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 **??** 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 **??**. The impatient or familiar reader however may skip this section without missing any necessary theory.

0.1 Ehrenfeucht-Fraïsse Games

Let us suppose we are given two linear orderings called A and B, 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 duplicator 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 [5, §6.1].

Example 0.1.1. 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.

---

[1]The names "spoiler" and "duplicator" were coined by Joel Spencer [2, §6]. The literature [5] this section of the text is mainly based on uses "Player I" and "Player II" instead.

Figure 0.1: Example 0.1.1 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 0.1.2. 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 o.1.3 (Ehrenfeucht-Fraïsse Game). [5, 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 a *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, the function 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 $c_1, \ldots, c_k \in A \cup B$, the additional sequences of functions 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 0.1.4 (EF_n-equivalence)**.** [5, 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 all Ehrenfeucht-Fraïsse games are determined. This means that one, and only one, of the players will always have a winning strategy [2, §6].

We saw in Example 0.1.2 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 it is no coincidence that Ehrenfeucht-Fraïsse games seem to arise so naturally in the field. We do indeed have a theorem relating them to model theory proper.

Theorem 0.1.5. *[4, 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 0.1.5 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 subsection titled "Ehrenfeucht-Fraïsse Games" in [4, §2.4]. The rest of this chapter should provide an adequate background for understanding it.

0.2 Syntax of Formal Languages

The rest of this chapter follows mostly from [1, §§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 [4].

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

Definition 0.2.1. [4, Definition 1.1.1] 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 C of constant symbols.

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

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

where ω is the cardinality of the natural numbers. We say that $\mathcal{L}$ is either *countable* or *uncountable* depending on whether $\|\mathcal{L}\|$ is countable or uncountable [1, §1.3].

The language of pure sets is the empty language $\mathcal{L}_{Set} = \varnothing$. In order to make meaningful statements regarding sets however, we need the set inclusion symbol ∈, which we view as a relation. Therefore the language of set theory is $\mathcal{L}_{\textbf{ZFC}} = \{\in\}$ [4, p.8][3, 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 $\mathcal{L}$ [1, 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 0.2.2. [4, 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 each $f \in \mathcal{F}$ and terms $t_1, \ldots, t_n \in T$.

Notice that we did not include relation symbols R of $\mathcal{L}$ in Definition 0.2.2. We can think of terms as stand-in for values in the language. In contrast, we include relation symbols in Definition 0.2.4 of formulas below. Formulas always carry with them a notion of being either true or false, which is not the case for terms. This is reflected in the way we think about relations as well: a relation either holds or does not hold for a certain set of values.

Example 0.2.3. 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

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

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

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

Note that this term 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 **??**. Note that our definition of a formula in $\mathcal{L}_{\mathbf{ZFC}}$ coincides with how we defined it before.

Definition 0.2.4 (Formula in a Language). [4, 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 $\mathcal{W}$ 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

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

and quantifiers

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

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*.

An example of a formula $\phi(v)$ with a free variable v would be $\exists w \, (v = w)$. A simple example of a sentence would be $\forall x \, \exists y \, R(x, y)$, where R is a binary relation symbol.

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 $\neg \, (\neg \phi \wedge \neg \psi)$. We can then express $\phi \rightarrow \psi$ as $\neg \phi \vee \psi$ and $\phi \leftrightarrow \psi$ as $(\phi \rightarrow \psi) \wedge (\psi \rightarrow \phi)$. Lastly, for a variable v, the formula $\forall v \, \phi(v)$ is considered an abbreviation of $\neg \, (\exists v \, \neg \phi)$[2] [4, 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 is true for all formulas in a language, we can use a proof by induction utilizing Definition 0.2.4. 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.

Proposition 0.2.5. *[1, Proposition 1.3.4.] The set of all formulas of a language $\mathcal{L}$ has cardinality $\|\mathcal{L}\|$.*

0.2.1 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*, since any given sentence 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, the elements for which the relation 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 $\neg \, (\forall v) \, \neg \phi(v)$ [1, 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 [1, 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 0.2.6 (Tautology). [1, pp.4-8, 24] Sentential logic consists of sentence symbols, also called propositional variables, together with the logical symbol

$$\neg \,(\text{not}) \quad \text{and} \quad \wedge \,(\text{and}).$$

Every propositional variable S is a formula in sentential logic. Further, if ϕ is a formula in sentential logic, then $\neg \phi$ is also a formula. Then, if ψ is another formula, so is $\psi \wedge \phi$.

Let $\overline{a} = \langle a_1, \ldots, a_n \rangle$ be a sequence of truth values, meaning each a_i either has the value t for *true* or f for *false*. For a formula ψ in sentential logic containing the sentence symbols $S_1, \ldots, S_n$, an assignment of $\overline{a}$ 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 $\overline{a}$.

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. This means that if ψ has propositional variables $S_1, \ldots, S_n$, we can simultaneously and uniformly substitute these variables with formulas $\rho_1, \ldots, \rho_n$ in $\mathcal{L}$ to arrive back at ϕ.

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

Example 0.2.7. The formula

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

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

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 get

$$\left(\phi_1 \rightarrow \phi_2 \right) \vee \left(\phi_2 \rightarrow \phi_1 \right), \tag{0.2.8}$$

which is equal to

$$\Big((v_1 + 1 < v_2) \rightarrow (v_2 \cdot v_1 = v_1) \Big) \vee \Big((v_2 \cdot v_1 = v_1) \rightarrow (v_1 + 1 < v_2) \Big).$$

In any given structure both ϕ_1 and ϕ_2 are either true or false, depending on v_1 and v_2. However since τ is a tautology in sentential logic, it is true for all possible assignments of truth values. As such (0.2.8) is true in every $\mathcal{L}_{or}$-structure, regardless if ϕ_1 and ϕ_2 are true or false, and is therefore 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 0.2.9. [1, pp. 1.3.5–1.3.9.] We have three groups of *logical axioms* and two *rules of inference* given below. We start with the logical axioms.

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 free in ψ, but not in ϕ. Then

   $$\forall v \, (\phi \rightarrow \psi(v)) \rightarrow (\phi \rightarrow \forall v \, \psi(v))$$

   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 ϕ 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

   $$\forall v \, \phi(v) \rightarrow \psi$$

   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) $x = y \rightarrow$
   $$\Big(t(v_1, \ldots, v_{i-1}, x, v_{i+1}, \ldots, v_n) = t(v_1, \ldots, v_{i-1}, y, v_{i+1}, \ldots, v_n) \Big),$$

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

are logical axioms.

The *Rules of Inference* are then

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

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

5. **Rule of Generalization**

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

Definition 0.2.10 (Proof). [1, 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 0.2.10 introduces concepts from proof theory, we will in general not be quite as thorough when proving sentences in model theory. However, this type of formality is needed for the next definition and following theorem.

Definition 0.2.11 (Consistency). [1, 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 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.

The definition of inconsistency can be intuitively explained by the fact that if all sentences can be proven, clearly contradictory sentences are also included. So if for example T is an inconsistent set of sentences in a language $\mathcal{L}$, then both $T \vdash (t_1 = t_2)$ and $T \vdash \neg (t_1 = t_2)$ for terms t_1, t_2 in $\mathcal{L}$. The following proposition formalizes this some more, we present it without proof.

Proposition 0.2.12. *[1, Proposition 1.3.10] Let Σ be a set of sentences in a language $\mathcal{L}$. Then*

1. *the set Σ is consistent if and only if every finite subset of Σ is consistent;*

2. *the set $\Sigma \cup \{\sigma\}$ is inconsistent for a sentence σ if and only if $\Sigma \vdash \neg\sigma$;*

3. *if Σ is maximally consistent and σ, τ are arbitrary sentences in $\mathcal{L}$, we have that*

   (i) *$\Sigma \vdash \sigma$ if and only if $\sigma \in \Sigma$,*

   (ii) *$\sigma \notin \Sigma$ if and only if $\neg\sigma \in \Sigma$,*

   (iii) *$\sigma \wedge \tau \in \Sigma$ if and only if $\sigma \in \Sigma$ and $\tau \in \Sigma$;*

4. *a formula τ is deductible from $\Sigma \cup \{\sigma\}$ for a sentence σ if and only if $\Sigma \vdash \sigma \rightarrow \tau$.*

   The last point is also known as the "Deduction Theorem".

With this proposition in mind we continue on to the final result relating only to formal languages. After this we move on to actually using these languages to construct structures and models.

Theorem 0.2.13 (Lindenbaum's Theorem). *[1, Lemma 1.2.9], see also [1, 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. By Proposition 0.2.5 we have that any given language has $\|\mathcal{L}\|$ formulas. We arrange the sentences of $\mathcal{L}$ in some list

$$\phi_1, \phi_2, \ldots, \phi_\alpha, \ldots$$

where the specific order is irrelevant, but in a way where each ϕ_α is associated with an ordinal α. This is possible since we are working within **ZFC** and as such every set has a well-order.

Based on our existing consistent set of sentences Σ we now construct an increasing chain

$$\Sigma = \Sigma_0 \subset \Sigma_1 \subset \Sigma_2 \subset \cdots \subset \Sigma_\alpha \subset \cdots$$

using this sequence of sentences.

We define Σ_1 as $\Sigma_0 \cup \phi_0$ if the set $\Sigma_0 \cup \phi_0$ is consistent and choose $\Sigma_1 = \Sigma_0$ otherwise. In general, at any given step α, we let $\Sigma_{\alpha+1} = \Sigma_\alpha \cup \{\phi_\alpha\}$ if and only if $\Sigma_\alpha \cup \{\phi_\alpha\}$ is consistent. If $\Sigma_\alpha \cup \{\phi_\alpha\}$ is inconsistent we let $\Sigma_{\alpha+1} = \Sigma_\alpha$. Finally, we let Γ be the union of all sets Σ_α.

The claim is now that Γ is consistent and indeed maximally consistent.

We first assume that Γ is inconsistent in order to contrive a contradiction. Then, by Proposition 0.2.12, there must exist a deduction $\Gamma \vdash (\sigma \wedge \neg\sigma)$ for some sentence σ in $\mathcal{L}$. Assume that θ_0, $\theta_1, \ldots, \theta_q$ are the sentences in Γ used for this deduction. There then exists some ordinal α such that θ_0, $\theta_1, \ldots, \theta_q$ all belong to Σ_α. The set of sentences Σ_α is consistent by construction however and we arrive at a contradiction. Hence Γ must be consistent.

Assume now that Γ is not maximally consistent. Then Γ is the proper subset of some other consistent set of sentences Δ. Let ϕ_β be some sentence in Δ. Since both Δ

and Γ are consistent and $\Gamma \subset \Delta$, we must have that $\Sigma_\beta \cup \{\phi_\beta\}$ also is consistent. But if that is the case, we would have that $\Sigma_{\beta+1} = \Sigma_\beta \cup \{\phi_\beta\}$ and therefore $\phi_\beta \in \Gamma$. As such Γ must be maximally consistent. $\qquad\square$

0.3 Structures, Theories and Models

Definition 0.3.1. A $\mathcal{L}$-*structure* $\mathcal{M}$ is given by the following:

1. a non-empty 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 $|M|$ [4, Definition 1.1.2].

We sometimes also identify a structure $\mathcal{M}$ by writing $\left(M, f^{\mathcal{M}}, R^{\mathcal{M}}, c^{\mathcal{M}}\right)$ [4] or $(M, \mathcal{I})$, where $\mathcal{I}$ is the function mapping symbols of $\mathcal{L}$ to their respective interpretations in M [1, p.20].

Definition 0.3.2. [4, 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 $a_1,\ldots, a_{n_f}$ in M,

(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 $a_1,\ldots, a_{n_R}$ in M,

(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 $\mathcal{N}$. A surjective $\mathcal{L}$-embedding, meaning the function μ is a bijection, is called an $\mathcal{L}$-*isomorphism*.

Example 0.3.3. [4, 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 an $\mathcal{L}_G$-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 $\overline{a} = a_1, \ldots, a_n$. 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 0.2.4 of formulas in defining what truth means for a structure, we have to be similarly careful when defining term valuation.

Definition 0.3.4. [1, Definition 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 an $\mathcal{L}$-structure, we define the *value of the term* $t(v_1, \ldots, v_n)$ *at* $\overline{a} = \langle a_1, \ldots, a_n \rangle \in M^n$ the following way:

1. if t is v_i, then $t(\overline{a}) = a_i$,

2. if t is the constant symbol c in $\mathcal{L}$, then $t(\overline{a})$ is given by the interpretation $c^{\mathcal{M}}$,

3. if t is the function $f(t_1, \ldots, t_n)$ in $\mathcal{L}$, then $t(\overline{a})$ is given by the interpretation $f^{\mathcal{M}}(t_1(\overline{a}), \ldots, t_n(\overline{a}))$.

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

Definition 0.3.5. [4, Definition 1.1.6] Let $\mathcal{M}$ be an $\mathcal{L}$-structure and let ϕ be a formula in $\mathcal{L}$ with free variables $\overline{v} = \langle v_1, \ldots, v_n \rangle$. We inductively define $\mathcal{M} \models \phi(\overline{a})$ for $\overline{a} = \langle a_1, \ldots, a_n \rangle \in M^n$.

1. If ϕ is $t_1 = t_2$, then $\mathcal{M} \models \phi(\overline{a})$ if and only if $t_1(\overline{a}) = t_2(\overline{a})$,

2. If ϕ is $R\left(t_1, \ldots, t_{n_R}\right)$, then $\mathcal{M} \models \phi(\overline{a})$ if and only if $\left\langle t_1(\overline{a}), \ldots, t_{n_R}(\overline{a}) \right\rangle \in R^{\mathcal{M}}$,

3. If ϕ is $\neg \psi$, then $\mathcal{M} \models \phi(\overline{a})$ if and only if $\mathcal{M} \not\models \psi(\overline{a})$,

4. If ϕ is $\psi \wedge \rho$, then $\mathcal{M} \models \phi(\overline{a})$ if and only if $\mathcal{M} \models \psi(\overline{a})$ and $\mathcal{M} \models \rho(\overline{a})$,

5. If ϕ is $\exists v_i \, \psi(\overline{v}, v_i)$, then $\mathcal{M} \models \phi(\overline{a})$ if and only if there exists some element $b \in M$ such that $\mathcal{M} \models \psi(\overline{a}, b)$

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

The following propositions make sure that Definition 0.3.5 is well defined. As such we ensure that model satisfaction of a formula is only dependent on the variables which are actually free in said formula. We also have that we can restrict $\mathcal{L}$-structures

to a substructure and still have certain types of formulas be true. Both of the following propositions follow from induction on Definitions 0.2.2 and 0.2.4 of terms and formulas.

Proposition 0.3.6. *[1, Proposition 1.3.16] Let $\mathcal{M} = \left(M, f^{\mathcal{M}}, R^{\mathcal{M}}, c^{\mathcal{M}} \right)$ be an $\mathcal{L}$-structure. Then the following statements hold:*

1. *Let $t(v_0, \ldots, v_p)$ be a term dependent on the variables $v_0, \ldots, v_p$ and let $\overline{x} = \left\langle x_0, \ldots, x_q \right\rangle$ and $\overline{y} = \left\langle y_0, \ldots, y_r \right\rangle$ be sequences in M, such that $p \leq q, p \leq r$ and $x_i = y_i$ whenever v_i is a variable of t. Then*

$$t(\overline{x}) = t(\overline{y}).$$

2. *Let ϕ be a formula in $\mathcal{L}$ with free and bound variables among $v_0, \ldots, v_p$ and let $\overline{x} = \left\langle x_0, \ldots, x_q \right\rangle$ and $\overline{y} = \left\langle y_0, \ldots, y_r \right\rangle$ be sequences in M, such that $p \leq q, p \leq r$ and $x_i = y_i$ whenever v_i is a variable of t. Then*

$$\mathcal{M} \models \phi(\overline{x}) \quad \textit{if and only if} \quad \mathcal{M} \models \phi(\overline{y}).$$

Proposition 0.3.7. *[4, Proposition 1.1.8] Suppose that $\mathcal{M}$ and $\mathcal{N}$ are two $\mathcal{L}$-structures with underlying sets M and N respectively and that $\mathcal{M}$ is a substructure of $\mathcal{N}$. Then, for $\overline{a} \in M^n$ and a quantifier-free formula $\phi(\overline{v})$ of $\mathcal{L}$ with n free variables, $\mathcal{M} \models \phi(\overline{a})$ if and only if $\mathcal{N} \models \phi(\overline{a})$.*

We can also define a type of equivalence between $\mathcal{L}$-structures.

Definition 0.3.8. [4, Definition 1.1.9] Two $\mathcal{L}$-structures $\mathcal{M}$ and $\mathcal{N}$ are called *elementary equivalent* if

$$\mathcal{M} \models \phi \quad \text{if and only if} \quad \mathcal{N} \models \phi$$

for all sentences ϕ of $\mathcal{L}$.

This is important, since like with algebraic structures we are more interested in the properties an $\mathcal{L}$-structure has rather how we name its elements. Definition 0.3.8 is very powerful when comparing models, this is used in Theorem 0.1.5 for example. The notion of an $\mathcal{L}$-isomorphism is more powerful, as the following proposition shows. Two isomorphic $\mathcal{L}$-structures are always elementary equivalent, however the converse only holds when the structures are finite [1, Proposition 1.3.19].

Proposition 0.3.9. *[4, Theorem 1.1.10] Let $\mathcal{M}$ and $\mathcal{N}$ be two $\mathcal{L}$-structures and suppose that $h : \mathcal{M} \rightarrow \mathcal{N}$ is an isomorphism. Then $\mathcal{M}$ and $\mathcal{N}$ are also elementary equivalent.*

The proof of Proposition 0.3.9 follows from Definition 0.2.4 of formulas, we omit it here for reasons of brevity. The full proof can be found in [4]. An important takeaway from it however, is that it shows that terms are "well-behaved" with respect to model satisfaction.

Suppose that $\phi(\bar{v})$ is a formula with free variables $\bar{v} = \langle v_0, \ldots, v_p \rangle$ in the language $\mathcal{L}$ and let $t_1(\bar{v}), \ldots, t_p(\bar{v})$ be terms, such that none of the variables of the t_is occur bound in ϕ. If $\mathcal{M}$ is an $\mathcal{L}$-structure, terms being well-behaved then means that $\mathcal{M} \models \left(\phi(t_0, \ldots, t_p) \right)(\bar{x})$ if and only if $\mathcal{M} \models \phi\left(t_0(\bar{x}), \ldots, t_p(\bar{x}) \right)$, Here $\bar{x} = \langle x_0, \ldots, x_p \rangle$ is a sequence in M and $\phi(t_0, \ldots, t_p)$ is the formula in $\mathcal{L}$ where each variable v_i is substituted with the corresponding term t_i for $i = 0, \ldots, p$.

0.3.1 Theories

Definition 0.3.10 (Theory). [4, §.1.2][1, §1.4] We call a set T of sentences in a language $\mathcal{L}$ a *theory* in $\mathcal{L}$. Further, we say that an $\mathcal{L}$-structure $\mathcal{M}$ is a *model* of T, if $\mathcal{M} \models \phi$ for all sentences $\phi \in T$.

Since T is a set of sentences, the notions of a *subtheory* and an *extension* of a theory are defined as usual.

For a theory T and a sentence ϕ in a language $\mathcal{L}$, we say that ϕ is a *logical consequence* of T if $\mathcal{M} \models \phi$ whenever $\mathcal{M} \models T$ [4, Definition 1.2.12].

We now provide some examples of models of different theories. One such example is of course the theory of groupoids, which the proof in our final chapter centres around.

Example 0.3.11. [4, Example 1.2.2] Let $\mathcal{L} = \{<\}$ be the language consisting only of the binary relation symbol $<$. We can then form the theory T_{lin} of strict linear orders by requiring it to contain the following sentences:

(i) $\forall x \quad \neg (x < x),$

(ii) $\forall x, y, z \quad \left(x < y \wedge y < z \right) \rightarrow x < z,$

(iii) $\forall x, y \quad \left(x < y \right) \vee \left(x = y \right) \vee \left(y < x \right).$

Removing sentence (iii) for example would give us the theory of non-linear strict orderings. Alternatively, we can expand T to be the theory of dense linear orders by adding the sentences

$$\forall x, y \left(x < y \rightarrow \exists z \, x < z < y \right)$$

and

$$\forall x \, \exists y, z \, \left(y < x < z \right).$$

Example 0.3.12. [4, Example 1.2.5] Let $\mathcal{L}_g = \{\cdot, \ e\}$ be the language of groups, where $\cdot$ is a binary function and e is a constant. The theory T_g of groups is then given by the following sentences:

(i) $\forall x, \ y, \ z \quad (x \cdot y) \cdot z = x \cdot (y \cdot z)$,

(ii) $\forall x \quad e \cdot x = x \cdot e = x$,

(iii) $\forall x \exists x^{-1} \quad x \cdot x^{-1} = x^{-1} \cdot x = e$.

Hence models of T_g is are models of group theory. If we add the sentence

$$\forall x, \ y \quad x \cdot y = y \cdot x,$$

to the theory above, we get the theory of abelian groups.

Say we remove condition (iii), then the theory is that of *monoids*.

Say we let T be the theory of groups of order 2. This means that T contains the sentences (i) - (iii) above, along with the sentence $\forall x \quad x \cdot x = e$. We claim that

$$T \not\models \exists x \, \exists y \, \exists z \, \left(x \neq y \wedge y \neq z \wedge x \neq z\right). \tag{0.3.13}$$

We then have that

$$\mathbb{Z}/2\mathbb{Z} \models T \wedge \exists x \, \exists y \, \exists z \, \left(x \neq y \wedge y \neq z \wedge x \neq z\right).$$

Since this is the only group of order 2, the above statement (0.3.13) holds [4, Proposition 1.2.13].

0.4 The Löwenheim-Skolem Theorem

We now explore the completeness and compactness regarding the models of a given theory. This bridges the gap between syntax and semantics and as such the Extended Completeness Theorem tells us that a theory is consistent if and only if it has a model. The Compactness Theorem states that models of a set of sentences is closed under taking subsets: if a set of sentences has a model, then every finite subset of that set also has a model.

What we focus on here of course is to prove the namesake of this chapter, the Löwenheim-Skolem Theorem. The proof of this is very straightforward and mostly a corollary of the two aforementioned theorems.

Definition 0.4.1. [1, p.61] Let T be a set of sentences in a language $\mathcal{L}$ and let $C \subseteq \mathcal{C}$ be a set of constant symbols of $\mathcal{L}$. We then call C a *set of witnesses* for T if and only if there exists some constant $c \in C$, such that

$$T \vdash \left(\exists x \, \phi(x) \to \phi(c)\right)$$

holds for all formulas ϕ of $\mathcal{L}$ which have at most a single free variable x.

It should be noted that this definition relates entirely to the syntax of $\mathcal{L}$, it does not have anything to do with $\mathcal{L}$-structures. Hence the expression $\phi(c)$ also means that we have replaced all free occurrences of x in $\phi(x)$ by the constant c. It is *not* the evaluation of a formula as seen in Definition 0.3.5.

Definition 0.4.1 is quite technical and also very catered to the needs of our proofs. We use it mainly in the following two lemmata, the proofs of which are quite lengthy and tedious as a result. This is the heavy lifting we have to do before applying Lemmata 0.4.2 and 0.4.3 in the proofs of Theorems 0.4.10 and 0.4.11. Those proofs are in turn comparatively straightforward; most of the complexity is contained in the lemmata.

Lemma 0.4.2. *[1, Lemma 2.1.1] Let T be a consistent set of sentences of a language $\mathcal{L}$ and let C be a set of new constant symbols not in $\mathcal{L}$ and of power $|C| = \|\mathcal{L}\|$. Then let $\overline{\mathcal{L}} = \mathcal{L} \cup C$ be the simple expansion of $\mathcal{L}$ formed by adding C.*

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

Proof. We will first define the set C with cardinality $\|\mathcal{L}\|$. Then, similarly to how we did in the proof of Theorem 0.2.13, we define a sequence of consistent terms T_ξ based on the elements of C and show that the desired set of sentences $\overline{T}$ is the union $\bigcup_{\xi < \alpha} T_\xi$.

Let α be the cardinal $\|\mathcal{L}\|$. Then, for each cardinal $\beta < \alpha$ we let c_β be a constant symbol, such that $c_\beta \neq c_\gamma$ holds for all cardinals γ where $\beta < \gamma < \alpha$. We can now define the set $C = \{c_\beta \mid \beta < \alpha\}$ and let $\overline{\mathcal{L}}$ be the simple extension $\overline{\mathcal{L}} = \mathcal{L} \cup C$.

Since $\|\mathcal{L}\|$ is at least countably infinite by the definition of the power of a language, we have that $\|\overline{\mathcal{L}}\| = \alpha$. Further, since α is also an ordinal, we can define the sequence $\{\phi_\xi \mid \xi < \alpha\}$ of all formulas ϕ in $\overline{\mathcal{L}}$ which have at most one free variable.

We now inductively define another sequence

$$T = T_0 \subset T_1 \subset \cdots \subset T_\xi \subset \cdots,$$

of sets of sentences together with a sequence $(d_\xi \mid \xi < \alpha)$ of constants in C, such that these satisfy the following properties:

(i) for each ξ, the set T_ξ is consistent in $\overline{\mathcal{L}}$;

(ii) if ξ is a successor ordinal $\xi = \eta + 1$, then $T_\xi = T_\eta \cup \{\exists x_\eta \, \phi_\eta(x_\eta) \to \phi_\eta(d_\eta)\}$, where x_η is the free variable of ϕ_η if it exists and an arbitrary variable v otherwise;

(iii) if ξ is a non-zero limit ordinal, then $T_\xi = \bigcup_{\eta < \xi} T_\eta$.

Assume now that we have defined T_η as above. We show that T_ζ is consistent for $\zeta = \eta + 1$.

Firstly we note that T_η contains a strictly lesser than α amount of sentences which are not valid formulas in the original language $\mathcal{L}$. These are all sentences which include symbols from C. Secondly, observe that each of these sentences only contain a finite number of constant symbols from C. Because of these two assertions, we can assume that there exists some constant d_η in the sequence $(d_\xi \mid \xi < \alpha)$, which has not previously occurred in any of the T_ξ, for $\xi < \zeta$.

Now let

$$T_\zeta = T_\eta \cup \left\{ \exists x_\eta\, \phi_\eta(x_\eta) \to \phi_\eta(d_\eta) \right\}$$

and we assume that T_ζ is inconsistent in order to derive a contradiction. By Proposition 0.2.12, this means that

$$T_\eta \vdash \neg\big(\exists x_\eta\, \phi_\eta(x_\eta) \to \phi_\eta(d_\eta)\big).$$

Then, using that $\phi \lor \psi$ and $\phi \to \psi$ are defined as $\neg\left(\neg\phi \land \neg\psi\right)$ and $\left(\neg\phi \lor \psi\right)$ respectively, we have that

$$
\begin{aligned}
& T_\eta \vdash \neg\big(\exists x_\eta\, \phi_\eta(x_\eta) \to \phi_\eta(d_\eta)\big) \\
\Longleftrightarrow\ & T_\eta \vdash \neg\big(\neg\left(\exists x_\eta\, \phi_\eta(x_\eta)\right) \lor \phi_\eta(d_\eta)\big) \\
\Longleftrightarrow\ & T_\eta \vdash \exists x_\eta\, \phi_\eta(x_\eta) \land \neg\phi_\eta(d_\eta).
\end{aligned}
$$

Since d_η does not occur in T_η, we can utilize the Rule of Generalization in the last statement and obtain

$$T_\eta \vdash \forall x_\eta \left(\exists x_\eta\, \phi_\eta(x_\eta) \land \neg\phi_\eta(x_\eta) \right)$$

which is a more general formulation of

$$T_\eta \vdash \left(\exists x_\eta\, \phi_\eta(x_\eta) \right) \land \neg\left(\exists x_\eta\, \phi_\eta(x_\eta) \right).$$

This is clearly a contradiction and hence T_ζ must be consistent.

We now assume that ζ is a limit ordinal and that T_η is properly defined for all $\eta < \zeta$. Then, since each T_η is consistent, so is $T_\zeta = \bigcup_{\eta < \zeta} T_\eta$. As such $(T_\xi \mid \xi < \alpha)$ is a well-defined sequence.

We now let $\overline{T}$ be the union $\bigcup_{\xi < \alpha} T_\xi$. Then, as we just showed in the definition of $(T_\xi \mid \xi < \alpha)$, $\overline{T}$ is made up of only consistent sets of sentences and is therefore also consistent in $\overline{\mathcal{L}}$, as an extension of T.

Finally we show that C is the set of witnesses of $\overline{T}$ in $\overline{\mathcal{L}}$. Suppose that ϕ is a formula with at most one free variable x in $\overline{\mathcal{L}}$. We can assume that ϕ is ϕ_ξ and that $x = x_\xi$. As such the sentence

$$\exists x_\xi\, \phi_\xi(x_\xi) \to \phi_\xi(d_\xi)$$

is a member of $T_{\xi+1}$ and thus also of $\overline{T}$. Since the formula ϕ was arbitrary, the set C is indeed a set of witnesses for $\overline{T}$. $\qquad\square$

Lemma 0.4.3. *[1, Lemma 2.1.2] Let T be a consistent set of sentences and let C be a set of witnesses for T in $\mathcal{L}$. Then there exists a model $\mathcal{M}$ satisfying the theory of T, such that every element of $\mathcal{M}$ is an interpretation of a constant $c \in C$.*

Proof. Suppose that T' is an extension of T, i.e. $T \subset T'$. Then T' also has C for a set of witnesses in $\mathcal{L}$ and if some $\mathcal{L}$-structure is a model of T' it is also a model of T. Because of this and since by Theorem 0.2.13 we can always extend a consistent set of sentences to a maximally consistent one, we can assume without loss of generality that the set T is maximally consistent.

Let $c, d \in C$ be two constants. We define a relation $c \sim d$ if and only if the formula $c = d$ is in T. This is indeed an equivalence relation: since T is maximally consistent we have that the relation $\sim$ satisfies

  (i) $c \sim c$,

  (ii) if $c \sim d$ and $d \sim e$, then $c \sim e$,

  (iii) if $c \sim d$, then $d \sim c$

for all $c, d, e \in C$. For elements $c \in C$ we denote the equivalence class of c by

$$\tilde{c} = \{d \in C \mid c \sim d\}.$$

We now get to constructing an $\mathcal{L}$-structure $\mathcal{M}$ with the underlying set A, the elements of which are the equivalence classes $\tilde{c}$ for $c \in C$. We construct this structure in steps and then show that it is actually a model of T.

  1. **Interpretations of relations $R \in \mathcal{R}$.**

    For each n_R-ary relation symbol R in $\mathcal{L}$, we define a relation $R' \subseteq C^{n_R}$ on the set C such that

$$\langle c_1, \ldots, c_{n_R} \rangle \in R' \iff R(c_1, \ldots, c_{n_R}) \in T. \qquad (0.4.4)$$

    We have that $\sim$ is a *congruence relation* for the relation R'. This means that if we have $\langle c_1, \ldots, c_{n_R} \rangle \in R'$ and $c_i \sim d_i$ for elements $c_i, d_i, \in C, i = 1, \ldots, n_{n_R}$, it follows that $\langle d_1, \ldots, d_{n_R} \rangle \in R'$.

    More formally the Identity Axioms ensure that

$$\vdash R(c_1, \ldots, c_{n_R}) \wedge (c_1 = d_1) \wedge \cdots \wedge (c_{n_R} = d_{n_R}) \rightarrow R(d_1, \ldots, d_{n_R}).$$

    We now define the interpretation $R^{\mathcal{M}}$ of R on A by

$$\langle \tilde{c}_1, \ldots, \tilde{c}_{n_R} \rangle \in R^{\mathcal{M}} \iff R(c_1, \ldots, c_{n_R}) \in T$$

    and have by (0.4.4) that $R^{\mathcal{M}}$ is not dependent on the choice of representative c for each equivalence class $\tilde{c}$.

2. **Interpretations of constants** $d \in C$

Let d be a constant symbol of $\mathcal{L}$. It then follows from first-order logic that

$$\vdash \exists v \, (d = v)$$

and since T is maximally consistent this also means that $\big(\exists v \, (d = v)\big) \in T$.

Further, since C is a set of witnesses in T, there exists a constant $c \in C$, such that $(d = c) \in T$. Hence we choose the interpretation of d in $\mathcal{M}$ to be the equivalence class $\tilde{c}$. In the case that $d \in C$, it of course follows that the interpretation of d is $\tilde{d}$.

The constant c might not be unique, however the equivalence class is. Using the Identity Axioms we have that

$$\vdash (d = c \wedge d = c') \rightarrow c = c'$$

such that the interpretation of d is not dependent on the representative of the equivalence class.

3. **Interpretations of functions** $f \in \mathcal{F}$

We proceed similarly to steps 1 and 2. We have that

$$\vdash \exists v \, f(c_1, \ldots, c_{n_f}) = v$$

holds by first-order logic for constants $c_1, \ldots, c_{n_f} \in C$ and a function $f \in \mathcal{F}$. Then, as T has witnesses in C, it follows that $\big(f(c_1, \ldots, c_{n_f}) = c\big) \in T$ for some constant $c \in C$. By the Identity Axioms

$$\vdash f(c_1, \ldots, c_{n_f}) = c \wedge (c_1 = d_1) \wedge \cdots \wedge (c_{n_R} = d_{n_R}) \wedge (c = d)$$
$$\rightarrow f(d_1, \ldots, d_{n_f}) = d \quad (0.4.5)$$

and as such the value of the function is unique with respect to the equivalence classes of C.

Finally, we define the interpretation of f in $\mathcal{M}$ by

$$f^{\mathcal{M}}(\tilde{c}_1, \ldots, \tilde{c}_{n_f}) = \tilde{c} \iff \big(f(c_1, \ldots, c_{n_f}) = c\big) \in T.$$

We have constructed an $\mathcal{L}$-structure $\mathcal{M}$, of which the underlying set A consists entirely out of the equivalence classes $\tilde{c}$ of C with respect to the relation $\sim$. As such every $\tilde{c} \in A$ is the interpretation of some constant symbol $c \in C$.

We now show that $\mathcal{M}$ is indeed a model of T. Let t be a term in $\mathcal{L}$ with no free variables, we show that

$$\mathcal{M} \models t = c \iff (t = c) \in T. \quad (0.4.6)$$

By Definition 0.2.2 a term without free variables is either of the form $t = d$ for a constant $d \in C$ or $t = f(t_1, \ldots, t_{f_n})$ for a function $f \in \mathcal{F}$ and terms $t_1, \ldots, t_{f_n}$. We have that (0.4.6) already holds for $t = d$ and $t = f(d_1, \ldots, d_{f_n})$ by steps 2 and 3 above, where $d, d_1, \ldots, d_{f_n} \in C$ are constants. Additionally (0.4.5) assures us that all functions are well defined in $\mathcal{M}$, meaning that we can extend (0.4.6) to hold for all terms t.

We now also have that

$$\mathcal{M} \models t_1 = t_2 \iff (t_1 = t_2) \in T, \qquad (0.4.7)$$

$$\mathcal{M} \models R^{\mathcal{M}}\left(t_1, \ldots, t_{n_R}\right) \iff R(t_1, \ldots, t_{n_R}) \in T \qquad (0.4.8)$$

for bound terms $t_1, t_2, \ldots, t_{n_R}$ following from step 1 above and the fact that T has witnesses in C.

With this we have covered model satisfaction for all atomic sentences in $\mathcal{L}$. We now proceed with proving

$$\mathcal{M} \models \phi \iff \phi \in T \qquad (0.4.9)$$

for sentences ϕ using induction on Definition 0.2.4. For sentences $\phi, \psi \in \mathcal{L}$ it follows from Definition 0.2.4 that

$$\mathcal{M} \models \neg \phi \iff (\neg \phi) \in T$$

$$\mathcal{M} \models \phi \wedge \psi \iff (\phi \wedge \psi) \in T.$$

Now assume that ϕ is the sentence $\exists v\, \psi(v)$ and that $\mathcal{M} \models \phi$. Then there exists some $\tilde{c} \in A$ such that $\mathcal{M} \models \psi(\tilde{c})$.

Due to how we defined the model $\mathcal{M}$ this means that there also exists some constant symbol $c \in C$ in $\mathcal{L}$, such that $\mathcal{M} \models \psi(c)$. This in turn, also by construction, means that $\psi(c) \in T$. Note that by $\psi(c)$ we mean the sentence in $\mathcal{L}$ where all occurrences of v have been replaced by the constant c and *not* the evaluation of ϕ with respect to c.

As such we have that

$$\vdash \psi(c) \rightarrow \exists v\, \psi(v)$$

and since ϕ *is* $\exists v\, \psi(v)$ also $\phi \in T$.

Now assume that $\phi \in T$, where again ϕ is $\exists v\, \psi(v)$, we want to prove that $\mathcal{M}$ satisfies ϕ. Because C is a set of witnesses to T this means that

$$T \vdash \exists v\, \psi(v) \rightarrow \psi(c).$$

The set T is maximally consistent and hence $\psi(c) \in T$ and $\mathcal{M} \models \psi(c)$. Therefore there exists an equivalence class $\tilde{c} \in A$ such that $\mathcal{M} \models \psi(\tilde{c})$.

Finally we have that

$$\mathcal{M} \models \phi$$

and hence $\mathcal{M}$ is a model of the theory T. $\qquad \square$

Theorem 0.4.10 (Extended Completeness Theorem). *[1, 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.*

Proof. Let $\mathcal{M}$ be a model of Σ and assume towards a contradiction that Σ is inconsistent. Then by Proposition 0.2.12, there exists a sentence $\phi \in \Sigma$, such that $\Sigma \vdash \neg\phi$. This also means that $\mathcal{M} \models \phi$ and $\mathcal{M} \models \neg\phi$. However this is a violation of Definition 0.3.5 and therefore Σ must be consistent.

We now assume that Σ is a consistent set of sentences and want to show that a model $\mathcal{M}$ of Σ exists. Consider the extensions $\overline{\Sigma}$ and $\overline{\mathcal{L}}$ of Σ and $\mathcal{L}$ respectively, such that $\overline{\mathcal{L}}$ has witnesses in $\overline{\Sigma}$, which is possible by Lemma 0.4.2.

Using Lemma 0.4.3 we construct a model $\mathcal{N}$ of $\overline{\Sigma}$. Now, $\mathcal{N}$ is an $\overline{\mathcal{L}}$-structure. Hence we let $\mathcal{M}$ be the restriction of $\mathcal{N}$ to $\mathcal{L}$. Because Σ does not contain any sentences of $\overline{\mathcal{L}}$, we have that $\mathcal{M}$ is an $\mathcal{L}$-structure and a model of Σ. $\square$

Theorem 0.4.11 (Compactness Theorem). *[1, Theorem 1.3.22] A set of sentences Σ has a model if and only if every finite subset of Σ has a model.*

Proof. By Proposition 0.2.12 the set Σ is consistent if and only if every finite subset of Σ is consistent. This combined with Theorem 0.4.10 means that if every finite subset of Σ has a model, it is also consistent. Hence Σ itself is consistent and has a model. By that same line of reasoning every finite subset of Σ has a model if Σ is consistent. $\square$

We now come to our final theorem and the namesake of this section. Assume we have a theory T in a language $\mathcal{L}$. The Löwenheim-Skolem Theorem then relates the power of models of a T to the power of $\mathcal{L}$ itself.

We treat the theorem in two separate parts. The Downward Löwenheim-Skolem Theorem assures us that there exist models of T with power *at most* $\|\mathcal{L}\|$. As such, if we have a model of T with cardinality greater than $\|\mathcal{L}\|$, we can *move down* and know that a model with lower power exists.

Similarly, the Upward Löwenheim-Skolem Theorem states that if T has an infinite model, it in fact has models of arbitrary infinite cardinality. This means that if we are able to model T using, say, the natural numbers as the underlying set, we also know that an uncountable model of T exists. This is the property we later make use of in the proof on Theorem **??**. As such the Upward Löwenheim-Skolem Theorem allows us to *move up* in cardinality for models of T.

Theorem 0.4.12 (Downward Löwenheim-Skolem Theorem). *[1, Corollary 2.1.4] Every consistent theory T in a language $\mathcal{L}$ has a model of power at most $\|\mathcal{L}\|$.*

Proof. We modify the proof of Theorem 0.4.10 using arguments from Lemma 0.4.2.

Let $\|\mathcal{L}\| = \alpha$. We again consider the simple extension of $\mathcal{L}$ by a set of constants, such that $\overline{\mathcal{L}} = \mathcal{L} \cup C$. Then by Lemma 0.4.2 T can be expanded to a consistent set of sentences $\overline{T}$ in $\overline{\mathcal{L}}$, such that $\overline{T}$ has C as a set of witnesses. As in the proof of Lemma

0.4.2 we can construct this set of constants C in a way where the cardinality of C is $|C| = \alpha$. This means that the power of $\overline{\mathcal{L}}$ also is α.

By Lemma 0.4.3 the set $\overline{T}$ has a model $\mathcal{N}$ in $\overline{\mathcal{L}}$. Hence we consider the model $\mathcal{M}$ which is the restriction of $\mathcal{N}$ to $\mathcal{L}$. We have already shown in the proof of Theorem 0.4.10 that $\mathcal{M}$ is indeed a model of T. Further, since $\lVert \mathcal{L} \rVert = \alpha$ we have that $\mathcal{M}$ has power of at most α. $\qquad \square$

Theorem 0.4.13 (Upward Löwenheim-Skolem Theorem). *[1, Corollary 2.1.6] If a theory T has infinite models, then it has infinite models of any given power $\alpha \geq \lVert \mathcal{L} \rVert$.*

Proof. Consider a set $C = \left\{ c_\xi \mid \xi < \alpha \right\}$ of constant symbols c_ξ which are not part of the language $\mathcal{L}$. We define a set of sentences Σ by

$$\Sigma = T \cup \left\{ \neg(c_\xi = c_\zeta) \mid \xi < \zeta < \alpha \right\}.$$

Further, note that every finite subset $\Sigma' \subset \Sigma$ also only contains a finite number of constants c_ξ. Hence every infinite model $\mathcal{M}$ of T can be extended to a model $\mathcal{M}'$ of Σ'.

It then follows from Theorem 0.4.11 that Σ itself also has a model $\mathcal{N}$. By the [Downward Löwenheim-Skolem Theorem](#) this model $\mathcal{N}$ has power of at most

$$\left\lVert \mathcal{L} \cup \left\{ c_\xi \mid \xi < \alpha \right\} \right\rVert = \alpha.$$

However, the underlying set N of $\mathcal{N}$ must have unique interpretations for each symbol in $\left\{ c_\xi \mid \xi < \alpha \right\}$. As such we have that the cardinality of N satisfies

$$\alpha \leq |N|$$

and indeed see that $\mathcal{N}$ exactly has power $|N|$. $\qquad \square$

Lastly it should be pointed out, that we do indeed use the Axiom of Choice in the last proof. The way we arrange the constants in a sequence is not always possible without **AC**. Additionally the comparability of sets is not given in **ZF**, which is the property that $|X| \leq |Y|$ or $|X| \geq |Y|$ holds for all set X and Y. This is something the theorem implicitly assumes in its formulation.