Mathematical Logic
Fall 2004
Professor R. Moosa

Chris Almost

Contents

1 Introduction

Mathematical Logic is the study of the type of reasoning done by mathematicians. (i.e. proofs, as opposed to observation) Axioms are the first unprovable laws. They are statements about certain "basic concepts" (undefined first concepts). There is usually some sort of "soft" justification for believing these first principles. After this, we define derived concepts and prove theorems about the concepts. All of this together is called an axiom system. Mathematical logic is really the study of axiom systems.

We must formalize mathematically the notions of statements, proofs, structure, and truth in a given structure, with respect to a given language. Once we have formalized these notions, we may prove theorems about them. Logic is really metamathematics. Often the study of the study of a field contributes back to the field itself, and this is the case with logic. We often discover connections to core areas of math itself (number theory, geometry, analysis, and algebra).

There is a dichotomy in logic. Given a statement (theorem/axiom/whatever), there is the syntax of the statement (what is written down on the board) and the semantics (what the statement really means). In some sense, we want to study the abstract semantics, but it is usually much easier to study the concrete syntax. Surprisingly, sometimes this works.

Now for some examples. Here we understand that the variables range over the natural numbers and the operations are as usual.

(i) $\exists y(x = y + y)$ says that x is even.

(ii) $(1 < p) \wedge \forall r\ \forall s(p = rs \Rightarrow (r = 1) \vee (s = 1))$ says that p is prime

(iii) $\forall x((1 + 1 < x) \wedge \text{even}(x) \Rightarrow \exists p\ \exists q(\text{prime}(p) \wedge \text{prime}(q) \wedge (x = p + q)))$ is the Goldbach conjecture

We don't know if the GC is true in $\mathbb{N}$. As a sentence, it has an interpretation in $\mathbb{R}$, but it is obviously false!

In a given axiom system Σ , there are two possible criteria for judging whether or not we accept a statement σ. Syntactically, we accept σ if it is a theorem of the system, that is, it can be proved. Sematically, we accept σ if the meaing of σ is true under every interpretation which is compatible with the axioms of σ (this is called logical or semantic consequence). The completeness theorem says that these agree.

2 First-order logic

2.1 Languages

Definition 2.1. A *language* L is a disjoint union of sets L^{rel} and L^{fun} where

(i) L^{rel} is the set of *relation symbols*, each R of which has an associated *arity* $a(R) \in \mathbb{N}$

(ii) L^{fun} is the set of *function symbols*, each F of which has an associated *arity* $a(F) \in \mathbb{N}$

A constant symbol is a function symbol of arity zero.

Example 2.2. (i) Language of group: $L_{group} = \{1, ^{-1}, \cdot\}$. There is a constant function 1, a unary function $^{-1}$, and a binary function $\cdot$

(ii) Additive language of Abelian groups: $L_{ab} = \{0, -, +\}$

(iii) Language of order: $L_{order} = \{<\}$ where $<$ is a binary relation

(iv) Language of ordered Abelian groups: $L_{abo} = \{<, 0, -, +\}$

(v) Language of rings: $L_{ring} = \{0, 1, -, +, \cdot\}$

(vi) Language of semirings: $L_{semiring} = \{0, 1, +, \cdot\}$

(vii) Language of ordered rings: $L_{o-ring} = \{<, 0, 1, -, +, \cdot\}$

(viii) Language of real vector spaces: $L_{vs(\mathbb{R})} = \{0, (\mu_r : r \in \mathbb{R}), -, +\}$ where μ_r is a unary function symbol for each $r \in \mathbb{R}$

Languages are not invested with meaning, they are just symbols! Languages form the basis for the syntax of logic.

2.2 Structures

Definition 2.3. Let L denote some language. A *structure* $\mathcal{A}$ for L (also called an L-structure) is a triple $(A, (R^{\mathcal{A}})_{R \in L^{rel}}, (F^{\mathcal{A}})_{F \in L^{fun}})$ consisting of the following: A nonempty set A, called the *underlying set* of $\mathcal{A}$ (or the *universe* of $\mathcal{A}$), for each $R \in L^{rel}$ a set $R^{\mathcal{A}} \subseteq A^{a(R)}$, and for each $F \in L^{fun}$ a function $F^{\mathcal{A}} : A^{a(F)} \to A$. These are called the *interpretations* of the relations and the functions, respectively. The interpretations of the functions and the relations are called the *primatives* of $\mathcal{A}$.

Remark. If c is a constant symbol then $c^{\mathcal{A}} : A^0 \to A$. Since A^0 is a single point, $c^{\mathcal{A}}$ is completely determined by it's range. Through abuse of notation, we say $c^{\mathcal{A}} \in A$.

Notation. When the structure is clear from the context, we will abuse the notation and use the same notation for a relation/function symbol and it's interpretation.

Example 2.4. (i) Every groups is an L_{group}-structure by interpreting $1, ^{-1}, \cdot$ as the identity, inverse, and multiplication of the group, respectively. e.g. $\mathcal{R} = (R^{\times}, 1, ^{-1}, \cdot)$ We can also make $\mathbb{N}$ an L_{group}-structure by interpreting 1 as the number 1, $\cdot$ as multiplication and $^{-1}$ as the zero map ($n^{-1} = 0$), $\chi = (\mathbb{N}, 1^{\chi}, ^{-1^{\chi}}, \cdot^{\chi})$. Note that this is not actually a group.

(ii) If $\mathcal{A} = (A, 0, -, +)$ is an Abelian group then $\mathcal{A}$ is an L_{ab}-structure in the natural way. For example, $(\mathbb{R}^+, 0, -, +)$

(iii) An $L_{vs(\mathbb{R})}$-structure: $\mathcal{A} = (\mathbb{R}[x], 0^{\mathcal{A}}, \mu_r^{\mathcal{A}} : P \mapsto rP, -^{\mathcal{A}}, +^{\mathcal{A}})$

2.3 Terms

Definition 2.5. Var $= \{v_0, v_1, v_2, \ldots\}$ is a countably infinite set of symbols whose elements are called variables. We assume that $v_m \neq v_n$ for $m \neq n$ and we assume that no variable appears as a function or relation symbol in any language that we will ever consider.

Remark. We often use x, y, z, etc. as variables. It doesn't really matter. For most of what we do, it isn't necessary that Var be countably infinite. We will point out in proofs where one needs to make changes in the proofs if Var is not countably infinite.

Definition 2.6. Suppose that L is a language. An *L-term* is a word on the alphabet $L^{fun} \cup \text{Var}$ (where "word" has the obvious (finite) meaning) such that:

(i) every variable is a term (as a word of length 1)

(ii) if $F \in L^{fun}$ is n-ary and $t_1, \ldots, t_n$ are terms, then $F\, t_1 \cdots t_n$ is a term

Example 2.7. $L = L_{ring}$, then $\cdot + 11 + x - y$ is a valid L_{ring}-term. We would normally write this as $(1 + 1) \cdot (x - y)$. We will be writing our terms in the "natural" way (using parentheses and infix notation) even though it is not strictly correct.

Lemma 2.8. *Suppose that $t_1, \dots, t_m$ and $v_1, \dots, v_n$ are L-terms and that $t_1 \cdots t_m = v_1 \cdots v_n$. Then $n = m$ and $t_i = v_i \ \forall\, i$.*

Proof. By induction on the length of $v_1 \cdots v_n =: l$. If $l = 0$ then both words are empty and we are done. If $m = 0$ then we are also done, so we may assume that $m > 0$. Suppose that $l > 0$ and the lemma is true for all words with smaller length. So the first letter of t_1 and v_1 is the same. There are two possibilities, either t_1 is a variable, or $t_1 = F\, a_1, \dots, a_k$ for some F, a_i's, and k. If $t_1 = F\, a_1, \dots, a_k$, then the first symbol of v_1 is F, and so $v_1 = F\, b_1, \dots, b_k$ for some terms b_i. Thus we get $Fa_1 \cdots a_k t_1 \cdots t_m = Fb_1 \cdots b_k v_1 \cdots v_n$ We may strip of the F and get that $a_1 \cdots a_k t_1 \cdots t_m = b_1 \cdots b_k v_1 \cdots v_n$. By the induction hypothesis $k + n = k + m$, so $n = m$ and $a_i = b_i$ for $i = 1, \dots k$ and $t_j = v_j$ for $j = 2, \dots n$. It follows that $t_1 = v_1$ and we are done, as the other case is analogous. □

Proposition 2.9. (Unique readability of terms) *Suppose that L is a language and t is an L-term. Then either t is a vabiable or t can be written uniquely as $F\, t_1 \cdots t_n$ for some n-ary function symbol F and L-terms t_i, $i = 1, 2, \dots, n$.*

Proof. Suppose that t is not a variable. Then $t = F\, t_1 \cdots, t_n$ for some n-ary function symbol F and terms t_i. If there is another way of writing t as $G\, v_1 \cdots v_m$, then $F\, t_1 \cdots t_n = G\, v_1 \cdots v_m$ (as words) and so $F = G$ (as function symbols). Thus $n = m$, and by the lemma $t_i = v_i \ \forall\, i$. □

Notation. We will often write $t = t(x_1, \dots, x_n)$ to indicate that the variables in t come from the list $x_1, \dots, x_n$. Note that it is not necessary that each of $x_1, \dots, x_n$ appear in t.

Definition 2.10. (Interpreting terms) Suppose that $\mathcal{A}$ is an L-structure and $t = t(x)$ is an L-term, where $x = (x_1, \dots, x_m)$ ($m \in \mathbb{N}$ is fixed). We associate to $t(x)$ a function $t^{\mathcal{A}} : A^m \to A$ given by

(i) if t is the variable x_i then define $t^{\mathcal{A}}(a) = a_i \ \forall\, a \in A^m$

(ii) if $t = F\, t_1 \cdots t_n$ for some n-ary function F then define $t^{\mathcal{A}}(a) = F^{\mathcal{A}}(t_1^{\mathcal{A}}(a), \dots, t_n^{\mathcal{A}}(a)) \ \forall\, a \in A^m$

We call $t^{\mathcal{A}}$ the *interpretation* of $t(x)$.

Remark. The definition is well-founded (works) because of Unique readability. Notice that $t^{\mathcal{A}}$ depends not only on t but also on $x = (x_1, \dots, x_m)$.

Example 2.11. $\mathcal{R} = (\mathbb{R}, 0, 1, -, +, \cdot)$ is an L_{ring}-structure. The term $\cdot + 11 + x - y$ is interpreted in $\mathcal{R}$ as the function $t^{\mathcal{R}} : \mathbb{R}^2 \to \mathbb{R} : (a, b) \mapsto 2(a - b)$.

2.4 Substructures

Definition 2.12. Suppose that $\mathcal{A}, \mathcal{B}$ are L-structures. Then $\mathcal{A}$ is a *substructure* of $\mathcal{B}$, denoted by $\mathcal{A} \subseteq \mathcal{B}$, if

(i) the underlying set A of $\mathcal{A}$ is a subset of the underlying set B of $\mathcal{B}$

(ii) for all relations $R \in L^{rel}$ of arity m, $R^{\mathcal{A}} = R^{\mathcal{B}} \cap A^m$

(iii) for all functions $F \in L^{fun}$ of arity n, $F^{\mathcal{A}} = F^{\mathcal{B}}|_{A^n}$. In particular, $F^{\mathcal{B}}(A^n) \subseteq A$.

Remark. Suppose $\mathcal{B}$ is an L-structure and $A \subseteq B$ is some subset of the universe of $\mathcal{B}$. If $F^{\mathcal{B}}(A^n) \subseteq A \ \forall\, F \in L^{fun}$ then A is the underlying set of a structure $\mathcal{A} \subseteq \mathcal{B}$. In particular,

$$\mathcal{A} = (A, (R^{\mathcal{B}} \cap A^{a(R)})_{R \in L^{rel}}, (F^{\mathcal{B}}|_{A^{a(F)}})_{F \in L^{fun}})$$

Example 2.13. (i) $(\mathbb{Z}, 0, 1, -, +, \cdot) \subseteq (\mathbb{Q}, 0, 1, -, +, \cdot) \subseteq (\mathbb{R}, 0, 1, -, +, \cdot)$

(ii) $(\mathbb{N}, <, 0, 1, +, \cdot) \subseteq (\mathbb{Z}, <, 0, 1, +, \cdot)$

(iii) But $\mathbb{N}$ is not the underlying set of any substructure of $(\mathbb{Z}, 0, 1, -, +, \cdot)$.

2.5 Interpreting L-formulas

Definition 2.14. The *logical symbols* are the variables (Var) together with

$$\top,\ \bot,\ \neg,\ \wedge,\ \vee,\ \exists,\ \forall,\ =$$

Definition 2.15. (i) An *atomic L-formula* is a word on the alphabet $L \cup \mathrm{Var} \cup \{\top, \bot, =\}$ of one of the following forms

(a) $\top$

(b) $\bot$

(c) $R\,t_1 \cdots t_m$, where $R \in L^{rel}$ and the t_i are L-terms

(d) $= t_1 t_2$, where t_1, t_2 are L-terms

(ii) An *L-formula* is a word on the alphabet of L union the logical symbols, defined inductively by

(a) every atomic formula is a formula

(b) if φ, ψ are formulae then so are $\wedge\varphi\psi$, $\vee\varphi\psi$, and $\neg\varphi$

(c) if φ is a formula and x is a variable then $\exists\, x\varphi$ and $\forall\, x\varphi$ are formulae

Remark. (i) We will write the formulas using infix notation for readability

(ii) We will use $\varphi \to \psi$ to abbreviate the formula $\vee\neg\varphi\psi$ (i.e. $\neg\varphi \vee \psi$)

Definition 2.16. Suppose that φ is an L-formula. Write φ as a string $s_1 \cdots s_m$, where each s_i is in L or is a logical symbol. A *subformula* of φ is a subword of the form $s_i \cdots s_k$, where $1 \leq i \leq k \leq m$ that is itself a formula. An occurrence of a variable x in the j^{th} place (so $x = s_j$) is said to be *bound* if there is a subformula of φ, $s_i \cdots s_k$, such that $1 \leq i \leq j \leq k \leq m$ of the form $\forall\, x\psi$ or $\exists\, x\psi$ for some formula ψ. An occurrence that is not bound is said to be *free*.

Example 2.17. Consider the formula $\exists\, x(x \cdot y = z) \wedge x^{-1} = 1$. Then the first two occurrences of x are bound, and the last occurrence of x, y and z are free.

Definition 2.18. An *L-sentence* is an L-formula with no free occurrence of variables.

Proposition 2.19. (Unique readability of formulae) *Suppose L is a language and φ is an L-formula. Then φ can be written uniquely as exactly one of the following*

(i) $\top$

(ii) $\bot$

(iii) $R\,t_1 \cdots t_m$, *where* $R \in L^{rel}$ *and the* t_i *are* L*-terms*

(iv) $= t_1 t_2$, *where* t_1, t_2 *are* L*-terms*

(v) $\neg\varphi$, *where* φ *is a formula*

(vi) $\wedge\varphi\psi$, *where* φ *and* ψ *are formulas*

(vii) $\vee\varphi\psi$, *where* φ *and* ψ *are formulas*

(viii) $\exists\, x\varphi$, *where* φ *is a formula and* x *is a variable*

(ix) $\forall x\varphi$, where φ is a formula and x is a variable

Notation. Suppose that φ is an L-formula. We often write $\varphi = \varphi(x_1, \ldots, x_n)$ to indicate that all of the free variables of φ are amoung the variables $x_1, \ldots, x_n$. It is not necessary that each of these variables actually appears as a free variable. It is assumed that they are distinct variables, however.

Definition 2.20. Suppose that $t_1, \ldots, t_n$ are L-terms and $x_1, \ldots, x_n$ are variables.

(i) If t is a term, then by $t(t_1/x_1, \ldots t_n/x_n)$ we mean the word obtained from t by replacing x_i with t_i simultaneously.

(ii) If φ is a formula, then by $\varphi(t_1/x_1, \ldots, t_n/x_n)$ we mean the word obtained from φ by replacing each free occurrence of x_i with t_i simultaneously.

Lemma 2.21. *(i) $t(t_1/x_1, \ldots t_n/x_n)$ is a term*

(ii) $\varphi(t_1/x_1, \ldots, t_n/x_n)$ is a formula
Also note that φ atomic implies that $\varphi(t_1/x_1, \ldots, t_n/x_n)$ is atomic.

Proof. (Sketch) First prove (i) by induction on length. This implies (ii) for φ atomic. Next, prove (ii) for general φ by induction on length. For example, say φ is $\exists x\psi$. There are two cases

(i) $x \notin \{x_1, \ldots, x_n\}$. Then $\varphi(t_1/x_1, \ldots, t_n/x_n)$ is $\exists x\psi(t_1/x_1, \ldots, t_n/x_n)$ and by induction, the right hand side is a formula

(ii) Without loss of generality, $x = x_1$. Then φ is $\exists x_1\psi$. All occurrences of x_1 in φ are bounded, so by definition $\varphi(t_1/x_1, \ldots, t_n/x_n)$ is $\varphi(t_2/x_2, \ldots, t_n/x_n)$ and by induction we are done

□

We will write $\varphi(t_1, \ldots, t_n)$ as shorthand for $\varphi(t_1/x_1, \ldots, t_n/x_n)$. Suppose we have a formula $\varphi(x_1, \ldots, x_n)$ and a stucture $\mathcal{A} = (A, \ldots)$. For elements of the structure $a_1, \ldots a_n \in A$, what do we mean by $\varphi(a_1, \ldots, a_n)$?

Definition 2.22. Suppose that $\mathcal{A} = (A, \ldots)$ is an L-structure. For each $a \in A$, we choose a new constant symbol $\underline{a}$ called the *name* of a. The language obtained from L by adding the names $\underline{a}$ for each $a \in A$ is denoted by $L_A := L \cup \{\underline{a} \mid a \in A\}$. The L-structure $\mathcal{A}$ can be made into an L_A-structure, $\mathcal{A}_A$, by keeping the same interpretation of symbols in L as $\mathcal{A}$ had, and interpreting each name $\underline{a}$ as a, (i.e. $\underline{a}^{\mathcal{A}_A} = a$).

It is understood that different names are used for different elements. It is also understood that the name of a is not a function, relation, or logical symbol. Note that L_A depends on $\mathcal{A}$. If $t(x_1, \ldots, x_n)$ is an L-term and $a_1, \ldots, a_n \in A$, then $t(\underline{a_1}, \ldots, \underline{a_n})$ is a variable-free L_A-term. Similarily, if $\varphi(x_1, \ldots, x_n)$ is an L-formula, then $\varphi(\underline{a_1}, \ldots, \underline{a_n})$ is an L_A-sentence. Note that every L-formula (resp. sentence) is an L_A-formula (resp. sentence). Every L_A-term $t(x_1, \ldots, x_n)$ is of the form $s(x_1, \ldots, x_n, \underline{a_1}, \ldots, \underline{a_m})$ for some L-term s and elements $a_1, \ldots, a_m \in A$. Then $t^{\mathcal{A}_A} : A^n \to A$, whereas $s^{\mathcal{A}} : A^{n+m} \to A$.

2.6 Truth

Definition 2.23. Let L be a language and $\mathcal{A}$ be an L-structure. We define inductively what it means for an L_A-sentence to be true in $\mathcal{A}$. If σ is an L_A-sentence, we write $\mathcal{A} \models \sigma$ to mean that σ is true in $\mathcal{A}$.

(i) $\mathcal{A} \models \top$, and $\mathcal{A} \not\models \bot$

(ii) $\mathcal{A} \models R\,t_1 \cdots t_m$ if and only if $(t_1^{\mathcal{A}}, \ldots, t_m^{\mathcal{A}}) \in R^{\mathcal{A}}$ where $t_1, \ldots, t_m$ are variable free L_A-terms

(iii) $\mathcal{A} \models t_1 = t_2$ if and only if $t_1^{\mathcal{A}} = t_2^{\mathcal{A}}$ where t_1, t_2 are variable free L_A-terms.

(iv) $\sigma = \neg\sigma_1$, $\mathcal{A} \models \sigma$ if and only if $\mathcal{A} \not\models \sigma_1$

(v) $\sigma = \sigma_1 \wedge \sigma_2$, $\mathcal{A} \models \sigma$ if and only if $\mathcal{A} \models \sigma_1$ and $\mathcal{A} \models \sigma_2$

(vi) $\sigma = \sigma_1 \vee \sigma_2$, $\mathcal{A} \models \sigma$ if and only if $\mathcal{A} \models \sigma_1$ or $\mathcal{A} \models \sigma_2$

(vii) $\sigma = \exists x\varphi(x)$, $\mathcal{A} \models \sigma$ if and only if there exists $a \in A$ such that $\mathcal{A} \models \varphi(\underline{a})$

(viii) $\sigma = \forall x\varphi(x)$, $\mathcal{A} \models \sigma$ if and only if for all $a \in A$ we have $\mathcal{A} \models \varphi(\underline{a})$

In particular, we have defined what it means for an L-sentence to be true in $\mathcal{A}$. We also say that $\mathcal{A}$ *satisfies* σ or that σ *holds in* $\mathcal{A}$ for $\mathcal{A} \models \sigma$. If $\mathcal{A} \models \sigma = \varphi(\underline{a_1}, \ldots, \underline{a_n})$ for φ an L-formula and $(a_1, \ldots, a_n) \in A^n$, then we say $(a_1, \ldots, a_n)$ *realizes* φ in $\mathcal{A}$, or $(a_1, \ldots, a_n)$ is *true of* φ in $\mathcal{A}$.

Definition 2.24. Given an L_A-formula $\varphi(x_1 \ldots, x_n)$, we let $\varphi^{\mathcal{A}_A}$ be the set of n-tuples that realize φ (i.e. $\{(a_1, \ldots, a_n) \in A^n \mid \mathcal{A} \models \varphi(\underline{a_1}, \ldots, \underline{a_n})\}$). We say that $\varphi^{\mathcal{A}}$ is *defined* by φ. A set $S \subseteq A^N$ is *definable in* $\mathcal{A}$ if $S = \varphi^{\mathcal{A}}$ for some L_A-formula $\varphi(x_1 \ldots, x_n)$. Moreover, if φ can be chosen to be of the form $\varphi(x_1 \ldots, x_n) = \psi(x_1 \ldots, x_n, \underline{b}_1, \ldots \underline{b}_m)$ where ψ is an L-formula and $b_1, \ldots b_m \in B \subseteq A$ then we say that φ is *B-definable*. In this case we may also say that S is *definable with parameters from B*. When $B = \varnothing$ we say that S is *0-definable*.

Example 2.25. In $(\mathbb{R}, <, 0, 1, +, -, \cdot)$, $S = \{x \in \mathbb{R} \mid x < \sqrt{2}\} \subseteq \mathbb{R}$ is definable. For example, it is defined by $\varphi(x) := x < \sqrt{2}$, an $L_{\mathbb{R}}$-formula. It is actually $\{\sqrt{2}\}$-definable. We can do better, by taking

$$\psi(x) := (x \cdot x < 1 + 1) \vee (x < 0)$$

ψ is an L-formula and $S = \psi^{\mathcal{A}}$, so S is 0-definable.

Example 2.26. In $(\mathbb{R}, <, 0, +, -)$, $S = \{x \in \mathbb{R} \mid x < \sqrt{2}\} \subseteq \mathbb{R}$ is $\{\sqrt{2}\}$-definable, but not 0-definable (at least by ψ). We will be able to prove this for certain later.

3 Model Theory

Definition 3.1. Suppose Σ is a collection of L-sentences. A *model* of Σ is an L-structure $\mathcal{A}$ such that each sentence in Σ is *true* in $\mathcal{A}$.

Example 3.2. (i) Groups are exactly the L_{group}-structures that are models of

$$GR := \{\forall x(x \cdot 1 = x \wedge 1 \cdot x = x), \forall x(x \cdot x^{-1} = 1 \wedge x^{-1} \cdot x = 1), \forall xyz(x \cdot (y \cdot z) = (x \cdot y) \cdot z)\}$$

(ii) Abelian groups are exactly the L_{group}-structures that are models of $GR \cup \{\forall xy(x \cdot y = y \cdot x)\}$. Alternatively, Abelian groups are the L_{ab}-structures that are models of

$$AB := \{\forall x(x + 0 = x), \forall x(x + (-x) = 0), \forall xyz(x + (y + z) = (x + y) + z), \forall xy(x + y = y + x)\}$$

(iii) Rings are exactly the L_{ring}-structures that are models of the ring axioms,

$$RI = AB \cup \{\forall x(x \cdot 1 = x \wedge 1 \cdot x = x), \forall xyz(x \cdot (y \cdot z) = (x \cdot y) \cdot z)\}$$

(iv) Fields are the models of the field axioms,

$$F = RI \cup \{\forall x \exists y(x \cdot y = 1 \wedge y \cdot x = 1), \forall xy(x \cdot y = y \cdot x), 1 \neq 0\}$$

(v) Algebraically closed fields are the models of

$$ACF = F \cup \{\forall u_1, \dots u_n \exists x(x^n + u_1 x^{n-1} + \cdots + u_n = 0) \mid n = 2, 3, \dots\}$$

Definition 3.3. We say that an L-sentence σ is a *logical consequence* of a collection of L-sentences Σ, written $\Sigma \models \sigma$, if σ is true in every model of Σ.

Example 3.4. $GR \models \forall xyz(x \cdot y = x \cdot z \to y = z)$

The completeness theorem will say that $\Sigma \models \sigma$ if and only if σ is "proveable" from Σ, once we know what proveable means.

Definition 3.5. Suppose that $\mathcal{A} = (A, \dots)$ is an L-structure. An $\mathcal{A}$-*instance* of an L-formula $\varphi(x_1, \dots, x_n)$ is an L_A-sentence of the form $\varphi(\underline{a}_1, \dots, \underline{a}_n)$, where $a_1, \dots, a_n \in A$. This is independent of the choice of $(x_1, \dots, x_n)$, but this requires proof. We say that an L-formula is *valid* in $\mathcal{A}$ if every $\mathcal{A}$-instance is true, denoted $\mathcal{A} \models \varphi$. φ is *satisfiable* in $\mathcal{A}$ if some $\mathcal{A}$-instance is true. An L-formula φ is a *logical consequence* of a collection of L-sentences Σ, written $\Sigma \models \varphi$, if φ is valid for all $\mathcal{A} \models \Sigma$. ($\mathcal{A} \models \Sigma$ means that $\mathcal{A}$ is a model of Σ.)

Exercise. $\varphi(x_1, \dots, x_n)$ is valid if an only if $\mathcal{A} \models \forall x_1, \dots, x_n \varphi$, if and only if $\varphi^{\mathcal{A}} = A^n$

Lemma 3.6. *$\mathcal{A} = (A, \dots)$ an L-structure. Let t be a variable free L_A-term. with $t^{\mathcal{A}} = a \in A$. Then*

(i) If $s(x)$ is an L_A-term then $s(t)$ is a variable free L_A-term, and $s(t)^{\mathcal{A}} = s(\underline{a})^{\mathcal{A}} = s^{\mathcal{A}}(a)$

(ii) If $\varphi(x)$ is an L_A-formula then $\varphi(t)$ is an L_A-sentence, and $\mathcal{A} \models \varphi(t)$ if and only if $\mathcal{A} \models \varphi(\underline{a})$

4 "Proofs"

4.1 Logical Axioms

Definition 4.1. L is a language. The *logicals axioms* of L are the *propositional axioms*, the *equality axioms*, and the *quantifier axioms*. A propositional axiom of L is a formula of one of the following forms, where φ, ψ, θ are L-formulas

(i) $\top$

(ii) $\varphi \to (\psi \vee \varphi)$; $\varphi \to (\varphi \vee \psi)$

(iii) $\neg\varphi \to (\neg\psi \to \neg(\varphi \vee \psi))$

(iv) $(\varphi \wedge \psi) \to \varphi$; $(\varphi \wedge \psi) \to \psi$

(v) $\varphi \to (\psi \to (\varphi \wedge \psi))$

(vi) $\varphi \to (\neg\varphi \to \bot)$

(vii) $(\varphi \to (\psi \to \theta)) \to ((\varphi \to \psi) \to (\varphi \to \theta))$

(viii) $(\neg\varphi \to \bot) \to \varphi$

Equality axioms of L are the following L-formulas, where x's, y's, and z's are variables

(i) $x = x$

(ii) $x = y \to y = x$

(iii) $(x = y) \wedge (y = z) \to x = z$

(iv) $(x_1 = y_1) \wedge \cdots \wedge (x_n = y_n) \wedge R\,x_1 \ldots x_n \to R\,y_1 \ldots y_n$ where $R \in L^{rel}$ is n-ary

(v) $(x_1 = y_1) \wedge \cdots \wedge (x_n = y_n) \to F\,x_1 \ldots x_n = F\,y_1 \ldots y_n$ where $F \in L^{fun}$ is n-ary

If φ is an L-formula, t an L-term, and y a variable, we say that t is free for y in φ if no variable in t becomes bound upon replacing free occurrences of y by t in φ. That is, for all variables x in t, there is no subformula of φ of the form $\forall\, x\psi$ or $\exists\, x\psi$ where ψ has an occurrence of y that is free in φ. The quantifier axoims are L-formulas of the form

(i) $\varphi(t/y) \to \exists\, y\varphi$

(ii) $\forall\, y\varphi \to \varphi(t/y)$

where φ is any L-formula, t is an L-term which is free for any y in φ.

You must check that every logical axiom of L is valid in every L-structure!

Definition 4.2. The logical rules of L are:

(MP) From φ and $\varphi \to \psi$ infer ψ

(G) If x does not occur freely in φ then

(a) from $\varphi \to \psi$ infer $\varphi \to \forall\, x\psi$

(b) from $\psi \to \varphi$ infer $\exists\, x\psi \to \varphi$

Check that if the hypotheses of a logical rule are valid in a given L-structure then so are the conclusions.

Definition 4.3. A *(formal) proof* of an L-formula φ from a set of L-sentences Σ is a sequence $\varphi_1, \ldots, \varphi_k$ of L-formulae such that

(i) $\varphi = \varphi_k$

(ii) For all $1 \leq j \leq k$, φ_j is either

(a) a logical axiom

(b) a sentence in Σ

(c) there are $1 \leq r, s < j$ such that φ_j can be infered from φ_r and φ_s by (MP) or from φ_r by (G).

We say that Σ proves φ and write $\Sigma \vdash \varphi$ if there exists a proof of φ from Σ.

Note that these are not the only possible logical axioms and logical rules.

Lemma 4.4. *For any L-formulas φ, ψ, θ,*

(i) Law of the Excluded Middle: $\vdash \varphi \vee \neg\varphi$

(ii) Contraction rule: $\vdash \varphi \vee \varphi \to \varphi$

(iii) Associativity: $\vdash (\varphi \vee \psi) \vee \theta \to \varphi \vee (\psi \vee \theta)$

(iv) Cut rule: $\vdash (\varphi \vee \psi) \wedge (\neg\varphi \vee \theta) \to (\psi \vee \theta)$

Proof. Exercise.

(i) Propostion axiom scheme 2 says that $\gamma \to \beta \vee \gamma$. In particular, $\varphi \to (\neg(\neg\varphi \vee \varphi) \vee \varphi)$ and $\varphi \to (\neg\varphi \vee \varphi)$. But the first is equivalent to $\varphi \to ((\neg\varphi \vee \varphi) \to \varphi)$ and the second $\varphi \to (\varphi \to \varphi)$. Now from another axiom, $(\gamma \to (\beta \to \pi)) \to ((\gamma \to \beta) \to (\gamma \to \pi))$, and so (sub it just φ, apply (MP)).

□

4.2 Propositional Logic

Definition 4.5. Given L-formulas $\varphi_1, \ldots, \varphi_n$, let $\mathrm{Prop}(\varphi_1, \ldots, \varphi_n)$ be the set of L-formulas obtained inductively as follows

(i) $\top, \bot, \varphi_1, \ldots, \varphi_n \in \mathrm{Prop}(\varphi_1, \ldots, \varphi_n)$

(ii) If $\theta, \psi \in \mathrm{Prop}(\varphi_1, \ldots, \varphi_n)$, so are $\neg\theta, \theta \wedge \psi, \theta \vee \psi$

A *truth assignment* on $\{\varphi_1, \ldots, \varphi_n\}$ is a map $t : \{\varphi_1, \ldots, \varphi_n\} \to \{0, 1\}$. We extend t to $\mathrm{Prop}(\varphi_1, \ldots, \varphi_n)$ inductively as follows $\hat{t} : \mathrm{Prop}(\varphi_1, \ldots, \varphi_n) \to \{0, 1\}$

(i) $\hat{t}(\top) = 1, \hat{t}(\bot) = 0$

(ii) $\hat{t}(\neg\theta) = 1 - \hat{t}(\theta)$

(iii) $\hat{t}(\theta \wedge \psi) = \min\{\hat{t}(\theta), \hat{t}(\psi)\}$

(iv) $\hat{t}(\theta \vee \psi) = \max\{\hat{t}(\theta), \hat{t}(\psi)\}$

An L-formula φ is an *L-tautology* if there exist L-formulae $\varphi_1, \ldots, \varphi_n$ such that $\varphi \in \mathrm{Prop}(\varphi_1, \ldots, \varphi_n)$ and for every truth assignment $t : \{\varphi_1, \ldots, \varphi_n\} \to \{0, 1\}$, $\hat{t}(\varphi) = 1$.

Note that all of the L-formulas in Lemma 4.4 above are L-tautologies, and also all of our propositional axioms are L-tautologies.

Theorem. (Tautology Theorem) *$\vdash \varphi$ for any L-tautology φ.*

Proof. See text. This proof requires some tricks. □

Lemma 4.6. *Let Σ be a set of L-sentences. Then*

(i) If $\Sigma \vdash \varphi$ then $\Sigma \vdash \forall x\varphi$

(ii) If $\Sigma \vdash \varphi$ and t is an L-term which is free for x in φ then

$$\Sigma \vdash \varphi(t/x)$$

(iii) If $\Sigma \vdash \varphi$ and $t_1, \ldots, t_n$ are L-terms whose variables do not occur bound in φ then

$$\Sigma \vdash \varphi(t_1/x_1, \ldots t_n/x_n)$$

Proof. (i) $\varphi \to (\neg\forall x\varphi \to \varphi)$ is an L-tautology on $\{\varphi, \forall x\varphi\}$. So Σ proves it. Since $\Sigma \vdash \varphi$, by (MP) $\Sigma \vdash \neg\forall x\varphi \to \varphi$. So by (G), $\Sigma \vdash \neg\forall x\varphi \to \forall x\varphi$. But $(\neg\forall x\varphi \to \forall x\varphi) \to \forall x\varphi$ is a tautology, and by (MP), $\Sigma \vdash \forall x\varphi$

(ii) $\Sigma \vdash \varphi$ gives $\Sigma \vdash \forall x\varphi$. $\forall x\varphi \to \varphi(t/x)$ is a quantifier axiom since t is free for x in φ. Therefore by (MP) $\Sigma \vdash \varphi(t/x)$.

(iii) Let $y_1, \ldots y_n$ be fresh variables. By using (ii) repeatedly, $\Sigma \vdash \varphi(y_1/x_1, \ldots y_n/x_n)$. Setting

$$\psi = \varphi(y_1/x_1, \ldots y_n/x_n)$$

by (ii) repeatedly, $\Sigma \vdash \psi(t_1/y_1, \ldots t_n/y_n)$. Now observe

$$\psi(t_1/y_1, \ldots t_n/y_n) = \varphi(t_1/x_1, \ldots t_n/x_n)$$

□

Lemma. (Deduction Lemma) *If $\Sigma \cup \{\sigma\} \vdash \varphi$ then $\Sigma \vdash \sigma \to \varphi$, where Σ is a set of L-sentences, σ is an L-sentence, and φ is any L-formula.*

Proof. By induction on the length of the proof of φ from $\Sigma \cup \{\sigma\}$. Suppose the proof is of length one. If φ is a logical axiom then $\Sigma \vdash \varphi$ and $\varphi \to (\sigma \to \varphi)$ and we are done by (MP). If $\varphi \in \Sigma \cup \{\sigma\}$ then either $\varphi \in \Sigma$ and we are done as before, or $\varphi = \sigma$ and we are done since $\varphi \to \varphi$ is a tautology.

If φ is obtained from (MP) applied to θ and $\theta \to \varphi$ then by the induction hypothesis $\Sigma \vdash \sigma \to \theta$ and $\Sigma \vdash \sigma \to (\theta \to \varphi)$.

$$(\sigma \to (\theta \to \varphi)) \to ((\sigma \to \theta) \to (\sigma \to \varphi))$$

is a propositional axiom. Using (MP) we have

$$\Sigma \vdash (\sigma \to \theta) \to (\sigma \to \varphi)$$

and applying it again gives us $\Sigma \vdash \sigma \to \varphi$.

Suppose that φ is obtained by (G). There are two cases

(i) $\varphi = \varphi' \to \forall x \psi$ where x does not occur freely in φ', and $\varphi' \to \psi$ appears earlier in the proof of φ from $\Sigma \cup \{\sigma\}$. By the induction hypothesis, $\Sigma \vdash \sigma \to (\varphi' \to \psi)$.

$$(\sigma \to (\varphi' \to \psi)) \to ((\sigma \wedge \varphi') \to \psi)$$

is a tautology. By (MP) $\Sigma \vdash (\sigma \wedge \varphi') \to \psi$ and x does not occur freely in $\sigma \wedge \varphi'$. By (G) $\Sigma \vdash \sigma \wedge \varphi' \to \forall x \psi$.

$$((\sigma \wedge \varphi') \to \forall x \psi) \to (\sigma \to (\varphi' \to \forall x \psi))$$

is a tautology. By (MP) $\Sigma \vdash \sigma \to \varphi$.

(ii) $\varphi = \exists x \psi \to \varphi'$ where x does not occur freely in φ', and $\psi \to \varphi'$ appears earlier in the proof of φ from $\Sigma \cup \{\sigma\}$. Exercise.

□

Definition 4.7. A set of L-sentences Σ is said to be consistent if $\Sigma \nvdash \bot$. Otherwise, Σ is said to be inconsistent.

Corollary 4.8. *$\Sigma \vdash \sigma$ if and only if $\Sigma \cup \{\neg\sigma\}$ is inconsistent.*

Proof. Assume that $\Sigma \vdash \sigma$. $\sigma \to (\neg\sigma \to \bot)$ is a propositional axiom. By (MP) $\Sigma \vdash \neg\sigma \to \bot$, and thus $\Sigma \cup \{\neg\sigma\} \vdash \neg\sigma \to \bot$. It is clear that $\Sigma \cup \{\neg\sigma\} \vdash \neg\sigma$, and so by (MP) $\Sigma \cup \{\neg\sigma\} \vdash \bot$. Therefore $\Sigma \cup \{\neg\sigma\}$ is inconsistent.

Now suppose that $\Sigma \cup \{\neg\sigma\} \vdash \bot$. Then by the deduction lemma $\Sigma \vdash \neg\sigma \to \bot$. But $(\neg\sigma \to \bot) \to \sigma$ is a propositional axiom. Therefore by (MP) $\Sigma \vdash \sigma$. □

4.3 Completeness

Proposition 4.9. *If $\Sigma \vdash \varphi$ then $\Sigma \models \varphi$, where Σ is a set of L-sentences and φ is an L-formula.*

Proof. Suppose $\Sigma \vdash \varphi$. Let $\mathcal{A} \models \Sigma$. Let $\varphi_1, \ldots, \varphi_k$ be a proof of φ from Σ. We show that $\mathcal{A} \models \varphi_j$ for each j. (In particular, $\mathcal{A} \models \varphi_k$ gives that $\mathcal{A} \models \varphi$.) Every logical axiom is valid in $\mathcal{A}$. If the hypotheses of any logical rule are valid in $\mathcal{A}$ then so are it's conclusions. Every sentence in Σ is true in $\mathcal{A}$, so by induction, $\mathcal{A} \models \varphi_j$ for all j. □

More interestingly,

Theorem. (Completeness Theorem, Gödel (1930)) $\Sigma \vdash \varphi$ *if and only if* $\Sigma \models \varphi$, *where* Σ *is a set of* L*-sentences and* φ *is an* L*-formula.*

Corollary. (Compactness Theorem) *If* $\Sigma \models \varphi$ *then there exists* $\Sigma_0 \subseteq \Sigma$ *with* Σ_0 *finite such that* $\Sigma_0 \models \varphi$.

Proof. $\Sigma \models \varphi$ implies that $\Sigma \vdash \varphi$ so there is a proof. Proofs are finite, so there is a finite subset $\Sigma_0 \subseteq \Sigma$ such that $\Sigma_0 \vdash \varphi$, which implies that $\Sigma_0 \models \varphi$. (Let Σ_0 be the set of the L-sentences that actually appear in the proof of φ from Σ.) □

Recall that Abelian groups are exactly the L_{ab}-structures which are models the Abelian group axioms AB. We say that the class of Abelian groups is "elementary".

Example 4.10. There is no collection of L_{group}-sentences, Σ, such that the models of Σ are exactly the finite groups. That is, the class of finite groups is not elementary. Suppose Σ exists. Let

$$\Sigma' := \Sigma \cup \{\exists\, x_1, \dots, x_n \wedge_{i \neq j} (x_i \neq x_j) \mid n = 1, 2, 3, \dots\}$$

Then Σ' has no models. It follows that $\Sigma' \models \bot$, vacuously. By compactness, there is some $N > 0$ such that

$$\Sigma'_N := \Sigma \cup \{\exists\, x_1, \dots, x_n \wedge_{i \neq j} (x_i \neq x_j) \mid n = 1, 2, \dots, N\}$$

and $\Sigma'_N \models \bot$. But any group of size at least N is a model of Σ'_N, but not a model of $\bot$. This is a contradiction.

Theorem. (Completeness Theorem, second form) Σ *is consistent if and only if* Σ *has a model.*

Proposition 4.11. *The two forms of completeness are equivalent. That is, the following are equivalent*

(i) $\Sigma \vdash \varphi$ *if an only if* $\Sigma \models \varphi$

(ii) Σ *is consistent if and only if* Σ *has a model*

Proof. (i) ⇒ (ii) Suppose that Σ is consistent. $\Sigma \nvdash \bot$, so $\Sigma \not\models \bot$ by (i). Thus Σ has model, for otherwise $\Sigma \models \bot$ vacuously. Conversely, suppose Σ has a model. Then $\Sigma \not\models \bot$, so $\Sigma \nvdash \bot$, again by (i).
(ii) ⇒ (i) Exercise. □

Theorem. (Compactness Theorem, second form) *If every finite subset of* Σ *has model then* Σ *has a model.*

4.4 Proof of the Completeness Theorem

We will prove that Σ is consistent if an only if Σ has a model. We've already proved that if Σ has a model then $\Sigma \not\models \bot$, so $\Sigma \nvdash \bot$ and Σ is consistent.

Suppose now that Σ is consistent. We need to find a model. Let T_L be the set of all variable free L-terms. Define $\sim_\Sigma$ by $t_1 \sim_\Sigma t_2$ if and only if $\Sigma \vdash t_1 = t_2$. The idea we are going use is that $T_L / \sim_\Sigma$ should be the universe of a model of Σ.

Lemma 4.12. *(i) For each* L*-term,* $\vdash t = t$

(ii) t, t' L*-terms, if* $\Sigma \vdash t = t'$ *then* $\Sigma \vdash t' = t$

(iii) t_1, t_2, t_3 L*-terms, if* $\Sigma \vdash t_1 = t_2$ *and* $\Sigma \vdash t_2 = t_3$ *then* $\Sigma \vdash t_1 = t_3$

(iv) $R \in L^{rel}$ m*-ary and* $t_1, t'_1, \dots, t_m, t'_m$ L*-terms such that* $\Sigma \vdash t_i = t'_i$ *for each* i*, then* $\Sigma \vdash R\, t_1 \dots t_m$ *implies that* $\Sigma \vdash R\, t'_1 \dots t'_m$

(v) *$F \in L^{fun}$ m-ary and $t_1, t_1', \ldots, t_m, t_m'$ L-terms such that $\Sigma \vdash t_i = t_i'$ for each i, then $\Sigma \vdash F\,t_1 \ldots t_m = F\,t_1' \ldots t_m'$*

Proof. These are just restatements of some of the logical axioms. Fill in the details. □

Parts (i) to (iii) enusure that $\sim_\Sigma$ is an equivalence relation on T_L. Suppose L has at least one constant symbol, so that T_L is non-empty.

Definition 4.13. Define an L-structure $\mathcal{A}_\Sigma$ as follows

- universe of $\mathcal{A}_\Sigma$ is $A_\Sigma = T_L/\sim_\Sigma$
- for $R \in L^{rel}$ m-ary, define $R^{\mathcal{A}_\Sigma} \subseteq (A_\Sigma)^m$ by $([t_1], \ldots [t_m]) \in R^{\mathcal{A}_\Sigma}$ if and only if $\Sigma \vdash R\,t_1 \ldots t_m$
- for $F \in L^{rel}$ m-ary, define $F^{\mathcal{A}_\Sigma}([t_1], \ldots [t_m]) = [t]$ if and only if $\Sigma \vdash F\,t_1 \ldots t_m = t$

Notice that the interpretations of the relation and function symbols is well founded because of parts (vi) and (v) of the lemma.

Lemma 4.14. *Suppose L has a constant symbol and Σ is consistent. Then*

(i) *for all $t \in T_L$, $t^{\mathcal{A}_\Sigma} = [t]$*

(ii) *for each atomic sentence σ, $\Sigma \vdash \sigma$ if and only if $\mathcal{A}_\Sigma \models \sigma$*

Proof. (i) Trivial induction.

(ii) Follows from (i) by induction. For example, $\sigma = R\,t_1 \ldots t_m$ for terms in T_L. Then $\Sigma \vdash \sigma$ if and only if $([t_1], \ldots, [t_m]) \in R^{\mathcal{A}_\Sigma}$, which is equvalent to having $(t_1^{\mathcal{A}_\Sigma}, \ldots, t_m^{\mathcal{A}_\Sigma}) \in R^{\mathcal{A}_\Sigma}$. This happens if and only if $\mathcal{A}_\Sigma \models R\,t_1 \ldots t_m$, by definition of $\models$. The other cases are analogous.

□

We wish to extend the above claim to all sentences. There are two obstacles. The first is that for any sentence σ, either $\mathcal{A}_\Sigma \models \sigma$ or $\mathcal{A}_\Sigma \models \neg\sigma$, by definition of $\models$. This is not necessarily the case for provability in Σ. That is, maybe $\Sigma \nvdash \sigma$ and $\Sigma \nvdash \neg\sigma$.

Definition 4.15. A set Σ of sentences is said to be *complete* if for every sentence σ either $\Sigma \vdash \sigma$ or $\Sigma \vdash \neg\sigma$.

For example, ACF is incomplete, as $ACF \nvdash \forall x(x + x = 0)$ and $ACF \nvdash \exists x(x + x \neq 0)$. Thus it is impossible for part (ii) to be true for all σ.

If $\sigma = \exists x\varphi(x)$, where φ is atomic, suppose that $\Sigma \vdash \exists x\varphi(x)$. For $\mathcal{A}_\Sigma \models \sigma$, one needs an element $a \in A_\Sigma$ such that $\mathcal{A}_\Sigma \models \varphi(\underline{a})$, where $\varphi(\underline{a})$ is an $L_{\mathcal{A}_\Sigma}$-atomic sentence. But $a = [t]$ for some $t \in T_L$. Now $a^{\mathcal{A}_\Sigma} = a = [t] = t^{\mathcal{A}_\Sigma}$. By a previous lemma, $\mathcal{A}_\Sigma \models \varphi(\underline{a})$ if and only if $\mathcal{A}_\Sigma \models \varphi(t)$. $\varphi(t)$ is an atomic sentence, so $\mathcal{A}_\Sigma \models \varphi(t)$ if and only if $\Sigma \vdash \varphi(t)$. So in summary, in order to extend the above claim to $\sigma = \exists x\varphi(x)$ we need that whenever $\Sigma \vdash \exists x\varphi(x)$ there is a variable free term $t \in T_L$ such that $\Sigma \vdash \varphi(t)$. This is a property of Σ that is not necessarily going to be the case.

Definition 4.16. A Σ-*witness* for the sentence $\exists x\varphi(x)$, where $\varphi(x)$ is any L-formula, is $t \in T_L$ such that $\Sigma \vdash \varphi(t)$. We say that Σ has witnesses if whenever $\Sigma \vdash \exists x\varphi(x)$ then there exists a Σ-witness for $\exists x\varphi(x)$.

For example, ACF does not have witnesses, as the variable-free terms of ACF are exactly the integers, and so $\exists x(x \cdot x = 1 + 1)$ has no witness.

Theorem 4.17. *Suppose L has a constant symbol and Σ is consistent. Then the following are equivalent*

(i) *for each L-sentence σ, $\Sigma \vdash \sigma$ if and only if $\mathcal{A}_\Sigma \models \sigma$*

(ii) Σ is complete and has witnesses.

In particular, if Σ is consistent, complete, and has witnesses, then $\mathcal{A}_\Sigma \models \Sigma$.

Proof. We have already seen that (i) implies (ii). We need to show that (ii) is sufficient for (i). Assume (ii) and use induction on the number of logical symbols. If σ is atomic, then this is the above lemma. If σ is one of $\neg\sigma_1$ or $\sigma_1 \vee \sigma_2$ or $\sigma_1 \wedge \sigma_2$ then this is easy by induction. (Try it.) If $\sigma = \exists x\varphi(x)$, then suppose that $\Sigma \vdash \sigma$. Σ has witnesses, so there is $t \in T_L$ such that $\Sigma \vdash \varphi(t)$. By the induction hypothesis, $\mathcal{A}_\Sigma \models \varphi(\underline{t^{\mathcal{A}_\Sigma}})$, so $\mathcal{A}_\Sigma \models \sigma$. And so on...

If $\sigma = \forall x\varphi(x)$ then ... □

Lemma 4.18. *If Σ is consistent then $\Sigma \subseteq \Sigma^c$, where Σ^c is a complete set of L-sentences.*

Proof. Let $\mathcal{P}$ be the set of all consistent sets of L-sentences containing Σ. Partially order it by inclusion. Given a chain, the union is also consistent (as proofs are finite) and hence an upper bound for the chain. Thus by Zorn's Lemma, $\mathcal{P}$ has a maximal element Σ^c. Σ^c is complete because given any σ, either $\sigma \in \Sigma^c$ or $\neg\sigma \in \Sigma$. (If $\Sigma^c \nvdash \sigma$ then, by a corollary above, $\Sigma^c \cup \{\neg\sigma\}$ is not inconsistent, so $\Sigma^c \cup \{\neg\sigma\} \in \mathcal{P}$ and hence by maximality $\neg\sigma \in \Sigma^c$.) □

Lemma 4.19. *Let Σ be a consistent set of L-sentences and c a fresh constant symbol not in L. Let $L_c = L \cup \{c\}$. Let $\varphi(y)$ be any L-formula.*

(i) If $\Sigma \vdash_{L_c} \varphi(c)$ then $\Sigma \vdash_L \varphi(y)$

(ii) If $\Sigma \vdash_L \exists y\varphi(y)$ then $\Sigma \cup \{\varphi(c)\}$ is a consistent set of L_c-sentences

Proof. (Sketch) Go through the proof of $\varphi(c)$ from Σ and replace every occurrence of c by a new variable z ($\neq y$ and does not occur in the proof). Check that we have a proof of $\varphi(z/y)$ from Σ in L. Notice that the only occurrences of c in this proof are in $\varphi(c)$ and in the logical axioms. Thus $\Sigma \vdash_L \varphi(z)$. For the second part, assume not. Then $\Sigma \cup \{\varphi(c)\} \vdash_{L_c} \bot$, which implies $\Sigma \vdash_{L_c} \varphi(c) \to \bot$ by the deduction lemma, which implies $\Sigma \vdash_L \varphi(y) \to \bot$. By (G) $\Sigma \vdash_L \exists y\varphi(y) \to \bot$, but $\Sigma \vdash_L \exists y\varphi(y)$, which implies that $\Sigma \vdash_L \bot$, a contradiction. □

Corollary 4.20. *Let Σ be consistent and let*

$$\Lambda = \{\sigma = \exists x\varphi(x) \mid \varphi(x) \text{ is an } L\text{-formula } \varphi(x) \text{ and } \Sigma \vdash \sigma\}$$

Let $\{c_\sigma \mid \sigma \in \Lambda\}$ be a set of new constant symbols such that $c_\sigma \neq c_{\sigma'}$ for any $\sigma \neq \sigma' \in \Lambda$. Let

$$L^w := L \cup \{c_\sigma \mid \sigma \in \Lambda\} \text{ and } \Sigma^w := \Sigma \cup \{\varphi(c_\sigma) \mid \exists x\varphi(x) \in \Lambda\}$$

Then Σ^w is a consistent set of L^w-sentences.

Proof. Repeated use of Lemma 4.18 (finitely many times, by contradiction). □

Note that we have not yet shown that Σ^w has witnesses!

We now complete the proof of completeness. Let Σ be a set of L-sentences. Set $L_0 = L$ and $\Sigma_0 = \Sigma$. Define

$$L_{n+1} = \begin{cases} L_n & \text{if } n \text{ is even} \\ L_n^w & \text{if } n \text{ is odd} \end{cases} \text{ and } \Sigma_{n+1} = \begin{cases} \Sigma_n^c & \text{if } n \text{ is even} \\ \Sigma_n^w & \text{if } n \text{ is odd} \end{cases}$$

Then $L_n \subseteq L_{n+1}$ and $\Sigma_n \subseteq \Sigma_{n+1}$ for all n, and Σ_n is a consistent set of L_n-sentences. Let

$$\Sigma_\infty := \bigcup_{n \geq 0} \Sigma_n \text{ and } L_\infty := \bigcup_{n \geq 0} L_n$$

Claim. Σ_∞ is a consistent set of L_∞-sentences

Proof. Any proof of $\bot$ from Σ_∞ in L_∞ is finite and so is a proof of $\bot$ from Σ_N from L_N for some N. □

Claim. Σ_∞ is complete.

Proof. Let σ be an L_∞-sentence. Then σ is an L_N-sentence for some N even. $L_{N+1} = L_N$ and $\Sigma_{N+1} = \Sigma_N^c$ is a complete set of L_N-sentences, so either $\Sigma_N^c \vdash_{L_N} \sigma$ or $\Sigma_N^c \vdash_{L_N} \neg\sigma$. Therefore either $\Sigma_\infty \vdash_{L_\infty} \sigma$ or $\Sigma_\infty \vdash_{L_\infty} \neg\sigma$. □

Claim. Σ_∞ has witnesses.

Proof. $\Sigma_\infty \vdash_{L_\infty} \exists x\varphi(x)$, where φ is an L_∞-formula implies that $\Sigma_N \vdash_{L_N} \exists x\varphi(x)$ for some large N odd. $L_{N+1} = L_N^w$ and $\Sigma_{N+1} = \Sigma_N^w$ is a consistent set of L_N^w-sentences. By construction of L_N^w from L_N, there is a constant symbol $c \in L_N^w$ such that $\varphi(c) \in \Sigma_N^w$ and $\Sigma_N^w \vdash_{L_N^w} \varphi(c)$. This implies that $\Sigma_\infty \vdash_{L_\infty} \varphi(c)$, so $\varphi(x)$ has a witness in L_∞. □

From these claims and the theorem above, $\mathcal{A}_{\Sigma_\infty} \models \Sigma_\infty$.

Definition 4.21. Suppose that L^* is a language extending L and $\mathcal{A} = (A, \dots)$ is an L^*-structure. The reduct of $\mathcal{A}$ to L, denoted by $\mathcal{A}|_L$ is the L-structure with universe A and each symbol in L is interpreted exactly as it was in $\mathcal{A}$. Essentially, we are forgetting the interpretation of the symbols in $L^* \setminus L$.

We must check 2 things

(i) $t^{\mathcal{A}|_L} = t^{\mathcal{A}}$ for any L-term

(ii) For any L_A-sentence σ, $\mathcal{A}|_L \models_L \sigma$ if and only if $\mathcal{A} \models_{L^*} \sigma$

Then $\mathcal{A}_{\Sigma_\infty} \models \Sigma_\infty$ implies that $\mathcal{A}_{\Sigma_\infty}|_L \models \Sigma$, so $\mathcal{A}_{\Sigma_\infty}|_L$ is a model of Σ.

5 Model Theory

Corollary. (Compactness Theorem) *Let Σ be a set of L-sentences such that every finite subset has a model. Then Σ has a model.*

Corollary. (Countable Löwenheim-Skolem Theorem) *Suppose that L is a countable language and Σ a set of L-sentences has a model. Then Σ has a countable model. (By the cardinality of a structure we mean the cardinality of its universe.)*

Proof. Σ is consistent. Var is countable and L is countable, hence the set of L-sentences is countable. In particular, $L \cup \{c_\sigma \mid \Sigma \vdash \sigma \text{ where } \sigma = \exists x\varphi(x)\}$ is countable. This implies that Σ^w is countable and so is L^w. Therefore L_∞ and T_{L_∞} are countable, and so $T_{L_\infty}/\sim_{\Sigma_\infty}$ is countable, and it is the underlying universe of $\mathcal{A}_{\Sigma_\infty}|_L \models \Sigma$. □

Theorem. (Generalized Löwenheim-Skolem Theorem) *Suppose L has size at most κ, and suppose that Σ, a set of L-sentences, has an infinite model. Then Σ has a model of size κ.*

This theorem is important because it implies that model theory cannot distinguish between cardinals. Model theory is not set theory.

Proof. Let $\{c_\lambda \mid \lambda < \kappa\}$ be a family of new constant symbols, pairwise distinct and not in L. Let $L' = L \cup \{c_\lambda \mid \lambda < \kappa\}$ and let $\Sigma' = \Sigma \cup \{c_\lambda \neq c_{\lambda'} \mid \lambda < \lambda' < \kappa\}$. We claim that Σ' is consistent. If not, then some finite subcollection $\Lambda \subseteq \Sigma'$ is inconsistent. Let $c_{\lambda_1}, \ldots, c_{\lambda_n}$ be the new constants that actually appear in Λ. Then $\Lambda \subseteq \Sigma \cup \{c_{\lambda_i} \neq c_{\lambda_j} \mid i \neq j, 1 \leq i < j \leq n\}$. Σ has an infinite model $\mathcal{A}$. Make $\mathcal{A}$ into an $L \cup \{c_{\lambda_1}, \ldots, c_{\lambda_n}\}$-structure by interpreting $c_{\lambda_1}, \ldots, c_{\lambda_n}$ as distinct elements (fix any such choice). Call this structure $\mathcal{A}_n$. Then $\mathcal{A}_n \models \Sigma \cup \{c_{\lambda_i} \neq c_{\lambda_j} \mid i \neq j, 1 \leq i < j \leq n\}$, so $\mathcal{A}_n \models \Lambda$, which is a contradiction.

Now L' has size at most κ, so there are at most κ many L'-sentences, which implies that L'^w has cardinality at most κ, and so L'_∞ has cardinality at most κ. As before, $T_{L'_\infty} / \sim_{\Sigma'_\infty} = \mathcal{A}_{\Sigma'_\infty}|_L \models \Sigma'$. Therefore Σ' has a model of size at most κ. But any model of Σ' has κ-many elements, so Σ' has a model of size exactly κ, call it $\mathcal{A}$. Then $\mathcal{A} \models \Sigma$. □

Definition 5.1. Let $\mathcal{A} = (A, \ldots)$, $\mathcal{B} = (B, \ldots)$ be L-structures. A *homomorphism* $h : \mathcal{A} \to \mathcal{B}$ is a map $h : A \to B$ such that

(i) for each m-ary relation symbol $R \in L^{rel}$ and any $(a_1, \ldots, a_m) \in A^m$

$$(a_1, \ldots, a_m) \in R^{\mathcal{A}} \text{ implies } (h(a_1), \ldots, h(a_n)) \in R^{\mathcal{B}}$$

(ii) for each m-ary function symbol $F \in L^{fun}$ and any $(a_1, \ldots, a_m) \in A^m$

$$h(F^{\mathcal{A}}(a_1, \ldots, a_m)) = F^{\mathcal{B}}(h(a_1), \ldots, h(a_n))$$

If you replace the implication above with "if and only if" in (i) then we say that h is a strong homomorphism. An *embedding* is an injective strong homomorphism. An *isomorphism* is a surjective embedding. An *automorphism* of $\mathcal{A}$ is an isomorphism from $\mathcal{A}$ to itself. If there exists an isomorphism $h : \mathcal{A} \to \mathcal{B}$ then we say $\mathcal{A}$ and $\mathcal{B}$ are *isomorphic* and we write $\mathcal{A} \cong \mathcal{B}$.

For example, if $\mathcal{A} \subseteq \mathcal{B}$ then the inclusion map $\mathcal{A} \hookrightarrow \mathcal{B}$ is an embedding. The ring homomorphisms from algebra are exactly the (strong) homomorphisms of L_{ring}-structures, and similarily for groups, fields, vector spaces, etc. We think of automorphisms as acting in the natural way on A^n for all n and on subsets of the universe, for example $h(a_1, \ldots, a_n) = (h(a_1), \ldots, h(a_n))$.

Proposition 5.2. *Suppose that $h : \mathcal{A} \to \mathcal{B}$ is an isomorphism, $\varphi = \varphi(x_1, \ldots, x_n)$ is an L-formula and $a = (a_1, \ldots, a_1) \in A^n$. Then*

$$\mathcal{A} \models \varphi(\underline{a}) \text{ if and only if } \mathcal{B} \models \varphi(\underline{h(a)})$$

In particular, a sentence is true in $\mathcal{A}$ if and only if it is true in $\mathcal{B}$.

Proof. We claim that if $t = t(x_1, \ldots, x_n)$ is an L-term and $a = (a_1, \ldots, a_1) \in A^n$ then $h(t^{\mathcal{A}}(a)) = t^{\mathcal{B}}(h(a))$, for any homomorphism $h : \mathcal{A} \to \mathcal{B}$. This is a straightforward induction on the length of t. $t = x_i$ implies that $h(t^{\mathcal{A}}(a)) = h(a_i) = t^{\mathcal{B}}(h(a))$. If $t = F\, t_1 \cdots t_m$ then we use the fact that h commutes with F and induction to get our conclusion.

For the proof of the proposition, use proof by induction on the number of logical symbols in φ. If φ is $\top$ or $\bot$, then we are clearly done. If $t = R\, t_1 \cdots t_m$ then by definition, $\mathcal{A} \models \varphi(\underline{a})$ if and only if $(t_1^{\mathcal{A}}(a), \ldots, t_m^{\mathcal{A}}(a)) \in R^{\mathcal{A}}$, which happens if and only if $(h(t_1^{\mathcal{A}}(a)), \ldots, h(t_m^{\mathcal{A}}(a))) \in R^{\mathcal{B}}$. By the above claim, $\mathcal{A} \models \varphi(\underline{a})$ if and only if $(t_1^{\mathcal{B}}(h(a)), \ldots, t_m^{\mathcal{B}}(h(a))) \in R^{\mathcal{B}}$, that is, $\mathcal{B} \models \varphi(\underline{h(a)})$. If φ is $t_1 = t_2$ then the conclusion follows because h is injective. Now if φ is not atomic there then are five cases. Negation, disjunction, and conjunction are easy. If φ is $\exists y \psi(x_1, \ldots, x_n, y)$ then $\mathcal{A} \models \varphi(\underline{a})$ if and only if there exists $c \in A$ such that $\mathcal{A} \models \psi(\underline{a}_1, \ldots, \underline{a}_1, \underline{c})$. This happens if and only if there is $b \in B$ such that $\mathcal{B} \models \psi(\underline{ha_1}, \ldots, \underline{ha_1}, \underline{b})$ by induction and the fact that h is surjective. For the last case, use the contrapositive and the case above. □

Notation. Define $\mathrm{Aut}(\mathcal{A})$ to be the group of automorphisms of $\mathcal{A}$ (under composition). If $C \subseteq A$ then $\mathrm{Aut}_C(\mathcal{A})$ denotes the automorphisms of $\mathcal{A}$ which fix C pointwise.

Corollary 5.3. *If $C \subseteq A$ and S is a C-definable set in $\mathcal{A} = (A, \ldots)$ then $h(S) = S$ for all $h \in Aut_C(\mathcal{A})$.*

Proof. Write $S = \varphi^{\mathcal{A}} \subseteq A^n$ where $\varphi = \varphi(\underline{c}_1, \ldots, \underline{c}_m, x_1, \ldots, x_n)$, where $c_1, \ldots, c_m \in C$. For any $h \in \mathrm{Aut}(\mathcal{A})$ then $h(S)$ is defined by $\varphi(\underline{h(c_1)}, \ldots, \underline{h(c_m)}, x_1, \ldots, x_n)$. (Check this.) So an automorphism applied to a definable set is again a definable set, defined by the formula obtained by applying the automorphism to the parameters. In particular, if h fixes C then $h(S)$ is defined by $\varphi(\underline{c}_1, \ldots, \underline{c}_m, x_1, \ldots, x_n)$, and so $h(S) = S$. □

Example 5.4. $\{\sqrt{2}\}$ is not 0-definable in $(\mathbb{R}, <, 0, 1)$ since there are automorphisms of $(\mathbb{R}, <, 0, 1)$ that move $\sqrt{2}$. But $\sqrt{2}$ is 0-definable in $(\mathbb{R}, +, \times, -, ^{-1}, 0, 1)$.

Definition 5.5. By an *L-theory* (or *theory*) we mean a set of L-sentences. A theory Σ is *complete* if $\Sigma \vdash \sigma$ or $\Sigma \vdash \neg\sigma$ for all L-sentences σ. Equivalently, given any two models $\mathcal{A}$, $\mathcal{B}$, of Σ then

$$\{\sigma \mid \mathcal{A} \models \sigma\} = \{\sigma \mid \mathcal{B} \models \sigma\} \ (= \{\sigma \mid \Sigma \vdash \sigma\})$$

Theorem. (Vaught test) *Suppose a L-theory Σ has a model and all models of Σ are infinite. If there exists $\kappa \geq card(L)$ such that any two models of size κ are isomorphic, then Σ is complete.*

Definition 5.6. We say that a theory is *κ-categorical* (or *categorical in powerκ*) if any two models of size κ are isomorphic.

Proof. Suppose that Σ is not complete, say $\Sigma \nvdash \sigma$ and $\Sigma \nvdash \neg\sigma$. Then $\sigma_1 = \Sigma \cup \{\sigma\}$ and $\Sigma_2 = \Sigma \cup \{\neg\sigma\}$ are both consistent and they both have infinite models. By the generalized L-S, they have models of size κ. Say $\mathcal{A}_1 \models \Sigma_1$ and $\mathcal{A}_2 \models \Sigma_2$. But $\mathcal{A}_1 \models \Sigma$ and $\mathcal{A}_2 \models \Sigma$, so $\mathcal{A}_1 \cong \mathcal{A}_2$. In particular, $\mathcal{A}_1 \models \sigma$ implies $\mathcal{A}_2 \models \sigma$ and $\mathcal{A}_2 \models \neg\sigma$, which is a contradiction. □

Example 5.7. $L = L_{order} = \{<\}$. Let DLO be the theory

$$\begin{gathered}\{\ \forall x(x \not< x),\\ \forall xyz(x < y \wedge y < z \to x < z),\\ \forall xy(x < y \vee x = y \vee y < x),\\ \forall xy(x < y \to \exists z(x < z < y)),\\ \forall x \exists y \exists z(x < y \wedge z < x\ )\}\end{gathered}$$

This is the theory of dense linear orderings without endpoints.

Theorem 5.8. (Cantor) *Any two countable dense linear orderings without endpoints are isomorphic. That is, DLO is $\aleph_0$-categorical.*

Proof. Typical "back-and-forth" argument. Let $(A, <)$ and $(B, <)$ be countable models of DLO. Say $A = \{a_n \mid n \in \mathbb{N}\}$ and $B = \{b_n \mid n \in \mathbb{N}\}$. We inductively define partial maps $\{h_n \mid n \in \mathbb{N}\}$ from A to B such that

(i) h_n is a bijection between a finite subset of A and a finite subset of B which preserves $<$

(ii) h_{n+1} extends h_n (denoted $h_n \subseteq h_{n+1}$)

(iii) $\bigcup_{n\in\mathbb{N}} \mathrm{dom}(h_n) = A$ and $\bigcup_{n\in\mathbb{N}} \mathrm{range}(h_n) = B$

Once we have these we set $h = \bigcup_{n\in\mathbb{N}} h_n : A \to B$, which is a $<$-preserving bijection. By $<$-preserving we mean $a < b \iff h(a) < h(b)$, and so h is an isomorphism.

Take $h_0 : \{a_0\} \to \{b_0\} : a_0 \mapsto b_0$. Suppose that we have defined $h_0, \ldots, h_n$ satisfying (i) and (ii).

"forth" If n is even, let k be the least such that $a_k \notin \mathrm{dom}(h_n)$. Let l be the least such that b_l is situated with respect to $\mathrm{range}(h_n)$ exactly as a_k is situated with respect to $\mathrm{dom}(h_n)$. That is, $a < a_k$ if and only if $h(a) < b_l$ for all $a \in \mathrm{dom}(h_n)$. Such an element $b_l \in B$ exists because of denseness and no endpoints. Set

$$h_{n+1} : \mathrm{dom}(h_n) \cup \{a_k\} \to \mathrm{range}(h_n) \cup \{b_l\}$$

by extending h_n and taking a_k to b_l. So h_{n+1} satisfies (i) and (ii).

"back" If n is odd, choose l least such that $b_l \notin \mathrm{range}(h_n)$. Let k be least such that $b < b_l$ if and only if $h_n^{-1}(b) < a_k$ for all $b \in \mathrm{range}(h_n)$. Set h_{n+1} exactly as above. Note that (iii) holds for $\{h_n \mid n \in \mathbb{N}\}$ as the domain and range strictly increase at each stage and we didn't "skip" anything. (Prove this, use induction.) In fact, we have shown that given any $r_1 < \cdots < r_k \in A$ and $s_1 < \cdots < s_k \in B$ there is an isomorphism $h : (A, <) \to (B, <)$ such that $h(r_i) = s_i$ for $i = 1, \ldots, k$, by taking $h_0(r_i) = s_i$ for each i and continuing as above.

□

Corollary 5.9. *DLO is complete.*

For example, there are no L-sentences distinguishing $(\mathbb{Q}, <)$ from $(\mathbb{R}, <)$.

Proof. By Vaught's test. (DLO has only infinite models, and it has at least one model (e.g. $(\mathbb{Q}, <)$) and $\aleph_0 \geq \mathrm{card}(L) = 1$ and all models of size $\aleph_0$ isomorphic.) □

Example 5.10. Fix a field F and let VS_F^∞ be the theory of infinite vector spaces over F in the language $L_{vs(F)}$. First, a few facts. If V and W are F-vector spaces then

(i) V has a basis $B \subseteq V$ such that every $v \in V$ can be written uniquely as $v = \sum_{b\in B} f_b b$ where $\{f_b \mid b \in B\}$ are scalars and only finitely many of them are non-zero.

(ii) Any two bases have the same cardinality

(iii) If B is a basis for V and C is a basis for W then any bijection $B \to C$ extends to an isomorphism $V \to W$. (Notice that isomorphism in the linear algebra sense is the same as isomorphism in the sense of $L_{vs(F)}$-structures.)

(iv) If B is a basis for V then $\mathrm{card}(V) = \mathrm{card}(B) \times \mathrm{card}(F)$ provided one of F or B is infinite. If both F and B are finite, then $\mathrm{card}(V) = \mathrm{card}(B)^{\mathrm{card}(F)}$ In particular, if $\mathrm{card}(B) > \mathrm{card}(F)$ then $\mathrm{card}(V) = \mathrm{card}(B)$ provided B is infinite.

Theorem 5.11. *VS_F^∞ is complete.*

Proof. Let $\kappa > \mathrm{card}(F)$ be an infinite cardinal. Then any two models of VS_F^∞ of cardinality κ have the same dimension, κ, by part (iv) above. By part (iii), they are isomorphic. By Vaught's test VS_F^∞ is complete. □

A similar idea shows that ACP_p is complete (see assignment 2 question 3).

Theorem 5.12. (Morely's Theorem) *Let L be a countable language. If Σ is an L-theory which is κ-categorical for some uncountable cardinal κ then Σ is λ-categorical for all $\lambda > \aleph_0$.*

6 Definability Theory

Notation. If $\mathcal{A}$ is an L-structure then we write $|\mathcal{A}|$ to denote it's universe.

Definition 6.1. Let $\mathcal{A} = (A, \dots)$ be an L-structure. A (partial) map $f : A^n \to A^m$ is *definable* if its graph $\Gamma(f) := \{(a, b) \in A^{n+m} \mid f(a) = b\}$ is definable. Similarily, for any $C \subseteq A$ we say that f is *C-definable* or *definable over C* if its graph is C-definable.

Example 6.2. (i) Addition is not definable in $(\mathbb{Q}, <)$.

Proof. Suppose that $\Gamma(+)$ is defined by the $L_{\mathbb{Q}}$-formula $\phi(r_1, \dots, r_k, x_1, x_2, x_3)$ and take $R := \{r_1, \dots, r_k\}$, so that $\Gamma(+)$ is R-definable. Without loss of generality we may assume $r_1 < \cdots < r_k$. Let $a > b > |r_k| + 1$. The back-and-forth construction in the proof of Cantor's theorem shows that since $r_1 < \cdots < r_k < b < a < a + b$ and $r_1 < \cdots < r_k < b < a < a + b + 1$ there is an automorphism h of $(\mathbb{Q}, <)$ which fixes $R \cup \{a, b\}$ and takes $a + b$ to $a + b + 1$. It follows that $h \in \mathrm{Aut}_R(\mathbb{Q})$ with $h(a, b, a + b) = (a, b, a + b + 1)$. Therefore $h(\Gamma(+)) \neq \Gamma(+)$, so $\Gamma(+)$ is not R-definable. □

(ii) Multiplication on $\mathbb{R}$ is not 0-definable in $\mathcal{R} = (\mathbb{R}, +, -, 0)$

Proof. Suppose that $\Gamma(\times)$ is defined by some L-formula $\phi(x_1, x_2, x_3)$. Let $h : \mathbb{R} \to \mathbb{R} : x \mapsto 2x$. This is an automorphism of $\mathcal{R}$. But $h(1, 2, 2) = (2, 4, 4)$ so $h(\Gamma(\times)) \neq \Gamma(\times)$. □

Definition 6.3. Let Σ be an L-theory. We say that two formulas φ and ψ are *Σ-equivalent* if $\Sigma \models \varphi \leftrightarrow \psi$. Equivalently, if in every model $\mathcal{A} \models \Sigma$, φ and ψ define the same set, $\varphi^{\mathcal{A}} = \psi^{\mathcal{A}}$.

Definition 6.4. An L-theory Σ is said to have *quantifier elimination* (or *admit* QE) if every L-formula $\varphi(x_1, \dots, x_n)$ is Σ-equivalent to a quantifier-free L-formula $\psi(x_1, \dots, x_n)$. In particular, every L-sentence is Σ-equivalent to a quantifier-free L-sentence.

Lemma 6.5. *Suppose that Σ is consistent and there exists an L-structure which embeds in every model of Σ. Then if Σ admits QE, it is complete.*

Proof. Suppose that $\mathcal{C}$ is an L-structure embedding in every model of Σ. Let $\mathcal{A}$, $\mathcal{B}$ be models of Σ. We will show that $\mathcal{A}$, $\mathcal{B}$ satisfy the same L sentences. Let σ be an L-sentence. By QE there is an L-sentence γ such that $\Sigma \models \sigma \leftrightarrow \gamma$. Therefore $\mathcal{A} \models \sigma$ if and only if $\mathcal{A} \models \gamma$, if and only if $\mathcal{C} \models \gamma$ (since $\mathcal{C} \subseteq \mathcal{A}$ and γ is quantifier-free, see assignment 1), if and only if $\mathcal{B} \models \gamma$, if and only if $\mathcal{B} \models \sigma$. □

Example 6.6. ACF_0 is consistent and $(\mathbb{Q}, 0, 1, +, \times, -)$ embeds in every field of characteristic 0. If ACF_0 admits QE then it is complete.

Example 6.7. In $(\mathbb{R}, <, 0, 1, +, \times, -)$ the L-formula $\exists y(ay^2 + by + c = 0)$ is equivalent to

$$(a \neq 0 \wedge b^2 - 4ac \geq 0) \vee (a = 0 \wedge b \neq 0) \vee (a = 0 \wedge b = 0 \wedge c = 0)$$

Definition 6.8. A *primative existential formula* is a formula of the form $\exists y \phi(x_1, \dots, x_n, y)$, where y is a single variable and $\phi(x_1, \dots, x_n, y)$ is a conjunction of atomic or negated atomic formulae. (Atomic and negated atomic formulae are often called literals.)

Lemma 6.9. *Σ admits QE if and only if every primative existential formula $\psi(x_1, \dots, x_n)$ is Σ-equivalent to a quantifier-free formula $\theta(x_1, \dots, x_n)$.*

Proof. The forward implication is obvious. Let $\varphi(x_1, \dots, x_n)$. Proof of the reverse implication by induction on the number of symbols in φ. If φ is atomic we are done. If φ is one of $\neg\varphi_1$, $\varphi_1 \wedge \varphi_2$, or $\varphi_1 \vee \varphi_2$ then it follows by induction. If φ is $\exists y \gamma(x_1, \dots, x_n, y)$ where γ is an L-formula, then by in the induction hypothesis γ is Σ-equivalent to some quantifier-free formula $\gamma^{qf}(x_1, \dots, x_n, y)$. So γ^{qf} is Σ-equivalent to a disjunction of conjunctions of literals

$$\gamma^{qf} = \bigvee_{i=1}^{m} \gamma_i^{qf}(x_1, \dots, x_n, y)$$

Now $\Sigma \models \varphi \leftrightarrow \bigvee_{i=1}^m \exists y(\gamma_i^{qf}(x_1, \dots, x_n, y))$. So $\varphi = \bigvee_{i=1}^m \exists y(\gamma_i^{qf}(x_1, \dots, x_n, y))$ since $\exists$ commutes with $\vee$. By assumption for each i there is a quantifier-free formula $\theta(x_1, \dots, x_n, y)$ that is Σ-equivalent to $\exists y \gamma_i^{qf}(x_1, \dots, x_n, y)$. Finish this. □

Definition 6.10. Suppose that $\mathcal{A} = (A, \dots)$ and $C \subseteq A$. The *substructure generated by* C, denoted $\langle C \rangle$, is the L-structure given by

- universe of $\langle C \rangle$ is $\{t^{\mathcal{A}}(c_1, \dots, c_n) \mid t(x_1, \dots, x_n)$ is an L-term, $c_1, \dots, c_n \in C\}$
- $R \in L^{rel}$ m-ary, take $R^{\langle C \rangle} := R^{\mathcal{A}} \cap |\langle C \rangle|^m$
- $F \in L^{fun}$ m-ary, take $F^{\langle C \rangle} := F^{\mathcal{A}}|_{|\langle C \rangle|^m}$

Check that $|\langle C \rangle| \subseteq A$ is the smallest subset of A containing C and closed under functions. Also check that $\langle C \rangle \subseteq \mathcal{A}$. We say that a substructure $\mathcal{C} \subseteq \mathcal{A}$ is finitely generated if $\mathcal{C} = \langle C \rangle$ for some $C \subseteq A$ finite.

Example 6.11. $\mathcal{R} = (R, 0, 1, +, -, \times)$ commutative ring, $A \subseteq R$, then $\langle R \rangle = (\mathbb{Z}[A], 0, 1, +, -, \times)$ where $\mathbb{Z}[A] = \{P(a_1, \dots, a_n) \mid a_1, \dots, a_n \in A,\ P(X_1, \dots, X_n) \in \mathbb{Z}[X_1, \dots, X_n]\}$.

Lemma 6.12. *Suppose that $\mathcal{A} = (A, \dots)$ and $\mathcal{B} = (B, \dots)$ are L-structures and $a_1, \dots, a_n \in A$ and $b_1, \dots, b_n \in B$. The following are equivalent:*

(i) For any quantifier-free formula $\phi(x_1, \dots, x_n)$, $\mathcal{A} \models \phi(\underline{a}_1, \dots, \underline{a}_n)$ if and only if $\mathcal{B} \models \phi(\underline{b}_1, \dots, \underline{b}_n)$

(ii) There exists an isomorphism $h : \langle \{a_1, \dots, a_n\} \rangle \to \langle \{b_1, \dots, b_n\} \rangle : a_i \mapsto b_i$

Proof. Define h as follows. $h(t^{\mathcal{A}}(a_{i_1}, \dots, a_{i_m})) := t^{\mathcal{B}}(b_{i_1}, \dots, b_{i_m})$. In particular, $h(a_i) = b_i$. Use (i) to show that h is an isomorphism from $\langle \{a_1, \dots, a_n\} \rangle$ to $\langle \{b_1, \dots, b_n\} \rangle$. h is well-defined and a bijection since if $e, f \in \langle \{a_1, \dots, a_n\} \rangle$ then $e = t^{\mathcal{A}}(a_{i_1}, \dots, a_{i_m})$, $f = s^{\mathcal{A}}(a_{i_1}, \dots, a_{i_{m'}})$. Then $e = f$ if and only if $\mathcal{A} \models t^{\mathcal{A}}(a_{i_1}, \dots, a_{i_m}) = s^{\mathcal{A}}(a_{i_1}, \dots, a_{i_{m'}})$, which happens if and only if $\mathcal{B} \models t^{\mathcal{A}}(b_{i_1}, \dots, b_{i_m}) = s^{\mathcal{A}}(b_{i_1}, \dots, b_{i_{m'}})$ by (i). This says that $h(e) = h(f)$, so h is well-defined. Check that h is a strong homomorphism.

Let $\mathcal{A}_0 = \langle \{a_1, \dots, a_n\} \rangle$ and $\mathcal{B}_0 = \langle \{b_1, \dots, b_n\} \rangle$. If $\phi(x_1, \dots, x_n)$ is any quantifier-free formula, then $\mathcal{A} \models \phi(\underline{a}_1, \dots, \underline{a}_n)$ if and only if $\mathcal{A}_0 \models \phi(\underline{a}_1, \dots, \underline{a}_n)$ since $\phi(\underline{a}_1, \dots, \underline{a}_n)$ is a quantifier-free $L_{\mathcal{A}_0}$-sentence and $\mathcal{A}_0 \subseteq \mathcal{A}$. This happens if and only if $\mathcal{B}_0 \models \phi(\underline{b}_1, \dots, \underline{b}_n)$ as there is an isomorphism between $\mathcal{A}_0$ and $\mathcal{B}_0$ taking a_i to b_i, and this is equivalent to $\mathcal{B}_0 \models \phi(\underline{b}_1, \dots, \underline{b}_n)$. □

Theorem 6.13. *Let Σ be an L-theory and $\phi(x_1, \dots, x_n)$ an L-formula. The following are equivalent:*

(i) ϕ is Σ-equivalent to a quantifier-free formula $\theta(x_1, \dots, x_n)$

(ii) Given any $\mathcal{A}, \mathcal{B} \models \Sigma$ and finitely generated substructures $\mathcal{A}_0 \subseteq \mathcal{A}$ and $\mathcal{B}_0 \subseteq \mathcal{B}$ with an isomorphism h between them, $\mathcal{A} \models \phi(\underline{a}_1, \dots, \underline{a}_n)$ implies that $\mathcal{B} \models \phi(\underline{ha_1}, \dots, \underline{ha_n})$ for any $a_1, \dots, a_n \in |\mathcal{A}_0|$

Proof. See text. □

Putting the last two propositions together yields

Corollary 6.14. (Criterion for QE) *Given*

(i) A conjunction of atomic or negated atomic L-formulae $\phi(x_1, \dots, x_m, y)$

(ii) Models $\mathcal{A}, \mathcal{B}$ *of* Σ

(iii) Substructures $\mathcal{A}_0 = \langle \{a_1, \dots, a_n\} \rangle \subseteq \mathcal{A}$, $\mathcal{B}_0 = \langle \{b_1, \dots, b_n\} \rangle \subseteq \mathcal{B}$

(iv) An isomorphism $h : \mathcal{A}_0 \to \mathcal{B}_0$ *taking* a_i *to* b_i *for* $i = 1, \dots, n$

Then Σ *has QE if* $\mathcal{A} \models \exists y \phi(\underline{a_{i_1}}, \dots, \underline{a_{i_m}}, y)$ *implies* $\mathcal{B} \models \exists y \phi(\underline{b_{i_1}}, \dots, \underline{b_{i_m}}, y)$.

Proof. We need only eliminate primative existential formulas. The last theorem gives us a sufficient condition for doing so. □

Proposition 6.15. *DLO has QE*

Proof. Let $\mathcal{A} = (A, \dots)$, $\mathcal{B} = (B, \dots)$ be models of DLO, $a_1 < \cdots < a_n \in A$, and $b_1 < \cdots < b_n \in B$. Let $\mathcal{A}_0$ and $\mathcal{B}_0$ be generated by the a_i's and b_i's respectively, and let h be an isomorphism between them taking a_i to b_i for each $i - 1, \dots, n$. Let $\phi(x_1, \dots, x_m, y)$ be a conjuction of atomic and negated atomic formulas. We need to show that if $\mathcal{A} \models \exists\, y \phi(\underline{a_{i_1}}, \dots, \underline{a_{i_m}}, y)$ then $\mathcal{B} \models \exists\, y \phi(\underline{b_{i_1}}, \dots, \underline{b_{i_m}}, y)$. Let $a \in A$ be such that $\mathcal{A} \models \phi(\underline{a_{i_1}}, \dots, \underline{a_{i_m}}, \underline{a})$. As we have seen in the proof that DLO is $\aleph_0$-categorical, there is some $b \in B$ such that h extends to $\mathcal{A}_0' := \langle \{a_1, \dots, a_n, a\} \rangle \cong \mathcal{B}_0' := \langle \{b_1, \dots, b_n, b\} \rangle$ that takes a to b. Then $\mathcal{A}_0' \models \phi(\underline{a_{i_1}}, \dots, \underline{a_{i_m}}, \underline{a})$ since it is a quantifier-free $L_{|\mathcal{A}_0|}$-sentence. Therefore $\mathcal{B}_0' \models \phi(\underline{b_{i_1}}, \dots, \underline{b_{i_m}}, \underline{b})$, so $\mathcal{B} \models \phi(\underline{b_{i_1}}, \dots, \underline{b_{i_m}}, \underline{b})$. Therefore $\mathcal{A} \models \exists\, y \phi(\underline{a_{i_1}}, \dots, \underline{a_{i_m}}, y)$ implies $\mathcal{B} \models \exists\, y \phi(\underline{b_{i_1}}, \dots, \underline{b_{i_m}}, y)$. Thus DLO admits QE. □

Corollary 6.16. *The definable subsets of* A *in* $(A, <) \models DLO$ *are exactly the finite unions of points and intervals. Here an interval is open and possibly infinite.*

Proof. It's clear that the finite unions of points and intervals are definable. For the other direction, by QE, it suffices to show that the sets defined by atomic formulas are finite unions of points or intervals (Check this). □

Definition 6.17. A theory Σ in a language containing $<$ extending the theory of total orderings is called *o-minimal* if every definable subset of $\mathcal{A}$ is a finite union of points and intervals for any $\mathcal{A} \models \Sigma$.

Example 6.18. The theory of $(\mathbb{R}, 0, 1, <, +, -, \cdot)$ is o-minimal

Example 6.19. The theory of DLO is o-minimal

Proposition 6.20. *Let* $p = 0$ *or a prime.* ACF_p *has QE.*

Proof. Set $R = \mathbb{Z}$ if $p = 0$ and $R = \mathbb{F}_p$ otherwise. Let $\mathcal{K} = (K, 0, 1, +, -, \cdot) \models ACF_p$ and $a_1, \dots, a_n \in \mathcal{K}$. Then $\langle \{a_1, \dots, a_n\} \rangle = R[a_1, \dots, a_n]$. Let $\mathcal{L} = (L, 0, 1, +, -, \cdot) \models ACF_p$ and $b_1, \dots, b_n \in \mathcal{L}$ Then $\langle \{b_1, \dots, b_n\} \rangle = R[b_1, \dots, b_n]$. Let h be an isomorphism of rings $R[a_1, \dots, a_n] \to R[b_1, \dots, b_n]$.

We know h extends to an isomorphism h' of fraction fields $\mathbb{F}(a_1, \dots, a_n) \to \mathbb{F}(b_1, \dots, b_n)$. Here $\mathbb{F} = \mathbb{Q}$ if $p = 0$ and $\mathbb{F} = \mathbb{F}_p$ otherwise. Moreover, h' extends to an isomorphism h'' of algebraic closures $\overline{\mathbb{F}(a_1, \dots, a_n)} \to \overline{\mathbb{F}(b_1, \dots, b_n)}$ Consult Hungerford for details.

Let $\phi(x_1, \dots, x_n, y)$ be a conjunction of atomic and negated atomic L-formulas. Suppose $\mathcal{K} \models \exists y \phi(\underline{a_1}, \dots, \underline{a_n}, y)$; we need to show $\mathcal{L} \models \exists y \phi(\underline{b_1}, \dots, \underline{b_n}, y)$.

Let $c \in K$ such that $\mathcal{K} \models \phi(\underline{a_1}, \dots, \underline{a_n}, \underline{c})$.

Case 1: If $c \in \overline{\mathbb{F}(a_1, \dots, a_n)}$ then $h''(c)$ will satisfy $\phi(\underline{b_1}, \dots, \underline{b_n}, y)$ in $\overline{\mathbb{F}(b_1, \dots, b_n)}$. Since $\overline{\mathbb{F}(b_1, \dots, b_n)} \subseteq \mathcal{L}$ we are done.

Case 2: If $c \notin \overline{\mathbb{F}(a_1, \ldots, a_n)}$ we may in trouble, since h'' may not extend to c. Consider $L = \overline{\mathbb{F}(b_1, \ldots, b_n)}$. Since $\mathcal{K} \models \phi(\underline{a_1}, \ldots, \underline{a_n}, \underline{c})$, ϕ must be a conjunction of formulas coming from the following possibilities.

(a) $\psi(x_1, \ldots, x_n)$ where ψ is an atomic or negated atomic L-formula

(b) $\{Q_i(x_1, \ldots, x_n, y) = 0\}_{i=1\ldots\ell}$ where the Q_i are polynomials in $\mathbb{Z}[X_1, \ldots, X_n, Y]$ so $Q_i(a_1, \ldots, a_n, Y) \in \mathbb{F}(a_1, \ldots, a_n)[Y]$.

(c) $\{P_i(x_1, \ldots, x_n, y) \neq 0\}_{i=1\ldots m}$ where the P_i are non-zero polynomials

Let $d \in L$ such that $P_i(\underline{b_1}, \ldots, \underline{b_n}, d) \neq 0$ for $i = 1, \ldots, r$. This is easy since we have r finite, the P_i expressible as polynomials in one variable, and our language L infinite. And so we are done. Remember, we wanted to show $\mathcal{L} \models \exists y \phi(\underline{b_1}, \ldots, \underline{b_n}, y)$. From (a), this follows immediately - we have no y to worry about. From (b), we note h'' maps the roots of the $Q_i(a_1, \ldots, a_n, Y)$ to roots of $Q_i(b_1, \ldots, b_n, Y)$, so we can choose our c to be in the algebraic closure. From (c), everything follows from our choice of d.

□

Corollary 6.21. *ACF_p is complete.*

Proposition 6.22. *Suppose Σ has QE. Given models of Σ, $\mathcal{A} \subseteq \mathcal{B}$, and any L_A-sentence θ, $\mathcal{A} \models \theta$ if and only if $\mathcal{B} \models \theta$.*

Proof. Write $\theta = \phi(\underline{a_1}, \ldots, \underline{a_n})$ where $\phi(x_1, \ldots, x_n)$ is an L-formula. By QE there is a quantifier-free L-formula ψ such that $\Sigma \models \phi$ if and only if $\Sigma \models \psi$. From a previous assignment relating structures to substructures, the result follows. □

Definition 6.23. We say Σ is *model-complete* if whenever $\mathcal{A} \subseteq \mathcal{B}$ are models of Σ and θ is an $L_{|\mathcal{A}|}$-sentence

$$\mathcal{A} \models \theta \text{ if and only if } \mathcal{B} \models \theta$$

We note that QE implies model-completeness.

Example 6.24. The theory of fields of characteristic 0 does not have QE. Consider $\exists x(x^2 = 1)$ in $\mathbb{R} \subseteq \mathbb{C}$.

Corollary 6.25. *Every definable subset of K in $(K, 0, 1, +, -, \times) \models ACF_p$ is either finite or cofinite.*

Proof. The sets defined by atomic fomulas in one variable are of the form $P(X) = 0$ or $\bot$ or $\top$ where $P \in K[X]$. But a non-zero polynomial has only finitely many roots. The class of sets obtained from the atomically defined sets using $\cap$, $\cup$, and complement are exactly the finite and co-finite sets, by QE. □

Definition 6.26. If a theory Σ is such that in every model $\mathcal{A} \models \Sigma$, every definable subset of $|\mathcal{A}|$ is finite or cofinite then we say that Σ is *strongly minimal*.

Definition 6.27. $\mathcal{A} \subseteq \mathcal{B}$ we say that $\mathcal{A}$ is an *elementary substructure* of $\mathcal{B}$, denoted $\mathcal{A} \preceq \mathcal{B}$, if for every $L_{|\mathcal{A}|}$-sentence θ

$$\mathcal{A} \models \theta \text{ if and only if } \mathcal{B} \models \theta$$

A theory Σ is model-complete if for any models $\mathcal{A} \subseteq \mathcal{B}$ of Σ then $\mathcal{A} \preceq \mathcal{B}$.

We showed QE implies model-complete. For example, $(\mathbb{R}, +, -, 0, 1, \times) \subseteq (\mathbb{C}, +, -, 0, 1, \times)$ and not elementary. This implies that the theory of fields of characteristic zero does not have QE.

Corollary. (Hilbert's Nullstellensatz) *Suppose that K is an algebraically closed field. If $I \subseteq K[X_1, \ldots, X_n]$ is a prime ideal then there exists $a = (a_1, \ldots, a_n) \in K^n$ such that $P(a) = 0$ for all $P \in I$.*

Proof. Let $X = (X_1, \ldots, X_n)$. $K[X]/I$ is an integral domain, as I is a prime ideal. Let L be the fraction field of $K[X]/I$. Then $K \subseteq K[X]/I \subseteq L \subseteq L^{alg}$. By model-completeness $K \preceq L^{alg}$. Now I is finitely generated since $K[X]$ is Noetherian, say $I = \langle P_1, \ldots, P_m \rangle$.

$$L^{alg} \models \exists x_1, \ldots, x_n \left(\bigwedge_{i=1}^{m} P_i(x_1, \ldots, x_n) = 0 \right)$$

witnessed by $(X_1 + I, \ldots, X_n + I) =: (b_1, \ldots, b_n) \in L^{alg}$. Thus

$$K \models \exists x_1, \ldots, x_n \left(\bigwedge_{i=1}^{m} P_i(x_1, \ldots, x_n) = 0 \right)$$

because K is an elementary substructure. □

Corollary. (Lefschetz Principle) *Let σ be an L_{ring}-sentence. The following are equivalent*

(i) σ is true in $\mathbb{C}$

(ii) σ is true in some algebraically closed field of characteristic zero

(iii) σ is true in all algebraically closed fields of characteristic zero

(iv) For infinitely many primes p, σ is true in some algebraically closed field of characteristic p

(v) For infinitely many primes p, σ is true in all algebraically closed fields of characteristic p

(vi) For all but finitely many primes p, σ is true in some algebraically closed field of characteristic p

(vii) For all but finitely many primes p, σ is true in all algebraically closed fields of characteristic p

Proof. Exercise. □

Lemma 6.28. *Σ has QE if and only if every completion of Σ has QE.*

Proof. $\Rightarrow$ is clear. Let $\varphi(x_1, \ldots, x_n)$ be an L-formula. Let $c_1, \ldots, c_n$ be new constant symbols and $L^* = L \cup \{c_1, \ldots, c_n\}$. Define

$$\Lambda := \{\theta \mid \Sigma \cup \{\varphi(c_1, \ldots, c_n)\} \models \theta \text{ and } \theta \text{ is quantifier-free}\}$$

Claim. $\Sigma \cup \Lambda \models \varphi(c_1, \ldots, c_n)$

Let $\mathcal{A}^* \models \Sigma \cup \Lambda$ and $\mathcal{A} = \mathcal{A}^*|_L \models \Sigma$. Then the theory of $\mathcal{A}$, $\mathrm{Th}(\mathcal{A})$, is complete and it contains Σ. So $\mathrm{Th}(\mathcal{A})$ admits QE. Let $\theta(x_1, \ldots, x_n)$ be a quantifier-free L-formula such that $\mathrm{Th}(\mathcal{A}) \models \varphi \leftrightarrow \theta$. Then $\neg\theta(c_1, \ldots, c_n) \notin \Lambda$, so $\theta(c_1, \ldots, c_n) \in \Lambda$ and so $\mathcal{A}^* \models \theta(c_1, \ldots, c_n)$. Thus $\mathcal{A}^* \models \theta(c_1, \ldots, c_n) \rightarrow \varphi(c_1, \ldots, c_n)$, so $\mathcal{A}^* \models \varphi(c_1, \ldots, c_n)$.

By compactness, there is some $\theta(c_1, \ldots, c_n)$ such that $\Sigma \cup \{\theta\} \models \varphi(c_1, \ldots, c_n)$, which implies that $\Sigma \models \theta(c_1, \ldots, c_n) \rightarrow \varphi(c_1, \ldots, c_n)$. But also $\Sigma \cup \{\varphi\} \models \theta(c_1, \ldots, c_n)$ by definition of Λ, so $\Sigma \models \theta(c_1, \ldots, c_n) \leftrightarrow \varphi(c_1, \ldots, c_n)$. □

It follows that ACF has QE. Suppose that $\Lambda \supseteq ACF$ and Λ is complete. Let $\mathcal{A} \models \Lambda$, with p as the characteristic of $\mathcal{A}$. Then for all $\mathcal{B} \models \Lambda$, the characteristic of $\mathcal{B}$ is p, so Λ is a model of ACF_p. Since ACF_p has QE, Λ has QE.

7 Real Closed Fields

In studying the theory of $\mathbb{R}$ it turns out that the "right" language is $L_{o-ring} = \{<, 0, 1, -, +, \times\}$. Recall that $<$ is 0-definable in $(\mathbb{R}, 0, 1, -, +, \times)$ by defining $x \leq y$ to be the set defined by the formula $\exists z(y - x = z^2)$.

Definition 7.1. *RCOF* is the L_{o-ring}-theory consisting of

(i) Axioms of fields

(ii) Axioms of total ordering

(iii) $\forall xyz(x < y \to x + z < y + z)$ and $\forall xy(0 < x \wedge 0 < y \to 0 < xy)$

(iv) For each $n \in \mathbb{N}$ odd $\forall x_1 \dots x_n \exists y(y^n + x_1 y^{n-1} + \cdots + x_{n-1}y + x_n = 0)$

(v) $\forall x(0 < x \to \exists y(x = y^2))$

Definition 7.2. A *real ordered closure* of an ordered field $(F, <)$ is a real closed ordered field $(R, <)$ such that $(F, <) \subseteq (R, <)$ and $F \subseteq R \subseteq F^{alg}$.

For example, $(\overline{\mathbb{Q}}, <)$ is the real order closure of $(\mathbb{Q}, <)$.

Theorem 7.3. (Artin-Schreir) *Up to L_{o-ring}-isomorphism every ordered field has a unique real ordered closure.*

Theorem 7.4. (Intermediate Value Property) *Suppose $R \models RCOF$ and $P \in R[x]$ is any polynomial. If $a, b \in R$, $a < b$, are such that $P(a) < 0$ and $P(b) > 0$ then there is $c \in R$, $a < c < b$, such that $P(c) = 0$.*

Theorem 7.5. *RCOF has QE.*

Proof. Let $\mathcal{A} = (K, 0, 1, <, +, -, \times)$ and $\mathcal{B} = (K', 0, 1, <, +, -, \times)$ both be models of *RCOF*. Let $\mathcal{A}_0 = (\mathbb{Q}[a_1, \dots, a_n], 0, 1, <, +, -, \times) \subseteq \mathcal{A}$ and $\mathcal{B}_0 = (\mathbb{Q}[b_1, \dots, b_n], 0, 1, <, +, -, \times) \subseteq \mathcal{B}$ be finitely generated substructures and $h : \mathcal{A}_0 \to \mathcal{B}_0$ an L_{o-ring}-isomorphism taking a_i to b_i for $i = 1, \dots, n$.

Claim. h extends to an ordered field isomorphism $h' : \mathbb{Q}(a_1, \dots, a_n) \to \mathbb{Q}(b_1, \dots, b_n)$ that takes $\frac{x}{y} \mapsto \frac{h(x)}{h(y)}$.

It is clear that h' is a field isomorphism, and it is not hard to see using axiom (iii) that it is an L_{o-ring}-isomorphism. For example, suppose that $x, y \in \mathbb{Q}[a_1, \dots, b_n]$ and $y > 0$.

$$\frac{x}{y} > 0 \iff x > 0 \iff \frac{h(x)}{h(y)} > 0 \text{ (since } h(y) > 0) \iff h'\left(\frac{x}{y}\right) > 0$$

Claim. h' extends to an L_{o-ring}-isomorphism $h'' : (F, <) \to (F', <)$, where $F = \mathbb{Q}(a_1, \dots, a_n)^{alg} \cap K$ and $F' = \mathbb{Q}(b_1, \dots, b_n)^{alg} \cap K'$, the real ordered closures of these fields.

The proof follows by the uniqueness of real ordered closure.

Let $\phi(x_1, \dots, x_n, y)$ be a conjunction of atomic and negated atomic formulae. Suppose that $\mathcal{A} \models \exists y \phi(\underline{a_1}, \dots, \underline{a_n}, y)$. We want $\mathcal{B} \models \exists y \phi(\underline{b_1}, \dots, \underline{b_n}, y)$. Let $c \in K$ be such that $\mathcal{A} \models \phi(\underline{a_1}, \dots, \underline{a_n}, \underline{c})$. If $c \in F$ then clearly $h''(c) \in F' \subseteq K'$ and is such that $\mathcal{B} \models \phi(\underline{b_1}, \dots, \underline{b_n}, \underline{h''(c)})$.

We may assume that $c \in K \setminus F$. Let $P_1(y), \dots, P_l(y) \in F[y]$ be the polynomials that appear in $\phi(\underline{a_1}, \dots, \underline{a_n}, y)$.

Claim. There is a $d \in F'$ such that $P_j(c) = 0$ if and only if $P_j^{h''}(d) = 0$ and $P_j(c) < 0$ if and only if $P_j^{h''}(d) < 0$ for all $j = 1, \dots, l$.

We may assume that none of the P_i are the zero polynomial, as that case is obvious. Since $c \in K \setminus F$, c is not in the algebraic closure of $\mathbb{Q}(a_1, \dots, a_n)$ and hence for all $j = 1, \dots, l$ we have $P_j(c) \neq 0$.

Case 1: Suppose that for some i, j there is a root r of P_i and s of P_j with $r < c < s$. Furthermore, choose r and s so that none of the roots of $P_1, \ldots, P_l$ fall in the interval (r, s). By the intermediate value property, none of $P_1, \ldots, P_l$ change sign in the interval (r, s). Thus

$$\operatorname{sign}(P_j(c)) = \operatorname{sign}\left(P_j\left(\frac{r+s}{2}\right)\right) = \operatorname{sign}\left(P_j^{h''}\left(\frac{h''(r)+h''(s)}{2}\right)\right)$$

for all $j = 1, \ldots, l$. Let $d = \frac{h''(r)+h''(s)}{2}$. Then it satisfies the conclusions of the claim.

Case 2: ?? This proof is broken. Ask Cameron why.

By the third claim, $\mathcal{B} \models \phi(\underline{b_1}, \ldots, \underline{b_n}, \underline{d})$. (why?) An atomic formula with parameters from F is of the form (or equivalent to the form) $P(y) = 0$ or $P(y) < 0$, where $P(y) \in F[y]$. If $P_1, \ldots, P_l$ appear in $\phi(\underline{a_1}, \ldots, \underline{a_n}, y)$ then $P_1^{h''}, \ldots, P_l^{h''}$ are the polynomials that appear in $\phi(\underline{b_1}, \ldots, \underline{b_n}, y)$. □

Corollary 7.6. *RCOF is o-minimal.*

Corollary 7.7. *RCOF is complete.*

Proof. It suffices to show that there exists an L_{o-ring}-structure which embeds in every model of *RCOF*. Let $\mathcal{Z} = (\mathbb{Z}, <, 0, 1, +, -, \times)$ (with the usual ordered ring structure on $\mathbb{Z}$). If $\mathcal{A} = (R, \ldots) \models RCOF$ then the characteristic of R must be zero. We must show that $<^{\mathcal{A}}$ restricts to $<$ on the integers. To do this we will show that the positive elements of $\mathbb{Z}$ are positive elements in R. Well, $0 < 1$ by the axioms, so $0 < 2$, $0 < 3$,... in $\mathcal{A}$. Therefore $\mathcal{Z} \subseteq \mathcal{A}$. □

It follows that *RCOF* axiomatises $\operatorname{Th}(\mathbb{R}, <, 0, 1, +, -, \times)$. What about $\operatorname{Th}(\mathbb{R}, 0, 1, +, -, \times)$? Note that it does not have QE, as the definable sets are the finite or cofinite sets, but if it had QE then intervals would be definable, a contradiction.

Lemma 7.8. *In a model of RCOF the positive elements are exactly the non-zero squares.*

Proof. Positive elements are squares by the axioms. Suppose that $x < 0$ and $x = b^2$ in some $\mathcal{R} \models RCOF$. Then $b < 0$, so $0 < -b$, but then $0 < (-b)^2 = b^2 = x$, a contradiction. □

Definition 7.9. A field F is called *formally real* if -1 is not a sum of squares.

It turns out that F is formally real if and only if there exists an order $<$ such that $(F, <)$ is an ordered field. Moreover, if F is formally real and $a \in F$ is not a sum of squares then there exists an ordering on F such that $a < 0$.

Theorem. (Hilbert's 17th Problem, Artin) *Let $f(X_1, \ldots, X_n) \in \mathbb{R}(X_1, \ldots, X_n)$ where $X_1, \ldots, X_n$ are indeterminates. If $f(a_1, \ldots, a_n) \geq 0$ for all $a_1, \ldots, a_n \in \mathbb{R}$ (i.e. f is positive semi-definite) then f is a sum of squares in $\mathbb{R}(X_1, \ldots, X_n)$.*

Proof. Assignment 4. □

Lemma 7.10. *Suppose that $(F, <)$ is an ordered field. The following are equivalent*

(i) $(F, <) \models RCOF$

(ii) For any $a \in F$ either a or $-a$ is a square, and every odd degree polynomial has a root.

Proof. Suppose that $(F, <) \models RCOF$. If $a < 0$ then $0 < -a$, so by *RCOF* axioms $-a$ is a square. The axioms of *RCOF* say that every odd degree polynomial has a root. Suppose that for any $a \in F$ either a or $-a$ is a square, and every odd degree polynomial has a root. In any ordered field negative elements are not squares, so every positive element is a square. □

Definition 7.11. *RCF* is the L_{ring}-theory consisting of

(i) Axioms of a field

(ii) -1 is not a sum of squares, $\neg\exists\, x_1 \dots x_m(-1 = x_1^2 + \cdots + x_m^2)$

(iii) For all odd n, $\forall\, x_1 \dots x_m \exists\, y(y^n + x_1 y^{n-1} + \cdots + x_{n-1}y + x_n = 0)$

(iv) $\forall\, x(\exists\, y(x = y^2) \vee \exists\, (-x = y^2))$

This is the theory of real closed fields.

Examples are $(\mathbb{R}, 0, 1, +, -, \times)$ and $(\overline{\mathbb{Q}}, 0, 1, +, -, \times)$ (real algebraic numbers).

Proposition 7.12. *(i) If $\mathcal{R} \models RCOF$ then $\mathcal{R}|_{L_{ring}} \models RCF$*

(ii) Conversely, if $\mathcal{R} \models RCF$ then the L_{o-ring}-structure obtained by interpreting $x < y$ by $\exists\, y(y - x = z^2)$ is a model of RCOF.

Proof. Exercise. □

Corollary 7.13. *RCF is model-complete and complete.*

Proof. Suppose that $\mathcal{A} \subseteq \mathcal{B}$ where $\mathcal{A} \models RCF$ and $\mathcal{B} \models RCF$. We want to show that $\mathcal{A} \preceq \mathcal{B}$. By Proposition 7.12, $(\mathcal{A}, <) \models RCOF$ and $(\mathcal{B}, <) \models RCOF$. Note that $(\mathcal{A}, <) \subseteq (\mathcal{B}, <)$ as L_{o-ring}-structures. We just need to check that $<^{\mathcal{B}}$ restricts to $<^{\mathcal{A}}$. But in both $\mathcal{A}$ and $\mathcal{B}$ the positives are the non-zero squares, and $\mathcal{A}$ and $\mathcal{B}$ agree on what the non-zero squares are as $\mathcal{A} \subseteq \mathcal{B}$. By QE for $RCOF$, $RCOF$ is model-complete. So $(\mathcal{A}, <) \preceq (\mathcal{B}, <)$, and hence $\mathcal{A} \preceq \mathcal{B}$. So RCF is model-complete.

For completeness, if $\mathcal{A} \models RCF$ and $\mathcal{B} \models RCF$ and σ is any L_{ring}-sentence

$$\mathcal{A} \models \iff (\mathcal{A}, <) \models \sigma \iff (\mathcal{B}, <) \models \sigma \iff \mathcal{B} \models \sigma$$

□

It follows that RCF axiomatises $\mathrm{Th}(\mathbb{R}, 0, 1, +, -, \times)$.

8 Computability, Undecidability, Incompleteness

Definition 8.1. For any L-theory Σ, the *theory of* Σ, $\mathrm{Th}(\Sigma) = \{\sigma\ L\text{-sentence} \mid \Sigma \vdash \sigma\}$. If $\mathcal{A}$ is an L-structure, $\mathrm{Th}(\mathcal{A}) = \{\sigma\ L\text{-sentence} \mid \mathcal{A} \models \sigma\}$

Remark. The following are equivalent

(i) Σ is complete

(ii) For any, equivalently for some, $\mathcal{A} \models \Sigma$ and $\mathrm{Th}(\Sigma) = \mathrm{Th}(\mathcal{A})$

(iii) For any $\mathcal{A}, \mathcal{B} \models \Sigma$, $\mathrm{Th}(\mathcal{A}) = \mathrm{Th}(\mathcal{B})$

The language we will be working with from now on is $L = \{<, 0, S, +, \cdot\}$, the language of arithmetic. Let the L-structure $\mathfrak{N} = (\mathbb{N}, <, 0, S, +, \cdot)$ be the natural numbers with the usual interpretation. $S : \mathbb{N} \to \mathbb{N} : x \mapsto x + 1$ is the successor function.

Informally, a set is calculable if there is a method for determining whether an element is in the set or not in finitely many steps. Similarily, a function $f : A \to B$ is calculable if there is a method such that given $a \in A$ we can obtain $f(a)$ in finitely many steps. This is not a mathematical definition. Gödel's Incompleteness theorem says (essentially) that there is no calculable set of L-sentences Σ such that $\mathrm{Th}(\Sigma) = \mathrm{Th}(\mathfrak{N})$.

It suffices and is convienent to formalise the notion of calculability of functions $f : \mathbb{N}^n \to \mathbb{N}$. If we have such a notion then $R \subseteq \mathbb{N}^n$ is calculable if its characteristic function $\chi_R : \mathbb{N}^n \to \mathbb{N}$ is calculable.

Example 8.2. (i) $+ : \mathbb{N}^2 \to \mathbb{N}$ and $\cdot : \mathbb{N}^2 \to \mathbb{N}$ are calculable. We learned how to calculate these in grade school

(ii) $\chi_{\leq} : \mathbb{N}^2 \to \mathbb{N}$ is calculable.

(iii) For any $1 \leq i \leq n$, let $I_i^N : \mathbb{N}^n \to \mathbb{N} : a \to a_i$, the i^{th} coordinate function. Then I_I^N is calculable.

(iv) Suppose that $H_1, \ldots, H_k : \mathbb{N}^n \to \mathbb{N}$ and $G : \mathbb{N}^k \to \mathbb{N}$ are calculable. Then $F = G(H_1, \ldots H_k) : \mathbb{N}^n \to \mathbb{N}$ is calculable.

(v) Suppose that we all agree that $G : \mathbb{N}^{n+1} \to \mathbb{N}$ is calculable, and we kow that for each $a \in \mathbb{N}^n$ there is $x \in \mathbb{N}$ such that $G(a, x) = 0$. Then consider the function $F : \mathbb{N}^n \to \mathbb{N}$ such that $F(a)$ is the least x such that $G(a, x) = 0$, denoted $\mu x(G(a, x) = 0)$. Then F is calculable.

8.1 Computablility

Definition 8.3. The computable (formal notion of calculable) functions (or recursive functions) are the functions $\mathbb{N}^n \to \mathbb{N}$ ($n = 0, 1, 2, \ldots$) obtained inductively by applying the following rules

(R1) $+ : \mathbb{N}^2 \to \mathbb{N}, \cdot : \mathbb{N}^2 \to \mathbb{N}, \chi_{\leq} : \mathbb{N}^2 \to \mathbb{N}$, $I_i^N : \mathbb{N}^n \to \mathbb{N}$ for all n, $1 \leq i \leq n$ are all computable

(R2) If $H_1, \ldots, H_k : \mathbb{N}^n \to \mathbb{N}$ and $G : \mathbb{N}^k \to \mathbb{N}$ are computable then $F = G(H_1, \ldots H_k) : \mathbb{N}^n \to \mathbb{N}$ is computable.

(R3) If $G : \mathbb{N}^{n+1} \to \mathbb{N}$ is calculable and we know that for each $a \in \mathbb{N}^n$ there is $x \in \mathbb{N}$ such that $G(a, x) = 0$, then $F : \mathbb{N}^n \to \mathbb{N}$ given by $F(a) := \mu x(G(a, x) = 0)$ is computable.

A relation $R \subseteq \mathbb{N}^n$ is computable if its characteristic function $\chi_R : \mathbb{N}^n \to \mathbb{N}$ is computable.

Clearly every computable function is calculable. The Church-Turing Thesis says that every calculable function $F : \mathbb{N} \to \mathbb{N}$ is computable. (It follows for $F : \mathbb{N}^n \to \mathbb{N}$, as we will see later.) This is not a precise mathematical statement. There is evidence for this claim

(i) We can, and will, show that many, many calculable functions are computable.

(ii) It has never failed. No one has produced a function which we agree is calculable that has proven to not be computable. In fact, any candidate F has been proven to be computable.

(iii) Many alternative formalizations of calculable have been proposed (e.g. "calculable by a formal machine"), but they all produce the same set of functions; namely the computable ones.

Lemma 8.4. *The functions $\chi_{\geq} : \mathbb{N}^2 \to \mathbb{N}$, $c_k^n : \mathbb{N}^n \to \mathbb{N} : a \mapsto k$, and $\chi_{=} : \mathbb{N}^2 \to \mathbb{N}$ are computable.*

Proof. $\chi_{\geq}(m, n) = \chi_{\leq}(n, m) = \chi_{\leq}(I_2^2(m, n), I_1^2(m, n))$, so $\chi_{\geq} = \chi_{\leq}(I_2^2, I_1^2)$ by R1, R2, $\chi_{\geq}$ is computable.

For $c_0^n : \mathbb{N}^n \to \mathbb{N} : a \mapsto 0$, $c_0^n(a) = \mu x(I_{n+1}^{n+1}(a, x) = 0)$. $I_{n+1}^{n+1} : \mathbb{N}^{n+1} \to \mathbb{R}$ is computable by R1. For all $a \in \mathbb{N}^n$ there is an x such that $I_{n+1}^{n+1}(a, x) = 0$ (namely $x = 0$). By R3 c_0^n is computable. Now note that $c_{k+1}^n = \mu x(c_k^n(a) < x) = \mu x(\chi_{\geq}(c_k^{n+1}(a, x), I_{n+1}^{n+1}(a, x)) = 0)$. Letting $G : \mathbb{N}^{n+1} \to \mathbb{N}$, $G = \chi_{\geq}(c_k^{n+1}, I_{n+1}^{n+1})$ we see by induction, R1, R2, R3, and the first part of the lemma that the constant functions are computable.

$\chi_{=} = \chi_{\leq} \cdot \chi_{\geq}$, so by R1 $\chi_{=}$ is computable. □

Lemma 8.5. *(i) If $P, Q \subseteq \mathbb{N}^n$ are computable relations then so are*

(a) $\neg P = \mathbb{N}^n \setminus P$

(b) $P \vee Q = P \cup Q$

(c) $P \wedge Q = P \cap Q$

(d) $P \to Q = (\mathbb{N}^n \setminus P) \cup Q$

(e) $P \leftrightarrow Q = [(\mathbb{N}^n \setminus P) \cup Q] \cap [(\mathbb{N}^n \setminus Q) \cup P]$

(ii) $<, \leq, =, \geq, >, \neq$ *are all computable.*

Proof. Exercise. It suffices to show for $\neg P$ and $P \wedge Q$. □

Lemma 8.6. *Suppose that $H_1, \ldots, H_k : \mathbb{N}^n \to \mathbb{N}$ are computable. Suppose that $R \subseteq \mathbb{N}^k$ is computable. then $R(H_1, \ldots, H_n) \subseteq \mathbb{N}^n$ is computable where $a \in R(H_1, \ldots, H_k) \iff (H_1(a), \ldots, H_k(a)) \in R$.*

Proof. $\chi_{R(H_1,\ldots,H_n)} = \chi_R(H_1, \ldots, H_k)$. □

Lemma 8.7. (i) (Definition of functions by cases) *Let $R_1, \ldots, R_k \subseteq \mathbb{N}^n$ be computable and suppose that for every $a \in \mathbb{N}^n$ exactly one of $R_1(a), \ldots, R_k(a)$ is true. If $G_1, \ldots, G_k : \mathbb{N}^n \to \mathbb{N}$ are computable then $F : \mathbb{N}^n \to \mathbb{N}$ defined by*

$$F(a) = \begin{cases} G_1(a) & \text{if } R_1(a) \\ \vdots & \\ G_k(a) & \text{if } R_k(a) \end{cases}$$

is computable.

(ii) (Definition of relations by cases) *$R_1, \ldots, R_k$ as before. Suppose that $P_1, \ldots, P_k \subseteq \mathbb{N}^n$ are computable. Then $P \subseteq \mathbb{N}^n$ defined by*

$$P(a) \iff \begin{cases} P_1(a) & \text{if } R_1(a) \\ \vdots & \\ P_k(a) & \text{if } R_k(a) \end{cases}$$

is computable

Proof. (i) $F = \chi_{R_1} \cdot G_1 + \cdots + \chi_{R_k} \cdot G_k$. Check this.

(ii) $P = P_1 \wedge R_1 \vee \cdots \vee R_k \wedge P_k$. Check this too.

□

Lemma 8.8. *Let $R \subseteq \mathbb{N}^{n+1}$ be computable. Suppose that for all $a \in \mathbb{N}^n$ there is $x \in \mathbb{N}$ such that $R(a, x)$. Then $F : \mathbb{N}^n \to \mathbb{N} : a \mapsto \mu x(R(a, x))$ is computable.*

Proof. $F(a) = \mu x(\chi_R(a, x) \neq 0) = \mu x(\chi_=(\chi_r(a, x), c_0^{n+1}(a, x)) = 0)$, which is computable by R3. □

Lemma 8.9. *$F : \mathbb{N}^n \to \mathbb{N}$ is computable if and only if its graph $\Gamma(F) \subseteq \mathbb{N}^{n+1}$ is computable.*

Proof. $\chi_{\Gamma(F)} = \chi_=(F(I_1^{n+1}, I_n^{n+1}), I_{n+1}^{n+1})$.
$F(a) = \mu x(\Gamma(F)(a, x))$. □

Lemma 8.10. *Let $R \subseteq \mathbb{N}^{n+1}$ be computable. Suppose $P, Q \subseteq \mathbb{N}^{n+1}$ are defined by: for all $a \in \mathbb{N}^n$, $y \in \mathbb{N}$*

(i) *$P(a, y)$ if and only if there is an $x < y$ such that $R(a, x)$*

(ii) *$Q(a, y)$ if and only if for all $x < y$, $R(a, x)$*

Then P and Q are computable.

Proof. (i) $P(a,y) \iff \mu x(R(a,x) \vee x = I^{n+1}_{n+1}(a,y)) < I^{n+1}_{n+1}(a,y)$

(ii) $P(a,y) \iff \neg$(there exists $x < y$ such that $\neg R(a,x)$), so by (1) we are done.

□

We use $\exists\, x < y$ as short for "there exists $x < y$". This is not to be confused with $\exists$ from a language.

Lemma 8.11. *The function* $\dot{-} : \mathbb{N}^2 \to \mathbb{N}$ *is computable, where*

$$\dot{-}(x,y) = \begin{cases} x - y & \text{if } x \geq y \\ 0 & \text{if } x < y \end{cases}$$

Lemma 8.12. *Pair* $: \mathbb{N}^2 \to \mathbb{N}$ *is defined by* $Pair(x,y) = \frac{(x+y)^2+(x+y)}{2} + x$. *Pair is computable and bijective.*

Proof. Exercise. Do it. □

Corollary 8.13. *Left, Right* $: \mathbb{N} \to \mathbb{N}$ *defined by* $Left(a) = I^2_1(Pair^{-1}(a))$ *and* $Right(a) = I^2_2(Pair^{-1}(a))$ *are both computable.*

Definition 8.14. *(Gödel's β function)* $\beta : \mathbb{N}^2 \to \mathbb{N}$ defined by

$$\beta(a,i) = \mu x(x \equiv \text{Left}(a) \pmod{1 + (i+1)\text{Right}(a)})$$

Note that the teriary relation $a \equiv b \pmod{c}$ is computable because

$$a \equiv b \pmod{c} \iff \exists\, x < a+1\ a = x \cdot c + b \vee \exists\, x < b+1\ b = x \cdot c + a$$

Proposition 8.15. *(i) β is computable*

(ii) $\beta(a,i) \leq a \dot{-} 1$

(iii) For any sequence $(a_0, \ldots, a_{n-1}) \in \mathbb{N}^n$ *there is an* $a \in \mathbb{N}$ *such that* $\beta(a,i) = a_i$

Proof. (i) See note above.

(ii) $\beta(a,i) \leq \text{Left}(a) \leq a$. If $a > 0$ then $\text{Left}(a) < a$, so $\beta(a,i) \leq a \dot{-} 1$.

(iii) Suppose that $(a_0, \ldots, a_{n-1}) \in \mathbb{N}^n$. Let $N \in \mathbb{N}$ be such that $N > a_i$ for $i = 0, 1, \ldots, n-1$ and N is a multiple of all the primes less than n. Then $1 + N, 1 + 2N, \ldots, 1 + nN$ are relatively prime. By the Chinese Remainder Theorem there is an M such that

$$\begin{aligned} M &\equiv a_0 \pmod{1+N} \\ M &\equiv a_1 \pmod{1+2N} \\ &\vdots \\ M &\equiv a_{n-1} \pmod{1+nN} \end{aligned}$$

Let $a = \text{Pair}(M,N)$. So $\text{Left}(a) = M$ and when it is divided by $1 + (1+i)\text{Right}(a) = 1 + (1+i)N$ we get a_i. Thus a satisfies the conclusion of the proposition.

□

Definition 8.16. Given $(a_1, \ldots, a_n) \in \mathbb{N}^n$ we define

$$\langle a_1, \ldots, a_n \rangle = \mu x(\beta(x,0) = n, \beta(x,1) = a_1, \ldots, \beta(x,n) = a_n)$$

So $\langle a_1, \ldots, a_n \rangle$ is the least natural number which encodes $(n, a_1, \ldots, a_n)$ in β. Note that $\langle\rangle = 0$.

Lemma 8.17. *(i) Define* $Seq := \{a \in \mathbb{N} \mid a = \langle a_1, \ldots, a_n \rangle \text{ for some } a_1, \ldots, a_n \in \mathbb{N}\} \subseteq \mathbb{N}$ *is computable.*

(ii) $(a_1, \ldots, a_n) \mapsto \langle a_1, \ldots, a_n \rangle : \mathbb{N}^n \to \mathbb{N}$ *is computable*

(iii) $\ell : \mathbb{N} \to \mathbb{N}$ *given by* $\ell(a) = \beta(a, 0)$ *is computable. This is known as the length function.*

(iv) $(a, i) \mapsto \beta(a, i+1) : \mathbb{N}^2 \to \mathbb{N}$ *is computable. This is the* i^{th} *coordinate function, and we denote it* $(a)_i = \beta(a, i+1)$.

(v) $In : \mathbb{N}^2 \to \mathbb{N} : (a, i) \mapsto \mu x(\ell(x) = i \wedge \forall\, j < i\ (x)_j = (a)_j)$ *is computable. This gets the initial segment of a seqence* a.

(vi) $* : \mathbb{N}^2 \to \mathbb{N} : (a, b) \mapsto \mu x(\ell(x) = \ell(a) + \ell(b) \wedge \forall\, i < \ell(a)\,(x)_i = (a)_i \text{ and } \forall\, j < \ell(b)\,(x)_{\ell(a)+j} = (a)_j)$ *is computable. This is sequence concatenation.*

Notice that $\beta(a, i) \leq a \dot{-} 1$ for all a and i.

Proof. See text. Most of these functions have already been given explicitly and are hence seen to be computable. □

Lemma 8.18. *Given* $F : \mathbb{N}^{n+1} \to \mathbb{N}$, *define* $\overline{F} : \mathbb{N}^{n+1} \to \mathbb{N}$ *for* $a \in \mathbb{N}^n$ *and* $b \in \mathbb{N}$ *by* $\overline{F}(a, b) = \langle F(a,0), F(a,1), \ldots, F(a, b-1) \rangle$. *Then* F *is computable if and only if* $\overline{F}$ *is computable.*

Proposition 8.19. *Suppose* $G : \mathbb{N}^{n+2} \to \mathbb{N}$ *is computable. Let* $F : \mathbb{N}^{n+1} \to \mathbb{N}$ *be such that for* $a \in \mathbb{N}^n$, $F(a, 0) = G(a, 0, 0)$ *and* $F(a, b+1) = G(a, b+1, \langle F(a,0), \ldots, F(a,b) \rangle)$. *That is,* $F(a, b) = G(a, b, \overline{F}(a, b))$ *is computable.*

If F is defined recursively using computable data then it is computable. For example, the function $m \mapsto 2^m$ is computable.

Let $A : \mathbb{N}^{n+1} \to \mathbb{N}$ and $B : \mathbb{N}^{n+2} \to \mathbb{N}$. Suppose $F : \mathbb{N}^{n+1} \to \mathbb{N}$ is defined by forall $a \in \mathbb{N}^n$, $F(a, 0) = A(a)$ and $F(a, b+1) = B(a, b, F(a, b))$. We say that F is defined by primitive recursion from A and B.

Corollary 8.20. *If* A *and* B *are computable then so is* F.

So also sets $R \subseteq \mathbb{N}^n$ defined recursively from computable data are computable.

8.2 Gödel Numbering

We want to talk about computable L-theories, formulae, *et cetera*. For the purposes of this we will assume that L is a finite language, but this assumption is not always necessary. Assign to each symbol s from the language, logical symbols, or variables a natural number, called the symbol number as follows:

(i) $s = v_i \in \text{Var} = \{v_1, v_2, \ldots\}$ set $SN(s) = 2i$

(ii) Otherwise assign $SN(s)$ to be an odd number. Notice that there are only finitely many symbols in the language and logical symbols.

Definition 8.21. The Gödel number $\ulcorner t \urcorner$ of an L-term t is defined recursively as follows

$$\ulcorner t \urcorner := \begin{cases} \langle SN(t) \rangle & \text{if } t = v_i \\ \langle SN(F), \ulcorner t_1 \urcorner, \ldots, \ulcorner t_n \urcorner \rangle & \text{if } t = F\, t_1 \ldots t_n \end{cases}$$

For a formula φ,

$$\ulcorner \varphi \urcorner := \begin{cases} \langle SN(\top) \rangle & \text{if } \varphi = \top \\ \langle SN(\bot) \rangle & \text{if } \varphi = \bot \\ \langle SN(=), \ulcorner t_1 \urcorner, \ulcorner t_2 \urcorner \rangle & \text{if } \varphi \text{ is } t_1 = t_2 \\ \langle SN(R), \ulcorner t_1 \urcorner, \ldots, \ulcorner t_n \urcorner \rangle & \text{if } \varphi = R\, t_1 \ldots t_n \\ \langle SN(\neg), \ulcorner \psi \urcorner \rangle & \text{if } \varphi = \neg\psi \\ \langle SN(\wedge), \ulcorner \psi \urcorner, \ulcorner \theta \urcorner \rangle & \text{if } \varphi = \psi \wedge \theta \\ \langle SN(\vee), \ulcorner \psi \urcorner, \ulcorner \theta \urcorner \rangle & \text{if } \varphi = \psi \vee \theta \\ \langle SN(\exists), SN(x), \ulcorner \psi \urcorner \rangle & \text{if } \varphi = \exists x \psi \\ \langle SN(\forall), SN(x), \ulcorner \psi \urcorner \rangle & \text{if } \varphi = \forall x \psi \end{cases}$$

Lemma 8.22. *The following subsets of* $\mathbb{N}$ *are computable:*

(i) $Vble := \{\ulcorner x \urcorner \mid x \text{ is a variable}\} \subseteq \mathbb{N}$

(ii) $Term := \{\ulcorner t \urcorner \mid t \text{ is an } L\text{-term}\} \subseteq \mathbb{N}$

(iii) $AFor := \{\ulcorner \varphi \urcorner \mid \varphi \text{ is an atomic } L\text{-formula}\} \subseteq \mathbb{N}$

(iv) $For := \{\ulcorner \varphi \urcorner \mid \varphi \text{ is an } L\text{-formula}\} \subseteq \mathbb{N}$

Proof. (i) $\mathrm{Vble}(a)$ if and only if $a \in \mathrm{Seq}$ and $\ell(a) = 1$ and $\exists y < a\ (a)_0 = 2y$. (Note that $y < (a)_0 = \beta(a,1) \leq a \dot{-} 1 \leq a$.)
The rest are exercises. Do them. □

Lemma 8.23. *There exists a computable function* $Sub : \mathbb{N}^3 \to \mathbb{N}$ *which satisfies for any terms* t *and* s*, variable* x*, and formula* φ*.*

$$Sub(\ulcorner t \urcorner, \ulcorner x \urcorner, \ulcorner s \urcorner) = \ulcorner t(s/x) \urcorner$$
$$Sub(\ulcorner \varphi \urcorner, \ulcorner x \urcorner, \ulcorner s \urcorner) = \ulcorner \varphi(s/x) \urcorner$$

Proof. See text. □

Lemma 8.24. *The following are computable.*

(i) $Fr := \{(\ulcorner \varphi \urcorner, \ulcorner x \urcorner) \mid x \text{ is free in } \varphi\} \subseteq \mathbb{N}^2$

(ii) $FrSub := \{(\ulcorner \varphi \urcorner, \ulcorner x \urcorner, \ulcorner t \urcorner) \mid t \text{ is free for } x \text{ in } \varphi\} \subseteq \mathbb{N}^3$

(iii) $PrAx := \{\ulcorner \varphi \urcorner \mid \varphi \text{ is a propositional axiom}\} \subseteq \mathbb{N}$

(iv) $EqAx := \{\ulcorner \varphi \urcorner \mid \varphi \text{ is an equality axiom}\} \subseteq \mathbb{N}$

(v) $Quant := \{\ulcorner \varphi \urcorner \mid \varphi \text{ is a quantifier axiom}\} \subseteq \mathbb{N}$

(vi) $MP := \{(\ulcorner \varphi \urcorner, \ulcorner \varphi \to \psi \urcorner, \ulcorner \psi \urcorner) \mid \varphi, \psi \text{ are } L\text{-formulae}\} \subseteq \mathbb{N}^3$

(vii) $Gen := \{(\ulcorner \varphi \urcorner, \ulcorner \psi \urcorner) \mid \psi \text{ follows from } \varphi \text{ by generalization}\} \subseteq \mathbb{N}^2$

(viii) Sent $:= \{\ulcorner\sigma\urcorner \mid \sigma$ *is an L-sentence*$\} \subseteq \mathbb{N}$

We can encode the syntax of L in $\mathbb{N}$ in an effective manner. Suppose that Σ is an L-theory. We say that Σ is computable if $\ulcorner\Sigma\urcorner := \{\ulcorner\sigma\urcorner \mid \sigma \in \Sigma\} \subseteq \mathbb{N}$ is computable. Define Prf_Σ to be the set of all proofs from Σ. More precisely, it is the set $\{\langle\ulcorner\varphi_1\urcorner, \ldots, \ulcorner\varphi_n\urcorner\rangle \mid (\varphi_1, \ldots, \varphi_n)$ is a proof of φ_n from $\Sigma\}$.

Lemma 8.25. *If Σ is computable then Prf_Σ is computable.*

If Σ is computable, is $\mathrm{Th}(\Sigma)$ computable?

$$a \in \mathrm{Th}(\sigma) \iff a \text{ is a sentence and there is } b \in \mathrm{Prf}_\Sigma \text{ such that } (b)_{\ell(b)} = a$$

This is not necessarily computable because there is no bound on the length of the proof of a.

Definition 8.26. A relation $S \subseteq \mathbb{N}$ is computably generated if there exists a computable relation $R \subseteq \mathbb{N}^2$ such that

$$S(a) \iff \exists\, b\ R(a,b)$$

8.3 Löb's Theorem

$\Sigma \supseteq N$, $\mathfrak{N} \models \Sigma$, $L = \{0, \cdot, +, S, <\}$

Found σ L-sentence such that $\Sigma \vdash \sigma \leftrightarrow \forall x \neg \mathrm{Pr}_\Sigma(x, S^{\ulcorner\sigma\urcorner}0)$. Such a σ exists by the "fixed-point" theorem. In $\mathfrak{N}$, $S^{\ulcorner\sigma\urcorner}0 = \ulcorner\sigma\urcorner$. In $\mathfrak{N}$, σ says "I am unprovable from Σ". $\mathfrak{N} \models \sigma$ and $\Sigma \nvdash \sigma$ (check this), so σ witnesses Gödel Incompleteness theorem. σ is called a Gödel sentence. By the fixed-point theorem, there is also τ L-sentence such that $\Sigma \vdash \tau \leftrightarrow \exists x \mathrm{Pr}_\Sigma(x, S^{\ulcorner\tau\urcorner}0)$. In $\mathfrak{N}$, τ says "I am provable from Σ". Heuristically, there is no problem with $\mathfrak{N} \models \tau$, $\mathfrak{N} \models \neg\tau$, $\Sigma \vdash \tau$, $\Sigma \nvdash \tau$. Löb proved that $\Sigma \vdash \tau$ (and hence $\Sigma \models \tau$).

Theorem 8.27. (Löb) *Let θ be any sentence such that $\Sigma \vdash \exists x Pr_\Sigma(x, S^{\ulcorner\tau\urcorner}0) \to \theta$ (this is called* soundness*, provable implies true) then $\Sigma \vdash \theta$.*

8.4 More definability

Definition 8.28. L,L' two languages, $\mathcal{A}$ and L-structure and $\mathcal{B}$ an L'-structure. We say that $\mathcal{A}$ is definable in $\mathcal{B}$ if there exists an injective map $\delta : |\mathcal{A}| \to |\mathcal{B}|^k$ for some $k \geq 0$, such that

- $\delta(|\mathcal{A}|) \subseteq |\mathcal{B}|^k$ is definable in $\mathcal{B}$ say defined by the $L_{|\mathcal{B}|}$-formula δU
- For every $R \in L^{rel}$, n-ary, $\delta(R^{\mathcal{A}}) \subseteq |\mathcal{B}|^{nk}$ is definable in $\mathcal{B}$ say defined by the $L_{|\mathcal{B}|}$-formula δR
- For every $F \in L^{fun}$, n-ary, $\delta(\Gamma(F^{\mathcal{A}})) \subseteq |\mathcal{B}|^{(n+1)k}$ is definable in $\mathcal{B}$ say defined by the $L_{|\mathcal{B}|}$-formula δF

Example 8.29. $\mathfrak{N} = (\mathbb{N}, 0, +, \cdot, S, <)$ is definable in $(\mathbb{Z}, +, 0, 1, \times, -)$. Take $\delta = id$, $k = 1$. By Lagrange's Theorem $x \in \mathbb{N} \iff \exists abcd(x = a^2 + b^2 + c^2 + d^2)$. $+$ is same, $\cdot$ is $\times$, $S(x) = x + 1$, $x < y \iff (y - x) \in \mathbb{N}$.

If δU and all the δR's and δF's are L'-formulae then we say that δ is a 0-definition of $\mathcal{A}$ in $\mathcal{B}$ (or $\mathcal{A}$ is 0-definable in $\mathcal{B}$).

Theorem 8.30. (Tarski) *If $\mathcal{A}$ is strongly undecidable and $\mathcal{A}$ is definable in $\mathcal{B}$, then $\mathcal{B}$ is strongly undecidable.*

Corollary 8.31. *$L = L_{ring}$ and RI = theory of rings is undecidable by the example above (since $\mathfrak{N}$ is strongly undecidable and definable in $\mathbb{Z}$, a ring)*

Theorem 8.32. (Julia Robinson) *$\mathcal{Z}$ is definable in $\mathcal{Q}$ (as a field).*

Hence $\mathcal{Q}$ is strongly undecidable, so the the theory of fields is undecidable.