subseteq A$  so that  $\{\lambda^n(K) : K \subseteq A \wedge K \text{ is compact in } \mathcal{T}_{\mathbb{R}}\} \neq \emptyset$ . Hence  $\sup(\{\lambda^n(K) : K \subseteq A \wedge K \text{ is compact in } \mathcal{T}_{\mathbb{R}}\})$  is well defined.*

**Proof.** First if  $K$  is compact with  $K \subseteq A$  then we have  $\lambda^n(K) \leq \lambda^n(A)$  so that

$$\sup(\{\lambda^n(K) : K \subseteq A \wedge K \text{ compact}\}) \leq \lambda^n(A) \quad (19.199)$$

For the opposite inequality we prove this first for the bounded case. So assume that  $A$  is bounded then [see 11.77] there exists a  $M \in \mathbb{R}$  such that  $\forall a \in A$  we have  $\|a\|_{\max} \leq M$ . Hence

$$A \subseteq \bar{B}_{\|\cdot\|_{\max}}(0, M) \quad (19.200)$$

From 11.56 it follows that  $\bar{B}_{\|\cdot\|_{\max}}(0, M)$  is closed and as it is bounded it follows from 11.300 that  $\bar{B}_{\|\cdot\|_{\max}}(0, M)$  is compact. Hence if we take  $\bar{B}_{\|\cdot\|_{\max}}(0, M) = C$  then

$$A \subseteq C \text{ where } C \text{ is compact closed and bounded} \quad (19.201)$$

Further if  $y \in \bar{B}_{\|\cdot\|_{\max}}(0, M) = C$  then  $\|y\|_{\max} \leq M$  or  $\forall i \in \{1, \dots, n\}$   $0 \leq |y_i| \leq M < M + 1 \Rightarrow -(M + 1) < y_i < M + 1$  or  $y \in [-(M + 1), M + 1]^n$ . So  $A \subseteq C \subseteq [-(M + 1), M + 1]^n$  or  $\lambda^n(A) \leq \lambda^n(C) \leq \lambda^n([-(M + 1), M + 1]^n) = v^n(\prod_{i \in \{1, \dots, n\}} [-(M + 1), M + 1]) = \prod_{i=1}^n (2 \cdot (M + 1)) < \infty$  proving that

$$\lambda^n(A) \leq \lambda^n(C) < \infty \quad (19.202)$$

As  $C$  is closed we have that  $C \in \mathcal{L}[\mathbb{R}^n]$  [see 19.58] and as also  $A \in \mathcal{L}[\mathbb{R}^n]$  it follows that  $C \setminus A \in \mathcal{L}[\mathbb{R}^n]$ . Take  $\varepsilon > 0$  then by 19.91 we have

$$\exists U \text{ open with } C \setminus A \subseteq U \text{ such that } \lambda^n(U) < \lambda^n(C \setminus A) + \varepsilon \quad (19.203)$$

As  $U$  is open we have that  $L = C \setminus U$  is closed and as  $L \subseteq C$  [a compact set], we have by 11.255 that

$$L = C \setminus U \text{ is compact,closed and } L \in \mathcal{L}[\mathbb{R}^n] \quad (19.204)$$

Further as  $C \subseteq (C \setminus U) \cup U = L \cup U$  it follows [see 19.66] that

$$\lambda^n(C) \leq \lambda^n(L) + \lambda^n(U) \quad (19.205)$$

Now as  $A \subseteq C$  [see 19.201] and  $\lambda^n(A) < \infty$  [see 19.202] we have by 19.64 that

$$\lambda^n(C \setminus A) = \lambda^n(C) - \lambda^n(A) \quad (19.206)$$

Next

$$\begin{aligned} \lambda^n(C) &\stackrel{19.205}{\leq} \lambda^n(L) + \lambda^n(U) \\ &\stackrel{19.203}{\leq} \lambda^n(L) + \lambda^n(C \setminus A) + \varepsilon \\ &\stackrel{19.206}{=} \lambda^n(L) + \lambda^n(C) - \lambda^n(A) + \varepsilon \end{aligned}$$

which as  $\lambda^n(C), \lambda^n(A) < \infty$  proves using 17.24 that  $\lambda^n(A) \leq \lambda^n(L) + \varepsilon$ . As  $\varepsilon > 0$  was chosen arbitrary it follows from 9.57 that

$$\lambda^n(A) \leq \lambda^n(L) \quad (19.207)$$

Further if  $x \in L = C \setminus U$  then  $x \in C \wedge x \notin U$ , if  $x \notin A$  then  $x \in C \setminus A \subseteq_{19.203} U$  contradicting  $x \notin U$  hence we must have  $x \in A$ . So  $L \subseteq A$  and as  $L$  is compact [see 19.204] we have that  $\lambda^n(L) \in \{\lambda^n(K) | K \subseteq A \wedge K \text{ compact}\}$  proving that  $\lambda^n(L) \leq \sup(\{\lambda^n(K) | K \subseteq A \wedge K \text{ compact}\})$ . Using 19.207 we have then that  $\lambda^n(A) \leq \sup(\{\lambda^n(K) | K \subseteq A \wedge K \text{ compact}\})$ . Combining this with 19.199 proves

$$\text{If } A \in \mathcal{L}[\mathbb{R}^n] \text{ is bounded then } \lambda^n(A) = \sup(\{\lambda^n(K) | K \subseteq A \wedge K \text{ is compact}\}) \quad (19.208)$$

Next for the general case, consider the case where  $A \in \mathcal{L}[\mathbb{R}^n]$  is not necessarily bounded. Then

$$\begin{aligned} A &= A \cap \mathbb{R}^n \\ &\stackrel{19.82}{=} A \cap \left( \bigcup_{i \in \mathbb{N}_0} [-i, i]^n \right) \\ &= \bigcup_{i \in \mathbb{N}_0} (A \cap [-i, i]^n) \\ &= \bigcup_{i \in \mathbb{N}_0} A_i \text{ where } A_i = A \cap [-i, i]^n \in \mathcal{L}[\mathbb{R}^n] \end{aligned}$$

If  $x \in A_i$  then  $\forall j \in \{1, \dots, n\}$  we have  $-i \leq x_j < i$  so that  $|x_j| \leq i$ , hence  $\|x\|_{\max} \leq i$  which proves that  $A_i$  is bounded. Further  $A_i = A \cap [-i, i]^n \subseteq A \cap [-i+1, i+1]^n = A_{i+1}$ . To summarize

$$\text{If } A \in \mathcal{L}[\mathbb{R}^n] \text{ then } \exists \{A_i\}_{i \in \mathbb{N}} \subseteq \mathcal{L}[\mathbb{R}^n] \text{ with } A_i \subseteq A_{i+1} \wedge A_i \text{ is bounded so that } A = \bigcup_{i \in \mathbb{N}_0} A_i \quad (19.209)$$

Using 19.67 on the above gives

$$\lambda^n(A) = \sup(\{\lambda^n(A_i) | i \in \mathbb{N}_0\}) \quad (19.210)$$

Assume now that  $\sup(\{\lambda^n(K) | K \text{ compact} \wedge K \subseteq A\}) < \lambda^n(A) \stackrel{19.210}{=} \sup(\{\lambda^n(A_i) | i \in \mathbb{N}_0\})$ . Then there exists a  $i \in \mathbb{N}_0$  such that

$$\sup(\{\lambda^n(K) | K \text{ compact} \wedge K \subseteq A\}) < \lambda^n(A_i)$$

As  $A_i \in \mathcal{L}[\mathbb{R}^n]$  is bounded [see 19.209] we have by 19.208 that  $\lambda^n(A_i) = \sup(\{\lambda^n(K) | K \text{ compact} \wedge K \subseteq A_i\})$  giving using the above that

$$\sup(\{\lambda^n(K) | K \text{ compact} \wedge K \subseteq A\}) < \sup(\{\lambda^n(K) | K \text{ compact} \wedge K \subseteq A_i\}) = \lambda^n(A_i)$$

So there exists a compact set  $L$  with  $L \subseteq A_i$  such that

$$\sup(\{\lambda^n(K) | K \text{ compact} \wedge K \subseteq A\}) < \lambda^n(L)$$

As  $A_i \subseteq A$  we have  $L \subseteq A$  so that  $L \in \{\lambda^n(K) | K \text{ compact} \wedge K \subseteq A\}$  and reach the contradiction

$$\lambda^n(L) \leq \sup(\{\lambda^n(K) | K \text{ compact} \wedge K \subseteq A\}) < \lambda^n(L)$$

Hence the assumption  $\sup(\{\lambda^n(K) | K \text{ compact} \wedge K \subseteq A\}) < \lambda^n(A)$  is wrong proving that  $\lambda^n(A) \leq \sup(\{\lambda^n(K) | K \text{ compact} \wedge K \subseteq A\})$ . So taking in account 19.199 we have

$$\lambda^n(A) = \sup(\{\lambda^n(K) | K \text{ compact} \wedge K \subseteq A\})$$

in the general case.  $\square$

As a example of a application of the above propositions we prove that the Lebesgue measure on  $\mathbb{R}^n$  can be seen as a extension of any measures on  $\mathcal{B}[\mathbb{R}^n]$ .

**Proposition 19.93.** *Let  $n \in \mathbb{N}_0$ . If  $\mu$  is a measure on  $\mathcal{B}[\mathbb{R}^n]$  such that  $\forall R \in \mathcal{D}^n$  [see 19.46] we have  $v^n(R) = \mu(R)$  then  $\mu = (\lambda^n)|_{\mathcal{B}[\mathbb{R}^n]}$  [or  $\forall A \in \mathcal{B}[\mathbb{R}^n]$  we have  $\mu(A) = \lambda^n(A)$ ]*

**Proof.** Let  $U \in \mathcal{T}_{\mathbb{R}^n}$  be an open set then by 19.50 there exists a pairwise disjoint family  $\{D_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{D}^n$  of half open rectangles such that  $U = \bigsqcup_{i \in \mathbb{N}_0} D_i$ . So

$$\begin{aligned}\mu(U) &= \mu\left(\bigsqcup_{i \in \mathbb{N}_0} D_i\right) \\ &= \sum_{i=1}^{\infty} \mu(D_i) \\ &= \sum_{i=1}^{\infty} v^n(D_i) \\ &\stackrel{19.88}{=} \sum_{i=1}^n \lambda^n(D_i) \\ &= \lambda^n\left(\bigsqcup_{i \in \mathbb{N}_0} D_i\right) \\ &= \lambda^n(U)\end{aligned}$$

proving

$$\forall U \in \mathcal{T}_{\mathbb{R}^n} \text{ we have } \mu(U) = \lambda^n(U) \quad (19.211)$$

Let  $A \in \mathcal{B}[\mathbb{R}^n] \subseteq \mathcal{L}[\mathbb{R}^n]$  then if  $U \in \mathcal{T}_{\mathbb{R}}$  such that  $A \subseteq U$  then we have  $\mu(A) \leq \mu(U) \stackrel{19.211}{=} \lambda^n(U)$  proving that  $\mu(A) \leq \inf(\{\lambda^n(U) | U \in \mathcal{T}_{\mathbb{R}^n} \wedge A \subseteq U\}) \stackrel{19.91}{=} \lambda^n(A)$ . So we have

$$\forall A \in \mathcal{B}[\mathbb{R}^n] \mu(A) \leq \lambda^n(A) \quad (19.212)$$

Now to replace this inequality by an equality. Let  $A \in \mathcal{B}[\mathbb{R}^n]$  be such that  $\exists a, b \in \mathbb{R}^n$  with  $a \leq b$  such that  $A \subseteq ]a, b[$ . Then as  $\lambda^n(]a, b[) \leq \lambda^n(]a, b[) = \prod_{i=1}^{\infty} (b_i - a_i) < \infty$  we have

$$\begin{aligned}\mu(]a, b[) &\leq \lambda^n(]a, b[) < \infty \\ \mu(A) &\leq_{19.212} \lambda^n(A) \leq \lambda^n(]a, b[) < \infty \\ \mu(]a, b[ \setminus A) &\leq_{19.212} \lambda^n(]a, b[ \setminus A) \leq \lambda^n(]a, b[) < \infty\end{aligned}$$

proving that

$$\mu(]a, b[), \mu(A), \mu(]a, b[ \setminus A), \lambda^n(]a, b[), \lambda^n(A), \lambda^n(]a, b[ \setminus A) < \infty \quad (19.213)$$

Further we have as  $A \stackrel{19.211}{=} (]a, b[ \setminus A) \bigsqcup A$  that

$$\mu(]a, b[) = \mu(]a, b[ \setminus A) + \mu(A) \leq_{19.212} \lambda^n(]a, b[ \setminus A) + \lambda^n(A) = \lambda^n(]a, b[) \stackrel{19.211}{=} \mu(]a, b[)$$

so that

$$\mu(]a, b[ \setminus A) + \mu(A) = \lambda^n(]a, b[ \setminus A) + \lambda^n(A) \quad (19.214)$$

If now  $\mu(A) < \lambda^n(A)$  then  $\mu(]a, b[ \setminus A) + \mu(A) < \mu(]a, b[ \setminus A) + \lambda^n(A) \leq_{19.212} \lambda^n(]a, b[ \setminus A) + \lambda^n(A)$  contradicting 19.214. Hence  $\lambda^n(A) \leq \mu(A)$  which combined with 19.212 proves  $\lambda^n(A) = \mu(A)$ . To summarize

$$\forall A \in \mathcal{B}[\mathbb{R}^n] \text{ such that } \exists a, b \in \mathbb{R}^n \text{ with } a \leq b \text{ and } A \subseteq ]a, b[ \text{ then } \mu(A) = \lambda^n(A) \quad (19.215)$$

Now for the general case let  $A \in \mathcal{B}[\mathbb{R}^n]$  then

$$\begin{aligned}A &= A \cap \mathbb{R}^n \\ &= A \cap \left( \bigcup_{i \in \mathbb{N}_0} ]-i, i[^n \right) \\ &= \bigcup_{i \in \mathbb{N}_0} (A \cap ]-i, i[^n) \\ &= \bigcup_{i \in \mathbb{N}_0} A_i \text{ where } A_i = A \cap ]-i, i[^n \in \mathcal{B}[\mathbb{R}^n] \wedge A_i \subseteq ]-i, i[^n\end{aligned}$$

Using 19.65 there exists a  $\{B_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{B}[\mathbb{R}^n]$  such that  $A = \bigcup_{i \in \mathbb{N}_0} A_i = \bigsqcup_{i \in \mathbb{N}_0} B_i$  with  $B_i \subseteq A_i \subseteq ]-i, i[^n$ . Hence we have that

$$\begin{aligned}\mu(A) &= \mu\left(\bigsqcup_{i \in \mathbb{N}_0} B_i\right) \\ &= \sum_{i=1}^{\infty} \mu(B_i) \\ &\stackrel{19.215}{=} \sum_{i=1}^{\infty} \lambda^n(B_i) \\ &= \lambda^n\left(\bigsqcup_{i \in \mathbb{N}_0} B_i\right) \\ &= \lambda^n(A)\end{aligned}$$

So  $\forall A \in \mathcal{B}[\mathbb{R}^n]$  we have  $\mu(A) = \lambda^n(A)$  proving that

$$\mu = (\lambda^n)|_{\mathcal{B}[\mathbb{R}^n]}$$

□

**Definition 19.94.** Let  $n \in \mathbb{N}_0$ ,  $A \subseteq \mathbb{R}^n$  and  $a \in \mathbb{R}^n$  then  $A + a = \{y + a \mid a \in \mathbb{R}^n\}$

**Proposition 19.95.** Let  $n \in \mathbb{N}_0$  and  $a \in \mathbb{R}^n$  then

1. If  $A \subseteq \mathbb{R}^n$  then  $x \in A + a$  if and only if  $x - a \in A$
2. If  $A \subseteq \mathbb{R}^n$  then  $(A + a) + (-a) = A$
3. If  $A, B \subseteq \mathbb{R}^n$  then  $A \subseteq B \Leftrightarrow A + a \subseteq B + a$
4. If  $\{A_i\}_{i \in I} \subseteq \mathcal{P}(\mathbb{R}^n)$  and  $\bigcup_{i \in I} A_i$  then  $(\bigcup_{i \in I} A_i) + a \subseteq \bigcup_{i \in I} (A_i + a)$
5. If  $A, B \subseteq \mathbb{R}^n$  then  $(A + a) \cap B = (A \cap (B + (-a))) + a$
6. If  $A, B \in \mathbb{R}^n$  then  $(A + a) \setminus B = (A \setminus (B + (-a))) + a$
7. If  $a, b, c \in \mathbb{R}^n$  with  $a \leq b$  then  $[a, b] + c = [a + c, b + c]$

**Proof.**

1.  $\Rightarrow$ . If  $x \in A + a$  then  $\exists y \in A$  such that  $x = y + a$  hence  $x - a = y \in A$   
 $\Leftarrow$ . If  $x - a \in A$  then  $x = (x - a) + a \in A + a$
2. If  $x \in (A + a) + (-a)$  then  $\exists y \in A + a$  such that  $x = y + (-a)$ . As  $y \in A + a$  there exists a  $z \in A$  such that  $y = z + a$ . Hence  $x = (z + a) + (-a) = z \in A$ . Further if  $x \in A$  then  $x + a \in A + a$  hence  $x = (x + a) + (-a) \in (A + a) + (-a)$
3.  $\Rightarrow$ . If  $x \in A + a$  then there exists a  $y \in A \subseteq B$  such that  $x = y + a$  which as  $y \in B$  proves that  $x \in B + a$  proving that  $A + a \subseteq B + a$   
 $\Leftarrow$ . If  $A \stackrel{(2)}{=} (A + a) + (-a) \subseteq (B + a) + (-a) \stackrel{(2)}{=} B$
4. If  $x \in (\bigcup_{i \in I} A_i) + a$  then there exists a  $y \in \bigcup_{i \in I} A_i$  such that  $x = y + a$ . So there exists a  $i \in I$  such that  $y \in A_i$  hence  $x \in A_i + a$  proving that  $x \in \bigcup_{i \in I} (A_i + a)$
- 5.

$$\begin{aligned}x \in (A + a) \bigcap B &\Leftrightarrow x \in A + a \wedge x \in B \\ &\stackrel{(1)}{\Leftrightarrow} (x - a) \in A \wedge (x - a) - (-a) \in B \\ &\stackrel{(1)}{\Leftrightarrow} (x - a) \in A \wedge (x - a) \in B + (-a) \\ &\Leftrightarrow (x - a) \in A \bigcap (B + (-a)) \\ &\stackrel{(1)}{\Leftrightarrow} x \in (A \bigcap (B + (-a))) + a\end{aligned}$$

6.

$$\begin{aligned}
x \in (A + a) \setminus B &\Leftrightarrow x \in A + a \wedge x \notin B \\
&\stackrel{(1)}{\Leftrightarrow} x - a \in A \wedge (x - a) + a \notin B \\
&\Leftrightarrow (x - a) \in A \wedge (x - a) \notin (B + (-a)) \\
&\Leftrightarrow (x - a) \in (A \bigcap (B + (-a))) \\
&\Leftrightarrow x \in (A \bigcap (B + (-a)))
\end{aligned}$$

7. We have

$$\begin{aligned}
x \in [a, b] + c &\stackrel{(1)}{\Leftrightarrow} x - c \in [a, b] \\
&\Leftrightarrow \forall i \in \{1, \dots, n\} a_i \leq x_i - c_i < b_i \\
&\Leftrightarrow \forall i \in \{1, \dots, n\} a_i + c_i \leq x_i < b_i + c_i \\
&\Leftrightarrow x \in [a + c, b + c]
\end{aligned}$$

□

We proceed now to prove that  $\varphi^n$  is translation invariant. First we prove this for  $v^n$  and half open rectangles.

**Lemma 19.96.** *Let  $n \in \mathbb{N}_0$ ,  $R \in \mathcal{R}^n$  and  $a \in \mathbb{R}^n$  then  $v^n(R) = v^n(R + a)$*

**Proof.** Let  $R \in \mathcal{R}^n$  then  $\exists c, d \in \mathbb{R}^n$  with  $c \leq d$  such that  $R = [c, d]$ . Then we have  $R + a \stackrel{19.95}{=} [c + a, d + a]$  hence  $v^n(R + a)$  exists and

$$\begin{aligned}
v^n(R + a) &= v^n([c + a, d + a]) \\
&= \prod_{i=1}^n (d_i + a_i - c_i - a_i) \\
&= \prod_{i=1}^n (d_i - c_i) \\
&= v^n(R)
\end{aligned}$$

□

**Lemma 19.97.** *Let  $n \in \mathbb{N}_0$ , then  $\forall A \subseteq \mathbb{R}^n$  and  $\forall a \in \mathbb{R}^n$  we have  $\varphi^n(A) = \varphi^n(A + a)$*

**Proof.** Given  $A \subseteq \mathbb{R}^n$  define  $\mathcal{C}_A = \{\sum_{i=1}^{\infty} v^n(R_i) \mid \{R_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{R}^n$  such that  $A \subseteq \bigcup_{i \in \mathbb{N}_0} R_i\}$  then if  $x \in \mathcal{C}_A$  there exists a  $\{R_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{R}^n$  with  $A \subseteq \bigcup_{i \in \mathbb{N}_0} R_i$  such that  $x = \sum_{i=1}^{\infty} v^n(R_i)$ . Using 19.96 we have that  $x = \sum_{i=1}^{\infty} v^n(R_i + a)$ , further by 19.95 we have  $A + a \subseteq \bigcup_{i \in \mathbb{N}_0} (R_i + a)$  hence  $x \in \mathcal{C}_{(A+a)}$ . So

$$\forall A \subseteq \mathbb{R}^n, \forall a \in \mathbb{R}^n \text{ we have } \mathcal{C}_A \subseteq \mathcal{C}_{(A+a)} \quad (19.216)$$

Next we have  $\mathcal{C}_{A+a} \subseteq_{19.216} \mathcal{C}_{(A+a)+(-a)} \stackrel{19.95}{=} \mathcal{C}_A$ . Using 19.216 again we conclude that  $\mathcal{C}_A = \mathcal{C}_{a+a}$  hence we have that

$$\varphi^n(A) = \inf(\mathcal{C}_A) = \inf(\mathcal{C}_{A+a}) = \varphi^n(A + a) \quad \square$$

Next we prove that every Lebesgue measurable set is translation invariant.

**Lemma 19.98.** *Let  $n \in \mathbb{N}_0$  then  $\forall a \in \mathbb{R}^n$  and  $A \in \mathcal{P}(\mathbb{R}^n)$  then  $A \in \mathcal{L}[\mathbb{R}^n]$  if and only if  $A + a \in \mathcal{L}[\mathbb{R}^n]$*

**Proof.** Take  $a \in \mathbb{R}^n$ . Assume that  $A \in \mathcal{L}[\mathbb{R}^n]$  then

$$\forall B \subseteq \mathbb{R}^n \text{ we have } \varphi^n(B) = \varphi^n(B \bigcap A) + \varphi^n(B \setminus A) \quad (19.217)$$

Hence for  $B \subseteq \mathbb{R}^n$  we have

$$\begin{aligned} \varphi^n(B \cap (A + a)) + \varphi^n(B \setminus (A + a)) &\stackrel{19.97}{=} \varphi^n((B \cap (A + a)) + (-a)) + \varphi^n((B \setminus (A + a)) + (-a)) \\ &\stackrel{19.95}{=} \varphi^n((B + (-a)) \cap A) + \varphi^n((B + (-a)) \setminus A) \\ &\stackrel{19.217}{=} \varphi^n(B) \end{aligned}$$

which proves that  $A + a \in \mathcal{L}[\mathbb{R}^n]$ . If on the other hand  $A + a \in \mathcal{L}[\mathbb{R}^n]$  then  $(A + a) + (-a) \in \mathcal{L}[\mathbb{R}^n]$  which by 19.95 proves that  $A \in \mathcal{L}[\mathbb{R}^n]$ .  $\square$

To summarize 19.97 and 19.98 we have the following theorem.

**Theorem 19.99.** *Let  $n \in \mathbb{N}_0$  then given  $a \in \mathbb{R}^n$  we have*

1.  $A \in \mathcal{L}[\mathbb{R}^n]$  if and only if  $A + a \in \mathcal{L}[\mathbb{R}^n]$
2.  $\forall A \in \mathcal{P}(\mathbb{R}^n)$  we have  $\varphi^n(A + a) = \varphi^n(A)$  which as  $\lambda^n = (\varphi^n)|_{\mathcal{L}[\mathbb{R}^n]}$  means that  $\forall A \in \mathcal{L}[\mathbb{R}^n]$   $\lambda^n(A + a) = \lambda^n(A)$

**Proof.** See 19.97 and 19.98.  $\square$

We prove now that every set in  $\mathcal{B}[\mathbb{R}^n]$  is translation invariant.

**Lemma 19.100.** *Let  $n \in \mathbb{N}_0$  then  $\forall A \in \mathcal{B}[\mathbb{R}^n]$  we have that  $A + a \in \mathcal{B}[\mathbb{R}^n]$*

**Proof.** First given  $a \in \mathbb{R}^n$  we will prove that  $\mathcal{S}_a = \{A \in \mathcal{P}[\mathbb{R}^n] | A + a \in \mathcal{B}[\mathbb{R}^n]\}$  is a  $\sigma$ -algebra.

1. As  $\mathbb{R}^n + a = \mathbb{R}^n \in \mathcal{B}[\mathbb{R}^n]$  [ $\mathbb{R}^n$  is a vector space] we have that  $\mathbb{R}^n \in \mathcal{S}_a$
2. Let  $A \in \mathcal{S}_a$  then  $A + a \in \mathcal{B}[\mathbb{R}^n]$ . Now

$$\begin{aligned} x \in (\mathbb{R}^n \setminus A) + a &\stackrel{19.95}{\Leftrightarrow} (x - a) \in \mathbb{R}^n \setminus A \\ &\Leftrightarrow (x - a) \in \mathbb{R}^n \wedge (x - a) \notin A \\ &\stackrel{\mathbb{R}^n \text{ is a vectorspace}}{\Leftrightarrow} x \in \mathbb{R}^n \wedge (x - a) \notin A \\ &\Leftrightarrow x \in \mathbb{R}^n \wedge x \notin (A + a) \\ &\Leftrightarrow x \in \mathbb{R}^n \setminus (A + a) \end{aligned}$$

proving that  $(\mathbb{R}^n \setminus A) + a = (\mathbb{R}^n \setminus (A + a)) \in \mathcal{B}[\mathbb{R}^n]$ , hence  $(\mathbb{R}^n \setminus A) \in \mathcal{S}_a$ .

3. If  $\{A_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{S}_a$  then  $\forall i \in \mathbb{N}_0$  we have  $A_i + a \in \mathcal{B}[\mathbb{R}^n]$ . hence

$$\left( \bigcup_{i \in \mathbb{N}_0} A_i \right) + a \stackrel{19.95}{=} \left( \bigcup_{i \in \mathbb{N}_0} (A_i + a) \right) \in \mathcal{B}[\mathbb{R}^n]$$

proving that  $\bigcup_{i \in \mathbb{N}_0} A_i \in \mathcal{S}_a$

Hence

$$\forall a \in \mathbb{R}^n \text{ we have that } \mathcal{S}_a \text{ is a } \sigma\text{-algebra} \quad (19.218)$$

Take now  $U \in \mathcal{T}_{\mathbb{R}^n}$  then given  $x \in U + a$  we have that  $(x - a) \in U$  and as  $U$  is open there exists a  $\delta > 0$  such that  $(x - a) \in B_{\|\cdot\|_{\max}}((x - a), \delta) \subseteq U$  or

$$x \in B_{\|\cdot\|_{\max}}((x - a), \delta) + a \subseteq_{19.95} U + a \quad (19.219)$$

Now

$$\begin{aligned} y \in B_{\|\cdot\|_{\max}}(x, \delta) &\Leftrightarrow \|x - y\|_{\max} < \delta \\ &\Leftrightarrow \|(x - a) - (y - a)\|_{\max} < \delta \\ &\Leftrightarrow y - a \in B_{\|\cdot\|_{\max}}(x - a, \delta) \\ &\Leftrightarrow y \in B_{\|\cdot\|_{\max}}(x - a, \delta) + a \end{aligned}$$

combining the above with 19.219 gives  $x \in B_{\|\cdot\|_{\max}}(x, \delta) \subseteq U + a$  proving that  $U + a \in \mathcal{T}_{\mathbb{R}} \subseteq \mathcal{B}[\mathbb{R}^n]$ . Hence we have that  $\mathcal{T}_{\mathbb{R}^n} \subseteq \mathcal{S}_a$ . As  $\mathcal{B}[\mathbb{R}^n] = \sigma(\mathcal{T}_{\mathbb{R}^n})$  is the smallest  $\sigma$ -algebra covering  $\mathcal{T}_{\mathbb{R}}$  we have that  $\mathcal{B}(\mathbb{R}^n) \subseteq \mathcal{S}_a$  or  $\forall A \in \mathcal{B}[\mathbb{R}^n]$  we have  $A + a \in \mathcal{B}[\mathbb{R}^n]$ . Hence

$$\forall a \in \mathbb{R}^n \text{ and } \forall A \in \mathcal{B}[\mathbb{R}^n] \text{ we have } A + a \in \mathcal{B}[\mathbb{R}^n] \quad \square$$

Next we will show that translation invariant measures on  $\mathcal{B}[\mathbb{R}^n]$  are identical to the Lebesgue measure multiplied by a strict positive constant. First we prove how  $[0, 1[$  can be written as a disjoint union of Dyadic intervals and  $[0, 1[^n$  as a disjoint union of Dyadic cubes

**Lemma 19.101.** *Let  $m \in \mathbb{N}$  then  $[0, 1[ = \bigsqcup_{i \in \{0, \dots, 2^m - 1\}} \left[ \frac{i}{2^m}, \frac{i+1}{2^m} \right[$*

**Proof.** First if  $i \in \{0, \dots, 2^m - 1\}$  then  $0 \leq \frac{i}{2^m} < \frac{i+1}{2^m} \leq \frac{(2^m-1)+1}{2^m} = 1$  proving that  $\left[ \frac{i}{2^m}, \frac{i+1}{2^m} \right[ \subseteq [0, 1[$ . Hence

$$\bigcup_{i \in \{0, \dots, 2^m - 1\}} \left[ \frac{i}{2^m}, \frac{i+1}{2^m} \right[ \subseteq [0, 1[ \quad (19.220)$$

If  $x \in [0, 1[$  then  $0 \leq x < 1$ . So  $0 \leq 2^m \cdot x$  we have using 9.56 that there exists a  $n \in \mathbb{N}$  such that  $n \leq 2^m \cdot x < n + 1$ . After division by  $2^m > 0$  we have

$$\frac{n}{2^m} \leq x < \frac{n+1}{2^m}$$

If  $2^m - 1 < n$  then  $2^m < n + 1$  or  $1 < \frac{n+1}{2^m} < x$  contradicting  $x < 1$ , hence  $n \in \{1, \dots, 2^m - 1\}$  proving that  $x \in \bigcup_{i \in \{1, \dots, 2^m - 1\}} \left[ \frac{i}{2^m}, \frac{i+1}{2^m} \right[$ . So  $[0, 1[ \subseteq \bigcup_{i \in \{1, \dots, m\}} \left[ \frac{i}{2^m}, \frac{i+1}{2^m} \right[$  which together with 19.220 gives

$$[0, 1[ = \bigcup_{i \in \{0, \dots, 2^m - 1\}} \left[ \frac{i}{2^m}, \frac{i+1}{2^m} \right[ \quad (19.221)$$

Finally if  $i, j \in \{0, \dots, 2^m - 1\}$  with  $i \neq j$  then we can assume that  $i < j$  [otherwise interchange  $i$  and  $j$ ] so that  $i + 1 \leq j$ . Then if  $x \in \left[ \frac{i}{2^m}, \frac{i+1}{2^m} \right[ \cap \left[ \frac{j}{2^m}, \frac{j+1}{2^m} \right[$  we have  $\frac{i}{2^m} \leq x < \frac{i+1}{2^m} \leq \frac{j}{2^m} \leq x < \frac{j+1}{2^m}$  giving the contradiction  $x < x$ . Hence  $\left[ \frac{i}{2^m}, \frac{i+1}{2^m} \right[ \cap \left[ \frac{j}{2^m}, \frac{j+1}{2^m} \right[ = \emptyset$  or using 19.221

$$[0, 1[ = \bigsqcup_{i \in \{0, \dots, 2^m - 1\}} \left[ \frac{i}{2^m}, \frac{i+1}{2^m} \right[ \quad \square$$

**Lemma 19.102.** *Let  $n \in \mathbb{N}_0$  and  $m \in \mathbb{N}$  then  $[0, 1[ = \bigsqcup_{\rho \in \{0, \dots, 2^m - 1\}^n} \left[ \frac{\rho}{2^m}, \frac{\rho+1}{2^m} \right[$*

**Proof.** If  $x \in \bigcup_{\rho \in \{0, \dots, 2^m - 1\}^n} \left[ \frac{\rho}{2^m}, \frac{\rho+1}{2^m} \right[$  then  $\exists \rho \in \{0, \dots, 2^m - 1\}^n$  such that  $\forall i \in \{1, \dots, n\}$  we have  $\rho_i \in \{0, \dots, 2^m - 1\}$  and  $x_i \in \left[ \frac{\rho_i}{2^m}, \frac{\rho_i+1}{2^m} \right[$ , applying then 19.101 gives  $x_i = [0, 1[$ . From this it follows that

$$\bigcup_{\rho \in \{0, \dots, 2^m - 1\}^n} \left[ \frac{\rho}{2^m}, \frac{\rho+1}{2^m} \right[ \subseteq [0, 1[ \quad (19.222)$$

If  $x \in [0, 1[^n$  then  $\forall i \in \{1, \dots, n\}$   $x_i \in [0, 1[$   $\stackrel{19.101}{=} \bigsqcup_{j \in \{0, \dots, 2^m - 1\}} \left[ \frac{j}{2^m}, \frac{j+1}{2^m} \right[$  so there exists a  $\rho_i \in \{0, \dots, 2^m - 1\}$  such that  $x_i \in \left[ \frac{\rho_i}{2^m}, \frac{\rho_i+1}{2^m} \right[$ . If we define  $\rho \in \{0, \dots, 2^m - 1\}^n$  by  $\rho = (\rho_1, \dots, \rho_n)$  then  $x \in \prod_{i \in \{1, \dots, n\}} \left[ \frac{\rho_i}{2^m}, \frac{\rho_i+1}{2^m} \right[ = \left[ \frac{\rho}{2^m}, \frac{\rho+1}{2^m} \right[$  proving that  $[0, 1[^n \subseteq \bigcup_{\rho \in \{0, \dots, 2^m - 1\}^n} \left[ \frac{\rho}{2^m}, \frac{\rho+1}{2^m} \right[$ . Combining this with 19.222 proves that

$$[0, 1[ = \bigsqcup_{\rho \in \{0, \dots, 2^m - 1\}^n} \left[ \frac{\rho}{2^m}, \frac{\rho+1}{2^m} \right[ \quad (19.223)$$

Finally if  $\exists \rho, \sigma \in \{0, \dots, 2^m - 1\}^n$  with  $\rho \neq \sigma$  then  $\exists i \in \{1, \dots, n\}$  such that  $\rho_i \neq \sigma_i$ . If now  $x \in \left[ \frac{\rho}{2^m}, \frac{\rho+1}{2^m} \right[ \cap \left[ \frac{\sigma}{2^m}, \frac{\sigma+1}{2^m} \right[$  then  $x_i \in \left[ \frac{\rho_i}{2^m}, \frac{\rho_i+1}{2^m} \right[ \cap \left[ \frac{\sigma_i}{2^m}, \frac{\sigma_i+1}{2^m} \right[ \stackrel{\rho_i \neq \sigma_i \wedge 19.101}{=} \emptyset$  leading to a contradiction. So we must have that  $\left[ \frac{\rho}{2^m}, \frac{\rho+1}{2^m} \right[ \cap \left[ \frac{\sigma}{2^m}, \frac{\sigma+1}{2^m} \right[ = \emptyset$  proving finally that

$$[0, 1[^n = \bigsqcup_{\rho \in \{0, \dots, 2^m - 1\}^n} \left[ \frac{\rho}{2^m}, \frac{\rho+1}{2^m} \right[$$

□

**Lemma 19.103.** Let  $n \in \mathbb{N}_0$  then and  $\langle \mathbb{R}^n, \mathcal{B}[\mathbb{R}^n], \mu \rangle$  be a measure space such that

1.  $\mu$  is non zero  $[\exists A \in \mathcal{B}[\mathbb{R}^n] \text{ such that } 0 < \mu(A)]$
2. If  $A \in \mathcal{B}[\mathbb{R}^n]$  is bounded then  $\mu(A) < \infty$
3. If  $A \in \mathcal{B}[\mathbb{R}^n]$  and  $a \in \mathbb{R}^n$  then  $\mu(A+a) = \mu(A)$  [note that this make sense because of 19.100]

then for  $c = \mu([0, 1]^n)$  we have that  $0 < c$  and  $\mu = c \cdot (\lambda^n)|_{\mathcal{B}[\mathbb{R}^n]}$  [or  $\forall A \in \mathcal{B}[\mathbb{R}^n]$  we have  $\mu(A) = c \cdot \lambda^n(A)$ ]

**Proof.** First if  $m \in \mathbb{N}$  and  $z \in \mathbb{Z}$  we have

$$\left[ \frac{z}{2^m}, \frac{z+1}{2^m} \right]^n = \left[ \frac{z}{2^m} + 0, \frac{z}{2^m} + \left( \frac{1}{2^m} \right) \right]^n \stackrel{19.95}{=} \left[ 0, \frac{1}{2^m} \right]^n + \left( \underbrace{\frac{z}{2^m}, \dots, \frac{z}{2^m}}_n \right) \quad (19.224)$$

Hence

$$\mu \left( \left[ \frac{z}{2^m}, \frac{z+1}{2^m} \right]^n \right) = \mu \left( \left[ 0, \frac{1}{2^m} \right]^n + \left( \underbrace{\frac{1}{2^m}, \dots, \frac{1}{2^m}}_n \right) \right) \quad (19.225)$$

$$\stackrel{(3), \left[ \frac{z}{2^m}, \frac{z+1}{2^m} \right]^n \in \mathcal{B}[\mathbb{R}^n]}{=} \mu \left( \left[ 0, \frac{1}{2^m} \right]^n \right) \quad (19.225)$$

$$\leq \left( \frac{1}{2^m} \right) \leq (1) \quad \mu([0, 1]^n) \quad (19.226)$$

As  $\mathbb{R}^n$  is open we have by 19.50 that there exists a  $\{D_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{D}^m$  such that  $\mathbb{R}^n \subseteq \bigcup_{i \in \mathbb{N}_0} D_i$ .

Let  $i \in \mathbb{N}_0$  then  $\exists m_i \in \mathbb{N}$  and  $z_i$  such that  $D_i = \left[ \frac{z}{2^m}, \frac{z+1}{2^m} \right]^n \stackrel{19.224}{=} \left[ 0, \frac{1}{2^m} \right]^n + \left( \underbrace{\frac{z}{2^m}, \dots, \frac{z}{2^m}}_n \right)$  so that

$\mu(D_i) = \mu \left( \left[ 0, \frac{1}{2^m} \right]^n + \left( \underbrace{\frac{z}{2^m}, \dots, \frac{z}{2^m}}_n \right) \right) \leq 19.226 \mu([0, 1]^n)$ . Hence if  $\mu([0, 1]^n) = 0$  we have  $\mu(D_i) = 0$  then  $\mu(\mathbb{R}^n) \leq \mu(\bigcup_{i \in \mathbb{N}_0} D_i) \leq \sum_{i=1}^{\infty} \mu(D_i) = 0$  or for every  $A \in \mathcal{B}[\mathbb{R}^n]$  we have  $\mu(A) \leq \mu(\mathbb{R}^n) = 0$  contradicting the fact that  $\mu$  is non zero. So we have

$$c = \mu([0, 1]^n) > 0 \quad (19.227)$$

Define now

$$\nu: \mathcal{B}[\mathbb{R}^n] \rightarrow [0, \infty] \text{ by } \nu(A) = \frac{1}{c} \cdot \mu(A)$$

then we have  $\nu(\emptyset) = \frac{1}{c} \cdot \mu(\emptyset) = 0$  and if  $\{A_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{B}[\mathbb{R}^n]$  is pairwise disjoint then

$$\begin{aligned} \nu \left( \bigsqcup_{i \in \mathbb{N}_0} A_i \right) &= \frac{1}{c} \cdot \mu \left( \bigsqcup_{i \in \mathbb{N}_0} A_i \right) \\ &= \frac{1}{c} \cdot \sum_{i=1}^{\infty} \mu(A_i) \\ &= \sum_{i=1}^{\infty} \frac{1}{c} \cdot \mu(A_i) \\ &= \sum_{i=1}^{\infty} \nu(A_i) \end{aligned}$$

Further if  $A \in \mathcal{B}[\mathbb{R}^n]$ ,  $a \in \mathbb{R}^n$  then  $\nu(A + \varepsilon) = \frac{1}{c} \cdot \mu(A + a) = \frac{1}{c} \cdot \mu(A)$ . To summarize we have that

$$\nu: \mathcal{B}[\mathbb{R}^n] \rightarrow [0, \infty] \text{ is a translation invariant measure on } \mathcal{B}[\mathbb{R}^n] \quad (19.228)$$

Next we have

$$\begin{aligned}
1 &= \frac{1}{c} \cdot c \\
&= \frac{1}{c} \cdot \mu([0, 1]) \\
&= \nu([0, c]) \\
&\stackrel{19.102}{=} \nu\left(\bigsqcup_{\rho \in \{0, \dots, 2^m-1\}^n} \left[\frac{\rho}{2^m}, \frac{\rho+1}{2^m}\right]\right) \\
&= \sum_{\rho \in \{0, \dots, 2^m-1\}^n} \nu\left(\left[\frac{\rho}{2^m}, \frac{\rho+1}{2^m}\right]\right) \\
&\stackrel{19.225}{=} \sum_{\rho \in \{0, \dots, 2^m-1\}^n} \nu\left(\left[0, \frac{1}{2^m}\right]^n\right) \\
&= \#(\{0, \dots, 2^m-1\}) \cdot \nu\left(\left[0, \frac{1}{2^m}\right]^n\right)
\end{aligned}$$

proving that

$$\nu\left(\left[0, \frac{1}{2^m}\right]^n\right) = \frac{1}{\#(\{0, \dots, 2^m-1\})} \quad (19.229)$$

Likewise

$$\begin{aligned}
1 &= \prod_{i=1}^n (1-0) \\
&= \lambda^n([0, 1]^n) \\
&\stackrel{19.102}{=} \lambda^n\left(\bigsqcup_{\rho \in \{0, \dots, 2^m-1\}^n} \left[\frac{\rho}{2^m}, \frac{\rho+1}{2^m}\right]\right) \\
&= \sum_{\rho \in \{0, \dots, 2^m-1\}^n} \lambda^n\left(\left[\frac{\rho}{2^m}, \frac{\rho+1}{2^m}\right]\right) \\
&= \sum_{\rho \in \{0, \dots, 2^m-1\}^n} \lambda^n\left(\left[\frac{\rho}{2^m}, \frac{\rho+1}{2^m}\right]\right) \\
&= \sum_{\rho \in \{0, \dots, 2^m-1\}^n} \left(\frac{1}{2^m}\right) \\
&= \#(\{0, \dots, 2^m-1\}) \cdot \frac{1}{2^m} \\
&= \#(\{0, \dots, 2^m-1\}) \cdot \lambda^n\left(\left[0, \frac{1}{2^m}\right]^n\right)
\end{aligned}$$

proving that

$$\lambda^n\left(\left[0, \frac{1}{2^m}\right]^n\right) = \frac{1}{\#(\{0, \dots, 2^m-1\})} \quad (19.230)$$

Finally If  $D \in \mathcal{D}^n$  then there exists a  $z \in \mathbb{Z}^n$ ,  $m \in \mathbb{N}$  so that  $D = \left[\frac{z}{2^m}, \frac{z+1}{2^m}\right]$  then we have

$$\begin{aligned}
\nu(D) &= \nu\left(\left[\frac{z}{2^m}, \frac{z+1}{2^m}\right]\right) \\
&\stackrel{19.225}{=} \nu\left(\left[0, \frac{1}{2^m}\right]^n\right) \\
&\stackrel{19.229, 19.230}{=} \lambda^n\left(\left[0, \frac{1}{2^m}\right]^n\right) \\
&\stackrel{19.99}{=} \lambda^n\left(\left[\frac{z}{2^m}, \frac{z+1}{2^m}\right]\right) \\
&= \lambda^n(D)
\end{aligned}$$

So  $\forall D \in \mathcal{D}^n$  we have  $\nu(D) = \lambda^n(D)$ . Applying 19.93 proves then that  $\forall A \in \mathcal{B}[\mathbb{R}^n]$   $\nu(A) = \lambda^n(A)$  or

$$\forall A \in \mathcal{B}[\mathbb{R}^n] \quad \mu(A) = c \cdot \lambda^n(A) \quad \square$$

### 19.3.6 Negligibility, Completeness and Regularity

**Definition 19.104.** ( $\mu$ -negligible set) Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space then  $S \subseteq X$  is  $\mu$ -negligible iff there exists a  $A \in \mathcal{A}$  with  $\mu(A) = 0$  such that  $S \subseteq A$ . The set of all  $\mu$ -negligible sets is noted as  $\mathcal{N}_\mu$  so  $\mathcal{N}_\mu = \{S \in \mathcal{P}(X) \mid \exists A \in \mathcal{A} \text{ with } \mu(A) = 0 \text{ such that } S \subseteq A\}$ .

**Lemma 19.105.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space then we have

1.  $\emptyset \in \mathcal{N}_\mu$
2. If  $S \in \mathcal{N}_\mu$  then if  $T \subseteq S$  we have  $T \in \mathcal{N}_\mu$
3. If  $\{S_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{N}_\mu$  then  $\bigcup_{i \in \mathbb{N}_0} S_i \in \mathcal{N}_\mu$  and  $\bigcap_{i \in \mathbb{N}_0} S_i \in \mathcal{N}_\mu$

**Proof.**

1. This is trivial as  $\emptyset \subseteq \emptyset$  and  $\mu(\emptyset) = 0$
2. As  $S$  is  $\mu$ -negligible there exists a  $A \in \mathcal{A}$  with  $\mu(A) = 0$  such that  $S \subseteq A$  which as  $T \subseteq S \Rightarrow T \subseteq A$  proves that  $T$  is  $\mu$ -negligible.
3. As  $\{S_i\}_{i \in \mathbb{N}} \subseteq \mathcal{N}_\mu$  there exists a  $\{A_i\}_{i \in \mathbb{N}} \subseteq \mathcal{A}$  with  $\forall i \in \mathbb{N} S_i \subseteq A_i$  and  $\mu(A_i) = 0$ . Further using 19.66 we have that  $\mu(\bigcup_{i \in \mathbb{N}_0} A_i) \leq \sum_{i=1}^{\infty} \mu(A_i) = 0$  which as  $\bigcup_{i \in \mathbb{N}_0} S_i \subseteq \bigcup_{i \in \mathbb{N}_0} A_i$  proves that

$$\bigcup_{i \in \mathbb{N}_0} S_i \in \mathcal{N}_\mu$$

Finally as  $\bigcap_{i \in \mathbb{N}_0} S_i \subseteq S_0$  it follows from (2) that  $\bigcap_{i \in \mathbb{N}_0} S_i \in \mathcal{N}_\mu$   $\square$

**Definition 19.106.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space then  $\mu$  is **complete** if every  $\mu$ -negligible set is in  $\mathcal{A}$ . In other words  $\mu$  is **complete** if  $\mathcal{N}_\mu \subseteq \mathcal{A}$ . If the measure of  $\langle X, \mathcal{A}, \mu \rangle$  is **complete** then  $\langle X, \mathcal{A}, \mu \rangle$  is called a **complete measure space**.

**Proposition 19.107.** Let  $X$  be a set and  $\varphi$  a outer measure on  $X$  then  $\langle X, \mathcal{M}_\varphi[X], \varphi|_{\mathcal{M}_\varphi[X]} \rangle$  [see 19.75] is a complete measure space.

**Proof.** If  $S \in \mathcal{N}_{\varphi|_{\mathcal{M}_\varphi[X]}}$  then  $\exists A \in \mathcal{M}_\varphi[X]$  with  $\varphi(A) = \varphi|_{\mathcal{M}_\varphi[X]}(A) = 0$  and  $S \subseteq A$ . Hence  $\varphi(S) \leq \varphi(A)$  so that  $\varphi(S) = 0$ . Applying then 19.74 proves that  $S \in \mathcal{N}_{\varphi|_{\mathcal{M}_\varphi[X]}}$ .  $\square$

Using the definition of Lebesgue measure spaces [see 19.81 and 19.86] we have then that

**Corollary 19.108.**  $\langle \mathbb{R}, \mathcal{L}[\mathbb{R}], \lambda \rangle$  and  $\langle \mathbb{R}^n, \mathcal{L}[\mathbb{R}^n], \lambda^n \rangle$ ,  $n \in \mathbb{N}_0$  are complete measure spaces.

We proceed now to show how to extend a measure space to a complete measure space.

**Lemma 19.109.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space,  $A \subseteq X$  and  $E, F \in \mathcal{A}$  such that  $E \subseteq A \subseteq F$  and  $\mu(F \setminus E) = 0$  then

$$\mu(F) = \mu(E) = \sup(\{\mu(B) \mid B \in \mathcal{A} \text{ and } B \subseteq A\})$$

**Proof.** As  $E \subseteq F$  we have that  $F = (F \setminus E) \sqcup E$  [see 1.28] hence  $\mu(F) = \mu(F \setminus E) + \mu(E) = 0 + \mu(E)$  giving

$$\mu(E) = \mu(F) \quad (19.231)$$

Further if  $B \in \mathcal{A}$  with  $B \subseteq A$  then  $B \subseteq F$  so  $\mu(B) \leq \mu(F) = \mu(E)$  hence  $\sup(\{\mu(B) \mid B \in \mathcal{A} \wedge B \subseteq A\}) \leq \mu(E)$ . As also  $E \in \mathcal{A}$  and  $E \subseteq A$  we have  $\mu(E) \leq \sup(\{\mu(B) \mid B \in \mathcal{A} \wedge B \subseteq A\})$ . So we have

$$\mu(F) = \mu(E) = \sup(\{\mu(B) \mid B \in \mathcal{A} \wedge B \subseteq A\}) \quad \square$$

**Definition 19.110.** Let  $\langle X, \mathcal{A} \rangle$  be a measurable space and  $\mu: \mathcal{A} \rightarrow [0, \infty]$  a measure then  $\mathcal{A}_\mu$  is defined by  $\mathcal{A}_\mu = \{A \subseteq X \mid \exists E, F \in \mathcal{A} \text{ such that } E \subseteq A \subseteq F \text{ and } \mu(F \setminus E) = 0\}$ . The set  $\mathcal{A}_\mu$  is called the completion of  $\mathcal{A}$  under  $\mu$ . Further we define

$$\bar{\mu}: \mathcal{A}_\mu \rightarrow [0, \infty] \text{ by } \bar{\mu}(A) = \mu(E) = \mu(F) \text{ where } E, F \in \mathcal{A}, \mu(F \setminus E) = 0 \text{ and } E \subseteq A \subseteq F$$

We must of course prove that this definition is independent of the choice of  $E, F$

**Proof.** Let  $E', F' \in \mathcal{A}$  be another pair of sets with  $E' \subseteq A \subseteq F'$  and  $\mu(F' \setminus E') = 0$  then we have

$$\begin{aligned} \mu(E') &\stackrel{19.109}{=} \mu(F') \\ &\stackrel{19.109}{=} \sup(\{\mu(B) \mid B \in \mathcal{A} \wedge B \subseteq A\}) \\ &\stackrel{19.109}{=} \mu(F) \\ &\stackrel{19.109}{=} \mu(E) \end{aligned}$$

□

**Proposition 19.111.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space then  $\mathcal{A}_\mu$  is a  $\sigma$ -algebra on  $X$  with  $\mathcal{A} \subseteq \mathcal{A}_\mu$ . Further  $\bar{\mu}: \mathcal{A}_\mu \rightarrow [0, \infty]$  is a measure with  $(\bar{\mu})|_{\mathcal{A}} = \mu$  and finally  $\langle X, \mathcal{A}_\mu, \bar{\mu} \rangle$  is complete. We call  $\langle X, \mathcal{A}_\mu, \mu \rangle$  the **completion** of  $\langle X, \mathcal{A}, \mu \rangle$ .

**Proof.** If  $A \in \mathcal{A}$  then  $A \subseteq A \subseteq A$  and  $\mu(A \setminus A) = \mu(\emptyset) = 0$  so that  $A \in \mathcal{A}_\mu$ , hence

$$\mathcal{A} \subseteq \mathcal{A}_\mu \quad (19.232)$$

As  $\emptyset \in \mathcal{A}$  we have by 19.232 that

$$\emptyset \in \mathcal{A}_\mu \quad (19.233)$$

If  $A \in \mathcal{A}_\mu$  then there exists  $E, F \in \mathcal{A}$  such that  $E \subseteq A \subseteq F$  and  $\mu(F \setminus E) = 0$ . Using 1.28 it follows that  $X \setminus F \subseteq X \setminus A \subseteq X \setminus E$  where as  $\mathcal{A}$  is a  $\sigma$ -algebra we have that  $X \setminus E, X \setminus F \in \mathcal{A}$ . Further

$$\begin{aligned} x \in (X \setminus E) \setminus (X \setminus F) &\Leftrightarrow x \in X \setminus E \wedge x \notin X \setminus F \\ &\Leftrightarrow (x \in X \wedge x \notin E) \wedge \neg(x \in X \wedge x \notin F) \\ &\Leftrightarrow (x \in X \wedge x \notin E) \wedge (x \notin X \vee x \in F) \\ &\Leftrightarrow (x \in X \wedge x \notin E \wedge x \notin F) \vee (x \in X \wedge x \notin E \wedge x \in F) \\ &\Leftrightarrow x \in X \wedge x \notin E \wedge x \in F \\ &\stackrel{F \subseteq X}{\Leftrightarrow} x \notin E \wedge x \in F \\ &\Leftrightarrow F \setminus E \end{aligned}$$

Hence  $(X \setminus E) \setminus (X \setminus F) = F \setminus E$  giving  $\mu((X \setminus E) \setminus (X \setminus F)) = \mu(F \setminus E) = 0$ . So we conclude that

$$\forall A \in \mathcal{A}_\mu \text{ we have } X \setminus A \in \mathcal{A}_\mu \quad (19.234)$$

Let  $\{A_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{A}_\mu$  then there exists  $\{E_i\}_{i \in \mathbb{N}_0}, \{F_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{A}$  such that  $\forall i \in \mathbb{N}_0 E_i \subseteq A_i \subseteq F_i$  and  $\mu(F_i \setminus E_i) = 0$ . Then we have

$$\bigcup_{i \in \mathbb{N}_0} E_i \subseteq \bigcup_{i \in \mathbb{N}_0} A_i \subseteq \bigcup_{i \in \mathbb{N}_0} F_i \quad (19.235)$$

where  $\bigcup_{i \in \mathbb{N}_0} E_i, \bigcup_{i \in \mathbb{N}_0} F_i \in \mathcal{A}$ . Further if  $x \in (\bigcup_{i \in \mathbb{N}_0} F_i) \setminus (\bigcup_{j \in \mathbb{N}_0} E_j)$  then  $x \in \bigcup_{i \in \mathbb{N}_0} F_i$  and  $x \notin \bigcup_{j \in \mathbb{N}_0} E_j$  hence  $\exists i \in \mathbb{N}_0$  such that  $x \in F_i$  and  $x \notin \bigcup_{j \in \mathbb{N}_0} E_j \Rightarrow x \notin E_i$  so that  $x \in F_i \setminus E_i$ , proving that  $(\bigcup_{i \in \mathbb{N}_0} F_i) \setminus (\bigcup_{j \in \mathbb{N}_0} E_j) \subseteq \bigcup_{i \in \mathbb{N}_0} (F_i \setminus E_i)$ . Hence

$$\mu\left(\left(\bigcup_{i \in \mathbb{N}_0} F_i\right) \setminus \left(\bigcup_{j \in \mathbb{N}_0} E_j\right)\right) \leq \mu\left(\bigcup_{i \in \mathbb{N}_0} (F_i \setminus E_i)\right) \leq \sum_{i \in \mathbb{N}_0} \mu(F_i \setminus E_i) = 0. \quad (19.236)$$

So we conclude that

$$\forall \{A_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{A}_\mu \text{ we have } \bigcup_{i \in \mathbb{N}_0} A_i \in \mathcal{A}_0 \quad (19.237)$$

From 19.233, 19.234 and 19.237 it follows that

$$\mathcal{A}_\mu \text{ is a } \sigma\text{-algebra} \quad (19.238)$$

Now as  $\emptyset \in \mathcal{A}$  and  $\emptyset \subseteq \emptyset \subseteq \emptyset$  we have that  $\bar{\mu}(\emptyset) = \mu(\emptyset) = 0$  proving that

$$\bar{\mu}(\emptyset) = 0 \quad (19.239)$$

If  $\{A_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{A}_\mu$  is a pairwise disjoint family of sets in  $\mathcal{A}_\mu$  then there exists  $\{E_i\}_{i \in \mathbb{N}_0}, \{F_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{A}$  such that  $\forall i \in \mathbb{N}_0 E_i \subseteq A_i \subseteq F_i$  and  $\mu(F_i \setminus E_i) = 0$ . Furthermore if  $i, j \in \mathbb{N}_0$  with  $i \neq j$  then  $E_i \cap E_j \subseteq A_i \cap A_j = \emptyset$  hence  $\{E_i\}_{i \in \mathbb{N}_0}$  is pairwise disjoint, So

$$\begin{aligned} \bar{\mu}\left(\bigsqcup_{i \in \mathbb{N}_0} A_i\right) &\stackrel{19.235, 19.236}{=} \mu\left(\bigsqcup_{i \in \mathbb{N}_0} E_i\right) \\ &\stackrel{\mu \text{ is a measure}}{=} \sum_{i=1}^{\infty} \mu(E_i) \\ &= \sum_{i=1}^{\infty} \bar{\mu}(E_i) \end{aligned}$$

The above together with 19.239 proves that  $\bar{\mu}$  is a measure, hence

$$\langle X, \mathcal{A}_\mu, \bar{\mu} \rangle \text{ is a measure space} \quad (19.240)$$

If  $A \in \mathcal{A}$  then as  $A \subseteq A \subseteq A$  and  $\mu(A \setminus A) = \mu(\emptyset) = 0$  we have that  $\bar{\mu}(A) = \mu(A)$  proving that

$$(\bar{\mu})_{|\mathcal{A}} = \mu \quad (19.241)$$

Let  $A \subseteq X$  such that  $\exists B \in \mathcal{A}_\mu$  with  $A \subseteq B$  and  $\bar{\mu}(B) = 0$  then there exists a  $E, F \in \mathcal{A}$  with  $E \subseteq B \subseteq F$  and  $\mu(F \setminus E) = 0$ , hence  $\mu(F) = \bar{\mu}(B) = 0$ . Next as  $\emptyset \subseteq A \subseteq B \subseteq F$  and  $\mu(F \setminus \emptyset) = \mu(F) = 0$  it follows that  $A \in \mathcal{A}_\mu$ . Hence

$$\langle X, \mathcal{A}_\mu, \bar{\mu} \rangle \text{ is complete} \quad \square$$

We will prove now that  $\langle \mathbb{R}^n, \mathcal{L}[\mathbb{R}^n], \lambda^n \rangle$  is the completion of  $\langle \mathbb{R}^n, \mathcal{B}[\mathbb{R}^n], (\lambda^n)_{|\mathcal{B}[\mathbb{R}^n]} \rangle$ . First we need a little lemma.

**Lemma 19.112.** *Let  $n \in \mathbb{N}_0$  and  $A \in \mathcal{L}[\mathbb{R}^n]$  [so  $A$  is Lebesgue measurable] then there exists  $E, F \in \mathcal{B}[\mathbb{R}^n]$  [so  $E, F$  are Borel sets] such that  $E \subseteq A \subseteq F$  and  $\lambda^n(F \setminus E) = 0$ .*

**Proof.** First we prove the lemma for the finite case. So let  $A \in \mathcal{L}[\mathbb{R}^n]$  be such that  $\lambda^n(A) < \infty$ . Take  $m \in \mathbb{N}_0$ . As  $\lambda^n(A) = \sup_{19.92} (\lambda^n(K) | K \subseteq A \wedge K \text{ is compact}) < \infty$ , hence  $\lambda^n(A) - \frac{1}{2 \cdot m} < \sup_{19.92} (\lambda^n(K) | K \subseteq A \wedge K \text{ is compact})$ , there exists a compact set  $K_m$  such that

$$K_m \subseteq A \text{ and } \lambda^n(A) - \frac{1}{2 \cdot m} < \lambda^n(K_m) \quad (19.242)$$

Further as  $\lambda^n(A) = \inf_{19.91} (\{\lambda^n(U) | A \subseteq U \wedge U \text{ open}\}) < \infty$ , hence  $\inf(\{\lambda^n(U) | A \subseteq U \wedge U \text{ open}\}) < \lambda^n(A) + \frac{1}{2 \cdot m}$ , there exists a  $U_m$  open such that

$$A \subseteq U_m \text{ and } \lambda^n(U_m) < \lambda^n(A) + \frac{1}{2 \cdot m} \quad (19.243)$$

Define now  $E = \bigcup_{m \in \mathbb{N}_0} K_m \in \mathcal{B}[\mathbb{R}^n]$  [compact sets are closed by 11.254] and  $F = \bigcap_{m \in \mathbb{N}_0} U_m \in \mathcal{B}[\mathbb{R}^n]$ . Then using 19.242, 19.243 it follows that

$$E \subseteq A \subseteq F \quad (19.244)$$

Further given  $m \in \mathbb{N}_0$  we have

$$\begin{aligned}
 \lambda^n(F \setminus E) &\leq_{F \subseteq U_m} \lambda^n(U_m \setminus E) \\
 &\leq_{K_m \subseteq E} \lambda^n(U_m \setminus K_m) \\
 &\stackrel{\lambda^n(K_m) \leq \lambda^n(A) < \infty \text{ and 19.64}}{\leq} \lambda^n(U_m) - \lambda^n(K_m) \\
 &\stackrel{\lambda^n(K_m) \leq \lambda^n(A) < \infty}{=} \lambda^n(U_m) - \lambda^n(A) + \lambda^n(A) - \lambda^n(K_m) \\
 &\stackrel{19.242, 19.243 \text{ and } \lambda^n(K_m) \leq \lambda^n(A) < \infty}{=} \frac{1}{2 \cdot m} + \frac{1}{2 \cdot m} \\
 &= \frac{1}{m}
 \end{aligned}$$

As  $m$  was chosen arbitrary we have by 9.57 that  $\lambda^n(F \setminus E) \leq 0$  hence we have (taking 19.244 in account) that

$$\text{If } A \in \mathcal{L}[\mathbb{R}^n] \text{ with } \lambda^n(A) < \infty \text{ then } \exists E, F \in \mathcal{B}[\mathbb{R}^n] \text{ such that } E \subseteq A \subseteq F \text{ and } \lambda^n(F \setminus E) = 0 \quad (19.245)$$

Next we prove the general case. Let  $A \in \mathcal{L}[\mathbb{R}^n]$  then as  $\mathbb{R}^n = \bigcup_{i \in \mathbb{N}_0} [-i, i]$  we have  $A = A \cap \mathbb{R}^n = A \cap (\bigcup_{i \in \mathbb{N}_0} [-i, i]) = \bigcup_{i \in \mathbb{N}_0} (A \cap [-i, i]) = \bigcup_{i \in \mathbb{N}_0} A_i$  where  $A_i = A \cap [-i, i] \in \mathcal{L}[\mathbb{R}^n]$ . As  $\lambda^n(A_i) \leq \lambda^n([-i, i]) = v^n([-i, i]) < \infty$  we have by 19.245 the existence of  $E_i, F_i \in \mathcal{B}[\mathbb{R}^n]$  with  $E_i \subseteq A_i \subseteq F_i$  and  $\lambda^n(F_i \setminus E_i) = 0$ . Take  $E = \bigcup_{i \in \mathbb{N}_0} E_i$  and  $F = \bigcup_{i \in \mathbb{N}_0} F_i$  then

$$E, F \in \mathcal{B}[\mathbb{R}^n] \wedge E \subseteq A \subseteq F \quad (19.246)$$

Further

$$\begin{aligned}
 \lambda^n(F \setminus E) &= \lambda^n\left(\left(\bigcup_{i \in \mathbb{N}_0} F_i\right) \setminus \left(\bigcup_{j \in \mathbb{N}_0} E_j\right)\right) \\
 &\stackrel{1.102}{=} \lambda^n\left(\bigcup_{i \in \mathbb{N}_0} \left(F_i \setminus \left(\bigcup_{j \in \mathbb{N}_0} E_j\right)\right)\right) \\
 &\leq \sum_{i=1}^{\infty} \lambda^n\left(F_i \setminus \left(\bigcup_{j \in \mathbb{N}_0} E_j\right)\right) \\
 &\leq \sum_{i=1}^{\infty} \lambda^n(F_i \setminus E_i) \\
 &= 0
 \end{aligned}$$

proving that

$$\lambda^n(F \setminus E) = 0 \quad (19.247)$$

The lemma is then proved by 19.246 and 19.247 □

**Theorem 19.113.** *Let  $n \in \mathbb{N}_0$  then  $\langle \mathbb{R}^n, \mathcal{L}[\mathbb{R}^n], \lambda^n \rangle$  is the completion of  $\langle \mathbb{R}^n, \mathcal{B}[\mathbb{R}^n], (\lambda^n)|_{\mathcal{B}[\mathbb{R}^n]} \rangle$*

**Proof.** Take  $\mu = (\lambda^n)|_{\mathcal{B}[\mathbb{R}^n]}$  then using 19.111 we have that

$$\mathcal{B}[\mathbb{R}^n] \subseteq (\mathcal{B}[\mathbb{R}^n])_{|\mu} \stackrel{\text{def}}{=} \{A \subseteq X \mid \exists E, F \in \mathcal{B}[\mathbb{R}^n] \text{ such that } E \subseteq A \subseteq F \text{ and } \mu(F \setminus E) = 0\}$$

If  $A \in \mathcal{L}[\mathbb{R}^n]$  then by 19.112 there exists  $E, F \in \mathcal{B}[\mathbb{R}^n]$  with  $E \subseteq A \subseteq F$  and  $0 = \lambda^n(F \setminus E) = (\lambda^n)|_{\mathcal{B}[\mathbb{R}^n]}(F \setminus E) = \mu(F \setminus E)$  proving that  $A \in (\mathcal{B}[\mathbb{R}^n])_{|\mu}$ . So

$$\mathcal{L}[\mathbb{R}^n] \subseteq (\mathcal{B}[\mathbb{R}^n])_{|\mu} \quad (19.248)$$

For the opposite inclusion let  $A \in (\mathcal{B}[\mathbb{R}^n])_{|\mu}$  then there exists a  $E, F \in \mathcal{B}[\mathbb{R}^n]$  with  $E \subseteq A \subseteq F$  and  $\mu(F \setminus E) = 0$ . Hence  $\lambda^n(A \setminus E) \leq_{A \subseteq F} \lambda^n(F \setminus E) = (\lambda^n)|_{\mathcal{B}[\mathbb{R}^n]}(F \setminus E) = \mu(F \setminus E) = 0$ , as  $\lambda^n$  is complete [see 19.108], it follows that  $A \setminus E \in \mathcal{L}[\mathbb{R}^n]$ . As  $E \subseteq A$  we have  $A = E \bigcup (A \setminus E) \in \mathcal{L}[\mathbb{R}^n]$ . This proves  $(\mathcal{B}[\mathbb{R}^n])_{|\mu} \subseteq \mathcal{L}[\mathbb{R}^n]$  which together with 19.248 proves

$$\mathcal{L}[\mathbb{R}^n] = (\mathcal{B}[\mathbb{R}^n])_{|\mu} \quad (19.249)$$

Next we prove that  $\lambda^n = \bar{\mu}$ . Take  $A \in \mathcal{L}[\mathbb{R}^n]$  then there exists  $E, F \in \mathcal{B}[\mathbb{R}^n]$  with  $E \subseteq A \subseteq F$  and  $\mu(F \setminus E) = 0$ . So  $0 = \mu(F \setminus E) = (\lambda^n)|_{\mathcal{B}[\mathbb{R}^n]}(F \setminus E) = \lambda^n(F \setminus E)$  proving

$$\lambda^n(F \setminus E) = 0 \quad (19.250)$$

Hence

$$\begin{aligned} \lambda^n(A) &\stackrel{E \subseteq A}{=} \lambda^n((A \setminus E) \bigsqcup E) \\ &= \lambda^n(A \setminus E) + \lambda^n(E) \\ &\leq \lambda^n(F \setminus E) + \lambda^n(E) \\ &\stackrel{19.250}{=} \lambda^n(E) \\ &= (\lambda^n)|_{\mathcal{B}[\mathbb{R}^n]}(E) \\ &= \mu(E) \\ &\stackrel{\text{def}}{=} \bar{\mu}(E) \end{aligned}$$

proving that

$$\lambda^n = \bar{\mu} \quad (19.251)$$

Using 19.249, 19.251 and the definition of  $\mu = (\lambda^n)|_{\mathcal{B}[\mathbb{R}^n]}$  proves that  $\langle \mathbb{R}^n, \mathcal{L}[\mathbb{R}^n], \lambda^n \rangle = \langle \mathbb{R}^n, (\mathcal{B}[\mathbb{R}^n])|_{\lambda^n}, (\lambda^n)|_{\mathcal{B}[\mathbb{R}^n]} \rangle$  or in other words  $\langle \mathbb{R}^n, \mathcal{L}[\mathbb{R}^n], \lambda^n \rangle$  is the completion of  $\langle \mathbb{R}^n, \mathcal{B}[\mathbb{R}^n], (\lambda^n)|_{\mathcal{B}[\mathbb{R}^n]} \rangle$ .  $\square$

**Definition 19.114.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space then the induced **outer measure**  $\mu^*$  and induced **inner measure**  $\mu_*$  are defined by

$$\begin{aligned} \mu^* : \mathcal{P}(X) &\rightarrow [0, \infty] \text{ where } \mu^*(A) = \inf(\{\mu(B) | A \subseteq B \text{ and } B \in \mathcal{A}\}) \\ \mu_* : \mathcal{P}(X) &\rightarrow [0, \infty] \text{ where } \mu_*(A) = \sup(\{\mu(B) | B \subseteq A \text{ and } B \in \mathcal{A}\}) \end{aligned}$$

**Note 19.115.** As  $\emptyset, X \in \mathcal{A}$  and  $\forall A \in \mathcal{P}(X)$   $\emptyset \subseteq A \subseteq X$  we have  $\{\mu(B) | A \subseteq B \text{ and } B \in \mathcal{A}\} \neq \emptyset \neq \{\mu(B) | B \subseteq A \text{ and } B \in \mathcal{A}\}$  so that the above infimums and supremums are well defined.

**Theorem 19.116.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space then we have

1.  $\forall A \in \mathcal{P}(X)$  we have  $\mu_*(A) \leq \mu^*(A)$
2.  $\forall A \in \mathcal{A}$  we have  $\mu_*(A) \leq \mu(A) \leq \mu^*(A)$
3.  $\mu^*$  is a outer measure on  $X$  [see 19.69]

**Proof.**

1. Let  $A \in \mathcal{P}(X)$  and take  $B \in \mathcal{A}$  such that  $A \subseteq B$  then for every  $B' \in \mathcal{A}$  with  $B' \subseteq A$  we have  $B' \subseteq B$  so that  $\mu(B') \leq \mu(B)$ , hence  $\mu_*(A) = \sup(\{\mu(B') | B' \subseteq A \wedge B' \in \mathcal{A}\}) \leq \mu(B)$ . As this is true for every  $B \in \mathcal{A}$  with  $A \subseteq B$  it follows that  $\mu_*(A) \leq \inf(\{\mu(B') | A \subseteq B' \wedge B' \in \mathcal{A}\}) = \mu^*(A)$ . Proving that

$$\mu_*(A) \leq \mu^*(A)$$

2. Let  $A \in \mathcal{A}$  then for every  $B \in \mathcal{A}$  with  $A \subseteq B$  we have  $\mu(A) \leq \mu(B)$  so that  $\mu(A) \leq \inf(\{\mu(B) | A \subseteq B \wedge B \in \mathcal{A}\}) = \mu^*(A)$ . Further  $\forall B \in \mathcal{A}$  with  $B \subseteq A$  we have  $\mu(B) \leq \mu(A)$  do that  $\mu_*(A) = \sup(\{\mu(B) | B \subseteq A \wedge B \in \mathcal{A}\}) \leq \mu(A)$ .
3. First as  $\emptyset \in \mathcal{A}$  and  $\emptyset \subseteq \emptyset$  we have that  $\mu^*(\emptyset) = \inf(\{\mu(B) | \emptyset \subseteq B \wedge B \in \mathcal{A}\}) \leq \mu(\emptyset) = 0$  proving that

$$\mu^*(\emptyset) = 0 \quad (19.252)$$

Second let  $A, C \in \mathcal{P}(X)$  with  $A \subseteq C$  then if  $\lambda \in \{\mu(B) | C \subseteq B \wedge B \in \mathcal{A}\}$  there exists a  $B \in \mathcal{A}$  such that  $C \subseteq B$  and  $\lambda = \mu(B)$ . As  $A \subseteq C$  we have  $A \subseteq B$  hence  $\lambda = \mu(B) \in \{\mu(B) | A \subseteq B \wedge B \in \mathcal{A}\}$  proving that  $\{\mu(B) | C \subseteq B \wedge B \in \mathcal{A}\} \subseteq \{\mu(B) | A \subseteq B \wedge B \in \mathcal{A}\}$ . Hence using 2.175 we have  $\mu^*(A) = \{\mu(B) | A \subseteq B \wedge B \in \mathcal{A}\} \leq \{\mu(B) | C \subseteq B \wedge B \in \mathcal{A}\} = \mu^*(C)$ . So

$$\forall A, C \in \mathcal{P}(X) \text{ with } A \subseteq C \text{ we have } \mu^*(A) \leq \mu^*(C) \quad (19.253)$$

Let  $\{A_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{P}(X)$  then we must consider for  $\sum_{i=1}^{\infty} \mu(A_i)$  two cases

$\sum_{i=1}^{\infty} \mu^*(A_i) = \infty$ . Then trivially  $\mu^*(\bigcup_{i \in \mathbb{N}_0} A_i) \leq \sum_{i=1}^{\infty} \mu^*(A_i)$

$\sum_{i=1}^{\infty} \mu^*(A_i) < \infty$ . Then  $\forall i \in \mathbb{N}_0$  we have  $\mu^*(A_i) < \infty$ . Choose  $\varepsilon > 0$  then  $\forall i \in \mathbb{N}_0$  we have for  $\mu^*(A_i) = \inf(\{\mu(B) | A_i \subseteq B \wedge B \in \mathcal{A}\})$  that  $\inf(\{\mu(B) | A_i \subseteq B \wedge B \in \mathcal{A}\}) < \mu^*(A_i) + \frac{\varepsilon}{2^i}$ , so there exists a  $B_i \in \mathcal{A}$  with  $A_i \subseteq B_i$  such that  $\mu(B_i) < \mu^*(A_i) + \frac{\varepsilon}{2^i}$ . Then

$$\bigcup_{i \in \mathbb{N}_0} A_i \subseteq \bigcup_{i \in \mathbb{N}_0} B_i \in \mathcal{A} \quad (19.254)$$

giving

$$\begin{aligned} \mu^*\left(\bigcup_{i \in \mathbb{N}_0} A_i\right) &\leq \inf\left(\left\{\mu(B) \mid \bigcup_{i \in \mathbb{N}_0} A_i \subseteq B \wedge B \in \mathcal{A}\right\}\right) \\ &\leq \mu\left(\bigcup_{i \in \mathbb{N}_0} B_i\right) \\ &\leq \sum_{i=1}^{\infty} \mu(B_i) \\ &= \sum_{i=1}^{\infty} \left(\mu^*(A_i) + \frac{\varepsilon}{2^i}\right) \\ &< \sum_{i=1}^{\infty} \mu^*(A_i) + \varepsilon \cdot \sum_{i=1}^{\infty} \frac{1}{2^i} \\ &\stackrel{11.415}{=} \sum_{i=1}^{\infty} \mu^*(A_i) + \varepsilon \end{aligned}$$

As  $\varepsilon$  was chosen arbitrary we have [see 9.57] that

$$\mu^*\left(\bigcup_{i \in \mathbb{N}_0} A_i\right) \leq \sum_{i=1}^{\infty} \mu^*(A_i) \quad (19.255)$$

From 19.252, 19.253 and 19.255 it follows that

$$\mu^* \text{ is a outer measure on } X$$

□

**Theorem 19.117.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space and let  $A \subseteq X$  with  $\mu^*(A) < \infty$  then  $A \in \mathcal{A}_\mu$  if and only if  $\mu^*(A) = \mu_*(A)$

**Proof.**

⇒. Let  $A \in \mathcal{A}_\mu$  then there exists  $E, F \in \mathcal{A}$  with  $E \subseteq A \subseteq F$  and  $\mu(F \setminus E) = 0$ . Then

$$\begin{aligned} \mu(E) &\leq \sup(\{\mu(B) | B \subseteq A \wedge B \in \mathcal{A}\}) \\ &= \mu_*(A) \\ &\stackrel{19.116}{\leq} \mu^*(A) \\ &= \inf(\{\mu(B) | A \subseteq B \wedge B \in \mathcal{A}\}) \\ &\leq \mu(F) \\ &\stackrel{19.109}{=} \mu(E) \end{aligned}$$

proving that

$$\mu_*(A) = \mu^*(A)$$

$\Leftarrow$ . Let  $A \subseteq X$  such that  $\mu_*(A) = \mu^*(A) < \infty$ . So given  $n \in \mathbb{N}_0$  we have that  $\mu^*(A) < \mu^*(A) + \frac{1}{2 \cdot n}$  and  $\mu^*(A) - \frac{1}{2 \cdot n} < \mu^*(A) = \mu_*(A)$ . As  $\mu^*(A) = \inf(\{\mu(B) | A \subseteq B \wedge B \in \mathcal{A}\})$  and  $\mu_*(A) = \sup(\{\mu(B) | B \subseteq A \wedge B \in \mathcal{A}\})$ . There exists  $E_n, F_n \in \mathcal{A}$  such that

$$A \subseteq F_n \wedge \mu(F_n) < \mu^*(A) + \frac{1}{2 \cdot n} \quad (19.256)$$

$$E_n \subseteq A \wedge \mu^*(A) - \frac{1}{2 \cdot n} < \mu(E_n) \quad (19.257)$$

Take  $E = \bigcup_{n \in \mathbb{N}_0} E_n \in \mathcal{A}$ ,  $F = \bigcap_{n \in \mathbb{N}_0} F_n$  then

$$E \subseteq A \subseteq F \quad (19.258)$$

Given  $n \in \mathbb{N}_0$  then we have  $\mu(E_n) \leq 19.116 \mu^*(E_n) \leq \mu^*(A) < \infty$  proving that

$$\mu(E_n) < \infty \quad (19.259)$$

Further we have

$$\begin{aligned} \mu(F \setminus E) &\leq_{F \subseteq F_n} \mu(F_n \setminus E) \\ &\leq_{E_n \subseteq E} \mu(F_n \setminus E_n) \\ &\stackrel{\mu(E_n) < \infty \wedge 19.64}{=} \mu(F_n) - \mu(E_n) \\ &\stackrel{\mu^*(A) < \infty}{=} \mu(F_n) - \mu^*(A) + \mu^*(A) - \mu(E_n) \\ &\stackrel{19.256, 19.257}{<} \frac{1}{2 \cdot n} + \frac{1}{2 \cdot n} \\ &= \frac{1}{n} \end{aligned}$$

As  $n$  was chosen arbitrary we have [see 9.57] that  $\mu(F \setminus E) \leq 0$  giving

$$\mu(F \setminus E) = 0 \quad (19.260)$$

Using 19.258, 19.260 and the fact that  $E, F \in \mathcal{A}$  we have that

$$A \in \mathcal{A}_\mu \quad \square$$

We introduce now the concept of regularity.

**Definition 19.118.** Let  $\langle X, \mathcal{T} \rangle$  be a topological space and  $\langle X, \mathcal{A}, \mu \rangle$  a measure space then

1.  $A \in \mathcal{A}$  is **inner regular** if  $\mu(A) = \sup(\{\mu(F) | F \in \mathcal{A} \wedge F \text{ is compact} \wedge F \subseteq A\})$
2.  $A \in \mathcal{A}$  is **outer regular** if  $\mu(A) = \inf(\{\mu(U) | U \in \mathcal{A} \wedge U \in \mathcal{T} \wedge A \subseteq U\})$
3.  $A \in \mathcal{A}$  is **regular** if it is **inner** and **outer** regular
4.  $\mu$  is **inner regular** if  $\forall A \in \mathcal{A}$  we have that  $A$  is **inner regular**
5.  $\mu$  is **outer regular** if  $\forall A \in \mathcal{A}$  we have that  $A$  is **outer regular**
6.  $\mu$  is **regular** if  $\mu$  is **inner** and **outer** regular

**Note 19.119.** The condition  $F \in \mathcal{A}$  in (1) is needed to ensure that  $\mu(F)$  is defined. Likely the condition  $U \in \mathcal{A}$  in (2) is needed to ensure that  $\mu(U)$  exists.

**Note 19.120.** Some books say that  $\mu$  is inner regular if  $\forall A \in \mathcal{T}$  we have that  $A$  is inner regular, which is a weaker condition than (5). Also some authors assume that  $\mu$  must be finite to be regular.

**Proposition 19.121.** Let  $n \in \mathbb{N}_0$  then the Lebesgue measure in  $\langle \mathbb{R}^n, \mathcal{L}[\mathbb{R}^n], (\lambda^n)_{|\mathcal{B}[\mathbb{R}^n]} \rangle$  is regular

**Proof.** This follows from 19.92 and 19.91.  $\square$

We show now that finite measures on  $\mathcal{B}[\mathbb{R}^n]$  are regular. First we prove a preliminary lemma.

**Lemma 19.122.** Let  $n \in \mathbb{N}_0$ ,  $\langle \mathbb{R}^n, \mathcal{B}[\mathbb{R}^n] \rangle$  the measurable space defined in 19.20] and  $\mu$  a measure on  $\mathcal{B}[\mathbb{R}^n]$  such that  $\forall A \in \mathcal{B}[\mathbb{R}^n]$  we have  $\mu(A) < \infty$  [ $\mu$  is a finite measure] then  $\forall A \in \mathcal{B}[\mathbb{R}^n]$  we have that

1.  $\mu(A) = \inf(\{\mu(U) | U \text{ open and } A \subseteq U\})$
2.  $\mu(A) = \sup(\{\mu(C) | C \text{ closed and } C \subseteq A\})$

**Note 19.123.** As open and closed sets are in  $\mathcal{B}[\mathbb{R}^n]$  [see 19.58],  $\emptyset \subseteq A$  and  $A \subseteq X$  the above definitions makes sense.

**Proof.** Let  $\mathcal{R}$  be the set of Borel sets that satisfies (1) and (2), more specific

$$\mathcal{R} = \{A \in \mathcal{B}[\mathbb{R}^n] | \inf(\{\mu(U) | U \text{ open and } A \subseteq U\}) = \sup(\{\mu(F) | F \text{ closed and } F \subseteq A\}) = \mu(A)\}$$

First we prove that  $\mathcal{R}$  contains all the open sets. Let  $V$  be a open set then using 19.59 there exists a sequence  $\{G_i\}_{i \in \mathbb{N}_0}$  of closed sets such that  $V = \bigcup_{i \in \mathbb{N}_0} G_i$ . Define  $\{F_i\}_{i \in \mathbb{N}_0}$  by  $F_i = \bigcup_{j \in \{1, \dots, i\}} G_j$  then as  $F_i$  is a finite union of closed sets,  $F_i$  is closed, also trivially  $F_i \subseteq F_{i+1}$ . Further as  $G_i \subseteq F_i$  we have  $\bigcup_{i \in \mathbb{N}_0} G_i \subseteq \bigcup_{i \in \mathbb{N}_0} F_i$  and if  $x \in \bigcup_{i \in \mathbb{N}_0} F_i$  there exists a  $i \in \mathbb{N}_0$  such that  $x \in F_i = \bigcup_{j \in \{1, \dots, i\}} G_j$ , hence  $\exists j \in \{1, \dots, i\}$  such that  $x \in G_j$  proving that  $\bigcup_{i \in \mathbb{N}_0} F_i \subseteq \bigcup_{i \in \mathbb{N}_0} G_i$ . So  $\bigcup_{i \in \mathbb{N}_0} G_i = \bigcup_{i \in \mathbb{N}_0} F_i$ . To summarize we have

$$V = \bigcup_{i \in \mathbb{N}_0} F_i \text{ with } \forall i \in \mathbb{N}_0 \text{ we have } F_i \subseteq F_{i+1} \text{ and } F_i \text{ is closed hence } F_i \in \mathcal{B}[\mathbb{R}^n] \quad (19.261)$$

So

$$\begin{aligned} \mu(V) &\stackrel{19.261}{=} \mu\left(\bigcup_{i \in \mathbb{N}_0} F_i\right) \\ &\stackrel{19.67}{=} \sup(\{\mu(F_i) | i \in \mathbb{N}_0\}) \end{aligned} \quad (19.262)$$

If  $F$  is closed and  $F \subseteq V$  then  $\mu(F) \leq \mu(V)$  so that

$$\sup(\{\mu(F) | F \text{ closed and } F \subseteq V\}) \leq \mu(V) \quad (19.263)$$

As  $F \subseteq \bigcup_{i \in \mathbb{N}_0} F_i = V$  and  $F_i$  is closed we have that  $\mu(F_i) \leq \sup(\{\mu(F) | F \text{ closed and } F \subseteq V\})$  so that

$$\sup(\{\mu(F_i) | i \in \mathbb{N}_0\}) \leq \sup(\{\mu(F) | F \text{ closed and } F \subseteq V\}) \quad (19.264)$$

So we have that  $\mu(V) \stackrel{19.262}{=} \sup(\{\mu(F_i) | i \in \mathbb{N}_0\}) \leq \sup(\{\mu(F) | F \text{ closed and } F \subseteq V\}) \leq \mu(V)$  which proves that

$$\mu(V) = \sup(\{\mu(F) | F \text{ closed and } F \subseteq V\}) \quad (19.265)$$

Further as  $V$  is open and  $V \subseteq V$  we have that  $\inf(\{\mu(U) | U \text{ open and } V \subseteq U\}) \leq \mu(V)$ , Further if  $U$  is open such that  $V \subseteq U$  then  $\mu(V) \leq \mu(U)$  hence  $\mu(V) \leq \inf(\{\mu(U) | U \text{ open and } V \subseteq U\})$  proving that

$$\mu(V) = \inf(\{\mu(U) | U \text{ open and } V \subseteq U\}) \quad (19.266)$$

From 19.264 and 19.266 and the definition of  $\mathcal{R}$  it follows that  $V \in \mathcal{R}$  proving as  $V$  was chosen arbitrary

$$\mathcal{T}_{\mathbb{R}^n} \subseteq \mathcal{R} \quad (19.267)$$

Define now

$$\mathcal{S} = \{A \in \mathcal{B}[\mathbb{R}^n] | \forall \varepsilon > 0 \quad \exists U, F \text{ where } U \text{ is open, } F \text{ closed, } F \subseteq A \subseteq U \text{ and } \mu(U \setminus F) < \varepsilon\}$$

and prove that  $\mathcal{R} = \mathcal{S}$ .

Let  $A \in \mathcal{R}$  and  $\varepsilon > 0$ . As  $\mu(A) = \sup(\{\mu(F) | F \text{ closed and } F \subseteq A\})$  there exists a  $F$  closed with  $F \subseteq A$  such that

$$\mu(A) - \frac{\varepsilon}{2} < \mu(F) \quad (19.268)$$

Further as  $\mu(A) = \inf(\{\mu(U) | U \text{ open and } A \subseteq U\})$  there exists a  $U$  open with  $A \subseteq U$  such that

$$\mu(U) < \mu(A) + \frac{\varepsilon}{2} \quad (19.269)$$

Now as  $\mu$  is finite we have by 19.64 that

$$\begin{aligned}\mu(U \setminus F) &= \mu(U) - \mu(F) \\ &\stackrel{19.269}{<} \mu(A) + \frac{\varepsilon}{2} - \mu(F) \\ &\stackrel{19.268}{<} \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\ &= \varepsilon\end{aligned}$$

hence we have

$$\mathcal{R} \subseteq \mathcal{S} \quad (19.270)$$

For the opposite inclusion. Let  $A \in \mathcal{S} \subseteq \mathcal{B}[\mathbb{R}^n]$  then if  $C$  is closed with  $C \subseteq A$  we have  $\mu(C) \leq \mu(A)$  so that  $\sup(\{\mu(C) | C \subseteq A \wedge C \text{ closed}\}) \leq \mu(A)$ . Further if  $U$  is open with  $A \subseteq U$  then  $\mu(A) \leq \mu(U)$  so that  $\mu(A) \leq \inf(\{\mu(U) | A \subseteq U \wedge U \text{ closed}\})$ . To summarize we have

$$\sup(\{\mu(C) | C \subseteq A \wedge C \text{ closed}\}) \leq \mu(A) \leq \inf(\{\mu(U) | A \subseteq U \wedge U \text{ open}\}) \quad (19.271)$$

Assume now that  $\inf(\{\mu(U) | A \subseteq U \wedge U \text{ open}\}) > \sup(\{\mu(C) | C \subseteq A \wedge C \text{ closed}\})$  then  $0 < \inf(\{\mu(U) | A \subseteq U \wedge U \text{ open}\}) - \sup(\{\mu(C) | C \subseteq A \wedge C \text{ closed}\}) = \varepsilon$ . As  $A \in \mathcal{S}$  there exists a open  $U$ , closed  $C$  so that

$$C \subseteq A \subseteq U \text{ and } \mu(U \setminus C) < \varepsilon \text{ or as } \mu \text{ is finite } \mu(U) < \mu(C) + \varepsilon \quad (19.272)$$

So

$$\begin{aligned}\inf(\{\mu(U) | A \subseteq U \wedge U \text{ open}\}) &\leq \mu(U) \\ &\stackrel{19.272}{<} \mu(C) + \varepsilon \\ &\leq \sup(\{\mu(C) | C \subseteq A \wedge C \text{ closed}\}) + \varepsilon \\ &\leq \sup(\{\mu(C) | C \subseteq A \wedge C \text{ closed}\}) + \inf(\{\mu(U) | A \subseteq U \wedge U \text{ open}\}) - \sup(\{\mu(C) | C \subseteq A \wedge C \text{ closed}\}) \\ &= \inf(\{\mu(U) | A \subseteq U \wedge U \text{ open}\})\end{aligned}$$

giving the contradiction  $\inf(\{\mu(U) | A \subseteq U \wedge U \text{ open}\}) < \inf(\{\mu(U) | A \subseteq U \wedge U \text{ open}\})$ . Hence  $\inf(\{\mu(U) | A \subseteq U \wedge U \text{ open}\}) \leq \sup(\{\mu(C) | C \subseteq A \wedge C \text{ closed}\})$  which together with 19.271 proves that  $\sup(\{\mu(C) | C \subseteq A \wedge C \text{ closed}\}) = \mu(A) = \inf(\{\mu(U) | A \subseteq U \wedge U \text{ open}\})$ . So  $A \in \mathcal{R}$  proving  $\mathcal{S} \subseteq \mathcal{R}$  which together with 19.270 gives  $\mathcal{S} = \mathcal{R}$  or using the definition of  $\mathcal{S}$ .

$$\mathcal{R} = \{A \in \mathcal{B}[\mathbb{R}^n] \forall \varepsilon > 0 \quad \exists U, F \text{ where } U \text{ open, } F \text{ closed, } F \subseteq A \subseteq U \wedge \mu(U \setminus F) < \varepsilon\} \quad (19.273)$$

We are now ready to prove that  $\mathcal{R}$  is a  $\sigma$ -algebra:

1. As  $\emptyset$  is open we have that  $\emptyset \in \mathcal{T}_{\mathbb{R}^n}$  and as  $\mathcal{T}_{\mathbb{R}^n} \subseteq \mathcal{R}$  [see 19.267] it follows that

$$\emptyset \in \mathcal{R} \quad (19.274)$$

2. Let  $A \in \mathcal{R}$  and take  $\varepsilon > 0$ . Using 19.273 there exists a open  $U$ , a closed  $F$  such that  $F \subseteq A \subseteq U$  and  $\mu(U \setminus F) < \varepsilon$ . Taking the complement we have

$$X \setminus U \text{ is closed, } X \setminus F \text{ is open and } X \setminus U \subseteq X \setminus A \subseteq X \setminus F \quad (19.275)$$

Further

$$\begin{aligned}x \in (X \setminus F) \setminus (X \setminus U) &\Leftrightarrow x \in X \setminus F \wedge \neg(x \in X \setminus U) \\ &\Leftrightarrow (x \in X \wedge x \notin F) \wedge \neg(x \in X \wedge x \notin U) \\ &\Leftrightarrow (x \in X \wedge x \notin F) \wedge (x \notin X \vee x \in U) \\ &\Leftrightarrow (x \in X \wedge x \notin F \wedge x \notin X) \vee (x \in X \wedge x \notin F \wedge x \in U) \\ &\Leftrightarrow x \in X \wedge x \notin F \wedge x \in U \\ &\stackrel{U \subseteq X}{\Leftrightarrow} x \in U \wedge x \notin F \\ &\Leftrightarrow x \in U \setminus F\end{aligned}$$

proving that  $(X \setminus F) \setminus (X \setminus U) = U \setminus F$ . Hence

$$\mu((X \setminus F) \setminus (X \setminus U)) = \mu(U \setminus F) < \varepsilon \quad (19.276)$$

Using 19.276, 19.275 in 19.273 proves that  $X \setminus A \in \mathcal{R}$  so

$$\forall A \in \mathcal{R} \text{ we have } X \setminus A \in \mathcal{R} \quad (19.277)$$

3. Let  $\{A_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{R}$  and take  $\varepsilon > 0$  then using 19.273 we have that  $\forall i \in \mathbb{N}_0$  there exists a open  $U_i$  and a closed  $F_i$  with  $F_i \subseteq A_i \subseteq U_i$  such that  $\mu(U_i \setminus F_i) < \frac{\varepsilon}{2^i}$ . Define  $U = \bigcup_{i \in \mathbb{N}_0} U_i$  and  $F = \bigcup_{i \in \mathbb{N}_0} F_i$  then  $U$  is open [note that  $F$  does not have to be closed]. Further  $F = \bigcup_{i \in \mathbb{N}_0} F_i \subseteq \bigcup_{i \in \mathbb{N}_0} A_i \subseteq \bigcup_{i \in \mathbb{N}_0} U_i = U$  hence

$$F \subseteq \bigcup_{i \in \mathbb{N}_0} A_i \subseteq U \text{ where } U \text{ is open} \quad (19.278)$$

Next

$$\begin{aligned} \mu(U \setminus F) &= \mu\left(\left(\bigcup_{i \in \mathbb{N}_0} U_i\right) \setminus \left(\bigcup_{i \in \mathbb{N}_0} F_i\right)\right) \\ &\stackrel{1.102}{=} \mu\left(\bigcup_{i \in \mathbb{N}_0} \left(U_i \setminus \left(\bigcup_{i \in \mathbb{N}_0} F_i\right)\right)\right) \\ &\leq \mu\left(\bigcup_{i \in \mathbb{N}_0} (U_i \setminus F_i)\right) \\ &\leq \sum_{i=1}^{\infty} \mu(U_i \setminus F_i) \\ &< \sum_{i=1}^{\infty} \frac{\varepsilon}{2^i} \\ &\stackrel{11.415}{=} \varepsilon \end{aligned}$$

giving

$$\mu(U \setminus F) < \varepsilon \quad (19.279)$$

Unfortunately we can not yet apply 19.273 as  $F = \bigcup_{i \in \mathbb{N}_0} F_i$  does not have to be closed. However if  $i \in \mathbb{N}_0$  then  $\bigcup_{j \in \{1, \dots, n\}} F_j$  is closed [as it is a finite union and using 11.13]. Define the family  $\{G_i\}_{i \in \mathbb{N}_0}$  by  $G_i = U \setminus (\bigcup_{j \in \{1, \dots, i\}} F_j) \in \mathcal{B}[\mathbb{R}^n]$  then  $G_i = U \setminus (\bigcup_{j \in \{1, \dots, i\}} F_j) \supseteq U \setminus (\bigcup_{j \in \{1, \dots, i+1\}} F_j) = G_{i+1}$ . Also  $G_i = U \setminus (\bigcup_{j \in \{1, \dots, i\}} F_j) \supseteq U \setminus (\bigcup_{j \in \mathbb{N}_0} F_j)$  so that

$$\bigcap_{i \in \mathbb{N}_0} G_i \supseteq U \setminus \left(\bigcup_{i \in \mathbb{N}_0} F_i\right)$$

Further if  $x \in \bigcap_{i \in \mathbb{N}_0} G_i$  then  $\forall i \in \mathbb{N}_0$  we have  $x \in G_i = U \setminus (\bigcup_{j \in \{1, \dots, i\}} F_j)$  so that  $x \in U \wedge x \notin F_i \Rightarrow x \notin \bigcup_{i \in \mathbb{N}_0} F_i$  giving  $x \in U \setminus (\bigcup_{i \in \mathbb{N}_0} F_i)$  proving that

$$\bigcap_{i \in \mathbb{N}_0} G_i \subseteq U \setminus \left(\bigcup_{i \in \mathbb{N}_0} F_i\right)$$

To summarize we have

$$\{G_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{B}[\mathbb{R}^n], \forall i \in \mathbb{N}_0 G_{i+1} \subseteq G_i \text{ and } \bigcap_{i \in \mathbb{N}_0} G_i = U \setminus \left(\bigcup_{i \in \mathbb{N}_0} F_i\right) = U \setminus F$$

which together with the fact that  $\mu$  is finite allows us to apply 19.67 so that

$$\mu(U \setminus F) = \inf(\{\mu(G_i) | i \in \mathbb{N}_0\})$$

As  $\mu(U \setminus F) < \varepsilon$  [see 19.279] there exists a  $m \in \mathbb{N}_0$  such that  $\mu(G_m) < \varepsilon$  or  $\mu(U \setminus (\bigcup_{i \in \{1, \dots, m\}} F_i)) < \varepsilon$ . Hence if we take  $C = \bigcup_{i \in \{1, \dots, m\}} F_i$  [a closed set by 11.13] then  $C \subseteq \bigcup_{i \in \mathbb{N}_0} F_i \subseteq_{19.278} \bigcup_{i \in \mathbb{N}_0} A_i \subseteq U$  and  $\mu(U \setminus C) < \varepsilon$ , so by 19.273 it follows that

$$\bigcup_{i \in \mathbb{N}_0} A_i \in \mathcal{R} \quad (19.280)$$

From 19.274, 19.277 and the above [19.280] it follows that  $\mathcal{R}$  is a  $\sigma$ -algebra which by 19.267 contains  $\mathcal{T}_{\mathbb{R}^n}$  hence  $\mathcal{B}[\mathbb{R}^n] = \sigma[\mathcal{T}_{\mathbb{R}^n}] \subseteq \mathcal{R}$ . Or using the definition of  $\mathcal{R}$  it follows that

$$\forall A \in \mathcal{B}[\mathbb{R}^n] \mu(A) = \sup(\{\mu(F) | F \text{ closed and } F \subseteq A\}) = \inf(\{\mu(U) | U \text{ open } A \subseteq U\})$$

proving the lemma.  $\square$

We use the above lemma to prove the following theorem

**Theorem 19.124.** *Let  $n \in \mathbb{N}_0$  and  $\langle \mathbb{R}^n, \mathcal{B}[\mathbb{R}^n], \mu \rangle$  the measurable space defined in 19.20] and  $\mu$  a measure on  $\mathcal{B}[\mathbb{R}^n]$  such that  $\forall A \in \mathcal{B}[\mathbb{R}^n]$  we have  $\mu(A) < \infty$  then  $\mu$  is regular [see 19.118] or in other words*

$$\begin{aligned} \mu(A) &= \inf(\{\mu(U) | U \in \mathcal{B}[\mathbb{R}^n] \wedge U \in \mathcal{T} \wedge A \subseteq U\}) \\ &= \sup(\{\mu(F) | F \in \mathcal{B}[\mathbb{R}^n] \wedge F \text{ is compact} \wedge F \subseteq A\}) \end{aligned}$$

**Note 19.125.** As  $\mathcal{T}_{\mathbb{R}^n} \subseteq \mathcal{B}[\mathbb{R}^n]$  and compact sets are closed [see 11.254] and thus elements of  $\mathcal{B}[\mathbb{R}^n]$  [see 19.58] we must only prove that

$$\mu(A) = \inf(\{\mu(U) | U \text{ open} \wedge A \subseteq U\}) = \sup(\{\mu(C) | C \text{ compact} \wedge C \subseteq A\})$$

**Proof.** Using 19.122 we have that

$$\mu(A) = \inf(\{\mu(U) | U \text{ open} \wedge A \subseteq U\}) \quad (19.281)$$

proving the first part. For the second part, note that if  $C$  is compact with  $C \subseteq A$  then  $\mu(C) \leq \mu(A)$  proving that

$$\sup(\{\mu(C) | C \text{ compact} \wedge C \subseteq A\}) \leq \mu(A) \quad (19.282)$$

Now for the opposite inclusion take  $\varepsilon > 0$  then, as  $\mu(A) \stackrel{19.122}{=} \sup(\{\mu(F) | F \text{ closed} \wedge F \subseteq A\})$ , there exists a closed set  $F$  with  $F \subseteq A$  and

$$\mu(A) - \varepsilon < \mu(F) \quad (19.283)$$

Now as  $F = \mathbb{R}^n \cap F \stackrel{19.82}{=} (\bigcup_{i \in \mathbb{N}_0} [-i, i]^n) \cap A = \bigcup_{i \in \mathbb{N}_0} ([-i, i]^n \cap F) = \bigcup_{i \in \mathbb{N}} F_i$  where  $F_i = [-i, i]^n \cap F$ . If  $x \in F_i$  then  $\forall j \in \{1, \dots, n\} |x_j| < i$  giving  $\|x\|_{\max} \leq i$  proving that  $F_i$  is bounded. As  $F_i$  is also closed [union of two closed sets] we have by the Heine-Borel theorem [see 11.300].that  $F_i$  is compact. Further  $F_i = [-i, i]^n \cap F \subseteq [-(i+1), i+1]^n \cap F = F_{i+1}$ . To summarize

$$\exists \{F_i\}_{i \in \mathbb{N}_0} \text{ with } \forall i \in \mathbb{N}_0 F_i \text{ is compact and } F_i \subseteq F_{i+1} \text{ such that } F = \bigcup_{i \in \mathbb{N}_0} F_i$$

Using the above on the properties of a measure [see 19.67] we have that

$$\mu(F) = \sup(\{\mu(F_i) | i \in \mathbb{N}_0\}) \quad (19.284)$$

As by 19.283 and 19.284 we have that  $\mu(A) - \varepsilon < \sup(\{\mu(F_i) | i \in \mathbb{N}_0\})$ , it follows that  $\exists m \in \mathbb{N}_0$  such that  $\mu(A) - \varepsilon < \mu(F_m)$ . As  $F_m \subseteq F$  and  $F$  is closed we have that  $\mu(F_m) \leq \sup(\{\mu(C) | C \text{ is compact} \wedge C \subseteq A\})$ . Hence  $\mu(A) - \varepsilon < \sup(\{\mu(C) | C \text{ is compact} \wedge C \subseteq A\})$ , as  $\varepsilon$  was chosen arbitrary, we have by 9.57 that  $\mu(A) \leq \sup(\{\mu(C) | C \text{ is compact} \wedge C \subseteq A\})$ . Combining this with 19.282 proves

$$\mu(A) = \sup(\{\mu(C) | C \text{ is compact} \wedge C \subseteq A\}) \quad (19.285)$$

The theorem is then proved by 19.281 and 19.285.  $\square$

## 19.4 Measurable functions

### 19.4.1 Dynkin Classes

This section describes a technique that is often useful for verifying the equality of measures and measurable functions.

**Definition 19.126.** Let  $X$  be a set then  $\mathcal{D} \subseteq \mathcal{P}(X)$  is a **Dynkin class** on  $X$  if

1.  $X \in \mathcal{D}$
2.  $\forall A, B \in \mathcal{D}$  with  $B \subseteq A$  we have  $A \setminus B \subseteq \mathcal{D}$
3. If  $\{A_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{D}$  is such that  $\forall i \in \mathbb{N}_0 A_i \subseteq A_{i+1}$  then  $\bigcup_{i \in \mathbb{N}_0} A_i \in \mathcal{D}$

**Example 19.127.** Let  $\langle X, \mathcal{A} \rangle$  be a measurable space [meaning that  $\mathcal{A}$  is a  $\sigma$ -algebra on  $X$ ] then  $\mathcal{A}$  is a Dynkin class on  $X$

**Proof.** This follows from the definition of a  $\sigma$ -algebra.  $\square$

**Example 19.128.** Let  $X$  be a set then  $\mathcal{P}(X)$  is a Dynkin class on  $X$

**Proof.**

1.  $X \in \mathcal{P}(X)$  as  $X \subseteq X$
2. If  $A, B \in \mathcal{P}(X)$  then  $A \setminus B \subseteq A \subseteq X$  so that  $A \setminus B \in \mathcal{P}(X)$
3. If  $\{A_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{P}(X)$  with  $\forall i \in \mathbb{N}_0 A_i \subseteq A_{i+1}$  then  $\forall i \in \mathbb{N}_0$  we have  $A_i \subseteq X$  hence  $\bigcup_{i \in \mathbb{N}_0} A_i \subseteq X$  proving that  $\bigcup_{i \in \mathbb{N}_0} A_i \in \mathcal{P}(X)$

$\square$

**Lemma 19.129.** Let  $X$  be a set and  $\mathcal{X}$  a non empty set of Dynkin classes on  $X$  then  $\bigcap_{\mathcal{D} \in \mathcal{X}} \mathcal{D}$  is a Dynkin class on  $X$ .

**Proof.**

1. As  $\forall \mathcal{D} \in \mathcal{X}$  we have that  $X \in \mathcal{D}$  it follows that  $X \in \bigcap_{\mathcal{D} \in \mathcal{X}} \mathcal{D}$
2. Let  $A, B \in \bigcap_{\mathcal{D} \in \mathcal{X}} \mathcal{D}$  then  $\forall \mathcal{D} \in \mathcal{X}$  we have that  $A, B \in \mathcal{D} \Rightarrow A \setminus B \in \mathcal{D}$  hence  $A \setminus B \in \mathcal{X}$
3. Let  $\{A_i\}_{i \in \mathbb{N}_0} \subseteq \bigcap_{\mathcal{D} \in \mathcal{X}} \mathcal{D}$  with  $\forall i \in \mathbb{N}_0 A_i \subseteq A_{i+1}$  then given  $i \in \mathbb{N}_0$  we have  $A_i \in \bigcap_{\mathcal{D} \in \mathcal{X}} \mathcal{D}$  so that given a  $\mathcal{D} \in \mathcal{X}$  we have  $A_i \in \mathcal{D}$  proving that  $\forall \mathcal{D} \in \mathcal{X}$  we have  $\{A_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{D} \Rightarrow \bigcup_{i \in \mathbb{N}_0} A_i \in \mathcal{D}$  hence  $\bigcup_{i \in \mathbb{N}_0} A_i \in \bigcap_{\mathcal{D} \in \mathcal{X}} \mathcal{D}$

$\square$

Just as in the case of  $\sigma$ -algebras [see 19.17] we can find the smallest Dynkin class covering a collection of subsets of a given set.

**Theorem 19.130.** Let  $X$  be a set and  $\mathcal{A} \subseteq \mathcal{P}(X)$  then there exists a **unique** Dynkin class  $\mathcal{D}$  on  $X$  such that  $\mathcal{A} \subseteq \mathcal{D}$  and for every Dynkin class  $\mathcal{D}'$  on  $X$  with  $\mathcal{A} \subseteq \mathcal{D}'$  we have that  $\mathcal{D} \subseteq \mathcal{D}'$ . this unique Dynkin class is noted as  $\mathcal{d}[\mathcal{A}]$  and is called the **Dynkin class generated by  $\mathcal{A}$** .

**Proof.** Define  $\mathcal{X} = \{\mathcal{C} \subseteq \mathcal{P}(X) | \mathcal{A} \subseteq \mathcal{C} \text{ and } \mathcal{C} \text{ is a Dynkin class}\}$  then as  $\mathcal{P}(X) \in \mathcal{X}$  [see 19.128]  $\mathcal{X} \neq \emptyset$ , hence by 19.129 we have that  $\mathcal{D} = \bigcap_{\mathcal{C} \in \mathcal{X}} \mathcal{C}$  is a Dynkin class. As  $\forall \mathcal{C} \in \mathcal{X}$  we have that  $\mathcal{A} \subseteq \mathcal{C}$  it follows that  $\mathcal{A} \subseteq \bigcap_{\mathcal{C} \in \mathcal{X}} \mathcal{C} = \mathcal{D}$ . If  $\mathcal{D}'$  is another Dynkin class with  $\mathcal{A} \subseteq \mathcal{D}'$  then  $\mathcal{D}' \in \mathcal{X}$  hence  $\mathcal{D} = \bigcap_{\mathcal{C} \in \mathcal{X}} \mathcal{C} \subseteq \mathcal{D}'$ . Finally let  $\mathcal{D}''$  be another Dynkin class with  $\mathcal{A} \subseteq \mathcal{D}''$  such that it contains every Dynkin class that contains  $\mathcal{A}$  we have that  $\mathcal{D} \subseteq \mathcal{D}''$  [ $\mathcal{D}$  is a Dynkin class containing  $\mathcal{A}$ ] and  $\mathcal{D}'' \subseteq \mathcal{D}$  [ $\mathcal{D}''$  is a Dynkin class containing  $\mathcal{A}$ ] so that  $\mathcal{D}'' = \mathcal{D}$  proving uniqueness.  $\square$

**Definition 19.131.** Let  $X$  be a set then  $\mathcal{A} \subseteq \mathcal{P}(X)$  is a  $\pi$ -system on  $X$  if  $\forall A, B \in \mathcal{A}$  we have  $A \cap B \in \mathcal{A}$ .

**Note 19.132.** Let  $X$  be a set and  $\mathcal{A}$  a  $\pi$ -system on  $X$  then for  $\{A_i\}_{i \in I}$  where  $I$  is finite and non empty we have that  $\bigcap_{i \in I} A_i \in \mathcal{A}$

**Proof.** This follows from 19.7.  $\square$

We prove now that every Dynkin class that is also a  $\pi$ -system is a  $\sigma$ -algebra

**Lemma 19.133.** *Let  $X$  be a set and  $\mathcal{D}$  a Dynkin class on  $X$  that is a  $\pi$ -system on  $X$  then  $\mathcal{D}$  is a  $\sigma$ -algebra*

**Proof.** Let's check the conditions for a  $\sigma$ -algebra [see  $\sigma$ -algebra]

1.  $X \in \mathcal{D}$  by definition of a Dynkin class on  $X$
2. Let  $A \in \mathcal{D}$  then as  $X \in \mathcal{D}$  and  $A \subseteq X$  we have by the definition of a Dynkin class that  $X \setminus A \in \mathcal{D}$
3. First if  $A, B \in \mathcal{D}$  then by (2)  $(X \setminus A), (X \setminus B) \in \mathcal{D}$  so, as  $\mathcal{D}$  is a  $\pi$ -system, we have that  $X \setminus (A \cup B) =_{1.28} (X \setminus A) \cap (X \setminus B) \in \mathcal{D}$ . Further as  $A \cup B =_{1.19} X \setminus (X \setminus (A \cup B)) \in \mathcal{D}$  we have proved that  $\forall A, B \in \mathcal{D} A \cup B \in \mathcal{D}$ . Applying then 19.7 gives

$$\forall \{A_i\}_{i \in I} \subseteq \mathcal{D} \text{ where } I \text{ is finite non empty we have that } \bigcup_{i \in I} A_i \in \mathcal{D}$$

Next let  $\{A_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{D}$  define  $\{B_i\}_{i \in \mathbb{N}_0}$  by  $B_i = \bigcup_{j \in \{1, \dots, i\}} A_j$ . Then using 19.5 we have that  $\forall i \in \mathbb{N}_0 B_i \subseteq B_{i+1}$  and  $\bigcup_{i \in \mathbb{N}_0} A_i = \bigcup_{i \in \mathbb{N}_0} B_i \in \mathcal{D}$  [as  $\mathcal{D}$  is a Dynkin system on  $X$ ]. So we have that

$$\forall \{A_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{D} \text{ we have } \bigcup_{i \in \mathbb{N}_0} A_i \in \mathcal{D} \quad \square$$

**Theorem 19.134.** *Let  $X$  a set and  $\mathcal{A} \subseteq \mathcal{P}(X)$  a  $\pi$ -system then  $\mathcal{d}[\mathcal{A}] = \sigma[\mathcal{A}]$  or in other words the generated sigma algebra is equal to the generated Dynkin class.*

**Proof.** First we prove that  $\mathcal{d}[\mathcal{A}]$  is itself a  $\pi$ -system because we can then apply the previous lemma 19.133. For this define

$$\mathcal{D}_1 = \{A \in \mathcal{d}[\mathcal{A}] \mid \forall C \in \mathcal{A} \text{ we have } A \cap C \in \mathcal{d}[\mathcal{A}]\} \quad (19.286)$$

If  $A \in \mathcal{A} \subseteq \mathcal{d}[\mathcal{A}]$  then, as  $\mathcal{A}$  is a  $\pi$ -system, we have that  $\forall C \in \mathcal{A}$  that  $A \cap C \in \mathcal{A} \subseteq \mathcal{d}[\mathcal{A}]$  which proves that  $A \in \mathcal{D}_1$ . Hence

$$\mathcal{A} \subseteq \mathcal{D}_1 \quad (19.287)$$

As  $X \in \mathcal{d}[\mathcal{A}]$  and  $\forall C \in \mathcal{A}$  we have  $X \cap C = C \in \mathcal{A} \subseteq \mathcal{d}[\mathcal{A}]$  which proves that

$$X \in \mathcal{D}_1 \quad (19.288)$$

If  $A, B \in \mathcal{D}_1$  with  $B \subseteq A$  then  $\forall C \in \mathcal{A} (A \cap C), (B \cap C) \in \mathcal{D}_1$ , further using 19.2, we have  $(A \setminus B) \cap C = (A \cap C) \setminus (B \cap C)$  which, as  $\mathcal{d}[\mathcal{A}]$  is a Dynkin system and  $B \cap C \subseteq A \cap C$  so that  $(A \cap C) \setminus (B \cap C) \in \mathcal{d}[\mathcal{A}]$ , proves that  $A \setminus B \in \mathcal{D}_1$ . Hence we have

$$\forall A, B \in \mathcal{D}_1 \text{ then } A \setminus B \in \mathcal{D}_1 \quad (19.289)$$

Further if  $\{A_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{D}_1$  is a increasing sequence of sets in  $\mathcal{D}_1$  and  $C \in \mathcal{A}$  then  $\forall i \in \mathbb{N}_0$  we have  $A_i \cap C \in \mathcal{d}[\mathcal{A}]$  so that  $\{A_i \cap C\}_{i \in \mathbb{N}_0}$  is a increasing sequence of sets in  $\mathcal{d}[\mathcal{A}]$ . So  $\bigcup_{i \in \mathbb{N}_0} (A_i \cap C) \in \mathcal{d}[\mathcal{A}]$ . As  $(\bigcup_{i \in \mathbb{N}_0} A_i) \cap C =_{1.101} \bigcup_{i \in \mathbb{N}_0} (A_i \cap C) \in \mathcal{d}[\mathcal{A}]$ , it follows that  $(\bigcup_{i \in \mathbb{N}_0} A_i) \cap C \in \mathcal{d}[\mathcal{A}]$ . So we have

$$\forall \{A_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{D}_1 \text{ with } \forall i \in \mathbb{N}_0 A_i \subseteq A_{i+1} \text{ we have } \bigcup_{i \in \mathbb{N}_0} A_i \in \mathcal{D}_1 \quad (19.290)$$

From 19.287, 19.288, 19.289 and 19.290 it follows that  $\mathcal{D}_1$  is a Dynkin class containing  $\mathcal{A}$  hence we have that

$$\mathcal{d}[\mathcal{A}] \subseteq \mathcal{D}_1 \quad (19.291)$$

Next define

$$\mathcal{D}_2 = \{A \in \mathcal{d}[\mathcal{A}] \mid \forall C \in \mathcal{d}[\mathcal{A}] \text{ we have } A \cap C \in \mathcal{d}[\mathcal{A}]\} \quad (19.292)$$

If  $A \in \mathcal{A} \subseteq \mathcal{d}[\mathcal{A}]$  and  $C \in \mathcal{d}[\mathcal{A}] \Rightarrow C \in \mathcal{D}_1$  then using the definition of  $\mathcal{D}_1$  we have  $A \cap C \in \mathcal{d}[\mathcal{A}]$  which proves that  $A \in \mathcal{D}_2$ . Hence

$$\mathcal{A} \subseteq \mathcal{D}_2 \quad (19.293)$$

Further as  $X \in \mathcal{d}[\mathcal{A}]$  and  $\forall C \in \mathcal{d}[\mathcal{A}]$   $X \cap C = C \in \mathcal{d}[\mathcal{A}]$  it follows that

$$X \in \mathcal{D}_2 \quad (19.294)$$

If  $A, B \in \mathcal{D}_2$  with  $B \subseteq A$  and  $C \in \mathcal{d}[\mathcal{A}]$  then  $B \cap C \subseteq A \cap C$  and by definition  $A \cap C, B \cap C \in \mathcal{d}[\mathcal{A}]$ , so as  $(A \setminus B) \cap C \stackrel{19.2}{=} (A \cap C) \setminus (B \cap C) \in \mathcal{d}[\mathcal{A}]$  it follows that

$$\forall A, B \in \mathcal{D}_2 \text{ we have } A \setminus B \in \mathcal{D}_2 \quad (19.295)$$

Further if  $\{A_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{D}_2$  is increasing and  $C \in \mathcal{d}[\mathcal{A}]$  then  $\forall i \in \mathbb{N}_0$  we have  $A_i \cap C \in \mathcal{d}[\mathcal{A}]$ , as  $(\bigcup_{i \in \mathbb{N}_0} A_i) \cap C \stackrel{1.101}{=} \bigcup_{i \in \mathbb{N}_0} (A_i \cap C) \in \mathcal{d}[\mathcal{A}]$  it follows, as  $\{A_i \cap C\}_{i \in \mathbb{N}_0}$  is clearly increasing, that  $(\bigcup_{i \in \mathbb{N}_0} A_i) \cap C \in \mathcal{d}[\mathcal{A}]$ . So we have

$$\forall \{A_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{D}_2 \text{ with } \forall i \in \mathbb{N}_0 A_i \subseteq A_{i+1} \text{ we have } \bigcup_{i \in \mathbb{N}_0} A_i \in \mathcal{D}_2 \quad (19.296)$$

Using 19.293, 19.294, 19.295 and 19.296 we have that  $\mathcal{D}_2$  is a Dynkin system containing  $\mathcal{A}$  so that

$$\mathcal{d}[\mathcal{A}] \subseteq \mathcal{D}_2 \quad (19.297)$$

If  $A, B \in \mathcal{d}[\mathcal{A}]$  then by the above  $A \in \mathcal{D}_2$  and still  $B \in \mathcal{d}[\mathcal{A}]$ , so that using the definition of  $\mathcal{D}_2$  it follows that  $A \cap B \in \mathcal{d}[\mathcal{A}]$ . Hence the Dynkin system  $\mathcal{d}[\mathcal{A}]$  is a  $\pi$ -system. Next we can apply 19.133 proving that  $\mathcal{d}[\mathcal{A}]$  is a  $\sigma$ -algebra which as  $\mathcal{A} \subseteq \mathcal{d}[\mathcal{A}]$  gives

$$\sigma[\mathcal{A}] \subseteq \mathcal{d}[\mathcal{A}] \quad (19.298)$$

Finally as every  $\sigma$ -algebra is a Dynkin system [see 19.127] we have that  $\sigma[\mathcal{A}]$  is a Dynkin class containing  $\mathcal{A}$  which proves that  $\mathcal{d}[\mathcal{A}] \subseteq \sigma[\mathcal{A}]$ , combining this with 19.298 gives

$$\sigma[\mathcal{A}] = \mathcal{d}[\mathcal{A}] \quad \square$$

**Corollary 19.135.** *Let  $\langle X, \mathcal{A} \rangle$  be a measurable space,  $\mathcal{C}$  a  $\pi$ -system on  $X$  such that  $\mathcal{A} = \sigma[\mathcal{C}]$  [hence  $\mathcal{C} \subseteq \mathcal{A}$ ]. If  $\mu, \nu$  are two finite measures on  $\mathcal{A}$  such that  $\mu(X) = \nu(X)$  and  $\forall C \in \mathcal{C} \mu(C) = \nu(C)$  then  $\mu = \nu$*

**Proof.** Define  $\mathcal{D} = \{A \in \mathcal{A} \mid \mu(A) = \nu(A)\}$  then

$$\mathcal{C} \subseteq \mathcal{D} \quad (19.299)$$

Further we can easily proof that  $\mathcal{D}$  is a Dynkin system

1. By assumption  $\mu(X) = \nu(X)$  proving that  $X \in \mathcal{D}$
2. If  $A, B \in \mathcal{D}$  with  $B \subseteq A$  then  $\mu(A) = \nu(A)$ ,  $\mu(B) = \nu(B)$  and

$$\begin{aligned} \mu(A \setminus B) &\stackrel{\mu \text{ is finite and 19.64}}{=} \mu(A) - \mu(B) \\ &= \nu(A) - \nu(B) \\ &\stackrel{\nu \text{ is finite and 19.64}}{=} \nu(A \setminus B) \end{aligned}$$

proving that  $A \setminus B \in \mathcal{D}$

3. If  $\{A_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{D}$  is increasing then  $\forall i \in \mathbb{N}_0$  we have that  $\mu(A_i) = \nu(A_i)$ . As  $\{A_i\}_{i \in \mathbb{N}_0}$  is increasing we have by 19.67 that  $\mu(\bigcup_{i \in \mathbb{N}_0} A_i) = \lim_{i \rightarrow \infty} \mu(A_i) = \lim_{i \rightarrow \infty} \nu(A_i) = \nu(\bigcup_{i \in \mathbb{N}_0} A_i)$  proving that  $\bigcup_{i \in \mathbb{N}_0} A_i \in \mathcal{D}$

As  $\mathcal{C} \subseteq \mathcal{D}$  [A Dynkin system] we have that  $\mathcal{d}[\mathcal{C}] \subseteq \mathcal{D}$ . Further as  $\mathcal{C}$  is a  $\pi$ -system we have by the previous theorem [see 19.134] that  $\sigma[\mathcal{C}] = \mathcal{d}[\mathcal{A}]$  proving that  $\mathcal{A} = \sigma[\mathcal{C}] \subseteq \mathcal{D}$ . Hence we have  $\forall A \in \mathcal{A}$  that  $\mu(A) = \nu(A)$  which finally proves that

$$\mu = \nu \quad \square$$

The above corollary works only for finite measures and is not very useful as most interesting measures are not finite, this is corrected in the following corollary.

**Corollary 19.136.** Let  $\langle X, \mathcal{A} \rangle$  be a measurable space,  $\mathcal{C}$  a  $\pi$ -system on  $X$  such that  $\mathcal{A} = \sigma[\mathcal{C}]$  [hence  $\mathcal{C} \subseteq \mathcal{A}$ ]. If  $\mu, \nu$  are measures on  $\mathcal{A}$  such that  $\forall C \in \mathcal{C}$  we have  $\mu(C) = \nu(C)$  and there exists a  $\{C_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{C}$  with  $\bigcup_{i \in \mathbb{N}_0} C_i = X$  and  $\forall i \in \mathbb{N}_0 \quad C_i \subseteq C_{i+1}$ ,  $\mu(C_i) < \infty$  then we have that  $\mu = \nu$

**Proof.** For each  $n \in \mathbb{N}_0$  define  $\mu_n: \mathcal{A} \rightarrow [0, \infty]$  by  $\mu_n(A) = \mu(A \cap C_n)$  and  $\nu_n: \mathcal{A} \rightarrow [0, \infty]$  by  $\nu_n(A) = \nu(A \cap C_n)$ . Then we have  $\forall n \in \mathbb{N}_0$  that

1.  $\mu_n(\emptyset) = \mu(\emptyset \cap C_n) = \mu(\emptyset) = 0$  and  $\nu_n(\emptyset) = \nu(\emptyset \cap C_n) = \nu(\emptyset) = 0$
2. If  $\{A_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{A}$  is a pairwise disjoint sequence of sets in  $\mathcal{A}$  then  $\{A_i \cap C_n\}_{i \in \mathbb{N}_0}$  is also a pairwise disjoint sequence of sets in  $\mathcal{A}$ . So

$$\mu_n\left(\bigsqcup_{i \in \mathbb{N}_0} A_i\right) = \mu\left(\left(\bigsqcup_{i \in \mathbb{N}_0} A_i\right) \cap C_n\right) = \mu\left(\bigsqcup_{i \in \mathbb{N}_0} (A_i \cap C_i)\right) = \sum_{i=1}^{\infty} \mu(A_i \cap C_i) = \sum_{i=1}^{\infty} \mu_n(A_i)$$

and

$$\nu_n\left(\bigsqcup_{i \in \mathbb{N}_0} A_i\right) = \nu\left(\left(\bigsqcup_{i \in \mathbb{N}_0} A_i\right) \cap C_n\right) = \nu\left(\bigsqcup_{i \in \mathbb{N}_0} (A_i \cap C_i)\right) = \sum_{i=1}^{\infty} \nu(A_i \cap C_i) = \sum_{i=1}^{\infty} \nu_n(A_i)$$

From (1) and (2) it follows that  $\forall n \in \mathbb{N}_0 \mu_n, \nu_n$  are measures on  $\mathcal{A}$ . Further for  $n \in \mathbb{N}_0$  and  $A \in \mathcal{A}$  we have that  $\mu_n(A) = \mu(A \cap C_n) \leq \mu(C_n) < \infty$  and  $\nu_n(A) = \nu(A \cap C_n) \leq \nu(C_n) < \infty$ . Hence we have that

$$\forall n \in \mathbb{N}_0 \mu_n, \nu_n \text{ are finite measures on } \mathcal{A} \quad (19.300)$$

Using then the previous corollary 19.135 we have that

$$\forall n \in \mathbb{N}_0 \mu_n = \nu_n \quad (19.301)$$

Finally if  $A \in \mathcal{A}$  then  $A = A \cap X = A \cap (\bigcup_{i \in \mathbb{N}_0} C_i) = \bigcup_{i \in \mathbb{N}_0} (A \cap C_i)$  where  $\{A \cap C_i\}_{i \in \mathbb{N}_0}$  is increasing hence we can apply 19.67 giving

$$\begin{aligned} \mu(A) &= \mu\left(\bigcup_{i \in \mathbb{N}_0} (A \cap C_i)\right) \\ &\stackrel{19.67}{=} \lim_{i \rightarrow \infty} \mu(A \cap C_i) \\ &= \lim_{i \rightarrow \infty} \mu_i(A) \\ &\stackrel{19.301}{=} \lim_{i \rightarrow \infty} \nu_i(A) \\ &= \lim_{i \rightarrow \infty} \nu(A \cap C_i) \\ &\stackrel{19.67}{=} \mu\left(\bigcup_{i \in \mathbb{N}_0} (A \cap C_i)\right) \\ &= \nu(A) \end{aligned}$$

which proves that

$$\mu = \nu$$

□

## 19.5 Measurable functions

### 19.5.1 Measurable functions

**Definition 19.137.** Let  $\langle X, \mathcal{A} \rangle, \langle Y, \mathcal{B} \rangle$  be two measurable spaces spaces,  $A \in \mathcal{A}$  and  $f: A \rightarrow Y$  a function then  $f$  [the graph of the function] is  **$\mathcal{A}, \mathcal{B}$ -measurable** if  $\forall B \in \mathcal{B}$  we have that  $f^{-1}(B) \in \mathcal{A}$ . The set of  $\mathcal{A}, \mathcal{B}$ -measurable function graphs is noted as  $\mathcal{M}[A, \mathcal{A}; Y, \mathcal{B}]$ , so

$$\mathcal{M}[A, \mathcal{A}; Y, \mathcal{B}] = \{f \in Y^X \mid f \text{ is } \mathcal{A}, \mathcal{B}\text{-measurable}\}$$

A function whose graph is  $\mathcal{A}, \mathcal{B}$ -measurable is called a  $\mathcal{A}, \mathcal{B}$ -measurable function.

**Note 19.138.** A set can have many  $\sigma$ -algebras defined on it, so the same function can at the same time be measurable and not measurable, depending on the  $\sigma$ -algebras used.

**Definition 19.139.** Let  $\langle X, \mathcal{T}_X \rangle, \langle Y, \mathcal{T}_Y \rangle$  be a topological spaces,  $f: X \rightarrow Y$  a function then  $f$  is **Borel measurable** if  $f$  is  $\mathcal{B}[X, \mathcal{T}_X], \mathcal{B}[Y, \mathcal{T}_Y]$ -measurable where  $\mathcal{B}[X, \mathcal{T}_X] = \sigma[\mathcal{T}_X]$  and  $\mathcal{B}[Y, \mathcal{T}_Y] = \sigma[\mathcal{T}_Y]$

In case of generated  $\sigma$ -algebras we have an easier way of checking that a function is measurable as is shown in the next theorem.

**Theorem 19.140.** Let  $\langle X, \mathcal{A} \rangle$  be a measurable space,  $A \in \mathcal{A}$ ,  $Y$  a set and  $\mathcal{B} \subseteq \mathcal{P}(Y)$  so that  $\langle Y, \sigma[\mathcal{B}] \rangle$  is a measurable space,  $f: A \rightarrow Y$  a function then  $f$  is  $\mathcal{A}, \sigma[\mathcal{B}]$ -measurable if and only if  $\forall B \in \mathcal{B}$  we have  $f^{-1}(B) \in \mathcal{A}$ .

**Proof.**

$\Rightarrow$ . If  $f$  is  $\mathcal{A}, \sigma[\mathcal{B}]$ -measurable then if  $B \in \mathcal{B}$  we have as  $\mathcal{B} \subseteq \sigma[\mathcal{B}]$  that  $f^{-1}(B) \in \mathcal{A}$

$\Leftarrow$ . Assume that  $\forall B \in \mathcal{B}$  we have that  $f^{-1}(B) \in \mathcal{A}$ . Define now  $\mathcal{H} = \{B \in \mathcal{P}(Y) | f^{-1}(B) \in \mathcal{A}\}$  then by the assumption we have that

$$\mathcal{B} \subseteq \mathcal{H} \quad (19.302)$$

As  $f^{-1}(Y) = A \in \mathcal{A}$  we have

$$Y \in \mathcal{H} \quad (19.303)$$

If  $B \in \mathcal{H}$  then  $f^{-1}(B) \in \mathcal{A}$  and  $f^{-1}(Y \setminus B) \stackrel{2.58}{=} A \setminus f^{-1}(B) \in \mathcal{A}$  [as  $\mathcal{A}$  is a sigma algebra and  $A \in \mathcal{A}$ ] giving

$$\forall B \in \mathcal{H} \text{ we have } Y \setminus B \in \mathcal{H} \quad (19.304)$$

Let  $\{B_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{H}$  then as  $f^{-1}(\bigcup_{i \in \mathbb{N}_0} B_i) \stackrel{2.62}{=} \bigcup_{i \in \mathbb{N}_0} f^{-1}(B_i) \in \mathcal{A}$  [as  $\mathcal{A}$  is a sigma algebra] proving

$$\forall \{B_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{H} \text{ we have } \bigcup_{i \in \mathbb{N}_0} B_i \in \mathcal{A} \quad (19.305)$$

From 19.303, 19.304 and 19.305 it follows that  $\mathcal{H}$  is a  $\sigma$ -algebra on  $Y$  that by 19.302 contains  $\mathcal{B}$ . Hence  $\sigma[\mathcal{B}] \subseteq \mathcal{H}$  which means that  $\forall B \in \sigma[\mathcal{B}]$   $f^{-1}(B) \in \mathcal{A}$  proving that  $f$  is a  $\mathcal{A}, \sigma[\mathcal{B}]$ -measurable function.  $\square$

**Proposition 19.141.** Let  $\langle X, \mathcal{A} \rangle, \langle Y, \mathcal{B} \rangle$  be measurable spaces,  $A, B \in \mathcal{A}$  with  $B \subseteq A$  and  $f: A \rightarrow Y$  a  $\mathcal{A}, \mathcal{B}$ -measurable function then  $f|_B: B \rightarrow Y$  is  $\mathcal{A}, \mathcal{B}$ -measurable

**Proof.** Let  $C \in \mathcal{B}$  then  $(f|_B)^{-1}(C) \stackrel{2.30}{=} B \cap f^{-1}(C) \in \mathcal{A}$  [as  $B \in \mathcal{A}$  and  $f^{-1}(C) \in \mathcal{A}$  as  $f = \mathcal{A}, \mathcal{B}$ -measurable]  $\square$

**Proposition 19.142.** Let  $\langle X, \mathcal{A} \rangle, \langle Y, \mathcal{B} \rangle$  be measurable spaces,  $A \in \mathcal{A}$ ,  $\{B_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{A}$  such that  $\bigcup_{i \in \mathbb{N}_0} B_i = A$  and  $f: A \rightarrow Y$  be such that  $\forall i \in \mathbb{N}_0 f|_{B_i}$  is  $\mathcal{A}, \mathcal{B}$ -measurable then  $f$  is  $\mathcal{A}, \mathcal{B}$ -measurable.

**Proof.** Let  $C \in \mathcal{B}$  then

$$\begin{aligned} f^{-1}(C) &= A \bigcap f^{-1}(C) \\ &= \left( \bigcup_{i \in \mathbb{N}_0} B_i \right) \bigcap f^{-1}(C) \\ &\stackrel{1.101}{=} \bigcup_{i \in \mathbb{N}_0} (B_i \bigcap f^{-1}(C)) \\ &\stackrel{2.30}{=} \bigcup_{i \in \mathbb{N}_0} (f|_{B_i})^{-1}(C) \end{aligned}$$

As  $\forall i \in \mathbb{N}_0 f|_{B_i}$  is  $\mathcal{A}, \mathcal{B}$ -measurable, we have that  $(f|_{B_i})^{-1}(C) \in \mathcal{A}$  so that by the above it follows that  $f^{-1}(C) \in \mathcal{A}$ .  $\square$

**Proposition 19.143.** *Let  $\langle X, \mathcal{T}_X \rangle, \langle Y, \mathcal{T}_Y \rangle$  and  $f: X \rightarrow Y$  a continuous function then  $f$  is  $\mathcal{B}[X, \mathcal{T}_X], \mathcal{B}[Y, \mathcal{T}_Y]$ -measurable. In other words continuous functions are Borel measurable.*

**Proof.** If  $U \in \mathcal{T}_Y$  then as  $f$  is continuous we have  $f^{-1}(U) \in \mathcal{T}_X \subseteq \mathcal{B}[X, \mathcal{T}_X]$  which using 19.140 proves that  $f$  is  $\mathcal{B}[X, \mathcal{T}_X], \mathcal{B}[Y, \mathcal{T}_Y]$ -measurable.  $\square$

**Proposition 19.144.** *Let  $\langle X, \mathcal{A}_X \rangle, \langle Y, \mathcal{A}_Y \rangle, \langle Z, \mathcal{A}_Z \rangle$ , measurable spaces,  $A \in \mathcal{A}_X, f: A \rightarrow Y$  a  $\mathcal{A}_X, \mathcal{A}_Y$ -measurable function,  $B \in \mathcal{A}_Y$  such that  $f(A) \subseteq B$  and  $g: B \rightarrow Z$  a  $\mathcal{A}_Y, \mathcal{A}_Z$ -measurable function then  $g \circ f$  is a  $\mathcal{A}_X, \mathcal{A}_Z$  measurable function.*

**Proof.** Take  $A \in \mathcal{A}_Z$  then  $g^{-1}(A) \in \mathcal{A}_Y$  so that  $f^{-1}(g^{-1}(A)) \in \mathcal{A}_X$ . Finally  $(g \circ f)^{-1}(A) \stackrel{2.59}{=} f^{-1}(g^{-1}(A)) \in \mathcal{A}_X$  proving the proposition.  $\square$

We have two interesting cases of measurable functions, when the image of a function consists of real numbers or extended real numbers.

**Definition 19.145.** *Let  $\langle X, \mathcal{A} \rangle$  be a measurable space,  $A \in \mathcal{A}$  then if  $f: A \rightarrow \bar{\mathbb{R}}$  is  $\mathcal{A}, \mathcal{B}[\bar{\mathbb{R}}]$ -measurable [see 19.31] then we say that  $f$  is  $\mathcal{A}$ -measurable. The set of  $\mathcal{A}$ -measurable function graphs is noted as  $\overline{\mathcal{M}[A, \mathcal{A}]}$ , so*

$$\overline{\mathcal{M}[A, \mathcal{A}]} = \{f \in \bar{\mathbb{R}}^A \mid f \text{ is } \mathcal{A}, \mathcal{B}[\bar{\mathbb{R}}]\text{-measurable}\} = \mathcal{M}[A, \mathcal{A}; \bar{\mathbb{R}}, \mathcal{B}[\bar{\mathbb{R}}]] \subseteq \bar{\mathbb{R}}^A$$

If  $\langle X, \mathcal{A} \rangle = \langle \mathbb{R}, \mathcal{L}[\mathbb{R}] \rangle, A \in \mathcal{L}[\mathbb{R}]$  /or  $\langle X, \mathcal{A} \rangle = \langle \mathbb{R}^n, \mathcal{L}[\mathbb{R}^n] \rangle, A \in \mathcal{L}[\mathbb{R}^n]$  / the Lebesgue measurable sets on  $\mathbb{R}$ /or  $\mathbb{R}^n$  /see 19.81 and 19.86/ then a function  $f: A \rightarrow \bar{\mathbb{R}}$  /or  $f: A \rightarrow \bar{\mathbb{R}}^n$  / is Lebesgue-measurable if  $f$  is  $\mathcal{L}[\mathbb{R}]$ -measurable /or  $f$  is  $\mathcal{L}[\mathbb{R}^n]$  measurable/.

We have an equivalent definition for real functions

**Definition 19.146.** *Let  $\langle X, \mathcal{A} \rangle$  be a measurable space and  $A \in \mathcal{A}$  then if  $f: A \rightarrow \mathbb{R}$  is  $\mathcal{A}, \mathcal{B}[\mathbb{R}]$ -measurable [see 19.26] we say that  $f$  is  $\mathcal{A}$ -measurable. The set of all  $\mathcal{A}$ -measurable function graphs is noted as  $\mathcal{M}[A, \mathcal{A}]$ , so*

$$\mathcal{M}[A, \mathcal{A}] = \{f \in \mathbb{R}^A \mid f \text{ is } \mathcal{A}, \mathcal{B}[\mathbb{R}]\text{-measurable}\} = \mathcal{M}[A, \mathcal{A}; \mathbb{R}, \mathcal{B}\{\mathbb{R}\}] \subseteq \mathbb{R}^A$$

**Note 19.147.** Altough we have proved that the composition of measurable functions is measurable, this does not hold for Lebesgue measurable function. Notice the following differences for the composition of measurable functions.

1. If  $\langle X, \mathcal{A}_X \rangle, \langle Y, \mathcal{A}_Y \rangle$  and  $\langle Z, \mathcal{A}_Z \rangle$  are measurable sets then if  $f: X \rightarrow Y$  is  $\mathcal{A}_X, \mathcal{A}_Y$ -measurable and  $g: Y \rightarrow Z$  is  $\mathcal{A}_Y, \mathcal{A}_Z$  then  $g \circ f$  is  $\mathcal{A}_X, \mathcal{A}_Z$  measurable [see 19.144]
2. If  $f, g: \mathbb{R} \rightarrow \mathbb{R}$  are Lebesgue measurable then  $f$  is  $\mathcal{L}[\mathbb{R}], \mathcal{B}[\mathbb{R}]$ -measurable and  $g$  is  $\mathcal{L}[\mathbb{R}], \mathcal{B}[\mathbb{R}]$ -measurable. As  $g$  is not  $\mathcal{B}[\mathbb{R}], \mathcal{B}[\mathbb{R}]$ -measurable we can not use 19.144 to proof that  $g \circ f$  is  $\mathcal{L}[\mathbb{R}], \mathcal{B}[\mathbb{R}]$ -measurable [hence Lebesgue measurable]. It is guaranteed that  $\forall A \in \mathcal{B}[\mathbb{R}] g^{-1}(A) \in \mathcal{L}[\mathbb{R}]$ , which does not guarantee that  $g^{-1}(A) \in \mathcal{B}[\mathbb{R}]$ . If however  $g$  is not only Lebesgue measurable but also Borel measurable then  $g$  is not only  $\mathcal{L}[\mathbb{R}], \mathcal{B}[\mathbb{R}]$ -measurable but also  $\mathcal{B}[\mathbb{R}], \mathcal{B}[\mathbb{R}]$ -measurable so that we can apply 19.144 proving that  $g \circ f$  is  $\mathcal{L}[\mathbb{R}], \mathcal{B}[\mathbb{R}]$ -measurable hence Lebesgue measurable. The difference is that we have on  $\mathbb{R}$  different  $\sigma$ -algebras, either  $\mathcal{L}[\mathbb{R}]$  or  $\mathcal{B}[\mathcal{A}]$ .
3. If  $f, g: \mathbb{R} \rightarrow \mathbb{R}$  are Borel measurable then  $f, g$  are  $\mathcal{B}[\mathbb{R}], \mathcal{B}[\mathbb{R}]$ -measurable so that we can use 19.144 to prove that  $g \circ f$  is  $\mathcal{B}[\mathbb{R}], \mathcal{B}[\mathbb{R}]$ -measurable and thus Borel measurable.

Note that up to now we have to duplicate our statements between real functions and extended real functions. The following proposition allows us to significantly reduce the number of proofs and statements for measurability of real functions or extended real functions.

**Proposition 19.148.** Let  $\langle X, \mathcal{A} \rangle$  and  $A \in \mathcal{A}$  then

1. If  $f \in \overline{\mathcal{M}}[A, \mathcal{A}]$  is such that  $f(A) \subseteq \mathbb{R}$  then  $f \in \mathcal{M}[A, \mathcal{A}]$
2.  $\mathcal{M}[A, \mathcal{A}] \subseteq \overline{\mathcal{M}[A, \mathcal{A}]}$

**Proof.**

1. If  $B \in \mathcal{B}[\mathbb{R}]$  then as  $B \cap \mathbb{R} = B$  we have by 19.32 that  $B \in \mathcal{B}[\overline{\mathbb{R}}]$ , so as  $f \in \overline{\mathcal{M}[A, \mathcal{A}]}$  it follows that  $f^{-1}(B) \in \mathcal{A}$ , which proves that  $f \in \mathcal{M}[A, \mathcal{A}]$
2. If  $B \in \mathcal{B}[\overline{\mathbb{R}}]$  then by 19.32  $B \cap \mathbb{R} \in \mathcal{B}[\mathbb{R}]$  and as  $f \in \mathcal{M}[A, \mathcal{A}]$  we have that  $f^{-1}(B \cap \mathbb{R}) \in \mathcal{A}$ , further  $f^{-1}(B \cap \mathbb{R}) \stackrel{2.62}{=} f^{-1}(B) \cap f^{-1}(\mathbb{R}) = f^{-1}(B) \cap A = f^{-1}(B)$ . So that  $f^{-1}(B) \in \mathcal{A}$  which proves that  $(f: A \rightarrow \overline{\mathbb{R}}) \in \overline{\mathcal{M}[\mathbb{R}]}$ .  $\square$

The following definition will be used to find an alternative condition for a (extended) real function to be  $\mathcal{A}$ -measurable.

**Definition 19.149.** Let  $X$  be a set then we define

1. If  $f: X \rightarrow \overline{\mathbb{R}}, g: X \rightarrow \overline{\mathbb{R}}$  then given  $x \in \mathbb{R}$  we define

$$\begin{aligned} \{f < x\} &\stackrel{\text{def}}{=} f^{-1}([-\infty, x]) = \{t \in X \mid f(t) < x\} \\ \{f \leq x\} &\stackrel{\text{def}}{=} f^{-1}([-\infty, x]) = \{t \in X \mid f(t) \leq x\} \\ \{f > x\} &\stackrel{\text{def}}{=} f^{-1}(]x, \infty]) = \{t \in X \mid x < f(t)\} \\ \{f \geq x\} &\stackrel{\text{def}}{=} f^{-1}([x, \infty]) = \{t \in X \mid x \leq f(t)\} \\ \{f = x\} &\stackrel{\text{def}}{=} f^{-1}(\{x\}) = \{t \in X \mid x = f(t)\} \\ \{f \leq g\} &\stackrel{\text{def}}{=} \{x \in X \mid f(x) \leq g(x)\} \\ \{f < g\} &\stackrel{\text{def}}{=} \{x \in X \mid f(x) < g(x)\} \\ \{f = g\} &\stackrel{\text{def}}{=} \{x \in X \mid f(x) = g(x)\} \end{aligned}$$

2. If  $f: X \rightarrow \mathbb{R}, g: X \rightarrow \mathbb{R}$  then given  $x \in \mathbb{R}$  we define

$$\begin{aligned} \{f < x\} &\stackrel{\text{def}}{=} f^{-1}(]-\infty, x]) = \{t \in X \mid f(t) < x\} \\ \{f \leq x\} &\stackrel{\text{def}}{=} f^{-1}(]-\infty, x]) = \{t \in X \mid f(t) \leq x\} \\ \{f > x\} &\stackrel{\text{def}}{=} f^{-1}(]x, \infty]) = \{t \in X \mid x < f(t)\} \\ \{f \geq x\} &\stackrel{\text{def}}{=} f^{-1}([x, \infty]) = \{t \in X \mid x \leq f(t)\} \\ \{f = x\} &\stackrel{\text{def}}{=} f^{-1}(\{x\}) = \{t \in X \mid x = f(t)\} \\ \{f \leq g\} &\stackrel{\text{def}}{=} \{x \in X \mid f(x) \leq g(x)\} \\ \{f < g\} &\stackrel{\text{def}}{=} \{x \in X \mid f(x) < g(x)\} \\ \{f = g\} &\stackrel{\text{def}}{=} \{x \in X \mid f(x) = g(x)\} \end{aligned}$$

**Lemma 19.150.** Let  $X$  be a set,  $f: X \rightarrow \overline{\mathbb{R}}$  and  $x \in \mathbb{R}$  then

1.  $\{f \geq x\} = \bigcap_{n \in \mathbb{N}_0} \{f > x - \frac{1}{n}\}$
2.  $X \setminus \{f \geq x\} = \{f < x\}$
3.  $\{f \leq x\} = \bigcap_{n \in \mathbb{N}_0} \{f < x + \frac{1}{n}\}$
4.  $X \setminus \{f \leq x\} = \{f > x\}$

**Proof.**

1. Let  $n \in \mathbb{N}$  then if  $t \in \{f \geq x\}$  we have that  $x \leq f(t) \Rightarrow x - \frac{1}{n} < f(t) \Rightarrow t \in \{x - \frac{1}{n} < f\}$  proving

$$\{f \geq x\} \subseteq \bigcap_{n \in \mathbb{N}_0} \left\{ f > x - \frac{1}{n} \right\} \quad (19.306)$$

If  $t \in \bigcap_{n \in \mathbb{N}_0} \{f > x - \frac{1}{n}\}$  then  $\forall n \in \mathbb{N}_0$  we have  $x - \frac{1}{n} < f(t)$ . Assume that  $f(t) < x$  then by 17.8 we have that there exists a  $n \in \mathbb{N}_0$  such that  $f(t) + \frac{1}{n} < x \Rightarrow f(t) < x - \frac{1}{n}$  a contradiction so we must have  $x \leq f(t)$  or  $t \in \{f \geq x\}$ . So  $\bigcap_{n \in \mathbb{N}_0} \{f > x - \frac{1}{n}\} \subseteq \{f \geq x\}$ . Combining this with 19.306 proves

$$\{f \geq x\} = \bigcap_{n \in \mathbb{N}_0} \left\{ f > x - \frac{1}{n} \right\}$$

2.

$$\begin{aligned} X \setminus \{f \geq x\} &= X \setminus f^{-1}([x, \infty)) \\ &= f^{-1}(\bar{\mathbb{R}}) \setminus f^{-1}([x, \infty)) \\ &\stackrel{2.58}{=} f^{-1}(\bar{\mathbb{R}} \setminus [x, \infty)) \\ &= f^{-1}([-\infty, x]) \\ &= \{f < x\} \end{aligned}$$

3. Let  $n \in \mathbb{N}_0$  then if  $t \in \{f \leq x\}$  we have that  $f(t) \leq x \Rightarrow f(t) < x + \frac{1}{n}$  or  $t \in \{f < x + \frac{1}{n}\}$ . So

$$\{f \leq x\} \subseteq \bigcap_{n \in \mathbb{N}_0} \left\{ f < x + \frac{1}{n} \right\} \quad (19.307)$$

Let  $t \in \bigcap_{n \in \mathbb{N}_0} \{f < x + \frac{1}{n}\}$  then  $\forall n \in \mathbb{N}_0$  we have  $f(t) < x + \frac{1}{n}$ . Assume that  $f(t) > x$  then by 17.8 there exists a  $n \in \mathbb{N}_0$  such that  $x + \frac{1}{n} < f(t)$  a contradiction hence  $f(t) \leq x$  or  $t \in \{f \leq x\}$ . Combining this with 19.307 proves that

$$\{f \leq x\} \subseteq \bigcap_{n \in \mathbb{N}_0} \left\{ f < x + \frac{1}{n} \right\}$$

4.

$$\begin{aligned} X \setminus \{f \leq x\} &= X \setminus f^{-1}([-\infty, x]) \\ &= f^{-1}(\bar{\mathbb{R}}) \setminus f^{-1}([-\infty, x]) \\ &\stackrel{2.58}{=} f^{-1}(\bar{\mathbb{R}} \setminus [-\infty, x]) \\ &= f^{-1}([x, \infty)) \\ &= \{x < f\} \\ &\square \end{aligned}$$

**Theorem 19.151.** Let  $\langle X, \mathcal{A} \rangle$  be a measurable set,  $A \in \mathcal{A}$  and  $f: A \rightarrow \bar{\mathbb{R}}$  then we have the following equivalences:

1.  $f$  is  $\mathcal{A}$ -measurable
2.  $\forall x \in \mathbb{R}$  we have  $\{f > x\} \in \mathcal{A}$
3.  $\forall x \in \mathbb{R}$  we have  $\{f \geq x\} \in \mathcal{A}$
4.  $\forall x \in \mathbb{R}$  we have  $\{f < x\} \in \mathcal{A}$
5.  $\forall x \in \mathbb{R}$  we have  $\{f \leq x\} \in \mathcal{A}$

**Proof.**

**1  $\Rightarrow$  2.** Take  $x \in \mathbb{R}$  then  $[x, \infty] \in \sigma[\{a, \infty] | a \in \mathbb{R}\}] \stackrel{19.32}{=} \mathcal{B}[\bar{\mathbb{R}}]$  we have as  $f$  is  $\mathcal{A}$ -measurable that  $\{f > x\} = f^{-1}([x, \infty]) \in \mathcal{A}$

**2  $\Rightarrow$  1.** As  $\forall x \in \mathbb{R}$  we have  $f^{-1}([a, \infty)) = \{f > x\} \in \mathcal{A}$ , it follows from 19.140 and  $\sigma[\{a, \infty] | a \in \mathbb{R}\}] = \mathcal{B}[\bar{\mathbb{R}}]$  that  $f$  is  $\mathcal{A}, \mathcal{B}[\bar{\mathbb{R}}]$ -measurable or in other words  $f$  is  $\mathcal{A}$ -measurable.

**2  $\Rightarrow$  3.** Let  $x \in \mathbb{R}$  then using (2) we have  $\forall n \in \mathbb{N}_0$  that  $\{f > x - \frac{1}{n}\} \in \mathcal{A}$  hence  $\bigcap_{n \in \mathbb{N}_0} \{f > x - \frac{1}{n}\} \in \mathcal{A}$ . Using 19.150  $\{f \geq x\} = \bigcap_{n \in \mathbb{N}_0} \{f > x - \frac{1}{n}\}$  proving that  $\{f \geq x\} \in \mathcal{A}$ .

**3  $\Rightarrow$  4.** Let  $x \in \mathbb{R}$  then  $\{f < x\} \stackrel{19.150}{=} A \setminus \{f \geq x\} \in \mathcal{A}$  [as  $A, \{f \geq x\} \in \mathcal{A}$ , together with 19.10]

**4  $\Rightarrow$  5.** Let  $x \in \mathbb{R}$  then by (4) we have that  $\forall n \in \mathbb{N}_0$  that  $\{f < x + \frac{1}{n}\} \in \mathcal{A}$  hence  $\bigcap_{n \in \mathbb{N}_0} \{f < x + \frac{1}{n}\} \in \mathcal{A}$ . Using 19.150  $\{f \leq x\} = \bigcap_{n \in \mathbb{N}_0} \{f < x + \frac{1}{n}\}$  proving that  $\{f \leq x\} \in \mathcal{A}$

**5  $\Rightarrow$  1.** Let  $x \in \mathbb{R}$  then  $\{x < f\} \underset{19.150}{=} A \setminus \{f \leq x\} \in \mathcal{A}$  [as  $A \in \mathcal{A}$  together with 19.10]  $\square$

We have a similar result for real functions and the proofs are essentially the same with the exception of some details.

**Lemma 19.152.** *Let  $X$  be a set,  $f: X \rightarrow \mathbb{R}$  and  $x \in X$  then*

1.  $\{f \geq x\} = \bigcap_{n \in \mathbb{N}_0} \{f > x - \frac{1}{n}\}$
2.  $X \setminus \{f \geq x\} = \{f < x\}$
3.  $\{f \leq x\} = \bigcap_{n \in \mathbb{N}_0} \{f < x + \frac{1}{n}\}$
4.  $X \setminus \{f \leq x\} = \{f > x\}$

**Proof.**

1. Let  $n \in \mathbb{N}$  then if  $t \in \{f \geq x\}$  we have that  $x \leq f(t) \Rightarrow x - \frac{1}{n} < f(t) \Rightarrow t \in \{f > x - \frac{1}{n}\}$  proving

$$\{f \geq x\} \subseteq \bigcap_{n \in \mathbb{N}_0} \left\{ f > x - \frac{1}{n} \right\} \quad (19.308)$$

If  $t \in \bigcap_{n \in \mathbb{N}_0} \{f > x - \frac{1}{n}\}$  then  $\forall n \in \mathbb{N}_0$  we have  $x - \frac{1}{n} < f(t)$ . Assume that  $f(t) < x$  then by 9.56 we have that there exists a  $n \in \mathbb{N}_0$  such that  $f(t) + \frac{1}{n} < x \Rightarrow f(t) < x - \frac{1}{n}$  a contradiction, so we must have  $x \leq f(t)$  or  $t \in \{f \geq x\}$ . Hence  $\bigcap_{n \in \mathbb{N}_0} \{f > x - \frac{1}{n}\} \subseteq \{f \geq x\}$ . Combining this with 19.308 proves

$$\{f \geq x\} = \bigcap_{n \in \mathbb{N}_0} \left\{ f > x - \frac{1}{n} \right\}$$

2.

$$\begin{aligned} X \setminus \{f \geq x\} &= X \setminus f^{-1}([x, \infty[) \\ &= f^{-1}(\mathbb{R}) \setminus f^{-1}([x, \infty[) \\ &\stackrel{2.58}{=} f^{-1}(\mathbb{R} \setminus [x, \infty[) \\ &= f^{-1}(]-\infty, x[) \\ &= \{f < x\} \end{aligned}$$

3. Let  $n \in \mathbb{N}_0$  then if  $t \in \{f \leq x\}$  we have that  $f(t) \leq x \Rightarrow f(t) < x + \frac{1}{n}$  or  $t \in \{f < x + \frac{1}{n}\}$ . So

$$\{f \leq x\} \subseteq \bigcap_{n \in \mathbb{N}_0} \left\{ f < x + \frac{1}{n} \right\} \quad (19.309)$$

Let  $t \in \bigcap_{n \in \mathbb{N}_0} \{f < x + \frac{1}{n}\}$  then  $\forall n \in \mathbb{N}_0$  we have that  $f(t) < x + \frac{1}{n}$ . Assume that  $x < f(t)$  then by 9.56 there exists a  $n \in \mathbb{N}_0$  such that  $x + \frac{1}{n} < f(t)$  a contradiction hence  $f(t) \leq x$  or  $t \in \{f \leq x\}$ . Combining this with 19.309 proves that

$$\{f \leq x\} \subseteq \bigcap_{n \in \mathbb{N}_0} \left\{ f < x + \frac{1}{n} \right\}$$

4.

$$\begin{aligned} X \setminus \{f \leq x\} &= X \setminus f^{-1}(]-\infty, x[) \\ &= f^{-1}(\mathbb{R}) \setminus f^{-1}(]-\infty, x[) \\ &\stackrel{2.58}{=} f^{-1}(\mathbb{R} \setminus ]-\infty, x[) \\ &= f^{-1}(]x, \infty[) \\ &= \{f > x\} \end{aligned}$$

$\square$

**Theorem 19.153.** Let  $\langle X, \mathcal{A} \rangle$  be a measurable set,  $A \in \mathcal{A}$  and  $f: A \rightarrow \mathbb{R}$  then we have the following equivalences:

1.  $f$  is  $\mathcal{A}$ -measurable
2.  $\forall x \in \mathbb{R}$  we have  $\{f > x\} \in \mathcal{A}$
3.  $\forall x \in \mathbb{R}$  we have  $\{f \geq x\} \in \mathcal{A}$
4.  $\forall x \in \mathbb{R}$  we have  $\{f < x\} \in \mathcal{A}$
5.  $\forall x \in \mathbb{R}$  we have  $\{f \leq x\} \in \mathcal{A}$

**Proof.**

**1  $\Rightarrow$  2.** Take  $x \in \mathbb{R}$  then  $]x, \infty[ \in \sigma[\{a, \infty | a \in \mathbb{R}\}]$   $\stackrel{19.26}{=} \mathcal{B}[\mathbb{R}]$  we have as  $f$  is  $\mathcal{A}$ -measurable that  $\{f > x\} = f^{-1}(]x, \infty[) \in \mathcal{A}$

**2  $\Rightarrow$  1.** As  $\forall x \in \mathbb{R}$  we have  $f^{-1}(]x, \infty[) = \{f > x\} \in \mathcal{A}$ , it follows from 19.140 and  $\sigma[\{a, \infty | a \in \mathbb{R}\}] \stackrel{19.26}{=} \mathcal{B}[\mathbb{R}]$  that  $f$  is  $\mathcal{A}, \mathcal{B}[\mathbb{R}]$ -measurable or in other words  $f$  is  $\mathcal{A}$ -measurable.

**2  $\Rightarrow$  3.** Let  $x \in \mathbb{R}$  then using (2) we have  $\forall n \in \mathbb{N}_0$  that  $\{f > x - \frac{1}{n}\} \in \mathcal{A}$  hence  $\bigcap_{n \in \mathbb{N}_0} \{f > x - \frac{1}{n}\} \in \mathcal{A}$ . Using 19.152  $\{f \geq x\} = \bigcap_{n \in \mathbb{N}_0} \{f > x - \frac{1}{n}\}$  proving that  $\{f \geq x\} \in \mathcal{A}$

**3  $\Rightarrow$  4.** Let  $x \in \mathbb{R}$  then  $\{f < x\} \stackrel{19.152}{=} A \setminus \{f \geq x\} \in \mathcal{A}$  [as  $A, \{f \geq x\} \in \mathcal{A}$  together with 19.10]

**4  $\Rightarrow$  5.** Let  $x \in \mathbb{R}$  then by (4) we have that  $\forall n \in \mathbb{N}_0$  that  $\{f < x + \frac{1}{n}\} \in \mathcal{A}$  hence  $\bigcap_{n \in \mathbb{N}_0} \{f < x + \frac{1}{n}\} \in \mathcal{A}$ . Using 19.152  $\{f \leq x\} = \bigcap_{n \in \mathbb{N}_0} \{f < x + \frac{1}{n}\}$  proving that  $\{f \leq x\} \in \mathcal{A}$

**5  $\Rightarrow$  1.** Let  $x \in \mathbb{R}$  then  $\{f > x\} \stackrel{19.152}{=} A \setminus \{f \leq x\} \in \mathcal{A}$  [as  $\{f \leq x\}, A \in \mathcal{A}$  together with 19.10]  $\square$

### 19.5.2 Example of measurable functions

**Example 19.154.**  $\langle X, \mathcal{A} \rangle, \langle Y, \mathcal{B} \rangle$  be measurable spaces then  $\emptyset: \emptyset \rightarrow Y$  defined by is  $\mathcal{A}, \mathcal{B}$ -measurable.

**Proof.** Let  $B \in \mathcal{B}$  then as  $\emptyset^{-1}(B) \stackrel{2.4}{=} \{x \in \emptyset | \exists y \in Y \vdash (x, y) \in \emptyset\} = \emptyset \in \mathcal{A}$   $\square$

**Example 19.155.** Let  $\langle X, \mathcal{A} \rangle$  be a measurable space,  $A \in \mathcal{A}$  and  $C_0: X \rightarrow \mathbb{R}$  defined by  $C_0(x) = 0$  then  $C_0$  is  $\mathcal{A}$ -measurable [or  $C_0 \subseteq \mathcal{M}[A, \mathcal{A}] \subseteq \overline{\mathcal{M}[A, \mathcal{A}]}$ ]. F

**Proof.** Take  $B \in \mathcal{B}[\overline{\mathbb{R}}]$  the we have either

**0  $\in B$ .** then

$$\begin{aligned} x \in C_0^{-1}(B) &\Leftrightarrow x \in A \wedge C_0(x) \in B \\ &\stackrel{C_0(x)=0 \in B}{\Leftrightarrow} x \in A \end{aligned}$$

proving that  $C_0^{-1}(B) = A \in \mathcal{A}$

**0  $\notin B$ .** then if  $x \in C_0^{-1}(B)$  we have  $0 = C_0(x) \in B$  a contradiction, hence  $C_0^{-1}(B) = \emptyset \in \mathcal{A}$

As in all case  $C_0^{-1}(B) \in \mathcal{A}$  it follows that  $C_0$  is  $\mathcal{A}$ -measurable.  $\square$

**Example 19.156.** Let  $\langle X, \mathcal{T} \rangle$  be a topological set,  $U$  a open set then

1. If  $f: U \rightarrow \overline{\mathbb{R}}$  is continuous [using the subspace topology  $\mathcal{T}_U = \{V \cap U | V \in \mathcal{T}\}$  [see 11.6]] then  $f$  is Borel measurable [more specific  $f$  is  $\mathcal{B}[X, \mathcal{T}], \mathcal{T}_{\overline{\mathbb{R}}}$ -measurable].
2. If  $f: U \rightarrow \mathbb{R}$  is continuous [using the subspace topology  $\mathcal{T}_U = \{V \cap U | V \in \mathcal{T}\}$  [see 11.6]] then  $f$  is Borel measurable [more specific  $f$  is  $\mathcal{B}[X, \mathcal{T}], \mathcal{T}_{\mathbb{R}}$ -measurable].

**Proof.**

1. Let  $V \in \mathcal{T}_{\mathbb{R}}$  then by continuity we have that  $f^{-1}(V) \in \mathcal{T}_U$ , so there exists a  $W \in \mathcal{T}$  such that  $f^{-1}(V) = W \cap U \in \mathcal{T}$  [as  $U, W \in \mathcal{T}$ ], hence  $f^{-1}(V) \in \mathcal{T} \subseteq \mathcal{B}[X, \mathcal{T}]$ . Finally using 19.140 it follow that  $f$  is  $\mathcal{B}[X, \mathcal{T}]$ ,  $\mathcal{T}_{\mathbb{R}}$ -measurable.
2. Let  $V \in \mathcal{T}_{\mathbb{R}}$  then by continuity we have that  $f^{-1}(V) \in \mathcal{T}_U$ , so there exists a  $W \in \mathcal{T}$  such that  $f^{-1}(V) = W \cap U \in \mathcal{T}$  [as  $U, W \in \mathcal{T}$ ] hence  $f^{-1}(V) \in \mathcal{T} \subseteq \mathcal{B}[X, \mathcal{T}]$ . Finally using 19.140 it follow that  $f$  is  $\mathcal{B}[X, \mathcal{T}]$ ,  $\mathcal{T}_{\mathbb{R}}$ -measurable.  $\square$

**Example 19.157.** Let  $\langle X, \mathcal{A} \rangle$  be a measurable space,  $A \in \mathcal{A}$  then  $\mathcal{X}_A \in \mathcal{M}[X, \mathcal{A}]$  where  $\mathcal{X}_A \in \mathbb{R}^X$  is defined by  $\mathcal{X}_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \in X \setminus A \end{cases}$

**Proof.** Let  $x \in \mathbb{R}$  then we must consider two cases:

**$1 \leq x$ .** then if  $t \in \{x < \mathcal{X}_A\}$  we have  $1 \leq \mathcal{X}_A(t)$  so we must have  $\mathcal{X}(t) = 1$  or  $t \in A$  proving that  $\{x < \mathcal{X}_A\} \subseteq A$ . Further if  $t \in A$  then  $\mathcal{X}_A(t) = 1 \geq 1$  so that  $A \subseteq \{x < \mathcal{X}_A\}$ . Hence  $\{x < \mathcal{X}_A\} = A \in \mathcal{A}$

**$x < 1$ .** then if  $t \in \{x < \mathcal{X}_A\}$  we have  $\mathcal{X}_A(t) < 1$  so we must have  $\mathcal{X}(t) = 0$  or  $t \in X \setminus A$  proving that  $\{x < \mathcal{X}_A\} \subseteq X \setminus A$ . Further if  $t \in X \setminus A$  then  $\mathcal{X}_A(t) = 0 < 1$  so that  $\{x < \mathcal{X}_A\} \subseteq X \setminus A$ . Hence  $\{x < \mathcal{X}_A\} = X \setminus A \in \mathcal{A}$ .

As in all cases  $\{x < \mathcal{X}_A\} \in \mathcal{A}$  we have by 19.151 (for (1)) [or 19.153 for (2)].  $\square$

**Lemma 19.158.** Let  $\langle X, \mathcal{A} \rangle$  be a measurable space  $f \in \overline{\mathcal{M}[X, \mathcal{A}]}$  be a  $\mathcal{A}$ -measurable function,  $A \in \mathcal{A}$  then  $\mathcal{X}_A \cdot f$  defined by  $(\mathcal{X}_A \cdot f)(x)$  is  $\mathcal{A}$ -measurable [in other words  $\mathcal{X}_A \cdot f \in \overline{\mathcal{M}[X, \mathcal{A}]}$ ]

**Proof.** Let  $x \in \mathbb{R}$  then we must consider the following cases for  $x$ :

**$x \geq 0$ .** Let  $t \in \{\mathcal{X}_A \cdot f > x\}$  we have  $\mathcal{X}_A(t) \cdot f(t) = (\mathcal{X}_A \cdot f)(x) > x \geq 0$  so that  $t \in A$  [otherwise  $\mathcal{X}_A(t) \cdot f(t) = 0$ ] and  $f(t) = \mathcal{X}_A(t) \cdot f(t) = (\mathcal{X}_A \cdot f)(x) > x$  giving  $t \in A \cap \{f > x\}$ . So we have

$$\{\mathcal{X}_A \cdot f > x\} \subseteq A \cap \{f > x\} \quad (19.310)$$

If  $t \in A \cap \{f > x\}$  then  $\mathcal{X}_A(t) = 1 \wedge f(t) > x$  so  $(\mathcal{X}_A \cdot f)(t) = \mathcal{X}_A(t) \cdot f(t) = f(t) > x$  or  $t \in \{\mathcal{X}_A \cdot f > x\}$  proving that  $A \cap \{f > x\} \subseteq \{\mathcal{X}_A \cdot f > x\}$ . Which combined with 19.310 proves

$$\{\mathcal{X}_A \cdot f > x\} = A \cap \{f > x\}$$

As  $A \in \mathcal{A}$  and  $\{f > x\} \in \mathcal{A}$  [as  $f$  is  $\mathcal{A}$ -measurable] it follows from the above that

$$\{\mathcal{X}_A \cdot f > x\} \in \mathcal{A}$$

**$x < 0$ .** Let  $t \in \{\mathcal{X}_A \cdot f > x\}$  then we have either

**$t \in A$ .** then  $f(t) = \mathcal{X}_A(t) \cdot f(t) = (\mathcal{X}_A \cdot f)(t) > x$  proving that  $t \in \{f > x\} \subseteq \{f > x\} \cup (X \setminus A)$

**$t \in X \setminus A$ .** then  $t \in \{f > x\} \cup (X \setminus A)$

hence

$$\{\mathcal{X}_A \cdot f > x\} \subseteq t \in \{f > x\} \subseteq \{f > x\} \cup (X \setminus A) \quad (19.311)$$

If  $t \in \{f > x\} \cup (X \setminus A)$  then either

**$t \in \{f > x\}$ .** Then if  $t \in X \setminus A$  we have  $(\mathcal{X}_A \cdot f)(t) = \mathcal{X}_A(t) \cdot f(t) = 0 > x$  and if  $t \in A$  we have  $(\mathcal{X}_A \cdot f)(t) = \mathcal{X}_A(t) \cdot f(t) = f(t) > x$  so that in all cases  $(\mathcal{X}_A \cdot f)(t) > x$ . Hence

$$t \in \{\mathcal{X}_A \cdot f > x\}$$

$t \in X \setminus A$ . Then  $(\mathcal{X}_A \cdot f)(t) = \mathcal{X}_A(t) \cdot f(t) = 0 > x$  or

$$t \in \{\mathcal{X}_A \cdot f > x\}$$

As in all cases we have  $t \in \{\mathcal{X}_A \cdot f > x\}$  it follows that  $\{f > x\} \cup (X \setminus A) \subseteq \{\mathcal{X}_A \cdot f > x\}$  it follows from 19.311 that  $\{f > x\} \cup (X \setminus A) = \{\mathcal{X}_A \cdot f > x\}$ . As  $X \setminus A \in \mathcal{A}$  and  $\{f > x\} \in \mathcal{A}$  it follows that

$$\{\mathcal{X}_A \cdot f > x\} \in \mathcal{A}$$

So in all cases we have  $\{\mathcal{X}_A \cdot f > x\}$  which proves by 19.151 that  $\mathcal{X}_A \cdot f$  is  $\mathcal{A}$ -measurable.  $\square$

**Definition 19.159.** Let  $X$  be a non empty set then a function  $f: X \rightarrow \bar{\mathbb{R}}$  is a **simple function** if  $f(X)$  is finite [is non empty as  $X$  is non empty].

**Proposition 19.160.** Let  $X$  be a set,  $\alpha \in \mathbb{R}$  and  $f: X \rightarrow \bar{\mathbb{R}}$ ,  $g: X \rightarrow \bar{\mathbb{R}}$  be simple functions then we have

1.  $\alpha \cdot f: X \rightarrow \bar{\mathbb{R}}$  defined by  $(\alpha \cdot f)(x) = \alpha \cdot f(x)$  is a simple function
2. If  $f(X) \subseteq \mathbb{R}$  and  $g(X) \subseteq \mathbb{R}$  then  $f + g: X \rightarrow \bar{\mathbb{R}}$  defined by  $(f + g)(x) = f(x) + g(x)$  is a simple function. Note: that the requirement  $f(X), g(X) \subseteq \mathbb{R}$  is needed to ensure that the sum is always defined.

**Proof.**

1. We have to consider two cases for  $\alpha$

$\alpha = 0$ . then  $(\alpha \cdot f)(X) = \{0\}$  which is finite and non empty

$\alpha \neq 0$ . as  $f$  is a simple function there exists a  $n \in \mathbb{N}_0$  and a bijection  $\beta: \{1, \dots, n\} \rightarrow f(X)$ .

Define now  $\alpha_f: \{1, \dots, n\} \rightarrow (\alpha \cdot f)(X)$  by  $\alpha_f(i) = \alpha \cdot \beta(i)$ . If  $y \in (\alpha \cdot f)(X)$  then there exists a  $x \in X$  such that  $y = (\alpha \cdot f)(x) = \alpha \cdot f(x) \Rightarrow f(x) = \frac{y}{\alpha}$ , as  $f(x) \in f(X)$  there exists a  $i \in \{1, \dots, n\}$  such that  $\beta(i) = f(x) = \frac{y}{\alpha} \Rightarrow y = \alpha \cdot \beta(i) = \alpha_f(i)$  proving that  $\alpha_f$  is a surjection. So using 5.48 we have that  $(\alpha \cdot f)(X)$  is finite.

so in all cases  $(\alpha \cdot f)(X)$  is finite proving that  $\alpha \cdot f$  is a simple function.

2. As  $f$  and  $g$  are simple functions we have that  $f(X)$  and  $g(X)$  are finite. As the product of finite sets is finite [see 5.44] we have that

$$f(X) \times g(X) \text{ is finite}$$

Define now

$$s: f(X) \times g(Y) \rightarrow \{x + y \mid (x, y) \in f(X) \times g(Y)\} \text{ by } s(x, y) = x + y$$

Then if  $z \in \{x + y \mid (x, y) \in f(X) \times g(Y)\}$  there exists a  $x \in f(X)$  and a  $y \in g(Y)$  such that  $z = x + y = s(x, y)$  which proves that  $s$  is surjective. Applying then 5.48 proves that

$$\{x + y \mid (x, y) \in f(X) \times g(Y)\} \text{ is finite}$$

Let  $y \in (f + g)(X)$  then there exists a  $x \in X$  such that  $y = f(x) + g(x) \in \{x + y \mid (x, y) \in f(X) \times g(Y)\}$ . Hence  $(f + g)(X) \subseteq \{x + y \mid (x, y) \in f(X) \times g(Y)\}$  which, as we have proved that  $\{x + y \mid (x, y) \in f(X) \times g(Y)\}$  is finite, proves using 5.36 that  $(f + g)(X)$  is finite.. This proves that  $f + g$  is a simple function.  $\square$

**Example 19.161.** Let  $\langle X, \mathcal{A} \rangle$  be a measurable space,  $A \in \mathcal{A}$  and  $f: A \rightarrow \bar{\mathbb{R}}$  a simple function then  $f$  is  $\mathcal{A}$ -measurable if  $\forall y \in f(X)$  we have that  $\{f = y\} = f^{-1}(\{y\}) \in \mathcal{A}$ .

**Proof.** As  $f$  is a simple function we have that  $f(A)$  is finite. Let  $B \in \mathcal{B}[\bar{\mathbb{R}}]$  then  $f^{-1}(B \cap f(A)) \stackrel{2.62}{=} f^{-1}(B) \cap f^{-1}(f(A)) = f^{-1}(B) \cap A = f^{-1}(B)$  proving that

$$f^{-1}(B \cap f(A)) = f^{-1}(B) \quad (19.312)$$

As  $f(A)$  is finite we have that  $B \cap f(A)$  is finite, so we have either

$$B \cap f(A) = \emptyset. \text{ then } f^{-1}(B) \stackrel{19.312}{=} f^{-1}(B \cap f(A)) = f^{-1}(\emptyset) = \emptyset \in \mathcal{A}$$

$$B \cap f(A) \neq \emptyset. \text{ then}$$

$$\begin{aligned} f^{-1}(B) &\stackrel{19.312}{=} f^{-1}(B \cap f(A)) \\ &= f^{-1}\left(\bigcup_{y \in B \cap f(A)} \{y\}\right) \\ &= \bigcup_{y \in B \cap f(A)} f^{-1}(\{y\}) \end{aligned}$$

which as  $f^{-1}(\{y\}) \in \mathcal{A}$  and  $B \cap f(A)$  is finite proves by 19.10 that  $f^{-1}(B) \in \mathcal{A}$

So in all cases we have that  $f^{-1}(B) \in \mathcal{A}$  proving that  $f$  is  $\mathcal{A}$ -measurable.  $\square$

### 19.5.3 Properties of measurable functions

**Proposition 19.162.** *Let  $\langle X, \mathcal{A} \rangle$  be a measurable space,  $A \in \mathcal{A}$  and let  $f: X \rightarrow \overline{\mathbb{R}}$  and  $g: X \rightarrow \overline{\mathbb{R}}$  be  $\mathcal{A}$ -measurable functions then*

1.  $\{f < g\} \in \mathcal{A}$
2.  $\{f \leq g\} \in \mathcal{A}$
3.  $\{f = g\} \in \mathcal{A}$
4.  $\{f \neq g\} \in \mathcal{A}$
5. If  $x \in \overline{\mathbb{R}}$  then  $\{f = x\} \in \mathcal{A}$

**Proof.**

1. Let  $t \in \{f < g\}$  then  $f(t) < g(t)$ . Using the density of the rational numbers in the extended real numbers [see 17.9] there exists a  $q \in \mathbb{Q}$  such that  $f(t) < q < g(t)$  or  $t \in \{f < q\} \wedge t \in \{g > q\}$ . So we have

$$\{f < g\} \subseteq \bigcup_{q \in \mathbb{Q}} (\{f < q\} \cap \{g > q\}) \quad (19.313)$$

If  $t \in \bigcup_{q \in \mathbb{Q}} (\{f < q\} \cap \{g > q\})$  then  $\exists q \in \mathbb{Q}$  such that  $t \in \{f < q\} \cap \{g > q\} \Rightarrow f(t) < q < g(t) \Rightarrow f(t) < g(t) \Rightarrow t \in \{f < g\}$ . So we have  $\bigcup_{q \in \mathbb{Q}} (\{f < q\} \cap \{g > q\}) \subseteq \{f < g\}$ . Combining this with 19.313 we have

$$\{f < g\} = \bigcup_{q \in \mathbb{Q}} (\{f < q\} \cap \{g > q\}) \quad (19.314)$$

As  $\forall q \in \mathbb{Q}$  we have by 19.151 that  $\{f < q\}, \{g > q\} \in \mathcal{A} \Rightarrow \{f < q\} \cap \{g > q\} \in \mathcal{A}$  and  $\mathbb{Q}$  is denumerable [see 7.40] we have by 19.14 that  $\bigcup_{q \in \mathbb{Q}} (\{f < q\} \cap \{g > q\}) \in \mathcal{A}$ . Hence by 19.314 we conclude that

$$\{f < g\} \in \mathcal{A}$$

2. We have

$$\begin{aligned} t \in (X \setminus \{g < f\}) &\Leftrightarrow t \in X \wedge \neg(g(t) < f(t)) \\ &\Leftrightarrow t \in X \wedge f(t) \leq g(t) \\ &\Leftrightarrow t \in \{f \leq g\} \end{aligned}$$

proving that  $X \setminus \{g < f\} = \{f \leq g\}$ . Using (1) we have  $\{g < f\} \in \mathcal{A}$  hence  $X \setminus \{g < f\} \in \mathcal{A}$  which proves that

$$\{f \leq g\} \in \mathcal{A}$$

3. We have

$$\begin{aligned}
 t \in (\{f \leq g\} \setminus \{f < g\}) &\Leftrightarrow t \in \{f \leq g\} \wedge t \notin \{f < g\} \\
 &\Leftrightarrow t \in X \wedge f(t) \leq g(t) \wedge \neg(f(t) < g(t)) \\
 &\Leftrightarrow t \in X \wedge f(t) \leq g(t) \wedge g(t) \leq f(t) \\
 &\Leftrightarrow t \in X \wedge f(t) = g(t) \\
 &\Leftrightarrow t \in \{f = g\}
 \end{aligned}$$

proving that  $\{f = g\} = \{f \leq g\} \setminus \{f > g\} \in \mathcal{A}$  [using (1), (2) and 19.10].

4. We have

$$\begin{aligned}
 t \in X \setminus \{f = g\} &\Leftrightarrow t \in X \wedge \neg(f(t) = g(t)) \\
 &\Leftrightarrow t \in X \wedge f(t) \neq g(t) \\
 &\Leftrightarrow t \in \{f \neq g\}
 \end{aligned}$$

proving that  $\{f \neq g\} = X \setminus \{f = g\} \in \mathcal{A}$  [using (3)]

5. Let  $x \in \overline{\mathbb{R}}$  then as  $\{x\}$  is closed [see 17.40], we have by 19.30 that  $\{x\} \in \mathcal{B}[\overline{\mathbb{R}}]$  which as  $f$  is  $\mathcal{A}$ -measurable proves that  $\{f = x\} = f^{-1}(\{x\}) \in \mathcal{A}$ .  $\square$

**Definition 19.163.** Let  $X$  be a set,  $k \in \mathbb{N}_0$  and  $\{f_n\}_{n \in \{k, \dots, \infty\}} \subseteq \overline{\mathbb{R}}$  a sequence of function graphs from functions from  $X$  to  $\overline{\mathbb{R}}$  then we define

1.  $\sup_{n \in \{k, \dots, \infty\}} f_n: X \rightarrow \overline{\mathbb{R}}$  where  $(\sup_n f_n)(x) = \sup(\{f_n(x) | n \in \{k, \dots, \infty\}\})$
2.  $\inf_{n \in \{k, \dots, \infty\}} f_n: X \rightarrow \overline{\mathbb{R}}$  where  $(\inf_n f_n)(x) = \inf(\{f_n(x) | n \in \{k, \dots, \infty\}\})$
3.  $\limsup_{n \rightarrow \infty} f_n: X \rightarrow \overline{\mathbb{R}}$  where  $(\limsup_{n \rightarrow \infty} f_n)(x) = \limsup_{n \rightarrow \infty} f_n(x)$
4.  $\liminf_{n \rightarrow \infty} f_n: X \rightarrow \overline{\mathbb{R}}$  where  $(\liminf_{n \rightarrow \infty} f_n)(x) = \liminf_{n \rightarrow \infty} f_n(x)$
5. If  $\forall x \in X \lim_{n \rightarrow \infty} f_n(x)$  exists then we define  $\lim_{n \rightarrow \infty} f_n: X \rightarrow \overline{\mathbb{R}}$  by  $(\lim_{n \rightarrow \infty} f_n)(x) = \lim_{n \rightarrow \infty} f_n(x)$

**Lemma 19.164.** Let  $X$  be a set,  $k \in \mathbb{N}_0$  and  $\{f_n\}_{n \in \{k, \dots, \infty\}} \subseteq \overline{\mathbb{R}}^X$  a sequence of function graphs of functions from  $X$  to  $\overline{\mathbb{R}}$  then

1.  $\limsup_{n \rightarrow \infty} f_n = \inf_{n \in \{k, \dots, \infty\}} g_n$  where  $g_n = \sup_{l \in \{n, \dots, \infty\}} f_l$
2.  $\liminf_{n \rightarrow \infty} f_n = \sup_{n \in \{k, \dots, \infty\}} h_n$  where  $h_n = \inf_{l \in \{n, \dots, \infty\}} f_l$

**Proof.**

1. Let  $x \in X$ . If  $n \in \mathbb{N}_0$  then  $g_n(x) = \sup(\{f_l(x) | l \in \{n, \dots, \infty\}\})$ . So

$$\begin{aligned}
 (\limsup_{n \rightarrow \infty} f_n)(x) &\stackrel{\text{def}}{=} \inf(\{\sup(\{f_l(x) | l \in \{n, \dots, \infty\}\}) | n \in \{k, \dots, \infty\}\}) \\
 &= \inf(\{g_n(x) | n \in \{k, \dots, \infty\}\}) \\
 &= (\inf_{n \in \{k, \dots, n\}} g_n)(x)
 \end{aligned}$$

proving that

$$\limsup_{n \rightarrow \infty} f_n = \inf_{n \in \{k, \dots, \infty\}} g_n$$

2. Let  $x \in X$ . If  $n \in \mathbb{N}_0$  then  $h_n(x) = \inf(\{f_l(x) | l \in \{n, \dots, \infty\}\})$ . So

$$\begin{aligned}
 (\liminf_{n \rightarrow \infty} f_n)(x) &\stackrel{\text{def}}{=} \sup(\{\inf(\{f_l(x) | l \in \{n, \dots, \infty\}\}) | n \in \{k, \dots, \infty\}\}) \\
 &= \inf(\{h_n(x) | n \in \{k, \dots, \infty\}\}) \\
 &= (\inf_{n \in \{k, \dots, n\}} h_n)(x)
 \end{aligned}$$

proving that

$$\liminf_{n \rightarrow \infty} f_n = \sup_{n \in \{k, \dots, \infty\}} g_n$$

□

**Proposition 19.165.** Let  $\langle X, \mathcal{A} \rangle$  be a measurable space,  $A \in \mathcal{A}$  and  $f: A \rightarrow \overline{\mathbb{R}}$ ,  $g: A \rightarrow \overline{\mathbb{R}}$  be  $\mathcal{A}$ -measurable functions then

1.  $f \vee g: A \rightarrow \overline{\mathbb{R}}$  defined by  $(f \vee g)(x) = \max(f(x), g(x))$  is  $\mathcal{A}$ -measurable.
2.  $f \wedge g: A \rightarrow \overline{\mathbb{R}}$  defined by  $(f \wedge g)(x) = \min(f(x), g(x))$  is  $\mathcal{A}$ -measurable.

**Proof.**

1. We have Let  $x \in \mathbb{R}$  then

$$\begin{aligned} t \in \{(f \vee g) < x\} &\Leftrightarrow \max(f(t), g(t)) < x \\ &\Leftrightarrow f(t) < x \wedge g(t) < x \\ &\Leftrightarrow t \in \{f < x\} \wedge t \in \{g < x\} \\ &\Leftrightarrow t \in (\{f < x\} \cap \{g < x\}) \end{aligned}$$

proving that  $\{(f \vee g) < x\} = \{f < x\} \cap \{g < x\}$ . As by 19.151  $\{f < x\}, \{g < x\} \in \mathcal{A}$  it follows that  $\{f < x\} \cap \{g < x\} \in \mathcal{A}$ . Using 19.151 again proves that  $f \vee g$  is  $\mathcal{A}$ -measurable.

2. We have Let  $x \in X$  then

$$\begin{aligned} t \in \{(f \wedge g) > x\} &\Leftrightarrow \min(f(t), g(t)) > x \\ &\Leftrightarrow f(t) > x \wedge g(t) > x \\ &\Leftrightarrow t \in \{f > x\} \wedge t \in \{g > x\} \\ &\Leftrightarrow t \in (\{f > x\} \cap \{g > x\}) \end{aligned}$$

proving that  $\{(f \wedge g) > x\} = \{f > x\} \cap \{g > x\}$ . As by 19.151  $\{f > x\}, \{g > x\} \in \mathcal{A}$  it follows that  $\{f > x\} \cap \{g > x\} \in \mathcal{A}$ . Using 19.151 again proves that  $f \wedge g$  is  $\mathcal{A}$ -measurable. □

**Definition 19.166.** Let  $X$  be a set,  $n \in \mathbb{N}_0$ ,  $\emptyset \neq I \subseteq \mathbb{N}_0$  finite,  $\{f_i\}_{i \in \{1, \dots, n\}} \subseteq \overline{\mathbb{R}}^X$  then  $\max(\{f_i | i \in I\})$  is defined by  $\max(\{f_i | i \in I\})(x) = \max(\{f_i(x) | i \in I\})$

**Lemma 19.167.** Let  $X$  be a set,  $n \in \{2, \dots, \infty\}$  then if  $\{f_i\}_{i \in \{1, \dots, n\}} \subseteq \overline{\mathbb{R}}^X$  we have

1.  $\forall k \in \{1, \dots, n\}$   $\max(\{f_i | i \in \{1, \dots, n\}\}) = \max(\max(\{f_i | i \in \{1, \dots, n\} \setminus \{k\}\}), f_k)$
2. If  $\{f_i\}_{i \in \{1, \dots, n\}} \subseteq \mathbb{R}^X$  then  $\max(\{f_i | i \in \{1, \dots, n\}\}) \in \mathbb{R}^X$

**Proof.**

1. Let  $x \in A$  then if  $m = \max(\{f_i | i \in \{1, \dots, n\}\})(x) = \max(\{f_i(x) | i \in \{1, \dots, n\}\})$  we have that  $m \in \{\{f_i(x) | i \in \{1, \dots, n\}\}\}$  and  $\forall i \in \{1, \dots, n\}$   $f_i(x) \leq m$ . So  $\exists i \in \{1, \dots, n\}$  such that  $m = f_i(x)$  and we have two cases to consider:

$i \in \{1, \dots, n\} \setminus \{k\}$ . then as  $\forall i \in \{1, \dots, n\}$   $f_i(x) \leq m$  we have  $m = \max(\{f_i(x) | i \in \{1, \dots, n\} \setminus \{k\}\})$  and  $f_k(x) \leq m$  so that  $m = \max(\max(\{f_i(x) | i \in \{1, \dots, n\} \setminus \{k\}\}), f_k(x))$ . proving that

$$\max(\{f_i | i \in \{1, \dots, n\}\})(x) = \max(\max(\{f_i(x) | i \in \{1, \dots, n\}\}), f_k(x))$$

$i = k$ . then as  $\forall i \in \{1, \dots, n\}$   $f_i(x) \leq m = f_i(x) = f_k(x)$  we have  $\max(\{f_i(x) | i \in \{1, \dots, n\} \setminus \{k\}\}) \leq m$  so that  $m = \max(\max(\{f_i(x) | i \in \{1, \dots, n\} \setminus \{k\}\}), f_k(x))$ , proving that

$$\max(\{f_i | i \in \{1, \dots, n\}\})(x) = \max(\max(\{f_i(x) | i \in \{1, \dots, n\}\}), f_k(x))$$

Hence

$$\max(\{f_i | i \in \{1, \dots, n\}\}) = \max(\max(\{f_i | i \in \{1, \dots, n\} \setminus \{k\}\}), f_k)$$

2. Let  $m = \max(\{f_i | i \in \{1, \dots, n\}\})(x) = \max(\{f_i(x) | i \in \{1, \dots, n\}\})$  then  $m \in \{f_i(x) | i \in \{1, \dots, n\}\}$  such that  $\exists i \in \{1, \dots, n\}$  such that  $m = f_i(x) \in \mathbb{R}$ . □

**Corollary 19.168.** Let  $\langle X, \mathcal{A} \rangle$  be a measurable space,  $A \in \mathcal{A}$ ,  $n \in \mathbb{N}_0$  and  $\{f_i\}_{i \in \{1, \dots, n\}} \subseteq \overline{\mathcal{M}[A, \mathcal{A}]}$  then  $\max(\{f_i | i \in \{1, \dots, n\}\}) \in \overline{\mathcal{M}[A, \mathcal{A}]}$ . Further if  $\{f_i\}_{i \in \{1, \dots, n\}} \subseteq \mathcal{M}[A, \mathcal{A}]$  then  $\max(\{f_i | i \in \{1, \dots, n\}\}) \in \mathcal{M}[A, \mathcal{A}]$ .

**Proof.** The first part is proved by induction, let

$$\mathcal{S} = \{n \in \mathbb{N}_0 | \text{If } \{f_i\}_{i \in \{1, \dots, n\}} \subseteq \overline{\mathcal{M}[A, \mathcal{A}]} \text{ then } \max(\{f_i | i \in \{1, \dots, n\}\}) \in \overline{\mathcal{M}[A, \mathcal{A}]}\}$$

then we have:

**1**  $\in \mathcal{S}$ . If  $\{f_i\}_{i \in \{1, \dots, 1\}} \subseteq \overline{\mathcal{M}[A, \mathcal{A}]}$  then  $\forall x \in A$  we have

$$\max(\{f_i | i \in \{1, \dots, n\}\})(x) = \max(\{f_i(x) | i \in \{1, \dots, 1\}\}) = \max(\{f_1(x)\}) = f_1(x)$$

proving that  $\max(\{f_i | i \in \{1, \dots, n\}\}) = f_1 \in \overline{\mathcal{M}[A, \mathcal{A}]}$  so  $1 \in \mathcal{S}$

**n**  $\in \mathcal{S} \Rightarrow n + 1 \in \mathcal{S}$ . Let  $\{f_i\}_{i \in \{1, \dots, n+1\}} \subseteq \overline{\mathcal{M}[A, \mathcal{A}]}$ . As  $n \in \mathcal{S}$  we have that  $\max(\{f_i | i \in \{1, \dots, n\}\}) \in \overline{\mathcal{M}[A, \mathcal{A}]}$  which, as  $f_{n+1} \in \overline{\mathcal{M}[A, \mathcal{A}]}$ , gives using 19.165 that

$$\max(\max(\{f_i | i \in \{1, \dots, n\}\}), f_{n+1}) \in \overline{\mathcal{M}[A, \mathcal{A}]}$$

finally using the previous lemma 19.167 we have that

$$\max(\{f_i | i \in \{1, \dots, n+1\}\}) = \max(\max(\{f_i | i \in \{1, \dots, n\}\}), f_{n+1})$$

proving that  $n + 1 \in \mathcal{S}$

Mathematical induction proves then that  $\forall n \in \mathbb{N}_0$  we have  $\max(\{f_i | i \in \{1, \dots, n\}\}) \in \overline{\mathcal{M}[A, \mathcal{A}]}$ .

For the final part if  $\{f_i\}_{i \in \{1, \dots, n\}} \subseteq \mathcal{M}[A, \mathcal{A}] \subseteq \overline{\mathcal{M}[A, \mathcal{A}]}$  then  $\forall i \in \{1, \dots, n\}$   $f_i \in \mathbb{R}^X$ , so that by 19.167  $\max(\{f_i | i \in \{1, \dots, n\}\}) \in \mathbb{R}^X$  which combined with the first part proves that  $\max(\{f_i | i \in \{1, \dots, n\}\}) \in \mathcal{M}[A, \mathcal{A}]$ .  $\square$

**Proposition 19.169.** Let  $\langle X, \mathcal{A} \rangle$  be a measurable space,  $k \in \mathbb{N}_0$ ,  $A \in \mathcal{A}$  and  $\{f_n\}_{n \in \{k, \dots, \infty\}} \subseteq \overline{\mathcal{M}[A, \mathcal{A}]}$  a sequence of functions that are  $\mathcal{A}$ -measurable then we have:

1.  $\sup_{n \in \{k, \dots, \infty\}} f_n: A \rightarrow \overline{\mathbb{R}}$  is  $\mathcal{A}$ -measurable [in other words  $\sup_{n \in \{k, \dots, \infty\}} f_n \in \overline{\mathcal{M}[A, \mathcal{A}]}$ ]
2.  $\inf_{n \in \{k, \dots, \infty\}} f_n: A \rightarrow \overline{\mathbb{R}}$  is  $\mathcal{A}$ -measurable [in other words  $\sup_{n \in \{k, \dots, \infty\}} f_n \in \overline{\mathcal{M}[A, \mathcal{A}]}$ ].
3.  $\limsup_{n \rightarrow \infty} f_n: A \rightarrow \overline{\mathbb{R}}$  is  $\mathcal{A}$ -measurable [in other words  $\sup_{n \in \{k, \dots, \infty\}} f_n \in \overline{\mathcal{M}[A, \mathcal{A}]}$ ].
4.  $\liminf_{n \rightarrow \infty} f_n: A \rightarrow \overline{\mathbb{R}}$  is  $\mathcal{A}$ -measurable [in other words  $\sup_{n \in \{k, \dots, \infty\}} f_n \in \overline{\mathcal{M}[A, \mathcal{A}]}$ ].
5. Let  $D = \left\{ x \in A \mid \left( \limsup_{n \rightarrow \infty} f_n \right)(x) = \left( \limsup_{n \rightarrow \infty} f_n \right)(x) \right\}$  [this is the set where by 17.63  $\lim_{n \rightarrow \infty} f_n$  is defined] then  $D \in \mathcal{A}$  and  $\lim_{n \rightarrow \infty} f_n: D \rightarrow \overline{\mathbb{R}}$  is  $\mathcal{A}$ -measurable [in other words  $\lim_{n \rightarrow \infty} f_n \in \overline{\mathcal{M}[D, \mathcal{A}]}$ ]

**Proof.**

1. Let  $x \in \mathbb{R}$ . If  $t \in \left\{ \left( \sup_{n \in \{k, \dots, \infty\}} f_n \right) \leq x \right\}$  then  $\sup(\{f_n(t) | n \in \{k, \dots, \infty\}\}) \leq x$ . So  $\forall n \in \mathbb{N}_0$  we have  $f_n(t) \leq x \Rightarrow t \in \{f_n \leq x\}$ . Hence

$$\left\{ \left( \sup_{n \in \{k, \dots, \infty\}} f_n \right) \leq x \right\} \subseteq \bigcap_{n \in \{k, \dots, \infty\}} \{f_n \leq x\} \quad (19.315)$$

If  $t \in \bigcap_{n \in \{k, \dots, \infty\}} \{f_n \leq x\}$  we have  $\forall n \in \{k, \dots, \infty\}$  that  $t \in \{f_n \leq x\} \Rightarrow f_n(t) \leq x$ . Assume that  $x < \sup(\{f_n(t) | n \in \{k, \dots, \infty\}\})$  then there exists a  $n \in \{k, \dots, \infty\}$  such that  $x < f_n(t)$  leading to the contradiction  $x < x$ . Hence  $\sup(\{f_n(t) | n \in \{k, \dots, \infty\}\}) \leq x$  or  $t \in \left\{ \left( \sup_{n \in \{k, \dots, \infty\}} f_n \right) \leq x \right\}$  proving  $\bigcap_{n \in \{k, \dots, \infty\}} \{f_n \leq x\} \subseteq \left\{ \left( \sup_{n \in \{k, \dots, \infty\}} f_n \right) \leq x \right\}$ . Combining this with 19.315 gives

$$\left\{ \left( \sup_{n \in \{k, \dots, \infty\}} f_n \right) \leq x \right\} = \bigcap_{n \in \{k, \dots, \infty\}} \{f_n \leq x\} \quad (19.316)$$

As  $\forall n \in \{k, \dots, n\}$   $f_n$  is  $\mathcal{A}$ -measurable, hence  $\{f_n \leq x\} \in \mathcal{A}$ , applying 19.14 proves  $\bigcap_{n \in \{k, \dots, \infty\}} \{f_n \leq x\} \in \mathcal{A}$  or using 19.316 that  $\left\{ \left( \sup_{n \in \{k, \dots, \infty\}} f_n \right) \leq x \right\} \in \mathcal{A}$ . Using 19.151 it follows that

$$\sup_{n \in \{k, \dots, \infty\}} f_n \text{ is } \mathcal{A}\text{-measurable}$$

2. Let  $x \in \mathbb{R}$ . If  $t \in \left\{ \left( \inf_{n \in \{k, \dots, \infty\}} f_n \right) < x \right\}$  then  $\inf(\{f_n(t) | n \in \{k, \dots, n\}\}) < x$ . So  $\exists n \in \{k, \dots, \infty\}$  such that  $f_n(t) < x \Rightarrow t \in \{f_n < x\}$ . Hence

$$\left\{ \left( \inf_{n \in \{k, \dots, \infty\}} f_n \right) < x \right\} \subseteq \bigcup_{n \in \{k, \dots, \infty\}} \{f_n < x\} \quad (19.317)$$

If  $t \in \bigcup_{n \in \{k, \dots, \infty\}} \{f_n < x\}$  then  $\exists n \in \{k, \dots, \infty\}$  such that  $t \in \{f_n < x\} \Rightarrow f_n(t) < x$ , hence  $\inf(\{f_n(t) | n \in \{k, \dots, n\}\}) \leq f_n(t) < x$ , proving that  $t \in \left\{ \left( \inf_{n \in \{k, \dots, \infty\}} f_n \right) < x \right\}$ . So  $\bigcup_{n \in \{k, \dots, \infty\}} \{f_n < x\} \subseteq \left\{ \left( \inf_{n \in \{k, \dots, \infty\}} f_n \right) < x \right\}$  which by 19.317 gives

$$\left\{ \left( \inf_{n \in \{k, \dots, \infty\}} f_n \right) < x \right\} = \bigcup_{n \in \{k, \dots, \infty\}} \{f_n < x\} \quad (19.318)$$

As  $\forall n \in \{k, \dots, n\}$   $f_n$  is  $\mathcal{A}$ -measurable, hence  $\{f_n < x\} \in \mathcal{A}$ , applying then 19.14 proves  $\bigcup_{n \in \{k, \dots, \infty\}} \{f_n < x\} \in \mathcal{A}$  or using 19.318 that  $\left\{ \left( \inf_{n \in \{k, \dots, \infty\}} f_n \right) \leq x \right\} \in \mathcal{A}$ . Using 19.151 it follows that

$$\inf_{n \in \{k, \dots, \infty\}} f_n \text{ is } \mathcal{A}\text{-measurable}$$

3.  $\forall n \in \{k, \dots, \infty\}$  we have by (1) that  $g_n = \sup_{l \in \{n, \dots, \infty\}} f_l$  is  $\mathcal{A}$ -measurable. Using (2) we have then that  $\inf_{n \in \{k, \dots, \infty\}} g_n$  is  $\mathcal{A}$ -measurable. As  $\limsup_{n \rightarrow \infty} f_n \stackrel{19.164}{=} \inf_{n \in \{k, \dots, \infty\}} g_n$  we have proved that

$$\limsup_{n \rightarrow \infty} f_n \text{ is } \mathcal{A}\text{-measurable}$$

4.  $\forall n \in \{k, \dots, \infty\}$  we have by (2) that  $h_n = \inf_{l \in \{n, \dots, \infty\}} f_l$  is  $\mathcal{A}$ -measurable. Using (1) we have then that  $\sup_{n \in \{k, \dots, \infty\}} h_n$  is  $\mathcal{A}$ -measurable. As  $\liminf_{n \rightarrow \infty} f_n \stackrel{19.164}{=} \sup_{n \in \{k, \dots, \infty\}} h_n$  we have proved that

$$\liminf_{n \rightarrow \infty} f_n \text{ is } \mathcal{A}\text{-measurable}$$

5. From (3),(4) it follows that  $\limsup_{n \rightarrow \infty} f_n$  and  $\liminf_{n \rightarrow \infty} f_n$  are  $\mathcal{A}$ -measurable, combining this with 19.162, proves that

$$D = \left\{ \limsup_{n \rightarrow \infty} f_n = \limsup_{n \rightarrow \infty} f_n \right\} \in \mathcal{A} \quad (19.319)$$

Let  $x \in \overline{\mathbb{R}}$  then

$$\begin{aligned} t \in D \cap \left\{ \left( \limsup_{n \rightarrow \infty} f_n \right) \leq x \right\} &\Leftrightarrow t \in D \wedge \left( \limsup_{n \rightarrow \infty} f_n \right)(t) \leq x \\ t \in D \Rightarrow \lim_{n \rightarrow \infty} f_m(t) &\stackrel{\Leftrightarrow}{=} \limsup_{n \rightarrow \infty} f_n(t) \quad t \in D \wedge \left( \lim_{n \rightarrow \infty} f_n \right)(t) \leq x \\ &\Leftrightarrow t \in \left\{ \left( \lim_{n \rightarrow \infty} f_n \right) \leq x \right\} \end{aligned}$$

which proves that  $\left\{ \left( \lim_{n \rightarrow \infty} f_n \right) \leq x \right\} = D \cap \left\{ \left( \limsup_{n \rightarrow \infty} f_n \right) \leq x \right\} \in \mathcal{A}$  [as  $D \in \mathcal{A}$  and  $\left\{ \left( \limsup_{n \rightarrow \infty} f_n \right) \leq x \right\} \in \mathcal{A}$  [see 19.151]]. Finally using 19.151 it follows that

$$\lim_{n \rightarrow \infty} f_n: D \rightarrow \overline{\mathbb{R}} \text{ is } \mathcal{A}\text{-measurable}$$

□

**Proposition 19.170.** Let  $\langle X, \mathcal{A} \rangle$  be a measurable space,  $A \in \mathcal{A}$ ,  $\alpha \in \mathbb{R}$  and  $f: A \rightarrow \overline{\mathbb{R}}$ ,  $g: A \rightarrow \overline{\mathbb{R}}$   $\mathcal{A}$ -measurable functions then

1. If  $f(A) \subseteq [0, \infty]$ ,  $g(A) \subseteq [0, \infty]$  then  $f + g: A \rightarrow \overline{\mathbb{R}}$  defined by  $(f + g)(x) = f(x) + g(x)$  is  $\mathcal{A}$ -measurable
2.  $\alpha \cdot f: A \rightarrow \overline{\mathbb{R}}$  defined by  $(\alpha \cdot f)(x) = \alpha \cdot f(x)$  is  $\mathcal{A}$ -measurable

**Proof.**

1. Let  $x \in \mathbb{R}$ . If  $t \in \{f + g < x\}$  then  $f(t) + g(t) < x \underset{0 \leq f(t) + g(t) < x \Rightarrow g(t) \in \mathbb{R}}{\Rightarrow} f(t) < x - g(t)$  and using 17.9 there exists a  $q \in \mathbb{Q}$  such that  $f(t) < q < x - g(t) \Rightarrow f(t) < q \wedge g(t) < x - q \Rightarrow t \in \{f < q\} \cap \{g < x - q\}$ . Hence

$$\{f + g < x\} \subseteq \bigcup_{q \in \mathbb{Q}} (\{f < q\} \cap \{g < x - q\}) \quad (19.320)$$

Further if  $t \in \bigcup_{q \in \mathbb{Q}} (\{f < q\} \cap \{g < x - q\})$  there exists a  $q \in \mathbb{Q}$  such that  $t \in \{f < q\} \wedge t \in \{g < x - q\}$  or  $f(t) < q \wedge g(t) < x - q \Rightarrow f(t) + g(t) < q + x - q = x$ , hence  $t \in \{f + g < x\}$ . So  $\bigcup_{q \in \mathbb{Q}} (\{f < q\} \cap \{g < x - q\}) \subseteq \{f + g < x\}$  which using 19.320 proves that

$$\{f + g < x\} = \bigcup_{q \in \mathbb{Q}} (\{f < q\} \cap \{g < x - q\}) \quad (19.321)$$

As  $f, g$  are  $\mathcal{A}$ -measurable we have that  $\{f < q\} \cap \{g < x - q\} \in \mathcal{A}$ , further, as  $\mathbb{Q}$  is denumerable [see 7.40], we can use 19.14 to conclude that  $\bigcup_{q \in \mathbb{Q}} (\{f < q\} \cap \{g < x - q\}) \in \mathcal{A}$ . Hence using 19.151 it follows that

$$f + g: A \rightarrow \overline{\mathbb{R}} \text{ is } \mathcal{A}\text{-measurable}$$

2. We have two cases to consider for  $\alpha$ :

**$\alpha = 0$ .** then  $\forall x \in A$   $(\alpha \cdot f)(A) = \{0\}$  proving by 19.155 that  $\alpha \cdot f$  is  $\mathcal{A}$ -measurable.

**$0 < \alpha$ .** Let  $x \in \mathbb{R}$  then

$$\begin{aligned} t \in \{\alpha \cdot f < x\} &\Leftrightarrow \alpha \cdot f(t) < x \\ &\Leftrightarrow f(t) < \frac{x}{\alpha} \\ &\Leftrightarrow t \in \left\{ f < \frac{x}{\alpha} \right\} \end{aligned}$$

proving that  $\{\alpha \cdot f < x\} = \left\{ f < \frac{x}{\alpha} \right\} \in \mathcal{A}$  [as  $f$  is  $\mathcal{A}$ -measurable]. Using 19.151 we have then that  $\alpha \cdot f$  is  $\mathcal{A}$ -measurable.

**$\alpha < 0$ .** Let  $x \in \mathbb{R}$  then

$$\begin{aligned} t \in \{\alpha \cdot f < x\} &\Leftrightarrow \alpha \cdot f(t) < x \\ &\Leftrightarrow f(t) > \frac{x}{\alpha} \\ &\Leftrightarrow t \in \left\{ f > \frac{x}{\alpha} \right\} \end{aligned}$$

proving that  $\{\alpha \cdot f < x\} = \left\{ f > \frac{x}{\alpha} \right\} \in \mathcal{A}$  [as  $f$  is  $\mathcal{A}$ -measurable]. Using 19.151 we have then that  $\alpha \cdot f$  is  $\mathcal{A}$ -measurable.  $\square$

Using mathematical induction we have that the finite sum of measurable functions is measurable.

**Corollary 19.171.** Let  $\langle X, \mathcal{A} \rangle$  be a measurable space,  $n \in \mathbb{N}_0$  and  $\{f_i\}_{i \in \{1, \dots, n\}} \subseteq \overline{\mathcal{M}[X, \mathcal{A}]}$  a finite set of  $\mathcal{A}$ -measurable functions from  $X$  to  $\overline{\mathbb{R}}$  with  $f_i(X) \subseteq [0, \infty]$  and  $\sum_{i=1}^n f_i: X \rightarrow \overline{\mathbb{R}}$  defined by  $(\sum_{i=1}^n f_i)(x) = \sum_{i=1}^n f_i(x)$  then  $\sum_{i=1}^n f_i$  is  $\mathcal{A}$ -measurable.

**Proof.** This is trivially proved by induction, so define  $\mathcal{S} = \{i \in \{1, \dots, n\} \mid \text{if } i \leq n \text{ then } \sum_{j=1}^i f_j \text{ is } \mathcal{A}\text{-measurable}\}$  then we have:

**$1 \in \mathcal{S}$ .** This follows from the fact that  $\sum_{j=1}^1 f_j = f_1$  and that  $f_1$  is  $\mathcal{A}$ -measurable.

$i \in \mathcal{S} \Rightarrow i+1 \in \mathcal{S}$ . Let  $i \in \mathcal{S}$  then if  $i+1 \leq n$  we have  $(\sum_{j=1}^{i+1} f_j)(x) = \sum_{j=1}^{i+1} f_j(x) = f_{i+1}(x) + \sum_{j=1}^i f_j(x)$ . So  $\sum_{j=1}^{i+1} f_j = f_{i+1} + \sum_{j=1}^i f_j$  which, as  $i \in \mathcal{S}$  is the sum of two  $\mathcal{A}$ -measurable functions, is  $\mathcal{A}$ -measurable [using 19.170]. Hence  $i+1 \in \mathcal{S}$   $\square$

**Note 19.172.** In the above proposition and corollary we assume that  $f, g$  are non negative. This is essential to make sure that  $f+g$  is well defined. If we restrict ourself to real valued functions we can extend the properties that that measurable functions have.

**Proposition 19.173.** Let  $\langle X, \mathcal{A} \rangle$  be a measurable space,  $\alpha \in \mathbb{R}$ ,  $A \in \mathcal{A}$  and  $f: A \rightarrow \mathbb{R}$ ,  $g: A \rightarrow \mathbb{R}$   $\mathcal{A}$ -measurable functions [meaning  $f, g \in \mathcal{M}[A, \mathcal{A}]$ ] then we have

1.  $f+g: X \rightarrow \mathbb{R}$  defined by  $(f+g)(x) = f(x) + g(x)$  is  $\mathcal{A}$ -measurable [in other words  $f+g \in \mathcal{M}[A, \mathcal{A}]$ ].
2.  $\alpha \cdot f: X \rightarrow \mathbb{R}$  defined by  $(\alpha \cdot f)(x) = \alpha \cdot f(x)$  is a  $\mathcal{A}$ -measurable [in other words  $\alpha \cdot f \in \mathcal{M}[A, \mathcal{A}]$ ].
3.  $f-g: X \rightarrow \mathbb{R}$  defined by  $(f-g)(x) = f(x) - g(x)$  is  $\mathcal{A}$ -measurable [in other words  $f-g \in \mathcal{M}[A, \mathcal{A}]$ ].
4.  $f^2: X \rightarrow \mathbb{R}$  defined by  $f^2(x) = (f(x))^2$  is  $\mathcal{A}$ -measurable [in other words  $f^2 \in \mathcal{M}[A, \mathcal{A}]$ ].
5.  $f \cdot g: X \rightarrow \mathbb{R}$  defined by  $(f \cdot g)(x) = f(x) \cdot g(x)$  is  $\mathcal{A}$ -measurable [in other words  $f \cdot g \in \mathcal{M}[A, \mathcal{A}]$ ].
6.  $D = \{x \in A \mid g(x) \neq 0\} \in \mathcal{A}$  and  $\frac{f}{g}: D \rightarrow \mathbb{R}$  defined by  $\left(\frac{f}{g}\right)(x) = \frac{f(x)}{g(x)}$  is  $\mathcal{A}$ -measurable [in other words  $\frac{f}{g} \in \mathcal{M}[D, \mathcal{A}]$ ].

### Proof.

1. The proof is the same as in the previous proposition. Let  $x \in \mathbb{R}$ . If  $t \in \{f+g < x\}$  then  $f(t) + g(t) < x \Rightarrow f(t) < x - g(t)$  and using 9.58 there exists a  $q \in \mathbb{Q}$  such that  $f(t) < q < x - g(t) \Rightarrow f(t) < q \wedge g(t) < x - q \Rightarrow t \in \{f < q\} \cap \{g < x - q\}$ . Hence

$$\{f+g < x\} \subseteq \bigcup_{q \in \mathbb{Q}} (\{f < q\} \cap \{g < x - q\}) \quad (19.322)$$

Further if  $t \in \bigcup_{q \in \mathbb{Q}} (\{f < q\} \cap \{g < x - q\})$  there exists a  $q \in \mathbb{Q}$  such that  $t \in \{f < q\} \wedge t \in \{g < x - q\}$  or  $f(t) < q \wedge g(t) < x - q \Rightarrow f(t) + g(t) < q + x - q = x$ , hence  $t \in \{f+g < x\}$ . So  $\bigcup_{q \in \mathbb{Q}} (\{f < q\} \cap \{g < x - q\}) \subseteq \{f+g < x\}$  which using 19.322 proves that

$$\{f+g < x\} = \bigcup_{q \in \mathbb{Q}} (\{f < q\} \cap \{g < x - q\}) \quad (19.323)$$

As  $f, g$  are  $\mathcal{A}$ -measurable we have that  $\{f < q\} \cap \{g < x - q\} \in \mathcal{A}$ , further, as  $\mathbb{Q}$  is denumerable [see 7.40], we can use 19.14 to conclude that  $\bigcup_{q \in \mathbb{Q}} (\{f < q\} \cap \{g < x - q\}) \in \mathcal{A}$ . Hence using 19.151 it follows that

$$f+g: A \rightarrow \bar{\mathbb{R}} \text{ is } \mathcal{A}\text{-measurable}$$

2. We have to consider the following cases for  $\alpha \in \mathbb{R}$ :

**$\alpha = 0$ .** Then  $\forall x \in A$  we have  $(\alpha \cdot f)(x) = 0$  which by 19.155 means that  $\alpha \cdot f$  is  $\mathcal{A}$ -measurable.

**$0 < \alpha$ .** Let  $x \in \mathbb{R}$  then  $t \in \{\alpha \cdot f < x\} \Leftrightarrow \alpha \cdot f(t) < x \Leftrightarrow f(t) < \frac{x}{\alpha} \Leftrightarrow t \in \{f < \frac{x}{\alpha}\}$  proving that  $\{\alpha \cdot f < x\} = \{f < \frac{x}{\alpha}\} \in \mathcal{A}$  [as  $f$  is  $\mathcal{A}$ -measurable and 19.151]. Hence using 19.151 it follows that  $\alpha \cdot f$  is  $\mathcal{A}$ -measurable.

**$\alpha > 0$ .** Let  $x \in \mathbb{R}$  then  $t \in \{\alpha \cdot f < x\} \Leftrightarrow \alpha \cdot f(t) < x \Leftrightarrow f(t) > \frac{x}{\alpha} \Leftrightarrow t \in \{f > \frac{x}{\alpha}\}$  proving that  $\{\alpha \cdot f < x\} = \{f > \frac{x}{\alpha}\} \in \mathcal{A}$  [as  $f$  is  $\mathcal{A}$ -measurable and 19.151]. Hence using 19.151 it follows that  $\alpha \cdot f$  is  $\mathcal{A}$ -measurable.

3. Let  $x \in X$  then  $(f - g)(x) = f(x) - g(x) = f(x) + (-1) \cdot g(x) = f(x) + ((-1) \cdot g)(x) = (f + (-1) \cdot g)(x)$  which proves that  $f + (-1) \cdot g$ . Using (2) we have that  $(-1) \cdot g$  is  $\mathcal{A}$ -measurable. So applying (1) proves then that  $f - g = f + (-1 \cdot g)$  is  $\mathcal{A}$ -measurable.

4. Let  $x \in \mathbb{R}$  then we have two cases to consider:

**$x \leq 0$ .** If  $t \in \{f^2 < x\}$  we have  $f^2(t) < x \leq 0$  which, as  $0 \leq f^2(x)$ , is a contradiction. Hence  $\{f^2 < x\} = \emptyset \in \mathcal{A}$

**$0 < x$ .** Then

$$\begin{aligned} t \in \{f^2 < x\} &\Leftrightarrow f^2(t) < x \\ &\stackrel{9.73}{\Leftrightarrow} -\sqrt{x} < f(t) < \sqrt{x} \\ &\Leftrightarrow t \in \{f > -\sqrt{x}\} \wedge t \in \{f < \sqrt{x}\} \\ &\Leftrightarrow t \in \{f > -\sqrt{x}\} \bigcap \{f < \sqrt{x}\} \end{aligned}$$

proving that  $\{f^2 < x\} = \{f > -\sqrt{x}\} \bigcap \{f < \sqrt{x}\}$ . As  $f$  is  $\mathcal{A}$ -measurable it follows from 19.151 that  $\{f > -\sqrt{x}\}, \{f < \sqrt{x}\} \in \mathcal{A}$ , Hence  $\{f^2 < x\} \in \mathcal{A}$ .

As in all the cases  $\{f^2 < x\} \in \mathcal{A}$  it follows using 19.151 again that  $f^2$  is  $\mathcal{A}$ -measurable

5. As  $f, g$  are  $\mathcal{A}$ -measurable we have by (1) that  $f + g$  is  $\mathcal{A}$ -measurable, Applying (4) we have that  $f^2, g^2$  and  $(f + g)^2$  are  $\mathcal{A}$ -measurable. Using (3) twice we have that  $((f + g)^2 - f^2) - g^2$  is  $\mathcal{A}$  measurable, by (2) we have then that

$$\frac{1}{2} \cdot (((f + g)^2 - f^2) - g^2) \text{ is } \mathcal{A}\text{-measurable} \quad (19.324)$$

Finally if  $x \in A$

$$\begin{aligned} \left( \frac{1}{2} \cdot (((f + g)^2 - f^2) - g^2) \right)(x) &= \frac{1}{2} \cdot (((f(x) + g(x))^2 - f^2(x)) - g^2(x)) \\ &= \frac{1}{2} \cdot (f^2(x) + g^2(x) + 2 \cdot f(x) - f^2(x) - g^2(x)) \\ &= f(x) \cdot g(x) \end{aligned}$$

proving that  $f \cdot g = \frac{1}{2} \cdot (((f + g)^2 - f^2) - g^2)$  and by 19.324 that

$f \cdot g$  is  $\mathcal{A}$ -measurable

6. Let  $D = \{x \in A | g(x) \neq 0\}$  then

$$\begin{aligned} t \in \{x \in A | g(x) \neq 0\} &\Leftrightarrow t \in A \wedge g(t) \neq 0 \\ &\Leftrightarrow t \in A \wedge (g(t) < 0 \vee g(t) > 0) \\ &\Leftrightarrow (t \in A \wedge g(t) < 0) \vee (t \in A \wedge g(t) > 0) \\ &\Leftrightarrow t \in \{g < 0\} \vee t \in \{g > 0\} \\ &\Leftrightarrow t \in \{g < 0\} \bigcup \{g > 0\} \end{aligned}$$

proving using the measurability of  $g$  together with 19.151 that

$$D = \{x \in A | g(x) \neq 0\} = \{g < 0\} \bigcup \{g > 0\} \in \mathcal{A} \quad (19.325)$$

If  $t \in (\{t \in A | g(t) > 0\} \bigcap \{t \in A | f(t) < x \cdot g(t)\}) \bigcup (\{t \in A | g(t) < 0\} \bigcap \{t \in A | f(t) > x \cdot g(t)\})$  then we have either:

**$\{t \in A | g(t) > 0\} \bigcap \{t \in A | f(t) < x \cdot g(t)\}$ .** then  $t \in A \wedge g(t) > 0 \wedge f(t) < x \cdot g(t)$ , so that  $t \in A \wedge \frac{f(t)}{g(t)} < x$ , proving as  $g(t) \neq 0$ , that  $t \in \{t \in D | \frac{f(t)}{g(t)} < x\}$ .

**$\{t \in A | g(t) < 0\} \bigcap \{t \in A | f(t) > x \cdot g(t)\}$ .** then  $t \in A \wedge g(t) < 0 \wedge f(t) > x \cdot g(t)$  so that  $t \in A \wedge \frac{f(t)}{g(t)} < x$  proving as  $g(t) \neq 0$ , that  $t \in \{t \in D | \frac{f(t)}{g(t)} < x\}$ .

So we have

$$\left( \{t \in A | g(t) > 0\} \cap \{t \in A | f(t) < x \cdot g(t)\} \right) \cup \left( \{t \in A | g(t) < 0\} \cap \{t \in A | f(t) > x \cdot g(t)\} \right) \subseteq \left\{ t \in D | \frac{f(t)}{g(t)} < x \right\} \quad (19.326)$$

For the opposite inclusion, let  $t \in \left\{ t \in D | \frac{f(t)}{g(t)} < x \right\}$  then  $g(t) \neq 0$  and  $\frac{f(t)}{g(t)} < x$  so we must consider the following cases fro  $g(t) \neq 0$ :

**$g(t) > 0$ .** then  $f(t) < x \cdot g(t)$  or  $t \in \{t \in A | g(t) > 0\} \cap \{t \in A | f(t) < x \cdot g(t)\}$

**$g(t) < 0$ .** then  $f(t) > x \cdot g(t)$  or  $t \in \{t \in A | g(t) < 0\} \cap \{t \in A | f(t) > x \cdot g(t)\}$

proving that  $\left\{ t \in D | \frac{f(t)}{g(t)} < x \right\} \subseteq \left( \{t \in A | g(t) > 0\} \cap \{t \in A | f(t) < x \cdot g(t)\} \right) \cup \left( \{t \in A | g(t) < 0\} \cap \{t \in A | f(t) > x \cdot g(t)\} \right)$ . So combining this with 19.326 proves that

$$\left( \{t \in A | g(t) > 0\} \cap \{t \in A | f(t) < x \cdot g(t)\} \right) \cup \left( \{t \in A | g(t) < 0\} \cap \{t \in A | f(t) > x \cdot g(t)\} \right) = \left\{ t \in D | \frac{f(t)}{g(t)} < x \right\} \quad (19.327)$$

Now as  $g$  is  $\mathcal{A}$ -measurable it follows from 19.151 that  $\{t \in A | g(t) > 0\} \in \mathcal{A}$ ,  $\{t \in A | g(t) < 0\} \in \mathcal{A}$ . As  $f, g$  are  $\mathcal{A}$ -measurable and  $x \cdot g$  is  $\mathcal{A}$ -measurable [using (2)] it follows from 19.162 that  $\{t \in A | f(t) > x \cdot g(t)\}, \{t \in A | f(t) < x \cdot g(t)\} \in \mathcal{A}$  which proves that  $\left\{ t \in D | \frac{f(t)}{g(t)} < x \right\} \in \mathcal{A}$ . Using 19.151 we get finally that

$$\frac{f}{g}: D \rightarrow \mathbb{R} \text{ is } \mathcal{A}\text{-measurable} \quad \square$$

**Corollary 19.174.** Let  $\langle X, \mathcal{A} \rangle$  be a measurable space,  $n \in \mathbb{N}_0$  and  $\{f_i\}_{i \in \{1, \dots, n\}} \subseteq \mathcal{M}[A, \mathcal{A}]$  a finite set of  $\mathcal{A}$ -measurable function graphs from  $X$  to  $\mathbb{R}$  then  $\sum_{i=1}^n f_i: X \rightarrow \mathbb{R}$  defined by  $(\sum_{i=1}^n f_i)(x) = \sum_{i=1}^n f_i(x)$  is  $\mathcal{A}$ -measurable [in other words  $\sum_{i=1}^n f_i \in \mathcal{M}[A, \mathcal{A}]$ ]

**Proof.** This is trivially proved by induction, so define  $\mathcal{S} = \{i \in \{1, \dots, n\} | \text{if } i \leq n \text{ then } \sum_{j=1}^i f_j \text{ is } \mathcal{A}\text{-measurable}\}$  then we have:

**$1 \in \mathcal{S}$ .** This follows from the fact that  $\sum_{j=1}^1 f_j = f_1$  and that  $f_1$  is  $\mathcal{A}$ -measurable.

**$i \in \mathcal{S} \Rightarrow i+1 \in \mathcal{S}$ .** Let  $i \in \mathcal{S}$  then if  $i+1 \leq n$  we have  $(\sum_{j=1}^{i+1} f_j)(x) = \sum_{j=1}^{i+1} f_j(x) = f_{i+1}(x) + \sum_{j=1}^i f_j(x)$ . So  $\sum_{j=1}^{i+1} f_j = f_{i+1} + \sum_{j=1}^i f_j$  which, as  $i \in \mathcal{S}$  is the sum of two  $\mathcal{A}$ -measurable functions, is  $\mathcal{A}$ -measurable by 19.173. Hence  $i+1 \in \mathcal{S}$   $\square$

**Definition 19.175.** Let  $X$  be a set and  $f: X \rightarrow \overline{\mathbb{R}}$  a extended real valued function then we define

1.  $f^+: X \rightarrow \overline{\mathbb{R}}$  by  $f^+(x) = \max(f(x), 0)$  [so  $f^+(X) \subseteq [0, \infty]$ ]
2.  $f^-: X \rightarrow \overline{\mathbb{R}}$  by  $f^-(x) = -\min(f(x), 0)$  [so  $f^-(X) \subseteq [0, \infty]$ ]
3.  $|f|: X \rightarrow \overline{\mathbb{R}}$  by  $|f|(x) = |f(x)|$  [so  $|f|(X) \subseteq [0, \infty]$ ]

**Lemma 19.176.** Let  $X$  be a set and  $f: X \rightarrow \overline{\mathbb{R}}$ ,  $g: X \rightarrow \overline{\mathbb{R}}$  extended real functions then

1.  $|f| = f^+ + f^-$  [note that the sum is well defined]
2.  $f = f^+ - f^-$  [note that the difference is well defined]
3. If  $\alpha \in [0, \infty[$  then  $(\alpha \cdot f)^+ = \alpha \cdot f^+$  and  $(\alpha \cdot f)^- = \alpha \cdot f^-$
4. If  $\alpha \in ]-\infty, 0]$  then  $(\alpha \cdot f)^+ = -\alpha \cdot f^-$  and  $(\alpha \cdot f)^- = -\alpha \cdot f^+$
5. If  $f(X) \subseteq \mathbb{R}$  and  $g(X) \subseteq \mathbb{R}$  then we have  $\forall x \in X$ 
  - a.  $(f + g)^+(x) \leq f^+(x) + g^+(x) = (f^+ + g^+)(x)$

- b.  $(f + g)^-(x) \leq f^-(x) + g^-(x) = (f^- + g^-)(x)$
6.  $\forall x \in X$  we have  $f^+(x) \leq |f(x)|$
7.  $\forall x \in X$  we have  $f^-(x) \leq |f(x)|$
8.  $f^+ = f \wedge C_0$
9.  $f^- = -(f \vee C_0)$

**Proof.**

1. Let  $x \in X$  then for  $x$  we have either

$f(x) \leq 0$ . then  $f^+(x) = \max(f(x), 0) = 0$  and  $f^-(x) = -\min(f(x), 0) = -f(x) = |f(x)|$   
 so that  $f^+(x) + f^-(x) = |f(x)|$

$0 < f(x)$ . then  $f^+(x) = \max(f(x), 0) = f(x)$  and  $f^-(x) = -\min(f(x), 0) = 0$  so that  
 $f^+(x) + f^-(x) = f(x) = |f(x)|$

2. Let  $x \in X$  then for  $x$  we have either

$f(x) \leq 0$ . then  $f^+(x) = \max(f(x), 0) = 0$  and  $f^-(x) = -\min(f(x), 0) = -f(x)$  so that  
 $f(x) = f^+(x) - f^-(x)$ .

$0 < f(x)$ . then  $f^+(x) = \max(f(x), 0) = f(x)$  and  $f^-(x) = -\min(f(x), 0) = 0$  so that  
 $f(x) = f^+(x) - f^{-1}(x)$

3. Let  $x \in X$  then we have either

$f(x) \leq 0$ . then

$$\begin{aligned} (\alpha \cdot f)^+(x) &= (\alpha \cdot f)^+(x) \\ &= \max(\alpha \cdot f(x), 0) \\ &= 0 \\ &= \alpha \cdot 0 \\ &= \alpha \cdot \max(f(x), 0) \\ &= \alpha \cdot f^+(x) \end{aligned}$$

and

$$\begin{aligned} (\alpha \cdot f)^{-1}(x) &= -\min(\alpha \cdot f(x), 0) \\ &= -\alpha \cdot f(x) \\ &= \alpha \cdot (-f(x)) \\ &= \alpha \cdot (-\min(f(x), 0)) \\ &= \alpha \cdot f^-(x) \end{aligned}$$

$0 < f(x)$ . then

$$\begin{aligned} (\alpha \cdot f)^+(x) &= \max(\alpha \cdot f(x), 0) \\ &= \alpha \cdot f(x) \\ &= \alpha \cdot \max(f(x), 0) \\ &= \alpha \cdot f^+(x) \end{aligned}$$

and

$$\begin{aligned} (\alpha \cdot f)^-(x) &= -\min(\alpha \cdot f(x), 0) \\ &= 0 \\ &= \alpha \cdot 0 \\ &= \alpha \cdot (-\min(f(x), 0)) \\ &= \alpha \cdot f^-(x) \end{aligned}$$

4. Let  $x \in X$  then for  $x$  we have either

$f(x) \leq 0$ . then

$$\begin{aligned} (\alpha \cdot f)^+(x) &= \max(\alpha \cdot f(x), 0) \\ &= \alpha \cdot f(x) \\ &= \alpha \cdot \min(f(x), 0) \\ &= -\alpha \cdot f^-(x) \end{aligned}$$

and

$$\begin{aligned} (\alpha \cdot f)^-(x) &= -\min(\alpha \cdot f(x), 0) \\ &= 0 \\ &= -\alpha \cdot 0 \\ &= -\alpha \cdot \max(f(x), 0) \\ &= -\alpha \cdot f^+(x) \end{aligned}$$

$0 < f(x)$ . then

$$\begin{aligned} (\alpha \cdot f)^+(x) &= \max(\alpha \cdot f(x), 0) \\ &= 0 \\ &= \alpha \cdot 0 \\ &= \alpha \cdot (\min(f(x), 0)) \\ &= -\alpha \cdot f^-(x) \end{aligned}$$

and

$$\begin{aligned} (\alpha \cdot f)^-(x) &= -\min(\alpha \cdot f(x), 0) \\ &= -\alpha \cdot f(x) \\ &= -\alpha \cdot \max(f(x), 0) \\ &= -\alpha \cdot f^+(x) \end{aligned}$$

5. Let  $x \in X$  then

a. As  $f(x) \leq \max(f(x), 0) = f^+(x)$  and  $g(x) \leq \max(g(x), 0) = g^+(x)$  we have

$$\begin{aligned} (f+g)(x) &= f(x) + g(x) \\ &\leq f^+(x) + g^+(x) \end{aligned}$$

which as  $0 \leq f^+(x) + g^+(x)$  proves that  $\max((f+g)(x), 0) \leq f^+(x) + g^+(x)$  so that

$$\begin{aligned} (f+g)^+(x) &= \max((f+g)(x), 0) \\ &\leq f^+(x) + g^+(x) \end{aligned}$$

b. Then  $-f^-(x) = \min(f(x), 0) \leq f(x)$  and  $-g^-(x) = \min(g(x), 0) \leq g(x)$  so that

$$\begin{aligned} -f^-(x) - g^-(x) &\leq f(x) + g(x) \\ &\leq (f+g)(x) \end{aligned}$$

which as  $-f^-(x) - g^-(x) \leq 0$  proves that  $-f^-(x) - g^-(x) \leq \min((f+g)(x), 0)$  so that

$$\begin{aligned} (f+g)^-(x) &= -\min((f+g)(x), 0) \\ &\leq f^-(x) + g^-(x) \end{aligned}$$

6. Let  $x \in X$  then  $f(x) \leq |f(x)|$  and  $0 \leq |f(x)|$  so that  $f^+(x) = \max(f(x), 0) \leq |f(x)|$

7. Let  $x \in X$  then  $-f(x) \leq |f(x)| \Rightarrow -|f(x)| \leq f(x)$  and  $-|f(x)| \leq 0$  so that  $\min(f(x), 0) \geq -|f(x)|$ . Hence  $f^-(x) = -\min(f(x), 0) \leq |f(x)|$

8. Let  $x \in X$  then  $f^+(x) = \max(f(x), 0) = \max(f(x), C_0(x)) = (f \wedge C_0)(x)$
9. Let  $x \in X$  then  $f^-(x) = -\min(f(x), 0) = -\min(f(x), C_0(x)) = -(f \vee C_0)(x)$

□

**Proposition 19.177.** Let  $\langle X, \mathcal{A} \rangle$  be a measurable space,  $A \in \mathcal{A}$  and  $f: A \rightarrow \overline{\mathbb{R}}$  a  $\mathcal{A}$ -measurable function then

1.  $f^+: A \rightarrow \overline{\mathbb{R}}$  is  $\mathcal{A}$ -measurable.
2.  $f^-: A \rightarrow \overline{\mathbb{R}}$  is  $\mathcal{A}$ -measurable.
3.  $|f|: A \rightarrow \overline{\mathbb{R}}$  is  $\mathcal{A}$ -measurable.

**Proof.** First note that by 19.155  $C_0: A \rightarrow \overline{\mathbb{R}}$  defined by  $C_0(x) = 0$  is  $\mathcal{A}$ -measurable.

1. As  $f^+ = f \wedge C_0$  and  $f, C_0$  are  $\mathcal{A}$ -measurable it follows from 19.165 that  $f^+$  is  $\mathcal{A}$ -measurable.
2. As  $f, C_0$  are  $\mathcal{A}$ -measurable it follows from 19.165 that  $f \vee C_0$  is  $\mathcal{A}$ -measurable. Using 19.170 it follows that  $f^- = -(f \vee C_0)$  is  $\mathcal{A}$ -measurable.
3. As  $|f| \stackrel{19.176}{=} f_+ + f_-$  it follows from (2),(2) and 19.173 that  $|f|$  is  $\mathcal{A}$ -measurable. □

The following theorem will be essential to find a equivalent definition of the measure integral.

**Theorem 19.178.** Let  $\langle X, \mathcal{A} \rangle$  be a measurable space,  $A \in \mathcal{A}$  and  $f: A \rightarrow \overline{\mathbb{R}}$  with  $f \in \overline{\mathcal{M}[A, \mathcal{A}]}$  and  $f(A) \subseteq [0, \infty]$  [a  $\mathcal{A}$ -measurable non negative extended real valued function] then there exists a sequence  $\{f_i\}_{i \in \mathbb{N}_0}$  of  **$\mathcal{A}$ -measurable simple functions** such that  $\forall x \in A$ .

1.  $\forall i \in \mathbb{N}_0 f_i(x) \leq f_{i+1}(x)$  and  $0 \leq f_i(x) < \infty$
2.  $f(x) = \sup(n \in \mathbb{N}_0 | f_n(x) |) \stackrel{(1)+17.71}{=} \lim_{n \rightarrow \infty} f_n(x)$  [or in other words  $f = \lim_{n \rightarrow \infty} f_n = \sup_{n \in \{1, \dots, \infty\}} f_n$ ]

**Proof.** Let  $n \in \mathbb{N}_0$  and define

$$I_n = \{i \in \mathbb{N}_0 | i \leq n \cdot 2^n\} = \{1, 2, \dots, n \cdot 2^n\} \quad (19.328)$$

Define for  $k \in I_n$

$$A_{n,k} = \left\{ x \in A \mid \frac{k-1}{2^n} \leq f(x) < \frac{k}{2^n} \right\} = f^{-1}\left(\left[\frac{k-1}{2^n}, \frac{k}{2^n}\right]\right) \quad (19.329)$$

Now

$$\begin{aligned} t \in A_{n,k} &\Leftrightarrow t \in A \wedge \frac{k-1}{2^n} \leq f(t) < \frac{k}{2^n} \\ &\Leftrightarrow \left( t \in A \wedge \frac{k-1}{2^n} \leq f(t) \right) \wedge \left( t \in A \wedge f(t) < \frac{k}{2^n} \right) \\ &\Leftrightarrow t \in \left\{ f \geq \frac{k-1}{2^n} \right\} \wedge t \in \left\{ f < \frac{k}{2^n} \right\} \\ &\Leftrightarrow t \in \left\{ f \geq \frac{k-1}{2^n} \right\} \cap \left\{ f < \frac{k}{2^n} \right\} \end{aligned}$$

proving that

$$A_{n,k} = \left\{ f \geq \frac{k-1}{2^n} \right\} \cap \left\{ f < \frac{k}{2^n} \right\} \quad (19.330)$$

As  $f$  is  $\mathcal{A}$ -measurable we have by 19.151 that  $\left\{ f \geq \frac{k-1}{2^n} \right\}, \left\{ f < \frac{k}{2^n} \right\} \in \mathcal{A}$ , so it follows that

$$\forall k \in I_n \text{ we have that } A_{n,k} \in \mathcal{A} \quad (19.331)$$

Let  $l, k \in I_n$  with  $l < k \Rightarrow l+1 \leq k \Rightarrow l \leq k-1$ , assume that  $t \in A_{n,l} \cap A_{n,k}$  then  $\frac{l-1}{2^n} \leq f(t) < \frac{l}{2^n} \leq f(t) < \frac{k}{2^n}$  giving the contradiction  $f(t) < f(t)$ , hence we must have that  $A_{n,l} \cap A_{n,k} = \emptyset$ . So

$$\forall l, k \in I_n \text{ with } l \neq k \text{ we have } A_{n,k} \cap A_{n,l} = \emptyset \quad (19.332)$$

If  $x \in \bigcup_{k \in I_n} A_{n,k}$  then there exists a  $k \in I_n$  such that  $x \in A_{n,k}$  and this  $k$  is unique because of 19.332. This allows us to define  $\{f_n\}_{n \in \mathbb{N}_0} \subseteq \overline{\mathbb{R}}^A$  by

$$f_n(x) = \begin{cases} \frac{k-1}{2^n} & \text{if } x \in \bigcup_{k \in I_n} A_{n,k} \text{ where } k \text{ is the unique } k \text{ such that } x \in A_{n,k} \\ n & \text{if } x \in A \setminus \bigcup_{k \in I_n} A_{n,k} \end{cases} < \infty \quad (19.333)$$

Using the above definition it is clear that  $f_n(A) \subseteq \{n\} \cup \left( \bigcup_{k \in I_n} \left\{ \frac{k-1}{2^n} \right\} \right)$  which is finite [see 5.39], taking this in account together with 19.331 proves

$$\forall n \in \mathbb{N}_0 \text{ } f_n \text{ is a simple } \mathcal{A}\text{-measurable function} \quad (19.334)$$

Let  $n \in \mathbb{N}_0$ , and  $x \in f^{-1}([0, n]) \Rightarrow 0 \leq f(x) < n$ . Define  $B_x = \left\{ k \in I_n \mid f(x) < \frac{k}{2^n} \right\} \subseteq I_n$  a finite set. As  $n \cdot 2^n \in I_n$  and  $f(x) < n = \frac{n \cdot 2^n}{2^n}$  we have that  $n \cdot 2^n \in B_x$  proving that  $B_x \neq \emptyset$ , so  $l = \min(B_x)$  exists. We have to consider two cases for  $l$ :

**$l = 1$ .** then  $\frac{1-1}{2^n} = 0 \leq f(x) < \frac{1}{2^n} \Rightarrow x \in A_{n,1} \subseteq \bigcup_{k \in I_n} A_{n,k}$

**$1 < l$ .** then  $1 \leq l-1 < l \leq n \cdot 2^n$  proving that  $l-1 \in I_n$ , and as  $l-1 < \min(B_x)$  we must have that  $\frac{l-1}{2^n} \leq f(x) < \frac{l}{2^n} \Rightarrow x \in A_{n,l} \subseteq \bigcup_{k \in I_n} A_{n,k}$

So we have proved that  $f^{-1}([0, n]) \subseteq \bigcup_{k \in I_n} A_{n,k}$ . Also for  $k \in I_n$  we have that  $A_{n,k} = f^{-1}\left(\left[\frac{k-1}{2^n}, \frac{k}{2^n}\right]\right) \subseteq f^{-1}([0, n])$ , hence

$$\forall n \in \mathbb{N}_0 \text{ we have } f^{-1}([0, n]) = \bigcup_{k \in I_n} A_{n,k} \quad (19.335)$$

Let  $x \in A$  and  $n \in \mathbb{N}$  then either:

**$n \leq f(x)$ .** then  $x \notin f^{-1}([0, n]) \stackrel{19.335}{=} \bigcup_{k \in I_n} A_{n,k} \Rightarrow x \in A \setminus \bigcup_{k \in I_n} A_{n,k}$ , so using 19.333 we have that  $f_n(x) = n \leq f(x)$

**$0 \leq f(x) < n$ .** then  $x \in f^{-1}([0, n]) \stackrel{19.335}{=} \bigcup_{k \in I_n} A_{n,k}$ , so there exists a unique  $k \in I_n$  such that  $x \in A_{n,k}$ , using 19.333 it follows that  $f_n(x) = \frac{k-1}{2^n} \leq f(x) < \frac{k}{2^n}$

To summarize

$$\forall x \in A, \forall n \in \mathbb{N}_0 \text{ we have } f_n(x) \leq f(x) \text{ and } f_n(x) = \begin{cases} n & \text{if } n \leq f(x) \\ \frac{k-1}{2^n} & \text{if } f(x) < n \text{ where } k \in I_n \end{cases} \quad (19.336)$$

and

$$\forall n \in \mathbb{N}_0 \text{ if } x \in A \text{ is such that } f(x) < n \text{ then } \exists k \in I_n \text{ such that } f_n(x) = \frac{k-1}{2^n} \leq f(x) < \frac{k}{2^n} \quad (19.337)$$

Let now  $x \in A$  then we must consider the following cases:

**$n+1 \leq f(x)$ .** then  $n, n+1 \leq f(x)$  and by 19.336  $f_n(x) = n < n+1 = f_{n+1}(x)$  so

$$f_n(x) < f_{n+1}(x)$$

**$n \leq f(x) < n+1$ .** then by 19.336, 19.337  $n = f_n(x)$  and  $\exists k \in I_{n+1}$  with  $f_{n+1}(x) = \frac{k-1}{2^{n+1}} \leq f(x) < \frac{k}{2^{n+1}}$ . As  $n \leq f(x)$  we have that  $n < \frac{k}{2^{n+1}} \Rightarrow n \cdot 2^{n+1} < k \Rightarrow n \cdot 2^{n+1} \leq k-1 \Rightarrow n \leq \frac{k-1}{2^{n+1}} = f_{n+1}(x)$  proving that

$$f_n(x) \leq f_{n+1}(x)$$

**$0 \leq f(x) < n$ .** then  $x \in f^{-1}([0, n])$ ,  $x \in f^{-1}([0, n+1])$ , so using 19.335  $\exists k \in I_n, l \in I_{n+1}$  such that  $x \in A_{n,k} \cap A_{n+1,l}$ , hence  $\frac{k-1}{2^n} \leq f(x) < \frac{k}{2^n} \wedge \frac{l-1}{2^{n+1}} \leq f(x) < \frac{l}{2^{n+1}}$ . From this it follows that  $\frac{k-1}{2^n} < \frac{l}{2^{n+1}} \Rightarrow \frac{2 \cdot k - 2}{2^{n+1}} < \frac{l}{2^{n+1}} \Rightarrow 2 \cdot k - 2 < l \Rightarrow 2 \cdot k - 2 \leq l-1$  giving  $\frac{k-1}{2^n} = \frac{2 \cdot k - 2}{2^{n+1}} \leq \frac{l-1}{2^{n+1}}$ . Using 19.333 we have that  $f_{n+1}(x) = \frac{l-1}{2^{n+1}}, f_n(x) = \frac{k-1}{2^n}$  proving that

$$f_n(x) \leq f_{n+1}(x)$$

So

$$\forall x \in A \text{ we have } \forall n \in \mathbb{N}_0 \text{ that } f_n(x) \leq f_{n+1}(x) \text{ and } 0 \leq f_n(x) < \infty \quad (19.338)$$

Let  $x \in A$ . Using 19.336 it follows that

$$\sup(\{f_n(x) | n \in \mathbb{N}_0\}) \leq f(x) \quad (19.339)$$

Assume now that  $\sup(\{f_n(x) | n \in \mathbb{N}_0\}) < f(x)$  [so that  $\sup(\{f_n(x) | n \in \mathbb{N}_0\}) \in [0, \infty)$ ]. For  $f(x)$  we have either:

**$f(x) = \infty$ .** then as  $\sup(\{f_n(x) | n \in \mathbb{N}_0\}) < \infty$  we have by 9.56 that  $\exists n \in \mathbb{N}_0$  such that  $\sup(\{f_n(x) | n \in \mathbb{N}_0\}) < n < \infty = f(x)$ . Using 19.336 we have that  $f_n(x) = n$ , so  $\sup(\{f_n(x) | n \in \mathbb{N}_0\}) < n = f_n(x)$  a contradiction.

**$f(x) < \infty$ .** Take  $\delta = f(x) - \sup(\{f_n(x) | n \in \mathbb{N}_0\}) > 0$ , using 9.41 there exists a  $n_1 \in \mathbb{N}_0$  such that  $\frac{1}{2^{n_1}} < \delta$ , using 9.56 there exists a  $n_2 \in \mathbb{N}_0$  such that  $f(x) < n_2$ . Take  $n = \max(n_1, n_2)$  then  $f(x) < n$  so that by 19.337 there exists a  $k \in I_n$  such that  $f_n(x) = \frac{k-1}{2^n} \leq f(x) < \frac{k}{2^n}$ . then  $f(x) - f_n(x) < \frac{k}{2^n} - \frac{k-1}{2^n} = \frac{1}{2^n} \leq \frac{1}{2^{n_1}} < \delta = f(x) - \sup(\{f_n(x) | n \in \mathbb{N}_0\})$ , so that  $\sup(\{f_n(x) | n \in \mathbb{N}_0\}) < f_n(x)$  a contradiction.

As we have in all cases a contradiction we must have that  $f(x) \leq \sup(\{f_n(x) | n \in \mathbb{N}_0\})$  which combined with 19.339 proves that

$$f(x) = \sup(\{f_n(x) | n \in \mathbb{N}_0\}) \quad (19.340)$$

The theorem is now proved by 19.334, 19.338 and 19.340.  $\square$

**Corollary 19.179.** Let  $\langle X, \mathcal{A} \rangle$  be a measurable space and  $f: X \rightarrow \overline{\mathbb{R}}$  a  $\mathcal{A}$ -measurable function then there exists a sequence  $\{f_i\}_{i \in \mathbb{N}_0}$  of  **$\mathcal{A}$ -measurable simple functions** between  $X$  and  $\overline{\mathbb{R}}$  such that

1.  $\forall n \in \mathbb{N}_0 f_n(X) \subseteq \mathbb{R}$
2.  $\forall x \in X f(x) = \lim_{n \rightarrow \infty} f_i(x)$  [or in other words  $f = \lim_{n \rightarrow \infty} f_n$ ]

**Proof.** Using 19.176 we have  $f = f^+ - f^-$  [where the difference is well defined] and as  $f$  is  $\mathcal{A}$ -measurable it follows from 19.177 that  $f^+$  and  $f^-$  are  $\mathcal{A}$ -measurable. Using the fact that  $f^+$ ,  $f^-$  are non negative we can apply the previous theorem 19.178 to find two sequences of  $\mathcal{A}$ -measurable simple functions  $\{f_n^+\}_{n \in \mathbb{N}_0}$ ,  $\{f_n^-\}_{n \in \mathbb{N}_0}$  with  $\forall n \in \mathbb{N}_0$  and  $x \in X$ ,  $f_n^+(x) \in \mathbb{R}$ ,  $f_n^-(x) \in \mathbb{R}$ ,  $f^+(x) = \lim_{n \rightarrow \infty} f_n^+(x)$  and  $f^-(x) = \lim_{n \rightarrow \infty} f_n^-(x)$ . Define now  $(f_n)_{n \in \mathbb{N}_0}$  by  $f_n = f_n^+ - f_n^-$  then by 19.160 and 19.170 it follows that  $f_n$  is a  $\mathcal{A}$ -measurable simple function. As  $\forall x \in X f^+(x) - f^-(x)$  is well defined and  $\forall n \in \mathbb{N}_0 f_n^+(x) - f_n^-(x)$  is well defined [difference of real numbers] we have using 17.77 that

$$\begin{aligned} \lim_{n \rightarrow \infty} f_n(x) &\stackrel{\text{def}}{=} \lim_{n \rightarrow \infty} (f_n^+(x) + (-1 \cdot f_n^-(x))) \\ &\stackrel{17.77}{=} \lim_{n \rightarrow \infty} f_n^+(x) + (-1) \cdot \lim_{n \rightarrow \infty} f_n^-(x) \\ &= f^+(x) + (-1) \cdot f^-(x) \\ &= f^+(x) - f^-(x) \\ &= (f^+ - f^-)(x) \\ &= f(x) \end{aligned}$$

So there exists a sequence of real simple functions  $\{f_n\}_{n \in \mathbb{N}_0}$  such that  $\lim_{n \rightarrow \infty} f_n(x) = f(x)$   $\square$

#### 19.5.4 Properties that hold everywhere

Using the concept of  $\mu$ -negligible sets [see 19.104] we can introduce the concept of almost everywhere predicates and complete measures.

**Definition 19.180.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space and  $P(x)$  a predicate then  $P(x)$  is  **$\mu$ -a.e. true** if  $\{x \in X | \neg(P(x))\} \in \mathcal{N}_\mu$  [see 19.104]. In other words  $P(x)$  is  **$\mu$ -a.e. true** if there exists a  $N \in \mathcal{A}$  with  $\mu(N) = 0$  such that  $\{x \in X | \neg(P(x))\} \subseteq N$ .

Using the above we can define the following

**Definition 19.181.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space and  $f: X \rightarrow \overline{\mathbb{R}}$ ,  $g: X \rightarrow \overline{\mathbb{R}}$  then

1.  $f = g$   **$\mu$ -a.e.** if there exists a  $N \in \mathcal{A}$  with  $\mu(N) = 0$  such that  $\{x \in X \mid f(x) \neq g(x)\} \subseteq N$ .
2.  $f \geq g$   **$\mu$ -a.e.** if there exists a  $N \in \mathcal{A}$  with  $\mu(N) = 0$  such that  $\{x \in X \mid f(x) < g(x)\} \subseteq N$ .
3.  $f = 0$   **$\mu$ -a.e.** if there exists a  $N \in \mathcal{A}$  with  $\mu(N) = 0$  such that  $\{x \in X \mid f(x) = 0\} \subseteq N$
4. Let  $\{f_n\}_{n \in \mathbb{N}_0}$  a sequence of functions from  $X$  to  $\overline{\mathbb{R}}$  then  $\{f_n\}_{n \in \mathbb{N}_0}$  converges  **$\mu$ -a.e.** to  $f$  if there exists a  $N \in \mathcal{A}$  with  $\mu(N) = 0$  such that  $\{x \in X \mid \{f_n(x)\}_{n \in \mathbb{N}_0} \text{ does not converges to } f(x)\} \subseteq N$ .

**Proposition 19.182.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space,  $A \subseteq X$  such that  $\exists N \in \mathcal{A}$  with  $\mu(N) = 0$  so that  $A \subseteq N$  and  $f: X \rightarrow \overline{\mathbb{R}}$ . If we define  $\mathcal{X}_{X \setminus A} \cdot f: X \rightarrow \overline{\mathbb{R}}$  by  $(\mathcal{X}_{X \setminus A} \cdot f)(x) = \mathcal{X}_{X \setminus A}(x) \cdot f(x)$  [see 2.14 for a definition of  $\mathcal{X}_{X \setminus A}$ ] then we have that

$$\mathcal{X}_{X \setminus A} \cdot f = f \quad \mu\text{-a.e.}$$

**Proof.** Let  $x \in \{x \in X \mid (\mathcal{X}_{X \setminus A} \cdot f)(x) \neq f(x)\}$  then  $x \in X$  and  $(\mathcal{X}_{X \setminus A} \cdot f)(x) \neq f(x)$ . Assume that  $x \notin A$  then  $x \in X \setminus A$  and  $(\mathcal{X}_{X \setminus A} \cdot f)(x) = \mathcal{X}_{X \setminus A}(x) \cdot f(x) = 1 \cdot f(x) = f(x)$  contradicting  $(\mathcal{X}_{X \setminus A} \cdot f)(x) \neq f(x)$ , so we must have that  $x \in A$ . Hence  $\{x \in X \mid (\mathcal{X}_{X \setminus A} \cdot f)(x) \neq f(x)\} \subseteq A \subseteq N$  proving that

$$\mathcal{X}_{X \setminus A} \cdot f = f \quad \mu\text{-a.e.}$$

□

**Proposition 19.183.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space and  $f: X \rightarrow \overline{\mathbb{R}}$ ,  $g: X \rightarrow \overline{\mathbb{R}}$ ,  $h: X \rightarrow \overline{\mathbb{R}}$ ,  $k: X \rightarrow \overline{\mathbb{R}}$

1. If  $f = g$   $\mu$ -a.e. and  $g = h$   $\mu$ -a.e. then  $f = h$   $\mu$ -a.e.
2. If  $f \leq g$   $\mu$ -a.e. and  $g \leq h$   $\mu$ -a.e. then  $f \leq h$   $\mu$ -a.e..
3. If  $f \leq g$   $\mu$ -a.e. and  $g \leq h$   $\mu$ -a.e. then  $f = g$   $\mu$ -a.e..
4. If  $\alpha \in \mathbb{R} \setminus \{0\}$  then if  $f = g$   $\mu$ -a.e. then  $\alpha \cdot f = \alpha \cdot g$   $\mu$ -a.e..
5. If  $\forall x \in X$  we have that  $f(x) + g(x)$  and  $h(x) + k(x)$  is well defined and in addition  $f = h$   $\mu$ -a.e. and  $g = k$   $\mu$ -a.e. then  $f + g = h + k$   $\mu$ -a.e.

**Proof.**

1. As  $f = g$   $\mu$ -a.e. and  $g = h$   $\mu$ -a.e. there exists  $N_1, N_2 \in \mathcal{A}$  with  $\mu(N_1) = 0 = \mu(N_2)$  such that  $\{x \in X \mid f(x) \neq g(x)\} \subseteq N_1$  and  $\{x \in X \mid g(x) \neq h(x)\} \subseteq N_2$ . Let  $x \in \{x \in X \mid f(x) \neq h(x)\}$  then  $f(x) \neq g(x) \vee g(x) \neq h(x)$  [for if  $f(x) = g(x) \wedge g(x) = h(x) \Rightarrow f(x) = h(x)$ ] so that  $x \in \{x \in X \mid f(x) \neq g(x)\} \cup \{x \in X \mid g(x) \neq h(x)\} \subseteq N_1 \cup N_2$ . As  $\mu(N_1 \cup N_2) \leq \mu(N_1) + \mu(N_2) = 0$  it follows that  $f = h$   $\mu$ -a.e..
2. As  $f \leq g$   $\mu$ -a.e. and  $g \leq h$   $\mu$ -a.e. there exists  $N_1, N_2 \in \mathcal{A}$  with  $\mu(N_1) = 0 = \mu(N_2)$  such that  $\{x \in X \mid f(x) \leq g(x)\} \subseteq N_1$  and  $\{x \in X \mid g(x) \leq h(x)\} \subseteq N_2$ . Let  $x \in \{x \in X \mid f(x) \leq h(x)\}$  then  $f(x) \leq g(x) \vee g(x) \leq h(x)$  [for if  $f(x) \leq g(x) \wedge g(x) \leq h(x) \Rightarrow f(x) \leq h(x)$ ] so that  $x \in \{x \in X \mid f(x) \leq g(x)\} \cup \{x \in X \mid g(x) \leq h(x)\} \subseteq N_1 \cup N_2$ . As  $\mu(N_1 \cup N_2) \leq \mu(N_1) + \mu(N_2) = 0$  it follows that  $f \leq h$   $\mu$ -a.e..
3. As  $f \leq g$   $\mu$ -a.e. and  $g \leq h$   $\mu$ -a.e. there exists  $N_1, N_2 \in \mathcal{A}$  with  $\mu(N_1) = 0 = \mu(N_2)$  such that  $\{x \in X \mid f(x) \leq g(x)\} \subseteq N_1$  and  $\{x \in X \mid g(x) \leq h(x)\} \subseteq N_2$ . Let  $x \in \{x \in X \mid f(x) \neq g(x)\}$  then  $f(x) \leq g(x) \vee g(x) \leq h(x)$  [for if  $f(x) \leq g(x) \wedge g(x) \leq h(x) \Rightarrow f(x) = g(x)$ ] so that  $x \in \{x \in X \mid f(x) \leq g(x)\} \cup \{x \in X \mid g(x) \leq h(x)\} \subseteq N_1 \cup N_2$ . As  $\mu(N_1 \cup N_2) \leq \mu(N_1) + \mu(N_2) = 0$  it follows that  $f = h$   $\mu$ -a.e..
4. As  $f = g$   $\mu$ -a.e. there exists a  $N \in \mathcal{A}$  with  $\mu(N) = 0$  such that  $\{x \in X \mid f(x) \neq g(x)\} \subseteq N$ . Let now  $x \in \{x \in X \mid \alpha \cdot f(x) \neq \alpha \cdot g(x)\}$  then  $\alpha \cdot f(x) \neq \alpha \cdot g(x) \xrightarrow{\alpha \neq 0} f(x) \neq g(x)$  proving that  $\{x \in X \mid \alpha \cdot f(x) \neq \alpha \cdot g(x)\} \subseteq \{x \in X \mid f(x) \neq g(x)\} \subseteq N$ . So  $\alpha \cdot f \neq \alpha \cdot g$ .

5. As  $f = h$   $\mu$ -a.e. and  $g = k$   $\mu$ -a.e. there exists  $N_1, N_2 \in \mathcal{A}$  with  $\mu(N_1) = 0 = \mu(N_2)$  such that  $\{x \in X \mid f(x) \neq h(x)\} \subseteq N_1$  and  $\{x \in X \mid g(x) \neq k(x)\} \subseteq N_2$ . Let  $x \in \{x \in X \mid f(x) + g(x) \neq h(x) + k(x)\}$  then  $f(x) \neq h(x) \vee g(x) \neq k(x)$  [for if  $f(x) = h(x) \wedge g(x) = k(x) \Rightarrow f(x) + g(x) = h(x) + k(x)$ ] so that  $x \in \{x \in X \mid f(x) \neq h(x)\} \cup \{x \in X \mid g(x) \neq k(x)\} \subseteq N_1 \cup N_2$ . As  $\mu(N_1 \cup N_2) \leq \mu(N_1) + \mu(N_2) = 0$  it follows that  $f + g = h + k$   $\mu$ -a.e..  $\square$

**Proposition 19.184.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a **complete** measure space [see 19.106] and let  $f: X \rightarrow \overline{\mathbb{R}}$  a  $\mathcal{A}$ -measurable function and  $g: X \rightarrow \overline{\mathbb{R}}$  such that  $f(x) = g(x)$   $\mu$ -a.e. then  $g(x)$  is  $\mathcal{A}$ -measurable.

**Proof.** As  $f = g$   $\mu$ -a.e. there exists a  $N \in \mathcal{A}$  with  $\mu(N) = 0$  and  $\{x \in X \mid f(x) \neq g(x)\} \subseteq N$ . So

$$x \in X \setminus N \Rightarrow f(x) = g(x) \quad (19.341)$$

. Let  $t \in \mathbb{R}$  then

$$\begin{aligned} x \in \{g \leq t\} &\Leftrightarrow x \in X \wedge g(x) \leq t \\ &\Leftrightarrow (x \in X \setminus N \vee x \in N) \wedge g(x) \leq t \\ &\Leftrightarrow (x \in X \setminus N \wedge g(x) \leq t) \vee (x \in N \wedge g(x) \leq t) \\ &\stackrel{19.341}{\Leftrightarrow} (x \in X \setminus N \wedge f(x) \leq t) \cup (x \in N \wedge g(x) \leq t) \\ &\stackrel{X \setminus N, N \subseteq X}{\Leftrightarrow} (x \in X \setminus N \wedge x \in X \wedge f(x) \leq t) \cup (x \in N \wedge x \in X \wedge g(x) \leq t) \\ &\Leftrightarrow x \in (X \setminus N) \cap \{f \leq t\} \wedge x \in N \cap \{g \leq t\} \\ &\Leftrightarrow x \in (X \setminus N) \cap \{f \leq t\} \cup (N \cap \{g \leq t\}) \end{aligned} \quad (19.342)$$

As  $N \cap \{g \leq t\} \subseteq N$ ,  $\mu(N) = 0$  and  $\mu$  is complete [see 19.106 for a definition] we have  $N \cap \{g \leq t\} \in \mathcal{A}$ . As  $f$  is  $\mu$ -measurable we have by 19.151 that  $\{f \leq t\} \in \mathcal{A}$ , also as  $N \in \mathcal{A} \Rightarrow X \setminus N \in \mathcal{A}$ ,  $((X \setminus N) \cap \{f \leq t\}) \in \mathcal{A}$ . So using 19.342 we have that  $\{g \leq t\} \in \mathcal{A}$  which by 19.151 proves that

$g$  is  $\mathcal{A}$ -measurable  $\square$

**Corollary 19.185.** Let  $\langle X, \mathcal{A}, \mu \rangle$  a complete measure space,  $\{f_n\}_{n \in \mathbb{N}_0} \subseteq \overline{\mathcal{M}[X, \mathcal{A}]}$  be a sequence of  $\mathcal{A}$ -measurable function graphs from  $X \rightarrow \overline{\mathbb{R}}$ ,  $f: X \rightarrow \overline{\mathbb{R}}$  such that  $\{f_n(x)\}_{n \in \mathbb{N}_0}$  converges  $\mu$ -a.e. to  $f$  then  $f$  is  $\mathcal{A}$ -measurable.

**Proof.** Take  $B = \{x \in X \mid \{f_n(x)\}_{n \in \mathbb{N}_0} \text{ does not converges to } f(x)\}$ . As  $\{f_n(x)\}_{n \in \mathbb{N}_0}$  converges  $\mu$ -a.e. to  $f$  there exists a  $N \in \mathcal{A}$  with  $B \subseteq N$  and  $\mu(N) = 0$ . If  $x \in X \setminus B$  then  $f(x) = \lim_{n \rightarrow \infty} f_n(x) = \liminf_{n \rightarrow \infty} f_n(x)$  so we have that  $f = \liminf_{n \rightarrow \infty} f_n$   $\mu$ -a.e.. As by 19.169  $\liminf_{n \rightarrow \infty} f_n$  is  $\mathcal{A}$ -measurable [given that  $\forall n \in \mathbb{N}_0$   $f_n$  is  $\mathcal{A}$ -measurable] we have using the previous proposition 19.184 that

$f$  is  $\mathcal{A}$ -measurable  $\square$

The following lemma show another way of representing a simple function that will be usefull to define the integral on a simple function.

**Lemma 19.186.** Let  $X$  be a set and  $f: X \rightarrow \overline{\mathbb{R}}$  be a simple function then there exists a  $n \in \mathbb{N}_0$  and a bijection  $\beta: \{1, \dots, n\} \rightarrow f(X)$  such that for  $\{A_i\}_{i \in \{1, \dots, n\}}$  defined by  $A_i = f^{-1}(\{\beta(i)\})$  we have

1.  $X = \bigsqcup_{i \in \{1, \dots, n\}} A_i$  / $X$  is the pairwise disjoint union of  $A_i$ 's
2.  $\forall x \in X$  we have that  $f(x) = \sum_{i=1}^n \beta(i) \cdot \chi_{A_i}(x)$  /or in other words  $f = \sum_{i=1}^n \beta(i) \cdot \chi_{A_i}$ .

**Proof.** As  $f$  is a simple function we have by definition that  $f(X)$  is finite and non empty, so there exists a  $n \in \mathbb{N}_0$  and a bijection  $\beta: \{1, \dots, n\} \rightarrow f(X)$ . Define now  $\{A_i\}_{i \in \{1, \dots, n\}}$  by  $f^{-1}(\{\beta(i)\})$  then as  $A_i = f^{-1}(\{\beta(i)\}) \subseteq X$  we have  $\bigcup_{i \in \{1, \dots, n\}} A_i \subseteq X$ . Now if  $x \in X$  then  $f(x) \in f(X)$  so that  $\exists i \in \{1, \dots, n\}$  with  $f(x) = \beta(i) \Rightarrow x \in f^{-1}(\{\beta(i)\}) = A_i$  proving that  $X \subseteq \bigcup_{i \in \{1, \dots, n\}} A_i$ . Hence we have

$$X = \bigcup_{i \in \{1, \dots, n\}} A_i \quad (19.343)$$

Further if  $i, j \in \{1, \dots, n\}$  with  $i \neq j$  then if  $x \in A_i \cap A_j$  we have  $x \in f^{-1}(\{\beta(i)\}) \wedge x \in f^{-1}(\{\beta(j)\})$  or  $\beta(i) = f(x) = \beta(j)$   $\beta$  is a bijection  $\Rightarrow i = j$  a contradiction. So we must have that  $A_i \cap A_j = \emptyset$  which using 19.343 proves

$$X = \bigsqcup_{i \in \{1, \dots, n\}} A_i \quad (19.344)$$

If now  $x \in X$  then by the above there exists a  $i \in \{1, \dots, n\}$  such that  $x \in A_i \Rightarrow f(x) = \beta(i)$  and  $\forall j \in \{1, \dots, n\} \setminus \{i\} x \notin A_j$ . So

$$\begin{aligned} f(x) &= \beta(i) \\ &\stackrel{x \in A_i}{=} \beta(i) \cdot \mathcal{X}_{A_i}(x) \\ &\stackrel{x \notin A_j \text{ if } j \in \{1, \dots, n\} \setminus \{i\}}{=} \beta(i) \cdot \mathcal{X}_{A_i}(x) + \sum_{j \in \{1, \dots, n\} \setminus \{i\}} \beta(j) \cdot \mathcal{X}_{A_i}(x) \\ &= \sum_{j \in \{1, \dots, n\}} \beta(j) \cdot \mathcal{X}_{A_j}(x) \end{aligned}$$

proving the lemma.  $\square$

**Lemma 19.187.** Let  $X$  be a set,  $n \in \mathbb{N}_0$ ,  $\{A_i\}_{i \in \{1, \dots, n\}}$ ,  $\{B_i\}_{i \in \{1, \dots, n\}} \subseteq \mathcal{P}(X)$  such that  $\forall i \in \{1, \dots, n\}$  we have  $A_i \subseteq B_i$  and  $\{a_i\}_{i \in \{1, \dots, n\}} \subseteq [0, \infty]$  then if we define

$$\begin{aligned} f: X \rightarrow \overline{\mathbb{R}} \text{ by } f(x) &= \sum_{i=1}^n \alpha_i \cdot \mathcal{X}_{A_i}(x) \text{ [where } \mathcal{X}_{A_i} \text{ is the characteristics function [see 2.14]]} \\ g: X \rightarrow \overline{\mathbb{R}} \text{ by } g(x) &= \sum_{i=1}^n \alpha_i \cdot \mathcal{X}_{B_i}(x) \text{ [where } \mathcal{X}_{B_i} \text{ is the characteristics function [see 2.14]]} \end{aligned}$$

we have

1.  $\forall x \in X$  we have  $f(x) \leq g(x)$
2.  $\forall x \in X \setminus (\bigcup_{i \in \{1, \dots, n\}} (B_i \setminus A_i))$  we have  $f(x) = g(x)$

**Proof.** First given  $x \in X$  define

$$A_x = \{i \in \{1, \dots, n\} | x \in A_i\} \text{ and } B_x = \{i \in \{1, \dots, n\} | x \in B_i\}$$

If  $i \in A_x$  we have by definition that  $x \in A_i$ , hence as  $A_i \subseteq B_i$  we have  $x \in B_i \Rightarrow i \in B_x$  proving that

$$A_x \subseteq B_x \quad (19.345)$$

Next we have

$$\begin{aligned} f(x) &= \sum_{i \in \{1, \dots, n\}} \alpha_i \cdot \mathcal{X}_{A_i}(x) \\ &= \sum_{i \in \{1, \dots, n\} \setminus A_x} \alpha_i \cdot \mathcal{X}_{A_i}(x) + \sum_{i \in A_x} \alpha_i \cdot \mathcal{X}_{A_i}(x) \\ &= 0 + \sum_{i \in A_x} \alpha_i \\ &\stackrel{A_x \subseteq B_x}{\leq} \sum_{i \in B_x} \alpha_i \\ &= \sum_{i \in \{1, \dots, n\} \setminus B_x} \alpha_i \cdot \mathcal{X}_{B_i}(x) + \sum_{i \in B_x} \alpha_i \cdot \mathcal{X}_{B_i}(x) \\ &= \sum_{i \in \{1, \dots, n\}} \alpha_i \cdot \mathcal{X}_{B_i}(x) \\ &= g(x) \end{aligned}$$

proving

$$\forall x \in X \text{ we have } f(x) \leq g(x) \quad (19.346)$$

This concludes part (1) of the lemma. For part (2) assume that  $x \in X \setminus (\bigcup_{i \in \{1, \dots, n\}} (B_i \setminus A_i))$  so that  $x \notin \bigcup_{i \in \{1, \dots, n\}} (B_i \setminus A_i)$ . Let  $j \in B_x$  we have by definition  $x \in B_j$ , if  $x \notin A_j$  then  $x \in B_j \setminus A_j$  contradicting  $x \notin \bigcup_{i \in \{1, \dots, n\}} (B_i \setminus A_i)$ , so we must have that  $x \in A_j$  or  $j \in A_x$ , which proves that  $B_x \subseteq A_x$ . Combining this with 19.345 proves that

$$\forall x \in X \setminus \left( \bigcup_{i \in \{1, \dots, n\}} (B_i \setminus A_i) \right) \text{ we have } A_x = B_x \quad (19.347)$$

Hence for  $x \in X \setminus (\bigcup_{i \in \{1, \dots, n\}} (B_i \setminus A_i))$

$$\begin{aligned} f(x) &= \sum_{i \in \{1, \dots, n\}} \alpha_i \cdot \mathcal{X}_{A_i}(x) \\ &= \sum_{i \in \{1, \dots, n\} \setminus A_x} \alpha_i \cdot \mathcal{X}_{A_i}(x) + \sum_{i \in A_x} \alpha_i \cdot \mathcal{X}_{A_i}(x) \\ &= \sum_{i \in A_x} \alpha_i \\ &\stackrel{A_x = B_x}{=} \sum_{i \in B_x} \alpha_i \\ &= \sum_{i \in \{1, \dots, n\} \setminus B_x} \alpha_i \cdot \mathcal{X}_{B_i}(x) + \sum_{i \in B_x} \alpha_i \cdot \mathcal{X}_{B_i}(x) \\ &= \sum_{i \in \{1, \dots, n\}} \alpha_i \cdot \mathcal{X}_{B_i}(x) \\ &= g(x) \end{aligned}$$

proving

$$\forall x \in X \setminus \left( \bigcup_{i \in \{1, \dots, n\}} (B_i \setminus A_i) \right) \text{ we have } f(x) = g(x) \quad \square$$

**Proposition 19.188.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space,  $\mathcal{A}_\mu$  the completion of  $\mathcal{A}$  under  $\mu$  [see 19.110, 19.111] and  $f: X \rightarrow \overline{\mathbb{R}}$  a function then the following are equivalent

1.  $\exists g, h: X \rightarrow \overline{\mathbb{R}}$  such that
  - a.  $g, h$  are  $\mathcal{A}$ -measurable
  - b.  $\forall x \in X \ g(x) \leq f(x) \leq h(x)$
  - c.  $g = h$   $\mu$ -a.e.
2.  $f$  is  $\mathcal{A}_\mu$ -measurable

**Proof.**

**1  $\Rightarrow$  2.** Let  $x \in \mathbb{R}$  then as  $g$  is  $\mathcal{A}$ -measurable we have by 19.151 that  $\{g \leq x\} \in \mathcal{A}$ , further as by 19.111  $\mathcal{A} \subseteq \mathcal{A}_\mu$  we have that  $\{g \leq x\} \in \mathcal{A}_\mu$ . So using 19.151 we have that

$$g \text{ is } \mathcal{A}_\mu\text{-measurable} \quad (19.348)$$

Take  $B = \langle x \in X \mid g(x) \neq h(x) \rangle$  then by (1.c) there exists a  $N \in \mathcal{A}$  with  $\mu(N) = 0$  and  $B \subseteq N$ . Using 19.111 we have that  $\mathcal{A} \subseteq \mathcal{A}_\mu$  and  $(\bar{\mu})_{|\mathcal{A}} = \mu$  so that

$$N \in \mathcal{A}_\mu \text{ and } \bar{\mu}(N) = \mu(N) = 0 \quad (19.349)$$

If  $x \in X \setminus N \subseteq X \setminus B$  then  $g(x) = h(x) \stackrel{1.b}{\Rightarrow} g(x) = f(x)$ , so if  $x \in X$  is such that  $f(x) \neq g(x)$  then  $x \notin X \setminus N \Rightarrow x \in X \setminus (X \setminus N) = N$  so

$$\{x \in X \mid f(x) \neq g(x)\} \subseteq N \quad (19.350)$$

Combining 19.349, 19.350 proves that  $f = g$   $\bar{\mu}$ -almost everywhere. As  $g$  is  $\mathcal{A}_\mu$ -measurable [see 19.348] we can apply 19.184 proving that

$$f \text{ is } \mathcal{A}_\mu\text{-measurable}$$

**2⇒1.** We prove this in three steps, first we prove the proposition for non negative real simple functions then for simple functions and finally the general case.

1. Let  $s: X \rightarrow \overline{\mathbb{R}}$  be a real non negative  $\mathcal{A}_\mu$ -measurable simple function. Using 19.186 there exists a  $n \in \mathbb{N}_0$  and a bijection  $\alpha: \{1, \dots, n\} \rightarrow s(X) \subseteq [0, \infty]$  such that, if we define  $\{A_i\}_{i \in \{1, \dots, n\}}$  by  $A_i = s^{-1}(\{\alpha(i)\})$ , it follows

$$X = \bigsqcup_{i \in \{1, \dots, n\}} A_i \text{ and } \forall x \in X \text{ we have } s(x) = \sum_{i=1}^n \alpha(i) \cdot \chi_{A_i}(x) \quad (19.351)$$

As  $s$  is  $\mathcal{A}_\mu$ -measurable we have by 19.161 that  $\forall i \in \{1, \dots, n\}$   $A_i \in \mathcal{A}_\mu$  which using the definition of  $\mathcal{A}_\mu$  [see 19.110] gives

$$\forall i \in \{1, \dots, n\} \text{ there exists } C_i, B_i \in \mathcal{A} \text{ such that } C_i \subseteq A_i \subseteq B_i \text{ and } \mu(B_i \setminus C_i) = 0 \quad (19.352)$$

Define now  $g = \sum_{i=1}^n \alpha(i) \cdot \chi_{C_i}$  and  $h = \sum_{i=1}^n \alpha(i) \cdot \chi_{B_i}$  then by 19.170, 19.171 and 19.157 we have that

$$g, h \text{ are } \mathcal{A}\text{-measurable} \quad (19.353)$$

Using the previous lemma 19.187 and 19.352 it follows that

$$\forall x \in X \text{ we have } g(x) \leq s(x) \leq h(x) \quad (19.354)$$

Using the previous lemma 19.187 and 19.352 again we find that  $\forall x \in X \setminus (\bigcup_{i \in \{1, \dots, n\}} (B_i \setminus C_i))$   $g(x) = h(x)$ . So if  $x \in X$  is such that  $g(x) \neq h(x)$  we must have that  $x \notin X \setminus (\bigcup_{i \in \{1, \dots, n\}} (B_i \setminus C_i))$  or  $\{x \in X | g(x) \neq h(x)\} \subseteq \bigcup_{i \in \{1, \dots, n\}} (B_i \setminus C_i)$ . As  $\mu(\bigcup_{i \in \{1, \dots, n\}} (B_i \setminus C_i)) \leq 19.66 \sum_{i=1}^n \mu(B_i \setminus C_i) \stackrel{19.352}{=} 0$  we conclude that

$$g = h \text{ } \mu\text{-a.e.} \quad (19.355)$$

2. Let  $s: X \rightarrow \overline{\mathbb{R}}$  be a real valued  $\mathcal{A}_\mu$ -measurable function [not necessary non negative] then using 19.176 we have that  $s = s^+ - s^-$ . As  $s^+(X) = \{\max(s(x), 0) | x \in X\} \subseteq s(X) \cup \{0\}$  a finite set in  $\mathbb{R}$  and  $s^-(X) = \{-\min(s(x), 0) | x \in X\} \subseteq \{-y | y \in s(X)\} \cup \{0\}$  a finite set in  $\mathbb{R}$ , it follows that  $s^+$  and  $s^-$  are simple functions. Further using 19.177 and the  $\mathcal{A}_\mu$ -measurability of  $s$  we have that  $s^+$ ,  $s^-$  are  $\mathcal{A}_\mu$ -measurable. So we can use (1) [19.353, 19.354, 19.355] to find  $\mathcal{A}$ -measurable function  $g^+, g^-, h^+, h^-$  such that

$$\forall x \in X \text{ } g^+(x) \leq s^+(x) \leq h^+(x) \text{ and } g^-(x) \leq s^-(x) \leq h^-(x) \quad (19.356)$$

$$h^+ = g^+ \text{ } \mu\text{-a.e. and } h^- = g^- \text{ } \mu\text{-a.e.} \quad (19.357)$$

Let  $x \in X$  then we have (taking in account that all the terms in the sum are finite)

$$g^+(x) - h^-(x) \leq s^+(x) - s^-(x) \leq h^+(x) - g^-(x) \quad (19.358)$$

If we define  $g = g^+ - h^-$  and  $h = h^+ - g^-$  then using 19.173 and the above we have that

$$g, h \text{ are } \mathcal{A}\text{-measurable and } \forall x \in X \text{ } g(x) \leq s(x) \leq h(x) \quad (19.359)$$

Further as  $h^+ = g^+ \mu\text{-a.e.}$  and  $h^- = g^- \mu\text{-a.e.}$  we have by 19.183 and 19.357 that  $g^+ - h^- = h^+ - g^- \mu\text{-a.e.}$  proving that

$$g = h \text{ holds } \mu\text{-a.e.} \quad (19.360)$$

3. Now for the general case. Let  $f: X \rightarrow \overline{\mathbb{R}}$  be a  $\mathcal{A}_\mu$ -measurable function. Using 19.179 there exists a sequence  $\{f_n\}_{n \in \mathbb{N}_0}$  of  $\mathcal{A}_\mu$ -measurable real valued simple functions such that

$$\forall x \in X \text{ } f(x) = \lim_{n \rightarrow \infty} f_n(x) \quad (19.361)$$

Let  $n \in \mathbb{N}_0$  then using 19.359 and 19.360 there exists a  $\mathcal{A}$ -measurable functions  $g_n$ ,  $h_n$  such that

$$\forall x \in X \text{ } g_n(x) \leq f_n(x) \leq h_n(x) \text{ and } g_n = h_n \mu\text{-a.e.} \quad (19.362)$$

Define now  $g$  and  $h$  as  $g = \limsup_{n \rightarrow \infty} g_n$  and  $h = \liminf_{n \rightarrow \infty} h_n$  then using 19.169

$$g, h \text{ are } \mathcal{A}\text{-measurable} \quad (19.363)$$

Let  $x \in X$  then we have

$$\begin{aligned} g(x) &= \limsup_{n \rightarrow \infty} g_n(x) \\ &\leq_{19.362 \text{ and } 17.55} \limsup_{n \rightarrow \infty} f_n(x) \\ &\stackrel{19.361 \text{ def of lim}}{=} f(x) \\ &\stackrel{19.361 \text{ def of lim}}{=} \liminf_{n \rightarrow \infty} f_n(x) \\ &\leq_{19.362 \text{ and } 17.55} \liminf_{n \rightarrow \infty} h_n(x) \\ &= h(x) \end{aligned}$$

proving that

$$\forall x \in X \quad g(x) \leq f(x) \leq h(x) \quad (19.364)$$

Let  $n \in \mathbb{N}_0$  then as  $g_n = h_n$  is  $\mu$ -almost anywhere there exists  $s \in \mathcal{A}$  with  $\mu(N_s) = 0$  such that  $A_n = \{x \in X \mid g_n(x) \neq h_n(x)\} \subseteq N_s$ . Let  $x \in X \setminus (\bigcup_{i \in \mathbb{N}_0} N_i) = \bigcap_{i \in \mathbb{N}_0} (X \setminus N_i)$  then  $\forall i \in \mathbb{N}_0$  we have  $x \in X \setminus N_i$  proving that  $g_i(x) = h_i(x)$  so that

$$\begin{aligned} h(x) &= \liminf_{n \rightarrow \infty} h_n(x) \\ &= \liminf_{n \rightarrow \infty} g_n(x) \\ &\leq_{17.61} \limsup_{n \rightarrow \infty} g_n(x) \\ &= g(x) \end{aligned}$$

which as by 19.364  $g(x) \leq h(x)$  proves that  $h(x) = g(x)$ . Hence

$$\forall x \in X \setminus \left( \bigcup_{i \in \mathbb{N}_0} N_i \right) \text{ we have } h(x) = g(x)$$

So if  $x \in \{x \in X \mid g(x) \neq h(x)\}$  we must have that  $x \notin X \setminus (\bigcup_{i \in \mathbb{N}_0} N_i)$  or  $x \in \bigcup_{i \in \mathbb{N}_0} N_i$  proving that  $\{x \in X \mid g(x) \neq h(x)\} \subseteq \bigcup_{i \in \mathbb{N}_0} N_i$  which as  $\mu(\bigcup_{i \in \mathbb{N}_0} N_i) \leq \sum_{i=1}^{\infty} \mu(N_i) = 0$  proves that:

$$g = h \quad \mu\text{-a.e.} \quad (19.365)$$

Finally  $(1 \Rightarrow 2)$  is proved by 19.363, 19.364 and 19.365.  $\square$

## 19.6 The measure integral

We are now ready to define the measure integral on measurable functions. We do this in three steps, first we define the integral on real valued non negative simple functions, then on non negative extended real measurable functions and finally on real valued measurable functions.

### 19.6.1 Integral of simple functions

Recap the definition of a simple function [see 19.159], we define now the set of measurable simple functions on a measurable space as follows:

**Definition 19.189.** Let  $\langle X, \mathcal{A} \rangle$  be a measurable space then the set of measurable simple real valued functions  $\mathcal{S}[X, \mathcal{A}] \subseteq \overline{\mathcal{M}[X, \mathcal{A}]}$  is defined as

$$\begin{aligned} \mathcal{S}[X, \mathcal{A}] &= \{f \in \mathbb{R}^X \mid f \text{ is a simple function and } \mathcal{A}\text{-measurable}\} \\ &\stackrel{19.159 \text{ and } 19.161}{=} \{f \in \mathbb{R}^X \mid f(X) \text{ is finite and } \forall y \in f(X) \quad f^{-1}(y) \in \mathcal{A}\} \end{aligned}$$

Further the set of non negative real valued simple functions is noted as  $\mathcal{S}^+[X, \mathcal{A}]$  hence

$$\mathcal{S}^+[X, \mathcal{A}] = \{f \in \mathcal{S}[X, \mathcal{A}] \mid f(X) \subseteq [0, \infty[\}$$

**Proposition 19.190.** Let  $\langle X, \mathcal{A} \rangle$  be a measurable space and  $f \in \mathcal{S}[X, \mathcal{A}]$  then there exists a  $n \in \mathbb{N}_0$  and families  $\{A_i\}_{i \in \{1, \dots, n\}} \subseteq \mathcal{A}$ ,  $\{a_i\}_{i \in \{1, \dots, n\}} \subseteq \mathbb{R}$  such that  $\forall i, j \in \{1, \dots, n\}$  with  $i \neq j$  we have  $A_i \cap A_j = \emptyset$  and

$$f = \sum_{i=1}^n a_i \cdot \chi_{A_i}$$

Further if in addition  $f \in \mathcal{S}^+[X, \mathcal{A}]$  then  $\{a_i\}_{i \in \{1, \dots, n\}} \subseteq [0, \infty[$

**Proof.** Using 19.186 there exists a  $n \in \mathbb{N}_0$  and a bijection  $\beta: \{1, \dots, n\} \rightarrow f(X)$  such that for  $\{A_i\}_{i \in \{1, \dots, n\}}$  defined by  $A_i = f^{-1}(\{\beta(i)\})$  we have

1.  $X = \bigsqcup_{i \in \{1, \dots, n\}} A_i$  [ $X$  is the pairwise disjoint union of  $A_i$ 's]
2.  $\forall x \in X$  we have that  $f(x) = \sum_{i=1}^n \beta(i) \cdot \chi_{A_i}(x)$

Define  $\{a_i\}_{i \in \{1, \dots, n\}}$  by  $a_i = \beta(i)$  then by (1) we have  $\forall x \in X$  that  $f(x) = \sum_{i=1}^n a_i \cdot \chi_{A_i}(x)$  proving that

$$\sum_{i=1}^n a_i \cdot \chi_{A_i}$$

Using 19.161 we have  $\forall i \in \{1, \dots, n\}$  that  $A_i = f^{-1}(\{\beta(i)\}) \in \mathcal{A}$  so that

$$\{A_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{A}$$

Finally if  $f \in \mathcal{S}^+[X, \mathcal{A}]$  we have that  $f(X) \subseteq [0, \infty[$ , so if  $i \in \mathbb{N}_0$  then  $a_i = \beta(i) \in f(X) \subseteq [0, \infty[$  proving that

$$\{a_i\}_{i \in \{1, \dots, n\}} \subseteq [0, \infty[$$

□

We have also the opposite of the previous proposition.

**Proposition 19.191.** Let  $\langle X, \mathcal{A} \rangle$  be a measurable space,  $n \in \mathbb{N}_0$ ,  $\{A_i\}_{i \in \{1, \dots, n\}} \subseteq \mathcal{A}$  a pairwise disjoint family and  $\{a_i\}_{i \in \{1, \dots, n\}} \subseteq [0, \infty[$  then for  $f = \sum_{i=1}^n a_i \cdot \chi_{A_i}$  we have

1. If  $i \in \{1, \dots, n\}$  then  $\forall x \in A_i$  we have  $f(x) = a_i$
2.  $\forall x \in X \setminus (\bigcup_{i \in \{1, \dots, n\}} A_i)$  we have  $f(x) = 0$
3.  $f \in \mathcal{S}^+[X, \mathcal{A}]$

**Proof.**

1. Let  $i \in \{1, \dots, n\}$  then if  $x \in A_i$  we have that  $\chi_{A_i}(x) = 1$  and  $\forall j \in \{1, \dots, n\} \setminus \{i\}$   $x \notin A_j \Rightarrow \chi_{A_j}(x) = 0$ . Hence

$$\begin{aligned} f(x) &= \sum_{j \in \{1, \dots, n\}} a_j \cdot \chi_{A_j}(x) \\ &= \sum_{j \in \{i\}} a_j \cdot \chi_{A_j}(x) + \sum_{j \in \{1, \dots, n\} \setminus \{i\}} a_j \cdot \chi_{A_j}(x) \\ &= a_i + 0 \\ &= a_i \end{aligned}$$

2. If  $x \in X \setminus (\bigcup_{i \in \{1, \dots, n\}} A_i)$  then  $\forall i \in \{1, \dots, n\}$   $x \notin A_i \Rightarrow \chi_{A_i}(x) = 0$  so that  $f(x) = \sum_{i=1}^n a_i \cdot \chi_{A_i}(x) = \sum_{i=1}^n 0 = 0$ .

3. Let  $x \in X$  then we have either  $x \in X \setminus (\bigcup_{i \in \{1, \dots, n\}} A_i)$  giving  $f(x) = 0$  or  $x \in \bigcup_{i \in \{1, \dots, n\}} A_i$  so that  $\exists i \in \{1, \dots, n\}$  with  $x \in A_i \Rightarrow f(x) = a_i$ . Hence  $f(X) \subseteq \{a_i | i \in \{1, \dots, n\}\} \cup \{0\}$  a finite set proving that

$$f \text{ is a simple function and } f(X) \subseteq [0, \dots, \infty[ \tag{19.366}$$

Let  $y \in f(X)$ . For  $y$  we have then either

**y = 0.** Define  $J = \{i \in \{1, \dots, n\} | a_i = 0\}$ . Let  $x \in f^{-1}(\{y\}) = f^{-1}(\{0\})$  then we have either

$x \in \bigcup_{i \in \{1, \dots, n\}} A_i$ . then there exists a  $i \in A_i$  such that  $0 = f(x) = a_i$  hence  $i \in J$

proving that  $x \in \bigcup_{i \in J} A_i \subseteq (\bigcup_{i \in J} A_i) \cup (X \setminus (\bigcup_{i \in \{1, \dots, n\}} A_i))$

$x \in X \setminus (\bigcup_{i \in \{1, \dots, n\}} A_i)$ . then  $x \in (\bigcup_{i \in J} A_i) \cup (X \setminus (\bigcup_{i \in \{1, \dots, n\}} A_i))$   
proving

$$f^{-1}(\{y\}) \subseteq \left( \bigcup_{i \in J} A_i \right) \cup \left( X \setminus \left( \bigcup_{i \in \{1, \dots, n\}} A_i \right) \right) \quad (19.367)$$

For the opposite including let  $x \in (\bigcup_{i \in J} A_i) \cup (X \setminus (\bigcup_{i \in \{1, \dots, n\}} A_i))$  then either

$x \in \bigcup_{i \in J} A_i$ . then  $\exists i \in J$  such that  $x \in A_i$  so that  $f(x) = a_i = 0$  or  $x \in f^{-1}(\{y\})$

$x \in X \setminus (\bigcup_{i \in \{1, \dots, n\}} A_i)$ . then using (2) we have  $f(x) = 0$  or  $x \in f^{-1}(\{y\})$

which proves that  $(\bigcup_{i \in J} A_i) \cup (X \setminus (\bigcup_{i \in \{1, \dots, n\}} A_i)) \subseteq f^{-1}(\{y\})$ . Combining this with 19.367 proves that  $f^{-1}(\{y\}) = (\bigcup_{i \in J} A_i) \cup (X \setminus (\bigcup_{i \in \{1, \dots, n\}} A_i))$ . As  $\{A_i\}_{i \in \{1, \dots, n\}} \subseteq \mathcal{A}$  we have that  $(\bigcup_{i \in J} A_i) \cup (X \setminus (\bigcup_{i \in \{1, \dots, n\}} A_i)) \in \mathcal{A}$ , hence

$$f^{-1}(\{y\}) \in \mathcal{A}$$

$y \neq 0$ . Let  $K = \{i \in J \mid a_i = y\}$  and define  $B = \bigcup_{i \in K} A_i$  then as  $\{A_i\}_{i \in \{1, \dots, n\}} \subseteq \mathcal{A}$  we have that

$$B \in \mathcal{A} \quad (19.368)$$

If  $x \in B$  then there exists a  $i \in K$  such that  $x \in A_i$  hence  $f(x) = a_i = y$  proving that

$$B \subseteq f^{-1}(\{y\}) \quad (19.369)$$

For the opposite inclusion let  $x \in f^{-1}(\{y\})$  then  $f(x) = y$ . Assume that  $x \notin B$  then either

$x \in \bigcup_{i \in \{1, \dots, n\} \setminus K} A_i$ . so that  $\exists i \in \{1, \dots, n\} \setminus K$  such that  $x \in A_i$  hence  $f(x) = a_i \neq y$  a contradiction.

$x \in X \setminus (\bigcup_{i \in \{1, \dots, n\}} A_i)$ . then  $f(x) = 0 \neq y$  a contradiction.

As in all cases we reach a contradiction it follows that  $x \in B$  proving that  $f^{-1}(\{y\}) \subseteq B$  which combined with 19.369 proves

$$f^{-1}(\{y\}) = B \in \mathcal{A}$$

So its proved that

$$\forall y \in f(X) \text{ we have } f^{-1}(\{y\}) \in \mathcal{S} \quad (19.370)$$

As  $f$  is a simple real valued non negative function [see 19.366],  $\mathcal{A}$ -measurable [see 19.370 and 19.161] it follows that

$$f \in \mathcal{S}^+[X, \mathcal{A}] \quad \square$$

**Example 19.192.** Let  $\langle X, \mathcal{A} \rangle$  be a measurable function,  $a \in [0, \infty[$  and  $C_a \in \mathbb{R}^X$  defined by  $C_a(x) = a$  then  $C_a \in \mathcal{S}^+[X, \mathcal{A}]$ .

**Proof.** Define  $\{a_i\}_{i \in \{1, \dots, 1\}}$  by  $a_i = a$  and  $\{A_i\}_{i \in \{1, \dots, 1\}}$  by  $A_i = X$  then as  $X \in \mathcal{A}$  we have by the previous proposition [see 19.191] that  $\sum_{i=1}^1 a_i \cdot \mathcal{X}_{A_i} \in \mathcal{S}^+[X, \mathcal{A}]$ . Finally if  $x \in X$  then  $(\sum_{i=1}^1 a_i \cdot \mathcal{X}_{A_i})(x) = \sum_{i=1}^1 a_i \cdot \mathcal{X}_{A_i}(x) = a = C_a(x)$  proving that  $C_a = \sum_{i=1}^1 a_i \cdot \mathcal{X}_{A_i} \in \mathcal{S}^+[X, \mathcal{A}]$ .  $\square$

**Proposition 19.193.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space and  $f \in \mathcal{S}^+[X, \mathcal{A}]$ . Assume that  $\exists n, m \in \mathbb{N}_0$ , pairwise disjoint families of measurable sets  $\{A_i\}_{i \in \{1, \dots, n\}} \subseteq \mathcal{A}$ ,  $\{B_i\}_{i \in \{1, \dots, m\}} \subseteq \mathcal{A}$  and  $\{a_i\}_{i \in \{1, \dots, n\}}$ ,  $\{b_i\}_{i \in \{1, \dots, m\}}$  such that

$$\sum_{i=1}^n a_i \cdot \mathcal{X}_{A_i} = f = \sum_{i=1}^m b_i \cdot \mathcal{X}_{B_i}$$

then

$$\sum_{i=1}^n a_i \cdot \mu(A_i) = \sum_{i=1}^m b_i \cdot \mu(B_i)$$

**Note 19.194.** The condition that  $f \in S^+[X, \mathcal{A}]$  is needed because this ensures that  $\{a_i\}_{i \in \{1, \dots, n\}}$ ,  $\{b_i\}_{i \in \{1, \dots, m\}}$  are non negative so that  $\sum_{i=1}^n a_i \cdot \mu(A_i)$  and  $\sum_{i=1}^m b_i \cdot \mu(B_i)$  are well defined [as  $\mu(A_i), \mu(B_i)$  can be  $\infty$ ].

**Proof.** First define  $I = \{i \in \{1, \dots, n\} \mid A_i \neq \emptyset \wedge a_i \neq 0\}$  and  $J = \{i \in \{1, \dots, m\} \mid B_i \neq \emptyset \wedge b_i \neq 0\}$  then we have for  $x \in X$  that

$$\begin{aligned} \left( \sum_{i=1}^n a_i \cdot \mathcal{X}_{A_i} \right)(x) &= \sum_{i=1}^n a_i \cdot \mathcal{X}_{A_i}(x) \\ &= \sum_{i \in \{1, \dots, n\}} a_i \cdot \mathcal{X}_{A_i}(x) \\ &= \sum_{i \in I} a_i \cdot \mathcal{X}_{A_i}(x) + \sum_{i \in \{1, \dots, n\} \setminus I} a_i \cdot \mathcal{X}_{A_i}(x) \\ &= \sum_{i \in I} a_i \cdot \mathcal{X}_{A_i}(x) + 0 \\ &= \sum_{i \in I} a_i \cdot \mathcal{X}_{A_i}(x) \end{aligned}$$

and

$$\begin{aligned} \left( \sum_{i=1}^m b_i \cdot \mathcal{X}_{B_i} \right)(x) &= \sum_{i=1}^m b_i \cdot \mathcal{X}_{B_i}(x) \\ &= \sum_{i \in \{1, \dots, m\}} b_i \cdot \mathcal{X}_{B_i}(x) \\ &= \sum_{i \in J} b_i \cdot \mathcal{X}_{B_i}(x) + \sum_{i \in \{1, \dots, m\} \setminus J} b_i \cdot \mathcal{X}_{B_i}(x) \\ &= \sum_{i \in J} b_i \cdot \mathcal{X}_{B_i}(x) + 0 \\ &= \sum_{i \in J} b_i \cdot \mathcal{X}_{B_i}(x) \end{aligned}$$

proving that

$$\sum_{i \in I} a_i \cdot \mathcal{X}_{A_i} = f = \sum_{i \in J} b_i \cdot \mathcal{X}_{B_i} \quad (19.371)$$

Let  $x \in \bigcup_{i \in I} A_i$  then there exists a  $i \in I$  such that  $x \in A_i$  and as  $\{A_i\}_{i \in I}$  is pairwise disjoint we have that  $f(x) = \sum_{i \in I} a_i \cdot \mathcal{X}_{A_i}(x) \stackrel{19.191}{=} a_i \neq 0$ . Assume that  $x \notin \bigcup_{i \in J} B_i$  then we have  $f(x) = 0$  a contradiction so we must have that  $x \in \bigcup_{i \in I} B_i$ . This proves that  $\bigcup_{i \in I} A_i \subseteq \bigcup_{i \in J} B_i$ . In a similar way we have that  $\bigcup_{i \in J} B_i \subseteq \bigcup_{i \in I} A_i$ . Hence we have

$$\bigsqcup_{i \in I} A_i = \bigsqcup_{i \in J} B_i \quad (19.372)$$

Let  $i \in \{1, \dots, n\}$  then

$$A_i = A_i \bigcap \left( \bigsqcup_{j \in I} A_j \right) \stackrel{19.372}{=} A_i \bigcap \left( \bigsqcup_{j \in J} B_j \right) = \bigsqcup_{j \in J} (A_i \bigcap B_j) \quad (19.373)$$

Let  $i \in \{1, \dots, m\}$  then

$$B_i = B_i \bigcap \left( \bigsqcup_{j \in I} B_j \right) \stackrel{19.372}{=} B_i \bigcap \left( \bigsqcup_{j \in I} A_j \right) = \bigsqcup_{j \in I} (A_j \bigcap B_i) \quad (19.374)$$

Define  $K = \{(i, j) \in \{1, \dots, n\} \times \{1, \dots, m\} \mid A_i \cap B_j \neq \emptyset\}$ . Then if  $(i, j) \in K$  we have  $A_i \cap B_j \neq \emptyset$  hence  $\exists x \in A_i \cap B_j$ , So

$$a_i \stackrel{19.191 \wedge x \in A_i}{=} \sum_{j \in I} a_j \cdot \mathcal{X}_{A_i}(x) = f(x) = \sum_{j \in I} b_j \cdot \mathcal{X}_{B_j}(x) \stackrel{19.191 \wedge x \in B_j}{=} b_j$$

proving that

$$\forall (i, j) \in K \text{ we have } a_i = b_j \quad (19.375)$$

Now

$$\begin{aligned}
\sum_{i=1}^n a_i \cdot \mu(A_i) &= \sum_{i \in I} a_i \cdot \mu(A_i) + \sum_{i \in \{1, \dots, n\} \setminus I} a_i \cdot \mu(A_i) \\
&= \sum_{i \in I} a_i \cdot \mu(A_i) \\
&\stackrel{19.373}{=} \sum_{i \in I} a_i \cdot \mu\left(\bigsqcup_{j \in J} (A_i \cap B_j)\right) \\
&= \sum_{i \in I} a_i \cdot \left(\sum_{j \in J} \mu(A_i \cap B_j)\right) \\
&= \sum_{i \in I} \left(\sum_{j \in J} a_i \cdot \mu(A_i \cap B_j)\right) \\
&\stackrel{10.47}{=} \sum_{(i, j) \in I \times J} a_i \cdot \mu(A_i \cap B_j) \\
&= \sum_{(i, j) \in (I \times J) \setminus K} a_i \cdot \mu(A_i \cap B_j) + \sum_{(i, j) \in K} a_i \cdot \mu(A_i \cap B_j) \\
&\stackrel{(i, j) \in (I \times J) \setminus K \Rightarrow A_i \cap B_j = \emptyset}{=} \sum_{(i, j) \in K} a_i \cdot \mu(A_i \cap B_j) \\
&\stackrel{19.375}{=} \sum_{(i, j) \in K} b_j \cdot \mu(A_i \cap B_j) \\
&\stackrel{(i, j) \in (I \times J) \setminus K \Rightarrow A_i \cap B_j = \emptyset}{=} \sum_{(i, j) \in (I \times J) \setminus K} b_j \cdot \mu(A_i \cap B_j) + \sum_{(i, j) \in K} b_j \cdot \mu(A_i \cap B_j) \\
&= \sum_{(i, j) \in I \times J} b_j \cdot \mu(A_i \cap B_j) \\
&\stackrel{10.47}{=} \sum_{j \in J} \left(\sum_{i \in I} b_j \cdot \mu(A_i \cap B_j)\right) \\
&= \sum_{j \in J} b_j \cdot \left(\sum_{i \in I} \mu(A_i \cap B_j)\right) \\
&= \sum_{j \in J} b_j \cdot \mu\left(\bigsqcup_{i \in I} A_i \cap B_j\right) \\
&\stackrel{19.374}{=} \sum_{j \in J} b_j \cdot \mu(B_j) \\
&= \sum_{j \in J} b_j \cdot \mu(B_j) + \sum_{j \in \{1, \dots, m\} \setminus J} b_j \cdot \mu(B_j) \\
&= \sum_{j=1}^m b_j \cdot \mu(B_j)
\end{aligned}$$

proving that

$$\sum_{i=1}^n a_i \cdot \mu(A_i) = \sum_{i=1}^m b_i \cdot \mu(B_i) \quad \square$$

The two above propositions 19.190 and 19.193 ensures that the following definition is well defined (existence and uniqueness).

**Definition 19.195.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space then the integral  $\int^S: \mathcal{S}^+[X, \mathcal{A}] \rightarrow [0, \infty]$  is defined by  $f \rightarrow \int^S f d\mu$  where

$$\int^S f d\mu = \sum_{i=1}^n a_i \cdot \mu(A_i)$$

where  $n \in \mathbb{N}_0$ ,  $\{A_i\}_{i \in \{1, \dots, n\}} \subseteq \mathcal{A}$  pairwise disjoint and  $\{a_i\}_{i \in \{1, \dots, n\}} \subseteq [0, \infty[$  such that

$$f = \sum_{i=1}^n a_i \cdot \chi_{A_i}$$

To compare two simple functions we need the following lemma.

**Lemma 19.196.** Let  $\langle X, \mathcal{A} \rangle$  be a measure space and  $f, g \in \mathcal{S}^+[X, \mathcal{A}]$  then there exists a  $n \in \mathbb{N}_0$ ,  $\{C_i\}_{i \in \{1, \dots, n\}} \subseteq \mathcal{A}$  of pairwise disjoint measurable sets and  $\{f_i\}_{i \in \{1, \dots, n\}} \subseteq [0, \infty[$ ,  $\{g_i\}_{i \in \{1, \dots, n\}} \subseteq [0, \infty[$  such that

$$\begin{aligned} X &= \bigsqcup_{i \in \{1, \dots, n\}} C_i \\ f &= \sum_{i=1}^n f_i \cdot \chi_{C_i} \\ g &= \sum_{i=1}^n g_i \cdot \chi_{C_i} \end{aligned}$$

**Proof.** As  $f, g \in \mathcal{S}^+[X, \mathcal{A}]$  we have using 19.190 the existance of two pairwsise disjoint measurable sets  $\{A'_i\}_{i \in \{1, \dots, n'\}} \subseteq \mathcal{A}$ ,  $\{B'_i\}_{i \in \{1, \dots, m'\}} \subseteq \mathcal{A}$  and two families of non negative numbers  $\{a'_i\}_{i \in \{1, \dots, n'\}} \subseteq [0, \infty[$ ,  $\{b'_i\}_{i \in \{1, \dots, m'\}} \subseteq [0, \infty[$  such that

$$\begin{aligned} f &= \sum_{i=1}^{n'} a'_i \cdot \chi_{A'_i} \\ g &= \sum_{i=1}^{m'} b'_i \cdot \chi_{B'_i} \end{aligned}$$

Define now  $\{A_i\}_{i \in \{1, \dots, n'+1\}} \subseteq \mathcal{A}$ ,  $\{B_i\}_{i \in \{1, \dots, m'+1\}} \subseteq \mathcal{A}$  and  $\{a_i\}_{i \in \{1, \dots, n'+1\}} \subseteq [0, \infty[$ ,  $\{b_i\}_{i \in \{1, \dots, m'+1\}} \subseteq [0, \infty[$  as follows:

$$\begin{aligned} A_i &= \begin{cases} X \setminus \bigcup_{i \in \{1, \dots, n'\}} A'_i & \text{if } i = n'+1 \\ A'_i & \text{if } i \in \{1, \dots, n'\} \end{cases} \\ B_i &= \begin{cases} X \setminus \bigcup_{i \in \{1, \dots, m'\}} B'_i & \text{if } i = m'+1 \\ B'_i & \text{if } i \in \{1, \dots, m'\} \end{cases} \\ a_i &= \begin{cases} 0 & \text{if } i = n'+1 \\ a'_i & \text{if } i \in \{1, \dots, n'\} \end{cases} \\ b_i &= \begin{cases} 0 & \text{if } i = m'+1 \\ b'_i & \text{if } i \in \{1, \dots, m'\} \end{cases} \end{aligned}$$

then clearly  $\{A_i\}_{i \in \{1, \dots, n'+1\}}$ ,  $\{B_i\}_{i \in \{1, \dots, m'+1\}}$  are pairwise disjoint and

$$X = \bigsqcup_{i \in \{1, \dots, n'+1\}} A_i = \bigsqcup_{i \in \{1, \dots, m'+1\}} B_i \quad (19.376)$$

From the above we have then that

$$\forall i \in \{1, \dots, n'+1\} A_i = A_i \cap X = \bigsqcup_{j \in \{1, \dots, m'+1\}} (A_i \cap B_j) \quad (19.377)$$

$$\forall i \in \{1, \dots, m'+1\} B_i = B_i \cap X = \bigsqcup_{j \in \{1, \dots, n'+1\}} (A_j \cap B_i) \quad (19.378)$$

so that

$$X = \bigsqcup_{i=1}^{n'+1} \left( \bigsqcup_{j=1}^{m'+1} (A_i \cap B_j) \right) = \bigsqcup_{i=1}^{m'+1} \left( \bigsqcup_{j=1}^{n'+1} (A_j \cap B_i) \right) \quad (19.379)$$

Using 5.44 we have that  $\{1, \dots, n' + 1\} \times \{1, \dots, m' + 1\}$  is finite and there exists a bijection

$$\beta: \{1, \dots, (n' + 1) \cdot (m' + 1)\} \rightarrow \{1, \dots, n' + 1\} \cdot \{1, \dots, m' + 1\}$$

Using the above bijection take  $n = (n' + 1) \cdot (m' + 1)$  and define  $\{C_i\}_{i \in \{1, \dots, n\}} \subseteq \mathcal{A}$ ,  $\{f_i\}_{i \in \{1, \dots, n\}}$  and  $\{g_i\}_{i \in \{1, \dots, n\}}$  by

$$C_i = A_{\beta(i)_1} \cap B_{\beta(i)_2}, \quad f_i = a_{\beta(i)_1}, \quad g_i = b_{\beta(i)_2}$$

Take  $x \in X$  then by 19.379 there exists a  $i \in \{1, \dots, n' + 1\}$  such that  $x \in \bigcup_{j=1}^{m'+1} (A_i \cap B_j)$ , hence there exists a  $j \in \{1, \dots, m' + 1\}$  such that  $x \in A_i \cap B_j$ . As  $\beta$  is a bijection there exists a  $k \in \{1, \dots, n\}$  such that  $\beta(k) = (i, j)$ , hence  $x \in A_{\beta(k)_1} \cap B_{\beta(k)_2} = C_k$ . This proves that  $X \subseteq \bigcup_{i \in \{1, \dots, n\}} C_i$  which as trivially  $\bigcup_{i \in \{1, \dots, n\}} C_i \subseteq X$  proves that

$$X = \bigcup_{i \in \{1, \dots, n\}} C_i. \quad (19.380)$$

Further if  $k, l \in \{1, \dots, n\}$  with  $k \neq l$  then as  $\beta$  is a bijection we have that  $\beta(k) \neq \beta(l)$ . If now  $x \in C_k \cap C_l$  then  $x \in (A_{\beta(k)_1} \cap B_{\beta(k)_2}) \cap (A_{\beta(l)_1} \cap B_{\beta(l)_2})$  so that  $x \in A_{\beta(k)_1} \cap B_{\beta(l)_2} \wedge x \in A_{\beta(l)_1} \cap B_{\beta(l)_2}$ , as  $\{A_i\}_{i \in \{1, \dots, n'+1\}}$ ,  $\{B_i\}_{i \in \{1, \dots, m'+1\}}$  are pairwise disjoint we have that  $\beta(k)_1 = \beta(l)_1 \wedge \beta(k)_2 = \beta(l)_2$ , leading to the contradiction  $\beta(k) = \beta(l)$ . Hence we have that

$$\{C_i\}_{i \in \{1, \dots, n\}} \subseteq \mathcal{A} \text{ is pairwise disjoint} \quad (19.381)$$

Let  $x \in X$  then by 19.380 there exists a  $k \in \{1, \dots, n\}$  such that  $x \in C_k$  and take  $\beta(k) = (i, j)$ , then  $C_k = A_i \cap B_j$ . For  $i$  we must consider the following cases:

**$i = n' + 1$ .** then  $x \in A_{n'+1} = X \setminus (\bigcup_{r \in \{1, \dots, n'\}} A'_r)$  and using 19.191  $f(x) = 0$ . Further using 19.191 again, it follows from  $x \in C_k$  that  $(\sum_{r=1}^n f_r \cdot \chi_{C_r})(x) = f_k = a_{\beta(k)_1} = a_{n'+1} = 0$  proving that

$$\left( \sum_{r=1}^n f_r \cdot \chi_{C_r} \right)(x) = f(x)$$

**$i \in \{1, \dots, n'\}$ .** then  $x \in A_i$ , so using 19.191 we have  $f(x) = a'_i$  and  $(\sum_{r=1}^n f_r \cdot \chi_{C_r})(x) \underset{x \in C_k}{=} f_k = a_{\beta(k)_1} = a_i = a'_i$ , Hence

$$\left( \sum_{r=1}^n f_r \cdot \chi_{C_r} \right)(x) = f(x)$$

So we conclude that

$$f = \sum_{r=1}^n f_r \cdot \chi_{C_r} \quad (19.382)$$

For  $j$  we must consider the following cases:

**$j = m' + 1$ .** then  $x \in B_{m'+1} = X \setminus (\bigcup_{r \in \{1, \dots, n'\}} B'_r)$  so that using 19.191  $g(x) = 0$ . Further using 19.191 again, it follows from  $x \in C_k$  that  $(\sum_{r=1}^n g_r \cdot \chi_{C_r})(x) = g_k = b_{\beta(k)_2} = b_{m'+1} = 0$  proving that

$$\left( \sum_{r=1}^n g_r \cdot \chi_{C_r} \right)(x) = g(x)$$

**$j \in \{1, \dots, m'\}$ .** then  $x \in B_j$  so using 19.191 we have  $g(x) = b'_j$  and  $(\sum_{r=1}^n g_r \cdot \chi_{C_r})(x) \underset{x \in C_k}{=} g_k = b_{\beta(k)_2} = b_j = b'_j$ , Hence

$$\left( \sum_{r=1}^n g_r \cdot \chi_{C_r} \right)(x) = g(x)$$

So we conclude that

$$g = \sum_{r=1}^n g_r \cdot \chi_{C_r} \quad (19.383)$$

The lemma is then proved by 19.380, 19.381, 19.382 and 19.383.  $\square$

The integral of non negative real valued simple functions has the following properties.

**Proposition 19.197.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space then we have

1. If  $\alpha \in [0, \infty[$  and  $f \in \mathcal{S}^+[X, \mathcal{A}]$  then  $\alpha \cdot f \in \mathcal{S}^+[X, \mathcal{A}]$  and  $\int^{\mathcal{S}} (\alpha \cdot f) d\mu = \alpha \cdot \int^{\mathcal{S}} f d\mu$
2. If  $f, g \in \mathcal{S}^+[X, \mathcal{A}]$  then
  - a.  $f + g \in \mathcal{S}^+[X, \mathcal{A}]$  and  $\int^{\mathcal{S}} (f + g) d\mu = \int^{\mathcal{S}} f d\mu + \int^{\mathcal{S}} g d\mu$
  - b. if  $\forall x \in X f(x) \leq g(x)$  then  $\int^{\mathcal{S}} f d\mu \leq \int^{\mathcal{S}} g d\mu$
  - c.  $f \wedge g \in \mathcal{S}^+[X, \mathcal{A}]$ ,  $\int^{\mathcal{S}} (f \wedge g) d\mu \leq \int^{\mathcal{S}} f d\mu$  and  $\int^{\mathcal{S}} (f \wedge g) d\mu \leq \int^{\mathcal{S}} g d\mu$
3. If  $f \in \mathcal{S}^+[X, \mathcal{A}]$  is such that  $f = 0$   $\mu$ -a.e. then  $\int^{\mathcal{S}} f d\mu = 0$

**Proof.**

1. As  $f \in \mathcal{S}^+(X, \mathcal{A})$  there exists by 19.190 a  $n \in \mathbb{N}_0$ , a pairwise disjoint family  $\{A_i\}_{i \in \{1, \dots, n\}} \subseteq \mathcal{A}$  and a  $\{a_i\}_{i \in \{1, \dots, n\}} \subseteq [0, \infty[$  such that  $f = \sum_{i=1}^n a_i \cdot \mathcal{X}_{A_i}$ . Then  $\{\alpha \cdot a_i\}_{i \in \{1, \dots, n\}} \subseteq [0, \infty[$  and  $\forall x \in X$  we have

$$\begin{aligned} (\alpha \cdot f)(x) &= \alpha \cdot f(x) \\ &= \alpha \cdot \sum_{i=1}^n a_i \cdot \mathcal{X}_{A_i}(x) \\ &= \sum_{i=1}^n (\alpha \cdot a_i) \cdot \mathcal{X}_{A_i}(x) \\ &= \left( \sum_{i=1}^n (\alpha \cdot a_i) \cdot \mathcal{X}_{A_i} \right)(x) \end{aligned}$$

proving that

$$\alpha \cdot f = \sum_{i=1}^n (\alpha \cdot a_i) \cdot \mathcal{X}_{A_i}$$

Using 19.191 it follows then that

$$\alpha \cdot f \in \mathcal{S}^+[X, \mathcal{A}]$$

Finally we have

$$\begin{aligned} \int^{\mathcal{S}} \alpha \cdot f d\mu &= \sum_{i=1}^n (\alpha \cdot a_i) \cdot \mu(A_i) \\ &= a \cdot \sum_{i=1}^n a_i \cdot \mu(A_i) \\ &= \alpha \cdot \int^{\mathcal{S}} f d\mu \end{aligned}$$

2. As  $f, g \in \mathcal{S}^+[X, \mathcal{A}]$  we have, using the previous lemma [see 19.196], that there exists a pairwise disjoint family of measurable sets  $\{C_i\}_{i \in \{1, \dots, n\}}$  and families of non negative numbers  $\{f_i\}_{i \in \{1, \dots, n\}}$ ,  $\{g_i\}_{i \in \{1, \dots, n\}}$  such that

$$\begin{aligned} X &= \bigsqcup_{i=1}^n C_i \\ f &= \sum_{i=1}^n f_i \cdot \mathcal{X}_{C_i} \\ g &= \sum_{i=1}^n g_i \cdot \mathcal{X}_{C_i} \end{aligned}$$

- a. Consider now  $\{f_i + g_i\}_{i \in \{1, \dots, n\}} \subseteq [0, \infty]$  then using 19.191 we have that

$$\sum_{i=1}^n (f_i + g_i) \cdot \mathcal{X}_{C_i} \in \mathcal{S}^+[X, \mathcal{A}] \tag{19.384}$$

Further if  $x \in X$  then there exists a  $i \in \{1, \dots, n\}$  such that  $x \in C_i$ . Hence

$$\begin{aligned} \left( \sum_{i=1}^n (f_i + g_i) \cdot \mathcal{X}_{C_i} \right)(x) &\stackrel{19.191}{=} f_i + g_i \\ &\stackrel{19.191}{=} \left( \sum_{i=1}^n f_i \cdot \mathcal{X}_{C_i} \right)(x) + \left( \sum_{i=1}^n g_i \cdot \mathcal{X}_{C_i} \right)(x) \\ &= f(x) + g(x) \\ &\stackrel{\text{def}}{=} (f + g)(x) \end{aligned}$$

proving that

$$f + g = \sum_{i=1}^n (f_i + g_i) \cdot \mathcal{X}_{C_i} \in \mathcal{S}[X, \mathcal{A}] \quad (19.385)$$

Finally we have that

$$\begin{aligned} \int^{\mathcal{S}} (f + g) d\mu &\stackrel{19.385}{=} \sum_{i=1}^n (f_i + g_i) \cdot \mu(C_i) \\ &= \sum_{i=1}^n f_i \cdot \mu(C_i) + \sum_{i=1}^n g_i \cdot \mu(C_i) \\ &= \int^{\mathcal{S}} f d\mu + \int^{\mathcal{S}} g d\mu \end{aligned}$$

- b. Assume that  $\forall x \in X$  we have that  $f(x) \leq g(x)$ . Define  $I = \{i \in \{1, \dots, n\} \mid C_i \neq \emptyset\}$  then  $\forall i \in I$  there exists a  $x \in C_i$  so that

$$f_i \stackrel{19.191}{=} f(x) \leq g(x) \stackrel{19.191}{=} g_i. \quad (19.386)$$

Next

$$\begin{aligned} \int^{\mathcal{S}} f d\mu &= \sum_{i=1}^n f_i \cdot \mu(C_i) \\ &= \sum_{i \in \{1, \dots, n\} \setminus I} f_i \cdot \mu(C_i) + \sum_{i \in I} f_i \cdot \mu(C_i) \\ &= 0 + \sum_{i \in I} f_i \cdot \mu(C_i) \\ &\stackrel{19.386}{\leq} \sum_{i \in I} g_i \cdot \mu(C_i) \\ &= \sum_{i \in I} g_i \cdot \mu(C_i) + 0 \\ &= \sum_{i \in \{1, \dots, n\} \setminus I} g_i \cdot \mu(C_i) + \sum_{i \in I} g_i \cdot \mu(C_i) \\ &= \int^{\mathcal{S}} g d\mu \end{aligned}$$

- c. Take  $\{\max(f_i, g_i)\}_{i \in \mathbb{N}_0} \subseteq [0, \infty[$  then using 19.191 we have that

$$\sum_{i=1}^n \min(f_i, g_i) \cdot \mathcal{X}_{C_i} \in \mathcal{S}^+[X, \mathcal{A}] \quad (19.387)$$

Let  $x \in X$  then there exists a  $i \in \{1, \dots, n\}$  such that  $x \in C_i$  so that

$$\begin{aligned} (f \wedge g)(x) &= \min(f(x), g(x)) \\ &\stackrel{19.191}{=} \min(f_i, g_i)(x) \\ &\stackrel{19.191}{=} \left( \sum_{i=1}^n \min(f_i, g_i) \cdot \mathcal{X}_{C_i} \right)(x) \end{aligned}$$

proving that  $f \wedge g = \sum_{i=1}^n \min(f_i, g_i) \cdot \chi_{C_i}$  which by 19.387 proves that

$$f \wedge g \in \mathcal{S}^+[X, \mathcal{A}]$$

Further as  $\forall x \in X$  we have that  $\min(f(x), g(x)) \leq f(x), g(x)$  we have by (b) that

$$\int^S (f \wedge g) d\mu \leq \int^S f d\mu \text{ and } \int^S (f \wedge g) d\mu \leq \int^S g d\mu$$

3. As  $f = 0$   $\mu$ -a.e. we have

$$\exists N \in \mathcal{A} \text{ with } \mu(N) = 0 \text{ such that } \{x \in X \mid 0 \neq f(x)\} \subseteq N$$

As  $f \in \mathcal{S}^+[X, \mathcal{A}]$  there exists by 19.190 a  $n \in \mathbb{N}_0$ , a pairwise disjoint family  $\{A_i\}_{i \in \{1, \dots, n\}} \subseteq \mathcal{A}$  and a  $\{a_i\}_{i \in \{1, \dots, n\}} \subseteq [0, \infty[$  such that  $f = \sum_{i=1}^n a_i \cdot \chi_{A_i}$ . Let  $i \in \{1, \dots, n\}$  then we have for  $a_i$  either:

**$a_i = 0$ .** Then  $a_i \cdot \mu(A_i) = 0$

**$0 < a_i$ .** Then if  $x \in A_i$  we have  $f(x) \stackrel{19.191}{=} a_i > 0$  so that  $x \in \{x \in X \mid 0 \neq f(x)\} \subseteq N$  proving that  $A_i \subseteq N$  hence  $\mu(A_i) \leq \mu(N) = 0$  proving that  $a_i \cdot \mu(A_i) = 0$

So  $\forall i \in \{1, \dots, n\}$  we have that  $a_i \cdot \mu(A_i) = 0$  proving that

$$\int^S f d\mu = \sum_{i=1}^n a_i \cdot \mu(A_i) = 0$$

□

**Corollary 19.198.** Let  $\langle X, \mathcal{A}, \mu \rangle$  a measure space,  $n \in \mathbb{N}_0$  and  $\{f_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{S}^+[X, \mathcal{A}]$  then

$$\max(\{f_i \mid i \in \{1, \dots, n\}\}) \in \mathcal{S}^+[X, \mathcal{A}]$$

**Proof.** We proof this by induction so let

$$\mathcal{S} = \{n \in \mathbb{N}_0 \mid \text{If } \{f_i\}_{i \in \{1, \dots, n\}} \subseteq \mathcal{S}^+[X, \mathcal{A}] \text{ then } \max(\{f_i \mid i \in \{1, \dots, n\}\}) \in \mathcal{S}^+[X, \mathcal{A}]\}$$

then we have

**$1 \in \mathcal{S}$ .** If  $\{f_i\}_{i \in \{1, \dots, 1\}} \subseteq \mathcal{S}^+[X, \mathcal{A}]$  then  $\max(\{f_i \mid i \in \{1, \dots, n\}\}) = f_1 \in \mathcal{S}^+[X, \mathcal{A}]$  proving that  $1 \in \mathcal{S}$

**$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$ .** Let  $\{f_i\}_{i \in \{1, \dots, n+1\}} \subseteq \mathcal{S}^+[X, \mathcal{A}]$  then as  $n \in \mathcal{S}$  we have  $\max(\{f_i \mid i \in \{1, \dots, n\}\}) \in \mathcal{S}$  so that

$$\max(\{f_i \mid i \in \{1, \dots, n+1\}\}) \stackrel{19.167}{=} \max(\max(\{f_i \mid i \in \{1, \dots, n\}\}), f_{n+1}) \in \mathcal{S}[X, \mathcal{A}]$$

proving that  $n+1 \in \mathcal{S}$

□

**Proposition 19.199.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space,  $f \in \mathcal{S}^+[X, \mathcal{A}]$  and  $\{f_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{S}^+[X, \mathcal{A}]$  such that  $\forall x \in X$

$$\forall i \in \mathbb{N}_0 \quad f_i(x) \leq f_{i+1}(x)$$

$$f(x) = \lim_{i \rightarrow \infty} f_i(x)$$

then

$$\int^S f d\mu = \lim_{n \rightarrow \infty} \int^S f_n d\mu$$

**Proof.** Let  $x \in X$  then  $\forall i \in \mathbb{N}_0$  we have  $f_i(x) \leq f_{i+1}(x)$  so that by the previous proposition 19.197 it follows that  $\int^S f_i d\mu \leq \int^S f_{i+1} d\mu$ . Hence

$$\forall x \in X \quad \{f_i(x)\}_{i \in \mathbb{N}_0} \text{ and } \left\{ \int^S f_i d\mu \right\}_{i \in \mathbb{N}_0} \text{ are increasing sequences} \quad (19.388)$$

Applying then 17.71 on the above we have that

$$\lim_{i \rightarrow \infty} \int^S f_i d\mu = \sup \left( \left\{ \int^S f_i d\mu \mid i \in \mathbb{N}_0 \right\} \right) \geq 0 \quad (19.389)$$

and

$$\forall x \in X \text{ we have } f(x) = \lim_{i \rightarrow \infty} f_i(x) = \sup (\{f_i(x) \mid i \in \mathbb{N}_0\}) \quad (19.390)$$

From the above we have given  $x \in X$  that  $\forall i \in \mathbb{N}_0 \ f_i(x) \leq f(x)$ , so using 19.197 it follows that  $\int^S f_i d\mu \leq \int^S f d\mu$ , hence  $\sup (\{ \int^S f_i d\mu \mid i \in \mathbb{N}_0 \}) \leq \int^S f d\mu$ . Combining this with 19.389 gives

$$0 \leq \lim_{i \rightarrow \infty} \int^S f_i d\mu \leq \int^S f d\mu \quad (19.391)$$

For the opposite inequality, let  $\varepsilon \in \mathbb{R}$  be such that  $0 < \varepsilon < 1$  so that  $0 < 1 - \varepsilon$ . As  $f \in \mathcal{S}^+[X, \mathcal{A}]$  there exists by 19.190 a  $n \in \mathbb{N}_0$ , a pairwise disjoint family of measurable sets  $\{A_i\}_{i \in \{1, \dots, n\}}$  and a family of non negative real numbers  $\{a_i\}_{i \in \{1, \dots, n\}}$  so that  $f = \sum_{i=1}^n a_i \cdot \chi_{A_i}$ . Then we have two cases to consider for  $\{a_i\}_{i \in \{1, \dots, n\}}$ :

$\forall i \in \{1, \dots, n\} a_i = 0$ . then  $\int^S f d\mu = \sum_{i=1}^n a_i \cdot \mu(A_i) = 0$  so that using 19.391 we have  $\lim_{i \rightarrow \infty} \int^S f_i d\mu = 0$  giving

$$\lim_{i \rightarrow \infty} \int^S f_i d\mu = \int^S f d\mu = 0$$

$\exists i \in \{1, \dots, n\}$  with  $a_i \neq 0$ . then  $I = \{i \in \{1, \dots, n\} \mid a_i \neq 0\} \neq \emptyset$  and as it is finite we have a  $m \in \mathbb{N}_0$  and a bijection  $\beta: \{1, \dots, m\} \rightarrow I$ . Let  $x \in X$  then

$$\begin{aligned} f(x) &= \sum_{i \in \{1, \dots, n\}} a_i \cdot \chi_{A_i}(x) \\ &= \sum_{i \in \{1, \dots, n\} \setminus I} a_i \cdot \chi_{A_i}(x) + \sum_{i \in I} a_i \cdot \chi_{A_i}(x) \\ &= \sum_{i \in I} a_i \cdot \chi_{A_i}(x) \\ &\stackrel{10.43}{=} \sum_{i \in \{1, \dots, m\}} a_{\beta(i)} \cdot \chi_{A_{\beta(i)}}(x) \\ &= \sum_{i=1}^m a_{\beta(i)} \cdot \chi_{A_{\beta(i)}}(x) \end{aligned}$$

proving if we define  $b_i = a_{\beta(i)}$  and  $B_i = A_{\beta(i)}$  that

$$f = \sum_{i=1}^m b_i \cdot \chi_{B_i} \text{ and } \forall i \in \{1, \dots, m\} \ b_i > 0 \quad (19.392)$$

Take  $\varepsilon \in \mathbb{R}$  such that  $0 < \varepsilon < 1$  so that  $0 < 1 - \varepsilon$ . Let  $i \in \{1, \dots, m\}$  and define  $\{B_{i,j}\}_{j \in \mathbb{N}_0}$  by

$$B_{i,j} = \{x \in B_i \mid f_j(x) \geq (1 - \varepsilon) \cdot b_i\} = B_i \cap \{f_j \geq (1 - \varepsilon) \cdot b_i\} \in \mathcal{A} \text{ [by 19.153]} \quad (19.393)$$

Next if  $x \in B_i$  then  $\sup (\{f_j(x) \mid j \in \mathbb{N}_0\}) \stackrel{19.390}{=} f(x) \stackrel{19.191, 19.392}{=} b_i > (1 - \varepsilon) \cdot b_i$  [as  $0 < b_i \wedge 0 < 1 - \varepsilon < 1$ ], hence using the definition of a supremum there exists a  $k \in \mathbb{N}_0$  such that  $f_k(x) > (1 - \varepsilon) \cdot b_i$  or  $x \in B_{i,k} \Rightarrow B_i \subseteq \bigsqcup_{k \in \mathbb{N}_0} B_{i,k}$ . This together with the fact that  $B_{i,k} \subseteq B_i$  gives

$$\forall i \in \{1, \dots, m\} \text{ we have that } B_i = \bigsqcup_{k \in \mathbb{N}_0} B_{i,k} \quad (19.394)$$

Let  $i \in \{1, \dots, m\}$  and  $k \in \mathbb{N}_0$  then if  $x \in B_{i,k}$  we have that

$$f_k(x) \geq (1 - \varepsilon) \cdot b_i \underset{\{f_i\}_{i \in \mathbb{N}_0} \text{ is increasing}}{\Rightarrow} f_{k+1}(x) \geq (1 - \varepsilon) \cdot b_i$$

proving that  $x \in B_{i,k+1}$ , hence

$$\forall k \in \mathbb{N}_0 \ B_{i,k} \subseteq B_{i,k+1} \quad (19.395)$$

The conclusion 19.395 allows us to use 19.67 (1) giving

$$\mu(B_i) \underset{19.394}{=} \mu\left(\bigsqcup_{k \in \mathbb{N}_0} B_{i,k}\right) \underset{19.67}{=} \lim_{k \rightarrow \infty} \mu(B_{i,k}) = \sup(\{\mu(B_{i,k}) | k \in \mathbb{N}_0\}) \quad (19.396)$$

Given  $k \in \mathbb{N}_0$  define  $g_k$  by

$$g_k = \sum_{i=1}^m (1 - \varepsilon) \cdot b_i \cdot \chi_{B_{i,k}} \quad (19.397)$$

Using 19.191 together with the fact that  $B_{i,k} \in \mathcal{A}$  [see 19.393] proves

$$\{g_k\}_{k \in \mathbb{N}_0} \subseteq \mathcal{S}^+[X, \mathcal{A}] \quad (19.398)$$

Further if  $x \in X$  and  $k \in \mathbb{N}_0$  then either:

$x \in \bigcup_{i \in \{1, \dots, m\}} B_{i,k}$ . Then there exists a  $i \in \{1, \dots, m\}$  such that  $x \in B_{i,k}$  so that  $(1 - \varepsilon) \cdot b_i < f_k(x)$ , which as  $g_k(x) \underset{19.397, 19.191}{=} (1 - \varepsilon) \cdot b_i$  proves

$$g_k(x) < f_k(x)$$

Further as  $B_{i,k} \subseteq B_{i,k+1}$  we have  $x \in B_{i,k+1} \underset{19.397, 19.191}{\Rightarrow} g_{k+1}(x) = (1 - \varepsilon) \cdot b_i = g_k(x)$  hence

$$g_k(x) \leq g_{k+1}(x)$$

$x \in X \setminus (\bigcup_{i \in \{1, \dots, m\}} B_{i,k})$ . then

$$g_k(x) \underset{19.397, 19.191}{=} 0 \leq f_k(x)$$

and

$$g_k(x) = 0 \leq g_{k+1}(x)$$

So we have

$$\forall x \in X, \forall k \in \mathbb{N}_0 \ g_k(x) \leq f_k(x) \text{ and } g_k(x) \leq g_{k+1}(x) \quad (19.399)$$

Using the above we have by 17.71 that  $\lim_{k \rightarrow \infty} \int^{\mathcal{S}} g_k d\mu$  exists. Let's calculate this limit

$$\begin{aligned} \lim_{k \rightarrow \infty} \int^{\mathcal{S}} g_k d\mu &\underset{\text{def}}{=} \lim_{k \rightarrow \infty} \left( \sum_{i=1}^m (1 - \varepsilon) \cdot b_i \cdot \mu(B_{k,i}) \right) \\ &\underset{17.90}{=} \sum_{i=1}^m \lim_{k \rightarrow \infty} ((1 - \varepsilon) \cdot b_i \cdot \mu(B_{k,i})) \\ &\underset{17.77}{=} \sum_{i=1}^m (1 - \varepsilon) \cdot b_i \cdot \lim_{k \rightarrow \infty} \mu(B_{k,i}) \\ &\underset{19.396}{=} \sum_{i=1}^m (1 - \varepsilon) \cdot b_i \cdot \mu(B_i) \\ &= (1 - \varepsilon) \cdot \sum_{i=1}^m b_i \cdot \mu(B_i) \\ &\underset{19.392}{=} (1 - \varepsilon) \cdot \int^{\mathcal{S}} f d\mu \end{aligned} \quad (19.400)$$

By 19.399 we have using 19.197 that  $\int^{\mathcal{S}} g_k d\mu \leq \int^{\mathcal{S}} f_k d\mu$ , applying then 17.77 gives  $\lim_{k \rightarrow \infty} \int^{\mathcal{S}} g_k d\mu \leq \lim_{k \rightarrow \infty} \int^{\mathcal{S}} f_k d\mu$ , combining this with the above [see 19.400] gives

$$(1 - \varepsilon) \cdot \int^{\mathcal{S}} f d\mu \leq \lim_{k \rightarrow \infty} \int^{\mathcal{S}} f_k d\mu \quad (19.401)$$

For  $\int^S f d\mu$  we must now consider the following cases:

$\int^S f d\mu = \infty$ . then as  $0 < 1 - \varepsilon$  we have that  $(1 - \varepsilon) \cdot \int^S f d\mu = \infty$  so that by 19.401 we have that  $\lim_{k \rightarrow \infty} \int^S f_k d\mu = \infty$  proving that

$$\int^S f d\mu = \lim_{k \rightarrow \infty} \int^S f_k d\mu \Rightarrow \int^S f d\mu \leq \lim_{k \rightarrow \infty} \int^S f_k d\mu$$

$\int^S f d\mu < \infty$ . then by 19.391 we have that  $\lim_{k \rightarrow \infty} \int^S f_k d\mu < \infty$ . So using 19.402  $\int^S f d\mu \leq \lim_{k \rightarrow \infty} \int^S f_k d\mu + \varepsilon \cdot \int^S f d\mu$  which as  $0 < \varepsilon < 1$  is chosen arbitrary proves by 9.57 that

$$\int^S f d\mu \leq \lim_{k \rightarrow \infty} \int^S f_k d\mu$$

So in all case we have  $\int^S f d\mu \leq \lim_{k \rightarrow \infty} \int^S f_k d\mu$ , the opposite inequality of 19.391. Hence

$$\int^S f d\mu = \lim_{k \rightarrow \infty} \int^S f_k d\mu$$

□

### 19.6.2 Integral of non negative functions

We extend now the integral on non negative simple function to non negative measurable functions.

**Definition 19.200.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space then we define  $\mathcal{M}^+[X, \mathcal{A}]$  and  $\overline{\mathcal{M}^+[X, \mathcal{A}]}$  as follows:

$$\mathcal{M}^+[X, \mathcal{A}] = \{f \in \mathcal{M}[X, \mathcal{A}] \mid f(x) \subseteq [0, \infty]\} \subseteq \mathcal{M}[X, \mathcal{A}] \text{ / see 19.146/}$$

$$\overline{\mathcal{M}^+[X, \mathcal{A}]} = \{f \in \overline{\mathcal{M}[X, \mathcal{A}]} \mid f(X) \subseteq [0, \infty]\} \subseteq \overline{\mathcal{M}[X, \mathcal{A}]} \text{ / see 19.145/}$$

**Note 19.201.** Using 19.148 we have  $\mathcal{M}[X, \mathcal{A}] \subseteq \overline{\mathcal{M}[X, \mathcal{A}]}$  proving that  $\mathcal{M}^+[X, \mathcal{A}] \subseteq \overline{\mathcal{M}^+[X, \mathcal{A}]}$

**Definition 19.202.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space and  $f \in \overline{\mathcal{M}^+[X, \mathcal{A}]}$  [in other words  $f$  is a non negative  $\mathcal{A}$ -measurable function]. Then  $\int^+ : \overline{\mathcal{M}^+[X, \mathcal{A}]} \rightarrow [0, \infty]$  is defined by  $f \rightarrow \int^+ f d\mu$  where

$$\int^+ f d\mu = \sup \left( \left\{ \int^S g d\mu \mid g \in \mathcal{S}^+[X, \mathcal{A}] \text{ and } \forall x \in X \models g(x) \leq f(x) \right\} \right)$$

If  $A \in \mathcal{A}$  then we defined  $\int_A^+ f d\mu$  by

$$\int_A^+ f d\mu = \int^+ (f \cdot \chi_A) d\mu$$

**Proof.** Of course we must prove that  $\{\int^S g d\mu \mid g \in \mathcal{S}^+[X, \mathcal{A}] \text{ and } \forall x \in X \models g(x) \leq f(x)\}$  is not empty to make this definition valid. Using 19.192 we have that  $C_0 \in \mathcal{S}^+[X, \mathcal{A}]$  and  $C_0(x) = 0 \leq f(x)$  proving that  $\{\int^S g d\mu \mid g \in \mathcal{S}^+[X, \mathcal{A}] \text{ and } \forall x \in X \models g(x) \leq f(x)\} \neq \emptyset$  □

**Note 19.203.** Note that in the definition of  $\int^+$  we don't use the fact that  $f$  must be  $\mathcal{A}$ -measurable. Actual the definition works for every non negative function. However the requirement that  $f$  is  $\mathcal{A}$ -measurable gives us nice limit properties to the integral.

We show now that  $\int^+$  is indeed a extension of  $\int^S$

**Proposition 19.204.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measurable space and  $f \in \mathcal{S}^+[X, \mathcal{A}]$  then

$$\int^+ f d\mu = \int^S f d\mu$$

**Proof.** As  $f \in \mathcal{S}^+[X, \mathcal{A}]$  and  $\forall x \in X f(x) \leq f(x)$  we have that  $\int^S f d\mu \in \{\int^S g d\mu | g \in \mathcal{S}^+[X, \mathcal{A}] \text{ and } \forall x \in X \models g(x) \leq f(x)\}$  so that

$$\int^S f d\mu \leq \sup \left( \left\{ \int^S g d\mu | g \in \mathcal{S}^+[X, \mathcal{A}] \text{ and } \forall x \in X \models g(x) \leq f(x) \right\} \right) = \int^+ f d\mu$$

For the opposite inequality, if  $g \in \mathcal{S}^+[X, \mathcal{A}]$  with  $\forall x \in X g(x) \leq f(x)$  then by 19.197 we have that  $\int^S g d\mu \leq \int^S f d\mu$  so that

$$\int^+ f d\mu \stackrel{\text{def}}{=} \sup \left( \left\{ \int^S g d\mu | g \in \mathcal{S}^+[X, \mathcal{A}] \text{ and } \forall x \in X \models g(x) \leq f(x) \right\} \right) \leq \int^+ f d\mu.$$

Combining the two above inequalities proves

$$\int^+ f d\mu = \int^S f d\mu$$

□

**Example 19.205.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space then for  $C_0: X \rightarrow \overline{\mathbb{R}}$  defined by  $C_0(x) = 0$  we have that  $\int^+ C_0 d\mu = 0$

**Proof.** Let  $x \in X$  then take  $\{A_i\}_{i \in \{1, \dots, 1\}}$  and  $\{a_i\}_{i \in \{1, \dots, 1\}}$  defined by  $A_1 = X$  and  $a_1 = 0$  then  $(\sum_{i=1}^1 a_i \cdot \mathcal{X}_{A_i})(x) = a_1 \cdot \mathcal{X}_{A_1}(x) = 0 \cdot 1 = 0 = C_0(x)$  proving that

$$C_0 = \sum_{i=1}^1 a_i \cdot \mathcal{X}_{A_i} \in \mathcal{S}^+[X, \mathcal{A}]$$

Hence

$$\begin{aligned} \int^+ C_0 d\mu &\stackrel{19.204}{=} \int^S C_0 d\mu \\ &= \sum_{i=1}^1 a_i \cdot \mu(A_i) \\ &= a_1 \cdot \mu(A_1) \\ &= 0 \cdot \mu(X) \\ &= 0 \end{aligned}$$

□

**Example 19.206.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space and  $f \in \overline{\mathcal{M}^+[X, \mathcal{A}]}$  such that  $f = 0$   $\mu$ -a.e. then  $\int f d\mu = 0$

**Proof.** As  $f = 0$   $\mu$ -a.e. there exists a  $N \in \mathcal{A}$  with  $\mu(N) = 0$  such that  $\{x \in X | 0 < f(x)\} \subseteq N$ . Let now  $g \in \mathcal{S}^+[X, \mathcal{A}]$  be such that  $g(x) \leq f(x)$  then if  $x \in \{x \in X | 0 < g(x)\}$  we have  $0 < g(x) \leq f(x)$  so that  $x \in \{x \in X | 0 < f(x)\} \subseteq N$ . Hence we have that  $\{x \in X | 0 < g(x)\} \subseteq N$  proving that  $g = 0$   $\mu$ -a.e.. Applying then 19.197 proves that  $\int^+ g d\mu = 0$ . So

$$\left\{ \int g d\mu | g \in \mathcal{S}^+[X, \mathcal{A}] \text{ with } \forall x \in X g(x) \leq f(x) \right\} = \{0\}$$

which finally gives

$$\int^+ f d\mu = \sup \left( \left\{ \int g d\mu | g \in \mathcal{S}^+[X, \mathcal{A}] \text{ with } \forall x \in X g(x) \leq f(x) \right\} \right) = 0$$

□

Next we show that  $\int^+$  has the same properties as  $\int^S$

**Theorem 19.207.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space and let  $f \in \overline{\mathcal{M}^+[X, \mathcal{A}]}$  [a non negative  $\mathcal{A}$ -measurable function] and  $\{f_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{S}^+[X, \mathcal{A}]$  such that  $f(x) = \lim_{n \rightarrow \infty} f_n(x)$  and  $\forall n \in \mathbb{N}_0$  and  $f_n(x) \leq f_{n+1}(x)$  so then  $\lim_{n \rightarrow \infty} \int^S f_n d\mu$  exists and

$$\int^+ f d\mu = \sup \left( \left\{ \int^S f_n d\mu | n \in \mathbb{N}_0 \right\} \right) = \lim_{n \rightarrow \infty} \int^S f_n d\mu \stackrel{19.204}{=} \lim_{n \rightarrow \infty} \int^+ f_n d\mu$$

**Proof.** As  $\forall x \in X$  we have that  $\forall n \in \mathbb{N}_0$   $f_n(x) \leq f_{n+1}(x)$  it follows from 17.71 that  $\lim_{n \rightarrow \infty} f_n(x)$  exists and

$$\forall x \in X \text{ we have } \sup(\{f_n(x) | n \in \mathbb{N}_0\}) = \lim_{n \rightarrow \infty} f_n(x) = f(x) \quad (19.402)$$

Using 19.197 it follows that  $\int^{\mathcal{S}} f_n d\mu \leq \int^{\mathcal{S}} f_{n+1} d\mu$ , so that by 17.71 we have that  $\lim_{n \rightarrow \infty} \int^{\mathcal{S}} f_n d\mu$  exists and

$$\lim_{n \rightarrow \infty} \int^{\mathcal{S}} f_n d\mu = \sup\left(\left\{\int^{\mathcal{S}} f_n d\mu | n \in \mathbb{N}_0\right\}\right) \quad (19.403)$$

From 19.402 it follows that  $\forall x \in X$  we have  $f_n(x) \leq f(x) \forall n \in \mathbb{N}_0$  so that

$$\int^{\mathcal{S}} f_n d\mu \leq \sup\left(\left\{\int^{\mathcal{S}} g d\mu | g \in \mathcal{S}^+[X, \mathcal{A}] \text{ and } \forall x \in X g(x) \leq f(x)\right\}\right) \stackrel{\text{def}}{=} \int^+ f d\mu$$

proving that  $\sup(\{\int^{\mathcal{S}} f_n d\mu | n \in \mathbb{N}_0\}) \leq \int^+ f d\mu$ . Combining this with 19.403 proves

$$\lim_{n \rightarrow \infty} \int^{\mathcal{S}} f_n d\mu \leq \int^+ f d\mu \quad (19.404)$$

For the opposite inclusion let  $g \in \mathcal{S}^+[X, \mathcal{A}]$  be such that  $\forall x \in X g(x) \leq f(x)$ . Then using 19.197 we have that

$$\{g \wedge f_n\}_{n \in \mathbb{N}_0} \subseteq \mathcal{S}^+[X, \mathcal{A}] \text{ and } \forall n \in \mathbb{N}_0 \int^{\mathcal{S}} (g \wedge f_n) d\mu \leq \int^{\mathcal{S}} f_n d\mu \quad (19.405)$$

Let  $x \in X$  then given  $n \in \mathbb{N}_0$  we have

$$(g \wedge f_n)(x) = \min(g(x), f_n(x)) \leq \min(g(x), f_{n+1}(x)) = (g \wedge f_{n+1})(x) \quad (19.406)$$

so using 17.71 it follows that  $\lim_{n \rightarrow \infty} (g \wedge f_n)(x)$  exists and

$$\lim_{n \rightarrow \infty} (g \wedge f_n)(x) = \sup(\{(g \wedge f_n)(x) | n \in \mathbb{N}_0\}) \quad (19.407)$$

As  $(g \wedge f_n)(x) = \min(g(x), f_n(x)) \leq g(x)$  we have that

$$\sup(\{(g \wedge f_n)(x) | n \in \mathbb{N}_0\}) \leq g(x) \quad (19.408)$$

As  $g(x) \leq f(x)$  we have for  $g(x)$  either

$g(x) = f(x)$ . then  $g(x) = f(x) \stackrel{19.402}{=} \sup(\{f_n(x) | n \in \mathbb{N}_0\})$  so that  $\forall n \in \mathbb{N}_0$  we have  $f_n(x) \leq g(x)$ , or  $(g \wedge f_n)(x) = \min(g(x), f_n(x)) = f_n(x)$  proving that

$$\sup(\{(g \wedge f_n)(x) | n \in \mathbb{N}_0\}) = \sup(\{f_n(x) | n \in \mathbb{N}_0\}) \stackrel{19.402}{=} f(x) = g(x)$$

$g(x) < f(x)$ . then  $\exists n \in \mathbb{N}_0$  such that  $g(x) < f_n(x) \leq f(x)$ , hence  $(g \wedge f_n)(x) = \min(g(x), f_n(x)) = g(x)$  so that  $g(x) \leq \sup(\{(g \wedge f_n)(x) | n \in \mathbb{N}_0\})$  which by 19.408 proves that

$$\sup(\{(g \wedge f_n)(x) | n \in \mathbb{N}_0\}) = g(x)$$

So in all cases we have  $\sup(\{(g \wedge f_n)(x) | n \in \mathbb{N}_0\}) = g(x)$  which by 19.407 proves that

$$\lim_{n \rightarrow \infty} (g \wedge f_n)(x) = g(x) \quad (19.409)$$

Using 19.405, 19.406 and 19.409 we have by 19.199 that

$$\int^{\mathcal{S}} g(x) d\mu = \lim_{n \rightarrow \infty} \int^{\mathcal{S}} (g \wedge f_n) d\mu \quad (19.410)$$

Note that by 19.405  $\int^{\mathcal{S}} (g \wedge f_n) d\mu \leq \int^{\mathcal{S}} f_n d\mu$ , so that using the properties of the limit [see 17.77 we have that  $\lim_{n \rightarrow \infty} \int^{\mathcal{S}} (g \wedge f_n) d\mu \leq \lim_{n \rightarrow \infty} \int^{\mathcal{S}} f_n d\mu$ . Combining this with 19.410 gives  $\int^{\mathcal{S}} g d\mu = \lim_{n \rightarrow \infty} \int^{\mathcal{S}} (g \wedge f_n) d\mu \leq \lim_{n \rightarrow \infty} \int^{\mathcal{S}} f_n d\mu$ . So

$$\int^+ f d\mu \stackrel{\text{def}}{=} \sup\left(\left\{\int^{\mathcal{S}} g d\mu | g \in \mathcal{S}^+[X, \mathcal{A}] \text{ with } \forall x \in X g(x) \leq f(x)\right\}\right) \leq \lim_{n \rightarrow \infty} \int^{\mathcal{S}} f_n d\mu$$

which combined with 19.404 proves finally

$$\int^+ f d\mu = \lim_{n \rightarrow \infty} \int^S f_n d\mu \stackrel{19.403}{=} \sup \left( \left\{ \int^S f_n d\mu \mid n \in \mathbb{N}_0 \right\} \right)$$

□

In the following corollary we make use of the requirement of measurability to give an alternative definition of  $\int^+$ .

**Corollary 19.208.** *Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space and  $f: X \rightarrow \overline{\mathbb{R}}$  such that  $f \in \overline{\mathcal{M}[X, \mathcal{A}]}$  then there exists a sequence  $\{f_n\}_{n \in \mathbb{N}_0} \subseteq \mathcal{S}^+[X, \mathcal{A}]$ , with  $\forall n \in \mathbb{N}_0$  and  $\forall x \in X f_n(x) \leq f_{n+1}(x)$ , such that*

$$\forall x \in X \text{ we have } f(x) = \lim_{n \rightarrow \infty} f_n(x)$$

$$\int^+ f d\mu = \lim_{n \rightarrow \infty} \int^S f_n d\mu = \sup \left( \left\{ \int^S f_n d\mu \mid n \in \mathbb{N}_0 \right\} \right)$$

**Proof.** As  $f$  is  $\mathcal{A}$ -measurable and non negative we can use 19.178 to find a sequence  $\{f_n\}_{n \in \mathbb{N}_0} \subseteq \mathcal{S}^+[X, \mathcal{A}]$  such that  $\forall x \in X$  we have  $\forall n \in \mathbb{N}_0 f_n(x) \leq f_{n+1}(x)$  and

$$f(x) = \lim_{n \rightarrow \infty} f_n(x) = \sup (\{f_n(x) \mid n \in \mathbb{N}_0\}).$$

Using then the previous theorem [see 19.207] we have that

$$\int^+ f d\mu = \lim_{n \rightarrow \infty} \int^S f_n d\mu = \sup \left( \left\{ \int^S f_n d\mu \mid n \in \mathbb{N}_0 \right\} \right)$$

□

The above corollary let's us to easily transfer the properties of the integral of simple functions to properties of the integral of non negative measurable functions. 19.197

**Proposition 19.209.** *Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space and  $f, g \in \overline{\mathcal{M}^+[X, \mathcal{A}]}$  then we have*

1.  $\forall \alpha \in [0, \infty[$  we have  $\alpha \cdot f \in \overline{\mathcal{M}^+[X, \mathcal{A}]}$  and  $\int^+ \alpha \cdot f d\mu = \alpha \cdot \int^+ f d\mu$
2.  $f + g \in \overline{\mathcal{M}^+[X, \mathcal{A}]}$  and  $\int^+ (f + g) d\mu = \int^+ f d\mu + \int^+ g d\mu$
3. If  $\forall x \in X$  we have  $f(x) \leq g(x)$  then  $\int^+ f d\mu \leq \int^+ g d\mu$

**Proof.** First using the previous corollary [see 19.208] there exists  $\{f_i\}_{i \in \mathbb{N}_0}, \{g_i\}_{i \in \mathbb{N}} \subseteq \mathcal{S}^+[X, \mathcal{A}]$  of increasing simple functions so that

$$f = \lim_{n \rightarrow \infty} f_n, g = \lim_{n \rightarrow \infty} g_n, \int^+ f d\mu = \lim_{n \rightarrow \infty} \int^S f_n d\mu \text{ and } \int^+ g d\mu = \lim_{n \rightarrow \infty} \int^S g_n d\mu \quad (19.411)$$

then we have:

1. Using 19.170 and the fact that  $0 \leq \alpha$  it follows that  $\alpha \cdot f \in \overline{\mathcal{M}^+[X, \mathcal{A}]}$ . Further using 19.197 and the fact that  $0 \leq \alpha$  we have that

$$\{\alpha \cdot f_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{S}^+[X, \mathcal{A}]$$

Given  $x \in X$  we have  $\forall n \in \mathbb{N}_0$

$$\alpha \cdot f_n(x) \leq \alpha \cdot f_{n+1}(x)$$

Further

$$\begin{aligned} (\alpha \cdot f)(x) &= \alpha \cdot f(x) \\ &= \alpha \cdot \lim_{n \rightarrow \infty} f_n(x) \\ &\stackrel{17.77}{=} \lim_{n \rightarrow \infty} (\alpha \cdot f_n(x)) \\ &= \lim_{n \rightarrow \infty} (\alpha \cdot f_n)(x) \end{aligned}$$

Hence using 19.207

$$\begin{aligned}
 \int^+(\alpha \cdot f) d\mu &= \lim_{n \rightarrow \infty} \int^{\mathcal{S}} (\alpha \cdot f_n) d\mu \\
 &\stackrel{19.197}{=} \lim_{n \rightarrow \infty} \left( \alpha \cdot \int^{\mathcal{S}} f_n d\mu \right) \\
 &\stackrel{17.77}{=} \alpha \cdot \lim_{n \rightarrow \infty} \int^{\mathcal{S}} f_n d\mu \\
 &\stackrel{19.411}{=} \alpha \cdot \int^{\mathcal{S}} f d\mu
 \end{aligned}$$

2. Using 19.170 we have that  $f + g \in \overline{\mathcal{M}^+[X, \mathcal{A}]}$ . Further using 19.197 we have that

$$\{f_i + g_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{S}^+[X, \mathcal{A}]$$

Given  $x \in X$  we have  $\forall n \in \mathbb{N}_0$

$$\begin{aligned}
 (f_n + g_n)(x) &= f_n(x) + g_n(x) \\
 &\leq f_{n+1}(x) + g_{n+1}(x) \\
 &= (f_{n+1} + g_{n+1})(x)
 \end{aligned}$$

Further

$$\begin{aligned}
 (f + g)(x) &= f(x) + g(x) \\
 &= \lim_{n \rightarrow \infty} f_n(x) + \lim_{n \rightarrow \infty} g_n(x) \\
 &\stackrel{17.77}{=} \lim_{n \rightarrow \infty} (f_n + g_n)(x)
 \end{aligned}$$

Hence using 19.207

$$\begin{aligned}
 \int^+(f + g) d\mu &= \lim_{n \rightarrow \infty} \int^+(f_n + g_n) d\mu \\
 &\stackrel{19.197}{=} \lim_{n \rightarrow \infty} \left( \int^+ f_n d\mu + \int^+ g_n d\mu \right) \\
 &\stackrel{17.77}{=} \lim_{n \rightarrow \infty} \int^+ f_n d\mu + \lim_{n \rightarrow \infty} \int^+ g_n d\mu \\
 &\stackrel{19.411}{=} \int^+ f d\mu + \int^+ g d\mu
 \end{aligned}$$

3. Assume that  $\forall x \in X f(x) \leq g(x)$ . Let  $s \in \left\{ \int^{\mathcal{S}} h d\mu \mid h \in \mathcal{S}^+[X, \mathcal{A}] \right\}$  and  $\forall x \in X$  we have  $h(x) \leq f(x)$  then there exists a  $h \in \mathcal{S}^+[X, \mathcal{A}]$  with  $\forall x \in X h(x) \leq f(x) \leq g(x)$  such that  $s = \int^{\mathcal{S}} h d\mu$ , which proves that  $s \in \left\{ \int^{\mathcal{S}} h d\mu \mid h \in \mathcal{S}^+[X, \mathcal{A}] \text{ and } \forall x \in X \text{ we have } h(x) \leq g(x) \right\}$  or

$$\begin{cases} \int^{\mathcal{S}} h d\mu \mid h \in \mathcal{S}^+[X, \mathcal{A}] \text{ and } \forall x \in X \text{ we have } h(x) \leq f(x) \end{cases} \subseteq$$

$$\begin{cases} \int^{\mathcal{S}} h d\mu \mid h \in \mathcal{S}^+[X, \mathcal{A}] \text{ and } \forall x \in X \text{ we have } h(x) \leq g(x) \end{cases}$$

hence using 2.175

$$\begin{aligned}
 \int^+ f d\mu &= \\
 \sup \left( \left\{ \int^{\mathcal{S}} h d\mu \mid h \in \mathcal{S}^+[X, \mathcal{A}] \text{ and } \forall x \in X \text{ we have } h(x) \leq f(x) \right\} \right) &\leq \\
 \sup \left( \left\{ \int^{\mathcal{S}} h d\mu \mid h \in \mathcal{S}^+[X, \mathcal{A}] \text{ and } \forall x \in X \text{ we have } h(x) \leq g(x) \right\} \right) &= \\
 \int^+ g d\mu
 \end{aligned}$$

□

**Corollary 19.210.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space,  $n \in \mathbb{N}_0$  and  $\{f_i\}_{i \in \mathbb{N}_0} \subseteq \overline{\mathcal{M}^+[X, \mathcal{A}]}$  then if we define  $\sum_{i=1}^n f_i$  by  $(\sum_{i=1}^n f_i)(x) = \sum_{i=1}^n f_i(x)$  we have

$$\sum_{i=1}^n f_i \in \overline{\mathcal{M}^+[X, \mathcal{A}]}$$

and

$$\int^+ \left( \sum_{i=1}^n f_i \right) d\mu = \sum_{i=1}^n \int^+ f_i d\mu$$

**Proof.** This is trivially proved by induction. So let

$$\mathcal{S} = \left\{ n \in \mathbb{N}_0 \mid \text{If } \{f_i\}_{i \in n} \subseteq \overline{\mathcal{M}^+[X, \mathcal{A}]} \text{ then } \sum_{i=1}^n f_i \in \overline{\mathcal{M}^+[X, \mathcal{A}]} \text{ and } \int^+ \left( \sum_{i=1}^n f_i \right) d\mu = \sum_{i=1}^n \int^+ f_i d\mu \right\}$$

then we have:

**1  $\in \mathcal{S}$ .** Let  $\{f_i\}_{i \in \{1, \dots, 1\}} \subseteq \overline{\mathcal{M}^+[X, \mathcal{A}]}$  then  $\sum_{i=1}^1 f_i = f_1 \in \overline{\mathcal{M}^+[X, \mathcal{A}]}$  and  $\int^+ (\sum_{i=1}^1 f_i) d\mu = \int^+ f_1 d\mu = \sum_{i=1}^1 \int^+ f_i d\mu$  proving that  $1 \in \mathcal{S}$ .

**$n \in \mathcal{S} \Rightarrow n+1 \in \mathcal{S}$ .** Let  $\{f_i\}_{i \in \{1, \dots, n\}} \subseteq \overline{\mathcal{M}^+[X, \mathcal{A}]}$  then as  $n \in \mathcal{S}$  we have that  $\sum_{i=1}^n f_i \in \overline{\mathcal{M}^+[X, \mathcal{A}]}$  hence

$$\sum_{i=1}^{n+1} f_i = \left( \sum_{i=1}^n f_i \right) + f_{n+1} \in \overline{\mathcal{M}^+[X, \mathcal{A}]} \quad [\text{using 19.209(2)}]$$

Further we have

$$\begin{aligned} \int^+ \left( \sum_{i=1}^{n+1} f_i \right) d\mu &= \int^+ \left( \left( \sum_{i=1}^n f_i \right) + f_{n+1} \right) d\mu \\ &\stackrel{19.209(2)}{=} \int^+ \left( \sum_{i=1}^n f_i \right) d\mu + \int^+ f_{n+1} d\mu \\ &\stackrel{n \in \mathcal{S}}{=} \left( \sum_{i=1}^n \int^+ f_i d\mu \right) + \int^+ f_{n+1} d\mu \\ &= \sum_{i=1}^{n+1} \int^+ f_i d\mu \end{aligned}$$

proving that  $n+1 \in \mathcal{S}$ . □

### 19.6.3 Integral of integrable functions

Now we can define the integral on functions that can have positive or negative values. First note that if  $f \in \overline{\mathcal{M}[X, \mathcal{A}]}$  then by 19.177 we have that  $f^+, f^-$  are  $\mathcal{A}$ -measurable which as  $f^+(X) \subseteq [0, \infty] \wedge f^-(X) \subseteq [0, \infty]$  [see 19.175] proves that  $f^+, f^- \in \overline{\mathcal{M}^+[X, \mathcal{A}]}$ . This allows for the following definition.

**Definition 19.211.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space then  $f \in \overline{\mathbb{R}^X}$  is  **$\mu$ -integrable** if  $f \in \overline{\mathcal{M}[X, \mathcal{A}]}$  and  $\int^+ f^+ d\mu$  and  $\int^+ f^- d\mu$  is a real number [so that  $\int^+ f^+ d\mu - \int^+ f^- d\mu$  is well defined]. The integral  $\int f d\mu$  is then defined by

$$\int f d\mu \stackrel{\text{def}}{=} \int^+ f^+ d\mu - \int^- f^- d\mu \in \mathbb{R}$$

The set of all  $\mu$ -integrable function graphs is noted as  $\mathcal{L}[X, \mathcal{A}, \mu, \overline{\mathbb{R}}]$  so

$$\mathcal{L}[X, \mathcal{A}, \mu, \overline{\mathbb{R}}] = \left\{ f \in \overline{\mathcal{M}[X, \mathcal{A}]} \mid \int^+ f^+ d\mu \in \mathbb{R} \vee \int^- f^- d\mu \in \mathbb{R} \right\}$$

We have a similar set of real valued  $\mu$ -integrable functions  $\mathcal{L}[X, \mathcal{A}, \mu, \mathbb{R}]$  defined by

$$\mathcal{L}[X, \mathcal{A}, \mu, \mathbb{R}] = \left\{ f \in \mathcal{M}[X, \mathcal{A}] \mid \int^+ f^+ d\mu \in \mathbb{R} \vee \int^+ f^- d\mu \in \mathbb{R} \right\}$$

**Note 19.212.** As  $\mathcal{M}[X, \mathcal{A}] \subseteq \overline{\mathcal{M}[X, \mathcal{A}]}$  [see 19.148] we have that

$$\mathcal{L}[X, \mathcal{A}, \mu, \mathbb{R}] \subseteq \mathcal{L}[X, \mathcal{A}, \overline{\mathbb{R}}]$$

**Note 19.213.** In contrast to  $\int^+ f d\mu$  which can be  $\infty$  we have that  $\int f d\mu$  is always a real number.

The above defines integrability on functions on  $X$ , if  $A$  is a measurable set of a measure space we can define integrability on this set as follows.

**Definition 19.214.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space,  $A \in \mathcal{A}$  and  $f \in \overline{\mathbb{R}}^X$  then  $f$  is  **$\mu$ -integrable on  $A$**  if  $f \cdot \chi_A$  is  $\mu$ -integrable and  $\int_A f d\mu$  is defined by

$$\int_A f d\mu = \int f \cdot \chi_A d\mu$$

**Definition 19.215.** Let  $X$  be a set,  $A \subseteq X$  and  $f: A \rightarrow \overline{\mathbb{R}}$  be a function then  $f_{A \rightarrow X}: X \rightarrow \overline{\mathbb{R}}$  is defined by

$$f_{A \rightarrow X}(x) = \begin{cases} f(x) & \text{if } x \in A \\ 0 & \text{if } x \in X \setminus A \end{cases}$$

**Note 19.216.** Using the above definition it is trivial to see that  $(f_{A \rightarrow X})|_A = f$

The above definition allows us to define  $\mu$ -integrability and the integral on functions with domain a measurable set of a measure space.

**Definition 19.217.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space,  $A \in \mathcal{A}$  and  $f \in \overline{\mathbb{R}}^A$  then  $f$  is  **$\mu$ -integrable on  $A$**  if  $f_{A \rightarrow X}$  is  **$\mu$ -integrable**. Further the integral of  $f$  over  $A$  is defined as the integral of  $f_{X \rightarrow A}$  over  $A$ , hence

$$\int_A f d\mu = \int_A f_{X \rightarrow A} d\mu$$

If  $\mu(X \setminus A) = 0$  then we note  $\int_A f d\mu$  as  $\int f d\mu$  so

$$\text{If } \mu(X \setminus A) = 0 \text{ then } \int f d\mu = \int_A f_{X \rightarrow A} d\mu$$

In the above definition  $f$  is  $\mu$ -integrable if  $f_{X \rightarrow A}$  is  $\mathcal{A}$ -measurable and  $\int^+ (f_{X \rightarrow A})^+ d\mu$  and  $\int^+ (f_{X \rightarrow A})^- d\mu$  are finite. We can reduce this to the condition that  $f$  must be  $\mathcal{A}$ -measurable and  $\int^+ (f_{X \rightarrow A})^+ d\mu$  and  $\int^+ (f_{X \rightarrow A})^- d\mu$  is finite using the following proposition.

**Proposition 19.218.** Let  $\langle X, \mathcal{A} \rangle$  be a measurable space,  $A \in \mathcal{A}$  and  $f: A \rightarrow \overline{\mathbb{R}}$  be a  $\mathcal{A}$ -measurable function then  $f_{A \rightarrow X}$  is  $\mathcal{A}$ -measurable [or  $f_{A \rightarrow X} \in \overline{\mathcal{M}[X, \mathcal{A}]}$ ].

**Proof.** We use 19.151 to prove this. Let  $a \in \mathbb{R}$  then we have as  $f$  is  $\mathcal{A}$ -measurable that

$$\{x \in A \mid f(x) > a\} = \{f > a\} \in \mathcal{A} \quad (19.412)$$

For  $a$  we have either

**$a \geq 0$ .** Let  $x \in \{x \in X \mid f_{A \rightarrow X}(x) > a\}$  then we can not have  $x \in X \setminus A$  [for then  $f_{A \rightarrow X}(x) = 0 \leq a \Rightarrow f_{A \rightarrow X}(x) > a$ ] hence  $x \in A$ , so that  $f(x) = f_{A \rightarrow X}(x) > a$  proving  $x \in \{x \in A \mid f(x) > a\}$ . Further if  $x \in \{x \in A \mid f(x) > a\}$  then  $x \in A$  and  $f(x) > a$  so that  $f_{A \rightarrow X}(x) = \sup_{x \in A} f(x) > a$  proving that  $x \in \{x \in X \mid f_{A \rightarrow X}(x) > a\}$ . Hence  $\{x \in X \mid f_{A \rightarrow X}(x) > a\} = \{x \in A \mid f(x) > a\} \in \mathcal{A}$  which proves that

$$\{x \in X \mid f_{A \rightarrow X}(x) > a\} \in \mathcal{A}$$

**a < 0.** Let  $x \in \{x \in X | f_{A \rightarrow X}(x) > a\}$  then  $f_{A \rightarrow X}(x) > a$  and for  $x \in X$  we have either

**x ∈ A.** Then  $f(x) = f_{A \rightarrow x}(x) > 0$  proving that  $x \in \{x \in A | f(x) > a\}$  so that

$$x \in \{x \in A | f(x) > a\} \bigcup (X \setminus A)$$

**x ∈ X \ A.** Then

$$x \in \{x \in A | f(x) > a\} \bigcup (X \setminus A)$$

From this we conclude that

$$\{x \in X | f_{A \rightarrow X}(x) > a\} \subseteq \{x \in A | f(x) > a\} \bigcup (X \setminus A) \quad (19.413)$$

If  $x \in \{x \in A | f(x) > a\} \bigcup (X \setminus A)$  then either

**x ∈ X \ A.**  $f_{A \rightarrow X}(x) = 0 > a$  proving that

$$x \in \{x \in X | f_{A \rightarrow X}(x) > a\}$$

**x ∈ A.** then  $f_{A \rightarrow X}(x) = f(x) > a$  proving that

$$x \in \{x \in X | f_{A \rightarrow X}(x) > a\}$$

Hence we conclude that  $\{x \in A | f(x) > a\} \bigcup (X \setminus A) \subseteq \{x \in X | f_{A \rightarrow X}(x) > a\}$  which together with 19.413 proves that  $\{x \in X | f_{A \rightarrow X}(x) > a\} = \{x \in A | f(x) > a\} \bigcup (X \setminus A) \in \mathcal{A}$  [by 19.412]. Hence

$$\{x \in X | f_{A \rightarrow X}(x) > a\} \in \mathcal{A}$$

As in all cases we have  $\{f_{A \rightarrow X} > a\} = \{x \in X | f_{A \rightarrow X}(x) > a\} \in \mathcal{A}$  we conclude that  $f_{X \rightarrow A}$  is  $\mathcal{A}$ -measurable.  $\square$

Using the above proposition we can simplify the condition of a function graph with domain a measurable set [see 19.217] as follows.

**Proposition 19.219.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space,  $A \in \mathcal{A}$  then  $f \in \bar{\mathbb{R}}^A$  is  $\mu$ -measurable if  $f$  is  $\mathcal{A}$ -measurable and  $\int^+(f_{A \rightarrow X})^+ d\mu, \int^-(f_{A \rightarrow X})^- d\mu$  are finite.

**Proof.** See 19.217 and 19.218  $\square$

Next we show that for a function defined on a measure space the integral over the measure space is the same as the integral.

**Proposition 19.220.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space and  $f \in \bar{\mathbb{R}}^X$  then the following are equivalent

1.  $f$  is  $\mu$ -integrable
2.  $f$  is  $\mu$ -integrable on  $X$  [using the definition 19.214]
3.  $f$  is  $\mu$ -integrable on  $X$  [using the definition 19.217]

further we have that

$$\int f d\mu = \int_X f d\mu \text{ [using the definition 19.214]}$$

$$\int f d\mu = \int_X f d\mu \text{ [using the definition 19.217]}$$

**Proof.** First note that given  $x \in X$  we have

$$\begin{aligned} f(x) &= \begin{cases} f(x) & \text{if } x \in X \\ 0 & \text{if } x \in X \setminus X = \emptyset \end{cases} \\ &= f_{X \rightarrow X}(x) \\ &= f(x) \cdot \chi_X(x) \end{aligned}$$

proving that

$$f = f_{X \rightarrow X} = f \cdot \mathcal{X}_X \quad (19.414)$$

The rest of the proof is then trivial:

**1  $\Rightarrow$  2.** As  $f$  is  $\mu$ -integrable we have that  $f$  is  $\mathcal{A}$ -measurable and  $\int^+ f^+ d\mu, \int^+ f^- d\mu$  are finite.

Hence using 19.414  $f \cdot \mathcal{X}_X$  is  $\mathcal{A}$ -measurable and  $\int^+ (f \cdot \mathcal{X}_X)^+ d\mu, \int^+ (f \cdot \mathcal{X}_X)^- d\mu$  are finite, proving that  $f$  is  $\mu$ -integrable on  $X$  [using the definition 19.214]

**2  $\Rightarrow$  3.** If  $f \cdot \mathcal{X}_X$  is  $\mathcal{A}$ -measurable and  $\int^+ (f \cdot \mathcal{X}_X)^+ d\mu, \int^+ (f \cdot \mathcal{X}_X)^- d\mu$  are finite then using

19.414 we have that  $f_{X \rightarrow X}$  is  $\mathcal{A}$ -measurable and  $\int^+ (f_{X \rightarrow X})^+ d\mu, \int^+ (f_{X \rightarrow X})^- d\mu$  are finite.

So  $f$  is  $\mu$ -integrable on  $X$  [using the definition 19.217]

**3  $\Rightarrow$  1.** If  $f_{X \rightarrow X}$  is  $\mathcal{A}$ -measurable and  $\int^+ (f_{X \rightarrow X})^+ d\mu, \int^+ (f_{X \rightarrow X})^- d\mu$  are finite then using

19.414 we have that  $f$  is  $\mathcal{A}$ -measurable  $\int^+ f^+ d\mu, \int^+ f^- d\mu$  are finite. So  $f$  is  $\mu$ -integrable.

Finally we have

$$\int f d\mu \stackrel{19.414}{=} \int f \cdot \mathcal{X}_X d\mu = \int_X f d\mu \quad [\text{using the definition 19.214}]$$

$$\int f d\mu \stackrel{19.414}{=} \int f \cdot \mathcal{X}_X d\mu \stackrel{19.414}{=} \int f_{X \rightarrow X} \cdot \mathcal{X}_X d\mu = \int_X f_{X \rightarrow X} d\mu = \int_X f d\mu \quad [\text{using the definition 19.217}]$$

□

The following proposition shows that  $\int$  is a extension of  $\int^+$

**Proposition 19.221.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space  $f \in \overline{\mathbb{R}^X}$  with  $f(X) \subseteq [0, \infty]$  then  $f$   $\mu$ -integrable [see 19.211] iff  $f \in \overline{\mathcal{M}^+[X, \mathcal{A}]}$  and  $\int^+ f d\mu$  is finite. Further if  $f$  is  $\mu$ -integrable then  $\int f d\mu = \int^+ f d\mu < \infty$

**Proof.** First note that as  $f(X) \subseteq [0, \infty]$  we have  $\forall x \in X$  that  $f^+(x) = \max(f(x), 0) = f(x)$  and  $f^-(x) = -\min(f(x), 0) = 0$  proving that

$$f^+ = f \wedge f^- = C_0 \quad (19.415)$$

If  $f$  is  $\mu$ -integrable then  $f$  is  $\mathcal{A}$ -measurable hence, as  $f(X) \subseteq [0, \infty]$ , it follows that  $f \in \mathcal{M}^+[X, \mathcal{A}]$ . Further  $\int^+ f^+ d\mu$  and  $\int^+ f^- d\mu$  must be finite, from which using 19.415 it follows that  $\int^+ f d\mu$  is finite. On the other hand if  $f \in \mathcal{M}^+[X, \mathcal{A}]$  then  $f$  is  $\mathcal{A}$ -measurable and if  $\int^+ f d\mu < \infty$  we have by 19.415 that  $\int^+ f^+ d\mu < \infty$  and  $\int^+ f^- d\mu = \int^+ C_0 d\mu \stackrel{19.205}{=} 0 < \infty$ , proving that  $f$  is  $\mu$ -integrable. Finally if  $f$  is  $\mu$ -integrable then

$$\begin{aligned} \int f d\mu &\stackrel{\text{def}}{=} \int^+ f^+ d\mu - \int^+ f^- d\mu \\ &\stackrel{19.415}{=} \int^+ f d\mu - \int^+ C_0 d\mu \\ &\stackrel{19.205}{=} \int^+ f d\mu - 0 \\ &= \int^+ f d\mu \end{aligned}$$

□

**Proposition 19.222.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space and  $f \in \overline{\mathcal{M}[X, \mathcal{A}]}$  such that  $f = 0$   $\mu$ -a.e. then  $f$  is  $\mu$ -integrable and  $\int f d\mu = 0$

**Proof.** As  $f = 0$   $\mu$ -a.e. we have

$$\exists N \in \mathcal{A} \text{ with } \mu(N) = 0 \text{ such that } \{x \in X \mid f(x) \neq 0\} \subseteq N$$

If  $x \in \{x \in X \mid f^+(x) \neq 0\}$  then  $\max(f(x), 0) \neq 0$  so that  $f(x) \neq 0$  proving that  $x \in \{x \in X \mid f(x) \neq 0\} \subseteq N$ . Hence  $\{x \in X \mid f^+(x) \neq 0\} \subseteq N$  proving that  $f^+ = 0$   $\mu$ -a.e.. From 19.206 we have that

$$\int^+ f^+ d\mu = 0 < \infty \quad (19.416)$$

If  $x \in \{x \in X \mid f^-(x) \neq 0\}$  then  $-\min(f(x), 0) \neq 0$  so that  $f(x) \neq 0$  proving that  $x \in \{x \in X \mid f(x) \neq 0\} \subseteq N$ . Hence  $\{x \in X \mid f(x) \neq 0\} \subseteq N$  or  $f^- = 0$   $\mu$ -a.e.. From 19.206 we have that

$$\int^+ f^- d\mu = 0 < \infty \quad (19.417)$$

Finally 19.416 and 19.417 it follows that

$$f \text{ is } \mu\text{-integrable and } \int f d\mu = \int^+ f^+ d\mu - \int^+ f^- d\mu = 0 - 0 = 0 \quad \square$$

#### 19.6.4 Properties of the integral

**Proposition 19.223.** *Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measures space,  $g \in \mathcal{L}$  a non negative real valued  $\mu$ -integrable function and  $f \in \overline{\mathcal{M}[X, \mathcal{A}]}$  a extended real valued  $\mathcal{A}$ -measurable function such that  $|f| \leq g$   $\mu$ -a.e. then  $f \in \mathcal{L}[X, \mathcal{A}, \mu, \overline{\mathbb{R}}]$  and  $\int f d\mu \leq \int g d\mu$ .*

**Proof.** We prove the theorem in two stages:

1. Assume that  $\forall x \in X$  we have  $|f(x)| \leq g(x)$ . Then  $g$  is  $\mu$ -integrable and by definition we have  $g \in \overline{\mathcal{M}[X, \mathcal{A}]}$ , further as  $g$  is non negative and real valued we have:

$$g \in \mathcal{M}^+[X, \mathcal{A}]$$

As  $f \in \overline{\mathcal{M}[X, \mathcal{A}]}$  we have by 19.177 that  $f^+, f^- \in \overline{\mathcal{M}^+(X, \mathcal{A})}$ , further  $\forall x \in X$  we have using 19.176 that  $f^+(x), f^-(x) \leq |f(x)| \leq g(x)$ , So using 19.209 it follows that  $\int^+ f^+ d\mu \leq \int^+ g d\mu$  and  $\int^+ f^- d\mu \leq \int g d\mu$ . As  $g$  is non negative and integrable we have by 19.221 that  $\int^+ g d\mu < \infty$ . Hence  $\int^+ f^+ d\mu < \infty$  and  $\int^+ f^- d\mu < \infty$  which proves that

$$f \text{ is } \mu\text{-integrable or equivalent } f \in \mathcal{L}[X, \mathcal{A}, \mu, \overline{\mathbb{R}}]$$

As  $\forall x \in X$  we have  $|f(x)| \leq g(x) < \infty$  [as  $g$  is real valued] we have that  $f$  is real valued, hence  $f \in \mathcal{L}$  and from  $f(x) \leq |f(x)| \leq g(x)$  it follows from 19.225 that

$$\int f d\mu \leq \int g d\mu$$

2. Assume that  $|f| \leq g$   $\mu$ -a.e. then if we define  $A = \{x \in X \mid -(|f(x)| \leq g(x))\}$  there exists a  $N \in \mathcal{A}$  with  $\mu(N) = 0$  such that  $A \subseteq N$ . Let  $x \in X$  then we have either:

$x \in N$ . Then  $|\mathcal{X}_{X \setminus N} \cdot f|(x) = |\mathcal{X}_{X \setminus N}(x) \cdot f(x)| = |0 \cdot f(x)| = 0 = 0 \leq g(x)$

$x \in X \setminus N$ . Then  $\mathcal{X}_{X \setminus N}(x) = 1$  and  $x \notin A$  so that  $|f(x)| \leq g(x)$  hence  $|f_1(x)| = |f(x)| \leq g(x)$

proving that

$$\forall x \in X \text{ we have } |\mathcal{X}_{X \setminus N} \cdot f|(x) \leq g(x)$$

As  $N \in \mathcal{A} \Rightarrow X \setminus N \in \mathcal{A}$  we have by 19.158 that  $\mathcal{X}_{X \setminus N} \cdot f \in \overline{\mathcal{M}[X, \mathcal{A}]}$ , this and the above allows us to apply (1) giving

$$\mathcal{X}_A \cdot f \in \mathcal{L}[X, \mathcal{A}, \mu, \overline{\mathbb{R}}] \text{ and } \int (\mathcal{X}_A \cdot f) d\mu \leq \int g d\mu$$

Let  $x \in \{x \in X \mid (\mathcal{X}_A \cdot f)(x) \neq f(x)\}$  then  $x \in X \setminus A$  [for if  $(\mathcal{X}_{X \setminus A} \cdot f)(x) = \mathcal{X}_A(x) \cdot f(x) = f(x)$ ] proving that  $\{x \in X \mid (\mathcal{X}_A \cdot f)(x) \neq f(x)\} \subseteq X \setminus A \subseteq N$ . Hence  $f = \mathcal{X}_A \cdot f$   $\mu$ -a.e. so we can apply 19.228 giving that

$$f \in \mathcal{L}[X, \mathcal{A}, \mu, \overline{\mathbb{R}}] \text{ and } \int f d\mu = \int (\mathcal{X}_A \cdot f) d\mu \leq \int g d\mu \quad \square$$

**Lemma 19.224.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space and let  $f_1, g_1, f_2, g_2$  be non negative real valued  $\mu$ -integrable functions [see 19.211] such that  $f_1 - f_2 = g_1 - g_2$  then

$$\int^+ f_1 d\mu - \int^+ f_2 d\mu = \int f_1 d\mu - \int f_2 d\mu = \int g_1 d\mu - \int g_2 d\mu = \int^+ g_1 d\mu - \int^+ g_2 d\mu$$

**Proof.** Let  $x \in X$  then from  $f_1 - f_2 = g_1 - g_2$  we have

$$\begin{aligned} (f_1 - f_2)(x) &= (g_1 - g_2)(x) & \Rightarrow & f_1(x) - f_2(x) = g_1(x) - g_2(x) \\ & \xrightarrow{f_1(x), f_2(x), g_1(x), g_2(x) \in \mathbb{R}} f_1(x) + g_2(x) & \Rightarrow & f_1(x) + g_2(x) = f_2(x) + g_1(x) \\ & & & \Rightarrow (f_1 + g_2)(x) = (f_2 + g_1)(x) \end{aligned}$$

proving that

$$f_1 + g_2 = f_2 + g_1 \quad (19.418)$$

As  $f_1, f_2, g_1, g_2$  are  $\mu$ -integrable and non negative we have by the above proposition [see 19.221] that

$$\begin{aligned} f_1, g_1, f_2, g_2 &\in \overline{\mathcal{M}^+[X, \mathcal{A}]} \\ \int f_1 d\mu &= \int^+ f_1 d\mu < \infty \\ \int f_2 d\mu &= \int^+ f_2 d\mu < \infty \\ \int g_1 d\mu &= \int^+ g_1 d\mu < \infty \\ \int g_2 d\mu &= \int^+ g_2 d\mu < \infty \end{aligned}$$

Using 19.209 we have that

$$\int^+ f_1 d\mu + \int^+ g_2 d\mu \stackrel{19.209}{=} \int^+ (f_1 + g_2) d\mu \stackrel{19.418}{=} \int^+ (f_2 + g_1) d\mu \stackrel{19.209}{=} \int^+ f_2 d\mu + \int^+ g_1 d\mu$$

So we have that

$$\int^+ f_1 d\mu - \int^+ f_2 d\mu = \int^+ g_1 d\mu - \int^+ g_2 d\mu$$

which combined with 19.221 proves that

$$\int f_1 d\mu - \int f_2 d\mu = \int g_1 d\mu - \int g_2 d\mu \quad \square$$

**Proposition 19.225.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space and let  $f, g \in \mathcal{L}[X, \mathcal{A}, \mu, \mathbb{R}]$  [real valued  $\mu$ -integrable functions] then

1.  $\forall \alpha \in \mathbb{R}$  we have that  $\alpha \cdot f$  is  $\mu$ -integrable and  $\int (\alpha \cdot f) d\mu = \alpha \cdot \int f d\mu$
2.  $f + g$  is  $\mu$ -integrable and  $\int (f + g) d\mu = \int f d\mu + \int g d\mu$
3. If  $\forall x \in X \quad f(x) \leq g(x)$  holds then  $\int f d\mu \leq \int g d\mu$

**Proof.**

1. As  $f$  is  $\mu$ -integrable we have that  $f \in \mathcal{M}[X, \mathcal{A}]$  and  $\int^+ f d\mu$  is finite. Using 19.170 we have that

$$\alpha \cdot f \in \mathcal{M}[X, \mathcal{A}] \quad (19.419)$$

Consider the following cases for  $\alpha$

**$\alpha = 0$ .** then  $\alpha \cdot f = 0$  and non negative and  $\int^+ (\alpha \cdot f) d\mu = \int^+ C_0 d\mu \stackrel{19.205}{=} 0 < \infty$  so using 19.221 we have that  $\int (\alpha \cdot f) d\mu = \int^+ (\alpha \cdot f) d\mu = 0 = 0 \cdot \int f d\mu$

**0 <  $\alpha$ .** then we have that

$$(\alpha \cdot f)^+ \underset{19.176}{=} \alpha \cdot f^+ \text{ and } (\alpha \cdot f)^- \underset{19.176}{=} \alpha \cdot f^- \quad (19.420)$$

so

$$\begin{aligned} \int^+ (\alpha \cdot f)^+ d\mu &= \int^+ (\alpha \cdot f^+) d\mu \\ &\underset{19.209}{=} \alpha \cdot \int^+ f^+ d\mu < \infty \end{aligned} \quad (19.421)$$

$$\begin{aligned} \int^+ (\alpha \cdot f)^- d\mu &= \int^+ (\alpha \cdot f^-) d\mu \\ &\underset{19.209}{=} \alpha \cdot \int^+ f^- d\mu < \infty \end{aligned} \quad (19.422)$$

Using 19.419, 19.421 and 19.413 we have that

$$\alpha \cdot f \text{ is } \mu\text{-integrable}$$

and

$$\begin{aligned} \int (\alpha \cdot f) d\mu &= \int^+ (\alpha \cdot f)^+ d\mu + \int^+ (\alpha \cdot f)^- d\mu \\ &\underset{19.421, 19.422}{=} \alpha \cdot \int^+ f^+ d\mu + \alpha \cdot \int^+ f^- d\mu \\ &= \alpha \cdot \left( \int^+ f^+ d\mu + \int^+ f^- d\mu \right) \\ &= \alpha \cdot \int f d\mu \end{aligned}$$

**$\alpha < 0$ .** then we have that

$$(\alpha \cdot f)^+ \underset{19.176}{=} -\alpha \cdot f^- \text{ and } (\alpha \cdot f)^- \underset{19.176}{=} -\alpha \cdot f^+ \quad (19.423)$$

so

$$\begin{aligned} \int^+ (\alpha \cdot f)^+ d\mu &= \int^+ (-\alpha) \cdot f^- d\mu \\ &\underset{19.209}{=} (-\alpha) \cdot \int^+ f^- d\mu < \infty \end{aligned} \quad (19.424)$$

$$\begin{aligned} \int^+ (\alpha \cdot f)^- d\mu &= \int^+ (-\alpha) \cdot f^+ d\mu \\ &\underset{19.209}{=} (-\alpha) \cdot \int^+ f^+ d\mu < \infty \end{aligned} \quad (19.425)$$

Using 19.419, 19.424 and 19.425 we have that

$$\alpha \cdot f \text{ is } \mu\text{-integrable}$$

and

$$\begin{aligned} \int (\alpha \cdot f) d\mu &= \int^+ (\alpha \cdot f)^+ d\mu - \int^+ (\alpha \cdot f)^- d\mu \\ &\underset{19.424, 19.425}{=} (-\alpha) \cdot \int^+ f^- d\mu - (-\alpha) \cdot \int^+ f^+ d\mu \\ &= (-\alpha) \cdot \left( \int^+ f^- d\mu - \int^+ f^+ d\mu \right) \\ &= \alpha \cdot \left( \int^+ f^+ d\mu - \int^- f^- d\mu \right) \\ &= \alpha \cdot \int f d\mu \end{aligned}$$

2. As  $f, g$  are  $\mu$ -integrable we have that

$$f, g \in \mathcal{M}[X, \mathcal{A}] \xrightarrow{19.170} f + g \in \mathcal{M}[X, \mathcal{A}] \quad (19.426)$$

$$\int^+ f^+ d\mu < \infty \wedge \int^+ f^- d\mu < \infty \wedge \int^+ g^+ d\mu < \infty \wedge \int^+ g^- d\mu < \infty \quad (19.427)$$

Further as we have by 19.176 that  $\forall x \in X$

$$(f + g)^+(x) \leq f^+(x) + g^+(x) \text{ and } (f + g)^- \leq f^-(x) + g^-(x)$$

which using 19.209 proves that

$$\begin{aligned} \int^+ (f + g)^+ d\mu &\leq \int^+ (f^+ + g^+) d\mu = \int^+ f^+ + \int^+ g^+ d\mu < \infty \\ \int^+ (f + g)^- d\mu &\leq \int^+ (f^- + g^-) d\mu = \int^+ f^- d\mu + \int^+ g^- d\mu < \infty \end{aligned}$$

The above together with 19.426 proves that

$$f + g \text{ is } \mu\text{-integrable}$$

Next we have that

$$\begin{aligned} (f + g)^+ - (f + g)^- &\stackrel{19.176}{=} f + g \\ &\stackrel{19.176}{=} (f^+ - f^-) + (g^+ - g^-) \\ &= (f^+ + g^+) - (f^- + g^-) \end{aligned} \quad (19.428)$$

Finally

$$\begin{aligned} \int (f + g) d\mu &\stackrel{\text{def}}{=} \int^+ (f + g)^+ d\mu - \int^+ (f + g)^- d\mu \\ &\stackrel{19.224 \text{ and } 19.428}{=} \int^+ (f^+ + g^+) d\mu - \int^+ (f^- + g^-) \\ &= \int^+ f^+ d\mu + \int^+ g^+ d\mu - \int^+ f^- d\mu - \int^+ g^- d\mu \\ &\stackrel{\text{integrals are real}}{=} \left( \int^+ f^+ d\mu - \int^+ f^- d\mu \right) + \left( \int^+ g^+ d\mu - \int^+ g^- d\mu \right) \\ &= \int f d\mu + \int g d\mu \end{aligned}$$

3. If  $\forall x \in X f(x) \leq g(x)$  we have  $0 \leq g(x) - f(x) = (g - f)(x)$  hence  $g - f$  is non negative, further using (1),(2) we have that  $f - g$  is  $\mu$ -integrable. From 19.221 we have that

$$0 \leq \int^+ (g - f) d\mu = \int (g - f) d\mu \quad (19.429)$$

Further

$$\begin{aligned} \int g d\mu - \int f d\mu &\stackrel{(1)}{=} \int g d\mu + \int (-f) d\mu \\ &\stackrel{(2)}{=} \int (g - f) d\mu \\ &\stackrel{19.429}{\geq} 0 \end{aligned}$$

proving that

$$\int f d\mu \leq \int g d\mu \quad \square$$

**Proposition 19.226.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space and  $f \in \overline{\mathcal{M}[X, \mathcal{A}]}$  then  $f$  is  $\mu$ -integrable if and only if  $|f|$  is  $\mu$ -integrable. If  $f$  or  $|f|$  is  $\mu$ -integrable then  $|\int f d\mu| \leq \int |f| d\mu$

**Proof.** First we prove the equivalence:

$\Rightarrow$ . If  $f$  is  $\mu$ -integrable then  $\int^+ f^+ d\mu < \infty$  and  $\int^+ f^- d\mu < \infty$  so that by 19.221  $f^+$  and  $f^-$  are  $\mu$ -integrable with  $\int^+ f^+ d\mu = \int f^+ d\mu$  and  $\int^+ f^- d\mu = \int f^- d\mu$ . As  $|f| \stackrel{19.176}{=} f^+ + f^-$  it follows from 19.225 that  $|f|$  is  $\mu$ -integrable.

$\Leftarrow$ . Assume that  $|f|$  is  $\mu$ -integrable then as  $|f|$  is non negative we have by 19.221 that  $\int^+ |f| d\mu < \infty$ . If  $x \in X$  then by 19.176 we have  $f^+(x) \leq |f(x)|$  and  $f^-(x) \leq |f(x)|$ , so using 19.209 we have that  $\int^+ f^+ d\mu \leq \int^+ |f| d\mu < \infty$  and  $\int^+ f^- d\mu \leq \int^+ |f| d\mu < \infty$ , proving that  $f$  is  $\mu$ -integrable.

Finally

$$\begin{aligned} \left| \int f d\mu \right| &= \left| \int^+ f^+ d\mu - \int^+ f^- d\mu \right| \\ &\leq \left| \int^+ f^+ d\mu \right| + \left| \int^+ f^- d\mu \right| \\ &\leq \int^+ f^+ d\mu + \int^+ f^- d\mu \\ &\stackrel{19.209}{=} \int^+ (f^+ + f^-) d\mu \\ &= \int^+ |f| d\mu \\ &\stackrel{19.221}{=} \int |f| d\mu \end{aligned}$$

□

**Lemma 19.227.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space and let  $f, g \in \overline{\mathcal{M}^+[X, \mathcal{A}]}$  such that  $f(x) = g(x)$   $\mu$ -a.e. then

$$\int^+ f d\mu = \int^+ g d\mu$$

**Proof.** Define

$$H = \{x \in X \mid f(x) \neq g(x)\} = \{f \neq g\}$$

As  $f, g$  are  $\mathcal{A}$ -measurable we have by 19.162 that  $H \in \mathcal{A}$ , further as  $f = g$   $\mu$ -always everywhere there exists a  $N \in \mathcal{A}$  with  $\mu(N) = 0$  such that  $H \subseteq N$ , hence  $\mu(H) \leq \mu(N) = 0$ . To summarize

$$H \in \mathcal{A} \text{ and } \mu(H) = 0 \tag{19.430}$$

Define the non negative function  $h: X \rightarrow [0, \infty]$  by

$$h(x) = \begin{cases} \infty & \text{if } x \in H \\ 0 & \text{if } x \in X \setminus H \end{cases}$$

Let  $a \in \mathbb{R}$  then we have either

$0 \leq a$ . Then if  $x \in \{h > a\}$  we have  $h(x) > a \geq 0$  so that  $x \in H$  [otherwise  $h(x) = 0 \not> 0$ ] so that  $\{h > a\} \subseteq H$ . If  $x \in H$  then  $h(x) = \infty > a$  so that  $H \subseteq \{h > a\}$ . So we have that

$$\{h > a\} = H \in \mathcal{A} \tag{19.431}$$

$a < 0$ . Then if  $x \in X$  we have either  $x \in H$  so that  $h(x) = \infty > a$  or  $x \in X \setminus H$  giving  $h(x) = 0 > a$ . Hence  $X \subseteq \{h > a\} \subseteq X$  proving

$$\{h > a\} = X \in \mathcal{A}$$

As in all cases we have  $\{h > a\} \in \mathcal{A}$  it follows that

$$h \text{ is } \mathcal{A}\text{-measurable} \tag{19.432}$$

Let  $n \in \mathbb{N}_0$  and define  $h_n = n \cdot \chi_H = \sum_{i=1}^1 a_i \cdot \chi_{A_i}$  where  $a_1 = n$ ,  $A_1 = H$  then

$$h_n \in \mathcal{S}^+[X, \mathcal{A}] \text{ and } \int^S h_n d\mu = n \cdot \mu(H) \stackrel{19.430}{=} 0 \quad (19.433)$$

Further given  $n \in \mathbb{N}_0$  and  $x \in X$  we have

$$h_n(x) = \begin{cases} n & \text{if } x \in H \\ 0 & \text{if } x \in X \setminus H \end{cases} \leq \begin{cases} n+1 & \text{if } x \in H \\ 0 & \text{if } x \in X \setminus H \end{cases} = h_{n+1}(x) \quad (19.434)$$

Let  $x \in X$  then we have either

$$x \in H. \text{ then } \lim_{n \rightarrow \infty} h_n(x) = \lim_{n \rightarrow \infty} n \stackrel{17.75}{=} \infty = h(x)$$

$$x \in X \setminus H. \text{ then } \lim_{n \rightarrow \infty} h_n(x) = \lim_{n \rightarrow \infty} 0 = 0 = h(x)$$

proving that  $\lim_{n \rightarrow \infty} h_n(x) = h(x)$ , which together with 19.433 and 19.434 allows us to use proposition 19.207 to get that

$$\int^+ h d\mu = \lim_{n \rightarrow \infty} \int^S h_n d\mu \stackrel{19.433}{=} \lim_{n \rightarrow \infty} 0 = 0 \quad (19.435)$$

Further if  $x \in X$  then we have either

$$x \in H. \text{ then } f(x) \leq \infty = g(x) + \infty = g(x) + h(x) \text{ and } g(x) \leq \infty = f(x) + \infty = f(x) + h(x)$$

$$x \in X \setminus H. \text{ then } f(x) = g(x) + 0 = g(x) + h(x) \text{ and } g(x) = f(x) + 0 = f(x) + h(x)$$

which proves that in all cases  $f(x) \leq g(x) + h(x)$  and  $g(x) \leq f(x) + h(x)$ . So using 19.209 we have that  $\int^+ f d\mu \leq \int^+ g d\mu + \int^+ h d\mu \stackrel{19.435}{=} \int^+ g d\mu$  and  $\int^+ g d\mu \leq \int^+ f d\mu + \int^+ h d\mu \stackrel{19.435}{=} \int^+ f d\mu$ . Hence we have that

$$\int^+ g d\mu = \int^+ f d\mu \quad (19.436) \quad \square$$

**Proposition 19.228.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space and let  $f, g \in \overline{\mathcal{M}[X, \mathcal{A}]}$  such that  $f(x) = g(x)$   $\mu$ -a.e.. If  $f$  is  $\mu$ -integrable then  $g$  is  $\mu$ -integrable and  $\int f d\mu = \int g d\mu$

**Proof.** Define

$$H = \{x \in X \mid f(x) \neq g(x)\} = \{f \neq g\}$$

As  $f, g$  are  $\mathcal{A}$ -measurable we have by 19.162 that  $H \in \mathcal{A}$ , further as  $f = g$   $\mu$ -always everywhere there exists a  $N \in \mathcal{A}$  with  $\mu(N) = 0$  such that  $H \subseteq N$ , hence  $\mu(H) \leq \mu(N) = 0$ . To summarize

$$H \in \mathcal{A} \text{ and } \mu(H) = 0 \quad (19.437)$$

As  $f, g$  are  $\mathcal{A}$ -measurable we have by 19.177 that  $f^+, f^-, g^+, h^+$  are  $\mathcal{A}$ -measurable. Let  $x \in X$  then

$$\text{if } f^+(x) \neq g^+(x) \Rightarrow \max(f(x), 0) \neq \max(g(x), 0) \Rightarrow f(x) \neq g(x)$$

$$\text{if } f^-(x) = g^-(x) \Rightarrow \min(f(x), 0) \neq \min(g(x), 0) \Rightarrow f(x) \neq g(x)$$

proving that  $\{x \in X \mid f^+(x) \neq g^+(x)\} \subseteq H$  and  $\{x \in X \mid f^-(x) \neq g^-(x)\} \subseteq H$  which as by 19.437  $\mu(H) = 0$  proves that

$$f^+ = g^+ \text{ } \mu\text{-a.e. and } f^- = g^- \text{ } \mu\text{-a.e.}$$

Using then 19.227 we have [taking in account also that  $f$  is  $\mu$ -integrable]

$$\int^+ g^+ d\mu \stackrel{19.227}{=} \int^+ f^+ d\mu < \infty \text{ and } \int^+ g^- d\mu \stackrel{19.227}{=} \int^+ f^- d\mu < \infty$$

So we have that

$$g \text{ is } \mu\text{-integrable and } \int f d\mu = \int^+ f^+ d\mu - \int^+ f^- d\mu = \int^+ g^+ d\mu - \int^+ g^- d\mu = \int g d\mu \quad \square$$

**Proposition 19.229.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measurable space,  $t \in ]0, \infty[$  and  $f \in \overline{\mathcal{M}^+[X, \mathcal{A}]}$  [a non negative measurable extended real values function] then for  $A_t = \{f \geq t\} \stackrel{\text{def}}{=} \{x \in X \mid f(x) \geq t\}$  we have

$$\mu(A_t) \leq \frac{1}{t} \cdot \int_{A_t}^+ f d\mu \leq \frac{1}{t} \cdot \int^+ f d\mu$$

**Proof.** Consider the function  $t \cdot \mathcal{X}_{A_t} = \sum_{i=1}^1 a_i \cdot \mathcal{X}_{A_i}$  where  $a_1 = t$  and  $A_1 = A_t$  then as  $A_t = \{f \geq t\} \in \mathcal{A}$  [as  $f$  is  $\mathcal{A}$ -measurable] we have that

$$t \cdot \mathcal{X}_{A_t} \in \mathcal{S}^+[X, \mathcal{A}] \text{ and } \int^+ (t \cdot \mathcal{X}_{A_t}) d\mu = \sum_{i=1}^1 a_i \mu(A_i) = t \cdot \mu(A_t) < \infty \quad (19.438)$$

Let  $x \in X$  then either

$x \in A_t$ . then  $0 < t \leq f(x)$  so that  $0 \leq t \cdot \mathcal{X}_{A_t}(x) \leq f(x) \cdot \mathcal{X}_{A_t}(x) \leq f(x)$  [as  $\mathcal{X}_{A_t}(x) = 1$ ]

$x \in X \setminus A_t$ . then  $\mathcal{X}_{A_t}(x) = 0$  so that trivially  $0 \leq t \cdot \mathcal{X}_{A_t}(x) \leq f(x) \cdot \mathcal{X}_{A_t}(x) \leq f(x)$

proving that  $0 \leq t \cdot \mathcal{X}_{A_t} \leq f(x) \cdot \mathcal{X}_{A_t} \leq f(x)$  so that by 19.209

$$t \cdot \mu(A_t) \stackrel{19.438}{=} \int^+ (t \cdot \mathcal{X}_{A_t}) d\mu \leq \int^+ \mathcal{X}_{A_t} \cdot f d\mu \leq \int^+ f d\mu$$

proving that

$$\mu(A_t) \leq \frac{1}{t} \cdot \int_{A_t}^+ f d\mu \leq \frac{1}{t} \cdot \int^+ f d\mu \quad \square$$

For the next corollary we introduce the concept of  $\sigma$ -finite sets

**Definition 19.230.** Let  $\langle X, \mathcal{A}, \mu \rangle$  then  $A \subseteq X$  is  **$\sigma$ -finite under  $\mu$**  if there exists a sequence  $\{A_i\}_{i \in \mathbb{N}_0} \subseteq \mathcal{A}$  with  $\forall n \in \mathbb{N}_0 \mu(A_n) < \infty$  such that  $A = \bigcup_{n \in \mathbb{N}_0} A_n$

**Corollary 19.231.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space and let  $f \in \mathcal{L}[X, \mathcal{A}, \mu, \overline{\mathbb{R}}]$  [ $f$  is a extended real valued  $\mu$ -integrable function] then  $\{x \in X \mid f(x) \neq 0\}$  is  $\sigma$ -finite under  $\mu$

**Proof.** Let  $n \in \mathbb{N}_0$  and define

$$A_n = \left\{ x \in X \mid |f(x)| \geq \frac{1}{n} \right\} \quad (19.439)$$

As  $f$  is  $\mu$ -integrable [implying that  $f \in \overline{\mathcal{M}[X, \mathcal{A}]}$ ] we have by 19.226 that  $|f|$  is  $\mu$ -integrable so that  $\int^+ |f| d\mu < \infty$ . Combining this with 19.229 gives

$$\mu(A_n) \leq n \cdot \int^+ |f| d\mu < \infty$$

Let  $x \in \{x \in X \mid f(x) \neq 0\}$  then  $0 < |f(x)|$  so that by 17.8 there exists a  $n \in \mathbb{N}_0$  such that  $0 < \frac{1}{n} < |f(x)|$  proving that  $x \in A_n$ . Hence  $\{x \in X \mid f(x) \neq 0\} \subseteq \bigcup_{n \in \mathbb{N}_0} A_n$ . On the other side if  $x \in \bigcup_{n \in \mathbb{N}_0} A_n$  there exists a  $n \in \mathbb{N}_0$  such that  $x \in A_n$ , hence  $|f(x)| \geq \frac{1}{n} > 0$  giving  $f(x) \neq 0$  which proves that  $\bigcup_{n \in \mathbb{N}_0} A_n \subseteq \{x \in X \mid f(x) \neq 0\}$ . So

$$\{x \in X \mid f(x) \neq 0\} = \bigcup_{n \in \mathbb{N}_0} A_n \quad \square$$

**Corollary 19.232.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space,  $f \in \overline{\mathcal{M}[X, \mathcal{A}]}$  [hence by 19.177  $|f| \in \overline{\mathcal{M}^+[X, \mathcal{A}]}$ ] such that  $\int^+ |f| d\mu = 0$  then  $f(x) = 0$   $\mu$ -a.e.. Further if  $f$  is  $\mu$ -integrable with  $\int |f| d\mu = 0$  then  $f = 0$   $\mu$ -a.e.

**Proof.** For the first part let  $f \in \overline{\mathcal{M}[X, \mathcal{A}]}$  with  $\int^+ |f| = 0$ . Let  $n \in \mathbb{N}_0$  then using 19.229 we have that

$$\mu\left(\left\{x \in X \mid |f(x)| \geq \frac{1}{n}\right\}\right) \leq n \cdot \int^+ |f| d\mu \quad (19.440)$$

Let  $x \in \{x \in X \mid f(x) \neq 0\}$  then  $0 < |f(x)|$  so that by 17.8 there exists a  $n \in \mathbb{N}_0$  such that  $0 < \frac{1}{n} < |f(x)|$  proving that  $x \in A_n$ . Hence  $\{x \in X \mid f(x) \neq 0\} \subseteq \bigcup_{n \in \mathbb{N}_0} \{x \in X \mid |f(x)| \geq \frac{1}{n}\}$ . On the other hand if  $x \in \bigcup_{n \in \mathbb{N}_0} \{x \in X \mid |f(x)| \geq \frac{1}{n}\}$  there exists a  $n \in \mathbb{N}_0$  such that  $x \in \{x \in X \mid |f(x)| \geq \frac{1}{n}\}$ , hence  $|f(x)| \geq \frac{1}{n} > 0$  giving  $f(x) \neq 0$  which proves that  $\bigcup_{n \in \mathbb{N}_0} \{x \in X \mid |f(x)| \geq \frac{1}{n}\} \subseteq \{x \in X \mid f(x) \neq 0\}$ , So

$$\{x \in X \mid f(x) \neq 0\} = \bigcup_{n \in \mathbb{N}_0} \left\{ x \in X \mid |f(x)| \geq \frac{1}{n} \right\} \in \mathcal{A}$$

Finally

$$\mu(\{x \in X \mid f(x) \neq 0\}) = \mu\left(\left\{x \in X \mid |f(x)| \geq \frac{1}{n}\right\}\right) \leq \sum_{i=1}^{\infty} \mu\left(\left\{x \in X \mid |f(x)| \geq \frac{1}{n}\right\}\right) = 0$$

which proves that

$$f = 0 \text{ } \mu\text{-a.e.}$$

Finally if  $f$  is  $\mu$ -integrable with  $\int |f| d\mu = 0$  then by 19.221  $\int^+ |f| d\mu = \int^- |f| d\mu = 0$  so that by the first part we have that  $f = 0$   $\mu$ -a.e. again.  $\square$

**Corollary 19.233.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space and let  $f \in \mathcal{L}[X, \mathcal{A}, \mu, \bar{\mathbb{R}}]$  [ $f$  is a extended real valued  $\mu$ -integrable function] such that  $\forall A \in \mathcal{A}$  we have  $\int_A f d\mu \geq 0$  then  $f \geq 0$  holds  $\mu$ -a.e..

**Proof.** Let  $A = \{x \in X \mid 0 \leq f(x)\} = \{x \in X \mid f(x) < 0\} = \{f < 0\} \in \mathcal{A}$ , If  $x \in X$  then we have either  $x \in A$ . Then  $f(x) < 0$  and  $\mathcal{X}_A(x) = 1$  so that  $(\mathcal{X}_A \cdot f)(x) = f(x) < 0$  hence

$$(\mathcal{X}_A \cdot f)^+(x) = \max((\mathcal{X}_A \cdot f)(x), 0) = 0$$

$x \in X \setminus A$ . Then  $\mathcal{X}_A(x) = 0$  so that  $\max((\mathcal{X}_A \cdot f)(x), 0) = \max(0, 0) = 0$  hence

$$(\mathcal{X}_A \cdot f)^+(x) = \max((\mathcal{X}_A \cdot f)(x), 0) = 0$$

which proves that  $(\mathcal{X}_A \cdot f)^+ = C_0$  so that by 19.205

$$\int^+ (\mathcal{X}_A \cdot f)^+ d\mu = 0 \tag{19.441}$$

So

$$0 \leq \int_A f d\mu = \int (\mathcal{X}_A \cdot f) d\mu = \int^+ (\mathcal{X}_A \cdot f)^+ d\mu - \int^+ (\mathcal{X}_A \cdot f)^- d\mu = - \int^+ (\mathcal{X}_A \cdot f)^- d\mu \leq 0$$

proving that

$$\int^+ (\mathcal{X}_A \cdot f)^- d\mu = 0 \tag{19.442}$$

Applying 19.221 on 19.441 and 19.442 it follows that

$$\int (\mathcal{X}_A \cdot f)^+ d\mu = \int (\mathcal{X}_A \cdot f^-) d\mu = 0$$

so

$$\int |\mathcal{X}_A \cdot f| d\mu \stackrel{19.176}{=} \int ((\mathcal{X}_A \cdot f)^+ + (\mathcal{X}_A \cdot f)^-) \stackrel{19.225}{=} \int (\mathcal{X}_A \cdot f)^+ d\mu + \int (\mathcal{X}_A \cdot f)^- d\mu = 0$$

Using the above with the previous corollary 19.232 proves that  $\mathcal{X}_A \cdot f = 0$   $\mu$ -a.e.. Hence there exists a  $N \in \mathcal{N}$  with  $\mu(N) = 0$  such that  $\{x \in X \mid (\mathcal{X}_A \cdot f)(x) \neq 0\} \subseteq N$ . If  $x \in A$  then  $f(x) < 0$  and  $\mathcal{X}_A(x) = 1$  proving that  $(\mathcal{X}_A \cdot f)(x) < 0 \Rightarrow (\mathcal{X}_A \cdot f)(x) \neq 0$  hence  $x \in N$  giving  $A \subseteq N$ . Hence as  $A = \{x \in X \mid 0 \leq f(x)\}$  we have

$$f \geq 0 \text{ holds } \mu\text{-a.e.} \quad \square$$

**Corollary 19.234.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space and  $f \in \mathcal{L}[X, \mathcal{A}, \mu, \bar{\mathbb{R}}]$  [ $f$  is a extended real valued  $\mu$ -integrable function] then  $\mu(\{x \in X \mid |f(x)| = \infty\}) = 0$  so that  $|f| < \infty$   $\mu$ -a.e..

**Proof.** As  $f$  is  $\mu$ -integrable we have by 19.226 that  $|f|$  is  $\mu$ -integrable, using 19.229 we have that

$$\forall n \in \mathbb{N}_0 \text{ we have } \mu(\{x \in X \mid |f(x)| \geq n\}) \leq \frac{1}{n} \cdot \int^+ |f| d\mu \stackrel{19.221}{=} \frac{1}{n} \cdot \int |f| d\mu \quad (19.443)$$

Let  $n \in \mathbb{N}_0$  then if  $x \in \{x \in X \mid |f(x)| = \infty\}$  then  $|f(x)| = \infty > n$  so that  $\{x \in X \mid |f(x)| = \infty\} \subseteq \{x \in X \mid |f(x)| \geq n\}$ , as  $\{x \in X \mid f(x) = \infty\} = \{f = \infty\} \in \mathcal{A}$  [see 19.162], we have

$$\mu(\{x \in X \mid |f(x)| = \infty\}) \leq \mu(\{x \in X \mid |f(x)| \geq n\}) \leq \frac{1}{n} \cdot \int |f| d\mu$$

As  $n$  is chosen arbitrary and  $\int |f| d\mu < \infty$  we have by 9.57 that

$$\mu(\{x \in X \mid |f(x)| = \infty\}) = 0$$

proving that

$$|f| < \infty \text{ } \mu\text{-a.e.} \quad \square$$

**Corollary 19.235.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space and  $f \in \overline{\mathcal{M}[X, \mathcal{A}]}$  then  $f \in \mathcal{L}[X, \mathcal{A}, \mu, \overline{\mathbb{R}}]$  if and only if there exists a  $g \in \mathcal{L}[X, \mathcal{A}, \mu, \mathbb{R}]$  such that  $f(x) = g(x)$   $\mu$ -a.e..

In other words a  $\mathcal{A}$ -measurable extended real valued function is  $\mu$ -integrable if there exists a real valued  $\mu$ -integrable function that is equal to  $f$  almost everywhere.

**Proof.**

$\Leftarrow$ . Assume the there exists a function  $g \in \mathcal{L}[X, \mathcal{A}, \mu, \mathbb{R}]$  such that  $f(x) = g(x)$   $\mu$ -a.e., then as  $\mathcal{L}[X, \mathcal{A}, \mu, \mathbb{R}] \subseteq \mathcal{L}[X, \mathcal{A}, \mu, \overline{\mathbb{R}}]$  we have by 19.228 that  $f$  is  $\mu$ -integrable.

$\Rightarrow$ . Suppose that  $f \in \mathcal{L}[X, \mathcal{A}, \mu, \overline{\mathbb{R}}]$  then we must create a  $g$  that is real valued,  $\mu$ -integrable and  $\mu$ -a.e. to  $f$ . To do this define the set  $A$  of all elements of  $X$  where  $f(x)$  is either  $+\infty$  or  $-\infty$ .

$$A = \{x \in X \mid |f(x)| = \infty\} = \{x = \infty\}$$

then as  $f$  is  $\mathcal{A}$ -measurable [because  $f \in \mathcal{L}[X, \mathcal{A}, \mu, \overline{\mathbb{R}}]$ ] we have by 19.162 that  $A \in \mathcal{A}$  and thus also  $X \setminus A \in \mathcal{A}$ . Furthermore using 19.234 we have that  $\mu(A) = 0$ . To summarize

$$A \in \mathcal{A} \wedge X \setminus A \in \mathcal{A} \wedge \mu(A) = 0 \quad (19.444)$$

Define now  $g = \chi_{X \setminus A} \cdot f$  then as  $\chi_{X \setminus A}$  is  $\mathcal{A}$ -measurable [see 19.157] and  $f$  is  $\mathcal{A}$ -measurable it follows from 19.173 that

$$g = \chi_{X \setminus A} \cdot f \text{ is } \mathcal{A}\text{-measurable} \quad (19.445)$$

Further if  $x \in X$  then either  $x \in A$  giving  $g(x) \underset{x \notin X \setminus A}{=} 0 \in \mathbb{R}$  or  $x \in X \setminus A$  so that  $g(x) = f(x) \in \mathbb{R}$  proving together with 19.445 that

$$g \in \mathcal{M}[X, \mathcal{A}] \quad (19.446)$$

Further using 19.182, 19.444 we have that

$$f = g \text{ } \mu\text{-always measurable} \quad (19.447)$$

Finally using the above,  $g \in \mathcal{M}[X, \mathcal{A}] \subseteq \overline{\mathcal{M}[X, \mathcal{A}]}$  [see 19.446] and  $f$  is  $\mu$ -integrable allows us to use corollary 19.228 to prove that  $g$  is  $\mu$ -integrable or together with 19.446 that

$$g \in \mathcal{L} \quad \square$$

### 19.6.5 Limit theorems for integrable functions

**Lemma 19.236.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space,  $f \in \overline{\mathcal{M}^+[X, \mathcal{A}]}$  and  $\{f_i\}_{i \in \mathbb{N}_0} \subseteq \overline{\mathcal{M}^+[X, \mathcal{A}]}$  so that for  $x \in X$

$$1. \forall n \in \mathbb{N}_0 \quad f_n(x) \leq f_{n+1}(x)$$

$$2. f(x) = \lim_{n \rightarrow \infty} f_n(x)$$

then

$$\int^+ f d\mu = \lim_{n \rightarrow \infty} \int^+ f_n d\mu$$

**Proof.** Let  $x \in X$  then from (1) we have by 17.71 that  $f(x) = \sup(\{f_n(x) | n \in \mathbb{N}_0\})$  so that

$$\forall n \in \mathbb{N}_0 \text{ we have } f_n(x) \leq f(x) \quad (19.448)$$

From (2) and the above we have using 19.209 that

$$\forall n \in \mathbb{N}_0 \text{ we have } \int^+ f_n d\mu \leq \int^+ f_{n+1} d\mu \leq \int^+ f d\mu \quad (19.449)$$

Using the above together with 17.71 gives that  $\lim_{n \rightarrow \infty} \int^+ f_n d\mu$  exists and

$$\lim_{n \rightarrow \infty} \int^+ f_n d\mu = \sup \left( \left\{ \int^+ f_n d\mu | n \in \mathbb{N}_0 \right\} \right) \leq \int^+ f d\mu \quad (19.450)$$

For the reverse inequality take  $n \in \mathbb{N}_0$  then using 19.178 there exists a sequence  $\{g_{n,i}\}_{i \in \mathbb{N}_0} \subseteq \mathcal{S}^+[X, \mathcal{A}]$  such that for  $x \in X$  we have

$$\forall i \in \mathbb{N}_0 \quad g_{n,i}(x) \leq g_{n,i+1}(x), \quad \lim_{i \rightarrow \infty} g_{n,i}(x) = \sup(\{g_{n,i}(x) | i \in \mathbb{N}_0\}) = f_n(x) \quad (19.451)$$

Define now  $\{h_n\}_{n \in \mathbb{N}_0}$  by

$$h_n = \max(\{g_{i,n} | i \in \{1, \dots, n\}\}) \quad (19.452)$$

then by 19.198 we have that

$$\forall n \in \mathbb{N}_0 \text{ we have } h_n \in \mathcal{S}^+[X, \mathcal{A}] \quad (19.453)$$

Let  $x \in X$  and  $n \in \mathbb{N}_0$  then  $\forall i \in \{1, \dots, n\}$  we have

$$\begin{aligned} g_{i,n}(x) &\leq g_{i,n+1}(x) \\ &\leq \max(\{g_{i,n+1}(x) | i \in \{1, \dots, n+1\}\}) \\ &= \max(\{g_{i,n+1} | i \in \{1, \dots, n+1\}\})(x) \\ &= h_{n+1}(x) \end{aligned}$$

so that

$$h_n(x) = \max(\{g_{i,n} | i \in \{1, \dots, n\}\}) \leq h_{n+1}(x) \quad (19.454)$$

Further as  $\forall i \in \{1, \dots, n\}$   $g_{i,n}(x) \leq \sup(\{g_{i,j}(x) | j \in \mathbb{N}_0\}) = f_i(x) \leq f_n(x)$  so that

$$h_n(x) = \max(\{g_{i,n}(x) | i \in \{1, \dots, n\}\}) \leq f_n(x) \quad (19.455)$$

So for  $n \in \mathbb{N}_0$  we have  $h_n(x) \leq f_n(x) \leq 19.448 f(x)$ , giving  $\sup(\{h_n(x) | n \in \mathbb{N}_0\}) \leq f(x)$  which using 19.454 and 17.71 proves that

$$\lim_{n \rightarrow \infty} h_n(x) = \sup(\{h_n(x) | n \in \mathbb{N}_0\}) \leq f(x) \quad (19.456)$$

Assume that  $\sup(\{h_n(x) | n \in \mathbb{N}_0\}) < f(x)$  then as  $f(x) = \sup(\{f_n(x) | x \in \mathbb{N}\})$  there exists a  $n \in \mathbb{N}_0$  such that

$$\sup(\{h_n(x) | n \in \mathbb{N}_0\}) < f_n(x)$$

further as  $f_n(x) = \sup(\{g_{n,i}(x) | i \in \mathbb{N}_0\})$  there exists a  $k \in \mathbb{N}_0$  such that

$$\sup(\{h_n(x) | n \in \mathbb{N}_0\}) < g_{n,k}(x) \leq g_{n,\max(n,k)}(x)$$

which as

$$g_{n,\max(n,k)}(x) \leq \max(\{g_{i,\max(n,k)} | i \in \{1, \dots, \max(n, k)\}\}) = h_{n,\max(n,k)}(x)$$

leads to the contradiction  $\sup(\{h_n(x) | n \in \mathbb{N}_0\}) < h_{n,\max(n,k)}(x)$ . So the assumption must be wrong, giving  $f(x) \leq \sup(\{h_n(x) | n \in \mathbb{N}_0\})$  which combined with 19.456 proves that

$$\lim_{n \rightarrow \infty} h_n(x) = \sup(\{h_n(x) | n \in \mathbb{N}_0\}) = f(x) \quad (19.457)$$

Next as  $\{h_n\}_{n \in \mathbb{N}_0} \subseteq \mathcal{S}^+[X, \mathcal{A}]$  [see 19.453] we have using 19.454 and 19.457 that by 19.208

$$\begin{aligned} \int^+ f d\mu &\stackrel{19.208}{=} \lim_{n \rightarrow \infty} \int^S h_n d\mu \\ &\stackrel{19.204}{=} \lim_{n \rightarrow \infty} \int^+ h_n d\mu \\ &\leq \lim_{n \rightarrow \infty} \int^+ f_n d\mu \quad [h_n(x) \text{ see 19.455 and 19.209}] \end{aligned}$$

which together with 19.450 proves that

$$\int^+ f d\mu = \lim_{n \rightarrow \infty} \int^+ f_n d\mu$$

□

We release now slithly the conditions for the above lemma giving the Monotone Convergence Theorem.

**Theorem 19.237. (Monotone Convergence Theorem)** *Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space,  $n \in \mathbb{N}_0$ ,  $f \in \overline{\mathcal{M}^+[X, \mathcal{A}]}$  and  $\{f_i\}_{i \in \{1, \dots, n\}} \subseteq \overline{\mathcal{M}^+[X, \mathcal{A}]}$  such that*

1.  $\forall n \in \mathbb{N}_0 \ f_n \leq f_{n+1}$
2.  $f = \lim_{n \rightarrow \infty} f_n$

$\mu$ -a.e. then we have

$$\int^+ f d\mu = \lim_{n \rightarrow \infty} \int^+ f_n d\mu$$

**Proof.** Using the assumption there exists a  $N \in \mathcal{A}$  with  $\mu(N)$  such that

$$A = \left\{ x \in X \mid \neg \left( (\forall n \in \mathbb{N}_0 f_n(x) \leq f_{n+1}(x)) \wedge \left( f(x) = \lim_{n \rightarrow \infty} f_n(x) \right) \right) \right\} \subseteq N$$

Then if  $x \in X$  we have either

$x \in X \setminus A$ . then

$$\forall n \in \mathbb{N}_0 \ (\mathcal{X}_{X \setminus A} \cdot f)(x) = \mathcal{X}_{X \setminus A}(x) \cdot f_n(x) \leq \mathcal{X}_{X \setminus A}(x) \cdot f_{n+1}(x) = (\mathcal{X}_{X \setminus A} \cdot f_{n+1})(x)$$

and

$$\begin{aligned} (\mathcal{X}_A \cdot f)(x) &= \mathcal{X}_{X \setminus A}(x) \cdot f(x) \\ &= f(x) \\ &= \lim_{n \rightarrow \infty} f_n(x) \\ &= \lim_{n \rightarrow \infty} (\mathcal{X}_{X \setminus A}(x) \cdot f_n(x)) \\ &= \lim_{n \rightarrow \infty} (\mathcal{X}_{X \setminus A} \cdot f_n)(x) \end{aligned}$$

$x \in A$ . then

$$\forall n \in \mathbb{N}_0 \ (\mathcal{X}_{X \setminus A} \cdot f)(x) = 0 \cdot f_n(x) = 0 \leq 0 \cdot f_{n+1}(x) = \mathcal{X}_{X \setminus A}(x) \cdot f_{n+1}(x) = (\mathcal{X}_{X \setminus A} \cdot f_{n+1})(x)$$

and

$$\begin{aligned} (\mathcal{X}_{X \setminus A} \cdot f)(x) &= 0 \cdot f(x) \\ &= 0 \\ &= \lim_{n \rightarrow \infty} (0 \cdot f_n(x)) \\ &= \lim_{n \rightarrow \infty} (\mathcal{X}_{X \setminus A}(x) \cdot f_n(x)) \\ &= \lim_{n \rightarrow \infty} (\mathcal{X}_{X \setminus A} \cdot f_n)(x) \end{aligned}$$

Hence giving  $x \in X$  we have  $\forall n \in \mathbb{N}_0$  that  $(\mathcal{X}_{X \setminus A} \cdot f_n)(x) \leq (\mathcal{X}_{X \setminus A} \cdot f_{n+1})(x)$  and  $(\mathcal{X}_{X \setminus A} \cdot f)(x) = \lim_{n \rightarrow \infty} (\mathcal{X}_{X \setminus A} \cdot f_n)(x)$  which allows us to apply the previous lemma 19.236 giving

$$\int^+ (\mathcal{X}_{X \setminus A} \cdot f) d\mu = \lim_{n \rightarrow \infty} \int^+ (\mathcal{X}_{X \setminus A} \cdot f_n) d\mu \quad (19.458)$$

Using 19.182 it follows that

$$f = (\mathcal{X}_{X \setminus A} \cdot f) \text{ } \mu\text{-a.e.}$$

and

$$\forall n \in \mathbb{N}_0 \text{ } (\mathcal{X}_{X \setminus A} \cdot f_n) = f_n \text{ } \mu\text{-a.e.}$$

From this it follows that

$$\int^+ f d\mu \stackrel{19.227}{=} \int^+ (\mathcal{X}_{X \setminus A} \cdot f) d\mu \text{ and } \forall n \in \mathbb{N}_0 \int^+ (\mathcal{X}_{X \setminus A} \cdot f_n) d\mu = \int^+ f_n d\mu$$

which substituted in 19.458 proves finally that

$$\int^+ f d\mu = \lim_{n \rightarrow \infty} \int^+ f_n d\mu \quad \square$$

**Corollary 19.238. (Beppo Levi's Theorem)** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space,  $\{f_i\}_{i \in \mathbb{N}_0} \subseteq \overline{\mathcal{M}^+[X, \mathcal{A}]}$  then for  $\sum_{i=1}^{\infty} f_i$  defined by  $(\sum_{i=1}^n f_i)(x) = \sum_{i=1}^{\infty} f_i(x)$  [see 17.94] we have

$$\int^+ \left( \sum_{i=1}^{\infty} f_i \right) d\mu = \sum_{i=1}^{\infty} \int^+ f_i d\mu$$

**Proof.** Given  $x \in X$  we have that

$$\forall n \in \mathbb{N}_0 \sum_{i=1}^n f_i(x) \leq \sum_{i=1}^{n+1} f_i(x) \wedge \left( \sum_{i=1}^{\infty} f_i \right)(x) = \sum_{i=1}^{\infty} f_i(x) \stackrel{17.94}{=} \lim_{n \rightarrow \infty} \sum_{i=1}^n f_i(x)$$

allowing us to apply the Monotone Convergence Theorem [see 19.237] so that

$$\begin{aligned} \int^+ \left( \sum_{i=1}^{\infty} f_i \right) d\mu &\stackrel{\text{def}}{=} \int^+ \left( \lim_{n \rightarrow \infty} \left( \sum_{i=1}^n f_i \right) \right) d\mu \\ &\stackrel{19.237}{=} \lim_{n \rightarrow \infty} \int^+ \left( \sum_{i=1}^n f_i \right) d\mu \\ &\stackrel{19.210}{=} \lim_{n \rightarrow \infty} \left( \sum_{i=1}^n \int^+ f_i d\mu \right) \\ &= \sum_{i=1}^{\infty} \int^+ f_i d\mu \end{aligned} \quad \square$$

**Lemma 19.239. (Fatou's lemma)** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space and  $\{f_i\}_{i \in \mathbb{N}_0} \subseteq \overline{\mathcal{M}^+[X, \mathcal{A}]}$  then

$$\int^+ \left( \liminf_{n \rightarrow \infty} f_i \right) d\mu \leq \liminf_{n \rightarrow \infty} \int^+ f_i d\mu$$

**Note 19.240.** Using 19.169 we have that  $\liminf_{n \rightarrow \infty} f_i \in \overline{\mathcal{M}^+[X, \mathcal{A}]}$  so that  $\int^+ \left( \liminf_{n \rightarrow \infty} f_i \right) d\mu$  is defined.

**Note 19.241.** We can not assume that  $\lim_{n \rightarrow \infty} \int^+ f_i d\mu$  exists which is the reason  $\liminf_{n \rightarrow \infty} \int^+ f_i d\mu$  is used.

**Proof.** Define  $\{g_i\}_{i \in \mathbb{N}_0}$  by  $g_n = \inf(\{f_i | i \in \{n, \dots, \infty\}\})$  then using 19.169 we have that

$$\{g_n\}_{n \in \mathbb{N}_0} \subseteq \overline{\mathcal{M}^+[X, \mathcal{A}]} \quad (19.459)$$

Let  $x \in X$  then  $\{f_i(x) | i \in \{n+1, \dots, \infty\}\} \subseteq \{f_i(x) | i \in \{n, \dots, \infty\}\}$  so that by 2.175 we have  $\inf(\{f_i(x) | i \in \{n, \dots, \infty\}\}) \leq f_n(x)$  and  $\inf(\{f_i(x) | i \in \{n, \dots, \infty\}\}) \leq \inf(\{f_i(x) | i \in \{n+1, \dots, \infty\}\})$  which gives

$$\forall n \in \mathbb{N}_0 \quad g_n(x) \leq f_n(x) \text{ and } g_n(x) \leq g_{n+1}(x) \quad (19.460)$$

So we have

$$\begin{aligned} \lim_{n \rightarrow \infty} g_n(x) &\stackrel{17.71}{=} \sup(\{g_i(x) | i \in \{n, \dots, \infty\}\}) \\ &= \sup(\{\inf(\{f_i(x) | i \in \{n, \dots, \infty\}\}) | n \in \mathbb{N}_0\}) \\ &\stackrel{\text{def}}{=} \liminf_{n \rightarrow \infty} f_i(x) \end{aligned}$$

We can apply the Monotone Convergence Theorem [see 19.237] we have that

$$\begin{aligned} \int^+ \left( \liminf_{n \rightarrow \infty} f_i \right) d\mu &= \int^+ \left( \lim_{n \rightarrow \infty} g_n \right) d\mu \\ &\stackrel{19.237}{=} \lim_{n \rightarrow \infty} \int^+ g_n d\mu \\ &\stackrel{\text{def}}{=} \liminf_{n \rightarrow \infty} \int^+ g_n d\mu \end{aligned} \quad (19.461)$$

Using 19.460 and 19.209 we have  $\forall n \in \mathbb{N}_0$  that  $\int^+ g_n d\mu \leq \int^+ f_n d\mu$  allowing us to use 17.55 giving  $\liminf_{n \rightarrow \infty} \int^+ g_n d\mu \leq \liminf_{n \rightarrow \infty} \int^+ f_n d\mu$ . Substituting in 19.461 gives finally

$$\int^+ \left( \liminf_{n \rightarrow \infty} f_i \right) d\mu \leq \liminf_{n \rightarrow \infty} \int^+ f_n d\mu \quad \square$$

**Theorem 19.242. (Lebesgue's Dominated Convergence Theorem)** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space,  $g \in \mathcal{L}[X, \mathcal{A}, \mu, \overline{\mathbb{R}}]$  a non negative extended real valued  $\mu$ -integrable function,  $f \in \overline{\mathcal{M}[X, \mathcal{A}]}$  and  $\{f_i\}_{i \in \mathbb{N}_0} \subseteq \overline{\mathcal{M}[X, \mathcal{A}]}$  such that

$$f = \lim_{n \rightarrow \infty} f_n$$

and

$$\forall n \in \mathbb{N}_0 \quad |f_n| \leq g$$

holds  $\mu$ -a.e. then

1.  $f \in \mathcal{L}[X, \mathcal{A}, \mu, \overline{\mathbb{R}}]$
2.  $\{f_n\}_{n \in \mathbb{N}_0} \in \mathcal{L}[X, \mathcal{A}, \mu, \overline{\mathbb{R}}]$
3.  $\int f d\mu = \lim_{n \rightarrow \infty} \int f_n d\mu$

**Proof.** As  $g \in \mathcal{L}[X, \mathcal{A}, \mu, \overline{\mathbb{R}}]$  is non negative we have by 19.221 that

$$g \in \overline{\mathcal{M}^+[X, \mathcal{A}]} \text{ and } \int g d\mu = \int^+ g d\mu < \infty \quad (19.462)$$

From 19.234 and the non negativity of  $g$  we have that  $|g| = g < \infty$   $\mu$ -a.e. so there exists a  $N_1 \in \mathcal{A}$  with  $\mu(N_1)$  such that

$$A = \{x \in X | g(x) = \infty\} = \{x \in X | |g(x)| = \infty\} \subseteq N_1 \quad (19.463)$$

From the assumptions there exists a  $N_2 \in \mathcal{A}$  with  $\mu(N_2) = 0$  such that

$$B = \left\{ x \in X | \neg \left( f(x) = \lim_{n \rightarrow \infty} f_n(x) \wedge \forall n \in \mathbb{N}_0 |f_n(x)| \leq g(x) \right) \right\} \subseteq N_2 \quad (19.464)$$

Define  $N = N_1 \cup N_2$  then  $N \in \mathcal{A}$  and  $\mu(N) \leq \mu(N_1) + \mu(N_2) = 0 + 0 = 0$  giving

$$A \bigcup B \subseteq N \in \mathcal{A} \text{ with } \mu(N) = 0 \quad (19.465)$$

As  $N \in \mathcal{A}$ , so that  $X \setminus N \in \mathcal{A}$ , we have using 19.158 and 19.177 that

$$\begin{aligned} \mathcal{X}_{X \setminus N} \cdot f \in \overline{\mathcal{M}[X, \mathcal{A}]} & \wedge |\mathcal{X}_{X \setminus N} \cdot f| \in \overline{\mathcal{M}^+[X, \mathcal{A}]} \\ & \wedge \mathcal{X}_{X \setminus N} \cdot g \in \overline{\mathcal{M}^+[X, \mathcal{A}]} \\ & \wedge \forall n \in \mathbb{N}_0 \mathcal{X}_{X \setminus N} \cdot f_n \in \overline{\mathcal{M}[X, \mathcal{A}]} \end{aligned} \quad (19.466)$$

Let  $x \in X$  then we have either

$x \in N$ . then  $\mathcal{X}_{X \setminus A}(x) = 0$  so that  $\forall n \in \mathbb{N}_0$

$$|\mathcal{X}_{X \setminus N} \cdot f_n|(x) = 0 = (\mathcal{X}_{X \setminus N} \cdot g)(x)$$

and

$$(\mathcal{X}_{X \setminus N} \cdot f)(x) = 0 = \lim_{n \rightarrow \infty} (\mathcal{X}_{X \setminus N} \cdot f_n)(x) = \left( \lim_{n \rightarrow \infty} (\mathcal{X}_{X \setminus N} \cdot f_n) \right)(x)$$

and

$$(\mathcal{X}_{X \setminus N} \cdot g)(x) = 0 < \infty$$

$x \in X \setminus N$ . then  $x \notin A, x \notin B$  and  $\mathcal{X}_{X \setminus N}(x) = 1$  so that  $\forall n \in \mathbb{N}_0$

$$|\mathcal{X}_{X \setminus N} \cdot f_n|(x) = f_n(x) \leq_{x \notin B} g(x) = (\mathcal{X}_{X \setminus N} \cdot g)(x)$$

and

$$\begin{aligned} (\mathcal{X}_{X \setminus N} \cdot f)(x) &= f(x) \\ &\stackrel{x \notin B}{=} \lim_{n \rightarrow \infty} f_n(x) \\ &= \lim_{n \rightarrow \infty} (\mathcal{X}_{X \setminus N} \cdot f_n)(x) \\ &= \left( \lim_{n \rightarrow \infty} (\mathcal{X}_{X \setminus N} \cdot f_n) \right)(x) \end{aligned}$$

and

$$(\mathcal{X}_{X \setminus N} \cdot g)(x) = g(x) <_{x \notin A} \infty$$

which proves that  $\forall x \in X$

$$\begin{aligned} \forall n \in \mathbb{N}_0 |\mathcal{X}_{X \setminus N} \cdot f_n|(x) &\leq (\mathcal{X}_{X \setminus N} \cdot g)(x) < \infty \\ \left( \lim_{n \rightarrow \infty} (\mathcal{X}_{X \setminus N} \cdot f_n) \right)(x) &= (\mathcal{X}_{X \setminus N} \cdot f)(x) \end{aligned} \quad (19.467)$$

Using  $\mu(N) = 0$  we have by 19.182 that  $g = \mathcal{X}_{X \setminus N} \cdot g$   $\mu$ -a.e. Further we have by 19.228 and the fact that  $\mathcal{X}_{X \setminus A} \cdot g$  is real valued that:

$$\mathcal{X}_{X \setminus A} \cdot g \in \mathcal{L} \wedge \int (\mathcal{X}_{X \setminus N} \cdot g) d\mu = \int g d\mu < \infty \quad (19.468)$$

Using 19.467, 19.466, the above and 19.209 we have that

$$\int^+ |\mathcal{X}_{X \setminus N} \cdot f_n| d\mu \leq \int^+ (\mathcal{X}_{X \setminus N} \cdot g) d\mu \stackrel{19.221}{=} \int (\mathcal{X}_{X \setminus N} \cdot g) d\mu < \infty$$

From the above we have using 19.221 that  $|\mathcal{X}_{X \setminus N} \cdot f_n| \in \mathcal{L}[X, \mathcal{A}, \mu, \overline{\mathbb{R}}]$ , using 19.466 and 19.226 we have that  $\mathcal{X}_{X \setminus N} \cdot f_n \in \mathcal{L}[X, \mathcal{A}, \mu, \overline{\mathbb{R}}]$ . Combining this with the fact that by 19.182  $\mathcal{X}_{X \setminus N} \cdot f_n = f_n$   $\mu$ -a.e., 19.466 and 19.228 we have that  $f_n \in \mathcal{L}[X, \mathcal{A}, \mu, \overline{\mathbb{R}}]$  and  $\int f_n d\mu = \int (\mathcal{X}_{X \setminus A} \cdot f_n) d\mu$ . As by 19.467  $\mathcal{X}_{X \setminus N} \cdot f_n$  is real valued we have

$$\forall n \in \mathbb{N}_0 f_n \in \mathcal{L}[X, \mathcal{A}, \mu, \overline{\mathbb{R}}], \mathcal{X}_{X \setminus N} \cdot f_n \in \mathcal{L} \text{ and } \int f_n d\mu = \int (\mathcal{X}_{X \setminus A} \cdot f_n) d\mu \quad (19.469)$$

As by 19.467  $\lim_{n \rightarrow \infty} (\mathcal{X}_{X \setminus N} \cdot f_n)(x) = (\mathcal{X}_{X \setminus N} \cdot f)(x)$  we have by 17.76 that

$$|\mathcal{X}_{X \setminus N} \cdot f|(x) = \lim_{n \rightarrow \infty} |\mathcal{X}_{X \setminus N} \cdot f_n|(x) \quad (19.470)$$

As by 19.467  $|\mathcal{X}_{X \setminus N} \cdot f_n|(x) \leq (\mathcal{X}_{X \setminus N} \cdot g)(x)$  we have using 17.77 that

$$\lim_{n \rightarrow \infty} |\mathcal{X}_{X \setminus N} \cdot f_n|(x) \leq (\mathcal{X}_{X \setminus N} \cdot g)(x) < \infty$$

Hence using 19.470 and the above we conclude

$$\forall x \in X |(\mathcal{X}_{X \setminus N} \cdot f)(x)| \leq (\mathcal{X}_{X \setminus N} \cdot g)(x) < \infty \quad (19.471)$$

Combining the above with 19.466 and 19.209 we have  $\int^+ |\mathcal{X}_{X \setminus N} \cdot f| d\mu \leq \int^+ (\mathcal{X}_{X \setminus N} \cdot g) d\mu <_{19.468} \infty$ . Using 19.466 and 19.221 we have that  $|\mathcal{X}_{X \setminus N} \cdot f| \in \mathcal{L}[X, \mathcal{A}, \mu, \overline{\mathbb{R}}]$ . Further by 19.226 it follows that  $\mathcal{X}_{X \setminus A} \cdot f \in \mathcal{L}[X, \mathcal{A}, \mu, \overline{\mathbb{R}}]$  which as by 19.182  $\mathcal{X}_{X \setminus A} \cdot f = f$   $\mu$ -a.e. proves that

$$f \in \mathcal{L}[X, \mathcal{A}, \mu, \overline{\mathbb{R}}] \text{ and } \int f d\mu = \int (\mathcal{X}_{X \setminus N} \cdot f) d\mu \quad (19.472)$$

and as  $\mathcal{X}_{X \setminus A} \cdot f$  is real valued [see 19.467] we have that

$$\mathcal{X}_{X \setminus N} \cdot f \in \mathcal{L} \quad (19.473)$$

Given  $x \in X$  and  $n \in \mathbb{N}_0$  then from  $|(\mathcal{X}_{X \setminus N} \cdot f_n)(x)| \leq (\mathcal{X}_{X \setminus N} \cdot g)(x) < \infty$  [see 19.467] we have that

$$\begin{aligned} (\mathcal{X}_{X \setminus N} \cdot f_n)(x) &\leq (\mathcal{X}_{X \setminus N} \cdot g)(x) \\ -(\mathcal{X}_{X \setminus N} \cdot f_n)(x) &\leq (\mathcal{X}_{X \setminus N} \cdot g)(x) \end{aligned}$$

so that

$$\begin{aligned} 0 &\leq (\mathcal{X}_{X \setminus N} \cdot f_n + \mathcal{X}_{X \setminus N} \cdot g)(x) \\ 0 &\leq (\mathcal{X}_{X \setminus N} \cdot g - \mathcal{X}_{X \setminus N} \cdot f_n)(x) \end{aligned}$$

proving that

$$\{\mathcal{X}_{X \setminus N} \cdot f_n + \mathcal{X}_{X \setminus N} \cdot g\}_{n \in \mathbb{N}_0}, \{\mathcal{X}_{X \setminus N} \cdot g \mathcal{X}_{X \setminus N} \cdot f_n\}_{n \in \mathbb{N}_0} \text{ are non negative sequences} \quad (19.474)$$

Using 19.468, 19.469 and 19.225 we have also

$$\{\mathcal{X}_{X \setminus N} \cdot f_n + \mathcal{X}_{X \setminus N} \cdot g\}_{n \in \mathbb{N}_0} \in \mathcal{L} \wedge \{\mathcal{X}_{X \setminus N} \cdot g \mathcal{X}_{X \setminus N} \cdot f_n\}_{n \in \mathbb{N}_0} \in \mathcal{L} \quad (19.475)$$

Further using 17.67, 17.77 and 19.467 we have that

$$\begin{aligned} (\mathcal{X}_{X \setminus N} \cdot g)(x) + (\mathcal{X}_{X \setminus N} \cdot f)(x) &= \lim_{n \rightarrow \infty} (\mathcal{X}_{X \setminus N} \cdot g)(x) + \lim_{n \rightarrow \infty} (\mathcal{X}_{X \setminus N} \cdot f_n)(x) \\ &= \lim_{n \rightarrow \infty} ((\mathcal{X}_{X \setminus N} \cdot g)(x) + (\mathcal{X}_{X \setminus N} \cdot f_n)(x)) \\ &= \liminf_{n \rightarrow \infty} (\mathcal{X}_{X \setminus N} \cdot g + \mathcal{X}_{X \setminus N} \cdot f_n)(x) \\ (\mathcal{X}_{X \setminus N} \cdot g)(x) - (\mathcal{X}_{X \setminus N} \cdot f)(x) &= \lim_{n \rightarrow \infty} (\mathcal{X}_{X \setminus N} \cdot g)(x) - \lim_{n \rightarrow \infty} (\mathcal{X}_{X \setminus N} \cdot f_n)(x) \\ &= \lim_{n \rightarrow \infty} ((\mathcal{X}_{X \setminus N} \cdot g)(x) - (\mathcal{X}_{X \setminus N} \cdot f_n)(x)) \\ &= \liminf_{n \rightarrow \infty} (\mathcal{X}_{X \setminus N} \cdot g - \mathcal{X}_{X \setminus N} \cdot f_n)(x) \end{aligned} \quad (19.476)$$

Next we have using 19.473 and 19.468

$$\begin{aligned} \int (\mathcal{X}_{X \setminus N} \cdot g) d\mu + \int (\mathcal{X}_{X \setminus N} \cdot f) d\mu &\stackrel{19.225}{=} \int (\mathcal{X}_{X \setminus N} \cdot g + \mathcal{X}_{X \setminus N} \cdot f) d\mu \\ &\stackrel{19.476}{=} \int \liminf_{n \rightarrow \infty} (\mathcal{X}_{X \setminus N} \cdot g + \mathcal{X}_{X \setminus N} \cdot f_n) d\mu \\ &\leq_{19.239, 19.474} \liminf_{n \rightarrow \infty} \int (\mathcal{X}_{X \setminus N} \cdot g + \mathcal{X}_{X \setminus N} \cdot f_n) d\mu \\ &\stackrel{19.225}{=} \liminf_{n \rightarrow \infty} \left( \int (\mathcal{X}_{X \setminus N} \cdot g) d\mu + \int (\mathcal{X}_{X \setminus N} \cdot f_n) d\mu \right) \\ &\leq_{17.60} \liminf_{n \rightarrow \infty} \int (\mathcal{X}_{X \setminus N} \cdot g) d\mu + \liminf_{n \rightarrow \infty} \int (\mathcal{X}_{X \setminus N} \cdot f_n) d\mu \\ &\stackrel{17.53}{=} \int (\mathcal{X}_{X \setminus N} \cdot g) d\mu + \liminf_{n \rightarrow \infty} \int (\mathcal{X}_{X \setminus N} \cdot f_n) d\mu \end{aligned}$$

As the integrals are real numbers the above gives

$$\int (\mathcal{X}_{X \setminus N} \cdot f) d\mu \leq \liminf_{n \rightarrow \infty} \int (\mathcal{X}_{X \setminus N} \cdot f_n) d\mu \quad (19.477)$$

Similar we have that

$$\begin{aligned}
 \int (\mathcal{X}_{X \setminus N} \cdot g) d\mu - \int (\mathcal{X}_{X \setminus N} \cdot f) d\mu &\stackrel{19.225}{=} \int (\mathcal{X}_{X \setminus N} \cdot g - \mathcal{X}_{X \setminus N} \cdot f) d\mu \\
 &\stackrel{19.476}{=} \int \liminf_{n \rightarrow \infty} (\mathcal{X}_{X \setminus N} \cdot g - \mathcal{X}_{X \setminus N} \cdot f_n) d\mu \\
 &\leq_{19.239, 19.474} \liminf_{n \rightarrow \infty} \int (\mathcal{X}_{X \setminus N} \cdot g - \mathcal{X}_{X \setminus N} \cdot f_n) d\mu \\
 &\stackrel{19.225}{=} \liminf_{n \rightarrow \infty} \left( \int (\mathcal{X}_{X \setminus N} \cdot g) d\mu - \int (\mathcal{X}_{X \setminus N} \cdot f_n) d\mu \right) \\
 &\leq_{17.60} \liminf_{n \rightarrow \infty} \int (\mathcal{X}_{X \setminus N} \cdot g) d\mu + \liminf_{n \rightarrow \infty} - \int (\mathcal{X}_{X \setminus N} \cdot f_n) d\mu \\
 &\stackrel{17.58}{=} \liminf_{n \rightarrow \infty} \int (\mathcal{X}_{X \setminus N} \cdot g) d\mu - \limsup_{n \rightarrow \infty} \int (\mathcal{X}_{X \setminus N} \cdot f_n) d\mu \\
 &\stackrel{17.53}{=} \int (\mathcal{X}_{X \setminus N} \cdot g) d\mu - \limsup_{n \rightarrow \infty} \int (\mathcal{X}_{X \setminus N} \cdot f_n) d\mu
 \end{aligned}$$

As the integrals are real numbers the above gives

$$\limsup_{n \rightarrow \infty} \int (\mathcal{X}_{X \setminus N} \cdot f_n) d\mu \leq \int (\mathcal{X}_{X \setminus N} \cdot f) d\mu$$

which combined with 19.477 gives

$$\int (\mathcal{X}_{X \setminus N} \cdot f) d\mu \leq \liminf_{n \rightarrow \infty} \int (\mathcal{X}_{X \setminus N} \cdot f_n) d\mu \leq_{17.61} \limsup_{n \rightarrow \infty} \int (\mathcal{X}_{X \setminus N} \cdot f_n) d\mu \leq \int (\mathcal{X}_{X \setminus N} \cdot f) d\mu$$

proving that  $\lim_{n \rightarrow \infty} \int (\mathcal{X}_{X \setminus N} \cdot f_n) d\mu$  exists and  $\int (\mathcal{X}_{X \setminus N} \cdot f) d\mu = \lim_{n \rightarrow \infty} \int (\mathcal{X}_{X \setminus N} \cdot f_n) d\mu$ . Using then 19.472 and 19.469 we have

$$\int f d\mu = \lim_{n \rightarrow \infty} \int f_n d\mu \tag{19.478}$$

The theorem is then proved by 19.469, 19.472 and the above.  $\square$

### 19.6.6 Integral of complex functions

Using the defintion of the complex numbers [see 9.1], we see that the set  $\mathbb{C}$  of complex numbers is defined as  $\mathbb{C} = \mathbb{R} \times \mathbb{R} = \mathbb{R}^2$ . Further all the norms on  $\mathbb{C}$  are equivalent [see 11.86] and generate the same topology, so the topology based on the maximum norm  $\mathcal{T}_{\mathbb{R}^2}$  is also the canonical topology  $\mathcal{T}_{\mathbb{C}}$  on  $\mathbb{C}$  based on the the complex norm  $|z| = \sqrt{\operatorname{Re}(z)^2 + \operatorname{Img}(z)^2}$  [see 9.32, 9.17]. The difference between  $\mathbb{C}$  and  $\mathbb{R}$  lies the definition of a product in  $\mathbb{C}$ . Hence we can use  $\langle \mathbb{R}, \mathcal{B}[\mathbb{R}^2] \rangle$  as the definition of the measurable space  $\langle \mathbb{C}, \mathcal{B}[\mathbb{C}] \rangle$

**Definition 19.243.** *The measurable space for the complex numbers iis defined by as  $\langle \mathbb{C}, \mathcal{B}[\mathbb{C}] \rangle$  where  $\mathbb{C} = \mathbb{R}^2$  and  $\mathcal{B}[\mathbb{C}] = \sigma[\mathcal{T}_{\mathbb{C}}] = \sigma[\mathcal{T}_{\mathbb{R}^2}] = \mathcal{B}[\mathbb{R}^2]$*

**Definition 19.244.** *Let  $\langle X, \mathcal{A} \rangle$  be a measurable space then the set of all  $\mathcal{A}, \mathcal{B}[\mathbb{C}]$ -measurable function graphs is noted as  $\mathcal{M}[X, \mathcal{A}, \mathbb{C}]$  so*

$$\mathcal{M}[X, \mathcal{A}, \mathbb{C}] = \{f \in \mathbb{C}^X \mid f \text{ is } \mathcal{A}, \mathcal{B}[\mathbb{C}]\text{-measurable}\}$$

The next proposition shows the relation between  $\mathcal{B}[\mathbb{R}^n]$  functions and their components.

**Proposition 19.245.** *Let  $\langle X, \mathcal{A} \rangle$  be a measurable space and  $f: X \rightarrow \mathbb{R}^n$  then  $f$  is  $\mathcal{A}, \mathcal{B}[\mathbb{R}^n]$ -meeasurable if and only if  $\forall i \in \{1, \dots, n\}$   $f_i: X \rightarrow \mathbb{R}$  defined by  $f_i = \pi_i \circ f$  is  $\mathcal{A}$ -measurable [meaning  $\mathcal{A}, \mathcal{B}[\mathbb{R}]$ -measurable by convention].*

**Proof.**

$\Rightarrow$ . Given  $i \in \{1, \dots, n\}$  we have by 11.188 that  $\pi_i: \mathbb{R}^n \rightarrow \mathbb{R}$  is continuous in the toplogies  $\mathcal{T}_{\mathbb{R}^n}$  and  $\mathcal{T}_{\mathbb{R}}$ , so using 19.143 it follows that  $\pi_i$  is  $\mathcal{B}[\mathbb{R}^n], \mathcal{B}[\mathbb{R}]$ -measurable. As  $f$  is  $\mathcal{A}, \mathcal{B}[\mathbb{R}^n]$ -measurable we can finally use 19.144 to prove that  $f_i = \pi_i \circ f$  is  $\mathcal{A}, \mathcal{B}[\mathbb{R}]$ -measurable.

$\Leftarrow$ . Let  $A \in \mathcal{H}^n$  then  $\exists x \in \mathbb{R}^n$  so that  $A = \{y \in \mathbb{R}^n \mid \forall i \in \{1, \dots, n\} y_i \in ]-\infty, x_i[\}$  [see 19.53]. We have then that

$$\begin{aligned} y \in f^{-1}(A) &\Leftrightarrow f(y) \in A \\ &\Leftrightarrow \forall i \in \{1, \dots, n\} (f(y))_i \in ]-\infty, x_i[ \\ &\Leftrightarrow \forall i \in \{1, \dots, n\} f_i(y) \in ]-\infty, x_i[ \\ &\Leftrightarrow \forall i \in \{1, \dots, n\} y \in f_i^{-1}(]-\infty, x_i[) \\ &\Leftrightarrow \forall i \in \{1, \dots, n\} y \in \{f_i < x_i\} \\ &\Leftrightarrow y \in \bigcap_{i \in \{1, \dots, n\}} \{f_i < x_i\} \end{aligned}$$

Hence  $f^{-1}(A) = \bigcap_{i \in \{1, \dots, n\}} \{f_i < x_i\} \in \mathcal{A}$  [as  $\forall i \in \{1, \dots, n\}$   $f_i$  is  $\mathcal{A}$ -measurable we can apply 19.151]. So we have proved that

$$\forall A \in \mathcal{H}^n \text{ we have } f^{-1}(A) \in \mathcal{A}$$

As  $\mathcal{B}[\mathbb{R}^n] = \sigma[\mathcal{H}^n]$  we have by 19.140 that  $f$  is  $\mathcal{A}, \mathcal{B}[\mathbb{R}^n]$ -measurable.  $\square$

**Corollary 19.246.** Let  $\langle X, \mathcal{A} \rangle$  be a measurable space then  $f: X \rightarrow \mathbb{C}$  is  $\mathcal{A}, \mathcal{B}$ -measurable if and only if  $\text{Re}(f)$  and  $\text{Img}(f)$  are  $\mathcal{A}$ -measurable

**Proof.** This follows from the previous proposition [see 19.245] and the fact that  $\mathbb{C} = \mathbb{R}^2$ ,  $\text{Re}(f) = \pi_1 \circ f$  and  $\text{Img}(f) = \pi_2 \circ f$ .  $\square$

**Corollary 19.247.** Let  $\langle X, \mathcal{A} \rangle$  be a measurable space then

1.  $\forall f, g \in \mathcal{M}[X, \mathcal{A}, \mathbb{C}]$  we have that  $f + g \in \mathcal{M}[X, \mathcal{A}, \mathbb{C}]$
2.  $\forall \alpha \in \mathbb{C}$  we have that  $\alpha \cdot f \in \mathcal{M}[X, \mathcal{A}, \mathbb{C}]$

**Proof.**

1. As  $f, g \in \mathcal{M}[X, \mathcal{A}, \mathbb{C}]$  we have by the previous corollary [see 19.246] that

$$\text{Re}(f), \text{Img}(f), \text{Re}(g), \text{Img}(g) \in \mathcal{M}[X, \mathcal{A}] \quad (19.479)$$

Let  $x \in X$  then

$$\begin{aligned} (f + g)(x) &= f(x) + g(x) \\ &= \text{Re}(f(x)) + i \cdot \text{Img}(f(x)) + \text{Re}(g(x)) + i \cdot \text{Img}(g(x)) \\ &= (\text{Re}(f(x)) + \text{Re}(g(x))) + i \cdot (\text{Img}(f(x)) + \text{Img}(g(x))) \\ &= (\text{Re}(f) + \text{Re}(g))(x) + i \cdot (\text{Img}(f) + \text{Img}(g))(x) \end{aligned}$$

proving that

$$\text{Re}(f + g) = \text{Re}(f) + \text{Re}(g) \text{ and } \text{Img}(f + g) = \text{Img}(f) + \text{Img}(g)$$

Using the above, 19.479 and 19.170 we have that  $\text{Re}(f + g), \text{Img}(f + g) \in \mathcal{M}[X, \mathcal{A}]$  so that by the previous corollary [see 19.246] we have that

$$f + g \in \mathcal{M}[X, \mathcal{A}, \mathbb{C}]$$

2. As  $f \in \mathcal{M}[X, \mathcal{A}, \mathbb{C}]$  we have by the previous corollary [see 19.246] that

$$\text{Re}(f), \text{Img}(f) \in \mathcal{A} \quad (19.480)$$

Let  $x \in X$  then

$$\begin{aligned} (\alpha \cdot f)(x) &= \alpha \cdot f(x) \\ &= (\text{Re}(\alpha) + i \cdot \text{Img}(\alpha)) \cdot (\text{Re}(f(x)) + i \cdot \text{Img}(f(x))) \\ &= (\text{Re}(\alpha) \cdot \text{Re}(f(x)) - \text{Img}(\alpha) \cdot \text{Img}(f(x))) + i \cdot (\text{Re}(\alpha) \cdot \text{Img}(f(x)) + \text{Img}(\alpha) \cdot \text{Re}(f(x))) \\ &= (\text{Re}(\alpha) \cdot \text{Re}(f) - \text{Img}(\alpha) \cdot \text{Img}(f))(x) + i \cdot (\text{Re}(\alpha) \cdot \text{Img}(f) + \text{Img}(\alpha) \cdot \text{Re}(f))(x) \end{aligned}$$

proving that

$$\operatorname{Re}(\alpha \cdot f) = \operatorname{Re}(\alpha) \cdot \operatorname{Re}(f) - \operatorname{Img}(\alpha) \cdot \operatorname{Img}(f) \wedge \operatorname{Img}(\alpha \cdot f) = \operatorname{Re}(\alpha) \cdot \operatorname{Img}(f) + \operatorname{Img}(\alpha) \cdot \operatorname{Re}(f)$$

Using the above, 19.480 and 19.170 we have that  $\operatorname{Re}(\alpha \cdot f), \operatorname{Img}(\alpha \cdot f) \in \mathcal{M}[X, \mathcal{A}]$  which by the previous corollary [see 19.246] proves

$$\alpha \cdot f \in \mathcal{M}[X, \mathcal{A}, \mathbb{C}]$$

□

We can now define integrability on complex valued functions.

**Definition 19.248.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measurable space and  $f: X \rightarrow \mathbb{C}$  a function then  $f$  is  $\mu$ -integrable iff  $\operatorname{Re}(f) \in \mathcal{L}[X, \mathcal{A}, \mu, \mathbb{R}]$  and  $\operatorname{Img}(f) \in \mathcal{L}[X, \mathcal{A}, \mu, \mathbb{R}]$  where  $\operatorname{Re}(f)(x) = \operatorname{Re}(f(x))$  and  $\operatorname{Img}(f)(x) = \operatorname{Img}(f(x))$ . The integral  $\int f d\mu \in \mathbb{C}$  is then defined by

$$\int f d\mu = \int \operatorname{Re}(f) d\mu + i \cdot \int \operatorname{Img}(f) d\mu$$

The set of all complex integrable functions is noted as  $\mathcal{L}[X, \mathcal{A}, \mu, \mathbb{C}]$

$$\mathcal{L}[X, \mathcal{A}, \mu, \mathbb{C}] = \{f \in \mathbb{C}^X \mid f \text{ is } \mu\text{-integrable}\}$$

Integrals of complex valued functions have similar properties as integrals of real valued functions.

**Proposition 19.249.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measurable space,  $\alpha \in \mathbb{C}$  and  $f: X \rightarrow \mathbb{C}$ ,  $g: X \rightarrow \mathbb{C}$  complex functions such that  $f$  and  $g$  are  $\mu$ -integrable then:

1.  $f + g$  is  $\mu$ -integrable and

$$\int (f + g) d\mu = \int f d\mu + \int g d\mu$$

2.  $\alpha \cdot f$  is  $\mu$ -integrable and

$$\int (\alpha \cdot f) d\mu = \alpha \cdot \int f d\mu$$

**Proof.** As  $f, g$  are integrable we have using 19.225 that

$$\operatorname{Re}(f), \operatorname{Re}(g), \operatorname{Img}(f), \operatorname{Img}(g) \in \mathcal{L}$$

1. Let  $x \in X$  then  $\operatorname{Re}(f+g)(x) = \operatorname{Re}(f(x) + g(x)) = \operatorname{Re}(f(x)) + \operatorname{Re}(g(x)) = (\operatorname{Re}(f) + \operatorname{Re}(g))(x)$  and  $\operatorname{Img}(f+g)(x) = \operatorname{Img}(f(x) + g(x)) = \operatorname{Img}(f(x)) + \operatorname{Img}(g(x)) = (\operatorname{Img}(f) + \operatorname{Img}(g))(x)$  proving that

$$\operatorname{Re}(f+g) = \operatorname{Re}(f) + \operatorname{Re}(g) \text{ and } \operatorname{Img}(f+g) = \operatorname{Img}(f) + \operatorname{Img}(g)$$

Using 19.225 we have then that  $\operatorname{Re}(f+g), \operatorname{Img}(f+g) \in \mathcal{L}[X, \mathcal{A}, \mu, \mathbb{R}]$  so that  $f+g$  is  $\mu$ -integrable and

$$\begin{aligned} \int (f+g) d\mu &= \int \operatorname{Re}(f+g) d\mu + i \cdot \int \operatorname{Img}(f+g) d\mu \\ &= \int (\operatorname{Re}(f) + \operatorname{Re}(g)) d\mu + i \cdot \int (\operatorname{Img}(f) + \operatorname{Img}(g)) d\mu \\ &\stackrel{19.225}{=} \int \operatorname{Re}(f) d\mu + \int \operatorname{Re}(g) d\mu + i \cdot \left( \int \operatorname{Img}(f) d\mu + \int \operatorname{Img}(g) d\mu \right) \\ &= \left( \int \operatorname{Re}(f) d\mu + i \cdot \int \operatorname{Img}(f) d\mu \right) + \left( \int \operatorname{Re}(g) d\mu + i \cdot \int \operatorname{Img}(g) d\mu \right) \\ &= \int f d\mu + \int g d\mu \end{aligned}$$

2. Let  $x \in X$  then

$$\begin{aligned} (\alpha \cdot f)(x) &= (\operatorname{Re}(\alpha) + i \cdot \operatorname{Img}(\alpha)) \cdot (\operatorname{Re}(f(x)) + i \cdot \operatorname{Img}(f(x))) \\ &= \operatorname{Re}(\alpha) \cdot \operatorname{Re}(f(x)) - \operatorname{Img}(\alpha) \cdot \operatorname{Img}(f(x)) + i \cdot (\operatorname{Re}(\alpha) \cdot \operatorname{Img}(f(x)) + \operatorname{Img}(\alpha) \cdot \operatorname{Re}(f(x))) \end{aligned}$$

So that  $\operatorname{Re}(\alpha \cdot f(x)) = \operatorname{Re}(\alpha) \cdot \operatorname{Re}(f(x)) - \operatorname{Img}(\alpha) \cdot \operatorname{Img}(f(x))$  and  $\operatorname{Img}(\alpha \cdot f(x)) = \operatorname{Re}(\alpha) \cdot \operatorname{Img}(f(x)) + \operatorname{Img}(\alpha) \cdot \operatorname{Re}(f(x))$  proving that

$$\operatorname{Re}(\alpha \cdot f) = \operatorname{Re}(\alpha) \cdot \operatorname{Re}(f) - \operatorname{Img}(\alpha) \cdot \operatorname{Img}(f) \wedge \operatorname{Img}(\alpha \cdot f) = \operatorname{Re}(\alpha) \cdot \operatorname{Img}(f) + \operatorname{Img}(\alpha) \cdot \operatorname{Re}(f)$$

Using 19.225 we have that  $\operatorname{Re}(\alpha \cdot f), \operatorname{Img}(\alpha \cdot f) \in \mathcal{L}[X, \mathcal{A}, \mu, \mathbb{R}]$  so that  $\alpha \cdot f$  is  $\mu$ -integrable and

$$\begin{aligned} \int (\alpha \cdot f) d\mu &= \\ \int (\operatorname{Re}(\alpha) \cdot \operatorname{Re}(f) - \operatorname{Img}(\alpha) \cdot \operatorname{Img}(f)) d\mu + i \cdot \int (\operatorname{Re}(\alpha) \cdot \operatorname{Img}(f) + \operatorname{Img}(\alpha) \cdot \operatorname{Re}(f)) d\mu &\stackrel{19.225}{=} \\ \operatorname{Re}(\alpha) \cdot \int \operatorname{Re}(f) d\mu - \operatorname{Img}(\alpha) \cdot \int \operatorname{Img}(f) d\mu + i \cdot \operatorname{Re}(\alpha) \cdot \int \operatorname{Img}(f) d\mu + i \cdot \operatorname{Img}(\alpha) \cdot \int &= \\ \operatorname{Re}(f) d\mu &= \\ \operatorname{Re}(\alpha) \cdot \left( \int \operatorname{Re}(f) d\mu + i \cdot \int \operatorname{Img}(f) d\mu \right) + i \cdot \operatorname{Img}(\alpha) \cdot \left( \int \operatorname{Re}(f) d\mu + i \cdot \int \operatorname{Img}(f) d\mu \right) &= \\ (\operatorname{Re}(\alpha) + i \cdot \operatorname{Img}(f)) \cdot \left( \int \operatorname{Re}(f) d\mu + i \cdot \int \operatorname{Img}(f) d\mu \right) &= \\ \alpha \cdot \int f d\mu & \end{aligned}$$

□

**Lemma 19.250.** Let  $\langle X, \mathcal{A} \rangle$  be a measurable space and  $f: X \rightarrow \mathbb{C}$  a complex valued function such that  $f$  is  $\mathcal{A}, \mathcal{B}[\mathbb{C}]$ -measurable then

1.  $|f|$  is  $\mathcal{A}$ -measurable
2.  $\operatorname{Re}(f)$  is  $\mathcal{A}$ -measurable
3.  $\operatorname{Img}(f)$  is  $\mathcal{A}$ -measurable

**Proof.**

1. As  $||: \mathbb{C} \rightarrow \mathbb{R}$  defined by  $||(x) = |x|$  is continuous [using 11.162 the norm function is continuous] we have by 19.143 that  $||$  is  $\mathcal{B}[\mathbb{C}], \mathcal{B}[\mathbb{R}]$ -measurable. Further as  $f$  is  $\mathcal{A}, \mathcal{B}[\mathbb{C}]$  measurable we have by 19.144 that

$$|| \circ f \text{ is } \mathcal{A}, \mathcal{B}[\mathbb{R}]\text{-measurable}$$

Finally as  $\forall x \in X$  we have that  $((|| \circ f)(x) = ||(f(x)) = |f(x)| = |f|(x))$  it follows that  $|f| = || \circ f$  which proves that

$$|f| \text{ is } \mathcal{A}, \mathcal{B}[\mathbb{R}]\text{-measurable or } |f| \text{ is } \mathcal{A}\text{-measurable}$$

2. As  $\operatorname{Re}: \mathbb{C} \rightarrow \mathbb{R}$  defined by  $x \rightarrow \operatorname{Re}(x)$  is continuous [see 11.156] we have by 19.143 that  $\operatorname{Re}$  is  $\mathcal{B}[\mathbb{C}], \mathcal{B}[\mathbb{R}]$ -measurable. Further as  $f$  is  $\mathcal{A}, \mathcal{B}[\mathbb{C}]$  measurable we have by 19.144 that

$$\operatorname{Re} \circ f \text{ is } \mathcal{A}, \mathcal{B}[\mathbb{R}]\text{-measurable}$$

Finally as  $\forall x \in X$  we have that  $(\operatorname{Re} \circ f)(x) = \operatorname{Re}(f(x)) = \operatorname{Re}(f)(x)$  it follows that  $\operatorname{Re}(f) = \operatorname{Re} \circ f$  which proves that

$$\operatorname{Re}(f) \text{ is } \mathcal{A}, \mathcal{B}[\mathbb{R}]\text{-measurable or } \operatorname{Re}(f) \text{ is } \mathcal{A}\text{-measurable}$$

3. As  $\operatorname{Img}: \mathbb{C} \rightarrow \mathbb{R}$  defined by  $x \rightarrow \operatorname{Img}(x)$  is continuous [see 11.156] we have by 19.143 that  $\operatorname{Img}$  is  $\mathcal{B}[\mathbb{C}], \mathcal{B}[\mathbb{R}]$ -measurable. Further as  $f$  is  $\mathcal{A}, \mathcal{B}[\mathbb{C}]$  measurable we have by 19.144 that

$$\operatorname{Img} \circ f \text{ is } \mathcal{A}, \mathcal{B}[\mathbb{R}]\text{-measurable}$$

Finally as  $\forall x \in X$  we have that  $(\text{Img} \circ f)(x) = \text{Img}(f(x)) = \text{Img}(f)(x)$  it follows that  $\text{Img}(f) = \text{Img} \circ f$  proving

$\text{Img}(f)$  is  $\mathcal{A}, \mathcal{B}[\mathbb{R}]$ -measurable or  $\text{Img}(f)$  is  $\mathcal{A}$ -measurable  $\square$

**Proposition 19.251.** *Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space and  $f: X \rightarrow \mathbb{C}$  a complex valued function that is  $\mathcal{A}, \mathcal{B}[\mathbb{C}]$ -measurable. Then  $f$  is  $\mu$ -integrable if and only if  $|f|$  is  $\mu$ -integrable. If either  $f$  or  $|f|$  is  $\mu$ -integrable then  $|\int f d\mu| \leq \int |f| d\mu$*

**Proof.**

$\Rightarrow$ . Using 19.250 we have that  $|f|$  is  $\mathcal{A}$ -measurable. As  $f$  is  $\mu$ -integrable we have that  $\text{Re}(f), \text{Img}(f) \in \mathcal{L}[X, \mathcal{A}, \mu, \mathbb{R}]$  and by definition  $\text{Re}(f), \text{Img}(f) \in \mathcal{M}[X, \mathcal{A}]$ . Applying 19.226 it follows that  $|\text{Re}(f)|, |\text{Img}(f)|$  are  $\mu$ -integrable hence

$$\int^+ |\text{Re}(f)| d\mu \stackrel{19.221}{=} \int |\text{Re}(f)| d\mu < \infty \wedge \int^+ |\text{Img}(f)| d\mu \stackrel{19.221}{=} \int |\text{Img}(f)| d\mu < \infty \quad (19.481)$$

Let  $x \in X$  then  $|f|(x) = |f(x)| \leq 9.33 |\text{Re}(f(x))| + |\text{Img}(f(x))| = |\text{Re}(f)|(x) + |\text{Img}(f)|(x)$  and using 19.209 it follows that

$$\int^+ |f| d\mu \leq \int^+ (|\text{Re}(f)| + |\text{Img}(f)|) d\mu \stackrel{19.209}{=} \int^+ |\text{Re}(f)| d\mu + \int^+ |\text{Img}(f)| d\mu < \infty$$

which by 19.221, 19.481 proves that

$$|f| \text{ is } \mu\text{-integrable}$$

$\Leftarrow$ . As  $|f|$  is  $\mu$ -integrable we have that

$$|f| \in \mathcal{M}^+ \{X, \mathcal{A}\} \text{ and } \int^+ |f| d\mu \stackrel{19.221}{=} \int |f| d\mu < \infty$$

As  $f$  is  $\mathcal{B}[\mathbb{R}], \mathcal{B}[\mathbb{C}]$ -measurable it follows from 19.250 that  $\text{Re}(f)$  and  $\text{Img}(f)$  are  $\mathcal{A}$ -measurable. Let  $x \in X$  then  $|\text{Re}(f)|(x) = |\text{Re}(f(x))| \leq 9.33 |f(x)| = |f|(x)$  and  $|\text{Img}(f(x))| \leq 9.33 |f(x)| = |f|(x)$  proving by 19.223 that  $\text{Re}(f) \in \mathcal{L}[X, \mathcal{A}, \mu, \mathbb{R}]$  and  $\text{Img}(f) \in \mathcal{L}[X, \mathcal{A}, \mu, \mathbb{R}]$ . Hence by definition

$$f \text{ is } \mu\text{-integrable}$$

Define now

$$w = \begin{cases} 1 & \text{if } |\int f d\mu| = 0 \\ \frac{\int f d\mu}{|\int f d\mu|} & \text{if } |\int f d\mu| \neq 0 \end{cases}$$

Then for  $\int f d\mu$  we have either

$\int f d\mu = 0$ . then  $|\int f d\mu| = 0 \Rightarrow w = 1$  so that  $w \cdot |\int f d\mu| = 1 \cdot 0 = \int f d\mu$

$\int f d\mu \neq 0$ . then  $|\int f d\mu| \neq 0 \Rightarrow w = \frac{\int f d\mu}{|\int f d\mu|}$  so that  $w \cdot |\int f d\mu| = \int f d\mu$

giving

$$\int f d\mu = w \cdot \left| \int f d\mu \right| \text{ and } |w| = 1$$

Then we have

$$\begin{aligned} \left| \int f d\mu \right| &= \frac{1}{w} \cdot \int f d\mu \\ &\stackrel{19.249}{=} \int \frac{1}{w} \cdot f d\mu \\ &\stackrel{\text{def}}{=} \int \text{Re}\left(\frac{1}{w} \cdot f\right) d\mu + i \cdot \int \text{Img}\left(\frac{1}{w} \cdot f\right) d\mu \\ &\stackrel{|\int f d\mu| \in \mathbb{R}}{=} \int \text{Re}\left(\frac{1}{w} \cdot f\right) d\mu \end{aligned} \quad (19.482)$$

Let  $x \in X$  then using 9.33  $\operatorname{Re}\left(\frac{1}{w} \cdot f(x)\right) \leq \left|\frac{1}{w} \cdot f(x)\right| \stackrel{9.33}{=} \left|\frac{1}{w}\right| \cdot |f(x)| \stackrel{9.33}{=} \frac{1}{|w|} \cdot |f(x)| = |f|(x)$  which combined with 19.225 proves that  $\int \operatorname{Re}\left(\frac{1}{w} \cdot f\right) d\mu \leq \int |f| d\mu$ . Combining this result with 19.482 proves finally

$$\left| \int f d\mu \right| \leq \int |f| d\mu \quad \square$$

**Proposition 19.252.** *Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space and  $f \in \mathcal{L}[X, \mathcal{A}, \mu, \mathbb{C}]$  such that  $f = 0$   $\mu$ -a.e then  $\int f d\mu = 0$*

**Proof.** As  $f = 0$   $\mu$ -a.e. there exists a  $N \in \mathcal{N}_\mu$  with  $\mu(N) = 0$  such that  $\{x \in X \mid f(x) \neq 0\} \subseteq N$ . If  $x \in \{x \in X \mid \operatorname{Re}(f(x)) \neq 0\}$  then  $\operatorname{Re}(f(x)) \neq 0$  hence  $f(x) \neq 0$  so  $x \in \{x \in X \mid f(x) \neq 0\} \subseteq N$ . Similar if  $x \in \{x \in X \mid \operatorname{Img}(f(x)) \neq 0\}$  then  $\operatorname{Img}(f(x)) \neq 0$  hence  $f(x) \neq 0$  so  $x \in \{x \in X \mid f(x) \neq 0\} \subseteq N$ . So we have that  $\{x \in X \mid \operatorname{Re}(f(x)) \neq 0\} \subseteq N$  and  $\{x \in X \mid \operatorname{Img}(f(x)) \neq 0\} \subseteq N$  proving that  $\operatorname{Re}(f) = 0$   $\mu$ -a.e. and  $\operatorname{Img}(f)$  is  $\mu$ -a.e. Using 19.246 and the fact that  $f \in \mathcal{M}[X, \mathcal{A}, \mathbb{C}]$  we have that  $\operatorname{Re}(f), \operatorname{Img}(f) \in \mathcal{M}[X, \mathcal{A}]$ , so using 19.222  $\int \operatorname{Re}(f) d\mu = 0$  and  $\int \operatorname{Img}(f) d\mu = 0$ , we have finally

$$\int f d\mu = \int \operatorname{Re}(f) d\mu + i \cdot \int \operatorname{Img}(f) d\mu = 0 + i \cdot 0 = 0 \quad \square$$

## 19.7 The spaces $\mathcal{L}^p$ and $L^p$

**Definition 19.253.** *Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space and  $p \in [1, \infty[$  then  $\mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{R}]$  and  $\mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{C}]$  is defined by:*

1.  $\mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{R}] = \{f \in \mathcal{M}[X, \mathcal{A}] \mid |f|^p \text{ is } \mu\text{-integrable}\}$  where  $|f|^p(x) = |f(x)|^p$
2.  $\mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{C}] = \{f \in \mathcal{M}[X, \mathcal{A}, \mathbb{C}] \mid |f|^p \text{ is } \mu\text{-integrable}\}$  where  $|f|^p(x) = |f(x)|^p$

**Proposition 19.254.** *Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space then*

1.  $\mathcal{L}[X, \mathcal{A}, \mu, \mathbb{R}] = \mathcal{L}^1[X, \mathcal{A}, \mu, \mathbb{R}]$  [see 19.211 for a definition of  $\mathcal{L}[X, \mathcal{A}, \mu, \mathbb{R}]$ ].
2.  $\mathcal{L}[X, \mathcal{A}, \mu, \mathbb{C}] = \mathcal{L}^1[X, \mathcal{A}, \mu, \mathbb{C}]$  [see 19.248 for a definition of  $\mathcal{L}[X, \mathcal{A}, \mu, \mathbb{C}]$ ].

**Proof.**

1. Let  $f \in \mathcal{L}[X, \mathcal{A}, \mu, \mathbb{R}]$  then  $f \in \mathcal{M}[X, \mathcal{A}]$  and  $f$  is  $\mu$ -integrable, using 19.226 it follows that  $|f|^1 = |f|$  is  $\mu$ -integrable or  $f \in \mathcal{L}^1[X, \mathcal{A}, \mu, \mathbb{R}]$  proving that  $\mathcal{L}[X, \mathcal{A}, \mu, \mathbb{R}] \subseteq \mathcal{L}^1[X, \mathcal{A}, \mu, \mathbb{R}]$ . Next if  $f \in \mathcal{L}^1[X, \mathcal{A}, \mu, \mathbb{R}]$  then  $f \in \mathcal{M}[X, \mathcal{A}]$  and  $|f| = |f|^1$  is  $\mu$ -integrable, using 19.226 again it follows that  $f$  is  $\mu$ -integrable or  $f \in \mathcal{L}[X, \mathcal{A}, \mu, \mathbb{R}]$ , proving that  $\mathcal{L}^1[X, \mathcal{A}, \mu, \mathbb{R}] \subseteq \mathcal{L}[X, \mathcal{A}, \mu, \mathbb{R}]$ . Hence

$$\mathcal{L}[X, \mathcal{A}, \mu, \mathbb{R}] = \mathcal{L}^1[X, \mathcal{A}, \mu, \mathbb{R}]$$

2. Let  $f \in \mathcal{L}[X, \mathcal{A}, \mu, \mathbb{C}]$  then  $f \in \mathbb{C}^X$  and  $\operatorname{Re}(f), \operatorname{Img}(f)$  are  $\mu$ -integrable, hence by definition  $\operatorname{Re}(f), \operatorname{Img}(f)$  are  $\mathcal{A}$ -measurable. So using 19.246 it follows that  $f$  is  $\mathcal{A}, \mathcal{B}[\mathbb{C}]$ -measurable. We can then apply 19.251 proving that  $|f|^1 = |f|$  is  $\mu$ -integrable, hence  $f \in \mathcal{L}^1[X, \mathcal{A}, \mu, \mathbb{C}]$  proving that  $\mathcal{L}[X, \mathcal{A}, \mu, \mathbb{C}] \subseteq \mathcal{L}^1[X, \mathcal{A}, \mu, \mathbb{C}]$ . Further if  $f \in \mathcal{L}^1[X, \mathcal{A}, \mu, \mathbb{R}]$  then  $f$  is  $\mathcal{A}, \mathcal{B}[\mathbb{C}]$ -measurable and  $|f| = |f|^1$  is  $\mu$ -integrable, which by 19.251 proves that  $f$  is  $\mu$ -integrable, hence  $f \in \mathcal{L}[X, \mathcal{A}, \mu, \mathbb{C}]$ . So

$$\mathcal{L}[X, \mathcal{A}, \mu, \mathbb{C}] = \mathcal{L}^1[X, \mathcal{A}, \mu, \mathbb{C}] \quad \square$$

Next we want to check if for  $f, g \in \mathcal{L}^p[X, \mathcal{A}, \mathbb{C}]$  [or  $f, g \in \mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{R}]$ ] and  $\alpha \in \mathbb{R}$  we have that  $f + g \in \mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{C}]$  and  $\alpha \cdot f \in \mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{R}]$  [or  $f + g \in \mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{C}]$ ]. For this we need the following two lemmas.

**Lemma 19.255.** *Let  $p \in [1, \dots, \infty[$  then we have that*

1.  $||^p: \mathbb{R} \rightarrow \mathbb{R}$  defined by  $||^p(x) = |x|^p$  is  $\mathcal{B}[\mathbb{R}], \mathcal{B}[\mathbb{R}]$ -measurable
2. If  $f \in \mathcal{M}[X, \mathcal{A}]$  then  $|f|^p \in \mathcal{M}^+[X, \mathcal{A}]$
3.  $||^p: \mathbb{C} \rightarrow \mathbb{R}$  defined by  $||^p(z) = |z|^p$  is  $\mathcal{B}[\mathbb{C}], \mathcal{B}[\mathbb{R}]$ -measurable
4. If  $f \in \mathcal{M}[X, \mathcal{A}, \mathbb{C}]$  then  $|f|^p \in \mathcal{M}^+[X, \mathcal{A}]$

**Proof.**

1. First by 11.162 the norm function  $||: \mathbb{R} \rightarrow \mathbb{R}$  is continuous [using the topologies  $\mathcal{T}_{\mathbb{R}}$  and  $\mathcal{T}_{\mathbb{R}}$  for the domain and codomain]. Consider now the subspace topology  $\mathcal{T}_{\mathbb{R}} \cap [0, \infty[$  [see 11.6], then if  $U \in \mathcal{T}_{\mathbb{R}} \cap [0, \infty[$  there exists a  $U' \in \mathcal{T}_{\mathbb{R}}$  such that  $U = U' \cap [0, \infty[$ , so

$$||^{-1}(U) = ||^{-1}(U \cap [0, \infty[) \stackrel{2.62}{=} ||^{-1}(U) \cap ||^{-1}([0, \infty[)_{\text{range}(||) = [0, \infty[} ||^{-1}(U) \in \mathcal{T}_{\mathbb{R}}$$

which proves that

$$||: \mathbb{R} \rightarrow [0, \infty[ \text{ is continuous using the topologies } \mathcal{T}_{\mathbb{R}}, \mathcal{T}_{\mathbb{R}} \cap [0, \infty[$$

Further using 18.40 we have that  $[.]^p: [0, \infty[ \rightarrow \mathbb{R}$  is continuous [using the topologies  $\mathcal{T}_{\mathbb{R}} \cap [0, \infty[$ ]. Hence as the compositions of continuous functions is continuous it follows that

$$[.]^p \circ || \text{ is continuous using the topologies } \mathcal{T}_{\mathbb{R}}, \mathcal{T}_{\mathbb{R}}$$

Finally, as  $([.]^p \circ ||)(x) = [.]^p(|x|) = |x|^p = (||^p)(x)$ , hence  $||^p = [.]^p \circ ||$ . it follows that  $||^p$  is continuous [using the topologies  $\mathcal{T}_{\mathbb{R}}, \mathcal{T}_{\mathbb{R}}$ ]. Using then 19.143 we have

$$||^p \text{ is } \mathcal{B}[\mathbb{R}], \mathcal{B}[\mathbb{R}]\text{-measurable.}$$

2. Let  $x \in X$  then  $(||^p \circ f)(x) = ||^p(f(x)) = |f(x)|^p$  so that  $|f|^p = ||^p \circ f$ . As  $f$  is  $\mathcal{A}, \mathcal{B}[\mathbb{R}]$ -measurable and by (1)  $||^p$  is  $\mathcal{B}[\mathbb{R}], \mathcal{B}[\mathbb{R}]$  we have by 19.144 that  $|f|^p$  is  $\mathcal{A}, \mathcal{B}[\mathbb{R}]$ -measurable or  $|f|^p \in \mathcal{M}^+[X, \mathcal{A}]$ .

3. First by 11.162 the norm function  $||: \mathbb{C} \rightarrow \mathbb{R}$  is continuous [using the topologies  $\mathcal{T}_{\mathbb{C}}$  and  $\mathcal{T}_{\mathbb{R}}$  for the domain and codomain]. Consider now the subspace topology  $\mathcal{T}_{\mathbb{R}} \cap [0, \infty[$  [see 11.6], then if  $U \in \mathcal{T}_{\mathbb{R}} \cap [0, \infty[$  there exists a  $U' \in \mathcal{T}_{\mathbb{R}}$  such that  $U = U' \cap [0, \infty[$ , so

$$||^{-1}(U) = ||^{-1}(U \cap [0, \infty[) \stackrel{2.62}{=} ||^{-1}(U) \cap ||^{-1}([0, \infty[)_{\text{range}(||) = [0, \infty[} ||^{-1}(U) \in \mathcal{T}_{\mathbb{C}}$$

which proves that

$$||: \mathbb{C} \rightarrow [0, \infty[ \text{ is continuous using the topologies } \mathcal{T}_{\mathbb{C}}, \mathcal{T}_{\mathbb{R}} \cap [0, \infty[$$

Further using 18.40 we have that  $[.]^p: [0, \infty[ \rightarrow \mathbb{R}$  is continuous [using the topologies  $\mathcal{T}_{\mathbb{R}} \cap [0, \infty[$ ]. Hence as the compositions of continuous functions is continuous it follows that

$$[.]^p \circ || \text{ is continuous using the topologies } \mathcal{T}_{\mathbb{C}}, \mathcal{T}_{\mathbb{R}}$$

Finally, as  $([.]^p \circ ||)(x) = [.]^p(|x|) = |x|^p = (||^p)(x)$ , hence  $||^p = [.]^p \circ ||$ , it follows that  $||^p$  is continuous [using the topologies  $\mathcal{T}_{\mathbb{C}}, \mathcal{T}_{\mathbb{R}}$ ]. Using then 19.143 we have

$$||^p \text{ is } \mathcal{B}[\mathbb{C}], \mathcal{B}[\mathbb{R}]\text{-measurable.}$$

4. Let  $x \in X$  then  $(||^p \circ f)(x) = ||^p(f(x)) = |f(x)|^p$  so that  $|f|^p = ||^p \circ f$ . As  $f$  is  $\mathcal{A}, \mathcal{B}[\mathbb{C}]$ -measurable and by (3)  $||^p$  is  $\mathcal{B}[\mathbb{C}], \mathcal{B}[\mathbb{R}]$  we have by 19.144 that  $|f|^p$  is  $\mathcal{A}, \mathcal{B}[\mathbb{R}]$ -measurable or  $|f|^p \in \mathcal{M}^+[X, \mathcal{A}]$ .

□

**Lemma 19.256.** *Let  $p \in ]0, \dots, \infty[$  and  $z_1, z_2 \in \mathbb{R}$  /or  $z_1, z_2 \in \mathbb{C}$  / then  $|z_1 + z_2|^p \leq 2^p \cdot (|z_1|^p + |z_2|^p)$*

**Proof.** For  $|z_1|, |z_2|$  we must consider the following cases:

$|z_1| \leq |z_2|$ . Then  $|z_1 + z_2| \leq |z_1| + |z_2| \leq 2 \cdot |z_2|$ , as  $p \in ]0, \infty[$  we have by 18.39 (7),(8) that

$$|z_1 + z_2|^p \leq 2^p \cdot |z_2|^p \leq 2^p \cdot (|z_1|^p + |z_2|^p)$$

$|z_2| < |z_1|$ . Then  $|z_1 + z_2| \leq |z_1| + |z_2| < 2 \cdot |z_1|$ , as  $p \in ]0, \infty[$  we have by 18.39 (7),(8) that

$$|z_1 + z_2|^p < 2^p \cdot |z_1|^p \leq 2^p \cdot (|z_1|^p + |z_2|^p)$$

So in all cases we have

$$|z_1 + z_2|^p \leq 2^p \cdot (|z_1|^p + |z_2|^p) \quad \square$$

**Lemma 19.257.** Let  $p \in [1, \infty[$  and  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space then we have

1.  $\forall f, g \in \mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{R}]$  we have  $f + g \in \mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{R}]$
2.  $\forall f \in \mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{R}]$  and  $\alpha \in \mathbb{R}$  we have that  $\alpha \cdot f \in \mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{R}]$
3.  $\forall f, g \in \mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{C}]$  we have  $f + g \in \mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{C}]$
4.  $\forall f \in \mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{C}]$  and  $\mathbb{C}$  we have that  $\alpha \cdot f \in \mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{C}]$

**Proof.**

1. As  $f, g \in \mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{R}]$  we have that  $f, g \in \mathcal{M}[X, \mathcal{A}]$  and  $|f|^p, |g|^p$  are  $\mu$ -integrable. From 19.170 it follows that  $f + g \in \mathcal{M}[X, \mathcal{A}]$ , finally applying 19.255 gives

$$|f + g|^p \in \mathcal{M}^+[X, \mathcal{A}] \quad (19.483)$$

Let  $x \in X$  then

$$\begin{aligned} |f + g|^p(x) &= |f(x) + g(x)|^p \\ &\leq_{19.256} 2^p \cdot (|f(x)|^p + |g(x)|^p) \end{aligned}$$

hence using 19.209 we have that

$$\int^+ |f + g|^p d\mu \leq 2^p \cdot \left( \int^+ |f|^p d\mu + \int^+ |g|^p d\mu \right) \quad (19.484)$$

Now as  $|f|^p, |g|^p$  are  $\mu$ -integrable we have by 19.221 that  $\int^+ |f|^p d\mu < \infty$  and  $\int^+ |g|^p d\mu < \infty$ . So it follows from 19.484 that  $\int^+ |f + g|^p d\mu < \infty$ . Combining this with 19.483 and 19.221 proves that  $|f + g|^p$  is  $\mu$ -integrable or as  $f + g \in \mathcal{M}[X, \mathcal{A}]$

$$|f + g|^p \in \mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{R}]$$

2. As  $f \in \mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{R}]$  we have that  $f \in \mathcal{M}[X, \mathcal{A}]$  and  $|f|^p$  is  $\mu$ -integrable. From 19.170 it follows that  $\alpha \cdot f \in \mathcal{M}[X, \mathcal{A}]$ . Let  $x \in X$  then  $|\alpha \cdot f|^p(x) = |\alpha \cdot f(x)|^p \stackrel{9.33 \text{ and } 18.39}{=} |\alpha|^p \cdot |f(x)|^p$  proving that  $|\alpha \cdot f|^p = |\alpha|^p \cdot |f|^p$  which as  $|f|^p$  is  $\mu$ -integrable proves by 19.225 that  $|\alpha \cdot f|^p$  is  $\mu$ -integrable or as  $\alpha \cdot f \in \mathcal{M}[X, \mathcal{A}]$

$$\alpha \cdot f \in \mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{R}]$$

3. As  $f, g \in \mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{C}]$  we have that  $f, g \in \mathcal{M}[X, \mathcal{A}, \mathbb{C}]$  and  $|f|^p, |g|^p$  are  $\mu$ -integrable. From 19.247 it follows that  $f + g \in \mathcal{M}[X, \mathcal{A}, \mathbb{C}]$ , finally applying 19.255 gives

$$|f + g|^p \in \mathcal{M}^+[X, \mathcal{A}] \quad (19.485)$$

Let  $x \in X$  then

$$\begin{aligned} |f + g|^p(x) &= |f(x) + g(x)|^p \\ &\leq_{19.256} 2^p \cdot (|f(x)|^p + |g(x)|^p) \end{aligned}$$

hence using 19.209 we have that

$$\int^+ |f + g|^p d\mu \leq 2^p \cdot \left( \int^+ |f|^p d\mu + \int^+ |g|^p d\mu \right) \quad (19.486)$$

Now as  $|f|^p, |g|^p$  are  $\mu$ -integrable we have by 19.221 that  $\int^+ |f|^p d\mu < \infty$  and  $\int^+ |g|^p d\mu < \infty$ . So it follows from 19.486 that  $\int^+ |f+g|^p d\mu < \infty$ . Combining this with 19.485 and 19.221 proves that  $|f+g|^p$  is  $\mu$ -integrable or as  $f+g \in \mathcal{M}[X, \mathcal{A}, \mathbb{C}]$

$$f+g \in \mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{C}]$$

4. As  $f \in \mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{R}]$  we have that  $f \in \mathcal{M}[X, \mathcal{A}, \mathbb{C}]$  and  $|f|^p$  is  $\mu$ -integrable. From 19.247 it follows that  $\alpha \cdot f \in \mathcal{M}[X, \mathcal{A}, \mathbb{C}]$ . Let  $x \in X$  then  $|\alpha \cdot f(x)| = |\alpha| \cdot |f(x)|^p \stackrel{9.33 \text{ and } 18.39}{=} |\alpha|^p \cdot |f(x)|^p$  proving that  $|\alpha \cdot f|^p = |\alpha|^p \cdot |f|^p$  which as  $|f|^p$  is  $\mu$ -integrable proves by 19.225 that  $|\alpha \cdot f|^p$  is  $\mu$ -integrable or as  $\alpha \cdot f \in \mathcal{M}[X, \mathcal{A}, \mathbb{C}]$

$$\alpha \cdot f \in \mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{C}] \quad \square$$

**Definition 19.258.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space then  $\mathcal{L}^\infty[X, \mathcal{A}, \mu, \mathbb{R}]$  and  $\mathcal{L}^\infty[X, \mathcal{A}, \mu, \mathbb{C}]$  are defined by:

1.  $\mathcal{L}^\infty[X, \mathcal{A}, \mu, \mathbb{R}] = \{f \in \mathcal{M}[X, \mathcal{A}] \mid \exists M \in \mathbb{R} \text{ such that } \forall x \in X \text{ we have } |f(x)| \leq M\}$
2.  $\mathcal{L}^\infty[X, \mathcal{A}, \mu, \mathbb{C}] = \{f \in \mathcal{M}[X, \mathcal{A}, \mathbb{C}] \mid \exists M \in \mathbb{R} \text{ such that } \forall x \in X \text{ we have } |f(x)| \leq M\}$

**Lemma 19.259.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space then

1.  $\forall f, g \in \mathcal{L}^\infty[X, \mathcal{A}, \mu, \mathbb{R}] \text{ we have } f+g \in \mathcal{L}^\infty[X, \mathcal{A}, \mu, \mathbb{R}]$
2.  $\forall f \in \mathcal{L}^\infty[X, \mathcal{A}, \mu, \mathbb{R}] \text{ and } \alpha \in \mathbb{R} \text{ we have that } \alpha \cdot f \in \mathcal{L}^\infty[X, \mathcal{A}, \mu, \mathbb{R}]$
3.  $\forall f, g \in \mathcal{L}^\infty[X, \mathcal{A}, \mu, \mathbb{C}] \text{ we have } f+g \in \mathcal{L}^\infty[X, \mathcal{A}, \mu, \mathbb{C}]$
4.  $\forall f \in \mathcal{L}^\infty[X, \mathcal{A}, \mu, \mathbb{C}] \text{ and } \mathbb{C} \text{ we have that } \alpha \cdot f \in \mathcal{L}^\infty[X, \mathcal{A}, \mu, \mathbb{C}]$

**Proof.**

1. As  $f, g \in \mathcal{L}^\infty[X, \mathcal{A}, \mu, \mathbb{R}]$  we have that  $f, g \in \mathcal{M}[X, \mathcal{A}]$ , hence by 19.170  $f+g \in \mathcal{M}[X, \mathcal{A}]$ , and there exists a  $M_1, M_2$  such that  $\forall x \in X$  we have  $|f(x)| \leq M_1, |g(x)| \leq M_2$ . If  $M = M_1 + M_2$  then  $\forall x \in X$  we have  $|(f+g)(x)| = |f(x) + g(x)| \leq |f(x)| + |g(x)| \leq M_1 + M_2 = M$ . So we have that

$$f+g \in \mathcal{L}^\infty[X, \mathcal{A}, \mu, \mathbb{R}]$$

2. As  $f \in \mathcal{L}^\infty[X, \mathcal{A}, \mu, \mathbb{R}]$  we have that  $f \in \mathcal{M}[X, \mathcal{A}]$ , hence by 19.170  $\alpha \cdot f \in \mathcal{M}[X, \mathcal{A}]$ , and there exists a  $M' \in \mathbb{R}$  such that  $\forall x \in X$   $|f(x)| \leq M'$ . Take  $M = |\alpha| \cdot |M'| \in [0, \infty[$  then given  $x \in X$  we have  $|(\alpha \cdot f)(x)| = |\alpha \cdot f(x)| = |\alpha| \cdot |f(x)| \leq |\alpha| \cdot M' = M$ . So we have

$$\alpha \cdot f \in \mathcal{L}[X, \mathcal{A}, \mu, \mathbb{R}]$$

3. As  $f, g \in \mathcal{L}^\infty[X, \mathcal{A}, \mu, \mathbb{C}]$  we have that  $f, g \in \mathcal{M}[X, \mathcal{A}, \mathbb{C}]$ , hence by 19.247  $f+g \in \mathcal{M}[X, \mathcal{A}, \mathbb{C}]$ , and a  $M_1, M_2$  such that  $\forall x \in X$  we have  $|f(x)| \leq M_1, |g(x)| \leq M_2$ . If  $M = M_1 + M_2$  then  $\forall x \in X$  we have  $|(f+g)(x)| = |f(x) + g(x)| \leq |f(x)| + |g(x)| \leq M_1 + M_2 = M$ . So we have that

$$f+g \in \mathcal{L}^\infty[X, \mathcal{A}, \mu, \mathbb{C}]$$

4. As  $f \in \mathcal{L}^\infty[X, \mathcal{A}, \mu, \mathbb{C}]$  we have that  $f \in \mathcal{M}[X, \mathcal{A}, \mathbb{C}]$ , hence by 19.247  $\alpha \cdot f \in \mathcal{M}[X, \mathcal{A}, \mathbb{C}]$  and there exists a  $M' \in \mathbb{R}$  such that  $\forall x \in X$   $|f(x)| \leq M'$ . Take  $M = |\alpha| \cdot |M'| \in [0, \infty[$  then given  $x \in X$  we have  $|(\alpha \cdot f)(x)| = |\alpha \cdot f(x)| = |\alpha| \cdot |f(x)| \leq |\alpha| \cdot M' = M$ . So we have

$$\alpha \cdot f \in \mathcal{L}[X, \mathcal{A}, \mu, \mathbb{C}] \quad \square$$

Now to simplify notation we use the following convention if we just use the common properties of  $\mathbb{R}$  and  $\mathbb{C}$ .

**Notation 19.260.**  $\mathbb{K}$  means either  $\mathbb{C}$  or  $\mathbb{K}$ . So

1.  $\mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{K}]$  means either  $\mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{C}]$  or  $\mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{R}]$

2.  $\mathcal{L}^\infty[X, \mathcal{A}, \mu, \mathbb{K}]$  means either  $\mathcal{L}^\infty[X, \mathcal{A}, \mu, \mathbb{R}]$
3.  $\mathcal{M}[X, \mathcal{A}, \mathbb{K}]$  means either  $\mathcal{M}[X, \mathcal{A}]$  or  $\mathcal{M}[X, \mathcal{A}, \mathbb{C}]$

Now that the sum and scalar product are defined by the above propositions we can proof that  $\mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{R}]$  and  $\mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{C}]$  are actually vector spaces.

**Corollary 19.261.** Let  $p \in [1, \dots, \infty[$  then

1.  $\langle \mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{K}], +, \cdot \rangle$  is a vector space where  $\forall f, g \in \mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{K}]$  and  $\forall \alpha \in \mathbb{K}$  we have that  $f + g$  and  $\alpha \cdot f$  is defined by  $(f + g)(x) = f(x) + g(x)$  and  $(\alpha \cdot f)(x) = \alpha \cdot f(x)$ .
2.  $\langle \mathcal{L}^\infty[X, \mathcal{A}, \mu, \mathbb{K}], +, \cdot \rangle$  is a vector space where  $\forall f, g \in \mathcal{L}^\infty[X, \mathcal{A}, \mu, \mathbb{K}]$  and  $\alpha \in \mathbb{K}$  we have that  $f + g$  and  $\alpha \cdot f$  is defined by  $(f + g)(x) = f(x) + g(x)$  and  $(\alpha \cdot f)(x) = \alpha \cdot f(x)$ .

**Proof.** This follows from 10.115, 10.111 and 19.257, 19.259  $\square$

**Definition 19.262.** Let  $p \in [1, \dots, \infty[$  and  $\langle X, \mathcal{A}, \mu \rangle$  a measure space, define then  $\|\cdot\|_p$  by

$$\|\cdot\|_p: \mathcal{L}[X, \mathcal{A}, \mu, \mathbb{K}] \rightarrow [0, \infty[ \text{ by } \|f\|_p = \left( \int |f|^p d\mu \right)^{\frac{1}{p}}$$

**Proposition 19.263.** Let  $p \in [1, \dots, \infty[$ ,  $\langle X, \mathcal{A}, \mu \rangle$  a measure space and  $f \in \mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{K}]$  such that  $\|f\|_p = 0$  then  $f = 0$   $\mu$ -a.e.

**Proof.** As  $\|f\|_p = 0$  we have  $(\int |f|^p d\mu)^{\frac{1}{p}} = 0$ , giving using 18.39  $\int |f|^p d\mu = 0$  or as  $|(|f|^p)| = |f|^p$  that  $\int |(|f|^p)| d\mu = 0$ . By 19.232 it follows that  $|f|^p = 0$   $\mu$ -a.e. So there exists a  $N \in \mathcal{A}$  with  $\mu(N) = 0$  and  $\{x \in X \mid |f(x)| \neq 0\} \subseteq N$ . If now  $x \in \{x \in X \mid |f(x)| \neq 0\}$  then  $f(x) \neq 0 \Rightarrow |f(x)| \neq 0 \xrightarrow{18.39} |f(x)|^p \neq 0$  proving that  $\{x \in X \mid f(x) \neq 0\} \subseteq \{x \in X \mid |f(x)|^p \neq 0\} \subseteq N$  hence  $f = 0$   $\mu$ -a.e.  $\square$

Now to define  $\|\cdot\|_\infty$  on  $\mathcal{L}^\infty[X, \mathcal{A}, \mu, \mathbb{R}], \mathcal{L}^\infty[X, \mathcal{A}, \mu, \mathbb{C}]$  we introduce the concept of **locally  $\mu$ -null sets**.

**Definition 19.264.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space then  $N \subseteq X$  is **locally  $\mu$ -null** if  $\forall A \in \mathcal{A}$  with  $\mu(A) < \infty$  we have  $A \cap N \in \mathcal{N}_\mu$  [see 19.104 for the definition negligible sets  $\mathcal{N}_\mu$ ]

**Example 19.265.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space then  $\emptyset$  is **locally  $\mu$ -null**

**Proof.** Let  $A \in \mathcal{A}$  then  $A \cap \emptyset = \emptyset \in \mathcal{A}$  with  $\mu(\emptyset) = 0$  proving that  $\emptyset$  is locally  $\mu$ -null.  $\square$

We have the following properties of locally null sets

**Proposition 19.266.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measurable space then

1. Every negligible set is locally null [in other words if  $A \in \mathcal{N}_\mu$  then  $A$  is locally  $\mu$ -null]
2. If  $A \subseteq B$  and  $B$  is locally  $\mu$ -null then  $A$  is locally  $\mu$ -null
3. If  $\{A_i\}_{i \in \mathbb{N}_0}$  is a sequence of locally  $\mu$ -null sets then  $\bigcup_{i \in \mathbb{N}_0} A_i$  is locally  $\mu$ -null

**Proof.**

1. If  $B \in \mathcal{N}_\mu$  then there exists a  $N \in \mathcal{A}$  with  $\mu(N) = 0$  such that  $B \subseteq N$ . If now  $A \in \mathcal{A}$  with  $\mu(A) < \infty$  then  $A \cap B \subseteq B \subseteq N$  so that  $A \cap B \in \mathcal{N}_\mu$  proving that  $B$  is locally  $\mu$ -null.
2. Let  $D \in \mathcal{A}$  with  $\mu(D) < \infty$  then there exists a  $N \in \mathcal{A}$  with  $\mu(N) = 0$  such that  $D \cap B \subseteq N$ , as  $A \subseteq B$  we have  $D \cap A \subseteq A \cap B \subseteq N$  proving that  $A$  is locally  $\mu$ -null.
3. Let  $A \in \mathcal{A}$  with  $\mu(A) < \infty$  then as  $\forall i \in \mathbb{N}_0$  we have that  $A_i$  is locally  $\mu$ -null there exists a  $N_i \in \mathcal{A}$  with  $\mu(N_i) = 0$  such that  $A \cap A_i \subseteq N_i$ . Now  $A \cap (\bigcup_{i \in \mathbb{N}_0} A_i) = \bigcup_{i \in \mathbb{N}_0} (A \cap A_i) \subseteq \bigcup_{i \in \mathbb{N}_0} N_i$  which as  $\mu(\bigcup_{i \in \mathbb{N}_0} N_i) \leq \sum_{i=1}^{\infty} \mu(N_i) = 0$  proves that  $A \cap (\bigcup_{i \in \mathbb{N}_0} A_i) \in \mathcal{N}_\mu$  and finally that  $\bigcup_{i \in \mathbb{N}_0} A_i$  is locally  $\mu$ -null.  $\square$

**Definition 19.267.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space then we define  $\|\cdot\|_\infty: \mathcal{L}^\infty[X, \mathcal{A}, \mu, \mathbb{K}] \rightarrow [0, \infty[$  by

$$\|f\|_\infty = \inf(\{M \in [0, \infty[ \mid \{x \in X \mid |f(x)| > M\}\} \text{ is locally } \mu\text{-null})$$

**Proof.** We must of course proof that the infimum exists and is finite. If  $f \in \mathcal{L}^\infty[X, \mathcal{A}, \mu, \mathbb{K}]$  then  $f \in \mathcal{M}[X, \mathcal{A}, \mathbb{K}]$  and there exists a  $M \in \mathbb{R}$  such that  $\forall x \in X$  we have  $|f(x)| \leq M$ . Now if  $x \in \{x \in X \mid |f(x)| > M\}$  we have that  $|f(x)| > M$  contradicting  $|f(x)| \leq M$  hence  $\{x \in X \mid |f(x)| > M\} = \emptyset$  which is **locally  $\mu$ -null**. So  $\{M \in [0, \infty[ \mid \{x \in X \mid |f(x)| > M\}\} \text{ is locally } \mu\text{-null}\} \neq \emptyset$  and its infimum exists, where

$$0 \leq \inf(\{M \in [0, \infty[ \mid \{x \in X \mid |f(x)| \geq M\}\} \text{ is locally null}) < M < \infty$$

□

**Lemma 19.268.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measurable space and  $f \in \mathcal{L}^\infty[X, \mathcal{A}, \mu, \mathbb{K}]$  then

$$\{x \in X \mid |f(x)| > \|f\|_\infty\} \text{ is locally } \mu\text{-null}$$

**Proof.** Let  $n \in \mathbb{N}_0$  then as

$$\inf(\{M \in [0, \infty[ \mid \{x \in X \mid |f(x)| > M\}\} \text{ is locally } \mu\text{-null}) = \|f\|_\infty < \|f\|_\infty + \frac{1}{n}$$

there exists a  $M_n$  such that

$$\{x \in X \mid |f(x)| > M_n\} \text{ is locally } \mu\text{-null and } \|f\|_\infty \leq M_n < \|f\|_\infty + \frac{1}{n} \quad (19.487)$$

If  $x \in \{x \in X \mid |f(x)| > \|f\|_\infty\}$  then  $\|f\|_\infty < |f(x)|$  so that using 9.56 there exists a  $n \in \mathbb{N}_0$  such that  $\|f\|_\infty < \|f\|_\infty + \frac{1}{n} < |f(x)|$ , hence using 19.487 we have that  $M_n < \|f\|_\infty + \frac{1}{n} \Rightarrow M_n < |f(x)|$  proving that  $x \in \{x \in X \mid |f(x)| > M_n\}$ . So we have that

$$\{x \in X \mid |f(x)| > \|f\|_\infty\} \subseteq \bigcup_{n \in \mathbb{N}_0} \{x \in X \mid |f(x)| > M_n\} \quad (19.488)$$

If  $x \in \bigcup_{n \in \mathbb{N}_0} \{x \in X \mid |f(x)| > M_n\}$  there exists a  $n \in \mathbb{N}_0$  such that  $x \in \{x \in X \mid |f(x)| > M_n\}$  or  $|f(x)| > M_n \geq_{19.487} \|f\|_\infty$  giving  $x \in \{x \in X \mid |f(x)| > \|f\|_\infty\}$ . So

$$\bigcup_{n \in \mathbb{N}_0} \{x \in X \mid |f(x)| > M_n\} \subseteq \{x \in X \mid |f(x)| > \|f\|_\infty\}$$

which combined with 19.488 gives

$$\{x \in X \mid |f(x)| > \|f\|_\infty\} = \bigcup_{n \in \mathbb{N}_0} \{x \in X \mid |f(x)| > M_n\}$$

Using then 19.487 together with 19.266 on the above proves that

$$\{x \in X \mid |f(x)| > \|f\|_\infty\} \text{ is locally } \mu\text{-null}$$

□

**Theorem 19.269. (Hölder's Inequality)** Let  $p, q \in [1, \infty]$  such that  $\frac{1}{p} + \frac{1}{q} = 1$ ,  $\langle X, \mathcal{A}, \mu \rangle$  a measure space and  $f \in \mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{K}]$ ,  $g \in \mathcal{L}^q[X, \mathcal{A}, \mu, \mathbb{K}]$  then  $f \cdot g \in \mathcal{L}^1[X, \mathcal{A}, \mu, \mathbb{K}]$  and

$$\int |f \cdot g| d\mu \leq \|f\|_p \cdot \|g\|_q$$

**Proof.** We have for  $p, q$  the following cases to consider:

$p = \infty$ . Then as  $1 = \frac{1}{p} + \frac{1}{q} = \frac{1}{\infty} + \frac{1}{q} = \frac{1}{q}$  we have  $q = 1$ . Hence  $g \in \mathcal{L}^1[X, \mathcal{A}, \mu, \mathbb{K}]$ , so we have that  $\int |g| d\mu$  is  $\mu$ -integrable. By 19.231 we have that  $\{x \in X \mid |g(x)| \neq 0\}$  is  $\sigma$ -finite. So there exists a  $\{A_i\}_{i \in \mathbb{N}_0}$  with  $\forall i \in \mathbb{N}_0 \mu(A_i) < \infty$  such that

$$A = \{x \in X \mid g(x) \neq 0\} \subseteq \bigcup_{i \in \mathbb{N}_0} A_i \quad (19.489)$$

Further as  $f \in \mathcal{L}^\infty[X, \mathcal{A}, \mu, \mathbb{K}]$  we have by the previous lemma [see 19.268] that

$$B = \{x \in X \mid |f(x)| > \|f\|_\infty\} \text{ is locally } \mu\text{-null} \quad (19.490)$$

Now

$$\begin{aligned} A \cap B &\subseteq \left( \bigcup_{i \in \mathbb{N}_0} A_i \right) \cap \{x \in X \mid |f(x)| > \|f\|_\infty\} \\ &= \bigcup_{i \in \mathbb{N}_0} (A_i \cap \{x \in X \mid |f(x)| > \|f\|_\infty\}) \end{aligned} \quad (19.491)$$

Now as  $\{x \in X \mid |f(x)| > \|f\|_\infty\}$  is locally  $\mu$ -null and  $\mu(A_i) < \infty$  we have that  $A_i \cap \{x \in X \mid |f(x)| > \|f\|_\infty\} \in \mathcal{N}_\mu$ . So there exists a  $N_i \in \mathcal{A}$  with such that  $\mu(N_i) = 0$  such that  $A_i \cap \{x \in X \mid |f(x)| > \|f\|_\infty\} \subseteq N_i$ . Hence substituting this in 19.491 we have that

$$\{x \in X \mid g(x) \neq 0\} \cap \{x \in X \mid |f(x)| > \|f\|_\infty\} \subseteq \bigcup_{i \in \mathbb{N}_0} N_i \quad (19.492)$$

If  $x \in X$  is such that  $|f \cdot g|(x) = |f(x)| \cdot |g(x)| = |f(x) \cdot g(x)| > \|f\|_\infty \cdot |g(x)|$  then we must have  $g(x) \neq 0$  [otherwise we would have  $0 > 0$ ] so that after dividing by  $|g(x)|$  we have that  $|f(x)| > \|f\|_\infty$ , hence  $x \in \{x \in X \mid g(x) \neq 0\} \cap \{x \in X \mid |f(x)| > \|f\|_\infty\}$ . This proves using 19.492 that

$$\{x \in X \mid |f \cdot g| > \|f\|_\infty \cdot |g(x)|\} \subseteq \{x \in X \mid g(x) \neq 0\} \cap \{x \in X \mid |f(x)| > \|f\|_\infty\} \subseteq \bigcup_{i \in \mathbb{N}_0} N_i$$

which as  $\mu(\bigcup_{i \in \mathbb{N}_0} N_i) \leq \sum_{i=1}^{\infty} \mu(N_i) = 0$  gives

$$|f \cdot g| \leq \|f\|_\infty \cdot |g| \text{ } \mu\text{-a.e.} \quad (19.493)$$

As  $f \in \mathcal{L}^\infty[X, \mathcal{A}, \mu, \mathbb{K}]$ ,  $g \in \mathcal{L}^1[X, \mathcal{A}, \mu, \mathbb{K}]$  we have  $f, g \in \mathcal{M}[X, \mathcal{A}, \mu, \mathbb{K}]$ , so using 19.173, 19.177 we have that  $|f \cdot g| \in \mathcal{M}^+[X, \mathcal{A}, \mu, \mathbb{K}]$ . Further from  $g \in \mathcal{L}^1[X, \mathcal{A}, \mu, \mathbb{K}]$  it follows that  $|g|$  is  $\mu$ -integrable, hence by 19.225  $\|f\|_\infty \cdot |g|$  is  $\mu$ -integrable. Applying then 19.223 on 19.493 proves that

$$|f \cdot g| \in \mathcal{L}[X, \mathcal{A}, \mu, \mathbb{K}] \xrightarrow{19.226} f \cdot g \in \mathcal{L}^1[X, \mathcal{A}, \mu] \text{ and } \int |f \cdot g| d\mu \leq \|f\|_\infty \cdot \int |g| d\mu \quad (19.494)$$

As  $\|g\|_q = \|g\|_1 = (\int |g|^1 d\mu)^{\frac{1}{q}} = \int |g| d\mu$  we have by the above

$$f \cdot g \in \mathcal{L}^1[X, \mathcal{A}, \mu, \mathbb{K}] \text{ and } \int |f \cdot g| d\mu = \|f\|_p \cdot \|g\|_q$$

$q = \infty$ . This is the same as the  $p = \infty$  case if we interchange  $p$  with  $q$  giving

$$\int |f \cdot g| d\mu \leq \|f\|_p \cdot \|g\|_q$$

$p, q \in [1, \infty[$ . Let  $x \in X$  then we have either:

$$f(x) = \mathbf{0} \vee g(x) = \mathbf{0}. \text{ Then } |f(x) \cdot g(x)| = 0 \leq \frac{\|f(x)\|^p}{p} \cdot \frac{\|g(x)\|^p}{q}$$

$f(x) \neq \mathbf{0} \neq g(x)$ . Then using Young's inequality [see 18.42] we have  $\forall x \in X$  that

$$|f(x) \cdot g(x)| = |f(x)| \cdot |g(x)| \leq \frac{|f(x)|^p}{p} + \frac{|g(x)|^q}{q}$$

giving

$$\forall x \in X \text{ we have } |f(x) \cdot g(x)| \leq \frac{|f(x)|^p}{p} + \frac{|g(x)|^q}{q} \quad (19.495)$$

As  $f \in \mathcal{L}^p(X, \mathcal{A}, \mu, \mathbb{K})$ ,  $g \in \mathcal{L}^q(X, \mathcal{A}, \mu, \mathbb{K})$  it follows that  $|f|^p, |g|^q$  are  $\mu$ -integrable, hence by 19.225  $\frac{|f|^p}{p} + \frac{|g|^q}{q}$  is  $\mu$ -integrable. Further as  $f, g \in \mathcal{M}[X, \mathcal{A}, \mu, \mathbb{K}]$  we have by 19.173, 19.177 that  $|f \cdot g| \in \mathcal{M}[X, \mathcal{A}, \mathbb{K}]$ . Using then 19.495 on 19.223 proves that  $|f \cdot g|$  is  $\mu$ -integrable, hence

$$f \cdot g \in \mathcal{L}^1[X, \mathcal{A}, \mu, \mathbb{K}] \quad (19.496)$$

and

$$\int |f \cdot g| d\mu \leq \frac{1}{p} \cdot \int |f|^p d\mu + \frac{1}{q} \cdot \int |g|^q d\mu \quad (19.497)$$

Assume now that  $\|f\|_p = 1 = \|g\|_q$  then  $1 \stackrel{18.39}{=} (\|f\|_p)^p = \left( \left( \int |f|^p d\mu \right)^{\frac{1}{p}} \right)^p \stackrel{18.39}{=} \int |f|^p d\mu$  and  $1 \stackrel{18.39}{=} (\|g\|_q)^q = \left( \left( \int |g|^q d\mu \right)^{\frac{1}{q}} \right)^q \stackrel{18.39}{=} \int |g|^q d\mu$ . Substituting this in 19.497 gives

$$\int |f \cdot g| d\mu \leq \frac{1}{p} + \frac{1}{q} = 1 = \|f\|_p \cdot \|g\|_q$$

To summarize

If  $f \in \mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{K}], g \in \mathcal{L}^q[X, \mathcal{A}, \mu, \mathbb{K}], \|f\|_p = \|g\|_q = 1$  then  $f \cdot g \in \mathcal{L}^1[X, \mathcal{A}, \mu, \mathbb{K}]$  and

$$\int |f \cdot g| d\mu \leq \|f\|_p \cdot \|g\|_q = 1 \quad (19.498)$$

Now for the general case we have two possibilities to consider:

$\|f\|_p \neq 0 \wedge \|g\|_q \neq 0$ . Take then  $f' = \frac{f}{\|f\|_p}$  and  $g' = \frac{g}{\|g\|_q}$  then by 19.257  $f' \in \mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{K}], g' \in \mathcal{L}^q[X, \mathcal{A}, \mu, \mathbb{K}]$  and

$$\begin{aligned} \|f'\|_p &= \left( \int \left| \frac{f}{\|f\|_p} \right|^p d\mu \right)^{\frac{1}{p}} \\ &= \left( \int \left| \frac{f}{\|f\|_p} \right|^p d\mu \right)^{\frac{1}{p}} \\ &= \left( \frac{1}{(\|f\|_p)^p} \cdot \int |f|^p d\mu \right)^{\frac{1}{p}} \\ &= \left( \frac{(\|f\|_p)^p}{(\|f\|_p)^p} \right)^{\frac{1}{p}} \\ &\stackrel{18.39}{=} 1 \\ \|g'\|_q &= \left( \int \left| \frac{g}{\|g\|_q} \right|^q d\mu \right)^{\frac{1}{q}} \\ &= \left( \int \left| \frac{g}{\|g\|_q} \right|^q d\mu \right)^{\frac{1}{q}} \\ &= \left( \frac{1}{(\|g\|_q)^q} \cdot \int |g|^q d\mu \right)^{\frac{1}{q}} \\ &= \left( \frac{(\|g\|_q)^q}{(\|g\|_q)^q} \right)^{\frac{1}{q}} \\ &= 1 \end{aligned}$$

So we can apply 19.498 giving

$$f' \cdot g' \in \mathcal{L}^1[X, \mathcal{A}, \mu, \mathbb{K}] \text{ and } \int |f' \cdot g'| d\mu \leq \|f'\|_p \cdot \|g'\|_q = 1 \quad (19.499)$$

Using 19.257 we have then that  $f \cdot g = \|f\|_p \cdot \|g\|_q \cdot f' \cdot g' \in \mathcal{L}^1[X, \mathcal{A}, \mu, \mathbb{K}]$ . Further as

$$\begin{aligned} \int |f' \cdot g'| d\mu &= \int \left| \frac{1}{\|f\|_p \cdot \|g\|_q} \cdot f \cdot g \right| d\mu \\ &= \frac{1}{\|f\|_p \cdot \|g\|_q} \cdot \int |f \cdot g| d\mu \end{aligned}$$

which combined with 19.499 gives

$$f \cdot g \in \mathcal{L}^1[X, \mathcal{A}, \mu, \mathbb{K}] \text{ and } \int |f \cdot g| d\mu \leq \|f\|_p \cdot \|g\|_q$$

$\|f\|_p = 0$ . Then by 19.263 we have that  $f=0$   $\mu$ -a.e., so there exists a  $N \in \mathcal{A}$  with  $\mu(N)=0$  such that  $\{x \in X | f(x) \neq 0\} \subseteq N$ . If  $x \in \{x \in X | |f \cdot g|(x) \neq 0\}$  then  $f(x) \cdot g(x) \neq 0$  hence  $f(x) \neq 0$  proving that  $\{x \in X | |f \cdot g|(x) \neq 0\} \subseteq \{x \in X | f(x) \neq 0\} \subseteq N$ . So  $|f \cdot g|=0$   $\mu$ -a.e. Hence using 19.496 and 19.252 it follows that  $\int |f \cdot g| d\mu = 0$  proving that

$$\int |f \cdot g| d\mu = 0 \leq 0 = \|f\|_p \cdot \|g\|_q$$

$\|g\|_q = 0$ . Then 19.263 implies that  $g=0$   $\mu$ -a.e., so there exists a  $N \in \mathcal{A}$  with  $\mu(N)=0$  such that  $\{x \in X | g(x) \neq 0\} \subseteq N$ . If  $x \in \{x \in X | |f \cdot g|(x) \neq 0\}$  then  $f(x) \cdot g(x) \neq 0$  hence  $g(x) \neq 0$  proving that  $\{x \in X | |f \cdot g|(x) \neq 0\} \subseteq \{x \in X | f(x) \neq 0\} \subseteq N$ . So  $|f \cdot g|=0$   $\mu$ -a.e. Hence using 19.496 and 19.252 it follows that  $\int |f \cdot g| d\mu = 0$  proving that

$$\int |f \cdot g| d\mu = 0 \leq 0 = \|f\|_p \cdot \|g\|_q$$

So in all cases we have

$$\int |f \cdot g| d\mu \leq \|f\|_p \cdot \|g\|_q$$

which together with 19.496 proves the theorem for the case  $p, q \in [1, \infty[$ .  $\square$

**Theorem 19.270. (Minkowski's Inequality)** Let  $p \in [1, \infty]$ ,  $\langle X, \mathcal{A}, \mu \rangle$  a measure space then  $\forall f, g \in \mathcal{L}[X, \mathcal{A}, \mu, \mathbb{K}]$  then  $f+g \in \mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{K}]$  and  $\|f+g\|_p \leq \|f\|_p + \|g\|_p$

**Proof.** Using 19.257 we have that

$$f+g \in \mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{K}] \quad (19.500)$$

or using the definition of  $\mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{K}]$

$$f+g \in \mathcal{M}[X, \mathcal{A}, \mathbb{K}] \text{ and } |f+g|^p \text{ is } \mu\text{-integrable} \quad (19.501)$$

For  $p \in [1, \infty]$  we have to consider two cases:

**$p = \infty$ .** Define  $F = \{x \in X | |f(x)| > \|f\|_\infty\}$  and  $G = \{x \in X | |g(x)| > \|g\|_\infty\}$  then using 19.268 we have that  $F$  and  $G$  are locally  $\mu$ -null. If  $x \in X \setminus (F \cup G)$  then  $|f(x)| \leq \|f\|_\infty$  and  $|g(x)| \leq \|g\|_\infty$  so that  $|(f+g)(x)| \leq |f(x)| + |g(x)| \leq \|f\|_\infty + \|g\|_\infty$  proving that

$$X \setminus (F \cup G) \subseteq \{x \in X | |(f+g)(x)| \leq \|f\|_\infty + \|g\|_\infty\}$$

Using 1.28 we have than

$$X \setminus \{x \in X | |(f+g)(x)| \leq \|f\|_\infty + \|g\|_\infty\} \subseteq F \cup G$$

Hence  $\{x \in X | |(f+g)(x)| > \|f\|_\infty + \|g\|_\infty\} \subseteq F \cup G$  proving by 19.266 that

$$\{x \in X | |(f+g)(x)| > \|f\|_\infty + \|g\|_\infty\} \text{ is a locally } \mu\text{-null}$$

Hence  $\|f+g\|_\infty \leq \inf(\{M \in [0, \infty] | \{x \in X | |f(x)| > M\} \text{ is locally } \mu\text{-null}\}) \leq \|f\|_\infty + \|g\|_\infty$  proving that

$$\|f+g\|_p \leq \|f\|_p + \|g\|_p$$

**$p = 1$ .** Given  $x \in X$  we have  $\forall x \in X$  that  $|(f+g)x| \leq |f(x)| + |g(x)|$  so using 19.225  $\int |f+g| d\mu \leq \int |f| d\mu + \int |g| d\mu$  proving as  $p=1$  that

$$\|f+g\|_1 \leq \|f\|_1 + \|g\|_1$$

**$p \in ]1, \infty[$ .** Let  $q = \frac{p}{p-1}$  then  $\frac{1}{p} + \frac{1}{q} = \frac{1}{p} + \frac{p-1}{p} = \frac{p}{p} = 1$  and  $(p-1) \cdot q = p$ . So it follows that  $(|f+g|^{p-1})^q = |f+g|^p$  which is  $\mu$ -integrable [see 19.501] proving as  $f+g \in \mathcal{M}[X, \mathcal{A}, \mathbb{K}]$  that

$$|f+g|^{p-1} \in \mathcal{L}^q[X, \mathcal{A}, \mu, \mathbb{K}] \quad (19.502)$$

Now as  $f, g \in \mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{K}]$  and  $\frac{1}{p} + \frac{1}{q} = 1$  we have by 19.269, 19.502 that

$$f \cdot |f + g|^{p-1} \in \mathcal{L}^1[X, \mathcal{A}, \mu, \mathbb{K}] \text{ and } g \cdot |f + g|^{p-1} \in \mathcal{L}^1[X, \mathcal{A}, \mu, \mathbb{K}] \quad (19.503)$$

and

$$\int |f \cdot |f + g|^{p-1}| d\mu \leq \|f\|_p \| |f + g|^{p-1} \|_q \text{ and } \int |g \cdot |f + g|^{p-1}| d\mu \leq \|g\|_p \| |f + g|^{p-1} \|_q \quad (19.504)$$

Further as  $\forall x \in X$  we have that

$$\begin{aligned} (|f + g|(x))^p &\stackrel{18.39}{=} |f + g|(x) \cdot (|f + g|(x))^{p-1} \\ &\leq (|f|(x) + |g|(x)) \cdot (|f + g|(x))^{p-1} \\ &= |f|(x) \cdot (|f + g|(x))^{p-1} + |g|(x) \cdot (|f + g|(x))^{p-1} \end{aligned}$$

so that by 19.225 we have

$$\begin{aligned} \int |f + g|^p d\mu &\leq \int (|f| \cdot |f + g|^{p-1} + |g| \cdot |f + g|^{p-1}) d\mu \\ &\stackrel{19.225}{=} \int |f| \cdot |f + g|^{p-1} d\mu + \int |g| \cdot |f + g|^{p-1} d\mu \\ &\stackrel{19.504}{\leq} \|f\|_p \| |f + g|^{p-1} \|_q + \|g\|_p \| |f + g|^{p-1} \|_q \\ &= (\|f\|_p + \|g\|_p) \cdot \| |f + g|^{p-1} \|_q \\ &= (\|f\|_p + \|g\|_p) \cdot \left( \int |f + g|^{p-1} d\mu \right)^{\frac{1}{q}} \\ &\stackrel{|f+g|^{p-1}(x) \in [0, \infty[}{=} (\|f\|_p + \|g\|_p) \cdot \left( \int (|f + g|^{p-1})^q d\mu \right)^{\frac{1}{q}} \\ &= (\|f\|_p + \|g\|_p) \cdot \left( \int |f + g|^p d\mu \right)^{\frac{1}{q}} \quad (19.505) \end{aligned}$$

For  $\int |f + g|^p d\mu$  we have either:

$$\int |f + g|^p d\mu = \mathbf{0}, \text{ then } \|f + g\|_p = \left( \int |f + g|^p d\mu \right)^{\frac{1}{p}} = 0 \leq \|f\|_p + \|g\|_p \text{ giving}$$

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p$$

$$\int |f + g|^p d\mu \neq \mathbf{0}, \text{ then by 18.39 that } (|f + g|^p d\mu)^{\frac{1}{q}} \neq 0 \text{ so using 19.505 we have that}$$

$$\begin{aligned} \|f\|_p + \|g\|_p &\geq \frac{\int |f + g|^p d\mu}{\left( \int |f + g|^p d\mu \right)^{\frac{1}{q}}} \\ &\stackrel{18.39}{=} \left( \int |f + g|^p d\mu \right) \cdot \left( \int |f + g|^p d\mu \right)^{\frac{1}{q}} \\ &\stackrel{18.39}{=} \left( \int |f + g|^p d\mu \right)^{1 - \frac{1}{q}} \\ &= \left( \int |f + g|^p d\mu \right)^{\frac{1}{p}} \\ &= \|f + g\|_p \end{aligned}$$

proving that

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p$$

So in all cases we have

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p$$

□

**Corollary 19.271.** Let  $p \in [1, \infty]$  and  $\langle X, \mathcal{A}, \mu \rangle$  a measure space then  $\langle \mathcal{L}[X, \mathcal{A}, \mu, \mathbb{K}], \|\cdot\|_p \rangle$  is a pseudo normed space where  $\forall f, g \in \mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{K}]$  and  $\forall \alpha \in \mathbb{K}$  we have that  $f + g$  and  $\alpha \cdot f$  is defined by  $(f + g)(x) = f(x) + g(x)$  and  $(\alpha \cdot f)(x) = \alpha \cdot f(x)$ .

**Proof.** Using 19.261 we have that  $\langle \mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{K}], +, \cdot \rangle$  is a vector space, so we only have to prove that  $\|\cdot\|_p$  is a pseudo norm. First by definition [see 19.262, 19.267] we have that

$$\|\cdot\|_p: \mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{K}] \rightarrow [0, \infty[ \quad (19.506)$$

Further using the above Minkowski inequality [see 19.270] we have

$$\forall f, g \in \mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{K}] \text{ we have } \|f + g\|_p \leq \|f\|_p + \|g\|_p \quad (19.507)$$

Let  $f \in \mathcal{L}[X, \mathcal{A}, \mu, \mathbb{K}]$  and  $\alpha \in \mathbb{K}$  then for  $p$  we must consider the following cases:

$p \in [1, \infty[$ . Then

$$\begin{aligned} \|\alpha \cdot f\|_p &= \left( \int |\alpha \cdot f|^p d\mu \right)^{\frac{1}{p}} \\ &= \left( \int (|\alpha| \cdot |f|)^p d\mu \right)^{\frac{1}{p}} \\ &\stackrel{18.39}{=} \left( \int |\alpha|^p \cdot |f|^p d\mu \right)^{\frac{1}{p}} \\ &\stackrel{19.225}{=} \left( |\alpha|^p \cdot \int |f|^p d\mu \right)^{\frac{1}{p}} \\ &\stackrel{18.39}{=} (|\alpha|^p)^{\frac{1}{p}} \cdot \left( \int |f|^p d\mu \right)^{\frac{1}{p}} \\ &= (|\alpha|^p)^{\frac{1}{p}} \cdot \|f\|_p \\ &\stackrel{18.39}{=} |\alpha| \cdot \|f\|_p \end{aligned}$$

$p = \infty$ . For  $\alpha$  we have the following possibilities to consider:

**$\alpha = 0$ .** Then given  $M \in [0, \infty[$ , if  $x \in \{x \in X \mid |\alpha \cdot f|(x) > M\}$  we have  $0 = |\alpha \cdot f(x)| > 0$  a contradiction, so  $\{x \in X \mid |\alpha \cdot f|(x) > M\} = \emptyset$  and thus by 19.265 is locally  $\mu$ -null. So  $\|\alpha \cdot f\|_\infty = \inf(\{M \in [0, \infty[ \mid \{x \in X \mid |\alpha \cdot f|(x) > M\} \text{ is locally } \mu\text{-null}\}) = \inf([0, \infty[) = 0$ . Hence as  $\alpha = 0$

$$\|\alpha \cdot f\|_\infty = 0 = |\alpha| \cdot \|f\|_\infty$$

**$\alpha \neq 0$ .** Then  $|\alpha| \neq 0$  then given  $M \in [0, \infty[$  we have

$$\begin{aligned} x \in \{x \in X \mid |(\alpha \cdot f)(x)| > M\} &\Leftrightarrow |\alpha \cdot f|(x) > M \\ &\Leftrightarrow |\alpha| \cdot |f(x)| > M \\ &\Leftrightarrow |f(x)| > \frac{M}{|\alpha|} \\ &\Leftrightarrow x \in \left\{ x \in X \mid |f(x)| > \frac{M}{|\alpha|} \right\} \end{aligned}$$

proving that

$$\{x \in X \mid |(\alpha \cdot f)(x)| > M\} = \left\{ x \in X \mid |f(x)| > \frac{M}{|\alpha|} \right\} \quad (19.508)$$

Define

$$\begin{aligned} A &= \{M \in [0, \infty[ \mid \{x \in X \mid |\alpha \cdot f|(x) > M\} \text{ is locally } \mu\text{-null}\} \\ B &= \{M \in [0, \infty[ \mid \{x \in X \mid |f(x)| > \frac{M}{|\alpha|}\} \text{ is locally } \mu\text{-null}\} \end{aligned}$$

If  $M \in A$  then  $\{x \in X \mid |\alpha \cdot f(x)| > M\}$  is locally  $\mu$ -null so using 19.508 we have that  $\left\{x \in X \mid |f(x)| > \frac{M}{|\alpha|}\right\}$  is locally  $\mu$ -null hence  $M' = \frac{M}{|\alpha|} \in B$  or  $M \in \{|\alpha| \cdot M' \mid M' \in B\}$  giving

$$A \subseteq \{|\alpha| \cdot M' \mid M' \in B\} \quad (19.509)$$

On the other side, if  $M \in \{|\alpha| \cdot M' \mid M' \in B\}$  there exists a  $M' \in B$  such that  $M = |\alpha| \cdot M'$  or  $\frac{M}{|\alpha|} = M' \in B$  hence  $\left\{x \in X \mid |f(x)| > \frac{M}{|\alpha|}\right\}$  is locally  $\mu$ -null. Using 19.508 we have then that  $\{x \in X \mid |(\alpha \cdot f)(x)| > M\}$  is locally  $\mu$ -null or  $M \in A$ , proving that  $\{|\alpha| \cdot M' \mid M' \in B\} \subseteq A$ . Using 19.509 we have then that

$$A = \{|\alpha| \cdot M' \mid M' \in B\} \quad (19.510)$$

As  $\|\alpha \cdot f\|_\infty \stackrel{\text{def}}{=} \inf(A)$  we have by 19.510  $\forall M \in B$  that  $\|\alpha \cdot f\|_\infty \leq |\alpha| \cdot M$  so that  $\frac{\|\alpha \cdot f\|_\infty}{|\alpha|} \leq M$  hence  $\frac{\|\alpha \cdot f\|_\infty}{|\alpha|} \leq \inf(B) = \|f\|_\infty$  or

$$\|\alpha \cdot f\|_\infty \leq |\alpha| \cdot \|f\|_\infty \quad (19.511)$$

Further if  $M \in \{|\alpha| \cdot M' \mid M' \in B\}$  then  $M = |\alpha| \cdot M'$  where  $M' \in B$  so that  $\frac{M}{|\alpha|} \in B$  hence  $\|f\|_\infty = \inf(B) \leq \frac{M}{|\alpha|}$  or  $|\alpha| \cdot \|f\|_\infty \leq M$  so that

$$|\alpha| \cdot \|f\|_\infty \leq \inf(\{|\alpha| \cdot M' \mid M' \in B\}) \stackrel{19.510}{=} \inf(A) = \|\alpha \cdot f\|_\infty$$

which combined with 19.511 proves that

$$\|\alpha \cdot f\|_p = \|\alpha \cdot f\|_\infty = |\alpha| \cdot \|f\|_\infty = |\alpha| \cdot \|f\|_p$$

So we have for  $\forall \alpha \in \mathbb{K}$  that

$$\|\alpha \cdot f\|_p = |\alpha| \cdot \|f\|_p$$

Hence  $\forall p \in [1, \infty[$  we have that

$$\|\alpha \cdot f\|_p = |\alpha| \cdot \|f\|_p \quad (19.512)$$

Using 19.506, 19.507 and 19.512 it follows that

$$\|\cdot\|_p \text{ is a pseudo norm} \quad \square$$

It would be nice if  $\langle \mathcal{L}[X, \mathcal{A}, \mu], \|\cdot\|_p \rangle$  is a normed space. However if  $A \in \mathcal{A}$  is such that  $\mu(A) = 0$  and  $A \neq \emptyset$  then if  $x \in X$  we have that  $\mathcal{X}_A(x) = 1 \neq 0$  and  $\|\mathcal{X}_A\|_1 = \int |\mathcal{X}_A| d\mu = \int^+ |\mathcal{X}_A| d\mu = \int^S |\mathcal{X}_A| d\mu = \mu(A) = 0$ . So we can not assume that  $\|f\|_p$  means  $f = 0$ . To solve this problem we define the following.

**Definition 19.272.** Let  $\langle X, \mathcal{A}, \mu \rangle$  be a measure space,  $p \in [1, \infty[$  then  $\mathcal{N}_p[X, \mathcal{A}, \mu, \mathbb{K}]$  is defined by

$$\mathcal{N}_p[X, \mathcal{A}, \mu, \mathbb{K}] = \{f \in \mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{K}] \mid \|f\|_p\}$$

We have the following characterization of the  $\mathcal{N}_p[X, \mathcal{A}, \mu, \mathbb{K}]$ .

**Lemma 19.273.** Let  $\langle X, \mathcal{A}, \mu \rangle$  then we have

1. If  $p \in [1, \infty[$  then  $f \in \mathcal{N}_p[X, \mathcal{A}, \mu, \mathbb{K}]$  if and only if  $f \in \mathcal{M}[X, \mathcal{A}, \mathbb{K}]$  with  $f = 0$   $\mu$ -a.e.
2.  $f \in \mathcal{N}_\infty[X, \mathcal{A}, \mu, \mathbb{K}]$  if and only if  $f \in \mathcal{L}^\infty[X, \mathcal{A}, \mu, \mathbb{K}]$   $\{x \mid f(x) \neq 0\}$  is locally  $\mu$ -null

**Proof.** First we have for  $p \in [1, \infty[$  that

$$\begin{aligned} x \in \{x \in X \mid f(x) \neq 0\} &\Leftrightarrow f(x) \neq 0 \\ &\stackrel{\|\cdot\| \text{ is a norm}}{\Leftrightarrow} |f(x)| \neq 0 \\ &\stackrel{18.39}{\Leftrightarrow} |f(x)|^p \neq 0 \\ &\Leftrightarrow x \in \{x \in X \mid |f(x)|^p \neq 0\} \end{aligned}$$

proving that

$$\forall p \in [1, \infty] \{x \in X | f(x) \neq 0\} = \{x \in X | |f(x)|^p \neq 0\} \quad (19.513)$$

1.

$\Rightarrow$ . As  $f \in \mathcal{N}_p[X, \mathcal{A}, \mu, \mathbb{K}]$  we have  $\|f\|_p = 0$ , then  $(\int |f|^p d\mu)^{\frac{1}{p}} = 0$ , so by 18.39  $\int |f|^p d\mu = 0$ . As  $\forall x \in X$  we have  $|f(x)|^p \in [0, \infty] \Rightarrow |f(x)|^p = |f(x)|^p$  we have  $\int |f|^p d\mu = 0$ . Using 19.232 we it follows that  $|f|^p = 0$   $\mu$ -a.e., hence there exists a  $N \in \mathcal{A}$  with  $\mu(N) = 0$  such that  $\{x \in X | |f(x)|^p \neq 0\} \subseteq N$  or using 19.513 that  $\{x \in X | f(x) \neq 0\} \subseteq N$  proving that

$$f = 0 \text{ } \mu\text{-a.e.}$$

Further as  $\mathcal{N}_p[X, \mathcal{A}, \mu, \mathbb{K}] \subseteq \mathcal{L}[X, \mathcal{A}, \mu, \mathbb{K}]$  we have also that

$$f \in \mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{K}]$$

$\Leftarrow$ . If  $f = 0$   $\mu$ -a.e then there exists a  $N \in \mathcal{A}$  with  $\mu(N) = 0$  such that  $\{x \in X | f(x) \neq 0\} \subseteq N$ . Using 19.513 it follows that  $\{x \in X | |f(x)|^p \neq 0\} \subseteq N$  proving that  $|f|^p = 0$   $\mu$ -a.e., so we can use 19.222 to get that  $|f|^p$  is  $\mu$ -integrable and  $\int |f|^p d\mu = 0$ . As  $f \in \mathcal{M}[X, \mathcal{A}, \mathbb{K}]$  we have that

$$f \in \mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{K}]$$

Further  $(\int |f|^p d\mu)^{\frac{1}{p}} = (0)^{\frac{1}{p}} = 0$  which by definition gives

$$\|f\|_p = 0$$

hence

$$f \in \mathcal{N}_p[X, \mathcal{A}, \mu, \mathbb{K}]$$

2.

As by 19.268 we have that  $\{x \in X | |f(x)| > \|f\|_\infty\}$  is locally  $\mu$ -null we have, as  $\|f\|_\infty = 0$ , that  $\{x \in X | |f(x)| > 0\}$  is locally  $\mu$ -null or as  $\{x \in X | |f(x)| \neq 0\} = \{x \in X | |f(x)| > 0\}$

$$\{x \in X | |f(x)| \neq 0\} \text{ is locally } \mu\text{-null}$$

Further as  $\mathcal{N}_\infty[X, \mathcal{A}, \mu, \mathbb{K}] \subseteq \mathcal{L}^\infty[X, \mathcal{A}, \mu, \mathbb{K}]$  then

$$f \in \mathcal{L}^\infty[X, \mathcal{A}, \mu, \mathbb{K}]$$

$\Leftarrow$ . As  $\{x \in X | |f(x)| \neq 0\}$  is locally  $\mu$ -null we have, as  $\{x \in X | |f(x)| \neq 0\} = \{x \in X | |f(x)| > 0\}$ , that  $\{x \in X | |f(x)| > 0\}$  is locally  $\mu$ -null. Hence  $0 \in \{M \in [0, \infty] | \{x \in X | |f(x)| > M\} \text{ is locally } \mu\text{-null}\}$ , proving that  $\inf(\{M \in [0, \infty] | \{x \in X | |f(x)| > M\}\}) = 0$ , hence

$$\|f\|_\infty = 0$$

As  $f \in \mathcal{L}^\infty[X, \mathcal{A}, \mu, \mathbb{K}]$  we have by the above that

$$f \in \mathcal{N}^\infty[X, \mathcal{A}, \mu, \mathbb{K}]$$

□

The following will be essential to develop a normed space out of  $\mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{K}]$ .

**Lemma 19.274.** *Let  $p \in [1, \infty]$  and  $\langle X, \mathcal{A}, \mu \rangle$  a measure space then  $\langle \mathcal{N}_p[X, \mathcal{A}, \mu, \mathbb{K}], +, \cdot \rangle$  is a sub space from  $\langle \mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{K}], +, \cdot \rangle$  [see 19.261 for the definition of  $+$  and  $\cdot$ ].*

**Proof.**

1. If  $f, g \in \mathcal{N}_p[X, \mathcal{A}, \mu, \mathbb{K}]$  then  $\|f\|_p = \|g\|_p$ . So  $\|f + g\|_\infty \leq \|f\|_p + \|g\|_p = 0$  proving that  $\|f + g\| = 0$ . Hence

$$f + g \in \mathcal{N}_p[X, \mathcal{A}, \mu, \mathbb{K}]$$

2. If  $f \in \mathcal{N}_p[X, \mathcal{A}, \mu, \mathbb{K}]$  and  $\alpha \in \mathbb{K}$  then  $\|f\|_p = 0$  so that  $\|\alpha \cdot f\|_p = |\alpha| \cdot \|f\|_p = 0$  proving that

$$\alpha \cdot f \in \mathcal{N}_p[X, \mathcal{A}, \mu, \mathbb{K}] \quad \square$$

We can now use the concept of a factor space [see 10.193] to define  $L^p[X, \mathcal{A}, \mu, \mathbb{K}]$

**Definition 19.275.** Let  $p \in [0, \infty]$  then  $\langle L[X, \mathcal{A}, \mu, \mathbb{K}], +, \cdot \rangle$  is defined as the **factor space** defined by  $\langle \mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{K}], +, \cdot \rangle$  and  $\langle \mathcal{N}_p[X, \mathcal{A}, \mu, \mathbb{K}], +, \cdot \rangle$  [see 10.193]. To recap the definition of a **factor space** we have [taking care of the context for the meaning of  $+$ ,  $\cdot$  and 0

1. A equivalence relation  $\sim_p \in \mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{K}]$ , defined by

$$f \sim_p g \text{ if and only if } f - g \in \mathcal{N}_p[X, \mathcal{A}, \mu, \mathbb{K}]$$

2. Given  $f \in \mathcal{L}[X, \mathcal{A}, \mu, \mathbb{K}]$  we have that  $[f]_p$  is defined by

$$\begin{aligned} [f]_p &= \{g \in \mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{K}] \mid f \sim_p g\} \\ &= \{g \in \mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{K}] \mid f - g \in \mathcal{N}_p[X, \mathcal{A}, \mu, \mathbb{K}]\} \end{aligned}$$

3.  $L^p[X, \mathcal{A}, \mu, \mathbb{K}]$  is defined by

$$L^p[X, \mathcal{A}, \mu, \mathbb{K}] = \{[x]_p \mid x \in \mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{K}]\}$$

4. The sum operator  $+: L^p[X, \mathcal{A}, \mu, \mathbb{K}] \times L^p[X, \mathcal{A}, \mu, \mathbb{K}] \rightarrow L^p[X, \mathcal{A}, \mu, \mathbb{K}]$  is defined by

$$[f]_p + [g]_p = [f + g]_p$$

5. The scalar product operatior  $\cdot: \mathbb{K} \times L^p[X, \mathcal{A}, \mu, \mathbb{K}] \rightarrow L^p[X, \mathcal{A}, \mu, \mathbb{K}]$  is defined by

$$\alpha \cdot [f]_p = [\alpha \cdot f]_p$$

6. The neutral element 0 is defined by

$$0 = [0]_p$$

It remains to define a norm on  $L^p[X, \mathcal{A}, \mu, \mathbb{K}]$  to make it a normed space

**Theorem 19.276.** Let  $p \in [0, \infty[$  and  $\langle X, \mathcal{A}, \mu, \mathbb{K} \rangle$  then  $\|\|_{L_p}: L^p[X, \mathcal{A}, \mu, \mathbb{K}] \rightarrow [0, \infty[$  defined by

$$\|f\|_{L_p} = \|f'\|_p \text{ where } f = [f']_p$$

is well defined and is a norm on  $L^p[X, \mathcal{A}, \mu, \mathbb{K}]$  making  $\langle L^p[X, \mathcal{A}, \mu, \mathbb{K}], \|\|_{L_p} \rangle$  a normed space

**Proof.** Furst we must proof that  $\|\|_{L_p}$  is well defined. So take  $x \in L^p[X, \mathcal{A}, \mu, \mathbb{K}]$  them there exists a  $f \in \mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{K}]$  such that  $x = [f]_p$  and assume that there si another  $g \in \mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{K}]$  such that  $x = [g]_p$ . Then  $f - g \in \mathcal{N}_p[X, \mathcal{A}, \mu, \mathbb{K}]$  so that  $\|g - f\|_p = \|f - g\|_p = 0$ . As  $\|f\|_p \leq \|f - g + g\|_p \leq \|f - g\|_p + \|g\|_p = \|g\|_p$  and  $\|g\|_p = \|g - f + f\|_p \leq \|g - f\|_p + \|f\|_p = \|f\|_p$  we have  $\|f\|_p \leq \|g\|_p \leq \|f\|_p$  proving that  $\|f\|_p = \|g\|_p$ . So

$$\|\|_{L_p} \text{ is well defined}$$

Let  $f, g \in L^p[X, \mathcal{A}, \mu, \mathbb{K}]$  then  $f = [f']_p$ ,  $g = [g']_p$  and

$$\begin{aligned} \|f + g\|_p &= \|[f']_p + [g']_p\|_{L_p} \\ &= \|[f' + g']_p\|_{L_p} \\ &= \|f' + g'\|_p \\ &\leq \|f'\|_p + \|g'\|_p \\ &= \|[f']_p\|_{L_p} + \|[g']_p\|_{L_p} \\ &= \|f\|_{L_p} + \|g\|_{L_p} \end{aligned}$$

proving that

$$\forall f, g \in L^p[X, \mathcal{A}, \mu, \mathbb{K}] \text{ we have } \|f + g\|_{L_p} \leq \|f\|_{L_p} + \|g\|_{L_p}$$

If  $\alpha \in \mathbb{K}$  and  $f \in L^p[X, \mathcal{A}, \mu, \mathbb{K}]$  then  $f = [f']_p$

$$\begin{aligned}\|\alpha \cdot f\|_{L^p} &= \|\alpha \cdot [f']_p\|_{L^p} \\ &= \|[\alpha \cdot f']_p\|_{L^p} \\ &= \|\alpha \cdot f'\|_p \\ &= |\alpha| \cdot \|f'\|_p \\ &= |\alpha| \cdot \|f\|_{L^p}\end{aligned}$$

proving that

$$\forall \alpha \in \mathbb{K}, \forall f \in L^p[X, \mathcal{A}, \mu, \mathbb{K}] \text{ we have } \|\alpha \cdot f\|_p = |\alpha| \cdot \|f\|_p$$

Further if  $f \in L^p[X, \mathcal{A}, \mu, \mathbb{K}]$  with  $\|f\|_{L^p} = 0$  then there exists a  $f' \in \mathcal{L}^p[X, \mathcal{A}, \mu, \mathbb{K}]$  such that  $f = [f']_p$  and  $0 = \|f\|_{L^p} = \|[f']_p\|_{L^p} = \|f'\|_p$  so that  $f' \in \mathcal{N}_p[X, \mathcal{A}, \mu, \mathbb{K}]$  proving that  $f = [0]_p$  the neutral element in  $L^p[X, \mathcal{A}, \mu, \mathbb{K}]$ . So we have

$\forall f \in L^p[X, \mathcal{A}, \mu, \mathbb{K}]$  with  $\|f\|_{L^p} = 0$  we have  $f = [0]_p = 0$  [the neutral element in  $L^p[X, \mathcal{A}, \mu, \mathbb{K}]$ ]  $\square$

# Index

$\cup$	8	$\mathbb{Z}$	153
$\cap$	8	$\bullet_n$	445
$\langle \mathbb{Q}, + \rangle$	168	$\mathcal{B}(X, Y)$	467
$\sqrt{x}$	214	$\mathcal{B}_{\mathbb{R}}$	684
$\{n, \dots, m\}$	237	$\mathcal{B}_{\infty}$	684
$\langle V, \oplus, \odot \rangle$	275	$\mathcal{B}_{-\infty}$	684
$\{1, \dots, i-1, i+1, \dots, n\}$	306	$\mathcal{B}_{\bar{\mathbb{R}}}$	684
$\hat{\Delta}$	322	$\mathcal{B}$	392
$\hat{\Delta}_L$	328	$\mathcal{C}(X, Y)$	421
$\bar{B}_d(x, \varepsilon)$	398	$\mathcal{C}$	390
$\ \cdot\ $	402	$d(\mathcal{A})$	828
$\ L\ $	431	$\mathcal{D}^n$	774
$\langle x, y \rangle$	22	$\mathcal{D}_m^n$	774
$\binom{n}{k}$	520	$\mathcal{F}(X, F)$	290
$(i \rightarrow x)$	570	$\mathcal{G}L(X)$	636
$]_{-\infty}, a[i$	777	$\mathcal{M}(L)$	338
$\overline{\mathcal{M}}[X]$	833	$\mathcal{M}(L; \{e_i\}_{i \in \{1, \dots, n\}}; \{f_i\}_{i \in \{1, \dots, m\}})$	338
$\overline{\mathcal{M}^+[X, \mathcal{A}]}$	871	$\mathcal{M}^+[X, A]$	871
$\langle x, y, z \rangle$	23	$\mathcal{M}(n \times m, F)$	335
$\{G_i\}_{i \in I}$	24	$\mathcal{P}'(A)$	21
$\{n, \dots\}$	115	$\mathcal{P}_n(L)$	442
$\{a_0, \dots, a_n\}$	124	$\mathcal{P}(A)$	21
$\{a_1, \dots, a_n\}$	124	$\mathcal{P}_1 \boxplus \mathcal{P}_2$	502
$\ \cdot\ $	162	$\mathcal{R}^n$	771
$\frac{n}{k}$	167, 167	$\mathcal{S}(f; \mathbb{P})$	504
$\langle \mathbb{R}, +, \cdot \rangle$ is a field	188	$\mathcal{T}_{\bar{\mathbb{R}}}$	685
$\langle \mathbb{N}_0, \leq \rangle$ is conditional complete	111	$\mathcal{T}_{\text{box}}$	395
$\liminf_{n \rightarrow \infty} x_n$	475	$\cong$	68
$\limsup_{n \rightarrow \infty} x_n$	475	$i_n(i; j)$	287
$\liminf_{i \rightarrow \infty} x_i$	691	$\Delta_L$	326
$\limsup_{i \rightarrow \infty} x_i$	691	$\Delta$	318
$\#(A)$	128	$\emptyset$	8
$-\mathbb{N}_0_{\mathbb{Z}}$	158	$\mathfrak{U}(x)$	454
$1_A$	39	$\Gamma(n_1, \dots, n_k)$	380
$\alpha \cdot A$	279	$\in$	7
$\approx$	39	$\leqslant$	201
$\mathbb{C}_{\mathbb{N}_0}$	224	$<$	201
$\mathbb{C}$	217	$\mu(\mathcal{P})$	504
$\mathbb{N}_0_{\mathbb{Q}}$	174	$\mu^*$	821
$\mathbb{N}_0_{\mathbb{R}}$	207	$\mu_*$	821
$\mathbb{N}_0_{\mathbb{Z}}$	157	$\notin$	7
$\mathbb{P}$	504	$\oplus_{i \in \{1, \dots, n\}} V_i$	368
$\mathbb{Q}_{\mathbb{R}}$	180	$\partial_i f(x)$	575
$\mathbb{Q}$	167	$\pi_i$	55, 143
$\mathbb{R}_+$	183	$\pi$	747
$\mathbb{R}_-$	183	$\preccurlyeq$	284
$\mathbb{R}_{\mathbb{C}}$	219	$\sigma^{[i]}$	259
$\mathbb{R}$ is conditional complete	204	$\sigma[\mathcal{A}]$	759
$\mathbb{R}$	179	$\sigma f$	264
$\mathbb{R}_{\mathbb{C}+}, \mathbb{R}_{\mathbb{C}-}$	221	$\subset$	7
$\mathbb{Z}_{\mathbb{Q}}$	173	$\subseteq$	7, 59
$\mathbb{Z}_{\mathbb{R}}$	206	$\varphi^n$	797
		$A^\circ$	390
		$A_\mu$	817
		$A^c$	9

$A + B$	279	column rank of a matrix	337
$A - B$	279	common divisor	164
$A/R$	61	compact subsets	448
$A \simeq B$	68	compact topological space	447
$A \times B$	15	comparability	67
$A = B$	7	complete measures	817
abelian group	93	complete metric space	467
abelian semi-group	93	completion of a measure space	818
absolute convergence domain	723	complex inner spaces	410
absolute value	162, 211	complex norm	222
absorbing element of integers	155	complex spectral theorem	552
absorption laws	10	concave function	606
accumulation point	392	conditional complete classes	72
accumulation point of a filterbase	455	conjugate of a complex number	218
addition of natural numbers	107	conjugate transpose	419
adjoint( $L$ )	332	connected space	509
adjoint( $M$ )	351	connected subset	510
adjoint of a linear function	417	continuous function at a point	420
adjoint of a matrix	351	continuous functions	421
algebra	754	continuous mapping at a point	420
$\sigma$ -algebra	756	continuous mappings	421
almost everywhere	853	contraction	495
anti-symmetric	59	contraction of tensors	388
Archimedean property of $\mathbb{Z}$	162	convergence domain	723
Archimedean property of $\mathbb{Q}$	175	convergence in a metric space	463
Archimedean property of $\mathbb{N}_0$	114	convergence radius	724
axiom of choice	83	convergency in a topological space	463
axiom of class construction	8	convergent filterbase	455
axiom of extent	7	convex function	606
axiom of infinity	21	convex sets	615
axiom of pairing	14	$\cos(x)$	742
axiom of powersets	21	countable compact class	452
axiom of replacement	52	countable set	124
axiom of subsets	20	$\mathbb{D}$	828
axiom of unions	20	decreasing sequence	463, 689
$B_d(x, \varepsilon)$	398	Dedekind's Cut	179
$B^A$	52	DeMorgan's Laws	10
Baire Category Theorem	498	dense subset	397
Baire space	397	density theorem	210, 232
Banach fixed point theorem	496	derivative	567
Banach space	467	derived set	392
basis of a topology	392	$\det(L)$	326
basis of a vector space	284	$\det(M)$	343
binomial constant	520	determinant	326
Borel algebra	759	determinant function	318
Borel function	831	determinant of a matrix	343
bounded function	467	$D^n f$	582
bounded set	401	$Df(x)$	563
$c_{s,t}$	388	$\text{diam}(A)$	401
$C_{X,c}$	421	diameter of a set	401
$C^n$	582	$\infty$ -differentiability	582
$C^\infty$	582	differential of a constant function	566
$C_B$	37	differential of a linear function	567
$\mathcal{B}[\mathbf{X}, \mathcal{T}]$	759	direct internal sum	368
canonical decomposition of a function	62	direct internal sum of subspaces	297
Carathéodory theorem	788	distance function	425
cartesian power of sets	274	divergent limit	516
cartesian product of a family of sets	140	divides	164
Cauchy sequence	464	division algorithm	163
chain	67, 67	domain	18
chain rule	568	dual space	295
characteristics function	37	Dyadic cubes	774
closed ball	398	Dynkin class	827
closed sets	390	$e_A$	39
$\text{cols}(M)$	337	eigenvalue	541
column matrix	334	eigenvector	541

elementair transformations . . . . .	358	increasing function . . . . .	68
enumerable set . . . . .	124	increasing sequence . . . . .	463
equipotent . . . . .	117	increasing sequence . . . . .	689
equivalence relation . . . . .	60	induced order relation . . . . .	66
equivalent metrics . . . . .	400	induced topology . . . . .	427
equivalent norms . . . . .	409	$\inf(A)$ . . . . .	70
Euler's number . . . . .	734	$\inf_{\mathbb{R}}$ . . . . .	668
even numbers . . . . .	164	$\inf$ . . . . .	668
$\exp(z)$ . . . . .	728	infinite countable set . . . . .	124
extented reals . . . . .	665	infinite sets . . . . .	124
extreme value theorem . . . . .	449	initial segment . . . . .	67
extremum . . . . .	605	inner measure induced . . . . .	821
$f(C)$ . . . . .	35	inner regular sets . . . . .	821
$f^{-1}(D)$ . . . . .	35	inner set . . . . .	390
$f_{X \rightarrow A}$ . . . . .	876	integrable functions on a measure space . . . . .	876
$\int_a^b f$ . . . . .	505	integral domain . . . . .	96
$\int_a^b f$ . . . . .	506	intermediate value theorem . . . . .	511
$f'$ . . . . .	567	inverse function theorem . . . . .	651
$f^{(n)}$ . . . . .	582	inverse induced topology . . . . .	427
$f'_+(x)$ . . . . .	599	irrational numbers . . . . .	206
$f'_-(x)$ . . . . .	599	isometry . . . . .	400, 409
$f'_+(x)$ . . . . .	599	isomorphism . . . . .	68, 293
$f'_-(x)$ . . . . .	599	iteration . . . . .	104
$f^+$ . . . . .	871	i-the coordinate . . . . .	285
$\int_A f d\mu$ . . . . .	876, 876, 876	Jacobian . . . . .	578
$f_n \circ \dots \circ f_1$ . . . . .	116	$L(X, Y)$ . . . . .	433
faithful action . . . . .	95	$L(X_1, \dots, X_n; Y)$ . . . . .	438
family of classes . . . . .	24	$L(X^n; Y)$ . . . . .	439
family of elements . . . . .	58	$L^n(X_1, \dots, X_n; Y)$ . . . . .	439
family of elements with finite support . . . . .	245	$L^n(X; Y)$ . . . . .	439
field . . . . .	97	$L_{a,b}$ . . . . .	616
field with characterization zero . . . . .	106	$L_{(\dots a)}$ . . . . .	308
filterbase on a set . . . . .	456	Lagrange's theorem . . . . .	606
finite sets . . . . .	124	least element . . . . .	70
finite sum . . . . .	237	Lebesgue measurable set . . . . .	793
first countable topological space . . . . .	446	Lebesgue measurable space . . . . .	793
fully ordered class . . . . .	67	Lebesgue measure . . . . .	793
function . . . . .	36	left action . . . . .	95
function graph . . . . .	20	limit of a partial function . . . . .	398
fundamental system of neighborhoods . . . . .	446	limit point . . . . .	392
Fundamental Theorem of Calculus (Classical version) . . . . .	613	line segment . . . . .	616
Fundamental Theorem of Calculus . . . . .	614	linear combination . . . . .	280
gcd . . . . .	164	linear dependent family . . . . .	282
general linear group . . . . .	295	linear dependent set . . . . .	282
generalized deMorgan's Law . . . . .	30	linear independence . . . . .	282
generalized interval . . . . .	74	linear independent family . . . . .	282
generalized sum . . . . .	246, 246	linear independent set . . . . .	282
generating basis of a topology . . . . .	393	linear isomorphism . . . . .	293
$GL(X)$ . . . . .	295	linear mapping . . . . .	291
greatest element . . . . .	70	linear operator . . . . .	537
group . . . . .	93	linear ordered class . . . . .	67
group homeomorphism . . . . .	94	linear span . . . . .	280
group isomorphism . . . . .	94	local weak maximum . . . . .	605
$H^n$ . . . . .	637	local weak minimum . . . . .	605
Hausdorff space . . . . .	445	log . . . . .	735
Hausdorff's maximal principle . . . . .	87	lower bound . . . . .	70
Heine-Borel theorem . . . . .	460	lower half open space . . . . .	777
$\text{Hom}(X_1, \dots, X_n; Y)$ . . . . .	308	Lower half spaces . . . . .	778
$\text{Hom}(X^n; Y)$ . . . . .	312	$M^T$ . . . . .	336, 419
$\text{Hom}(X, Y)$ . . . . .	294	$[i \boxplus j]M$ . . . . .	346
homeomorphism . . . . .	426	$[<m]M$ . . . . .	350
$\text{Hom}(X . . . . .$	537	$[>m]M$ . . . . .	350
immediate successor . . . . .	80	$[i \oplus j]M$ . . . . .	351
inclusion function . . . . .	39	$M^H$ . . . . .	419
		$M(X, V)$ . . . . .	277
		mapping . . . . .	36

$\varepsilon$ -mapping . . . . .	564	rank of a matrix . . . . .	337
mathematical induction . . . . .	100	rational cuts . . . . .	179
matrix . . . . .	334	real product spaces . . . . .	410
maximal element . . . . .	70	reciprocal cut . . . . .	186
measurable function . . . . .	831	recursion . . . . .	101
$\mathcal{A}$ -measurable function . . . . .	832	recursion on $S_{n+1}$ - step form . . . . .	119
metric topology . . . . .	399	recursion on $\mathbb{N}_0$ step form . . . . .	106
minimal element . . . . .	70	reflexitivity . . . . .	59
multilinear mapping . . . . .	307	regular linear mapping . . . . .	292
multiplication of natural numbers . . . . .	108	regular space . . . . .	446
$n!$ . . . . .	120	reindexing of a family . . . . .	51
Napierian logarithm . . . . .	735	relation . . . . .	59
natural numbers . . . . .	99	restricted partial function . . . . .	39
negative cut . . . . .	180	restriction of a equivalence relation . . . . .	61
negative whole numbers . . . . .	158	Rieman Integral . . . . .	505
$\mu$ -negligible . . . . .	817	Riemannian sum . . . . .	504
neighborhood . . . . .	446	Riemann Integral . . . . .	506
neutral element of natural elements . . . . .	107	Riesz Representation theorem . . . . .	417
norm of a continuous linear function . . . . .	431	right action . . . . .	95
norm of a partition . . . . .	504	ring . . . . .	95
normal operator . . . . .	550	ring homomorphism . . . . .	97
normal space . . . . .	446	ring isomorphism . . . . .	97
normed space . . . . .	402	Rolle's theorem . . . . .	606
normed topology . . . . .	402	row matrix . . . . .	334
odd numbers . . . . .	164	row rank of a matrix . . . . .	337
one to one correspondence . . . . .	39	rows( $M$ ) . . . . .	337
open ball . . . . .	398	$S_n$ . . . . .	119
open functions . . . . .	422	$S_a$ . . . . .	67
open mapping theorem . . . . .	499	scalair product of matrices . . . . .	335
open neighborhood . . . . .	446	second countable topological space . . . . .	447
open set . . . . .	389	section . . . . .	80
order relation . . . . .	64	self-adjoint . . . . .	543
order relation in $\mathbb{Z}$ . . . . .	158	semi-group . . . . .	93
order relation on rational numbers . . . . .	172	sequence . . . . .	139
orthogonal complement . . . . .	556	sequential compact space . . . . .	450
orthogonality . . . . .	414	sign( $\sigma$ ) . . . . .	268
orthonormality . . . . .	414	sin( $x$ ) . . . . .	742
outer measure induced . . . . .	818	skew symmetry . . . . .	313
outer regular sets . . . . .	823	smooth functions . . . . .	582
$P_n$ . . . . .	254	stitching topologies . . . . .	394
pair of classes . . . . .	22	strictly increasing function . . . . .	68
partial function . . . . .	35	sub field . . . . .	98
partial mapping . . . . .	35	sub-group . . . . .	94
partial ordered set . . . . .	65	sub-ring . . . . .	96
partition of $[a, b]$ . . . . .	501	sub-semi-group . . . . .	93
p-contravariant,q-covariant tensor . . . . .	387	subspace of a vector space . . . . .	276
peano axioms . . . . .	100	subspace topology . . . . .	390
permutation . . . . .	241	successor . . . . .	21
permutation group . . . . .	255	successor set . . . . .	21
poset . . . . .	65	sum of matrices . . . . .	335
positive whole numbers . . . . .	157	sum of subspaces . . . . .	367
pre-order . . . . .	64	sup( $A$ ) . . . . .	70
pre-ordered class . . . . .	64	sup $_{\mathbb{R}}$ . . . . .	668
pretuple . . . . .	53	sup . . . . .	668
product of matrices . . . . .	335	support( $\{x_i\}_{i \in I}$ ) . . . . .	245
product tuple . . . . .	141	support of a family of elements . . . . .	245
projection function . . . . .	55	symmetric . . . . .	59
pseudo metric space . . . . .	398	tagged partition . . . . .	504
pseudo normed topology . . . . .	402	tensor product . . . . .	371
psuedo normed space . . . . .	402	tensor product of vectors . . . . .	373
quotient vector space . . . . .	296	the real numbers . . . . .	179
$R_{\{a_i\}_{i \in \mathbb{N}_0}}$ . . . . .	724	toplinear isomorphism . . . . .	447, 636
range . . . . .	18	topological space . . . . .	389
rank( $L$ ) . . . . .	302	topological vector space . . . . .	447
rank( $M$ ) . . . . .	337	topology . . . . .	389
rank of a linear mapping . . . . .	302	transfinite induction . . . . .	80

transitive . . . . .	59	$\sum_{i \in I}^{\oplus} V_i$ . . . . .	368
transitive action . . . . .	95	$V^*$ . . . . .	295
transpose of a matrix . . . . .	336, 419	$V_1 \oplus \dots \oplus V_n$ . . . . .	368, 368
transposition . . . . .	241	$v_1 \otimes \dots \otimes v_n$ . . . . .	373
tripple of classes . . . . .	23	$V_1 \otimes \dots \otimes V_n$ . . . . .	372, 373
Tychonoff's theorem . . . . .	460	vector space . . . . .	275
$U_x$ . . . . .	561	well-ordered class . . . . .	79
$U^{\perp}$ . . . . .	556	well-ordering theorem . . . . .	88
ultra filter . . . . .	457	$x_{i \rightarrow a}$ . . . . .	307
unary matrix . . . . .	419	$x_{\dots a}$ . . . . .	308
uniform continuous . . . . .	424	$\lim_{i \rightarrow \infty} x_i$ . . . . .	464
uniform convergence . . . . .	468	$\sum_{i=-\infty}^{\infty} x_i$ . . . . .	722
unit matrix . . . . .	335	$X^n$ . . . . .	274
unitary operator . . . . .	544	$x + A$ . . . . .	279
universal property of $\mathbb{N}_0$ . . . . .	103	$X \oplus Y$ . . . . .	297
upper bound . . . . .	70	$x_1 \otimes \dots \otimes x_n$ . . . . .	372
upper triangular matrix . . . . .	541	$X/Y$ . . . . .	296
$v_i$ . . . . .	285	zero divisor . . . . .	96
$\sum_{i \in \{1, \dots, n\}} V_i$ . . . . .	367	Zorn's Lemma . . . . .	87