PMATH 433 Course Notes

Set Theory and Model Theory

Rahim Moosa • Fall 2020 • University of Waterloo

Table of Contents

Part 1

Set Theory

1 First axioms

We use the natural numbers $0, 1, 2, \ldots$ to "count" finite sets. We may

- enumerate, list, and order, or
- measure size.

Our aim is to develop ordinals and cardinals to do this work for arbitrary (possibly infinite) sets.

We start by building the set of natural numbers. First, start with the undefined notions of a set, and membership. We would like to define

$$0 := \varnothing$$
$$1 := \{0\}$$
$$2 := \{0, 1\}$$

and more generally, $n + 1 = S(n) := n \cup \{n\}$, where $S(n)$ is said to be the successor of n.

Without any assumptions, these may not exist. We require some axioms.

Axiom 1.1 (Emptyset). There exists a set which has no members, called the empty set, and is denoted by $\varnothing$.

To produce 1 from 0, we need to know that if x is a set, then so is $\{x\}$.

Axiom 1.2 (Pairset). Given sets x, y, there is a set, denoted by $\{x, y\}$, with the property that its only members are x and y. That is, $t \in \{x, y\}$ if and only if $t = x$ or $t = y$.

Note that if $x = y$, then $t \in \{x, y\}$ if and only if $t = x$. To conclude that $\{x, y\} = \{x\}$, we require the following axiom.

Axiom 1.3 (Extension). For any two sets x and y, $x = y$ if and only if x and y have the same members.

For the general case, from n to get $S(n) = n \cup \{n\}$, we need the next axiom.

Axiom 1.4 (Unionset). Given a set x, there exists a set denoted by $\bigcup x$, whose members are precisely the members of the members of x. That is, $t \in \bigcup x$ if and only if $t \in y$ for some $y \in x$.

Then we can write $S(n) = \bigcup\{n, \{n\}\}$, since $t \in S(n)$ if and only if $t \in n$ or $t = n$. Hence if n exists, then by pairset (twice) and unionset, $S(n)$ exists. With these four axioms, we can now prove that every natural number exists.

What about the set of *all* natural numbers? Why don't we simply add an axiom saying that there is a set whose elements are precisely the natural numbers?

For example, we could write: "There exists a set N such that $t \in N$ if and only if $t = 0$ or $t = 1$ or $t = 2$ or $\cdots$." However, the right-hand side is not a **definite condition** on t.

Definition 1.5 (Definite conditions and operations). If x, y are sets or indeterminates standing for sets, then $x \in y$ and $x = y$ are definite (binary) conditions. If P and Q are definite conditions, then

- "not P", denoted $\neg P$,
- "P and Q", denoted $P \wedge Q$,
- "P or Q", denoted $P \vee Q$,
- "for all x, P", denoted $\forall x, P$, and
- "there exists x, P", denoted $\exists x, P$,

are all definite conditions. Only conditions arising as above in finitely many steps are definite conditions. We say an operation $H(x)$ is definite if the condition $y = H(x)$ is definite.

We want our existence axioms to be of the form: "There exists a set N such that $t \in N$ if and only if $P(t)$," where P is a definite condition.

2 The set of natural numbers

Observe that we can rewrite our axioms as definite conditions as follows.

AXIOM 2.1 (Emptyset). There exists a set $\varnothing$ satisfying $\neg\exists y, y \in \varnothing$.

AXIOM 2.2 (Pairset). Given two sets x, y, there is a set $\{x, y\}$ satisfying

$$\forall t\,(t \in \{x, y\} \leftrightarrow ((t = x) \vee (t = y))).$$

Note that "if P then Q", denoted $P \to Q$, is definite, as it can be expressed as $\neg P \vee Q$.

AXIOM 2.3 (Unionset). If x is a set, then there exists a set $\bigcup x$ satisfying

$$\forall t\,(t \in \bigcup x \leftrightarrow \exists y\,((y \in x) \wedge (t \in y))).$$

Of course, it is easy to see that the extension axiom is definite, so we simply restate it here.

AXIOM 2.4 (Extension). For any two sets x and y, $x = y$ if and only if x and y have the same members.

Now, we introduce new axioms.

AXIOM 2.5 (Infinity). There exists a set I that contains 0 and is preserved by the successor function. That is, I satisfies

$$(0 \in I) \wedge (\forall x\,(x \in I \to S(x) \in I)).$$

Observe that we can express the condition $S(x) \in I$ as

$$\exists y\,((y \in I) \wedge \forall t\,((t \in y) \leftrightarrow ((t \in x) \vee (t = x)))),$$

where we have $y = S(x)$. Hence, the infinity axiom is also definite.

Note that the set I from the infinity axiom is not uniquely determined. Certainly, every natural number is in I, but there is no reason why I cannot contain other elements.

We will call a set I **inductive** if it contains 0 and is closed under the successor function S. We want to find the "smallest" inductive set. Intuitively, if I is a fixed inductive set, then

$$\bigcap\{J \subseteq I : J \text{ is inductive}\}$$

is the set of all natural numbers. However, in order for this set to exist, we require more axioms, and we need to be able to take intersections of sets.

DEFINITION 2.6. Let x, y be sets. We write $x \subseteq y$ to say that every member of x is a member of y. That is,

$$\forall t\,(t \in x \to t \in y).$$

AXIOM 2.7 (Powerset). Given a set x, there exists a set $\mathcal{P}(x)$ satisfying

$$\forall t\,(t \in \mathcal{P}(x) \leftrightarrow \underbrace{\forall y\,(y \in t \to y \in x)}_{t \subseteq x}).$$

AXIOM 2.8 (Bounded Separation). Suppose x is a set and P is a definite condition. Then there exists a set y satisfying

$$\forall t\,(t \in y \leftrightarrow ((t \in x) \wedge P(t))).$$

We denote the above set as $y = \{t \in x \mid P(t)\} \subseteq x$.

Note that P must be definite. Moreover, the set is bounded, in the sense that we must start with a set x as a domain.

If we were allowed unbounded separation, then we could consider the set $R = \{t : t \notin t\}$. Observe that if $R \in R$, then $R \notin R$, and if $R \notin R$, then $R \in R$. Hence, R does not exist as a set. This is called **Russell's paradox**.

Exercise 2.9. Show that, given a non-empty set x, there exists a set $\bigcap x$ satisfying

$$\forall t\,(t \in \bigcap x \leftrightarrow \forall y\,(y \in x \to t \in y)).$$

What happens if $x = \varnothing$?

We can now construct the set of natural numbers ω.

Definition 2.10 (Set of Natural Numbers). Fix an inductive set I. Then the set of natural numbers ω is given by

$$\omega := \bigcap\{J \in \mathcal{P}(I) \mid (0 \in J) \wedge \forall t\,(t \in J \to S(t) \in J)\}.$$

Exercise 2.11. Prove that ω does not depend on the inductive set I.

Finally, we list an axiom which we will use later.

Axiom 2.12 (Replacement). Suppose P is a definite binary condition such that for every set x, there is a unique set y such that $P(x, y)$. Given any set A, there exists a set B with the property

$$y \in B \iff \exists x\,((x \in A) \wedge P(x, y)).$$

We can say this as "the image of a set under a definite operation exists as a set".

These eight axioms, together with "regularity" which says that every set has a $\in$-minimum element, form the axiom system known as Zermelo-Fraenkel set theory, denoted by **ZF**. We won't assume regularity in this course.

3 Classes and definite operators

Definition 3.1 (Class). A **class** is a collection of sets satisfying some definite property. We can apply unbounded separation to get a class. If P is a definite condition, then $[[z : P(z)]]$ is a class.

Remark 3.2.

1. All sets are classes. If x is a set, then $x = [[t : t \in x]]$.
2. Some classes are sets. For example, $[[z : z \in \omega]] = \omega$.
3. Not all classes are sets. Consider $R = [[t : t \notin t]]$. We showed earlier that this Russell class is not a set.

Definition 3.3 (Proper Class). A **proper class** is a class that is not a set.

Example 3.4. The universal class $U = [[t : t = t]]$ is not a set. Indeed, if U were a set, then $R = \{t \in U : t \notin t\}$ would be a set by bounded separation, a contradiction.

We say that two classes are equal if and only if they have the same members.

Note that membership is a binary relation between a set and a class. Consider $x \in y$, and note that x must be a set, and y must be a class. It does not make sense to talk about a class being a member of another class.

Definition 3.5 (Ordered Pair). Given sets x, y, we define the ordered pair

$$(x, y) := \{\{x\}, \{x, y\}\}.$$

This exists by the pairset axiom, and is an element of $\mathcal{P}(\mathcal{P}(X \cup Y))$, where $x \in X$ and $y \in Y$.

Exercise 3.6. Show that with this definition, we have the property that $(x, y) = (x', y')$ if and only if $x = x'$ and $y = y'$.

Definition 3.7 (Cartesian Product). Suppose X, Y are classes. The **Cartesian product** is defined as

$$X \times Y := [[z : z = (x, y) \text{ for some } x \in X,\ y \in Y]].$$

Note that $X \times Y$ is also a class. Indeed, if $X = [[x : P(x)]]$ and $Y = [[y : Q(y)]]$, then we can write

$$X \times Y = [[z \mid \exists x\, \exists y\, (P(x) \wedge Q(y) \wedge z = (x, y))]],$$

which is a definite condition.

Also note that if X, Y are sets, then so is $X \times Y$, since $\mathcal{P}(\mathcal{P}(X \cup Y))$ is a set by the powerset axiom, and we have

$$X \times Y = \{z \in \mathcal{P}(\mathcal{P}(X \cup Y)) : z = (x, y) \text{ for some } x \in X,\ y \in Y\}.$$

By bounded separation, this shows that $X \times Y$ is indeed a set.

Given classes X, Y by definite operation $f : X \to Y$, we mean a subclass $\Gamma(f) \subseteq X \times Y$ such that for all $x \in X$, there is a unique $y \in Y$ such that $(x, y) \in \Gamma(f)$. We are identifying the operation f with its graph $\Gamma(f)$. In a sense, this is a "vertical line test". We often write $f(x) = y$ instead of $(x, y) \in \Gamma(f)$.

Example 3.8. Let $S : \text{Sets} \to \text{Sets}$ be the successor function, where Sets is the class of all sets, namely $[[z : z = z]]$. Then we can write

$$\Gamma(S) = [[z \mid z = (x, y) \wedge \forall t\, (t \in y \leftrightarrow (t \in x) \vee (t = x))]],$$

and we have that $\Gamma(S) \subseteq \text{Sets} \times \text{Sets}$ is a subclass that satisfies the "vertical line test". Thus, S is a definite operation.

Remark 3.9. Suppose X and Y are sets and $f : X \to Y$ is a definite operation. Then $\Gamma(f) \subseteq X \times Y$ is a subset. In fact, if B is a set and A is a class with $A \subseteq B$, then A is a set.

PROOF. We have $A = [[z \mid P(z)]]$ with P definite, and $A = \{z \in B \mid P(z)\}$. Apply bounded separation. $\square$

Finally, we restate the replacement axiom using the language of classes.

AXIOM 3.10 (Replacement). Let Sets be the class of all sets. Suppose $f : \text{Sets} \to \text{Sets}$ is a definite operation and $A \in \text{Sets}$. Then there exists a set B satisfying

$$\forall t\,(t \in B \leftrightarrow \exists a\,((a \in A) \wedge (\underbrace{t = f(a)}_{(a,t)\in\Gamma(f)}))).$$

4 Ordering the natural numbers

THEOREM 4.1 (Induction Principle). Suppose that $J \subseteq \omega$ is such that $0 \in J$ and if $x \in J$ then $S(x) \in J$. Then $J = \omega$.

PROOF. By assumption, $J \subseteq \omega$. But by definition of ω, we have $\omega \subseteq J$. □

LEMMA 4.2. Suppose $n \in \omega$.

(a) If $x \in n$, then $x \in \omega$. That is, every element of ω is a subset of ω.

(b) If $x \in n$, then $x \subseteq n$.

(c) We have $n \notin n$.

(d) Either $n = 0$ or $0 \in n$.

(e) If $x \in n$, then either $S(x) \in n$ or $S(x) = n$.

PROOF. We prove parts (a) and (b). The proofs of the other properties are similar.

(a) Let $J := \{n \in \omega : n \subseteq \omega\}$. We want to show that $J = \omega$. Clearly, $0 \in J$ since $\varnothing \subseteq A$ for all sets A. Now, suppose $n \in J$. We have $S(n) = n \cup \{n\} \in \omega$. Since $n \subseteq \omega$ and $n \in \omega$, it follows that $S(n) \subseteq \omega$. By the induction principle, $J = \omega$.

(b) Let $J := \{n \in \omega : \forall x\,(x \in n \to x \subseteq n)\}$. As before, we want $J = \omega$. We have $0 \in J$ since there are no elements in $\varnothing$, and hence $\forall x\,(x \in \varnothing \to x \subseteq \varnothing)$ holds vacuously. Suppose $n \in J$. Consider $S(n) = n \cup \{n\}$. Note that $n \in S(n)$ and $n \subseteq S(n)$. Let $x \in S(n)$. If $x = n$, then $x \subseteq S(n)$. Otherwise, we have $x \in n$, in which case $x \subseteq n$ since $n \in J$, and hence $x \subseteq S(n)$. Thus, $S(n) \in J$, and so $J = \omega$ by induction. □

DEFINITION 4.3 (Strict Partial Ordering). A **strict partial ordering** on a set E is a binary relation R on E satisfying:

(i) **antireflexivity:** $\neg\, xRx$ for any $x \in E$.

(ii) **antisymmetry:** If xRy and yRx, then $x = y$.

(iii) **transitivity:** If xRy and yRz, then xRz.

PROPOSITION 4.4. $\in$ is a strict partial ordering on ω.

PROOF. Observe that part (c) of Lemma 4.2 is precisely antireflexivity.

Suppose $n, m \in \omega$ are such that $n \in m$ and $m \in n$. By part (b) of Lemma 4.2, we have $n \subseteq m$ and $m \subseteq n$, and hence $n = m$.

Finally, suppose $\ell, m, n \in \omega$ are such that $\ell \in m$ and $m \in n$. Then $m \subseteq n$ again by part (b) of Lemma 4.2, and thus $\ell \in n$. □

DEFINITION 4.5 (Linear Ordering). A strict partial ordering R on E is said to be **linear** (or **total**) if whenever $x, y \in E$, we either have xRy or yRx or $x = y$.

PROPOSITION 4.6. $(\omega, \in)$ is a linear ordering.

PROOF. Fix $n \in \omega$. Consider

$$J := n \cup \{m \in \omega : n \in m\} \cup \{n\}.$$

We want to show that $J = \omega$. Note that $J \subseteq \omega$ since $n \subseteq \omega$ by part (a) of Lemma 4.2.

Observe that if $n = 0$, then we may simply use part (d) of Lemma 4.2, so we may assume $n \neq 0$. Then by (d) of Lemma 4.2 again, we must have $0 \in n$, so $0 \in J$.

Suppose $m \in J$. If $m \in n$, then by part (e) of Lemma 4.2, either $S(m) \in n$ or $S(m) = n$. In either case, we see that $S(m) \in J$. If $n \in m$, then $n \in S(m)$, so $S(m) \in J$. Finally, if $m = n$, then $n \in S(n) = S(m)$, and thus $S(m) \in J$.

Therefore, we may conclude that J is inductive, and so $J = \omega$, as required. □

5 Ordinals

Definition 5.1 (Well-Ordering). A strict linear ordering $(E, <)$ is a **well-ordering** if every non-empty subset of E has a least element.

Proposition 5.2. $(\omega, \in)$ is a well-ordering.

Proof. Suppose that $X \subseteq \omega$ has no $\in$-least element. We will show that $X = \varnothing$. Consider

$$J := \{n \in \omega : S(n) \cap X = \varnothing\} \subseteq \omega.$$

We claim that $J = \omega$. First, we show that $0 \in J$. Suppose not. Then $S(0) \in X \neq \varnothing$, and since 0 is the only element in $S(0)$, it must be that $0 \in X$. But 0 is $\in$-least in ω by Lemma 4.2 part (d), and hence 0 is $\in$-least in X, a contradiction.

Now, fix $n \in J$. Consider $S(S(n)) \cap X$. Since $n \in J$, we see that $S(n) \cap X = \varnothing$. If $S(n) \in X$, then $S(n)$ is $\in$-least in X, so we must have $S(n) \notin X$. So $S(S(n)) \cap X = \varnothing$, and thus $S(n) \in J$, which proves our claim.

Let $n \in \omega$. Then $n \in J$, so we have that $S(n) \cap X = \varnothing$. But $n \in S(n)$, which implies that $n \notin X$. Since this holds for arbitrary $n \in \omega$, it must be that $X = \varnothing$. □

Definition 5.3 (Ordinal). An **ordinal** is a set α such that

(i) if $x \in \alpha$, then $x \subseteq \alpha$, and

(ii) $(\alpha, \in)$ is a well-ordering.

We denote by Ord the class of all ordinals. Later, we will show that Ord is a proper class.

Exercise 5.4. Verify that Ord is indeed a class. That is, show that being an ordinal is a definite condition.

Example 5.5.

(a) By Lemma 4.2 (a) and Proposition 5.2, ω is an ordinal.

(b) Every natural number is an ordinal. These are said to be **finite ordinals**. Indeed, fix $n \in \omega$. By Lemma 4.2 (a), we have $n \subseteq \omega$, so $(n, \in|_n)$ is a well-ordering, so (ii) holds. Moreover, (i) holds by Lemma 4.2 (b).

Exercise 5.6. Prove that not every subset of ω is a natural number. This will show that the converse of (i) in the definition of an ordinal does not necessarily hold.

When working with elements of an ordinal, we will sometimes write $x < y$ instead of $x \in y$.

Lemma 5.7. Suppose $\alpha, \beta \in \text{Ord}$, where $\alpha \subseteq \beta$ with $\alpha \neq \beta$. Then $\alpha \in \beta$.

Proof. Assume that $\alpha \subseteq \beta$ and $\alpha \neq \beta$, where $\alpha, \beta \in \text{Ord}$. Let $D = \beta \setminus \alpha = \{x \in \beta : x \notin \alpha\}$. Note that $D \neq \varnothing$ as $\alpha \neq \beta$. Let $d \in D$ be least. We will show that $\alpha = d$.

Note that by construction, we have $\alpha \subseteq \beta$, and moreover, $d \subseteq \beta$, since $d \in \beta$ and β is an ordinal.

Suppose that we did not have $d \subseteq \alpha$. Then there is an element $x \in d \setminus \alpha$. Then $x < d$ and $x \in d \setminus \alpha \subseteq \beta \setminus \alpha = D$, contradicting the fact that d is least in D. So $d \subseteq \alpha$.

On the other hand, let $x \in \alpha$. Since $x, d \in \beta$ and β is an ordinal, we either have $x < d$ or $d < x$ or $x = d$. First, notice that we cannot have $x = d$, since $x \in \alpha$ but $d \in \beta \setminus \alpha$. Also observe that if $d < x$, then together with $x \in \alpha$, noting that α is an ordinal, it follows from property (i) of the definition of an ordinal that $d \in \alpha$, a contradiction. Therefore, we must have $x < d$, so $\alpha \subseteq d$.

Thus, we have shown that $\alpha = d$, so $\alpha \in D$, and thus $\alpha \in \beta$. □

6 Ordering the ordinals

Proposition 6.1.

(a) Every member of an ordinal is an ordinal.

(b) No ordinal is a member of itself.

(c) If $\alpha \in \text{Ord}$, then $S(\alpha) \in \text{Ord}$.

(d) If $\alpha, \beta \in \text{Ord}$, then $\alpha \cap \beta \in \text{Ord}$.

Proof. We prove (a) and (b), and leave (c) and (d) as exercises.

(a) Suppose $\alpha \in \text{Ord}$ and $z \in \alpha$. Let $x \in z$ and suppose $y \in x$. Since $z \in \alpha$, we have $z \subseteq \alpha$, which implies that $x \in \alpha$. Then, since $x \in \alpha$, we get $x \subseteq \alpha$, and so $y \in \alpha$. So $x, y, z \in \alpha$, and $(\alpha, \in)$ is a strict well-ordering. Therefore, by transitivity, $y \in z$, and hence $x \subseteq z$. Since $z \subseteq \alpha$, it follows that $(z, \in)$ is a strict well-ordering as $(\alpha, \in)$ is. Thus, $z \in \text{Ord}$.

(b) Suppose $\alpha \in \text{Ord}$. Then $(\alpha, \in)$ is a strict well-ordering. If $\alpha \in \alpha$, then by antireflexivity of $(\alpha, \in)$, we get $\alpha \notin \alpha$, a contradiction. □

Proposition 6.2.

(a) If $\alpha, \beta \in \text{Ord}$, then either $\alpha = \beta$ or $\alpha \in \beta$ or $\beta \in \alpha$.

(b) If $E \subseteq \text{Ord}$ is a set of ordinals, then $(E, \in)$ is a strict well-ordering.

(c) If $E \subseteq \text{Ord}$ is a set of ordinals, then its supremum $\sup E := \bigcup E$ is an ordinal.

(d) Ord is a proper class.

Proof.

(a) Let $\alpha, \beta \in \text{Ord}$. Note that $\alpha \cap \beta \in \text{Ord}$ by Proposition 6.1 part (d), and $\alpha \cap \beta \subseteq \alpha$. Suppose $\alpha \cap \beta \neq \alpha$. Then by Lemma 5.7, we get $\alpha \cap \beta \in \alpha$. Similarly, if $\alpha \cap \beta \neq \beta$, then $\alpha \cap \beta \in \beta$. Together, these imply that $\alpha \cap \beta \in \alpha \cap \beta$, but this is a contradiction with Proposition 6.1 part (b), since $\alpha \cap \beta$ is an ordinal.

Thus, we must have $\alpha \cap \beta = \alpha$ or $\alpha \cap \beta = \beta$. If $\alpha \cap \beta = \alpha$, then $\alpha \subseteq \beta$, so either $\alpha = \beta$ or $\alpha \subsetneq \beta$, in which case Lemma 5.7 implies that $\alpha \in \beta$. On the other hand, if $\alpha \cap \beta = \beta$, then $\beta \subseteq \alpha$, which tells us either $\beta = \alpha$ or $\beta \subsetneq \alpha$. In the latter case, Lemma 5.7 gives $\beta \in \alpha$.

(b) Suppose $E \subseteq \text{Ord}$ is a set of ordinals. Consider $(E, \in)$.

- Antireflexivity: Proposition 6.1 part (b).
- Antisymmetry: Suppose $\alpha, \beta \in E$ are such that $\alpha \in \beta$ and $\beta \in \alpha$. Then $\alpha \subseteq \beta$ and $\beta \subseteq \alpha$, so $\alpha = \beta$.
- Transitivity: Suppose $\alpha, \beta, \gamma \in E$, where $\alpha \in \beta$ and $\beta \in \gamma$. Then $\beta \subseteq \gamma$, so $\alpha \in \gamma$.
- Linearity: Proposition 6.2 part (a).
- Well-ordered: Suppose $A \subseteq E$ where $A \neq \varnothing$. Let $\alpha \in A$. If $\alpha \cap A = \varnothing$, then α is the least element in A. Now, suppose $\alpha \cap A \neq \varnothing$. Let $A' := \alpha \cap A \subseteq \alpha$, which contains a least element, say $a \in A'$ (as α is an ordinal). Suppose $b \in A$ with $b \in a$. Since $a \in \alpha$, we have $b \in \alpha$, and hence $b \in \alpha \cap A = A'$, contradicting the fact that a is least in A'. Thus, a is least in A.

This shows that E is strictly well-ordered by membership.

(c) Let $E \subseteq \text{Ord}$ be a set of ordinals. Note that $\sup E$ is an ordinal by Proposition 6.1 part (a). Then by Proposition 6.2 part (b), $(\sup E, \in)$ is a strict well-ordering. Now, suppose $\alpha \in \sup E$. Then $\alpha \in \gamma$ for some $\gamma \in E$. Since γ is an ordinal, we have $\alpha \subseteq \gamma$. If $x \in \alpha$, then $x \in \gamma \subseteq \bigcup E = \sup E$. Hence $\alpha \subseteq \sup E$, and so $\sup E \in \text{Ord}$.

(d) Suppose Ord is a set. By Proposition 6.2 part (b), $(\text{Ord}, \in)$ is a strict well-ordering. Moreover, by Proposition 6.1 part (a), if $\alpha \in \text{Ord}$, then $\alpha \subseteq \text{Ord}$. Hence $\text{Ord} \in \text{Ord}$. But ordinals cannot be members of themselves, a contradiction. Hence, Ord is a proper class. □

As justified by Proposition 6.2, for $\alpha, \beta \in \text{Ord}$, the notation $\alpha < \beta$ will be synonymous with $\alpha \in \beta$.

7 Sups, successors, limits, and transfinite induction

LEMMA 7.1.

(a) If $\alpha \in \text{Ord}$, then $\alpha < S(\alpha)$ and there is no ordinal in between.

(b) Suppose E is a set of ordinals. Then $\sup E = \bigcup E$ is the least upper bound of E.

(c) Given a set of ordinals E, there exists a least ordinal not in E.

PROOF.

(a) Note that $S(\alpha) = \alpha \cup \{\alpha\}$, so we see immediately that $\alpha < S(\alpha)$, and if $x < S(\alpha)$, then $x < \alpha$ or $x = \alpha$.

(b) If $\sup E$ is not an upper bound for E, then there exists $\alpha \in E$ such that $\sup E < \alpha$. But $\alpha \subseteq \bigcup E = \sup E$, and hence $\sup E < \sup E$. Since $\sup E$ is an ordinal, this is a contradiction. To see that it is least, observe that if $\alpha < \sup E$, then $\alpha < \beta$ for some $\beta \in E$.

(c) Note that since $\text{Ord} \setminus E$ is a proper class, we cannot use well-ordering to take the least element of $\text{Ord} \setminus E$. However, if we can find $\alpha \in \text{Ord}$ such that $E \subsetneq \alpha$, then we can consider the set $\alpha \setminus E$. The least element of $\alpha \setminus E$ will then be the least element not in E.

If $\alpha = \sup E$, then note that $E \subseteq \alpha$ may not be strict.

If $\alpha = S(\sup E)$, then we may have $E = S(\sup E)$, in which case $\alpha \setminus E$ would be empty.

Finally, $\alpha = S(S(\sup E))$ works. We leave it as an exercise to verify this. □

DEFINITION 7.2 (Successor and limit ordinals). A **successor ordinal** is of the form $S(\alpha)$ for some $\alpha \in \text{Ord}$. An ordinal which is not a successor ordinal is said to be a **limit ordinal**.

EXAMPLE 7.3. Every non-zero $n \in \omega$ is a successor ordinal, while 0 and ω are limit ordinals.

Transfinite induction is a method to prove things about Ord.

THEOREM 7.4 (Transfinite induction). Suppose P is a definite condition satisfying:

$$\text{If } \alpha \text{ is an ordinal and } P(\beta) \text{ is true for all } \beta < \alpha, \text{ then } P(\alpha) \text{ is true.} \qquad (\star)$$

Then P is true of all ordinals.

PROOF. Suppose P is not true of all ordinals. Let $\alpha \in \text{Ord}$ be such that $\neg P(\alpha)$. Let

$$D = \{\underbrace{\beta \leq \alpha}_{\beta < S(\alpha)} : \neg P(\beta)\},$$

which is a set by bounded separation. Note that $\alpha \in D$, so D is non-empty. Let $\alpha_0 \in D$ be least. Hence for all $\beta < \alpha_0$, $P(\beta)$ is true. By $(\star)$, we see that $P(\alpha_0)$ is true, a contradiction. □

This is not the usual form of induction that we see. We often see it stated differently, in a more useful form. We leave it as an exercise to show that the following form is equivalent to the first.

COROLLARY 7.5 (Transfinite induction – second form). Suppose P is a definite condition satisfying:

(1) $P(0)$ is true.

(2) For all ordinals β, if $P(\beta)$ then $P(S(\beta))$.

(3) For all limit ordinals $\alpha > 0$, if $P(\beta)$ for all $\beta < \alpha$, then $P(\alpha)$.

Then P is true of all ordinals.

8 Transfinite recursion

In the previous section, we introduced transfinite induction, which is a method of proving things on Ord. On the other hand, transfinite recursion is a method to construct a definite operation on Ord with some desired property.

Let X be the class of all functions whose domain is an ordinal. That is,

$$X = [[\,\Gamma : \exists A, B,\ A \in \text{Ord},\ \Gamma \subseteq A \times B \text{ is the graph of a function}\,]].$$

We also introduce the following notation: if F is a definite operation on Ord and $\alpha \in \text{Ord}$, then $F \restriction_\alpha$ denotes the function obtained by restricting F to α.

THEOREM 8.1 (Transfinite recursion). Given a definite operation $G : X \to \text{Sets}$, there is a unique definite operation $F : \text{Ord} \to \text{Sets}$ satisfying

$$F(\alpha) = G(\underbrace{F \restriction_\alpha}_{\in X})$$

for all $\alpha \in \text{Ord}$.

(That is, if we know what F is on members of α, then G tells us what F should be on α.)

PROOF. We give a sketch of the proof.

It is straightforward to show uniqueness by transfinite induction. Let $F' : \text{Ord} \to \text{Sets}$ where $F'(\alpha) = G(F' \restriction_\alpha)$. Given $\alpha \in \text{Ord}$, suppose that $F(\beta) = F'(\beta)$ for all $\beta < \alpha$. Verify that this implies $F \restriction_\alpha = F' \restriction_\alpha$. Then, it follows that $F(\alpha) = G(F \restriction_\alpha) = G(F' \restriction_\alpha) = F'(\alpha)$. By transfinite induction, we have $F(\alpha) = F'(\alpha)$ for all ordinals α.

To show existence, we first introduce some terminology. A function t with domain α is an **α-function defined by G** if

$$t(\beta) = G(t \restriction_\beta)$$

for all $\beta < \alpha$. This is an approximation to the F we want to construct.

Observe that:

(1) If an α-function defined by G exists, then it is unique. (We can prove this by transfinite induction with the property $P(x) = x \geq \alpha \vee t(x) = t'(x)$.)

(2) If $\alpha < \beta$, t_α is an α-function defined by G, and t_β is a β-function defined by G, then $t_\alpha \subseteq t_\beta$ (that is, $\Gamma(t_\alpha) \subseteq \Gamma(t_\beta)$).

CLAIM. For all $\alpha \in \text{Ord}$, an α-function defined by G, t_α, exists.

PROOF OF CLAIM. We use the second form of transfinite induction. Clearly, $t_0 = \varnothing$. Given t_α for $\alpha \in \text{Ord}$, define

$$t_{S(\alpha)} = \begin{cases} t_\alpha & \text{on members of } \alpha \\ G(t_\alpha) & \alpha \text{ itself.} \end{cases}$$

Finally, if β is a limit ordinal and t_α is given for all $\alpha < \beta$, then define

$$t_\beta = \bigcup \{t_\alpha : \alpha < \beta\}.$$

Note that this is a set by replacement, as $\alpha \mapsto t_\alpha$ is a definite operation. Then t_β is a function by observation (2). Check that this is indeed a β-function defined by G. ■

Finally, take

$$F = \bigcup [[t_\alpha : \alpha \in \text{Ord}]].$$

Verify that $F(\alpha) = G(F \restriction_\alpha)$. □

We now state an easier form of transfinite recursion to work with.

COROLLARY 8.2 (Transfinite recursion – second form). Suppose that G_1 is a set, $G_2 : \text{Sets} \to \text{Sets}$ is a definite operation, and $G_3 : X \to \text{Sets}$ is a definite operation. Then there exists a unique definite operation $F : \text{Ord} \to \text{Sets}$ satisfying:

(1) $F(0) = G_1$.

(2) For all $\alpha \in \text{Ord}$, $F(S(\alpha)) = G_2(F(\alpha))$.

(3) For all limit ordinals $\alpha > 0$, $F(\alpha) = G_3(F \restriction_\alpha)$.

PROOF. Let $G : X \to \text{Sets}$ be given by

$$G(f) = \begin{cases} G_1 & \text{if } f = \varnothing \\ G_2(f(\alpha)) & \text{if } \text{dom}(f) = S(\alpha) \\ G_3(f) & \text{if } \text{dom}(f) \neq \varnothing, \text{ limit ordinal.} \end{cases}$$

Apply the first form of transfinite recursion to this to produce a unique $F : \text{Ord} \to \text{Sets}$ such that

$$F(\alpha) = G(F \restriction_\alpha)$$

for all $\alpha \in \text{Ord}$. It is straightforward to check that (1), (2), and (3) hold and that F is unique. □

9 Ordinal arithmetic

As an application of transfinite recursion, we can now define some arithmetic operations on ordinals.

Definition 9.1 (Ordinal addition). Fix $\beta \in \text{Ord}$ and define $\beta + \alpha$ for all $\alpha \in \text{Ord}$ as follows:

(1) $\beta + 0 := \beta$,

(2) $\beta + S(\alpha) := S(\beta + \alpha)$ for all $\alpha \in \text{Ord}$,

(3) $\beta + \alpha := \sup\{\beta + \gamma : \gamma < \alpha\}$ for all limit ordinals $\alpha > 0$.

In terms of transfinite recursion (second form), we have $G_1 = \beta \in \text{Sets}$, $G_2 = S : \text{Sets} \to \text{Sets}$, and $G_3 = \sup(\text{Im}\, f) : X \to \text{Sets}$ for f a function with domain an ordinal. This gives a unique definite operation $\text{Ord} \to \text{Sets}$ taking $\alpha \mapsto \beta + \alpha$, satisfying (1) to (3). We do this for each $\beta \in \text{Ord}$.

Notice that for $\beta \in \text{Ord}$, we have

$$\begin{aligned} \beta + 1 &= \beta + S(0) \\ &= S(\beta + 0) \text{ by (2)} \\ &= S(\beta) \text{ by (1)}. \end{aligned}$$

As such, we will tend to use $\alpha + 1$ to mean $S(\alpha)$. Also, note that ordinal addition is not commutative. We have

$$1 + \omega = \sup\{1 + n : n < \omega\} = \omega \neq S(\omega) = \omega + 1.$$

Remark 9.2. Ordinal addition concatenates the order type of the ordinals. More informally, for $\alpha, \beta \in \text{Ord}$ and taking $\alpha + \beta$, we are essentially placing β on top of α.

Definition 9.3 (Ordinal multiplication). Fix $\beta \in \text{Ord}$. Then define

(a) $\beta \cdot 0 := 0$,

(b) $\beta \cdot S(\alpha) := \beta \cdot \alpha + \beta$ for all $\alpha \in \text{Ord}$,

(c) $\beta \cdot \alpha := \sup\{\beta \cdot \gamma : \gamma < \alpha\}$ for all limit ordinals $\alpha > 0$.

In terms of transfinite recursion, we have $G_1 = 0$, $G_2(x) = x + \beta$, and $G_3 = \sup(\text{Im}\, f)$.

Note that we have

$$\begin{aligned} \omega \cdot 1 &= \omega \cdot S(0) \\ &= \omega \cdot 0 + \omega \text{ by (2)} \\ &= 0 + \omega \text{ by (1)} \\ &= \sup\{0 + n : n < \omega\} \\ &= \omega. \end{aligned}$$

In fact, we have $\beta \cdot 1 = \beta$ for all ordinals β. Also, notice that

$$\begin{aligned} \beta \cdot 2 &= \beta \cdot S(1) \\ &= \beta \cdot 1 + \beta \text{ by (2)} \\ &= \beta + \beta. \end{aligned}$$

We can extend this to obtain the ordinal

$$\omega \cdot \omega = \sup\{w \cdot n : n < \omega\}.$$

Finally, we observe that ordinal multiplication is also non-commutative. Indeed,

$$\omega \cdot 2 = \omega + \omega \neq \omega = \sup\{2 \cdot n : n < \omega\} = 2 \cdot \omega.$$

Exercise 9.4. What is the order type of $\beta \cdot \alpha$ in terms of the well-orderings β and α?

We end this section with some properties of ordinal addition and multiplication.

Proposition 9.5. Let $\alpha, \beta, \delta \in \text{Ord}$.

(a) $\alpha < \beta \Leftrightarrow \delta + \alpha < \delta + \beta$.

(b) $\alpha = \beta \Leftrightarrow \delta + \alpha = \delta + \beta$.

(c) $(\alpha + \beta) + \delta = \alpha + (\beta + \delta)$.

(d) If $\delta \neq 0$, then $\alpha < \beta \Leftrightarrow \delta\alpha < \delta\beta$.

(e) If $\delta \neq 0$, then $\alpha = \beta \Leftrightarrow \delta\alpha = \delta\beta$.

(f) $(\alpha\beta)\delta = \alpha(\beta\delta)$.

Proof. Exercise – use transfinite induction. □

10 Well-orderings and ordinals

From the work we have done, it might appear that ordinals are a very special class of well-orderings. However, we will prove that ordinals are in fact all the well-orderings. First, we begin with a couple lemmas which we will use to prove our main theorem.

LEMMA 10.1. Well-orderings are **rigid**; that is, the only automorphism is the identity.

PROOF. Suppose $(E, <)$ is a well-ordering, and let $f : (E, <) \to (E, <)$ be an automorphism (i.e. f is a bijective map and $f(a) < f(b)$ if and only if $a < b$). Let $D = \{x \in E : f(x) \neq x\}$. Assume towards a contradiction that $f \neq \mathrm{id}$. Then $D \neq \varnothing$. Let $a \in D$ be least. We consider two cases.

If $f(a) < a$, then $f(a) \notin D$, so $f(f(a)) = f(a)$. But f is injective, so $f(a) = a$, a contradiction.

If $a < f(a)$, then $f^{-1}(a) < a$, and hence $f^{-1}(a) \notin D$. From this, we see that $f^{-1}(a) = f(f^{-1}(a)) = a$. Applying f to both sides, we obtain $a = f(f^{-1}(a)) = f(a)$, which implies that $a \notin D$, a contradiction. □

LEMMA 10.2. A well-ordering is not isomorphic to any proper initial segment.

PROOF. Let $(E, <)$ be a well-ordering, and suppose that $b \in E$ so that $E_{<b} = \{x \in E : x < b\}$ is a proper initial segment. We want to show that $(E, <) \not\simeq (E_{<b}, <)$. Assume towards a contradiction that $f : (E, <) \to (E_{<b}, <)$ is an isomorphism. Let $D = \{x \in E : f(x) \neq x\}$. Note that $b \in D$ since $f(b) \in E_{<b} \not\ni b$, so $D \neq \varnothing$. So let $a \in D$ be least. We leave it as an exercise to show that both $f(a) > a$ and $f(a) < a$ lead to a contradiction. □

We are now ready to present the main result.

THEOREM 10.3. Every strict well-ordering is isomorphic to an ordinal. Moreover, the ordinal and the isomorphism are unique.

PROOF. Let $(E, <)$ be a well-ordering, and let $f : (E, <) \to (\alpha, \in)$ be an isomorphism, where $\alpha \in \mathrm{Ord}$. To see that f is unique, suppose that $g : (E, <) \to (\alpha, \in)$ is another isomorphism. Then

$$g^{-1} \circ f : (E, <) \to (E, <)$$

is an automorphism. By Lemma 10.1, we must have $g^{-1} \circ f = \mathrm{id}$, and hence $g = f$.

For the uniqueness of α, suppose that $g : (E, <) \to (\beta, \in)$ is an isomorphism, where $\beta \in \mathrm{Ord}$. If $\alpha \neq \beta$, then either $\alpha \in \beta$ or $\beta \in \alpha$. If $\alpha \in \beta$, then by Assignment 2 Question 4, we know that $(\alpha, \in) = (\beta_{\in\alpha}, \in)$ is an initial segment. Then

$$(\beta, \in) \xrightarrow[\simeq]{g^{-1}} (E, <) \xrightarrow[\simeq]{f} (\alpha, \in)$$

is an isomorphism, contradicting Lemma 10.2. The argument is analogous for the $\beta \in \alpha$ case, so we must have $\alpha = \beta$.

It remains to prove that if $(E, <)$ is a well-ordering, then there exists $\alpha \in \mathrm{Ord}$ and an isomorphism $f : (E, <) \to (\alpha, \in)$. Note that we may assume that $E \neq \varnothing$, for otherwise $(E, <) \simeq (0, \in)$. Consider the set

$$A := \{x \in E : (E_{<x}, <) \text{ is isomorphic to an ordinal}\}.$$

To see that A is a set, we leave it as an exercise to check that the condition presented above is definite, and apply bounded separation. Notice that if $e \in E$ is least, then $(E_{<e}, <) \simeq (0, \in)$, so $e \in A$, and thus $A \neq \varnothing$.

Let $f : A \to \mathrm{Ord}$ be the definite operation such that $(E_{<x}, <) \simeq (f(x), \in)$ for all $x \in A$. By the replacement axiom, $\mathrm{Im}(f)$ is a set of ordinals. Let $\alpha \in \mathrm{Ord}$ be the least ordinal not in $\mathrm{Im}(f)$. We prove that $f : A \to \mathrm{Im}(f)$ is an isomorphism between $(E, <)$ and $(\alpha, \in)$.

(1) *f preserves the ordering; in fact, if $x, y \in E$ with $y \in A$ and $x < y$, then $x \in A$ and $f(x) \in f(y)$.* Indeed, since $x < y$, we see that $E_{<x}$ is an initial segment of $E_{<y}$. If h is the isomorphism between $(E_{<y}, <)$ and $(f(y), \in)$, then

$$h(E_{<x}) = \{\alpha \in f(y) : \alpha \in h(x)\} = h(x).$$

Hence, h restricts to an isomorphism between $E_{<x}$ and the ordinal $h(x)$. It follows that $x \in A$, and by uniqueness, we have $h(x) = f(x)$. Thus, $f(x) \in f(y)$.

(2) $\alpha = \mathrm{Im}(f)$. Suppose that $\beta \in \alpha$. By our choice of α, we have $\beta \in \mathrm{Im}(f)$. Conversely, suppose that $\beta \in \mathrm{Im}(f)$. Let h be the isomorphism between $(E_{<x}, <)$ and $(\beta, \in)$ for some $x \in A$. Then $\beta \neq \alpha$. If $\alpha < \beta$, then $\alpha = h(y)$ for some $y < x$. By the proof of (1), we have $y \in A$ and $f(y) = h(y) = \alpha$, contradicting $\alpha \notin \mathrm{Im}(f)$. Thus, $\beta \in \alpha$, as desired.

(3) *f is injective.* Suppose that $f(x) = f(y)$. If $x < y$, then $(E_{<x}, <)$ is a proper initial segment of $(E_{<y}, <)$ which is isomorphic to $(E_{<y}, <)$, contradicting Lemma 10.2. An analogous argument can be made for the $y < x$ case. Hence, $x = y$.

(4) $A = E$. Suppose for a contradiction that $E \setminus A$ is non-empty, and let $x \in E \setminus A$ be least. By (1), no element less than x is in A. That is, $A = E_{<x}$. We have already proved that f is an isomorphism between $(A, <)$ and $(\alpha, \in)$, so this implies that $x \in A$, a contradiction.

This proves that f is an isomorphism between $(E, <)$ and $(\alpha, \in)$, as required. □

11 Equinumerosity

Definition 11.1 (Equinumerous). Two sets A and B are said to be **equinumerous** if there exists a bijective function $f : A \to B$.

Proposition 11.2 (Schroeder-Bernstein). Two sets A and B are equinumerous if and only if there exist injective functions $A \to B$ and $B \to A$.

Proof. The forward direction is immediate. Suppose that there exist injective functions $A \hookrightarrow B$ and $B \hookrightarrow A$. Then we may consider the injection

$$\underbrace{A \hookrightarrow B \hookrightarrow A}_{f}.$$

It suffices to prove that if $f : X \to X$ is an injective function from a set X to itself, and $X \supseteq Y \supseteq f(X)$, then X and Y are equinumerous. Indeed, observe that

$$X \supseteq Y \supseteq f(X) \supseteq f(Y) \supseteq f^2(X) \supseteq f^2(Y) \supseteq \cdots .$$

Consider the set $Z := (X \setminus Y) \cup (f(X) \setminus f(Y)) \cup \cdots$. Let $W := X \setminus Z$. Clearly, $X = Z \sqcup W$ is a disjoint union. On the other hand, we leave it as an exercise to check that $Y = f(Z) \sqcup W$. Then $g : X \to Y$ given by

$$g(x) = \begin{cases} f(x) & x \in Z \\ \mathrm{id}(x) & x \in W \end{cases}$$

is a bijection. □

Definition 11.3 (Finite, countable). A set is **finite** if it is equinumerous to some $n \in \omega$. A set is **countable** if it is finite or equinumerous with ω.

Lemma 11.4. Let $\alpha \in \mathrm{Ord}$ be infinite. Then α is equinumerous with $\alpha + 1$.

Proof. Define $f : \alpha + 1 \to \alpha$ as follows:

$$f(x) = \begin{cases} x + 1 & \text{if } x \in \omega \\ 0 & \text{if } x = \alpha \\ x & \text{otherwise.} \end{cases}$$

Since α is infinite, we have $\alpha \notin \omega$, so f is well-defined on all of $\alpha + 1$. It is surjective as its image clearly contains ω, and if $x \in \alpha \setminus \omega$, then $f(x) = x$. It is also clearly injective. □

The above lemma tells us that ordinals are not good at measuring sizes of sets. We need something better to do so.

Definition 11.5 (Cardinal). A **cardinal** is an ordinal that is not equinumerous with any strictly lesser ordinal. We denote by Card the class of all cardinals.

We leave it as an exercise to check that Card is a class. Note that Card is a subclass of Ord.

Remark 11.6. By Lemma 11.4, no infinite successor ordinal is a cardinal. Hence, every infinite cardinal is a limit ordinal. Conversely, not all limit ordinals are cardinals. For instance, $\omega \cdot 2$ is equinumerous with ω, but $\omega \cdot 2$ is a limit ordinal.

Example 11.7. The following can be proved using induction:

(a) Every $n \in \omega$ is a cardinal.

(b) ω itself is a cardinal.

Are there any other cardinals? The following proposition allows us to find uncountable cardinals, namely, cardinals larger than ω.

PROPOSITION 11.8. For every set E, there is a unique cardinal $h(E)$ which is the least ordinal not equinumerous with a subset of E.

PROOF. We show that

$$X = [[\alpha \in \text{Ord} \mid \alpha \text{ is equinumerous with a subset of } E]]$$

is a set. Consider

$$W = \{(A, \prec) : A \subseteq E \text{ and } \prec \text{ is a well-ordering of } A\},$$

which is a set by bounded separation, as each $(A, \prec) \in \mathcal{P}(E) \times \mathcal{P}(E \times E)$.

Let $F : W \to \text{Ord}$ be the definite operation on W such that $F(A, \prec)$ is the unique ordinal that is isomorphic to $(A, \prec)$, which can be done due to Theorem 10.3. By the replacement axiom, $\text{Im}(F) \subseteq \text{Ord}$ is a set.

We leave it as an exercise to show that $\text{Im}(F) = X$. Then, the least ordinal not in $\text{Im}(F)$ is our desired $h(E)$. Note that $h(E)$ is a cardinal, since if $\alpha < h(E)$, then $\alpha \notin X$, and hence α would be equinumerous with some subset of E. □

12 Axiom of choice

Last time, for every set E, we constructed cardinals $h(E)$. In particular, we have $h(\omega) > \omega$, so we have a cardinal which is uncountable.

However, we want more out of cardinals. We want every set to be equinumerous to a cardinal. This would imply that every set A has a bijection f such that

$$A \xrightarrow{f} \alpha \in \mathrm{Card} \subseteq \mathrm{Ord}\,.$$

In particular, as $\alpha \in \mathrm{Ord}$, we have the well-ordering $(\alpha, \in)$. We wish to use this to define a well-ordering on A. For $a, b \in A$, we want to define $a < b$ if and only if $f(a) \in f(b)$, so that $(A, <)$ is a well-ordering. But we cannot prove in **ZF** that every set admits a well-ordering. Hence, we require the axiom of choice.

Definition 12.1 (Choice function). For any set $\mathcal{F}$, a **choice function on** $\mathcal{F}$ is a function $c : \mathcal{F} \to \bigcup \mathcal{F}$ such that for all $F \in \mathcal{F}$, we have $c(F) \in F$.

Axiom 12.2 (Axiom of Choice). For any set $\mathcal{F}$, if $\varnothing \notin \mathcal{F}$, then there exists a choice function on $\mathcal{F}$.

Remark 12.3. The axiom of choice (AC) is a set existence axiom. For any set $\mathcal{F}$ not containing $\varnothing$, there is a set $\Gamma \subseteq \mathcal{F} \times \bigcup \mathcal{F}$ such that for all $F \in \mathcal{F}$, there is a unique $a \in \bigcup \mathcal{F}$ such that $(F, a) \in \Gamma$ and $a \in F$.

Remark 12.4. We do not always need AC to find a choice function. For example, suppose that we have the singleton set $\mathcal{F} = \{F\}$, where $F \neq \varnothing$. Let $a \in F$. Consider $\Gamma := \{(F, a)\}$. Observe that this is a set by **ZF**. Note that $\Gamma \subseteq \mathcal{F} \times F = \mathcal{F} \times \bigcup \mathcal{F}$, and we chose $a \in F$, so Γ is a choice function on $\mathcal{F}$.

Exercise 12.5. Show that if $\mathcal{F}$ is a finite set, then we can construct a choice function on $\mathcal{F}$ without using AC.

Exercise 12.6. Suppose $(A, <)$ is a well-ordering. Let $\mathcal{F} := \mathcal{P}(A) \setminus \{\varnothing\}$. Prove that $c : \mathcal{F} \to \bigcup \mathcal{F}$ where $c(F)$ is the $<$-least element of F for each $F \in \mathcal{F}$ is a choice function without using AC.

Theorem 12.7. In **ZF**, the following are equivalent:

(1) Axiom of Choice.

(2) Well-Ordering Principle: Every set admits a well-ordering.

(3) Zorn's Lemma: Suppose (E, R) is a strict poset that satisfies

$$\begin{gathered}\text{If } C \subseteq E \text{ is totally ordered by } R \text{ (i.e. it is a } \textbf{chain} \text{ in } (E, R)),\\ \text{then } C \text{ has an upper } R\text{-bound in } E.\end{gathered} \tag{$\star$}$$

Then (E, R) has a maximal element.

Proof. We prove (1) implies (2). The remaining proofs are in the complete notes (Theorem 2.8).

Suppose A is an arbitrary set. Let c be a choice function on $\mathcal{P}(A) \setminus \{\varnothing\}$, which exists by AC. Define, by transfinite recursion, a definite operation F on Ord by

$$F(\alpha) = \begin{cases} c(A \setminus \mathrm{Im}(F \restriction_\alpha)) & \text{if } A \setminus \mathrm{Im}(F \restriction_\alpha) \neq \varnothing \\ \theta & \text{otherwise} \end{cases}$$

where θ is an ordinal not in A fixed in advance.

If $F(\alpha) \neq \theta$ for all α, then $F : \mathrm{Ord} \to A$. In particular, $F \restriction_{h(A)} : h(A) \to A$, where $h(A)$ is the least ordinal not equinumerous with any subset of A, as in Proposition 11.8.

Check that $F \restriction_{h(A)}$ is injective (by transfinite induction).

But this implies that $h(A)$ is equinumerous with a subset of A, a contradiction.

Hence, we must have $F(\alpha) = \theta$ for some α. Let $\alpha \in \text{Ord}$ be least such that $F(\alpha) = \theta$. Hence, $F \restriction_\alpha : \alpha \to A$ is injective (check). Since $F(\alpha) = \theta$, we have $\text{Im}(F \restriction_\alpha) = A$, so it follows that $F \restriction_\alpha$ is a bijection.

Now, define $a < b$ in A by $F \restriction_\alpha^{-1} (a) \in F \restriction_\alpha{}^{-1}(b)$ in α. Then, $(A, <)$ is a well-ordering. □

There are many more statements which are equivalent to AC, such as the following:

(1) Every surjective function has a right inverse.

(2) All vector spaces have a basis.

13 Cardinality

PROPOSITION 13.1. Assume the axiom of choice. Every set is equinumerous to a unique cardinal.

PROOF. Let X be a set. By the axiom of choice and Theorem 12.7, there is a well-ordering $<$ on X. By Theorem 10.3, $(X, <)$ is order isomorphic with $(\alpha, \in)$ for some $\alpha \in \mathrm{Ord}$. Let $S := \{\beta \leq \alpha : X \text{ is equinumerous with } \beta\}$, and let α_0 be least in S. If $\beta < \alpha_0$, then $\beta \notin S$, so β is not equinumerous with X. As α_0 is equinumerous with X, we see that β is not equinumerous with α_0. Hence, α_0 is a cardinal. To see that this cardinal is unique, we note that by definition, distinct cardinals are not equinumerous. □

DEFINITION 13.2 (Cardinality). Let X be a set. The **cardinality** of X, denoted by $|X|$, is the unique cardinal that X is equinumerous to.

REMARK 13.3. Let X, Y be sets. Then $|X| = |Y|$ if and only if X and Y are equinumerous.

PROPOSITION 13.4. Let X, Y be sets. Then $|X| \leq |Y|$ if and only if there is an injective map $X \to Y$.

PROOF. Let $\kappa = |X|$ and $\lambda = |Y|$. If $\kappa \leq \lambda$, then $\kappa \subseteq \lambda$. Then there are bijections $f : X \to \kappa$ and $g : \lambda \to Y$, and hence there is an injection

$$X \xrightarrow{f} \kappa \subseteq \lambda \xrightarrow{g} Y.$$

Conversely, suppose there is an injective map $h : X \to Y$. Again, we have bijections $f : X \to \kappa$ and $g : \lambda \to Y$, and so we have an injection

$$\kappa \xrightarrow{f^{-1}} X \xrightarrow{h} Y \xrightarrow{g^{-1}} \lambda.$$

If $\lambda < \kappa$, then $\lambda \subseteq \kappa$. By Schroeder-Bernstein, we see that κ and λ are equinumerous. This is a contradiction as both κ and λ are cardinals. Hence, $\kappa \leq \lambda$. □

PROPOSITION 13.5. Let A, B be sets. If $f : A \to B$ is a function, then $|\operatorname{Im}(f)| \leq |A|$.

PROOF. Consider the definite operation that maps $b \in B$ to $f^{-1}(b) = \{a \in A : f(a) = b\}$ (this is called the fiber). Then the map $g : \operatorname{Im}(f) \to \mathcal{F} := \{f^{-1}(b) : b \in \operatorname{Im}(f)\}$ given by $b \mapsto f^{-1}(b)$ is also a definite operation. By the axiom of choice, there exists a choice function $c : \mathcal{F} \to \bigcup \mathcal{F} \subseteq A$. It follows that

$$s : \operatorname{Im}(f) \xrightarrow{g} \mathcal{F} \xrightarrow{c} A$$

is injective. By Proposition 13.4, $|\operatorname{Im}(f)| \leq |A|$. □

For the remainder of this course, we will assume the axiom of choice along with the Zermelo-Fraenkel axioms of set theory (**ZFC**) without explicitly saying so.

14 Enumerating cardinals

Let $\kappa \in \mathrm{Card}$. Recall that $h(\kappa)$ as in Proposition 11.8 is the least cardinal strictly bigger than κ. Indeed, if $\lambda \in \mathrm{Card}$ is such that $\lambda > \kappa$, then $|\lambda| = \lambda > \kappa = |\kappa|$. Hence, there is no injective function $\lambda \to \kappa$ by Proposition 13.4, and so $\lambda \geq h(\kappa)$.

NOTATION 14.1. As such, we denote $h(\kappa)$ by κ^+.

Note that we now have two different notions of successors between ordinals and cardinals. For κ infinite, we have $\kappa + 1 \neq \kappa^+$.

In this lecture, we will give an ordinal-valued enumeration of all infinite cardinals. In particular, we will have a strictly increasing definite operation $\mathrm{Ord} \to \mathrm{Card} : \alpha \mapsto \aleph_\alpha$ onto the infinite cardinals. We define this via transfinite recursion via:

(1) $\aleph_0 := \omega$.

(2) $\aleph_{\alpha+1} := \aleph_\alpha^+$ for all $\alpha \in \mathrm{Ord}$.

(3) $\aleph_\beta := \sup\{\aleph_\alpha : \alpha < \beta\}$ for all limit ordinals $\beta > 0$.

LEMMA 14.2. For all $\alpha \in \mathrm{Ord}$, $\aleph_\alpha$ is an infinite cardinal.

PROOF. It is obvious that these are infinite, and that $\aleph_\alpha$ is a cardinal at the zero and successor stages. It suffices to check that $\aleph_\beta$ is a cardinal when $\beta > 0$ is a limit ordinal. Let $\alpha \in \mathrm{Ord}$ be such that $\alpha < \aleph_\beta$. By definition, there is $\gamma < \beta$ such that $\alpha < \aleph_\gamma$. By our inductive hypothesis, $\aleph_\gamma$ is an infinite cardinal. Then $\alpha < \aleph_\gamma \leq \aleph_\beta$, and hence $|\alpha| < |\aleph_\gamma| = \aleph_\gamma \leq \aleph_\beta = |\aleph_\beta|$. Thus, α and $\aleph_\beta$ are not equinumerous, so $\aleph_\beta$ is a cardinal. □

LEMMA 14.3. Let $\alpha, \beta \in \mathrm{Ord}$ be such that $\alpha < \beta$. Then $\aleph_\alpha < \aleph_\beta$.

PROOF. We proceed by transfinite induction on β. There is nothing to prove for $\beta = 0$. For $\beta = \gamma + 1$, we have $\aleph_\beta = \aleph_\gamma^+ > \aleph_\gamma$, and the result follows from the inductive hypothesis. Finally, if β is a non-zero limit ordinal, then there exists $\gamma < \beta$ such that $\alpha < \gamma$. Then $\aleph_\alpha < \aleph_\gamma$ by the inductive hypothesis, and $\aleph_\gamma \leq \aleph_\beta$ since $\aleph_\gamma \subseteq \aleph_\beta$. Hence, we have $\aleph_\alpha < \aleph_\beta$, as desired. □

LEMMA 14.4. For all $\alpha \in \mathrm{Ord}$, $\alpha \leq \aleph_\alpha$. The inequality is strict if α is a successor ordinal.

PROOF. We proceed by transfinite induction on α. Clearly, $0 \leq \aleph_0$. For $\alpha = \beta + 1$, we know by the inductive hypothesis that $\beta \leq \aleph_\beta$, and hence $\alpha = \beta + 1 \leq \aleph_\beta + 1$. On the other hand, we have $\aleph_\beta + 1 < |\aleph_\beta|^+ = \aleph_\alpha$, so $\alpha < \aleph_\alpha$. Now, suppose that α is a non-zero limit ordinal. For every $\beta < \alpha$, we have that $\beta \leq \aleph_\beta < \aleph_\alpha$, where the first inequality is by the inductive hypothesis, and the second inequality follows from Lemma 14.3. Thus, $\alpha = \sup\{\beta : \beta < \alpha\} \leq \aleph_\alpha$. □

EXERCISE 14.5. The inequality in the above lemma may not be strict. Consider the sequence $\alpha_0 = 0$ and $\alpha_{n+1} = \aleph_{\alpha_n}$. Verify that for $\alpha := \sup\{\alpha_n : n < \omega\}$, we have $\aleph_\alpha = \alpha$. In fact, this works for any ordinal α_0, not just 0.

PROPOSITION 14.6. Every infinite cardinal is of the form $\aleph_\alpha$ for some $\alpha \in \mathrm{Ord}$.

PROOF. Suppose κ is an infinite cardinal. By the previous lemma, we have $\kappa \leq \aleph_\kappa < \aleph_{\kappa+1}$. Hence, it suffices to prove that for all ordinals β and every infinite cardinal $\kappa < \aleph_\beta$, there exists $\alpha \in \mathrm{Ord}$ such that $\kappa = \aleph_\alpha$. We proceed by transfinite induction on β. For $\beta = 0$, there is nothing to prove. Suppose that this holds for β. Let $\kappa < \aleph_{\beta+1} = \aleph_\beta^+$. Then, it must be that $\kappa \leq \aleph_\beta$. If $\kappa = \aleph_\beta$, we are done. If $\kappa < \aleph_\beta$, then the result follows from the inductive hypothesis. Finally, suppose $\beta > 0$ is a limit ordinal, and let $\kappa < \aleph_\beta$. Then $\kappa < \aleph_\gamma$ for some $\gamma < \beta$. By the inductive hypothesis, $\kappa = \aleph_\alpha$ for some $\alpha \in \mathrm{Ord}$. □

As a result of Proposition 14.6, we see that $\alpha \mapsto \aleph_\alpha$ is a strict ordinal enumeration of all infinite cardinals.

To finish off this lecture, we briefly discuss the continuum hypothesis.

THEOREM 14.7 (Cantor's diagonalization). For any set E, $|\mathcal{P}(E)| > |E|$.

PROOF. Observe that $x \mapsto \{x\}$ is an embedding of E into $\mathcal{P}(E)$, so we have $|E| \leq |\mathcal{P}(E)|$. Now, suppose for a contradiction that there exists a bijective function $f : E \to \mathcal{P}(E)$. Consider the set $\Delta := \{x \in E : x \notin f(x)\}$. Notice that $\Delta \in \mathcal{P}(E)$ and hence there is some $x_0 \in E$ such that $\Delta = f(x_0)$. If $x_0 \in \Delta$, then by definition, we have $x_0 \notin f(x_0) = \Delta$, a contradiction. Hence, we must have $x_0 \notin \Delta$. But then $x_0 \in f(x_0) = \Delta$, another contradiction. Thus, no such bijection f exists, and so $|E| < |\mathcal{P}(E)|$. □

Due to this theorem, we would like to ask what cardinal $|\mathcal{P}(\aleph_0)|$ is, based on the previous hierarchy we have constructed. The statement that $|\mathcal{P}(\aleph_0)| = \aleph_1$ is called the **continuum hypothesis** (CH). Moreover, the **generalized continuum hypothesis** (GCH) states that $|\mathcal{P}(\kappa)| = \kappa^+$ for all cardinals κ. These statements are independent of **ZFC**; that is, they cannot be proved in **ZFC**, nor can their negations. Unlike the axiom of choice, the (generalized) continuum hypothesis is not indispensable to most of contemporary mathematics. In this course, we will not assume CH and GCH, nor their negations.

15 Cardinal arithmetic

In the interest of time, the topics in this section were covered briefly. It is encouraged to read Section 3.3 in the complete notes thoroughly to understand all the details.

Definition 15.1 (Cardinal sum). Let $\kappa_1, \kappa_2 \in \text{Card}$. The **cardinal sum** of κ_1 and κ_2 is

$$\kappa_1 + \kappa_2 := |X_1 \cup X_2|,$$

where X_1 and X_2 are disjoint, $|X_1| = \kappa_1$, and $|X_2| = \kappa_2$.

Remark 15.2.

(a) We need to prove this is well-defined. That is, the choice of the disjoint sets X_1, X_2 does not matter. The proof can be seen in Lemma 3.9 of the notes.

(b) If X_1, X_2 are not necessarily disjoint sets, then $|X_1 \cup X_2| \leq |X_1| + |X_2|$. Indeed, let $X_1' = X_1 \times \{1\}$ and $X_2' = X_2 \times \{2\}$. These are clearly disjoint, so we see that

$$|X_1' \cup X_2'| = |X_1'| + |X_2'| = |X_1| + |X_2|.$$

Then, we can consider the surjective map $X_1' \cup X_2' \to X_1 \cup X_2 : (a, b) \mapsto a$. It follows from Proposition 13.4 that

$$|X_1 \cup X_2| \leq |X_1' \cup X_2'| = |X_1| + |X_2|.$$

(c) A warning: ordinal and cardinal sum are not the same. Nonetheless, we use the same symbol to denote both. Context will make it clear what is meant.

Definition 15.3 (Cardinal product). Let $\kappa_1, \kappa_2 \in \text{Card}$. The **cardinal product** of κ_1 and κ_2 is

$$\kappa_1 \cdot \kappa_2 := |X_1 \times X_2|,$$

where X_1, X_2 are sets such that $|X_1| = \kappa_1$ and $|X_2| = \kappa_2$.

Remark 15.4.

(a) As with cardinal sum, we need to prove well-definedness. In particular, we can simply pick κ_1 and κ_2 to be our sets, as we do not require them to be disjoint. The proof can be seen in Lemma 3.9.

(b) Cardinal product is not the same as ordinal product, so we have notational ambiguity.

By the following theorem, cardinal arithmetic of infinite cardinals trivializes to computing maxima.

Theorem 15.5. Let $\kappa_1, \kappa_2 \in \text{Card}$ with $\kappa_1 \leq \kappa_2$ and κ_2 infinite.

(a) $\kappa_1 + \kappa_2 = \kappa_2$.

(b) If $\kappa_1 \neq 0$, then $\kappa_1 \cdot \kappa_2 = \kappa_2$.

Proof. See notes (Corollary 3.14). □

We now extend the notions of cardinal sum and product to arbitrary (possibly infinite) sets of cardinals. First, we need to introduce some preliminaries on arbitrary sequences of sets.

Definition 15.6. Let I be a set.

(a) By an **I-sequence of sets**, we understand a function $f : I \to \text{Sets}$. We often use the notation $(X_i : i \in I)$, where $X_i := f(i)$.

(b) Let $(X_i : i \in I)$ be an I-sequence of sets. The **Cartesian product**, denoted by $\times_{i \in I} X_i$, is the set of all functions $f : I \to \bigcup_{i \in I} X_i$ such that $f(i) \in X_i$.

Definition 15.7 (Generalized cardinal sum). Let $(\kappa_i : i \in I)$ be an I-sequence of cardinals. The **generalized cardinal sum** is

$$\sum_{i\in I} \kappa_i := \left| \bigcup_{i\in I} X_i \right|,$$

where $(X_i : i \in I)$ is an I-sequence of sets that are pairwise disjoint and $|X_i| = \kappa_i$ for all $i \in I$.

Remark 15.8.

(a) We need to prove well-definedness. This can be seen in Lemma 3.17. Note that it is the same proof as the finite case, except we need to use the axiom of choice here.

(b) If $(X_i : i \in I)$ is given by $f : I \to \text{Sets} : i \mapsto X_i$, then $\bigcup_{i\in I} X_i := \bigcup \text{Im}(f)$.

Notation 15.9. Let $(\kappa_i : i \in I)$ be given by $f : I \to \text{Sets}$ with $f(i) = \kappa_i$ for all $i \in I$. We write $\sup_{i\in I} \kappa_i := \sup \text{Im}(f)$.

From the following theorem, infinite cardinal sums simply reduce to computing suprema.

Theorem 15.10. Suppose I is an infinite set and $(\kappa_i : i \in I)$ is an I-sequence of non-zero cardinals. Then:

(a) $\sup_{i\in I} \kappa_i$ is a cardinal.

(b) $\sum_{i\in I} \kappa_i = \max\{|I|, \sup_{i\in I} \kappa_i\}$.

Proof. The proof of (a) is Lemma 3.19 in the notes, and the proof of (b) is Proposition 3.20 in the notes. □

Definition 15.11 (Generalized cardinal product). Let $(\kappa_i : i \in I)$ be an I-sequence of cardinals. The **generalized cardinal product** is

$$\prod_{i\in I} \kappa_i := |\times_{i\in I} X_i|,$$

where $(X_i : i \in I)$ is a sequence of sets with $|X_i| = \kappa_i$ for all $i \in I$.

On the other hand, infinite cardinal products are more interesting, as they do not reduce to computing suprema. Let I be a set such that $|I| \geq 2$, and suppose $\kappa_i = 2$ for all $i \in I$. The map that assigns each $a = (a_i : i \in I) \in \times_{i\in I} 2$ to the set $\{i \in I : a_i = 1\} \in \mathcal{P}(I)$, is a bijection between $\times_{i\in I} 2$ and $\mathcal{P}(I)$ (check this). Hence, $\prod_{i\in I} 2 = |\mathcal{P}(I)| > |I|$, where the last inequality follows from Cantor's diagonalization (Theorem 14.7).

Definition 15.12 (Cardinal exponentiation). Let $\kappa, \lambda \in \text{Card}$. We define κ^λ to be the cardinality of the set of all functions $f : \lambda \to \kappa$.

Lemma 15.13. For $\kappa, \lambda \in \text{Card}$, we have $\prod_{i<\lambda} \kappa$.

Proof. The λ-th Cartesian power of a set is exactly the set of functions from λ to that set. □

In particular, we see that $2^\lambda = |\mathcal{P}(\lambda)|$.

Lemma 15.14. Let $\kappa, \lambda, \mu \in \text{Card}$.

(a) If $\lambda \leq \mu$, then $\kappa^\lambda \leq \kappa^\mu$ and $\lambda^\kappa \leq \mu^\kappa$.

(b) $\kappa^{\lambda+\mu} = \kappa^\lambda \cdot \kappa^\mu$.

(c) $(\kappa^\lambda)^\mu = \kappa^{\lambda\cdot\mu}$.

Proof. See notes (Lemma 3.23). □

The following theorem is a generalization of Cantor's diagonalization.

THEOREM 15.15 (König's Theorem). Suppose that $(\kappa_i : i \in I)$ and $(\lambda_i : i \in I)$ are both sequences of cardinals, and $\kappa_i < \lambda_i$ for all $i \in I$. Then

$$\sum_{i \in I} \kappa_i < \prod_{i \in I} \lambda_i.$$

PROOF. See notes (Theorem 3.25). □

This concludes the set theory portion of the course, and we begin model theory starting from the next section.

Part 2

Model Theory

16 Structures

Definition 16.1 (Structure). A **structure** $\mathcal{M}$ consists of:

- a non-empty set M, called the **universe** of $\mathcal{M}$;
- a sequence $(c_i : i \in I_C)$ of elements from $\mathcal{M}$, called the **constants** of $\mathcal{M}$;
- a sequence $(f_i : i \in I_F)$ of functions on powers of M, called the **basic functions** of $\mathcal{M}$ (i.e. each $f_i : M^{n_i} \to M$, and n_i is said to be the **arity** of f_i);
- a sequence $(R_i : i \in I_R)$ of subsets of powers of M, called the **basic relations** of $\mathcal{M}$ (i.e. each $R_i \subseteq M^{k_i}$, with k_i being the **arity** of R_i).

The **signature** of $\mathcal{M}$ is given by $((c_i : i \in I_C), (f_i : i \in I_F), (R_i : i \in I_R))$.

Remark 16.2. By convention, we have $M^0 = 1 = \{0\}$. Hence, if f is a basic relation of arity 0, then it is determined by $f(0) \in M$. In particular, 0-ary functions are just constants. Thus, we usually only deal with basic functions of arity at least 1.

Remark 16.3. Note that I_C, I_F, I_R may be empty, but the universe M must be non-empty.

Example 16.4. Consider the real numbers $\mathbb{R}$.

(a) If we are interested in the ordering on $\mathbb{R}$, we may consider the structure $\mathcal{M} = (\mathbb{R}, <)$, where $\mathbb{R}$ is the universe and $<$ is a basic binary relation, namely $\{(a, b) \in \mathbb{R}^2 : a < b\}$.

(b) If we are interested in the additive group of reals, then we may consider the structure $\mathcal{M} = (\mathbb{R}, 0, +, -)$ where $\mathbb{R}$ is the universe, 0 is a constant, $+$ is a binary basic function, and $-$ is a unary basic function that takes the additive inverse.

(c) If we want $\mathbb{R}$ as a ring, we may take $\mathcal{M} = (\mathbb{R}, 0, 1, +, -, \cdot)$ where 1 is a constant, $\cdot$ is a binary basic function, and everything else is as in (b).

(d) The ordered ring of reals is given by $\mathcal{M} = (\mathbb{R}, 0, 1, +, -, \cdot, <)$.

Note that all of these examples have the same universe, but different signatures.

Definition 16.5 (Expansion, Reduct). Suppose that $\mathcal{M}$ and $\mathcal{N}$ are structures. We say that $\mathcal{N}$ is an **expansion** of $\mathcal{M}$, or $\mathcal{M}$ is a **reduct** of $\mathcal{N}$, if they have the same universe and the signature of $\mathcal{M}$ is contained in the signature of $\mathcal{N}$.

Example 16.6. $(\mathbb{R}, 0, +, -)$ is a reduct of $(\mathbb{R}, 0, 1, +, -, \cdot)$.

An important theme in model theory is to ask questions such as the following:

(1) Can we recover $(\mathbb{R}, 0, 1, +, -, \cdot)$ from its reduct $(\mathbb{R}, 0, 1, +)$?

(2) Can we recover $(\mathbb{R}, 0, 1, +, -, \cdot, <)$ from $(\mathbb{R}, 0, 1, +, -, \cdot)$?

It turns out that the answer to (1) is no, while the answer to (2) is yes. However, we need to introduce definability, which will be later in this course.

To discuss structures with the same signature, it is useful to introduce the notion of languages.

Definition 16.7 (Language). A **language** L consists of the following sets of symbols:

- a set L^C of **constant symbols**;
- a set L^F of **function symbols** together with a positive integer n_f for each $f \in L^F$, called the **arity** of f;

- a set L^R of **relation symbols** together with a positive integer k_R for each $R \in L^R$, called the **arity** of R.

Definition 16.8 (L-structure). Let L be a language. An **L-structure** is a structure $\mathcal{M}$ together with bijective correspondences $L^C \leftrightarrow I_C$, $L^R \leftrightarrow I_R$, and $L^F \leftrightarrow I_F$ that preserve arity. Namely, each constant symbol $c \in L^C$ is associated with a constant $c^{\mathcal{M}} \in M$ of $\mathcal{M}$, each n-ary function symbol $f \in L^F$ is associated with an n-ary basic function $f^{\mathcal{M}} : M^n \to M$ of $\mathcal{M}$, and each k-ary relation symbol $R \in L^R$ is associated with a k-ary basic relation $R^{\mathcal{M}} \subseteq M^k$ of $\mathcal{M}$. These constants, functions, and relations $c^{\mathcal{M}}, f^{\mathcal{M}}, R^{\mathcal{M}}$ are said to be the **interpretations** in $\mathcal{M}$ of the corresponding symbols.

17 Embeddings

A language L is made up of constant, function, and relation symbols with arity. An L-structure is a non-empty set together with an interpretation of the symbols in L.

The only requirements to be an L-structure are that constant symbols are interpreted as elements of the universe, n-ary function symbols are interpreted as n-ary functions on the universe, and k-ary relation symbols are interpreted as subsets of the k-th power of the universe.

Example 17.1. Let $L = \{0, +, -\}$ be the language of additive groups, namely $L^C = \{0\}$, $L^F = \{+, -\}$ where $+$ is binary and $-$ is unary, and $L^R = \varnothing$. Then $\mathcal{R} = (\mathbb{R}, 0, +, -)$ and $\mathcal{Z} = (\mathbb{Z}/4\mathbb{Z}, 0, +, -)$ are both L-structures.

Note that this is an abuse of notation; we should really be writing $\mathcal{R} = (\mathbb{R}, 0^{\mathcal{R}}, +^{\mathcal{R}}, -^{\mathcal{R}})$ and $\mathcal{Z} = (\mathbb{Z}/4\mathbb{Z}, 0^{\mathcal{Z}}, +^{\mathcal{Z}}, -^{\mathcal{Z}})$. However, this is unwieldy, so we often drop the superscripts distinguishing between the symbols and their interpretations.

Every group can be viewed naturally as an L-structure, but not every L-structure is a group.

Example 17.2. Consider the same language $L = \{0, +, -\}$. Then $\mathcal{N} = (\mathbb{Z}, 0^{\mathcal{N}}, +^{\mathcal{N}}, -^{\mathcal{N}})$ where $0^{\mathcal{N}} = 1731$, $+^{\mathcal{N}} : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z} : (a, b) \mapsto \max\{a, b\}$, and $-^{\mathcal{N}} : \mathbb{Z} \to \mathbb{Z} : a \mapsto 0$ is an L-structure, but is not a group.

Previously, we discussed expansions and reducts, where the universe is unchanged but the signature is expanded or reduced. We can also think about structures in the opposite way, where the universe is expanded or reduced, but the signature (or rather, the language) is the same.

Definition 17.3 (L-embedding). Suppose that L is a language, and let $\mathcal{M}$ and $\mathcal{N}$ be L-structures. An **L-embedding** of $\mathcal{M}$ in $\mathcal{N}$, written $j : \mathcal{M} \to \mathcal{N}$, is an injective function $j : M \to N$ satisfying:

(i) for all $c \in L^C$, $j(c^{\mathcal{M}}) = c^{\mathcal{N}}$;

(ii) for all $f \in L^F$ and $a \in M^{n_f}$, $j(f^{\mathcal{M}}(a)) = f^{\mathcal{N}}(j(a))$;

(iii) for all $R \in L^R$ and $a \in M^{k_R}$, $a \in R^{\mathcal{M}}$ if and only if $j(a) \in R^{\mathcal{N}}$.

A surjective L-embedding is said to be an **L-isomorphism**.

Definition 17.4 (L-substructure). Let $\mathcal{M}$ and $\mathcal{N}$ be as in Definition 17.3. If $M \subseteq N$ and the inclusion map $M \to N$ is an L-embedding, then we call $\mathcal{M}$ an **L-substructure** of $\mathcal{N}$, or $\mathcal{N}$ an **L-extension** of $\mathcal{M}$. We denote this by $\mathcal{M} \subseteq \mathcal{N}$.

In particular, $\mathcal{M} \subseteq \mathcal{N}$ if and only if $M \subseteq N$, $c^{\mathcal{N}} = c^{\mathcal{M}}$ for all $c \in L^C$, $f^{\mathcal{N}} \restriction_{M^n} = f^{\mathcal{M}}$ for all n-ary function symbols $f \in L^R$, and $R^{\mathcal{N}} \cap M^k = R^{\mathcal{M}}$ for all k-ary relation symbols $R \in L^R$.

Exercise 17.5. Suppose that L is a language, $\mathcal{N}$ is an L-structure, and $A \subseteq N$. Then A is the universe of an L-substructure of $\mathcal{N}$ if and only if $A \neq \varnothing$, A contains all constants of $\mathcal{N}$, and A is preserved under all basic functions of $\mathcal{N}$. In this case, there is a unique L-substructure $\mathcal{A} \subseteq \mathcal{N}$ whose universe is A.

We now give some examples of structures and their substructures.

Example 17.6.

- For $\mathcal{M} = (\mathbb{R})$, every non-empty subset of $\mathbb{R}$ is a substructure.
- For $\mathcal{M} = (\mathbb{R}, 0, +)$, the substructures of $\mathcal{M}$ are the semigroups.
- For $\mathcal{M} = (\mathbb{R}, 0, +, -)$, the substructures are precisely the subgroups.
- For $\mathcal{M} = (\mathbb{R}, 0, 1, +, -, \cdot)$, the substructures are precisely the subrings.
- For $\mathcal{M} = (\mathbb{R}, <)$, note that the relations do not impose any conditions, and so the substructures are the non-empty subsets of $\mathbb{R}$ with the induced orderings.

Example 17.7. Fix a field F. The language of F-vector spaces is $L = \{0, +, -, (\lambda_a)_{a\in F}\}$, where each λ_a is a unary function symbol. If V is an F-vector space, we can make an L-structure $\mathcal{V}$ where:

- V is the universe;
- $0^{\mathcal{V}}$ is the zero vector;
- $+^{\mathcal{V}}$ denotes vector addition;
- $-^{\mathcal{V}}$ gives the negative of a vector;
- $\lambda_a^{\mathcal{V}} : V \to V : v \mapsto av$ denotes scalar multiplication.

Then $\mathcal{V}$ is an L-structure where the substructures are precisely the subspaces.

18 Terms

In our definition, we saw L-substructures as a special case of L-embeddings. However, due to the following exercise, it suffices to study substructures to study L-embeddings.

EXERCISE 18.1. Suppose $j : \mathcal{M} \to \mathcal{N}$ is an L-embedding. Let $A := j(M) \subseteq N$. Then there is a unique substructure $\mathcal{A} \subseteq \mathcal{N}$ with universe A, and such that $j : \mathcal{M} \to \mathcal{A}$ is an L-isomorphism.

We want to define L-formulae. These are used:

(1) To describe properties of L-structures, such as axioms. For instance, if $L = \{<\}$, then to be able to talk about posets among all L-structures, we use formulae.

(2) To define certain subsets of a structure. For example, consider $\mathcal{R} = (\mathbb{R}, 0, 1, +, -, \times)$ and suppose we want the set of multiplicative units in $\mathcal{R}$. We can use formulae to isolate these.

Before we begin with formulae, we need to start with terms. We will make use of a fixed infinite set Var of symbols called **variables**. We assume that Var is countable.

DEFINITION 18.2 (L-term). Fix a language L. The set of **L-terms** is the set of strings of symbols defined recursively as follows:

(i) Every variable is an L-term.

(ii) Every constant symbol of L is an L-term.

(iii) If $f \in L^F$ is n-ary and $t_1, \ldots, t_n$ are L-terms, then $f(t_1, \ldots, t_n)$ is an L-term.

We write $t = t(x_1, \ldots, x_n)$ to mean that the variables appearing in t come from the list $\{x_1, \ldots, x_n\}$. This is similar to how we write multivariable polynomials.

Note that L-terms are finite strings of symbols from $\text{Var} \cup L^C \cup L^F \cup \{(\,,\,)\} \cup \{\,,\,\}$.

We will often abuse notation for readability and write things more naturally. For example, consider the language $L = \{0, 1, +, -, \times\}$. Instead of the correctly written L-term

$$\times(+(x_1, -(x_2)), \times(1, x_2)),$$

we will write $(x_1 - x_2)(1x_2)$.

DEFINITION 18.3 (Interpreting terms). Suppose $\mathcal{M}$ is an L-structure and $t = t(x_1, \ldots, x_n)$ is an L-term. We define the **interpretation of t in $\mathcal{M}$** to be the function $t^{\mathcal{M}} : M^n \to M$ defined recursively as follows:

(i) If t is x_i for some $1 \leq i \leq n$, then $t^{\mathcal{M}}(a_1, \ldots, a_n) = a_i$ (i.e. a coordinate projection).

(ii) If t is some $c \in L^C$, then $t^{\mathcal{M}}(a_1, \ldots, a_n) = c^{\mathcal{M}}$ (i.e. a constant function).

(iii) If t is $f(t_1, \ldots, t_\ell)$ where $f \in L^F$ is ℓ-ary and $t_1, \ldots, t_\ell$ are L-terms, then $t_i = t_i(x_1, \ldots, x_n)$, and hence

$$t^{\mathcal{M}}(a_1, \ldots, a_n) = f^{\mathcal{M}}(t_1^{\mathcal{M}}(a_1, \ldots, a_n), \ldots, t_\ell^{\mathcal{M}}(a_1, \ldots, a_n)).$$

REMARK 18.4. Note that $t^{\mathcal{M}}$ depends not only on t and $\mathcal{M}$, but also on the presentation $t = t(x_1, \ldots, x_n)$.

EXAMPLE 18.5. Suppose that t is a variable x.

- If $t = t(x)$, then $t^{\mathcal{M}} : M \to M : a \mapsto a$ is the identity.
- If $t = t(x, y)$, then $t^{\mathcal{M}} : M^2 \to M : (a, b) \mapsto a$.
- If $t = t(y, x)$, then $t^{\mathcal{M}} : M^2 \to M : (a, b) \mapsto b$.

EXERCISE 18.6. Suppose that $\mathcal{M}$ is an L-structure and $A \subseteq M$. Then A is the universe of a substructure if and only if $A \neq \varnothing$ and $t^{\mathcal{M}}(A^n) \subseteq A$ for every L-term $t(x_1, \ldots, x_n)$.

REMARK 18.7. Let $\mathcal{T} := \{t^{\mathcal{M}} : t \text{ is an } L\text{-term}\}$. Then $\mathcal{T}$ is the smallest collection of functions on Cartesian products of M satisfying:

- all coordinate projections are in $\mathcal{T}$;
- all constant functions $M^n \to M : (a_1, \ldots, a_n) \mapsto c^{\mathcal{M}}$ for $c \in L^C$ are in $\mathcal{T}$;
- every function $f^{\mathcal{M}}$ where $f \in L^F$ is in $\mathcal{T}$;
- the collection is closed under composition.

19 Formulae

The following is the base case of the formulae which we would like to define.

DEFINITION 19.1 (Atomic L-formula). An **atomic L-formula** is a string of symbols from

$$L \cup \{(\,,\,)\} \cup \{\,,\,\} \cup \mathrm{Var} \cup \{=\}$$

of the form

(i) $(t = s)$, where t and s are L-terms, or

(ii) $R(t_1, \ldots, t_k)$, where $R \in L^R$ is k-ary and $t_1, \ldots, t_k$ are L-terms.

As with L-terms, we will abuse notation for readability. For instance, if $L = \{<, \times\}$, then we will write $y_1 < y_2^2$ instead of $<(y_1, \times(y_2, y_2))$.

DEFINITION 19.2 (L-formulae). The set of **L-formulae** is the smallest set of strings from

$$L \cup \{(\,,\,),=\} \cup \{\,,\,\} \cup \mathrm{Var} \cup \{\neg, \wedge, \vee, \forall, \exists\}$$

satisfying the following:

(i) Every atomic L-formula is an L-formula.

(ii) If ϕ and ψ are L-formulae, then so are $(\phi \wedge \psi)$, $(\phi \vee \psi)$, and $\neg\phi$.

(iii) If ϕ is an L-formula and $x \in \mathrm{Var}$, then $\forall x\, \phi$ and $\exists x\, \phi$ are also L-formulae.

We will write $(\phi \to \psi)$ to mean $(\neg\phi \vee \psi)$, and $(\phi \leftrightarrow \psi)$ to mean $(\phi \to \psi) \wedge (\psi \to \phi)$.

DEFINITION 19.3 (Free and bound variables). Suppose that ϕ is an L-formula and $x \in \mathrm{Var}$. An occurrence of x in ϕ is said to be **bound** of it appears in the scope of a quantifier $(\forall, \exists)$. An occurrence of x is ϕ is **free** if it is not bound.

EXAMPLE 19.4. Let $L = \{\in\}$. Consider the L-formula given by

$$(x \in y) \wedge \forall z\, (z \in x \to z \in y) \wedge \forall z\, (z \in y \to (z \in x) \vee (z = x)).$$

Then every occurrence of x and y is free, but every occurrence of z is bound. Viewed in the natural way of set theory, this says that y is a successor of x.

We write $\phi = \phi(x_1, \ldots, x_n)$ to mean that all free occurrences of variables in ϕ are from $\{x_1, \ldots, x_n\}$. As before, we do not require all of $x_1, \ldots, x_n$ to appear in ϕ. For instance, we may represent the formula in the above example as $\phi(x, y)$ or $\phi(x, y, w)$.

DEFINITION 19.5 (L-sentence). An L-formula with no free variables is an **L-sentence**.

For simplification, we will assume that no variable appears in an L-formula both bound and free.

EXAMPLE 19.6. Let $L = \{<, \times, 0\}$ and consider the completely valid L-formula

$$(x < 0) \wedge \exists x\, (x^2 = y).$$

Note that the first occurrence of x is free, while the second occurrence is bound. Next, consider

$$(x < 0) \wedge \exists z\, (z^2 = y).$$

While these formulae are not the same, we will see that they have the same meaning, which is why we can make the above simplification.

EXAMPLE 19.7. Consider the language of set theory $L = \{\in\}$.

- The L-terms are just the variables, as there are no constant or function symbols.
- The atomic L-formulae are precisely of the form $(x = y)$ or $(x \in y)$ where $x, y \in \mathrm{Var}$.
- The L-formulae are the definite properties in set theory. For instance, the **ZFC** axioms are L-formulae; in fact, since there are no free variables, they are L-sentences.

20 Truth

DEFINITION 20.1 (Satisfaction). Let $\mathcal{M}$ be an L-structure, suppose $\phi = \phi(x)$ is an L-formula where $x = (x_1, \ldots, x_n)$, and $a = (a_1, \ldots, a_n) \in M^n$. We define $\mathcal{M} \vDash \phi(a)$ inductively as follows:

(i) If ϕ is $(t_1 = t_2)$ where $t_1 = t_1(x)$ and $t_2 = t_2(x)$ are L-terms, then $\mathcal{M} \vDash \phi(a)$ if $t_1^{\mathcal{M}}(a) = t_2^{\mathcal{M}}(a)$.

(ii) If ϕ is $R(t_1, \ldots, t_k)$ where $R \in L^R$ is a k-ary relation symbol in $\mathcal{M}$ and each $t_i = t_i(x_1, \ldots, x_n)$, then $\mathcal{M} \vDash \phi(a)$ if $(t_1^{\mathcal{M}}(a), \ldots, t_k^{\mathcal{M}}(a)) \in R^{\mathcal{M}}$.

(iii) If ϕ is $\neg\psi$ where $\psi = \psi(x)$ is an L-formula, then $\mathcal{M} \vDash \phi(a)$ if $\mathcal{M} \nvDash \psi(a)$.

(iv) If ϕ is $(\psi \wedge \theta)$ where $\psi(x)$ and $\theta(x)$ are L-formulae, then $\mathcal{M} \vDash \phi(a)$ if $\mathcal{M} \vDash \psi(a)$ and $\mathcal{M} \vDash \theta(a)$.

(v) If ϕ is $(\psi \vee \theta)$ where $\psi(x)$ and $\theta(x)$ are L-formulae, then $\mathcal{M} \vDash \phi(a)$ if $\mathcal{M} \vDash \psi(a)$ or $\mathcal{M} \vDash \theta(a)$.

(vi) If ϕ is $\exists y\, \psi$ where $y \in \text{Var}$ and $\psi(x, y)$ is an L-formula, then $\mathcal{M} \vDash \phi(a)$ if there exists $b \in M$ such that $\mathcal{M} \vDash \psi(a, b)$.

(vii) If ϕ is $\forall y\, \psi$ where $y \in \text{Var}$ and $\psi(x, y)$ is an L-formula, then $\mathcal{M} \vDash \phi(a)$ if for every $b \in M$, we have $\mathcal{M} \vDash \psi(a, b)$.

If $\mathcal{M} \vDash \phi(a)$, then we say that $\mathcal{M}$ **satisfies** a, or that $\phi(a)$ **is true in** $\mathcal{M}$, or that a **realizes** $\phi(x)$ **in** $\mathcal{M}$. The set of all realizations of ϕ in $\mathcal{M}$,

$$\{a \in M^n : \mathcal{M} \vDash \phi(a)\},$$

is called the **set defined by** ϕ **in** $\mathcal{M}$, and is denoted by $\phi^{\mathcal{M}}$.

REMARK 20.2. If $n = 0$, then ϕ is an L-sentence. By convention, we have $M^0 = \{\varnothing\}$, so the only question is whether $\varnothing$ realizes $\phi(x)$ or not. In particular, ϕ is either true or false in $\mathcal{M}$.

EXAMPLE 20.3. Consider the L-formula $\phi(x) = \exists z\, (z^2 = x)$.

Let $\mathcal{R} = (\mathbb{R}, 0, 1, +, -, \times)$. Then $\mathcal{R} \vDash \phi(2)$, but $\mathcal{R} \vDash \neg\phi(-2)$. In fact, $\phi^{\mathcal{R}} = \mathbb{R}_{\geq 0}$. In particular, if σ is given by $\forall x\, \phi(x)$, then $\mathcal{R} \vDash \neg\sigma$.

Now, consider instead $\mathcal{Q} = (\mathbb{Q}, 0, 1, +, -, \times)$. We see that $\mathcal{Q} \vDash \neg\phi(2)$ this time, and we have

$$\phi^{\mathcal{Q}} = \left\{ \frac{n^2}{m^2} : n, m \in \mathbb{Z},\, m \neq 0 \right\}.$$

Finally, if $\mathcal{C} = (\mathbb{C}, 0, 1, +, -, \times)$, then in fact $\mathcal{C} \vDash \sigma$. That is, $\phi^{\mathcal{C}} = \mathbb{C}$.

EXAMPLE 20.4. Let $L = \{0, +, -\}$ and let $\mathcal{Z} = (\mathbb{Z}, 0, +, -) \subseteq (\mathbb{Q}, 0, +, -) = \mathcal{Q}$. Consider the atomic formula $\psi(x, y)$ given by $y + y = x$. For $a, b \in \mathbb{Z}$, notice that

$$\mathcal{Z} \vDash \psi(a, b) \iff a = 2b \iff \mathcal{Q} \vDash \psi(a, b).$$

That is, $\mathcal{Z}$ and $\mathcal{Q}$ agree on what the integer solutions of ψ are.

Now, let $\phi(x)$ be given by $\exists y\, \psi(x, y)$. Then $\mathcal{Q} \vDash \phi(1)$, but $\mathcal{Z} \vDash \neg\phi(1)$, so $\mathcal{Q}$ and $\mathcal{Z}$ do not agree on the integer realizations of $\phi(x)$.

21 Elementary substructures

The final example from the previous section gives some insight into how satisfaction for formulae is inherited by substructures.

PROPOSITION 21.1. Suppose that $\mathcal{M} \subseteq \mathcal{N}$ are L-structures, $\phi(x)$ is an L-formula where $x = (x_1, \ldots, x_n)$, and $a = (a_1, \ldots, a_n) \in M^n$. If ϕ is quantifier-free, then $\mathcal{M} \vDash \phi(a)$ if and only if $\mathcal{N} \vDash \phi(a)$.

PROOF. In order to prove something about all formulae, one usually has to begin by proving something about terms and then proceeding by induction on the complexity of the formula. The result about terms is itself usually proved by induction on the complexity of the term.

We will begin with the following claim on terms.

Claim. For every L-term $t(x)$, we have $t^{\mathcal{N}} \restriction_{M^n} = t^{\mathcal{M}}$.

Proof of Claim. We proceed by induction on the complexity of the term t.

- Suppose that $t = x_i$ for some $1 \leq i \leq n$. Then $t^{\mathcal{N}} \restriction_{M^n}$ is the i-th coordinate projection on N^n restricted to M^n. This is simply the i-th coordinate projection of M^n, which is $t^{\mathcal{M}}$.
- Suppose that $t = c$ where $c \in L^C$. Then $t^{\mathcal{N}} \restriction_{M^n}$ is the constant function on N^n with value $c^{\mathcal{N}}$, restricted to M^n. Since $\mathcal{M} \subseteq \mathcal{N}$, we have $c^{\mathcal{M}} = c^{\mathcal{N}}$, so this is just the constant function on M^n with value $c^{\mathcal{M}}$, which is $t^{\mathcal{M}}$.
- Suppose that $t = f(t_1, \ldots, t_\ell)$ where $f \in L^F$ is ℓ-ary and $t_1, \ldots, t_\ell$ are L-terms. Then

  $$t^{\mathcal{N}}(a) = f^{\mathcal{N}}(t_1^{\mathcal{N}}(a), \ldots, t_\ell^{\mathcal{N}}(a)) = f^{\mathcal{N}}(t_1^{\mathcal{M}}(a), \ldots, t_\ell^{\mathcal{M}}(a)) = f^{\mathcal{M}}(t_1^{\mathcal{M}}(a), \ldots, t_\ell^{\mathcal{M}}(a)) = t^{\mathcal{M}}(a),$$

  where the second equality is from the inductive hypothesis, and the third equality follows from the fact that $\mathcal{M} \subseteq \mathcal{N}$ and hence $f^{\mathcal{N}} \restriction_{M^n} = f^{\mathcal{M}}$.

This completes the proof of the claim. ■

We now prove the proposition by induction on the complexity of $\phi(x)$.

- Suppose that $\phi(x)$ is atomic; that is, $\phi(x)$ is $(t = s)$ for L-terms $t(x)$ and $s(x)$. Then

  $$\begin{aligned} \mathcal{M} \vDash \phi(a) &\iff t^{\mathcal{M}}(a) = s^{\mathcal{M}}(a) \\ &\iff t^{\mathcal{N}}(a) = s^{\mathcal{N}}(a) \text{ by Claim} \\ &\iff \mathcal{N} \vDash \phi(a). \end{aligned}$$

- Suppose that $\phi(x)$ is $R(t_1, \ldots, t_k)$ where $R \in L^R$ is k-ary and $t_1(x), \ldots, t_k(x)$ are L-terms. Then

  $$\begin{aligned} \mathcal{M} \vDash \phi(a) &\iff (t_1^{\mathcal{M}}(a), \ldots, t_k^{\mathcal{M}}(a)) \in R^{\mathcal{M}} \\ &\iff (t_1^{\mathcal{M}}(a), \ldots, t_k^{\mathcal{M}}(a)) \in R^{\mathcal{N}} \text{ since } \mathcal{M} \subseteq \mathcal{N} \text{ implies } R^{\mathcal{M}} = R^{\mathcal{N}} \cap M^k \\ &\iff (t_1^{\mathcal{N}}(a), \ldots, t_k^{\mathcal{N}}(a)) \in R^{\mathcal{N}} \text{ by Claim} \\ &\iff \mathcal{N} \vDash \phi(a). \end{aligned}$$

We have proved the base cases. Let $\psi(x)$ and $\theta(x)$ be quantifier-free formulae for which the result is known.

- Suppose $\phi(x)$ is $\neg\psi(x)$. Then

  $$\begin{aligned} \mathcal{M} \vDash \phi(a) &\iff \mathcal{M} \nvDash \psi(a) \\ &\iff \mathcal{N} \nvDash \psi(a) \text{ by inductive hypothesis} \\ &\iff \mathcal{N} \vDash \phi(a). \end{aligned}$$

- Suppose $\phi(x)$ is $(\psi(x) \wedge \theta(x))$. Then

$$\begin{aligned} \mathcal{M} \vDash \phi(a) &\iff \mathcal{M} \vDash \psi(a) \text{ and } \mathcal{M} \vDash \theta(a) \\ &\iff \mathcal{N} \vDash \psi(a) \text{ and } \mathcal{N} \vDash \theta(a) \text{ by inductive hypothesis} \\ &\iff \mathcal{N} \vDash \phi(a). \end{aligned}$$

- Suppose $\phi(x)$ is $(\psi(x) \vee \theta(x))$. We can prove this the exact same way as above, or we may simply note that $(\psi \vee \theta)$ is equivalent to $\neg(\neg\psi \wedge \neg\theta)$.

As $\phi(x)$ is quantifier-free, this completes the induction. □

PROPOSITION 21.2. Suppose that $\mathcal{M} \subseteq \mathcal{N}$ are L-structures, $\phi(x)$ is an L-formula where $x = (x_1, \ldots, x_n)$, and $a = (a_1, \ldots, a_n) \in M^n$.

(a) Suppose that $\phi(x)$ is existential (i.e. $\phi(x)$ is of the form $\exists y\, \psi(x, y)$ where $y = (y_1, \ldots, y_m)$ and ψ is quantifier free). If $\mathcal{M} \vDash \phi(a)$, then $\mathcal{N} \vDash \phi(a)$.

(b) Suppose that $\phi(x)$ is universal (i.e. $\phi(x)$ is of the form $\forall y\, \psi(x, y)$ where $y = (y_1, \ldots, y_m)$ and ψ is quantifier free). If $\mathcal{N} \vDash \phi(a)$, then $\mathcal{M} \vDash \phi(a)$.

PROOF. To prove (a), suppose that $\phi(x)$ is existential. Then

$$\begin{aligned} \mathcal{M} \vDash \phi(a) &\iff \text{There exists } b \in M^m \text{ such that } \mathcal{M} \vDash \psi(a, b) \\ &\iff \text{There exists } b \in M^m \text{ such that } \mathcal{N} \vDash \psi(a, b) \text{ by Proposition 21.1} \\ &\implies \text{There exists } b \in N^m \text{ such that } \mathcal{N} \vDash \psi(a, b) \text{ since } M \subseteq N \\ &\iff \mathcal{N} \vDash \phi(a). \end{aligned}$$

To prove (b), we use the fact that $\forall y$ is equivalent to $\neg\exists y\, \neg$. We have

$$\begin{aligned} \mathcal{N} \vDash \phi(a) &\iff \mathcal{N} \vDash \forall y\, \psi(a, y) \\ &\iff \mathcal{N} \vDash \neg\exists y\, \neg\psi(a, y) \\ &\iff \mathcal{N} \nvDash \exists y\, \neg\psi(a, y) \\ &\implies \mathcal{M} \nvDash \exists y\, \neg\psi(a, y) \text{ by the contrapositive of (a)} \\ &\iff \mathcal{M} \vDash \neg\exists y\, \exists\psi(a, y) \\ &\iff \mathcal{M} \vDash \phi(a). \end{aligned}$$

This completes the proof of the proposition. □

While Propositions 21.1 and 21.2 are about substructures $\mathcal{M} \subseteq \mathcal{N}$, the proofs clearly extend to L-embeddings $j : \mathcal{M} \to \mathcal{N}$. For instance, if ϕ is quantifier free and $a \in M^n$, we have $\mathcal{M} \vDash \phi(a)$ if and only if $\mathcal{N} \vDash \phi(j(a))$. We can strengthen the notion of an embedding to force Proposition 21.1 to hold for any formula, and not just quantifier free ones.

DEFINITION 21.3 (Elementary embedding, elementary substructure). Suppose that $\mathcal{M}$ and $\mathcal{N}$ are L-structures with universes M and N respectively. An L-embedding $j : \mathcal{M} \to \mathcal{N}$ is called an **elementary embedding** if for all L-formulae $\phi(x)$ where $x = (x_1, \ldots, x_n)$ and all $a \in M^n$, we have $\mathcal{M} \vDash \phi(a)$ if and only if $\mathcal{N} \vDash \phi(j(a))$.

If $M \subseteq N$ and the containment map $j : \mathcal{M} \to \mathcal{N}$ is an elementary embedding, then we say that $\mathcal{M}$ is an **elementary substructure** of $\mathcal{N}$, denoted by $\mathcal{M} \preceq \mathcal{N}$.

EXAMPLE 21.4. Let $\mathcal{Q} = (\mathbb{Q}, 0, +, -)$.

(a) From Example 20.4, we have that $\mathcal{Z} = (\mathbb{Z}, 0, +, -) \not\preceq \mathcal{Q}$. In particular, recall that we had the sentence $\phi(x) := \exists y\, (y + y = x)$, and we saw that $\mathcal{Q} \vDash \phi(1)$ but $\mathcal{Z} \vDash \neg\phi(1)$.

(b) In fact, $\mathcal{Q}$ has no proper elementary subgroups. Indeed, suppose that $\mathcal{G} = (G, 0, +, -) \preceq \mathcal{Q}$. We make the remark that in the definition of an elementary structure, we can take $n = 0$, so elementary substructures satisfy all the same L-sentences as the extension. Note that $\mathcal{Q} \vDash \exists x\,(x \neq 0)$. Now, since $\mathcal{G} \preceq \mathcal{Q}$, we obtain $\mathcal{G} \vDash \exists x\,(x \neq 0)$. That is, we have $a \in G$ where $a \neq 0$. Since $G \subseteq \mathbb{Q}$, we see that $a = \frac{n}{m}$ for some $n, m \in \mathbb{Z}$. We may assume without loss of generality that $n, m > 0$ as we may simply take the additive inverse if $a < 0$. As $a \in G$, we also have $n \in G$, since we may write

$$n = \underbrace{a + \cdots + a}_{m \text{ times}}.$$

Observe that we have

$$\mathcal{Q} \vDash \forall x \exists y\,(\underbrace{y + \cdots + y}_{n \text{ times}} = x).$$

That is, $\mathbb{Q}$ is n-divisible. It then follows that

$$\mathcal{Q} \vDash \exists y\,(\underbrace{y + \cdots + y}_{n \text{ times}} = n).$$

Since $\mathcal{G} \preceq \mathcal{Q}$, we obtain

$$\mathcal{G} \vDash \exists y\,(\underbrace{y + \cdots + y}_{n \text{ times}} = n).$$

Hence, there is $\ell \in G$ such that

$$\underbrace{\ell + \cdots + \ell}_{n \text{ times}} = n.$$

Necessarily, $\ell = 1$, so $1 \in G$. From this, we obtain $\mathbb{Z} \leq G \leq \mathbb{Q}$. Finally, for every $0 < k \in \mathbb{Z}$, let σ_k be the L-sentence

$$\sigma_k := \forall x \exists y\,(\underbrace{y + \cdots + y}_{k \text{ times}} = x).$$

Note that $\mathcal{Q} \vDash \sigma_k$ for all $k > 0$, and from $\mathcal{G} \preceq \mathcal{Q}$, we have that $\mathcal{G} \vDash \sigma_k$ for all $k > 0$ as well. Thus, $G = \mathbb{Q}$, which we wanted to show.

22 Tarski-Vaught

PROPOSITION 22.1. Isomorphisms are elementary embeddings.

PROOF. Suppose that $j : \mathcal{M} \to \mathcal{N}$ is an L-isomorphism of L-structures. We need to show that for all L-formulae $\phi(x_1, \ldots, x_n)$ and $a \in M^n$, we have

$$\mathcal{M} \vDash \phi(a) \iff \mathcal{N} \vDash \phi(j(a)). \tag{$\star$}$$

We prove this by induction on the complexity of ϕ. If ϕ is atomic, it is quantifier-free, so $(\star)$ holds for any L-embedding; in particular, it holds for j. The $\wedge, \vee, \neg$ cases are easy to check by induction. We may write $\forall$ has $\neg\exists\neg$, so it suffices to check the case where ϕ is $\exists y\, \psi(x_1, \ldots, x_n, y)$ where $(\star)$ holds for ψ. Indeed, we have

$$\begin{aligned}
\mathcal{M} \vDash \phi(a) &\iff \text{there is } b \in M \text{ such that } \mathcal{M} \vDash \psi(a, b) \\
&\iff \text{there is } b \in M \text{ such that } \mathcal{N} \vDash \psi(j(a), j(b)) && \text{by inductive hypothesis} \\
&\iff \text{there is } c \in N \text{ such that } \mathcal{N} \vDash \psi(j(a), c) && \text{by surjectivity of } j \\
&\iff \mathcal{N} \vDash \exists y\, \psi(j(a), y) \\
&\iff \mathcal{N} \vDash \phi(j(a)).
\end{aligned}$$

This completes the induction. □

We saw previously that being an elementary substructure is a very strong condition. Recall that the equation $y + y = 1$ has a solution in $\mathbb{Q}$ but not in $\mathbb{Z}$, and so $(\mathbb{Z}, 0, +, -) \not\preceq (\mathbb{Q}, 0, +, -)$ as in Example 21.4. The following proposition tells us that this is a typical way that substructures can fail to be elementary.

PROPOSITION 22.2 (Tarski-Vaught Test). Suppose that $\mathcal{M} \subseteq \mathcal{N}$ are L-structures. The following are equivalent:

(i) $\mathcal{M}$ is an elementary substructure of $\mathcal{N}$.

(ii) For every L-formula $\phi(x_1, \ldots, x_n, y)$ and $a \in M^n$,

$$\text{if } \mathcal{N} \vDash \exists y\, \phi(a, y), \text{ then there is } b \in M \text{ such that } \mathcal{N} \vDash \phi(a, b). \tag{$\dagger$}$$

PROOF. First, we prove that (i) implies (ii). Suppose that $\mathcal{M} \preceq \mathcal{N}$. Let $\psi(x) := \exists y\, \phi(x_1, \ldots, x_n, y)$ and $a \in M^n$. If $\mathcal{N} \vDash \exists y\, \phi(a, y)$, then $\mathcal{N} \vDash \psi(a)$. As $\mathcal{M} \preceq \mathcal{N}$, we obtain $\mathcal{M} \vDash \psi(a)$. In particular, $\mathcal{M} \vDash \exists y\, \phi(a, y)$, so there exists $b \in M$ such that $\mathcal{M} \vDash \phi(a, b)$. Using the fact that $\mathcal{M} \preceq \mathcal{N}$ again, it follows that $\mathcal{N} \vDash \phi(a, b)$, so $(\dagger)$ holds.

Conversely, assume (ii). We show that for every L-formula $\theta(x_1, \ldots, x_n)$ and $a \in M^n$, we have

$$\mathcal{M} \vDash \theta(a) \iff \mathcal{N} \vDash \theta(a). \tag{$\star$}$$

As usual, we proceed by induction on the complexity of θ. If θ is atomic, $(\star)$ follows from the fact that $\mathcal{M} \subseteq \mathcal{N}$ and θ is quantifier-free. The $\neg, \vee, \wedge$ cases follow easily from induction. Again, $\forall$ can be written as $\neg\exists\neg$, so we only need to prove that $(\star)$ holds when θ is $\exists y\, \phi(x_1, \ldots, x_n, y)$, where $(\star)$ holds for ϕ. Indeed, suppose $\mathcal{M} \vDash \theta(a)$. Then $\mathcal{M} \vDash \phi(a, b)$ for some $b \in M$, and by the inductive hypothesis, we obtain $\mathcal{N} \vDash \phi(a, b)$. Hence, $\mathcal{N} \vDash \theta(a)$. On the other hand, suppose $\mathcal{N} \vDash \theta(a)$. Then $\mathcal{N} \vDash \exists y\, \phi(a, y)$, and by $(\dagger)$, there is $b \in M$ such that $\mathcal{N} \vDash \phi(a, b)$. By the inductive hypothesis, $\mathcal{M} \vDash \phi(a, b)$, and hence $\mathcal{M} \vDash \theta(a)$. Thus, $(\star)$ holds for θ, and we are done. □

EXERCISE 22.3. Suppose that M is the universe of an L-structure $\mathcal{M}$, and $A \subseteq M$. Then A is the universe of an elementary substructure of $\mathcal{M}$ if and only if for every L-formula $\phi(x_1, \ldots, x_n, y)$ and $a \in A^n$, if $\mathcal{M} \vDash \exists y\, \phi(a, y)$, then there exists $b \in A$ such that $\mathcal{M} \vDash \phi(a, b)$. (Hint: The left-to-right implication is easy. For the right-to-left direction, show that A is the universe of a substructure of $\mathcal{M}$, and apply Tarski-Vaught.)

The following is one of the first theorems in model theory. We can prove it as a consequence of Exercise 22.3.

THEOREM 22.4 (Downward Löwenheim-Skolem). Suppose that $\mathcal{M}$ is an L-structure and $A \subseteq M$. Then there exists an elementary substructure of $\mathcal{M}$ that contains A and is of cardinality at most $|A| + |L| + \aleph_0$. In particular, if $A = \varnothing$ and L is countable, then every L-structure has a countable elementary substructure.

PROOF. Let $\kappa := |A| + |L| + \aleph_0$. We construct a countable chain

$$A = A_0 \subseteq A_1 \subseteq A_2 \subseteq \cdots \subseteq M$$

such that for any L-formula $\phi(x_1, \ldots, x_n, y)$ and $a \in A_i^n$, if $\mathcal{M} \vDash \exists y\, \phi(a, y)$, then there is $b \in A_{i+1}$ such that $\mathcal{M} \vDash \phi(a, b)$.

Note that every L-formula is a finite string from a set of symbols of size $|L| + \aleph_0 \leq \kappa$, and there are at most κ-many such finite strings since κ is infinite (by Theorem 15.5 and Theorem 15.10 part (b)). Similarly, there are at most κ-many finite tuples from each A_i since $|A_i| \leq \kappa$. Thus, the number of pairs (ϕ, a) is at most κ.

For any pair $\phi(x_1, \ldots, x_n, y)$ and $a \in A_i^n$ with $\mathcal{M} \vDash \exists y\, \phi(a, y)$, we choose a realization and include it in A_{i+1}. Then we have $A_i \subseteq A_{i+1}$, $|A_{i+1}| \leq \kappa$, and A_{i+1} satisfies the desired property.

Finally, let $B := \bigcup_i A_i$, and note that we have $|B| \leq \kappa$ as well. By Exercise 22.3, it follows that B is the universe of an elementary substructure of $\mathcal{M}$, which completes the proof. □

23 Definable sets and parameters

Consider $\mathcal{R} = (\mathbb{R}, 0, 1, +, \times, <)$. If $0 \leq n < m$ are integers, then we can define the open interval (n, m) in $\mathcal{R}$ with the formula

$$\phi(x) := (\underbrace{1 + \cdots + 1}_{n \text{ times}} < x) \wedge (x < \underbrace{1 + \cdots + 1}_{m \text{ times}}).$$

Indeed, we have $\phi^{\mathcal{R}} = (n, m)$. We may also define (q, r) for $q, r \in \mathbb{Q}$. For instance, $(0, \frac{1}{2})$ may be defined with the formula

$$(0 < x) \wedge \exists y\, ((y + y = 1) \wedge (x < y)).$$

However, we cannot define an interval such as $(0, \pi)$. We need to change the language to allow parameters from the universe. This motivates the following definition.

Definition 23.1. Suppose that $\mathcal{M}$ is an L-structure and $B \subseteq M$. We write

$$L_B := L \cup \{\underline{b} : b \in B\},$$

where each $\underline{b}$ is a new, distinct constant symbol. We define $\mathcal{M}_B$ to be the L_B-structure whose universe is M, the symbols from L are interpreted exactly as in $\mathcal{M}$, and for each $b \in B$,

$$\underline{b}^{\mathcal{M}_B} := b.$$

We see that $\mathcal{M}_B$ is the natural way of making $\mathcal{M}$ into an L_B-structure.

We often drop the underline and rely on context to distinguish between $b \in B$ and $b \in L_B^C$. Also, as the construction of $\mathcal{M}_B$ from $\mathcal{M}$ is quite natural, we sometimes drop the subscript and rely on context to determine if $\mathcal{M}$ is being viewed as an L-structure or an L_B-structure.

Remark 23.2. We may now rephrase Tarski-Vaught as follows. Suppose $\mathcal{M} \subseteq \mathcal{N}$. Then $\mathcal{M} \preceq \mathcal{N}$ if and only if for every L_M-formula $\phi(y)$ in a single variable, if $\mathcal{N} \vDash \exists y\, \phi(y)$, then there is $b \in M$ such that $\mathcal{N} \vDash \phi(b)$.

Definition 23.3 (Definable sets). Let $\mathcal{M}$ be an L-structure and $B \subseteq M$. A subset $X \subseteq M^n$ is **definable over** B (or **B-definable**) **in** $\mathcal{M}$ if there exists an L_B-formula $\phi(x_1, \ldots, x_n)$ such that

$$X = \{a \in M^n : \mathcal{M}_B \vDash \phi(a)\}.$$

In this case, we write $X = \phi^{\mathcal{M}}$ and say that ϕ **defines** X. We say that X is **definable** in $\mathcal{M}$ if it is M-definable, and that X is **0-definable** if it is definable over $\varnothing$.

Remark 23.4. Observe that if $\phi(x_1, \ldots, x_n)$ is an L_B-formula, then there exists an L-formula $\psi(x_1, \ldots, x_n, y_1, \ldots, y_m)$ and a tuple $b \in B^m$ such that $\phi(x_1, \ldots, x_n) = \psi(x_1, \ldots, x_n, b)$. Thus, $X \subseteq M^n$ is B-definable if and only if $X = \{a \in M^n : \mathcal{M} \vDash \psi(a, b)\}$ for some L-formula $\psi(x_1, \ldots, x_n, y_1, \ldots, y_m)$ and $b \in B^m$.

Before we state the following proposition, we will introduce the notation $\mathrm{Aut}_B(\mathcal{M})$ to mean the set of L-automorphisms of $\mathcal{M}$ that fix B pointwise. Note that this is also equal to the set of L_B-automorphisms of $\mathcal{M}_B$.

Proposition 23.5. Suppose $\mathcal{M}$ is an L-structure, $B \subseteq M$, and $X \subseteq M^n$. If X is B-definable, then for all automorphisms $j \in \mathrm{Aut}_B(\mathcal{M})$, we have $j(X) = X$.

Proof. If X is B-definable, then $X = \{a \in M^n : \mathcal{M} \vDash \phi(a, b)\}$ where $\phi(x_1, \ldots, x_n, y_1, \ldots, y_m)$ is an L-formula and $b = (b_1, \ldots, b_m) \in B^m$. For any $a \in M^n$, we see that

$$\begin{aligned} a \in X &\iff \mathcal{M} \vDash \phi(a, b) \\ &\iff \mathcal{M} \vDash \phi(j(a), j(b)) && \text{by Proposition 22.1} \\ &\iff \mathcal{M} \vDash \phi(j(a), b) && \text{since } j(b) = b \\ &\iff j(a) \in X. \end{aligned}$$

Thus, $j(X) = X$ as desired. □

Due to this proposition, with some knowledge of automorphisms of a structure, we may produce non-definable sets.

EXAMPLE 23.6. The interval $(0,1)$ is not 0-definable in $(\mathbb{R},<)$.

PROOF. Suppose that $(0,1)$ were 0-definable in $(\mathbb{R},<)$. Consider the map $j:\mathbb{R}\to\mathbb{R}:x\mapsto x+1$. This is indeed an automorphism of $(\mathbb{R},<)$, but $j((0,1))=(1,2)\neq(0,1)$, which contradicts Proposition 23.5. □

The above example was fairly trivial, so we will give a slightly more complicated one.

Let $f:X\to Y$ be a function where $X\subseteq M^n$ and $Y\subseteq M^m$. We say that f is **definable** in $\mathcal{M}$ if $\Gamma(f)=\{(a,f(a)):a\in X\}\subseteq M^{n+m}$, the graph of f, is definable.

EXAMPLE 23.7. Addition of real numbers is not definable in $(\mathbb{R},<)$.

PROOF. Suppose that $\Gamma(+)\subseteq\mathbb{R}^3$ is definable over $b_1,\dots,b_m\in\mathbb{R}$. We may assume without loss of generality that $b_1<b_2<\cdots<b_m$. Fix $c\in\mathbb{R}$ such that $c>b_m$. Now, define $j:\mathbb{R}\to\mathbb{R}$ by

$$j(x)=\begin{cases}x & x\leq c\\ x-\frac{x-c}{2} & x>c.\end{cases}$$

In particular, if $x>c$, then j maps x halfway to c. We see that j is an automorphism fixing $b_1,\dots,b_m$ pointwise. However, we have

$$j(\underbrace{c+1,c+1,2c+2}_{\in\Gamma(+)})=(\underbrace{c+\tfrac{1}{2},c+\tfrac{1}{2},\tfrac{3}{2}c+1}_{\notin\Gamma(+)}),$$

contradicting Proposition 23.5, so $+$ is not definable. □

However, this method relies on the existence of many automorphisms. For example, it was proved in Homework 6 that the only automorphism of $(\mathbb{R},0,1,+,-,\times,<)$ is the identity. Hence, to prove that the integers are not definable in this structure (which they aren't), we need to thoroughly understand what the definable sets look like.

24 Algebraic and semi-algebraic sets

Let $L = \{0, 1, +, -, \times\}$ be the language of rings and $\mathcal{R} = (R, 0, 1, +, -, \times)$, where R is commutative and unitary. We would like to know what the definable sets of $\mathcal{R}$ are.

Suppose that $P_1, \ldots, P_\ell \in R[X_1, \ldots, X_n]$. Then the zero set

$$V(P_1, \ldots, P_\ell) := \{a \in R^n : P_i(a) = 0 \text{ for all } 1 \leq i \leq n\}$$

is called an **algebraic** or **Zariski closed** subset of R^n. Note that these are all quantifier-free definable in $\mathcal{R}$. Indeed, we may consider the L_R-formula $\phi(x_1, \ldots, x_n)$ given by

$$\bigwedge_{i=1}^{\ell} (P_i(x_1, \ldots, x_n) = 0).$$

In fact, these, and their finite boolean combinations (intersection, union, complement) are the *only* quantifier-free definable sets in $\mathcal{R}$.

The atomic L_R-formulae are of the form $(t = s)$ where $t(x_1, \ldots, x_n)$ and $s(x_1, \ldots, x_n)$ are L_R-terms. In a similar fashion to Question 4 of Homework 6, we can show that the L_R-terms agree with the polynomials over R. Hence, the atomically definable sets in $\mathcal{R}$ are precisely the hypersurfaces: namely of the form $V(P)$ where $P \in R[X_1, \ldots, X_n]$. From this, we can see that the quantifier-free definable subsets of R^n are precisely the sets of the form

$$V_1 \setminus W_1 \cup \cdots \cup V_k \setminus W_k,$$

where $W_i \subseteq V_i \subseteq R^n$ are algebraic sets.

Exercise 24.1. Using disjunctive normal form (DNF), every quantifier-free formula is logically equivalent to one of the form

$$\bigvee_{i=1}^{r} \left(\bigwedge_{j=1}^{s_i} \phi_{i,j} \wedge \bigwedge_{j=1}^{t_i} \neg \psi_{i,j} \right)$$

where $\phi_{i,j}, \psi_{i,j}$ are atomic.

Such sets are called the **Zariski-constructible** sets. With this exercise, we have now proven that in any ring, the quantifier-free definable sets are the Zariski-constructible sets.

Fact 24.2 (Tarski). Suppose that R is an algebraically closed field. Then every definable set in $\mathcal{R} = (R, 0, 1, +, -, \times)$ is Zariski-constructible. That is, every definable set is quantifier-free definable, so $\mathcal{R}$ has quantifier elimination. (We prove this later.)

Corollary 24.3. If R is an algebraically closed field, then every definable subset of R in $\mathcal{R} = (R, 0, 1, +, -, \times)$ is either finite or cofinite (the complement is finite).

Consider $R = \mathbb{R}$, and $\mathcal{R} = (\mathbb{R}, 0, 1, +, -, \times)$. Let

$$\phi(x) := \exists y \, (y^2 = x).$$

Then $\phi^{\mathcal{R}} = \mathbb{R}_{\geq 0}$. This is neither finite nor cofinite, and hence is not quantifier-free definable in $\mathcal{R}$. Hence, $\mathcal{R}$ does not have quantifier elimination.

Fact 24.4 (Macintyre). If $\mathcal{R}$ is a ring that admits quantifier elimination, then it is an algebraically closed field. (We will not prove this; this is outside the scope of the course.)

Equipped with this fact, what can we say about $\mathbb{R}$? Note that $<$ on $\mathbb{R}$ is definable in $(\mathbb{R}, 0, 1, +, -, \times)$, since to encode $x < y$, we may consider the formula

$$\exists z \, (z^2 = y - x) \wedge \neg(x = y).$$

FACT 24.5 (Tarski). Every definable set in $(\mathbb{R}, 0, +, -, \times, <)$ is quantifier-free definable. (We prove this later.)

This says that if we include $<$ in our language, then every definable set is quantifier-free definable. In particular, $<$ was the only thing that we needed quantifiers for.

EXERCISE 24.6. The quantifier-free definable sets in an ordered ring $(R, 0, 1, +, -, \times, <)$ (i.e. $a < b$ implies $a + c < b + c$, and if $a < b$ with $c > 0$, then $ac < bc$) are finite boolean combinations of sets of the form $P(x_1, \ldots, x_n) = 0$, or $P(x_1, \ldots, x_n) > 0$, where $P \in R[X_1, \ldots, X_n]$.

The sets from Exercise 24.6 are called **semi-algebraic**.

COROLLARY 24.7. The definable sets in $(\mathbb{R}, 0, 1, +, -, \times)$ are precisely the semi-algebraic sets.

Note that these examples were very tame, in which we could easily describe the definable sets. This is not always the case. For instance, consider the ring of integers. There is no easy way of describing the definable sets other than repeating the definition; in fact, Gödel proved that all of mathematics could be encoded using the ring of integers, so this should make sense.

25 Theories and models

Previously, we looked at structures and analyzed their definable sets (via formulas and interpretations). On the other hand, formulas (or rather, sentences) can be used to axiomatise classes of structures. We will later see that these two topics are intimately related and complementary.

DEFINITION 25.1. Let L be a language.

- An **L-theory** is a set of L-sentences.
- Let T is an L-theory. A **model** of T is an L-structure $\mathcal{M}$ such that for every L-sentence $\sigma \in T$, we have $\mathcal{M} \vDash \sigma$. We denote this by $\mathcal{M} \vDash T$.
- An L-theory is said to be **consistent** if it has a model.
- A class $\mathcal{K}$ of L-structures is said to be **elementary** or **axiomatisable** if there exists an L-theory T such that $\mathcal{M} \in \mathcal{K}$ if and only if $\mathcal{M} \vDash T$.

EXAMPLE 25.2. Let $L = \{e, \cdot, ^{-1}\}$ be the language of groups (we do not use additive notation as this is typically reserved for abelian groups). The following are all axiomatisable.

- **Groups:** Each of the group axioms can be written as an L-sentence. In fact, these are finitely axiomatisable.
- **Abelian groups:** Take the group axioms, and add an L-sentence that says that the elements commute. These are also finitely axiomatisable.
- **Groups of a fixed exponent n:** We may add the L-sentence $\forall x\,(x^n = e)$, and these are also finitely axiomatisable.
- **Torsion-free groups:** For each $n \in \mathbb{N}$, we may consider the L-sentence
$$\sigma_n := \forall x\,((x^n = e) \to (x = e)).$$
These are infinitely axiomatisable, and we will prove that they are not finitely axiomatisable.
- **Divisible groups:** In particular, groups in which for every $n \in \mathbb{N}$, every element has an n-th root. Indeed, we may add the L-sentence
$$\tau_n := \forall x \exists y\,(y^n = x)$$
for all $n \in \mathbb{N}$. These are infinitely axiomatisable, but not finitely axiomatisable.

We will show later that the following are not axiomatisable.

- **Torsion groups:** Note that we can only quantify over the universe of the structure, and not $\mathbb{N}$, so $\forall x \exists n\,(x^n = e)$ is not an L-sentence. We also see that
$$\forall x\,((x = e) \vee (x^2 = e) \vee \cdots)$$
is not an L-sentence as it is not of finite length.
- **Finite groups:** It is possible to write an L-sentence expressing that there are exactly n elements in the structure, so the collection of finite groups of a particular size is axiomatisable. However, we cannot axiomatise the collection of all finite groups.

DEFINITION 25.3. Let $\mathcal{M}$ be an L-structure. The **theory of $\mathcal{M}$** is the L-theory given by
$$\mathrm{Th}(\mathcal{M}) := \{\sigma : \mathcal{M} \vDash \sigma\}.$$

If $\mathcal{M}$ and $\mathcal{N}$ are L-structures, then $\mathcal{M}$ is **elementarily equivalent** to $\mathcal{N}$ if $\mathrm{Th}(\mathcal{M}) = \mathrm{Th}(\mathcal{N})$. That is, for every L-sentence σ, $\mathcal{M} \vDash \sigma$ if and only if $\mathcal{N} \vDash \sigma$. We write $\mathcal{M} \equiv \mathcal{N}$.

REMARK 25.4. To prove that two L-structures are elementarily equivalent, one only needs to show one direction, as the other follows from the fact that L-sentences are closed under negation.

REMARK 25.5. Let $\mathcal{M}$ and $\mathcal{N}$ be L-structures.

(a) If $j : \mathcal{M} \to \mathcal{N}$ is an elementary embedding, then $\mathcal{M}$ is elementarily equivalent to $\mathcal{N}$ by taking $n = 0$ in the definition of an elementary embedding. Hence, if $\mathcal{M} \preceq \mathcal{N}$, then $\mathcal{M} \equiv \mathcal{N}$.

(b) Note that $\mathcal{M} \subseteq \mathcal{N}$ and $\mathcal{M} \equiv \mathcal{N}$ does not necessarily imply that $\mathcal{M} \preceq \mathcal{N}$. For example, consider $\mathcal{M} = (\omega \setminus \{0\}, <)$ and $\mathcal{N} = (\omega, <)$. We see that $\mathcal{M} \subseteq \mathcal{N}$. Moreover, the map $j : \mathcal{M} \to \mathcal{N}$ taking $n \in \omega \setminus \{0\}$ to its predecessor is an L-isomorphism, and hence an elementary embedding. So by part (a), $\mathcal{M} \equiv \mathcal{N}$. However, observe that $\mathcal{M} \not\preceq \mathcal{N}$ since the L_M-sentence $\exists y\,(y < 1)$ is true in $\mathcal{N}$, but not in $\mathcal{M}$.

(c) If $\mathcal{M} \subseteq \mathcal{N}$, then $\mathcal{M} \preceq \mathcal{N}$ (as L-structures) if and only if $\mathcal{M}_M \equiv \mathcal{N}_M$ (as L_M-structures).

The following proposition is a strengthening of part (c) of the previous remark.

PROPOSITION 25.6. Let $\mathcal{M}$ and $\mathcal{N}$ be L-structures. There exists an elementary embedding $j : \mathcal{M} \to \mathcal{N}$ if and only if $\mathcal{N}$ can be expanded to be a model of the L_M-theory $\mathrm{Th}(\mathcal{M}_M)$.

PROOF. Suppose that $j : \mathcal{M} \to \mathcal{N}$ is an elementary embedding. We expand $\mathcal{N}$ into an L_M-structure $\mathcal{N}'$ by setting $\underline{a}^{\mathcal{N}'} := j(a)$ for every $a \in M$, interpreting all symbols the same as $\mathcal{N}$, and keeping the universe the same. We leave it as an exercise to check that $\mathcal{N}' \vDash \mathrm{Th}(\mathcal{M}_M)$ using the fact that j is an elementary embedding.

Conversely, if $\mathcal{N}'$ is an expansion of $\mathcal{N}$ such that $\mathcal{N}' \vDash \mathrm{Th}(\mathcal{M}_M)$, then we may define the map $j : \mathcal{M} \to \mathcal{N}$ by $j(a) := \underline{a}^{\mathcal{N}'}$ for all $a \in M$. Verify that j is an elementary embedding. □

26 Entailment

Definition 26.1. Let T be an L-theory and σ be an L-sentence. We say that T **implies** (or **entails**) σ, denoted by $T \vDash \sigma$, if for every model $\mathcal{M} \vDash T$, we have $\mathcal{M} \vDash \sigma$. (We also say that σ is a **consequence** of T.) We say that T is **complete** if for every L-sentence σ, either $T \vDash \sigma$ or $T \vDash \neg\sigma$.

Example 26.2.

(a) Let $\mathcal{M}$ be an L-structure, and let $T = \mathrm{Th}(\mathcal{M})$. For every L-sentence σ, either $\mathcal{M} \vDash \sigma$ or $\mathcal{M} \vDash \neg\sigma$, so it follows that T is complete.

(b) Let $L = \{0, 1, +, -, \times\}$ be the language of rings. Let T_1 be the theory of rings. Note that T_1 is incomplete, as we may consider the sentence $\sigma := \forall x \forall y\,(xy = yx)$. Then $T_1 \nvDash \sigma$ as non-commutative rings exist, but $T_1 \nvDash \neg\sigma$ as well since commutative rings exist.

Let T_2 be the theory of fields. This is also incomplete; we may consider the sentence $\tau := \forall x \exists y\,(y^2 = x)$, which is true in $\mathbb{C}$ but false in $\mathbb{R}$.

Let T_3 be the theory of algebraically closed fields. This is incomplete as well, since the sentence $(1 + 1 = 0)$ is true in $\mathbb{Z}/2\mathbb{Z}$ but false in $\mathbb{C}$.

Finally, ACF_p, the theory of algebraically closed fields of characteristic p, where p is either prime or zero, is complete. (We prove this fact later.)

Lemma 26.3. Let T be a consistent L-theory, and let $\overline{T} = \{\sigma : T \vDash \sigma\}$. The following are equivalent.

(i) T is complete.

(ii) $\overline{T}$ is maximally consistent.

(iii) $\overline{T} = \mathrm{Th}(\mathcal{M})$ for some (or equivalently, any) $\mathcal{M} \vDash T$.

(iv) Any two models of T are elementarily equivalent.

Proof. (i) $\Rightarrow$ (ii). Suppose that T is complete. Note that $\overline{T}$ is consistent, since any model of T is a model of $\overline{T}$. Now, suppose that $S \supsetneq \overline{T}$. Let $\sigma \in S \setminus \overline{T}$. Then $T \nvDash \sigma$, and since T is complete, $T \vDash \neg\sigma$. Hence, $\neg\sigma \in \overline{T} \subseteq S$. As $\sigma, \neg\sigma \in S$, it follows that S has no models.

(ii) $\Rightarrow$ (iii). Suppose that $\overline{T}$ is maximally consistent. Let $\mathcal{M} \vDash T$ be arbitrary. Then $\mathcal{M} \vDash \overline{T}$, and hence $\mathrm{Th}(\mathcal{M}) \supseteq \overline{T}$. Now, $\mathrm{Th}(\mathcal{M})$ is consistent as it has $\mathcal{M}$ as a model, and $\overline{T}$ is maximally consistent, so $\mathrm{Th}(\mathcal{M}) = \overline{T}$. (This proves the stronger version of (iii) where $\overline{T} = \mathrm{Th}(\mathcal{M})$ for *all* models $\mathcal{M} \vDash T$.)

(iii) $\Rightarrow$ (iv) Suppose that $\overline{T} = \mathrm{Th}(\mathcal{M})$ for some $\mathcal{M} \vDash T$. (Here, we use the weaker version of (iii).) Let $\mathcal{N} \vDash T$. Then $\mathrm{Th}(\mathcal{N}) \supseteq \overline{T} = \mathrm{Th}(\mathcal{M})$. Since $\mathrm{Th}(\mathcal{M})$ is complete, it is maximally consistent (from (i) $\Rightarrow$ (ii) applied to $\mathrm{Th}(\mathcal{M})$), so $\mathrm{Th}(\mathcal{N}) = \mathrm{Th}(\mathcal{M})$. Thus, $\mathcal{N} \equiv \mathcal{M}$. Every model of T is elementarily equivalent to $\mathcal{M}$, and hence any two models of T are elementarily equivalent to each other.

(iv) $\Rightarrow$ (i). Assume that any two models of T are elementarily equivalent to each other. Suppose that σ is an L-sentence such that $T \nvDash \sigma$. Then there exists a model $\mathcal{M} \vDash T$ such that $\mathcal{M} \vDash \neg\sigma$. As any model of T is elementarily equivalent to $\mathcal{M}$, it follows that $\neg\sigma$ is true in every model of T. Thus, $T \vDash \neg\sigma$, so T is complete. □

The Compactness Theorem for first-order logic is of fundamental importance and is the starting point for model theory and its applications.

Theorem 26.4 (Compactness Theorem). Let L be a language and T be an L-theory. T is consistent if and only if every finite subset of T is consistent.

Generally, the Compactness Theorem is seen as an immediate consequence of Gödel's completeness theorem, which states that T is consistent if and only if there is no "formal derivation" of a contradiction using the sentences in T as assumptions. The fact that derivations are finite then implies the Compactness Theorem. However, this approach involves proof theory, which we would like to avoid. Instead, we will prove the Compactness Theorem using ultraproducts, which we will introduce in the next lecture.

27 Ultraproducts

Definition 27.1 (Filter). Let I be a non-empty set. A **filter** on I is a subset $\mathcal{F} \subseteq \mathcal{P}(I)$ satisfying:

(i) $I \in \mathcal{F}$ and $\varnothing \notin \mathcal{F}$.

(ii) If $A, B \in \mathcal{F}$, then $A \cap B \in \mathcal{F}$.

(iii) If $A \in \mathcal{F}$ and $A \subseteq B \subseteq I$, then $B \in \mathcal{F}$.

Filters give us a notion of "largeness" for subsets of I. We give some examples below.

Example 27.2.

(a) Let $I = \mathbb{R}$. Then $\mathcal{F} = \{A \subseteq \mathbb{R} : \mathbb{R} \setminus A \text{ has Lebesgue measure } 0\}$ is a filter on I.

(b) Let I be an infinite set, and let κ be a cardinal with $\aleph_0 \leq \kappa \leq |I|$. Then $\mathcal{F} = \{A \subseteq I : |I \setminus A| < \kappa\}$ is a filter on I. In particular, when $\kappa = \aleph_0$, then $\mathcal{F} = \{A \subseteq I : A \text{ is cofinite}\}$, and this is called the **Fréchet filter** on I.

(c) Let I be a non-empty set. A **principal filter** on I is a filter of the form $\mathcal{F} = \{A \subseteq I : x \in A\}$ for some fixed $x \in I$.

An **ultrafilter** is a maximal filter; that is, a filter that is not properly contained in any filter on I. Note that every principal filter is an ultrafilter. To see this, let $\mathcal{F} = \{A \subseteq I : x \in A\}$ be a principal filter. Then any larger filter would contain a set that does not contain x, and hence its intersection with $\{x\} \in \mathcal{F}$ would be the empty set, contradicting the fact that $\varnothing$ is not contained in any filter. It is slightly more difficult to describe the ultrafilters that are not principal filters. First, we prove some facts.

Proposition 27.3. Every filter can be extended to an ultrafilter.

Proof. Let I be a non-empty set and suppose that $\mathcal{F}$ is a filter on I. Let Λ be the set of filters on I that contain $\mathcal{F}$, and observe that Λ is partially ordered by $\subseteq$. Now, if $\mathcal{G}_1 \subseteq \mathcal{G}_2 \subseteq \cdots$ is a chain in Λ, then $\bigcup_{i \in \omega} \mathcal{G}_i \in \Lambda$. By Zorn's Lemma, we obtain a maximal element $\mathcal{U} \in \Lambda$, and $\mathcal{U} \supseteq \mathcal{F}$ is an ultrafilter, as required. □

Lemma 27.4. A filter $\mathcal{U}$ on a set I is an ultrafilter if and only if for every subset $A \subseteq I$, either $A \in \mathcal{U}$ or $I \setminus A \in \mathcal{U}$.

Proof. Suppose that $\mathcal{U}$ is an ultrafilter and let $A \subseteq I$. Suppose that $A \notin \mathcal{U}$. Then $\mathcal{F} = \{B \subseteq I : B \supseteq C \setminus A \text{ for some } C \in \mathcal{U}\}$ is a filter (check this). Now, $\mathcal{U} \subseteq \mathcal{F}$ by simply taking $B = C$ in the set above. By maximality, we have $\mathcal{U} = \mathcal{F}$, and it follows that $I \setminus A \in \mathcal{F} = \mathcal{U}$.

Conversely, assume that for every $A \subseteq I$, either $A \in \mathcal{U}$ or $I \setminus A \in \mathcal{U}$. Suppose to the contrary that there is a filter $\mathcal{F} \supsetneq \mathcal{U}$. Let $A \in \mathcal{F} \setminus \mathcal{U}$. As $A \notin \mathcal{U}$, we have $I \setminus A \in \mathcal{U}$. Now $A \cap (I \setminus A) = \varnothing \in \mathcal{F}$, a contradiction. □

With a little bit more work, we can describe the ultrafilters which are not principal filters. The following exercise can be proven from what we just showed above.

Exercise 27.5. An ultrafilter $\mathcal{U}$ is non-principal if and only if $\mathcal{U}$ extends the Fréchet filter.

Definition 27.6 (Ultraproduct). Let I be an infinite set. Let L be a language and let $(\mathcal{M}_i : i \in I)$ be a sequence of L-structures. Suppose that $\mathcal{U}$ is an ultrafilter on I. The **ultraproduct** of $(\mathcal{M}_i : i \in I)$ with respect to $\mathcal{U}$ is the L-structure

$$\mathcal{M} := \prod_{\mathcal{U}} \mathcal{M}_i$$

defined as follows:

- The universe of $\mathcal{M}$ is $M := (\prod_{i \in I} M_i)/E$ where E is the equivalence relation $(a_i : i \in I)\, E\, (b_i : i \in I)$ if $\{i \in I : a_i = b_i\} \in \mathcal{U}$, that is, the indices of the sequences agree.

- For every constant symbol $c \in L^C$, set $c^{\mathcal{M}} = [(c^{\mathcal{M}_i} : i \in I)]$, that is, the E-class of the sequence given by $c^{\mathcal{M}_i}$ for each $i \in I$.
- For every n-ary function symbol $f \in L^F$ and all $\alpha_1, \ldots, \alpha_n \in M$, define

$$f^{\mathcal{M}}(\alpha_1, \ldots, \alpha_n) := \beta,$$

  where if $\alpha_j = [(a_{ij} : i \in I)]$ for $1 \leq j \leq n$, then $\beta = [(f^{\mathcal{M}_i}(a_{i1}, \ldots, a_{in}) : i \in I)]$. Verify that this definition does not depend on the set of representatives $\alpha_1, \ldots, \alpha_n$.
- For every k-ary relation symbol $R \in L^R$ and $\alpha_1, \ldots, \alpha_k \in M$, define $R^{\mathcal{M}} \subseteq M^k$ by

$$(\alpha_1, \ldots, \alpha_k) \in R^{\mathcal{M}} \iff \{i \in I : (a_{i1}, \ldots, a_{ik}) \in R^{\mathcal{M}_i}\} \in \mathcal{U},$$

  where each $\alpha_j = [(a_{ij} : i \in I)]$ for $1 \leq j \leq k$. Similarly with function symbols, check that this definition does not depend on the set of representatives $\alpha_1, \ldots, \alpha_k$.

28 Łoś' Theorem

This lecture will be focused on proving Łoś' Theorem, of which we will explore its consequences and use it to prove the Compactness Theorem. Informally, it states that if $(\mathcal{M}_i : i \in I)$ is a sequence of L-structures, $\mathcal{U}$ an ultrafilter, and ϕ is an L-formula, then ϕ is true in the ultraproduct $\prod_{\mathcal{U}} \mathcal{M}_i$ if and only if ϕ is true in almost all $\mathcal{M}_i$.

THEOREM 28.1 (Łoś' Theorem). Let $\{\mathcal{M}_i : i \in I\}$ be a sequence of L-structures and $\mathcal{U}$ be an ultrafilter on I. Let $\mathcal{M} = \prod_{\mathcal{U}} \mathcal{M}_i$ be the ultraproduct of $\{\mathcal{M}_i : i \in I\}$ with respect to $\mathcal{U}$. Suppose that $\phi(x_1, \ldots, x_n)$ is an L-formula, and $g_1, \ldots, g_n \in \prod_{i \in I} M_i$. Then

$$\mathcal{M} \vDash \phi([g_1], \ldots, [g_n]) \iff \{i \in I : \mathcal{M}_i \vDash \phi(g_1(i), \ldots, g_n(i))\} \in \mathcal{U}.$$

In particular, if σ is an L-sentence, then

$$\mathcal{M} \vDash \sigma \iff \{i \in I : \mathcal{M}_i \vDash \sigma\} \in \mathcal{U}.$$

PROOF. We first state the following claim on terms.

CLAIM. For any L-term $t = t(x_1, \ldots, x_n)$ and $g_1, \ldots, g_n \in \prod_{i \in I} M_i$, we have

$$t^{\mathcal{M}}([g_1], \ldots, [g_n]) = [(t^{\mathcal{M}_i}(g_1(i), \ldots, g_n(i)))_{i \in I}].$$

PROOF OF CLAIM. This follows easily by induction on the complexity of the term and the definition of the ultraproduct, and as such, is left as an exercise. ■

With this claim, we may proceed by induction on the complexity of ϕ. Suppose that ϕ is atomic. If ϕ is of the form $(t_1 = t_2)$ where $t_1 = t_1(x_1, \ldots, x_n)$ and $t_2 = t_2(x_1, \ldots, x_n)$ are L-terms, then

$$\begin{aligned}
\mathcal{M} \vDash \phi([g_1], \ldots, [g_n]) &\iff t_1^{\mathcal{M}}([g_1], \ldots, [g_n]) = t_2^{\mathcal{M}}([g_1], \ldots, [g_n]) \\
&\overset{\text{Claim}}{\iff} [(t_1^{\mathcal{M}_i}(g_1(i), \ldots, g_n(i)))_{i \in I}] = [(t_2^{\mathcal{M}_i}(g_1(i), \ldots, g_n(i)))_{i \in I}] \\
&\iff \{i \in I : t_1^{\mathcal{M}_i}(g_1(i), \ldots, g_n(i)) = t_2^{\mathcal{M}_i}(g_1(i), \ldots, g_n(i))\} \in \mathcal{U} \\
&\iff \{i \in I : \mathcal{M}_i \vDash \phi(g_1(i), \ldots, g_n(i))\} \in \mathcal{U}.
\end{aligned}$$

On the other hand, if ϕ is $R(t_1, \ldots, t_\ell)$ where R is an ℓ-ary relation symbol and each $t_j = t_j(x_1, \ldots, x_n)$ is an L-term, then

$$\begin{aligned}
\mathcal{M} \vDash \phi([g_1], \ldots, [g_n]) &\iff (t_1^{\mathcal{M}}([g_1], \ldots, [g_n]), \ldots, t_\ell^{\mathcal{M}}([g_1], \ldots, [g_n])) \in R^{\mathcal{M}} \\
&\overset{\text{Claim}}{\iff} ([(t_1^{\mathcal{M}_i}(g_1(i), \ldots, g_n(i)))_{i \in I}], \ldots, [(t_\ell^{\mathcal{M}_i}(g_1(i), \ldots, g_n(i)))_{i \in I}]) \in R^{\mathcal{M}} \\
&\iff \{i \in I : (t_1^{\mathcal{M}_i}(g_1(i), \ldots, g_n(i)), \ldots, t_\ell^{\mathcal{M}_i}(g_1(i), \ldots, g_n(i))) \in R^{\mathcal{M}_i}\} \in \mathcal{U} \\
&\iff \{i \in I : \mathcal{M}_i \vDash \phi(g_1(i), \ldots, g_n(i))\} \in \mathcal{U}.
\end{aligned}$$

Now, suppose that $\phi(x_1, \ldots, x_n)$ is $\neg\psi(x_1, \ldots, x_n)$ where the result is known for ψ. We have

$$\begin{aligned}
\mathcal{M} \vDash \phi([g_1], \ldots, [g_n]) &\iff \mathcal{M} \nvDash \psi([g_1], \ldots, [g_n]) \\
&\iff \{i \in I : \mathcal{M}_i \vDash \psi(g_1(i), \ldots, g_n(i))\} \notin \mathcal{U} \text{ by IH} \\
&\iff I \setminus \{i \in I : \mathcal{M}_i \vDash \psi(g_1(i), \ldots, g_n(i))\} \in \mathcal{U} \text{ by Lemma 27.4} \\
&\iff \{i \in I : \mathcal{M}_i \nvDash \psi(g_1(i), \ldots, g_n(i))\} \in \mathcal{U} \\
&\iff \{i \in I : \mathcal{M}_i \vDash \phi(g_1(i), \ldots, g_n(i))\} \in \mathcal{U}.
\end{aligned}$$

We leave the $\wedge$ and $\vee$ cases as an exercise. In particular, the $\wedge$ case follows from the fact that ultrafilters are closed under intersections, and the $\vee$ case follows from the fact that ultrafilters are closed under unions.

As usual, the case of the universal quantifier is reduced to the existential quantifier. Suppose that $\phi(x_1, \ldots, x_n)$ is of the form $\exists y\, \psi(x_1, \ldots, x_n, y)$ where the result is known for ψ. Then $\mathcal{M} \vDash \phi([g_1], \ldots, [g_n])$ if and only if

there is some $h \in \prod_{i \in I} M_i$ such that $\mathcal{M} \vDash \psi([g_1], \ldots, [g_n], [h])$. By the inductive hypothesis, this is if and only if

$$X_h := \{i \in I : \mathcal{M}_i \vDash \psi(g_1(i), \ldots, g_n(i), h(i))\} \in \mathcal{U}.$$

Now, let $Y := \{i \in I : \mathcal{M}_i \vDash \phi(g_1(i), \ldots, g_n(i))\}$. We want that $\mathcal{M} \vDash \phi([g_1], \ldots, [g_n])$ if and only if $Y \in \mathcal{U}$. It suffices to show that there exists $h \in \prod_{i \in I} M_i$ such that $X_h \in \mathcal{U}$ if and only if $Y \in \mathcal{U}$. Indeed, for the left-to-right direction, suppose that there exists $h \in \prod_{i \in I} M_i$ such that $X_h \in \mathcal{U}$. Observe that $X_h \subseteq Y$, and since supersets are preserved under filters, it follows that $Y \in \mathcal{U}$. Conversely, suppose that $Y \in \mathcal{U}$. For each $i \in Y$, we have $\mathcal{M}_i \vDash \exists y \, \psi(g_1(i), \ldots, g_n(i), y)$. Let $a_i \in M_i$ witness this; that is, $\mathcal{M}_i \vDash \psi(g_1(i), \ldots, g_n(i), a_i)$. Now, define $h : I \to \bigcup_{i \in I} M_i$ by

$$h(i) := \begin{cases} a_i & i \in Y \\ b_i & i \notin Y, \end{cases}$$

where $b_i \in M_i$ for $i \notin Y$ is fixed arbitrarily. Note that $Y \subseteq X_h$ since if $i \in Y$, then $h(i) = a_i$, so that $\mathcal{M}_i \vDash \psi(g_1(i), \ldots, g_n(i), h(i))$, and hence $i \in X_h$. As before, supersets are closed under filters, so $X_h \in \mathcal{U}$. This completes the proof. □

29 Some consequences of Łoś' Theorem

First, we introduce a special case of the ultraproduct.

Definition 29.1 (Ultrapower). Suppose that $\mathcal{M}$ be an L-structure, I is a non-empty set, and $\mathcal{U}$ is an ultrafilter on I. The **ultrapower** of $\mathcal{M}$ with respect to $\mathcal{U}$ is the ultraproduct of $(\mathcal{M}_i : i \in I)$ where $\mathcal{M}_i = \mathcal{M}$ for all $i \in I$ with respect to $\mathcal{U}$. We write

$$\prod_{\mathcal{U}} \mathcal{M} =: \mathcal{M}^I/\mathcal{U}.$$

Now, let $\mathcal{M}$ be an L-structure. Consider the map $d : \mathcal{M} \to \mathcal{M}^I/\mathcal{U}$ which maps $a \in M$ to $[(a)_{i\in I}]$, called the **diagonal**.

Proposition 29.2. Suppose that $\mathcal{M}$ is an L-structure, I is a non-empty set, and $\mathcal{U}$ is an ultrafilter on I. Then the diagonal $d : \mathcal{M} \to \mathcal{M}^I/\mathcal{U}$ is an elementary embedding.

Proof. Let $\phi(x_1, \ldots, x_n)$ be an L-formula and suppose $a_1, \ldots, a_n \in M$. We want to show that

$$\mathcal{M} \vDash \phi(a_1, \ldots, a_n) \iff \mathcal{M}^I/\mathcal{U} \vDash \phi(d(a_1), \ldots, d(a_n)).$$

For each $a \in M$, we let $d_a \in \prod_{i\in I} M$ be the function $d_a(i) = a$ for all $i \in I$. Then the diagonal is given by $a \mapsto [d_a]$. For the left-to-right direction, we have

$$\begin{aligned}
\mathcal{M} \vDash \phi(a_1, \ldots, a_n) &\implies \mathcal{M} \vDash \phi(d_{a_1}(i), \ldots, d_{a_n}(i)) \text{ for all } i \in I \\
&\implies \{i \in I : \mathcal{M} \vDash \phi(d_{a_1}(i), \ldots, d_{a_n}(i))\} = I \in \mathcal{U} \\
&\implies \mathcal{M}^I/\mathcal{U} \vDash \phi([d_{a_1}], \ldots, [d_{a_n}]) \text{ by Łoś' Theorem.}
\end{aligned}$$

Conversely, if $\{i \in I : \mathcal{M} \vDash \phi(d_{a_1}(i), \ldots, d_{a_n}(i)\} \in \mathcal{U}$, then this set is non-empty and hence $\mathcal{M} \vDash \phi(d_{a_1}(i), \ldots, d_{a_n}(i))$ for some $i \in I$. Since each $d_{a_j}(i) = a_j$, we then obtain $\mathcal{M} \vDash \phi(a_1, \ldots, a_n)$ as desired. □

Due to this proposition, we can see that from identifying $\mathcal{M}$ with its image under the diagonal, we obtain $\mathcal{M} \preceq \mathcal{M}^I/\mathcal{U}$.

Definition 29.3 (Finite intersection property). Let X be a set and suppose that $F = (F_i : i \in I)$ is a non-empty family of subsets of X indexed by I. Then F has the **finite intersection property** if for every non-empty finite subset $J \subseteq I$, we have $\bigcap_{i\in J} F_i \neq \varnothing$.

Proposition 29.4. Suppose that $(\mathcal{M}_i : i < \omega)$ is a sequence of L-structures and $\mathcal{U}$ is a non-principal ultrafilter. Then $\mathcal{M} = \prod_{\mathcal{U}} \mathcal{M}_i$ is $\aleph_1$-compact; that is, given any countable collection $\{F_i : i < \omega\}$ of non-empty definable subsets of M^ℓ with the finite intersection property, it follows that $\bigcap_{i<\omega} F_i \neq \varnothing$.

Proof. We prove this proposition in the 0-definable case; the general case with parameters is on Assignment 10. For each $i < \omega$, suppose that F_i is defined by the L-formula $\phi_i(x)$.

We may assume that $\vDash \phi_{n+1} \to \phi_n$ (that is, this holds for any L-structure) by taking conjunctions. In particular, we may replace $F_0, F_1, F_2, \cdots$ by $F_0, F_1 \cap F_0, F_2 \cap F_1 \cap F_0, \cdots$. Moreover, we may also assume that $\phi_0 := (x = x)$ so that $F_0 = M^\ell$.

For each $i < \omega$, let

$$n_i := \max\{n \leq i : \mathcal{M}_i \vDash \exists x\, \phi_n(x)\}.$$

Note that this always exists as $n = 0$ works. Now, to prove the proposition, we want to find a sequence $(a_i)_{i<\omega}$ with $a_i \in M_i$ such that if $a = [(a_i)] \in M$, then $\mathcal{M} \vDash \phi_n(a)$ for all $n < \omega$.

Indeed, let $a_i \in M_i$ be such that $\mathcal{M}_i \vDash \phi_{n_i}(a_i)$. We will show that $a = [(a_i)]$ works. Fix some $n < \omega$. Then, consider the set

$$X_n := \{i : i \geq n \text{ and } \mathcal{M}_i \vDash \exists x\, \phi_n(x)\}.$$

Claim 1. $X_n \in \mathcal{U}$.

Proof of Claim 1. By the finite intersection property, we clearly have $F_n \neq \varnothing$. Thus, $\mathcal{M} \vDash \exists x\, \phi_n(x)$. By Łoś' Theorem, we see that $A := \{i < \omega : \mathcal{M}_i \vDash \exists x\, \phi_n(x)\} \in \mathcal{U}$. Then, since $\mathcal{U}$ is a non-principal ultrafilter, we know by Exercise 27.5 that $\mathcal{U}$ extends the Fréchet filter. Hence, $B := \{i < \omega : i \geq n\} \in \mathcal{U}$. Now, filters are closed under intersections, and so $X_n = A \cap B \in \mathcal{U}$. ■

Claim 2. $X_n \subseteq \{i < \omega : \mathcal{M}_i \vDash \phi_n(a_i)\}$.

Proof of Claim 2. Let $i \in X_n$. Then $\mathcal{M}_i \vDash \exists x\, \phi_n(x)$ and $i \geq n$. Observe that from $\mathcal{M}_i \vDash \exists x\, \phi_n(x)$, we have $n \leq n_i$. Thus, it follows that $\vDash \phi_{n_i} \to \phi_n$. By our choice of a_i, we have $\mathcal{M}_i \vDash \phi_{n_i}(a_i)$, and hence $\mathcal{M}_i \vDash \phi_n(a_i)$. ■

From these claims, noting that filters are closed under supersets, we obtain $\{i < \omega : \mathcal{M}_i \vDash \phi_n(a_i)\} \in \mathcal{U}$ for every $n < \omega$. By Łoś' Theorem, we have $\mathcal{M} \vDash \phi_n(a)$ for all $n < \omega$, as $a = [(a_i)]$. □

Remark 29.5. Note that $(\mathbb{Z}, <)$ is not $\aleph_1$-compact. Indeed, consider the collection $\{F_i : i < \omega\}$ where $F_i := \{m \in \mathbb{Z} : m > i\}$ for each $i < \omega$. This has the finite intersection property, but $\bigcap_{i<\omega} F_i = \varnothing$, as there is no greatest integer.

Remark 29.6. Consider the ordered ring of reals $\mathcal{R} = (\mathbb{R}, 0, 1, +, -, \times, <)$. Fix a non-principal ultrafilter $\mathcal{U}$ on ω. Then, from Proposition 29.2, we have $\mathcal{R} \preceq \mathcal{R}^\omega/\mathcal{U} =: \mathcal{R}^*$. We call $\mathcal{R}^*$ a "non-standard model" of $\mathcal{R}$. By Proposition 29.4, $\mathcal{R}^*$ is $\aleph_1$-compact. In particular, we can find that there are elements in $\mathcal{R}^*$ that is bigger than every integer, called infinite elements. Moreover, there exist infinitesimal elements, namely elements greater than 0 but less than $1/n$ for every $0 < n < \omega$. This is the beginning of non-standard analysis. Since $\mathcal{R}$ and $\mathcal{R}^*$ are elementarily equivalent, we may prove a statement in $\mathcal{R}^*$, and if it is a sentence, then it must also hold true in $\mathcal{R}$.

30 Compactness Theorem

Before we prove the Compactness Theorem, we will give one more application of Łoś' Theorem.

PROPOSITION 30.1. The class of finite groups is not axiomatisable in $L = \{0, +, -\}$.

PROOF. For each $n < \omega$, let $\mathcal{G}_n$ be a group of size n. For example, we may take $\mathcal{G}_n = (\mathbb{Z}/n\mathbb{Z}, 0, +, -)$ for each $n < \omega$, and we have a sequence $(\mathcal{G}_n : n < \omega)$ of L-structures. Now, fix a non-principal ultrafilter $\mathcal{U}$ on ω, and consider

$$\mathcal{G} = \prod_{\mathcal{U}} \mathcal{G}_n.$$

Let T be the theory of groups and let $\sigma \in T$. By Łoś' Theorem, we have

$$\mathcal{G} \vDash \sigma \iff \{i \in I : \mathcal{G}_i \vDash \sigma\} = I \in \mathcal{U}.$$

Hence, $\mathcal{G} \vDash \sigma$, and so $\mathcal{G} \vDash T$. That is, $\mathcal{G}$ is a group.

Next, for each $N < \omega$, let

$$\sigma_N := \exists x_1 \cdots \exists x_N \left(\bigwedge_{1 \le i < j \le N} (x_i \neq x_j) \right).$$

Each σ_N says that there exists at least N distinct elements. Then by Łoś' Theorem,

$$\mathcal{G} \vDash \sigma_N \iff \{n : \mathcal{G}_n \vDash \sigma_N\} = \{n : n \ge N\} \in \mathcal{U}.$$

Observe that $\{n : n \ge N\}$ is an element of the Fréchet filter, and that the Fréchet filter is a subset of $\mathcal{U}$ since $\mathcal{U}$ is non-principal. Thus, $\mathcal{G} \vDash \sigma_N$ for all $N < \omega$, and so $\mathcal{G}$ is infinite. (Exercise: In fact, since $\mathcal{G}$ is $\aleph_1$-compact, we have $|\mathcal{G}| > \aleph_0$.)

Finally, suppose that the class of finite groups was axiomatisable, say by the L-theory T_{fin}. Then $\mathcal{G}_n \vDash T_{\text{fin}}$ for all $n < \omega$. By Łoś' Theorem, we obtain $\mathcal{G} \vDash T_{\text{fin}}$, so $\mathcal{G}$ is finite, a contradiction. □

We now restate the Compactness Theorem, and use the tools that we have developed to prove it.

THEOREM 30.2 (Compactness Theorem). Let L be a language and T be an L-theory. T is consistent if and only if every finite subset of T is consistent.

PROOF. The forward direction is obvious; the backward direction is the only one that needs proving. Suppose that every finite subset of T is consistent. Let

$$I = \mathcal{P}^{\text{fin}}(T) := \{\Sigma \subseteq T : \Sigma \text{ is finite}\}.$$

(Note that this forms a lattice structure.) For each $\Sigma \in I$, there exists a model $\mathcal{M}_\Sigma$ of Σ by assumption. Then $(\mathcal{M}_\Sigma : \Sigma \in I)$ is a sequence of L-structures. Then, for all $\Sigma \in I$, let

$$X_\Sigma = \{\Sigma' \in I : \Sigma \subseteq \Sigma'\}.$$

Intuitively, this is the cone above Σ in I. Let $\mathcal{A} = \{X_\Sigma : \Sigma \in I\}$. Note that $\mathcal{A}$ is not a filter. However, we can see that $\varnothing \notin \mathcal{A}$, $I = X_\varnothing \in \mathcal{A}$, and for any $\Sigma, \Delta \in I$, we have $X_\Sigma \cap X_\Delta = X_{\Sigma \cup \Delta} \in \mathcal{A}$. Then, we can make a filter by

$$\mathcal{F} = \{Y \subseteq I : Y \supseteq X_\Sigma \text{ for some } \Sigma \in I\}.$$

We leave it as an exercise to check that this is indeed a filter on I. Let $\mathcal{U} \supseteq \mathcal{F}$ be an ultrafilter extending $\mathcal{F}$. (Exercise: If T is an infinite theory, then $\mathcal{U}$ is non-principal.)

Let $\mathcal{M} := \prod_{\mathcal{U}} \mathcal{M}_\Sigma$. We claim that $\mathcal{M} \vDash T$. Indeed, let $\sigma \in T$. Then

$$X_{\{\sigma\}} = \{\Sigma \in I : \sigma \in \Sigma\} \subseteq \{\Sigma \in I : \mathcal{M}_\Sigma \vDash \sigma\}.$$

Since $X_{\{\sigma\}} \in \mathcal{A} \subseteq \mathcal{F} \subseteq \mathcal{U}$, it follows that

$$\{\Sigma \in I : \mathcal{M}_\Sigma \vDash \sigma\} \in \mathcal{U}$$

as filters are closed under supersets. Finally, by Łoś' Theorem, $\mathcal{M} \vDash \sigma$. So $\mathcal{M} \vDash T$, and we may conclude that T is consistent. □

The following is an equivalent formulation of the Compactness Theorem.

COROLLARY 30.3. Let T be an L-theory and σ be an L-sentence. If $T \vDash \sigma$, then there is a finite subset $\Sigma \subseteq T$ such that $\Sigma \vDash \sigma$.

PROOF. Let $S = T \cup \{\neg\sigma\}$. Since $T \vDash \sigma$, we have that S is inconsistent. By the Compactness Theorem, there is a finite subset $\Delta \subseteq S$ which is inconsistent. It follows that $\Delta \cup \{\neg\sigma\}$ is inconsistent as well. But $\Delta \cup \{\neg\sigma\} = \Sigma \cup \{\neg\sigma\}$ for some finite $\Sigma \subseteq T$. Since $\Sigma \cup \{\neg\sigma\}$ is inconsistent, we must have $\Sigma \vDash \sigma$. □

31 First consequences of compactness

Having just proved the Compactness Theorem, we will show some typical applications of it to particular classes of structures.

Example 31.1. Let $L = \varnothing$. The class of infinite L-structures is not finitely axiomatisable.

Proof. Let T be the natural L-theory of infinite sets; that is, for each $n < \omega$, let

$$\tau_n := \exists x_1 \cdots \exists x_n \left(\bigwedge_{1 \leq i < j \leq n} (x_i \neq x_j) \right)$$

and set $T = \{\tau_n : n < \omega\}$. Suppose that the class of infinite L-structures was finitely axiomatisable. Let σ be an L-sentence such that $\mathcal{M} \vDash \sigma$ if and only if the universe of $\mathcal{M}$ is infinite. Clearly, $T \vDash \sigma$. By compactness, there is a finite subset $\Sigma \vDash T$ such that $\Sigma \vDash \sigma$. In particular, there exists $m < \omega$ such that $\Sigma \subseteq \{\tau_1, \ldots, \tau_m\}$. Then, we see that $\{\tau_1, \ldots, \tau_m\} \vDash \sigma$. Now, let $\mathcal{M}$ be a finite set of size $m + 1$. We have $\mathcal{M} \vDash \{\tau_1, \ldots, \tau_m\}$, but $\mathcal{M} \nvDash \sigma$, a contradiction. □

Example 31.2. Let $L = \{0, +, -\}$. The class $\mathcal{K}$ of abelian torsion groups is not elementary.

Proof. We will use compactness to prove this, but note that we may very well use ultraproducts and Łoś' Theorem as seen in Proposition 30.1.

Suppose for a contradiction that $\mathcal{K}$ were elementary, and let T be an axiomatisation of $\mathcal{K}$. Let $L' := L \cup \{c\}$ where c is a new constant symbol. Let $T' := T \cup \{\tau_n : 0 < n < \omega\}$ be the L'-theory where for each $0 < n < \omega$, we have

$$\tau_n := (\underbrace{c + \cdots + c}_{n \text{ times}} \neq 0).$$

We claim that T' is consistent. Indeed, suppose that $\Sigma \subseteq T'$ is a finite subset. Then $\Sigma \subseteq T \cup \{\tau_1, \ldots, \tau_m\}$ for some $m < \omega$. Let $\mathcal{M} = (\mathbb{Z}/(m+1)\mathbb{Z}, 0, +, -) \vDash T$. We expand $\mathcal{M}$ to an L'-structure $\mathcal{M}'$ by setting $c^{\mathcal{M}'} = 1 \pmod{m+1}$. Observe that $\mathcal{M}' \vDash T$. Moreover, $\mathcal{M}' \vDash \tau_i$ for all $i \leq m$, so $\mathcal{M}' \vDash \Sigma$. It follows from compactness that T' is consistent.

Finally, let $\mathcal{N} = (N, 0, +, -, a = c^{\mathcal{N}}) \vDash T'$ where $a \in N$. Then $(N, 0, +, -)$, the L-reduct of $\mathcal{N}$, is a model of T, and hence a torsion group. But $\mathcal{N} \vDash \tau_n$ for all $n < \omega$, so for all $0 < n < \omega$,

$$\underbrace{a + \cdots + a}_{n \text{ times}} \neq 0$$

in $\mathcal{N}$. Thus, a is not torsion, a contradiction. □

Proposition 31.3. Let $\mathcal{M}$ and $\mathcal{N}$ be L-structures. Then $\mathcal{M} \equiv \mathcal{N}$ if and only if there exists an L-structure $\mathcal{R}$ such that $\mathcal{M} \preceq \mathcal{R}$ and $\mathcal{N} \preceq \mathcal{R}$.

Proof. The right-to-left direction is clear. Conversely, suppose that $\mathcal{M} \equiv \mathcal{N}$. Let $L' = L_{M \sqcup N} := L \cup \{\underline{a} : a \in M\} \cup \{\underline{b} : b \in N\}$. Then $T' := \mathrm{Th}(\mathcal{M}_M) \cup \mathrm{Th}(\mathcal{M}_N)$ is an L'-theory. By Proposition 25.6, it suffices to show that T' is consistent. Indeed, if $\mathcal{R}' \vDash \mathcal{T}'$, let $\mathcal{R}$ be the L-reduct of $\mathcal{R}'$. Since $\mathcal{R}$ can be expanded to a model of $\mathrm{Th}(\mathcal{M}_M)$, Proposition 25.6 says that there exists an elementary embedding $\mathcal{M} \to \mathcal{R}$. Analogously, there exists an elementary embedding $\mathcal{N} \to \mathcal{R}$.

Let $\Sigma \subseteq T'$ be finite. Taking conjunctions, we may assume that $\Sigma = \{\phi(\underline{a_1}, \ldots, \underline{a_m}), \psi(\underline{b_1}, \ldots, \underline{b_n})\}$ where $\phi(x_1, \ldots, x_m)$, $\psi(y_1, \ldots, y_n)$ are L-formulas, $a_1, \ldots, a_m \in M$, $b_1, \ldots, b_n \in N$, and $\mathcal{M} \vDash \phi(a_1, \ldots, a_m)$, $\mathcal{N} \vDash \psi(b_1, \ldots, b_n)$.

In particular, $\mathcal{N} \vDash \exists y_1 \cdots \exists y_n \, \psi(y_1, \ldots, y_n)$ and since $\mathcal{M} \equiv \mathcal{N}$, we have $\mathcal{M} \vDash \exists y_1 \cdots \exists y_n \, \psi(y_1, \ldots, y_n)$ as well. Hence, there exist $a'_1, \ldots, a'_n \in M$ such that $\mathcal{M} \vDash \psi(a'_1, \ldots, a'_n)$. Let $\mathcal{M}'$ be the L'-structure expanding $\mathcal{M}$ such that $\underline{a}^{\mathcal{M}'} = a$ for all $a \in M$, $\underline{b_i}^{\mathcal{M}'} = a'_i$ for all $1 \leq i \leq n$, and any other $b \in N$ is interpreted arbitrarily in $\mathcal{M}'$. Then $\mathcal{M}' \vDash \Sigma$. By compactness, it follows that T' is consistent. □

32 Upward Löwenheim-Skolem and Vaught

We now turn our attention to more general and theoretical applications of compactness.

THEOREM 32.1 (Upward Löwenheim-Skolem). Suppose $\mathcal{M}$ is an infinite L-structure, and κ is a cardinal such that $\kappa \geq \max\{|M|, |L|\}$. Then there exists an elementary extension $\mathcal{N}$ of $\mathcal{M}$ such that $|N| = \kappa$.

PROOF. Let $L' := L_M \cup \{c_\alpha : \alpha < \kappa\}$ where each c_α is a new constant symbol. Let

$$T' := \mathrm{Th}(\mathcal{M}_M) \cup \{c_\alpha \neq c_\beta : \text{for all } \alpha, \beta < \kappa \text{ such that } \alpha \neq \beta\}.$$

We show that T' is consistent. To that end, let $\Sigma \subseteq T'$ be finite. Note that Σ involves only finitely many sentences of the form $c_\alpha \neq c_\beta$, so $\mathcal{M}_M$ can be expanded to a model of Σ as M is infinite. Thus, T' is consistent by compactness.

Let $\mathcal{R}' \vDash T'$ and let $\mathcal{R}$ be the L-reduct of $\mathcal{R}'$. Since $\mathcal{R}$ can be expanded to a model of $\mathrm{Th}(\mathcal{M}_M)$, it follows from Proposition 25.6 that there is an elementary embedding $\mathcal{M} \to \mathcal{R}$. Moreover, note that $|R| \geq \kappa$, and this is witnessed by interpretations of the c_α in $\mathcal{R}'$.

Finally, by Downward Löwenheim-Skolem, there is an L-structure $\mathcal{N} \preceq \mathcal{R}$ such that $|N| = \kappa$ and $\mathcal{M} \subseteq \mathcal{N}$. Since $\mathcal{M} \subseteq \mathcal{N} \preceq \mathcal{R}$ and $\mathcal{M} \preceq \mathcal{R}$, it follows that $\mathcal{M} \preceq \mathcal{N}$ (exercise), so we are done. □

COROLLARY 32.2 (Vaught's Test). Suppose that T is an L-theory with only infinite models. Moreover, suppose that for some infinite cardinal $\kappa \geq |L|$, all models of T of size κ are isomorphic (in which we say that T is **κ-categorical**). Then T is complete.

PROOF. Let $\mathcal{M}$ and $\mathcal{N}$ be two models of T. Using either Downward Löwenheim-Skolem or Upward Löwenheim-Skolem, we can find an L-structure $\mathcal{M}'$ such that either $\mathcal{M} \preceq \mathcal{M}'$ or $\mathcal{M}' \preceq \mathcal{M}$ with $|M'| = \kappa$. Similarly, we can find an L-structure $\mathcal{N}'$ such that either $\mathcal{N} \preceq \mathcal{N}'$ or $\mathcal{N}' \preceq \mathcal{N}$ with $|N'| = \kappa$. By κ-categoricity, $\mathcal{M}'$ is isomorphic to $\mathcal{N}'$. But $\mathcal{M}' \equiv \mathcal{M}$ and $\mathcal{N}' \equiv \mathcal{N}$, so $\mathcal{M} \equiv \mathcal{N}$. By Lemma 26.3, T is complete. □

The following is a classic example of an application of Vaught's Test.

EXAMPLE 32.3. Let $L = \{<\}$. We denote by DLO the L-theory of dense linear orderings without endpoints. Then DLO is complete.

PROOF. We show that DLO is $\aleph_0$-categorical. We will use a classic "back-and-forth" argument (also known as an Ehrenfeucht-Fraïssé game).

Suppose that $(E_1, <) \vDash \text{DLO}$ and $(E_2, <) \vDash \text{DLO}$ are both countable. We construct recursively finite sets $A_0 \subseteq A_1 \subseteq \cdots$ and $B_0 \subseteq B_1 \subseteq \cdots$ such that $E_1 = \bigcup_{i<\omega} A_i$ and $E_2 = \bigcup_{i<\omega} B_i$, with an order preserving bijection $f_i : A_i \to B_i$ where $f_i \subseteq f_{i+1}$ for all $i < \omega$. Once we do this, then

$$f := \bigcup_{i<\omega} f_i : E_1 \to E_2$$

is a order preserving bijection. In particular, f is an L-isomorphism from $(E_1, <)$ to $(E_2, <)$.

To give some intuition for what we are about to construct, we note that in the odd steps, we ensure that $\bigcup_{i<\omega} A_i = E_1$ ("forth"), and in the even steps, we ensure that $\bigcup_{i<\omega} B_i = E_2$ ("back").

First, enumerate $E_1 = \{a_i : i < \omega\}$ and $E_2 = \{b_i : i < \omega\}$.

Step 0: Let $A_0 = B_0 = \varnothing = f_0$.

Now, suppose that we have built an order preserving bijection $f_n : A_n \to B_n$ where $A_n \subseteq E_1$ and $B_n \subseteq E_2$ are finite subsets.

Step $n+1$: Suppose that $n+1$ is odd; that is, $n+1 = 2m+1$ for some $m < \omega$. We want to get a_m into A_{n+1}. If $a_m \in A_n$, then we set $A_{n+1} = A_n$, $B_{n+1} = B_n$, and $f_{n+1} = f_n$. If $a_m \notin A_n$, consider the relationship of a_m to A_n. We have one of the following three cases.

(1) a_m is less than every element in A_n.

(2) a_m is greater than every element in A_n.

(3) There are consecutive elements $\alpha < \beta$ in A_n such that $\alpha < a_m < \beta$.

Let $A_{n+1} = A_n \cup \{a_m\}$. We now choose $b \in E_2$ based on which case a_m falls in.

(1) Let $b \in E_2$ be less than every element of B_n, which exists since $(E_2, <)$ has no endpoints.

(2) Let $b \in E_2$ be greater than every element of B_n.

(3) Choose $b \in E_2$ such that $f_n(\alpha) < b < f_n(\beta)$, which is possible since f_n preserves order and $(E_2, <)$ is dense.

Then, set $B_{n+1} = B_n \cup \{b\}$ and $f_{n+1} = f_n \cup \{(a_m, b)\}$.

On the other hand, suppose $n+1$ is even, so $n+1 = 2m$ for some $m < \omega$. In this case, we want to get b_m into B_{n+1}. As before, if $b_m \in B_n$, then set $B_{n+1} = B_n$, $A_{n+1} = A_n$, and $f_{n+1} = f_n$. Otherwise, $b_m \notin B_n$, so b_m sits with respect to the elements of B_n in three possible cases analogously to cases (1), (2), and (3) above. Choose $a \in E_1$ according to which case b_m satisfies. Then, set $B_{n+1} = B_n \cup \{b_m\}$, $A_{n+1} = A_n \cup \{a\}$, and $f_{n+1} = f_n \cup \{(a, b_m)\}$.

Hence, we have shown that DLO is $\aleph_0$-categorical. Since every model of DLO is infinite and L is finite, it follows from Vaught's Test that DLO is complete. □

Remark 32.4. Note that $(\mathbb{Q}, <) \equiv (\mathbb{R}, <)$, so in particular, "completeness of $<$" cannot be expressed using L-sentences.

33 ACF_p is complete

In this lecture, we fix the language of rings $L = \{0, 1, +, -, \times\}$. Recall that in Example 26.2, we stated that ACF_p, the L-theory of algebraically closed fields of characteristic p (where p is either prime or zero), is complete. We now prove this using Vaught's Test.

LEMMA 33.1. Suppose that $K \vDash \mathrm{ACF}_p$ is uncountable. Then $\mathrm{trdeg}(K/\mathbb{F}) = |K|$, where $\mathbb{F} \subseteq K$ is the prime field; that is, $\mathbb{F} = \mathbb{Q}$ for $p = 0$ and $\mathbb{F} = \mathbb{Z}/p\mathbb{Z}$ for prime p.

PROOF. In general, if B is a finite set, then $|\mathbb{F}(B)^{\mathrm{alg}}| = \aleph_0$. To see this, observe that $|\mathbb{F}(B)| = \aleph_0$. Since there are only $\sum_{n<\omega} |\mathbb{F}(B)|^n = \aleph_0$ polynomials over $\mathbb{F}(B)$, each with finitely many roots, we obtain $|\mathbb{F}(B)^{\mathrm{alg}}| \leq \aleph_0$. On the other hand, $|\mathbb{F}(B)^{\mathrm{alg}}| \geq \aleph_0$ follows from the structure of finite fields.

If B is infinite, then $|\mathbb{F}(B)^{\mathrm{alg}}| = |B|$. Clearly, $\mathbb{F}(B)^{\mathrm{alg}} \supseteq B$, so $|\mathbb{F}(B)^{\mathrm{alg}}| \geq |B|$. Conversely, the number of polynomials over $\mathbb{F}(B)$ is bounded above by

$$\sum_{n<\omega} |\mathbb{F}(B)|^n = \sum_{n<\omega} |B|^n = |B|,$$

so $|\mathbb{F}(B)^{\mathrm{alg}}| \leq |B|$.

Finally, suppose K is uncountable. Let $B \subseteq K$ be a transcendence basis so that $K = \mathbb{F}(B)^{\mathrm{alg}}$. Then B is infinite, so we obtain

$$\mathrm{trdeg}(K/\mathbb{F}) = |B| = |\mathbb{F}(B)^{\mathrm{alg}}| = |K|. \qquad \square$$

THEOREM 33.2. Let κ be an uncountable cardinal. Then ACF_p is κ-categorical.

PROOF. Suppose that $K \vDash \mathrm{ACF}_p$ and $L \vDash \mathrm{ACF}_p$ where $|K| = |L| = \kappa$. Let $B \subseteq K$ be a transcendence basis over $\mathbb{F}$, and similarly, let $C \subseteq L$ be a transcendence basis over $\mathbb{F}$. By the previous lemma, we have that $|B| = |K| = \kappa = |L| = |C|$. Let $\alpha : B \to C$ be a bijection. Then this extends to an isomorphism

$$\mathbb{F}(B) \to \mathbb{F}(C) : \frac{P(b_1, \ldots, b_n)}{Q(b_1, \ldots, b_n)} \mapsto \frac{P(\alpha(b_1), \ldots, \alpha(b_n))}{Q(\alpha(b_1), \ldots, \alpha(b_n)}.$$

By the uniqueness of algebraic closure, this further extends to an isomorphism $K = \mathbb{F}(B)^{\mathrm{alg}} \to \mathbb{F}(C)^{\mathrm{alg}} = L$. $\square$

COROLLARY 33.3. ACF_p is complete.

PROOF. Every algebraically closed field is infinite, and the language of rings is finite. By the previous theorem and Vaught's Test, ACF_p is complete. $\square$

REMARK 33.4. As a consequence of this, $(\mathbb{Q}^{\mathrm{alg}}, 0, 1, +, -, \times) \equiv (\mathbb{C}, 0, 1, +, -, \times)$, so no sentence in the language of rings can distinguish between these two structures.

REMARK 33.5. While ACF_p is κ-categorical for any uncountable cardinal κ, note that it is not $\aleph_0$-categorical. To see this, let $\mathbb{F}$ be the prime field and consider

$$\mathbb{F}^{\mathrm{alg}} \ncong \mathbb{F}(t_1)^{\mathrm{alg}} \ncong \mathbb{F}(t_1, t_2)^{\mathrm{alg}} \ncong \cdots$$

where $t_1, t_2, \ldots$ are indeterminates. Each of these fields is countable.

EXERCISE 33.6. DLO is not $\aleph_1$-categorical.

REMARK 33.7. Morley's Categoricity Theorem states that if L is a countable language and T is an L-theory which is κ-categorical for some uncountable cardinal κ, then T is in fact λ-categorical for every uncountable cardinal λ. The proof of this is quite involved and is outside the scope of the course. However, this gives some insight into the situation we had with ACF_p, where ACF_p is κ-categorical for any uncountable cardinal κ. Moreover, due to the above exercise, DLO is not λ-categorical for any uncountable cardinal λ, because if it were, then it would also be $\aleph_1$-categorical.

As a consequence of the completeness of ACF_p, we can prove the following "metatheorem" in algebraic geometry. For the rest of the lecture, for any prime p, let $\mathbb{F}_p$ denote the prime field of p elements, $\mathbb{Z}/p\mathbb{Z}$.

THEOREM 33.8 (Lefschetz Principle). Let $L = \{0, 1, +, -, \times\}$ and suppose that σ is an L-sentence. The following are equivalent.

(i) $K \vDash \sigma$ for some $K \vDash \mathrm{ACF}_0$.

(ii) $K \vDash \sigma$ for all $K \vDash \mathrm{ACF}_0$.

(iii) $\mathbb{F}_p^{\mathrm{alg}} \vDash \sigma$ for all but finitely primes.

(iv) $\mathbb{F}_p^{\mathrm{alg}} \vDash \sigma$ for infinitely many primes.

PROOF. (i) $\Rightarrow$ (ii). This follows from the completeness of ACF_0.

(ii) $\Rightarrow$ (iii). We have $\mathrm{ACF}_0 \vDash \sigma$. By compactness, $\Sigma \vDash \sigma$ for some finite set $\Sigma \subseteq \mathrm{ACF}_0$. In particular, ACF_0 is made up of the axioms for algebraically closed fields, ACF, together with sentences of the form

$$\tau_n := \forall x \, (x \neq 0 \to \underbrace{x + \cdots + x}_{n \text{ times}} \neq 0)$$

for all $n < \omega$. For some $N < \omega$, we have $\Sigma \subseteq \mathrm{ACF} \cup \{\tau_1, \ldots, \tau_N\}$. Let $p > N$ be prime. Then $\mathrm{ACF}_p \vDash \tau_n$ for $1 \leq n \leq N$, and hence $\mathrm{ACF}_p \vDash \sigma$. Hence, $\mathbb{F}_p^{\mathrm{alg}} \vDash \sigma$ for all but finitely primes.

(iii) $\Rightarrow$ (iv). This is clear.

(iv) $\Rightarrow$ (i). We prove the contrapositive. Suppose that $K \vDash \mathrm{ACF}_0$ with $K \nvDash \sigma$. Then $K \vDash \neg\sigma$. By the completeness of ACF_0, we obtain $\mathrm{ACF}_0 \vDash \neg\sigma$. From (ii) implies (iii) which we just proved, $\mathrm{ACF}_p \vDash \neg\sigma$ for all but finitely many primes p. Thus, $\mathbb{F}_p^{\mathrm{alg}} \vDash \sigma$ for only finitely many primes. $\square$

34 Quantifier elimination

In Lecture 24, we looked at the definable sets of some particular structures, and briefly discussed the notion of quantifier elimination. We now explore this further.

DEFINITION 34.1 (T-equivalent). Let T be an L-theory. Let $\phi(x_1, \ldots, x_n)$ and $\psi(x_1, \ldots, x_n)$ be L-formulas. We say that ϕ and ψ are **T-equivalent** if

$$T \vDash \forall x_1 \cdots \forall x_n \, (\phi(x_1, \ldots, x_n) \leftrightarrow \psi(x_1, \ldots, x_n)).$$

Equivalently, for every model $\mathcal{M} \vDash T$, we have $\phi^{\mathcal{M}} = \psi^{\mathcal{M}}$.

DEFINITION 34.2 (Quantifier elimination). An L-theory T admits **quantifier elimination** if every formula is T-equivalent to a quantifier-free formula. In particular, in every model of T, all definable sets are quantifier-free definable.

We would like to develop some criterion for a theory to have quantifier elimination. Towards that end, we introduce some more basic notions.

DEFINITION 34.3. Suppose that $\mathcal{M}$ is an L-structure and $A \subseteq M$. The **substructure generated by** A is the smallest substructure of $\mathcal{M}$ whose universe contains A, if it exists. If this happens to be $\mathcal{M}$ itself, we say that A **generates** $\mathcal{M}$.

LEMMA 34.4. Suppose that $\mathcal{M}$ is an L-structure and $A \subseteq M$. Moreover, assume that A is non-empty or L contains a constant symbol. Then the substructure generated by A exists, with universe

$$\{t^{\mathcal{M}}(a_1, \ldots, a_n) : n < \omega,\, a_1, \ldots, a_n \in A,\, t(x_1, \ldots, x_n) \text{ is an } L\text{-term}\}.$$

PROOF. By Exercise 17.5, a non-empty subset is the universe of a substructure if and only if it contains all the constants and is preserved by all the basic functions. We show this is the case for

$$N := \{t^{\mathcal{M}}(a_1, \ldots, a_n) : n < \omega,\, a_1, \ldots, a_n \in A,\, t(x_1, \ldots, x_n) \text{ is an } L\text{-term}\}.$$

Considering the L-term x, it is clear that $A \subseteq N$. If $c \in L^C$, then c is itself an L-term and hence $c^{\mathcal{M}} \in N$. That is, N contains A and all constants of $\mathcal{M}$, so in particular, $N \neq \varnothing$. If $f \in L^F$ is ℓ-ary and $a_1, \ldots, a_\ell \in N$ where each $a_i = t_i^{\mathcal{M}}(a_{i,1}, \ldots, a_{i,n_i})$, then

$$f^{\mathcal{M}}(a_1, \ldots, a_\ell) = t^{\mathcal{M}}(a_{1,1}, \ldots, a_{1,n_1}, \ldots, a_{\ell,1}, \ldots, a_{\ell,n_\ell})$$

where $t = f(t_1, \ldots, t_\ell)$, so $f^{\mathcal{M}}(a_1, \ldots, a_\ell) \in N$. It follows that N is the universe of a substructure of $\mathcal{M}$, say $\mathcal{N}$. Now, to see that $\mathcal{N}$ is the substructure generated by A, we show that if $\mathcal{N}' \subseteq \mathcal{M}$ and $A \subseteq N'$, then $N \subseteq N'$. Indeed, for any L-term $t(x_1, \ldots, x_n)$ and $a_1, \ldots, a_n \in A \subseteq N'$, we have $t^{\mathcal{M}}(a_1, \ldots, a_n) = t^{\mathcal{N}'}(a_1, \ldots, a_n)$ since $\mathcal{N}' \subseteq \mathcal{M}$. Hence, $t^{\mathcal{M}}(a_1, \ldots, a_n) \in N'$. □

In Proposition 25.6, we saw that there is an elementary embedding from $\mathcal{M}$ to $\mathcal{N}$ if and only if $\mathcal{N}$ can be expanded to a model of $\mathrm{Th}(\mathcal{M}_M)$. We give a similar criterion for the existence of an embedding, but refine it to take into account generating sets.

LEMMA 34.5. Suppose $\mathcal{M}$ is an L-structure generated by $A \subseteq M$. Assume that A is non-empty or L contains a constant symbol. Consider the L_A-theory

$$\mathrm{qfTh}(\mathcal{M}_A) := \{\phi(\underline{a_1}, \ldots, \underline{a_n}) : n < \omega,\, a_1, \ldots, a_n \in A,\, \phi \text{ quantifier-free, and } \mathcal{M} \vDash \phi(a_1, \ldots, a_n)\}.$$

Suppose that $\mathcal{N}$ is an L-structure. Then there exists an L-embedding $j : \mathcal{M} \to \mathcal{N}$ if and only if $\mathcal{N}$ can be expanded to an L_A-structure $\mathcal{N}'$ such that $\mathcal{N}' \vDash \mathrm{qfTh}(\mathcal{M}_A)$.

PROOF. If such an embedding $j : \mathcal{M} \to \mathcal{N}$ exists, then we can expand $\mathcal{N}$ to an L_A-structure $\mathcal{N}'$ by $\underline{a}^{\mathcal{N}'} := j(a)$ for each $a \in A$. Then $\mathcal{N}' \vDash \mathrm{qfTh}(\mathcal{M}_A)$ by Proposition 21.1.

Conversely, let $\mathcal{N}' \vDash \mathrm{qfTh}(\mathcal{M}_A)$ be an expansion of $\mathcal{N}$. We define $j : \mathcal{M} \to \mathcal{N}$ as follows. Suppose $b \in M$. By Lemma 34.4, there is an L-term $t(x_1, \ldots, x_n)$ and $a \in A^n$ such that $b = t^{\mathcal{M}}(a)$. Then, $t(\underline{a})$ is an L_A-term, so set $j(b) := t(\underline{a})^{\mathcal{N}'}$. This map is injective. Indeed, if $b \neq b'$ are elements of M with $b = t^{\mathcal{M}}(a)$ and $b' = s^{\mathcal{M}}(a')$ where t and s are L-terms, $a \in A^n$, and $a' \in A^m$, then $(t(\underline{a}) \neq s(\underline{a'})) \in \mathrm{qfTh}(\mathcal{M}_A)$, so $j(b) \neq j(b')$. For $c \in L^C$, we have $j(c^{\mathcal{M}}) = c^{\mathcal{N}'}$ by definition, and $c^{\mathcal{N}'} = c^{\mathcal{N}}$ as $\mathcal{N}'$ expands $\mathcal{N}$. If $f \in L^F$ is n-ary and $b_1, \ldots, b_n \in M$, then writing $b_i = t_i^{\mathcal{M}}(a_i)$ where each t_i is an L-term and $a_i \in A^{n_i}$, we have that $j(b_i) = t_i(\underline{a_i})^{\mathcal{N}'}$. Moreover, $f^{\mathcal{M}}(b_1, \ldots, b_n) = f^{\mathcal{M}}(t_1^{\mathcal{M}}(a_1), \ldots, t_n^{\mathcal{M}}(a_n))$. Hence, it follows that

$$\begin{aligned} j(f(b_1, \ldots, b_n)) &= f(t_1(\underline{a_1}), \ldots, t_n(\underline{a_n}))^{\mathcal{N}'} \\ &= f^{\mathcal{N}'}(t_1(\underline{a_1})^{\mathcal{N}'}, \ldots, t_n(\underline{a_n})^{\mathcal{N}'}) \\ &= f^{\mathcal{N}}(j(b_1), \ldots, j(b_n)), \end{aligned}$$

where the last equality uses the fact that $\mathcal{N}'$ is an expansion of $\mathcal{N}$. Finally, suppose that $R \in L^R$ is n-ary and $b_1, \ldots, b_n \in M$. Again, writing $b_i = t_i^{\mathcal{M}}(a_i)$ where each t_i is an L-term and $a_i \in A^{n_i}$, we have

$$\begin{aligned} (b_1, \ldots, b_n) \in R^{\mathcal{M}} &\iff (t_1^{\mathcal{M}}(a_1), \ldots, t_n^{\mathcal{M}}(a_n)) \in R^{\mathcal{M}} \\ &\iff R(t_1(\underline{a_1}), \ldots, t_n(\underline{a_n})) \in \mathrm{qfTh}(\mathcal{M}_A) \\ &\iff (t_1(\underline{a_1})^{\mathcal{N}'}, \ldots, t_n(\underline{a_n})^{\mathcal{N}'}) \in R^{\mathcal{N}'} \\ &\iff (j(b_1), \ldots, j(b_n)) \in R^{\mathcal{N}}. \end{aligned}$$

Hence, j is an L-embedding, as required. □

We now give a criterion for eliminating quantifiers from a given formula.

THEOREM 34.6. Suppose T is an L-theory and $\phi(x_1, \ldots, x_n)$ is an L-formula. Suppose that $n > 0$ or L contains a constant symbol. The following are equivalent.

(i) $\phi(x_1, \ldots, x_n)$ is T-equivalent to some quantifier-free formula $\psi(x_1, \ldots, x_n)$.

(ii) Suppose $\mathcal{M}$ and $\mathcal{N}$ are both models of T, and that $\mathcal{A}$ is an L-substructure of both $\mathcal{M}$ and $\mathcal{N}$. Then for all $a_1, \ldots, a_n \in A$, $\mathcal{M} \vDash \phi(a_1, \ldots, a_n)$ if and only if $\mathcal{N} \vDash \phi(a_1, \ldots, a_n)$.

PROOF. Suppose that $\phi(x_1, \ldots, x_n)$ is T-equivalent to a quantifier-free formula $\psi(x_1, \ldots, x_n)$. Let $\mathcal{M}, \mathcal{N}$, and $\mathcal{A}$ be as in (ii). Then for any $a = (a_1, \ldots, a_n) \in A^n$, we have

$$\begin{aligned} \mathcal{M} \vDash \phi(a) &\iff \mathcal{M} \vDash \psi(a) \text{ as } \phi \text{ is } T\text{-equivalent to } \psi \text{ and } \mathcal{M} \vDash T \\ &\iff \mathcal{A} \vDash \psi(a) \text{ as } \mathcal{A} \subseteq \mathcal{M} \text{ and Proposition 21.1} \\ &\iff \mathcal{N} \vDash \psi(a) \text{ as } \mathcal{A} \subseteq \mathcal{N} \text{ and Proposition 21.1} \\ &\iff \mathcal{N} \vDash \phi(a) \text{ as } \phi \text{ is } T\text{-equivalent to } \psi \text{ and } \mathcal{N} \vDash T. \end{aligned}$$

Conversely, suppose (ii) holds. We want to find a quantifier-free formula that is T-equivalent to $\phi(x_1, \ldots, x_n)$. Consider the set

$$\Phi := \{\psi(x_1, \ldots, x_n) : \psi \text{ quantifier-free and } T \vDash \forall x_1 \cdots \forall x_n \, (\phi(x_1, \ldots, x_n) \to \psi(x_1, \ldots, x_n))\}.$$

Let $c_1, \ldots, c_n$ be new constant symbols and let $L' := L \cup \{c_1, \ldots, c_n\}$. We denote by $\Phi(c_1, \ldots, c_n)$ the set of L'-sentences $\{\psi(c_1, \ldots, c_n) : \psi \in \Psi\}$.

CLAIM. $T \cup \Psi(c_1, \ldots, c_n) \vDash \phi(c_1, \ldots, c_n)$.

PROOF OF CLAIM. Suppose not. Then there is a model $\mathcal{M} \vDash T$ and $a_1, \ldots, a_n \in M$ such that $\mathcal{M} \vDash \psi(a_1, \ldots, a_n)$ for all $\psi \in \Psi$ but $\mathcal{M} \vDash \neg\phi(a_1, \ldots, a_n)$. Consider the theory

$$T' := T \cup \mathrm{qfTh}(\mathcal{M}_{\{a_1, \ldots, a_n\}}) \cup \{\phi(\underline{a_1}, \ldots, \underline{a_n})\}$$

in the language $L \cup \{\underline{a_1}, \ldots, \underline{a_n}\}$.

SUBCLAIM. T' is consistent.

PROOF OF SUBCLAIM. If not, then by compactness, there is a quantifier-free formula $\psi(x_1, \ldots, x_n)$ with $\mathcal{M} \vDash \psi(a_1, \ldots, a_n)$ such that $T \cup \{\psi(\underline{a_1}, \ldots, \underline{a_n}), \phi(\underline{a_1}, \ldots, \underline{a_n})\}$ is inconsistent. Hence, we have

$$T \vDash \phi(\underline{a_1}, \ldots, \underline{a_n}) \to \neg\psi(\underline{a_1}, \ldots, \underline{a_n}).$$

But $\underline{a_1}, \ldots, \underline{a_n}$ are new constant symbols, so this implies that

$$T \vDash \forall x_1 \cdots \forall x_n \, (\phi(x_1, \ldots, x_n) \to \neg\psi(x_1, \ldots, x_n)).$$

In particular, this gives $\neg\psi \in \Psi$, so $\mathcal{M} \vDash \neg\psi(a_1, \ldots, a_n)$. But this is a contradiction as ψ was chosen such that $\mathcal{M} \vDash \psi(a_1, \ldots, a_n)$. ■

Having proved the subclaim, let $\mathcal{N}' \vDash T'$ and let $\mathcal{N}$ be the L-reduct of $\mathcal{N}'$, so $\mathcal{N} \vDash T$. Let $\mathcal{A}$ be the substructure of $\mathcal{M}$ generated by $\{a_1, \ldots, a_n\}$, which exists by Lemma 34.4, noting that $n > 0$ or L contains a constant symbol. Recall that $\mathcal{N}' \vDash \mathrm{qfTh}(\mathcal{M}_{\{a_1,\ldots,a_n\}})$. Then by Lemma 34.5, there exists an L-embedding $j : \mathcal{A} \to \mathcal{N}$. Identifying $\mathcal{A}$ with its image under j, we may assume that $\mathcal{A} \preceq \mathcal{N}$. Applying (ii), we have $\mathcal{M} \vDash \phi(a_1, \ldots, a_n)$. But this is a contradiction since we assumed that $\mathcal{M} \vDash \neg\phi(a_1, \ldots, a_n)$, which proves the claim. ■

Now, we have that $T \cup \Psi(c_1, \ldots, c_n) \vDash \phi(c_1, \ldots, c_n)$. Write $c = (c_1, \ldots, c_n)$. By compactness, there exist $\psi_1, \ldots, \psi_\ell \in \Psi$ such that $T \cup \{\psi_1(c), \ldots, \psi_\ell(c)\} \vDash \phi(c)$. Let

$$\psi(x_1, \ldots, x_n) := \bigwedge_{i=1}^{\ell} \psi_i(x_1, \ldots, x_n).$$

Then we have that $T \vDash \psi(c) \to \phi(c)$. But $c_1, \ldots, c_n$ are new constant symbols which do not appear in T, so this implies that

$$T \vDash \forall x_1 \cdots \forall x_n \, (\psi(x_1, \ldots, x_n) \to \phi(x_1, \ldots, x_n)).$$

On the other hand, we have $T \vDash \forall x_1 \cdots \forall x_n \, (\phi(x_1, \ldots, x_n) \to \psi(x_1, \ldots, x_n))$ since $\psi_i \in \Psi$ for all $1 \leq i \leq \ell$. Thus, ϕ is T-equivalent to ψ, completing the proof. □

35 A criterion for quantifier elimination

From Theorem 34.6, we can test whether a given formula is equivalent to a quantifier-free formula. We can now develop a criterion for a theory to admit quantifier elimination.

DEFINITION 35.1. An **L-literal** is an atomic L-formula or a negated atomic L-formula.

THEOREM 35.2 (Criterion for QE). Let T be an L-theory satisfying the following condition.

($\star$) Let $\mathcal{M}$ and $\mathcal{N}$ be two models of T and $\mathcal{A}$ a common L-substructure of $\mathcal{M}$ and $\mathcal{N}$. Let $\psi(y)$ be a conjunction of L_A-literals in a single free variable y. If $\psi(y)$ has a solution in $\mathcal{M}$, then it has a solution in $\mathcal{N}$.

Then T admits quantifier elimination.

PROOF. Suppose that $\phi(x)$ is an L-formula where $x = (x_1, \ldots, x_n)$ and $n > 0$. We show, using ($\star$) and by induction on the complexity of ϕ, that ϕ is T-equivalent to a quantifier-free formula in x. Clearly, if $\phi(x)$ is atomic, then it is already a quantifier-free formula. Moreover, note that being T-equivalent to a quantifier-free formula is closed under conjunctions, disjunctions, and negations. As usual, we can write $\forall$ in terms of $\neg$ and $\exists$. So it remains to consider the case where $\phi(x)$ is of the form $\exists y\, \psi(x, y)$ where y is a single variable in ψ and $\psi(x, y)$ is T-equivalent to a quantifier-free formula $\theta(x, y)$. Writing θ in disjunctive normal form, we have

$$\theta(x, y) = \bigvee_i \bigwedge_j \psi_{i,j}(x, y)$$

where each $\psi_{i,j}$ is an L-literal. It now suffices to show that $\exists y \left(\bigvee_i \bigwedge_j \psi_{i,j}(x, y)\right)$ is T-equivalent to a quantifier-free formula in x. For this, we will verify the condition given in Theorem 34.6. Let $\mathcal{M}$ and $\mathcal{N}$ be two models of T, and let $\mathcal{A}$ be a common L-substructure of $\mathcal{M}$ and $\mathcal{N}$. Let $a \in A^n$. If $\mathcal{M} \vDash \exists y \left(\bigvee_i \bigwedge_j \psi_{i,j}(a, y)\right)$, then in particular, $\mathcal{M} \vDash \exists y \left(\bigwedge_j \psi_{i,j}(a, y)\right)$ for some i. Now, observe that $\bigwedge_j \psi_{i,j}(a, y)$ is a conjunction of L_A-literals in y. By ($\star$), $\mathcal{N} \vDash \exists y \left(\bigwedge_j \psi_{i,j}(a, y)\right)$ for the same i, and hence $\mathcal{N} \vDash \exists y \left(\bigvee_i \bigwedge_j \psi_{i,j}(a, y)\right)$. By symmetry, the converse is also true. Hence, we obtain

$$\mathcal{M} \vDash \exists y \left(\bigvee_i \bigwedge_j \psi_{i,j}(x, y) \right) \iff \mathcal{N} \vDash \exists y \left(\bigvee_i \bigwedge_j \psi_{i,j}(x, y) \right).$$

By Theorem 34.6, it follows that $\exists y \left(\bigvee_i \bigwedge_j \psi_{i,j}(x, y)\right)$ is T-equivalent to a quantifier-free formula in x, which we wanted to show. □

To finish this lecture, we give a simple example where we apply this criterion.

EXAMPLE 35.3. Let $L = \varnothing$. The L-theory T of infinite sets admits quantifier elimination.

PROOF. We prove that T satisfies the condition ($\star$) in Theorem 35.2. Suppose $\mathcal{M}$ and $\mathcal{N}$ are infinite sets with a common non-empty subset $\mathcal{A}$. Let $\psi(y)$ be a conjunction of L_A-literals that has a solution in $\mathcal{M}$. We want to show that $\psi(y)$ has a solution in $\mathcal{N}$.

Note that an L_A-literal in the variable y is of one of the forms $y = y$, $y \neq y$, $y = a$, $y \neq b$, $a = b$, or $a \neq b$ where $a, b \in A$. We may throw out the literals of the form $a = b$ and $a \neq b$ out of $\psi(y)$ because if they are true in $\mathcal{M}$, then they are clearly also true in $\mathcal{N}$. We can also throw out $y = y$, as this is satisfied by everything. Also, note that $y \neq y$ cannot appear in $\psi(y)$ since it has no solutions. Hence, we may assume that $\psi(y)$ is of the form

$$\bigwedge_{i=1}^{r} (y = a_i) \wedge \bigwedge_{j=1}^{s} (y \neq b_i),$$

where $a_1, \ldots, a_r, b_1, \ldots, b_s \in A$. If $r \geq 1$, then a_1 is a solution to $\psi(y)$ in $\mathcal{M}$. But $a_1 \in \mathcal{A} \subseteq \mathcal{N}$, so a_1 is also a solution to $\psi(y)$ in $\mathcal{N}$. Otherwise, assume that $r = 0$. Then $\psi(y)$ is simply given by

$$\bigwedge_{j=1}^{s} (y \neq b_i).$$

Since $\mathcal{N} \vDash T$, we have that $\mathcal{N}$ is infinite, so $\psi(y)$ is satisfied in $\mathcal{N}$ by choosing some element $b \in N \setminus \{b_1, \ldots, b_s\}$. $\square$

36 Examples of quantifier elimination

We give a slightly more complicated example than before.

EXAMPLE 36.1. Let $L = \{<\}$. The L-theory DLO admits quantifier elimination.

PROOF. Suppose that $(M, <) \vDash \text{DLO}$ and $(N, <) \vDash \text{DLO}$. Let $(A, <)$ be a common substructure of $(M, <)$ and $(N, <)$. Let $\psi(y)$ be a conjunction of L_A-literals which has a realization in $(M, <)$. We want to show that $\psi(y)$ has a realization in $(N, <)$.

The possible conjuncts in $\psi(y)$ are ones of the form

- $y = y$
- $y \neq y$
- $y < y$
- $y \geq y$
- $y = a$
- $y \neq a$
- $y < a$
- $y \geq a$
- $a < y$
- $a \geq y$

as well as the L_A-literals not involving y. Clearly, the L_A-literals not involving y can be dropped, since if they are true in $\mathcal{M}$, then they are true in $\mathcal{A}$, and hence in $\mathcal{N}$. The first four from the above list can also be dropped because they are either realized by everything or never realized. If $y = a$, $y \geq a$, or $a \geq y$ appears in $\psi(y)$, then $a \in A \subseteq N$ is a solution, so we can also drop these. The only remaining literals we did not drop are of the form $y \neq a$, $y < a$, and $a < y$. Hence, we may assume that $\psi(y)$ is of the form

$$\bigwedge_{i=1}^{r} (y > a_i) \wedge \bigwedge_{j=1}^{s} (y \neq b_j) \wedge \bigwedge_{k=1}^{t} (y < c_k)$$

where $a_1, \ldots, a_r, b_1, \ldots, b_s, c_1, \ldots, c_t \in A$. Since this has a solution in $(M, <)$, it must be that $\max_{1 \leq i \leq r} a_i < \min_{1 \leq k \leq t} c_k$. Since $(N, <) \vDash \text{DLO}$, there are infinitely many elements in N strictly between these two, so choose one that is not equal to any of $b_1, \ldots, b_s$. This will satisfy $\psi(y)$. □

REMARK 36.2. As a consequence of this example, if $(M, <) \vDash \text{DLO}$, then the definable subsets of M are finite unions of points and open intervals. Such an structure in a language containing $<$ is said to be **o-minimal**. A theory T is then **o-minimal** if every model of T is o-minimal. In particular, we see that DLO is o-minimal. It can also be shown that $\text{Th}(\mathbb{R}, 0, 1, +, -, \times, <)$ is o-minimal.

The examples we have done were of complete theories. Let us look at an example in which the theory is not complete.

EXAMPLE 36.3. Let L be the language of rings. The L-theory ACF admits quantifier elimination.

PROOF. Suppose that $K \vDash \text{ACF}$ and $L \vDash \text{ACF}$. Let $R \subseteq K$ and $R \subseteq L$; in particular, R is a common subring of K and L. Let $\psi(y)$ be a conjunction of L_R-literals with a solution in K. We show that $\psi(y)$ has a solution in L.

First, we make some remarks. Note that $\text{char}(K) = \text{char}(L)$ as they both contain the common subring R. Moreover, R is an integral domain as it is the subring of a field, so we have that the fraction field $\text{Frac}(R)$ is a common subfield of K and L. Then, by the uniqueness of algebraic closure, $F := \text{Frac}(R)^{\text{alg}}$ is a common subfield of K and L, as K and L are algebraically closed.

Now, $\psi(y)$ is of the form

$$\left(\bigwedge_{i=1}^{r} P_i(y) = 0 \right) \wedge \left(\bigwedge_{j=1}^{s} Q_j(y) \neq 0 \right)$$

where $P_1, \ldots, P_r, Q_1, \ldots, Q_s \in R[y] \subseteq F[y]$. We may assume that $P_1, \ldots, P_r, Q_1, \ldots, Q_s$ are non-constant, for otherwise they can be dropped from $\psi(y)$. If $r > 0$, then a solution to $\psi(y)$ in K will have to be in

$\mathrm{Frac}(R)^{\mathrm{alg}} = F \subseteq L$, so we are done. Otherwise, $r = 0$, so we have that $\psi(y)$ is of the form

$$\bigwedge_{j=1}^{s} Q_j(y) \neq 0.$$

Every non-constant polynomial in a field has only finitely many solutions. Since L is infinite, there exists $c \in L$ such that $Q_j(c) \neq 0$ for all $1 \leq j \leq s$. Then we have $L \vDash \psi(c)$. □

Remark 36.4. As an immediate consequence, every definable set in a model of ACF is Zariski-constructible, which we noted in Fact 24.2.

37 Hilbert's Nullstellensatz

We now give an application of the fact that ACF admits quantifier elimination.

REMARK 37.1. If T admits quantifier elimination, then T is **model-complete**; that is, whenever $\mathcal{M} \vDash T$ and $\mathcal{N} \vDash T$, if $\mathcal{M} \subseteq \mathcal{N}$, then $\mathcal{M} \preceq \mathcal{N}$. Note that model-completeness is strictly weaker than quantifier elimination. For example, consider $T = \mathrm{Th}(\mathbb{R}, 0, 1, +, -, \times)$, which does not admit quantifier elimination (as seen in Lecture 24). However, it is a fact (which requires some more work) that T is model-complete.

EXAMPLE 37.2. Recall that DLO admits quantifier elimination, and hence is model-complete. Since $(\mathbb{Q}, <) \subseteq (\mathbb{R}, <)$, it follows that $(\mathbb{Q}, <) \preceq (\mathbb{R}, <)$. Similarly, since ACF admits quantifier elimination, we obtain $(\mathbb{Q}^{\mathrm{alg}}, 0, 1, +, -, \times) \preceq (\mathbb{C}, 0, 1, +, -, \times)$.

THEOREM 37.3 (Hilbert's Nullstellensatz). Suppose that $K \vDash \mathrm{ACF}$ and $P_1, \ldots, P_\ell \in K[X_1, \ldots, X_n]$. Suppose there does not exist $Q_1, \ldots, Q_\ell \in K[X_1, \ldots, X_n]$ such that

$$Q_1 P_1 + \cdots + Q_\ell P_\ell = 1. \tag{$\star$}$$

Then there is $a \in K^n$ such that $P_1(a) = P_2(a) = \cdots = P_\ell(a) = 0$.

PROOF. By model-completeness, it suffices to find $K \subseteq L \vDash \mathrm{ACF}$ such that

$$L \vDash \exists x \left(\bigwedge_{i=1}^{\ell} P_i(x) = 0 \right),$$

since $K \preceq L$. Note that the above is an L_K-sentence. By ($\star$), the ideal $I := (P_1, \ldots, P_\ell) \subseteq K[X_1, \ldots, X_n]$ is proper. Let $\mathfrak{M} \supseteq I$ be a maximal ideal. Then, $K[X_1, \ldots, X_n]/\mathfrak{M} =: F$ is a field. Letting π be the projection map, we find that

$$K \xrightarrow{\subseteq} K[X_1, \ldots, X_n] \xrightarrow{\pi} K[X_1, \ldots, X_n]/\mathfrak{M} = F$$

is an embedding, since $\mathfrak{M} \cap K = (0)$. Now, $K \subseteq F \subseteq F^{\mathrm{alg}} =: L \vDash \mathrm{ACF}$. For each $1 \leq i \leq n$, let $b_i := \pi(X_i) \in F \subseteq L$. Then, for $1 \leq j \leq \ell$,

$$\begin{aligned} P_j(b_1, \ldots, b_n) &= P_j(\pi(X_1), \ldots, \pi(X_n)) \\ &= \pi(P_j(X_1, \ldots, X_n)) \\ &= 0 \end{aligned}$$

since $P_j \in I \subseteq \mathfrak{M}$. Thus, $b = (b_1, \ldots, b_n)$ is a solution to all $P_j(x) = 0$ in L. □