Group Structure on arbitrary sets: An algebraic application of the Axiom of Choice

by Oskar Emmerich

6th June 2025

LUND
UNIVERSITY
Faculty of Science
Department of Mathematics

Bachelor's Thesis in Mathematics

Thesis advisor: Anitha Thillaisundaram

Abstract

The thesis should include an abstract that summarizes its contents; mathematical jargon can be utilized here. The typical length of an abstract is between 100 and 300 words.

Contents

Introduction

Historical Background

In 1902 Bertrand Russell showed with what is now known as *Russell's Paradox* that the previously used approach to set theory was inconsistent. Ernst Zermelo then created an axiomatic framework for set theory in 1905, motivated both by attempting to preserve results such as the theory of infinities by Georg Cantor, as well as avoiding paradoxes. These axioms, later modified by Abraham Fraenkel, became known as the nine *Zermelo-Fraenkel Axioms* (ZF) as well as the *Axiom of Choice* (AC)[Gol98, pp.66-70, 75].

The axiom of choice in particular is of special interest in many areas of mathematics, especially in algebra and topology, often in the form of the equivalent statement of *Zorn's Lemma*, which says that every non-empty partially ordered set with an upper bound has a maximal element [Jec78].

Finally in 1971 András Hajnal and Andor Kertész published a paper [HK72] which provided another equivalence to AC, namely that there exists a cancellative groupoid structure on every (uncountably infinite) set. This paper makes use of first-order model theory, an area of logic developed during the first half of the 20th century, which utilizes models of formal languages to obtain results. Kertész later expanded on this, providing an alternative algebraic partial proof in a lecture series given at the University of Jyväskylä [Ker75].

Thesis Structure

The aim of this thesis is to provide context to the paper [HK72] and to derive the theory needed for the proof of its main theorem:

Theorem 0.0.1. *The following sentences are equivalent in ZF:*

1. *Axiom of Choice*

2. *Every non-empty set admits a cancellative groupoid structure*

We start in the first chapter by briefly giving an overview of some of the necessary background knowledge needed for the rest of the text. This includes stating the ZF axioms as well as the Axiom of Choice itself.

Then, in the second chapter, we will explore orderings and well-orderings in the context of axiomatic set theory. Of special importance here will be Zorn's Lemma, a well-known equivalence of AC. We will finish this chapter by giving a proof for a lemma by Hartogs [Har15], which is also found in [HK72]. This lemma states that for any arbitrary set, there always exists an ordinal which no subset of that set can be injectively mapped to.

In the third chapter we will move on to an introduction to model theory. This is done with the aim of proving the upwards Löwenheim-Skolem Theorem, which states that a language with a countable model also has a uncountable model. Model theory is a very useful tool for applying results from logic to non-logic areas of mathematics, especially abstract algebra as we will see later. As Chang and Heisler put it in [CK90] (a very good historical introduction to model theory and the first comprehensive textbook for the subject),

Model Theory = Universal Algebra + Logic.

Finally, in the fourth and final chapter, we will give a detailed proof of the aforementioned theorem by Hajnal and Kertész. In this, we will apply the previous results by Hartogs and Löwenheim and Skolem from Chapters 2 and 3.

Chapter 1

Preliminaries

The convention in this thesis will be to say **ZF** when talking about Zermelo-Fraenkel set theory *without* the axiom of choice. When talking about the axiom of choice on its own we will say **AC**, and when talking about Zermelo-Fraenkel set theory together with the axiom of choice we use **ZFC**.

We will use the convention of including 0 at the beginning of the natural numbers $\mathbb{N}$, i.e. $\mathbb{N} = \{0, 1, 2 \ldots\}$. This is a *natural* choice, since we then can use $\mathbb{N}$ to mean the set described by the Axiom of Infinity. If we want to talk about strictly positive integers we use the notation $\mathbb{Z}^+ = \mathbb{N} \setminus \{0\}$.

Lastly, whenever we deal with the negation of some symbol, we cross it out to mean this, for example $a \neq b$ means $\neg (a = b)$.

1.1 First-Order Logic and Classes

When talking about first-order logic we mean the symbol

$$= \text{ (equals)}$$

in conjunction with the logical connectives

$$\neg \text{ (not)}, \ \wedge \text{ (and)}, \ \vee \text{ (or)}, \ \rightarrow \text{ (implies) and } \ \leftrightarrow \text{ (if and only if)}$$

as well as the quantifiers

$$\forall \text{ (for all) and } \exists \text{ (exists).}$$

Additionally, as we are talking about sets, we use the symbol $\in$ to denote set inclusion. [Jec78, pp.2-3]

In order to effectively talk about properties of set and set-like structures we need to somehow properly define formulas. We will go more into depth about formulas in Chapter 3, where we will introduce the notion of a formal language.

For now we are only concerned with the language of sets defined above.

Definition 1.1.1 (Formula of a Set)**.** An *atomic formula* in set theory is either

1. $x = y$, or
2. $x \in y$

A *formula* ϕ is any combination of atomic formulas with logical connectives and quantifiers.

The symbols x and y above are called variables and for any two variables an atomic formula is either true or false for each x and y. A variable *occurs freely* inside of a formula if it does not appear inside of a $\exists$ or $\forall$ quantifier, otherwise the variable is *bound*. We write $\phi(x_1, \ldots, x_n)$ for a formula with $n \in \mathbb{Z}^+$ free variables. A formula where every variable is bound is called a *sentence*. [Mar02, pp.10-11]

A sentence is either true or false and a formula with free variables is true or false for each choice of the free variables. Each of the **ZF** axioms below are examples of sentences and as axioms we assume them to be inherently true (within the framework of our theory). An example of a formula with a free variable would be

$$\phi(x) = \exists y \ (y \in x).$$

This formula is only false for the empty set, since it is the unique set which does not contain any elements. In that sense we think of formulas with free variables describing a *property*, something we make use of in *classes*.

Definition 1.1.2 (Class)**.** Let $\phi(x, p_1, \ldots, p_n)$ be a formula in first-order logic. Then a *class* $\mathbf{C}$ is defined as

$$\mathbf{C} = \{x \mid \phi(x, p_1, \ldots, p_n)\}.$$

The class $\mathbf{C}$ is called *definable from* $p_1, \ldots, p_n$. Furthermore, if x is the only free variable of ϕ, the class $\mathbf{C}$ is simply called *definable*. [Jec78, p.3]

In practice, we use classes as a tool to help us construct useful sets in **ZFC**, as elements of classes are always sets in the stricter sense. All sets are classes, but not all classes are sets. Assume that we have a fixed set s; we can then always construct the corresponding class $\mathbf{S} = \{x \mid x = s\}$. A class which is not a set is called a *proper class*.

We consider to two classes to be the same if they have the same elements. The familiar operations of *inclusion*, *union*, *intersection*, and *difference* are definable using formulas. As such for classes $\mathbf{C}$, $\mathbf{D}$,

$$\begin{aligned}
\mathbf{C} \subseteq \mathbf{D} &\iff \forall x \, (x \in \mathbf{C} \implies x \in \mathbf{D}) \\
\mathbf{C} \cup \mathbf{D} &= \{x \mid x \in \mathbf{C} \vee x \in \mathbf{D}\} \\
\mathbf{C} \cap \mathbf{D} &= \{x \mid x \in \mathbf{C} \wedge x \in \mathbf{D}\} \\
\mathbf{C} \setminus \mathbf{D} &= \{x \mid x \in \mathbf{C} \wedge x \notin \mathbf{D}\} \\
\bigcup \mathbf{C} &= \{x \mid x \in S \text{ for some } S \in \mathbf{C}\}
\end{aligned}$$

[Jec78, pp.3-4]

For the use in this text, classes are in a sense "naive sets-like object"; they to help us describe collections of sets without worrying about paradoxes. Consider for example the class $\mathbf{V} = \{x \mid x = x\}$, which is the universe of all sets and does not exist in pure set theory. Another important class which we will make use of later is the *empty class* $\varnothing = \{x \mid x \neq x\}$ (although this is also a set as we will see).

1.2 Zermelo-Fraenkel Axioms of Set Theory

We assume that the reader has some familiarity with axiomatic set theory, but for convenience and consistency we restate some of the necessary basics here. For a more thorough introduction of the topic, see [Gol98, §§4.3-4.5], alternatively [Jec78, §1.1] gives a more technical overview. The formulation of the axioms below is based on both textbooks.

1.2.1 Axiom of Extensionality

$$\forall x \, \forall y \, \big(x = y \iff \forall z \, (z \in x \iff z \in y)\big)$$

Two sets are equal if and only if they contain the same elements.[Gol98, §4.3, p.76]

1.2.2 Axiom of Pairs

$$\forall x \, \forall y \, \exists z \, \forall w \, \big(w \in z \iff (w = x \vee w = y)\big)$$

For any two sets, there is a set whose elements are precisely these sets.

We define an ordered pair $\langle x, y \rangle$ to be the set $\big\{\{x\}, \{x, y\}\big\}$. Further, ordered n-tuples are defined recursively as $\langle x_1, x_2, x_3, \ldots, x_n \rangle = \langle x_1, \langle x_2, x_3, \ldots, x_n \rangle\rangle$.[Gol98, §4.3, pp.76, 79-80]

Ordered pairs satisfy the property that for any sets x, y, u, v, if $\langle x, y \rangle = \langle u, v \rangle$, then $x = u$ and $y = v$.[Gol98, Theorem 4.2]

1.2.3 Axiom Schema of Separation

Let $\phi(z, p)$ be a formula in first-order logic with free variables z and x. Then

$$\forall x \, \forall p \, \exists y \, \forall z \, \big(z \in y \iff (z \in x \wedge \phi(z, p))\big). \tag{1.2.1}$$

For any sets x and p there exists a unique set consisting of all elements z in x for which $\phi(z, p)$ holds. This is an axiom schema, meaning an infinite collection of axioms, since (1.2.1) is a separate axiom for every formula $\phi(z, p)$.

Assume $\psi(z, p_1, \ldots, p_n)$ is a more general formula for which we want to utilize the axiom schema for. We can then let $\phi(z, p)$ be the formula

$$\phi(z, p) = \exists p_1 \cdots \exists p_n \left(\left(p = \langle p_1, \ldots, p_n \rangle \right) \text{ and } \psi(z, p_1, \ldots, p_n) \right),$$

where we have that $\phi(z, \langle p_1, \ldots, p_n \rangle)$ is true if and only if $\psi(z, p_1, \ldots, p_n)$ is true. Hence we can generalize the axiom schema to

$$\forall x \, \forall p_1 \cdots \forall p_n \, \exists y \, \forall z \left(z \in y \iff \left(z \in x \wedge \psi(z, p_1, \ldots, p_n) \right) \right). \tag{1.2.2}$$

Let $\mathbf{C} = \{ z \mid \psi(z, p_1, \ldots, p_n) \}$ be the class containing all sets z which satisfy the formula ψ for any free variables $p_1, \ldots, p_n$. We can then utilize (1.2.2) and have that

$$\forall x \, \exists y \, \left(\mathbf{C} \cap x = y \right),$$

describes the same set. This means that any subclass of a set is also a set, and naturally a subclass of a set is called a subset.

As a consequence the set theoretic operations *difference* $(x \setminus y) \subseteq x$ and *intersection* $(x \cap y) \subseteq x$ are also defined for any two sets x and y, since all sets also are classes. Further the intersection

$$\bigcap \mathbf{C} = \{ z \mid z \in x \text{ for every } x \in \mathbf{C} \}$$

of a class $\mathbf{C}$ is a set, since it is a subset of all of its elements, which are strictly sets. [Jec78, pp.5-6]

1.2.4 The Empty Set

$$\exists x \forall y \; y \notin x$$

There is a set with no elements. We call this set $\varnothing = \{\}$. [Gol98]

This statement stands out, as it is not an axiom, the existence of the empty set also arises from the Axiom Schema of Separation. We include it here for the sake of completeness.

Since we can define the empty class $\varnothing = \{ u \mid u \neq u \}$, the empty set is also a set. However this follows from $\varnothing$ being a subset of all sets and hence only under the assumption that at least one other set exists. The existence of that set, in turn, follows from the Axiom of Infinity.[Jec78, p.6]

1.2.5 Axiom of Unions

$$\forall x \exists y \forall z \left(z \in y \iff \exists w \left(z \in w \wedge w \in x \right) \right)$$

For any set x there is a set, denoted by $\bigcup x$, which is the union of all elements of x, meaning it contains all elements of each member of x.

For any sets, we recursively define their union as

$$\begin{aligned} x \cup y &= \bigcup \{x, y\} \\ x \cup y \cup z &= (x \cup y) \cup z \\ &\vdots \end{aligned}$$

as well as a set with more then two members as,

$$\{x_1, \ldots, x_n\} = \{x_1\} \cup \cdots \cup \{x_n\}$$

where the existence of $\{x, y\}$ and $\{x\} = \{x, x\}$ is justified by the Axiom of Pairs. [Jec78, p.6]

1.2.6 Axiom of Power Sets

$$\forall x \, \exists y \, \forall z \, (z \in y \iff z \subseteq x)$$

For any set x there is a set, denoted by $\mathcal{P}(x)$ and called the power set of x, consisting of all subsets $s \subseteq x$.

Using the axioms of Separation, Union and Power Set we can define the *cartesian product* of the sets x and y as

$$x \times y = \{\langle u, v\rangle \mid u \in x \vee v \in y\} \subseteq \mathcal{P}\left(\mathcal{P}\left(x \cup y\right)\right),$$

and for multiple sets $x_1, \ldots, x_n$ as

$$x^n = \underbrace{x_1 \times \cdots \times x_{n-1} \times x_n}_{n \text{ times}} = (x_1 \times \cdots \times x_{n-1}) \times x_n.$$

We call a set $R \subseteq X^n$ an *n-ary* relation over X. In general R is a set of tuples and if R is a binary relation, we write $x \, R \, y$ for $\langle x, y\rangle \in R$.

A binary relation $f \subseteq X \times Y$ is called a *function* (or *map*), if

$$(\langle x, y\rangle \in f \wedge \langle x, z\rangle \in f) \implies y = z$$

holds for all $x \in X$ and $y, z \in Y$, and if for all $u \in X$ there exists some $v \in Y$, such that $\langle u, v\rangle \in f$.

In the definition above, we call X the *domain* and Y the *codomain* of f. The set

$$\mathbf{Im}(f) = \{y \mid \exists x \in X \ (\langle x, y\rangle \in f)\}$$

is called the *image* of f. In general we write $f : X \to Y$ for $f \subseteq X \times Y$ and $f(x) = y$ whenever $\langle x, y\rangle \in f$, in the latter case saying that x *maps to* y in f.

A function $f : X \to Y$ is called a *surjection* if $\mathbf{Im}(f) = Y$ and an *injection*, if

$$(f(x_1) = y \wedge f(x_2) = y) \implies x_1 = x_2.$$

A function is *bijection*, if it is both a surjection and an injection. [Jec78, pp.7-10]

1.2.7 Axiom of Infinity

$$\exists x \left(\varnothing \in x \wedge \forall y \left(y \in x \implies y \cup \{y\} \in x\right)\right)$$

There is an inductive set. An inductive set contains both the empty set $\varnothing$ as well as the successor x^+ of every x in the set. In this context the successor of a set x is defined as $x^+ = x \cup \{x\}$. We will go into more detail on this in Section 2.3.

1.2.8 Axiom Schema of Replacement

Let $\phi(x,\ y,\ p)$ a formula in first-order logic with free variables x, y and p.

$$\begin{aligned} &\forall x\, \forall y\, \forall z\ \left(\phi(x,\ y,\ p) \wedge \phi(x,\ z,\ p) \implies y = z\right) \\ &\implies \forall x'\, \exists y'\, \forall y\ \left(y \in y' \iff (\exists x \in x')\, \phi(x,\ y,\ p)\right) \end{aligned} \tag{1.2.3}$$

Similar to the Axiom Schema of Separation, this is an axiom schema, meaning that (1.2.3) is a separate axiom for each formula $\phi(x,\ y,\ p)$. We can also generalize the Axiom Schema of Replacement in a similar way, by replacing $\phi(x,\ y,\ p)$ with $\phi(x,\ y,\ p_1, \ldots,\ p_n)$.

The axiom schema states that if $\mathbf{F} = \left\{(x\, y) \mid \phi(x,\ y,\ p)\right\}$ is a class function, then the image $\mathbf{Im}(\mathbf{F})$ is a set whenever we restrict the domain of $\mathbf{F}$ to a set. Consequently this restriction of $\mathbf{F}$ is also a function of sets. [Jec78, p.11]

1.2.9 Axiom of Foundation

$$\forall x \exists y \left(y \in x \wedge x \cap y = \varnothing\right)$$

Every set contains an $\in$-minimal element, we call this being *well-founded*.[Gol98, p.92] This also means there exist no infinitely descending chains of sets, such as $x_0 \ni x_1 \ni x_2 \ni \cdots$.[Gol98, Theorem 4.3, p.95]

1.3 The Axiom of Choice

To talk about the Axiom of Choice we need to first define what a choice function is, the concept which the axiom is centered around.

Definition 1.3.1 (Choice Function)**.** [Jec78, p.38] Let S be a family of non-empty sets. A function $f : S \to \bigcup S$ is called a *choice function* of S if

$$f(X) \in X$$

holds for all sets $X \in S$.

The Axiom of Choice is then defined as follows:

Definition 1.3.2 (Axiom of Choice)**.** [Jec78, p.38] There exists a choice function for every family of non-empty sets.

The Axiom of Choice is not always needed for showing that a choice function exists. Take for example $S = \mathcal{P}(\mathbb{N})$, under the usual order $<$ every subset of $\mathbb{N}$ has a least element. We can therefore construct a choice function $f : S \to \mathbb{N}$ by letting $f(N)$ be the unique least element of N for $N \in S$. This is however not possible for a family of possibly infinite subsets of $\mathbb{R}$; for example the open interval $(0,\ 1)$ does not contain a least element.

In general there does not always exist an external structure for sets which we can utilize to construct a choice function. The Axiom of Choice ensures that we can, but not how that choice function might look like. In fact **AC** is the only axiom of **ZFC** which states the existence of a mathematical object without explicitly defining it. This is a powerful tool, but can lead lead to fairly unintuitive results. As such, with **AC** there exists a way to order the real numbers where every subset has a least element (including open intervals like $(0,\ 1)$)!

Chapter 2

Orderings and Well-Orderings

2.1 Linear, Partial and Well-Orderings

We start by defining the two types of partial orderings, *strict* and *weak* ones. To give a more intuitive understanding of how these differ we use the notation $<$ and $\leq$ respectively, but R is also commonly used to denote a relation.

Definition 2.1.1 (Strict Partial Order)**.** [Gol98, p.165] Let X be a set and $< \subseteq X \times X$ a binary relation on X. Then $<$ is called a *(strict) partial order of* X, and $(X, <)$ called a *(strictly) partially ordered set*, if it is

(i) **irreflexive**: $\forall x \in X \ (x \not< x)$

(ii) **transitive**: $\forall x, y, z \in X \ \big((x < y \wedge y < z) \implies x < z\big)$

It is called *linear* if for all x, y in X, $x < y$ or $y < x$ or $x = y$.

Definition 2.1.2 (Weak Partial Order)**.** [Gol98, p.164] Let X be a set and $\leq \subseteq X \times X$ a binary relation on X. Then $\leq$ is called a *weak partial order* of X, and $(X, \leq)$ is called a *weakly partially ordered set*, if it is

(i) **reflexive**: $\forall x \in X \ (x \leq x)$

(ii) **transitive**: $\forall x, y, z \in X \ \big((x \leq y \wedge y \leq z) \implies x \leq z\big)$

(iii) **anti-symmetric**: $\forall x, y \in X \ \big((x \leq y) \wedge (y \leq x) \implies x = y\big)$

It is called *linear* if for all x, y in X, $x \leq y$ or $y \leq x$.

Definition 2.1.3. [Jec78, p.13] If $(X, <_X)$ and $(Y, <_Y)$ are two partially ordered sets, we call a function $f : X \to Y$ *order-preserving* if

$$x_1 <_X x_2 \iff f(x_1) <_Y f(x_2).$$

If X and Y are both linearly ordered, an order-preserving function is also said to be *increasing*.

The function f is called an *order-isomorphism* if f is both order-preserving and bijective. Whenever it is clear from context that we are talking about ordered sets we simply call f an *isomorphism* and write $X \simeq Y$. If f is order-preserving and injective, it is called an *order-embedding*. [Gol98, p.167]

A partially ordered set $(X, <)$ is sometimes also referred to simply as X by some abuse of notation when the relation $<$ is known. Additionally, whenever we talk about partially or linearly ordered sets without specifying which type, and where the type of partial order matters, we are referring to strict ones. [Jec78, p.12] Out of convenience, when talking about a strict partial order $<$, we sometimes refer to the term $(a < b \vee a = b)$ as $a \leq b$.

Clearly it is straightforward to define a weak partial order R' from a strict partial order R, letting $\langle x, y \rangle \in R'$ whenever $\langle x, y \rangle \in R$ or $x = y$.

Example 2.1.4. Let $(\mathbb{Q}, <)$ and $(\mathbb{R}, <)$ be the rational and real numbers with their respective usual order. Then, both $\mathbb{Q}$ and $\mathbb{R}$ are strictly partially and linearly ordered with respect to $<$. Additionally the function $f : \mathbb{Q} \to \mathbb{R}$ defined as $f(x) = x$ is an order-embedding, however due to the differences in cardinality the inverse f^{-1} is not a proper function.

Example 2.1.5. Consider the complex numbers $\mathbb{C}$ with $<_{\mathbb{R}}$ the usual order as defined on $\mathbb{R}$. Then $(\mathbb{C}, <_{\mathbb{R}})$ is a strict partial order, but not linear since $<_{\mathbb{R}}$ is only defined for strictly real numbers. Let us denote a complex number as an element of $\mathbb{R}^2$, so we express $z = a + bi$ as $\langle a, b \rangle$. We can then define a linear order on $\mathbb{R}^2$, by letting

$$\langle a_1, b_1 \rangle < \langle a_2, b_2 \rangle \text{ if } a_1 < a_2 \\ \text{or } a_1 = a_2 \wedge b_1 < b_2.$$

This is called the *lexicographical order* of $\mathbb{R}^2$ with respect to the order $<$ on $\mathbb{R}$ and is perhaps the most natural way to a Cartesian product. [Gol98, p.182] The *anti-lexicographical order* order of $\mathbb{R}^2$ would be

$$\langle a_1, b_1 \rangle < \langle a_2, b_2 \rangle \text{ if } b_1 < b_2 \\ \text{or } b_1 = b_2 \wedge a_1 < a_2.$$

If we express a complex number as $z = re^{\varphi i}$, we can define a linear order a different way:

$$z_1 = r_1 e^{\varphi_1 i} < z_2 = r_2 e^{\varphi_2 i} \text{ if } r_1 < r_2 \\ \text{or } r_1 = r_2 \wedge (\varphi_1 < \varphi_2 \bmod 2\pi).$$

If we view the number $z = re^{\varphi i}$ as the ordered pair $\langle r, \varphi \rangle$, this is then the lexicographical ordering of $\mathbb{R} \times [0, 2\pi)$.

Definition 2.1.6. [Jec78, p.12] An element a of an ordered set $(X, <)$ is the *least element* of X with respect to $<$, if $\forall x \in X\ (a < x \vee x = a)$. Similarly, an element z is called the *greatest element* of X if $\forall x \in X\ (x < a \vee x = a)$.

This notion of a least element lets us define a special kind of linearly ordered set:

Definition 2.1.7 (Well-Order)**.** [Jec78, p.13] A strict linear order $<$ of a set X is called a *well-ordering* if every subset of X has a least element.

Example 2.1.8. The natural numbers $\mathbb{N}$ are a well-ordered set with respect to their usual order. The least element of $\mathbb{N}$ is 0.

Example 2.1.9. The integers $\mathbb{Z}$ are not well-ordered under their usual order. They have no least element, and while of course every finite subset has a least element, this does not hold for all subsets of $\mathbb{Z}$ (for example the set of even integers).

Well-ordered sets are central to the axiomatic set theory at hand. In fact, one of the most important results we will treat here, is that every set can we well-ordered (with the Axiom of Choice).

Further, we will introduce the concept of ordinals as a way to properly classify all well-ordered sets. The next two lemmata are needed for the proof of Theorem 2.3.5, an important result with regards to ordinals. For this we need to define the initial section of an ordered set first:

Definition 2.1.10. If X is a well-ordered set and $s \in X$, we call the set $\{x \in X \mid x < s\}$ an *initial segment* of X.

Lemma 2.1.11. *[Jec78, Lemma 2.1, p.13] If $(W, <)$ is a well-ordered set and $f : W \to W$ is an increasing function, then $f(x) \geq x$ for each $x \in W$.*

Proof. In order to contrive a contradiction, we assume that $X = \{x \in W \mid f(x) < x\}$, the collection of elements of W not satisfying the lemma, is a non-empty set. We then let z be the least element of X and $w = f(z)$ its preimage in f. By the definition of X we this means that $f(w) < w$, contradicting the initial assumption that f is an increasing function. □

Lemma 2.1.12. *[Jec78, Lemma 2.2, p.13] No well-ordered set is isomorphic to an initial segment of itself.*

Proof. Assume for a contradiction that f is an order-isomorphism from an ordered set $(X, <)$ to an initial segment $(S, <) = \{x \in X \mid x < s\}$, for some $s \in X$ of itself. The image of f is then $\mathbf{Im}(f) = \{x \in X \mid x < s\} = S$, but we know this is not possible by Lemma 2.1.11. □

Theorem 2.1.13. *[Jec78, Theorem 1] Let $(W_1, <_1)$ and $(W_1, <_2)$ be well-ordered sets. Then one of the following holds:*

1. *W_1 is isomorphic to W_2,*
2. *W_1 is isomorphic to an initial segment of W_2,*
3. *W_2 is isomorphic to an initial segment of W_1.*

Proof. Let W_1 and W_2 be as in the statement of the theorem and let $W_i(u)$ be the initial segment $\{u \in W_i \mid u < v\}$ of W_i for $i \in \{1, 2\}$. We can then define the following set of ordered pairs:

$$f = \left\{\left\langle x, y\right\rangle \in W_1 \times W_2 \mid W_1(x) \simeq W_2(y)\right\}.$$

By Lemma 2.1.12 no element of either W_1 or W_2 can be a member of more than one ordered pair in f, since

$$\left\langle x, y_1\right\rangle, \left\langle x, y_1\right\rangle \in f \implies y_1 \simeq x \simeq y_2.$$

Hence f is a bijective function, however not necessarily one of which the domain and image are W_1 and W_2.

Let $h : W_1(u) \to W_2(v)$ be an isomorphism between two initial segments of W_1 and W_2. Then if we have $u' < u$ in W_1, it follows that $W_1(u') \simeq W_2(h(u'))$ and hence $\langle u', h(u')\rangle$ must be in f.

Based on these properties we can explore the following cases:

1. If $\mathbf{Dom}(f) = W_1$ is the domain of f and $\mathbf{Im}(f) = W_2$, we have that W_1 and W_2 must be isomorphic. Hence case 1 of the the theorem holds.

2. If $\mathbf{Im}(f) \neq W_2$, we have that $W_2 \setminus \mathbf{Im}(f)$ is non-empty and denote the least element of $W_2 \setminus \mathbf{Im}(f)$ by y_0. Then $\mathbf{Im}(f) = W_2(y_0)$, since for $y_1 < y_2$ in W_2, having $y_2 \in \mathbf{Im}(W_2)$ means that $y_1 \in \mathbf{Im}(W_2)$.

   Further $\mathbf{Dom}(f) = W_1$, because otherwise $\mathbf{Dom}(f) = W_1(w_0)$ for the least element x_0 of $W_1 \setminus \mathbf{Dom}(f)$. This in turn results in a contradiction as $W_1(x_0)$ is necessarily isomorphic to $W_2(y_0)$, meaning that $\left\langle x_0, y_0\right\rangle \in f$ and $x_0 \in \mathbf{Dom}(f)$.

   As such we have that $f : W_1 \to W_2(y_0)$ is an order-isomorphism and case 2 of the theorem holds.

3. If $\mathbf{Dom}(f) \neq W_1$, we have that $\mathbf{Dom}(f) = W_1(w_0)$ for the least element x_0 of the set $W_1 \setminus \mathbf{Dom}(f)$. Proceeding analogously to the case before we have that $\mathbf{Im}(f) = W_2$. Hence W_2 is order-isomorphic to an initial segment of W_1 and case 3 of the theorem holds.

By case 2 it is clear that these are the only possibilities for $\mathbf{Dom}(f)$ and $\mathbf{Im}(f)$ and by Lemma 2.1.12 the cases must be mutually exclusive. □

2.2 Properties of Linear Orderings

There are some more important concepts to define when discussing linear orderings, namely how we describe their properties. The sets $\mathbb{N}$, $\mathbb{Z}$ and $\mathbb{Q}$ are all countable, but their usual orderings clearly all differ. On the other hand, $\mathbb{Q}$ and $\mathbb{R}$ have different cardinalities, however the way both are ordered seems very similar.[1]

Definition 2.2.1. [Ros82, Definition 1.20] Let $(X, <)$ be a strictly linearly ordered set and $b \in X$ an element of X. Then an element $c \in X$ is called the (unique and immediate) *successor* of b, if

$$\forall x \in X \ (x < c \implies x < b \vee x = b) .$$

Similarly an element $a \in X$ is called the (unique and immediate) *predecessor* of b, if

$$\forall x \in X \ (a < x \implies x = b \vee b < x) .$$

Every element in $\mathbb{N}$ and $\mathbb{Z}$ has an immediate successor and every element in $\mathbb{Z}$ has an immediate predecessor. This is however not the case for elements of $\mathbb{Q}$, as the natural order of the rationals is *dense*.

Definition 2.2.2 (Dense orderings)**.** [Ros82, Definition 2.1] Let $(Y, <)$ be a strictly linearly ordered set. Then Y is called *dense*, if

$$\forall a_1, a_2 \in Y \ (a_1 < a_2 \implies \exists a \in Y \ (a_1 < a \wedge a < a_2)) .$$

We will not dwell on the concept of density too long. Especially the distinction between the two dense orderings of $\mathbb{Q}$ and $\mathbb{R}$ goes more into the direction of point-set topology and is beyond the scope of this text. For a treatment of this topic from the perspective of linear orderings we refer the curious reader to [Ros82, §2].

The broader discussion of the properties of linear orderings is important however, as we need a way to classify and compare ordered sets. For the classification we utilize order preserving functions; this is especially useful for the use-case of well-ordered sets all they always relate to each other in this way by Theorem 2.1.13.

Definition 2.2.3 (Order Type)**.** [Ros82, Definitions 1.12, 1.13] Let $(X, <_X)$ and $(Y, <_Y)$ be linear orderings. We say that X and Y have the same *order type*, if there exists an order-isomorphism $f : X \to Y$.

Assume that α is some linear ordering that we choose as a representative. If X and α are order-isomorphic we also say that X has *order type* α.

[1] The sets $\mathbb{Z}$, $\mathbb{Q}$ and $\mathbb{R}$ are used without formally defining them or their orders here. It is assumed the reader is familiar.

Example 2.2.4. Consider the positive integers $\mathbb{Z}^+ = \mathbb{N} \setminus \{0\}$. We can define the function $s : \mathbb{N} \to \mathbb{Z}^+$ by $s(n) = n + 1$ and have that s is an order-isomorphism:

$$n < m \iff n + 1 < m + 1.$$

Hence $\mathbb{Z}^+$ has the same order-type as $\mathbb{N}$.

Example 2.2.5. [Ros82, Exercise 2.3] We denote the order type of the rational numbers $\mathbb{Q}$ under the usual order by η. Then the punctured rationals $\mathbb{Q} \setminus \{0\}$ with the usual ordering also have order type η.

2.3 Ordinals

There are several ways to define the natural numbers $\mathbb{N}$, the way we do it here, and the way generally used in set theory, is to use the Axiom of Infinity, motivated by the Peano Axioms. This states that $\mathbb{N}$ is an *inductive set*, meaning that it contains 0, defined as $\varnothing = \{\}$, as well as the successor of every element in it, including of course 0 itself. [Gol98, p.39]

Definition 2.3.1. [Gol98, p.38] The successor of a set α is $\alpha^+ = \alpha \cup \{\alpha\}$. The successor of $0 = \varnothing$ is called 1 and the successor of 1 is called 2, etc.

The consequence of this is that $\mathbb{N}$ is the set we are familiar with: $\{0, 1, 2, 3, \ldots\}$. It also means that any natural number is defined as the set of all of its predecessors. For example $3 = \{0, 1, 2\} = \{\{\}, \{\{\}\}, \{\{\}, \{\{\}\}\}\}$ and $5 = \{0, 1, 2, 3, 4\}$.

Perhaps slightly more subtle then is that, under the usual ordering, $n < m \implies n \in m \wedge n \subset m$. This is an important property and the natural numbers, as well as the set $\mathbb{N}$ at large, are called *transitive* sets. The notion of a *transitive set* is not to be confused with that of a *transitive (binary) relation*, which is an unfortunate overlap in terminology.

Definition 2.3.2. [Jec78, p.14] A set T is called *transitive* if

$$\forall x \, (x \in T \implies x \subseteq T).$$

Definition 2.3.3. [Jec78, p.14] A set is called an *ordinal number* or *ordinal* if it is transitive and well-ordered by $\in$. We say $\alpha < \beta$ if and only if $\alpha \in \beta$.

Ordinals are denoted by lowercase Greek letters: $\alpha, \beta, \gamma, \ldots$. The ordinal associated with $(\mathbb{N}, \in)$ specifically is denoted by ω. We know that ω is indeed an ordinal by construction. It follows from the following lemma that every natural number also is an ordinal with respect to set inclusion.

Lemma 2.3.4. *[Jec78, Lemma 2.3, p.15]*

1. *The empty set $\emptyset$ is an ordinal.*
2. *If α is an ordinal and $\beta \in \alpha$, then β is an ordinal.*
3. *If α, β are ordinals and $\alpha \subset \beta$, then $\alpha \in \beta$.*
4. *If α, β are ordinals, then either $\alpha \subseteq \beta$ or $\beta \subseteq \alpha$.*

Proof.

1. The empty set has no non-empty subsets, hence it is transitive and well-ordered by $\in$.
2. If $\beta \in \alpha$, then $\beta \subseteq \alpha$ by definition. Since α is well-ordered and transitive, so is β.
3. Let γ be the least element of the set $\beta \setminus \alpha$. We show that $\alpha = \gamma$.

   The ordinal α is transitive by definition and from this it follows that there are no "gaps" in the order. Indeed α must be an initial segment of β. As an initial segment, we can describe α as the set $\{\xi \in \beta \mid \xi < \gamma\}$. Again by the definition of ordinals, this is the set γ itself and $\alpha = \beta$.
4. We know that the intersection $\alpha \cap \beta = \gamma$ must be an ordinal, since not least the empty set also is an ordinal. However anything other than $\alpha = \gamma$ or $\beta = \gamma$ results in a contradiction:

   Assume for this contradiction that $\gamma \in \alpha$. Then $\gamma \in \beta$ by the second point of the lemma. Because γ is defined as the intersection of α and β, this means that $\gamma \in \gamma$. Since γ is an ordinal, strictly linearly ordered, this is not possible. □

Theorem 2.3.5. *[Jec78, Theorem 2, p.15] Every well-ordered set is order-isomorphic to a unique ordinal.*

Proof. Let W be a well ordered set. We will show that W is order-isomorphic to an ordinal α, the uniqueness of α follows from Lemma 2.1.12.

Let $\mathbf{F}$ be the class function

$$\mathbf{F} = \{\langle x, \alpha\rangle \mid W(x) \simeq \alpha\},$$

which maps an element x to the ordinal α, only if the initial segment $\{u \in W \mid u < x\}$ given by x is order-isomorphic to α. Then, by the Axiom Schema of Replacement, the restriction $\mathbf{F}|_W$ is a function of sets, since $\mathbf{Dom}(\mathbf{F}|_W) \subseteq W$. As such $\mathbf{Im}(\mathbf{F}|_\mathbf{W})$ is also a set.

Additionally, we have that $\mathbf{F}(w)$ is defined for each $w \in W$: Consider for a contradiction the least element w_0 of W not isomorphic to an ordinal. Then $W(w_0) \simeq \beta$ for some ordinal β and consequently the least ordinal α_0 for which $\beta < \alpha_0$ holds must be isomorphic to w_0.

Finally we let γ be the least ordinal such that $\gamma \notin \mathbf{Im}(\mathbf{F}|_W)$ and have that γ is order-isomorphic to W as every $\alpha \in \gamma$ is isomorphic to an initial segment of W. □

This theorem makes it possible for us to associate the order type of a well-ordered sets with precisely the ordinal it is order-isomorphic to. In that sense we use the terms order type and ordinal interchangeably when talking about well-ordered sets.

The idea of ordinals was introduced using the natural numbers and while that is a useful comparison, we want to think about ordinals more as a generalization of $\mathbb{N}$, rather that a direct analog. As the name implies, ordinals describe magnitudes of *order* rather than *size*.

Consider for example the ordinal $\omega + 1 = \omega \cup \{\omega\}$, which describes the element coming after the "number" which is larger that every natural number. We can still define a bijective function from $\omega + 1$ to the natural numbers, so $\omega + 1$ is not larger in a *cardinal* sense, but rather relates to order. We will not deal with the proper notion of cardinality, the size of sets, in this text, however we will see in Theorem 2.5.2 that the distinction between cardinality and ordinals is not always clear cut.

In $\omega + 1$ and ω we can also recognize the two important types of ordinals, *successor ordinals* and *limit ordinals*.

Definition 2.3.6 (Successor Ordinal)**.** [Jec78, p.13] We call an ordinal α a *successor ordinal*, if it is the direct successor

$$\alpha = \beta^+ = \beta + 1$$

for some other ordinal β.

Definition 2.3.7 (Limit Ordinal)**.** [Jec78, Exercise 2.3] We call an ordinal α a *limit ordinal* if it is not a successor ordinal. Then $\alpha = \bigcup \alpha$ is the least upper bound of the set $\{\beta \mid \beta < \alpha\}$ and

$$\beta < \alpha \implies \beta + 1 < \alpha.$$

The latter follows since $\beta + 1 \not< \alpha$ would imply that both $\beta \in \alpha$ and $\beta \in \beta + 1$ hold. But since α is not a successor ordinal by definition we also have $\alpha \neq \beta + 1$. This means $\beta + 1 \in \alpha$ must be true by contradiction, because ordinals are well-ordered by set inclusion.

We also say consider 0 a limit ordinal and say that the least upper bound of 0 is itself.

We now introduce the final concept important for ordinals, that of *transfinite induction*. When we do a proof by (regular) induction, we show that some property holds for every natural number. Transfinite is the same concept extended to the ordinal numbers; loosely speaking we show not only that a property holds for a base case and every successive number, but also that if it holds for all ordinals smaller than some limit ordinal, that it holds for the limit ordinal as well. This concept is generalized in the following theorem:

Theorem 2.3.8 (Transfinite Induction)**.** *[Jec78, Theorem 3] Let* $\mathbf{C}$ *be a class containing only ordinals and let the following hold for* $\mathbf{C}$*:*

(i) $0 \in \mathbf{C}$*,*

(ii) If $\alpha \in \mathbf{C}$ *is an ordinal, then* $\alpha + 1 \in \mathbf{C}$*,*

(iii) If α *is a non-zero ordinal and* $\beta \in \mathbf{C}$ *holds for all* $\beta < \alpha$*, then* $\alpha \in \mathbf{C}$*.*

Then $\mathbf{C}$ *is the class of all ordinals.*

Proof. Let *Ord* be the class of all ordinals. Assume for a contradiction that the theorem does not hold and assume that α is the least ordinal contained in the class $Ord \setminus \mathbf{C}$.

If $\alpha = 0$ we immediately arrive at a contradiction, hence assume that α is some non-zero ordinal. If α is a successor ordinal we have that its direct predecessor must be a member of $\mathbf{C}$ and by the second criteria of the theorem we have that $\alpha \in \mathbf{C}$.

Similarly, if α is a limit ordinal we have that $\beta \in \mathbf{C}$ for all $\beta < \alpha$. Therefore, by the third criteria of the theorem, we must also have that $\alpha \in \mathbf{C}$, a contradiction. Hence $Ord = \mathbf{C}$. □

2.4 The Well-Ordering Theorem

The following, along with *Zorn's Lemma*, is one of the most fundamental results in set theory. There is a (bad) joke that goes:

> The *Axiom of Choice* is obviously true, the *Well-Ordering Theorem* obviously false, and who knows with *Zorn's Lemma*.

Definition 2.4.1 (Zermelo's Well-Ordering Theorem)**.** [Jec78, Theorem 15, p.39] Every set can be well ordered.

We do not provide a proof for Definition 2.4.1 in **ZFC** here directly. This theorem, as it turn out, is not just another regular theorem, and we will therefore also not treat it as one.

Indeed, the Well-Ordering Theorem is actually equivalent to the Axiom of Choice. This means that if either statement is assumed to be true (and it has to be assumed since we are talking about *axioms*), the other one can be proved from it. This is the same methodology we will use for proving our main result, Theorem 4.0.1, as well. There we will show equivalence of our main statement, that a group structure exists on all arbitrary sets, with the Well-Ordering Theorem. As such by transitivity, this main statement is also equivalent to the Axiom of Choice.

Theorem 2.4.2. *[Jec78, Theorem 15, p.39] The Well-ordering Theorem is equivalent to the Axiom of Choice.*

Proof. We provide a proof in two parts; first showing that the Well-Ordering Theorem is true in **ZFC**. Then, conversely, we prove **AC** in **ZF**, assuming that the Well-Ordering Theorem holds true.

1. **Axiom of Choice $\implies$ Well-Ordering Theorem**

   We proceed by transfinite induction.

   Let A be an arbitrary set and let $S = \mathcal{P}(A) \setminus \varnothing$ be the collection of all non-empty subsets of A. Let $f : S \to A$ be a choice function (as specified by the Axiom of Choice). We then define an ordinal sequence $(a_\alpha \mid \alpha < \theta)$ the following way:

   $$\begin{aligned} a_0 &= f(A) \\ a_\alpha &= f\left(A \setminus \left\{a_\xi \mid \xi < \alpha\right\}\right) \qquad \text{if } A \setminus \left\{a_\xi \mid \xi < \alpha\right\} \text{ is non-empty.} \end{aligned}$$

   Now let θ be the smallest ordinal such that $A = \left\{a_\xi \mid \xi < \theta\right\}$.

   We know that such an ordinal must exist, since the sequence $(a_\alpha \mid \alpha < \theta)$ is entirely defined by the choice function f. The function f maps every non-empty subset of A, i.e. members of S, to an element of that subset (in A).

   By defining the ordinal sequence the way we did, it is not possible for any element of A to occur in the sequence twice. Any subset of A, which is the input of the choice function for some element a_γ in the sequence, does not contain any elements a_α for $\alpha < \gamma$, and by definition f cannot map to any of these members.

   As such $\mathbf{Im}\left((a_\alpha \mid \alpha < \theta)\right) = A$ and $(a_\alpha \mid \alpha < \theta)$ enumerates A, meaning the sequence is a bijection.[2] Hence A can be well-ordered, the least element of any subset being the one which corresponds to the smallest ordinal in the sequence.

2. **Well-Ordering Theorem $\implies$ Axiom of Choice**

   Let S be a set of non-empty sets.

   The union $\bigcup S$ can be well-ordered by assumption and clearly $s \in S$ implies $s \subseteq \bigcup S$. We can then define the function $f : S \to \bigcup S$ to map any elements of S to its least element, according to the well-order of $\bigcup S$.

   Evidently f is a choice function and since the set S was arbitrary the Axiom of Choice holds. □

2.5 Hartogs' Lemma

We continue with the conclusion of this chapter, a lemma originally stated by Hartogs in a paper from 1915. Just as with Theorem 2.4.1, this is used in the proof of our final

[2] Recall that a sequence is just a function from $\mathbb{N}$, respectively an ordinal, to the set of its elements

result, Theorem 4.0.1. Since our goal is to prove an equivalence with the Axiom of Choice, we will work completely in **ZF** without **AC** in this section.

The lack of **AC** is also the reason the comparability of sets is not given, another equivalence of the Axiom of Choice. [Har15] As such we do not talk about cardinality in the traditional sense in the lemma, even though the terminology is very similar.

Lemma 2.5.1. *[Har15, Annex] Let M be an arbitrary set. Then there exists a set M_Ω consisting of all well-ordered subsets of M.*

Proof. We follow the proof from [Har15] and adjust it to fit our own set-theoretic definition of orderings.

Recall that an ordering of M is defined as a subset $R \subseteq M \times M$. We first construct the set M_ω of all possible well-orderings of the set M itself. Note that since we cannot assume the Well-ordering Theorem, the set M_ω might also be the empty set.

We can define formulas $\phi_1, \dots, \phi_4$ with free variables M and R, such that these hold true only for well-order relations R of a given set M (in the right context). Then

$$\begin{aligned}
\phi_1(R, M) &= \forall x \in M \ \langle x, x\rangle \notin R \\
\phi_2(R, M) &= \forall x, y, z \in M \ \left(\langle x, y\rangle \in R \wedge \langle y, z\rangle \in R\right) \to \langle x, z\rangle \in R \\
\phi_3(R, M) &= \forall x, y \in M \ \left(\langle x, y\rangle \in R \vee x = y \vee \langle y, x\rangle \in R\right) \\
\phi_4(R, M) &= \forall S \in \mathcal{P}(M) \, \exists m \in S \ (\forall x \in S \ (\langle m, x\rangle \in R \vee m = x))
\end{aligned}$$

are all valid formulas in **ZF**, with ϕ_1, ϕ_2, ϕ_3 utilizing only the Axiom of Pairs and ϕ_4 using the Axiom of Pairs and the Axiom of Power Sets.

Then, fixing the set M, we have that

$$M_\omega = \left\{R \in \mathcal{P}(M \times M) \mid \phi_1(R, M) \wedge \phi_2(R, M) \wedge \phi_3(R, M) \wedge \phi_4(R, M)\right\}$$

is the initially desired set of all well-orders of M and a valid construction following from the Axiom Schema of Replacement and Axiom of Power Sets.

We can also see that we are able to construct such a set X_ω for any arbitrary set X. Hence, finally, given a set M, the set of all possible well-orderings of any subsets m of M is given by the union

$$M_\Omega = \bigcup_{m \in \mathcal{P}(M)} m_\omega.$$

Note that M_Ω is non-empty, since we are always able to define a well-order for finite subsets of M. (Compare with the proof of Theorem 2.4.2, however we do not need **AC** for a choice function on finite sets.) □

With this lemma in mind, we now continue to the desired statement. We follow the outline given in [HK72] for the proof of this lemma, the more detailed steps are taken from [Har15].

Lemma 2.5.2. *[Har15][HK72, Lemma] Let A be an arbitrary set. Then there exists an ordinal α, such that no mapping from a subset $S \subseteq A$ to the ordinal α is a bijection.*

Proof. We let the ordinal α take the following value:

$$\alpha = \bigcup \left\{ \text{type}\,(X, R) + 1 \,|\, X \subseteq A, R \subseteq A \times A \wedge R \text{ well-orders } A \right\}. \tag{2.5.3}$$

We will show that α is such that it satisfies the lemma's statement.

We can see that the right hand side of the set is precisely the set described in Lemma 2.5.1. As such we can express the ordinal in (2.5.3) as

$$\alpha = \bigcup \left\{ \text{type}\,(R) + 1 \,|\, R \in A_\Omega \right\},$$

where type (R) is the order type of each well-order R[3] in the set A_Ω of all well-orderings of subsets of A.

Recall that we associate the order type of a well-ordered set with the unique ordinal it is isomorphic to, motivated by Theorem 2.3.5. Hence α describes the smallest ordinal which is larger (in an ordinal sense) than all well-orderable subsets of A. The latter property follows from the fact that such an ordinal is either an successor ordinal or a limit ordinal, the term "+1" and the union symbol in the definition of α make sure that both these cases are accounted for.

By construction we now have that every subset associated with an order R in A_Ω is isomorphic to an initial segment of α, but not to α itself. Further, there can not exist a bijective map from an arbitrary subset $S \subseteq A$ to α.

If such a map were to exist, we could define a well-ordering of S according to α; however we have already established that S either has no well-order or that there exists a bijection to an initial segment of α. This is a contradiction and we must have that no injective function from α to A exists, satisfying the requirements of the lemma. □

[3] This is an abuse of notation, since we usually refer to the ordering R of a set X using (X, R). Here the underlying set was left out in order to more clearly convey what ordinal we are talking about.

However, if we only view the set structure of an ordering and how we defined ordered pairs, we have that $\bigcup \langle x, y \rangle = \bigcup \{\{x\}, \{x, y\}\} = \{x, y\}$. In a linear ordering every element relates to every other element in some way, therefore we have that $\bigcup \bigcup R = X$.

For this reason it is justified to leave out the underlying set of the well-order R.

Chapter 3

Model Theory

The aim of this chapter is to give an introduction to the basic tools of model theory with which we will prove the Löwenheim-Skolem Theorem. This theorem, divided into two parts, is dependent on the Axiom of Choice, hence we will later use it to show that **AC** implies the existence of a group structure on all non-empty sets.

We were already concerned with axiomatizations in the previous chapters, however we will formalize these notions some more here. The set theory axioms outlined in Chapter 1 still hold true, but we will discover more about what is and is not true in for example group theory or other algebraic structures. We start the treatment of model theory with the motivating example of "Ehrenfeucht-Fraïsse games", exemplifying how the field relates to the linear orderings from Chapter 2. The impatient or familiar reader however may skip this section without missing any necessary theory.

3.1 Ehrenfeucht-Fraïsse Games

Let us suppose we are given two linear orderings, based on which we define a two player game. Player I is called *spoiler* and starts by picking a point on one of the two orderings. If they pick an element of A we call this a_1 and if they pick an element of B we call it b_1. After the spoiler has picked their point, player II will pick a point on the other linear ordering. We call player II *duplicator*.[1] Say spoiler picked a_1 in A, duplicator then has to pick some element b_1 of B.

The spoiler and duplicator go back and forth picking points on the linear orderings for a predetermined amount of turns n. Spoiler always gets to choose the ordering they pick a point on and duplicator has to pick something on the other ordering. The second player, duplicator, wins if the elements $a_1, \ldots, a_n$ are in the same order with respect to A as the elements $b_1, \ldots, b_n$ are with respect to B. Spoiler wins if duplic-

[1] The names "spoiler" and "duplicator" were coined by Joel Spencer.[HV24, §6] The literature [Ros82] this section of the text is mainly based on uses "Player I" and "Player II" instead.

ator loses, if the elements are not in the same order. This is how a single play of an *Ehrenfeucht-Fraïsse game*, also called a *Back-and-forth game*, in n turns is played out. [Ros82, §6.1]

Example 3.1.1. We will go through the play of a game with 3 steps. Let $A = \mathbb{Q}$ and $B = \mathbb{Z}$ under the usual order be the two linear orderings that are played on.

Spoiler starts and picks the point $a_1 = 0$ in $\mathbb{Q}$. Duplicator on their turn also picks the point $b_0 = 0$ but in $\mathbb{Z}$. On the next turn spoiler picks $a_2 = 1$ in $\mathbb{Z}$, duplicator again matches this and picks $b_2 = 1$ in $\mathbb{Q}$. Now however spoiler can and does pick b_3 in $\mathbb{Q}$ to be $\frac{1}{2}$. This is not possible for duplicator to match since they are now confined to picking a point a_3 in $\mathbb{Z}$. Hence no matter what point duplicator picks to be a_3, they will lose.

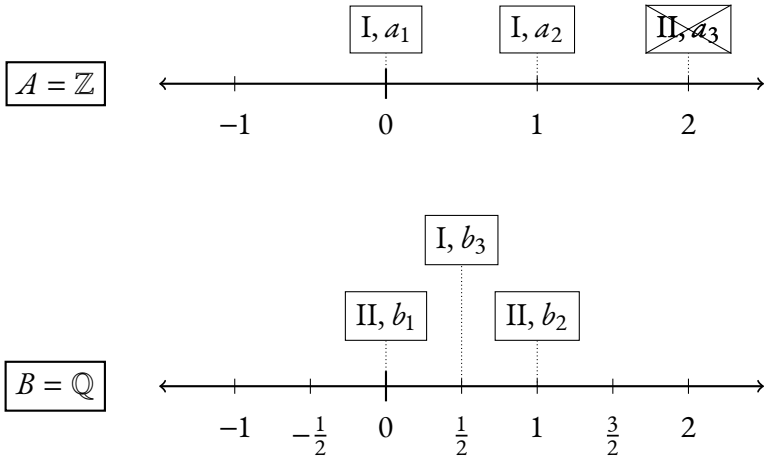

Figure 3.1: Example 3.1.1 illustrated. Points *spoiler* picked are prefixed by "I", points *duplicator* picked are prefixed by "II". It can be seen that duplicator is not able to pick a suitable element for b_3 in $\mathbb{Z}$ such that a_1, a_2 and a_3 have the same order as b_1, b_2 and b_3.

It should be clear that in order for spoiler to win, they need to exploit the differences of the linear orders in play. In the example above spoiler can alway pick their first two elements to be a_1 and a_2 in $\mathbb{Z}$ in a way where a_2 is the immediate successor of a_1. Then, because $\mathbb{Q}$ is dense, no matter which elements duplicator picks to be b_1 and b_2, spoiler can pick b_3 such that $b_1 < b_3 < b_2$. As a_2 is the immediate successor of a_1, there exists no element in $\mathbb{Z}$ satisfying $a_1 < a_3 < a_2$.

This is called a *winning strategy* for spoiler. A winning strategy for duplicator would be the converse, where duplicator can react according to every possible move by spoiler and still win. For duplicator to have a winning strategy, the linear orderings should be *similar* in some sort of way, in fact the existence of a winning strategy by

duplicator is a type of equivalence in itself. Such a winning strategy shown in the example below, where we play a game with two dense orderings.

Example 3.1.2. We play an Ehrenfeucht-Fraïsse game in n rounds.

Let $A = \mathbb{R}$ and $B = \mathbb{Q}$ under the usual order. We can define a winning strategy for duplicator inductively. Duplicator then picks a point according to one of the following cases on an arbitrary turn i:

1. It is the first turn and spoiler picks $a_1 \in A$.

   Duplicator then picks $b_1 = 0$.

2. Spoiler picks $a_i \in A$ such that $a_j < a_i < a_k$ holds for some $a_j,\ a_k \in A$.

   It can be assumed without loss of generality that a_j and a_k are the respectively largest and smallest points for which this property holds. Then there exist corresponding $b_j,\ b_k \in B$ with $b_j < b_k$. Hence duplicator can choose

   $$b_i = \frac{b_k - b_j}{2}.$$

3. Spoiler picks $a_i \in A$ such that a_i is either larger or smaller than all a_k for $k < i$.

   Duplicator can pick

   $$b_i = b_s - 1 \qquad \text{or} \qquad b_i = b_l + 1$$

   in B where b_s and b_l are the smallest and largest elements of $\{b_k \mid k < n\}$ respectively.

Since $\mathbb{Q}$ is a proper subset of $\mathbb{R}$ the same strategy works if spoiler chooses an element b_i in $B = \mathbb{Q}$. The strategy for the second turn follows immediately from the third case.

With this example in mind we now formally define the play of an Ehrenfeucht-Fraïsse game and what it means for duplicator to have a winning strategy.

Definition 3.1.3 (Ehrenfeucht-Fraïsse Game)**.** [Ros82, Definition 6.2] Let A and B be linear orderings and let $n \in \mathbb{Z}^+$ be a fixed positive integer. Then a *play of an Ehrenfeucht-Fraïsse game* $EF_n(A, B)$ is defined as an ordered sequence with $2n$ elements. For each positive integer $k \leq n$ player I, *spoiler*, chooses an element in either A or B. Player II, *duplicator*, then chooses an element of the ordering spoiler did not pick. We denote an element at the kth turn as a_k if it was picked from A and b_k if it was picked from B.

We say that duplicator has won a play of the game $EF_n(A, B)$ if for all positive integers $i, j \leq n$

$$a_i <_A a_j \iff b_i <_B b_j$$

with respect to the linear orders $<_A$ and $<_B$ of A and B. We say that spoiler has won a play of the game if duplicator did not win.

We say that there exists *winning strategy of duplicator* if there exists a sequence of functions $f_1, \ldots, f_n$, which fulfills the following requirements:

(i) The domain of each f_k is the set of all ordered k-tuples with elements in $A \cup B$;

(ii) For elements $c_1, \ldots, c_k \in A \cup B$, representing the first k moves by spoiler, f_k satisfies

$$f_k(c_1, \ldots, c_k) \in A \text{ if } c_k \in B \text{ and}$$
$$f_k(c_1, \ldots, c_k) \in B \text{ if } c_k \in A.$$

(iii) For elements $c_1, \ldots, c_k \in A \cup B$, f_k, a_k and b_k are defined as

$$a_k = \begin{cases} c_k & \text{if } c_k \in A \\ f_k(c_1, \ldots, c_k) & \text{if } c_k \in B \end{cases} \quad \text{and} \quad b_k = \begin{cases} c_k & \text{if } c_k \in B \\ f_k(c_1, \ldots, c_k) & \text{if } c_k \in A \end{cases}$$

for positive integers $k \leq n$. These terms, based on $f_1, \ldots, f_n$, then satisfy

$$a_i <_A a_j \iff b_i <_B b_j$$

for all $i, j \leq n$, where $i, j \in \mathbb{Z}^+$.

If no such sequence of functions exists, we say that there exists a *winning strategy of spoiler*.

The proper definition of a winning strategy lets us define the following type of equivalence:

Definition 3.1.4 (EF_n-equivalence)**.** [Ros82, Definition 6.8] We say that a linearly ordered set A is EF_n-equivalent to a linearly ordered set B, if there exists a winning strategy of duplicator in $EF_n(A, B)$.

The orderings A and B are called EF-equivalent if duplicator has a winning strategy in $EF_n(A, B)$ for all $n \in \mathbb{N}$.

This is indeed a proper type of equivalence, since by the Gale-Stewart Theorem from game theory all Ehrenfeucht-Fraïsse games are determined.[HV24, §6] This means that one, and only one, of the players will always have a winning strategy.

We saw in Example 3.1.2 that $\mathbb{R}$ and $\mathbb{Q}$ are EF-equivalent. We were able to find a winning strategy for duplicator because both ordered sets were dense. As such we could always find a point between any other two points in the rational numbers that could match the order of the possibly irrational points in $\mathbb{R}$. The sets $\mathbb{R}$ and $\mathbb{Q}$ are of course different, but they still *satisfy the same property* of densely linear orders.

This notion of what structures satisfy which properties is at the core of model theory. And its no coincidence that Ehrenfeucht-Fraïsse games seem give rise so naturally to the field. We do indeed have a theorem relating them to model theory proper.

Theorem 3.1.5. *[Maro2, Theorem 2.4.6] Let $\mathcal{M}$ and $\mathcal{N}$ be two $\mathcal{L}$-structures, where $\mathcal{L}$ is a finite language without any function symbols. Then duplicator has a winning strategy in $EF_n(\mathcal{M}, \mathcal{N})$ for all $n \in \mathbb{N}$ if and only if $\mathcal{M}$ and $\mathcal{N}$ satisfy exactly the same theories.*

We will not prove Theorem 3.1.5 here, but rather include it as an appetizer; a means show the usefulness of model theory applied to problems we may already be interested in. The reader interested in the proof of the theorem may read it in the sub-subsection titled "Ehrenfeucht-Fraïsse Games" in [Maro2, §2.4]. The rest of this chapter should provide an adequate background for understanding it.

3.2 Syntax of Formal Languages

The rest of this chapter follows mostly from [CK90, §§1.3, 1.4, 2.1], where the overall structure is preserved, although we reformulate and expand on the definitions, theorems and proofs to better fit our context. Especially some of the earlier definitions are also lifted from [Maro2].

We utilize set theoretic notation in the definition of languages, however we do not consider the set inclusion symbol $\in$ to be to be part of first-order logic. In general we are working within **ZFC** in this chapter.

Definition 3.2.1. A *formal language* $\mathcal{L}$ in first order logic is associated with the following sets:

1. A set $\mathcal{F}$ of function symbols f with n_f arguments each, where $n_f \in \mathbb{Z}^+$ is a positive integer;
2. a set $\mathcal{R}$ of n_R-ary relation symbols R, where $n_R \in \mathbb{Z}^+$ is a positive integer,
3. a set $\mathcal{C}$ of constant symbols. [Maro2, Definition 1.1.1]

The set-theoretic structure of $\mathcal{L}$ is $\mathcal{L} = \mathcal{F} \cup \mathcal{R} \cup \mathcal{C}$. If $|\mathcal{L}|$ is the cardinality of the set $\mathcal{L}$ we say that the *power* of the language $\mathcal{L}$ is

$$\|\mathcal{L}\| = |\mathcal{L}| \cup \omega.$$

We say that $\mathcal{L}$ is either countable or uncountable depending on whether $\|\mathcal{L}\|$ is countable or uncountable. [CK90, §1.3]

The language of pure sets is the empty language $\mathcal{L}_{Set} = \emptyset$. In order to make meaningful statements regarding sets however, we need the set inclusion symbol $\in$, which we view as a relation. Therefore the language of set theory is $\mathcal{L}_{\mathbf{ZFC}} = \{\in\}$. [Mar02, p.8][Jec78, p.80]

Since languages are given by the single set of the symbols they contain, we have that normal set operations are applicable. If $\mathcal{L}$ and $\mathcal{L}'$ are two languages, such that $\mathcal{L} \subset \mathcal{L}'$, we say that $\mathcal{L}'$ is the *expansion* of $\mathcal{L}$ and that $\mathcal{L}$ is the *reduction* of $\mathcal{L}'$. In the case that $\mathcal{L}'$ is of the form $\mathcal{L}' = \mathcal{L} \cup X$, where X is a set containing only constant symbols, we say that $\mathcal{L}'$ is the *simple expansion* of $\mathcal{L}$. [CK90, p.19]

In order to express statements in a language we also define *terms*. These are are expressions created from chaining together function symbols with variables and constants.

Definition 3.2.2. [Mar02, Definition 1.1.4] We say that T is the set of $\mathcal{L}$-terms, if T is the smallest set such that

(i) $c \in T$ for each constant $c \in \mathcal{C}$

(ii) $v_i \in T$ for variable symbols v_i, where $i = 1, 2, \ldots,$

(iii) $f(t_1, \ldots, t_n) \in T$ for $f \in \mathcal{F}$ and terms $t_1, \ldots, t_n \in T$.

Example 3.2.3. Consider the language of (unitary) rings $\mathcal{L}_r = \{+, -, \cdot, 0, 1\}$. The symbols $+, -, \cdot$ are all binary functions, and 0 and 1 are constants. Since a language does not possess any external structure itself, this is also the language of fields.

For variables v_1 and v_2 we can define the term

$$\cdot\,(+\,(v_1.\,1)\,,\ +\,(v_2,\ +\,(1,\ 1)))$$

in $\mathcal{L}_r$. In more familiar notation this would be

$$(v_1 + 1) \cdot (v_2 + 1 + 1)\,.$$

Note that this term has can have a different meaning depending on which algebraic structure we evaluate it in. In the finite field $\mathbb{F}_2$ for example, we have that $1 + 1 = 0$.

We define formulas in a language $\mathcal{L}$ analogously to how we did in Chapter 1. Note that our definition of a formula in $\mathcal{L}_{\mathbf{ZFC}}$ coincides with how we defined it before.

Definition 3.2.4 (Formula in a Language)**.** [Mar02, Definition 1.1.5] An *atomic formula* in a language $\mathcal{L}$ is either

1. $t_1 = t_2$ for terms $t_1,\ t_2 \in T$, or
2. $R(t_1, \ldots, t_{n_R})$ for $R \in \mathcal{R}$, and terms $t_1, \ldots, t_{n_R} \in T$.

The set of $\mathcal{L}$-formulas is then the smallest set such that

(i) for all $\phi \in \mathcal{W}$, the negation $\neg\phi$ is also in $\mathcal{W}$,

(ii) for all $\phi,\ \psi \in \mathcal{W}$, the formulas $\phi \wedge \psi$ and $\phi \vee \psi$ are also in $\mathcal{W}$,

(iii) for all $\phi \in \mathcal{W}$, the formulas $\forall v_i\, \phi$ and $\exists v_i\, \phi$ are also in $\mathcal{W}$.

with respect to the logical connectives

$$\neg \text{ (not)},\ \wedge \text{ (and)},\ \vee \text{ (or)},\ \rightarrow \text{ (implies) and } \leftrightarrow \text{ (if and only if)}$$

and quantifiers

$$\forall \text{ (for all) and } \exists \text{ (exists).}$$

Recall that variables occurring inside of a quantifier are called *bound* and that variables which are not bound are called *free*. A formula with no free variables is called a *sentence*.

Note that some of the logical connectives can be considered shorthand of others. For formulas ϕ and ψ we say that $\phi \vee \psi$ is defined as $\neg\left(\neg\phi \wedge \neg\psi\right)$. We can then express $\phi \rightarrow \psi$ as $\neg\phi \vee \psi$ and $\phi \leftrightarrow \psi$ as $\left(\phi \rightarrow \psi\right) \wedge \left(\psi \rightarrow \phi\right)$. Lastly, for a variable v, the formula $\forall v\, \phi(v)$ is considered an abbreviation of $\neg\left(\exists v\, \neg\phi\right)$[2]. [Mar02, Remark 1.1.7]

This is a more technical definition of formulas compared to how we defined them in the Preliminaries. It comes from the fact that whenever we want to prove something something is true for all formulas in a language, we can use a proof by induction utilizing Definition 3.2.4. Using the shorthand formulas we only have to consider the cases $\neg\phi$, $\phi \wedge \psi$ and $\exists v\, \phi(v)$, for formulas ϕ and ψ and some variable v.

3.2.1 Logical Axioms

When working in model theory we consider the difference between *syntax* and *semantics*. Syntax is everything relating to structure of sentences and formal languages. This includes what kind of symbols we are working with in a given language and how we can combine them to create valid sentences.

A language is purely *syntactical*, a given sentences in a language has no assigned meaning on its own. If we are working with a language describing number theory for example, we might have a relation symbol *gcd* which takes in three arguments and returns whether the third argument is the greatest common divisor of the first two. A formal language only includes the relation symbol itself though, elements for which relations would hold are not part of the language.

[2] Alternatively, assuming that $\forall$ is a quantifier, we could also define $\exists v\, \phi(v)$ as $\neg\,(\forall v)\ \neg\phi(v)$. [CK90, p.23]

Assigning a meaning to a relation symbol such as *gcd* is *semantics*. Having a relation be true or false for a some given value is a *semantical* property, this is where the namesake models come in. In a model the symbols of a language are assigned meaning and we can explore which sentences are or are not true under some given assumptions.[CK90, p.3]

There are some sentences which are true in every language however, $x = x$ for example. We formalize this notion with the following *logical axioms* and *rules of inference*. This is needed in order to make the syntactical notion of formulas into a *formal system*. In essence, we define the kind of formulas which have to be true in every model. This is also why the *axiom* label is justified.

Definition 3.2.5 (Tautology)**.** [CK90, pp.8, 24] Let $\mathcal{F}$ be a set of simple sentence statements and let ϕ be a formula in $\mathcal{F}$. Let $\bar{a} = a_1, \ldots a_n$ be a sequence of truth values, meaning each a_i either has the value t for *true* or f for *false*. For a such a sentence ψ in $\mathcal{F}$ containing sentence symbols $S_1, \ldots, S_n$, an assignment of $\bar{a}$ into ψ substitutes each S_i with a_i.

We say that ψ is a *tautology in sentential logic* if and only if the value of ψ is t for every assignment $\bar{a}$.

We say that a formula ϕ of a formal language $\mathcal{L}$ is *a tautology of* $\mathcal{L}$, if it is possible to obtain ϕ by substituting sentence symbols S_i with formulas of $\mathcal{L}$ in a tautology in sentential logic ψ.

We can think of a formula in $\mathcal{F}$ as a placeholders $S_1, \ldots, S_n$ chained together with logical connectives. Then, when we evaluate it with a sequence of truth values $\bar{a}$, the formula "terminates" either in t or f.

Example 3.2.6. The formula

$$\tau = (S_1 \implies S_2) \vee (S_2 \implies S_1)$$

is a tautology in sentential logic. No matter which value of $\bar{a} = \langle a_1, a_2 \rangle$ we choose in the set $\{\langle t, t\rangle, \langle t, f\rangle, \langle f, t\rangle, \langle f, f\rangle\}$ we always have that $\tau(\bar{a}) = t$.

Let $\mathcal{L}_{or} = \{+, -, \cdot, <, 0, 1\}$ be the language of ordered rings. Suppose that ϕ_1 and ϕ_2 are the two $\mathcal{L}_{or}$-formulas $v_1 + 1 < v_2$ and $v_2 \cdot v_1 = v_1$ respectively, each having (the same) free variables v_1 and v_2. If we substitute the formulas ϕ_1 and ϕ_2 into τ, we have that

$$\begin{aligned}\tau_r &= (\phi_1 \implies \phi_2) \vee (\phi_2 \implies \phi_1) \\ &= ((v_1 + 1 < v_2) \implies (v_2 \cdot v_1 = v_1)) \\ &\quad \vee ((v_2 \cdot v_1 = v_1) \implies (v_1 + 1 < v_2))\end{aligned}$$

is a tautology in $\mathcal{L}_{or}$. Note that for example

$$(v_1 < v_2) \vee (v_1 = v_2) \vee (v_2 < v_1)$$

is not a tautology in $\mathcal{L}_{or}$. Intuitively, we can explain this by observing that there are structures of the language which might not interpret the symbol $<$ as a linear order.

Definition 3.2.7. We have the three groups of *logical axioms* and two *rules of inference*.

1. **Sentential Axioms**

   Every tautology ψ of $\mathcal{L}$ is a logical axiom.

2. **Quantifier Axioms**

   a) Let ϕ, ψ be formulas in $\mathcal{L}$ and let v be a variable which is not free in ϕ. Then
   $$\forall v \left((\phi \implies \psi(v)) \implies (\phi \implies \forall v \, \psi(v)) \right)$$
   is a logical axiom.

   b) Let $\phi(v)$ be a formula in $\mathcal{L}$ with a free variable v. Let ψ be a formula in $\mathcal{L}$, such that ψ is the sentence where all free occurrences v in ψ have been substituted by some term t. This means that no variable of t is bound in ψ at the place it is introduced. Then
   $$\forall v \, \phi(v) \implies \psi$$
   is a logical axiom.

3. **Identity Axioms**

   Let x and y be variables and let $t(v_1, \ldots, v_n)$ and $\phi(v_1, \ldots, v_n)$ be a term and an atomic formula respectively, both dependent on free variables $v_1, \ldots, v_n$.

   Then

   a) $x = x$,

   b) $x = y \implies$
   $\left(t(v_1, \ldots, v_{i-1}, x, v_{i+1}, \ldots, v_n) = t(v_1, \ldots, v_{i-1}, y, v_{i+1}, \ldots, v_n) \right)$,

   c) $x = y \implies$
   $\left(\phi(v_1, \ldots, v_{i-1}, x, v_{i+1}, \ldots, v_n) \implies \phi(v_1, \ldots, v_{i-1}, y, v_{i+1}, \ldots, v_n) \right)$

   are logical axioms.

The *Rules of Inference* are then

4. **Rule of Detachment (Modus Ponens)**

   If ψ and $\phi \to \psi$ hold, then ψ holds.

5. **Rule of Generalization**

   If ψ holds, then $\forall x\, \psi(x)$ holds.

Definition 3.2.8 (Proof)**.** [CK90, p.25] Let $\mathcal{L}$ be a language and let Σ be a set of sentences in $\mathcal{L}$. We write $\vdash \phi$ to say that a sentence ϕ is a theory in $\mathcal{L}$. Further, we write $\Sigma \vdash \phi$ to say that the sentence ϕ *can be proven from* Σ (in conjunction with the logical axioms). Another way to say this is to say that ϕ *is deductible from* Σ or that *there exists a proof of* ϕ *from* Σ.

A theory can thought of as something that can be proven only from the logical axioms. While Definition 3.2.8 introduces concepts the proof theory, we will in general not be quite that formal when proving sentences in model theory. However, this type of formality is needed for the next definition.

Definition 3.2.9 (Consistency)**.** [CK90, p.25] A set of sentences Σ is *inconsistent* in a language $\mathcal{L}$ if and only if every formula in $\mathcal{L}$ can be proven from Σ. The set Σ is called *consistent* if it is not inconsistent. We have say that a sentence σ is consistent in $\mathcal{L}$ if the singleton set $\{\sigma\}$ is consistent.

We say that Σ is *maximally consistent* if and only if Σ is consistent and no set of sentences Σ' which has Σ as a proper subset is consistent.

Theorem 3.2.10 (Lindenbaum's Theorem)**.** *[CK90, Proposition 1.3.11] Let Σ be a consistent set of sentences in a language $\mathcal{L}$. Then Σ can be extended to a maximally consistent set of sentences in $\mathcal{L}$.*

Proof. [CK90, Lemma 1.2.9] TODO □

3.3 Structes, Theories and Models

Definition 3.3.1. A *$\mathcal{L}$-structure* or $\mathcal{M}$ is given by the following:

1. A nonempty set M,
2. a function $f^{\mathcal{M}} : M^{n_f} \to M$ for each $f \in \mathcal{F}$,
3. a set $R^{\mathcal{M}} \subseteq M^{n_R}$ for each $R \in \mathcal{R}$,
4. an element $c^{\mathcal{M}} \in M$ for each $c \in \mathcal{C}$.

The set M is referred to as the *universe*, *domain* or *underlying set* of $\mathcal{M}$ and $f^{\mathcal{M}}$, $R^{\mathcal{M}}$ and $c^{\mathcal{M}}$ are called the *interpretations* of $\mathcal{M}$. The *cardinality of* $\mathcal{M}$ is the cardinality of its underlying set $|M|$. [Mar02, Definition 1.1.2]

We sometimes also identify a structure $\mathcal{M}$ by writing $(M, f^{\mathcal{M}}, R^{\mathcal{M}}, c^{\mathcal{M}})$ [Mar02] or $(M, \mathcal{I})$, where $\mathcal{I}$ is the function mapping symbols of $\mathcal{L}$ to their respective interpretations in M. [CK90, p.20]

Definition 3.3.2. [Mar02, Definition 1.1.3] Let $\mathcal{M}$ and $\mathcal{N}$ be two $\mathcal{L}$-structures, having the underlying sets M and N respectively. We call a function $\mu : M \to N$ an $\mathcal{L}$-embedding if it is injective and preserves the interpretation of all symbols of $\mathcal{L}$.

More precisely this means that μ satisfies

(i) $\mu\left(f^{\mathcal{M}}(a_1, \ldots, a_{n_f})\right) = f^{\mathcal{N}}(a_1, \ldots, a_{n_f})$ for all $f \in \mathcal{F}$ and all sequences $a_1, \ldots, a_{n_f}$,

(ii) $\langle a_1, \ldots, a_{n_R} \rangle \in R^{\mathcal{M}}$ if and only if $\langle \mu(a_1), \ldots, \mu(a_{n_R}) \rangle \in R^{\mathcal{N}}$ for all $f \in \mathcal{F}$ and $a_1, \ldots, a_n$,

(iii) $\mu(c^{\mathcal{M}}) = c^n$ for all $c \in C$.

In the case that $N \subseteq M$, we say that $\mathcal{N}$ is a *substructure* of $\mathcal{M}$ and that $\mathcal{M}$ is an *extension* of $\mathcal{N}$.

Example 3.3.3. [Mar02, p.8] Let $\mathcal{L}_G = \{\cdot, e\}$ be the language of groups, $\cdot$ is a binary function symbol and e is the constant symbol which we usually associate with the identity element.

Then $\mathcal{N} = (\mathbb{N}, +, 0)$ is a $\mathcal{L}_G$-structure, however it is not actually a group.

As discussed earlier, structures are where we explore *semantics*. To effectively explore semantical ideas in different structures however, we first need to actually define what we consider truth in a structure to actually be.

Say we have a term t with (free) variables $v_1, \ldots, v_n$, and we want to know what truth value t takes in some structure when we evaluate it at $\bar{a} = a_1, \ldots, a_n$. One might think that this is function evaluation in the same sense which we usually use it, and they would be mostly correct. However, since we will use Definition 3.2.4 of formulas in defining what truth means for a structure, we have to be similarly careful when defining term valuation.

Definition 3.3.4. [CK90, 1.3.13] Let t be a term in a language $\mathcal{L}$ dependent on variables $v_1, \ldots, v_n$. If $\mathcal{M} = \left(M, f^{\mathcal{M}}, R^{\mathcal{M}}, c^{\mathcal{M}}\right)$ is a $\mathcal{L}$-structure, we define the *value of the term* $t(v_1, \ldots, v_n)$ *at* $\bar{a} = \langle a_1, \ldots, a_n \rangle \in M^n$ the following way:

1. If t is v_i, then $t(\bar{a}) = a_i$,
2. if t is the constant symbol c in $\mathcal{L}$, then $t(\bar{a})$ is given by the interpretation $c^{\mathcal{M}}$,
3. if t is the function $f(t_1, \ldots, t_n)$ in $\mathcal{L}$, then $t(\bar{a})$ is given by the interpretation $f^{\mathcal{M}}(t_1(\bar{a}), \ldots, t_n(\bar{a}))$.

We can then define what it means for a sentence to be true in a $\mathcal{L}$-structure:

Definition 3.3.5. [Maro2, Definition 1.1.6] Let $\mathcal{M}$ be an $\mathcal{L}$-structure and let ϕ be a formula in $\mathcal{L}$ with free variables $\bar{v} = \langle v_1, \ldots, v_n \rangle$. We inductively define $\mathcal{M} \models \phi(\bar{a})$ and for $\bar{a} = \langle a_1, \ldots, a_n \rangle \in M^n$.

1. If ϕ is $t_1 = t_2$, then $\mathcal{M} \models \phi(\bar{a})$ if and only if $t_1(\bar{a}) = t_2(\bar{a})$,
2. If ϕ is $R\left(t_1, \ldots, t_{n_R}\right)$, then $\mathcal{M} \models \phi(\bar{a})$ if and only if $\left\langle t_1(\bar{a}), \ldots, t_{n_R}(\bar{a}) \right\rangle \in R^{\mathcal{M}}$,
3. If ϕ is $\neg\psi$, then $\mathcal{M} \models \phi(\bar{a})$ if and only if $\mathcal{M} \not\models \psi(\bar{a})$,
4. If ϕ is $\psi \wedge \rho$, then $\mathcal{M} \models \phi(\bar{a})$ if and only if $\mathcal{M} \models \psi(\bar{a})$ and $\mathcal{M} \models \rho(\bar{a})$,
5. If ϕ is $\exists v_i \, \psi(\bar{v}, v_i)$, then $\mathcal{M} \models \phi(\bar{a})$ if and only if there exists some element $b \in M$ such that $\mathcal{M} \models \psi(\bar{a}, b)$

We say that $\mathcal{M}$ *satisfies* $\phi(\bar{a})$, whenever $\mathcal{M} \models \phi(\bar{a})$.

3.4 The Löwenheim-Skolem Theorem

Lemma 3.4.1. *[CK90, Lemma 2.1.1] Let T be a consistent set of sentences of $\mathcal{L}$. Let C be a set of new constant symbols of power $|C| = \left\|\mathcal{L}\right\|$ and let $\bar{\mathcal{L}} = \mathcal{L} \cup C$ be the simple expansion of $\mathcal{L}$ formed by adding C.*

Then T can be expanded to a consistent set of sentences $\bar{T}$ in $\bar{\mathcal{L}}$, which has C as a set of witnesses in $\bar{\mathcal{L}}$.

Lemma 3.4.2. *[CK90, Lemma 2.1.2] Let T be a set of sentences and let C be a set of witnesses of T in $\mathcal{L}$. Then T has a model $\mathfrak{U}$, such that every element of $\mathfrak{U}$ is an interpretation of a constant $c \in C$.*

Theorem 3.4.3 (Extended Completeness Theorem)**.** *[CK90, Theorem 1.3.21] Let Σ be a set of sentences in $\mathcal{L}$. Then Σ is consistent if and only if Σ has a model.*

Theorem 3.4.4 (Downward Löwenheim-Skolem Theorem)**.** *[CK90, Corollary 2.1.4] Every consistent theory T in $\mathcal{L}$ has a model of power at most $\left\|\mathcal{L}\right\|$.*

Theorem 3.4.5 (Compactness Theorem)**.** *[CK90, Theorem 1.3.22] A set of sentences Σ has a model if and only if every finite subset of Σ has a model.*

Theorem 3.4.6 (Upward Löwenheim-Skolem Theorem)**.** *[CK90, Corollary 2.1.6] If T has infinite models, then it has infinite models of any given power $\alpha \geq \left\|\mathcal{L}\right\|$.*

Chapter 4

Hajnal's and Kertész's Theorem

We now arrive at our main theorem:

Theorem 4.0.1. *[HK72] The following are equivalent in ZF:*

1. *Axiom of Choice*
2. *Every non-empty set admits a cancellative groupoid structure*

Proof. The theorem is proven in two steps, deriving a single direction implication for each sentence.

1. **Groupoid Structure on arbitrary Sets $\implies$ Axiom of Choice**

   We show that the existence of a groupoid structure on every non-empty set implies that every set can be well-ordered. By Theorem 2.4.2 this is equivalent to the Axiom of Choice.

   Let A be an arbitrary set and let α be an ordinal as described in Theorem 2.5.2 in Section 2.5. This means that there exists no bijective mapping from α to any subset of A (including A itself). We then let $(B,\ R)$ be a well-ordered set of type α and such that $A \cap B = \emptyset$.

   Now let C be the set $C = A \cup B$, by assumption there exists some operation $+$, such that $(C,\ +)$ is a cancellative groupoid. We will show that for every $x \in A$ there exists $y \in B$, such that $x + y \in B$ holds.

   Let us assume for a contradiction that the above claim does not hold. This would imply that some $a \in A$ exists for which $a + y \in A$ holds for all $y \in B$. Let $f : B \to A$ be the function defined by $f(y) = a + y$. We have that $+$ is a cancellative groupoid operation, hence f must be injective; a contradiction by Theorem 2.5.2, since we had assumed that B is of type α.

We let $D = B \times B$ be the well-ordered set with respect to the lexicographical ordering R' of R, and define a function $g : A \to D$ by

$$g(x) = \min_{R'} \{\langle u, v\rangle \in D \mid x + u = v\}.$$

The function g maps every element x of A to the least pair $\langle u, v\rangle$ in $B \times B$ satisfying $x + u = v$. From earlier in the proof we know that such a pair must exists and that g must in fact be injective. This again follows from $+$ being cancellative, since if x_1, x_2 are two elements of A, having $f(x_1) = f(x_2)$ would imply that

$$\begin{aligned} & x_1 + u = v = x_2 + u \\ \iff & x_1 = v + u^{-1} = x_2 \end{aligned}$$

for some pair $\langle u, v\rangle \in D$. Since $\mathbf{Im}(g)$ is a subset of D it itself is a well-ordered set. As such we can define a well-order R'' on A by letting $x_i R'' x_j$ whenever $g(x_i) R' g(x_j)$.

2. **Axiom of Choice $\implies$ Groupoid Structure on arbitrary Sets**

   ...

□

Bibliography

[CK90] Chen Chung Chang and H. Jerome Keisler. *Model Theory*. North-Holland Publishing Co., Amsterdam, 3. edition, 1990.

[Gol98] D.C. Goldrei. *Classic set theory*. CRC Press, Boca Raton, 1998.

[Har15] F. Hartogs. Über das Problem der Wohlordnung. *Math. Ann.*, 76(4):438–443, 1915.

[HK72] A. Hajnal and A. Kertész. Some new algebraic equivalents of the axiom of choice. *Publ. Math. Debrecen*, 19:339–340, 1972.

[HV24] Wilfrid Hodges and Jouko Väänänen. Logic and Games. In Edward N. Zalta and Uri Nodelman, editors, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Winter 2024 edition, 2024.

[Jec78] Thomas Jech. *Set Theory*. Pure and applied Mathematics. Academic Press, New York, 1978.

[Ker75] Andor Kertész. *Einführung in die transfinite Algebra*. Birkhäuser, Basel, cop. 1975.

[Mar02] David Marker. *Model theory*. Graduate texts in mathematics. Springer, New York, cop. 2002.

[Ros82] Joseph G. Rosenstein. *Linear Orderings*. Pure and applied Mathematics. Academic Press, New York, 1982.