MATHEMATICAL LOGIC 1
LECTURE NOTES

Contents

1. Model Theory 2
1.1. Introduction 2
1.2. Warm-up: Propositional Logic 3
1.3. First-Order Structures 4
1.4. Relationships between structures 9
2. Relations between structures (continued) 11
2.1. First-Order Theories 12
2.2. Definable Sets 14
2.3. Compactness Theorem 17
2.4. Formal proof 17
3. Direct proof of Compactness 18
3.1. Henkin Contruction 18
4. Complete Theories 24
4.1. Remark: 24
4.2. Example: 24
4.3. Example: 24
4.4. Example: 24
4.5. Example: 24
4.6. Remark: 24
4.7. Corollary: 25
5. Up and Down 25
5.1. Recall: 25
5.2. Example: 25
6. Downward Lowenheim-Skolem Theorem 26
6.1. Corollary 26
7. Downward Löwenheim-Skolem 28
8. Elementary Chains 28
9. Back-n-Forth 29
10. Back-n-Forth (Continued) 29
11. Quantifier Elimination 30
12. Computability Theory Introduction 33
13. Partial Recursive Functions 34
14. Turing Machines 37
15. Turing Machines (Continued) 38
15.1. Examples 38
15.2. Partial recursive functions vs. Turing computable functions 39
15.3. Universal Turing Machine 39

16. Computable and Computably Enumerable Sets 41
16.1. Normal Form Theorem 42
17. Computably enumerable (c.e.) sets 43
18. Creative sets 45
19. Construction of a simple set 45
20. Consequences of simple sets 46
21. Many-one Reducibility 47
22. Index sets 48
23. Many-one degrees 49
24. Index sets 50
25. Gödel's Incompleteness Theorem 52
25.1. Peano Arithmetic 52
25.2. Gödel Numbering 53
26. Decidable vs. Undecidable Theories (for a fixed $\mathcal{L}$) 56
27. Set Theory 57
28. Axioms of ZFC 57
29. Consequences of Extensionality and Comprehension 57
30. More Sets 58
31. Relations 59
32. Well-Orderings 59
33. Ordinals 61
34. Ordinals 64
35. Ordinal Arithmetic 65
36. Ordinal Arithmetic 66
37. Classes and Recursion 66
38. Classes and Recursion 67
39. Cardinals 67
40. Cardinal Arithmetic 69
40.1. Choice and cardinals 71
40.2. Cardinal exponentiation 72

1. Model Theory

1.1. **Introduction.** This course is the first of a two-part introduction to mathematical logic at the graduate level. The main topics we will cover are model theory, computability theory, and set theory. We will begin with an introduction to model theory.

Model theory is the study of mathematical structures through the lens of logic. In this class, we will restrict our attention to first-order logic, which is a sufficiently rich language that allows us to express quantified statements, such as statements of the form "For every x, there is a y such that ...." In general, one can study the model theory of any logic, such as higher order logics, continuous logic, modal logic, and so on, but these topics are well beyond the scope of this course.

What do we gain by studying mathematical structures using first-order logic? First, doing so yields insight into the expressive power of various logical languages. For instance, we can ask of a given structure: Which objects or collections of objects in this structure are

definable? Are there certain objects or collections of objects that are *not* definable? If so, what prevents them from being definable?

Second, the tools of model theory allow us to determine the consequences of a given set of axioms. This idea should be familiar: It is now well-known that the parallel postulate is not a consequence of the other axioms of Euclidean geometry. This was shown by exhibiting a non-Euclidean geometry in which all of the axioms of Euclidean geometry are true but the parallel postulate is false. In such a non-Euclidean geometry, points and lines are not understood as we typically understand them; for instance, in spherical geometry, we might take a point to be a pair of antipodal points (classically understood) on the surface of a sphere, and the lines are the great circles (circles of maximal circumference) on the surface of the sphere. In this setting, there are no parallel lines, as every two lines must intersect. This maneuver of reinterpreting points and lines was a prototype of what we now refer to as "giving a model." David Hilbert exploited this technique to prove that various axioms of geometry were *independent* of one another.

Third, using model theory, we can isolate interesting properties of mathematical theories (where a mathematical theory is just a collection of sentences in some first-order language). For instance, some theories are complete, which means roughly that they imply all of the first-order truths about a given structure. This is a nice feature for a theory to have, particularly if that theory can be listed by some algorithmic procedure. In this case, there is an algorithm for determining whether or not a given formula is a consequence of the theory. Significantly, as we will see, some theories such as the theory of Peano arithmetic are incomplete (however, in general, one needs the tools of computability theory to show that a given theory is incomplete).

1.2. **Warm-up: Propositional Logic.** Before we begin our study of model theory, it will be helpful to review some propositional logic. Propositional logic is a fairly straightforward logic, but it is not a very expressive logic. That is, one cannot express interesting mathematical statements in propositional logic (or at least not in such a way that allows us to understand the logical structure of such statements).

The language of propositional logic consists of infinitely many propositional symbols $P_1, P_2, \ldots, P, Q, R, \ldots$ and the following logical connectives:

- $\&$ "and" (sometimes $\wedge$),
- $\vee$ "or,"
- $\neg$ "not,"
- $\rightarrow$ "if-then," and
- $\leftrightarrow$ "if and only if"

We also explicitly include the left and right parentheses as part of the language to avoid ambiguity in more complex propositional formulas.

Propositional formulas are built up in the standard inductive way: Let A and B be propositional formulas. Then the following are propositional formulas:[1]

- $A \,\&\, B$;
- $A \vee B$;
- $\neg A$;

[1]Technically, we should include parentheses in each of these expressions as well, but this is not so important for our purposes.

- $A \to B$;
- $A \leftrightarrow B$.

We can fully understand a propositional formula by understanding its truth-conditions. Moreover, the truth-conditions of a propositional formula are given by its truth table. For instance, the following should be familiar:

P	Q	$P \,\&\, Q$	$P \vee Q$	$\neg P$	$P \to Q$	$P \leftrightarrow Q$
T	T	T	T	F	T	T
T	T	F	T	F	F	F
T	F	F	T	F	T	F
T	F	F	F	F	T	T

As a more complex example, consider the following:

P	Q	R	$P \vee (Q \,\&\, R)$
T	T	T	T
T	T	F	T
T	F	T	T
T	F	F	T
F	T	T	T
F	T	F	F
F	F	T	F
F	F	F	F

The key observation to make is this: to fully understand the truth-conditions of a propositional formula such as $P \vee (Q \,\&\, R)$, there are only eight possibilities to consider, where each possibility depends on the truth values assigned to P, Q, and R. In other words, each possibility depends on how we interpret P, Q, and R—as standing for a true proposition or a false proposition. Thus, we can think of each row of the above truth table as corresponding to a possible interpretation of the propositional symbols.

As we transition to first-order logic, we will similarly approach the truth-conditions of first-order sentences. However, when we extend from propositional logic to a more expressive, and hence more complex, language, one consequence is that it will be more complicated to describe the truth-conditional for formulas in our more expressive language.

1.3. **First-Order Structures.**

1.3.1. *First-order languages.* We would like to have a system in which we can express facts about mathematical structures in a completely formal way. As a motivating example, consider the task of expressing facts about the natural numbers $\mathbb{N}$. A "fact" might be something like "$3 = 3$", or "$1 + 2 = 3$". How might we express such statements?

From a syntactic perspective, we will require a number of symbols:

(i) $=$, $<$, (2-place predicate symbols)
(ii) $+$, $\times$, (2-place function symbols)
(iii) 0, 1, (distinguished constant symbols)

We are also going to need $\mathbb{N}$ itself, since in order to state facts about $\mathbb{N}$, we have to be able to refer to the elements of $\mathbb{N}$. $\mathbb{N}$ is called the *universe*, or *domain of discourse*.

Notice that once we have, for example, the symbols 1 and $+$, we can refer to any $n \in \mathbb{N}$ as $1 + \cdots + 1$. So although the universe $\mathbb{N}$ is infinite, we can use a finite number of symbols to pick out or "label" every element in $\mathbb{N}$. This motivates the following definition.

Definition 1.1. A *first-order language* $\mathcal{L}$ is given by the following:

(1) A set $\mathcal{F}$ of function symbols, and for each symbol f in $\mathcal{F}$, a positive number n_f.
(2) A set $\mathcal{R}$ of relation symbols, and for each symbol R in $\mathcal{R}$, a positive number n_R.
(3) A set $\mathcal{C}$ of constant symbols.

Remark. The positive numbers n_f and n_R indicate the *arity* of the associated function or relation symbol. That is, they indicate that the language will interpret f and R as a function/relation of n_f and n_R variables, respectively.

Remark. For a given first-order language, we allow that $\mathcal{F}$, $\mathcal{R}$, or $\mathcal{C}$ may be empty.

Example 1.2. Some first-order languages

(1) The language of arithmetic: $\{+, \times, 0, 1\}$
$\mathcal{F} = \{+, \times\}$, $\mathcal{R} = \emptyset$, $\mathcal{C} = \{0, 1\}$, $n_+ = 2$, $n_\times = 2$
(2) The language of groups: $\{\cdot, e\}$
$\mathcal{F} = \{\cdot\}$, $\mathcal{R} = \emptyset$, $\mathcal{C} = \{e\}$, $n_\cdot = 2$
(3) The language of graphs: $\{R\}$
$\mathcal{F} = \emptyset$, $\mathcal{R} = \{R\}$, $\mathcal{C} = \emptyset$, $n_R = 2$
(4) The language of sets: $\{\in\}$
$\mathcal{F} = \emptyset$, $\mathcal{R} = \{\in\}$, $\mathcal{C} = \emptyset$, $n_\in = 2$
(5) The language of pure sets: $\emptyset$
$\mathcal{F} = \emptyset$, $\mathcal{R} = \emptyset$, $\mathcal{C} = \emptyset$

1.3.2. *$\mathcal{L}$-structures.* Given a language $\mathcal{L}$, we can consider whether a given mathematical structure is "appropriate" for $\mathcal{L}$. Referring to the languages in Example 1.2, it is clear that e.g. the "language of groups" is not sufficient to express the same facts as the "language of arithmetic".

Obviously, the labels we gave these languages have given away the fact that we know in advance which mathematical structures are suited to them, but in general this will not be the case.

Definition 1.3. An *$\mathcal{L}$-structure* $\mathcal{M}$ is given by:

(1) a nonempty *universe* M (also called a *domain of discourse* or an *underlying set*);
(2) a function $f^{\mathcal{M}} : M^{n_f} \to M$ for each function symbol $f \in \mathcal{F}$;
(3) a relation $R^{\mathcal{M}} \subseteq M^{n_R}$ for each relation symbol $R \in \mathcal{R}$; and
(4) an element $c^{\mathcal{M}} \in M$ for each constant symbol $c \in \mathcal{C}$.

We say that $f^{\mathcal{M}}$, $R^{\mathcal{M}}$, and $c^{\mathcal{M}}$ are interpretations of the $\mathcal{L}$-symbols f, R, and c in the $\mathcal{L}$-structure $\mathcal{M}$. An $\mathcal{L}$-structure is commonly expressed as a tuple:

$$\mathcal{M} = (M, f^{\mathcal{M}}, R^{\mathcal{M}}, c^{\mathcal{M}} : f \in \mathcal{F}, R \in \mathcal{R}, c \in \mathcal{C})$$

Example 1.4. Let $\mathcal{L}_g = \{\cdot, e\}$, the language of groups as in Example 1.2. In general, an $\mathcal{L}_g$-structure will have the form $(G, \cdot_G, e_G)$. The following are $\mathcal{L}_g$-structures:

- $G_1 = (\mathbb{R}, \times, 1)$

- $G_2 = (\mathbb{Z}, +, 0)$
- $G_3 = (\mathbb{N}, +, 0)$

Note that G_3 is an $\mathcal{L}_g$-structure, although it is not a group. This is perfectly valid; an interpretation of a language implies nothing about truth or falsity.

The cardinality $|\mathcal{M}|$ of an $\mathcal{L}$-structure $\mathcal{M}$ is $|M|$, the cardinality of its universe.

1.3.3. *$\mathcal{L}$-formulas.* So far, we have only used our language $\mathcal{L}$ in a very basic way. Our original goal was to express facts, to describe properties of a structure. To do this, we must begin to build up a "grammar" for $\mathcal{L}$, which will require a number of additional symbols:

- the symbols in $\mathcal{L}$;
- variable symbols $v_0, v_1, \dots$;
- the equality symbol $=$;
- propositional connectives $\wedge$, $\vee$, $\neg$;
- quantifiers $\exists$, $\forall$; and
- parentheses (,).

Once our "grammar" is in place, we can begin to construct basic statements, called $\mathcal{L}$-terms, and use these to construct somewhat-less-basic statements called atomic $\mathcal{L}$-formulas. From the atomic $\mathcal{L}$-formulas, we can build $\mathcal{L}$-formulas.

Definition 1.5. The set of *$\mathcal{L}$-terms* is the smallest set T defined recursively such that

(1) $c \in T$ for each $c \in \mathcal{C}$;
(2) $v_i \in T$ for each variable symbol v_i $(i = 1, 2, \dots)$; and
(3) if $t_1, ..., t_{n_f} \in T$ and $f \in \mathcal{F}$ is a function symbol, then $f(t_1, ..., t_{n_f}) \in T$.

Example 1.6. Let $\mathcal{L} = \{+, \times, 0, 1\}$, the language of arithmetic. Then the following are examples of $\mathcal{L}$-terms:

- $\times(v_1, +(v_2, 1))$
- $+(1, +(1, +(1, 1)))$

$\mathcal{L}$-terms act like "nouns," in that they pick out some element of the universe. In the case of an $\mathcal{L}$-term consisting only of constant symbols, like $+(1, +(1, +(1, 1)))$, there will be only one element (in this case, the number 4). In the case of an $\mathcal{L}$-term which includes variable symbols, like $\times(v_1, +(v_2, 1))$, the $\mathcal{L}$-term can refer to multiple elements, depending on the values substituted in place of the variables.

Definition 1.7. An $\mathcal{L}$-formula ϕ is an *atomic* $\mathcal{L}$-formula if ϕ is of the form

(1) $t_1 = t_2$, where t_1, t_2 are terms; or
(2) $R(t_1, ..., t_{n_R})$, where $t_1, ..., t_{n_R}$ are terms.

An atomic $\mathcal{L}$-formula is a basic statement of the form "this is that," or "this is related to that." Once we have defined atomic $\mathcal{L}$-formulas, we can define the general set of all $\mathcal{L}$-formulas by closing under the logical connectives $\wedge, \vee, \neg$, and the quantifiers $\exists$ and $\forall$.

Definition 1.8. The set of *$\mathcal{L}$-formulas* is the smallest set $\mathcal{W}$ containing the atomic $\mathcal{L}$-formulas such that

(1) if $\phi \in \mathcal{W}$, then $\neg\phi \in \mathcal{W}$;
(2) if $\phi, \psi \in \mathcal{W}$, then $(\phi \wedge \psi) \in \mathcal{W}$ and $(\phi \vee \psi) \in \mathcal{W}$; and

(3) if $\phi \in \mathcal{W}$, then $\exists v_i \phi \in \mathcal{W}$ and $\forall v_i \phi \in \mathcal{W}$.

Example 1.9. Let $\mathcal{L} = \{+, \times, 0, 1\}$, the language of arithmetic. Then the following are examples of $\mathcal{L}$-formulas:

- $(v_1 = 0) \vee \neg(v_1 = 0)$
- $\exists v_2 \, (v_2 \times v_2 = v_1)$
- $\forall v_1 \exists v_2 \, (v_2 \times v_2 = v_1)$

Definition 1.10. A variable v is called a *free variable* if v is not in the scope of a quantifier $\exists v$ or $\forall v$. Otherwise, v is bound.

Example 1.11.

- $(v_1 = 0) \vee \neg(v_1 = 0)$ $\quad$ v_1 is free
- $\exists v_2 \, (v_2 \times v_2 = v_1)$ $\quad$ v_1 is free, v_2 is bound
- $\forall v_1 \exists v_2 \, (v_2 \times v_2 = v_1)$ $\quad$ v_1 and v_2 are bound

Definition 1.12. An $\mathcal{L}$-*sentence* is an $\mathcal{L}$-formula having no free variables.

Where an $\mathcal{L}$-formula picks out a property of a structure, an $\mathcal{L}$-sentence is either true or false in a structure. In this course, we will mainly be concerned with the interpretation of $\mathcal{L}$-sentences in $\mathcal{L}$-structures, rather than $\mathcal{L}$-formulas in general.

1.3.4. *Terms and Formulas.* Above we considered the language of groups, $\mathcal{L}_g = \{\cdot, e\}$, where $\cdot$ is a binary function symbol and e is a constant symbol. Recall that we also discussed two $\mathcal{L}_g$-structures, $(\mathbb{Z}, +, 0)$ and $(\mathbb{N}, +, 0)$. There is a sense in which $(\mathbb{Z}, +, 0)$ satisfies the axioms of the theory of groups, but $(\mathbb{N}, +, 0)$ does not. Our goal is to make this notion of satisfaction more precise.

Fix a first-order language $\mathcal{L}$. How do we interpret $\mathcal{L}$-terms in an $\mathcal{L}$-structure? Suppose that $\mathcal{M}$ is an $\mathcal{L}$-structure and t is a term built up from the variables

$$\bar{v} = (v_{i_1}, \ldots, v_{i_m}),$$

for some natural number m. We will interpret t as a function:

$$t^{\mathcal{M}} : M^m \to M.$$

Let s be a subterm of t, i.e. a term that occurs within t. Let $\bar{a} = (a_{i_1}, \ldots, a_{i_m}) \in M^m$. We inductively define $s^{\mathcal{M}}(\bar{a})$ as follows:

(1) if s is a constant symbol c, then $s^{\mathcal{M}}(\bar{a}) = c^{\mathcal{M}}$;
(2) if s is the variable v_{i_j} for $1 \leq j \leq m$, then $s^{\mathcal{M}}(\bar{a}) = a_{i_j}$; and
(3) if s is the term $f(t_1, \ldots, t_{n_f})$ for some $f \in \mathcal{F}$ and terms $t_1, \ldots, t_{n_f}$, then

$$s^{\mathcal{M}}(\bar{a}) = f^{\mathcal{M}}(t_1(\bar{a}), \ldots, t_{n_f}(\bar{a})).$$

Example 1.13. Let $\mathcal{L} = \{f, g, c\}$, where f is a unary function symbol, g is a binary function symbol and c is a constant symbol. Consider the following terms:

$$t_1 = g(v_1, c),$$
$$t_2 = f(g(v_1, c)), \text{ and}$$
$$t_3 = g(f(g(v_1, c)), v_2).$$

Let $\mathcal{M} = (\mathbb{N}, 2^x, +, 1)$. Then

$$t_1^{\mathcal{M}}(a_1, a_2) = g^{\mathcal{M}}(a_1, c^{\mathcal{M}}) = a_1 + 1,$$
$$t_2^{\mathcal{M}}(a_1, a_2) = f^{\mathcal{M}}(g^{\mathcal{M}}(a_1, c^{\mathcal{M}})) = 2^{a_1+1}, \text{ and}$$
$$t_3^{\mathcal{M}}(a_1, a_2) = g^{\mathcal{M}}(f^{\mathcal{M}}(g^{\mathcal{M}}((a_1, c^{\mathcal{M}})), a_2) = 2^{a_1+1} + a_2.$$

Now, how do we interpret $\mathcal{L}$-structures?

Definition 1.14. Let ϕ be an $\mathcal{L}$-formula with free variables $\bar{v} = (v_{i_1}, \ldots, v_{i_m})$ and let $\mathcal{M}$ be an $\mathcal{L}$-structure with universe M. For $\bar{a} = (a_{i_1}, \ldots, a_{i_m}) \in M^m$, we define $\mathcal{M} \models \phi(\bar{a})$ (read "$\mathcal{M}$ *satisfies* $\phi(\bar{a})$") inductively as follows:

(1) Atomic formulas:
- If ϕ is $t_1 = t_2$ for terms t_1, t_2, then $\mathcal{M} \models \phi(\bar{a}) \Leftrightarrow t_1^{\mathcal{M}} = t_2^{\mathcal{M}}$.
- If ϕ is $R(t_1, \ldots, t_{n_R})$ for $R \in \mathcal{R}$, then $\mathcal{M} \models \phi(\bar{a}) \Leftrightarrow (t_1^{\mathcal{M}}(\bar{a}), \ldots, t_{n_R}^{\mathcal{M}}(\bar{a})) \in R^{\mathcal{M}}$.

(2) General formulas:
- If ϕ is $\neg\psi$, then $\mathcal{M} \models \phi(\bar{a}) \Leftrightarrow \mathcal{M} \not\models \psi(\bar{a})$.
- If ϕ is $\psi \wedge \theta$, then $\mathcal{M} \models \phi(\bar{a}) \Leftrightarrow \mathcal{M} \models \psi(\bar{a})$ and $\mathcal{M} \models \theta(\bar{a})$.
- If ϕ is $\psi \vee \theta$, then $\mathcal{M} \models \phi(\bar{a}) \Leftrightarrow \mathcal{M} \models \psi(\bar{a})$ or $\mathcal{M} \models \theta(\bar{a})$.

(3) Quantifiers
- If ϕ is $\exists v_j\ \phi$, then $\mathcal{M} \models \phi(\bar{a}) \Leftrightarrow$ there is some $b \in M$ such that $\mathcal{M} \models \psi(\bar{a}, b)$.
- If ϕ is $\forall v_j\ \phi$, then $\mathcal{M} \models \phi(\bar{a}) \Leftrightarrow$ for all $b \in M$, $\mathcal{M} \models \psi(\bar{a}, b)$.

The statement "$\mathcal{M} \models \phi(\bar{a})$" may also be read "$\phi(\bar{a})$ is true in $\mathcal{M}$."

To reiterate: $\mathcal{L}$-sentences are either true or false in a given $\mathcal{L}$-structure, while $\mathcal{L}$-formulas express properties of elements in the domain (or tuples of elements of the domain).

Remark.

(1) Because of the expressive power of $\{\wedge, \exists\}$, the connectives $\vee$ and $\forall$ are redundant:

$$\phi \vee \psi \Leftrightarrow \neg(\neg\phi \wedge \neg\psi) \text{ and } \forall x\ \phi \Leftrightarrow \neg\exists\neg\phi.$$

Below, when we prove results via induction on $\mathcal{L}$-formulas, we need not consider cases when $\vee$ or $\forall$ are the main connectives.

(2) For variables, hereafter we will also use the letters w, x, y, z and so on, rather than just the indexed variables v_{i_j}.

(3) Also, note that $\forall$ and $\exists$ are first-order quantifiers, and thus only range over elements of the domain; alone, they lack the ability to make statements about subsets of the domain. Thus statements like "if a subset of $\mathbb{R}$ is nonempty and bounded above, it has a least upper bound" are inexpressible in first-order logic if we take our domain to be $\mathbb{R}$. However, second-order logic permits quantification over subsets of the domain, third-order logic permits quantification over subsets of a collection of subsets of the domain, and so on, but we will not be considering higher-order logics in this course.

1.4. **Relationships between structures.** In abstract algebra or linear algebra, one commonly studies relationships between certain structures, such as:

- groups and subgroups,
- vector spaces and subspaces, and
- maps between structures (such as homomorphisms, isomorphisms, linear transformations, etc.).

We now consider this phenomenon from a model-theoretic point of view.

Definition 1.15. Let $\mathcal{M}$, $\mathcal{N}$ be $\mathcal{L}$-structures for a fixed language $\mathcal{L}$, with universes M and N, respectively. An *$\mathcal{L}$-embedding* $\eta : \mathcal{M} \to \mathcal{N}$ (sometimes written as $\eta : M \to N$), is an injective map between the universes of $\mathcal{M}$ and $\mathcal{N}$ which preserves the interpretation of the symbols in $\mathcal{L}$. That is,

(1) $\eta(f^{\mathcal{M}}(a_1, ..., a_{n_f})) = f^{\mathcal{N}}(\eta(a_1), ..., \eta(a_{n_f}))$ for all function symbols $f \in \mathcal{F}$ and $a_1, ..., a_{n_f}$ in M;
(2) $(a_1, ..., a_{m_R}) \in R^{\mathcal{M}} \Leftrightarrow (\eta(a_1), ..., \eta(a_{m_R})) \in R^{\mathcal{N}}$ for all relation symbols R in $\mathcal{R}$ and $a_1, ..., a_{m_R}$ in M; and
(3) $\eta(c^{\mathcal{M}}) = c^{\mathcal{N}}$ for all constant symbols c in $\mathcal{C}$.

A special kind of $\mathcal{L}$-embedding is an $\mathcal{L}$-isomorphism.

Definition 1.16. Let $\mathcal{M}$ and $\mathcal{N}$ be $\mathcal{L}$-structures for a fixed language $\mathcal{L}$. If $\eta : \mathcal{M} \to \mathcal{N}$ is an $\mathcal{L}$-embedding and η is a bijection, then η is a *$\mathcal{L}$-isomorphism*.

Using the notion of an $\mathcal{L}$-embedding we can also define the substructure relation between two $\mathcal{L}$-structures.

Definition 1.17. Let $\mathcal{M}$ and $\mathcal{N}$ be $\mathcal{L}$-structures for a fixed language $\mathcal{L}$. If $M \subseteq N$, and the inclusion map that sends each element of M into N is an $\mathcal{L}$-embedding, we say that $\mathcal{M}$ is a *substructure* of $\mathcal{N}$ or that $\mathcal{N}$ is an *extension* of $\mathcal{M}$, and write $\mathcal{M} \subseteq \mathcal{N}$.

Let us consider some examples of embeddings and substructures.

Example 1.18.

(1) In the language of groups, $(\mathbb{Z}, +, 0)$ is a substructure of $(\mathbb{Q}, +, 0)$, which is a substructure of $(\mathbb{R}, +, 0)$.
(2) Again in the language of groups, if $\eta : \mathbb{Z} \to \mathbb{R}$ satisfies $\eta(x) = e^x$, then η is an $\mathcal{L}$-embedding from $(\mathbb{Z}, +, 0) \to (\mathbb{R}, \times, 1)$.

We can characterize the substructure relation in terms of the satisfaction of quantifier-free formulas.

Definition 1.19. The set of *quantifier-free $\mathcal{L}$-formulas* for a fixed language $\mathcal{L}$ is the smallest set of $\mathcal{L}$-formulas which contains the atomic $\mathcal{L}$-formulas and is closed under the logical connectives.

Proposition 1.20. Suppose $\mathcal{M} \subseteq \mathcal{N}$. Let $\bar{a} \in M$ and $\phi(\bar{v})$ be a quantifier-free $\mathcal{L}$-formula. Then $\mathcal{M} \models \phi(\bar{a}) \Leftrightarrow \mathcal{N} \models \phi(\bar{a})$.

Proof. We first prove the following claim:

Claim 1.21. Suppose $\mathcal{M} \subseteq \mathcal{N}$. If $t(\bar{v})$ is a term and $\bar{b} \in M$, then $t^{\mathcal{M}}(\bar{b}) = t^{\mathcal{N}}(\bar{b})$.

Proof of Claim. We proceed by induction on the complexity of terms.

(1) If t is a constant c, then $c^{\mathcal{M}} = c^{\mathcal{N}}$, since $\mathcal{M} \subseteq \mathcal{N}$.

(2) If t is the variable x_i, then $t^{\mathcal{M}}(\bar{b}) = b_i = t^{\mathcal{N}}(\bar{b})$, since $\mathcal{M} \subseteq \mathcal{N}$.

(3) If $t = f(t_1, ..., t_n)$, where f is an n-ary function symbol, $t_1, ..., t_n$ are terms, and $t_i^{\mathcal{M}}(\bar{b}) = t_i^{\mathcal{N}}(\bar{b})$, then observe that since $\mathcal{M} \subseteq \mathcal{N}$, it must be the case that $f^{\mathcal{M}} = f^{\mathcal{N}}{\restriction}_{M^n}$, the function obtained by restricting f to M^n. Therefore

$$\begin{aligned} t^{\mathcal{M}}(\bar{b}) &= f^{\mathcal{M}}(t_1^{\mathcal{M}}(\bar{b}), ..., t_n^{\mathcal{M}}(\bar{b})) \\ &= f^{\mathcal{N}}(t_1^{\mathcal{M}}(\bar{b}), ..., t_n^{\mathcal{M}}(\bar{b})) && (\text{since } f^{\mathcal{M}} = f^{\mathcal{N}}{\restriction}_{M^n}) \\ &= f^{\mathcal{N}}(t_1^{\mathcal{N}}(\bar{b}), ..., t_n^{\mathcal{N}}(\bar{b})) && (\text{by IH}) \\ &= t^{\mathcal{N}}(\bar{b}). \end{aligned}$$

□

We can now prove the proposition by induction on the complexity of formulas.

(1) If ϕ is $t_1 = t_2$, then

$$\begin{aligned} \mathcal{M} \models \phi(\bar{a}) &\Leftrightarrow t_1^{\mathcal{M}}(\bar{a}) = t_2^{\mathcal{M}}(\bar{a}) \\ &\Leftrightarrow t_1^{\mathcal{N}}(\bar{a}) = t_2^{\mathcal{M}}(\bar{a}) && (\text{by claim}) \\ &\Leftrightarrow \mathcal{N} \models \phi(\bar{a}). \end{aligned}$$

(2) If ϕ is $R(t_1, ..., t_n)$, then

$$\begin{aligned} \mathcal{M} \models \phi(\bar{a}) &\Leftrightarrow (t_1^{\mathcal{M}}(\bar{a}), ..., t_n^{\mathcal{M}}(\bar{a})) \in R^{\mathcal{M}} \\ &\Leftrightarrow (t_1^{\mathcal{M}}(\bar{a}), ..., t_n^{\mathcal{M}}(\bar{a})) \in R^{\mathcal{N}} && (\text{since } \mathcal{M} \subseteq \mathcal{N}) \\ &\Leftrightarrow (t_1^{\mathcal{N}}(\bar{a}), ..., t_n^{\mathcal{N}}(\bar{a})) \in R^{\mathcal{N}} && (\text{by claim}) \\ &\Leftrightarrow \mathcal{N} \models \phi(\bar{a}). \end{aligned}$$

Suppose that the result holds for some quantifier-free formula ψ, and ϕ is $\neg\psi$. Then

$$\begin{aligned} \mathcal{M} \models \phi(\bar{a}) &\Leftrightarrow \mathcal{M} \models \neg\psi(\bar{a}) \\ &\Leftrightarrow \mathcal{M} \not\models \psi(\bar{a}) \\ &\Leftrightarrow \mathcal{N} \not\models \psi(\bar{a}) && (\text{by IH}) \\ &\Leftrightarrow \mathcal{N} \models \neg\psi(\bar{a}) \\ &\Leftrightarrow \mathcal{N} \models \phi(\bar{a}). \end{aligned}$$

Finally, suppose that the result holds for quantifier-free formulas ψ_1 and ψ_2, and ϕ is $\psi_1 \wedge \psi_2$. Then

$$\begin{aligned} \mathcal{M} \models \phi(\bar{a}) &\Leftrightarrow \mathcal{M} \models \psi_1(\bar{a}) \wedge \mathcal{M} \models \psi_2(\bar{a}) \\ &\Leftrightarrow \mathcal{N} \models \psi_1(\bar{a}) \wedge \mathcal{N} \models \psi_2(\bar{a}) && (\text{by IH}) \\ &\Leftrightarrow \mathcal{N} \models \phi(\bar{a}). \end{aligned}$$

□

Definition 1.22. For a fixed language $\mathcal{L}$, $\mathcal{L}$-structures $\mathcal{M}$ and $\mathcal{N}$ are called *elementarily equivalent*, denoted $\mathcal{M} \equiv \mathcal{N}$, if $\mathcal{M} \models \phi \Leftrightarrow \mathcal{N} \models \phi$, for all $\mathcal{L}$-sentences ϕ.

2. Relations between structures (continued)

Theorem 2.1. Suppose $j : \mathcal{M} \to \mathcal{N}$ is an $\mathcal{L}$-isomorphism. Then $\mathcal{M} \equiv \mathcal{N}$.

Proof. We first prove the following claim.

Claim 2.2. Suppose $j : \mathcal{M} \to \mathcal{N}$ is an $\mathcal{L}$-isomorphism, and suppose that t is a term with free variables $\bar{v} = (v_1, ..., v_n)$. For $\bar{a} = (a_1, ..., a_n) \in M$, let $j(\bar{a}) = (j(a_1), ..., j(a_n))$. Then $j(t^{\mathcal{M}}(\bar{a})) = t^{\mathcal{N}}(\bar{a})$.

Proof of Claim. We proceed by induction on the complexity of terms.

(1) If $t = c$, then

$$\begin{aligned} j(t^{\mathcal{M}}(\bar{a})) &= j(c^{\mathcal{M}}) \\ &= c^{\mathcal{M}} && (j\text{'s an } \mathcal{L}\text{-embedding}) \\ &= t^{\mathcal{N}}(j(\bar{a})). \end{aligned}$$

(2) If $t = v_i$, then

$$\begin{aligned} j(t^{\mathcal{M}}(\bar{a})) &= j(a_i) \\ &= t^{\mathcal{N}}(j(\bar{a})) \\ &= j(a_i). \end{aligned}$$

(3) If $t = f(t_1, ..., t_n)$ then

$$\begin{aligned} j(t^{\mathcal{M}}(\bar{a}) &= j(f^{\mathcal{M}}(t_1^{\mathcal{M}}(\bar{a})), ..., f^{\mathcal{M}}(t_n^{\mathcal{M}}(\bar{a}))) \\ &= f^{\mathcal{N}}(j(t_1^{\mathcal{M}}(\bar{a}), ..., t_n^{\mathcal{M}}(\bar{a}))) \\ &= f^{\mathcal{N}}(j(t_1^{\mathcal{N}}(\bar{a}), ..., t_n^{\mathcal{N}}(\bar{a}))) && (\text{by IH}) \\ &= t^{\mathcal{N}}(j(\bar{a})). \end{aligned}$$

□

We will show by induction on formulas that $\mathcal{M} \models \phi(a_1, ..., a_n) \Leftrightarrow \mathcal{N} \models \phi(j(a_1), ..., j(a_n))$.

(1) If $\phi(v)$ is $t_1 = t_2$, then

$$\begin{aligned} \mathcal{M} \models \phi(\bar{a}) &\Leftrightarrow t_1^{\mathcal{M}}(\bar{a}) = t_2^{\mathcal{M}}(\bar{a}) \\ &\Leftrightarrow j(t_1^{\mathcal{M}}(\bar{a})) = j(t_2^{\mathcal{M}}(\bar{a})) && (j\text{'s a function}) \\ &\Leftrightarrow t_1^{\mathcal{M}}(j(\bar{a})) = t_2^{\mathcal{N}}(j(\bar{a})) \\ &\Leftrightarrow \mathcal{N} \models \phi(j(\bar{a})). \end{aligned}$$

(2) If ϕ is $R(t_1, ..., t_n)$, then

$$\begin{aligned} \mathcal{M} \models \phi(\bar{a}) &\Leftrightarrow (t_1^{\mathcal{M}}(\bar{a}), ..., t_n^{\mathcal{M}}(\bar{a})) \in R^{\mathcal{M}} \\ &\Leftrightarrow (j(t_1^{\mathcal{M}}(\bar{a})), ..., j(t_n^{\mathcal{M}}(\bar{a}))) \in R^{\mathcal{N}} && (j\text{'s an embedding}) \\ &\Leftrightarrow (t_1^{\mathcal{N}}(j(\bar{a})), ..., t_n^{\mathcal{N}}(j(\bar{a}))) \in R^{\mathcal{M}} && (\text{claim}) \\ &\Leftrightarrow \mathcal{N} \models \phi(j(\bar{a})). \end{aligned}$$

(3) If ϕ is $\neg\psi$, then

$$\begin{aligned}\mathcal{M} \models \phi(\bar{a}) &\Leftrightarrow \mathcal{M} \models \neg\psi(\bar{a}) \\ &\Leftrightarrow \mathcal{M} \not\models \psi(\bar{a}) \\ &\Leftrightarrow \mathcal{N} \not\models \psi(j(\bar{a})) && \text{(by IH)} \\ &\Leftrightarrow \mathcal{N} \models \neg\psi(j(\bar{a})) \\ &\Leftrightarrow \mathcal{N} \models \phi(j(\bar{a})).\end{aligned}$$

(4) If ϕ is $\psi_1 \wedge \psi_2$, then

$$\begin{aligned}\mathcal{M} \models \phi(\bar{a}) &\Leftrightarrow \mathcal{M} \models \psi_1(\bar{a}) \wedge \mathcal{M} \models \psi_2(\bar{a}) \\ &\Leftrightarrow \mathcal{N} \models \psi_1(j(\bar{a})) \wedge \mathcal{N} \models \psi_2(j(\bar{a})) && \text{(by IH)} \\ &\Leftrightarrow \mathcal{N} \models \phi(j(\bar{a})).\end{aligned}$$

(5) If $\phi(v)$ is $\exists w\psi(\bar{v}, w)$, then

$$\begin{aligned}\mathcal{M} \models \phi(\bar{a}) &\Leftrightarrow \mathcal{M} \models \psi(\bar{a}, b) \text{ for some } b \in M \\ &\Leftrightarrow \mathcal{N} \models \psi(j(\bar{a}, c)) \text{ for some } c \in N && (j\text{'s onto}) \\ &\Leftrightarrow \mathcal{N} \models \phi(j(\bar{a})).\end{aligned}$$

□

2.1. **First-Order Theories.**

Definition 2.3. Let $\mathcal{L}$ be a language. An *$\mathcal{L}$-theory* T is a collection of $\mathcal{L}$-sentences.

Definition 2.4. An $\mathcal{L}$-structure $\mathcal{M}$ is a *model* of an $\mathcal{L}$-theory T, written $\mathcal{M} \models T$, if $\mathcal{M} \models \phi$ for every $\phi \in T$.

Definition 2.5. An $\mathcal{L}$-theory T is *satisfiable*, or equivalently, *T has a model*, if there exists some $\mathcal{L}$-structure $\mathcal{M}$ such that $\mathcal{M} \models T$.

Definition 2.6. A class $\mathcal{K}$ of structures is called an *elementary class* if there is an $\mathcal{L}$-theory T so that $\mathcal{K}$ is the set $\{\mathcal{M} : \mathcal{M} \models T\}$.

Given an elementary class $\mathcal{K}$ and a corresponding theory T, we say that T *axiomatizes the class* $\mathcal{K}$. Let us consider some examples of elementary classes.

Example 2.7 (Infinite sets)**.** Let $\mathcal{L} = \emptyset$. Let $\phi_n = \exists x_1, ..., x_n(\bigwedge_{i<j\leq n} x_i \neq x_j)$. Then ϕ_n essentially says that there are n distinct elements in the universe. Let $T = \{\phi_n : n \in \mathbb{N}\}$. Then we have that $\mathcal{M} \models T \Leftrightarrow |\mathcal{M}|$ is infinite.

Example 2.8 (Partial orders and variants)**.** Let $\mathcal{L} = \{\leq\}$. Consider the following axioms.

(O_1) $\forall x \forall y((x \leq y \wedge y \leq x) \rightarrow y = x)$
(O_2) $\forall x(x \leq x)$
(O_3) $\forall x \forall y \forall z((x \leq y \wedge y \leq z) \rightarrow x \leq z)$
(O_4) $\forall x \forall y(x \leq y \vee y \leq x)$
(O_5) $\forall x \forall y \exists z(x \leq z \leq y \wedge z \neq x \neq y)$
(O_6) $\forall x \exists y(x \neq y \wedge x \leq y)$

(O_7) $\forall x \exists y (x \neq y \wedge y \leq x)$

The axioms (O_1)-(O_3) are the axioms of partial orders.

(O_1)-(O_4) yield the theory of a linear order.

(O_1)-(O_5) yield the theory of a dense linear order.

(O_1)-(O_7) yield the theory of a dense linear order without endpoints (DLOWE).

Example 2.9 (Groups)**.** Let $\mathcal{L} = \{\cdot, e\}$, where $\cdot$ is a binary function symbol and e is a constant symbol. Also let $x \cdot y$ stand for $\cdot(x, y)$. Then the class of groups is axiomatized by:

(G_1) $\forall x \ e \cdot x = x \cdot e = x$,

(G_2) $\forall x \forall y \forall z \ (x \cdot y) \cdot z = x \cdot (y \cdot z)$,

(G_3) $\forall x \exists y \ x \cdot y = y \cdot x = e$,

Adding the following statement to the previous 3, axiomatizes the class of Abelian groups:

(G_4) $\forall x \forall y \ x \cdot y = y \cdot x$.

Now let $\phi_n(x)$ be the $\mathcal{L}$-formula:

$$\overbrace{x \cdot x \cdot x \cdot \cdots \cdot x}^{n} = e$$

We will write $nx = e$ for the statement above. Adding

(G_5) $\{\forall x (x = e \vee \neg \phi_n(x) \mid n \geq 2\}$

to axioms (G_1)-(G_3) axiomatizes the collection of torsion-free groups.

We can also axiomatize the class of groups in which all elements have order $\leq N$ by adding to the axioms the sentence

(G_6) $\forall x \bigvee_{n \leq N} \phi_n(x)$.

Remark. The same idea cannot be used to axiomatize the class of torsion groups, since the sentence

$$\forall x \bigvee_{n \in N} \phi_n(x)$$

is infinitely long and hence is not first-order.

Example 2.10 (Rings and fields)**.** Let $\mathcal{L}_r = \{+, -, \cdot, 0, 1\}$. The symbols $+$, $-$, and $\cdot$ are binary function symbols and 0 and 1 are constant symbols. The following axiomatize the class of rings:

the axioms of additive abelian groups,

(R_1) $\forall x \forall y \forall z \ (x - y = z \longleftrightarrow x = y + z)$

(R_2) $\forall x \ x \cdot 0 = 0$,

(R_3) $\forall x \forall y \forall z \ (x \cdot y) \cdot z = x \cdot (y \cdot z)$,

(R_4) $\forall x \ x \cdot 1 = 1 \cdot x = x$,

(R_5) $\forall x \forall y \forall z \ x \cdot (y + z) = (x \cdot y) + (x \cdot z)$

(R_6) $\forall x \forall y \forall z \ (x + y) \cdot z = (x \cdot z) + (y \cdot z)$.

To axiomatize the class of fields, we add to the axioms above the following:

(F_1) $\forall x \forall y\ x \cdot y = y \cdot x$,
(F_2) $\forall x\ (x \neq 0 \longleftrightarrow \exists y\ x \cdot y = 1)$.
We axiomatize the class of algebraically closed fields (ACF) by adding to the above axioms of fields the sentence

(F_3) $\forall a_0 \ldots \forall a_{n-1} \exists x\ x^n + \sum_{i=0}^{n-1} a_i \cdot x^i = 0$.

for each $n \in \mathbb{Z}^+$.
Let ψ_p be the $\mathcal{L}_r$-sentence:

$$\forall x\ \underbrace{x + \cdots + x}_{p} = 0,$$

which ensures that any field satisfying it has characteristic p. For $p > 0$ a prime, $\mathrm{ACF}_p = \mathrm{ACF} \cup \{\psi_p\}$ is the theory of ACF of characteristic p and $\mathrm{ACF}_0 = \mathrm{ACF} \cup \{\neg\psi_p \mid p \in w\}$ is the theory of algebraically closed fields of characteristic 0.

Example 2.11 (Peano arithmetic)**.** Let $\mathcal{L} = \{+, \cdot, s, 0\}$. Here, $+$ and $\cdot$ are binary function symbols, s is a unary function symbol, and 0 is a constant symbol. The axioms for Peano arithmetic are:

(PA_1) $\forall x\ s(x) \neq 0$,
(1) $\forall x\ (x \neq 0 \longrightarrow \exists y\ s(y) = x)$,
(2) $\forall x\ x + 0 = x$,
(3) $\forall x \forall y\ x + s(y) = s(x + y)$,
(4) $\forall x\ x \cdot 0 = 0$,
(5) $\forall x \forall y\ x \cdot s(y) = (x \cdot y) + x$,
(6) the sentence $Ind(\phi)$ for each formula $\phi(v, \overline{w})$, where $Ind(\phi)$ is :

$$\forall \overline{w}\ [(\phi(0, \overline{w} \wedge \forall v\ (\phi(v, \overline{w}) \longrightarrow \phi(s(v), \overline{w}))) \longrightarrow \forall x\ \phi_(x, \overline{w})]$$

$Ind(\phi)$ asserts that if $\overline{a} \in M$ and $X = \{m \in M \mid \mathcal{M} \models \phi(m, \overline{a})\}$ and $0 \in X$ and $s(m) \in X$ whenever $m \in X$, then $X = M$.

2.1.1. *Logical Consequence.*

Definition 2.12. Let T be an $\mathcal{L}$-theory and ϕ and $\mathcal{L}$-sentence. ϕ is a logical consequence of T, written $T \models \phi$, if $\mathcal{M} \models T \Longrightarrow \mathcal{M} \models \phi$.

Proposition 2.13. Let T be the theory of groups where every element has order 2. Let ϕ be $\exists x_1 \exists x_2 \exists x_3\ (x_1 \neq x_2 \wedge x_2 \neq x_3 \wedge x_1 \neq x_3)$. Then $T \not\models \phi$.

Proof. $\mathbb{Z}/2\mathbb{Z} \models T$ and $\mathbb{Z}/2\mathbb{Z} \not\models \phi$. Hence $T \not\models \phi$.

□

2.2. **Definable Sets.**

Definition 2.14. Let $\mathcal{M} = (M, \ldots)$ be an $\mathcal{L}$-structure. We say $X \subseteq M^n$ is definable if there is an $\mathcal{L}$-formula $\phi(v_1, \ldots, v_n, w_1, \ldots, w_m)$ and $\overline{b} \in M^m$ such that $X = \{\overline{a} \in M^n \mid \mathcal{M} \models \phi(\overline{a}, \overline{b})$. We call $\overline{b}$ parameters. We say $\phi(\overline{x}, \overline{b})$ defines X. X is A-definable or *definable over* A if there is a formula $\psi(\overline{v}, w_1, \ldots, w_m)$ and parameters $\overline{b} \in A^m$ such that $\psi(\overline{v}, \overline{b})$ defines X.

Example 2.15. Define $<$ in $\mathbb{N}$, $\mathbb{R}$, $\mathbb{Z}$.

- Let $\mathcal{M} = (\mathbb{N}, +, \cdot, s, 0)$. Let $\phi(x, y)$ be $\exists z\ x + z = y \wedge z \neq 0$. Then $\mathcal{M} \models \phi(m, n) \Longleftrightarrow m < n$.

Remark. Once $<$ is defined, we can define $\{(m,n) \in \mathbb{N}^2 \mid m < n\}$.

- Let $\mathcal{M} = (\mathbb{R}, +, \cdot, -, 0, 1)$. Let $\phi(x,y)$ is $\exists z\ x + z^2 = y \wedge z \neq 0$. Then $\mathcal{M} \models \phi(x,y) \iff x < y$.
- Let $\mathcal{M} = (\mathbb{Z}, +, -, 0, 1)$. We will use Lagrange's theorem that states that every non-negative integer can be expressed as a sum of 4 squares. Then let $\phi(x,y)$ be: $\exists z_1 \exists z_2 \exists z_3 \exists z_4\ (z_1 \neq 0 \wedge y = x + z_1^2 + z_2^2 + z_3^2 + z_4^2)$. Then $\mathcal{M} \models \phi(m,n) \iff m < n$.

Remark. In all these cases, there were no parameters. Hence $<$ is $\emptyset$-definable.

Example 2.16. *Definable set with parameters*

Let $\mathcal{M} = (\mathbb{Q}, +, -, \cdot, 0, 1)$ where $\mathbb{Q}$ is the universe, $+, -, \cdot$ are binary function symbols, and 0 and 1 are constant symbols. Let $p(x) \in \mathbb{Q}[x]$. Let $Y = \{x \in \mathbb{Q} \mid p(x) = 0\}$. We claim Y is definable. Let $p(x) = \sum_{i=0}^{n} a_i x^i$ for $a_0, \ldots, a_n \in \mathbb{Q}$. Let $\phi(v, w_0, \ldots, w_n)$ be

$$w_n \cdot v^n + \cdots + w_1 \cdot v + w_0 = 0.$$

Then $Y = \{x \in \mathbb{Q} \mid \mathcal{M} \models \phi(x, a_0, \ldots, a_n)\}$. Let $A \supseteq \{a_0, \ldots, a_n\}$. Then Y is A-definable.

Proposition 2.17. Let $\mathcal{M}$ be an $\mathcal{L}$-structure with $\mathcal{L} = \{f \in \mathcal{F}, R \in \mathcal{R}, c \in \mathcal{C}\}$. Suppose that D_n is a collection of subsets of M^n for all $n \geq 1$, and $\mathcal{D} = (D_n \mid n \geq 1)$ is the smallest collection such that:

(1) $M^n \in D_n$;
(2) for all n-ary $f \in \mathcal{F}$, $\text{graph}(f^{\mathcal{M}}) \in D_{n+1}$;
(3) for all n-ary $R \in \mathcal{R}$, $R^{\mathcal{M}} \in D_n$;
(4) for all $i, j \leq n$, $\{(x_1, \ldots, x_n) \in M^n \mid x_i = x_j\} \in D_n$;
(5) if $X \in D_n$, then $M \times X \in D_{n+1}$;
(6) each D_n is closed under complement, finite intersections, and finite unions;
(7) if $X \in D_{n+1}$ and $\pi : M^{n+1} \to M^n$ is the projection $\pi(x_1, \ldots, x_{n+1}) = (x_1, \ldots, x_n)$, then $\pi(X) \in D_n$; and
(8) if $X \in D_{n+m}$ and $\overline{b} \in M^m$, then $\{\overline{a} \in M^n \mid (\overline{a}, \overline{b}) \in X\} \in D_n$.

Then $X \subseteq M^n$ is definable $\iff X \in D_n$.

Proof. ($\Longleftarrow$) We show that the definable sets satisfy 1-8. Since $\mathcal{D}$ is the smallest set satisfying 1-8, we will have $\mathcal{D} \subseteq$ definable sets.

(1) $M^n = \{(x_1, \ldots, x_n) \in M^n \mid \mathcal{M} \models x_1 = x_2\}$.
(2) $\text{graph}(f^{\mathcal{M}}) = \{(x_1, \ldots x_n, y) \in M^{n+1} \mid \mathcal{M} \models f^{\mathcal{M}}(x_1, \ldots, x_n) = y\}$.
(3) $R^{\mathcal{M}}$ is defined by R.
(4) $\{x \in M^n \mid x_i = x_j\}$ is defined by $v_i = v_j$.
(5) If $X \subseteq M^n$ is defined by $\phi(v_1, \ldots, v_n, \overline{a})$, $M \times X$ is defined by $\phi(v_2, \ldots, v_{n+1}, \overline{a})$. $M \times X = \{(x_1, \ldots, x_{n+1}) \in M^{n+1} \mid \mathcal{M} \models \phi(x_2, \ldots, x_{n+1}, \overline{a})\}$.
(6) If $X \subseteq M^n$ is defined by $\phi(\overline{v}, \overline{a})$ and $Y \subseteq M^n$ is defined by $\psi(\overline{v}, \overline{b})$, then
   - $M \setminus X$ is defined by $\neg\phi(\overline{v}, \overline{a})$.
   - $X \cap Y$ is defined by $\phi(\overline{v}, \overline{a}) \wedge \psi(\overline{v}, \overline{b})$.
   - $X \cup Y$ is defined by $\phi(\overline{v}, \overline{a}) \vee \psi(\overline{v}, \overline{b})$.
(7) If $X \subseteq M^{n+1}$ is defined by $\phi(v_0, \ldots, v_{n+1}, \overline{a})$, then $\pi(X)$ is defined by $\exists v_{n+1} \phi(v_1, \ldots, v_{n+1}, \overline{a})$.
(8) If $X \subseteq M^{n+m}$ is defined by $\phi(v_1, \ldots, v_{n+m}, \overline{c})$ and $\overline{b} \in M^m$, then $\{\overline{a} \in M^n \mid (\overline{a}, \overline{b}) \in X\}$ is defined by $\phi(v_1, \ldots, v_n, \overline{b}, \overline{c})$.

($\Longrightarrow$) Suppose $X \subseteq M^n$ is definable. Claim: If $t(v_1, \ldots, v_n)$ is a term, then $\{(\overline{x}, y) \in M^{n+1} \mid t^{\mathcal{M}}(\overline{x}) = y\} \in D_{n+1}$.

- If t is c: $\{(x_1, x_2) \in M^2 \mid x_1 = x_2\} \in D_2$ by 4. Let $c^{\mathcal{M}} = a \in M$. Then by 8, $\{x \in M \mid x = a\} \in D_1$. By n applications of 5, $\{(\overline{x}, a) \mid \overline{x} \in M^n\} \in D_{n+1}$.
- If t is v_i, $t^{\mathcal{M}}(x_1, \ldots, x_n) = x_i$. Let $\{(x_1, \ldots, x_{n+1}) \in M^{n+1} \mid x_i = x_{n+1}\} \in D_{n+1}$.
- Suppose t is $f(t_1, \ldots, t_m)$. By Induction let $G_i \in D_{n_i}$ be the graph of $t_i^{\mathcal{M}} : M^n \to M$, $n_i \leq n$. Let $G \in D_{m+1}$ be the graph of $f^{\mathcal{M}}$. Then the graph of $t^{\mathcal{M}}$ is $\{\overline{x}, y) \in M^{n+1} \mid \exists z_1 \ldots \exists z_n (\bigwedge_{i=1}^{m} (\overline{x}, z_i) \in G_i \wedge (\overline{z}, y) \in G\} \in D_{n+1}$.

Now we want to show that every $\emptyset$-definable set $X \subseteq M^n$ is in D_n. We do this by induction on atomic formulas.

- If ϕ is $t_1 = t_2$, then $\{\overline{x} \in M^n \mid t_1^{\mathcal{M}}(\overline{x}) = t_2^{\mathcal{M}}(\overline{x})\} = \{\overline{x} \in M^n \mid \exists y \exists z \; t_1^{\mathcal{M}}(\overline{x}) = y \wedge t_2^{\mathcal{M}}(\overline{x}) = z \wedge y = z\} \in D_n$ by 6, 4, and 7.
- If ϕ is $R(t_1, \ldots, t_n)$, then $\{\overline{x} \in M^n \mid \mathcal{M} \models \phi(\overline{x})\} = \{\overline{x} \in M^n \mid \exists z_1 \ldots \exists z_m \bigwedge_{i=1}^{m} t_i^{\mathcal{M}}(\overline{x}) = z_i \wedge \overline{z} \in R^{\mathcal{M}}\}$ by 7 and 3.

Therefore all $\emptyset$-definable sets by atomic formulas are in $\mathcal{D}$. Since $\mathcal{D}$ is closed under $\wedge, \vee, \neg, \exists$, all sets that are $\emptyset$-definable by formulas are also in $\mathcal{D}$. All definable sets are in $\mathcal{D}$, by 8.

□

How do we show that a set $S \subseteq M^n$ is NOT definable?

Definition 2.18. $\sigma : \mathcal{M} \to \mathcal{M}$ is an $\mathcal{L}$-automorphism of an $\mathcal{L}$-structure $\mathcal{M}$ if σ is an $\mathcal{L}$-isomorphism.

Proposition 2.19. Let $\mathcal{M}$ be an $\mathcal{L}$-structure. If $X \subseteq M^n$ is A-definable, then every $\mathcal{L}$-automorphism of σ of $\mathcal{M}$ that fixes A pointwise must fix X setwise, i.e., if $\sigma(a) = a$ for all $a \in A$, then $\sigma(X) = X$.

Proof. Let $\psi(\overline{v}, \overline{a})$ define X with $\overline{a} \in A$. Let σ be an $\mathcal{L}$-automorphism of $\mathcal{M}$ such that $\sigma(a_i) = a_i$ for $a_i \in \overline{a}$. Let $\overline{b} \in M^n$.

Earlier Theorem: If $j : \mathcal{M} \to \mathcal{N}$ is an $\mathcal{L}$-isomorphism, then $\mathcal{M} \models \phi(\overline{b}) \iff \mathcal{N} \models \phi(j(\overline{b}))$ for all $\mathcal{L}$-formulas ϕ.

Then $\mathcal{M} \models \psi(\overline{b}, \overline{a}) \iff \mathcal{M} \models \psi(\sigma(\overline{b}), \sigma(\overline{a}))$ (by the theorem) $\iff \mathcal{M} \models \psi(\sigma(\overline{b}), \overline{a})$ (by the way σ was defined). Then $\overline{b} \in X \iff \sigma(\overline{b}) \in X$. □

Corollary 2.20. $\mathbb{R}$ is not definable in $\mathbb{C}$.

Proof. Suppose $\mathbb{R}$ is A-definable for some finite $A \subseteq \mathbb{C}$. Pick $r \in \mathbb{R}$, $s \in \mathbb{C} \setminus \mathbb{R}$ transcendental over A. There exists an automorphism of $\mathbb{C}$ that fixes A and permutes r and s. Hence $\mathbb{R}$ is not definable over A.

□

2.3. **Compactness Theorem.** Question: Given an $\mathcal{L}$-theory T and $\mathcal{L}$-sentence ϕ, how can we show that $T \models \phi$?

One approach is to show that there is a formal proof of ϕ from T. Why does this suffice?

Theorem 2.21 (Soundness Theorem)**.** If there is a formal proof of ϕ from T, then $T \models \phi$.

2.4. **Formal proof.**

Definition 2.22. A proof of ϕ from T is a finite sequence of $\mathcal{L}$-formulas $\psi_1, \ldots, \psi_m$ such that $\psi_m = \phi$ and for each i, either $\psi_i \in T$ or ψ_i follows from $\psi_1, \ldots, \psi_{i-1}$ by some simple logical rules. We write $T \vdash \phi$ if there is a proof of ϕ from T.

Example 2.23. Some examples of simple logical rules:

- "From ϕ and ψ, conclude $\phi \wedge \psi$".
- "From $\phi \wedge \psi$, conclude ϕ".
- "From $\phi \longrightarrow \psi$ and ϕ, conclude ψ".

Some key properties of proofs:

(1) Proofs are finite.
(2) If $T \vdash \phi$, then $T \models \phi$.
(3) If T is a finite set of $\mathcal{L}$-sentences, then there is an algorithm that, given a sequence of $\mathcal{L}$-formulas Σ and an $\mathcal{L}$-sentence ϕ, will decide whether Σ is a proof ϕ from T.

Remark. This is not the same as saying there is an algorithm for determining whether $T \vdash \phi$.

Theorem 2.24. *(Gödel's Completness Theorem)* If $T \models \phi$, then $T \vdash \phi$, where T is an $\mathcal{L}$-theory and ϕ is an $\mathcal{L}$-sentence.

From the Completness Theorem and the Soundness Theorem, we can determine whether an $\mathcal{L}$-theory is satisfiable.

Definition 2.25. T is inconsistent if $T \vdash \phi \wedge \neg\phi$ for some sentence ϕ.

Corollary 2.26. T is consistent if and only if T is satisfiable.

Proof. Suppose T is inconsistent. Suppose also that $\mathcal{M} \models T$. So $T \vdash \phi \wedge \neg\phi$ for some $\mathcal{L}$-sentence ϕ, but by the Soundness Theorem, $T \models \phi \wedge \neg\phi$, so $\mathcal{M} \models \phi \wedge \neg\phi$. Hence $\mathcal{M} \models \phi$ and $\mathcal{M} \not\models \phi$. On the other hand, if T is not satisfiable, then every model of T is a model of $\phi \wedge \neg\phi$. So $T \models \phi \wedge \neg\phi$. By the Completness Theorem, $T \vdash \phi \wedge \neg\phi$, so T is inconsistent. $\square$

Theorem 2.27. *(Compactness Theorem)* An $\mathcal{L}$-theory T is satisfiable if and only if every finite subset of T is satisfiable (T is finitely satisfiable).

Remark. This is intimately related to topological compactness.

Proof. ($\Longrightarrow$) Let T be satisfiable. Then there exists a structure $\mathcal{M} \models T$. Let T_0 be a finite subset of T. Then $\mathcal{M} \models T_0$ as well.

($\Longleftarrow$) For the second direction, we prove the contrapositive. If T is not satisfiable, then by the corollary, T is inconsistent. Let ϕ be an $\mathcal{L}$-sentence such that $T \vdash \phi \wedge \neg\phi$. Let Σ be a proof of $\phi \wedge \neg\phi$ from T. Σ is finite, so it only contains finitely many $\mathcal{L}$-sentences from T. Call this set T_0. Then $T_0 \vdash \phi \wedge \neg\phi$. By the corollary, T_0 is not satisfiable. Hence T is not finitely satisfiable. $\square$

3. Direct proof of Compactness

Theorem 3.1 (Compactness Theorem)**.** Let T be an $\mathcal{L}$-theory. T is satisfiable if and only if T is finitely satisfiable.

3.1. **Henkin Contruction.** The idea is to have enough constants in our language to be able to talk about every element in the model we are building.

Definition 3.2 (Witness Property(WP))**.** An $\mathcal{L}$-theory T has the witness property if for every $\mathcal{L}$-formula $\phi(v)$ with one free formula, there is a constant symbol c such that $T \models [(\exists v\phi(v)) \to \phi(c)]$.

Definition 3.3 (Maximal)**.** An $\mathcal{L}$-theory T is maximal if for all $\mathcal{L}$-sentences ϕ, either $\phi \in T$ or $\neg\phi \in T$.

Lemma 3.4 (*)**.** Suppose T is a maximal and finitely satisfiable $\mathcal{L}$-theory. If $\triangle \subset T$ is finite and $\triangle \models \psi$, then $\psi \in T$.

Proof. Suppose T is an $\mathcal{L}$-theory that is maximal and finitely satisfiable. Let $\triangle \subset T$ be finite, and let $\triangle \models \psi$. Let $\psi \notin T$. By maximality of T, $\neg\psi \in T$. Then $\triangle \cup \neg\psi \subset T$ is finite but not satisfiable. □

Lemma 3.5. Suppose T is a maximal, finitely satisfiable $\mathcal{L}$-theory with the witness property. Then T has a model.

Let T be a maximal, finitely satisfiable $\mathcal{L}$-theory with the witness property. Let $\mathcal{C}$ be the collection of constants in $\mathcal{L}$. Let $c \backsim d$ hold for $c, d \in \mathcal{C} \iff c = d \in T \iff T \models c = d$.

Claim 3.6. $\backsim$ is an equivalence relation.

Proof. It is clear that $c \backsim c$, since $c = c \in T$.

Suppose $c \backsim d$. Then $c = d \in T$. Then $\{c = d\} \models d = c$. So by *, $d = c \in T$. Thus, $d \backsim c$.

Suppose $c \backsim d$ and $d \backsim e$. Then $c = d, d = e \in T$ and $\{c = d, d = e\} \models c = e$. Then by *, $c = e \in T$. □

The universe of $\mathcal{M}$ will be $\mathcal{C}/\backsim$. Let c^* denote the equivalence class of c, and let $c^{\mathcal{M}} = c^*$.

Claim 3.7. Let R be an n-ary relation symbol. Given $c_1, ..., c_n, d_1, ..., d_n$ such that $c_i = d_i$ for $i = 1, ..., n$, $R(\overline{c}) \in T \iff R(\overline{d}) \in T$.

Proof. Since $c_i = d_i$ for $i = 1, ..., n$, if $R(\overline{c}) \in T$, then $\{c_1 = d_1, ..., c_n = d_n, R(\overline{c})\} \models R(\overline{d})$. So by *, $R(\overline{d}) \in T$. Similarly if $R(\overline{d}) \in T$, $R(\overline{c}) \in T$. □

Denote $R^{\mathcal{M}} = \{(c_1^*, ..., c_n^*) : R(c_1, ...c_n) \in T\}$. By claim 2, $R^{\mathcal{M}}$ is well defined.

Let f be an n-ary function symbol and $c_1, ..., c_n$ be constants. Then $\emptyset \models \exists v(f(c_1, ..., c_n) = v)$. By WP, $T \models f(c_1, ..., c_n) = c_{n+1}$ for some $c_{n+1} \in \mathcal{C}$, which implies that $f(c_1, ..., c_n) = c_{n+1} \in T$. Let $d_1, ..., d_{n+1}$ satisfy $c_i \backsim d_i$ for $i = 1, ..., n, n+1$. Then $f(c_1, ..., c_n) = c_{n+1} \in T \iff f(d_1, ..., d_n) = d_{n+1} \in T$ (similiar proof to that of claim 2).

If for constants $c_1, ..., c_n, e_1, ..., e_n$, we have $c_i \backsim e_i$ for $i = 1, ...n$ and $f(e_1, ..., e_n) = e_{n+1} \in T$, then $\{c_1 = e_1, ..., c_n = e_n, f(\overline{c}) = c_{n+1}, f(\overline{e}) = e_{n+1}\} \models e_{n+1} = c_{n+1}$. Set $f^{\mathcal{M}}(c_1^*, ..., c_n^*) = d^*$ for some $d \in \mathcal{C}$, which occurs if and only if $f(c_1, ..., c_n) = d \in T$. So $f^{\mathcal{M}}$ is a well-defined function.

To show $\mathcal{M} \models T$, we show that terms behave correctly.

Claim 3.8. Suppose t is a term with free variables $v_1, ..., v_n$. If $c_1, ..., c_n, d \in \mathcal{C}$,

$$t(c_1, ..., c_n) = d \in T \iff t^{\mathcal{M}}(c_1^*, ..., c_n^*) = d^*$$

Proof. ($\Longrightarrow$) Proof by induction on terms. If t is a constant, c, we have $c = d \in T$. Then $t^{\mathcal{M}}(c_1^*, ..., c_n^*) = c^{\mathcal{M}} = c^* = d^*$. If t is v_i, then $c_i = d_i \in T$. Hence $t^{\mathcal{M}}(c_1^*, ..., c_n^*) = c_i^* = d^*$.

Suppose the claim holds for terms $t_1, ..., t_m$ and t is $f(t_1, ..., t_m)$. Then by WP of T and *, there are constants $d_1, ..., d_m$ such that $t_i(c_1, ..., c_n) = d_i \in T$ for $i = 1, ..., m$ and $f(d_1, ..., d_m) = d \in T$. By the inductive hypothesis, $t_i^{\mathcal{M}}(\overline{c}^*) = d_i^*$ for $i = 1, ..., m$. Thus, $f^*(d_1^*, ..., d_m^*) = d^* \Longrightarrow t^{\mathcal{M}}(c_1^*, ..., c_n^*) = d^*$.

($\Longleftarrow$) Suppose $t^{\mathcal{M}}(c_1^*, ..., c_n^*) = d^*$. By WP and *, there is an $e \in \mathcal{C}$ such that $t(c_1, .., c_n) = e \in T$. By the forward direction, $t^{\mathcal{M}}(c_1^*, ..., c_n^*) = e^*$. So $e^* = d^*$. Hence $e = d \in T$. Thus, $t(c_1, ..., c_n) = d \in T$ by *. □

Claim 3.9. For all $\mathcal{L}$-formulas $\phi(v_1, ..., v_n)$ and constants $c_1, ..., c_n$,

$$\mathcal{M} \models \phi(\overline{c}) \iff \phi(\overline{c}) \in T.$$

Proof. By induction on formulas.

If ϕ is $t_1 = t_2$, by WP and *, there are constants d_1 and d_2 such that $t_1(\overline{c}) = d_1 \in T$ and $t_2(\overline{c}) = d_2 \in T$. By claim 3, $t_i(\overline{c}^*) = d_i^*$ for $i = 1, 2$. So

$$\begin{aligned}
\mathcal{M} \models \phi(c^*) &\iff t_1^{\mathcal{M}}(\overline{c}^*) = t_2^{\mathcal{M}}(\overline{c}^*) \\
&\iff d_1^* = d_2^* \\
&\iff d_1 = d_2 \in T \\
&\iff t_1(\overline{c}) = t_2(\overline{c}) \in T.
\end{aligned}$$

Now if ϕ is $R(t_1, ..., t_n)$, by WP and *, $\exists d_1, ..., d_n$ such that $t_i(\overline{c}) = d_i \in T$ for $i = 1, ..., n$. By prev. claim, $t_i^*(\overline{c}^*) = d_i^*$ for $i = 1, ..., n$. So

$$\begin{aligned}
\mathcal{M} \models \phi(\overline{c}^*) &\iff \mathcal{M} \models R(t_1(\overline{c}^*), ..., t_n(\overline{c}^*)) \\
&\iff \mathcal{M} \models R(d_1^*, ..., d_n^*) \\
&\iff R(\overline{d}) \in T \\
&\iff \phi(\overline{c}) \in T.
\end{aligned}$$

Suppose the result holds for ϕ and ψ.

Negation: If $\mathcal{M} \models \neg\phi(\overline{c}^*)$, then $\mathcal{M} \not\models \phi(\overline{c}^*)$. By IH, $\phi(\overline{c}) \notin T$. But T is maximal so $\neg\phi(\overline{c}) \in T$. Then since T is finitely satisfiable,$\phi(\overline{c}^*) \notin T$. So

$$\begin{aligned}
\phi(\overline{c}^*) \notin T &\iff \mathcal{M} \not\models \phi(\overline{c}^*) \\
&\iff \mathcal{M} \models \neg\phi(\overline{c}^*).
\end{aligned}$$

And:

$$\begin{aligned}
\mathcal{M} \models (\phi \wedge \psi)(\overline{c}^*) &\iff \mathcal{M} \models \phi(\overline{c}^*) \textit{ and } \mathcal{M} \models \psi(\overline{c}^*) \\
&\iff \phi(\overline{c}) \in T \textit{ and } \phi(\overline{c}) \in T \\
&\iff (\phi \wedge \psi)(\overline{c}) \in T \textit{ (by } *).
\end{aligned}$$

Existential: If ϕ is $\exists v\phi(v, \overline{d})$,

$$\begin{aligned}\mathcal{M} \models \phi \; for some \; c^* \in M &\implies \mathcal{M} \models \phi(c^*, \overline{d}^*) \\ &\iff \phi(c, \overline{d}) \in T \\ &\implies \exists v\phi(v, \overline{d}^*) \in T \; (by *).\end{aligned}$$

□

Lemma 3.10. Let T be a finitely satisfiable $\mathcal{L}$-theory. Then there is a language $\mathcal{L}^* \supseteq \mathcal{L}$ and a finitely satisfiable $\mathcal{L}^*$-theory $T^* \supseteq T$ with the WP. Moreover, we can choose $\mathcal{L}^*$ so that $|\mathcal{L}^*| = |\mathcal{L}| + \aleph_0$.

Proof. Step 1: We show there exists a language $\mathcal{L}_1 \supseteq \mathcal{L}$ and a finitely satisfiable $\mathcal{L}_1$-theory $T_1 \supseteq T$ such that for any formula $\phi(v)$ with one free variable (w.o.f.v.), there exists a constant symbol c_ϕ such that $T_1 \models (\exists v\phi(v)) \to \phi(c_\phi)$.

For each $\phi(v)$ w.o.f.v., let c_ϕ be the new constant symbol and let $\mathcal{L}_1 = \mathcal{L} \cup \{c_\phi : \phi(v)$ and $\mathcal{L}$-formula w.o.f.v$\}$. For each such ϕ, let θ_ϕ be the $\mathcal{L}$-sentence $(\exists v\phi(v)) \to \phi(c_\phi)$. Let $T_1 = T \cup \{\theta_\phi : \phi(v)$ and $\mathcal{L}$-formula w.o.f.v$\}$.

Claim: T_1 is finitely satisfiable.

Proof. Let $\delta \subset T$ be finite. Then $\delta = \delta_o \cup \{\theta_{\phi_1}, ..., \theta_{\phi_n}\}$ where $\delta_0 \subset T$ is finite. T is finitely satisfiable, so $\mathcal{M} \models \delta_0$ for some $\mathcal{L}$-structure. We convert $\mathcal{M}$ into an $\mathcal{L} \cup \{c_{\phi_1}, ..., c_{\phi_n}\}$-structure $\mathcal{M}'$. If $\mathcal{M} \models \exists v\phi_i(v)$, then $\mathcal{M} \models \phi_i(a_i)$ for some $a_i \in M$, let $c_{\phi_i}^{\mathcal{M}'} = a_i$. If $\mathcal{M} \not\models \exists v\phi_i(v)$, we can let $c_{\phi_i}^{\mathcal{M}'}$ be anything in M. Then $\mathcal{M}' \models \theta_{\phi_i}$ for $i \leq n$, and $\mathcal{M}' \models \theta_0$ since $\mathcal{L}$-symbols have the same interpretations as in $\mathcal{M}$. □

We repeat this step to build a sequence of languages $\mathcal{L} \subseteq \mathcal{L}_1 \subseteq \mathcal{L}_2 \subseteq ...$ and a set of finitely satisfiable $\mathcal{L}_i$-theories $T \subseteq T_1 \subseteq T_2...$ such that if $\phi(v)$ is an $\mathcal{L}$-formula w.o.f.v, there exists a constant $c \in \mathcal{L}_{i+1}$ such that $T_{i+1} \models (\exists v\phi(v)) \to \phi(c)$. We set $\mathcal{L}^* = \bigcup \mathcal{L}_i$ and $T^* = \bigcup T_i$. □

Theorem 3.11. Let T be an $\mathcal{L}$-theory. Then T is satisfiable if and only if it is finitely satisfiable.

If T is a maximal, finitely satisfiable $\mathcal{L}$-theory with the witness property, then T has a model. T is *maximal* if for every $\mathcal{L}$-sentence ϕ, either $\phi \in T$ or $\neg\phi \in T$. T has the *witness property* if for every $\mathcal{L}$-formula $\phi(v)$ with one free variable, there is some constant c such that

$$T \models (\exists v\phi(v)) \to \phi(c).$$

We are now in the process of showing that a finitely satisfiable $\mathcal{L}$-theory T can be extended to a maximal theory T^* that is finitely satisfiable and has the witness property.

Theorem 3.12. Let T be a finitely satisfiable $\mathcal{L}$-theory. Then there is

- a language $\mathcal{L}^* \supseteq \mathcal{L}$ and
- a finitely satisfiable $\mathcal{L}^*$-theory $T^* \supseteq T$

such that any theory extending T^* has the witness property. Moreover, we can choose $\mathcal{L}^*$ in such a way that $|\mathcal{L}^*| = |\mathcal{L}| + \aleph_0$.

To prove this theorem, we defined a sequence of languages

$$\mathcal{L} \subseteq \mathcal{L}_1 \subseteq \mathcal{L}_2 \subseteq \ldots \subseteq \mathcal{L}_i \subseteq \ldots$$

and a sequence of theories

$$T \subseteq T_1 \subseteq T_2 \subseteq \ldots \subseteq T_i \subseteq \ldots$$

where each T_i is an $\mathcal{L}_i$-theory.

The key to the definition of each pair $(\mathcal{L}_i, T_i)$ is that we add constants to $\mathcal{L}_{i+1}$ and sentences of the form $(\exists v \phi(v)) \to \phi(c_\phi)$ to T_{i+1}, where $\exists v \phi(v)$ is an $\mathcal{L}_i$-sentence and c_ϕ is a new constant in $\mathcal{L}_{i+1}$.

We showed that each T_i is finitely satisfiable, which follows inductively from the fact that the original theory T is finitely satisfiable.

Now, let $\mathcal{L}^* = \bigcup_{i \in \omega} \mathcal{L}_i$ and $T^* = \bigcup_{i \in \omega} T_i$.

- Why is T^* finitely satisfiable?
  - If $\Delta \subseteq T^*$ is finite, then $\Delta \subseteq T_i$ for some i.
- Let $T' \supseteq T^*$ be any $\mathcal{L}^*$-theory. Why does T' have the witness property?
  - If $\exists v \phi(v)$ is an $\mathcal{L}^*$-formula, it is an $\mathcal{L}_i$-formula for some i.

Lastly, we verify that $|\mathcal{L}^*| = |\mathcal{L}| + \aleph_0$.

Note that if $\mathcal{L}$ is finite, there are countably many $\mathcal{L}$-sentences of the form $\exists v \phi(v)$, hence $|\mathcal{L}_1| = \aleph_0$.

If $|\mathcal{L}| = \kappa$ for some infinite κ, then there are κ many $\mathcal{L}$-sentences of the form $\exists v \phi(v)$, hence $|\mathcal{L}_1| = \kappa$ as well.

Similarly, if $|\mathcal{L}_i| = \kappa$, then $|\mathcal{L}_{i+1}| = \kappa$.

It follows by induction that $|\mathcal{L}_i| = |\mathcal{L}| + \aleph_0$ for every i.

Since $\mathcal{L}^*$ is a countable union of sets of size $|\mathcal{L}| + \aleph_0$, it follows that $|\mathcal{L}^*|$ also has size $|\mathcal{L}|$.

To extend T to a maximal theory, first we prove the following.

Lemma 3.13. Suppose that T is a finitely satisfiable $\mathcal{L}$-theory and ϕ is an $\mathcal{L}$-sentence. Then either $T \cup \{\phi\}$ or $T \cup \{\neg\phi\}$ is finitely satisfiable.

Proof. Suppose that $T \cup \{\phi\}$ is not finitely satisfiable.

Then there is some finite $\Delta \subseteq T$ such that $\Delta \models \neg\phi$.

We claim that $T \cup \{\neg\phi\}$ is finitely satisfiable.

Let $\Sigma \subseteq T$ be finite. Then $\Sigma \cup \Delta$ is satisfiable.

Since $\Delta \models \phi$, it follows that $\Sigma \cup \{\neg\phi\}$ is satisfiable.

Hence the claim follows. □

We will also make use of Zorn's Lemma to extend T to a maximal theory.

Definition 3.14. Let $(X, <)$ be a partial order.

- $C \subseteq X$ is a *chain* in X if C is linearly ordered by $<$; that is, for every $x, y \in C$, either $x \mathcal{L} eq y$ or $y \mathcal{L} eq x$.
- $x \in X$ is an *upper bound* of C if $c \mathcal{L} eq x$ for all $c \in C$.
- $x \in X$ is maximal if there is no $z \in X$ with $x < z$.

Theorem 3.15 (Zorn's Lemma)**.** If $(X, <)$ is a partial order such that every chain in X has an upper bound, then there is a maximal element of X.

Theorem 3.16. Let T be a finitely satisfiable $\mathcal{L}$-theory. Then there is a maximal, finitely satisfiable $\mathcal{L}$-theory $T' \supseteq T$.

Proof. Let I be the set of all finitely satisfiable $\mathcal{L}$-theories containing T, which we can partially order by $\subseteq$.

If $C \subseteq I$ is a chain, let $T_C = \bigcup\{\Sigma : \Sigma \in C\}$.

Let Δ be a finite subset of T_C. Then there is some $\Sigma \in C$ such that $\Delta \subseteq \Sigma$. Thus T_C is finitely satisfiable.

Since $T_C \supseteq \Sigma$ for all $\Sigma \in C$, T_C is an upper bound of C. Hence every chain in I has an upper bound.

By Zorn's Lemma, there is some $T' \in I$ that is maximal with respect to the ordering $\subseteq$.

Let ϕ be any $\mathcal{L}$-sentence. By the previous lemma, either $T' \cup \{\phi\}$ or $T' \cup \{\neg\phi\}$ is finitely satisfiable.

Since T' is maximal in the ordering $\subseteq$ on I, it follows that $\phi \in T$ or $\neg\phi \in T$.

Thus T is maximal.

□

Let ϕ be any $\mathcal{L}$-sentence. By the previous lemma, either $T' \cup \{\phi\}$ or $T' \cup \{\neg\phi\}$ is finitely satisfiable.

Since T' is maximal in the ordering $\subseteq$ on I, it follows that $\phi \in T$ or $\neg\phi \in T$.

Thus T is maximal.

Theorem 3.17. Let T be a finitely satisfiable $\mathcal{L}$-theory and κ an infinite cardinal with $\kappa \geq |\mathcal{L}|$. Then there is a model of T of cardinality at most κ.

Proof. There are four steps.

- Step 1: Extend $\mathcal{L}$ to a language $\mathcal{L}^*$ and T to a finitely satisfiable $\mathcal{L}^*$-theory T^* such that
  - any $\mathcal{L}^*$-theory extending T^* has the witness property and
  - $|\mathcal{L}^*| \leq \kappa$.
- Step 2: Let $T' \supseteq T^*$ be a finitely satisfiable, maximal $\mathcal{L}^*$-theory.
- Step 3: By the Henkin construction, there is a model $\mathcal{M} \models T'$ with $|M| \leq \kappa$ (since $|\mathcal{L}^*| \leq \kappa$).
- Step 4: Since $T' \supseteq T^* \supseteq T$, it follows that $\mathcal{M} \models T$.

□

Proposition 3.18. Let $\mathcal{L} = \{+, \cdot, <, 0, 1\}$ and $\mathrm{Th}(\mathbb{N})$ be the full $\mathcal{L}$-theory of the natural numbers. Then there is some $\mathcal{M} \models \mathrm{Th}(\mathbb{N})$ and $a \in M$ such that a is larger than every natural number.

Proof. Let $\mathcal{L}^* = \mathcal{L} \cup \{c\}$ where c is a new constant symbol and let

$$T = \mathrm{Th}(\mathbb{N}) \cup \{\underbrace{1 + 1 + \ldots + 1}_{n\text{times}} < c : n \in \mathbb{N}\}.$$

For any finite $\Delta \subseteq T$, by interpreting c as a sufficiently large $n \in \mathbb{N}$, we have $\mathbb{N} \models \Delta$.

T is thus finitely satisfiable, and so by the compactness theorem, it has a model $\mathcal{M}$.

If $c^{\mathcal{M}} = a \in M$, then a is greater than every natural number (represented in M by the objects picked out by the interpretation of $0, 1, 1+1, 1+1+1, \ldots$ in $\mathcal{M}$).

□

Proposition 3.19. Let $\mathcal{L}$ be a language containing $\{\cdot, e\}$, the language of groups, let T be an $\mathcal{L}$-theory extending the theory of groups, and let $\phi(v)$ be an $\mathcal{L}$-formula.

Suppose that for every n there is $G_n \models T$ and $g_n \in G_n$ with finite order greater than n such that $G_n \models \phi(g_n)$.

Then, there is $G \models T$ and $g \in G$ such that $G \models \phi(g)$ and g has infinite order.

In particular, there is no formula that defines the torsion points in all models of T.

Let $\mathcal{L}^* = \mathcal{L} \cup \{c\}$ where c is a new constant symbol and let

$$T^* = T \cup \{\phi(c)\} \cup \{\underbrace{c \cdot c \cdot \ldots \cdot c}_{n\text{times}} \neq e : n \in \mathbb{N}\}.$$

For any finite $\Delta \subseteq T^*$, there is some $m \in \mathbb{N}$ such that

$$\Delta \subseteq T \cup \{\phi(c)\} \cup \{\underbrace{c \cdot c \cdot \ldots \cdot c}_{n\text{times}} \neq e : n = 1, \ldots, m\}.$$

By hypothesis, there is some G_m with $g_m \in G_m$ of finite order greater than m such that $G_m \models T \cup \{\phi(g_m)\}$, and hence Δ is satisfiable.

Thus T^* is finitely satisfiable, and so by the compactness theorem, there is some $G \models T^*$.

If $g \in G$ is the interpretation of c in G, then $\phi(g)$ holds. However, g has infinite order and hence is not a torsion point.

Lemma 3.20. If $T \models \phi$, then $\Delta \models \phi$ for some finite $\Delta \subseteq T$.

Proof. Suppose for every finite $\Delta \subseteq T$, $\Delta \not\models \phi$.

Then for any finite $\Delta \subseteq T$, $\Delta \cup \{\neg\phi\}$ is satisfiable.

Hence $T \cup \{\neg\phi\}$ is finitely satisfiable, and so by the compactness theorem, $T \not\models \phi$. □

Theorem 3.21 (Completeness Theorem)**.** Let T be an $\mathcal{L}$-theory and ϕ an $\mathcal{L}$-sentence. If $T \models \phi$, then $T \vdash \phi$.

This is equivalent to:

Theorem 3.22 (Completeness Theorem 2.0)**.** If T is consistent, then T is satisfiable.

Lemma 3.23. (i) If T is inconsistent, then $T \vdash \phi$ for every $\mathcal{L}$-sentence ϕ.

(ii) $T \vdash \phi$ if and only if $T \cup \{\neg\phi\}$ is inconsistent.

Suppose $T \models \phi$ implies $T \vdash \phi$ (Completeness 1.0).

If T has no model, then $T \models \phi \wedge \neg\phi$ for some $\mathcal{L}$-sentence Φ.

By Completeness 1.0, it follows that $T \vdash \phi \wedge \neg\phi$, hence T is inconsistent.

Suppose if T is consistent, then T is satisfiable (Completeness 2.0).

If $T \models \phi$, we have two cases to consider:

- Case 1: T is inconsistent.

  Then by (i) of the previous lemma, $T \vdash \phi$.

- Case 2: T is consistent.

  Suppose in this case that $T \cup \{\neg\phi\}$ is consistent.
  Then by Completeness 2.0, $T \cup \{\neg\phi\}$ is satisfiable.
  However, this contradicts our assumption that $T \models \phi$.
  It follows that $T \cup \{\neg\phi\}$ is inconsistent, so by (ii) of the previous lemma, $T \vdash \phi$.

4. Complete Theories

Definition 4.1 (Complete)**.** An $\mathcal{L}$-Theory T is complete if for every sentence ϕ, $T \models \phi$ or $T \models \neg\phi$.

Definition 4.2 (Theory of a structure)**.** For an $\mathcal{L}$-structure $\mathcal{M}$, $Th(\mathcal{M}) = \{\phi : \phi \text{ an } \mathcal{L}\text{-sentence } \& \ \mathcal{M} \models \phi\}$.

4.1. **Remark:** In general $Th(\mathcal{M})$ is quite hard to work with. We want a simpler complete theory T such that $\mathcal{M} \models T$ and T is complete. In this case $\mathcal{M} \models T \iff T \models \phi$.

Proposition 4.3. Let T be an $\mathcal{L}$-theory with infinite models. If κ is an infinite cardinal, then T has a model of cardinality κ.

Proof.

Let $\mathcal{L}^* = \mathcal{L} \cup \{c_\alpha : \alpha < \kappa\}$ where c_α is a new constant symbol. Let T^* be $T \cup \{c_\alpha \neq c_\beta : \alpha \neq \beta \ \& \ \alpha < \kappa \ \& \ \beta < \kappa\}$ an $\mathcal{L}^*$-theory. T^* is finitely satisfiable. To see this, let $\Delta \subset T \cup \{c_\alpha \neq c_\beta : \alpha \neq \beta \ \& \ \alpha < I \ \& \ \beta < I\}$ where $I \leq \kappa$ and I is finite. Let $\mathcal{M} \models T$ be infinite. We can interpret the symbols $\{c_\alpha : \alpha \in I\}$ as $|I|$ distinct members of $\mathcal{M}$.

Hence $\mathcal{M} \models \Delta$. By compactness T^* is satisfiable. Let $\mathcal{M}' \models T^*$, then $\mathcal{M}' \models T$. Therefore $\mathcal{M}' \geq \kappa$. For the opposite direction to show equality, one can use a Henkin construction.

□

Definition 4.4 (Categoricity)**.** Let κ be an infinite cardinal and let T be an $\mathcal{L}$-theory with models of size κ. T is κ-categorical if any two models of cardinality κ are model ismorphic.

4.2. **Example:** Let $\mathcal{L} = \emptyset$. The theory of an infinite set is κ-categorical for all infinite κ. The isomorphism is just the bijection between the structures as there is no special language to preserve.

4.3. **Example:** Let $\mathcal{L} = \{E\}$ be the language for the theory of an equivalence relation with two infinite equivalence classes. T is $\aleph_0$-categorical, but is not κ-categorical for uncountable κ. To see how this could fail at an uncountable cardinal, consider the case where one model has both classes uncountable, while the other has one uncountable class while the other is countably infinite.

4.4. **Example:** If $\mathcal{M}_0$ and $\mathcal{M}_1$ are two models of a theory T that have cardinality κ but are not model isomorphic with each other, then T is not κ categorical.

4.5. **Example:** Algebraically closed fields of characteristic p, where p is prime or zero, are κ-categorical for all uncountable κ, but is not $\aleph_0$ categorical. The idea of how to prove this is to note that two algebraically close fields with the same characteristic and cardinality are isomorphic so long as they have the same transcendence degree.

4.6. **Remark:** $\{e\}$ and $\{\pi\}$ are algebraically independent of $\mathbb{Q}$. However, whether $\{e + \pi\}$ is algebraically independent of $\mathbb{Q}$ is an open question.

Proposition 4.5 (Vaught's Test)**.** Let T be a satisfiable $\mathcal{L}$-theory with no finite models which is κ-categorical for some infinite $\kappa \geq \mathcal{L}$. Then T is complete.

Proof.

Suppose T is not complete. Then there is some $\mathcal{L}$-sentence ϕ such that $T \not\models \phi$ and $T \not\models \neg\phi$. Let T_0 be $T \cup \{\phi\}$ and T_1 be $T \cup \{\neg\phi\}$. By an earlier result, there are models $\mathcal{M}_0$ and $\mathcal{M}_1$ both of cardinality κ such that they model the respective T_i. However, these two models are not isomorphic as they disagree on ϕ. This contradicts the κ-categoricity of the theory, so it must be the case that T is complete. □

4.7. **Corollary:** Let $\mathcal{L}$ be the language of rings. The following are equivalent for an $\mathcal{L}$-sentence ϕ.

(1) $\mathbb{C} \models \phi$
(2) $F \models \phi$ for every ACF_0.
(3) $F \models \phi$ for some ACF_0.
(4) There are infinitely many primes p such that $F \models \phi$ for some ACF_p.
(5) There exists a number m such that for all primes p above m, $F \models \phi$ for all ACF_p.

5. Up and Down

5.1. **Recall:** If $j : \mathcal{M} \to \mathcal{N}$ is an $\mathcal{L}$-embedding, then for all atomic formulas $\phi(\bar{v})$ and $\bar{a} \in M$,

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

This does not necessarily hold for all formulas! For intance, $j : \mathbb{Z} \to \mathbb{Z} \times \mathbb{Z}$ may fail with some existential formulas. We want to beef this up so we can have something similar for all formulas.

Definition 5.1 (Elementary Embedding)**.** An $\mathcal{L}$-embedding $j : \mathcal{M} \to \mathcal{N}$ is elementary if for **all** $\mathcal{L}$-formulas $\phi(\bar{v})$ where $\bar{a} \in \mathcal{M}$,

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

If $\mathcal{M} \subset \mathcal{N}$ is a substructure, then if $j : \mathcal{M} \to \mathcal{N}$ is elementary, we say that $\mathcal{M}$ is an elementary substructure of $\mathcal{N}$, or $\mathcal{N}$ is an elementary extension of $\mathcal{M}$, and we write $\mathcal{M} \prec \mathcal{N}$.

5.2. **Example:** Consider the rings $(\mathbb{Z}, +, \bullet, 0, 1)$,$(\mathbb{Q}, +, \bullet, 0, 1)$, and $(\mathbb{R}, +, \bullet, 0, 1)$.

$$(\mathbb{Z}, +, \bullet, 0, 1) \not\prec (\mathbb{Q}, +, \bullet, 0, 1) \not\prec (\mathbb{R}, +, \bullet, 0, 1)$$

This is because the relation ϕ defined by $\phi(x) = \exists x\ x + x = 1$ is modelled in $\mathbb{Q}$, but not in $\mathbb{Z}$. Similarly, the relation ψ defined by $\psi(x) = \exists x\ x \bullet x = 2$ holds in $\mathbb{R}$ but not in $\mathbb{Q}$.

Definition 5.2. ($\mathcal{L}_M$) Suppose $\mathcal{M}$ is an $\mathcal{L}$-structure. Then $\mathcal{L}_M$ is the language obtained from $\mathcal{L}$ by adding new constants of the form c_m for each $m \in M$.

Definition 5.3. ($Diag(\mathcal{M})$) The atomic diagram of $\mathcal{M}$, $Diag(\mathcal{M})$, is the collection $\{\phi(c_1, ...c_n :$ ϕ is either an atomic $\mathcal{L}$-formula and $\mathcal{M} \models \phi(c_{m_1}, ..., c_{m_n})$ or ϕ is the negation of an atomic formula and is satisfied$\}$

Definition 5.4. ($Diag_{el}(\mathcal{M})$) The elementary diagram of $\mathcal{M}$, $Diag_{el}(\mathcal{M})$, is defined to be the collection $\{\phi(c_{m_1}, ..., c_{m_n}) : \phi$ is an $\mathcal{L}$-formula and $\mathcal{M} \models \phi(m_1, ..., m_n)\}$

Proposition 5.5. (1) Suppose $\mathcal{N}$ is an $\mathcal{L}_M$ structure and $\mathcal{N} \models Diag(\mathcal{M})$. Then, viewing $\mathcal{N}$ as an $\mathcal{L}$-structure, there is an $\mathcal{L}$-embedding of $\mathcal{M}$ into $\mathcal{N}$.
(2) If $\mathcal{N} \models Diag_{el}(\mathcal{M})$ then this $\mathcal{L}$-embedding is elementary.

Proof. For part (1), let $j : \mathcal{M} \to \mathcal{N}$, $m \in M$, and set $j(m) = c_m^{\mathcal{N}}$. We need j to be injective and preserve the interpretations of constant, function and relation symbols. If $m_1, m_2 \in M$ and $m_1 \neq m_2$, then $c_{m_1} \neq c_{m_2} \in Diag(\mathcal{M})$. Note that $j(m_1) = c_{m_1}^{\mathcal{N}} \neq c_{m_2}^{\mathcal{N}} = j(m_2)$, so j is injective.

If f is a function symbol in $\mathcal{L}$, and $f^{\mathcal{M}}(m_1, ..., m_n) = m_{n+1}$ then $f(c_{m_1}, ..., c_{m_n}) = c_{m_{n+1}} \in Diag(\mathcal{M})$. Thus $f^{\mathcal{N}}(j(m_1), ..., j(m_n)) = j(m_{n+1})$, so we are done. We can do a very similar argument for the relation symbols.

For the constant symbols, $c^{\mathcal{M}} = m$ for some m. But $j(m) = c_m$ so we are done. Therefore j is an $\mathcal{L}$-embedding.

Now for part (2). If $\mathcal{N} \models Diag_{el}(\mathcal{M})$, $\mathcal{M} \models \phi(m_1, ..., m_n) \iff \phi(c_{m_1}, ..., c_{m_n}) \in Diag_{el}(\mathcal{M}) \iff \mathcal{N} \models \phi(j(m_1), ..., j(m_n))$ □

Proposition 5.6 (Upward Lowenheim-Skolem)**.** Let $\mathcal{M}$ be an infinite $\mathcal{L}$-structure, κ be an infinite cardinal satisfying $\kappa > |\mathcal{M}| + |\mathcal{L}|$. Then there is an $\mathcal{L}$-structure $\mathcal{N}$ of cardinality κ and $j : \mathcal{M} \to \mathcal{N}$ that is an elementary embedding.

Proof. $\mathcal{M} \models Diag_{el}(\mathcal{M})$, where $Diag_{el}(\mathcal{M})$ is an $\mathcal{L}_M$ theory with an infinite model. Let $\kappa \geq |\mathcal{M}| + |\mathcal{L}| = |\mathcal{L}_M|$. There is $\mathcal{N} \models Diag_{el}(\mathcal{M})$ of cardinality κ by a previous theorem. By lemma (ii), there exists an elementary embedding j from $\mathcal{M}$ into $\mathcal{N}$. □

In order to set the stage for the downward part of the Lowenheim-Skolem theorem, we will prove a proposition known as the Tarski-Vaught test.

Proposition 5.7 (Tarski-Vaught)**.** Suppose $\mathcal{M} \subset \mathcal{N}$. Then $\mathcal{M} \prec \mathcal{N} \iff$ for all formulas $\phi(v, \bar{w})$ and $\bar{a} \in M$, $\mathcal{N} \models \phi(b, \bar{a})$ for some $b \in N$ then $\mathcal{N} \models \phi(c, \bar{a})$ for some $c \in M$.

Proof. First suppose that $\mathcal{N}$ is an elementary extension of $\mathcal{M}$. Then if $\mathcal{N} \models \phi(b, \bar{a})$ for some $b \in N$, then $\mathcal{N} \models \exists x\ \phi(x, \bar{a})$, so $\mathcal{M} \models \phi(c, \bar{a})$ for some $c \in \mathcal{M}$.

For the opposite direction, we need to show that all $\mathcal{L}$-formula are preserved through the embedding. Since $\mathcal{M} \subset N$, this is true for atomic formulas. Proceeding by an induction on formulas, we see that this is true for formulas of the form $\phi \wedge \psi$ and $\neg\phi$ if the result holds for ϕ and for ψ. We will show the existential quantifier step.

Suppose the claim holds for $\psi(v, \bar{w})$. Let $\bar{a} \in \mathcal{M}$. If $\mathcal{M} \models \exists v\ \psi(v, \bar{a})$ for some $b \in M$. By the induction hypothesis, $\mathcal{N} \models \psi(b, \bar{a})$, so $\mathcal{N} \models \exists v\ \psi(v, \bar{a})$. If $\mathcal{N} \models \exists v\ \psi(v, \bar{a})$, then $\mathcal{N} \models \psi(b, \bar{a})$ for some $b \in N$.

If $\mathcal{N} \models \exists v\ \psi(v, \bar{a})$, then $\mathcal{N} \models \psi(b, \bar{a})$ for some $b \in N$.

Thus $\mathcal{M} \models \psi(c, \bar{a})$ for some $c \in M$ by assumption, so $\mathcal{M} \models \exists v\ \psi(v, \bar{a})$.

□

6. Downward Lowenheim-Skolem Theorem

6.1. **Corollary.** Let ϕ be a sentence in the language of rings. The following are equivalent:

(1) $\mathbb{C} \models \phi$
(2) $F \models \phi$ for all algebraically closed fields of characteristic zero.
(3) $F \models \phi$ for some algebraically closed fields of characteristic zero.
(4) There exists infinitely many primes p such that $F \models \phi$ for some algebraically closed field of characteristic p.
(5) For all primes p, $F \models \phi$ for all algebraically closed fields of characteristic p.

Proof. We have $\mathbb{C} \models \phi \iff ACF_0 \models \phi$ by the completeness of ACF_0. It follows that (1) $\iff$ (2) $\iff$ (3). It is also evident that (5) $\to$ (4).

To see (2) $\to$ (5), suppose $ACF_0 \models \phi$. Then by the consequences of the compactness theorem, $\exists \Delta \subset ACF_0$ such that $\Delta \models \phi$. Thuse for large enough p, $ACF_p \models \phi$. This yields (5).

Now for (4) $\to$ (2), by contrapositive. Suppose that $ACF_0 \not\models \phi$. By the completeness of the theory we have that $ACF_0 \models \neg\phi$. By the same argument for (2) $\to$ (5), $ACF_p \models \neg\phi$ for all sufficiently large p. This it is not the case that there exists infinitely many primes p such that $F \models \phi$ for some F of characteristic p.

□

Proposition 6.1 (Tarski-Vaught Test)**.** Suppose $\mathcal{M} \subset \mathcal{N}$. Then $\mathcal{M} \prec \mathcal{N} \iff$ for all formulas $\phi(v, \bar{w})$ and $\bar{a} \in M$, $\mathcal{N} \models \phi(b, \bar{a})$ for some $b \in N$ then $\mathcal{N} \models \phi(c, \bar{a})$ for some $c \in M$.

To show $\mathcal{M} \prec \mathcal{N}$ we proved $\mathcal{M} \models \psi(\bar{a}) \iff \mathcal{N} \models \psi(\bar{a})$ for all $\mathcal{L}$-formulas ϕ and $\bar{a} \in M^n$.
$\mathcal{M} \models \exists x\ \psi(x, \bar{a}) \iff \mathcal{N} \models \exists x\ \psi(x, \bar{a})$, so for $b \in N$ and some $c \in M$:
$\mathcal{N} \models \psi(b, \bar{a}) \to \mathcal{N} \models \psi(c, \bar{a}) \to \mathcal{M} \models \psi(c, \bar{a}) \to \mathcal{M} \models \exists x\ \psi(x, \bar{a})$
We could try to do this with a Henkin construction but there are problems with this approach, so we will do something with functions rather than constants.

Definition 6.2 (Built-in Skolem functions)**.** An $\mathcal{L}$-theory T has built-in Skolem functions if for all $\mathcal{L}$-formulas $\psi(v, w_1, ..., w_n)$, there is an n-ary $f \in \mathcal{F}$ such that
$T \models \forall \bar{w}(\exists v\ \phi(v, \bar{w}) \to \phi(f(\bar{w}, w)))$.

Proposition 6.3. Let T be an $\mathcal{L}$-theory. Then there is an extension $\mathcal{L}^* \geq \mathcal{L}$ and T^* a $\mathcal{L}^*$-theory such that $T \subset T^*$, where T^* has built-in Skolem-functions and if $\mathcal{M} \models T$, then $\mathcal{M}$ can be expanded to $\mathcal{M}^* \models T^*$. We can choose $\mathcal{L}^*$ such that $|\mathcal{L}^*| = |\mathcal{L}| + |\aleph_0|$. We call T^* the skolemization of T.

Proof. Define $\mathcal{L}_0 \subset \mathcal{L}_1 \subset ...$ where the union of the $\mathcal{L}_i$ is $\mathcal{L}$. Define T to be the union of $T_0 \subset T_1 \subset ...$ where each T_i is an $\mathcal{L}_i$ theory. Given an $\mathcal{L}_i$, let $\mathcal{L}_{i+1}$ be $\mathcal{L}_i \cup \{f_\phi : \phi(v, \bar{w})$ is an $\mathcal{L}_i$ formula and ϕ is n-ary$\}$. For an $\mathcal{L}_i$ formula $\phi(v, \bar{w})$ let Φ_ϕ be the formula
$\forall \bar{w}((\exists v\ \phi(v, \bar{w})) \to \phi(f_\phi(\bar{w}), \bar{w})$
Define $T_{i+1} = T_i \cup \{\Phi_\phi : \phi$ is an $\mathcal{L}_i$-formula$\}$. Now we will show that T has built-in Skolem functions.

Claim: If $\mathcal{M} \models T_i$ then we can interpret the function symbols in $\mathcal{L}_{i+1}/L$ so the $\mathcal{M} \models T_{i+1}$.

Proof of claim: If $\phi(v, \bar{w})$ is an $\mathcal{L}_i/L$-formula, define $g : M^n \to M$. Let $b_0 \in M$. For $\bar{a} \in M^n$, let $X_{\bar{a}} = \{b \in M : \mathcal{M} \models \phi(b, \bar{a})\}$.

If $X_{\bar{a}} \neq \emptyset$, let $g(\bar{a}) \in X_{\bar{a}}$. If $X_{\bar{a}} = \emptyset$ let $g(\bar{a}) = b_0$. If $\mathcal{M} \models \exists v\ \phi(v, \bar{a})$ then $\mathcal{M} \models \phi(g(\bar{a}), \bar{a})$, let $f_\phi^{\mathcal{M}} = g$. Therefore $\mathcal{M} \models \Phi_\phi$. Let $\mathcal{L}^* = \bigcup_{i \in \mathbb{N}} \mathcal{L}_i$ and $T^* = \bigcup_{i \in \mathbb{N}} T_i$.

Then $\Phi_\phi \in T_{i+1} \subset T^*$. Therefore T^* has built-in Skolem functions. Now we will iterate the claim.

For any $\mathcal{M} \models T$ we can interpret the symbols of $\mathcal{L}^*/\mathcal{L}$ to make $\mathcal{M} \models T$. We added one function to $\mathcal{L}_{i+1}$ for each $\mathcal{L}_i$-formula, so $|\mathcal{L}_{i+1}| = |\mathcal{L}_i| + \aleph_0$.

□

Proposition 6.4 (Downward Lowenheim Skolem)**.** Let $\mathcal{M}$ be an $\mathcal{L}$-structure and let $X \subset \mathcal{M}$. Then there is an elementary substructure $\mathcal{N} \prec \mathcal{M}$ such that $X \subset N$ and $|N| \leq |X| + |\mathcal{L}| + \aleph_0$.

Proof. By the previous lemma, we know that $Th(\mathcal{M})$ has built-in Skolem-functions.

Let $X_0 = X$. Given X_i, let $X_{i+1} = X_i \cup \{f^{\mathcal{M}}(\bar{a}) : f \text{ is n-ary, } \bar{a} \in X_i^n \text{ and } n \in \mathbb{N}\}$. Let $N = \bigcup_{i\in\mathbb{N}} X_i$. Then $|N| \leq |X| + |\mathcal{L}| + \aleph_0$. Define $\mathcal{N}$ by interpreting the constants, function, and relational symbols such that $\mathcal{N} \prec \mathcal{M}$. □

7. Downward Löwenheim-Skolem

Theorem 7.1 (Downward Löwenheim-Skolem)**.** Let $\mathcal{M}$ be an $\mathcal{L}$-structure and let $X \subseteq \mathcal{M}$. Then there exists $\mathcal{N} \prec \mathcal{M}$, such that $x \subseteq \mathcal{N}$ and $|\mathcal{N}| \leq |X| + |\mathcal{L}| + \aleph_0$.

Proof. Last time we showed that $\mathcal{M}$ has built in Skolem functions and let $X_{i+1} = X_i \cup \{f^{\mathcal{M}}(\bar{a}) : f \in \mathcal{F}, \bar{a} \in X_i^n\}$. Then we set $N = \bigcup X_i$ and showed $|\mathcal{N}| \leq |X| + |\mathcal{L}| + \aleph_0$. Now we will define $\mathcal{N}$ as follows

- If f is an n-ary function symbol and $\bar{a} \in N^n$, then $\bar{a} \in X_i$ for some i. Then it follows that $f^{\mathcal{M}}(\bar{a}) \in X_{i+1} \subseteq N$. Define $f^{\mathcal{N}} = f^{\mathcal{M}}|_N$.
- If R is an n-ary relation symbol, then similarly define $R^{\mathcal{N}} = R^{\mathcal{M}} \bigcap N^n$.
- If c is a constant symbol, consider the $\mathcal{L}$-formula $\exists v\ v = c$. Then there is a Skolem function f such that $f(x) = c^{\mathcal{M}}$ for any $x \in X \subseteq N$. Then define $c^{\mathcal{N}} = c^{\mathcal{M}} \in N$.

So $\mathcal{N}$ is an $\mathcal{L}$-structure and $\mathcal{N} \subseteq \mathcal{M}$. Lastly we shall show $\mathcal{N} \prec \mathcal{M}$. Let $\phi(v, \bar{w})$ be an $\mathcal{L}$-formula, $\bar{a} \in N$, and $\mathcal{M} \models \phi(b, \bar{a})$ for some $b \in M$. Then there is a function f_ϕ such that $\mathcal{M} \models \phi(f_\phi(\bar{a}), \bar{a})$ where $f_\phi(\bar{a}) \in N$. Then, by the Tarski-Vaught theorem, $\mathcal{N} \prec \mathcal{M}$. □

8. Elementary Chains

Definition 8.1. Let $(I, <)$ be a linear order. Suppose $\mathcal{M}_i$ is an $\mathcal{L}$-structure for each $i \in I$. Then $(\mathcal{M}_i : i \in I)$ is a *chain* if $\mathcal{M}_i \subseteq \mathcal{M}_j$ for all $i < j$. If $\mathcal{M}_i \prec \mathcal{M}_j$ for all $i < j$, we call $(\mathcal{M}_i : i \in I)$ an *elementary chain.*

If $(\mathcal{M}_i : i \in I)$ is a nonempty chain of structures, then the union of the chain, $\mathcal{M} = \bigcup_{i\in I} \mathcal{M}_i$ from $M = \bigcup_{i\in I} M_i$ is defined to be as follows:

- If f is an n-ary function symbol and $\bar{a} \in M^n$, then $\bar{a} \in M_i$ for some i, as $(I, <)$ is a linear order. As for all $i < j$, $\mathcal{M}_i \subseteq \mathcal{M}_j$, let $f^{\mathcal{M}} = \bigcup_{i\in I} f^{\mathcal{M}_i}$
- If R is an n-ary relation symbol and $\bar{a} \in M^n$, then similarly define $R^{\mathcal{M}} = \bigcup_{i\in I} R^{\mathcal{M}_i}$.
- If c is a constant symbol, then for $i < j$, $c^{\mathcal{M}_i} = c^{\mathcal{M}_j}$. So define $c^{\mathcal{M}} = c^{\mathcal{M}_i}$ for any $i \in I$.

Proposition 8.2. Suppose $(I, <)$ is a linear order and $(\mathcal{M}_i : i \in I)$ is an elementary chain. If $\mathcal{M} = \bigcup \mathcal{M}_i$, then $\mathcal{M}_i \prec \mathcal{M}$.

Proof. By induction, we show for all $i \in I$, all $\mathcal{L}$-formulas ϕ, and all $\bar{a} \in M_i^n$

$$\mathcal{M} \models \phi(\bar{a}) \Leftrightarrow \mathcal{M}_i \models \phi(\bar{a})$$

This is clear for atomic formulas, and for conjunctions and negations. We will show that this holds for quantifiers. Suppose ϕ is $\exists v\ \psi(v, \bar{a})$ and the claim holds for ψ, that is $\mathcal{M} \models \psi(\bar{a}) \Leftrightarrow \mathcal{M}_i \models \psi(\bar{a})$.
Assume $\mathcal{M}_i \models \exists v\ \psi(v, \bar{a})$. So, for some $b \in M_i$, $\mathcal{M}_i \models \psi(b, \bar{a})$. By the claim, $\mathcal{M} \models \psi(b, \bar{a})$ so $\mathcal{M} \models \exists v\ \psi(v, \bar{a})$.
Now assume $\mathcal{M} \models \exists v\ \psi(v, \bar{a})$. Then for some $b \in M$, $\mathcal{M} \models \psi(b, \bar{a})$. Now $b \in M_j$ for some j.

If $i > j$, then $\mathcal{M}_i \models \psi(b, \bar{a})$ by the claim on ψ. If $i < j$, as $\mathcal{M}_i \prec \mathcal{M}_j$ and $\mathcal{M}_j \models \exists v\ \psi(v, \bar{a})$, conclude $\mathcal{M}_i \models \exists v\ \psi(v, \bar{a})$ □

9. Back-n-Forth

Theorem 9.1. The set of Dense Linear Orders Without Endpoints (DLOWE) is $\aleph_0$-categorical and complete.

Proof. Let $(A, <)$, $(B, <)$ be countable models of DLOWE. Let $A = \{a_0, a_1, ...\}$ and $B = \{b_0, b_1, ...\}$ be enumerations of A and B. We build a sequence of maps $f_i : A_i \to B_i$ where $A_i \subseteq A$ and $B_i \subseteq B$ are finite, for all $i < j$, $f_i \subseteq f_j$, and if $x, y \in A_i$ and $x < y$ then $f_i(X) < f_I(y)$. We will choose A_i and B_i such that $A = \bigcup A_i$ and $B = \bigcup B_i$. Then $f = \bigcup f_i : A \to B$ is the desired isomorphism. We will proceed in stages to create each A_i and B_i. At odd stages we will ensure that $A = \bigcup A_i$ and at even stages we will ensure that $B = \bigcup B_i$.

Stage 0: $A_0 = B_0 = f_0 = \emptyset$

Stage $n + 1 = 2m + 1$**:** The odd stage. We ensure $a_m \in A_n$. If $a_m \in A_n$ already, let $A_{n+1} = A_n$, $B_{n+1} = B_n$, and $f_{n+1} = f_n$. Assume $a_m \notin A_n$. To add a_m to $dom(f_{n+1})$ we need to find $B \in B \backslash B_n$ such that $\alpha \in a_m$ if and only if $\forall \alpha \in A_n\ (f_n(\alpha) < b)$. on of the following cases will hold.

Case 1: $\forall \mathbf{x} \in \mathbf{A_n}\ (\mathbf{a_m} > \mathbf{x})$. Then we can choose a $b_m \in B \backslash B_n$ such that for every $y \in B_n,\ (b_m > y)$.

Case 2: $\forall \mathbf{x} \in \mathbf{A_n}\ (\mathbf{a_m} < \mathbf{x})$. Similarly we can choose a $b_m \in B \backslash B_n$ such that for every $y \in B_n,\ (b_m < y)$.

Case 3: Otherwise. Then there are $\alpha, \beta \in A_n$ such that $\alpha < a_m < \beta$ and for all $\gamma \in A_n,\ (\gamma \leq \alpha \vee \gamma \geq \beta)$. Then we choose b_m such that $f_n(\alpha) < b_m < f_n(\beta)$.

In each of these cases let $A_{n+1} = A \cup a_m$, $B_{n+1} = B \cup b_m$, and $f_{n+1} = fn \cup \{(a_m, b_m)\}$.

Stage $n + 1 = 2m$**:** The even stage is similar to the above odd stage.

□

10. Back-n-Forth (Continued)

We begin by reiterating and finishing the back-and-forth argument.

Theorem 10.1. DLOWE is $\aleph_0$-categorical.

Proof. We let $(A, <_A)$ and $(B, <_B)$ be models of DLOWE, and we enumerate A as $\{a_0, a_1, \ldots, \}$ and construct an enumeration for B as $\{b_0, b_1, \ldots, \ldots\}$ (countability, of course, allows this). For any $n \in \omega$, let $A_n = \{a_0, \ldots, a_n\}$ and assume that $B_n = \{b_0, \ldots, b_n\}$ is defined. Either a_{n+1} is to the left of everything (i.e. for all $i \leq n$, $a_{n+1} <_A a_i$), in which case we define b_{n+1} as some $b \in B$ to the left of B_n (i.e., all the points in B_n), which exists because $(B, <_B)$ has no endpoints. If a_{n+1} is to the right of all of the points in A_n, define b_{n+1} as some $b \in B$ to the right of B_n. If there are some $0 \leq i < j \leq n$ for which $a_i < a_{n+1} < a_j$, use density of B to find some point $b \in B$ such that $b_i < b < b_j$, and define $b_{n+1} = b$. Thus define a sequence of functions $\{f_n\}_{n \in \omega}$ recursively by $f_0(a_0) = b_0$ and $f_{n+1} = f_n \cup \{a_{n+1}, b_{n+1}\}$. Let f be the function $\cup_{n \in \omega} f_n$. Then f is our desired order isomorphism. □

Claim 10.2. DLOWE is complete.

Proof. By Vaught's test: DLOWE is $\aleph_0$-categorical, with no finite models. Hence DLOWE is complete. □

11. Quantifier Elimination

Sometimes, complexly-definable sets get odd. An example of a "complicated" set: in the structure $(\mathbb{N}, +, \times, <)$, the "quantifier-free" sets are those defined by polynomial equations and inequalities.

Definition 11.1. Let $\mathcal{M}$ be a $\mathcal{L}$-structure for some first-order language $\mathcal{L}$. We say that a set $A \subseteq M$ is a Σ_1^0 **set** if A is definable by a formula of the form $\exists x\ \phi(x, y)$, for a quantifier-free formula ϕ. We similarly say A is Π_1^0 is A is definable by a formula of the form $\forall x\ \phi(x, y)$, where ϕ is, again, quantifier-free. Building up (but not entirely), A is a Σ_2^0 set if A is definable by $\exists x \forall y\ \phi(x, y, z)$, for quantifier-free ϕ, and so on.

Sometimes we get lucky: quantified statements are equivalent to quantifier-free sentences.

Example 11.2. Let $\phi(a, b, c, x)$ be $ax^2 + bx + c = 0$. We then have

$$\mathbb{R} \models \phi(a, b, c, x) \leftrightarrow [(a \neq 0 \wedge b^2 - 4ac \geq 0) \vee (a = 0 \wedge (b = 0 \rightarrow c = 0)].$$

So $\exists x\ (ax^2 + bx + c = 0)$ is equivalent to $[(a \neq 0 \wedge b^2 - 4ac \geq 0) \vee (a = 0 \wedge (b = 0 \rightarrow c = 0)]$. We even have

$$\mathbb{C} \models \exists x\ \phi(a, b, c, x) \leftrightarrow (a \neq 0 \vee b \neq 0 \vee c \neq 0).$$

However, in $\mathbb{Q}$, $\phi(a, b, c, x)$ is NOT equivalent to a quantifier-free sentence.

Example 11.3. Let $\phi(a, b, c, d)$ be

$$\exists x \exists y \exists u \exists v\ (xa + yc = 1 \wedge xb + yd = 0 \wedge ua + vc = 0 \wedge ub + vd = 1).$$

This is the sentence:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \text{ is invertible.}$$

In any field F, $(F, +_F, \cdot_F, 0_F, 1_F) \models \phi(a, b, c, d) \leftrightarrow (ad - bc \neq 0)$.

Definition 11.4. An $\mathcal{L}$-theory T has **quantifier elimination** if, for every $\mathcal{L}$-formula ϕ, there exists a quantifier-free formula ψ such that

$$T \models \forall \bar{x}\ (\phi(\bar{x}) \leftrightarrow \psi(\bar{x})).$$

Lemma 11.5. Let $(A, <)$ and $(B, <)$ be models of countable DLOWE's, $a_1, \ldots, a_n \in A$ and $b_1, \ldots, b_n \in B$ and suppose that $a_1 < \ldots < a_n$ and $b_1 < \ldots < b_n$. There there exists a function $f : A \rightarrow B$ such that f is an isomorphism and $f(a_i) = b_i$ for all $1 \leq i \leq n$.

Proof. Use the back-and-forth argument. □

Theorem 11.6. DLOWE has quantifier elimination.

Proof. First, suppose that ϕ is an $\mathcal{L}_{\text{DLOWE}}$-sentence. If $\mathbb{Q} \models \phi$, then since DLOWE is complete, we have DLOWE $\models \phi$. We then have

$$\text{DLOWE} \models \forall x\ (\phi \leftrightarrow x = x).$$

If $\mathbb{Q} \models \neg\phi$, then DLOWE $\models \neg\phi$. So

$$\text{DLOWE} \models \forall x\ (\phi \leftrightarrow x \neq x).$$

Suppose that ϕ is an $\mathcal{L}_{\text{DLOWE}}$-formula with free variables $x_1, \ldots, x_n$, for $n \geq 1$. We show that there exists a quantifier-free formula ψ with free variables $x_1, \ldots, x_n$ such that

$$\mathbb{Q} \models \forall \bar{x}\ (\phi(\bar{x}) \leftrightarrow \psi(\bar{x})).$$

For a function $\sigma : \{(i,j) : 1 \leq i < j \leq n\} \to \{0,1,2\}$, let $\chi_\sigma(x_1, \ldots, x_n)$ be the formula

$$\bigwedge_{\sigma(i,j)=0} (x_i = x_j)\ \wedge \bigwedge_{\sigma(i,j)=1} (x_i < x_j)\ \wedge \bigwedge_{\sigma(i,j)=2} (x_i > x_j).$$

We call χ_σ a sign-condition which describes a possibly inconsistent arrangement of n elements in a linearly-ordered set.

Let ϕ be an $\mathcal{L}$-formula with $n \geq 1$ free variables. Let

$$\Lambda_\phi = \{\chi_\sigma : \exists \bar{a} \in \mathbb{Q}^n\ (\mathbb{Q} \models \chi_\sigma(\bar{a} \wedge \phi(\bar{a})\}.$$

We have two cases to consider:

(1) $\Lambda_\phi = \emptyset$; then $\mathbb{Q} \models \forall x\ \neg\phi(\bar{x})$. Thus $\mathbb{Q} \models \forall \bar{x}\ (\phi(\bar{x}) \leftrightarrow x_i \neq x_j)$.

(2) $\Lambda_\phi \neq \emptyset$: let $\psi_\phi(\bar{x})$ be $\bigvee_\sigma \chi_\sigma(\bar{x})$ (Λ_ϕ is finite). By choice of Λ_ϕ, $\mathbb{Q} \models \forall \bar{x}(\phi(\bar{x}) \to \psi_\phi(\bar{x}))$.
On the other hand, suppose that $\bar{b}$ satisfies $\mathbb{Q} \models \phi_\psi(\bar{b})$. Let $\sigma \in \Lambda_\phi$ be such that $\mathbb{Q} \models \chi_\sigma(\bar{b})$; such σ exists since ψ_ϕ is $\bigvee_\sigma \chi_\sigma$. Let $\bar{a} \in \mathbb{Q}^n$ be such that $\mathbb{Q} \models \phi(\bar{a}) \wedge \chi_\sigma(\bar{a})$. By the lemma, let $f : \mathbb{Q} \to \mathbb{Q}$ be an order isomorphism such that $f(a_i) = b_i$ for all $1 \leq i \leq n$. But then $\mathbb{Q} \models \phi(\bar{b})$ since f is an $\mathcal{L}_{\text{DLOWE}}$-isomorphism. Hence $\mathbb{Q} \models \forall \bar{x}\ (\psi_\phi(\bar{x}) \to \phi(\bar{x}))$, so $\mathbb{Q} \models \forall \bar{x}\ (\phi(\bar{x}) \leftrightarrow \psi_\phi(\bar{x}))$.

In both cases, it then follows that DLOWE $\models \forall \bar{x}\ (\phi(\bar{x}) \leftrightarrow \psi_\phi(\bar{x}))$ by completeness of DLOWE. Therefore DLOWE has quantifier elimination. □

Last time, we introduced the notion of *quantifier elimination*: an $\mathcal{L}$-theory T has *quantifier elimination* if, for every $\mathcal{L}$-formula $\phi(\bar{x})$ of n free variables, there exists a quantifier-free $\psi(\bar{x})$ such that $T \models \forall \bar{x}\ (\phi(\bar{x}) \leftrightarrow \psi(\bar{x}))$. We also proved that DLOWE has quantifier elimination.

Definition 11.7. An $\mathcal{L}$-theory is **model-complete** if, for all $\mathcal{L}$-structures $\mathcal{M}$ and $\mathcal{N}$, if $\mathcal{M}, \mathcal{N} \models T$ and $\mathcal{M} \subseteq \mathcal{N}$, then $\mathcal{M} \preceq \mathcal{N}$.

Proposition 11.8. Let T be an $\mathcal{L}$-theory. If T has quantifier elimination, then T is model-complete.

Proof. Let $\mathcal{M}$ and $\mathcal{N}$ be $\mathcal{L}$-structures and suppose that $\mathcal{M}, \mathcal{N} \models T$ and $\mathcal{M} \subseteq \mathcal{N}$. Let $\phi(\bar{x})$ be an $\mathcal{L}$-formula and let $\bar{a} \in M^n$. By quantifier elimination of T, there exists a quantifier-free formula, say $\psi(\bar{x})$, such that $T \models \forall \bar{x}\ (\phi(\bar{x}) \leftrightarrow \psi(\bar{x}))$. Since $\mathcal{M} \subseteq \mathcal{N}$, we have $\mathcal{M} \models \psi(\bar{b}) \iff \mathcal{N} \models \psi(\bar{b})$, for every $\bar{b} \in M^n$. So $\mathcal{M} \models \phi(\bar{a}) \iff \mathcal{M} \models \psi(\bar{a}) \iff \mathcal{N} \models \psi(\bar{a}) \iff \mathcal{N} \models \phi(\bar{a})$. As $\phi(\bar{x})$ and $\bar{a}$ were arbitrary, it follows that $\mathcal{M} \preceq \mathcal{N}$, as desired. □

Proposition 11.9. Let T be an $\mathcal{L}$-theory and suppose that T is model-complete. Suppose also that there is an $\mathcal{L}$-structure $\mathcal{M}_0 \vDash T$ that $\mathcal{L}$-embeds into every $\mathcal{M} \models T$. (We say that $\mathcal{M}_0$ is a **prime model**.) Then T is complete.

Proof. Let $\mathcal{M}_0 \vDash T$ be a prime model. By model-completeness, every $\mathcal{L}$-embedding $\mathcal{M}_0 \to \mathcal{M}$ for $\mathcal{M} \models T$ is elementary. Hence $\mathcal{M}_0 \equiv \mathcal{M}$. Thus every two models of T are elementarily equivalent by transitivity of $\equiv$.

Suppose T is not complete. Let ϕ be an $\mathcal{L}$-formula such that $T \nvDash \phi$ and $T \nvDash \neg\phi$. Let $\mathcal{M}, \mathcal{N}$ be $\mathcal{L}$-structures such that $\mathcal{M}, \mathcal{N} \vDash T$ but $\mathcal{M} \vDash \phi$ and $\mathcal{N} \vDash \neg\phi$. But then $\mathcal{M} \not\equiv \mathcal{N}$, contradicting that every two models of T are elementarily equivalent. □

We now have three completeness results. Let T be an $\mathcal{L}$-theory.

(1) (T is κ-categorical for some infinite κ and T has no finite models) $\implies$ T is complete.
(2) T has quantifier elimination $\implies$ T is model-complete.
(3) (T is model-complete and T has a prime model) $\implies$ T is complete.

Theorem 11.10. DLO with endpoints is not complete.

Proof. Let $\mathcal{M} = ([0,1], <)$. Let $\mathcal{N} = ([-1,1], <)$. Note that $\mathcal{M} \subseteq \mathcal{N}$ (i.e., $j : \mathcal{M} \hookrightarrow \mathcal{N}$ is an $\mathcal{L}$-embedding). Consider the formula $\phi(x)$,

$$\forall y\ (x < y \vee x = y).$$

We have $\mathcal{M} \vDash \phi(0)$ but $\mathcal{N} \nvDash \phi(0)$. Thus $\mathcal{M} \not\preceq \mathcal{N}$, so DLO with endpoints is not complete. □

Corollary 11.11. DLO with endpoints does not have quantifier elimination.

Let $\mathcal{L}$ be the expanded language (of DLO) $\{<, c_0, c_1, \ldots\}$. Let $T_3 = \text{DLOWE} \cup \{c_i < c_j : i < j\}$. We claim that T_3 is complete. First, we prove two preliminary results.

(1) For any language $\mathcal{L}$, two $\mathcal{L}$-structures $\mathcal{M}, \mathcal{N}$ are elementarily equivalent iff they are elementarily equivalent for every finite $\mathcal{L}' \subseteq \mathcal{L}$.

Proof. (): Suppose $\mathcal{M} \equiv \mathcal{N}$ for finite $\mathcal{L}' \subseteq \mathcal{L}$. For any $\mathcal{L}$-sentence ϕ, let $\mathcal{L}_\phi$ consist of the symbols of $\mathcal{L}$ occuring in ϕ. Since first-order sentences are of finite length, $\mathcal{L}_\phi$ is finite for any $\mathcal{L}$-sentence ϕ. Then since $\mathcal{M} \equiv \mathcal{N}$ for any $\mathcal{L}_\phi$-sentence, $\mathcal{M} \vDash \phi \iff \mathcal{N} \vDash \phi$. □

(2) If $\mathcal{L}$ is countable, T is an $\mathcal{L}$-theory with no finite models, and every two countable models of T are elementarily equivalent, then T is complete.

Proof. Suppose there is a formula, call it ϕ, such that $T \nvDash \phi$ and $T \nvDash \neg\phi$. Let $\mathcal{M}_0$ and $\mathcal{M}_1$ be $\mathcal{L}$-structures such that $\mathcal{M}_0, \mathcal{M}_1 \vDash T$ and $\mathcal{M}_0 \vDash \phi$ and $\mathcal{M}_1 \vDash \neq \phi$. By the Downward Lowenheim-Skolem theorem, let $\mathcal{N}_0 \preceq \mathcal{M}_0$, and $\mathcal{N}_1 \preceq \mathcal{M}_1$, be such that $\mathcal{N}_i \vDash T$ and $|N_i| \leq |\mathcal{L}| + \aleph_0 = \aleph_0$. Since T has no finite models, we have $|N_i| \geq \aleph_0$, so

$|N_i| = \aleph_0$. It follows from $\mathcal{N}_0 \preceq \mathcal{M}_0$ that $\mathcal{N}_0 \vDash \phi$. Similarly, $\mathcal{N}_1 \vDash \neg\phi$, so $\mathcal{N}_1 \nvDash \phi$. But this contradicts that any two countable models of T are elementarily equivalent. □

We are now ready to show that T_3 is complete.

Proof. Given countable $\mathcal{L}$-structures $\mathcal{M}, \mathcal{N} \vDash T$, and given a finite sublanguage $\mathcal{L}' \subseteq \mathcal{L}$, $\mathcal{M} \equiv \mathcal{N}$ as $\mathcal{L}'$-structures by restriction of back-and-forth, and defining the isomorphism for constants first. By (1), $\mathcal{M} \equiv \mathcal{N}$ as $\mathcal{L}$-structures. By (2), since $\mathcal{L}$ is countable, and T_3 has no finite models, it follows T_3 is complete. □

Definition 11.12. For any complete $\mathcal{L}$-theory T and any cardinal κ, let $I(T, \kappa)$ be the number of models of cardinality κ up to isomorphism. $I(T, \kappa)$ is the **spectrum** of the theory T.

Vaught: For $n \geq 3$, there exists a complete $\mathcal{L}$-theory T such that $I(T, \aleph_0) = n$. For $n = 2$, there is no such theory.

Question 11.13. Does there exist an $\mathcal{L}$-theory T such that $\aleph_0 < I(T, \aleph_0) < 2^{\aleph_0}$?

If the continuum hypothesis holds, then no. If not, then who knows? The *Vaught Conjecure* is no, independently of CH.

12. Computability Theory Introduction

Given an $\mathcal{L}$-theory, T, and an $\mathcal{L}$-sentence, ϕ, we would like to know whether $T \models \phi$. Better yet, we would like to have an algorithm based on T that, given ϕ, will output "yes" if $T \models \phi$ and "no" if $T \not\models \phi$. In 1900, Hilbert posed the question of whether there was a universal algorithm of this type to determine if there is an integer solution to a polynomial with finitely many unknowns. In 1928, Hilbert followed up this question with another: Is there an algorithm that, given an $\mathcal{L}$-sentence ϕ, would output "yes" if ϕ is true in all $\mathcal{L}$-structures and "no" otherwise? The answer to both of these questions is no, there does not exist such an algorithm.

In order to show no such algorithm exists, we need to formalize the notion of an algorithm, or a computable function. One type of computable function is the partial recursive function.

Definition 12.1 (Primitive Recursive Functions)**.** The collection of primitive recursive functions, $PRIM$, is the smallest class of number theoretic functions containing the following initial functions:

(1) Zero function: $c_0(n) = 0$
(2) Successor function: $s(n) = n + 1$
(3) Projection function: $\Pi^m_k(n_1, ...n_m) = n_k$

that is closed under the following schema for defining new functions:

(4) Composition: If $g : \mathbb{N}^{l+1} \to \mathbb{N}$ and $h_0, h_1, ..., h_l : \mathbb{N}^m \to \mathbb{N}$ are primitive recursive functions, then $f : \mathbb{N}^m \to \mathbb{N}$ defined by $f(\bar{h}) = g(h_0(\bar{n}), ..., h_l(\bar{n})$ is a primitive recursive function.
(5) Primitive recursion: If $g : \mathbb{N}^m \to \mathbb{N}$ and $h : \mathbb{N}^{m+2} \to \mathbb{N}$ are primitive recursive functions, then so is $f : \mathbb{N}^{m+1} \to \mathbb{N}$ defined by $f(\bar{k}, 0) = g(\bar{k})$ and $f(\bar{k}, n + 1) = h(\bar{k}, n, f(\bar{k}, n))$

Examples of functions in $PRIM$:

(1) Constant function: for a given $k \in \mathbb{N}$, $\forall n\ c_k(n) = k$.
By induction on k. For $k = 0$, $C_0 \in PRIM$. Suppose $c_k \in PRIM$. Then $c_{k+1}(n) = c_k(n) + 1 = s(c_k(n)) = s \circ c_k(n) \in PRIM$ by composition.

(2) Addition function: $\forall m \forall n\ (a(m, n) = m + n)$
Define $a(m, n)$ by primitive recursion as follows.
$a(m, 0) = \Pi^1_1(m)$
$a(m, n + 1) = a(m, n) + 1 = s(\Pi^3_3(m, n + 1, a(m, n))$

(3) Multiplication function: $\forall k \forall n\ (m(k, n) = k * n)$.
Define $m(k, n)$ by primitive recursion as follows.
$m(k, 0) = c_0(k)$
$m(k, n + 1) = m(k, n) + k = a(\Pi^3_3(k, n, m(k, n)), \Pi^3_1(k, n, m(k, n)))$

(4) Predecessor function: $\delta(m) = \begin{cases} m - 1 & \text{if } m > 0 \\ 0 & \text{if } m = 0 \end{cases}$
Define $\delta(m)$ by primitive recursion as follows.
$\delta(0) = c_0(n)$
$\delta(m + 1) = m = \Pi^2_1(m, \delta(m))$

(5) Recursive difference function: $m \dot{-} n = \begin{cases} m - n & \text{if } m \geq n \\ 0 & \text{if } m < n \end{cases}$
Define by primitive recursion as follows.
$m \dot{-} 0 = m$
$m \dot{-} (n + 1) = \delta(m \dot{-} n) = \delta(\Pi^3_3(m, n, m \dot{-} n))$

(6) Absolute difference function: $|m - n| = \begin{cases} m - n & \text{if } m \geq n \\ n - m & \text{if } n \geq m \end{cases}$

(7) Sign function: $sg(n) = \begin{cases} 0 & \text{if } n = 0 \\ 1 & \text{otherwise} \end{cases}$

(8) Remainder function: $rem(m, n) = \begin{cases} 0 & \text{if } m = 0 \\ \text{the remainder upon division of } n \text{by } m & \text{otherwise} \end{cases}$

(9) Characteristic function of $m|n$: $d(m, n) = \begin{cases} 1 & \text{if } m|n \\ 0 & \text{otherwise} \end{cases}$

(10) Bounded sums: If $f(\bar{m}, n) \in PRIM$, then so is $h(\bar{m}, p) = \sum_{n \leq p} f(\bar{m}, n)$.
Define by primitive recursion as follows.
$h(\bar{m}, 0) = f(\bar{m}, 0)$
$h(\bar{m}, p + 1) = f(\bar{m}, p + 1) + \sum_{n \leq p} f(\bar{m}, n) = f(\bar{m}, p + 1) + h(\bar{m}, p)$

13. Partial Recursive Functions

Definition 13.1. A function $f : A \to B$ is *total* if $f(x)$ is defined for every $x \in A$, denoted $f(x) \downarrow$. If it is undefined at x, we write $f(x) \uparrow$, and we say that f is *partial*.

Every function in $PRIM$ is total. We can extend $PRIM$ to a collection of (possibly partial) functions. We call these partial recursive functions.

$PRIM \subseteq$ Partial Recursive Functions

Definition 13.2. The collection of *partial recursive functions* is the smallest collection of number-theoretic functions that contains the primitive recursive functions and is closed under the following:
If $g(\overline{n}, m)$ is partial recursive, then so is f given by

$$f(\overline{n}) = \mu m[g(\overline{n}, m) = 0], \text{ where}$$
$$\mu m[g(\overline{n}, m) = 0] = m_0 \iff g(\overline{n}, m_0) = 0 \text{ and } \forall m < m_0(g(\overline{n}, m) \downarrow \neq 0).$$

The μ operator is an *unbounded search*. Compute $g(\overline{n}, 0), g(\overline{n}, 1), \ldots$ until m is found such that $g(\overline{n}, m) = 0$.

A function f defined by g and using the μ operator can fail in two different ways:

- $g(\overline{n}, k) \downarrow \neq 0 \forall k \in \mathbb{N}$
- $g(\overline{n}, 0) \downarrow \neq 0, \ldots, g(\overline{n}, k) \downarrow \neq 0, g(\overline{n}, k+1) \uparrow$

Theorem 13.3 (Church's Thesis)**.** The following two statements hold:

- f is recursive $\iff$ f is total effectively computable.
- f is partial recursive $\iff$ f is effectively computable.

This is the same as the Church-Turing thesis, without Turing machines. Here, "effectively computable" means there is some description of an algorithm in some language that can be used to compute any value of $f(x)$ where $f(x) \downarrow$.

Consequences:

(1) If f is not partial recursive, then it is not effectively computable, so there is no algorithm to determine the values of f.
(2) If we can give an intuitively reasonable description of an algorithm for computing f, then we can find a description of f as a partial recursive function. This allows us to sketch an algorithm, rather than building a partial recursive function from the gound up.

Definition 13.4. Let $S \subseteq \mathbb{N}$ be any subset and $R \subseteq \mathbb{N}^k$ be any relation.

(1) The *characteristic functions* of S and R are given by

$$\chi_S(x) = \begin{cases} 1 & \text{if } x \in S \\ 0 & \text{if } x \notin S \end{cases}$$

$$\chi_R(x) = \begin{cases} 1 & \text{if } R(x) \text{ holds} \\ 0 & \text{if } R(x) \text{ holds} \end{cases}$$

(2) S or R are primitive recursive if χ_S or χ_R are primitive recursive, respectively.

Example 13.5. The relation $n < m$ is recursive: $\chi_R(x) = sg(y \dot{-} x)$.

Proposition 13.6. If P, Q are (primitive) recursive relations, then so are $\neg P, P \wedge Q, P \vee Q$.

Proof.

$$\chi_P(x) = \begin{cases} 1 & \text{if } P(x) \text{ holds} \\ 0 & \text{if } \neg P(x) \text{ holds} \end{cases}$$

$\chi_{P\wedge Q}(x) = \chi_P(x) * \chi_Q(x)$
$\chi_{P\vee Q} = sg(\chi_P(x) + \chi_Q(x))$ □

Proposition 13.7. Every finite set is primitive recursive.

Proof. By induction on $n = |S|$.
If $n = 1$ then $S = \{a\}$ and $\chi_S(x) = \overline{sg}(|x - a|)$.
Suppose the proposition holds for sets of cardinality up to k.
Let $S = \{a_1, a_2, \ldots a_{k+1}\}$. Set $S_0 = \{a_1 \ldots a_k\}$ and $S_1 = \{a_{k+1}\}$. Then

$$\chi_S(x) = \chi_{S_0}(x) + \chi_{S_1}(x).$$

□

Theorem 13.8. Let f be a recursive function with an infinite range. Then there is an injective recursive g where range(g) = range(f).

Proof. Define a sequence $n_0 < n_1 < \cdots$, where $n_0 = 0$. For $i \geq 1$, let n_i be the least i such that $f(n_i) \neq f(n_j)$ for $j < i$. Let $g(i) = n_i$. For example,

$$f(1) = 19, f(2) = 19, f(3) = 19, f(4) = 12$$
$$g(1) = 19, g(4) = 12$$

If $g(i) = g(j)$, then $f(n_i) = f(n_j) \iff n_j = n_i$ by the minimality of n_i, so $i = j$.
Clearly, range$(g) \subseteq$ range(f).
Suppose $k \in$ range(f). There is a least n such that $f(n) = k$. Then $n_i = n$ for some i, and $g(i) = f(n_i) = k$.
Hence range(g) = range(f). □

Proposition 13.9. If $R(\overline{m}, n)$ is recursive and f is recursive, then $R(\overline{m}, f(n))$ is recursive.

Proof. Let $R^*(\overline{m}, n)$ be defined such that

$$R^*(\overline{m}, n) \text{ holds} \iff R(\overline{m}, f(n)) \text{ holds}$$

Then we have

$$\chi_{R^*}(m_1, \ldots, m_k, n) = \chi_R(id(m_1), \ldots, id(m_k), f(n)) \ .$$

□

Lemma 13.10. If $R(\overline{m}, n)$ is recursive, then the function f defined by $f(\overline{m}) = \mu n[R(\overline{m}, n) \text{ holds}]$ is partial recursive.

Proof. Observe $f(\overline{m}) = \mu n[\chi_{\neg R}(\overline{m}, n) = 0]$, and we know $\chi_{\neg R}$ is recursive as χ_R is. □

Example 13.11. Primes and Divisibility

(1) $D(n) = \#$ of divisors of n (including 1 and n) is primitive recursive. Recall that $m|n$ is primitive recursive, so we can define D as follows:

$$D(n) = \sum_{m \leq n} \chi_|(\overline{m}, n)$$

(2)

$$Pr(n) = \begin{cases} 1 & \text{if } n \text{ is prime} \\ 0 & \text{if } n \text{ is composite, } n = 0, n = 1 \end{cases}$$

This is primitive recursive. Note $Pr(n) = 1$ if and only if $D(n) \leq 2$ and $n \neq 0, 1$. We can rewrite $P(r)$ as,

$$Pr(n) = \overline{sg}(D(n) \dot{-} 2) \times sgn(n \dot{-} 1)$$

(3) Let p_n denote the n^{th} prime number, where $p_0 = 2$. The function $n \mapsto p_n$ is recursive, and p_n is total as there are infinitely any primes.

$$p_0 = 2$$
$$p_{n+1} = \mu z[z > p_n \text{ and } Pr(z)]$$

(4) Let $(m)_i$ be the exponent of p_i in the prime factorization of m. For example,

$$(2^2 3^9 5^7)_0 = 2$$
$$(2^2 3^9 5^7)_1 = 9$$
$$(2^2 3^9 5^7)_2 = 7$$

$(m)_i$ is a recursive function in m, i.

Sometimes we would like to code pairs $(x, y) \in \mathbb{N}^2$ as single elements $n \in \mathbb{N}$.

Cooper's pairing function:

$$\langle x, y \rangle = 2^x (2y + 1) \dot{-} 1$$

Better pairing function:

$$\langle x, y \rangle = \frac{1}{2}(x + y)(x + y + 1) + y$$

We can code k-tuples inductively using these pairing functions: $\langle n_1, \ldots, n_k \rangle = \langle \langle n_1 \ldots n_{k-1} \rangle, n_k \rangle$.

We now consider another model of computation.

14. Turing Machines

There are two basic parts, or "hardware", that make a Turing machine.

(1) A tape, subdivided into cells, infinitely extendable in both directions.
(2) A read/write head.

To write out the "program" of a TM, we use:

(1) Tape symbols $S_0, S_1, \ldots, S_n$ that can be written on the tape.
   - S_0 = "0"
   - S_1 = "1"
(2) A finite list of states $q_0, q_1, \ldots, q_n$.
(3) Action symbols that tell the read/write head what to do.
   - L: move to the left.
   - R: move to the right

- 0: erase the current cell
- 1: print 1 on the current cell.

The program of a Turing machine T consists of quadruples of the form
$Q = q_i S A q_j$ where q_i is the initial state, S is a symbol on the tape, A is an action, and q_j is the end state. If T is in the state q_i reading the symbol S, it takes action A and transitions to the state q_j.

This can be problematic. For instance, we have not defined where to HALT yet, nor have we guaranteed that quadruples are unique.

Definition 14.1. If T is in the state q_i reading symbol S, we say that Q is *applicable*, where $Q = q_i S A q_j$.
T HALTS if no quadruple is applicable.
A set X of quadruples is *consistent* if whenever we have both $q_i S A q_j$ and $q_i S A' q_k$, we have that $A = A'$ and $q_j = q_k$.

Definition 14.2. A *Turing machine* is a consistent set of quadruples.

How to compute with a Turing machine:

- Input Convention: To input $n \in \mathbb{N}$, place $n+1$ 1s on the tape, then set the head to the leftmost 1 in the state q_0.
- Output Convention: If a computation halts, output the number of 1s printed on the tape, denoted $\phi_T(n)$.

Definition 14.3. A function $f : \mathbb{N} \to \mathbb{N}$ is *Turing computable* if $f = \phi_T$ for some Turing Machine T.

Example 14.4. The best example I have ever seen:
$S(n) = n + 1$ is Turing Computable when $T = \emptyset$.

15. Turing Machines (Continued)

15.1. Examples.

Example 15.1. Show $C_0(n) = 0$ is Turing computable.

$q_0\ 1\ 0\ q_1$ (Subroutine for deleting 1's)
$q_1\ 0\ R\ q_0$ (Move right in search of another 1 to apply subroutine)

Example 15.2. Show $sg(n) = \begin{cases} 0 & \text{if } n = 0 \\ 1 & \text{if } n \neq 0 \end{cases}$ is Turing computable.

$q_0\ 1\ R\ q_1$
$q_1\ 0\ L\ q_2$
$q_2\ 1\ 0\ q_3$
$q_3\ 0\ R\ q_2$
$q_1\ 1\ 0\ q_3$

How do we compute functions of $k > 1$ variables? We will use the following convention:
We associate with T a partial function of $k > 0$ variables, denoted $\phi_T^{(k)}$, by letting $\phi_T^{(k)}(\bar{n})$ be

the output obtained from inputting $n_1, ..., n_k$ into T in the form:

$$\underbrace{1\ldots1}_{n_1+1}\ 0\ \underbrace{1\ldots1}_{n_2+1}\ 0\ \ldots\ 0\ \underbrace{1\ldots1}_{n_k+1}$$

We will set the reading head over the leftmost 1 in state q_0.

Example 15.3. Find a Turing machine that computes $C_0^{(k)}(n1, ..., n_k) = 0$ for any $k > 0$.

$q_0\ 1\ 0\ q_1$
$q_1\ 0\ R\ q_0$ } erase current block of 1's

$q_0\ 0\ R\ q_2$ } hunt for new block of 1's

$q_2\ 1\ 1\ q_0$ } transition to erase subroutine

15.2. **Partial recursive functions vs. Turing computable functions.**

- Projection functions are Turing computable.
- If $f, g : \mathbb{N} \to \mathbb{N}$ are Turing computable, then $f \circ g$ is Turing computable.
- Turing computable functions are closed under primitive recursion and the μ operator.
- Hence the partial recursive functions are a subset of the Turing computable functions.
- Turing proved the reverse inclusion.

Turing's thesis:
f is effectively computable $\Leftrightarrow$ f is Turing computable
Church-Turing thesis:
f is effectively computable $\Leftrightarrow$ f is partially recursive $\Leftrightarrow$ f is Turing computable

15.3. **Universal Turing Machine**
First step: Obtain a computable list of Turing machines by Gödel numbering.

(1) Code tape symbols, action symbols, internal states as follows:

$gn(L) = 2,$
$gn(R) = 3,$
$gn(q_i) = p_{2+2i},$
$gn(s_i) = p_{2+2i+1}.$

(2) Code quadruples: if $Q = q_iSAq_j$, then let $gn(Q) = 2^{gn(q_i)} * 3^{gn(S)} * 5^{gn(A)} * 7^{gn(q_n)}$.

(3) Code Turing programs: if $P = \{Q_0, ..., Q_r\}$, then let $gn(P) = 2^{gn(Q_0)} * ... * p_r^{gn(Q_r)}$.

Note: A program can have multiple Gödel numbers.
Observation: gn^{-1} is computable.

Definition 15.4. The e^{th} Turing machine P_e is defined by

$$P_e = \begin{cases} P & \text{if } gn^{-1}(e) = \text{ some Turing program } P \\ \emptyset & \text{otherwise} \end{cases}$$

$\phi_e^{(k)}$ is the k-place partial function computed by P_e.
$\phi_e^{(1)} = \phi_e$ is the e^{th} partial computable function.
We say that P_e has index e.

Theorem 15.5 (Enumeration Theorem)**.** $\phi_z(x)$ is a partial computable function of x and z.

Proof. We compute $\phi_z(x)$ by the following procedure:

Step 1: Using z, find the Turing program P_z.
Step 2: Give P_z the input x and wait for this computation to terminate.
Step 3: If the computation terminates, set $\phi_z(x)$ equal to the output written on the tape; otherwise, $\phi_z(x)\uparrow$.

□

Theorem 15.6. There is no computable list $\{f_e\}_{e\geq 0}$ of all (total) computable functions such that $f_z(x)$ is a computable function of x and z.

Proof. Suppose there is such a computable list. We will show that there is a total computable function not on the list. Let

$$h(x) = f_x(x) + 1.$$

Clearly h is total, since each f_e is, and h is computable since it is defined in terms of the computable list $\{f_e\}_{e\geq 0}$. Thus $h = f_z$ for some $z \in \mathbb{N}$. It follows that $f_z(z) = h(z) = f_z(z)+1$, which is impossible. Thus, no such computable list exists. □

Lemma 15.7 (Padding Lemma)**.** If f is a partial computable function, then there are infinitely many i such that $f = \phi_i$.

Proof. Let P be a Turing program for f. Let q_n be a new state that does not appear in P, and let $P^* = \{q_n 1 R q_0\} \cup P$. Then P^* computes f, as the new quadruple is never used, but $gn(P) \neq gn(P^*)$. □

Theorem 15.8 (The *s-m-n* Theorem)**.** Let $m, n \geq 1$. There is a computable function $s_n^m : \mathbb{N}^{m+1} \to \mathbb{N}$ such that for all $x, y_1, \ldots, y_m, z_1, \ldots, z_n \in \mathbb{N}$,

$$\phi_x^{(n+m)}(y_1, \ldots, y_m, z_1, \ldots, z_n) = \phi_{s_n^m(x,y_1,\ldots,y_m)}^{(n)}(z_1, \ldots, z_n).$$

We will mostly use the *s-m-n* Theorem in the following form.

Corollary 15.9. If $f(x, y)$ is a partial computable function, then there is a computable function g such that

$$f(x, y) = \phi_{g(x)}(y).$$

Theorem 15.10 (Turing 1936)**.** universal Turing machine

Proof. . □

Theorem 15.11 (Kleene's Fixed Point Theorem)**.** .

Proof. . □

Computable approximations

16. Computable and Computably Enumerable Sets

Recall that a set $S \subseteq \mathbb{N}$ is computable if and only if its characteristic function χ_S is a computable function.

Example 16.1. $\{n : n = 2k \text{ for some } k\}$ is computable.
$\{p : p \text{ is prime }\}$ is computable.

Is $S = \{x \in \mathbb{N} : \text{there is a string of at least } x \text{ 7's in the decimal expansion of } \pi\}$ computable? Yes!

- Case 1: If $S = \{0, \dots, n\}$ for some $n \in \mathbb{N}$, then S is finite, and hence is computable.
- Case 2: $S = \mathbb{N}$ is computable since $\chi_S = c_1(x)$.

Is $S = \{x \in \mathbb{N} : \text{there is a string of } \textit{exactly } x \text{ 7's in the decimal expansion of } \pi\}$ computable? This is still unknown, since we do not know how the digits in π are distributed. However, π itself is a *computable real number*. That is, we can write π as,

$$\pi = 4 - \frac{4}{3} + \frac{4}{5} - \frac{4}{7} + \cdots$$

We can check the expansion of π and produce a list $x_1, x_2, \dots$ for blocks of length 7 (not necessarily in order). Thus we can *computably enumerate* S. That is to say, we cannot compute χ_S, but we can list the elements in S.

Definition 16.2. $A \subseteq \mathbb{N}$ is *computably enumerable (c.e.)* if and only if $A = \emptyset$ or there exists a computable function f such that $A = \{f(0), f(1), \dots\}$.

Theorem 16.3. If $A \subseteq \mathbb{N}$ is computable, then A is c.e.

Proof. If $A = \emptyset$, there is nothing to show.
If $a \in A$, then

$$f(n) = \begin{cases} n & \text{if } n \in A \\ a & \text{if } n \notin A \end{cases}$$

$\therefore A = \text{range}(f)$ where f is computable. □

Theorem 16.4 (Complementation Theorem)**.** $A \subseteq \mathbb{N}$ is computable $\iff$ A, $\overline{A}$ are c.e.

Proof. ($\Rightarrow$): If A is computable, it's c.e. by the previous theorem. The characteristic function for $\overline{A}$ is $\chi_{\overline{A}} = 1 - \chi_A$. So $\overline{A}$ is computable and thus c.e.
($\Leftarrow$): Let $A = \text{range}(f)$ and $\overline{A} = \text{range}(g)$. We can check both functions in parallel for each $n \in \mathbb{N}$.

$$f(0), f(1), f(2), \dots$$
$$g(0), g(1), g(2), \dots$$

Every $n \in \mathbb{N}$ is on one of these lists. We run both of these functions in parallel until we find some $y(n)$ such that $f(y(n)) = n$.
Thus $f(y(n)) = n$ if and only if $n \in A$. □

Proposition 16.5. f is increasing if $f(x) < f(x+1)$ for all $x \in \mathbb{N}$. An infinite $A \subseteq \mathbb{N}$ is computable if and only if it is the range of an increasing computable function.

Proof. ($\Rightarrow$): If A is computable, let

$$f(0) = \mu y[y \in A]$$
$$f(n+1) = \mu y[y \in A \text{ and } y > f(n)]$$

($\Leftarrow$): If $a = \text{range}(f)$ where f is an increasing computable function, we see

$$\chi_A(x) = \begin{cases} 1 & \text{if } \exists y \leq x \text{ s.t. } f(y) = x \\ 0 & \text{otherwise} \end{cases}$$

$$x \in A \iff \exists y \leq x \text{ s.t. } f(y) = x$$

□

Note that if $P(x, y)$ is computable, then $\exists x \leq z P(x, y)$ is computable, since primitive recursive functions are closed under bounded quantifiers. That is, prim. rec. functions are closed under $\exists x \leq n$ and $\forall x \leq n$.

Proposition 16.6. Every infinite c.e. set has an infinite computable subset.

Proof. Let A be infinite c.e., where $A = \text{range}(f)$. Define

$$g(0) = f(0)$$
$$g(n+1) = \mu y[\exists m f(m) \wedge \forall i \leq n f(m) = g(i)]$$

So g is computable and increasing.
$\therefore B = \text{range}(g) \subseteq \text{range}(f) = A$ is computable by the previous proposition. □

Notice that here, range(f) is infinite, so f has an increasing subsequence. In this case, it's g. As g is bounded below by 0, range(g) is infinite.

Proposition 16.7. If A, B are c.e., then so is $A \cup B$.

Proof. Let $A = \text{range}(f)$ and $B = \text{range}(g)$. Let k be defined as follows:

$$k(x) = \begin{cases} f(n) & \text{if } x = 2n \\ g(n) & \text{if } x = 2n+1 \end{cases}$$

Thus $A \cup B = \text{range}(k)$. □

16.1. **Normal Form Theorem.**

Observe:

$$P_e \text{ halts on input } x \iff \phi_e(x) \downarrow \iff x \in dom(\phi_e)$$

Definition 16.8 (Halting Set and Halting Problem)**.** For each $e \in \mathbb{N}$, $W_e = dom(\phi_e)$ is the *Halting Set* for P_e. If W_e is computable, we say that the *Halting Problem* for P_e is *solvable*.

Notice:

$$W_{e,s} = dom(\phi_{e,s}) = \{x : \phi_{e,s}(x)\downarrow\}$$
$$\phi_{e,s} \text{ only halts if } x < s. \text{ Thus, } W_{e,s} \text{ is finite.}$$
$$W_e = \cup_{s\in\mathbb{N}} W_{e,s}$$

Finally,

$$\begin{aligned} x \in W_e \iff x \in dom(\phi_e) &\iff \phi_e(x)\downarrow \\ &\iff \exists s \phi_{e,s}(x)\downarrow \\ &\iff \exists s x \in W_{e,s} \\ &\iff x \in \cup_{s\in\mathbb{N}} W_{e,s} \end{aligned}$$

17. Computably enumerable (c.e.) sets

Definition 17.1. $W_{e,s} = \text{dom}(\varphi_{e,s})$ for $e, s \in \mathbb{N}$.

Claim 17.2. $W_{e,s}$ is finite (and hence computable) and $W_e = \bigcup_{s\in\mathbb{N}} W_{e,s}$.

Proof.

$$x \in W_e \Leftrightarrow x \in \text{dom}(\varphi_e) \Leftrightarrow \exists s(\varphi_{e,s}(x)\downarrow) \Leftrightarrow \exists s(x \in \text{dom}(\varphi_{e,s})) \Leftrightarrow x \in \bigcup_{s\in\mathbb{N}} W_{e,s}.$$

□

Proposition 17.3. $A \subseteq \mathbb{N}$ is $\Sigma_1^0 \Leftrightarrow \overline{A}$ is Π_1^0

Proof.

$$A \in \Sigma_1^0 \Leftrightarrow A = \{x : \exists y(R(x,y))\}\ R \text{ comp.} \Leftrightarrow \overline{A} = \{x : \forall y(\neg R(x,y)\}\ \neg R \text{ comp.} \Leftrightarrow \overline{A} \in \Pi_1^0.$$

□

$S = \{x : \exists y_1 \exists y_2 ... \exists y_n (R(x, y_1, y_2, ..., y_n))\}$. What is the complexity of this set?

$$\begin{aligned} S &= \{x : \exists z(z = \langle y_1, ..., y_n\rangle) \wedge R(x, y_1, ..., y_n)\} \\ &= \{x : \exists z(R(x, \pi_1^n(z), \pi_2^n(z)., \pi_n^n(z)))\} \\ &= \{x : \exists z(R(x, \pi_1(z), \pi_2(\pi_1(z), ..., \pi_2(\pi_2(...\pi_2(\pi_1(z))))))\} \\ &= \{x : \exists z(R(x, \pi_1(z), \pi_1 \circ \pi_2(z), ..., \pi_1 \circ \pi_2^{(n-1)}(z)))\} \end{aligned}$$

Thus S is Σ_1^0. Similarly, $\{\overline{x} : \exists \overline{y}(R(\overline{x}, \overline{y})\}$ is Σ_1^0.

We can define Σ_1^0 and Π_1^0 subsets of $\mathbb{N}$.

Theorem 17.4 (Normal Form Theorem for C.E. Sets)**.** The following are equivalent:

(1) A is c.e.
(2) A is Σ_1^0.
(3) $A = W_e$ for some $e \in \mathbb{N}$.

Proof.

(1) $\Rightarrow$ (2): Assume A is c.e. If $A = \emptyset$, then $x \in A \Leftrightarrow \exists s(s = x + 1)$. If $A \neq \emptyset$, then $A = \text{range}(f)$ for some computable f. Then $x \in A \Leftrightarrow \exists y(f(y) = x)$, which is a computable relation because f is total. Thus, A is Σ_1^0.

(2) $\Rightarrow$ (3): Assume A is Σ_1^0. Then $x \in A \Leftrightarrow \exists y(R(x,y))$ with R computable. Let

$$\psi(x) = \begin{cases} 0 & \exists y(R(x,y)) \\ \uparrow & \text{otherwise.} \end{cases}.$$

We see that ψ is partial computable, so $\psi = \varphi_e$ for some $e \in \mathbb{N}$. Then

$$x \in A \Leftrightarrow \exists y(R(x,y)) \Leftrightarrow \psi(x)\downarrow \Leftrightarrow \varphi_e(x)\downarrow .$$

Thus, $A = \text{dom}(\varphi_e) = W_e$ for some $e \in \mathbb{N}$.

(3) $\Rightarrow$ (1): Assume $A = W_e$ for some $e \in \mathbb{N}$. If $W_e = \emptyset$, A is c.e. If $W_e \neq \emptyset$, $\exists p \in A$. Let

$$f(\langle x,s\rangle) = \begin{cases} x & \varphi_{e,s}(x)\downarrow \\ p & \text{otherwise.} \end{cases}$$

Then f is a total computable function.

$$x \in A \Leftrightarrow \varphi_e(x)\downarrow \Leftrightarrow \exists s(\varphi_{e,s}(x)\downarrow) \Leftrightarrow f(\langle x,s\rangle)\downarrow = x \text{ for some } s \Leftrightarrow x \in \text{range}(f).$$

Thus, A is .c.e.

□

We have shown that if $R(x,y)$ and $S(x,y)$ are computable, $R(x,y) \wedge S(x,y)$, $\neg R(x,y)$, $R(x,y) \vee S(x,y)$ are computable. If R is computable, we know that $\{x : \exists(R(x,y)\}$ is computably enumerable. In general, is it computable? No.

Theorem 17.5. There is a c.e. set that is not computable.

Proof. Let $K = \{e : e \in W_e\} = \{e : \varphi_e(e)\downarrow\}$. K is Σ_1^0:

$$e \in K \Leftrightarrow \varphi_e(e)\downarrow \Leftrightarrow \exists s \varphi_{e,s}(e)\downarrow .$$

Hence, K is c.e. To show K is not computable, we show that $\overline{K}$ is not c.e.

Suppose $\overline{K}$ is c.e. Then $\overline{K} = W_e$ for some $e \in \mathbb{N}$.

$$e \in W_e \Leftrightarrow e \in \overline{K} \Leftrightarrow e \notin K \Leftrightarrow e \notin W_e. \perp$$

□

Corollary 17.6. There exists a Turing machine with an unsolvable halting problem.

Proof. If K is c.e., then $K = W_e$ for some e. W_e is not computable, so the halting problem for P_e is not solvable. □

Corollary 17.7. The halting problem for the universal Turing machine is unsolvable.

Proof. Let P_e be as above. This means $x \in W_e$ if and only if U halts on input (e,x). The solvability of the halting problem for U implies the solvability of the halting problem for P_e. □

18. Creative sets

Definition 18.1. $A \subseteq \mathbb{N}$ is creative if and only if

(1) A is c.e. and
(2) there is a computable function f such that for each $e \in \mathbb{N}$

$$W_e \subseteq \overline{A} \Rightarrow f(e) \in \overline{A} \setminus W_e.$$

We call f the creative function for A.

Observe: If A is c.e. but not computable, then for all $e \in \overline{A} \neq W_e$. If A is creative, there is a computable function that witnesses this: If $W_e \subseteq \overline{A}$, $f(e)$ produces some $y \in \overline{A} \setminus W_e$ which witnesses that $W_e \neq \overline{A}$.

Theorem 18.2. Creative sets exist.

Proof. We claim that K is creative.

(1) K is c.e., so $f(e) = e$.
(2) The creative function for K is the identity function: If $W_e \subseteq \overline{K}$, then $e \notin W_e$, since $e \in W_e \Rightarrow W_e \cap K \neq \emptyset$. So $e \notin W_e \Rightarrow e \in \overline{K}$. Hence $e \in \overline{K} \setminus W_e$.

□

Proposition 18.3. Let C be creative. There is an algorithm such that, given n members of $\overline{C}$ ($n \geq 0$), yields $n+1$ members of $\overline{C}$. Thus $\overline{C}$ contains an infinite c.e. subset.

Proof. Let $y_1, ..., y_n$ be a (possibly empty) list of members of $\overline{C}$, and let f be the creative function of C. Observe: Given a finite set of natural numbers $X = \{n_0, ..., n_k\}$ computably in $n_0, ..., n_k$, we can find the index of a Turing machine T such that $\mathrm{dom}(\varphi_T) = X$. Hence, we can effectively find an index i such that $X = W_i$.

Let $W_i = \{y_1, ..., y_n\} \subseteq \overline{C}$. Since C is creative, $f(i) \in \overline{C} \setminus W_i$. Let $y_{n+1} = f(i)$. Hence $\{y_1, ..., y_n, y_{n+1}(= f(i)\} \subseteq \overline{C}$.

To obtain an infinite c.e. subset of $\overline{C}$, let $W_{i_0} = \emptyset$ and enumerate i_0 into A. Now apply the above algorithm to $W_{i_1} = \{i_0\}$, $W_{i_2} = \{i_0, f(i_1)\}$, $W_{i_3} = \{i_0, f(i_1), f(i_2)\}$, etc., and $W_{i_n} \subseteq \overline{C}$ for $n \in \omega$. Let $S = \bigcup_{n \in \omega} W_{i_n}$. $S = \mathrm{g}$, where $g(0) = i_0$, $g(n+1) = f(i_{n+1})$. Hence S is infinite and c.e. □

19. Construction of a simple set

Let $g(n) = f(i_n)$. Then g enumerates S, and g is computable. Key: the sequence $(i_n)_{n \in \mathbb{N}}$ is computable.

Theorem 19.1 (Enumeration Theorem for C.E. Sets)**.** There is a c.e. set K_0 such that, for each $e \in \mathbb{N}$, $W_e = \{x : \langle x, e\rangle \in K_0\}$.

Proof. Let $K_0 = \{\langle x, e\rangle \, x \in W_e\}$. Why is K_0 c.e.?

$$\langle x, e\rangle \in K_0 \Leftrightarrow x \in W_e \Leftrightarrow \exists s(x \in W_{e,s})$$

which is a computable relation. So K_0 is Σ^0_1 by previous theorem. So by NFT, K_0 is c.e. □

Definition 19.2. A set S is simple if

(1) S is c.e.;

(2) $\overline{S}$ is infinite; and
(3) $|W_e| = \infty \Rightarrow S \cap W_e \neq \emptyset$.

$\overline{S}$ contains no infinite c.e. set. If S is simple, then S is not computable.

Theorem 19.3. There exists a simple set.

Proof. We computably enumerate S in stages 0,1,2,3,..., to satisfy the following requirements: For each $e \in \mathbb{N}$,

- $\mathcal{N}_e : |S \cap \{0, ..., 2e\}| \leq e$
- $\mathcal{P}_e : |W_e| = \infty \Rightarrow S \cap W_e \neq \emptyset$

If $\mathcal{N}_e$ is satisfied so that $|S \cap \{0, ..., 2e\}| \leq e$, then $|\overline{S} \cap \{0, ..., 2e\} > e$. If we satisfy $\mathcal{N}_e$ for infinitely many e, then $|\overline{S}| \geq |\overline{S} \cap \{0, ..., 2e\}| > e$.

Algorithm for enumerating S

(1) For each requirement $\mathcal{P}$ that is unsatisfied, wait for a stage s at which there is some number $x \in W_{e,s}$ with $x > 2e$.
(2) If such an x appears, enumerate x into S. At this stage, $\mathcal{P}_e$ becomes satisfied.

End of construction.

Verify that S is c.e. If W_e is infinite, let s be the least stage such that $x \in W_{e,s}$ with $x > 2e$. By construction, if $\mathcal{P}_e$ is not already satisfied, it becomes satisfied by part (2) of the Algorithm, so $S \cap W_e \neq \emptyset$. □

Lemma 19.4. $\mathcal{N}$ is satisfied for each $e \in \mathbb{N}$ into S.

Proof. We only enumerate $x \leq 2e$ on behalf of some W_i with $i \leq e$. For each such i, at most one $x \leq 2e$, so the lemma follows. □

20. Consequences of simple sets

Proposition 20.1.

(1) If A is simple and W is an infinite c.e. set, then $A \cap W$ is infinite and c.e.
(2) If A, B are simple, so is $A \cap B$.

($A \cap W \neq \emptyset$ because A is simple and W is infinite and c.e.)

Proof.

(1) A, W as above, Suppose that $|A \cap W| = n$. Consider $W \setminus (A \cap W)$. $W \setminus (A \cap W)$ is infinite and c.e. But $W \setminus (A \cap W) \subseteq \overline{A}$. $\perp$
(2) Suppose A, B are simple. $A \cap B$ is c.e. Let W be infinite and c.e. By (1), $A \cap W$ is infinite and c.e. By simplicity of B, $(A \cap W) \cap B \neq \emptyset$. Hence $(A \cap B) \cap W \neq \emptyset$. Lastly, $\overline{A \cap B}$ is infinite because $A \cap B \subseteq A \Rightarrow \overline{A \cap B} \supseteq \overline{A}$, and $\overline{A}$ is infinite because A is simple.

□

Proposition 20.2.

(1) Assuming the existence of a simple set, there is a simple set $S \supseteq \{2n : n \in \omega\}$.
(2) The collection of simple sets is not closed under unions.

Proof. Let A be simple. Define $S = \{2n : n \in W\} \cup \{2k + 1 : k \in A\}$. S is c.e. Given W_e infinite and c.e., consider the following two cases:

- Case 1: $\exists 2n \in W_e$ for some n. Done.
- Case 2: $\neq \exists 2n \in W_e$. Let $\hat{W}_e = \{k : 2k+1 \in W_e\}$, which is infinite and c.e. So $\exists k \in A_= \cap \hat{W}_e$ and so $2k+1 \in S \cap W_e$. $\overline{S}$ is infinite because $\overline{S} = \{2k+1 : k \in \overline{A}\}$ is infinite.

□

21. Many-one Reducibility

Definition 21.1 (Many-one reducibility)**.**
For $A, B \subseteq \mathbb{N}$, A is m-reducible to B, written $A \leq_m B$, if there exists a computable function f such that $\forall x \in \mathbb{N}, x \in A \iff f(x) \in B$.

Clearly, $A \leq_m B \implies \bar{A} \leq_m \bar{B}$.

Proposition 21.2. $\leq_m$ is reflexive and transitive.

Proof.

- $A \leq_m A$ via $f(x) = x$.
- If $A \leq_m B$ via f and $B \leq_m C$ via g, then $A \leq_m C$ via $g \circ f$, i.e. $x \in A \iff f(x) \in B \iff g(f(x)) \in C$

□

Theorem 21.3. (1) If $A \leq_m B$ and B is computable, then A is computable.
(2) If $A \leq_m B$ and B is computably enumerable, then A is computably enumerable.

Proof. (1) If $A \leq_m B$ via f, then $\chi_A = \chi_B \circ f$. Then $x \in A \iff f(x) \in B$.
(2) If $A \leq_m B$ via f and B is computably enumerable, then by the normal form theorem, $x \in B \iff \exists y R(x, y)$, R a computable relation. Hence $x \in A \iff \exists y R(f(x), y)$. Let $R(x, y)$ hold $\iff R(f(x), y)$ holds.

□

$k_0 = \{< x, y >: x \in W_y\}$ and $k = \{e : \phi_e(e) \downarrow\}$

Proposition 21.4. $k \leq_m k_0$
Corollary: K_0 is not computable.

Proof. We want to find a computable function f such that $x \in K \iff f(x) \in K_0$. Let $f(x) =< x, x >$. Then $x \in K \iff \phi_x(x) \downarrow \iff < x, x > \in K_0$. □

Proposition 21.5. Let A be computably enumerable. Then $A \leq_m K_0$.

Proof. $A = W_e$ for some e. Let $f(x) =< x, e >$. The $x \in A \iff x \in W_e \iff < x, e > \in K_0$. □

Definition 21.6 (m-complete)**.**
If B is computably enumerable and $A \leq_m B$ for all computably enumerable A, then B is m-complete.

Corollary: the relation $\phi_e(y) \downarrow$ on x,y is not computable.

Definition 21.7 (1-reducibility)**.**
A is 1-reducible to B, written $A \leq_1 B$, if $\exists f$ computable and 1-1 such that $\forall x x \in A \iff f(x) \in B$.

- A is computably enumerable $\implies A \leq_1 K_0(f(x) =< x, e >$ is 1-1)
- $A \leq_1 B \implies A \leq_m B$

Theorem 21.8. There are computably enumerable sets A,B such that $A \leq_m B$, but $A \not\leq_1 B$.

Proof. Let $A = \{2k : k \in \mathbb{N}\}$, $B \supseteq A$ simple.
Claim: $A \leq_m B$.
Let $b \notin B$. Define $f(x) = \begin{cases} x & \text{if } x \text{ is even} \\ b & \text{if } x \text{ is odd} \end{cases}$
Suppose $A \leq_1 B$ via g. Then $x \in A \iff g(x) \in B$. Then x is odd $\iff g(x) \notin B$. Let h(n) = g(2n+1) and let $S = \{h(0), h(1), ...\} = \{g(1), g(3), ...\} \subseteq \bar{B}$. This is infinite because g is one to one, and S is computably enumerable because S = range(h).

□

22. Index sets

Definition 22.1. Let $\mathbb{A}$ be a collection of computably enumerable sets (or partially computable functions). The index set of $\mathbb{A}$ is the set $A \subseteq \mathbb{N}$ of all indices of members of $\mathbb{A}$.

If A is an index set of some set $\mathbb{A}$ of computably enumerable sets, then A satisfies: $x \in A$ and $\phi_x = \phi_y$ and $W_x = W_y \implies y \in A$.

Example 22.2.
- $K_1 = \{e : W_e \neq \emptyset\}$
- $Tot = \{e : \phi_e \text{ is total}\}$
- $Fin = \{e : W_e \text{ is finite}\}$
- $Inf = \{e : W_e \text{ is infinite}\}$
- $Cof = \{e : W_e \text{ is cofinite}\}$

Theorem 22.3. K is not an index set.

Proof. We want to define a computable function f such that: $\{n\} = W_{f(n)}$. Let

$$\psi(x, y) = \begin{cases} 1 & \text{if } x = y \\ \uparrow & \text{otherwise} \end{cases}$$

By the S-M-N theorem, there exists a computable function f such that $\phi_{f(x)}(x) = \psi(x, y)$. For $n \in \mathbb{N}, dom(\phi_{f(n)}) = W_{f(n)} = \{n\}$. Since f is computable, by the fixed point theorem, there exists e such that $\{e\} = W_{f(e)} = W_e$. So $W_e = \{e\}$. $e \in W_e \implies e \in K$. By padding, let e' satisfy $W_{e'} = W_e$ and $e' \neq e$. Then notice $e' \notin W_{e'} = W_e = \{e\}$. So $e \in W_e \implies e \in K, W_e = W_{e'}$, but $e' \notin K$. □

Theorem 22.4 (Rice's theorem)**.** If A is an index set and $A \neq \emptyset, \mathbb{N}$, then either $K \leq_m A$ or $\bar{K} \leq_m A$. Thus A is not computable.

Proof. Let A be as above and choose $e \in A$ and $e' \notin A$. Then we have two cases to consider:
Case 1: $\emptyset$ has no index in A. Then define

$$\psi(x, y) = \begin{cases} \phi_e(y) & \text{if } x \in K \\ \uparrow & \text{if } x \notin K \end{cases}$$

Then by the S-M-N theorem, there exists a computable function f such that $\psi(x,y) = \phi_{f(x)}(y)$. Then

$$W_{f(x)} = \begin{cases} W_e & \text{if } x \in K \\ \emptyset & \text{if } x \notin K \end{cases}$$

Then $x \in K \implies W_{f(x)} = W_e \implies f(x) \in A$
and $x \notin K \implies W_{f(x)} = \emptyset \implies f(x) \notin A$.
Then $K \leq_m A$ via f.
Case 2: $\emptyset$ has no index in $\bar{A}$. This case is the same as case 1, but replace e with e' and A with $\bar{A}$. □

23. Many-one degrees

Definition 23.1. For $A, B \subseteq \mathbb{N}$, A is m-reducible to B, written $A \leq_m B$, if there is a computable f such that for all $x \in \mathbb{N}$, $x \in A \Leftrightarrow f(x) \in B$.

Clearly, $A \leq_m B$ implies $\overline{A} \leq_m \overline{B}$.

Proposition 23.2. $\leq_m$ is reflexive and transitive.

Proof. $A \leq_m A$ via $f(x) = x$.
If $A \leq_m B$ via f and $B \leq_m C$ via g, then $A \leq_m C$ via $g \circ f$.

$$x \in A \Leftrightarrow f(x) \in B \Leftrightarrow g(f(x)) \in C.$$

□

Theorem 23.3.

(1) If $A \leq_m B$ and B is computable, then A is computable.
(2) If $A \leq_m B$ and B is c.e., then A is c.e.

Proof.

(1) $A \leq_m B$ via f, then $\chi_A = \chi_B \circ f$ and $x \in A \Leftrightarrow f(x) \in B$.
(2) If $A \leq_m B$ via f and B is c.e., then by NFT:

$$x \in B \Leftrightarrow \exists y(R(x,y)) \text{ computable } R$$

$$x \in A \Leftrightarrow \exists y(R(f(x),y).$$

Let $\hat{R}(x,y)$ hold $\Leftrightarrow R(f(x),y)$ holds.

□

$K_0 = \{\langle x,y\rangle : x \in W_y\}$ $\varphi(y(x)\downarrow$

Proposition 23.4. $K \leq_m K_0$

Proof. Want f computable such that $x \in K \Leftrightarrow f(x) \in K_0$. Let $f(x) = \langle x,x\rangle$, so

$$x \in K \Leftrightarrow \varphi_x(x)\downarrow \Leftrightarrow \langle x,x\rangle \in K_0.$$

□

Corollary 23.5. K is not computable.

$K = \{e : \varphi_e(e)\downarrow\}$ c.e., noncomputable.

Proposition 23.6. Let A be c.e. Then $A \leq_m K_0$.

Proof. $A = W_e$ for some e. Let $f(x) = \langle x, e \rangle$.

$$x \in A \Leftrightarrow x \in W_e \Leftrightarrow \langle x, e \rangle \in K_0.$$

□

Corollary 23.7. The relation $\varphi_y(x) \downarrow$ on x, y is not computable.

Definition 23.8. A is 1-reducible to B, written $A \leq_1 B$, if there is a computable one-to-one f such that for all x, $x \in A \Leftrightarrow f(x) \in B$.

A c.e. implies that $A \leq_m B$. Is the converse true?

Theorem 23.9. There are c.e. sets A, B such that $A \leq_m B$ but $A \not\leq_1 B$.

Proof. Let $A = \{2n : n \in \mathbb{N}\}$ and let $B \supset A$ be simple. We claim that $A \leq_m B$. Let $b \notin B$. Define

$$f(x) = \begin{cases} x & x \text{ even} \\ b & x \text{ odd.} \end{cases}$$

Suppose $A \leq_1 B$ via g. Then $x \in A \Leftrightarrow g(x) \in B$ and x is odd $\Leftrightarrow g(x) \notin B$. Let $h(n) = g(2n+1)$. Let $S = \{h(0), ...\} = \{g(1), ...\} \subseteq \overline{B}$. This set is infinite because g is one-to-one. Also, S is c.e. because $S = \text{range}(h)$. $\perp$ □

.

24. Index sets

Definition 24.1. Let $\mathcal{A}$ be a collection of c.e. sets (or p.c.f.s). The index set of $\mathcal{A}$ is the set $A \in \mathbb{N}$ or all indices of members of $\mathcal{A}$. If A is an index set of some set $\mathcal{A}$ of c.e. sets, then A satisfies $x \in A$ and $\varphi_x = \varphi_y \Rightarrow y \in A$.

Examples 24.2.

- $K_1 = \{e : W_e \neq \emptyset\}$
- Tot $= \{e : \varphi_e \text{ is total}\}$
- Fin $= \{e : W_e \text{ is finite}\}$
- Inf $= \{e : W_e \text{ is infinite}\}$
- Cof $= \{e : W_e \text{ is cofinite}\}$

Theorem 24.3. K is not an index set.

Proof. Want to define a computable function f such that $\{n\} = W_{f(n)}$. Let

$$\psi(x, y) = \begin{cases} 1 & x = y \\ \uparrow & x \neq y. \end{cases}$$

By S-m-n Theorem, there is a computable f such that $\varphi_{f(n)}(y) = \psi(x, y)$. For $n \in \mathbb{N}$, $\text{dom}(\varphi_{f(n)}) = W_{f(n)} = \{n\}$. Now we want some e such that $e \in K$ and $W_e = W_{e'}$, but $e' \notin K$. We see that f is computable, so by the Fixed Point Theorem, there is some e such that $\{e\} = W_{f(e)} = W_e$. By the Padding Lemma, $e \in W_e$. Let e' satisfy $e \in K$. $W_{e'} = W_e$ and $e' \neq e$. Note: $e' \notin W_{e'} = W_e = \{e\}$, so $e \in W_e \Leftrightarrow e \in K$, $W_e = W_{e'}$ but $e' \notin K$. □

Theorem 24.4 (Rice's Theorem)**.** If A is an index set and $A \neq \emptyset, \mathbb{N}$, then $K \leq_m A$ or $\overline{K} \leq_m A$. Thus A is not computable.

Proof. Let A be as above and choose $e \in A$ and $e' \notin A$. We have two cases to consider.

- Case 1: φ has no index set in A. Define

$$\psi(x,y) = \begin{cases} \varphi_e(y) & x \in K \\ \uparrow & x \notin K. \end{cases}$$

  By S-m-n Theorem, there is a computable f such that $\psi(x,y) = \varphi_{f(x)}(x)$.

$$W_{f(x)} = \begin{cases} W_e & x \in K \\ \emptyset & x \notin K. \end{cases}$$

$$x \in K \Rightarrow W_{f(x)} = W_e \Rightarrow f(x) \in A$$
$$x \notin K \Rightarrow W_{f(x)} = \emptyset \Rightarrow f(x) \notin A$$

  $K \leq_m A$ via f.
- Case 2: φ has no index set in $\overline{A}$. Replace e with e' and A with $\overline{A}$.

□

Theorem 24.5 (Kleene's Fixed Point Theorem with Parameters)**.**

Proof. Let $f(x, \bar{y})$ be computable. There is a computable function $k(\bar{y})$ such that

$$\phi_{f(k(\bar{y}),\bar{y})} = \phi_{k(\bar{y})}$$

Now define

$$\psi(x,\bar{y},z) = \begin{cases} \phi_{\phi_x(x,\bar{y})}(z) & \text{if } \phi_x(x,\bar{y}) \downarrow, \\ \uparrow & \text{otherwise.} \end{cases}$$

By the S-M-N Theorem, there exists a computable function $h(x, \bar{y})$ such that

$$\phi_{h(x,\bar{y})}(z) = \psi(x,\bar{y},z) = \phi_{\phi_x(x,\bar{y})}(z)$$

Consider the computable function $f(h(x,\bar{y}),\bar{y})$. Thus, $f(h(x,\bar{y}),\bar{y}) = \phi_e(x,\bar{y})$ for some $e \in \mathbb{N}$. Let x = e.

$$\phi_{f(h(e,\bar{y}),\bar{y})} = \phi_{\phi_e(e,\bar{y})} = \phi_{h(e,\bar{y})}$$

Let $k(\bar{y}) = h(e,\bar{y})$.

□

Theorem 24.6. All creative sets are m-complete.

Proof. Let C be creative with creative function f and let A be any c.e. set. Define

$$\psi(x,y,z) = \begin{cases} z & \text{if } y \in A \text{and } f(x) = z, \\ \uparrow & \text{otherwise.} \end{cases}$$

Use S-M-N Theorem to obtain the computable function g such that

$$\phi_{g(x,y)}(z) = \psi(x,y,z)$$

Consider

$$W_{g(x,y)} = \begin{cases} \{f(x)\} & \text{if } y \in A, \\ \emptyset & \text{if } y \notin A. \end{cases}$$

By Kleene's Fixed Point Theorem with Parameters, there exists a computable function k(y) such that

$$W_{k(y)} = W_{g(k(y),y)}$$

Let $y \in A$. Then,

$$y \in A \implies \forall x f(x) \in W_{g(x,y)} \implies f(k(y)) \in W_{g(k(y),y)} = W_{k(y)} \implies W_{k(y)} \not\subseteq \overline{C} \implies f(k(y)) \in C$$

Let $y \notin A$. Then,

$$y \notin A \implies \forall x W_{g(x,y)} = \emptyset \implies W_{(g(k(y),y)} = W_{k(y)} = \emptyset \implies f(k(y)) \in \overline{C}$$

Therefore, $A \leq_m C$ via $f \circ k$. □

Corollary 24.7. The set of all creative sets is $0'_m$.

25. Gödel's Incompleteness Theorem

25.1. **Peano Arithmetic.** The language of Peano Arithmetic is $\mathcal{L}_{PA} = \{+, \cdot, S, 0\}$

(1) $S(x) \neq 0$
(2) $S(x) = S(y) \implies x = y$
(3) $x + 0 = x$
(4) $x + S(y) = S(x + y)$
(5) $x \cdot 0 = 0$
(6) $x \cdot S(y) = x \cdot y + x$
(7) If ϕ is a $\mathcal{L}_{PA}$-formula, then $(\phi(0)\ \&\ \forall x(\phi(x) \implies \phi(S(x))) \implies \forall \phi(x)$.

Definition 25.1. (1) $\mathcal{R} \subseteq \mathbb{N}^k$ is representable in PA if there exists an $\mathcal{L}_{PA}$-formula $\phi(x_1, ..., x_k)$ such that

$$\mathcal{R}(\bar{m}) \text{ holds} \iff \text{PA} \models \phi(\bar{m})$$
$$\neg\mathcal{R}(\bar{m}) \text{ holds} \iff \text{PA} \models \neg\phi(\bar{m})$$

(2) A k-place function f is representable in PA if graph(f) = {(x,y) : f(x)=y} is representable.
(3) A set S $\subseteq \mathbb{N}$ is representable in PA if the relation "$x \in S$" is representable in PA.

Fact 25.2. (1) $\mathcal{R}$ is representable in PA, $\neg\mathcal{R}$ is representable in PA.
(2) $\mathcal{P}$,$\mathcal{Q}$ representable in PA, then $\mathcal{P}\&\mathcal{Q}$ and $\mathcal{P} \vee \mathcal{Q}$ are representable in PA.
(3) S,T $\subseteq \mathbb{N}$ are representable in PA, S$\cap$T, S$\cup$T, $\neg$S, and $\neg$T are representable in PA.
(4) S $\subseteq \mathbb{N}$ is representable in PA $\iff$ X_S is representable in PA.

Theorem 25.3. All total computable functions are representable in PA. In particular, computable sets are representable in PA.

Q: What can we say about c.e. sets in PA? Or any first-order theory?

25.2. **Gödel Numbering.** Assign Gödel numbers to PA just as we assigned them to Turing Machines.

Example 25.4. Let $\psi = S_0, S_1, ..., S_n$ symbols. Then $\text{gn}(\psi) = 2^{gn(S_0)} \cdot 3^{gn(S_1)} \cdot p_n^{gn(S_N)}$. If $\overline{\psi}$ = ψ_1, ψ_2, ..., ψ_n, then $\text{gn}(\overline{\psi}) = 2^{gn(\psi_0)} \cdot 3^{gn(\psi_1)} \cdot p_n^{gn\psi_N)}$

We have the following computable relations:

- Form(m) $\iff$ gn^{-1}(m) is an $\mathcal{L}_{PA}$-formula
- Ax(m) $\iff$ gn^{-1}(m) is an axiom of PA
- MP(m,n,p) $\iff$ Form(m) & Form(n) & Form(p) & gn^{-1}(p) is derived from gn^{-1}(m) and gn^{-1}(n) via modus ponens
- Gen(m,n) $\iff$ Form(m) & Form(n) and gn^{-1}(n) is derived from gn^{-1}(m) via generalization ($\phi \vdash \forall x \phi$)
- Proof(m) $\iff$ gn^{-1} is a proof in PA

Definition 25.5.
- length(m) is largest i such that $p_i | m$.
- $(m)_i$ is the exponent of p_i in the prime factorization of m.

Definition 25.6 (computably axiomatizable)**.** A first order theory Γ is computably axiomatizable if it has a set of axioms, the Gödel numbers of which form a computable set.

Theorem 25.7. The set of T_{PA} of all Gödel numbers of the theorems of PA is computably enumerable.

Proof. $m \in T_{PA} \iff$ $\exists$p proof(p) and $p_{length(p)} = m$. Then, given a computably axiomatizable first order theory with computable rules of deduction, we can Gödel number our sentences and the above relations. □

Theorem 25.8. If Γ is a computably axiomatizable first order theory, then T_Γ is computably enumerable.

Theorem 25.9. Suppose $\Gamma$is a first order theory such that T_Γ is computably enumerable. Then Γ is computably axiomatizable.

Proof. T_Γ is computably enumerable. Let $T_\Gamma = range(f)$. Then let $\phi_n = gn^{-1}(f(n))$. Finally, let

- $\psi_0 = \phi_0$
- $\psi_1 = \phi_1 \cap \phi_1$

... □

Definition 25.10. (1) A $\mathcal{L}$-theory T is consistent if for each $\mathcal{L}$-sentencee ϕ, $\neg(T \vdash \phi \cap T \vdash \neg\phi)$

(2) An $\mathcal{L}_{PA}$-theory is ω-consistent if for each $\mathcal{L}_{PA}$ sentence ϕ, if $T \vdash \exists x \neg\phi(x)$, then $T \nvdash \phi(m)$ for some $m \in \mathbb{N}$.

Proposition 25.11. For an $\mathcal{L}_{PA}$-theory T, if T is ω-consistent, then T is consistent.

Proof. Suppose T is inconsistent. Then $T \vdash \exists x \neg\phi(x)$. Then $T \vdash \phi(m)$ for each m, violating ω-consistency. □

Example 25.12. If the standard model of arithmetic: $\eta = (\mathbb{N}, +, \times, S, 0)$ satisfies PA, then PA is ω-consistent.

Proof. Suppose $\eta \models PA$ and $\eta \models \phi(m)$ for every $m \in \mathbb{N}$. Then $\eta \models \neg(\exists x)\neg\phi(x)$, since $\eta \not\models (\exists x)\neg\phi(x)$. Then PA $\nvdash \exists x \neg\phi(x)$. □

Definition 25.13 (semi-representable)**.** $S \subseteq \mathbb{N}$ is semi-representable in PA if $\exists \mathcal{L}_{PA}$-formula ϕ such that $m \in S \iff PA \vdash \phi(m)$.

Theorem 25.14 (Semi-Representability theorem)**.** Assume PA is ω-consistent. Then for $S \subseteq \mathbb{N}$, the following are equivalent:

(1) S is computably enumerable.
(2) S is semi-representable in PA.
(3) $S \leq_m T_{PA}$

Proof. $(2 \implies 3)$:
Assume S is semi-representable in PA via ϕ. Then $n \in S \iff PA \vdash \phi(n)$. Then $n \in S \iff gn(\phi(n)) \in T_{PA}$. So $S \leq_m T_{PA}$. □

Proof. $(3 \implies 1)$:
T_{PA} is computably enumerable. Thus $S \leq_m T_{PA}$ must be computably enumerable. □

Note: $1 \implies 2$ is in the next lecture.

Theorem 25.15 (Semi-Representability Theorem)**.** Assume PA is ω-consistent. For S $\subseteq \mathbb{N}$, the following are equivalent:

(1) S is c.e.
(2) S is semi-representable in PA
(3) S $\leq_m$ T$_P A$

Proof. We proved (2) $\implies$ (3) and (3) $\implies$ (1) last class, so we proceed with (1) $\implies$ (2). Since S is c.e., S is Σ_0^1. That is, there exists a computable relation R such that m $\in$ S $\iff$ $\exists n$(R(m,n) holds). R is a computable relation, so it is representable in PA. By definition, there exists an $\mathcal{L}_{PA}$-formula ψ(x,y) such that R(x,y) $\iff$ PA $\vdash \psi$(x,y) & $\neg$R(x,y) $\iff$ PA $\vdash \neg\psi$(x,y). Let ϕ(x) be the formula $\exists x(\psi(x,y))$.

Claim: m $\in$ S $\iff$ PA $\vdash \phi(\underline{m})$.
($\implies$) Suppose m $\in$ S. Then $\exists n$(R(m,n) holds), so let R(m,n$_0$) hold for some n$_0 \in \mathbb{N}$. Then PA $\vdash \psi(\underline{m}, \underline{n_0})$. PA $\vdash \psi(\underline{m}, \underline{n_0}) \implies \exists x\ \psi(\underline{m}, x)$. So PA $\vdash \exists x\ \psi(\underline{m}, x)$, i.e. PA $\vdash \phi(\underline{m})$.
($\impliedby$) Suppose PA $\vdash \phi(\underline{m})$, i.e. PA $\vdash \exists x\ \psi(\underline{m}, x)$. Then PA $\vdash \exists x\ \neg\neg\psi(\underline{m}, x)$. By ω-consistency, $\exists n_0$ such that PA $\nvdash \psi(\underline{m},\underline{n_0})$. Since ψ represents R, R(m,n$_0$) holds, i.e. $\exists n$ R(m,n). Thus m $\in$ S. □

Corollary 25.16. T$_{PA}$ is m-complete, and hence creative.

Corollary 25.17. If Γ is a computably axiomatizable first-order theory, then T$_\Gamma \leq_m$ T$_{PA}$.

Proposition 25.18. Let Γ be a computably axiomatizable $\mathcal{L}_{PA}$-theory.

(1) If S $\subseteq \mathbb{N}$ is semi representable in Γ, then S is c.e.
(2) If Γ is ω-consistent and every computable relation is representable in Γ, then every c.e. set is semi-representable in Γ.

Recall: All computable sets are representable in PA.

Proposition 25.19. If PA is consistent and S $\subseteq \mathbb{N}$ is representable in PA, then S and $\overline{S}$ are semi-representable in PA.

$$S \subseteq \mathbb{N} \text{ is representable in PA iff } \begin{cases} m \in S \Longrightarrow PA \vdash \phi(\underline{m}) \\ m \notin S \Longrightarrow PA \vdash \neg\phi(\underline{m}) \end{cases}$$

$$S \subseteq \mathbb{N} \text{ is semi-representable in PA iff } m \in S \Longleftrightarrow PA \vdash \phi(\underline{m})$$

Corollary 25.20. The representable sets in PA are exactly the computable sets.

Lemma 25.21.

(1) Let Γ be a computably axiomatizable first order theory in which $K = \{e : \phi_e(e)\downarrow\}$ is semi-representable. Then there exists an $\mathcal{L}$-sentence ϕ such that $\Gamma \nvdash \phi$ and $\Gamma \nvdash \neg\phi$. (So Γ is incomplete.)
(2) If PA is ω-consistent, then PA is incomplete.

Proof. (1) Let ϕ semi-represent K in Γ. Then $n \in K \Longleftrightarrow \Gamma \vdash \phi(\underline{m})$. Since $n \notin K$ for some $n \in \mathbb{N}$, it must be the case that $\Gamma \nvdash \phi(\underline{n})$. Hence Γ is consistent.
Since $\overline{K}$ is not c.e., it cannot be semi-representable in Γ. So it is not the case that

$$n \in \overline{K} \Longleftrightarrow \Gamma \vdash \neg\phi(\underline{n})$$

However, if $\Gamma \vdash \neg\phi(\underline{n})$, then it must be the case that $n \notin K$. (Suppose $\Gamma \vdash \neg\phi(\underline{n})$ and $n \in K$. Then $\Gamma \vdash \phi(\underline{n})$, contradicting the consistency of Γ.) So ($\Longleftarrow$) is always true.
Thus, there exists $n \in K$ such that $\Gamma \nvdash \neg\phi(\underline{n})$. Moreover, we must have $\Gamma \nvdash \phi(\underline{n})$. (Otherwise, we would have $n \in K$, because ϕ semi-represents K in Γ.)

(2) Assume PA is ω-consistent. All c.e. sets are semi-representable in PA, so K is semi-representable in PA. Hence PA is incomplete. □

Remark.

(1) There is some $n \in \overline{K}$ such that PA $\nvdash \neg\phi(\underline{n})$, where ϕ semi-represents K in PA. So if $\mathbb{N} \models$ PA, then $\mathbb{N} \models \neg\phi(\underline{n})$ or $\mathbb{N} \models \phi(\underline{n})$. However, if $\mathbb{N} \models \phi(\underline{n})$, then $n \in K$. So $\mathbb{N} \models \neg\phi(\underline{n})$.

(2) There exist infinitely many n such that $n \in \overline{K}$ and PA $\nvdash \neg\phi(\underline{n})$ and PA $\nvdash \phi(\underline{n})$.

Consequence of Lemma: Let Γ be an ω-consistent $\mathcal{L}_{PA}$-theory in which every computable relation is representable. Then Γ is incomplete.

Theorem 25.22 (G1)**.** There is no ω-consistent computably axiomatizable $\mathcal{L}_{PA}$-theory Γ extending PA that is complete.

Proof. If $\Gamma \supseteq$ PA, then PA $\vdash \phi$ implies $\Gamma \vdash \phi$. Thus if ϕ represents a relation in PA, then ϕ represents a relation in Γ. All computable relations are representable in PA, so all computable relations are representable in Γ. Thus Γ is incomplete by the above. □

Theorem (G2): "PA cannot prove its own consistency"

Observe PA is consistent $\implies$ PA $\nvdash \neg(0 = 0)$. Let m = gn($\neg(0 = 0)$). Then PA is consistent $\iff$ $\neg\exists k$Proof(k) and $(\mathrm{k})_{length(k)} = \mathrm{m}$. Both are representable in PA.

Call the formula expressing that $\neg(0 = 0)$ is not provable in PA con(PA).

Theorem (G2): PA $\nvdash$ con(PA) if PA is consistent.

26. Decidable vs. Undecidable Theories (for a fixed $\mathcal{L}$)

Definition 26.1. Let Γ be a first-order theory. Then Γ is decidable if T_Γ is computable; that is, if we can computable decide, for each ϕ, whether or not $\Gamma \vdash \phi$.

We have seen $\mathrm{K} \leq_m \mathrm{T}_{PA} \implies \mathrm{T}_{PA}$ is not computable $\implies \mathrm{T}_{PA}$ is not decidable. Are there any decidable theories?

Proposition 26.2. If Γ is complete (i.e. for every ϕ, $\Gamma \vdash \phi$ or $\Gamma \vdash \neg\phi$) and computably axiomatizable, then Γ is decidable.

Proof. To determine whether $\mathrm{m} \in \mathrm{T}_\Gamma$ or $\mathrm{m} \notin \mathrm{T}\Gamma$,

(1) Compute $\mathrm{gn}^{-1}(\mathrm{m})$. Suppose $\mathrm{gn}(\phi) = \mathrm{m}$.
(2) Γ is computably axiomatizable, so enumerate proofs until we see $\Gamma \vdash \phi \implies \mathrm{m} \in \mathrm{T}_{PA}$, or $\Gamma \vdash \neg\phi \implies \mathrm{m} \notin \mathrm{T}_{PA}$.

□

DLOWE is complete and finitely axiomatizable. Hence computably axiomatizable $\implies$ DLOWE decidable. ACF_0 and ACF_P are complete and computably axiomatizable, so decidable.

27. Set Theory

Language $\mathcal{L} = \{e\}$
For any property ϕ of sets, $\{x : \phi(x)\}$ is a definable set. Q: Is S $\in$ S? If S $\in$ S, then S $\notin$ S. If S $\notin$ S, then S $\in$ S.
Axiom 0: Universe is nonempty

$$\exists x(x = x)$$

Axiom 1: Extensionality

$$\forall x \forall y(\forall y(z \in x \iff z \in y) \implies x = y)$$

Axiom 2: Foundation

$$\forall x[\exists y(y \in x) \implies \exists y(y \in x \wedge \neg \exists z(z \in x \wedge z \in y)]$$

Axiom 3: Comprehension (Scheme)

For each formula ϕ, with free variables among x, y, z, w_1,..., w_n

$$\forall z \forall w_1 ... \forall w_n \exists y[x \in y \iff x \in z \wedge \phi(x, w_1, ..., w_n)]$$

28. Axioms of ZFC

0. **Existence:** There is a set.
1. **Extensionality:** Defines equality of sets.
2. **Foundation:** Every nonempty set x has an element disjoint from x.
3. **Comprehension:** Set builder notation (rules out $\{x : x \notin x\}$).

Axioms 0-3 have already been described in detail.

4. **Pairing:** $\forall x \forall y \exists z(x \in z \wedge y \in z)$
Warning: x and y are not necessarily the only elements of z!
5. **Union:** $\forall \mathcal{F} \exists A \forall Y \forall x[(x \in Y \wedge Y \in \mathcal{F}) \to x \in A]$
Warning: A could be bigger than $\bigcup \mathcal{F}$!
6. **Replacement Scheme:** For any first-order formula φ with free variables among $x, y, A, w_1, ..., w_n$: $\forall A \forall w_1, ..., w_n([\forall x \in A \exists! y \varphi(x, y)] \to [\exists Y \forall x \in A \exists y \in Y \varphi(x, y)])$.
(In other words, ranges of functions exist.)

Note: Using Axioms 0, 1, and 3-6, we can define $\emptyset$, $\subseteq$, ordinal successor ($S(x) = x \cup \{x\}$), and the notion of a well-ordering.

7. **Infinity:** $\exists x[\emptyset \in x \wedge \forall y \in x(S(y) \in x)]$
8. **Power Set:** $\forall x \exists y \forall z(z \subseteq x \to z \in y)$
9. **Choice:** $\forall A \exists R(R$ well-orders $A)$

29. Consequences of Extensionality and Comprehension

Proposition 29.1. There is a unique set that contains no elements, denoted $\emptyset$.

Proof. By Axiom 0, there exists some set z. By comprehension, we can define $\{x \in z : x \neq x\}$. Call this $\emptyset$.
Suppose there exists some set y with no elements. Then $\forall z(z \in y \leftrightarrow z \in \emptyset)$. So by extensionality, $y = \emptyset$.

□

Theorem 29.2. There is no universal set, i.e. a set y such that $\forall x(x \in y)$.

Proof. Suppose $\exists y \forall x(x \in y)$. By comprehension, let $\{x \in y : x \notin x\} = S$.
S is a set, so $S \in y$. But now we have $S \in S \Leftrightarrow S \notin S$, which is a contradiction. So there is no universal set.

□

Write $A \subseteq B$ as shorthand for $\forall x(x \in A \to x \in B)$.

Remark. $\emptyset \subseteq A$ and $A \subseteq A$ for any A.

Remark. $\emptyset$ is the only set whose existence can be proved from Axioms 0, 1, 3. Proof: Let $\mathcal{M} = \{\emptyset\}$, $\in^{\mathcal{M}} = \in$. (So $\mathcal{M}$ is the model whose universe is $\{\emptyset\}$, with $\in^{\mathcal{M}}$ being the usual set membership.) Then $\mathcal{M} \vDash$ Axioms 0, 1, 3. But $\mathcal{M} \vDash \forall y(y = \emptyset)$. So we must have Axioms 0, 1, 3 $\nvdash$ (equivalently, $\nvdash$ or $\nvDash$) $\exists y(y \neq \emptyset)$.

30. More Sets

Now we need to define more sets using the other axioms. (Heuristic: "small sets exist".)

From x, y, we want to form $\{x, y\}$.
By pairing, $\exists z(x \in z \wedge y \in z)$, so by comprehension,
consider $\{v \in z : v = x \vee v = y\}$.
Call this $\{x, y\}$. Unique by extensionality.

Let $\{x\} = \{x, x\}$. So for every set x, the set $\{x\}$ exists.
In particular, we can build infinite sets, but this doesn't eliminate the need for the axiom of infinity.

We can define $\{x, \{x, y\}\} =: < x, y >$.
Check: $\forall x \forall y \forall x' \forall y'[(< x, y > = < x', y' >) \to (x = x' \wedge y = y')]$

Union

Given a family of sets $\mathcal{F}$, we can define
$\bigcup \mathcal{F} = \{x : \exists Y \in \mathcal{F}(x \in Y)\}$
$= \{x \in A : \exists Y \in \mathcal{F}(x \in Y)\}$ (using comprehension), where A is guaranteed to exist by the union axiom.

Intersection

For $\mathcal{F} \neq \emptyset$, let $\bigcap \mathcal{F} = \{x : \forall Y \in \mathcal{F}(x \in Y)\}$.
This exists, because we take $B \in \mathcal{F}$, and then we have (using comprehension)
$\bigcap \mathcal{F} = \{x \in B : \forall Y \in \mathcal{F}(x \in Y)\}$.

Let $A \cap B := \bigcap\{A, B\}$, $A \cup B := \bigcup\{A, B\}$, and $A \backslash B := \{x \in A : x \notin B\}$.

The replacement scheme is suggested by our heuristic: "small sets exist". (If $f : A \to B$, then $\text{ran}(f)$ is no bigger than A.)

Cartesian products are defined by replacement:
$A \times B = \{< x, y >: x \in A \wedge y \in B\}$, as follows.

First, for a fixed $y \in B$, we have $\forall x \in A \exists! z(z =< x, y >)$.
So apply replacement to define $\text{prod}(A, y) = \{z : \exists x \in A(z =< x, y >)\}$.
Next, $\forall y \in B \exists! z(z = \text{prod}(A, y))$. Note that z is unique by extensionality.
Apply replacement to define $\text{prod}(A, B) = \{\text{prod}(A, y) \colon \exists y \in B\}$.
Lastly, $A \times B = \bigcup \text{prod}(A, B)$.

31. Relations

Definition 31.1 (Relation)**.** A relation is a set R, all of whose members are ordered pairs.

Definition 31.2 (Domain)**.** $\text{dom}(R) = \{x : \exists y(< x, y > \in R)\}$

Definition 31.3 (Range)**.** $\text{ran}(R) = \{y : \exists x(< x, y > \in R)\}$

We write xRy instead of $< x, y > \in R$.

Definition 31.4 (Function)**.** "f is a function" means f is a relation and
$\forall x \in \text{dom}(f) \exists! y \in \text{ran}(f)(< x, y > \in f)$.

$f : A \to B$ means $\text{dom}(f) = A$ and $\text{ran}(f) \subseteq B$.

Definition 31.5 (Restriction and image)**.** If $C \subseteq A$, then $f \restriction C = f \cap (C \times B)$ (the restriction of f to C), and $f''(C) = \text{ran}(f \restriction C) = \{f(x) : x \in C\}$.

Definition 31.6 (Total ordering)**.** A (strict) total ordering is a pair $< A, R >$ such that A is a set, R is a relation on A, and the following hold:
1) R is transitive on A $(\forall x, y, z(xRy \wedge yRz \to xRz))$;
2) R is irreflexive on A $(\forall x(\neg xRx))$; and
3) R satisfies trichotomy $(\forall x, y(x = y \vee xRy \vee yRx))$.

Definition 31.7 (Isomorphism)**.** If R, S are relations on A, B respectively, then
$< A, R > \cong < B, S >$ means $\exists$ bijection $f : A \to B$ s.t. $\forall x, y \in A[xRy \leftrightarrow f(x)Sf(y)]$.
We say "f is an isomorphism from $< A, R >$ to $< B, S >$".

32. Well-Orderings

Definition 32.1 (Well-ordering)**.** R "well-orders A", or "$< A, R >$ is a well-ordering", if $< A, R >$ is a (strict) total ordering (henceforth, "strict" will be implicit) and every nonempty subset of A has an R-least element.

Definition 32.2 (Proper initial segment)**.** For a total ordering $< A, R >$ and $x \in A$, let $\text{pred}(A, x, R)$ be $\{y \in A : yRx\}$ ("a proper initial segment of A").

Lemma 32.3. If $< A, R >$ is a well-ordering, then for all $x \in A$, $< A, R > \ncong <\text{pred}(A, x, R), R >$.

Proof. Suppose $f :< A, R > \to <\text{pred}(A, x, R), R >$ is an isomorphism for some $x \in A$.

Consider the set $S = \{y \in A : f(y) \neq y\}$. Since $f(x) \neq x$ (because $f(x)Rx$ and R is irreflexive), we have $S \neq \emptyset$. So by well-ordering, there exists an R-least $y \in S$. By trichotomy, either $yRf(y)$ or $f(y)Ry$.

Suppose $f(y)Ry$. Then (since y is R-least in S) $f(y) \notin S$, so $f(f(y)) = f(y)$, so (since f is 1-1) $f(y) = y$, which contradicts our choice of y.

Suppose $yRf(y)$. Then yRx (since $f(y)Rx$). Hence since f is onto, $\exists z \neq y(f(z) = y \neq z)$, so $z \in S$. Then since y is R-least in S, we have yRz. But we also have $f(z)Rf(y)$, which contradicts the R-preservation of f.

□

Lemma 32.4. If $< A, R >, < B, S >$ are isomorphic well-orderings, then the isomorphism is unique.

Proof. Suppose $f, g :< A, R > \to < B, S >$ are distinct isomorphisms.
Let $T = \{x : f(x) \neq g(x)\}$. Since $T \neq \emptyset$ by assumption, let y be the R-least element of T.
WLOG, assume $f(y)Sg(y)$.
g onto $\Rightarrow \exists z[g(z) = f(y)]$, so $g(z)Sg(y) \Rightarrow zRy \Rightarrow z \notin T \Rightarrow f(z) = g(z) = f(y)$.
But zRy, so this contradicts the assumption that f is 1-1.

□

Theorem 32.5. Let $< A, R >, < B, S >$ be two well-orderings. Then exactly one of the following holds.
1) $< A, R > \cong < B, S >$
2) $\exists x \in A(<\text{pred}(A, x, R), R > \cong < B, S >)$
3) $\exists y \in B(< A, S > \cong <\text{pred}(B, y, S), S >)$

Proof. Let $f = \{< v, w >: v \in A \wedge w \in B \wedge <\text{pred}(A, v, R), R > \cong <\text{pred}(B, w, S), S >\}$.

Claim 1: f is an isomorphism from some initial segment of A to some initial segment of B.
Proof: Suppose $v \in \text{dom}(f)$. Then by definition, there exists $w \in B$ so that
$<\text{pred}(A, v, R), R > \cong <\text{pred}(B, w, S), S >$ via an isomorphism g.
Let $z \in A$ with zRv. Then $g \restriction z$ gives an isomorphism between $<\text{pred}(A, z, R), R >$ and $<\text{pred}(B, g(z), S), S >$, so $z \in \text{dom}(f)$.
Similarly, if $z \in B$ with zSw, then $z \in \text{ran}(f)$.

Claim 2: The initial segments cannot both be proper.
Proof: Suppose $f :<\text{pred}(A, x, R), R > \to <\text{pred}(B, y, S), S >$. Then by the definition of f, $< x, y > \in f$, which contradicts our assumption about $\text{dom}(f)$.

The theorem follows from Claims 1 and 2.

□

33. Ordinals

Definition 33.1 (Transitive)**.** A set x is transitive if every element of x is a subset of x.

Example 33.2. $\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$, and all of its elements individually.

Non-example: $\{\{\emptyset\}\}$ is not transitive, because $\{\emptyset\} \in \{\{\emptyset\}\}$, but $\{\emptyset\} \not\subseteq \{\{\emptyset\}\}$.

Definition 33.3 (Ordinal)**.** A set x is an ordinal if x is transitive and x is well-ordered by $\in$.

Note: "x is well-ordered by $\in$" means $< x, \in_x >$ is a well-ordering, where $\in_x = \{< y, z > \in x \times x : y \in z\}$.

Example 33.4. $\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}$

Non-example: $\{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}\}$ doesn't satisfy trichotomy!

Hereafter, we drop the $\in_x$. So "$x \cong < A, R >$" means "$< x, \in_x > \cong < A, R >$", and for $y \in x$, "pred(x, y)" means "pred$(x, y, \in_x)$".

Theorem 33.5. The following five statements hold.
(1) If x is an ordinal and $y \in x$, then y is an ordinal and $y = \text{pred}(x, y)$.
(2) If x and y are ordinals and $x \cong y$, then $x = y$.
(3) If x, y ordinals, exactly one of the following holds: $x = y$, $x \in y$, $y \in x$.
(4) If x, y, z ordinals, $x \in y$, and $y \in z$, then $x \in z$.
(5) If C is a nonempty set of ordinals, we have $\exists x \in C \forall y \in C(x \in y \vee x = y)$.

Proof.

(1) Let $y \in x$. Since x is an ordinal, $y \subseteq x$. Clearly, y is totally ordered by $\in$.

• y is well-ordered by $\in$: Let $S \subseteq y$ be nonempty. Since x is well-ordered by $\in$, $S \subseteq y \subseteq x$ implies that there is an $\in$-least element in S, as desired.

• y is transitive: To show $\forall z \in y(z \subseteq y)$, let $z \in y$ and take $w \in z$ arbitrary.
Together, $y \subseteq x$ and $z \in y$ imply $z \in x$. Also, x is transitive, so $w \in z \Rightarrow w \in x$.
So $w \in z \in y \in x$. Since x is strictly ordered by $\in$, either $y \in w$ or $w \in y$.
If $y \in w$, then $y \in w \in z \in y$, contradicting the strictness of the ordering.
So $w \in y$, which proves $z \subseteq y$.

• $y = \text{pred}(x, y)$: This follows from $z \in y \Leftrightarrow z \in \text{pred}(x, y)$.

(2) Suppose $x \cong y$ but $x \neq y$. WLOG, $\exists z \in x \backslash y$. By the well-ordering of x, let z be $\in$-least. (Henceforth, we simply write "least".)

By the transitivity of x, $z \in x \Rightarrow z \subseteq x$. So $\forall w \in z(w \in x)$. By our choice of z, this implies $\forall w \in z(w \in y)$. Hence $z \subseteq y$.

Note $y \neq z$. (Otherwise, we'd have $x \cong y$ and $y = z \in x$, which is impossible: by an earlier result, x can't be isomorphic to a proper initial segment.)

Hence $y \backslash z$ is nonempty. By the well-ordering of y, let w' be least in $y \backslash z$. Then $z = \text{pred}(y, w') =$ (by (1)) w', so $z \in y$, contradicting our choice of z.

Thus $x = y$.

(3) By an earlier result on well-orderings, we must have one of:

(i) $x \cong y$.
(ii) $x \cong$ an initial segment of y.
(iii) $y \cong$ an initial segment of x.

Each case gives us what we want:

(i) $x \cong y \Rightarrow x = y$ by (2).
(ii) $\exists w \in y(x \cong \text{pred}(y, w) = w)$ (by (1)) $\Rightarrow$ (by (2)) $x = w \in y$.
(iii) same as (ii); implies $y \in x$.

Furthermore, since $\in$ is strict, no two of $x = y, x \in y, y \in x$ can hold simultaneously.

(4) z is transitive, so $y \in z \Rightarrow y \subseteq z$, so $x \in y \Rightarrow x \in z$.

(5) Let $C \neq \emptyset$ be a set of ordinals.
We want to show $\exists x \in C \forall y \in C(x = y \vee x \in y)$.
Equivalently (by (3)), $\exists x \in C \forall y \in C(y \notin x)$.
Equivalently, $\exists x \in C(x \cap C = \emptyset)$.

Let $x \in C$. If $x \cap C = \emptyset$, we're done. Otherwise, let x' be least in $x \cap C$.
If $x' \cap C \neq \emptyset$, then $\exists y \in x' \cap C$, so $y \in x' \subseteq x$ and $y \in C$, so $y \in x \cap C$ and $y \in x'$, contradicting our choice of x'.
Hence $x' \cap C = \emptyset$, and we're done.

□

Theorem 33.6 (Burali-Forti Paradox)**.** $\neg \exists z \forall x$ (x is an ordinal $\rightarrow x \in z$)

Proof. Suppose there is such a z. Then ON $= \{x : x \text{ an ordinal}\}$ exists.
ON is transitive by **(1)**, and well-ordered by **(3)**, **(4)**, **(5)**.
Thus ON is an ordinal, so ON $\in$ ON, violating the strictness of $\in$ on ON.

□

Lemma 33.7. If A is a transitive set of ordinals, it is an ordinal (i.e. if A is a set of ordinals and $\forall x \in A \forall y \in x(y \in A)$, then A is an ordinal).

Proof. Suppose A is a transitive set of ordinals. Then by part 5 of the previous theorem, A is well-ordered. Thus, A is an ordinal. □

Theorem 33.8. If $\langle A, R\rangle$ is a well-ordering, then there exists a unique ordinal C such that

$$\langle A, R\rangle \simeq C$$

Proof. Since isomorphic ordinals are identical, uniqueness holds. To show existence, let $B = \{a \in A : \exists x(x \text{ is an ordinal} \wedge x \simeq \langle pred(A, a, R), R\rangle\}$. Then define a function f with domain B such that for all $a \in B$, f(a) is the unique x such that $x \simeq \langle pred(A, a, R), R\rangle\}$(since isomorphisms are identical, x is unique). By replacement, let $C = f(a)$.

Claim: C is a Cardinal.

Proof. C is a set of ordinals. So it suffices to show $\forall x \in C \forall y \in x(y \in C)$. Let $x \in C$ and $y \in x$. Then $x \simeq \langle pred(A, a, R), R\rangle\}$ for some $a \in A$. Then y is a well-ordering, so $y \simeq \langle pred(A, a, R), R\rangle \simeq x$ or $y \simeq$ to an initial segment of $\langle pred(A, a, R), R\rangle$ or $\langle pred(A, a, R), R\rangle \simeq$ to an initial segment of y. If $y \simeq x$, $x = y \in x$. A contradiction. If $\langle pred(A, a, R), R\rangle \simeq$ to an initial segment of y, $x \simeq$ an initial segment of $\langle pred(A, a, R), R\rangle$. Then $x \in y$ and $y \in x$, so $y \in y$. A contradiction. So $y \simeq$ to an initial segment of $\langle pred(A, a, R), R\rangle$. Then $\exists a' R a$ such that $y \simeq \langle pred(A, a', R), R\rangle$. So $y \in C$. So C is a cardinal.

Claim: $f : \langle B, R\rangle \to C$ is an isomorphism.

Proof. It is clear that f is a map. Also, f is onto by definition. And by choice of f(a) being a unique value, f is one-to-one. Furthermore, $aRa' \longrightarrow f(a) \in f(a')$, so f is order preserving.

Claim: Either B=A, or $B = pred(A, b, R)$ for some $b \in B$.

Proof. Suppose $B \neq A$. Then $\langle B, R\rangle$ is a well-ordering so $\langle B, R\rangle \simeq$ to an initial segment of A.

Now if $B = pred(A, b, R)$ for some $b \in B$, since $\langle B, R\rangle \simeq C$, $\langle pred(A, b, R), R\rangle \simeq C$. So be definition of B, $b \in B$.A contradiction. So A=B. Hence, $\langle A, R\rangle = \langle B, R\rangle \simeq C$.

□

Definition 33.9. If $\langle A, R\rangle$ is a well-ordering, $type\langle A, R\rangle$ is the unique ordinal C such that $\langle A, R\rangle \simeq C$.

Notation remark. Hereafter, we will use Greek letters to denote ordinals, and $\alpha \leq \beta$ will mean $\alpha \in \beta$.

Definition 33.10. If X is a set of ordinals, $sup(X) = \bigcup X$ and if $X \neq \emptyset$, $min(X) = \bigcap X$.

Lemma 33.11. (1) $\forall \alpha, \beta(\alpha \leq \beta \leftrightarrow \alpha \subseteq \beta)$.

(2) If X is a set of ordinals, $sup(X)$ is the least ordinal greater than or equal to all elements of X. If $X \neq \emptyset$, $min(X)$ is the least element of X.

Proof. (1) Given α,β, suppose $\alpha \leq \beta$. So $\alpha \in \beta$ or $\alpha = \beta$. By transitivity, $\alpha \subseteq \beta$. Now suppose $\alpha \subseteq \beta$. If $\beta \in \alpha$, $\beta \subseteq \alpha$, so $\alpha = \beta$. Then $\alpha \in \alpha$. A contradiction. So $\alpha \leq \beta$.

(2) First we show $sup(X) = \bigcup X$ is an ordinal. We know that $sup(X)$ is a collection of ordinals. So it suffices to show $\forall x \in sup(X) \forall y \in x(y \in sup(X))$. Let $x \in sup(X)$ and $y \in x$. Then $\exists \alpha \in X$ such that $x \in \alpha$. Then $x \subseteq \alpha$. So $y \in \alpha$. Hence, $y \in sup(X)$.

Now we show $sup(X)$ is greater than or equal to all elements of X. Let $\alpha \in X$. Then $\alpha \in sup(X)$, and $\alpha \leq \bigcup X$ by (1). Let $\gamma \geq \mu$ $\forall \mu \in X$. Then since $\alpha \in \bigcup X$, $\exists \beta \in X$ such that $\alpha \in \beta \in \gamma$. Then by transitivity, $\alpha \in \gamma$. So $\bigcup X \subseteq \gamma$.

Now for $min(X)$, let $X \neq \emptyset$.

Claim: $\bigcap X$ is an ordinal.

Proof. We have $\bigcap X$ is a collection of ordinals. So we show $\forall x \in \bigcap X \forall y \in x (y \in \bigcap X)$. Let $x \in \bigcap X$ and $y \in x$. Then $\forall \alpha \in X (x \in \alpha)$. So $y \in x \in \alpha, \implies \forall \alpha \in X (y \in \alpha)$ by transitivity. So $y \in \bigcap X$. Hence $\bigcap X$ is an ordinal.

Claim: Let α be the least element of X. Then $\alpha = \bigcap X$.

Proof. ($\subseteq$) If $\beta \in \bigcap X$, $\forall \gamma \in X (\beta \leq \gamma)$. In particular, since $\alpha \in X$, $\beta \leq \gamma$ and thus $\beta \in \alpha$.

($\supseteq$) If $\beta \in \alpha$, then since $\forall \gamma \in X (\alpha \in \gamma)$, $\beta \in \alpha \in \gamma$. Which implies $\beta \in \gamma$. Then $\forall \gamma \in X (\beta \in \gamma)$. Hence $\beta \in \bigcap X$.

□

Last time we built on our knowledge of ordinals, showing that every well-ordering is isomorphic to a unique ordinal as well as proving several results about the $<$ relation and how it relates to the $\in$ relation. This time we continue to build on the previous few lectures by defining the natural numbers as ordinals and introducing ordinal arithmetic.

34. Ordinals

We begin by defining the successor function and proving several of its properties.

Definition 34.1 (Successor Function)**.** We define the *successor function* S to be $S(\alpha) = \alpha \cup \{\alpha\}$

Lemma 34.2. For any α we have:

(1) $S(\alpha)$ is an ordinal
(2) $\alpha < S(\alpha)$
(3) $\forall \beta (\beta < S(\alpha) \iff \beta \leq \alpha)$

Proof. This lemma follows straightforwardly from the definitions. □

Definition 34.3. An ordinal α is a *successor ordinal* if $\exists \beta (\alpha = S(\beta))$. An ordinal α is a *limit ordinal* if $\alpha \neq \emptyset$ and α is not a successor ordinal.

Definition 34.4 (Natural Numbers)**.** We define the *natural numbers* inductively using the successor function starting with $\emptyset$ by $0 = \emptyset, 1 = S(0) = S(\emptyset), 2 = S(1) = S(S(\emptyset)) ... n = S(n-1)$. An ordinal α is a natural number if $\forall \beta \leq \alpha (\beta = \emptyset$ or β is a successor ordinal).

We now add to our axiomatic foundation of Set Theory by introducing the Axiom of Infinity, which intuitively states that there exists an infinite set.

Axiom 34.5 (Axiom of Infinity)**.** $\exists x (\emptyset \in x \wedge (\forall y \in x (S(y) \in x)))$

Claim 34.6. If x satisfies the Axiom of Infinity, then x contains all natural numbers.

Proof. Suppose x satisfies the Axiom of Infinity and $\exists n \notin x$ where n is a natural number. By definition $\exists m < n$ such that $S(m) = n$, so we have $m \notin x$ either. Now we know that $n \setminus x$ is nonempty, so let n' be least in $n \setminus x$. Then $\exists m' < n'$ with $S(m') = n'$ and $m' \in x$, so $n' \in x$. $\Rightarrow\Leftarrow$ □

Definition 34.7. We denote the set of all natural numbers by ω.

Observe that all ordinals smaller than ω are either $\emptyset$ or a successor ordinal and that if ω were a successor ordinal, then it would be the successor of some natural number, and hence

would itself be a natural number. Since ω contains all natural numbers, we would have $\omega \in \omega$ which would be a contradiction. Hence ω is a limit ordinal and is actually the first such ordinal.

Theorem 34.8 (Peano Postulates)**.** The following properties hold of ω :

(1) $\emptyset \in \omega$
(2) $\forall n \in \omega, S(n) \in \omega$
(3) $\forall n, m \in \omega((n \neq m) \rightarrow (S(n) \neq S(m))$
(4) $\forall x \subset \omega((\emptyset \in x \wedge (\forall n \in x(S(n) \in x))) \rightarrow x = \omega)$

35. Ordinal Arithmetic

We now introduce addition and multiplication on ordinals and show some of their properties, as well as explain where our intuition about arithmetic on the natural numbers might break down in the general case.

Definition 35.1 (Ordinal Addition)**.** For ordinals α and β, we define $\alpha + \beta = type(\langle \alpha \times \{0\}\rangle \cup \langle \beta \times \{1\}\rangle, R)$,
where $R = \{\langle\langle \xi, 0\rangle, \langle \eta, 0\rangle\rangle : \xi < \eta < \alpha\} \cup \{\langle\langle \xi, 1\rangle, \langle \eta, 1\rangle\rangle : \xi < \eta < \beta\} \cup \langle\langle \alpha \times \{0\}\rangle \times \langle \beta \times \{1\}\rangle\rangle$.

Intuitively, $\alpha + \beta$ corresponds to the ordinal type of all pairs $\langle \alpha, \{0\}\rangle \cup \langle \beta \times \{1\}\rangle$ equipped with the lexicographic ordering. Alternatively, α and β can be visualized as lines, with their sum being interpreted as appending one line to the other.

Lemma 35.2. For all ordinals α, β, γ, we have:

(1) $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$
(2) $\alpha + 0 = \alpha$
(3) $\alpha + 1 = S(\alpha)$
(4) $\alpha + S(\beta) = S(\alpha + \beta)$
(5) If β is a limit ordinal, then $\alpha + \beta = sup\{\alpha + \xi : \xi < \beta\}$

Note that $1 + \omega \neq \omega + 1$. Intuitively for $1 + \omega$ we have that ω is being shifted up by 1 whereas for $\omega + 1$ we have 1 being appended onto the end of ω. Hence addition is not commutative in general, although it is commutative on the natural numbers.

Definition 35.3 (Ordinal Multiplication)**.** For ordinals α and β, we define $\alpha \cdot \beta = type(\beta \times \alpha, R)$,
where R is the lexicographic ordering on $\beta \times \alpha$ i.e. $\langle \xi, \eta\rangle R \langle \xi', \eta'\rangle \iff (\xi < \xi') \vee (\xi = \xi' \wedge \eta < \eta')$.

Intuitively, $\alpha \cdot \beta$ is β copies of α.

Lemma 35.4. For all ordinals α, β, γ, we have:

(1) $\alpha \cdot (\beta \cdot \gamma) = (\alpha \cdot \beta) \cdot \gamma$
(2) $\alpha \cdot 0 = 0$
(3) $\alpha \cdot 1 = \alpha$
(4) $\alpha \cdot S(\beta) = \alpha \cdot \beta + \alpha$
(5) If β is a limit ordinal, then $\alpha \cdot \beta = sup\{\alpha \cdot \xi : \xi < \beta\}$
(6) $\alpha \cdot (\beta + \gamma) = \alpha \cdot \beta + \alpha \cdot \gamma$

Note that similarly to addition, multiplication is not commutative in the general case, although it is on the natural numbers. For example, we have that $\omega \cdot 2 \neq 2 \cdot \omega$ because, intuitively, the first is 2 copies of ω while the second is ω copies of 2. Also of note is that in the general case multiplication is not right-associative, i.e. $(\beta + \gamma) \cdot \alpha \neq \beta \cdot \alpha + \gamma \cdot \alpha$.

Last time we continued to build on our knowledge of ordinals by defining the natural numbers, the axiom of infinity, and introducing ordinal arithmetic. This time we finish up ordinal arithmetic and introduce classes and recursion, most importantly proving transfinite induction and transfinite recursion.

36. Ordinal Arithmetic

Definition 36.1.

(1) A^n is the set of all functions from n to A.
(2) $A^{<\omega} = \bigcup\{A^n : n \in \omega\}$.

Note that $A^2 \neq A \times A$, although there is a bijection between them, $f : \{0, 1\} \to A \in A^2$ is $f(0) = a_0, f(1) = a_1$ for $\langle a_0, a_1\rangle \in A \times A$.

To show A^n and $A^{<\omega}$ exist, we define $\phi(n, y) = \forall s(s \in y \iff s$ is a function from n to $A)$ and use the axioms of induction and replacement, then identify A^{n+1} with $A^n \times A$ to show $\forall n \exists y \phi(n, y)$. We then use extensionality to get $\forall n \exists y' \phi(n, y')$ and apply replacement to define $\{A^n : n \in \omega\}$ and then $A^{<\omega} := \bigcup\{A^n : n \in \omega\}$.

37. Classes and Recursion

Definition 37.1 (Class)**.** A collection of the form $\{x : \phi(x)\}$ is called a *class*. A *proper class* is a class that is not also a set.

Examples of proper classes include the universe $V = \{x : x = x\}$ and the ordinals $\mathbf{ON} = \{x : x$ is an ordinal $\}$.

For the remainder of these notes we write classes of ordinals in **boldface**.

Theorem 37.2 (Transfinite Induction)**.** If $\mathbf{C} \subseteq \mathbf{ON}$ is a class and $\mathbf{ON} \neq \emptyset$, then $\mathbf{C}$ has a least element.

Proof. Fix some $\alpha \in \mathbf{C}$. If α is least, then we are done, so assume α is not least. Now let β be least in $\alpha \cap \mathbf{C}$. We claim β is least in $\mathbf{C}$. □

Theorem 37.3 (Jech's Version)**.** Suppose $\mathbf{C}$ is a nonempty collection of ordinals satisfying

(1) $\emptyset \in \mathbf{C}$
(2) $\alpha \in \mathbf{C} \to S(\alpha) \in \mathbf{C}$
(3) If β is a limit ordinal and $\forall \alpha < \beta(\alpha \in \mathbf{C}) \to \beta \in \mathbf{C}$

Then $\mathbf{C} = \mathbf{ON}$.

Transfinite induction is a generalization of induction on the natural numbers to include ordinals and Jech's version is analogous to the well-cording principle. As you might expect, transfinite induction and Jech's version are equivalent. In general, a proof by transfinite induction on α establishes $\forall \alpha \psi(\alpha)$ by showing for each $\alpha(\forall \beta < \alpha \psi(\beta) \to \psi(S(\alpha))$. If $\exists \alpha \neg \psi(\alpha)$, then the least such α gives a contradiction, thus $\forall \alpha \psi(\alpha)$.

Theorem 37.4 (Transfinite Recursion)**.** If $F : V \to V$, then there is a unique function $G : \mathbf{ON} \to V$ such that $\forall \alpha G(\alpha) = F(G \restriction \alpha)(\bigstar)$.

Proof. Uniqueness: Suppose G_1 and G_2 both satisfy ★. We claim that $\forall\alpha G_1(\alpha) = G_2(\alpha)$. Suppose not. Then let α be the least such that $G_1(\alpha) \neq G_2(\alpha)$. But then $G_1(\alpha) = F(G_1 \restriction \alpha) = F(G_2 \restriction \alpha) = G_2(\alpha) \Rightarrow\Leftarrow$.
Existence: We define g to be a *δ-approximation* if g is a function with domain δ and $\forall\alpha < \delta g(\alpha) = F(g \restriction \delta)$.
Claim 1: If g is a δ-approximation and g' is a δ'-approximation, then $g \restriction (\delta \cap \delta') = g' \restriction (\delta \cap \delta')$. Without loss of generality suppose $\delta < \delta'$. Let $\alpha < \delta$ be least such that $g(\alpha) \neq g'(\alpha)$. Then $g(\alpha) = F(g \restriction \alpha) = F(g' \restriction \alpha) = g'(\alpha)$ and so $g(\alpha) = g'(\alpha)$ which is a contradiction. Hence the claim holds.
Claim 2: For each δ, there exists a δ-approximation. Suppose not. Let δ be the least such that there is no δ-approximation. By replacement we form the set $\{g_\alpha \restriction \alpha : \alpha < \delta\}$. Define $h = \{\langle \alpha, F(g_\alpha \restriction \alpha)\rangle : \alpha < \delta\}$. Then we have $\forall\alpha\forall\beta < \alpha\ h(\beta) = g_\alpha(\beta)$ and hence $h \restriction \alpha = g_\alpha \restriction \alpha$. Thus $\forall\alpha < \delta h(\alpha) = F(g_\alpha \restriction \alpha) = F(h \restriction \alpha)$ and hence h is a δ-approximation. $\Rightarrow\Leftarrow$.
So now we know that for every δ there exists a unique δ-approximation. Define $G(\alpha) = g(\alpha)$ where g is a δ-approximation for some $\delta > \alpha$. By the way that G is defined, it satisfies $\forall\alpha G(\alpha) = F(G \restriction \alpha)$ and the theorem holds. □

Transfinite recursion can be used to define the $+, \cdot$ operations.

Definition 37.5. We define α^β by recursion:

(1) $\alpha^0 = 1$
(2) $\alpha^{\beta+1} = \alpha^\beta \cdot \alpha$
(3) If β is a limit ordinal, then $\alpha^\beta = sup\{\alpha^\xi : \xi < \beta\}$

Note that ordinal exponentiation is *not* the same as cardinal exponentiation. For example, $2^\omega = sup\{2^n : n < \omega\} = \omega$, which is not true under cardinal exponentiation.

Last time we finished up ordinal arithmetic and introduced classes and recursion. We proved transfinite induction and recursion, which are parallels of induction and recursion over the natural numbers. This time we finish up classes and recursion and introduce cardinals and cardinal arithmetic.

38. Classes and Recursion

Theorem 38.1 (Cantor Normal Form)**.** Every $\alpha > \emptyset$ can be uniquely represented as $\alpha = \omega^{\beta_1} \cdot l_1 + \omega^{\beta_2} \cdot l_2 + ... + \omega^{\beta_n} \cdot l_n$, where $1 \leq n < \omega$ is a natural number, $\alpha \geq \beta_1 \geq \beta_2 \geq ... \geq \beta_n$ and $1 \leq l_i < \omega$ for $i = 1, 2, 3, ...n$.

Proof. By induction on α. For the base case, $\alpha = 1$ and so $\alpha = \omega^\emptyset \cdot 1$. Now suppose the result holds for all $\rho < \alpha$. Then, by a previous homework exercise, we have that there exists a unique δ and a unique ρ such that $\alpha = \omega^\beta \cdot \delta + \rho$. We have that $\delta < \omega$ because otherwise β is not least. Applying the induction hypothesis to ρ gives us our result. □

We note that this is very similar to the division algorithm for the natural numbers, but is generalized to include the ordinals.

39. Cardinals

Definition 39.1.

(1) $A \precsim B$ if and only if there is a one-to-one function $f : A \to B$
(2) $A \approx B$ if and only if there is a bijection $f : A \to B$
(3) $A \prec B$ if $A \precsim B$ and $B \not\precsim A$

Remark: The relation $\precsim$ is transitive and $\approx$ is an equivalence relation.
We present the next theorem without proof:

Theorem 39.2 (Shröder-Bernstein)**.** If $A \precsim B$ and $B \precsim A$, then $A \approx B$.

Note that if A can be well-ordered, then $A \approx \alpha$ for some ordinal α.

Definition 39.3. If A can be well-ordered, then the *cardinality* of A, denoted $|A|$, is the least α such that $A \approx \alpha$.

Note that the axiom of choice allows every set to be well-ordered, hence $|A|$ is defined for every A. Even without choice, since every ordinal α is corresponds to a well-ordering, $|\alpha|$ is defined for every α.

Definition 39.4. We define α to be a *cardinal* if and only if $\alpha = |\alpha|$. Equivalently, α is a *cardinal* if and only if $\forall \beta < \alpha \ \beta \not\approx \alpha$.

Hereafter we use the greek letters κ and λ to range over cardinals.

Lemma 39.5. If $|\alpha| \leq \beta \leq \alpha$, then $|\alpha| = |\beta|$.

Proof. We have that since $\beta \leq \alpha$, $\beta \subseteq \alpha$, so $\beta \precsim \alpha$. Then, since $\alpha \approx |\alpha| \subseteq \beta$, $\alpha \precsim \beta$. Then, by Schröder-Bernstein, $\alpha \approx \beta$ and hence $|\alpha| = |\beta|$. □

Lemma 39.6. If $n \in \omega$ then

(1) $n \not\approx n + 1$
(2) $\forall \alpha(\alpha \approx n \to \alpha = n)$

Proof.

(1) By induction.
(2) Suppose $\alpha \approx n$, i.e. $|\alpha| = n$. Since $|\alpha| \leq \alpha$, we have that $n \leq \alpha$. Now suppose that $n < \alpha$. Then $n < n+1 \leq \alpha$, and thus $|\alpha| = n$. This implies that $|\alpha| = |n+1| = n+1$, but we assumed that $|\alpha| = n$. $\Rightarrow\Leftarrow$

□

We have as a corollary that ω is a cardinal. If it were not, then it would have to be in bijection with some $n < \omega$, but $\omega \not\approx n \forall n \in \omega$. We also have that every $n \in \omega$ is a cardinal.

Definition 39.7.

(1) A is *finite* if $|A| < \omega$.
(2) A is *countable* if $|A| \leq \omega$.
(3) A is *infinite* if A is not finite.
(4) A is *uncountable* if A is not countable.

Definition 40.1.

(1) $\kappa \oplus \lambda = |\kappa \times 0 \cup \lambda \times 1|$
(2) $\kappa \otimes \lambda = |\kappa \times \lambda|$

Note that $\kappa \times \{0\} \cup \lambda \times \{1\} \approx \lambda \times \{0\} \cup \kappa \times \{1\}$, so $\kappa \oplus \lambda = \lambda \oplus \kappa$. By swapping coordinates, we also have that $\kappa \times \lambda \approx \lambda \times \kappa$, so $\kappa \otimes \lambda = \lambda \otimes \kappa$. Consequently we have that $|\kappa + \lambda| = |\lambda + \kappa| = \kappa \oplus \lambda$ and $|\kappa \cdot \lambda| = |\lambda \cdot \kappa| = \kappa \otimes \lambda$. To note the difference between cardinal arithmetic and ordinal arithmetic, we note that in cardinal arithmetic we have that $\omega \oplus 1 = |\omega + 1| = |1 + \omega| = |\omega| = \omega$ and similarly $\omega \otimes 2 = |\omega \cdot 2| = |2 \cdot \omega| = |\omega| = \omega$.

Lemma 40.2. For $n, m \in \omega$, $n \oplus m = n + m < \omega$ and $n \otimes m = n \times m < \omega$

Proof. We show $n + m < \omega$ by induction. When $m = 0, n + 0 = n \in \omega$. If $n + m \in \omega$, then $(n + m) + 1 = n + (m + 1) = n + S(m) = S(n + m) \in \omega$. For $n \times m \in \omega$, if $m = 0$ then $n \cdot 0 = 0 \in \omega$ and if $n \cdot m \in \omega$, then $n \cdot (m + 1) = n \cdot m + n \in \omega$. Since $n \oplus m \approx n + m$ and $n \otimes m \approx n \cdot m$, we apply part (2) of the previous lemma to arrive at the equality. □

Lemma 40.3. Every infinite cardinal is a limit ordinal

Proof. Suppose κ is an infinite cardinal and that $\kappa = \alpha + 1$ for some α. Then $\kappa = |\kappa| = |\alpha + 1| = |1 + \alpha| = |\alpha|$ which contradicts the fact that κ is least. □

Theorem 40.4. If κ is an infinite cardinal, then $\kappa \otimes \kappa = \kappa$.

Proof. We show this by transfinite induction on κ. Assume for all infinite cardinals $\lambda < \kappa$ that $\lambda \otimes \lambda = \lambda$. We claim that for $\alpha < \kappa$, $|\alpha \times \alpha| = |\alpha| \otimes |\alpha| < \kappa$. First, if α is finite, then $|\alpha| = \alpha$, so $|\alpha \times \alpha| = \alpha \otimes \alpha = |\alpha| \otimes |\alpha|$. Next, if α is infinite, then $|\alpha \times \alpha| = |\alpha| = |\alpha| \otimes |\alpha|$ by the inductive hypothesis.

We now define a well-ordering $\lhd$ on $\kappa \times \kappa$, which we then show has order type $\leq \kappa$.

$$\langle \alpha, \beta \rangle \lhd \langle \alpha', \beta' \rangle \Leftrightarrow \max(\alpha, \beta) < \max(\alpha', \beta') \\ \vee\ \max(\alpha, \beta) = \max(\alpha', \beta')\ \wedge\ (\alpha, \beta) \text{ precedes } (\alpha', \beta') \text{ lexicographically}$$

Observe that for $\langle \alpha, \beta \rangle \in \kappa \times \kappa$, $\langle \alpha, \beta \rangle$ has at most $|\max(\alpha, \beta) + 1 \times \max(\alpha, \beta) + 1| < \kappa$ $\lhd$-predecessors in κ. It follows that $\text{type}(\kappa \times \kappa, \lhd) \leq \kappa$, which implies that $|\kappa \times \kappa| \leq \kappa$. Lastly, since $\kappa \leq |\kappa \times \kappa|$, we conclude that $\kappa = |\kappa \times \kappa| = \kappa \otimes \kappa$.

□

Corollary 40.5. Let λ and κ be infinite cardinals.

(1) $\kappa \oplus \lambda = \kappa \otimes \lambda = \max(\kappa, \lambda)$.
(2) $|\kappa^{<\omega}| = \kappa$.

Proof. For (1), note that

$$\max(\kappa, \lambda) \precsim \kappa \oplus \lambda \precsim \kappa \otimes \lambda \precsim \max(\kappa, \lambda) \otimes \max(\kappa, \lambda) = \max(\kappa, \lambda).$$

For (2), for each n we can define by induction a one-to-one map $f_n : \kappa^n \to \kappa$ using the same idea as in the proof of Theorem 40.4 (that is, we can define a well-ordering $\lhd_n$ on $\underbrace{\kappa \times \ldots \times \kappa}_{n \text{ times}}$ and use the fact that $\kappa^n \approx \underbrace{\kappa \times \ldots \times \kappa}_{n \text{ times}}$). Having defined these maps, we then define a one-to-one map $f : \bigcup_{n \in \omega} \kappa^n \to \omega \times \kappa$, which implies that $|\kappa^{<\omega}| \leq \omega \otimes \kappa = \kappa$. □

It is consistent with the axioms considered ths far that ω is the only infinite cardinal. However, the existence of larger infinite cardinals follows from:

Axiom 8: Power Set: $\forall x \exists y \forall z [z \subseteq x \to z \in x]$

Definition 40.6. $\mathcal{P}(x) = \{z : z \subseteq x\}$.

Theorem 40.7 (Cantor)**.** $x \precnsim \mathcal{P}(x)$.

Proof. Given $f : x \to \mathcal{P}(x)$ one-to-one, we show that f is not onto. Let $S = \{x : x \notin f(x)\}$. $S \subseteq x$ but $S \notin \text{ran}(f)$. Indeed, if there is some $z \in x$ such that $f(z) = S$, we have if $z \in f(z)$ if and only if $z \in S$ if and only if $z \notin f(z)$. □

By the axiom of choice, we have $|\mathcal{P}(x)| > |x|$; in particular $|\mathcal{P}(\omega)| > \omega$. However, we can prove the existence of a cardinal larger than ω without using the axiom of choice.

Theorem 40.8. For every α there is some cardinal $\kappa > \alpha$.

Proof. Assume $\alpha \geq \omega$. Let $W = \{R \in \mathcal{P}(\alpha \times \alpha) : R \text{ well-orders } \alpha\}$. Next, let $S = \{\text{type}(\langle \alpha, R \rangle) : R \in W\}$, which exists by replacement.

We claim that $\sup(S)$ is a cardinal. First, $\sup(S)$ is an ordinal by definition of the supremum. Now suppose that there is some $\beta < \sup(S)$ such that $\beta \approx \sup(S)$. Then there is some $\gamma \in S$ such that $\beta \leq \gamma \leq \sup(S)$. Then by Lemma REF, we have $\gamma \approx \sup(S)$. Since γ is the order type of some well-order R on α, γ is bijective with α. Then the composition of the bijection between γ and α and that between γ and $\sup(S)$ induces a well-ordering of α of order type $\sup(S)$. This implies that $\sup(S) \in S$, which is impossible because $\in$ is irreflexive on ordinals.

Lastly, we claim that $\sup(S) > \alpha$. This follows immediately from the fact that the ordering on α given by $\in$ is a well-ordering of order type α, which yields $\alpha \in S$. □

Definition 40.9.

(i) For a given ordinal α, α^+ is the least cardinal strictly greater than α.

(ii) κ is a *successor cardinal* if $\kappa = \alpha^+$ for some α.

(iii) κ is a *limit cardinal* if $\kappa > \omega$ and κ is not a successor cardinal.

Definition 40.10. $\aleph_\alpha = \omega_\alpha$ is defined by transfinite recursion on α as follows:

(1) $\omega_0 = \omega$;

(2) $\omega_{\alpha+1} = (\omega_\alpha)^+$; and

(3) for γ a limit ordinal, $\omega_\gamma = \sup\{\omega_\alpha : \alpha < \gamma\}$.

Lemma 40.11.

(i) For each ordinal α, ω_α is a cardinal.

(ii) For every infinite cardinal κ, there is some α such that $\kappa = \omega_\alpha$.

(iii) If $\alpha < \beta$, then $\omega_\alpha < \omega_\beta$.

(iv) ω_α is a successor cardinal if and only if α is a successor ordinal.

(v) ω_α is a limit cardinal if and only if α is a limit ordinal.

Proof. (i) We proceed by transfinite recursion on α. For $\alpha = 0$, $\omega_0 = \omega$. It is clear that if ω_α is a cardinal, then $\omega_{\alpha+1} = (\omega_\alpha)^+$ is a cardinal.

Suppose that γ is a limit ordinal and that ω_α is a cardinal for all $\alpha < \gamma$. Suppose further that ω_γ is not a cardinal. Then there is some $\xi < \omega_\gamma$ such that $\xi \approx \omega_\gamma$. It follows that there is some $\alpha < \gamma$ such that $\xi \leq \omega_\alpha < \omega_{\alpha+1} < \omega_\gamma$ (where we can assume the second

inequality is strict since ω_α is a cardinal by hypothesis). It follows from Lemma REF that $|\xi| = |\omega_\alpha| = |\omega_{\alpha+1}| = |\omega_\gamma|$, which is impossible.

(ii) We again proceed by transfinite recursion on cardinals. First, $\omega = \omega_0$. Next, suppose that κ is a successor cardinal and for all $\lambda < \kappa$, $\lambda = \omega_\alpha$ for some α. Then there is some $\lambda < \kappa$ such that $\kappa = \lambda^+ = (\omega_\alpha)^+ = \omega_{\alpha+1}$.

Lastly, suppose that κ is a limit cardinal. We claim that for all α, $\alpha \leq \omega_\alpha$, which we show by transfinite recursion on α. Clearly $0 \leq \omega_0$. Suppose that $\alpha \leq \omega_\alpha$. Then $\alpha + 1 \leq \omega_\alpha + 1 < \omega_{\alpha+1}$. Suppose that α is a limit and for all $\beta < \alpha$, $\beta \leq \omega_\beta$. Since $\alpha = \sup\{\beta : \beta < \alpha\}$ and for every $\beta < \alpha$ we have $\beta \leq \omega_\beta \leq \omega_\alpha$, it follows that $\alpha \leq \omega_\alpha$.

By the above claim, we have $\kappa \leq \omega_\kappa$, from which it follows that $\{\gamma : \kappa \leq \omega_\gamma\}$ is non-empty. Let γ be the least ordinal such that $\kappa \leq \omega_\gamma$. We next claim that γ is a limit. Suppose not. Then $\gamma = \xi + 1$ for some ξ. Thus $\kappa \leq \omega_{\xi+1}$. It follows that either $\omega_\xi < \kappa$ or $\kappa \leq \omega_\xi$. The former implies that $\kappa = (\omega_\xi)^+$, which contradicts the fact that κ is a limit cardinal, while the latter contradicts the fact that $\gamma = \xi + 1$ is the least such that $\kappa \leq \omega_{\xi+1}$.

Since γ is a limit, $\omega_\gamma = \sup\{\omega_\alpha : \alpha < \gamma\}$. Note that if $\kappa < \omega_\gamma$, then there is some $\alpha < \gamma$ such that $\kappa \leq \omega_\alpha$, which contradicts our choice of γ. Thus $\kappa = \omega_\gamma$, which completes the proof of (ii).

(iii) We proceed by transfinite recursion on β. The case that $\beta = 0$ is trivial. Suppose for a fixed β and for every $\alpha < \beta$ we have $\omega_\alpha < \omega_\beta$. Then given any $\alpha < \beta + 1$, either $\alpha < \beta$, in which case $\omega_\alpha < \omega_\beta < \omega_{\beta+1}$, or $\alpha = \beta$, in which case $\omega_\alpha = \omega_\beta < \omega_{\beta+1}$. Lastly, if β is a limit, then as $\omega_\beta = \sup\{\omega_\alpha : \alpha < \beta\}$, the result immediately follows.

(iv) Suppose that ω_α is a successor cardinal. Then $\omega_\alpha = \kappa^+$ for some cardinal κ. By (ii), $\kappa = \omega_\beta$ for some β, hence $\omega_\alpha = (\omega_\beta)^+ = \omega_{\beta+1}$. Conversely, if $\alpha = \beta + 1$, then $\omega_\alpha = \omega_{\beta+1} = (\omega_\beta)^+$, which implies that ω_α is a successor cardinal.
(v) This follows immediately from (iv). □

40.1. **Choice and cardinals.**

Lemma 40.12. (AC) If there is a surjective function $f : Y \to X$, then $|Y| \leq |X|$.

Proof. Let R be a well-ordering on X. Define $g : Y \to X$ so that for $y \in Y$,

$$g(y) = \text{ the } R\text{-least element of } f^{-1}(\{y\}).$$

Clearly g is one-to-one, hence $Y \precsim X$. □

Lemma 40.13. (AC) If $\kappa \geq \omega$ and $(X_\alpha)_{\alpha<\kappa}$ is a family of sets satisfying $|X_\alpha| \leq \kappa$ for all $\alpha < \kappa$, then $\left|\bigcup_{\alpha<\kappa} X_\alpha\right| \leq \kappa$.

Proof. Define a well-ordering on $\mathcal{P}(\bigcup_{\alpha<\kappa} X_\alpha \times \kappa)$, which includes all functions from $X_\alpha \to \kappa$ for each $\alpha < \kappa$. For each such α, let f_α be the least function from X_α to κ in this order. Now define $f : \bigcup_{\alpha<\kappa} X_\alpha \to \kappa \times \kappa$ by setting $f(x) = \langle \alpha, f_\alpha(x)\rangle$, where α is the least such that $x \in X_\alpha$. One can readily verify that f is one-to-one, hence $\bigcup_{\alpha<\kappa} X_\alpha \precsim \kappa \times \kappa \precsim \kappa$. □

It is consistent with ZF that $\mathcal{P}(\omega)$ and ω_1 can both be expressed as the countable union of countable sets.

Theorem 40.14. (AC) Let κ be an infinite cardinal. Suppose $B \subseteq A$, $|B| \leq \kappa$, and $\mathcal{S}$ is a set of $\leq \kappa$ finitary functions (i.e. n-ary functions for some $n \in \omega$) on A. Then the closure of B under $\mathcal{S}$, the least $C \subseteq A$ such that $B \subseteq C$ and C is closed under all functions in $\mathcal{S}$) has cardinality $\leq \kappa$.

Proof. (Sketch) Define a sequence $C_0, C_1, \ldots$ such that $C_0 = B$ and $C_{n+1} = C_n \cup \{f''(C_n^{a_f}) : f \in \mathcal{S}\}$, where a_f is the arity of f. One then argues by induction and Lemma 40.13 that $|C_n| \leq \kappa$ for every n and hence $\left|\bigcup_{n\in\omega} C_n\right| \leq \kappa$. □

Example 40.15. Every infinite group contains a countably infinite subgroup: Given an infinite group G, take $B \subseteq G$ with $|B| = \omega$. Then closing under the group operation and inverses, both finitary operations, yields a subgroup that by Theorem 40.14 has cardinality ω.

40.2. **Cardinal exponentiation.**

Definition 40.16. For sets A, B, $A^B := \{f : f \text{ is a function with } \mathrm{dom}(f) = A \text{ and } \mathrm{ran}(f) = B\}$.

We sometimes write A^B as ${}^B A$. Note that A^B exists since $A^B \subseteq \mathcal{P}(B \times A)$.

Definition 40.17. (AC) $\kappa^\lambda := |{}^\lambda\kappa|$.

Hereafter, we will write κ^λ to denote the cardinal and ${}^\lambda\kappa$ to denote the corresponding set of functions.

Lemma 40.18. If $\lambda \geq \omega$ and $2 \leq \kappa \leq \lambda$, then ${}^\lambda\kappa \approx {}^\lambda 2 \approx \mathcal{P}(\lambda)$.

Proof. ${}^\lambda 2 \approx \mathcal{P}(\lambda)$ by identifying sets with their characteristic functions. To complete the proof, we have

$$ {}^\lambda 2 \precsim {}^\lambda\kappa \precsim {}^\lambda\lambda \precsim \mathcal{P}(\lambda \times \lambda) \approx \mathcal{P}(\lambda) \approx {}^\lambda 2,$$

where the second $\approx$ follows from the fact that $A \approx B$ implies $\mathcal{P}(A) \approx \mathcal{P}(B)$. □

Remark. Viewed as ordinal exponentiation, $2^\omega = \omega$; viewed as cardinal exponentiation, we have $2^\omega = |\mathcal{P}(\omega)| > \omega$.

Lemma 40.19. (AC) If κ, λ, and σ are cardinals, then

$$\kappa^{\lambda\oplus\sigma} = \kappa^\lambda \otimes \kappa^\sigma \text{ and } (\kappa^\lambda)^\sigma = \kappa^{\lambda\otimes\sigma}.$$

Proof. Without choice we prove that for sets A, B, and C, if $B \cap C = \emptyset$, then

$${}^{(B\cap C)}A \approx {}^B A \times {}^C A.$$

To establish the bijection, each $f : B \cup C \to A$ can be mapped to $\langle f \restriction B, f \restriction C\rangle$, and conversely, each pair $f_0 : B \to A$ and $f_1 : C \to A$ can be mapped to the function $f : B \cup C \to A$ defined by

$$f(x) = \begin{cases} f_0(x) & \text{if } x \in B \\ f_1(x) & \text{if } x \in C \end{cases}.$$

Applying this line of reasoning to κ, λ, and σ, and using choice, yields the conclusion. Next, without choice, we can prove that ${}^C({}^B A) \approx {}^{C\times B}A$. Given $f : C \to {}^B A$, let f_c be the function $f(c)$. Then to establish the bijection, each such function f is mapped to the g such that $g(c, b) = f_c(b)$. Conversely, each $h \in {}^{C\times B}A$ can be mapped to the function g such that $g(c) = h(c, \cdot) : B \to A$. □

Definition 40.20. (AC)
(i) Continuum Hypothesis (CH): $2^{\omega} = \omega_1$
(ii) Generalized Continuum Hypothesis (GCH): $2^{\omega_\alpha} = \omega_{\alpha+1}$

Assuming GCH, we can determine the value of κ^λ, which is the goal of the remainder of this section. First we introduce the notion of cofinality.

Definition 40.21.
(i) For a function $f : \alpha \to \beta$, f maps α *cofinally* if $\mathrm{ran}(f)$ is unbounded in β.
(ii) The cofinality of an ordinal β, denoted $\mathrm{cf}(\beta)$, is the least α such that there is a map from α cofinally into β.

Remark.
(1) $\mathrm{cf}(\beta) \leq \beta$ for every β.
(2) If β is a successor, then $\mathrm{cf}(\beta) = 1$.

Lemma 40.22. There is a cofinal map $f : \mathrm{cf}(\beta) \to \beta$ that is strictly increasing, i.e., $\xi < \eta$ implies $f(\xi) < f(\eta)$.

Proof. Let $g : \mathrm{cf}(\beta) \to \beta$ be cofinal. We define f recursively by

$$f(\eta) = \max\{g(\eta), \sup\{f(\xi) + 1 : \xi < \eta\}\}.$$

□

Lemma 40.23. If α is a limit ordinal and $f : \alpha \to \beta$ is a strictly increasing cofinal map, then $\mathrm{cf}(\alpha) = \mathrm{cf}(\beta)$.

Proof. First we show that $\mathrm{cf}(\beta) \leq \mathrm{cf}(\alpha)$ by mapping $\mathrm{cf}(\alpha)$ cofinally into β. Ket $g : \mathrm{cf}(\alpha) \to \alpha$ map $\mathrm{cf}(\alpha)$ cofinally into α. Then one can verify that $f \circ g$ maps $\mathrm{cf}(\alpha)$ cofinally into β.

Next, we show that $\mathrm{cf}(\alpha) \leq \mathrm{cf}(\beta)$. Let $g : \mathrm{cf}(\beta) \to \beta$ be cofinal. Then we define $h : \mathrm{cf}(\beta) \to \alpha$ by setting

$$h(\xi) = \text{ the least } \eta \text{ such that } f(\eta) > g(\xi).$$

Note that if $\mathrm{ran}(h)$ is bounded in α, then the domain of f is bounded in α. But since f is strictly increasing and cofinal, f cannot map some initial segment of α cofinally into β. Thus, h is cofinal, and the conclusion follows. □

By applying the previous two lemmas, we have:

Corollary 40.24. $\mathrm{cf}(\mathrm{cf}(\beta)) = \mathrm{cf}(\beta)$ for all β.

Definition 40.25. β is regular if β is a limit ordinal and $\mathrm{cf}(\beta) = \beta$.

Corollary 40.24 implies that $\mathrm{cf}(\beta)$ is regular for all limit ordinals β.

Lemma 40.26. If β is regular then β is a cardinal.

Proof. Suppose that $|\beta| < \beta$. Then we can map $|\beta|$ cofinally into β, which implies that $\mathrm{cf}(\beta) \leq |\beta| < \beta$, a contradiction. □

Lemma 40.27.
(i) ω is regular.
(ii) (AC) κ^+ is regular.

Proof. (i) is immediate. For (ii), suppose there is some $\alpha < \kappa^+$ and a cofinal $f : \alpha \to \kappa^+$. Then $\kappa^+ = \bigcup\{f(\xi) : \xi < \alpha\}$. As this is a union of $\leq \kappa$ many sets of cardinality $\leq \kappa$, it follows from Lemma 40.13 that $|\kappa^+| \leq \kappa$, which is impossible. □

It is consistent with ZF that $\mathrm{cf}(\omega_1) = \omega$. Moreover, it is open whether ZF can prove the existence of a cardinality of cardinality greater than ω.

Infinite cardinals that are not regular are called *singular*. For instance, ω_ω is singular: $\omega_\omega = \sup\{\omega_n : n \in \omega\}$, which implies that $\mathrm{cf}(\omega_\omega) = \omega$. More generally, we have:

Lemma 40.28. ?? If α is a limit ordinal, then $\mathrm{cf}(\omega_\alpha) = \mathrm{cf}(\alpha)$.

Proof. We define a cofinal map $f : \alpha \to \mathrm{cf}(\omega_\alpha)$ by setting $f(\gamma) = \omega_\gamma$ for each $\gamma < \alpha$. Since f is strictly increasing, it follows from Lemma 40.23 that $\mathrm{cf}(\omega_\alpha) = \mathrm{cf}(\alpha)$. □

Remark. The condition that $\omega_\alpha = \alpha$ is necessary but not sufficient for α to be regular. For necessity, observe that if ω_α is a regular limit cardinal, then by regularity, $\mathrm{cf}(\omega_\alpha) = \omega_\alpha$ and by Lemma **??**, $\mathrm{cf}(\omega_\alpha) = \mathrm{cf}(\alpha)$. Thus we have $\mathrm{cf}(\omega_\alpha) = \alpha$, which implies that $\omega_\alpha \leq \alpha$. However, in the proof of Lemma 40.11(ii), we showed that $\alpha \leq \omega_\alpha$. Thus $\omega_\alpha = \alpha$.

To see that the condition $\omega_\alpha = \alpha$ is not sufficient for α to be regular, we define a sequence $(\sigma_n)_{n\in\omega}$ of cardinals as follows:

- $\sigma_0 = \omega$; and
- $\sigma_{n+1} = \omega_{\sigma_n}$.

Let $\alpha = \sup\{\sigma_n : n \in \omega\}$. Then $\omega_\alpha = \sup\{\omega_{\sigma_n} : n \in \omega\} = \sup\{\sigma_n : n \in \omega\} = \alpha$. However, $\mathrm{cf}(\omega_\alpha) = \omega$, and hence α is singular.

The previous example shows that a regular cardinal must be very large. This motivates the following definition.

Definition 40.29.

(1) κ is *weakly inaccessible* if κ is a regular limit cardinal.
(2) (AC) κ is *strongly inaccessible* if $\kappa > \omega$, κ is regular, and for all $\lambda < \kappa$, $2^\lambda < \kappa$.

Clearly every strongly inaccessible cardinal is weakly inaccessible. Under the assumption of GCH, every weakly inaccessible cardinal is strongly inaccessible. To see this, given a weakly inaccessible cardinal κ, suppose that $2^{\omega_\alpha} \geq \kappa$ for some $\omega_\alpha < \kappa$. It follows by GCH that $2^{\omega_\alpha} = \omega_{\alpha+1} = (\omega_\alpha)^+$, which implies that $\kappa = (\omega_\alpha)^+$, a successor cardinal. This contradicts the fact that κ is a limit cardinal.

ZFC does not prove that weakly inaccessible cardinals exist. However, it is consistent with ZFC that 2^ω is weakly inaccessible, and it is consistent with ZFC that 2^ω is greater than the first weakly inaccessible cardinal.

Lemma 40.30 (König)**.** (AC) If κ is infinite and $\mathrm{cf}(\kappa) \leq \lambda$, then $\kappa^\lambda > \kappa$.

Proof. Fix any cofinal $f : \lambda \to \kappa$. Let $G : \kappa \to {}^\lambda\kappa$. We claim that G is not onto, which we will show by defining a function $h : \lambda \to \kappa$ such that $h \notin \mathrm{ran}(G)$. We set

$$h(\alpha) := \text{ the least element of } \kappa \setminus \{G(\mu)(\alpha) : \mu < f(\alpha)\}.$$

Suppose now that $h \in \mathrm{ran}(G)$. Then there is some $\mu < \kappa$ such that $G(\mu) = h$. Since f is cofinal, there is some $\alpha < \lambda$ such that $f(\alpha) > \mu$. Thus $h(\alpha) = G(\mu)(\alpha)$ and $h(\alpha) \in \kappa \setminus \{G(\mu)(\alpha) : \mu < f(\alpha)\}$, which yields a contradiction. Thus $h \notin \mathrm{ran}(G)$. □

The following is immediate:

Corollary 40.31. $(\omega_\omega)^\omega > \omega_\omega$.

Corollary 40.32. (AC) If $\lambda \geq \omega$, then $\mathrm{cf}(2^\lambda) > \lambda$.

Proof. By the contrapositive of Lemma 40.30, $\kappa^\lambda \leq \kappa$ implies that $\mathrm{cf}(\kappa) > \lambda$. Now,

$$(2^\lambda)^\lambda = 2^{\lambda \otimes \lambda} = 2^\lambda,$$

hence $(2^\lambda)^\lambda \leq 2^\lambda$. By our initial observation, it follows that $\mathrm{cf}(2^\lambda) > \lambda$. □

Lemma 40.33. (AC+GCH) Assume that $\kappa, \lambda \geq 2$ and at least one is infinite. Then

(1) $\kappa \leq \lambda$ implies $\kappa^\lambda = \lambda^+$;
(2) $\kappa > \lambda \geq \mathrm{cf}(\kappa)$ implies $\kappa^\lambda = \kappa^+$; and
(3) $\lambda < \mathrm{cf}(\kappa)$ implies $\kappa^\lambda = \kappa$.

Proof. (1) Since $2 \leq \kappa \leq \lambda$, by Lemma 40.18 we have

$$\kappa^\lambda = 2^\lambda = \lambda^+,$$

where the second equality follows from GCH.

(2) By Lemma 40.30, $\kappa > \lambda \geq \mathrm{cf}(\kappa)$ implies that $\kappa^\lambda > \kappa$. Thus

$$\kappa < \kappa^\lambda \leq \kappa^\kappa = 2^\kappa = \kappa^+,$$

where the last equality follows from GCH. Thus we have $\kappa^\lambda = \kappa^+$.

(3) If $\lambda < \mathrm{cf}(\kappa)$, then every function in ${}^\lambda\kappa$. Thus we have ${}^\lambda\kappa = \bigcup\{{}^\lambda\alpha : \alpha < \kappa\}$. We claim that for each $\alpha < \kappa$, $|{}^\lambda\alpha| \leq \max(\alpha, \lambda)^+ \leq \kappa$. We consider two cases:

Case 1: $\alpha \leq \lambda$. In this case $|{}^\lambda\alpha| \leq |{}^\lambda\lambda| = \lambda^\lambda = 2^\lambda = \lambda^+ \leq \kappa$.

Case 2: $\lambda \leq \alpha$. First observe that $\alpha^+ = |\alpha|^+$. Now, we have

$$|{}^\lambda\alpha| \leq |{}^\alpha\alpha| = |{}^\alpha 2| = |\mathcal{P}(\alpha)| = |\mathcal{P}(|\alpha|)| = 2^{|\alpha|} = |\alpha|^+ = \alpha^+ \leq \kappa.$$

Thus the claim follows. From the claim it follows that ${}^\lambda\kappa$ can be expressed as the union of $< \kappa$ many sets of size at most κ, which implies that $\kappa^\lambda = \kappa$. □

Definition 40.34. (AC) $\beth_\alpha$ is defined by transfinite recursion on α.

(1) $\beth_0 = \omega$;
(2) $\beth_{\alpha+1} = 2^{\beth_\alpha}$; and
(3) for γ limit, $\beth_\gamma = \sup\{\beth_\alpha : \alpha < \gamma\}$.

We can recast the GCH as the statement that $\forall\alpha(\beth_\alpha = \omega_\alpha)$.