PMATH433/733 — Model Theory and Set Theory

Classnotes for Fall 2018

by

Johnson Ng

BMath (Hons), Pure Mathematics major, Actuarial Science Minor

University of Waterloo

Table of Contents

List of Definitions

List of Theorems

Foreword

This course has a ratio of about 1:3 for naive set theory to model theory.

Part I

Set Theory

1 Lecture 1 Sep 06th

1.1 Introduction to Set Theory

In this course, we shall focus only on **practical set theory**, which is more commonly knowned as **naive set theory**. In practical set theory, we look at set theory as a language of mathematics. Some of the examples of which we look into in this flavour of set theory are (transfinite) induction and recursion, and the measuring of the sizes of sets.

Another approach to set theory, one that is deemed required in order to learn set theory is a more formal way, is to look at set theory as the foundations of mathematics. Such an approach is more axiomatic, rigorous, and grounding as compared to practical set theory. This course will try to work around going into these topics, as they can take a life of their own, and within the context of this course, the topics that will be explored using this approach are not required.

1.2 Ordinals

1.2.1 Zermelo-Fraenkel Axioms

We use the natural numbers, i.e.

$$0, 1, 2, 3, 4, \ldots$$

to **count** finite sets. There are two related meanings attached to the word "count" here:

- enumeration; and
- measuring (of sizes)

In order to facilitate the introduction to certain axioms that we shall need, let our current goal be to develop an infinitary generalization of the natural numbers, so as to be able to enumerate and measure arbitrary sets.

We have that

$$\begin{aligned} \text{enumeration} &\to \text{ordinals} \\ \text{measuring} &\to \text{cardinals} \end{aligned}$$

where $\to$ represents "leads to" here.

To CONSTRUCT the natural numbers, we require 3 basic notions that shall remain undefined but understood:

- a set;
- membership, denoted by $\in$; and
- equality.

One such construction is

$$0 := \emptyset, \text{ the empty set}$$
$$1 := \{0\} = \{\emptyset\}, \text{ the set whose only member is } 0$$
$$2 := \{0, 1\} = \{\emptyset, \{\emptyset\}\}, \text{ the set whose only members are } 0 \text{ and } 1.$$

Definition 1 (Successor)

Given a natural number n, the ***successor*** *of n is the natural number next to n, which can be obtained by*

$$S(n) := n \cup \{n\}.$$

We can use the definition of a successor to construct the rest of the natural numbers.

Example 1.2.1

Just to verify to ourselves that the definition indeed works, observe that

$$S(1) = 2 = \{\emptyset, \{\emptyset\}\} = \{\emptyset\} \cup \{\{\emptyset\}\}.$$

So to construct the natural number 3, we see that

$$\begin{aligned} S(2) = 3 = \{0, 1, 2\} = 2 \cup \{2\} &= \{\emptyset, \{\emptyset\}\} \cup \{\{\emptyset, \{\emptyset\}\}\} \\ &= \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\} \end{aligned}$$

Looking at these, we start wondering to ourselves: how do we know that $\emptyset$ exists in the first place? How do we know that we can use $\cup$ and what does it even mean? Now it is meaningless if we cannot take that $\emptyset$ always exists, nor is it meaningful if we cannot take the $\cup$ of sets. And so, to allow us to continue, or even start with these notions, we require axioms.

Axiom 1 (Empty Set Axiom)

There exists a set, denoted by $\emptyset$, with no members.

With this axiom, we can indeed construct 0. To get 1 from 0, we have that 1 is a set whose only member is zero, and so if we take a member from 1, that member must be 0.

Axiom 2 (Pairset Axiom)

Given set x, y, there exists a set, denoted by $\{x, y\}$, whose only members are x and y. In other words,

$$t \in \{x, y\} \leftrightarrow (t = x \lor t = y)$$

Now note that in Axiom 2, if $x = y$, then the set $\{x, y\}$ has only x as its member. For example, we realize that $1 = \{0, 0\} = \{0\}$. But why exactly does this equality make sense? What exactly does "realize" mean?

Axiom 3 (Axiom of Extension)

Given sets x, y, $x = y$ if and only if x and y have the same members.

Now, using the above 3 axioms, we are guaranteed that

$$\begin{aligned} 0 &= \emptyset \text{ exists by the Empty Set Axiom} \\ 1 &= \{\emptyset\} \text{ exists by the Pairset Axiom} \\ 2 &= \{\emptyset, \{\emptyset\}\} \text{ exists by the Pairset Axiom} \end{aligned}$$

Now we've constructed 3 to be the set whose only members are $0, 1$, and 2. So far, within our axioms, there is no such thing as $\{0, 1, 2\}$, which is what our 3 is supposed to be. We now require the following axiom:

Axiom 4 (Union Set Axiom)

Given a set x, there exists a set denoted by $\cup x$, whose members are precisely the members of the members of x, i.e.

$$t \in \cup x \leftrightarrow (t \in y \text{ for some } y \in x)$$

So, by this axiom, we have that given any n, $S(n) = \cup\{n, \{n\}\}$, or in other words,

$$t \in S(n) \leftrightarrow t \in n \vee t = n.$$

With all of the above axioms, we can iteratively construct each and every natural number in a rigorous manner. However, our goal is to construct infinitely many of them.

It is tempting to simply take the infinitude of natural numbers simply as an axiom, i.e.

> *There exists a set whose members are precisely the natural numbers.*

There is a certain rule to which we set down axioms, and that is, axioms must be expressable in a "finitary" manner, i.e. they must be expressible using first-order logic.

Definition 2 (Definite Condition)

We define a ***definite condition*** *as follows:*

- *$x \in y$ and $x = y$ are definite conditions, where x and y are both indeterminants, standing for sets, or are sets themselves;*
- *if P and Q are definite conditions, then so are*
  - *not P, denoted as $\neg P$;*
  - *P and Q, denoted as $P \wedge Q$;*

- *P or Q, denoted as $P \vee Q$;*
- *for all x, P, denoted as $\forall x P$; and*
- *there exists x, P, denoted as $\exists x P$.*

Example 1.2.2

$$x \in 1, 0 \in 2, 2 \in 0$$

are all definite conditions. Note, however, that $2 \in 0$ is false.

Note

*"If P then Q", which is also written as $P \to Q$, is also a definite condition since it is "**equivalent**"[1] to the statement $\neg P \vee Q$.*

Consequently, $P if and only if Q$, which is also expressed as $P \leftrightarrow Q$, can be written as

$$(\neg P \vee Q) \wedge (\neg Q \vee P)$$

[1] We have yet to define what equivalent statements are but we shall take this for granted for now.

Now, with this definition, and first-order logic notations in mind, we can write:

- Empty Set Axiom: $\exists x \, \forall t \, \neg(t \in x)$
- Pairset Axiom: $\forall x \, \forall y \, \exists p \, \forall t \, (t \in p \leftrightarrow ((t = x) \vee (t = y)))$
- Union Set Axiom: $\forall x \, \exists z \, \forall t \, ((t \in z) \leftrightarrow (\exists y \, ((y \in x) \wedge (t \in y))))$

Exercise 1.2.1
Write Axiom 3 in first-order logic notation.

Solution

$$\forall x \, \forall y$$
$$(x = y) \leftrightarrow (\forall t \, ((t \in x) \leftrightarrow (t \in y)))$$

Note that the statement that we proposed as an axiom for the set of natural numbers in page 18 is not definite, although that itself is not obvious.

For example, we may try to write

$$\exists x \, (\forall t \, ((t \in x) \leftrightarrow ((t = 0) \vee (t = 1) \vee (t = 2) \vee \ldots)))$$

and then notice that we do not have the notion of "..." within the "tools" that we are allowed to use.

2 *Lecture 2 Sep 11th*

2.1 *Ordinals (Continued)*

2.1.1 *Zermelo-Fraenkel Axioms (Continued)*

We stopped at the discussion about allowing for an infinite set, so that we can construct our set of infinite natural numbers. The idea here is to take *the smallest set that contains* 0 *and is preserved by the successor function*[1]

[1] Q: Why the smallest set?

Axiom 5 (Infinity Axiom)

There exists a set I *that contains* 0 *and is preserved by the successor function. We may express this as*

$$\exists I((0 \in I) \land \forall x(x \in I \to S(x) \in I))$$

where we have defined that $S(x) \in I$ *means*

$$\exists y(\forall t(t \in y \leftrightarrow (t \in x) \lor (t = x)) \land (y \in I))$$

We call I *the* ***successor set***.

Since we want the smallest of such successor sets, we can try taking the intersection of all successor sets. But before we can do that, we require more axiomatic statements.

Definition 3 (Subsets)

$x \subseteq y$ means that every element of x is an element of y, i.e.

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

With a definition of a subset, we can define the Powerset Axiom.

Axiom 6 (Powerset Axiom)

Given a set x, there exists a set $\mathcal{P}(x)$ that contains all subsets of x, i.e.

$$\forall t((t \subseteq \mathcal{P}(x)) \leftrightarrow (t \subseteq x))$$

We also require the following axiom.

Axiom 7 ((Bounded) Separation Axiom)

Given a set x and a definition condition P, there exists a set whose elements are precisely the members of x that satisfies P, i.e.

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

where

$$y = \{z \in x \mid P(z)\}.$$

There are two important aspects to the Bounded Separation Axiom:

- it is bounded by the set x; and
- P is a definite condition.

Exercise 2.1.1 (Set Intersection)

Prove that given a non-empty set x, there exists a set $\cap x$ satisfying

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

Proof

to be solved

Definition 4 (Natural Numbers)

Let I be a successor set. The set of natural numbers is[2]

$$\omega := \cap\{J \subseteq I : J \text{ is a successor set }\}$$

[2] We can also write $J \subseteq I$ as $J \in \mathcal{P}(I)$ and invoke the Bounded Separation Axiom.

Note

J being a successor set can be expressed by the definite condition

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

so we can write the definite condition in the above definition by

$$\omega := \cap\{J \subseteq I : (0 \in J) \vee \forall x\, (x \in J \to S(x) \in J)\}$$

Exercise 2.1.2

Show that the definition of ω does not actually depend on I, i.e. if given I_1 and I_2 such that we have

$$\omega_1 = \cap\{J \subseteq I_1 : J \text{ is a successor set }\}$$
$$\omega_2 = \cap\{J \subseteq I_2 : J \text{ is a successor set }\}$$

we have

$$\omega_1 = \omega_2.$$

Proof

to be solved

Another useful axiom that we will use later is the following:

Axiom 8 (Replacement Axiom)

Suppose P is a binary definite condition[3] *such that for every set x, there is a unique y satisfying $P(x, y)$. Given a set A, there is a set B such that $t \in B$ if and only if there is an $a \in A$ with $P(a, t)$.*

[3] A **binary definite condition** has only two variables.

Note

The slogan for the Replacement Axiom is:

> *The image of a set under a definite operation exists.*

These eight axioms, along with another ninth axiom called the Regularity Axiom[4], constitutes the **Zermelo-Fraenkel Set Theory**.

Note that all axioms, save the Extensionality, assert the existence of sets.

[4] We shall not discuss too much about this. According to the lecture and the lecture notes, the Regularity Axiom states that every set has a minimal element. On Wikipedia, the axiom states that every set has an element that does not intersect with the set itself.

2.1.2 *Classes*

There are times where we are interested in a collection of sets that do not form a set themselves.

Example 2.1.1 (Russell's Paradox)

There is no set containing all sets.

Proof

Suppose such a set exists, and call it U. Now consider the set

$$R := \{x \in U : x \notin x\},$$

which exists by Bounded Separation. Observe that

$$\begin{aligned} R \in R &\implies R \notin R \quad \lightning \\ \implies R \notin R &\implies R \in R \quad \lightning \end{aligned}$$

Thus such a set U cannot exist.

To talk about such collections, that may or may not be sets, we define *classes*.

Definition 5 (Class)

A class is any collection of sets defined by definite property, i.e. given any difinite condition P,

$$\llbracket \ z \mid P(z) \ \rrbracket$$

is the class of all sets satisfying P.

Here, instead of Bounded Separation, we have what is called **unbounded separation**.

Note

We shall use $[\![\quad]\!]$ rather than $\{ \quad \}$ to emphasize that we are talking about classes, i.e. we may be talking about non-sets.

Example 2.1.2

$$\mathrm{Set} := [\![\ z \mid z = z \ ]\!]$$

is the universal class of all sets.

Note

- *Every set is a class.*

Proof

Suppose x is a set. We may write

$$x = [\![\ z \mid z \in x \ ]\!] .$$

□

- *Some classes are not sets; these are called* ***proper classes****. E.g. the universal class of all sets, and*

$$\mathit{Russell} := [\![\ z \mid z \notin z \ ]\!] .$$

3 Lecture 3 Sep 13th

3.1 Ordinals (Continued 2)

3.1.1 Cartesian Products and Function

Definition 6 (Ordered Pairs)

*Given sets x, y, an **ordered pair** of x and y is defined as*[1]

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

[1] This invokes the Pairset Axiom thrice. Why did we not define an ordered pair as

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

instead?

Note

Note that we must have

$$((x, y) = (x', y')) \iff (x = x' \wedge y = y').$$

Proof

The ($\Longleftarrow$) direction is clear by Extensionality. For the other direction, we shall break it into 2 cases:

Case 1: $x = y$. Then $\{x, y\} = \{x\}$ by Extensionality, and so

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

Therefore, we have that

$$\{\{x\}\} = (x, y) = (x', y') = \{\{x'\}, \{x', y'\}\}$$

So we have

$$\{x\} = \{x'\} \implies x = x'$$

and

$$\{x\} = \{x', y'\} \implies y' = x = y.$$

Thus we have

$$x = x' \land y = y'.$$

Case 2: Suppose $x \neq y$ and $x' \neq y'$ [2] We have

[2] If any of them are equal, Case 1 would apply.

$$\{\{x\}, \{x, y\}\} = \{\{x'\}, \{x', y'\}\}$$

Then

$$\{x\} = \{x'\} \lor \{x\} = \{x', y'\}$$

The latter leads to a contradiction, since it would imply

$$x' = x = y'.$$

Thus $x = x'$. Also, we have

$$\{x, y\} = \{x'\} \lor \{x, y\} = \{x', y'\}$$

Now the former leads to a contradiction since it would imply that

$$x = x' = y.$$

Now since $x = x'$, it must be that $y = y'$, otherwise $y = x' = x$ would contradict our assumption. Therefore, we have that

$$x = x' \land y = y'.$$

□

With ordered pairs, we can build Cartesian products:

Definition 7 (Cartesian Product)

Given classes X and Y, the ***Cartesian Product*** *of X and Y is defined as*

$$X \times Y := [\![\; z : z = (x, y), x \in X, y \in Y \;]\!]$$

Note

We can express this definition using definite conditions;

$$\forall x, y\Big((x \in X) \wedge (y \in Y) \wedge \Big(\exists a, b(\forall t(t \in a \leftrightarrow t = x)) \wedge \forall t(t \in b \leftrightarrow (t = x) \vee (t = y))\Big) \wedge$$
$$\forall t\Big(t \in z \leftrightarrow ((t = a) \vee (t = b))\Big)\Big)$$

Note

- *A Cartesian product is a class.*
- *If A is a set and B is a class, and $B \subseteq A$, then B is also a set. This is easy to show: observe that by Extentionality,*

  $$B = \{a \in A \mid a \in B\}.$$

  By Bounded Separation Axiom, B is a set[3]*.*

[3] This statement can be rephrased as: subclasses of a set are subsets.

Consequently, Cartesian products of sets are sets themselves; if X and Y are sets, we want to show that $X \times Y$ is a set so it is sufficient to show that it is contained in one. Recall that

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

and $\{x, y\} \subset X \cup Y$ which means $\{x, y\} \in \mathcal{P}(X)$, and we observe that $\{x\} \in \mathcal{P}(X \cup Y)$. So $(x, y) \in \mathcal{P}(X \cup Y)$. Therefore, $X \times Y \subset \mathcal{P}(\mathcal{P}(X \cup Y))$, and we show to ourselves that $X \times Y$ is indeed a set.

Definition 8 (Definite Operation)

*Given classes X and Y, a **definite operation** $f : X \to Y$ is a subclass $\Gamma(f) \subseteq X \times Y$ such that*

$$\forall x \in X\ \exists! y \in Y\ (x, y) \in \Gamma(f).$$

Note

We write $f(x) = y$ to mean $(x, y) \in \Gamma(f)$. We also refer to $\Gamma(f)$ as the ***graph of** f*.

Example 3.1.1

The successor function $S : \text{Set} \to \text{Set}$ is a definite operation such that

$$S(x) = x \cup \{x\}$$

This is true since is can be expressed as

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

To show that S is a definite operation, we need to show that S is a definite condition.

Note

If X and Y are sets and f is a definite operation, then $\Gamma(f) \subseteq X \times Y$ is a set. In such a case, we call f a function.

Definition 9 (Functions)

A function is a definite operation $f : X \to Y$ where X and Y are both sets.

We can now restate the Replacement Axiom.

Axiom 9 (Replacement Axiom (Restated))

If $f : X \to Y$ is a definite operation, and $A \subseteq X$ is a set, then $\exists B \subseteq Y$ that is a set such that $t \in B$ if and only if $t = f(a)$ for some $a \in A$.

3.1.2 The Natural Numbers

Theorem 10 (Induction Principle)

Suppose $J \subseteq \omega$, $0 \in J$ and whenever $n \in J$, $S(n) \in J$. Then $J = \omega$.

Proof

By assumption, J is a successor set, therefore $\omega \subseteq J$ by definition. Thus, sinec $J \subseteq \omega$, we have $J = \omega$. □

Lemma 11 (Properties of the Natural Numbers)

Suppose $n \in \omega$. We have

1. $n \subseteq \omega$;
2. $\forall m \in n \quad m \subseteq n$;
3. $n \notin n$;
4. $n = 0 \veebar 0 \in n$; *and*
5. $y \in n \implies S(y) \in n \veebar S(y) = n$.

Proof

1. Let[4]

   $$J := \{n \in \omega : n \subseteq \omega\} \subseteq \omega.$$

   Note that $\emptyset \subseteq \omega$ and so $0 \subseteq \omega$. By membership, $0 \in J$.

   Suppose $m \in J$. Consider $S(m) = m \cup \{m\}$. Since $J \subseteq \omega$, $m \in \omega$. Since $m \in \omega$, $\{m\} \subseteq \omega$. Therefore $S(m) = m \cup \{m\} \subseteq \omega$, and so $S(m) \in J$. So J is a successor set. And thus by Induction Principle, $J = \omega$.

2. Let

   $$J := \{n \in \omega : \forall m \in n, m \subseteq n\}.$$

   It is vacuously true that $0 \in J$ since $\emptyset$ is a subset of every $n \in J$. Suppose $n \in J$. Then $\forall m \in n$, we have $m \subseteq n$. Consider $S(n) = n \cup \{n\}$. Note that $n \in S(n)$ and $n \subseteq S(n)$. For $x \in S(n)$ such that $x \neq n$, we must have that $x \in n$. By assumption, $x \subseteq n \subseteq S(n)$. Therefore, $S(n) \in J$, and so J is a successor set. By the Induction Principle, $J = \omega$.

[4] We construct this J and show that it is a successor set. Note that if $J = \omega$, our proof is complete.

3. Let
$$J := \{n \in \omega : n \notin n\}.$$
We have $0 = \emptyset \notin \emptyset$. So $0 \in J$.

Let $n \in J$. Consider $S(n) = n \cup \{n\}$. In particular, note that $n \in S(n)$. Suppose, for contradiction, that $S(n) \in S(n)$. Then $S(n) = n$ or $S(n) \in n$.

$S(n) = n \implies n \in S(n) = n \lightning n \notin n$.

$S(n) \in n \implies S(n) \subseteq n$ by part 2 $\implies n \in n \lightning n \notin n$.

Thus $S(n) \notin S(n)$ and so $S(n) \in J$. So J is a successor set, and so by the Induction Principle, $J = \omega$.

4. It suffices to show that
$$\omega = \{0\} \cup \{n \in \omega : 0 \in n\}.$$
Let J = RHS. We have that $0 \in J$. Suppose $n \in J$ such that $n \neq 0$. Then $0 \in n$. Since $n \subseteq S(n) = n \cup \{n\}$, we have that $0 \in S(n)$. Therefore, $S(n) \in J$. So J is a successor set, and so by the Induction Principle, $J = \omega$ as required.

5. Let
$$J := \{n \in \omega : y \in n \implies S(y) \in n \veebar S(y) = n\}.$$
$0 \in J$ is vacuously true, since there are no $y \in 0$. Suppose $n \in J$. Let $y \in S(n) = n \cup \{n\}$. We have two choices: either $y \in n$ or $y = n$. If $y \in n$, then $S(y) \in n \veebar S(y) = n$, since $n \in J$. We have that

$S(y) \in n \subseteq S(n)$ in which case we are done; and

$(sy) \subseteq n \in S(n)$.

Otherwise, if $y \notin n$, then $y = n$. Then we simply have $S(y) = S(n)$. Thus J is a succesor set and so by the Induction Principle, $J = \omega$.

□

3.1.3 *Well-Orderings*

Definition 10 (Strict Partially Ordered Set)

A ***strict partially ordered set*** *(or* ***strict poset***[5]*) is a set* E *together with* $R \subseteq E^2 = E \times E$ *such that*

1. *(**anti-reflexive**)* $\forall a \in E \quad (a, a) \notin R$;
2. *(**anti-symmetric**)* $\forall a, b \in E \quad (a, b) \in R \wedge (b, a) \in R \implies a = b$; *and*
3. *(**transitivity**)* $\forall a, b, c \in E \quad (a, b), (b, c) \in R \implies (a, c) \in R$.

[5] This is my unofficial terminology

Definition 11 (Strict Totally Ordered Set)

A strict poset is ***total*** *(or* ***linear****) if*

$$\forall a, b \in E \quad (a, b) \in R \veebar (b, a) \in R$$

Definition 12 (Well-Order)

A strict linear order is ***well-ordered*** *if*

$$\forall X \subseteq E(X \neq \emptyset) \quad \exists a \in X \quad \forall b \in X(b \neq a) \quad (a, b) \in R$$

i.e. every nonempty subset of E *has a* ***least element****.*

We shall prove the following next lecture.[6]

[6] Anti-reflexivity and Anti-symmetry were proven in this lecture, but I am moving it to the next for ease of reading.

Proposition (ω is Strictly Well-ordered)

$(\omega, \in)$ *is a strict well-ordering.*

4 Lecture 4 Sep 18th

4.1 Ordinals (Continued 3)

4.1.1 Well-Orderings (Continued)

Lemma (Lemma 11)
Suppose $n \in \omega$. We have

1. $n \subseteq \omega$;
2. $\forall m \in n \quad m \subseteq n$;
3. $n \notin n$;
4. $n = 0 \veebar 0 \in n$; and
5. $y \in n \implies S(y) \in n \veebar S(y) = n$.

Proposition 12 (ω is Strictly Well-Ordered)

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

Proof
By Lemma 11, we have that $\forall n \in \omega, n \notin n$. (**anti-reflexivity** ✔).

$\forall n, m \in \omega$, suppose, for contradiction, that $n \in m$ and $m \in n$. Again, by Lemma 11, we have $n \subseteq m$ and $m \subseteq n$, which implies that $n = m$. Thus, we have $n \in m = n$ and $m \in n = m$, a contradiction to the fact that $n \notin n$ and $m \notin m$ (**anti-symmetry** ✔).

$\forall x, y, z \in \omega$ such that $x \in y$ and $y \in z$, by Lemma 11, $y \in z \implies y \subseteq z \implies x \in z$ (**transitivity** ✔).

To show totality of the relation, let $n \in \omega$. WTS for any $m \in \omega$, either

$$m \in n, \quad m = n, \text{ or } n \in m.$$

Let[1]

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

Case 1: $n = 0$. In this case, we have[2]

$$J = \emptyset \cup \{\emptyset\} \cup \{m \in \omega : 0 \in m\}$$

As a consequence of Lemma 11 (4), we have that $J = \omega$.

[1] We construct J such that J will contain all the possible cases, and use this fact to prove that $J = \omega$ so these 3 cases are the only scenarios that can happen.
[2] Note that $0 = \emptyset$.

Case 2: $n \neq 0$. Again, by Lemma 11 (4), since $n \neq 0$, we must have $0 \in n \subseteq J$ and so $0 \in J$. Now suppose that $m \in J$.

Case 2(a): $m \in n$. Then by Lemma 11 (5), $S(m) \in n$ or $S(m) = n$.
$S(m) \in n \implies S(m) \in J$
$S(m) \in n \implies S(m) \in J$

Case 2(b): $m = n$. Then $S(m) = S(n) = n \cup \{n\}$. And so $n \in S(m)$, which implies $S(m) \in J$.

Case 2(c): $n \in m$ Then since $S(m) = m \cup \{m\}$, we have that $m \in m \subseteq S(m)$. Therefore $S(m) \in J$.

Therefore, J is a sucessor subset of ω. Thus by the Induction Principle, $J = \omega$. (**totality ✔**)

To prove that $\in$ is a well-ordering, suppose $X \subseteq \omega$ is non-empty. Suppose, for contradiction, that X has no $\in$-least element. Now consider

$$J = \{n \in \omega : S(n) \cap X = \emptyset\}$$

Claim: J is a successor set.[3]

[3] Since we want to prove that $\in$ is a well-ordering, we can suppose that there is a non-empty subset of ω that is not empty, and has no $\in$-least element. The core idea here is that, by the construction of J, if $J = \omega$, then all elements of ω would be disjoint from X, forcing X to be the empty set.

By Lemma 11 (4), 0 is the $\in$-least element of ω. If $0 \in X$, then 0 would be $\in$-least in X, contradicting our supposition. Thus $0 \notin X$, And so

$$S(0) \cap X = (0 \cup \{0\}) \cap X = \{0\} \cap X = \emptyset$$

since $0 \notin X$. Thus $0 \in J$.

Suppose $n \in J$. By construction of J, we have $S(n) \cap X = \emptyset$. Observe that

$$S(S(n)) \cap X = (S(n) \cup \{S(n)\}) \cap X.$$

Now if RHS of the above is non-empty (aiming for contradiction), then we may have $S(n) \in X$. Then $S(n)$ would be the $\in$-least element in X, a contradiction. If $m \in S(n)$, we have that $m \notin X$ since $S(n) \cap X = \emptyset$. Thus $SS(n) \cap X = \emptyset$ and so $S(n) \in J$. Therefore, by the Induction Principle, $J = \omega$.

We observe that $\forall n \in \omega$,

$$\emptyset = S(n) \cap X) = (n \cup \{n\}) \cap X$$

$\implies n \notin X$, and so we must have $X = \emptyset$ (**well-ordered ✔**). $\square$

Note
Given $n, m \in \omega$, we often write $n < m$ to mean $n \in m$.

Definition 13 (Ordinals)

An ***ordinal*** *is a set α satisfying:*

1. $x \in \alpha \implies x \subseteq \alpha$;
2. $(\alpha, \in)$ *is a strict well-ordering.*

Example 4.1.1

ω is an ordinal: $\forall n \in \omega$, by Lemma 11, $n \subseteq \omega$, and ω is proven to have a strict well-ordering under $\in$.

Example 4.1.2

Every natural number is an ordinal (*finite ordinals*): by Lemma 11 (2), the first property is satisfied; well-ordering follows from the property of ω.

Let Ord denote the class of all ordinals. We shall show later that Ord is a proper class.

Exercise 4.1.1

Verify that for a set to be an ordinal is a definite condition.

Observe that

$$\forall t (t \in \mathrm{Ord} \leftrightarrow ((x \in t \implies x \subseteq t) \land ((t, \in) \text{ is a strict well-ordering })))$$

where $(x \in t \implies x \subset t)$ is the definite condition

$$\forall x (x \in t \to \forall a (a \in x \to a \in t))$$

and $(t, \in)$ is a strict well-ordering is the definite condition

$$\forall s (s \subseteq t \land s \neq \emptyset \to \exists a (a \in s \to \forall b (b \in s \land b \neq a \to (a, b) \in (\in))))$$

Lemma 13 (Proper Subsets of an Ordinal Are Its Elements)

If $\alpha, \beta \in \mathrm{Ord}$ and $\alpha \subsetneq \beta$, then $\alpha \in \beta$.

Proof

We shall prove that α is the least element in β that is not in α itself.[4]

Let $D := \beta \setminus \alpha = \{x \in \beta : x \notin \alpha\} \subset \beta$ [5]. Since $\alpha \subsetneq \beta$, $D \neq \emptyset$. Since $\beta \in \mathrm{Ord}$, $(\beta, \in)$ has a strict well-ordering, and so D has a least element, d. Note that $d \in \beta$, and since $\beta \in \mathrm{Ord}$, $d \subseteq \beta$.

Claim: $\alpha = d$.[6] WTS $\alpha \subseteq d$. $\forall x \in \alpha$, we have $x, d \in \beta$. Then since $(\beta, \in)$ is a strict well-ordering, we have either

$$x < d, \quad x = d, \quad \text{or } d < x$$

Note that $x \neq d$, otherwise $x = d \in D = \beta \setminus \alpha$.

[7] If $d < x$, then $d \in x$ (by our notation). Now since $\alpha \in \mathrm{Ord}$, $x < \alpha \implies x \in\subseteq$, and so $d \in \alpha$, which is yet another contradiction ($d \in D = \beta \setminus \alpha$).

Thus we must have $x < d$, i.e. $x \in d$. So $\alpha \subseteq d$.

WTS $d \subseteq \alpha$. Suppose not. Then let $x \in d \setminus \alpha$. Then since $d \in D = \beta \setminus \alpha$, we have $x \in \beta \setminus \alpha$, which then contradicts the minimality of d. Therefore, $d = \alpha$ as required. □

[4] We shall construct a subset of $\beta \setminus \alpha$ and show that α is its element.

[5] Exists by Bounded Separation Axiom.

[6] If $\alpha = d$, then α is the said least element.

[7] This is an errorneous proof.

Warning

$$d < x \wedge x < \alpha \underset{\text{transitivity}}{\implies} d < \alpha \implies d \in \alpha$$

This argument is errorneous because we do not yet know if $\alpha \in \beta$.

Proposition 14 (Properties of Ordinals)

1. *Every member of an ordinal is an ordinal.*
2. $\alpha \in \mathrm{Ord} \implies \alpha \notin \alpha$.
3. $\alpha \in \mathrm{Ord} \implies S(\alpha) \in \mathrm{Ord}$.
4. $\alpha, \beta \in \mathrm{Ord} \implies \alpha \cap \beta \in \mathrm{Ord}$.
5. $\alpha, \beta \in \mathrm{Ord} \implies \alpha \in \beta \vee \alpha = \beta \vee \beta \in \alpha$.
6. $E \subseteq \mathrm{Ord}$ *a subset* $\implies (E, \in)$ *is a strict well-ordering.*[8]

Some of proofs of these properties are available in the course notes.

Exercise 4.1.2

Prove Item 3, Item 4, and Item 5 of Proposition 14.

[8] **I think** that such an E need not be an ordinal itself. For example, $E = \{1, 5, 10\} \subset \mathrm{Ord}$, but $4 \in 5$ and $4 \notin E$, and so $5 \in E$ but $5 \nsubseteq E$.

7. Ord *is a proper class.*

Proof

1. Suppose $x \in \beta \in \text{Ord}$. WTS $x \in \text{Ord}$, and we shall show that x satisfies Definition 13.

   Since $\beta \in \text{Ord}$, $x \in \beta \implies x \subseteq \beta$. Thus $(x, \in)$ is a strict well-ordering (through inheriting the property). So it suffices to show that $y \in x \implies y \subseteq x$. So let $y \in x$, and let $t \in x$ [9]. Observe that

   $$t \in y \implies t < y$$
   $$y \in x \implies y < x$$

   and $t, y, x \in \beta \in \text{Ord}$. Therefore, by transitivity, we have $t < y < x \implies t \in x$.

2. Suppose not, i.e. $\alpha \in \alpha$. Then $\alpha \subseteq \alpha \in \text{Ord}$, and so $(\alpha, \in)$ is a strict well-ordering, i.e. $\alpha \notin \alpha$, a contradiction.

6. Suppose $A \subseteq E$ and $A \neq \emptyset$. Let $\alpha \in A$.

   Case 1: $\alpha \cap A = \emptyset$. Then $\forall \beta \in \alpha \implies \beta \notin A$. Therefore α is $\in$-least in A.

   Case 2: $a \cap A \neq \emptyset$. Let $A' = \alpha \cap A \subseteq \alpha$. Since $\alpha \in A \subseteq E \subseteq \text{Ord}$, we have $(\alpha, \in)$ is a strict well-ordering, and so A' has a strict well-ordering as well, and thus it must have a $\in$-least element, x. Then x is the $\in$-least element in A.

7. If Ord is a set, then by Item 6, $(\text{Ord}, \in)$ is a strict well-ordering. Also, by Item 1, every element of Ord is a subset of Ord. Therefore, Ord satisfies Definition 13, and so $\text{Ord} \in \text{Ord}$, which contradicts Item 2. Therefore $\text{Ord} \notin \text{Set}$.

□

[9] To show that $y \subseteq x$, we need to show that $\forall t \in y$, $t \in x$.

5 Lecture 5 Sep 20th

5.1 Ordinals (Continued 4)

Note

If $A, B \in \mathrm{Ord}$, we will write $A < B$ to mean $A \in B$.

Proposition 15 (Properties of Ordinals 2)

1. *If $\alpha \in \mathrm{Ord}$, then $\alpha < S(\alpha)$, and there is nothing in between.*
2. *Let $E \subseteq \mathrm{Ord}$, where $E \neq \emptyset$ is a set, and $\sup E := \cup E$. Then $\sup E \in \mathrm{Ord}$, and it is a least upper bound for E.*[1]
3. *If $E \subseteq \mathrm{Ord}$ is a subset, then there is a least ordinal that is not in E.*

[1] I noted down from the lectures that this is "not necessarily strict", but I do not remember what it means now. **(Clarification required.)**

Perhaps this related to my question; can $E = \sup E$?

This is not necessarily true. If $E \notin \mathrm{Ord}$, then $E \neq \sup E$.

Proof

1. Since $S(\alpha) = \alpha \cup \{\alpha\}$, $\alpha \in S(\alpha)$ and so $\alpha < S(\alpha)$.

   It suffices to show that $\forall x < S(\alpha)$, we have $x \leq \alpha$. Let $x < S(\alpha)$, i.e. $x \in S(\alpha)$. So $x \in \alpha$ or $x = \alpha$, i.e. $x < \alpha$ or $x = \alpha$.

2. By definition, $\forall x \in E \subseteq \mathrm{Ord}$, we have that $x \subseteq \mathrm{Ord}$. Since $\cup E \subseteq E$, we have that $\cup E \subseteq \mathrm{Ord}$ is a subset. Thus by Proposition 14 Item 6, $(\cup E, \in)$ is a strict well-ordering.

   [2] Suppose $\alpha \in \cup E$, then $\exists e \in E$ such that $\alpha \in e \subseteq E \subseteq \mathrm{Ord}$. So e is an ordinal and so $\alpha \subseteq e$. [3] $\forall x \in \alpha$, we have $x \in e \in E$, and so $x \in \cup E$ by definition. Therefore $\alpha \subseteq \cup E$.

   And so, we have shown that $\cup E = \sup E \in \mathrm{Ord}$.

Exercise 5.1.1

Prove Proposition 15 Item 3.

Recommended strategy: $\alpha \in \mathrm{Ord}$ such that $E \subsetneq \alpha$ and take the least element of $\alpha \setminus E$ (which is non-empty). Prove that this least element is the least ordinal that is not in E.

You can take $\alpha = SS(\sup E)$. Verify that this works.

[2] This part shows that $\cup E$ is also an ordinal.

[3] Now we show that $\alpha \subseteq \cup E$.

Claim 1: $\sup E$ is an upper bound for E.

Suppose, for contradiction, that $\exists e \in E$ such that $\sup E < e$. Then since $\sup E$ and e are both ordinals, we have $\sup E \in e \in E$. Then by definition of $\cup$, we have that $\sup E \in \cup E = \sup E$, but by Proposition 14 Item 2, $\sup E \notin \sup E$, a contradiction.

Thus $\sup E$ is an upper bound as claimed.

Claim 2: $\sup E$ is the supremum (least upper bound).

$\forall \alpha < \sup E$, we have that $\alpha \in \sup E = \cup E$, and so $\exists e \in E$ such that $\alpha \in e$. Then $\alpha < e \in E$, i.e. α is not an upper bound of E.

Definition 14 (Successor Ordinal)

*The **successor ordinal** is an ordinal of the form $S(\alpha)$ for some $\alpha \in \mathrm{Ord}$.*

Definition 15 (Limit Ordinal)

*A **limit ordinal** is an ordinal that is not a successor.*

Example 5.1.1

0 and ω are both limit ordinals; 0 is vacuosly a limit ordinal, and ω is not a successor of any $\alpha \in Ord$ [4].

On the other hand, for $n \in \omega$ such that $n \neq 0$, $\exists \cup n \in \omega$ such that $S(\cup n) = n$ [5].

[4] Need a more careful proof, which I cannot do. The idea is to show that any such ordinal α will be an element of ω, and so will its successor $S(\alpha)$, and $\omega \notin \omega$.
[5] See A1.

Exercise 5.1.2

Prove that $S(\omega)$ is a successor ordinal.

Solution
We have that $\omega \in \mathrm{Ord}$, and so $S(\omega)$ is a successor ordinal.

5.1.1 *Transfinite Induction & Recursion*

Theorem 16 (Transfinite Induction Theorem v1)

Suppose P is a definite condition, with the property

$$\forall \alpha \in \text{Ord} \wedge (\forall \beta < \alpha\ P(\beta)) \implies P(\alpha). \tag{5.1}$$

Then P is true of all ordinals.

Proof

$P(0)$ is vacuously true, since there are no elements that are less than 0. Suppose $P(\alpha)$ is false for some $\alpha \in \text{Ord}$ such that $\alpha > 0$. By the Bounded Separation Axiom,

$$D := \{\beta \leq \alpha : \neg P(\beta)\}$$

is a set [6]. Note that $D \neq \emptyset$, since $\alpha \in D$. Since $\alpha \in \text{Ord}$, we have $D \subseteq \alpha \subseteq \text{Ord}$, and so $(D, \in)$ has a strict well-ordering. Let $\alpha_0 \in D$ be $\in$-least. Then $\forall \beta < \alpha_0$, we have that $\neg P(\beta)$, which contradicts the assumption Equation (5.1). Thus $P(\alpha)$ is true for all ordinals. □

[6] Note that $\beta \leq \alpha \iff \beta < \alpha \vee \beta = \alpha \iff \beta \in S(\alpha)$

Theorem 17 (Transfinite Induction Theorem v2)

Suppose P is a definite condition satisfying

1. $P(0)$;
2. $\forall \beta \in \text{Ord}\ P(\beta) \implies P(S(\beta))$; *and*
3. *If $\alpha \in \text{Ord}$ is a limit ordinal and $\forall \beta < \alpha$, $P(\beta)$, then $P(\alpha)$.*

Then P is true of all ordinals.

This statement strongly resembles the Induction Princple that we have learnt in the earlier years of university. In contrast, v1 resembles Strong Induction Principle. It can be shown that v1 $\iff$ v2. v1 $\implies$ v2 is proven in this lecture.

Exercise 5.1.3
Prove that Theorem 17 $\implies$ Theorem 16.

Proof

It suffices to show that P satisfies Equation (5.1), i.e. $\forall \alpha \in \text{Ord}$, we want to prove that $\forall \beta < \alpha$, if $P(\beta)$, then $P(\alpha)$.

When $\alpha = 0$, we have $P(0)$ and so Equation (5.1) is satisfies. When $\alpha > 0$ is a limit ordinal, our assumption immediately satisfies Equation (5.1). Now suppose $\alpha > 0$ is a successor ordinal, and suppose that $\alpha = S(\gamma)$ for some $\gamma \in \text{Ord}$. By the assumption in

Equation (5.1), we have that

$$\forall \beta < \gamma \; P(\beta) \implies P(\gamma)$$

and so by condition (2), we have $P(S(\gamma))$ since $\gamma \in \mathrm{Ord}$. Thus we have $P(\alpha) = P(S(\gamma))$. □

We shall prove the following in the next lecture:

Theorem (Transfinite Recursion)

Let X be a class of all definite operations whose domain is an ordinal. Given a definite operation

$$G : X \to \mathrm{Set}$$

$\exists! F : \mathrm{Ord} \to \mathrm{Set}$, *a definite operation, such that $F(\alpha = F(F \restriction_\alpha))$, for all $\alpha \in \mathrm{Ord}$.*

We want to use Transfinite Recursion to construct definite operations on ordinals such that they have properties that we are familiar with (and hence desire).

Note (Notation - Restriction)

Let $H : U \to Y$ be a definite operation on classes U, Y, and $Z \subseteq U$ a subclass. $H \restriction_Z$ is the definite operation

$$H \restriction_Z : Z \to Y$$

obtained by restricting H onto Z.

Note

In the theorem, we stated that F has its domain on Ord*. We know that for $\alpha \in \mathrm{Ord}$, $\alpha \subseteq \mathrm{Ord}$, and so $F \restriction_\alpha$ makes sense; in particular,*

$$F \restriction_\alpha : \alpha \to \mathrm{Set}.$$

Note that $F \restriction_\alpha \in X$, and so $G(F \restriction_\alpha)$ is valid and makes sense.

6 Lecture 6 Sep 25th

6.1 Ordinals (Continued 5)

6.1.1 Transfinite Induction & Recursion (Continued)

Theorem 18 (Transfinite Recursion v1)

Let X be a class of all definite operations whose domain is an ordinal. Given a definite operation

$$G : X \to \text{Set}$$

$\exists! F : \text{Ord} \to \text{Set}$, *a definite operation, such that* $F(\alpha) = G(F \restriction_\alpha))$, *for all* $\alpha \in \text{Ord}$.

Before proving the theorem, we shall note the following definition.

Definition 16 (α-function)

*Using definitions in Theorem 18, a function t with domain in the ordinals is called an **α-function** defined by G if*

$$\forall \beta < \alpha \quad t(\beta) = G(t \restriction_\beta).$$

Proof

We shall first prove for **uniqueness**. Suppose F and F' are two definition operations such that

$$F : \text{Ord} \to \text{Set} \quad F' : \text{Ord} \to \text{Set}$$
$$F(\alpha) = G(F \restriction_\alpha) \quad F'(\alpha) = G(F' \restriction_\alpha)$$

[1]Suppose $\forall \beta < \alpha \in \text{Ord}$, we have $F(\beta) = F'(\beta)$. Note that

$$F(\beta) = F'(\beta) \iff F \restriction_\alpha = F' \restriction_\alpha$$
$$\implies F(\alpha) = G(F \restriction_\alpha) = G(F' \restriction_\alpha) = F'(\alpha)$$

[1] Here, we want to use Transfinite Induction v1 to show that they are unique.

Thus, uniqueness of F is guaranteed.

To prove existence, firstly, we note that the α-functions defined in Definition 16 are approximations to the F that we want. However, before going further, we need to show that they are also unique and that we can form a chain of extensions on these functions over Ord.

Uniqueness of t_α Let t, t' be a α-functions defined by G. WTS $\forall \beta < \alpha$, $t(\beta) = t'(\beta)$. If an α-function defined by G exists, we shall denote it as t_α.

Consider the definite condition

$$P(x) := (x \geq \alpha) \vee (t(x) = t'(x)).$$

Suppose that $\forall \gamma < \beta$, $P(\gamma)$ holds, i.e. $\gamma \geq \alpha$ or $t(\gamma) = t'(\gamma)$, which implies that $t \restriction_\beta = t' \restriction_\beta$. Therefore $t(\beta) = t'(\beta)$, i.e. $P(\beta)$ holds. Thus t_α is unique if it exists by Transfinite Induction.

t_α as a chain of extensions Now $\forall \beta < \alpha \in \text{Ord}$, we have that $\beta \subseteq \alpha$. If t_α and t_β exist, then

$$\Gamma(t_\beta) \subseteq \Gamma(t_\alpha),$$

or in other words

$$t_\alpha \restriction_\beta = t_\beta.$$

We shall denote this relation as $t_\beta \subseteq t_\alpha$.

Existence of t_α The existence of t_α is a definite condition: by the Replacement Axiom, a function that maps $\alpha \mapsto t_\alpha$ is definite, and by Bounded Separation Axiom, the set

$$\Gamma(t_\alpha) = \{(\beta, G(t_\alpha \restriction_\beta) \mid \beta < \alpha, t_\alpha(\beta) = G(t_\alpha \restriction_\beta)\} \subseteq \text{Ord} \times \text{Set}$$

exists. [2]Now $t_0 = t_\emptyset$ is vacuously true. Suppose for any successor ordinal α, t_α exists. Since $\alpha \in \text{Ord}$, we have that there is nothing between α and its successor $S(\alpha)$, and $\alpha < S(\alpha)$. Thus

$$t_{S(\alpha)} = t_\alpha \cup \{(\alpha, G(t_\alpha))\},$$

which exists by Union Set Axiom. We see that $t_{S(\alpha)}$ extends t_α onto α itself.

Suppose $\alpha > 0$ is a limit ordinal. Since t_α is a definite condition by Replacement, by Bounded Separation, we have that[3]

$$t_\alpha = \bigcup_{\beta<\alpha} t_\beta = \cup\{t_\beta : \beta < \alpha\}.$$

Note that t_α is, indeed, an α-function defined by G:

- t_α is a function on α: we have that $\forall \beta < \alpha$,the t_β's form a chain of extensions;
- t_α is an α-function defined by G: $\forall \beta < \alpha$, since α is a limit ordinal, $S(\beta) < \alpha$, and so

$$t_\alpha(\beta) = t_{S(\beta)}(\beta) = G(t_{S(\beta)} \restriction_\alpha) = G(t_\beta \restriction_\alpha).$$

And so by Transfinite Induction, $\forall \alpha \in Ord$, t_α exists as required.

Construction of F Now for any $\beta < \alpha$, we have a chain of extensions

$$t_0 \subseteq t_1 \subseteq t_2 \subseteq \ldots \subseteq t_\beta \subseteq t_\alpha \subseteq \ldots$$

Let[4]

$$F := \bigcup_{\alpha \in \text{Ord}} t_\alpha = \cup [\![\ t_\alpha \mid \alpha \in \text{Ord} \ ]\!]$$

[2] We use Transfinite Induction v2 to CTP.

[3] **Verify the motivation in using or the reason behind getting this.**

[4] I will leave the proof unfinished here. **Need to verify my understanding**

Corollary 19 (Transfinite Recursion v2)

Given $G_1 \in \text{Set}$, $G_2 : \text{Set} \to \text{Set}$ *a definite operation,* $G_3 : X \to \text{Set}$ *a definite operation, where X is the class of all definite opeartion whose domain is an ordinal. Then*

$$\exists! F : \text{Ord} \to \text{Set}$$

such that

1. $F(0) = G_1$;
2. $\forall \alpha \in \mathrm{Ord} \quad F(S(\alpha)) = G_2(F(\alpha))$; and
3. $\forall \beta > 0$ a limit ordinal, $F(\beta) = G_3(F \restriction_\beta)$.

Proof

The result is clear by Theorem 16 with $G : X \to Set$ defined by

$$G(f) = \begin{cases} G_1 & f = \emptyset \\ G_2(f(\alpha)) & \mathrm{Dom}(f) = S(\alpha) \\ G_3(f) & \mathrm{Dom}(f) > 0 \text{ a limit ordinal} \end{cases}$$

□

6.1.2 Ordinal Arithmetic

6.1.2.1 Ordinal Addition

Definition 17 (Ordinal Addition)

Let $\beta \in \mathrm{Ord}$. For any $\alpha \in \mathrm{Ord}$, we define

$$\beta + \alpha$$

using Transfinite Recursion[5] *on α as follows:*

- $\beta + 0 := \beta$;
- *if α is a successor ordinal, then* $\beta + S(\alpha) := S(\beta + \alpha)$; *and*
- *if $\alpha > 0$ is a limit ordinal, then* $\beta + \alpha := \sup\{\beta + \gamma : \gamma < \alpha\}$.

[5] Note that we are using Corollary 19 with

$G_1 = \beta$
$G_2 = S : \mathrm{Set} \to \mathrm{Set}$
$G_3 : X \to \mathrm{Set}$ by $G_3(f) = \sup \mathrm{Img}(f)$

Exercise 6.1.1

Using both the Induction Principle and Transfinite Induction, prove that $\beta + \alpha \in \mathrm{Ord}$.

Example 6.1.1

We have that

$$0, 1, 2, ..., \omega, \omega + 1, \omega + 2, ..., \omega + \omega$$

Observe that

$$\omega + 1 = \omega + S(0) = S(\omega + 0) = S(\omega)$$

and so $\omega + 1$ is a successor to ω. In general, we have that $\forall \alpha \in \mathrm{Ord}$, $\alpha + 1 = S(\alpha)$. For example,

$$\omega + 2 = \omega + S(1) = S(\omega + 1).$$

On the other hand, note that

$$\omega + \omega = \sup\{\omega + n : n \in \omega\}.$$

Unlike regular addition, ordinal addition is not commutative. For instance, while $\omega + 1 = S(\omega)$,

$$1 + \omega = \sup\{1 + n : n \in \omega\} = \omega.$$

Exercise 6.1.2

Prove that ordinal addition is only commutative for "finite" ordinals.

6.1.2.2 *Ordinal Multiplication*

Definition 18 (Ordinal Multiplication)

Let $\beta \in \mathrm{Ord}$. For any $\alpha \in Ord$, we define

$$\beta \cdot \alpha$$

using Transfinite Recusion[6] as follows:

- $\beta \cdot 0 := 0$;
- *if α is a successor ordinal, $\beta \cdot S(\alpha) := \beta\alpha + \beta$*
- *if $\alpha > 0$ is a limit ordinal, $\beta \cdot \alpha := \sup\{\beta \cdot \gamma : \gamma < \alpha\}$.*

[6] Here, we use

$$G_1 = 0$$
$$G_2 : \mathrm{Set} \to \mathrm{Set} \text{ by } G_2(x) = x + \beta$$
$$G_3(f) = \sup \mathrm{Img}(f)$$

Example 6.1.2

We have

$$\omega \cdot 1 = \omega \cdot S(0) = \omega \cdot 0 + \omega = 0 + \omega = \omega.$$

Exercise 6.1.3

Prove that in general, $\forall \beta \in \mathrm{Ord}$, *we have* $\beta \cdot 1 = \beta$.

Exercise 6.1.4

Prove that $\forall \alpha, \beta \in \mathrm{Ord}$, $\alpha \cdot \beta \in \mathrm{Ord}$.

Example 6.1.3

We have

$$\omega \cdot 2 = \omega \cdot S(1) = \omega \cdot 1 + \omega = \omega + \omega.$$

Exercise 6.1.5

Prove that in general, $\forall \beta \in \mathrm{Ord}$, *we have* $\beta \cdot 2 = \beta + \beta$.

Note that ordinal multiplication, like its addition counterpart, is not necessarily commutative.

Example 6.1.4

While we have

$$1 \cdot \omega = \sup\{1 \cdot n \mid n \in \omega\} = \omega = \omega \cdot 1,$$

observe that

$$2 \cdot \omega = \sup\{2 \cdot n \mid n \in \omega\} = \omega \neq \omega + \omega = \omega \cdot 2.$$

Proposition 20 (Properties of Ordinal Addition and Ordinal Multiplication)

Let $\alpha, \beta, \delta \in \mathrm{Ord}$.

- $\alpha < \beta \iff \delta + \alpha < \delta + \beta$;
- $\alpha = \beta \iff \delta + \alpha = \delta + \beta$;
- (***(associativity)***) $(\alpha + \beta) + \delta = \alpha + (\beta + \delta)$;
- *if* $\delta \neq 0$, *then* $\alpha < \beta \iff \delta\alpha < \delta\beta$;
- *if* $\delta \neq 0$, *then* $\alpha = \beta \iff \delta\alpha = \delta\beta$;
- $(\alpha\beta)\delta = \alpha(\beta\delta)$.

Exercise 6.1.6
Prove Proposition 20.

6.1.2.3 Ordinal Exponentiation

Definition 19 (Ordinal Exponentiation)

Let $\beta \in \mathrm{Ord}$. For any $\alpha \in \mathrm{Ord}$, define

$$\beta^\alpha$$

using Transfinite Recursion by

- $\beta^0 := 1$;
- *if α is a successor ordinal, then $\beta^{S(\alpha)} := \beta^\alpha \cdot \beta$;*
- *if $\alpha > 0$ is a limit ordinal, then $\beta^\alpha := \sup\{\beta^\gamma \mid \gamma < \alpha\}$.*

In the next lecture, we shall study the following theorem:

Theorem (Strict Well-Ordered Sets are Isomorphic to a Unique Ordinal)
Every strict well-ordering is isomorphic to an ordinal. Both the ordinal and the isomorphism are unique.

The definition of isomorphism is:

Definition 20 (Isomorphism)

*Let E, F be sets and R, S be relations defined on each set respectively so. We say that (E, R) and (F, S) are **isomorphic**, which we denote by*

$$(E, R) \simeq (F, S),$$

if $\exists f : E \to F$, a bijection, such that

$$e_1 R e_2 \iff f_1 S f_2.$$

*Such an f is called an **isomorphism**.*

7 Lecture 7 Sep 27th

7.1 Ordinals (Continued 6)

7.1.1 Well-Orderings and Ordinals

Before proving the theorem stated at the end of last lecture, we require the following 2 lemmas:

Lemma 21 (Rigidity of Well-Orderings)

Well-orderings are rigid, i.e. the only automorphism[1] is the identity.

[1] An **automorphism** is an isomorphism from a set to itself.

Proof

Suppose $(E, <)$ is a well-ordering, and $f : E \to E$ an automorphism. Let[2]

$$D = \{x \in E \mid f(x) \neq x\}.$$

[2] We look at the fellas that were 'moved'.

Suppose for contradiction that f is not the identity map, i.e. $D \neq \emptyset$. Then, since $D \subseteq E$, D has a well-ordering, and so we can pick $\alpha \in D$ to be the least element.

Case 1: $f(a) < a$. Then $f(a) \notin D$ since a is least. But then that would mean

$$f(f(a)) = f(a) \implies f(a) = a$$

since f is a bijection. This contradicts the choice that $a \in D$.

Case 2: $a < f(a)$. Since f is a bijection, its inverse exists, and so

$$\begin{aligned} f^{-1}(a) < a &\implies f^{-1}(a) \notin D \\ &\iff ff^{-1}(a) = f^{-1}(a) \iff a = f^{-1}(a) \end{aligned}$$

which contradicts the choice that $a \in D$, yet again. Thus there is no such element $a \in D$, forcing $D = \emptyset$, and so f must be the identity map. □

Lemma 22 (Strict Well-Ordering $\not\simeq$ Any of Its Proper Initial Segment)

A strict well-ordering is not isomorphic to any proper initial segment[3] of itself.

[3] **Definition 21 (Initial Segment)** *For a strict well-order $(E, <)$, an* ***initial segment*** *is a subset of the form*

$$\{x \in E : x < b\}$$

for some $b \in E$.

Proof
Let $(E, <)$ be a strict well-ordering. $\forall b \in E$,

$$I_b := \{x \in E : x < b\}$$

has an induced well-order, in particular $(I_b, <)$.

Suppose for contradiction that there exists an isomorphism $f : E \to I_b$. Let

$$D = \{x \in E \mid f(x) \neq x\}$$

$b \notin I_b$ and so $b \notin D$. **Proof is left incomplete until I verify the proof with the prof.**

Theorem 23 (Strict Well-Ordered Sets are Isomorphic to a Unique Ordinal)

Every strict well-ordering is isomorphic to an ordinal. Both the ordinal and the isomorphism are unique.

Proof
Uniqueness

Let (E, R) be a strict well-ordering. Suppose that

$$(E, R) \simeq (\alpha, \in) \text{ and } (E, R) \simeq (\beta, \in) \tag{7.1}$$

where $\alpha, \beta \in \mathrm{Ord}$. Then, either

$$\alpha < \beta, \quad \alpha = \beta \quad \text{or} \quad \beta < \alpha.$$

Suppose $\alpha < \beta$ (this argument also works for $\beta < \alpha$, by simply swapping the inequality on α and β). Then α is a proper initial segment of β. From Equation (7.1), we have that $(\alpha, \in) \simeq (\beta, \in)$ via (E, R), but this contradicts Lemma 22. Thus we must have $\alpha = \beta$.

Existence

If $E = \emptyset$, then we can choose $0 \in \mathrm{Ord}$ to be the ordinal of which E is isomorphic to. Suppose $E \neq \emptyset$. Denote an initial segment of E by

$$I_x := \{y \in E \mid y < x\}.$$

Let

$$A = \{x \in E \mid \exists \beta \in \mathrm{Ord} \ \ (I_x, R) \simeq (\beta, \in)\}.$$

Notice that $(I_x, R) \simeq (\beta, \in)$ is a definite condition: the both I_x and β are sets, and so by Replacement, there is a graph, $\Gamma(f)$, from I_x to β; injectivity of an element in $\Gamma(f)$ is expressible as

$$\forall y_1 \forall y_2 \forall \beta_1 \forall \beta_2 (y_1, y_2 \in I_x \land \beta_1, \beta_2 \in \mathrm{Ord}((y_1, \beta_1) = (y_2, \beta_2) \leftrightarrow y_1 = y_2));$$

surjectivity is expressible as

$$\forall \beta (\beta \in \alpha \to \exists y (y \in I_x \to (y, \beta) \in \Gamma(f))).$$

Thus by Bounded Separation, A is a set. Also, A is nonempty, since the least element of E will be isomorphic to 0.

By our uniqueness proof above, let f be a function on A, where $f(x)$ is the unique ordinal that is isomorphic to (I_x, R). By Replacement,

$$\mathrm{Img}(f) = \{f(x) \in \mathrm{Ord} \mid x \in A\} \subset \mathrm{Ord}$$

is a set. By Proposition 15 Item 3, $\exists \alpha \in \mathrm{Ord} \setminus \mathrm{Img}(f)$ that is $\in$-least. We want to show that $f : A \to \mathrm{Img}(f)$ is an isomorphism between (E, R) and $(\alpha, \in)$.[4]

[4] This is why we need to show that

1. f is order-preserving, which is one of the requirements of an isomorphism (by our definition in Definition 20);
2. f is injective;
3. $\alpha = \mathrm{Img}(f)$;
4. $A = E$,

where the last 2 items will force f to be surjective.

<u>f is order-preserving</u>: We shall also show here that A is **downward closed**. $\forall x, y \in E$, we want to show that $xRy \land y \in A \implies x \in A$. By the assumption, we have that

$$f(x) \overset{h}{\simeq} I_X \subsetneq I_y \overset{h}{\simeq} f(y)$$

Since $f(x), f(y) \in \mathrm{Ord}$, we have either

$$f(x) < f(y), \quad f(x) = f(y), \quad \text{or} \quad f(y) < f(x).$$

If $f(x) = f(y)$, then $I_x \simeq I_y$ which contradicts Lemma 22. If $f(y) < f(x)$, then

$$h(I_x) \subseteq h(I_y) = f(y) < f(x) = h(I_x),$$

which is a contradiction. Thus we must have $f(x) < f(y)$, i.e. f preserves order as claimed, and $f(x)$ is an initial segment of $f(y)$, i.e. $f(x) \in \mathrm{Ord}$, which implies that $x \in A$.

$\underline{\alpha = \mathrm{Img}(f)}$: Suppose $\beta \in \alpha \implies \beta < \alpha \implies \beta \in \mathrm{Img}(f)$ by choice of α being the least. Therefore, $\alpha \in \mathrm{Img}(f)$.

Now suppose that $\beta \in \mathrm{Img}(f)$. Then $\exists x \in A$ such that $(I_x, R) \simeq (\beta, \in)$. Again, we have 3 possibilities; either

$$\alpha < \beta, \quad \alpha = \beta \quad \text{or} \quad \beta < \alpha.$$

Now $\alpha < \beta \implies \alpha \in \beta \implies \alpha = h(I_y)$ where yRx and $I_y \subset I_X$. Since A is downward closed, $\alpha \in \mathrm{Img}(f)$, a contradiction. We also have that $\alpha \neq \beta$ since $\beta \in \mathrm{Img}(f)$ and $\alpha \in \mathrm{Ord} \setminus \mathrm{Img}(f)$, i.e. $\alpha \notin \mathrm{Img}(f)$. Thus $\beta < \alpha$, and so $\beta \in \alpha$. Therefore $\alpha = \mathrm{Img}(f)$.

$\underline{f \text{ is injective}}$: Suppose that $f(x) = f(y)$. $xRy \implies I_x$ is an initial segment of I_y, which contradicts Lemma 22. The argument is similar for if yRx. Thus we must have $x = y$.

$\underline{A = E}$: It suffices to show that $E \setminus A = \emptyset$. Suppose for contradiction that $E \setminus A \neq \emptyset$, i.e. $E = A$. Then $\exists x \in E \setminus A$, since $E \setminus A \subset E$ which has a strict well-ordering. Since f preserves order, for any $y \in E$, $xRy \implies y \notin A$. On the other hand, $yRx \implies y \in A$. Thus $I_x = A$. However, since f is an isomorphism between (A, R) and $(\alpha, \in)$ (since we assume that $A = E$), we have that $I_x = A \simeq \alpha$, i.e. $x \in A$, which is a contradiction to the choice that $x \in E \setminus A$.

This completes the proof. $\square$

7.2 Cardinals

While ordinals allow us to enumerate, we cannot use it to "measure". For example, $\omega \neq \omega + 1$, but they have the same size.

Definition 22 (Equinumerous)

*Two sets A and B have the same size, or **equinumerous**, if there is a bijection from A to B. We denote this relation by $|A| = |B|$.*[5]

[5] Note that we have yet to define $|\cdot|$.

The following is a well-known theorem that makes proving equinumerosity a lot easier.

Lemma 24 (Schröder-Bernstein Theorem)

Given two sets A and B, $|A| = |B|$ if and only if there exists injections in both directions, i.e. an injection from A to B, and an injection from B to A.

Proof

The ($\implies$) direction is easy, since a bijection exists. So it suffices to show the ($\impliedby$) direction. We shall use $A \hookrightarrow B$ to say that there is an injection from A to B.

Suppose

$$A \hookrightarrow B \overset{g}{\hookrightarrow} A$$

Then $\exists f : A \to A$ an injective map, and we would have

$$f(A) \subseteq g(B) \subseteq A. \tag{7.2}$$

From here, it suffices to show that for an injective ap $f : X \to X$, if we have

$$f(X) \subseteq Y \subseteq X,$$

then $|Y| = |X|$. From our observation in Equation (7.2), we have

$$X \supseteq Y \supseteq f(X) \supseteq f(Y) \supseteq f^2(X) \supseteq f^2(Y) \supseteq f^3(X) \supseteq \dots$$

Let

$$Z = X \setminus Y \;\dot{\cup}\; f(X) \setminus f(Y) \;\dot{\cup}\; f^2(X) \setminus f^2(Y) \;\dot{\cup}\; \ldots$$

and

$$W = X \setminus Z$$

Then

$$X = Z \;\dot{\cup}\; W$$

Claim: For sets A and B such that $B \subseteq A$, we have that[6]

$$f(A \setminus B) = f(A) \setminus f(B).$$

[6] Forgive me if the proof of this claim is a little sloppy.

Note that since f is injective, $f : B \to f(B)$ is a bijective map. Suppose $\exists x \in A \setminus B$ such that $f(x) \in f(B)$. Since $f : B \to f(B)$ is bijective, $\exists b \in B$ such that $f(b) = f(x)$, but f is injective. Thus the claim is true.

Using a similar argument, it can be shown that $f(A \;\dot{\cup}\; B) = f(A) \;\dot{\cup}\; f(B)$.

Observe that

$$\begin{aligned} f(Z) &= f(X \setminus Y) \;\dot{\cup}\; f(f(X) \setminus f(Y)) \;\dot{\cup}\; f(f^2(X) \setminus f^2(Y)) \;\dot{\cup}\; \ldots \\ &= f(X) \setminus f(Y) \;\dot{\cup}\; f^2(X) \setminus f^2(Y) \;\dot{\cup}\; f^3(X) \setminus f^3(Y) \;\dot{\cup}\; \ldots \end{aligned}$$

and note that

$$f(Z) = Z \setminus (X \setminus Y).$$

Since $W = X \setminus Z$, we have $W \subseteq Y$. Since $Z \cap W = \emptyset$, we still have $f(Z) \cap W = \emptyset$. Also, note that $(X \setminus Y) \cap W = \emptyset$. Thus, we have

$$Y = X \setminus (X \setminus Y) = (Z \;\dot{\cup}\; W) \setminus (X \setminus Y) = f(Z) \;\dot{\cup}\; W$$

Let $g : X \to Y$ such that

$$g(A) = \begin{cases} A & A \subseteq W \\ f(A) & A \subseteq Z \end{cases}$$

Clearly so, g is bijective. □

Example 7.2.1

We claimed that $|\omega| = |\omega + 1|$.

We can simply use the identity map from $\omega \to \omega + 1$. For $\omega + 1 \to \omega$, consider the mapping

$$f(\alpha) = \begin{cases} S(\alpha) & \alpha \in \omega \\ 0 & \alpha = \omega \end{cases}$$

This map is clearly injective by properties of elements of ω.

Definition 23 (Cardinal)

*A **cardinal** is an ordinal α with the property that $\forall \beta < \alpha, |\alpha| \neq |\beta|$*

In the next lecture, we shall see that the collection of cardinals is a proper class, and is a subclass of the ordinals.

Definition 24 (Finite & Countable)

*A set A is **finite** if $|A| = |n|$ for some $n \in \omega$. A is **countable** if A is finite or $|A| = |\omega|$.*

8 Lecture 8 Oct 02nd

8.1 Cardinals (Continued)

Note

If $\kappa \in \text{Card}$ *and* κ *is infinite, then* κ *is a limit ordinal. In other words, successor ordinals are either finite or are not Cardinals. This is true since* $\forall \alpha \in \text{Ord}$ *such that* $\alpha \geq \omega$*, clearly we have* $\alpha \hookrightarrow S(\alpha)$*, and we can define a function* $f : S(\alpha) \to \alpha$ *such that*

$$f(\beta) = \begin{cases} S(\beta) & \beta < \omega \\ \beta & \omega \leq \beta < \alpha \\ 0 & \beta = \alpha \end{cases},$$

which is injective, and so by Schröder-Bernstein, $|S(\alpha)| = |\alpha|$.

Proposition 25 (The Least Cardinality Not Equinumerous to Subsets of a Set)

$\forall E \in \text{Set} \; \exists \alpha \in \text{Ord} \; \forall e \subseteq E$

$$|e| \neq |\alpha|$$

and there exists a least such $h(E) \in \text{Ord}$.

Note

Note that $h(E) \in \text{Card}$.

Proof

Suppose not, i.e. $\exists \alpha \in Ord$ such that $|h(E)| = |\alpha|$ and $\alpha < h(E)$. Since $h(E)$ is the least, we must have $|\alpha| = |A|$ or some $A \subseteq E$. Then $|h(E)| = |A|$, which is a contradiction to the definition of $h(E)$. □

In particular, we have that $h(\omega)$ is an uncountable cardinal.

Now to prove Proposition 25.

Proof

It suffices to prove the existence of $h(E)$. Consider the class

$$H = [\![\ \alpha \in \mathrm{Ord} \mid \exists e \subseteq E \ |e| = |\alpha| \ ]\!] .$$

If we can show that H is a set, then the least element not in H shall be our $h(E)$. Consider

$$W := \{(A, R) \mid A \subseteq E, \ (A, R) \text{ is a strict well-ordering } \}$$

which is a set by Replacement, i.e.

$$W \subseteq \mathcal{P}(E) \times \mathcal{P}(E \times E).$$

Note that $W \neq \emptyset$, since $\emptyset \subseteq E$ and the empty relation would be a well-ordering of $\emptyset$. By Theorem 23, $\exists f : W \to \mathrm{Ord}$ such that $f(A, R)$ is a unique ordinal. By Replacement, $\mathrm{Img}(f) \subseteq \mathrm{Ord}$ is a set.

Claim: $\mathrm{Img}(f) = H$: It is clear that $\mathrm{Img}(f) \subseteq H$, since all elements of $\mathrm{Img}(f)$ are isomorphic to some subset of E by definition. Now let $\alpha \in H$. Then $\exists g : \alpha \to A$ a bijection, for some $A \subseteq E$. Then define $\prec$ on A by

$$a \prec b \iff g^{-1}(a) < g^{-1}(b).$$

Then $(\alpha, <) \overset{g}{\underset{\simeq}{\to}} (A, \prec)$. Therefore, $(A, \prec) \in W$ and $f(A, \prec) = (\alpha, <)$, i.e. $H \subseteq \mathrm{Img}(f)$. Thus $H = \mathrm{Img}(f)$ and so H is a set as required. □

Remark
This revelation tells us that Card *is a proper class.*

To use Card to measure all sets, we need every set to be equinumerous with a cardinal. In particular, every set would then have to be equinumerous to an ordinal. This would then require evrey set to have a strict well-ordering, which is something that we cannot prove with our axioms thus far.

8.1.1 Axiom of Choice

Definition 25 (Choice Function)

*Suppose $\mathcal{F}$ is a set. A **choice function** on $\mathcal{F}$ is a function*

$$c : \mathcal{F} \to \cup\mathcal{F} \text{ such that } \forall F \in \mathcal{F} \;\; c(F) \in F.$$

Note
If $\emptyset \in \mathcal{F}$, then $\mathcal{F}$ has no choice function, since nothing belongs in $\emptyset$.

Axiom 26 (Axiom of Choice)

Every $\mathcal{F} \in \mathrm{Set}$ such that $\emptyset \notin \mathcal{F}$ admits a choice function.[1]

[1] This is, again, an existential axiom.

Note
Unlike the other axioms, while the Axiom of Choice asserts the existence of choice functions on sets, the choice function need not be unique. Recall that in other axioms, the sets of which we assert their existence are unique.

Theorem 27 (Axiom of Choice and Its Equivalents)

TFAE

1. *Axiom of Choice*
2. *Well-ordering Principle: Every set admits a well-ordering.*
3. *Zorn's Lemma: If (E, R) is a strict poset with the property that* ***every totally ordered subset of*** E ***has an upper bound****, i.e.*

$$\forall A \subset E (\forall a, b \in A \; aRb \lor bRa \lor a = b) \; \forall a \in A \exists e \in E (aRe \lor a = e).$$

Then (E, R) has a maximal elements, i.e. $\exists z \in E$ such that $\forall x \in E$, $\neg zRx$.

Proof

(1) $\implies$ (2): Let $A \in \mathrm{Set}$. If $A = \emptyset$, then there is nothing to do and the statement is vacuously true. So suppose $A \neq \emptyset$. By the asusmption, fix a choice function c on $\mathcal{F} := \mathcal{P}(A) \setminus \{\emptyset\}$. Let $\theta \in \mathrm{Ord} \setminus A$. Define a definite operation $F : \mathrm{Ord} \to \mathrm{Set}$ such that

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

Note that F exists by Transfinite Recursion[2].

[2] I am not sure how.

Claim 1: F halts, i.e. $\exists \alpha \in \mathrm{Ord}$ such that $F(\alpha) = \theta$ and $\forall \beta \in \mathrm{Ord}$ such that $\alpha < \beta$, $F(\beta) = \theta$.

Suppose not. Then F must be injective and has codomain $\cup \mathcal{F} = \cup(\mathcal{P}(A) \setminus \{\emptyset\}) = A$, i.e. we have that $F : \mathrm{Ord} \to A$ injective. Then by Proposition 25, there exists $h(A)$ that is the least ordinal that is not equinumerous with any subset of A. We may consider $F \restriction_{h(A)} : h(A) \to A$ since $h(A) \subset \mathrm{Ord}$. Now $\forall \alpha < \beta \in h(A)$, by our supposition, $F(\beta) \neq \theta$, and so $F(\beta) = c(A \setminus \mathrm{Img}(F \restriction_\beta)) \in (A \setminus \mathrm{Img}(F \restriction_\beta))$. Thus $F(\beta) \notin \mathrm{Img}(F \restriction_\beta)$. But since $\alpha < \beta$, it must be that $F(\alpha) \in \mathrm{Img}(F \restriction_\beta)$. Then $F(\alpha) \neq F(\beta)$. Consequently, since $\forall \beta \in h(A)$, we have that $F(\beta) \neq \theta$, and so we created an injection from $h(A)$ to A. This is impossible by the definition of $h(A)$. Thus it must be the case that $\exists \beta \in h(A)$ such that $F(\beta) = \theta$.

Let α be the least such β. The previous paragraph showed that $F \restriction_\alpha$ is an injection from α to A. It remains to show that the map is surjective. Suppose $\mathrm{Img}(F\restriction_\alpha) \neq A$. Then $A \setminus \mathrm{Img}(F\restriction_\alpha) \neq \emptyset$. Then

$$F(\alpha) = c(A \setminus \mathrm{Img}(F\restriction_\alpha)) \in A$$

which is a contradiction since $F(\alpha) = \theta$. So $A = \mathrm{Img}(F\restriction_\alpha)$ and so $F \restriction \alpha$ is surjective.

Therefore $F\restriction_\alpha$ is a bijection from α to A, and so A has an induced strict well-ordering from α.

(3) $\implies$ (1): Let $\mathcal{F}$ be a set such that $\emptyset \notin \mathcal{F}$. Let Λ be the set of all partial choice functions on $\mathcal{F}$, identified with their graphs, i.e. $\forall f \in \Lambda$, $f : \mathcal{G} \to \cup\mathcal{G}$ such that $f(G) \in G$ for all $G \in \mathcal{G}$, and $\mathcal{G} \subseteq \mathcal{F}$. Note that Λ is indeed a set since the graphs exist by Replacement, and Λ is therefore a set from Bounded Separation. $\Lambda \neq \emptyset$, since the function $f(F) = x$ exists for $F \in \mathcal{F} \neq \emptyset$, and $F \neq \emptyset$.

Now $(\Lambda, \subseteq)$ is a poset, where we order the functions by extensions. For every Θ that is a totally ordered subset of Λ, we have that the union, $\cup\Theta$ is the upper bound of Θ. Thus the assumptions for Zorn's Lemma are satisfied, and so there exists a maximal function

$$f : \mathcal{G} \to \cup\mathcal{G} \text{ in } \Lambda.$$

To prove that this f is a choice function on $\mathcal{F}$, we want to prove that $f : \mathcal{F} \to \cup\mathcal{F}$, i.e. we need to show that $\mathrm{Dom}(f) = \mathcal{F}$. Suppose not, i.e. $\exists F \in \mathcal{F}$ such that $F \notin \mathrm{Dom}(f)$. Then since $F \neq \emptyset$, $\exists x \in F$, and so $f \cup \{(F, x)\}$ is a larger partial choice function on $\mathcal{F}$, contradicting the maximality of f in Λ. Thus $\mathrm{Dom}(f) = \mathcal{F}$ as claimed, and so f is a choice function on $\mathcal{F}$, proving the Axiom of Choice.

(2) $\implies$ (3): Suppose (E, R) is a strict poset. By (2), let $<$ be a strict well-ordering on E. Now by ☕ Theorem 23, $\exists! \alpha \in \mathrm{Ord}$ such that $(E, R) \simeq (\alpha, \in)$ through a unique isomorphism. Therefore, we may assume that $E \in \mathrm{Ord}$. Suppose that (E, R) satisfies the assumptions of Zorn's Lemma.

Assume, to the contrary, that (E, R) does not have an R-maximal element. From 💧 Proposition 25, $\exists h(E) \in Ord$ such that $\forall \beta < h(E)$, $\forall A \subseteq E$, $|\beta| \neq |A|$[3]. Recursively so, define $F : h(E) \to E$ by

$$F(0) = e \text{ for some } e \in E$$
$$F(S(\beta)) = < \text{-least element } \gamma \text{ of } E \text{ such that } F(\beta) R \gamma$$

[3] The strategy here is to use, once again, 💧 Proposition 25 to arrive at a contradiction that is similar to when we were proving (1) $\implies$ (2).

and for $\beta > 0$ a limit ordinal,

$$F(\beta) = \begin{cases} < \text{-least } \gamma \text{ such that } F(\zeta) R \gamma \text{ for all } \zeta < \beta \text{ if such } \gamma \text{ exists} \\ e \qquad \text{otherwise} \end{cases}$$

Note that this function is well-defined, since $F(\beta)R\gamma$ properly distinguishes the ordinals, for there are no R-maximal element in E, and $<$ is a strict well-ordering on E.

Now since $h(E) \in \text{Card}$, it is a limit ordinal, to show that F is injective, it suffices to show that $\forall \beta < h(E)$, $F \restriction_\beta$ is strictly order-preserving, i.e. $\forall x < y < \beta$, $F(x)RF(y)$. We shall prove this by Transfinite Induction.

$\beta = 0$ is vacuously true. If $\beta > 0$ is a limit ordinal, then $\beta = \bigcup_{\gamma < \beta} \gamma$, and so the strict ordering is preserved as given by the induction hypothesis. For β a successor ordinal, consider $F \restriction_{S(\beta)}$. $\forall x < y < S(\beta)$, if $y \neq \beta$, then we are done by the induction hypothesis. Suppose $y = \beta$. Since β is a successor ordinal, $\exists \gamma \in \text{Ord}$ such that $S(\gamma) = \beta$. Since $\gamma < S(\gamma)$ (by Proposition 15 Item 1), either $x = \gamma$ or $x < \gamma$. If $x < \gamma$, then our proof is complete by the induction hypothesis. If $x = \gamma$, then since $x = \gamma < S(\gamma) = \beta$, regardless if γ is a limit ordinal or successor ordinal, we have that $F(\gamma)RF(\beta)$.

Thus, by Transfinite Induction, we have that $F \restriction_\beta$ is strictly order-preserving for any $\beta < h(E)$, implying that $F : h(E) \hookrightarrow E$, hence contradicting the definition of $h(E)$. Thus, an R-maximal must exist. □

9 Lecture 9 Oct 04th

9.1 Cardinals (Continued 2)

9.1.1 Axiom of Choice (Continued)

From hereon, unless stated otherwise, we shall assume AC.

Proposition 28 (Using Cardinals to Measure Sets)

Assume AC. Every set is equinumerous with a cardinal.

Proof

Let $A \in \text{Set}$. By the Well-Ordering Principle, A is well-orderable, i.e. there is a strict well-ordering on A. By Theorem 23, $\exists! \alpha \in \text{Ord}$ such that $(A, <) \simeq (\alpha, \in)$. Let

$$S = \{\beta \leq \alpha \mid |\beta| = |\alpha|\},$$

which is a set by Bounded Separation. Note that $S \neq \emptyset$ since $\alpha \in S$. Let β be the least such ordinal in S. By this minimal choice, $\beta \in \text{Card}$, $|\beta| = |\alpha| = |A|$. □

And now our notation of $|A| = |B|$ makes sense provided the following definition.

Definition 26 (Cardinality)

Let $A \in \text{Set}$. $|A|$, in which we shall call the ***cardinality*** *of A, is the (unique) cardinal which is equinumerous with A.*

Proposition 29 (Lesser Cardinality)

$\forall A, B \in \text{Set} \quad |A| \leq |B| \iff \exists f : A \hookrightarrow B.$

Proof

Let $\kappa = |A|$ and $\lambda = |B|$. If $\kappa \leq \lambda$, then

$$A \overset{\text{bijection}}{\to} \kappa \overset{id}{\hookrightarrow} \lambda \overset{\text{bijection}}{\to} B$$

By composition of the 3 functions, $A \hookrightarrow B$.

Conversely, suppose $\exists h : A \hookrightarrow B$. Suppose to the contrary that $\lambda < \kappa$. Then $\lambda \subsetneq \kappa$. Then

$$\kappa \overset{\text{bijection}}{\to} A \overset{h}{\hookrightarrow} B \overset{\text{bijection}}{\to} \lambda$$

and so there exists an injection from $\kappa \to \lambda$. By Schröder-Bernstein, we have $|\kappa| = |\lambda|$, which contradicts the fact that $\kappa \in \text{Card}$. Thus $\kappa \leq \lambda$. □

Corollary 30 (Cardinalities are Always Comparable)

$\forall A, B \in \text{Set} \quad A \hookrightarrow B \vee B \hookrightarrow A.$

Proof

WLOG, suppose $\neg(B \hookrightarrow A)$. Then $\neg(|B| \leq |A|)$, i.e. $|A| < |B|$, i.e. (not by Proposition 29), $\exists f : A \hookrightarrow B$. □

Proposition 31 (Functions are "Lossy Compressions")

Suppose $f : A \to B$. *Then* $|\text{Img}(f)| \leq |A|$.

Proof

Let

$$\mathcal{F} = \{f^{-1}(y) \mid y \in \operatorname{Img}(f)\}$$

be the set of **fibres**[1] of f. Note that $\emptyset \notin \mathcal{F}$, as otherwise we would be saying that $\exists y \in \operatorname{Img}(f)$ that has no pre-image. Let $h : \operatorname{Img}(f) \to A$ by

$$h(y) := c(f^{-1}(y)) \in A$$

where $c : \mathcal{F} \to \cup\mathcal{F}$ is a choice function that exists by AC. Clearly so, h is injective, and so by Proposition 29, $|\operatorname{Img}(f)| \leq |A|$. $\square$

[1]

Definition 27 (Fibres)

*Let $f : A \to B$. For $y \in \operatorname{Img}(f)$, the **fibre** of y is defined as*

$$f^{-1}(y) = \{x \in A \mid f(x) = y\}.$$

*The fibres are also commonly called the **pullback**.*

Proposition 32 (Countable Union of Countable Sets is Countable)

Let $A \in \mathrm{Set}$ be countable, and every $a \in A$ is also countable. Then $\cup A$ is countable.

Proof

to be added

9.1.2 *Hierarchy of Infinite Cardinals*

Note (Notation)

Let $\kappa \in \mathrm{Card}$. Let $\kappa^+ := h(\kappa)$, which is the least ordinal not equinumerous with any subset of κ. We proved that $\forall E \in \mathrm{Set}$, $h(E) \in \mathrm{Card}$. Therefore, $\kappa^+ \in \mathrm{Card}$.

Remark

κ^+ is the least cardinal that contains κ. This follows immediately from the definition of $h(\kappa)$.

And so we observe that $\kappa \mapsto \kappa^+$ is a "successor" operation on cardinals.[2]

[2] Note that $\kappa + 1$ is not necessarily κ^+.

Definition 28 (Cardinal Numbers)

Using Transfinite Recursion, define the following ordinal-enumerated collection of cardinals

- $\aleph_0 = \omega$;
- $\forall \alpha \in \mathrm{Ord}$ *that is a successor ordinal,* $\aleph_{S(\alpha)} = \aleph_{\alpha+1} = \aleph_\alpha^+$; *and*
- $\forall \alpha > 0$ *that is a limit ordinal,* $\aleph_\alpha := \sup\{\aleph_\beta \mid \beta < \alpha\}$.

The $\aleph_\alpha$*'s are called* ***cardinal numbers****.*

Lemma 33 (Cardinal Numbers are Cardinals)

If $\alpha \in \mathrm{Ord}$, *then* $\aleph_\alpha \in \mathrm{Card}$.

Proof

We shall use Transfinite Induction. The result is clear for $\alpha = 0$, since $\aleph_0 = \omega$ and $\forall n \in \omega, |n| \neq |\omega|$. For successor ordinals α, since $h(\alpha)$ is an ordinal, we have that $\aleph_\alpha^+ = h(\aleph_\alpha)$ is also a cardinal. Now for $\alpha > 0$ a limit ordinal, suppose that $\beta < \aleph_\alpha$. Since $\aleph_\alpha = \sup\{\aleph_\gamma \mid \gamma < \alpha\}$, $\exists \gamma < \alpha$ such that $\beta < \aleph_\gamma$. Since $\aleph_\gamma$ is a cardinal by the Inductive Hypothesis, and $\beta < \aleph_\gamma < \aleph_\alpha$, we have that

$$|\beta| < |\aleph_\gamma| \leq |\aleph_\alpha|.$$

Thus $\aleph_\alpha$ is not equinumerous with any lesser ordinal, i.e. it is a cardinal. □

Lemma 34 (Ordinals Index the Cardinal Numbers)

$\forall \alpha < \beta \in \mathrm{Ord}$, *we have* $\aleph_\alpha < \aleph_\beta$.

Proof

We shall use Transfinite Induction on β. $\beta = 0$ is true vacuously so. For β a successor ordinal, suppose $\forall \alpha < \beta$, $\aleph_\alpha < \aleph_\beta$. Now $\aleph_{\beta+1} = \aleph_\beta^+ = h(\aleph_\beta)$, which we have that $\aleph_\beta < h(\aleph_\beta)$ as proven before.

Now suppose that $\beta > 0$ is a limit ordinal. Then $\aleph_\beta = \sup\{\aleph_\alpha \mid \alpha < \beta\}$ and $\forall \gamma < \alpha < \beta$, $\aleph_\gamma < \aleph_\alpha$. Since β is a limit ordinal, $\exists \zeta < \beta$ such that $\alpha < \zeta$. By the Induction Hypothesis, we have that $\aleph_\alpha < \aleph_\zeta$, and $\aleph_\zeta \leq \aleph_\beta$. Thus $\aleph_\zeta \subseteq \aleph_\beta$, and so $\aleph_\alpha < \aleph_\beta$. □

Lemma 35 (Infinite Cardinals are Distant)

$\forall \alpha \in \mathrm{Ord}$, $\alpha \leq \aleph_\alpha$. *The inequality is strict if α is a successor ordinal.*

Proof

Again, we shall use Transfinite Induction on α. For $\alpha = 0$, we have $0 < \aleph_0 = \omega$, and so $\alpha = 0$ holds. For α a successor ordinal, suppose $\alpha \leq \aleph_\alpha$. Thus $\alpha + 1 \leq \aleph_\alpha + 1$. By definition of $h(\aleph_\alpha)$ and definition of cardinal numbers

$$\alpha + 1 \leq \aleph_\alpha + 1 < \aleph_\alpha^+ = \aleph_{\alpha+1}.$$

Thus the statement holds for $\alpha + 1$, and indeed, the inequality is strict for successor ordinals. For $\alpha > 0$ a limit ordinal, suppose that $\forall \beta < \alpha$, we have $\beta \leq \aleph_\beta$. By Lemma 34, we have

$$\beta \leq \aleph_\beta < \aleph_\alpha.$$

Thus

$$\alpha = \sup\{\beta \mid \beta < \alpha\} < \aleph_\alpha$$

as required. □

Proposition 36 (All Infinite Cardinals are Indexed by the Ordinals)

Every infinite cardinal is of the form $\aleph_\alpha$ for some $\alpha \in$ Ord.

Proof

Let $\kappa \in$ Card. By Lemma 35, we have $\kappa \leq \aleph_\kappa < \aleph_{\kappa+1}$. And so we can show that $\forall \beta \in$ Ord, $\forall \kappa < \aleph_\beta$, $\exists \alpha < \beta$ such that $\kappa = \aleph_\alpha$. We shall use Transfinite Induction on β.

Since there are no infinite cardinals strictly below ω, $\beta = 0$ is trivially true. Suppose $\beta = \gamma + 1$ is a successor ordinal, where $\gamma \in$ Ord, and suppose $\kappa < \aleph_\beta$. Since $\gamma < \beta$, Lemma 34 implies that $\aleph_\gamma < \aleph_\beta = \aleph_{\gamma+1} = \aleph_\gamma^+$, and by definition, there are no cardinals between $\aleph_\gamma$ and $\aleph_\beta$. Thus $\kappa \leq \aleph_\gamma$. We thus have that either $\kappa = \aleph_\gamma$ or, by the Induction Hypothesis, $\exists \alpha < \gamma$ such that $\kappa = \aleph_\alpha$. Thus the statement holds for successor ordinals.

Let $\beta > 0$ be a limit ordinal and $\kappa < \aleph_\beta = \sup\{\aleph_\gamma \mid \gamma < \beta\}$. Then $\exists \gamma < \beta$ such that $\kappa < \aleph_\gamma$, and so by the Induction Hypothesis, $\exists \alpha < \gamma$ such that $\kappa = \aleph_\alpha$. □

Consequently, we have ourselves an *ordinal-valued, order-preserving complete indexing* of the infinite cardinals.

Exercise 9.1.1

For the inequality in Lemma 35, show that the equality can occur. In particular, consider the sequence of ordinals defined recursively by $\alpha_0 = 0$*, and* $\alpha_{n+1} = \aleph_{\alpha_n}$*, and verify that* $\alpha = \aleph_\alpha$*. In fact, this works if we start with any ordinal* α_0*, not just* 0*.*

10 *Lecture 10 Oct 11th*

10.1 *Cardinals (Continued 3)*

10.1.1 *Cardinal Arithmetic*

Contents in this lecture should be extended upon. A lot of the contents need to be explored for it was presented tersely so without all the details of which we need.

10.1.1.1 *Cardinal Summation*

Definition 29 (Cardinal Sum)

$\forall \kappa_1, \kappa_2 \in \mathrm{Card}$. *Let the* ***cardinal sum***

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

where $X_1, X_2 \in \mathrm{Set}$ *such that*

$$|X_1| = \kappa_1 \text{ and } |X_2| = \kappa_2$$

and $X_1 \cap X_2 = \emptyset$.

Remark

This definition does not depend on the choice of X_1 *and* X_2*, i.e. if we have another* X_1' *and* X_2' *such that*

$$|X_i'| = |X_i| = \kappa_i \quad i = 1, 2$$

and $X_1' \cap X_2' = \emptyset$*, then*

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

This shows that the cardinal summation is well-defined.

Exercise 10.1.1
Prove this remark.

Note

If $X, Y \in \mathrm{Set}$ are arbitarily chosen, then

$$|X \cup Y| \leq |X| + |Y|$$

Warning

The cardinal summation is different from ordinal summation. For example, $\aleph_0 + \aleph_0$, where the $+$ represents the ordinal summation, is not a cardinal, as we have shown that $|\omega + \omega| = |\omega|$.

Therefore, there is a need to explicitly mention the context of which $+$ is used.

10.1.1.2 Cardinal Product

Definition 30 (Cardinal Product)

$\forall \kappa_1, \kappa_2 \in \mathrm{Card}$, *the* ***cardinal product***

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

where $X_1, X_2 \in \mathrm{Set}$ with $|X_i| = \kappa_i$, for $i = 1, 2$.

Exercise 10.1.2

Prove that Definition 30 is well-defined.

Remark

We may as well choose

$$X_1 = \kappa_1 \text{ and } X_2 = \kappa_2,$$

and so $\kappa_1 \kappa_2 = |\kappa_1 \times \kappa_2|$.[1]

[1] We cannot do so for cardinal sums, for $\kappa_1 \cap \kappa_2 \neq \emptyset$.

Exercise 10.1.3

Prove that the cardinal sum and product agrees with ordinal sum and products on the finite ordinals. This is the usual arithmetic on natural numbers.

Theorem 37 (Dominance of the Larger Cardinal)

Let $\kappa_1, \kappa_2 \in \text{Card}$ not both finite. Then

1. $\kappa_1 + \kappa_2 = \max\{\kappa_1, \kappa_2\}$; *and*
2. *if neither κ_1 or κ_2 is 0, then $\kappa_1 \kappa_2 = \max\{\kappa_1, \kappa_2\}$.*

Exercise 10.1.4
Prove/read Theorem 37.

We can generalize the notions of cardinal sum and cardinal product. But first, a definition.

Definition 31 (I-sequence)

*Let $I \in \text{Set}$. By an **I-sequence** of sets, we mean a definite operation*[2]

$$f : I \to \text{Set}.$$

We write such sequences as

$$(x_i : i \in I)$$

where $x_i := f(i)$.

[2] Note that $f : I \to \text{Img}(f)$ is a function by the Replacement Axiom.

Definition 32 (Generalized Cardinal Sum)

*Suppose $(\kappa_i : i \in I)$ is a sequence of cardinals. We define the **(generalized) cardinal sum** to be*

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

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

Exercise 10.1.5

Check that Definition 32 is well-defined.

Note

Observe that

$$\bigcup_{i \in I} X_i = \cup \{ X_i \mid i \in I \} = \cup \operatorname{Img}(f)$$

where $(X_i : i \in I)$ *is a sequence of functions and* $f : I \to \mathrm{Set}$ *is given by* $f(i) = X_i$.

Theorem 38 (Properties of Cardinal Sum)

Let $I \in \mathrm{Set}$ *be infinite, and* $(\kappa_i : i \in I)$ *a sequence of cardinals not all zero. Then*

1. $\sup_{i \in I} \kappa_i \in \mathrm{Card}$; *and*
2. $\sum_{i \in I} \kappa_i = \max\{ |I|, \sup_{i \in I} \kappa_i \}$.

Exercise 10.1.6
Prove/read *Theorem 38*

Example 10.1.1

We have that

$$\sum_{0 < n \in \omega} n = \max\{ |\omega \setminus \{0\}|, \sup_{0 < n \in \omega} n \} = \max\{ \aleph_0, \omega \} = \aleph_0.$$

Definition 33 (Generalized Cardinal Product)

Suppose $(\kappa_i : i \in I)$ *is a sequence of cardinals. Then the* ***cardinal product*** *is defined as*

$$\prod_{i \in I} \kappa_i := |X_1 \times X_2 \times \ldots| = \left| \bigtimes_{i \in I} X_i \right|$$

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

Exercise 10.1.7

Check that *Definition 33* *is well-defined.*

Note

Note that $\bigtimes_{i \in I} X_i$ *is properly defined, as*

$$\begin{aligned} \bigtimes_{i \in I} X_i &:= \{(a_i : i \in I) \mid \forall i \in I, a_i \in X_i\} \\ &= \left\{ f : I \to \bigcup_{i \in I} X_i, f(i) = X_i \right\}. \end{aligned}$$

Remark

Once again, we might as well take $X_i = \kappa_i$*, just as we did when we defined pair products.*

Example 10.1.2

Suppose $\kappa_i = 2$ for all $i \in I$. Define a function such that

$$\bigtimes_{i \in I} 2 \to \mathcal{P}(I)$$

given by

$$(a_i : i \in I) \mapsto \{i \in I : a_i = 1\}.$$

Note that $2 = \{0, 1\}$, and so each $a_i = 0$ or 1. Clearly so, this is a bijection[3]. Consequently, we have

$$\prod_{i \in I} 2 = \left| \bigtimes_{i \in I} 2 \right| = |\mathcal{P}(I)|$$

[3] This is the usual correspondence between subsets and characteristic functions.

We claim that

$$|I| < |\mathcal{P}(I)|.$$

10.1.2 *An Interlude on the Continuum Hypothesis*

Theorem 39 (Cantor's Diagonalization)

$\forall I \in$ Set, *we have* $|I| < |\mathcal{P}(I)|$.

Proof

Clearly we have that $I \hookrightarrow \mathcal{P}(I)$ through the map $i \mapsto \{i\}$. Thus $|I| \leq |\mathcal{P}(I)|$.

Suppose to the contrary that there exists a bijection $f : I \to \mathcal{P}(I)$.

Let[4]

$$\Delta = \{i \in I : i \notin f(i)\} \subseteq I.$$

which is a set by Bounded Separation. Thus $\Delta \in \mathcal{P}(I)$. Then $\exists i_0 \in I$ such that $f(i_0) = \Delta$. Now if $i_0 \in \Delta$, then $i_0 \in f(i_0) = \Delta$, but this contradicts the membership condition which states that $i_0 \notin f(i_0)$. If $i_0 \notin \Delta$, then $i_0 \notin f(i_0) = \Delta$, but by the membership condition, it must be that $i_0 \in \Delta$, yet another contradiction. Thus such a bijection does not exist. □

[4] This definition looks awfully familiar to Russell's Paradox.

This proves our earlier claim. In fact, so long as $|I| > 2$,

$$\prod_{i \in I} 2 \neq \max\{|I|, \sup_{i \in I} 2\} = \max\{|I|, 2\},$$

since $RHS = |I|$ while $LHS = |\mathcal{P}(I)|$.

Definition 34 (Cardinal Exponentiation)

*$\forall \kappa, \lambda \in \text{Card}$, the **cardinal exponentiation** is defined as*

$$\kappa^\lambda := |\text{Fun}(\lambda, \kappa)|,$$

where $\text{Fun}(\lambda, \kappa)$ is the set of all functions from λ to κ.

Exercise 10.1.8

Prove that

$$\prod_{i < \lambda} \kappa = \kappa^\lambda.$$

As a consequence, we have that $2^\lambda = |\mathcal{P}(\lambda)|$. Then what is $|\mathcal{P}(\aleph_0)|$?

Axiom 40 (Continuum Hypothesis)

We have that

$$\aleph_0 < 2^{\aleph_0} = |\mathcal{P}(\aleph_0)|,$$

i.e. $2^{\aleph_0} = \aleph_1$.

More generally,

Axiom 41 (Generalized Continuum Hypothesis)

$\forall \kappa \in \mathrm{Card}$, *we have*

$$2^\kappa = \kappa^+.$$

In this course, we will not assume the Continuum Hypothesis (nor for the general case).

It has been proven by Paul Cohen (1963)[5] that the Continuum Hypothesis is independent from ZFC.

[5] Cohen, P. J. (1963). The independence of the continuum hypothesis. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC221287/

Part II

Model Theory

11 Lecture 11 Oct 16th

The course will not cover for cofinality and coregularity, but it may be helpful to read through the material.

We shall now venture into **Model Theory**.

11.1 First-order Logic

11.1.1 Structure

Definition 35 (Structure)

A structure $\mathcal{M}$ consists of the following data:

1. *A non-empty set M, called the* ***universe*** *of $\mathcal{M}$;*
2. *A sequence $(c_i : i \in I_{\text{con}})$, where I_{con} is an I-sequence of constants, of distinguished elements of M, called the* ***constants*** *of $\mathcal{M}$;*
3. *A sequence of M-valued functions on powers of M,*

   $$(f_i : M^{n_i} \to M : i \in I_{\text{fun}}),$$

   called the ***basic functions*** *of $\mathcal{M}$, and for each $i \in I_{\text{fun}}$, $n_i < \omega$ is called the* ***arity*** *of f_i;*
4. *A sequence of subsets of powers of M,*

   $$(R_i \subset M^{m_i} : i \in I_{\text{rel}})$$

   called the ***basic relations*** *of $\mathcal{M}$, and for each $i \in I_{\text{rel}}$, $m_i < \omega$ is called the* ***arity*** *of R_i.*

Note

1. *Note that we defined $M \neq \emptyset$, as there is nothing too interesting to study from an empty universe.*

2. *Also, note that we shall always use the corresponding capital alphabet as the universe of the structure (e.g. the universe of the structure $\mathcal{M}$ is M).*

3. *While $M \neq \emptyset$, we allow I_{con}, I_{fun} and I_{rel} to be $\emptyset$. In such a case, the structure that we have is about a* ***pure set****.*

4. *If f is a $0-$ary basic function, then we have*

$$f : M^0 \to M.$$

By convention, we have that $M^0 = 1 = \{0\}$, and so we have $f(0) \in M$. Thus, such an f is "essentially" a basic constant. Therefore, we shall usually assume that the arity of f is strictly positive.

Example 11.1.1

Consider $\mathbb{R}$, the set of real numbers[1]. The following are some of the structures that we can study on $\mathbb{R}$:

1. as a pure set: $\mathcal{R} = \mathbb{R}$, with $I_{\text{con}} = I_{\text{fun}} = I_{\text{rel}} = \emptyset$.
2. as an ordered set: $\mathcal{R} = (\mathbb{R}, <)$, with the basic **binary** relation $<$, and no basic functions or constants.
3. as an additive group: $\mathcal{R} = (\mathbb{R}, 0, +, -)$, with
   - 0 as a constant;
   - $+$ as a basic **binary** function; and
   - $-$ as a basic **unary** function.
4. as a ring: $\mathcal{R} = (\mathbb{R}, 0, 1, +, -, \cdot)$.
5. as a $\mathbb{Q}$-vector space: $\mathcal{R} = (\mathbb{R}, 0, +, -, (\lambda_q)_{q \in \mathbb{Q}})$, where $\lambda_q : \mathbb{R} \to \mathbb{R}$ is scalar multiplication by q.
6. as an ordered ring: $\mathcal{R} = (\mathbb{R}, 0, 1, +, -, \cdot, <)$.

[1] We did not construct $\mathbb{R}$ in our section on Set Theory. I may write up a full construction from $\mathbb{N}$ to $\mathbb{R}$ on my site.

Definition 36 (Expansion & Reduct)

*Suppose $\mathcal{M}$ and $\mathcal{N}$ are structures with the same universe. We say that $\mathcal{M}$ is an **expansion** of $\mathcal{N}$ or that $\mathcal{N}$ is a **reduct** of $\mathcal{M}$ if the basic constants, basic functions, and basic relations of $\mathcal{N}$ are contained in those of $\mathcal{M}$.*

Example 11.1.2

- $(\mathbb{R}, 0, +, -)$ is a reduct of $(\mathbb{R}, 0, 1, +, -, \cdot)$;
- $(\mathbb{R}, 0, 1, +, -, \cdot, <)$ is an expansion of $(\mathbb{R}, 0, 1, +, -, \cdot)$.

11.1.2 Language

Example 11.1.3

Consider the following 2 structures

$$\mathcal{R} = (\mathbb{R}, 0, +, -)$$
$$\mathcal{Z} = \left(\mathbb{Z}/2\mathbb{Z}, 0, +, - \right)$$

Notice that the $0, +$ and $-$ do not not actually share the same meaning, since, e.g., $+$ in $\mathcal{R}$ is addition on the reals, while $+$ in $\mathcal{Z}$ is addition on the integers modulo 2.

Definition 37 (Language)

*A **language** $\mathcal{L}$ consists of 3 sets of symbols:*

1. *a set $\mathcal{L}^{\text{con}}$ of constant symbols;*
2. *a set $\mathcal{L}^{\text{fun}}$ of function symbols, where each function symbol comes with an arity, which is a natural number;*
3. *a set $\mathcal{L}^{\text{rel}}$ of relation symbols, where each relation symbol comes with an arity, which is a natural number.*

Definition 38 ($\mathcal{L}$-structure)

*An **$\mathcal{L}$-structure** is a structure $\mathcal{M}$ with bijections between*

$$\begin{aligned} \mathcal{L}^{\text{con}} &\to I_{\text{con}} & c &\mapsto c^{\mathcal{M}} \\ \mathcal{L}^{\text{fun}} &\to I_{\text{fun}} & f &\mapsto f^{\mathcal{M}} \\ \mathcal{L}^{\text{rel}} &\to I_{\text{rel}} & R &\mapsto R^{\mathcal{M}} \end{aligned}$$

that preserves arity. We call

$$c^{\mathcal{M}}, f^{\mathcal{M}}, \text{ and } R^{\mathcal{M}}$$

*the **Interpretation** of c, f, R in $\mathcal{M}$.*

So in our previous example, we have that $\mathcal{R}$ and $\mathcal{Z}$ are both $\mathcal{L}$-structures, where $\mathcal{L} = \{0, +, -\}$ is the language with

- one constant symbol, 0;
- one binary function symbol, $+$;
- one unary function symbol, $-$.

This particular language is often referred to as the **language of additive groups**, for obvious[2] reasons.

[2] Obvious, if you have studied Group Theory.

Remark

- *To be precise, we should really write*

$$\mathcal{R} = (\mathbb{R}, 0^{\mathcal{R}}, +^{\mathcal{R}}, -^{\mathcal{R}}),$$

  so as to not obfuscate the symbols themselves and their interpretations in $\mathcal{R}$, but we shall forgive ourselves for this abuse of notation, for we shall leave this for the context to resolve this confusion.

- *Every group is naturally an $\mathcal{L}$-structure, where $\mathcal{L} = \{0, +, -\}$. The converse is definitely false. E.g. the following $\mathcal{L}$-structure*

$$\mathcal{M} = (\mathbb{Z}, 0^{\mathcal{M}}, +^{\mathcal{M}}, -^{\mathcal{M}})$$

  given by

$$\begin{aligned} 0^{\mathcal{M}} &= 12 \\ +^{\mathcal{M}} &: \mathbb{Z}^2 \to \mathbb{Z} \quad (a, b) \mapsto -2 \\ -^{\mathcal{M}} &: \mathbb{Z} \to \mathbb{Z} \quad a \mapsto 2^2 \end{aligned}$$

is not a group, since there are no identities nor inverses. But it is indeed an $\mathcal{L}$-structure.

Example 11.1.4

Let F be a field, and let $\mathcal{L}$ be the language of F-vector spaces, i.e.

$$\mathcal{L}^{\text{con}} = \{0\}$$
$$\mathcal{L}^{\text{fun}} = \{+, -, (\lambda_f)_{f \in F}\}$$
$$\mathcal{L}^{\text{rel}} = \emptyset$$

where $+$ is a binary function, $-$ a unary function, and λ_f a unary function for each $f \in F$. Then any F-vector spaces V is an $\mathcal{L}$-structure:

0	is interpreted as the zero vector
$+$	is interpreted as vector addition
$-$	is interpreted as the additive inverse of a vector
λ_f	is interpreted as scalar multiplication by f

The converse is, however, **not true**, and we shall study the reasons behind why this is not true later on.

Definition 39 ($\mathcal{L}$-Embedding)

Suppose $\mathcal{L}$ is a language, and $\mathcal{M}$ and $\mathcal{N}$ are $\mathcal{L}$-structures. An ***$\mathcal{L}$-embedding****, $j : \mathcal{M} \to \mathcal{N}$, is an injective function $j : M \to N$ such that*

1. $\forall c \in \mathcal{L}^{\text{con}} \quad j(c^{\mathcal{M}}) = c^{\mathcal{N}}$;
2. $\forall f \in \mathcal{L}^{\text{fun}}$ *each with arity n_f, and $\forall a_1, \ldots, a_{n_f} \in M$, we have*
$$j(f^{\mathcal{M}}(a_1, \ldots, a_{n_f})) = f^{\mathcal{N}}(j(a_1), \ldots, j(a_{n_f}))$$
3. $\forall R \in \mathcal{L}^{\text{rel}}$ *each with arity m_R, and $\forall (a_1, \ldots, a_{m_R}) \in M^{m_R}$, we have*
$$(a_1, \ldots, a_{m_R}) \in R^{\mathcal{M}} \iff (j(a_1), \ldots, j(a_{m_R})) \in R^{\mathcal{N}}.$$

Definition 40 (Substructure)

Continuing with the above assumptions and notation, if $M \subseteq N$, then we say that $\mathcal{M}$ is a ***substructure*** *on $\mathcal{N}$ if the identity map id : $M \to N$ is an $\mathcal{L}$-embedding. We denote this notion, without confusion, by $\mathcal{M} \subset \mathcal{N}$.*

We may also say that $\mathcal{N}$ is an ***extension*** *of $\mathcal{M}$.*

12 Lecture 12 Oct 18th

12.1 First-order Logic (Continued)

12.1.1 Language (Continued)

Note

Notice that $\mathcal{L}$-structures $\mathcal{M} \subseteq \mathcal{N}$ iff

- $M \subseteq N$;
- $\forall c \in \mathcal{L}^{\text{con}} \quad c^{\mathcal{M}} = c^{\mathcal{N}}$;
- $\forall f \in \mathcal{L}^{\text{fun}}$, *each with arity* n_f, $f^{\mathcal{M}} = f^{\mathcal{M}} \restriction_{M^{n_f}}$;
- $\forall R \in \mathcal{L}^{\text{rel}}$, *each with arity* m_R, $R^{\mathcal{M}} \subseteq R^{\mathcal{N}} \cap M^{m_R}$.

Exercise 12.1.1

Suppose $\mathcal{N}$ is an $\mathcal{L}$-structure with $A \subset N$, where $A \neq \emptyset$. Show that A is the universe of some (unique) substructure of $\mathcal{N}$ iff

- $\forall c \in \mathcal{L}^{\text{con}} \quad c^{\mathcal{N}} \in A$;
- $\forall f \in \mathcal{L}^{\text{fun}}$ *that is n-ary,* $f^{\mathcal{N}}(A^n) \subset A$.

The choice of the language determine what the substructures are.

Example 12.1.1

The following is table showing an example of a substructure in the given structures:

Structure	Substructures
$\mathbb{R}$	all non-empty subsets
$(\mathbb{R}, 0, +)$	all submonoids of $\mathbb{R}$ [1]
$(\mathbb{R}, 0, +, -)$	all subgroups
$(\mathbb{R}, 0, 1, +, -, \cdot)$	all subrings
$(\mathbb{R}, <)$	all non-empty subsets with induced order
$(\mathbb{R}, 0, +, -, (\lambda_q)_{q\in\mathbb{Q}})$	all $\mathbb{Q}$-subspaces

[1] Not that substructures of $(\mathbb{R}, 0, +)$ is not necessarily a group, and this is why we have always specified for $-$ for the language of additive groups. However, this is not necessary for the unique inverse in the language of fields, of which we shall clarify in a later section.

Definition 41 ($\mathcal{L}$-isomorphism)

An $\mathcal{L}$-isomorphism is a surjective $\mathcal{L}$-embedding.

Exercise 12.1.2

Let $j : \mathcal{M} \to \mathcal{N}$ be an $\mathcal{L}$-embedding. Then j induces an $\mathcal{L}$-isomorphism between $\mathcal{M}$ and a unique substructure $\mathcal{M}' \subseteq \mathcal{N}$.

12.1.2 *Terms & Formulas*

From hereon, we shall have the following countable infinite set,

$$\mathrm{Var} = \{x_0, x_1, \ldots\},$$

which we shall call as our set of **variables**[2].

[2] This is not to be confused with the **variance** in probability theory.

Definition 42 ($\mathcal{L}$-terms)

*Let $\mathcal{L}$ be a language. The set of **$\mathcal{L}$-terms** is the smallest set of strings of symbols from*

$$\mathcal{L} \cup \mathrm{Var} \cup \{(,)\} \cup \{,\} \quad ^3$$

satisfying

- *every variable is an $\mathcal{L}$-term;*
- *every constant of $\mathcal{L}^{\text{con}}$ is also an $\mathcal{L}$-term;*
- [4]*if $t_1, t_2, \ldots, t_n$ are $\mathcal{L}$-terms and $f \in \mathcal{L}^{\text{fun}}$ is n-ary, then $f(t_1, \ldots, t_n)$ is also an $\mathcal{L}$-term.*

[3] By $\{(,)\}$ and $\{,\}$, we mean the punctuation symbols "(", ",", ")".

[4] This is a recursive formula.

We often write a term t as

$$t = t(x_1, \ldots, x_n)$$

to mean that the variables in t come from $\{x_1, \ldots, x_n\}$.

Remark
All $x_1, \ldots, x_n$ need not always appear in t.

We shall use the following example to show another abuse of notation that we shall gladly do in this course:

Example 12.1.2

Let $\mathcal{L} = \{0, 1, +, -, \times\}$. The following is an $\mathcal{L}$-term:

$$\times(+(x_0, -(x_1)), \times(1, x_2)).$$

It is fairly straightforward to verify that the above is indeed an $\mathcal{L}$-term: using items (1), (2) and (3) in our definition above, by the given order,

1. all $x_0, x_1, 1, x_2$ are $\mathcal{L}$-terms by (1) and (2);
2. $-$ is a unary function, and so by (3), we have that $-(x_1)$ is an $\mathcal{L}$-term;
3. $+$ is a binary function, and by (3), we have that $+(x_0, -(x_1))$ is an $\mathcal{L}$-term;
4. $\times$ is a binary function, and so by (3), we have that $\times(1, x_2)$ is an $\mathcal{L}$-term;
5. finally, by (3), $\times(+(x_0, -(x_1)), \times(1, x_2))$ is an $\mathcal{L}$-term.

We shall "informally" write this as

$$(x_0 + (-x_1))(1x_2).$$

Note that we do not simply write $x_2 = 1x_2$, since we may not have that 1 acts as a "multiplicative identity" of sorts, as we would have in a ring.

Definition 43 (Interpretation)

Suppose $\mathcal{M}$ is an $\mathcal{L}$-structure, and $t = t(x_1, \ldots, x_n)$ is an $\mathcal{L}$-term. We define the ***itnerpretation*** *of t in $\mathcal{M}$ to be the function*

$$t^{\mathcal{M}} : M^n \to M,$$

defined recursively by

- *If t is x_i for some $1 \leq i \leq n$, then*

$$t^{\mathcal{M}} : M^n \to M \quad (a_1, \ldots, a_n) \mapsto a_i.$$

- *If t is c for some $c \in \mathcal{L}^{\text{con}}$, then*

$$t^{\mathcal{M}} : M^n \to M \quad (a_1, \ldots, a_n) \mapsto c^{\mathcal{M}}.$$

- *If t is $f(t_1, \ldots, t_l)$, where $t_1, \ldots, t_l$ are $\mathcal{L}$-terms and f is an l-ary function symbol, then $t^{\mathcal{M}} : M^n \to M$*

$$(a_1, \ldots, a_n) \mapsto f(t_1^{\mathcal{M}}(a_1, \ldots, a_n), t_2^{\mathcal{M}}(a_1, \ldots, a_n), \ldots, t_l^{\mathcal{M}}(a_1, \ldots, a_n))$$

Note

Note that the function $t^{\mathcal{M}}$ depends not just on t but on its presentation $t = t(x_1, \ldots, x_n)$.

Example 12.1.3

Let $\mathcal{L} = \{0, +, -\}$ and $\mathcal{R} = (\mathbb{R}, 0, +, -)$ an $\mathcal{L}$-structure. Let t be the term $x \in \text{Var}$. Then for $y \in \text{Var}$, the following are some interpretations of t in $\mathcal{R}$:

1. if $t = t(x)$, then we may interpret $t^{\mathcal{R}} : \mathbb{R} \to \mathbb{R}$ as the identity map, i.e. $r \mapsto r$;
2. if $t = t(x, y)$, then we may interpret $t^{\mathcal{R}} : \mathbb{R}^2 \to \mathbb{R}$ as taking the first component of the tuple, i.e. $(a, b) \mapsto a$;
3. if $t = t(x, y)$, then we may interpret $t^{\mathcal{R}} : \mathbb{R}^2 \to \mathbb{R}$ as taking the second component of the tuple, i.e. $(a, b) \mapsto b$.

Exercise 12.1.3

Prove the following: suppose $\mathcal{M}$ is an $\mathcal{L}$-structure, and $A \subseteq M$ such that $A \neq \emptyset$, then A is the universe of the substructure iff $\forall t = t(x_1, \ldots, x_n)$ that are $\mathcal{L}$-terms, we have $t^{\mathcal{M}}(A^n) \subseteq A$.

Example 12.1.4

Let $\mathcal{L} = \{0, 1, +, -, \times\}$ and $\mathcal{Z} = (\mathbb{Z}, =, 1, +, -, \times)$. Let $P \in \mathbb{Z}[x]$ where P is $x^2 + 2y - 1$. Then let $t_P = t_P(x, y)$ be the $\mathcal{L}$-term

$$x^2 + (y + y) + (-1).$$

Then the interpretation of t_P is

$$t_P^{\mathcal{Z}} : \mathbb{Z}^2 \to \mathbb{Z} \quad (a, b) \mapsto a^2 + 2b + 1.$$

Definition 44 (Atomic $\mathcal{L}$-Formulas)

*An **atomic $\mathcal{L}$-formula** is a (finite) string of symbols from*

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

of the form:

1. *$(t = s)$ where t, s are $\mathcal{L}$-terms;*
2. *$R(t_1, \ldots, t_l)$ where $t_1, \ldots, t_l$ are $\mathcal{L}$-terms, and $R \in \mathcal{L}^{\mathrm{rel}}$ with arity $l \in \omega$.*

Definition 45 ($\mathcal{L}$-formulas)

*The set of **$\mathcal{L}$-formulas** is the smallest set of (finite) strings of symbols from*

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

satisfying:

1. *every atomic $\mathcal{L}$-formula is an $\mathcal{L}$-formula;*
2. *if ϕ, ψ are $\mathcal{L}$-formulae, then so is $(\phi \wedge \psi)$, $\neg\phi$, and $(\phi \vee \psi)$;*
3. *if ϕ is an $\mathcal{L}$-formula and $x \in \mathrm{Var}$, then $\exists x \phi$ and $\forall x \phi$ are $\mathcal{L}$-formulae.*

Remark

1. *We shall use the following abbreviations:*
   - $\phi \to \psi$ *for* $(\neg\phi \vee \psi)$;
   - $\phi \leftrightarrow \psi$ *for* $(\phi \to \psi) \wedge (\psi \to \phi)$.
2. *We shall gleefully make the following abuse of notation: for the language* $\mathcal{L} = \{\times, <\}$, *we write* $x < y^2$ *instead of* $<(x, \times(y, y))$.

Definition 46 (Bound and Free Variables)

Suppose ϕ *is an* $\mathcal{L}$*-formula. An occurrence of a variable* x *in* ϕ *is called* ***bound*** *if it appears inside the scope of a quantifer (i.e. in the existence of* $\exists$ *and* $\forall$*). Otherwise, they are called* ***free****.*

Example 12.1.5

Let $\mathcal{L} = \{\in\}$, and $x, y, z \in \mathrm{Var}$. The following is an $\mathcal{L}$-formula:

$$(x \in y) \wedge \forall z \Big((z \in y) \to \big((z \in x) \vee (z = x) \big) \Big)$$

We have that x, y are the free variables while z is bound.

We write $\phi = \phi(x_1, \ldots, x_n)$ to mean that the free variables of ϕ come from $\{x_1, \ldots, x_n\}$, but it is not necessary that all of the variables appear in the $\mathcal{L}$-formula.

For the sake of convenience, we hall assume that no variable can be simultaneously bound and free in an $\mathcal{L}$-formula.

13 Lecture 13 Oct 23rd

13.1 First-order Logic (Continued 2)

13.1.1 Terms & Formulas (Continued)

Example 13.1.1

Let $\mathcal{L} = \{0, 1, +, -, \cdot\}$, and $x, y \in \mathrm{Var}$. We have that $x^2 = \cdot(x, x)$ and y are $\mathcal{L}$-terms. Given $\mathcal{R} = (\mathbb{R}, 0, 1, +, -, \cdot)$, we can interpret x^2, using (x, y) as

$$\left(x^2\right)^{\mathcal{R}} : \mathbb{R} \to \mathbb{R} \text{ such that } (r, s) \mapsto r^2$$

and interpret y, using (x, y) as

$$y^{\mathcal{R}} : \mathbb{R} \to \mathbb{R} \text{ such that } (r, s) \mapsto s.$$

Now since x^2 and y are $\mathcal{L}$-terms, we have that $x^2 = y$ is an **atomic formula**[1]. In the $\mathcal{L}$-formula $\exists x(x^2 = y)$, the only free variable is y, which is not "specified" by a quantifier.

On the other hand, for the $\mathcal{L}$-formula $\forall y; \exists x(x^2 = y)$, both x and y are bound.

[1] Atomic formulas are quantifier-free and so all the variables are free.

Remark

Notice that an $\mathcal{L}$-formula with free variable say something about the free variable. In our example of $\exists x(x^2 = y)$, the $\mathcal{L}$-formula "says" that y has a square root.

In the $\mathcal{L}$-formula $\forall y\ \exists x(x^2 = y)$, the meaning is different: it "says" that every element (in $\mathbb{R}$) has a square root. This is an important observation for the next definition.

Definition 47 ($\mathcal{L}$-Sentences)

An $\mathcal{L}$-formula with no free variable is called an ***$\mathcal{L}$-sentence****.*

Note

An $\mathcal{L}$-sentence has a truth value.

Definition 48 (Satisfaction / Realization)

Given a language $\mathcal{L}$ and an $\mathcal{L}$-structure $\mathcal{M}$, for some $n < \omega$, let

$$\phi = \phi(x_1, \ldots, x_n)$$

be an $\mathcal{L}$-formula[2], $\bar{a} = (a_1, \dots, a_n) \in M^n$ *and* $\bar{x} = (x_1, \dots x_n)$.

We define that $\bar{a}$ ***satisfies*** *(or* ***realizes****) the formula $\phi(\bar{x})$ in $\mathcal{M}$*[3]*, of which we denote by $\mathcal{M} \models \phi(\bar{a})$, through the following recursive definition, which we iterate on the complexity of the formula $\phi(\bar{x})$:*

1. *If $\phi(\bar{x})$ is of the form $t_1 = t_2$, where*

$$t_1 = t_1(\bar{x}) \text{ and } t_2 = t_2(\bar{x})$$

are $\mathcal{L}$-terms, then

$$\mathcal{M} \models \phi(\bar{a}) \text{ if } t_1^{\mathcal{M}}(\bar{a}) = t_2^{\mathcal{M}}(\bar{a}),$$

where $t_1^{\mathcal{M}}$ and $t_2^{\mathcal{M}}$ are realizations of t_1 and t_2 in $\mathcal{M}$, respectively.

2. *If $\phi(\bar{x})$ is of the form*

$$R(t_1, \dots, t_l), \text{ where } t_i = t_i(\bar{x}) \text{ are } \mathcal{L}\text{-terms, } R \in \mathcal{L}^{\mathrm{rel}}$$

with arity $l < \omega$, then we define $\mathcal{M} \models \phi(\bar{a})$ if

$$\left(t_1^{\mathcal{M}}(\bar{a}),\, t_2^{\mathcal{M}}(\bar{a}),\, \dots,\, t_l^{\mathcal{M}}(\bar{a})\right) \in R^{\mathcal{M}} \subseteq \mathcal{M}^l$$

3. *If $\phi(\bar{x})$ is of the form*

$$\phi_1(\bar{x}) \wedge \phi_2(\bar{x}),$$

where $\phi_1(\bar{x}), \phi_2(\bar{x})$ are known $\mathcal{L}$-formulas, then we define

$$\mathcal{M} \models \phi(\bar{a}) \iff \mathcal{M} \models \phi_1(\bar{a}) \wedge \mathcal{M} \models \phi_2(\bar{a}).$$

4. *If $\phi(\bar{x})$ is of the form*

$$\phi_1(\bar{x}) \vee \phi_2(\bar{x}),$$

where $\phi_1(\bar{x}), \phi_2(\bar{x})$ are known $\mathcal{L}$-formulas, then we define

$$\mathcal{M} \models \phi(\bar{a}) \iff \mathcal{M} \models \phi_1(\bar{a}) \vee \mathcal{M} \models \phi_2(\bar{a}).$$

5. *If $\phi(\bar{x})$ is of the form*

$$\neg\psi(\bar{x}),$$

where $\psi(\bar{x})$ is a known $\mathcal{L}$-formula, then we define[4]

$$\mathcal{M} \models \phi(\bar{a}) \iff \text{ it is not the case that } \mathcal{M} \models \phi(\bar{a}),$$

of which the latter shall be denoted as $\mathcal{M} \not\models \psi(\bar{a})$.

[2] If $n = 0$, then ϕ is simply an $\mathcal{L}$-sentence.

[3] We may also say that $\phi(\bar{a})$ is true in $\mathcal{M}$.

Notice that in the recursive definition, we start with **atomic $\mathcal{L}$-formulas** first before going to the **$\mathcal{L}$-formulas**.

[4] Here, we conveniently assumed the Law of Excluded Middle.

6. *If $\phi(\bar{x})$ is of the form*
$$\exists y \psi(\bar{x}, y),$$
where $\psi(\bar{x}, y)$ is an $\mathcal{L}$-formula, and y a (single) variable, then we define
$$\mathcal{M} \models \phi(\bar{a}) \iff \text{there exists } b \in M \text{ such that } \mathcal{M} \models \psi(\bar{a}, b).$$

7. *If $\phi(\bar{x})$ is of the form*
$$\forall y \psi(\bar{x}, y),$$
where $\psi(\bar{x}, y)$ is an $\mathcal{L}$-formula, and y a (single) variable, then we define
$$\mathcal{M} \models \phi(\bar{a}) \iff \text{for every} b \in M \text{ such that } \mathcal{M} \models \psi(\bar{a}, b).$$

Note

*The **set of all realizations** of an $\mathcal{L}$-formula ϕ in the $\mathcal{L}$-structure $\mathcal{M}$ is denoted as*
$$\phi^{\mathcal{M}} := \{\bar{a} \in M^n : \mathcal{M} \models \phi(\bar{a})\}.$$
*We may also call this set the **subset of M^n defined by ϕ**.*

Finally, we can define **definability**.

Definition 49 ($\mathcal{L}$-Definable)

*A subset $S \subset M^n$ is said to be **$\mathcal{L}$-definable** if $S = \phi^{\mathcal{M}}$ for some $\mathcal{L}$-formula $\phi = \phi(\bar{x}) = \phi(x_1, \ldots, x_n)$.*

Remark

If $n = 0$, we have that ϕ is an $\mathcal{L}$-sentence, and since $M^0 = 1 = \{0\}$, there are two possible scenarios and that is
$$\text{either } \mathcal{M} \models \phi \text{ or } \mathcal{M} \models \neg\phi.$$
Notice that $\phi^{\mathcal{M}}$ is either $1 = \{0\}$ or $0 = \emptyset$.

Example 13.1.2

Let $\mathcal{L} = \{0, 1, +, -, \times\}$, $\psi(x, y)$ be the atomic $\mathcal{L}$-formula $x^2 = y$, and $\phi(y)$ be the $\mathcal{L}$-formula $\exists x(x^2 = y)$. It is clear that we can also write $\phi(y) = \exists x \psi(x, y)$.

Let $\mathcal{R} = (\mathbb{R}, 0, 1, +, -, \times)$. We have that

$$\mathcal{R} \models \neg\phi(-1) \text{ and } \mathcal{R} \models \phi(2).$$

In fact, $\phi^{\mathcal{R}} = \mathbb{R}_{\geq 0} \subseteq \mathbb{R}$.

On the other hand, for the $\mathcal{L}$-structure $\mathcal{Q} = (\mathbb{Q}, 0, 1, +, -, \times)$, we have that

$$\mathcal{Q} \models \neg\phi(-1) \text{ and } \mathcal{Q} \models \neg\phi(2).$$

It is worth noting that

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

For the $\mathcal{L}$-structure $\mathcal{C} = (\mathbb{C}, 0, 1, +, -, \times)$, we have that $\phi^{\mathcal{C}} = \mathbb{C}$, or we may also write that $\mathcal{C} \models \forall y\ \phi(y)$ [5].

[5] We may further simplify our notations by letting

$$\sigma = \forall y \exists x(x^2 = y),$$

which we can then simply write $\mathcal{C} \models \sigma$.

Example 13.1.3

Consider the same language as in the last example, and let $\phi(y, z) = \exists x(x^2 = y - z)$. In the $\mathcal{L}$-structure $\mathcal{R} = (\mathbb{R}, 0, 1, +, -, \times)$, it is clear that

$$\phi^{\mathcal{R}} = \left\{ (a, b) \in \mathbb{R}^2 : b \leq a \right\} \subseteq \mathbb{R}^2,$$

and we observe that the set of solutions is also the binary relation $\geq$, i.e. $\geq$ is definable in the $\mathcal{L}$-structure $\mathcal{R}$.

Similarly so, we can show that $\leq$ is $\mathcal{L}$-definable in $\mathcal{R}$. [6]

[6] In other words, this example shows to us that we can define $\leq$ and $\geq$ using just the symbols $=$ and $-$.

Proposition 42 (Structure Traversal with respect to Quantifiers)

Let $\mathcal{M} \subseteq \mathcal{N}$ be $\mathcal{L}$-structures, where $\mathcal{M}$ is the $\mathcal{L}$-substructure, and variables $\bar{x} = (x_1, \ldots, x_n)$. Let $\phi(\bar{x})$ be an $\mathcal{L}$-formula, and $\bar{a} = (a_1, \ldots, a_n) \in M^n$. Then

1. *if ϕ is a* ***quantifier-free*** *formula, then*

$$\mathcal{M} \models \phi(\bar{a}) \iff \mathcal{N} \models \phi(\bar{a}).$$

2. *if ϕ is **universal**, i.e. $\phi(\bar{x})$ is of the form*

$$\forall y_1 \forall y_2 \dots \forall y_m \psi(\bar{x}, y_1, y_2, \dots, y_m),$$

where ψ is quantifier-free, then

$$\mathcal{M} \models \phi(\bar{a}) \impliedby \mathcal{N} \models \phi(\bar{a}).$$

3. *if ϕ is **existential**, i.e. $\phi(\bar{x})$ is of the form*

$$\exists y - 1 \exists y_2 \dots \exists y_m \psi(\bar{x}, y_1, y - 2, \dots, y_m),$$

where ψ is quantifer-free, then

$$\mathcal{M} \models \phi(\bar{a}) \implies \mathcal{N} \models \phi(\bar{a}).$$

I shall directly quote from the course notes from Professor Moosa:

> *This proposition has a very typical proof. In order to prove something about all formulas 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.*

Proof

1. **Claim**: We shall first show that if $t(\bar{x})$ is an $\mathcal{L}$-term, then $t^{\mathcal{M}} = t^{\mathcal{N}} \restriction_{M^n}$. We shall prove this claim by using induction on the complexity of t.

   - If t is a variable x_i, then we have

     $$t^{\mathcal{M}} : M^n \to M \text{ given by } (a_1, \dots, a_n) \mapsto a_i \qquad (13.1)$$
     $$t^{\mathcal{N}} : N^n \to N \text{ given by } (b_1, \dots, b_n) \mapsto b_i$$

     Since $M^n \subseteq N^n$, we have that the restriction of $t^{\mathcal{N}}$ to M^n will give us (13.1).

   - If t is a constant symbol, i.e. some $c \in \mathcal{L}^{\text{con}}$, then we simply have $c^{\mathcal{M}} = c^{\mathcal{N}}$, for $M \subseteq N$.

   - If t is $f(t_1, \dots t_l)$, where $f \in \mathcal{L}^{\text{fun}}$, where $t_1, \dots, t_l$ are $\mathcal{L}$-terms, then

     $$\begin{aligned} t^{\mathcal{M}}(\bar{a}) &= f^{\mathcal{M}}(t_1^{\mathcal{M}}(\bar{a}), \dots, t_l^{\mathcal{M}}(\bar{a})) \\ &= f^{\mathcal{M}}(t_1^{\mathcal{N}}(\bar{a}), \dots, t_l^{\mathcal{N}}(\bar{a})) \quad \because \text{ IH} \\ &= f^{\mathcal{N}}(t_1^{\mathcal{N}}(\bar{a}), \dots, t_l^{\mathcal{N}}(\bar{a})) \quad \because \text{ note on page 91} \\ &= t^{\mathcal{N}}(\bar{a}). \end{aligned}$$

     Therefore $t^{\mathcal{N}} \restriction_{M^n} = t^{\mathcal{M}}$ as claimed $\dashv$.

Now to prove our original statement. We shall prove the statement by performing induction on ϕ.

- $\phi(\bar{x})$ **is** $t_1 = t_2$: We have

$$\mathcal{M} \models \phi(\bar{a}) \overset{\text{defn}}{\iff} t_1^{\mathcal{M}}(\bar{a}) = t_2^{\mathcal{M}}(\bar{a}) \overset{\text{claim}}{\iff} t_1^{\mathcal{N}}(\bar{a}) = t_2^{\mathcal{N}}(\bar{a}) \overset{\text{defn}}{\iff} \mathcal{N} \models \phi(\bar{a}).$$

- $\phi(\bar{x})$ **is** $R(t_1, \ldots t_l)$: Observe that

$$\begin{aligned} \mathcal{M} \models \phi(\bar{a}) &\overset{\text{defn}}{\iff} \left(t_1^{\mathcal{M}}(\bar{a}), \ldots, t_l^{\mathcal{M}}(\bar{a})\right) \in R^{\mathcal{M}} \\ &\overset{\text{claim}}{\iff} \left(t_1^{\mathcal{N}}(\bar{a}), \ldots, t_l^{\mathcal{N}}(\bar{a})\right) \in R^{\mathcal{M}} \\ &\iff \left(t_1^{\mathcal{N}}(\bar{a}), \ldots, t_l^{\mathcal{N}}(\bar{a})\right) \in R^{\mathcal{N}} \quad \because R^{\mathcal{M}} = R^{\mathcal{N}} \cap M^l \\ &\overset{\text{defn}}{\iff} \mathcal{N} \models \phi(\bar{a}). \end{aligned}$$

From this part onwards, suppose ϕ_1, ϕ_2 are quantifier-free $\mathcal{L}$-formulas.

- $\phi(\bar{x}) = \phi_1(\bar{x}) \vee \phi_2(\bar{x})$: We have

$$\begin{aligned} \mathcal{M} \models \phi(\bar{a}) &\iff \mathcal{M} \models \phi(\bar{a}) \text{ or } \mathcal{M} \models \phi_2(\bar{a}) \\ &\iff \mathcal{N} \models \phi(\bar{a}) \text{ or } \mathcal{N} \models \phi_2(\bar{a}) \quad \because \text{ IH} \\ &\iff \mathcal{N} \models \phi(\bar{a}). \end{aligned}$$

- $\phi(\bar{x}) = \phi_1(\bar{x}) \wedge \phi_2(\bar{x})$: We have

$$\begin{aligned} \mathcal{M} \models \phi(\bar{a}) &\iff \mathcal{M} \models \phi_1(\bar{a}) \text{ and } \mathcal{M} \models \phi_2(\bar{a}) \\ &\iff \mathcal{N} \models \phi_1(\bar{a}) \text{ and } \mathcal{N} \models \phi_2(\bar{a}) \quad \because \text{ IH} \\ &\iff \mathcal{N} \models \phi(\bar{a}). \end{aligned}$$

- $\phi(\bar{x}) = \neg\phi_1(\bar{x})$: We have

$$\begin{aligned} \mathcal{M} \models \phi(\bar{a}) &\iff \mathcal{M} \not\models \phi_1(\bar{a}) \\ &\iff \mathcal{N} \not\models \phi_1(\bar{a}) \quad \because \text{ IH} \\ &\iff \mathcal{N} \models \phi(\bar{a}). \end{aligned}$$

There is no need to check for cases where ϕ is universal or existential, since our assumption is that ϕ is quantifier-free. This completes the proof.

□

Exercise 13.1.1

Prove (2) and (3) of Proposition 42.

14 Lecture 14 Oct 25th

14.1 First-order Logic (Continued 3)

14.1.1 Terms & Formulas (Continued 2)

Example 14.1.1

Let $\mathcal{L}$ be the language of rings[1]. We have that in $\mathcal{Z}$, where

$$(\mathbb{Z}, 0, 1, +, -, \times) = \mathcal{Z} \subseteq \mathcal{Q} = (\mathbb{Q}, 0, 1, +, -, \times),$$

for the $\mathcal{L}$-formula $\phi(x) = \exists y(x = 2y)$, we observe that

$$\mathcal{Z} \models \neg\phi(1) \text{ but } \mathcal{Q} \models \phi(1).$$

[1] From hereon, for the sake of simplicity, I shall gleefully use such declarations for a language wherever it is convenient, and when the context is clear that I am merely using the common symbols that are used in the said theory.

This shows to us that existential formulas that

- are satisfied in an extension may not be satisfied in the substructure;
- is not satisfied in the substructure may be satisfied in the extension.

14.1.2 Elementary Embeddings

Considering Proposition 42, it is interesting for us to consider substructures that satisfies all the formulas of its extension, including formulas that are either universal or existential.

Definition 50 (Elementary Embeddings)

Suppose $\mathcal{M} \subseteq \mathcal{N}$ are $\mathcal{L}$-structures, and $\mathcal{M}$ is a substructure of $\mathcal{N}$. An $\mathcal{L}$-embedding $j : \mathcal{M} \to \mathcal{N}$ is an ***elementary embedding*** *if for any*

$\mathcal{L}$-formula ϕ and $\bar{a} \in M^n$,

$$\mathcal{M} \models \phi(\bar{a}) \iff \mathcal{N} \models \phi(j(\bar{a})).$$

*We call such an $\mathcal{M}$ an **elementary substructure** of $\mathcal{N}$, and we denote $\mathcal{M} \preceq \mathcal{N}$.*

In simpler words, any formula that is true in $\mathcal{N}$ is true in its elementary substructure $\mathcal{M}$,

Corollary 43 (Isomorphisms are Elementary Embeddings)

Every isomorphism is an elementary embedding.

Proof

We already know the result for quantifier-free statements from Proposition 42, and so it suffices to prove this statement by induction on the number of quantifiers, which we shall call n. In fact, it suffices to prove for the case of an existential $\mathcal{L}$-formula, since we can write $\forall$ as $\neg\exists\neg$.

Suppose $g : \mathcal{M} \to \mathcal{N}$ is an $\mathcal{L}$-isomorphism, where $\mathcal{M} \subseteq \mathcal{N}$. There is nothing to show for $n = 0$. Suppose that the statement holds for $n = m$. Consider an $\mathcal{L}$-formula of the form $\phi(\bar{x}) = \exists y \psi(\bar{x}, y)$, where $\bar{x} \in M^{m+1}$, and ψ an $\mathcal{L}$-formula with lower complexity than ϕ. Then for $a \in M^{m+1}$,

$$\begin{aligned} \mathcal{M} \models \phi(\bar{a}) &\iff \mathcal{M} \models \exists b \psi(\bar{a}, b) \\ &\iff \mathcal{N} \models \exists b \psi(g(\bar{a}), g(b)) \quad \because \text{ IH} \\ &\iff \mathcal{N} \models \exists c \psi(g(\bar{a}), c) \quad \because f \text{ is bijective} \\ &\iff \mathcal{N} \models \phi(g(\bar{a})). \end{aligned}$$

This completes the proof. □

Example 14.1.2

Let $\mathcal{Q} = (\mathbb{Q}, <, 0)$ be an $\mathcal{L} = \{<, 0\}$-structure. Show that the graph of addition is not definable in $\mathcal{L}$.

Solution

[2]We need make this inference from automorphisms on $\mathcal{Q}$. Let $\sigma : \mathbb{Q} \to \mathbb{Q}$ be an automorphism. Then we have that

$$\begin{gathered} x < y \iff \sigma(x) < \sigma(y) \\ \sigma(0) = 0 \\ \text{and } \sigma \text{ is a bijection} \end{gathered}$$

[2] Question: why did we look at the automorphisms? **Answer:** By Corollary 46, automorphisms that fixes, in this case, $\emptyset$ pointwise, which is trivially true, should remap the defined set into an element of its own. We use this fact to obtain a contradiction to show that addition is not definable.

For example, the map $x \mapsto ax$, for some $a > 0$, is an automorphism on $\mathbb{Q}$.

Consider the function

$$\phi(x) = \begin{cases} 2x+1 & x < -1 \\ x & x \in [-1,1] \\ 2x-1 & x > 1 \end{cases}$$

which has a graph as in Figure 14.1.

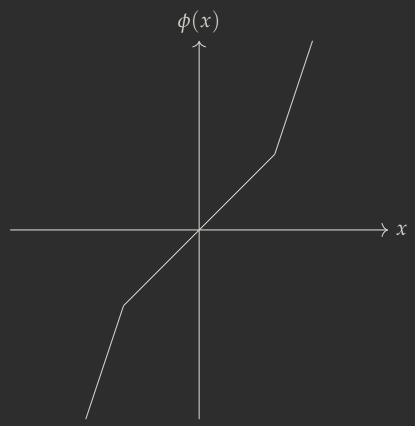

Figure 14.1: Graph of $\phi(x)$ in Example 14.1.2

It is rather clear that ϕ is indeed an automorphism, i.e. it has all the properties that we listed above. However, such an automorphism does not preserve addition itself, since

$$\begin{gathered} \phi(1+2) = \phi(3) = 5 \\ \phi(1) + \phi(2) = 4. \end{gathered}$$

Proposition 44 (Tarski-Vaught Test)

Suppose $\mathcal{M} \subseteq \mathcal{N}$. TFAE

1. $\mathcal{M} \preceq \mathcal{N}$
2. *For every $\mathcal{L}$-formula $\phi(\bar{x}, y)$ and all n-tuples $\bar{a} \in M^n$, if $\mathcal{N} \models \exists y \phi(\bar{a}, y)$, then there exists $b \in M$ such that $\mathcal{N} \models \phi(\bar{a}, b)$.*

Proof

$(1) \implies (2)$: Suppose $\mathcal{M} \preceq \mathcal{N}$. Then $\exists j : \mathcal{M} \to \mathcal{N}$ an elementary embedding such that for all $\mathcal{L}$-formula ϕ and $\bar{a} \in M^n$,

$$\mathcal{M} \models \phi(\bar{a}) \iff \mathcal{N} \models \phi(j(\bar{a})).$$

In particular, the identity map is such an embedding. Let $\psi(x) = \exists y \phi(\bar{x}, y)$. Then

$$\begin{aligned} \mathcal{N} \models \psi(\bar{a}) &\overset{\text{defn } 50}{\iff} \mathcal{M} \models \psi(\bar{a}) \\ &\iff \mathcal{M} \models \exists b \phi(\bar{a}, b) \\ &\overset{\text{defn } 50}{\iff} \mathcal{N} \models \exists b \phi(\bar{a}, b) \end{aligned}$$

This completes ($\implies$). $\dashv$

$(2) \implies (1)$: We shall use induction on the complexity of the $\mathcal{L}$-formula ϕ.

- If $\phi(\bar{a})$ is quantifier-free, then since $\mathcal{M} \subseteq \mathcal{N}$, by Proposition 42, we have $\mathcal{M} \models \phi(\bar{a}) \iff \mathcal{N} \models \phi(\bar{a})$. This case also covers the atomic formulas.

- If $\phi(\bar{a})$ is of the form $\phi_1(\bar{x}) \vee \phi_2(\bar{x})$, where $\phi_1(\bar{x}), \phi_2(\bar{x})$ are $\mathcal{L}$-formulas whose results are known, then

$$\begin{aligned} \mathcal{M} \models \phi(\bar{a}) &\iff \mathcal{M} \models \phi_1(\bar{a}) \text{ or } \mathcal{M} \models \phi_2(\bar{a}) \\ &\iff \mathcal{N} \models \phi_1(\bar{a}) \text{ or } \mathcal{N} \models \phi_2(\bar{a}) \quad \because \text{ IH} \\ &\iff \mathcal{N} \models \phi(\bar{a}). \end{aligned}$$

- If $\phi(\bar{x})$ is of the form $\phi_1(\bar{x}) \wedge \phi_2(\bar{x})$, the proof is the same as the previous item.

- If $\phi(\bar{x})$ is of the form $\neg\psi(\bar{x})$ for some $\mathcal{L}$-formula ψ, then

$$\begin{aligned} \mathcal{M} \models \phi(\bar{a}) &\iff \mathcal{M} \models \neg\psi(\bar{a}) \\ &\iff \mathcal{N} \models \neg\psi(\bar{a}) \quad \because \text{ IH} \\ &\iff \mathcal{N} \models \phi(\bar{a}). \end{aligned}$$

- As stated before in another proof, it suffices to prove for the existential case, since $\forall$ can be written as $\neg\exists\neg$. Suppose $\phi(\bar{x})$ is of the form $\exists y \psi(\bar{x}, y)$, for some $\mathcal{L}$-formula $\psi(\bar{x}, y)$ whose result we already know. We know that $\mathcal{M} \models \phi(\bar{a}) \implies \mathcal{N} \models \phi(\bar{a})$ by

Proposition 42. For the converse,

$$\begin{aligned}\mathcal{N} \models \phi(\bar{a}) &\iff \mathcal{N} \models \exists y \psi(\bar{a}, y) \\ &\implies \text{there exists } b \in M\ \mathcal{N} \models \psi(\bar{a}, b) \\ &\overset{\text{IH}}{\implies} \text{there exists } b \in M\ \mathcal{M} \models \psi(\bar{a}, b) \\ &\implies \mathcal{M} \models \exists y \psi(\bar{a}, y) \\ &\iff \mathcal{M} \models \phi(\bar{a}).\end{aligned}$$

This completes the proof. □

Remark

The Tarski-Vaught Test gives us an alternate mechanism to check if a substructure is elementary.

Theorem 45 (Downward Löwenheim-Skolem)

Suppose $\mathcal{M}$ is an $\mathcal{L}$-structure, and $A \subseteq M$. Then there exists an elementary substructure of $\mathcal{M}$ that contains A and is of cardinality at most $\max\{|A|, |\mathcal{L}|, \aleph_0\}$.

In particular, if $\mathcal{L}$ is countable, then every $\mathcal{L}$-structure has a countable elementary substructure.

Proof

Let $\kappa = \max\{|A|, |\mathcal{L}|, \aleph_0\}$. Define, recursively so, a countable chain of subsets of M,

$$A = A_0 \subseteq A_1 \subseteq \ldots \subseteq A_n \subseteq \ldots,$$

such that each of their cardinality is less than κ, such that for each $n \geq 0$, if $\phi(\bar{x}, y)$ is an $\mathcal{L}$-formula, and $\bar{a} \in A_n$ such that $\mathcal{M} \models \exists y \phi(\bar{a}, y)$, then it is clear that there exists $b \in A_{n+1}$ with $\mathcal{M} \models \phi(\bar{a}, b)$.

Given an A_n, we shall show a way to construct A_{n+1}. Requires clarification

Example 14.1.3

Consider the language $\mathcal{L}$ of rings and the $\mathcal{L}$-structure $\mathcal{C} = (\mathbb{C}, 0, 1, +, -, \times)$. We know that $\mathcal{Q} = (\mathbb{Q}, 0, 1, +, -, \times)$ is an $\mathcal{L}$-substructure of $\mathcal{C}$. It can be shown that the structure that is the **algebraic closure** of $\mathbb{Q}$ is an elementary $\mathcal{L}$-substructure of $\mathcal{C}$.[3]

Why? How can we show this?

[3]

15 Lecture 15 Oct 30th

15.1 First-order Logic (Continued 5)

15.1.1 Elementary Embeddings (Continued)

Example 15.1.1

$\mathcal{Q} = (\mathbb{Q}, 0, +, -)$ has no proper elementary subgroups.

Proof

Let $G \preceq \mathcal{Q}$. For a fixed $n > 0$, we know that

$$\mathcal{Q} \models \forall x \, \forall y \, (\underbrace{y + y + \ldots + y}_{n \text{ times}} = x)$$

$$\mathcal{Q} \models \exists x \, (x \neq 0)$$

So G must be some non-trivial divisible subgroup of $\mathcal{Q}$. Let $\frac{n}{m} \in G$, with $n \neq 0$. WMA[1] $n, m > 0$. Then $n \in G$ by n divisibility in G. Thus $1 \in G$ Thus $\mathbb{Z} \leq G \leq \mathbb{Q}$, and so[2] $G = \mathbb{Q}$ by divisibility.

[1] Short for **We May Assume**. We may indeed assume that both n and m are strictly positive, or we can just factor out -1.

How does this follow?

[2]

Remark

By Downward Löwenheim-Skolem, $(\mathbb{R}, 0, +, -)$ *has many proper elementary subgroups, e.g.* $\mathbb{Q}$.[3]

How so? Is it simply by taking $q + r$ for $q \in \mathbb{Q}$ and $r \in \mathbb{R} \setminus \mathbb{Q}$.

15.1.2 Parameters and Definable Sets

Example 15.1.2

In $(\mathbb{R}, <)$, the subset $(0, 1)$ is not $\mathcal{L}$-definable: we could have said that $(0 < x) \wedge (x < 1)$, but $0, 1 \notin \mathcal{L}$.

We would now like to rid ourselves of such a restriction.

Definition 51 (Parameters)

Suppose $\mathcal{L}$ is a language, $\mathcal{M}$ an $\mathcal{L}$-structure, and $B \subseteq M$. Let

$$\mathcal{L}_B := \mathcal{L} \cup \{\bar{b} : b \in B\}$$

be the language $\mathcal{L}$ extended with new constant symbols $\bar{b} \in B$. These additional symbols are called ***parameters****.*

We can consequently talk about $\mathcal{L}_B$-structures,

$$\mathcal{M}_B = (\mathcal{M}, \bar{b}^{\mathcal{M}} = b)_{b \in B}.$$

Remark
Since each of the $\bar{b}^{\mathcal{M}} = b \in M$, we will often just write $\mathcal{M}$ instead of putting a subscript of B for the $\mathcal{L}_B$-structure.[4]

3

Example 15.1.3

With parameters, in $(\mathbb{R}, <)_{\{0,1\}}$, we can use the $\mathcal{L}_{\{}0,1\}$-formula

$$(\bar{0} < x) \land (x < \bar{1}),$$

or more easily written as

$$(0 < x) \land (x < 1),$$

to define the interval $(0, 1)$.

Remark

1. *If $\phi(x_1, \ldots, x_n)$ is an $\mathcal{L}_B$-formula, then there exists an $\mathcal{L}$-formula*

$$\psi(x_1, \ldots, x_n, y_1, \ldots, y_m)$$

for some $m < \omega$, and $b_1, \ldots, b_m \in B$ such that

$$\phi(x_1, \ldots, x_n) = \psi(x_1, \ldots, x_n, \bar{b}_1, \ldots \bar{b}_m).$$

Moreover, given $a_1, \ldots, a_n \in M$,

$$\mathcal{M} \models \phi(a_1, \ldots, a_n) \iff \mathcal{M} \models \psi(a_1, \ldots, a_n, b_1, \ldots, b_m).$$

As mentioned in an earlier comment, we will often write

$$\psi(\vec{x}, \vec{\bar{b}}) \text{ as } \psi(\vec{x}, \vec{b}).$$

2. *Suppose $\mathcal{N} \subseteq \mathcal{M}$. Then $\mathcal{N} \preceq \mathcal{M}$ iff for every $\mathcal{L}_N$-sentence $\sigma = \phi(\bar{b}_1, \ldots, \bar{b}_m)$, for some $b_1, \ldots, b_m \in N$,*

$$\mathcal{N} \models \sigma \iff \mathcal{M} \models \sigma.$$

In particular, notice that

$$\mathcal{N} \models \sigma \iff \mathcal{N}_N \models \sigma \iff \mathcal{N}_N \models \phi(b_1, \ldots, b_m)$$
$$\mathcal{M} \models \sigma \iff \mathcal{M} \models \phi(b_1, \ldots, b_m).$$

Definition 52 (B-Definable)

*Let $\mathcal{M}$ be an $\mathcal{L}$-structure for some language $\mathcal{L}$, and $B \subseteq M$ a subset. A let $X \subseteq M^n$, for some $n < \omega$, is **D-definable** (or **definable over B**) if there is an $\mathcal{L}_B$-formula $\phi(x_1, \ldots, x_n)$ such that*

$$X = \{(a_1, \ldots, a_n) \in M^n : \mathcal{M}_B \models \phi(a_1, \ldots, a_n)\}.$$

*We say that X is **0-definable** if it is **$\emptyset$-definable**. We say that X is definable if it is **M-definable**, i.e. definable with parameters from anywhere in the universe. We say that X is **quantifier-free definable** (respectively **existentially definable** or **universally definable**) if there is a quantifier-free (respectively existential or universal) formula ϕ such that $X = \phi^{\mathcal{M}}$.*

Note

- *If X is definable, then it is B-definable for some $B \subseteq M$.*
- *A function $f : X \to Y$ is B-definable if $X \subseteq M^n$, $Y \subseteq M^m$ are B-definable, and*

  $$\Gamma(f) \subseteq X \times Y \subseteq M^{n+m}$$

  is B-definable.

Example 15.1.4

Let $\mathcal{L} = \{0, 1, +, -, \times\}$, and $\mathcal{R} = (\mathbb{R}, 0, 1, +, -, \times))$. We shall assume (in this course, "perversely" so) that a ring is commutative and unitary. What are the definable sets in $\mathcal{R}$?

Example 15.1.5 (Zero Sets of Polynomials)

Suppose $\mathcal{M} = (F, 0, 1, +, -, \times)$ where R is a field. Suppose $P_1, \ldots, P_l \in R[X_1, \ldots, X_n]$. The zero set of $\{P_1, \ldots, P_l\}$

$$V(P_1, \ldots, P_l) = \{a \in R^n : P_i(a) = 0, i = 1, \ldots, l\}$$

is called a **variety**. Such sets are called **algebraic sets** of R^n. They form the closed sets of a topology on R^n, called a **Zariski** topology, and the closed sets are called **Zariski's closed subsets**. We can define such a set using the $\mathcal{L}_R$-formula

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

It is clear that Zariski's closed subsets of R^n are quantifier-free definable in $\mathcal{R}$, with parameters from R.

More generally, any finnite **boolean combination**[5] of Zariski's closed subsets are quantifier-free definable. Sets that can be expressible as such are said to be **Zariski-constructible**.

[4] Is this saying that from hereon, we will somewhat be more loose on the constants of the language, in that we can use any of the constants from the underlying universe?

In fact, Zariski's closed subsets are exactly all of the quantifier-free definable sets.

Exercise 15.1.1

Let $B \subseteq R$ and S a subring generated by B. The $\mathcal{L}_B$-terms of $\mathcal{R}$ are precisely the polynomial functions over S, i.e. for any $\mathcal{L}_B$-term, $t(x_1, \ldots, x_n)$, there is $P_t \in S[X_1, \ldots, X_n]$ such that $t^{\mathcal{R}} = P_t$, as functions on R^n.

In fact, for any ring $A \supseteq S$, $t^{\mathcal{A}} = P_t$, where $\mathcal{A} = (A, 0, 1, +, -, \times)$.

Suppoer ϕ is an atomic $\mathcal{L}_R$-formula. Then ϕ can be

$$t(x_1, \ldots, x_n) = s(x_1, \ldots, x_n),$$

where t, s are $\mathcal{L}_R$-terms. Then by the exercise above, we have that

$$\phi^{\mathcal{R}} = \bigvee (P_t - P_s),$$

i.e. it is Zariski closed (or Z-closed). We see that quantifier-free formulas are Zariski-constructible (Z-constructible).

The following is a "fact" of which we shall see again later on.

Theorem

If $R = F$ is an algebraically closed field, then every definable set in $(F, 0, 1, +, -, \times)$ is Z-constructible.

This is a **quantifier-elimination theorem**, of which we shall study slightly later on in the course.

Remark

With regards to the unnumbered theorem above, the only definable subsetes of F are the finite and cofinite[6] sets.

[5] Boolean combinations include taking unions, intersections, and complements.

Example 15.1.6

In $(\mathbb{R}, 0, 1, +, -, \times)$, $(0, 1)$ is definable by $(0 < x) \wedge (x < 1)$ and $<$ is definable in the structure. However, we have that $<$ is not a quantifier-free definable set. This is because $\mathbb{R}$ as a field is not algebraically closed.

Fortunately, just so that we can have some sort of closure[7], we shall present the following theorem without proof.

[6]

Definition 53 (Cofinite)
A cofinite set is a set whose complement is a finite set.

Theorem (Macintyre)

If F is a field and the definable sets in $(F, 0, 1, +, -, \times)$ are all Z-constructible, then F is algebraically closed.

This gives us some sort of comfort knowing that we have some method of checking if a field is algebraically closed.

16 Lecture 16 Nov 01st

16.1 First-order Logic (Continued 6)

16.1.1 Parameters and Definable Sets (Continued)

Corollary 46 (Automorphisms on B-definable Sets)

Suppose $X \subseteq M^n$ be a B-definable set in $\mathcal{M}$, where $B \subseteq M$. Let

$$j \in \mathrm{Aut}_B(\mathcal{M}) := \{\sigma \in \mathrm{Aut}(\mathcal{M}) \mid \sigma \restriction_B = id_B\},$$

such that j acts on M^n coordinate-wise. Then $j(X) = X$.

Proof

Say X is defined by $\phi(\vec{x}, \vec{b})$ where $\phi(\vec{x}, \vec{y})$ is a $\mathcal{L}$-formula, $\vec{x} = (x_1, \ldots, x_n)$, $\vec{y} = (y_1, \ldots y_m)$, and $\vec{b} = (b_1, \ldots b_m) \in B^m$. In particular, $X = \{\vec{a} \in M^n \mid \mathcal{M} \models \phi(\vec{a}, \vec{b})$. Since j is elementary, we have

$$\begin{aligned} \vec{a} \in X &\iff \mathcal{M} \models \phi(\vec{a}, \vec{b}) \\ &\iff \mathcal{M} \models \phi(j(\vec{a}), j(\vec{b})) \\ &\iff \mathcal{M} \models \phi(j(\vec{a}), \vec{b}) \quad \because j \text{ fixes } B \text{ pointwise} \\ &\iff j(\vec{a}) \in X \end{aligned}$$

Therefore, $j(X) = X$ as required. □

16.1.2 Theories

Definition 54 ($\mathcal{L}$-Theory)

*An **$\mathcal{L}$-Theory** is simply a set of $\mathcal{L}$-sentences.*

Definition 55 (Model)

*A **model** of an $\mathcal{L}$-theory T is an $\mathcal{L}$-structure $\mathcal{M}$ such that $\mathcal{M} \models \sigma$ for all $\sigma \in T$. We denote this by $\mathcal{M} \models T$. A theory is said to be **consistent** if it has a model.*

Definition 56 ($\mathcal{L}$-elementary / $\mathcal{L}$-axiomatizable)

*A class K of $\mathcal{L}$-structures is said to be **$\mathcal{L}$-elementary** (or **$\mathcal{L}$-axiomatizble**) if there is an $\mathcal{L}$-theory T such that*

$$\mathcal{M} \in K \iff \mathcal{M} \models T.$$

Given T, the models of T is

$$\mathrm{Mod}(T) := [\![\ \mathcal{M} \mid \mathcal{M} \models T \ ]\!],$$

so K is $\mathcal{L}$-elementary if $K = \mathrm{Mod}(T)$.

Example 16.1.1

Let $\mathcal{L} = \{e, \cdot, \mathrm{inv}\}$. The following classes are elementary:[1]

- groups
- abelian groups
- divisible groups $(\forall x \, \exists y \, (y^n = x))$
  (note: this is infinitely axiomatizable)
- torsian free groups $(\forall x \, (x^n = e \to x = e))$
  (note: this is also infinitely axiomatizble)
- for any n, the class of groups of exponent n is elementary $(\forall x \, (x^n = e))$

[1] How do we know this?

The following are not elementary (which we shall prove later)

- all torsion groups
- all finite groups

Definition 57 (Theories of a Structure)

If $\mathcal{M}$ is an $\mathcal{L}$-structure, the theories of this structure is

$$\mathrm{Th}(\mathcal{M}) := \{\sigma : \mathcal{M} \models \sigma\},$$

where σ is an $\mathcal{L}$-sentence.

Definition 58 ($\mathcal{L}$-elementarily Equivalent)

Given two $\mathcal{L}$-structures $\mathcal{M}$ and $\mathcal{N}$, we sya that $\mathcal{M}$ is ***$\mathcal{L}$-elementarily equivalent*** *to $\mathcal{N}$ if*

$$\mathrm{Th}(\mathcal{M}) = \mathrm{Th}(\mathcal{N}).$$

We denote this by $\mathcal{M} \equiv \mathcal{N}$.

Note

If $j : M \to N$ is an elementary embedding, then $\mathcal{M} \equiv \mathcal{N}$. In particular, if $\mathcal{M} \simeq \mathcal{N}$, then $\mathcal{M} \equiv \mathcal{N}$.

However, the converse is not true: we know from the Downward Löwenheim Skolem that $\mathcal{M} \equiv \mathcal{N}$ but it is not necessary that $\mathcal{M} \simeq \mathcal{N}$.

Proposition 47

Let $\mathcal{M}$ and $\mathcal{N}$ be $\mathcal{L}$-structures. There exists an $\mathcal{L}$-elementary embedding $j : \mathcal{M} \to \mathcal{N}$ iff $\mathcal{N}$ has an expansion $\mathcal{N}'$ to $\mathcal{L}_M$ such that $\mathcal{N}' \models \mathrm{Th}(\mathcal{M}_M)$.

Proof

Given $j : \mathcal{M} \to \mathcal{N}$, let $\mathcal{N}'$ be the $\mathcal{L}_B$-structure that expands $\mathcal{N}$, given by

$$\mathcal{N}' = (\mathcal{N}, \bar{b}^{\mathcal{N}'} = j(b)),$$

for all $b \in M$. Then by the definition of an elementary $\mathcal{L}$-embedding, we must have $\mathcal{N}' \models \mathrm{Th}(\mathcal{M}_M)$.

Conversely, if there is an expansion $\mathcal{N}'$ of $\mathcal{N}$ on $\mathcal{L}_M$ such that $\mathcal{N}' \models \mathrm{Th}(\mathcal{M}_M)$, then we simply need to consider the mapping

$$j : \mathcal{M} \to \mathcal{N} \text{ by } b \mapsto b^{\mathcal{N}'},$$

which will give us the desired elementary $\mathcal{L}$-embedding. □

Example 16.1.2 ($\subseteq + \equiv \not\preceq$)

Let $\mathcal{N} = (\omega, <)$ and $\mathcal{M} = (\omega \setminus \{0\}, <)$ be $\mathcal{L} = \{<\}$ structures. Clearly so, $\mathcal{M} \subseteq \mathcal{N}$. They are also isomorphic: $j : \mathcal{N} \to \mathcal{M}$ by $j(n) = n + 1$ is an isomorphism. However, notice that

$$\mathcal{M} \models \neg \exists x \ (x < 1)$$
$$\mathcal{N} \models \exists x \ (x < 1),$$

so $\mathcal{M} \not\preceq \mathcal{N}$.

Definition 59 (Implication)

Suppose T is an $\mathcal{L}$-theory, and σ an $\mathcal{L}$-sentence. We say that T ***implies*** *σ (or T* ***entails*** *σ), denoted by $T \models \sigma$, if for every $\mathcal{M} \models T$, we have $\mathcal{M} \models \sigma$.*

Definition 60 (Complete)

An $\mathcal{L}$-theory T is said to be complete if for every $\mathcal{L}$-sentence σ,

$$\textit{either } T \models \sigma \textit{ or } T \models \neg\sigma.$$

Note
In words, a theory is complete if every $\mathcal{L}$-formula or its negation can be derived.

Example 16.1.3

$\text{Th}(\mathcal{M})$ is complete for any $\mathcal{L}$-structure $\mathcal{M}$.

Proof
For every $\mathcal{L}$-sentence σ, we have either $\sigma \in \text{Th}(\mathcal{M})$ or $\sigma \notin \text{Th}(\mathcal{M})$, which implies $\mathcal{M} \models \sigma$ and $\mathcal{M} \models \neg\sigma$, respectively.

Example 16.1.4

Let's go on a quest for a complete theory: Let $\mathcal{L} = \{0, 1, +, -, \times\}$ and T be the theory of rings. Then, let

$$\sigma = \forall x\ (x \neq 0 \to \exists y\ (xy = 1)),$$

which asserts that *every non-zero element has a multiplicative inverse*. However, in the models

$$\mathcal{Z} = (\mathbb{Z}, 0, 1, +, -, \times)$$
$$\mathcal{Q} = (\mathbb{Q}, 0, 1, +, -, \times)$$

we know that

$$\mathcal{Z} \models \neg\sigma \text{ while } \mathcal{Q} \models \sigma.$$

So $T \not\models \sigma$ and $T \not\models \neg\sigma$, i.e. T is incomplete.

Let's add the property of multiplicative inverses to our theory and call the new theory T_1. This is the theory of fields. Consider the $\mathcal{L}$-sentence

$$\sigma = \exists x\ (x^2 + 1 = 0),$$

which asserts that *every element has a square root*. However, we know that this is true in $\mathbb{C}$ but not in $\mathbb{R}$. Thus T_1 is still incomplete.

Let T_2 be the theory of algebraically closed fields (ACF), i.e. fields where every element is a root of some polynomial, and in particular every element will have a square root, mitigating the incompleteness above. However, the $\mathcal{L}$-sentence $\sigma = 1 + 1 = 0$ is true in $\overline{H}_2^{\text{alg}}$ but not true in $\mathbb{C}$.

Let T_3 be the theory of algebraically closed fields with characteristic p, with p either being prime or 0. This is a complete theory, but we are not ready to prove this.

Lemma 48 (Equivalence to a Complete Theory)

Suppose T is a consistent $\mathcal{L}$-theory. TFAE

1. *T is complete;*
2. *$\overline{T} := \{\sigma \mid T \models \sigma\}$, called **the set of consequences of T**, is maximally consistent;*
3. *$\overline{T} = \text{Th}(\mathcal{M})$ for some (equivalently any)*[2] *$\mathcal{M} \models T$;*
4. *Any 2 models of T are elementarily equivalent.*

What does this mean?

Proof

$(1) \implies (2)$: Suppose $\overline{T}$ is not maximal, i.e. $\overline{T} \subsetneq S$ for some S an $\mathcal{L}$-theory. Let $\sigma \in S \setminus \overline{T}$. Then $T \not\models \sigma$. But T is complete, and so $T \models \neg\sigma$. Thus $\neg\sigma \in \overline{T} \subsetneq S$. Then we have that $\{\sigma, \neg\sigma\} \subset S$ and so S has no models, i.e. S is inconsistent.

$(2) \implies (3)$: If $\mathcal{M}$ is a model of T, then $\mathcal{M}$ is a model of $\overline{T}$, thus $\overline{T} \subseteq \text{Th}(\mathcal{M})$. Since $\overline{T}$ is maximal, we must have $\overline{T} = \text{Th}(\mathcal{M})$.

$(3) \implies (4)$: Suppose $\mathcal{M} \models T$ such that $\overline{T} = \text{Th}(\mathcal{M})$. Let $\mathcal{N} \models T$. Then $\mathcal{N} \models \overline{T}$, and as above, we have $\overline{T} = \text{Th}(\mathcal{N})$. Therefore $\text{Th}(\mathcal{M}) = \text{Th}(\mathcal{N})$, i.e. $\mathcal{M} \equiv \mathcal{N}$.

$(4) \implies (1)$: Suppose $T \not\models \sigma$, where σ is an $\mathcal{L}$-sentence. Then there exists $\mathcal{M} \models T$ such that $\mathcal{M} \models \neg\sigma$. Let $\mathcal{N} \models T$. By assumption, since any two models of T are elementarily equivalent, we must also have $\mathcal{N} \models \neg\sigma$. Therefore $T \models \neg\sigma$. □

17 Lecture 17 Nov 06th

17.1 First-order Logic (Continued 7)

17.1.1 Theories (Continued)

Definition 61 (Partial $\mathcal{L}$-elementary Map)

Let $\mathcal{M}$ and $\mathcal{N}$ be $\mathcal{L}$-structures, $A \subseteq M$, and $f : A \to N$ a function. f is a ***partial $\mathcal{L}$-elementary map*** *(pem) if for all $\mathcal{L}$-formulae $\phi(x_1, \ldots, x_n)$, and for all $a_1, \ldots, a_n \in A$,*

$$\mathcal{M} \models \phi(a_1, \ldots, a_n) \iff \mathcal{N} \models \phi(f(a_1), \ldots, f(a_n)).$$

Remark

- $f : \emptyset \to N$ *is a pem iff* $\mathcal{M} \equiv \mathcal{N}$.
- $f : M \to N$ *is a pem iff* $\mathcal{M} \preceq \mathcal{N}$ *(i.e. f is an elementary embedding).*

Proposition 49 (Elementary Equivalence in Finite Stuctures)

Suppose at least one of $\mathcal{M}$ and $\mathcal{N}$ finite. Then

$$\mathcal{M} \equiv \mathcal{N} \iff \mathcal{M} \simeq \mathcal{N}.$$

Proof

Without loss of generality, suppose M is finite. Suppose $\mathcal{M} \equiv \mathcal{N}$ and $f : \emptyset \to N$ is a pem. We shall extend f, one-by-one, until $\mathrm{Dom}(f) = M$, so that f remains a pem and is an isomorphism

(Note: We must have $|M| = |N| = n$ for some $n < \omega$, since one of them is finite).

By induction, suppose that we have a pem

$$f : \{a_1, \dots, a_k\} \to N$$

for $0 \leq k < n$, and $\{a_1, \dots, a_k\} \subseteq M$. Let $X_{k+1,1}, \dots, X_{k+1,m_{k+1}} \subseteq M$ be the $\{a_1, \dots, a_k\}$-definable sets that contains a_{k+1}. Note that there are finitely many such subsets since $\mathcal{P}(M)$ is finite. Let

$$\phi_{k+1,j}(a_1, \dots, a_k, x) \quad 1 \leq j \leq m_{k+1}$$

be some (fixed) $\mathcal{L}$-formula that defines $X_{k+1,j}$, respectively so. With that, we have that

$$\mathcal{M} \models \bigwedge_{i=1}^{m_{k+1}} \phi_{k+1,j}(a_1, \dots, a_k, a_{k+1}),$$

which implies that

$$\mathcal{M} \models \exists x \left(\bigwedge_{i=1}^{m_{k+1}} \phi_{k+1,j}(a_1, \dots, a_k, x) \right).$$

Thus by the induction hypothesis and f being a pem,

$$\mathcal{N} \models \exists x \left(\bigwedge_{i=1}^{m_{k+1}} \phi_{k+1,j}(f(a_1), \dots, f(a_k), x) \right).$$

Let $b \in N$ satisfy this formula[1] and define $f' : \{a_1, \dots, a_k, a_{k+1}\} \to N$ such that

$$a_1 \mapsto f(a_i) \quad \text{for } 1 \leq i \leq k$$
$$a_{k+1} \mapsto b.$$

It is clear that f' is a pem and an isomorphism. Thus by induction, we have that $\mathcal{M} \simeq \mathcal{N}$. □

What are these other properties, and why should b not satisfy these other properties?

[1] During lecture, it was mentioned that b should not satisfy other properties.

Remark

This gives us an excuse to "ignore" finite structures when it comes to elementary equivalence, since it is no different from studying about isomorphisms.

17.2 Compactness

Theorem (Compactness Theorem)
Let $\mathcal{L}$ be a lagnauge, and T an $\mathcal{L}$-theory. T is consistent iff every finite subset of T is consistent.

The following is an equivalent formulation.

Corollary 50 (Corollary of Compactness Theorem)

Supppose T is an $\mathcal{L}$-theory for some langauge $\mathcal{L}$, and σ some $\mathcal{L}$-sentence. We have $T \models \sigma$ iff there exists a finite subset $\Sigma \subseteq T$ such that $\Sigma \models \sigma$.

Proof
It is clear that $T \models \sigma$ iff $T \cup \{\neg\sigma\}$ is inconsistent. □

In this course, we shall study a different proof for the Compactness Theorem, one that uses **ultraproducts**.

17.2.1 A proof of compactness using ultraproducts

Motivation Observe that $T = \bigcup_{n<\omega} \Sigma_n$, where $\Sigma_n \subseteq T$ is finite, and

$$\Sigma_0 \subseteq \Sigma_1 \subseteq \Sigma_2 \subseteq \ldots.$$

Since each Σ_n is consistent, there exists a model $\mathcal{M}_n \models \Sigma_n$. Ideally, the "limit" of these $\mathcal{M}_n$'s should be a model of T.

Definition 62 (Filter)

Let $I \in \mathrm{Set}$ and $\mathcal{F} \subseteq \mathcal{P}(I)$. We say that $\mathcal{F}$ is a filter on I if

1. $I \in \mathcal{F}$ *and* $\emptyset \notin \mathcal{F}$;
2. $A, B \in \mathcal{F} \implies A \cap B \in \mathcal{F}$;
3. $A \in \mathcal{F} \wedge A \subseteq B \implies B \in \mathcal{F}$.

Remark

It is reasonable to think of filters as a notion of "largeness" for subsets of I; indeed, since any superset of a smaller set in the filter is also in the filter itself.

Example 17.2.1

1. $\{\mathbb{R} \setminus X \mid X$ has Lebesgue measure $0\}$ is a filter on $\mathbb{R}$.
2. Let $I \in \mathit{Set}$, and $\kappa \leq |I| \in$ Card. Then

$$\mathcal{F} = \{A \subset I : |I \setminus A| < \kappa\}$$

   is a filter. In particular, if $I = \omega$, and $\kappa = \aleph_0$. Then the above filter, which is the set of **cofinite sets**, is called a **Fréchet filter**.

3. Let $I \in \text{Set} \setminus \{\emptyset\}$, and $a \in I$. The set

$$\mathcal{F}_a = \{A \subseteq I \mid a \in A\}$$

   is called a **principal filter** on I.

Definition 63 (Ultrafilter)

*An **ultrafilter** on I is a filter that is maximal.*

Example 17.2.2

It is not difficult to notice that principal filters are maximal: Suppose $\mathcal{F}_a$ is a principle filter on I based on $a \in I$, and suppose to the contrary that $\mathcal{F}_a \subsetneq \mathcal{G}$ for some filter $\mathcal{G}$ on I. Then let $A \in \mathcal{G} \setminus \mathcal{F}_a$. In particular, $a \notin A$. However, $\{a\} \in \mathcal{F}_a \subseteq G$. Then $\{a\} \cap A = \emptyset \in \mathcal{G}$ since $\mathcal{G}$ is a filter, but a filter should not contain $\emptyset$.

In general, it is rather diffcult to describe a non-principal ultrafilter. However, there is a tool that tells us the existence of ultrafilters at our disposal: Zorn's Lemma!

Lemma 51 (Existence of Ultrafilters)

Every filter on I extends to an ultrafilter on I.

Proof

Let $\mathcal{F}$ be a filter on I. Consider

$$\mathcal{Z} = \{\mathcal{H} \subseteq \mathcal{P}(I) \mid \mathcal{H} \text{ a filter }, \mathcal{F} \subseteq \mathcal{H}\}.$$

It is rather clear that $(\mathcal{Z}, \subseteq)$ is a poset, closed under unions of chains, and the union of filters is a filter. Thus the chains have an upper bound, and by Zorn's Lemma, there is a maximal filter, which is an ultrafilter. ☐

Lemma 52 (Equivalent Characterization of an Ultrafilter)

Let $\mathcal{F}$ be a filter on I. $\mathcal{F}$ is an ultrafilter iff

$$\forall A \subseteq I \, (A \subseteq \mathcal{F}) \vee (I \setminus A \in \mathcal{F}).$$

Proof

$(\Longleftarrow)$ Suppose to the contrary that $\mathcal{F} \subsetneq \mathcal{G}$ is a filter. Then let $A \in \mathcal{F} \setminus \mathcal{F}$. Then by assumption, we have $I \setminus A \in \mathcal{F} \in \mathcal{G}$, but that implies

$$\emptyset = A \cap (I \setminus A) \in \mathcal{G},$$

and so $\mathcal{G}$ cannot be a filter. Thus $\mathcal{F}$ is an ultrafilter.

$(\Longrightarrow)$ Suppose $\mathcal{F}$ is an ultrafilter. Let $A \subseteq I$, and suppose $A \notin \mathcal{F}$. WTS $I \setminus A \in \mathcal{F}$. Let

$$\mathcal{G} := \{X \subseteq I \mid Y \setminus A \subseteq X, \text{ for some } Y \in \mathcal{F}\}.$$

We shall verify that $\mathcal{G}$ is a filter: since $I \supseteq Y \setminus A$ for any $Y \in \mathcal{F}$, $I \in \mathcal{G}$. Since $\emptyset$ is not a superset of any set, $\emptyset \notin \mathcal{G}$. Suppose $B, C \in \mathcal{G}$. Then $B \supseteq Y_1 \setminus A$ and $C \supseteq Y_2 \setminus A$, for some $Y_1, Y_2 \in \mathcal{F}$. It is clear that $B \cap C = (Y_1 \cap Y_2) \setminus A$. Thus $B \cap C \in \mathcal{G}$. Suppose $B \in \mathcal{G}$ and $B \subseteq C$. Then there exists $Y_1 \in \mathcal{F}$ such that $C \supseteq B \supseteq Y_1 \setminus A$. Thus $C \in \mathcal{G}$. This proves our claim that $\mathcal{G}$ is a filter.

Quite clearly so, since $A \notin \mathcal{F}$, $\mathcal{F} \subseteq \mathcal{G}$. Since $\mathcal{F}$ is an ultrafilter, we have that $\mathcal{F} = \mathcal{G}$. Therefore $I \setminus A \in \mathcal{G} = \mathcal{F}$, as required. ☐

We are now ready to define what an **ultraproduct** is.

Definition 64 (Ultraproduct)

*Let $(\mathcal{M}_i : i \in I)$ be a sequence of $\mathcal{L}$-structures indexed by $I \neq \emptyset$, and $\mathcal{U}$ an ultrafilter. An **ultrafilter** is an $\mathcal{L}$-structure which is defined as follows: we denote the ultraproduct $\mathcal{M}$ as*

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

where

- *its universe is*
$$M = \bigtimes_{i \in I} M_i / E,$$
where E is an equivalence relation defined as
$$(a_i : i \in I) E (b_i : i \in I) \iff \{i \in I : a_i = b_i\} \in \mathcal{U};$$
- *its constants are*
$$c^{\mathcal{M}} = [(c^{\mathcal{M}_i} : i \in I)]$$
for $c \in \mathcal{L}^{\text{con}}$;
- *its n-ary functions are*
$$f^{\mathcal{M}}([\alpha_1], \ldots, [\alpha_n]) = [(f^{\mathcal{M}_i}(a_1(i), \ldots, a_n(i)) : i \in I)]$$
where $[\alpha_1], \ldots, [\alpha_n] \in M$ and $\alpha_j = (a_j(i) : i \in I)$, where $a_j(i) \in M_i$;
- *its k-ary relations are*
$$([\alpha_1], \ldots, [\alpha_n]) \in R^{\mathcal{M}} \iff \{i \in I : (a_1(i), \ldots, a_n(i)) \in R^{\mathcal{M}_i}\} \in \mathcal{U}.$$

Exercise 17.2.1

Verify that the functions and relations of the ultraproduct $\mathcal{M}$ is well-defined; in particular, the definition of the functions and relations are independent of the elements $[\alpha_i] \in M$.

Note

Note that in the definition of an ultraproduct, we may notice that, almost deceivingly so, $\mathcal{U}$ *need not be an ultrafilter.*

18 Lecture 18 Nov 08th

18.1 Compactness (Continued)

18.1.1 A Proof of Compactness Using Ultraproducts (Continued)

Recall our last note in the last lecture.

Example 18.1.1

Suppose that $\mathcal{F} = \{I\}$ is a filter (which it trivially is), then E is the trivial equivalence relation, and $\mathcal{M}$ would just be the "cartesian product" of the $\mathcal{M}_i$'s.

For instance, suppose that we have the language $\mathcal{L} = \{0, 1, +, -, \times\}$ and $I = \omega$. Then $\mathcal{F} = \{\omega\}$. Consider the $\mathcal{L}$-structure $\mathcal{M} = (\mathbb{Z}, 0, 1, +, -, \times)$. Then in

$$R = \bigtimes_{i \in I} (\mathbb{Z}, 0, 1, +, -, \times),$$

we have

$$(1, 0, 1, 0, \ldots) \cdot (0, 1, 0, 1, \ldots) = (0, 0, 0, 0, \ldots)$$

but neither $(1, 0, 1, 0, \ldots)$ nor $(0, 1, 0, 1, \ldots)$ are $(0, 0, 0, 0, \ldots)$. Therefore R is not an **integral domain**.

Suppose that we have, instead, constructed an ultraproduct

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

with an ultrafilter $\mathcal{U}$. Then in comparison, we have

$$[(1, 0, 1, 0, \ldots)] \cdot [(0, 1, 0, 1, \ldots)] = [(0, 0, 0, 0, \ldots)] = 0_{\mathcal{S}}.$$

Now if $[(1, 0, 1, 0, \ldots)] \neq 0_{\mathcal{S}}$, then any of the elements that is zero on the even entries would not be in $\mathcal{U}$. If $[(0, 1, 0, 1, \ldots)] \neq 0_{\mathcal{S}}$, then any of the elements that is zero on the odd entries would not be in $\mathcal{U}$. But these sets are complements of each other, and so one of them must be in $\mathcal{U}$ since $\mathcal{U}$ is an ultrafilter.

Theorem 53 (Łoś)

Suppose we have a sequence of $\mathcal{L}$-structures $(\mathcal{M}_i : i \in I)$, an ultrafilter $\mathcal{U}$, an ultraproduct of the $\mathcal{M}_i$'s $\mathcal{M} = \prod_{\mathcal{U}}$, an $\mathcal{L}$-formula $\phi(x_1, \ldots, x_n)$, and elements $[\alpha_1], \ldots, [\alpha_n] \in M$, where $\alpha_1, \ldots, \alpha_n \in \bigtimes_{i \in I} M_i$. Then

$$\mathcal{M} \models \phi([\alpha_1], \ldots, [\alpha_n]) \iff \{i \in I : \mathcal{M}_i \models \phi(\alpha_1(i), \ldots, \alpha_n(i))\} \in \mathcal{U}.$$

In particular, if $n = 0$, i.e. for any $\mathcal{L}$-sentence σ, we have

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

Proof

We shall first prove the following statement about $\mathcal{L}$-terms, of which we will need for the rest of the proof.

Claim If $t(x_1, \ldots, x_n)$ is an $\mathcal{L}$-term, then for any $[\alpha_1], \ldots, [\alpha_n] \in \mathcal{M}$, we have

$$t^{\mathcal{M}}([\alpha_1], \ldots, [\alpha_n]) = [(t^{\mathcal{M}_i}(\alpha_1(i), \ldots, \alpha_n(i)) : i \in I)].$$

Proof of claim This is clearly true for when our $\mathcal{L}$-terms is either constants for functions, simply from the way of which we have defined them. If our $\mathcal{L}$-term is a variable, i.e. if $t(x_i) = x_i$, then the result follows as the x_i's are variables for exactly any of the $[\alpha_j]$'s. $\dashv$

We shall prove Łoś's theorem by the complexity of $\phi(x_1, \ldots, x_n)$.

Atomic formulae If ϕ is of the form

$$t(x_1, \ldots, x_n) = s(x_1, \ldots, x_n),$$

then we have

$$\begin{aligned}
&\mathcal{M} \models \phi([\alpha_1], \dots, [\alpha_n]) \\
&\iff t^{\mathcal{M}}([\alpha_1], \dots, [\alpha_n]) = s^{\mathcal{M}}([\alpha_1], \dots, [\alpha_n]) \\
&\iff [(t^{\mathcal{M}_i}(\alpha_1(i), \dots, \alpha_n(i)) : i \in I)] = [(s^{\mathcal{M}_i}(\alpha_1(i), \dots, \alpha_n(i)) : i \in I)] \\
&\iff \{i \in I : t^{\mathcal{M}_i}(\alpha_1(i), \dots, \alpha_n(i)) = s^{\mathcal{M}_i}(\alpha_1(i), \dots, \alpha_n(i))\} \in \mathcal{U} \\
&\iff \{i \in I : \mathcal{M}_i \models \phi(\alpha_1(i), \dots, \alpha_n(i))\} \in \mathcal{U}.
\end{aligned}$$

If ϕ is of the form $(t_1, \dots, t_k) \in R^k$, where t_j's are $\mathcal{L}$-terms, then

$$\begin{aligned}
&\mathcal{M} \models \phi([\alpha_1], \dots, [\alpha_n]) \\
&\iff (t_1^{\mathcal{M}}([\alpha_1], \dots, [\alpha_n]), \dots, t_k^{\mathcal{M}}([\alpha_1], \dots, [\alpha_n])) \in R^{\mathcal{M}} \subseteq M^k \\
&\iff \{i \in I : (t_1^{\mathcal{M}_i}(\alpha_1(i), \dots, \alpha_n(i)), \dots t_n^{\mathcal{M}_i}(\alpha_1(i), \dots, \alpha_n(i))) \in R^{\mathcal{M}_i}\} \in \mathcal{U} \\
&\iff \{i \in I : \mathcal{M}_i \models \phi(\alpha_1(i), \dots, \alpha_n(i))\} \in \mathcal{U},
\end{aligned}$$

where the second $\iff$ is by the interpretation of R in $\mathcal{M}$.

Logical Operators If ϕ is of the form $\neg\psi$ for some $\mathcal{L}$-formula $\psi(x_1, \dots, x_n)$, then

$$\begin{aligned}
&\mathcal{M} \models \phi([\alpha_1], \dots, [\alpha_n]) \\
&\iff \mathcal{M} \not\models \psi([\alpha_1], \dots, [\alpha_n]) \\
&\overset{IH}{\iff} \{i \in I : \mathcal{M}_i \models \psi(\alpha_1(i), \dots, \alpha_n(i))\} \notin \mathcal{U} \\
&\iff \{i \in I : \mathcal{M}_i \not\models \psi(\alpha_1(i), \dots, \alpha_n(i))\} \in \mathcal{U} \quad \because \mathcal{U} \text{ is an ultrafilter} \\
&\iff \{i \in I : \mathcal{M}_i \models \phi(\alpha_1(i), \dots, \alpha_n(i))\} \in \mathcal{U}.
\end{aligned}$$

If ϕ is of the form $\phi_1 \vee \phi_2$, where $\phi_1(x_1, \dots, x_n), \phi_2(x_1, \dots, x_n)$ are $\mathcal{L}$-formulas, then

$$\begin{aligned}
&\mathcal{M} \models \phi([\alpha_1], \dots, [\alpha_n]) \\
&\iff \mathcal{M} \models \phi_1([\alpha_1], \dots, [\alpha_n]) \text{ and } \mathcal{M} \models \phi_2([\alpha_1], \dots, [\alpha_n]) \\
&\iff \{i \in I : \mathcal{M}_i \models \phi_1(\alpha_1(i), \dots, \alpha_n(i))\}, \\
&\qquad \{i \in I : \mathcal{M}_i \models \phi_2(\alpha_1(i), \dots, \alpha_n(i))\} \in \mathcal{U} \\
&\iff \{i \in I : \mathcal{M}_i \models \phi_1(\alpha_1(i), \dots, \alpha_n(i)), \\
&\qquad \mathcal{M}_i \models \phi_2(\alpha_1(i), \dots, \alpha_n(i))\} \in \mathcal{U} \\
&\iff \{i \in I : \mathcal{M}_i \models \phi(\alpha_1(i), \dots, \alpha_n(i))\} \in \mathcal{U}.
\end{aligned}$$

There is no need to show for the case of $\wedge$, as it is equivalent to showing $\neg \vee \neg$.

Quantifiers As mentioned in a few of the previous proofs in earlier propositions and theorems, it suffices to only prove for the existential case. Suppose ϕ is of the form $\exists y \psi(y)$, for some $\mathcal{L}$-formula $\psi(x_1, \ldots, x_n, y)$. Then

$$\begin{aligned} &\mathcal{M} \models \phi([\alpha_1], \ldots, [\alpha_n]) \\ &\iff \text{there exists } [\beta] \in M \;\; \mathcal{M} \models \psi([\alpha_1], \ldots, [\alpha_n], [\beta]) \\ &\iff \text{there exists } [\beta] \in M \\ &\qquad A_\beta := \{i \in I : \mathcal{M}_i \models \psi(\alpha_1(i), \ldots, \alpha_n(i))\} \in \mathcal{U} \end{aligned}$$

Now our goal is to show that

$$A_\beta \in \mathcal{U}$$

$$\iff$$

$$B := \{i \in I : \text{ there exists } b_i \in M_i \;\; \mathcal{M}_i \models \psi(\alpha_1(i), \ldots, \alpha_n(i), b_i)\} \in \mathcal{U}$$

We can, in fact, show that $A_\beta = B$. It is clear that $A_\beta \subset B$, since we may just label $b_i = \beta(i)$. For $B \subseteq A_\beta$, define $\beta : I \to \bigcup_{i \in I} M_i$ by

$$\beta(i) = \begin{cases} b_i & i \in B \\ e_i & i \notin B \end{cases}$$

where e_i is just some dummy constant. Then $\beta \in \underset{i \in I}{\times} M_i$.

Now if $i \in B$, we have that $\beta(i) = b_i$ such that $\mathcal{M}_i \models \psi(\alpha_1(i), \ldots, \alpha_n(i), b_i)$, and so

$$\mathcal{M}_i \models \psi(\alpha_1(i), \ldots, \alpha_n(i), \beta(i)),$$

i.e. $i \in A_\beta$. This completes the proof. $\square$

Example 18.1.2

The class of finite groups is not axiomatizable.

Proof

Let $G_n = (Z/n\mathbb{Z}, 0, +, -)$. Let $\mathcal{U}$ be an ultrafilter that extends the Fréchet Filter. Then, let

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

Since each of the G_n's satisfies the axioms of groups, by Łoś' Theorem, we have that $\mathcal{G}$ is also a group. Also, $\mathcal{G}$ is infinite. Let σ_n be the $\mathcal{L}$-sentence

$$\sigma_n = \exists x_1, \ldots, x_n \left(\bigwedge_{1 \leq i < j \leq n} x_i \neq x_j \right),$$

which is the sentence that says that "I have n elements (and I may have more)". Then we have

$$\{m < \omega : G_m \models \sigma_n\} \supseteq \{m < \omega : m \geq n\} \in \mathcal{U}$$

since the set in the middle is cofinite. Thus the first set is in $\mathcal{U}$. In particular, we have that

$$\mathcal{G} \models \sigma_n \implies |G| \geq n$$

for any $n < \omega$. Then there exists an $\mathcal{L}$-sentence σ such that it says that "I have at least $\aleph_0$-many elements (and I may have more)". This sentence cannot be satisfied by any of the finite structures.

19 Lecture 19 Nov 13th

19.1 Compactness (Continued 2)

19.1.1 A Proof of Comapactness Using Ultraproducts (Continued 2)

We shall introduce one of the special cases of an ultraproduct:

Definition 65 (Ultrapower)

Given $I \neq \emptyset$, $\mathcal{U}$ an ultrafilter on I, and a seqeunce $(\mathcal{M}_i : i \in I)$ of $\mathcal{L}$-structures such that $\mathcal{M}_i = \mathcal{M}$, some $\mathcal{L}$-structure. Then

$$\mathcal{M}' = \prod_{\mathcal{U}} \mathcal{M}$$

is called the ***ultrapower*** *of $\mathcal{M}$.*

Corollary 54 (Corollary of Łoś)

Given an $\mathcal{L}$-structure $\mathcal{M}$, the diagonal map

$$\delta : \mathcal{M} \to \prod_{\mathcal{U}} \mathcal{M} \text{ given by } a \mapsto [(a, a, \ldots)]$$

is an elementary $\mathcal{L}$-embedding.

Note

From this corollary, we see that ultrapowers are elementary extensions of the single $\mathcal{L}$-structure.

Also, if $\mathcal{U}$ is principal, then $\mathcal{M} \simeq \prod_{\mathcal{U}} \mathcal{M}$ [1]

[1] This may be an interesting or at least a relatively simple exercise.

Exercise 19.1.1
Show that if $\mathcal{U}$ is principal in Corollary 54, then $\mathcal{M} \simeq \prod_{\mathcal{U}} \mathcal{M}$.

Proof

To show that δ is an elementary $\mathcal{L}$-embedding, we shall show that $\mathcal{M} \preceq \prod_{\mathcal{U}} \mathcal{M}$. Let $\phi(x_1, \ldots, x_n)$ be an $\mathcal{L}$-formula, and let $d : M \to M^I$ be defined by

$$d(a_i) = (a_i, a_i, \ldots).$$

Then from $\mathcal{M}$, we have

$$\begin{aligned} &\mathcal{M} \models \phi(a_1, \ldots, a_n) \\ &\implies \{i \in I : \mathcal{M} \models \phi(d(a_1)(i), \ldots, d(a_n)(i))\} \\ &\overset{(*)}{=} \{i \in I : \mathcal{M} \models \phi(a_1, \ldots, a_n)\} \overset{(**)}{=} I \in \mathcal{U} \\ &\overset{\text{Łoś}}{\iff} \prod_{\mathcal{U}} \mathcal{M} \models \phi([d(a_1)], \ldots, [d(a_n)]) \\ &\iff \prod_{\mathcal{U}} \mathcal{M} \models \phi(\delta(a_1), \ldots, \delta(a_n)), \end{aligned}$$

where $(*)$ is because the two realizations give the same i's, and $(**)$ is because the i's do not matter in the condition, allowing all $i \in I$ to be in the set. This gives us that $\mathcal{L}$-formulas from the original structure is true in its ultrapower.

From $\prod_{\mathcal{U}} \mathcal{M}$, we have

$$\begin{aligned} &\prod_{\mathcal{U}} \mathcal{M} \models \phi(\delta(a_1), \ldots, \delta(a_n)) \\ &\overset{\text{Łoś}}{\iff} \{i \in I : \mathcal{M} \models \phi(d(a_1)(i), \ldots, d(a_n)(i))\} \in \mathcal{U}. \end{aligned}$$

Then, for any i_0 in the above set, we have

$$\mathcal{M} \models \phi(d(a_1)(i_0), \ldots, d(a_n)(i_0)) \implies \mathcal{M} \models \phi(a_1, \ldots, a_n).$$

This completes the proof. □

$\aleph_1$-compactness We shall take a little detour and visit an aside that utilizes Łoś' Theorem.

Definition 66 ($\aleph_1$-Compact)

We say that a set A is ***$\aleph_1$-compact*** *if any countable collection of definable* *non-empty sets in A, with the* ***finite intersection property***[2]*, then the collection has a non-empty intersection.*

[2] The finite intersection property is such that every finite non-empty subcollection non-empty sets has a non-empty intersection. See also notes on PMATH 351 (Real Analysis).

Proposition 55 (Ultraproducts with a Non-Principal Ultrafilter is $\aleph_1$-compact)

Suppose $(\mathcal{M}_i : i < \omega)$ is a sequence of $\mathcal{L}$-structures, and $\mathcal{U}$ a non-principal ultrafilter on ω. Then $\prod_{\mathcal{U}} \mathcal{M}_i$ is $\aleph_1$-compact.

Note

The given proof is not one that I can fully agree with*, since in the proof we may somehow assume that the sets are nested, which is not given in our assumption, and may not actually be a necessary condition. The provided proof will be recorded for archiving purposes.*

Proof

Let $\mathcal{M} = \prod_{\mathcal{U}} \mathcal{M}_i$, and $(F_i : i < \omega)$ a sequence of definable non-empty sets in $\mathcal{M}$, such that $F_i \neq \emptyset$. **Suppose that $F_j \subseteq F_i$ for $i < j$.** We may assume that $F_0 = M^n$, given by the $\mathcal{L}$-formula $x = x$. For $0 < i < \omega$, let $\phi_i(x)$ be the $\mathcal{L}$-formula that defines F_i. For each i, let

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

where we know that the set above is non-empty as F_0 is defined. Then we have that $0 \leq n_i \leq i$. Also for each i, let $a_i \in M_i$ be such that $\mathcal{M}_i \models \phi_{n_i}(a_i)$. Now let $a = (a_i : i < \omega)$. We want to show that for all $n < \omega$,

$$\mathcal{M} \models \phi_n([a])$$

so that

$$[a] \in \bigcap_{n<\omega} F_n.$$

Come back to this proof after the exam.

Leaving it behind for now to focus on the more important stuff. See also https://math.stackexchange.com/questions/3027612/ultraproducts-with-a-non-principal-ultrafilter-i [solved].

Example 19.1.1

$(\mathbb{Z}, <)$ is not $\aleph_1$-compact but $\prod_{\mathcal{U}}(\mathbb{Z}, <)$, where $\mathcal{U}$ is a non-principal ultrafilter, is $\aleph_1$-compact.

We are now ready to tackle on the problem which we have set out to solve.

Theorem 56 (Compactness Theorem)

Let T be an infinite $\mathcal{L}$-theory. If every finite subset of T is consistent, then T is consistent.

Proof

Let I be the collection of finite subsets of T. Then by assumption, we have that for $i \in I$,

$$\mathcal{M}_i \models i.$$

Consider the collection of supersets of i

$$X_i := \{i' \in I : i \subseteq i'\}.$$

The collection of these X_i's gives us the structure of a **lattice**:

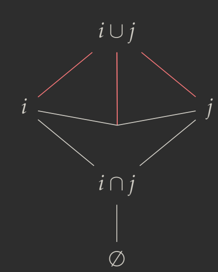

Figure 19.1: Lattice structure from the X_i's. Red lines are the parts captured by X_i.

[3] Consider

$$\mathcal{F} := \{Y \subseteq I : Y \supset X_i,\ i \in I\}.$$

Notice that $\mathcal{F}$ is a filter, as

- $I \in \mathcal{F}, \emptyset \in \mathcal{F}$;
- $\forall Y, Z \in \mathcal{F}\ \ Y \cap Z \supset X_{i \cup j} \implies Y \cap Z \in \mathcal{F}$; and
- $\forall Y \in \mathcal{F} \land Z \supset Y \implies Z \supset Y \supset X_i$ for some $i \in I \implies Z \in \mathcal{F}$.

Let $\mathcal{U} \supset \mathcal{F}$ be an ultrafilter[4]. Let $\mathcal{M} = \prod_{\mathcal{U}} M_i$.

To show that $\mathcal{M} \models T$, let $\sigma \in T$. Then $\{\sigma\} \in I$ and so $X_{\{\sigma\}} \in \mathcal{U}$ by construction. Then

$$X_{\{\sigma\}} = \{i \subset T \text{ finite } : \sigma \in i\}.$$

By assumption, for each $i \in X_{\{\sigma\}}$, we have $\mathcal{M}_i \models i$, and so in particular $\mathcal{M}_i \models \sigma$. Then in particular,

$$X_{\{\sigma\}} \subset S_\sigma := \{i \in I : \mathcal{M}_i \models \sigma\} \implies S_\sigma \in \mathcal{U}.$$

Need to show an example of $\mathcal{V}$ failing to "capture" supersets.

[3] Now if we simply consider the collection of these collections:

$$\mathcal{V} = \{X_i : i \in I\},$$

we notice that $\mathcal{V}$ is almost a filter, but not quite: we have

- $I = X_\emptyset \in \mathcal{V}$ and $\emptyset \notin \mathcal{V}$; and
- for $X_i, X_j \in \mathcal{V}$, we have $X_i \cap X_j = X_{i \cup j} \in \mathcal{V}$.

[4]

Exercise 19.1.2
Show that $\mathcal{U}$ is a non-principal ultrafilter.

Thus by Łoś we have $\mathcal{M} \models \sigma$ and so $\mathcal{M} \models T$ as required. $\square$

20 Lecture 20 Nov 15th

20.1 Compactness (Continued 3)

20.1.1 First Applications of Compactness

Example 20.1.1

Consider the empty language and K the class of all infinite $\mathcal{L}$-structures (as pure sets). Show that K is axiomatizable but not finitely so.

Proof

Bibliography

Cohen, P. J. (1963). The independence of the continuum hypothesis. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC221287/.

Moosa, R. (n.d.). *Set Theory and Model Theory*. (n.p.), 5th edition.

Index

Todo list