Model Theory for Algebra and Algebraic Geometry

David Marker

Spring 2010–Orsay

1 Language, Structures and Theories

In mathematical logic, we use first-order languages to describe mathematical structures. Intuitively, a structure is a set that we wish to study equipped with a collection of distinguished functions, relations, and elements. We then choose a language where we can talk about the distinguished functions, relations, and elements and nothing more. For example, when we study the ordered field of real numbers with the exponential function, we study the structure $(\mathbb{R}, +, \cdot, \exp, <, 0, 1)$, where the underlying set is the set of real numbers, and we distinguish the binary functions addition and multiplication, the unary function $x \mapsto e^x$, the binary order relation, and the real numbers 0 and 1. To describe this structure, we would use a language where we have symbols for $+, \cdot, \exp, <$ $, 0, 1$ and can write statements such as $\forall x \forall y \ \exp(x) \cdot \exp(y) = \exp(x + y)$ and $\forall x \ (x > 0 \rightarrow \exists y \ \exp(y) = x)$. We interpret these statements as the assertions "$e^x e^y = e^{x+y}$ for all x and y" and "for all positive x, there is a y such that $e^y = x$."

For another example, we might consider the structure $(\mathbb{N}, +, 0, 1)$ of the natural numbers with addition and distinguished elements 0 and 1. The natural language for studying this structure is the language where we have a binary function symbol for addition and constant symbols for 0 and 1. We would write sentences such as $\forall x \exists y \ (x = y + y \ \vee \ x = y + y + 1)$, which we interpret as the assertion that "every number is either even or 1 plus an even number."

Definition 1.1 A *language* $\mathcal{L}$ is given by specifying the following data:
i) a set of function symbols $\mathcal{F}$ and positive integers n_f for each $f \in \mathcal{F}$;
ii) a set of relation symbols $\mathcal{R}$ and positive integers n_R for each $R \in \mathcal{R}$;
iii) a set of constant symbols $\mathcal{C}$.

The numbers n_f and n_R tell us that f is a function of n_f variables and R is an n_R-ary relation.

Any or all of the sets $\mathcal{F}$, $\mathcal{R}$, and $\mathcal{C}$ may be empty. Examples of languages include:

i) the language of rings $\mathcal{L}_\mathrm{r} = \{+, -, \cdot, 0, 1\}$, where $+, -$ and $\cdot$ are binary function symbols and 0 and 1 are constants;

ii) the language of ordered rings $\mathcal{L}_{\rm or} = \mathcal{L}_{\rm r} \cup \{<\}$, where $<$ is a binary relation symbol;

iii) the language of pure sets $\mathcal{L} = \emptyset$;

iv) the language of graphs is $\mathcal{L} = \{R\}$ where R is a binary relation symbol.

Next, we describe the structures where $\mathcal{L}$ is the appropriate language.

Definition 1.2 An *$\mathcal{L}$-structure* $\mathcal{M}$ is given by the following data:

i) a nonempty set M called the *universe*, *domain*, or *underlying set* of $\mathcal{M}$;

ii) a function $f^{\mathcal{M}} : M^{n_f} \to M$ for each $f \in \mathcal{F}$;

iii) a set $R^{\mathcal{M}} \subseteq M^{n_R}$ for each $R \in \mathcal{R}$;

iv) an element $c^{\mathcal{M}} \in M$ for each $c \in \mathcal{C}$.

We refer to $f^{\mathcal{M}}$, $R^{\mathcal{M}}$, and $c^{\mathcal{M}}$ as the *interpretations* of the symbols f, R, and c. We often write the structure as $\mathcal{M} = (M, f^{\mathcal{M}}, R^{\mathcal{M}}, c^{\mathcal{M}} : f \in \mathcal{F}, R \in \mathcal{R}$, and $c \in \mathcal{C})$. We will use the notation $A, B, M, N, \ldots$ to refer to the underlying sets of the structures $\mathcal{A}, \mathcal{B}, \mathcal{M}, \mathcal{N}, \ldots$.

For example, suppose that we are studying groups. We might use the language $\mathcal{L}_{\rm g} = \{\cdot, e\}$, where $\cdot$ is a binary function symbol and e is a constant symbol. An $\mathcal{L}_{\rm g}$-structure $\mathcal{G} = (G, \cdot^{\mathcal{G}}, e^{\mathcal{G}})$ will be a set G equipped with a binary relation $\cdot^{\mathcal{G}}$ and a distinguished element $e^{\mathcal{G}}$. For example, $\mathcal{G} = (\mathbb{R}, \cdot, 1)$ is an $\mathcal{L}_{\rm g}$-structure where we interpret $\cdot$ as multiplication and e as 1; that is, $\cdot^{\mathcal{G}} = \cdot$ and $e^{\mathcal{G}} = 1$. Also, $\mathcal{N} = (\mathbb{N}, +, 0)$ is an $\mathcal{L}_{\rm g}$-structure where $\cdot^{\mathcal{N}} = +$ and $e^{\mathcal{G}} = 0$. Of course, $\mathcal{N}$ is not a group, but it is an $\mathcal{L}_{\rm g}$-structure.

Usually, we will choose languages that closely correspond to the structure that we wish to study. For example, if we want to study the real numbers as an ordered field, we would use the language of ordered rings $\mathcal{L}_{\rm or}$ and give each symbol its natural interpretation.

We will study maps that preserve the interpretation of $\mathcal{L}$.

Definition 1.3 Suppose that $\mathcal{M}$ and $\mathcal{N}$ are $\mathcal{L}$-structures with universes M and N, respectively. An *$\mathcal{L}$-embedding* $\eta : \mathcal{M} \to \mathcal{N}$ is a one-to-one map $\eta : M \to N$ that preserves the interpretation of all of the symbols of $\mathcal{L}$. More precisely:

i) $\eta(f^{\mathcal{M}}(a_1, \ldots, a_{n_f})) = f^{\mathcal{N}}(\eta(a_1), \ldots, \eta(a_{n_f}))$ for all $f \in \mathcal{F}$ and $a_1, \ldots, a_n \in M$;

ii) $(a_1, \ldots, a_{m_R}) \in R^{\mathcal{M}}$ if and only if $(\eta(a_1), \ldots, \eta(a_{m_R})) \in R^{\mathcal{N}}$ for all $R \in \mathcal{R}$ and $a_1, \ldots, a_{m_j} \in M$;

iii) $\eta(c^{\mathcal{M}}) = c^{\mathcal{N}}$ for $c \in \mathcal{C}$.

A bijective $\mathcal{L}$-embedding is called an *$\mathcal{L}$-isomorphism.* If $M \subseteq N$ and the inclusion map is an $\mathcal{L}$-embedding, we say either that $\mathcal{M}$ is a *substructure* of $\mathcal{N}$ or that $\mathcal{N}$ is an *extension* of $\mathcal{M}$.

For example:

i) $(\mathbb{Z}, +, 0)$ is a substructure of $(\mathbb{R}, +, 0)$.

ii) If $\eta : \mathbb{Z} \to \mathbb{R}$ is the function $\eta(x) = e^x$, then η is an $\mathcal{L}_{\rm g}$-embedding of $(\mathbb{Z}, +, 0)$ into $(\mathbb{R}, \cdot, 1)$.

The *cardinality of* $\mathcal{M}$ is $|M|$, the cardinality of the universe of $\mathcal{M}$. If $\eta : \mathcal{M} \to \mathcal{N}$ is an embedding then the cardinality of $\mathcal{N}$ is at least the cardinality of $\mathcal{M}$.

We use the language $\mathcal{L}$ to create formulas describing properties of $\mathcal{L}$-structures. Formulas will be strings of symbols built using the symbols of $\mathcal{L}$, variable symbols $v_1, v_2, \ldots$, the equality symbol $=$, the Boolean connectives $\wedge$, $\vee$, and $\neg$, which we read as "and," "or," and "not", the quantifiers $\exists$ and $\forall$, which we read as "there exists" and "for all", and parentheses (,).

Definition 1.4 The set of *$\mathcal{L}$-terms* is the smallest set $\mathcal{T}$ such that

i) $c \in \mathcal{T}$ for each constant symbol $c \in \mathcal{C}$,

ii) each variable symbol $v_i \in \mathcal{T}$ for $i = 1, 2, \ldots$, and

iii) if $t_1, \ldots, t_{n_f} \in \mathcal{T}$ and $f \in \mathcal{F}$, then $f(t_1, \ldots, t_{n_f}) \in \mathcal{T}$.

For example, $\cdot(v_1, -(v_3, 1))$, $\cdot(+(v_1, v_2), +(v_3, 1))$ and $+(1, +(1, +(1, 1)))$ are $\mathcal{L}_{\mathrm{r}}$-terms. For simplicity, we will usually write these terms in the more standard notation $v_1(v_3 - 1)$, $(v_1 + v_2)(v_3 + 1)$, and $1 + (1 + (1 + 1))$ when no confusion arises. In the $\mathcal{L}_{\mathrm{r}}$-structure $(\mathbb{Z}, +, \cdot, 0, 1)$, we think of the term $1 + (1 + (1 + 1))$ as a name for the element 4, while $(v_1 + v_2)(v_3 + 1)$ is a name for the function $(x, y, z) \mapsto (x + y)(z + 1)$. This can be done in any $\mathcal{L}$-structure.

Suppose that $\mathcal{M}$ is an $\mathcal{L}$-structure and that t is a term built using variables from $\overline{v} = (v_{i_1}, \ldots, v_{i_m})$. We want to interpret t as a function $t^{\mathcal{M}} : M^m \to M$. For s a subterm of t and $\overline{a} = (a_{i_1}, \ldots, a_{i_m}) \in M$, we inductively define $s^{\mathcal{M}}(\overline{a})$ as follows.

i) If s is a constant symbol c, then $s^{\mathcal{M}}(\overline{a}) = c^{\mathcal{M}}$.

ii) If s is the variable v_{i_j}, then $s^{\mathcal{M}}(\overline{a}) = a_{i_j}$.

iii) If s is the term $f(t_1, \ldots, t_{n_f})$, where f is a function symbol of $\mathcal{L}$ and $t_1, \ldots, t_{n_f}$ are terms, then $s^{\mathcal{M}}(\overline{a}) = f^{\mathcal{M}}(t_1^{\mathcal{M}}(\overline{a}), \ldots, t_{n_f}^{\mathcal{M}}(\overline{a}))$.

The function $t^{\mathcal{M}}$ is defined by $\overline{a} \mapsto t^{\mathcal{M}}(\overline{a})$.

For example, let $\mathcal{L} = \{f, g, c\}$, where f is a unary function symbol, g is a binary function symbol, and c is a constant symbol. We will consider the $\mathcal{L}$-terms $t_1 = g(v_1, c)$, $t_2 = f(g(c, f(v_1)))$, and $t_3 = g(f(g(v_1, v_2)), g(v_1, f(v_2)))$. Let $\mathcal{M}$ be the $\mathcal{L}$-structure $(\mathbb{R}, \exp, +, 1)$; that is, $f^{\mathcal{M}} = \exp$, $g^{\mathcal{M}} = +$, and $c^{\mathcal{M}} = 1$.

Then

$$t_1^{\mathcal{M}}(a_1) = a_1 + 1,$$

$$t_2^{\mathcal{M}}(a_1) = e^{1+e^{a_1}}, \text{ and}$$

$$t_3^{\mathcal{M}}(a_1, a_2) = e^{a_1 + a_2} + (a_1 + e^{a_2}).$$

We are now ready to define $\mathcal{L}$-formulas.

Definition 1.5 We say that ϕ is an *atomic $\mathcal{L}$-formula* if ϕ is either

i) $t_1 = t_2$, where t_1 and t_2 are terms, or

ii) $R(t_1, \ldots, t_{n_R})$, where $R \in \mathcal{R}$ and $t_1, \ldots, t_{n_R}$ are terms.

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

i) if ϕ is in $\mathcal{W}$, then $\neg\phi$ is in $\mathcal{W}$,

ii) if ϕ and ψ are in $\mathcal{W}$, then $(\phi \wedge \psi)$ and $(\phi \vee \psi)$ are in $\mathcal{W}$, and

iii) if ϕ is in $\mathcal{W}$, then $\exists v_i\ \phi$ and $\forall v_i\ \phi$ are in $\mathcal{W}$.

Here are three examples of $\mathcal{L}_{\rm or}$-formulas.

- $v_1 = 0 \vee v_1 > 0$.
- $\exists v_2\ v_2 \cdot v_2 = v_1$.
- $\forall v_1\ (v_1 = 0 \vee \exists v_2\ v_2 \cdot v_1 = 1)$.

Intuitively, the first formula asserts that $v_1 \geq 0$, the second asserts that v_1 is a square, and the third asserts that every nonzero element has a multiplicative inverse. We would like to define what it means for a formula to be true in a structure, but these examples already show one difficulty. While in any $\mathcal{L}_{\rm or}$-structure the third formula will either be true or false, the first two formulas express a property that may or may not be true of particular elements of the structure. In the $\mathcal{L}_{\rm or}$-structure $(\mathbb{Z}, +, -, \cdot, <, 0, 1)$, the second formula would be true of 9 but false of 8.

We say that a variable v *occurs freely* in a formula ϕ if it is not inside a $\exists v$ or $\forall v$ quantifier; otherwise, we say that it is *bound.*[1] For example v_1 is free in the first two formulas and bound in the third, whereas v_2 is bound in both formulas. We call a formula a *sentence* if it has no free variables.

Let $\mathcal{M}$ be an $\mathcal{L}$-structure. We will see that each $\mathcal{L}$-sentence is either true or false in $\mathcal{M}$. On the other hand, if ϕ is a formula with free variables $v_1, \ldots, v_n$, we will think of ϕ as expressing a property of elements of M^n. We often write $\phi(v_1, \ldots, v_n)$ to make explicit the free variables in ϕ. We must define what it means for $\phi(v_1, \ldots, v_n)$ to hold of $(a_1, \ldots, a_n) \in M^n$.

Definition 1.6 Let ϕ be a formula with free variables from $\overline{v} = (v_{i_1}, \ldots, v_{i_m})$, and let $\overline{a} = (a_{i_1}, \ldots, a_{i_m}) \in M^m$. We inductively define $\mathcal{M} \models \phi(\overline{a})$ as follows.

i) If ϕ is $t_1 = t_2$, then $\mathcal{M} \models \phi(\overline{a})$ if $t_1^{\mathcal{M}}(\overline{a}) = t_2^{\mathcal{M}}(\overline{a})$.

ii) If ϕ is $R(t_1, \ldots, t_{n_R})$, then $\mathcal{M} \models \phi(\overline{a})$ if $(t_1^{\mathcal{M}}(\overline{a}), \ldots, t_{n_R}^{\mathcal{M}}(\overline{a})) \in R^{\mathcal{M}}$.

iii) If ϕ is $\neg\psi$, then $\mathcal{M} \models \phi(\overline{a})$ if $\mathcal{M} \not\models \psi(\overline{a})$.

iv) If ϕ is $(\psi \wedge \theta)$, then $\mathcal{M} \models \phi(\overline{a})$ if $\mathcal{M} \models \psi(\overline{a})$ and $\mathcal{M} \models \theta(\overline{a})$.

v) If ϕ is $(\psi \vee \theta)$, then $\mathcal{M} \models \phi(\overline{a})$ if $\mathcal{M} \models \psi(\overline{a})$ or $\mathcal{M} \models \theta(\overline{a})$.

vi) If ϕ is $\exists v_j \psi(\overline{v}, v_j)$, then $\mathcal{M} \models \phi(\overline{a})$ if there is $b \in M$ such that $\mathcal{M} \models \psi(\overline{a}, b)$.

vii) If ϕ is $\forall v_j \psi(\overline{v}, v_j)$, then $\mathcal{M} \models \phi(\overline{a})$ if $\mathcal{M} \models \psi(\overline{a}, b)$ for all $b \in M$.

If $\mathcal{M} \models \phi(\overline{a})$ we say that $\mathcal{M}$ *satisfies* $\phi(\overline{a})$ or $\phi(\overline{a})$ is *true* in $\mathcal{M}$.

Remarks 1.7 • There are a number of useful abbreviations that we will use: $\phi \rightarrow \psi$ is an abbreviation for $\neg\phi \vee \psi$, and $\phi \leftrightarrow \psi$ is an abbreviation for $(\phi \rightarrow \psi) \wedge (\psi \rightarrow \phi)$. In fact, we did not really need to include the symbols $\vee$ and $\forall$. We could have considered $\phi \vee \psi$ as an abbreviation for $\neg(\neg\phi \wedge \neg\psi)$ and $\forall v \phi$ as an abbreviation for $\neg(\exists v \neg\phi)$. Viewing these as abbreviations will be an advantage

[1] To simplify some bookkeeping we will tacitly restrict our attention to formulas where in each subformula no variable v_i has both free and bound occurrences. For example we will not consider formulas such as $(v_1 > 0 \vee \exists v_1\ v_1 \cdot v_1 = v_2)$, because this formula could be replaced by the clearer formula $v_1 > 0 \vee \exists v_3\ v_3 \cdot v_3 = v_2$ with the same meaning. There are some areas of mathematical logic where one wants to be frugal with variables, but we will not consider such issues here. See [**?**] for a definition of satisfaction for arbitrary formulas.

when we are proving theorems by induction on formulas because it eliminates the $\vee$ and $\forall$ cases.

We also will use the abbreviations $\bigwedge_{i=1}^{n} \psi_i$ and $\bigvee_{i=1}^{n} \psi_i$ for $\psi_1 \wedge \ldots \wedge \psi_n$ and $\psi_1 \vee \ldots \vee \psi_n$, respectively.

• In addition to $v_1, v_2, \ldots$, we will use $w, x, y, z, \ldots$ as variable symbols.

• It is important to note that the quantifiers $\exists$ and $\forall$ range only over elements of the model. For example the statement that an ordering is complete (i.e., every bounded subset has a least upper bound) cannot be expressed as a formula because we cannot quantify over subsets. The fact that we are limited to quantification over elements of the structure is what makes it "first-order" logic.

When proving results about satisfaction in models, we often must do an induction on the construction of formulas. The next proposition asserts that if a formula without quantifiers is true in some structure, then it is true in every extension. It is proved by induction on quantifier-free formulas.

Proposition 1.8 *Suppose that $\mathcal{M}$ is a substructure of $\mathcal{N}$, $\overline{a} \in M$, and $\phi(\overline{v})$ is a quantifier-free formula. Then, $\mathcal{M} \models \phi(\overline{a})$ if and only if $\mathcal{N} \models \phi(\overline{a})$.*

Proof

Claim If $t(\overline{v})$ is a term and $\bar{b} \in M$, then $t^{\mathcal{M}}(\bar{b}) = t^{\mathcal{N}}(\bar{b})$. This is proved by induction on terms.

If t is the constant symbol c, then $c^{\mathcal{M}} = c^{\mathcal{N}}$.

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

Suppose that $t = f(t_1, \ldots, t_n)$, where f is an n-ary function symbol, $t_1, \ldots, t_n$ are terms, and $t_i^{\mathcal{M}}(\bar{b}) = t_i^{\mathcal{N}}(\bar{b})$ for $i = 1, \ldots, n$. Because $\mathcal{M} \subseteq \mathcal{N}$, $f^{\mathcal{M}} = f^{\mathcal{N}}|M^n$. Thus,

$$\begin{array}{rcl} t^{\mathcal{M}}(\bar{b}) & = & f^{\mathcal{M}}(t_1^{\mathcal{M}}(\bar{b}), \ldots, t_n^{\mathcal{M}}(\bar{b})) \\ & = & f^{\mathcal{N}}(t_1^{\mathcal{M}}(\bar{b}), \ldots, t_n^{\mathcal{M}}(\bar{b})) \\ & = & f^{\mathcal{N}}(t_1^{\mathcal{N}}(\bar{b}), \ldots, t_n^{\mathcal{N}}(\bar{b})) \\ & = & t^{\mathcal{N}}(\bar{b}). \end{array}$$

We now prove the proposition by induction on formulas.

If ϕ is $t_1 = t_2$, then

$$\mathcal{M} \models \phi(\overline{a}) \Leftrightarrow t_1^{\mathcal{M}}(\overline{a}) = t_2^{\mathcal{M}}(\overline{a}) \Leftrightarrow t_1^{\mathcal{N}}(\overline{a}) = t_2^{\mathcal{N}}(\overline{a}) \Leftrightarrow \mathcal{N} \models \phi(\overline{a}).$$

If ϕ is $R(t_1, \ldots, t_n)$, where R is an n-ary relation symbol, then

$$\begin{array}{rcl} \mathcal{M} \models \phi(\overline{a}) & \Leftrightarrow & (t_1^{\mathcal{M}}(\overline{a}), \ldots, t_n^{\mathcal{M}}(\overline{a})) \in R^{\mathcal{M}} \\ & \Leftrightarrow & (t_1^{\mathcal{M}}(\overline{a}), \ldots, t_n^{\mathcal{M}}(\overline{a})) \in R^{\mathcal{N}} \\ & \Leftrightarrow & (t_1^{\mathcal{N}}(\overline{a}), \ldots, t_n^{\mathcal{N}}(\overline{a})) \in R^{\mathcal{N}} \\ & \Leftrightarrow & \mathcal{N} \models \phi(\overline{a}). \end{array}$$

Thus, the proposition is true for all atomic formulas.

Suppose that the proposition is true for ψ and that ϕ is $\neg\psi$. Then,

$$\mathcal{M} \models \phi(\overline{a}) \Leftrightarrow \mathcal{M} \not\models \psi(\overline{a}) \Leftrightarrow \mathcal{N} \not\models \psi(\overline{a}) \Leftrightarrow \mathcal{N} \models \phi(\overline{a}).$$

Finally, suppose that the proposition is true for ψ_0 and ψ_1 and that ϕ is $\psi_0 \wedge \psi_1$. Then,

$$\begin{array}{rcl} \mathcal{M} \models \phi(\overline{a}) & \Leftrightarrow & \mathcal{M} \models \psi_0(\overline{a}) \text{ and } \mathcal{M} \models \psi_1(\overline{a}) \\ & \Leftrightarrow & \mathcal{N} \models \psi_0(\overline{a}) \text{ and } \mathcal{M} \models \psi_1(\overline{a}) \\ & \Leftrightarrow & \mathcal{N} \models \phi(\overline{a}). \end{array}$$

We have shown that the proposition holds for all atomic formulas and that if it holds for ϕ and ψ, then it also holds for $\neg\phi$ and $\phi \wedge \psi$. Because the set of quantifier-free formulas is the smallest set of formulas containing the atomic formulas and closed under negation and conjunction, the proposition is true for all quantifier-free formulas.

Elementary Equivalence and Isomorphism

We next consider structures that satisfy the same sentences.

Definition 1.9 We say that two $\mathcal{L}$-structures $\mathcal{M}$ and $\mathcal{N}$ are *elementarily equivalent* and write $\mathcal{M} \equiv \mathcal{N}$ if

$$\mathcal{M} \models \phi \text{ if and only if } \mathcal{N} \models \phi$$

for all $\mathcal{L}$-sentences ϕ.

We let Th($\mathcal{M}$), the *full theory of* $\mathcal{M}$, be the set of $\mathcal{L}$-sentences ϕ such that $\mathcal{M} \models \phi$. It is easy to see that $\mathcal{M} \equiv \mathcal{N}$ if and only if Th($\mathcal{M}$)= Th($\mathcal{N}$). Our next result shows that Th($\mathcal{M}$) is an isomorphism invariant of $\mathcal{M}$. The proof uses the important technique of "induction on formulas."

Theorem 1.10 *Suppose that $j : \mathcal{M} \to \mathcal{N}$ is an isomorphism. Then, $\mathcal{M} \equiv \mathcal{N}$.*

Proof We show by induction on formulas that $\mathcal{M} \models \phi(a_1, \ldots, a_n)$ if and only if $\mathcal{N} \models \phi(j(a_1), \ldots, j(a_n))$ for all formulas ϕ.

We first must show that terms behave well.

Claim Suppose that t is a term and the free variables in t are from $\overline{v} = (v_1, \ldots, v_n)$. For $\overline{a} = (a_1, \ldots, a_n) \in M$, we let $j(\overline{a})$ denote $(j(a_1), \ldots, j(a_n))$. Then $j(t^{\mathcal{M}}(\overline{a})) = t^{\mathcal{N}}(j(\overline{a}))$.

We prove this by induction on terms.

i) If $t = c$, then $j(t^{\mathcal{M}}(\overline{a})) = j(c^{\mathcal{M}}) = c^{\mathcal{N}} = t^{\mathcal{N}}(j(\overline{a}))$.

ii) If $t = v_i$, then $j(t^{\mathcal{M}}(\overline{a})) = j(a_i) = t^{\mathcal{N}}(j(a_i))$.

iii) If $t = f(t_1, \ldots, t_m)$, then

$$\begin{array}{rcl} j(t^{\mathcal{M}}(\overline{a})) & = & j(f^{\mathcal{M}}(t_1^{\mathcal{M}}(\overline{a}), \ldots, t_m^{\mathcal{M}}(\overline{a}))) \\ & = & f^{\mathcal{N}}(j(t_1^{\mathcal{M}}(\overline{a})), \ldots, j(t_m^{\mathcal{M}}(\overline{a}))) \\ & = & f^{\mathcal{N}}(t_1^{\mathcal{N}}(j(\overline{a})), \ldots, t_m^{\mathcal{N}}(j(\overline{a}))) \\ & = & t^{\mathcal{N}}(j(\overline{a})). \end{array}$$

We proceed by induction on formulas.

i) If $\phi(\overline{v})$ is $t_1 = t_2$, then

$$\begin{aligned}
\mathcal{M} \models \phi(\overline{a}) &\Leftrightarrow t_1^{\mathcal{M}}(\overline{a}) = t_2^{\mathcal{M}}(\overline{a}) \\
&\Leftrightarrow j(t_1^{\mathcal{M}}(\overline{a})) = j(t_2^{\mathcal{M}}(\overline{a})) \text{ because } j \text{ is injective} \\
&\Leftrightarrow t_1^{\mathcal{N}}(j(\overline{a})) = t_2^{\mathcal{N}}(j(\overline{a})) \\
&\Leftrightarrow \mathcal{N} \models \phi(j(\overline{a})).
\end{aligned}$$

ii) If $\phi(\overline{v})$ is $R(t_1, \ldots, t_n)$, then

$$\begin{aligned}
\mathcal{M} \models \phi(\overline{a}) &\Leftrightarrow (t_1^{\mathcal{M}}(\overline{a}), \ldots, t_n^{\mathcal{M}}(\overline{a})) \in R^{\mathcal{M}} \\
&\Leftrightarrow (j(t_1^{\mathcal{M}}(\overline{a})), \ldots, j(t_n^{\mathcal{M}}(\overline{a}))) \in R^{\mathcal{N}} \\
&\Leftrightarrow (t_1^{\mathcal{N}}(j(\overline{a})), \ldots, t_n^{\mathcal{N}}(j(\overline{a}))) \in R^{\mathcal{N}} \\
&\Leftrightarrow \mathcal{N} \models \phi(j(\overline{a})).
\end{aligned}$$

iii) If ϕ is $\neg\psi$, then by induction

$$\mathcal{M} \models \phi(\overline{a}) \Leftrightarrow \mathcal{M} \not\models \psi(\overline{a}) \Leftrightarrow \mathcal{N} \not\models \psi(j(\overline{a})) \Leftrightarrow \mathcal{N} \models \phi(j(\overline{a})).$$

iv) If ϕ is $\psi \wedge \theta$, then

$$\begin{aligned}
\mathcal{M} \models \phi(\overline{a}) &\Leftrightarrow \mathcal{M} \models \psi(\overline{a}) \text{ and } \mathcal{M} \models \theta(\overline{a}) \\
&\Leftrightarrow \mathcal{N} \models \psi(j(\overline{a})) \text{ and } \mathcal{N} \models \theta(j(\overline{a})) \Leftrightarrow \mathcal{N} \models \phi(j(\overline{a})).
\end{aligned}$$

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

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

Theories

Let $\mathcal{L}$ be a language. An *$\mathcal{L}$-theory* T is simply a set of $\mathcal{L}$-sentences. We say that $\mathcal{M}$ is a *model* of T and write $\mathcal{M} \models T$ if $\mathcal{M} \models \phi$ for all sentences $\phi \in T$.

The set $T = \{\forall x\ x = 0, \exists x\ x \neq 0\}$ is a theory. Because the two sentences in T are contradictory, there are no models of T. We say that a theory is *satisfiable* if it has a model.

We say that a class of $\mathcal{L}$-structures $\mathcal{K}$ is an *elementary class* if there is an $\mathcal{L}$-theory T such that $\mathcal{K} = \{\mathcal{M} : \mathcal{M} \models T\}$.

One way to get a theory is to take $\mathrm{Th}(\mathcal{M})$, the full theory of an $\mathcal{L}$-structure $\mathcal{M}$. In this case, the elementary class of models of $\mathrm{Th}(\mathcal{M})$ is exactly the class of $\mathcal{L}$-structures elementarily equivalent to $\mathcal{M}$. More typically, we have a class of structures in mind and try to write a set of properties T describing these structures. We call these sentences *axioms* for the elementary class.

We give a few basic examples of theories and elementary classes that we will return to frequently.

Example 1.11 *Infinite Sets*

Let $\mathcal{L} = \emptyset$.

Consider the $\mathcal{L}$-theory where we have, for each n, the sentence ϕ_n given by

$$\exists x_1 \exists x_2 \ldots \exists x_n \bigwedge_{i<j\leq n} x_i \neq x_j.$$

The sentence ϕ_n asserts that there are at least n distinct elements, and an $\mathcal{L}$-structure $\mathcal{M}$ with universe M is a model of T if and only if M is infinite.

Example 1.12 *Linear Orders*

Let $\mathcal{L} = \{<\}$, where $<$ is a binary relation symbol. The class of linear orders is axiomatized by the $\mathcal{L}$-sentences
$\forall x\ \neg(x < x)$,
$\forall x \forall y \forall z\ ((x < y \wedge y < z) \rightarrow x < z)$,
$\forall x \forall y\ (x < y \vee x = y \vee y < x)$.

There are a number of interesting extensions of the theory of linear orders. For example, we could add the sentence

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

to get the theory of dense linear orders, or we could instead add the sentence

$$\forall x \exists y\ (x < y \wedge \forall z (x < z \rightarrow (z = y \vee y < z)))$$

to get the theory of linear orders where every element has a unique successor. We could also add sentences that either assert or deny the existence of top or bottom elements.

Example 1.13 *Equivalence Relations*

Let $\mathcal{L} = \{E\}$, where E is a binary relation symbol. The theory of equivalence relations is given by the sentences
$\forall x\ E(x,x)$,
$\forall x \forall y (E(x,y) \rightarrow E(y,x))$,
$\forall x \forall y \forall z ((E(x,y) \wedge E(y,z)) \rightarrow E(x,z))$.

If we added the sentence

$$\forall x \exists y (x \neq y \wedge E(x,y) \wedge \forall z\ (E(x,z) \rightarrow (z = x \vee z = y)))$$

we would have the theory of equivalence relations where every equivalence class has exactly two elements. If instead we added the sentence

$$\exists x \exists y (\neg E(x,y) \wedge \forall z (E(x,z) \vee E(y,z)))$$

and the infinitely many sentences

$$\forall x \exists x_1 \exists x_2 \ldots \exists x_n \left(\bigwedge_{i<j\leq n} x_i \neq x_j \wedge \bigwedge_{i=1}^{n} E(x, x_i) \right)$$

we would axiomatize the class of equivalence relations with exactly two classes, both of which are infinite.

Example 1.14 *Graphs*

Let $\mathcal{L} = \{R\}$ where R is a binary relation. We restrict our attention to irreflexive graphs. These are axiomatized by the two sentences
$\forall x\ \neg R(x,x)$,
$\forall x \forall y\ (R(x,y) \rightarrow R(y,x))$.

Example 1.15 *Groups*

Let $\mathcal{L} = \{\cdot, e\}$, where $\cdot$ is a binary function symbol and e is a constant symbol. We will write $x \cdot y$ rather than $\cdot(x,y)$. The class of groups is axiomatized by
$\forall x\ e \cdot x = x \cdot e = x$,
$\forall x \forall y \forall z\ x \cdot (y \cdot z) = (x \cdot y) \cdot z$,
$\forall x \exists y\ x \cdot y = y \cdot x = e$.
We could also axiomatize the class of Abelian groups by adding $\forall x \forall y\ x \cdot y = y \cdot x$.

Let $\phi_n(x)$ be the $\mathcal{L}$-formula

$$\underbrace{x \cdot x \cdots x}_{n-\text{times}} = e;$$

which asserts that $nx = e$.

We could axiomatize the class of torsion-free groups by adding $\{\forall x\ (x = e \vee \neg \phi_n(x)) : n \geq 2\}$ to the axioms for groups. Alternatively, we could axiomatize the class of groups where every element has order at most N by adding to the axioms for groups the sentence

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

Note that the same idea will not work to axiomatize the class of torsion groups because the corresponding sentence would be infinitely long. In the next chapter, we will see that the class of torsion groups is not elementary.

Let $\psi_n(x,y)$ be the formula

$$\underbrace{x \cdot x \cdots x}_{n-\text{times}} = y;$$

which asserts that $x^n = y$. We can axiomatize the class of divisible groups by adding the axioms $\{\forall y \exists x\ \psi_n(x,y) : n \geq 2\}$.

It will often be useful to deal with additive groups instead of multiplicative groups. The class of additive groups is the collection structures in the language $\mathcal{L} = \{+, 0\}$, axiomatized as above replacing $\cdot$ by $+$ and e by 0.

Example 1.16 *Ordered Abelian Groups*

Let $\mathcal{L} = \{+, <, 0\}$, where $+$ is a binary function symbol, $<$ is a binary relation symbol, and 0 is a constant symbol. The axioms for ordered groups are
the axioms for additive groups,
the axioms for linear orders, and
$\forall x \forall y \forall z (x < y \rightarrow x + z < y + z)$.

Example 1.17 *Left R-modules*

Let R be a ring with multiplicative identity 1. Let $\mathcal{L} = \{+, 0\} \cup \{r : r \in R\}$ where $+$ is a binary function symbol, 0 is a constant, and r is a unary function symbol for $r \in R$. In an R-module, we will interpret r as scalar multiplication by R. The axioms for left R-modules are
the axioms for additive commutative groups,
$\forall x\ r(x + y) = r(x) + r(y)$ for each $r \in R$,
$\forall x\ (r + s)(x) = r(x) + s(x)$ for each $r, s \in R$,
$\forall x\ r(s(x)) = rs(x)$ for $r, s \in R$,
$\forall x\ 1(x) = x$.

Example 1.18 *Rings and Fields*

Let $\mathcal{L}_\mathrm{r}$ be the language of rings $\{+, -, \cdot, 0, 1\}$, where $+$, $-$, and $\cdot$ are binary function symbols and 0 and 1 are constants. The axioms for rings are given by
the axioms for additive commutative groups,
$\forall x \forall y \forall z\ (x - y = z \leftrightarrow x = y + z)$,
$\forall x\ x \cdot 0 = 0$,
$\forall x \forall y \forall z\ (x \cdot (y \cdot z) = (x \cdot y) \cdot z)$,
$\forall x\ x \cdot 1 = 1 \cdot x = x$,
$\forall x \forall y \forall z\ x \cdot (y + z) = (x \cdot y) + (x \cdot z)$,
$\forall x \forall y \forall z\ (x + y) \cdot z = (x \cdot z) + (y \cdot z)$.

The second axiom is only necessary because we include $-$ in the language (this will be useful later). We axiomatize the class of fields by adding the axioms
$\forall x \forall y\ x \cdot y = y \cdot x$,
$\forall x\ (x \neq 0 \rightarrow \exists y\ x \cdot y = 1)$.

We axiomatize the class of algebraically closed fields by adding to the field axioms the sentences

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

for $n = 1, 2, \ldots$. Let ACF be the axioms for algebraically closed fields.

Let ψ_p be the $\mathcal{L}_\mathrm{r}$-sentence $\forall x \underbrace{x + \ldots + x}_{p\mathrm{-times}} = 0$, which asserts that a field has characteristic p. For $p > 0$ a prime, let $\mathrm{ACF}_p = \mathrm{ACF} \cup \{\psi_p\}$ and $\mathrm{ACF}_0 = \mathrm{ACF} \cup \{\neg\psi_p : p > 0\}$, be the theories of algebraically closed fields of characteristic p and characteristic zero, respectively.

Example 1.19 *Ordered Fields*

Let $\mathcal{L}_{\rm or} = \mathcal{L}_{\rm r} \cup \{<\}$. The class of ordered fields is axiomatized by the axioms for fields,
the axioms for linear orders,
$\forall x \forall y \forall z\ (x < y \rightarrow x + z < y + z)$,
$\forall x \forall y \forall z\ ((x < y \wedge z > 0) \rightarrow x \cdot z < y \cdot z)$.

Example 1.20 *Differential Fields*

Let $\mathcal{L} = \mathcal{L}_{\rm r} \cup \{\delta\}$, where δ is a unary function symbol. The class of differential fields is axiomatized by
the axioms of fields,
$\forall x \forall y\ \delta(x + y) = \delta(x) + \delta(y)$,
$\forall x \forall y\ \delta(x \cdot y) = x \cdot \delta(y) + y \cdot \delta(x)$.

Example 1.21 *Peano Arithmetic*

Let $\mathcal{L} = \{+, \cdot, s, 0\}$, where $+$ and $\cdot$ are binary functions, s is a unary function, and 0 is a constant. We think of s as the successor function $x \mapsto x + 1$. The Peano axioms for arithmetic are the sentences
$\forall x\ s(x) \neq 0$,
$\forall x\ (x \neq 0 \rightarrow \exists y\ s(y) = x)$,
$\forall x\ x + 0 = x$,
$\forall x\ \forall y\ x + (s(y)) = s(x + y)$,
$\forall x\ \ x \cdot 0 = 0$,
$\forall x \forall y\ x \cdot s(y) = (x \cdot y) + x$,
and the axioms $\text{Ind}(\phi)$ for each formula $\phi(v, \overline{w})$, where $\text{Ind}(\phi)$ is the sentence
$\forall \overline{w}\ [(\phi(0, \overline{w}) \wedge \forall v\ (\phi(v, \overline{w}) \rightarrow \phi(s(v), \overline{w}))) \rightarrow \forall x\ \phi(x, \overline{w})]$.

The axiom $\text{Ind}(\phi)$ formalizes an instance of induction. It asserts that if $\overline{a} \in M$, $X = \{m \in M : \mathcal{M} \models \phi(m, \overline{a})\}$, $0 \in X$, and $s(m) \in X$ whenever $m \in X$, then $X = M$.

Logical Consequence

Definition 1.22 Let T be an $\mathcal{L}$-theory and ϕ an $\mathcal{L}$-sentence. We say that ϕ is a *logical consequence* of T and write $T \models \phi$ if $\mathcal{M} \models \phi$ whenever $\mathcal{M} \models T$.

We give two examples.

Proposition 1.23 *a) Let* $\mathcal{L} = \{+, <, 0\}$ *and let* T *be the theory of ordered Abelian groups. Then,* $\forall x (x \neq 0 \rightarrow x + x \neq 0)$ *is a logical consequence of* T.

b) Let T *be the theory of groups where every element has order 2. Then,* $T \not\models \exists x_1 \exists x_2 \exists x_3 (x_1 \neq x_2 \wedge x_2 \neq x_3 \wedge x_1 \neq x_3)$.

Proof

a) Suppose that $\mathcal{M} = (M, +, <, 0)$ is an ordered Abelian group. Let $a \in M \setminus \{0\}$. We must show that $a + a \neq 0$. Because $(M, <)$ is a linear order $a < 0$

or $0 < a$. If $a < 0$, then $a + a < 0 + a = a < 0$. Because $\neg(0 < 0)$, $a + a \neq 0$. If $0 < a$, then $0 < a = 0 + a < a + a$ and again $a + a \neq 0$.

b) Clearly, $\mathbb{Z}/2\mathbb{Z} \models T \wedge \neg\exists x_1 \exists x_2 \exists x_3 (x_1 \neq x_2 \wedge x_2 \neq x_3 \wedge x_1 \neq x_3)$.

In general, to show that $T \models \phi$, we give an informal mathematical proof as above that $\mathcal{M} \models \phi$ whenever $\mathcal{M} \models T$. To show that $T \not\models \phi$, we usually construct a counterexample.

Definable Sets

Definition 1.24 Let $\mathcal{M} = (M, \ldots)$ be an $\mathcal{L}$-structure. We say that $X \subseteq M^n$ is *definable* if and only if there is an $\mathcal{L}$-formula $\phi(v_1, \ldots, v_n, w_1, \ldots, w_m)$ and $\overline{b} \in M^m$ such that $X = \{\overline{a} \in M^n : \mathcal{M} \models \phi(\overline{a}, \overline{b})\}$. We say that $\phi(\overline{v}, \overline{b})$ *defines* X. We say that X is *A-definable* or *definable over A* if there is a formula $\psi(\overline{v}, w_1, \ldots, w_l)$ and $\overline{b} \in A^l$ such that $\psi(\overline{v}, \overline{b})$ defines X.

We give a number of examples using $\mathcal{L}_{\mathrm{r}}$, the language of rings.

- Let $\mathcal{M} = (R, +, -, \cdot, 0, 1)$ be a ring. Let $p(X) \in R[X]$. Then, $Y = \{x \in R : p(x) = 0\}$ is definable. Suppose that $p(X) = \sum_{i=0}^{m} a_i X^i$. Let $\phi(v, w_0, \ldots, w_n)$ be the formula

$$w_n \cdot \underbrace{v \cdots v}_{n-\text{times}} + \ldots + w_1 \cdot v + w_0 = 0$$

(in the future, when no confusion arises, we will abbreviate such a formula as "$w_n v^n + \ldots + w_1 v + w_0 = 0$"). Then, $\phi(v, a_0, \ldots, a_n)$ defines Y. Indeed, Y is A-definable for any $A \supseteq \{a_0, \ldots, a_n\}$.

- Let $\mathcal{M} = (\mathbb{R}, +, -, \cdot, 0, 1)$ be the field of real numbers. Let $\phi(x, y)$ be the formula

$$\exists z (z \neq 0 \wedge y = x + z^2).$$

Because $a < b$ if and only if $\mathcal{M} \models \phi(a, b)$, the ordering is $\emptyset$-definable.

- Let $\mathcal{M} = (\mathbb{Z}, +, -, \cdot, 0, 1)$ be the ring of integers. Let $X = \{(m, n) \in \mathbb{Z}^2 : m < n\}$. Then, X is definable (indeed $\emptyset$-definable). By Lagrange's Theorem, every nonnegative integer is the sum of four squares. Thus, if we let $\phi(x, y)$ be the formula

$$\exists z_1 \exists z_2 \exists z_3 \exists z_4 (z_1 \neq 0 \wedge y = x + z_1^2 + z_2^2 + z_3^2 + z_4^2),$$

then $X = \{(m, n) \in \mathbb{Z}^2 : \mathcal{M} \models \phi(m, n)\}$.

- Let F be a field and $\mathcal{M} = (F[X], +, -, \cdot, 0, 1)$ be the ring of polynomials over F. Then F is definable in $\mathcal{M}$. Indeed, F is the set of units of $F[X]$ and is defined by the formula $x = 0 \vee \exists y\ xy = 1$.

- Let $\mathcal{M} = (\mathbb{C}(X), +, -, \cdot, 0, 1)$ be the field of complex rational functions in one variable. We claim that $\mathbb{C}$ is defined in $\mathbb{C}(X)$ by the formula

$$\exists x \exists y\ y^2 = v \wedge x^3 + 1 = v.$$

For any $z \in \mathbb{C}$ we can find x and y such that $y^2 = x^3 + 1 = z$. Suppose that h is a nonconstant rational function and that there are nonconstant rational functions f and g such that $h = g^2 = f^3 + 1$. Then $t \mapsto (f(t), g(t))$ is a nonconstant rational function from an open subset of $\mathbb{C}$ into the curve E given by the equation $y^2 = x^3 + 1$. But E is an elliptic curve and it is known (see for example [**?**]) that there are no such functions.

A similar argument shows that $\mathbb{C}$ is the set of rational functions f such that f and $f + 1$ are both fourth powers. These ideas generalize to show that $\mathbb{C}$ is definable in any finite algebraic extension of $\mathbb{C}(X)$.

• Let $\mathcal{M} = (\mathbb{Q}_p, +, -, \cdot, 0, 1)$ be the field of p-adic numbers. Then $\mathbb{Z}_p$ the ring of p-adic integers is definable. Suppose $p \neq 2$ (we leave $\mathbb{Q}_2$ for Exercise **??**) and $\phi(x)$ is the formula $\exists y\ y^2 = px^2 + 1$. We claim that $\phi(x)$ defines $\mathbb{Z}_p$.

First, suppose that $y^2 = pa^2 + 1$. Let v denote the p-adic valuation. Because $v(pa^2) = 2v(a) + 1$, if $v(a) < 0$, then $v(pa^2)$ is an odd negative integer and $v(y^2) = v(pa^2 + 1) = v(pa^2)$. On the other hand, $v(y^2) = 2v(y)$, an even integer. Thus, if $\mathcal{M} \models \phi(a)$, then $v(a) \geq 0$ so $a \in \mathbb{Z}_p$.

On the other hand, suppose that $a \in \mathbb{Z}_p$. Let $F(X) = X^2 - (pa^2 + 1)$. Let $\overline{F}$ be the reduction of F mod p. Because $v(a) \geq 0$, $v(pa) > 0$ and $\overline{F}(X) = X^2 - 1$ and $\overline{F}' = 2X$. Thus, $\overline{F}(1) = 0$ and $\overline{F}'(1) \neq 0$ so, by Hensel's Lemma, there is $b \in \mathbb{Z}_p$ such that $F(b) = 0$. Hence $\mathcal{M} \models \phi(a)$.

• Let $\mathcal{M} = (\mathbb{Q}, +, -, \cdot, 0, 1)$ be the field of rational numbers. Let $\phi(x, y, z)$ be the formula

$$\exists a \exists b \exists c\ xyz^2 + 2 = a^2 + xy^2 - yc^2$$

and let $\psi(x)$ be the formula

$$\forall y \forall z\ ([\phi(y, z, 0) \wedge (\forall w(\phi(y, z, w) \rightarrow \phi(y, z, w + 1)))] \rightarrow \phi(y, z, x)).$$

A remarkable result of Julia Robinson (see [**?**]) shows that $\psi(x)$ defines the integers in $\mathbb{Q}$.

• Consider the natural numbers $\mathbb{N}$ as an $\mathcal{L} = \{+, \cdot, 0, 1\}$ structure. The definable sets are quite complex. For example, there is an $\mathcal{L}$-formula $T(e, x, s)$ such that $\mathbb{N} \models T(e, x, s)$ if and only if the Turing machine with program coded by e halts on input x in at most s steps (see, for example, [**?**]). Thus, the Turing machine with program e halts on input x if and only if $\mathbb{N} \models \exists s\ T(e, x, s)$, so the set of halting computations is definable. It is well known that this set is not computable (see, for example, [**?**]). This leads to an interesting conclusion.

Proposition 1.25 *The full $\mathcal{L}$-theory of the natural numbers is undecidable (i.e., there is no algorithm that when given an $\mathcal{L}$-sentence ψ as input will always halt answering "yes" if $\mathbb{N} \models \psi$ and "no" if $\mathbb{N} \models \neg\psi$).*

Proof For each e and x, let $\phi_{e,x}$ be the $\mathcal{L}$-sentence

$$\exists s\ T(\underbrace{1 + \ldots + 1}_{e-\text{times}}, \underbrace{1 + \ldots + 1}_{x-\text{times}}, s).$$

If there were such an algorithm we could decide whether the program coded by e halts on input x by asking whether $\mathbb{N} \models \phi_{e,x}$.

Recursively enumerable sets have simple mathematical definitions. By the Matijasevič–Robinson–Davis–Putnam solution to Hilbert's 10th Problem (see [?]) for any recursively enumerable set $A \subseteq \mathbb{N}^n$ there is a polynomial

$$p(X_1, \ldots, X_n, Y_1, \ldots, Y_m) \in \mathbb{Z}[\overline{X}, \overline{Y}]$$

such that

$$A = \{\overline{x} \in \mathbb{N}^n : \mathbb{N} \models \exists y_1 \ldots \exists y_m \; p(\overline{x}, \overline{y}) = 0\}.$$

The following example will be useful later.

Lemma 1.26 *Let $\mathcal{L}_{\mathrm{r}}$ be the language of ordered rings and $(\mathbb{R}, +, -, \cdot, <, 0, 1)$ be the ordered field of real numbers. Suppose that $X \subseteq \mathbb{R}^n$ is A-definable. Then, the topological closure of X is also A-definable.*

Proof Let $\phi(v_1, \ldots, v_n, \overline{a})$ define X. Let $\psi(v_1, \ldots, v_n, \overline{w})$ be the formula

$$\forall \epsilon \left[\epsilon > 0 \rightarrow \exists y_1, \ldots, y_n \; (\phi(\overline{y}, \overline{w}) \wedge \sum_{i=1}^{n} (v_i - y_i)^2 < \epsilon) \right].$$

Then, $\overline{b}$ is in the closure of X if and only if $\mathcal{M} \models \psi(\overline{b}, \overline{a})$.

How do we show that $X \subset M^n$ is not definable? The following proposition will often be useful.

Proposition 1.27 *Let $\mathcal{M}$ be an $\mathcal{L}$-structure. If $X \subset M^n$ is A-definable, then every $\mathcal{L}$-automorphism of $\mathcal{M}$ that fixes A pointwise fixes X setwise (that is, if σ is an automorphism of M and $\sigma(a) = a$ for all $a \in A$, then $\sigma(X) = X$).*

Proof Let $\psi(\overline{v}, \overline{a})$ be the $\mathcal{L}$-formula defining X where $\overline{a} \in A$. Let σ be an automorphism of $\mathcal{M}$ with $\sigma(\overline{a}) = \overline{a}$, and let $\overline{b} \in M^n$.

In the proof of Theorem 1.10, we showed that if $j : \mathcal{M} \rightarrow \mathcal{N}$ is an isomorphism, then $\mathcal{M} \models \phi(\overline{a})$ if and only if $\mathcal{N} \models \phi(j(\overline{a}))$. Thus

$$\mathcal{M} \models \psi(\overline{b}, \overline{a}) \leftrightarrow \mathcal{M} \models \psi(\sigma(\overline{b}), \sigma(\overline{a})) \Leftrightarrow \mathcal{M} \models \psi(\sigma(\overline{b}), \overline{a}).$$

In other words, $\overline{b} \in X$ if and only if $\sigma(\overline{b}) \in X$ as desired.

We give a sample application.

Corollary 1.28 *The set of real numbers is not definable in the field of complex numbers.*

Proof If $\mathbb{R}$ were definable, then it would be definable over a finite $A \subset \mathbb{C}$. Let $r, s \in \mathbb{C}$ be algebraically independent over A with $r \in \mathbb{R}$ and $s \notin \mathbb{R}$. There is an automorphism σ of $\mathbb{C}$ such that $\sigma | A$ is the identity and $\sigma(r) = s$. Thus, $\sigma(\mathbb{R}) \neq \mathbb{R}$ and $\mathbb{R}$ is not definable over A.

This proof worked because $\mathbb{C}$ has many automorphisms. The situation is much different for $\mathbb{R}$. Any automorphism of the real field must fix the rational numbers. Because the ordering is definable it must be preserved by any automorphism. Because the rationals are dense in $\mathbb{R}$, the only automorphism of the real field is the identity. Most subsets of $\mathbb{R}$ are undefinable (there are $2^{2^{\aleph_0}}$ subsets of $\mathbb{R}$ and only $2^{\aleph_0}$ possible definitions), but we cannot use Proposition 1.27 to show any particular set is undefinable. In fact, the converse to Proposition 1.27 holds for sufficiently rich models.

2 The Compactness Theorem

Let T be an $\mathcal{L}$-theory and ϕ an $\mathcal{L}$-sentence. To show that $T \models \phi$, we must show that ϕ holds in every model of T. Checking all models of T sounds like a daunting task, but in practice we usually show that $T \models \phi$ by giving an informal mathematical proof that ϕ is true in every model of T. One of the first great achievements of mathematical logic was giving a rigorous definition of "proof" that completely captures the notion of "logical consequence."

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

"from ϕ and ψ conclude $\phi \wedge \psi$," or

"from $\phi \wedge \psi$ conclude ϕ."

It will not be important for our purposes to go into the details of the proof system, but we stress the following points. (See [**?**], for example, for complete details of one possible proof system.)

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

Note that the last point does not say that there is an algorithm that will decide if $T \vdash \phi$. It only says that there is an algorithm that can check each purported proof.

We say that a language $\mathcal{L}$ is *recursive* if there is an algorithm that decides whether a sequence of symbols is an $\mathcal{L}$-formula. We say that an $\mathcal{L}$-theory T is recursive if there is an algorithm that, when given an $\mathcal{L}$-sentence ϕ as input, decides whether $\phi \in T$.

Proposition 2.1 *If $\mathcal{L}$ is a recursive language and T is a recursive $\mathcal{L}$-theory, then $\{\phi : T \vdash \phi\}$ is recursively enumerable; that is, there is an algorithm, that when given ϕ as input will halt accepting if $T \vdash \phi$ and not halt if $T \not\vdash \phi$.*

Proof There is $\sigma_0, \sigma_1, \sigma_2, \ldots$, a computable listing of all finite sequences of $\mathcal{L}$-formulas. At stage i of our algorithm, we check to see whether σ_i is a proof of ψ from T. This involves checking that each formula either is in T (which we can check because T is recursive) or follows by a logical rule from earlier formulas in the sequence σ_i and that the last formula is ϕ. If σ_i is a proof of ϕ from T, then we halt accepting; otherwise we go on to stage $i + 1$.

Remarkably, the finitistic syntactic notion of "proof" completely captures the semantic notion of "logical consequence."

Theorem 2.2 (Gödel's Completeness Theorem) *Let T be an $\mathcal{L}$-theory and ϕ an $\mathcal{L}$-sentence, then $T \models \phi$ if and only if $T \vdash \phi$.*

The Completeness Theorem gives a criterion for testing whether an $\mathcal{L}$-theory is satisfiable. We say that an $\mathcal{L}$-theory T is *inconsistent* if $T \vdash (\phi \wedge \neg\phi)$ for some sentence ϕ; otherwise we say that T is *consistent.* Because our proof system is sound, any satisfiable theory is consistent. The Completeness Theorem implies that the converse is true.

Corollary 2.3 *T is consistent if and only if T is satisfiable.*

Proof Suppose that T is not satisfiable. Because there are no models of T, every model of T is a model of $(\phi \wedge \neg\phi)$. Thus, $T \models (\phi \wedge \neg\phi)$ and by the Completeness Theorem $T \vdash (\phi \wedge \neg\phi)$.

This has a deceptively simple consequence.

Theorem 2.4 (Compactness Theorem) *T is satisfiable if and only if every finite subset of T is satisfiable.*

Proof Clearly, if T is satisfiable, then every subset of T is satisfiable. On the other hand, if T is not satisfiable, then T is inconsistent. Let σ be a proof of a contradiction from T. Because σ is finite, only finitely many assumptions from T are used in the proof. Thus, there is a finite $T_0 \subseteq T$ such that σ is a proof of a contradiction from T_0. But then T_0 is a finite unsatisfiable subset of T.

Although it is a simple consequence of the Completeness Theorem and the finite nature of proof, the Compactness Theorem is the cornerstone of model theory. Because it will not be useful for us to understand the exact nature of our proof system, we will not prove the Completeness Theorem. Instead, in the next section, we will give a second proof of the Compactness Theorem that does not appeal directly to the Completeness Theorem.

Basic Applications of Compactness

We conclude this section with several standard applications of the Compactness Theorem.

Corollary 2.5 *Suppose T has arbitrarily large finite models, then T has an infinite model.*

Proof Let ϕ_n be the sentence:

$$\exists v_1 \ldots \exists v_n \bigwedge_{i<j\leq n} v_i \neq v_j .$$

Let $T^* = T \cup \{\phi_n : n = 1, 2, \ldots\}$. Clearly any model of T^* is an infinite model of T. If $\Delta \subset T^*$ is finite, then for some N, $\Delta \subset T \cup \{\phi_1, \ldots, \phi_N\}$. There is $\mathcal{A} \models T$ with $|\mathcal{A}| \geq N$, thus $\mathcal{A} \models \Delta$. By the Compactness Theorem, T^* has a model.

Proposition 2.6 *Let $\mathcal{L} = \{\cdot, +, <, 0, 1\}$ and let* $\mathrm{Th}(\mathbb{N})$ *be the full $\mathcal{L}$-theory of the natural numbers. There is $\mathcal{M} \models \mathrm{Th}(\mathbb{N})$ and $a \in M$ such that a is larger than every natural number.*

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

$$T = \mathrm{Th}(\mathbb{N}) \cup \{\underbrace{1 + 1 + \ldots + 1}_{n-\text{times}} < c : \text{for } n = 1, 2, \ldots\}.$$

If Δ is a finite subset of T, we can make $\mathbb{N}$ a model of Δ by interpreting c as a suitably large natural number. Thus, T is finitely satisfiable and there is $\mathcal{M} \models T$. If $a \in M$ is the interpretation of c, then a is larger than every natural number.

Proposition 2.7 *Let $\mathcal{L}$ be a language containing $\{\cdot, e\}$, the language of groups, let T be an $\mathcal{L}$-theory extending the theory of groups, and let $\phi(v)$ be an $\mathcal{L}$-formula. Suppose that for all n there is $G_n \models T$ and $g_n \in G_n$ with finite order greater than n such that $G_n \models \phi(g_n)$. Then, there is $G \models T$ and $g \in G$ such that $G \models \phi(g)$ and g has infinite order. In particular, there is no formula that defines the torsion points in all models of T.*

Proof Let $\mathcal{L}^* = \mathcal{L} \cup \{c\}$, where c is a new constant symbol. Let T^* be the $\mathcal{L}$-theory

$$T \cup \{\phi(c)\} \cup \{\underbrace{c \cdot c \cdots c}_{n-\text{times}} \neq e : n = 1, 2, \ldots\}.$$

If G is a model of T^* and g is the interpretation of c in G then $G \models \phi(g)$ and g has infinite order. Hence, it suffices to show that T^* is satisfiable.

Let $\Delta \subseteq T^*$ be finite. Then

$$\Delta \subseteq T \cup \{\phi(c)\} \cup \{\underbrace{c \cdot c \cdots c}_{n-\text{times}} \neq e : n = 1, 2, \ldots, m\}$$

for some m. View G_m as an $\mathcal{L}^*$ structure by interpreting c as the element g_m. Because $G_m \models T \cup \{\phi(g_m)\}$ and g_m has order greater than m, $G_m \models \Delta$. Thus, T^* is finitely satisfiable and hence, by the Compactness Theorem, satisfiable.

Example 2.8 *Four Coloring Graphs*

Let $G = (V, E)$ be a graph such that every finite subgraph can be four colored.[2] We claim that G can be four colored. Let $\mathcal{L} = \{R, B, Y, G\} \cup \{c_v : v \in V\}$. Let Γ be the $\mathcal{L}$-theory with axioms:

i) $\forall x\ [(R(x) \wedge \neg B(x) \wedge \neg Y(x) \wedge \neg G(x)) \vee \ldots \vee (\neg R(x) \wedge \neg B(x) \wedge \neg Y(x) \wedge G(x))]$

ii) if $(v, w) \in E$ add the axiom: $\neg(R(c_v) \wedge R(c_w)) \wedge \ldots \wedge \neg(G(c_v) \wedge G(c_w))$.

[2] That is, we can color the vertices with four colors so that no adjacent vertices have the same color. For example, the Four Color Theorem says that every finite planar graph can be four colored.

If Δ is a finite subset of Γ, let V_Δ be the verticies such that c_v is used in Δ. Since the restriction of G to V_Δ is four colorable, Δ is consistent. Thus Γ is consistent. Let $\mathcal{A} \models \Gamma$.

Color G by coloring v as $\mathcal{A}$ colors c_v.

Theorem 2.9 (Upward Löwenheim–Skolem Theorem) *Suppose Γ is an $\mathcal{L}$-theory. If Γ has an infinite model, then it has a model of cardinality κ for every $\kappa \geq \max(|\mathcal{L}|, \aleph_0)$.*

Proof Let I be a set of cardinality κ. Let $\mathcal{L}^* = \mathcal{L} \cup \{c_\alpha : \alpha \in I\}$. Let

$$\Gamma^* = \Gamma \cup \{c_\alpha \neq c_\beta : \alpha \neq \beta\}.$$

If Δ is a finite subset of Γ^*, then in any infinite model $\mathcal{A}$ of Γ we can interpret the constants such that $\mathcal{A} \models \Delta$. Thus Γ has a model of size at most κ. But certainly any model of Γ^* has size at least κ (the map $\alpha \mapsto \widehat{c}_\alpha$ is one to one).

The next lemma is an easy consequence of the Completeness Theorem, but it also can be deduced from the Compactness Theorem.

Lemma 2.10 *If $T \models \phi$, then $\Delta \models \phi$ for some finite $\Delta \subseteq T$.*

Proof Suppose not. Let $\Delta \subseteq T$ be finite. Because $\Delta \not\models \phi$, $\Delta \cup \{\neg\phi\}$ is satisfiable. Thus, $T \cup \{\neg\phi\}$ is finitely satisfiable and, by the Compactness Theorem, $T \not\models \phi$.

3 Ultraproducts and Compactness

In this section we will give an alternative proof of the Compactness Theorem using ultraproducts, an algebraic method of *averaging* structures.

Let I be an infinite set. We let

$$\mathcal{P}(I) = \{A : A \subseteq I\}$$

be the *power set* of I.

Definition 3.1 We say that $\mathcal{F} \subseteq \mathcal{P}(I)$ is a *filter* if
i) $I \in \mathcal{F}$, $\emptyset \not\in \mathcal{F}$;
ii) If $A \in \mathcal{F}$ and $A \subseteq B$, then $B \in \mathcal{F}$;
iii) If $A, B \in \mathcal{F}$ then $A \cap B \in \mathcal{F}$.

We say that $\mathcal{F}$ is an *ultrafilter* if in addition,
iv) for all $A \subseteq I$ either $A \in \mathcal{F}$ or $I \setminus A \in \mathcal{F}$.

Example 3.2 *Cof* $= \{A \subseteq I : I \setminus A$ *is finite*$\}$ *is a filter.*

Example 3.3 *Let* $I = \mathbb{R}$ *then* $\mathcal{F} = \{A : \mathbb{R} \setminus A$ *has Lebesgue measure zero*$\}$, *is a filter.*

If $\mathbb{F}$ is a filter, we think of elements of $\mathcal{F}$ as l*large*, so if $A \in \mathcal{F}$ we think of A as large and that $i \in A$ for *almost all* $i \in I$.

We can think of an ultrafilter $\mathbb{F}$ as finitely additive two valued measures $\mu : \mathcal{P}(I) \to \{0, 1\}$, where $\mu(A) = 1$ if and only if $A \in \mathbb{F}$.

Lemma 3.4 *If* $\mathcal{F} \subseteq \mathcal{P}(I)$ *is a filter,* $A \subseteq I$ *and* $I \setminus A \not\in \mathcal{F}$, *then*

$$\mathcal{F}' = \{C : \textit{ there is } B \in \mathcal{F}, C \supseteq A \cap B\}$$

is an ultrafilter and $A \in \mathcal{F}'$.

Proof Since $I \supseteq I \cap A$, $I \in \mathcal{F}'$.

If $\emptyset \in \mathcal{F}'$, then there is $B \in \mathcal{F}$ such that $A \cap B = \emptyset$. But then $B \subseteq I \setminus A$ and $I \setminus A \in \mathcal{F}$, a contradiction.

It is easy to see that $\mathcal{F}'$ is closed under superset.

If $C_1, C_2 \in \mathcal{F}'$ there are $B_1, B_2 \in \mathcal{F}$ such that $C_i \supseteq B_i \cap A$. Then $C_1 \cap C_2 \supseteq B_1 \cap B_2 \cap A$, so $C_1 \cap C_2 \in \mathcal{F}'$.

Corollary 3.5 *If* $\mathcal{F} \subseteq \mathcal{P}(I)$ *is a filter, then there is an ultrafilter* $\mathcal{U} \supseteq \mathcal{F}$.

Proof Let $\mathcal{I} = \{\mathcal{F}' : \mathcal{F} \subseteq \mathcal{F}' \subseteq \mathcal{P}(I)$ is a filter$\}$.

If $(X, <)$ is a linearly ordered set, $\mathcal{F}_x \in \mathcal{I}$ for $x \in X$ and $\mathcal{F}_x \subseteq \mathcal{F}_y$ for $x < y$, then $\mathcal{F}^* = \bigcup_{x \in X} \mathcal{F}_x$ is a filter. Thus we can apply Zorn's Lemma to find $\mathcal{U} \in \mathcal{I}$ maximal. Suppose $A \subseteq I$. If $I \setminus A \not\in \mathcal{U}$, then, by the Lemma and the maximality of $\mathcal{U}$, $A \in U$.

Corollary 3.6 *There are non-principal ultrafilters.*

Proof Let $\mathcal{U} \supseteq \mathrm{Cof}$ be an ultrafilter. Then $\mathcal{U}$ contains no finite sets.

Our proof of the existence of non-prinicipal ultrafilters is non-constructive as it depends heavily on the Axiom of Choice. Unfortunately, some use of choice is unavoidable.

We will use ultrafilters to give a new construction of models. Let $\mathcal{L}$ be a first order language. Suppose that $\mathcal{M}_i$ is an $\mathcal{L}$-structure for all $i \in I$ with universe M_i. Let $\mathcal{U} \subseteq \mathcal{P}(\mathcal{I})$ be an ultrafilter.

We define $\sim$ on $\prod_{i\in I} M_i$ by

$$f \sim g \Leftrightarrow \{i \in I : f(i) = g(i)\} \in \mathcal{U}.$$

Lemma 3.7 $\sim$ *is an equivalence relation*

Proof Let $f, g, h \in \prod_{i\in I} M_i$. Clearly $f \sim f$ and if $f \sim g$, then $g \sim f$.

Suppose $f \sim g$ and $g \sim h$. Since

$$\{i : f(i) = h(i)\} \supseteq \{i : f(i) = g(i)\} \cap \{i : g(i) = h(i)\} \in \mathcal{U},$$

$f \sim h$.

For $f \in \prod_{i\in I}$, let $[f]$ be the $\sim$-equivalence class of f and let

$$M = \left\{[f] : f \in \prod_{i\in I} M_i\right\}.$$

We will interpret the symbols of $\mathcal{L}$ in M to construct an $\mathcal{L}$-structure $\mathcal{M}$, which we also denote $\prod M_i/\mathcal{U}$.

If c is a constant symbol of $\mathcal{L}$, let $f \in \prod M_i$ be the function $f(i) = c^{\mathcal{M}_i}$ and let $c^{\mathcal{M}} = [f]$.

Let R be an n-ary relation symbol of $\mathcal{L}$.

Lemma 3.8 $f_1, \ldots, f_n, g_1, \ldots, g_n \in \prod M_i$ *such that* $f_j \sim g_j$ *for all* $j = 1, \ldots, n$. *Then*

$$\{i \in I : (f_1(i), \ldots, f_n(i)) \in R^{\mathcal{M}_i}\} \in \mathcal{U} \Leftrightarrow \{i \in I : (g_1(i), \ldots, g_n(i)) \in R^{\mathcal{M}_i}\} \in \mathcal{U}.$$

Proof Suppose $\{i \in I : (f_1(i), \ldots, f_n(i)) \in R^{\mathcal{M}_i}\} \in U$. Then $\{i \in I : (g_1(i), \ldots, g_n(i)) \in R^{\mathcal{M}_i}\}$ contains

$$\{i \in I : (f_1(i), \ldots, f_n(i)) \in R^{\mathcal{M}_i}\} \cap \{i \in I : g_1(i) = f_1(i)\} \cap \ldots \cap \{i \in I : g_n(i) = f_n(i)\}.$$

Since $\mathcal{U}$ is a filter this later set is in $\mathcal{U}$.

The other direction is symmetric.

We define

$$R^{\mathcal{M}} = \{([f_1], \ldots, [f_n]) : \{i \in I : (f_1(i), \ldots, f_n(i)) \in \mathbb{R}^{\mathcal{M}_i}\} \in \mathcal{U}\}.$$

By the Lemma, this is well-defined and does not depend on the choice of representatives for the equivalence classes.

Let F be an n-ary function symbol of $\mathcal{L}$. Let $f_1,\dots,f_n,g_1,\dots,g_n \in \prod M_i$ with $f_j \sim g_j$ for $j = 1,\dots,n$. Define $f_{n+1}, g_{n+1} \in \prod M_i$ by

$$f_{n+1}(i) = F(f_1(i),\dots,f_n(i)) \text{ and } g_{n+1}(i) = F(g_1(i),\dots,g_n(i)).$$

Exercise 3.9 Argue as in Lemma 3.8 that $f_{n+1} \sim g_{n+1}$.

We define $F^{\mathcal{M}} : M^n \to M$ by

$$F([f_1],\dots,[f_n]) = [g]$$

where $g(i) = F(f_1(i),\dots,f_n(i))$. By Exercise 3.9 this is well defined and does not depend on choice of representatives.

We have now completely defined the structure $\mathcal{M} = \prod M_i/U$. We call $\mathcal{M}$ an *ultraproduct* of $(\mathcal{M}_i : i \in I)$

The following exercise is an easy induction on terms.

Exercise 3.10 If t is an $\mathcal{L}$-term, then $t^{\mathcal{M}}(f_1,\dots,f_n) = [g]$ where $g(i) = t^{\mathcal{M}_i}(f_1(i),\dots,f_n(i))$.

We can now state the Fundamental Theorem of Ultraproducts.

Theorem 3.11 (Łos's Theorem) *Let $\phi(v_1,\dots,v_n)$ be any $\mathcal{L}$-formula Then*

$$\mathcal{M} \models \phi([f_1],\dots,[f_n]) \Leftrightarrow \{i : \mathcal{M}_i \models \phi(f_1(i),\dots,f_n(i))\} \in \mathcal{U}.$$

Proof We prove this by induction on complexity of formulas

1) Suppose ϕ is $t_1 = t_2$ where t_1 and t_2 are terms.

Define $g_j(i) = t_j^{\mathcal{M}_i}(f_1(i),\dots,f_n(i))$. Then

$$\mathcal{M} \models t_1([f_1],\dots,[f_n]) = t_2([f_1],\dots,[f_n]) \Leftrightarrow [g_1] = [g_2]$$

$$\Leftrightarrow \{i : t_1^{M_i}(f_1(i),\dots,f_n(i)) = t_2^{\mathcal{M}_i}(f_1(i),\dots,f_n(i)\} \in \mathcal{U}$$

as desired.

2) Suppose ϕ is $R(t_1,\dots,t_m)$.

For $j = 1,\dots,m$ let $g_j(i) = t_i^{\mathcal{M}_i}(f_1(i),\dots,f_n(i))$. Then

$$\begin{aligned}\mathcal{M} \models \phi([f_1],\dots,[f_n]) &\Leftrightarrow \{i : (g_1(i),\dots,g_n(i)) \in R^{\mathcal{M}_i}\} \in \mathcal{U}\\ &\Leftrightarrow \{i : \mathcal{M}_i \models \phi(f_1(i),\dots,f_n(i))\} \in \mathcal{U}\end{aligned}$$

3) Suppose the theorem is true for θ and ψ, and ϕ is $\theta \wedge \psi$. (We suppress the parameters $[f_1],\dots,[f_n]$)

Then

$$\begin{aligned}\mathcal{M} \models \phi &\Leftrightarrow \mathcal{M} \models \psi \text{ and } \mathcal{M} \models \theta\\ &\Leftrightarrow \{i : \mathcal{M}_i \models \psi\} \in \mathcal{U} \text{ and } \{i : \mathcal{M}_i \models \psi\} \in \mathcal{U}\end{aligned}$$

$$\Leftrightarrow \quad \{i : \mathcal{M}_i \models \psi \wedge \theta\} \in \mathcal{U}$$

4) Suppose the theorem is true for ψ and ϕ is $\neg\psi$ Then

$$\begin{aligned} \mathcal{M} \models \phi &\Leftrightarrow \mathcal{M} \not\models \psi \\ &\Leftrightarrow \{i : \mathcal{M}_i \models \psi\} \notin \mathcal{U} \\ &\Leftrightarrow \{i : \mathcal{M}_i \models \neg\psi\} \in \mathcal{U} \end{aligned}$$

5) Suppose the theorem is true for $\psi(v)$ and ϕ is $\exists v\ \psi(v)$.

If $\mathcal{M} \models \exists v\ \psi(v)$, then there is g such that $\mathcal{M} \models \psi([g])$. But then

$$\{i : \mathcal{M}_i \models \exists v\ \psi(v)\} \supseteq \{i : \mathcal{M}_i \models \psi(g(i))\} \in \mathcal{U}$$

On the other hand if $A = \{i : \mathcal{M}_i \models \exists v\ \psi(v)\} \in \mathcal{U}$ define $g \in \prod M_i$ such that $\mathcal{M}_i \models \psi(g(i))$ for all $i \in A$. Then $\mathcal{M} \models \psi([g])$, so $\mathcal{M} \models \phi$.

Note that step 4) is the only place in the construction that we used that $\mathcal{U}$ is an ultrafilter rather than just a filter.

Exercise 3.12 Let $\mathcal{U}$ be a non-princpal ultrafilter on the set of prime numbers. For each prime p, let $\mathbb{F}_p^{\mathrm{alg}}$ be the algebraic closure of $\mathbb{F}_p$ the field with p elements. Prove that $\prod \mathbb{F}_p/\mathcal{U}$ is an algebraically closed field of characteristic 0.

Another Proof of Compactness

We can use Łos's Theorem to give a proof of the Compactness Theorem that avoids the Completeness Theorem.

Let Γ be an $\mathcal{L}$-theory such that every finite $\Delta \subseteq \Gamma$ has a model. Let I be the collection of finite subsets of Γ.

For $\phi \in \Gamma$ let

$$X_\phi = \{\Delta \in I : \Delta \models \phi\}$$

and let

$$\mathcal{F} = \{Y \subseteq I : X_\phi \subseteq Y \text{ for some } \phi \in \Gamma\}.$$

We claim that $\mathcal{F}$ is a filter. It is easy to see that $I \in \mathcal{F}$, $\emptyset \notin \mathcal{F}$ and $\mathcal{F}$ is closed under superset. Also if $Y_1, Y_2 \in \mathcal{F}$ there are ϕ_1, ϕ_2 such that $X_{\phi_i} \subseteq Y_i$. Then $X_{\phi_1 \wedge \phi_2} = X_{\phi_1} \cap X_{\phi_2}$, so

$$X_{\phi_1 \wedge \phi_2} \subseteq Y_1 \cap Y_2$$

and $Y_1 \cap Y_2 \in \mathcal{F}$

Let $\mathcal{U} \supseteq \mathcal{F}$ be an ultrafilter. For $\Delta \in I$, let $\mathcal{M}_\Delta \models \Delta$ and let $\mathcal{M} = \prod \mathcal{M}_\Delta/\mathcal{U}$. Since $X_\phi \in \mathcal{U}$ for all $\phi \in \Gamma$, by łos's Theorem $\mathcal{M} \models \Gamma$.

Ultrapowers and Elementary Extensions

Fix $\mathcal{M}$ and $\mathcal{L}$ structure and let $\mathcal{U}$ be an ultrafilter on an infinite set I. An interesting special case of the ultraproduct construction is when we take all of the $\mathcal{M}_i = \mathcal{M}$. In this case we let $\mathcal{M}^* = \mathcal{M}^I/U$.

Exercise 3.13 Prove that if $\mathcal{M}$ is finite or $\mathcal{U}$ is principal, then $\mathcal{M} \cong \mathcal{M}^*$.

For each $a \in M$, let $f_a : I \to M$ be the constant function $f_a(i) = a$. If $a \neq b$, then $[f_a] \neq [f_b]$. By Los's Theorem if $a_1, \ldots, a_n \in \mathcal{M}$ and ϕ is an $\mathcal{L}$-formula, then

$$\mathcal{M} \models \phi(a_1, \ldots, a_n) \Leftrightarrow \mathcal{M}^* \models \phi([f_{a_1}], \ldots, [f_{a_n}])$$

Identifying $\mathcal{M}$ and it's image under the embedding $a \mapsto [f_a]$ we can think of $\mathcal{M}$ as substructure of $\mathcal{M}^*$. Then for $a_1, \ldots, a_n \in M$.

$$\mathcal{M} \models \phi(a_1, \ldots, a_n) \Leftrightarrow \mathcal{M}^* \models \phi(a_1, \ldots, a_n).$$

Definition 3.14 If $\mathcal{M} \subseteq \mathcal{N}$ we say that $\mathcal{N}$ is an *elementary extension* of $\mathcal{M}$ and write $\mathcal{M} \prec \mathcal{N}$ if

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

for all $\overline{a} \in M$.

We have argued that $\mathcal{M}^*$ is an elementary extension of $\mathcal{M}$. This is only interesting if we can also prove $\mathcal{M}^*$ properly extends $\mathcal{M}$.

Proposition 3.15 *If $|I| \leq |\mathcal{M}|$ and $\mathcal{U}$ is a non-principal ultrafilter, then $\mathcal{M}^*$ is a proper extension of $\mathcal{M}$.*

Proof Let $f : I \to M$ be injective. Then for all $a \in M$, $|\{i : f(i) = f_a(i)\}| \leq 1$. Since $\mathcal{U}$ is non-principal, $f \not\sim f_a$. Thus $[f] \in M^* \setminus M$.

Cardinalities of Ultraproducts

Suppose we have $(\mathcal{M}_i : i \in I)$ and an ultrafilter $\mathcal{U} \subseteq \mathcal{P}(I)$.

Exercise 3.16 Suppose $\{i \in I : |\mathcal{M}_i| = n\} \in \mathcal{U}$, then $|\prod \mathcal{M}_i/\mathcal{U}| = n$

Exercise 3.17 If we also have $(\mathcal{N}_i : i \in I)$ and $\{i : |\mathcal{M}_i| = |\mathcal{N}_i|\} \in \mathcal{U}$, then $|\prod \mathcal{M}_i/U| = |\prod \mathcal{N}_i/U|$.

Exercise 3.18 If $\lambda \leq |\mathcal{M}_i| \leq \kappa$ for all $i \in I$, then

$$\lambda \leq \prod \mathcal{M}_i/\mathcal{U} \leq \kappa^{|I|}.$$

For the rest of these Exercises we will assume $I = \mathbb{N}$.

Exercise 3.19 Suppose that for all $n \in \mathbb{N}$, $\{i : |\mathcal{M}_i| = n\} \notin \mathcal{U}$ and $\mathcal{U}$ is non-principal.
a) Show there is a family X of functions $f : \mathbb{N} \to \mathbb{N}$ such that:

i) $|X| = 2^{\aleph_0}$

ii) for each $f \in X$ $f(n) < 2^n$

iii)$f \neq g \in X$, then $\{n : f(n) = g(n)\}$ is finite.

[Hint: For $\alpha : \mathbb{N} \to \{0,1\}$ let $f_\alpha(n) = \sum_{i=0}^{n-1} \alpha(i)2^i$].

b) Show there is a partition $I = \bigcup_{n=0}^{\infty} A_n$ such that

i) each $A_n \notin \mathcal{U}$

ii) if $i \in A_n$, then $|\mathcal{M}_i| \geq 2^i$.

[Hint: Let $A_n = \{i : 2^n \leq |M_i| < 2^{n+1}$ or $i = n$ and $|\mathcal{M}_i| \geq \aleph_0\}$.]

For $i \in I$ let $n(i)$ be unique such that $i \in A_{n(i)}$. For $i \in I$ choose $(m_{i,j} : 0 \leq j < 2^{n(i)})$ distinct elements of M_i. For $f \in X$, let $\alpha_f \in \prod M_i$ such that $\alpha_f(i) = m_{i,f(n(i))}$.

c) Prove that if $f \neq g \in X$, then $\alpha_f \not\sim \alpha_g$. Conclude that $|\prod \mathcal{M}_i/U| \geq 2^{\aleph_0}$.

Corollary 3.20 *Suppose that $\mathcal{U}$ is a non-prinicpal ultrafilter on $\mathbb{N}$, $|\mathcal{M}_n| \leq \aleph_0$ for all n, and $\{n : |\mathcal{M}_n| = m\} \notin U$ for any m, Then $|\prod \mathcal{M}_i/U| = 2^{\aleph_0}$.*

Exercise 3.21 Let $\mathcal{U}$ be a non-principal ultrafilter on the set of primes. Prove $\prod \mathbb{F}_p^{\mathrm{alg}}/\mathcal{U}$ is isomorphic to $\mathbb{C}$ the field of complex numbers.