

MATH681
Mathematical Logic

Pingbang Hu

February 26, 2023

Abstract

This is a graduate-level mathematical logic course taught by [Matthew Harrison-Trainor](#), aiming to obtain insights into all other branches of mathematics, such as algebraic geometry, analysis, etc. Specifically, we will cover model theory beyond the basic foundational ideas of logic.

While there are no required textbooks, some books do cover part of the material in the class. For example, Marker's *Model Theory: An Introduction* [[Mar02](#)], Hodges's *A Shorter Model Theory* [[HH97](#)], and Hinman's *Fundamentals of Mathematical Logic* [[Hin05](#)].



Contents

1	Language, Logic, and Structures	2
1.1	Syntax and Semantics	3
1.2	Theories	11
1.3	Completeness and Compactness	14
2	The Beginning of Model Theory	28
2.1	Complete Theories	28
2.2	Up and Down	37
2.3	Back and Forth	40

Chapter 1

Language, Logic, and Structures

Lecture 1: Introduction to Mathematical Logic

The goal of mathematical logic is to obtain insights into other areas of mathematics – algebra, analysis, combinatorics, and so on, by formalizing the **process** of mathematics. 5 Jan. 14:30

Remark. More concretely, there are different branches:

- (a) Model Theory: Study subsets of an object defined by a **formula** (i.e., first-order logic).
- (b) Computability Theory / Recursion Theory: Formalizing what it means to have an algorithm and studying relative computability.
- (c) Set Theory: Study the structure of the mathematical universe.
- (d) Proof Theory: Study the syntactic nature of **proofs**.

In this class, we study model theory in nature; specifically, we will cover

- basic definitions of logic:
 - What is a **formula**?
 - What does it mean for a **formula** to be **true**?
 - What is a **proof**?
- **Soundness** & completeness theorems:
 - Anything **provable** is **true**.
 - Anything **true** is **provable**.
- Compactness theorem:
 - Non-standard objects exist.
- Using compactness theorem for applications:
 - **Chevalley's theorem**.

The main theme of this course will be *syntax* v.s. *semantics*:

Syntax	v.s.	Semantics
proofs		truth
form of a formula		mathematical structures
number and type of quantifiers		isomorphisms, embeddings

1.1 Syntax and Semantics

1.1.1 Languages and Structures

Let's start with the fundamental object, [language](#).

Definition 1.1.1 (Language). A *language* \mathcal{L} consists of:

- a set \mathcal{F} of function symbols f with arities n_f ;
- a set \mathcal{R} of relation symbols R with arities n_R ;
- a set \mathcal{C} of constant symbols c .

A [language](#) is also sometimes called a *signature*, in which case we use σ rather than \mathcal{L} .

Note. A constant is the same as a 0-ary function.

Remark. Any or all sets in [Definition 1.1.1](#) might be empty.

Example (Graph). The [language](#) of graphs, $\mathcal{L}_{\text{graph}} = \{E\}$ where E is a binary (2-ary) relation symbol.

Example (Ring). The [language](#) of rings, $\mathcal{L}_{\text{ring}} = \{0, 1, +, \cdot, -\}$, where $0, 1$ are constants, $+, \cdot$ are binary functions, and $-$ is a unary function.

Example (Ordered ring). The [language](#) of ordered rings, $\mathcal{L}_{\text{ord}} = \mathcal{L}_{\text{ring}} \cup \{\leq\}$ where \leq is the binary relation for an ordered ring.

Then, given a [language](#), we can now interpret it in the following way.

Definition 1.1.2 (Structure). Given a [language](#) \mathcal{L} , an \mathcal{L} -*structure* \mathcal{M} consists of:

- a non-empty set M called the *universe*, *domain*, or *underlying set* of \mathcal{M} ;
- for each function symbol $f \in \mathcal{F}$, a function $f^{\mathcal{M}}: M^{n_f} \rightarrow M$;
- for each relation symbol $R \in \mathcal{R}$, a relation $R^{\mathcal{M}} \subseteq M^{n_R}$;
- for each constant symbol $c \in \mathcal{C}$, an element $c^{\mathcal{M}} \in M$.

Notation (Interpretation). The *interpretation* of symbols f, R, c in \mathcal{M} is $f^{\mathcal{M}}, R^{\mathcal{M}}, c^{\mathcal{M}}$, respectively.

Basically, a [structure](#) gives meaning to the symbols from the [language](#), and we often write

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

Notation. We usually use $\mathcal{M}, \mathcal{N}, \dots, \mathcal{A}, \mathcal{B}, \dots$ to refer to [structures](#), and M, N, \dots, A, B, \dots for the domains.^a

^aSome people use $|\mathcal{M}|$ for the domain of \mathcal{M} .

It's time to look at some examples.

Example. The rationals \mathbb{Q} and integers \mathbb{Z} are both $\mathcal{L}_{\text{ring}}$ -structures.

Proof. Clearly, the domain is the set of rationals, and naively, we let $+^{\mathbb{Q}} = +$ in \mathbb{Q} , $0^{\mathbb{Q}} = 0$ in

\mathbb{Q} , $1^{\mathbb{Q}} = 1$ in \mathbb{Q} , etc. In this way, $\mathbb{Q} = (\mathbb{Q}, 0, 1, +, \cdot, -)$ is an $\mathcal{L}_{\text{ring}}$ -structure. Similarly, $\mathbb{Z} = (\mathbb{Z}, 0, 1, +, \cdot, -)$ is as well. \circledast

While the language we have seen are all intuitively correct with their name, e.g., $\mathcal{L}_{\text{ring}}$, \mathcal{L}_{ord} , and $\mathcal{L}_{\text{graph}}$, they are really just the high-level abstraction of the objects in the subscript.

Example. Nothing forces an $\mathcal{L}_{\text{ring}}$ -structure to be a ring.

Proof. Since an $\mathcal{L}_{\text{ring}}$ -structure is just any structure with two binary functions, a unary function, and two constants interpreting the symbols of the language; hence we can define an $\mathcal{L}_{\text{ring}}$ -structure \mathcal{M} as

- $\mathcal{M} = \{0, 5, 11\}$;
- $0^{\mathcal{M}} = 5$;
- $1^{\mathcal{M}} = 11$;
- $+^{\mathcal{M}}$ is the constant function 0;
- $\cdot^{\mathcal{M}}$ is the function 5;
- $-^{\mathcal{M}}$ is the identity.

This is clearly not a ring since it fails nearly every axiom of a ring. \circledast

Note. Later, we will talk about theories that let us restrict to structures we want.

1.1.2 Embeddings and Isomorphisms

We can now consider the relation between structures.

Definition 1.1.3 (Embedding). Let \mathcal{M} and \mathcal{N} be \mathcal{L} -structures. A map $\eta: \mathcal{M} \rightarrow \mathcal{N}$ is an \mathcal{L} -embedding if it is one-to-one and preserves the interpretation of all symbols of \mathcal{L} :

- (a) for each function symbol $f \in \mathcal{F}$ of arity n_f , and $a_1, \dots, a_{n_f} \in M$,

$$\eta(f^{\mathcal{M}}(a_1, \dots, a_{n_f})) = f^{\mathcal{N}}(\eta(a_1), \dots, \eta(a_{n_f}));$$

- (b) for each relation symbol $R \in \mathcal{R}$ of arity n_R , and $a_1, \dots, a_{n_R} \in M$,

$$(a_1, \dots, a_{n_R}) \in R^{\mathcal{M}} \Leftrightarrow (\eta(a_1), \dots, \eta(a_{n_R})) \in R^{\mathcal{N}};$$

- (c) for each constant symbol $c \in \mathcal{C}$, $c^{\mathcal{M}} = c^{\mathcal{N}}$.

From the definition, an \mathcal{L} -embedding is an injection, and naturally, we have the following.

Definition 1.1.4 (Isomorphism). An \mathcal{L} -isomorphism is a bijective \mathcal{L} -embedding.

Definition 1.1.5 (Automorphism). An \mathcal{L} -automorphism of \mathcal{M} is an \mathcal{L} -isomorphism from \mathcal{M} to \mathcal{M} .

Definition. Let \mathcal{M} and \mathcal{N} be \mathcal{L} -structures. Suppose $M \subseteq N$ and the inclusion map $\iota: \mathcal{M} \hookrightarrow \mathcal{N}$ is an \mathcal{L} -embedding.

Definition 1.1.6 (Substructure). \mathcal{M} is a substructure of \mathcal{N} .

Definition 1.1.7 (Extension). \mathcal{N} is an extension of \mathcal{M} .

Example. Ring embeddings are $\mathcal{L}_{\text{ring}}$ -embeddings.

This generalizes the notions of embedding and isomorphism for many mathematical structures.

Remark. Asking that η be injective is the same as (b) in Definition 1.1.3 for the relation $=$ since

$$a = b \in \mathcal{M} \Leftrightarrow \eta(a) = \eta(b) \in \mathcal{N}.$$

The notion of substructure is language sensitive. For groups, there are two possible languages:

- (a) $\mathcal{L}_1 = \{e, \cdot\}$;
- (b) $\mathcal{L}_2 = \{e, \cdot, {}^{-1}\}$, i.e., with the unary inverse operation.

While both seem valid at the first glance, we should use the second one.

To see why, if we use \mathcal{L}_2 , the substructure of a group is the same thing as a subgroup. But if we use \mathcal{L}_1 , then $(\mathbb{N}, +, 0)$ is a substructure of $(\mathbb{Z}, +, 0)$, while \mathbb{N} is not a group for sure.¹

Similarly, we include $-$ in $\mathcal{L}_{\text{ring}}$ for a similar reason as in the previous example.

Example. An $\mathcal{L}_{\text{ring}}$ -substructure of a field will be a subring, not a subfield. If we want subfields, use $\mathcal{L}_{\text{ring}} \cup \{{}^{-1}\}$.^a

^aWe can set $0^{-1} = 0$, but never use this.

Lecture 2: Formulas and First-Order Logic

We start by asking that given a function symbol f of arity n , could we replace f with an $(n+1)$ -ary R relation to represent its graph? 10 Jan. 14:30

Example. Let \mathcal{L} be a language with only relation symbols. Let \mathcal{A} be an \mathcal{L} -structure. For any $B \subseteq A$, there is a substructure \mathcal{B} of \mathcal{A} with domain B .

Proof. For each relation symbol R , letting $R^{\mathcal{B}} = R^{\mathcal{A}} \cap B^{n_R}$ will make \mathcal{B} a substructure of \mathcal{A} . \circledast

The above is not true for function symbols though.

Example. If $G = (\mathbb{Z}, 0, +)$, then \mathbb{N} is not the domain of a subgroup. So if we took $\mathcal{L} = \{0, +, {}^{-1}\}$, where 0 is the unary relation, $+$ is the ternary relation, and ${}^{-1}$ is the binary relation, an \mathcal{L} -substructure of a group might not be a subgroup.

1.1.3 Terms

Intuitive, an \mathcal{L} -formula is an expression built using the symbols in a language \mathcal{L} , $=$, the logical connectives \wedge, \vee, \neg , and variable symbols v_1, v_2, \dots, x, y, z , and also quantifiers \exists and \forall .

Definition 1.1.8 (Term). Given a language \mathcal{L} , the set of \mathcal{L} -terms are defined inductively by:

- (a) each constant symbol is a term;
- (b) each variable symbol v_1, \dots is a term;
- (c) if f is a function symbol, and t_1, \dots, t_{n_f} are terms, then $f(t_1, \dots, t_{n_f})$ is a term.

If \mathcal{M} is an \mathcal{L} -structure, and t is a term involving only variables among v_1, \dots, v_n , then t has an interpretation $t^{\mathcal{M}}: M^n \rightarrow M$ as a function as follows. On input $a_1, \dots, a_n \in M$,

- (a) if t is a constant c , $t^{\mathcal{M}}(a_1, \dots, a_n) = c^{\mathcal{M}}$.
- (b) if t is a variable v_i , $t^{\mathcal{M}}(a_1, \dots, a_n) = v_i$;

¹Simply observe that both $(\mathbb{N}, 0, +), (\mathbb{Z}, 0, +)$ are \mathcal{L}_1 -structures.

(c) if t is $f(s_1, \dots, s_k)$, then $t^{\mathcal{M}}(a_1, \dots, a_n) = f^{\mathcal{M}}(s_1^{\mathcal{M}}(a_1, \dots, a_n), \dots, s_k^{\mathcal{M}}(a_1, \dots, a_n))$.

Intuition. We are basically substituting for variables and evaluating the expression.

Example. In $(\mathbb{R}, 0, 1, +, \cdot, -)$, a **term** is essentially just a polynomial with integer coefficients, assuming we interpret them in a ring. Technically, a **term** looks like

$$\cdot(+(1, 1), +(x, y)),$$

but we will write **terms** the natural way, i.e.,

$$(1 + 1)(x + y).$$

Also, we will use \underline{n} or n to represent the **term** $\underline{n} = \underbrace{1 + 1 + \dots + 1}_{n \text{ times}}$. So we could write the above **term** as $2 \cdot (x + y)$.

1.1.4 Formulas

Definition 1.1.9 (Formula). The set of \mathcal{L} -formulas is defined inductively:

- (a) If s, t are **terms**, then $s = t$ is a *formula*.
- (b) If R is a relation symbol of arity n_R and s_1, \dots, s_{n_R} are **terms**, then $R(s_1, \dots, s_{n_R})$ is a *formula*.
- (c) If f is a **formula**, then $\neg f$ is a *formula*.
- (d) If φ and ψ are **formulas**, then $\varphi \wedge \psi$ and $\varphi \vee \psi$ are *formulas*.
- (e) If φ is a **formula** and v_i are variables, then $\exists v_i \varphi$ and $\forall v_i \varphi$ are *formulas*.

Notation (Atomic formula). **Definition 1.1.9 (a)** and **(b)** are called *atomic formulas*.

Notation (Quantifier-free formula). **Definition 1.1.9 (a)**, **(b)**, **(c)**, and **(d)** are called *quantifier-free*.

This logic is called *first-order logic* (FO logic), since the quantifiers range over elements of the **structures**, but not over, e.g., subsets.

Example. We can say that an element x of a ring has a square root by $\exists y \ y^2 = x$.

Example. A group is torsion of order 2 can be said by $\forall x \ x \cdot x = e$.

Example. We can write down all the field/group/... axioms as **formulas**.

Notice that for the first example, the **formula** $\exists y \ y^2 = x$ only has meaning if we assign what x is. In this case, we say that y is *bound* by $\exists y$. But this is local:

Example. Consider

$$y = 1 \wedge \exists y \ y^2 = x,$$

while the first appearance of y is free, the second appearance of y is bound by (in the scope of) $\exists y$.

While our definitions work perfectly fine with the above example, but sometimes we don't want this to happen. In such a case, we simply replace the bound instances of y with a new variable z . This idea of variables being free or bound is defined formally as follows.

Definition 1.1.10 (Free variable). The *free variables* $\text{FV}(\varphi)$ of a **formula** φ are defined inductively:

- (a) $\text{FV}(s = t)$ is the set of variables showing up in s or t .
- (b) $\text{FV}(R(s_1, \dots, s_{n_R}))$ is the set of variables showing up in s_1, \dots, s_{n_R} .
- (c) $\text{FV}(\neg\varphi) = \text{FV}(\varphi)$.
- (d) $\text{FV}(\varphi \wedge \psi) = \text{FV}(\varphi \vee \psi) = \text{FV}(\varphi) \cup \text{FV}(\psi)$.
- (e) $\text{FV}(\exists x \varphi) = \text{FV}(\forall x \varphi) = \text{FV}(\varphi) \setminus \{x\}$.

Example. $\text{FV}(\exists y y^2 = x) = \{x\}$.

Example. $\text{FV}(\forall x x \cdot x = e) = \emptyset$.

Definition 1.1.11 (Sentence). A **formula** φ is called a *sentence* if it has no **free variables**.

Notation. If φ is a **formula** with **free variables** among x_1, \dots, x_n we often write $\varphi(x_1, \dots, x_n)$.

Remark. So given $\varphi(x_1, \dots, x_n)$, we know that φ has no other **free variables** than x_1, \dots, x_n .

Example. It's valid to write $\varphi(x, y, z) := x = y$.

1.1.5 Truths

Finally, we define the notion of **truth**.

Definition 1.1.12 (Truth). Given an \mathcal{L} -**structure** \mathcal{M} , let $\varphi(x_1, \dots, x_n)$ be an \mathcal{L} -**formula** and let $a_1, \dots, a_n \in M$. Then we say φ is *true* of \bar{a} in \mathcal{M} ,^a denoted as $\mathcal{M} \models \varphi(\bar{a})$, as follows:

- (a) If φ is $s = t$, then $\mathcal{M} \models \varphi(\bar{a})$ if $s^{\mathcal{M}}(\bar{a}) = t^{\mathcal{M}}(\bar{a})$.
- (b) If φ is $R(t_1, \dots, t_{n_R})$, then $\mathcal{M} \models \varphi(\bar{a})$ if $(t_1^{\mathcal{M}}(\bar{a}), \dots, t_{n_R}^{\mathcal{M}}(\bar{a})) \in R^{\mathcal{M}}$.
- (c) If φ is $\neg\psi$, then $\mathcal{M} \models \varphi(\bar{a})$ if $\mathcal{M} \not\models \psi(\bar{a})$.
- (d) If φ is $\psi_1 \wedge \psi_2$, then $\mathcal{M} \models \varphi(\bar{a})$ if $\mathcal{M} \models \psi_1(\bar{a})$ and $\mathcal{M} \models \psi_2(\bar{a})$.
- (e) If φ is $\psi_1 \vee \psi_2$, then $\mathcal{M} \models \varphi(\bar{a})$ if $\mathcal{M} \models \psi_1(\bar{a})$ or $\mathcal{M} \models \psi_2(\bar{a})$.
- (f) If φ is $\exists y \psi(\bar{x}, y)$, then $\mathcal{M} \models \varphi(\bar{a})$ if there's $b \in M$ such that $\mathcal{M} \models \psi(\bar{a}, b)$.
- (g) If φ is $\forall y \psi(\bar{x}, y)$, then $\mathcal{M} \models \varphi(\bar{a})$ if for all $b \in M$ such that $\mathcal{M} \models \psi(\bar{a}, b)$.

^aOr \mathcal{M} satisfies $\varphi(\bar{a})$.

Remark. Every **formula** is **true**, or its negation is.

Lecture 3: Logical Consequence and Equivalence

Notation (Material implication). The *material implication* $\varphi \rightarrow \psi$ between two **formulas** φ, ψ is an abbreviation of $\neg\varphi \vee \psi$.

12 Jan. 14:30

Notation. We use $\varphi \leftrightarrow \psi$ as an abbreviation of $((\varphi \rightarrow \psi) \wedge (\psi \rightarrow \varphi))$.

Essentially, \rightarrow and \leftrightarrow is different from \Rightarrow and \Leftrightarrow , where the former are only shown in [formula](#). Now, consider the [language of graphs](#) $\mathcal{L}_{\text{graph}} = \{E\}$, let's see some examples.

Example. An undirected graph can be written as

$$\forall x \forall y (xEy \rightarrow yEx).$$

Example. A vertex has at least three neighbors can be written as

$$\varphi(x) := \exists u \exists v \exists w (xEu \wedge xEv \wedge xEw \wedge u \neq v \wedge v \neq w \wedge u \neq w)$$

in non-reflexive graphs.

Example. For a vertex has exactly three neighbors,

$$\psi(x) := \exists u \exists v \exists w \forall y (xEu \wedge xEv \wedge xEw \wedge u \neq v \wedge v \neq w \wedge u \neq w \wedge (y = u \vee y = v \vee y = w \vee \neg yEx)).$$

Problem. Can we say that x has an even number of neighbors?

Answer. We can't. Some things are not expressible in FO logic. ⊛

Example. For a vertex x has a path of length 4 to y ,

$$\Theta(x, y) := \exists u \exists v \exists w (xEu \wedge uEv \wedge vEw \wedge wEy).$$

We can also express that there is a path of length at most 4.

Problem. Can we say that there is a path from x to y ?

Answer. We still can't! Not in FO logic (using [compactness theorem](#)). ⊛

Remark. When we prove results by induction on [formulas](#), we only need to prove for \neg, \wedge, \exists , instead of for both \wedge, \vee , and both \exists and \forall .

Proof. Since we can view $\varphi \vee \psi$ as an abbreviation for $\neg(\neg\varphi \wedge \neg\psi)$ and $\forall x \varphi$ as an abbreviation for $\neg(\exists x \neg\varphi)$. ⊛

Remark (Sheffer stroke). In fact, we can get \wedge, \vee, \neg from one logical connective, e.g., the *sheffer stroke* \uparrow , which is defined as

$$\varphi \uparrow \psi := \neg(\varphi \wedge \psi),$$

and we can use \uparrow to define \neg, \vee, \wedge .

Notation. Let Φ be a (possibly infinite) set of [sentences](#), we write $\mathcal{M} \models \Phi$ if $\mathcal{M} \models \varphi$ for all $\varphi \in \Phi$.

Definition 1.1.13 (Logical consequence). Let Φ be a set of [sentences](#), and φ be a [sentence](#). We say that φ is a *logical consequence* of Φ , written $\Phi \models \varphi$, if $\mathcal{M} \models \varphi$ whenever $\mathcal{M} \models \Phi$.

If $\Phi = \emptyset$ is the empty set, [Definition 1.1.13](#) is written as $\models \varphi$, i.e., φ is [true](#) in all \mathcal{L} -structures.²

²Recall that we always have a [language](#) \mathcal{L} implicitly.

Definition 1.1.14 (Equivalent). Given two formulas φ, ψ , $\varphi(\bar{x})$ and $\psi(\bar{x})$ are *equivalent* if

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

Problem. Two sentences φ and ψ are *equivalent* if and only if $\varphi \models \psi$ and $\psi \models \varphi$.

DIY

As previously seen. \mathcal{A} is a *substructure* of \mathcal{B} , or $\mathcal{A} \subseteq \mathcal{B}$, means that $A \subseteq B$ and $\text{id}: A \hookrightarrow B$ is an \mathcal{L} -embedding.

Proposition 1.1.1. Suppose that \mathcal{A} is a *substructure* of \mathcal{B} , and $\varphi(\bar{x})$ is a *quantifier-free formula*. Let $\bar{a} \in \mathcal{A}$,^a then $\mathcal{A} \models \varphi(\bar{a})$ if and only if $\mathcal{B} \models \varphi(\bar{a})$.

^aFormally, we need to write \mathcal{A} to be the Cartesian product with a fixed length.

Proof. We start with *terms* by proving that if t is a *term* and $\bar{b} \in \mathcal{A}$, then $t^{\mathcal{A}}(\bar{b}) = t^{\mathcal{B}}(\bar{b})$. The proof is induction on *terms*.

- (a) If t is a constant symbol c , then $t^{\mathcal{A}}(\bar{b}) = c^{\mathcal{A}} = c^{\mathcal{B}} = t^{\mathcal{B}}(\bar{b})$.
- (b) If t is a variable x_i , then $t^{\mathcal{A}}(\bar{b}) = b_i = t^{\mathcal{B}}(\bar{b})$.
- (c) If t is a function symbol $f(s_1, \dots, s_n)$ where s_i are *terms*, then $t^{\mathcal{A}}(\bar{b}) = f^{\mathcal{A}}(s_1^{\mathcal{A}}(\bar{b}), \dots, s_n^{\mathcal{A}}(\bar{b}))$.
By the induction hypothesis, $s_i^{\mathcal{A}}(\bar{b}) = s_i^{\mathcal{B}}(\bar{b}) \in \mathcal{A}$, and hence

$$t^{\mathcal{B}}(\bar{b}) = f^{\mathcal{B}}(s_1^{\mathcal{B}}(\bar{b}), \dots, s_n^{\mathcal{B}}(\bar{b})) = f^{\mathcal{A}}(s_1^{\mathcal{A}}(\bar{b}), \dots, s_n^{\mathcal{A}}(\bar{b})) = t^{\mathcal{A}}(\bar{b}),$$

i.e., $f^{\mathcal{B}} \upharpoonright_{\mathcal{A}} = f^{\mathcal{A}}$, so $t^{\mathcal{A}}(\bar{b}) = t^{\mathcal{B}}(\bar{b})$.

Now we turn to *formulas*, and prove that for φ *quantifier-free*, then $\mathcal{A} \models \varphi(\bar{a}) \Leftrightarrow \mathcal{B} \models \varphi(\bar{a})$ for $\bar{a} \in \mathcal{A}$. The proof is, again, induction on *formulas*.^a

- (a) If φ is $s = t$, then $s^{\mathcal{A}}(\bar{a}) = s^{\mathcal{B}}(\bar{a})$ and $t^{\mathcal{A}}(\bar{a}) = t^{\mathcal{B}}(\bar{a})$, so

$$\mathcal{A} \models \varphi(\bar{a}) \Leftrightarrow s^{\mathcal{A}}(\bar{a}) = t^{\mathcal{A}}(\bar{a}) \Leftrightarrow s^{\mathcal{B}}(\bar{a}) = t^{\mathcal{B}}(\bar{a}) \Leftrightarrow \mathcal{B} \models \varphi(\bar{a}).$$

- (b) If φ is $R(s_1, \dots, s_n)$, then

$$\mathcal{A} \models \varphi(\bar{a}) \Leftrightarrow (s_1^{\mathcal{A}}(\bar{a}), \dots, s_n^{\mathcal{A}}(\bar{a})) \in R^{\mathcal{A}} \Leftrightarrow (s_1^{\mathcal{B}}(\bar{a}), \dots, s_n^{\mathcal{B}}(\bar{a})) \in R^{\mathcal{B}} \Leftrightarrow \mathcal{B} \models \varphi(\bar{a}).$$

- (c) If φ is $\neg\psi$,

$$\mathcal{A} \models \varphi(\bar{a}) \Leftrightarrow \mathcal{A} \not\models \psi(\bar{a}) \Leftrightarrow \mathcal{B} \not\models \psi(\bar{a}) \Leftrightarrow \mathcal{B} \models \varphi(\bar{a}),$$

where we use the induction hypothesis in the second \Leftrightarrow .

- (d) If φ is $\psi_1 \vee \psi_2$,

$$\mathcal{A} \models \varphi(\bar{a}) \Leftrightarrow \mathcal{A} \models \psi_1(\bar{a}) \text{ or } \mathcal{A} \models \psi_2(\bar{a}) \Leftrightarrow \mathcal{B} \models \psi_1(\bar{a}) \text{ or } \mathcal{B} \models \psi_2(\bar{a}) \Leftrightarrow \mathcal{B} \models \varphi(\bar{a}),$$

where we use the induction hypothesis in the second \Leftrightarrow .

■

^aRecall that we only need to show one of \vee or \wedge , and here we pick \vee and treat \wedge as an abbreviation.

As previously seen (Characteristic). Given a field K , the *characteristic* p of K is the number of 1 you need to add 1 in order to get 0, i.e., $\underbrace{1 + 1 + \dots + 1}_p = 0$.

Example. Let L be a subfield of K , for each $p > 0$, $\varphi_p := \underbrace{1 + 1 + \dots + 1}_p = 0$, which says the characteristic p . φ_p is **quantifier-free**, so

$$L \models \varphi_p \Leftrightarrow K \models \varphi_p.$$

Example. Consider $\mathbb{Z} = (\mathbb{Z}, 0, 1, +, -, \cdot)$, and let $\varphi(x) := \neg \exists y \ y + y = x$. We see that $\mathbb{Z} \models \varphi(1)$ but $\mathbb{Q} \models \neg \varphi(1)$.

Proposition 1.1.2. Suppose that \mathcal{A} is a **substructure** of \mathcal{B} , and $\varphi(\bar{x}, y_1, \dots, y_n)$ is a **quantifier-free formula**. Let $\bar{a} \in \mathcal{A}$, then

- (a) if $\mathcal{A} \models \exists y_1 \dots \exists y_n \ \varphi(\bar{a}, y_1, \dots, y_n)$, then $\mathcal{B} \models \exists y_1 \dots \exists y_n \ \varphi(\bar{a}, y_1, \dots, y_n)$;
- (b) if $\mathcal{B} \models \forall y_1 \dots \forall y_n \ \varphi(\bar{a}, y_1, \dots, y_n)$, then $\mathcal{A} \models \forall y_1 \dots \forall y_n \ \varphi(\bar{a}, y_1, \dots, y_n)$.

Proof. Suppose that $\mathcal{A} \models \exists y_1 \dots \exists y_n \ \varphi(\bar{a}, y_1, \dots, y_n)$, so there are $b_1, \dots, b_n \in \mathcal{A}$ such that $\mathcal{A} \models \varphi(\bar{a}, b_1, \dots, b_n)$. Since φ is **quantifier-free**, so $\mathcal{B} \models \varphi(\bar{a}, b_1, \dots, b_n)$ from **Proposition 1.1.1**, and hence $\mathcal{B} \models \exists y_1 \dots \exists y_n \ \varphi(\bar{a}, y_1, \dots, y_n)$.

On the other hand, it's easy to see that (b) is implied by (a). ■

Notation (Existential). In **Proposition 1.1.2**, formulas as in (a) are called *existential* (\exists_1 or \exists) formulas.

Notation (Universal). In **Proposition 1.1.2**, formulas as in (b) are called *universal* (\forall_1 or \forall) formulas.

Example. Recall $\mathcal{L}_1 = \{e, \cdot\}$, $\mathcal{L}_2 = \{e, \cdot, {}^{-1}\}$.

- Associativity: $\forall x \forall y \forall z \ (xy)z = x(yz)$.
- Identity: $\forall x \ ex = xe$.

These are \forall -formulas in either language.

- Inverses in \mathcal{L}_1 : $\forall x \exists y \ xy = yx = e$, which is **not** an \forall -formula.
- Inverses in \mathcal{L}_2 : $\forall x \ xx^{-1} = x^{-1}x = e$, which is an \forall -formula.

Hence, group axioms in \mathcal{L}_1 are not universal, but in \mathcal{L}_2 they are.

The above discrepancy is the reason why \mathcal{L}_2 is better than \mathcal{L}_1 , i.e., \mathcal{L}_1 -substructure might not be a group.

Problem. Show that $\forall x \exists y \ xy = yx = e$ in the above example is not **equivalent** to an \forall -formula.

Lecture 4: Theories and Axioms

Example. Let $\mathcal{L}_1 = \{E\}$, where E is a binary relation representing edge relation; and $\mathcal{L}_2 = \{V, E, I\}$, where V, E are unary relations and I is a binary relation representing incidence such that $I(v, e)$ for $v \in V$, $e \in E$ means that v is a vertex on edge e . Then,

- Let G be a graph, viewed as an \mathcal{L}_1 -structure. A **substructure** of G is an induced subgraph $H \subseteq G$ such that any edge in G between two vertices of H is in H .
- If we view G as an \mathcal{L}_2 -substructure, a **substructure** is a subgraph H such that H has some vertices and edges from G .^a

17 Jan. 14:30

^aBut there might be edges in H with no vertices, which can be fixed by having two functions $I_1(e) = v$, $I_2(e) = w$ when $e: v \rightarrow w$.

The difference is that for \mathcal{L}_1 , having an edge is **quantifier-free**, while in \mathcal{L}_2 is **existential**. To elaborate a bit further, for \mathcal{L}_2 , vEw is **quantifier-free**, while in \mathcal{L}_2 ,

$$\exists (v \in V \wedge w \in V \wedge e \in E \wedge I(v, e) \wedge I(w, e))$$

is not **quantifier-free**.

1.2 Theories

Let's start by the notion of **theory**.

Definition 1.2.1 (Theory). An \mathcal{L} -theory is a set of **\mathcal{L} -sentences**.

Definition 1.2.2 (Model). \mathcal{M} is a *model* of a **theory** T , written as $\mathcal{M} \models T$, if $\mathcal{M} \models \varphi$ for all $\varphi \in T$.

Note. Not every **theory** has a **model**, e.g., $\{\exists x \, x \neq x\}$.

The above note motivates the following.

Definition 1.2.3 (Satisfiable). A **theory** is *satisfiable* if it has a **model**.

Definition 1.2.4 (Elementary class). A class \mathcal{K} of **\mathcal{L} -structures** \mathcal{M} is called an *elementary class* if there is an **\mathcal{L} -theory** T such that

$$\mathcal{K} = \{\mathcal{M} \mid \mathcal{M} \models T\}.$$

One way to get an **elementary class** is to take an **\mathcal{L} -structure** \mathcal{M} and take the **full theory**.

Definition 1.2.5 (Full theory). The *full theory* $\text{Th}(\mathcal{M})$ of an **\mathcal{L} -structure** \mathcal{M} is defined as $\text{Th}(\mathcal{M}) = \{\varphi \mid \mathcal{M} \models \varphi\}$.

From the definition, $\mathcal{M} \models \text{Th}(\mathcal{M})$, and $\text{Th}(\mathcal{M})$ characterizes the **structures** satisfying the same **sentences** as \mathcal{M} .

Definition 1.2.6 (Complete). A **theory** T is *complete* if for any **sentence** φ , either $\varphi \in T$ or $\neg\varphi \in T$.

Remark. $\text{Th}(\mathcal{M})$ is **complete**.

Definition 1.2.7 (Elementarily equivalent). \mathcal{M} and \mathcal{N} are *elementarily equivalent* $\mathcal{M} \equiv \mathcal{N}$ if for all **sentences** φ ,

$$\mathcal{M} \models \varphi \Leftrightarrow \mathcal{N} \models \varphi.$$

Remark. There are $\mathcal{N} \models \text{Th}(\mathbb{N})$, but \mathcal{N} is not isomorphic to \mathbb{N} . \mathcal{N} is called a *non-standard model of arithmetic*, and \mathcal{N} might have *infinite element* larger than all of \mathbb{N} . Here, $\mathbb{N} = (\mathbb{N}, 0, 1, +, \cdot, -)$

Example. $\mathbb{Z} \oplus \mathbb{Z} \not\equiv \mathbb{Z}$ as groups.

The other way to define a **theory** is to write down axioms.

Example (Infinite set). Let $\mathcal{L} = \emptyset$, and let T consist of

$$\varphi_n := \exists x_1 \dots \exists x_n \bigwedge_{i \neq j} x_i \neq x_j.$$

Example (Linear order). Let $\mathcal{L} = \{\leq\}$, and let T consist of the axioms of linear orders, e.g.,

$$\forall x \forall y (x \leq y \wedge y \leq x \rightarrow x = y).$$

There are other interesting theories of linear orders, e.g., dense ones.

Example (Dense linear order). Consider

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

where we use $a < b$ as shorthand of saying $a \leq b \wedge a \neq b$.

Example (Group). In $\mathcal{L}_{\text{group}} = \{e, \cdot, {}^{-1}\}$, let T be the group axioms.

Other theories of groups include Abelson group, divisible, etc.

Definition 1.2.8 (Finitely axiomatizable). A theory is *finitely axiomatizable* if it has a finite set of axioms.

Given a theory, consider $T^{\models} = \{\varphi \mid T \models \varphi\}$,³ so $\mathcal{M} \models T$ if and only if $\mathcal{M} \models T^{\models}$. Often we think of T and T^{\models} as the same. A theory T is *finitely axiomatizable* if there is a finite Φ such that $T^{\models} = \Phi^{\models}$.

1.2.1 Elementary Embeddings

Let's now consider the following notion.

Definition 1.2.9 (Elementary embedding). Let \mathcal{M} and \mathcal{N} be \mathcal{L} -structures, and $f: \mathcal{M} \rightarrow \mathcal{N}$ an \mathcal{L} -embedding. Then f is an *elementary embedding* if for any formula $\varphi(\bar{x})$ and $\bar{a} \in M$,

$$\mathcal{M} \models \varphi(\bar{a}) \Leftrightarrow \mathcal{N} \models \varphi(f(\bar{a})).$$

Definition 1.2.10 (Elementary substructure). If $f: \mathcal{M} \hookrightarrow \mathcal{N}$ is a *elementary embedding* where \mathcal{M} is a *substructure* of \mathcal{N} , then \mathcal{M} is an *elementary substructure* of \mathcal{N} , written as $\mathcal{M} \preceq \mathcal{N}$.

Example. As groups, $\mathbb{Z} \hookrightarrow \mathbb{Q}$ is not *elementary*. In fact, $\mathbb{Z} \not\equiv \mathbb{Q}$. Whereas, if $f: \mathcal{M} \hookrightarrow \mathcal{N}$ is an *elementary embedding*, $\mathcal{M} \equiv \mathcal{N}$.^a

^aAnd also much more is true.

Proposition 1.2.1. Every *isomorphism* is an *elementary embedding*.

Proof. Let $f: \mathcal{M} \rightarrow \mathcal{N}$ be an *isomorphism*. We will argue by induction on formulas φ , that for all $\bar{a} \in M$,

$$\mathcal{M} \models \varphi(\bar{a}) \Leftrightarrow \mathcal{N} \models \varphi(f(\bar{a})).$$

Firstly, observe that all cases except quantifiers are the same as [Proposition 1.1.1](#). For quantifiers, suppose that $\varphi(\bar{x})$ is $\exists y \psi(\bar{x}, y)$ and $\mathcal{M} \models \varphi(\bar{a})$. This means that there is $b \in M$ such that $\mathcal{M} \models \psi(\bar{a}, b)$. By the induction hypothesis, $\mathcal{N} \models \psi(f(\bar{a}), f(b))$, so $\mathcal{N} \models \varphi(f(\bar{a}))$.

Now suppose $\mathcal{N} \models \varphi(f(\bar{a}))$, then there is $c \in N$ such that $\mathcal{N} \models \psi(f(\bar{a}), c)$. Since f is an *isomorphism*, so there is a $b \in M$ such that $f(b) = c$. By the induction hypothesis, $\mathcal{M} \models \psi(\bar{a}, b)$,

³Recall [Definition 1.1.13](#).

so $\mathcal{M} \models \varphi(\bar{a})$. ■

Corollary 1.2.1. If $\mathcal{M} \cong \mathcal{N}$, then $\mathcal{M} \equiv \mathcal{N}$.

1.2.2 Definable Sets

Consider the following.

Definition 1.2.11 (Definable). Let \mathcal{M} be an \mathcal{L} -structure, then $X \subseteq M^n$ is *definable* if there is a formula $\varphi(x_1, \dots, x_n, \bar{y})$ and $\bar{b} \in M$ such that

$$X = \{\bar{a} \in M^n \mid \mathcal{M} \models \varphi(\bar{a}, \bar{b})\}.$$

Notation (Define). We say that $\varphi(\bar{x}, \bar{b})$ *defines* X over \bar{b} , written as $X = \varphi(\mathcal{M}, \bar{b})$.

Notation (Parameter). The tuple \bar{b} is called the *parameters* when X is *definable* over \bar{b} .

Remark. Sometimes X is *definable* without *parameters*, or *definable* over \emptyset .

Example. Take $\mathbb{R} = (\mathbb{R}, 0, 1, +, \cdot, -)$ in $\mathcal{L}_{\text{ring}}$, then

$$\leq = \{(a, b) : a \leq b\}$$

is *definable*.

Example. Let $\mathbb{Z} = (\mathbb{Z}, +, -, \cdot, 0, 1)$, then \mathbb{N} is \emptyset -*definable* in \mathbb{Z} by^a

$$\mathbb{N} = \{z \in \mathbb{Z} : \exists u, v, x, y \ u^2 + v^2 + x^2 + y^2 = z\}.$$

^aFrom the **Langrange's four-square theorem**, which says that every natural number is the sum of four squares.

Example. \mathbb{Z} is \emptyset -*definable* in $\mathbb{Q} = (\mathbb{Q}, +, -, \cdot, 0, 1)$. This is a result of Julia Robinson [Rob49], and the formulation is very complicated.

Problem. How does one show that a set is not *definable*? For example, \mathbb{R} is not *definable* in $\mathbb{C} = (\mathbb{C}, 0, 1, +, \cdot, -)$.

Lecture 5: Hilbert-Style Deductive System

We start by asking whether \mathbb{R} is *definable* in $\mathbb{C} = (\mathbb{C}, 0, 1, +, \cdot, -)$?

19 Jan. 14:30

Proposition 1.2.2. Let \mathcal{M} be an \mathcal{L} -structure, and let $X \subseteq M^n$ be a set which is *definable* over \bar{a} . Then any *automorphism* of \mathcal{M} that fixes \bar{a} pointwise^a fixes X setwise.^b

^aIf $\bar{a} = (a_1, \dots, a_m)$, then $f(a_i) = a_i$.

^bIf $b \in X$, then $f(b) \in X$.

Proof. Let f be an *automorphism* of \mathcal{M} fixing \bar{a} pointwise, and $X = \{\bar{b} \in M^n : \mathcal{M} \models \varphi(\bar{b}, \bar{a})\}$. Fix \bar{b} , and suppose $\bar{b} \in X$, so $\mathcal{M} \models \varphi(\bar{b}, \bar{a})$. Because f is an *elementary embedding* from Proposition 1.2.1,

$$\mathcal{M} \models \varphi(f(\bar{b}), f(\bar{a})) \Rightarrow \mathcal{M} \models \varphi(f(\bar{b}), \bar{a}),$$

hence $f(\bar{b}) \in X$. Similarly, if $\bar{b} \notin X$, $\mathcal{M} \models \neg\varphi(\bar{b}, \bar{a}) \Rightarrow \mathcal{M} \models \neg\varphi(f(\bar{b}), \bar{a})$, so $f(\bar{b}) \notin X$. ■

Remark. If X is \emptyset -definable, it is fixed setwise by any automorphism.

Example. \mathbb{N} is fixed setwise by any automorphism of the ring \mathbb{Z} . In fact, the only automorphism of \mathbb{Z} is the identity.

Example. \mathbb{N} is not \emptyset -definable in $\mathbb{Z} = (\mathbb{Z}, 0, +)$.

Proof. Consider an automorphism $f(x) = -x$ of the group \mathbb{Z} , which does not fix \mathbb{N} setwise. *

Problem. Is \mathbb{N} definable in $\mathbb{Z} = (\mathbb{Z}, 0, +)$ over some parameters \bar{a} ?

Answer. For example, if $\bar{a} = (1)$, then f does not fix 1. In fact, any automorphism fixing 1 also fixes all of \mathbb{Z} , but \mathbb{N} is not definable in $(\mathbb{Z}, 0, +)$. To prove this we need compactness. *

As previously seen. Given a field F , then $F(a) \cong F(b)$ if a and b have the same minimal polynomial over F or if both do not satisfy any polynomial over F .

Example. $\mathbb{Q}(\pi) \cong \mathbb{Q}(e)$ because π and e are both transcendental.

We now return to the big question: is \mathbb{R} definable in $\mathbb{C} = (\mathbb{C}, 0, 1, +, \cdot, -)$? If $f: \mathbb{Q}(a) \rightarrow \mathbb{Q}(b)$ such that $a \mapsto b$, then there is an automorphism $\hat{f}: \mathbb{C} \rightarrow \mathbb{C}$ such that $a \mapsto b$, i.e., \hat{f} extends f . In other words, we need to find such an f with $a \in \mathbb{R}$ and $b \notin \mathbb{R}$.

Example. $a = \pi$, $b = i\pi$ are both transcendental.

Example. a is a real $\sqrt[4]{2}$, b is a complex $\sqrt[4]{2}$.

The above two examples show that \mathbb{R} is not \emptyset -definable in \mathbb{C} . In fact, \mathbb{R} is not definable over any \bar{a} because there are elements of \mathbb{R} and $\mathbb{C} \setminus \mathbb{R}$ transcendental over any \bar{a} .

Intuition. There are so many a, b such that given any \bar{a} , we can still find a pair that works.

1.3 Completeness and Compactness

In this section, we're going to formalize proofs.

1.3.1 Proofs

There are all sorts of different proof systems, and the one we use is the so-called Hilbert-style deductive system. Before that, we first see some common notions.

Notation (Schema). A *schema* is written in symbols for formulas, variables, etc.

Example. $\varphi \rightarrow (\psi \rightarrow \varphi)$ is a schema, i.e., an infinite set with all possible choices of φ and ψ .

Specifically, every logical axiom is written in schema, meaning that any instance of a symbol for a formula, e.g., φ , can be replaced by any formula.

Definition 1.3.1 (Generalization). A formula φ is a *generalization* of a formula ψ if φ is $\forall x_1 \dots \forall x_n \psi$ where x_1, \dots, x_n are variables.

Notation (Hypothesis). *Hypotheses* are formulas that we may assume in a proof.

Definition 1.3.2 (Proof). A *proof* is a sequence of formulas $\{\varphi_i\}_{i=1}^n$ such that φ_n is the conclusion, and each formula is either an *axiom* or is obtained from the previous formulas by a *rule of inference*.

Moreover, for a proof based on a set of hypotheses Γ , then in addition to a *logical axiom*, we can assert a formula $\varphi \in \Gamma$. If we prove ψ using Γ as *hypotheses*, we write $\Gamma \vdash \psi$.

Definition 1.3.3 (Valid). If we *prove* ψ without *hypotheses*, we write $\vdash \psi$ and say ψ is *valid*.

Definition 1.3.4 (Logical axioms). The *logical axioms* are the following formulas written in *schema*, as well as all of their *generalizations*:

Definition 1.3.5 (Propositional axioms). The *propositional axioms* are

- (A1) $\varphi \rightarrow (\psi \rightarrow \varphi)$.
- (A2) $(\varphi \rightarrow (\psi \rightarrow \theta)) \rightarrow ((\varphi \rightarrow \psi) \rightarrow (\varphi \rightarrow \theta))$.
- (A3) $(\neg\varphi \rightarrow \neg\psi) \rightarrow ((\neg\varphi \rightarrow \psi) \rightarrow \varphi)$.
- (A4) $\forall x \varphi(x, \dots) \rightarrow \varphi(t, \dots)$ where t is any term.
- (A5) $[\forall x (\varphi \rightarrow \psi)] \rightarrow [(\forall x \varphi) \rightarrow (\forall x \psi)]$.
- (A6) $\varphi \rightarrow \forall x \varphi$, where x is not free in φ .

Definition 1.3.6 (Axioms for equality). The *axioms for equality* is

- (A7) for any terms t, u, v, \dots , function symbols f , and relation symbols R ,
 - (a) $t = t$.
 - (b) $t = u \rightarrow u = t$.
 - (c) $(t = u \wedge u = v) \rightarrow (t = v)$.
 - (d) $(u_1 = t_1 \wedge \dots \wedge u_{n_f} = t_{n_f}) \rightarrow f(u_1, \dots, u_{n_f}) = f(t_1, \dots, t_{n_f})$.
 - (e) $(u_1 = t_1 \wedge \dots \wedge u_{n_R} = t_{n_R}) \rightarrow (R(u_1, \dots, u_{n_R}) \leftrightarrow R(t_1, \dots, t_{n_R}))$.

Definition 1.3.7 (Rule of inference). From φ and $\varphi \rightarrow \psi$, deduces ψ .^a

^aThis is called *modus ponens*.

These formulas might have free variables.

Example. A proof from calculus of a limit, e.g., $\forall \epsilon \exists \delta \dots$. And we start by stating

1. let $\epsilon > 0$,
2. choose $\delta = \epsilon$,
- ⋮

$$n. |f(x) - f(y)| < \epsilon.$$

We should interpret **free variables** as anything.

As previously seen (Propositional logic). $(p \wedge q) \vee (r \wedge \neg q)$.

Remark. We can check whether the **propositional axioms** are **true** with a truth table.

Definition 1.3.8 (Propositional tautology). A *propositional tautology* is a boolean combination \vee, \wedge, \neg of **formulas** $\varphi_1, \dots, \varphi_n$ which is **true** via a truth table assigning true or false to each of $\varphi_1, \dots, \varphi_n$.

So instead of using **propositional axioms**, we could instead allow as **logical axioms** any **propositional tautology**. To prove **completeness**, we will need 5 **propositional tautologies**. We will **prove** some of these, but take others on faith.

Remark. **Propositional axioms** are enough to **prove** all **propositional tautologies**.

Notation. We write $\Gamma \vdash_{\mathcal{L}} \varphi$ if there is a **proof** of φ from Γ in the **language** \mathcal{L} .

Note. Passing to a larger **language** will not let you **prove** more, so we can just write \vdash .

Lecture 6: Soundness Theorem

To see why **propositional axioms** are enough to **prove** all **propositional tautologies**, we see one example. 24 Jan. 14:30

Problem. **Prove** $\varphi \rightarrow \varphi$ using **propositional axioms**.

Answer. We see that

1. $\varphi \rightarrow ((\psi \rightarrow \varphi) \rightarrow \varphi)$ from (A1), where ψ is any **formula** (possibly $\psi = \varphi$).
2. $[\varphi \rightarrow ((\psi \rightarrow \varphi) \rightarrow \varphi)] \rightarrow [(\varphi \rightarrow (\psi \rightarrow \varphi)) \rightarrow (\varphi \rightarrow \varphi)]$ from (A2).
3. $(\varphi \rightarrow (\psi \rightarrow \varphi)) \rightarrow (\varphi \rightarrow \varphi)$ from (MP) and the two above.
4. $\varphi \rightarrow (\psi \rightarrow \varphi)$ from (A1).
5. $\varphi \rightarrow \varphi$ from (MP) and the two above.

⊛

In general, we can **prove**

- | | |
|---|---|
| (a) $\varphi \rightarrow \varphi$; | (d) $(\varphi \rightarrow \psi) \rightarrow ((\neg \varphi \rightarrow \psi) \rightarrow \psi)$; |
| (b) $\varphi \rightarrow \neg \neg \varphi$; | |
| (c) $\neg \neg \varphi \rightarrow \varphi$; | (e) $\varphi \rightarrow (\psi \rightarrow (\varphi \rightarrow \psi))$, |
| and so on. | |

Note. As we said, we may replace **propositional axioms** by every **propositional tautologies**.

Some **proof** system also have a second rule about universal quantifiers, but in our system, we have built this into the axioms. We can prove, as a theorem, what the other proof systems take as a rule.

Theorem 1.3.1. If $\Gamma \vdash \varphi$, and x does not occur **freely** in Γ , then $\Gamma \vdash \forall x \varphi$.

Proof. Fix Γ and x , we use *induction on proofs*. Consider the set $\{\varphi \mid \Gamma \vdash \forall x \varphi\}$, we will show that this set contains all the **logical axioms**, **formulas** from Γ , and is closed under **modus ponens**.^a

- (a) If φ is a **logical axiom**, so is its **generalization** $\forall x \varphi$, so $\Gamma \vdash \forall x \varphi$.
- (b) If $\varphi \in \Gamma$, then x is not **free** in φ , so from (A6), $\varphi \rightarrow \forall x \varphi$, and from (MP), $\forall x \varphi$. The above are based on Γ , hence $\Gamma \vdash \forall x \varphi$.
- (c) Suppose $\Gamma \vdash \forall x \varphi$ and $\Gamma \vdash \forall x (\varphi \rightarrow \psi)$, we want to show that $\Gamma \vdash \forall x \psi$.
 1. By (A5), $\forall x (\varphi \rightarrow \psi) \rightarrow (\forall x \varphi \rightarrow \forall x \psi)$, Γ **proves** this.
 2. By (MP), $\Gamma \vdash \forall x \varphi \rightarrow \forall x \psi$.
 3. By (MP) again, $\Gamma \vdash \forall x \psi$.

■

^aThus, if $\Gamma \vdash \theta$, then $\theta \in \{\varphi \mid \Gamma \vdash \forall x \varphi\}$.

Corollary 1.3.1. If $\vdash \varphi$, then $\vdash \forall x \varphi$. So the **generalization** of anything **valid** is also **valid**.

We now ask a critical question: is our **proof** system a good one?

1.3.2 Soundness Theorem

The first thing we should check is whether our **proofs** are **sound**.

Definition 1.3.9 (Sound). A **proof** system is *sound* if any **provable sentence** φ is **true**.

The idea is that if an **L-sentence** φ is **provable**, then it is **true** in all **L-structures**, i.e., every thing we **prove** should be **true**, in other words, we can't **prove** wrong things.

Lemma 1.3.1 (Soundness). If Γ is a set of **L-sentences** and φ is a **sentence**, and $\Gamma \vdash_{\mathcal{L}} \varphi$, then $\Gamma \models \varphi$.

Proof. Suppose that $\Gamma \vdash \varphi$, let $\psi_1, \psi_2, \dots, \psi_n = \varphi$ be such a **proof**.^a Let $\bar{x} = (x_1, \dots, x_m)$ be the **free variable** that appears in the ψ_i . Let \mathcal{M} be an **L-structure**, $\mathcal{M} \models \Gamma$. To show $\mathcal{M} \models \varphi$, we show that by induction on i , for all $\bar{a} \in M^m$, $\mathcal{M} \models \psi_i(\bar{a})$. For ψ_i , we have three cases.

- (a) If $\psi_i \in \Gamma$, then $\mathcal{M} \models \Gamma$ so $\mathcal{M} \models \psi_i$.
- (b) If ψ_i is a (**generalization** of) a **logical axiom**, then we can check that $\mathcal{M} \models \psi_i(\bar{a})$. For example, if ψ_i is (A1), $\theta \rightarrow (\gamma \rightarrow \theta)$, it's easy to check that

$$\mathcal{M} \models \theta(\bar{a}) \rightarrow (\gamma(\bar{a}) \rightarrow \theta(\bar{a})).$$

- (c) If there are $j, k < i$ such that ψ_k is $\psi_j \rightarrow \psi_i$, from inductive hypothesis, for all \bar{a} , $\mathcal{M} \models \psi_j(\bar{a}), \mathcal{M} \models \psi_k(\bar{a})$, then $\mathcal{M} \models \psi_j(\bar{a}) \rightarrow \psi_i(\bar{a})$. Checking our definition of **truth**, we get $\mathcal{M} \models \psi_i(\bar{a})$.

■

^aSome ψ_i might be **formulas**, but φ should be a **sentence**.

There are remarks to make about some obvious properties of $\vdash_{\mathcal{L}}$.

Remark. If $\varphi \in \Gamma$, then $\Gamma \vdash \varphi$.

Remark. If $\Delta \subseteq \Gamma$, and $\Delta \vdash \varphi$, then $\Gamma \vdash \varphi$.

Remark. If $\Gamma \vdash_{\mathcal{L}} \varphi$, and $\mathcal{L}^+ \supseteq \mathcal{L}$, then $\Gamma \vdash_{\mathcal{L}^+} \varphi$.

Remark. If $\Gamma \vdash \varphi$, then there is a finite $\Delta \subseteq \Gamma$ such that $\Delta \vdash \varphi$.

We can prove the following.

Theorem 1.3.2 (Deduction theorem). For any set of formulas Γ , formulas θ and ψ ,

$$\Gamma \cup \{\theta\} \vdash \psi \Leftrightarrow \Gamma \vdash \theta \rightarrow \psi.$$

Proof. The backward direction is easier. Suppose $\Gamma \vdash \theta \rightarrow \psi$, then $\Gamma \cup \{\theta\} \vdash \psi$ since we can have a proof like:

1. θ
- \vdots (the proof of $\Gamma \vdash \theta \rightarrow \psi$)
- n . $\theta \rightarrow \psi$
- $n+1$. ψ .

Now, suppose that $\Gamma \cup \{\theta\} \vdash \psi$, then there is a proof $\psi_1, \dots, \psi_n = \psi$ from $\Gamma \cup \{\theta\}$. We argue inductively that $\Gamma \vdash \theta \rightarrow \psi_i$. For i , we have three cases.

- (a) If $\psi_i \in \Gamma$ or it is a logical axiom. By (A1), $\psi_i \rightarrow (\theta \rightarrow \psi_i)$, so $\Gamma \vdash \theta \rightarrow \psi_i$.
- (b) If $\psi_i = \theta$. Then $\Gamma \vdash \theta \rightarrow \theta$ by (A1) and (A2) from here, hence $\Gamma \vdash \theta \rightarrow \psi_i$.
- (c) If ψ_i follows from $\psi_j, \psi_k = \psi_j \rightarrow \psi_i$, using (MP) with $j, k < i$.
 1. From the induction hypothesis, $\Gamma \vdash \theta \rightarrow \psi_j$ and $\Gamma \vdash \theta \rightarrow (\psi_j \rightarrow \psi_i)$.
 2. By (A2), $\Gamma \vdash [\theta \rightarrow (\psi_j \rightarrow \psi_i)] \rightarrow [(\theta \rightarrow \psi_j) \rightarrow (\theta \rightarrow \psi_i)]$.
 3. By (MP), $\Gamma \vdash (\theta \rightarrow \psi_j) \rightarrow (\theta \rightarrow \psi_i)$.
 4. By (MP), $\Gamma \vdash \theta \rightarrow \psi_i$.

■

Lecture 7: Soundness, Completeness, and Compactness

Proposition 1.3.1 (Contraposition). If $\Gamma \cup \{\varphi\} \vdash \neg\psi$, then $\Gamma \cup \{\psi\} \vdash \neg\varphi$.

Proof. Suppose $\Gamma \cup \{\varphi\} \vdash \neg\psi$, by the deduction theorem says that $\Gamma \vdash \varphi \rightarrow \neg\psi$. From (A1), (A2), and (A3), we can prove $(\varphi \rightarrow \neg\psi) \rightarrow (\psi \rightarrow \neg\varphi)$. By (MP), $\Gamma \vdash \psi \rightarrow \neg\varphi$, then from the deduction theorem, $\Gamma \cup \{\psi\} \vdash \neg\varphi$. ■

Now we introduce an important notion.

Definition 1.3.10 (Consistent). A theory T is *consistent* if for all φ , it is not the case that $T \vdash \varphi$ and $T \vdash \neg\varphi$.

Definition 1.3.11 (Inconsistent). If a theory T is not consistent, then it's *inconsistent*.

We could make the same definition for a set of formulas.

Proposition 1.3.2 (Proof by contradiction). If $\Gamma \cup \{\varphi\}$ is inconsistent, then $\Gamma \vdash \neg\varphi$.

26 Jan. 14:30

Proof. There is ψ such that $\Gamma \cup \{\varphi\} \vdash \psi$ and $\Gamma \cup \{\varphi\} \vdash \neg\psi$, so $\Gamma \vdash \varphi \rightarrow \psi$ and $\Gamma \vdash \varphi \rightarrow \neg\psi$ by the deduction theorem. Using (A1), (A2), and (A3), we can prove that

$$(\varphi \rightarrow \psi) \rightarrow ((\varphi \rightarrow \neg\psi) \rightarrow \neg\varphi).$$

By (MP), $\Gamma \vdash (\varphi \rightarrow \neg\psi) \rightarrow \neg\varphi$, and by (MP) again, we have $\Gamma \vdash \neg\varphi$. ■

Proposition 1.3.3. If a theory T is consistent, and φ is a sentence, then either $T \cup \{\varphi\}$ or $T \cup \{\neg\varphi\}$ is consistent.

Proof. If they were both inconsistent, $T \vdash \neg\varphi$ and $T \vdash \neg\neg\varphi$, so T would be inconsistent. ✗ ■

Note. The above is also true for formulas.

Remark. If T is inconsistent, then $T \vdash \varphi$ for any φ .

Proof. If T is inconsistent, then $T \cup \{\neg\varphi\}$ is inconsistent for all φ . Hence, from proof by contradiction, $T \vdash \neg\neg\varphi$ for all φ , which is just $T \vdash \varphi$. ⊛

Definition 1.3.12 (Maximal). A theory T is *maximal* if it is consistent and for all sentences φ , either $\varphi \in T$ or $\neg\varphi \in T$.

In particular, if $T \vdash \varphi$, then $\varphi \in T$.

Intuition. Basically, a maximal consistent theory has opinion on everything.

Now, we want to see that given a consistent theory, whether we can extend it to a maximal one. To do this, we need the following.

Definition. Let (P, \leq) be a partially ordered set.

Definition 1.3.13 (Chain). A *chain* is a set $C \subseteq P$ such that for every $p, q \in C$, either $p \leq q$ or $q \leq p$.

Definition 1.3.14 (Upper bound). If $X \subseteq P$ is a set, an *upper bound* for X is an element $p \in P$ such that $p \geq q$ for all $q \in X$.

Definition 1.3.15 (Maximal). An element $p \in P$ is *maximal* if there is no $q \in P$ with $q > p$.

Note. Note that a maximal element might not be greater than everything, there is just nothing greater than it.

Theorem 1.3.3 (Zorn's lemma). Let (P, \leq) be a partially ordered set. If every non-empty chain in P has an upper bound, then P has a maximal element.

Theorem 1.3.4. Any consistent theory T can be extended to a maximal consistent theory $T' \supseteq T$.

Proof. We first consider the case that T is countable by considering \mathcal{L} is countable since if \mathcal{L} is countable, then there are only countable many formulas since there are only countable many formulas of each length.

Claim. The result holds for \mathcal{L} being countable.

Proof. Firstly, list out all sentences $\varphi_1, \varphi_2, \dots$, start with $T_0 = T$. Given T_i consistent, one of $T_i \cup \{\varphi_i\}$ or $T_i \cup \{\neg\varphi_i\}$ is consistent from Proposition 1.3.3. Let T_{i+1} be one of these that is consistent. Let $T^* = \bigcup_i T_i$, which is maximal, and we now show that T^* is consistent.

Suppose not, then $T^* \vdash \theta$ and $T^* \vdash \neg\theta$ for some θ . In this case, there is some T_i such that $T_i \vdash \theta$ and $T_i \vdash \neg\theta$ because proofs are finite, with T_i being consistent, a contradiction $\nmid \quad \circledast$

Claim. The result holds for arbitrary \mathcal{L} .

Proof. For arbitrary \mathcal{L} , let (P, \leq) be the set of consistent theories extending T_i ordered by inclusion. Let C be a non-empty chain, and let $T^* = \bigcup_{T' \in C} T' \supseteq T$.

We see that T^* is consistent because if $T^* \vdash \theta$ and $T^* \vdash \neg\theta$, there are finitely many formulas used in those proofs, from, say, $T_1, \dots, T_n \in C$. Because C is a chain, by reordering, we may assume that $T_1 \subseteq \dots \subseteq T_n$. So $T_n \vdash \theta$ and $T_n \vdash \neg\theta$, contradicting the consistency of T_n , so T^* is consistent, i.e., $T^* \in P$. Furthermore, T^* is an upper bound on C ,^a so (P, \leq) has a maximal consistent theory $T^* \supseteq T$ from Zorn's lemma.

If T^* is not maximal, then there is φ where $\varphi \notin T^*$, $\neg\varphi \notin T^*$. From Proposition 1.3.3, one of $T^* \cup \{\varphi\}$ or $T^* \cup \{\neg\varphi\}$ is consistent, hence in P , contradicting to T^* being maximal $\nmid \quad \circledast$

^aNote that C is arbitrary.

Remark. We can do that same proof for any \mathcal{L} using transfinite recursion for the uncountable case. Motivated by Lemma 1.3.1 and Theorem 1.3.4, we close this section with the following.

Theorem 1.3.5 (Soundness). Let T be a theory and φ be a sentence.

- (a) If $T \vdash \varphi$, then $T \models \varphi$.
- (b) If T is satisfiable, then it is consistent.

Proof. (a) is exactly Theorem 1.3.5. For (b), let $\mathcal{M} \models T$, suppose that T was inconsistent, then $T \vdash \varphi$ and $T \vdash \neg\varphi$ for some φ . By (a), $T \models \varphi$ and $T \models \neg\varphi$, so $\mathcal{M} \models \varphi$ and $\mathcal{M} \models \neg\varphi$. But $\mathcal{M} \models \neg\varphi$ means $\mathcal{M} \not\models \varphi$, so this is a contradiction, hence T is consistent. \blacksquare

1.3.3 Completeness and Compactness Theorems

After knowing our proof system is sound, we now ask the converse: is our proof system complete?

Definition 1.3.16 (Complete). A proof system is complete if any true sentence φ is provable.

And indeed, this is the case.

Theorem 1.3.6 (Completeness). Let T be a theory and φ be a sentence.

- (a) If $T \models \varphi$, then $T \vdash \varphi$.
- (b) If T is consistent, then it is satisfiable.

(b) implies (a) is easy to see. Suppose that $T \models \varphi$, so $T \cup \{\neg\varphi\}$ is not satisfiable. By (b), $T \cup \{\neg\varphi\}$ is inconsistent. By proof by contradiction, $T \vdash \varphi$. One important consequence of the completeness theorem is the compactness theorem.

Theorem 1.3.7 (Compactness). Let T be a theory and φ be a sentence.

- (a) If $T \models \varphi$, then there is a finite $T_0 \subseteq T$ such that $T_0 \models \varphi$.

(b) T is **satisfiable** if and only if every finite subset of T is **satisfiable**.

Proof. Consider the following.

(a*) If $T \vdash \varphi$, then there is a finite $T_0 \subseteq T$ such that $T_0 \vdash \varphi$.

(b*) If T is **consistent** if and only if every finite subset of T is **consistent**.

We see that (a*) and (b*) are true because **proofs** are finite, and **soundness** and **completeness** translate directly between (a) and (a*) (and (b) and (b*)). ■

Remark. The **compactness theorem** does have something to do with topological compactness; consider the topological space of **complete satisfiable theories**, with the basic open sets being the sets

$$U_\varphi := \{T : T \models \varphi\},$$

then this topological space is compact.

Let's see one cool example using **compactness**.

Example. Let $\mathcal{L} = \{0, 1, +, \cdot, -, <\}$, and $\mathcal{L}^* = \mathcal{L} \cup \{c\}$, where c is a new constant symbol. Let

$$T = \text{Th}_{\mathcal{L}}(\mathbb{N}) \cup \{c > \underline{n} \mid n \in \mathbb{N}\},$$

then T is finitely **satisfiable**.

Proof. Given $T_0 \subseteq T$ finite, $T_0 \subseteq \text{Th}_{\mathcal{L}}(\mathbb{N}) \cup \{c > \underline{n}, \dots, c > \underline{n}_\ell\}$, and may assume they are equal and show that T_0 is **satisfiable**. Let \mathcal{N} be the $\mathcal{L} \cup \{c\}$ -structure which is the **expansion** of the \mathcal{L} -structure \mathbb{N} , with

$$c^{\mathcal{N}} = 1 + \max(n_1, \dots, n_\ell),$$

then $\mathcal{N} \models T_0$, and T_0 is **satisfiable**. By **compactness**, T is **satisfiable**, say $\mathcal{A} \models T$. Then $\mathcal{A} \equiv \mathbb{N}$ and \mathcal{A} contains an element $c^{\mathcal{A}}$ bigger than $1, 1+1, 1+1+1, \dots$, but $\mathcal{A} \not\equiv \mathbb{N}$, so \mathcal{A} is a non-standard model of arithmetic. ⊛

We now start a long journey toward proving **completeness theorem**, specifically (b).

Lecture 8: Henkin Constants

1.3.4 Henkin Construction

31 Jan. 14:30

To prove **Theorem 1.3.6 (b)**, we need an additional definition and a technical lemma due to Henkin.

Definition 1.3.17 (Henkin constant). An \mathcal{L}^* -theory T^* has *Henkin constants* if for each formula $\varphi(x)$ with one **free variable**, there is a constant symbol $c \in \mathcal{L}^*$ such that

$$(\exists x \varphi(x)) \rightarrow \varphi(c) \text{ is in } T^*.$$

We see that the above is equivalent to

$$(\neg \forall x \varphi(x)) \rightarrow \neg \varphi(c) \text{ is in } T^*,$$

and we will use this version (\forall) and view \exists being a shorthand for $\neg \forall \neg$; also, we will use \rightarrow and \neg as primitive, and \wedge, \vee are shorthand.

Lemma 1.3.2. If $\Gamma \vdash \varphi(c)$, and c does not occur in Γ or in $\varphi(x)$, then there is a variable y not appearing in $\varphi(x)$, such that $\Gamma \vdash \forall y \varphi(y)$. Moreover, there is a **proof** of $\forall y \varphi(y)$ in which c does not appear.

Proof. Let $\alpha_1(c), \dots, \alpha_n(c) = \varphi(c)$ be a **proof** of $\varphi(c)$ from Γ , and let y be a variable not appearing in this **proof**. We claim that $\alpha_1(y), \dots, \alpha_n(y) = \varphi(y)$ is still a valid **proof** of $\varphi(y)$. There are three

cases to consider (for each $i = 1, \dots, n$):

- (a) If $\alpha_i(c)$ is in Γ , then c does not actually occur in $\alpha_i(c)$ because it does not appear in Γ . So $\alpha_i(y)$ is the same as $\alpha_i(c)$, hence in Γ .
- (b) If $\alpha_i(c)$ is a **logical axiom**, then $\alpha_i(y)$ is a **logical axiom** as well. For most of these it is easy to check, but for (A6), i.e., $\varphi \rightarrow \forall x \varphi$ if x is not **free** in φ , there is a little more. But y did not appear in $\alpha_i(c)$, so $y \neq x$, and substituting y for c will not stop x from being not **free**.
- (c) If $\alpha_i(c)$ follows by (MP) from $\alpha_j(c)$ and $\alpha_k(c) = \alpha_j(c) \rightarrow \alpha_i(c)$ for $j, k < i$, then $\alpha_i(y)$ follows by (MP) from $\alpha_j(y)$ and $\alpha_k(y) = \alpha_j(y) \rightarrow \alpha_i(y)$.

So $\Gamma \vdash \varphi(y)$ and the **proof** does not involve c . Let $\Phi \subseteq \Gamma$ be the subset of Γ that was used in the **proof**, so y does not appear in Φ , hence $\Phi \vdash \varphi(y)$ and $\Phi \vdash \forall y \varphi(y)$, so $\Gamma \vdash \forall y \varphi(y)$. ■

So Lemma 1.3.2 implies that we have $\Gamma \vdash \varphi(y)$ and the **proof** does not involve c . And sometimes, we want to be able to choose the variable y from above.

Corollary 1.3.2. If $\Gamma \vdash \varphi(c)$, and c does not occur in Γ or in $\varphi(x)$, then $\Gamma \vdash \forall x \varphi(x)$. Moreover, there is a **proof** of $\forall x \varphi(x)$ not involving c .^a

^aHere, x is any variable that does not appear in $\varphi(c)$.

Proof. We know that for some y , $\Gamma \vdash \forall y \varphi(y)$, (A4) says $\forall y \varphi(y) \rightarrow \varphi(x)$. So $\forall y \varphi(y) \vdash \varphi(x)$ since x does not appear in $\forall y \varphi(y)$, $\forall y \varphi(y) \vdash \forall x \varphi(x)$. ■

Note. x might appear in Γ .

Theorem 1.3.8. Let T be a **consistent \mathcal{L} -theory**. There is a **language** $\mathcal{L}^* \supseteq \mathcal{L}$ and $T^* \supseteq T$ a **consistent \mathcal{L}^* -theory** such that T^* has **Henkin constants**. We can choose \mathcal{L}^* such that $|\mathcal{L}^*| = |\mathcal{L}| + \aleph_0$, and all new symbols in \mathcal{L}^* are constants.

Proof. Let $\mathcal{L}_0 = \mathcal{L}$ and $T_0 = T$. Let \mathcal{L}_1 be the **expansion** of \mathcal{L}_0 by adding a new constant symbol c_φ for each **\mathcal{L}_0 -formula** φ w.r.t. the **Henkin** construction. First, we show that after this procedure, T_0 is still a **consistent \mathcal{L}_1 -theory**.

Remark. Technically, \vdash is really $\vdash_{\mathcal{L}}$, so this is a key step for seeing that it does not matter.

Claim. T_0 is still a **consistent \mathcal{L}_1 -theory** after the **expansion** of \mathcal{L}_0 .

Proof. If not, there is a **proof** of a **contradiction** from T_0 , and which uses only finitely many of the new constants symbols. By Corollary 1.3.2, we can replace these constants one-by-one by variables, e.g., if the original **contradiction** was $\varphi(c_1, \dots, c_n)$ and $\neg \varphi(c_1, \dots, c_n)$, then T_0 proves $\forall x_1, \dots, \forall x_n \varphi(x_1, \dots, x_n)$ and $\forall x_1, \dots, \forall x_n \neg \varphi(x_1, \dots, x_n)$. Moreover, these **proofs** take place in \mathcal{L}_0 , so by (A4), $T_0 \vdash_{\mathcal{L}_0} \varphi(x_1, \dots, x_n)$, and $T_0 \vdash_{\mathcal{L}_0} \neg \varphi(x_1, \dots, x_n) \not\vdash$ ⊗

To construct T_1 w.r.t. the **Henkin** construction, it's natural to consider the following: if φ is of the form $\neg \forall x \psi(x)$, then let

$$\theta_\varphi := (\neg \forall x \psi(x)) \rightarrow \neg \psi(c_\varphi), \text{ i.e., } (\exists x \neg \psi(x)) \rightarrow \neg \psi(c_\varphi),$$

otherwise, let $\theta_\varphi := \forall x (x = x)$ (trivially **true**). Let $\Theta = \{\theta_\varphi \mid \varphi \text{ an } \mathcal{L}_0\text{-formula}\}$, and we let that $T_1 = T_0 \cup \Theta$. We claim that T_1 is still **consistent**.

Claim. $T_1 = T_0 \cup \Theta$ is a **consistent \mathcal{L}_1 -language** after the **expansion** of \mathcal{L}_0 .

Proof. If not, then there are $\varphi_1, \dots, \varphi_{m+1}$ such that $T_0 \cup \{\theta_{\varphi_1}, \dots, \theta_{\varphi_m}, \theta_{\varphi_{m+1}}\}$ is **inconsistent**. Taking m to be as small as possible, $T_0 \cup \{\theta_{\varphi_i}\}_{i=1}^m$ is **consistent**, so $T_0 \cup \{\theta_{\varphi_i}\}_{i=1}^m \vdash \neg\theta_{\varphi_{m+1}}$ with φ_{m+1} being of the form $\neg\forall x \psi(x)$, $\theta_{\varphi_{m+1}}$ is $\neg\forall x \psi(x) \rightarrow \neg\psi(c_\varphi)$. By (A1), (A2), (A3),

$$T_0 \cup \{\theta_{\varphi_1}, \dots, \theta_{\varphi_m}\} \vdash \neg\forall x \psi(x) \text{ and } T_0 \cup \{\theta_{\varphi_1}, \dots, \theta_{\varphi_m}\} \vdash \psi(c_{\varphi_{m+1}}).$$

Since $c_{\varphi_{m+1}}$ does not appear in $T_0 \cup \{\theta_{\varphi_i}\}_{i=1}^m$, so $T_0 \cup \{\theta_{\varphi_i}\}_{i=1}^m \vdash \forall x \psi(x)$, i.e., $T_0 \cup \{\theta_{\varphi_i}\}_{i=1}^m$ is **inconsistent**, contradicting to the fact that m is the smallest choice \nless^a \otimes

^aIf $m = 0$, then we violate the **consistency** of T_0 .

It might be that T_1 does not have **Henkin constants** since there are new \mathcal{L}_1 -formulas which are not \mathcal{L}_0 -formulas. But we know that T_1 does have **Henkin constants** for \mathcal{L}_0 -formulas, hence we can repeat that process and keep fixing things. In general, given T_i and \mathcal{L}_i , define a T_{i+1} and \mathcal{L}_{i+1} in the above way. Since each T_i is **consistent**, so $T^* = \bigcup T_i$ is an $\mathcal{L}^* = \bigcup \mathcal{L}_i$ -theory. Note that T^* is **consistent** as a nested union of **consistent theories**, and T^* has **Henkin constants** because every \mathcal{L}^* -formula φ is an \mathcal{L}_i -formula for some i , and $\theta_\varphi \in T_{i+1} \subseteq T^*$.

Intuition. This is like “chasing its own tail,” which basically fixes new errors introduced every time and then takes the union in the end.

Finally, we want to show that $|\mathcal{L}^*| = |\mathcal{L}| + \aleph_0$. Given \mathcal{L}_i , we define \mathcal{L}_{i+1} to be \mathcal{L}_i plus new constants c_φ for φ on \mathcal{L}_i -formula. Then, we have

$$|\mathcal{L}_{i+1}| \leq |\mathcal{L}_i| + \underbrace{|\mathcal{L}_i|}_{\# \text{ of } \mathcal{L}_i\text{-formulas}} + \aleph_0 = |\mathcal{L}_i| + \aleph_0.$$

So for all i , $|\mathcal{L}_i| \leq |\mathcal{L}| + \aleph_0$, and $\mathcal{L}^* = \bigcup_i \mathcal{L}_i$ is a countable union, so $|\mathcal{L}^*| \leq |\mathcal{L}| + \aleph_0$, and in fact, $|\mathcal{L}^*| = |\mathcal{L}| + \aleph_0$. \blacksquare

After proving Theorem 1.3.8, we see that to prove Theorem 1.3.6 (b), we can proceed by:

1. extend T^* to a **maximal theory** T^{**} ;⁴
2. turn T^{**} into a **model**. The elements of the **model** are constant symbols from \mathcal{L}^* , modulo the equivalence relation $c \sim d$ if $c = d$ is in T^{**} , i.e., $T^{**} \vdash c = d$.

Thankfully, the first step is easy from Theorem 1.3.4, so we just need to show the second step, and we’re done.

Lecture 9: Proving the Completeness Theorem

To finish the proof of Theorem 1.3.6 (b), we follow the plan mentioned last lecture, and prove the following. 2 Feb. 14:30

Theorem 1.3.9. If T is a **maximal consistent** \mathcal{L} -theory with **Henkin constants**, then T has a **model**.

Proof. The **model** we build is called a “canonical model.” Let \mathcal{C} be the set of constants in \mathcal{L} , and define an equivalence relation \sim on \mathcal{C} by $c \sim d$ if and only if $c = d$ is in T .

Claim. The relation \sim on \mathcal{C} defined by $c \sim d \Leftrightarrow c = d \in T$ is an equivalence relation.

⁴Which still has **Henkin constants**.

Proof. We check the axioms for being an equivalence relation.

- (a) $c \sim c$ because $c = c$ is in T by (A7) (a).^a
- (b) If $c \sim d$, then $c = d$ is in T so $d = c$ is in T by (A7) (b), i.e., $d \sim c$.
- (c) If $c \sim d$ and $d \sim e$, then $c = d$ and $d = e \in T \Rightarrow c = e \in T$ by (A7) (c), so $c \sim e$.

⊗

^aOtherwise, $c \neq c$ is in T from the maximality, so $T \vdash c \neq c$ with $T \vdash c = c$, so T would be inconsistent.

Let $[c]$ be the equivalence class of c . Define an \mathcal{L} -structure \mathcal{M} with domain $M = \mathcal{C} / \sim = \{[c] \mid c \in \mathcal{C}\}$, with functions, relations, and constants defined as follows:

- (a) $c^{\mathcal{M}} = [c]$.
- (b) $R^{\mathcal{M}}([c_1], \dots, [c_n])$ **true** if $R(c_1, \dots, c_n)$ is in T . This is well-defined by (A7) (e).
- (c) $f^{\mathcal{M}}([c_1], \dots, [c_n]) = [d]$ if $f(c_1, \dots, c_n) = d$ is in T . Such a d exists because $\exists x f(c_1, \dots, c_n) = x$, i.e., $\neg \forall x f(c_1, \dots, c_n) \neq x$, is in T .^b If this is in T , then there is a Henkin constant d with $f(c_1, \dots, c_n) = d$ in T . To show that this is well-defined, from (A7) (d), i.e.,

$$(t_1 = u_1 \wedge \dots \wedge t_n = u_n) \rightarrow f(t_1, \dots, t_n) = f(u_1, \dots, u_n).$$

So if $[c_1] = [d_1], \dots, [c_n] = [d_n]$, then $c_1 = d_1, \dots, c_n = d_n$ are in T . So $f(c_1, \dots, c_n) = f(d_1, \dots, d_n)$ is in T by (A7) (d). If a and b are constants such that $f(c_1, \dots, c_n) = a$ and $f(d_1, \dots, d_n) = b$ are in T , so $a = b$ is in T by (A7) (c), i.e., the transitivity of $=$.

Now we need to show that $\mathcal{M} \models T$, i.e., we claim that

$$\mathcal{M} \models \varphi([c_1], \dots, [c_n]) \Leftrightarrow \varphi(c_1, \dots, c_n) \text{ is in } T.$$

We prove this by induction on terms and then formulas.

1. **Terms:** Show that $t^{\mathcal{M}}([c_1], \dots, [c_n]) = [d]$ if and only if $t(c_1, \dots, c_n) = d$ is in T .

- (a) If t is a constant e , $t^{\mathcal{M}}([c_1], \dots, [c_n]) = e^{\mathcal{M}} = [e]$, and

$$[e] = t^{\mathcal{M}}([c_1], \dots, [c_n]) = [d] \Leftrightarrow [e] = [d] \Leftrightarrow e = d \text{ is in } T.$$

- (b) If t is x_i , $t^{\mathcal{M}}([c_1], \dots, [c_n]) = [c_i]$. This is equal to $[d]$ if and only if $c_i = d$ is in T .
- (c) Suppose that $t(x_1, \dots, x_n) = f(s_1(x_1, \dots, x_n), \dots, s_m(x_1, \dots, x_n))$. Let

$$[d_i] = s_i^{\mathcal{M}}([c_1], \dots, [c_n]),$$

by the inductive hypothesis, $d_i = s_i(c_1, \dots, c_n)$ is in T . Let $[e] = f^{\mathcal{M}}([d_1], \dots, [d_m]) = t^{\mathcal{M}}([c_1], \dots, [c_n])$. By the definition of f , $e = f(d_1, \dots, d_m)$ is in T . By (A7) (d),

$$e = f(s_1(c_1, \dots, c_n), \dots, s_m(c_1, \dots, c_n))$$

is in T . This is the direction (\Rightarrow).

Now suppose that $t(c_1, \dots, c_n) = e'$ is in T . We want to show that $[e] = [e']$, i.e., $e = e'$ is in T . Since $e = t(c_1, \dots, c_n)$ is in T , and $e' = t(c_1, \dots, c_n)$ is in T . By (A7) (c), $e = e'$ is in T , so $[e'] = [e] = t^{\mathcal{M}}([c_1], \dots, [c_n])$. This is (\Leftarrow).

2. **Formulas:** Show that $\mathcal{M} \models \varphi([c_1], \dots, [c_n])$ if and only if $\varphi(c_1, \dots, c_n)$ is in T .^c

- (a) If φ is $s(x_1, \dots, x_n) = t(x_1, \dots, x_n)$:

(\Rightarrow) If $\mathcal{M} \models s([c_1], \dots, [c_n]) = t([c_1], \dots, [c_n])$,

$$s^{\mathcal{M}}([c_1], \dots, [c_n]) = t^{\mathcal{M}}([c_1], \dots, [c_n]).$$

Let $[d]$ be this element equal to the above, so $d = s(c_1, \dots, c_n)$ and $d = t(c_1, \dots, c_n)$ are in T so $\underbrace{s(c_1, \dots, c_n) = t(c_1, \dots, c_n)}_{\varphi(c_1, \dots, c_n)}$ is in T by (A7) (c).

(\Leftarrow) If $s(c_1, \dots, c_n) = t(c_1, \dots, c_n)$ is in T , let

$$[d] = s^{\mathcal{M}}([c_1], \dots, [c_n]) \text{ and } [e] = t^{\mathcal{M}}([c_1], \dots, [c_n]),$$

so $d = s(c_1, \dots, c_n)$ and $e = t(c_1, \dots, c_n)$ are in t , so $d = e$ is in t , and $[e] = [d]$.

(b) If φ is $R(t_1(\bar{x}), \dots, t_m(\bar{x}))$: Let $[d_i] = t_i^{\mathcal{M}}([c_1], \dots, [c_n])$,

$$\begin{array}{ccc} R^{\mathcal{M}}([d_1], \dots, [d_m]) \text{ is true} & \Longleftrightarrow & R(d_1, \dots, d_m) \text{ is in } T \\ \Updownarrow & & \Updownarrow \\ R^{\mathcal{M}}(t_1^{\mathcal{M}}[\bar{c}], \dots, t_m^{\mathcal{M}}[\bar{c}]) \text{ is true} & & R(t_1(\bar{c}), \dots, t_m(\bar{c})) \text{ is in } T \\ \Updownarrow & & \\ \mathcal{M} \models \varphi([c_1], \dots, [c_n]) & & \end{array}$$

(c) If φ is $\neg\psi$: Then

$$\mathcal{M} \models \varphi(\bar{c}) \Leftrightarrow \mathcal{M} \not\models \psi(\bar{c}) \Leftrightarrow \psi(\bar{c}) \text{ is not in } T \Leftrightarrow \varphi(\bar{c}) \text{ is in } T$$

where the last \Leftrightarrow follows from the fact that T is **maximal** and **consistent**.

(d) If φ is $\psi \rightarrow \theta$:

- If $\psi(\bar{c}) \rightarrow \theta(\bar{c})$ is in T : then if $\psi(\bar{c})$ is in T , then $\theta(\bar{c})$ is in T by (MP). then by the induction hypotheses, if $\mathcal{M} \models \psi(\bar{c})$, then $\mathcal{M} \models \theta(\bar{c})$.
- If $\mathcal{M} \models \psi(\bar{c}) \rightarrow \theta(\bar{c})$: then either $\mathcal{M} \models \theta(\bar{c})$ or $\mathcal{M} \models \neg\psi(\bar{c})$. So either
 - i. $\theta(\bar{c})$ is in T : by (A1), $\theta(\bar{c}) \rightarrow (\psi(\bar{c}) \rightarrow \theta(\bar{c}))$, so $\psi(\bar{c}) \rightarrow \theta(\bar{c})$ is in T .
 - ii. $\neg\psi(\bar{c})$ is in T : $T \cup \{\psi(\bar{c})\}$ is now **inconsistent**, so $T \cup \{\psi(\bar{c})\} \vdash \theta(\bar{c})$. From the **deductive theorem**, $T \vdash \psi(\bar{c}) \rightarrow \theta(\bar{c})$. Because T is **maximal** and **consistent**, $\psi(\bar{c}) \rightarrow \theta(\bar{c})$ is in T .

^bOtherwise, $\forall x f(c_1, \dots, c_n) \neq x$ is in T . By (A4), $f(c_1, \dots, c_n) \neq f(c_1, \dots, c_n)$ is in T , contradicts to (A7) (a).

^cIn particular, for a **sentence** φ , $\mathcal{M} \models \varphi \Leftrightarrow \varphi$ is in T , and so $\mathcal{M} \models T$.

Lecture 10: Introduction to Model Theory

Let's start by finishing the proof of Theorem 1.3.9.

7 Feb. 14:30

Proof of Theorem 1.3.9 (Continued). There's one final case left:

- (e) If φ is $\forall x \psi(x, \bar{y})$: Because T has **Henkin constants**, there is d such that $\neg\forall x \psi(x, \bar{c}) \rightarrow \neg\psi(d, \bar{c})$ is in T .
- If $\varphi(c_1, \dots, c_n)$ is not in T , i.e., $\forall x \psi(x, \bar{c})$ is in T , then since T is **maximal**, $\neg\forall x \psi(x, \bar{c})$ is in T . So by (MP), $\neg\psi(d, \bar{c})$ is in T . So, $\mathcal{M} \models \neg\psi([d], [\bar{c}])$ by induction hypotheses, hence $\mathcal{M} \models \neg\forall x \psi(x, [\bar{c}])$, i.e., $\mathcal{M} \not\models \varphi([\bar{c}])$.
 - If $\mathcal{M} \models \varphi([\bar{c}])$, then $\mathcal{M} \models \forall x \psi(x, [\bar{c}])$, so there is $[e]$ such that $\mathcal{M} \models \neg\psi([e], [\bar{c}])$. Hence, $\neg\psi(e, \bar{c})$ is in T . Suppose for a contradiction that $\varphi(\bar{c})$, i.e., $\forall x \psi(x, \bar{c})$ is in T , by (A4), $\forall x \psi(x, \bar{c}) \rightarrow \psi(e, \bar{c})$, so $\psi(e, \bar{c})$ is in T by **maximality** and by **consistency**. But then T is **inconsistent**, a contradiction \nmid Hence $\varphi(\bar{c})$ is not in T .

Thus, $\mathcal{M} \models T$, so T is **satisfiable**, proving the theorem. ■

Remark. We see that when proving the above, when we talk about \mathcal{M} , the witness comes for free, while for T , we need **Henkin constants** for getting a witness.

Now, we can complete the proof of **completeness theorem** by putting everything together.

Claim. The **completeness theorem (b)** holds.

Proof. We see that

1. **Theorem 1.3.8:** There is a **consistent** $T^* \supseteq T$ and \mathcal{L}^* -theory (with $\mathcal{L}^* \supseteq \mathcal{L}$) and T^* has **Henkin constants**.
2. **Theorem 1.3.4:** There is a **maximal consistent** \mathcal{L}^* -theory $T^{**} \supseteq T^*$, where T^{**} still has **Henkin constants**.
3. **Theorem 1.3.9:** T^{**} has a **model** \mathcal{M}^* an \mathcal{L}^* -structure. Let \mathcal{M} be the **reduct** of \mathcal{M}^* to an \mathcal{L} -structure.

Hence, $\mathcal{M} \models T$. *

As previously seen (Problem set 1). Let $\mathcal{L}^* \supseteq \mathcal{L}$. If \mathcal{M}^* is an \mathcal{L}^* -structure, then by ignoring the **interpretation** of the symbols in $\mathcal{L}^* - \mathcal{L}$, we get an \mathcal{L} -structure \mathcal{M} .

Notation (Reduct). \mathcal{M} is a *reduct* of \mathcal{M}^* .

Notation (Expansion). \mathcal{M}^* is an *expansion* of \mathcal{M} .

Remark. We see that \vdash and \models are the same.

1.3.5 Consequences of Completeness Theorem

Now, let's step back and look at the proof of the **completeness theorem**, and ask the following.

Problem. When we did the **Henkin** construction of $\mathcal{M}^* \models T^{**}$, how big was M ?

This can be answered by the following.

Theorem 1.3.10. If T is a **satisfiable** \mathcal{L} -theory, then it has a **model** of size at most $|\mathcal{L}| + \aleph_0$.

Proof. Since $|M| \leq |\mathcal{L}^*|$ since $\mathcal{M} = \mathcal{C} / \sim$, and in step one, $|\mathcal{L}^*| \leq |\mathcal{L}| + \aleph_0$, so $|M| \leq |\mathcal{L}| + \aleph_0$. ■

Example. Let $\mathcal{L} = \{f\}$, T says that f is injective but not surjective.

Example. Let $\mathcal{L} = \{\leq\}$, T says that \leq is a linear order with no greatest element.

Example. Let $\mathcal{L} = \emptyset$, T says that there are at least n elements for each n .

As previously seen. \vdash and \models are actually $\vdash_{\mathcal{L}}^a$ and $\models_{\mathcal{L}}^b$

^aProofs can only use \mathcal{L} -formulas.

^bOnly looking at \mathcal{L} .

Remark. Suppose $\mathcal{L} \supseteq \mathcal{L}_0$, and Γ a set of \mathcal{L}_0 -sentences, φ on \mathcal{L}_0 -sentence.

- (a) $\Gamma \models_{\mathcal{L}_0} \varphi \Leftrightarrow \Gamma \models_{\mathcal{L}_1} \varphi$.
- (b) $\Gamma \vdash_{\mathcal{L}_0} \varphi \Leftrightarrow \Gamma \vdash_{\mathcal{L}_1} \varphi$.

Proof. (a) and (b) are equivalent by the [completeness theorem](#), and we prove (a).

Suppose $\Gamma \models_{\mathcal{L}_0} \varphi$. Suppose \mathcal{M}_1 is an \mathcal{L}_1 -structure such that $\mathcal{M}_1 \models \Gamma$. Let \mathcal{M}_0 be the [reduct](#) of \mathcal{M}_1 to \mathcal{L}_0 and $\mathcal{M}_0 \models \Gamma$, so $\mathcal{M}_0 \models \varphi$, then $\mathcal{M}_1 \models \varphi$, thus $\Gamma \models_{\mathcal{L}_1} \varphi$.

Now, suppose $\Gamma \models_{\mathcal{L}_1} \varphi$. Suppose \mathcal{M}_0 is an \mathcal{L}_0 -structure with $\mathcal{M}_0 \models \Gamma$. Expand \mathcal{M}_0 to an \mathcal{L}_1 -structure \mathcal{M}_1 in any way. $\mathcal{M}_1 \models \Gamma$, so $\mathcal{M}_1 \models \varphi$. Thus, $\mathcal{M}_0 \models \varphi$, so $\Gamma \models_{\mathcal{L}_0} \varphi$. ⊛

Definition 1.3.18 (Computably enumerable). A set is *computably enumerable* or *computable listable* if there is a program that lists out its elements.

What is important about the [proof system](#)?

- (1) [Soundness](#) and [completeness](#), $\vdash \Leftrightarrow \models$.
- (2) [Proofs](#) are finite, and use only finitely many hypotheses \Rightarrow [compactness](#).
- (3) Computational properties. If \mathcal{L} is finite, or computable (complete list of symbols and their arities).
 - (a) We can compute with [formulas](#).
 - (b) Given a [formula](#), it's computable to check whether it's a [logical axiom](#).
 - (c) It's computable to check whether a [proof](#) is valid.
 - (d) If Γ is a [computably enumerable](#) set of [sentences](#), $\{\varphi : \Gamma \vdash \varphi\}$ is also [computably enumerable](#).⁵
 - (e) There is no program that given φ can decide whether $\vdash \varphi$ at least for $\mathcal{L} = \{E\}$, E binary.

⁵We can list out all the valid [proofs](#) from Γ of any φ .

Chapter 2

The Beginning of Model Theory

2.1 Complete Theories

Let's start with a proposition.

Proposition 2.1.1. Let T be an \mathcal{L} -theory with an infinite model, and let κ be an infinite cardinal with $\kappa \geq |\mathcal{L}|$. Then T has a model of cardinality κ .

Proof. Let \mathcal{C} be a set of κ -many new constants, and let $\mathcal{L}^* = \mathcal{L} \cup \mathcal{C}$. Let

$$T^* = T \cup \{c \neq d \mid c, d \in \mathcal{C} \text{ distinct}\}.$$

If $\mathcal{M} \models T^*$, then $|M| \geq \kappa$; also, if T^* is satisfiable, it has a model of size at most $|\mathcal{L}^*| = \kappa$ since

$$\kappa = |\mathcal{C}| \leq |\mathcal{L}^*| \leq |\mathcal{C}| + |\mathcal{L}| \leq \kappa + \kappa = \kappa$$

from Theorem 1.3.10. Hence, if T^* is satisfiable, it has a model \mathcal{M} with $|M| = \kappa$.

Claim. T^* is satisfiable.

Proof. It's enough to show that every finite $\Gamma \subseteq T^*$ is satisfiable from the compactness theorem. Let \mathcal{M} be infinite, and $\Gamma \subseteq T^*$ finite, then we can write

$$\Gamma \subseteq T \cup \{c_i \neq c_j \mid i, j = 1, \dots, n, i \neq j\}$$

for $c_1, \dots, c_n \in \mathcal{C}$ since only finitely many c_i are involved. Without loss of generality, $\Gamma = T \cup \{c_i \neq c_j \mid i, j = 1, \dots, n, i \neq j\}$. Pick $a_1, \dots, a_n \in M$, distinct, we then turn \mathcal{M} into an \mathcal{L}^* -structure \mathcal{M}^* with $c_i^{\mathcal{M}^*} = a_i$,^a resulting in $\mathcal{M}^* \models \Gamma$. ⊗

^aAnd each other $d \in \mathcal{C}$ with $d^{\mathcal{M}^*} = a_1$.

■

Lecture 11: Algebraically Closed Fields

2.1.1 A Detour to Algebraically Closed Fields

9 Feb. 14:30

Algebraically closed fields are a great example of a *tame* theory (as opposed to e.g., \mathbb{N} , which are not *tame*). We detour to discuss some important and related definitions for the future discussion.

Rings

All rings R we refer to will be commutative.

Definition 2.1.1 (Ideal). Let R be a ring. An *ideal* I of R is a set $I \subseteq R$ such that

- (a) $0 \in I$;
- (b) if $a, b \in I$, then $a + b \in I$;
- (c) if $a \in I$ and $r \in R$, $ra \in I$.

Intuition. An *ideal* is trying to act as a set of “zeros” (in order to be further mod out).

Definition 2.1.2 (Proper). An *ideal* is *proper* if $1 \notin I$, equivalently, $I \neq R$.

Definition. Let I be a *proper ideal*.

Definition 2.1.3 (Radical). I is *radical* if $a^n \in I$, then $a \in I$.

Definition 2.1.4 (Prime). I is *prime* if $ab \in I$, then $a \in I$ or $b \in I$.

Definition 2.1.5 (Maximal). I is *maximal* if there is no *proper ideal* $J \supsetneq I$.

Remark. *Maximal* \supseteq *Prime* \supseteq *Radical*.

Definition 2.1.6 (Polynomial ring). Let R be a ring. Then $R[x_1, \dots, x_n]$ is the *polynomial ring* with coefficients in R on indeterminates x_1, \dots, x_n .

Example. Let K be a field, $S \subseteq K^n$, and $I \subseteq K[x_1, \dots, x_n]$ defined as

$$I = \{f(\bar{x}) \mid f(\bar{s}) = 0 \text{ for all } \bar{s} \in S\}.$$

Then I is a *radical ideal*.

Definition 2.1.7 (Ideal generation). Let R be a ring. The *ideal* I *generated* by the set $\{x_1, \dots, x_n \in R\}$, denoted as $I = (x_1, \dots, x_n)$, is given by

$$I = \{r_1x_1 + \dots + r_nx_n \mid r_i \in R\}.$$

Intuition. The *ideal generated* by $\{x_i\}$ is the “smallest” *ideal* containing all x_i ’s.

Definition 2.1.8 (Principal ideal). An *ideal* is a *principal ideal* if it’s *generated* by a single element.

Definition 2.1.9 (Principal ideal ring). A ring R is a *principal ideal ring* if all its *ideals* are *principal*.

As previously seen (Zero divisor). If $a, b \neq 0$, but $ab = 0$, then a and b are *zero divisors* of the ring R .

Definition 2.1.10 (Integral domain). A nontrivial ring with no *zero divisors* is called an *integral domain*.^a

^aSome authors will just call *domain*.

Definition 2.1.11 (Principal ideal domain). An **integral domain** where all **ideals** are **principal** is called a *principal ideal domain* or *PID*.

Theorem 2.1.1. $K[x]$ is a **PID**, i.e., every **ideal** is **generated** by one element as $I = (f(x)) = \{g(x)f(x) \mid g(x) \in K[x]\}$.^a

^aIt's clear that $K[x]$ is an **integral domain**.

Proof. We can let g be the polynomial of the least degree in I . Then for any other $h \in I$, by long division, $h = gs + r$, with $\deg(r) < \deg(g)$. But then $r = h - gs \in I$, so if r has lower degree than g , $r = 0$, hence $h = gs \in (g)$. ■

If it's too much to ask for an **ideal generated** by a single element, then we might as well consider the finite case.

Definition 2.1.12 (Noetherian ring). A ring R is *Noetherian* if every **ideal** I of R is finitely generated.

Remark. Equivalently, there is no infinite proper ascending chain of **ideals**.

Theorem 2.1.2 (Hilbert basis theorem). If R is a **Noetherian ring**, then $R[x]$ is also **Noetherian**. In particular, $K[x_1, \dots, x_n]$ is **Noetherian** and so every **ideal** in $K[x_1, \dots, x_n]$ is finitely generated.

As previously seen (Ring homomorphism). Let R, S be rings. A *ring homomorphism* $\varphi: R \rightarrow S$ is a map satisfies

- (a) $\varphi(x +_R y) = \varphi(x) +_S \varphi(y)$ for $x, y \in R$;
- (b) $\varphi(x \times_R y) = \varphi(x) \times_S \varphi(y)$ for $x, y \in R$;
- (c) $\varphi(1_R) = 1_S$.

Theorem 2.1.3. If $\alpha: R \rightarrow S$ is a **ring homomorphism**, then $\ker \alpha$ is an **ideal** of R , and the induced map $\bar{\alpha}: R / \ker \alpha \rightarrow S$ is injective.

Theorem 2.1.4. Let R be a ring, and I an **ideal** of R .^a

- (a) R / I is an **integral domain** if and only if I is a **prime**.
- (b) R / I is a field if and only if I is **maximal**.

^aThen $\pi: R \rightarrow R / I$ is a **ring homomorphism** with kernel I .

Field Extensions

Now, we can talk about **field extension**.

Definition 2.1.13 (Field extension). If $K \subseteq L$ is a subfield of L , we call L / K a *field extension*.

Given a **field extension** L / K , then we have that L is a K -vector space, which suggests the following natural notion.

Definition 2.1.14 (Degree). The *degree* $[L: K]$ of L / K is the dimension of the K -vector space L .

Notation (Finite extension). If $[L: K]$ is finite, then we say L / K is a *finite extension*.

Example. \mathbb{C} is a [field extension](#) over \mathbb{R} with $[\mathbb{C} : \mathbb{R}] = 2$.

Proof. Since \mathbb{C} is an \mathbb{R} -vector space with basis $\{1, i\}$. *

Example. $\mathbb{Q}(\sqrt{2})$ is a [field extension](#) over \mathbb{Q} with $[\mathbb{Q}(\sqrt{2}) : \mathbb{Q}] = 2$.

Proof. Since $\mathbb{Q}(\sqrt{2})$ is a \mathbb{Q} -vector space with basis $\{1, \sqrt{2}\}$. *

The following is the powerful way to calculate the [degree](#) of a [field extension](#) if it can be constructed by a “tower” of [field extensions](#).

Theorem 2.1.5. If M / L and L / K are [field extensions](#), then $[M : K] = [M : L][L : K]$.

Algebraically Closed Fields

We care about [field extensions](#) L / K that are [algebraic](#). This start from defining what does it mean by a single element $a \in L$ is [algebraic](#) over K .

Definition. Let L / K be a [field extension](#), and $a \in L$.

Definition 2.1.15 (Algebraic). If there is a non-zero $f(x) \in K[x]$ such that $f(a) = 0$, then a is *algebraic* over K .

Definition 2.1.16 (Transcendental). If a is not [algebraic](#), then it is *transcendental* over K .

Definition 2.1.17 (Minimal polynomial). If a is [algebraic](#) over K , there is a non-zero, monic^a $f(x) \in K[x]$ of least degree such that $f(a) = 0$ which we call the *minimal polynomial* of a over K .

^aThis is a common practice.

Intuition. An [algebraic](#) number a is the root of some polynomials f in this [polynomial ring](#), and we can find the [minimal](#) such f .

As previously seen (Irreducible). A non-zero non-unit of an [integral domain](#) R is *irreducible* if it cannot be written as the product of two non-units.

Note. A [minimal polynomial](#) is [irreducible](#).

Remark. If $f(x)$ is a [minimal polynomial](#), then $(f(x)) = \{g(x) \in K[x] \mid g(a) = 0\}$.

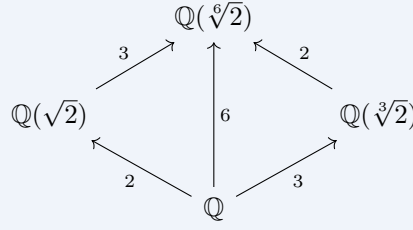
Example. Consider a [field extension](#) \mathbb{R} / \mathbb{Q} with $a = \sqrt{2} \in \mathbb{R}$. Then the [minimal polynomial](#) is $f(x) = x^2 - 2$.

Theorem 2.1.6. Let L / K be a [field extension](#) and $a \in L$, then a is [algebraic](#) over K if and only if $n = [K(a) : K] < \infty$. Furthermore, if a is [algebraic](#) over K , then n is the degree of the [minimal polynomial](#) of a , and $1, a, \dots, a^{n-1}$ is a basis for $K(a)$ as a K -vector space.

Proof idea. Think about $f(a) = a^n + r_{n-1}a^{n-1} + \dots + r_1a + r_0 = 0$. ■

The following example illustrates how can we combine [Theorem 2.1.5](#) and [Theorem 2.1.6](#),

Example. Let $f(x) = x^2 - 2$, $\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\}$.



Theorem 2.1.7. Let L/K be a field extension, $a \in L$, and $f(x) \in K[x]$ be the minimal polynomial of a over K .

(a) $K[x]/(f(x)) \cong K(a)$.^a

(b) If $b \in L$ has the same minimal polynomial as a , then $K(a) \cong K[x]/(f(x)) \cong K(b)$.

^aLet $x \in K[x]$, then $\bar{x} = x + (f(x)) \in K[x]/(f(x))$, i.e., \bar{x} is a root of f , hence the isomorphism is given by $\bar{x} \mapsto a$.

Example. Let $a = \sqrt{2}$, $b = -\sqrt{2}$, and $f(x) = x^2 - 2$ with $K = \mathbb{Q}$. Then

$$\begin{aligned} \mathbb{Q}(\sqrt{2}) &\cong \mathbb{Q}[x]/(x^2 - 2) \cong \mathbb{Q}(-\sqrt{2}); \\ a + b\sqrt{2} &\mapsto [a + bx] \mapsto a - b\sqrt{2}. \end{aligned}$$

Then, it's now natural to talk about a **algebraic extension**.

Definition 2.1.18 (Algebraic extension). Let L/K be a field extension. Then L is an *algebraic extension* of K if all $a \in L$ are algebraic over K .

If a is algebraic over K , then $K(a)/K$ is algebraic: If $b \in K(a)$, then $K(b) \subseteq K(a)$, so $[K(b):K] \leq [K(a):K] < \infty$, so b is algebraic over K .

Theorem 2.1.8. If M/L and L/K are algebraic extensions, then M/K is an algebraic extension.

Proof. Let $a \in M$, and let $b_1, \dots, b_n \in L$ be the coefficients of the minimal polynomial of a over L . Then b_1, \dots, b_n are algebraic over K . Since

$$\begin{aligned} [K(a):K] &\leq [K(a, b_1, \dots, b_n):K] \\ &= [K(a, b_1, \dots, b_n):K(b_1, \dots, b_n)] \cdot [K(b_1, \dots, b_n):K(b_2, \dots, b_n)] \cdots [K(b_n):K]. \end{aligned}$$

Since each of these is a finite extension, so $[K(a):K] < \infty$. ■

Definition 2.1.19 (Algebraically closed). A field L is *algebraically closed* if any non-constant $f(x) \in L[x]$ has a root in L .

Definition 2.1.20 (Algebraic closure). If L/K , then L is an *algebraic closure* of K if L is algebraically closed and an algebraic extension of K .

Remark. Over an algebraically closed field K , any polynomial $f(x) \in K[x]$ factors completely into $f(x) = (x - a_1) \cdots (x - a_n)$ for $n = \deg f$.

Example. \mathbb{C} is algebraically closed, while \mathbb{R} is not.

Example. \mathbb{C} is the algebraic closure of \mathbb{R} , and $[\mathbb{C} : \mathbb{R}] = 2$.

Example. $\mathbb{Q}^{\text{alg}} = \{a \in \mathbb{C} \mid a \text{ is algebraic over } \mathbb{Q}\}$ is the algebraic closure of \mathbb{Q} .

If L is algebraically closed, any $f(x) \in L[x]$ factors completely as $f(x) = (x - a_1) \cdots (x - a_n)$ and a_1, \dots, a_n are the only roots of f .

Theorem 2.1.9. Every field K has an algebraic closure. If L/K and M/K are algebraic closures over K , then $L \cong_K M$.^a

^aThere exists $\alpha: L \rightarrow M$ such that $\alpha(a) = a$ for $a \in K$.

Proof. First, we show the existence. Let f_1, f_2, \dots be (non-constant) polynomials over K . Start with $K = K_0$, let $g_1(x)$ be an irreducible factor of $f_1(x)$ and consider

$$K_1 := K_0[x] / (g_1(x)).$$

Since g_1 is irreducible, $(g_1(x))$ is maximal, so K_1 is a field with a root of f_1 . Now, we build

$$K = K_0 \subseteq K_1 \subseteq K_2 \subseteq \dots \subseteq K^* = \bigcup_i K_i$$

in the same way such that K_i contains a root of $f_i(x)$. Since any $f(x) \in K$ has a root in K^* , so K^*/K is algebraic. Now, we do the same construction for K^* to get

$$K \subseteq K^* \subseteq K^{**} \subseteq K^{***} \subseteq \dots \subseteq L = \bigcup K^{*...},$$

then L is algebraically closed since any non-constant polynomial with coefficients in L actually has coefficients in one of the $K^{*...}$, so it has a root in the next field. Now we prove the uniqueness.

Lemma 2.1.1. An algebraically closed field L has no proper algebraic extensions M .

Proof. If $a \in M$ is algebraic over L for some M , the minimal polynomial $f(x)$ of a factors completely (irreducible), so $f(x) = x - r$ for $r \in L$ with $f(a) = 0$, i.e., $a = r$, so $M = L$. ■

Lemma 2.1.2. Let L/K algebraic, M/K algebraically closed. Then there is an embedding $\alpha: L \rightarrow M$ fixing K .

Proof. Consider the case that $L = K(a)$ with a algebraic over K , and let $f(x)$ be the minimal polynomial of a over K . Then there is a root $b \in M$ of f with $K(a) \cong K[x]/(f) \cong K(b) \subseteq L$ from Theorem 2.1.7. Let this isomorphism be our α . ■

^aOnce this is done, repeat iteratively and get the general case by using Zorn's lemma or transfinite induction.

Hence, if L/K and M/K are algebraic closures over K , there is an embedding $\alpha: L \rightarrow M$ over K .

Finally, since $M/\alpha(L)$ is an algebraic extension, and $\alpha(L) \cong L$ is algebraically closed, by Lemma 2.1.1, $M = \alpha(L)$, so α is an isomorphism $L \rightarrow M$ over K . ■

Lecture 12: The ACF Theory and Categorical

Definition 2.1.21 (Characteristic). A field F has finite characteristic $p > 0$ if $\underbrace{1 + \dots + 1}_{p \text{ times}} = 0$.

14 Feb. 14:30

Remark. p is always prime, otherwise, F has characteristic $p = 0$, i.e., $1 + \dots + 1 \neq 0$, always.

It's now easy to come up with the following notion naturally.

Definition 2.1.22 (Prime field). The *prime field* \mathbb{F}_p in *characteristic* p such that $\mathbb{F}_p = \mathbb{Q}$ if $p = 0$, $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ if $p > 0$.

Definition 2.1.23 (Transcendence basis). Let L/K be a *field extension*. A set $S \subseteq L$ is called a *transcendence basis* of L/K if S is algebraically independent^a and L is an *algebraic extension* of $K(S)$, i.e., S is maximal.

^aNo $a_1, \dots, a_n \in S$ have non-zero polynomial $f(x_1, \dots, x_n) \in K[\bar{x}]$ with $f(a_1, \dots, a_n) = 0$.

Remark. Every *field extension* has a *transcendence basis* (and all *transcendence basis* have the same size).

Proof. On a combinatorial level, this is exactly the same as the proof that any two bases for a vector space have the same cardinality. \circledast

Example. Let $K(t_1, \dots, t_n)$ be the fraction field of $K[x_1, \dots, x_n]$, then $\{t_1, \dots, t_n\}$ is a *transcendence basis* for $K(t_1, \dots, t_n)$ over K .

Definition 2.1.24 (Transcendence degree). The *transcendence degree* of L over K is the cardinality of any *transcendence basis*.

If we do not specify K , then K is the *prime field* $K = \mathbb{F}_p$.

Theorem 2.1.10. Any two *algebraically closed* fields of the same *characteristic* p and *transcendence degree* are isomorphic.

Proof. Let L, K be those fields, with *transcendence basis* S, T over \mathbb{F}_p with $|S| = |T|$. L is the *algebraic closure* of $\mathbb{F}_p(S)$ and K is the *algebraic closure* of $\mathbb{F}_p(T)$. There is a bijection $f: S \rightarrow T$, and then f extends to $\bar{f}: \mathbb{F}_p(S) \rightarrow \mathbb{F}_p(T)$ such that

$$\bar{f}\left(\frac{\sum_{\alpha} r_{\alpha} \bar{x}^{\alpha}}{\sum_{\alpha} s_{\alpha} \bar{x}^{\alpha}}\right) = \frac{\sum_{\alpha} r_{\alpha} f(\bar{x})^{\alpha}}{\sum_{\alpha} s_{\alpha} f(\bar{x})^{\alpha}},$$

where $r_{\alpha}, s_{\alpha} \in \mathbb{F}_p$ and \bar{x}^{α} is some monomial from S , e.g., $x_1^2 x_2$ for $x_1, x_2 \in S$.^a

$\mathbb{F}_p(S)$ and $\mathbb{F}_p(T)$ are the same (up to isomorphism), but the *algebraic closures* are unique from [Theorem 2.1.9](#), so $K \cong L$ via an isomorphism extending \bar{f} . \blacksquare

^a α can be thought as a tuple, in the case of $x_1^2 x_2$, $\alpha = (2, 1)$.

The above proof actually shows more.

Corollary 2.1.1. If L/K and M/K are *field extensions* with *transcendence bases* S and T , and $\alpha: S \rightarrow T$ is a bijection, then α extends to an isomorphism $L \cong_K M$.

If we apply this inside a single *algebraically closed* field, we have the following.

Theorem 2.1.11. Let K be the *algebraic closure* of k , and L, M be subfields of K which *extend* k . Suppose that $\alpha: M \rightarrow L$ is an isomorphism fixing k , then α extends to an automorphism of K .

2.1.2 The ACF Theory

Finally, we are ready to introduce the *theory* we're going to study, which is called ACF. It turns out that the *models* of which are exactly the *algebraically closed* fields with nice properties we're going to discuss.

Definition 2.1.25 (ACF). ACF is the theory of algebraically closed fields consists of field axioms and formulas that for every $n \geq 1$,

$$\forall a_0 \dots \forall a_n (a_n \neq 0 \rightarrow \exists b \, a_n b^n + a_{n-1} b^{n-1} + \dots + a_0 = 0).$$

Remark. The models of ACF are exactly the algebraically closed fields with the language $\mathcal{L} = \mathcal{L}_{\text{ring}} = \{0, 1, +, -, \cdot\}$.

Notation (ACF_p) . For a prime $p > 0$, let $\text{ACF}_p := \text{ACF} \cup \{\underbrace{1 + \dots + 1}_p = 0\}$.

Notation (ACF_0) . Let $\text{ACF}_0 := \text{ACF} \cup \{\underbrace{1 + \dots + 1}_n \neq 0 \mid n \in \mathbb{N}\}$.

Definition 2.1.26 (Categorical). Let κ be an infinite cardinal and T be an \mathcal{L} -theory. T is κ -categorical if any $\mathcal{M}, \mathcal{N} \models T$ of size κ have $\mathcal{M} \cong \mathcal{N}$.

Definition 2.1.27 (Countably categorical). If κ is countable, then T is countably categorical.

Definition 2.1.28 (Uncountably categorical). If κ is uncountable, then T is uncountably categorical.

We see that for being uncountably categorical, we only need one uncountable κ .

Example. (\mathbb{Q}, \leq) is countably categorical.

Lemma 2.1.3. If K has transcendence degree λ , then $|K| = \lambda + \aleph_0$.

Proof. K is algebraic over $\mathbb{F}_p(S)$, where S is a transcendence basis of size λ . By counting, $|\mathbb{F}_p(S)| = \lambda + \aleph_0$, so $|\mathbb{F}_p(S)[x]| = \lambda + \aleph_0$. But since each element of K satisfies some polynomials, and each polynomial has finitely many roots in K , so $|K| = \lambda + \aleph_0$. ■

Theorem 2.1.12. Fix p . ACF_p is κ -categorical for every uncountable κ .

Proof. Let L, K be ACF_p for size κ , then L, K have transcendence degree κ , and hence are isomorphic from Theorem 2.1.10. With the application of Lemma 2.1.3, we're done. ■

Example. \mathbb{Q}^{alg} , the algebraic closure of \mathbb{Q} , has size \aleph_0 , and has transcendence degree is 0.

Example. $\mathbb{Q}(t)^{\text{alg}}$, the algebraic closure of $\mathbb{Q}(t) \cong \mathbb{Q}(\pi)$, has size \aleph_0 , and has transcendence degree is 1.

Proof. We see that

$$\mathbb{Q}(t)^{\text{alg}} = \{z \in \mathbb{C} \mid z \text{ is algebraic over } \mathbb{Q}(\pi)\}.$$

These are countable, but not isomorphic. ACF_0 is not countably categorical. The same with ACF_0 for $p > 0$. ⊗

Note. ACF is not uncountably categorical.

Theorem 2.1.13 (Vaught's test). Let T be a satisfiable \mathcal{L} -theory with no finite models. If T is κ -categorical for some infinite $\kappa \geq |\mathcal{L}|$, then T is complete.

Proof. Suppose T was not complete, so pick φ with $T \not\models \varphi$ and $T \not\models \neg\varphi$, and hence $T \cup \{\varphi\}$ and $T \cup \{\neg\varphi\}$ are satisfiable. By a consequence of the proof of completeness theorem (with a compactness argument),

- $T \cup \{\varphi\}$ has a model \mathcal{M} of size κ , and
- $T \cup \{\neg\varphi\}$ has a model \mathcal{N} of size κ .

But T is κ -categorical, so $\mathcal{M} \cong \mathcal{N}$, which is a contradiction \nmid ■

Corollary 2.1.2. ACF_p is complete for each p .

The axioms for ACF_p completely determines all first-order facts about algebraically closed fields of characteristic p .

Remark (Fact). The axioms for ACF or ACF_p can be listed computably. So $\{\varphi \mid \text{ACF} \models \varphi\}$ and $\{\varphi \mid \text{ACF}_p \models \varphi\}$ can be listed computably.

Definition 2.1.29 (Decidable). A theory T is *decidable* if there is a program that given φ , it determines whether $T \models \varphi$.

Remark. ACF_p is decidable.

Proof. Given φ , either $\text{ACF}_p \models \varphi$ or $\text{ACF}_p \models \neg\varphi$ since ACF_p is complete. By looking for a proof of φ and a proof of $\neg\varphi$, eventually we will find one, telling us whether $\text{ACF}_p \models \varphi$. \circledast

Theorem 2.1.14. ACF is decidable.

Proof. Given φ , simultaneously

- Look for a proof of $\text{ACF} \vdash \varphi$, and
- Look for p such that $\text{ACF}_p \vdash \neg\varphi$ (so $\text{ACF} \not\models \varphi$).^a

The first case is fine. Suppose $\text{ACF} \not\models \varphi$, so there is $\mathcal{M} \models \text{ACF}$, $\mathcal{M} \models \neg\varphi$. There is p such that $\mathcal{M} \models \text{ACF}_p$. Since ACF_p is complete, $\text{ACF}_p \models \neg\varphi$, so the search of the second case will halt, so the whole search will eventually halt. ■

^aWe don't know $\text{ACF} \models \neg\varphi$.

Lecture 13: Upward Löwenheim-Skolem Theorem

Corollary 2.1.3 (Leftschetz principle). Let \mathcal{L} be the language of rings. For an \mathcal{L} -sentence φ , the following are equivalent:

- $\mathbb{C} \models \varphi$
- every algebraically closed field of characteristic 0 $\models \varphi$
- some algebraically closed fields of characteristic 0 $\models \varphi$
- for all sufficient large positive p , φ is true in all algebraically closed fields of characteristic p
- for infinitely many positive p , φ is true in all algebraically closed fields of characteristic p

16 Feb. 14:30

Proof. We only show the first three, others are left as homework. Let $K \models \text{ACF}_0$, then since it's complete,

$$K \models \varphi \Leftrightarrow \text{ACF}_0 \models \varphi.$$

Theorem 2.1.15 (Ax-Grothendieck theorem). Let $f: \mathbb{C}^n \rightarrow \mathbb{C}^n$ be a polynomial map.^a If f is injective, then it's surjective. More generally, this is true for any $K \models \text{ACF}_p$ for any p .

^aI.e., $f(\bar{x}) = (f_1(\bar{x}), \dots, f_n(\bar{x}))$ where f_1, \dots, f_n are polynomials.

Proof. The claim can be expressed by the sentences, so by Leftschetz principle, it's enough to prove that if for $K = \overline{\mathbb{F}_p}$, for each $p > 0$. Let $f: \overline{\mathbb{F}_p}^n \rightarrow \overline{\mathbb{F}_p}^n$ be an injective polynomial map and $\bar{y} \in \overline{\mathbb{F}_p}^n$. Then there is a finite subfield $L \subseteq \overline{\mathbb{F}_p}$ which contains \bar{y} and the coefficients of f . Then, f restricts to an injective function $L^n \rightarrow L^n$, which is surjective because L^n is finite, so $\exists \bar{x} \in L^n$ such that $f(\bar{x}) = \bar{y}$. ■

2.2 Up and Down

Definition. Let \mathcal{M} be an \mathcal{L} -structure. Let $\mathcal{L}_M \supseteq \mathcal{L}$ be the expanded language with a new constant symbol \underline{a} for each $a \in M$.

Definition 2.2.1 (Atomic diagram). The *atomic diagram* of \mathcal{M} is the \mathcal{L}_M -theory

$$\text{Diag}(\mathcal{M}) := \{\varphi(\underline{a}_1, \dots, \underline{a}_n) \mid \mathcal{M} \models \varphi \text{ and } \varphi \text{ is atomic or negated of atomic}\}.$$

Definition 2.2.2 (Elementary diagram). The *elementary diagram* of \mathcal{M} is the \mathcal{L}_M -theory

$$\text{Diag}_{\text{el}}(\mathcal{M}) := \{\varphi(\underline{a}_1, \dots, \underline{a}_n) \mid \mathcal{M} \models \varphi \text{ and } \varphi \text{ an } \mathcal{L}\text{-formula}\}.$$

Note. There's a canonical way of expanding \mathcal{M} to an \mathcal{L}_M -structure with $\underline{a}^{\mathcal{M}} := a$, i.e., we write a for both the symbol and the element.

Lemma 2.2.1. Let \mathcal{N} be an \mathcal{L}_M -structure.

- (a) If $\mathcal{N} \models \text{Diag}(\mathcal{M})$ then, viewing \mathcal{N} as an \mathcal{L} -structure, there is an embedding $f: \mathcal{M} \rightarrow \mathcal{N}$.
- (b) If $\mathcal{N} \models \text{Diag}_{\text{el}}(\mathcal{M})$, then there is an elementary \mathcal{L} -embedding of \mathcal{M} into \mathcal{N} .

Proof. Take $f(a) = \underline{a}^{\mathcal{N}}$, then $\mathcal{N} \models \text{Diag}(\mathcal{M})$ means exactly that f is an embedding, and $\mathcal{N} \models \text{Diag}_{\text{el}}(\mathcal{M})$ means that f is an elementary embedding. ■

Theorem 2.2.1 (Upward Löwenheim-Skolem theorem). Let \mathcal{M} be an infinite \mathcal{L} -structure and let κ be an infinite cardinal $\kappa \geq |\mathcal{M}| + |\mathcal{L}|$. Then there is an \mathcal{L} -structure \mathcal{N} of cardinality κ such that $j: \mathcal{M} \rightarrow \mathcal{N}$ is elementary.

Proof. $\text{Diag}_{\text{el}}(\mathcal{M})$ is satisfiable since $\mathcal{M} \models \text{Diag}_{\text{el}}(\mathcal{M})$, so by Proposition 2.1.1, it has a model \mathcal{N} of cardinality $\kappa \geq |\mathcal{L}_M|$, and by Lemma 2.2.1, there is an elementary embedding $\mathcal{M} \rightarrow \mathcal{N}$. ■

Proposition 2.2.1 (Tarski-Vaught Test). Let \mathcal{M} be a substructure of \mathcal{N} . Then \mathcal{M} is an elementary substructure of \mathcal{N} if and only if for any formula $\varphi(x, \bar{y})$ and $\bar{a} \in M^n$, if there is $b \in N$ such that $\mathcal{N} \models \varphi(b, \bar{a})$, then there is $c \in M$ such that $\mathcal{N} \models \varphi(c, \bar{a})$.

Proof. The forward direction follows from the fact that \mathcal{M} is an **elementary substructure**, so the **truth** of $\exists x \varphi(x, \bar{y})$ is proved.

For the backward direction, suppose the condition holds. We show that $\mathcal{M} \models \varphi(\bar{a}) \Leftrightarrow \mathcal{N} \models \varphi(\bar{a})$ by induction on φ . Suppose the claim holds for φ, ψ . Then,

$$\mathcal{M} \models (\varphi \wedge \psi)(\bar{a}) \Leftrightarrow \mathcal{M} \models \varphi(\bar{a}) \text{ and } \mathcal{M} \models \psi(\bar{a}) \Leftrightarrow \mathcal{N} \models \varphi(\bar{a}) \text{ and } \mathcal{N} \models \psi(\bar{a}) \Leftrightarrow \mathcal{N} \models (\varphi \wedge \psi)(\bar{a}).$$

Finally, suppose the claim holds for $\varphi(x, \bar{y})$, then

$$\mathcal{M} \models \exists x \varphi(x, \bar{a}) \Leftrightarrow \exists b \in M \mathcal{M} \models \varphi(b, \bar{a}) \Leftrightarrow \exists b \in M \mathcal{N} \models \varphi(b, \bar{a})$$

by induction hypotheses. Conversely, $\mathcal{N} \models \exists x \varphi(x, \bar{a})$, then $\exists b \in N$ such that $\mathcal{N} \models \varphi(b, \bar{a})$ by the condition from the statement, so $\exists c \in M$ such that $\mathcal{N} \models \varphi(c, \bar{a})$. By the induction hypotheses, we further have $\mathcal{M} \models \varphi(c, \bar{a})$, hence $\mathcal{M} \models \exists x \varphi(x, \bar{a})$. ■

Example. The ring \mathbb{Z} is a **substructure** of \mathbb{Q} , but $\mathbb{Q} \models \exists x (x + x = 1)$ while $\mathbb{Z} \not\models \exists x (x + x = 1)$.

Lecture 14: Downward Löwenheim-Skolem theorem

21 Feb. 14:30

Definition 2.2.3 (Built-in Skolem functions). We say an **\mathcal{L} -theory** T has *built-in Skolem functions* if for all **\mathcal{L} -formulas** $\varphi(x, y_1, \dots, y_n)$, there is a function symbol f such that

$$T \models \forall \bar{y} (\exists x \varphi(x, \bar{y}) \rightarrow \varphi(f(\bar{y}), \bar{y})).$$

Lemma 2.2.2. Let T be an **\mathcal{L} -theory**, then there are $\mathcal{L}^* \supseteq \mathcal{L}$ and $T^* \supseteq T$ an **\mathcal{L}^* -theory** such that T^* has **built-in Skolem functions**. Moreover, if $\mathcal{M} \models T$, then we can expand \mathcal{M} to $\mathcal{M}^* \models T^*$.

Proof. Start with $\mathcal{L}_0 = \mathcal{L}$ and $T_0 = T$, we build $\mathcal{L}^* = \bigcup_i \mathcal{L}_i$ and $T^* = \bigcup_i T_i$. Given \mathcal{L}_i and T_i , define

$$\mathcal{L}_{i+1} = \mathcal{L}_i \cup \{f_\varphi \mid \varphi(x, \bar{y}) \text{ is an } \mathcal{L}_i\text{-formulas}\}$$

where the arity of f_φ is the same as \bar{y} , and

$$T_{i+1} = T_i \cup \{\forall \bar{y} (\exists x \varphi(x, \bar{y}) \rightarrow \varphi(f_\varphi(\bar{y}), \bar{y}))\}.$$

Now, we argue that if $\mathcal{M} \models T_i$, we can expand it to a **model** \mathcal{M}^* of T_{i+1} . Pick $c \in M$ a “default value.” Given φ and \bar{a} , define $f_\varphi^{\mathcal{M}^*}(\bar{a})$ to be some b with $\mathcal{M} \models \varphi(b, \bar{a})$ if such a b exists, or c otherwise. Then, $\mathcal{M}^* \models T_{i+1}$.

Now, T^* has **built-in Skolem functions**. Suppose $\mathcal{M} \models T$, i.e., $\mathcal{M} = \mathcal{M}_0 \models T_0$. Then \mathcal{M}_0 has an **expansion** $\mathcal{M}_1 \models T_1$, which has an **expansion** $\mathcal{M}_2 \models T_2$, etc. Then, $\mathcal{M}^* = \bigcup_i \mathcal{M}_i$ is a **model** of T^* , and by counting, we have $|\mathcal{L}^*| = |\mathcal{L}| + \aleph_0$. ■

Notation (Skolemization). We call T^* in **Lemma 2.2.2** a *Skolemization* of T .

Theorem 2.2.2 (Downward Löwenheim-Skolem theorem). Let \mathcal{M} be an **\mathcal{L} -structure** and $X \subseteq M$. Then there is an **elementary substructure** $X \subseteq \mathcal{N}$ of \mathcal{M} of cardinality $|\mathcal{N}| \leq |\mathcal{L}| + \aleph_0 + |X|$.

Proof. By **expanding** the **language**, we get \mathcal{M}^* and **\mathcal{L}^* -structure** with $\text{Th}(\mathcal{M}^*)$ has **built-in Skolem functions** (where $T = \text{Th}(\mathcal{M})$ in **Lemma 2.2.2**). Replacing \mathcal{M} by \mathcal{M}^* , etc., we may assume that we already had **built-in Skolem functions**.

Start with $X_0 = X \cup \{c^{\mathcal{M}} \mid c \text{ a constant symbol}\}$. Given X_i , define X_{i+1} as

$$X_{i+1} = X_i \cup \{f(\bar{a}) \mid f \text{ a function symbol, } \bar{a} \in X_i\}.$$

Let $N = \bigcup_i X_i$, let \mathcal{N} be the **substructure** of \mathcal{M} with domain N .^a Now, to show that \mathcal{N} is an **elementary substructure** of \mathcal{M} , we use the **Tarski-Vaught test**. Suppose that we have an **\mathcal{L} -formula**

$\varphi(x, \bar{y})$, $\bar{a} \in N$, $b \in M$ such that $\mathcal{M} \models \varphi(b, \bar{a})$ (we must replace b by $c \in N$). Since $\mathcal{M} \models \exists x \varphi(x, \bar{a})$, so $\mathcal{M} \models \varphi(f_\varphi(\bar{a}), \bar{a})$. But since $\bar{a} \in N$, so $f_\varphi(\bar{a}) \in N$, so the **Tarski-Vaught test** says \mathcal{N} is an **elementary substructure** of \mathcal{M} .

Finally, we see that $|X_0| \leq |X| + |\mathcal{L}| + \aleph_0$, and since N is a countable union, so

$$|N| \leq |X| + |\mathcal{L}| + \aleph_0.$$

■

^aHere, \mathcal{N} is called the *substructure generated by X* .

Example (Countable real closed field). Consider $\mathbb{R} = (\mathbb{R}, 0, 1, +, -, \cdot, \leq)$. Let $X \subseteq \mathbb{R}$ be countable, e.g., $X = \emptyset$ or $X = \{\pi, e\}$. Then there is $X \subseteq R \preceq \mathbb{R}$ such that R is countable. In particular, $\text{Th}(R) = \text{Th}(\mathbb{R})$. In this way, R is a *countable real closed field*:

- (a) -1 is not a sum of squares;
- (b) for all a , there is b such that $a = b^2$ or $a = -b^2$;
- (c) every odd degree polynomial has a root.

Example (Skolem's paradox). Let $\mathcal{L} = \{\in\}$, where \in a binary relation symbol. Let $T = \text{ZFC}$. Suppose that ZFC is a **satisfiable**,^a and let $\mathcal{M} \models T$. Then there is $\mathcal{N} \preceq \mathcal{M}$ with \mathcal{N} countable. Then,

$$\mathcal{N} \models \text{"there is no bijection between } \mathbb{R}^{\mathcal{N}} \text{ and } \mathbb{N}^{\mathcal{N}}\text{"}.$$

\mathcal{N} thinks that it contains an uncountable set $\mathbb{R}^{\mathcal{N}}$, but $\{a \in N \mid \mathcal{N} \models a \in \mathbb{R}^{\mathcal{N}}\} \subseteq N$ is countable. This is called *Skolem's paradox*.

^aFrom **Gödel's incompleteness theorem**, in ZFC, one can't prove that ZFC is **consistent**.

Definition 2.2.4 (Universally axiomatizable). Let T be an \mathcal{L} -theory, then T is *universally axiomatizable* if there is a set Γ of **universal sentences** such that $T \models \Gamma$ and $\Gamma \models T$.

Theorem 2.2.3. Let T be an \mathcal{L} -theory. T is **universally axiomatized** if and only if whenever $\mathcal{N} \models T$ and $\mathcal{M} \subseteq \mathcal{N}$, then $\mathcal{M} \models T$.

Proof. We already know the forward direction. Now, to prove the backward direction, suppose that if $\mathcal{N} \models T$, $\mathcal{M} \subseteq \mathcal{N}$, then $\mathcal{M} \models T$. Define

$$\Gamma = \{\varphi \text{ universal} \mid T \models \varphi\},$$

then $T \models \Gamma$. Now, we show that $\Gamma \models T$. We may assume that T is **satisfiable**^a and let $\mathcal{M} \models \Gamma$. We must prove that $\mathcal{M} \models T$. We will do this by finding $\mathcal{N} \supseteq \mathcal{M}$, $\mathcal{N} \models T$, which implies $\mathcal{M} \models T$. We build such an \mathcal{N} by showing that $\text{Diag}(\mathcal{M}) \cup T$ is **satisfiable** with **compactness theorem**. Let $\Delta \subseteq \text{Diag}(\mathcal{M}) \cup T$ be finite, then there is a finite set of **atomic** or negated **atomic formulas** $\varphi_1, \dots, \varphi_\ell$ and $m_1, \dots, m_k \in M$ such that

$$\Delta \subseteq \{\varphi_1(\bar{m}), \dots, \varphi_\ell(\bar{m})\} \cup T.$$

We may assume that they are actually equal. To show that Δ is **satisfiable**, it is enough to show that

$$\{\exists x_1 \dots \exists x_k (\varphi_1(\bar{x}) \wedge \dots \wedge \varphi_\ell(\bar{x}))\} \cup T$$

is **satisfiable**. If not, then $T \models \forall x_1 \dots \forall x_k \neg(\varphi_1(\bar{x}) \wedge \dots \wedge \varphi_\ell(\bar{x}))$. But this is **universal**, hence in Γ , so it is **true** in \mathcal{M} , i.e., $\mathcal{M} \models \varphi_1(\bar{m}) \wedge \dots \wedge \varphi_\ell(\bar{m})$ and $\mathcal{M} \models \forall x_1 \dots \forall x_k \neg(\varphi_1(\bar{x}) \wedge \dots \wedge \varphi_\ell(\bar{x}))$, a contradiction \nmid . Hence, Δ is **satisfiable**, so any finite subset is **satisfiable**, by **compactness theorem**, we're done. ■

^aSince otherwise $\Gamma \ni \forall x \, x \neq x$.

Lecture 15

23 Feb. 14:30

Proposition 2.2.2. Suppose $\mathcal{M}_1 \preceq \mathcal{M}_2 \preceq \dots$, and let $\mathcal{M} = \bigcup_i \mathcal{M}_i$. Then $\mathcal{M}_i \preceq \mathcal{M}$ for all i .^a

^aCountability is not necessary.

Proof. By induction on [formulas](#), we prove that for all i and $\bar{a} \in M_i$, $\mathcal{M}_i \models \varphi(\bar{a})$ if and only if $\mathcal{M} \models \varphi(\bar{a})$:

- (a) For φ [atomic](#), this is true since φ is [quantifier-free](#).
- (b) For \neg, \vee, \wedge , exactly the same as the [Tarski-Vaught test](#).
- (c) If φ is $\exists y \psi(\bar{x}, y)$:
 - If $\mathcal{M}_i \models \exists y \psi(\bar{a}, y)$, there is $b \in M_i$ such that $\mathcal{M}_i \models \psi(\bar{a}, b)$. Then by the induction hypothesis, $\mathcal{M} \models \psi(\bar{a}, b)$, so $\mathcal{M} \models \exists y \psi(\bar{a}, y)$.
 - If $\mathcal{M} \models \exists y \psi(\bar{a}, y)$, then there is $b \in M$ such that $\mathcal{M} \models \psi(\bar{a}, b)$. There is $j \geq i$ such that $b \in M_j$. By the induction hypothesis, $\mathcal{M}_j \models \psi(\bar{a}, b)$, so $\mathcal{M}_j \models \exists y \psi(\bar{a}, y)$. But $\mathcal{M}_i \preceq \mathcal{M}_j$, so $\mathcal{M}_i \models \exists y \psi(\bar{a}, y)$.

■

2.3 Back and Forth

2.3.1 Dense Linear Order Theory

Definition 2.3.1 (DLO). Let $\mathcal{L} = \{\leq\}$. The [theory](#) DLO of dense linear orders (without endpoints) has the axioms:

- (a) \leq is a linear order;
- (b) $\forall x \forall y (x < y \rightarrow \exists z x < z < y)$;
- (c) $\forall x \exists y \exists z (y < x < z)$.

Example. (\mathbb{Q}, \leq) and (\mathbb{R}, \leq) are DLO.

Example. Given $\mathcal{M}_1, \mathcal{M}_2$ two DLO's, $\mathcal{M}_1 + \mathcal{M}_2$ is also a DLO.

Example. $\mathbb{Q} + \mathbb{R} \not\cong \mathbb{R}$, so DLO is not $|\mathbb{R}| = 2^{\aleph_0}$ -categorical. In fact, not κ -categorical for any $\kappa \geq 2^{\aleph_0}$. Also, $\mathbb{Q} + \mathcal{M} \not\cong \mathbb{R} + \mathcal{M}$ for any \mathcal{M} . On the other hand, $\mathbb{Q} + \mathbb{Q} \cong \mathbb{Q}$ by considering picking an irrational number in \mathbb{Q} and split \mathbb{Q} into two.

Theorem 2.3.1. DLO is [countably categorical](#) and hence [complete](#).

Proof. Let (A, \leq) and (B, \leq) be two countable DLO's. Let a_1, a_2, \dots list A and b_1, b_2, \dots list B . We will build $f: A \rightarrow B$ an [isomorphism](#) stage-by-stage. At state i , $A_i \subseteq A$ finite, $B_i \subseteq B$ finite, and $f_i: A_i \rightarrow B_i$ a bijection. We will have $f_i \subseteq f_{i+1}$, $A_i \subseteq A_{i+1}$, and $B_i \subseteq B_{i+1}$, and each f_i will be a “partial [isomorphism](#),” i.e., if $a < a'$, then $f(a) < f(a')$. Finally, we need to make sure that $\bigcup_i A_i = A$, and $\bigcup_i B_i = B$, so $f = \bigcup_i f_i$ is an [isomorphism](#) from $A \rightarrow B$.

- Stage 0: $A_0 = \emptyset$, $B_0 = \emptyset$, $f_0 = \emptyset$.
- Stage $i + 1 = 2k + 1$: put a_k into A_{i+1} . If $a_k \in A_i$, then do nothing, i.e., $A_{i+1} = A_i$, $B_{i+1} = B_i$, $f_{i+1} = f_i$; otherwise, we have three possibilities:

- a_k is less than all of A_i : choose $b \in B$ less than all of B_i , and $f_{i+1}(a_k) = b$, and let $A_{i+1} = A_i \cup \{a_k\}$, $B_{i+1} = B_i \cup \{b\}$.^a
- There are a and a' in A_i such that a_k is between a and a' , say $a < a_k < a'$, and there is no other element of A_i between a and a' since A_i is finite: pick b with $f_i(a) < b < f_i(a')$, set $f_{i+1}(a_k) = b$, etc.^b
- a_k greater than all of A_i : similar to the first case.

- Stage $i + 1 = 2k + 2$: put b_k into $B_{i+1} = \text{Im}(f_{i+1})$. This is exactly the same, but in the other direction.

At the end, we get $f: A \rightarrow B$ an **isomorphism**, and hence DLO is **complete** because it is **countably categorical**. ■

^aSuch b exists since B_i is finite, and (B, \leq) has no left endpoint.

^bSuch b exists since B_i is finite and (B, \leq) is dense.

Corollary 2.3.1. $\mathbb{Q} + \mathbb{R} \equiv \mathbb{R}$.

Proof. Since $\text{Th}(\mathbb{Q} + \mathbb{R}) = \text{Th}(\mathbb{R}) = \{\varphi \mid \text{DLO} \models \varphi\}$. ■

Definition 2.3.2 (Complete). A linear order is *complete* if every subset bounded above has a least upper bound.

Corollary 2.3.2. There is no first order **sentence** φ such that $\mathcal{M} \models \varphi$ if and only if \mathcal{M} is a **complete** linear order.

2.3.2 Random Graph Theory

Definition 2.3.3 (Random graph). Fix countably infinitely many vertices v_1, v_2, \dots . Fix p such that $0 < p < 1$. For each pair of vertices, “flip a coin”: with probability p , put an edge; with $1 - p$, no edge.

Now, the question is, what graph do we get?

Remark. With probability 1, we get the same graph up to isomorphism, no matter what p is.

Let $\mathcal{L} = \{E\}$, where E is a binary relation. The **theory** T of **random graphs** has axioms:

- $\forall x \neg xEx$ and $\forall x \forall y (xEy \rightarrow yEx)$;
- $\exists x \exists y x \neq y$;
- for each n , ψ_n define as

$$\psi_n := \forall x_1 \dots \forall x_n \forall y_1 \dots \forall y_n \left[\bigwedge_{i=1}^n \bigwedge_{j=1}^n x_i \neq y_j \rightarrow \exists z \left(\bigwedge_{i=1}^n x_i E z \wedge \neg y_i E z \wedge z \neq x_i \wedge z \neq y_i \right) \right]$$

Intuition. Think of ψ_n as an “extension axiom.” And this axiom happens with probability $p^{|X|} \cdot (1 - p)^{|Y|}$ for a given z for some X, Y .

Note. If $m \leq n$, then $\psi_n \models \psi_m$.

Appendix

Bibliography

- [HH97] W. Hodges and S.M.S.W. Hodges. *A Shorter Model Theory*. Cambridge University Press, 1997. ISBN: 9780521587136. URL: <https://books.google.com/books?id=S6QYeuo4p1EC>.
- [Hin05] P.G. Hinman. *Fundamentals of Mathematical Logic*. Taylor & Francis, 2005. ISBN: 9781568812625. URL: <https://books.google.com/books?id=xA6D8o72qAgC>.
- [Mar02] D. Marker. *Model Theory : An Introduction*. Graduate Texts in Mathematics. Springer New York, 2002. ISBN: 9780387987606. URL: <https://books.google.com/books?id=gkvogoiEnuYC>.
- [Rob49] Julia Robinson. “Definability and decision problems in arithmetic”. In: *The Journal of Symbolic Logic* 14.2 (1949), pp. 98–114. DOI: [10.2307/2266510](https://doi.org/10.2307/2266510).