

MATH681
Mathematical Logic

Pingbang Hu

March 27, 2024

Abstract

This is a graduate-level mathematical logic course taught by [Matthew Harrison-Trainor](#) at University of Michigan, 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	Languages and Structures	3
1.2	Embeddings and Isomorphisms	4
1.3	Terms	5
1.4	Formulas	6
1.5	Truths	7
2	Soundness, Completeness, and Compactness	12
2.1	Theories	12
2.2	Elementary Embeddings	13
2.3	Definable Sets	14
2.4	Proofs	15
2.5	Soundness Theorem	18
2.6	Completeness and Compactness Theorems	21
3	The Beginning of Model Theory	29
3.1	Complete Theories	29
3.2	A Detour to Algebraically Closed Fields	29
3.3	The ACF Theory	36
3.4	Up and Down	38
3.5	Back and Forth	43
4	Quantifier Elimination and Algebraic Applications	47
4.1	Quantifier Elimination	47
4.2	Definable and Constructible Sets	51
4.3	Algebraic Closure	52
4.4	Types	53
4.5	Other Examples of Quantifier Elimination	55
5	Fraïssé Limits	57
5.1	Substructures' Properties	57
5.2	“Baby” Fraïssé Theorem	58
5.3	Fraïssé Theorem	58
6	Ultrapowers	64
6.1	Ultrafilters	64
6.2	Ultraproducts	66
6.3	Ultrapowers	67
7	Epilogue	76
7.1	An Review: Map of the Universe	76
7.2	An Outlook of Real Closed Fields	76

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. 11:30

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

And this is what this chapter aiming to address. We will turn to other topics based on these.

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. *

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. *

Note. Later, we will talk about theories that let us restrict to [structures](#) we want.

1.2 Embeddings and Isomorphisms

We can now consider the relation between [structures](#).

Definition 1.2.1 (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 1.2.2 (Isomorphism). An \mathcal{L} -isomorphism is a bijective \mathcal{L} -embedding.

Definition 1.2.3 (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 1.2.4 (Substructure). \mathcal{M} is a *substructure* of \mathcal{N} .

Definition 1.2.5 (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 1.2.1 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. 11:30

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.

1.3 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 1.3.1 (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 , $t^{\mathcal{M}}(a_1, \dots, a_n) = c^{\mathcal{M}}$.

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

- (b) if t is a variable v_i , $t^{\mathcal{M}}(a_1, \dots, a_n) = v_i$;
 (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 $2 \cdot (x + y)$.

1.4 Formulas

A **term** is just a building block of **formulas**, as we now see.

Definition 1.4.1 (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). Definition 1.4.1 (a) and (b) are called *atomic*.

Notation (Quantifier-free). Definition 1.4.1 (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**.

1.4.1 Bounded and Free Variables

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 1.4.2 (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 1.4.3 (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 $\varphi(x, y, z) := x = y$.

1.5 Truths

Finally, we define the notion of *truth*.

Definition 1.5.1 (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

1.5.1 Implications

12 Jan. 11:30

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

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. \otimes

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). \otimes

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 $\neg(\exists x \neg\varphi)$. \otimes

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 $\varphi \in \Phi$.

1.5.2 Logical Consequences and Equivalent

Definition 1.5.2 (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 1.5.2 is written as $\models \varphi$, i.e., φ is *true* in all \mathcal{L} -structures.²

Definition 1.5.3 (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 1.5.1. 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 .

²Recall that we always have a language \mathcal{L} implicitly.

(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 1.5.2. 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 1.5.1**, 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 1.5.2**, **formulas** as in (a) are called *existential* (\exists_1 or \exists) *formulas*.

Notation (Universal). In **Proposition 1.5.2**, **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

17 Jan. 11:30

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

^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 $e: v \rightarrow w$.

Remark. 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**.

Chapter 2

Soundness, Completeness, and Compactness

In this chapter, we're going to formalize [proofs](#), including what do we mean by “having a proof” of a statement, and study properties of which.

2.1 Theories

Let's start by the notion of [theory](#).

Definition 2.1.1 (Theory). An \mathcal{L} -theory is a set of [\$\mathcal{L}\$ -sentences](#).

Definition 2.1.2 (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 2.1.3 (Satisfiable). A [theory](#) is *satisfiable* if it has a [model](#).

Definition 2.1.4 (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 2.1.5 (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 2.1.6 (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 2.1.7 (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 (Non-standard model of arithmetic). 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. $\mathbb{Z} \oplus \mathbb{Z} \not\cong \mathbb{Z}$ 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 2.1.8 (Finitely axiomatizable). A **theory** is *finitely axiomatizable* if it has a finite set of axioms.

Given a **theory**, consider $T^{\models} = \{\varphi \mid T \models \varphi\}$,¹ 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 $T^{\models} = \Phi^{\models}$.

2.2 Elementary Embeddings

Let's now consider the following notion.

Definition 2.2.1 (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 2.2.2 (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.

¹Recall **Definition 1.5.2**.

Proposition 2.2.1. 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 1.5.1**. 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, $\mathcal{M} \models \psi(\bar{a}, b)$, so $\mathcal{M} \models \varphi(\bar{a})$. ■

Corollary 2.2.1. If $\mathcal{M} \cong \mathcal{N}$, then $\mathcal{M} \equiv \mathcal{N}$.

2.3 Definable Sets

Consider the following.

Definition 2.3.1 (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. 11:30

Proposition 2.3.1. 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 2.2.1**,

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

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.

2.4 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 2.4.1 (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 2.4.2 (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 2.4.3 (Valid). If we **prove** ψ without **hypotheses**, we write $\vdash \psi$ and say ψ is *valid*.

Definition 2.4.4 (Logical axioms). The *logical axioms* are the following **formulas** written in **schema**, as well as all of their **generalizations**:

Definition 2.4.5 (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 2.4.6 (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 2.4.7 (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 $\delta = \epsilon$,

$$\vdots$$

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

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 2.4.8 (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. 11:30

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 2.4.1. 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 2.4.1. 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?

2.5 Soundness Theorem

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

Definition 2.5.1 (Sound). A **proof** system is *sound* if any **provable sentence** φ is **true**.

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

Lemma 2.5.1 (Soundness). If Γ is a set of **\mathcal{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 **\mathcal{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 2.5.1 (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 ψ_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 2.5.1 (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 2.5.2 (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 2.5.3 (Inconsistent). If a theory T is not consistent, then it's *inconsistent*.

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

26 Jan. 11:30

Proposition 2.5.2 (Proof by contradiction). If $\Gamma \cup \{\varphi\}$ is **inconsistent**, then $\Gamma \vdash \neg\varphi$.

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 2.5.3. 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 2.5.4 (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 2.5.5 (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 2.5.6 (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 2.5.7 (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 2.5.2 (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 2.5.3. 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 2.5.3. 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 \otimes

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 2.5.3, one of $T^* \cup \{\varphi\}$ or $T^* \cup \{\neg\varphi\}$ is consistent, hence in P , contradicting to T^* being maximal \nmid \otimes

^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 2.5.1 and Theorem 2.5.3, we close this section with the following.

Theorem 2.5.4 (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 2.5.4. 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

2.6 Completeness and Compactness Theorems

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

Definition 2.6.1 (Complete). A proof system is complete if any true sentence φ is provable.

And indeed, this is the case.

Theorem 2.6.1 (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 2.6.2 (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 (Construction of non-standard model of arithmetic). Let $\mathcal{L} = \{0, 1, +, \cdot, -, <\}$, and $\mathcal{L}^* = \mathcal{L} \cup \{c\}$, where c is a new constant symbol. Then

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

is finitely **satisfiable**.

Proof. Given $T_0 \subseteq T$ finite, $T_0 \subseteq \text{Th}_{\mathcal{L}}(\mathbb{N}) \cup \{c > \underline{n}_1, \dots, c > \underline{n}_\ell\}$, 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

2.6.1 Henkin Construction

31 Jan. 11:30

To prove **Theorem 2.6.1 (b)**, we need an additional definition and a technical lemma due to Henkin.

Definition 2.6.2 (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 2.6.1. 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 2.6.1 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 2.6.1. 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 2.6.3. 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 2.6.1, 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 \nmid^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

2.6.2 Proof of Completeness Theorem

After proving Theorem 2.6.3, we see that to prove Theorem 2.6.1 (b), we can proceed by:

1. extend T^* to a maximal theory T^{**} ;²
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 2.5.3, 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 2.6.1 (b), we follow the plan mentioned last lecture, and prove the following. 2 Feb. 11:30

Theorem 2.6.4. 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.

²Which 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 $[d_i] = t_i^{\mathcal{M}}([c_1], \dots, [c_n])$,

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

(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 2.6.4.

7 Feb. 11:30

Proof of Theorem 2.6.4 (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 2.6.3:** There is a **consistent** $T^* \supseteq T$ and \mathcal{L}^* -**theory** (with $\mathcal{L}^* \supseteq \mathcal{L}$) and T^* has **Henkin constants**.
2. **Theorem 2.5.3:** There is a **maximal consistent** \mathcal{L}^* -**theory** $T^{**} \supseteq T^*$, where T^{**} still has **Henkin constants**.
3. **Theorem 2.6.4:** 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.

2.6.3 Consequences of Completeness Theorem

Size of Models

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 2.6.5. 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 .

Single Stroke and Double Stroke Style Deduction

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 $\Gamma \models_{\mathcal{L}_0} \varphi$. *

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**.

Computational Properties

Consider the following.

Definition 2.6.3 (Computably enumerable). A set is *computably enumerable* or *computable listable* if there is a program that lists out its elements.

If \mathcal{L} is finite, or computable (complete list of symbols and their arities), we have the following.

- (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**.³
- (e) There is no program that given φ can decide whether $\vdash \varphi$ at least for $\mathcal{L} = \{E\}$, E binary.

³We can list out all the valid **proofs** from Γ of any φ .

Chapter 3

The Beginning of Model Theory

We now discuss various properties of the [models](#) of some [theories](#) in our interest. In particular, we care about the size of the [models](#), and how different [models](#) with different size relate to each other.

3.1 Complete Theories

Let's start with a proposition.

Proposition 3.1.1. 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 2.6.5](#). 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

3.2 A Detour to Algebraically Closed Fields

9 Feb. 11:30

[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.

3.2.1 Rings

All rings R we refer to will be commutative.

Definition 3.2.1 (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 3.2.2 (Proper). An *ideal* is *proper* if $1 \notin I$, equivalently, $I \neq R$.

Definition. Let I be a *proper ideal*.

Definition 3.2.3 (Radical). I is *radical* if $a^n \in I$, then $a \in I$.

Definition 3.2.4 (Prime). I is *prime* if $ab \in I$, then $a \in I$ or $b \in I$.

Definition 3.2.5 (Maximal). I is *maximal* if there is no *proper ideal* $J \supsetneq I$.

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

Definition 3.2.6 (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 3.2.7 (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 3.2.8 (Principal ideal). An *ideal* is a *principal ideal* if it’s *generated* by a single element.

Definition 3.2.9 (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 3.2.10 (Integral domain). A nontrivial ring with no **zero divisors** is called an *integral domain*.^a

^aSome authors will just call *domain*.

Definition 3.2.11 (Principal ideal domain). An **integral domain** where all **ideals** are **principal** is called a *principal ideal domain* or *PID*.

Theorem 3.2.1. $K[x]$ 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 3.2.12 (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 3.2.2 (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 3.2.3. 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 3.2.4. 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 .

3.2.2 Field Extensions

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

Definition 3.2.13 (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 3.2.14 (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 3.2.5. If M / L and L / K are [field extensions](#), then $[M: K] = [M: L][L: K]$.

3.2.3 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 3.2.15 (Algebraic). If there is a non-zero $f(x) \in K[x]$ such that $f(a) = 0$, then a is *algebraic* over K .

Definition 3.2.16 (Transcendental). If a is not [algebraic](#), then it is *transcendental* over K .

Definition 3.2.17 (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 3.2.6. 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 $f(a) = a^n + r_{n-1}a^{n-1} + \dots + r_1a + r_01 = 0$. ■

The following example illustrates how can we combine **Theorem 3.2.5** and **Theorem 3.2.6**,

Example. Let $f(x) = x^2 - 2$, $\mathbb{Q}(\sqrt{2}) = \{a1 + b\sqrt{2} \mid a, b \in \mathbb{Q}\}$.



Theorem 3.2.7. 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 3.2.18 (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 3.2.8. 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 3.2.19 (Algebraically closed). A field L is **algebraically closed** if any non-constant $f(x) \in L[x]$ has a root in L .

Definition 3.2.20 (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 3.2.9. 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

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

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

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

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 3.2.1. 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 3.2.2. Let L/K algebraic, M/K algebraically closed. Then there is an embedding $\alpha: L \rightarrow M$ fixing K .

Proof. Consider the case that $L = K(a)$ ^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 3.2.7. 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 3.2.1, $M = \alpha(L)$, so α is an isomorphism $L \rightarrow M$ over K . ■

Lecture 12: The ACF Theory and Categorical

Definition 3.2.21 (Characteristic). A field F has finite *characteristic* $p > 0$ if $\underbrace{1 + \cdots + 1}_{p \text{ times}} = 0$.

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

The following notion comes up naturally.

Definition 3.2.22 (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 3.2.23 (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. *

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 3.2.24 (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 3.2.10. 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 3.2.9](#), so $K \cong L$ via an isomorphism extending \bar{f} . ■

^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 3.2.1. 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 3.2.11. 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 .

3.3 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 3.3.1 (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 3.3.2 (Categorical). Let κ be an infinite cardinal and T be an [L-theory](#). We say T is κ -categorical if for any $\mathcal{M}, \mathcal{N} \models T$ of size κ , we have $\mathcal{M} \cong \mathcal{N}$.

Definition 3.3.3 (Countably categorical). If κ is countable, then T is *countably categorical*.

Definition 3.3.4 (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 3.3.1. If K has [transcendence degree](#) λ , then $|K| = \lambda + \aleph_0$.

Proof. Let K be [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 3.3.1. For each p , ACF_p is κ -categorical for every uncountable κ .

Proof. Let L, K be a [model](#) of ACF_p for size κ . From [Theorem 3.2.10](#), if L, K have [transcendence degree](#) κ , then they are isomorphic. With the application of [Lemma 3.3.1](#), we're done. ■

Example. \mathbb{Q}^{alg} , the [algebraic closure](#) of \mathbb{Q} , has size \aleph_0 with [transcendence degree](#) 0.

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

The above implies $\mathbb{Q}^{\text{alg}} \not\cong \mathbb{Q}(t)^{\text{alg}}$, which lead to the following.

Remark. ACF_0 is not [countably categorical](#). The same with ACF_p for $p > 0$.

Proof. Notice that

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

which is countable. With $\mathbb{Q}^{\text{alg}} \not\cong \mathbb{Q}(t)^{\text{alg}}$, we have the result. \circledast

Note. ACF is not uncountably categorical.

Theorem 3.3.2 (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 is not complete, then we can pick φ with $T \not\models \varphi$ and $T \not\models \neg\varphi$, i.e., $T \cup \{\varphi\}$ 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 \blacksquare

Corollary 3.3.1. ACF_p is complete for each p .

Proof. Follows immediately by Theorem 3.3.1 and Vaught's test. \blacksquare

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

Remark. $\{\varphi \mid \text{ACF} \models \varphi\}$ and $\{\varphi \mid \text{ACF}_p \models \varphi\}$ can be listed computably.

Proof. Since the axioms for ACF or ACF_p can be listed computably. \circledast

Definition 3.3.5 (Decidable). A theory T is decidable if there is a program that given φ , it determines whether $T \models \varphi$ or $T \not\models \varphi$.

Corollary 3.3.2. ACF_p is decidable for each p .

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 $\text{ACF}_p \models \varphi$. \blacksquare

Corollary 3.3.3. ACF is decidable.

Proof. Given φ , the algorithm simultaneously

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

If $\text{ACF} \models \varphi$, then we will halt at the first case. Now suppose $\text{ACF} \not\models \varphi$, i.e., there is $\mathcal{M} \models \text{ACF}$ such that $\mathcal{M} \models \neg\varphi$, and also, 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, hence the whole search will eventually halt. \blacksquare

^aIt might not be true that $\text{ACF} \models \neg\varphi$, we don't know.

Lecture 13: Upward Löwenheim-Skolem Theorem

Another consequence of completeness is that since $\mathbb{C} \models \text{ACF}_0$, if K is any algebraically closed field of characteristic 0, $\mathbb{C} \equiv K$, i.e., we have the following. 16 Feb. 11:30

Remark. The sentences true of \mathbb{C} are exactly the same as the sentences true of any algebraically closed field.

Essentially, the idea is that if one proves an algebraic statement about the complex numbers by analytic techniques of complex analysis, then there will be a proof of the same algebraic statement using purely algebraic tools, which works in any algebraically closed field of characteristic 0. The compactness theorem also gives connections to fields of finite characteristic.

Theorem 3.3.3 (Leftschetz principle). Let \mathcal{L} be the language of rings. For an \mathcal{L} -sentence φ , the following are equivalent:

- (i) φ is true in \mathbb{C} ;
- (ii) φ is true in every algebraically closed field of characteristic 0;
- (iii) φ is true in some algebraically closed fields of characteristic 0;
- (iv) there is a number n such that φ is true in all algebraically closed fields of characteristic $p > n$;
- (v) for each number n , φ is true in all algebraically closed fields of characteristic $p > n$.

Proof. Let $K \models \text{ACF}_0$, then since it's complete, we know that $K \models \varphi \Leftrightarrow \text{ACF}_0 \models \varphi$, which proves the first three. Others are left as homework. ■

We can use the Leftschetz principle to prove the following.

Theorem 3.3.4 (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 $f(\bar{x}) = \bar{y}$. ■

3.4 Up and Down

After introducing categorical, and seeing that ACF as an example, we now study something else: when does a model exists w.r.t. some original model, and what size is it?

3.4.1 Diagrams

One way to capture the structure in terms of theory is using the so-called “diagrams”.

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

Definition 3.4.1 (Atomic diagram). The atomic diagram of \mathcal{M} is the $\mathcal{L}_{\mathcal{M}}$ -theory

$$\text{Diag}(\mathcal{M}) := \{\varphi(\underline{a}_1, \dots, \underline{a}_n) \mid \mathcal{M} \models \varphi(m_1, \dots, m_n) \text{ and } \varphi \text{ is atomic or negated of atomic}\}.$$

Definition 3.4.2 (Elementary diagram). The elementary diagram of \mathcal{M} is the $\mathcal{L}_{\mathcal{M}}$ -theory

$$\text{Diag}_{\text{el}}(\mathcal{M}) := \{\varphi(\underline{a}_1, \dots, \underline{a}_n) \mid \mathcal{M} \models \varphi(m_1, \dots, m_n) \text{ and } \varphi \text{ an } \mathcal{L}\text{-formula}\}.$$

Intuition. Basically both $\text{Diag}(\mathcal{M})$ and $\text{Diag}_{\text{el}}(\mathcal{M})$ contain the information about the **structure** but in the form of a **theory**.

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

Lemma 3.4.1. Let \mathcal{N} be an $\mathcal{L}_{\mathcal{M}}$ -**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**. ■

3.4.2 Upward Löwenheim-Skolem theorem

Theorem 3.4.1 (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 $\mathcal{M} \models \text{Diag}_{\text{el}}(\mathcal{M})$, by **Proposition 3.1.1** it has a **model** \mathcal{N} of cardinality $\kappa \geq |\mathcal{L}_{\mathcal{M}}| = |\mathcal{M}| + |\mathcal{L}|$, so an **elementary embedding** exists $\mathcal{M} \rightarrow \mathcal{N}$ by **Lemma 3.4.1**. ■

Intuition. The **upward Löwenheim-Skolem theorem** says that every **structure** is an **elementary sub-structure** of many much bigger **structures**.

As previously seen. Our very first application of the **compactness theorem**: in **the construction of the non-standard model of arithmetic**, we built $\mathcal{N} \models \text{Th}(\mathbb{N})$ not **isomorphic** to \mathbb{N} , which is exactly like this. Every element of \mathbb{N} can already be expressed as a **term** of the form $1 + \dots + 1$ without adding any new constants.

3.4.3 Downward Löwenheim-Skolem theorem

A “downward” version of the **upward Löwenheim-Skolem theorem** also exists, which says that big **models** contain smaller **elementary substructures**. This will take some more work to prove, so we first need a test for this.

Proposition 3.4.1 (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 since $\mathcal{M} \models \exists x \varphi(x, \bar{a})$ if and only if $\mathcal{N} \models \exists x \varphi(x, \bar{a})$.

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 . Since \mathcal{M} is a **substructure** of \mathcal{N} , this is true for all **quantifier-free formulas**, and in particular for the **atomic formulas**.

Suppose that the claim is true for ψ , then

$$\mathcal{M} \models \neg\psi(\bar{a}) \Leftrightarrow \mathcal{M} \not\models \psi(\bar{a}) \Leftrightarrow \mathcal{N} \not\models \psi(\bar{a}) \Leftrightarrow \mathcal{N} \models \neg\psi(\bar{a}),$$

so the claim is also true for $\neg\psi$. Similarly, 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

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

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

We will also need to introduce the [Skolemizations](#) of [theories](#).

21 Feb. 11:30

Definition 3.4.3 (Built-in Skolem function). 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

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

Intuition. This is like a parametrized version of [Henkin constants](#).

Lemma 3.4.2. 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 $\mathcal{M}^* \models T^*$. Finally, we can choose \mathcal{L}^* such that $|\mathcal{L}^*| = |\mathcal{L}| + \aleph_0$.

Proof. Start with $\mathcal{L}_0 = \mathcal{L}$ and $T_0 = T$, we build $\mathcal{L}_0 \subseteq \mathcal{L}_1 \subseteq \dots$ and $T_0 \subseteq T_1 \subseteq \dots$ and let $\mathcal{L}^* = \bigcup_i \mathcal{L}_i$ and $T^* = \bigcup_i T_i$. Given \mathcal{L}_i and T_i , define

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

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

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

We now argue that if $\mathcal{M}_i \models T_i$, we can [expand](#) it to a [model](#) \mathcal{M}_{i+1} of T_{i+1} . Pick $c \in M_i$ a “default value.” Given φ and \bar{a} , define $f_\varphi^{\mathcal{M}_{i+1}}(\bar{a})$ to be some b with $\mathcal{M}_i \models \varphi(b, \bar{a})$ if such a b exists, or c otherwise.^a Then, $\mathcal{M}_{i+1} \models T_{i+1}$. From this construction, we see that T^* has [built-in Skolem functions](#) since any \mathcal{L}^* -formula φ is in some \mathcal{L}_i and has a [Skolem function](#) in \mathcal{L}_{i+1} .

Now, suppose $\mathcal{M} \models T$, i.e., $\mathcal{M} = \mathcal{M}_0 \models T_0$. From above, \mathcal{M}_0 has an [expansion](#) $\mathcal{M}_1 \models T_1$, which has an [expansion](#) $\mathcal{M}_2 \models T_2$, etc. By [expanding](#) \mathcal{M} iteratively, we get a [model](#) \mathcal{M}^* of T^* .

Finally, at each step, we add one symbol for each \mathcal{L}_i -formula to \mathcal{L}_i , hence by counting, $|\mathcal{L}_{i+1}| = |\mathcal{L}_i| + \aleph_0$, hence $|\mathcal{L}^*| = |\mathcal{L}| + \aleph_0$. ■

^aIf such b doesn't exist, then the left-hand side is false, so we don't really care about the right-hand side.

Note. We see that this is a similar argument to when we added [Henkin constants](#), though it's simpler now because we can work semantically.

Notation (Skolemization). We call T^* in [Lemma 3.4.2](#) a *Skolemization* of T .

Theorem 3.4.2 (Downward Löwenheim-Skolem theorem). Let \mathcal{M} be an \mathcal{L} -structure and $X \subseteq M$. Then there is an [elementary substructure](#) \mathcal{N} of \mathcal{M} with $X \subseteq \mathcal{N}$ and $|N| \leq |X| + |\mathcal{L}| + \aleph_0$.

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 3.4.2). Hence, by 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

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

Let $N = \bigcup_i X_i$, and let \mathcal{N} be the substructure of \mathcal{M} with domain N . This can be done by:

- for each function symbol f , let $f^{\mathcal{N}}$ be the restriction of $f^{\mathcal{M}}$ to N ;^a
- for each relation symbol R , let $R^{\mathcal{N}}$ be the restriction of $R^{\mathcal{M}}$ to N ;
- for each constant symbol c , there is a Skolem function f such that $f(x) = c^{\mathcal{M}} \in N$ for all $x \in M$, so let $c^{\mathcal{N}} = c^{\mathcal{M}}$.

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 $\mathcal{M} \models \varphi(b, \bar{a})$, and 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, since $|X_0| \leq |X| + |\mathcal{L}| + \aleph_0$, with N being a countable union, $|N| \leq |X| + |\mathcal{L}| + \aleph_0$. ■

^aFrom the definition of N , it is closed under the applications of functions f , so the restriction takes values in N .

Notation (Generated substructure). \mathcal{N} in the downward Löwenheim-Skolem theorem 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 $X \subseteq \mathcal{R} \preceq \mathbb{R}$ such that \mathcal{R} is countable. In particular, $\text{Th}(\mathcal{R}) = \text{Th}(\mathbb{R})$ and \mathcal{R} is a countable real closed field, i.e.,

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

There is a whole theory of real closed fields just like for algebraically closed fields.

Intuition. Countable real closed fields are as algebraically closed as they can be while still being orderable.

Example (Skolem's paradox). Let $\mathcal{L} = \{\in\}$ be the language of set theory, where \in a binary relation symbol. Let $T = \text{ZFC}$, and suppose ZFC is satisfiable,^a i.e., there is a model \mathcal{M} such that $\mathcal{M} \models T$.

Then, there is a countable $\mathcal{N} \preceq \mathcal{M}$, in particular, there is a countable model of ZFC. We can then write down the sentence

$$\varphi := \text{“there is no bijection between } \mathbb{R}^{\mathcal{N}} \text{ and } \mathbb{N}^{\mathcal{N}}\text{”}$$

such that $\mathcal{N} \models \varphi$, in \mathcal{L} . Observe that \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.

We finish this section with two useful facts.

Definition 3.4.4 (Universally axiomatizable). Let T be an \mathcal{L} -theory, then T is *universally axiomatizable*.

zable if there is a set Γ of **universal sentences** such that $T \models \Gamma$ and $\Gamma \models T$.^a

^aI.e., $\mathcal{M} \models T$ if and only if $\mathcal{M} \models \Gamma$.

Theorem 3.4.3. 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. The forward direction is easy: suppose that T is **universally axiomatized** by Γ . If $\mathcal{M} \subseteq \mathcal{N}$ and $\mathcal{N} \models T$, then $\mathcal{N} \models \Gamma$, and since Γ consists only of **universal formulas**, $\mathcal{M} \models \Gamma$, so $\mathcal{M} \models T$.

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 need to show that $\Gamma \models T$. We may assume that T is **satisfiable**^a and let $\mathcal{M} \models \Gamma$, so we now want to prove that $\mathcal{M} \models T$. This can be done by finding $\mathcal{N} \supseteq \mathcal{M}$ and $\mathcal{N} \models T$, then from our assumption we have $\mathcal{M} \models T$. We build such an \mathcal{N} by showing that $\text{Diag}(\mathcal{M}) \cup T$ is **satisfiable**, and take the corresponding **model** to be \mathcal{N} . Then, trivially we have $\mathcal{N} \models T$.

This can be done by **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

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

assume that they are actually equal. To show that Δ is **satisfiable**, it is enough to show^b 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 since T is **satisfiable**,

$$T \models \forall x_1 \dots \forall x_k \neg(\varphi_1(\bar{x}) \wedge \dots \wedge \varphi_\ell(\bar{x})).$$

Since this is **universal**, hence in Γ , so it is **true** in \mathcal{M} . But it's also not **true** in \mathcal{M} since $\mathcal{M} \models \varphi_1(\bar{m}) \wedge \dots \wedge \varphi_\ell(\bar{m})$, a contradiction \nmid Hence, Δ is **satisfiable**, so any finite subset is **satisfiable**, by **compactness theorem**, we're done. ■

^aSince otherwise $\Gamma \ni \forall x x \neq x$, and hence $\Gamma \models T$ trivially.

^bSince the constant symbols m_1, \dots, m_k do not appear in T , we can interpret them as the witness to the \exists 's.

Lecture 15: The Random Graph Theory

Proposition 3.4.2. 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**, for all i and $\bar{a} \in M_i$, we show that $\mathcal{M}_i \models \varphi(\bar{a})$ if and only if $\mathcal{M} \models \varphi(\bar{a})$.

- (a) For φ is **atomic**, this is true since φ is **quantifier-free** and $\mathcal{M}_i \subseteq \mathcal{M}$.
- (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 $\mathcal{M} \models \psi(\bar{a}, b)$. Since $M = \bigcup_j \mathcal{M}_j$, there is $j \geq i$ such that $b \in M_j$. By the induction hypothesis, $\mathcal{M}_j \models \psi(\bar{a}, b)$, so $\mathcal{M}_j \models \exists y \psi(\bar{a}, y)$. Finally, since $\mathcal{M}_i \preceq \mathcal{M}_j$, so $\mathcal{M}_i \models \exists y \psi(\bar{a}, y)$.

23 Feb. 11:30

3.5 Back and Forth

We have examples of [uncountably categorical theories](#), but no examples of [countably categorical theories](#).

3.5.1 Dense Linear Order Theory

The simplest example of a [countably categorical theory](#) is the [theory](#) of “linear orders (without endpoints),” denoted as [DLO](#).

Definition 3.5.1 (DLO). Let $\mathcal{L} = \{\leq\}$. The [theory](#) of *dense linear orders (without endpoints)*, denoted as DLO, has the axioms:

- (a) \leq is a linear order;
- (b) $\forall x \forall y (x < y \rightarrow \exists z x < z < y)$ (the density axiom);
- (c) $\forall x \exists y \exists z (y < x < z)$ (the no-endpoints axiom).

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

To create a new [dense linear orders](#), given $\mathcal{M}_1, \mathcal{M}_2$ two DLO’s, define $\mathcal{M}_1 + \mathcal{M}_2$ with domain $M \sqcup N$ and has each element of M less than each element of N , and within M and N , the orderings are the same as in \mathcal{M} and \mathcal{N} . This is also a DLO.

Example. $\mathbb{Q} + \mathbb{Q}$ and $\mathbb{R} + \mathbb{R}$ are both DLO’s.

Example. $\mathbb{R} + \mathbb{R} \not\cong \mathbb{R}$.

Proof. Since \mathbb{R} has the least upper bound property while $\mathbb{R} + \mathbb{R}$ does not (there is no least upper bound for the first copy). *

Example. $\mathbb{Q} + \mathbb{Q} \cong \mathbb{Q}$.

Proof. For example, take some irrational, e.g., π . Then $\mathbb{Q} = \{x \mid x < \pi\} \cup \{x \mid x > \pi\}$, and we observe that we have

$$\{x \mid x < \pi\} \cong \mathbb{Q} \cong \{x \mid x > \pi\},$$

and hence piecing them together we have $\mathbb{Q} + \mathbb{Q} \cong \mathbb{Q}$. *

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 $\kappa \geq 2^{\aleph_0}$.

We now show that DLO is actually [countably categorical](#).

Theorem 3.5.1. The [theory](#) DLO is [countably categorical](#) and hence [complete](#).

Proof. Let (A, \leq) and (B, \leq) be two countable DLO’s, and let a_1, a_2, \dots and b_1, b_2, \dots be a listing of A and B , respectively. We build an [isomorphism](#) $f: A \rightarrow B$ stage-by-stage: at stage i , we have

- finite sets $A_i \subseteq A$ and $B_i \subseteq B$, and
- a bijection $f_i: A_i \rightarrow B_i$ called a *partial embedding*:^a if $a < a' \in A_i$, then $f_i(a) < f_i(a')$.

In this way, $f_i \subseteq f_{i+1}$, $A_i \subseteq A_{i+1}$, and $B_i \subseteq B_{i+1}$, and we need to make sure that

- $\bigcup_i A_i = A$, i.e., each element of A is in the domain of f (ensured by odd stages);
- $\bigcup_i B_i = B$, i.e., each element of B is in the range of f (ensured by even stages),

so $f = \bigcup_i f_i$ is a bijection from $A \rightarrow B$. Then since each f_i is a partial **embedding**, so f is an **\mathcal{L} -embedding**, hence an **isomorphism**. This will prove that DLO is **countably categorical**.

The construction works as follows.

- Stage 0: $A_0 = \emptyset$, $B_0 = \emptyset$, $f_0 = \emptyset$.
- Stage $i + 1 = 2k + 1$: the goal is to make sure $a_k \in A_{i+1} = \text{dom}(f_{i+1})$:
 - if $a_k \in A_i$ already, then do nothing, i.e., $A_{i+1} = A_i$, $B_{i+1} = B_i$, $f_{i+1} = f_i$;
 - otherwise, $a_k \notin A_i$, define $f_{i+1} \supseteq f_i$ by adding a_k to $A_{i+1} = A_i \cup \{a_k\}$, and for elements $a \in A_i$, $f_{i+1}(a) = f_i(a)$. Now, we have three possibilities:
 - * a_k is less than all of A_i : choose $b \in B$ less than all of B_i ;^b
 - * a_k is greater than all of A_i : similar to above;
 - * there are a and a' in A_i such that $a < a_k < a'$ with no other elements of A_i between a and a' since A_i is finite: pick b with $f_i(a) < b < f_i(a')$.^c

In all cases, we can choose b and let $B_{i+1} = B_i \cup \{b\}$ with $f_{i+1}(a_k) = b$.
- Stage $i + 1 = 2k + 2$: the goal is to make sure $b_k \in B_{i+1} = \text{Im}(f_{i+1})$: this is exactly the same, but in the other direction (e.g., working with f_i^{-1} rather than f_i).

Now everything is checked, so DLO is **countably categorical** (hence **complete** by **Vaught test**). ■

^aIf f_i maps \bar{a} to \bar{b} , then \bar{a} satisfies the same **atomic** and negated **atomic formulas** in (A, \leq) that \bar{b} does in (B, \leq) .

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

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

Note (Back-and-forth). We see that the above is the so-called *back-and-forth* argument.

Corollary 3.5.1. $\mathbb{Q} + \mathbb{R} \equiv \mathbb{R}$.

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

Definition 3.5.2 (Complete). A linear order is *complete* if every subset bounded above has a least upper bound.

Corollary 3.5.2. There is no first order **sentence** φ such that $\mathcal{M} \models \varphi$ if and only if \mathcal{M} is a **complete** linear order.

3.5.2 Random Graph Theory

Another example of a **countably categorical theory** is the **theory of random graph**.

Definition 3.5.3 (Random graph). A *random graph* we will consider is constructed as follows. Firstly, fix countably infinitely many vertices v_1, v_2, \dots , and 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? It turns out to be interesting enough, so we will look into it:

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

Then, consider the following **theory**.

Definition 3.5.4 (Random graph theory). Let $\mathcal{L} = \{E\}$, where E is a binary relation. The *random graph theory* T has axioms:

(a) $\forall x \neg xEx$ and $\forall x \forall y (xEy \rightarrow yEx)$ (irreflexive, undirected);

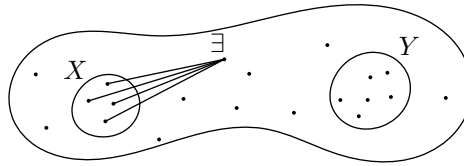
(b) $\exists x \exists y x \neq y$;

(c) for each n , define ψ_n 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 (Extension axiom). Think of ψ_n as an *extension axiom*: the property that for any finite disjoint sets X and Y , there is a vertex with an edge to each $x \in X$ and no edge to each $y \in Y$. This axiom happens with probability $p^{|X|} \cdot (1-p)^{|Y|}$ for a given z for some X, Y .

We see that a **model** of T is a graph with at least two elements with the **extension property**.



Note. In ψ_n 's, we allow there to be repetitions among the x 's and the y 's. In particular, if $m \leq n$, then $\psi_n \models \psi_m$.

Lecture 16: Quantifier Elimination

Remark. T is actually **countably categorical**.

7 Mar. 11:30

Proof. We will show this on homework. But this is basically **Theorem 3.5.1**. ⊛

Now we show that it is also **satisfiable**. Fix $N, p \in (0, 1)$, and vertices $\{1, \dots, N\}$. Generate a **random graph** G_N on vertices $\{1, \dots, N\}$ by, for each $i, j \leq N, i \neq j$, putting an edge between i and j with probability p and no edge with probability $1 - p$. Let \mathcal{G}_N be the class of all such graphs, and let $p_N(\varphi)$ be the probability that a **randomly generated graph** on N vertices satisfies φ .

Example. If $p = 1/2$, all graphs in \mathcal{G}_N are equally likely.

Observe the following.

Proposition 3.5.1. For every n , then $\lim_{N \rightarrow \infty} p_N(\psi_n) = 1$.

Proof. Fix n , and let G be a **random graph** in $\mathcal{G}_N, N > 2n$. Fix x_1, \dots, x_n , and y_1, \dots, y_n , and z in G , all distinct. Consider the probability q that

$$\neg \left(\bigwedge_{i=1}^n E(x_i, z) \wedge \neg E(y_i, z) \right),$$

which is just $1 - p^n(1-p)^n := q$ as we already know. The probability that

$$G \models \underbrace{\neg \exists z}_{\forall z \neg} \left(\bigwedge_{i=1}^n E(x_i, z) \wedge \neg E(y_i, z) \right)$$

is q^{N-2n} , where $N - 2n$ is the number of possible z 's.^a Let M be the number of different choices

of x_i 's and y_i 's, then we have

$$p_N(\neg\psi_n) \leq M \cdot q^{N-2n} \leq N^{2n} q^{N-2n}$$

by the union bound.^b Since $0 < q < 1$, so $p_N(\neg\psi_n) \rightarrow 0$, i.e., $p_N(\psi_n) \rightarrow 1$ as $N \rightarrow \infty$. ■

^aNotice that we already assume that $x_i \neq z \wedge y_i \neq z$.

^bThis bound is not allowing the x 's and y 's to have repetitions. But this doesn't change anything.

Then, we have the following.

Theorem 3.5.2. T is **satisfiable**.

Proof. From Proposition 3.5.1, $\lim_{N \rightarrow \infty} p_N(\psi_n) = 1$. In particular, for each n , there is N such that $p_N(\psi_n) > 0$, i.e., there is at least one $G \in \mathcal{G}_N$ such that $G \models \psi_n$, hence $G \models \psi_m$ for $m \leq n$.^a This means that for every finite $T^* \subseteq T$ is **satisfiable**, so T is **satisfiable** by the **compactness theorem**. ■

^aRemember that for $n > m$, $\models \psi_n \rightarrow \psi_m$.

Moreover, we have the following.

Theorem 3.5.3 (Zero-one law for graphs). For any \mathcal{L} -sentence φ , we either have $\lim_{N \rightarrow \infty} p_N(\varphi) = 0$ or $\lim_{N \rightarrow \infty} p_N(\varphi) = 1$. Moreover, T axiomatizes $\{\varphi : \lim_{N \rightarrow \infty} p_N(\varphi) = 1\}$, the “almost sure theory for graphs”, which is **decidable** and **complete**.

Proof. Since T is **countably categorical**, and has only infinite models, so it is **complete** by the **Vaught's test**. If $T \models \varphi$, there is some n such that $\{\psi_n, (a), (b)\} \models \varphi$. Since $p_N(\psi_n) \leq p_N(\varphi)$, so $\lim_{N \rightarrow \infty} p_N(\varphi) = 1$ and the left-hand side goes to 1 as well. On the other hand, if $T \not\models \varphi$, $T \models \neg\varphi$, so $\lim_{N \rightarrow \infty} p_N(\neg\varphi) = 1$, i.e., $\lim_{N \rightarrow \infty} p_N(\varphi) = 0$ from the same argument. ■

Before ending this section, we note that there are other things we may explore.

Remark. There are also Fraïssé constructions, and Ryll-Nardzewski theorem, etc.

Chapter 4

Quantifier Elimination and Algebraic Applications

4.1 Quantifier Elimination

Let's start with a definition.

Definition 4.1.1 (Quantifier elimination). A theory T has (admits) *quantifier elimination* if for every formula $\varphi(\bar{x})$ (with \bar{x} containing at least one variable), there is a quantifier-free $\psi(\bar{x})$ such that

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

The easiest example will be that DLO admits quantifier elimination. To prove this, we need some preliminaries.

Note. DLO has no constant symbols, so it has no quantifier-free sentences. For example,

$$\forall x \exists y y < x \leftrightarrow x = x,$$

where the right-hand side has a free variable. Another solution would be to allow \top and \perp for the true and false sentences.

Lemma 4.1.1. Let (A, \leq) and (B, \leq) be countable DLOs. Let $a_1, \dots, a_n \in A$ have $a_1 < a_2 < \dots < a_n$ and $b_1, \dots, b_n \in B$ have $b_1 < \dots < b_n$. Then there is an isomorphism $f: A \rightarrow B$ which maps $a_i \mapsto b_i$. Hence,

$$A \models \varphi(\bar{a}) \Leftrightarrow B \models \varphi(\bar{b}).$$

Proof. This is the same as the back-and-forth argument, but now we start with $\bar{a} \mapsto \bar{b}$. ■

Theorem 4.1.1. DLO admits quantifier elimination.

Proof. Fix $\varphi(\bar{x})$. If φ is actually a sentence then we're done since either

- DLO $\models \varphi$, so DLO $\models \forall x (\varphi \leftrightarrow x = x)$, or
- DLO $\models \neg \varphi$, so DLO $\models \forall x (\varphi \leftrightarrow x \neq x)$.

Now suppose $\varphi(\bar{x})$ has at least one free variable, $\bar{x} = (x_1, \dots, x_n)$. Since DLO is complete, it's enough to find $\psi(\bar{x})$ with $\mathbb{Q} \models \forall \bar{x} (\varphi(\bar{x}) \leftrightarrow \psi(\bar{x}))$. Let σ be a map from pairs i, j to $\{1, 2, 3\}$. Define^a

$$\theta_\sigma(x_1, \dots, x_n) := \bigwedge_{\sigma(i,j)=1} x_i < x_j \wedge \bigwedge_{\sigma(i,j)=2} x_i = x_j \wedge \bigwedge_{\sigma(i,j)=3} x_j < x_i.$$

If $\mathbb{Q} \models \theta_\sigma(\bar{a}) \wedge \theta_\sigma(\bar{b})$, then \bar{a} and \bar{b} satisfy the same formulas by Lemma 4.1.1, so^b

$$\Sigma := \{\sigma \mid \mathbb{Q} \models \exists \bar{x} (\theta_\sigma(\bar{x}) \wedge \varphi(\bar{x}))\} = \{\sigma \mid \mathbb{Q} \models \forall \bar{x} (\theta_\sigma(\bar{x}) \rightarrow \varphi(\bar{x}))\}.$$

If $\Sigma = \emptyset$, then $\varphi(x) \leftrightarrow x \neq x$; if $\Sigma \neq \emptyset$, let $\psi(\bar{x}) = \bigvee_{\sigma \in \Sigma} \theta_\sigma(\bar{x})$, then $\mathbb{Q} \models \forall \bar{x} (\varphi(\bar{x}) \leftrightarrow \psi(\bar{x}))$. ■

^aIt's clear that some θ_σ might be inconsistent.

^bThe second equality follows from the fact that if one such \bar{x} exists, all such \bar{x} work.

Lecture 17: Quantifier Elimination

Example. $\varphi(x, y) := \exists u (u > x \wedge u < y)$. This is equivalent, in DLO, to $\psi(x, y) := x < y$.

9 Mar. 11:30

Example. In Problem Set 2, we have looked at $\mathcal{L} = \emptyset$. There, we showed that if $A \subseteq B$ and both infinite, then $A \preceq B$. The same idea show that $T = \{\exists x_1 \dots \exists x_n \bigwedge_{i \neq j} x_i \neq x_j \mid n \in \mathbb{N}\}$ admits quantifier elimination.

Proposition 4.1.1. Let T be a theory that admits quantifier elimination. If $\mathcal{A}, \mathcal{B} \models T$ and $\mathcal{A} \subseteq \mathcal{B}$, then $\mathcal{A} \preceq \mathcal{B}$.

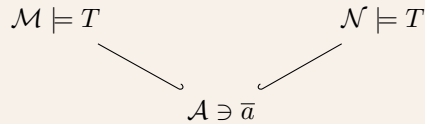
Proof. Given $\phi(\bar{x})$ and $\bar{a} \in A$, then $\phi(x)$ is equivalent to a quantifier-free $\psi(\bar{x})$ (modulo T). Then,

$$\mathcal{A} \models \phi(\bar{a}) \Leftrightarrow \mathcal{A} \models \psi(\bar{a}) \Leftrightarrow \mathcal{B} \models \psi(\bar{a}) \Leftrightarrow \mathcal{B} \models \phi(\bar{a}).$$

Next, we want to show that ACF admits quantifier elimination. First, we need a test for quantifier elimination.

Theorem 4.1.2. Let \mathcal{L} include at least one constant symbol c . Let T be an \mathcal{L} -theory, and $\phi(\bar{x})$ an \mathcal{L} -formula. Then the following are equivalent.

- (a) There is a quantifier-free $\psi(\bar{x})$ such that $T \models \forall \bar{x} (\phi(\bar{x}) \leftrightarrow \psi(\bar{x}))$.
- (b) If $\mathcal{M}, \mathcal{N} \models T$, and \mathcal{A} is a common substructure, then for all $\bar{a} \in A$, $\mathcal{M} \models \phi(\bar{a})$ if and only if $\mathcal{N} \models \phi(\bar{a})$.



Proof. (a) implies (b) is easy, since we have

$$\mathcal{M} \models \phi(\bar{a}) \Leftrightarrow \mathcal{M} \models \psi(\bar{a}) \Leftrightarrow \mathcal{A} \models \psi(\bar{a}) \Leftrightarrow \mathcal{N} \models \psi(\bar{a}) \Leftrightarrow \mathcal{N} \models \phi(\bar{a}).$$

To show (b) implies (a), we see that first, there are two easy cases:

- $T \models \forall \bar{x} \phi(\bar{x})$: take ψ to be $c = c$;
- $T \models \forall \bar{x} \neg \phi(\bar{x})$: take ψ to be $c \neq c$.

Now, suppose we are not in the above two cases. Let $\Gamma(\bar{x})$ be the set of all quantifier-free formulas $\psi(\bar{x})$ such that $T \models \forall \bar{x} (\phi(\bar{x}) \rightarrow \psi(\bar{x}))$, and let $\bar{d} = (d_1, \dots, d_n)$ be new constant symbols.

Claim. It's enough to show $T \cup \Gamma(\bar{d}) \models \phi(\bar{d})$.

Proof. If we can do this, then by **compactness**, there are $\psi_1(\bar{x}), \dots, \psi_m(\bar{x})$ such that

$$T \models [\psi_1(\bar{d}) \wedge \dots \wedge \psi_m(\bar{d})] \rightarrow \phi(\bar{d})$$

By the choice of Γ , the \leftarrow direction also holds, hence

$$T \models [\psi_1(\bar{d}) \wedge \dots \wedge \psi_m(\bar{d})] \leftrightarrow \phi(\bar{d}),$$

i.e., $T \models \forall \bar{x} (\psi(\bar{x}) \leftrightarrow \phi(\bar{d}))$ where $\psi(x) := \psi_1(\bar{x}) \wedge \dots \wedge \psi_m(\bar{x})$. ⊗

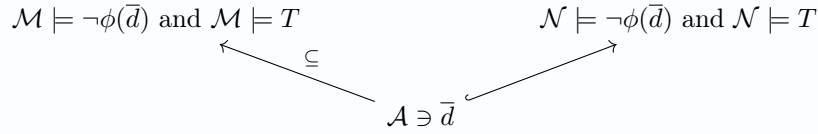
Claim. $T \cup \Gamma(\bar{d}) \models \phi(\bar{d})$.

Proof. Suppose not. Then $T \cup \Gamma(\bar{d}) \cup \{\neg\phi(\bar{d})\}$ is **satisfiable**.^a Let \mathcal{M} be a **model** of $T \cup \Gamma(\bar{d}) \cup \{\neg\phi(\bar{d})\}$, and let $\mathcal{A} \subseteq \mathcal{M}$ be the “submodel”^b generated by \bar{d} . Notice that since $c^{\mathcal{A}} \in \mathcal{A}$, so $\mathcal{A} \neq \emptyset$. Every element of \mathcal{A} is $t^{\mathcal{A}}(\bar{d}) = t^{\mathcal{M}}(\bar{d})$ for some **term** t . Also, $\mathcal{A} \models \Gamma(\bar{d})$ because $\mathcal{M} \models \Gamma(\bar{d})$ and $\Gamma(\bar{d})$ is **quantifier-free**. We now show that $T \cup \text{Diag}(\mathcal{A}) \cup \{\phi(\bar{d})\}$ is **satisfiable**, and take \mathcal{N} to be a **model**. If not, some finite subset is not **satisfiable**. There are **quantifier-free formulas** $\gamma_1(\bar{d}), \dots, \gamma_\ell(\bar{d})$ in $\text{Diag}(\mathcal{A})$ such that^c

$$T \models [\gamma_1(\bar{d}) \wedge \dots \wedge \gamma_\ell(\bar{d})] \rightarrow \neg\phi(\bar{d}).$$

Equivalently, $T \models \phi(\bar{d}) \rightarrow \neg[\gamma_1(\bar{d}) \wedge \dots \wedge \gamma_\ell(\bar{d})]$, so $T \models \forall \bar{x} (\phi(\bar{x}) \rightarrow [\neg\gamma_1(\bar{x}) \vee \dots \vee \neg\gamma_\ell(\bar{x})])$. This implies that $\neg\gamma_1(\bar{x}) \vee \dots \vee \neg\gamma_\ell(\bar{x})$ is in $\Gamma(\bar{x})$, hence $\mathcal{A} \models \neg\gamma_1(\bar{d}) \vee \dots \vee \neg\gamma_\ell(\bar{d})$. But $\mathcal{A} \models \gamma_1(\bar{d}) \wedge \dots \wedge \gamma_\ell(\bar{d})$, contradicts to the fact that $\gamma_1(\bar{d}), \dots, \gamma_\ell(\bar{d})$ is in $\text{Diag}(\mathcal{A})$.

Hence, there is $\mathcal{N} \models T \cup \text{Diag}(\mathcal{A}) \cup \{\phi(\bar{d})\}$. This contradicts to (b):



Hence, $T \cup \Gamma(\bar{d}) \models \phi(\bar{d})$. ⊗

^aNotice that $T \cup \Gamma(\bar{d})$ needs to be **satisfiable**, which is true from our assumption.

^b \mathcal{A} is the smallest **model** containing \bar{d} , and \bar{a} might not be a **model** of T .

^cWe can take $\gamma_i(\bar{d})$ to be just about \bar{d} but not $\gamma_i(\bar{a})$ for $\bar{a} \in \mathcal{A}$ is because each $\bar{a} \in \mathcal{A}$ is $t^{\mathcal{A}}(\bar{d})$, so we can replace \bar{a} by this **term**.

We now show that we can eliminate one \exists at a time. ■

Lemma 4.1.2. Let T be an **\mathcal{L} -theory**. Suppose that for every **quantifier-free formula** $\gamma(\bar{x}, y)$, there is a **quantifier-free formula** $\psi(\bar{x})$ such that

$$T \models \forall \bar{x} (\exists y \gamma(\bar{x}, y) \leftrightarrow \psi(\bar{x})),$$

then T has **quantifier elimination**.

Proof. By induction on **formulas** with the hypotheses of **Theorem 4.1.2** for the \exists quantifier case. ■

By putting together the last two results, i.e., **Lemma 4.1.2** and **Theorem 4.1.2**, we get the following test for **quantifier elimination**. Compare this, for example, to **Vaught's test** for **elementary-substructure**.

Corollary 4.1.1. Let T be an **\mathcal{L} -theory** with at least one constant. Suppose that for all **quantifier-free** $\psi(\bar{x}, y)$, if $\mathcal{M}, \mathcal{N} \models T$, and $\mathcal{A} \subseteq \mathcal{M}$ and $\mathcal{A} \subseteq \mathcal{N}$, $\bar{a} \in \mathcal{A}$ and $b \in M^{\mathcal{A}}$ is such that $\mathcal{M} \models \psi(\bar{a}, b)$, then

there is $c \in N$ such that $\mathcal{N} \models \psi(\bar{a}, c)$. Then T admits **quantifier elimination**.



^aThis is (b) for $\exists y \psi(\bar{x}, y)$.

4.1.1 Quantifier Elimination for ACF

Now, we start with our favorite ACF, and show that it admits **quantifier elimination**.

Theorem 4.1.3. ACF admits **quantifier elimination**.

Proof. We use **Corollary 4.1.1**. Suppose that $\mathcal{M}, \mathcal{N} \models \text{ACF}$, and $\mathcal{A} \subseteq \mathcal{M}, \mathcal{N}$ an **integral domain**. Let $\psi(\bar{x}, y)$ be a **quantifier-free formula** and $\bar{a} \in \mathcal{A}$. Suppose there is $b \in \mathcal{M}$ such that $\mathcal{M} \models \psi(\bar{a}, b)$.

Firstly, we replace \mathcal{N} by the **algebraic closure** of \mathcal{A} , then we can also assume that $\mathcal{N} \subseteq \mathcal{M}$ since any **algebraic closure** of \mathcal{A} embeds in any **algebraically closed** field containing \mathcal{A} .



Problem. What does ψ look like?

Answer. Since it's **quantifier-free**, we may assume that it is $\theta_1(\bar{x}, y) \vee \theta_2(\bar{x}, y) \vee \dots$, where each θ_i is a conjunction of **atomic** and negated **atomic formula**. We can replace ψ by whichever θ_i is satisfied by b . ⊗

Now, since in the **language** of rings, an **atomic formula** $\gamma(\bar{x}, y)$ is equivalent to a **formula** of the form $p(\bar{x}, y) = 0$, where $p \in \mathbb{Z}[\bar{X}, Y]$, so $\psi(\bar{x}, y)$ is equal to

$$p_1(\bar{x}, y) = 0 \wedge \dots \wedge p_k(\bar{x}, y) = 0 \wedge q_1(\bar{x}, y) \neq 0 \wedge q_\ell(\bar{x}, y) \neq 0$$

for $p_i, q_i \in \mathbb{Z}[\bar{X}, Y]$. Then $\psi(\bar{a}, y)$ says $p_1(\bar{a}, y) = 0 \wedge \dots$ where $p_i(\bar{a}, y)$ are now in $\mathcal{A}[y]$, i.e., polynomial in y with coefficient in \mathcal{A} . If any $p_i(\bar{a}, y)$ is non-trivial, then b is a solution, so $b \in \mathcal{N}$ since \mathcal{N} is **algebraically closed** (and we just take $c = b$ for applying **Corollary 4.1.1**). Otherwise, assume $p_i(\bar{a}, y)$ is trivial, so $\psi(\bar{a}, y)$ is just

$$q_i(\bar{a}, y) \neq 0 \wedge \dots \wedge q_\ell(\bar{a}, y) \neq 0.$$

Since b satisfies this, each $q_i(\bar{a}, y)$ is non-trivial, so $q_i(\bar{a}, y) = 0$ has only finitely many solutions. But \mathcal{N} is infinite, so there is a $c \in \mathcal{N}$ that is not a solution to any $q_i(\bar{a}, y) = 0$, so $\mathcal{N} \models \psi(\bar{a}, c)$. ■

Lecture 18: Quantifier Elimination for Algebraically Closed Fields

Problem 4.1.1. What do we get from **quantifier elimination**?

14 Mar. 11:30

Answer. Understand the **definable** sets, they are defined by **quantifier-free formulas**. ⊗

Remark. We see that

$$\varphi(\bar{x}) := \exists y \ p_n(\bar{x})y^n + \dots + p_1(\bar{x})y + p_0(\bar{x}) = 0$$

is equivalent to

$$\varphi(\bar{x}) := p_n(\bar{x}) \neq 0 \vee \dots \vee p_1(\bar{x}) \neq 0 \vee p_0(\bar{x}) = 0.$$

Definition 4.1.2 (Cofinite). A *cofinite* subset of a set X is a subset A such that $|A^c| < \infty$.

Proposition 4.1.2. The **definable**^a subsets of an **algebraically closed** field K are exactly the finite and **cofinite** sets.

^aUsing parameters.

Proof. Let $\{a_1, \dots, a_n\}$ be a finite set. This is **definable** by $x = a_1 \vee \dots \vee x = a_n$ using a_1, \dots, a_n as parameters; also, the complement of $\{a_1, \dots, a_n\}$ is **definable** by $x \neq a_1 \wedge \dots \wedge x \neq a_n$.^a

On the other hand, let $X = \{x \in K \mid K \models \varphi(x, \bar{a})\}$ be a **definable** subset of K . By **quantifier elimination**, we may assume that φ is **quantifier-free**, so φ is a boolean combination of **atomic** and negated **atomic formulas**. Notice that an **atomic formula** is of the form

$$p_n(\bar{a})x^n + \dots + p_1(\bar{a})x + p_0(\bar{a}) = 0,$$

hence this **atomic formula defines** either a finite set or all of K . A negated **atomic formula defines** a **cofinite** set or \emptyset . Boolean combinations of finite and **cofinite** sets are finite or **cofinite**, so X is finite or **cofinite**. ■

^aWe did not use anything about fields here.

Remark. Proposition 4.1.2 is not true for $X \subseteq K^2$.

Proof. $X = \{(x, y) \mid x^2 + y^2 + 1\}$. ⊛

4.2 Definable and Constructible Sets

Proposition 4.1.2 shows some desirable properties, so we come up with the following definition.

Definition 4.2.1 (Strongly minimal). A **theory** T is *strongly minimal* if for any $\mathcal{M} \models T$, and $X \subseteq M$ **definable**, X is either finite or **cofinite**.

Now let us consider **definable** sets of higher arities.

Definition 4.2.2 (Algebraic). Let K be a field, and $X \subseteq K^n$. We say that X is *algebraic* if there is a set S of polynomials over K such that X is the zero set of S .

As previously seen. Recall Definition 3.2.15 and compared it to the above.

Example. $X = \{(x, y) \mid x^2 + y^2 = 1\}$ is **algebraic** since $S = \{x^2 + y^2 - 1\}$.

The complement of an **algebraic set** is usually not **algebraic**.

Definition 4.2.3 (Constructible). The *constructible* sets are the boolean combinations of **algebraic** sets.

Remark. The **constructible** sets are exactly the **definable** sets in $K \models \text{ACF}$.

Proof. The **definable** sets, by **quantifier elimination**, boolean combinations of sets **defined** by **atomic formulas**, which are **algebraic**. So **definable** implies **constructible**.

On the other hand, it is enough to see that **algebraic set** are **definable**. The issue is that $S \subseteq K[\bar{X}]$ might be infinite. Let I be the **ideal generated** by S . Then the set $X \subseteq K^n$ of common zeros of S is also the set of common zeros of I .^a By the **Hilbert's basis theorem**, $I = (f_1, \dots, f_m)$ is finitely

generated. Hence,

$$X = \{\bar{a} \in K^n \mid f_1(\bar{a}) = 0 \wedge \cdots \wedge f_m(\bar{a}) = 0\},$$

hence it's definable. ⊛

^aEach $f \in I$ is $f = r_1 g_1 + \cdots + r_n g_n$, where $g_1, \dots, g_n \in S$.

Theorem 4.2.1 (Chevalley's theorem). Let K be an algebraically closed field and $X \subseteq K^n$ be constructible. Let $p: K^n \rightarrow K^m$ be a polynomial map, i.e., $p(\bar{x}) = (q_1(\bar{x}), \dots, q_m(\bar{x}))$ for polynomials q_i . Then, $p(X)$, the image of X under p , is also constructible.

Proof. Since we know that constructible is the same as definable, so consider

$$p(X) = \{\bar{y} \mid \exists \bar{x} (\bar{x} \in X \wedge p(\bar{x}) = \bar{y})\}$$

where for $\bar{x} \in X$, there is a formula expressing this, and for $p(\bar{x}) = \bar{y}$, $y_1 = q_1(\bar{x}) \wedge \cdots \wedge y_m = q_m(\bar{x})$. Hence, $p(X)$ is definable (hence constructible), since X was. ■

Example. $p(x_1, x_2, x_3) = (x_1, x_3)$.

Theorem 4.2.2 (Weak Hilbert's Nullstellensatz). Let K be algebraically closed, and $f_1, \dots, f_n \in K[\bar{x}]$. Then there is $\bar{a} \in K^m$ such that $f_1(\bar{a}) = \cdots = f_n(\bar{a}) = 0$ if and only if $1 \notin (f_1, \dots, f_n)$, i.e., there are no $r_1, \dots, r_n \in K[\bar{x}]$ such that $1 = r_1 f_1 + \cdots + r_n f_n$.

Proof. If $1 \in (f_1, \dots, f_n)$, there are $r_1, \dots, r_n \in K[\bar{x}]$ such that $1 = r_1 f_1 + \cdots + r_n f_n$. If \bar{a} was a common zero of f_1, \dots, f_n , then

$$1 = r_1(\bar{a})f_1(\bar{a}) + \cdots + r_n(\bar{a})f_n(\bar{a}) = 0,$$

so no such \bar{a} exists.

Now, suppose $1 \notin (f_1, \dots, f_n)$, so $(f_1, \dots, f_n) \neq K[\bar{x}]$. Let I be a maximal ideal containing f_1, \dots, f_n , and let $L = K[\bar{x}] / I$, which is a field extension of K , $K \hookrightarrow L$. There is $\bar{a} \in L^m$ which is a common root of f_1, \dots, f_n , namely $a_i = x_i + I$. Let M be the algebraic closure of L , with $\bar{a} \in M^n$. By quantifier elimination, $K \succeq M$. $M \models \exists \bar{y} f_1(\bar{y}) = 0 \wedge \cdots \wedge f_n(\bar{y}) = 0$, which is a formula about elements of K (the coefficients). Because $K \succeq M$, $K \models \exists \bar{y} f_1(\bar{y}) = 0 \wedge \cdots \wedge f_n(\bar{y}) = 0$, which says that f_1, \dots, f_n have a common zero of K . ■

Note. We leave the full version in the note, which relates to algebraic geometry (if you care).

Remark. The weak Hilbert's Nullstellensatz says that whether $1 \in (f_1, \dots, f_n)$ is the only barrier for f_1, \dots, f_n having a common zero.

4.3 Algebraic Closure

As previously seen (Definable closure). The *definable closure* $\text{dcl}(A)$ of A is the set of all $b \in M$ which are definable over A .

On the homework, we looked at dcl , which is not really useful. In strongly minimal theories, we look at acl instead, which is very important.

Definition 4.3.1 ((Model theoretical) algebraic closure). Let \mathcal{M} be a structure and $A \subseteq M$. Then the (model-theoretic) algebraic closure $\text{acl}(A)$ of A is the set of all $a \in M$ such that there are $\bar{b} \in A$ and a formula $\varphi(x, \bar{b})$ such that $\mathcal{M} \models \varphi(a, \bar{b})$ and there are only finitely many other a' with $\mathcal{M} \models \varphi(a', \bar{b})$.

Note. There are some properties that acl always satisfies.

- $\text{dcl}(A) \subseteq \text{acl}(A)$.
- $A \subseteq \text{acl}(A)$.
- If $A \subseteq B$, then $\text{acl}(A) \subseteq \text{acl}(B)$.
- $\text{acl}(A) = \text{acl}(\text{acl}(A))$.^a
- If $a \in \text{acl}(A)$, then there is a finite $F \subseteq A$ such that $a \in \text{acl}(F)$.

^aThis is a good exercise!

On homework 5, we will show that in **models** of a **strongly minimal theory**, acl also satisfies the “exchange property”:

Remark (Exchange property). If $a \in \text{acl}(X \cup \{b\})$ and $a \notin \text{acl}(X)$, then $b \in \text{acl}(X \cup \{a\})$. This makes acl a *pregeometry* or *matroid*.

In **algebraically closed** fields, $a \in \text{acl}(X)$ just means that a is **algebraic** over (the field generated by) X , i.e., there is $p(y) \in F_X[\bar{x}]$ such that $p(a) = 0$. In vector spaces, $a \in \text{acl}(X)$ if and only if $a \in \text{span}(X)$.

Intuition. This makes **model** has notions like dimensions, independence, etc.

Lecture 19: acl in Algebraically Closed Fields

Theorem 4.3.1. Let $K \models \text{ACF}$, and $A \subseteq K$. The $\text{acl}(A)$ is the set of all elements $b \in K$ which are **algebraic** over A .^a

^aI.e., that satisfy a non-zero polynomial with coefficients in the ring generated by A .

Proof. If $p(x)$ is a non-zero polynomial over A , with $p(b) = 0$, then $\{c \mid p(c) = 0\}$, so $p(x) = 0$ witnesses that $b \in \text{acl}(A)$.

Suppose that $b \in \text{acl}(A)$, as witnessed by $\varphi(x, \bar{a})$ for $\bar{a} \in A$. We may assume that φ is **quantifier-free**. Write

$$\varphi(x, \bar{a}) = \psi_1(x, \bar{a}) \vee \cdots \vee \psi_n(x, \bar{a}),$$

where each ψ_i is a conjunction of **atomic** or negated **atomic formulas**. We may replace φ with some ψ_i , choosing one that b satisfies. Then,

$$\varphi(x, \bar{a}) = p_1(x, \bar{a}) \wedge \cdots \wedge p_k(x, \bar{a}) = 0 \wedge q_1(x, \bar{a}) \neq 0 \wedge \cdots \wedge q_\ell(x, \bar{a}) \neq 0.$$

In order for this to have finitely many solutions, we must have $k \geq 1$ and some $p_i(x, \bar{a})$ non-zero polynomial in x . But then b satisfies a polynomial over A . ■

16 Mar. 11:30

4.4 Types

Now, unless we specify, we will now work in a **model** of a **strongly minimal theory**.

Definition 4.4.1 (Independent). A set X is *independent* if for all $a \in X$, $a \notin \text{acl}(X - \{a\})$.

Intuition. Think about the case of vector spaces.

Definition 4.4.2 (Basis). A set X is a *basis* of Y if it's the maximal **independent** set in Y .

Note. In homework, we will show that each set Y always contains a **basis**.

Definition 4.4.3 (Dimension). The *dimension* $\dim Y$ of Y is this cardinality of its **basis**.

Remark. If X_1, X_2 are two **bases** for Y , then $|X_1| = |X_2|$, i.e., **Definition 4.4.3** is well-defined.

Definition 4.4.4 (Type). Let \mathcal{M} be an **\mathcal{L} -structure**, $\bar{c} \in M$, $A \subseteq M$. The *type* of \bar{c} over A (in \mathcal{M}) is

$$\text{tp}^{\mathcal{M}}(\bar{c}/A) = \{\varphi(\bar{x}, \bar{a}) \mid \mathcal{M} \models \varphi(\bar{c}, \bar{a}) \text{ and } \bar{a} \in A\}.$$

Notation. Where we omit $/A$, we just mean $/\emptyset$.

Lemma 4.4.1. Let T be a **complete strongly minimal theory**, and $\mathcal{M}, \mathcal{N} \models T$. Let $\bar{a} \in M$ and $\bar{b} \in N$ be **independent** tuples of the same size, then $\text{tp}^{\mathcal{M}}(\bar{a}) = \text{tp}^{\mathcal{N}}(\bar{b})$, i.e., $\mathcal{M} \models \varphi(\bar{a}) \Leftrightarrow \mathcal{N} \models \varphi(\bar{b})$.

Proof. We do an induction on the lengths of \bar{a}, \bar{b} . Start with $n = 1$: let a, b be **independent**, then

$$\text{tp}(a) = \{\varphi(x) \mid \varphi(x) \text{ has \textbf{cofinitely} many solutions in } \mathcal{M}\}$$

since a is **independent**, $a \notin \text{acl}(\emptyset)$, i.e., $\varphi(\mathcal{M}) := \{c \in M \mid \mathcal{M} \models \varphi(c)\}$ is infinite for all φ . And since \mathcal{M} is strongly minimal, it's **cofinite**.

Example. In ACF, $\text{tp}(a) = \{p(a) \neq 0 \mid p \in \mathbb{Q}[x], p \neq 0\}$ (or “generated by” these). Also, $(1+1)x$ being 0 depends on the **characteristic**.

Similarly, we have

$$\text{tp}(b) = \{\varphi(x) \mid \varphi(x) \text{ has \textbf{cofinitely} many solutions in } \mathcal{N}\}.$$

Now, suppose that φ has k non-solutions in \mathcal{M} , $\mathcal{M} \models \exists^{=k} x \neg \varphi(x)$, with T being **complete**, $T \models \exists^{=k} x \neg \varphi(x)$, hence $\mathcal{N} \models \exists^{=k} x \neg \varphi(x)$. So if φ has **cofinitely** many (all but k) solutions in \mathcal{M} , the same is true in \mathcal{N} . Hence, $\text{tp}(a) = \text{tp}(b)$.

Example. The **completeness** is important: ACF is a non-example.

For $n+1$, let $\bar{a}a'$ and $\bar{b}b'$ be **independent** $(n+1)$ -tuples with $\text{tp}(\bar{a}) = \text{tp}(\bar{b})$. Suppose $\mathcal{M} \models \varphi(\bar{a}, a')$. Since $\mathcal{M} \models T$ is **strongly minimal** and $a' \notin \text{acl}(\bar{a})$, $\varphi(\bar{a}, \mathcal{M}) = \{c \in M \mid \mathcal{M} \models \varphi(\bar{a}, c)\}$ is **cofinite** with complement of size k . Then $\mathcal{M} \models \exists^{=k} x \neg \varphi(\bar{a}, x)$, so $\exists^{=k} x \neg \varphi(\bar{b}, x) \in \text{tp}(\bar{a}) = \text{tp}(\bar{b})$, i.e., $\mathcal{N} \models \exists^{=k} x \neg \varphi(\bar{b}, x)$. Then, since $b' \notin \text{acl}(\bar{b})$, so $b' \in \varphi(\bar{b}, \mathcal{N})$,^a hence $\mathcal{N} \models \varphi(\bar{b}, b')$. ■

^aSince $\varphi(\bar{b}, \mathcal{N})$ is **cofinite**, and if b' is in the complement of $\varphi(\bar{b}, \mathcal{N})$, which is finite, $b' \in \text{acl}(\bar{b})$ by $\neg \varphi(\bar{b}, x) \notin$

Lemma 4.4.2. Let T be a **strongly minimal theory**. If $\mathcal{M} \models T$ and $c \in \text{acl}(A)$, then there is a **formula** $\varphi(x, \bar{a}) \in \text{tp}(c/A)$ such that for any other $c' \in M$ with $\mathcal{M} \models \varphi(c', \bar{a})$, $\text{tp}(c/A) = \text{tp}(c'/A)$.

Proof. Let $\varphi(x, \bar{a}) \in \text{tp}(c/A)$ be such that

$$|\{x \in M \mid \mathcal{M} \models \varphi(x, \bar{a})\}| = k$$

is minimal. We claim that $\varphi(x, \bar{a})$ witnesses the statement. If not, there are some c' and a **formula** $\psi(x, \bar{a}') \in \text{tp}(c'/A)$ with $\mathcal{M} \models \varphi(c', \bar{a}) \wedge \neg \psi(c', \bar{a}')$. So

$$\{x \in M \mid \mathcal{M} \models \varphi(x, \bar{a}) \wedge \psi(x, \bar{a}')\} \subsetneq \{x \in M \mid \mathcal{M} \models \varphi(x, \bar{a})\}$$

since c' is in the right-hand side but not the left-hand side. This implies that the left-hand side has cardinality $< k$, a contradiction, so φ does imply ψ . ■

Theorem 4.4.1. Let T be a **complete strongly minimal theory**. If $\mathcal{M}, \mathcal{N} \models T$, then $\mathcal{M} \cong \mathcal{N}$ if and only if $\dim \mathcal{M} = \dim \mathcal{N}$.

Proof. Suppose $\dim \mathcal{M} = \dim \mathcal{N}$ with A, B being the **bases** for \mathcal{M}, \mathcal{N} . Then, we have $|A| = |B|$, so there exists a bijection $f: A \rightarrow B$, which is a **partial elementary map** since for $f, \bar{a} \in A$, $\text{tp}(\bar{a}) = \text{tp}(f(\bar{a}))$ from **Lemma 4.4.1**.

Definition 4.4.5 (Paritla elementary map). A map $g: U \subseteq M \rightarrow N$ is a *partial elementary map* if $\mathcal{M} \models \varphi(\bar{a}) \Leftrightarrow \mathcal{N} \models \varphi(g(\bar{a}))$ for $\bar{a} \in \text{dom } g$.^a

^aThis implies injectivity since if $\bar{a} \neq \bar{a}'$ with $g(\bar{a}) = g(\bar{a}')$, there exists some φ differentiate them.

By **Zorn's lemma**, there exists a maximal **partial elementary map** g from $\mathcal{M} \rightarrow \mathcal{N}$ extending f .

Claim. g is an **isomorphism**, i.e., $\text{dom } g = M$ ($\text{Im } g = N$ is automatic).

Proof. Assume that $c \in M - \text{dom } g$. We know that $c \in \text{acl}(A) \subseteq \text{acl}(\text{dom } g)$ since A is a **basis**, so let $\varphi(x, \bar{d})$ be the **formula** from **Lemma 4.4.2** that isolates $\text{tp}(c/\text{dom } g)$, i.e., whenever $\mathcal{M} \models \varphi(c', \bar{d})$, $\text{tp}(c'/\text{dom } g) = \text{tp}(c/\text{dom } g)$. Then, $\mathcal{M} \models \exists x \varphi(x, \bar{d})$, implying $\mathcal{N} \models \exists x \varphi(x, g(\bar{d}))$. Let $c' \in N$ witness this. Then $\text{tp}^{\mathcal{M}}(c/\text{dom } g) = \text{tp}^{\mathcal{N}}(c'/\text{Im } g)$ after identifying $\text{dom } g$ and $\text{Im } g$.^a Then we can define $g(c) = c'$ to extend g but remain **partial elementary**, contradicting to the maximality of g , so $\text{dom } g = M$.

Following the same argument, we can show that $\text{Im } g = N$, hence g is an **isomorphism**. \otimes

^aSuppose $\mathcal{M} \models \psi(c, \bar{b})$ for $\bar{b} \in \text{dom } g$, then in \mathcal{M} , $\mathcal{M} \models \forall x \varphi(x, \bar{a}) \rightarrow \psi(x, \bar{b})$. Since g is **elementary**, $\mathcal{N} \models \forall x \varphi(x, g(\bar{a})) \rightarrow \psi(x, g(\bar{b}))$, so $\mathcal{N} \models \psi(c', \bar{b})$.

The other direction is trivial, since if $\mathcal{M} \cong \mathcal{N}$, then clearly $\dim \mathcal{M} = \dim \mathcal{N}$. \blacksquare

Lecture 20: Fraïssé Theorem

4.5 Other Examples of Quantifier Elimination

21 Mar. 11:30

As previously seen. ACF, DLO, and also the **random graph theory**.^a

^aThis is very similar to the argument we gave for DLO.

We discuss some other examples of **theories** which admit **quantifier elimination**.

4.5.1 Vector Spaces

On the problem set, we will show that the **theory** of infinite (dimensional) vector spaces admits **quantifier elimination** and is **strongly minimal**.

4.5.2 Torsion-Free Abelian Groups

As previously seen (Torsion-free). An Abelian group G is said to be *torsion-free* if no element other than e has finite order.

Definition 4.5.1 (Divisible). A **torsion-free** group G is *divisible* if for every $n \in \mathbb{N}$ and $g \in G$, there is $h \in G$ such that

$$nh = \underbrace{h + \cdots + h}_{n \text{ times}} = g.$$

Since G is **torsion-free**, for each g , the corresponding h is unique.¹ The **theory** DAG of **torsion-free divisible** Abelian groups (i.e., \mathbb{Q} -vector spaces in **language** $\{+, 0, -\}$) admits **quantifier elimination**.

¹As if $nh = g = nh'$, then $n(h - h') = 0$ so $h = h'$.

4.5.3 Ordered Torsion-Free Divisible Abelian Groups

On top of DAG, if we also add an ordering, we get ODAG, the theory of ordered torsion-free divisible Abelian groups, which also admits quantifier elimination.

4.5.4 Presburger Arithmetic

Presburger arithmetic is $\text{Th}(\mathbb{N}, 0, 1, +)$.

4.5.5 Real Closed Fields

As previously seen (Real-closed field). A *real-closed field* is a field F that has the same first-order properties as the field of real numbers.

The theory of real-closed fields, denoted as RCF, does not admit quantifier elimination in the language $\mathcal{L} = \{0, 1, +, -, \cdot\}$. This is because the formula

$$\varphi(x, y) = \exists z \, x - y = z^2$$

is not equivalent to a quantifier-free formula.

Intuition. This formula is defining the ordering of the field.

Claim. $\varphi(x, y) = \exists z \, x - y = z^2$ is not equivalent to a quantifier-free formula in $\mathcal{L} = \{0, 1, +, -, \cdot\}$.

Proof. The first way one might try to prove this is to show that there are real-closed fields $\mathcal{M} \subseteq \mathcal{N}$ and $a, b \in \mathcal{N}$ such that $\mathcal{N} \models \varphi(a, b)$ but $\mathcal{M} \models \neg\varphi(a, b)$. But this strategy will not work, since $\varphi(x, y)$ is also equivalent, in any real-closed field, to the universal formula $\forall z \, y - x \neq z^2$. This is because only one of $x - y$ and $y - x$ can have a square root in a real-closed field. Instead, one should think about the test for quantifier elimination.

Proposition 4.5.1. There is no quantifier-free formula $\psi(x, y)$ such that RCF proves that $\psi(x, y)$ and $\varphi(x, y)$ are equivalent.

⊛

But if we add in the ordering as a symbol in the language, then in the language $\mathcal{L}_<$, the theory $\text{Th}(\mathbb{R}, 0, 1, +, -, \cdot, <)$ of real-closed ordered fields does admit quantifier elimination. This was shown by Tarski in the 1940s as part of showing that the theory of \mathbb{R} is decidable.

Example. An example of eliminating quantifiers from a formula that we already know is

$$\text{RCF} \models \exists x \, (a \neq 0 \wedge ax^2 + bx + c = 0) \leftrightarrow a \neq 0 \wedge b^2 - 4ac \geq 0.$$

After proving quantifier elimination, we can analyze the definable sets like we did for algebraically closed fields.

Remark (*o-minimal*). The theory of real-closed fields is not strongly minimal, but instead what is called *o-minimal*.^a

^aThere is a well-developed theory of *o-minimality* which would be an entire course in itself.

Proof. The definable sets in one variable are finite unions of intervals and points.

⊛

Remark. Tarski-Seidenberg algorithm can solve all Euclidean geometry problems.

Chapter 5

Fraïssé Limits

Our two examples of [countably categorical theories](#), DLO and the [random graph](#), are both characterized by an [extension axiom](#).¹ This influences a general way of construction [countably categorical structures](#).

5.1 Substructures' Properties

Definition 5.1.1 (Generated substructure). Given an \mathcal{L} -structure \mathcal{M} , and a set $A \subseteq M$, we write $\langle A \rangle$ for the [substructure](#) of M generated by A , defined as the smallest [substructure](#) of \mathcal{M} whose domain contains A .

Intuition. Equivalently, it's the [substructure](#) of \mathcal{M} containing A and all the constants, and is closed under the application of functions.

Note. Compare [Definition 5.1.1](#) to the [substructure generated](#) in the [downward Löwenheim-Skolem theorem](#).

And it's natural to talk about finiteness.

Definition 5.1.2 (Finitely generated). A [substructure](#) \mathcal{N} of \mathcal{M} is *finitely generated* if it is $\mathcal{N} = \langle A \rangle$ for some finite $A \subseteq M$.

Now, the question is how should we build [structures](#)? Idea is that to build “universal” countable [structures](#) with “all possible finitary behaviors”. Assume (for now) \mathcal{L} is relational.

Definition 5.1.3 (Age). For an \mathcal{L} -structure \mathcal{M} , the *age* of \mathcal{M} , $\text{Age}(\mathcal{M})$, is the class of all finite [\$\mathcal{L}\$ -substructure](#) which extend into \mathcal{M} .

Definition 5.1.4 (Hereditary property). A class \mathbb{K} of finite \mathcal{L} -structure has the *hereditary property* if for all $\mathcal{B} \in \mathbb{K}$ and \mathcal{A} which extends into \mathcal{B} , $\mathcal{A} \in \mathbb{K}$.

Intuition. If it's downward-closed under [embedding](#).

Definition 5.1.5 (Joint embedding property). A class \mathbb{K} of finite \mathcal{L} -structure has the *joint embedding property* if for all $\mathcal{A}, \mathcal{B} \in \mathbb{K}$, there exists $\mathcal{C} \in \mathbb{K}$ with [embeddings](#) $\mathcal{A} \hookrightarrow \mathcal{C} \hookleftarrow \mathcal{B}$:



¹Though not explicit for DLO, but essentially given $a < b$, we can “extend” this order by finding c such that $a < c < b$.

5.2 “Baby” Fraïssé Theorem

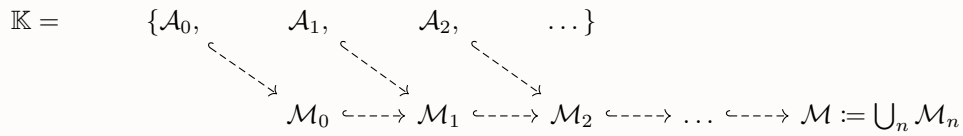
Consider the following.

Theorem 5.2.1 (Baby Fraïssé theorem). A class \mathbb{K} of finite \mathcal{L} -structure is $\text{Age}(\mathcal{M})$ for some countable \mathcal{M} if and only if

- \mathbb{K} is countable up to isomorphism;
- $\mathbb{K} \neq \emptyset$;
- \mathbb{K} has the hereditary property;
- \mathbb{K} has the joint embedding property.

Proof. The forward direction is clear. For the backward direction, let $\mathbb{K} = \{\mathcal{A}_0, \mathcal{A}_1, \dots\}$, we construct $\mathcal{M}_0 \subseteq \mathcal{M}_1 \subseteq \mathcal{M}_2 \subseteq \dots$ inductively as follows:

- $\mathcal{M}_0 := \mathcal{A}_0$.
- \mathcal{M}_{n+1} is chosen from \mathbb{K} using joint embedding property such that $\mathcal{M}_n \hookrightarrow \mathcal{M}_{n+1} \hookleftarrow \mathcal{A}_{n+1}$.



Let $\mathcal{M} := \bigcup_n \mathcal{M}_n$. Then,

- $\mathbb{K} \subseteq \text{Age}(\mathcal{M})$: because each $\mathcal{A}_n \in \mathbb{K}$ is embedded into $\mathcal{M}_n \subseteq \mathcal{M}$, so \mathcal{A}_n is embedded into \mathcal{M} ;
- $\text{Age}(\mathcal{M}) \subseteq \mathbb{K}$: for a finite subset $\mathcal{N} \subseteq \mathcal{M}$, we have $\mathcal{N} \subseteq \mathcal{M}_n$ for some n , so by hereditary property, $\mathcal{M}_n \in \mathbb{K}$, hence $\mathcal{N} \in \mathbb{K}$.

■

Example. $\text{Age}(\mathbb{Q}, <) = \text{Age}(\mathbb{N}, <) = \text{Age}(\mathbb{Z}, <) = \{\text{all finite linear orders}\}$, and in fact, it's also the age of any infinite linear order.

Example. $\text{Age}(\text{random graph theory}) = \{\text{all finite graphs}\} = \text{Age}(\coprod \{\text{finite graphs}\})$.



5.3 Fraïssé Theorem

Fraïssé asked when the class \mathbb{K} determines a single structure. It turns out that in addition to the hereditary property and the joint embedding property, we need a third property of a class \mathbb{K} .

Definition 5.3.1 (Ultrahomogeneous). A countable structure \mathcal{M} is *ultrahomogeneous* if for any finite subsets $\mathcal{A}, \mathcal{B} \subseteq \mathcal{M}$ and isomorphism $g: \mathcal{A} \cong \mathcal{B}$, there is an automorphism $\tilde{g}: \mathcal{M} \cong \mathcal{M}$ extending g .

$$\begin{array}{ccc} \mathcal{A} & \xrightarrow{g} & \mathcal{B} \\ \downarrow & & \downarrow \\ \mathcal{M} & \xrightarrow{\tilde{g}} & \mathcal{M} \end{array}$$

Definition 5.3.2 (Amalgamation property). A class \mathbb{K} of finite \mathcal{L} -structure has the *amalgamation property* if for all $\mathcal{A}, \mathcal{B}, \mathcal{C} \in \mathbb{K}$ such that



such that $\tilde{f} \circ f = \tilde{g} \circ g$, i.e., the following diagram commutes:



Intuition. We can “glue” \mathcal{B} and \mathcal{C} along their “common part” \mathcal{A} to get \mathcal{D} .

Definition 5.3.3 (Extension property). A countable structure \mathcal{M} has the *extension property* w.r.t. a class \mathbb{K} of (finite) structure if for all $\mathcal{A}, \mathcal{B} \in \mathbb{K}$ and $f: \mathcal{A} \rightarrow \mathcal{M}$ and $g: \mathcal{A} \hookrightarrow \mathcal{B}$, there exists $h: \mathcal{B} \hookrightarrow \mathcal{M}$ such that $h \circ g = f$.

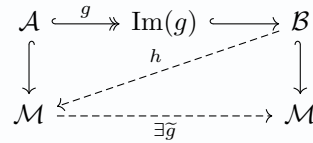


Intuition. This is a direct generalization of *extension axiom* (as the name suggests).

Note. \mathcal{M} *ultrahomogeneous* implies \mathcal{M} has *extension property* w.r.t. $\text{Age}(\mathcal{M})$.^a

^aIn homework, we will show the converse. More generally, \mathcal{M} has *extension property* w.r.t. $\text{Age}(\mathcal{N})$ implies \mathcal{N} *embeds* into \mathcal{M} , so \mathcal{M} is a *Fraïssé limit*, hence $\text{Age}(\mathcal{N}) \subseteq \text{Age}(\mathcal{M}) \Rightarrow \mathcal{N} \hookrightarrow \mathcal{M}$.

Proof. Without loss of generality, let $\mathcal{A}, \mathcal{B} \subseteq \mathcal{M}$ and f is the inclusion, we have



where $h := \tilde{g}^{-1}|_{\mathcal{B}}$.

⊛

We now see the generalized version of *baby Fraïssé theorem*.

Theorem 5.3.1 (Fraïssé theorem). A class \mathbb{K} of finite \mathcal{L} -structure is $\text{Age}(\mathcal{M})$ for an *ultrahomogeneous* countable \mathcal{M} if and only if

- \mathbb{K} is countable up to *isomorphism*;
- $\mathbb{K} \neq \emptyset$;
- \mathbb{K} has the *hereditary property*;
- \mathbb{K} has the *joint embedding property*;

- \mathbb{K} has the [amalgamation property](#).

Moreover, in that case, there exists a unique-up-to-[isomorphism](#) countable [ultrahomogeneous](#) \mathcal{M} such that $\text{Age}(\mathcal{M}) = \mathbb{K}$.

Definition 5.3.4 (Fraïssé class). A class \mathbb{K} with properties described in [Fraïssé theorem](#) is called a *Fraïssé class*.

Definition 5.3.5 (Fraïssé limit). The countable [ultrahomogeneous](#) \mathcal{M} such that $\text{Age}(\mathcal{M}) = \mathbb{K}$ is the *Fraïssé limit* of \mathbb{K} , denoted as $\text{Flm}(\mathbb{K})$.

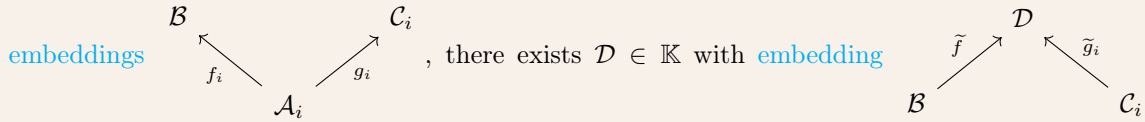
Lecture 21: Proof of Fraïssé Theorem

As previously seen. We want to build an infinite structure from finite pieces.

23 Mar. 11:30

To prove the [Fraïssé theorem](#), we need the following lemma.

Lemma 5.3.1. Let \mathbb{K} have the [amalgamation property](#). Then for $\mathcal{B}, \mathcal{A}_1, \dots, \mathcal{A}_n, \mathcal{C}_1, \dots, \mathcal{C}_n \in \mathbb{K}$ with



such that $\tilde{f} \circ f_i = \tilde{g}_i \circ g_i$, i.e.,

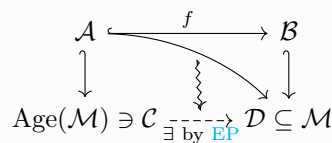


Proof. By induction on n with [amalgamation property](#), we have



Now, we try to prove the [Fraïssé theorem](#).

Proof of Theorem 5.3.1. To prove the forward direction, we need only to check [AP](#) for $\text{Age}(\mathcal{M})$. We use [EP](#) to extend $\mathcal{A} \hookrightarrow \mathcal{B} \hookrightarrow \mathcal{M}$ along $\mathcal{A} \hookrightarrow \mathcal{C}$. Let \mathcal{D} be the union of images of $\mathcal{C} \hookrightarrow \mathcal{M}$ and $\mathcal{B} \hookrightarrow \mathcal{M}$.



For the backward direction, we'll build $\mathcal{M}_0 \subseteq \mathcal{M}_1 \subseteq \dots$ as in [baby Fraïssé theorem](#) to ensure

\mathcal{M} obeys countably many “conditions”. A “condition” means:

- a **structure** $\mathcal{A} \in \mathbb{K}$ (means we need to **embed** \mathcal{A} into \mathcal{M}_n at some stage),
- a tuple $(\mathcal{A}, \mathcal{B}, g)$ where $\mathcal{A}, \mathcal{B} \in \mathbb{K}$ and g is an **embedding** $\mathcal{A} \hookrightarrow \mathcal{B}$ (means we need to **amalgamate** g into \mathcal{M}_n).

Then, we list all conditions

$$(\mathcal{A}_0, (\mathcal{A}_1, \mathcal{B}, g_1), \mathcal{A}_2, (\mathcal{A}_3, \mathcal{B}_3, g_3), \dots)$$

such that each condition of second type occurs infinitely often. Consider

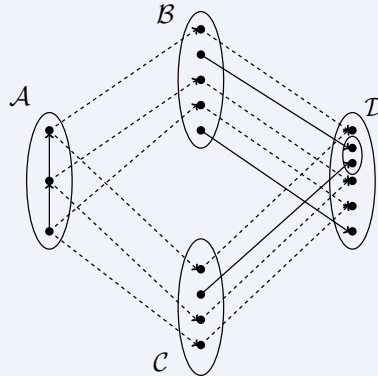


where $f_{1,i}$ are all **embeddings** $\mathcal{A}_1 \hookrightarrow \mathcal{M}_i$, for example. Then as before, $\text{Age}(\mathcal{M}) = \mathbb{K}$. To check **extension property** w.r.t. \mathbb{K} , we have

$$\begin{array}{ccc} \mathbb{K} \ni \mathcal{A} & \xrightarrow{f} & \mathcal{M}_n \subseteq \mathcal{M} \\ \downarrow & & \\ \mathbb{K} \ni \mathcal{B} & & \end{array}$$

since each triple we listed are infinitely often, at some $m \geq b$, $(\mathcal{A}_m, \mathcal{B}_m, g_m) = (\mathcal{A}, \mathcal{B}, g)$, so we just **amalgamate** this into \mathcal{M}_{m+1} . ■

Example. Let \mathbb{K} be finite linear orders which has **amalgamation property**:



Note. \mathbb{K} has the strong [amalgamation property](#) (SAP) if it's always possible to [amalgamate](#) such that $\text{Im}(\tilde{f}) \cap \text{Im}(\tilde{g}) = \text{Im}(\tilde{f} \circ f)$.^a



^aRecall that $\tilde{f} \circ f = \tilde{g} \circ g$.

Theorem 5.3.2. A [Fraïssé limit](#) has $\text{Age}(\mathcal{M})$ with SAP if and only if \mathcal{M} has trivial acl, i.e., $\text{acl}(A) = A$ for all $A \subseteq M$.

Lecture 22: Ryll-Nardzewski Theorem

Example. $(\mathbb{Q}, <) = \text{Flm}(\text{finite linear orders})$.

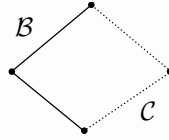
28 Mar. 11:30

Example. [random graph](#) = $\text{Flm}(\text{finite graphs})$.

Example. random K_d graph = $\text{Flm}(\text{finite } K_d\text{-free graphs})$.

Example. Finite square-free graphs do not have [amalgamation property](#).

Proof. Consider the following:



We see that after “amalgamate” \mathcal{B} with \mathcal{C} , the result becomes non-square-free! ⊗

Example. random poset = $\text{Flm}(\text{finite posets})$.

Example. $\mathbb{N} = \text{Flm}(\text{finite sets})$.

Example. $(\mathbb{N}^2, \sim) = \text{Flm}(\text{finite sets equipped with } \sim) \text{ where } (a, b) \sim (c, d) \Leftrightarrow a = c$.

Theorem 5.3.3. Let \mathcal{L} be a finite relational [language](#). Then every [ultrahomogeneous](#) \mathcal{L} -structure $\text{Th}(\mathcal{M})$ (i.e., $\mathcal{M} = \text{Flm}(\text{Age}(\mathcal{M}))$) is [countably categorical](#) and has [quantifier elimination](#).

Proof. Let $\mathbb{K} = \text{Age}(\mathcal{M})$, then $\mathcal{N} \cong \mathcal{M}$ if and only if $\text{Age}(\mathcal{N}) = \mathbb{K}$ and \mathcal{N} has [extension property](#) w.r.t. \mathbb{K} ^a from the [Fraïssé theorem](#).

For each finite \mathcal{L} -structure $\mathcal{A} = \{a_1, \dots, a_n\}$, let $\phi_{\mathcal{A}}(x_1, \dots, x_n)$ be an \mathcal{L} -formula such that for all \mathcal{N} , $\mathcal{N} \models \phi(\bar{b})$ if and only if $a_i \mapsto b_i$ is an [isomorphism](#) and $\mathcal{A} \hookrightarrow \mathcal{N}$. Then take the axioms, we want that $\forall \bar{x} \phi_{\mathcal{A}}(\bar{x}) \rightarrow “\mathcal{A} \in \mathbb{K}”$, but this is not first-order, so we instead consider $\forall \bar{x} \neg \phi_{\mathcal{A}}(\bar{x})$ for each $\mathcal{A} \notin \mathbb{K}$. Moreover, we consider $\forall \bar{x} \phi_{\mathcal{A}}(\bar{x}) \rightarrow \exists \bar{y} \phi_{\mathcal{B}}(\bar{x}, \bar{y})$ for each $\mathcal{A}, \mathcal{B} \in \mathbb{K}$, for $\mathcal{A} \subseteq \mathcal{B}$. ■

^aThis implies $\text{Age}(\mathcal{N}) = \mathbb{K}$, by [extension property](#) applied to $\mathcal{N} \hookleftarrow \emptyset \hookrightarrow \mathcal{A}$, hence $\mathcal{A} \hookrightarrow \mathcal{N}$.

Remark. We can also add $\exists \bar{x} \phi_{\mathcal{A}}(\bar{x})$ for each $\mathcal{A} \in \mathbb{K}$ to overcome the fact that we're considering \emptyset as an \mathcal{L} -structure.

The following theorem tells us that what do countably categorical models look like. We will not prove this.

Theorem 5.3.4 (Ryll-Nardzewski theorem). A countable structure \mathcal{M} has $\text{Th}(\mathcal{M})$ countably categorical if and only if the following equivalent conditions hold.

- (i) For each $n \in \mathbb{N}$, there are finitely many automorphism orbits of n -tuples $\bar{a} \in M^n$. In other words, the action $\text{Aut}(\mathcal{M}) \curvearrowright M^n$ has finitely many orbits.^a
- (ii) For each $n \in \mathbb{N}$, there are only finitely many types of n -tuples $\bar{a} \in M^n$.

^aWe say the action is oligomorphic.

Now, if we change every “finite” to “finitely generated” so far, we can show the same thing (for, e.g., the Fraïssé theorem).² However, if we look into the proof of Theorem 5.3.3, we see that we really need \mathcal{A} to be finite, hence we consider the following.

Definition 5.3.6 (Uniformly locally finite). A class \mathbb{K} is *uniformly locally finite* if for all $n \in \mathbb{N}$, there exists $k \in \mathbb{N}$ such that every n -generated structure in \mathbb{K} has size $\leq k$.

Example. For a finite field F , an F -vector space of $\dim \leq n$ has size $\leq |F|^n$.

Then, the statement of Theorem 5.3.3 need to be changed as follows.

Theorem 5.3.5. Let \mathcal{L} be a finite language. Then every ultrahomogeneous \mathcal{L} -structure $\text{Th}(\mathcal{M})$ such that $\text{Age}(\mathcal{M})$ is uniformly locally finite is countably categorical and has quantifier elimination.

Example. $F^{\oplus \omega} = \text{Flm}(\text{finite dimensional } F\text{-vector spaces})$ for all countable field F .

Example. $\mathbb{F}_p = \text{Flm}(\text{finite characteristic } p \text{ fields})$ for prime p .

²Though the proofs are slightly different.

Chapter 6

Ultrapowers

Lecture 23: Introduction to Ultraproducts

6.1 Ultrafilters

30 Mar. 11:30

6.1.1 Filters

Definition 6.1.1 (Filter). Let I be a non-empty set, and $\mathcal{P}(I)$ be the power set of I . A *filter* D on I is a collection $D \subseteq \mathcal{P}(I)$ such that

- (a) $I \in D$ and $\emptyset \notin D$;
- (b) if $A, B \in D$, then $A \cap B \in D$;
- (c) if $A \in D$, and $B \supseteq A$, then $B \in D$.

Intuition. Think of D as a collection of *very large subsets*, i.e., the set containing everything is large:

- (a) the empty set is not big;
- (b) anything bigger than a large set is large;
- (c) the intersection of two large sets is still large.

The most important examples of **filters** are the following.

Example (Principal filter). Let I be any set, $a \in I$. Then $D_a = \{A \subseteq I \mid a \in A\}$ is the *principal filter* generated by a . It says that the large sets are exactly those containing a .

Example (Frechet filter). Let I be any infinite set, and let D be the collection of all **cofinite** subsets of I . This is the *Frechet filter*.

In particular, a **Frechet filter** D is not **principal**, and any **filter** $D' \supseteq D$ is not **principal** because for $a \in I$, $I - \{a\} \in D \subseteq D'$, so if $\{a\} \in D'$, then $\emptyset = \{a\} \cap (I - \{a\}) \in D' \nsubseteq$ so $\{a\} \notin D'$.

Example (Extend). If D is a **filter** on I , and $X \subseteq I$ has $X \notin D$, then we can *extend* D to a **filter** $E \supseteq D$ containing $I - X$, namely $E = \{Y \subseteq I \mid Z - X \subseteq Y \text{ for some } Z \in D\}$.

6.1.2 Ultrafilters

Consider the following.

Definition 6.1.2 (Ultrafilter). A **filter** D on I is an *ultrafilter* if for all $X \subseteq I$, either $X \in D$ or

$I - X \in D$.

Note. It can't be both as $X \cap (I - X) = \emptyset \notin D$.

Intuition. One way to think of an **ultrafilter** is a finitely additive $\{0, 1\}$ -valued measure on $\mathcal{P}(I)$.^a

^aI.e., it takes only the values 0 on sets not in the **ultrafilter**, and 1 otherwise.

Example. The **principal filters** are (not very interesting) **ultrafilters**.

To see that there are interesting **ultrafilters**, we show that we can **extend** any **filter** to an **ultrafilter**; and in particular, if we **extend** the **Frechet filter**, we will get a non-**principal ultrafilter**.

Theorem 6.1.1. Let D be a **filter** on I , then there is \mathcal{U} an **ultrafilter** on I extending D .

Proof. Use **Zorn's lemma**. An **ultrafilter** will be a **maximal filter**. Let $P = \{E \text{ filter on } I \mid E \supseteq D\}$, and we order P by inclusion.

Consider a non-empty **chain** C of **filters** $E \supseteq D$, then $D^* = \bigcup_{E \in C} E$ is a **filter** extending D .^a D^* is then an **upper bound** for C , so **Zorn's lemma** applies. Let \mathcal{U} be a **maximal** element of \mathcal{P} , and we argue that \mathcal{U} is an **ultrafilter** extending D .

Suppose \mathcal{U} is not an **ultrafilter**, so there is $X \subseteq I$ such that $X \notin \mathcal{U}$, $I - X \notin \mathcal{U}$. From the **extending example**, we know that there is $\mathcal{U}' \supseteq \mathcal{U}$ a **filter** containing $I - X$, contradicting the **maximality** of \mathcal{U} , so \mathcal{U} is an **ultrafilter**. ■

^a $D^* \supseteq D$ since $E \in C$ did. D^* is a **filter** since if $A, B \in D^*$, then there is $E \in C$, $A, B \in E$, so $A \cap B \in E \subseteq D^*$.

Intuition. Think of an **ultrafilter** is a sort of “voting”: the set I as a set of voters, and a set $X \subseteq I$ of voters says “yes”, then “yes” wins if X is in the **ultrafilter**; and if $I - X$ is in the **ultrafilter** instead, then “no” wins.^a The closure of the **ultrafilter** under intersections is a very strong property which fails for elections with finitely many voters.

^aThe “ultra” part of the **ultrafilter** means that either “yes” or “no” will always win.

Remark. An **ultrafilter** on a finite set is **principal**.

Proof. Let $I = \{a_1, \dots, a_n\}$, and \mathcal{U} be an **ultrafilter**. For each a_i , either $\{a_i\} \in \mathcal{U}$ or $I - \{a_i\} \in \mathcal{U}$.

- If we're in the first case for any i , then \mathcal{U} is **principal**.
- Otherwise, $(I - \{a_1\}) \cap (I - \{a_2\}) \cap \dots \cap (I - \{a_n\}) = \emptyset \in \mathcal{U}$, a contradiction.

⊛

The above proof shows more.

Remark. If \mathcal{U} is a non-**principal ultrafilter** on I , \mathcal{U} extends the **Frechet filter**.

Ultimately, with the voting analogy, it's possible that (with finitely many voters) a majority of the voters prefer ice cream to cake, and that a majority of the voters prefer ice cream to candy, but that only a minority of voters prefer ice cream to both cake and candy.

Note. This is called the Condorcet paradox of voting,^a and is related to **Arrow's impossibility theorem**. One proof of which is to show that any “good” method of voting must be an **ultrafilter**, and that any **ultrafilter** on a finite set is **principal**, i.e, we must have a dictator.

^aTake SI652 to learn more!

6.2 Ultraproducts

We use this voting analogy to construct new **models** called **ultraproducts**.

Definition 6.2.1 (Ultraproduct). Let \mathcal{U} be an **ultrafilter** on I and for each $i \in I$, let \mathcal{M}_i be an **\mathcal{L} -structure**. The *ultraproduct* of the \mathcal{M}_i 's is an **\mathcal{L} -structure** $\prod_{i \in I} \mathcal{M}_i / \mathcal{U}$ defined as follows.

Let $\prod_{i \in I} M_i$ be the tuples $(a_i)_{i \in I}$ indexed by i ,^a and define an equivalence relation \sim (or $\sim_{\mathcal{U}}$) on $\prod_{i \in I} M_i$ by $(a_i)_{i \in I} \sim (b_i)_{i \in I}$ if $\{i \in I \mid a_i = b_i\} \in \mathcal{U}$. Then, the domain of $\prod_{i \in I} \mathcal{M}_i / \mathcal{U}$ will be $\prod_{i \in I} M_i / \sim$, the equivalence classes mod $\sim_{\mathcal{U}}$.

(a) For each constant $c \in \mathcal{L}$,

$$c^{\prod_{i \in I} \mathcal{M}_i / \mathcal{U}} = (c^{\mathcal{M}_i})_{i \in I} / \sim =: [(c^{\mathcal{M}_i})_{i \in I}].$$

(b) For each function symbol $f \in \mathcal{L}$ of arity n , and $[(a_i^1)], \dots, [(a_i^n)] \in \prod_{i \in I} \mathcal{M}_i / \sim$,

$$f^{\prod_{i \in I} \mathcal{M}_i / \mathcal{U}}([(a_i^1)], \dots, [(a_i^n)]) = (f^{\mathcal{M}_i}(a_i^1, \dots, a_i^n))_{i \in I} / \sim =: [(f^{\mathcal{M}_i}(a_i^1, \dots, a_i^n))_{i \in I}].$$

(c) For each relation symbol $R \in \mathcal{L}$ of arity n , and $[(a_i^1)], \dots, [(a_i^n)]$,

$$R^{\prod_{i \in I} \mathcal{M}_i / \mathcal{U}} \ni ([a_i^1]), \dots, [(a_i^n)] \Leftrightarrow \{i \mid R^{\mathcal{M}_i}(a_i^1, \dots, a_i^n)\} \in \mathcal{U}.$$

^aSometimes it's easier to think of (a_i) as a function $a(i)$ from I to $\bigcup M_i$ with $a(i) \in M_i$.

Intuition. Think of the \mathcal{M}_i as voting on what happens in the **ultraproduct**.

There are a number of things that one must check to see that **Definition 6.2.1** makes sense.

Claim. \sim (or $\sim_{\mathcal{U}}$) is indeed an equivalence relation.

Proof. The only thing that is nontrivial is the transitivity. Since $\{i \mid a_i = b_i\} \cap \{i \mid b_i = c_i\} \subseteq \{i \mid a_i = c_i\}$, so from the definition of **ultrafilter**, if the left-hand side in \mathcal{U} ($(a_i) \sim (b_i)$ and $(b_i) \sim (c_i)$), then the right-hand side is in \mathcal{U} as well ($(a_i) \sim (c_i)$), i.e., we have transitivity. \circledast

Claim. $f^{\prod_{i \in I} \mathcal{M}_i / \mathcal{U}}$ is well-defined.

Proof. Let $(a_i^1) \sim (b_i^1), \dots, (a_i^n) \sim (b_i^n)$, i.e., $X_j = \{i \in I \mid a_i^j = b_i^j\} \in \mathcal{U}$. Then $X = X_1 \cap \dots \cap X_n \in \mathcal{U}$. Thus, for $i \in X$, $f^{\mathcal{M}_i}(a_i^1, \dots, a_i^n) = f^{\mathcal{M}_i}(b_i^1, \dots, b_i^n)$, hence

$$[(f^{\mathcal{M}_i}(a_i^1, \dots, a_i^n))_{i \in I}] = [(f^{\mathcal{M}_i}(b_i^1, \dots, b_i^n))_{i \in I}],$$

i.e., $f^{\prod_{i \in I} \mathcal{M}_i / \mathcal{U}}$ is well-defined. \circledast

Note. $R^{\prod_{i \in I} \mathcal{M}_i / \mathcal{U}}$ is well-defined.

Proof. We know that

$$R^{\mathcal{M}}([a_i^1], \dots, [a_i^n]) \Leftrightarrow \{i \mid R^{\mathcal{M}_i}(a_i^1, \dots, a_i^n)\} \in \mathcal{U}$$

and

$$\neg R^{\mathcal{M}}([a_i^1], \dots, [a_i^n]) \Leftrightarrow \{i \mid \neg R^{\mathcal{M}_i}(a_i^1, \dots, a_i^n)\} \in \mathcal{U},$$

and since \mathcal{U} is an **ultrafilter**, exactly one will be the case. \circledast

Lecture 24: Ultrapowers and the Łoś's Theorem

6.3 Ultrapowers

4 Apr. 11:30

Notation. Now, we write $[a_i]$ for $[(a_i)_{i \in I}]$.

Notation (Almost everywhere). When something happens on a set in the [ultrafilter](#), we say that it happens *almost everywhere*.

Example. If $a \sim b$, then a and b are equal [almost everywhere](#).

[Ultraproducts](#) are only interesting when the [ultrafilter](#) is non-[principal](#); otherwise, the [ultraproduct](#) is just isomorphic to one of the facts, i.e., the dictator.

Definition 6.3.1 (Ultrapower). Fix a single [structure](#) \mathcal{M} and take $\mathcal{M}_i = \mathcal{M}$ for every $i \in I$. The *ultrapower* is defined as $\prod_{i \in I} \mathcal{M} / \mathcal{U}$.

Remark. $\mathcal{M} \hookrightarrow \prod \mathcal{M} / \mathcal{U}$ is an [embedding](#).

Proof. Consider the constant sequence embedding $a \mapsto [(a)_{i \in I}]$. ⊗

Let's first see one example.

Example (Non-standard reals). Let \mathcal{U} be a non-[principal ultrafilter](#) on $I = \mathbb{N}$, and let $\mathbb{R}^* = \prod \mathbb{R} / \mathcal{U}$.

Then, we see that $\mathbb{R} \hookrightarrow \mathbb{R}^*$ by $r \mapsto [r]$, i.e., the constant sequence r . This is a field embedding of fields.^a However, \mathbb{R}^* has other elements, e.g., $[(n)_{n \in \mathbb{N}}]$ and $[(1/n)_{n \in \mathbb{N}}]$, and for each $r \in \mathbb{R}$, $[(n)] \geq r = [r]$.^b Similarly, $[(1/n)]$ is infinitesimal, i.e., $0 < [1/n] < r$ for all $r \in \mathbb{R}^+$. It's possible to develop calculus by using infinitesimal elements of $\prod_{i \in \mathbb{N}} \mathbb{R} / \mathcal{U}$ instead of limits.

^aSince, for example, $[r] + [s] = [r + s]$, and $1 \mapsto [1]$ is the multiplicative identity of \mathbb{R}^* .

^bSince $\{n \mid n \geq r\}$ is [cofinite](#) and \mathcal{U} is non-[principal](#), hence in \mathcal{U} . This is what it means to have $[(n)] \geq [r]$ in \mathbb{R}^* .

Example. Let $a_i = 1$ for odd i , or 2 for even i . Then either $[a_i] = 1$ or $[a_i] = 2$ depending on \mathcal{U} .

Example. $[n + 1] = [n] + [1]$, $[n^2] = [n][n]$. We have $[n^2] \geq [n + a] = [n] + a$ for all $a \in \mathbb{N}$.

6.3.1 Łoś's theorem

The power of [ultraproducts](#) is that what is [true](#) in an [ultraproduct](#) is exactly what is [true](#) in almost all coordinates. This is the [Łoś's theorem](#).¹

Theorem 6.3.1 (Łoś's theorem). Let \mathcal{L} be a [language](#), and \mathcal{U} an [ultrafilter](#) on I . And for $i \in I$, we have an \mathcal{L} -[structure](#) \mathcal{A}_i . Then, for each $a_i^1, \dots, a_i^n \in A_i$,

$$\prod_{i \in I} \mathcal{A}_i / \mathcal{U} \models \varphi([a_i^1], \dots, [a_i^n]) \Leftrightarrow \{i \in I \mid \mathcal{A}_i \models \varphi(a_i^1, \dots, a_i^n)\} \in \mathcal{U}.$$

Proof. We do an induction on [formulas](#). For simplicity, assume that the [language](#) is relational by replacing n -ary functions by $n + 1$ -ary relations.^a For [atomic formulas](#) (equality and relations), it's straightforward from the definitions.

¹Pronounced as “wash”.

- Suppose the claim holds for φ and ψ . Then

$$\{i \mid \mathcal{A}_i \models \varphi(a_i^1, \dots, a_i^n)\} \cap \{i \mid \mathcal{A}_i \models \psi(a_i^1, \dots, a_i^n)\} = \{i \mid \mathcal{A}_i \models (\varphi \wedge \psi)(a_i^1, \dots, a_i^n)\}, \quad (6.1)$$

implying that

$$\prod_{i \in I} \mathcal{A}_i / \mathcal{U} \models \varphi(a_i^1, \dots, a_i^n) \text{ and } \prod_{i \in I} \mathcal{A}_i / \mathcal{U} \models \psi(a_i^1, \dots, a_i^n) \Leftrightarrow \prod_{i \in I} \mathcal{A}_i / \mathcal{U} \models (\varphi \wedge \psi)(a_i^1, \dots, a_i^n).$$

And from the induction hypotheses, the left-hand side of Equation 6.1 are both in \mathcal{U} , hence the right-hand side.

- Suppose it holds for φ . Then

$$\begin{aligned} \prod_{i \in I} \mathcal{A}_i / \mathcal{U} \models \neg \varphi(a_i^1, \dots, a_i^n) &\Leftrightarrow \prod_{i \in I} \mathcal{A}_i / \mathcal{U} \not\models \varphi(a_i^1, \dots, a_i^n) \\ &\Leftrightarrow \{i \mid \mathcal{A}_i \models \varphi(a_i^1, \dots, a_i^n)\} \notin \mathcal{U} \\ &\Leftrightarrow \{i \mid \mathcal{A}_i \models \neg \varphi(a_i^1, \dots, a_i^n)\} \in \mathcal{U}. \end{aligned}$$

- Suppose it holds for φ , and suppose $\prod_{i \in I} \mathcal{A}_i / \mathcal{U} \models \exists x \varphi(x, a_i^1, \dots, a_i^n)$. Pick $[b_i]$ such that $\prod_{i \in I} \mathcal{A}_i / \mathcal{U} \models \varphi([b_i], a_i^1, \dots, a_i^n)$. Then by the induction hypotheses,

$$\{i \mid \mathcal{A}_i \models \varphi(b_i, a_i^1, \dots, a_i^n)\} \in \mathcal{U},$$

hence $\{i \mid \mathcal{A} \models \exists x \varphi(x, a_i^1, \dots, a_i^n)\} \in \mathcal{U}$ since it contains $\{i \mid \mathcal{A}_i \models \varphi(b_i, a_i^1, \dots, a_i^n)\}$.

On the other hand, suppose $\{i \mid \mathcal{A}_i \models \exists x \varphi(x, a_i^1, \dots, a_i^n)\} \in \mathcal{U}$. For each i in this set, pick $b_i \in \mathcal{A}_i$ such that $\mathcal{A}_i \models \varphi(b_i, a_i^1, \dots, a_i^n)$. For all other i , pick any b_i . Then,

$$\{i \mid \mathcal{A}_i \models \varphi(b_i, a_i^1, \dots, a_i^n)\} \in \mathcal{U}.$$

By the induction hypotheses, $\prod_{i \in I} \mathcal{A}_i / \mathcal{U} \models \varphi([b_i], a_i^1, \dots, a_i^n)$, so

$$\prod_{i \in I} \mathcal{A}_i / \mathcal{U} \models \exists x \varphi(x, a_i^1, \dots, a_i^n).$$

■

^aIf there are function symbols, one must also do induction on [terms](#).

We then have the following immediate consequences.

Note (Non-standard reals). $\mathcal{M} \hookrightarrow \prod \mathcal{M} / \mathcal{U}$ is an [elementary embedding](#), i.e.,

$$\mathcal{M} \models \varphi(a_1, \dots, a_n) \Leftrightarrow \prod \mathcal{M} / \mathcal{U} \models \varphi([a_1], \dots, [a_n]).$$

Example. $\mathbb{R} \hookrightarrow \mathbb{R}^*$ is [elementary](#), and $\mathbb{R} \equiv \mathbb{R}^*$.^a

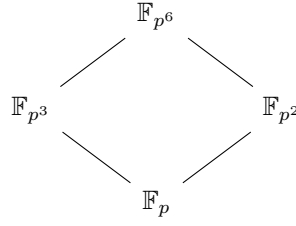
^a \mathbb{R}^* is also a real closed field.

Example (Pseudofinite field). Let \mathcal{U} be a non-principal ultrafilter on I , where I is the set of primes p . Let $F = \prod \mathbb{F}_p / \mathcal{U}$, which is a field from the [Łoś's theorem](#).^a Furthermore, F has [characteristic](#) 0 because for each q , $\{p \mid \mathbb{F}_p \models \underbrace{1 + \dots + 1}_q\} = \{q\} \notin \mathcal{U}$.

F is infinite, and if φ is a [sentence true](#) in any finite field, $\mathbb{F}_p \models \varphi$ for all p , so $F \models \varphi$. We call F a *pseudofinite field*.

^aSince we know that F has every property that \mathbb{F}_p has.

Note. $\mathbb{F}_p^{\text{alg}} = \bigcup \mathbb{F}_{p^n}$.



6.3.2 Another Proof of the Compactness Theorem

One application to [Łoś's theorem](#) is a proof of the [compactness theorem](#) without passing through the [completeness theorem](#).

As previously seen (Compactness theorem). Suppose that T is a finitely [satisfiable theory](#). Then T is [satisfiable](#).

Another proof of Theorem 2.6.2. Let T be an infinite \mathcal{L} -theory which is finitely [satisfiable](#). Let I be the collection of finite subsets of T . Then, for each $\Delta \in I$, let $X_\Delta = \{\Lambda \in I \mid \Delta \subseteq \Lambda\}$. Let

$$D = \{Y \subseteq I \mid X_\Delta \subseteq Y \text{ for some } \Delta \in I\}.$$

Claim. D is a [filter](#).

Proof. We see that

- (a) $\emptyset \notin D$ since each $X_\Delta \neq \emptyset$; and $I \in D$ since $X_\Delta \subseteq I$ for any Δ ;
- (b) if $Y \in D$ and $Z \supseteq Y$, then $X_\Delta \subseteq Y$ implies $X_\Delta \subseteq Z$ for some Δ ;
- (c) if $Y \in D$ and $Z \in D$, then there are Δ, Λ such that $X_\Delta \subseteq Y$ and $X_\Lambda \subseteq Z$. Then

$$X_{\Delta \cup \Lambda} = X_\Delta \cap X_\Lambda \subseteq Y \cap Z,$$

so $Y \cap Z \in D$ for $\Delta \cup \Lambda$.

⊗

Let \mathcal{U} be an [ultrafilter](#) on I extending D . For each $\Delta \in I$, choose $\mathcal{M}_\Delta \models \Delta$ since T is finitely [satisfiable](#). Let $\mathcal{M} = \prod_{\Delta \in I} \mathcal{M}_\Delta / \mathcal{U}$.

Claim. $\mathcal{M} \models T$.

Proof. Let $\varphi \in T$, then $X_{\{\varphi\}} \in D \subseteq \mathcal{U}$ and $\{\Delta \in I \mid \mathcal{M}_\Delta \models \varphi\} \supseteq \{\Delta \in I \mid \Delta \ni \varphi\} = X_{\{\varphi\}} \in \mathcal{U}$. From [Łoś's theorem](#),

$$\{\Delta \in I \mid \mathcal{M}_\Delta \models \varphi\} \in \mathcal{U} \Leftrightarrow \mathcal{M} \models \varphi,$$

so $\mathcal{M} \models \varphi$, hence $\mathcal{M} \models T$.

⊗

■

Intuition.

Fill this up

If \mathcal{L} is countable, we could list the [sentences](#) $\varphi_1, \varphi_2, \dots$ of T . Then, we pick $\mathcal{M}_i \models \varphi_1 \wedge \dots \wedge \varphi_i$. If \mathcal{U} is any non-[principal ultrafilter](#), $\prod \mathcal{M}_i / \mathcal{U} \models \varphi_i$ for all i .

Lecture 25: Applications of Ultrapowers in Algebra

6.3.3 Applications in Algebra

6 Apr. 11:30

What else are [ultraproducts](#) good for? On the homework, we will see *saturation*, that's one. Moreover, we can prove the [Keisler-Shelah theorem](#)

Theorem 6.3.2 (Keisler-Shelah theorem). Two \mathcal{L} -structures are [elementarily equivalent](#) if and only if they have some^a [isomorphic ultrapowers](#).

^aThere exists an [ultrafilter](#) on I such that $\prod \mathcal{M} / \mathcal{U} \cong \prod \mathcal{N} / \mathcal{U}$.

Proof. By [Łoś's theorem](#), $\mathcal{M} \hookrightarrow \prod \mathcal{M} / \mathcal{U}$ is an [elementary embedding](#). To prove the other direction, it'll involve choosing index set and [ultrapower](#), which is beyond our scope. ■

The last, perhaps surprising, we can prove bounds in algebra, and we will focus on this one. Consider the following.

Problem 6.3.1. If K is a field, $f_1, \dots, f_n, g \in K[x_1, \dots, x_n]$ have degree $\leq d$. If $g \in (f_1, \dots, f_n)$, i.e., $g = h_1 f_1 + \dots + h_n f_n$, is there some number $m = m(d, n)$ such that we can choose h_1, \dots, h_n of degree $\leq m$?

To set up [Problem 6.3.1](#), we let

- K_1, K_2, \dots be fields;
- $\bar{x} = \{x_1, \dots, x_n\}$;
- \mathcal{U} be a non-[principal ultrafilter](#) on \mathbb{N} ;
- $K = \prod K_i / \mathcal{U}$ (which is a field);
- $K[\bar{x}] = K[x_1, \dots, x_n]$ the polynomial ring over K ;
- $R = \prod K_i[x_1, \dots, x_n] / \mathcal{U}$, which is a ring.

To understand what we're doing, observe that elements of $K[\bar{x}]$ look like

$$\sum_{\text{monomial } M} [c_M(i)]M$$

where M are monomials $x_1^{r_1} \dots x_n^{r_n}$, and each $c_M(i) \in K_i$; similarly, elements of R look like

$$\left[\sum_{\text{monomial } M} c_M(i)M \right]$$

where each $\sum_M c_M(i)M \in K_i[\bar{x}]$.

Example. $[x_1^i] \in R$.

Claim. $K[\bar{x}]$ is a subring of R .

Proof. $\sum_M [c_M(i)]M \in K[\bar{x}]$ is identified with $[\sum_M c_M(i)M]$.

Note. $[x_1^i]$ is not in the image, i.e., not in $K[\bar{x}]$.

Let $\mathbb{N}^* = \prod \mathbb{N} / \mathcal{U}$, and define the *total degree* $\deg_i: K_i[\bar{x}] \rightarrow \mathbb{N}$ such that $\deg_i(x_1^{r_1} \cdot x_2^{r_2} \dots x_n^{r_n}) = r_1 + r_2 + \dots + r_n$. This induces

$$\deg^*: R \rightarrow \mathbb{N}^*, \quad [f_i] \mapsto [\deg_i(f_i)].$$

Example. $\deg^*([x_1^i]) = [i] \in \mathbb{N}^* - \mathbb{N}$.^a

^aRememberer that $\mathbb{N} \succeq \mathbb{N}^*$.

Claim. $K[\bar{x}] = \{f \in R \mid \deg^*(f) \in \mathbb{N}\}$.

Proof. We see that $K[\bar{x}] \subseteq \{f \in R \mid \deg^*(f) \in \mathbb{N}\}$ since if $f \in K[\bar{x}]$, $\deg^*(f) = \deg(f)$. To see this, let $f = \sum_M [c_M(i)]M$, if $[c_M(i)] \neq 0$, then there is a set $A_M \in \mathcal{U}$ such that for $i \in A_M$, $c_M(i) \neq 0$. Let $A = \bigcap_{[c_M(i)] \neq 0} A_M \in \mathcal{U}$, then

$$\deg_i \left(\sum_M c_M(i)M \right) = \deg(f)$$

for $i \in A$.

To show the other containment, let $f = [\sum_M c_M(i)M] \in R$ have $\deg^*(f) \in \mathbb{N}$, say d . So there is $A \in \mathcal{U}$ such that for $i \in A$, $\sum_M c_M(i)M$ has $\deg_i = d$. Then, let $g \in K[\bar{x}]$ be

$$\sum_{M: \deg(M) \leq d} [c_M(i)]M.$$

Then, $g \mapsto f$ under $K[\bar{x}] \hookrightarrow R$. ⊗

Intuition. Consider $a_1x \in K_1[x]$, $a_2x \in K_2[x]$, $a_3x \in K_3[x]$, etc., then we have $[a_i]x$. ⊗

To see why we introduce all these (especially R) for [Problem 6.3.1](#), [Theorem 6.3.3](#) states that if we can relate K and R as in [Theorem 6.3.3 \(b\)](#), then [Problem 6.3.1](#) is solved as stated in [Theorem 6.3.3 \(a\)](#).

Theorem 6.3.3. The following are equivalent.

- (a) There is $m = m(d, n)$ such that for any field K , and polynomials $f_1, \dots, f_\ell, g \in K[x_1, \dots, x_n]$ of $\deg \leq d$, if $g \in (f_1, \dots, f_\ell)$, then there are $h_1, \dots, h_\ell \in K[x_1, \dots, x_n]$ of $\deg \leq m$ such that $g = h_1f_1 + \dots + h_\ell f_\ell$.
- (b) For any fields K_i , $K = \prod K_i / \mathcal{U}$, $R = \prod K_i[\bar{x}] / \mathcal{U}$, if I is an [ideal](#) of $K[\bar{x}]$, then $IR \cap K[\bar{x}] = I$.^a

^a $IR = \{a_1r_1 + \dots + a_\ell r_\ell \mid a_i \in I, r_i \in R\}$. If $I = (f_1, \dots, f_\ell)$ in $K[\bar{x}]$, then $IR = \{r_1f_1 + \dots + r_\ell f_\ell \mid r_i \in R\}$, i.e., the [ideal](#) generated by I in R .

Proof. First, we show that (b) implies (a). Suppose not, fix d, n such that there is no such m . For each $m = 1, 2, \dots$, there is a field K_m , polynomials $f_1^m, \dots, f_\ell^m, g^m \in K_m[x_1, \dots, x_n]$ of $\deg \leq d$ such that $g^m \in (f_1^m, \dots, f_\ell^m)$ but there are no h_1^m, \dots, h_ℓ^m of $\deg \leq m$ with $g^m = h_1^m f_1^m + \dots + h_\ell^m f_\ell^m$.

Remark. Since m doesn't depend on ℓ , and $\deg \leq d$ implies that we can always assume the same ℓ .

Do the "setup" to K_m , let $K = \prod K_m / \mathcal{U}$, etc., and let $f_1 = [f_1^m] \in R, \dots, f_\ell = [f_\ell^m] \in R, g = [g^m] \in R$. By the fact that they have $\deg \leq d$, $f_1, \dots, f_\ell, g \in K[\bar{x}]$. Now, let $I = (f_1, \dots, f_\ell) \subseteq K[\bar{x}]$.

Claim. $g \notin I$.

Proof. If $g \in I$, there are $h_1 = [h_1^m], \dots, h_\ell = [h_\ell^m] \in K[\bar{x}]$ such that $g = h_1f_1 + \dots + h_\ell f_\ell$. So for most m , $g^m = h_1^m f_1^m + \dots + h_\ell^m f_\ell^m$, and $\deg_m(h_i^m) = \deg^*(h_i) \in \mathbb{N}$. Let m larger than all of these degrees such that the above holds. Since \mathcal{U} is non-principal, but we chose $g^m, f_1^m, \dots, f_\ell^m$ so that this didn't happen, a contradiction. ⊗

Claim. $g \in IR$.

Proof. For each m , $g^m \in (f_1^m, \dots, f_\ell^m)$, so choose h_1^m, \dots, h_ℓ^m such that

$$g^m = h_1^m f_1^m + \dots + h_\ell^m f_\ell^m.$$

Let $h_i = [h_i^m]$, then $g = h_1 f_1 + \dots + h_\ell f_\ell$ in R .^a So $g \in IR$. ⊗

^aNote that $h_i \in R$, but maybe not in $K[\bar{x}]$.

Hence, $g \in IR \cap K[\bar{x}]$, but $g \notin I$, contradicts (b), hence (b) implies (a). The other direction is in the same spirit, we basically just reverse the above argument. ■

Notation (Faithfully flat). R is *faithfully flat* over $K[\bar{x}]$ if $IR \cap K[\bar{x}] = I$.

Intuition. R being *faithfully flat* says that if $g \in K[\bar{x}]$ has $g = r_1 f_1 + \dots + r_\ell f_\ell$, $r_i \in R$, then $g = h_1 f_1 + \dots + h_\ell f_\ell$ for some $h_i \in K[\bar{x}]$.

Hence, to show Theorem 6.3.3 (a), we show Theorem 6.3.3 (b). But this is somewhat algebraically involved, so we first do something easier. Recall that *weak Hilbert's Nullstellensatz*.

Theorem 6.3.4 (Nullstellensatz). Let K be *algebraically closed*, and $g, f_1, \dots, f_n \in K[\bar{x}]$. Then every common zero of f_1, \dots, f_n is a zero of g if and only if $g^m = (f_1, \dots, f_n)$ for some m .

Proof. We use the so-called *Rabinowitsch trick*, i.e., the *weak* case in fact implies the *strong* case.

To prove the forward direction, suppose that every common zero of f 's is a zero of g . Then, $f_1, \dots, f_\ell, 1 - x_0 g$ has no common zero in $K[x_0, x_1, \dots, x_n]$. By *weak Hilbert's Nullstellensatz*,

$$1 = h_1 f_1 + \dots + h_\ell f_\ell + h_0 (1 - x_0 g)$$

for $h_i \in K[x_0, x_1, \dots, x_n]$. Substitute $x_0 = 1/g$, then

$$1 = h_0 (1/g, \dots) (1 - 1) + h_1 (1/g, \dots) + h_1 (1/g, \dots) f_1 + \dots + h_\ell (1/g, \dots) f_\ell.$$

Multiplying by g^m for some large m to clear the denominators, we have

$$g^m = \underbrace{[g^m h_1 (1/g, \dots)]}_{\in K[x_1, \dots, x_n]} f_1 + \dots + \underbrace{[g^m h_\ell (1/g, \dots)]}_{\in K[x_1, \dots, x_n]} f_\ell,$$

so $g^m \in (f_1, \dots, f_\ell)$. ■

Note. The *weak Hilbert's Nullstellensatz* is the case that $g = 1$.

Lecture 26: A Detour to Bounded Version of Nullstellensatz

As previously seen. The setup is:

11 Apr. 11:30

- \mathcal{U} be a non-principal ultrafilter on \mathbb{N} ;
- K_1, K_2, \dots be a sequence of fields;
- $R = \prod K_i[x_1, \dots, x_n] / \mathcal{U}$, which is a ring;
- $K[x_1, \dots, x_n] \hookrightarrow R$, with image $\{f \in R : \deg^*(f) \in \mathbb{N}\}$, where $\deg^* : R \rightarrow \mathbb{N}^* = \prod \mathbb{N} / \mathcal{U}$.

We're toward proving Theorem 6.3.3 (b). However, we should further develop some algebraic intuition to understand the setup.

Theorem 6.3.5. Given n and d , there is $m = m(d, n)$ such that for any algebraically closed field F and $f, g_1, \dots, g_\ell \in F[x_1, \dots, x_n]$ with $\deg \leq d$, if every common zero of g_1, \dots, g_ℓ is a zero of f , then $f^m \in (g_1, \dots, g_\ell)$.

Proof. Suppose that for a fixed n, d , no such m exists. So for each m , there are an algebraically closed field K_m and $f_m, g_{1,m}, \dots, g_{\ell,m}$ in $K_m[x_1, \dots, x_n]$ with degree at most d , such that every common zero of $g_{1,m}, \dots, g_{\ell,m}$ is a zero of f_m , but

$$f_m^m \notin (g_{1,m}, \dots, g_{\ell,m}).$$

Form $f = [f_m]$ and $g_i = [g_{i,m}]$, which are elements of R , but actually in $K[x_1, \dots, x_n]$ since $\deg^*(f) \leq d$ and $\deg^*(g_i) \leq d$. We want to apply the Nullstellensatz in $K[x_1, \dots, x_n]$ to f and g_1, \dots, g_ℓ . So we want to know that every common zero $a = [a_i] \in K$ of g_1, \dots, g_ℓ is a zero of f . The set $\{m \mid g_{1,m}(a_m) = \dots = g_{\ell,m}(a_m) = 0\}$ is in the ultrafilter.^a For each m in the set, a_m is a common zero of $g_{1,m}, \dots, g_{\ell,m}$, and so it is a zero of f_m , i.e., $f_m(a_m) = 0$. This is true for almost every m , so $f(a) = 0$ by Łoś's theorem. So there is some $t \in \mathbb{N}$ such that

$$f^t \in (g_1, \dots, g_\ell).$$

By Łoś's theorem, for almost every m , $f_m^t \in (g_{1,m}, \dots, g_{\ell,m})$ in $K_m[x_1, \dots, x_n]$. There is $m > t$ for which this is true, then

$$f_m^t \in (g_{1,m}, \dots, g_{\ell,m}) \rightarrow f_m^m = f_m^t \cdot f_m^{m-t} \in (g_{1,m}, \dots, g_{\ell,m}),$$

which contradicts how we choose everything at the beginning (e.g., m). ■

^aSince Łoś's theorem says $g_i(a) = 0$ if and only if $\{m \mid g_{i,m}(a_m) = 0\} \in \mathcal{U}$.

Intuition. This is a bounded version of the Nullstellensatz.

Note. For some u_m , $f_m^{u_m} \in (g_{1,m}, \dots, g_{\ell,m})$. But $u = [u_m] \in \mathbb{N}^*$ and $u \notin \mathbb{N}$, i.e., $f^u \in (g_1, \dots, g_\ell)$.

In some sense, the solutions are kind of “fundamental”.

Theorem 6.3.6. If $f_1, \dots, f_\ell \in K[x_1, \dots, x_n]$, then any solution of

$$f_1 y_1 + \dots + f_\ell y_\ell = 0$$

in R is an R -linear combination of solutions in $K[x_1, \dots, x_n]$.^a

^aIn this case, R is called a flat $K[x_1, \dots, x_n]$ -module.

Proof. Let $g = (g_1, \dots, g_\ell)$ be a solution in R of $f_1 y_1 + \dots + f_\ell y_\ell = 0$, we want to show that g is an R -linear combination of solutions in $K[x_1, \dots, x_n]$. Without loss of generality, assume that $f_1 \neq 0$, and we want to know that f_1 is monic in x_n , i.e.,

$$f_1 = x_n^d + g_{d-1} x_n^{d-1} + \dots + g_0,$$

where $g_i \in K[x_1, \dots, x_{n-1}]$. To get this, we need to do a change of variables such that

$$\begin{aligned} x_1 &\mapsto x_1 + x_n^t \\ x_2 &\mapsto x_2 + x_n^t \\ &\vdots \\ x_n &\mapsto x_n. \end{aligned}$$

For example, $x_1 x_2 x_3 \mapsto \dots + x_n^{3t}$, where we can choose a large enough t to make f_1 into f'_1 , monic.

Note. We need to make sure that we can go backwards. We will take care of this later.

$f_1 y_1 + f_\ell y_\ell$ has solutions $(-f_2, f_1, 0, \dots), (-f_3, 0, f_1, 0, \dots)$, and so on. By long division in R , there is h, r such that $g_2 = f_1 h + r$ where $\deg^*(r) < \deg^*(f_1)$. Either by [Los's theorem^a](#) or for each i , $g_2^i = f_1^i h^i + r^i$ where $\deg_{x_n}(r^i) < \deg_{x_n}(f_1^i)$ by long division in $K_1[x_1, \dots][x_n]$. Now take $(g_1, \dots, g_\ell) - h \cdot (-f_2, f_1, 0, \dots) = (g'_1, g'_2, \dots)$ is a solution of our homogeneous equation. Note that $g'_2 = r$, so $\deg_{x_n}^*(g'_2) < \deg_{x_n}^*(f_1)$, which is finite. Moreover, g'_3, g'_4, \dots did not change.

Keep doing this with $(-f_3, 0, f_1, 0, \dots)$ and so on, then we get a solution $g' = (g'_1, \dots, g'_\ell)^b$ with $\deg_{x_n}^*(g'_i) \in \mathbb{N}$ for $i > 1$. Since $f_1 g'_1 + \dots + f_\ell g'_\ell$, we know that

$$f_1 g'_1 = -(f_2 g'_2 + \dots + f_\ell g'_\ell),$$

where the right-hand side is finite degree in x_n , and f_1 also has finite degree in x_n , we see that $\deg_{x_n}^*(g'_1)$ is finite. So g' is a solution in $(\prod K_i[x_1, \dots, x_{n-1}] / \mathcal{U})[x_n]$. Furthermore, g is an R -linear combination of g' and solution in $K[x_1, \dots, x_n]$.

Now, repeat this but with g' and x_{n-1} . But we have to make sure everything has finite degree in x_n still, i.e., we use $(\prod K_i[x_1, \dots, x_{n-1}] / \mathcal{U})[x_n]$ instead of R . Eventually, g is a linear combination of solutions in $K[x_1, \dots, x_n]$ after n changes of variables. Finally, we reverse the changes of variables. ■

^aTo do this, we need to consider $(K_i[x_1, \dots, x_n], \mathbb{N}, +, -, \cdot, <, \deg)$.

^bWe have abused the notation here.

Again, a bounded version of [Theorem 6.3.6](#) exists, just as the case of [Nullstellensatz](#).

Theorem 6.3.7. Given n and d , there is an $m = m(n, d)$ such that for any $f_1, \dots, f_\ell \in K[x_1, \dots, x_n]$ of $\deg \leq d$, any solution of

$$f_1 y_1 + \dots + f_\ell y_\ell = 0$$

is a $K[x_1, \dots, x_n]$ -linear combination of solutions of $\deg \leq m$.

Proof. This follows from the same idea as [Theorem 6.3.6](#). ■

Next, we think about

$$f_1 y_1 + \dots + f_\ell y_\ell = g,$$

i.e., we are asking whether $g \in (f_1, \dots, f_\ell)$.

Lecture 27: Proving the Algebraic Bounds

As previously seen (Goal). If $I \subseteq K[x_1, \dots, x_n]$ then $IR \cap K[x_1, \dots, x_n] = I$.

13 Apr. 11:30

Theorem 6.3.8. If m is a [maximal ideal](#) of $K[x_1, \dots, x_n]$, then $mR \neq R$.

Proof. Let $m = (f_1, \dots, f_\ell)$. Then f_1, \dots, f_ℓ have a common zero in an [extension](#) of K , namely, $K[m_1, \dots, m_n] / m$. So, f_1, \dots, f_ℓ have a zero in any [algebraic closure](#) of K .

Now if $mR = R$, $1 = h_1 f_1 + \dots + h_\ell f_\ell$ where h_1, \dots, h_ℓ in R . So for almost every i , $1 = h_1^i f_1^i + \dots + h_\ell^i f_\ell^i$ in $K_i[x_1, \dots, x_n]$. Then f_1^i, \dots, f_ℓ^i can't have a common zero in any [extension](#) of K_i , including in \overline{K}_i , the [algebraic closure](#) of K_i .

Let $\overline{K} = \prod \overline{K}_i / \mathcal{U}$. This is an [algebraically closed](#) field containing K . Then, f_1, \dots, f_ℓ have a common zero $\overline{a} = [\overline{a}^i]$ in \overline{K} , i.e., $f_1(\overline{a}) = \dots = f_\ell(\overline{a}) = 0$. So for almost every i , $f_1^i(\overline{a}^i) = \dots = f_\ell^i(\overline{a}^i) = 0$ in \overline{K}_i . Hence, for almost every i , f_1^i, \dots, f_ℓ^i both have and don't have a common zero, a contradiction, implying $mR \neq R$. ■

Theorem 6.3.9. Let A be a subring of B such that

- (a) if $f_1, \dots, f_\ell \in A$, any solution in B of $f_1 y_1 + \dots + f_\ell y_\ell = 0$ is a linear combination of solutions in A ;
- (b) if m a [maximal ideal](#) of A , then $mB \neq B$,

then if I is an **ideal** of A , $IB \cap A = I$.

Proof. Consider the inclusion map $A/I \rightarrow B/IB$, which is well-defined since $I \subseteq IB$. If this is injective, then $I = IB \cap A$.

Take $x \in A$, $x \notin I$. Consider $f: A \rightarrow A/I$ such that $r \mapsto rx$. Observe that $f(1) \notin I$, i.e., $1 \notin \ker f$. Let $J = \{a \in A \mid f(a) = 0\} = \{a \in A \mid ax \in I\}$, then $1 \notin J$ since $x \notin I$. Moreover, J is an **ideal** of A , so $J \subsetneq A$. Then, $JB \neq B$ because $J \subseteq m$ a **maximal ideal** of A , and $mB \neq B$.

Now, consider the following, where we want to show that $B/BJ \rightarrow Bx/BI$ defined by $r \mapsto rx$ is a bijection.

$$\begin{array}{ccc} A/J & \hookrightarrow & Ax/I \\ \downarrow & & \downarrow \\ B/JB & \hookrightarrow & Bx/IB \end{array}$$

where $Ax/I = \{rx \mid r \in A\}/I$. Note that B/BJ is non-trivial since $B \neq JB$. We need to show that this is well-defined, i.e., $JB \subseteq IB$. If $r \in BJ$, $r = b_1j_1 + \dots + b_nj_n$ for $j_1, \dots, j_n \in J$, then $rx = b_1j_1x + \dots + b_nj_nx$ where we know that $j_ix \in I$ for all i , so $rx \in IB$.

This is also injective. Suppose $r \in B$ is such that $rx \in IB$, we want to show that $r \in JB$. Since $rx \in IB$, there are $f_1, \dots, f_\ell \in I$ and $h_1, \dots, h_\ell \in B$ with

$$rx = f_1h_1 + \dots + f_\ell h_\ell.$$

Think of (r, h_1, \dots, h_ℓ) as a solution of $-y_0x + y_1f_1 + \dots + y_\ell f_\ell = 0$, which is a homogeneous equation over A . So (r, h_1, \dots, h_ℓ) is a B -linear combination of $(r^i, h_1^i, \dots, h_\ell^i) \in A^{\ell+1}$, which are solution of the same equation. For each i , $r^i x = f_1h_1^i + \dots + h_\ell h_\ell^i \in I$. Since $r^i x \in I$, $r^i \in J$. Then $r = \sum_i g_i r^i$ where $g_i \in B$, so $r \in JB$. This shows that the diagram does commute.

Finally, since $JB \neq B$ and $B/JB \cong Bx/IB$, $x \notin IB$. Recall that $x \in A$ was an arbitrary element $\notin I$, so $A/I \rightarrow B/BI$ is injective. ■

From [Theorem 6.3.9](#), take $A = K[x_1, \dots, x_n]$ and $B = R$, we have the following.

Corollary 6.3.1. If $I \subseteq K[x_1, \dots, x_n]$ then $IR \cap K[x_1, \dots, x_n] = I$.

This is exactly [Theorem 6.3.3 \(b\)](#)! From [Theorem 6.3.3](#), we can finally put [Theorem 6.3.3 \(a\)](#) as a theorem.

Theorem 6.3.10. Given n and d , there is $m = m(n, d)$ such that for any field K and $f_1, \dots, f_\ell, g \in K[x_1, \dots, x_n]$ of $\deg \leq d$, if $g \in (f_1, \dots, f_\ell)$, then there are $h_1, \dots, h_\ell \in K[x_1, \dots, x_n]$ of $\deg \leq m$ such that $g = h_1f_1 + \dots + h_\ell f_\ell$.

This solves [Problem 6.3.1](#).

Problem. We've proved these bounds exist, but can we compute them?

Answer. Yes! This can be done by

- (a) an algebraic proof, or
- (b) *proof mining*: if we look very carefully at the [ultraproduct](#) proof, we can find bounds.

⊛

Remark. If we can compute bounds, we can solve these problems using linear algebra.

Chapter 7

Epilogue

7.1 An Review: Map of the Universe

To wrap up this course, we review what we have learned in this course, especially various [models](#). Specifically, we have a glance on [forkinganddividing](#), and discuss some [models](#).

Lecture 28: An Outlook of Real Closed Fields

7.2 An Outlook of Real Closed Fields

18 Apr. 11:30

Recall the [theory](#) of [real-closed fields](#) $\text{RCF} = \text{Th}(\mathbb{R}, 0, 1, +, -, \cdot, \leq)$ where RCF can be axiomatized by

- field axioms;
- ordered field axioms;
- -1 is not a sum of squares;
- for all x , either x or $-x$ is a square;
- every polynomial of odd degree has a root.

Remark. If we omit the ordered field axioms, we get $\text{Th}(\mathbb{R}, 0, 1, +, -, \cdot)$.

Theorem 7.2.1. RCF admits [quantifier elimination](#).

On the other hand, $\text{Th}(\mathbb{R}, 0, 1, +, -, \cdot)$ does not.

Example. $\exists z \ y - x = z^2 \Leftrightarrow y \geq x$ is not equivalent to a [quantifier free formula](#) in the algebraic language.

7.2.1 Model Complete

However, $\text{Th}(\mathbb{R}, 0, 1, +, -, \cdot)$ is [model complete](#).

Definition 7.2.1 (Model complete). A [theory](#) is *model complete* if every [formula](#) is equivalent to an [existential formula](#) (and a [universal formula](#)).

The last one can be seen by, for example,

$$\exists z \ y - x = z^2 \Leftrightarrow y \geq x \Leftrightarrow x = y \vee \neg x \geq y \Leftrightarrow x = y \vee \neg(\exists z \ x - y = z^2).$$

Lemma 7.2.1. T is **model complete** if and only if whenever $\mathcal{M}, \mathcal{N} \models T$, $\mathcal{M} \subseteq \mathcal{N}$, then $\mathcal{M} \preceq \mathcal{N}$.

7.2.2 Semialgebraic Sets

Theorem 7.2.1 says that the **definable** sets in a **model** of RCF are exactly the boolean combinations of sets of the form $\{\bar{x} \mid p(\bar{x}) > 0\}$.

Definition 7.2.2 (Semialgebraic). These are the *semialgebraic sets*.

Corollary 7.2.1. If $F \models \text{RCF}$, and $A \subseteq F^n$ is **semialgebraic**, then the topological closure of A is **semialgebraic**.

Abraham Robinson's simple proof of **Hilbert's 17th problem** (originally solved by Martin):

Problem 7.2.1 (Hilbert's 17th problem). Let F be a **real-closed field**, $f(\bar{x}) \in F(x_1, \dots, x_n)$ be a rational function. If $f(\bar{a}) \geq 0$ for all $\bar{a} \in F$, then $f(\bar{x})$ is a sum of squares of rational functions.

Proof. To prove this, we use **model completeness**. ■

Example. $(x - y^2)^2 + (x^2y + z - z^5)^2$.

Theorem 7.2.2. RCF has definable **Skolem functions**: Given $\varphi(\bar{x}, \bar{y})$, there is a definable function $f(\bar{y})$ such that

$$\exists \bar{x} \varphi(\bar{x}, \bar{y}) \rightarrow \varphi(f(\bar{y}), \bar{y}),$$

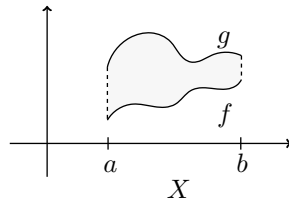
or equivalently, if $\{\bar{a} \mid \mathcal{M} \models \varphi(\bar{a}, \bar{b})\} \neq \emptyset$, then $f(\bar{b}) \in \{\bar{a} \mid \mathcal{M} \models \varphi(\bar{a}, \bar{b})\}$.

Example. $\varphi(x, y, z) := y < x < z$, so $X_{b,c} = \{a \mid \mathcal{M} \models \varphi(a, b, c)\} = (b, c)$ for $f(y, z) = (y + z)/2$.

Definition 7.2.3 (Cell). A *cell* is defined inductively as

- single points are 0-cells;
- intervals $(a, b) \subseteq F$ are 1-cells, $a, b \in F \cup \{\pm\infty\}$;
- if $X \subseteq F^n$ is an **m-cell**, and $f: X \rightarrow F$ is a continuous definable function, then $Y = \{(\bar{x}, f(\bar{x})) \mid \bar{x} \in X\}$ is an **m-cell**;
- if $X \subseteq F^n$ is an **m-cell**, $f, g: X \rightarrow F$ are continuous definable functions^a with $f(x) < g(x)$, then $Y = \{(\bar{x}, y) \mid \bar{x} \in X, f(\bar{x}) < y < g(\bar{x})\}$ is an **m + 1-cell**.

^aOr $g = +\infty$ or $f = -\infty$.



Theorem 7.2.3. Every **definable** set $X \subseteq F^n$ for $F \models \text{RCF}$ is a finite union of **cells**.

Example. $0 \leq x \leq 1 \wedge 0 \leq y \leq 1 \wedge 0 \leq z \leq 1$ defines a cube.

Example. $x^2 + y^2 \leq 1$ defines a circle.

Remark. $\text{Th}(\mathbb{R}_{\text{exp}}) := \text{Th}(\mathbb{R}, 0, 1, +, -, \cdot, \leq, \exp)$ is also o -minimal.

Remark. $\text{Th}(\mathbb{R}, 0, 1, +, -, \cdot, \leq, \sin, \cos)$ is not o -minimal.

Proof. Consider $\{x \mid \sin(x) = 0\}$, which is not a finite union of points and intervals. ⊛

Remark. $\text{Th}(\mathbb{R}, 0, 1, +, -, \cdot, \leq, \sin \upharpoonright_{[0,2\pi]}, \cos \upharpoonright_{[0,2\pi]})$ is o -minimal.

Definition 7.2.4 (Height). The *height* between a, b coprime is defined as $\text{height}(a/b) := \max(a, b)$.

Theorem 7.2.4 (Pila-Wilke). Let $X \subseteq \mathbb{R}^n$ be *definable* in an o -minimal expansion of \mathbb{R} . Let X^{alg} is the union of all infinite *semialgebraic* subsets of X , and $X^{\text{tr}} = X - X^{\text{alg}}$. Moreover, let $N(X^{\text{tr}}, T)$ be the number of rational points on X^{tr} of *height* at most T . Then for all ϵ , there is a c such that for all T ,

$$N(X^{\text{tr}}, T) \leq cT^\epsilon.$$

Intuition. Few rational points on the transcendental part of X .

Example. $X = \{(x, y, z) \in \mathbb{R}^3 \mid 1 < x, y < 2, z = x^y\}$. For q rational, $1 < q < 2$,

$$X_q = \{(x, q, z) \in \mathbb{R}^3 \mid 1 < x < 2, z = x^q\}$$

is *definable* purely algebraically. For example, $z = x^{a/b} \Leftrightarrow z = \sqrt[b]{x \cdots x} \Leftrightarrow \underbrace{z \cdots z}_{b \text{ times}} = \underbrace{x \cdots x}_{a \text{ times}}$. Then $X^{\text{alg}} = \bigcup_q X_q$. X^{tr} has no rational points.

For now, let T a *complete \mathcal{L} -theory*.

Notation. An n -*type* is a *type* where $\bar{x} = (x_1, \dots, x_n)$.

If $\mathcal{M} \models T$, $A \subseteq M$, we can also talk about n -types over A , i.e., set of *formulas* using A which are finitely realizable in \mathcal{M} (or actually realizable in some $\mathcal{N} \succeq \mathcal{M}$).

Example. Consider DLO (\mathbb{Q}, \leq) . The *1-types* over \mathbb{Q} are “cuts” in \mathbb{Q} .

Definition 7.2.5 (Complete). A *type* p is *complete* if for all φ , $\varphi \in p$ or $\neg\varphi \in p$.

Definition 7.2.6 (Isolated). A *type* $p(\bar{x})$ is *isolated* if there is a *formula* $\varphi(\bar{x})$ such that

- (a) $\varphi(\bar{x}) \cup T$ is *consistent* implies $T \models \exists \bar{x} \varphi(\bar{x})$;
- (b) $T \cup \{\varphi(\bar{x})\} \models p(\bar{x})$.

If $p(\bar{x})$ is *complete*, then $\varphi(\bar{x}) \in p(\bar{x})$.

Example. $\text{tp}(\pi/\mathbb{Q})$ is non-*isolated*.

Example. $\text{tp}(2/\mathbb{Q})$ is *isolated*.

An *isolated type* must be realized in all *models*.

Theorem 7.2.5 (Omitting types). Let \mathcal{L} be countable, T be complete, $p_1(\bar{x}), p_2(\bar{x})$ countably many non-isolated types. Then there is $\mathcal{M} \models T$ that omits^a each p_i .

^aOmit means that there is no $\bar{a} \in M$ satisfies p_i .

Example. Consider $\mathcal{L} = \{R_1, R_2, \dots\}$ be all unary. Then T says any finite combinations of the R_i is possible.

Definition 7.2.7 (Atomic). A model \mathcal{M} is *atomic* of T if \mathcal{M} only realizes isolated type.

Definition 7.2.8 (Prime). A model \mathcal{M} is *prime* of T if for all $\mathcal{N} \models T$, $\mathcal{M} \hookrightarrow \mathcal{N}$ elementary.

Theorem 7.2.6. Being atomic is equivalent to being prime.

Countably many types implies that atomic model exists. There might also be saturated models which realizes lots of types.

Remark. There is no T with only two countable models.

Proof. Suppose not, T has two models \mathcal{M} and \mathcal{N} . Then \mathcal{M} is atomic/prime, and \mathcal{N} is saturated, i.e., \mathcal{N} realizes some non-isolated types $\mathcal{N} \models p(\bar{a})$. $\text{Th}(\mathcal{N}, \bar{a}, M)$ is not countably categorical. Take \mathcal{A} prime/atomic “over \bar{a} ”, argued that $\mathcal{A} \not\cong \mathcal{N}$, $\mathcal{A} \not\cong \mathcal{M}$. *

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).