Contents

Chapter 1

Set Theory

我携一樽酒，独上江祖石。
自从天地开，更长几千尺。
举杯向天笑，天回日西照。
永愿坐此石，长垂严陵钓。
寄谢山中人，可与尔同调。

李白

1.1 Ordinal

Definition 1.1.1. A ***set*** is a collection of objects and for a set S we say $x \in S$ if x is an object inside S. In particular, sets must satisfy the following axioms:

1. (Empty Set Axiom) There exists a set which has no memebers, it is called the empty set and denoted by $\varnothing$.
2. (Extension Axiom) For each sets X, Y, we say $X = Y$ if and only if X and Y have the same memebers, i.e. $x \in X$ if and only if $x \in Y$.
3. (Pair Set Axiom) Given sets X, Y, there exists a set, denoted by $\{X, Y\}$, with the property only $X, Y \in \{X, Y\}$, i.e. $t \in \{X, Y\}$ if and only if $t = X$ or $t = Y$.
4. (Union Axiom) Given a set X, there exists a set, denoted by $\bigcup_X$, whose members are precisely the memebers of the memebers of X, i.e. $t \in \bigcup_X$ if and only if $t \in Y$ for some $Y \in X$.

Example 1.1.2 (Natural Numbers)**.** We define the natural numbers as follows:

1. $0 := \varnothing$, this is guaranteed by empty set axiom.
2. $1 := \{0\}$, this is guaranteed by pair set axiom, in particular, we note by pair set axiom there exists a set $\{0, 0\}$ and by extension axiom we see this is equal to $\{0\}$.

3. $2 := \{0, 1\}$ and so on.

In particular, suppose n is already defiend, then we define $S(n)$, the ***successor*** of n, to be the union of n and $\{n\}$, i.e. $S(n) = n \cup \{n\}$, which is equal $\bigcup_{\{n,\{n\}\}}$ where we know $\{n, \{n\}\}$ exists by pair set axiom.

At this point, we defined natural numbers, but we also want to know the set of all natural numbers, this requires more axioms.

Definition 1.1.3. ***definite conditions*** are following[1]:$x \in Y$ and $X = Y$ are both definite conditions, and if P, Q are definite conditions, then

1. "not P", i.e. $\neg P$, is definite condition.
2. "P and Q", i.e. $P \wedge Q$, is definite.
3. "P or Q", i.e. $P \vee Q$, is definite.
4. "for all x in P", i.e. $\forall x \in P$, is definite.
5. "there exists $x \in P$", i.e. $\exists x \in P$, is definite.

Then, only conditions arising as above in finitely many steps are definite conditions.

Remark 1.1.4. Note if P then Q is defined as $\neg P \vee Q$ and is denoted by $P \Rightarrow Q$. Hence, it should be clear what it means when we write $P \Leftrightarrow Q$.

Example 1.1.5. We will show all the four axioms defined above are definite conditions:

1. Empty set axiom: there exists a set $\varnothing$ such that $\neg(\exists Y, Y \in \varnothing)$.
2. Pair set axiom: given sets X, Y, there is a set $\{X, Y\}$ satisfying $\forall t(t \in \{x, y\} \Leftrightarrow (t = x) \vee (t = y))$.
3. Union axiom: if X is a set, then there exists a set $\bigcup_X$ satisfying $\forall t(t \in U_X \Leftrightarrow \exists Y((y \in X) \wedge (t \in Y)))$.

Definition 1.1.6. We say a operation $H(x)$ is ***definite*** if the condition $y = H(x)$ is definite.

Definition 1.1.7. One more set axiom:

1. (Infinity Axiom) There exists a set I that contains 0 and is preserved by the successor function S. Viz, I satisfy: $0 \in I$, and $\forall x \in I, S(x) \in I$. In particular, $S(x)$ is defined as $\exists y((y \in I) \wedge \forall t((t \in Y) \Leftrightarrow ((t \in x) \vee (t = x))))$ in our context. This shows $S(x)$ is a definite operation.

Definition 1.1.8. We call a set I ***inductive*** if it contains 0 and is closed under the successor operation S.

Example 1.1.9. Given a nonempty set X, show there exists a set $\bigcap_X$ satisfying $\forall t(t \in \bigcap_X \Leftrightarrow \forall y(y \in X \Rightarrow t \in Y))$.

Solution. Hmmm ♠

Remark 1.1.10. At this point, we may also write $\bigcup X$ or $\bigcap X$ to mean union and intersection, instead of $\bigcup_X$ or $\bigcap_X$. In addition, we see $X \cup Y$ is just $\bigcup\{X, Y\}$ and similar for intersection.

---

[1]this will be a temporary definition

Definition 1.1.11. We say $X \subseteq Y$ if $\forall t(t \in X \Rightarrow t \in Y)$.

Definition 1.1.12. Two more set axiom:

1. (Power Set Axiom) Given a set X there exists a set $\mathcal{P}(X)$ satisfying $\forall t(t \in \mathcal{P}(x) \Leftrightarrow \forall y(y \in t \Rightarrow y \in X))$.
2. (Bounded Separation Axiom) Suppose X is a set and P is a definite condition, then there exists a set Y satisfying $\forall t(t \in Y \Leftrightarrow ((t \in X) \wedge P(t)))$. We use the following notation $Y = \{t \in X : P(t)\}$.

Example 1.1.13. Note bounded in the above axiom is essential. Consider the Russell's paradox: consider $R := \{t : t \notin t\}$, this only differ from bounded separation axiom that we lack ambient set, i.e. our X in the above axiom. This lead to contradiction as if R is a set, then is $R \in R$? If $R \in R$ then $R \notin R$ and if $R \notin R$ then $R \in R$. Hence R is not a set.

Example 1.1.14. Now we construct the set of natural numbers, denoted by ω, as follows. Fix an inductive set I, define

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

One should try to show our definition of ω does not depend on I (note all we want is $S(0), S(1), \ldots$ and all inductive sets contains 0, start from there and remember ω is the smallest of all such inductive sets).

Definition 1.1.15. One more set axiom:

1. (Replacement Axiom) Suppose P is a definite binary condition such that for every set X there is a unique set Y such that $P(X, Y)$. Then, given any set A, there exists a set B with the property that $y \in B \Leftrightarrow \exists x((x \in A) \wedge P(x, y))$.

Definition 1.1.16. Now we have defined the Zermelo-Fraenkel set theory, which is the following eight axioms we defined above:

1. (Empty Set Axiom) There exists a set which has no memebers, it is called the empty set and denoted by $\varnothing$.
2. (Extension Axiom) For each sets X, Y, we say $X = Y$ if and only if X and Y have the same memebers, i.e. $x \in X$ if and only if $x \in Y$.
3. (Pair Set Axiom) Given sets X, Y, there exists a set, denoted by $\{X, Y\}$, with the property only $X, Y \in \{X, Y\}$, i.e. $t \in \{X, Y\}$ if and only if $t = X$ or $t = Y$.
4. (Union Axiom) Given a set X, there exists a set, denoted by $\bigcup_X$, whose members are precisely the memebers of the memebers of X, i.e. $t \in \bigcup_X$ if and only if $t \in Y$ for some $Y \in X$.
5. (Infinity Axiom) There exists a set I that contains 0 and is preserved by the successor function S. Viz, I satisfy: $0 \in I$, and $\forall x \in I, S(x) \in I$. In particular, $S(x)$ is defined as $\exists y((y \in I) \wedge \forall t((t \in Y) \Leftrightarrow ((t \in x) \vee (t = x))))$ in our context. This shows $S(x)$ is a definite operation.
6. (Power Set Axiom) Given a set X there exists a set $\mathcal{P}(X)$ satisfying $\forall t(t \in \mathcal{P}(x) \Leftrightarrow \forall y(y \in t \Rightarrow y \in X))$.
7. (Bounded Separation Axiom) Suppose X is a set and P is a definite condition, then there exists a set Y satisfying $\forall t(t \in Y \Leftrightarrow ((t \in X) \wedge P(t)))$. We use the following notation $Y = \{t \in X : P(t)\}$.
8. (Replacement Axiom) Suppose P is a definite binary condition such that for

every set X there is a unique set Y such that $P(X, Y)$. Then, given any set A, there exists a set B with the property that $y \in B \Leftrightarrow \exists x((x \in A) \wedge P(x, y))$.

together with one more axiom called the regularity axiom (we will not use that).

1.2 Classes and Definite Operations

Definition 1.2.1. A ***class*** is a collection of sets satisfying some definite condition.

Remark 1.2.2. This is just what we get when we apply unbounded separation. If P is a definite condition, then $[Z : P(Z)]$ is not necessary a set, but it is a class.

First of all, all sets are classes. Indeed, if x is a set, then $[t : t \in x]$ is a class which is identical to x. Second, some classes are sets. For example, $[z : z \in \omega]$ is a set, but we cannot make sure, when we write something down, it will always be set, which is our third point. Third, not all classes are sets. Consider $R = [t : t \notin t]$, the Russell class is not a set.

Definition 1.2.3. A ***proper class*** is a class that is not a set.

Example 1.2.4. Consider the universal class $\mathcal{U} = [t : t = t]$. This is not a set because if U is a set then $R = [t : t \notin t]$ would be a set by consider the bounded separation, i.e. we see $R = \{t \in U : t \notin t\}$ as set if U is a set. Hence U is not a set and a proper class.

Definition 1.2.5. Two classes are ***equal*** if and only if they have the same members.

Remark 1.2.6. Membership is a binary relation between a set and a class, i.e. $x \in Y$ means a set x inside a class Y (proper or not). It does not make sense to talk about a class as a member of another class.

Definition 1.2.7. Given sets x, y, we define the ***ordered pair*** $(x, y) := \{\{x\}, \{x, y\}\}$.

Example 1.2.8. Show that $(x, y) = (x', y')$ if and only if $x = x'$ and $y = y'$.

Solution. Hmmmmm. ♠

Definition 1.2.9. Suppose X, Y are classes, the ***Cartesian product*** $X \times Y$ is defined by $\{z : \exists x \exists y((x \in X) \wedge (y \in Y) \wedge (z = (x, y)))\}$.

Remark 1.2.10. If X, Y are sets, then $X \times Y$ should be a set as well. Indeed, by bounded separation axiom we see $X \times Y = \{z \in \mathcal{P}(\mathcal{P}(X \cup Y)) : \exists x, \exists y, x \in X \wedge y \in Y \wedge z = (x, y)\}$.

Definition 1.2.11. Given classes X, Y, by definite operation $f : X \to Y$, what we mean is a subclass $\Gamma(f)$ of Cartesian product $X \times Y$ such that for all $x \in X$ there exists unique $y \in Y$ such that $(x, y) \in \Gamma(f)$.

Remark 1.2.12. Hence, we are identifying the operation f with its graph $\Gamma(f)$. We often write $f(x) = y$ instead of $(x, y) \in \Gamma(f)$. In other word, we defined $f : X \to Y$ to be some particular subclass of the Cartesian product.

Example 1.2.13. Consider $S : \mathcal{U} \to \mathcal{U}$ where $\mathcal{U}$ is the universal class, i.e. class of all sets, and S is the successor function, i.e. $S(x) = x \cup \{x\}$. We will show S is a definition operation.

We see $\Gamma(S) = [z : z = (x, y) \wedge \forall t (t \in Y \Leftrightarrow (t \in x) \vee (t = x))]$ and $\Gamma(S)$ is a subclass of $U \times U$ that satisfy the "vertical line test", i.e. uniqueness of y in $S(x) = y$. Hence S is a definite operation as desired.

Remark 1.2.14. Suppose X, Y are sets and $f : X \to Y$ is a definite operation, then $\Gamma(f)$ is a subset of $X \times Y$. In fact, if B is a set and A is a class and $A \subseteq B$, then A is set. This is because the only reason that a class is not a set would be we don't have a bound, but if you live inside a set, then we get bounded separation instead of unbounded separation.

Indeed, let $A \subseteq B$ where B is set and A is class. Suppose $A = [z : P(z)]$ with P definite, then $A = \{z \in B : P(z)\}$ and hence by bounded separation we see A is a set.

Definition 1.2.15. Replacement Axiom Redefiend:

1. (Replacement Axiom) Suppose $f : \mathcal{U} \to \mathcal{U}$ where $\mathcal{U}$ is the universal class and f is definite operation, and $A \in \mathcal{U}$. Then there exists a set B satisfying $\forall t (t \in B \Leftrightarrow \exists a ((a \in A) \wedge (t = f(a))))$

1.3 Ordering on ω

Proposition 1.3.1 (Induction Principle). *Suppose $J \subseteq \omega$ and J is inductive set, then*

$$J = \omega$$

Proof. $J \subseteq \omega$ by assumption. Then by definition we see $\omega \subseteq J$ as desired. ♡

Lemma 1.3.2. *Suppose $n \in \omega$:*

1. *If $x \in n$, then $x \in \omega$. Viz, every element of ω is a subset of ω.*
2. *If $x \in n$ then $x \subseteq n$.*
3. *$n \notin n$.*
4. *$n = 0$ or $0 \in n$.*
5. *If $x \in n$ then $S(x) \in n$ or $S(x) = n$.*

Proof. We only prove some of them, but all of the proofs uses Induction Principle.

Let us try to prove (a). Let $J = \{n \in \omega : n \subseteq \omega\}$, it suffice to prove $J = \omega$. First note $0 \in J$ because $\varnothing \in A$ for all set A and hence $0 = \varnothing \subseteq A$. Suppose $n \in J$, then $S(n) = n \cup \{n\} \in \omega$. Since $n \subseteq \omega$ and $n \in \omega$ it follows $S(n) \subseteq \omega$. Thus J is inductive set in ω and hence $J = \omega$.

Now we prove (b). Let $J = \{n \in \omega : \forall x(x \in n \Rightarrow x \subseteq n)\}$. First note $0 \in J$ because it is true vacuously. Suppose $n \in J$ and say $S(n) = n \cup \{n\}$. Note $n \in S(n)$ and $n \subseteq S(n)$. Let $x \in S(n)$, if $x = n$ then $x \subseteq S(n)$ as desired. Otherwise, $x \in n \Rightarrow x \subseteq n$ since $n \in J$ and hence $x \subseteq S(n)$. Hence $J = \omega$ as desired. ♡

Remark 1.3.3. When we think about $(1), (2)$ and (3) carefully, these conditions imply that $(\omega, \in)$ is a strict poset[1] as we will see from the following proposition.

Proposition 1.3.4. *$(\omega, \in)$ is a strict partial ordered set.*

Proof. By (3) from the above lemma we see $n \notin n$. Now suppose $n, m \in \omega$ such that $n \in m$ and $m \in n$, by part (2) of above lemma we see $n \subseteq m$ and $m \subseteq n$ and hence $n = m$. Now suppose $l, m, n \in \omega$ such that $l \in m, m \in n$. Then we see $l \subseteq m$ and $m \subseteq n$, hence $l \subseteq n$ and so $l \in n$ as desired. ♡

Definition 1.3.5. A strict poset (E, R) is ***linear*** (or ***total***), if $x, y \in E$ we have xRy or yRx or $x = y$.

Proposition 1.3.6. *$(\omega, \in)$ is a linaer ordering.*

Proof. Fix $n \in \omega$, consider $J := n \cup \{m \in \omega : n \in m\} \cup \{n\} \subseteq \omega$ (this is because $n \subseteq \omega$ by part (1) of Lemma 1.3.2), we want to show $J = \omega$. Note if $n = 0$ then this is just part (4) of the above lemma 1.3.2. Hence we may assume $n \neq 0$.

If $n \neq 0$, then by Lemma 1.3.2 part (d) we see $0 \in n$, i.e. $0 \in J$. Now suppose $m \in J$ be arbitrary, then $m \in n, m \in \{m \in \omega : n \in m\}$ or $m \in \{n\}$. Consider three cases:

1. If $m \in n$, then by part (5) of Lemma 1.3.2 we see $S(m) \in n$ or $S(m) = n$. If $S(m) \in n$ then $S(m) \in J$ as desired. If $S(m) = n$ then $S(m) \in J$ as well.
2. If $n \in m$, then $n \in S(m)$, so $S(m) \in J$.
3. If $m = n$ then $S(m) = S(n)$ and hence $n \in S(m)$ as desired.

Thus J is inductive and the proof follows.

♡

Definition 1.3.7. A strict linear ordering (E, R) is well-ordered if every nonempty subset of E has a least element.

Proposition 1.3.8. *$(\omega, \in)$ is a well-ordering.*

Proof. Let $X \subseteq \omega$ which has no least element, we will prove $X = \varnothing$. Consider $J = \{n \in \omega : S(n) \cap X = \varnothing\}$. We will show $J = \omega$. First, we show $0 \in J$. Indeed, if not, then $S(0) \cap X \neq \varnothing$. However, this imply $0 \in X$ where 0 is a least element in all of ω, i.e. X has a least element, a contradiction. Hence $0 \in J$.

---

[1]Recall a strict partial ordering is a binary relation R on a set E satisfying the following condition: $\neg(xRx)$ for all $x \in E$, $xRy \wedge yRx \Rightarrow x = y$, and $xRy \wedge yRz \Rightarrow xRz$

Now fix $n \in J$, we need to show $S(n) \in J$. Consider $S(S(n)) \cap X$, since $n \in J$ we see $S(n) \cap X = \varnothing$. Hence, the only possible case is $S(n) \in X$. Together we see $S(n)$ is actually a least element in X, a contradiction. Thus we see $S(n) \notin X$. This means $S(S(n)) \cap X = \varnothing$ and so $S(n) \in J$ as desired. Hence we see $J = \omega$ by Induction Principle.

Now let any $n \in \omega$, since $J = \omega$ we see $n \in J$, then $S(n) \cap X = \varnothing$. However we see $n \in S(n)$ so $n \notin X$, i.e. $X = \varnothing$ as desired. ♡

1.4 Ordinal

Definition 1.4.1. An ***ordinal*** is a set α such that $x \in \alpha$ then $x \subseteq \alpha$ and $(\alpha, \in)$ is a well-ordering.

Example 1.4.2. We see ω is an ordinal and every $n \in \omega$ is also an ordinal (note well-ordering is inheritance to restriction so n is well-ordering, and $x \in n \Rightarrow x \subseteq n$ by our lemma 1.3.2.(2)). Those two ordinals are called finite ordinals.

Example 1.4.3. Show that not every subset of ω is a natural number.

Lemma 1.4.4. *Suppose β is an ordinal and $\alpha \subseteq \beta$ and $\alpha \neq \beta$. If α is an ordinal then $\alpha \in \beta$.*

Proof. Assume α is proper subset of β and both of them are ordinal. In general, when working with elements of an ordinal in this note, we will sometimes write $x < y$ instead of $x \in y$.

Let $D = \beta \backslash \alpha := \{x \in \beta : x \notin \alpha\}$, tthen $D \neq \varnothing$ as $\alpha \neq \beta$. Let $d \in D$ be a least element, as $D \subseteq \beta$ while β is well-ordering. We will show $\alpha = d$.

First we show $d \subseteq \alpha$. Suppose not, then we can find $x \in d \backslash \alpha$, then $x < d$ and $x \in d \backslash \alpha \subseteq \beta \backslash \alpha = D$, i.e. we get a contradiction as d should be a least element.

Now we show $\alpha \subseteq d$. Let $x \in \alpha$, since $x, d \in \beta$ where β is an ordinal, we see either $x < d$ or $d < x$ or $x = d$. First we see $x = d$ lead to contradiction as $x \in \alpha$ but $d \in \beta \backslash \alpha$, so they cannot be equal. Second, we see $d < x$ is impossible. We first present a ***WRONG*** argument: since $d < x$ and $x < \alpha$, so by transitivity we see $d < \alpha$, i.e. $d \in \alpha$, hence a contradiction. The cavity here is that we are assuming α is a member of β, then we can use transitivity, but thats exactly what we are trying to prove! Now we present the right argument: If $d < x$ then $d \in x$. However, we see $x \in \alpha$ by definition. Thus, we see $x \subseteq \alpha$ since α is an ordinal. Hence we see $d \in \alpha$, a contradiction.

Hence we must have $x < d$, i.e. $x \in d$ and so every element of α is inside d as desired. At this point, we have shown $\alpha = d$ and since $d \in D \subseteq \beta$, we get $\alpha \in \beta$ as desired. ♡

Definition 1.4.5. Let Ord denote the class of all ordinals (we need to check that being an ordinal is a definite condition).

Proposition 1.4.6.

1. *Every member of an ordinal is an ordinal.*
2. *No ordinal is a mebmer of itself.*
3. *If* $\alpha \in \mathrm{Ord}$ *then* $S(\alpha) \in \mathrm{Ord}$.
4. *If* $\alpha, \beta \in \mathrm{Ord}$ *then* $\alpha \cap \beta \in \mathrm{Ord}$.

Proof. Suppose α is ordinal and $z \in \alpha$. Suppose $x \in z$ and suppose $y \in x$. Since $z \in \alpha$, we see $z \subseteq \alpha$, and hence $x \in \alpha$. Similarly we see $y \in \alpha$ as well. So $x, y, z \in \alpha$ and $(\alpha, \in)$ is strict well-ordering, and so by transitivity we get $y \in z$. Hence, we have shown if $x \in z$ then $x \subseteq z$ as desired. Then we need to show $(z, \in)$ is well-ordering. However, since $z \subseteq \alpha$ we see $(z, \in)$ is well-ordering as α is. Hence $z \in \mathrm{Ord}$ as desired.

Now let $\alpha \in \mathrm{Ord}$. However, if $\alpha \in \alpha$ then $(\alpha, \in)$ is not strict ordering, a contradiction (if $\alpha \in \alpha$ then we see $\alpha \notin \alpha$ by strict ordering definition).

The rest are left as exercise. ♡

Proposition 1.4.7.

1. *If* $\alpha, \beta \in \mathrm{Ord}$ *then either* $\alpha = \beta$ *or* $\alpha \in \beta$ *or* $\beta \in \alpha$.
2. *Every set of ordinals is strictly well-ordered by* $\in$.
3. *Let* $E \subseteq \mathrm{Ord}$ *be a subset, then*

$$\sup E := \bigcup E \in \mathrm{Ord}$$

4. Ord *is a proper class.*

Proof. Note $\alpha \cap \beta \in \mathrm{Ord}$, $\alpha \cap \beta \subseteq \alpha$. Suppose $\alpha \cap \beta \neq \alpha$, then by a lemma we see $\alpha \cap \beta \in \alpha$. Similarly, if $\alpha \cap \beta \neq \beta$, then $\alpha \cap \beta \in \beta$. Hence we get $\alpha \cap \beta \in \alpha \cap \beta$, a contradiction. This says either $\alpha \cap \beta = \alpha$ or $\alpha \cap \beta = \beta$. By $\alpha \cap \beta = \alpha$ we see $\alpha \subseteq \beta$, hence either $\alpha = \beta$ or $\alpha \subsetneq \beta$. If $\alpha \subsetneq \beta$ then $\alpha \in \beta$ as desired. Do similar analysis to the case $\alpha \cap \beta = \beta$, we see we get either $\beta = \alpha$ or $\beta \subsetneq \alpha$ and hence $\beta \in \alpha$ as desired. This proves (1).

Now we prove (2). Let $E \subseteq \mathrm{Ord}$ is a set. Consider $(E, \in)$, then antireflexivity is by 1.4.6.2, i.e. we see no ordinal is a member of itself so $\alpha \notin \alpha$ as desired. To see anti-symmetry, suppose $\alpha, \beta \in E$ and $\alpha \in \beta$ and $\beta \in \alpha$, then we see $\alpha \subseteq \beta$ and $\beta \subseteq \alpha$ and hence they are equal. Finally, let $\alpha, \beta, \gamma \in E$, then we see $\alpha \in \beta$ and $\beta \in \gamma$ imply $\beta \subseteq \gamma$ and hence $\alpha \in \gamma$ as desired.

To see linearity, this is just part (1) of this proposition.

Finally, suppose $A \subseteq E$ and suppose $A \neq \varnothing$. Let $\alpha \in A$, now consider the following cases:

1. $\alpha \cap A = \varnothing$: this means α is least in A.

2. $\alpha \cap A \neq \varnothing$: let $A' = \alpha \cap A \subseteq \alpha$ is nonempty, so A' has a least element, say $a \in A'$ as α is an ordinal. Suppose $b \in A$ with $b \in a$, then $a \in \alpha$ imply $b \in \alpha$. Hence $b \in \alpha \cap A$ and so $b \in A'$. This is a contradiction as a is least in A". Hence a is least in A as desired.

Now we prove (3). First, $\sup E$ is a set of ordinals as we see memebers of an ordinal is ordinal (by 1.4.6.1). By part (2), we see $(\sup E, \in)$ is a strict well-ordering. Suppose $\alpha \in \sup E$, then $\alpha \in \gamma$ for all $\gamma \in E$. Since $\alpha \in \gamma$ we see $\alpha \subseteq \gamma$. If $x \in \alpha$ then $x \in \gamma \subseteq \bigcup E = \sup E$. This shows $\alpha \subseteq \sup E$ and hence $\sup E$ is an ordinal as desired.

Suppose Ord is a set, then by part (2), we see $(\mathrm{Ord}, \in)$ is a strict well-ordering. By 1.4.6.(1), if $a \in \mathrm{Ord}$ then $a \subseteq \mathrm{Ord}$. Therefore, the set $\mathrm{Ord} \in \mathrm{Ord}$. This is a contradiction to the fact that no ordinal is a member of itself.

♡

Remark 1.4.8. We will often write $\alpha < \beta$ to mean $\alpha \in \beta$ for $\alpha, \beta \in \mathrm{Ord}$.

Lemma 1.4.9.

1. *Let $\alpha \in \mathrm{Ord}$, then $\alpha < S(\alpha)$ and there is no ordinal in between.*
2. *Suppose E is a set of ordinals, then $\sup E$ is a least upper bound for E.*
3. *Given any set of ordinals E, there exists a least ordinal $\alpha \notin E$.*

Proof. To see (1), note by definition $S(\alpha) = \alpha \cup \{\alpha\}$ and so $\alpha \in S(\alpha)$ as desired. Next, if $\alpha \subseteq \beta \subseteq S(\alpha)$, then we see $\beta \in S(\alpha)$, so either $\beta < \alpha$ or $\beta = \alpha$ by definition of $S(\alpha)$ and the fact $\beta \subseteq S(\alpha)$ imply $\beta \in S(\alpha)$ or ordinals.

Next we show (2). Given $\alpha \in E$, hence $\alpha \subseteq \bigcup E$ where we see since $\bigcup E$ is an ordinal. Therefore we get $\alpha \in \bigcup E = \sup E$ as desired. To show it is least upper bound, let $\alpha < \sup E$, then we see $\alpha < \beta$ for some $\beta \in E$ by definition of union and the proof follows as we see anything less than $\sup E$ is not an upper bound.

To show (3), we cannot use well-ordering to get least elements of $\mathrm{Ord} \setminus E$ because this may be a proper class. Now we prove it using the right way. If we cannot find $\alpha \in \mathrm{Ord}$ such that $E \subsetneq \alpha$, then we can consider the set $\alpha \setminus E$. The least element of $\alpha \setminus E$ will be a least ordinal not in E. So, what α should we take? The first two natural answer should be $\sup E$ or $S(\sup E)$, but none of them works. However, taking $S(S(\sup E))$, this will give us the desired α as one should check.

♡

Definition 1.4.10. A ***successor ordinal*** is an ordinal of the form $S(\alpha)$ for some $\alpha \in \mathrm{Ord}$. If α is not a successor ordinal.

Example 1.4.11. Show $0, \omega$ are limit ordinals and every nonzero $n \in \omega$ is a successor ordinal.

1.5 Transfinite Induction/Recursion

Remark 1.5.1. Induction is a method to prove certain definite conditions holds about all Ord.

Theorem 1.5.2 (Transfinite Induction)**.** *Suppose $P(x)$ is definite condition such that if α is an ordinal and $P(\beta)$ is true for all $\beta < \alpha$ then $P(\alpha)$ is true. Then $P(\alpha)$ is true to all ordinals.*

Proof. Suppose P is not true for all ordinals and let $\alpha \in \mathrm{Ord}$ so $\neg P(\alpha)$. Let $D = \{\beta \leq \alpha : \neg P(\beta)\}$ be a set (this is set because $\beta \leq \alpha$ is bounded separation, i.e. $\beta \leq \alpha$ if and only if $\beta < S(\alpha)$ if and only if $\beta \in S(\alpha)$, which is a set). This set D is not empty because it contains α.

Let $\alpha_0 \in D$ be least (such α_0 exists as D is a set of ordinals and all sets of ordinal is well-ordered). Then for all $\beta < \alpha_0$ we must have $P(\beta)$ is true. However this imply $P(\alpha_0)$ is true and hence we get a contradiction. ♡

Corollary 1.5.2.1 (Transfinite Induction, Second Form)**.** *Let P be definite condition satisfying:*

1. *$P(0)$ is true.*
2. *For all ordinals β, if $P(\beta)$ is true then $P(S(\beta))$ is true.*
3. *For all limit ordinals, if $P(\beta)$ is true for all $\beta < \alpha$ then $P(\alpha)$ is true.*

Then P is true for all ordinals.

Proof. Immediate. ♡

Remark 1.5.3. Recursion is to construct a definite operation on Ord with some desired property.

Definition 1.5.4. Let $\mathcal{X}$ be the class of all functions whose domain is an ordinal. In other word,

$$\mathcal{X} := [\Gamma : \exists X, Y, X \in \mathrm{Ord}, \Gamma \subseteq X \times Y \text{ is a graph of function}]$$

Theorem 1.5.5. *Given definite operation $G : \mathcal{X} \to \mathcal{U}$ where $\mathcal{U}$ is the class of all sets. Then there is a unique definite operation $F : \mathrm{Ord} \to \mathcal{U}$ such that $F(\alpha) = G(F|_\alpha)$ for all $\alpha \in \mathrm{Ord}$.*

Proof. First, to prove this condition uniquely determine F is straightforward transfinite induction. Say $F' : \mathrm{Ord} \to \mathit{Sets}$ also satisfy this condition $F'(\alpha) = G(F'|_\alpha)$. Now given α, suppose $F(\beta) = F'(\beta)$ for all $\beta < \alpha$. Then we see $F|_\alpha = F'|_\alpha$. Hence we see $F(\alpha) = G(F|_\alpha) = G(F'|_\alpha) = F'(\alpha)$ as desired. By transfinite induction this tell us $F(\alpha) = F'(\alpha)$ for all $\alpha \in \mathrm{Ord}$.

Now we prove the existence. First we say a function t with domain an ordinal α is an α-function defined by G if for all $\beta < \alpha$ we have $t(\beta) = G(t|_\beta)$. So such functions are approximations to the operation F that we are after. First observe an α-function defined by G must be unique (using transfinite induction), and that for

$\alpha \leq \beta$, the α-function defined by G extends to the β-function defined by G, i.e. the graph of t_α is contained in t_β.

Now we show for each $\alpha \in \mathrm{Ord}$ there exists an α-function defined by G. To show this, we use second form of transfinite induction. For $\alpha = 0$, the empty function vacuously satisfies the condition. Suppose $\alpha \neq 0$ and there is an α-function t_α, then define

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

where the union here is really by identifying the function with its graph. Clearly t is a function with domain $S(\alpha)$. To see this is the $S(\alpha)$-function defined by G, note $t(\alpha) = G(t_\alpha) = G(t|_\alpha)$ while for $\beta < \alpha$ we have $t(\beta) = t_\alpha(\beta) = G(t_\alpha|_\beta) = G(t|_\beta)$.

Finally, suppose $\alpha \neq 0$ is a limit ordinal and that for each $\beta < \alpha$ there is a β-function, say t_β, that is defined by G.

We claim the operation $t \mapsto t_\alpha$ is a definite operation. Indeed, $y = t_\alpha$ if and only if y is a function whose domain is α and for all $x < \alpha$, $y(x) = G(y|_x)$. In order to verify that this is a definite condition it suffice to show being a function with domain α is definite, and that given a function y the operation $x \mapsto y|_x$ is definite. Both are left as an straightforward exercise.

Now, since the operation $\beta \mapsto t_\beta$ is definite, we see by replacement axiom we get $T := \{t_\beta : \beta < \alpha\}$ is a set. Let $t = \bigcup T$, to verify that t is the α-function defined by G, we need to show:

1. t is a function on α. This follows from the above observation that t_β form a chain of extension.
2. t is the α-function defined by G: for all $\beta < \alpha$, note $S(\beta) < \alpha$ as α is limit ordinal, and $t(\beta) = t_{S(\beta)}(\beta) = G(t_{S(\beta)}|_\beta) = G(t|_\beta)$.

Hence, at this point, we have proved our claim that for each $\alpha \in \mathrm{Ord}$ there exists an α-function defined by G. Let us denote by t_α the α-function defined by G.

Now define $F(\alpha) = t_{S(\alpha)}(\alpha)$ for all ordinals α. By above claim and the fact S is definite, we see F is a definite operation. Note that for any $\beta < \alpha$, since $t_{S(\alpha)}$ extends $t_{S(\beta)}$ we have $t_{S(\alpha)}(\beta) = t_{S(\beta)}(\beta) = F(\beta)$. That is, $F|_\alpha = t_{S(\alpha)}|_\alpha$ for all $\alpha \in \mathrm{Ord}$. Hence

$$F(\alpha) = t_{S(\alpha)}(\alpha) = G(t_{S(\alpha)}|_\alpha) = G(F|_\alpha)$$

as desired. ♡

Corollary 1.5.5.1 (Transfinite Recursion, Second Form). *Suppose G_1 is a set, $G_2 : \mathcal{U} \to \mathcal{U}$ is a definite operation between class of sets to class of sets. Let $G_3 : \mathcal{X} \to U$ where $\mathcal{X}$ is the definite operation where $\mathcal{X}$ is the class of all functions on ordinals.*

Then there exists a unique definite operation $F : \mathrm{Ord} \to \mathcal{U}$ satisfying:

1. *$F(0) = G_1$.*
2. *For all ordinals α, $F(S(\alpha)) = G_2(F(\alpha))$.*
3. *For all limit ordinals $\alpha > 0$, $F(\alpha) = G_3(F|_\alpha)$.*

Proof. Let $G : \mathcal{X} \to \mathcal{U}$ be

$$G(f) = \begin{cases} G_1, & \text{if } t = \varnothing \\ G_2(f(\alpha)), & \text{if the domain of } f = S(\alpha) \\ G_3(f), & \text{if domain of } f \text{ is limit ordinal} \end{cases}$$

Now apply transfinite recursion to this G to produce this unique F so $F(\alpha) = G(F|_\alpha)$ for all $\alpha \in \mathrm{Ord}$. ♡

1.6 Ordinal Arithmetic

Definition 1.6.1. Fix $\beta \in \mathrm{Ord}$, we will define $\alpha + \beta$ for all $\alpha \in \mathrm{Ord}$ by transfinite recursion as follows:

1. $\beta + 0 := \beta$.
2. $\beta + S(\alpha) = S(\beta + \alpha)$.
3. If α is limit ordinal, then we define $\beta + \alpha = \sup\{\beta + \gamma : \gamma < \alpha\}$.

Remark 1.6.2. In terms of the second version of transfinite recursion, $G_1 = \beta$, G_2 is suppose be a the successor function, which is a definite operation between $\mathcal{U}$ and $\mathcal{U}$. Then G_3 is the sup of $\operatorname{Im} f$ for f a function with domain an ordinal. Hence, by the transfinite recursion, there is a unique definite operation from Ord to $\mathcal{U}$ which maps α to $\beta + \alpha$. We do this to every single β, then we get addition on Ord.

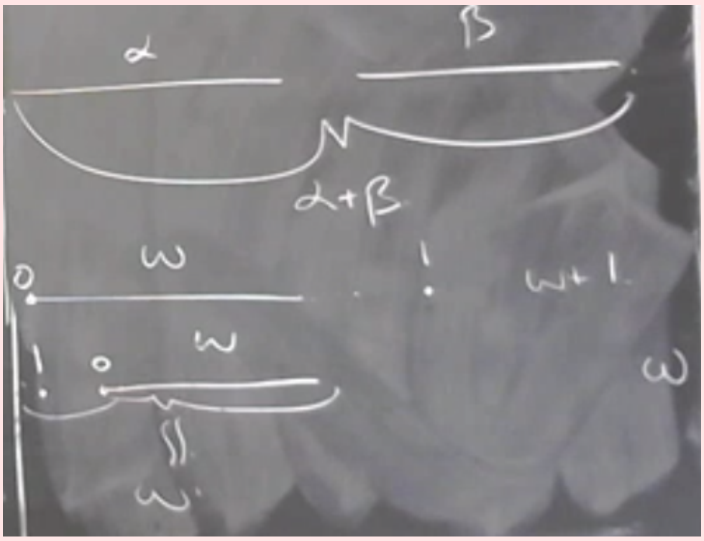
Example 1.6.3. We have $\beta + 1 = \beta + S(0) = S(\beta + 0) = S(\beta)$. Also, we have $0, 1, 2, 3, \ldots, \omega, \omega + 1, \ldots, \omega + \omega, \ldots$ where we see

$$\omega + \omega = \sup\{\omega + n : n < \omega\}$$

However, we note this addition is not commutative! Indeed, observe

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

This is because ordinal sum concatenates the order type of the ordinals:



Definition 1.6.4. Fix $\beta \in \mathrm{Ord}$, then we define:

1. $\beta \cdot 0 := 0$.
2. $\beta \cdot S(\alpha) := \beta\alpha + \beta$.
3. For limit ordinal α we have $\beta \cdot \alpha := \sup\{\beta \cdot \gamma : \gamma < \alpha\}$.

Remark 1.6.5. In other word, we have $G_1 = 0$, $G_2(x) = x + \beta$, and $G_3 = \sup(\operatorname{Im} f)$.

Example 1.6.6. We have

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

In fact, we have $\beta \cdot 1 = \beta$ for all ordinals β. On the other hand,

$$\beta \cdot 2 = \beta \cdot S(1) = \beta \cdot 1 + \beta = \beta + \beta$$

This can go on for all $n \in \omega$. Now, we note

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

We finally remark that multiplication is not commutative as well. For instance, we see

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

Example 1.6.7. Exercise: what is the order type of $\beta \cdot \alpha$ in terms of the well-ordering β and α.

Proposition 1.6.8. *Suppose $\alpha, \beta, \delta \in \mathrm{Ord}$, then:*

1. *$\alpha < \beta$ if and only if $\delta + \alpha < \delta + \beta$.*
2. *$\alpha = \beta$ if and only if $\delta + \alpha = \delta + \beta$.*
3. *$(\alpha + \beta) + \delta = \alpha + (\beta + \delta)$.*
4. *If $\delta \neq 0$ then $\alpha < \beta$ if and only if $\delta\alpha < \delta\beta$.*
5. *If $\delta \neq 0$ then $\alpha = \beta$ iff $\delta\alpha = \delta\beta$.*
6. *$(\alpha\beta)\delta = \alpha(\beta\delta)$*

1.7 Well-Ordering and Ordinals

Lemma 1.7.1. *Well-ordering is rigid, i.e. the only automorphism is the identity.*

Proof. Suppose $(E, <)$ is an well-ordering and let $f : (E, <) \to (E, <)$ be an automorphism, i.e. f is a bijection and $f(x) < f(y) \Leftrightarrow x < y$. Let $D = \{x \in E : f(x) \neq x\}$ and assume $f \neq \mathrm{Id}$ for a contradiction, i.e. D is not empty. Let $a \in D$ be a least element (exists since well-ordering).

If $f(a) < a$ then we see $f(a) \notin D$ and hence $f(f(a)) = f(a)$ since D are the set of things not fixed. However, since f is bijection, we get $f(a) = a$ by apply inverse to it and so $a \notin D$, a contradiction.

If $f(a) > a$. Apply f^{-1} we get $f^{-1}(a) < a$. Thus we see $f^{-1}(a)$ is not in D and so $f(f^{-1}(a)) = f^{-1}(a)$ so we get $a = f^{-1}(a)$ and so $a = f(a)$, a contradiction. ♡

Lemma 1.7.2. *A well-ordering $(E, <)$ is not isomorphic to any proper initial segment $E_b := \{x \in E : x < b\}$ for any $b \in E$.*

Proof. We want to show $(E, <) \ncong (E_b, <)$ for some b. Suppose $f : (E, <) \to (E < b, <)$ is an isomorphism for contradiction. Now let $D = \{x \in E : f(x) \neq x\}$ and this is not empty because $f(b) \in E_b$ where $b \notin E_b$. Now let $a \in D$ be a least element, one can show both possibilities are impossible. ♡

Lemma 1.7.3. *Let $(E, <)$ be a well-ordering and $f : (E, <) \to (\alpha, \in)$ be an isomorphism between orders where $\alpha \in \mathrm{Ord}$. Then f and α are both unique.*

Proof. If there is another $g : (E, <) \to (\alpha, \in)$, then we get $g^{-1} \circ f$ must be an isomorphism of orders between E and E. However the only automorphism for E is the identity, i.e. $g = f$ as desired.

Suppose $g : (E, <) \to (\beta, \in)$ is an isomorphism where $\beta \in \mathrm{Ord}$. If $\alpha \neq \beta$ then either $\alpha \in \beta$ or $\beta \in \alpha$, WLOG assume $\alpha \in \beta$. However then by homework we see $(\alpha, \in) = (\beta_\alpha, \in)$ where β_α is an strict initial segment in β. Now note composite $g^{-1} : \beta \to E$ and $f : E \to \alpha$ give us an isomorphism of β to α where α is an initial segment of β, i.e. this is impossible as we showed in above lemma. Hence we indeed have both f and α are unique. ♡

Theorem 1.7.4. *Every strict well-ordering is isomorphic to an ordinal. Moreover, the ordinal and the isomorphism is unique.*

Proof. By above lemma, it suffice to show existence of some isomorphism. Note we may assume E is not empty as the empty well-ordering is isomorphic to $0 = \varnothing \in \mathrm{Ord}$. Consider the set $A = \{x \in E : \exists \alpha_x \in \mathrm{Ord}, (E_x, <) \cong (\alpha_x, \in)\}$ and we note A is not empty because for a least element $e \in E$ we have $E_e \cong 0$. Now why is A is a set? Well, A is bounded and we only need to show $P(x)$ is definite condition where $P(x)$ if and only if $\exists \alpha_x \in \mathrm{Ord}, (E_x, <) \cong (\alpha_x, \in)$. This is left as an exercise.

Now let $f : E \to \mathrm{Ord}$ be the definite operation such that $(E_x, <) \cong (f(x), \in)$ for all $x \in A$. We see easily this is well-defined and one should check this is definite operation.

By replacement, $\mathrm{Im}(f)$ is a set of ordinals. Let $\alpha \in \mathrm{Ord}$ be least such that $\alpha \notin \mathrm{Im}(f)$. We claim $f : A \to \mathrm{Im}(f)$, $\mathrm{Im}(f) = \alpha$, $E = A$ and f is an isomorphism. ♡

Chapter 2

Cardinality

风回小院庭芜绿，柳眼春相续。凭阑半日独无言，依旧竹声新月似当年。
笙歌未散尊罍在，池面冰初解。烛明香暗画堂深，满鬓青霜残雪思难任。

李煜

2.1 Equinumerosity

Definition 2.1.1. Two set A, B are said to be ***equinumerous*** if there exists a bijection $f : A \to B$.

Proposition 2.1.2 (Schroeder-Bernstein). *A and B are equinumerous if and only if there exists injection $f : A \to B$ and $g : B \to A$.*

Proof. Say we have a composition $f : A \hookrightarrow B \hookrightarrow A$, so it suffice to prove that if $f : X \to X$ is injection from a set to itself, and $X \supseteq Y \supseteq f(X)$ then X and Y are equinumerous.

To show this, consider $Z = \bigcup_{i \geq 0} f^i(X) \backslash f^i(Y)$ and let $W = X \backslash Z$. Then we see $X = Z \cup W$ and moreover, this is a disjoint union. We should check we have disjoint union $Y = f(Z) \cup W$. Then we consider the map $i : X \to Y$ given by $z \mapsto f(z)$ and $w \mapsto w$ for $z \in Z$ and $w \in W$. One should check this is a bijection. ♡

Definition 2.1.3. We say a set A is:

1. ***finite*** if it is equinumerous to some $n \in \omega$.

2. ***countable*** if A is equinumerous with ω or finite.
3. ***uncountable*** if A is not countable.

Lemma 2.1.4. *Let $\alpha \in \mathrm{Ord}$ be infinite, then α and α are equinumerous.*

Proof. Define $f : \alpha + 1 \to \alpha$ to be

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

and this is clearly a bijection. ♡

Definition 2.1.5. A ***cardinal*** is an ordinal that is not equinumerous with any strictly lesser ordinal.

Remark 2.1.6. By above lemma we get all infinite successor ordinals are not cardinal. Viz, all infinite cardinals are limit ordinals. However, the converse is false, i.e. not all limit ordinals are cardinals. E.g. take $\omega \cdot 2$ and this is equinumerous with ω.

Example 2.1.7. Show $n \in \omega$ and ω are both cardinals.

Proposition 2.1.8. *Every set E there is a unique cardinal $h(E)$, which is the least ordinal not equinumerous with any subset of E.*

Proof. We show the class $X = [\alpha \in \mathrm{Ord} : \alpha \text{ is equinumerous with a subset of } E]$ is a set. Consider the collection of well-orderings on A, $W := \{(A, <) : A \subseteq E \text{ and } < \text{ is well-ordering on } A\}$, which is a set by bounded separation (as $(A, <) \in \mathcal{P}(E) \times \mathcal{P}(E \times E)$).

Let F be the definite operation on W such that $F(A, <)$ is the unique ordinal that is isomorphic to $(A, <)$. Now by replacement, we see the image of F, $\mathrm{Im}(F) \subseteq \mathrm{Ord}$, is a set. One should check $\mathrm{Im}\, F = X$ and hence we see X is a set as desired.

Now we let $h(E)$ be the least ordinal not in X. This is exactly what we are looking and the only thing left to check is to show $h(E)$ is cardinal. However, if $\alpha < h(E)$ then $\alpha \notin X$ and so α is equinumerous with some subset A of E. Hence, we cannot have a bijection between α and $h(E)$ as that will give us a contradiction. ♡

Remark 2.1.9. By the above proposition, we see we constructed cardinal $h(E)$ for every set E. In particular, $h(\omega) > \omega$, i.e. this is our first uncountable cardinal. However, that is not good enough.

Now we want every set to be equinumerous with some cardinal. However, this would imply every set is equinumerous with an ordinal and this bijection would actually give us a well-ordering on the set. In other word, if we can show every set is equinumerous with some cardinal, we get the conclusion that every set admits a well-ordering.

However, you may have seen this before, and this is exactly the well-ordering principle, which is something we cannot prove within ZF axiom, i.e. we need axiom of choice.

2.2 Axiom Of Choice

Definition 2.2.1. For any set F, a ***choice function on*** F is a function $c : F \to \bigcup F$ such that $\forall f \in F$, we have $c(f) \in f$.

Definition 2.2.2 (Axiom of Choice/AC)**.** Note this is an axiom. For any set F, if $\varnothing \notin F$, then there exists a choice function on F.

Remark 2.2.3. First we note AC is a set existence axiom, i.e. there is a set $\Gamma \subseteq F \times \bigcup F$ such that for all $f \in F$ we can find unique $a \in \bigcup F$ such that $(f, a) \in \Gamma$ and $a \in f$.

Second, we don't always need AC to find a choice function.

Example 2.2.4. Suppose $F = \{f\}$ with $f \neq \varnothing$. Then let $a \in f$, consider the set $\Gamma = (f, a)$. This is a choice function and we can do this in ZF, i.e. we don't need AC. In general, in finite set, we probably don't need AC.

On the other hand, say $(A, <)$ is a well-ordering, let $F = \mathcal{P}(A) \backslash \{\varnothing\}$. One should try to prove that there exists $c : F \to \bigcup F$ where $c(f)$ is the $<$-least element of f. We don't need AC for this as well.

Theorem 2.2.5. *In ZF, the following are equivalent:*

1. *Axiom of Choice.*
2. *Well-ordering Principle: every set admits a well-ordering.*
3. *Zorn's lemma: suppose (E, R) is a poset that satisfy condition $(*)$: if $C \subseteq E$ is a total ordering by R (i.e. is a chain in (E, R)), then C has an upper R-bound in E. Then (E, R) has a maximal element.*

Proof. We shall prove $(1) \Rightarrow (2)$. Suppose A is arbitrary set and let $c : \mathcal{P}(A) \backslash \{\varnothing\} \to \bigcup \mathcal{P}(A) \backslash \{\varnothing\}$ be a choice function. Define, by transfinite recursion, a definite operation F on Ord

$$F(\alpha) = \begin{cases} c(A \backslash \operatorname{Im}(F|_\alpha)), & \text{if } A \backslash \operatorname{Im}(F|_\alpha) \neq \varnothing \\ \theta, & \text{else} \end{cases}$$

where $\theta \in \text{Ord}$ is not in A fixed in advance.

If $F(\alpha) \neq \theta$ for any α, then $F : \text{Ord} \to A$. In particular, $F|_{h(A)} : h(A) \to A$ where $h(A)$ is the least ordinal not equinumerous with any subset of A. One should check $F|_{h(A)}$ is injective but that implies $h(A)$ must be equinumerous with some subset of A, a contradiction. Hence we must be able to find $F(\alpha) = \theta$ for some α. Let α be least ordinal such that $F(\alpha) = \theta$.

So we note $F|_\alpha : \alpha \to A$ is also injective. Since $F(\alpha) = \theta$, we must have $\mathrm{Im}(F|_\alpha) = A$, i.e. we actually get $F|_\alpha$ is a bijection. Therefore, we define $a < b$ in A by the order relation given by $F|_\alpha^{-1}(a)$ and $F|_\alpha^{-1}(b)$ in α.

Now let us assume the well-ordering principle and prove Zorn's lemma. Suppose $(E, \prec)$ is a poset satisfying the stated condition on chains. By well-ordering principle, there exists a well-ordering on E, say $<$. WLOG we may assume both orderings are strict. Since it admits well-ordering, we may assume E is isomorphic to an ordinal, i.e. $E \in \mathrm{Ord}$ and $<$ is in fact just $\in$. Now let $h(E)$ be the cardinal given by Proposition 2.1.8, that is, $h(E)$ is the least ordinal that not equinumerous with any subset of E. We will verify Zorn's lemma as follows: we will assume a contradiction that E has no $\prec$-maximal element and use that to construct an embedding $h(E)$ into E, which would indeed be a contradiction. Let $F : h(E) \to E$ be defined by transfinite recursion as follows: $F(0) = 0$, for all ordinal $\alpha < h(E)$ we get $F(\alpha + 1)$ to be the $<$-least $\beta \in E$ so $F(\alpha) \prec \beta$. For limit ordinal $\alpha < h(E)$ we define $F(\alpha)$ to be the $<$-least β such that $\forall \gamma < \alpha, F(\gamma) \prec \beta$, if such β exists. If such β does not exists, define $F(\alpha) = 0$. This function is well-defined since E has no $\prec$-maximal element by assumption and $<$ is a well-ordering on E. Since $h(E)$ is a limit ordinal, to show F is injective it suffice to prove the following claim:

Claim: For all $\alpha < h(E)$, $F|_\alpha$ is strictly order preserving, i.e. for all $x < y < \alpha$ we have $F(x) \prec F(y)$. We will prove this by transfinite induction on $\alpha < h(E)$. For $\alpha = 0$ there is nothing to prove, if α is limit ordinal then $\alpha = \bigcup_{\beta<\alpha} \beta$ and hence the claim follows by induction hypothesis. Finally, suppose $\alpha \in \mathrm{Ord}$ and consider $F|_{\alpha+1}$. There are two possibilities, either α itself is a successor or a limit. If $\alpha = \beta + 1$ then $F|_{\alpha+1} = F|_{\gamma:\gamma<\beta+1}$ and by definition we have $F(\beta) \prec F(\beta + 1)$. This, together with induction hypothesis, shows $F|_{\alpha+1}$ is strictly order preserving. Thus, suppose α is a limit ordinal, by induction hypothesis we see $F|_\alpha$ is strictly order preserving, and hence $\mathrm{Im}(F|_\alpha)$ forms a totally $\prec$-ordered set. It follows that there exists an upper bound to this set, and hence by definition $F(\alpha)$ is such a bound. So $F(\gamma) \prec F(\alpha)$ for all $\gamma < \alpha$, hence $F|_{\alpha+1}$ is strictly order preserving as desired. This proves the claim.

Finally let us assume Zorn's lemma and drive the axiom of choice. Suppose $\mathcal{F}$ is a set of non-empty sets and let us consider the set Λ of all partial choice functions on $\mathcal{F}$, identified with their graphs. That is, the elements of Λ are sets of the form

$$\{(G, x) : G \in \mathcal{G}, x \in G\}$$

where $\mathcal{G} \subseteq \mathcal{F}$. Note that Λ is non-empty, it contains, for example, $\{(F, x)\}$ for each $F \in \mathcal{F}$ and $x \in F$. Now Λ forms a poset under $\subseteq$. Moreover, if Θ is a totally ordered subset of Λ, then $\bigcup \Theta$ is quite easily seen to be a partial choice function on $\mathcal{F}$, and hence an upper bound for Θ in Λ. So by assumption of Zorn's lemma we get a maximal element, say f_∞. I claim f_∞ is a choice function on $\mathcal{F}$. Suppose not, then there exists some $F \in \mathcal{F}$ such that F is not in the domain of f_∞. However, then by fix $x \in F$, we see $f_\infty \cup \{(F, x)\}$ is a strictly larger partial choice function on $\mathcal{F}$, contradicting maximality of f_∞. Thus f_∞ must be a choice function as desired. ♡

2.3 Cardinality

Proposition 2.3.1. *Assume AC, then:*

1. *Every set is equinumerous to a cardinal.*
2. *The cardinal in part* (1) *is unique.*

Proof. (1): Let X be a set, then by well-ordering principle, there is a well-ordering on X, say $(X,<)$. Now we see X must be isomorphic to some ordinal since X admits well-ordering, i.e. we may think X as an ordinal. Thus we could consider $S := \{\beta \le X : X \text{ is equinumerous to } \beta\}$. Then we can find $\alpha_0 \in S$ to be least and we see α_0 is a cardinal. Indeed, if $\beta < \alpha_0$, then $\beta \notin S$, so β is not equinumerous with X, but α_0 is, so β is not equinumerous with α_0, i.e. α_0 is a cardinal as desired.

(2): Immediate by the construction of α_0. ♡

Definition 2.3.2. If X is a set, then by $|X|$, the ***cardinality*** of X, we mean the unique cardinal that X is equinumerous to.

Remark 2.3.3. We note $|X| = |Y|$ if and only if X and Y are equinumerous. We also remark that once we start to talk about cardinality, we will normally assume AC.

Proposition 2.3.4. *Let X, Y be sets, then $|X| \le |Y|$ if and only if there is an injective map from X to Y.*

Proof. Let $|X| = \kappa$ and $|Y| = \lambda$ where $\kappa, \lambda \in \mathrm{Card}$ is the class of all cardinals. If $\kappa \le \lambda$ then $\kappa \subseteq \lambda$, hence we see

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

give us a injection from X to Y where f, g are bijection. Conversely, suppose there is an injective map $h : X \to Y$, then consider the diagram

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

give us an injection from κ to λ. If $\lambda < \kappa$ then $\lambda \subseteq \kappa$, thus by Schroeder-Bernstein, we must have κ and λ are equinumerous, which contradicts the fact $\lambda < \kappa$ and they are both cardinals. Therefore, we must have $\lambda \nless \kappa$ and so $\kappa \le \lambda$ as desired. ♡

Proposition 2.3.5. *If $f : A \to B$ is a function, then $|\mathrm{Im}(f)| \le |A| = |\mathrm{Dom}(f)|$.*

Proof. Consider the map $b \mapsto f^{-1}(b) = \{a \in A : f(a) = b\}$ where $b \in B$. This is a definite operation g

$$\mathrm{Im}(f) \to \mathcal{F} = \{f^{-1}(b) : b \in \mathrm{Im}(f)\}$$

under the map $g(b) = f^{-1}(b)$. Thus, we can find a choice function on $\mathcal{F}$, i.e. we get $c : \mathcal{F} \to \bigcup \mathcal{F}$ where we see $\bigcup \mathcal{F} \subseteq A$. Hence, we get a function $\mathrm{Im}(f)$ to $\mathcal{F}$, then from $\mathcal{F}$ to A. Thus consider the map $s(b) = c(g(b)) = c(f^{-1}(b))$. This is injective as one should check. Thus we get $|\mathrm{Im}(f)| \le |A|$ as desired. ♡

Proposition 2.3.6. *Let A, B be sets, then:*

1. *We either have an injection from A to B or an injection from B to A.*
2. *A countable union of countable set is countable.*

Proof. Omitted. ♡

Remark 2.3.7. For the rest of the course we will assume AC, and this with ZF is called ZFC axioms.

We also remark the operation h on cardinal. Suppose $\kappa \in \mathrm{Card}$, then $h(\kappa)$ is the least cardinal not equinumerous with any subsets of κ. In fact, this $h(\kappa)$ is the least cardinal greater than κ, i.e. this is like the successor operation on cardinals.

Indeed, if $\lambda > \kappa$ are two cardinals, then $|\lambda| = \lambda > \kappa = |\kappa|$. Hence there is no injection $\lambda \to \kappa$, hence λ is not equinumerous with any subset of κ and so $\lambda \geq h(\kappa)$ by the definition of $h(\kappa)$.

Definition 2.3.8. We denote $h(\kappa)$ by κ^+ if $\kappa \in \mathrm{Card}$.

Remark 2.3.9. We note $\kappa + 1$ is completely different from κ^+. In the following, we will give an ordinal-valued enumeration of all infinite cardinals, i.e. we will give a strictly increasing definite operation from Ord onto Card.

Definition 2.3.10. Define, recursively, $\aleph_0 = \omega$, $\aleph_{\alpha+1} = \aleph_\alpha^+$ and if α is limit ordinal then

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

Lemma 2.3.11. *For all $\alpha \in \mathrm{Ord}$ we have $\aleph_\alpha$ is an infinite cardinal.*

Proof. We only need to check $\aleph_\beta$ with β being limit ordinal. Let $\alpha \in \mathrm{Ord}$ with $\alpha < \aleph_\beta$, then there is $\gamma < \beta$ such that $\alpha < \aleph_\gamma$. Thus by induction hypothesis we see $\aleph_\gamma$ is an infinite cardinal and $\alpha < \aleph_\gamma \leq \aleph_\beta$. Hence we get

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

so α and $\aleph_\beta$ are not equinumerous. Hence $\aleph_\beta$ is a cardinal as desired. ♡

Lemma 2.3.12. *If $\alpha < \beta$ are two ordinals, then*

$$\aleph_\alpha < \aleph_\beta$$

Proof. By induction, exercise. ♡

Lemma 2.3.13. *For all $\alpha \in \mathrm{Ord}$ we have $\alpha \leq \aleph_\alpha$.*

Proof. By induction, exercise. We add a word on the $\leq$ in the lemma, as this inequality is strict for successor ordinals, it may fail for limit ordinals. ♡

Example 2.3.14. Consider

$$0, \aleph_0, \aleph_{\aleph_0}, \aleph_{\aleph_{\aleph_0}}, \dots$$

Viz, $\alpha_0 = 0$ and $\alpha_{n+1} = \aleph_{\alpha_n}$, one should try to check $\alpha = \sup\{\alpha_n : n < \omega\}$ then we have

$$\aleph_\alpha = \alpha$$

Proposition 2.3.15. *Every infinite cardinal is of the form $\aleph_\alpha$ for some $\alpha \in \mathrm{Ord}$.*

Proof. Suppose $\kappa \in \mathrm{Card}$ be infinite, then $\kappa \le \aleph_\kappa < \aleph_{\kappa+1}$. Thus, it suffice to prove for all ordinals β and every infinite cardinal $\kappa < \aleph_\beta$, there is $\alpha \in \mathrm{Ord}$ such that $\kappa = \aleph_\alpha$. We prove this by induction on β.

If $\beta = 0$, then the claim is trivial. Suppose our claim holds for β. Let $\kappa < \aleph_{\beta+1} = \aleph_\beta^+$. Hence we have $\kappa \le \aleph_\beta$ and we either have $\kappa = \aleph_\beta$ (we are done!) or $\kappa < \aleph_\beta$ (by induction hypothesis we are done)。

Now suppose $\beta > 0$ is a limit ordinal. Let $\kappa < \aleph_\beta$. Then we have $\kappa < \aleph_\gamma$ for some $\gamma < \beta$ and hence by induction hypothesis we are done. ♡

Remark 2.3.16 (Continuum Hypothesis)**.** The starting point is Cantor's diagonalisation, i.e. for any set E, we have $|E| < |\mathcal{P}(E)|$. Thus, the question is, how big is $\mathcal{P}(E)$, the power set of E? In particular, the Continuum hypothesis (CH) is asking this question about $\aleph_0$. In particular, Continuum hypothesis claims

$$|\mathcal{P}(\aleph_0)| = \aleph_1$$

It is well-known that this Continuum hypothesis is independent of ZFC and this is done by Cohen's technique called forcing in the 60's. In this course, we will not assume CH nor ¬CH.

We also have the generalized continuum hypothesis, which states

$$|\mathcal{P}(\kappa)| = \kappa^+$$

2.4 Cardinal Arithmetic

Definition 2.4.1. Let $\kappa_1, \kappa_2 \in \mathrm{Card}$, the ***cardinal sum*** is

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

where X_1, X_2 are disjoint sets with $|X_i| = \kappa_i$.

Remark 2.4.2.

1. One should try to prove this is well-defined.

2. For any sets X_1, X_2, we have $|X_1 \cup X_2| \le |X_1| + |X_2|$. Indeed, let $X_i' = X_i \times \{i\}$, then we see $|X_1' \cup X_2'| = |X_1'| + |X_2'| = |X_1| + |X_2|$. On the other hand, we have surjection $X_1' \cup X_2' \to X_1 \cup X_2$ given by $(a, b) \mapsto a$. Hence $|X_1 \cup X_2| \le |X_1' \cup X_2'| = |X_1| + |X_2|$.
3. We note except the finite case, ordinal sum is different from cardinal sum.

Definition 2.4.3. The ***cardinal product*** of κ_1 and κ_2 is

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

where X_1, X_2 are sets with $|X_i| = \kappa_i$.

Remark 2.4.4.

1. Prove the cardinal product is well-defined.
2. Note cardinal product is not the same as ordinal product. There is a notational ambiguity here.

Theorem 2.4.5. *Let $\kappa_1, \kappa_2 \in \mathrm{Card}$ not both finite, then:*

1. $\kappa_1 + \kappa_2 = \max\{\kappa_1, \kappa_2\}$.
2. *If neither is zero, then* $\kappa_1\kappa_2 = \max\{\kappa_1, \kappa_2\}$.

Definition 2.4.6. Let I be a set, by ***I-sequence*** of sets $(X_i : i \in I)$ we mean a function

$$f : I \to \mathcal{U}$$

where $X_i = f(i)$ and $\mathcal{U}$ is the class of sets.

Definition 2.4.7. Suppose $(\kappa_i : i \in I)$ is an I-sequence of cardinals, the ***generalized cardinal sum*** is

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

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

Remark 2.4.8.

1. Prove this is well-defined.
2. If $(X_i : i \in I)$ is given by $f : I \to \mathcal{U}$ via $i \mapsto X_i$, then $\bigcup_{i \in I} X_i := \bigcup \mathrm{Im}(f)$.

Theorem 2.4.9. *Suppose I is infinite and $(\kappa_i : i \in I)$ is a sequence of non-zero cardinal, then:*

1. $\sup_{i \in I} \kappa_i$ *is a cardinal.*
2. $\sum_{i \in I} = \max\{|I|, \sup_{i \in I} \kappa_i\}$

Remark 2.4.10. If $(\kappa_i : i \in I)$ is given by $f : I \to \mathcal{U}$ with $f(i) = \kappa_i$, then $\sup_{i \in I} \kappa_i = \sup \mathrm{Im}(f)$.

Definition 2.4.11. Let $(\kappa_i : i \in I)$ be a sequence of cardinals. The ***generalized cardinal product*** is

$$\prod_{i \in I} \kappa_i := |F|$$

where F is the set of all functions $g : I \to \bigcup_{i \in I} X_i$ such that $g(i) \in X_i$ where $(X_i : i \in I)$ is a sequence of sets with $|X_i| = \kappa_i$. In other word, F is the set of all I-sequence $(x_i : i \in I)$ where $x_i \in X_i$, which is also denoted by $\times_{i \in I} X_i$.

Example 2.4.12. Suppose $k_i = 2$ for all i. Then

$$\prod_{i \in I} 2 = |\times_{i\in I} 2|$$

where $(a_i : i \in I) \in \times_{i \in I} 2$ and we have $a_i = 0$ or $a_i = 1$. Hence $\times_{i\in I} 2$ is just $\mathcal{P}(I)$ via the map $(a_i : i \in I) \mapsto \{i \in I : a_i = 1\}$. One should check this map is bijection. If $|I| > 1$ then $\prod_{i\in I} 2 > 2$.

Definition 2.4.13. Let $\kappa, \lambda \in \mathrm{Card}$, the ***cardinal exponentiation***

$$\kappa^\lambda := |\operatorname{Mor}(\lambda, \kappa)|$$

where $\operatorname{Mor}(\lambda, \kappa)$ is the set of functions $f : \lambda \to \kappa$. In other word, we have $\operatorname{Mor}(\lambda, \kappa) = \prod_{i<\lambda} \kappa$.

Example 2.4.14. We have $2^\lambda = |\mathcal{P}(\lambda)|$ by above example computation. We should try to compare this to $\sum_{i<\lambda} \kappa$, which is equal $\max\{\lambda, \kappa\}$ if $\lambda > \aleph_0$ with $\kappa \neq 0$.

Lemma 2.4.15.

1. *$\lambda \leq \mu \Rightarrow \kappa^\lambda \leq \kappa^\mu$ and $\lambda^\kappa \leq \mu^\kappa$.*
2. *$\kappa^{\lambda+\mu} = \kappa^\lambda \cdot \kappa^\mu$.*
3. *$(\kappa^\lambda)^\mu = \kappa^{\lambda\mu}$.*

Theorem 2.4.16 (Konig's Theorem)**.** *Let $(\kappa_i : i \in I)$ and $(\lambda_i : i \in I)$ are two sequence of cardinals, with $\kappa_i < \lambda_i$. Then*

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

Chapter 3

Model Theory

况闻内金盘，尽在卫霍室。
中堂舞神仙，烟雾蒙玉质。
暖客貂鼠裘，悲管逐清瑟。
劝客驼蹄羹，霜橙压香桔。

杜甫

3.1 Structure

Definition 3.1.1. A ***structure*** $\mathcal{M}$ consists of:

1. A nonempty set, M, called the ***universe*** of $\mathcal{M}$.
2. A sequence $(c_i : i \in I_c)$ of elements from M,called the ***constants*** of $\mathcal{M}$.
3. A sequence $(f_i : i \in I_F)$ of functions on powers of M called the ***basic functions*** of $\mathcal{M}$. Viz, for each $i \in I_F$ we have $f_i : M^{n_i} \to M$ with n_i defined to be the ***arity*** of f_i.
4. A sequence $(R_i : i \in I_R)$ of subsets of powers of M called the ***basic relations*** of $\mathcal{M}$. Viz, for each i we have $R_i \subseteq M^{k_i}$ with k_i called the ***arity*** of R_i.

Definition 3.1.2. Let $\mathcal{M}$ be a structure, the ***signature*** of $\mathcal{M}$ is $((c_i : I_c), (f_i : i \in I_F), (R_i : i \in I_R))$.

Remark 3.1.3. By convention, $M^0 = 1$ for any sets. So if f is a basic function of arity 0 then it is determined by $f(0) \in M$. Hence 0-ary functions are just constants and we usually only deal with the basic functions of arity bigger or equal to 1.

Also, we remark I_c, I_F, I_R can be empty, but M is nonempty. In particular, we may write the structure $\mathcal{M}$ as $\mathcal{M}(M, I_c, I_F, I_R)$.

Example 3.1.4. Consider $\mathbb{R}$:

1. If we are interested in the ordering on $\mathbb{R}$, then we consider the structure $\mathcal{M}(\mathbb{R}, \varnothing, \varnothing, <) = \mathcal{M}(\mathbb{R}, <)$ where $<$ is a basic binary relation, namely $\{(a,b) \in \mathbb{R}^2 : a < b\}$ and $\mathbb{R}$ is the universe.
2. If we are interested in the additve group of reals, then we consider the structure $\mathcal{M}(\mathbb{R}, 0, \{+,-\}, \varnothing) = \mathcal{M}(\mathbb{R}, 0, +, -)$ where 0 is a constant, $+$ is a binary basic function and $-$ is an unary basic function which sends a to $-a$.
3. Similarly, if we are interested in the ring of reals, then structure is $\mathcal{M}(\mathbb{R}, 0, 1, +, -, \cdot)$.
4. Of course we have the ordered ring structure $\mathcal{M}(\mathbb{R}, 0, 1, +, -, *, <)$.

Definition 3.1.5. Suppose $\mathcal{M}, \mathcal{N}$ are structures. We say $\mathcal{N}$ is an ***expansion*** of $\mathcal{M}$, or $\mathcal{M}$ is a ***reduct***, if they have the same universe and the signature of $\mathcal{M}$ is contained in $\mathcal{N}$.

Example 3.1.6. We see $\mathcal{M}(\mathbb{R}, 0, +, -)$ is a reduct of $\mathcal{M}(\mathbb{R}, 0, 1, +, -, \cdot)$.

Remark 3.1.7. An important theme in model theory is to ask question like:

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

To discuss structures in same signature but different universes, it is useful to introduce "languages".

Definition 3.1.8. A ***language*** consists of the following sets of symbols:

1. A set L^c of ***constant symbols***.
2. A set L^F of ***function symbols*** together with a positive integer n_f for each $f \in L^F$ called its arity.
3. A set L^R of ***relation symbols*** together with a positive integer n_r for each $r \in L^R$ called its arity.

Definition 3.1.9. Suppose L is a language, an L-***structure*** is a structure $\mathcal{M}$ together with bijections between the three sets of symbols of L and three sequence of signatures of $\mathcal{M}$ that preserve arity.

Example 3.1.10. Let $L = \{0, +, -\}$ be the language of groups, i.e. the set of constant symbols only contain 0, the function symbols are $\{+,-\}$ and relation symbols are empty set. Then $\mathbb{R} = (\mathbb{R}, 0, +, -)$ and $\mathbb{Z}/4\mathbb{Z} = (\mathbb{Z}/4\mathbb{Z}, 0, +, -)$ are both L-structures. Also, every group can be viewed as a L-structure but not every L-structure is a group.

For example, consider $\mathcal{N} = (\mathbb{Z}, 0^{\mathcal{N}}, +^{\mathcal{N}}, -^{\mathcal{N}})$ where $0^{\mathcal{N}} = 114514$, $+^{\mathcal{N}}(a,b) = \max(a,b)$ and $-^{\mathcal{N}}(a) = 0$. Then this is not a group, but it is L-structure.

Definition 3.1.11. Suppose $\mathcal{M}, \mathcal{N}$ are L-structures. An L-***embedding*** $j : \mathcal{M} \to \mathcal{N}$ is an injective function $j : M \to N$ satisfying:

1. $j(c^{\mathcal{M}}) = c^{\mathcal{N}}$ for all constant symbol $c \in L^c$.
2. For all $f \in L^F$, say f is n-ary, then $j(f^{\mathcal{M}}(a_1, \ldots, a_n)) = f^{\mathcal{N}}(a_1, \ldots, a_n)$ for all $a_1, \ldots, a_n \in M$.
3. For every $R \in L^R$, say k-ary, for all $a_1, \ldots, a_k \in M$ we have $(a_1, \ldots, a_k) \in R^{\mathcal{M}}$ if and only if $(j(a_1), \ldots, j(a_k)) \in R^{\mathcal{N}}$.

Definition 3.1.12. An L-***isomorphism*** is a surjective L-embedding.

Definition 3.1.13. Let $\mathcal{M}, \mathcal{N}$ be two L-structures such that $M \subseteq N$ and $j : M \to N$

the inclusion map is an L-embedding, then we call $\mathcal{M}$ an L-***substructure*** of $\mathcal{N}$, or $\mathcal{N}$ is an L-extension. We use $\mathcal{M} \subseteq \mathcal{N}$ to denote this.

Remark 3.1.14. So, $\mathcal{M} \subseteq \mathcal{N}$ if and only if all interpretations of constant symbols of $\mathcal{M}$ is equal the interpretation of $\mathcal{N}$, for all $f \in L^F$ we have $f^{\mathcal{N}}|_M = f^{\mathcal{N}}$, and $R^{\mathcal{N}} \cap M = R^{\mathcal{M}}$.

Example 3.1.15 (Exercise)**.** Suppose $\mathcal{N}$ is an L-structure and $A \subseteq N$. Then A is the universe of an L-substructure if and only if:

1. $A \neq \varnothing$.
2. A contains all constants of $\mathcal{N}$.
3. A is preserved under all basic functions of $\mathcal{N}$.

In this case, there is a unique L-substructure $\mathcal{A} \subseteq \mathcal{N}$ whose universe is A.

Example 3.1.16. We have the following table

Structure	Substructure
$(\mathbb{R})$, i.e. $L = \varnothing$	Nonempty subsets of $\mathbb{R}$
$(\mathbb{R}, 0, +)$	Monoids of $(\mathbb{R}, 0, +)$
$(\mathbb{R}, 0, +, -)$	Subgroups of the group $(\mathbb{R}, +)$
$(\mathbb{R}, 0, 1, +, -, \cdot)$	Subrings of $(\mathbb{R}, +, \cdot)$
$(\mathbb{R}, <)$	Nonempty subsets of $\mathbb{R}$ with induced ordering

Example 3.1.17. Fix a field F, we try to define the language of F-vector spaces is $L = \{0, +, -, (\lambda_a)_{a \in F}\}$ where λ_a is an unary function symbol.

If V is an F-vector space, we can make it an L-structure $\mathcal{V}$ by:

1. Universe equal V.
2. $0^{\mathcal{V}}$ is the zero vector.
3. $+^{\mathcal{V}}$ is vector addition.
4. $-^{\mathcal{V}}$ is the negative of a vector.
5. $\lambda_a^{\mathcal{V}}(v) = av$ for all $v \in V, a \in F$.

Then $\mathcal{V}$ is an L-structure.

Example 3.1.18 (Exercise)**.** Suppose $j : \mathcal{M} \to \mathcal{N}$ is an L-embedding. Let $A = j(M) \subseteq N$, then there is an unique substructure $\mathcal{A} \subseteq \mathcal{N}$ with universe A such that $j : \mathcal{M} \to \mathcal{A}$ such that it is an isomorphism.

The point of this exercise is that we see to study L-embedding it suffices to study substructures.

3.2 Terms and Formulas

Remark 3.2.1. We want to define L-formulas. These are used:

1. To describe properties of L-structures, i.e. axioms. For example, $L = \{<\}$, to isolate about posets among all L-structures, we use formulas.
2. To define certain subsets of the structure. For example, if we have $(R, 0, 1, +, -, \times)$ is a ring, then we want to look at the set of inverse of R

Remark 3.2.2. We will make use of a fixed infinite set Var of symbols called ***variables***. We will assume Var is countable.

Definition 3.2.3. Fix a language L. The set of ***L-terms*** is the set of strings of symbols defined recursively as follows[1]:

1. Every variables is an L-term.
2. Every constant symbol of L is an L-term.
3. If $f \in L^F$ is n-ary, and $t_1, ..., t_n$ are L-terms t hen $f(t_1, ..., t_n)$ is an L-term.

Definition 3.2.4. We write $t = t(x_1, ..., x_n)$ to mean that variables appearing in t come from the list $\{x_1, ..., x_n\}$.

Remark 3.2.5. Note L-terms are actually finite strings of symbols from $\mathrm{Var} \cup L^c \cup L^F \cup \{(,)\} \cup \{,\}$.

Remark 3.2.6 (Abuse of Notation)**.** For readability, we write things more naturally. For example, for $L = \{0, 1, +, -, \times\}$ instead of correctly write

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

we often write

$$(x_1 - x_2)(1x_2)$$

Note we cannot write $1x_2$ as x_2 because in some interpretation 1 may not be the identity.

Definition 3.2.7. Suppose $\mathcal{M}$ is an L-structure and $t = t(x_1, ..., x_n)$ is an L-term. We define the ***interpretation of t in $\mathcal{M}$*** to be the function

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

defined recursively as follows:

1. If $t = x_i$ for some $i = 1, ..., n$, then $t^{\mathcal{M}}(a_1, ..., a_n) = a_i$.
2. If $t = c$ for some $c \in L^c$, then $t^{\mathcal{M}}(a_1, ..., a_n) = c^{\mathcal{M}}$, i.e. it is a constant function.
3. If t is $f(t_1, ..., t_l)$ where $f \in L^F$ is l-ary, $t_1, ..., t_l$ are L-terms, then $t_i = t_i(x_1, ..., x_n)$ and
$$t^{\mathcal{M}}(a_1, ..., a_n) = f^{\mathcal{M}}(t_1^{\mathcal{M}}(a_1, ..., a_n), ..., t_l^{\mathcal{M}}(a_1, ..., a_n))$$

Remark 3.2.8. Note $t^{\mathcal{M}}$ depends not only on t and $\mathcal{M}$ but also on the presentation $t = t(x_1, ..., x_n)$, i.e. if we change the order of variable or number of variables we may get different function.

Example 3.2.9. Suppose t is a variable x. Then:

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

---

[1] alternatively this should be the smallest set satisfying the three properties

Example 3.2.10 (Exercise)**.** Let $\mathcal{M}$ be an L-structure, $A \subseteq M$. Then A is the universe of a substructure if and only if:

1. $A \neq \varnothing$
2. $t^{\mathcal{M}}(A^n) \subseteq A$ for every L-term $t(x_1, ..., x_n)$.

Remark 3.2.11. Let $\mathcal{T} := \{t^{\mathcal{M}} : t \text{ an L-term}\}$, then $\mathcal{T}$ is the smallest collection of functions on Cartesian powers of M satisfying:

1. All coordinate projections are in $\mathcal{T}$.
2. All constant functions $M^n \to M$ are in $\mathcal{T}$.
3. Closed under composition.

Definition 3.2.12. An ***atomic L-formula*** is a string of symbols from $L \cup \{(,)\} \cup \{,\} \cup \{=\} \cup \text{Var}$ of the form:

1. $(t = s)$ where t, s are L-terms.
2. $R(t_1, ..., t_k)$ where $R \in L^R$ is k-ary, $t_1, ..., t_k$ are L-terms.

Remark 3.2.13 (Abuse of Notation)**.** For example, if $L = \{<, \times\}$ we write $y_1 < y_2^2$ instead of the correct atomic formula $<(y_1, \times(y_2, y_2))$ where we recall $<$ is binary relation.

Definition 3.2.14. The set of ***L-formulas*** is the smallest set of string of symbols from $L \cup \{(,), =\} \cup \{,\} \cup \text{Var} \cup \{\neg, \wedge, \vee, \forall, \exists\}$ satisfying:

1. Every atomic L-formulas is an L-formula.
2. If ϕ, ψ are L-formulas, then so are $\phi \wedge \psi$, $\phi \vee \psi$ and $\neg\phi$.
3. If ϕ is an L-formula, and $x \in \text{Var}$ then $\forall x \phi$ and $\exists x \phi$ are also formulas.

Definition 3.2.15. Suppose ϕ is an L-formula and $x \in \text{Var}$. An occurrence of x in ϕ is said to be ***bound*** if it appears in the ***scope*** of quantifiers $\forall$ or $\exists$. Otherwise we say the occurrence is ***free***.

Example 3.2.16. Let $L = \{\in\}$. Then

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

where we write $\phi \to \psi$ to mean $\neg\phi \vee \psi$. The intuition of the above formula is, $x \in y$, x is a subset of y, and elements of y are either x or elements of x. In other word, even this is meaningless right now, the intuition says y is the successor of x.

In particular, we see in the above formula, the first x and y are free. The z followed by it is bounded, and x, y in the second part is still free. Finally, in the third part, x, y are still free and z is bounded.

In other word, this formula is really about x and y and not about z, i.e. z is a dummy variable can be replaced.

Remark 3.2.17. We write $\phi = \phi(x_1, ..., x_n)$ to mean that all free occurrences of variables are from $\{x_1, ..., x_n\}$. Note this do not require all of $x_1, ..., x_n$ to appear.

Definition 3.2.18. An L-formula with no free variable is called an ***L-sentence***.

Remark 3.2.19. We will assume that no variable appears in an L-formula both bound and free.

Example 3.2.20. In $L = \{<, \times, 0\}$, then the formula $(x < 0) \wedge \exists x(x^2 = y)$. The first occurrence of x is free and the second x is bounded. This may be complicated to interpret so we can rewrite this as

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

In the above new formula, we get x and y are free and z is bounded. This two are different formulas, but when we assign meanings to both formulas, they will have the same meaning.

Example 3.2.21. In $L = \{\in\}$, the terms are just variables Var. The atomic formulas are $(x = y)$ or $(x \in y)$. Thus we see, the definite properties in the set theory are those expressible by L-formulas. In particular, all ZFC axioms are L-formulas and they are actually all L-sentences.

3.3 Truth and Elementary Substructures

Definition 3.3.1. Let L be a language, $\mathcal{M}$ a L-structure, let $\phi(x_1, ..., x_n)$ be L-formula. Then let $a = (a_1, ..., a_n) \in M^n$, we define $\mathcal{M} \vDash \phi(a)$, and say a ***satisfies*** or ***realizes*** in $\mathcal{M}$ or a is true in $\mathcal{M}$, if:

1. If ϕ is $(t_1 = t_2)$ where $t_1 = t_1(x_1, ..., x_n)$ and $t_2 = t_2(x_1, ..., x_n)$ are L-terms. Then $\mathcal{M} \vDash \phi(a_1, ..., a_n)$ if $t_1^{\mathcal{M}}(a_1, ..., a_n) = t_1^{\mathcal{M}}(a_1, ..., a_n)$ where now $t_1^{\mathcal{M}}, t_2^{\mathcal{M}}$ now have real meanings.
2. If ϕ is $R(t_1, ..., t_k)$ where $t_i = t_i(x_1, ..., x_n)$ and R is k-ary relation symbol. Then $\mathcal{M} \vDash \phi(a_1, ..., a_n)$ if $(t_1^{\mathcal{M}}(a_1, ..., a_n), ..., t_k^{\mathcal{M}}(a_1, ..., a_n)) \in R^{\mathcal{M}}$.
3. If ϕ is $\psi \wedge \theta$ for some L-formulas $\psi(x_1, ..., x_n)$ and $\theta(x_1, ..., x_n)$, then $\mathcal{F} \vDash \phi(a)$ if and only if $\mathcal{M} \vDash \psi(a)$ and $\mathcal{M} \vDash \theta(a)$.
4. Similarly we define what $\psi \vee \theta$ as what you would expect, i.e. if $\phi = \psi \vee \theta$ then $\mathcal{M} \vDash \phi(a)$ if and only if $\mathcal{M} \vDash \psi(a)$ or $\mathcal{M} \vDash \theta(a)$.
5. If ϕ is $\neg\psi$ then $\mathcal{M} \vDash \phi(a)$ if $\mathcal{M} \nvDash \psi(a)$, i.e. iff $\psi(a)$ is not realized in $\mathcal{M}$.
6. If ϕ is $\exists y\psi$ then $\psi = \psi(x_1, ..., x_n, y)$ and $\mathcal{M} \vDash \phi(a)$ if there exists some $b \in M$ such that $\mathcal{M} \vDash \psi(a, b)$.
7. If ϕ is $\forall y\psi$ then $\mathcal{M} \vDash \phi(a)$ if for every $b \in M$, we have $\mathcal{M} \vDash \psi(a, b)$.

Definition 3.3.2. Let L be language, $\mathcal{M}$ a L-structure, $\phi = \phi(x_1, ..., x_n)$ be L-formula. Define

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

We call this the set ***defined*** by $\phi(x_1, ..., x_n)$ in $\mathcal{M}$.

Remark 3.3.3. What happen if $n = 0$? Then ϕ is an L-sentence, and either $\mathcal{M} \vDash \phi$ or not, i.e. ϕ is either true or not.

Example 3.3.4. Consider $\mathcal{R} = (\mathbb{R}, 0, 1, +, -, \times)$ and $\phi(x) \coloneqq \exists z(z^2 = x)$. Then $\mathcal{R} \vDash \phi(2)$ but $\mathcal{R} \vDash \neg\phi(-2)$. In fact, $\phi^{\mathcal{R}} = \mathbb{R}_{\geq 0}$. In particular, if $\sigma \coloneqq \forall x\phi(x)$, which is an L-sentence, then $\mathcal{R} \vDash \neg\sigma$.

Consider $\mathcal{Q} = (\mathbb{Q}, 0, 1, +, -, \times)$ and keep ϕ the same. Then we get $\mathcal{Q} \vDash \neg\phi(2)$ and $\mathcal{Q} \vDash \neg\phi(-2)$ are both true. In particular, we get

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

If $\mathcal{C} = (\mathbb{C}, 0, 1, +, -, \times)$, then $\mathcal{C} \vDash \sigma$, i.e. $\phi^{\mathcal{C}} = \mathbb{C}$.

Example 3.3.5. Now let $\mathcal{Z} = (\mathbb{Z}, 0, +, -) \subseteq (\mathbb{Q}, 0, +, -) = \mathcal{Q}$ with the language $L = \{0, +, -\}$ of groups. Consider the atomic formula $\psi(x, y) : y + y = x$. For $a, b \in \mathbb{Z}$, we have $\mathcal{Z} \vDash \psi(a, b)$ if and only if $a = 2b$ if and only if $\mathcal{Q} \vDash \psi(a, b)$. In other word, $\mathcal{Z}$ and $\mathcal{Q}$ agree on what the integer solutions are. Then consider $\phi(x) \coloneqq \exists y \psi(x, y)$, we see $\mathcal{Q} \vDash \phi(1)$. However, in $\mathcal{Z}$ this is false, i.e. $\mathcal{Z} \vDash \phi(1)$. In other word, now $\mathcal{Z}$ and $\mathcal{Q}$ do not agree on the integer realization of $\phi(x)$. Viz, there is a big difference between ϕ and ψ.

Proposition 3.3.6. *Suppose $\mathcal{M} \subseteq \mathcal{N}$ be two L-structures, and $\phi(x)$ is an L-formula, and $x = (x_1, \ldots, x_n)$. Let $a = (a_1, \ldots, a_n) \in M^n$, then if ϕ is quantifiers free, then $\mathcal{M} \vDash \phi(a)$ if and only if $\mathcal{N} \vDash \phi(a)$.*

Proof. Start by a claim on terms.

We claim for every L-term $t(x)$, we have $t^{\mathcal{N}}|_{M^n} = t^{\mathcal{M}}$. We use induction on the complexity on the term t. First suppose $t = x_i$ for $i = 1, \ldots, n$. Then $t^{\mathcal{N}}|_{M^n}$ is the ith coordinate projection on N^n restricted to M^n. Thus it is equal the ith coordinate projection on M^n. This is exactly $t^{\mathcal{M}}$.

Next suppose $t = c \in L^c$. Then we see since $c^{\mathcal{M}} = c^{\mathcal{N}}$ as $\mathcal{M}$ is a substructure of $\mathcal{N}$. Hence we get ${}^{\mathcal{N}}|_{M^n}$ is the constant function on N^n restricted to M^n, which is exactly the constant function $t^{\mathcal{M}}$ as desired.

Now let $t = f(t_1, \ldots, t_l)$ where $f \in L^F$ be l-ary and $t_1, \ldots, t_l$ are all L-terms that are in our above case. Let $a \in M^n$ then we get

$$t^{\mathcal{N}}(a) = f^{\mathcal{N}}(t_1^{\mathcal{N}}(a_1), \ldots, t_l^{\mathcal{N}}(a))$$

where by induction hypothesis we get $t_i^{\mathcal{N}}(a) \in \mathcal{M}$ and hence we get

$$f^{\mathcal{N}}(t_1^{\mathcal{N}}(a), \ldots, f_l^{\mathcal{N}}(a)) = f^{\mathcal{N}}(t_1^{\mathcal{M}}(a), \ldots, t_l^{\mathcal{M}}(a))$$

and since $\mathcal{M} \subseteq \mathcal{N}$ we get $f^{\mathcal{N}}|_{M^n} = f^{\mathcal{M}}$, we can conclude that

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

This concludes our claim about terms.

Now we do induction on complexity of $\phi(x)$. If $\phi(x)$ is atomic, then $\phi(x)$ is $(t = s)$ for some L-terms $t(x)$ and $s(x)$. In other word, $\mathcal{M} \vDash \phi(a)$ if and only if $t^{\mathcal{M}}(a) = s^{\mathcal{M}}(a)$ if and only if $t^{\mathcal{N}}(a) = s^{\mathcal{N}}(a)$ by the claim about terms above and hence if and only if $\mathcal{N} \vDash \phi(a)$. The other possibilities is $\phi(x) = R(t_1, \ldots, t_k)$ where R is k-ary relation symbol and $t_1, \ldots, t_k$ are L-terms. Then we get

$$\begin{aligned}
\mathcal{M} \vDash \phi(a) &\Leftrightarrow (t_1^{\mathcal{M}}(a), \ldots, t_k^{\mathcal{M}}(a)) \in R^{\mathcal{M}} = R^{\mathcal{N}} \cap M^k \\
&\Leftrightarrow (t_1^{\mathcal{M}}(a), \ldots, t_k^{\mathcal{M}}(a)) \in R^{\mathcal{N}} \\
&\Leftrightarrow (t_1^{\mathcal{N}}(a), \ldots, t_k^{\mathcal{N}}(a)) \in R^{\mathcal{N}} \\
&\Leftrightarrow \mathcal{N} \vDash \phi(a)
\end{aligned}$$

Next, if $\phi(x)$ is $\neg\psi(x)$. Then we see

$$\begin{aligned} \mathcal{M} \vDash \phi(a) &\Leftrightarrow \mathcal{M} \nvDash \psi(a) \\ &\Leftrightarrow \mathcal{N} \nvDash \psi(a) \\ &\Leftrightarrow \mathcal{N} \vDash \phi(a) \end{aligned}$$

If $\phi(x)$ is $\psi(x) \wedge \theta(x)$, we have

$$\begin{aligned} \mathcal{M} \vDash \phi(x) &\Leftrightarrow \mathcal{M} \vDash \psi(a) \text{ and } \mathcal{M} \vDash \theta(a) \\ &\Leftrightarrow \mathcal{N} \vDash \psi(a) \text{ and } \mathcal{N} \vDash \theta(a) \\ &\Leftrightarrow \mathcal{N} \vDash \phi(a) \end{aligned}$$

The proof for $\vee$ is similar (or to note $P \vee Q$ is the same as $\neg(\neg P \wedge \neg Q)$ and by above parts). ♡

Proposition 3.3.7. *Let $\mathcal{M} \subseteq \mathcal{N}$, $x = (x_1, ..., x_n)$, $\phi(x)$ is L-formula, $a = (a_1, ..., a_n) \in M^n$, then:*

1. *If $\phi(x)$ is existential then $\mathcal{M} \vDash \phi(a) \Rightarrow \mathcal{N} \vDash \phi(a)$.*
2. *If $\phi(x)$ is universal then $\mathcal{N} \vDash \phi(a) \Rightarrow \mathcal{M} \vDash \phi(a)$.*

Proof. Suppose $\phi(x)$ is $\exists y \psi(x, y)$ with $y = (y_1, ..., y_m)$ where ψ is quantifier free (q.f.). Then

$$\begin{aligned} \mathcal{M} \vDash \phi(a) &\Leftrightarrow \exists b \in M^m, \mathcal{M} \vDash \psi(a, b) \\ &\Leftrightarrow \exists b \in M^m \mathcal{N} \vDash \psi(a, b) \text{ since } \psi \text{ is q.f. and by above Prop} \\ &\Rightarrow \exists b \in N^m, \mathcal{N} \vDash \psi(a, b) \\ &\Rightarrow \mathcal{N} \vDash \phi(a) \end{aligned}$$

Then part 2 is just follows from $\forall y$ is same as $\neg\exists y$.

Or we can check step by step for part 2. Suppose $\phi(x)$ is $\forall y \psi(x, y)$ where $y = (y_1, ..., y_m)$ and ψ is q.f., then we have

$$\begin{aligned} \mathcal{N} \vDash \phi(a) &\Leftrightarrow \mathcal{N} \vDash \forall y \psi(a, y) \\ &\Leftrightarrow \mathcal{N} \vDash \neg\exists y \neg\psi(a, y) \\ &\Leftrightarrow \mathcal{N} \nvDash \exists y \neg\psi(a, y) \text{ where the last part is q.f.} \\ &\text{by existential case and contrapositive we have} \\ &\Rightarrow \mathcal{M} \nvDash \exists y \neg\psi(a, y) \\ &\Leftrightarrow \mathcal{M} \vDash \neg\exists y \neg\psi(a, y) \\ &\Leftrightarrow \mathcal{M} \vDash \phi(a) \end{aligned}$$

♡

Remark 3.3.8. What we just did for $\mathcal{M} \subseteq \mathcal{N}$ is also true for L-embeddings $j : \mathcal{M} \to \mathcal{N}$. For example, if ϕ is q.f. and $a \in M^n$, then

$$\mathcal{M} \vDash \phi(a) \Leftrightarrow \mathcal{N} \vDash \phi(j(a))$$

Definition 3.3.9. An L-embedding $j : \mathcal{M} \to \mathcal{N}$ is ***elementary*** if for all L-formulas $\phi(x)$, with $x = (x_1, ..., x_n)$ and all $a \in M^n$, we have

$$\mathcal{M} \models \phi(a) \Leftrightarrow \mathcal{N} \models \phi(j(a))$$

If $j : \mathcal{M} \to \mathcal{N}$ is elementary and $j : M \to N$ is the inclusion map, then we say that $\mathcal{M}$ is an elementary substructure of $\mathcal{N}$, denoted by $\mathcal{M} \preccurlyeq \mathcal{N}$.

Example 3.3.10. Consider $\mathcal{Q} = (\mathbb{Q}, 0, +, -)$. Then:

1. We have $\mathcal{Z} = (\mathbb{Z}, 0, +, -) \not\preccurlyeq \mathcal{Q}$. Indeed, let $\phi(x) : \exists y(y + y = x)$ then $\mathcal{Q} \models \phi(1)$ but $\mathcal{Z} \models \neg\phi(1)$.
2. In fact, $\mathcal{Q}$ has no proper elementary subgroup. Suppose for contradiction, $\mathcal{G} = (G, 0, +, -)$ is an elementary substructure of $\mathcal{Q}$. Then in $\mathcal{Q}$, we have $\mathcal{Q} \models \exists x(x \neq 0)$. Note in the definition of elementary structure you can take $n = 0$, i.e. elementary substructures satisfy all the same L-sentences all the extensions. Since $\exists x(x \neq 0)$ is L-sentence and $\mathcal{G} \preccurlyeq \mathcal{Q}$, we should have $\mathcal{G} \models \exists x(x \neq 0)$. So we have $a \in \mathcal{G}$ where $a \neq 0$. However, since $G \subseteq \mathbb{Q}$, we must have $a = \frac{n}{m}$ where $n, m \in \mathbb{Z}$ where WLOG we may assume $n, m > 0$. However, note since $\mathcal{G}$ is closed under addition, hence $n \in \mathcal{G}$ as we just add m times of a. Now, we note $\mathcal{Q}$ is n-divisible, i.e. $\mathcal{Q} \models \sigma_n$ with $\forall x \exists y(y + ... + y = x)$ where we add n times of y. This is a L-sentence about fixed n. Hence σ_n must also be true for $\mathcal{G}$ and so there exists $y \in \mathcal{G}$ such that $y + ... + y = n$ where we add n times of y, since $n \in \mathcal{G}$ as we just showed. However, this forces $y = 1$, i.e. $1 \in \mathcal{G}$ and hence $\mathcal{G}$ contains all integers. However, now use divisibility again, we can get all of $\mathcal{Q}$ using integers, so $\mathcal{G}$ must actually be $\mathcal{Q}$, a contradiction.

3.4 Tarski-Vaught

Proposition 3.4.1. *Isomorphism $j : \mathcal{M} \to \mathcal{N}$ is elementary.*

Proof. Claim: for all L-term $t(x)$, we have $j(t^{\mathcal{M}}(a)) = t^{\mathcal{N}}(j(a))$. We will do this by induction on complexity. If $t = t(x_1, ..., x_n) = x_i$, let $a = (a_1, ..., a_n)$ be arbitrary with $a = j(b)$ where $b = (b_1, ..., b_n) \in M^n$ (exists since j is bijection). Then we get $j(t^{\mathcal{M}}(b)) = j(b_i) = a_i = t^{\mathcal{N}}(j(b))$ as desired. Next, suppose $t = c$, then $t^{\mathcal{M}}(a) = c^{\mathcal{M}}$ and $t^{\mathcal{N}}(a) = c^{\mathcal{N}}$ and hence

$$j(t^{\mathcal{M}}(a)) = j(c^{\mathcal{M}}) = c^{\mathcal{N}} = t^{\mathcal{N}}(j(a))$$

as desired. Finally, suppose $t = f(t_1, ..., t_l)$ with $j(t_i^{\mathcal{M}}(a)) = t_i^{\mathcal{N}}(j(a))$ for each i. Then we get

$$\begin{aligned} j(t^{\mathcal{M}}(a)) &= j(f^{\mathcal{M}}(t_1^{\mathcal{M}}(a), ..., t_l^{\mathcal{M}}(a)) \\ &= f^{\mathcal{N}}(j(t_1^{\mathcal{M}}(a)), ..., j(t_l^{\mathcal{M}}(a))) \\ &= f^{\mathcal{N}}(t_1^{\mathcal{N}}(j(a)), ..., t_l^{\mathcal{N}}(j(a))) \\ &= t^{\mathcal{N}}(j(a)) \end{aligned}$$

This concludes our claim for L-terms.

Now we deal with atomic L-formulas. In this case, say $\phi = \phi(x) := (t = s)$ where t, s are L-terms, we get $\mathcal{M} \vDash \phi(a)$ if and only if $t^{\mathcal{M}}(a) = s^{\mathcal{M}}(a)$ if and only if $j(t^{\mathcal{M}}(a)) = j(s^{\mathcal{M}}(a))$ if and only if $t^{\mathcal{N}}(j(a)) = s^{\mathcal{N}}(j(a))$ if and only if $\mathcal{N} \vDash \phi(j(a))$. Similarly, if $\phi = \phi(x) := R(t_1, ..., t_l)$, then we have $\mathcal{N} \vDash \phi(b)$ if and only if $(t_1^{\mathcal{N}}(b), ..., t_l^{\mathcal{N}}(a)) \in R^{\mathcal{N}}$. However, let $j(a) = b$ we get $t_i^{\mathcal{N}}(b) = t_i^{\mathcal{N}}(j(a)) = j(t_i^{\mathcal{M}}(a))$ and hence $(t_1^{\mathcal{N}}(b), ..., t_l^{\mathcal{N}}(a)) \in R^{\mathcal{N}}$ if and only if $(j(t_1^{\mathcal{M}}(a)), ..., j(t_l^{\mathcal{M}}(a))) \in R^{\mathcal{N}}$ if and only if $(t_1^{\mathcal{M}}(a), ..., t_l^{\mathcal{M}}(a)) \in R^{\mathcal{M}}$ by definition of L-embedding. Thus we get $\mathcal{M} \vDash \phi(a)$ if and only if $\mathcal{N} \vDash \phi(j(a))$ as desired.

Next, say $\phi = \phi(x) := \neg\psi$ where ψ has less complexity, i.e. $\mathcal{M} \vDash \psi(a)$ if and only if $\mathcal{N} \vDash \psi(j(a))$. Then we see $\mathcal{M} \vDash \phi(a)$ if and only if $\mathcal{M} \nvDash \psi(a)$ if and only if $\mathcal{N} \nvDash \psi(j(a))$ if and only if $\mathcal{N} \vDash \phi(j(a))$ as desired.

Next, say $\phi = \phi(x) = \psi \wedge \theta$ with ψ, θ all with less complexity. Then we get $\mathcal{M} \vDash \phi(a)$ if and only if $\mathcal{M} \vDash \psi(a)$ and $\mathcal{M} \vDash \theta(a)$ if and only if $\mathcal{N} \vDash \psi(j(a))$ and $\mathcal{N} \vDash \theta(j(a))$ if and only if $\mathcal{N} \vDash \phi(j(a))$ as desired.

Hence by the above we get $\psi \vee \theta$ case is also done since $\psi \vee \theta$ is logically equivalent to $\neg(\neg\psi \wedge \neg\theta)$.

We will now show $\exists y \psi(x, y)$ holds. Then as $\forall$ is $\neg\exists\neg$ we get all the cases handled. Say $\phi(x) = \exists y \psi(x, y)$, then we get

$$\begin{aligned}
\mathcal{M} \vDash \phi(a) &\iff \mathcal{M} \vDash \psi(a, b) \text{ for some } b \in M \\
&\iff \mathcal{N} \vDash \psi(j(a), j(b)) \text{ for some } b \in M \quad \text{by the induction hypothesis} \\
&\iff \mathcal{N} \vDash \psi(j(a), c) \text{ for some } c \in N \quad \text{as } j \text{ is surjective} \\
&\iff \mathcal{N} \vDash \phi(j(a))
\end{aligned}$$

This concludes our proof. ♡

Theorem 3.4.2 (Tarski-Vaught Test). *Let $\mathcal{M} \subseteq \mathcal{N}$, the following are equivalent:*

1. $\mathcal{M} \preccurlyeq \mathcal{N}$
2. *For every formula $\phi(x_1, ..., x_n, y)$ and $a \in M^n$, if $\mathcal{N} \vDash \exists y \phi(a, y)$ then there is $b \in M$ such that $\mathcal{N} \vDash \phi(a, b)$.*

Proof. We first show $(1) \Rightarrow (2)$. Consider the formula $\psi(x_1, ..., x_n) : \exists y \phi(x_1, ..., x_n, y)$. If $\mathcal{N} \vDash \exists y \phi(a, y)$, then we have $\mathcal{N} \vDash \psi(a)$. Hence we get $\mathcal{M} \vDash \psi(a)$ and so $\mathcal{M} \vDash \exists \phi(a, y)$ as $\mathcal{M} \preccurlyeq \mathcal{N}$. Hence we can find $b \in M$ such that $\mathcal{M} \vDash \phi(a, b)$ and so $\mathcal{N} \vDash \phi(a, b)$. This shows (2) holds as desired.

Conversely, assume (2). To prove (1) we show that for every $\theta(x_1, ..., x_n)$ and $a \in M^n$, we have

$$\mathcal{M} \vDash \theta(a) \Leftrightarrow \mathcal{N} \vDash \theta(a)$$

We use induction on complexity of θ. If θ is atomic, it holds because $\mathcal{M} \subseteq \mathcal{N}$. The case for $\wedge, \neg, \vee$ are all handled by above lemmas (since they are quantifier free).

Thus we show $\theta := \exists y \phi(x_1, ..., x_n, y)$. Suppose $\mathcal{M} \models \theta(a)$, then $\mathcal{M} \models \phi(a, b)$ for some $b \in M$ and hence by induction hypothesis we have $\mathcal{N} \models \phi(a, b)$. Therefore, $\mathcal{N} \models \theta(a)$ as desired.

Now suppose $\mathcal{N} \models \theta(a)$ and so $\mathcal{N} \models \exists y \phi(a, y)$ and so there is $b \in M$ such that $\mathcal{N} \models \phi(a, b)$ since (2) holds. Now by induction hypothesis we have $\mathcal{M} \models \phi(a, b)$ and so $\mathcal{M} \models \theta(a)$. ♡

Corollary 3.4.2.1 (Downward-Lowenheim-Skolem). *Suppose $\mathcal{M}$ is an L-structure and $A \subseteq M$. Then there exists an elementary substructure of $\mathcal{N}$ that contains A and is of cardinality $|A| + |L| + \aleph_0 := \kappa$. In particular, if $A = \varnothing$ and L is countable, then every L-structure has an countable elementary substructure.*

Proof. We construct a chain $A = A_0 \subseteq A_1 \subseteq A_2 \subseteq ... \subseteq M$ such that for every formula $\phi(x_1, ..., x_n, y)$ and $a \in A_i^n$, if $\mathcal{M} \models \exists \phi(a, y)$ then there is $b \in A_{i+1}$ such that $\mathcal{M} \models \phi(a, b)$.

So, how many pairs of (ϕ, a) are there at the ith stage? Well, we have at most κ many. Hence, each $|A_i| \leq \kappa$. Now let

$$B = \bigcup_{i \geq 0} A_i$$

and we have $|B| \leq \kappa$ and it is the universe of an elementary substructure of $\mathcal{M}$ as one could try to show(a hint is the exercise after this corollary). ♡

Example 3.4.3 (Exercise). Suppose $A \subseteq M$, where M is the universe of L-structure $\mathcal{M}$. Then A is the universe of an elementary substructure if and only if, for every L-formula $\phi(x, y)$ with $x = (x_1, ..., x_n)$ and every $a \in A^n$, if $\mathcal{M} \models \exists y \phi(a, y)$ then there is $b \in A$ such that $\mathcal{M} \models \phi(a, b)$.

We note A is just a subset, however, the condition above encodes the fact A contains all constants (by taking the sentence $\exists y (y = c)$) and the fact A is closed under all functions (by taking $\phi := \forall x \exists y, f(x) = y$).

Example 3.4.4. Let $\mathcal{R} = (\mathbb{R} : 0, 1, +, -, \times, <)$. If $0 \leq n < m$ are integers, then (n, m) is defined in $\mathcal{R}$ by

$$\phi(x) : (\sum_{i=1}^{n} 1 < x) \wedge (x < \sum_{i=1}^{m} 1)$$

Viz, we have $\phi^{\mathcal{R}} = (n, m)$. Similarly we can construct interval (q, r) for $q, r \in \mathbb{Q}$.

However, we cannot define the open interval $(0, \pi)$ using our current language. Thus, we need to change the language to allow ***parameters*** from the universe.

Definition 3.4.5. Let $\mathcal{M}$ be L-structure and $B \subseteq M$. Then we let $L_B := L \cup \{\underline{b} : b \in B\}$ where $\underline{b}$ means new constant symbols (i.e. all distinct). Then define $\mathcal{M}_B$ to be the L_B-structure where the universe is $\mathcal{M}$ and the symbols from L are exactly the same, and for each $b \in B$ we define

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

Example 3.4.6. Back to the example, in order to get the interval $(0,\pi)$ we can just include a new constant symbol $\underline{\pi}$. We usually drop the subscript and treat $\mathcal{M}$ as an L_B structure.

Remark 3.4.7. With our new definition, Tarski-Vaught becomes the following: Say $\mathcal{M} \subseteq \mathcal{N}$, then $\mathcal{M} \preccurlyeq \mathcal{N}$ if and only if for every L_M-formula $\phi(y)$ in a single variable if $\mathcal{N} \models \exists y \phi(y)$ then there is $b \in M$, $\mathcal{N} \models \phi(b)$.

Definition 3.4.8. Let $\mathcal{M}$ be L-structure, $B \subseteq M$. A subset $X \subseteq M^n$ is ***definable over*** B (or B-definable) if there is a L_B-formula $\phi(x_1, ..., x_n)$ such that $X = \{a \in M^n : \mathcal{M} \models \phi(a)\}$.

We say X is ***definable in*** M if it is L_M-definable (equivalently, X is B-definable for some finite set $B \subseteq M$).

We say X is 0-***definable*** if it is definable over $\varnothing$.

Remark 3.4.9. Note if X is B-definable, then there exists an L-formula $\phi = \phi(x_1, ..., x_n, y_1, ..., y_m)$ and some $b_1, ..., b_m \in M$ such that

$$X = \{a \in M^n : \mathcal{M} \models \phi(a, b_1, ..., b_m)\}$$

The L_B-formula for X would be $\psi(x_1, ..., x_n) \coloneqq \phi(x_1, ..., x_n, \underline{b_1}, ..., \underline{b_m})$.

Proposition 3.4.10. *Let $\mathcal{M}$ be L-structure, let $B \subseteq M$, $X \subseteq M^n$. If X is B-definable then for all $j \in \mathrm{Aut}_B(\mathcal{M}) \coloneqq$\{the set of all L-automorphism of $\mathcal{M}$ that fix B pointwise\}=\{set of L_B-automorphism of $\mathcal{M}_B$\}, we have $j(X) = X$ when consider j as set map, i.e. j fixes X set-wise.*

Proof. Say $X = \{a \in M^n : \mathcal{M} \models \phi(a, b)\}$ where $\phi(x_1, ..., x_n, y_1, ..., y_m)$ is L-formula with $b = (b_1, ..., b_m) \in B^m$. For any $a \in M^n$, we have

$$\mathcal{M} \models \phi(a, b) \Leftrightarrow \mathcal{M} \models \phi(j(a), j(b))$$

as j is isomorphism and every isomorphism is elementary. However, we also get

$$\mathcal{M} \models \phi(j(a), j(b)) \Leftrightarrow \mathcal{M} \models \phi(j(a), b)$$

as $j(b) = b$. Thus we get

$$a \in X \Leftrightarrow j(a) \in X$$

and the proof follows. ♡

Example 3.4.11. We have $(0,1)$ is not 0-definable in $(\mathbb{R}, <)$ (it is definable by the $L_{\mathbb{R}}$-formula $(\underline{0} < x \wedge x < \underline{1})$).

Let $j : \mathbb{R} \to \mathbb{R}$ be the automorphism $j(x) = x + 1$, then j is an automorphism of $(\mathbb{R}, <)$. However, $j((0,1)) \neq (0,1)$ where we note $\mathrm{Aut}_{\varnothing}(\mathcal{M}) = \mathrm{Aut}(\mathcal{M})$, i.e. the last proposition tell us if $(0,1)$ is 0-definable then $(0,1)$ should be fixed by all automorphism. We just showed this is not the case and hence it is a contradiction.

Definition 3.4.12. Let $f : X \to Y$ be a function where $X \subseteq M^n$ and $Y \subseteq M^m$. Then we say f is ***definable in*** $\mathcal{M}$ if $\Gamma(f) \subseteq M^{n+m}$ is definable.

Example 3.4.13. We will show $+$ is not definable in $(\mathbb{R}, <)$. Suppose $\Gamma(+) \subseteq \mathbb{R}^3$ is definable over $b_1, ..., b_m \in \mathbb{R}$. WLOG we may assume $b_1 < ... < b_m$. Now fix $c \in \mathbb{R}$ such that $c > b_m$. Define $j : \mathbb{R} \to \mathbb{R}$ defined by

$$j(x) := \begin{cases} x, & \text{if } x \leq c \\ x - \frac{x-c}{2}, & \text{if } x > c \end{cases}$$

Then $j \in \mathrm{Aut}_{\{b_1,...,b_m\}}(\mathbb{R}, <)$. However, j does not fix $\Gamma(+)$, i.e. we have

$$j(c+1, c+1, 2c+2) = (c + \frac{1}{2}, c + \frac{1}{2}, \frac{3}{2}c + 1) \notin \Gamma(+)$$

Hence a contradiction.

3.5 Examples: Algebraic/Semi-Algebraic Sets

Example 3.5.1. Consider the language of rings $L = (0, 1, +, -, \times)$. Then $\mathcal{R} = (R, 0, 1, +, -, \times)$ where R is a commutative and unitary, what are the definable sets in $\mathcal{R}$?

Suppose $P_1...P_l \in R[X_1, ..., X_n]$ be the ring of polynomial in n variables. Then

$$V(P_1, ..., P_l) := \{a \in R^n : \forall i, P_i(a) = 0\}$$

is called an ***algebraic*** or ***Zariski closed*** subset of R^n. Clearly those are all (quantifier free) definable in R, consider L_R-formula $\phi(x_1, ..., x_n)$ given by

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

In fact, these algebraic sets and their finite boolean combinations (i.e. finite intersections, unions, and complements) are the only quantifier free definable sets in R.

Indeed, atomic L_R-formulas are of the form $(t = s)$ where $t(x_1, ..., x_n), s(x_1, ..., x_n)$ are L_R-terms. As in a homework exercise, L_R-terms agree with polynomials over R. So the atomically definable sets in $\mathcal{R}$ are precisely the hypersurfaces: $V(P)$ where $P \in R[X_1, ..., X_n]$. Hence the quantifier free definable subsets of R^n are exactly of the form

$$V_1 \backslash W_1 \cup ... \cup V_k \backslash W_k$$

where $W_i \subseteq V_i \subseteq R^n$ are algebraic sets (using disjuncted normal form: every q.f. formula is logically equivalent to one of the form

$$\bigvee_{i=1}^{r} (\bigwedge_{i=1}^{s} \phi_i \wedge \bigwedge_{i=1}^{t} \neg\psi_i)$$

where ϕ_i, ψ_i are atomic).

Such sets are called the ***Zariski-constructible sets***. We have proven that in any ring, q.f. definable sets are just Zariski constructible sets.

Remark 3.5.2. In the above example, we dealt with q.f. formulas. Now what about formulas with quantifiers? Well, we need the following fact.

Fact: If R is algebraically closed field, then every definable set in

$$\mathcal{R} := (R, 0, 1, +, -, \times)$$

is Zariski-constructible sets (i.e. every definable set is q.f. definable, i.e. R has quantifier elimination). As a corollary of this fact, every definable subsets of R in $\mathcal{R}$, where R is algebraically closed, is either finite or cofinite.

Example 3.5.3. Consider $R = \mathbb{R}$ and $\mathcal{R} = (R, 0, 1, +, -, \times)$. Let $\phi(x) : \exists y(y^2 = x)$, then $\phi^{\mathcal{R}} = \mathbb{R}_{\geq 0}$. This is not finite nor cofinite and hence it is not q.f. definable in $\mathcal{R}$. Hence $\mathcal{R}$ does not have quantifier elimination. In fact, we have a theorem of Macintyre as follows (we will not prove this).

If $\mathcal{R}$ is a ring that admits quantifier elimination, then it is an algebraically closed field.

Remark 3.5.4. What can we say about $\mathbb{R}$? Note $<$ on $\mathbb{R}$ is definable in $(\mathbb{R}, 0, 1, +, -, \times)$ since $x < y$ if and only if $\exists z(z^2 = y - x) \wedge y \neq x$.

We have the following fact: every definable set in $(\mathbb{R}, 0, 1, +, -, \times, <)$ is quantifier free definable.

Example 3.5.5 (Exercise)**.** The q.f. definable sets in an ordered ring $(R, 0, 1, +, -, \times, <)$, i.e. $a < b \Rightarrow a + c < b + c$ and $a < b, c > 0 \Rightarrow ac < bc$, are finite boolean combination of sets of the form

$$P(x_1, ..., x_l) = 0$$

or

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

where P, P_i are polynomials. Those are called semi-algebraic sets.

Remark 3.5.6. With this new definition, we see the definable sets in $(\mathbb{R}, 0, 1, +, -, \times)$ are precisely the semi-algebraic sets.

3.6 Theory and Models

Definition 3.6.1. Let L be a language. An ***L-theory*** is a set of L-sentences.

Definition 3.6.2. Let T be an L-theory, a ***model*** of T is an L-structure $\mathcal{M}$ such that for every $\sigma \in T$, we have $\mathcal{M} \vDash \sigma$. We denote this by $\mathcal{M} \vDash T$.

Definition 3.6.3. A theory is ***consistent*** if it has a model.

Definition 3.6.4. A class $\mathcal{K}$ of L-structures, we say this class is ***elementary*** or ***axiomatisable*** if there exists an L-theory T such that $\mathcal{M} \in \mathcal{K}$ if and only if $\mathcal{M} \vDash T$.

Example 3.6.5. Let $L = \{e, \cdot, -1\}$ be the language of groups where e is the identity, $\cdot$ is multiplication of groups and -1 is the inverse function of the group and we use a^{-1} to denote $-1(a)$. Then the following are elementary:

1. The class of groups (this is finitely axiomatisable, i.e. the L-theory is finite here).
2. The class of abelian groups (finitely axiomatisable).
3. Groups of a fixed order n, i.e. we add the sentence $\forall x(x^n = e)$ to the L-theory of group axioms.
4. Torsion-free groups, i.e. we add the following sentence $\sigma_n : \forall x(x^n = e \Rightarrow x = e)$ to the L-theory of group axioms where $n \in \omega$. This is not finitely axiomatisable at present based on how we write down the L-theory. We can show the class of torsion free groups is actually not finitely axiomatisable no matter what L-theory we choose but we do not have the tools right now.
5. Divisible groups, i.e. we add $\tau_n : \forall x \exists y(y^n = x)$ for all $n \in \omega$. This is also infinite axiomatisable but we cannot prove right now.

In the future, we will show the following are not elementary:

1. The class of torsion groups, i.e. $\forall x \exists n(x^n = e)$ is not an L-sentence because we cannot quantify the natural numbers using language of groups. Thus we must do $\forall x((x = e) \vee (x^2 = e) \vee (x^3 = e) \vee \ldots)$, but this is not an L-sentence since we get infinite disjunction.
2. The class of finite groups. To include finite groups of all orders, we also need infinite disjunction, which is not allowed.

Definition 3.6.6. Let $\mathcal{M}$ be L-structure, the ***theory of*** $\mathcal{M}$ is the L-theory $\mathrm{Th}(\mathcal{M}) := \{\sigma : \mathcal{M} \vDash \sigma\}$.

Definition 3.6.7. If $\mathcal{M}, \mathcal{N}$ are L-structures, then $\mathcal{M}$ is said to be ***elementarily equivalent*** to $\mathcal{N}$ if $\mathrm{Th}(\mathcal{M}) = \mathrm{Th}(\mathcal{N})$, i.e. $\mathcal{M} \vDash \sigma \Leftrightarrow \mathcal{N} \vDash \sigma$ for all L-sentences σ. We use $\mathcal{M} \equiv \mathcal{N}$ to denote this.

Remark 3.6.8. If $j : \mathcal{M} \to \mathcal{N}$ is elementary embedding, then $\mathcal{M} \equiv \mathcal{N}$. Hence $\mathcal{M} \preccurlyeq \mathcal{N}$ then $\mathcal{M} \equiv \mathcal{N}$. However, the converse is false.

Example 3.6.9. Note $\mathcal{M} \subseteq \mathcal{N}$ and $\mathcal{M} \equiv \mathcal{N}$ does not imply $\mathcal{M} \preccurlyeq \mathcal{N}$. Take $\mathcal{M} = (\omega \backslash \{0\}, <)$ and $\mathcal{N} = (\omega, <)$. Then $\mathcal{M} \subseteq \mathcal{N}$ and on the other hand, we note $j : \mathcal{M} \to \mathcal{N}$ given by $n \to n - 1$ is isomorphism and every isomorphism is elementary embedding. Thus we have $\mathcal{M} \equiv \mathcal{N}$ as well. However, $\mathcal{M}$ is not elementary substructure of $\mathcal{N}$ since $\exists y(y < 1)$ is true in $\mathcal{N}$ but false in $\mathcal{M}$ where $\exists y(y < 1)$ is L_M-sentence.

Remark 3.6.10. On the other hand, suppose $\mathcal{M} \subseteq \mathcal{N}$ is a substructure, then we have $\mathcal{M} \preccurlyeq \mathcal{N}$ as L-structures if and only if $\mathcal{M}_M \equiv \mathcal{N}_M$ as L_M-structures.

Proposition 3.6.11. *Lte $\mathcal{M}, \mathcal{N}$ be L-structures. There exists elementary embedding $j : \mathcal{M} \to \mathcal{N}$ if and only if $\mathcal{N}$ can be expanded to a model of the L_M-theory $\mathrm{Th}(\mathcal{M}_M)$.*

Proof. Suppose $j : \mathcal{M} \to \mathcal{N}$ is an elementary embedding. Make $\mathcal{N}$ into an L_M-structure as follows: for every new constant symbol $\underline{a}$, i.e. $a \in M$, we define

$$\underline{a}^{\mathcal{N}} := j(a)$$

and interpret the other symbols as $\mathcal{N}$ did before. Call this new L_M-structure to be $\mathcal{N}'$. Then the universe of $\mathcal{N}'$ is equal $\mathcal{N}$ and one could use the property of elementary embedding to prove $\mathcal{N}'$ is a theory of $\mathrm{Th}(\mathcal{M}_M)$.

Conversely, suppose $\mathcal{N}'$ is an expansion of $\mathcal{N}$ such that $\mathcal{N}' \models \mathrm{Th}(\mathcal{M}_M)$. Then we define $j : \mathcal{M} \to \mathcal{N}$ as follows: for $a \in M$, then we define $j(a) = \underline{a}^{\mathcal{N}'}$. Now since $\mathcal{N}' \models \mathrm{Th}(\mathcal{M}_M)$, you should check we indeed get elementary embedding. ♡

Definition 3.6.12. Let T be an L-theory, σ be a L-sentence. We say that T ***implies*** (or ***entails***) σ, denoted by $T \models \sigma$, if for every model $\mathcal{M} \models T$ we have $\mathcal{M} \models \sigma$. We also says σ is a ***consequence*** of T.

Definition 3.6.13. We say a L-theory T is ***complete*** if for any L***-sentence*** σ, either $T \models \sigma$ or $T \models \neg\sigma$.

Example 3.6.14.

1. Let $\mathcal{M}$ be L-structure, let $T = \mathrm{Th}(\mathcal{M})$. Then for every σ, either $\sigma \in T$ or $\neg\sigma \in T$. This means $\mathrm{Th}(\mathcal{M})$ is always complete.
2. Consider the language of rings L and let T_1 be the theory of rings (i.e. contains axioms of rings). Then T_1 is incomplete. For example, for $\sigma : \forall x \forall y (xy = yx)$, then $T \not\models \sigma$ since we have non-commutative rings. On the other hand we also have $T \not\models \neg\sigma$ since we have commutative rings.
3. Let L be the language of rings and T_2 be the theory of fields. Then T_2 is incomplete as we consider $\tau : \forall x \exists y (y^2 = x)$. Then in $\mathbb{C}$ this is true and in $\mathbb{R}$ this is false. Thus we cannot model τ and cannot model $\neg\tau$.
4. Let L be the language of rings and T_3 be the theory of algebraically closed field. Consider the sentence $1 + 1 = 0$, then in $\mathbb{F}_2$ this is true but in $\mathbb{F}_3$ this is false. Hence T_3 is not complete.
5. Finally, consider ACF_p be the theory of algebraically closed fields of characteristic p where p is either prime or zero. This theory is complete.

Lemma 3.6.15. *Let T be consistent T-theory, let $\overline{T} := \{\sigma : T \models \sigma\}$, the following are equivalent:*

1. *T is complete.*
2. *$\overline{T}$ is maximally consistent, i.e. no bigger collection is consistent.*
3. *$\overline{T} = \mathrm{Th}(\mathcal{M})$ for some $\mathcal{M} \models T$.*
4. *If $\mathcal{M} \models T$ and $\mathcal{N} \models T$ then $\mathcal{M} \equiv \mathcal{N}$.*

Proof. (1) $\Rightarrow$ (2): note $\overline{T}$ is consistent since every model of T is a model of $\overline{T}$. Let S properly contain $\overline{T}$, we want to show S is inconsistent. Let $\sigma \in S \backslash \overline{T}$, then $T \not\models \sigma$ by definition of $\overline{T}$. Since T is complete, we have $T \models \neg\sigma$, so $\neg\sigma \in \overline{T} \subseteq S$. Hence $\sigma, \neg\sigma \in S$ and so S has no models.

(2) $\Rightarrow$ (3): Let $\mathcal{M} \models T$ be arbitrary. So $\mathcal{M} \models \overline{T}$ and hence $\overline{T} \subseteq \mathrm{Th}(\mathcal{M})$. However, by (2), $\overline{T}$ is maximally consistent while $\mathrm{Th}(\mathcal{M})$ is consistent (has $\mathcal{M}$ as a model). This means $\mathrm{Th}(\mathcal{M}) \subseteq \overline{T}$ as well.

(3) $\Rightarrow$ (4): suppose $\mathcal{M} \models T$ and $\overline{T} = \mathrm{Th}(\mathcal{M})$. Let $\mathcal{N}$ be another model of T, then $\overline{T} \subseteq \mathrm{Th}(\mathcal{N})$ where $\overline{T} = \mathrm{Th}(\mathcal{M})$. Since $\mathrm{Th}(\mathcal{M})$ is complete and we proved (1) $\Rightarrow$ (2), we get $\mathrm{Th}(\mathcal{M})$ is maximally consistent. Therefore, $\mathrm{Th}(\mathcal{N}) = \mathrm{Th}(\mathcal{M})$ as desired. This shows $\mathcal{M} \equiv \mathcal{N}$. This means, every model of T is elementarily equivalent to $\mathcal{M}$ and hence any two models of T must be elementarily equivalent to each other.

(4) $\Rightarrow$ (1). Say $T \nvDash \sigma$. Then there is some $\mathcal{M} \vDash T$ such that $\mathcal{M} \vDash \neg\sigma$. By (4), if $\mathcal{N} \vDash T$ is any other model then $\mathcal{N} \vDash \neg\sigma$. Thus $T \vDash \neg\sigma$ and hence T is complete as desired. ♡

Theorem 3.6.16 (Compactness Theorem). *Let T be an L-theory, then T is consistent if and only if every finite subset of T is consistent.*

Proof. We will prove this after, when we introduce ultraproducts. ♡

3.7 Ultraproducts

Definition 3.7.1. Let I be a set, a ***filter*** on I is a subset $\mathscr{F} \subseteq \mathcal{P}(I)$ satisfying:

1. $I \in \mathcal{F}$, $\varnothing \notin \mathscr{F}$.
2. $A, B \in \mathscr{F} \Rightarrow A \cap B \in \mathscr{F}$.
3. $A \in \mathscr{F}, A \subseteq B \subseteq I \Rightarrow B \in \mathscr{F}$.

Example 3.7.2.

1. If $I = \mathbb{R}$, then let $\mathscr{F} = \{A \subseteq \mathbb{R} : \mu(\mathbb{R} \backslash A) = 0\}$ where μ is the Lebesgue measure.
2. If I is any set, let κ be a fixed cardinal such that $\aleph_0 \le \kappa \le |I|$. Then define $\mathscr{F} = \{A \subseteq I : |I \backslash A| < \kappa\}$. When $\kappa = \omega$, then $\mathscr{F} = \{A \subseteq I : A \text{ is cofinite}\}$. This is called the ***Frechet filter*** on I and it always exists if I is infinite.
3. Let $x \in I$, consider $\mathscr{F} := \{A \subseteq I : x \in A\}$. This is a filter and we call them the principle filters.

Definition 3.7.3. Let L be a language, I be a set, and suppose we have a sequence of L-structures $(\mathcal{M}_i : i \in I)$. Suppose $\mathscr{F}$ is a filter on I, define a new L-structure $\Pi := \prod_{\mathscr{F}} \mathcal{M}_i$ as follows:

1. The universe of $\prod_{\mathscr{F}} \mathcal{M}_i$ is $\prod_{i \in I} M_i / E$ where E is the equivalence relation $(a_i : i \in I) \sim (b_i : i \in I)$ if and only if $\{i \in I : a_i = b_i\} \in \mathscr{F}$.
2. If $c \in L^c$, then $c^\Pi := [(c^{\mathcal{M}_i} : i \in I)]$ where $[a]$ means the E-class of the sequence a.
3. If $f \in L^F$ is n-ary function symbol. Let $a_1 := [(\alpha_{1i} : i \in I)], \ldots, a_n := [(\alpha_{ni} : i \in I)] \in \Pi$, then we define
$$f^\Pi(a_1, \ldots, a_n) := \big[(f^{\mathcal{M}_i}(\alpha_{1i}, \ldots, \alpha_{ni}) : i \in I)\big]$$
One should check this does not depend on the choice of representatives α_{ij} and only on the $a_1, \ldots, a_n$.
4. Let $R \in L^R$ be n-ary relation symbol. Let $a_1 := [(\alpha_{1i} : i \in I)], \ldots, a_n := [(\alpha_{ni} : i \in I)] \in \Pi$, then $(a_1, \ldots, a_n) \in R^\Pi$ if and only if
$$(i \in I : (a_{1i}, \ldots, a_{ni}) \in R^{\mathcal{M}_i}) \in \mathscr{F}$$
Check this is well-defined.

Then $\Pi = \prod_{\mathscr{F}} \mathcal{M}_i$ is an L-structure.

Definition 3.7.4. An ***ultrafilter*** is a maximal filter $\mathscr{U}$, i.e. if $\mathscr{F}$ is another filter such that $\mathscr{U} \subseteq \mathscr{F} \subseteq \mathcal{P}(I)$, then $\mathscr{F} = \mathscr{U}$.

Lemma 3.7.5. *Let $\mathscr{U}$ be a filter on I, then $\mathscr{U}$ is ultrafilter if and only if for every $A \subseteq I$, either $A \in \mathscr{U}$ or $I \backslash A \in \mathscr{U}$.*

Proof. ($\Rightarrow$): Suppose $A \subseteq I$. If $A \notin \mathscr{U}$, then consider $\mathscr{F} := \{B \subseteq I : \exists C \in \mathscr{U}, C \backslash A \subseteq B\}$. One can check this $\mathscr{F}$ is a filter using the fact $A \notin \mathscr{U}$. Clearly $\mathscr{F}$ contains $\mathscr{U}$ since we can take $B = C$ and so every $C \in \mathscr{U}$ is in $\mathscr{F}$. By maximality, then we must have $\mathscr{U} = \mathscr{F}$. However, note $I \backslash A \in \mathscr{F} = \mathscr{U}$, i.e. $I \backslash A \in \mathscr{U}$ as desired.

($\Leftarrow$): Suppose $\mathscr{F}$ is a filter properly contains $\mathscr{U}$, we will seek a contradiction. Let $A \in \mathscr{F} \backslash \mathscr{U}$, then we must have $I \backslash A \in \mathscr{U}$ by assumption. Then $A \cap (I \backslash A) \in \mathscr{F}$ and so $\varnothing \in \mathscr{F}$, which is a contradiction by the definition of a filter. ♡

Remark 3.7.6. Suppose $\mathscr{F}$ is a filter, we can always produce a maximal filter contains $\mathscr{F}$. This is just the Zorn's lemma. Indeed, consider Λ be the set of filters on I that contains $\mathscr{F}$. Then this is a partially ordered by $\subseteq$. If $\mathscr{G}_1 \subseteq \mathscr{G}_2 \subseteq \ldots$ is a chain in Λ, then $\bigcup_{i \in \omega} \mathscr{G}_i \in \Lambda$ and this is an upper bound for the chain. Hence we get a maximal element $\mathscr{U} \in \Lambda$ that contains $\mathscr{F}$ as desired.

Example 3.7.7.

1. Every principle filters are ultrafilters.
2. (Exercise): Suppose $\mathscr{U}$ is an ultrafilter on I, then $\mathscr{U}$ is non-principle if and only if $\mathscr{U}$ extends the Frechet filter.

Theorem 3.7.8 (Los' Theorem)**.** *Suppose $\mathcal{U}$ is an ultrafilter on I, let $\mathcal{M} = \prod_{\mathcal{U}} \mathcal{M}_i$ be the ultraproduct of $(\mathcal{M}_i : i \in I)$ with respect to $\mathcal{U}$. Suppose $\phi(x_1, \ldots, x_n)$ is an L-formula and*[1] *$g_1, \ldots, g_n \in \prod_{i \in I} M_i$. Then $\mathcal{M} \vDash \phi([g_1], \ldots, [g_n])$ if and only if*

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

In other word, $\mathcal{M} \vDash \phi([g_1], \ldots, [g_n])$ if and only if it is true in almost all $\mathcal{M}_i$. In particular, taking $n = 0$, then we see if $\phi = \sigma$ is a sentence, then $\mathcal{M} \vDash \sigma$ if and only if $\{i \in I : \mathcal{M}_i \vDash \sigma\} \in \mathcal{U}$.

Proof. We will proceed by induction on complexity of ϕ. For notational simplicity we will assume $n = 1$. We will start with a claim on terms.

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

$$t^{\mathcal{M}}([g_1], \ldots, [g_n]) = [(t^{\mathcal{M}_i}(g_i(1), \ldots, g_n(i)))_{i \in I}]$$

To see this claim, we use induction on complexity of the term. This is left as an exercise but it is almost just the definition.

Now we can proof the theorem. Start with atomic formula, say $\phi : t_1 = t_2$ where $t_1 = t_1(x)$ and $t_2 = t_2(x)$. Then, we see $\mathcal{M} \vDash \phi([g])$ if and only if $t_1^{\mathcal{M}}([g]) = t_2^{\mathcal{M}}([g])$

[1] In the following, we note we can think sequences as functions. In other word, those $g_1, \ldots, g_n$ we are about to define are functions $g_i : I \to \bigcup_{i \in I} M_i, g_i(k) \in M_k$

if and only if $[(t_1^{\mathcal{M}_i}(g))_{i\in I}] = [(t_2^{\mathcal{M}_i}(g_i))_{i\in I}]$ if and only if $\{i \in I : t_1^{\mathcal{I}_i}(g_i) = t_2^{\mathcal{M}_i}(g_i)\} \in \mathcal{U}$ if and only if $\{i \in I : \mathcal{M}_i \models \phi(g_i)\} \in \mathcal{U}$.

Now say $\phi := R(t_1, ..., t_l)$ where $t_j = t_j(x)$ are L-terms, R is l-ary relations. Then we see $\mathcal{M} \models \phi([g])$ iff $(t_1^{\mathcal{M}}([g]), ..., t_l^{\mathcal{M}}([g])) \in R^{\mathcal{M}}$ iff $([(t_1^{\mathcal{M}_i}(g_i))], ..., [(t_l^{\mathcal{M}_i}(g_i))]) \in R^{\mathcal{M}}$ iff $\{i \in I : (t_1^{\mathcal{M}_i}(g_i), ..., t_l^{\mathcal{M}_i}(g_i)) \in R^{\mathcal{M}_i}\} \in \mathcal{U}$ iff $\{i \in I : \mathcal{M}_i \models \phi(g_i)\} \in \mathcal{U}$.

Now say $\phi = \neg\psi$ (this is where we need ultrafilter, not just filter). In this case, we have $\mathcal{M} \models \phi([g])$ iff $\mathcal{M} \not\models \psi([g])$ where we can use induction hypothesis and so this happens if and only if $\{i \in I : \mathcal{M}_i \models \psi(g_i)\} \notin \mathcal{U}$. Since $\mathcal{U}$ is ultrafilter, we see this happens if and only if $I \backslash \{i \in I : \mathcal{M}_i \models \psi(g_i)\} \in \mathcal{U}$ if and only if $\{i \in I : \mathcal{M}_i \not\models \psi(g_i)\} \in \mathcal{U}$ if and only if $\{i \in I : \mathcal{M}_i \models \phi(g_i)\} \in \mathcal{U}$ if and only if $\mathcal{M} \models \phi$. The $\wedge, \vee$ case are left as exercise.

Now say $\phi(x) := \exists y \psi(x, y)$. Then we see $\mathcal{M} \models \phi([g])$ if and only if there is $h \in \prod_{i\in I} M_i$ such that $\mathcal{M} \models \psi([g], [h])$. Now by induction hypothesis, we see this happens if and only if there is $h \in \prod_{i\in I} M_i$, $\{i \in I : \mathcal{M}_i \models \psi(g_i, h_i)\} \in \mathcal{U}$. Call this set X_h, i.e. $X_h = \{i \in I : \mathcal{M}_i \models \psi(g_i, h_i)\}$. Now let $Y = \{i \in I : \mathcal{M}_i \models \phi(g_i)\}$, we want to prove $\mathcal{M} \models \phi([g])$ iff $Y \in \mathcal{U}$.

Thus, what we are left to show is there exists $h \in \prod_{i\in I} M_i$ such that $X_h \in \mathcal{U}$ if and only if $Y \in \mathcal{U}$. First, say there exists $h \in M$ and $X_h \in \mathcal{U}$. Then note $X_h \subseteq Y$ by definition and hence $Y \in \mathcal{U}$ by definition of filter. Now suppose $Y \in \mathcal{U}$, then for each $i \in Y$, we have $\mathcal{M}_i \models \exists y \psi(g_i, y)$, say $a_i \in M_i$ we have $\mathcal{M} \models \psi(g_i, a_i)$. Now define $h : I \to \bigcup_{i\in I} M_i$ by setting $h_i := \begin{cases} a_i, & \text{if } i \in Y \\ b_i, & \text{if } i \notin Y \end{cases}$ where $b_i \in \mathcal{M}_i$ is arbitrary. This define $h \in M = \prod_{i\in I} M_i$. We claim $Y \subseteq X_h$ since if $i \in Y$ then $h_i = a_i$ and so $\mathcal{M}_i \models \psi(g_i, h_i)$, so $i \in X_h$. Hence we must have $X_h \in \mathcal{U}$ as $\mathcal{U}$ is a filter. ♡

Definition 3.7.9. Let $\mathcal{M}$ be L-structure, I a set, $\mathcal{U}$ a ultrafilter on I. Then the ultraproduct $(\mathcal{M}_i : i \in I)$ with $\mathcal{M}_i = \mathcal{M}$, with respect to $\mathcal{U}$, is defined to be the ***ultrapower*** $\mathcal{M}^I/\mathcal{U}$ of $\mathcal{M}$, i.e. the ultrapower of $\mathcal{M}$ is

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

Definition 3.7.10. Let $\mathcal{M}$ be L-structure, I a set, $\mathcal{U}$ a ultrafilter on I. Define $d : \mathcal{M} \to \mathcal{M}^I/\mathcal{U}$ given by

$$d(a) = [(a, a, a, a, ...)]$$

This is the ***diagonal embedding***.

Proposition 3.7.11. *$d : \mathcal{M} \to \mathcal{M}^I/\mathcal{U}$ is an elementary embedding.*

Proof. Let $\phi(x_1, ..., x_n)$ be an L-formula, $a_1, ..., a_n \in M$. We want to show $\mathcal{M} \models \phi(a_1, ..., a_n)$ iff $\mathcal{M}^I/\mathcal{U} \models \phi(d(a_1), ..., d(a_n))$. However, by Los' theorem, it tell us that $\mathcal{M}^I/\mathcal{U} \models \phi(d(a_1), ..., d(a_n))$ if and only if $\{i \in I : \mathcal{M} \models \phi(a_1, ..., a_n)\} \in \mathcal{U}$ if and only if $\mathcal{M} \models \phi(a_1, ..., a_n)$ since elements of $\mathcal{U}$ are not empty and $I \in \mathcal{U}$. ♡

Remark 3.7.12. We can identifying $\mathcal{M}$ with the image under d, we have $\mathcal{M} \preccurlyeq \mathcal{M}^I/\mathcal{U}$.

Definition 3.7.13. Let $\{M_i : i \in I\}$ be a countable collection of sets, i.e. $|I| \le \omega$, then we say $\{M_i : i \in I\}$ has ***finite intersection property*** (FIP), if for any fintie $i_1, ..., i_k \in I$ we have $M_{i_1} \cap ... \cap M_{i_k} \neq \varnothing$.

Definition 3.7.14. We say a L-structure $\mathcal{M}$ is $\aleph_1$***-compact***, if given any countable set of 0-definable subset of M^n, say $\{F_i : i \in I\}$, if it has FIP, then $\bigcap_{i\in I} F_i \neq \varnothing$.

Proposition 3.7.15. *Suppose $(\mathcal{M}_i : i \in I)$ is a sequence of L-structures with $I = \omega$, let $\mathcal{U}$ be non-principle ultrafilter, then $\mathcal{M} := \prod_{\mathcal{U}} \mathcal{M}_i$ is $\aleph_1$-compact.*

Proof. Let F_i be defined by L-formula $\phi_i(x)$ in M^l. We claim we may assume $\vDash (\phi_{n+1} \Rightarrow \phi_n)$. We just take conjunction, i.e. replace $F_0, F_1, F_2, \ldots$ by $F_0, F_0 \cap F_1, F_0 \cap F_1 \cap F_2, \ldots$.

We may also assume ϕ_0 is $x = x$ by add it in the front. Thus we may assume $F_0 = M^l$. For each $i < \omega$, let $n_i := \max\{n \le i : \mathcal{M}_i \vDash \exists x \phi_n(x)\}$, note this exists because $n = 0$ satisfy this. We want to show, a sequence $(a_i)_{i\in I}$ with $a_i \in M_i$, if $a = [(a_i)] \in M$ then $\mathcal{M} \vDash \phi_n(a)$ for all n.

Let $a_i \in M_i$ be such that $\mathcal{M}_i \vDash \phi_{n_i}(a_i)$, we want to show $a = [(a_i)]$ works. Fix n, then define

$$X_n = \{i : i \ge n, M_i \vDash \exists x \phi_n(x)\}$$

We claim $X_n \in \mathcal{U}$. To see this, note $F_n \neq \varnothing$ and so $\mathcal{M} \vDash \exists x \phi_n(x)$ and so by Los' theorem we get

$$\{i \in I : \mathcal{M}_i \vDash \exists x \phi_n(x)\} \in \mathcal{U}$$

and since $\mathcal{U}$ is not principle, we get U extends the Frechet filter and hence $\{i : i \ge n\} \in \mathcal{U}$. Hence we get X_n is the intersection of two sets in $\mathcal{U}$ and so X_n is in $\mathcal{U}$ as desired.

We claim $X_n \subseteq \{i \in I : \mathcal{M}_i \vDash \phi_n(a_i)\}$. Let $i \in X_n$, so $\mathcal{M}_i \vDash \exists x \phi_n(x)$ and $i \ge n$. However, note n_i is the largest number with the property $\mathcal{M}_i \vDash \exists x \phi_n(x)$ so we must have $n \le n_i$. Hence we get $\phi_{n_i} \Rightarrow \phi_n$. However, note $\mathcal{M}_i \vDash \phi_{n_i}(a_i)$ by the choice of a_i and hence we get $\mathcal{M}_i \vDash \phi_n(a_i)$.

Hence, we get

$$\{i \in I : \mathcal{M}_i \vDash \phi_n(a_i)\} \in \mathcal{U}$$

for all n. Thus now by Los' theorem we get $\mathcal{M} \vDash \phi_n(a)$ for all n, as $a = [(a_i)]$. ♡

Remark 3.7.16. Now take $\mathcal{R} = (\mathbb{R}, 0, 1, +, -, \cdot, <)$. Fix a non-principle ultrafilter $\mathcal{U}$ on ω. Now consider $\mathcal{R} \preccurlyeq \mathcal{R}^I/\mathcal{U} = \mathcal{R}^*$. Then $\mathcal{R}^*$ is a non-standard model of $\mathcal{R}$. By the proposition proved before, we get $\mathcal{R}^*$ is $\aleph_1$-compact.

Thus, we get an element in $\mathcal{R}^*$ is bigger than all integers (by consider the sequence of 0-definable set given by $F_i := \{n \in \omega : n \ge i\}$ for $i \in \omega$). Those are call infinite elements.

We also have infinitesimal elements that are >0 but less than $\frac{1}{n}$ for all $n>0$.

Proposition 3.7.17. *The class of finite groups is not elementary/axiomatisable in* $L=\{0,+,-\}$.

Proof. For each $n<\omega$ let G_n be a group of size n. For example, $G_n=\mathbb{Z}/n\mathbb{Z}$. Thus we get a sequence of groups $\{G_n : i\in\omega\}$. Fix a non-principle ultrafilter $\mathcal{U}$ on ω and consider $G=\prod_{\mathcal{U}} G_n$. Let T be the theory of groups, let $\sigma\in T$, by Los' theorem, we see $G\vDash\sigma$ if and only if $\{i\in I : G_i\vDash\sigma\}\in\mathcal{U}$. However, note $\{i\in I : G_i\vDash\sigma\}=I$ as each G_i is a group, hence we get G is in fact a group.

Now for each n, let σ_n be the sentence says there exists at least n distinct elements. Now by Los' we get

$$G\vDash\sigma_N \Leftrightarrow \{n : G_n\vDash\sigma_N\}\in\mathcal{U}$$

where we note $\{n : G_n\vDash\sigma_N\}=\{n : n\geq N\}$, which is in Frechet filter which is contained in $\mathcal{U}$. Hence we must have $G\vDash\sigma_N$ for all $N\in\omega$. This shows G is infinite group.

Now suppose for a contradiction, say we have a theory of finite groups, say T_I. However, then we get $\{i\in I : G_i\vDash T_I\}=I\in\mathcal{U}$, so by Los' theorem we must have $G\vDash T_I$, i.e. G must be finite group, a contradiction. ♡

3.8 Compactness Theorem

Theorem 3.8.1 (Compactness Theorem)**.** *Let T be an L-theory, then T is consistent iff every finite subset of T is consistent.*

Proof. We only need to prove every finite subset of T is consistent imply T is consistent as the other direction is trivial.

Let $I:=\mathcal{P}^{fin}(T)=\{\Sigma\subseteq T : |\Sigma|<\omega\}$, and for each $\Sigma\in I$, let $\mathcal{M}_\Sigma\vDash\Sigma$ by our assumption and so we get a sequence $(\mathcal{M}_\Sigma : \Sigma\in I)$.

We will construct an ultrafilter on I. For each $\Sigma\in I$, let $X_\Sigma:=\{\Pi\in I : \Sigma\subseteq\Pi\}$ and we let $A=\{X_\Sigma : \Sigma\in I\}$. This is not a filter but we have:

1. $\varnothing\notin A$, $IX_\varnothing\in A$
2. $X_\Sigma\cap X_\Delta=X_{\Sigma\cap\Delta}\in A$

Thus, we can make A into a filter by define

$$\mathcal{F}:=\{Y\subseteq I : \exists\Sigma\in I, X_\Sigma\subseteq Y\}$$

However, note every filter is contained in some ultrafilter, thus we get an ultrafilter $\mathcal{U}$ containing $\mathcal{F}$. One could try to show $\mathcal{U}$ is not an principle filter, but this is not needed in the proof.

Now let
$$\mathcal{M} := \prod_{\mathcal{U}} \mathcal{M}_\Sigma$$
and we want to show $\mathcal{M} \models T$.

To that end, let $\sigma \in T$ be arbitrary, then we have $X_\sigma = \{\Sigma \in I : \sigma \in \Sigma\} \subseteq \{\Sigma \in I : \mathcal{M}_\Sigma \models \sigma\}$. Since $X_\sigma \in A \subseteq \mathcal{F} \subseteq \mathcal{U}$ we get
$$\{\Sigma \in I : \mathcal{M}_\Sigma \models \sigma\} \in \mathcal{U}$$
and hence we get $\mathcal{M} \models \sigma$ by Los' theorem. This concludes the proof. ♡

Corollary 3.8.1.1. *If $T \models \sigma$ then there is a finite subset $\Sigma \subseteq T$ such that $\Sigma \models \sigma$.*

Proof. Consider $S = T \cup \{\neg\sigma\}$, then we have S is inconsistent as $T \models \sigma$. Hence by compactness there is a finite $\Delta \subseteq S$ such that Δ is inconsistent. Now note $\Delta \cup \{\neg\sigma\}$ is also inconsistent. However, $\Delta \cup \{\neg\sigma\} = \Sigma \cup \{\neg\sigma\}$ for some finite $\Sigma \subseteq T$.

Since $\Sigma \cup \{\neg\sigma\}$ is inconsistent, we must have $\Sigma \models \sigma$. ♡

Theorem 3.8.2 (Upward Lowenheim Skolem)**.** *Suppose $\mathcal{M}$ is an infinite L-structure and κ is a cardinal greater than or equal* $\max\{|M|, |L|\}$. *Then there exists $\mathcal{N}$ such that $\mathcal{M} \preccurlyeq \mathcal{N}$ such that $|N| = \kappa$.*

Proof. Let $L' = L_M \cup \{c_\alpha : \alpha < \kappa\}$ where c_α are all constants. Let $T' = \mathrm{Th}(\mathcal{M}_M) \cup \{c_\alpha \neq c_\beta : \forall \alpha, \beta < \kappa(\alpha \neq \beta)\}$. To show T' is consistent, it suffice to show each finite $\Sigma \subseteq T'$ is consistent.

Then Σ involves only finitely many c_α and so, $\mathcal{M}_M$ can be expanded to a model of Σ. This uses the fact $\mathcal{M}$ is infinite so we can interpret those finitely many c_α to be distinct elements.

Hence there is a model $\mathcal{R}' \models T'$ and let $\mathcal{R}$ be the L-reduct of $\mathcal{R}'$. Since $\mathcal{R}$ can be expanded to a model of $\mathrm{Th}(\mathcal{M}_M)$ by 4.45 in the course note. There is an elementary embedding $\mathcal{M} \preccurlyeq \mathcal{R}$. Also $|R| \geq \kappa$ and so by the downward Lowenheim Skolem there is a elementary substructure $\mathcal{N} \preccurlyeq \mathcal{R}$ such that $|N| = \kappa$ and $M \subseteq N$. Clearly $\mathcal{M} \subseteq \mathcal{N}$ but we cannot conclude $\mathcal{M}$ is elementary substructure right now.

However, by the way we constructed $\mathcal{R}$ we see $\mathcal{M} \preccurlyeq \mathcal{R}$. Hence, by the fact $\mathcal{M} \preccurlyeq \mathcal{R}$, $\mathcal{M} \subseteq \mathcal{N}$ and $\mathcal{N} \preccurlyeq \mathcal{R}$ then we get $\mathcal{M} \preccurlyeq \mathcal{N}$ as desired. ♡

Corollary 3.8.2.1 (Vaught's Test)**.** *Let T be an L-theory with only infinite models. Suppose that for some infinite cardinal $\kappa \geq |L|$, all models of T of size κ are isomorphic. Then T is complete.*

Proof. Let $\mathcal{M} \models T$ and $\mathcal{N} \models T$. Then use either Upward or Downward Lowenheim Skolem, we have $\mathcal{M}'$ such that either $\mathcal{M} \preccurlyeq \mathcal{M}'$ or $\mathcal{M}' \preccurlyeq \mathcal{M}$ with $|M'| = \kappa$. Similarly we get $\mathcal{N}' \preccurlyeq \mathcal{N}$ or $\mathcal{N} \preccurlyeq \mathcal{N}'$ with $|N'| = \kappa$.

Hence, by assumption, we get $\mathcal{N}'$ and $\mathcal{M}'$ are isomorphic. However, note $\mathcal{N}$ is elementarily equivalent to $\mathcal{N}'$ and $\mathcal{M}$ is elementarily equivalent to $\mathcal{M}'$, therefore, we get $\mathcal{M} \equiv \mathcal{N}$ as desired. ♡

Definition 3.8.3. Let T be an L-theory with only infinite models, then we say T is ***κ-categorical*** if κ is an infinite cardinal with $\kappa \geq |L|$, such that all models of T of size κ are isomorphic.

Example 3.8.4. Let $L = \{<\}$. Let DLO be the L-theory of dense linear ordering without endpoints. We claim DLO is $\aleph_0$-categorical.

Suppose $(E_1, <)$ and $(E_2, <)$ are two dense linaer ordering without endpoints are both countable. We build the isomorphism as follows: consider a finite chain of elements $A_0 \subseteq A_1 \subseteq A_2 \subseteq A_3 \subseteq \ldots$ such that $E_1 = \bigcup_{i\geq 0} A_i$ and similarly $B_0 \subseteq B_1 \subseteq \ldots$ such that $E_2 = \bigcup_{i\geq 0} B_i$. Suppose for all $i < \omega$ there is an order preserving bijection $f_i : A_i \to B_i$ such that f_{i+1} extends f_i for all i. Then we can consider the function $f \coloneqq \bigcup_{i<\omega} f_i$, which will be from E_1 to E_2 and it will be a bijection that preserves order. This concludes we find an L-isomorphism as desired.

Hence, now we just need to construct such f_i and A_i, B_i.

In the odd step, we ensure that $\bigcup_{i<\omega} A_i = E_1$. In the even step, we ensure $\bigcup_{i<\omega} B_i = E_2$. In particular, let $E_1 = \{a_i : i < \omega\}$ and $E_2 = \{b_i : i < \omega\}$.

Then, at step 0, we have $A_0 = B_0 = \varnothing = f_i$. Suppose we have built order preserving bijection $f_n : A_n \to B_n$.

Say $n+1$ is odd, say $n+1 = 2m+1$. We want to a_m into A_{n+1}. Hence if $a_m \in A_n$ already, then we set $A_{n+1} = A_n$, $B_{n+1} = B_n$ and $f_{n+1} = f_n$. If a_m is not in A_n, consider the relationship of a_m to A_n, we have one of the three cases:

1. $a_m < A_n$
2. $a_m > A_n$
3. there is $\alpha < \beta$ in A_n such that $\alpha < a_m < \beta$.

Let $A_{n+1} = A_n \cup \{a_m\}$. To extend f_n:

1. In case one, let $b \in E_2$ be less than every element in B_n and set $B_{n+1} = B_n \cup \{b\}$ and extend f_n by setting $f_{n+1}(a_n) = b$. Note we can find such $b \in E_2$ since E_2 is a model of DLO, i.e. it has no endpoints.
2. In case two, let $b \in E_2$ be so that $b > B_n$. Let $B_{n+1} = B_n \cup \{b\}$ and $f_{n+1}(a_n) = b$ as well.
3. In case three, let $b \in E_2$ be so that $f_n(\alpha) < b < f_n(\beta)$. This is possible since f_n preserves order and E_2 is dense. Let $B_{n+1} = B_n \cup \{b\}$ and $f_{n+1}(a_n) = b$.

This concludes the odd case.

Now say $n + 1 = 2m$ is even, then do the same thing we did above but to B_m this time.

This concludes DLO is $\aleph_0$-categorical and hence by Vaught's test we have DLO is complete.

Example 3.8.5. By the above example, we see $(\mathbb{Q}, <)$ is elementarily equivalent to

$(\mathbb{R}, <)$ since both of them are DLO. In particular, note $\mathbb{R}$ is complete (topological space) and $\mathbb{Q}$ is not. However, since they are elementarily equivalent, this tell us, the completeness of ordering is not expressible as L-sentence.

Example 3.8.6. We will show ACF_p is complete where p is zero or a prime and $L = \{0, 1, +, -, \cdot\}$.

First, we note if K is a uncountable field such that $K \vDash \mathrm{ACF}_p$. Then $\operatorname{tr.deg}(K/\mathbb{F}) = |K|$ where $\mathbb{F}$ is the prime field, i.e. smallest field of characteristic p.

To see this, note in general, if B is a finite set then $|\overline{\mathcal{F}(B)}| = \aleph_0$. On the other hand, if B is infinite, then $|\overline{\mathcal{F}(B)}| = |B|$. Therefore, if K is uncountable, let $B \subseteq K$ be a tr. basis so that $K = \overline{\mathbb{F}(B)}$, then we get $|K| = |B| = \operatorname{tr.deg}(K/\mathbb{F})$ as desired.

Theorem 3.8.7. *Let κ be uncountable cardinal. Then ACF_p is κ-categorical.*

Proof. Say $K, L \vDash \mathrm{ACF}_p$ with $|K| = |L| = \kappa$. Let B be tr.basis of K over $\mathbb{F}$ and $C \subseteq L$ be tr.basis of L over $\mathbb{F}$. Then we have $|B| = |K| = \kappa = |C| = |L|$.

Therefore, let $\alpha : B \to C$ be a bijection, then this extends to a isomorphism of fields from $\mathbb{F}(B)$ to $\mathbb{F}(C)$. Hence by uniqueness of algebraic closure, we get $K = \overline{\mathbb{F}(B)} \cong \overline{\mathbb{F}(C)} = L$ and so the proof follows. ♡

Corollary 3.8.7.1. *ACF_p is complete.*

Proof. By Vaught's test and the fact all algebraic closed fields are infinite. ♡

Remark 3.8.8. Note ACF_p is not $\aleph_0$-categorical. Also, since ACF_p is complete, we see we cannot write down a L-sentence telling $\overline{\mathbb{Q}}$ is different from $\mathbb{C}$. In other word, there is no single sentence saying an element is transcendental.

Theorem 3.8.9 (Morley's Categoricity Theorem). *Let L be countable, T an L-theory. If T is κ-categorical for some uncountable κ, then T is λ-categorical for all uncountable λ.*

Theorem 3.8.10 (Lefschetz Principle). *Let $L = \{0, 1, +, -, \cdot\}$, σ an L-sentence, then the following are equivalent:*

1. *$K \vDash \sigma$ for some $K \vDash ACF_0$.*
2. *$K \vDash \sigma$ for any $K \vDash ACF_0$.*
3. *$\overline{\mathbb{F}_p} \vDash \sigma$ for all but finitely many primes p.*
4. *$\overline{\mathbb{F}_p} \vDash \sigma$ for infinitely many primes p.*

Proof. (1) $\Rightarrow$ (2) is due to ACF_0 is complete.

(2) $\Rightarrow$ (3): note (2) says $\mathrm{ACF}_0 \vDash \sigma$. Hence by compactness, we see $\Sigma \vDash \sigma$ for some finite set $\Sigma \subseteq \mathrm{ACF}_0$. However, note Σ is made up by sentences from ACF and sentences saying the field is characteristic zero, i.e. $\Sigma \subseteq \mathrm{ACF} \cup \{\tau_1, \ldots, \tau_N\}$ where τ_n says $\forall x (x \neq 0 \Rightarrow n \cdot x \neq 0)$. Thus let $p > N$, then we see $\mathrm{ACF}_p \vDash \sigma$ as $\mathrm{ACF}_p \vDash \tau_n$ for all $1 \leq n \leq N$.

(3) $\Rightarrow$ (4): Immediate.

$\neg$(1) $\Rightarrow$ $\neg$(4): suppose $K \vDash \mathrm{ACF}_0$ and $K \nvDash \sigma$. Thus $K \vDash \neg\sigma$. Since ACF_0 is complete, we get $\mathrm{ACF}_p \vDash \neg\sigma$. However, since (1) $\Rightarrow$ (3) we see $\mathrm{ACF}_p \vDash \neg\sigma$ for all but finitely many p. Thus $\overline{\mathbb{F}_p} \vDash \neg\sigma$ for all but finitely many p. Hence $\overline{\mathbb{F}_p} \vDash \sigma$ for only finitely many p. Thus we get (4) $\Rightarrow$ (1) as desired. ♡

3.9 Quantifiers Elimination

Definition 3.9.1. Let T be an L-theory, $\phi(x_1, ..., x_n)$ is T***-equivalent*** to $\psi(x_1, ..., x_n)$ if $T \vDash \forall x(\phi(x) \Leftrightarrow \psi(x))$, i.e. for all $\mathcal{M} \vDash T$, we have $\phi^{\mathcal{M}} = \psi^{\mathcal{M}}$.

Definition 3.9.2. A theory T admits ***quantifier elimination***(QE) if every formula is T-equivalent to a quantifier free formula, i.e. in every model of T, all definable sets are quantifier-free definable.

Theorem 3.9.3. *Let T be L-theory and $\phi(x_1, ..., x_n)$ be L-formula. Suppose $n > 0$ or L has a constant symbol. The following are equivalent:*

1. *ϕ is T-equivalent to some q.f. formula.*
2. *Given $\mathcal{M}, \mathcal{N} \vDash T$, and a common substructure $\mathcal{A} \subseteq \mathcal{M}$ and $\mathcal{A} \subseteq \mathcal{N}$, then for any $a_1, ..., a_n \in \mathcal{A}$ we have $\mathcal{M} \vDash \phi(a_1, ..., a_n)$ if and only if $\mathcal{N} \vDash \phi(a_1, ..., a_n)$.*

Proof. (1) $\Rightarrow$ (2): Let $\psi(x_1, ..., x_n)$ be q.f. formula so ϕ is T-equivalent to ψ. Then we see $\mathcal{M} \vDash \phi(a_1, ..., a_n)$ if and only if $\mathcal{M} \vDash \psi(a_1, ..., a_n)$. However, since ψ is q.f. we have $\mathcal{M} \vDash \psi(a_1, ..., a_n)$ happens if and only if $\mathcal{A} \vDash \psi(a_1, ..., a_n)$ as $\mathcal{A} \subseteq \mathcal{M}$. With the same argument, since $\mathcal{A} \vDash \psi(a_1, ..., a_n)$ we get $\mathcal{N} \vDash \psi(a_1, ..., a_n)$ and hence we get $\mathcal{N} \vDash \phi(a_1, ..., a_n)$.

Let
$$\Psi = \{\psi(x) : \psi \text{ is q.f. and } T \vDash \forall x(\phi(x) \Rightarrow \psi(x))\}$$
with $x = (x_1, ..., x_n)$. Now let $L' = L \cup \{c_1, ..., c_n\}$ where those c_i's are new constant symbols.

Then we claim: if (2) holds, then $T \cup \Psi(c_1, ..., c_n) \vDash \phi(c_1, ..., c_n)$ in the language of L'.

Before we prove the claim, we show how this implies (1). By (2) and the claim, we see $T \cup \psi(c_1, ..., c_n) \vDash \phi(c_1, ..., c_n)$. Then by compactness, there is $\psi \in \Psi$ such that $T \cup \{\psi(c_1, ..., c_n)\} \vDash \phi(c_1, ..., c_n)$. However, this means $T \vDash \psi(c_1, ..., c_n) \Rightarrow \phi(c_1, ..., c_n)$. Now, since we just added the new constant symbols $c_1, ..., c_n$ and T is an L-theory, this implies $T \vDash \forall x(\psi(x_1, ..., x_n) \Rightarrow \phi(x_1, ..., x_n))$. Thus we get ϕ is T-equivalent to ψ as desired.

Now it remains to prove the claim.

Suppose the claim is false and we derive a contradiction. Then there is $\mathcal{M} \vDash T$ and $a_1, ..., a_n \in M$ such that $\mathcal{M} \vDash \psi(a_1, ..., a_n)$ for all $\psi \in \Psi$ and $\mathcal{M} \vDash \neg\phi(a_1, ..., a_n)$.

Consider

$$T' := T \cup qf\,\mathrm{Th}(\mathcal{M}_{\{a_1,\dots,a_n\}}) \cup \{\phi(\underline{a_1},\dots,\underline{a_n})\}$$

in the language $L \cup \{\underline{a_1},\dots,\underline{a_n}\}$ with $\underline{a_i}$ new constant symbols. In particular,

$$qf\,\mathrm{Th}(\mathcal{M}_{\{a_1,\dots,a_n\}}) = \{\psi(\underline{a_1},\dots,\underline{a_n}) : \psi(x_1,\dots,x_n) \text{ is q.f. and } \mathcal{M} \vDash \psi(a_1,\dots,a_n)\}$$

Subclaim: T' is consistent.

Proof of subclaim: if not then by compactness there is a q.f. $\psi(x_1,\dots,x_n)$ with $\mathcal{M} \vDash \psi(a_1,\dots,a_n)$ such that

$$T \cup \{\psi(\underline{a_1},\dots,\underline{a_n}), \phi(\underline{a_1},\dots,\underline{a_n})\}$$

is inconsistent. Therefore, we get

$$T \vDash \phi(\underline{a_1},\dots,\underline{a_n}) \Rightarrow \neg\psi(\underline{a_1},\dots,\underline{a_n})$$

Hence we get $T \vDash \forall x(\phi(x) \Rightarrow \neg\psi(x))$ and hence $\neg\psi \in \Psi$ by definition. Then $\mathcal{M} \vDash \neg\psi(a_1,\dots,a_n)$, which contradicts $\mathcal{M} \vDash \psi(a_1,\dots,a_n)$. This proves the subclaim.

Now we back to the proof of the claim. Let $\mathcal{N}' \vDash T'$ and let $\mathcal{N}$ be the L-reduct of $\mathcal{N}'$. Thus we have $\mathcal{N} \vDash T$. Let $\mathcal{A}$ be the substructure of $\mathcal{M}$ generated by $\{a_1,\dots,a_n\}$, i.e. it is the smallest substructure of $\mathcal{M}$ contains the set $\{a_1,\dots,a_n\}$. Note this exists since $n > 0$ or $L^c \neq \varnothing$. Then, it is not hard to see we get the universe of $\mathcal{A}$ is equal $A = \{t^{\mathcal{M}}(a_1,\dots,a_n) : t(x_1,\dots,x_n) \text{ is L-terms}\}$ (left as an exercise).

Thus we get $\mathcal{M} \vDash T$ and $\mathcal{A} \subseteq \mathcal{M}$, we would like $\mathcal{A}$ to be a substructure of $\mathcal{N}$ as well. Now, recall $\mathcal{N}' \vDash qf\,\mathrm{Th}(\mathcal{M}_{\{a_1,\dots,a_n\}})$, let $j : A \to N$ given by

$$t^{\mathcal{M}}(a_1,\dots,a_n) \mapsto t^{\mathcal{N}}(\underline{a_1}^{\mathcal{N}'},\dots,\underline{a_n}^{\mathcal{N}'})$$

We leave it as an exercise to show j is an L-embedding. This tells us $\mathcal{A}$ is a substructure of $\mathcal{N}$ as desired.

Since $\phi(\underline{a_1},\dots,\underline{a_n}) \in T'$, we have $\mathcal{N} \vDash \phi(a_1,\dots,a_n)$. Apply our assumption (2) to get $\mathcal{M} \vDash \phi(a_1,\dots,a_n)$. However, this is a contradiction as we assumed $\mathcal{M} \vDash \neg\phi(a_1,\dots,a_n)$.

This concludes the proof of the claim and hence the proof of the theorem. ♡

Definition 3.9.4. Let L be a language, then an ***L-literal*** is an atomic formula or an negated atomic formula.

Definition 3.9.5. Let T be L-theory, we say T is ***eliminable*** if, for all $\mathcal{M} \vDash T$ and $\mathcal{N} \vDash T$ with common substructure $\mathcal{A} \subseteq \mathcal{M}$ and $\mathcal{A} \subseteq \mathcal{N}$, and L-formula $\psi(y)$ is a conjunction of L_A-literals in a single variable y, then if $\psi(y)$ has a solution in $\mathcal{M}$ then it has a solution in $\mathcal{N}$.

Theorem 3.9.6 (Criterion for QE). *Let T be L-theory. If T is eliminable then T admits quantifier elimination.*

Proof. Suppose $\phi(x)$ with $x = (x_1, ..., x)$ is an L-formula. We show using the fact T is eliminable and by induction on complexity, that ϕ is T-equivalent to a q.f. formula in x.

Now:

1. If $\phi(x)$ is atomic then we are done.
2. $\wedge, \vee, \neg$ are also immediate since being T-equivalent to q.f. is closed under these operations.
3. $\forall$ is equivalent to $\neg\exists\neg$.

Hence, we are only left to show $\phi(x)$ is of the form $\exists y \psi(x, y)$ then it is T-equivalent to q.f. formula. By induction hypothesis we have ψ is T-equivalent to a q.f. formula $\tilde{\psi}(x, y)$.

Writing $\tilde{\psi}$ in disjuncted normal form we may assume $\tilde{\psi}(x, y)$ is of the form

$$\bigvee_i \bigwedge_j \psi_{ij}(x, y)$$

where ψ_{ij} are L-literals. So it suffices to show that $\exists(\bigvee_i \bigwedge_j \psi_{ij}(x, y))$ is T-equivalent to q.f. formula in x.

We use the characterization from the last theorem. Let $\mathcal{M}, \mathcal{N}, \mathcal{A}$ be given, then we see $\mathcal{M} \vDash \exists y(\bigvee_i \bigwedge_j \psi_{ij}(a, y))$ implies $\mathcal{M} \vDash \exists y(\wedge_j \psi_{ij}(a, y))$ for some i. However, note $\bigwedge_j \psi_{ij}(x, y)$ is just a conjunction of L_A-literals and recall the definition of eliminable, we get $\mathcal{N} \vDash \exists(\bigwedge_j \psi_{ij}(a, y))$ for some particular i. This means $\mathcal{N} \vDash \exists(\bigvee_i \bigwedge_j \psi_{ij}(x, y))$. Then the converse is also true by symmetry and so by last theorem, we get ϕ is T-equivalent to q.f. formula as desired.

♡

Example 3.9.7. Let $L = \varnothing$ and T be the theory of infinite sets. We claim T admits QE.

Solution. We apply the criterion and check if it is eliminable. Let $\mathcal{M} \vDash T$ and $\mathcal{N} \vDash T$ with $\mathcal{A} \subseteq \mathcal{M}, \mathcal{N}$ be a common substructure. Now consider $\psi(y)$ be a conjunction of L_A-literals, which we assume it has a solution in $\mathcal{M}$, we need to show it has a solution in $\mathcal{N}$.

What do the conjunctions in $\psi(y)$ look like? Well, we have:

1. $y = y$ and $y \neq y$.
2. $y = a$ and $y \neq a$ for $a \in A$.
3. $a = b$ and $a \neq b$ for $a \in A$.

Those are all the possible L_A-literals we get. However, we can throw out the third term in $\psi(y)$ since we don't care about it as it add no constraint on the feasibility of our system of eqations. Similarly we can throw the first term (1) out since it either is sovled by everything (the case $y = y$) or we get no solution at all (the case $y \neq y$).

Hence, the only meaningful equations we get is $y = a$ and $y \neq a$ for $a \in A$. In other word, we may assume $\psi(y)$ is of the form

$$\bigwedge_{i=1}^{r} y = a_i \wedge \bigwedge_{j=1}^{s} y \neq b_j$$

with $a_i, b_j \in A$.

If $n \geq 1$ then a_1 is a solution to $\psi(y)$ in $\mathcal{M}$ where $a_1 \in A$ is also in $\mathcal{N}$. Hence we are done.

If $n = 0$ then take any a that $a \neq b_j$ for all j, such element in A exists as A is infinite, thus we get a solution in $\mathcal{N}$ as well and we are done.

♠

Example 3.9.8. DLO admits QE.

Solution. Say $(M, <)$ and $(N, <)$ models DLO with $(A, <)$ a common substructure. Let $\psi(y)$ a conjunction of L_A-literals which has a realization in M, we want to show $\psi(y)$ has a realization in N.

Again, we write down possible conjuncts in $\psi(y)$:

1. $y = y$ and $y \neq y$.
2. $y < y$ and $y \geq y$.
3. $y = a$ for $a \in A$.
4. $y \neq a$ for $a \in A$.
5. $y < a$ for $a \in A$.
6. $y \geq a$ for $a \in A$.
7. $a < y$ for $a \in A$.
8. $a \geq y$ for $a \in A$.
9. Literals not contain y.

Like before, we can throw away (1) and (2). If (3) appears then we immediately get a solution, hence we may assume they do not appear. We have (4) and (5) are allowed. For (6) we can re-write in terms of (3) and (7) and so we can throw away. For (7) we have to keep. For (8) we can rewrite as $y = a$ or $y < a$ and so we can throw away. Finally for (9) it adds no constraint and so we can throw away.

In other word, we are left with $(4), (5)$ and (7). Hence we may assume $\psi(y)$ is of the form

$$\bigwedge_{i} y > q_i \wedge \bigwedge_{j} y \neq b_j \wedge \bigwedge_{k} y < c_k$$

and since we have a solution to this in M, we must have $a_i < c_k$ for all i, k. However, by density of M and the fact it has no end points, we must be able to find infinitely many elements between $a_i < c_k$ in A, i.e. we can avoid those finitely many $y \neq b_j$. In other word, we get a desired solution in $\mathcal{N}$ as A is subset of $\mathcal{N}$.

If $(M, <)$ is a model of DLO, then the definable subsets of M^1 are finite unions of points and intervals. Such a theory in language containing $<$ are called ***o-minimal***, i.e. if the theory's definable susbets of M^1 are just finite unions of points and integers by the $<$ relation, then it is o-minimal. ♠

Example 3.9.9. ACF admits QE for $p = 0$ or prime.

Solution. Let K, L be two models of ACF and R a substructure (subring) of both K and L. Let $\psi(y)$ be conjunctions of L_R-literals with a solution in K, we want to show $\psi(y)$ has a solution in L.

First, note $\operatorname{Char}(K) = \operatorname{Char}(L)$ as they share a common subring R. Second, note R is integral domain and hence we get $\operatorname{Frac}(R)$ is contained in both K and L. Thrid, observe by uniqueness of algebraic closure we get $F := \overline{\operatorname{Frac}(R)}$ is a common subfield of K and L.

In particular, note $\psi(y)$ are of the form

$$\left(\bigwedge_{i=1}^{r} P_i(y) = 0\right) \wedge \left(\bigwedge_{j=1}^{s} Q_j(y) \neq 0\right)$$

with $P_i, Q_j \in R[y] \subseteq F[y]$ where we may assume P_i, Q_j are not constant polynomials.

If $r \neq 0$ then a solution $\psi(y)$ in K will have to be in $\overline{\operatorname{Frac}(R)}$, which is a subfield of L as well and hence we are done. If $r = 0$, then we are only excluding finitely many solutions and sicne we have infinitely many elements in the algebraic closure of $\operatorname{Frac}(R)$, we are bound to find some elements avoids them all. Hence we are done as well. ♠

Corollary 3.9.9.1. *Every definable sets in a model of ACF is Zariski-constructible.*

Definition 3.9.10. T is ***model-complete*** if $\mathcal{M} \models T$ and $\mathcal{N} \models T$ with $\mathcal{M} \subseteq \mathcal{N}$ then $\mathcal{M} \preccurlyeq \mathcal{N}$.

Remark 3.9.11. If T is QE then Q is model-complete. The converse is false.

In particular, we then get:

1. $(\mathbb{Q}, <) \preccurlyeq (\mathbb{R}, <)$ since DLO admits QE.
2. $(\overline{\mathbb{Q}}, 0, 1, +, -, \times) \preccurlyeq (\overline{\mathbb{C}}, 0, 1, +, -, \times)$ since ACF admits QE.

Theorem 3.9.12 (Hilbert's Weak Nullstellensatz)**.** *Let $K \models ACF$ and $P_1, ..., P_l \in K[x_1, ..., x_n]$. Suppose there does not exists $Q_1, ..., Q_l \in K[x_1, ..., x_n]$ such that $Q_1P_1 + Q_2P_2 + ... + Q_lP_l = 1$ then there is $a \in K^n$ such that $P_i(a) = 0$ for all i.*

Proof. By model-completeness, it suffices to find field extension L extending K with $L \models$ ACF, such that $L \models \exists x(\bigwedge_{i=1}^{l} P_i(x) = 0)$, which is an L_K-sentence. This is because $K \subseteq L$ and since ACF admits QE we get $K \preccurlyeq L$ and so the sentence is also true in K and we get $K \models \exists x(\bigwedge_{i=1}^{l} P_i(x) = 0)$ and the proof follows.

To find such L, we consider the ideal I generated by $P_1, ..., P_l$ in $K[x_1, ..., x_n]$. Note this ideal I must be proper by the assumption we have no $Q_1, ..., Q_l$ so $\sum Q_i P_i = 1$. Then let M be the proper maximal ideal containing I and consider the field

$F = K[x_1, ..., x_n]/M$, which is a field extension of K. Then the field $L = \overline{F}$ will the job as each $P_1, ..., P_l$ are inside M and proof follows.

♡