



**NAMIBIA UNIVERSITY
OF SCIENCE AND TECHNOLOGY**

THE STEINITZ EXCHANGE THEOREM AND ITS
APPLICATIONS

A THESIS SUBMITTED IN FULFILMENT OF THE
REQUIREMENTS FOR THE DEGREE OF MASTER OF
SCIENCE IN APPLIED MATHEMATICS

OF

THE NAMIBIA UNIVERSITY OF SCIENCE AND
TECHNOLOGY

BY

TOBIAS KAENANDUNGE

211114383

AUGUST 2020

SUPERVISOR: PROF. GÜNTER HEIMBECK

CO-SUPERVISOR: DR. ONESMUS SHUUNGULA

Abstract

It seems that during the last decades, no research was done which is related to the Steinitz exchange theorem. However, the generalised Steinitz exchange theorem has been investigated in books and articles . The generalized Steinitz exchange theorem is not a theorem of linear algebra but for reaching generalization of the Steinitz exchange theorem which has applications for example in field theory, in the theory of abelian groups and in module theory.

The objective of this study was to prove the Steinitz exchange theorem of linear algebra for arbitrary vector spaces over arbitrary division rings. Nearly all books on linear algebra which have the Steinitz exchange theorem explicitly state and prove this theorem only for finitely generated vector spaces. Only one exception can be found. In another source, the Steinitz exchange theorem is proved under the additional assumption, that the linearly independent subset is finite.

In this study the exchange theorem of Steinitz is proved in full generality with the means of linear algebra. The statement of the theorem of Steinitz is a statement of the following type: under certain conditions there exists a set with certain properties. The question when this set is uniquely determined could be completely solved. In addition, an application of the theorem of Steinitz is presented. This is the classical application which was given already by Graßmann: Any two bases of a vector space are equipotent.

The first chapter is about the basic concepts of the study. The second chapter reviews the relevant literature and outlines the methodology used in the study. The literature review is mainly about the generalized theorem of Steinitz, but also include the versions of the Steinitz exchange theorem found in books of linear algebra. The third chapter presents the results of the study with proofs. The study is concluded in the last chapter with proposals for further study.

Acknowledgement

Foremost, I would like to express my sincere gratitude to my supervisor Prof. Günter Heimbeck for the continuous support for my MSc. study and research, for his patience, motivation, enthusiasm, mentorship and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better supervisor and mentor for MSc. study.

Beside my supervisor, I would like to thank my co-supervisor Dr. Onesmus Shuungula for all the help and guidance and making time for me to consult him during his busy schedules. His help really meant a lot.

My sincere thanks also goes to my colleagues Dr. Gemechu Dibaba, Dr. Dismas Ntirampeba and Mr. Frans Ndinodiva for their help with Latex. Thank you to all my colleagues in the department of Mathematics and Statistics at the Namibia University of Science and Technology for their support and motivation during my study.

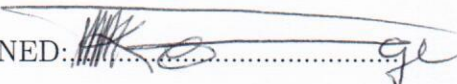
Last but not least, I would like to thank my parents for bringing me up and for nurturing me. Thanks to all the family members for the prayers and many thanks to my friends for the continuous support.

Dedication

This thesis is dedicated to my family.

Declaration

I, TOBIAS KAENANDUNGE, do hereby declare that this thesis is my own unaided work being submitted for the degree of Master of Science in Applied Mathematics at the Namibia University of Science and Technology. It has not been submitted for any degree or examination at any other university.

SIGNED: 

DATE: 27/08/2020

CERTIFICATION

This is to certify that this thesis titled “ **THE STEINITZ EXCHANGE THEOREM AND ITS APPLICATIONS** ” was undertaken by **Tobias Kaenandunge** at the Department of Mathematics, Namibia University of Science and Technology in partial fulfilment for the requirements of the award of the Master of Science degree in Applied Mathematics.

Prof G. Heimbeck

Supervisor

Signature: *G Heimbeck*

Date: *27/08/2020*

Mr. B.E. Obabueki

Head of Department

Signature: *B.E. Obabueki*

Date: *27/8/2020*

Prof F. Massamba

External Examiner

Signature: *F. Massamba*

Date: *27/08/2020*

Contents

1	Introduction and basic concepts	1
1.1	Introduction	1
1.1.1	Background	1
1.1.2	Statement of the problem	1
1.1.3	Objective of the study	2
1.2	Basic concepts from set theory	2
1.2.1	Ordered pairs	2
1.2.2	Ordered sets and functions.	3
1.2.3	Equipotent sets	14
1.3	Basic concepts from group theory and ring theory	17
1.4	Basic concepts from Linear Algebra	27
2	Review of relevant literature and methodology used in this research	38
2.1	Review of relevant literature	38
2.1.1	The Steinitz exchange theorem	38
2.1.2	The generalized Steinitz exchange theorem	39
2.2	Methodology used in this research	41
2.2.1	Methods from set theory	41
2.2.2	Methods from linear algebra and algebra	42
3	Results of the study	43
3.1	The Steinitz exchange theorem	43
3.2	Application of Steinitz exchange theorem	48
4	Conclusions and recommendations	49
4.1	Conclusions	49

4.2 Recommendations	50
Bibliography	51

Chapter 1

Introduction and basic concepts

1.1 Introduction

1.1.1 Background

This study is devoted to the exchange theorem of Steinitz. This theorem is a theorem of linear algebra which appeared first in Graßmann's book in 1844. Since this book could not be obtained, Graßmann's theorem is quoted as it is stated in (Graßmann, 1878). Because of the unusual terminology used in Graßmann's book - in those days linear algebra did not exist - the version of this theorem as it is stated in (van der Waerden, 1930) is quoted. van der Waerden's version comes very close to what Graßmann's book contains. In this context, two systems of vectors are called equivalent if they have the same span.

Exchange theorem of Steinitz (van der Waerden, 1930)

Let y_1, \dots, y_s be linearly independent vectors and x_1, \dots, x_r vectors such that each y_j is linearly dependent on x_1, \dots, x_r . Then there exists a system of exactly s vectors $x_{i_1}, x_{i_2}, \dots, x_{i_s}$ which can be exchanged with y_1, \dots, y_s such that the new system is equivalent to x_1, \dots, x_r . In particular $s \leq r$.

This theorem appeared again in an article of Steinitz in 1913 and, since then, it is called the exchange theorem of Steinitz.

1.1.2 Statement of the problem

The Steinitz exchange theorem contains the following assumptions

(1) The vector space is a real vector space.

(2) The two systems of vectors are finite.

One can ask whether these two assumptions are needed, and this is the problem that would be solved. In this regard, (van der Waerden, 1930) made an important contribution. He proved that the Steinitz exchange theorem is true for a vector space over a division ring, provided the vector space is of finite dimension.

1.1.3 Objective of the study

The objective of the study is to solve the problem which has been stated in 1.1.2. The outcome is that the two assumptions of the Steinitz exchange theorem can be deleted.

1.2 Basic concepts from set theory

This section discusses the basic definitions and theorems from set theory which will be used in this study. In some cases, the facts presented are not readily available in the literature. In other cases, simplified proofs are given or additions are made. Well ordered sets will be discussed because of the articles (Hughes, 1962–1964) and (Hughes, 1965-1966).

Because the concepts *binary relation* and *function* depend on ordered pairs, the set-theoretical definition of an ordered pair is given. This definition was introduced by Kuratowski. Also, the set-theoretical definition of a function is required. In this thesis, if A and B , then $A \subset B$ means A is a subset of B even if $A = B$.

1.2.1 Ordered pairs

Definition 1.2.1. Let A and B be sets.

$$(A, B) := \{\{A\}, \{A, B\}\}$$

(A, B) is called the ordered pair with components A and B .

In particular, ordered pairs are sets.

Proposition 1.2.1. Let A, B, C, D be sets. Then

$$(A, B) = (C, D) \iff A = C \text{ and } B = D.$$

Proof. (\Leftarrow) Suppose $A = C$ and $B = D$. Since (A, B) and (C, D) have the same first component and the same second component, $(A, B) = (C, D)$.

(\Rightarrow) Suppose $(A, B) = (C, D)$. Now $\{A\} = \bigcap_{X \in (A, B)} X = \bigcap_{X \in (C, D)} X = \{C\}$. Hence $A = C$.

Now it will be shown that $B = D$.

Case 1: Suppose $A = B$. Since $(A, B) = (C, D)$ and (A, B) is a singleton, (C, D) is a singleton. Hence $C = D$. Since $A = C$, all four sets A, B, C, D are equal. Therefore $B = D$.

Case 2: Suppose $A \neq B$. Since (A, B) is not a singleton, (C, D) is not a singleton. Hence $C \neq D$.

Further $\{A, B\} = \bigcup_{X \in (A, B)} X = \bigcup_{X \in (C, D)} X = \{C, D\}$. Since $B \in \{C, D\}$ and $B \neq C$, $B = D$. ■

From the proof of the previous proposition, one sees that if (A, B) is an ordered pair, then

$\bigcap_{X \in (A, B)} X = \{A\}$. Therefore A is uniquely determined by (A, B) . If $A = B$, B is uniquely

determined by (A, B) because A is uniquely determined. Suppose $A \neq B$. Then $\bigcup_{X \in (A, B)} X =$

$\{A, B\}$. Therefore B is uniquely determined by (A, B) . If (A, B) is an ordered pair, one can call A the first component of (A, B) and B the second component of (A, B) .

This is used when one introduces the projections of the cartesian product $X \times Y$ of two sets X and Y onto the factors X and Y .

$$p : X \times Y \rightarrow X \text{ defined by } p(a, b) := a \text{ and } q : X \times Y \rightarrow Y \text{ defined by } q(a, b) := b.$$

1.2.2 Ordered sets and functions.

Definition 1.2.2. An order relation on a set X is a binary relation on X which is reflexive, antisymmetric and transitive. A set together with an order relation is called an ordered set.

In the following it is not assumed that the ordered set is non-empty. In general, an ordered set (X, \leq) may contain elements $x, y \in X$ such that $x \not\leq y$ and $y \not\leq x$. If this is not the case, (X, \leq) is called a totally ordered or linearly ordered set, or a chain, and the order is called a total order or a linear order.

To order a set X , one chooses an order relation on X and consider this set together with this order relation. Every set can be ordered, for example, inclusion \subset is an order relation on every power set.

If (X, \leq) is an ordered set and $Y \subset X$, then \leq induces in Y an order relation. This order relation is the intersection of \leq and $Y^2 = Y \times Y$. In particular, every subset of an ordered set can be viewed as an ordered set.

Definition 1.2.3. A maximal element of an ordered set X is an element $m \in X$ such that if $x \in X$ and $m \leq x$, then $m = x$.

It is added that one must distinguish between maximal elements and the maximum of an ordered set. If an ordered set has a maximum, the maximum is a maximal element, and it is the only maximal element. In particular, if an ordered set has more than one maximal element, then there is no maximum. However, if an ordered set has exactly one maximal element, the maximal element need not be the maximum as the following example shows. Let X be an infinite set and $A \subset X$ a non-empty subset such that $X - A$ is infinite. Consider

$$\Omega := \mathcal{P}(A) \cup \{Y \subset X - A \mid Y \text{ finite}\}$$

ordered by inclusion. A is a maximal element and every finite subset of $X - A$ is not a maximal element. Therefore, Ω has exactly one maximal element, but it is not the maximum.

Similarly, one introduces minimal elements of an ordered set.

Definition 1.2.4. A minimal element of an ordered set X is an element $m \in X$ such that if $x \in X$ and $x \leq m$ then $x = m$.

For minimal elements, similar statements to those of maximal elements are true. An ordered set need not have a minimal element and an ordered set can have more than one minimal element.

Nearly every book on set theory uses the following proposition without providing a proof.

Proposition 1.2.2. *Every non-empty finite ordered set has maximal elements.*

Proof. By induction on the cardinality.

Let X be an ordered set such that $|X| = 1$. Then X is a singleton, say $X = \{a\}$. Then a is a maximal element of X . Let $n \in \mathbb{N}$ and assume that every ordered set of cardinality n has a maximal element. Let X be an ordered set such that $|X| = n + 1$. Then there exists $a \in X$ because $X \neq \emptyset$. Form $X' := X - \{a\}$. By induction hypothesis, X' has a maximal element $m \in X'$. If $m \not\leq a$, then m is a maximal element of X . Suppose $m \leq a$. Since $m \neq a$, $m < a$. Let $x \in X$ such that $a \leq x$. Then $m < x$. Since $x \notin X'$, $x = a$. Hence a is a maximal element of X . ■

An infinite ordered set need not have maximal elements. If X is an infinite set, then the finite subsets of X , ordered by inclusion, is an example of an ordered set which has no maximal elements for the following reason. Take $\Omega := \{Y \subset X \mid Y \text{ finite}\}$ ordered by inclusion. Let $M \in \Omega$. Since X is infinite and M is finite, there exist some $x \in X - M$. Since M is a finite subset of X , $M \cup \{x\}$ is finite. Hence $M \cup \{x\} \in \Omega$. $M \cup \{x\}$ is strictly bigger than M . Since M was any element of Ω , Ω has no maximal elements.

In case of infinite ordered sets, it seems that in most cases maximal elements are obtained by applying Zorn's lemma.

Zorn's lemma.

Let (X, \leq) be an ordered set. If every chain in X is bounded above, then (X, \leq) has maximal elements.

An ordered set which satisfies the hypothesis of Zorn's lemma is non-empty because \emptyset is a chain in the ordered set and bounded above. It remains to mention that if X is finite, Zorn's lemma can be proved, compare Proposition 1.2.2.

Theorem 1.2.1 (Hausdorff Maximality Principle). *Let (X, \leq) be an ordered set and $\Gamma_0 \subset X$ a chain. Then there exists a maximal chain in X which contains Γ_0 .*

Proof. Take

$$\Omega := \{\Gamma \subset X \mid \Gamma \text{ a chain, } \Gamma_0 \subset \Gamma\}$$

ordered by inclusion.

Since $\Gamma_0 \in \Omega$, $\Omega \neq \emptyset$. Let C be a non-empty chain in Ω and consider

$$\bar{\Gamma} = \bigcup_{\Gamma \in C} \Gamma.$$

Since $\Gamma \subset X$ for all $\Gamma \in C$, $\bar{\Gamma} \subset X$. Since $C \neq \emptyset$, C contains an element, say $\Gamma_1 \in \Omega$. Since $\Gamma_0 \subset \Gamma_1$ and $\Gamma_1 \subset \bar{\Gamma}$, $\Gamma_0 \subset \bar{\Gamma}$. Let $a, b \in \bar{\Gamma}$. Since $\bar{\Gamma} = \bigcup_{\Gamma \in C} \Gamma$, there exists $\Gamma_1, \Gamma_2 \in C$ such that $a \in \Gamma_1$, $b \in \Gamma_2$. Since C is a chain, $\Gamma_1 \subset \Gamma_2$ or $\Gamma_2 \subset \Gamma_1$. Hence $a, b \in \Gamma_2$ or $a, b \in \Gamma_1$. Since Γ_1, Γ_2 are chains, a and b are comparable. Therefore, $\bar{\Gamma}$ is a chain. Therefore $\bar{\Gamma} \in \Omega$. Since $\Gamma \subset \bar{\Gamma}$ for all $\Gamma \in C$, $\bar{\Gamma}$ is an upper bound of C . Therefore Zorn's lemma is applicable. Hence Ω has maximal elements. Let $\Gamma^* \in \Omega$ be a maximal element of Ω . Then Γ^* is a chain in X and $\Gamma_0 \subset \Gamma^*$. Let Γ be a chain in X such that $\Gamma^* \subset \Gamma$. Since Γ is a chain in X and $\Gamma_0 \subset \Gamma$, $\Gamma \in \Omega$. Since Γ^* is a maximal element of Ω , $\Gamma = \Gamma^*$. Therefore Γ^* is a maximal chain in X which contains Γ_0 . ■

From Hausdorff Maximality Principle, one obtains Zorn's lemma in a very simple manner.

Let (X, \leq) be an ordered set which satisfy the hypothesis of Zorn's lemma. Since \emptyset is a chain in X , by the Hausdorff maximality principle, there exists a maximal chain Γ in X . By the hypothesis of Zorn's lemma, Γ is bounded above, i.e. there exists $a \in X$ such that $x \leq a$ for all $x \in \Gamma$. Then $\Gamma \cup \{a\}$ is a chain in X . Since Γ is a maximal chain, $a \in \Gamma$. Therefore a is a maximum of Γ . Therefore, a is a maximal element of X .

The following seems to be worthwhile mentioning. If a maximal chain is bounded above, then it has a greatest element. In particular, if an ordered set satisfies the condition of Zorn's lemma, then every maximal chain has a greatest element.

Definition 1.2.5. A function is a set f of ordered pairs such that if $(x, y), (x, y') \in f$, then $y = y'$.

Some definitions and notations are mentioned as follows.

If f is a function, then $X := \{x \mid \text{there exists some } y \text{ such that } (x, y) \in f\}$ is called the domain of f . The domain of the function f will also be denoted by D_f . If $(x, y) \in f$, one writes $y = f(x)$ or $y = x^f$. $y = f(x)$ or $y = x^f$ is called the image of x with respect to f . The range of f is the set $f(X) := \{f(x) \mid x \in X\}$.

A codomain of f is a superset of the range of f . The domain and the range of a function are uniquely determined by the function. Any superset of the range of a function is a codomain of that function. In particular every function has many codomains. If X is the domain of f and Y is a codomain of f , one writes $f : X \rightarrow Y$. Note that $f = \{(x, f(x)) \mid x \in X\}$. In particular a function is equal to its graph.

Every subset of a function is a function. In particular, if f and g are functions, then $f \cap g$ is a function. \emptyset is a function, called the empty function. Let f and g be functions such that $f \subset g$. Then f is called a restriction of g and g is called an extension of f . If $f : X \rightarrow Y$ and $A \subset X$, then $f|_A := f \cap (A \times Y)$ is called the restriction of f to A .

In general, the union of two functions is not a function.

Proposition 1.2.3. *Let f and g be functions. Then*

$$f \cup g \text{ is a function} \iff f(x) = g(x) \text{ for all } x \in D_f \cap D_g.$$

Proof. Suppose $f \cup g$ is a function. Let $x \in D_f \cap D_g$. Since $x \in D_f$ and $x \in D_g$, $(x, f(x)) \in f$ and $(x, g(x)) \in g$. Therefore $(x, f(x)), (x, g(x)) \in f \cup g$. Since $f \cup g$ is a function, $f(x) = g(x)$. To

prove the converse implication, suppose $f(x) = g(x)$ for all $x \in D_f \cap D_g$. Let $(x, y), (x, y') \in f \cup g$. If both $(x, y), (x, y') \in f$ or both $(x, y), (x, y') \in g$, then $y = y'$ because f and g are functions. Suppose $(x, y) \in f$ and $(x, y') \in g$. Then $y = f(x)$ and $y' = g(x)$. Since $x \in D_f \cap D_g$, $f(x) = g(x)$ and hence $y = y'$. Therefore $f \cup g$ is a function. ■

Proposition 1.2.4. *Let Γ be a chain of functions and $g := \bigcup_{f \in \Gamma} f$. Then g is a function.*

$$D_g = \bigcup_{f \in \Gamma} D_f \text{ and } g(D_g) = \bigcup_{f \in \Gamma} f(D_f).$$

Proof. Let $a \in g$. Then there exists $f \in \Gamma$ such that $a \in f$. Since f is a function, a is an ordered pair. Hence, g is a set of ordered pairs. Let $(x, y), (x, y') \in g$. Then there exist $f_1, f_2 \in \Gamma$ such that $(x, y) \in f_1$ and $(x, y') \in f_2$. Since Γ is a chain, $f_1 \subset f_2$ or $f_2 \subset f_1$. Hence $(x, y), (x, y') \in f_2$ or $(x, y), (x, y') \in f_1$. Since f_1 and f_2 are functions, $y = y'$. Therefore g is a function.

Let $x \in D_g$. Since $(x, g(x)) \in g$, there exists $f \in \Gamma$ such that $(x, g(x)) \in f$. Then $x \in D_f$. Therefore, $D_g \subset \bigcup_{f \in \Gamma} D_f$. Suppose $x \in \bigcup_{f \in \Gamma} D_f$. Then there exists $f \in \Gamma$ such that $x \in D_f$. Since $f \subset g$, $x \in D_g$. Hence $\bigcup_{f \in \Gamma} D_f \subset D_g$. This implies that $D_g = \bigcup_{f \in \Gamma} D_f$. The proof for the range is similar. ■

Definition 1.2.6. Let X be a set and \leq, \leq' order relations on X .

If $\leq \subset \leq'$, then \leq is called finer than \leq' and \leq' is called coarser than \leq (Bourbaki, 1968).

It seems that these definitions are only used by (Bourbaki, 1968).

Since every order relation on X is reflexive, equality $=$ is the finest order relation on X .

The following theorem shows that every order relation is contained in a coarsest order relation.

Theorem 1.2.2. *Let (X, \leq_0) be an ordered set*

$$\Omega := \{ \leq \mid \leq \text{ an order relation on } X \text{ such that } \leq_0 \subset \leq \}$$

ordered by inclusion. Ω has maximal elements and these are maximal order relations on X .

Proof. $\Omega \neq \emptyset$ because $\leq_0 \in \Omega$. Let $\Gamma \subset \Omega$ be a non-empty chain and consider

$$\tilde{\leq} := \bigcup_{\leq \in \Gamma} \leq.$$

Now it will be shown that $\tilde{\leq}$ is an order relation on X . It is clear that $\tilde{\leq}$ is a set of ordered pairs. Let $x \in X$. Since $\Gamma \neq \emptyset$, there exists $\leq \in \Gamma$. Since \leq is an order relation, $(x, x) \in \leq$. Since $\leq \subset \tilde{\leq}$, $(x, x) \in \tilde{\leq}$. Therefore, $\tilde{\leq}$ is reflexive. Let $(x, y), (y, x) \in \tilde{\leq}$. Then there exist

$\leq_1, \leq_2 \in \Gamma$ such that $(x, y) \in \leq_1$ and $(y, x) \in \leq_2$. Since Γ is a chain, $\leq_1 \subset \leq_2$ or $\leq_2 \subset \leq_1$. Hence, $(x, y), (y, x) \in \leq_2$ or $(x, y), (y, x) \in \leq_1$. Since \leq_1 and \leq_2 are order relations, $x = y$. Therefore, $\tilde{\leq}$ is anti-symmetric. Let $(x, y), (y, z) \in \tilde{\leq}$. There exist $\leq_1, \leq_2 \in \Gamma$ such that $(x, y) \in \leq_1, (y, z) \in \leq_2$. As above, $(x, y), (y, z) \in \leq_1$ or $(x, y), (y, z) \in \leq_2$. Since \leq_1 and \leq_2 are order relations, $(x, z) \in \leq_1$ or $(x, z) \in \leq_2$. Hence $(x, z) \in \tilde{\leq}$. Therefore, $\tilde{\leq}$ is transitive. Thus $\tilde{\leq}$ is an order relation on X and this order relation is coarser than \leq_0 . Therefore $\tilde{\leq} \in \Omega$. Since $\leq \subset \tilde{\leq}$ for all $\leq \in \Gamma$, $\tilde{\leq}$ is an upper bound of Γ . By Zorn's lemma, Ω has maximal elements.

Let \leq' be a maximal element of Ω and \leq an order relation on X such that $\leq' \subset \leq$. Since $\leq_0 \subset \leq', \leq_0 \subset \leq$. Therefore $\leq \in \Omega$. Since \leq' is a maximal element of Ω , $\leq' = \leq$. Therefore \leq' is a maximal order relation on X . ■

Theorem 1.2.3. *Let (X, \leq) be an ordered set and $r, s \in X$ non-comparable elements. Then the following hold. The union of \leq and $\{(x, y) \in X^2 \mid x \leq r \text{ and } s \leq y\}$ is an order relation on X :*

$$\leq' := \leq \cup \{(x, y) \in X^2 \mid x \leq r, s \leq y\}.$$

\leq' is the finest order relation on X which is strictly coarser than \leq and which satisfies $r \leq' s$.

Proof. Since $\leq \subset \leq'$, \leq' is reflexive. Let $x, y \in X$ such that $x \leq' y$ and $y \leq' x$. Then $x \leq y$ or $x \leq r$ and $s \leq y$, and $y \leq x$ or $y \leq r$ and $s \leq x$. Assume $x \leq r$ and $s \leq y$. If $y \leq x$, then $s \leq r$ which is a contradiction. Suppose $y \leq r$ and $s \leq x$. Then $s \leq r$, which is again a contradiction. Since $x \leq r$ and $s \leq y$ is not a possibility, $x \leq y$. Assume $y \leq r$ and $s \leq x$. Then $s \leq r$ which is a contradiction. Thus $y \leq x$. Therefore $x = y$. Hence \leq' is antisymmetric. Let $x, y, z \in X$ such that $x \leq' y, y \leq' z$. Then $x \leq y$ or $x \leq r, s \leq y$ and $y \leq z$ or $y \leq r, s \leq z$. If $x \leq y$ and $y \leq z, x \leq z$. Hence $x \leq' z$. Suppose $x \leq y$ and $y \leq r$ and $s \leq z$. Then $x \leq' z$. Suppose $x \leq r, s \leq y$ and $y \leq z$. Then $x \leq r$ and $s \leq z$. Hence $x \leq' z$. Suppose $x \leq r$ and $s \leq y$ and $y \leq r$ and $s \leq z$. This is impossible because $s \leq r$ which is a contradiction. Therefore, \leq' is transitive. Since $r \leq r$ and $s \leq s, r \leq' s$. Let $\tilde{\leq}$ be an order relation on X such that $\leq \subset \tilde{\leq}$ and $(r, s) \in \tilde{\leq}$. Let $(x, y) \in X \times X$ such that $x \leq r$ and $s \leq y$. Then, $(x, r), (r, s), (s, y) \in \tilde{\leq}$. Since $\tilde{\leq}$ is transitive, $(x, y) \in \tilde{\leq}$. Hence $\leq' \subset \tilde{\leq}$. It follows that \leq' is the finest order relation on X which is strictly coarser than \leq . ■

Theorem 1.2.4. *Let (X, \leq) be an ordered set. Then \leq is a coarsest order relation on X if and only if \leq is linear.*

Proof. Suppose \leq is a coarsest order relation on X . Let $r, s \in X$. If $r \not\leq s$ and $s \not\leq r$, then by the previous theorem, \leq is not a coarsest order relation on X . Therefore $r \leq s$ or $s \leq r$. Hence \leq is linear. ■

To prove the converse implication, suppose \leq is linear. Let \leq' be an order relation on X such that $\leq \subset \leq'$. Let $(r, s) \in \leq'$. If $r = s$, $(r, s) \in \leq$. Suppose $r \neq s$. Since \leq is linear, $(r, s) \in \leq$. $(s, r) \in \leq$ leads to $(s, r) \in \leq'$ which is not possible. Therefore $(r, s) \in \leq$. Thus $\leq' \subset \leq$ and this means $\leq = \leq'$. Hence \leq is maximal. ■

Theorem 1.2.5. *Let \leq be an order relation on X . Then $\bigcap_{\substack{\leq \subset \leq_1 \\ \leq_1 \text{ linear}}} \leq_1 = \leq$.*

Proof. Let

$$\tilde{\leq} := \bigcap_{\substack{\leq \subset \leq_1 \\ \leq_1 \text{ linear}}} \leq_1.$$

$\tilde{\leq}$ exists because by theorems 1.2.2 and 1.2.4 there exists an order relation which is coarser than \leq which is linear. Clearly, $\leq \subset \tilde{\leq}$. Let $r, s \in X$ such that $r \not\leq s$ and $s \not\leq r$. There exists an order relation \leq' on X such that $\leq \subset \leq'$ and $r \leq' s$. There exists a linear order \leq_1 such that $\leq' \subset \leq_1$. Then $r \leq_1 s$. This implies $s \not\leq_1 r$. Similarly, there exists an order relations \leq'' on X such that $\leq \subset \leq''$ and $s \leq'' r$. There exists a linear order \leq_2 such that $\leq'' \subset \leq_2$. Since $s \leq'' r$, $s \leq_2 r$. This implies $r \not\leq_2 s$. Therefore, $(r, s), (s, r) \notin \tilde{\leq}$. This means that $X^2 - \leq \subset X^2 - \tilde{\leq}$. Therefore, $\tilde{\leq} \subset \leq$ and hence $\tilde{\leq} = \leq$. ■

The following will be used in this section.

Definition 1.2.7. Let (X, \leq) be an ordered set and $a \in X$. Then,

$$s(a) = \{x \in X | x < a\}, \bar{s}(a) = \{x \in X | x \leq a\},$$

where $s(a)$ is called the initial segment determined by a and $\bar{s}(a)$ is called the weak initial segment determined by a according to, for example, (Halmos, 1960).

Clearly, $s(a) = \bar{s}(a) - \{a\}$.

Definition 1.2.8. Let (X, \leq) be an ordered set.

$$A \subset X \text{ a section of } X \iff s(a) \subset A \text{ for all } a \in A.$$

In some books, sections are also called down sets.

Remark. (1) $s(a)$, $\bar{s}(a)$ are sections.

Proof. Let $x \in s(a)$ and $y \in X$ such that $y \leq x$. Since $x < a$, $y < a$. Hence $y \in s(a)$. Therefore $s(a)$ is a section. The proof for $\bar{s}(a)$ is similar. ■

(2) $\{\bar{s}(a) \mid a \in X\}$ is isomorphic to X .

Proof. Define $\varphi : X \rightarrow \{\bar{s}(a) \mid a \in X\}$ defined by $\varphi(a) = \bar{s}(a)$. Let $a, b \in X$ with $a \leq b$. Let $x \in X$ with $x \leq a$. Then, by transitivity, $x \leq b$. Hence $\varphi(a) \subset \varphi(b)$. Let $a, b \in X$ with $\varphi(a) = \varphi(b)$. Since $a \in \varphi(a) = \varphi(b)$, $a \leq b$. Since $b \in \varphi(b) = \varphi(a)$, $b \leq a$. Hence $a = b$. Thus φ is injective. It is clear from the definition of φ that φ is surjective. Therefore φ is an isomorphism. ■

Proposition 1.2.5. Let (X, \leq) be an ordered set and $(A_i)_I$ a non-empty family of sections of (X, \leq) . Then $\bigcap_I A_i$ and $\bigcup_I A_i$ are sections of (X, \leq) .

Proof. Let $a \in \bigcap_I A_i$ and $x \in X$ such that $x \leq a$. Let $j \in I$. Then $a \in A_j$. Since $x \leq a$ and A_j is a section, $x \in A_j$. Hence $x \in \bigcap_I A_i$. Therefore $\bigcap_I A_i$ is a section of X .

Let $a \in \bigcup_I A_i$ and $x \in X$ such $x \leq a$. Then there exists $j \in I$ such that $a \in A_j$. Since $x \leq a$ and A_j is a section, $x \in A_j$. Hence $x \in \bigcup_I A_i$. Therefore $\bigcup_I A_i$ is a section of X . ■

There is an interesting relationship between linear orders on a set X and maximal chains of the powerset of X . According to (Lüneburg, 1989), Hessenberg observed this in 1906. It seems that Hessenberg has never published this fact.

Theorem 1.2.6. (1) Let (X, \leq) be a linearly ordered set. Then the set of all sections of (X, \leq) is a maximal chain of $\mathcal{P}(X)$.

(2) Let X be a set and Γ a maximal chain of $\mathcal{P}(X)$. If $x, y \in X$, consider

$$x \leq y \iff \text{if } A \in \Gamma \text{ such that } y \in A, \text{ then } x \in A.$$

\leq is a linear order and Γ is the set of all the sections of (X, \leq) .

Proof. (1) Let

$$\Gamma := \{A \subset X \mid A \text{ a section of } (X, \leq)\}.$$

Let $A, B \in \Gamma$ such that $A \not\subset B$. Then there exists $z \in A - B$. Since $z \in A$ and A is a section of X , $\bar{s}(z) \subset A$. Since $z \notin B$ and B is a section, every element of B is smaller than z . Hence $B \subset \bar{s}(z)$. Therefore, $B \subset A$. Thus Γ is a chain. Let $D \subset X$ such that $\Gamma \cup \{D\}$ is a chain. Let $d \in D$. Since $s(d)$ is a section of X , $s(d) \in \Gamma$. Since $\Gamma \cup \{D\}$ is a chain, $s(d) \subset D$ or $D \subset s(d)$. Since $d \notin s(d)$ and $d \in D$, $D \subset s(d)$ is not possible. Hence $s(d) \subset D$. Therefore D is a section. Thus $\Gamma \cup \{D\} = \Gamma$. Therefore, Γ is a maximal chain of $\mathcal{P}(X)$.

- (2) By proposition 1.4.5, $I_x \in \Gamma$. If $x, y \in A$, then $x \leq y$ if and only if every section which contains y also contains x . Let $x, y \in X$ such that $x \leq y$. Since $\{A \in \Gamma \mid A \in A\} \subset \{A \in \Gamma \mid x \in A\}$, $I_x \subset I_y$. Suppose $I_x \subset I_y$. Let $A \in \Gamma$ such that $y \in A$. Then $I_y \subset A$. Therefore $x \in A$. Hence $x \leq y$. This means $x \leq y \iff I_x \subset I_y$. Since $I_x \subset I_x$, $x \leq x$ for all $x \in X$. Therefore \leq is reflexive. Suppose $x, y \in X$ such that $x \leq y$ and $y \leq x$. Since $I_x \subset I_y$ and $I_y \subset I_x$, $I_x = I_y$. Consider $I_x - \{x\}$. Let $D \in \Gamma$. Suppose $x \in D$. Then $I_x \subset D$ which implies $I_x - \{x\} \subset D$. Suppose $x \notin D$. Since I_x and D are comparable, $D \subset I_x$. Since $x \notin D$, $D \subset I_x - \{x\}$. This implies that $I_x - \{x\} \in \Gamma$. Similarly, $I_y - \{y\} \in \Gamma$. Now $I_x - \{x\} \subset I_y - \{y\}$ or $I_y - \{y\} \subset I_x - \{x\}$. This is a contradiction if $x \neq y$. Therefore $x = y$ and hence \leq is anti-symmetric. Suppose $x, y, z \in X$ such that $x \leq y$ and $y \leq z$. Since $I_x \subset I_y$ and $I_y \subset I_z$, $I_x \subset I_z$. Hence $x \leq z$ and this implies that \leq is transitive. Let $x, y \in X$. Since $I_x, I_y \in \Gamma$, $I_x \subset I_y$ or $I_y \subset I_x$. Hence $x \leq y$ or $y \leq x$. Therefore, \leq is linear. Let $A \in \Gamma$ and $x \in A$. Then $I_x \subset A$. Therefore, A is a section of (X, \leq) . ■

The following theorems 1.2.7 and 1.2.8 are taken from Lüneburg (1989).

Theorem 1.2.7. *Let X be a set and Γ a maximal chain of $\mathcal{P}(X)$. Let $A \in \Gamma - \{X\}$. Take*

$$I_A := \bigcap_{\substack{X \in \Gamma \\ A \subsetneq X}} X.$$

Then $I_A = A$ or $|I_A - A| = 1$.

Proof. It is clear that $I_A \supset A$. Suppose $I_A \neq A$. Let $Y \in \Gamma$. Since Γ is a chain, A and Y are comparable. Therefore, $Y \subset A$ or $A \subsetneq Y$. Hence $Y \subset A$ or $I_A \subset Y$. There exists $x \in I_A - A$. Since $Y \in \Gamma$, $Y \subset A$ or $A \cup \{x\} \subset I_A \subset Y$. Therefore, $A \cup \{x\}$ and Y are comparable. Hence $\Gamma \cup \{A \cup \{x\}\}$ is a chain. Since Γ is a maximal chain, $\Gamma \cup \{A \cup \{x\}\} = \Gamma$. Therefore, $A \cup \{x\} \in \Gamma$. Since $A \subsetneq A \cup \{x\}$, $I_A \subset A \cup \{x\}$. Since $A \subsetneq I_A$, $I_A = A \cup \{x\}$. Therefore, $|I_A - A| = 1$. ■

As stated earlier, well ordered sets are discussed in this study because of the articles of Hughes, (Hughes, 1962–1964) and (Hughes, 1965-1966).

Definition 1.2.9. A well ordered set is an ordered set such that every non-empty subset has a minimum.

Remark. *Every well-ordered set is linearly ordered.*

Proof. Let X be well-ordered. Let $a, b \in X$. Take $Y := \{a, b\}$. Y is a non-empty subset of X . Since X is well-ordered, $m := \min(Y) = \min\{a, b\}$ exists. If $m = a$, then $a \leq b$. If $m = b$, $b \leq a$. Hence X is linearly ordered. ■

Theorem 1.2.8. *Let (X, \leq) be a linearly ordered set. The following statements are equivalent.*

- a) (X, \leq) is well-ordered.
- b) If A is a section of (X, \leq) and $A \neq X$, $I_A \neq A$.
- c) If A is a section of (X, \leq) and $A \neq X$, $|I_A - A| = 1$.

Proof. a) \Rightarrow b) Suppose (X, \leq) is well-ordered. Let A be a section of (X, \leq) such that $A \neq X$. Since $X - A \neq \emptyset$ and (X, \leq) is well-ordered, $s := \min(X - A)$ exists. Let Y be a section of (X, \leq) such that $A \subsetneq Y$. Then $Y \cap (X - A) \neq \emptyset$. Let $x \in Y \cap (X - A)$. Then $s \leq x$. Since Y is a section and $s \leq x$, $s \in Y$. This implies $s \in I_A$. Since $s \notin A$, $I_A \neq A$.

b) \Rightarrow c) From theorem 1.2.7, $|I_A - A| = 1$.

c) \Rightarrow a) Suppose $|I_A - A| = 1$ for each section A of (X, \leq) such that $A \neq X$. Let $Y \subset X$ such that $Y \neq \emptyset$. Take

$$B := \{a \in X - Y \mid a \leq x \text{ for all } x \in Y\}.$$

Let $a \in B$ and $b \in X$ such that $b \leq a$. Let $x \in Y$. Since $a \leq x$, $b \leq x$. Assume $b \in Y$. Since $b \leq a$ and $a \leq x$ for all $x \in Y$, $a \leq b$. Hence $a = b$. This implies $a \in Y$ which is a contradiction. Therefore, $b \notin Y$. Since $b \in X - Y$, $b \in B$. Therefore, B is a section. Since $Y \neq \emptyset$, $B \neq X$. By hypothesis $I_B - B$ is a singleton, say $I_B - B = \{z\}$. Let $x \in Y$. Suppose $x \leq z$. Since I_B is a section and $z \in I_B$, $x \in I_B$. Since $Y \cap B = \emptyset$, $x \notin B$. Therefore $x = z$. Hence $z \leq x$ for all $x \in Y$. Since $z \notin B$, $z \in Y$. Therefore z is a minimum of Y . ■

Theorem 1.2.9 (Transfinite induction). *Let (X, \leq) be a well-ordered set and $I \subset X$ which has the following property: if $a \in X$ and $s(a) \subset I$, then $a \in I$. Then $I = X$*

Proof. Assume $I \neq X$. Then $X - I \neq \emptyset$. Since X is well-ordered, $m := \min(X - I)$ exists. Let $x \in s(m)$. Since $x < m$, $x \notin X - I$. Since $x \in X$, $x \in I$. Therefore, $s(m) \subset I$. By hypothesis $m \in I$. This is a contradiction because $m \in X - I$. Therefore $I = X$. ■

The following theorem is of fundamental importance.

Theorem of Zermelo. Every set can be well-ordered.

The proofs of this theorem and the next theorem will be skipped.

Let W be a well-ordered set, X any set and $a \in W$. A function $t : s(a) \rightarrow X$ is called a sequence in X of type a .

Transfinite recursion. Let f be a function with co-domain X whose domain is the set of all sequences in X of type a for any $a \in W$. Then there exists exactly one function $U : W \rightarrow X$ such that $U(a) = f(U|_{s(a)})$ for all $a \in W$.

Definition 1.2.10. Let X be a set. A partition of X is a subset $\Pi \subset \mathcal{P}(X)$ such that:

- (1) $\bigcup_{A \in \Pi} A = X$
- (2) If $A, B \in \Pi$, then $A \cap B = \emptyset$ or $A = B$
- (3) $\emptyset \notin \Pi$

The elements of a partition are called *components*.

Note that the empty set is the partition of the empty set.

Definition 1.2.11. Let Π be a partition of a set X . A transversal or a complete system of representatives of Π is a subset $T \subset X$ such that $|T \cap A| = 1$ for all $A \in \Pi$.

In the following theorem, the Axiom of Choice will be used. Therefore the axiom of choice is stated.

Axiom of Choice.

Let $(X_i)_I$ be a family of non-empty sets. Then there exists a function $f : I \rightarrow \bigcup_I X_i$ such that $f(i) \in X_i$ for all $i \in I$.

f is also called a choice function. Zorn's lemma is equivalent to the Axiom of Choice. This means the following: If one assumes that Zorn's lemma is true, one can prove the Axiom of Choice and if one assumes that the Axiom of Choice is true, one can prove Zorn's lemma. For a proof compare, for example, (Halmos, 1960).

The proof of the next theorem is an application of the Axiom of Choice.

Theorem 1.2.10. *Every partition has a transversal.*

Proof. Let Π be a partition of X . $(A)_{A \in \Pi}$ is a family of non-empty sets because $\emptyset \notin \Pi$. By the axiom of choice, there exists a function $f : \Pi \rightarrow X$ such that $f(A) \in A$ for all $A \in \Pi$. Choose

$$T := f(\Pi).$$

Clearly, $T \subset X$. Let $A \in \Pi$. Then $f(A) \in A$ and $f(A) \in T$. Hence $f(A) \in T \cap A$. Let $Z \in T \cap A$. Since $Z \in T$, there exists $B \in \Pi$ such that $f(B) = Z$. Since $Z \in B$ and $Z \in A$, $Z \in A \cap B$. Since A and B are components of a partition, $A = B$. Since $Z = f(A)$, $T \cap A = \{f(A)\}$. Thus T is a transversal. ■

1.2.3 Equipotent sets

Definition 1.2.12. Let X, Y be sets.

X is equipotent to Y : \iff there exists a bijection $f : X \rightarrow Y$.

This means two sets are said to be equipotent if and only if they contain the same number of elements. Note that this also makes sense if X and Y are infinite sets.

It is remarkable that the relation equipotent was studied long before set theory came into existence. Set theory started in 1872 with an article of Georg Cantor. Around 1650, the physicist Galilei published an article where it was verified that the set \mathbb{N} of the natural numbers is equipotent to the set \mathbb{Z} of the integers, inspite of the fact that \mathbb{N} is a proper subset of \mathbb{Z} .

‘Equipotent’ is an equivalence relation on the class of all sets. Take note that the collection of all sets is not a set. In particular, the equivalence class of a non-empty set is not a set.

If one must prove that two sets are equipotent, in many cases one uses the theorem of Schröder-Bernstein. The proof of this theorem will be skipped here.

Schröder-Bernstein theorem (Halmos, 1960). Let X and Y be sets such that X is equipotent to a subset of Y and Y is equipotent to a subset of X . Then X and Y are equipotent.

The theorem of Schröder-Bernstein is a theorem about arbitrary sets. However the proof of this theorem does not depend on the Axiom of Choice.

For the proof of the following theorem it will be used that if X and Y are any two sets, then X is equipotent to a subset of Y or Y is equipotent to a subset of X .

Definition 1.2.13. A set X is called countable if and only if it is equipotent to a subset of \mathbb{N} . If X is countable, then X is finite or equipotent to \mathbb{N} . Sets which are equipotent to \mathbb{N} are called countably infinite or denumerable.

Theorem 1.2.11. *Every infinite set contains a countably infinite subset.*

Proof. Let X be infinite. Then X is equipotent to a subset of \mathbb{N} or \mathbb{N} is equipotent to a subset of X . If X is equipotent to a subset of \mathbb{N} , then X is countably infinite. Hence X is a countably infinite subset of X . If \mathbb{N} is equipotent to a subset of X , then X contains a countably infinite subset. ■

To prove the next theorem, some preparations are needed. In particular, the product theorem of set theory is required.

Product Theorem of Set Theory.

If X is infinite, then X^2 is equipotent to X .

The following proposition is a simple consequence of the product theorem.

Proposition 1.2.6. *Let X be an infinite set.*

1. X^n is equipotent to X for all $n \in \mathbb{N}$.
2. $\bigcup_{n=1}^{\infty} X^n$ is equipotent to X .

Proof. 1. Proof by induction on n .

X^1 is equipotent to X .

Let $n \in \mathbb{N}$ and assume X^n is equipotent to X . X^{n+1} is equipotent to $X^n \times X$. Since X^n is equipotent to X , $X^n \times X$ is equipotent to $X \times X$. Since $X \times X$ is equipotent to X , X^{n+1} is equipotent to X .

2. Since X^n is equipotent to X and X is equipotent to $X \times \{n\}$, X^n is equipotent to $X \times \{n\}$ for all $n \in \mathbb{N}$. Therefore $\bigcup_{n=1}^{\infty} X^n$ is equipotent to $\bigcup_{n=1}^{\infty} (X \times \{n\}) = X \times \mathbb{N}$. Since X is infinite, \mathbb{N} is equipotent to a subset of X . Therefore, $\bigcup_{n=1}^{\infty} X^n$ is equipotent to a subset of $X \times X$. Since $X \times X$ is equipotent to X , $\bigcup_{n=1}^{\infty} X^n$ is equipotent to a subset of X . Since X is equipotent to X^1 and $X^1 \subset \bigcup_{n=1}^{\infty} X^n$, X is equipotent to a subset of $\bigcup_{n=1}^{\infty} X^n$. By the theorem of Schröder-Bernstein, $\bigcup_{n=1}^{\infty} X^n$ is equipotent to X . ■

The following notation will be used: if X is any set,

$$\text{Fin}(X) := \{A \subset X \mid A \text{ finite}\}.$$

Theorem 1.2.12. *If X is infinite, $\text{Fin}(X)$ is equipotent to X .*

Proof. Let $f : \bigcup_{n=1}^{\infty} X^n \rightarrow \text{Fin}(X)$ be defined by $f(a) = \{a_1, \dots, a_m\}$

$f(\bigcup_{n=1}^{\infty} X^n) = \text{Fin}(X) - \{\emptyset\}$. Hence $\text{Fin}(X) - \{\emptyset\}$ is equipotent to a subset of $\bigcup_{n=1}^{\infty} X^n$. Since $\text{Fin}(X)$ is infinite, $\text{Fin}(X)$ is equipotent to $\text{Fin}(X) - \{\emptyset\}$. Therefore, $\text{Fin}(X)$ is equipotent to a subset of $\bigcup_{n=1}^{\infty} X^n$. Since $\bigcup_{n=1}^{\infty} X^n$ is equipotent to X , $\text{Fin}(X)$ is equipotent to a subset of X . $x \in X \mapsto \{x\} \in \text{Fin}(X)$ is injective. Therefore, X is equipotent to a subset of $\text{Fin}(X)$. Thus by the theorem of Schröder-Bernstein, $\text{Fin}(X)$ and X are equipotent. ■

The following fact will be used in the proof of the next theorem: the range of any function is equipotent to a subset of the domain of the function.

The next theorem is known but not stated explicitly in books. It seems to be an interesting theorem. Later in this study, it will be used to prove the theorem of Löwig.

Theorem 1.2.13. *Let X be infinite and Γ a collection of finite subsets of X such that $\bigcup_{A \in \Gamma} A = X$. Then Γ is equipotent to X .*

Proof. By Theorem 1.2.12, $\text{Fin}(X)$ is equipotent to X . Since $\Gamma \subset \text{Fin}(X)$, Γ is equipotent to a subset of X . In order to prove that X is equipotent to a subset of Γ , take

$$S := \{A \times \{A\} \mid A \in \Gamma\}$$

and consider

$$C := \bigcup_{B \in S} B = \{(x, A) \mid A \in \Gamma, x \in A\}.$$

Let $x \in X$. Since $\bigcup_{A \in \Gamma} A = X$, $x \in \bigcup_{A \in \Gamma} A$. Therefore there exists $A \in \Gamma$ such that $x \in A$. Then $\varphi(x, A) := x$ defines a function from C to X . It is clear that φ is surjective. Therefore, X is equipotent to a subset of C . Let $A \in \Gamma$. Since Γ is infinite, by theorem 1.4.13, Γ contains a subset $B \subset \Gamma$ which is equipotent to \mathbb{N} . Then B contains a subset $B_1 \subset B$ which is equipotent to A . Since A is equipotent to a subset of Γ , there exist an injection $\psi : A \rightarrow \Gamma$. $(x, A) \in A \times \{A\} \rightarrow (\psi(x), A) \in \Gamma \times \Gamma$ is injective because φ is injective. Since the elements of S are disjoint in pairs, out of these functions, one makes an injection from C to $\Gamma \times \Gamma$. Therefore C is equipotent to a subset of $\Gamma \times \Gamma$. Since $\Gamma \times \Gamma$ is equipotent to Γ , C is equipotent to a subset of Γ . This implies that X is equipotent to a subset of Γ . Therefore, by the theorem of Schröder-Bernstein, Γ is equipotent to X . ■

1.3 Basic concepts from group theory and ring theory

For the definition of a vector space, one must know what groups and rings are. This section contains the very basics from group theory and ring theory. Definitions in this section can be found in (Cohn, 1974).

Definition 1.3.1. Let X be a set. A binary operation on X is a function $\cdot : X \times X \rightarrow X$. If $(a, b) \in X \times X$, the image of (a, b) under \cdot is denoted by ab .

A set X together with a binary operation \cdot is called a groupoid and is denoted by (X, \cdot) or just X . Binary operations are also denoted by $+$. Depending on the notation, one has multiplicatively written groupoids and additively written groupoids. The concept of a binary operation is a very general concept, on a finite set X , there exists $|X|^{|X|^2}$ binary operations. Therefore one is interested in properties a binary operation may have.

Definition 1.3.2. An identity element of a groupoid (X, \cdot) is an element $e \in X$ such that $ex = x = xe$ for all $x \in X$.

Proposition 1.3.1. *Every groupoid (X, \cdot) has at most one identity element.*

Proof. Suppose $e, e' \in X$ are identity elements. Then $ee' = e'$ because e is an identity element. Also, $ee' = e$ because e' is an identity element. Therefore, $e = e'$. ■

One of the most important properties of a binary operation may have is associativity.

Definition 1.3.3. Let (X, \cdot) be a groupoid. \cdot is associative if and only if $a(bc) = (ab)c$ for all $a, b, c \in X$.

In a groupoid, a product of more than two elements is not defined. To make it meaningful, one must insert brackets. The brackets constitute an instruction how to compute the product. If a binary operation is associative, then a product of more than two factors does not depend on the manner of computation.

Definition 1.3.4. A semigroup is an associative groupoid. A monoid is a semigroup with an identity element.

Definition 1.3.5. Let (X, \cdot) be a monoid and $a \in X$. An inverse of $a \in X$ is an element $b \in X$ such that $ab = e = ba$.

Proposition 1.3.2. *Every element of a monoid (X, \cdot) has at most one inverse.*

Proof. Let $a \in X$ and suppose $b, b' \in X$ are inverses of a . Then $b = be = b(ab') = (ba)b' = eb' = b'$. ■

Definition 1.3.6. A group G is a set with a binary operation \cdot such that:

- i) \cdot is associative
- ii) there exists $e \in G$ such that $ae = a$ for all $a \in G$
- iii) for each $a \in G$ there exists $b \in G$ such that $ab = e$.

In most books of algebra, ii) and iii) above are stated as follow:

- ii) G has an identity element.
- iii) Every element of G has an inverse.

This is unnecessary because of the next proposition.

Proposition 1.3.3. *Every group has exactly one identity element and every element of the group has exactly one inverse.*

Proof. Let G be a group. The uniqueness property follow from propositions 1.3.1 and 1.3.2. Let $a \in G$. By iii), there exists $b \in G$ such that $ab = e$. Again by iii), there exists $c \in G$ such that $bc = e$. Then $ec = (ab)c = a(bc) = ae = a$. Multiplying this relation by e from the left, one gets $ea = e(ec) = (ee)c = ec = a$. Therefore, e is an identity element. Since $ec = a$, $c = a$. Therefore b is an inverse of a . ■

If M be a monoid and $a \in M$, then a is called invertible or a unit if the inverse of a exists. The inverse of a is denoted by a^{-1} . The invertible elements of M form a group $U(M)$ called the units group of M .

Let G be a group. A subset of G is called a complex.

In the following definition, \subset means subset of.

Definition 1.3.7. Let $X, Y \subset G$.

$$XY := \{xy \mid x \in X, y \in Y\}$$

is called the complex product of X and Y .

The complex product is associative.

Proposition 1.3.4. *Let G be a group and $S \subset G$. Then (S, \cdot) is a group if and only if i) $xy \in S$ for all $x, y \in S$, ii) $x^{-1} \in S$ for all $x \in S$ and iii) $S \neq \emptyset$.*

Proof. Suppose S is a group. Then $ab \in S$ for all $a, b \in S$. Since S is a group, there exists an identity $e' \in S$. Multiplying $e'e' = e'$ by e'^{-1} results in $e' = e$. Let $x \in S$. There exists $y \in S$ such that $xy = e'$. Since $e' = e$, $xy = e$. Hence $y = x^{-1}$. Therefore $x^{-1} \in S$. Since S is a group, $S \neq \emptyset$. To prove the converse inclusion, because of i), \cdot is a binary operation on S and it is associative. Since $S \neq \emptyset$, there exists $x \in S$. Since $x^{-1} \in S$, $xx^{-1} = e \in S$. Every element of S has an inverse. Therefore, (S, \cdot) is a group. ■

Definition 1.3.8. Let G be a group. A subgroup of G is a subset S of G such that:

- i) $xy \in S$ for all $x, y \in S$
- ii) $x^{-1} \in S$ for all $x \in S$
- iii) $S \neq \emptyset$

If S is a subgroup of G , one writes $S \leq G$.

Definition 1.3.9. Let G be a group, $S \leq G$ and $x \in G$.

$$xS := \{x\}S = \{xs | s \in S\}, \quad Sx := S\{x\} = \{sx | s \in S\}$$

xS is called the left coset of S represented by x and Sx is called the right coset of S represented by x .

It is simple to prove that the left cosets of G and the right cosets of G form partitions of G . In general these two partitions are not equal. They are equal if and only if the subgroup is a normal subgroup.

Definition 1.3.10. Let G be a group. A normal subgroup of G is a subgroup $N \leq G$ such that $xN = Nx$ for all $x \in G$.

If N is a normal subgroup of G one writes $N \trianglelefteq G$.

A group (G, \cdot) is called abelian if \cdot is commutative. That is $ab = ba$ for all $a, b \in G$. It is evident that every subgroup of an abelian group is normal.

In group theory, one normally uses exponential notations. If x is an element of a group G and α is a function with domain G , then x^α denotes the image of x with respect to α .

Definition 1.3.11. Let G and H be groups. A homomorphism from G to H is a function $\alpha : G \rightarrow H$ such that $(xy)^\alpha = x^\alpha y^\alpha$ for all $x, y \in G$.

An isomorphism from G to H is a bijective homomorphism from G to H .

$$G \cong H \text{ (} G \text{ isomorphic to } H) \iff \text{there exists an isomorphism from } G \text{ to } H.$$

The basic facts about isomorphisms are the following.

Remarks.

1. Let $\alpha : G \rightarrow H$ be an isomorphism. Then $\alpha^{-1} : H \rightarrow G$ is an isomorphism.

Proof. $\alpha^{-1} : H \rightarrow G$ is bijective. Let $x, y \in H$.

$$(xy)^{\alpha^{-1}} = ((x^{\alpha^{-1}})^{\alpha} (y^{\alpha^{-1}})^{\alpha})^{\alpha^{-1}} = ((x^{\alpha^{-1}} y^{\alpha^{-1}})^{\alpha})^{\alpha^{-1}} = x^{\alpha^{-1}} y^{\alpha^{-1}}. \quad \blacksquare$$

2. Let G, H, K be groups, $\alpha : G \rightarrow H$ and $\beta : H \rightarrow K$ be isomorphisms. Then $\alpha\beta : G \rightarrow K$ is an isomorphism.

Proof. $\alpha\beta$ is bijective since it is a composition of two bijective functions. $(xy)^{\alpha\beta} = ((xy)^\alpha)^\beta = (x^\alpha y^\alpha)^\beta = (x^\alpha)^\beta (y^\alpha)^\beta = x^{\alpha\beta} y^{\alpha\beta}$ for all $x, y \in G$. ■

3. Let G be a group. $id : G \rightarrow G$ is an isomorphism.

What has been stated above means that “isomorphic” is an equivalence relation.

Definition 1.3.12. Let $\alpha : G \rightarrow H$ be a homomorphism.

$$\text{im } \alpha := G^\alpha, \text{ ker } \alpha := \{x \in G \mid x^\alpha = e\}.$$

$\text{im } \alpha$ is called the image of α and $\text{ker } \alpha$ is called the kernel of α .

Remarks.

1. $\text{im } \alpha \leq H$.
2. $\{e\}^{\alpha^{-1}} = \{x \in G \mid x^\alpha \in \{e\}\} = \{x \in G \mid x^\alpha = e\} = \text{ker } \alpha$.
3. α injective $\iff \text{ker } \alpha = \{e\}$.

Proof. (\Rightarrow) Suppose α is injective. Let $x \in \ker \alpha$. Then $x^\alpha = e = e^\alpha$. Since α is injective, $x = e$. Therefore, $\ker \alpha = \{e\}$.

(\Leftarrow) Suppose $\ker \alpha = \{e\}$. Let $x, y \in G$ such that $x^\alpha = y^\alpha$. Then $e = (x^\alpha)^{-1}x^\alpha = (x^\alpha)^{-1}y^\alpha = (x^{-1}y)^\alpha$. Since $x^{-1}y \in \ker \alpha = \{e\}$, $x^{-1}y = e$. Hence $x = y$. ■

4. $\ker \alpha \trianglelefteq G$

Proof. Let $K := \ker \alpha$ and $x \in G$. Let $a \in K$. $(x^{-1}ax)^\alpha = (x^{-1})^\alpha a^\alpha x^\alpha = (x^{-1})^\alpha e x^\alpha = (x^{-1})^\alpha x^\alpha = e$. Hence $x^{-1}ax \in K$. This implies $x^{-1}Kx \subset K$. Hence $Kx \subset xK$. This is also true for x^{-1} : $Kx^{-1} \subset x^{-1}K$. Multiplying by x from both sides: $xK \subset Kx$. Therefore $Kx = xK$. ■

Definition 1.3.13. Let G be a group and suppose $N \trianglelefteq G$.

$$G/N := \{xN \mid x \in G\}.$$

G/N with the complex product as the binary operation is a group called the factor group or quotient group of G modulo N .

Let G be a group and $X \subset G$. The intersection of all subgroups of G which contain X is also a subgroup of G . This subgroup is denoted $\langle X \rangle$ and is called the subgroup generated by X . $\langle X \rangle$ is the smallest subgroup of G containing X .

Definition 1.3.14. A cyclic group is a group that can be generated by a single element.

Let G be a group and $x \in G$.

$$\langle x \rangle = \langle \{x\} \rangle = \bigcap_{x \in S \leq G} S.$$

Remarks.

1. Let $x \in \mathbb{Z}$. Then $\langle x \rangle = \{kx \mid k \in \mathbb{Z}\}$.

Proof. Since $x \in \langle x \rangle$ and $\langle x \rangle$ is a subgroup, $kx \in \langle x \rangle$ for all $k \in \mathbb{Z}$. Therefore, $\{kx \mid k \in \mathbb{Z}\} \subset \langle x \rangle$. Let $k, l \in \mathbb{Z}$. Then $kx - lx = (k - l)x$ and $-kx = (-k)x$. Therefore, $\{kx \mid k \in \mathbb{Z}\}$ is a subgroup of \mathbb{Z} which contains x . Now $\langle x \rangle \subset \{kx \mid k \in \mathbb{Z}\}$ and hence the equality follows. ■

2. Every cyclic group is abelian.

Proof. Suppose G is cyclic. Then there exists $x \in G$ such that $G = \langle x \rangle$. Let $a, b \in G$. Then there exists $r, s \in \mathbb{Z}$ such that $a = x^r$, $b = x^s$. $ab = x^r x^s = x^{r+s} = x^{s+r} = x^s x^r = ba$. ■

If G is a group, then $|G|$ is called the order of G . If G is infinite, one writes $|G| = \infty$ and one says that G is of infinity order.

Definition 1.3.15. Let G be a group and $x \in G$.

$$o(x) := |\langle x \rangle|$$

is called the order of x .

The order of an element of a group is a natural number or infinity.

Definition 1.3.16. Let G be a group.

$$\text{Aut } G := \{\alpha : G \rightarrow G \mid \alpha \text{ a bijective homomorphism}\}.$$

The elements of $\text{Aut } G$ are called automorphisms of G .

Definition 1.3.17. Let G be a group.

$$Z(G) := \{a \in G \mid ax = xa \text{ for all } x \in G\}.$$

$Z(G)$ is called the centre of G .

The centre of G is a normal subgroup of G .

Now a brief compilation of basic concepts about rings follows.

Definition 1.3.18. A ring is a set R with two binary operations $+$, \cdot which satisfy the following conditions:

1. $(R, +)$ is an abelian group
2. (R, \cdot) is a monoid
3. $a(b + c) = ab + ac$ and $(a + b)c = ac + bc$ for all $a, b, c \in R$.

Remarks.

- (1) If \cdot is commutative, R is called commutative.
- (2) $(R, +)$ is called the additive group of R and denoted by R^+ . The zero of $(R, +)$ is called the zero of R and denoted by 0 .
- (3) The identity of (R, \cdot) is called the identity of R and denoted by 1 .
- (4) If $a \in R$, then $x \in R \mapsto ax \in R$ and $x \in R \mapsto xa \in R$ are endomorphisms of R^+ . In particular, $a0 = 0 = 0a$, $a(-b) = -ab = (-a)b$.
- (5) Assume $1 = 0$. Let $x \in R$. $x = 1x = 0x = 0$. Hence $R = \{0\}$. $R = \{0\}$ is called a zero ring.
- (6) Let $a, b \in R$ such that $ab = ba$. Then $(a + b)^n = \sum_{k=0}^n \binom{n}{k} a^k b^{n-k}$ for all $n \in \mathbb{N}_0$.

Definition 1.3.19. The units of (R, \cdot) are called the units of the ring R . $R^\times = \{a \in R \mid a \text{ is a unit}\}$ is called the units group of the ring.

Simple examples are, units group of \mathbb{R} is $\mathbb{R} - \{0\}$, units group of \mathbb{Z} is $\{-1, 1\}$.

Definition 1.3.20. Let $a \in R$.

a is a zero divisor of $R \iff a \neq 0$ and there exists $b \in R - \{0\}$ such that $ab = 0$ or $ba = 0$.

Definition 1.3.21. R is an entire ring $\iff \begin{cases} (1) R \neq \{0\} \\ (2) R \text{ has no zero divisors.} \end{cases}$

Definition 1.3.22.

$$\text{char}R = \begin{cases} 0 & \text{if } o_{R^+}(1) = \infty \\ o_{R^+}(1) & \text{if } o_{R^+}(1) < \infty \end{cases}$$

$\text{Char}R$ is called the characteristic of R .

Definition 1.3.23. Let A be an additively written abelian group.

$$\text{End}A := \{\alpha : A \rightarrow A \mid \alpha \text{ an endomorphism}\}.$$

If $\alpha, \beta \in \text{End}A$, then $\alpha + \beta, \alpha\beta : A \rightarrow A$ are defined by $x^{\alpha+\beta} = x^\alpha + x^\beta$, $x^{\alpha\beta} = (x^\alpha)^\beta$. $\text{End}A$ with these operations is a ring called the endomorphism ring of an abelian group A .

Definition 1.3.24. A subring of R is a subset $S \subset R$ such that

1. $x - y \in S$ for all $x, y \in S$,

2. $xy \in S$ for all $x, y \in S$,

3. $1 \in S$.

Remark. *Every subring of a ring is a ring.*

Proof. Let R be a ring and S a subring of R . Since $xy \in S$ for all $x, y \in S$, \cdot is a binary operation on S . Since $1 \in S$, $0 = 1 - 1 \in S$. If $y \in S$, $-y = 0 - y \in S$. If $x, y \in S$, $x + y = x - (-y) \in S$. Addition is a binary operation on S . Since S is a subgroup of R^+ , $(S, +)$ is an abelian group. (S, \cdot) is a monoid. The two distributive laws are true. Therefore S is a ring. ■

Definition 1.3.25. Let R, S be rings. A homomorphism is a function $f : R \rightarrow S$ which has the following properties:

1. $f(x + y) = f(x) + f(y)$ and $f(xy) = f(x)f(y)$ for all $x, y \in R$,

2. $f(1) = 1$.

Remarks. Suppose $f : R \rightarrow S$ is a homomorphism.

(1) Since $f(x + y) = f(x) + f(y)$ for all $x, y \in R$, $f : R^+ \rightarrow S^+$ is a homomorphism. In particular, $f(0) = 0$, $f(-x) = -f(x)$ and $f(kx) = kf(x)$ ($x \in R$, $k \in \mathbb{Z}$).

(2) Let $a \in R^\times$. Then $aa^{-1} = 1$ and $a^{-1}a = 1$.

$1 = f(1) = f(aa^{-1}) = f(a)f(a^{-1})$ and $1 = f(1) = f(a^{-1}a) = f(a^{-1})f(a)$. Hence $f(a) \in S^\times$. $x \in R^\times \mapsto f(x) \in S^\times$ is a homomorphism. In particular $f(a^{-1}) = f(a)^{-1}$ for all $a \in R^\times$.

Proposition 1.3.5. *Let R, S be rings and $f : R \rightarrow S$ a ring homomorphism.*

(a) *If $T \subset R$ is a subring, then $f(T) \subset S$ is a subring.*

(b) *If $V \subset S$ is a subring, then $f^{-1}(V) \subset R$ is a subring.*

Proof. (a) Suppose T is a subring of R . Let $y, y' \in f(T)$. Then there exist $x, x' \in T$ such that $f(x) = y$ and $f(x') = y'$. Since T is a subring, $x - x' \in T$ and $f(x - x') = f(x) - f(x') = y - y'$. Hence, $y - y' \in f(T)$. Since T is a subring, $xx' \in T$. $f(xx') = f(x)f(x') = yy'$. Hence $yy' \in f(T)$. Since $1 \in T$, $1 = f(1) \in f(T)$.

(b) Suppose V is a subring of S . Let $x, y \in f^{-1}(V)$. Since $f(x), f(y) \in V$ and V is a subring, $f(x - y) = f(x) - f(y) \in V$. Hence $x - y \in f^{-1}(V)$. This means that $f(xy) = f(x)f(y) \in V$. Hence $xy \in f^{-1}(V)$. Since $f(1) = 1 \in V$, $1 \in f^{-1}(V)$. ■

Definition 1.3.26. Let $f : R \rightarrow S$ be a ring homomorphism.

$$\text{im } f := f(R), \quad \ker f := \{x \in R \mid f(x) = 0\}.$$

$\text{im } f$ is called the image of f and $\ker f$ is called the kernel of f .

Remarks

1. $\text{im } f$ is a subring of S .
2. $\ker f$ is the kernel of $f : R^+ \rightarrow S^+$. Hence $\ker f$ is a subgroup of R^+ . If $\ker f$ is a subring then S is a zero ring. The converse is also true.

Definition 1.3.27. An ideal of R is a subset $I \subset R$ such that

1. $x + y \in I$ for all $x, y \in I$
2. If $x \in I$ and $r \in R$, then $rx, xr \in I$
3. $I \neq \emptyset$

Remark.

$$f \text{ injective} \iff \ker f = \{0\}.$$

The proof of this remark is done similarly as in group theory.

Proposition 1.3.6. Let R be a ring and $(I_\lambda)_\Lambda$ be a non empty family of ideals of R . Then

$\bigcap_{\Lambda} I_\lambda$ is an ideal.

Proof. Let $x, y \in \bigcap_{\Lambda} I_\lambda$. Since $x, y \in I_\lambda$ and I_λ is an ideal, $x + y \in I_\lambda$ for all $\lambda \in \Lambda$. Hence, $x + y \in \bigcap_{\Lambda} I_\lambda$. Let $x \in \bigcap_{\Lambda} I_\lambda$ and $r \in R$. Since $x \in I_\lambda$ and I_λ is an ideal, $rx, xr \in I_\lambda$ for all $\lambda \in \Lambda$. Hence $rx, xr \in \bigcap_{\Lambda} I_\lambda$. Since $0 \in I_\lambda$ for all $\lambda \in \Lambda$, $0 \in \bigcap_{\Lambda} I_\lambda$. ■

Definition 1.3.28. Let $X \subset R$.

$$(X) := \bigcap_{\substack{X \subset I \subset R \\ I \text{ ideal of } R}} I.$$

(X) is called the ideal generated by X .

Definition 1.3.29. Let $a \in R$. The ideal generated by $\{a\}$ is called the principal ideal and denoted by (a) .

Definition 1.3.30.

$$Z(R) := \{z \in R \mid xz = zx \text{ for all } x \in R\}.$$

$Z(R)$ is called the centre of R .

Remarks.

1. $Z(R)$ is sometimes denoted by $C(R)$.
2. $Z(R)$ is a subring of R .

Proof. Let $z, w \in Z(R)$. $x(z - w) = xz - xw = zx - wx = (z - w)x$ for all $x \in R$. Hence $z - w \in Z(R)$. $x(zw) = (xz)w = (zx)w = z(xw) = z(wx) = (zw)x$ for all $x \in R$. Hence $zw \in Z(R)$. Since $x1 = x = 1x$ for all $x \in R$, $1 \in Z(R)$. Therefore $Z(R)$ is a subring of R . ■

Proposition 1.3.7. Let $a \in Z(R)$. Then $(a) = \{ra \mid r \in R\}$

Proof. Take

$$I := \{ra \mid r \in R\}.$$

Let $x, y \in I$. Then there exist $r, s \in R$ such that $x = ra$, $y = sa$. $x + y = ra + sa = (r + s)a \in I$. Let $t \in R$. $tx = t(ra) = (tr)a \in I$, $xt = (ra)t = r(at) = r(ta) = (rt)a \in I$. $a = 1a \in I$. Thus I is an ideal of R . Since $a \in I$, $(a) \subset I$. Let $r \in R$. Since (a) is an ideal containing a , $ra \in (a)$. Hence $I \subset (a)$. Therefore, $I = (a)$. ■

Definition 1.3.31. Let R be a ring and I an ideal of R . The cosets of I are called the residue classes of I . The set

$$R/I := \{a + I \mid a \in R\}$$

of all residue classes of I with the following operations

$$(a + I) + (b + I) := a + b + I$$

$$(a + I)(b + I) := ab + I \text{ where } a, b \in R$$

is called the residue class ring of R modulo I .

$\mathbb{Z}_n := \mathbb{Z}/(n)$ is called the residue class ring of \mathbb{Z} modulo n .

Proposition 1.3.8. Let $n \in \mathbb{N}$. Then $\mathbb{Z}_n = \{0 + (n), 1 + (n), \dots, n - 1 + (n)\}$. Furthermore, $|\mathbb{Z}_n| = n$.

Proof. $\{0 + (n), 1 + (n), \dots, n - 1 + (n)\} \subset \mathbb{Z}_n$ is clear. Let $k \in \mathbb{Z}$. Division with remainder: There exist $q, r \in \mathbb{Z}$ such that $0 \leq r < n$ and $k = qn + r$. $k + (n) = qn + r + (n) = r + (n)$. To prove that $|\mathbb{Z}_n| = n$, let $0 \leq k, l \leq n - 1$ and $k + (n) = l + (n)$. Then $k - l \in (n)$. Since $n|k - l$ and $|k - l| < n$, $k - l = 0$. Hence $k = l$. ■

Definition 1.3.32. A division ring is a ring R such that $R^\times = R - \{0\}$. A field is a commutative division ring.

Remark. Every division ring is an entire ring.

Proof. Let R be a division ring. Since $R^\times = R - \{0\}$, $R \neq \{0\}$. Let $a, b \in R$ such that $ab = 0$ and $a \neq 0$. $0 = a^{-1}(ab) = (a^{-1}a)b = b$. Similarly, if $ab = 0$ and $b \neq 0$, then $a = 0$. Therefore R has no zero divisors. ■

Remark. If n is a prime then \mathbb{Z}_n is a field.

Proof. Suppose n is a prime. Let $1 \leq k \leq n - 1$. Since $\gcd\{k, n\} = 1$, there exist $r, s \in \mathbb{Z}$ such that $rk + sn = 1$. $1 + (n) = rk + sn + (n) = rk + (n) + sn + (n) = rk + (n) = (r + (n))(k + (n))$. Therefore, $k + (n)$ is a unit. ■

Definition 1.3.33. Let $p \in \mathbb{N}$ be a prime.

$$GF(p) := \mathbb{Z}_p$$

is called Galois field of order p .

1.4 Basic concepts from Linear Algebra

The starting point is the definition of a vector space.

Definition 1.4.1. Let K be a division ring and V an additively written abelian group. V is called a K -left vector space if an operation $(\lambda, x) \in K \times V \rightarrow \lambda x \in V$ is given which satisfies the following conditions:

- | | |
|--|---|
| (i) $\lambda(x + y) = \lambda x + \lambda y$ | (ii) $(\lambda + \mu)x = \lambda x + \mu x$ |
| (iii) $(\lambda\mu)x = \lambda(\mu x)$ | (iv) $1x = x$ |

for all $\lambda, \mu \in K$ and $x, y \in V$.

The fact that the operation $(\lambda, x) \in K \times V \rightarrow \lambda x \in V$ is given which satisfies (i)-(iv) is expressed saying that the division ring K operates on the abelian group V . This is a classical description of a vector space due to E. Noether.

The zero of the abelian group $(V, +)$ is called the zero vector of V and denoted by 0 .

Definition 1.4.2. Let V be a K -left vector space and $\lambda \in K$. Let $h(\lambda) : V \rightarrow V$ be defined by

$$h(\lambda)(x) := \lambda x.$$

$h(\lambda)$ is called the homothety with factor λ .

h_λ is sometimes used for $h(\lambda)$.

Theorem 1.4.1. (1) $h : K \rightarrow \text{End}(V, +)$ is a ring homomorphism. Furthermore, $V \neq \{0\}$ if and only if h is injective.

(2) $C(h(K))$ is the ring of the endomorphisms of the vector space V .

Proof. (1) Let $x, y \in V$ and $\lambda \in K$. $h(\lambda)(x + y) = \lambda x + \lambda y = h(\lambda)(x) + h(\lambda)(y)$. Hence $h(\lambda) \in \text{End}(V, +)$. Let $\lambda, \mu \in K$. $h(\lambda + \mu)(x) = (\lambda + \mu)x = \lambda x + \mu x = h(\lambda)(x) + h(\mu)(x) = (h(\lambda) + h(\mu))(x)$ for all $x \in V$. Hence $h(\lambda + \mu) = h(\lambda) + h(\mu)$. $h(\lambda\mu)(x) = (\lambda\mu)x = \lambda(\mu x) = h(\lambda)(h(\mu)x) = (h(\lambda) \circ h(\mu))(x)$ for all $x \in V$. Hence $h(\lambda\mu) = h(\lambda) \circ h(\mu)$. $h(1)(x) = 1x = x = \text{id}_V(x)$ for all $x \in V$. Hence $h(1) = \text{id}_V$. Therefore, $h : K \rightarrow \text{End}(V, +)$ is a ring homomorphism.

Suppose h is injective. Since K is a division ring, $0, 1 \in K$ and $0 \neq 1$. Since $h(0) = 0$ and $h(1) = \text{id}_V$, $V \neq \{0\}$. Suppose $V \neq \{0\}$. Since h is a ring homomorphism, $\ker h$ is an ideal of K . Since $h(1) = \text{id}_V$, $1 \notin \ker h$. Therefore $\ker h \neq K$. Since (0) and K are the only ideals of K , $\ker h = (0)$. Therefore, h is injective.

(2) Let $f \in C(h(K))$. Since $f \in \text{End}(V, +)$, $f(x + y) = f(x) + f(y)$ for all $x, y \in V$. Let $\lambda \in K$ and $x \in V$. $f(\lambda x) = f(h(\lambda)(x)) = (f \circ h(\lambda))(x) = (h(\lambda) \circ f)(x) = h(\lambda)(f(x)) = \lambda f(x)$. Therefore f is linear. Now suppose $f : V \rightarrow V$ is linear. Since $f(x + y) = f(x) + f(y)$ for all $x, y \in V$, $f \in \text{End}(V, +)$. Let $\lambda \in K$. $(f \circ h(\lambda))(x) = f(h(\lambda)(x)) = f(\lambda x) = \lambda f(x) = h(\lambda)(f(x)) = (h(\lambda) \circ f)(x)$ for all $x \in V$. Hence $f \circ h(\lambda) = h(\lambda) \circ f$. Therefore $f \in C(h(K))$. ■

A non-zero vector space is the same as a non-zero abelian group together with a subring of the endomorphism ring of the abelian group which is a division ring. This is sometimes used as a description of a vector space.

Theorem 1.4.2. (1) If $\lambda \in K$ and $x \in V$, then

$$\lambda x = 0 \iff \lambda = 0 \text{ or } x = 0$$

(2) Let $\lambda \in K$ and $x \in V$. Then $(-\lambda)x = -\lambda x = \lambda(-x)$.

Proof. (1) \Leftarrow Suppose $\lambda = 0$. Let $a \in V$. If $\alpha, \beta \in K$, $(\alpha + \beta)a = \alpha a + \beta a$. Hence $\alpha \in K \rightarrow \alpha a \in V$ is a homomorphism from K^+ to $(V, +)$. Therefore $0a = 0$. Let $\mu \in K$. If $a, b \in V$, $\mu(a + b) = \mu a + \mu b$. Hence $a \in V \rightarrow \mu a \in V$ is an endomorphism. Therefore, $\mu 0 = 0$.

\Rightarrow Suppose $\lambda x = 0$ and $\lambda \neq 0$. Since K is a division ring, λ^{-1} exists. $0 = \lambda^{-1}(\lambda x) = (\lambda^{-1}\lambda)x = 1x = x$.

(2) $\lambda x + (-\lambda)x = (\lambda + (-\lambda))x = 0x = 0$. Hence $(-\lambda)x = -\lambda x$.

$\lambda x + \lambda(-x) = \lambda(x + (-x)) = \lambda 0 = 0$. Hence $\lambda(-x) = -\lambda x$. ■

Definition 1.4.3. Let V be a K -left vector space. A subspace S of V is a subset $S \subset V$ such that

- (1) $x + y \in S$ for all $x, y \in S$,
- (2) $\lambda x \in S$ for all $\lambda \in K, x \in S$,
- (3) $S \neq \emptyset$.

If S is a subspace of V , one writes $S \leq V$. $S < V$ means that S is a proper subspace of V .

Remark. Every subspace of a K -left vector space is a K -left vector space.

Proof. Suppose $S \leq V$. Because of (1), addition is a binary operation on S . This operation is associative and commutative. Since $S \neq \emptyset$, there exists $x \in S$. $0 = 0x \in S$ and $a + 0 = a$ for all $a \in S$. Let $a \in S$. Then $-a = (-1)a \in S$ and $a + (-a) = 0$. Therefore $(S, +)$ is an abelian group. $(\lambda, x) \in K \times S \rightarrow \lambda x \in S$ satisfies the four axioms of a vector space. ■

Theorem 1.4.3. Let Γ be a non-empty set of subspaces of V . Then $\bigcap_{S \in \Gamma} S$ is a subspace of V .

Proof. $\bigcap_{S \in \Gamma} S \subset V$ because $\Gamma \neq \emptyset$. Let $x, y \in \bigcap_{S \in \Gamma} S$. Then $x, y \in S$ for all $S \in \Gamma$. Since $S \leq V$, $x + y \in S$ for all $S \in \Gamma$. Hence $x + y \in \bigcap_{S \in \Gamma} S$. Let $\lambda \in K$ and $x \in \bigcap_{S \in \Gamma} S$. Since $x \in S$ and $S \leq V$, $\lambda x \in S$ for all $S \in \Gamma$. Hence $\lambda x \in \bigcap_{S \in \Gamma} S$. Since $0 \in S$ for all $S \in \Gamma$, $0 \in \bigcap_{S \in \Gamma} S$. Hence $\bigcap_{S \in \Gamma} S \neq \emptyset$. Therefore, $\bigcap_{S \in \Gamma} S \leq V$. ■

The following definition is as it is in (Greub, 1967).

Definition 1.4.4. Let V be a K -left vector space and $X \subset V$.

$$[X] := \bigcap_{X \subset S \leq V} S$$

$[X]$ is called the span of X or the subspace generated by X .

Remark. $[\emptyset] = \{0\}$

Proof. $\{0\}$ is a subspace and $\emptyset \subset \{0\}$. Hence $[\emptyset] \leq \{0\}$. Since $[\emptyset] \neq \emptyset$, $[\emptyset] = \{0\}$. ■

Let V be a vector space and $X, Y \subset V$. Then one defines $X + Y$ as follows

$$X + Y = \{x + y \mid x \in X, y \in Y\}.$$

If S, T are subspaces of V then $S + T$ is a subspace of V .

Theorem 1.4.4. Let $S, T \leq V$. Then $S + T \leq V$ and $S + T = [S \cup T]$.

Proof. Let $x, x' \in S + T$. Then there exist $s, s' \in S$, $t, t' \in T$ such that $x = s + t$, $x' = s' + t'$. $x + x' = (s + t) + (s' + t') = (s + s') + (t + t')$. Since $s + s' \in S$ and $t + t' \in T$, $x + x' \in S + T$. Let $\lambda \in K$. $\lambda x = \lambda s + \lambda t$. Since $\lambda s \in S$ and $\lambda t \in T$, $\lambda x \in S + T$. $S + T \neq \emptyset$ as $S, T \neq \emptyset$. Therefore, $S + T \leq V$. Since $S \subset S + T$ and $T \subset S + T$, $S \cup T \subset S + T$. Since $S + T$ is a subspace, $[S \cup T] \leq S + T$. It is clear that $S, T \leq [S \cup T]$. Since $[S \cup T]$ is a subspace, $S + T \leq [S \cup T]$. Therefore, $S + T = [S \cup T]$. ■

Remark. Let $R, S, T \leq V$. Then $R \cap S + R \cap T \leq R \cap (S + T)$.

Proof. Since $S \leq S + T$, $R \cap S \leq R \cap (S + T)$. Since $T \leq S + T$, $R \cap T \leq R \cap (S + T)$. Since $R \cap (S + T)$ is a subspace, $R \cap S + R \cap T \leq R \cap (S + T)$. ■

Theorem 1.4.5 (Dedekind's rule). Let $R, S, T \leq V$. If $S \leq R$, then

$$R \cap (S + T) = S + R \cap T.$$

Proof. Suppose $S \leq R$. Since $S \leq R$ and $S \leq S + T$, $S \leq R \cap (S + T)$. Since $R \cap T \leq R$ and $R \cap T \leq S + T$, $R \cap T \leq R \cap (S + T)$. Since $R \cap (S + T)$ is a subspace, $S + R \cap T \leq R \cap (S + T)$. Let $x \in R \cap (S + T)$. Since $x \in R$ and $s \in S \leq R$, $t \in R$. Since $t \in R \cap T$ and $x = s + t$, $x \in S + R \cap T$. Therefore, $R \cap (S + T) \leq S + R \cap T$. ■

Theorem 1.4.6. *Let $X \subset V$. If $X \neq \emptyset$, then*

$$[X] = \left\{ \sum_{i=1}^n \lambda_i x_i \mid \lambda_1, \dots, \lambda_n \in K, x_1, \dots, x_n \in X \right\}.$$

Proof. Let

$$S := \left\{ \sum_{i=1}^n \lambda_i x_i \mid \lambda_1, \dots, \lambda_n \in K, x_1, \dots, x_n \in X \right\}.$$

Now it will be shown that $[X] = S$.

Let $a, b \in S$. Then there exist $\lambda_1, \dots, \lambda_m, \mu_1, \dots, \mu_n \in K$ and $x_1, \dots, x_m, y_1, \dots, y_n \in X$ such that $a = \sum_{i=1}^m \lambda_i x_i$, $b = \sum_{i=1}^n \mu_i y_i$. $a + b = \sum_{i=1}^m \lambda_i x_i + \sum_{i=1}^n \mu_i y_i$ is a linear combination of elements of X . Therefore $a + b \in S$. Let $\alpha \in K$. $\alpha a = \alpha \sum_{i=1}^m \lambda_i x_i = \sum_{i=1}^m \alpha(\lambda_i x_i) = \sum_{i=1}^m (\alpha \lambda_i) x_i \in S$. $S \neq \emptyset$ as $X \neq \emptyset$. Hence $S \leq V$. Let $a \in S$. Then there exist $\lambda_1, \dots, \lambda_n \in K$, $x_1, \dots, x_n \in X$ such that $a = \sum_{i=1}^n \lambda_i x_i$. Since $x_1, \dots, x_n \in [X]$ and $[X] \leq V$, $a \in [X]$. Hence $S \leq [X]$. Since $X \subset S$ and $S \leq V$, $[X] \leq S$. Therefore $S = [X]$. ■

Definition 1.4.5. Let V be a K -left vector space. A spanning set of V is a subset $S \subset V$ such that $[S] = V$.

Lemma 1.4.7. *Let V be a K -left vector space.*

(a) $[X \cup Y] = [X] + [Y]$ for all $X, Y \subset V$.

(b) If Γ is a non-empty chain of subspaces of V , then $\bigcup_{U \in \Gamma} U$ is a subspace of V .

(c) If Γ is a non-empty chain of subsets of V , then $[\bigcup_{A \in \Gamma} A] = \bigcup_{A \in \Gamma} [A]$

Proof. (a) Since $X \subset X \cup Y$, $[X] \leq [X \cup Y]$. Similarly, $[Y] \leq [X \cup Y]$. Since $[X \cup Y]$ is a subspace, $[X] + [Y] \leq [X \cup Y]$. Since $X \subset [X] \leq [X] + [Y]$ and $Y \subset [Y] \leq [X] + [Y]$, $X \cup Y \subset [X] + [Y]$. Since $X \cup Y \subset [X] + [Y]$ and $[X] + [Y]$ is a subspace, $[X \cup Y] \leq [X] + [Y]$. Hence $[X \cup Y] = [X] + [Y]$.

(b) Let $x, y \in \bigcup_{U \in \Gamma} U$ and $\lambda \in K$. Then there exists $U_1, U_2 \in \Gamma$ such that $x \in U_1$ and $y \in U_2$. Since Γ is a chain, $U_1 \leq U_2$ or $U_2 \leq U_1$. Then $x, y \in U_1$ or $x, y \in U_2$. Since U_1 and U_2 are subspaces, $x + y \in U_1$ or $x + y \in U_2$. Since $U_1, U_2 \subset \bigcup_{U \in \Gamma} U$, $x + y \in \bigcup_{U \in \Gamma} U$. Since $x \in U_1$, $\lambda x \in U_1$.

Since $U_1 \subset \bigcup_{U \in \Gamma} U$, $\lambda x \in \bigcup_{U \in \Gamma} U$. $\bigcup_{U \in \Gamma} U \neq \emptyset$ and subspaces are non-empty.

(c) If $B \in \Gamma$, $B \subset \bigcup_{A \in \Gamma} A$. Hence $[B] \leq [\bigcup_{A \in \Gamma} A]$. Therefore $\bigcup_{A \in \Gamma} [A] \subset [\bigcup_{A \in \Gamma} A]$. Let $B, C \in \Gamma$. Since Γ is a chain, $B \subset C$ or $C \subset B$. Hence $[B] \leq [C]$ or $[C] \leq [B]$. Hence $([B])_{B \in \Gamma}$ is a chain of subspaces and nonempty. By b), $\bigcup_{B \in \Gamma} [B]$ is a subspace. Let $C \in \Gamma$. Then $C \subset [C] \subset \bigcup_{B \in \Gamma} [B]$. Hence $\bigcup_{C \in \Gamma} C \subset \bigcup_{B \in \Gamma} [B]$. Since $\bigcup_{B \in \Gamma} [B]$ is a subspace, $[\bigcup_{C \in \Gamma} C] \subset \bigcup_{B \in \Gamma} [B]$. Let $B \in \Gamma$. Since $B \subset \bigcup_{C \in \Gamma} C$, $[B] \leq [\bigcup_{C \in \Gamma} C]$. Therefore, $\bigcup_{B \in \Gamma} [B] \leq [\bigcup_{C \in \Gamma} C]$. ■

Definition 1.4.6. Let V be a K -left vector space and $X \subset V$. X is called linearly independent if and only if X has the following property: if $x_1, \dots, x_n \in X$ are distinct and $\lambda_1, \dots, \lambda_n \in K$ such that $\sum_{i=1}^n \lambda_i x_i = 0$, then $\lambda_1 = \lambda_2 = \dots = \lambda_n = 0$.

If X is not linearly independent then X is called linearly dependent.

Remarks.

(1) Suppose $X \subset V$ is linearly dependent. Then there exist distinct $x_1, \dots, x_n \in X$ and $\lambda_1, \dots, \lambda_n \in K$, not all zero such that $\sum_{i=1}^n \lambda_i x_i = 0$.

(2) \emptyset is linearly independent.

(3) Every set of vectors which contains 0 is linearly dependent.

(4) Every subset of a linearly independent set is linearly independent.

The definitions in this section are taken from (Greub, 1967).

Definition 1.4.7. Let V be a K -left vector space. A basis of V is a linearly independent spanning set.

One can show that a basis of a K -left vector space V is a minimal spanning set of V .

Definition 1.4.8. Let V be a K -left vector space.

V finitely generated $:\iff$ there exists a finite $S \subset V$ such that $[S] = V$.

Remark. Later on it will be shown that any two bases of a vector space are equipotent, therefore one can introduce the dimension of a vector space.

Definition 1.4.9. Let V be a K -left vector space and B a basis of V . Then $|B|$ is called the dimension of V .

$$\dim V := |B|.$$

Theorem 1.4.8. *Let B be a basis of V and take*

$$\Sigma(B) := \{[X] \mid X \subset B\}.$$

Then $(\mathcal{P}(B), \subset)$ and $(\Sigma(B), \leq)$ are ordered sets. $\varphi : \mathcal{P}(B) \rightarrow \Sigma(B)$ defined by $\varphi(X) := [X]$ is an isomorphism and $\varphi^{-1}(U) = U \cap B$ for all $U \in \Sigma(B)$. Further, if $X \subset B$, then $[X] \cap B = X$.

Proof. Let $X, Y \subset B$ such that $X \subset Y$. $\varphi(X) = [X] \leq [Y] = \varphi(Y)$. Let $X, Y \subset B$ such that $X \not\subset Y$. Then there exists $a \in X - Y$. Assume $a \in \varphi(Y) = [Y]$. Then there exist distinct $y_1, \dots, y_n \in Y$ and $\lambda_1, \dots, \lambda_n \in K - \{0\}$ such that $a = \sum_{k=1}^n \lambda_k y_k$. This implies $(-1)a + \sum_{k=1}^n \lambda_k y_k = 0$. Hence $\{a, y_1, \dots, y_n\}$ is linearly dependent. This set is a subset of B , therefore there is a contradiction. Hence $a \notin \varphi(Y)$. Therefore, $\varphi(X) \not\leq \varphi(Y)$. Now, $X \subset Y$ if and only if $\varphi(X) \leq \varphi(Y)$ for all $X, Y \subset B$. Suppose $X, X' \subset B$ such that $\varphi(X) = \varphi(X')$. Since $\varphi(X) \leq \varphi(X')$, $X \subset X'$. Since $\varphi(X') \leq \varphi(X)$, $X' \subset X$. Therefore $X = X'$. Thus φ is injective. By definition of $\Sigma(B)$, φ is surjective. Therefore, φ is bijective and hence $\varphi : \mathcal{P}(B) \rightarrow \Sigma(B)$ is an isomorphism. Let $U \in \Sigma(B)$. Since $U \cap B \subset B$, $\varphi(U \cap B) = [U \cap B] \leq U$. Since $U \cap B$ is a spanning set of U , $[U \cap B] = U$. Therefore $\varphi(U \cap B) = U$. Hence, $\varphi^{-1}(U) = U \cap B$. Suppose $X \subset B$. Then $X \subset [X] \cap B$ and $[X] \leq [[X] \cap B]$. Since $[X] \cap B \subset [X]$, $[[X] \cap B] \leq [X]$. Therefore, $[X] = [[X] \cap B]$. Since X and $[X] \cap B$ are bases of $[X]$ and $X \subset [X] \cap B$, $X = [X] \cap B$. ■

Theorem 1.4.9. *Let $(X_i)_I$ be a family in $\mathcal{P}(B)$. Then*

(1)

$$\varphi\left(\bigcup_I X_i\right) = \sum_I \varphi(X_i).$$

(2) If $I \neq \emptyset$,

$$\varphi\left(\bigcap_I X_i\right) = \bigcap_I \varphi(X_i).$$

Proof. (1) Let $j \in I$. Since $X_j \subset \bigcup_I X_i$, $\varphi(X_j) \subset \varphi\left(\bigcup_I X_i\right)$. Since $\varphi\left(\bigcup_I X_i\right)$ is a subspace which contains all terms of $\sum_I \varphi(X_i)$, $\sum_I \varphi(X_i) \leq \varphi\left(\bigcup_I X_i\right)$. Let $j \in I$. Since $X_j \subset \varphi(X_j)$ and $\varphi(X_j) \leq \sum_I \varphi(X_i)$, $X_j \subset \sum_I \varphi(X_i)$. Since this true for all $j \in I$, $\bigcup_I X_i \subset \sum_I \varphi(X_i)$. Therefore, $\varphi\left(\bigcup_I X_i\right) \leq \sum_I \varphi(X_i)$. Therefore, $\varphi\left(\bigcup_I X_i\right) = \sum_I \varphi(X_i)$.

(2) Let $j \in I$. Since $\bigcap_I X_i \subset X_j$, $\varphi\left(\bigcap_I X_i\right) \leq \varphi(X_j)$. Therefore, $\varphi\left(\bigcap_I X_i\right) \leq \bigcap_I \varphi(X_i)$. Let

$w \in \bigcap_I \varphi(X_i)$. Since $w \in \varphi(X_i)$, there exists $x_i \in X_i$ such that $\varphi(x_i) = w$ for all $i \in I$. Since φ is injective, $(X_i)_I$ is constant. Let $X := X_i (i \in I)$. Then $w = \varphi(X) \in \varphi(X_i)$ for all $i \in I$. Hence $w \in \varphi(\bigcap_I X_i)$ and this implies $\bigcap_I \varphi(X_i) \subset \varphi(\bigcap_I X_i)$. Therefore $\bigcap_I \varphi(X_i) = \varphi(\bigcap_I X_i)$ follows. ■

Theorem 1.4.10. *Let X be a non-empty subset of V and let $a \in [X]$. Then there exist $x_1, \dots, x_n \in X$ and $\lambda_1, \dots, \lambda_n \in K$ such that*

$$a = \sum_{i=1}^n \lambda_i x_i.$$

Let $1 \leq l \leq n$ such that $\lambda_l \neq 0$. Then

$$[(X - \{x_l\}) \cup \{a\}] = [X]$$

Proof. Since $X - \{x_l\} \subset [X]$ and $a \in [X]$, $[(X - \{x_l\}) \cup \{a\}] \leq [X]$. $a = \sum_{i=1}^n \lambda_i x_i = \sum_{i=1, i \neq l}^n \lambda_i x_i + \lambda_l x_l$.

Then $x_l = \lambda_l^{-1} a - \lambda_l^{-1} \sum_{i=1, i \neq l}^n \lambda_i x_i \in [(X - \{x_l\}) \cup \{a\}]$. Since $(X - \{x_l\}) \cup \{a\}$ contain X , $[X] \leq [(X - \{x_l\}) \cup \{a\}]$. Hence the proof is complete. ■

Definition 1.4.10. Let V be a K -left vector space and $R \leq V$. A complement of R in V is a subspace $S \leq V$ such that $R + S = V$ and $R \cap S = \{0\}$. If S is a complement of R in V , one writes $R \oplus S = V$.

In order books, the complement of a subspace is called a supplement.

Proposition 1.4.1. *Let V be a K -left vector space, $U \leq V$ and $W \leq V$ a complement of U in V . Then $W \cong V/U$. In particular, any two complements of U in V are isomorphic.*

Proof. Let $f : W \rightarrow V/U$ be defined by $f(x) := x + U$. Let $x, y \in W$ and $\lambda \in K$. Then $f(x + y) = x + y + U = (x + U) + (y + U) = f(x) + f(y)$, $f(\lambda x) = \lambda x + U = \lambda(x + U) = \lambda f(x)$. Hence f is linear. Let $x \in \ker f$. Then $f(x) = x + U$ and $f(x) = 0 = U$. Since $x + U = U$, $x \in U$. Since $x \in W$, $x \in U \cap W = \{0\}$. Hence $x = 0$. Since $\ker f = \{0\}$, f is injective. Let $y + U \in V/U$. Since $y \in V = U + W$, there exist $u \in U$, $w \in W$ such that $y = u + w$. $y + U = u + w + U = (u + U) + (w + U) = U + w + U = w + U = f(w)$. Hence f is surjective. Since f is linear and bijective, f is an isomorphism. ■

Theorem 1.4.11. *Let $R \leq V$, $X \subset V$ such that $R + [X] = V$. Then there exists $Y \subset X$ such that $[Y]$ is a complement of R in V .*

Proof. Let

$$\Omega := \{Z \subset X \mid R \cap [Z] = \{0\}\}$$

, ordered by inclusion. $\Omega \neq \emptyset$ as $\emptyset \in \Omega$. Let Γ be a non-empty chain in Ω . Take

$$W := \bigcup_{Z \in \Gamma} Z.$$

It is clear that $W \subset X$. If $W = \emptyset$, $R \cap [W] = R \cap \{0\} = \{0\}$. Hence $W \in \Omega$. Suppose $W \neq \emptyset$. Let $x \in R \cap [W]$. Since $x \in [W]$, there exist $w_1, \dots, w_n \in W$ and $\lambda_1, \dots, \lambda_n \in K$ such that $x = \sum_{i=1}^n \lambda_i w_i$. For each $1 \leq i \leq n$, there exists $Z_i \in \Gamma$ such that $w_i \in Z_i$. Since Γ is a chain and $\{Z_1, Z_2, \dots, Z_n\}$ is finite, $\{Z_1, Z_2, \dots, Z_n\}$ has a greatest element Z^* . Then $w_1, w_2, \dots, w_n \in Z^*$. Now $x \in R \cap [Z^*] = \{0\}$. Therefore, $R \cap [W] = \{0\}$. Hence $W \in \Omega$. Since $Z \subset W$ for all $Z \in \Gamma$, W is an upper bound of Γ . By Zorn's lemma, Ω has a maximal element Y . Since $Y \in \Omega$, $Y \subset X$ and $R \cap [Y] = \{0\}$. Let $x \in X - Y$ and form

$$Y' := Y \cup \{x\}.$$

Then $Y' \subset X$ and Y' strictly bigger than Y . But Y is a maximal element of Ω . Hence Y' is not an element of Ω . Thus $R \cap [Y'] \neq \{0\}$. Choose $z \in R \cap [Y'] - \{0\}$. This z is a linear combination of some of the vectors in Y' . Since $R \cap [Y] = \{0\}$, in this representation, x occurs. Hence there exist $y_1, \dots, y_n \in Y$, $\lambda_1, \dots, \lambda_n, \lambda \in K$ such that $z = \sum_{i=1}^n \lambda_i y_i + \lambda x$. If $\lambda = 0$, $z = 0$ which is a contradiction because $z \in R \cap [Y'] - \{0\}$. Therefore $\lambda \neq 0$. Now

$$x = -\lambda^{-1} \sum_{i=1}^n \lambda_i y_i + \lambda^{-1} z \in [Y] + R.$$

Since $X - Y \subset R + [Y]$, $R + [X] \subset R + [Y]$. Hence $R + [Y] = V$. Thus $[Y]$ is a complement of R in V . ■

Later on it will be shown that Theorem 1.4.11. is closely related to the Steinitz exchange theorem. To simplify the referring to Theorem 1.4.11., this theorem will be named. In the following its will be called the complement theorem. The proof of the following theorem will be skipped.

Theorem 1.4.12. *Let V be a K -left vector space and B a basis of V . Let $U \subset V$. Then*

$$U \in \Sigma(B) \iff U \text{ has only one complement in } \Sigma(B).$$

Definition 1.4.11. Let V be a K -left vector space. A hyperplane of V is a maximal proper subspace of V .

Theorem 1.4.13. Let V be a K -left vector space and $a \in V - \{0\}$. Then there exists a hyperplane of V such that $a \notin H$

Proof.

$$\Omega := \{U \leq V \mid a \notin U\}$$

, ordered by inclusion. $\Omega \neq \emptyset$ as $\{0\} \in \Omega$. Let $\Gamma \subset \Omega$ be a non empty chain and

$$W := \bigcup_{U \in \Gamma} U.$$

Since W is equal to the union of a collection of subspaces of V , $W \subset V$. Let $x, y \in W$. Then there exist $U_1, U_2 \in \Gamma$ such that $x \in U_1$ and $y \in U_2$. Since Γ is a chain, $U_1 \leq U_2$ or $U_2 \leq U_1$. Hence $x, y \in U_1$ or $x, y \in U_2$. Since U_1 and U_2 are subspaces, $x + y \in U_1$ or $x + y \in U_2$. Since $U_1, U_2 \subset W$, $x + y \in W$. Let $\lambda \in K$. Since $x \in U_1$, $\lambda x \in U_1$. Hence $\lambda x \in W$. $W \neq \emptyset$ because $\Gamma \neq \emptyset$ and subspaces are non empty. Therefore, $W \leq V$. Since $a \notin U$ for all $U \in \Gamma$, $a \notin W$. Therefore, $W \in \Omega$. Since $U \leq W$ for all $U \in \Omega$, W is an upper bound of Γ . By Zorn's lemma, Ω has maximal elements. Let $H \in \Omega$ be a maximal element. Since $a \notin H$, $H < V$. Let $L \leq V$ such that $H < L$. Since $L \notin \Omega$, $a \in L$. Then $H + [a] \leq L$. Assume $H + [a] \neq L$. Then there exists $b \in L - (H + [a])$. Since $H + [b] \notin \Omega$, $a \in H + [b]$. There exists $h \in H$ and $\lambda \in K$ such that $a = h + \lambda b$. If $\lambda = 0$, then $a = h \in H$ which is a contradiction. Therefore $\lambda \neq 0$. Thus $b = \lambda^{-1}(-h + a) \in H + [a]$ which is again a contradiction. Therefore $H + [a] = L$. Since H is a proper subspace of V , $H + [a] = V$. Therefore, H is a hyperplane. \blacksquare

Definition 1.4.12. A linear form of a K -left vector space V is a linear mapping from V to K .

Theorem 1.4.14. Let V be a K -left vector space, $H \leq V$ a hyperplane and $a \in V - H$. There exists exactly one linear form $\alpha : V \rightarrow K$ such $\alpha(a) = 1$ and $\ker \alpha = H$.

Proof. Existence. Since $a \notin H$, $H < H + [a] \leq V$. Since H is a hyperplane, $H + [a] = V$. Since $H \cap [a] = \{0\}$, $[a]$ is a complement of H in V . Let $x \in V$. Since $x \in H + [a]$, there exist $h \in H$ and $\lambda \in K$ such that $x = h + \lambda a$. Suppose $x = h' + \lambda' a$ where $h' \in H$ and $\lambda' \in K$. Since $h + \lambda a = h' + \lambda' a$, $h - h' = (\lambda' - \lambda)a$. Since $a \notin H$, $\lambda' - \lambda = 0$, i.e. $\lambda' = \lambda$. Then $h' = h$. Therefore $\alpha(x) := \lambda$ defines a function $\alpha : V \rightarrow K$. Let $x, x' \in V$. Then there exist $h, h' \in H$, $\lambda, \lambda' \in K$ such that $x = h + \lambda a$, $x' = h' + \lambda' a$. $x + x' = (h + \lambda a) + (h' + \lambda' a) = (h + h') + (\lambda + \lambda')a$. Hence

$\alpha(x + x') = \lambda + \lambda' = \alpha(x) + \alpha(x')$. Let $\mu \in K$. $\mu x = \mu(h + \lambda a) = \mu h + \mu(\lambda a) = \mu h + (\mu\lambda)a$. Therefore, $\alpha(\mu x) = \mu\lambda = \mu\alpha(x)$. Hence α is linear. Since $a = 0 + 1a$, $\alpha(a) = 1$. If $h \in H$, $h = h + 0a$. Hence $\alpha(h) = 0$. thus $H \leq \ker\alpha$. $\ker\alpha \neq V$ because $\alpha(a) = 1$. Since H is a hyperplane, $\ker\alpha = H$.

Uniqueness. Let $\beta : V \rightarrow K$ be a linear form such that $\beta(a) = 1$ and $\ker\beta = H$. Let $x \in V$. Then there exists $h \in H$ and $\lambda \in K$ such that $x = h + \lambda a$. $\beta(x) = \beta(h + \lambda a) = \beta(h) + \lambda\beta(a) = 0 + \lambda \cdot 1 = \lambda = \alpha(x)$. Therefore, $\beta = \alpha$. ■

Remarks.

Let V be a K -left vector space and B a basis of V and $b \in B$.

(1) Then $[B - \{b}]$ is a hyperplane.

Proof. Since B is a minimal spanning set of V , $[B - \{b}] < V$. Since $[B - \{b}] + [b] = [(B - \{b\}) \cup \{b\}] = [B] = V$, one obtains $W = W \cap V = W \cap ([B - \{b\}] + [b]) = [B - \{b\}] + W \cap [b]$ for a subspace $W \leq$ which satisfy $[- \{b\}] < W \leq V$. Since $W \cap [b] \neq \{0\}$, $[b] \leq W$. Therefore $W = V$. Hence $[B - \{b}]$ is a hyperplane. ■

(2) Since $[B - \{b}]$ is a hyperplane and $b \notin [B - \{b}]$, there exists a linear form $\epsilon_b : V \rightarrow K$ such that $\epsilon_b(b) = 1$ and $\ker \epsilon_b = [B - \{b}]$. $(\epsilon_b)_B$ is the family of the coordinate forms of V with respect to B .

(3) Let $x \in V - \{0\}$. Since B is a basis of V , there exist distinct $b_1, \dots, b_n \in B$ and $\lambda_1, \dots, \lambda_n \in K$ such that $x = \sum_{i=1}^n \lambda_i b_i$. Let $1 \leq j \leq n$. $\epsilon_{b_j}(x) = \epsilon_{b_j}(\sum_{i=1}^n \lambda_i b_i) = \sum_{i=1}^n \lambda_i \epsilon_{b_j}(b_i) = \lambda_j \epsilon_{b_j}(b_j) = \lambda_j$. Hence $x = \sum_{i=1}^n \epsilon_{b_i}(x) b_i$. If $b \in B - \{b_1, \dots, b_n\}$, $\epsilon_b(x) = 0$.

Only finitely many terms of $(\epsilon_b(x))_B$ are non-zero. If $x \in V$, one writes $x = \sum_{b \in B} \epsilon_b(x) b$.

Chapter 2

Review of relevant literature and methodology used in this research

2.1 Review of relevant literature

2.1.1 The Steinitz exchange theorem

Here are versions of the Steinitz exchange theorem found in the literature. (Cohn, 1974) states the exchange theorem of Steinitz as follow:

Let x_1, x_2, \dots, x_r be a linearly independent set of elements of a vector space V and let Y be a spanning set consisting of s elements. Then $r \leq s$ and we can find a spanning set of the form $\{x_1, x_2, \dots, x_r, y_{r+1}, \dots, y_s\}$ where $y_i \in Y$ for $i = r + 1, \dots, s$.

Twenty nine years later, the same author (Cohn, 2003) stated the Steinitz exchange theorem as follows:

If $\{v_1, \dots, v_m\}$ is a set of m linearly independent vectors in a vector space V and $\{w_1, \dots, w_n\}$ spans V , then $m \leq n$ and possibly after reordering the w_i , the set $\{v_1, \dots, v_m, w_{m+1}, \dots, w_n\}$ spans V .

Another version is in (Oeljeklaus & Remmert, 1974) which is stated in German as follows:

Austauschsatz (E. Steinitz) Es seien I und J disjunkte endliche Mengen und $(v_i)_{i \in I}, (v_j)_{j \in J}$ Familien in V mit folgenden Eigenschaften:

1) Die Familie $(v_i)_{i \in I}$ erzeugt V ,

2) Die Familie $(v_j)_{j \in J}$ ist frei.

Dann gibt es eine Teilmenge I' von I , so daß die 'Vereinigungsfamilie' $(v_k)_{k \in I' \cup J}$ eine Basis von V ist.

This is translated to English as follow:

Let I and J be disjoint finite sets and $(v_i)_{i \in I}$, $(v_j)_{j \in J}$ families in V which satisfy the following conditions:

1) The Family $(v_i)_{i \in I}$ generates V

2) The Family $(v_j)_{j \in J}$ is Linearly independent.

Then there exists a subset I' of I such that $(v_k)_{k \in I' \cup J}$ is a basis of V .

In this version $|J| \leq |I|$ is missing.

As in (Cohn, 2003) and (Cohn, 1974), the vector space is finitely generated. All the versions of the Steinitz exchange theorem which have been found in the literature are stated for finitely generated vector spaces except one case. The exception is (Tietz, 1973). In Greub (1967), this version appears as an exercise. The following is a translation from (Tietz, 1973).

Let V be a vector space, $S \subset V$ a spanning set and $T \subset V$ a finite linearly independent set. Then there exists $T' \subset S$ such that T' is equipotent to T and $(S - T') \cup T$ is a spanning set.

2.1.2 The generalized Steinitz exchange theorem

Those articles on the Steinitz exchange theorem which have been examined make no statements whether the Steinitz exchange theorem is true for infinite dimensional vector spaces. Eventually, the generalized Steinitz exchange theorem was found. This theorem has been discussed and proved in (Cohn, 2003), (Jacobson, 1951–1964), (Lenz, 1976), (Fuchs, 1970) and (Lüneburg, 1989). The presentations in these books are similar and there are no doubts that the proofs are correct. Since the generalized Steinitz exchange theorem implies the Steinitz exchange theorem, the Steinitz exchange theorem is true for vector spaces of infinite dimensions, compare (Lüneburg, 1989). What remains is to briefly comment on the articles of Hughes, which are (Hughes, 1962–1964) and (Hughes, 1965–1966), and the article of Graczyńska, (Graczyńska, 2010). In (Hughes, 1962–1964), dependent spaces were introduced. The introduction of these

spaces went in two steps. One consider an arbitrary set S and a set Δ of finite subsets of S which contains atleast two elements. A subset of S is called dependent if it is a superset of an element of Δ . A subset of S which is not dependent is called independent or a basis. Now one gets into a first conflict with (Graczynska, 2010). In this article, it is claimed that the existance of bases is a consequence of transitivity which has not been defined yet. This is false and it can be proven by the following.

Proposition 2.1.1. *Let S be a set and Δ a set of finite subset of S containing at least two elements. If Γ is an independent subset of S , then there exists a basis of S such that $\Gamma \subset \Delta$.*

Proof. Suppose $\Gamma \subset S$ is independent. Form

$$\Omega := \{\Lambda \subset S \mid \Gamma \subset \Lambda, \Lambda \text{ an indepent chain}\}$$

ordered by inclusion. $\Omega \neq \emptyset$ as $\Gamma \in \Omega$. Let C be a nonempty chain in Ω and form

$$\Psi := \bigcup_{\Sigma \in C} \Sigma.$$

Since no element of C contains an element of Δ , Ψ is indepent. Since C is non-empty, $\Gamma \subset \Psi$. Hence $\Psi \in \Omega$. Since $\Sigma \subset \Psi$ for all $\Sigma \in C$, Ψ is an upper bound of C . Therefore, by Zorn's lemma, Ω has maximal elements. Let M be a maximal element of Ω and $N \subset S$ a chain of independent subsets of S such that $M \subset N$. Since $N \in \Omega$ and M is a maximal element of Ω , $N = M$. Therefore, M is a maximal element and then a basis. ■

Now the second step of the construction of a covering space comes. A relation between elements of S and subsets of S was defined as follows. Let $x \in S$ and $A \subset S$. Then x depends on A , denoted by $x \sim \sum A$ if either $x \in A$ or $x \notin A$ and there exist finitely many elements of A say $x_1, \dots, x_n \in A$ such that $\{x, x_1, \dots, x_n\} \in \Delta$. Now one assumes that \sim is transtive. This means that if A and B are subsets of S and $x \in S$ such that $x \sim \Sigma A$ and $a \sim \Sigma B$ for all $a \in A$ then $x \sim \Sigma B$.

In (Hughes, 1962–1964), the following theorem was proven.

If A is basis and B an independent subset (both being well ordered) of the dependence space S , then there is a definite subset A' of A , such that $B + (A - A')$ is also a basis of S , and a definite one-one correspondence between A' and B .

In this context, one will ask what a definite subset A' of A is and what a definite one-one correspondence between A' and B is.

In (Hughes, 1965-1966), the following theorem, which is called the improved Steinitz' exchange theorem in this article was proven.

If A is a basis and B an independent subset, both being well ordered, of the dependence space S , there exists an explicitly defined, one-one mapping φ of B onto A' , a subset of A , such that φ is the identity map on $B \cap A$ and $B + (A - A')$ is a basis of S .

In this context, one will ask what an explicitly defined, one-one mapping is.

The problem of Hughes' articles is the notations he is using. If one looks at relation (1) in (Hughes, 1962–1964), on the left, one has an $n + 1$ tuple of elements of S and on the right, a set of subsets of S . But an $n + 1$ tuple is not an element of Δ . On page 113 of (Hughes, 1962–1964), there is ΣA . It is known that A is a subset of S but Σ is not defined and hence ΣA is not defined. In this article, $+$ has two meanings. It may mean the union of two disjoint sets or part of a relation like relation (2). The same applies to \sim .

Let $A \subset S$ and $x \in S$. If x is dependent on A in the sense of Hughes and $x \notin A$, then there exists $x_0, x_1, \dots, x_n \in A$ such that $\{x_0, x_1, \dots, x_n\} \in \Delta$. In the sense Cohn et. al. in addition it is assumed that $\{x_1, \dots, x_n\}$ is independent.

(Graczynska, 2010) tried to prove that the two definitions are equivalent. On page 155 of (Graczynska, 2010), the last 4 lines of the proof of lemma 8 are not correct.

2.2 Methodology used in this research

The methods used in this study are from set theory, linear algebra and algebra.

2.2.1 Methods from set theory

- The set theoretical definition of a function

This is needed in connection with section 4.1, where the set Ω is introduced. As defined in chapter 1, functions are sets of ordered pairs. They are used in this study to prove that two sets are equipotent if there exists an injective function between them.

- Ordered sets, maximal elements and Zorn's lemma

For vector spaces of infinite dimensions, one needs Zorn's lemma. Zorn's lemma works

with ordered sets. To prove that a set has maximal elements, the set need to be ordered. As seen in section 4.1, the set Ω of functions needs to be ordered and it is ordered by inclusion. Zorn's lemma is also needed to prove basic facts about vector spaces, like the existence of a basis or existence of complements for subspaces.

2.2.2 Methods from linear algebra and algebra

Vector spaces of arbitrary dimensions over arbitrary division rings are considered in this study. The Steinitz exchange theorem will be derived from the following fact: every spanning set of a vector space contains a subset which generates a complement to a given subspace. To get the full Steinitz exchange theorem, the theorem of Löwig is required, which states that any two bases of a vector space are equipotent.

Chapter 3

Results of the study

This chapter presents the results of the study and their proofs.

3.1 The Steinitz exchange theorem

To begin with, the Steinitz exchange theorem will be stated for arbitrary vector spaces. This version is equivalent to the versions in the literature.

Theorem 3.1.1 (The exchange theorem of Steinitz). *Let V be a K -left vector space, $S \subset V$ a spanning set of V and $I \subset V$ a linearly independent subset. Then there exists a subset $A \subset S$ such that A is equipotent to I and $I \cup (S - A)$ is a spanning set of V .*

It remains to compare the above theorem to what Graßmann has. In case of Graßmann, V is a real vector space, S is a finite subset of V and I is a finite linearly independent subset of the span of S . Graßmann concludes that S contains a subset $A \subset S$ which is equipotent to I and the span of $I \cup (S - A)$ is equal to the span of S .

There exists an obvious modification of the Steinitz exchange theorem. If S is a basis of V then the conclusion of the theorem holds because a basis is a spanning set. On the other hand, if one knows that the Steinitz exchange theorem holds for a basis, then the above version of the Steinitz exchange theorem is true because every spanning set of V contains a basis of V .

Now it will be explored how to prove the Steinitz exchange theorem. A first possibility is to set up an ordered set of functions. In this context, the domain of the function f will be denoted by D_f .

$$\Omega := \{f : D_f \rightarrow S \mid D_f \subset I, f \text{ injective}, D_f \cup (S - f(D_f)) \text{ spanning set of } V\}.$$

This set will be ordered by inclusion. If $f, g \in \Omega$, then $f \subset g$ means that g is an extension of f . The focus is on maximal elements of Ω , as it will become apparent soon.

Theorem 3.1.2. *Let $f \in \Omega$. Then f is a maximal element of Ω if and only if $D_f = I$.*

Proof. Suppose $f \in \Omega$ and $D_f = I$. Let $g \in \Omega$ such that $f \subset g$. Since $I = D_f \subset D_g$ and $D_g \subset I$, $D_g = D_f$. Since $f \subset g$, $g = f$. Therefore f is a maximal element of Ω .

To prove the converse implication, let $f \in \Omega$ such that $D_f \neq I$. Since D_f is a proper subset of I , there exists some $x \in I - D_f$. Since $D_f \cup (S - f(D_f))$ is a spanning set of V , $[D_f] + [S - f(D_f)] = V$. Hence there exist $u \in [D_f]$, $v \in [S - f(D_f)]$ such that $x = u + v$. Since I is linearly independent and $x \in I - D_f$, $x \notin [D_f]$. Therefore $v \neq 0$. Since $v \in [S - f(D_f)]$, there exist $y_1, \dots, y_n \in S - f(D_f)$ and $\lambda_1, \dots, \lambda_n \in K$ such that $v = \sum_{i=1}^n \lambda_i y_i$. Since $v \neq 0$, there exists $1 \leq j \leq n$ such that $\lambda_j y_j \neq 0$. Hence $x = u + v = u + \sum_{i=1}^n \lambda_i y_i = u + \lambda_j y_j + \sum_{\substack{i=1, \\ i \neq j}}^n \lambda_i y_i$ and this leads to

$$-\lambda_j y_j = u - x + \sum_{\substack{i=1, \\ i \neq j}}^n \lambda_i y_i.$$

Now

$$\bar{f} := f \cup \{(x, y_j)\}$$

will be considered. \bar{f} is a function because $x \notin D_f$. Since $y_j \notin f(D_f)$, \bar{f} is injective. Further, $D_{\bar{f}} = D_f \cup \{x\} \subset I$. Since $u \in [D_f]$ and $x \in D_{\bar{f}}$, $u - x \in [D_{\bar{f}}]$. Since $\{y_i \mid 1 \leq i \leq n, i \neq j\} \subset S - \bar{f}(D_{\bar{f}})$, $-\lambda_j y_j = u - x + \sum_{\substack{i=1, \\ i \neq j}}^n \lambda_i y_i \in [D_{\bar{f}}] + [S - \bar{f}(D_{\bar{f}})]$. Since $\lambda_j \neq 0$, $y_j \in [D_{\bar{f}}] + [S - \bar{f}(D_{\bar{f}})]$. Since $S - f(D_f) \subset [D_{\bar{f}}] + [S - \bar{f}(D_{\bar{f}})]$, $[D_{\bar{f}}] + [S - \bar{f}(D_{\bar{f}})] = V$. Therefore, $\bar{f} \in \Omega$. Since $f \subsetneq \bar{f}$, f is not a maximal element of Ω . ■

Suppose $g \in \Omega$ is a maximal element. Then $D_g = I$. Choose $A := g(I)$. Then $A \subset S$ and A is equipotent to I because g is injective and $I \cup (S - A) = I \cup (S - g(I))$ is a spanning set of V . Therefore A is a subset of S which has the properties which are stated in the Steinitz exchange theorem. The converse is also true. Let $I \subset V$ be linearly independent, $S \subset V$ a spanning set and $A \subset S$ such that $I \cup (S - A)$ is a spanning set. Since I is equipotent to A , there exists

a bijection $f : I \rightarrow A$. f is injective and $D_f = I$. Further, $I \cup (S - A)$ is a spanning set. Therefore f is a maximal element of Ω .

Proposition 3.1.1. *If I is finite, Ω has maximal elements.*

Proof. Suppose I is finite. If $f \in \Omega$, then $D_f \subset I$ and hence, $|D_f| \leq |I|$. Consider

$$\Gamma := \{D_f \mid f \in \Omega\}.$$

Since Ω is non-empty, Γ has maximal elements. Let $D_g \in \Gamma$ be a maximal element of Γ . Then g is a maximal element of Ω . ■

If I is infinite, Ω has maximal elements. Since the Steinitz exchange theorem is true, Ω has maximal elements. All attempts to prove that every non-empty chain in Ω is bounded above have failed.

That the Steinitz exchange theorem is true if I is finite and V is an arbitrary vector space is a known fact as it appears in (Tietz, 1973), page 48 ff. However, Tietz's proof is totally different. It is a proof by induction on the cardinality of I .

The following is an application of the Steinitz exchange theorem to finitely generated vector spaces. the statements which will be proved, can be found in nearly every book on linear algebra.

Theorem 3.1.3. *Let V be a finitely generated K -left vector space. Then the following statements holds.*

- a) *Every linearly independent subset of V is finite.*
- b) *Every subspace of V is finitely generated.*
- c) *Any two bases of V are equipotent.*

Proof. a) Since V is finitely generated, there exists a finite subset $S \subset V$ such that $[S] = V$. Let $I \subset V$ be linearly independent and let I_0 be a finite subset of I . By the exchange theorem of Steinitz, there exists $A \subset S$ such that A is equipotent to I_0 and $I_0 \cup (S - A)$ is a spanning set of V . Now $|I_0| = |A| \leq |S|$. Since every finite subset of I contains atmost $|S|$ elements, I is finite.

b) Let $U \leq V$ and let B be a basis of U . Since B is linearly independent, B is finite. Therefore, U is finitely generated.

c) Let B and B' be bases of V . By a), B and B' are finite. B is linearly independent and B' is a spanning set. By the Steinitz exchange theorem, there exists $A \subset B'$ such that A is equipotent to B and $B \cup (B' - A)$ is a spanning set. In particular, $|A| \leq |B'|$. Since $|A| = |B|$, one obtains $|B| \leq |B'|$. Interchanging B and B' results in $|B'| \leq |B|$. Therefore, $|B| = |B'|$. ■

The next theorem was first proved by Löwig and appeared in (Löwig, 1934).

Theorem 3.1.4 (Theorem of Löwig). *Any two bases of a K -left vector space V are equipotent.*

Proof. Let B and B' be bases of V . If B and B' are finite, By Theorem 3.1.3. c), $|B| = |B'|$ and hence B and B' are equipotent. Now suppose B and B' are infinite. Let $(\epsilon_b)_B$ be the family of the coordinate forms of V with respect to B . If $c \in B'$, consider

$$B_c := \{b \in B \mid \epsilon_b(c) \neq 0\}.$$

B_c is finite for all $c \in B'$. Assume $\{B_c \mid c \in B'\}$ does not cover B . Then $\bigcup_{c \in B'} B_c$ is contained in a hyperplane of V . This is a contradiction because this hyperplane contains B' . Therefore, $\{B_c \mid c \in B'\}$ covers B . By theorem 1.2.12., $\{B_c \mid c \in B'\}$ is equipotent to B . Since $x \in B' \rightarrow B_x \in \{B_c \mid c \in B'\}$ is surjective, $\{B_c \mid c \in B'\}$ is equipotent to a subset of B' . Therefore, B is equipotent to a subset of B' . Interchanging B and B' results in B' is equipotent to a subset of B . By the theorem of Schröoder-Bernstein, B and B' are equipotent. ■

Note that in the proof of the theorem of Löwig, cardinal numbers have not been used.

Now the proof of the Steinitz exchange theorem for arbitrary vector spaces follows.

Proof of the Steinitz exchange theorem. Let V be a K -left vector space, $I \subset V$ linearly independent and $S \subset V$ a spanning set of V . Since $[I] + [S] = V$, by the complement theorem, there exists $C \subset S$ such that $[I] \oplus [C] = V$. Since $[C] + [S - C] = V$, by the complement theorem, there exists $D \subset S - C$ such that $[C] \oplus [D] = V$. Since $[I]$ and $[C]$ have a common complement, $[I] \cong [D]$. There exists a basis A for $[D]$ such that $A \subset D$. Since $D \subset S$, $A \subset S$. Thus $[I] \cong [A]$ and I and A are bases. By the theorem of Löwig, $|I| = |A|$. Since $D \subset S - C$, $C \subset S - D$. Since $S - D \subset S - A$, $C \subset S - A$. Since $[I] \oplus [C] = V$, $[I \cup (S - A)] = V$. ■

From the Steinitz exchange theorem, one obtains a subset $A \subset S$ which satisfies two conditions. One can ask the question: When is this set A uniquely determined? This question is answered by the following theorem.

Theorem 3.1.5. *A is uniquely determined if and only if $I = \emptyset$ or $I \neq \emptyset$ and finite, S is a basis of V and $[[I] \cap S] = [I]$*

It will be shown that the condition is sufficient.

Proof. Suppose $I = \emptyset$. Since A is equipotent to I , $A = \emptyset$. Now suppose that $I \neq \emptyset$ and finite, S is a basis of V and $[[I] \cap S] = [I]$. Since $A \subset S$ and S is a basis of V , $[A] \oplus [S - A] = V$. Since $I \cup (S - A)$ is a spanning set, $V = [I \cup (S - A)] = [I] + [S - A] = [[I] \cap S] + [S - A] = [([I] \cap S) \cup (S - A)]$. Since S is a basis, $(([I] \cap S) \cup (S - A)) = S$. This implies $A \subset [I] \cap S$. Since $|[I] \cap S| = \dim[I] = |I| = |A|$, $A = [I] \cap S$.

To prove the converse implication, suppose A is uniquely determined. Assume I is infinite. Since A is infinite, there exists $a \in A$. Consider $A' := A - \{a\}$. Since A' is equipotent to A , A' is equipotent to I . Since $I \cup (S - A) \subset I \cup (S - A')$ and $I \cup (S - A)$ is a spanning set, $I \cup (S - A')$ is a spanning set. This is a contradiction because $A' \neq A$. Therefore, I is finite. One may assume that $I \neq \emptyset$. Assume $S - A$ is linearly dependent. Then $S - A$ is not a minimal spanning set of $[S - A]$. Hence there exist $c \in S - A$ such that $[S - (A \cup \{c\})] = [S - A]$. With $a \in A$, form $A' := (A - \{a\}) \cup \{c\}$. Then $|A'| = |I|$. Since $A' \subset A \cup \{c\}$, $S - (A \cup \{c\}) \subset S - A'$. $V = [I] + [S - (A \cup \{c\})] \leq [I] + [S - A'] = [I \cup (S - A')]$. Hence $I \cup (S - A')$ is a spanning set. This is a contradiction because $A' \neq A$. Therefore, $S - A$ is linearly independent.

Assume $[I] \cap [S - A] \neq \{0\}$. Choose $z \in ([I] \cap [S - A]) - \{0\}$. Since $z \in [S - A]$ there exist $s_1, \dots, s_n \in S - A$ and $\lambda_1, \dots, \lambda_n \in K^\times$ such that $z = \sum_{i=1}^n \lambda_i s_i$. Let $a \in A$ and consider $A' := (A - \{a\}) \cup \{s_1\}$. Then $A' \subset S$ and $|A'| = |I|$. One has $z - \sum_{i=2}^n \lambda_i s_i = \lambda_1 s_1$. Since $\lambda_1 s_1 \in [I] + [S - A]$ and $\lambda_1 \neq 0$, $s_1 \in [I] + [S - A']$. Since $S - A' \subset [I] + [S - A']$, $(S - A') \cup \{s_1\} = (S - ((A - \{a\}) \cup \{s_1\})) \cup \{s_1\} = S - (A - a) \supset S - A$. Since $S - A \subset [I] + [S - A']$ and $I \subset [I] + [S - A']$, $[I] + [S - A'] = V$. Hence $I \cup (S - A')$ is a spanning set. This is a contradiction because $A' \neq A$. Therefore, $[I] \cap [S - A] = \{0\}$. Hence $[I] \oplus [S - A] = V$. Since S is a spanning set, $[A] + [S - A] = V$. There exists a linearly independent $E \subset A$ such that $[E] \oplus [S - A] = V$. Since $[I]$ and $[E]$ have a common complement, $[I] \cong [E]$. Hence $|I| = \dim [I] = \dim [E] = |E| \leq |A| = |I|$. Hence $|E| = |A|$. Since $E \subset A$, $E = A$. This implies $[A] \oplus [S - A] = V$. Since A and $S - A$ are linearly independent, S is a basis of V .

Assume $[[I] \cap S] \neq [I]$. By theorem 1.4.12., there exists $F \subset S$ such that $F \neq S - A$ and $[I] \oplus [F] = V$. Consider $A' := S - F$. Obviously $A' \subset S$ and $F = S - A'$. Since $[I] + [F] = V$, $I \cup F = I \cup (S - A')$ is a spanning set. $A' \neq A$ because $F \neq S - A$. Since $[I] \oplus [S - A'] = V$ and $[A'] \oplus [S - A'] = V$, $[I] = [A']$. Hence $|I| = |A'|$. This is a contradiction. Therefore, $[[I] \cap S] = [I]$. ■

$[[I] \cap S] = [I]$ means that $[I]$ can be generated by a subset of the basis S .

3.2 Application of Steinitz exchange theorem

Some books on linear algebra claim that there exist many applications of the Steinitz exchange theorem. But all the books which have been examined have only one application namely to prove that two bases of a finitely generated vector space are equipotent. Since the Steinitz exchange theorem is true for arbitrary vector spaces, one obtains the theorem of Löwig.

Theorem 3.2.1. *Any two bases of a K -left vector space V are equipotent.*

Proof. Let B and B' be bases of V . Since B is linearly independent and B' is a spanning set, there exists a subset $A \subset B'$ such that A is equipotent to B and $I \cup (B' - A)$ is a spanning set. Since A is equipotent to B and $A \subset B'$, B is equipotent to a subset B' . Interchanging of B and B' results in B' is equipotent to a subset of B . By the theorem of Schrönder-Bernstein, B and B' are equipotent. ■

The above proof of the thorem of Löwig is a transparent proof which does not use cardinal numbers.

Chapter 4

Conclusions and recommendations

4.1 Conclusions

For the classical version of the Steinitz exchange theorem, i.e. the Steinitz exchange theorem for finitely generated vector spaces, a simplified proof has been given. In (Tietz, 1973), the theorem is proved by induction on the cardinality of the linearly independent set. The alternative is that one can set up an ordered set of functions. From maximal elements of this ordered set, one obtains the Steinitz exchange theorem.

In section 4.1, it was proven that the Steinitz exchange theorem holds for arbitrary vectors over arbitrary division rings. It was proved that the Steinitz exchange theorem implies the theorem of Löwig and complement theorem and that the theorem of Löwig together with the complement theorem implies the exchange theorem of Steinitz. However it remains open whether for the proof of the Steinitz exchange theorem the theorem of Löwig is needed.

The only application of the Steinitz exchange theorem which could be found in the literature is to prove that any two bases of a finitely generated vector space are equipotent. In addition to that, in this study the Steinitz exchange theorem is used to prove that every subspace of a finitely generated vector space is finitely generated.

The Steinitz exchange theorem is a statement of the following type: under certain conditions, there exists a set which satisfies given conditions. In this study it has been investigated when this set is uniquely determined. It was found that the subset $A \subset I$ of the linearly independent set is uniquely determined if and only if either the linearly independent set is empty or it is non-empty but finite, the spanning set is a basis and the span of the linearly independent set

is generated by a subset of the basis.

4.2 Recommendations

- The results in this study can be used for a book on Linear Algebra.
- A topic for further research is to investigate whether the Steinitz exchange theorem can be derived from the complement theorem.
- One could investigate maximal chains of arbitrary vector spaces. There could be a relationship between maximal chains of vector spaces and ordered bases.
- One may also consider topological vector spaces and investigate the Steinitz exchange theorem from this perspective.

References

- Bourbaki, N. (1968). *Elements of mathematics: Theory of sets* (Vol. 1). Addison-Wesley.
- Cohn, P. M. (1974). *Algebra* (Vol. 1). John Wiley & Sons, Chichester-New York-Brisbane-Toronto.
- Cohn, P. M. (2003). *Basic algebra*. Springer-Verlag, London.
- Fuchs, L. (1970). *Infinite abelian groups*. Academic press, New York and London.
- Graczyńska, E. (2010). Dependence spaces. *Bulletin of the section of logic*, 39(3), 153–160.
- Graßmann, H. (1878). *Die lineale ausdehnungslehre*. Verlag von Otto Wigand, Leipzig.
- Greub, W. H. (1967). *Linear algebra*. Springer-Verlag, Berlin-Heidelberg-New York.
- Halmos, P. R. (1960). *Naive set theory. the university series in undergraduate mathematics*. D. Van Nostrand Co., Princeton-NJ-Toronto-London-New York.
- Hughes, N. J. S. (1962–1964). Steinitz exchange theorem for infinite bases. *Compositio Mathematica*, 15, 113–118.
- Hughes, N. J. S. (1965–1966). Steinitz exchange theorem for infinite bases ii. *Compositio Mathematica*, 17, 152–155.
- Jacobson, N. (1951–1964). *Lectures in abstract algebra*. Springer-Verlag, New York.
- Lenz, H. (1976). *Grundlagen der elementarmathematik*. Carl Haser Verlag, München-Wien.
- Löwig, H. (1934). Über die dimension linearer räume. , 18-23.
- Lüneburg, H. (1989). *Tools and fundamental constructions of combinatorial mathematic*. BI-Wiss.-Verlag, Zurich.

Oeljeklaus, E., & Remmert, R. (1974). *Lineare algebra* (Vol. I). Springer-Verlag, Berlin-Heidelberg-New York.

Tietz, H. (1973). *Lineare geometrie*. Vandenhoeck und Ruprecht.

van der Waerden, B. L. (1930). *Algebra*. Springer-Verlag, Berlin-Heidelberg-New York.